The EMBO Journal (2012) 31, 707–719 | © 2012 European Molecular Biology Organization | All Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

Tail-anchor targeting by a Get3 tetramer: the structure of an archaeal homologue

Christian JM Suloway, Michael E Rome and William M Clemons Jr*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA

Efficient delivery of membrane proteins is a critical cellular process. The recently elucidated GET (*G*uided *E*ntry of *T*A proteins) pathway is responsible for the targeted delivery of tail-anchored (TA) membrane proteins to the endoplasmic reticulum. The central player is the ATPase Get3, which in its free form exists as a dimer. Biochemical evidence suggests a role for a tetramer of Get3. Here, we present the first crystal structure of an archaeal Get3 homologue that exists as a tetramer and is capable of TA protein binding. The tetramer generates a hydrophobic chamber that we propose binds the TA protein. We use small-angle X-ray scattering to provide the first structural information of a fungal Get3/TA protein complex showing that the overall molecular envelope is consistent with the archaeal tetramer structure. Moreover, we show that this fungal tetramer complex is capable of TA insertion. This allows us to suggest a model where a tetramer of Get3 sequesters a TA protein during targeting to the membrane.
The EMBO Journal (2012) **31,** 707–719. doi:10.1038/emboj.2011.433; Published online 29 November 2011
Subject Categories: membranes & transport; structural biology
Keywords: Arr4p; bioSAXS; GET pathway; protein transport; TRC40

Introduction

Targeted delivery of membrane proteins is a highly regulated process. The ubiquitous co-translational signal recognition particle (SRP) pathway is responsible for the delivery of the majority of membrane proteins (Shan and Walter, 2005). Tail-anchored (TA) membrane proteins are exceptions. They are defined topologically as having a single transmembrane domain (TM) near their C-terminus (Borgese *et al*, 2003). Found in cytoplasmically associated membranes of all organisms, they account for 2–3% of open reading frames in humans and nearly 1% in yeast and prokaryotes (Beilharz *et al*, 2003; Kalbfleisch *et al*, 2007; Kriechbaumer *et al*, 2009; Pedrazzini, 2009; Borgese and Righi, 2010). The signal for TA protein membrane delivery, the C-terminal TM, is not accessible for targeting by the SRP pathway (Kutay *et al*, 1993); therefore, they must be delivered via a different route.

In eukaryotes, pathways for TA protein delivery to the ER have recently been elucidated (reviewed in Rabu *et al*, 2009 and Simpson *et al*, 2010). The majority of TA proteins are targeted via the GET (*G*uided *E*ntry of *T*A proteins) pathway. This pathway, here described for yeast, progresses from a Sgt2/Get4/Get5 sorting complex (Bag6 complex in mammals; Mariappan *et al*, 2010) that delivers an appropriate TA protein to the ATPase Get3 (TRC40 in mammals) (Battle *et al*, 2010; Chartron *et al*, 2010; Wang *et al*, 2010), which then targets the protein to the ER membrane via Get1/Get2 (Schuldiner *et al*, 2008). The central protein Get3 was the first component discovered that directly participates in TA targeting (Stefanovic and Hegde, 2007). Get3 binds specifically to the TM of a TA substrate and is essential for efficient delivery to the ER (Stefanovic and Hegde, 2007; Favaloro *et al*, 2008; Schuldiner *et al*, 2008).

Based on the structural studies, a functional model for Get3 has been proposed where nucleotide state modulates an 'open' versus 'closed' homo-dimer (Bozkurt *et al*, 2009; Hu *et al*, 2009; Mateja *et al*, 2009; Suloway *et al*, 2009; Yamagata *et al*, 2010). As proposed, the 'closed' dimer uses a helical subdomain (HSD) to form a hydrophobic groove for binding the TM of the TA protein. However, biochemical studies in mammalian extracts showed that TA proteins form complexes with Get3 compatible with a higher order complex (Stefanovic and Hegde, 2007; Favaloro *et al*, 2008). Furthermore, recombinant expression of Get3 with a TA substrate yields a complex capable of TA membrane insertion *in vitro*. In this case, the complex contains a Get3 tetramer suggesting this is the oligomeric state of the targeting complex (Bozkurt *et al*, 2009; Favaloro *et al*, 2010). The difference between the dimer and tetramer models suggested by either structure or biochemistry remains to be reconciled.

A homologue of Get3 was recently identified in archaea that had previously been annotated as an ArsA, a structurally related bacterial ATPase involved in arsenate export (Borgese and Righi, 2010). Based on homology alone, distinguishing an ArsA from a Get3 is difficult; however, several key differences have been identified. The simplest is that Get3 is a homodimer while the ArsA monomer contains a tandem repeat, forming a pseudo-dimer. Get3 lacks the identified metal coordinating residues of ArsA (Boskovic *et al*, 1996; Stefanovic and Hegde, 2007) but contains a unique 'Get3 motif' that is required for TA binding (Mateja *et al*, 2009; Supplementary Figure S1). Additionally, Get3 homologues typically contain a pair of cysteines at their dimer interface that coordinate zinc and are essential for function (CxxC motif) (Metz *et al*, 2006; Stefanovic and Hegde, 2007).

The need for a specific TA targeting protein, such as Get3, was thought to be unique to eukaryotes that contain membrane bound organelles and, thereby, multiple membranes for insertion. The presumption has been that in prokaryotes there is no specialized machinery for delivery of TA proteins,

*Corresponding author. Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 E California Blvd MC 114-96, Pasadena, CA 91125, USA.
Tel.: +1 626 3951796; Fax: +1 626 3958826;
E-mail: clemons@caltech.edu

Received: 3 May 2011; accepted: 9 November 2011; published online: 29 November 2011

as they would not require targeting to a specific membrane, indeed none have been found in eubacteria. The identification of a Get3 homologue in archaea brings this idea into question. Of the currently sequenced archaeal genomes roughly 50% contain a putative Get3. These can be classified into two groups based on the presence of the CxxC motif. They are found in methanogens, halophiles and thermophiles implying that organisms in these extreme environments have an additional level of complexity in membrane insertion.

Here, we present the first structure of an archaeal Get3 from *Methanocaldococcus jannaschii* (*Mj*Get3). The structure is of a symmetrical homo-tetramer and features a hydrophobic chamber that we postulate sequesters the hydrophobic TA. We demonstrate that the archaeal Get3 is capable of TA binding, that the solution state of a fungal Get3/TA complex is similar to the tetrameric *Mj*Get3 and this fungal tetramer complex is capable of membrane insertion. This allows us to postulate a new model for TA targeting by Get3.

