CHARACTERIZATION OF THE INTERACTION OF GET3 WITH THE MEMBRANE BOUND RECEPTORS GET1 AND GET2

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

vorgelegt dem Fachbereich

Biochemie, Chemie und Pharmazie (FB 14)

der Goethe - Universität

Frankfurt am Main

von

Jean Aymard NZIGOU MANDOUCKOU

Frankfurt 2017

D 30

“Wisdom entereth not into a malicious mind, and science without conscience is but the ruin of the soul.”

***« Science sans conscience n'est que ruine de l'âme*. »**

François Rabelais

Vom Fachbereich Biochemie, Chemie und Pharmazie (FB14)
der Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr Michael Karas
Gutachter: Prof. Dr. Volker Dötsch
Prof. Dr. Clemens Glaubitz

Datum der Disputation:

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

CAML. : calcium-modulating cyclophilin ligand
CCD. : charge-coupled devices
CEN. : yeast centromere sequence
CMOS. : complementary metal-oxide-semiconductor
Cs. : spherical aberration
CTF. : contrast transfer function
DDM. : n-Dodecyl ß-D-maltoside
DM. : n-decyl ß-D-maltoside
EFTEM. : energy-filtered transmission electron microscopy
EM. : Electron microscopy
ER. : endoplasmic reticulum
ESFRI. : European Strategy Forum on Research infrastructures
FOM. : figure-of merit
FSC. : Fourier Shell correction
GAL1. : galactokinase
GFP. : green fluorescence protein
GPI. : glycosylphosphatidylinositol
GraFix. :Gradient Fixation
GTP. : Guanosine-5'-triphosphate
HCl. : hydrochloric acid
HSP. : heat shock protein 70
IBD. : insertion box
IPTG. : isopropyl-ß-D-thiogalactopyranoside
LB. : Luria-Bertani
LDAO. : N,N-dimethydodecylamine n-oxide
LiAc. : lithium acetate
LILBID. : Laser Induced Liquid Bead Ion Desorption
Linac. : linear accelerator
MALDI. : Matrix-assisted laser desorption/ionization
MST. : microscale Thermophoresis
NBS. : nucleotide-binding sites
Ni-NTA. : nickel-nitrilotriacetic acid
PCR. : polymerase chain reaction
SDS. : sodium-dodecyl-sulfate
SDS-PAGE. : Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis
Sgt2. : small Glu-rich tetratricopeptide repeat-containing 2
SNARE. : soluble N-ethylmaleimide-sensitive factor attachment protein receptor
SNR. : signal-to-noise ratio
SR. : SRP receptor
SRP. : signal recognition particle
TA. : Tail-anchored
TB. : Terrific broth
TEM. : transmission electron microscopy
TEV. : Tobacco Etch Virus

ABBREVIATIONS

TMD. : transmembrane domain
TRC. : TMD recognition complex
VAMP. : vesicle-associated membrane protein
WRB. : tryptophan-rich basic protein
YIp. : yeast integrating plasmids
YPD. : yeast peptone dextrose

SUMMARY

The focus of this research was to understand the molecular mechanism that lies behind the insertion of tail-anchored membrane proteins into the ER membrane of yeast cells. State-of-art instruments such as LILBID, and Cryo-EM, combined with the introduction of direct electron detectors, were used to analyze the proteins that capture tail-anchored proteins near the ER membrane and help their releases from a chaperone, an ATPase named Get3. Get3 escorts TA proteins to the ER membrane, where both Get3 and the TA proteins interact sequentially to Get3 membrane bound receptors Get1 and Get2. Get1 and Get2 are homologs of mammalian WRB and CAML.

The native host was used to separately produce Get1, Get2, and the Get2/Get1 single chain constructs. The studies showed that when Get1 is expressed alone, Get1 does not seems to be located in the ER membrane but rather in microbodies like shape organelles (or peroxisome). Interestingly, Get1 seems to be located in the ER membrane when it is linked to Get2 as single chain construct.

The localization study of Get2/Get1 fused to GFP shows from the fluorescence intensity that Get2/Get1.GFP has a tube-like morphology or membrane-enclosed sacs (cisterna), implying that Get2/Get1 is actually targeted to the ER membrane and is likely functional. In other words, Get1 and Get2 stabilize each other in the ER membrane.

The expression of Get2/Get1 was found to be already optimum when expressed as single chain construct because the fluorescence counts did not improve when additives such as DMSO or histidine were added. However, when Get1 and Get2 are expressed separately, additives improve their protein production yield. In 1 liter culture, Get1 yield is increased by about 3 mg and Get2 by 1.8 mg. This can be explained by the space that Get1 and Get2 should occupy within the ER membrane as they must coexist with other membrane components to maintain the homeostasis of the cell. Hence, if there were no gain for single chain construct expression, it meant that Get2/Get1 was already well expressed on its own in ER membrane and has reached its optimum expression without the help of additives. The Get2/Get1 overexpression is more stable, tolerated and less toxic for the cells to express it at a high level.

DDM has proved to be the best detergent from the detergents tested to solubilize Get1, Get2, and Get2/Get1.

Thereafter, Get1, Get2 (data not shown), and Get2/Get1 were successfully purified in DDM micelles.

Furthermore, for the first time using LILBID, the actual study has shown that Get1 and Get2 are predominantly a heterotetramer (2xGet1 and 2xGet2) but higher oligomerization may exist as well.

Get3 binds to Get1 in a biphasic way with a specific strong binding of an affinity of 57 nM and the second of 740 nM nonspecific indicative of heterogeneity within the interaction between Get1 and Get3. This heterogeneity is caused by the presence of different conformation of either protein. However, in order to characterize a high-resolution structure model of a specific target one needs highly homogenous and identical molecules of the target protein or complex in solution. The homogeneity increases the chances of growing crystals during crystallography as the good homogeneity will likely generate a perfect packing of unit cells stack (also known as crystal lattice) in the three-dimensional spaces. The same truth goes for the single particles analysis Cryo-EM, especially for smaller complexes where having less or no conformation alterations of specific targets will enable the researcher to classify the particles in 2D and 3D, therefore improving the signal-to-noise-ratio that will ultimately lead to high-resolution structure determination.

Get1, Get2/Get1 and chimeric variants (tGet2/Get1, T4l.Get2/Get1, T4l.Get2.apocyte.Get1) were crystallized but none of the crystals could diffract due to heterogeneity.

This heterogeneity was not only occurring upon the binding of Get3 to its membrane receptors, but seems to be already present within the receptors themselves through possibly different conformation.

In this Ph.D. thesis, the heterogeneity of purified Get2 and Get1 as complex or individually in detergent is then, so far, the limiting factor for obtaining a high-resolution structure model of Get1 and Get2. As mentioned above, the heterogeneity observed was

not due to the quality of the sample preparation but rather to the effect of different conformations that could have been native, or just because of the micelle used, as it was proven by the 3-D heterogeneity classification by Cryo-EM.

In general, crosslinking is one way to keep the integrity of protein complexes, however it appeared not to improve the sample quality when it was analyzed in micelles. Often the integrity of some membrane proteins is affected when they are solubilized and purified in detergents.

Finally, in this study, the structural map of Get2 and Get1 complex linked with chimeric protein T4 lysozyme and apocytochrome C b_{562}RIL gene was obtained at 10 Å. However, this single chain construct has a density map corresponding to heterodimer species (one Get1 and Get2). Therefore, based on those data the tertiary structure of Get2/Get1 in micelle is poorly defined. It could be that the membrane extraction in DDM and the purification destabilizes the structure of the complex so future works should only focus on getting the structure in lipid system environment.

ZUSAMMENFASSUNG

Der Fokus dieser Forschungsarbeit lag auf dem Verständnis der molekularen Mechanismen, die der Insertion von *tail-anchored* Membranproteinen in die ER Membran von Hefezellen zugrunde liegen. Modernste biophysikalische Techniken wie LILBID und Cryo-EM wurden zur Analyse von Proteinen benutzt, welche *tail-anchored* Proteine in der Nähe der ER Membran festhalten und ihre Freilassung von Get3 begünstigen, einem Chaperone mit ATPase-Aktivität. Get3 eskortiert TA Proteine zur ER Membran, wo sowohl Get3 als auch die TA Proteine sequentiell mit den Get3 membrangebunden Rezeptoren Get1 und Get2 interagieren. Get1 und Get2 haben ihre Homologe in Säugetieren in WRB und CAML.

Der native Wirt wurde benutzt um getrennt Get1, Get2 und das Get2/Get1 Fusionskonstrukt zu produzieren. Die Studien zeigten, dass, Get1, wenn separat exprimiert, nicht an der ER Membran lokalisiert zu sein scheint, sondern vielmelr in Mikrokörpern wie *shape organelles* (oder Peroxisomen). Interssanterweise scheint Get1 allerdings an der ER Membran lokalisiert zu sein, wenn es als Fusionskonstrukt mit Get2 verbunden ist.

Die Lokalisation eines GFP fusionierten Get2/Get1 Konstruktes wurde durch die GFP Fluoreszenz analysiert und zeigte röhrenförmige Morphologien oder membranumschlossene *sacs* (Zisternen), was andeutet, dass Get2/Get1 zur ER Membran lokalisiert und wahrscheinlich funktional ist. In anderen Worten, Get1 und Get2 stabilisieren sich gegenseitig in der ER Membran.

Die Expression von Get2/Get1 als Fusionskonstrukt war optimal und konnte nicht durch Zugabe von DMSO oder Histidin weiter erhöht werden, wie sich durch Fluoreszenzzählungen ermitteln lies. Allerdings konnte die Proteinproduktionsausbeute durch Zugabe von Additiven gesteigert werden, wenn Get1 und Get2 separat exprimiert wurden. Pro liter Kultur konnte die Ausbeute von Get1 um etwa 3 mg und die von Get2 um etwa 1.8 mg gesteigert werden.

Eine mögliche Erklärung ist der Abstand, den Get1 und Get2 in der ER Membran einnehmen müssen, da sie zusammen mit anderen Membrankomponenten koexistieren müssen, um die Homöostase der Zelle aufrecht zu erhalten.

Aus dem nicht Vorhandensein einer Expressionssteigerung für das Fusionskonstrukt lässt sich somit schlussfolgern, dass das Konstrukt Get2/Get1 sein Expressionsoptimum an der ER Membran selbstständig und ohne Hilfe von Additiven erreicht hat.

Die Überexpression von Get2/Get1 ist stabiler, tolerierter und weniger toxisch für die Zellen bei Expression auf einem hohen Level.

Von den getesteten Detergenzien hat sich DDM als dasjenige herausgestellt, welches Get1, Get2 und Get2/Get1 am besten solubilisiert.

Hiernach wurden Get1, Get2 (Daten nicht gezeigt) und Get2/Get1 erfolgreich in DDM Mizellen aufgereinigt.

Weiterhin, wurde erstmalig durch LILBID in dieser Studie gezeigt, das Get1 und Get2 predominant als Heterotetramer (2xGet1 und 2xGet2) vorliegen, allerdings könnten auch höhere Oligomierisierungen existieren.

Get3 bindet Get1 biphasisch mit einer starken ersten, spezifischen Affinität von 57 nM und einer zweiten unspezifischen Affinität von 740 nM, die eine Heterognenität der Interaktion zwischen Get1 und Get3 impliziert. Diese Heterogenität wird durch das Vorhhandensein unterschiedlicher Konformationen der jeweiligen Proteine verursacht. Allerdings müssen für die Charakterisierung eines hochauflösenden Strukturmodels eines bestimmten Zieles die Moleküle des Zielproteines oder Komplexes in der Lösung hoch homogen vorliegen.

Diese Homogenität erhöht die Wahrscheinlichkeit eines wachsenden Kristalls während der Kristallographie da eine gute Homogenität für eine perfekte Packung der Einheitszellen (auch *cristal lattice* genannt) im dreidimensionalen Raum sorgt.

Das Gleiche stimmt für Einzelpartikelanalysen in der Cryo-EM, besonders für kleinere Komplexe, bei denen das Vorhandensein von nur wenigen oder keinen Konfirmationsalternativen eines spezifischen Ziels es dem Forscher ermöglicht, die

Partikel einfach in 2D und 3D zu klassifizieren, was das Signal-zu-Rausch-Verhältnis verbessert und letztendlich zu einer hoch aufgelösten Struktur führt.

Get1, Get2/Get1 und Chimerenvarianten (tGet2/Get1, T4l.Get2/Get1, T4l.Get2.apocyte.Get1) wurden kristallisiert aber aufgrund der Heterogenität konnte keiner der Kristalle diffraktieren.

Diese Heterogenität trat nicht nur nach der Bindung von Get3 zu seinen Membranrezeptoren auf, sondern scheint schon in den Rezeptoren selbst präsent zu sein, möglicherweise verursacht durch unterschiedliche Konfirmationen.

In dieser Doktorarbeit ist die Heterogenität von aufgereinigtem Get2 und Get1 als Komplex oder individuell in Detergenz bisher der limitierende Faktor für ein hoch-auflösendes Strukturmodell von Get1 und Get2.

Wie vorher angemerkt, war die beobachtete Heterogenität nicht verursacht durch die Qualität der Probenvorbereitung sondern vielmehr ein Effekt von unterschiedlichen Konfirmationen, welche nativen Ursprungs oder durch die Mizellen verursacht worden sein könnten, wie die 3-D Heterogenitätsklassifikation durch Cryo-EM aufzeigte.

Im Allgemeinen ist *crosslinking* der Weg um die Integrität von Proteinkomplexen zu erhalten, allerdings schien es nicht der Verbesserung der Analyse in der Mizelle zu helfen. Die Integrität einiger Membranproteine ist oft beeinträchtigt, wenn sie in Detergenz solublisiert und aufgereinigt werden.

Letzendlich wurde in dieser Studie die strukturelle Karte des Get2 und Get1 Komplexes, verbunden durch die chimären Proteine T4 Lysozym und Apocytochrome C b_{562}RIL, mit einer Auflösung von 10 Å erhalten. Allerdings hat dieses Fusionskonstrukt die Dichtekarte von heterodimeren Spezies (ein Get1 und Get2), deshalb ist die Tertiärstruktur von Get2/Get1 in Mizellen basierend auf diesen Daten schwerlich definiert. Es könnte sein, das die Membranextraktion in DDM und die Aufreinigung die Struktur des Komplexes destabilisieren, deshalb sollten zukünftige Arbeiten den Fokus nur darauf legen, die Struktur nur in Lipidsystemumgebungen zu ermitteln.

Chapter A: Introduction

Proteins play a key role in the functioning of living organisms and their importance was described for the first time in 1926 by Sumner (Sumner, 1926). Ever since this time, resolving structures of proteins has been the milestone of understanding bimolecular mechanisms that lie behind the functioning of living organisms.

Proteins can be classified into three main classes: the globular proteins or soluble proteins, fibrous proteins and the membrane proteins associated with lipid bilayer membranes (Alberts *et al.*, 2002).

Up to very recently, X-ray crystallography and NMR were the leading tools to resolve the atomic structures of proteins (Rhodes, 2000; Drenth, 1999). But today, with the new horizon that electron microscopy is opening (Kuhlbrandt, 2014), the hope of solving structures of the most difficult proteins and protein complexes by Cryo-EM has been revived as never before.

Understanding how those proteins are synthesized and delivered to their target cells compartment is a key issue that may help to profoundly understand mechanisms that control the functioning of living organisms.

The introduction of this work gives a general insight of membrane proteins, moreover how they are targeted and inserted within the membrane. Then, a particular description of a class of membrane protein known as tail anchored membrane proteins, their topology and how they are inserted into the membrane.

1.1 Membrane Proteins

Membrane proteins are essential for the functioning of any living organism. They are involved in the movement of molecules such as vitamins, ions, nutrients and drugs across the lipid bilayer. They also play a role in transduction of energy and signals, cell-cell communication (Whitelegge, 2013). They can be channels, transporters, receptors and ion pumps. Additionally, they are major drug targets.

Membrane proteins represent 20-30% of all genes in most of the genomes (Krogh *et al.*, 2001).

Membrane proteins are defined as integral membrane proteins when they are inserted in the lipid bilayer in such way that one or more non-polar domains, also known as a transmembrane domain (TMD), span the hydrophobic moieties of the lipid bilayer (Figure1.1). The TMD can be a ß-sheet or α-helix.

The integral membrane proteins can be sub-classified in four types; the integral monotopic proteins which consists of one α-helical polypeptide attached to one side of the lipid bilayer, the integral bitopic type I proteins which are characterized by single pass TMD and have in addition their carboxyl-terminus present in the cytosol, while the integral bitopic type II proteins are single pass TMD where the amino-terminus is found in the cytosol, and the last type of integral membrane proteins which consist of more than one TMD spanning the lipid bilayer and known as the integral polytopic proteins. These are divided in two sub-groups: the helix bundles present in all types of biological membranes and the beta barrel which are found in the outer membranes of bacteria, mitochondria and chloroplasts.

The membrane proteins are peripheral when they are attached to the surface of the lipid bilayer non-covalently, through either other membrane proteins, or through lipids via an α helix or beta sheet (called lipid-anchored proteins) or via glycosylphosphatidylinositol (GPI).

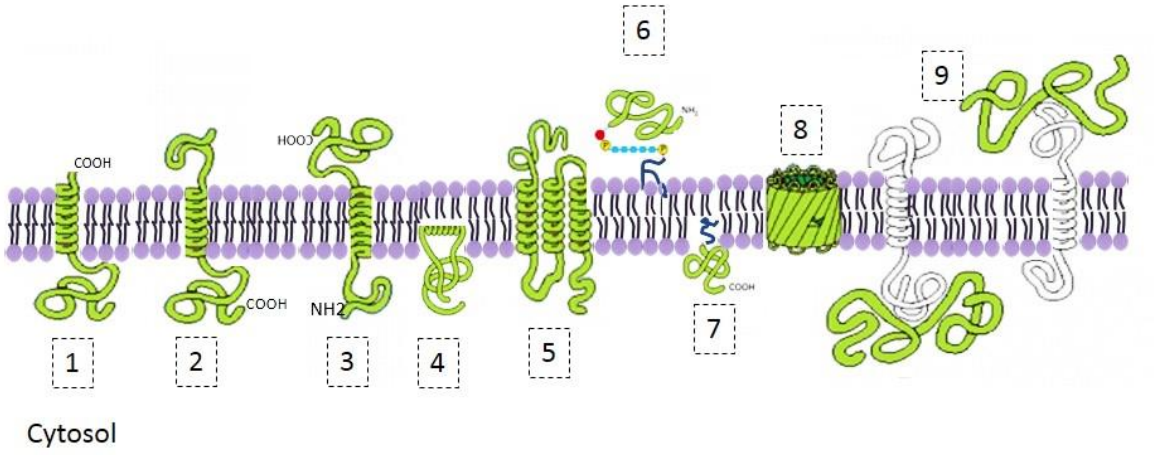


Figure 1.1: Schematic overview of integral membrane protein within the lipid bilayer. 1) Tail anchored protein, 2) Type I bitopic integral membrane protein, 3) Type 2 bitopic integral membrane protein , 4) Monotopic integral membrane protein, 5) Polytopic integral membrane protein with helix bundles, 6) GPI anchor 7) fatty acid anchor and 9) peripheral membrane proteins attached weakly to another integral membrane protein (Figure adapted from Alberts *et al*, 2002).

Lipid Bilayer

The amphiphilic nature of lipid bilayer contributes to the stability of the functional membrane proteins and the impermeability of the biological membranes allows most molecules to pass only with the assistance of the membrane proteins.

The lipid bilayer consists of two opposed hydrophobic chains connected to polar head groups which is in contact with aqueous environment. The transmembrane or integral proteins are stabilized in the membrane via hydrophobic interactions between the hydrophobic tails of the lipid bilayer and the hydrophobic domain of the transmembrane proteins.

1.2 Tail-Anchored Membrane Proteins

The early stage of the post biosynthesis of eukaryotic integral membrane proteins is their targeting, followed by their insertion into the membrane of either the endoplasmic reticulum (ER), the mitochondria or the chloroplasts before they are transported to other organelle membranes that do not have any membrane insertion apparatus such as vesicles, Golgi complex, or plasma membrane (Kutay *et al.*, 1993). The majority of

integral membrane proteins have an amino-terminal signal sequence or a hydrophobic signal at the N-terminus that triggers the integral membrane protein to anchor to the lipid bilayer. A minority of integral membrane proteins known as Tail-anchored (TA) proteins (Figure 1.1) have a tail of single hydrophobic transmembrane region near the C-terminus. The single hydrophobic C-terminal sequence of TA proteins is a signal that helps the targeting of TA proteins to the membrane before their insertion to the membrane (Plath and Rapoport, 2000). A classical single TMD is about more or less 20 residues, and the TMD of TA proteins are found within 50 residues of the C-terminus (Kriechbaumer *et al.,* 2009). Due to this specific topology of TA proteins, their biosynthesis diverge from the majority of integral membrane proteins; the biosynthesis of TA proteins is completed while still in the cytosol, and releases from ribosome in the cytosol. This process is known as Post-translation insertion of integral membrane proteins which is different from the well-known classical Co-translation insertion. TA proteins are found almost in all kingdoms of life (Table 1.2.1).

TABLE 1.2.1: NUMBER OF TA PROTEINS ACROSS EVOLUTIONARY ORGANISMS (ADAPTED FROM *BORGESE AND RIGHI, 2010*)

ORGANISMS	Number of TA proteins	References
HOMO. SAPIENS	**411**	Kalbfleisch *et al.,* 2007
SACCHAROYCES CEREVISIAE	**55**	*Beilharz et al., 2003*
ARABIDOPSIS. THALIANA	**454**	*Kriechbaumer al, 2009*
ESCHERICHIA. COLI	**11**	*Borgese and Righi, 2010*
METHANOCOCCUS. MARIPALUDIS	**12**	*Borgese and Righi, 2010*
RICKETTSIA. PROWAZEKII	**7**	*Borgese and Righi, 2010*

TA proteins are involved in numerous biological processes such as vesicular transport and fusion (Ungar and Hughson, 2003), electron transfer (Pedrazzini *et al*., 2000), protein translocation (Osborne et al., 2005) and in the regulation apoptosis (Hockenbery *et al*., 1990). TA proteins are also associated to other varieties of biological processes. They are primary localized in compartments such as the chloroplast, mitochondria and the ER from where they are transported to the membranes of other organelles such as the Golgi, the endosomes, the nuclear envelope, the vacuole, synaptic vesicles, the lysosomes, the

peroxisomes and the plasma membranes. TA proteins are found not only along the evolution of organisms but also localized in all types of membranes where they have different functions as summarized in Table 1.2.2.

TABLE 1.2.2: LOCALIZATION AND FUNCTIONS OF TA PROTEINS

LOCALIZATION	TA proteins	Function
ER	Cytochrome b5	Enzymatic
	Heme oxygenase I and II	Enzymatic
	Sec61γ and sec61β	Translocation
	UBC6	Enzymatic
	Bcl-2	Regulation of apoptosis
	FRT1, FRT2	Translocation
TRANSPORT VESICLES GOLGI COMPLEX	Giantin	Tethering proteins
PEROXISOMES	Pex15p	Tanslocation
TRANS-GOLGI NETWORK	Us9 protein of α hespes virus	Constituent of viral envelope
TARGET MEMBRANES FOR VESICULAR FUSION	Syntaxins, synaptobrevins	SNARE protein

1.3 Biogenesis of membrane proteins

After or during the biogenesis of integral membrane proteins, the targeting and the insertion of integral membrane proteins into membranes are two fundamental steps that help all integral membrane proteins occupy their position within their target cell or organelle membrane to consequently accomplish their countless functions. Those two mechanisms require hydrophobic peptide signals at a certain position of the secondary structure of the protein sequence. The targeting signal triggers the mode of the targeting and insertion of the membrane proteins to the membrane. Both process often require the assistance of chaperones that maintain the stability of the nascent hydrophobic polypeptide exiting the ribosomal tunnel. Thereafter, the insertion is completed and the translocation toward other organelles can then take place. The integral membrane proteins are inserted into the membrane by either the co-translation insertion mode or by the post-translation insertion mode.

1.3.1 Co-Translation Insertion of Membrane Proteins

The co-translation insertion pathway is extraordinarily conserved from bacteria to mammalian cells. Bacteria (Dalbey and Chen, 2004) and eukaryotic cells (Keenan *et al.*, 2001) signal recognition particle (SRP) molecules are involved in recognizing a targeting signal sequence that directs the nascent integral membrane protein from the ribosome to the membrane. The co-translation insertion pathway is also known as SRP pathway.

In bacteria, the minimal SRP complex consists of a 4.5S RNA, a GTPase Ffh protein (de Gier *et al.*, 1996; Dalbey and Chen, 2004) which has guanosine-5'-triphosphate (GTP) binding domain (Figure 1.3.1.1) and YidC (Samuelson *et al.*, 2000). The SRP complex interacts with SRP receptor, FtsY (Miller *et al.*, 1994) in a GTP-depend manner.

FtsY is an essential protein for the biosynthesis of integral membrane proteins in bacteria (Saluanova and Bibi, 1997) and it also has a GTP binding domain (Miller *et al.*, 1994). Ffh and FtsY have structural similarities (Figure 1.3.1). The interaction of the SRP complex with FtsY through the NG domain stimulates the hydrolysis of GTP, while any SRP lacking NG domain fails to target to the membrane.

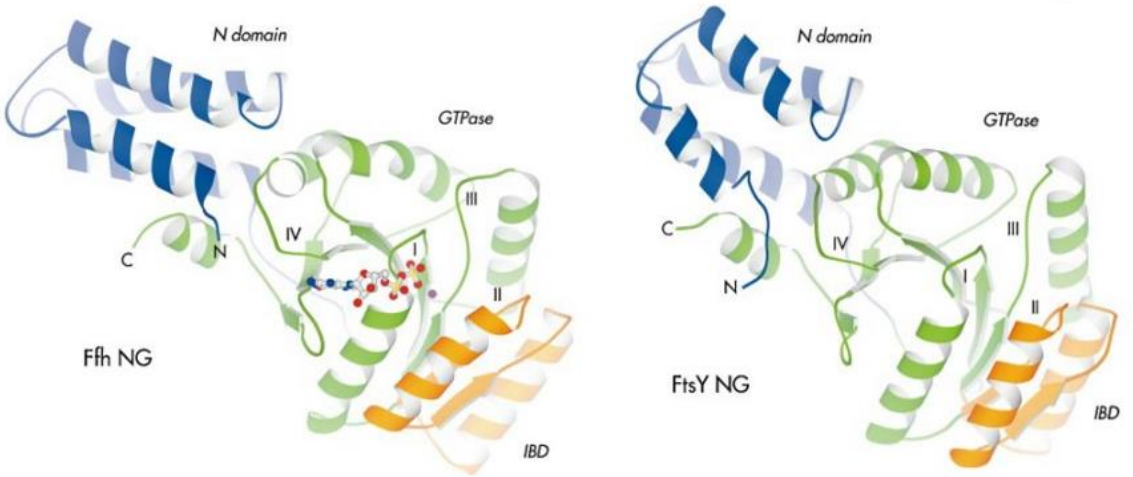


Figure 1.3.1.1: Structural similarity of NG domain of Ffh (T. aquaticus) bound to Mg2+ GDP and NG domain of FtsY in apo form (E. coli). Ffh and FtsY have three domains. The N domain (blue), the G domain (green) with GTPase fold with four conserved sequence motif (I-IV) arranged around the nucleotide-binding site and the insertion box (orange) (Keenan *et al*., 2001).

A SecYEG translocation channel is present within the membrane to actively drive the efficient delivery and unloading of the complex of the ribosome-nascent protein to the target membrane (Akopian *et al.*, 2009) in order to initiate the insertion of the nascent polypeptide (Park *et al.*, 2014). The interaction of the complex ribosome-SRP-nascent protein with SecYEG initiates the conformation changes of the complex SRP-FtsY leading to the hydrolysis of GTP by both Ffh and FtsY.

The early translocation intensively involves SecY, a subunit of SecYEG. SecY has 10 α-helical TMD and a molecular weight of 48 KDa (Rensing and Maier, 1994). Structural evidence shows that SecE stabilizes SecY. *E. coli* SecE has 14 KDa and 3 α-helical TMD, and is essential for protein transport and cell viability (Schatz *et al.*, 1991). SecG has two α-helical TMD connected by a cytoplasmic loop and it is 12 KDa protein. Within the membrane, the nascent chain binds to the cytoplasmic surface of SecY and opens the SecY channel. The nascent chain is then inserted as a loop into SecY channel with the hydrophobic signal sequence placed into the open lateral gate (Park *et al.*, 2014).

The *in vivo* and *in vitro* site-directed cross-linking approach of the SecY-YidC interface showed that YidC is in contact with all four transmembrane domains of the lateral gate (Sachelaru *et al.*, 2013), which also shows the importance of YidC during the insertion of the membrane proteins

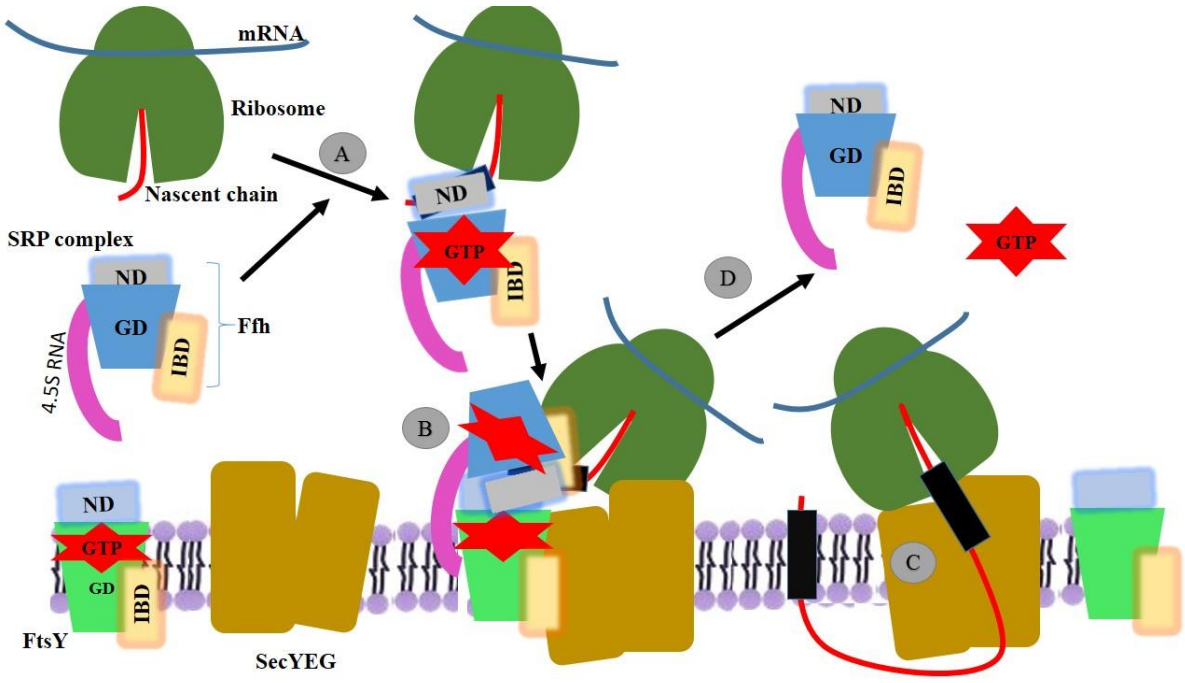


Figure 1.3.1.2: Overview of the SRP pathway schematic in *E. coli*. SRP complex binds first to the nascent peptide (A) and the complex is targeted to the membrane where it binds first to FtsY bound to GTP. From their FtsY interacts to FfH through their N domains (ND) (B), and subsequently the complex ribosome-nascent peptide-SRP-GTP-FtsY through the G domain (GD), induces GTPase activity of each protein causing the dissociation of the complex (D). The complex ribosome-nascent peptide is transferred to the translocase SecYEG where the translation is resumed (C). The overview of the SRP pathway was prepared from the literature cited in the chapter.

The mammalian SRP complex counterpart consists of six members namely SRP 72, 68, 54, 19, 14, and 9, and the 7SL RNA (Figure 1.3.1.6). The 7SL RNA has a domain called Alu domain, and is required for the elongation arrest of the ribonucleoprotein complex (Poritz *et al.*, 1988). In *E. coli* 4.5S RNA there is no Alu domain present.

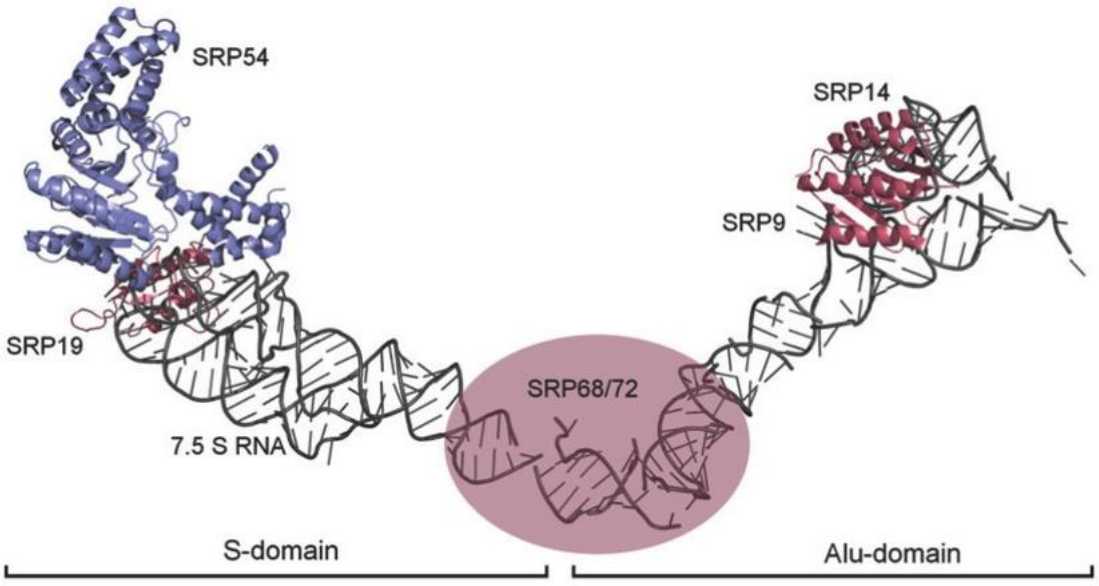


Figure 1.3.1.3: SRP structure in mammalian cells (Denk *et al*., 2014).

SRP54 is a homologous of Ffh from *E. Coli* and it binds first to the nascent protein chain (Keenan *et al.,* 2001) like in bacteria. SRP54 consists of two domains, an amino-terminal domain of 33 KD that has a GTP-binding site and a carboxy-terminal domain of 22 kD (Zopf *et al.,* 1993).

The mammalian SRP receptor (SR) is a heterodimeric protein composed of a 69 KDa α-subnit (SRα) and a 30 KD β-subunit (SRβ), both of which have GTPase activity (Miller *et al.,* 1995). SRα is homologous to the bacteria FtsY, as well as GTPase activity. It is closely related in sequence, structure, and function to the GTPase domain of SRP54. However, SRβ is an integral membrane protein anchored to the ER membrane with just a single transmembrane region. SRβ belongs to a GTPase sub-family related to ARF and Sar1. SRα and SRβ are associated. In eukaryotic cells, SRβ is required for the functioning of the SR as mutation in the GTPase domain destroys the function of the SR *in vivo* (Ogg *et al.,* 1998).

As conclusion for co-translation insertion, it is important to mention that in bacteria, as well as in mammalian cells, GTP binding is required for the initial targeting steps. The GTP hydrolysis is induced by the interactions of Ffh and FtsY in bacteria. In eukaryotic cells, nevertheless, in order to associate with SRα, the SRβ must be in GTP bound form to bind the ribosome cargo complex therefor leading to the hydrolysis of GTP by SRβ

and the releases of SRP from SR can occur simultaneously (Keenan *et al.*, 2005). In mammalian cells, the nascent protein is inserted into the ER membrane by several proteins complexes; the translocase Sec61 complex, which consists of Sec61α, Sec61β and Sec61γ (Sec61p, Sbh1p and Sss1p in *S. cerevisiae*) corresponding to the bacteria SecYEG. Sec61 complex acts as a ribosome receptor as well as a signal peptide receptor forming a pore for the membrane protein insertion (Zimmermann *et al.*, 2006).

1.3.2 Post-Translation insertion of Membrane Proteins

During the protein biosynthesis, the entire ribosome can accommodate about 40 amino acids during the biosynthesis of protein while the TMD consist of 20 to 25 amino acids.

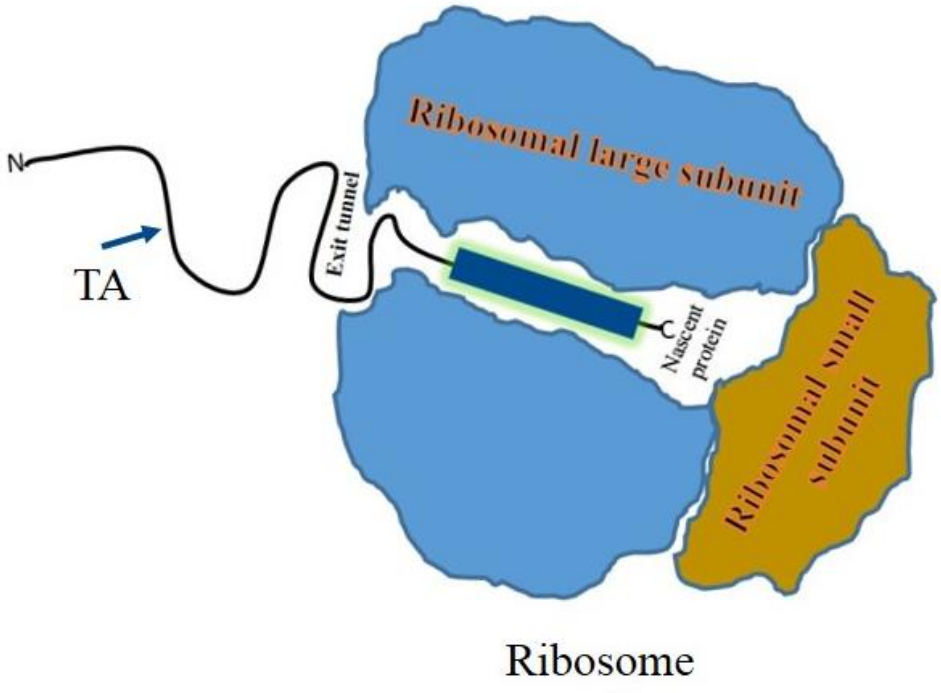


Figure 1.3.2: Ribosome association during the biosynthesis of TA proteins. Integral membrane proteins that have a tail of single hydrophobic transmembrane region near the C-terminus.

Considering that the TA proteins have only a single TMD near the C-terminus, meaning within the last 50 amino acids residues, the biosynthesis of TA proteins finishes while it is still in the ribosome (Figure 1.3.2), without being able to be recognized by the SRP molecules for the TA proteins membrane targeting. Though, to stabilize the newly

synthesized TA protein in the cytosol, various chaperones bind to the new TA protein prior to its post-translation insertion to the membrane.

In the mid-90s, evidence of the post-translational insertion of integral membrane proteins was shown for the first time *in vitro* (Kutay *et al.,* 1995). As opposed to the SRP pathway which uses GTP as energy source for the co-translational insertion of integral membrane protein to occur, ATP is required for the post-translation insertion. One of the soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins, synaptobrevin (TA protein), which is also a vesicle-associated membrane protein (VAMP), was used to show ATP is required for post-translation insertion. Synaptobrevin was not directly inserted into the synatique-like vesicles (target organelle) in the *in vitro* experiment, but it was firstly inserted to the ER before it was transported through the Golgi (Kutay *et al.,* 1995).

As most molecular details of the mechanisms involved in post-translation insertion of membrane proteins have focused on eukaryote, the present description will focus on the TA post-translation insertion in eukaryotic cells.

There are four major possible pathways for the post-translation insertion of TA proteins, namely the unassisted mechanism that does not required any insertion machinery, the heat shock protein 70 (HSP) pathway, the SRP pathway and the Guided entry of tail anchored proteins or GET (TMD recognition complex 40 or TRC40 in mammal) pathway (Hegde and Keenan, 2011).

1.3.2.1 Unassisted Post-translation Insertion

Using TA protein cytochrome b5, it was initially shown that selective depletion of components involved in co-translational protein translocation had no effect on the insertion of cytochrome b5. This was a strong hint that the targeting of TA proteins in the ER may also occur in a ribosome-independent pathway. Interestingly, without GTP nor ATP, when the cytochrome b5 was tagged with N-glycosylation consensus site (b5-Nglyc) and all components of signal peptide-driven translocation (Sec61p, Sec62p, Sec63p, or Bip/Kar2p, or Sbh1p/1p and Lhs1p in yeast) were non-functional, the efficient translocation could still happen (Yabal *et al.,* 2003), therefore showing that the C-

terminal tail of TA proteins can be translocated without the participation of Sec61 complex. It was experimentally shown later in details that TA proteins with a moderate hydrophobicity of the TMD can enter the lipid bilayer without the assistance of any insertion machinery at the membrane (Brambillasca *et al.*, 2005).

1.3.2.2 Molecular Chaperones Hsp40/Hsc70 Pathway

Hsp70 molecular chaperones and other chaperones are intensively involved in the biogenesis of proteins by guiding their folding and also aiding their transport to their target site or compartment. For instance, during protein folding, Hsp70 binds to the exposed hydrophobic sequence and prevents their irreversible aggregation (Rüdiger *et al.*, 1997; Flynn *et al.*, 1991; Langer *et al.*, 1992). While Hsp40 stimulates the ATPase activity of Hsp70 (Liberek *et al.*, 1991), it preferentially binds hydrophobic peptides (Rüdiger *et al.*, 2001). Additionally, Hsp70 facilitates the translocation of secretory and mitochondrial precursor polypeptides (Deshaies *et al.*, 1988). Hsc70 and Hsp70 belongs to Hsp70s family. They have distinct ATPase domains that differ given that the first is constitutively expressed while the second, the Hsp70 is only expressed when exposed to stress (Tutar *et al.*, 2006).

In the process of understanding the mechanisms involved in the biogenesis of TA proteins, the interaction of TA proteins with two cytosolic chaperones Hsp40 and Hsc70 was first found. Hsp40/Hsc70 complex showed that it facilitate the insertion of the sec61β, a TA protein, in an ATP-dependent manner (Abell *et al.*, 2007). The inhibition of the complex Hsp40/Hsc70 showed that the complex is required for the integration of a subset of TA proteins.

As mentioned earlier for the unassisted pathway as well as other pathways, the hydrophobicity of TMD of TA proteins is the deciding factor on the usage of the Hsp40/Hsc70 pathway or not, therefore TA proteins with a low hydrophobic index use the Hsp40/Hsc70 route for membrane targeting (Rabu *et al.*, 2008).

1.3.2.3 Post-translation Insertion TA Proteins by SRP Pathway

Although the SRP route is mainly used for the delivery of secretory and most integral membrane proteins by co-translation insertion, a couple of reports suggested that

eukaryotic SRP molecules can also support the delivery of TA proteins to the ER (Abell *et al.*, 2004, and 2002) in a GTP, and SRP-receptor dependent mode. Indeed Abell and colleagues using TA proteins Sec61β and synaptobrevin 2 showed that SRP molecules were interestingly associated with those two TA proteins after their complete biosynthesis. Functionally, they found that TA proteins can be targeted to ER membranes by a route that requires SRP, SRP receptor and GTP. Although ribosome did not play a role throughout the targeting to membrane in this process, the SRP presence was required during the release of the TA proteins from the ribosome (Abell *et al.*, 2004).

1.4. Get Pathway

Almost 10 years ago the cytosolic TMD recognition complex (TRC) that targets TA proteins in a post-translational insertion manner into the ER membrane was revealed (Stefanovic and Hegde, 2007). A highly conserved cytosolic 40 KDa ATPase named TRC40 was identified as the central player of cytosolic targeting machinery (Stefanovic and Hedge, 2007; Favaloro *et al.*, 2008). TRC40 is also known as Asna-1. Asna-1 is homologous to the *Escherichia coli* ArsA and *Saccharomyces cerevisiae* Arr4. ArsA is a catalytic subunit of ArsAB pump that confers resistance to arsenicals. In presence of arsenites, ArsA interacts with ArsB, and then couple the hydrolysis of ATP with the extrusion of the metalloids (Chen *et al.*, 1986; Mukhopadhyay *et al.*, 2006).

The first structure of *E. coli* ArsA shows the ATPase consists of two homologues domain of L-shape (dimer) and two nucleotide-binding sites (NBS), and by binding of arsenite its ATPase activity is activated. The NBS are located at the interface of both domain (Zhou *et al.*, 2000). ArsA was stabilized by Sb (III) and Mg^{2+}ADP.

In yeast, Arr4p was renamed Get3, Arr4p was earlier shown to have physical interaction with Mdm39 (or Get1) and Rmd7 (or Get2) (Schuldiner *et al.*, 2005). This discovery led to the identification of other three players Get4, Get5 (Jonikas *et al.*, 2009) and the small Glu-rich tetratricopeptide repeat-containing 2 (Sgt2) (Costanzo *et al.*, 2010) involved in the biogenesis of TA proteins using the Get pathway.

The key players in the Get pathway for the targeting and the insertion of TA proteins are Sgt2, Get3, Get4, and Get5 respectively for TA targeting also known as the early step

(targeting) and then Get1 and Get2 for the incorporation of the TA proteins to the ER or the late step (Insertion).

The ATPase Get3 is a central player during the targeting. Many Get3 structures are available in the literature, Get3 was initially revealed to have two main structural states as illustrated by the yeast Get3 structures (Mateja *et al.,* 2009) and others intermediate states. Functional Get3 is mainly a dimer that is stabilized by zinc ion. Like ArsA, each Get3 monomer has a NBS site which consists of the dimerization interface and a methionine-rich part, helical domain. The two main states are indeed either the nucleotide-free (open) or nucleotide-bound state (closed). Each Get3 has 13 alpha helices and 7 beta strands (Figure 1.4.2, Hedge *et al.*, 2011).

Furthermore, other structural studies on archaea homologues (Get3 from *Methanocaldococcus jannaschii* (MjGet3)) suggest that Get3 is Tetramer (Figure 1.4.3) (Suloway *et al.*, 2012). Suloway and colleagues at that time claimed and hypothesized that the Get3 tetramer was likely the functional state that binds to the TA substrate while it completely shields and maintains the solubility of TMD of the TA proteins.

However, it was shown recently that this is not the case in gradient sucrose experiments and the functional species of Get3 that binds to TA proteins, targeting them to the ER membrane, is only the dimeric form. In that set of experiments it is shown that the Get3 dimer and the complex Get3-TA protein co-migrate about the same volume in comparison to the higher organized Get3 complexes (higher order tetramer) which migrate at higher volume. Besides, they showed that Get3-D57 has the same structural conformation as the closed ADP.AlF4$^-$-bound Get3, hence the conclusive remarks that the higher-order assembly is unlikely involved in the targeting complex formation (Mateja *et al.,* 2015).

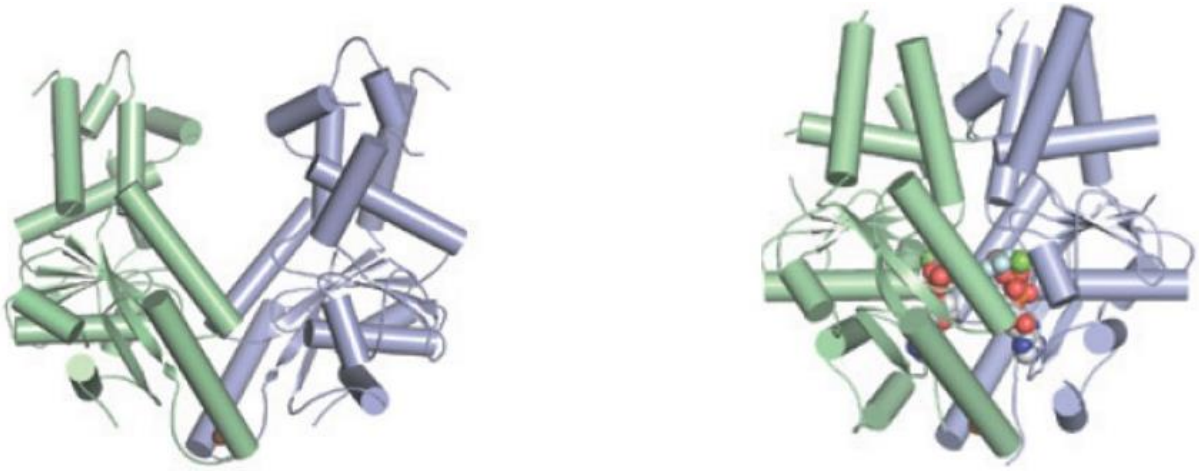


Figure 1.4.2: Symmetric homodimer Get3. The *S. pombe* open nucleotide-free Get3 (left) and the *S. cerevisiae* closed ADP.AlF4$^-$-bound Get3 (right) (Mateja *et al.*, 2009; Hedge *et al.*, 2011). Each monomer is divided into two subdomains: the ATPase core subdomain and the α-helical subdomain.

During the TA protein targeting, besides interchanging from active dimer to the tetramer, Get3 ATPase assumes multiple conformations (Get3 ATPase cycle). Namely, the `open´, `semi-open´, `semi-closed´, `closed´ dimers and the tetrameric conformations throughout its activation by ATP (Mateja *et al.*, 2009; Rome *et al.*, 2013).

Upon nucleotide binding, Get3 changes its conformation. The dimer interface of Get3 has a relatively large hydrophobic region in contrast to the nucleotide-free state where the same region is disrupted and replaced by a large charged cleft of about 20 Å (Mateja *et al.*, 2009).

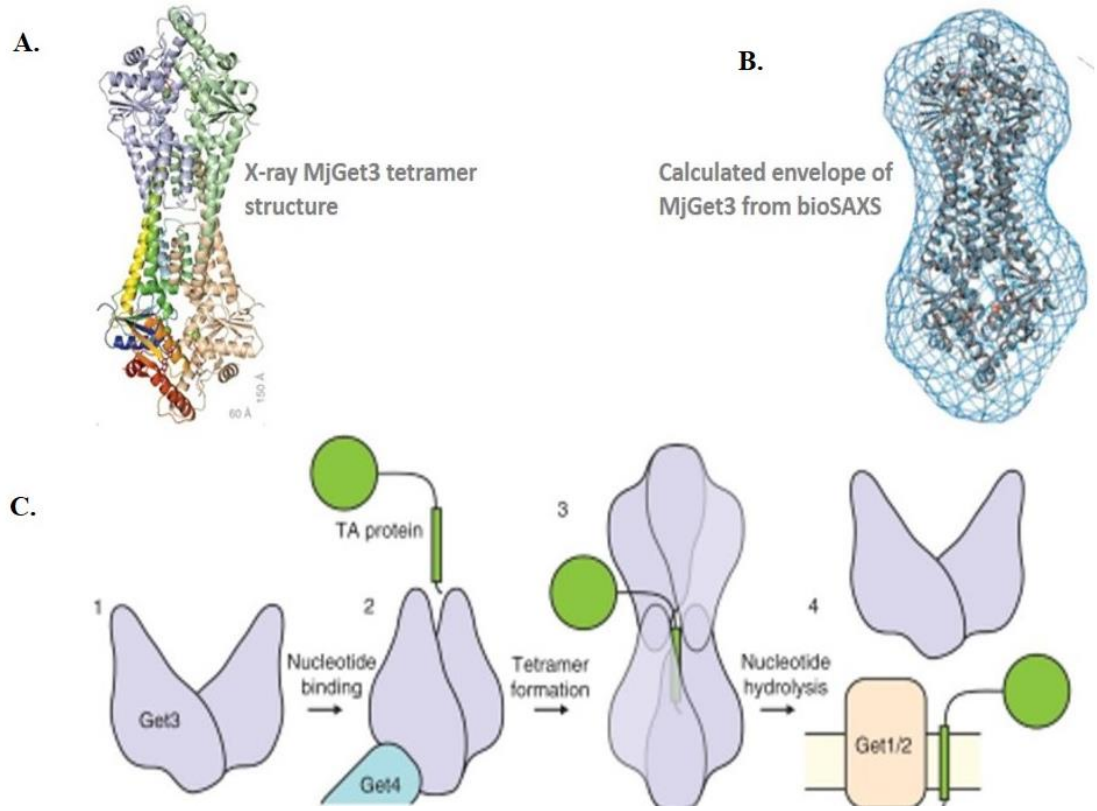


Figure 1.3.3: Fungal Get3 tetramer model and working hypothesis of TA targeting to the membrane by Get3. A. is the crystal structure and B is the calculated envelop of MjGet3 from bioSAXS data, both structural information shows that the fungal Get3 is likely tetrameric. C. is a model showing how the dimeric Get3 in open conformation upon nucleotides bind is stabilized by a second dimer when it binds the TA protein, forming a tetramer of Get3. The hydrophobic binding chamber completely protects the hydrophobic part of the TA protein prior its assisted insertion by Get1 and Get2 complex. Figures adapted and taken from Suloway *et al.*, 2012.

