The client spectrum of Get3, an evolutionarily conserved chaperone of membrane proteins

Dissertation

for the award of the degree

"Doctor of Philosophy"
Division of Mathematics and Natural Sciences
of the Georg-August-Universität Göttingen

within the doctoral program Molecular Biology

of the Georg-August University School of Science (GAUSS)

submitted by

Ákos Farkas
from Budapest, Hungary

Göttingen 2019

Members of the Thesis Committee:

Prof. Dr. Blanche Schwappach-Pignataro (1st reviewer)

Department of Molecular Biology, University Medical Center Göttingen

Dr. Alexander Stein (2nd reviewer)

Research Group of Membrane Protein Biochemistry, Max Planck Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Peter Rehling

Department of Cellular Biochemistry, University Medical Center Göttingen

Further members of the examination board:

Dr. Alexis Caspar Faesen

Research Group of Biochemistry of Signal Dynamics, Max Planck Institute for Biophysical Chemistry, Göttingen

Dr. Hans Dieter Schmitt

Department of Neurobiology, Max Planck Institute for Biophysical Chemistry, Göttingen

Prof. Dr. Tobias Moser

Institute for Auditory Neuroscience, University Medical Center Göttingen

Date of oral examination: 20th November 2019

Affidavit

Herewith I declare that I prepared the PhD dissertation 'The client spectrum of Get3, an evolutionarily conserved chaperone of membrane proteins' on my own and with no other sources and aids than quoted.

Göttingen, September 2019

Ákos Farkas

Publications associated with the dissertation

Farkas, Á., De Laurentiis, E.I., and Schwappach, B. (2019). The natural history of Get3-like chaperones. Traffic *20*, 311–324.

Table of contents

Abstract

Tail-anchored proteins constitute a physiologically important class of membrane proteins. Due to their topology, they are targeted post-translationally to membranes. Although Get3 is responsible for the membrane targeting of several ER-bound TA proteins, our knowledge on its *in-vivo* client spectrum remains limited. The oxidation-induced chaperone activity of Get3 detected *in-vitro* suggests that Get3 may handle further substrates besides TA proteins. My bioinformatic analysis shows that Get3-like chaperones are more widespread than previously expected and suggests that a general chaperone activity is the ancestral function of Get3. Using budding yeast as a model system, I show that although Get3 can target many TA proteins, only a fraction of them are obligate substrates of Get3. I provide evidence that an essential TA protein, Sed5, is a super-client of Get3 that interacts with it before and also after membrane targeting. I also show that Get3 can target proteins with multiple transmembrane segments located proximal to their C-terminus as part of the GET pathway. Finally, I describe membrane proteins with amphipathic helices as potential chaperone substrates of Get3. Therefore, the current study significantly expands the range of potential Get3 clients *in vivo* and indicates that Get3 is an important chaperone of membrane proteins both before and after membrane targeting.

List of Abbreviations

5-FOA   5-fluoroorotic acid

AH   amphipathic helix

AID-tag small degron inducing degradation of tagged proteins in the presence of IAA

ADP   adenosine diphosphate

ATP   adenosine triphosphate

CAH   C-terminal amphipathic helix

C'-TMS   proteins with multiple TMSs close to the C-terminus

Cytb5   mammalian cytochrome b5

DNA   deoxyribonucleic acid

DTT   dithiothreitol

EDTA   ethylenediaminetetraacetic acid

ER   endoplasmic reticulum

EMC   ER membrane protein complex

FDR   false discovery rate

GET   guided entry of TA proteins

GFP   green fluorescent protein

GPI   glycosylphosphatidylinositol

HEPES   4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

IAA   indole-3-acetic acid, an auxin

LiAc   lithium acetate

MS   mass spectrometry

NADH   protonated form of nicotinamide adenine dinucleotide

NAH   N-terminal amphipathic helix

NBS	nucleotide binding site
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
*Sc*Get3	budding yeast Get3
SDS	sodium dodecyl sulfate
SGA	synthetic genetic analysis
Sgt2	small glutamine-rich tetratricopeptide repeat-containing protein 2
SIMIBI	Signal recognition particle, MinD, BioD
SNARE	soluble N-ethylmaleimide-sensitive-factor attachment receptor
SND	SRP-independent targeting
SR	SRP receptor
SRP	signal recognition particle
StageTip	stop-and-go extraction tip
TEV	Tobacco Etch Virus
TM	transmembrane
TMS	transmembrane segment
TPR	tetratricopeptide repeat
TRIS	tris(hydroxymethyl)aminomethane
TA	tail-anchored
U	unit

List of Figures, Supplementary Figures and Tables

Chapter One - Introduction

Chapter Two – Manuscript I

Chapter Three – Manuscript II

Chapter One

Introduction

1. Classes of membrane proteins

A uniform feature of all living cells is the presence of a lipid bilayer, the cell membrane separating the internal cytoplasm from the external environment. Similarly, all living cells produce proteins that are inserted into or stay associated with the cell membrane, but the way membrane proteins are anchored to the membrane can be quite diverse (Fig 1).

1.1. Integral membrane protein

Integral membrane proteins embed themselves into the hydrophobic portion of the lipid bilayer with the help of hydrophobic amino acids. Some proteins contain amphipathic helices with a sufficiently hydrophobic patch which allows them to contact the internal hydrophobic environment of the membrane without passing through it completely (Giménez-Andrés et al., 2018). An example for these monotopic membrane proteins is provided by the mitochondrial carnitine palmitoyltransferase II, which plays a role in the beta-oxidation of fatty acids (Allen et al., 2019; Hsiao et al., 2006)

Most integral membrane proteins, however, contain one or more membrane spanning transmembrane segments (TMS), which are hydrophobic enough to bypass the lipid bilayer. Since TMSs are prone to aggregate in aqueous environments like the cytosol, most of them require chaperones that keep them soluble and deliver them for insertion into the target membrane. There are several pathways that ensure the proper targeting of membrane proteins and many of these are evolutionarily highly conserved.

A special class of integral membrane proteins are tail-anchored (TA) proteins. They are characterized by a single TMS at the C-terminus with their N-terminal domain facing the cytosol. To classify as a TA protein, the number of amino acid residues after the TMS should not exceed approximately 30 amino acids (Borgese, 2003; Kutay et al., 1993), although the protein with the longest luminal sequence that is still considered a TA protein is Sec20, which has approximately 90 amino acid residues after the TMS (Ast et al., 2013). Due to the TMS located at the C-terminus, translation ends before the the TMS emerges from the ribosome, and TA proteins can only be inserted into membranes post-translationally. Based on estimates, there are at least around 40 TA proteins of the

secretory pathway in yeast (Ast et al., 2013) and several hundreds in mammals (Kalbfleisch et al., 2007), making them an important category of membrane proteins.

1.2. Peripheral membrane proteins

Many membrane proteins are only peripherally associated with the membrane and do not form an integral part of it. These proteins can contain domains that can bind the polar head groups of the lipid bilayer without making contact with the internal hydrophobic core of the membrane (Moravcevic et al., 2012). Alternatively, peripheral membrane proteins can bind other integral membrane proteins which act as membrane anchors for them. For instance, several subunits of the evolutionarily conserved vacuolar ATPase complex only associate with the vacuolar membrane via binding integral membrane components of the complex (Cotter et al., 2015).

1.3. Lipid-anchored membrane proteins

The addition of a lipid moiety to a protein can provide an anchor for membrane association, a prominent example for which is prenylation (Wang and Casey, 2016). Prenylation can happen when a protein carries a CaaX sequence at its C-terminus, where "a" stands for aliphatic amino acids. This signal is recognized by prenyltransferases which transfer either a farnesyl or a geranyl-geranyl moiety depending on the specific CaaX sequence present (Caplin et al., 1994). For instance, the yeast soluble N-ethylmaleimide-sensitive-factor attachment receptor (SNARE) protein Ykt6 has a C-terminal CaaX motif, CIIM in this case, which results in the addition of a farnesyl moiety (McNew et al., 1997).

Some secreted proteins remain membrane attached through the processing of their C-terminus resulting in the addition of a glycolipid moiety in its place. This glycosylphosphatidylinositol (GPI) anchor is added once the protein has translocated through the endoplasmic reticulum (ER), and is therefore exclusively found on extracellular proteins (Orlean and Menon, 2007). A unique feature of these proteins is the presence of a signal sequence at the N-terminus which allows translocation into the ER lumen, and a variably hydrophobic sequence at the C-terminus acting as a signal for the addition of the GPI anchor (Orlean and Menon, 2007).
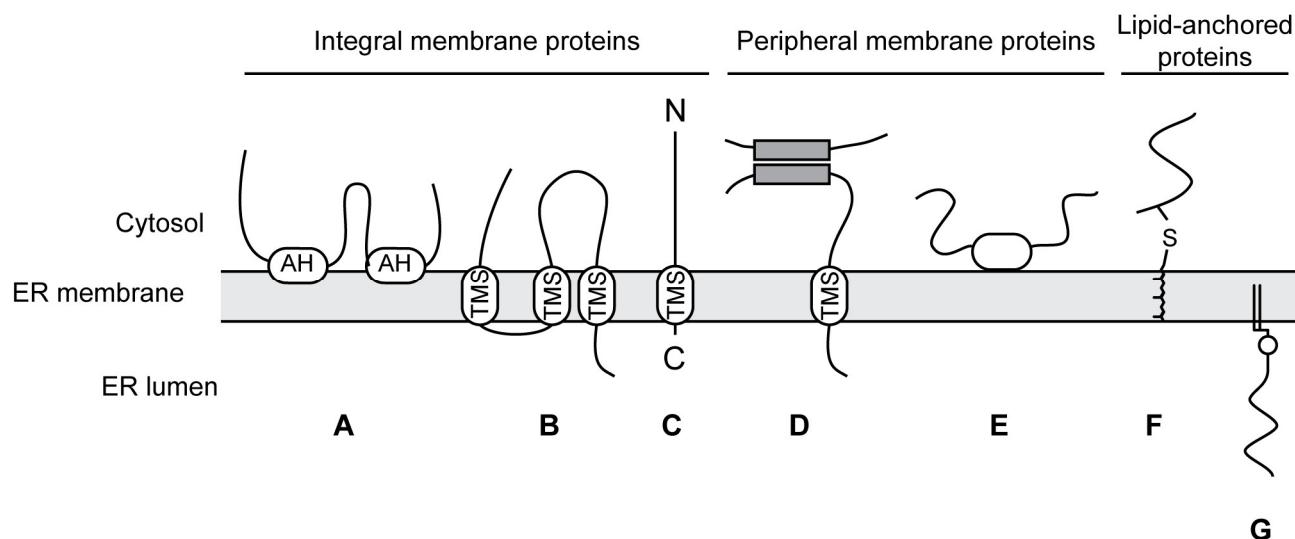
Figure 1: *Illustration of the different classes of membrane proteins. A. Monotopic integral membrane protein with amphipathic helices (AH). B. Integral membrane protein with TMSs. C. TA protein. D. Peripheral membrane protein binding an integral membrane protein. E. Peripheral membrane protein with a domain binding polar head groups. F. Farnesylated protein. G. GPI-anchored protein.*

2. Targeting pathways of membrane proteins

In eukaryotic cells, membrane proteins can be inserted either into the ER, mitochondria (Harbauer et al., 2014), peroxisomes (Mayerhofer, 2016), or chloroplasts in plants (Lee et al., 2014). There are several pathways that exist for delivering membrane proteins to the ER, and each of them cater to a different yet partially overlapping range of substrates. From the ER, membrane proteins can reach all compartments of the secretory pathway, the plasma membrane as well as the endolysosomal system.

2.1. SRP pathway

Proteins that carry a short N-terminal hydrophobic sequence, called the signal sequence, are recognized co-translationally by the signal recognition particle (SRP). This ribonucleoprotein complex delivers its substrates to its membrane receptor (SR) and hands over the translating amino acid chain to the Sec61 translocon, which acts as a channel into the ER lumen for the newly synthesized protein. Remarkably, the pathway is evolutionarily conserved even in bacteria. The main protein component of the SRP, Srp54 in eukaryotes and Ffh in bacteria, both use their methionine-rich M-domain to bind the emerging hydrophobic signal sequence. Besides the SRP, the translocon channel is

also conserved in bacteria where its homolog, SecY, helps insert proteins into the bacterial plasma membrane (Akopian et al., 2013).

Recent results suggest that many proteins with N-terminal signal sequences are targeted efficiently in the absence of SRP, whereas other proteins with TMSs, irrespective of where they are located in the protein, depend on SRP for their targeting (Costa et al., 2018). This highlights the need to experimentally verify the biogenesis pathway membrane proteins use in each individual case, as there seem to be several exceptions to the general rules associated with the different pathways.

2.2. The posttranslational translocon

Signal-sequence bearing proteins that the SRP fails to recognize before translation finishes have to be inserted into the membrane post-translationally. In this case, cytosolic chaperones recognize the hydrophobic signal sequence and deliver it to the Sec61 translocon. However, since the ribosome and the ongoing translation does not push the protein through the translocon, a force is required to pull the nascent chain through the translocon. This is provided by BiP, an ATPase and chaperone in the ER lumen that acts as a "ratchet" to pull the protein through the translocon (Rapoport, 2007).

2.3. SND pathway

Proteins that lack a signal sequence and contain a TMS only in their middle part utilize a different pathway to reach the membrane. Although the mechanistic details remain unknown, three proteins, Snd1, Snd2, and Snd3 seem to cooperate to ensure the SRP- independent targeting (SND) of these proteins (Aviram et al., 2016). While Snd1 is cytosolic and interacts with the ribosome, Snd2 and Snd3 are integral membrane proteins that interact with the Sec61 translocon. This pathway displays a broad substrate specificity and is able to compensate for the loss of SRP in vivo and may even recognize proteins with C-terminal TMSs. The SND system has been so far characterized only in yeast, and only one of the three proteins have been found in mammals (Haßdenteufel et al., 2017). Nevertheless, its ability to acts as a backup for other targeting pathways thanks to its broad substrate range makes it an important player to consider when interpreting results from the manipulation of other targeting pathways (Aviram et al., 2016).

2.4. EMC pathway

The ER membrane protein complex (EMC) is an evolutionarily conserved complex of the ER membrane that can insert TA proteins with low TMS hydrophobicity into the ER (Guna et al., 2018). As an additional function, it also helps orient the first TMS of polytopic membrane proteins and thus supports membrane protein integration through the Sec61 translocon (Chitwood et al., 2018; Shurtleff et al., 2018). However, since all studies so far are from mammalian systems, it is not known whether it has the same functions in yeast as well.

2.5. GET pathway

The Guided Entry of TA proteins (GET) pathway consisting of at least six proteins in yeast ensures the proper targeting of many TA proteins. In short, TA proteins are first recognized by a chaperone, the Small Glutamine-rich Tetratricopeptide repeat-containing protein 2 (Sgt2), then handed over to another chaperone, Get3 with the help of a complex made up of Get4/Get5 acting as a bridge. Get3 then delivers TA proteins to its receptor complex consisting of Get1/Get2 at the ER, which inserts the TA proteins into the ER membrane (Mateja and Keenan, 2018).

2.6. Unassisted insertion

Interestingly, not all membrane proteins seem to need assistance to insert into membranes. The mammalian TA protein cytochrome b5 (Cytb5) has been found to be able to spontaneously insert into protein-free liposomes (Brambillasca et al., 2005) and rough microsomes without TRC40, the mammalian homolog of Get3 (Stefanovic and Hegde, 2007). Nonetheless, considering that other pathways can also insert TA proteins as described above, we cannot exclude the possibility that Cytb5 uses another pathway for its insertion as the low hydrophobicity of its TMS could make it an ideal substrate of the EMC pathway.

3. The GET pathway

As noted above, the GET pathway is a major route TA proteins can take to reach the ER. Most of its elements are conserved between different eukaryotes and thanks to available structures of most involved proteins, the mechanistic details of the pathway are well established once a TA protein is committed to it (Fig 2).
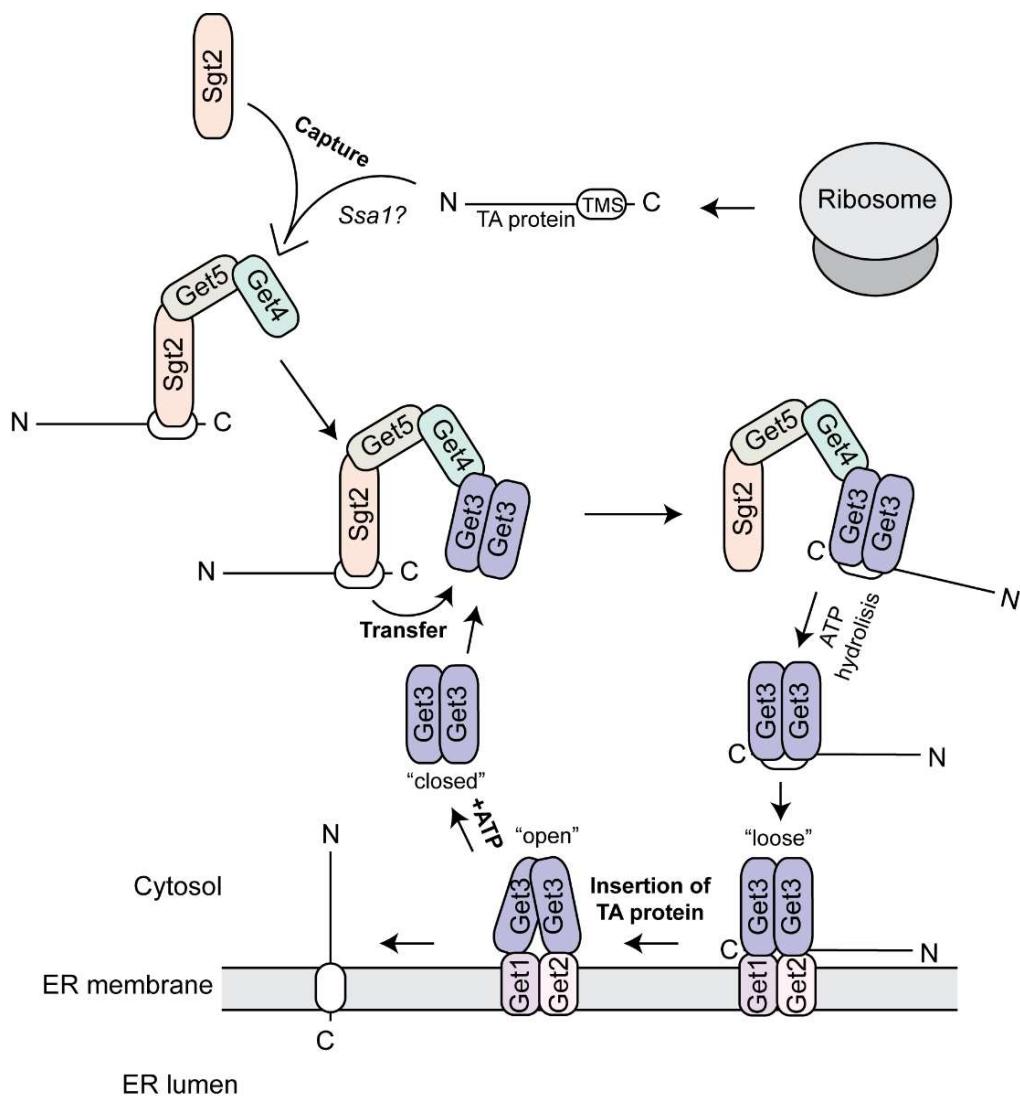
Figure 2: Overview of the GET pathway. *After translation, TA proteins are probably captured by Ssa1, then handed over to Sgt2. TA-bound Sgt2 associates with Get4-Get5, which makes the transfer of the TA protein from Sgt2 to Get3 in closed conformation possible. After the TA protein has been transferred to Get3, Get3 dissociates from the pretargeting complex, and contacts its receptor complex at the ER consisting of Get1-Get2. By this point, ATP has been hydrolyzed by Get3, and its closed conformation starts to loosen. The TA protein is inserted into the membrane, Get3 assumes its open confirmation, and after dissociation of ADP and the binding of ATP, Get3 dissociates from the receptor complex and assumes its closed conformation again. Based on Mateja and Keenan, 2018. For further references, see text.*

3.1. The pretargeting complex

The initial step of the pathway involves the recognition of a substrate TA protein by Sgt2. Sgt2 has an N-terminal domain that interacts with Get5, and a C-terminal methionine-rich domain that can bind TA proteins (Wang et al., 2010). The two domains are linked by a tetratricopeptide repeat (TPR) domain, which allows Sgt2 to interact with other chaperones including Ssa1, Hsp104 and Ybr137w (Kohl et al., 2011). In line with the proposed role of the TRP domain, a recent study has indicated that Sgt2 can receive TA substrates from Ssa1 (Cho and Shan, 2018).

To hand over the TA substrate to Get3, Sgt2 contacts a complex composed of Get4 and Get5. Get5 has a ubiquitin-like domain in the middle, which allows it to interact with the N-terminus of Sgt2 (Chang et al., 2010), whereas its N-terminus is used to contact Get4. The C-terminus of Get5 allows it to homodimerize, however, the *in-vivo* relevance of this is not known (Chartron et al., 2010). Get4 in turn uses its C-terminal domain to bind Get5, and its N-terminal portion is reserved for interacting with Get3 (Chang et al., 2010; Chartron et al., 2010). Thus, the pretargeting complex composed of Sgt2, Get5, and Get4 acts a loading platform for TA substrates with a binding interface that allows Get3 to receive the TA protein for insertion.

3.2. Get3 and its ER receptor Get1, Get2

Get3 belongs structurally to the Signal recognition particle, MinD, BioD (SIMIBI) family of nucleotide triphosphate binding proteins and thus shares several conserved elements required for binding and hydrolyzing ATP (Bange and Sinning, 2013). Uniquely in its family, however, it also possesses a region that folds into a hydrophobic groove, providing an interaction surface for the TMS of TA proteins (Mateja et al., 2009). It forms homodimers in cells which are stabilized by coordinating a zinc ion with the help of a CxxC motif (Metz et al., 2006), and this homodimerization is a prerequisite for its TA protein targeting function (Mateja et al., 2009).

Its ATP binding, hydrolysis, and release cycle is intricately linked with its function as a TA biogenesis factor. Based on in vitro experiments, the nucleotide-free, i.e. apo-form of Get3 has only low binding affinity to the pretargeting complex, whereas its ATP-bound binds it strongly (Gristick et al., 2014). ATP binding also causes Get3 to switch from an "open" to a "closed" confirmation, which favour the assembly of the hydrophobic groove necessary for TA substrate binding (Bozkurt et al., 2009; Mateja et al., 2009). At

the same time, Get3's binding to the pretargeting complex inhibits its ATPase activity (Rome et al., 2013), keeping it primed to receive the TA substrate. After receiving the substrate from Sgt2, Get3 dissociates from the pretargeting complex and hydrolyzes ATP (Rome et al., 2014). As Get3 loaded with the TA protein reaches its receptor complex consisting of Get1 and Get2 at the ER, they help Get3 release ADP, turn back into its open confirmation and insert the TA protein into the membrane (Rome et al., 2013; Stefer et al., 2011; Wang et al., 2011). Get1 and Get2 form a *bona-fide* insertase complex (Wang et al., 2014), with Get1 being part of the evolutionarily conserves Oxa1 superfamily of membrane protein biogenesis factors (Anghel et al., 2017).

3.3. The client spectrum of Get3

Although the TA protein targeting activity of Get3 has been clearly demonstrated *in vitro* and *in vivo* for some TA proteins (Schuldiner et al., 2008), its full *in-vivo* client spectrum remains elusive. Although the TMS hydrophobicity of TA proteins has been suggested to be the main determinant of Get3 dependence in mammals (Guna et al., 2018), Pex15, a TA protein in yeast with very low TMS hydrophobicity is a well-characterized substrate of Get3 (Li et al., 2019; Schuldiner et al., 2008). Therefore, there must be other factors as well that drive Get3 dependence.

Besides TA proteins, newly synthesized GPI-anchored proteins also possess a C-terminal hydrophobic TMS, which will be removed as the GPI anchor is attached to the protein later on (Orlean and Menon, 2007). Several GPI-anchored proteins have N-terminal signal sequences with low hydrophobicity, and are thus targeted to the ER in an SRP-independent manner (Ast et al., 2013). The subset of these proteins that also contain a highly hydrophobic C-terminal sequence requires Get3 for chaperoning their TMS (Ast et al., 2013). However, the complete range of GPI-anchored substrates of Get3 is not known.

Furthermore, the mammalian homolog of Get3, TRC40 can deliver short secretory proteins such as apelin and statherin to the Sec61 translocon (Johnson et al., 2012). This shows that Get3/TRC40 can have roles in protein biogenesis outside of the scope of the GET pathway.

3.4. Additional functions of Get3

Get3 performs other functions besides its role in TA protein targeting. On the one hand, it can bind the cytosolic C-terminal portion of the chloride transporter Gef1. This interaction depends on the presence of copper in the cytosol (Metz et al., 2006), which may reflect the sensitivity of Get3 to oxidation as discovered in the context of its in-vitro chaperone activity (Voth et al., 2014). On the other hand, it was also reported to be co-regulated with the proteasome and the Cdc48-Npl4-Ufd1 complex, indicating a potential role in protein homeostasis (Auld et al., 2006). Interestingly, Get3 forms foci together with the pretargeting complex and other chaperones when cells suffer acute glucose starvation, suggesting that it may be an active part of a chaperoning platform assembled when ATP levels drop in the cells (Powis et al., 2013).

Upon oxidation, Get3 can form multimers and get activated as a general chaperone and prevent the aggregation of chemically denatured substrates in vitro. This chaperone activity is ATP-independent and can be reversed by reducing the protein and providing ATP simultaneously. Such conformational rearrangements are also known from bacterial Hsp33 (Graumann et al., 2001). Although the physiological client spectrum of Get3 as a stress-induced chaperone is unknown, it highlights the idea that Get3 may have an additional chaperone function unrelated to TA protein targeting (Voth et al., 2014).

4. Aims of the thesis

The number of known biogenetic pathways membrane proteins can use has increased considerably, yet our knowledge of the client spectrum of each of these pathways remains limited in many cases. A prime example of this problem is illustrated by TA proteins, which were long thought to mainly use the GET pathway, but several points of evidence suggest that other pathways are also involved in their biogenesis. The discovery of an ATPase independent chaperone activity of the central component of the GET pathway, Get3 suggests that it has further in vivo functions besides TA targeting. Considering the above, the aims of this study were the following:

- To gain a better understanding of the evolutionary relationships between Get3-like proteins in all domains of life in order to help delineate its different in vivo functions in eukaryotes.

- To characterize Get3's *in-vivo* TA protein client spectrum with the help of microscopy-based systematic screens in budding yeast.

- To uncover novel non-TA protein classes of Get3 substrates with regards to its function as a membrane targeting factor and a chaperone, and characterize their interaction with Get3 biochemically.

Chapter Two

The natural history of Get3-like chaperones

As published in:

Farkas, Á., De Laurentiis, E.I., and Schwappach, B. (2019). The natural history of Get3-like chaperones. Traffic *20*, 311–324.

Shared under Creative Commons licence CC BY-NC-ND 4.0.

See: https://creativecommons.org/licenses/by-nc-nd/4.0/

To access Table S1, please see the online version of the manuscript:

https://onlinelibrary.wiley.com/doi/full/10.1111/tra.12643

Author contributions

Figure	Author	Contribution
Fig 1-5, Table S1-4	Ákos Farkas	Carried out all bioinformatic analysis as described in Materials and Methods, visualized all figures and tables, wrote text
	Evelina Ines de Laurentiis	Wrote introduction jointly with Ákos Farkas, supported text with ideas and by proofreading it.
	Blanche Schwappach	Provided general support in conceptualization of experiments, supported visualization of figures and assembly of text with critical comments, writing, and by proofreading it.

Traffic WILEY

The natural history of Get3-like chaperones

Ákos Farkas[1] 🄳 | Evelina Ines De Laurentiis[1] 🄳 | Blanche Schwappach[1,2] 🄳

[1]Department of Molecular Biology, Göttingen University Medical Center, Göttingen, Germany

[2]Max-Planck Institute for Biophysical Chemistry, Göttingen, Germany

Correspondence
Blanche Schwappach, Department of Molecular Biology, Göttingen University Medical Center Humboldtallee, 23 D-37073 Göttingen, Germany.
Email: blanche.schwappach@med.uni-goettingen.de

Get3 in yeast or TRC40 in mammals is an ATPase that, in eukaryotes, is a central element of the GET or TRC pathway involved in the targeting of tail-anchored proteins. Get3 has also been shown to possess chaperone holdase activity. A bioinformatic assessment was performed across all domains of life on functionally important regions of Get3 including the TRC40-insert and the hydrophobic groove essential for tail-anchored protein binding. We find that such a hydrophobic groove is much more common in bacterial Get3 homologs than previously appreciated based on a directed comparison of bacterial ArsA and yeast Get3. Furthermore, our analysis shows that the region containing the TRC40-insert varies in length and methionine content to an unexpected extent within eukaryotes and also between different phylogenetic groups. In fact, since the TRC40-insert is present in all domains of life, we suggest that its presence does not automatically predict a tail-anchored protein targeting function. This opens up a new perspective on the function of organellar Get3 homologs in plants which feature the TRC40-insert but have not been demonstrated to function in tail-anchored protein targeting. Our analysis also highlights a large diversity of the ways Get3 homologs dimerize. Thus, based on the structural features of Get3 homologs, these proteins may have an unexplored functional diversity in all domains of life.

KEYWORDS
bacteria, Chlorophyta, Embryophyta, endoplasmic reticulum, Get3p, molecular chaperone, Rhodophyta, tail-anchored protein

1 | INTRODUCTION

Tail-anchored (TA) proteins are a class of membrane proteins that contain a C-terminal hydrophobic transmembrane segment (TMS) and a functional N-terminal cytosolic domain.[1,2] TA proteins are a diverse group of eukaryotic membrane proteins found among others in the secretory pathway,[3] nuclear envelope,[4] peroxisomes,[5] mitochondria[6] and in chloroplasts.[7] They have a wide range of functions, such as assistance in vesicular trafficking,[3] protein translocation[8] and degradation[9] of membrane proteins. The function of TA proteins has been shown to be essential in all domains of life and their transport to the correct biological membrane, or protein targeting, needs to be efficient and

accurate as targeting errors can have detrimental cellular effects. Additionally, TMSs are prone to aggregation and their spontaneous insertion into lipid bilayers may be slow in vivo. Therefore, in order to ensure efficient and organelle-specific insertion of TA proteins and to prevent the aggregation of TMSs in the cytoplasm, most studies to date suggest that the targeting and insertion of TA proteins involves one or more cytosolic factors.

The mechanism through which TA proteins are targeted and inserted is distinct from the co-translational signal recognition particle (SRP)-facilitated process by which most membrane proteins with N-terminal or internal signals are targeted. Indeed, because the C-terminal TMS of a TA protein emerges from the ribosome at the end of translation, TA proteins are targeted and inserted through post-

translational mechanisms. One such pathway, the guided entry of TA proteins (GET), identified a little over 10 years ago, has been shown to mediate the proper delivery of several TA proteins in mammals,[10,11] budding yeast[12] and more recently in plants.[13,14]

Extensive biochemical and structural studies performed over the last decade have characterized the targeting of TA proteins utilizing the yeast GET pathway (as reviewed in[15]). Initially, a pre-targeting complex, consisting of a small glutamine-rich tetratricopeptide repeat containing protein Sgt2, and Get4 and Get5 in yeast, or Bag6, SGTA, TRC35 and UBL4A in mammals, captures the TA protein following its release from the ribosome, then transfers it to the ATPase Get3 in budding yeast, or TRC40 in mammals.[15–17] The TA-bound Get3/TRC40 protects and delivers the TA protein to the ER membrane, where its receptor complex comprised of Get1 and Get2 in yeast or WRB and CAML in mammals stimulates its subsequent release into the membrane.[18–21]

Despite the apparent complexity and necessity of the GET pathway to prevent aggregation of hydrophobic proteins, depletion of GET pathway components in budding yeast (*Saccharomyces cerevisiae*) and *Arabidopsis thaliana* is not lethal.[12,14] Yet the functional importance of the GET pathway is highlighted by the fact that GET pathway deletion *S. cerevisiae* strains show increased heat and oxidative stress sensitivity,[12,22] and the depletion of TRC40 is embryonically lethal in mice.[23] Since Get3 was shown to possess chaperone holdase activity upon oxidation[24] these phenotypes may to some extent reflect a chaperone activity not involved in targeting TA proteins during biogenesis.

2 | GET3 HOMOLOGS IN THE DIFFERENT DOMAINS OF LIFE

Some phylogenetic aspects of other GET pathway components have been recently discussed, in particular the evolutionary relationships between components of the pretargeting complex comprising Get4, Get5 and Bag6[17] and the legacy of membrane protein biogenesis factors similar to bacterial Oxa1 that also include Get1.[18] Here, we combine a review of the literature on Get3 structure and function with a comprehensive bioinformatic analysis of the structural elements of the protein involved in TA protein binding. This integration focusses on properties of Get3-like proteins in all domains of life that render the hydrophobic cage versatile and should be considered for both functions of these proteins.

A systematic search for Get3- and ArsA-homologous proteins in the KEGG and OrthoDB databases combined with further BLAST analysis yielded 2208 sequences (Supporting Information Table S1), from which 51 representative sequences were chosen to construct a phylogenetic tree (Figure 1). This analysis reveals a functionally unexplored diversity of Get3-like proteins (Table 1). Focusing on structural aspects of different homologs such as domain organization or the presence of sequence motifs and comparing them with known structures and functions of Get3 homologs, we would like to highlight that Get3-like chaperones from different kingdoms are more similar

to each other than previously recognized based on a comparison of eukaryotic Get3 or TRC40 with prokaryotic ArsA.[25] At the same time, they are remarkably diverse with respect to their modes of (pseudo) dimerization and structural features outside the well conserved ATPase domain.

3 | STRUCTURAL ORGANIZATION OF ARSA AND GET3 PROTEINS

A bacterial homolog of Get3, ArsA confers resistance to arsenite in *Escherichia coli*[26] and shows high structural similarity to Get3 (Figure 2A,B). ArsA folds such that two highly similar domains in tandem form a metal binding site and two nucleotide binding sites (NBS) at their interface.[27] The NBS is similar to those found in other members of the Signal recognition particle, MinD, BioD (SIMIBI) class of P-loop NTPases and contains conserved structural elements necessary for ATP hydrolysis including the P-loop, Switch I and II and the A-loop.[28–30] The metal binding site involves three functionally essential cysteine residues, however, these residues are not conserved in eukaryotic Get3 homologs[27] (Figure 2A, ball-and-stick model residues).