Results

Structure of MjGet3

*Mj*Get3 was purified and crystallized after recombinant expression in *E. coli*. We noted that the protein appeared to be a tetramer by size exclusion chromatography (SEC), further discussed below. The best crystals grew in the presence of ADP or the nucleotide analogue ADP·AlF_x in two space groups. The two crystal forms were P2 diffracting to 3.2 Å and $P2_1$ diffracting to 3.3 Å grown in ADP and ADP·AlF_x, respectively. Both structures were solved by molecular replacement (MR), the $P2_1$ structure using a nucleotide-hydrolase domain (NHD) from *Af*Get3 (3IBG) as the search model and the P2 structure using an NHD from the $P2_1$ structure. The remainders of both structures were built independently. The final refined structures had free R-factors of 29.6% for the P2 and 28.6% for the $P2_1$ forms. Data collection statistics are presented in Supplementary Table S1. The two structures are very similar with an RMSD of 0.8 Å over all Cα. There are four copies of *Mj*Get3 in the asymmetric units forming a homo-tetramer (Figure 1A and B; Supplementary Figure S2). The overall structure results in a dumbbell-shaped particle ~150 Å long. Unless noted, all figures will use the structure from the $P2_1$ crystal form.

Individual subunits of *Mj*Get3 are very similar to those from fungal Get3 structures (Figure 1C; Supplementary Figure S3). They closely align in the NHD with an RMSD of 0.8 Å for residues 24–96, 149–171 and 234–333 of *Mj*Get3 to the transition state 'closed' *Sc*Get3 (2WOJ). There is clear density for ADP and a magnesium ion; however, at this resolution we could not clearly resolve the presence of an AlF_x in the $P2_1$ form (Figure 1D). There are a number of features of the NHD including switch I and II loops, which are responsible for transmitting changes in the nucleotide state to conformational changes related to function (Sprang, 1997; Leipe *et al*, 2002). For Get3, it has been suggested that they play a role in modulating between the open and closed state of Get3 as a dimer (Bozkurt *et al*, 2009; Mateja *et al*, 2009; Suloway *et al*, 2009). Here, the switch I region is similar in

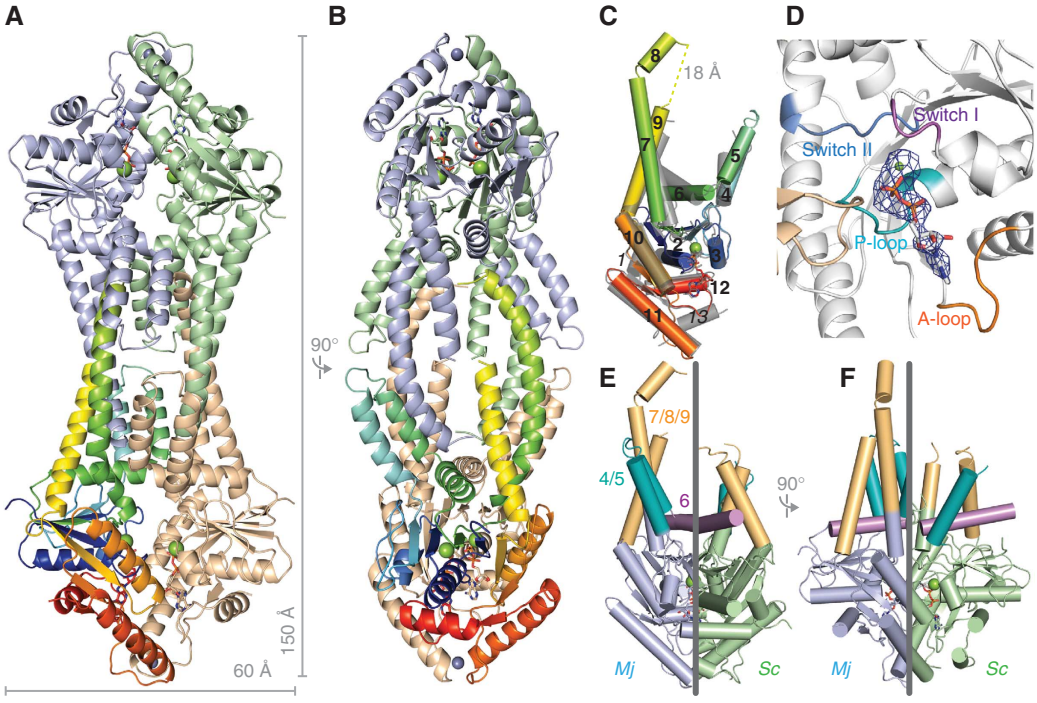

Figure 1 The structure of *Mj*Get3. (**A**) The overall structure of *Mj*Get3 in cartoon representation with each subunit coloured differently and one subunit colour ramped blue to red from N- to C-terminus with approximate measurements on the side. Nucleotides are represented as sticks and ions as spheres. (**B**) 90° rotation relative to (**A**). (**C**) A monomer of *Mj*Get3 colour ramped as in (**A**) overlaid with *Sc*Get3 (2WOJ-A) shown in grey. Helix 1 is not resolved in *Mj*Get3 and helix 13 is not obviously present. Dashed line connects helix 8 to helix 9. (**D**) A view of the nucleotide binding pocket highlighting hydrolase features: P-loop (cyan), switch I (purple), switch II (blue) and A-loop (orange). The opposing subunit in tan. ADP is in sticks coloured by atoms. Mg^{2+} as green sphere. $2F_0-F_C$ density for the nucleotide is shown as a blue mesh contoured at 2.5σ. (**E**) A split-view comparison of the dimers of *Mj*Get3 (lavender) and *Sc*Get3 (2WOJ—green). The dimers each have two-fold symmetry in the views shown. For clarity, only half of each structure from the overlay is shown to give a direct comparison. *Mj*Get3 on the left and *Sc*Get3 on the right. Coloured helices are 4/5 (teal), 6 (purple) and 7/8/9 (light orange). (**F**) A 90° rotation relative to (**E**).

conformation to structures of 'closed' fungal Get3 (Figure 1C; Supplementary Figure S3). Switch II is connected to helix 7. This helix has moved away from the NHD to form the tetramer cage pulling the switch II loop into a conformation not compatible with hydrolysis (Figure 1D; Supplementary Figure S4A and B). The organization of this region is closer to the conformation of the structure of the 'closed' fungal Get3 bound to ADP (3IQX) (Supplementary Figure S3).