1.4.1 Pre-targeting of TA Proteins

In the initial loading step, Get4 and Get5 escort TA proteins to Get3 (Wang *et al.*, 2010). Beforehand, after the biosynthesis and the complete exit of TA proteins from the ribosome take place, TA proteins are kept soluble through their initial direct interaction with the heat-shock protein co-chaperone, Sgt2. The complex Sgt2/TA protein is then transferred to Get3 by the mediation of Get4/Get5. The closed homodimer of Get3 bound to ATP is further stabilized by Get4 in a C2 symmetry manner (Gristick *et al.*, 2015).

Interestingly structural data shows that one Sgt2 dimer interacts with one Get5 monomer (Tung *et al.*, 2013), assuming that it is by this way that the complex Get4-Get5-Get3 binds ATP and the Sgt2-TA complex comes in proximity to deliver the substrate (TA

proteins) onto Get3 (Wang *et al.*, 2010). Therefore, Get4 stabilizes Get3 on the one hand, while on the other Ge5 interacts with Sgt2 to hand in the TA proteins to Get3.

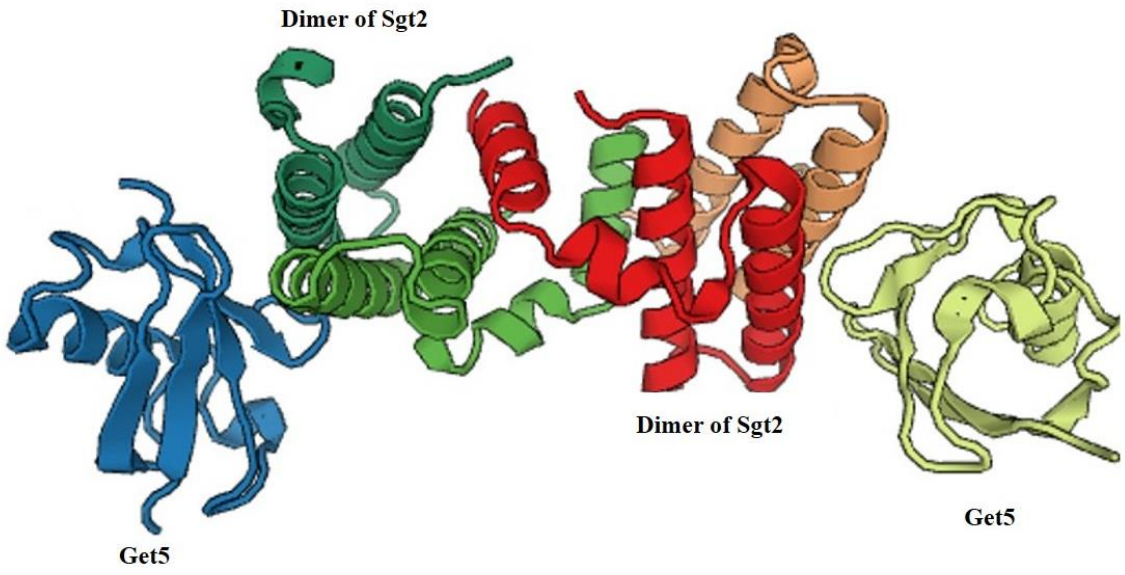


Figure 1.4.1.1: Overall structure of each Get5 monomer with each Sgt2 dimer (Tung et al., 2013). In the structure TA protein is not seen. This interaction is the link that enable the transfer of the TA proteins from Sgt2 to Get3 bound to ATP. While Get3 interacts with Get4 in the one hand, Get5 interacts with Sgt2 in the other.

Once both pre-targeting complexes (Get4-Get5-Get3 bound ATP and Sgt2-TA protein) have been brought in close proximity, the substrate (TA-proteins) is then transferred to Get3 (Figure 1.4.1.1). The structures of Get3 homodimer bound to TA proteins showed how the α-helical TMD is embedded within the hydrophobic groove (Figure 1.4.1.2), Mateja *et al.*, 2015).

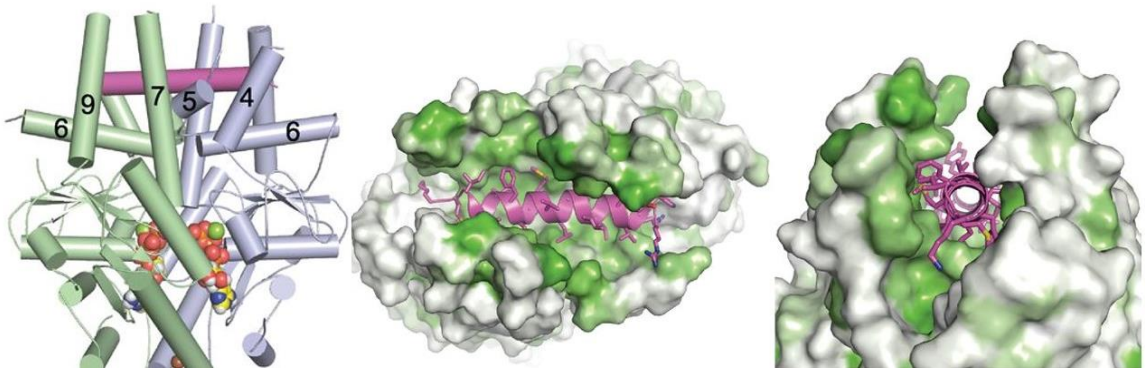


Figure 1.4.1.1: Structure of *S. cerevisiae* Get3-truncated Pep12 TA protein bound to nucleotides and the surface representation of the TA proteins within the hydrophobic groove of Get3 homodimer (Mateja *et al.*, 2015).

1.4.2 The targeting of TA proteins

Once the targeting complex is formed, ATP is hydrolyzed to ADP + Pi then the complex Get3-TA protein can then bind to Get2 via its N-terminal flexible domain (Stefer *et al.*, 2011). Two Get2 bind symmetrically to form two closed Get3 (Figure 1.4.2.1) in such way that each Get2 binds the homodimer Get3 away from the dimer interface (Stefer *et al.*, 2011). Three conserved and negatively charged residues of Get3 Asp 265, Glu 307 and Asp 308 of Get3 are involved in direct interaction with the conserved RERR motif of the alpha helices 1 of Get2 (Mariappan *et al.*, 2011; Stefer *et al.*, 2011). Beside the negatively charged residues of Get3 involved in Get2 interaction, the additional DELYED motif is crucial for TA protein insertion. The DELYED motif of Get3 is wrapped by Get2 (Figure 1.4.2.1). Although the structural data were without bound TA, this model might be applied for the interaction of Get2 with Get3-TA.

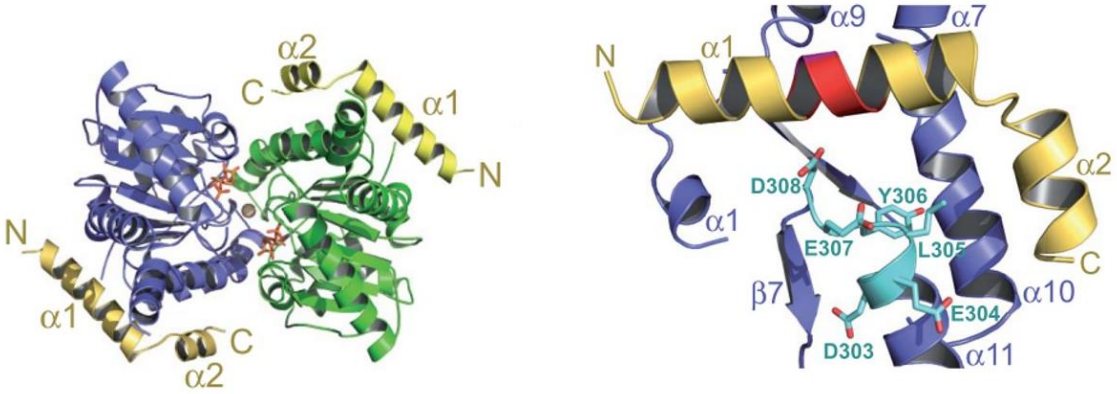


Figure 1.4.2.1: Structure of Get3-cytosolic Get2. The left structure shows how Get2 (yellow) is bound to the outside of closed Get3 (green and blue) bound to ADP-Mg2+. The right structure is Ge3/Get2 interphase. The conserved RERR motifs (red) of helix α1 of Get2 form ionic interactions with conserved negatively charged residues D265 , E307, D308 and a DELYED of Get3 (Stefer *et al.*, 2011).

After the targeting complex is recruited to Get1 by Get2, then the homodimer of Get1 can interact with homodimer of Get3 in order to insert and release the TA protein into the membrane of the ER. Each cytosolic domain of Get1 homodimer has an antiparallel coiled-coil structure (helices α1 and α2) that binds deep into the interface of the homodimer of Get3, opening the homodimer of Get3 throughout this same process (Stefer *et al.*, 2011). The flexible TA proteins binding domain (TABD) is well defined in the complex (Figure 1.4.2.2). One monomer of Get3 has the same interface with Get1 (Figure 1.4.2.2), also called interface I (Stefer *et al.*, 2011). This interface I is formed by the helix α2 (Y65, A66, and W68) of Get1 and the helices α10 (F246, L249, and Y250) and α11 (Y298 and L305) of the NBS (or nucleotide binding domain, NBD) of Get3 (Figure 1.5.8). A DELYED motif of Get3 is also involved in the interaction with Get1. The interface II shows a small contact area of the other monomer of Get3 (helix α4 and α5 of Get3 TABD) with Get1 and the six-residue loop formed by Get1 helices (Mariappan *et al.*, 2011; Stefer *et al.*, 2011). However, overlapping binding sites between Get1 and Get2 suggest that Get1 and Get2 cannot occupy the same site on Get3. Most the past work showed that the ATP hydrolysis by Get3 is needed to expose the Get1 binding site because when ATP is bound to Get3, the shared binding site of Get1 into Get3 is buried while the same Get1 binding site is exposed to solvent in the Mg^{2+} -ADP bound state (Mariappan *et al.*, 2011).

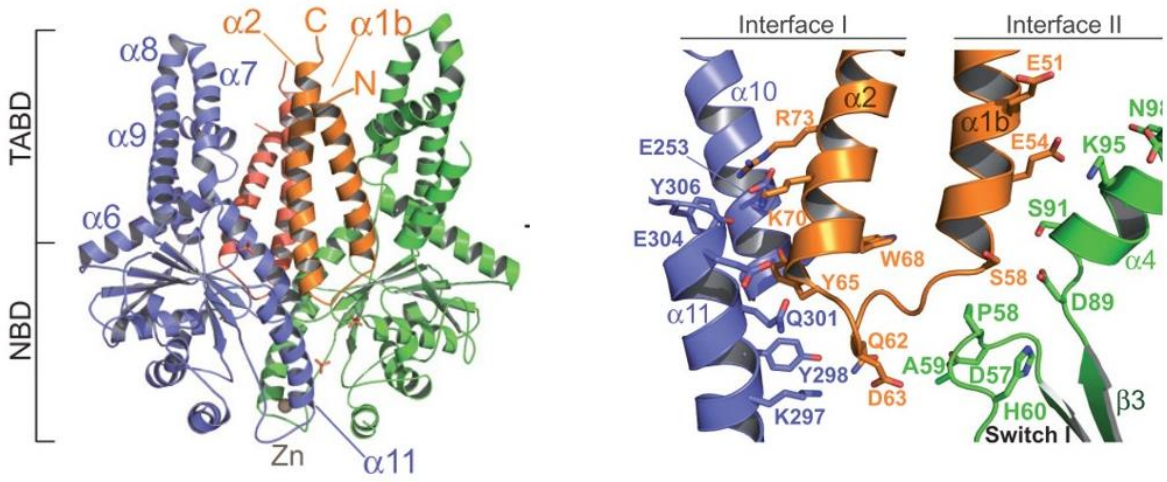


Figure 1.5.8: Structure of Get3-cytosolic Get1. Get3 is in blue and green while Get1 is in orange. The left figure shows the overall structure of Get3 and the cytosolic Get1 with the Zn ion Get3-bound. The right figure interfaces (I and II) shared between Get3 and the cytosolic Get1 (Stefer *et al.*, 2011).

1.4.3 TA proteins insertion by Get1 and Get2

As mention earlier Get1 and Get2 are the Get3 ER receptors, therefore they are responsible for the insertion of TA proteins in the ER of yeast cells. Get1 and Get2 have three TMDs each. In addition, Get2 has a long N-terminal flexible part. Get1 consists of 235 amino acid residues which corresponds to a size of 27 Kda. Get2 has 285 amino acid residues and a size of 31.5 KDa.

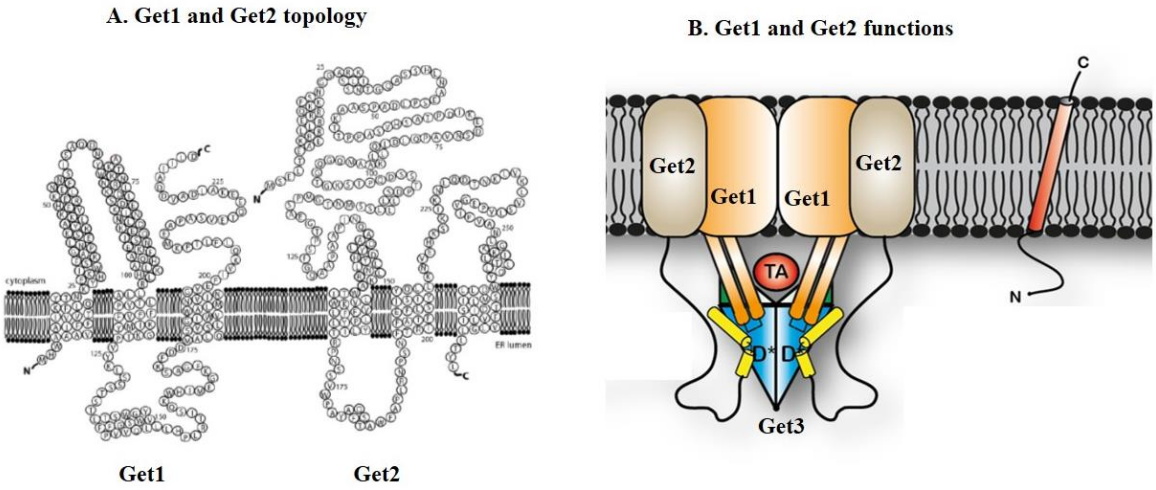


Figure 1.4.3: Topology and the function Get1 and Get2. A. shows how Get1 and Get2 TMDs are imbedded topologically within the ER membrane while B. is the schematic of the complex Get2/Get1 and Get3 bound TA protein before the insertion, Get2 and Get1 facilitating the insertion by capturing the TA proteins (Stefer *et al.,* 2011).

Get1 and Get2 are homologous of the tryptophan-rich basic protein (WRB) and the calcium-modulating cyclophilin ligand (CAML) in mammals. (Colombo *et al.*, 2015; Wang *et al.,* 2014). Before the beginning of this thesis, despite the lack of experiment data showing that Get1 and Get2 are both dimer, it is generally assumed they are forming a heterotetramer based on the studies done with the cytosolic part. In contrast to the delivery of TA proteins to the ER that has been well characterized, a lot of biochemical and structural studies remain to be done so as to understand the insertion into the ER membrane. Recent biochemical and biomolecular studies have revealed one important function of Get1 and Get2. Get1 and Get2 TMDs have insertase function that facilitates the entry of the TA protein into the lipid bilayer (Wang *et al.,* 2014). Mutations in the TMDs of Get1 and Get2 showed that TA proteins cannot be discharged by Get3 despite intact cytosolic domains of Get1 and Get2 that bring the Get3 bound to TA in proximity of the lipid bilayer. The TMDs of Get1 and Get2 complex capture the TMD of the TA proteins and facilitate their incorporation into the lipid bilayer. Therefore, during the insertion there is a collaboration between the cytosolic domain and the TMDs of Get1 and Get2 that permit an efficient release of the TA proteins from Get3 (Wang *et al.*, 2014).

1.6 The aims of this study

A lot of work has been done using the unstructured cytosolic domains of Get1 and Get2 with regard to their interaction with Get3.

Up to today, nothing was published regarding the stoichiometry of Get1 and Get2, and little structural information on the interaction of Get1 and Get2, or their interaction with Get3, is known.

In order to understand the molecular mechanisms that control the insertion of TA proteins at the ER of yeast, the aim was first to express and purify Get1 and Get2, secondly to define the homogeneity of purified Ge1, Get2 and the Get2/Get1 complex. Finally, the stoichiometry of the complexes Get1/Get2 and Get1/Get2/Get3 should be investigated. Then, after those aims were achieved, a structural characterization of the complexes Get1/Get2 and Get1/Get2/Get3 using X-crystallography or cryo electron microscopy should be started.

Chapter B: Materials

2.1 Chemicals and Yeast Strains

- NEB T7 expression cells (Biolab)
- N,N-Dimethyldodecylamine N-oxide (LDAO; Anatrace, cat. no. D360)
- n-Dodecyl-β-D-maltopyranoside (DDM; Anatrace, cat. no. D310)
- Fos-choline 12 (Anatrace, cat. no. F308)
- Polyethylene(9)dodecyl ether (C12E9; Anatrace, cat. no. AP0129)
- -Pure yeast-enhanced green fluorescent protein (yEGFP) (S65G; S72A)
- Yeast nitrogen base without amino acids (BD, cat. no. 291920)
- Yeast synthetic drop-out medium supplement without Ura (Sigma, cat. no. Y1501)
- D-(+)-Glucose (Sigma, cat. no. G7021)
- D-(+)-Galactose (Sigma, cat. no. G0625)
- D-Sorbitol (Sigma, cat. no. S1876)
- ATP determination kit (biaffin, Germany)
- 1x complete protease inhibitor cocktail tablets (Roche, mannheim, Germany)
- Ni-NTA resin (Qiagen)
- Strep-Tactin XT resin (iba)
- Bicinchoninic acid (BCA) protein assay kit (Pierce, cat. no. 23225)
- Imidazole, minimum 99% (Sigma, cat. no. I2399)
- Ni-NTA superflow resin (Qiagen, cat. no. 30430)
- DTT (Sigma, cat. no. 43815)
- Complete protease inhibitor cocktail tablets (Roche, cat. no. 04 693 132001)
- TEV protease
- Coomassie brilliant blue R-250 (Sigma, cat. no. B-7920)
- Fluorescent protein standard (Invitrogen, cat. no. LC 5928)
- Protein ladder (Invitrogen, cat. no. LC 5625)
- PEG 3350 (Sigma, cat. no. P3640)
- Lithium acetate (Sigma, cat. no. L4158)
- Single-stranded carrier DNA, salmon sperm (Sigma, cat. no. D1626)

- Bacteriological agar (Sigma, cat. no. A5306)
- Yeast peptone dextrose media (YPD media; Sigma, cat. no. P7750)
- DMSO (Sigma, cat. no. D2438)
- URA media
- Yeast suspension buffer (YSB)
- Solubilization buffer (SB) for in-gel fluorescence
- Cell resuspension buffer
- NulcleoSpin Plasmid method from Macherey-Nagel (Macherey-Nagel, Düren Germany). Equilibration buffer
- Membrane resuspension buffer
- Acrylamid (Rotiphorese Gel, carl ROTH)
- Anti-mouse IgG alkaline phosphatise (Sigma)
- Coomassie Brilliant blue G-250 (AppliChem)
- EDTA (Carl, Roth)
- IPTG (Carl, ROTH)
- Ni-NTA column material (Qiagen)
- NaCl (Carl, ROTH)
- SDS (Carl ROTH)
- SigmaMarker Low Range (sigma)
- SYBR® Green I (Invitrogen)
- TEMED (Carl ROTH)
- Tryptone (Carl ROTH)
- Tween 20 (Merck)
- T4 DNA ligase (Jena Bioscience)
- yeast extract (Carl Roth)
- EMS Ultra Fine Tweezers, Style 5TTH, SA
- SB100BN TEM Grid Storage Box, numbered, 100 Capacit
- BSA free (100µg) Cat. Nr. 34660 (Qiagen)
- Anti Mouse (secoundary antibody) Produced in goat (Sigma A9917-1 ml)

2.2 Equipment

- Äkta FPLC (GE Healthcare)
- 300 kV Titan Krios (FEI, Eindhoven, The Netherlands)
- 4k x 4k K2 Summit detector (Gatan Inc., California, USA)
- DigitalMicrographTM (Gatan Inc., California, USA)
- FastPrep-24 classic instrument (MP Biomedicals)
- Gradient former instrument from BioComp (Gradient Master 107, BioComp Instruments)
- Gradient fractionator (Brandel)
- Lumi Imager F1 instrument (Roche)
- 10 L fermentation instrument (Biostat C, Braun Biotech international)
- Microplate reader GENios Pre spectrofluometer (TECAN instrument)
- PVDF Millipore membrane to gel dimension (pore width 0,45 µm)
- Rigaku FR-E+ SuperBright generator
- Rigaku MicroMax-007HF generator
- R1.2/1.3 UltrAuFoil grid (Quantifoil
- SW60 rotors (Beckman)
- TECNAI F30 microscope (300KV FEG, FEI Tecnai TEM)
- TEM Grids, carbon Film coated, 300 Mesh, Cu (Science services, Germany)
- Ultrafree-MC, GV 0.22 µm filter (Merck Millipore)
- Vitrobot Mark IV (FEI)
- Whatman gel blotting paper (Carl Roth)
- Zeis LSM 510 confocal Microscope (Carl Zeiss, Jena)Centrifuge 5804 R (Eppendorf)
- Centrifuge 5415 R (Eppendorf)
- Ultracentrifuge (Beckmann Coulter)
- Frac-950 fraction collector with Rack C (GE Healthcare)
- 5-ml His-trap columns (GE Healthcare)
- Constant Systems TS series cell disruptor (Constant Systems)
- Dialysis tubing, 12–14 kDa molecular weight cut-off (Spectrumlabs)
- Nunc 96-well black optical bottom plate (Nunc)

- HiLoad 16/600 Superdex 200 PG
- Shaking incubator with temperature control (New Brunswick Scientific)
- Superose 6 10/300 GL Tricorn gel filtration column (GE Healthcare)
- Superdex 200 10/300 GL Tricorn gel filtration column (GE Healthcare)
- Superdex 200 PC 3.2/30 column for analytical studies (GE Healthcare).
- Thermocycler (Tpersonal, Biometra)
- Tunair 2.5-l full baffled shaker flasks (VWR)
- TLA 120.1 rotors (Beckman)
- Acid washed glass beads, 500 μm (Sigma)
- 50-ml Aerated capped tubes (TPP)
- Preparative ultracentrifuge, Beckman Coulter Optima L-XP series with Beckman
- 45Ti rotor (Beckman)
- 70.1 Ti rotor, Fixed Angle (Beckman).
- Benchtop ultracentrifuge, Beckman Coulter Optima MAX series with TLA-55 and
- 1.5-ml Polyallomer microcentrifuge tubes (Beckman)
- Centrifugal filter devices (Millipore/Vivascience)
- Confocal microscope (Leica TCS SP2 upright confocal microscope; Leica)
- DarkReader transilluminator (Clare Chemical)
- DarkReader viewing glasses (Clare Chemical)
- Transformation cuvettee (BIO-RAD)
- MinElute Gel Extraction (QIAGEN)
- multitron® incubator shaker (INFORS-HT)
- PCR purification Kit (Jena Bioscience)
- RC5C centrifuge instrument (Sorvall)
- Thermocycler (Eppendorf)
- Thermomixer comfort (Eppendorf)
- Whatman gel blotting paper (Carl Roth)

2.3 Reagents and media

All solutions were prepared with Milli Q water and sterilized by filtering was performed using 0.22 µm filters.

- IPTG stock: 1 M in dH_2O.
- Kanamycin stock, 1000-fold: 25 mg/mL kanamycin sulfate in dH_2O.
- LB medium (1 liter): 10 g bactotryptone, 5 g yeast extract, 10 g NaCl. Sterilize by autoclaving.
- Resuspension buffer for T7-RNA Polymerase: 30 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM EDTA, 10 mM
- β-mercaptoethanol, 5 % glycerol, protease inhibitor cocktail.
- Bactotryptone, 100 g yeast extract, 50 g NaCl.
- DTT: 500 mM stock in dH_2O.
- HEPES: 2.4 M stock. Adjust to pH 8.0 with 10 M KOH.
- Magnesium acetate tetrahydrate, Mg (OAc) 2:1 M solution in H_2O.
- Buffers for DNA-Agarose gels, SDS-PAGE and Western Blotting.
- Agarose: e.g. 1 % (w/v) agarose boiled in 1-fold TAE buffer. Store at RT.
- Ammoniumperoxid sulfate, APS: 10 % stock solution by dissolving 100 mg/mL ammoniumperoxid sulfate in dH_2O. Store at 4°C
- Blotting buffer (Towbin): Dissolve 25 mM Tris, 192 mM Glycin, 3.5 mM (1 %) SDS, 15 % MeOH in dH_2O. pH should self-adjust to pH 8.3. Store at 4 °C.Coomassie brilliant blue-staining solution for SDS gels: 50 % (v/v) ethanol (96 %), 10 % (v/v) acetic acid (100 %) and 0.1 % (w/v) Coomassie Brilliant Blue G250. Dissolve in dH_2O and store at RT in a dark bottle to avoid exposure to light.
- DNA-loading dye, 6-fold: 40 % (w/v) sucrose, 0.25 % (w/v) bromphenol blue and 0.25 % xylene cyanol
- ECL 1: 100 mM Tris-HCl, pH 7.5, 2.9 mM Luminol, 0.4 % p-coumaric acid in H 2 O. Store at 4°C. Luminol and p-coumaric acid are dissolved in 100 % DMSO.
- ECL 2: 100 mM Tris-HCl, pH 8.0, 0.06 % H_2O_2 in H_2O. Store at 4°C.
- PBS wash buffer for western blots (10-fold): Dissolve 1.37 M NaCl, 0.03 M KCl, 80 mM Na 2 HPO_4 and 15 mM KH_4PO_4 in H_2O. Add 0.05 % Tween 20 to the 1-fold dilution. Store at 4°C.

- SDS-PAGE sample buffer, 5-fold: 25 % (w/v) glycerol, 25 % (v/v) β-mercaptoethanol, 7.5 % (w/v) SDS, 0.1 % (w/v) coomassie G250, 300 mM Tris-HCl, pH 6.8. Store at RT.
- TAE buffer, 50-fold: Dissolve 2 M Tris, 5.7 % acetic acid (v/v) and 50 mM EDTA in H 2 O. Store at RT.
- Tricine gel buffers: Anode buffer, 10-fold: 1 M Tris-HCl, pH 8.9. Store at RT. Cathode buffer, 10-fold: 1 M Tris, 1 M tricine, 1 % SDS (w/v). The pH should self-adjust to 8.25. Store at RT.
- Gel buffer, 3-fold: 3 M Tris-HCl, pH 8.45, 0.3 % SDS (w/v). Store at RT.
- 1 M Na_2CO_2/$NaHCO_2$ (pH10) for 1 L the sodium carbonate and the sodium bicarbonate were prepared separately and mix in ratio 6:4. To prepare 1 M Na_2CO_2, 105.99 g of anhydrous Na_2CO_2 were dissolved in water and diluted to 1 L. For 1 M of sodium bicarbonate 84 g $NaHCO_2$ were dissolved in water and diluted to 1 L. Then for 1 L buffer of 1M 1 M Na_2CO_2/$NaHCO_2$ 600 ml Na_2CO_2 mix with 400 ml of $NaHCO_2$
- 1 M Tris-HCl (pH 7.6)
- 1 M Hepes-NaOH (pH 7.5)

Drop out Mix (Without the selection marker) stock solution.

0.5 g adenine, 10 g Leucine, 2 g p-aminobenzoic acid, 2 g myo-inositol, and 2 g of amino acids mix without the selection marker component.

Yeast growth media (1 L)

6.7 g bacto yeast nitrogen base (without amino acid), 2 g drop out mix, and additionally either:

100 mg glucose for expression cultrure, 20 g for pre-culture for growth, 20 g agar for agar plate growth

Terrific broth (TB) media (1 L)

12 g Tryptone, 24 g Yeast extract 5 g Glycerol, The final volume was adjusted up to 900 ml with distilled water.

10 x TB Salt (1 L): 0.17 M KH_2PO_4, 0.72 M K_2HPO_4, the 900 ml media and the 10x TB salt were autoclaved and cooled prior to be used.

Pre-culture media (mL)

0.2 ml of 5M NaCl, 0.5 ml of 10% glucose, 1 ml of 10x TB salt, 8.3 ml TB media

Overexpresion media (890 mL)

20 ml of 5M NaCl, 50 ml of 10% glucose, 100 ml of 10x TB salt, 720 ml TB media

2 x sodium-dodecyl-sulfate (SDS) sample buffer (100 mL)

10 ml of 1.5 Tris/HCl, pH 6.8, 12 ml of 10% SDS, 30 ml glycerol, 15 ml Beta-mercaptoethanol, Small amount of Bromophenol blue, add distilled water to a total volume of 100 ml

2xSDS sample buffer for cell disruption (20 mL)

4 g glycerol, 0.8 g SDS, 1.2 g Tris, 40 µl of 0.5 M EDTA, 2 ml of 1 M DTT, 2.2 ml of 2 M HCl, Small amount of Bromophenol blue, add distilled water to 20 ml.

10 x stock SDS running buffer

10 g SDS, 30.3 g Tris, 144.2 g Glycin, add distilled water to a total volume of l l

TABLE 2.1: LIST OF PRIMERS

NAME	Sequence 5' to 3'
FORWARD GET1GFP8HIS (GF1)	accccggattctagaactagtggatcccccatgcattgggcagcagcggta
REVERSE GET1GFP8HIS (GR1)	aaattgaccttgaaaatataaattttccccatctaaaataatggcatcgtc
FORWARD GET2GFP8HIS (GF2)	accccggattctagaactagtggatcccccatgtctgaattaacagaggcg
REVERSE GET2GFP8HIS (GR2)	aaattgaccttgaaaatataaattttcccctaagtacgtcaataagcctaa
FORWARD GET3GFP8HIS (GF2)	accccggattctagaactagtggatcccccatggatttaaccgtggaacc
REVERSE GET3GFP8HIS (GR2)	aaattgaccttgaaaatataaattttcccccttccttatcttctaactcataaatg
GET1-FWR (GET1)	ggttggatccatgcattgggcagcagcgg
GET1-RV (GET1)	cggtataagcttatctaaaataatggcatcgtc
GET2-FWR (GET2)	cgttggatccatgtctgaattaacagagg
GET2-RV (GET2)	cgttggatccatgtctgaattaacagagg
TGET2-FWR (TRUNCATED GET2)	cgttggatccatgcctgcagccccgatcaatcaagc
GET3-FWR (GET3)	gttaatatacctcttatactttaacgtcaaggagaaaaacatggatttaaccgtggaacc
GET3-RV (GET3)	ctcgagttaatgatgatgatgatggtggtggtgatggtgttccttatcttctaactcataaat
T4L-FWR1 (FOR STREP STAG)	ttaacgtcaaggagaaaaaaccccggattctagaactagtatgaatatatttgaaatgttacg
T4L-RV1 (FOR STREP TAG)	ggtagtccaaagcagcttgattgatcggggctgcaggcatcgcggcatacgcgtcccaagtgcc
T4L-FWR2 (FOR 10X HIS)	taacgtcaaggagaaaaaaccccggattctagaactagtggatccatgaatatatttgaaatgttacg
T4L-RV2 (FOR 10X HIS)	agtaatcatggtagtccaaagcagcttgattgatcggggccgcggcatacgcgtcccaagtgcc
GET2-APOCYTE-FWR	tcgtttgtgctaattgtcttaggcttattgacgtactta gctgcagctgacctggaagacaactgggaaac
APOCYTE-GET1-RV	ccacaataaagaatatcgctaccgctgctgcccaatgcattgcagcgaggtatttctggatgtaagcg

TABLE 2.2 : SDS-PAGE PREPARATION

CHEMICALS	**Separating gel (100 ml)**				**STACKING GEL (60 ML)**
	10%	12%	15%	18%	
30% ACRYLAMID	33.75 ml	40.0 ml	50.0 ml	60.0 ml	10 ML
1.5 M TRIS, PH 8.8	25.0 ml	25.0 ml	25.0 ml	25.0 ml	-
10% SDS	1 ml	1 ml	1 ml	1 ml	0.4 ML
TEMED	120 µl	120 µl	120 µl	120 µl	60 ML
10% APS	600 µl	600 µl	600 µl	600 µl	300 ML
1 M TRIS, PH 6.8	-	-	-	-	6.25 ML
DISTILLED WATER	39.53 ML	33.28 ML	23.28 ML	13.28 ML	32.99 ML

TABLE 2.3: SACCHAROMYCES CEREVISIAE STRAINS USED

STRAINS	Genotyping
CEN-PK259-1D	Mata; ura3-52; his3-Δ1; leu2-3,112; trp1-289; MAL2-8^{c} ; SUC2; pep4(41,1160)::loxP-Kan-loxP
BY.PK1318-18C	Mata; ura3-52; his3-Δ1; leu2-Δ0; TRP1; LYS2; met15-Δ0; pep4::uptag-kanMX4-downtag frt2::uptag-kanMX4-downtag
BY.PK1248-8A	Mata; ura3Δ0; his3-Δ1; leu2-Δ0; TRP1; LYS2; met15; get1::uptag-kanMX4-downtag get2::uptag-kanMX4-downtag
FGY217	MATa; ura3-52; lys2Δ201; pep4Δ; David Drew *et al*., 2008

SOB-medium (1 L)

20 g Tryptone, 5 g yeast extract, 0.584 g NaCl, 0.186 g KCl

LB-medium (1 L)

10 g Tryptone, 5 g yeast extract, 10 g NaCl, 20 g agar plates, the media were dissolved to 1 l with distilled water and autoclaved.

ECL 1

5ml 1M Tris (e.g. pH 7.5), 500µl luminal, 220µl Cumaric acid, add distilled water to a total volume of 50 ml, 15 mM KH_2PO_4

ECL 2

5ml 1M Tris, 100µl H2O2 (30%), add distilled water to a total volume of 50 ml

PBS (10x) 1 Liter

1,37 M NaCl, 30 mM KCl, 80 mM Na_2HPO_4

2.4 Software

Scipion: Scipion is a software framework for integrating several 3D electron microscopy software packages through a workflow-based approach where it allows the execution of reusable, standardized, traceable and reproducible image-processing protocols (de la Rosa-Trevín et al., 2016). Scipion supports almost all packages as follow Bsoft, CTFIND, EMAN, Frealign, Gautomatch, gCTF, gEMpicker, IMAGIC-4D, Localrec, Magdistortion, Motioncorr/dosefgpu, Motioncor2, Relion, ResMap, SIMPLE, SPIDER, and Xmipp.

UCSF Chimera: Chimera (v1.11.2) is a free program for interactive 3-D data visualization, analysis of molecular structures and related data such as density maps, supramolecular assemblies, sequence alignments, docking results, trajectories and conformational ensembles. Chimera is developed by the Resources for Biocomputing Visualization and Informatics, funded by the National Institutes of Health (USA).

Ctffind3: is a free program that use algorithm that fits the amplitude modulations visible in power spectrum with an estimation of the contrast transfer function (CTF) of experimental images (Mindell and Grigorieff, 2003).

EMAN2.1: EMAN2.1 is a complete image processing suite for quantitative analysis of greyscale images with a primary focus on processing data on transmission electron microscopes, with complete workflows for performing high resolution single particle reconstruction, 2-D and 3-D heterogeneity analysis, random conical tilt reconstruction and subtomogram averaging. (Bell *et al.*, 2016).

IMAGIC-5: (v110325; ImageScience, Berlin, GER) is a commercially available software package for the processing of electron microscopic images, with 2-D classification, 3-D reconstruction and 3-D refinement. It integrates angular reconstitution algorithms, as well as projection matching functionality using cross-correlation-based alignment algorithms without weighting functions.

RELION-1.4: RELION stand for REgularized LIkelihood OptimizatioN, is an open-source computer program for the refinement of macromolecular structures by single-particle analysis of electron cryo-microscopy (cryo-EM) data.

PyMol: Microplate reader control and data analysis software Software for the production of high quality 3D images of small molecules and biological macromolecules.

UNICORN 5.11: Software package for control, supervision of chromatograpy systems and processing of chromatograms.

Chapter C: Methods

3.1 Preparations of competent yeast Saccharomyces cerevisiae cells

Yeast is a common host organism for the production of recombinant proteins. Yeast cells are eukaryotes with comparable machinery to the native human source cells of many proteins of interest. There are numerous advantages for using yeast as eukaryotic system to produce native proteins, for instance it is fast system, easy and cheap process relatively to other eukaryotic systems. *Pichia pastoris* and *Saccharoromyces cerevisiae* are the most important components of a matrix of membrane protein production hosts for eukaryotic proteins with yeast system, and also because many eukaryotic membrane proteins for activity rely on the presence of specific post-translation modification such as acylation, phosphorylation and glycosylation (Freigassner *et al.,* 2009). Some of the most difficult human membrane proteins have been produced in high yields in yeast and it did help to determine their high resolved structures (Nyblom *et al.*, 2007). Also pharmaceuticals have efficiently applied yeast system for protein production such as insulin production for example to adapted to their high demand (Kjeldsen *et al.*, 2002). One of the attractiveness of working with *S. cerevisiae* is that there is a wide selection of vectors that can be used for various purposes. *S. cerevisiae* expression plasmids are useful tools for gene cloning and molecular biology as they are shuttle vectors, they can be propagated in both yeast and bacteria (Sikorski and Hieter, 1988)

The procedure used for the competent yeast preparation was adapted from initially described procedure (Drew *et al.*, 2008):

5 ml of yeast peptone dextrose (YPD) (10 g bacto yeast extract, 20 g bacto peptone, 20 g glucose for 1 L) medium were incubated with a colony of the appropriate yeast strain in a 50-ml capped at 200 *r.p.m* overnight at 30 °C. 2 ml of the overnight culture were diluted into 50 ml of YPD in a 250 ml flask and cultured for 6 hours at 280 r.p.m at 30 °C. Then the cells were centrifuged at 3,000 g for 5 minutes at 4 °C, the supernatants were discarded and the cell pellets resuspended with 25 ml sterile dH_2O and centrifuged at 3,000 g for 5 minutes 4 °C, the supernatants were then discarded and the cell pellets resuspended in 1 ml of 100 mM lithium acetate (LiAc). The suspensions were transferred into a 1.5 ml tube and centrifuged at 8,000 g for 15 second. The cell pellets were resuspended in 400 µl of 100 mM LiAc.

3.2 Transformation

3.2.1 Transformation of *Saccharomyces cerevisiae* cells

There are two major types of yeast vectors: The mitotically stable yeast replicating plasmid with only 1-2 copies per cell containing a yeast centromere and ARS sequence (*CEN*) (Sikorski and Hieter, 1988). The other ones are with 2µ origin of replication and have about 20 copies per cell (Christianson *et al.*, 1992). For this specific study and also most research that involve heterologous protein production, uses 2µ plasmids. The galactokinase (GAL1) promoter is the highest expression for controller and inducible expression of heterogeneous genes (Christianson *et al.*, 1992; Mumberg *et al.*, 1994). The refined features of yeast vector were basically constructed in two series of shuttle vectors based on the backbone of the phagemid pBluescript (Stratagene, La Jolla, CA) and they were called pRS vectors (Christianson *et al.*, 1992). The pRS vectors have the *lacZ* α-fragment for blue/white color screening, T7/T3 promoters for *in vitro* RNA transcription, and f1 phage origin for production of single-stranded DNA. The *CEN* plasmids are in pRS410 series while the *2µ* plasmids are within the pRS420 series. The pRS420 series have higher expression than the pRS410 series (Christianson *et al.*, 1992; Mumberg *et al.*, 1994). The series are yeast integrating plasmids (YIp). The pRS420 series have four different selection *His3* (pR423), *TRP1* (PRS424), *LEU2* (pRS425), or *URA3* (pRS426 Based on the researcher needs, one has a choice between copy number, promoter and marker.

Once the competent yeast cells were prepared on same day of the transformation and the vector containing the target gene were ready; the competent cells were transformed by heat shock. In one hand the 50 µl of competent cell suspension were mixed with 5 µl of 10 mg/ml single-stranded carrier DNA (sigma) and 240 µl of 50% (wt/vol) PEG 3350 and vortexed for 5 seconds. In the other hand the genetic material to be used for transformation depends on the goal, either digested vector was mixed with PCR products (and with or without oligo) for gap repair cloning or just a single or more vectors containing different selection marker for co-overexpression; 3 µl of 25 ng/µl *Sma*I- digested vector was mixed with 5 µl of 150 ng/µl PCR product or any oligo while for just a transformation of the cells 1 µl of 25 ng/µl of the expression yeast vector already containing the gene of interest was used directly; the DNA cocktail is then brought to a total volume of 100 µl with sterile dH_2O and vortexed for 5 seconds. Both the competent cells and the DNA mixtures were mixed together by 5 seconds vortex followed by a 30 minutes incubation at 30 °C and transferred to 42 °C for 25 minutes for heat shock, after all the mixture was transferred to ice for 5 minutes. The cells were

thereafter centrifuged at 8,000g for 15 seconds; the supernatant was decanted and replaced by 300 µl of YDP media and pre-cultured for an hour. The culture was then centrifuged again for 8,000 g at 15 seconds, the supernatant was removed and the pellet was dissolved in 50 µl sterile distilled water then transformed yeast Cells were plated on the respective selective agar drop out media and incubated for 2-3 at 30 °C.

3.2.1 *E. coli* Transformation

All *E.coli* were transformed by heat shock procedure the same way; all plasmid of the pRS420 series carry ampicillin gene selection marker while the pET24d carrying Get3 gene contain kanamycin selection marker. The plasmids were used to transform *DH5α* (for the all kind of plasmid isolation purpose) and NEB T7 expression cells (Biolab) for Get3 expression. 1 µl of the DNA was added into 50 µl of competent cells then the mixture was left for 30 minutes on ice. Thereafter the tube was transferred to 42 °C for 45 seconds and returned to ice for 5 minutes. 300 µl of LB were then added and mixed for 45 minutes at 37 °C with shaking at 900 rpm. 50 µl of the transformation mixture were spread on plate consisting of either 50 µg/ml of kanamycin or 100 µg/ml of ampicillin depending on the antibiotic gene resistance carried by plasmid used and subsequently the transformed cells were incubated at 37 °C overnight. The following day a colony was picked to inoculate 10 ml liquid culture LB or TB media for DNA preparation or protein expression in the same condition as above. The plates were stored at 4 °C for later use.

3.3 Construction of the *Saccharomyces cerevisiae* expression plasmids

3.3.1 DNA preparation

The DNA of the yeast target proteins expressed were incorporated mainly in three different yeast shuttle vector backbone, namely the pRS423, pRS425 and pRS426 backbone which contain respectively the *HIS3*, *LEU3* and *URA3* markers for selection and maintenance (Christianson *et al.*, 1992; Mumberg *et al.*, 1994).

3.3.1.1. Yeast GFP fusion based expression construct

The preparations of the green fluorescence protein (GFP) based constructs were generated based on the initial description of the construction of multi-copy yeast expression plasmids by homologous recombination in *S. cerevisiae* (Drew and Hyun, 2012).

The 2µ yeast expression vector genotype comprises a *GAL1* promoter from the pRS426GAL1 backbone, an *URA3 marker* gene, a sequence gene encoding GFP, an octa-histidine tag sequence for western blot detection and affinity chromatography purification, an ampicillin

resistance gene for *E. coli* propagation and a Tobacco Etch Virus (TEV) sequence for GFP removal after the fusion purification.

Before the homologous recombination, the GFP based backbone vector was first digested by *Sma*I while the target sequences of Get1, Get2, Get3 and Get2/Get1 genes were being amplified by the polymerase chain reaction (PCR) (Tables 3.3.1; 3.3.2; 3.3.3; 3.3.4; and 3.3.5). Each primer consists of 35 bp 5'- overhangs sequences complementing the upstream and the downstream of either side of the *SmaI* restriction site within the original pDDGFP-2 vector. The primer design enables the fusion of the gene of interest at the upfront the gene encoding GFP at the N-terminus using gap repair cloning.

The resulting GFP fusion constructs were then used to evaluate the expression quality of the protein using GFP as reporter gene. Once the expression of the fusion of the target and GFP was established, new yeast expression vectors consisting of the gene of the protein to be expressed without GFP were then generated. The new expression yeast vectors lacking GFP coding gene have instead a strep tag or a 10-his tag.

3.3.2. Polymerase chain reaction

Each DNA of interest for amplification was mixed according to Table 3.3.1 or Table 3.3.2. Then the amplification take in a thermocycler as described in Table 3.3.4 or Table 3.3.5.

All the primers used for the amplification are also described in Table 3.3.6

TABLE 3.3.1: PCR REACTION MIX (100 µL) WITH VENT POLYMERASE (NEB)

REAGENTS	Final concentration	Volume in µl
DNA	100 ng	1
20 PMOL FORWARD PRIMER	1 µM	10
20 PMOL REVERSE PRIMER	1 µM	10
DNTP	200 µM	1.5
VENT POLYMERASE	1 units	1
10X POLYMERASE BUFFER	-	10
NUCLEASE-FREE WATER	-	66.5

TABLE 3.3.2: PCR REACTION MIX (100 µL) WITH Q5 HIGH-FIDELITY POLYMERASE (NEB)

REAGENTS	Final concentration	Volume in µl
TEMPLATE DNA	<1000 ng	1
10 PMOL FORWARD PRIMER	0.5 µM	10
10 PMOL REVERSE PRIMER	0.5 µM	10
10 MM DNTPS	200 µM	2
5X Q5 HIGH GC ENHANCER	1X	20
5X Q5 BUFFER	1X	20
Q5 HIGH-FIDELITY POLYMERASE	0.02 U/µl	1
NUCLEASE-FREE WATER	-	36

TABLE 3.3.3: PCR PROGRAM FOR VENT POLYMERASE

STEPS		Temperature in [°C]	Time in minutes
INITIAL DENATURATION		95	2
AMPLIFICATION	Denaturation	95	1
35 CYCLES	Annealing	56	1
	Elongation	72	Time depends on the DNA length
FINAL ELONGATION		72	10
STORAGE		4	α

TABLE 3.3.4: TOUCHDOWN PCR PROGRAM FOR Q5 HIGH FIDELITY

STEPS		Temperature [°C]	Time in seconds
INITIAL DENATURATION		98	300
INITIAL ANNEALING		58	30
INITIAL ELONGATION		72	40
10 CYCLES	Denaturation	98	10
	Annealing	Ramp to 54°C	1°C/Second
	Elongation	72	Time depends on the DNA length
30 CYCLES	Denaturation	98	10
	Annealing	51	30
	Elongation	72	60
FINAL ELONGATION		72	120
STORAGE		4	α

3.3.3. Ligation

The vectors and genes of interest were always ligated according to table 3.3 at room temperature overnight.

TABLE 3.3.5: LIGATION MIX

COMPONENTS	Volume	Concentration
DNA (OLIGO)	1 µl	
BSA	0.1 µl	1x
LIGASE BUFFER	1 µl	1x
VECTOR	2 µl	50 ng
DISTILLED WATER	4.9 µl	
LIGASE	1 µl	

3.3.4 Construction of new expression vector without GFP fusion

To further simplify the expression system, GFP.10xHis tag was removed. Two anti-parallel oligos encoding for 10xHis and two anti-parallel encoding for strep tag were ordered. They were annealed as described in Table 3.3.6 to create the strep cloning cassette and the 10xHis cloning cassette using the thermocycler (Tpersonal, Biometra).

Each cassette was flanked with *Bam*HI-*Hin*dIII restriction enzymes at the 5' end and with *Xho*1 at the 3' end (5'-*Bam*HI-*Hin*dIII-tag-*Xho*1-3') resulting 5'-GATC sequence overhang and AGCT-3' overhang.

TABLE 3.3.6: OLIGOS ANNEALING

STEPS	Temperature	Time in minutes
1	96	2
2	72	2
3	37	2
4	25	2
4	4	60

3.3.4.1. Ligation of the tag sequence into pDDGFP-2 Backbone

The pDDGFP-2 vector was sequentially digested by *Xho*I and *Bam*HI (New England Biolabs, NEB) (Table 3.2). The cleaved pDDGFP-2 gene was purified from the gel using Qiagen gel extraction kit then independently ligated with each cloning cassette for two hours at 37 °C.

The resulting pJANY-S1 (strep-tagged expression plasmid) and pJANY-H1 (His-tagged expression plasmid) expression vectors were used to transform DH5α *E. coli* cells, and then the plasmids were isolated and sequenced to confirm that the cloning was successful.

TABLE 3.3.4.1: SEQUENTIAL DIGESTION OF PDDGFP-2 WITH *BAM*HI PLUS *XHO*I

RESTRICTION ENZYME	1 units (1 µl) *Bam*H1, first 20 minutes) + 1 units (1 µl) *Xho1* after 20 minutes
DNA	1 µg
10X NEBUFFER	4.9 µl buffer (first 20 minutes) + 0.1 µl of buffer after 20 minutes
REACTION VOLUME	49 µl (20 minutes) + 1 µl restriction enzyme after 20 minutes
INCUBATION TIME	60 minutes
INCUBATION TEMPERATURE	Room temperature

3.4 Plasmid isolation

3.4.1 Plasmid isolation from *S. cerevisiae* cells

The DNA plasmid were isolated from yeast using the Zymoprep™ Yeast Plasmid Miniprep II kit (**Zymoprep™ II**, zymo research, Irvine, USA) according to the manufacturer instructions. The Zymoprep™ II protocol is a simple and efficient yeast plasmid preparation based on the *E. coli* alkaline lysis method with a provided Zymolyase™ enzyme which is added in the first solution. There is no need for glass beads, or phenol and the plasmid from yeast cells was each time reliably recovered from colonies, patches on plates, or liquid cultures. The isolated plasmid was then directly used for *DH5α E. coli* propagation or PCR to check the correct gene of the target protein.

3.4.2 Plasmid DNA preparation from *E. coli*

In order to prepare sufficient DNA for sequencing, *E.coli* or yeast cells transformation for target protein expression, the subsequent two steps of DH5α transformation by the target vector and the purification of the Plasmid DNA were done.

5 ml of saturated *E. coli* LB culture media were used to isolate Plasmid DNA using NulcleoSpin Plasmid method from Macherey-Nagel (Macherey-Nagel, Düren, Germany). This procedure consists of precipitating first the protein and then liberating the DNA from *E.Coli* by SDS/alkaline lysis then binding the plasmid DNA to a silica membrane, genomic DNA, and cell debris is subsequently pelleted by centrifugation. Furthermore, contaminations such as salts and soluble macromolecular cellular components are later removed by provided ethanol solution. The quality of the produced DNA is then assessed by measuring spectrophotometrically the DNA concentration at λ = 260 nm using a NanoDrop 1000

Spectrophotometer (Peqlab Biotechnologie GmbH, Erlangen, Germany) which contained already pre-established parameters that includes the Lambert-Beer's law stipulating that an absorption of 1 corresponds to a concentration of 50 µg/mL of double-stranded DNA.

3.5 Colony PCR

In order to identify which colony contains the target gene, we dissolve the colony into 15 µl drop out liquid media; we removed 5 µl for fast PCR. The 5 µl of the cells is heated at 95 °C for 10 minutes and used as template for colony PCR. The PCR is carried in 25 µl total volume with the same parameters as described in previous chapters by reducing the volume accordingly to the dilution factor to 25 µl. After the PCR confirmed the correct target, the remaining 10 µl of the culture can be used to inoculate 10 ml drop out selective media for plasmid extraction from yeast cells.