Unlike bacterial ArsA, Get3 in budding yeast (*Sc*Get3) and other fungi and animals has a single Get3-homology domain. Two *Sc*Get3 monomers assemble into rotationally symmetrical homodimers to form a structure analogous to the arrangement of the two domains found in ArsA (Figure 2B). In ArsA, a short helix involved in coordinating the metal ion (orange in Figure 2A) folds into a groove, whereas the same region forms extended helices in Get3 (helix 7, 9; orange in Figure 2B), and also contains an additional stretch of amino acids dubbed the TRC40-insert.[25] Thus, a large hydrophobic surface is created (Figure 2B, bottom row) allowing Get3 to accommodate the TMSs of TA proteins. At the same time, the helix contained within the TRC40-insert (helix 8, not visible in the structure) is thought to act as a lid that closes on captured TMSs, thus shielding them from the solvent.[31]

Get3 homologs with the ability to bind TA proteins have also been found in archaea, and one out of the four archaeal homologs studied so far could deliver captured substrates to the membrane.[32,33] In bacteria, the only currently known Get3 homologs with a hydrophobic groove belong to photosynthetic bacteria and they also have an α-crystallin domain at the C-terminus (red in Figure 2C).[34] α-crystallin domains are key components of heat shock proteins and are essential for their chaperone function.[35] Although such Get3 homologs are also found in land plants,[34] their function remains unknown. Moreover, land plants, Chlorophytes and red algae have been proposed or shown to have several Get3 homologs without an α-crystallin domain as well, some of them in chloroplasts and mitochondria.[14]

4 | CONSERVATION OF HELIX 8, THE "LID" CLOSING THE HYDROPHOBIC GROOVE

Get3 has several hydrophobic residues necessary for its interaction with TA proteins, and they mostly converge on the C-terminal portion
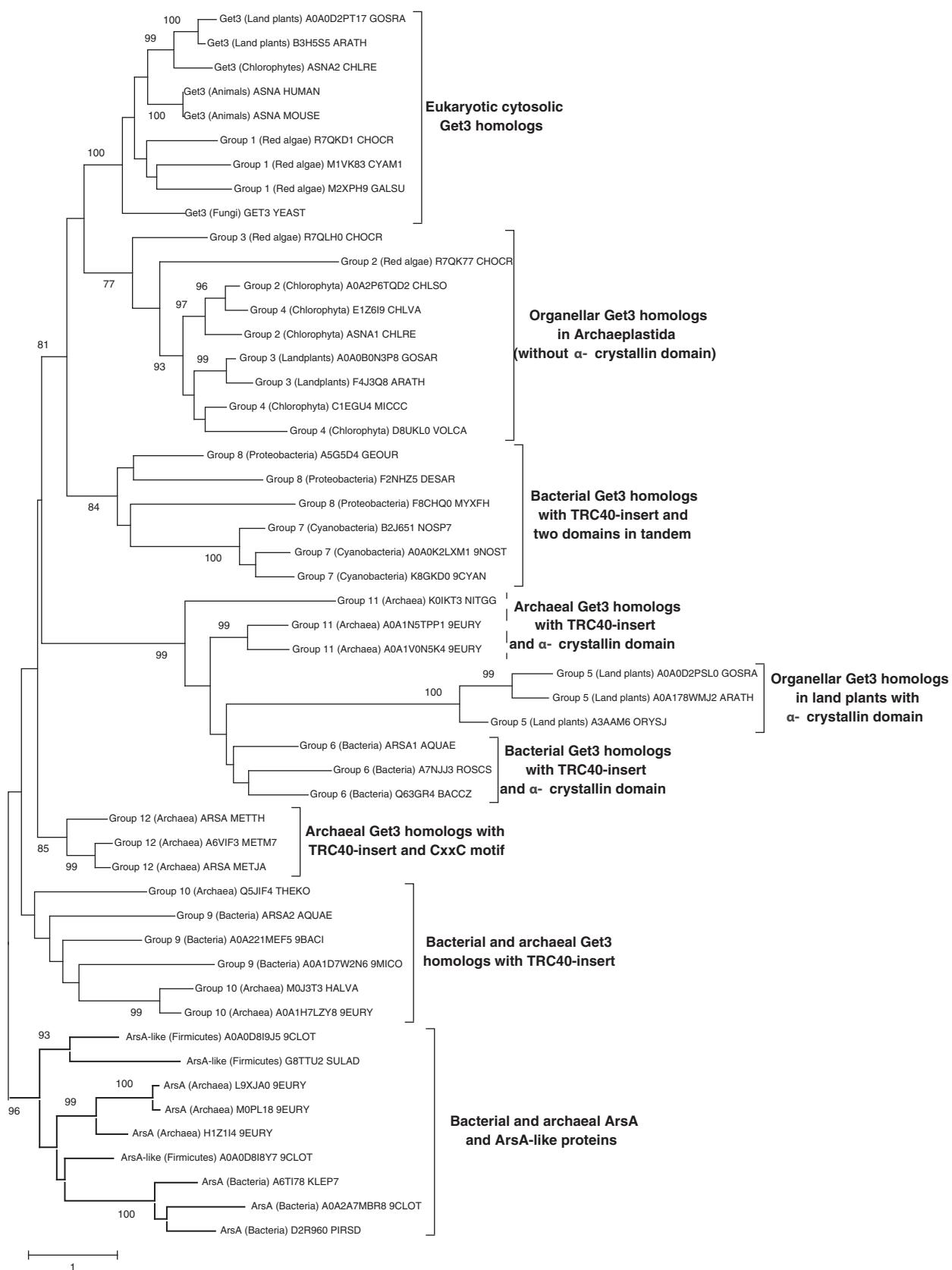
FIGURE 1   Maximum likelihood rooted phylogenetic tree of three representative sequences of each group of Get3 homologs as defined in Table 1. Percentage of trees in which the sequences clustered together after applying 1000 bootstraps are indicated at nodes if the value is higher than 70%. Scale bar indicates number of substitutions per site

TABLE 1 Summary of structural features present in the different groups of the phylogenetic tree (Figure 1). Based on the sequences listed in Table S1

	Species	Protein	Identified sequences in group	Localization	Hydrophobic groove/TRC40-insert	CxC motif	CxxC motif	α-crystallin domain	Get3 domains	Get1/2/4 binding residues	Other features/Comments
Eukaryotes	Animals, Fungi	Get3/TRC40	629	Cytoplasmic	Yes	Frequent	Yes	No	Single	Yes	In 597 out of 629 sequences
	Land plants, Chlorophytes	Get3/TRC40	87	Cytoplasmic	Yes	Yes	No	No	Single	Yes	C-terminal charged patch
	Red algae	Group 1	4	Cytoplasmic	Yes	Frequent	No	No	Double	Yes	In three out of four sequences
	Red algae, Chlorophytes	Group 2	10	Organellar	Yes	No	No	No	Double	No	
	Land plants, Red algae	Group 3	90	Organellar	Yes	No	Yes	No	Single	No	
	Chlorophytes	Group 4	8	Organellar	Yes	No	No	No	Single	No	
	Land plants	Group 5	48	Plastidial	Yes	No	No	Yes	Single	No	
Bacteria	Mainly Actinobacteria, Chloroflexi, Chlorobi Cyanobacteria, Firmicutes, some Acidobacteria, Aquificae, Bacteroidetes, Fusobacteria, Spirochaetes, Proteobacteria	Group 6	301	Cytoplasmic	Yes	No	No	Yes	Single	No	Rarely: nucleotide binding site missing (such homologs excluded from current analysis)
	Cyanobacteria	Group 7	29	Cytoplasmic	Yes	No	No	No	Double	No	
	Myxococcales, some Proteobacteria	Group 8	17	Cytoplasmic	Yes	No	Yes	No	Double	No	
	Mainly Actinobacteria, Firmicutes, Proteobacteria, some Aquificae, Spirochaetes, Synergistetes	Group 9	131	Cytoplasmic	Yes	No	No	No	Single	No	
	Some Firmicutes	ArsA-like	18	Cytoplasmic	No	No	No	No	Single	No	Part of arsenite resistance operon
	Mainly Actinobacteria, Bacteroidetes, Proteobacteria, Firmicutes, Planctomycetes	ArsA	432	Cytoplasmic	No	No	No	No	Double	No	Part of arsenite resistance operon
Archaea	Euryarcheota	ArsA	27	Cytoplasmic	No	No	No	No	Double	No	
	Euryarcheota, Crenarcheota	Group 10	155	Cytoplasmic	Yes	No	No	No	Single	No	Some of the Get1/2/4 binding residues are

TABLE 1 (Continued)

Species	Protein	Identified sequences in group	Localization	Hydrophobic groove/TRC40-insert	CxC motif	CxxC motif	α-crystallin domain	Get3 domains	Get1/2/4 binding residues	Other features/Comments
Euryarcheota, Thaumarcheota,	Group 11	21	Cytoplasmic	Yes	No	No	Yes	Single	No	present in some homologs
Euryarcheota, Crenarcheota	Group 12	200	Cytoplasmic	Yes	No	Yes	No	Single	No	

of helix 7 and the short helix 8 following it.[25] Recently, it has emerged that helix 8 is needed to ensure an efficient transfer of substrates from upstream components to Get3, but it has no major effect on the dissociation of substrates already captured by Get3.[36] In the structures of eukaryotic Get3 homologs, the region around helix 8 is poorly defined because of its high flexibility and is heavily influenced by the overall conformation of the protein.[25,31] Although helix 8 forms a helix separate from helix 7 in fungal Get3 structures (Figure 3A), these two helices appear to line up or even merge completely in structures of archaeal homologs of Get3.[32,33] The Get3 homolog of the archaeon *Methanocaldococcus jannaschii* (*Mj*Get3) exists either as a dimer, similar to *S. cerevisiae* Get3 (*Sc*Get3) or in a tetrameric form, a dimer of dimers, the assembly of which is mediated by the region corresponding to helix 8 in *Sc*Get3 (Figure 3B).[33] Although *Sc*Get3 and its human homolog TRC40 both form tetramers under specific conditions,[37,38] these tetrameric structures remain structurally unsolved and the role of helix 8 in their assembly also remains unknown. Nevertheless, it is intriguing that based on secondary structure predictions, the regions corresponding to helix 7 and 8 of *Sc*Get3 would be expected to form a single helix as seen in *Mj*Get3 (Figure 3A), yet whether this region can indeed assume two distinct conformations remains to be seen.

Although the region linking helix 7 and 9 appears to be moderately conserved, especially at helix 8 in eukaryotes, its length varies considerably within and between phylogenetic groups (Figure 3C). Indeed, although the average length of the stretch homologous to the linker between helix 7 and 9 in *Sc*Get3 is approximately 21 to 22 amino acids in eukaryotes, there are notable exceptions as well. For instance, *Sc*Get3 only has 15 amino acids in this region while the archaeal *Mj*Get3 has 25, showing that from a functional perspective, substantial variation is allowed in this region. Interestingly, unlike in bacterial ArsA (Figure 3A), the length of this region in bacterial Get3 homologs with an α-crystallin domain is comparable to that observed in eukaryotes. For example, while the region in Firmicutes is often 21 amino acids long, just like in animals, in Cyanobacteria it is as long as in *Mj*Get3 and in many Chlorobi bacteria almost as short as in *Sc*Get3. However, whether this indicates any functional similarity is not known.

Besides the length of the linker between helix 7 and 9, its amino acid composition also shows variation within and between phyla (Figure 3D). It has been suggested that the methionine-rich nature of the hydrophobic groove is important for the accommodation of the TMS.[25] This hypothesis is further strengthened by the analogy with SRP, where the methionine-rich M domain of Srp54 is essential for binding the signal peptide.[39] In Get3 helix 8, there are two and three methionine residues in *Sc*Get3 and human TRC40, respectively, and their combined loss in *Sc*Get3 leads to decreased substrate binding.[25] Consistent with the idea that the presence of the methionine residues is related to the TA targeting function, in homologs not expected to be involved in TA protein targeting (bacterial and plastidial-mitochondrial Get3 homologs without an α-crystallin domain in land plants), there is mostly no or just a single methionine in the corresponding region. However, looking at Get3 homologs known to

(A)   (B)   (C)

ArsA
E. coli

Get3
S. cerevisiae
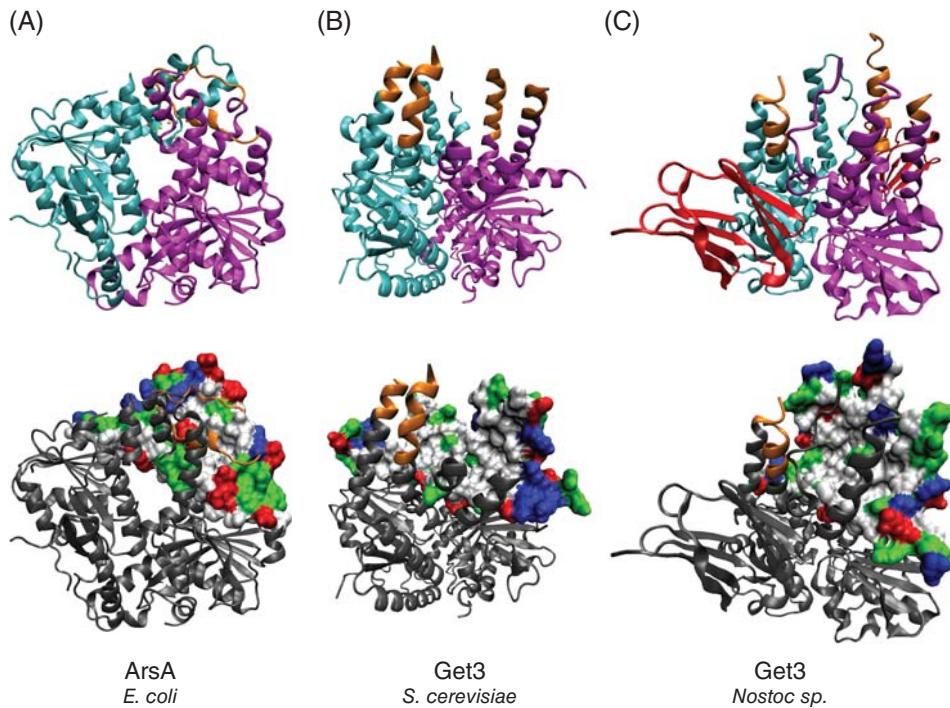
Get3
Nostoc sp.

FIGURE 2   Structure of selected Get3 homologs. Top: individual domains (A) or subunits (B and C) are marked in cyan and magenta. Bottom: hydrophobic groove or homologous region shown in surface view. Hydrophobic and nonpolar residues are shown in white, polar residues shown in green, acidic residues shown in red and basic residues are shown in blue. To allow a better view of the interior of the groove, only half of the groove is shown in surface view (B and C). A, Structure of *E. coli* ArsA (PDB ID: 1F48). The region unique to ArsA is highlighted in orange. Heavy metal ion coordinating cysteines are shown as ball-and-stick models. B, Structure of *S. cerevisiae* Get3 (PDB ID: 4XTR). The region homologous to the one marked in orange in A is also marked in orange here. C, Structure of Get3 from a *Nostoc* species (PDB ID: 3IGF). The α-crystallin domain is depicted in red

bind or target TA proteins, it becomes clear that although eukaryotic and archaeal homologs tend to have at least one or more methionine residues in this stretch, there are several species without any as well (Figure 3D). Taken together, although helix 8 may have become enriched in methionine in certain species to support TA protein targeting, the presence of methionine residues does not seem to be a requirement for helix 8 to fulfill its function.

5 | HELICES LINING THE HYDROPHOBIC GROOVE

One of the defining features of Get3 with respect to bacterial ArsA is the presence of the TRC40-insert, which corresponds to helix 8 in *Sc*Get3 and the amino acids linking it to helix 9 (Figure 4A).[25] The TRC40 insert with an extended helix 7 and 9, together with helices 4, 5 and 6 creates a hydrophobic area so that TMSs can be accommodated and shielded from solvents in the resulting groove. Mutational studies have revealed that some of the hydrophobic residues of helix 7 and 8 are important for substrate binding by Get3.[25] Interestingly, while the residues that show the strongest effect in mutational studies of *Sc*Get3 are not conserved in bacterial ArsA, other hydrophobic residues in helix 7 are universally conserved in eukaryotic Get3 homologs and bacterial ArsA as well (Figure 4A).[40] Therefore, the presence of these crucial hydrophobic residues and the TRC40-insert may be indicative of functional similarity between eukaryotic Get3 and any given bacterial homolog. Indeed, several bacterial phyla have Get3 homologs with an α-crystallin domain that also have the TRC40-insert, and the surrounding helices often contain periodic hydrophobic amino acids that are even positionally conserved (Figure 4A). Although previously only described in photosynthetic

bacteria and land plants,[34] such Get3 homologs can be found in diverse groups of non-photosynthetic bacteria as well, including Actinobacteria, Aquificae, Bacteroidetes, Firmicutes, Fusobacteria and Proteobacteria.

The helix lying at the bottom of the hydrophobic groove, helix 6, shows a high overall similarity between bacterial ArsA and Get3, and mutations of hydrophobic residues in this helix mostly affect the ATPase activity of Get3 but not substrate binding.[25] Since one of the cysteines involved in coordinating the metal ion is close to the N-terminal part of this helix, its presence, coupled with the lack of the Get3-/TRC40-insert, is expected to be a strong indicative feature of ArsA homologs (Figure 4B). Indeed, most such ArsA homologs in our analysis are highly similar to *E. coli* ArsA and have two domains in tandem. However, some Firmicutes bacteria seem to be unique in that they possess two copies of such ArsA homologs, but each with only a single domain instead of two (Table 1). In this case, one of them is similar to the first domain of *E. coli* ArsA, and the other is similar to the second. However, because there is no obvious feature that would mediate dimerization, it is unknown whether they actually do form dimers and function as a bona fide ArsA in vivo.

Another special feature in helix 6 is found in plastidial and mitochondrial Get3 homologs without an α-crystallin domain in land plants, Chlorophyta and red algae, that is, the Archaeplastida clade. Besides the Get3 homologs already shown to localize to the chloroplast and mitochondria,[14] similar organellar Get3 homologs are predicted to exist in other groups within the Archaeplastida clade as well (Table S2). In spite of overall sequence similarity to *Sc*Get3, many of these homologs have several proline residues at the N-terminus of helix 6 (Figure 4B), the relevance of which is currently unknown. Furthermore, such homologs uniformly lack the CxC motif found on the beta strand following helix 9, a feature strongly, although not
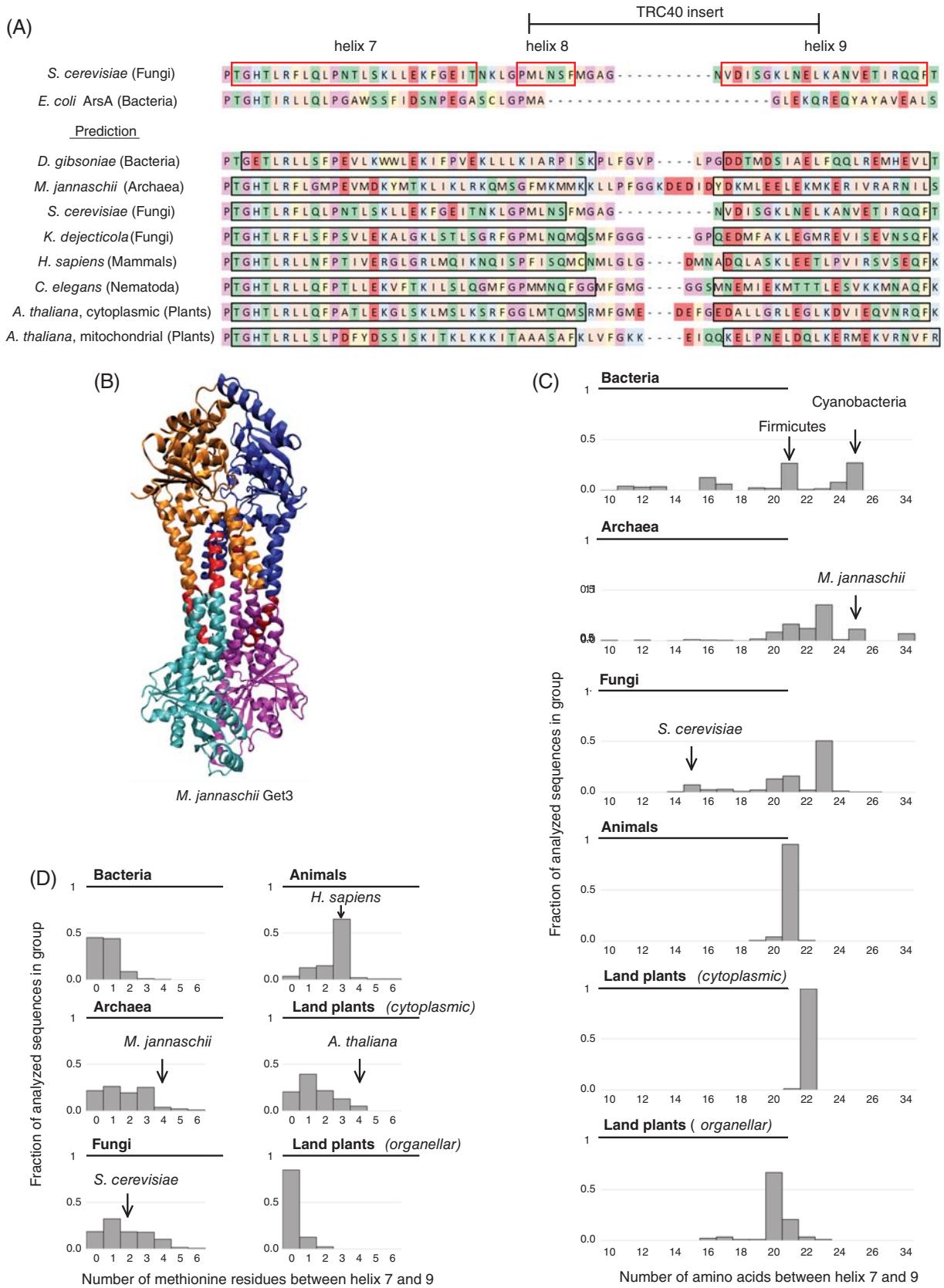
FIGURE 3 Legend on next coloumn.

universally conserved among eukaryotic cytoplasmic Get3 homologs (Figure 4C). Considering that these homologs also lack key residues required for binding Get1, Get2 and Get4 (Figure 4C, Table 1), it is clear that such organellar Get3 homologs should fulfill a related yet distinct function compared to cytoplasmic Get3 homologs.

Looking at the hydrophobic groove as a whole, its methionine-rich nature has been thought to be a feature related to the TA protein targeting function of Get3.[25] As stated above, the presence of methionine residues in helix 8 is probably not a prerequisite for TA protein targeting. However, counting all the methionine residues that could potentially flank the hydrophobic groove (from helix 4 to helix 9), it becomes clear that despite considerable variety, all fungi have at least four methionine residues in this region (ScGet3 has six), and most vertebrates have three times as many (Figure 4D). On the other hand, many bacterial Get3 homologs with an α-crystallin domain also have multiple methionine residues (Table S3), with several Firmicutes homologs having as many as eight (Figure 4D). As an exception, homologs in Actinobacteria tend to have fewer or no methionine residues at all (Figure 4D). Taken together, the fact that many bacterial homologs have as many methionine residues as some fungi do, and that there has been no indication so far that these homologs target TA proteins in bacteria, it is likely that the methionine-rich nature of the hydrophobic groove had already been present before the TA protein targeting function of Get3 was acquired. Then, as eukaryotic Get3 became more and more specialized to target TA proteins, it may have acquired further methionine residues in the groove to facilitate the binding of TMSs.

6 | GET3 HOMOLOGS IN THE EUKARYOTIC GROUP ARCHAEPLASTIDA

In yeast and most other eukaryotes, Get3 functions as a rotationally symmetrical homodimer because dimerization is necessary for both its ATPase activity and the formation of the TMS binding hydrophobic groove.[22,25] In fungi and mammals, a conserved CxxC motif in each subunit aligns to coordinate a zinc ion (Figure 5A), which is necessary for dimer formation.[22,25] Bacterial ArsA homologs are very similar structurally, but the two halves of the dimer are produced as two domains in tandem in a single polypeptide chain. Each domain corresponds to a subunit in a eukaryotic dimeric Get3.[27] It is likely that because of the two domains being part of a single protein, the

interaction between them is stable enough so that no CxxC is required in ArsA homologs.

Surprisingly, cytoplasmic homologs in land plants and Chlorophytes lack the CxxC motif, yet they are functional in targeting TA proteins[13] and they form dimers.[14] An analysis of the sequence of these homologs provides clues as to how dimerization could happen in the absence of the CxxC motif. On the one hand, a pronounced acidic patch composed of three to five acidic residues is located in land plant homologs adjacent to the site where the CxxC motif would be (Figure 5A). Barring a few exceptions, an ExxE motif is found in this sequence, and such motifs are known to be able to coordinate iron ions.[41,42] Therefore, although land plant homologs are lacking a CxxC motif, they could still utilize metal ions to stabilize the dimer. On the other hand, most land plant and Chlorophyte homologs have a short, ca. 30 amino acid long, strongly charged extension missing in all other phyla, which may be involved in dimerization (Figure 5B and C). Finally, cytoplasmic Get3 homologs in red algae are distinct from land plants or Chlorophytes, in that they form a single polypeptide chain containing two domains in tandem, just like bacterial ArsA, and they similarly lack the CxxC motif as well (Figure 5A and C).

Compared to cytoplasmic homologs, predicted organellar Get3 variants without an α-crystallin domain in Archaeplastida display an even greater diversity. Namely, these proteins lack the CxC motif and key Get1, Get2, Get4 binding residues and they often contain extra prolines in helix 6. Even so, they are still hypothesized to dimerize and use several different ways to achieve this (Figure 5C). In organellar homologs in land plants and red algae, a CxxC motif is present, and is likely used to form a dimer. However, red algae have other homologs as well, as do Chlorophytes, that contain two domains in a single protein. Intriguingly, additional homologs of Get3 can be found in Chlorophytes that have no apparent dimerization motif, which does not exclude the possibility that they still form dimers in unexpected ways.

It has to be noted that an organellar homolog from *Chlamydomonas reinhardtii* predicted here to be organellar has been previously proposed to be cytoplasmic.[43] However, the protein is highly similar to other homologs in land plants that have been shown to be organellar,[14] and homologous proteins from other Chlorophytes are consistently predicted to be organellar as well (Table S2). Therefore, the localization of these homologs in Chlorophytes remains uncertain for the moment. The picture is further complicated by the fact that some of the Get3 homologs in Archaeplastida are highly similar in sequence to other organellar homologs, yet are predicted to be

FIGURE 3 Comparison of the TRC40-insert between species. A, Known secondary structure of ScGet3 (top) compared with the predicted structure of the same region in different species (bottom, predicted helices marked with black frame). Hydrophobic residues shown in peach, aromatic residues in ochre, basic residues in blue, acidic residues in red, hydrophilic residues in green, proline and glycine in mauve, cysteine in yellow. B, Structure of *M. jannaschii* Get3 (PDB ID: 3UG6). Subunits are marked with cyan, magenta, orange and blue. The region homologous to the region between helix 7 and 9 in ScGet3 is shown in red. C, Distribution of the length of the region homologous to the sequence between helix 7 and 9 in ScGet3 among the sequences used for the current analysis. All bins containing at least 1% of the sequences are shown in the chart. Number of analyzed sequences: Bacteria—299; Archaea—376; Fungi—489; Animals—140; Land plants (cytoplasmic) —78; Land plants (organellar, excluding α-crystallin domain Get3 homologs)—87. D, Distribution of the number of methionine residues in the region homologous to the sequence between helix 7 and 9 in ScGet3 among the sequences used for the current analysis. All bins containing at least 1% of the sequences are shown in the chart. The number of sequences analyzed are as in C
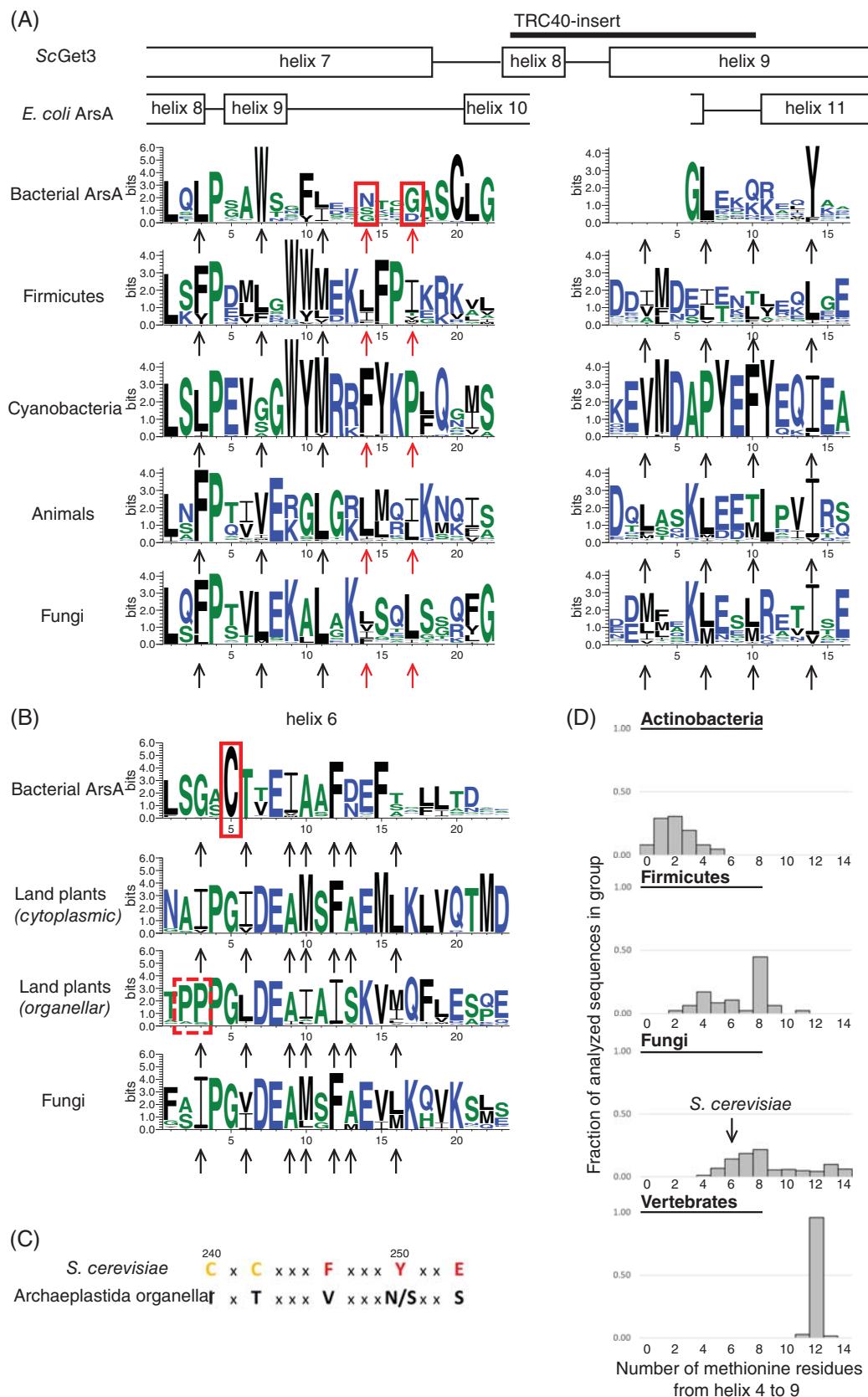
FIGURE 4 Legend on next coloumn.

cytoplasmic because of a lack of a transit peptide (Table S2). This could either indicate a further cytoplasmic group of such homologs or simply reflect an inaccurate bioinformatic prediction of the N-terminus of the proteins based on genomic sequences.

Besides the above-mentioned homologs, land plants also have a plastidial Get3 homolog that is closer in similarity to cyanobacterial homologs than it is to eukaryotic ones (Figures 5C and 1 and Table 1).[34] Accordingly, this is the only eukaryotic Get3 homolog currently known to have an α-crystallin domain at its C-terminus like the one seen in the structure of *Nostoc*Get3 (Figure 2C).

The fact that land plants have Get3 homologs in the chloroplast with hypothetically two different ways to dimerize (one with an α-crystallin domain, one with a CxxC-motif) raises the question of what advantage having two such close homologs may bring. A possibility would be that the different modes of dimerization allow the organism to regulate the activity or various functions of the protein. Nonetheless, as no study has been carried out on these proteins to date, their function remains elusive as of now.

7 | PREVIOUSLY UNNOTICED BACTERIAL GET3 HOMOLOGS

As mentioned above, several major groups of bacteria have a Get3 homolog with the TRC40-insert similar to *Nostoc*Get3 (Figure 2C), with an α-crystallin domain attached to the C-terminus and a nucleotide binding site present, which is missing in *Nostoc*Get3 (Figure 5D, Table 1). Although it is known that α-crystallin domains can mediate dimerization and act as a chaperone,[35] it is not clear from the available structure of *Nostoc*Get3 whether or how it contributes to the stabilization of the Get3 dimer. The possibility that it may have a different function is supported by the fact that several groups of bacteria have Get3 homologs with the TRC40-insert but no α-crystallin domain (Figure 5D and Table 1). These are highly similar to archaeal Get3 homologs lacking a CxxC motif, at least one of which has been demonstrated to be able to form dimers and bind TA proteins.[33] Therefore, it is highly likely that they can also dimerize and bind hydrophobic sequences.

Furthermore, uniquely among bacteria, Cyanobacteria, Myxococcal species and some further Proteobacteria contain Get3 homologs with two domains in tandem where both domains contain a TRC40-insert (Figure 5D). Except for Cyanobacteria, they also have the CxxC motif,

which makes these homologs unique not only among bacteria but in all domains of life.

It is currently unclear what the functions of these Get3 homologs are. Taking into consideration that all of the above-mentioned bacterial homologs have the TRC40-insert, current theory would predict their involvement in TA protein biogenesis. Since they have not been characterized yet, we can only rely on predictions. If they indeed insert TA proteins into the membrane, one would expect that bacterial species with more TA proteins would be more likely to have a TRC40-insert containing Get3 homolog than species that have fewer TA proteins. Indeed, Proteobacteria and Cyanobacteria with Get3 homologs arranged as two domains in tandem tend to have more predicted TA proteins than other bacteria (empty circles in Figure 5E and Table S4). However, comparing the abundance of predicted TA proteins between different bacterial species and the presence or absence of other Get3 homologs with a TRC40-insert reveals no correlation (Figure 5E and Table S4). Furthermore, it has been shown that other chaperones are responsible for TA protein targeting in bacteria, at least in *E. coli*.[44]

Considering that *Sc*Get3 can act as a more general chaperone under specific conditions,[38] that a large group of bacterial Get3 homologs have an α-crystallin domain with expected chaperone activity, and that all the TRC40-insert-containing homologs mentioned above are predicted to have a hydrophobic groove, it is possible that these homologs act more as general chaperones than TA protein targeting factors. From this perspective, the TA protein targeting activity of cytoplasmic eukaryotic Get3 homologs could represent an adaptation of an ancient, more general chaperoning function. On the same note, it would be interesting to know whether the above-mentioned Get3 homologs in chloroplasts and mitochondria have a similar function to those found in bacteria or represent a third group of Get3-like chaperones with unexpected functions. As summarized in Table 1, it is clear that Get3-like chaperones are widespread and structurally diverse and much remains to be discovered about the dynamic structure and function of these proteins.