As seen before, two of the monomers come together to form a dimer stabilized by a cysteine-coordinated zinc (Figure 1A and B). There are two of these dimers in each tetramer. The structures are very similar and were partially constrained by non-crystallographic symmetry (NCS) during refinement. The orientations of the monomers in the dimer are most consistent with the transition state 'closed' form of *Sc*Get3 (Figure 1E and F; Supplementary Figure S3).

The most dramatic difference between tetrameric *Mj*Get3 and the dimeric 'closed' *Sc*Get3 is the helices that surround the putative TA binding groove (Figure 1E and F). Helix 6, which lies at the base of the groove, matches the conformation of the transition state 'closed' *Sc*Get3 (2WOJ), tilted relative to other fungal structures. Helix 6 is shorter in the *Mj*Get3 tetramer than in the *Sc*Get3 dimer and has moved in the direction of the dimer interface. This shortens the proposed hydrophobic binding groove by ~10 Å (Supplementary Figure S3; Mateja *et al*, 2009). The end of helix 6 becomes more exposed relative to the 'open' form (2WOO) (Supplementary Figure S3) consistent with hydrogen/deuterium exchange experiments that show this helix exchanges hydrogen more rapidly after TA binding (Bozkurt *et al*, 2009). The five helices flanking the groove (helices 4, 5, 7, 8 and 9) are in a similar orientation to that seen in the 'closed' fungal structures (Figure 1E and F). They are extended and more ordered, with only eight residues connecting helix 8 to helix 9 missing backbone density.

Instead of a hydrophobic binding groove, the *Mj*Get3 tetramer uses the flanking helices to form the walls of a cage generating a hydrophobic central chamber (Figure 1A and B). Here, and in closed fungal Get3 structures, the loop formed by helices 4 and 5 (4/5 loop) is on the opposite side of helix 6 from the loop formed by helices 7 through 9 (7/8/9 loop) (Figure 1C). In *Mj*Get3, the 4/5 loop tilts away from the binding groove. In doing so it forms a three-helix bundle to helix 8 extended from a subunit across the tetramer stabilized by hydrophobic interactions (Figure 1A and B). Calculated by PISA (Krissinel and Henrick, 2007), the tetramer buries an additional ~6000 Å^2 relative to the dimers (~11 000 Å^2) with the three-helix bundle providing most of the contacts.

The chamber formed by the tetramer is predominantly hydrophobic (Figure 2A). The size of the internal cavity is ~30 Å across the middle to ~40 Å down the long axis (Supplementary Figure S5). This is compatible with the dimensions of a single TM with volume remaining for additional helices. The chamber is lined by helices 5, 7, 8 and 9. Unlike in the dimer model, helix 4 is on the periphery as part of the three-helix bundle (Figure 2B). The two helix 8s from the dimer extends into the groove of the opposing dimer (Figure 2B). The result is that they cover the floor and block the ends of the groove. This configuration prevents direct contact from helix 6 to the chamber. The cage has openings at the site where the disordered loop between helix 8 and 9 is missing (Figure 2C), which could provide an access point

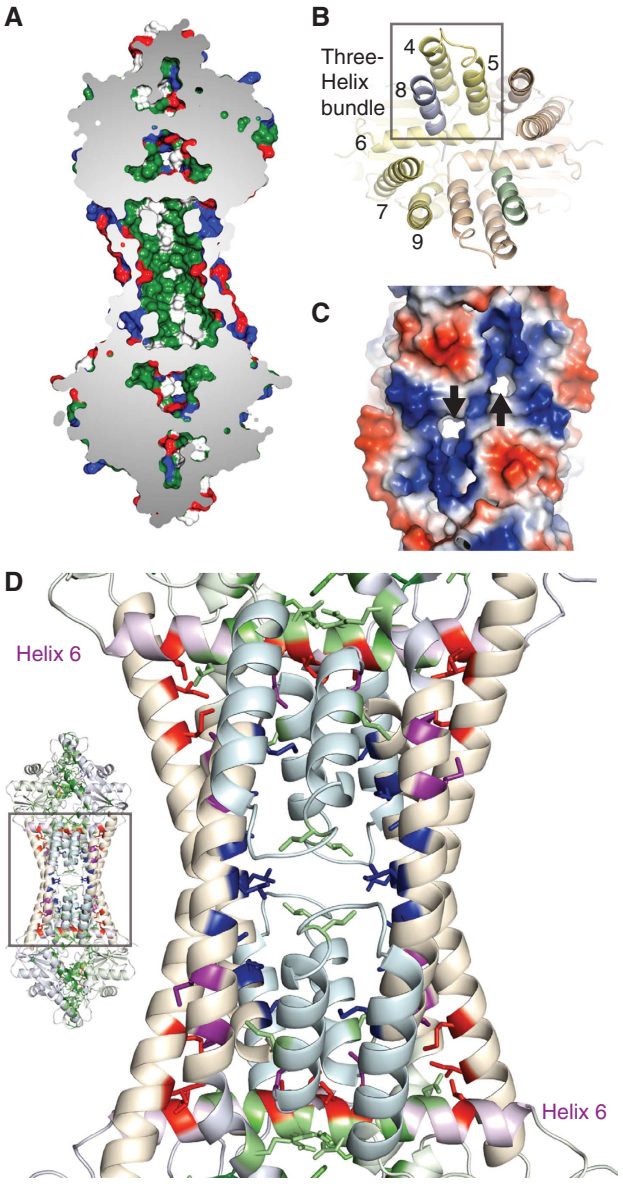

Figure 2 Central cavity. (**A**) Surface representation of *Mj*Get3 cut through the middle. Residues coloured based on type: positive charge (blue), negative charge (red) and hydrophobic (green). (**B**) Similarly to Figure 1A rotated 90° forward showing the cage walls with the foreground removed. (**C**) External view of the cavity in surface representation coloured based on electrostatic potential (negative—red to positive—blue). Holes indicated by arrow. (**D**) External view of the central cavity highlighting mutants from previous studies, inset shows zoom in region coloured similarly to Figure 1E. Mutants resulting in a negative growth phenotype by Suloway *et al* (2009) are shown in green. Mutants from Mateja *et al* (2009) are coloured according to loss of nucleotide hydrolysis (red), TA binding (blue) or both (purple).

between the internal chamber and the cytoplasm. This opening, when viewed based on electrostatic potential, has an overall positive charge. The general charge is conserved in eukaryotes and is consistent with a discrimination point preventing binding of mitochondrial TA proteins (Mateja *et al*, 2009).