3.6 Get proteins expression

All the strains of *S. cerevisiae* used have a *pep4* deletion. This specific deletion inhibits Pep4 protease activity and reduces the level of vacuolar hydrolases (Woolford *et al.*, 1986). In the presence of *pep4* deletion the membrane yields were enhanced (Newstead *et al., 2007*; Drew *et al.*, 2008).

3.6.1 Initial Get1 and Get2/Get1-GFP expression test

After the transformation of single colonies of the yeast transformed cells were used to inoculate separately in 10 ml pre-culture growth media that consist 67 mg uracil drop out media (-URA media), 20 mg of bacto yeast nitrogen base (without amino acid) 2% glucose. The cultures were grown 1-2 days until they grow to the log phase. Per pre-culture cells were divided in four aliquots, one was stored at – 20 °C for future plasmid purification and DNA sequence, the other aliquots were then diluted to an $O.D_{600}$ of 0.7 up to 10 ml culture media and proteins were subsequently expressed by directly adding 2% of galactose final concentration to the growth media in presence of 0.1% glucose. The protein expression occurs for 18 hours at 30 ^{0}C with 200 r.p.m.

Thereafter the first 10 ml culture containing the expressed target protein were centrifuged for 3000 *x g* for 5 minutes and the pellet was resuspended in 200 µl of suspension buffer (50 mM Tris-HCl (pH 7.6), 10 % glycerol, 1x complete protease inhibitor cocktail tablets (Roche, mannheim, Germany)). The suspension was the transfer to a black Nun 96 well plate, the GFP fluorescence emission was measured at 512 nm after an excitation at 488 nm using the microplate reader GENios Pre spectrofluometer (TECAN instrument). After fluorescence

reading the suspension was transferred to 1.5 ml Eppendorf tube, glass beads (sigma) were added up to 500 µl plus 500 µl suspension buffer, the yeast cells were broken by strong bench vortex for 15 minutes thereafter the unbroken cells were pellet down at 22,000 *x g* in desktop centrifuge for 5 seconds at 4 °C, the supernatant was transferred to a clean tube, additional 500 µl suspension were added to the pellet and the centrifugation step was repeated, and the final supernatant was transferred to the initial for SDS-PAGE and western blot analysis. 1 mL of the supernatant was centrifuged for an hour at 4 °C at 20,000 *x g*. The crude membrane was suspended in 50 µl suspension buffer and then 15 µl of the suspension was mixed with 15 µl of sample buffer (50 mM Tris-HCl (pH 7.6), 5 % glycerol, 5 mM EDTA (pH 8.0), 4% SDS, 5 mM DTT, 0.0% bromophenol blue). 15 µl were loaded separately on Sodium Dodecyl Sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for coomassie brilliant blue staining and western blot identification.

3.6.2 Quality assessment of the expressed Get1- and Get2/Get1-GFP by confocal microscopy

After the expression of couple of colonies, the colonies that have the corrected genes were identified by western blot upon the GFP expression test. The cell batch with the highest whole cell fluorescence count was selected for membrane localization assessment by confocal microscopy using the Zeis LSM 510 confocal Microscope (Carl Zeiss, Jena) at Frankfurt center for advanced light microscopy. A second set of the best colony was harvested like previously described and the cells were resuspended in drop out culture (-uracil) media containing 50% glycerol, 1 µl of the cell suspension was transferred to a microscope slide then recovered by a coverslip. The sample was then illuminated by Köhler illumination with a 10 x magnification lens then the focal plan was set to zero. After the lens oil was added and the magnification was changed to higher magnification oil-immersion lens. The overall localization of the Get membrane proteins was estimated through the blue light on. The argon laser was used for the excitation at 488 nm to capture a detailed localization of Get-GFP fusion which had detection emission between 505 nm and 535 nm.

3.6.3 Optimization of the overexpression induction

The yeast cells were inoculated in Falcon™ 50mL Conical Centrifuge Tubes containing 10 ml -Ura media with 2% glucose. The culture was incubated overnight in an orbital shaker set at 30 ^{0}C at 280 r.p.m. The next day one tenth of the overnight culture was mixed with nine tenth of culture media for expression to a final $O.D_{600}$=0.7. Each expression culture media consisted of 10 ml –Ura medium with 0.1% glucose, 2% galactose. The following additives were added

in separate tubes: with separately No additive, 10% (wt/v) glycerol, 0.04% (wt/vol) histidine, 2.5 mM trehalose and 2.5% (Vol/vol) dimethyl sulfoxide (DMSO)). The GFP fluorescence was then recorded at various time points after the induction (0, 2, 4, 8, 12, 14, 18, 20, and 22h).

3.7 Detergent Screening

Detergents

Detergents belong to a class of substances called surface-active agents or surfactants (Moroi, 1992). Surfactants attract aqueous solutions at their surface molecules where they are exposed to aqueous environment. Likewise lipid bilayer, detergents consist of a non-polar hydrophobic portion and a polar hydrophilic portion. The polar, hydrophilic part is called hydrophilic or lipophobic group also referred as the head. The non-polar part is called hydrophobic or lipophilic group and referred as the tail. The heads make detergents slightly soluble in aqueous solutions by forming hydrogen bonds and electrostatics interaction with water molecules. In contrast to the non-polar tail is unable to form such interactions. Detergents are amphiphilic compounds cluster in solution which form structures referred as micelles. The concentration range above which they form is known as critical micelle concentration (CMC) (Moroi, 1992).

Detergents can be classified in four main groups: Ionic, Bile Acid salts, Non-ionic and Zwitterionic detergents.

Ionic Detergents

The ionic detergents contain a head group with a net charge (cationic or anionic). For instance sodium dodecyl sulfate (SDS) is a negatively charged detergent (cationic) and cetyl trymethyl-ammonium bromide (CTAB) which carries positive charge (anionic). Those type of detergents are harsh and tend to unfold the protein (Iwata *et al.*, 2003) as they disrupt both inter and intra-molecular protein-protein interactions.

Zwitterionic Detergents

Zwitterionic detergents contain both a positive and negative charge in their hydrophilic head group. They have an apolar tail and a polar head but they are electrically neutral. They are not only effective at extraction of membrane proteins but also keep their native structures. They can however disrupt protein-protein interactions like ionic detergents. Zwitterionic

detergents include series of phosphocholine-related detergents (Fos-Choline, FOS Mea and Cyfos), and CHAPS/CHAPSO.

Bile Acid Salt Detergents

Bile acid salt detergents are also ionic detergents (anionic) with well-defined polar head group and contained a rigid steroidal hydrphobic group. They consist of anionic carboxyl group at the end of short alkyl chain and a hydroxyl groups on the steroids structure. Sodium salts of cholic acid and deoxycholic acid are examples of bile acid salt detergents.

Non-Ionic Detergents

The non-ionic detergents consist of uncharged hydrophilic head groups. They are mild and non-denaturing as they only disrupt protein-lipid and lipid-lipid interactions rather than protein-protein interactions. The hydrophilic head can be either polyoxyethylene glycols (Triton series and tween series), sugar based detergents (octylglucopyranoside), maltosides such as the dodecyl- β-D-maltoside (DDM), and cycloalkylglycosides (cymal series). This class of detergents represents the most used for purification and structural determination of membrane proteins.

3.7.1 Cells culture for expression

From a colony showing the highest GFP expression, 10 ml –URA media containing 2% glucose was inoculated and incubated at 30 °C at 280 r.p.m. The next day the pre-cultures were transferred into culture media in 1:10 ratio of the pre-culture and 9:10 expression media making the induction at the $O.D_{600}$ of about 0.7. The cultures were incubated for 18 hours. The growth media were supplemented with 0.1% of glucose (sigma), 2% of galactose (sigma) plus the corresponding additive that had the highest fluorescence reading during optimization procedure. After 18 hours culture, the cells were harvest by centrifugation at 4000 g for 10 minutes at 4 °C. Then the pellets were weighted and frozen -20 °C.

3.7.2 Membrane preparation

The pellets were resuspended in a volume of 2.5 ml of cell suspension buffer of 50 mM Tris-HCl (pH 7.6, 1 mM EDTA, and 0.6 M sorbitol per liter culture. 2.5 ml of glass beads were the equilibrated with 2.5 ml of the cell suspension buffer, the mixture was kept at 4 °C. 1x complete protease inhibitor cocktail tablets and Dnase were then added. The cells were broken using the FastPrep-24 classic instrument (MP Biomedicals) at 6 m/s for 40 seconds, and incubating on ice then repeating the process 4 times.

Thereafter the broken cells were centrifuge at 10,000 g to remove the unbroken cells and debris. The supernatants were collected and the breakage efficiency was checked by GFP fluorescence reading before and after the cell breakage. Then the membranes were collected after 2 hours ultracentrifugation at 150,000 g using 70.1 Ti Rotor, Fixed Angle (Beckman). The supernatants were discarded and the pellets were resuspended to 0.6 ml of the membrane suspension buffer 20 mM Tris-HCl (pH 7.6), 0.3 sucrose, 0.1 mM $CaCl_2$ per 100 ml culture volume using appropriate disposal syringe with 30G (0,3 mm x 13 mm) gauge needle. Then 100 µl of membrane suspension was transferred to 96-well plate to measure the GFP fluorescence and evaluate the amount of the membrane protein.

3.7.3 Membrane solubilization for analytical study

Membrane proteins are insoluble in aqueous solution. In order to assist membrane protein preserving their native structure in solution lipids are replaced by detergents. In absence of lipid; detergents are introduced to bring membrane proteins to a thermodynamically stable isotropic solution of membrane proteins. As mimic of the association lipid bilayer-membrane proteins, membrane proteins associate to detergents via hydrophobic interactions. In this new environment membrane proteins are surrounded by a layer of detergent where the hydrophilic heads are exposed to aqueous solution. The hydrophobic interactions or forces between hydrophobic regions induces the formation of micelles in solution and when a non-polar substance or group is mixed with aqueous solution, water molecules stability is disrupted, and therefore the water molecules must re-organize their hydrogen bond network around non-polar groups. As both consequences the hydrogen interactions cause the hydrophobic molecules to shield themselves from water molecules in order to keep their native structure. Therefore using this law of attraction regarding the affinity between regions, the hydrophobic regions close to themselves while the hydrophilic region of the membranes proteins interact with the aqueous solution.

To determine in which detergent Get1 is well solubilized approximately 3.5 mg/ml of membrane suspension were transferred into 900 µl in 1.5 ml Beckman polyallomer microcentrifuge tubes, 100 µl of 1% of a specific detergent were added. For the detergent screening I used only 4 detergents were used: N,N-dimethydodecylamine n-oxide (LDAO), Foscholine 12, n-Dodecyl ß-D-maltoside (DDM), and the n-decyl ß-D-maltoside (DM). The membrane suspension and the detergent were incubated for 1 hour by gently mixing. 100 µl of the mixture was then removed for the GFP quantification and the remaining 900 µl of the membrane suspension were centrifuge in a benchtop ultracentrifuge at 100,000 g at 4°C with

the 70.1 Ti Rotor, Fixed Angle (Beckman). The clarified supernatant was transferred into a new 1.5 ml tube. 100 µl of the supernatant were used for the GFP fluorescence quantification and another 100 µl of the detergent-solubilized sample were filtered by the Ultrafree-MC, GV 0.22 µm filter (Merck Millipore) and injected on the Superdex 200 PC 3.2/30 column for analytical studies using Äkta FPLC (GE Healthcare). Prior sample injection the Superdex 200 PC 3.2/30 column was equilibrated by 1.5 column volume of 20 mM Tris-HCl (pH 7.6) 0.03% DDM. 15 µl of each relevant fraction was run on SDS-PAGE for analysis of the correct band size.

3.8 Large scale overexpression

All the optimized conditions during the expression tests were upscale for large scale specifically to each protein. The pre-culture for large scale overexpression as described previously. 1 L in 2.5 mL flask and 10 L fermentation (Biostat C, Braun Biotech international) were used for protein overexpression.

3.9 Cells suspension and membrane preparation

After the large scale expression the cells were in general resuspended without EDTA, with 0.6 sorbitol and 100 mM of the buffer in which the protein is purified (several buffer were used until the correct buffer were finalized). Then as previously described the glass beads were the equilibrated with the cell resuspension buffer. Then mix at 4 °C 1 volume of beads with 1 volumes of cell suspension plus add 1x complete protease inhibitor cocktail tablets and DNase. The cells were broken using the FastPrep-24 classic instrument (MP Biomedicals) at 6 m/s for 40 seconds, and incubating on ice then repeating the process 4 times. Thereafter, the glass beads were let to sediment then a centrifugation step at 10,000 g was done to remove the unbroken cells and debris. Then the membrane was collected after 2 hours ultracentrifugation at 55 000 r.p.m. using the 70.1 Ti Rotor, Fixed Angle (Beckman). The supernatants were discarded and the pellets were suspended to in 6 ml of 50 mM of chosen buffer, 0.5 M NaCl, 0.3 M sucrose, 0.1 mM $CaCl_2$ membrane suspension buffer per 1 L culture volume. The membrane was frozen in liquid nitrogen and stored at -80 °C or directly used.

3.10 Membrane protein solubilization

First the membrane suspension was resuspended with an appropriate disposal syringe with gauge needle. During the work several buffers from wide pH range were tested. MES (pH5.6), Tris-HCl (pH8), Hepes-NaOH (pH7.4), or Na_2CO_2/$NaHCO_2$ (pH10) buffer were separately used depending on purpose of the experiment. The membrane suspension was diluted to a final concentration of 3 mg/ml with an equilibrium buffer of 100 mM of the

buffer, 500 mM NaCl, 10 % glycerol, 0.1 % DDM. Then 1 % powder of DDM final concentration was added for solubilization. The mixture was incubated for 1 hour at 4°C by gently mixing with a magnetic stirrer. The insoluble material was removed as pellet with 55 000 r.p.m ultracentrifucation for 1 hour. In the case GFP fusion 100 µl of samples were collected before and after each step for fluorescence quantification using microplate reader GENios Pre spectrofluometer (TECAN instrument)

3.11 Membrane protein purification

The purification of the membrane proteins were mainly done in two major steps, the affinity chromatography and size exclusion chromatography.

3.11.1 Affinity chromatography purification

During the affinity chromatography purification either Ni-NTA resin (Qiagen) or Strep-Tactin XT resin (iba) were used in a single step or in a two sub-steps purification.

3.11.1.1 Ni-NTA affinity purification.

Per each milligram of proteins 1 ml of Ni-NTA (2 ml of 50% slurry) were added to the solubilized membrane proteins, five column volumes of the equilibrium buffer and 10 mM imidazole were added. The mixture was incubated between 2 and 3 hours at 4°C. After the equilibration the slurry was poured into a plastic gravity flow purification column. Then column was washed with 30 column volume of 10 mM imidazole in the equilibrium buffer. Followed by 35 column volumes of 20 mM imidazole in a buffer of 100 mM of chosen buffer, 350 mM NaCl, 10 % glycerol, 0.05 % DDM. Then again 35 column volumes of 40 mM imidazole in a buffer of 100 mM of chosen buffer, 250 mM NaCl, 2 % glycerol, 0.05 % DDM. The pure protein was eluted with 5 column volumes of 250 mM imidazole in a buffer of 100 mM of chosen buffer, 150 mM NaCl, 2 % glycerol, 0.03 % DDM.

3.11.1.2 Twin-Strep-tag affinity purification

The Iba protocol was adapted to the target Get membrane proteins. For each mg of membrane protein 0.5 ml resin were used, the column was equilibrated with 2 column volumes of 100 mM of chosen buffer, 500 mM NaCl, 10 % glycerol, 0.1 % DDM. Then the solubilized membrane protein was applied to the column. The column was washed two times with one column volume of the buffer (100 mM of chosen buffer, 500 mM NaCl, 10 % glycerol, 0.1 % DDM). Then the columns were washed two times with one column volume of a solution buffer of 100 mM of chosen buffer, 350 mM NaCl, 10 % glycerol, 0.05 % DDM. Finally the

columns were washed two times with one column volume of a solution buffer of 100 mM of chosen buffer, 250 mM NaCl, 2 % glycerol, 0.05 % DDM.

The pure protein was eluted through various elution steps; in the first step the elution 1 was eluted with 0.6 column volume of elution buffer 1 of 100 mM of chosen buffer, 250 mM NaCl, 2 % glycerol, 0.05 % DDM and 50 mM biotin. The second elution step, the elutioon was eluted with 1.6 column volumes of elution buffer 2 of 100 mM of chosen buffer, 150 mM NaCl, 0.03 % DDM and 50 mM biotin. Then finally in the third elution step the protein fraction was eluted two times with 0.8 column volume of elution buffer 3 of 100 mM of chosen buffer, 150 mM NaCl, 0.03 % DDM and 50 mM biotin.

3.11.2 Step two size exclusion chromatography purification

Prior any application on the any column the purified proteins were filtered by the Ultrafree-MC, GV 0.22 μm filter (Merck Millipore). The purified proteins were further purified on the Äkta FPLC (GE Healthcare). The application on gel filtration accordingly to the recommendation of the usage of each column with regard to the flow rate, the pH, the sample concentration and the volume of the sample. For the single protein purification the Superdex 75 10/300 GL and Superdex 75 16/60 (GE Healthcare Life Sciences) columns were used while Superdex 200 10/300 GL and the Superpose 6 10/300 (GE Healthcare Life Sciences) were used for single chain Get1 and Get2 purified complex alone or in complex with Get3.

3.12 Get3 Expression and purification

3.12.1 Expression of Get3.

Get3 was expressed and purified according to the protocol published by Stefer and colleagues (Stefer *et al.,* 2011). Get3 was expressed in T7 express competent *E. coli* (NEB) with the pET24d derivative vector containing Get3 gene and kanamycin resistance. A transformed T7 cells were pre-cultured overnight in 10 ml LB medium at 37 °C and using speed of 180 r.p.m. The following day the culture was transferred to TB medium and incubated for 37 °C with shaking 180 r.p.m until about OD600 of 0.4. Then the culture was cooled to 17 °C followed by the induction of Get3 expression with 0.5 mM isopropyl-ß-D-thiogalactopyranoside (IPTG). The cells were culture for 16 hours and harvested by centrifugation at 5,000 r.p.m for 10 minutes at 4 °C. Each liter culture was resuspended with 40 ml of lysis buffer (50 mM Tris-HCl (pH 7.5), 500 mM NaCl supplemented with DNase and protease inhibitors cocktail). The cells were disrupted using by 6 sonification cycles of 30 seconds each, interrupted by additional 30 seconds incubation on ice. The cellular debris were removed by a centrifugation of 16 000 r.p.m for 30 minutes at 4 °C.

3.12.2 Purification of Get3

Get3 was purified in three steps, Ni-NTA affinity chromatography purification, overnight dialysis coupled with tobacco etch virus (TEV) digestion, reverse affinity chromatography and the size exclusion chromatography.

For a milligram of protein 1 ml of Ni-NTA (2 ml of 50% slurry) were added to the supernatant of the previous step and then five column volumes of the equilibrium buffer (50 mM Tris-HCl (pH 8), 500 mM NaCl, 2% glycerol and 2 mM ß-mercaptoethanol) were added. The mixture was incubated from 2-3 hours at 4°C. After the equilibration the slurry was poured into a plastic gravity flow purification column. Then column was washed with 10 column volumes of 10 mM imidazole in the equilibrium buffer. Followed by 35 column volumes of 50 mM Tris-HCl (pH 8), 150 mM NaCl, 2% glycerol, 2 mM ß-mercaptoethanol and 10 mM imidazole. Followed by another 35 column volumes of (50 mM Tris-HCl (pH8), 150 mM NaCl, 2% glycerol, 2 mM ß-mercaptoethanol and 50 mM imidazole). The pure protein was eluted with 5 column volumes (50 mM Tris-HCl (pH8), 150 mM NaCl, 2% glycerol, 2 mM ß-mercaptoethanol and 250 mM imidazole).

The dialysis using spectrum laboratories membrane coupled with the TEV protease cleavage of the His-tagged. For every 24 mg of Get3 protein 1 mg of TEV were added. The dialysis occurred overnight against 3 L of dialysis buffer 20 mM Hepes-NaOH (pH 7.5), 150 mM NaCl, 2% glycerol, 2 mM $MgCl_2$ and 1 mM 1,4-Dithiothreitol.

The reverse affinity chromatography enables to remove the His-tagged TEV protease, the cleaved Get3 and other remaining contaminant proteins during the step tree. The 5 ml His-trap column was equilibrated with 2 column volumes of 20 mM Hepes-NaOH (pH 7.5), 150 mM NaCl, 2% glycerol, 2 mM $MgCl_2$, 1 mM 1,4-Dithiothreitol and 10 mM imidazole. Then the dialyzed sample was passed through the His-Trap column at flow rate of 2 ml/min. The Get3 protein was collected from the flow-through. The His-tagged TEV protease was eluted from the column. All the fractions were analyzed by SDS-PAGE.

- Finally the separation of the protein was achieved by size exclusion using the dialysis buffer.

3.13 Protein detection and identification

3.13.1 Coomassie staining

70 mg of Coomassie Brilliant blue G-250 (AppliChem) was added to 1 liter of distilled water and stirred for one hour. Thereafter 3 ml of 32% concentrated hydrochloric acid (HCl) was

added and stirred for an additional 15 minutes. The solution was then applied on the gel and heated for 45 second at 460 W in a microwave. The gel was incubated for 15 minutes on shaker. Thereafter the staining solution was discarded and the gel was washed once with distilled water. The staining solution was applied again, the previous step was repeated. To increase the staining performance, abundant amount of water was applied. The gel was heated at 460 W in a microwave for 30 second and incubated for 15 minutes on a shaker. However if the investigated proteins should be identify by mass spectrometry, the staining was performed at RT for 1 hour without heating (first step) and several hours for the second staining step.

3.13.2 Western Blot

The SDS-PAGE gel, 5 pieces of Whatman gel blotting paper (Carl Roth), and the PVDF Millipore membrane to gel dimension (pore width 0,45 µm) were prepared with all about the same dimensions, were incubated with transfer buffer (20% methanol in 1x SDS running buffer). The membrane was activated for 1 minute in 100% methanol, then rinsed well with dH_2O to eliminate the rest of methanol. The Whatman Papers and blotting sandwich sponges of the Western Blot unit were equilibrated by soaking with 25 mM Tris, 192 mM glycine, 3.5 mM SDS and 15% (v/v) methanol of transfer buffer. The blotting sandwich between anode and cathode consists of 3 Whatman papers, the membrane, and the gel and further 2 Whatman papers. The air bubbles and the excess of buffer were removed prior closing the blotting unit. The blot was run at a current of 340 mA for 45 minutes using the Bio-Rad instrument. After the blot was completed the PVDF Millipore membrane was incubated for one hour with blocking solution (5% non-fat dry milk in TBS-T (20 mM Tris, 150 mM NaCl, pH 7.5 and 0.05% Tween)). Thereafter the membrane was washed three times for 10 minutes at RT in TBS-T. The primary anti-His antibody (Qiagen) was added in dilution 1:5,000 and incubated for one hour at room temperature while shaking. Subsequently the membrane was washed as previously. The anti-mouse IgG conjugated antibody (sigma) was added in a dilution 1:5000 plus the anti-strep tag conjugate against the marker in a dilution of 1:7,500 (1.3 µl in 10 ml) to 10 ml TBS plus 50 mg milk powder and incubated for one hour at RT while shaking. The membrane was washed two time for 10 minutes with TBS-T at RT. The blots were developed using chemiluminescence reaction and detected with the Lumi Imager F1 instrument (Roche). The membrane was incubated with 5 ml developing solution 1 (100 mM Tris pH 7.5 (5 ml from 1 M Tris to prepare 50 ml working solution), one tenth volume of final reaction volume of luminal (for example 500 µl for 50 ml reaction volume), 220 µl of Cumaric acid for 50 ml reaction volume) for one minute. Then 5 ml of developing solution 2 (100 mM Tris, 100 µl of 30% of H_2O_2) was added.

3.13.3 Silver staining

The SDS-PAGE gel was incubated for 30 minutes in 100 ml fixing solution (40 ml ethanol, 10 ml glacial acetic acid and 50 ml distilled water). Thereafter the proteins within the gel were oxidized for minutes with 100 ml incubating solution (30 ml ethanol, 6.8 g sodium acetate, 0.2 g sodium thiosulfalte and 70 ml distilled water). Subsequently the gel was washed three times in distilled water for 5 minutes and stained for minutes with 50 ml silver staining solution (50 mg silver nitrate, 50 ml distilled water and 30 µl formaldehyde solution freshly added before using) for 30 minutes on a shaker. Finally the gel was transferred into 100 ml the developing solution (2.6 g sodium carbonate, 100 ml distilled water and 30 µl formaldehyde solution freshly added before using). When sufficient signal intensity was observed, the gel was transferred into 100 ml of stop solution (1,46 g EDTA distilled water).

3.14 Circular dichroism spectroscopy

The CD spectrum of purified Get1 was measured by circular dichroism spectroscopy (CD). Get1 concentration was diluted to 5 µmol/l in buffer (50 mM NaCl and 20 mM Tris-HCl (pH8)). The melting curve of Get1 was measured with a Jasco J-810 spectropolarimeter from the Prof. Dr. Harald Schwalbe laboratory (Goethe University Frankfurt, Germany). The experiment was carried out at standard sensitivity with a band width of 1 nm, a response of 1 second and a scanning speed of 50 nm/min at 20°C in a cuvette of 1 mm cell length. The CD spectrum of Get1 was recorded from 190 to 260 nm. The data represent an average of three accumulations and were base-line corrected by subtraction of a buffer spectrum recorded under identical conditions. Melting curves were measured at a wavelength of 222 nm with a slope of 2°C/min from 10 to 95°C.

3.15 Microscale thermophoresis

The microscale Thermophoresis (MST) detects changes in the hydration shell, charge or size of the molecules by measuring changes of the mobility of molecules in microscopic temperatures gradients (nano Temper technologies). The Monolith NT.115 was used to monitor the thermophoresis of fluorescently labelled Get1-NTblue titrations experiments by non-labeled binding partner Get3. Get1 concentration was kept to 35 nM while the titrant Get3 concentrations ranges between 5 µM and 0.15 nM. The experiment was carried out in 50 mM Tris-HCl (pH 7.6), 150 mM NaCl, 2 mM $MgCl_2$ and 0.03% DDM. After short incubation the samples were loaded into the MST hydrophilic treated glass capillaries and the MST-analysis was performed.

3.16 Mass spectrometry

3.16.1 MALDI-TOF/TOF

The purified identities of Get1, Get2 and Get3 from SDS-PAGE gel slices were confirmed by the Matrix-assisted laser desorption/ionization (MALDI)-time-of-flight (TOF)/TOF through various laboratories: Applied Biomics (USA, project number DOVO141219), the laboratory of Prof. Dr. Karas (Institute for Pharmaceutical Chemistry, Goethe University Frankfurt, Germany) and Dr. Julian Langer (Max Plank institute of Biophysics, Frankfurt, Germany).

3.16.2 Laser Induced Liquid Bead Ion Desorption (LILBID)

The determination of the stoichiometry Get1, Get2 and Get3 complexes were analyzed by a mass spectrometry ionization method LILBID in the laboratory of Prof. Dr. Morgner (Goethe University Frankfurt, Germany). LILBID is a method used to analyze large integral membrane protein complexes and their subunits. The ions in LILBID are IR-laser desorbed from microdroplets containing membrane proteins complexes in detergent. LILBID is highly sensitive, and very efficient in sample handling. One can uses wide range of buffers. The ions detection is narrow and low-charge state distributions (Morgner *et al.*, 2007).

3.17 ATP determination

The level of ATP was determined upon the increase of the concentration of Get2 and Get1 complex using the ATP determination kit (biaffin, Germany). The amounts of ATP were quantitatively determined upon the increases of Get2 and Get1 complex and changes in the concentration. The chemiluminescence at 560 (pH 7.8) is produced by the oxidation of D-luciferin catalyzed by luciferase in ATP-dependent manner:

$$luciferin + ATP + O_2 \xrightarrow{Mg^{2+},luciferase} oxyluciferin + AMP + pyrophosphate + CO_2 + light$$

The D-Luciferin was prepared in 500 µl reaction buffer as well as 150 µl of DTT reaction buffer provided by the kit and the final reaction of 9755 µl reaction buffer, 100 µl DTT, 100 µl D-Luciferin and 45 µl of luciferase were mix together. The resulting mixture was aliquoted, frozen and kept at -20°C. After thawing an aliquot of the final reagent it was allowed to reach the room temperature. 50 µl of the final reagent was mixed with 50 µl of protein solution that may contain ATP. The luminescent signal was measured after 10 minutes. The luminescence background was subtracted using just the buffer to determine the change in luminescence.

3.18 Get2/Get1 complex preparation by GraFix for single-Particle Cryo-EM

The Gradient Fixation (GraFix) tool has been established couple of years ago as tool to purify and stabilize macromolecular complexes for single particles cryo-EM (Kastner *et al.*, 2008). GraFix offers an opportunity to obtain homogeneous samples that often researchers use to have a reliable and high-resolution structure. GraFix is a combination of sedimentation of complexes in a gradient density (for instance, glycerol, sucrose, trehalose can be used to create the gradient density) using a density gradient ultracentrifugation, with weak intramolecular chemical cross-linking (Figure 3.18). GraFix result in the formation of a stabilized monodisperse complexes which are then prepared for negative stain or unstained cryo-EM (Kastner *et al.*, 2008).

Get2/Get1.GFP, T4l.Get2/Get1 and Get3 were all eluted during the affinity chromatography purification in Hepes buffer (150 mM Hepes-NaOH (pH7.5), 150 mM NaCl, 2% glycerol and 500 mM imidazole). The initial cross-linking test was performed with T4l.Get2/Get1 and Get3 complex using glutaraladehyde as crosslinker in concentrations of 0.25% and 0.125%. The complex was mix for 1 hour in presence of 0.1% DDM. Three different amounts of the complex; 1.6 µg, 3.2 µg and 16 µg of 30 µl of the reaction were prepared and 10 µl were remove after 5 minutes, 30 minutes and 60 minutes were the reaction was quenched by 100 mM Tris-HCl, pH 7.4 at room temperature.

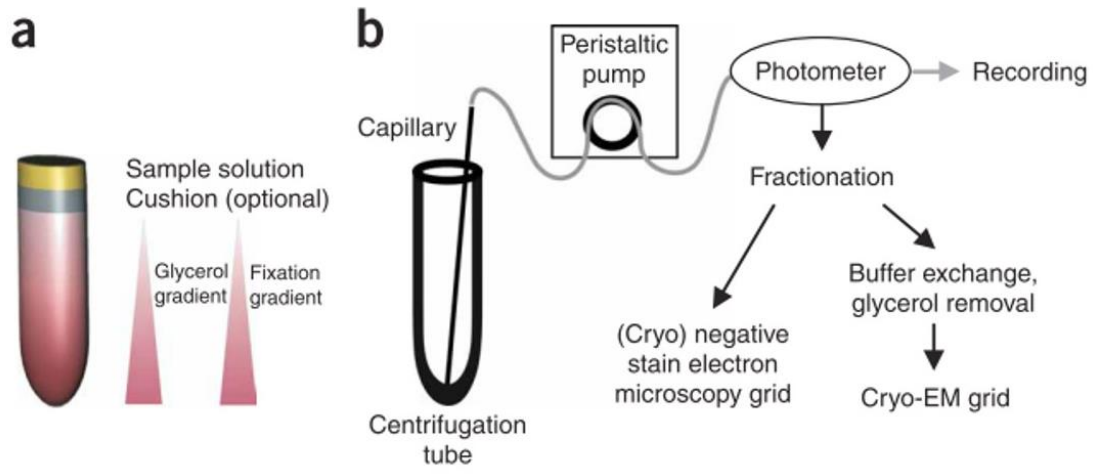


Figure 3.18: GraFix setup. a. Preparation of the Grafix gradient, cross-linking reagent is added to the denser solution (glycerol here); b. Show how the GraFix gradients are fractionated from the bottom to the top (Kastner *et al.*, 2008).

There after we tested the interaction by crosslinking with the water soluble, homobifunctional imidoester crosslinkers which do not alter the charge of protein namely the dimethyl adipimidate-2HCl (DMA, spacer Arm length 8.6Å), the dimethyl pimelimidate-2HCl (DMP spacer Arm length 9.2Å) and the dimethyl adipimidate-2HCl (DMA, spacer Arm length 8.6Å), the dimethyl subermidate-2HCl (DMS spacer Arm length 11.2Å). Once we had tested the appropriate concentration of the crosslinkers we extended and applied to GraFix. 4 ml of 10 to 30% glycerol gradients were prepared using the gradient former instrument from BioComp (Gradient Master 107, BioComp Instruments) then 400 µl of sample were added then were centrifuged for overnight at 40 000 rpm using SW60 rotor Beckman. The gradients were run with and without fixation. Then the gradient was fractionated using the gradient fractionator (Teledyne Isco) at 1V with 20% pump and detected at 254 nm. The samples were used for negative staining.

3.20 Transmission electron microscopy studies

During the recent years Electron microscopy (EM) has become one the most powerful and valuable tool that enables to acquire nice projection images of very small biological objects. Since one of the first great revolution of the structural biology using EM that lead to the structural model of the tail of the T4 bacteriophage at 30 Å using computational 3-D reconstruction of the transmission electron microscopy (TEM), and projections (DeRossier And Klug, 1968), with better electron detectors a new revolution has rocketed the field of the structural biology by storm in recent years, known as the Resolution Revolution (Kühlbrandt, 2014). From the experimental density maps it is now possible to obtain directly the structure of large macromolecules near atomic resolution using electron microscopy (Li *et al.*, 2013). During the Cryo-EM experiment the data collected could be impair by high-dose that primes the radiation damage of exposed biological samples which lead to noisy images, therefore this limiting factor is overcome by low electron doses that are in general within of the range of 10-30 $e^-/Å^2$ on the same sample (Fujiyoshi, 2013: Van Heel *et al.*, 2000).

New generation of sensitive electron detectors have contributed to the improvement of the imaging of macromolecules by powerfully recording high-resolution data; it has improved the quality of the image-recording medium and the image blurring caused by the instability of the sample stage or motion induced by the illuminating electron beam (Li *et al.*, 2013). The photographic film or the scintillator based CCD had the disadvantages of first converting electrons into photons and reconverting them to photoelectron, the new electrons detectors camera have the ability to do so directly. While he CCD was not sensitive enough for high-

resolution data collection and the photographic film did not have the fast electronic readout and high-data throughput (Kühlbrandt, 2014), the new generation of complementary metal-oxide-semiconductor (CMOS) cameras solve the solution of both limiting factors (Li *et al.,* 2013).

In the 3D reconstruction, the phase and amplitude of the specimen are directly obtained from the micrographs and computation system advance to digitize image processing of the Fourier transform of the image which has resulted in development of single particle analysis from TEM projections of randomly scattered particles (van Heel and Frank, 1981).

There are three main EM methods to study biological samples: The negative staining, the 2D electron diffraction and Cryo-EM.

In the negative staining, to improve the signal-to-noise ratio (SNR) the samples are stained with heavy metal salts which are used to scatter electrons strongly by coating biological sample with electron dense molecules. The negative staining enhances the contrast of the target specimen giving the overall shape and surface. Negative staining is easy to use and fast method; therefore the resolutions are not most the time good.

The samples to be analyzed in the 2D electron diffraction are not stained, they are instead 2D crystals. The crystalline materials are embedded in a thin layer and cooled for diffraction collection

In the Cryo-EM, the biological samples to be studied by EM at cryogenic temperature are vitrified by rapid freezing. Cooling the biological sample helps to preserve structure of the specimen studied by preventing the evaporation of water from the molecules in microscope's vacuum and keeping the radiation damage localized (Van Heel *et al.,* 2000).

3.20.1 Single particle negative staining

Before sample application, the samples were filtered using the centrifugal filter the Ultrafree-MC device (Millipore) with a pore size of 0.22 μm. TEM Grids, carbon film coated, 300 Mesh, Cu (Science services, Germany), were used. The grids were first charged with electrons, then 4 μl of 5 μg/ml of proteins were applied onto the grid, left one minute adsorption, then soaked with 2 % uranyl acetate to fully cover the proteins with uranyl acetate, the grids were then incubated for 2 minutes then any remaining liquid was blotted off from the side with filter paper (Whatman paper) and the air dried for 20 minutes in the fume hood.

3.19 Crystallization studies

Membrane protein crystals are classified in three classes (Iwata *et al.,* 2000): The 2D crystals, 3D crystals type I and 3D crystals type II

The 2D crystals are reconstituted biological membrane systems after simultaneous addition and removal of the detergent used to solubilize and purify the membrane proteins. The 2D crystals are used in electron microscopy.

The type I 3D crystals are stacked 2D crystals. The type I crystals are brought together by hydrophobic interactions of connected layers of 2D crystals, they can be grown for instance lipidic cubic phases

The type II 3D crystals are the most common membrane protein crystals obtained. Type II crystals are obtained the same way as with soluble proteins (example the use of precipitants). However in order to obtain the best quality crystals, the preliminary step of selecting the detergent is crucial.

For this studies type II 3D crystals were grown.

In membrane protein crystallography there are microbatch, vapor diffusion, dialyze, and lipidic phase techniques to obtained crystals. The vapor diffusion and lipid cubic phase techniques for the studies.

The Vapor diffusion techniques consist of diffusion and evaporation of water solution as mean to achieve the supersaturation state of the pure protein. Often the precipitant reagent solution and the protein solution are mixed in a 1:1 ratio but this ratio can be changed in trial and error process; the drop is either suspended and sealed over the well (hanging drop) or just sit on top of bed surface that allow the evaporation and the equilibrium to occur (sitting drop). At the beginning the concentration of the precipitant reagent is higher in the well solution, this will induce water to evaporate from the drop until the system reaches the vapor equilibration then reaching the supersaturation in the main time. Drops are prepared in siliconized microscope glass cover slip. The sandwich drop technique represents a mean of controlling the vapor diffusion while a sitting drop technique could be used to overcome the limitation of hanging big volume drops during the optimization

The lipidic cubic phase is quasi-solid membrane system. It is made of lipid, water, and protein all together forming a structured, transparent, and complex three-dimensional lipidic arrangement (Landau and Rosenbusch, 1996). The technique was design to ease the

protein/detergent complex optimization screening once the membrane protein has been purified in a specific detergent. This technique eliminate additional detergent screening for crystallization as the crystals are grown in lipid environment. The system is castoff to form spontaneously a three-dimensional continuous bilayer as such matrices provide nucleation sites, a kind of seeding that prime the growth of protein crystals by lateral diffusion of the protein molecules in the bilayer (Landau and Rosenbusch, 1996). Incorporated membrane protein keeps its activity and functional structure in lipid cubic phase. The matrix of the system is made of two compartments; one compartment is a membrane system with an infinite three-dimensional periodic minimal surface (two circular bilayer unite in the protein can insert). This compartment is solidified or stabiles by the interpenetration of a system of continuous aqueous networks (Landau and Rosenbusch, 1996).

Once the crystals have grown, they were mounted at room temperature test for diffraction or salt crystals. The data collection were always done at very low temperature (-196° C) to increase the molecular order and the diffraction resolution of the crystals. The cooling of Crystal with liquid nitrogen protects the crystals from damaging effect of X-ray radiation. This procedure was done spontaneously and quickly through cryo-protection by the mother liquor in a process called flash freezing or shock cooling. This was done to always prevent the crystallization of the water molecules to protect the protein crystal from ice formation. Before the shock cooling the protein crystal was transferred first to a solvent containing an anti-freeze that acts as cryo-protectant in conditions similar to those where the crystal was grown. Substance such as Glycerol, MPD, and low molecular weight of PEG which can act as cryo-protectant.

The crystallization were done in collaboration within Instruct one the European Strategy Forum on Research infrastructures (ESFRI) lead by Prof. Dr. Dr. h.c. Hartmut Michel from the Max-Planck-Institute of Biophysics (Frankfurt). The initial crystal screenings were setup using the automated platform for high-throuput protein crystallization from Rigaku (CrystalMation). The CrystaMation system is equipped with two liquid handling robots: An Alchemist II for making the crystallization screens and a Phonix Re for crystallization plate set up. The crystal growths were further screened by vapor diffusion sitting drop and the lipid cubic phase techniques screening. The crystallization of Get1, Get2/Get1 single chain, T4l.Get2/Get1 single chain, truncatedGet2/Get1 single chain, T4l.Get2.apocytochrome.Get1 (or T4l.tGet2.apocyte.Get1) single chain, all with and without the binding partner Get3 were carried out mostly at concentration higher than 4 mg/ml with a sample size of 100 nL of

protein overlaid with 100 nL mother liquor and incubated at 4 °C or 18 °C in 96 well plate formats with commercial and in-house prepared membrane protein screens. For MemGold, MemGold2, MemStart and Memplus, Memsys and sigma membrane kits (Molecular Dimensions), The MbClass Suite and the MbClass II suite (Qiagen) were used vapor diffusion crystallization set up. While the lipid cubic phase crystallization screens were setup manually using the Phoenix robot on the Cubic Phase I Suite and Cubic Phase II Suite (Qiagen). The growth of the crystals were imaged and monitored via the CrystalTrakweb. Over 5760 conditions were screens.

3.19.1 Crystal Optimization

The initial crystal hits were optimized by varying different parameter such as the salt, the additive, the pH on the 24 well plates in order to obtain high-quality crystals. For this procedure the hanging drop vapor diffusion was used to try to grow larger and well diffracting crystals. In total about of 720 (24x30) different conditions were optimized in 0.5 µl of protein mix with 0.5 µl of optimizing initial conditions.

Most of crystals were first tested for salt crystal using one of the two in-house X-ray station: The very powerful Rigaku FR-E+ SuperBright generator equipped with a VariMax-VHF multilayer optic combined with a Saturn 944+ CCD detector or the Rigaku MicroMax-007HF generator with a VariMax-HR optic connected to a Rigaku R-Axis IV image plate. The other crystals were quickly cryo-protected and directly flash-frozen in liquid nitrogen and later tested for diffraction data collection at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) or at Deutsches Elektronen-Synchrotron (DESY, Hamburg, Germany) or at the Swiss Lightsource SLS (Paul Scherrer Institut (PSI), Villigen, Switzerland).

3.20.2 Single-particle Cryo-EM

For single-particle cryo-EM, a 4 µl aliquot of 100 µg/ml purified T4l.tGet2.apocyte.Get1 sample was applied to a glow-discharged (20 s) R1.2/1.3 UltrAuFoil grid (Quantifoil), and plunge-frozen in liquid ethane (Vitrobot Mark IV (FEI) at 95% humidity, 5 °C, 8.5 s blotting time, blot force -1). Dose-fractionated movies (30 frames, 0.25 s each) were recorded using DigitalMicrographTM (Gatan Inc., California, USA) at a nominal magnification of 130,000× (1.05 Å/pixel) using a 300 kV Titan Krios (FEI, Eindhoven, The Netherlands) in nanoprobe energy-filtered transmission electron microscopy (EFTEM) mode using a GATAN GIF Quantum® SE post-column energy filter set to 30 eV slit width. A 4k x 4k K2 Summit detector (Gatan Inc., California, USA) was operated in dose-fractionation counting mode with

a dose rate of ~7.5 electrons per pixel per second (0.25 s single frame exposure) and a total dose of ~56 electrons per Å2. Defocus values ranged from -0.7 to -4 μm.

3.20.3 Image processing

3.20.3.1 Motion correction

Before processing the images, the electron beam-induced sample motion that might significantly degrades the resolution and image blurring was corrected to improve the resolution of data and where intrinsic image information can be restored to high resolution (Li *et al*., 2013).

3.20.3.2 Bandpass filter

A bandpass filter also known as Gaussian filter was then applied to remove information above and below certain resolution levels from the reconstruction. The information of low frequencies as they describe very coarse structure above molecular information (high-pass filter) and the high frequencies (low-pass filter) describing structure features below the resolution level were eliminated. The bandpass filtering helps to remove noise from the reconstruction process.

3.20.3.3 CTF correction

The contrast transfer function (CTF) is very important factor of an image formation in the TEM. As during data acquisition the biological samples work like weak phase object by changing the phase of the beam in optical systems, therefore when one collect data, one has to deliberately defocus the images in order to see the sample. The CTF then modulates the frequency spectrum of the sample by the changed defocus and an aberration due to lens also known as spherical aberration (Cs). In the Fourier transform function of an image the modulations of the CTF can be seen as thone-rings. The CTF of all images need to be then corrected to determine the original positions for all images and multiply the values with -1. Therefore transferring image function information that became negative values during the data acquisition to positive side of the x-axis in order to have valuable informations during the image processing. The CTFfind3, a program developed in the MRC laboratory was used (Mindell and Grigorieff, 2003).

3.20.3.4 Particle picking

During the course of the work particles were picked using three different packages depending on the goal:

EMAN2.1: eman2-boxer was used to pick particles by first setting the particle size and diameter. The Swarn tool for picking is then selected. In this mode one has to select couple of particles and the algorithm proposes you more particles which one can remove if they are not truly particles. Then when one move to the next micrograph, the particles are picked automatically and there also the particles can be removed. By selecting the "gauss" tool one can set the Gaussian kernel width and check with the best threshold. Those parameters were in general improved couple of times until satisfaction.

Xmipp particles picking from Scipion. It consists in two steps. During the first step one pick the particles manually by clicking on manual tool and executing the protocol. The particles sizes are set here. Same as EMAN Swarn tool the algorithm will train a classifier and will propose some coordinates automatically. One can remove or add new particles or even erasing large areas of contamination. In general the performance of the classifier after 30 micrographs was really good. At this point I registered the output coordinates by clicking on the "coordinates red button. Then I closed the GUI and open the "xmipp3-automatic" box and choice the previous execution of the manual and supervised as input. The second step is completely automatic picking based on the optimized picking of the first step.

RELION particles picking: For relion the 2D classes are first generated by manual picking particles. Then use the 2D classes from manual pick as template. At this level Relion auto-picking was started by first computing the figure-of merit (FOM) maps, adjusting the parameters such the picking threshold and the minimum inter-particle distance and then picking the rest of the micrographs. For 3 micrographs representing the all data set were selected, example a low defocus, high defocus and thick ices as input and as well as the reference class averages. To perform the cross-correlation the angular sampling was used. The contrast was inverted and the CTF corrected was set to "Yes" prior executing. Thereafter one open the relion-autopicking and set it as the FOM. Parameters such as picking threshold and particle distance were adjusted few times until satisfaction. When every parameter was optimized the automatic picking was started using now all the micrographs. At the end of the process through their coordinates particles were then extracted.

EMAN2.1 was at the end the preferred particles picking tool and their coordinates were used to extract the particles on RELION. However at this point each package has its pros and cons such as calculating additional CTF at particles level and selecting particles based on their signal to noise ratio in EMAN2.1 or assigning a z-score value for each particle where particles with high Z-score are disabled using Xmipp3 or in Relion with particle sorting one can sort particles from the good to the bad ones in a file called rlnParticleSelectZScore.

3.20.3.5 Picking particles for Random Conical Tilt

In order to produce starting model from random conical tilt where each specimen stage of the TECNAI F30 microscope (300KV FEG, FEI Tecnai TEM) was tilted 45° and new low-power image was taken. To pick semi automatically the particle pairs, either Xmipp package or EMAN2.1 package was used according to the developer instructions. Once the coordinate pairs were picked, the particle pairs were then extracted.

3.20.3.6 2D classification

During the 2D classification, the particles are placed into 2D similar classes, aligned, and centred in comparable positions in order to obtain an average of the 2D image of good particles. It allows to have different view or conformation of our particles of interest, but it also helps to discard bad particles.

Several approaches are used by different package to align 2D images. Although Xmipp, EMAN2.1, IMAGIC (for negative staining) and RELION packages were used to generate 2D class averages, for Cryo-EM RELION was mainly used for 2D classification processing. The 2D classification was set in such way that a maximum of 200 particles included in each class and the regulation factor T=2 was used. After the classification the best classes were selected and bad classes were discarded. The 2D classification was repeated additional 2 times to clean up the data.

3.20.3.7 Initial volume estimation

Prior 3D reconstruction of any final 3D map it is necessary to have in hand an estimated low resolution initial model. For this purpose different methods from different packages were comparatively used to generate a consensus starting model: EMAN Random Conical Tilt, Xmipp Random Conical Tilt, Xmipp 3D-RANSAC, Xmipp Reconstruct Significant, and EMAN Initial Model. In the low resolution structure presented in this work, no symmetry was applied during the reconstruction process from the starting model to the final 3D map.

3.20.3.8 3D classification of structural heterogeneity

For the unsupervised 3D classification to separate particles on their heterogeneity in RELION, the starting volume from Xmipp Reconstruct Significant was used. As the size of the Get2/Get1 complex is relatively small, low-pass filter of 12 angstroms was used. 50 iterations were performed with a regularization factor T=4, No symmetry C1 was kept. The 3D classification was done with mask generated by relion command **relion_mask_create**. The 3D classification was initial done with 5 classes then focus only on 4 classes.

3.20.3.9 High-resolution 3D refinement

Each 3D map from the 3D classification was refined with a low-pass filter of 12 angstroms and other parameters were not changed. The RELION refinement is based on gold-standard approach to prevent overfitting and over-estimation of the resolution. The Data were separated in halves and refined independent reconstructions against each half-set. Therefore the Fourier Shell correction (FSC) between the two independent reconstructions produces a reliable resolution estimate. In that way the noise that may be refined after each iteration has chances to be prevented.

3.20.3.10 Movies Refinement

The movie processing is the continuation of the 3D refinement. There each movie, a separate stack with movie frames of all particles per micrograph were stored in the same directory with the extracted particle stacks. The movie-frames of each particles were aligned to the fitted motion tracks, and each single frames are used during the RELION reconstruction to estimate a B-factor for each movie frame which result on averaging the movie frames of each particle with frequency-dependent weights according to their relative B-factors. This procedure results in improved SNR.

3.20.3.11 Particles polishing

The particle polishing was performed on the data output file from the 3D movie refinement.

3.20.3.12 High-resolution 3D refinement

Because the polished particles are better aligned than the original ones and that the 3D auto-refinement were already made; an additional 3D auto-refinement was performed to In order to improve the resolution.

3.20.3.13 Post-processing

To prevent overfitting during the 3D auto-refinement the two independently refined reconstructions are not masked when the FSC curve is calculated, this likely lead to an underestimation of the true resolution of the reconstructed signal because this signal is limited to the central region of the map and the surrounding solvent region just contributes to noise. Therefore to reduce the noise, the solvent region was masked. The phases of Fourier constituents of the two half-reconstructions with special frequencies higher than a given cut-off were randomized to correct the FSC curves.

Chapter D: Results

4.1 GFP fusion: A tool for the optimization of the expression, solubilization and purification of Get1 and Get2

To obtain pure Get1 and Get2 that were used for structural studies, the expression, the solubilization and the purification were initially established using GFP fusions and FGY217 yeast strains (Drew *et al.*, 2008).

The genes of Get1, Get2 and the single chain construct of linked Get2 and Get1 (Figure 4.1) were amplified with 5´ and 3´overhang 35 nucleotides matching a linearized p426 backbone vector by *Sma*1 restriction enzyme containing GFP fusion (Figure 4.2).

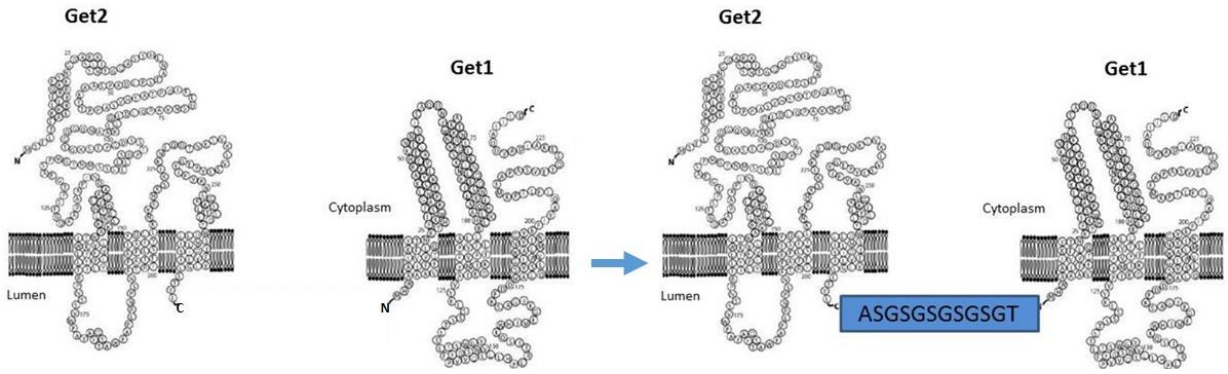


Figure 4.1: Single chain construct between Get2 and Get1. The topology of Get1, Get2, and Get2/Get1 were adapted from Stefer *et al.*, 2011.