8 | MATERIALS AND METHODS

8.1 | Retrieval and processing of sequences and structures

Identifiers of Get3 and ArsA homologs were retrieved from KEGG Database (https://www.genome.jp/kegg/) and OrthoDB[45] (https://

FIGURE 4 Consensus sequence and important features of the helices flanking the hydrophobic groove and of the region C-terminally adjacent to it. A, Consensus sequence of the region homologous to *Sc*Get3 helix 7 and 9 in different groups of Get3 homologs. Residues flanking the hydrophobic groove in *Sc*Get3 are marked with an arrow. Residues shown to be important for TA protein binding are marked with a red arrow. Corresponding residues are highlighted with a red rectangle in the ArsA consensus sequence. B, Consensus sequence of the bottom of the hydrophobic groove (*Sc*Get3 helix 6) in different groups of Get3 homologs. Residues facing the hydrophobic groove in *Sc*Get3 are marked with an arrow. Heavy metal ion coordinating cysteine in ArsA and additional proline residues in organellar Get3 homologs highlighted by red boxes. C, Comparison of the CxC motif and key Get1/Get2/Get4 residues between *Sc*Get3 and organellar homologs of Get3 in land plants. D, Distribution of the number of methionine residues in the region homologous to the sequence from helix 4 to 9 in *Sc*Get3 among the sequences used for the current analysis. All bins containing at least 1% of the sequences are shown in the chart. Number of analyzed sequences: Actinobacteria—62; Firmicutes—47; Fungi—489; Vertebrates—70
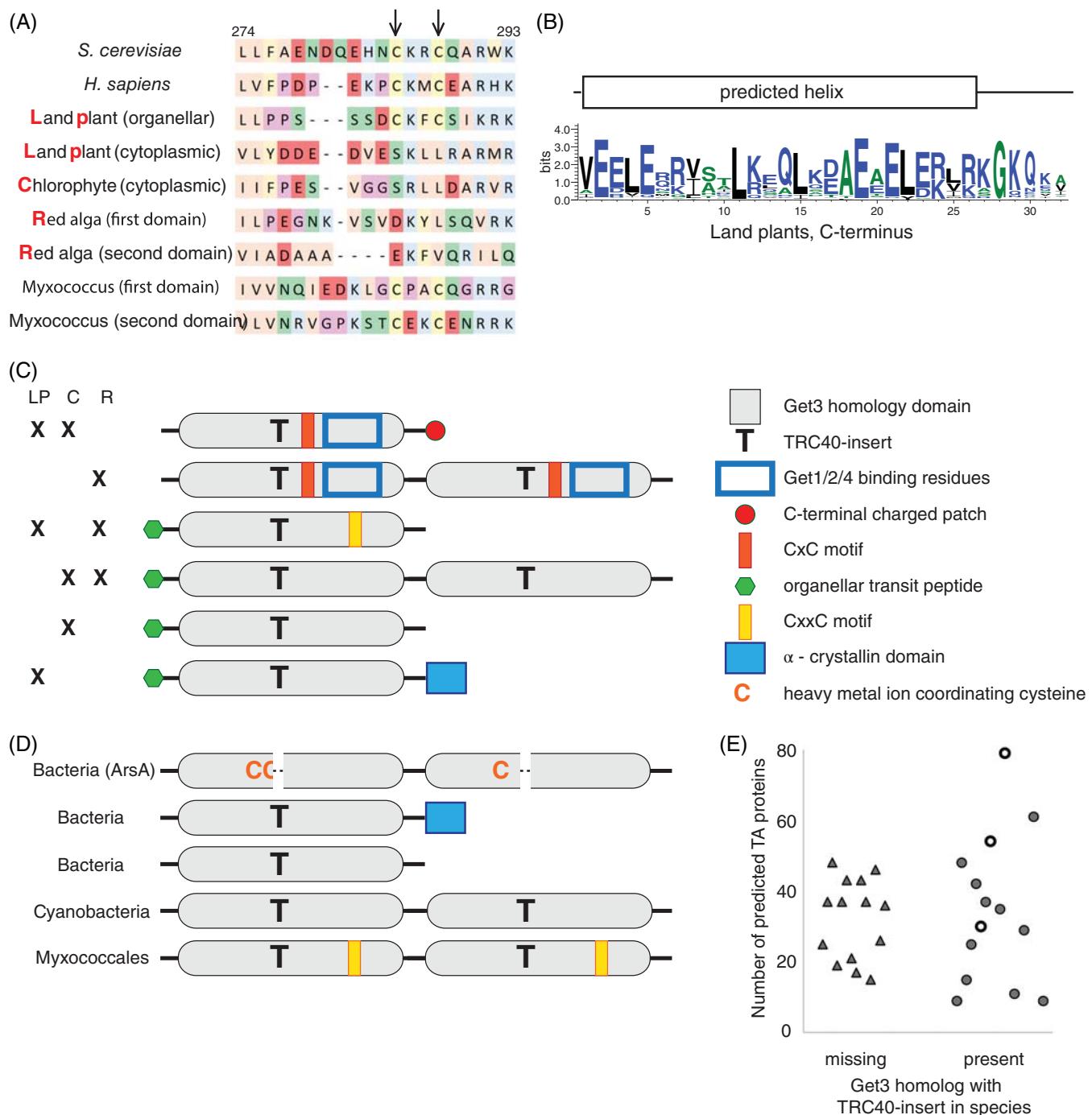
FIGURE 5 Get3 homologs use various strategies to form dimers. A, Comparison of the sequence adjacent to the CxxC motif in *Sc*Get3 and homologs from other organisms. B, Consensus sequence and secondary structure prediction of the charged C-terminal helix found in cytoplasmic Get3 homologs in land plants. C, Graphical representation of main structural features of land plant (LP), chlorophyte (C) and red algal (R) Get3 homologs. D, Graphical representation of main structural features of bacterial Get3 homologs. E, Comparison of the presence or absence of a TRC40-insert containing Get3 homolog in bacterial species with the number of predicted TA proteins in the given species. Empty circles represent Proteobacteria and Cyanobacteria with TRC40-insert containing Get3 homologs arranged as two domains in a single polypeptide

www.orthodb.org). Identifiers of Get3 homologs with an α-crystallin domain in land plants were retrieved using a blast search in land plants using the sequence of *Nostoc*Get3. Sequences were retrieved from Uniprot (www.uniprot.org) based on the identifiers collected from the databases and the blast search and aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Alignment was visualized and manually adjusted using Jalview.[46] Incomplete sequences were filtered out based on missing major regions compared to other homologs.

Secondary structure predictions were carried out using JPred4 (http://www.compbio.dundee.ac.uk/jpred4/).[47] Logos of consensus sequences were visualized using WebLogo 3.[48] Structures were retrieved from RCSB (https://www.rcsb.org/) and visualized using VMD (http://www.ks.uiuc.edu/Research/vmd/).[49]

8.2 | Construction of the phylogenetic tree

All analyses related to the phylogenetic tree were carried out in Mega X.[50] Sequences were aligned using the MUSCLE algorithm with default settings and manually adjusted when necessary. The evolutionary history was inferred by using the Maximum Likelihood method and Whelan and Goldman + Freq. model.[51] The tree with the highest log likelihood (−46 225.64) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches if they clustered together in more than 70% of the trees. Initial trees for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (five categories [+G, parameter = 1.3658]). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 51 amino acid sequences. There were a total of 1055 positions in the final dataset.

8.3 | Prediction of TA proteins in bacteria

The proteome of each species was download from Uniprot (www.uniprot.org). Transmembrane domains in the whole proteome were predicted using TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).[52] Proteins with a single TMS and less than 30 amino acids between the TMS and the C-terminus were considered candidates. These were tested for the presence of an N-terminal signal sequence using SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) with a cutoff value set to the value recommended by the software for the given bacterial phylum.[53]

ORCID

Ákos Farkas https://orcid.org/0000-0002-7295-3391
Evelina Ines De Laurentiis https://orcid.org/0000-0002-6469-1627
Blanche Schwappach https://orcid.org/0000-0003-0225-6432

REFERENCES

1. Kutay U, Hartmann E, Rapoport TA. A class of membrane proteins with a C-terminal anchor. *Trends Cell Biol*. 1993;3:72-75.
2. Borgese N, Fasana E. Targeting pathways of C-tail-anchored proteins. *Biochim Biophys Acta Biomembr*. 2011;1808(3):937-946.
3. Hardwick KG, Pelham HRB. SED5 encodes a 39-kD integral membrane protein required for vesicular transport between the ER and the Golgi complex. *J Cell Biol*. 1992;119(3):513-521.
4. Manilal S, Man NT, Sewry CA, Morris GE. The Emery-Dreifuss muscular dystrophy protein, emerin, is a nuclear membrane protein. *Hum Mol Genet*. 1996;5(6):801-808. https://doi.org/10.1093/hmg/5.6.801.
5. Elgersma Y, Kwast L, van den Berg M, et al. Overexpression of Pex15p, a phosphorylated peroxisomal integral membrane protein required for peroxisome assembly in S.Cerevisiae, causes proliferation of the endoplasmic reticulum membrane. *EMBO J*. 1997;16(24):7326-7341. https://doi.org/10.1093/emboj/16.24.7326.
6. Mozdy AD, McCaffery JM, Shaw JM. Dnm1p GTPase-mediated mitochondrial fission is a multi-step process requiring the novel integral membrane component Fis1p. *J Cell Biol*. 2000;151(2):367-380. https://doi.org/10.1083/jcb.151.2.367.
7. von Loeffelholz O, Kriechbaumer V, Ewan RA, et al. OEP61 is a chaperone receptor at the plastid outer envelope. *Biochem J*. 2011;438(1):143-153. https://doi.org/10.1042/BJ20110448.
8. Finke K, Plath K, Panzner S, et al. A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of *S. cerevisiae*. 1996;15(7):1482-1494.
9. Sommer T, Jentsch S. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature*. 1993;365(6442):176-179. https://doi.org/10.1038/365176a0.
10. Stefanovic S, Hegde RS. Identification of a targeting factor for post-translational membrane protein insertion into the ER. *Cell*. 2007;128(6):1147-1159. https://doi.org/10.1016/j.cell.2007.01.036.
11. Favaloro V, Spasic M, Schwappach B, Dobberstein B. Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J Cell Sci*. 2008;121(11):1832-1840. https://doi.org/10.1242/jcs.020321.
12. Schuldiner M, Metz J, Schmid V, et al. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell*. 2008;134(4):634-645. https://doi.org/10.1016/j.cell.2008.06.025.
13. Srivistava R, Zalisko BE, Keenan RJ, Howell SH. The GET system inserts the tail-anchored SYP72 protein into endoplasmic reticulum membranes. *Plant Physiol*. December 2016;2016:00928-01145. https://doi.org/10.1104/pp.16.00928.
14. Xing S, Mehlhorn DG, Wallmeroth N, et al. Loss of GET pathway orthologs in *Arabidopsis thaliana* causes root hair growth defects and affects SNARE abundance. *Proc Natl Acad Sci*. 2017;114(8):E1544-E1553. https://doi.org/10.1073/pnas.1619525114.
15. Mateja A, Keenan RJ. A structural perspective on tail-anchored protein biogenesis by the GET pathway. *Curr Opin Struct Biol*. 2018;51:195-202. https://doi.org/10.1016/j.sbi.2018.07.009.
16. Mock J-Y, Chartron JW, Zaslaver M, Xu Y, Ye Y, Clemons WM. Bag6 complex contains a minimal tail-anchor–targeting module and a mock BAG domain. *Proc Natl Acad Sci*. 2015;112(1):106-111. https://doi.org/10.1073/pnas.1402745112.
17. Mock J-Y, Xu Y, Ye Y, Clemons WM. Structural basis for regulation of the nucleo-cytoplasmic distribution of Bag6 by TRC35. *Proc Natl Acad Sci*. 2017;114(44):11679-11684. https://doi.org/10.1073/pnas.1702940114.
18. Anghel SA, McGilvray PT, Hegde RS, Keenan RJ. Identification of Oxa1 homologs operating in the eukaryotic endoplasmic reticulum. *Cell Rep*. 2017;21(13):3708-3716. https://doi.org/10.1016/j.celrep.2017.12.006.

19. Wang F, Chan C, Weir NR, Denic V. The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature*. 2014;512(7515):441-444. https://doi.org/10.1038/nature13471.

20. Vilardi F, Lorenz H, Dobberstein B. WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. *J Cell Sci*. 2011;124(8):1301-1307. https://doi.org/10.1242/jcs.084277.

21. Yamamoto Y, Sakisaka T. Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. *Mol Cell*. 2012;48(3):387-397. https://doi.org/10.1016/j.molcel.2012.08.028.

22. Metz J, Wachter A, Schmidt B, Bujnicki JM, Schwappach B. The yeast Arr4p ATPase binds the chloride transporter Gef1p when copper is available in the cytosol. *J Biol Chem*. 2006;281(1):410-417. https://doi.org/10.1074/jbc.M507481200.

23. Mukhopadhyay R, Ho Y-S, Swiatek PJ, Rosen BP, Bhattacharjee H. Targeted disruption of the mouse *Asna1* gene results in embryonic lethality. *FEBS Lett*. 2006;580(16):3889-3894. https://doi.org/10.1016/j.febslet.2006.06.017.

24. Voth W, Schick M, Gates S, et al. The protein targeting factor Get3 functions as ATP-independent chaperone under oxidative stress conditions. *Mol Cell*. 2014;56(1):116-127. https://doi.org/10.1016/j.molcel.2014.08.017.

25. Mateja A, Szlachcic A, Downing ME, et al. The structural basis of tail-anchored membrane protein recognition by Get3. *Nature*. 2009;461(7262):361-366. https://doi.org/10.1038/nature08319.

26. Chen C-M, Misra TK, Silver S, Rosen BP. Nucleotide sequence of the structural genes for an anion pump. *J Biol Chem*. 1986;261(32):15030-15038.

27. Zhou T, Radaev S, Rosen BP, Gatti DL. Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump. *EMBO J*. 2000;19(17):4838-4845. https://doi.org/10.1093/emboj/19.17.4838.

28. Bange G, Sinning I. SIMIBI twins in protein targeting and localization. *Nat Struct Mol Biol*. 2013;20(7):776-780. https://doi.org/10.1038/nsmb.2605.

29. Zhou T, Radaev S, Rosen BP, Gatti DL. Conformational changes in four regions of the *Escherichia coli* ArsA ATPase link ATP hydrolysis to ion translocation. *J Biol Chem*. 2001;276(32):30414-30422. https://doi.org/10.1074/jbc.M103671200.

30. Leipe DD, Wolf YI, Koonin EV, Aravind L. Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol*. 2002;317(1):41-72. https://doi.org/10.1006/jmbi.2001.5378.

31. Mateja A, Paduch M, Chang H-Y, et al. Structure of the Get3 targeting factor in complex with its membrane protein cargo. *Science*. 2015;347(6226):1152-1155. https://doi.org/10.1126/science.1261671.

32. Sherrill J, Mariappan M, Dominik P, Hegde RS, Keenan RJ. A conserved archaeal pathway for tail-anchored membrane protein insertion. *Traffic*. 2011;12(9):1119-1123. https://doi.org/10.1111/j.1600-0854.2011.01229.x.

33. Suloway CJM, Rome ME, Clemons WM. Tail-anchor targeting by a Get3 tetramer: the structure of an archaeal homologue. *EMBO J*. 2012;31(3):707-719.

34. Chartron JW, Clemons WM, Suloway CJM. The complex process of GETting tail-anchored membrane proteins to the ER. *Curr Opin Struct Biol*. 2012;22(2):217-224. https://doi.org/10.1016/j.sbi.2012.03.001.

35. Hochberg GKA, Ecroyd H, Liu C, et al. The structured core domain of B-crystallin can prevent amyloid fibrillation and associated toxicity. *Proc Natl Acad Sci*. 2014;111(16):E1562-E1570. https://doi.org/10.1073/pnas.1322673111.

36. Chio US, Chung S, Weiss S, Shan S. A chaperone lid ensures efficient and privileged client transfer during tail-anchored protein targeting. *Cell Rep*. 2019;26(1):37-44.e7. https://doi.org/10.1016/j.celrep.2018.12.035.

37. Kurdi-Haidar B, Heath D, Aebi S, Howell SB. Biochemical characterization of the human arsenite-stimulated ATPase (hASNA-I). *J Biol Chem*. 1998;273(35):22173-22176. https://doi.org/10.1074/jbc.273.35.22173.

38. Bozkurt G, Stjepanovic G, Vilardi F, et al. Structural insights into tail-anchored protein binding and membrane insertion by Get3. *Proc Natl Acad Sci*. 2009;106(50):21131-21136. https://doi.org/10.1073/pnas.0910223106.

39. Lütcke H, High S, Römisch K, Ashford AJ, Dobberstein B. The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. *EMBO J*. 1992;11(4):1543-1551.

40. Borgese N, Righi M. Remote origins of tail-anchored proteins. *Traffic*. 2010;11(7):877-885. https://doi.org/10.1111/j.1600-0854.2010.01068.x.

41. Wösten MMSM, Kox LFF, Chamnongpol S, Soncini FC, Groisman EA. A signal transduction system that responds to extracellular iron. *Cell*. 2000;103(1):113-125. https://doi.org/10.1016/S0092-8674(00)00092-1.

42. Stearman R, Yuan DS, Yamaguchi-Iwai Y, Klausner RD, Dancis A. A permease-oxidase complex involved in High-affinity iron uptake in yeast. *Science*. 1996;271(5255):1552-1557. https://doi.org/10.1126/science.271.5255.1552.

43. Formighieri C, Cazzaniga S, Kuras R, Bassi R. Biogenesis of photosynthetic complexes in the chloroplast of *Chlamydomonas reinhardtii* requires ARSA1, a homolog of prokaryotic arsenite transporter and eukaryotic TRC40 for guided entry of tail-anchored proteins. *Plant J*. 2013;73(5):850-861. https://doi.org/10.1111/tpj.12077.

44. Peschke M, Le Goff M, Koningstein GM, et al. SRP, FtsY, DnaK and YidC are required for the biogenesis of the *E. coli* tail-anchored membrane proteins DjlC and Flk. *J Mol Biol*. 2018;430(3):389-403. https://doi.org/10.1016/j.jmb.2017.12.004.

45. Kriventseva EV, Kuznetsov D, Tegenfeldt F, et al. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res*. 2019;47(D1):D807-D811. https://doi.org/10.1093/nar/gky1053.

46. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25(9):1189-1191. https://doi.org/10.1093/bioinformatics/btp033.

47. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res*. 2015;43(W1):W389-W394. https://doi.org/10.1093/nar/gkv332.

48. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. *Genome Res*. 2004;14(6):1188-1190. https://doi.org/10.1101/gr.849004.

49. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-38. https://doi.org/10.1016/0263-7855(96)00018-5.

50. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 2018;35(6):1547-1549. https://doi.org/10.1093/molbev/msy096.

51. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*. 2001;18(5):691-699. https://doi.org/10.1093/oxfordjournals.molbev.a003851.

52. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J Mol Biol*. 2001;305(3):567-580. https://doi.org/10.1006/jmbi.2000.4315.

53. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785-786. https://doi.org/10.1038/nmeth.1701.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

Title: **The Natural History of Get3-like Chaperones**

Supplemental Information

Ákos Farkas[1], Evelina Ines De Laurentiis[1], Blanche Schwappach[1,2]

1   Universitätsmedizin Göttingen Institut für Molekularbiologie Humboldtallee 23 D-37073 Göttingen, Germany

2   Max-Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany

Supplemental materials

Table S1: Relates to Table 1. Contains the identifiers of the sequences based on which Table 1 was created. Column E marks the sequences used for the phylogenetic tree in Figure 1.

Table S2: Relates to Figure 4B and Figure 4C. Contains the identifiers of the sequences Figure 4B and 4C refers to.

Table S3: Relates to Figure 4D. Contains the corresponding data in bacterial phyla not shown in Figure 4D.

Table S4: Related to Figure 5E. Contains the data and the names of the taxons used in Figure 4E.

Supplemental Table Legends

Table S1: List of Get3 homologs used in the current study. The classification of each homolog based on the groups listed in Table 1 and the species they are found in are provided.

Table S2: List of Get3 homologs with a potential plastidial or mitochondrial localization used in the current study. The classification of each homolog based on the groups listed in Table 1 and their predicted localization based on TargetP 1.1 are provided. Abbreviations of predictions: C – chloroplast; M – mitochondria; - – no organelle predicted

Table S3: Distribution of the number of methionine residues in the region homologous to the sequence from helix 4 to 9 in *Sc*Get3 among bacteria containing Get3 homologs with an α-crystallin domain.

Table S4: Analysis of presence or absence of different types of Get3 or ArsA homologues, number of proteins, membrane proteins, and tail-anchored proteins in the indicated bacterial taxa.

Table S2

Group based on Table 1	Uniprot ID	Taxonomy	TargetP 1.1 prediction	Group based on Table 1	Uniprot ID	Taxonomy	TargetP 1.1 prediction
Group 4	D8UKL0_VOLCA	Chlorophyta	C	Group 3	A0A0D2RYE9_GOSRA	Embryophyta	M
Group 4	K8F9L6_9CHLO	Chlorophyta	C	Group 3	A0A059BHV7_EUCGR	Embryophyta	C
Group 4	A0A090M8H3_OSTTA	Chlorophyta	M	Group 3	A0A059AC89_EUCGR	Embryophyta	M
Group 4	A4S2F3_OSTLU	Chlorophyta	-	Group 3	V4SPJ7_9ROSI	Embryophyta	-
Group 4	C1EGU4_MICCC	Chlorophyta	-	Group 3	A0A1U7ZL42_NELNU	Embryophyta	C
Group 4	C1MP51_MICPC	Chlorophyta	M	Group 3	A0A200R5Q6_9MAGN	Embryophyta	C
Group 4	E1Z6I9_CHLVA	Chlorophyta	-	Group 3	A0A2G5D279_AQUCA	Embryophyta	C
Group 4	I0YX28_COCSC	Chlorophyta	M	Group 3	A0A2H3XHX9_PHODC	Embryophyta	C
Group 3	M1VA07_CYAM1	Rhodophyta	M	Group 3	A0A2H3YRJ4_PHODC	Embryophyta	M
Group 3	M2XZS1_GALSU	Rhodophyta	-	Group 3	A0A2H9ZUZ8_9ASPA	Embryophyta	M
Group 3	R7QLH0_CHOCR	Rhodophyta	-	Group 3	I1IFB4_BRADI	Embryophyta	M
Group 3	A0A176VLW2_MARPO	Embryophyta	C	Group 3	I1IYH4_BRADI	Embryophyta	M
Group 3	D8SQW5_SELML	Embryophyta	M	Group 3	J3LH15_ORYBR	Embryophyta	M
Group 3	D8RTG3_SELML	Embryophyta	C	Group 3	C5XYT4_SORBI	Embryophyta	M
Group 3	W1NHA6_AMBTC	Embryophyta	C	Group 3	C5Y9U4_SORBI	Embryophyta	M
Group 3	A0A2G9GYH7_9LAMI	Embryophyta	C	Group 3	B4FXM9_MAIZE	Embryophyta	M
Group 3	A0A2G9HKT1_9LAMI	Embryophyta	M	Group 3	B6TZB7_MAIZE	Embryophyta	M
Group 3	S8E2Y5_9LAMI	Embryophyta	M	Group 3	K3YT13_SETIT	Embryophyta	M
Group 3	A0A022Q8P4_ERYGU	Embryophyta	C	Group 3	A0A1E5WCP2_9POAL	Embryophyta	M
Group 3	A0A022RIQ3_ERYGU	Embryophyta	C	Group 3	M0TT67_MUSAM	Embryophyta	C
Group 3	A0A1J6ITG7_NICAT	Embryophyta	C	Group 3	A0A0K9PK69_ZOSMR	Embryophyta	C
Group 3	A0A1U7XLU8_NICSY	Embryophyta	C	Group 2	A0A250XGS9_9CHLO	Chlorophyta	M
Group 3	A0A1U8F339_CAPAN	Embryophyta	C	Group 2	A0A2J8A6T7_9CHLO	Chlorophyta	M
Group 3	A0A1S3Z6A3_TOBAC	Embryophyta	C	Group 2	ASNA1_CHLRE	Chlorophyta	M
Group 3	A0A1U7XLE1_NICSY	Embryophyta	C	Group 2	A0A150GLK6_GONPE	Chlorophyta	C
Group 3	A0A1S3X3L0_TOBAC	Embryophyta	M	Group 2	A0A087SRG6_AUXPR	Chlorophyta	M
Group 3	A0A1S4CJ63_TOBAC	Embryophyta	C	Group 2	A0A2P6TQD2_CHLSO	Chlorophyta	C
Group 3	K4DAD9_SOLLC	Embryophyta	C	Group 2	M2XUU5_GALSU	Rhodophyta	M
Group 3	M1AE76_SOLTU	Embryophyta	M	Group 2	M1UWD0_CYAM1	Rhodophyta	-
Group 3	M1AE77_SOLTU	Embryophyta	M	Group 2	A0A1X6PKT4_PORUM	Rhodophyta	C
Group 3	A0A0A0KWR9_CUCSA	Embryophyta	C	Group 2	R7QK77_CHOCR	Rhodophyta	C
Group 3	A0A1S3BFJ9_CUCME	Embryophyta	C	Group 5	I1ID51_BRADI	Embryophyta	C
Group 3	A0A2I4FAM5_JUGRE	Embryophyta	C	Group 5	J3LGE0_ORYBR	Embryophyta	C
Group 3	A0A2I4GLQ3_JUGRE	Embryophyta	-	Group 5	A3AAM6_ORYSJ	Embryophyta	C
Group 3	A0A2I4GLR1_JUGRE	Embryophyta	C	Group 5	K3YSP5_SETIT	Embryophyta	M
Group 3	A0A1S2YK80_CICAR	Embryophyta	M	Group 5	C5XZY5_SORBI	Embryophyta	M
Group 3	A0A1S3E4U6_CICAR	Embryophyta	C	Group 5	M7YIF9_TRIUA	Embryophyta	-
Group 3	A0A072UVY6_MEDTR	Embryophyta	C	Group 5	A0A1D6HI20_MAIZE	Embryophyta	C
Group 3	G7I505_MEDTR	Embryophyta	C	Group 5	A0A2I0AHV9_9ASPA	Embryophyta	M
Group 3	A0A072VNY3_MEDTR	Embryophyta	C	Group 5	A0A2H3XRX0_PHODC	Embryophyta	C
Group 3	A0A2K3NMY2_TRIPR	Embryophyta	M	Group 5	A0A0K9NU83_ZOSMR	Embryophyta	C
Group 3	A0A2K3LCV4_TRIPR	Embryophyta	M	Group 5	A0A1J6IDG8_NICAT	Embryophyta	C
Group 3	I1LCZ2_SOYBN	Embryophyta	C	Group 5	A0A1S4CYI1_TOBAC	Embryophyta	C
Group 3	I1NHH3_SOYBN	Embryophyta	C	Group 5	A0A1S3YPY0_TOBAC	Embryophyta	C
Group 3	A0A1S3V658_VIGRR	Embryophyta	C	Group 5	A0A1U8ETJ7_CAPAN	Embryophyta	C
Group 3	V7BGY9_PHAVU	Embryophyta	C	Group 5	A0A2G3B6J8_CAPCH	Embryophyta	C
Group 3	B9SY64_RICCO	Embryophyta	C	Group 5	K4BXG6_SOLLC	Embryophyta	C
Group 3	A0A067LJE1_JATCU	Embryophyta	C	Group 5	D7KQU0_ARALL	Embryophyta	C
Group 3	A0A067LHD2_JATCU	Embryophyta	M	Group 5	A0A178WMJ2_ARATH	Embryophyta	C
Group 3	A9PGI6_POPTR	Embryophyta	C	Group 5	A0A087HR01_ARAAL	Embryophyta	C
Group 3	A0A1Q3B3K7_CEPFO	Embryophyta	C	Group 5	R0GLF8_9BRAS	Embryophyta	C
Group 3	A0A1Q3B3X1_CEPFO	Embryophyta	M	Group 5	A0A1S2Z867_CICAR	Embryophyta	C
Group 3	A0A2P5D3Y8_TREOI	Embryophyta	C	Group 5	V4K9A9_EUTSA	Embryophyta	C
Group 3	W9SYC9_9ROSA	Embryophyta	C	Group 5	K7K7G9_SOYBN	Embryophyta	C
Group 3	W9S2I5_9ROSA	Embryophyta	M	Group 5	A0A1J7IV84_LUPAN	Embryophyta	C
Group 3	M5X6A8_PRUPE	Embryophyta	C	Group 5	G7JX27_MEDTR	Embryophyta	C
Group 3	A0A087H8G3_ARAAL	Embryophyta	C	Group 5	V7C5C5_PHAVU	Embryophyta	C
Group 3	D7L8X8_ARALL	Embryophyta	C	Group 5	A0A2K3PL22_TRIPR	Embryophyta	C
Group 3	F4J3Q8_ARATH	Embryophyta	C	Group 5	M5XS14_PRUPE	Embryophyta	C
Group 3	A0A0D3AER6_BRAOL	Embryophyta	C	Group 5	W9R3Z1_9ROSA	Embryophyta	M
Group 3	V4M2J4_EUTSA	Embryophyta	C	Group 5	A0A2P5DQ30_PARAD	Embryophyta	M
Group 3	A0A087GCJ7_ARAAL	Embryophyta	M	Group 5	A0A2P5ENF6_TREOI	Embryophyta	M
Group 3	D7MUD3_ARALL	Embryophyta	M	Group 5	A0A1R3J603_COCAP	Embryophyta	C
Group 3	Q5XF80_ARATH	Embryophyta	M	Group 5	A0A1R3KIF2_9ROSI	Embryophyta	C
Group 3	A0A0D3AKN9_BRAOL	Embryophyta	M	Group 5	S8DUW4_9LAMI	Embryophyta	-
Group 3	R0EWE9_9BRAS	Embryophyta	M	Group 5	A0A0D2PSL0_GOSRA	Embryophyta	C
Group 3	V4N939_EUTSA	Embryophyta	M	Group 5	A0A067KAR8_JATCU	Embryophyta	C
Group 3	A0A1R3HAK3_COCAP	Embryophyta	-	Group 5	W1P8W2_AMBTC	Embryophyta	C
Group 3	A0A1R3J9V5_9ROSI	Embryophyta	-	Group 5	A0A2G5D7X1_AQUCA	Embryophyta	C
Group 3	A0A2I0IRW4_PUNGR	Embryophyta	M	Group 5	A0A1Q3CBM7_CEPFO	Embryophyta	C
Group 3	A0A2I0J6S8_PUNGR	Embryophyta	M	Group 5	A0A1S3BET7_CUCME	Embryophyta	C
Group 3	A0A0B0N3P8_GOSAR	Embryophyta	C	Group 5	A0A059BF85_EUCGR	Embryophyta	C
Group 3	A0A0D2SBJ4_GOSRA	Embryophyta	C	Group 5	A0A2I4FWX0_JUGRE	Embryophyta	C
Group 3	A0A1U8N7T4_GOSHI	Embryophyta	C	Group 5	A0A200Q3L2_9MAGN	Embryophyta	C
Group 3	A0A0B0PPN1_GOSAR	Embryophyta	C	Group 5	A0A176VIM6_MARPO	Embryophyta	C
Group 3	A0A0D2RS72_GOSRA	Embryophyta	C	Group 5	A0A1U8A7V7_NELNU	Embryophyta	C
Group 3	A0A1U8KDC9_GOSHI	Embryophyta	C	Group 5	D8QQV2_SELML	Embryophyta	-
				Group 5	D8R829_SELML	Embryophyta	-

C: chloroplast; M: mitochondria; -: no organelle predicted

Table S3

Alpha-crystallin domain containing Get3 homologs in bacteria

Bacterial phyla	Number of Met residues between helix 4 and 9											Total
	0	1	2	3	4	5	6	7	8	9	11	
Acidobacteria							2					2
Actinobacteria	5	18	19	12	5	3						62
Aquificae				2	2	1	1					6
Bacteroidetes						2	1					3
Chlorobi				1	5	11	17	1	37	6		78
Chloroflexi			2	4		2	1		2	3		14
Cyanobacteria			3	1	24	16	30					74
Firmicutes			1	3	8	4	5	1	21	3	1	47
Fusobacteria				1	7							8
Proteobacteria		1			1	1		1				4
Spirochaetes			1		1							2
Thermodesulfobacteria					1							1
Total	5	19	26	24	54	40	57	3	60	12	1	301

Table S4

Uniprot taxon ID	Phylum	Get3 homolog with TRC40-insert	ArsA	Comment	Total proteins	Membrane proteins	TA proteins
Algibacter alginicilyticus	Bacteroidetes		yes		3236	741	25
Aquifex aeolicus	Aquificae	yes			1560	318	9
Bacillus cereus (strain ZK / E33L)	Firmicutes	yes	yes		5639	1688	48
Bacteroides thetaiotaomicron (strain ATCC 29148 / DSM 2079 / NCTC 10582 / E50 / VPI-5482)	Bacteroidetes		yes		4783	1176	37
Burkholderia multivorans (strain ATCC 17616 / 249)	Proteobacteria		yes		6043	1439	48
Chloracidobacterium thermophilum (strain B)	Acidobacteria	yes			3054	736	25
Chlorobaculum tepidum (strain ATCC 49652 / DSM 12025 / NBRC 103806 / TLS) (Chlorobium tepidum)	Chlorobi	yes			2251	464	42
Clostridium neonatale	Firmicutes		yes		4004	1024	19
Cutibacterium granulosum DSM 20700	Actinobacteria		yes		5074	1228	37
Desulfallas gibsoniae DSM 7213	Firmicutes	yes			4314	938	61
Desulfobacca acetoxidans (strain ATCC 700848 / DSM 11109 / ASRB2)	Proteobacteria	yes		double domain Get3 homolog	2862	714	30
Fusobacterium nucleatum subsp. nucleatum (strain ATCC 25586 / CIP 101130 / JCM 8532 / LMG 13131)	Fusobacteria	yes			2050	479	37
Klebsiella pneumoniae subsp. pneumoniae (strain ATCC 700721 / MGH 78578)	Proteobacteria		yes		5127	1245	43
Lactobacillus plantarum (strain ATCC BAA-793 / NCIMB 8826 / WCFS1)	Firmicutes		yes		3088	857	21
Leptospirillum ferrooxidans (strain C2-3)	Nitrospirae		yes		2414	583	17
Lysinibacillus macroides	Firmicutes		yes		4250	1097	26
Magnetospirillum caucaseum	Proteobacteria		yes		4651	873	37
Megasphaera cerevisiae DSM 20462	Firmicutes		yes		2843	622	36
Methylobacillus flagellatus (strain KT / ATCC 51484 / DSM 6875)	Proteobacteria		yes		2610	661	15
Myxococcus fulvus (strain ATCC BAA-855 / HW-1)	Proteobacteria	yes		double domain Get3 homolog	7269	1308	54
Nocardia farcinica (strain IFM 10152)	Actinobacteria	yes			5944	1267	35
Nostoc punctiforme (strain ATCC 29133 / PCC 73102)	Cyanobacteria	yes		double domain Get3 homolog	6608	1542	79
Pirellula staleyi (strain ATCC 27377 / DSM 6068 / ICPB 4128) (Pirella staleyi)	Planctomycetes		yes		4712	1304	46
Planctopirus limnophila (strain ATCC 43296 / DSM 3776 / IFAM 1008 / 290) (Planctomyces limnophilus)	Planctomycetes		yes		4259	1122	43
Planococcus plakortidis	Firmicutes	yes			3056	843	11
Roseiflexus castenholzii (strain DSM 13941 / HLO8)	Chloroflexi	yes			4325	1211	29
Salinicoccus halodurans	Firmicutes	yes	yes		2634	678	15
Treponema caldarium (strain ATCC 51460 / DSM 7334 / H1) (Spirochaeta caldaria)	Spirochaetes	yes			2755	755	9

Chapter Three

Get3 chaperones diverse classes of membrane proteins before and after membrane targeting

Manuscript in preparation.

See next page for authors and their contributions.

Authors and their contributions

Figure	Author	Contribution
Fig 1A-B, Fig S1	Ákos Farkas	Crossed the library with the query strain, carried out all microscopy, evaluated images, visualized results
Fig 1C-D	Ákos Farkas	Created query strains, crossed SWAT strains with query strains, carried out microscopy, created script to analyze images
Fig 1E	Ákos Farkas	Visualized the scheme
Fig 1F-G	Anne Clancy	SGA crossed SWAT strains with query strain, carried out microscopy, evaluated images
Fig S2	Anne Clancy	Performed experiment
Fig 1F-G	Ákos Farkas	Evaluated images, visualized results
Fig 2A, Fig S3A,B	Ákos Farkas	Executed protocol up until digest of samples, silver stained the gel, visualized the results
Fig 2A, Fig S3B	Jakob Bader	Performed mass spectrometry analysis, processed and statistically analysed raw data
Fig 2B-C	Ákos Farkas	Created modified library, carried out microscopy, evaluated images
Fig 2D	Ákos Farkas	Carried out spotting assay
Fig 2E, Fig S4	Ákos Farkas	Prepared samples for lipidomic analysis, analyzed results from mass spectrometry of samples
Fig 2E, Fig S4	Lipotype GmbH	Performed mass spectrometry analysis of samples
Fig 3A-B Fig S5	Dr. Kathrin Ulrich	Produced indicated proteins, carried out measurements
Fig 3A-B	Ákos Farkas	Visualized results
Fig 3C	Ákos Farkas	Carried out spotting assay
Fig 3D-E	Ákos Farkas	Created modified library, carried out microscopy, evaluated images
Fig 4A-E	Ákos Farkas	Processed samples, carried out microscopy, visualized results
Fig 5A-I	Ákos Farkas	Processed samples, carried out microscopy, visualized results
Fig 6A-E	Ákos Farkas	Created list of TA proteins and carried out bioinformatic analysis, created strains, processed samples, carried out microscopy, visualized results
Fig 7A-F, Fig S5-6	Ákos Farkas	Processed samples, visualized results
	Anne Clancy, Ákos Farkas	Created strains and plasmids as indicated in Table 2, 5.