Extensive mutagenesis has been performed on *Sc*Get3 by our laboratory and others (Figure 2D; Supplementary Figure S6; Mateja *et al*, 2009; Suloway *et al*, 2009). Mateja

et al (2009) focused on the putative TA binding groove and identified three classes of mutations: those that affect TA binding (blue), nucleotide hydrolysis (red) or both (purple). The location of residues that affect nucleotide hydrolysis predominantly cluster to the base of the groove on helix 6; however, in the dimer model these residues did not affect TA binding despite being components of the putative groove. In the tetramer structure, the reason for this becomes clear as these residues do not directly contact the hydrophobic chamber and, therefore, would not be predicted to affect binding (Supplementary Figure S6B and C). They do lie directly below the chamber and would be expected to communicate the state of the chamber to the NHD. The mutants that affect TA binding all coat the interior of the hydrophobic chamber (Figure 2D; Supplementary Figure S6C).

Archaeal Get3 binds TA proteins

The high sequence and structural homology between fungal and archaeal Get3 suggest that the archaeal protein is capable of binding TA proteins. We established a method for purifying a Get3/TA protein complex heterologously in *E. coli* using a two-step purification procedure where both Get3 and the TA protein contain affinity tags (Figure 3A; similarly to Bozkurt *et al*, 2009; Favaloro *et al*, 2010 and Yamagata *et al*, 2010). Using this method, we could reliably purify *Sc*Get3 bound to a variety of TA substrates. We analysed the complex using inductively coupled plasma mass spectrometry (ICP-MS) to identify bound ions (Yamagata *et al*, 2010). We found an approximate 2:1 ratio of Get3 to zinc but only trace amounts of magnesium suggesting that there are no appreciable amounts of bound nucleotide, consistent with what has been reported before. An example of this complex is the purification of *Sc*Get3 bound to Sbh1, the yeast homologue of mammalian Sec61β and a demonstrated GET pathway substrate (Figure 3B, lane 1; Schuldiner *et al*, 2008). Interestingly, when we co-expressed *Mj*Get3 with Sbh1 we were able to purify a stable chimeric complex confirming that the archaeal homologue is capable of binding a TA protein (Figure 3B, lane 2). As a control, co-expression of *Mj*Get3 with a TM deletion of Sbh1 was unable to form a complex by this method, demonstrating that *Mj*Get3 specifically bound the TA (Supplementary Figure S7A and B).

Next, we tested to see if the archaeal Get3 was capable of binding a native substrate. We first co-expressed *Mj*Get3 with *Mj*Secβ, the archaeal homologue of Sec61β whose structure has been solved and is predicted to be a TA protein (van den Berg *et al*, 2004; Borgese and Righi, 2010). Using our two-step purification, we were able to obtain a stable complex of the two proteins (Figure 3B, lane 3). Neither component was recovered when expressed alone and purified by the same two-step procedure.

A number of additional TA proteins have been identified bioinformatically from *Methanococcus maripaludis*, a related species to *M. jannaschii* (Borgese and Righi, 2010). Some of these have homologues in *M. jannaschii* including SecE, another Sec channel component and MtrA and MtrB, TA subunits of tetrahydromethanopterin S-methyltransferase. All three proteins could be purified as a complex bound to *Mj*Get3 (Figure 3B, lanes 4–6). We wanted to test another archaeal homologue and chose to work with the Get3 from *Thermococcus kodakaraensis* (*Tk*Get3). This Get3 falls into the class of archaeal homologues missing the pair of zinc-coordinating cysteines. Using the two-step method, *Tk*Get3 is capable of forming stable complexes with the same set of TA proteins tested for *Mj*Get3 (Figure 3B, lanes 7–11).

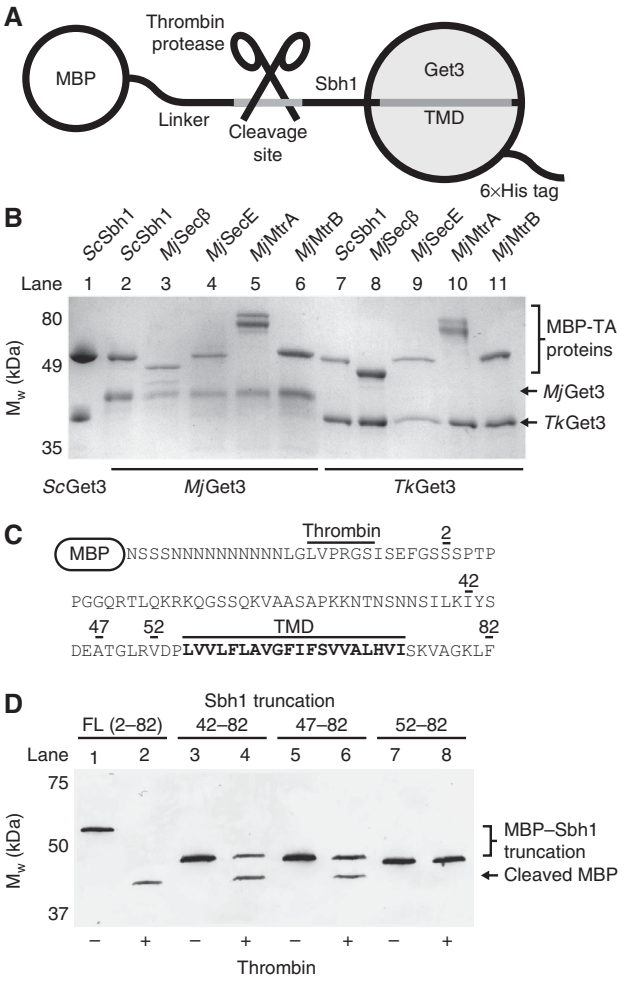

Figure 3 TA protein binding by Get3. (**A**) Diagram of recombinantly expressed complex showing the two affinity tags used for purification (MBP, maltose binding protein). Position of thrombin cleavage site indicated. (**B**) Coomassie-stained SDS–PAGE of Get3/TA protein complexes from various species purified by recombinant co-expression. (**C**) Sequence of Sbh1 fusion to MBP with residues from truncations indicated. The TMD is in bold. (**D**) SDS–PAGE and western blot of the *Sc*Get3/Sbh1 truncation complexes pre- and post-thrombin cleavage with αMBP antibody against the MBP–Sbh1 fusion. Accessibility of protease site results in a shift of the MBP–Sbh1 fusion to a lower MBP band. Residues numbers of Sbh1 are indicated. Figure source data can be found in Supplementary data.