The GFP expression vectors containing the gene of interest were prepared by homologous recombination after the transformation of the yeast strains

Thereafter the transformation of the expression, the membrane solubilization and purification were tested and optimized using GFP based vectors included in the P426Gal backbone vector. In general, the fluorescence intensity of the expressed fusion protein at different time points in comparison to non-induced expression was measured between 505 to 515 nm emission wavelengths.

Provided that GFP was cloned downstream of the gene of interest, any GFP intensity measured had direct relation with the upstream cloned gene to be expressed. When using a microplate reader GENios Pre spectrofluometer (TECAN instrument), the concentration of GFP was determined by the following formula:

Concentration of GFP$= fluorescence\ intensity \div 3.4 \div 1000\ mg/ml$ (1)

Then, the amount of the protein was:

$$[\text{mp}] = [GFP] \times Molecular\ weight\ of\ GFP \div Protein\ Molecular\ weight \quad (2)$$

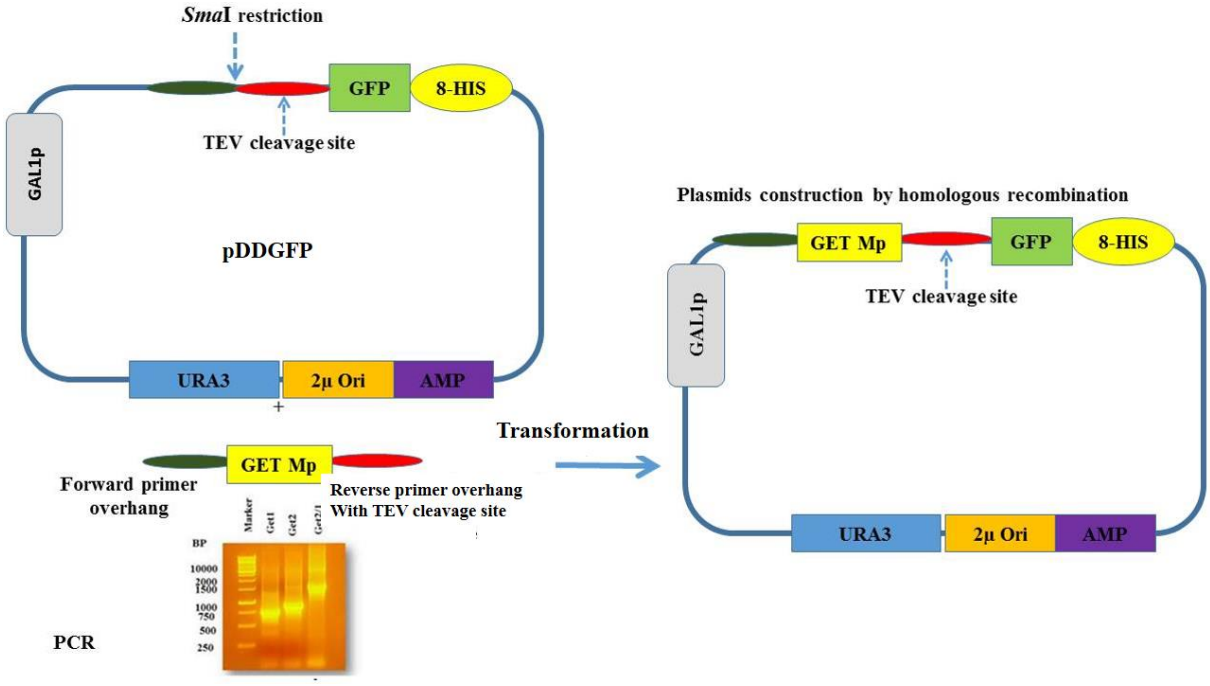


Figure 4.2: Schematic illustrating of the contruction of Get1- and Get2- fused to GFP fusion plasmids. The vectors were used for the expression, the localization, the solubilization and purification of Get receptors. Each gene encoding the protein of interest was placed upstream of GFP.

4.2.1 Selection of the best expressing clones by the intensity fluorescence counts

Seven independents colonies were first picked and cultured overnight in uracil drop out media. Before the induction of the protein expression, each of the seven cultures were separated in two aliquots. One aliquot was used to induce 10 ml drop out culture medium by addition of 2% galactose and the other one was stored for future use. After overnight culture in the presence of galactose, the expressed protein was measured by whole cell fluorescence count intensity and the two best fluorescence intensities were selected as shown in Table 4.2.1.

Get1 highest fluorescence intensities were 2595.3 and 1855.5, while Get2 intensities were 1963.8 and 2489.8. The intensity of the Get2/Get1 was significantly higher than Get1 and Get2 with a read intensity of 7896.5 and 7465.

TABLE 4.2.1. EXPRESSION TEST USING GFP MEASUREMENT

Whole fluorescence cell counts of the control: No induction	Get1	Get2	Get2/Get1
	311.5	320.5	306.75
Whole cell Fluorescence counts for 7 independent yeast transformants	1677.5	**1963.8**	6523
	1002.8	1162	**7896.5**
	1058.8	1803.5	5785.3
	1009.5	1350.8	**7465**
	2595.3	799.5	4093
	1774	**2489.8**	4408.8
	1855.5	1331	4556.5

By applying the formula (1), the corresponding GFP fusion concentrations were found to be 0.76 mg/ml and 0.45 mg/ml for Get1, 0.48 mg/ml and 0.64 mg/ml for Get2 and finally 2.2 mg/ml and 2.1 mg/ml for Get1/Get2. The molecular weights of GFP, Get1, Get2 and Get1/Get2 are 27 KDa, 29 KDa, 31 KDa, and 61 KDa, respectively. The corresponding overexpressed amounts of Get1, Get2, and Get2/Get1 were deducted from the formula (2) by:

- For [Get1]= $0.76 \times (27/29)\frac{mg}{ml} = 0.7\ mg/ml$ and the second one is 0.42 mg/ml.
- For [Get2]= $0.48 \times (27/31)\frac{mg}{ml} = 0.42\ mg/ml$ and 0.56 mg/ml
- [Get2/Get1]= $2.2 \times (27/61)\frac{mg}{ml} = 0.97\ mg/$ and 0.93 mg/ml.

Therefore, Get2/Get1 expressed as single chain has more expressed protein which seems that Get1 and Get2 are more stable in a single chain construct or when they are interacting with each other compared to when they are overexpressed individually. After the expression test, the same cells were collected for in gel fluorescence (Biometra).

4.2.2. Quality Assessment by Confocal Microscopy

Get1 and Get2 have been shown to have cooperative binding to Get3-TA (Wang *et al*., 2011), however the cytosolic N-terminus of Get2 was shown to tether Get3/TA complex to the membrane and Get3 did display high affinity for cytosolic Get1 in both the apo (Kd = 17 nM) and ADP (Kd =31 nM) states (Stefer *et al.*, 2011). Therefore, in this thesis, for the reconstitution of the complex Get1/Get2/Get3 the focus was mainly on stabilizing the interaction of Get3 with Get1 to acquire structural information of the overall complex.

Henceforth, after the expression test of Get3 receptors, 100 µl of the same sample were used for assessing the correct targeting of the ER by the subcellular localization of the expressed Get1 and Get2/Get1 using GFP fusion. The correct targeting of Get1- and Get2/Get1- GFP fusions are shown in Figure 4.2.2 a. and Figure 4.2.2. c. During the confocal microscopy, the samples were then illuminated by Köhler illumination with a 10x magnification lens. Afterwards, then the focal plan was set to zero. Finally, the lens oil was added and the magnification was changed to higher magnification oil-immersion lens.

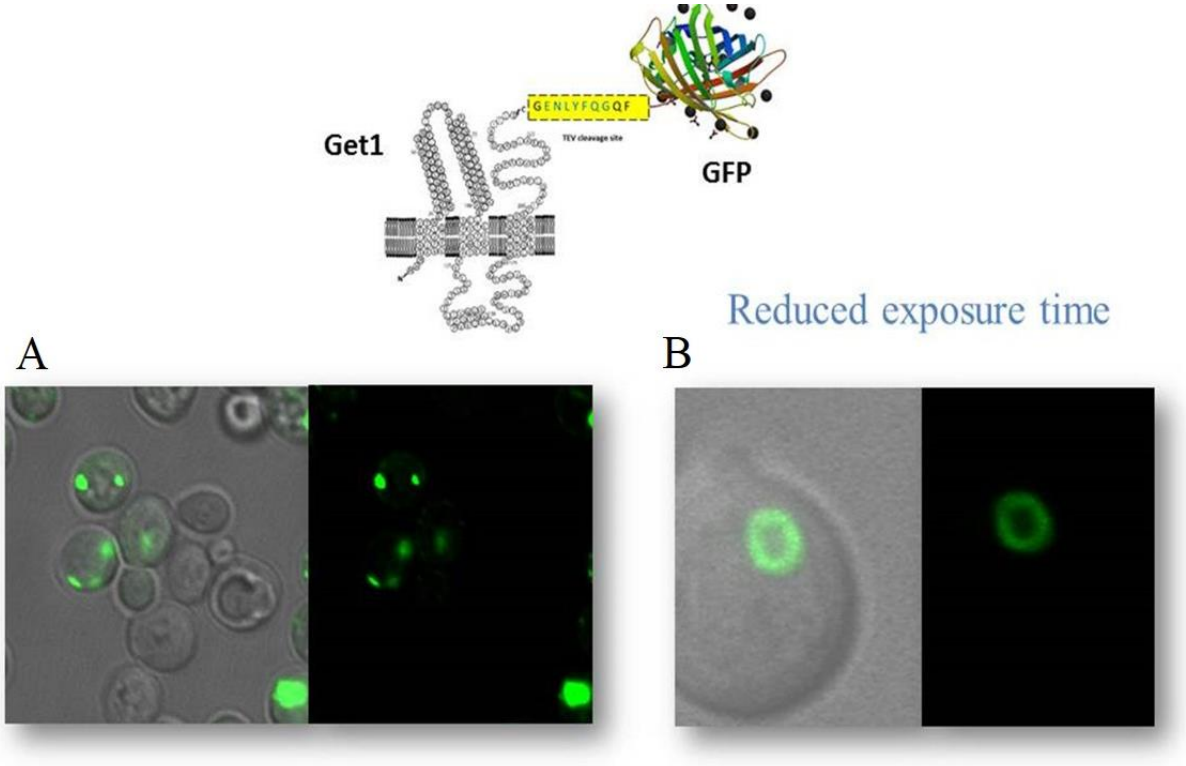


Figure 4.2.2 a: Quality assessment of Get1.GFP by confocal microscopy. The localization of the expressed Get1. The left figure is an overall view yeast cells containing the plasmid pJNG1-GFP, the grey with green detection is the differential interference contrast and the dark image is the overlay image detected at visible wavelength with GFP emission signal at 512 nm. The figure shows a lot cells are not fit or are dead. The right figure was taken by reducing the exposure time, (the topology of Get1 is adapted from *Stefer et al 2011*, and the structure of the crystal structure of the superfolded GFP, protein data bank accession number 2B3P).

The quality of Get1-GFP was assessed first; although Get1 fused to GFP was not localized in the cytosol nor in the plasma membrane, the confocal images showed that Get1 was targeted to an organelle membrane (Figure 4.2.2a), but this membrane was not a plasma membrane or ER membrane. It seems that it was, however, peroxisome (microbodies), vesicle, vacuole or even a nucleus like shape. All the eventualities are possible. Still, by closely investigating the sequence of Get1, two yeast peroxisomal targeting signals SKF and YKL are present near the TMD entry in the cytoplasm or lumen (Figure 4.2.2 b).

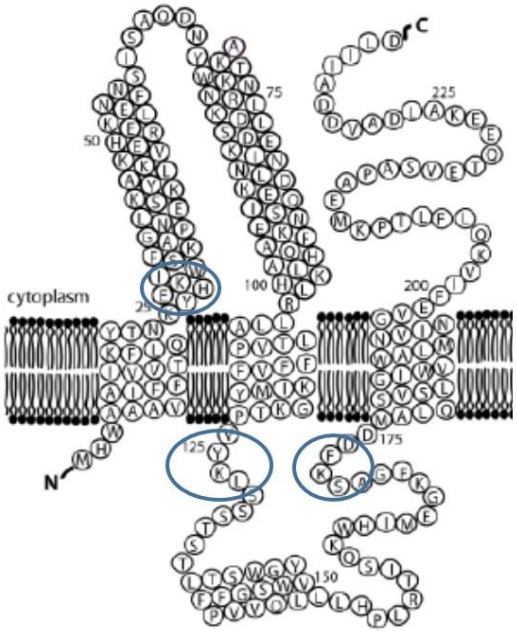


Figure 4.2.2 b: The topology of Get1. Two yeast peroxisomal targeting signals SKF and YKL (blue circle). The topology of Get1, is adapted from Stefer *et al.*, 2011.

However, Get2/Get1 seems to be targeted to the right ER membrane as the figure 4.2.2 c shows. GFP shows healthy cells with cisternae like shape which seems to be consistent with the expression test reported in Table 4.2.1.

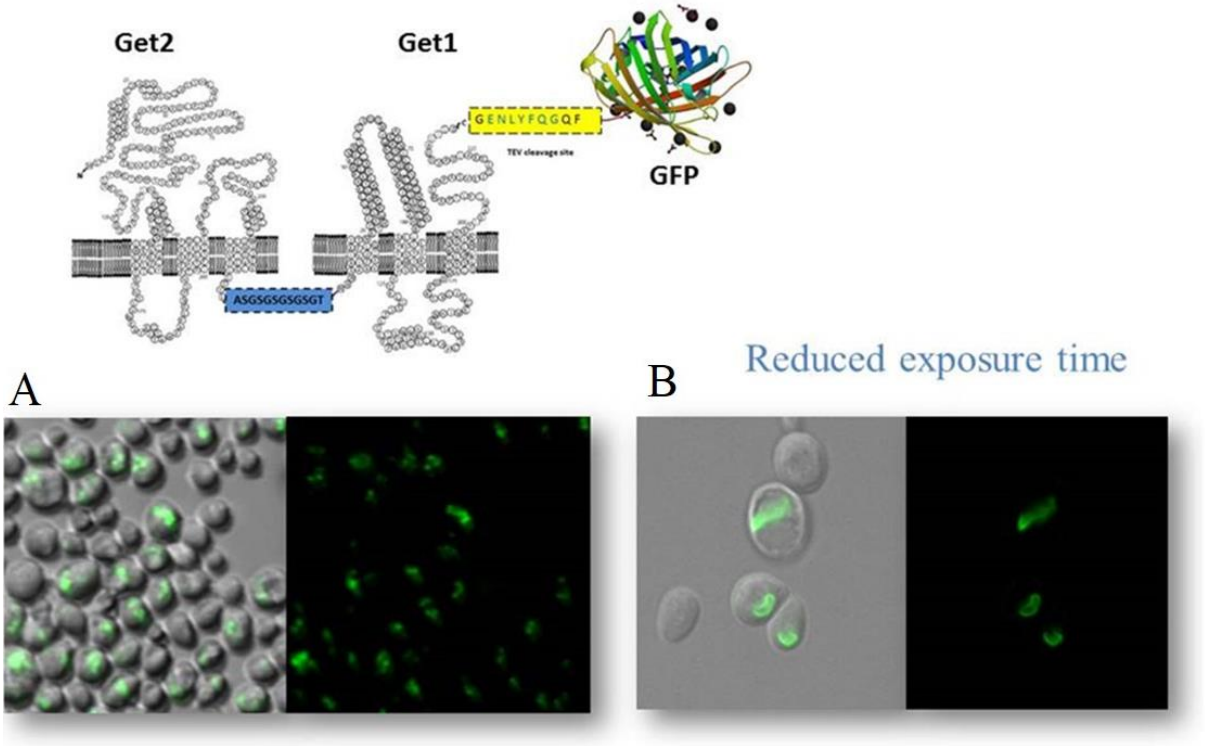


Figure 4.2.2 c: Quality assessement of Get2/Get1.GFP by confocal microscopy. The localization of the expressed Get2/Get1. The left figure is an overall view yeast cells containing the plasmid pJNG21-GFP where the expression of Get2/Get1 was induced, the grey with green detection is the differential interference contrast and the dark image is the overlay image detected at visible wavelength with GFP emission signal at 512 nm

4.2.3 Optimization of the cultivation conditions upon overexpression

To enhance the yields of Get1, Get2, and Get2/Get1, the effect of chemical additives was analyzed. Chemical chaperones can improve the functional yields of proteins. When yeast cells are stressed upon the addition of chemical chaperones they will produce molecular chaperones in order to prevent the unfolding or to help the misfolded proteins. For this study DMSO, glycerol, histidine, and trehalose were used during the induction alone or as mixtures.

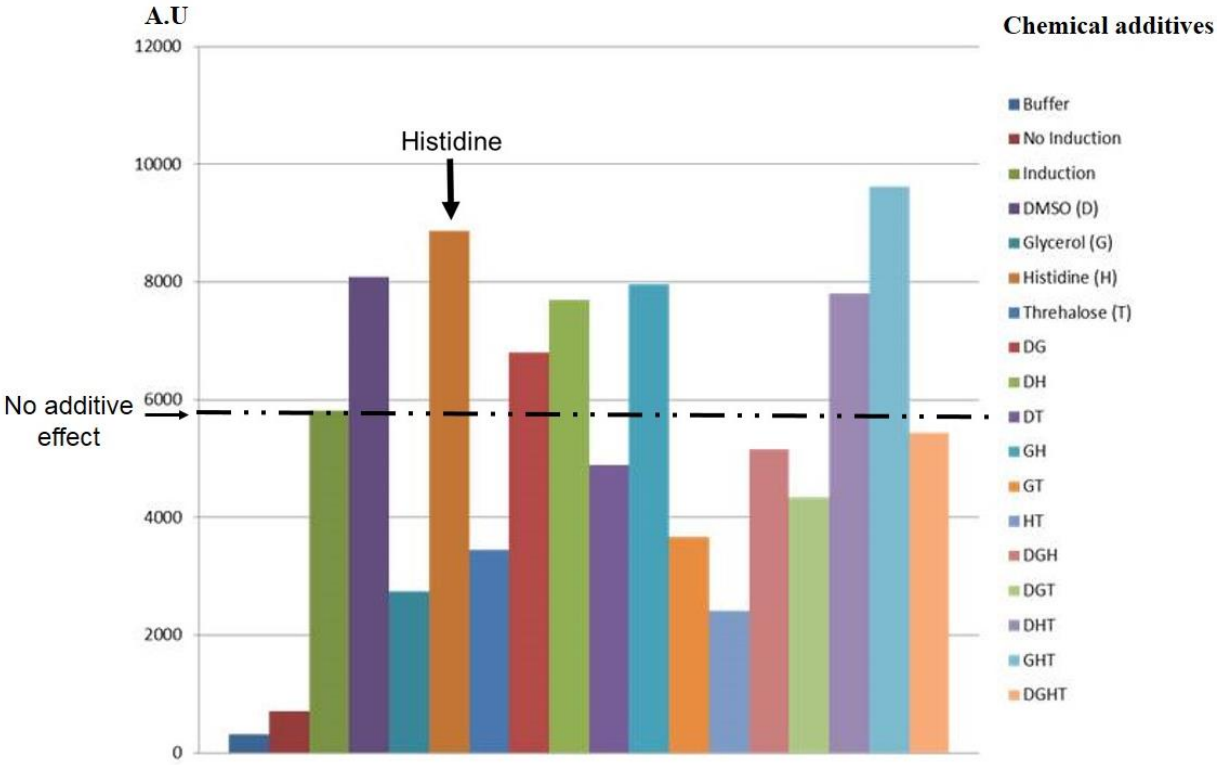


Figure 4.2.3 a. Effect of Chemical chaperone on Get1.GFP expression in yeast. Get1-GFP expression when induced in presence of DMSO, glycerol, histidine, trehalose, and mixture of those chemical chaperones to each other. Each increase of 1000 a.u has about 0.30 mg/ml increase in concentration of GFP which represent about the same increase of Get1 protein (0.28 mg/ml).

The chemical additives histidine and DMSO have an effect on the expression of Get1, at first the effect does not seem significant but it is about 2000 a.u, which is about an 0.6 mg/ml increase in concentration. This is considerable as each pellet of a liter culure is dissolved in 5 ml, in addition gives a gain of 3 mg of expressed protein and for membrane proteins, which turns out to be a huge gain.

The effect of those chemical additives were tested for Get2.GFP (Figure 4.2.3 b.) and Get2/Get1.GFP (Figure 4.2.3 c) as well. It was found that histidine alone increases the yield by 1.8 mg of protein, but no significant change was observed when Get2/Get1.GFP expressing cells were treated with chemical additives. However, there is slight effect with histidine but in the case of Get2/Get1.GFP the expression by itself is already relatively high in comparaison to Get1.GFP and Get2.GFP. One explanation might be that the space in the ER membrane is limited and the single chain Get2/Get1.GFP likely occupies more space. At this point, it was just important to see that we have an effect of chemical chaperonnes and this is why there is no error bar as the trend was that histidine has an effect on all Get receptor constructs.

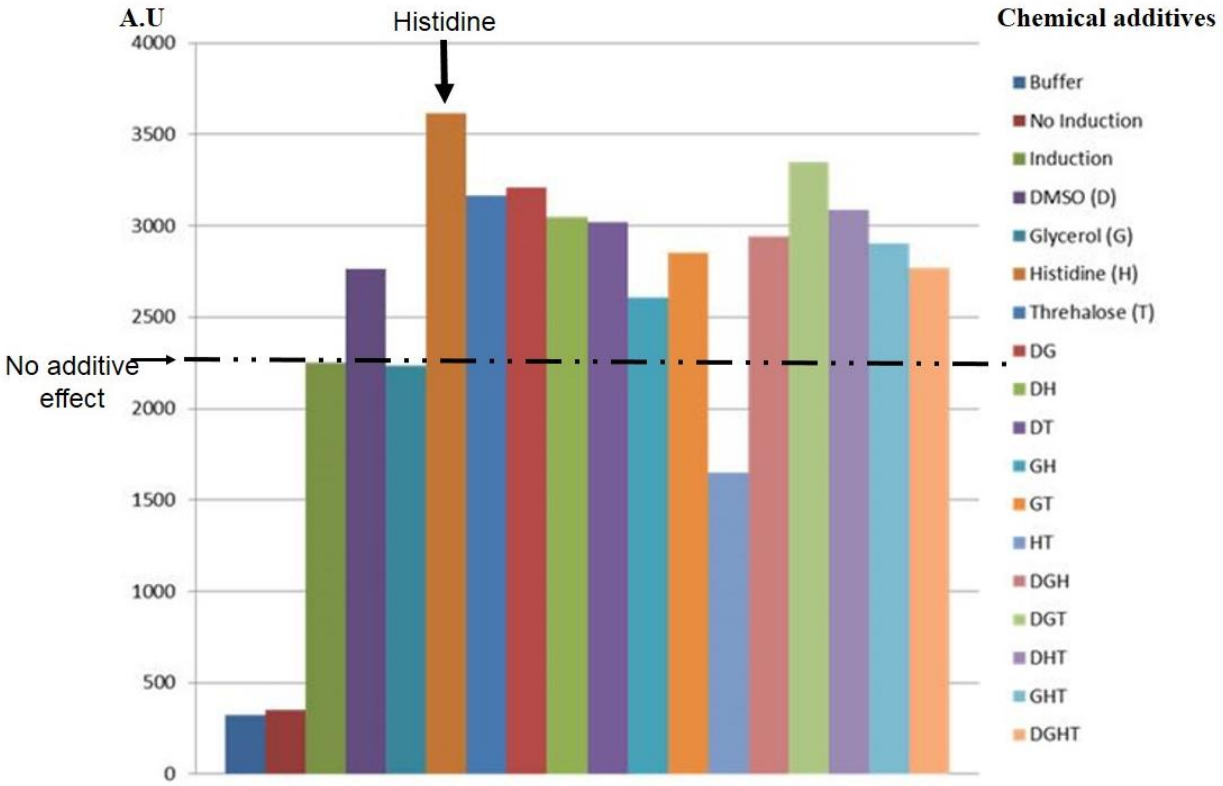


Figure 4.2.3 b. Effect of Chemical chaperone on Get2.GFP expression in yeast. Get2-GFP expression when induced in presence of DMSO, glycerol, histidine, trehalose, and mixture of those chemical chaperones to each other. Histidine on its own has increased the concentration of GFP to 0.4 mg/ml therefore Get2 to 0.35 mg/ml and a gain of about 1.8 mg per liter culture.

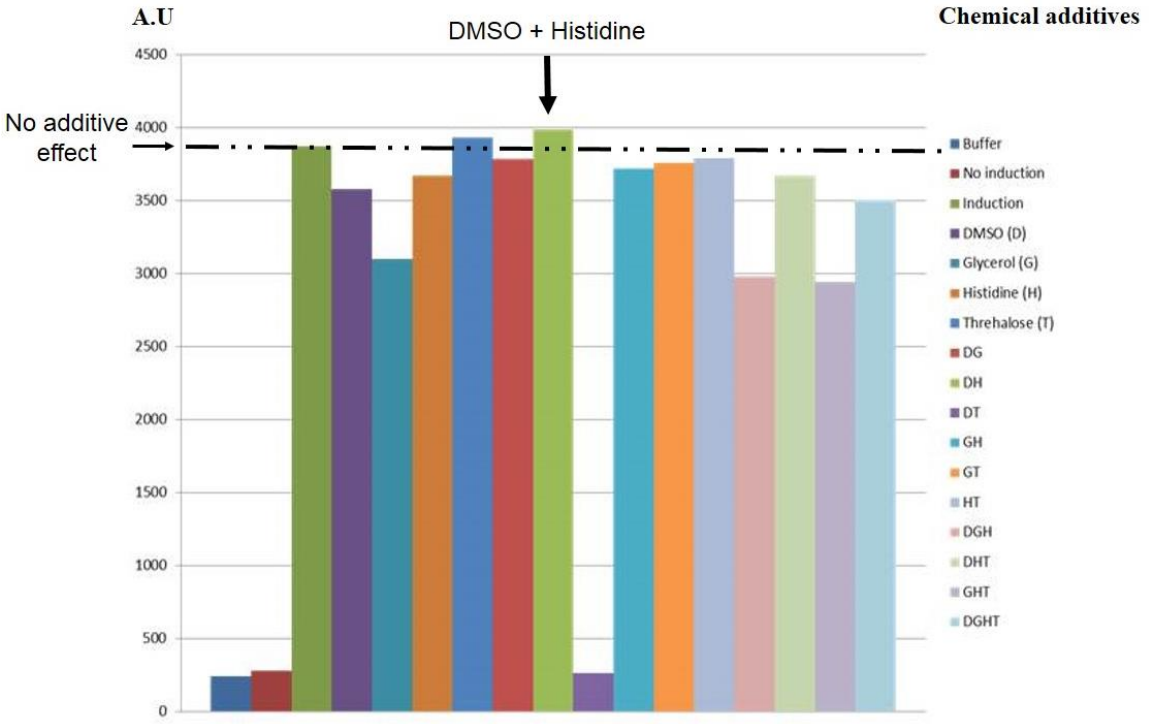


Figure 4.2.3 c. Effect of Chemical chaperone on Get2/Get1.GFP expression in yeast. Get2/Get1-GFP expression when induced in presence of DMSO, glycerol, histidine, trehalose, and mixture of those chemical chaperones to each other. No real effect of chemical additive is observed for the single chain Get2/Get1.GFP. However very slight effect of histidine.

4.2.4 Detergent Screening

From the pre-established conditions above, a small scale preparation for screening was used for solubilization analysis. As the study goal was to reconstitute mainly the interaction between Get3 and Get1, it was assumed that the detergent where Get1 was stable would also be appropriate for Get2 and Get2/Get1. Therefore, Get1 was expressed and followed by the membrane preparation. The membranes were separately solubilized in presence of Foscholine 12, LDAO, DDM, and DM as described in the method section. Then each sample was analyzed by analytical gel filtration. The results show that Get1 eluted as a monodisperse peak and DDM was likely the most suitable detergent to continue the work compared to other detergents tested (Figure 4.2.4). However, Foscholine 12 may enable the separation of two species that might form a complex in DDM (Figure 4.2.4). The 1.22 ml elution peak represents the main elution peak of Get1.GFP, while the peak eluting at 2.09 ml represents the residual cleaved Get1.GFP together with buffer components such as salt. DDM was then selected for solubilization of Get1 and Get2/Get1 receptors.

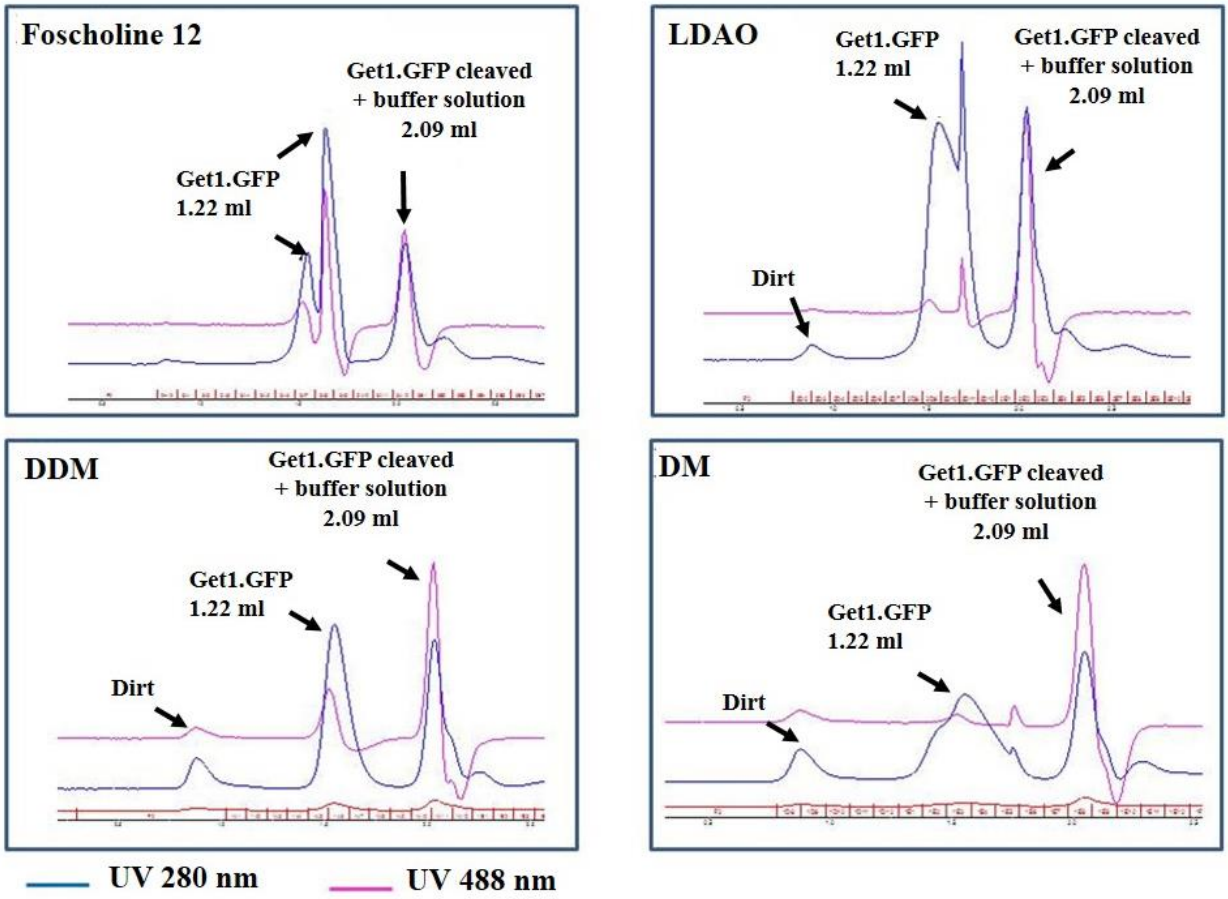


Figure 4.2.4: Detergent screening: Foscholine 12, LDAO, DDM, and DM were used to test the solubilization of Get1. As the Äkta system does not have a wavelength detection around 505 nm, 488 nm was used to detect the whereabouts of GFP fluorescence intensity signal. The proteins were excluded at a flow rate of 0.005 ml/min using the Superdex 200 PC 3.2/30 column for analytical studies using Äkta FPLC (GE Healthcare). Except for the top left (DDM) figure, the first peak was from residual components of the previous run. Get1.GFP is excluded mainly at 1.22 ml with shoulder or shoulder (DM) or a second peak (LDAO or Foscholine 12). The sample prepared in DDM seems to be more homogenous.

4.2.5 Fast Ni-NTA Affinity Purification Test

Once DDM was selected as the detergent for solubilization, all procedures were then executed and tested for large scale overexpression. Get1 and Get2/Get1 fused with GFP were overexpressed in a liter culture; the membrane was first extracted and then the proteins were solubilized in 1% DDM. Thereafter, a quick nickel-nitrilotriacetic acid (Ni-NTA) affinity purification followed by western blot identification was performed. The results showed that the solubilized Get1 and Get1/Get2 single chain can also be purified. However, the western blot analysis demonstrated that GFP was cleaved off (Figure 4.2.5), therefore it was necessary to simplify the system.

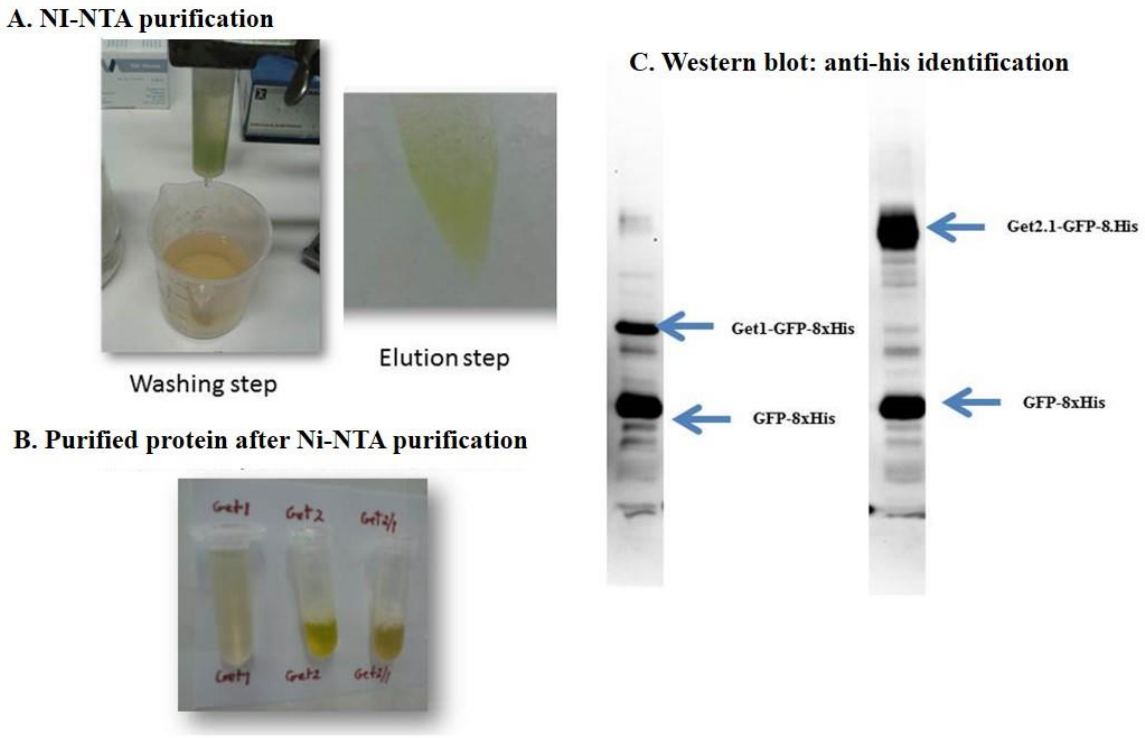


Figure 4.2.5: Immolilized-Metal Affinity Chromatography. After the overexpression and the solubilization of Get1.GFP, Get2.GFP and Get2/Get1.GFP, the 3 fusion proteins with GFP were used to test their purification by using Ni-NTA affinity chromatography A. where they were fire bound to Ni-NTA beads by gently mixing them in presence of 10 mM imidazole, 0.1% DDM for an hour followed by a first 10 column volume wash in presence of 10 mM imidazole and 0.1% DDM, a second 20 column volume wash in presence of 20 mM imidazole and 0.05% DDM, a third wash with 20 column volume in presence of 45 mM imidazole and 0.03% DDM, finally the proteins were eluted with 250 mM imidazole and 0.03% DDM. In B show the purified protein after elution. C. The proteins were then identified by anti-his western blot.

After it was notified that the fusion proteins were instable due to the cleavage of GFP, all procedures were optimized by removing GFP. The complete pDDGFP-2 vector was adapted by replacing the GFP.8xHis by either a 10xHis tag or a strep tag.

4.3 Optimization of the expression vectors by GFP fusion removal

Once the expression, the solubilization and the Immobilized-Metal Affinity Chromatography (IMAC) purification were optimized using the GFP system, the GFP.8xHis from the initial vector pDDGFP-2 were replaced by oligos encoding for either a strep-tag or 10xHis-tag. The series of expression vectors were extended to those with different yeast selection markers like *HIS3*, *LEU2* or *TRP1*. The procedure of how those expression vectors were obtained is shown as follows:

Construction of the cloning cassettes

Two sets of complementary oligos of strep-tag and 10xHis-tag were used to create the strep cloning cassette and the 10xHis cloning cassette (Figure 4.3).

Each cassette was flanked with *Bam*HI-*Hin*dIII restriction enzymes at the 5' end and with *Xho*I at the 3' end (5'-*Bam*HI-*Hin*dIII-tag-*Xho*I-3') resulting in 5'-GATC sequence overhang and AGCT-3' overhang (Figure 4.3).

The pJANY-S1 and pJANY-H1 expression vectors were generated as backbone for protein expression. JANY stands for Jean Aymard NZIGOU yeast plasmid and S1 strep-tag and 1 for *URA3* marker, likewise the H1 for 10xHis tag in vector that carry *URA3* selection marker (Figure 4.3)

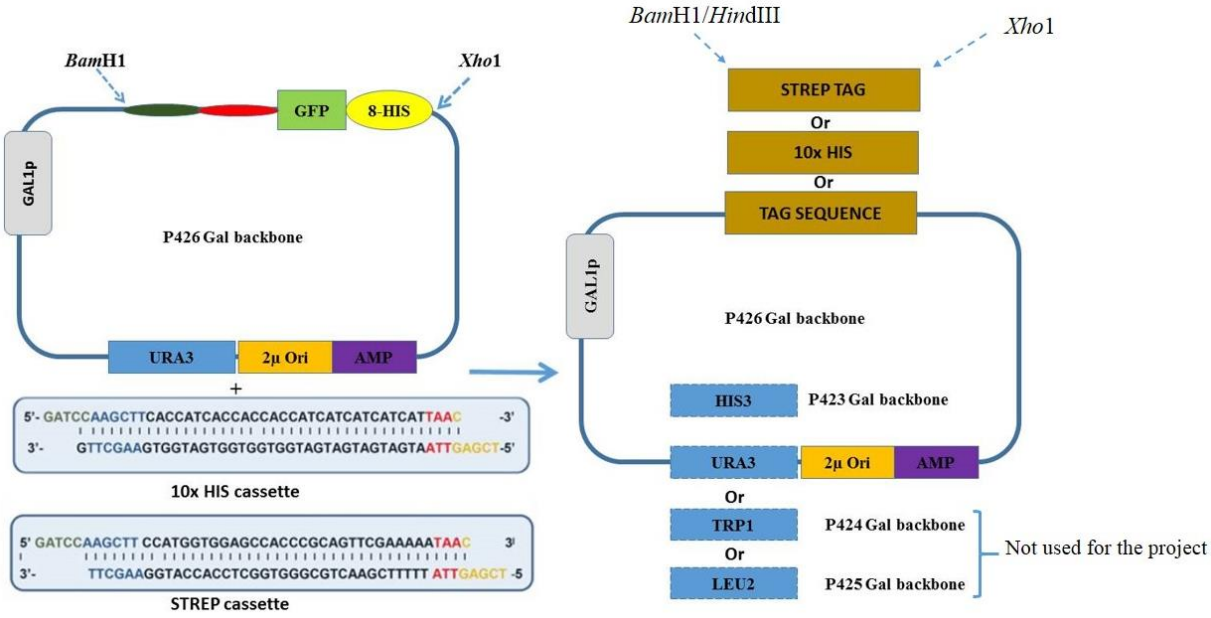


Figure 4.3: Schematic of the construction of the optimized expression vectors pJANYs. The pDDGFP vector was restricted by *Bam*HI and *Xho*I and replaced either by strep tag or 10xHis tag, they were named respectively pJANY-S1 and pJANY-H1. Also, vectors carrying the *HIS3* selection marker were constructed based on those two expression vectors

4.4. Interaction of Get3 with its membrane bound receptors

Prior to the interaction studies, each receptor and Get3 were prepared separately. Get3 was prepared according to standard existing protocols (Stefer *et al.*, 2011). Get3 has been shown to interact strongly with the cytosolic and unstructured part of Get1 and Get2 (Stefer *et al.*, 2011) but nothing was done with the full protein in micelles. Therefore, before proceeding with the structural studies, the interaction of Get3 with its membrane bound receptors was tested. To prepare Get3 membrane bound receptors, pJANY vectors (with either 10xHis tag or strep tag) were used to carry the gene encoding for each receptor, and the vectors were constructed by homologous recombination as described in the methods section.

4.4.1 Interaction studies of Get3 with Get1

In the aim of reconstructing the complex Get2/Get1/Get3, the initiative was to first show that the interaction of Get3 with Get1 can be obtained without the cytosolic part of Get2. The interaction of Get3 and Get1 was studied and the result of this interaction could be extended to any single chain construct lacking the N-terminal part of Get2 or any construct of the single chain where the N-terminal part was substituted by a chimeric peptide.

4.4.1.1 Expression and purification of Get3

Get3 was expressed in T7 express competent *E. coli* (NEB) with the pET24d derivative vector containing the Get3 gene and kanamycin resistance. Get3 was purified in three steps as described in the methods section. Get3 shows a nice monodisperse peak with a shoulder. Get3 has two oligomerization states, the functional dimeric state and the tetrameric state. The shoulder seen in the Get3 simply could represent the minor tetramer (Figure 4.4.1.1). Although the SEC images presented are from HiLoad 16/600 Superdex 200 pg (GE Healthcare), the gel filtration was also done using the 10/300 GL superdex 200 (GE Healthcare).

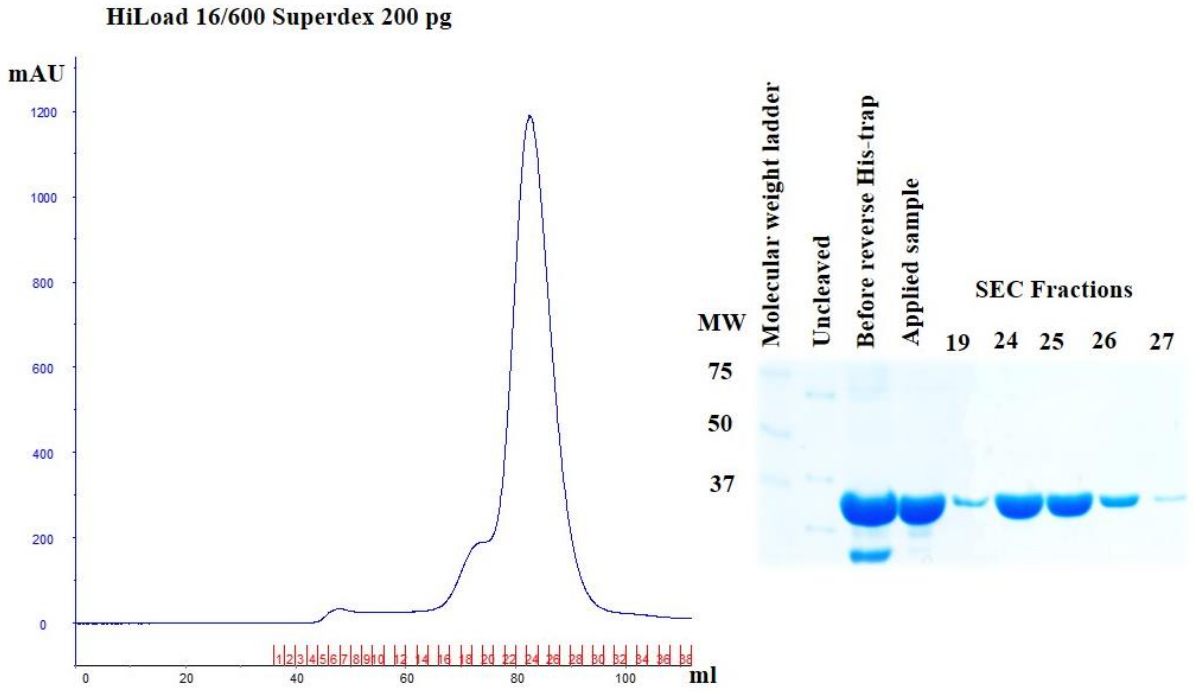


Figure 4.4.1.1 SEC of Get3. Get3 was purified in 3 main steps, Ni-NTA affinity chromatography, TEV cleavage followed by the Ni-NTA reverse affinity chromatography and the size exclusion. 19, 24, 25, 26, and 27 are fractions from the gel filtration.

4.4.1.2 Expression, purification, and thermostability of Get1

To express and purify Get1, the vectors carrying Get1 were first prepared, named pJNG1-S1 (Get1-strep) and pJNG1-H1 (Get1-10xhis). The index 1 in S1 and H1 are the vectors carrying Ura3 marker but in addition vectors carrying the His3 marker were also prepared and indexed S2 (strep) and H2 (10xhis). To analyze the interaction of Get1 with Get3, Get3 was obtained as described in the previous chapter but sometimes the His-tag of Get3 was not digested for the tandem purification purpose. For the interaction studies, as Get3 was purified often with his tag, pJNG1-S1 was used for expression. After the expression of Get1, Get1 was solubilized as previously described and then purified by strep purification (Figure 4.4.1.2 A). The thermostability of Get1 was further analyzed by CD spectroscopy (Figure 4.4.1.2 B).

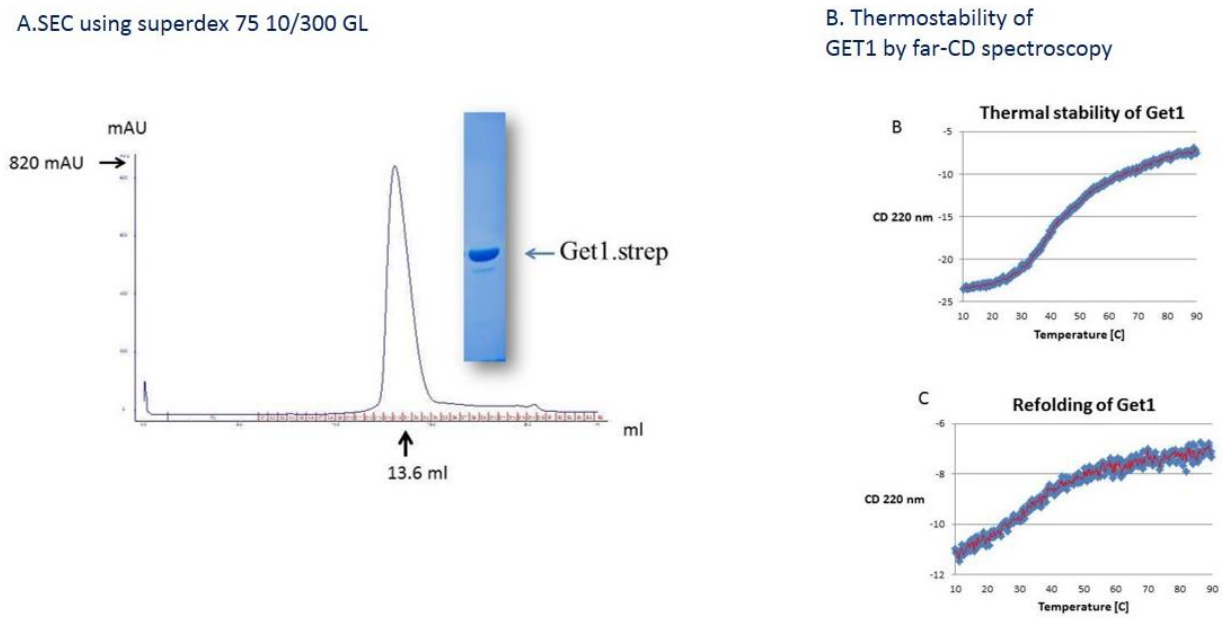


Figure 4.4.1.2 Get1 purification and its thermostability. Get1 was purified at pH 7.6 and shows a pure monodisperse peak (A). Get1 was purified using Superdex 75 10/300 GL column and elute around 13.6 ml. Get1 has stability of 45 degree (B) and can be refolded (C).

Analysis of the secondary structure and thermostability of Get1 by far-CD spectroscopy was recorded between 190 and 260 nm. The secondary structure of Get1 shows a spectrum of α-helical protein (the data is not shown because of very low signal). The change ellipticity at 220 nm as indicative of the thermostability was analyzed through the thermal denaturation of Get1 between 10 and 90^0C. The melting point of Get1 was found to be 45^0C. Get1 can be partially refolded.

4.4.1.3 Crystallization of Get1

In the meantime, pure Get1 was used for crystallization. 768 conditions were screened for crystallization, some crystalline material appeared only three months later but the crystals which appeared after 100 days did not look like regular protein crystals (Figure 4.4.1.3). Get1 did not diffract and considering the fact that Get1 has a Tm of 45 degree Celsius (Figure 4.4.1.B), no further testing were done on them as they were not stable.

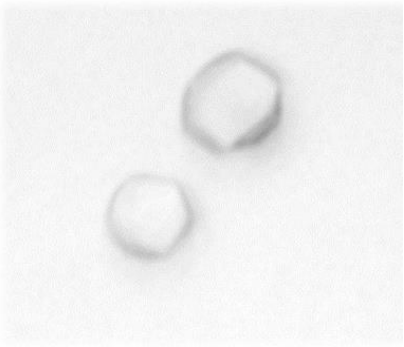


Figure 4.4.1.3 Crystallization Get1. Get1 crystals do not show a regular protein crystal shape. They did grow in presence of 0.05 M magnesium acetate, 24 % (v/v) polyethylene glycol 400 and 0.05 M sodium acetate.

4.4.1.4 Interaction of Get3 and Get1

Get1 and Get3 were prepared separately, then they were brought together during the solubilization of Get1 in presence of 1% DDM. After the solubilization they were purified in tandem, first by Ni-NTA chromatography affinity, then strep affinity chromatography. They can be purified as a complex (Figure 4.4.1.4.1) as SDS-PAGE after the second purification step showed a nice interaction between Get3 and Get1.

At this point the question was the following: How strong was the interaction between full Get1 and Get3? The affinity studies for direct interaction between the cytosolic Get1 and Get3 were done before, but nothing was done with the full protein Get1. Therefore, a MST assay was performed to show the direct interaction between Get1 and Get3.