1. Introduction

Newly synthesized membrane proteins require targeting pathways to reach their specific membrane destination. Due to the hydrophobic nature of the transmembrane segments (TMS), they are prone to aggregate in the cytosol and therefore require chaperoning during their delivery to their target membrane. Many membrane proteins are co-translationally inserted into the membrane with the help of the signal recognition particle (SRP) pathway (Akopian et al., 2013). In addition, the evolutionarily conserved GET pathway is one of the major routes that non-mitochondrial TA proteins can take to reach the membrane (Schuldiner et al., 2008; Stefanovic and Hegde, 2007). However, other pathways such as the recently identified SRP-independent (SND) pathway are also known to be able to deliver them to their destination (Aviram et al., 2016).

The first element of the GET pathway to make contact with TA proteins is a pretargeting complex comprised of Sgt2, Get5 and Get4. They recognize and hand over TA substrates to the pathway's central component, Get3, an ATPase which chaperones TA proteins and inserts them into the membrane via its receptor complex made up of Get1 and Get2 (Mateja and Keenan, 2018). Although the Get3 dependence of individual TA proteins have been addressed in several studies (Guna et al., 2018; Schuldiner et al., 2008), the *in-vivo* client spectrum of Get3 remains insufficiently understood. Some studies approached the question *in vivo* systematically either in yeast (Li et al., 2019) or in mammalian cells (Coy-Vergara et al., 2019). Still, even these focused solely on TA proteins, although it has become clear that Get3 can handle other substrates as well, such as GPI-anchored (Ast et al., 2013) and short secretory proteins (Johnson et al., 2012).

In vitro, Get3 can also acquire general chaperone activity upon oxidation and prevent the aggregation of chemically denatured substrates (Voth et al., 2014). Unlike its TA-targeting activity (Stefanovic and Hegde, 2007), this function of Get3 is independent of its ability to hydrolyze ATP. Interestingly, Get3 homologs are found in all domains of life even in species where it is not expected to be responsible for TA delivery to membranes (Farkas et al., 2019). Therefore it is possible that Get3 has a more ancient chaperone function not restricted to TA protein insertion.

In this study, we conducted systematic visual screens using budding yeast to delineate the range of possible substrates of Get3. We provide biochemical evidence that

Get3 interacts with and probably chaperones novel classes of membrane proteins. Finally, we also show that Sed5 is a unique TA protein with which Get3 interacts before and after its membrane insertion.

2. Results

2.1. Loss of Get3 affects a diverse range of proteins directly and indirectly

Systematic microscopy-based studies in yeast have been limited by the fact that previously available yeast libraries used C-terminal protein tags, which can perturb the localization of several groups of proteins, including GPI-anchored and TA proteins. The recently published SWAT library, which contains N-terminally GFP-tagged strains, circumvents these issues and enables the study of proteins carrying C-terminal targeting information (Weill et al., 2018). Therefore, to uncover novel targeting substrates of Get3, we knocked out GET3 in the secretome subset of the SWAT library, which covers 95 % of the predicted membrane proteins of the secretory pathway (Yofe et al., 2016). We compared the resulting strains with the unmodified library using automated microscopy.

Surprisingly, our screen revealed a wide range of proteins affected in *get3* strains, most of which were not TA proteins (Fig 1A). We observed three major phenotypes: proteins with decreased or diffuse signal in *get3* strains, proteins with increased signal in *get3* strains, or proteins that had a different cellular localization in *get3* strains compared to wt (Fig 1B, Fig S1). Besides the previously characterized Get3 substrate Sed5 (Schuldiner et al., 2008), we found several novel TA proteins that became diffuse in the *get* mutant strains, such as Coy1, Lam5, Mga2, Mtc4. Some TA proteins appeared in mitochondria (Far10, Frt1), which also confirms previous results (Li et al., 2019; Schuldiner et al., 2008). Interestingly, the TA protein Mps2, which is normally located in the ER and forms a single punctum, formed multiple foci in *get3* cells (Fig 1B). The normally punctate localization pattern of the TA protein Bos1 changed to an occasionally formed brighter punctum (Fig 1B), and the ER-mitochondrial TA protein Hfd1 showed brighter mitochondrial staining in *get3* cells (Fig S1A).

Besides the TA proteins, several other membrane proteins were affected. Some of these resembled TA proteins because they had two or three TMSs close to the C-terminus (C'-TMS, Fig 1A), but others had no obvious unifying feature in their topology or

had no predicted TMS at all. Additionally, we found most peroxisomal proteins affected by the loss of Get3 (Table 1). However, due to the complex nature of peroxisomal biogenesis and its dependence on several cellular processes, we decided not to explore the causes of the observed effects of Get3 on these proteins further. We also found mitochondrial proteins mislocalized by the N-terminal GFP tag affected in *get3* cells as expected (Table 1, Vitali et al., 2018). We noted that loss of Get3 had an effect on some proteins with a predicted luminal N-terminus whose biogenesis was likely impaired due to the N-terminal GFP tag (Table 1). Since we intended to focus on the native client spectrum of Get3 in this study and both of these groups could plausibly represent artefacts of the N-terminal GFP tagging, we did not analyse these proteins any further.

The current model of the GET pathway implies that Get3 can only target TA proteins with the help of the pretargeting complex. Interestingly, Sed5 shows a graded response depending on whether Sgt2, Get4/Get5, or Get3 is knocked out (Battle et al., 2010; Jonikas et al., 2009; Kohl et al., 2011), and we wondered whether other TA proteins also follow a similar pattern. To test this, we knocked out each cytosolic GET pathway component in SWAT strains expressing GFP-Sed5, GFP-Coy1, or GFP-Mtc4, and quantified the distribution of the signal intensity of the pixels in the cells of the different strains. Compared to a diffuse cytosolic signal, strong puncta skew the distribution of pixel signal intensity in cells to the right, resulting in high skewness. Since all three proteins show a punctate pattern in wt that becomes diffuse in *get3* strains, we used the skewness of the distribution of pixel intensities to approximate how punctate the localization pattern was. We confirm that Sed5 showed a graded response to the loss of cytosolic GET pathways components, and Coy1, Mtc4 showed a similar pattern as well (Fig 1C). We wondered whether proteins that change their localization in *get3* cells but did not become diffuse behaved similarly. We quantified two such proteins, Opi3 and Mps2, both of which form multiple puncta in the absence of Get3 (Fig 1B). Like Sed5, these proteins also showed a graded response depending on the component of the pretargeting complex knocked out (Fig 1D). Taken together, these results indicate that many of the proteins affected in *get3* cells are also affected by the loss of upstream GET pathway components. However, they also hint at the possibility that the presence of Sgt2 is not an obligate requirement for the functionality of the pathway *in vivo*, as the extent of the effect of loss of Sgt2 varies highly between the different substrates tested.

Being affected by the loss of Get3 does not necessarily mean that a protein interacts with or is the substrate of the GET pathway. Since several TA proteins are essential, their impaired membrane targeting could cause secondary effects observed upon loss of Get3 in case of non-TA proteins. Several of the proteins we found affected in *get3* cells are involved in Golgi function or are related to lipid metabolism (marked with red and blue in Fig 1A, respectively). We cannot exclude the possibility that the perturbation of one or some key proteins in the given category are responsible for the effects observed in the others. For example, Sed5 is an essential protein that is vital for proper Golgi function (Hardwick and Pelham, 1992) and is one of the best characterized clients of Get3 (Powis et al., 2013; Schuldiner et al., 2008). To test the possibility that the impaired biogenesis of Sed5 causes the effects seen in other Golgi proteins, we made use of the auxin-induced degradation (AID) system (Nishimura et al., 2009). In short, the AID-method involves tagging a protein of interest with a small degron (AID-tag), which induces the ubiquitylation and subsequent degradation of the target protein by an exogenous E3-ligase, Tir1, in the presence of indole-3-acetic acid (IAA, Fig 1E). We created a strain containing N-terminally AID-tagged Sed5 and Tir1, and crossed it with several SWAT library strains carrying the GFP-tag on Get3, Golgi protein hits and Ubx3, an integral membrane protein that formed foci in *get3* cells. Upon addition of auxin, we observed a rapid degradation of Sed5 (Fig S2), which resulted in approximately 90% decrease in Sed5 levels in between 3 to 5 hours. We therefore imaged the strains without added IAA and four hours after addition of IAA, and concluded that several Golgi proteins replicated the phenotype observed in *get3* cells (Fig 1F), including the TA protein Bos1. On the other hand, Ubx3 and Coy1 did not react to the loss of Sed5 (Fig 1G). These results indicate that TA and non-TA proteins affected by the loss of Get3 can indeed be secondary hits downstream of primary substrates of Get3 such as Sed5. However, others are independent Sed5 and are either related to other primary hits or are primary hits themselves.

2.2. A diverse set of membrane proteins colocalizes with a substrate-trapping mutant of Get3

Since we observed that knocking out Get3 affected several proteins indirectly and other pathways can compensate for and suppress the effects of the loss of Get3 (Aviram

et al., 2016), we sought to employ a method that allows us to identify clients that can but do not have to use Get3. Get3 D57E, which is an ATPase-deficient mutant, forms foci in yeast and traps the TA protein Sed5 (Powis et al., 2013). Moreover, its human homolog, TRC40 D74E has been successfully used to trap clients in human cells (Coy-Vergara et al., 2019). Although it would be expected based on current models of the GET pathway (Gristick et al., 2014), TRC40 D74E did not enrich the pretargeting complex, probably due to its overexpression (Coy-Vergara et al., 2019). To test whether yeast Get3 D57E indeed interacts more strongly than Get3 with the pretargeting complex, we expressed cytosolic GFP, Get3-GFP wt, and Get3-GFP D57E from a plasmid and performed an anti-GFP immunoprecipitation in the presence of detergent (Fig S3A). Mass spectrometry (MS) analysis of the samples revealed that the membrane receptor complex of Get3 comprised of Get1, Get2 co-immunoprecipitated equally efficiently with Get3 wt and D57E (Fig 2A, Fig S3B). We also found two further membrane proteins strongly enriched in both wt and D57E samples, Gsf2, Mnn1. Since they were much less enriched over background than Get1 or Get2, we hypothesize that these could represent low affinity membrane receptors of Get3, but the *in-vivo* relevance of this interaction remains unknown. Interestingly, the only TA protein that showed an enrichment with Get3 D57E over Get3 wt in the presence of detergent was Tlg2, although we could not verify this interaction in directed pulldown experiments using a tagged form of Tlg2 (data not shown). Lastly, the pretargeting complex made up of Get4, Get5, Sgt2 was clearly enriched in Get3 D57E samples, and Ybr137w, a proposed interactor of the pretargeting complex (Kohl et al., 2011; Wang et al., 2010), was only enriched above background with Get3 D57E. This confirms that under the physiological expression levels obtained in yeast, Get3 D57E indeed interacts with the pretargeting complex more strongly than Get3 wt. Combined with the fact that it forms cytosolic foci (Powis et al., 2013), we concluded it could be a suitable tool to discover novel substrates of Get3 *in vivo*.

Lack of the receptor components Get1 and Get2 also induces the formation of Get3 puncta (Schuldiner et al., 2008). To further enhance puncta formation by Get3 D57E, we knocked out *GET1*, *GET2* in addition to *GET3* in the subset of the SWAT library strains used in our *get3* screen. We then introduced mTagBFP2-Get3 D57E, a blue fluorescent protein tagged version of Get3 D57E (Subach et al., 2011), into this modified strain collection and confirmed that a model substrate TA protein, Pep12 colocalizes with Get3 D57E under these conditions (Fig 2B). Surprisingly, high-throughput microscopy of

the strain collection revealed that the majority of proteins affected by the loss of Get3 (Fig 1A) did not colocalize with Get3 D57E (Fig 2C). At the same time, we found 48 proteins in total that were not affected in *get3* cells but colocalized with Get3 D57E. Importantly, more than half of these proteins are non-TA membrane proteins that do not belong to any class of Get3 substrates described so far. Whether Get3 interacts with these non-TA proteins as part of the GET pathway or independently of it cannot be decided based on this assay. However, irrespective of the mode of interaction, these results suggest that Get3 traps and hence interacts with more membrane proteins than previously suggested.

2.3. Get3 plays a role in ergosterol synthesis in budding yeast

As observed for proteins affected in *get3* cells (Fig 1A), we have also found several proteins colocalizing with Get3 D57E that are involved in Golgi function or lipid metabolism (red and blue proteins in Fig 2C). Whereas proteins involved in Golgi function are TA proteins or their interaction partner (Sly1), we have found three genes involved in sterol metabolism, one of which is a TA protein (Erg9) and the other two are either monotopic or peripheral membrane proteins (Erg1, Erg27). Considering that these genes are essential in ergosterol biosynthesis (Bhattacharya et al., 2018) and one of them, the evolutionarily conserved Erg1, is the target of the drug terbinafine (Padyana et al., 2019; Ryder, 1992), we wondered whether loss of Get3 increases the terbinafine sensitivity of yeast cells. Indeed, we found that a *get3* strain had increased sensitivity to terbinafine considerably, whereas *get4*, *get5*, *sgt2* strains were unaffected (Fig 2D), which could indicate a novel role for Get3 in the ergosterol biosynthesis pathway.

Taking into consideration the proteins involved in lipid metabolism that were either affected in *get3* cells or colocalized with Get3 D57E, we wanted to determine whether the loss of Get3 induces changes in the lipid composition of yeast cells. To this end, we compared the lipid composition of wt and *get3* strains grown to the logarithmic phase either in rich or synthetic complete media. Although some classes of lipids were only changed in *get3* cells compared to wt in one of the two media (Fig S4), four classes showed statistically significant differences between wt and *get3* in both media (Fig 2E). The decreased ergosterol ester content in *get3* cells is in line with the above findings suggesting a role for Get3 as an interaction partner of some proteins in the ergosterol biosynthesis pathway. Similarly, the decreased triacylglycerol content of *get3* cells further

corroborates the idea that lipid droplets are affected in *get3* cells as evidenced by some lipid droplet proteins affected by the loss of Get3 (Lpl1, Pet10, Mtc4; Fig 1A). The increased phosphatidylcholine levels could be related to the observed mislocalization of Opi3 (Fig 1B), which is responsible for phosphatidylcholine synthesis (Kodaki and Yamashita, 1987). Lastly, the enzyme synthesizing inositol phosphorylceramide is located in the Golgi (Levine et al., 2000), and the perturbed Golgi function observed in *get3* cells could therefore also affect the synthesis of this lipid group as well.

2.4. An ATP-binding deficient mutant of Get3 accumulates further substrates

Get3's substrate binding is suggested to be enhanced when it is ATP-bound (Bozkurt et al., 2009; Mateja et al., 2009), which explains why the ATP hydrolysis deficient mutant Get3 D57E acts as a trapping mutant. We therefore wondered whether an ATP-binding deficient mutant would show decreased substrate binding *in vivo*. Furthermore, Get3's oxidation-induced chaperone activity is ATP independent, and we also wanted to test whether loss of ATP binding would affect its chaperone activity. Based on homology to similar ATPases, mutating the conserved Lys31 to alanine (K31A) should abolish ATP-binding (Saraste et al., 1990), yet surprisingly, this mutant has been shown to bind TA proteins in vitro (Yamagata et al., 2010). Thus, we decided to characterize this mutant in vitro and in vivo as well.

Indeed, consistent with the expected effect of the mutations, the ATPase activity of Get3 K31A and Get3 D57E were both severely diminished compared to Get3 wt (Fig 3A). Furthermore, their chaperone activity was comparable to that of Get3 wt upon oxidation (Fig S5). As shown previously, the oxidation-activated chaperone activity of Get3 can be switched off by reducing the protein and, importantly, supplying it with ATP (Voth et al., 2014). However, we found that unlike Get3 wt or D57E, once activated, the chaperone activity of Get3 K31A was strongly resistant to deactivation irrespective of the presence or absence of ATP (Fig 3B). Although comparable oxidative conditions are unlikely in living cells under normal growth conditions, this raises the possibility that even small amounts of Get3 K31A that get activated in living cells fail to switch off their chaperone activity. If that is the case, we would expect the *in-vivo* phenotype of Get3 K31A to be more pronounced than that of Get3 D57E due to its dysregulated chaperone activity.

Since overexpressing Get3 in the absence of its receptor complex is detrimental to cell growth, probably due to the accumulation of substrates with Get3 (Schuldiner et al., 2008), we wondered whether Get3 D57E and K31A could have a similar effect even in the presence of Get1, Get2. Expressing the different Get3 mutants from a plasmid containing a galactose-activated promoter, we found that while overexpressing Get3 D57E caused only a mild growth defect, Get3 K31A strongly inhibited cell growth (Fig 3C). To test whether substrate binding at the hydrophobic groove of Get3 is responsible for the effect, we combined D57E and K31A with further mutations that disrupt the hydrophobic TA-binding groove of Get3 (Coy-Vergara et al., 2019; Mateja et al., 2009). Indeed, we found that the mutation F190D I193D (FIDD) combined with either D57E or K31A completely relieved the negative effects of both mutants on cell growth (Fig 3C). This result shows that substrate trapping by Get3 D57E and especially Get3 K31A is the driving force behind the observed growth defects. Interestingly, substrate trapping seemingly does not require ATP binding, although it is thought to be necessary for the occlusion of TA proteins in the Get3 complex.

We then introduced mTagBFP2-Get3 K31A into the same collection of SWAT library strains used in the D57E screen and compared the list of proteins colocalizing with Get3 D57E and Get3 K31A. We found that mTagBFP2-Get3 K31A formed stronger foci in cells than mTagBFP2-Get3 D57E (Fig 3D). Not only did it colocalize with all proteins also trapped by Get3 D57E, but it trapped some of them more efficiently (Fig 3D). Furthermore, we found two proteins that colocalized with the K31A mutant but not with Get3 D57E (Fig 3E), confirming that Get3 K31A is an even more efficient trap than Get3 D57E. Because of the ATP-binding deficiency of the mutant, this observation is only partially consistent with current models of the GET pathway and provides the first correlation between Get3's *in-vitro* chaperone activity and its effects on living cells.

2.5. Get3 interacts with integral membrane proteins containing two TMSs close to the C-terminus

We found several proteins that were either affected in *get3* cells or colocalized with Get3 D57E and feature multiple TMSs. However, theses TMSs were located close to the C-terminus (Fig 1A, Fig 2C). Due to the similar position of these TMSs to a TA, we tested whether they interacted with Get3 in a similar manner to TA proteins.

We chose Vph2 as a model substrate for this class of proteins with multiple C-terminal TMSs. Vph2 is a factor involved in the assembly of the vacuolar ATPase (Jackson and Stevens, 1997). It possesses two TMSs located close to the C-terminus (Fig 4A) and colocalized with Get3 D57E (Fig 4B) in the logarithmic phase. To determine whether Get3 D57E can indeed trap Vph2, as was the case for TA proteins in human cells (Coy-Vergara et al., 2019), we co-expressed N-terminally HA-tagged Vph2 with Get3-TEV-GFP wt and D57E. Get3-TEV-GFP had a Tobacco Etch Virus (TEV) protease cleavage site integrated between Get3 and GFP to allow the elution from the beads with TEV protease (Rigaut et al., 1999). This approach minimises the background of an immunoprecipitation performed in the absence of detergent in order to preserve hydrophobic Get3-substrate interactions. Using this method, we detected a clear enrichment of HA-Vph2 with Get3 D57E over wt (Fig 4C-D).

Get5 forms cytosolic foci even in the absence of Get3 in stationary phase and these foci accumulate the mislocalized TA protein Sed5 (Jonikas et al., 2009). Hence, we asked whether Vph2 – like Sed5 – colocalizes with Get5 when Get3 is not present. As a control, we confirmed that Vph2 also colocalized with Get3 D57E in the stationary phase (Fig 4E). Similarly, when mTagBFP2-Get5 was expressed in the absence of Get3 in the stationary phase, Vph2 showed clear colocalization (Fig 4E). This confirms that Vph2 behaves like a TA protein substrate of Get3, and opens the possibility that other proteins with similar topology we found to colocalize with Get3 D57E represent a novel class of Get3-dependent targeting substrates.

2.6. Get3 interacts with the monotopic membrane proteins Atf1 and Atf2

Among the groups of proteins affected in *get3* cells or colocalizing with Get3 D57E we also found monotopic and peripheral membrane proteins. Their interaction with Get3, however, cannot be explained within the model framework of the canonical TA protein targeting pathway. To gain a better understanding of the nature of this interaction, we set out to study Atf2, a monotopic membrane protein we found to colocalize with Get3 D57E in the logarithmic phase.

Atf2 has two amphipathic helices, both of which are required for its membrane localization (Fig 5A-B) irrespective of whether Atf2 is N- or C-terminally GFP-tagged (Lin and Wheeldon, 2014). We also noted that its paralog, Atf1, despite being similar in

topology, did not colocalize with Get3 D57E (Fig 5A, C). When co-expressed with Get3-TEV-GFP from a plasmid, N-terminally HA-tagged Atf2 co-immunoprecipitated specifically with Get3-TEV-GFP wt and D57E (Fig 5D). However, unlike Vph2 and most TA proteins investigated in the mammalian system (Coy-Vergara et al., 2019), Atf2 was not enriched with Get3 D57E compared to wt (Fig 5D). This finding suggests that it is not a targeting substrate of the GET pathway, but rather interacts with Get3 in a non-canonical manner.

Next, we sought to determine whether any of the amphipathic helices of Atf2 are required for its interaction with Get3. Interestingly, HA-Atf2 lacking either the N- or the C-terminal amphipathic helix (NAH and CAH, respectively) failed to co-immunoprecipitate with Get3-TEV-GFP (Fig 5E). This may suggest that the interaction between Atf2 and Get3 takes place after Atf2 reaches the membrane and Get3 recognizes a motif in Atf2 in a context-dependent manner.

Since TA proteins contact a hydrophobic substrate-binding groove of Get3 (Mateja et al., 2009), we wondered whether Atf2 utilized the same binding surface on Get3 as well. We found that the amount of HA-Atf2 co-immunoprecipitating with Get3-TEV-GFP FIDD, a mutant with a disrupted hydrophobic substrate-binding groove, decreased substantially compared to Get3 wt (Fig 5F-G), which confirms that Atf2 also contacts Get3 via the hydrophobic substrate-binding groove.

Co-immunoprecipitation experiments with HA-Atf1, the paralog of Atf2, and Get3-TEV-GFP constructs revealed that Atf1 shows a similar biochemical interaction profile with Get3 as Atf2 (Fig 5H, I). Although both co-purified equally well with Get3-TEV-GFP wt and D57E, the interaction was abolished by the FIDD mutation. This highlights the fact that a lack of colocalization between Get3 D57E and a potential substrate does not exclude the possibility that they interact outside of the GET pathway. Moreover, Atf1 and Atf2 represent a new class of interaction partners of Get3 which utilizes the hydrophobic interface also used by TA proteins. However, their interaction may take place after and not before them reaching the membrane since both hydrophobic segments, and thus membrane localization, are required for the interaction.

2.7. The Get3 dependence of TA proteins is influenced by their N-terminal domain

Out of all the TA proteins we tested in our screens, we found less than 25 % mislocalized in *get3* cells, but in total approximately 70 % were either affected in *get3* cells or captured by Get3 D57E (Fig 6A). Some TA proteins that we found neither affected by the loss of Get3 nor captured by Get3 D57E (marked with an asterisk in Fig 6A) were previously shown to mislocalized in *get3* cells (Li et al., 2019; Schuldiner et al., 2008). This result highlights that the Get3-dependence of TA proteins is affected by their expression levels (Li et al., 2019) and the exact growth conditions (Powis et al., 2013; Schuldiner et al., 2008), even when they show no visible membrane targeting defects at low expression levels and in the logarithmic growth phase. Therefore, our results show that under normal growth conditions and without strong overexpression, the majority of TA proteins can reach the membrane in the absence of Get3. At the same time, when a trapping mutant of Get3 is expressed in the cells, this approach reveals that most of them do use the GET pathway to a certain extent, which becomes apparent as a localization to Get3-positive foci.

The hydrophobicity of the TMS of TA proteins has emerged to be a major determinant of their dependence on Get3 (Coy-Vergara et al., 2019; Guna et al., 2018). The hydrophobicity of the TMS of mammalian SEC61B, which corresponds to a score of 21.6 on the Transmembrane (TM) tendency scale (Zhao and London, 2006), has been proposed to be the limit above which TA proteins are targeted by Get3/TRC40 (Guna et al., 2018). To test whether the same principles hold true in yeast, we plotted the hydrophobicity of the TMS (Fig 6B) calculated either with the GRAVY (Kyte and Doolittle, 1982) or the TM tendency score system (Zhao and London, 2006) for the yeast TA proteins included in our screen. Both systems are based on the hydrophobicity of the amino acid residues, but the TM tendency scale also takes into account the probability of amino acids to be found in TMSs based on empirical evidence. Interestingly, TA proteins that are affected in *get3* cells do not show a clear dependence on the hydrophobicity of the TMS as several of them have a low value on both hydrophobicity scales (Fig 6 B, blue and red dots).

However, a line can be drawn at around a TM tendency score of 20, above which, barring a few exceptions, TA proteins are either affected in *get3* cells or captured by Get3 D57E. One of these exceptions is Bet1, but it was captured by Get3 K31A. TA proteins

with lower TM-tendency TMSs were on the other hand not captured by Get3 D57E as a rule. The notable exception in this range is Fis1, a mitochondrial TA protein that has never been found to be affected by loss of Get3. The GRAVY score of the TMS also provided a good prediction for TA proteins with higher scores, especially above 50, as such TA proteins are affected in *get3* cells or captured by Get3 D57E as a rule. Taken together, a TM tendency score higher than 20, especially when combined with a GRAVY score higher than 50, provides a fairly accurate prediction for TA proteins to be either affected by the loss of Get3 or get captured by Get3 D57E in yeast cells.

We noticed that those TA proteins that both mislocalized in *get3* cells and were captured by Get3 D57E formed no clear group in the hydrophobicity plot (Fig 6B, red dots). Since Sed5, an essential protein and well-characterized substrate of Get3, was also in this group, we tested whether the N-terminal domains of these proteins play a role in their strong dependence on Get3. We chose Gos1, another SNARE protein that showed no Get3 dependence in our screens as a control and created chimeras between Sed5 and Gos1 by swapping their cytosolic domains and TMSs. When expressed from a plasmid, GFP-Sed5 strongly mislocalized in *get3* cells, but we could not observe major differences for GFP-Gos1 between wt and *get3* cells (Fig 6C). However, when we swapped their TMSs, the behaviour of the resulting chimeras was determined by the N-terminal domain and not the TMS. Namely, the construct carrying the cytosolic domain of Sed5 and the TMS of Gos1 (GFP-Sed5-Gos1TM) still strongly mislocalized, but the cytosolic domain of Gos1 with the TMS of Sed5 (GFP-Gos1-Sed5TM) was largely unaffected in *get3* cells (Fig 6C). Furthermore, when coexpressed with mTagBPF2-Get3 D57E, GFP-Sed5-Gos1TM still colocalized with Get3 D57E but GFP-Gos1-Sed5TM did not (Fig 6D). This shows that the N-terminal domain of TA proteins can have a profound influence on their Get3 dependence, and this may be the reason why several TA proteins with high TMS hydrophobicity were affected by the loss of Get3, yet did not colocalize with Get3 D57E.

To further dissect the role of the N-terminal domain of Sed5 in its dependence on Get3, we substituted its TMS with the lipidation sequence of Ykt6, a SNARE protein which is C-terminally lipidated (McNew et al., 1997). The attached lipid moiety provides a membrane anchor for Sed5-ykt6, which is thus no longer a TA protein. When coexpressed with mTagBFP2-Get3 D57E, GFP-Sed5-ykt6 did not colocalize with Get3 in the logarithmic phase (Fig 6E). Since glucose starvation has been shown to induce

colocalization between Get3 and mislocalized Sed5 (Powis et al., 2013), we glucose-starved the cells and found that GFP-Sed5-ykt6 colocalized with mTagBFP2-Get3 D57E under these conditions. When we substituted the TMS of Gos1 with the same lipidation sequence, unlike Sed5, we could not find any colocalization between GFP-Gos1-ykt6 and mTagBFP2-Get3 D57E either in the logarithmic phase or during glucose starvation (Fig 6E). Therefore, the N-terminal domain of Sed5 seems to have the unique property of being recognized by Get3 under stress conditions even if it is not part of a TA protein but anchored to the membrane by a lipid anchor. Notably, expressing the soluble N-terminus of Sed5 with the TMS of Gos1 lacking its last amino acid (GFP-Sed5-Gost1TM(-1)) abolished the construct's membrane localization, probably due to the TMS being too short for membrane integration, and did not lead to any colocalization with Get3 D57E (Fig 6F). This implies that the mechanism leading to this colocalization requires membrane targeting and some unknown feature of the N-terminal domain of Sed5.

2.8. Sed5 is a super-client of Get3 and interacts with it before and after getting targeted to the membrane

To further assess whether Get3 interacts with Sed5 after having been targeted to the membrane, we added an opsin-tag to the C-terminus of Sed5 (Sed5-op) to be able to monitor its membrane integration through the glycosylation of the tag (Fig 7A), as has been used with other TA proteins in *in vitro* (Schuldiner et al., 2008; Stefer et al., 2011) and *in viv*o assays (Coy-Vergara et al., 2019). As expected, an extra band and a faint higher-molecular weight smear appeared corresponding to glycosylated Sed5-op (Sed5-op-glyc), the latter or which probably representing complex glycosylation, when we expressed Sed5-op (Fig 7B). Only the non-glycosylated Sed5-op was visible when a glycosylation deficient Sed5-op mutant was expressed (Fig 7B). N-glycosylation can be removed from glycoproteins by the cytosolic peptide N-glycanase Png1 when proteins are extracted from membranes (Suzuki et al., 2000), for instance when they are targeted to degradation (Fig 7 A). Therefore, to test whether any proportion of Sed5 was a target of Png1 and hence extracted from the membrane, we knocked out *PNG1* in wt and *get3* strains. To make sure that endogenous Sed5 does not compete for Get3 binding with Sed5-op, we expressed Sed5-op from a plasmid and deleted the chromosomal copy of *SED5*.

Sed5-op in logarithmically growing cells showed only minimal glycosylation (Fig S6). This could be the result of the combination of Sed5 cycling between the ER and Golgi (Wooding and Pelham, 1998) and most of the Sed5 population being freshly synthesized due to the fast division of cells, both of which increase the chance of Sed5 evading glycosylation in the ER. Since cell division and protein synthesis slows down as cells reach stationary phase (Werner-Washburne and Singer, 1993), thus giving more chance for the existing Sed5 to get glycosylated in the ER, we also tested Sed5-op glycosylation in the stationary phase (Fig 7C). Indeed, Sed5-op showed clear glycosylation in the stationary phase and consistent with a defect in primary targeting to the membrane, the amount of glycosylated Sed5 was lower in *get3* than in wt cells (Fig 7D). Surprisingly, *png1* and *png1 get3* cells showed increased Sed5-op glycosylation compared to wt and *get3*, respectively (Fig 7D). This result suggests that a pool of Sed5-op was extracted from the membrane and deglycosylated by *PNG1*. Sed5 glycosylation levels in *png1 get3* cells was restored to the levels measured in wt cells. This finding suggests that there is a pool of Sed5 that is extracted and stabilized by Get3 but degraded in the absence of Get3.

To determine whether the observed effects on the amount of the glycosylated form of Sed5-op reflect a physical interaction between Get3 and Sed5-op in the stationary phase, we immunoprecipitated Get3-TEV-GFP in *png1* cells expressing Sed5-op. Under these conditions, both glycosylated and non-glycosylated Sed5-op co-eluted with Get3 (Fig 7E), which demonstrates an interaction between Get3 and Sed5-op after its membrane insertion in the stationary phase. On the other hand, when Get3-TEV-GFP D57E was expressed, most co-eluting Sed5-op was non-glycosylated (Fig 7E), which is in line with results from mammalian cells (Coy-Vergara et al., 2019) and strongly implies that it was trapped before membrane insertion. Furthermore, Sed5-op failed to co-immunoprecipitate with the TA-binding deficient FIDD mutant of Get3 (Fig 7F), confirming that this interaction uses the TA substrate binding site of Get3. Lastly, when the immunoprecipitation was performed in strains expressing Png1, only non-glycosylated Sed5-op co-eluted with Get3, showing that the interaction takes place when the glycosylated opsin tag is exposed to the cytosol, after extraction of Sed5-op from the membrane. Taken together, these results demonstrate that Sed5 is not only a TA substrate of Get3, but also contacts it in the stationary phase after its extraction from the

membrane, which may affect its degradation kinetics and could possibly lead to its reinsertion.

3. Discussion

In our screens, most TA proteins showed no obvious defect in *get3* cells (Fig 6A), albeit some of them are known to mislocalize when strongly overexpressed or tested in stationary phase cells (Schuldiner et al., 2008). This underscores the robustness of membrane targeting *in vivo* thanks to the overlapping substrate specificity of alternative targeting pathways (Ast et al., 2013; Aviram et al., 2016). Based on the combined results of TA proteins affected in *get3* cells and those captured by Get3 D57E, it became clear that high TMS hydrophobicity is indeed a good indicator for Get3 dependence (Fig 6B). Intriguingly, the opposite was not true, and TA proteins with low TMS hydrophobicity could still be affected by the loss of Get3, although they were highly unlikely to colocalize with Get3 D57E. This could indicate that unlike in mammalian systems (Guna et al., 2018) and despite the apparent lack of colocalization with Get3 D57E, yeast Get3 can target TA proteins with low TMS hydrophobicity. Some of these TA proteins, namely Vps64, Gos1, Sss1, have been found to mislocalize to mitochondria only in *get3 msp1* double mutant cells (Li et al., 2019), explaining why they were not affected in our screens. However, this also means there could be further TA and non-TA substrates of Get3 that were neither affected in *get3* cells, nor colocalized with Get3 D57E in our screens.

Some of the non-TA proteins that did colocalize with Get3 D57E resembled TA proteins in that all their TMSs were found close to the C-terminus. Such proteins may also have to be post-translationally targeted to the membrane due to the position of their TMSs. We show that one of these proteins, Vph2 is enriched by Get3 D57E and therefore propose that these proteins represent a novel class of GET-pathway clients (Fig 8).

Our screens also revealed some monotopic and peripheral membrane proteins as potential Get3 clients, including Atf2. Importantly, we found that Get3's interaction with these proteins depends on the same hydrophobic Get3 surface that the protein uses to chaperone TA proteins. However, since the prerequisite of the interaction between Atf2 and Get3 seemed to be the prior membrane localization of Atf2, we can exclude the possibility that Get3 is involved in targeting Atf2 to the membrane. Instead, Get3 could serve as a chaperone for Atf2 once it arrives at the membrane. Although we did not

observe any changes in Atf2 in *get3* cells, we cannot exclude that its function is affected. Moreover, other cytosolic chaperones could mask the effects of loss of Get3, just like alternative pathways target TA proteins in the absence of it. Such proteins could represent members of a class of substrates that contact Get3 outside of the scope of the GET pathway (Fig 8).

Our results show that several non-TA proteins' localization or steady-state levels are affected by the loss of Get3 (Fig 1A). However, the effects on many of these proteins are probably indirect due to the disturbed biogenesis of some key Get3 substrates. One such substrate, Sed5, is known to be able to suppress the Kar2-secretion phenotype of *get3* cells resulting from impaired Golgi function when overexpressed (Schuldiner et al., 2008). Conversely, we found that the transient loss of Sed5 affects several, although not all of the proteins also affected by the loss of Get3 (Fig 1F, G), providing further confirmation that the loss of Sed5 function from the Golgi is a hallmark of impaired GET pathway function. Our results indicate that one of the reasons Sed5 is so uniquely Get3 dependent is that it interacts with Get3 not only before but also after membrane insertion. These findings imply that a population of Sed5 is extracted from the membrane in stationary phase, which can be then recognized by Get3 (Fig 8). Although we do not know the details of the mechanism, our results would support a model in which Get3 protects Sed5 from degradation and allows it to be reinserted into the membrane later.