The binding chamber only sequesters the TM

Get3 specifically recognizes and binds the hydrophobic TA (Stefanovic and Hegde, 2007). The nature of the tetramer's hydrophobic chamber suggests it is capable of sequestering a TA; however, it would be unable to fit a typical soluble domain. We would postulate that while the TA is in the chamber the soluble domain extends out through the gap in the cage. The consequence would be that Get3 would only protect a minimal amount of the TA protein from solvent. Our double affinity tagged system provides a means to test this. Here, we took the N-terminal fusion of MBP to Sbh1 and introduced a thrombin protease site between the two proteins, which we then co-expressed with *Sc*Get3 and purified

(Figure 3A and C). For the full-length MBP-Sbh1 fusion, the addition of thrombin resulted in a gel shift of MBP from running as a fusion to MBP alone (Figure 3D, lane 1 versus lane 2). TM topology predicting software, TMHMM 2.0 (Krogh *et al*, 2001), predicts the TM to extend from residue 55 to 74 (Figure 3C). We generated constructs with progressively shorter N-terminal soluble domains of Sbh1 and tested them in the assay. A complex of *Sc*Get3 with an MBP fusion of Sbh1 truncated to eight amino acids N-terminal of the TM could be cleaved (Figure 3D, lane 5 versus 6); however, Sbh1 truncated to three amino acids N-terminal to the TM could not (Figure 3D, lane 7 versus 8). Therefore, Get3 sequesters only a few amino acids in addition to the hydrophobic TM. This result could also be consistent with the dimer model where the groove covers only the TA.

Get3 can form a tetramer in solution

The current model for the function of fungal Get3 is based on Get3 always maintaining a dimeric state. *Sc*Get3 purified after expression in *E. coli* is predominantly a dimer by SEC; however, a small pool always purified as a tetramer (Figure 4A and B, solid blue trace). This tetramer pool was stable enough to be rerun over the column (Figure 4B, dashed line). We noted that the tetramer fraction would degrade over time to dimer, while we would never see conversion of the dimer to tetramer. We suspected, as noted below, that the tetramer fraction is stabilized by interactions with hydrophobic peptides. In this case, tetrameric *Sc*Get3 might be bound to hydrophobic *E. coli* peptides. We searched for evidence of these by mass spectrometry but were unable to find any *E. coli* peptides. This may not be surprising, as the expected hydrophobic peptides are typically hard to identify by liquid chromatography tandem mass spectrometry (Wu and Yates, 2003) and may be in low abundance.

In contrast to *Sc*Get3, the solution state of *Mj*Get3 is primarily a tetramer by SEC (Figure 4A, green trace), which is consistent with the crystal structure. We were interested to see if this was a general feature of archaeal Get3 homologues. Get3 from the mesophile *M. maripaludis* (*Mm*Get3) was predominantly tetrameric in solution (Figure 4A, red trace). *Tk*Get3, which lacks the coordinating cysteines, is similar to *Sc*Get3 in that the protein elutes as both a tetramer and a dimer (Figure 4A, cyan trace).

The interactions forming the tetramer in the *Mj*Get3 crystal structure are primarily hydrophobic packing between helix 8 on one subunit and helices 4 and 5 from the opposing subunit (Figure 4C). This suggests that the tetramer should be sensitive to detergent. To test this, we dialysed the protein against a 1% (w/v) solution of the small micelle detergent *N*-octyl-β-D-glucopyranoside (β-OG) just above the critical micelle concentration. After overnight dialysis, the majority of the *Mj*Get3 shifted to a peak consistent with a dimer (Figure 4D, compare green trace with black trace). When we then dialysed out the detergent, the protein returned to the tetramer (Figure 4D, black dashed trace) confirming that in these conditions this was the stable state. When we diluted out the detergent in the presence of ATP we saw no effect on the oligomeric state, although we did note an increase in absorption consistent with nucleotide binding (260/280 ratio went from 0.81 without nucleotide to 1.02 in the presence of ATP). This was intriguing as adding ATP to purified tetrameric *Mj*Get3 would cause the protein to precipitate. For further confirmation of the stability and hydrophobic nature of the interaction, we performed SEC of MjGet3 in the presence of high salt (1 M NaCl) or denaturant (1 M urea). Neither affected the oligomeric state (Supplementary Figure S8).

Helix 8 had been suggested to play an important role in TA binding, possibly as a cover to the groove proposed in the dimer model (Mateja *et al*, 2009). Deletion of this helix resulted in a loss of TA binding (Yamagata *et al*, 2010). In the tetramer model, this helix stabilizes the interface; therefore, it is critical to formation of the TA binding hydrophobic chamber. To test this directly, we introduced mutations in helix 8 of *Mj*Get3 near the interface of the three-helix bundle (Figure 4C). Two of these, F192D and M196D, directly disrupt the hydrophobic interface and both of them shift predominantly to a dimer peak by SEC (Figure 4E, red and purple trace). A third mutation, M193D, was not at the interface (Figure 4C) and this mutant strongly favoured the tetramer (Figure 4E, orange trace). Partial deletion of helix 8 resulted in a loss of the tetramer peak as well (Figure 4E, yellow trace).

Fungal Get3 tetramer is capable of TA membrane insertion in vitro

A biological role for the tetramer complex is difficult to prove directly. We decided to address this by demonstrating that our purified Get3/TA tetramer complex was capable of insertion into purified yeast microsomes. It had previously been shown that both a fungal Get3/TA tetramer complex (Bozkurt *et al*, 2009) and a mammalian Get3/TA tetramer complex (Favaloro *et al*, 2010) are capable of insertion into mammalian microsomes. In both cases, membrane integration was verified by glycosylation of the C-terminus of the TA protein, which can only happen if the C-terminus has entered into the lumen of the microsomes. We chose to perform a similar assay using all fungal components.