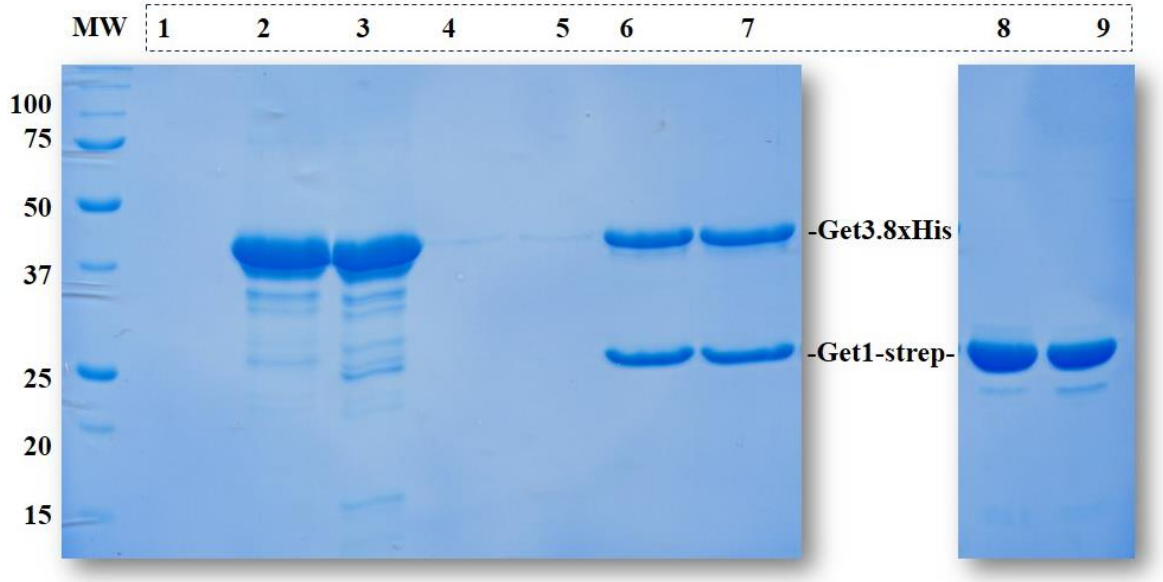


Figure 4.4.1.4.1: Get1 and Get3 interaction. The complex was first purified by His-tag affinity chromatography then by strep-tag purification. 1. Flow through; 2. Heated SDS sample of the washed fraction 1; 3. Washed fraction 1 without heating the sample; 4. Heated SDS sample of the washed fraction 5; 5. Washed fraction 5 without heating the sample; 6. Heated SDS sample of the elution Get1/Get3 complex; 7. SDS sample of the elution Get1/Get3 complex without heating the sample. Heated SDS sample of only purified Get1. SDS sample of only purified Get1 without heating the sample.

The Monolith NT.115 (NanoTemper) was used to determine the molecular interaction of full length Get1 and Get3 (Figure 4.4.1.4.2) in solution. The Monolith system did help to measure the affinity between Get1 and Get3. This system uses fluorescent dyes. In the titration curve, the fluorescence intensity is on the y axis and the concentration of the titrant [nm] on the x axis.

Initially, no bumps were observed in the thermophoresis with temperature jump (T-jump) measurement which implies that the samples were of good quality. Then, two changes were observed in the fluorescence intensity upon the binding of Get3. The binding of Get3 to Get1 confirmed the high affinity between the two proteins, in the nanomolar range. The experiment highlights two binding events occurring (Figure 4.4.1.4.2 and Figure 4.4.1.4.3). One specific binding with an affinity of 57 nM and a second unspecific binding of 740 nM.

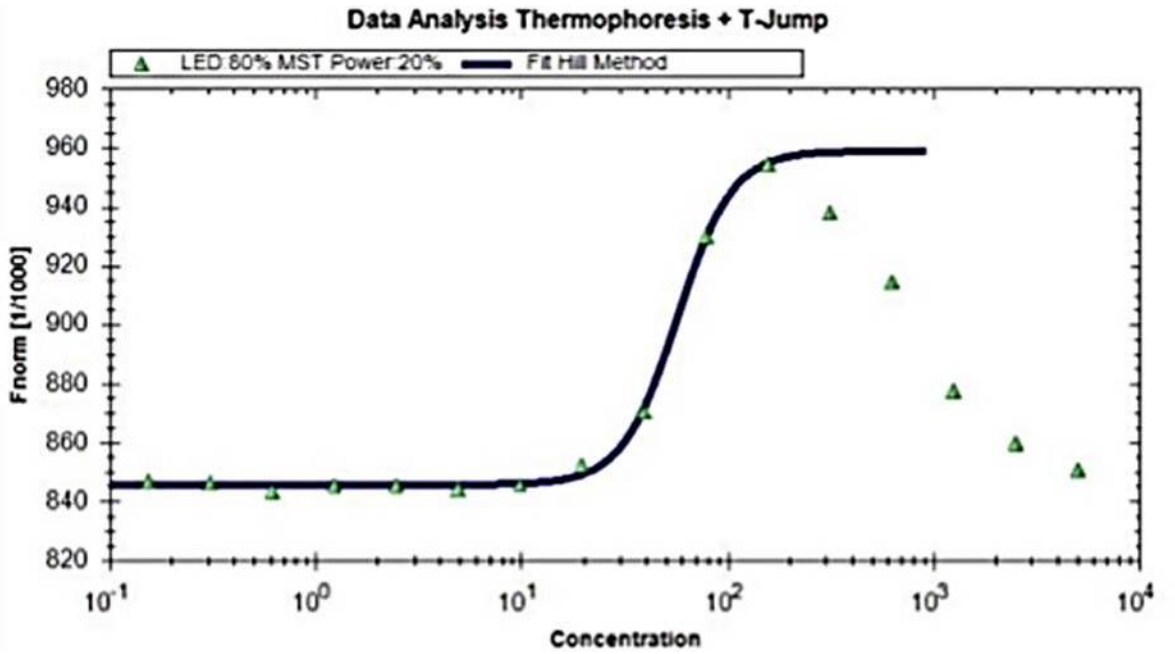


Figure 4.4.1.4.2: Titration of fluorescently labeled Get1-NTblue with non-labeled titrant Get3. The scanning was executed at the capillary position 80% light-emitting diode (LED) MST power of 20%. The concentrations on the x-Axis are plotted in [nM] while the y axis is fluorescence intensity counts. T-Jump is the thermophoresis with jump, it reports the binding events. No bumps are seen in the jump

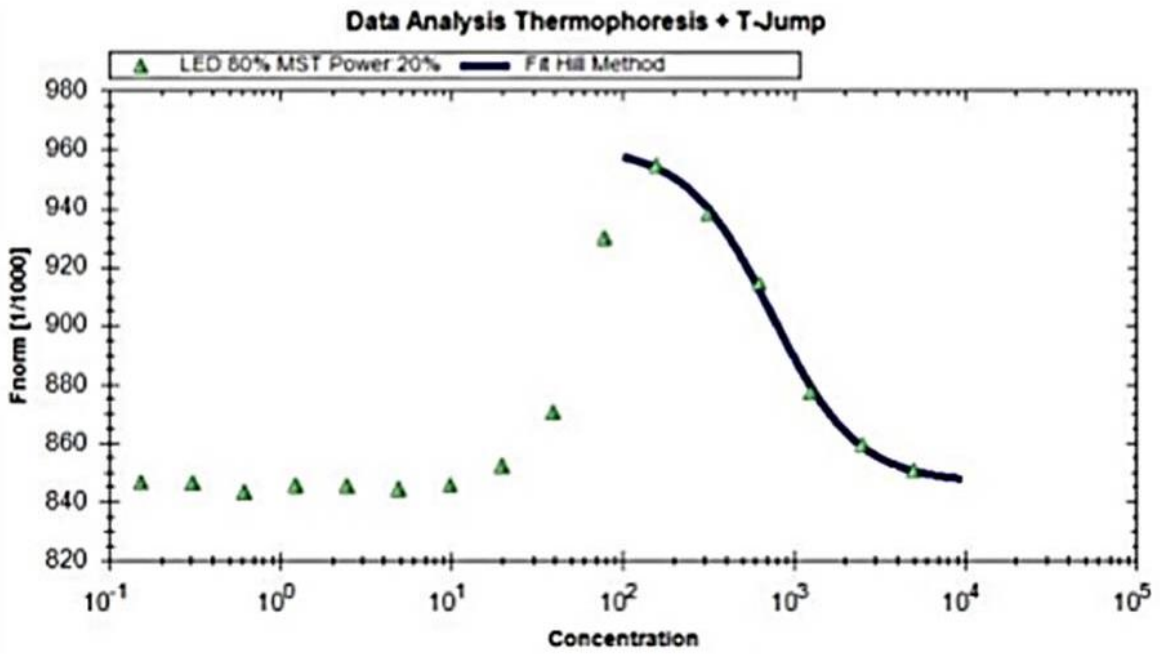


Figure 4.4.1.4.3: Biphasic behaviour of the binding between Get3 and Get1. The titration shows that there are two binding events through the biphasic binding isotherm: two IC_{50} were determined 57 nM (A) and the second of 740 nM (B).

The mass action law yields a shift of the 50% point depending on the concentration of offered binding partner, Get1. The results of the binding of Get3 to Get1 shows a biphasic behavior at equilibrium implying binding events which indicate heterogeneity in the binding.

4.4.2 Interaction studies of Get3 with Get2/Get1

The heterogeneity within the interaction between Get1 and Get3 suggested that the addition of Get2 would further stabilize that complex. As Get1 and Get2 interact within the membrane and would have been problematic to be purified when separately expressed, a single chain construct was created (Figure 4.1). Binding experiments with the Get2/Get1 fused with GFP confirmed that the single chain construct can interact with Get3 (Figure 4.4.2).

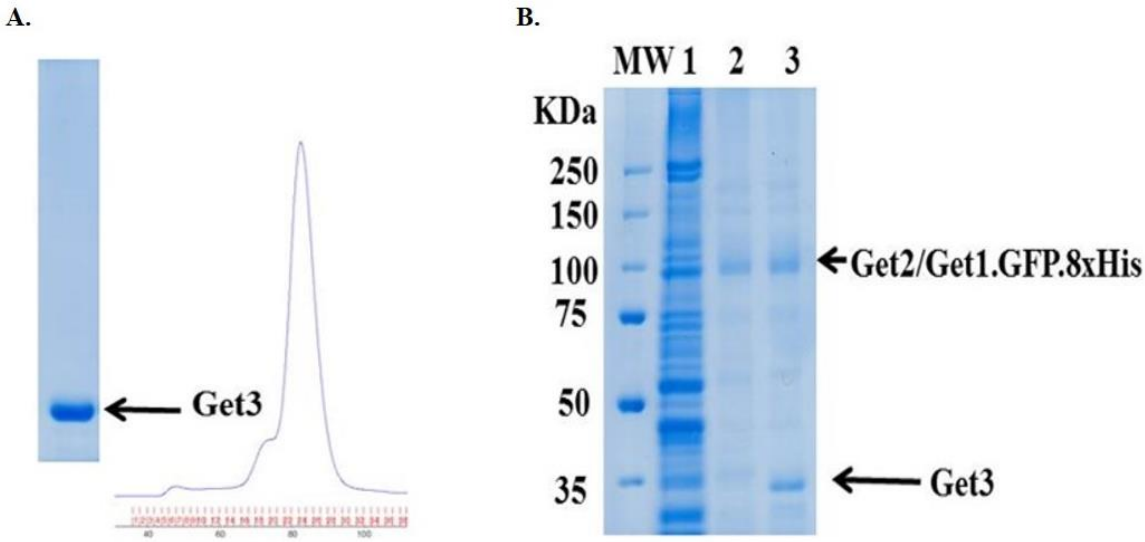


Figure 4.4.2: Get3 and Get2/Get1.8xHis interaction. Purified Get3 (A) was added in excess before Get2/Get1.GFP.8xHis His-tag affinity chromatography step and was purified with Get3. B. 1 flow through during the washed; 2. Control purified Get2/Get1.GFP.8xHis.

4.5 Characterization of the complex of Get2 and Get1

After heterogeneity was observed in the interaction of Get1 and Get3, and also because the expression of the single chain seemed to be better than Get1 only or Get2 only (chapter 4.2.1), it was thought that the interaction of Get2/Get1 will also improve the heterogeneity, therefore the complex Get2 and Get1 was characterized as single chain.

.

4.5.1 Purification of the single chain Get2/Get1 construct

The initial purification was done at pH 7.6 using the strep-tagged construct (Figure 4.5.1.1). The elution peak of the chromatogram was not symmetric on gel filtration.

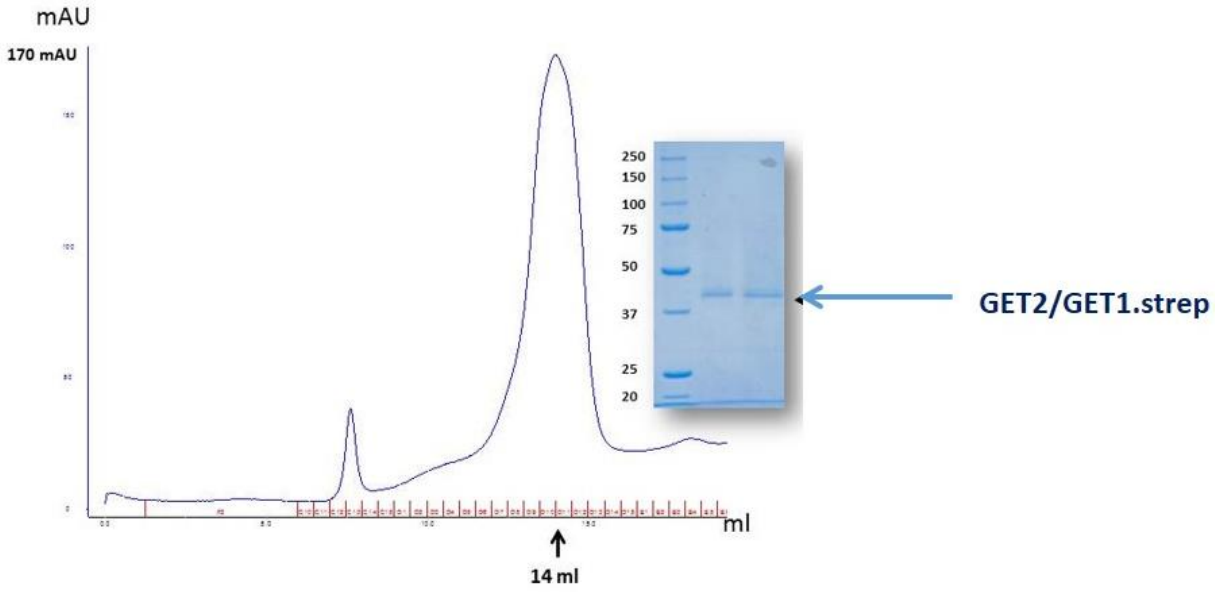


Figure 4.5.1.1: Purification of the full length Get2/Get1 at pH 7.6. The single chain construct was purified first using the Superdex 200 10/300 GL column and elute around 14 ml. Get2/Get1 is not really monodisperse at pH 7.6.

However, when the Get2/Get1 single chain protein was purified at pH 10 (Figure 4.5.1.2), the chromatogram looked more homogenous, the gel filtration peak was monodisperse and symmetric, but the SDS PAGE and western blot analysis showed multiple smaller bands. When all the bands were identified by MS/MS, they were part of the Get2/Get1 protein

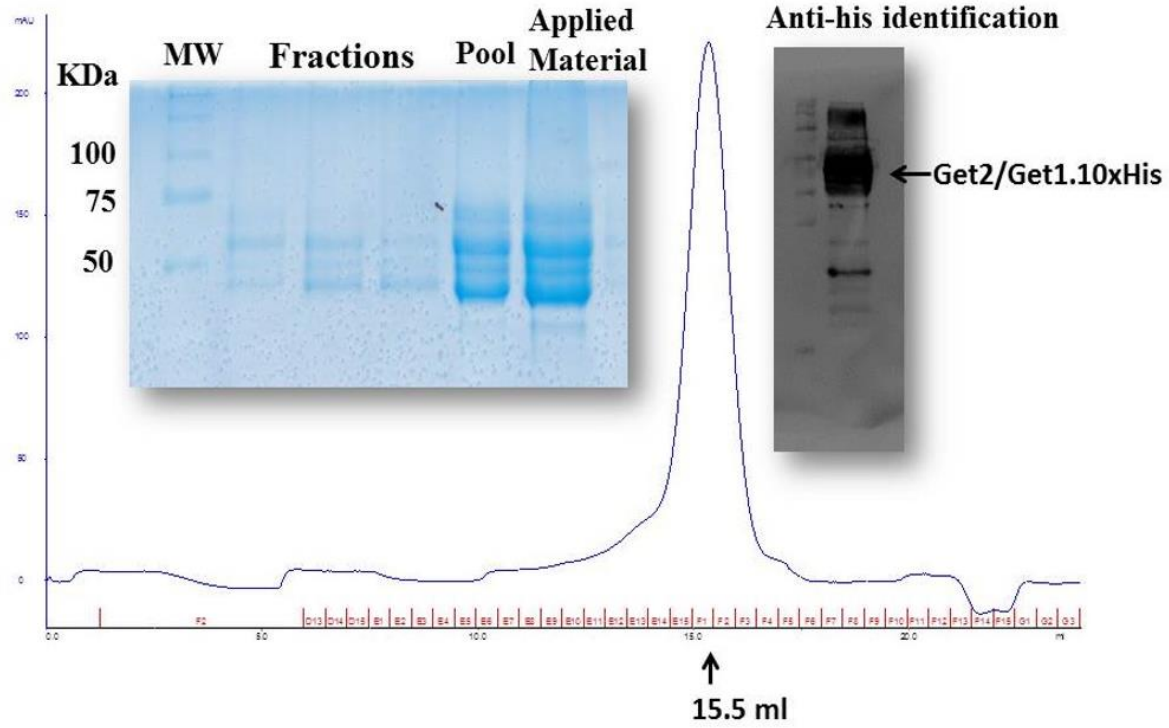


Figure 4.5.1.2: Optimization of the purification of the full length Get2/Get1 . The single chain construct purification in $Na_2CO_2/NaHCO_2$ buffer pH 10. Get2/Get1 elute as a symmetric monodisperse peak on gel filtration around 15.5 ml. The fractions on the SDS PAGE are the fractions around 15.5 ml (not visible from the chromatogram here but they are fractions F1, F2, and F3). The pool is F1, F2, and F3 fractions. The western blot identification image on the left is from the pool fractions

4.5.2 Stoichiometry and crystallization studies of the single chain Get2/Get1 protein

After the purification of the Single chain Get2/Get1 protein (Figure 4.5.1.2), the stoichiometry of Get2/Get1 was determined by LILBID (Figure 4.5.2) in collaboration of Prof. Dr. Morgner (Goethe University, Frankfurt, Germany). LILBID is a method used to analyze large integral membrane protein complexes and their subunits. The ions in LILBID are IR-laser desorbed from micro droplets containing membrane protein complexes in detergent. LILBID is highly sensitive, and very efficient in sample handling. One can use a wide range of buffers (Morgner *et al.*, 2007).

LILBID analysis showed two signals of 50 KDa, 61 KDa, a higher molecular weight signal and two lower molecular weight signals (Figure 4.5.2).

Considering the three following facts:

Firstly, the DNA sequence encoding for the His-tag used for purification was positioned downstream of the sequence encoding for Get2/Get1. Secondly, based on the topology of Get2/Get1 (Figure 4.1), the cytosolic N-terminus of Get2 is unstructured, flexible and

therefore prone to be cleaved off. Thirdly, the identification of all SDS PAGE multiple bands of the purified Get2/Get1 **(**Figure 4.5.1.2**)** by MS/MS showed that they are either Get1 or Get2 which imply that they are part of the single chain Get2/Get1 protein.

Thus, it can be concluded that the signal of 61 KDa is the intact monomeric Get2/Get1, the signal of 50 KDa is the Get2/Get1 without the unstructured cytosolic N-terminal part of Get2. The first lower peak is the signal of the negatively double charged protein corresponding to the monomer signal and the second lower peak is the signal of the negatively triple charged protein of the dimer. The higher signal has lost the unstructured N-terminal part of Get2. At this point the N-terminal cleavage of Get2 seems to be the cause of the heterogeneous mixture of the single chain heterodimer Get2/Get1 (Figure 4.5.2).

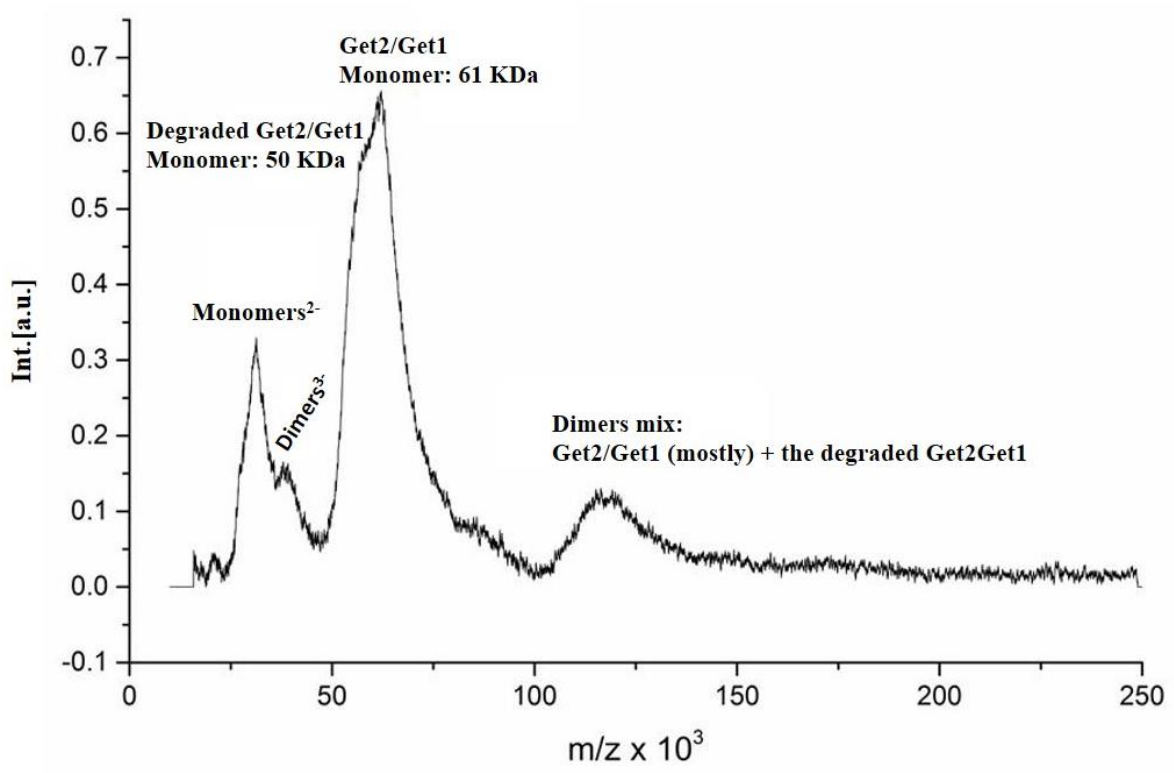


Figure 4.5.2: LILBID analysis of Get2/Get1 single chain construct. The stoichiometry analysis of the single chain Get2/Get1 shows that there are two signals of monomeric Get2/Get1 (50 KDa and 61 KDa) and dimers mix signals. One is the degraded Get2/Get1 corresponding to a molecular weight of 50 KDa and the second is the signal of the intact Get2/Get1 of 61 KDa. The dimers signal is mostly from the intact Get2/Get1, but also there is dimers mixes between both types of Get2/Get1 monomers.

In the meantime, the crystallization of Get2/Get1 was set as described in the method chapter 3.19. Crystals of Get2/Get1 were initially obtained and optimized around the following conditions on 1µl sitting drops: 200 mM $CaCl_2$, 50 mM glycine pH9, 30% PEG 400. Although the crystals appeared after 3 days, none of them diffracted which makes sense when

we consider the heterogeneity of the dimer (Figure 4.5.2). At this point the N-terminal cleavage of Get2 within the Get2/Get1 protein expressed as single chain seems to be the cause of the heterogeneous mixture of the single chain heterodimer Get2/Get1.

4.5.3 LILBID analysis of truncated Get2/Get1 without the unstructured N-terminal cytosolic Get2 domain

Based on the purification of Get2/Get1 **(**Figure 4.5.1.2), the LILBID analysis of Get2/Get1 (Figure 4.5.2) and the outcome of the crystallization test of Get2/Get1 data, it was thought that the complete truncation of the unstructured N-terminal cytosolic Get2 part will likely improve the homogeneity and might consequently facilitate the biochemical and the crystalization studies. Hereafter, the first 128 amino acid residues of Get2/Get1 were removed (Figure 4.5.3 A) and the construct was cloned into the pJANY-S1 plasmid resulting in the new pJNGt21-S1 plasmid. In addition to the single chain contruct of Get2/Get1, the N-terminal part of Get2 was deleted in the Get2 expresion plasmid as well.

In parallel, the highly crystallizable T4 lysozyme (T4L) was fused with the N-terminus of Get2/Get1 replacing the unstructured cytosolic part of Get2 (Figure 4.5.3 B)**,** with the idea that it will facilitate the crystallization of Get2/Get1. The same approach was used for human beta2 adrenergic receptor, a G-protein coupled receptor (GPCR) (Zou *et al.*, 2012). The details of the construction can be found in the appendix.

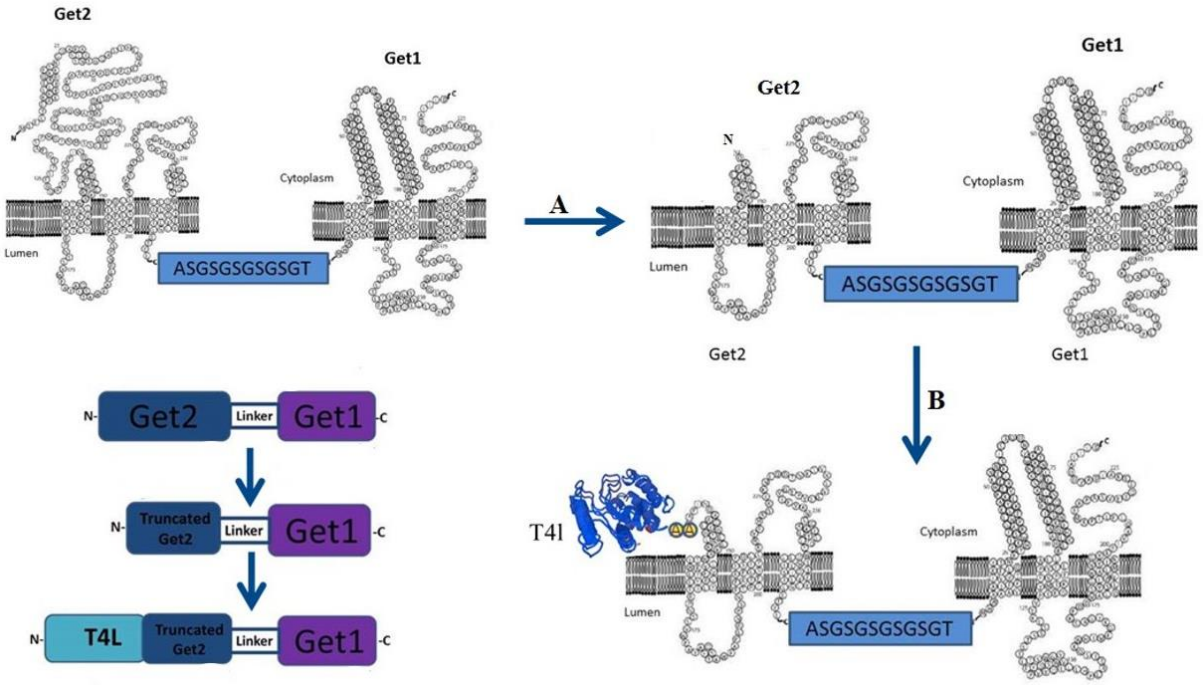


Figure 4.5.3.1: Topology of the truncated_Get2/Get1 (or tGet2/Get1) and T4l.Get2/Get1. A) tGet2/Get1 was cloned into pJANY-S1 resulting the vector pJNGt21-S1. B) From the vector pJNGt-21-S1, the T4 lysozyme sequence was added to generate T4l.Get2/Get1 single chain construct. Four vectors carrying T4l.Get2/Get1 were prepared. Two carry Ura3 marker, pJNGT421-S1 and pJNT4LG21-H1. The other two carry the His3 marker; pJNT4LG21-S2 and pJNT4LG21-H2. The topology of tGet2/Get1 and T4l.Get2/Get1 were adapted from Stefer *et al.*, 2011.

Purification of the single Chains truncated_Get2/Get1 and T4l.Get2/Get1

The tGet2/Get1 and T4l.Get2/Get1 single chain constructs with 10xHis-tags were expressed and then purified. With respect to the multiple degradation bands previously observed with the intact Get2/Get1 (Figure 4.5.1.2)**,** the new variants tGet2/Get1 and T4l.Get2/Get1 were more stable (Figure 4.5.3.2).

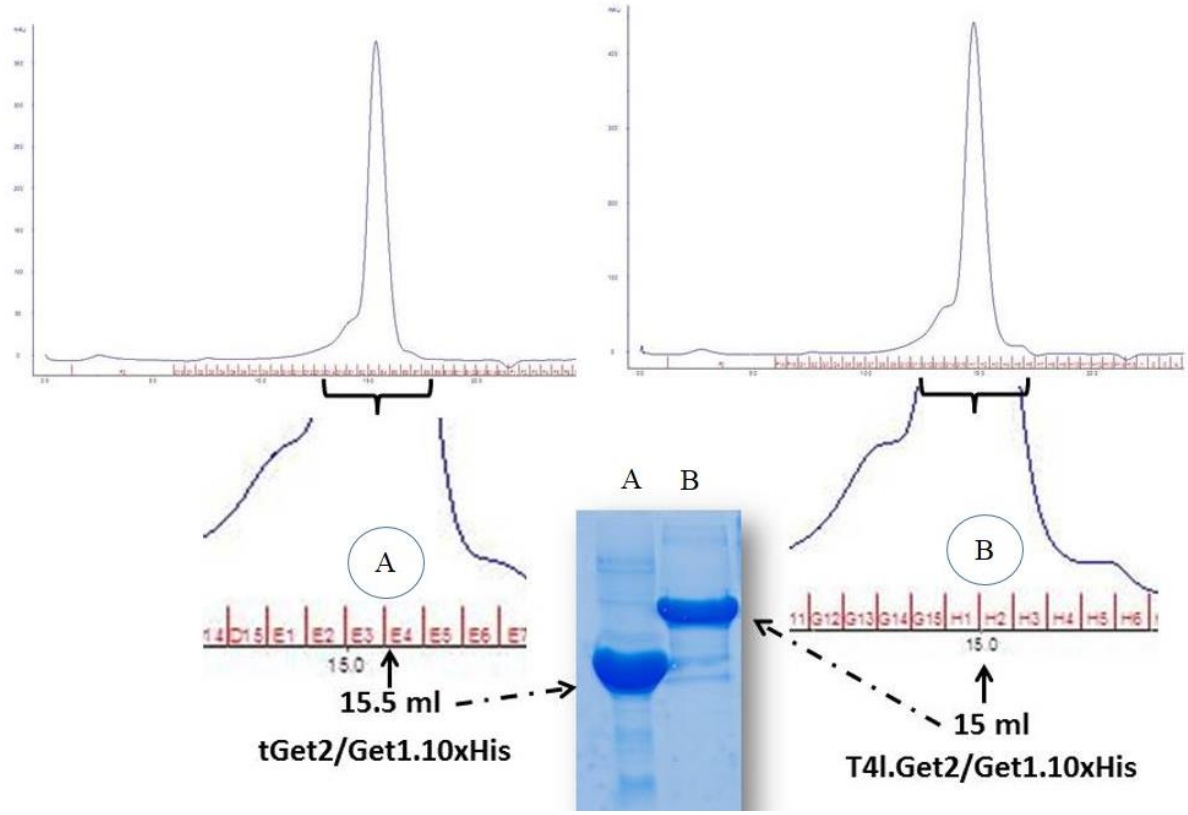


Figure 4.5.3.2: Purification of tGet2/Get1 and T4l.Get2/Get1. Both proteins were purified during the SEC at pH10 using superpose 6 10/300 GL. Their chromatograms are monodisperse and their peaks are symmetric. The SEC of tGet2/Get1 has its main peak at 15.5 ml and the fractions E3, E4, and E3 of tGet2/Get1 (A) were pooled together and run on SDS PAGE. Likewise, T4l.Get2/Get1 has its main peak at 15 ml and the G15, H1, and H2 of T4l.Get2/Get1 (B) were pooled together. Both chromatograms have shoulder suggesting a higher oligomerization state compared to their predominant elution peak.

Stoichiometry of the single Chains truncated_Get2/Get1

LILBID analysis of tGet2/Get1 was used to determine if the new variants were indeed improved versions of the full length Get2/Get1, and that the unstructured cytosolic N-terminal part of Get2 was really the cause of the mixture between the full Get2/Get1 (61 KDa) and the degraded (50 KDa) form. The LILBID analysis showed that the tGet2/Get1 is cleaner and more homogenous. (Figure 4.5.3.3).

LILBID results also confirmed that the single chain Get2/Get1 is really a dimer (heterotetramer of Get2/Get1; two Get2, and two Get1) but higher oligomers may exist as well because of the presence of shoulders seen on the chromatogram of the purification of tGet2/Get1 and T4l.Get2/Get1 constructs (Figure 4.5.3.2**)**. In addition, the results show that the degradation of Get2/Get1 (from 61 KDa to 50 KDa) was happening within the unstructured cytosolic N-terminal domain of Get2.

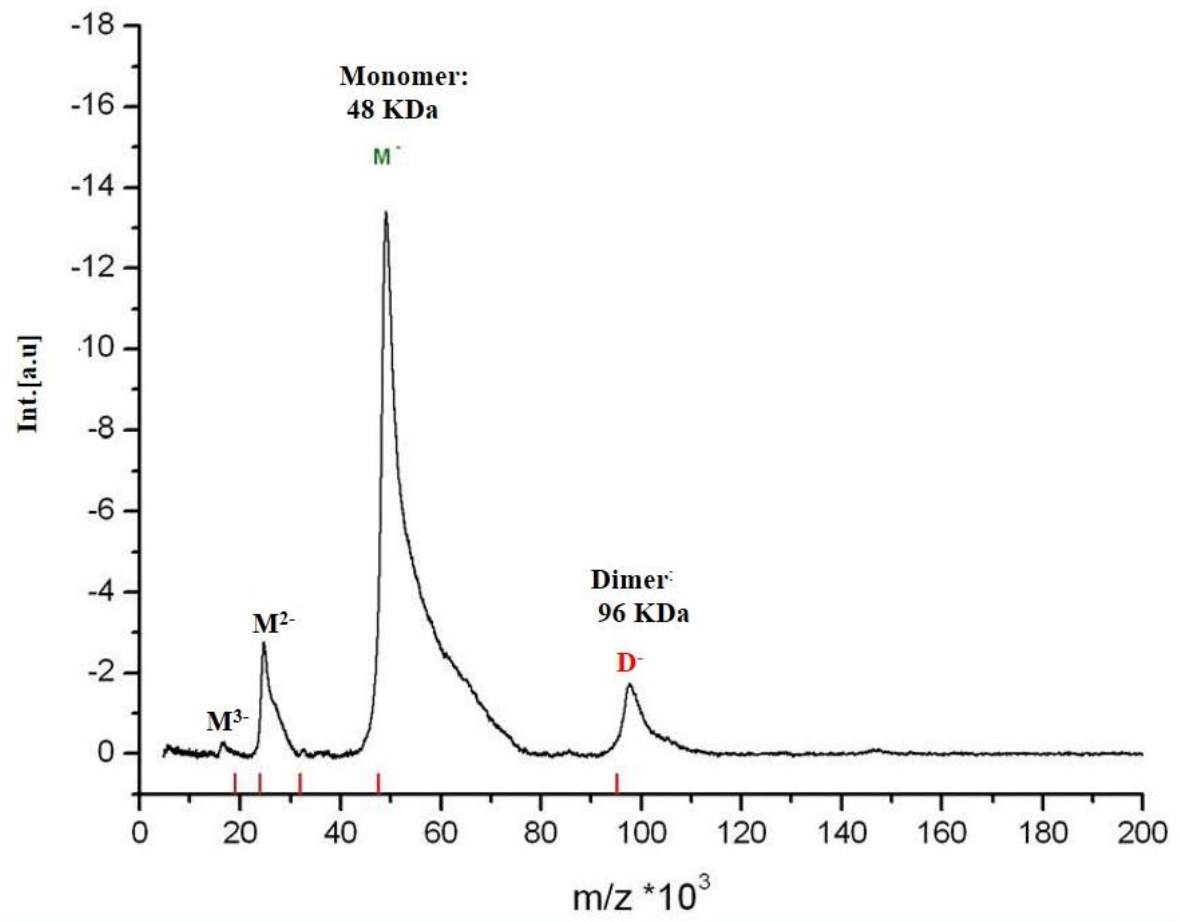


Figure 4.5.3.3: LILBID analysis of tGet2/Get1 single chain construct. The analysis shows the existence of a clear dimer of tGet2/Get1. M^{2-} and M^{3-} are double and triple respective negative charges of the monomer.

4.5.4 Crystallization of the single Chain variants tGet2/Get1 and T4l.Get2/Get1

To obtain the Get3/Get2/Get1 complex for structural studies, and since Get3 was His-tagged, the Get2/Get1 variant constructs were prepared with strep-tags (pJNGt21-S1 and pJNT4LG21-S1) and used for overexpression and co-purification with His-tagged Get3. For both single chain variants no difference was observed when they were purified using their strep-tag in comparison to 10xHis-tag (data not shown).

Moreover, prior to structural studies of the single chain carrying the highly crystallizable T4l (Zou *et al.*, 2012) substitution of the unstructured cytosolic N-terminal of Get2, the interaction of T4l.Get2/Get1 with Get3 was analyzed to see if the chimeric T4l was interfering with the binding of Get3 to Get1 within the single chain T4l.Get2/Get1. The interaction of Get3 and T4l.Get2/Get1 was evaluated using the pulldown of Get3 by T4l.Get2/Get1 (Figure 4.5.4.1). In that experiment, Get3 was co-purified by Ni-NTA without the removal of His-tag by TEV cleavage. Get3 was added during the solubilization of T4l.Get2/Get1 before the purification, the mixture was then ultracentrifuged to remove the unsolubilized T4l.Get2/Get1 which

pelleted down. The supernatant was used for strep chromatography purification (Figure 4.5.4.1). The results show a Get3 band co-purified with T4l.Get2/Get1. The Band was identified by mass spectrometry as Get3.

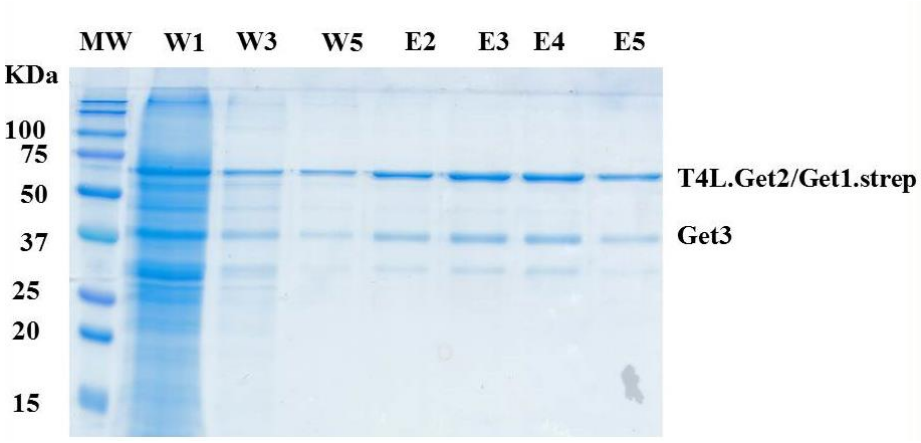


Figure 4.5.4.1: Get3 and T4l.Get2/Get1.strep interaction. Ni-NTA purified Get3-His was added in excess during the solubilization of T4l.Get2/Get1.strep then purified by affinity chromatography purification. Because of the number of lane the sample loading to washing fractions 1, 3, and 5 (W1, W2, W3) from the 5 washing fractions, and the elution to fractions 2, 3, 4, 5 (E2, E3, E4, and E5) from the 6 elution fractions. Get3 was co-purified with T4l.Get2/Get1.strep.

Because Get3 was co-purified with the strep affinity chromatography purification, it was now clear that T4l within the single chain construct T4l.Get2/Get1 does not interfere with the binding of Get3 to Get1.

Crystallization of the single Chains T4l.Get2/Get1 and truncated_Get2/Get1

The tGet2/Get1 (Figure 4.5.4.2 A. and B) and T4l.Get2/Get1 **(**Fig**u**re 4.5.4.2 C. and D) were used in crystallization experiments. However, none of the crystals that were obtained diffracted (home generator source). Interestingly, all the crystals grew only in presence of calcium.

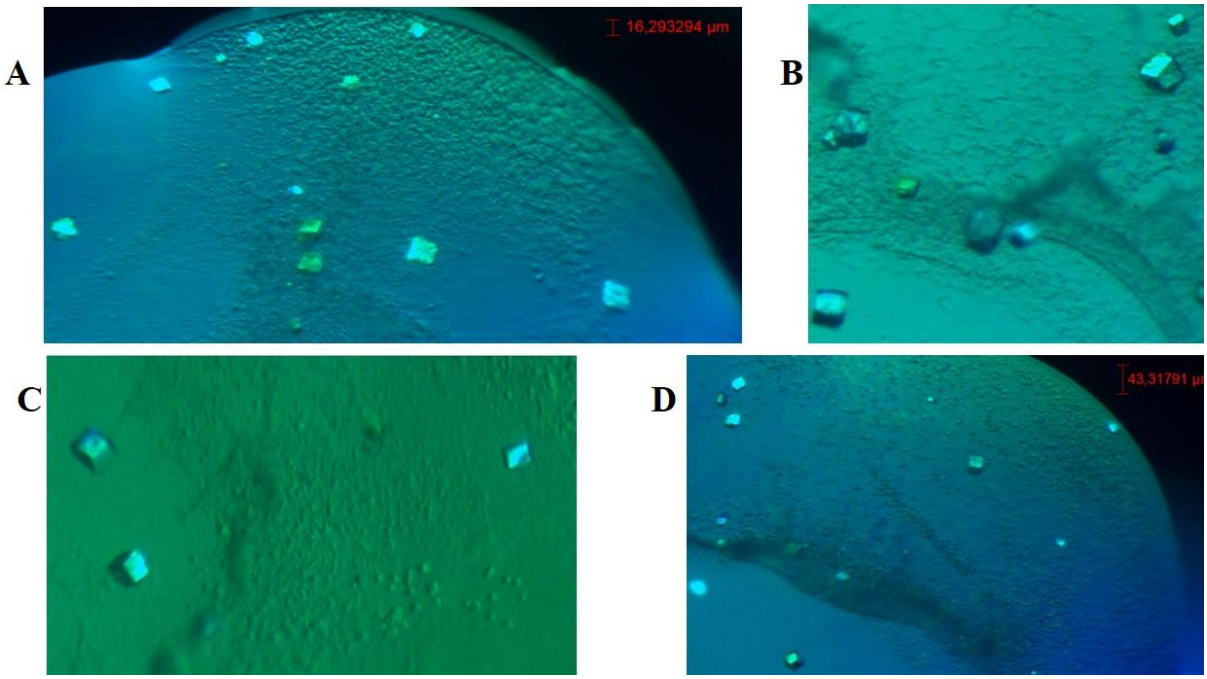


Figure 4.5.4.2: Crystals of tGet2/Get1 and T4l.Get2/Get1. A. crystals of t.Get2/Get1 grown in presence of 200 mM $CaCl_2$, 100 mM Tris pH8, 44% PEG400; B. crystals of t.Get2/Get1 grown in in presence 200 mM $CaCl_2$, 100 mM Mes pH 6, 26% PEGME350; C. crystals of T4l.Get2/Get1 grown in presence of 200 mM $CaCl_2$, 100 mM Mes pH 6, 33% PEG 400; D. crystals of T4l.Get2/Get1 grown in presence of 200 mM $CaCl_2$, 100 mM Hepes pH7.5, 53% PEG400.

4.5.5 Thermostabilization of T4l.Get2/Get1 by the apocytochrome b_{562}RIL linker

During the crystallization of tGet2/Get1 and T4l.Get2/Get1 it was observed that the crystals were not stable. Other research groups have shown with several GPCRs crystallization experiments that thermostabilized apocytochrome b_{562}RIL has superior characteristics than the T4l and increases the thermostability of GPCRs by 13 °C in comparison to T4l (Chun *et al.*, 2012). Besides the expressed T4l at the N-terminus of tGet2/Get1, it did not interfere with Get3 binding (chapter 4.5.4). Due to the distance between the N- and C- termini of thermostabilized apocytochrome b_{562}RIL (13.7 Å), and to improve the thermostability of the actual T4l.Get2/Get1, plus using the high crystallization capability advantage of both chimeric proteins (T4l and thermostabilized apocytochrome b_{562}RIL), the linker (ASGSGSGSGSGT) between Get2 and Get1 (Wang *et al.*, 2014) was replaced by the thermostabilized apocytochrome b_{562}RIL (Chun *et al.*, 2012).

The new vector carrying the single chain construct was digested by *Nhe*1 which is located within the linker sequence between the coding sequence of Get2 and Get1. Then the thermostabilized apocytochrome b_{562}RIL gene was amplified in such way that the overhang of the forward primer has a matching sequence with the last 40 nucleotides of Get2 (C-terminus) and the reverse primer has also an overhang sequence to the first 40 nucleotides of Get1

(Table 2.1). The single chain T4l.Get2/Get1 construct with the apocytochrome C linker was constructed by gap repair as previously described.

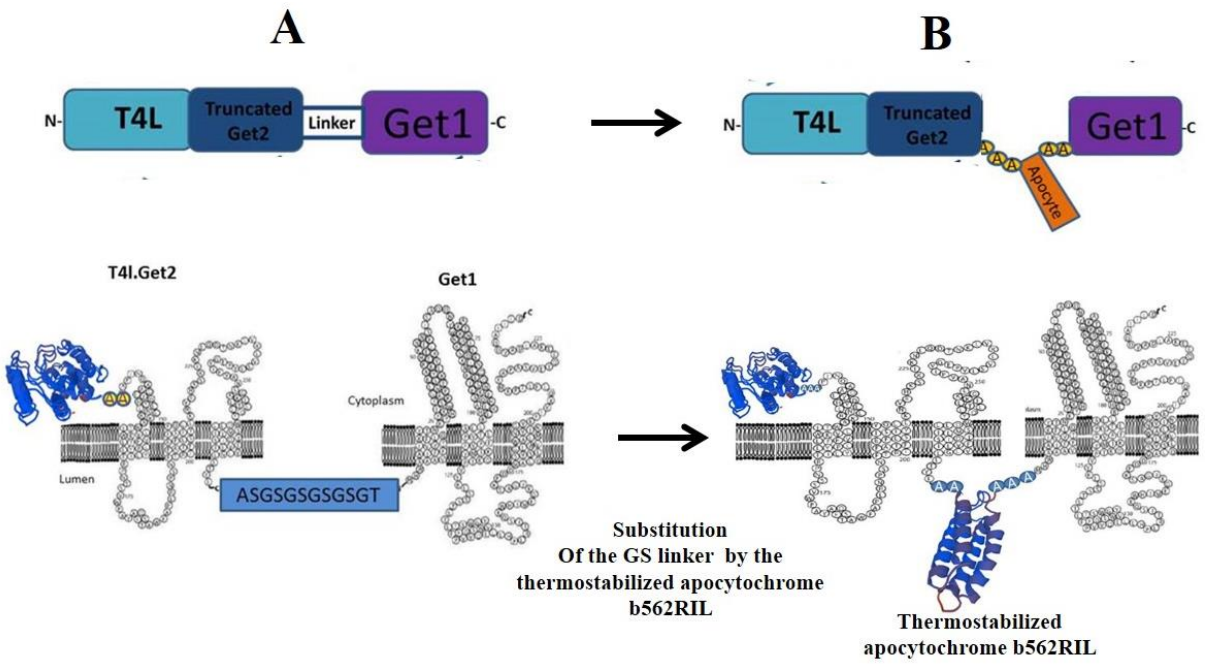


Figure 4.5.5: Thermostabilization of Get2/Get1. The GS linker between Get2/Get1 was replaced by the thermostabilized apocytochrome b_{562}RIL to improve the stability of the single chain construct. The resulting single construct was name T4l.tGet2.apocyte.Get1 and was cloned into pJANY-H2 resulting the vector pJNT4LtG2apoG1-H2 carrying HIS3 selection marker. The topology of the T4l.tGet2.apocyte.Get1 were adapted from the single chain construct Get2/Get1 Stefer *et al.*, 2011.

Purification of T4l.tGet2.apocyte.Get1

The optimum construct for the study consists of the single chain T4l.tGet2.apocyte.Get1 (Figure 4.5.5). Purified T4l.tGet2.apocyte.Get1.10xHis was cleaner, obvious degradation products are missing and the size exclusion chromatogram looks symmetric compared to other single chain construct variants previously purified (Figure 4.5.5.1). This single chain construct was purified using Na_2CO_2/$NaHCO_2$ (pH10) buffer as well.

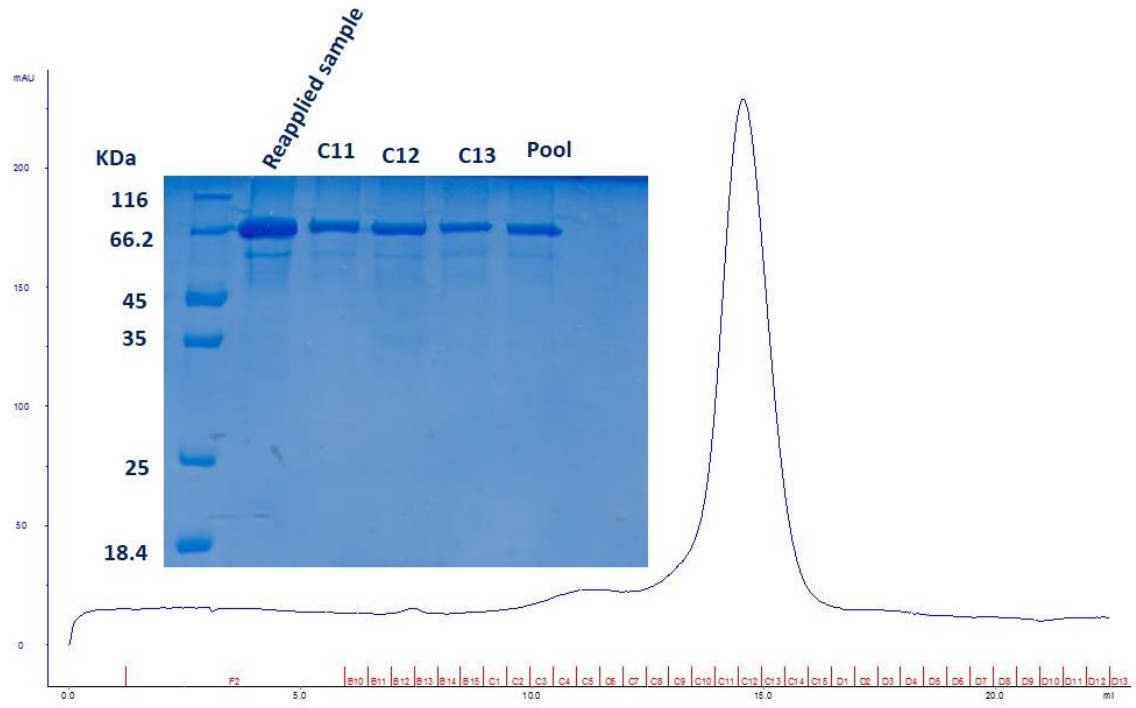


Figure 4.5.5.1: Purification of T4l.tGet2.apocyte.Get1. The purification of T4l.tGet2.apocyte.Get1.10xHis single chain construct shows a monodisperse peak at volume exclusion 14.5 ml and a clean SDS-PAGE bands.

Crystallization of the single chain T4l.tGet2.apocyte.Get1.10xHis

The thermostable T4l.tGet2.apocyte.Get1.10xHis did not improved the crystallization in terms of diffraction although the lipid cubic phase crystallization was also used (data not shown).