It is likely that more substrates of Get3 remain to be discovered regarding its activity as a membrane protein chaperone. Although the precise nature and relevance of Get3's interaction with non-TA proteins remains to be elucidated, the proteins we identified as new Get3 interaction partners indicate that Get3 can bind and probably chaperone other proteins using its hydrophobic groove.

4. Materials and methods

4.1. Yeast strains and plasmids

Plasmids were constructed either by cutting the plasmid backbone and the insert amplified via polymerase chain reaction with restriction enzymes creating appropriate sticky ends and ligating them with T4 ligase (Thermo Fisher) or by using the Gibson assembly method (Gibson et al., 2009). All restriction enzymes used during cloning were from Thermo Fisher. After amplification in *Escherichia coli* ElectroTenBlue (Agilent; Catalog #: 200159) bacteria, plasmids were confirmed by sequencing. All plasmids used in this study are listed in Table 2. The cloning strategy and primers used to create them are listed in Table 3 and 4.

All yeast strains used in this study are listed in Table 5. Genomic modification and transformation of strains with plasmids was done based on the lithium-acetate (LiAc) method (Gietz et al., 1992). Briefly, logarithimic phase cells were pelleted and resuspended in 120 µL LiAc buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM LiAc), then mixed with 14 µL carrier DNA (10 mg / mL) and plasmid or DNA construct tagging or knocking out genes. 750 µL PEG buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 100 mM LiAc, 40% polyethylene glycol (PEG) 3350) was added, and samples incubated for 1 h at 42 °C shaking at 900 rpm. For plasmid transformations, samples were pelleted, resuspended in 50 µl autoclaved water, and spread on appropriate selection plates. When genomic modification was carried out, strains were instead pelleted and resuspended in rich media, then incubated for 1 h at 30 °C shaking at 900 rpm. Samples were then pelleted, resuspended in 50 µl autoclaved water, then spread on appropriate selection plates.

Genetically modified yeast strains were constructed by inserting a DNA cassette into a specific locus via homologous recombination as specified in Table 5. Integration of the construct into the locus was verified by performing polymerase chain reaction (PCR) on genomic DNA with primers specific to the gene modified. Tagging of Sed5 with the 6xMyc-AID tag and expression of Tir1 was verified by Western Blotting by Anne Clancy.

To generate modified versions of the SWAT library, synthetic genetic array-based (SGA) crosses were carried out as described previously (Cohen and Schuldiner, 2011; Tong and Boone, 2006). In short, SWAT library strains were mated with a query strain carrying the specific mutations on rich medium plates (see Table 5), the resulting diploid

strains were selected for all applicable markers, then induced to sporulate on nitrogen-starvation plates. After sporulation, haploid strains were first selected for auxotrophic, then for antibiotic markers to yield the final modified haploid strain. As a counterselection against any remaining diploid cells, all plates after sporulation contained the toxic amino acid analogs Canavanine and Thyalisine.

Diploid strains expressing mTagBFP2-tagged proteins in the SWAT background were created by crossing the modified SWAT library strains with a *MATalpha* strain (Table 5) expressing the mTagBFP2-tagged protein from a plasmid. Strains were imaged after selecting for diploid cells on synthetic dropout plates lacking leucine and uracil.

Strains lacking chromosomal *SED5* were created as described in Table 5. To remove p416 mCherry-Sed5 from cells, strains were streaked out on synthetic dropout plates lacking leucine supplemented with 5-fluoroorotic acid as a counterselection marker against p416 mCherry-Sed5 and allowed to grow for two days at 30 °C. Single colonies were restreaked once onto plates of the same composition, then tested for loss of mCherry-signal via microscopy.

4.2. Reversible oxidation of Get3 – Performed by Dr. Kathrin Ulrich

Get3 was produced recombinantly as described previously (Voth et al., 2014). Recombinant Get3 was diluted to 5 µM in HEPES buffer (40 mM HEPES-KOH, pH 7.5) and treated with 5 mM DTT, 5 µM $ZnCl_2$, and 0.5 mM Mg-ATP for 6 h at 30 °C. The protein was washed with HEPES buffer using an Amicon filter with a 30 kDa cut-off until the flow through was free of thiols and run through a gel filtration spin column afterwards. For oxidation, Get3 was diluted to 50 µM in HEPES buffer and treated with 2 mM H_2O_2 and 50 µM $CuCl_2$ for 4 min at 37 °C. The oxidants were immediately removed using a gel filtration spin column. The thiol concentration of reduced and oxidized Get3 was determined by the reaction with 5,5'-dithiobis-(2-nitrobenzoic acid) ($\varepsilon 412$ = 13.6 $mM^{-1}cm^{-1}$). For the re-reduction, 5 µM Get3 was mixed with 5 mM DTT and 5 µM $ZnCl_2$ in HEPES buffer and incubated for various times in the presence or absence of 0.5 mM Mg-ATP or Mg-ADP.

4.3. ATPase activity assay – Performed by Dr. Kathrin Ulrich

The ATPase activity of 4 mM or 8 mM Get3 was investigated using an NADH-coupled enzymatic assay containing 1 mM phosphoenolpyruvate, 12 U pyruvate kinase,

0.5 mM NADH, and 22 U lactate dehydrogenase in 100 mM HEPES, 10 mM MgCl$_2$, 20% Glycerol, pH 7.0 (Kiianitsa et al., 2003). The reaction was initiated by adding 2 mM ATP and the decrease in the NADH absorption was monitored at 340 nm in a 96-well BMG FLUOstar Omega microplate reader.

4.4. Chaperone activity assay – Performed by Dr. Kathrin Ulrich

Get3 chaperone activity was measured by following the aggregation of chemically denatured citrate synthase as described (Voth et al., 2014). Briefly, 12 µM citrate synthase (Sigma-Aldrich) was denatured in 40 mM HEPES, pH 7.5 containing 6 M guanidine hydrochloride overnight at room temperature. Citrate synthase was diluted to a final concentration of 0.075 µM in HEPES buffer at 30 °C under continuous stirring and light scattering was monitored ($\lambda_{ex/em}$ = 360 nm) in the absence (control) or presence of a four-fold molar excess of Get3 using a Hitachi F4500 fluorescence spectrophotometer equipped with a temperature-controlled cuvette holder and stirrer.

4.5. Spotting assays

Strains were grown to logarithmic phase, then normalized OD$_{600}$ 0.8 and resuspended in water. 5 µl of a 1 : 10 dilution series was spotted onto appropriate plates, which were incubated at 30 °C for at least 2 days before imaging. To control the expression of the p426 constructs, strains were cultured in synthetic dropout media lacking uracil and containing 2 % raffinose instead of glucose, then spotted onto synthetic dropout plates lacking uracil and containing either 2 % raffinose or 2 % galactose.

For the terbinafine sensitivity assay, rich media plates containing 50 µg / ml terbinafine were made using a stock solution of terbinafine dissolved in methanol.

4.6. Protein MS

4.6.1. Immunoprecipitation for MS

Processing of samples is based on a protocol described previously (Ripaud et al., 2014). Strains were grown to OD$_{600}$ 0.8 in synthetic dropout media lacking leucine. 500 mL of each strain was harvested and the cell pellet stored at -80 °C. Pellets were thawed and resuspended in 3 mL lysis buffer (25 mM Tric-HCL pH 7.5, 50 mM KCl, 10 mM MgCl2, 5 % glycerol, 1 % IGEPAL CA-630, 1 mM PMSF, 5 µg/mL Leupeptin/Pepstatin, 1 tablet / 20 mL Roche cOmplete EDTA-free protease inhibitor cocktail), then crushed by hand in

liquid nitrogen. Thawed lysates were centrifuged at 500 g for 5 min, the supernatant was transferred to a new tube and centrifuged at 500 g for another 5 min. 500 µl of the supernatant was treated with 50 U of Bezonase (Sigma-Aldrich) for 1 h at 4 °C on a rotating wheel. 50 µl µMACS Anti-GFP MicroBeads (Miltenyi Biotec) were added, and incubated on a rotating wheel for 1h at 4 °C. Samples were loaded onto µ Columns (Miltenyi Biotec) in a µMACS Separator (Miltenyi Biotec), washed four times with 200 µl Wash Buffer I (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 % glycerol, 0.05 % glycerol), then four times with 200 µl Wash Buffer II (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 % glycerol). Samples were digested on the column first with 0.15 µg LysC in 25 µl digestion buffer I (50 mM Tris-HCl pH 7.4, 8 M Urea, 1 mM DTT) for 1 h at room temperature. Columns were washed with 25 µl digestion buffer II (50 mM Tris-HCl pH 7.4, 5 mM Chloroacetamide), and residual sample on the column was further digested with 0.15 µg trypsin in 25 µl digestion buffer II for 1 h at room temperature. Columns were washed with 50 µl digestion buffer II, yielding 100 µl eluate in total, which was mixed with 50 µl 300 mM thiourea. After the addition of 0.5 µg trypsin, eluates were further digested overnight at room temperature, then analysed by MS.

4.6.2. StageTip peptide clean-up – Performed by Jakob Bader

The digests were conditioned for peptide stop-and-go extraction tip (StageTip) clean-up by adding 15µl of 5.5% (v/v) trifluoroacetic acid, 22% (v/v) acetonitrile. The samples were incubated on ice for 10 min, centrifuged at 20 000 g at 4°C for 10 min and the supernatant was transferred to another tube to remove any possible precipitate or particle. StageTip clean-up was performed with three plugs of C18 material (Kulak et al., 2014). In brief, StageTips were washed with 100 µL buffer B (50% (v/v) acetonitrile, 0.5% (v/v) acetic acid), 100 µL methanol and two times with 100 µL buffer A (2% (v/v) acetonitrile, 0.5% (v/v) acetic acid). Half the volume of the sample were loaded onto the StageTip. The StageTips were again washed two times with 200 µL buffer A and centrifuged to dryness. The peptides were eluted with 60 µL buffer B and the eluate evaporated in a vacuum concentrator to dryness. The peptides were resuspended in 10 µL A* (2% (v/v) acetonitrile, 0.5% (v/v) acetic acid, 0.1% (v/v) trifluoroacetic acid) and stored at -20 °C.

4.6.3. LC-MS/MS analysis – Performed by Jakob Bader

For LC-MS/MS analysis, 1 µL of peptide sample was loaded onto an in-house packed liquid chromatography column (75-µm inner diameter, 50-cm length, and 1.9-µm C18

particles [Dr. MaischGmbH]). Peptides were separated on an EASY-nLC 1200 HPLC system (Thermo Fisher Scientific) in a gradient of buffer A (0.5% (v/v) formic acid) to buffer B (0.5% (v/v) formic acid, 80% (v/v) acetonitrile) at a flow rate of 300 nL/min and a temperature of 60°C. The gradient started at 5% B and the concentration of B was increased linearly to 30% over 65 min, increased to 60% B over 5 min, increased to 95% B over 5 min, kept at 95% B for 5 min, reduced to 5% B over 5 min and kept at 5% B for 5 min. A Quadrupole Orbitrap mass spectrometer (Q Exactive HF-X; Thermo Fisher Scientific) was coupled to the liquid chromatography via a nano-electrospray source. The mass spectrometer was operated in a data-dependent mode. The survey scan range was set from 300 to 1,650 m/z, with a resolution of 60,000 at m/z 200. Up to the 12 most abundant isotope patterns with a charge of two to five were isolated with a 1.4 Th window and subjected to collision-induced dissociation fragmentation at a normalized collision energy of 27 and subsequent MS/MS analysis at a resolution of 15,000 at m/z 200. The automatic gain control target was set to $3*10^6$ charges and 10^5 charges for MS and MS/MS events, respectively. Injection times were limited to 20 ms and 60 ms for MS and MS/MS events, respectively. Dynamic exclusion to minimize resequencing was set to 30 s. Data were acquired using Xcalibur software (Thermo Fisher Scientific).

4.6.4. MS data analysis – Performed by Jakob Bader

To process MS raw data, we employed the MaxQuant software (version 1.6.0.15) using the Andromeda engine to search MS features against the uniprot reference proteome of Saccharomyces cerevisiae strain S288C of canonical and isoform sequences downloaded in March 2018 (Tyanova et al., 2016). The uniprot proteome was amended manually with the Get3 D57E and the Aequoria Victoria GFP. Enzyme specificity was set to trypsin, allowing cleavage N-terminally to proline and up to two mis-cleavages. Carbamidomethylation was set as fixed modification, and acetylation on the protein N-termius and methionine oxidation were permitted variable modifications. A false discovery rate (FDR) cutoff of 1% was applied at the peptide and protein level. The MaxQuant feature "Match between runs" which allows the transfer of peptide identifications in the absence of sequencing after nonlinear retention time alignment was enabled. Protein abundance was normalized with MaxLFQ label-free normalization yielding LFQ-intensities.

4.6.5. Bioinformatic analysis - Performed by Jakob Bader

For bioinformatic analysis and visualization, we used the Perseus software environment version 1.5.2.11 (Tyanova, Temu, Sinitcyn, et al., 2016). Common contaminants, reverse sequences and proteins only identified by site localization were removed from further analysis. Protein LFQ-intensities were log2-transformed. The dataset was filtered for proteins that have been observed in at least 66% of replicates of at least one experimental condition, i.e. the bait used in the immunoprecipitation. Missing values were imputed from a normal distribution using the default Perseus settings (sample-wise imputation, down shift = 1.8, width = 0.3). One of the four replicates each of both the Get3 and the Get3 D57E samples presented as outlier in a principal component analysis and was excluded from further analysis. Due to the protein groups concept, Get3-GFP and Get3-GFP D57E appear as separate proteins in the dataset, and for identification in the dataset Get3-GFP D57E was assigned the mock uniprot ID Q121542. In the dataset and the plots, "Get3" (Q12154) refers to the abundance of the relevant Get3 (wt or D57E), thereby permitting a direct comparison of their abundance. In turn, "Get3 D57E" (Q121542) refers to the mutation-specific variant. Due to the missing value imputation of Get3 D57E in replicates of Get3 wt cells that do not contain Get3 D57E, fold changes involving Get3 D57E are finite instead of infinite. Statistical significance was computed using a two-tailed Student's T-test with unequal variance between samples in Excel by Ákos Farkas.

4.7. Immunoprecipitation with TEV cleavage

Strains were grown to OD_{600} 0.8 in synthetic dropout media lacking appropriate auxotrophic selection markers (leucine or uracil). 400 mL of each strain was harvested and the cell pellet stored at -80 °C. Pellets were thawed and resuspended in 3 mL lysis buffer (25 mM Tric-HCL pH 7.5, 50 mM KCl, 10 mM MgCl2, 5 % glycerol, 1 mM PMSF, 5 µg/mL Leupeptin/Pepstatin, 1 tablet / 20 mL Roche cOmplete EDTA-free protease inhibitor cocktail), then crushed by hand in liquid nitrogen. Thawed lysates were centrifuged at 500 g for 5 min, the supernatant was transferred to a new tube and centrifuged at 500 g for another 5 min. 50 µL µMACS Anti-GFP MicroBeads (Miltenyi Biotec) was added to 750 µL supernatant, and incubated on a rotating wheel for 1h at 4 °C. Samples were loaded onto µ Columns (Miltenyi Biotec) in a µMACS Separator (Miltenyi Biotec), washed twice with 200 µL Wash Buffer II (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 % glycerol), then twice with 200 µL TEV cleavage buffer (50 mM Tris-HCl pH

7.5, 150 mM NaCl). 25 µL TEV cleavage buffer containing TEV protease was added onto the columns and incubated for 1 h at room temperature. Cleaved protein was eluted with 50 µL TEV cleavage buffer, mixed with SDS sample buffer (final concentration: 50 mM Tris pH 6.8, 1% SDS, 0.1% w/v Bromophenol blue, 5% Glycerol, 100 mM DTT) and analysed by SDS-PAGE followed by Western blotting.

4.8. Lipid analysis

4.8.1. Preparation of samples for lipid analysis

Cells were grown to logarithmic phase in rich or synthetic complete media. 15 OD_{600} units of cells were pelleted at room temp, washed twice in 20 mL 155 mM Ammonium bicarbonate buffer in 50 mL tubes. Cells were resuspended in 1 mL Ammonium bicarbonate, transferred to a 1.5 mL tube, then pelleted at room temperature. The supernatant was removed and the pellet was frozen in liquid nitrogen, then stored overnight at -80°. The following morning pellets were thawed on ice and resuspended in 1 ml cold Ammonium bicarbonate buffer. 200 µl zirconia beads (0.5 mm diameter) were added, and the samples were shaken using a vortex at 4° three times for 5 minutes, with 5 minutes on ice between the rounds. The lysate including beads was frozen in liquid nitrogen and stored at -80°. Samples were pelleted (2 000 rpm in tabletop centrifuge), the supernatant taken, frozen and sent for lipidomic analysis.

4.8.2. Lipid extraction for MS lipidomics – Performed by Lipotype GmbH

MS-based lipid analysis was performed by Lipotype GmbH (Dresden, Germany) as described (Ejsing et al., 2009; Klose et al., 2012). Lipids were extracted using a two-step chloroform/methanol procedure (Ejsing et al., 2009). Samples were spiked with internal lipid standard mixture containing: cytidine diacylglycerol 17:0/18:1, ceramide 18:1;2/17:0, diacylglycerol 17:0/17:0, lysophosphatidate 17:0, lyso-phosphatidylcholine 12:0, lysophosphatidylethanolamine 17:1, lyso-phosphatidylinositol 17:1, lysophosphatidylserine 17:1, phosphatidate 17:0/14:1, phosphatidylcholine 17:0/14:1, phosphatidylethanolamine 17:0/14:1, phosphatidylglycerol 17:0/14:1, phosphatidylinositol 17:0/14:1, phosphatidylserine 17:0/14:1, ergosterol ester 13:0, triacylglycerol 17:0/17:0/17:0, stigmastatrienol, inositolphosphorylceramide 44:0;2, mannosylinositolphosphorylceramide 44:0;2 and mannosyl-di-(inositolphosphoryl)ceramide 44:0;2. After extraction, the organic phase was transferred to an infusion plate and dried in a speed vacuum concentrator. 1st step dry extract was

re-suspended in 7.5 mM ammonium acetate in chloroform/methanol/propanol (1:2:4, V:V:V) and 2nd step dry extract in 33% ethanol solution of methylamine in chloroform/methanol (0.003:5:1; V:V:V). All liquid handling steps were performed using Hamilton Robotics STARlet robotic platform with the Anti Droplet Control feature for organic solvents pipetting.

4.8.3. MS data acquisition – Performed by Lipotype GmbH

Samples were analyzed by direct infusion on a QExactive mass spectrometer (Thermo Scientific) equipped with a TriVersa NanoMate ion source (Advion Biosciences). Samples were analyzed in both positive and negative ion modes with a resolution of $Rm/z=200=280000$ for MS and $Rm/z=200=17500$ for MSMS experiments, in a single acquisition. MSMS was triggered by an inclusion list encompassing corresponding MS mass ranges scanned in 1 Da increments (Surma et al., 2015). Both MS and MSMS data were combined to monitor ergosterol ester, diacylglycerol and triacylglycerol ions as ammonium adducts; phosphatidylcholine as an acetate adduct; and cardiolipin, phosphatidate, phosphatidylethanolamine, phosphatidylglycerol, phosphatidylinositol and phosphatidylserine as deprotonated anions. MS only was used to monitor lysophosphatidate, lysophosphatidylethanolamine, lyso-phosphatidylinositol, lysophosphatidylserine, inositolphosphorylceramide, mannosylinositol-phosphorylceramide, mannosyl-di- (inositolphosphoryl)ceramide as deprotonated anions; ceramide and lyso-phosphatidylcholine as acetate adducts.

4.8.4. Data analysis and post-processing – Performed by Lipotype GmbH

Data were analyzed with in-house developed lipid identification software based on LipidXplorer (Herzog et al., 2011, 2012). Data post-processing and normalization were performed using an in-house developed data management system. Only lipid identifications with a signal-to-noise ratio >5, and a signal intensity 5-fold higher than in corresponding blank samples were considered for further data analysis.

4.9. Microscopy of yeast strains and libraries

Images in Fig 1F-G were acquired as described previously (Rivera-Monroy et al., 2016). Namely, stationary-phase cells were diluted into 384-well glass-bottom microtiter plates (Matriplate, Brooks Life Science Systems) containing low fluorescence minimal media (Reference number: CYN6502; FORMEDIUM Ltd.) supplemented with

methionine, histidine, uracil and leucine, and allowed to grow for four hours at 30 °C without or in the presence of 250 mM IAA. The plates were automatically imaged at 30 °C on an Imaging Machine 03-dual (Acquifer) widefield high-content screening microscope, equipped with a white LED array for bright field imaging, an LED fluorescence excitation light source, an sCMOS camera (2048 x 2048 pixels) and a stationary plate holder in combination with movable optics. Images were acquired with 470 nm filter cubes (Ex 469/35, Em 525/39, dichroic 497) or without filter cube for brightfield images using 3 z-slices (dz = 1 μ m) and a 40x CFI Super Plan Fluor ELWD N.A. 0.60 (Nikon) objective. The focal plane was automatically detected in the brightfield channel using a yeast autofocus algorithm. Experiment was performed by Anne Clancy.

Images for the quantification shown in Fig 1C were acquired identically by Ákos Farkas but in the absence of IAA and taking three z-slices in the brightfield, and one in the fluorescent channel.

Images of yeast libraries were acquired by Ákos Farkas as follows. Stationary-phase cells were diluted into 384-well glass-bottom microtiter plates (Matriplate, Brooks Life Science Systems) containing low fluorescence minimal media (Reference number: CYN6502; FORMEDIUM Ltd.) supplemented with 2 % glucose and amino acids as required by the specific strains, and allowed to grow for four hours at 30 °C. The plates were imaged at room temperature automatically with a Nikon Ti2 2-E inverted microscope equipped with a computer-controlled stage, a Lumencor Spectra X light source, a pco.edge 5.5 M-AIR-CL-PCO sCMOS camera (2560 x 2156 pixels). Focal plane was detected with the Perfect Focus System (Nikon). Images were acquired with a 100x CFI Plan Apo Lambda 100x/1.45 oil objective using settings appropriate for GFP (Ex 470/24, Em 520/35, Dichroic 488) or mTagBFP2 (Ex 395/25, Em433/24, Dichroic 405).

All other images of yeast cells were acquired by Ákos Farkas as described above with the Nikon microscope, except that the strains were allowed to grow to logarithmic phase in appropriate selection media in glass tubes while shaking at 30 °C.

All microscopy images were processed with ImageJ (Schneider et al., 2012).

4.10. NaOH lysis

Sample preparation was adapted from the previously described NaOH lysis protocol (Kushnirov, 2000). Briefly, 750 μl cells in the logarithmic phase were pelleted,

then resuspended in 1 mL 250 mM NaOH. Samples were incubated on ice for 10 min, centrifuged for 1 min at 5 000 g, and resuspended in SDS-PAGE sample buffer. After incubating the samples for 5 at room temperature, they were centrifuged for 1 min at 5 000 g, and the supernatant loaded onto an SDS-PAGE gel. Yeast strains processed via NaOH lysis were grown in synthetic dropout media.

4.11. Silver staining

Silver staining was performed with the SilverQuest staining kit (Invitrogen) according to the manufacturer's instructions using the basic staining protocol.

4.12. Quantification of images

A script was written in KNIME (https://www.knime.com) to automatically identify cells and quantify the distribution of the fluorescence within them similarly to a method described previously (Jonikas et al., 2009). In short, using an ImageJ (Schneider et al., 2012)macro in KNIME, the difference of two brightfield images (+/- 1 µM from focal plane) was used to identify segments (cells). Background areas containing no irregularities were identified based on areas of image with low variance. Previously identified cells were removed if they overlapped with the calculated background area. Cells were also filtered for circularity, variance in the brightfield signal within the cell, the range of brightfield signal values within cells in in order to remove cells from analysis that contained irregularities or were stacked on each other. Outliers in size, i.e. too big or small cells were also excluded. Fluorescent channel background was estimated based on the fluorescent signal observed in the background areas identified with the brightfield images, then subtracted from the fluorescent channel image. Cells identified in the brightfield channels, as described above, were used to calculate the pixel distribution for each cell in the fluorescent channel. Outliers with too little variance in the fluorescent signal or too high signal intensity were removed in order to exclude dead cells. Using the different statistical parameters obtained for the distribution of pixel intensities within each cell, the average value for each statistic was calculated for the whole image. The KNIME workflow containing the script is available upon request.

Nine images were acquired automatically as described above for each strain, and low-quality images were removed manually. The skewness of the distribution of pixel intensities in the cells were measured using the KNIME script described above and averaged for each image (>500 cells per frame), and the average of the images was

calculated for each strain. Error bars in Fig 1C represent average +/- standard error of the mean.

4.13. Western blotting

Samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane (Amershan Protran). Membranes were first blocked in blocking buffer (5 % milk in Tris-buffered saline, 0.1 % Tween-20), then incubated in blocking buffer containing primary antibodies overnight at 4 °C or for three hours at room temperature. After incubation with fluorescent secondary antibodies in blocking buffer, immunoblots were washed three times for 10 minutes in blocking buffer without milk, then scanned with a LI-COR Odyssey scanner. Primary and secondary antibodies are listed in Table 7 and 8. Immunoblots were quantified in ImageStudioLite 5.2 (LI-COR). For the quantification of Sed5-op glycosylation, signals were summed between the band marked with "Sed5-op-glyc" and the 100 kDA marker, excluding the non-specific band also visible when Sed5-opNQ was tested.

4.14. Bioinformatic identification of TA proteins

To compile the list of potential TA proteins, we took the results of the different TMS prediction methods compiled in the TopologYeast database (Weill et al., 2019), which covers the whole yeast proteome. Besides the already known TA proteins (Ast et al., 2013), we selected those that could be successfully N-terminally GFP-tagged in the SWAT library and had no signal sequence or mitochondrial target peptide predicted but for which at least five of the nine methods predicted a single TMS. Of these, we included those that had no more than 90 amino acid residues after the predicted TMS since that is the longest C-terminal extension among the currently known TA proteins (Sec20). We further filtered out those that are known to localize to the inner mitochondrial membrane. We found two further TA proteins that had been predicted previously (Beilharz, 2003) but did not show up in our analysis (Tom5, Tom7), and we added two proteins which are predicted to be TA proteins based on their human homologs (Erg9, Hfd1). We excluded those TA proteins which have a known lumenal N-terminus (Ost4, Pbn1) or that did not survive the SGA procedure (Vam3).

To make the hydrophobicity of their TMS comparable, we selected the 20 amino acid window with the highest hydrophobicity centered around the predicted TMS and

calculated the TM tendency (Zhao and London, 2006) and GRAVY score (Kyte and Doolittle, 1982) for this sequence.

List of figures

Figure 1. Direct and indirect effects of losing Get3 do not exclusively converge on TA proteins.

A. Hits from a screen monitoring the localization and signal intensity of ca. 1,500 N-terminally GFP-tagged proteins in the absence of Get3 organized by phenotype and topology. Proteins localized to the Golgi apparatus or cycling between the ER and the Golgi are indicated in red, proteins involved in different aspects of lipid metabolism in blue. Arrows mark proteins for which images are shown in (B). TMS: Transmembrane segment; C'-TMS: Proteins with multiple TMSs close to the C-terminus.

B. Representative images of specific GFP-tagged proteins illustrating the phenotypic categories used in (A).

C. Changes in the relative intensity of the punctate pattern observed for GFP-Sed5, GFP-Coy1, and GFP-Mtc4 reveal milder effects of losing Sgt2 than losing Get4 or Get5. GFP-Sed5 is a positive control for which this graded dependency has been observed before. 8 or 9 images of each strain were quantified, each containing several hundreds of cells. WT represents a strain with the *MET15* locus replaced with the same resistance cassette that was used to generate the other knock-out strains.

D. Dependence of the mis-localization of GFP-Opi3 and GFP-Msp2 on components of the cytosolic GET complex reveals milder effects of losing Sgt2 than losing Get4 or Get5.

E. Scheme illustrating the AID system for acute downregulation of a tagged protein after addition of the plant hormone IAA.

F. Effect of the acute down-regulation of AID-Sed5 on the localization of Golgi proteins suggests that their Get3-dependence (A,B) may be indirect due to the Get3-dependence of GFP-Sed5 (D). Experiment performed by Anne Clancy. Scale bar: 2 μM.

G. Effect of the acute down-regulation of AID-Sed5 on the localization of GFP-Ubx3 (A,B) and GFP-Coy1 (A;B;D) suggests that they may directly depend on Get3 in their steady-state distribution. Experiment performed by Anne Clancy. Scale bar: 2 μM.

Figure 2. A trapping approach reveals TA proteins and membrane proteins of other topologies trapped by Get3.

A. Analysis of the Get3 interactome by GFP-affinity purification in the presence of detergent and MS reveals that the ATPase impaired trap mutant Get3 D57E enriches the cytosolic GET components Sgt2, Get4, Get5, and Ybr137w. Proteins eluted from a GFP affinity matrix were identified and plotted according to their relative enrichment with either Get3-GFP fusion protein. The data point labelled Get3 D57E reflects the peptide including the altered residue whereas the data point labelled Get3 on the diagonal indicated that very similar amounts of either protein were detected in the eluates. MS performed by Jakob Bader.

B. Illustration of the principle by which hits were identified in this screen, i.e. colocalization of N-terminal GFP fusion proteins from the library with a fusion of Get3 to the blue protein mTagBFP2. Arrows indicate colocalization. Scale bar: 2 µM.

C. Results of a microscopy-based screen in which *get1 get2 get3* mutants of the same set of SWAT library strains used for the screen described in Fig 1A was crossed with a *get1 get2 get3* strain expressing mTagBFP2-Get3 D57E from a plasmid. Proteins colocalizing with mTagBFP2-Get3 D57E are listed according to their topological category.

D. Impaired growth of a *get3* strain on terbinafine corroborates a role for Get3 in the function of enzymes involved in sterol synthesis such as Erg1, Erg9, and Erg27. All strains are from the SGA query strain background.

E. Lipidomic analysis of *get3* cells grown in full or minimal medium reveals Get3's relevance to sterol and triacylglycerol metabolism. In addition, the levels of inositol-phosphorylceramide and phosphatidyl-choline were affected while the majority of glycerol- and sphingo-phospholipids was unaffected. Mass spectrometric lipid analysis performed by Lipotech

Figure 3. Impaired deactivation of the Get3 chaperone results in the trapping of additional proteins and is toxic to the cell.

A. Changing lysine 31 to alanine results in a reduction in ATPase activity similar to that of Get3 D57E. Experiment performed by Dr. Kathrin Ulrich.

B. Reversal of the oxidation-activated chaperone activity of Get3 K31A is severely impaired and is unaffected by ATP unlike in the case of Get3 wt and D57E. Experiment performed by Dr. Kathrin Ulrich.

C. Overexpression of Get3 K31A is toxic to the cells, whereas Get3 D57E only shows mild growth impairment. This effect depends on a functional TA-binding groove.

D-E. Examples of proteins trapped by Get3 K31A but only to a lesser extent (D) or not (E) by Get3 D57E. The same trap screen was performed with the K31A variant as the one describe in Fig 2C. While all hits identified with the D57E variant were also found to colocalize with Get3 K31A, this variant identified additional hits as shown here. Scale bar: 2 µM

Figure 4. Get3 D57E can trap proteins containing multiple TMSs close to the C-terminus.

A. Topological scheme of Vph2, an assembly factor involved in V-ATPase biogenesis.

B. GFP-Vph2 colocalizes with mTagBFP2-Get3 D57E in the logarithmic phase.

C. GFP-affinity purification in the absence of detergent reveals a physical interaction between HA-tagged Vph2 and Get3 D57E. The immunoblot was incubated with anti-HA and anti-Get3 antibodies. The asterisk marks a non-specific band.

D. Quantification of the experiment shown in (C).

E. GFP-Vph2 accumulates in foci marked by mTagBFP2-Get3 D57E in the stationary phase. In the absence of Get3, GFP-Vph2 still accumulates in foci marked by mTagBFP2-Get5 in the stationary phase.

Figure 5. Get3 interacts with proteins anchored to the membrane via amphipathic alpha-helices.

A. Topological scheme of the alcohol O-acetyltransferases Atf1 and Atf2 illustrating the N- (NAH) and C-terminally (CAH) located amphipathic alpha-helices.

B. Deletion of either NAH or CAH abolishes membrane localization of GFP-Atf2. Scale bar: 2 µM.

C. GFP-Atf2 colocalizes with mTagBFP2-Get3 D57E but GFP-Atf1 does not. Scale bar: 2 µM.

D. GFP-affinity purification in the absence of detergent reveals a physical interaction between HA-tagged Atf2 with both Get3 wt and Get3 D57E. Immunoblot was incubated with anti-HA and anti-Get3 antibodies. The asterisk marks a non-specific band.

E. Deletion of either NAH or CAH abolishes the interaction of HA-Atf2 with Get3. Immunoblot was incubated with anti-HA and anti-Get3 antibodies. The asterisk marks a non-specific band.

F. Interaction between Atf2 and Get3 depends on the TA-binding groove of Get3. Immunoblot was incubated with anti-HA and anti-Get3 antibodies. The asterisk marks a non-specific band.

G. Quantification of the experiment shown in (F).

H. GFP-affinity purification in the absence of detergent reveals a physical interaction between HA-tagged Atf1 with Get3 wt that depends on the TA binding groove of Get3 and is not increased with Get3 D57E. Immunoblot was incubated with anti-HA and anti-Get3 antibodies. The asterisk marks a non-specific band.

I. Quantification of the experiment shown in (H).

Figure 6. Sed5 is a super-client of Get3, which is dictated by its N-terminus.

A. TA proteins tested in the current study and their dependence on Get3. Sed5 is one of the few TA proteins that are affected by the loss of Get3 and trapped by Get3 D57E. In contrast, Gos1 is an example of a TA protein that is neither affected nor trapped by Get3 when expressed chromosomally.

B. Yeast TA proteins plotted based on the TM tendency and GRAVY score of their TMS. Colored dots represent TA proteins depicted in (A).

C. Get3-dependence of GFP-Sed5 localization maps to the N-terminus of Sed5. Localization of GFP-Sed5 and GFP-Gos1 in a wt and get3 deletion strain was compared to two chimeras swapping the N-terminal regions of the proteins and their TMS. Scale bar: 2 µM.

D. Colocalization of Sed5 with Get3 D57E and the lack thereof of Gos1 is governed by the N-terminus, as revealed by chimeras described in (C) coexpresed with mTagBFP2-Get3 D57E. Scale bar: 2 μM. Arrows indicate colocalization. Perpendicular bars indicate lack of colocalization.

E. Chimeras containing the N-terminus of Sed5 or Gos1 and the lipidation sequence of Ykt6 do not colocalize with Get3 D57E in the logarithmic phase. However, GFP-Sed5-ykt6 colocalizes with Get3 D57E after inducing glucose starvation, whereas GFP-Gos1-ykt6 does not. Scale bar: 2 μM. Arrows indicate colocalization. Perpendicular bars indicate lack of colocalization.

F. A chimera of the N-terminal domain of Sed5 and the TMS of Gos1 lacking the last amino acid, a tryptophan residue, localizes to the cytosol and does not colocalize with Get3 D57E after glucose starvation. Perpendicular bars indicate lack of colocalization. Scale bar: 2 μM.

Figure 7. Get3 interacts with Sed5 after its membrane insertion.

A. Scheme showing the glycosylation of an opsin-tagged Sed5 construct (Sed5-op). If extracted from the membrane, Sed5-op can be deglycosylated by the cytosolic peptide N-glycanase Png1 and then targeted for degradation. Thus, loss of Png1 can reveal a membrane-extracted population of Sed5.