We generated an MBP-tagged Sbh1 with a glycosylation site from opsin at its C-terminus (MBP–Sbh1-op). This purified as a stable tetrameric complex with *Sc*Get3. Using this complex, we observed successful insertion of MBP–Sbh1-op into *S. cerevisiae* Δ*get3* microsomes by glycosylation of the C-terminal opsin tag, confirmed by subsequent deglycosylation with the endoglycosidase EndoH after disruption of the microsomes (Figure 5A, lanes 3 and 4). As previously seen (Favaloro *et al*, 2010), no insertion was observed when microsomes were pre-treated with trypsin (Figure 5A, lane 1). Furthermore, insertion is sensitive to the binding of nucleotide, here disrupted by the addition of EDTA to the reaction (Figure 5A, lane 2). MBP–Sbh1-op could be purified without Get3 and alone failed to insert (Figure 5A, lane 5). This shows that the *Sc*Get3/TA tetramer complex is on a functional insertion pathway. We were interested to see if we could get transfer from an archaeal complex into our yeast microsomes. We purified a stable *Mj*Get3/MBP–Sbh1-op tetramer complex. This complex was unable to insert the yeast substrate into *S. cerevisiae* Δ*get3* microsomes (Supplementary Figure S9A).

The inability of the archaeal *Mj*Get3/yeast TA tetramer complex to insert into yeast microsomes implies that there are significant differences in the putative archaeal pathway at the membrane. Indeed, there are no obvious homologues of either Get1 or Get2 in any archaea. To further explore the species dependence of our insertion results, we decided

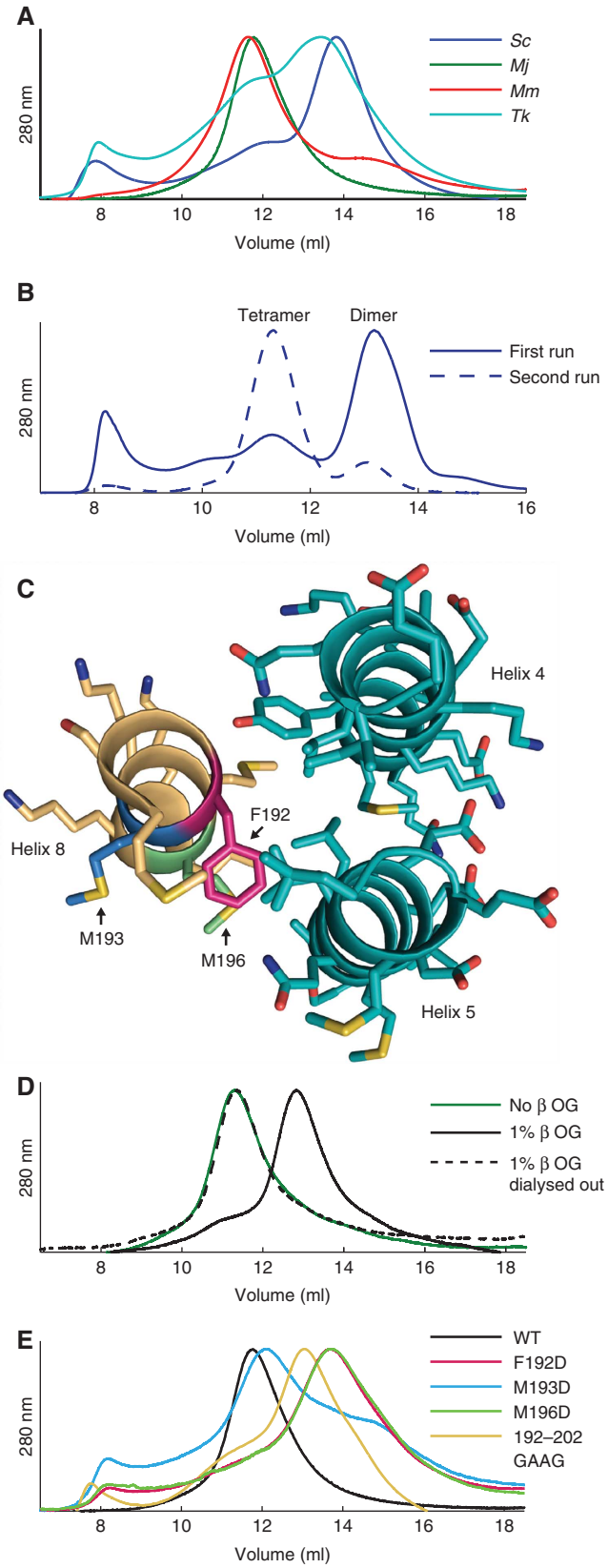

Figure 4 Oligomeric state of Get3. (**A**) SEC of Get3 from *S. cerevisiae* (blue) and three archaeal species: *M. jannaschii* (green), *M. maripaludis* (red) and *T. kodakaraensis* (cyan). Tetramers and dimers eluted around 12 and 14 ml, respectively. (**B**) SEC run on different columns of *Sc*Get3 after affinity purification (solid line). The tetramer peak was pooled and rerun (dashed line). (**C**) The three-helix bundle that stabilizes Get3 tetramers coloured as in Figure 2B with mutated residues highlighted. (**D**) SEC of *Mj*Get3 in the absence (green) and presence (black) of 1% β-OG and a sample where detergent was dialysed out (dashed). Detergent shifts the peak to a volume corresponding to a dimer (13 ml). After dialysis, the peak shifts back to a tetramer (11 ml). (**E**) SEC of *Mj*Get3 mutants. The 192–202 GAAG trace (yellow) corresponds to a deletion of helix 8.