4.5.6 Summary of the molecular cloning using *S. cerevisieae*

During the course of this study the expression vectors evolved in function and improved the expression yield (Figure 4.5.6). First Get1, Get2, and Get2/Get1 were cloned into a pDDGFP vector to assess the targeting and the localization to the ER membrane, pJNG1GFP (Get1.GFP.8xHis), pJNG2GFP (Get2.GFP.8xHis), and pJNG21GFP (Get2/Get1.GFP.8xHis) were obtained (1). From pDDGFP the expression backbone pJANY-S1 (with strep tag) and pJANY-H1 (with 10xHis tag) were generated (2), they also carry the Ura3 selection marker. At this point the pJANY-S1 and pJANY-H1 were used as backbone to either insert Get1, Get2, and Get2/Get1 (3 and 4) or replace the Ura3 selection marker to His3 (S2 or H2 naming) in combination of the gene of interest to be expressed.

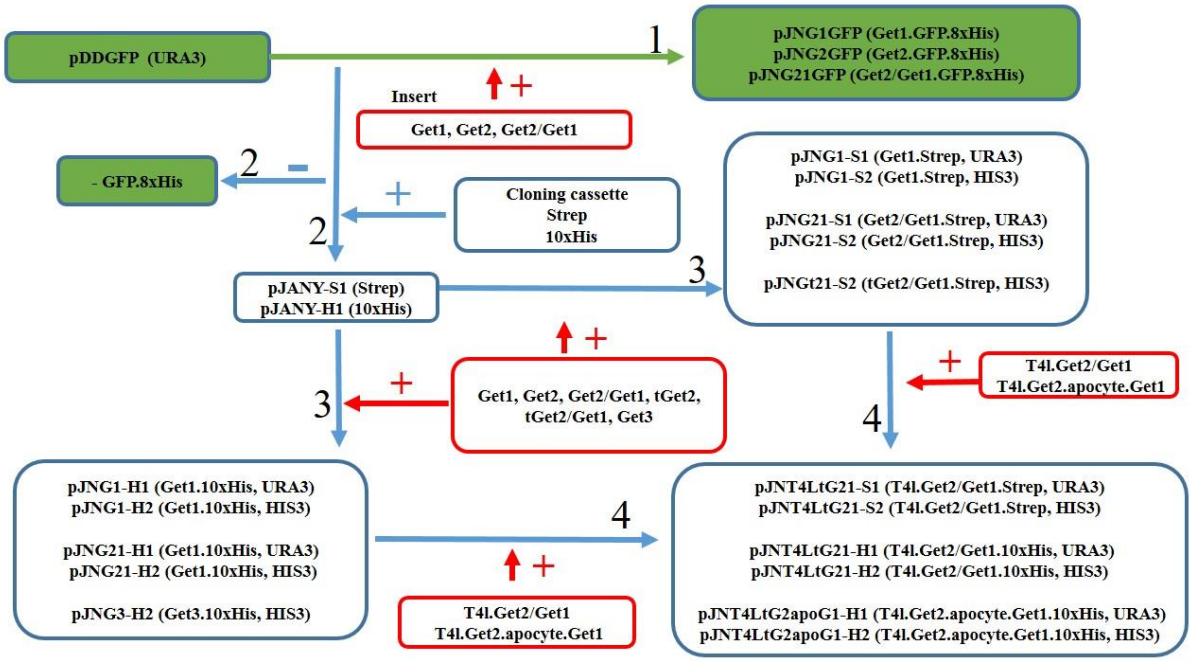
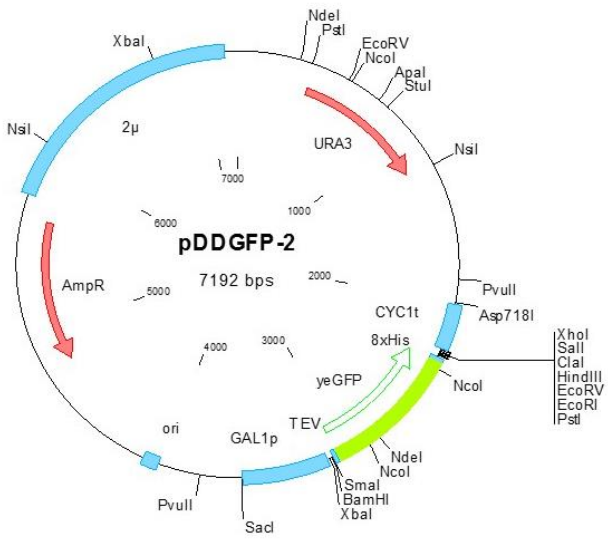
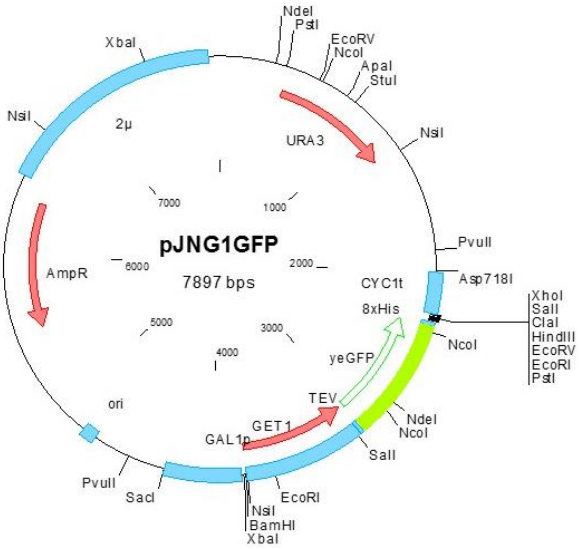
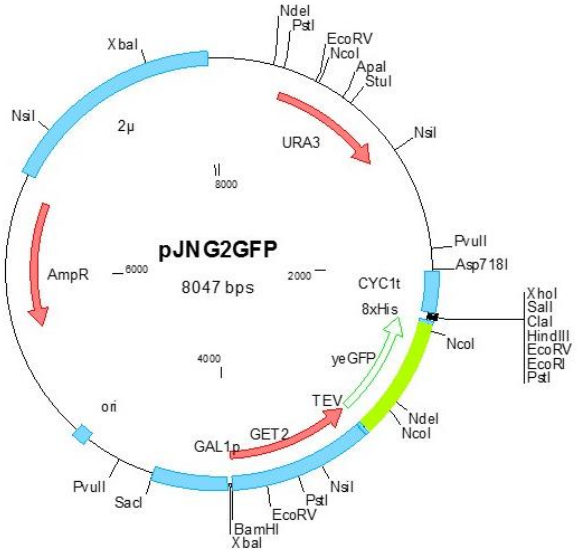


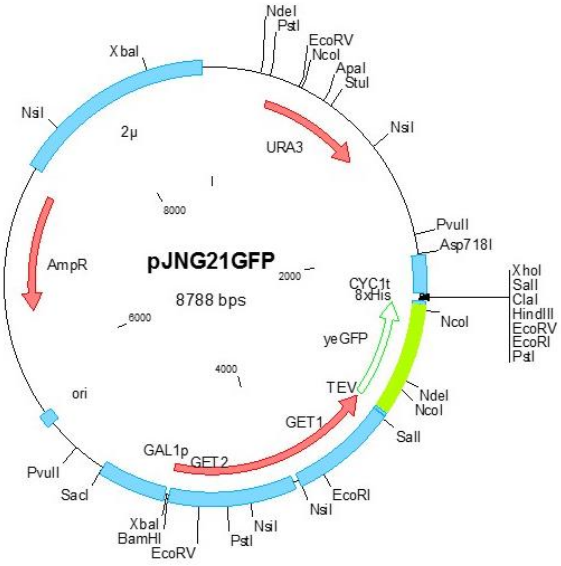
Figure 4.5.6: Overview of the main constructs used. 1. Construction of Get fused to GFP constructs in pDDGFP-2 by homologous recombination. 2. Construction of the expression backbone vectors without GFP, the vectors have first *URA3* marker (S1 and H1) and have either both strep-tag or 10xHis-tag. 3. Construction of expression vectors carrying either Get1, Get2, or Get2/Get1 together with *URA3* marker or *HIS3* marker (S2 or H2). 4. Insertion of the chimeric proteins T4l and thermostabilized apocytochrome b_{562}RIL into the expression vectors carrying the *HIS3* selection marker.

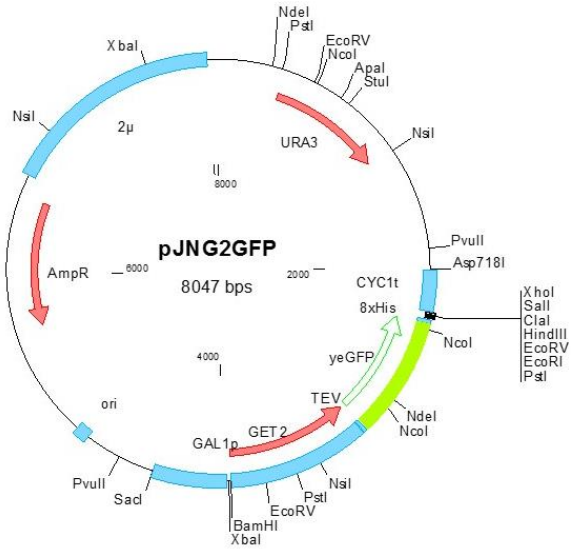
Plasmid maps generated

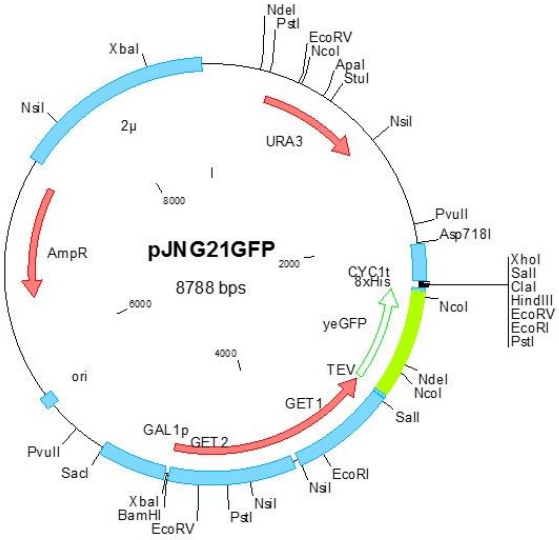


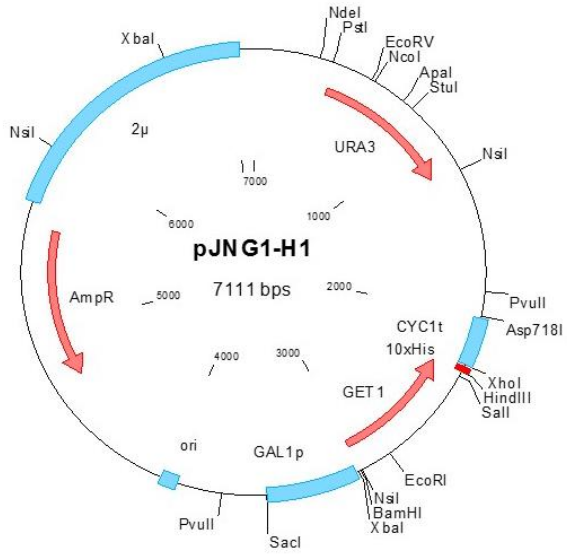


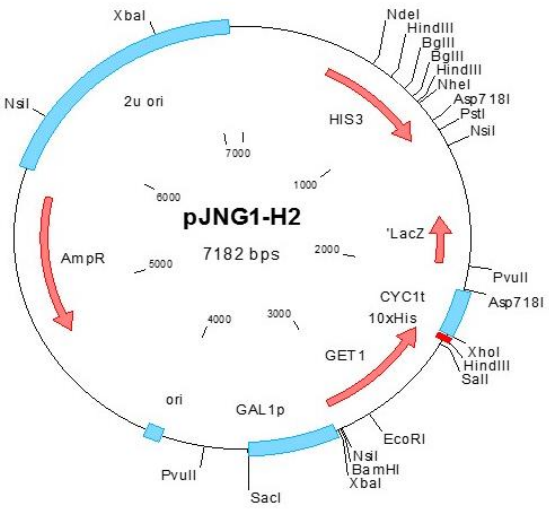


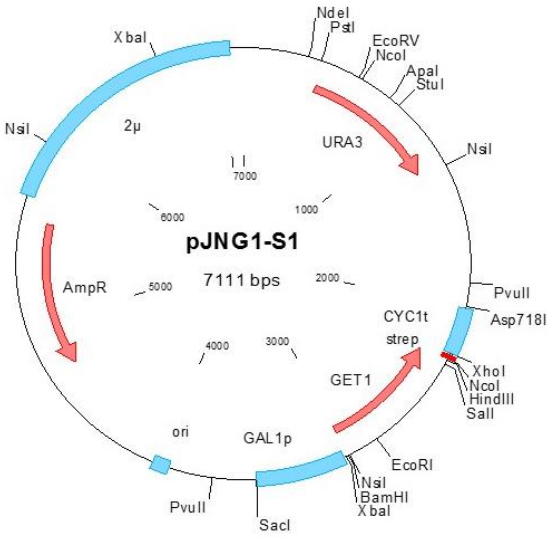


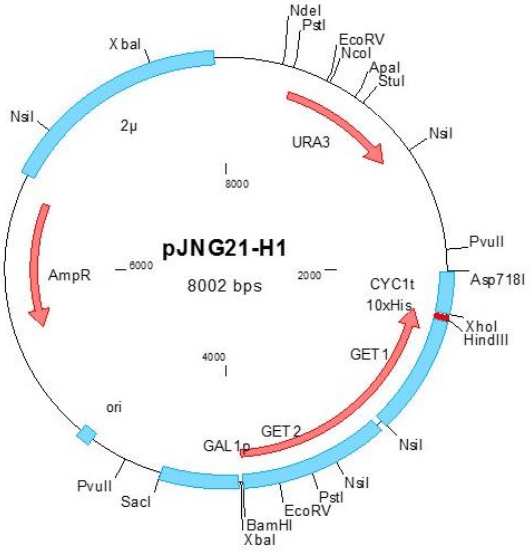

pJNG2GFP
8047 bps
2µ
URA3
AmpR
ori
GAL1p
GET2
TEV
yeGFP
8xHis
CYC1t
XbaI
NsiI
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
PvuII
Asp718I
XhoI
SalI
ClaI
HindIII
EcoRI
SacI
BamHI

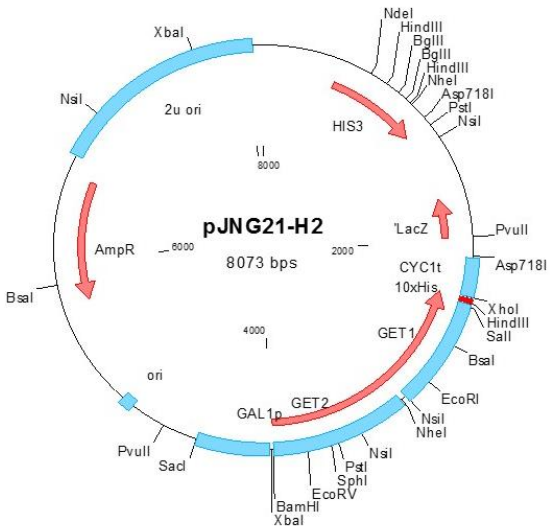

pJNG21GFP
8788 bps
2µ
URA3
AmpR
ori
GAL1p
GET2
TEV
yeGFP
8xHis
CYC1t
GET1
XbaI
NsiI
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
PvuII
Asp718I
XhoI
SalI
ClaI
HindIII
EcoRI
SacI
BamHI

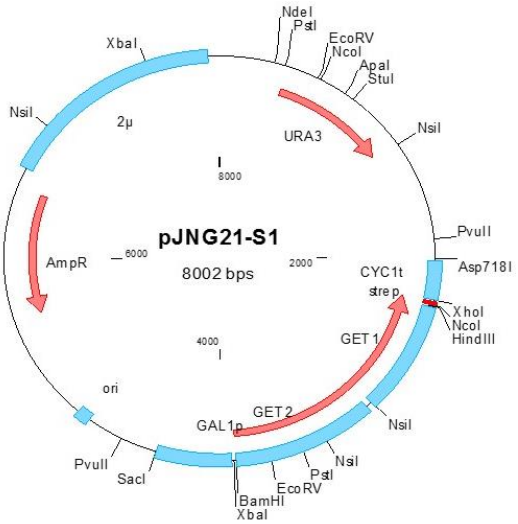

pJNG1-H1
7111 bps
2µ
URA3
AmpR
ori
GAL1p
GET1
10xHis
CYC1t
XbaI
NsiI
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
PvuII
Asp718I
XhoI
HindIII
SalI
EcoRI
SacI
BamHI

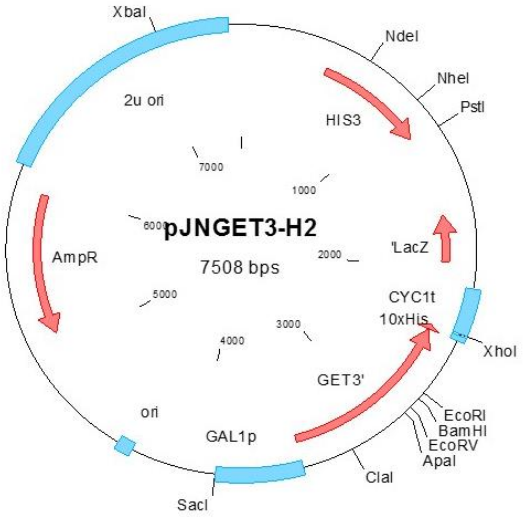

pJNG1-H2
7182 bps
2u ori
HIS3
AmpR
ori
GAL1p
GET1
10xHis
CYC1t
'LacZ
XbaI
NsiI
NdeI
HindIII
BglII
NheI
Asp718I
PstI
PvuII
XhoI
SalI
EcoRI
SacI
BamHI

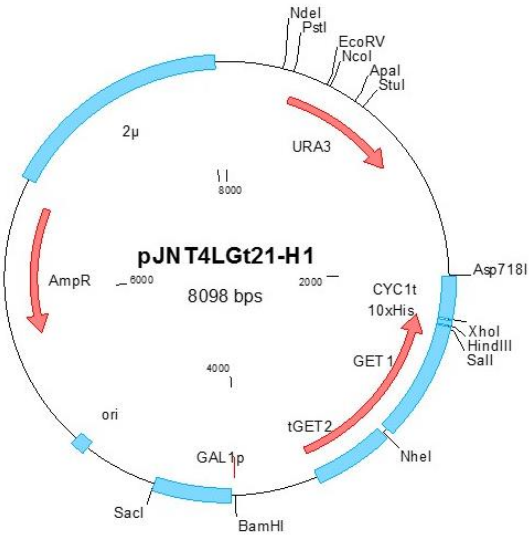

pJNG1-S1
7111 bps
2µ
URA3
AmpR
GET1
CYC1t
strep
GAL1p
ori
XbaI
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
NsiI
PvuII
Asp718I
XhoI
NcoI
HindIII
SalI
EcoRI
BamHI
SacI


pJNG21-H1
8002 bps
2µ
URA3
AmpR
GET1
CYC1t
10xHis
GET2
GAL1p
ori
XbaI
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
NsiI
PvuII
Asp718I
XhoI
HindIII
SacI
BamHI


pJNG21-H2
8073 bps
2u ori
HIS3
AmpR
'LacZ
CYC1t
10xHis
GET1
GET2
GAL1p
ori
BsaI
XbaI
NdeI
HindIII
BglII
NheI
Asp718I
PstI
NsiI
PvuII
SacI
SphI
EcoRV
BamHI
EcoRI


pJNG21-S1
8002 bps
2µ
URA3
AmpR
GET1
CYC1t
strep
GET2
GAL1p
ori
XbaI
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
NsiI
PvuII
Asp718I
XhoI
HindIII
SacI
BamHI


pJNGET3-H2
7508 bps
2u ori
HIS3
AmpR
'LacZ
CYC1t
10xHis
GET3'
GAL1p
ori
XbaI
NdeI
NheI
PstI
XhoI
EcoRI
BamHI
EcoRV
ApaI
ClaI
SacI


pJNT4LGt21-H1
8098 bps
2µ
URA3
AmpR
GET1
CYC1t
10xHis
tGET2
GAL1p
ori
NdeI
PstI
EcoRV
NcoI
ApaI
StuI
Asp718I
XhoI
HindIII
SalI
NheI
SacI
BamHI

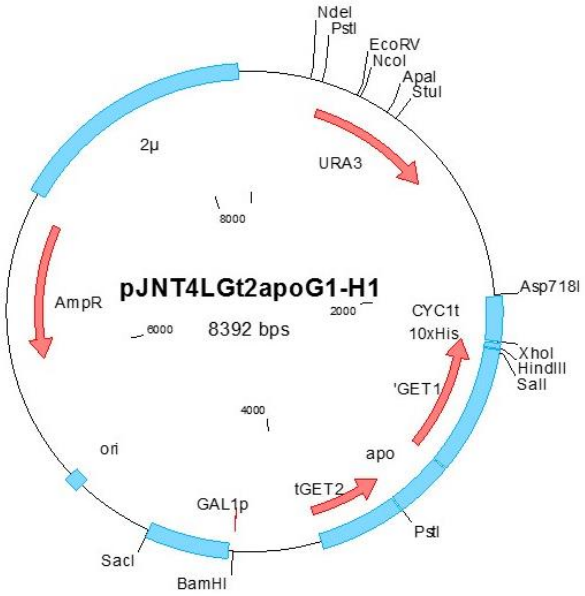


Figure 4.5.6.1: Plasmid maps

4.5.7 Comparison of the elution volumes of the single chain variants upon SEC

The comparison of the SEC chromatograms of all the single chain constructs (Figure 4.5.7), showed that the truncated_Get2/Get1 (tGet2/Get1) had the same elution volume peak of 15.5 ml as the full length Get2/Get1 which confirmed the degradation of the full length Get2/Get1 described in the chapters 4.5.1 and 4.5.2. The T4l.Get2/Get1 and T4l.tGet2.apocyte.Get1 elution volumes were normal (Table 4.5.7). This comparison of the SEC chromatogram led to the assumption that both the T4l.Get2/Get1 and T4l.tGet2.apocyte.Get1 variant were dimers in solution (two monomers each).

TABLE 4.5.7: THE ELUTION VOLUMES OF THE SINGLE CHAIN VARIANTS ON SEC

PROTEIN	Protein size (A.A)	Molecular weight (KDa)	Elution volume on superpose 6 10/300 GL (mL)
TGET2/GET1	419	48	15.5
GET2/GET1	548	61	15.5
T4L.GET2/GET1	580	67	15
T4L.TGET2.APOCYTE.GET1	674	77	14.5

4.6 Electron microscopy

Since the new generation of sensitive direct electron detectors became available, electron microscopy (EM) was used to generate structural models of membrane protein complexes (Kühlbrandt, 2014). EM is now considered as the best alternative tool to collect structural data of membrane protein complexes with sizes above 100 KDa through 3-D reconstruction. In the case of Get2/Get1, the single chain construct T4l.tGet2.apocyte.Get1 was selected because it

was much cleaner with symmetric SEC chromatogram. It was also assumed that T4l.tGet2.apocyte.Get1 was also a dimer based on the comparison elution volumes of the Get2/Get1 variants from gel filtration (Table 4.5.7) and on stoichiometry studies of Get2/Get1 and tGet2/Get1 (chapters 4.5.2 and 4.5.3). The dimer of T4l.tGet2.apocyte.Get1 (two monomer of single chain construct) has a size of 154 KDa.

There are three main EM methods used to study biological samples: The negative staining, the 2D electron diffraction and single particle Cryo-EM.

For this particular study, only the negative staining and Cryo-EM techniques were used. Negative staining is generally used first to obtain the overall shape because of high contrast of the target specimen. 2 % uranyl acetate was used to stain the Get2/Get1 variants. Once the negative staining was done in another preparation, the samples were vitrified by rapid freezing to help to preserve the structural integrity of T4l.tGet2.apocyte.Get1.

4.6.1 Negative staining

Before Cryo-EM, the overall shape and surface of T4l.tGet2.apocyte.Get1 was analysed by negative staining single particle analysis (Figure 4.6.1) and a random conical tilt (RCT) of stained images (Figure 4.6.3).

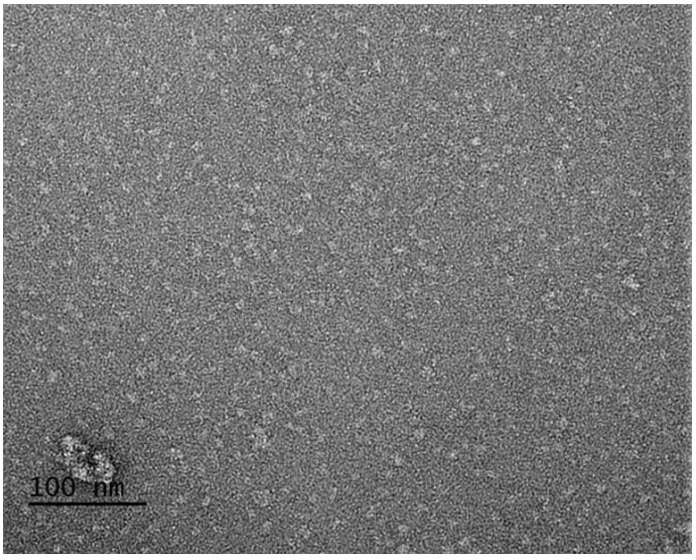


Figure 4.6.1: Negative staining of T4l.tGet2.apocyte.Get1. The sample specimen were stained with 2 % uranyl acetate.

After the images of negative staining samples were taken, particles were picked and the 2D classification was generated (Figure 4.6.2). Hand tilted images by 45° were also prepared (Figure 4.6.3). A low resolution starting model from untilted data (Figure 4.6.1 and Figure 4.6.2) and from the random conical tilted data was generated (Figure 4.6.4 and Figure 4.6.5).

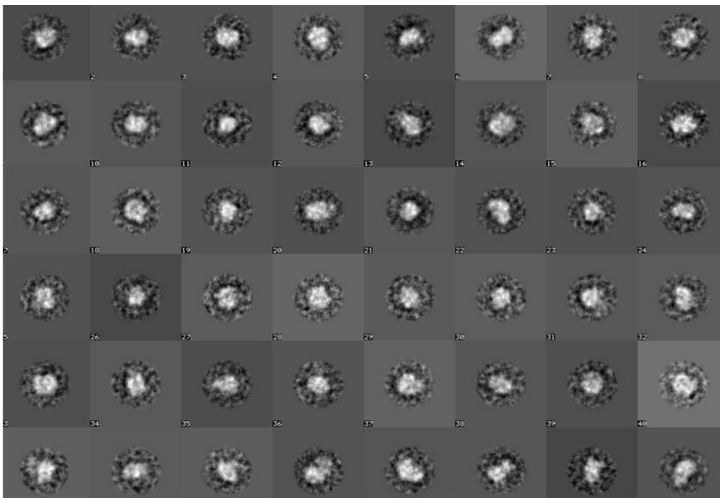


Figure 4.6.2. 2D Class average negative staining T4l.tGet2.apocyte.Get1.

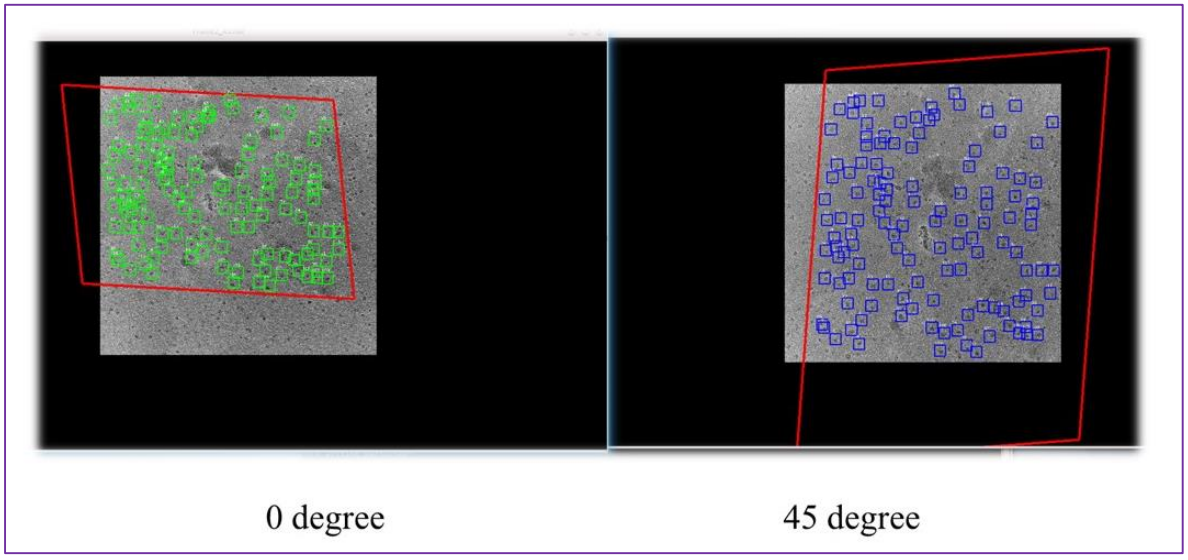


Figure 4.6.3. Example of tilted images of T4L.Get2.apcyte.Get1 for model validation.

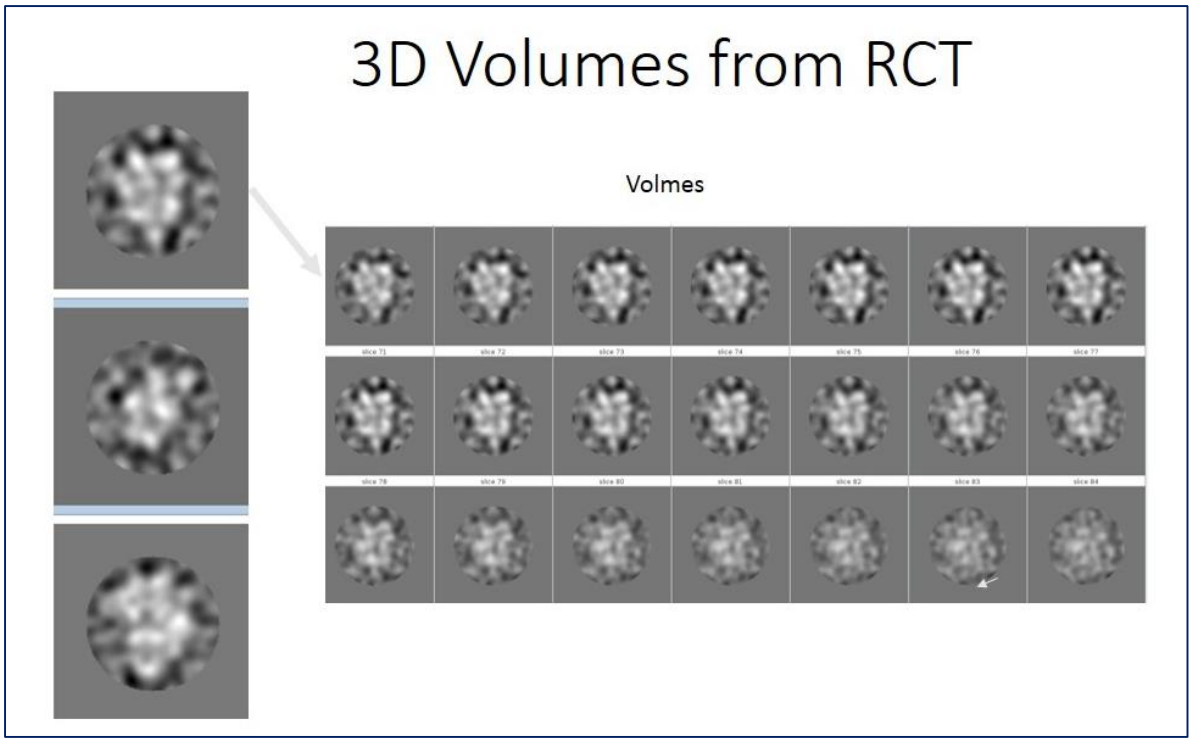


Figure 4.6.4. Volumes generated from the RCT.

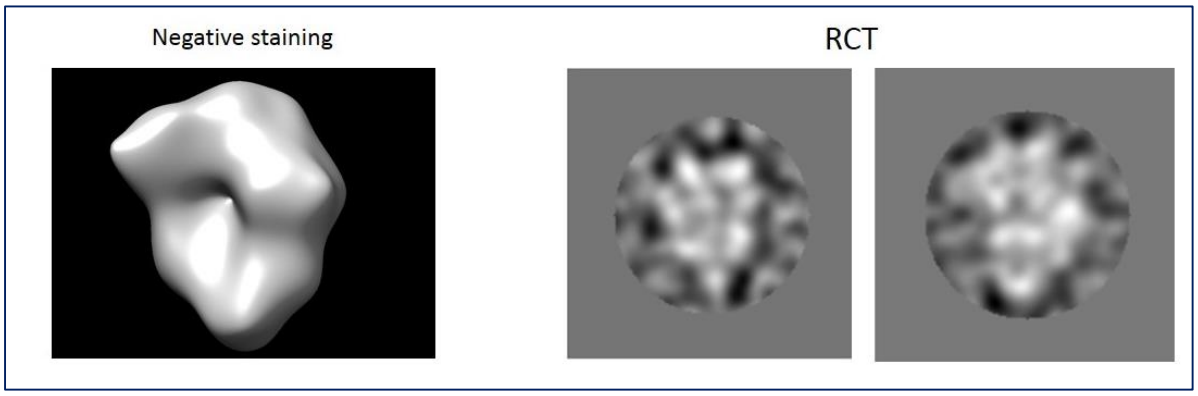


Figure 4.6.5. Starting model negative. Left the starting model from negative staining compare to two RCT volumes to illustrate the shape similarity.

Once it was sure that the sample preparation was good and that the particles of T4l.tGet2.apocyte.Get1 could be picked (Figure 4.6.5), a new sample preparation was done following the same protocol, and the T4l.tGet2.apocyte.Get1 sample specimen was vitrified for Cryo-EM analysis.

4.6.2 Cryo-EM analysis of T4l.tGet2.apocyte.Get1

The size of the complex is at the Cryo-EM detection limit. Consequently, T4l.tGet2.apocyte.Get1 samples were prepared to be vitrify. The first sample to be analyzed was transferred to an amphipol environment. The particles were visible (Figure 3.6.2.1). The CTF was then calculated. The particles were picked and classified into 2D Classes (Figure 3.6.2.2).

It was observed that there were an average of 25 particles per frame which would have required a lot of resources to process the data. Instead T4l.tGet2.apocyte.Get1 purified in DDM was used (Figure 3.6.2.3). The DDM sample was then vitrified and analyzed by Cryo-EM. The preparation in DDM was better, clean and homogenous **(**Figure 4.6.2.3 A). The thon rings of the power spectrum did not show any asygmatism, hence they were perfect circles **(**Figure 4.6.2.3 A).

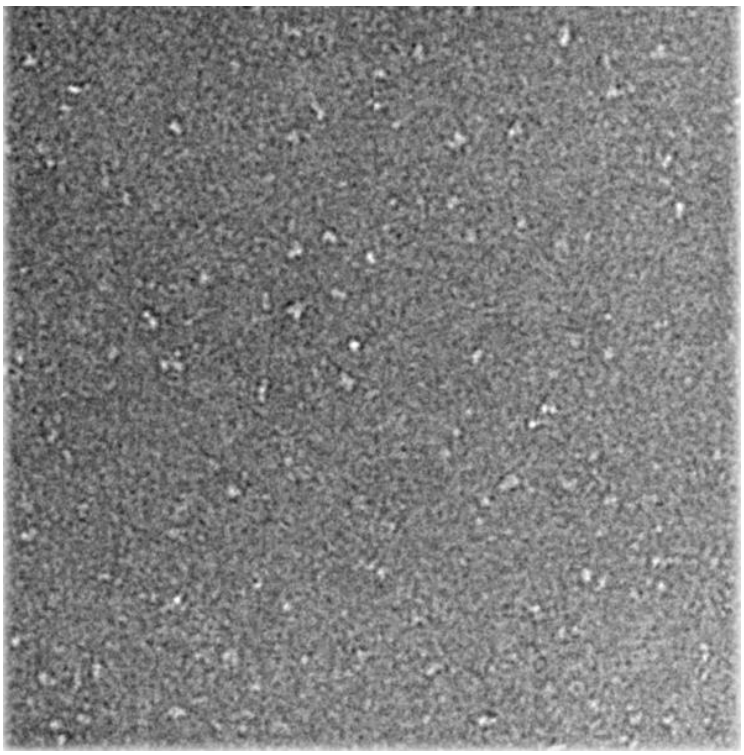


Figure 4.6.2.1: T4l.Get.apocyte.Get1 in amphipol.

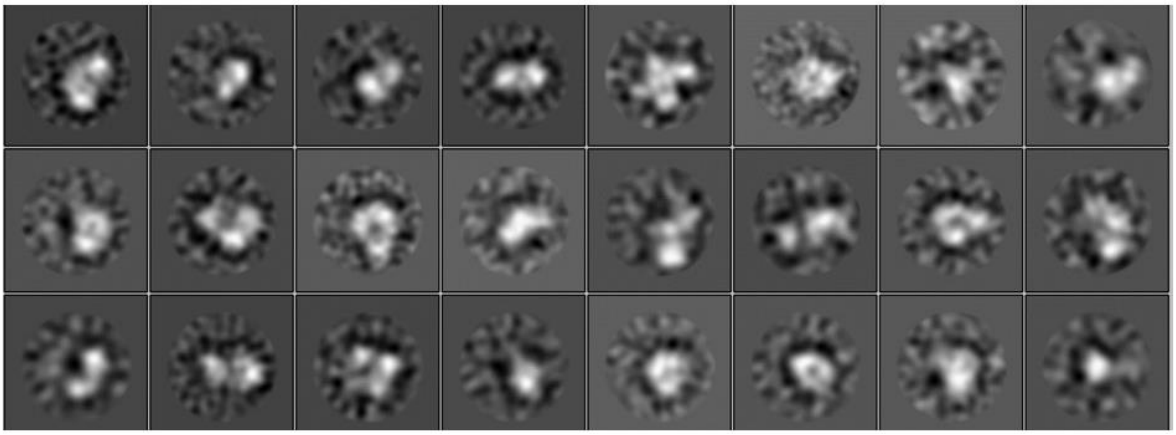


Figure 4.6.2.2: 2D class average of T4l.Get.apocyte.Get1 in amphipol.

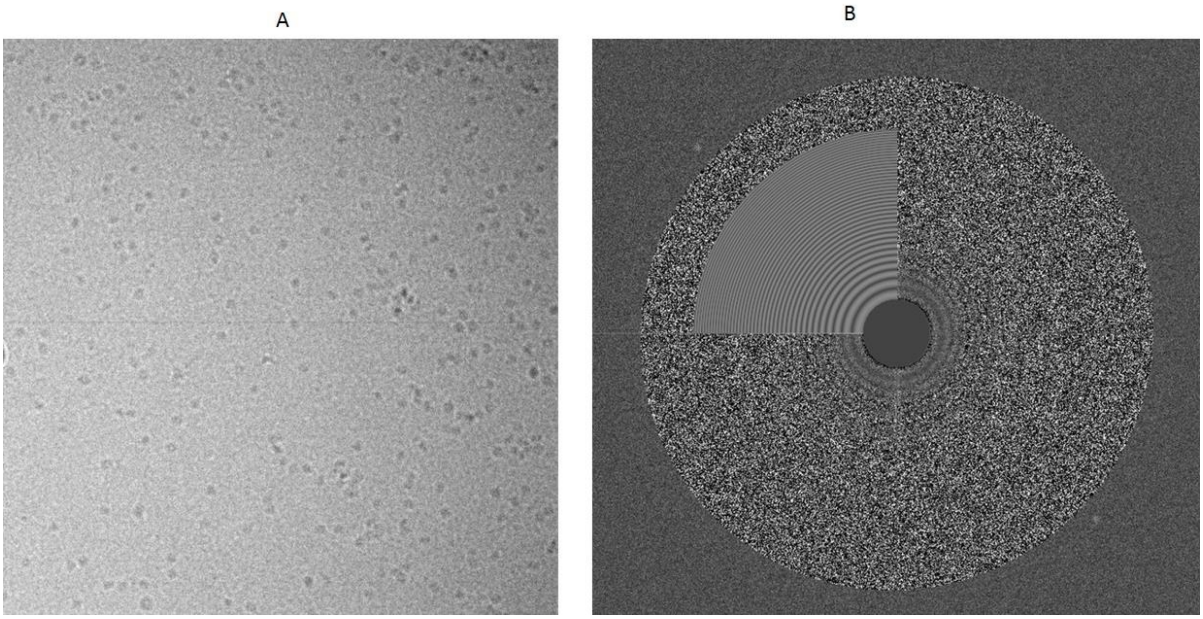


Figure 4.6.2.3: T4l.Get.apocyte.Get1 in DDM. A. micrographs with its Fourier transform, the power spectrum, shows thon rings pattern.

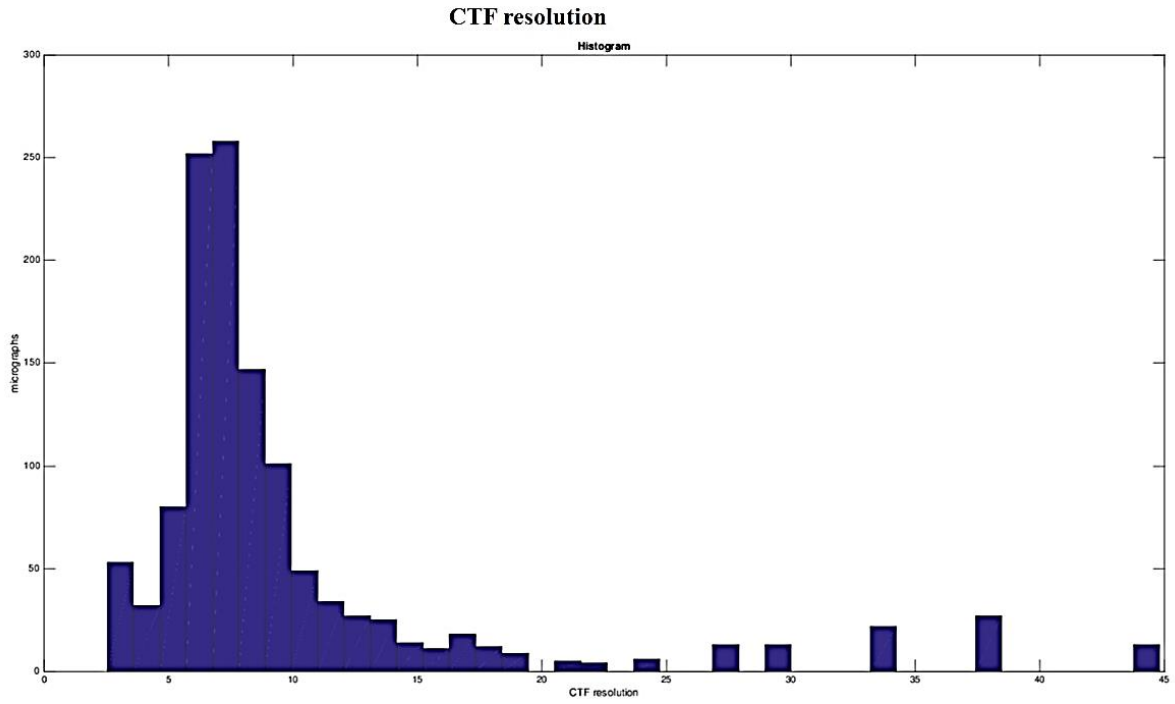


Figure 4.6.2.4: CTF resolution evaluation shows that 50% of the micrographs has the resolution between 6-7 Å.

914 frames were taken and the CTF of each frame was determined. The Figure 3.6.2.4 shows the trend of the CTF resolution estimation of all micrographs (Figure 3.6.2.4). For over 50% of the frames the CTF resolution was found between 6 and 7 Å. Thereafter, the particles were picked and classified into 2D class averages (Figure 3.6.2.5) and the initial model was generated (Figure 4.6.2.7). The heterogeneity of particles was further analyze during the 3D classification. The results indeed showed heterogeneity in the population (Figure 4.6.2.7).

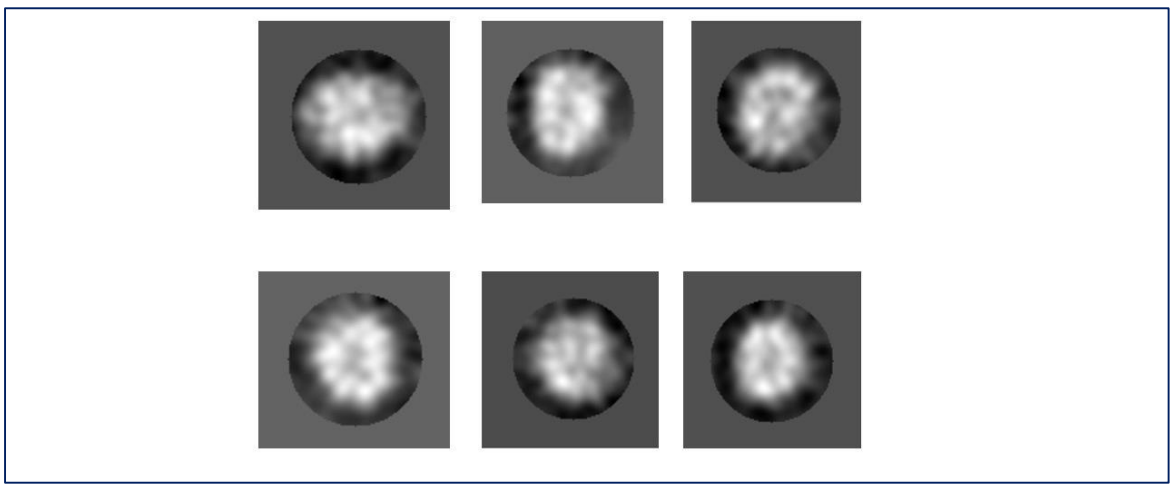


Figure 4.6.2.5: 2D class average of T4l.Get.apocyte.Get1 in DDM.

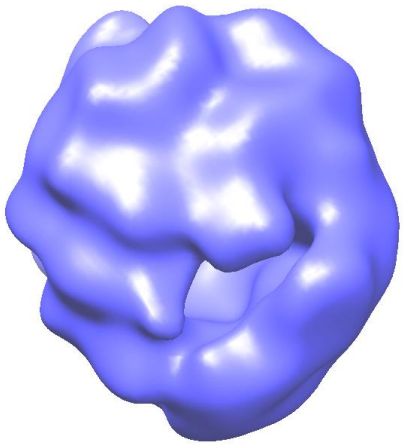


Figure 4.6.2.6: Starting model from Cryo-EM of T4l.Get.apocyte.Get1.

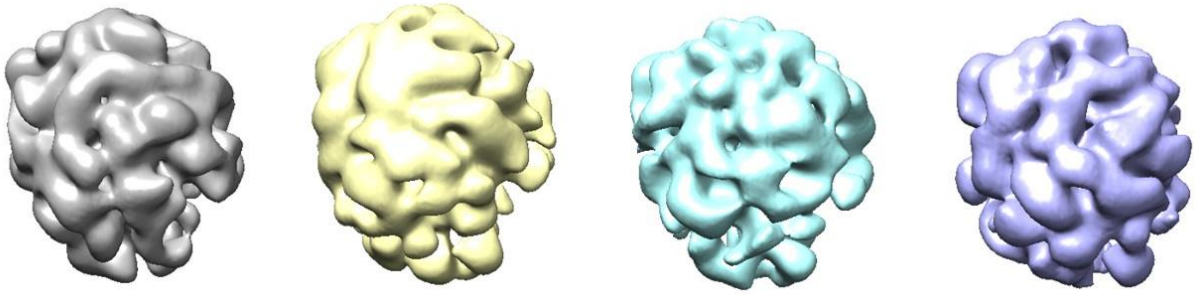


Figure 4.6.2.7: 3D classes of T4l.tGet2.apocyte.Get1.

The grey map was refined to a resolution of 10 Å and gave a final angular accuracy of 5.1 degree. Closer look at map 3 by fitting the structures of T4l (18 KDa) and apocytochrome b_{562}RIL (12 KDa) suggested that the density observed is the density of monomeric single chain complex T4l.tGet2.apocyte.Get1 (Figure 4.14.2.8).

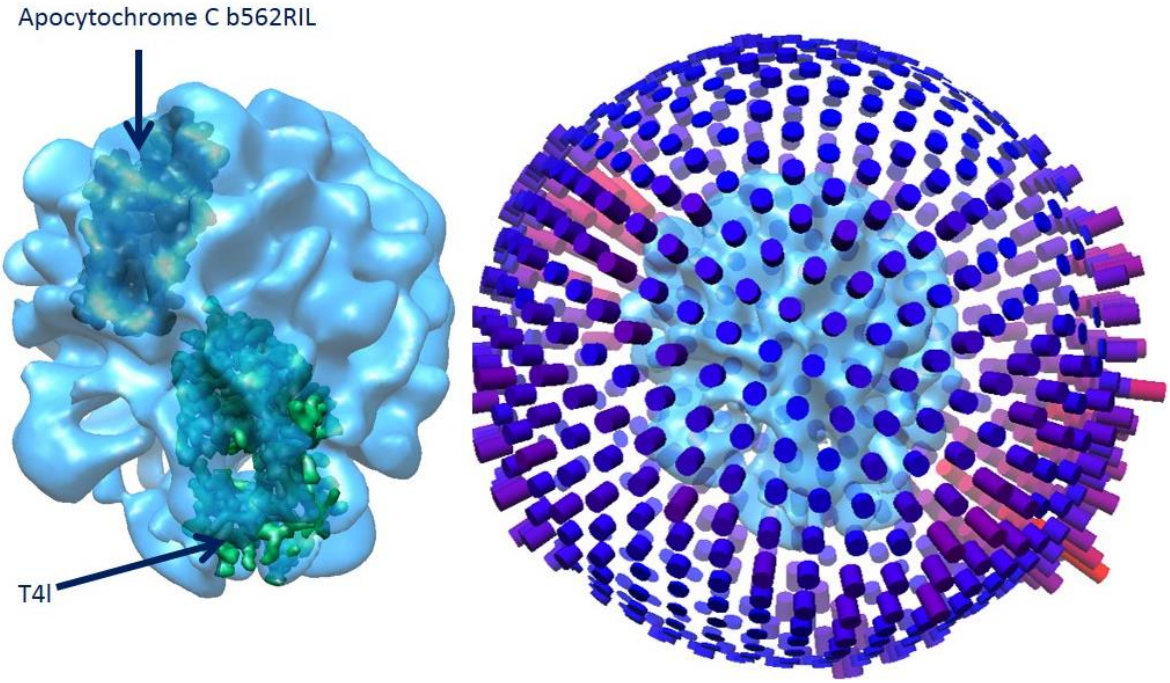


Figure 4.6.2.8: Fitting maps of T4l.tGet2.apocyte.Get1, T4l structure and apocytochrome b562RIL structure. Left figure shows the fitting T4l structure (18 KDa) and apocytochrome b562RIL structure (12 KDa) to the density map of T4l.tGet2.apocyte.Get1 and the cryo-EM density map with the orientation distribution

Chapter E: Discussion

5.1. *Saccharomyces cerevisiae* as host for eukaryotic membrane proteins expression

In the post-genomics science era the production of large quantities of eukaryotic membrane proteins remains still a challenge. The lack of sufficient quantities of stable and functional membrane proteins is a limiting factor for the biochemical and structural characterization of many membrane proteins. This is why, despite the importance of membrane proteins as important drug targets for pharmaceutical industry, structural information of membrane proteins is lacking behind the soluble proteins counterpart.

In an idealistic world, it is preferable to overexpress membrane proteins or any protein from their native hosts. So, during this thesis, the yeast Get3 receptors, namely Get1, Get2 and all single chain Get2/Get1 variants, were expressed in their native yeast host, *S. cerevisiae.* Also, the yeast GFP system developed by David Drew and colleagues was used as basis to monitor and optimize the expression, the localization to the ER membrane, and the purification of Get3 receptors.

The molecular biology behind the yeast expression vector is very attractive using *S. cerevisiae* as host for eukaryotic membrane proteins expression (Drew et al., 2008).