B. Comparison of Sed5-op and Sed5-op NQ reveals that a second Sed5-positive band reflects the glycosylated form of Sed5. A faint higher molecular weight smear was also detected probably representing complex glycosylation of Sed5-op. Immunoblot was incubated with anti-Sed5 and anti-Pgk1 antibody.

C. Extent of Sed5-op glycosylation in the indicated strains in the stationary growth phase. Immunoblot was incubated with anti-Sed5 and anti-Pgk1 antibody.

D. Quantification of the experiments shown in (D).

E. Get3 physically interacts with the glycosylated pool of Sed5. Comparison between Get3 wt and D57E corroborates trapping of a biogenetic precursor by Get3 D57E as well as interaction between Get3 wt and Sed5-op after membrane targeting in the stationary growth phase. Immunoblot was incubated with anti-Sed5 and anti-Get3 antibody. The asterisk indicates a non-specific band.

F. Interaction between Get3 wt and the glycosylated pool of Sed5-op depends on the TA binding groove of Get3. In the presence of Png1 only non-glycosylated Sed5-op immunoprecipitates with Get3, indicating that the interaction between glycosylated Sed5-op and Get3 takes place after Sed5-op has been extracted from the membrane. Immunoblot was incubated with anti-Sed5 and anti-Get3 antibody. The asterisk indicates a non-specific band.

Figure 8. Get3 chaperones a diverse class of membrane proteins.

Get3 can deliver TA proteins to the ER membrane (1) and facilitate the biogenesis of GPI-anchored proteins (2). Get3 can also target proteins with multiple TMSs close to their C-terminus to the ER (3). Get3 can also contact membrane proteins that lack TMSs and interact with the membrane via amphipathic helices (4). Sed5 is a super-client of Get3 that can interact with it in the stationary phase after reaching the membrane in order to be stabilized by Get3, which could lead to its eventual degradation or Get3-dependent reinsertion into the membrane (5).

Figure 1

A

B

C

D

E

F

G

Figure 2

Figure 3

A

B

C

D

E

Figure 4

A

Vph2 1 ———————— C ———————— [TM1] L [TM2] — C — 215
135-155 168-186

B

Logarithmic phase
get1/get1 get2/get2 get3/get3

GFP-Vph2   mTagBFP2-Get3
D57E

C

HA-Vph2
Get3-TEV-GFP

wt   D57E   kDa

HA →   — 25

Get3 *   — 55
    →   — 40

IP: GFP

n = 3

D

HA-Vph2 / Get3 normalized to wt

wt   D57E

E

Stationary phase

GFP-Vph2   mTagBFP2

get1/get1 get2/get2 get3/get3   mTagBFP2-
Get3 D57E

GET1/get1 GET2/get2 get3/get3   mTagBFP2-
Get5

Figure 5

Figure 6

A

TA proteins screened

Affected in *get3*	Colocalizing with Get3 D57E			
		Not affected in *get3*		
Coy1	Bos1	Csm4	Snc1	Bet1	Sss1*
Far10		Cyb5	Snc2	Frt2	Tlg2*
Frt1	Phm6	Dpm1	Sso1	Gem1	Tom5
Hfd1		Erg9	Sso2	**Gos1***	Tom6
Lam5	Rmd1	Fis1	Syn8	Kar1	Tom7
Mga2		Hlj1	Tlg1	Lam6	Tom22
Mps2	**Sed5**	Pep12	Ubc6	Meo1	Tsc3
Mtc4		Pmp2	Use1	Nyv1	Ufe1
Pex15	Ylr297w	Sbh1	Vti1	Pgc1	Vps64*
Ygl230c		Scs2	Ygl041c-b	Prm3	Ycl021w-a
		Sec22	Ygl194c-a	Rmd8	Ydl012c
		Sft1	Ysy6	Sec20	Ydr034w-b
				Spg1	Ykl065w-a
				Spt23	Yor034c-a

B

C

D

E

F

Figure 7

Figure 8

List of supplementary figures and tables

Supplementary Figure 1. Loss of Get3 affects a wide range of proteins. Representative images of all proteins affected in *get3* cells except for proteins listed in Table 1. All proteins are N-terminally GFP-tagged. Strains marked with an asterisk contain a signal sequence – sfGFP tagging cassette unlike the others to ensure proper targeting. WT for these strains represents a SWAT strain crossed with an SGA query strain containing the same antibiotic resistance cassette at the *MET15* locus as the *get3* strains.

Supplementary Figure 2. AID-Sed5 is efficiently degraded in the presence of IAA. AID-Sed5 is efficiently degraded in the presence of IAA when Tir1 is expressed. The SWAT Get3 strain has the same genetic background as the ones used in Fig 1F, G. AID-Sed5 carried an N-terminal Myc-tag, which was used for detection. Immunoblot was incubated with anti-Myc and anti-Bmh1/Bmh2 antibody. Experiment performed by Anne Clancy.

Supplementary Figure 3. Get3 D57E enriches the pretargeting complex in yeast compared to Get3 wt

A. Silver staining of a gel showing the eluates after immunoprecipitating the indicated GFP-tagged Get3 variants.

B. MS identifies the pretargeting complex strongly and statistically significantly enriched in Get3 D57E eluates compared to Get3 wt. MS performed by Jakob Bader.

Supplementary Figure 4. Lipidomic analysis of wt and *get3* cells in rich and minimal medium. Groups of lipids that were not significantly altered in *get3* cells compared to wt in both rich and minimal medium.

Supplementary Figure 5. Chaperone activity of oxidized Get3. Chaperone acitivity of Get3 wt, D57E, and K31A at time point zero as shown in Fig. 3B, normalized to Get3 wt. Experiment was performed by Dr. Kathrin Ulrich

Supplementary Figure 6. Sed5-op is weakly glycosylated in the logarithmic phase. Sed5-op is weakly glycosylated in the logarithmic phase irrespective of the genomic background. Immunoblot was incubated with anti-Sed5 and anti-Pgk1 antibodies.

Table 1. Table containing proteins affected in *get3* cells but excluded from further analysis. The list includes mitochondrial proteins mislocalized due to the N-terminal GFP tag, proteins perturbed by the N-terminal GFP tag and peroxisomal proteins.

Table 2. Plasmids used in this study.

Table 3. Cloning strategy used for each plasmid.

Table 4. Primers used to create plasmids listed in Table 2.

Table 5. Strains used in this study.

Table 6. Primers used for amplifying constructs to genetically modify yeast strains.

Table 7. Primary antibodies used in this study.

Table 8. Secondary antibodies used in this study.

Supplementary figure 1
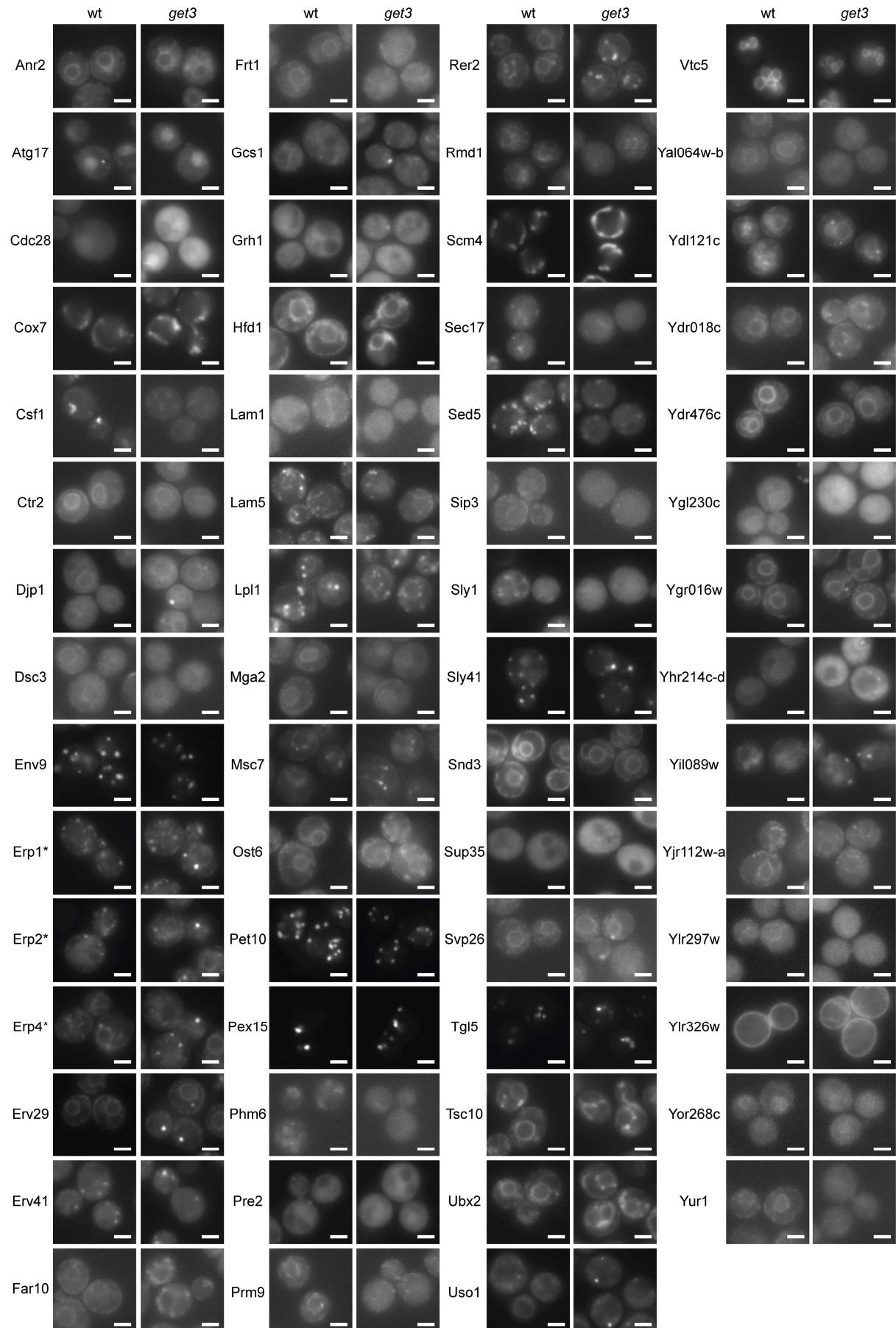
Supplementary Figure 2
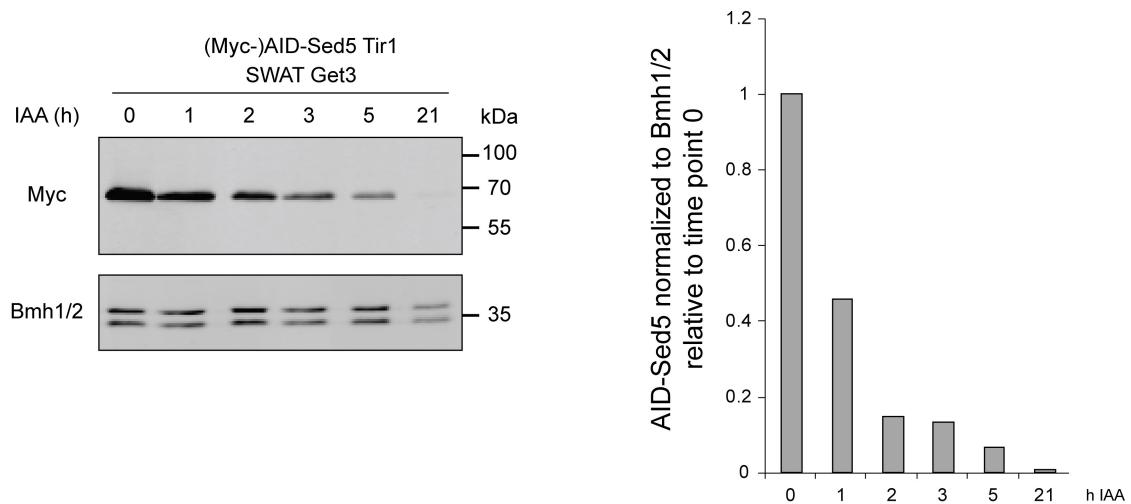
Supplementary Figure 3
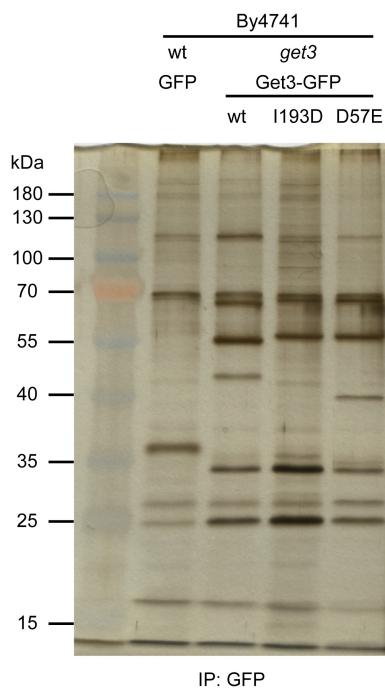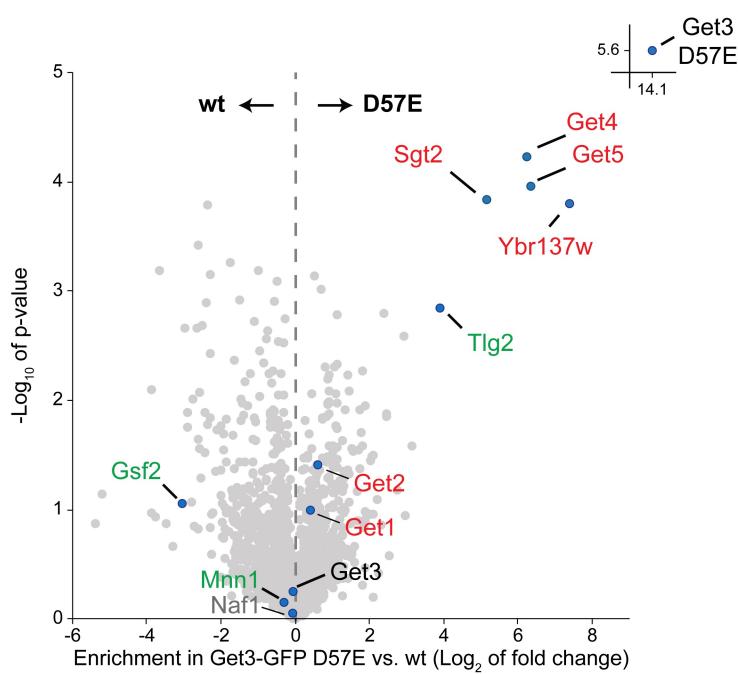
Supplementary Figure 4
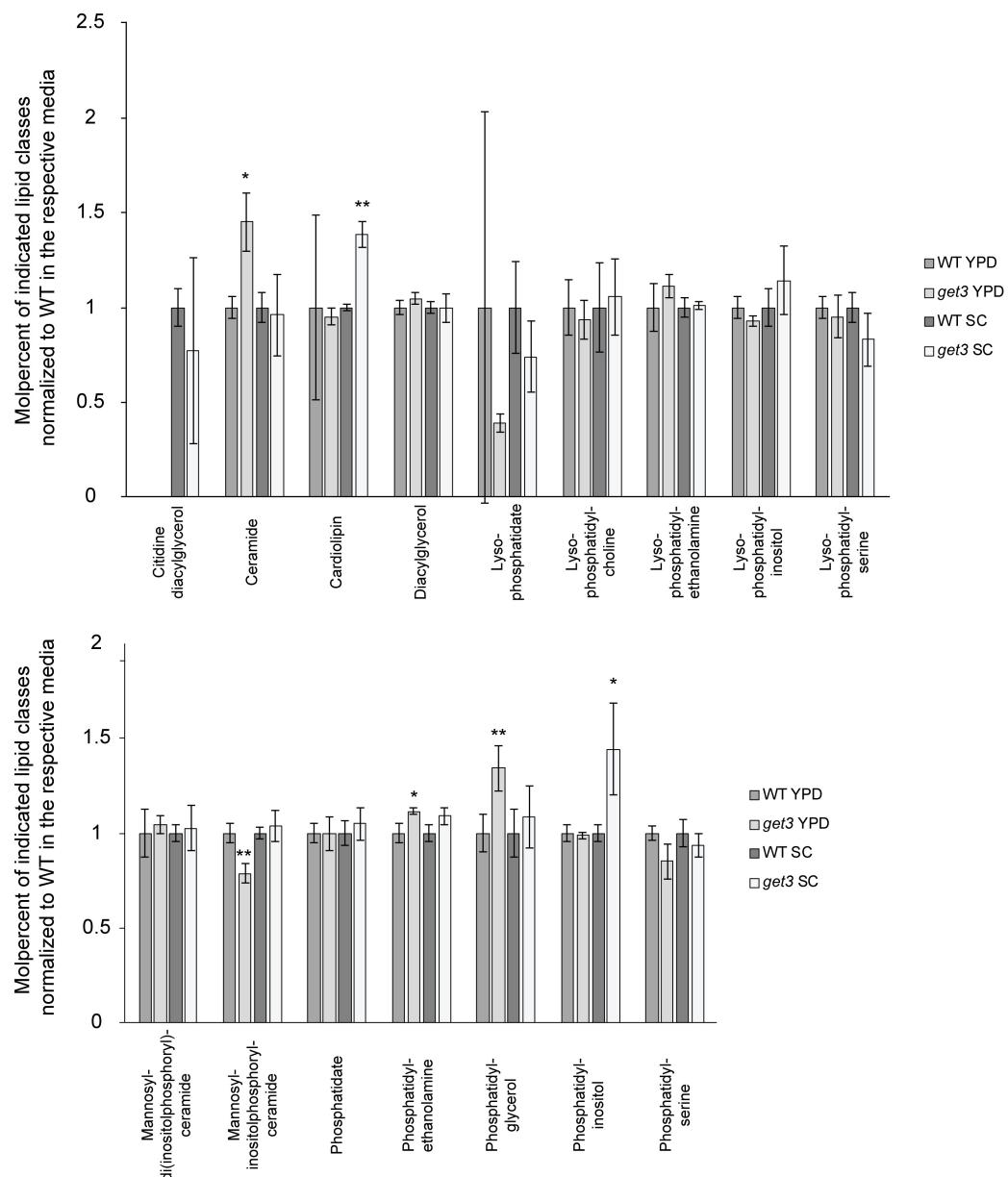
Supplementary Figure 5
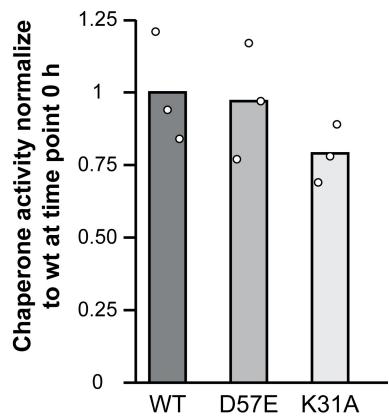
Supplementary Figure 6
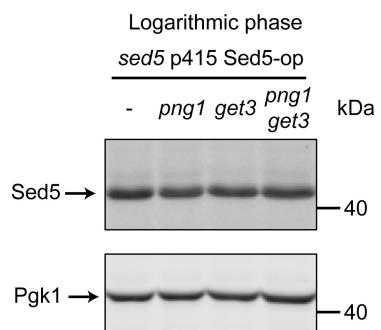
Table 1

Gene	Category	Localization	
		wt	**Δget3**
MRX11	perturbed mitochondrial	faint mitochondria	mitochondria and puncta
AIM11	perturbed mitochondrial	some puncta and mitochondria	more puncta appear
AIM20	perturbed mitochondrial	faint ER, some mitochondria, puncta	only puncta and mitochondria
CRD1	perturbed mitochondrial	faint ER or mitochondria	bright mitochondria
MRX9	perturbed mitochondrial	ER-puncta in wt	mitochondria
YBL039W-B	perturbed mitochondrial	Mitochondria, puncta in wt	brighter puncta
YPR117W	perturbed mitochondrial	several fainter puncta	most puncta lost, only few faint puncta
YOL162W	perturbed mitochondrial	faint ER	several small puncta
SHH3	perturbed mitochondrial	ER, mitochondria	only mitochondria
SYM1	perturbed mitochondrial	mitochondria and some ER	only mitochondria
FUN14	perturbed mitochondrial	ER	mitochondria
MIM1	perturbed mitochondrial	mitochondria	brighter mitochondria
NDI1	perturbed mitochondrial	faint mitochondria	brighter mitochondria
FMP32	perturbed mitochondrial	faint mitochondria	brighter mitochondria
PEX25	affected peroxisomal	puncta	fainter puncta and ER
PNC1	affected peroxisomal	puncta (less strong than other peroxisomal matrix proteins)	fainter puncta
GTO1	affected peroxisomal	puncta	fainter puncta
PEX4	affected peroxisomal	puncta (less strong than other peroxisomal matrix proteins)	fainter puncta
AAT2	affected peroxisomal	puncta (less strong than other peroxisomal matrix proteins)	fainter puncta
ANT1	affected peroxisomal	puncta	fainter puncta and ER often appears
BET4	affected peroxisomal	cytosol, faint punctum	only cytosol
CAT2	affected peroxisomal	puncta	fainter puncta
CIT2	affected peroxisomal	puncta	fainter puncta
CTA1	affected peroxisomal	puncta	fainter puncta
DCI1	affected peroxisomal	puncta	fainter puncta
ECI1	affected peroxisomal	puncta	fainter puncta
FAA2	affected peroxisomal	puncta	fainter puncta
FOX2	affected peroxisomal	puncta	fainter puncta

Gene	Category	Localization	
		wt	Δget3
GPD1	affected peroxisomal	puncta	fainter puncta
IDP3	affected peroxisomal	puncta	fainter puncta
LPX1	affected peroxisomal	puncta	fainter puncta
MDH3	affected peroxisomal	puncta	fainter puncta
MLS1	affected peroxisomal	small, clear puncta	fewer and much fainter puncta
NPY1	affected peroxisomal	puncta	much weaker puncta
PCS60	affected peroxisomal	puncta	fainter puncta
PEX11	affected peroxisomal	puncta	fainter puncta
PEX12	affected peroxisomal	puncta	fainter puncta
PEX13	affected peroxisomal	puncta (less strong than other peroxisomal matrix proteins)	fainter puncta
PEX14	affected peroxisomal	puncta	fainter puncta
PEX17	affected peroxisomal	puncta	fainter puncta
PEX18	affected peroxisomal	puncta (less strong than other peroxisomal matrix proteins)	fainter puncta
PEX21	affected peroxisomal	weak puncta and cytosol	fewer and fainter puncta
PEX27	affected peroxisomal	strong punctum	punctum fainter in, several small puncta appear often
PEX34	affected peroxisomal	puncta	more puncta
PEX6	affected peroxisomal	puncta (1-2) and cytosol	more puncta (2-4)
PEX7	affected peroxisomal	puncta	fainter puncta
PEX8	affected peroxisomal	puncta	fainter puncta
POT1	affected peroxisomal	puncta	fainter puncta
POX1	affected peroxisomal	puncta	fainter puncta
PXA2	affected peroxisomal	puncta	fainter puncta
PXP1	affected peroxisomal	puncta	fainter puncta
SPS19	affected peroxisomal	puncta	fainter puncta
STR3	affected peroxisomal	puncta	fainter puncta
TES1	affected peroxisomal	puncta	fainter puncta
YMR018W	affected peroxisomal	puncta	puncta lost or fainter
CUE4	perturbed other membrane	ER, nucleus	ER mostly lost, punctum appears sometimes
SNL1	perturbed other membrane	ER, puncta	only puncta
CUE1	perturbed other membrane	ER, punctum	cytosol, faint punctum
CBR1	perturbed other membrane	faint ER, nucleus, rarely punctum	only punctum, sometimes puncta

Gene	Category	Localization	
		wt	**Δget3**
DPL1	perturbed other membrane	puncta, sometimes faint ER	cytosol or rarely punctum
UBP1	perturbed other membrane	faint ER	cytosol
STB4	perturbed other membrane	faint ER, some puncta	nucleus
ERG3	perturbed other membrane	faint ER	cytosolic in most cells, rarely faint ER
LCB1	perturbed other membrane	ER, rarely punctum	fainter ER, often punctum
OST4	perturbed other membrane	ER, mitochondria	only mitochondria

Table 2

Name	Reference / Source	Description	Used in figure	Made by
p415Met25	(Mumberg et al., 1994)	Yeast expression plasmid for expressing proteins under control of the *MET25* promoter.	None; used for creating other plasmids	
p415Met25 yGFP	This study	Yeast expression plasmid for expressing cytosolic yeast enhanced GFP as found in pKT128 (Sheff and Thorn, 2004).	Fig 2A	Ákos Farkas
p415Met25 Get3-yGFP	Schwappach group	Yeast expression plasmid for expressing Get3 tagged with yGFP C-terminally.	Fig 2A	
P415Met25 Get3-yGFP I193D	Schwappach group	Yeast expression plasmid for expressing Get3 I193D tagged with yGFP C-terminally	Not part of manuscript (Fig S3A)	
p415Met25 Get3-yGFP D57E	(Powis et al., 2013)	Yeast expression plasmid for expressing Get3 D57E tagged with yGFP C-terminally.	Fig 2A	
p415Met25 mTagBFP2-Get3 D57E	Schwappach group	Yeast expression plasmid for expressing Get3 D57E tagged with mTagBFP2 N-terminally.	Fig 2B; Fig 3D; Fig 4B, E; Fig 5C	Ákos Farkas
p415Met25 mTagBFP2-Get3 K31A	This study	Yeast expression plasmid for expressing Get3 K31A tagged with mTagBFP2 N-terminally.	Fig 3D-E	Ákos Farkas
p415Met25 mTagBFP2-Get5	This study	Yeast expression plasmid for expressing Get5 tagged with mTagBFP2 N-terminally.	Fig 4E	Ákos Farkas
p415Met25 HA-Vph2	This study	Yeast expression plasmid for expressing Vph2 tagged with HA N-terminally.	Fig 4C	Ákos Farkas
p415Met25 HA-Atf2	This study	Yeast expression plasmid for expressing Atf2 tagged with HA N-terminally.	Fig 5D-G	Ákos Farkas
p415Met25 HA-Atf2 dNAH	This study	Yeast expression plasmid for expressing Atf2 lacking NAH tagged with HA N-terminally.	Fig 5E	Ákos Farkas
p415Met25 HA-Atf2 dCAH	This study	Yeast expression plasmid for expressing Atf2 lacking CAH tagged with HA N-terminally.	Fig 5E	Ákos Farkas
p415Met25 HA-Atf1	This study	Yeast expression plasmid for expressing Atf1 tagged with HA N-terminally.	Fig 5H-I	Ákos Farkas
p415Met25 sfGFP-Sed5	This study	Yeast expression plasmid for expressing Sed5 tagged with sfGFP N-terminally.	Fig 6C	Ákos Farkas
p415Met25 sfGFP-Sed5-Gos1TM	This study	Yeast expression plasmid for expressing the chimera of the cytosolic domain of Sed5 and the TMS of Gos1 tagged with sfGFP N-terminally.	Fig 6C-D	Ákos Farkas
p415Met25 sfGFP-Gos1	This study	Yeast expression plasmid for expressing Gos1 tagged with sfGFP N-terminally.	Fig 6C	Ákos Farkas
p415Met25 sfGFP-Gos1-Sed5TM	This study	Yeast expression plasmid for expressing the chimera of the cytosolic domain of Gos1 and the TMS of Sed5 tagged with sfGFP N-terminally.	Fig 6C-D	Ákos Farkas
P415Met25 sfGFP-Sed5-Gos1TM(-1)	This study	Yeast expression plasmid for expressing the chimera of the cytosolic domain of Sed5 and the TMS of Gos1 lacking its last amino acid (tryptophane) tagged with sfGFP N-terminally.	Fig 6F	Ákos Farkas
p415Met25 sfGFP-Sed5-ykt6	This study	Yeast expression plasmid for expressing Sed5 with its TMS replaced by the lipidation sequence of Ykt6 tagged with sfGFP N-terminally.	Fig 6E	Ákos Farkas
p415Met25 sfGFP-Gos1-ykt6	This study	Yeast expression plasmid for expressing Gos1 with its TMS replaced by the lipidation sequence of Ykt6 tagged with sfGFP N-terminally.	Fig 6E	Ákos Farkas
p415Met25 Sed5-op	This study	Yeast expression plasmid for expressing Sed5 carrying a C-terminal opsin tag.	Fig 7C-F	Ákos Farkas

Name	Reference	Description	Used in figure	Made by
p416Met25	(Mumberg et al., 1994)	Yeast expression plasmid for expression of proteins under control of the MET25 promoter.	None; used for creating other plasmids	
P416Met25 mCherry-Sed5	This study	Yeast expression plasmid for expressing Sed5 with an N-terminal mCherry tag.	None; used for creating strains.	Ákos Farkas
p416Met25 yGFP	This study	Yeast expression plasmid for expressing cytosolic yeast enhanced GFP as found in pKT128 (Sheff and Thorn, 2004).	Fig 5D, H-I; Fig 7E	Ákos Farkas
p416Met25 Get3-TEV-yGFP	This study	Yeast expression plasmid for expressing Get3 tagged with yGFP C-terminally. The linker sequence included the recognition sequence of TEV protease.	Fig 4C-D; Fig 5D-I; Fig 7E-F	Ákos Farkas
p416Met25 Get3-TEV-yGFP D57E	This study	Yeast expression plasmid for expressing Get3 D57E tagged with yGFP C-terminally. The linker sequence included the recognition sequence of TEV protease.	Fig 4C-D; Fig 5D, H-I; Fig 7E	Ákos Farkas
p416Met25 Get3-TEV-yGFP FIDD	This study	Yeast expression plasmid for expressing Get3 F190D I193D tagged with yGFP C-terminally. The linker sequence included the recognition sequence of TEV protease.	Fig 5F-I; Fig 7F	Ákos Farkas
p416Met25 sfGFP-Atf2 dNAH	This study	Yeast expression plasmid for expressing Atf2 lacking NAH tagged with sfGFP N-terminally.	Fig 5B	Ákos Farkas
p416Met25 sfGFP-Atf2 dCAH	This study	Yeast expression plasmid for expressing Atf2 lacking CAH tagged with sfGFP N-terminally.	Fig 5B	Ákos Farkas
p416Met25 mTagBFP2-Get3 D57E	Schwappach group	Yeast expression plasmid for expressing Get3 D57E tagged with mTagBFP2 N-terminally.	Fig 6D-E	Ákos Farkas
p416Met25 Sed5-op	This study	Yeast expression plasmid for expressing Sed5 carrying a C-terminal opsin tag.	Fig 7B	Ákos Farkas
p416Met25 Sed5-opNQ	This study	Yeast expression plasmid for expressing Sed5 carrying a C-terminal opsin tag in which the glycosylation acceptor asparagine was mutated to glutamine.	Fig 7B	Ákos Farkas
p426 GAL1	(Mumberg et al., 1994)	Yeast expression plasmid that allows expression of proteins under control of the GAL1 promoter.	Fig 3C	
p426 GAL1 Get3	This study	Yeast expression plasmid for expressing Get3.	Fig 3C	Ákos Farkas
p426 GAL1 Get3 D57E	This study	Yeast expression plasmid for expressing Get3 D57E.	Fig 3C	Ákos Farkas
p426 GAL1 Get3 D57E FIDD	This study	Yeast expression plasmid for expressing Get3 D57E F190D I193D	Fig 3C	Ákos Farkas
p426 GAL1 Get3 K31A	This study	Yeast expression plasmid for expressing Get3 K31A	Fig 3C	Ákos Farkas
p426 GAL1 Get3 K31A FIDD	This study	Yeast expression plasmid for expressing Get3 K31A F190D I193D	Fig 3C	Ákos Farkas
pCEV KAN NAT BLE	Schwappach group	Plasmid encoding a Nat^R (*ScTEFpr::nat::ScADH1t*); Ble^R (*ScPGK1pr::ble::ScCyc1t*); Kan^R (*AgTEFpr::kan::AgTeft*) resistance cassette.	None; used for creating strains	
pAG25	(Goldstein and McCusker, 1999)	Plasmid encoding a Nat^R resistance cassette	None; used for creating strains	

Table 3

Name	Insert/Inserts	Plasmid backbone	Restriction enzymes used	Method
p415Met25 yGFP	AF1 and AF2 was used to amplify GFP-coding sequence from pKT128.	p415Met25	HindIII, XhoI	T4 ligase
p415Met25 mTagBFP2-Get3 D57E	Two constructs amplified: 1) AF21 + AF22 from a plasmid carrying mTagBFP2, then cut with HindIII, NcoI; 2) AF23 + AF24 from p416 Get3 D57E (Voth et al., 2014), then cut with NcoI, XhoI	p415Met25	Backbone: HindIII, Xho; Insert: NcoI	T4 ligase
p415Met25 mTagBFP2-Get3 K31A	First, two constructs were made: 1) AF25 + AF26 with p416 Get3 (Voth et al., 2014); 2) AF23 + AF27 with p416 Get3 (Voth et al., 2014). Using these as templates, the final construct 4) was made with AF23 + AF25, then cut with NcoI, XhoI. The other insert was from cutting p415Met25 mTagBFP2-Get3 D57E with XbaI and XhoI.	p415Met25	Backbone: HindIII, Xho; Insert with NcoI	T4 ligase
p415Met25 mTagBFP2-Get5	Two constructs: 1) p415Met25 mTagBFP2-Get3 D57E with AF14 + AF15; 2) from By4741 genomic DNA with AF16 and AF17	p415Met25	XbaI, XhoI	Gibson assembly
p415Met25 HA-Vph2	AF3 and AF4 was used to amplify HA-Vph2 cassette from By4741 genomic DNA.	p415Met25	XbaI, XhoI	Gibson assembly
p415Met25 HA-Atf2	AF5 and AF12 was used amplify HA-Atf2 cassette from By4741 genomic DNA	p415Met25	XbaI, XhoI	Gibson assembly
p415Met25 HA-Atf2 dNAH	Two construct were amplified from p415Met25 HA-Atf2 with the following primer pairs: 1) AF5 and AF7 2) AF8 and AF12	p415Met25	XbaI, XhoI	Gibson assembly
p415Met25 HA-Atf2 dCAH	AF5 and AF13 was used to amplify HA-Atf2 dCAH cassette from p416Met25sfGFP-Atf2 dCAH	p415Met25	XbaI, XhoI	Gibson assembly
p415Met25 HA-Atf1	AF6 and AF10 was used to amplify HA-Atf1 cassette from By4741 genomic DNA	p415Met25	XbaI, XhoI	Gibson assembly
p415Met25 sfGFP-Sed5	Insert amplified with AF41 + AF42 from genomic DNA from SWAT Sed5. The backbone used was p415Met25 mCherry-Sed5 generated exactly as described for p416Met25 mCherry-Sed5 but in p415 background.	p415Met25	HindIII, XhoI	T4 ligase
p415Met25 sfGFP-Sed5-Gos1TM	Insert amplified with AF41 + AF50 from p415Met25 sfGFP-Sed5.	p415Met25	HindIII, XhoI	T4 ligase
p415Met25 sfGFP-Gos1	Insert 1) cut from p415Met25 sfGFP-Sed5 with HindIII, EcoRI; 2) amplified from By4741 genomic DNA with AF39 + AF48, cut with EcoRI, XhoI.	p415Met25	Backbone: HindIII, XhoI; Insert: EcoRI	T4 ligase
p415Met25 sfGFP-Gos1-Sed5TM	Insert 1) cut from p415Met25 sfGFP-Sed5 with HindIII, EcoRI; 2) amplified from By4741 genomic DNA with AF39 + AF49, cut with EcoRI, XhoI.	p415Met25	Backbone: HindIII, XhoI; Insert: EcoRI	T4 ligase
p415Met25 sfGFP-Sed5-Gos1TM(-1)	Insert amplified with AF41 + AF51 from p415Met25 sfGFP-Sed5.	p415Met25	HindIII, XhoI	T4 ligase