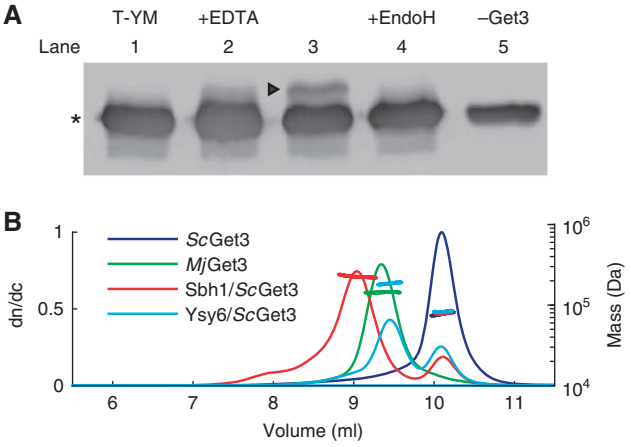

Figure 5 Functional studies of Get3/TA protein tetramer complex and SEC-MALLS. (**A**) *In-vitro* membrane integration of an MBP fusion to Sbh1 with a C-terminal opsin tag into microsomes purified from *S. cerevisiae* Δ*get3* by a purified *Sc*Get3/MBP–Sbh1-op complex. A western blot against MBP of *in-vitro* translocation assays into trypsinized yeast microsomes (T-YM—lane 1), in the presence of EDTA (lane 2), standard *in-vitro* translocation conditions before (lane 3) and after EndoH treatment (lane 4) and MBP–Sbh1 purified without Get3 (lane 5). Star indicates MBP–Sbh1-op and the arrow points to band shifted by glycosylation of MBP–Sbh1-op after membrane integration. (**B**) Molecular weights of Get3 and Get3/TA protein complexes measured by SEC coupled to MALLS. Traces of differential index of refraction (dn/dc) and calculated molecular weights are shown. Figure source data can be found in Supplementary data.

to use a reconstituted *in-vitro* translation system that has previously been used to demonstrate insertion by eukaryotic Get3 variants (Stefanovic and Hegde, 2007). Using lysates from the *S. cerevisiae* Δ*get3* strain, we were able to reconstitute insertion into purified microsomes using a model substrate (N-terminal truncation of MBP–Sbh1-op) dependent on the presence of *Sc*Get3 (Supplementary Figure S9B). Based on this, we decided to see if any of the archaeal Get3 homologues could facilitate insertion. Similarly to the purified complex assay, *Mj*Get3 was unable to facilitate insertion, as was *Mm*Get3 or *Tk*Get3 (Supplementary Figure S9B and C). This suggests that, under these conditions, archaeal Get3 homologues are unable to support insertion into fungal microsomes.

Solution characterization of Get3 and Get3–TA protein complexes

Complexes between fungal or mammalian Get3 bound to TA substrates are consistent with a tetramer of Get3 (Bozkurt *et al*, 2009; Favaloro *et al*, 2010). We decided to assay the oligomeric state of various complexes using multi-angle laser light scattering (MALLS), which provides a relative molecular weight of the particle. Using this method, both the purified *Sc*Get3 dimer and the tetrameric *Mj*Get3 were consistent with calculated molecular weights, 80.8 and 153 kDa, respectively (Figure 5B; Table I). Using purified *Sc*Get3/TA complexes, the experimental molecular weights were consistent with a tetramer of *Sc*Get3; however, the stoichiometry of the TA proteins was inconclusive (Figure 5B; Table I). This suggests that more than one TA protein is bound per Get3 tetramer possibly because the complex is formed in the absence of other GET partners. The size of the chamber should be able to accommodate multiple TA proteins (Supplementary Figure S5).

Table I MALLS and SAXS statistics

MALLS	*Sc*Get3	*Mj*Get3	*Sc*Get3/Ysy6	*Sc*Get3/Sbh1
Get3 monomer (kDa)	41.3	38.8	40.2	39.4
Get3 oligomer (kDa)	82.5	155	161	157
M_w measured (kDa)	80.8	153	198	238
Difference (kDa)	−2.5	−2.0	+37.1	+80.6
TA M_w (kDa)	—	—	8.8	9.3
Difference/TA M_w	—	—	4.2	8.6
SAXS		*Mj*Get3 (ADP·AlF_x)	*Sc*Get3/Ysy6	*Sc*Get3/Sbh1 47–82
R_g theoretical (Å)		44.2	—	—
R_g Guinier (Å)		47.0	49.8	52.1
R_g GNOM (Å)		47.5	49.6	51.0
D_{max}		165	173	163

The data are consistent with the *Mj*Get3 tetramer having a similar conformation to the Get3/TA complex. We expected that the elongated dumbbell structure and the stability of the purified complexes could be exploited in an analysis using biological small-angle X-ray scattering (bioSAXS). The benefit of this technique is that it can provide measures of dimensions in solution along with allowing for the calculation of low-resolution molecular envelopes (Putnam *et al*, 2007). bioSAXS curves of *Mj*Get3 and *Sc*Get3/TA complexes show similar dimensions such as the radii of gyration (R_g) and maximum dimension (D_{max}) (Table I). The unbiased overall shapes of the pair-distribution functions, P(r), are also similar with a primary peak followed by a shoulder peak (Figure 6A), strongly suggesting a multi-domain protein consistent with the crystal structure of *Mj*Get3. This further indicates that the fungal and archaeal tetramers have similar overall architectures.

In addition to overall dimensions, bioSAXS can use the experimentally measured distance distributions (Figure 6A) in a variety of refinement procedures to obtain *ab initio* molecular envelopes (Putnam *et al*, 2007). We used DAMMIN (Svergun, 1999) to calculate molecular envelopes of each of our complexes. Knowing that the Get3 tetramer has internal symmetry, we imposed P22 symmetry on our model. This did not affect the overall dimensions of any of the complexes; however, it was necessary to establish a consistent envelope. We calculated a molecular envelope for *Mj*Get3 (Figure 6B; Supplementary Figure S10). The *ab initio* fit to the data returned a dumbbell-shaped envelope consistent with the crystal structure. Viewed down the two-fold along the long axis of the tetramer, we measured a crossing angle of the widest point in each of the dimers. In the crystal structure, this crossing angle is ∼30°. Performing a similar measure for the molecular envelope results in a crossing angle of ∼40°. This suggests that the crystal structure stabilizes a slightly twisted form of the *Mj*Get3 compared with the solution state. We generated a molecular envelope for both the *Sc*Get3/Ysy6 complex and a truncated *Sc*Get3/$Sbh1_{47–82}$ complex (Figure 6C and D; Supplementary Figure S10). These resulted in very similar dumbbell-shaped structures confirming that the *Mj*Get3 tetramer is a good model for the fungal Get3/TA tetramer complex.