The conditional and the regulatory GAL1 promoter system was used to control the expression of Get3 receptors. Glucose is the preferred energy carbon source in *S. cerevisiae*. When it is present, galactose and other carbon sources for metabolism are repressed. In absence of glucose, the addition of galactose in the system can induce the overexpression of membrane proteins under the control of the strong GAL1 promoter.

Moreover, the gene cloning into standard vectors by homologous recombination or gap repair is done without using restriction enzymes, therefore ligating restricted DNA in a tube. The cloned and isolated yeast plasmid can be propagated both in yeast as well as bacteria. In addition, the system offers a wide choice of vector copy number and selection markers for yeast (chapter 3.2.1).

The GFP fusion approach did indeed help to ease the optimization of the expression and the solubilization of the corresponding Get3 receptors (see chapter 4.2). Upon the expression of Get1, Get2, and Get2/Get1 fused to GFP, the quality of the expressed proteins were efficiently evaluated and screened using the fast screening advantages of the GFP based technology.

As a first step, the overexpression of Get1, Get2, and Get2/Get1, was monitored by measuring the fluorescence counts at specific time points of the cultivation and the concentration in mg per ml was accurately estimated from 10 ml culture. For instance, in Table 4.2.1one set of experiments is presented showing the fluorescence counts after the cultivation of Get3 receptors. One of the main advantages observed using the GFP tool was that it saved resources and time during the screening procedure without the need of doing an overexpression. The information provided by the expression test from the 10 ml culture did enable to easily expand from 10 ml to 1 liter cultures. However, it is important to notice that the yield from the small-scale expression is not always proportional by just upscaling the volume of the culture. Many parameters such as the aeration in the 2 500 ml culture flask compared to 10 ml falcon tube can alter the expected yield.

Nevertheless, the main information of this preliminary expression test was that Get1, Get2, and Get2/Get1 can be expressed in the native host in sufficient quantity for structural studies. Also, the single chain constructs were better expressed than Get1 and Get2. This finding was an initial hint that Get1 and Get2 stabilize each other within the ER membrane. Plus Get2/Get1 overexpression as single chain would be also stable, tolerated and less toxic for the cells.

5.2 Chemical chaperones improve the product of the expression of Get3 receptors

Once the expression was confirmed, the best clones were selected by their highest fluorescence intensity (in green in Table 4.2.1) and used to test if the expression can be improved by addition of chemical additives into the culture media during the protein expression.

The concept of using chemical additives, also called chemical chaperones, originated in the early 2000. Chemical chaperones were used to increase the stability of native proteins *in E. coli* upon their addition during the induction of protein expression (Diamant *et al.*, 2001). The aim was to chemically stress the cells just before the expression of the recombinant proteins. The stressed cells would have then likely increased the production of molecular chaperones. The produced molecular chaperones facilitate the correct folding of the recombinant proteins by preventing their aggregation, and disassembling protein aggregates. The technique was applied in the field of membrane protein expression, and it did help to increase the yield of functional mammalian G protein coupled receptors (GPCRs) several fold over standard expression conditions (André *et al.*, 2006). In this study, when chemical additives were used for Get3 receptors, changes were observed in the yield of Get1, Get2, and Get2/Get1. For

instance when Get1 expressing cells were treated with DMSO upon induction, an increase of about 2000 a.u fluorescence counts was observed which is equivalent of 0.6 mg/ml and represents 3 mg gain in a liter culture (Figure 4.2.3 a). This is a significant increase for membrane proteins. By analogy Get2 had a gain of 1.8 mg per liter culture with histidine (Figure 4.2.3 b); However there was no gain for the single-chain construct (Figure 4.2.4.3 c), which can be explained by the fact Get2/Get1 was already well expressed and well folded on its own confirming that Get1 and Get2 stabilize each other.

5.3 Functional expressed Get2/Get1 is targeted and inserted into the ER membrane

The next question was if the expressed Get1 and Get2/Get1 fused to GFP were properly folded and inserted into the membrane. Since the Get proteins are cloned upstream of GFP, GFP is often used as indicator to monitor membrane protein insertion versus inclusion body formation, as GFP in inclusion bodies will not fluoresce (Drew *et al*, 2001). Hence the fluorescence intensities observed (Figure 4.2.2 a and Figure 4.2.2. c) during the expression of the recombinants fusion proteins Get1.GFP and Get2/Get1.GFP suggested that they were correctly folded. Get1 and Get2/Get were also targeted to the membrane. In order to confirm the quality of the Get3 receptors and also that they were not targeted to any membrane but within the ER membrane organelle, the confocal microscopy was used to assess Get1 and Get2/Get1 localizations.

The localization of Get2 was not analyzed because the aim of the study was to reconstitute the interaction of Get1 and Get3 to form a stable complex Get2/Get1/Get3.

Get1 is found fluorescent around an organelle that has a shape of spherical vesicles which leaves room for interpretation as those could be seen as part of the Golgi body. However when the topology and the secondary structure of Get1.GFP were analyzed (Figure 4.2.2 b), it was recognized that Get1 exhibits two yeast peroxisome targeting signals SKF and one YKL (Van Ael and Fransen, 2006). They are present near the TMD entry in the cytoplasm or lumen (Figure 4.2.2 b) but are not the end of the sequence. Then by looking more closely the confocal images of Get1.GFP, it is noticeable that the GFP signal was not localized in the cytosol nor in the plasma membrane. The confocal images showed that Get1.GFP was targeted to an organelle with a membrane that was not a plasma membrane or ER membrane. From the appearance it seems that it was, peroxisome (microbodies), vesicle, vacuole or even nucleus like shape. But based on the two yeast peroxisomal targeting signals SKF and one YKL the odds are that Get1 alone is targeted to the peroxisome.

The localization of Get2/Get1 fused to GFP shows through the fluorescence intensity observed that Get2/Get1.GFP has a tube-like morphology or membrane-enclosed sacs (cisterna), implying that Get2/Get1 is actually targeted to the ER membrane and is likely functional (Figure 4.2.2 c).

5.4 DDM is the most suitable detergent to solubilize and purify Get3 receptors

The next step was to screen the best detergent for solubilization. A batch from clones used for Get3 receptors was also used for the detergents screening experiment. The analysis was limited to Get1.GFP and the results were simply extrapolated to Get2/Get1.GFP. The monodispersity during the analytical gel filtration was the indicator for stability in a particular detergent.

Foscholine 12, LDAO, DDM, and DM detergents were used and screened the monodispersity by analytical size-exclusion chromatography. DDM was found to be the best detergent because the analytical size exclusion of Get1.GFP solubilized and purified in DDM shows a symmetric peak at 1.22 ml elution (Figure 4.2.4). One liter large scale preparation test confirmed the results from analytical analysis. The proteins were successfully solubilized and purified in DDM by Ni-NTA affinity chromatography (Figures 4.2.5 A and 4.2.5 B).

Despite this outcome, the membrane proteins fused to GFP showed that part of the C-terminus containing the GFP was cleaved off (Figure 4.2.5 C). Get3 receptors showed they were instable in micelles when fused to GFP and caused degradation of the proteins.

Therefore, as the protocol was established, it was useful to substitute the GFP.8xHis-tag with a simple strep-tag or 10xHis-tag. Two new expression vectors carrying a *URA3* selection marker were then constructed, namely pJANY-S1 (Strep) and pJANY-H1 (10xHis) (Figure 4.3 and Figure 4.5.6).

5.5 Get1 behaves well at physiological pH 7.6

Following the procedure that was developed using GFP fusion, Get1 was cloned into pJANY-S1 and pJANY-H1 vectors.

Get1 was successfully expressed, solubilized and purified in DDM. The same results were achieved with either strep-tag or 10xHis-tag constructs. Get1 purified at pH 7.6 (Figure 4.4.1.2 A) showed on gel filtration a symmetric monodisperse peak.

The main focus in the study was to get an insight of the molecular interactions of the Get complex in the ER membrane, and for this the reconstitution of the interaction of Get3 with

Get1 was crucial. This interaction between Get1 and Get3 was confirmed to be direct (Figures 4.4.1.4.1, 4.4.1.4.2, and 4.4.1.4.2) and therefore the purified Get1 was likely functional.

5.6 Get2 and Get1 is predominantly a dimer in solution

Get2 and Get1 were linked by ASGAGGSEGGGSEGGTSGAT and are functional (Wang *et al.*, 2014). It was necessary to show the interaction *in vitro* with proteins before further proceeding; the first single chain in hand was used to test the interaction of the single chain construct Get2/Get1 with Get3 (Figure 4.4.2). Thereafter, Get2/Get1 strep-tag was purified under the same condition used for Get1 strep-tag.

The purification of Get2/Get1 at pH 7.6 was not monodisperse at first from SEC (Figure 4.5.1.1). The Get3 receptors have all basic pI (Figure 5.6).

TABLE 5.6: THEORETICAL COMPUTED OF GET3 RECEPTORS PI

PROTEIN	Protein size plus 10xHis (A.A)	Molecular weight (KDa)	pI
GET2	290	33	9.52
GET1	247	28	9.25
TGET2/GET1	419	48	9.43
GET2/GET1	548	61	9.40

To optimize the purification of Get2/Get1, the pI of Get2/Get1 was considered for its purification. That leads to the purification of Get2/Get1 in Na_2CO_2/$NaHCO_2$ buffer at pH 10. At pH 10 Get2/Get1 showed a symmetric monodisperse peak (Figure 4.5.1.2) but the SDS-PAGE and western blot showed some degradation of the single chain and it was confirmed by mass spectrometry. The large cytosolic flexible N-terminal domain of Get2 was considered to be potentially problematic at this point.

Get2/Get1 was subject to stoichiometry analysis by LILBID. Two signals of monomeric Get2/Get1 (50 KDa and 61 KDa) were observed which consequently leads to a corresponding heterogeneous signal of the dimer signal. Nonetheless, it was the first evidence that Get1 and Get1 form heterotetramers (two monomers of Get2/Get1).

As the results were not clear and it looked as the N-terminal cytosolic part of Get2 was degraded during overexpression, the single chain construct was stabilized by removing the residues from 1 to 128 from the N-terminal part of Get2 (Figure 4.5.3.1 A). The expression and purification of the truncated tGet2/Get1 (Figure 4.5.3.2 A) showed fewer degradation

products. The main fraction of the elution peak was used to analyze again the stoichiometry by LILBID. The analysis of the new results showed that the tGet2/Get1 was cleaner and more homogenous.

There was no further ambiguity about the stoichiometry of Get2/Get1 (Figure 4.5.3.3). LILBID results have confirmed that the single chain Get2/Get1 forms a dimer (heterotetramer of Get2/Get1; two Get2, and two Get1).

Furthermore, the results showed that the degradation of Get2/Get1 (from 61 KDa to 50 KDa) was happening within the unstructured cytosolic N-terminus of Get2. This result did confirm that presence of the N-terminal part of Get2 is inadequate for structural studies because the degradation of Get2/Get1 is prone to heterogeneity.

Based on the LILBID results Get2 and Get1 form a heterotetramer (2xGet1 and 2xGet2) in solution.

5.7 T4 lysozyme and thermostabilized apocytochrome b_{562}RIL chimeric proteins help to stabilize the single chain construct Get2/Get1

To improve the stability for structural studies, the unstructured N-terminal part was replaced with T4 lysozyme (Figure 4.5.3.1 B). T4l variants were used in other studies to optimize the crystallization and promote alternative packing interactions. They were tested first on M3 muscarinic receptor (M3) crystallization, and the structure of T4L.M3 crystals was solved at resolution of 2.8 Å (Thorsen *et al.,* 2014). The N-terminal modifications of the flexible N-terminus of Get2 appeared to have improved the stability of the single chain construct (Figure 4.5.3.2).

Before proceeding with more analysis the interaction between T4l.Get2/Get1 with Get3 was evaluated (Figure 4.5.4.1). Thereafter, the crystallization of both the tGet2/Get1 and T4l.Get2/Get1 was investigated (Figure 4.5.3.3). However, none of the crystals diffracted.

Afterwards, the next idea was to find a way of thermostabilizing the single chain constructs. For this goal the GS linker sequence between Get2 and Get1 ASGAGGSEGGGSEGGTSGAT (Wang *et al.,* 2014) was substituted with the thermostabilized apocytochrome b_{562}RIL (Figure 4.5.5) sequence that was shown to have advantages over T4 lysozyme when applied to GPCRs (Chun *et al.,* 2012). The new optimum single chain construct has both chimeric proteins T4 lysozyme and thermostabilized apocytochrome b_{562}RIL. The pure T4l.tGet2.apocyte.Get1 showed better behavior on SEC as well as on SDS-PAGE (Figure 4.5.5.1). This result implies that T4l.tGet2.apocyte.Get1 is more stable and the optimum

construction. The crystals of T4l.tGet2.apocyte.Get1 were larger but they did not diffract either.

5.8 Get3 receptors in solution are heterogeneous

To study the interaction of the complex of Get1/Get2/Get3, the approach was to study the affinity between Get1 and Get3. No such studies were so far done with the full membrane protein Get1. Both proteins were purified separately and the interaction between Get1 and Get3 was investigated by MST (Figure 4.4.1.4.2 and 4.4.1.4.3). MST is a very sensitive technique, MST measures the dissociation constants and can unveil different binding mechanism by distinct thermophoresis signals. The great asset of MST is that it analyzes the molecular interactions in solution, a surface immobilization like in the case of surface plasmon resonance (SPR) is not needed. The level of homogeneity in the number of binding site can also be revealed with precision by MST. The result not only showed a strong binding of Get3 to Get1, but also a biphasic behavior of the binding with two different affinities (K_D of 57 nM and the second of 740 nM). The second binding event of 740 mM is likely caused by heterogeneous mixture that can be either carry over contaminations or different oligomerization state. Previous studies on the small unstructured N-terminus of Get1 had revealed a binding affinity of 17 nM with Get3 in the apo state and 31 nM in the ADP bound state. Their affinity dramatically decreased in the presence of ATP (Stefer *et al.*, 2011).

Additionally, when any Get2/Get1 variant (Figure 4.5.4.1) is bound to Get3 by tandem affinity purification, one additional protein band co-elutes around 30 KDa yet this band was identified as a Get2/Get1 degradation product. When the Get1/Get2/Get3 complexes were analyzed by LILBID (Figure 7.9, appendix) and GraFix (Figure 7.6.1-7.6.5 appendix), despite reconstituting the complex from pure proteins, it was observed that the heterogeneity was recurrent which most of the time led to the destabilization of Get3 for unknown reasons. This suggested that the integrity of Get3 receptors could be affected when kept in micelles.

Heterogeneity in micelles could also be an explanation why none the crystals diffracted.

5.9 Cryo-EM of Get1 and Get2 complex

Crystallization is often considered as the gold standard studies to have deep understanding of molecular mechanisms at atomic level in biology, hence all purified Get3 receptors were crystalized (Figure 4.4.1.3 and Figure 4.5.4.2). Most crystals grow after two days between 5 to 30 μm. Neither microseeding nor lipid cubic phase could improve their quality. One could argue that those crystals were salt crystals but nothing pointed in this direction because those

crystals were big enough for home source salt crystal testing. The salt crystal testing showed that the crystals were not salt. The other point is that the crystals decayed as soon as they reached their maximum size; they usually went from good looking crystals to irregular shape. The majority of crystals observed grew in presence of calcium. Also Get2 is a Calcium-Modulating Cyclophilin Ligand (CamL) mammalian homologs. CamL, together with the tryptophan-rich basic protein (WRP), are TRC40 receptors. They are necessary to facilitate the targeting and the insertion of TA proteins into the ER membrane of mammalian cells (Fabio Vilardi *et al.*, 2014; Colombo). It now seems that calcium might play a part during TA protein targeting and insertion in yeast as in mammalian cells. The heterogeneity can be caused by the existence of variable conformations of Get3 receptors within the ER membrane, all conformations are necessary for simultaneously inserting the TA proteins to the ER membrane as well as the removal of bound Get3 from the membrane. It might be interesting to see the effect of calcium at the molecular level, and how it affects the insertion or the removal of Get3 from the ER membrane.

As it was not possible to collect any diffraction data from any crystals, the heterogeneity of the Get1 and the single chain Get2/Get1 is the likely reason for which the crystals did not diffract.

The purification of T4l.Get2/Get1 showed a monodisperse peak. Also, because Get2/Get1 is heterotetramer, it was thought that the heterogeneous ensemble can be separated into individual 3D classes using electron microscopy. First T4l.tGet2.apocyte.Get1single chains were prepared and solubilized in 1% DDM and purified. Half of the sample was reconstituted in amphipol. The DDM sample was first stained by 2% uranyl acetate for an electron microscopy negative stain study (Figure 4.6.1). Individual particles were picked and classified (Figure 4.6.2 and Figure 4.6.3 for random conical tilt (RCT)) then the starting negative stain model was prepared (Figure 4.6.4 and Figure 4.6.5). Initially, 17500 particles were manually picked from 207 micrographs using the EMAN2.1 software. From the picked particles a 2D class average was generated in IMAGIC (Figure 4.6.2) or in other packages. Each specimen stage was tilted 45 degree for RCT. Finally, a negative staining starting model was generated (Figure 4.6.3).

Thereafter, the samples in DDM and amphipol were vitrified. The amphipol sample was analyzed first. 400 images of T4l.tGet2.apocyte.Get1 in amphipol were recorded. Despite the small size of the protein, individual particles were seen (Figure 4.6.2.1). Furthermore, despite the relatively good quality of the 2d Class average (Figure 4.6.2.2), the majority of frames had

only between 15 and 20 particles and it was therefore decided to use the DDM sample. The DDM sample showed better results (Figure 4.6.2.3) with an average of over 200 particles per frame and no aggregates. The thon rings shown by the power spectrum does not have any astigmatisms meaning distorted rings. After the CTF was calculated, it shows that more than 50% of the data have information between 6 and 7 Å (Figure 4.6.2.4) at the micrograph level and about 50 out of 914 micrographs have a resolution between 3 and 4 Å, about one quarter between 8 and 10Å and the residual number below 10Å. Considering those CTF resolution estimation one could predict to have a structure with a resolution higher than 8 Å. Following the CTF estimation

210 000 particles were picked in similar 2D class (Figure 4.6.2.5). From the class averages, the starting Cryo-EM model was generated by applying no symmetry using the signification reconstruct from Scipion package (Figure 4.6.2.6).

However, again during the 3D reconstruction a heterogeneity analysis (Figure 4.6.2.7) gave 3D classes that did not converge during the refinement to a high resolution structure. It seems that within each class there is still heterogeneity despite the relatively nice resolution (10 Å) for this size 77 KDa. Still, to actually confirm the higher oligomer state Get1-strep was co-expressed with Get1-10xHis and Get2/Get1-strep with Get2/Get1-10xHis. Then, the complexes were pulled down by Ni-NTA purification and were afterwards identified by anti-strep antibodies (Figure 7.5). The results confirm that Get2/Get1 form heterotetramers. As the purified Get3 receptors are not stable in DDM and suitable for X-ray crystallography or high resolution Cryo-EM, Get3 receptors should be further reconstructed in lipid environment. As a result, one might express and purify them as a stable big complex with either a TA protein or together with a Get3.TA protein complex, while Get3 receptors are transferred to a lipid environment first, prior to any biophysical studies. This will likely help to reduce the heterogeneity observed when Get3 receptors are maintained in DDM.

5.10 ATP are present in pure Get3 receptors prepared samples

To check the probable cause of the disruption of Get3 when in complex with Get2/Get1 variants or Get1, the presence of ATP in solution was monitored (Figure 7.7, appendix). The ATP carryover during the preparation was checked by determining the change of ATP upon the increase of Get2/Get1 variants concentration in solution. For this study, the optimum construct the T4l.tGet2.apocyte.Get1 single chain and the full length single chain Get2/Get1 were used. In this experiment, the presence of ATP with concentration of 1.5 μM was detected. The effect is more significant with the native Get2/Get1. The unexplainable part of

the results was that ATP by itself is very instable to be carried over and that the intensity was stronger in presence of 1 mM calcium but reduced with 10 mM Hydroxylamine (Figure 7.7, appendix). Although there is no clear explanation for these results so far, an interesting point is that mass spectrometry identified impurities that are associated with mitochondria molecules involved in the synthesis of ATP such as mitochondrial cytochrome c oxidase protein 20 (COX20) (Figure 7.8 A, appendix), ATP synthesis protein 25 (Figure 7.8 B, appendix), mitochondrial cytochrome c oxidase subunit 6 (COX6) (Figure 7.8 C, appendix). Indeed, during the generation of cellular ATP through the mitochondrial oxidative phosphorylation, electrons are transferred from the NADH and FADH2 to the mitochondrial respiratory chain which produce a proton gradient across the inner mitochondrial membrane enabling the production of ATP by the F1F0 ATP synthase, and at the end of the electron transfer is the cytochrome c oxidase, which reduces molecular water to oxygen (Lorenzi *et al.*, 2016). More recently Get1 and Get2 have been shown to play an important role in the maintenance of the mitochondrial morphology and the levels of cardiolipin (Joshi, Fei, and Greenberg, 2016). Although this is not enough to explain why the mitochondrial proteins linked to the production of ATP are impurities observed in the Get3 receptors preparations.

Chapter F: OUTLOOK

Wang and colleagues have demonstrated that the TMD of Get2 and Get1 capture the TMD of TA proteins to help the releases of the TA proteins from Get3 (Wang *et al.*, 2015). Using the same mutations that led to the finding that Get2/Get1 is an insertase that facilitates the insertion of TA proteins, the complex Get2/Get1 TMD and the TMD of TA protein Sec22 can be produced *in vitro* in a lipid environment. The same approach was successfully performed to have the structure of the complex between ribosome-nascent chain-SecY complexes. A cysteine within the nascent chain was engineered and crosslinked to a cysteine in the plug of the Sec Y channel (Figure 1.4.1.4) (Park *et al.*, 2014). By selecting the important Get2/Get1 cysteine mutation, and doing a tandem purification, a stable conformation of the Get2/Get1-TA complex might be prepared in lipid like environment and could be studied by electron microscopy so as it could provide high resolution structure information.

LITERATURE

Abell, B. M., Jung, M., Oliver, J. D., Knight, B. C., Tyedmers, J., Zimmermann, R., & High, S. (2003). Tail-anchored and signal-anchored proteins utilize overlapping pathways during membrane insertion. *J Biol Chem, 278*(8), 5669-5678.

Abell, B. M., Pool, M. R., Schlenker, O., Sinning, I., & High, S. (2004). Signal recognition particle mediates post-translational targeting in eukaryotes. *EMBO J, 23*(14), 2755-2764.

Akopian, D., Dalal, K., Shen, K., Duong, F., & Shan, S. O. (2013). SecYEG activates GTPases to drive the completion of cotranslational protein targeting. *J Cell Biol, 200*(4), 397-405.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter P. (2002). Internal organization of the cell, membrane structure. Molecular Biology of the Cell. Fourth edition, Garland publishing

Andre, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G., Magnin, T., . . . Reinhart, C. (2006). Enhancing functional production of G protein-coupled receptors in Pichia pastoris to levels required for structural studies via a single expression screen. *Protein Sci, 15*(5), 1115-1126.

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics, 22*(2), 195-201.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter P. (2002). Internal organization of the cell, membrane structure. Molecular Biology of the Cell. Fourth edition, Garland publishing

Bäuerlein, E., Skrzipczyk, H. J., & Küchler, B. (1982). Butylhydroxylamine inhibits H+-driven ATP synthesis of the TF1 .Fo-ATPase incorporated into liposomes. *FEBS Lett, 141*, 173-175.

Beilharz, T., Egan, B., Silver, P. A., Hofmann, K., & Lithgow, T. (2003). Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in Saccharomyces cerevisiae. *J Biol Chem, 278*(10), 8219-8223.

Bell, J. M., Chen, M., Baldwin, P. R., & Ludtke, S. J. (2016). High resolution single particle refinement in EMAN2.1. *Methods, 100*, 25-34.

Bill, R. M., Henderson, P. J., Iwata, S., Kunji, E. R., Michel, H., Neutze, R., . . . Vogel, H. (2011). Overcoming barriers to membrane protein structure determination. *Nat Biotechnol, 29*(4), 335-340.

Borgese, N., Brambillasca, S., & Colombo, S. (2007). How tails guide tail-anchored proteins to their destinations. *Curr Opin Cell Biol, 19*(4), 368-375.

Borgese, N., Colombo, S., & Pedrazzini, E. (2003). The tale of tail-anchored proteins: coming from the cytosol and looking for a membrane. *J Cell Biol, 161*(6), 1013-1019.

Borgese, N., & Fasana, E. (2011). Targeting pathways of C-tail-anchored proteins. *Biochim Biophys Acta, 1808*(3), 937-946.

Borgese, N., & Righi, M. (2010). Remote origins of tail-anchored proteins. *Traffic, 11*(7), 877-885.

Brambillasca, S., Yabal, M., Makarow, M., & Borgese, N. (2006). Unassisted translocation of large polypeptide domains across phospholipid bilayers. *J Cell Biol, 175*(5), 767-777.

Brambillasca, S., Yabal, M., Soffientini, P., Stefanovic, S., Makarow, M., Hegde, R. S., & Borgese, N. (2005). Transmembrane topogenesis of a tail-anchored protein is modulated by membrane lipid composition. *EMBO J, 24*(14), 2533-2542.

Bulbarelli, A., Sprocati, T., Barberi, M., Pedrazzini, E., and Borgese, N. (2002). Trafficking of tail-anchored proteins: transport from the endoplasmic reticulum to the plasma membrane and sorting between surface domains in polarised epithelial cells. *J Cell Sci, 115*, 1689-1702.

Chartron, J. W., Clemons, W. M., Jr., & Suloway, C. J. (2012). The complex process of GETting tail-anchored membrane proteins to the ER. *Curr Opin Struct Biol, 22*(2), 217-224.

Chartron, J. W., Suloway, C. J., Zaslaver, M., & Clemons, W. M., Jr. (2010). Structural characterization of the Get4/Get5 complex and its interaction with Get3. *Proc Natl Acad Sci U S A, 107*(27), 12127-12132.

Chartron, J. W., VanderVelde, D. G., & Clemons, W. M., Jr. (2012). Structures of the Sgt2/SGTA dimerization domain with the Get5/UBL4A UBL domain reveal an interaction that forms a conserved dynamic interface. *Cell Rep, 2*(6), 1620-1632.

Chartron, J. W., VanderVelde, D. G., Rao, M., & Clemons, W. M., Jr. (2012). Get5 carboxyl-terminal domain is a novel dimerization motif that tethers an extended Get4/Get5 complex. *J Biol Chem, 287*(11), 8310-8317.

Chen, C. M., Misrat, T. K., Silver, S., & Rosen, B. P. (1986). Nucleotide Sequence of the Structural Genes for an Anion Pump: The plasmid-encoded arsenical resistance operon. *J Biol Chem, 261*, 15030-15038.

Christianson, T. .., Sikorski, R.S., Dante, M., Shero, J.H., and Hieter, P. (1992). Multifunctional yeast high-copy-number shuttle vectors. *Gene, 110*, 119-122.

Chun, E., Thompson, A. A., Liu, W., Roth, C. B., Griffith, M. T., Katritch, V., . . . Stevens, R. C. (2012). Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure, 20*, 967-976.

Colombo, S. F., Cardani, S., Maroli, A., Vitiello, A., Soffientini, P., Crespi, A., . . . Borgese, N. (2016). Tail-anchored Protein Insertion in Mammals: FUNCTION AND RECIPROCAL INTERACTIONS OF THE TWO SUBUNITS OF THE TRC40 RECEPTOR. *J Biol Chem, 291*(29), 15292-15306.

Dalbey, R. E., & Chen, M. (2004). Sec-translocase mediated membrane protein biogenesis. *Biochim Biophys Acta, 1694*(1-3), 37-53. doi:10.1016/j.bbamcr.2004.03.009

de Gier, J. W. L., Mansournia, P., Valent, Q. A., Phillips , G. J., Luirink, J., & von Heijne, J. (1996). Assembly of a cytoplasmic membrane protein in Escherichia coli is dependent on the signal recognition particle. *FEBS Lett, 399*, 307-309.

de la Rosa-Trevin, J. M., Quintana, A., Del Cano, L., Zaldivar, A., Foche, I., Gutierrez, J., . . . Carazo, J. M. (2016). Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol, 195*(1), 93-99.

de Marco, A., Vigh, L., Diamant, S., & Goloubinoff, P. (2000). Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. *Cell Stress Chaperones, 10*, 329-239.

Denic, V., Dotsch, V., & Sinning, I. (2013). Endoplasmic reticulum targeting and insertion of tail-anchored membrane proteins by the GET pathway. *Cold Spring Harb Perspect Biol, 5*(8), a013334.

Denks, K., Vogt, A., Sachelaru, I., Petriman, N. A., Kudva, R., & Koch, H. G. (2014). The Sec translocon mediated protein transport in prokaryotes and eukaryotes. *Mol Membr Biol, 31*(2-3), 58-84.

Dennerlein, S., & Rehling, P. (2015). Human mitochondrial COX1 assembly into cytochrome c oxidase at a glance. *J Cell Sci, 128*(5), 833-837. doi:10.1242/jcs.161729

Deshaies, R. J., & Schekman, R. (1989). SEC62 Encodes a Putative Membrane Protein Required for Protein Translocation into the Yeast Endoplasmic Reticulum. *The Journal of Cell Biology, 109*, 2653-2664.

Drenth, J. (1999). Principles of protein X-Ray crystallography, Springer advanced texts in biology. 2nd edition

Drew, D., Newstead, S., Sonoda, Y., Kim, H., von Heijne, G., & Iwata, S. (2008). GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in Saccharomyces cerevisiae. *Nat Protoc, 3*(5), 784-798.

Drew, D.E., von Heijne, G., Nordlund, P., & de Gier, J-W. L. (2001). Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. Febs Letters, 507, 220-224

Favaloro, V., Spasic, M., Schwappach, B., & Dobberstein, B. (2008). Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J Cell Sci, 121*(11), 1832-1840.

Flick, J.S., & Johnston (1990). Two Systems of Glucose Repression of the GAL] Promoter in Saccharomyces cerevisiae. MOLECULAR AND CELLULAR BIOLOGY, 4757-4769.

Freigassner, M., Pichler, H., & Glieder, A. (2009). Tuning microbial hosts for membrane protein production. *Microb Cell Fact, 8*, 69.

Gristick, H. B., Rao, M., Chartron, J. W., Rome, M. E., Shan, S. O., & Clemons, W. M., Jr. (2014). Crystal structure of ATP-bound Get3-Get4-Get5 complex reveals regulation of Get3 by Get4. *Nat Struct Mol Biol, 21*(5), 437-442.

Gristick, H. B., Rome, M. E., Chartron, J. W., Rao, M., Hess, S., Shan, S. O., & Clemons, W. M., Jr. (2015). Mechanism of Assembly of a Substrate Transfer Complex during Tail-anchored Protein Targeting. *J Biol Chem, 290*(50), 30006-30017.

Hattori, M., Hibbs, R. E., & Gouaux, E. (2012). A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure, 20*(8), 1293-1299.

Hegde, R. S., & Keenan, R. J. (2011). Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat Rev Mol Cell Biol, 12*(12), 787-798. doi:10.1038/nrm3226

Hegde, R. S., & Ploegh, H. L. (2010). Quality and quantity control at the endoplasmic reticulum. *Curr Opin Cell Biol, 22*(4), 437-446.

Hieter, R. S. S. a. P. (1989). A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces ceratisiae. *Genetics, 122*, 19-27.

High, S., & Abell, B. M. (2004). Tail-anchored protein biosynthesis at the endoplasmic reticulum: the same but different. *Biochemical Society Transactions, 32*, 660-662.

Huttemann, M., Pecina, P., Rainbolt, M., Sanderson, T. H., Kagan, V. E., Samavati, L., . . . Lee, I. (2011). The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis. *Mitochondrion, 11*(3), 369-381.

Iwata, S., et al. (2003). Methods and results in crystallization of membrane proteins. IUL Biotechnology series

Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H. M., Duhr, S., Baaske, P., & Breitsprecher, D. (2014). MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure, 1077,* 101-113.

Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P., & Duhr, S. (2011). Molecular interaction studies using microscale thermophoresis. *Assay Drug Dev Technol, 9*(4), 342-353.

Johnston, M. (1987) A Model Fungal Gene Regulatory Mechanism: the GAL Genes of Saccharomyces Cerevisiae. MICROBIOLOGICAL REVIEWS, 458-476.

Johnson, N., Powis, K., & High, S. (2013). Post-translational translocation into the endoplasmic reticulum. *Biochim Biophys Acta, 1833*(11), 2403-2409.

Johnston, M., Flick, J. S., & Pexton, T. (1994). Multiple mechanisms provide rapid and stringent glucose repression of GAL gene expression in Saccharomyces cerevisiae. *Molecular and Cellular Biology, 14*(6), 3834-3841.

Joshi, A.S., Fei, N., and Greenberg, M.L. The GET complex is required for maintenance of mitochondrial morphology and normal cardiolipin levels. FEMS Yeast Res., in press.

Jungnickel, B., & Rapoport, T. A. (1995). A Posttargeting Signal Sequence Recognition Event in the Endoplasmic Reticulum Membrane. *Cell, 82*, 261-270.

Kastner, B., Fischer, N., Golas, M. M., Sander, B., Dube, P., Boehringer, D., . . . Stark, H. (2008). GraFix: sample preparation for single-particle electron cryomicroscopy. *Nat Methods, 5*(1), 53-55.

Katay, U., Hartmann, E and Tom A. Rapoport. (1993). A class of membrane proteins with a Co.terminal anchor. TRENDS IN CELL BIOLOGY, 3.

Keenan, R. J., Freymann, D. M., Stroud, R. M., & Walter, P. (2001). The signal recognition particle. *Annu Rev Biochem, 70,* 755-775. Kelly, L., Pieper, U., Eswar, N., Hays, F.

A., Li, M., Roe-Zurz, Z., . . . Sali, A. (2009). A survey of integral alpha-helical membrane proteins. *J Struct Funct Genomics, 10*(4), 269-280.

Kelly, S. M., & Price, N. C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Current Protein and Peptide Science, 1*, 349-384.

Kjeldsen, T., Ludvigsen, S., Diers, I., Balschmidt, P., Sorensen, A. R., & Kaarsholm, N. C. (2002). Engineering-enhanced protein secretory expression in yeast with application to insulin. *J Biol Chem, 277*(21), 18245-18248.

Kriechbaumer, V., Shaw, R., Mukherjee, J., Bowsher, C. G., Harrison, A. M., & Abell, B. M. (2009). Subcellular distribution of tail-anchored proteins in Arabidopsis. *Traffic, 10*(12), 1753-1764.

Krogh, A., Larsson, B., von Heijne, G., & Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol, 305*(3), 567-580. 5

Kuhlbrandt, W. (2014). Biochemistry. The resolution revolution. *Science, 343*(6178), 1443-1444.

Kupiecki , F. P., & Coon, M. J. (1959). Bicarbonate- and Hydroxylamine-dependent Degradation of Adenosine Triphosphate. *J Biol Chem, 234*, 2428-2432.

Kutay, U., Ahnert-Hilgerl, G., Hartmann, E., Wiedenmann, B., & Rapoport, T. A. (1995). Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. *EMBO J, 14*, 217-223.

Kutay, U., Hartmann, E., & Rapoport, T. A. (1993). A class of membrane proteins with a C-terminal anchor. *Trends Cell Biol, 3*(3), 72-75.

Le Parc, A., Leonil, J., & Chanat, E. (2010). AlphaS1-casein, which is essential for efficient ER-to-Golgi casein transport, is also present in a tightly membrane-associated form. *BMC Cell Biol, 11*, 65. doi:10.1186/1471-2121-11-65

Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., . . . Cheng, Y. (2013). Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods, 10*(6), 584-590.

LIBEREK, K., MARSZALEK, J., ANG, D., GEORGOPOULOS, C., & ZYLICZ, M. (1991). Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc Natl Acad Sci U S A, 88*.

Lingwood, D., & Simons, K. (2010). Lipid rafts as a membrane-organizing principle. *Science, 327*(5961), 46-50.

Lorenzi, I., Oeljeklaus, S., Ronsor, C., Bareth, B., Warscheid, B., Rehling, P., & Dennerlein, S. (2016). The ribosome-associated Mba1 escorts Cox2 from insertion machinery to maturing assembly intermediates. *Mol Cell Biol*. doi:10.1128/MCB.00361-16

Mandon, E. C., Trueman, S. F., & Gilmore, R. (2009). Translocation of proteins through the Sec61 and SecYEG channels. *Curr Opin Cell Biol, 21*(4), 501-507.

Mariappan, M., Mateja, A., Dobosz, M., Bove, E., Hegde, R. S., & Keenan, R. J. (2011). The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature, 477*(7362), 61-66.

Marty, N. J., Teresinski, H. J., Hwang, Y. T., Clendening, E. A., Gidda, S. K., Sliwinska, E., . . . Mullen, R. T. (2014). New insights into the targeting of a subset of tail-anchored proteins to the outer mitochondrial membrane. *Front Plant Sci, 5*, 426.

Mateja, A., Paduch, M., Chang, H. Y., Szydlowska, A., Kossiakoff, A. A., Hegde, R. S., & Keenan, R. J. (2015). Protein targeting. Structure of the Get3 targeting factor in complex with its membrane protein cargo. *Science, 347*(6226), 1152-1155.

Mateja, A., Szlachcic, A., Downing, M. E., Dobosz, M., Mariappan, M., Hegde, R. S., & Keenan, R. J. (2009). The structural basis of tail-anchored membrane protein recognition by Get3. *Nature, 461*(7262), 361-366.

McPherson, A. (2004). Introduction to protein crystallization. *Methods, 34*(3), 254-265.

Mercanti, V., & Cosson, P. (2010). Resistance of Dictyostelium discoideum membranes to saponin permeabilization. *BMC Res Notes, 3*, 120.

Miller, J. D., Bernstein, H. D., & WALTER, P. (1994). Interaction of E.coli Fff/4.5S ribonucleoprotein and FtsY mimics that of mammalian signal recognition particle and receptor. *Nature, 367*, 657-659.

Miller, J. D., Tajima, S., Lauffer, L., & WALTER, P. (1995). The fl Subunit of the Signal Recognition Particle Receptor Is a Transmembrane GTPase that Anchors the a Subunit, a Peripheral Membrane GTPase, to the Endoplasmic Reticulum Membrane. *The Journal of Cell Biology, 128*, 273-282.

Mindell, J. A., & Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of Structural Biology, 142*(3), 334-347.

Molleken, K., Schmidt, E., & Hegemann, J. H. (2010). Members of the Pmp protein family of Chlamydia pneumoniae mediate adhesion to human cells via short repetitive peptide motifs. *Mol Microbiol, 78*(4), 1004-1017.

Moreno, A., SantoDomingo, J., Fonteriz, R. I., Lobaton, C. D., Montero, M., & Alvarez, J. (2010). A confocal study on the visualization of chromaffin cell secretory vesicles with fluorescent targeted probes and acidic dyes. *J Struct Biol, 172*(3), 261-269.

Morgner, N., Kleinschroth, T., Barth, H. D., Ludwig, B., & Brutschy, B. (2007). A novel approach to analyze membrane proteins by laser mass spectrometry: from protein subunits to the integral complex. *J Am Soc Mass Spectrom, 18*(8), 1429-1438.

Morth, J. P., Pedersen, B. P., Toustrup-Jensen, M. S., Sorensen, T. L., Petersen, J., Andersen, J. P., . . . Nissen, P. (2007). Crystal structure of the sodium-potassium pump. *Nature, 450*(7172), 1043-1049.

Moroi, Y. (1992). Micelles: Theoretical and applied aspects. Plenum Press

Mukhopadhyay, R., Ho, Y. S., Swiatek, P. J., Rosen, B. P., & Bhattacharjee, H. (2006). Targeted disruption of the mouse Asna1 gene results in embryonic lethality. *FEBS Lett, 580*(16), 3889-3894.

Mumberg, D., Mailer, R., & Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene, 156*, 119-122.

Newstead, S., Kim, H., von Heijne, G., Iwata, S., & Drew, D. (2007). High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A, 104*(35), 13936-13941.

Nyblom, M., Oberg, F., Lindkvist-Petersson, K., Hallgren, K., Findlay, H., Wikstrom, J., . . . Hedfalk, K. (2007). Exceptional overproduction of a functional human membrane protein. *Protein Expr Purif, 56*(1), 110-120.

Nyblom, M., Poulsen, H., Gourdon, P., Reinhard, L., Andersson, M., Lindahl, E., . . . Nissen, P. (2013). Crystal structure of Na+, K(+)-ATPase in the Na(+)-bound state. *Science, 342*(6154), 123-127.

Oganesyan, N., Ankoudinova, I., Kim, S. H., & Kim, R. (2007). Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expr Purif, 52*(2), 280-285.

Ogg, S. C., Barz, W. P., & Walter, P. (1998). A Functional GTPase Domain, but not its Transmembrane Domain, is Required for Function of the SRP Receptor -subuni. *The Journal of Cell Biology, 142*, 341–354.

Palamarczyk, G., Drake, R., Haley, B., & Lennarz, W. J. (1990). Evidence that the synthesis of glucosylphosphodolichol in yeast involves a 35-kDa membrane protein. *Proc Natl Acad Sci U S A, 87*(7), 2666-2670.

Panzner, S., Dreier, L., Hartmann, E., Kostka, S., & Rapoport, T. A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell, 81*(4), 561-570.

Park, E., Menetret, J. F., Gumbart, J. C., Ludtke, S. J., Li, W., Whynot, A., . . . Akey, C. W. (2014). Structure of the SecY channel during initiation of protein translocation. *Nature, 506*(7486), 102-106.

Park, K. S., & Kim, J. S. (2006). Engineering of GAL1 promoter-driven expression system with artificial transcription factors. *Biochem Biophys Res Commun, 351*(2), 412-417.

Pedersen, B. P., Buch-Pedersen, M. J., Morth, J. P., Palmgren, M. G., & Nissen, P. (2007). Crystal structure of the plasma membrane proton pump. *Nature, 450*(7172), 1111-1114.

Pedrazzini, E. (2009). Tail-Anchored Proteins in Plants. *Journal of Plant Biology, 52*(2), 88-101.

Pestov, N. B., & Rydstrom, J. (2007). Purification of recombinant membrane proteins tagged with calmodulin-binding domains by affinity chromatography on calmodulin-agarose: example of nicotinamide nucleotide transhydrogenase. *Nat Protoc, 2*(1), 198-202.

Poritz, M. A., Strub, K., & Walter, P. (1988). Human SRP RNA and E. coli 4.5S RNA contain a highly homologous structural domain. Cell, 55(1), 4-6.

Rabu, C., Schmid, V., Schwappach, B., & High, S. (2009). Biogenesis of tail-anchored proteins: the beginning for the end? J Cell Sci, 122(Pt 20), 3605-3612. doi:10.1242/jcs.041210

Rabu, C., Wipf, P., Brodsky, J. L., & High, S. (2008). A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum. J Biol Chem, 283(41), 27504-27513.

Rapoport, T. A. (1992). Transport of Proteins Across the Endoplasmic Reticulum Membrane. Science, 258.

Rensing, S. A., & Maier, U. G. (1994). The SecY protein family: comparative analysis and phylogenetic relationships. Mol Phylogenet Evol, 3(3),

Richards, F. M., & Vithayathil, P. J. (1959). the preparation of the subtilisin-modified Ribonuclease and the separation of the peptide and protein component. THE JOURNAL OF BIOLOGICAL CHEMISTRY, 234, 1459-1465.

Rhodes, G. (2000) Crystallography made crystal clear: A guide for users of macromolecules models. Academic press. 2nd edition.

Rome, M. E., Rao, M., Clemons, W. M., & Shan, S. O. (2013). Precise timing of ATPase activation drives targeting of tail-anchored proteins. Proc Natl Acad Sci U S A, 110(19), 7666-7671.

Rudiger, S., Buchberger, A., & Bukau, B. (1997). Interaction of Hsp70 chaperones with substrates. Nat Struct Biol, 4(5), 342-349.

Samuelson, J. C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., . . . Dalbey, R. E. (2000). YidC mediates membrane protein insertion in bacteria. Nature, 406(6796), 637-641.

Schatz, P. J., Biekerl, K. L., Ottemann, K. M., Silhavy, T. J., & Beckwith, J. (1991). One of three transmembrane stretches is sufficient for the functioning of the SecE protein, a membrane component of the E.coli secretion machinery. EMBO J, 10, 1 749 - 1757,.

Scheres, S. H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol, 180(3), 519-530.

Schmidt, I., Look, C., Bock, E., & Jetten, M. S. (2004). Ammonium and hydroxylamine uptake and accumulation in Nitrosomonas. Microbiology, 150(Pt 5), 1405-1412.

Schuldiner, M., Collins, S. R., Thompson, N. J., Denic, V., Bhamidipati, A., Punna, T., . . . Krogan, N. J. (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell, 123(3), 507-519.

Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H. D., . . . Weissman, J. S. (2008). The GET complex mediates insertion of tail-anchored proteins into the ER membrane. Cell, 134(4), 634-645.

Seidel, S. A., Dijkman, P. M., Lea, W. A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., . . . Duhr, S. (2013). Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. Methods, 59(3), 301-315.

Seluanov, A., & Bib, E. (1997). FtsY, the Prokaryotic Signal Recognition Particle Receptor

Homologue, Is Essential for Biogenesis of Membrane Proteins. THE JOURNAL OF BIOLOGICAL CHEMISTRY, 272, 2053–2055.

Shao, S., & Hegde, R. S. (2011). Membrane protein insertion at the endoplasmic reticulum. Annu Rev Cell Dev Biol, 27, 25-56.

Sikora, C. W., & Turner, R. J. (2005). SMR proteins SugE and EmrE bind ligand with similar affinity and stoichiometry. Biochem Biophys Res Commun, 335(1), 105-111.

Sikorski, R. S., & Hieter, P. (1989). A System of Shuttle Vectors and Yeast Host Strains Designed for Efficient Manipulation of DNA in Saccharomyces ceratisiae. Genetics, 122, 19-27.

Simpson, P. J., Schwappach, B., Dohlman, H. G., & Isaacson, R. L. (2010). Structures of Get3, Get4, and Get5 provide new models for TA membrane protein targeting. Structure, 18(8), 897-902.

Stefanovic, S., & Hegde, R. S. (2007). Identification of a targeting factor for posttranslational membrane protein insertion into the ER. Cell, 128(6), 1147-1159.

Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y. Y., Schwarz, D., . . . Sinning, I. (2011). Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. Science, 333(6043), 758-762.

Sui, Y., Liu, J., Wisniewski, M., Droby, S., Norelli, J., & Hershkovitz, V. (2012). Pretreatment of the yeast antagonist, Candida oleophila, with glycine betaine increases oxidative stress tolerance in the microenvironment of apple wounds. Int J Food Microbiol, 157(1), 45-51.

Suloway, C. J., Rome, M. E., & Clemons, W. M., Jr. (2012). Tail-anchor targeting by a Get3 tetramer: the structure of an archaeal homologue. EMBO J, 31(3), 707-719.

Sumner, J.B.(1926). The isolation and the crystallization of the enzyme urease. Preliminary paper. J Biol Chem, 69: 435-441

Tanaka, Y., Sugano, Y., Takemoto, M., Mori, T., Furukawa, A., Kusakizako, T., . . . Tsukazaki, T. (2015). Crystal Structures of SecYEG in Lipidic Cubic Phase Elucidate a Precise Resting and a Peptide-Bound State. Cell Rep, 13(8), 1561-1568.

Thielmann, Y., Koepke, J., & Michel, H. (2012). The ESFRI Instruct Core Centre Frankfurt: automated high-throughput crystallization suited for membrane proteins and more. J Struct Funct Genomics, 13(2), 63-69.

Thorsen, T. S., Matt, R., Weis, W. I., & Kobilka, B. K. (2014). Modified T4 Lysozyme Fusion Proteins Facilitate G Protein-Coupled Receptor Crystallogenesis. Structure, 22(11), 1657-1664.

Tung, J. Y., Li, Y. C., Lin, T. W., & Hsiao, C. D. (2013). Structure of the Sgt2 dimerization domain complexed with the Get5 UBL domain involved in the targeting of tail-anchored membrane proteins to the endoplasmic reticulum. Acta Crystallogr D Biol Crystallogr, 69(Pt 10), 2081-2090.

Tutar, Y., Song, Y., & Masison, D. C. (2006). Primate chaperones Hsc70 (constitutive) and Hsp70 (induced) differ functionally in supporting growth and prion propagation in Saccharomyces cerevisiae. Genetics, 172(2), 851-861.

Ulrike Kutay, G. A.-H., Enno Hartmann, B. W. a., & A.Rapoport, T. (1995). Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. EMBO J, 14, 217-223.

Ungar, D., & Hughson, F. M. (2003). SNARE protein structure and function. Annu Rev Cell Dev Biol, 19, 493-517.

Van Ael, E., & Fransen, M. (2006). Targeting signals in peroxisomal membrane proteins. Biochimica et Biophysica Acta 1763, 1629-1638

van den Ent, F., & Lowe, J. (2006). RF cloning: a restriction-free method for inserting target genes into plasmids. J Biochem Biophys Methods, 67(1), 67-74.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T., . . . Patwardhan, A. (2000). Single-particle electron cryo-microscopy: towards atomic resolution. Q Rev Biophys, 33(4), 307-369. Vilardi, F., Lorenz, H., & Dobberstein, B. (2011). WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. J Cell Sci, 124(Pt 8), 1301-1307. Vilardi, F., Stephan, M., Clancy, A., Janshoff, A., & Schwappach, B. (2014). WRB and CAML are necessary and sufficient to mediate tail-anchored protein targeting to the ER membrane. PLoS One, 9(1), e85033.

Voorhees, R. M., & Hegde, R. S. (2016). Structure of the Sec61 channel opened by a signal sequence. Science, 351, 88-91.

Walter, P., & Blobel, G. (1980). Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. Proc Natl Acad Sci U S A, 77(12), 7112-7116.

Wang, F., Brown, E. C., Mak, G., Zhuang, J., & Denic, V. (2010). A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. Mol Cell, 40(1), 159-171.

Wang, F., Chan, C., Weir, N. R., & Denic, V. (2014). The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. Nature, 512(7515), 441-444. doi:10.1038/nature13471

Wang, F., Whynot, A., Tung, M., & Denic, V. (2011). The mechanism of tail-anchored protein insertion into the ER membrane. Mol Cell, 43(5), 738-750.

Watts, L. A. W. a. T. D. (1990). Yeast Acetyl-CoA Carboxylase: In Vitro Phosphorylation By Mammalian Yeast Protein Kinases.

Welch, W. J., & Brown, C. R. (1996). Influence of molecular and chemical chaperones on protein folding. Cell Stress Chaperones, 1(2), 109-115.

Whitelegge, J. P. (2013). Integral membrane proteins and bilayer proteomics. Anal Chem, 85(5), 2558-2568.

Woolford, C. A., Daniels, L. B., Park, F. J., Jones, E. W., Van Arsdell, J. N., & Innis, M. A. (1986). The PEP4 gene encodes an aspartyl protease implicated in the posttranslational regulation of Saccharomyces cerevisiae vacuolar hydrolases. Molecular and Cellular Biology, 6(7), 2500-2510.

Woolford, C. A., Daniels, L. B., Park, F. J., Jones, E. W., Van Arsdell, J. N., & Innis, M. A. (1986). The PEP4 Gene Encodes an Aspartyl Protease Implicated in the Posttranslational Regulation of Saccharomyces cerevisiae Vacuolar Hydrolases. Molecular and Cellular Biology, 2500-2510.

Yabal, M., Brambillasca, S., Soffientini, P., Pedrazzini, E., Borgese, N., & Makarow, M. (2003). Translocation of the C terminus of a tail-anchored protein across the endoplasmic reticulum membrane in yeast mutants defective in signal peptide-driven translocation. J Biol Chem, 278(5), 3489-3496.

Yamamoto, Y., & Sakisaka, T. (2012). Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. Mol Cell, 48(3), 387-397.

Zhou, T., Radaev, S., Rosen, B. P., & Gatti, D. L. (2000). Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump. EMBO J, 19(17), 4838-4845.

Zimmermann, R., Eyrisch, S., Ahmad, M., & Helms, V. (2011). Protein translocation across the ER membrane. Biochim Biophys Acta, 1808(3), 912-924.