Name	Insert/Inserts	Plasmid backbone	Restriction enzymes used	Method
p415Met25 sfGFP-Sed5-ykt6	Insert amplified with AF41 + AF52 from p415Met25 sfGFP-Sed5.	p415Met25	HindIII, XhoI	T4 ligase
p415Met25 sfGFP-Gos1-ykt6	Insert 1) cut from p415Met25 sfGFP-Sed5 with HindIII, EcoRI; 2) amplified from By4741 genomic DNA with AF39 + AF40, cut with EcoRI, XhoI.	p415Met25	Backbone: HindIII, XhoI; Insert: EcoRI	T4 ligase
p415Met25 Sed5-op	p416Met25 Sed5-op was cut with XbaI and HindIII	p415Met25	XbaI, HindIII	T4 ligase
p416Met25 mCherry-Sed5	Two inserts were amplified: 1) AF45 + AF46 with p415 mCherry-Sed5 (Powis et al., 2013) cut with XbaI, BamHI; 2) AF47 + AF42 with By4741 genomic DNA cut with BamHI, XhoI.	p416Met25	XbaI, BamHI, XhoI	T4 ligase
p416Met25 yGFP	p415Met25 yGFP was cut with HindIII, XhoI	p416Met25	HindIII, XhoI	T4 ligase
p416Met25 Get3-TEV-yGFP	Constructs: 1) was amplified from p416 Get3 (Voth et al., 2014) with AF33 + AF34. 2) was made by cutting 1) with BcuI, EcoRI and taking the 700 bp piece. 3) was made by cutting 1) with EcoRI and BamHI and taking the 300 bp piece. 4) was made with AF34 + AF36 from pKT128, then cut with BamHI and XhoI. 2), 3), and 4) were ligated.	p416Met25	Backbone: BcuI, XhoI; Insert: EcoRI, BamHI	T4 ligase
p416Met25 Get3-TEV-yGFP D57E	Insert was cut from p426 GAL1 Get3 D57E with BcuI and EcoRI	p416Met25	BcuI, EcoRI	
p416Met25 Get3-TEV-yGFP FIDD	Amplified two constructs from p416Met25 Get3-TEV-yGFP; 1) AF37 + AF30; 2) AF31 + AF38. p416Met25 Get3-TEV-yGFP was used as backbone.	p416Met25	BcuI, EcoRI	Gibson assembly
p416Met25 sfGFP-Atf2 dNAH	Insert was amplified using AF11 + AF12 with p415Met25 HA-Atf2 dNAH. Backbone was p416Met25 sfGFP-Atf2 dCAH.	p416Met25	EcoRI, XhoI	Gibson assembly
p416Met25 sfGFP-Atf2 dCAH	Insert was amplified using AF11 + AF13 with p415Met25 HA-Atf2 dCAH. Backbone was p416Met25 sfGFP-Sed5, which had been made identically to p415Met25 sfGFP-Sed5 but in the p416 background.	p416Met25	EcoRI, XhoI	Gibson assembly
p416Met25 mTagBFP2-Get3 D57E	Insert was cut from p415Met25 mTagBFP2-Get3 D57E with XbaI, XhoI.	p416Met25	XbaI, XhoI	T4 ligase
p416Met25 Sed5-op	AF18 and AF19 was used to amplify Sed5-op from By4741 genomic DNA.	p416Met25	XbaI, HindIII	T4 ligase
p416Met25 Sed5-opNQ	AF18 and AF20 was used to amplify Sed5-opNQ from p416Met25 Sed5-op.	p416Met25	XbaI, HindIII	T4 ligase
p426 GAL1 Get3	Insert was cut from p416 Get3 (Voth et al., 2014)	p426 GAL1	BcuI, XhoI	T4 ligase
p426 GAL1 Get3 D57E	Insert was cut from p416 Get3 D57E (Voth et al., 2014)	p426 GAL1	BcuI, XhoI	T4 ligase
p426 GAL1 Get3 D57E FIDD	Two constructs were amplified from p426 GAL1 Get3 D57E with the following primer pairs: 1) AF29 + AF30 2) AF31 + AF32	p426 GAL1	BcuI, XhoI	Gibson assembly

Name	Insert/Inserts	Plasmid backbone	Restriction enzymes used	Method
p426 GAL1 Get3 K31A	Insert was amplified from p415Met25 mTagBFP2-Get3 K31A using AF28 and AF25.	p426 GAL1	BcuI, XhoI	T4 ligase
p426 GAL1 Get3 K31A FIDD	Two constructs were amplified from p426 GAL1 Get3 K31A with the following primer pairs: 1) AF29 + AF30 2) AF31 + AF32	p426 GAL1	BcuI, XhoI	Gibson assembly

Table 4

ID	Name	Sequence (5' to 3')
AF1	HindIII_yGFP_F	agtaataagcttATGTCTAAAGGTGAAGAATTATTCACTG
AF2	yGFPstop_XhoI_R	tgtcgactcgagTTATTTGTACAATTCATCCATACCATGG
AF3	GibsonMet25prXbaI_HA_Vph2_F	ttacccccatccatactctagaatgccatacccatcgatgttccagattacgctATGTTCGAAATTAAACTGAATGATAGAATAAC
AF4	Vph2stop_XhoI_GibsonCyc1t_R	gaatgtaagcgtgacataactaattacatgactcgagTTACAGTGTAATCTTGCTTAGAACTTTC
AF5	Met25pr_XbaI_HA_Atf2_F	CATAGATACAATTCTATTACCCCCATCCATACTCTAGAATGccatacccatcgatgttccagattacgctATGGAAGATATAGAAGGATACGAACCAC
AF6	Met25pr_XbaI_HA_Atf1_F	CATAGATACAATTCTATTACCCCCATCCATACTCTAGAATGccatacccatcgatgttccagattacgctATGAATGAAATCGATGAGAAAAATCAGG
AF7	Atf2_dNAH_R	GTAAAATTCGAGTACATTTTCTGACGGTCTATCAACTCTTGAGTGATATGTGG
AF8	Atf2_dNAH_F	CATATCACTCAAGAGTTGATAGACCGTCAGAAAATGTACTCGAATTTTACTGTTTACGC
AF9	linkerS4_Atf1_F	gtggtggtgcgacagagaattcatcgATGAATGAAATCGATGAGAAAAATCAGG
AF10	Atf1stop_XhoI_Cyc1Gibson_R	gaatgtaagcgtgacataactaattacatgactcgagCTAAGGGCCTAAAAGGAGAGCTTTG
AF11	linkerS4_Atf2_F	gtggtggtgcgacagagaattcatcgATGGAAGATATAGAAGGATACGAACCAC
AF12	Atf2stop_XhoI_Cyc1Gibson_R	gaatgtaagcgtgacataactaattacatgactcgagTTAAAGCGACGCAAATTCGCCG
AF13	Atf2dCAHstop_XhoI_Cyc1Gibson_R	gaatgtaagcgtgacataactaattacatgactcgagTTAATCCCGTAGAGTGCCCTGAACC
AF14	Met25prGibson_XbaI_mTagBFP2_F	CATAGATACAATTCTATTACCCCCATCCATACTCTAGAATGGTGTCTAAGGGCGAAGAG
AF15	mTagBFP2_linkerS4_R	ATcgatgaattctctgtcgcaccaccaccaccagaggaaccgccATTAAGCTTGTGCCCCAGTTTG
AF16	linkerS4_Get5_F	gtggtggtgcgacagagaattcatcgATGAGCACATCCGCCAGCGGTC
AF17	Get5stop_XhoI_Cyc1tGibson_R	gaatgtaagcgtgacataactaattacatgactcgagTTATTTGGCCAGAGACCAGCC
AF18	XbaI-XmaI-Sed5_F	TACTCTAGACCCGGGATGAACATAAAGGATAGAACTTCAG
AF19	op-link-Sed5_R	GATAAGCTTTTAGCCCGTCTTGTTGGAGAAAGGCACGTAGAAGTTTGGGCCggcgccggcGGATCCGCCGCCGCCATTGACTAAAACCCAAATAACG
AF20	opsinNQstop-HindIII_R	ATCGATAAGCTTTTAGCCCGTCTTCTGGGAGAAAGGCACGTAG
AF21	HinDIII_mTagBFP2_F	CCATCCTCTAGAAAGCTTATGGTGTCTAAGGGCGAAGAG
AF22	mTagBFP2_G4S_BamHI_R	GTTCATGGATCCGCCGCCGCCCGAGCCGCCGCCGCCATTAAGCTTGTGCCCCAGTTTG
AF23	G4S_Get3_F	GGCGGCGGCGGCTCGGGCGGCGGCGGATCCatggatttaaccgtggaacctaatttg
AF24	XhoI_Cyc1term_R	GTGACATAACTAATTACATGACTCGAGG
AF25	Get3stop_HindIII_R1	ACTAATTACATGACTCGAGGTCGACGGTATCGATAAGCTTctattccttatcttctaactcataaatgac

ID	Name	Sequence (5' to 3')
AF26	Get3_K31A_F	GGTGGTGTTGGAGCGACTACTTCATCATGTTCC
AF27	Get3_K31A_R	GGAACATGATGAAGTAGTCGCTCCAACACCACC
AF28	SmaI_Get3_F2	CTCTAGAACTAGTGGATCCCCCGGGatggatttaaccgtggaacctaatttg
AF29	GALpr_F	ggtaattaatcagcgaagcg
AF30	Get3_F190D_I193D_R	ccaatttgttggtGTCttcaccgTCcttttccaaaagcttgg
AF31	Get3_F190D_I193D_F	ccaagcttttggaaaagGAcggtgaaGACaccaacaaattgg
AF32	Cyc1term_R2	CCTTTTCGGTTAGAGCGGATG
AF33	BcuI_Get3_F	TCTAGAACTAGTGTCGACatggatttaaccgtggaacctaatttg
AF34	BamHI_TEV_G4S_yGFP_F	GGTGGCGGATCCGAGAATTTGTATTTTCAGGGTGGCGGCGGCGGCTCGGGCGGCGGCGGTTCCatgtctaaaggtgaagaattattcactgg
AF35	Get3_G4S_TEV_R	CGCCACCCTGAAAATACAAATTCTCGGATCCGCCACCGCCGGAGCCACCGCCACCttccttatcttctaactcataaatgactttgc
AF36	yGFPstop_XhoI_R	tgtcgactcgagTTATTTGTACAATTCATCCATACCATGG
AF37	Met25pr_XbaI_SpeI_Get3_F	CATAGATACAATTCTATTACCCCCATCCATACTCTAGAACTAGTATGGATTTAACCGTGGAACCTAATTTG
AF38	Get3_i_R	ggaaatcaattcctgaattagtctttcag
AF39	EcoRI_linkerS4_Gos1_F	gacagagaattcatcgATGAGCTCACAACCGTCTTTCG
AF40	Gos1_ykt6_XhoI_R	ACATGACTCGAGTTACATGATGATGCAACACGAATTGGAGTTTTTCTTCCTTCTCGTGTTAATCTTC
AF41	HinDIII_sfGFP_F	agtaataagcttATGTCCAAGGGTGAAGAGCTATTTAC
AF42	C-Sed5_XhoI_R	ACATGACTCGAGTTAATTGACTAAAACCCAAATAACGAAAAATACAAAG
AF43	Sed5noTM_Gos1TMWstop_XhoI_R	ACATGACTCGAGTTACCATGTGAAAAACAAAAACAGTATACAAAGGGTGGTTATCGTGGCCAATACAAACGCTCTATTACTCTTTATCCTGTCGAAGTATTTC
AF44	EcoRI_linkerS4_Sed5_F	gacagagaattcatcgATGAACATAAAGGATAGAACTTCAGAATTTC
AF45	HinDIII_mCherry_F	CCATCCTCTAGAAAGCTTATGGTGAGCAAGGGCGAGGAG
AF46	mCherry_G4S_BamHI_R	GTTCATGGATCCGCCGCCGCCCGAGCCGCCGCCGCCCTTGGACAGCTCGTCCATGCCG
AF47	BamHI_N-Sed5_F	GGCGGCGGATCCATGAACATAAAGGATAGAACTTCAGAATTTCAAC
AF48	Gos1stop_XhoI_R	ACATGACTCGAGTTATTACCATGTGAAAAACAAAAACAGTATACAAAG
AF49	Gos1noTM_Sed5TMstop_XhoI_R	ACATGACTCGAGTTATTAATTGACTAAAACCCAAATAACGAAAAATACAAAGATTATAAAAAAAACCTTTGCGGCTAACCAGTTTTTCTTCCTTCTCGTGTTAATCTTC
AF50	Sed5noTM_Gos1TMWstop_XhoI_R	ACATGACTCGAGTTACCATGTGAAAAACAAAAACAGTATACAAAGGGTGGTTATCGTGGCCAATACAAACGCTCTATTACTCTTTATCCTGTCGAAGTATTTC
AF51	Sed5noTM_Gos1TMstop_XhoI_R	ACATGACTCGAGTTATGTGAAAAACAAAAACAGTATACAAAGGGTGGTTATCGTGGCCAATACAAACGCTCTATTACTCTTTATCCTGTCGAAGTATTTC
AF52	Sed5noTM_ykt6lipidstop_R	ACATGACTCGAGTTACATGATGATGCAACACGAATTGGATCTATTACTCTTTATCCTGTCGAAGTATTTC

Table 5

Name	Genotype	Source	Used in figure	Construction	Made by
SWAT library	By4741 *MATa* SWAT-sfGFP-X *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	(Yofe et al., 2016)	Fig 1A; Fig S1A		
SGA query	*MATalpha his3Δ1, leu2Δ0, LYS2+, met15Δ0, ura3Δ0, can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2* (S288C background)	Kind gift from Maya Schuldiner	None; Used to create other strains		
By4741	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	(Brachmann et al., 1998)	Fig 5B, D, H-I; Fig 6C; Fig 7B		
By4741 *Δget3*	By4741 *get3Δ::Kan^R*	(Schuldiner et al., 2008)	Fig 3C; Fig 4C; Fig 5D-I; Fig 6C-E		
SWAT library *Δget3*	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, get3Δ::Nat^R*	This study	Fig 1A-D; Fig S1A	Made with SGA method by crossing SWAT library with SGA query *Δget3*	Ákos Farkas
SWAT library strains *Δmet15*	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, met15Δ::Nat^R*	This study	Fig 1C	Made with SGA method by crossing SWAT library strains with SGA query *Δmet15*	Ákos Farkas
SWAT library strains *Δget4*	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, get4Δ::Nat^R*	This study	Fig 1C	Made with SGA method by crossing SWAT library with SGA query *Δget4*	Ákos Farkas
SWAT library strains *Δget5*	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, get5Δ::Nat^R*	This study	Fig 1C	Made with SGA method by crossing SWAT library strains with SGA query *Δget5*	Ákos Farkas
SWAT library strains *Δsgt2*	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, sgt2Δ::Nat^R*	This study	Fig 1C	Made with SGA method by crossing SWAT library strains with SGA query *Δsgt2*	Ákos Farkas

Name	Genotype	Source	Used in figure	Construction	Made by
SWAT library strains AID-SED5 Tir1	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, ScCYC1t*(rev)::*KAN^R*(rev)::*ScPGK1pr*(rev)::*ScTEFpr::AtTIR1::ScADHt::SpNop1pr*::6xmyc-AID-Sed5	This study	Fig 1E-F	Made with SGA method by crossing SWAT library strains with SGA query AID-Sed5 Tir1	Anne Clancy
SWAT library *Δget1 Δget2 Δget3*	SWAT library; *can1Δ::STE2pr-spHIS5, lyp1Δ::STE3pr-LEU2, get1Δ::Kan^R get2Δ::Nat^R get3Δ::Ble^R*	This study	None, used to create other strains	Made with SGA method by crossing SWAT library with SGA query *Δget1 Δget2 Δget3*	Ákos Farkas
SWAT library *Δget1/Δget1 Δget2/Δget2 Δget3/Δget3* p415 mTagBFP2-Get3 D57E	SWAT library; *CAN1/can1Δ::STE2pr-spHIS5, LYP1/lyp1Δ::STE3pr-LEU2, get1Δ::Kan^R/get1Δ::Kan^R, get2Δ::Nat^R/get2Δ::Nat^R get3Δ::Ble^R/get3Δ::Ble^R*, p415Met25 mTagBFP2-Get3 D57E	This study	Fig 2B-C; Fig 3D-E; Fig 4B, E; Fig 5C	SWAT library *Δget1 Δget2 Δget3* was mated with alpha *Δget1 Δget2 Δget3* transformed with p415Met25 mTagBFP2-Get3 D57E; diploid cells were copied onto synthetic dropout plates lacking leucine and uracil.	Ákos Farkas
SWAT library *Δget1/Δget1 Δget2/Δget2 Δget3/Δget3* p415 mTagBFP2-Get3 K31A	SWAT library; *CAN1/can1Δ::STE2pr-spHIS5, LYP1/lyp1Δ::STE3pr-LEU2, get1Δ::Kan^R/get1Δ::Kan^R, get2Δ::Nat^R/get2Δ::Nat^R, get3Δ::Ble^R/get3Δ::Ble^R*, p415Met25 mTagBFP2-Get3 K31A	This study	Fig 3D-E	SWAT library *Δget1 Δget2 Δget3* was mated with alpha *Δget1 Δget2 Δget3* transformed with p415Met25 mTagBFP2-Get3 K31A; diploid cells were copied onto synthetic dropout plates lacking leucine and uracil.	Ákos Farkas

Name	Genotype	Source	Used in figure	Construction	Made by
SWAT library strains *GET1/Δget1 GET2/Δget2 Δget3/Δget3* p415 mTagBFP2-Get5	SWAT library; *CAN1/can1Δ::STE 2pr-spHIS5, LYP1/lyp1Δ::STE3 pr-LEU2, GET1/get1Δ::Kan^R GET2/get2Δ::Nat^R, get3Δ::Kan^R/get3Δ: :Ble^R*, p415Met25 mTagBFP2-Get5	This study	Fig 4F	SWAT library *Δget1 Δget2 Δget3* was mated with alpha *Δget3* transformed with p415Met25 mTagBFP2-Get5; diploid cells were copied onto synthetic dropout plates lacking leucine and uracil.	Ákos Farkas
By4741 *Δsed5* p416 mCherry-Sed5	By4741 *sed5Δ::Nat^R* p416Met25 mCherry-Sed5	This study	None; used to create other strains	By4741 was transformed with p416Met25 mCherry-Sed5. The *SED5* locus was replaced with Nat^R amplified from pCEV KAN NAT BLE.	Ákos Farkas
By4741 *Δsed5* Sed5-op	By4741 *sed5Δ::Nat^R* p415Met25 Sed5-op	This study	Fig S6A; Fig 7C-D	By4741 *Δsed5* p416 mCherry-Sed5 was transformed with p415Met25 Sed5-op. The final strain was made from the resulting strain with the 5-FOA method.	Ákos Farkas
By4741 *Δsed5 Δget3* Sed5-op	By4741 *sed5Δ::Nat^R get3Δ::Ble^R* p415Met25 Sed5-op	This study	Fig S6A; Fig 7C-D, F	The GET3 locus was replaced *Ble^R* amplified from pCEV KAN NAT BLE in By4741 *Δsed5* p416 mCherry-Sed5. The resulting strain was transformed with p415Met25 Sed5-op. The final strain was made from this strain with the 5-FOA method.	Ákos Farkas
By4741 *Δsed5 Δpng1* Sed5-op	By4741 *sed5Δ::Nat^R png1Δ::Kan^R* p415Met25 Sed5-op	This study	Fig S6A; Fig 7C-E	The *PNG1* locus was replaced with the Kan^R cassette amplified from pCEV KAN NAT BLE in By4741 *Δsed5* p416 mCherry-Sed5. The resulting strain was transformed with p415Met25 Sed5-op. The final strain was made from this strain with the 5-FOA method.	Ákos Farkas

Name	Genotype	Source	Used in figure	Construction	Made by
By4741 *Δsed5 Δget3 Δpng1* Sed5-op	By4741 *sed5Δ::Nat^R get3Δ::Ble^R png1Δ::Kan^R* p415Met25 Sed5-op	This study	Fig S6A; Fig 7C-E, F	The *PNG1* locus was replaced with the *Kan^R* cassette, the *GET3* locus with *Ble^R* from pCEV KAN NAT BLE in By4741 *Δsed5* p416Met25 mCherry-Sed5. The resulting strain was transformed with p415Met25 Sed5-op. The final strain was made from this strain with the 5-FOA method.	Ákos Farkas
SGA query *Δmet15*	SGA query *met15Δ::Nat^R*	This study	Fig 2D	The *MET15* locus was replaced with the *Nat^R* cassette amplified from pAG25.	Ákos Farkas
SGA query *Δget3*	SGA query *get3Δ::Nat^R*	This study	Fig 2D	The *GET3* locus was replaced with the *Nat^R* cassette amplified from pAG25	Anne Clancy
SGA query *Δget4*	SGA query *get4Δ::Nat^R*	This study	Fig 2D	The *GET4* locus was replaced with the *Nat^R* cassette amplified from pAG25.	Ákos Farkas
SGA query *Δget5*	SGA query *get5Δ::Nat^R*	This study	Fig 2D	The *GET5* locus was replaced with the *Nat^R* cassette amplified from pAG25.	Ákos Farkas
SGA query *Δsgt2*	SGA query *sgt2Δ::Nat^R*	This study	Fig 2D	The *SGT2* locus was replaced with the *Nat^R* cassette amplified from pAG25.	Ákos Farkas
SGA query *Δget1 Δget2*	SGA query *get1Δ::Kan^R get2Δ::Nat^R*	This study	None; used to create other strains	The *GET1* locus was replaced by the Kan^R cassette, the *GET2* locus with the Nat^R cassette.	Anne Clancy

Name	Genotype	Source	Used in figure	Construction	Made by
SGA query *Δget1 Δget2 Δget3*	SGA query *get1Δ::Kan^R get2Δ::Nat^R get3Δ::Ble^R*	This study	None; used to create other strains	The *GET3* locus was replaced with the *Ble^R* cassette amplified from genomic DNA from By4741 *Δget1 Δget2 Δget3* described in Schuldiner et al., 2008.	Ákos Farkas
SGA query AID-Sed5 Tir1	SGA query *ScCYC1t::Kan^R::Sc PGK1pr::ScTEFpr:: AtTIR1::ScADH1t:: SpNOP1pr::*6xmyc-AID-Sed5	This study	None; used to create other strains	Sed5 was tagged with the indicated cassette. The cassette had been cloned beforehand into a bacterial cloning plasmid and amplified from it.	Anne Clancy
alpha *Δget3*	*MATalpha his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 get3Δ::Kan^R*	This study	None; used to create other strains	Made by mating By4741 *Δget3* with By4742, then sporulating the resulting diploid strain and selecting the correct progeny.	Jutta Metz
alpha *Δget1 Δget2 Δget3*	*MATalpha his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 get1Δ::Kan^R get2Δ::Nat^R get3Δ::Ble^R*	This study	None; used to create other strains	In alpha *Δget3,* the *GET1* locus was replaced with the *Kan^R* cassette, the *GET2* locus with the *Nat^R* cassette, the *Get3* locus with the *Ble^R* cassette amplified from genomic DNA from By4741 *Δget1 Δget2 Δget3* described in Schuldiner et al., 2008.	Ákos Farkas

Table 6

Name	Sequence	Used to create strain
Sed5_ko-NAT_AP2096_F	CCTCTGTTTCCAACGACCATTTCACCATACCACACAACTCCCcgcacacaccatagcttcaaaatg	By4741 *Δsed5* p416Met25 mCherry-Sed5
Sed5_ko-NAT_AP2096_R	GAGAAGGAAAAAAAATAAATAATCAAATATGTATCTTTTTTTCTTTTACTTTGCTTgagcgacctcatgctatacctg	By4741 *Δsed5* p416Met25 mCherry-Sed5
Get3_ko-BLE_AP2096_F	AAACGTACGACAAGAACAAGAAGATCATCACATTGTAATTcctggaagtaccttcaaagaatg	SGA query Δget1 Δget2 Δget3, alpha *Δget1 Δget2 Δget3*
Get3_ko-BLE_AP2096_R	TTATATGTCGTATGTATCTATTTATGGTATTCAGGGGCTTcagctggatcttcgagcgtc	SGA query Δget1 Δget2 Δget3, alpha *Δget1 Δget2 Δget3*
Png1_ko-KAN_AP2096_F	AGATATAGAAGACGTAAAACAACCTGAAGAAGGTAACAGGTGGAGTAGCAtaggtctagagatctgtttagcttgc	By4741 *Δsed5 Δpng1* Sed5-op, By4741 Δsed5 Δget3 Δpng1 Sed5-op
Png1_ko-KAN_AP2096_R	ATGTTCTTAGATAAAATTCCTTACACATAATAATCTGTTCTTTTATTTCTattaagggttctcgagagctcg	By4741 *Δsed5 Δpng1* Sed5-op, By4741 Δsed5 Δget3 Δpng1 Sed5-op
Met15_fko-NAT_F	ACGTGAAGCTGTCGATATTGGGGAACTGTGGTGGTTGGCAcacatacgatttaggtgacac	SGA query *Δmet15*
Met15_fko-NAT_R	CACTATTGATTGCTTAAAAGGGCAATCCGACTATATCTGAaatacgactcactatagggag	SGA query *Δmet15*
Get4_ko-NAT_F	AGTAAACATCATAAAGGGACATAAATAATAATAACAAGCTcacatacgatttaggtgacac	SGA query *Δget4*
Get4_ko-NAT_R	CGCAAACATATTTATCTATTCCTTCGCAAATATGCTCTTTaatacgactcactatagggag	SGA query *Δget4*
Get5_ko-NAT_F	ATAAACTAGCGAAGAATAATAACTTTATACAAAATTAATCcacatacgatttaggtgacac	SGA query *Δget5*
Get5_ko-NAT_R	GTGTAAAATAACAAGTATGTACGTACTAACTATACTAATCaatacgactcactatagggag	SGA query *Δget5*
Sgt2_ko-NAT_F	CTGACCAAGTGATATCTTATTAATACAAATCTACTGTACGcacatacgatttaggtgacac	SGA query *Δsgt2*
Sgt2_ko-NAT_R	CTACATAACATGTATTGCATTAAAGGCTTATTTCAGTCCAaatacgactcactatagggag	SGA query *Δsgt2*
Get1_flank_F	TTGCACGTACCAACTACCTC	alpha *Δget1 Δget2 Δget3*
Get1_flank_R	GAGACGGAGGACATTAAGAATACTG	alpha *Δget1 Δget2 Δget3*
Get2_flank_F	GGAGGGAGGGAGAAGTTTGG	alpha *Δget1 Δget2 Δget3*
Get2_flank_R	ACCCATTGTGTCGTTGGAATC	alpha *Δget1 Δget2 Δget3*
get3_f_F	GGATAAAACGGTGGCAACTTCAAACAAGTTGAGGGA	alpha *Δget1 Δget2 Δget3*
get3_flank_R	GTTATATGTCGTATGTATCTATTTATGGTATTCAGGGGCTTC	alpha *Δget1 Δget2 Δget3*

Table 7

Primary antibodies			
Raised against	**Species**	**Dilution**	**Reference**
Myc	mouse	1 : 2 000	Santa Cruz, ID: sc-40
Bmh1/Bmh2	rabbit	1 : 2 000	(Michelsen et al., 2006)
Sed5	rabbit	1 : 5 000	(Schuldiner et al., 2008)
Get3	guinea pig	1 : 5 000	(Metz et al., 2006)

Table 8

Secondary antibodies				
Raised in	**Raised against**	**Dilution**	**Conjugated to**	**Company**
donkey	mouse	1 : 10 000	IRDye 800CW	LI-COR
donkey	rabbit	1 : 10 000	IRDye 680LT	LI-COR
donkey	rabbit	1 : 10 000	IRDye 800CW	LI-COR
goat	mouse IgG1	1 : 10 000	IRDye 680LT	LI-COR
donkey	guinea pig	1 : 10 000	IRDye 800CW	LI-COR
donkey	guinea pig	1 : 10 000	IRDye 680LT	LI-COR

Chapter Four

Discussion

1. An evolutionary perspective on Get3-like chaperones

The history of research on the family of Get3-like chaperones started with the initial discovery of a bacterial ATPase involved in arsenite resistance, ArsA, which was proposed to represent the catalytic subunit of the arsenite pump of the bacterial plasma membrane (Chen et al., 1986). Later its eukaryotic homolog in budding yeast was found to bind the chloride channel Gef1 when copper was present in the cytosol (Metz et al., 2006). In fact, *get3* cells showed a pronounced copper sensitivity (Metz et al., 2006), further supporting the idea that Get3 and its homologs may be involved in heavy metal resistance. Get3's discovery as a cytosolic chaperone delivering TA proteins to the ER membrane (Schuldiner et al., 2008; Stefanovic and Hegde, 2007), however, showed that eukaryotic Get3 is more different from bacterial ArsA than initially expected. When the structure of Get3 was solved (Bozkurt et al., 2009; Mateja et al., 2009; Suloway et al., 2009), it became clear that the most prominent difference between ArsA and eukaryotic Get3 is the presence of a hydrophobic groove used for TA protein binding that is absent in its bacterial homologs. The methionine-rich character of this hydrophobic groove is reminiscent of that found in Srp54 (Mateja et al., 2009), and is generally thought to be important for binding its substrates. In line with this idea, Get3 homologs found in Archaea, which are closer relatives of eukaryotes than bacteria, also contain the methionine-rich hydrophobic groove and have been shown to be capable of binding TA proteins (Sherrill et al., 2011; Suloway et al., 2012). Photosynthetic bacteria were also found to have Get3 homologs resembling eukaryotic Get3, however, the presence of an α-crystallin domain at their C-terminus sets them apart from their eukaryotic homologs (Chartron et al., 2012).

I conducted a bioinformatic analysis of publicly available bacterial proteomes that surprisingly showed that Get3 homologs that possess a hydrophobic cage are found in several bacterial phyla. Furthermore, considering that other chaperones seem to be required for the biogenesis of TA proteins in bacteria (Peschke et al., 2018) and that no correlation was found between the presence or absence of Get3 homologues and the number of predicted TA proteins in a given species, bacterial Get3-like proteins probably

fulfil another role. Considering the presence of an extensive hydrophobic groove, it is tempting to speculate that they act as more general chaperones.

My extensive comparison of eukaryotic and bacterial Get3 homologues also revealed that the methionine content of the hydrophobic groove varies considerably even between closely related species. Get3 homologs in model species, and in general in vertebrates are indeed rich in methionine residues, however, this is not a feature shared by all eukaryotes. Therefore, the unifying feature shared by Get3 homologs in all domains of life is a hydrophobic groove with variable methionine content, suggesting that they all share similar functions as well. Possibly, the original function of the ancestor of eukaryotic Get3 was a more general chaperone. This probably remains its function in bacteria today, whereas it in eukaryotes Get3-like proteins became more and more specialised in assisting TA proteins in their biogenesis. The need for a highly specialised machinery of TA targeting in eukaryotes and especially multicellular organisms is explained by their highly complex membrane system and the fact that they may express hundreds of TA proteins more than bacteria (Kalbfleisch et al., 2007). The ATPase-independent chaperone function of Get3, which was found to be activated *in vitro* upon oxidation (Voth et al., 2014), supports the idea that eukaryotic Get3 may have additional functions besides TA targeting.

2. Direct and indirect effects of loss of Get3

Our automated screen using an N-terminally GFP-tagged library supported the idea that not only TA proteins are affected upon loss of Get3. However, the effects of a loss of a gene are often pleiotropic and such experiments cannot reveal on their own whether Get3 actually interacts with any of these proteins or whether they are simply indirectly affected via a direct effect on a substrate of Get3. A transient knock-down of Sed5, an essential protein and a strongly Get3-dependent TA protein revealed that some of the observed effects can be explained through Sed5's loss from the Golgi due to the absence of Get3. I found three StART-like domain containing proteins affected in *get3* cells, Lam1, Lam5 and Sip3 (Gatta et al., 2015), however, only Lam5 is a TA protein. They all have similar functions in transferring sterol between membranes and they localize to membrane contact sites (Gatta et al., 2015). Therefore, it is impossible to tell based on these results whether they are all indirectly affected, or whether Lam5 is

affected because it is a TA protein and thus affects the others indirectly. It is also possible that they all interact with Get3 directly through its alternative chaperone activity, but to decide which of these possibilities is true, more directed experiments will be necessary.

Another crucial aspect of any such high-content screen is the level and regulation of expression of the tested proteins. For instance, the expression level of TA substrates influences whether any defects are visible in *get3* cells (Li et al., 2019). In the SWAT library used in our screens the moderate *NOP1* promoter drives the expression of each tagged protein and thus – in most cases – provides a modest overexpression compared to using the native promoter of each gene (Weill et al., 2018). It is possible that there are further affected proteins which would only be revealed if a stronger overexpression was employed. Furthermore, since many proteins only tolerate either an N- or a C-terminal fusion to a tag without being functionally perturbed, a similar screen with a C-terminally GFP-tagged library may reveal even further hits (Meurer et al., 2018). Nevertheless, for TA proteins, the N-terminal fusion library that I screened is ideally suited since C-terminal GFP fusion abolishes the TA topology.

3. Novel insights into the pretargeting complex

Besides distinguishing between direct and indirect effects, I also tested whether the hits of my screen were unique to loss of Get3 or hits upon disturbing GET-pathway function in general. Based on the current model of GET pathway, Get3 should only be able to capture and deliver substrates to its receptor when the pretargeting complex composed of Sgt2, Get5, and Get4 is present (Shao et al., 2017). However, just like it has been suggested in the case of Sed5 (Kohl et al., 2011), the tested TA and non-TA proteins all showed a graded effect depending on whether Sgt2, Get4/Get5, or Get3 was knocked out. This strongly suggests that in contrast to the current models based on *in-vitro* evidence, Get3 can probably receive substrates from other chaperones in addition to the pretargeting complex *in vivo* or it can recognize them without the help of other proteins. Furthermore, since the effects of knocking out *SGT2* were clearly smaller than those of losing Get4 or Get5, it seems Get4 and Get5 can recruit other chaperones to hand over TA substrates to Get3. This is in line with the finding that Sgt2 is not required to recruit Get3 to glucose starvation induced foci but Get4 and Get5 are (Powis et al., 2013).

Based on current models of Get3's TA protein targeting cycle, ATP-bound Get3 should interact with the pretargeting complex more strongly than either its ADP-bound or its apo form. Confirming these expectations, immunoprecipited Get3-GFP D57E, which is an ATPase deficient mutant of Get3 (Voth et al., 2014) and thus expected to be in a mostly ATP-bound state, enriched all the components of the pretargeting complex more strongly than Get3 wt. Interestingly, Ybr137w was found to elute specifically with Get3 D57E but not with Get3 wt. Ybr137w is an interactor of Sgt2 (Kohl et al., 2011; Wang et al., 2010) and can form multimeric complexes to function as a chaperone (Yeh et al., 2014). Yet, its in vivo function remains unknown. Although no obvious defect in TA protein biogenesis has been observed in its absence (Kohl et al., 2011) and is not conserved in vertebrates, our results confirm it being part of the yeast pretargeting complex. A possible explanation for the apparent lack of TA-protein related phenotypes is that its functions may only become apparent when cells are exposed to stress or only when substrates are strongly overexpressed. Nevertheless, further experiments are needed to understand the role Ybr137w in the GET pathway in budding yeast.

4. Get3 interacts with different classes of membrane proteins

To get a better indication of what proteins Get3 may interact with directly, we expressed Get3 D57E tagged with a blue fluorescent protein, which forms foci in cells and has been successfully used to trap substrates in the mammalian system (Coy-Vergara et al., 2019; Powis et al., 2013). The automated screen of these strains revealed a wide range of membrane proteins colocalizing with Get3 D57E providing further evidence that the *in-vivo* client range of Get3 extends beyond TA proteins. Based on the topology of the colocalizing proteins, many of them had multiple C-terminal TMSs, others were peripheral membrane proteins, yet some had no obvious topological features besides being integral membrane proteins.