Zimmermann, R., Muller, L., & Wullich, B. (2006). Protein transport into the endoplasmic reticulum: mechanisms and pathologies. Trends Mol Med, 12(12), 567-573.

Zopf, D., Bernstein, H. D., & WALTER, P. (1993). GTPase Domain of the 54-kD Subunit of the Mammalian Signal Recognition Particle Is Required for Protein Translocation But Not for Signal Sequence Binding. The Journal of Cell Biology, 120, 1113-1121.

Zou, Y., Weis, W. I., & Kobilka, B. K. (2012). N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. PLoS One, 7(10), e46039.

APPENDIX

7.1 Construction of Vector for Expression of Get1, Get2, truncated_Get2, Get2/Get1 and truncated_Get2/Get Vectors

Get1, Get2, truncated_Get2, the single chain Get2/Get1 and its truncated_Get2/Get1 version were amplified according to the protocols in Tables 3.2, 3.3, 3.4, 3.5 and 3.6. In the truncated version of Get2 the first 128 amino acid residues were removed within the design of primers. The initial single chain Get2 and Get1 (Get2/Get1) DNA was gifted by Prof. Dr Vladimir Denic (Harvard University), and the construct includes ASGSGSGSGSGT linker sequence between Get2 at the N-terminus and Get1. The forward primer of the N-terminus protein anneal to the N-terminus of Get2 while the reverse primer to the C-terminus of Get1 during the amplification of the single chain. The primers used for each target are described in Table 3.9. The PCR product of each specific gene was digested (*Bam*H1 and *Hin*dIII) then ligated with a specific restriction enzyme of cleaved pJANY-S1 (for strep tagged protein purification) or of pJANY-H1 (for 10X-his tagged protein purification). After an hour, 1 µl of the ligation mix was used to transform DH5α cells. Each transformation reaction was plated on Luria-Bertani (LB) agar in presence of ampicillin for overnight cultuvation. The next day colonies were picked separately to inoculate 8 ml of LB media in presence of ampicillin for mini preparation. The resulting expression vectors are named pJNG1-S1 (Get1-strep), pJNG21-S1 (Get2/Get1-strep), pJNG2-S1 (Get2-strep), pJNG1-H1 (Get1-10xhis), pJNG21-H1 (Get2/Get1-10xhis), pJNG2-H1 (Get2-10xhis), pJNGt2-S1 (truncated_Get2-strep) and pJNGt21-S1 (truncated_Get2/Get1-strep). The vectors used have pRS426 backbone consisting of *URA3* marker selection in yeast to further increase the possibilities the *GAL1* promoter gene, the gene of interest and the *CYC1* terminator gene of the transcription (Patrick Russo and Fred Sherman, 1989) were cleaved by *Pvu*II restriction enzyme and transferred to the vector backbone *pRS423* digested by *Eco*RI and *Xho*I. Therefore the new vector pJNG1-S2 (Get1-strep), pJNG21-S2 (Get2/Get1-strep), pJNG2-S2 (Get2-strep), pJNG1-H2 (Get1-10xhis), pJNG21-H2 (Get2/Get1-10xhis), pJNG2-H2 (Get2-10xhis), pJNGt2-S2 (truncated_Get2-strep) and pJNGt21-S2 (truncated_Get2/Get1-strep) consist of *HIS3* marker for yeast selection and maintenance.

7.2 Get3 based expression vector in yeast

Get3 was introduced into backbone pRS423, therefore using Get3 primers from table 10 to amplify Get3 and the PCR product was introduced into BamH1/Sal1 digested pJNG21-H2 using gap-repair cloning. The pJNGet3-H2 vector was obtained.

Construction of expressing T4L.Get2 and T4L.Get2/Get1 vectors

As Get2 has long N-terminal flexible part, in order to improve the odds of well order crystals the N-terminal region from residues 1 to 128 was replaced by n-terminal T4 Lysozyme. This approach has shown to facilitate the crystallization of G-protein coupled receptors (Zou *et al.*, 2012). pJNGt2-S1, pJNGt21-S1, pJNGt2-H1, and pJNGt21-H1 were digested by *Bam*H1 to generate by homologous recombination new yeast expression plasmids that carry respectively the fusion proteins T4l.tGet2-strep, T4l.tGet2/Get1-strep, T4l.tGet2-10xHis, T4l.tGet2/Get1-10xHis. The primers of T4L amplification are described in table 2.1

7.3 Preparation of the TA protein pep12 vector: pJNPep12-H3

PRS425 backbone was used to generate the vector pJNPep12-H3 which enables the expression of the transmembrane part of the TA protein pep12. The ***pRS425*** backbone carry *LEU2* marker for yeast selection and maintenance. The vector was cleaved by BamHI and HindIII. The Gal1 promoter was then amplified with 40 bp homologous sequence overhang to the pRS425 cleaved vector in one side and another primer 40 bp homologous sequence to pep12 (Table 3.12). The pJNPep12-H3 was created by Gap repair process in mixing the pRS425 cleaved, the Gal promoter PCR product and three additional homologous oligos (Figure 3.1).

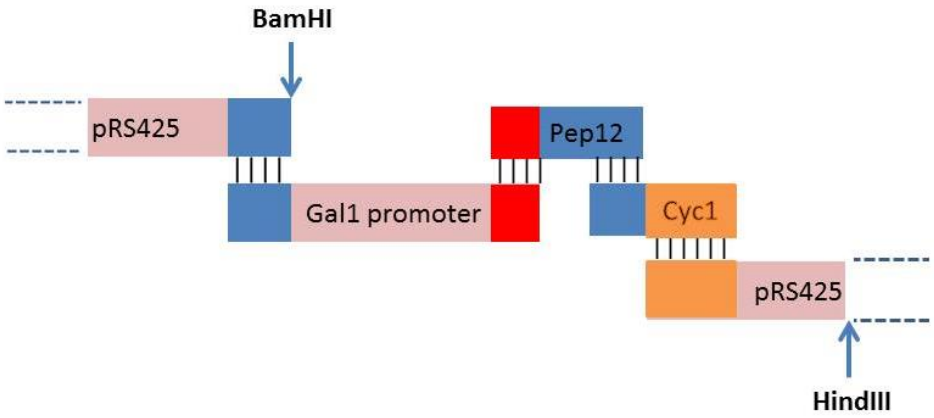


Figure 7.3: Schematic of the preparation pJNPep12-S3 vector by homologous recombination. pRS425 is the BamHI, HindII cleaved vector backbone, Gal promoter the PCR product and 3 distinct oligos overlapping to each other; the Pep12 has the homologous sequence with the 40 bp overhang of the 3' primer of Gal1 promoter PCR product, Cyc1 is 38 base pair transcription terminator plus an extension sequence matching 40 bp of the 3' end of the Pep12, last oligo is the mixture of cyc1 sequence plus 40 bp of homologous sequence to cleaved pRS425.

TABLE 7.3: PRIMER OF GAL PROMOTER AND THE TRUNCATED PEP12

PRIMER		Sequences
GAL PROMOTER	Forward	5´-cagtgagcgcgcgtaatacgactcactatagggcgaattgagtacggattagaagccgcc-3´
	Reverse	5´-aagaagcattacgagaagcacaatcaacaaatacaccctcatatgggatggtgatcaagatc-3´
TRUNCATED PEP12		5´-atgagggtgtatttgttgattgtgcttctcgtaatgcttctttttatttttctcattatgaaattgtaa-3´
PEP12/CYC1T		5´-ttctcgtaatgcttctttttatttttctcattatgaaattgtaaaaaaccttgcttgagaaggttttgggacgctcgaaggcttaat-3´
CYC1T/PRS425		5´-aaaaccttgcttgagaaggttttgggacgctcgaaggctttaatttgcggtggagctccagcttttgttccctttagtgaggttaat-3´

7.4. Site-directed mutagenesis of Get2/Get1

In order to stabilize Get1 against its exposure to any proteases that may have destabilized it, several mutants of Get1 within the single chain construct of Get2/Get1 were prepared.

TABLE 7.4: MUTATION OF GET1 WITHIN THE SINGLE CHAIN CONSTRUCT GET2/GET1

SINGLE MUTANTS	K23M; K27M; K31M; K173M; S30A; S172A
DOUBLE MUTANTS	K23,173M; K27,173M; K31,173M; S30;172A
TRIPLE MUTANTS	K23,27,173M; K23,31,173M; K27,31, 173M

To generate each mutant of Get1, two single primer PCR reactions were performed in parallel as described using the SPRING method (Edelheit *et al.*, 2009). Q5 high-fidelity DNA polymerase was used for PCR as the extension time is reduced to 20 seconds per kb in comparison to 2 minutes per kb for the commonly used Pfu ultra polymerase. Then both PCR were combined, denatured at 95 °C then both the parental DNA and the newly synthesized strands were randomly annealed by slow cooling according to parameters in table 3.14.

TABLE 7.4.1: DENATURING AND SLOW COOLING OF MIXTURE OF PARENTAL DNA AND SYNTHESIZED DNA CARRYING POINT MUTATION

TEMPERATURE IN °C	Time in minutes
95	5
90	1
80	1
70	0.5
60	0.5
50	0.5
40	0.5
37	0.5

In order to digest the methylated parental DNA in the reaction above, 40 U of restriction enzyme *dpnI* were added to the 50 µl reaction mixture and incubated overnight at 37 °C. The following day 10 µl the *DpnI* treated sample were used to transform competent *E.coli DH5α* plasmid DNA preparation.

7.5 Oligomerization of Get1 and Get2/Get1

In co-expression experiment Get1-strep was co-expressed with Get1-10xHis in one hand. In another one Get2/Get1-strep single chain was co-expressed with Get2/Get1-10xHis single chain. The resulting complexes were solubilized and purified in Ni-NTA affinity chromatography and identified by anti-strep antibody (Figure 7.3). This result shows that Get1 and Get2/Get1 single chain oligomerizes, probably in dimer as the LILBID confirmed.

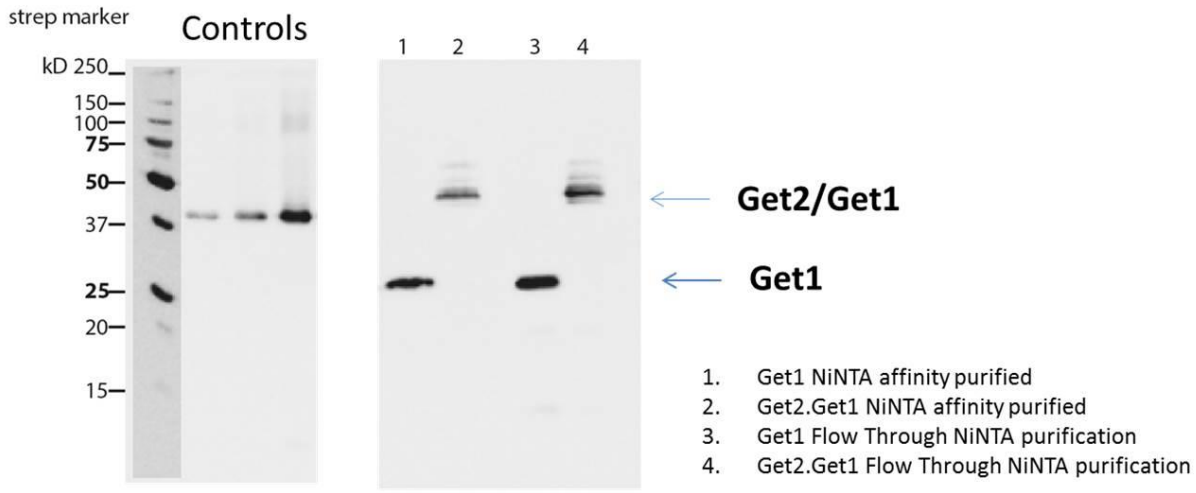


Figure 7.5: Co-expression Get1.strep with Get1.10xHis and Get2/Get1.strep with Get2/Get1.10xHis. The co-expressions were done using CEN.PK258.1D yeast cells and pull-down of co-expressed Get proteins via IMAC (Ni-NTA affinity) followed western blot with anti-strep identification.

7.6 Complex Get2/Get1 by GraFix

GraFix analysis was performed to reduce the heterogeneity of Get2/Get1/Get3 complex. The complex was first purified and thereafter the concentration to be used was initially tested with super reactive crosslinker, glutaraldehyde then different crosslinkers were tested prior Grafix (Figure 7.6.1). The complex was cross-linked within 5 minutes with glutaraldehyde, while the DMA, DMP and DMS started cross-linking only after 60 minutes. Amongst the water soluble homobifunctional imidoester crosslinkers DMP seems to be the best. Therefore glutaradehyde and DMP were selected for Grafix study. Grafix of Get2/Get1/Get1 complex shows two different fractionation profile by measuring the optical density (Figure 7.6.2) where the amphipol preparation seems to be less stable complex. Then a less reactive crosslinker, DMP, was used to study the complex with Grafix. The experiment shows that we have different species of the complex of different molecular weights using DMP crosslinker.

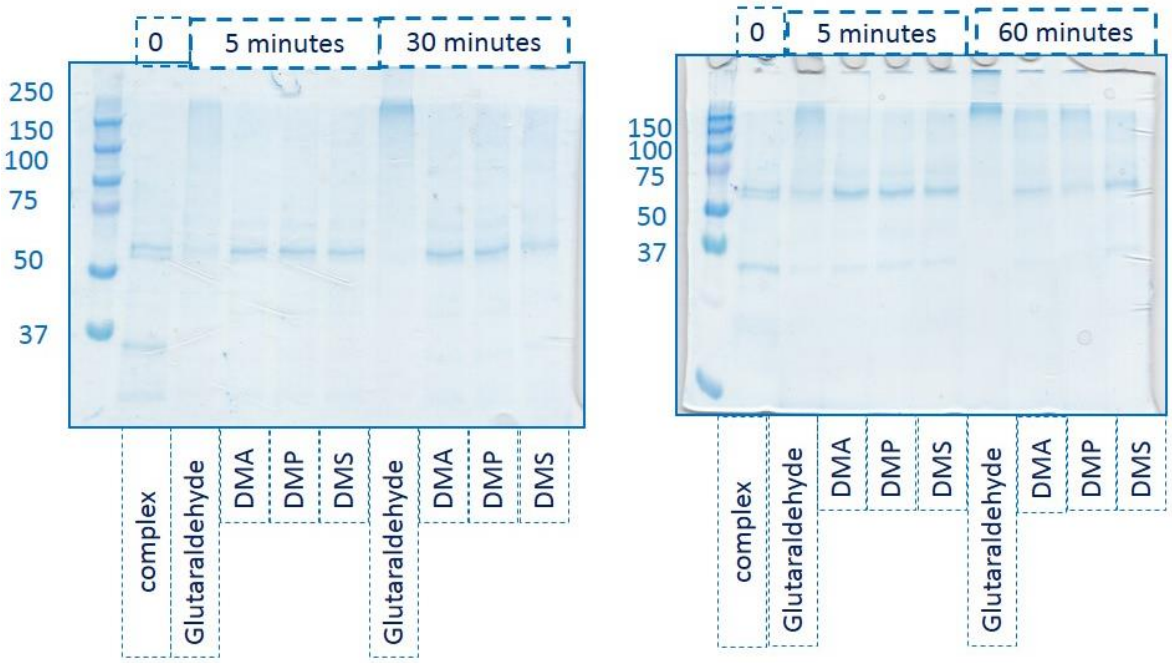


Figure 7.6.1: Crosslinking analysis. After 5 minutes 0.02% glutaraldehyde is very reactive. DMA, DMP and DMS start reacting only after 60 minutes.

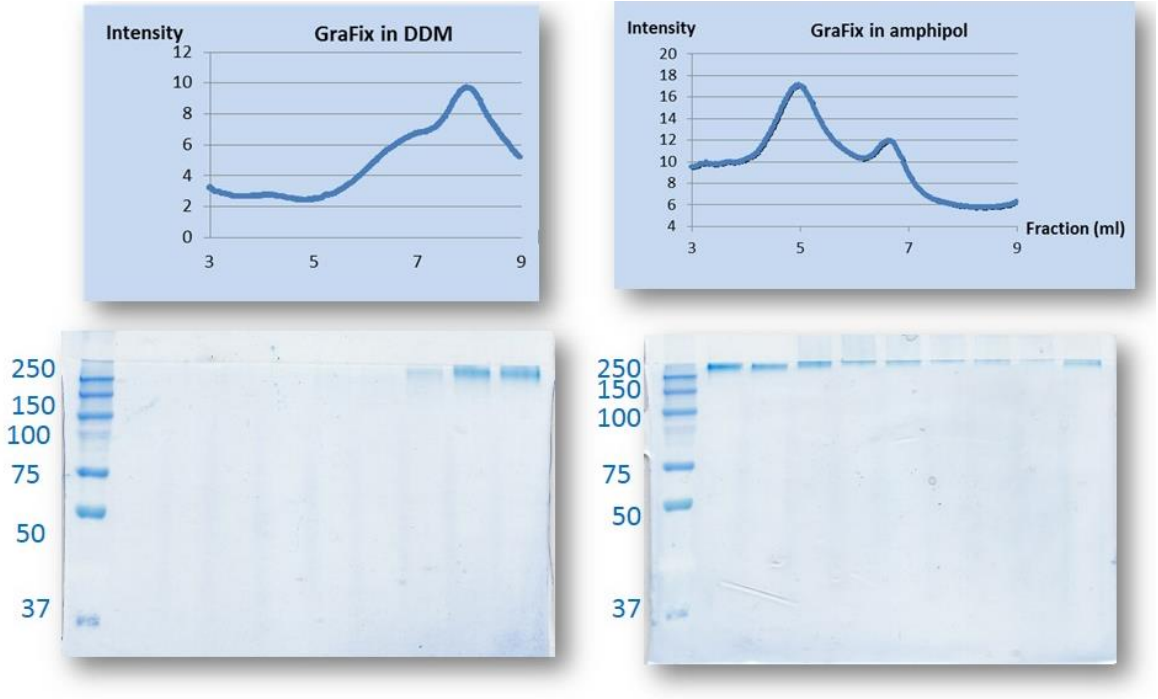


Figure 7.6.2: Grafix using 0.02% glutaraldehyde when the complex is DDM or reconstituted in amphipol.

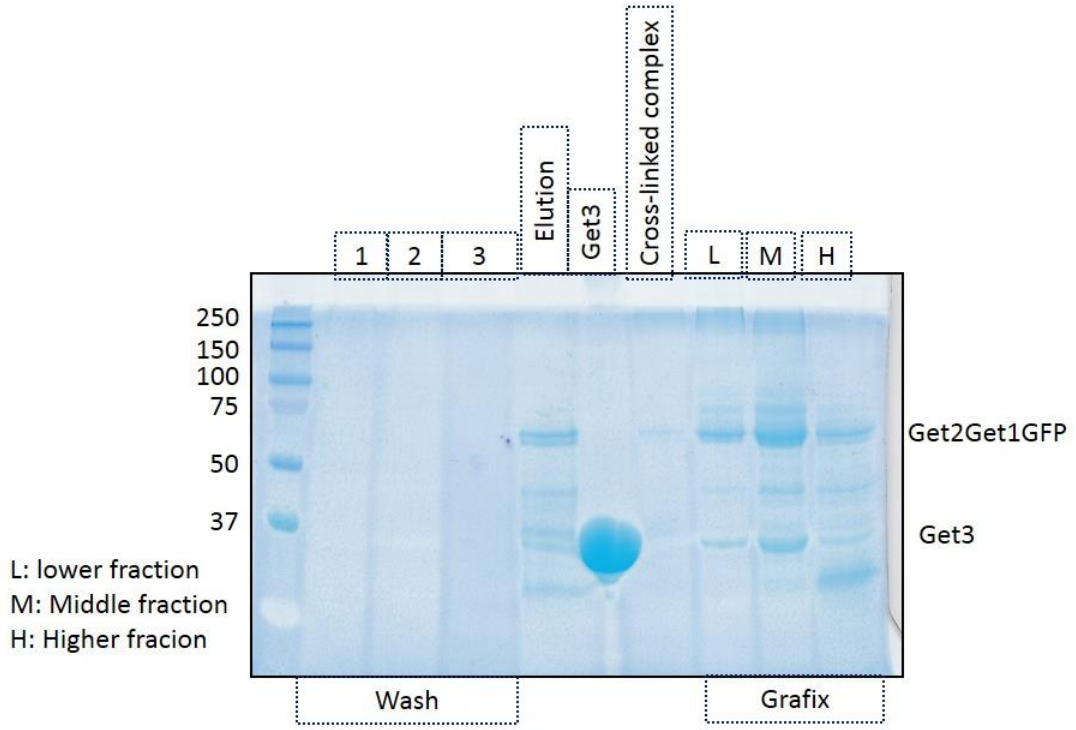
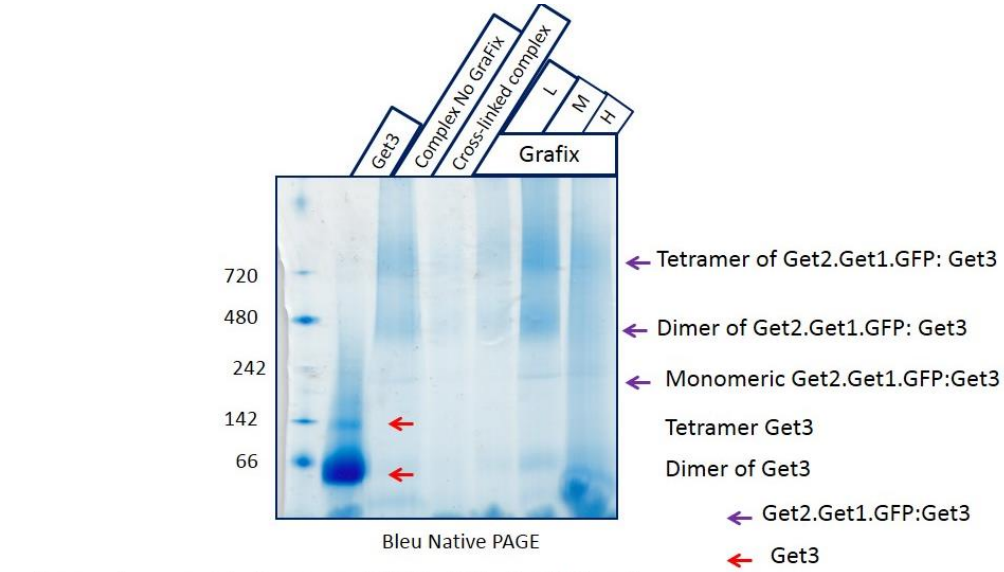


Figure 7.6.3: GraFix using DMP crosslinker SDS PAGE



For membrane protein the apparent MW is divided by 1.8 therefore:

- Monomeric Get2.Get1.GFP:Get3= **128 KDa**
- Dimeric Get2.Get1.GFP:Get3=**256 KDa**
- Tetrameric Get2.Get1.GFP:Get3= **512 KDa**
- Dimeric Get3: **78 KDa**

Figure 7.6.4: GraFix using DMP crosslinker by native page**.** The oligomeric nomination here is based on single chain construct. Monomeric here will simply be a monomer of each protein. The apparent molecular (observed) is divided by 1.8 to have the correct molecular weight for membrane protein.

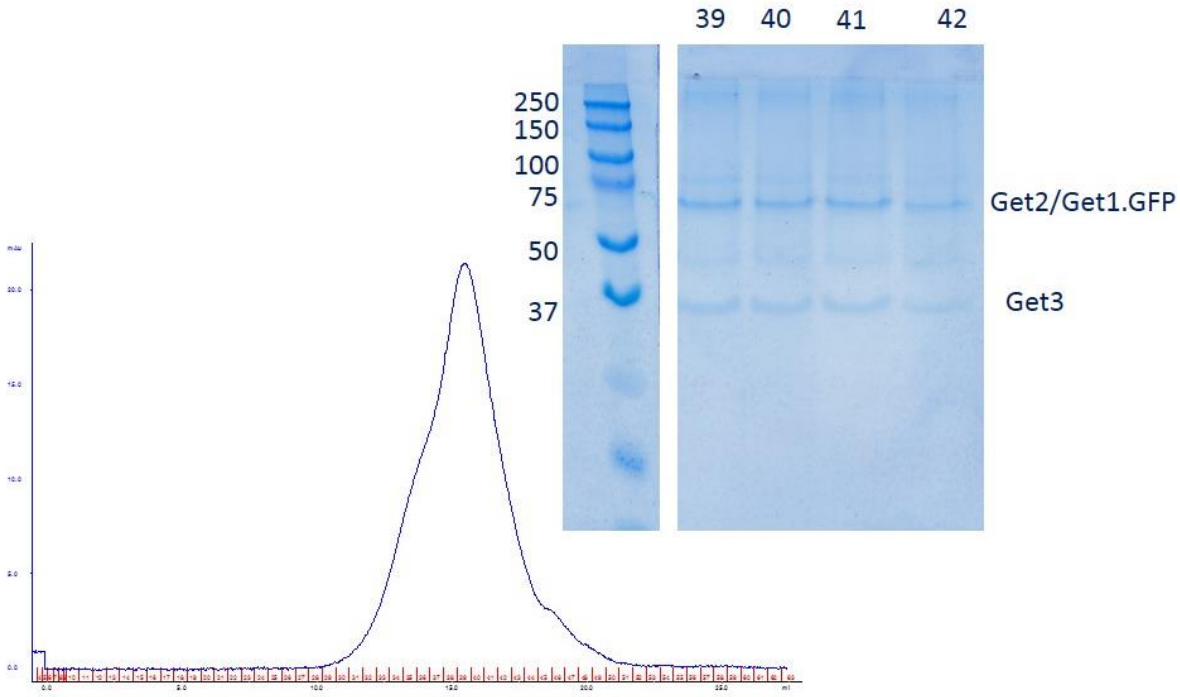


Figure 7.6.5. SEC of the Grafix middle fraction of Get2/Get1.GFP**.** It seems that despite the GraFix and higher molecular weight shown on the native gel the GraFix did not crosslink Get2/Get1.GFP to higher oligomeric state.

7.7 ATP Determination in Presence of Get2/Get1 in solution

As the ATP make the complex Get2/Get1/Get3 fall apart, it was measured in an independent experiment if ATP was present in solution of pure Get2/Get1, therefore likely binding to Get2/Get1 when ATP presence was identified, the assay for the same amount of proteins that the wild type Get2/Get1 single chain construct had more ATP present compared to the T4l.tGet2.apocyte.Get1, the effect was more pronounce in presence 1 mM calcium but reduced in presence of 10 mM hydroxylamine solution (H2NO).

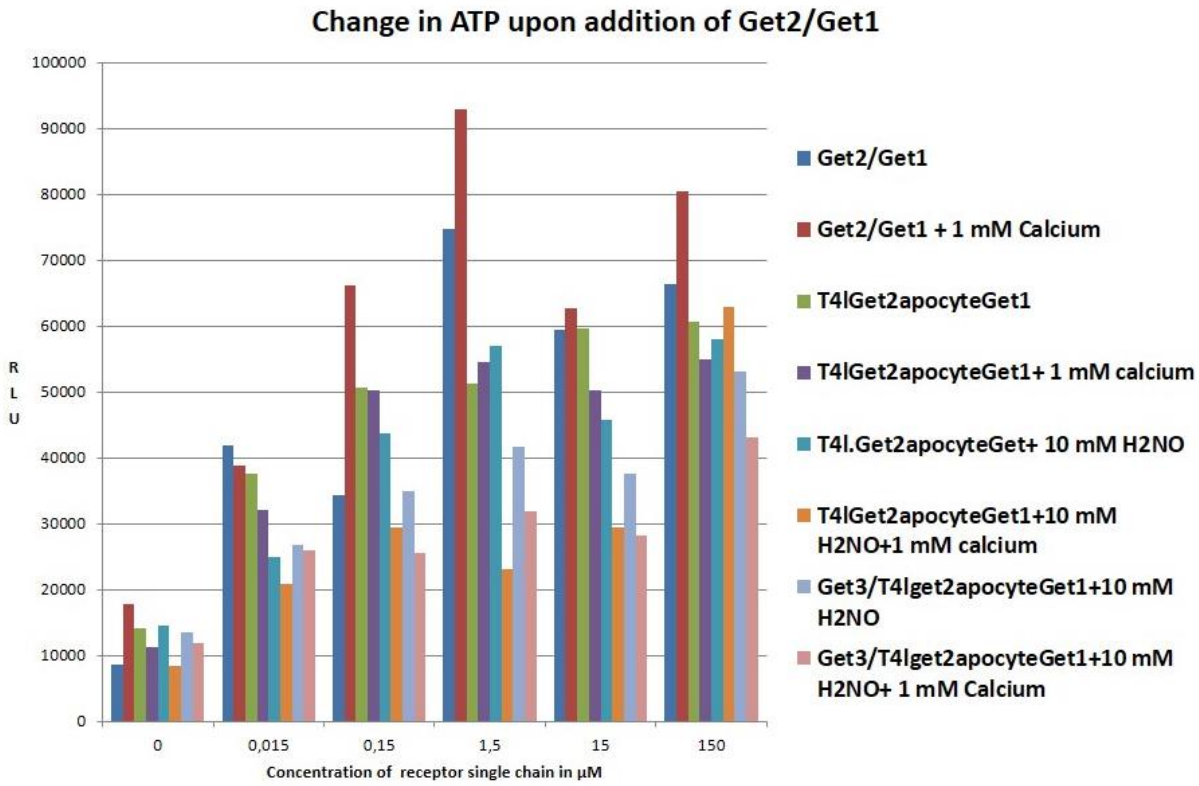
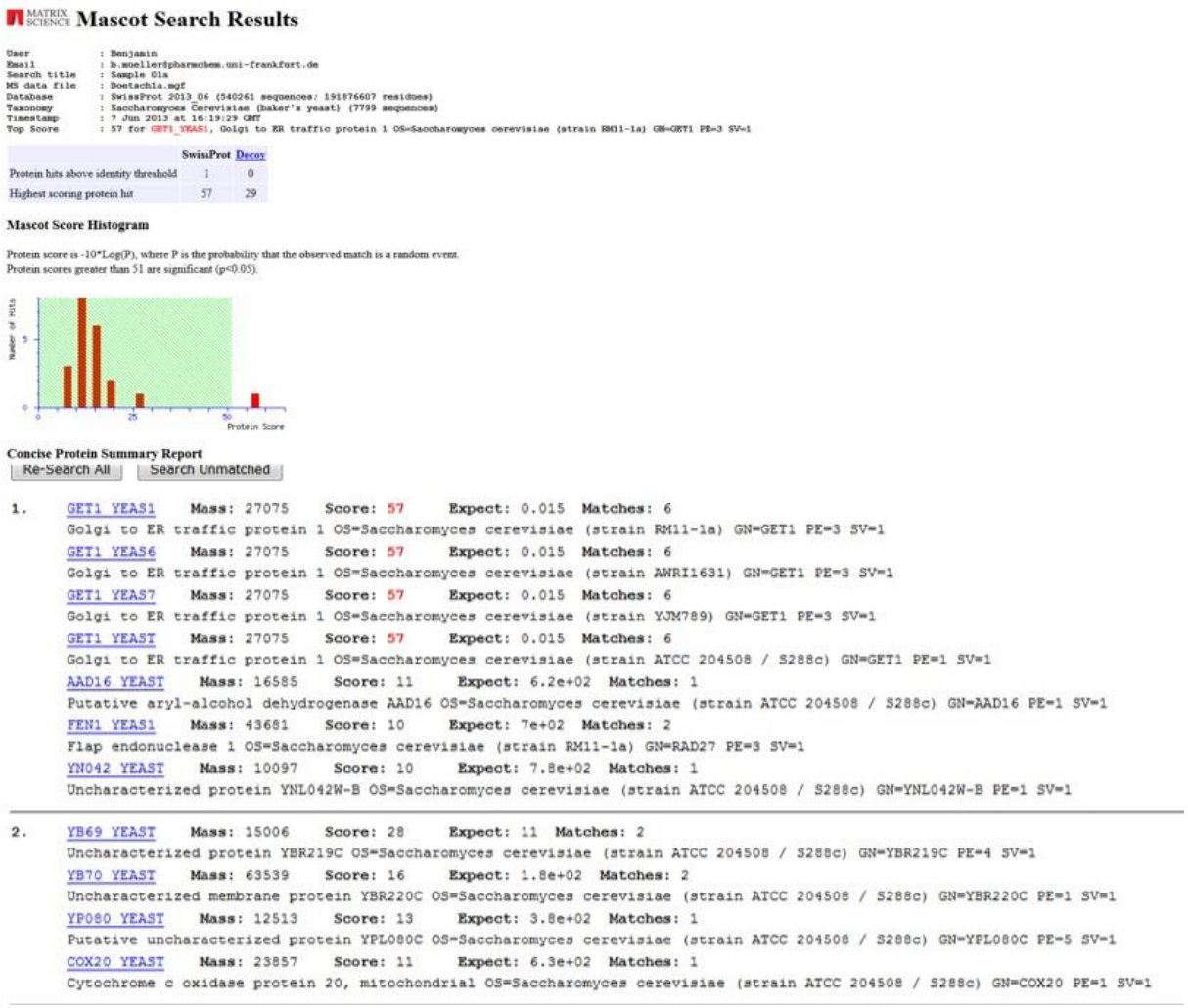


Figure 7.7: ATP determination in Get2/Get1. ATP was quantified in of the single chain that has the full Get2 and Get1 or in presence the T4l.tGet2.apocyte.Get1. All the buffer were ATP free. And the concentration of the proteins were progressively improved.

7.8 Mass Spectrometry

Although during the course of my thesis I have identified a lot of SDS-PAGE gel bands by mass spectrometry, hereby I present a screenshot of the relevant for the results I had. I had three independent center as described in the methods to identify Get proteins, we also found some that might be relevant. .



MATRIX SCIENCE **Mascot Search Results**

```
User            : Benjamin
Email           : b.moeller@pharmchem.uni-frankfurt.de
Search title    : Sample G1a
MS data file    : Doetach1a.mgf
Database        : SwissProt 2013_06 (540261 sequences; 191876607 residues)
Taxonomy        : Saccharomyces Cerevisiae (baker's yeast) (7799 sequences)
Timestamp       : 7 Jun 2013 at 16:19:29 GMT
Top Score       : 57 for GET1_YEAS1, Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain RM11-1a) GN=GET1 PE=3 SV=1
```

	SwissProt	Decoy
Protein hits above identity threshold	1	0
Highest scoring protein hit	57	29

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 51 are significant (p<0.05).

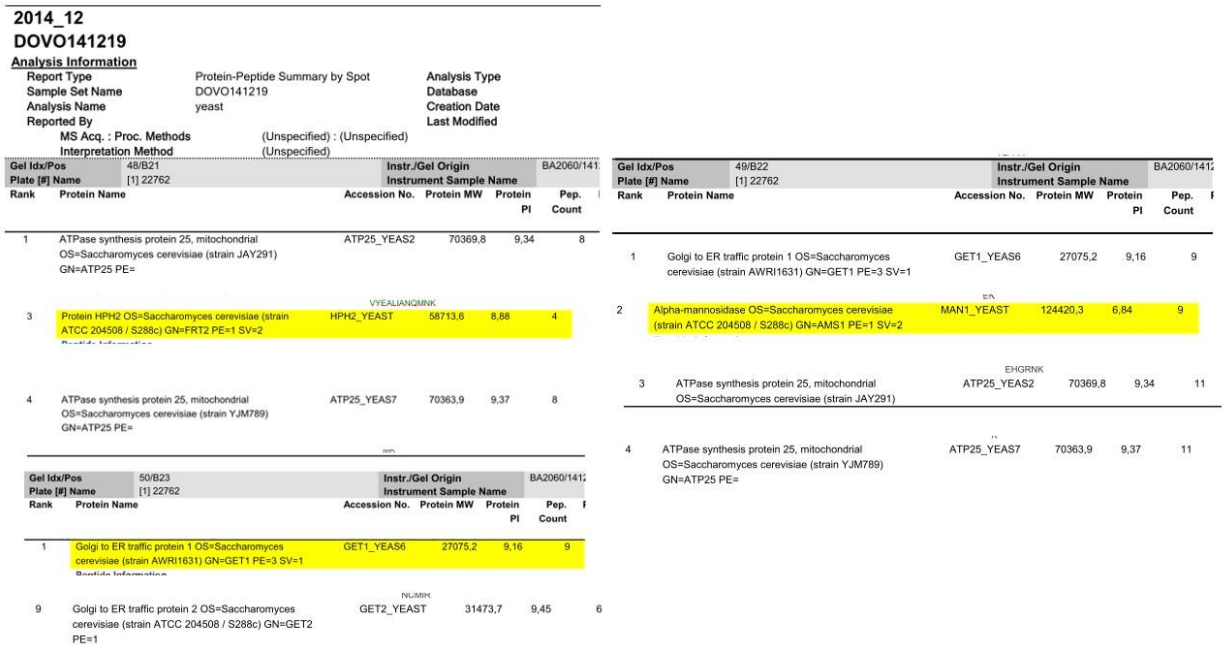
Concise Protein Summary Report

Re-Search All | Search Unmatched

```
1.  GET1_YEAS1   Mass: 27075    Score: 57    Expect: 0.015  Matches: 6
    Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain RM11-1a) GN=GET1 PE=3 SV=1
    GET1_YEAS6   Mass: 27075    Score: 57    Expect: 0.015  Matches: 6
    Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain AWRI1631) GN=GET1 PE=3 SV=1
    GET1_YEAS7   Mass: 27075    Score: 57    Expect: 0.015  Matches: 6
    Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain YJM789) GN=GET1 PE=3 SV=1
    GET1_YEAST   Mass: 27075    Score: 57    Expect: 0.015  Matches: 6
    Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=GET1 PE=1 SV=1
    AAD16_YEAST  Mass: 16585    Score: 11    Expect: 6.2e+02  Matches: 1
    Putative aryl-alcohol dehydrogenase AAD16 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=AAD16 PE=1 SV=1
    FEN1_YEAS1   Mass: 43681    Score: 10    Expect: 7e+02  Matches: 2
    Flap endonuclease 1 OS=Saccharomyces cerevisiae (strain RM11-1a) GN=RAD27 PE=3 SV=1
    YN042_YEAST  Mass: 10097    Score: 10    Expect: 7.8e+02  Matches: 1
    Uncharacterized protein YNL042W-B OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YNL042W-B PE=1 SV=1

2.  YB69_YEAST   Mass: 15006    Score: 28    Expect: 11  Matches: 2
    Uncharacterized protein YBR219C OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YBR219C PE=4 SV=1
    YB70_YEAST   Mass: 63539    Score: 16    Expect: 1.8e+02  Matches: 2
    Uncharacterized membrane protein YBR220C OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YBR220C PE=1 SV=1
    YP080_YEAST  Mass: 12513    Score: 13    Expect: 3.8e+02  Matches: 1
    Putative uncharacterized protein YPL080C OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=YPL080C PE=5 SV=1
    COX20_YEAST  Mass: 23857    Score: 11    Expect: 6.3e+02  Matches: 1
    Cytochrome c oxidase protein 20, mitochondrial OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=COX20 PE=1 SV=1
```

Figure 7.8 A: Mass spectrometry of Get1 purified at the research Group of Prof. Dr. M. Karas (Frankfurt university). Get1 was identified as well as a mitochondrial cytochrome c oxidase protein 20 (COX20).



2014_12
DOVO141219

Analysis Information

Report Type	Protein-Peptide Summary by Spot	Analysis Type	
Sample Set Name	DOVO141219	Database	
Analysis Name	yeast	Creation Date	
Reported By		Last Modified	
MS Acq. : Proc. Methods	(Unspecified) : (Unspecified)		
Interpretation Method	(Unspecified)		

Gel Idx/Pos 48/B21 — Instr./Gel Origin — BA2060/141
Plate [#] Name [1] 22762 — Instrument Sample Name

Rank	Protein Name	Accession No.	Protein MW	Protein PI	Pep. Count
1	ATPase synthesis protein 25, mitochondrial OS=Saccharomyces cerevisiae (strain JAY291) GN=ATP25 PE=	ATP25_YEAS2	70369,8	9,34	8
3	Protein HPH2 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=FRT2 PE=1 SV=2	HPH2_YEAST	58713,6	8,88	4
4	ATPase synthesis protein 25, mitochondrial OS=Saccharomyces cerevisiae (strain YJM789) GN=ATP25 PE=	ATP25_YEAS7	70363,9	9,37	8

Gel Idx/Pos 50/B23 — Instr./Gel Origin — BA2060/141
Plate [#] Name [1] 22762 — Instrument Sample Name

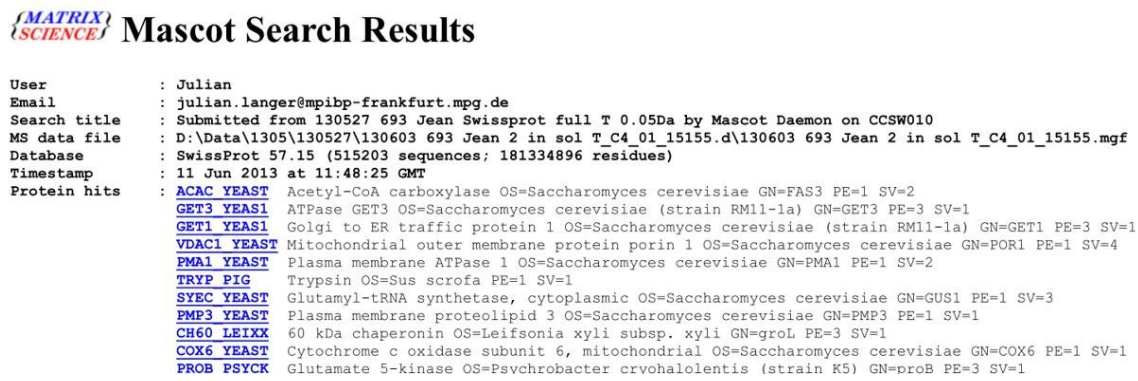
Rank	Protein Name	Accession No.	Protein MW	Protein PI	Pep. Count
1	Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain AWRI1631) GN=GET1 PE=3 SV=1	GET1_YEAS6	27075,2	9,16	9
9	Golgi to ER traffic protein 2 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=GET2 PE=1	GET2_YEAST	31473,7	9,45	6

Gel Idx/Pos 49/B22 — Instr./Gel Origin — BA2060/141
Plate [#] Name [1] 22762 — Instrument Sample Name

Rank	Protein Name	Accession No.	Protein MW	Protein PI	Pep. Count
1	Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain AWRI1631) GN=GET1 PE=3 SV=1	GET1_YEAS6	27075,2	9,16	9
2	Alpha-mannosidase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=AMS1 PE=1 SV=2	MAN1_YEAST	124420,3	6,84	9
3	ATPase synthesis protein 25, mitochondrial OS=Saccharomyces cerevisiae (strain JAY291)	ATP25_YEAS2	70369,8	9,34	11
4	ATPase synthesis protein 25, mitochondrial OS=Saccharomyces cerevisiae (strain YJM789) GN=ATP25 PE=	ATP25_YEAS7	70363,9	9,37	11

Figure 7.8 B: Mass spectrometry of Get2/Get1 at applied biomics.

After Get2/Get1 were purified, all bands were send for identification. Get2/Get1, HPH2 (FRT2), and alpha mannosidase were identified in significant manner but ATP synthesis protein 25 mitochondrial was often present in all the sample send.



MATRIX SCIENCE **Mascot Search Results**

```
User          : Julian
Email         : julian.langer@mpibp-frankfurt.mpg.de
Search title  : Submitted from 130527 693 Jean Swissprot full T 0.05Da by Mascot Daemon on CCSW010
MS data file  : D:\Data\1305\130527\130603 693 Jean 2 in sol T_C4_01_15155.d\130603 693 Jean 2 in sol T_C4_01_15155.mgf
Database      : SwissProt 57.15 (515203 sequences; 181334896 residues)
Timestamp     : 11 Jun 2013 at 11:48:25 GMT
Protein hits  : ACAC_YEAST  Acetyl-CoA carboxylase OS=Saccharomyces cerevisiae GN=FAS3 PE=1 SV=2
                GET3_YEAS1  ATPase GET3 OS=Saccharomyces cerevisiae (strain RM11-1a) GN=GET3 PE=3 SV=1
                GET1_YEAS1  Golgi to ER traffic protein 1 OS=Saccharomyces cerevisiae (strain RM11-1a) GN=GET1 PE=3 SV=1
                VDAC1_YEAST Mitochondrial outer membrane protein porin 1 OS=Saccharomyces cerevisiae GN=POR1 PE=1 SV=4
                PMA1_YEAST  Plasma membrane ATPase 1 OS=Saccharomyces cerevisiae GN=PMA1 PE=1 SV=2
                TRYP_PIG    Trypsin OS=Sus scrofa PE=1 SV=1
                SYEC_YEAST  Glutamyl-tRNA synthetase, cytoplasmic OS=Saccharomyces cerevisiae GN=GUS1 PE=1 SV=3
                PMP3_YEAST  Plasma membrane proteolipid 3 OS=Saccharomyces cerevisiae GN=PMP3 PE=1 SV=1
                CH60_LEIXX  60 kDa chaperonin OS=Leifsonia xyli subsp. xyli GN=groL PE=3 SV=1
                COX6_YEAST  Cytochrome c oxidase subunit 6, mitochondrial OS=Saccharomyces cerevisiae GN=COX6 PE=1 SV=1
                PROB_PSYCK  Glutamate 5-kinase OS=Psychrobacter cryohalolentis (strain K5) GN=proB PE=3 SV=1
```

Figure 7.8 C: Mass spectrometry of Get2/Get1/Get3 complex purified at Max-Planck for Biophysics Frankfurt. The complex was identified as well as a mitochondrial cytochrome c oxidase subunit 6 (COX6).

7.9 Heterogeneous complex of Get1/Get2/Get3 analysis by LILBID

The analysis of the complex Get1/Get2/Get3 shows a heterogeneous mixture where it is not to assign the stoichiometry (Figure 7.9).

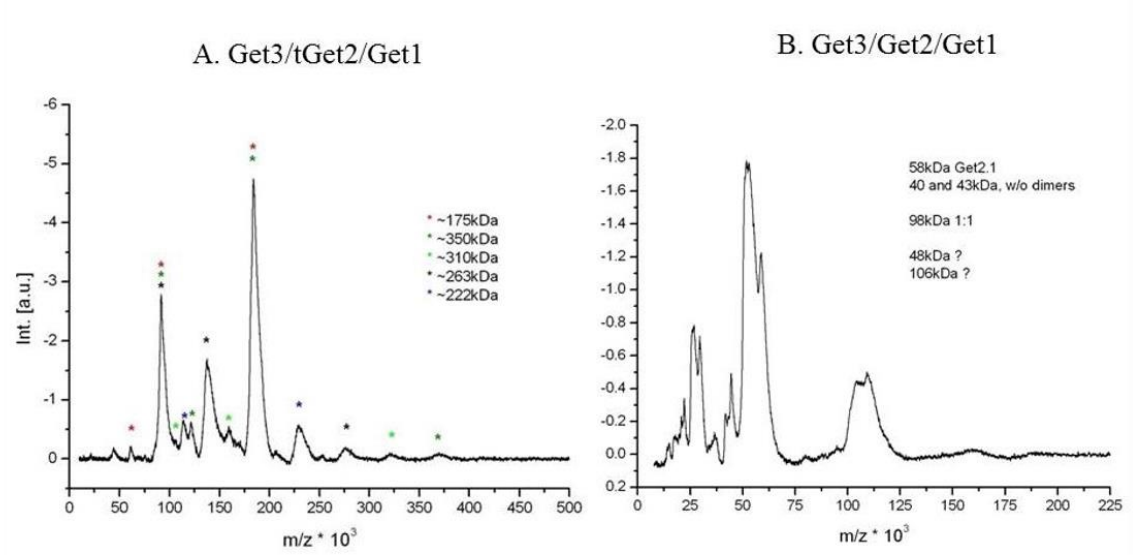


Figure 7.9: LILBID analysis of complex between Get1, Get2 and Get3.

The analysis shows that only the complex of Get3 with the tGet2/Get1 seems to have higher order oligomer while the complex of Get3 with the full length is completely destabilized and it is impossible to assign the corresponding observed peaks with any of the species of the complex.

ACKNOWLEDGEMENTS

It has been a great honour to be part of the group lead by Prof. Dr. Dötsch who has become a source of inspiration on how to lead by example.

I am grateful for this unique scientific and life experience, no word can match my gratitude. I could not have asked better than this opportunity to have completed my thesis in Dötsch group. I will be looking forward continuing the close collaboration on Get project maybe many more projects in future.

I had wonderful time working in the Dötsch group, I am still amazed by the kindness, softness of Mrs Sigrid Oğuzer-Fachinger. The kindness of Dr. Frank Bernhard, Dr. Vladimir Rogov, Mrs Natalja Rogova, and Mrs Birgit Schäfer.

I also will not forget Mr Manfred Strupf assistance regarding all issues related to the dysfunctional equipment and software, Manfred was always nice and there when I need to fix computer and technical issues. Everyone was always nice unless I am blind I cannot recall any bad day with others colleagues.

It would be unfair to omit all my fellow PhDs, PostDocs and others colleagues in my acknowledgement, Eric, Erika, Fang, both Susanne, Ralf, Rathnish, Aisha, Bea, Edith, Gregor, Peter, Floriant, Jakob, Daniel, Christopher, Katharina, Christian, Andreas, Jan, Jessica, Laura, Robert, Julia, both Sebastian, Marcel, and Dr. Frank Löhr.

As I am writing these acknowledgements, I feel like leaving once more a part of me here, it just pains me to think I will be leaving the lab but it is how life goes.

The results achieved in this work would not be possible without the close collaborations that Prof. Dr. Dötsch set for us to progress in the research.

I have in mind first Dr Peter Kötter. Peter was as good as people in our lab, although when I started the project, I did apprehend the yeast system alone to establish it in our laboratory, but Peter is helpful and knowledgeable regarding everything in yeast. Most of things I know from yeast system I know it from Peter as he fine tune what I discover by helping to simplify a lot of proceeding part. Also everyone in Prof. Dr. Entian, Karl-Dieter laboratory adopted me, as most of the time I was there, using the incubators and fermentors. I would cite Samy, the other Peter, both Stephanie and other people I can recall names but who were nonetheless nice as well.

Dr Margot Frangakis and Prof. Dr. Achilleas Frangakis who have supported me for the Cryo-EM work. Margot has also given me so much that no words or gifts can compensate it. I also learn all my Cryo-EM skills in Prof. Dr. Achilleas Frangakis. I was also adopted in Prof. Dr. Achilleas Frangakis laboratory in such way that I felt that I belong to their laboratory. For that I would also thank both Michael, Anja, Anna, Johnson, Valentin and Fabrisio.

I would thank Prof. Dr. Nina Morgner and Dr. Jan Hoffmann who were always available to measure my sample for LILBID analysis.

In the MPI my close collaboration with Dr. Yvonne Thielmann, Mrs Barbara Rathmann who was always kind and ready to set up my crystallization plates, Dr. Ulrich Ermler, Dr. Julian Langer, Prof. Dr. Hartmut Michel for allowing us to work in his laboratory.

Prof. Dr. Werner Kühlbrandt, Dr. Janet Vonck, Dr. Maria Müller for accepting to start electron microscopy work and supported to be annoy me as I was pushing things to go faster without really assessing the difficulty of the task.

I would thank Prof. Dr. Robert Tampé for allowing me to use his FatPrep-24 benchtop homogenizer and other instruments in his laboratory.

Prof. Dr. Clemens Glaubitz for not only accepting being member of my PhD committee but also because he also allowed me to use the centrifuges in his laboratory. Also thank everyone in his laboratory for their kindness.

I would thank our collaborators on the project from other universities, Prof. Dr. Irmgard Sinning, Prof. Dr. Vladimir Denic.

This list of acknowledgement is endless considering all the countless good interaction with everyone I came across on campus.

I thank SFB 807 Transport and Communication across Biological Membrane (TRAM), the Deutsche Forschungsgemeinschaft (DFG) for the material and financial support.

I will end up this note by thanking my parents, my beloved mother who has help me to strongly be grounded, my father, my bothers Junior, Martin and Paul. My great friends who have always been there during the difficult and good time those years, namely, Dr. Bivigou Koumba Achille and Dr. Stefen Schmitt.