Interestingly, many of the membrane proteins in the latter group are expected to have luminal N-termini (Ost6, Opy2, Pin2), are known to localize differently in cells based on whether they are N- or C-terminally GFP-tagged (Opy2, Pin2), or they produce a faint signal with an N-terminal GFP tag (Yet2, Vba3). We cannot exclude the possibility that this may indicate perturbed biogenesis due to the position of the tag and their trapping by Get3 D57E may be the result of the specific position of the tag. This scenario is supported

by the fact that a similar phenomenon has been observed with mitochondrial proteins mislocalized to the ER due to an N-terminal GFP tag (Vitali et al., 2018). Alternatively, it is tempting to speculate that some of these are *bona-fide* chaperone substrates of Get3 unrelated to the TA targeting function of the GET pathway.

Since *in vitro* Get3 is known to possess an ATPase-independent chaperone activity, we examined the behaviour of a Get3 mutant not expected to be able to bind ATP, Get3 K31A (Yamagata et al., 2010). Our collaborator, Dr. Kathrin Ulrich, found that once activated through oxidation, the chaperone activity of Get3 K31A could not be switched off under reducing conditions unlike it was observed with Get3 wt or D57E. Since this could result in a shift towards a chaperone-active form of Get3 *in vivo*, I tested the range of proteins captured by Get3 K31A as well. I found that Get3 K31A captured all the proteins that Get3 D57E did, although it formed more prominent foci and colocalization was more pronounced with the trapped proteins. We also found some proteins that colocalized specifically with Get3 K31A and not with Get3 D57E, which could either be the result of them being chaperone substrates of Get3, or simply being more weakly interacting substrates that only become visible when strong Get3 foci are formed in the cells. Overexpressing Get3 D57E to induce stronger puncta formation could help determine which of the two scenarios is true. Consistent with the idea that trapping substrates with Get3 is deleterious to the cells, when either Get3 D57E or K31A was overexpressed using the same galactose-inducible promoter, I found that both impaired the growth of cells. However, the effect of Get3 K31A was clearly stronger, which suggests that irrespective of expression levels, Get3 K31A possesses an additional *in-vivo* activity, which could be related to the altered regulation of its chaperone activity. Nonetheless, the negative impact of Get3 D57E or K31A on cell growth was completely abolished when either mutation was combined with further mutations affecting the hydrophobic substrate binding surface. This shows that irrespective of the different functions of Get3, the hydrophobic groove known to bind TA proteins is essential for its interaction with substrates. This is in line with the idea of an evolutionarily conserved hydrophobic substrate-binding interface which is also found in bacterial homologs. Therefore, it is possible that the TA binding and the alternative chaperone activity of Get3 are manifestations of the same basic mode of interaction of Get3 with its substrates but that the difference between TA proteins and other substrates may lie in how these clients gain access to the hydrophobic cage.

5. Novel classes of Get3 clients

A prominent group among the proteins found to colocalize with Get3 D57E was proteins with multiple TMSs close to the C-terminus. These proteins may be subject to the same constraints during their biogenesis as TA proteins. Therefore, it is possible that Get3 plays a role in their biogenesis as well. Our model substrate, Vph2 was indeed enriched when I immunoprecipitated Get3 D57E compared to Get3 wt, as was shown for TA substrates in mammalian cells (Coy-Vergara et al., 2019). Furthermore, Vph2 colocalized with the pretargeting complex marked by mTagBFP2-Get5, which is also a feature of TA substrates (Jonikas et al., 2009). Taken together, these results indicate that Get3 is not selective to single TMSs as found in TA proteins and acts as a more general targeting factor for proteins with C-terminal TMSs. This further underscores the interplay between the SND proteins and the GET pathway as it has been shown that the same TMS can recruit different targeting pathways depending on its relative position in the tested construct (Aviram et al., 2016).

A further interesting group of colocalizing proteins with Get3 D57E consisted of peripheral membrane proteins. Since such proteins utilize hydrophobic patches to interact with membranes, I wondered whether Get3 could interact with these proteins as well. I found that one of these proteins, Atf2, which uses two amphipathic helices to contact the membrane, co-purified strongly with Get3 wt but was not further enriched in Get3 D57E. Since this interaction was abolished with Get3 FIDD, it likely occurs through the same hydrophobic groove that TA proteins use. Although not found in our colocalization screen, its paralog, Atf1 behaved similarly to Atf2 biochemically. This emphasizes that the lack of observed colocalization does not necessarily indicate an absence of interaction and there may be many more proteins interacting with Get3 that we could not uncover with the help of Get3 D57E. Furthermore, both the N- and the C-terminal amphipathic helix of Atf2 were necessary for its interaction with Get3, suggesting that Get3 recognizes a sequence in Atf2 in a context-dependent manner, i.e. only once Atf2 has reached the membrane. This opens up the possibility that Get3 acts a chaperone for proteins after being targeted to the membrane as opposed to TA proteins, which are recognized by Get3 before their membrane insertion.

Furthermore, three proteins involved in ergosterol biosynthesis, Erg1, Erg9, Erg27 were found colocalizing with Get3 D57E, of which only Erg9 is a TA protein. In case the

other two interact with Get3 outside of the scope of the GET pathway, this would suggest that ergosterol synthesis is affected by both Get3's TA targeting function and its alternative chaperone function. The lipidomic analysis of *get3* cells provided further support for the idea that ergosterol synthesis is affected in *get3* cells. Therefore, Get3 could play an important role in the lipid homeostasis of cells via its TA protein targeting and chaperone activity required for proteins involved in lipid metabolism.

6. Get3 dependence of TA proteins is determined by the hydrophobicity of the TMS and other properties of the N-terminal cytosolic domain

The most widely studied class of Get3 substrates are TA proteins and the Get3 dependence of several such proteins has been characterized *in vitro* as well as *in vivo* as mentioned before. My results are in line with previous findings for many TA proteins, however, some of those found to depend on Get3 by others (Li et al., 2019; Schuldiner et al., 2008) did not react to the loss of Get3 of colocalized with Get3 D57E in our screens. A likely explanation for this discrepancy lies in the expression levels employed in the different studies, because Get3 dependence was mostly observed when TA substrates were strongly overexpressed (Li et al., 2019; Schuldiner et al., 2008). This highlights that there are only few TA proteins that are obligate Get3 substrates at low expression levels. Importantly, only few were found to mislocalize to mitochondria even when the gene responsible for the clearance of mislocalized TA proteins from mitochondria, *MSP1*, was knocked out in combination with *GET3* (Li et al., 2019). Furthermore, I found that the hydrophobicity of the TMS of TA proteins as measured by the TM tendency scale provides a good prediction for those that are captured by Get3 D57E, which confirms findings from mammalian systems (Chitwood et al., 2018; Coy-Vergara et al., 2019). Yet there was no correlation between TMS hydrophobicity and whether a TA protein mislocalized in the absence of Get3. This shows that other factors besides TMS hydrophobicity can have a major influence on Get3 dependence.

There were only a select few TA proteins that both mislocalized in *get3* cells and were trapped by Get3 D57E, among which Sed5 is the best characterized evolutionarily conserved substrate also affected in mammalian cells (Rivera-Monroy et al., 2016). To determine what role the N-terminal domain of Sed5 plays in its Get3 dependence, I created chimeras by swapping the TMSs of Sed5 and Gos1, a TA protein unaffected in

my screens. To our surprise, the factor determining colocalization with Get3 D57E and dependence on Get3 for targeting turned out to be the N-terminal cytosolic portion and not the TMS. Even when the lipidation sequence of Ykt6 was used in place of the TMS, Sed5-ykt6 still colocalized with Get3 D57E under glucose starvation but Gos1-ykt6 did not. This suggests there is a unique property inherent in the N-terminal domain of Sed5 that drives its Get3 dependence.

To determine whether Sed5 interacts with Get3 after a primary round of targeting to the membrane, I tagged Sed5 with a short sequence derived from mammalian opsin that allows glycosylation in the ER lumen (Pedrazzini et al., 2000). Although only minimal glycosylation was observed in the logarithmic growth phase probably due to the Golgi localization of Sed5, glycosylation became prominent in the stationary phase. Therefore, we conducted all further experiments in the stationary phase. As expected, we found that the glycosylation of Sed5 was largely dependent on Get3, in accordance with Get3 targeting Sed5 to the ER as a TA protein. However, a knock-out of *PNG1*, which encodes a cytosolic N-glycanase capable of removing glycosylation from proteins extracted from the membrane (Suzuki et al., 2000), increased Sed5 glycosylation in both wt and *get3* cells. The fact that the effect of losing Png1 in *get3* cells was proportionally higher than that in wt cells, indicates that Get3 may play a role in the life cycle of Sed5 after its initial insertion into the membrane. This can help explain why Sed5 and its mammalian counterpart syntaxin 5 are among the few TA proteins that are uniquely sensitive to the loss of Get3 and found to be Get3 dependent across different species.

Immunoprecipitating Get3-GFP showed that Get3 interacts with a glycosylated pool of Sed5, which is reduced when Get3-GFP D57E is pulled down instead. This indicates that Get3 interacts with Sed5 as a TA protein as evidenced by the unglycosylated Sed5 co-eluting with Get3 D57E, but it also interacts with Sed5 after its membrane insertion demonstrated by the glycosylated Sed5 co-eluting with Get3 wt. Importantly, the latter interaction was found to depend on an intact hydrophobic groove of Get3, as the Get3 FIDD largely abolished this interaction. Although the fate of the glycosylated Sed5 pool interacting with Get3 remains unknown, it could be potentially chaperoned by Get3 and stored for later degradation or possibly even reinsertion. This would provide an opportunity for the cells to store proteins when their function is not required for later use instead of degrading them directly.

Chapter Five

Summary and outlook

Get3's function in TA protein targeting has long been assumed to be the driving factor behind the diverse phenotypes observed in *get3* cells. Here, I demonstrate that only a portion of TA proteins are acutely affected in the absence of Get3, although a major part of all TA proteins can be captured by Get3 when a trapping mutant is expressed. We find a diverse range of membrane proteins colocalizing with a trapping mutant, which represents a hitherto unknown pool of potential substrates. We demonstrate that manipulating the ATP-binding capacity of Get3 can affect the properties of its chaperone function *in vitro*, which creates a strong substrate-trapping mutant *in vivo* and results in a severe growth defect when expressed in cells.

We characterize novel classes of Get3 substrates. This includes proteins with multiple TMSs located close to the C-terminus, which likely represents clients of the whole GET pathway. Furthermore, some peripheral membrane proteins, as demonstrated here for Atf2 and Atf1, can interact with Get3 through its hydrophobic binding groove, but unlike TA proteins are not enriched by Get3 D57E. The fact that their membrane localization is necessary for their interaction with Get3 suggests that they are chaperoned by Get3 after membrane targeting. Other peripheral membrane proteins trapped by Get3 D57E are involved in ergosterol synthesis, and a lipidomic analysis confirmed that ergosterol metabolism is affected in *get3* cells.

The screening of TA proteins revealed that TMS hydrophobicity is an important factor in determining capture by Get3 D57E, however, it does not correlate with whether a TA protein mislocalized in the absence of Get3. Of the few TA proteins that both mislocalize in *get3* cells and are captured by Get3 D57E, we dissect the role of the N-terminal domain of Sed5 to show that Get3 can interact with TA proteins both before and after membrane targeting.

The evolutionary comparison of Get3 homologs across bacteria, archaea and eukaryotes revealed that Get3 homologs with a hydrophobic cage are found across all domains of life and thus represent an ancient class of proteins. Since they are unlikely to be involved in TA targeting in bacteria, my analysis raises the possibility that Get3-like proteins act as more general chaperones that were adapted in eukaryotes for TA targeting while preserving their original chaperone function.

Compared to the TA protein clients of Get3, we know little about additional chaperone substrates *in vivo* which probably includes some of the proteins I found in this study. Further experiments are needed to clarify the specific interaction taking place between such chaperone clients and Get3 and the physiological relevance thereof. The exact nature of the interaction between Sed5 and Get3 also warrants further studies due its evolutionary conservation and the essential function of Sed5. Taken together, this study provides the first insights into the *in-vivo* physiological relevance of Get3's chaperone function outside the scope of the GET pathway and draws attention to the multiple roles Get3 plays in the context of chaperoning proteins and their biogenesis.

References

For references for the manuscript "The natural history of Get3-like chaperones", please see Chapter Two.

Akopian, D., Shen, K., Zhang, X., and Shan, S. (2013). Signal Recognition Particle: An Essential Protein-Targeting Machine. Annual Review of Biochemistry *82*, 693–721.

Allen, K.N., Entova, S., Ray, L.C., and Imperiali, B. (2019). Monotopic Membrane Proteins Join the Fold. Trends in Biochemical Sciences *44*, 7–20.

Anghel, S.A., McGilvray, P.T., Hegde, R.S., and Keenan, R.J. (2017). Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum. Cell Reports *21*, 3708–3716.

Ast, T., Cohen, G., and Schuldiner, M. (2013). A Network of Cytosolic Factors Targets SRP-Independent Proteins to the Endoplasmic Reticulum. Cell *152*, 1134–1145.

Auld, K.L., Hitchcock, A.L., Doherty, H.K., Frietze, S., Huang, L.S., and Silver, P.A. (2006). The Conserved ATPase Get3/Arr4 Modulates the Activity of Membrane-Associated Proteins in Saccharomyces cerevisiae. Genetics *174*, 215–227.

Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., Haßdenteufel, S., Dudek, J., Jung, M., Schorr, S., et al. (2016). The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. Nature *540*, 134–138.

Bange, G., and Sinning, I. (2013). SIMIBI twins in protein targeting and localization. Nature Structural & Molecular Biology *20*, 776–780.

Battle, A., Jonikas, M.C., Walter, P., Weissman, J.S., and Koller, D. (2010). Automated identification of pathways from quantitative genetic interaction data. Molecular Systems Biology *6*.

Beilharz, T. (2003). Bipartite Signals Mediate Subcellular Targeting of Tail-anchored Membrane Proteins in Saccharomyces cerevisiae. Journal of Biological Chemistry *278*, 8219–8223.

Bhattacharya, S., Esquivel, B.D., and White, T.C. (2018). Overexpression or Deletion of Ergosterol Biosynthesis Genes Alters Doubling Time, Response to Stress Agents, and Drug Susceptibility in *Saccharomyces cerevisiae*. MBio *9*.

Borgese, N. (2003). The tale of tail-anchored proteins: coming from the cytosol and looking for a membrane. The Journal of Cell Biology *161*, 1013–1019.

Bozkurt, G., Stjepanovic, G., Vilardi, F., Amlacher, S., Wild, K., Bange, G., Favaloro, V., Rippe, K., Hurt, E., Dobberstein, B., et al. (2009). Structural insights into tail-anchored

protein binding and membrane insertion by Get3. Proceedings of the National Academy of Sciences *106*, 21131–21136.

Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115–132.

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

Caplin, B.E., Hettich, L.A., and Marshall, M.S. (1994). Substrate characterization of the Saccharomyces cerevisiae protein farnesyltransferase and type-I protein geranylgeranyltransferase. Biochimica et Biophysica Acta *1205*, 39–48.

Chang, Y.-W., Chuang, Y.-C., Ho, Y.-C., Cheng, M.-Y., Sun, Y.-J., Hsiao, C.-D., and Wang, C. (2010). Crystal Structure of Get4-Get5 Complex and Its Interactions with Sgt2, Get3, and Ydj1. Journal of Biological Chemistry *285*, 9962–9970.

Chartron, J.W., Suloway, C.J.M., Zaslaver, M., and Clemons, W.M. (2010). Structural characterization of the Get4/Get5 complex and its interaction with Get3. Proceedings of the National Academy of Sciences *107*, 12127–12132.

Chartron, J.W., Clemons, W.M., and Suloway, C.J.M. (2012). The complex process of GETting tail-anchored membrane proteins to the ER. Current Opinion in Structural Biology *22*, 217–224.

Chen, C.-M., Misra, T.K., Silver, S., and Rosen, B.P. (1986). Nucleotide sequence of the structural genes for an anion pump. Journal of Biological Chemistry *261*, 15030–15038.

Chitwood, P.J., Juszkiewicz, S., Guna, A., Shao, S., and Hegde, R.S. (2018). EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. Cell *175*, 1507-1519.e16.

Cho, H., and Shan, S. (2018). Substrate relay in an Hsp70-cochaperone cascade safeguards tail-anchored membrane protein targeting. The EMBO Journal e99264.

Cohen, Y., and Schuldiner, M. (2011). Advanced Methods for High-Throughput Microscopy Screening of Genetically Modified Yeast Libraries. In Network Biology, G. Cagney, and A. Emili, eds. (Totowa, NJ: Humana Press), pp. 127–159.

Costa, E.A., Subramanian, K., Nunnari, J., and Weissman, J.S. (2018). Defining the physiological role of SRP in protein-targeting efficiency and specificity. Science *359*, 689–692.

Cotter, K., Stransky, L., McGuire, C., and Forgac, M. (2015). Recent Insights into the Structure, Regulation, and Function of the V-ATPases. Trends in Biochemical Sciences *40*, 611–622.

Coy-Vergara, J., Rivera-Monroy, J., Urlaub, H., Lenz, C., and Schwappach, B. (2019). A trap mutant reveals the physiological client spectrum of TRC40. Journal of Cell Science *132*, jcs230094.

Ejsing, C.S., Sampaio, J.L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R.W., Simons, K., and Shevchenko, A. (2009). Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. Proceedings of the National Academy of Sciences *106*, 2136–2141.

Farkas, Á., De Laurentiis, E.I., and Schwappach, B. (2019). The natural history of Get3-like chaperones. Traffic *20*, 311–324.

Gatta, A.T., Wong, L.H., Sere, Y.Y., Calderón-Noreña, D.M., Cockcroft, S., Menon, A.K., and Levine, T.P. (2015). A new family of StART domain proteins at membrane contact sites has a role in ER-PM sterol transport. ELife *4*.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature Methods *6*, 343–345.

Gietz, D., Jean, A.S., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Research *20*, 1425–1425.

Giménez-Andrés, M., Čopič, A., and Antonny, B. (2018). The Many Faces of Amphipathic Helices. Biomolecules *8*, 45.

Goldstein, A.L., and McCusker, J.H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast *15*, 1541–1553.

Graumann, J., Lilie, H., Tang, X., Tucker, K.A., Hoffmann, J.H., Vijayalakshmi, J., Saper, M., Bardwell, J.C.A., and Jakob, U. (2001). Activation of the Redox-Regulated Molecular Chaperone Hsp33—A Two-Step Mechanism. Structure *9*, 377–387.

Gristick, H.B., Rao, M., Chartron, J.W., Rome, M.E., Shan, S., and Clemons, W.M. (2014). Crystal structure of ATP-bound Get3–Get4–Get5 complex reveals regulation of Get3 by Get4. Nature Structural & Molecular Biology *21*, 437–442.

Guna, A., Volkmar, N., Christianson, J.C., and Hegde, R.S. (2018). The ER membrane protein complex is a transmembrane domain insertase. Science *359*, 470–473.

Harbauer, A.B., Zahedi, R.P., Sickmann, A., Pfanner, N., and Meisinger, C. (2014). The Protein Import Machinery of Mitochondria—A Regulatory Hub in Metabolism, Stress, and Disease. Cell Metabolism *19*, 357–372.

Hardwick, K.G., and Pelham, H.R.B. (1992). SED5 encodes a 39-kD integral membrane protein required for vesicular transport between the ER and the Golgi complex. The Journal of Cell Biology *119*, 513–521.

Haßdenteufel, S., Sicking, M., Schorr, S., Aviram, N., Fecher-Trost, C., Schuldiner, M., Jung, M., Zimmermann, R., and Lang, S. (2017). hSnd2 protein represents an alternative targeting factor to the endoplasmic reticulum in human cells. FEBS Letters *591*, 3211–3224.

Herzog, R., Schwudke, D., Schuhmann, K., Sampaio, J.L., Bornstein, S.R., Schroeder, M., and Shevchenko, A. (2011). A novel informatics concept for high-throughput shotgun lipidomics based on the molecular fragmentation query language. Genome Biology *12*, R8.

Herzog, R., Schuhmann, K., Schwudke, D., Sampaio, J.L., Bornstein, S.R., Schroeder, M., and Shevchenko, A. (2012). LipidXplorer: A Software for Consensual Cross-Platform Lipidomics. PLoS ONE *7*, e29851.

Hsiao, Y.-S., Jogl, G., Esser, V., and Tong, L. (2006). Crystal structure of rat carnitine palmitoyltransferase II (CPT-II). Biochemical and Biophysical Research Communications *346*, 974–980.

Jackson, D.D., and Stevens, T.H. (1997). VMA12 Encodes a Yeast Endoplasmic Reticulum Protein Required for Vacuolar H+-ATPase Assembly. Journal of Biological Chemistry *272*, 25928–25934.

Johnson, N., Vilardi, F., Lang, S., Leznicki, P., Zimmermann, R., and High, S. (2012). TRC40 can deliver short secretory proteins to the Sec61 translocon. Journal of Cell Science *125*, 4414–4414.

Jonikas, M.C., Collins, S.R., Denic, V., Oh, E., Quan, E.M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J.S., et al. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science *323*, 1693–1697.

Kalbfleisch, T., Cambon, A., and Wattenberg, B.W. (2007). A Bioinformatics Approach to Identifying Tail-Anchored Proteins in the Human Genome. Traffic *8*, 1687–1694.

Kiianitsa, K., Solinger, J.A., and Heyer, W.-D. (2003). NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. Analytical Biochemistry *321*, 266–271.

Klose, C., Surma, M.A., Gerl, M.J., Meyenhofer, F., Shevchenko, A., and Simons, K. (2012). Flexibility of a Eukaryotic Lipidome – Insights from Yeast Lipidomics. PLoS ONE *7*, e35063.

Kodaki, T., and Yamashita, S. (1987). Yeast Phosphatidylethanolamine Methylation Pathway. Journal of Biological Chemistry *262*, 15428–15435.

Kohl, C., Tessarz, P., von der Malsburg, K., Zahn, R., Bukau, B., and Mogk, A. (2011). Cooperative and independent activities of Sgt2 and Get5 in the targeting of tail-anchored proteins. Biological Chemistry *392*, 601–608.

Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nature Methods *11*, 319–324.

Kushnirov, V.V. (2000). Rapid and reliable protein extraction from yeast. Yeast *16*, 857–860.

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

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology *157*, 105–132.

Lee, J., Kim, D.H., and Hwang, I. (2014). Specific targeting of proteins to outer envelope membranes of endosymbiotic organelles, chloroplasts, and mitochondria. Frontiers in Plant Science *5*.

Levine, T.P., Wiggins, C.A.R., and Munro, S. (2000). Inositol Phosphorylceramide Synthase Is Located in the Golgi Apparatus of Saccharomyces cerevisiae. Molecular Biology of the Cell *11*, 2267–2281.

Li, L., Zheng, J., Wu, X., and Jiang, H. (2019). Mitochondrial AAA-ATPase Msp1 detects mislocalized tail-anchored proteins through a dual-recognition mechanism. EMBO Reports *20*, e46989.

Lin, J.-L., and Wheeldon, I. (2014). Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae. PLoS ONE *9*, e104141.

Mateja, A., and Keenan, R.J. (2018). A structural perspective on tail-anchored protein biogenesis by the GET pathway. Current Opinion in Structural Biology *51*, 195–202.

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

Mayerhofer, P.U. (2016). Targeting and insertion of peroxisomal membrane proteins: ER trafficking versus direct delivery to peroxisomes. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1863*, 870–880.

McNew, J.A., Søgaard, M., Lampen, N.M., Machida, S., Ye, R.R., Lacomis, L., Tempst, P., Rothman, J.E., and Söllner, T.H. (1997). Ykt6p, a Prenylated SNARE Essential for Endoplasmic Reticulum-Golgi Transport. Journal of Biological Chemistry *272*, 17776–17783.

Metz, J., Wachter, A., Schmidt, B., Bujnicki, J.M., and Schwappach, B. (2006). The Yeast Arr4p ATPase Binds the Chloride Transporter Gef1p When Copper Is Available in the Cytosol. Journal of Biological Chemistry *281*, 410–417.

Meurer, M., Duan, Y., Sass, E., Kats, I., Herbst, K., Buchmuller, B.C., Dederer, V., Huber, F., Kirrmaier, D., Štefl, M., et al. (2018). Genome-wide C-SWAT library for high-throughput yeast genome tagging. Nature Methods *15*, 598–600.

Michelsen, K., Mrowiec, T., Duderstadt, K.E., Frey, S., Minor, D.L., Mayer, M.P., and Schwappach, B. (2006). A Multimeric Membrane Protein Reveals 14-3-3 Isoform Specificity in Forward Transport in Yeast: Arg-Based Sorting Motifs in Yeast. Traffic *7*, 903–916.

Moravcevic, K., Oxley, C.L., and Lemmon, M.A. (2012). Conditional Peripheral Membrane Proteins: Facing up to Limited Specificity. Structure *20*, 15–27.

Mumberg, D., Müller, R., and Funk, M. (1994). Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. Nucleic Acids Research *22*, 5767–5768.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., and Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. Nature Methods *6*, 917–922.

Orlean, P., and Menon, A.K. (2007). Thematic review series: Lipid Posttranslational Modifications. GPI anchoring of protein in yeast and mammalian cells, or: how we learned to stop worrying and love glycophospholipids. Journal of Lipid Research *48*, 993–1011.

Padyana, A.K., Gross, S., Jin, L., Cianchetta, G., Narayanaswamy, R., Wang, F., Wang, R., Fang, C., Lv, X., Biller, S.A., et al. (2019). Structure and inhibition mechanism of the catalytic domain of human squalene epoxidase. Nature Communications *10*.

Pedrazzini, E., Villa, A., Longhi, R., Bulbarelli, A., and Borgese, N. (2000). Mechanism of Residence of Cytochrome B(5), a Tail-Anchored Protein, in the Endoplasmic Reticulum. The Journal of Cell Biology *148*, 899–914.

Peschke, M., Le Goff, M., Koningstein, G.M., Karyolaimos, A., de Gier, J.-W., van Ulsen, P., and Luirink, J. (2018). SRP, FtsY, DnaK and YidC Are Required for the Biogenesis of the E. coli Tail-Anchored Membrane Proteins DjlC and Flk. Journal of Molecular Biology *430*, 389–403.

Powis, K., Schrul, B., Tienson, H., Gostimskaya, I., Breker, M., High, S., Schuldiner, M., Jakob, U., and Schwappach, B. (2013). Get3 is a holdase chaperone and moves to deposition sites for aggregated proteins when membrane targeting is blocked. Journal of Cell Science *126*, 473–483.

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

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Séraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature Biotechnology *17*, 1030–1032.

Ripaud, L., Chumakova, V., Antonin, M., Hastie, A.R., Pinkert, S., Körner, R., Ruff, K.M., Pappu, R.V., Hornburg, D., Mann, M., et al. (2014). Overexpression of Q-rich prion-like proteins suppresses polyQ cytotoxicity and alters the polyQ interactome. Proceedings of the National Academy of Sciences *111*, 18219–18224.

Rivera-Monroy, J., Musiol, L., Unthan-Fechner, K., Farkas, Á., Clancy, A., Coy-Vergara, J., Weill, U., Gockel, S., Lin, S.-Y., Corey, D.P., et al. (2016). Mice lacking WRB reveal differential biogenesis requirements of tail-anchored proteins in vivo. Scientific Reports *6*, 39464.

Rome, M.E., Rao, M., Clemons, W.M., and Shan, S. (2013). Precise timing of ATPase activation drives targeting of tail-anchored proteins. Proceedings of the National Academy of Sciences *110*, 7666–7671.

Rome, M.E., Chio, U.S., Gristick, H., and Shan, S. (2014). Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway. Proceedings of the National Academy of Sciences *111*, E4929–E4935.

Ryder, N.S. (1992). Terbinafine: Mode of action and properties of the squalene epoxidase inhibition. British Journal of Dermatology *126*, 2–7.

Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop - a common motif in ATP- and GFP-binding proteins. Trends in Biochemical Sciences *15*, 430–434.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature Methods *9*, 671–675.

Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., and Weissman, J.S. (2008). The GET Complex Mediates Insertion of Tail-Anchored Proteins into the ER Membrane. Cell *134*, 634–645.

Shao, S., Rodrigo-Brenni, M.C., Kivlen, M.H., and Hegde, R.S. (2017). Mechanistic basis for a molecular triage reaction. Science *355*, 298–302.

Sheff, M.A., and Thorn, K.S. (2004). Optimized cassettes for fluorescent protein tagging inSaccharomyces cerevisiae. Yeast *21*, 661–670.

Sherrill, J., Mariappan, M., Dominik, P., Hegde, R.S., and Keenan, R.J. (2011). A Conserved Archaeal Pathway for Tail-Anchored Membrane Protein Insertion. Traffic *12*, 1119–1123.

Shurtleff, M.J., Itzhak, D.N., Hussmann, J.A., Schirle Oakdale, N.T., Costa, E.A., Jonikas, M., Weibezahn, J., Popova, K.D., Jan, C.H., Sinitcyn, P., et al. (2018). The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. ELife *7*.

Stefanovic, S., and Hegde, R.S. (2007). Identification of a Targeting Factor for Posttranslational Membrane Protein Insertion into the ER. Cell *128*, 1147–1159.

Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y.-Y., Schwarz, D., Bomke, J., Hein, C., Lohr, F., Bernhard, F., et al. (2011). Structural Basis for Tail-Anchored Membrane Protein Biogenesis by the Get3-Receptor Complex. Science *333*, 758–762.

Subach, O.M., Cranfill, P.J., Davidson, M.W., and Verkhusha, V.V. (2011). An Enhanced Monomeric Blue Fluorescent Protein with the High Chemical Stability of the Chromophore. PLoS ONE *6*, e28674.

Suloway, C.J.M., Chartron, J.W., Zaslaver, M., and Clemons, W.M. (2009). Model for eukaryotic tail-anchored protein binding based on the structure of Get3. Proceedings of the National Academy of Sciences *106*, 14849–14854.

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

Surma, M.A., Herzog, R., Vasilj, A., Klose, C., Christinat, N., Morin-Rivron, D., Simons, K., Masoodi, M., and Sampaio, J.L. (2015). An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids: Automated shotgun lipidomics for blood plasma analysis. European Journal of Lipid Science and Technology *117*, 1540–1549.

Suzuki, T., Park, H., Hollingsworth, N.M., Sternglanz, R., and Lennarz, W.J. (2000). PNG1 , a Yeast Gene Encoding a Highly Conserved Peptide: N-Glycanase. The Journal of Cell Biology *149*, 1039–1052.

Tong, A.H.Y., and Boone, C. (2006). Synthetic genetic array analysis in Saccharomyces cerevisiae. Methods in Molecular Biology *313*, 171–192.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods *13*, 731–740.

Vitali, D.G., Sinzel, M., Bulthuis, E.P., Kolb, A., Zabel, S., Mehlhorn, D.G., Figueiredo Costa, B., Farkas, Á., Clancy, A., Schuldiner, M., et al. (2018). The GET pathway can increase the risk of mitochondrial outer membrane proteins to be mistargeted to the ER. Journal of Cell Science *131*, jcs211110.

Voth, W., Schick, M., Gates, S., Li, S., Vilardi, F., Gostimskaya, I., Southworth, D.R., Schwappach, B., and Jakob, U. (2014). The Protein Targeting Factor Get3 Functions as ATP-Independent Chaperone under Oxidative Stress Conditions. Molecular Cell *56*, 116–127.

Wang, M., and Casey, P.J. (2016). Protein prenylation: unique fats make their mark on biology. Nature Reviews Molecular Cell Biology *17*, 110–122.

Wang, F., Brown, E.C., Mak, G., Zhuang, J., and Denic, V. (2010). A Chaperone Cascade Sorts Proteins for Posttranslational Membrane Insertion into the Endoplasmic Reticulum. Molecular Cell *40*, 159–171.

Wang, F., Whynot, A., Tung, M., and Denic, V. (2011). The Mechanism of Tail-Anchored Protein Insertion into the ER Membrane. Molecular Cell *43*, 738–750.

Wang, F., Chan, C., Weir, N.R., and Denic, V. (2014). The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. Nature *512*, 441–444.

Weill, U., Yofe, I., Sass, E., Stynen, B., Davidi, D., Natarajan, J., Ben-Menachem, R., Avihou, Z., Goldman, O., Harpaz, N., et al. (2018). Genome-wide SWAp-Tag yeast libraries for proteome exploration. Nature Methods *15*, 617–622.

Weill, U., Cohen, N., Fadel, A., Ben-Dor, S., and Schuldiner, M. (2019). Protein Topology Prediction Algorithms Systematically Investigated in the Yeast *Saccharomyces cerevisiae*. BioEssays *41*, 1800252.

Werner-Washburne, M., and Singer, R.A. (1993). Stationary Phase in the Yeast Saccharomyces cerevisiae. *57*, 19.

Wooding, S., and Pelham, H.R.B. (1998). The Dynamics of Golgi Protein Traffic Visualized in Living Yeast Cells. Molecular Biology of the Cell *9*, 2667–2680.

Yamagata, A., Mimura, H., Sato, Y., Yamashita, M., Yoshikawa, A., and Fukai, S. (2010). Structural insight into the membrane insertion of tail-anchored proteins by Get3. Genes to Cells *15*, 29–41.

Yeh, Y.-H., Lin, T.-W., Li, Y.-C., Tung, J.-Y., Lin, C.-Y., and Hsiao, C.-D. (2014). Structural and Functional Characterization of Ybr137wp Implicates Its Involvement in the Targeting of Tail-Anchored Proteins to Membranes. Molecular and Cellular Biology *34*, 4500–4512.

Yofe, I., Weill, U., Meurer, M., Chuartzman, S., Zalckvar, E., Goldman, O., Ben-Dor, S., Schütze, C., Wiedemann, N., Knop, M., et al. (2016). One library to make them all: streamlining the creation of yeast libraries via a SWAp-Tag strategy. Nature Methods *13*, 371–378.

Zhao, G., and London, E. (2006). An amino acid "transmembrane tendency" scale that approaches the theoretical limit to accuracy for prediction of transmembrane helices: Relationship to biological hydrophobicity. Protein Science *15*, 1987–2001.

Acknowledgements

Curriculum Vitae

Personal information

Name and surname	Ákos Farkas
Date of birth	22.12.1988
Place of birth	Budapest, Hungary

Educational background

2014 – current	**PhD student at Georg-August-Universität Göttingen**
	– International Max Planck Research School
	Molecular Biology MSc/PhD
	Master's final examinations grade: excellent (1.0)
	PhD project: Identifying the in-vivo substrate spectrum of the GET pathway.
2011 - 2014	**Eötvös Loránd University, Budapest, Hungary**
	B.Sc. in Biology
	Degree classification: with honours
2007 – 2011	**Corvinus University of Budapest, Hungary**
	BA in International Business
	Overall classification of qualification: excellent (5)
	Degree awarded in 2012

Scholarships and prizes

09/2014 – 08/2015	**Stipend by the International Max Planck Research School**
09/2009 – 06/2010	**Scholarship of the Hungarian Republic**
	At Corvinus University of Budapest
	Awarded for excellent academic results, and taking part in extracurricular social, scientific activities
12/2012	**First place in animal anatomy competition**
	At Eötvös Loránd University, Budapest
	Local competition covering invertebrate and vertebrate anatomy

List of publications

Farkas, Á., De Laurentiis, E.I., and Schwappach, B. (2019). The natural history of Get3-like chaperones. Traffic 20, 311–324.

Rivera-Monroy, J., Musiol, L., Unthan-Fechner, K., **Farkas, Á.,** Clancy, A., Coy-Vergara, J., Weill, U., Gockel, S., Lin, S.-Y., Corey, D.P., et al. (2016). Mice lacking WRB reveal differential biogenesis requirements of tail-anchored proteins in vivo. Scientific Reports *6*, 39464.

Vitali, D.G., Sinzel, M., Bulthuis, E.P., Kolb, A., Zabel, S., Mehlhorn, D.G., Figueiredo Costa, B., **Farkas, Á.,** Clancy, A., Schuldiner, M., et al. (2018). The GET pathway can increase the risk of mitochondrial outer membrane proteins to be mistargeted to the ER. Journal of Cell Science 131, jcs211110.