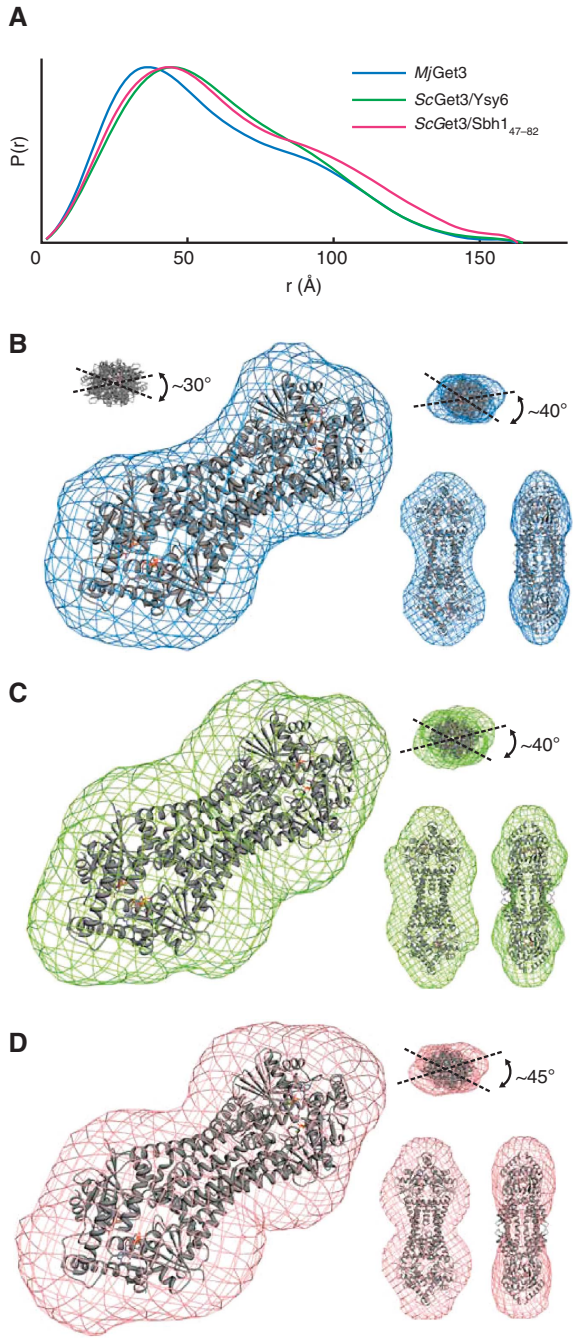

Figure 6 Size and shape of Get3/TA protein complexes. (**A**) Pair-distribution functions from bioSAXS of *Mj*Get3 (blue), *Sc*Get3/Ysy6 (green) and *Sc*Get3/Sbh1$_{47–82}$ (salmon). (**B**) Calculated envelope of *Mj*Get3 from bioSAXS data (blue mesh) with *Mj*Get3 coordinates fit to the overall envelope. Insets are three views 90° rotated around the vertical and horizontal axis. Lines are drawn through the widest point of each lobe and a crossing angle is calculated along the long axis. Top left inset shows the crossing angle for the structure alone. (**C**) Similarly to (**B**) for the envelope of *Sc*Get3/Ysy6 complex (green mesh). (**D**) Same as (**C**) for the envelope of the *Sc*Get3/Sbh1$_{47–82}$ complex (salmon mesh).

Discussion

The details of the GET pathway are rapidly being discovered. So far, Get3 is the best characterized of the components, yet there are still a number of important questions that remain to be answered. Here, we solve the structure of an archaeal Get3. The structure is a tetramer with a hydrophobic chamber that we postulate sequesters TA. We demonstrate that archaeal homologues are capable of TA protein binding. Moreover, we show the first structural information of fungal Get3/TA protein complexes, which is consistent with the biochemical data supporting a tetramer model for TA binding.

The presence of a Get3 homologue in archaea is exciting and suggests a novel membrane protein-targeting pathway in this domain of life. The lack of homologues of other GET components implies that the pathway, if it exists, will be significantly different. Based on the structural homology, the fact that archaeal Get3s can bind TA proteins is not surprising; however, it supports the possibility of an archaeal TA targeting pathway. It is also interesting that an archaeal homologue that does not contain the CXXC motif, *Tk*Get3, is capable of both oligomerization and TA binding. This motif is essential in fungal Get3 homologues and may suggest an evolutionary path for these proteins. Unlike certain TA proteins (e.g., Secβ and SecE), Get3 homologues are not found universally in archaea. This hints that the presumed pathway may not be essential or may be required for specific substrates. The fact that very closely related archaea differ in having a Get3 homologue is a question for further study (e.g., *Pyrococcus abysii* versus *Pyrococcus furiosus* and *Pyrococcus horikoshii*).

Tetramers of Get3 have been seen in a variety of contexts suggesting that this state plays a functional role in the GET pathway. In the initial functional identification of TRC40 (the mammalian homologue of Get3), the protein isolated from *in-vitro* translation in a reticulocyte lysate was seen in a large complex ranging in size compatible with the tetramer (Stefanovic and Hegde, 2007). Similarly, the crosslinking of a mammalian Get3/TA complex was consistent with tetramers by SDS–PAGE (Favaloro *et al*, 2008). Bozkurt *et al* (2009) purified a fungal Get3/TA protein complex that they determined was a Get3 tetramer by analytical ultra-centrifugation and SEC. This complex, in an *in-vitro* insertion assay using mammalian ER microsomes, is competent for membrane insertion. Favaloro *et al* (2010) completed a similar study using a mammalian Get3/TA complex. In this case, the size of the complex was compatible with a tetramer by SEC and was also competent for membrane insertion. We have now replicated these results in a purified fungal system. Thus, in the cases where functional insertion was demonstrated, the oligomeric state of Get3 is a tetramer.

It is clear that eukaryotic Get3 can exist as a stable dimer and it is likely this state plays a functional role. We, and others, have demonstrated that Get3 is a tetramer when bound to a TA substrate and this complex is capable of TA insertion. This conflicts with the dimer model suggested based on earlier structures. As we now present a contrasting model, it is useful to posit the differences in the two models. The dimer model of Get3 TA binding suggests that the HSD captures the length of a TM along a hydrophobic groove with a floor provided by helix 6 and helices 4, 5, 7 and 9 providing the walls. In the unbound transition state structure, the dimensions of the groove seem compatible with a hydrophobic TM helix; however, while covering three sides of the protein the groove leaves one face of the protein exposed to solvent. It has been suggested that helix 8 solves this by becoming ordered upon TA binding to cover the exposed face.

Consistent with this, mutations in the walls of the groove and deletion of helix 8 prevent TA binding (Mateja *et al*, 2009; Yamagata *et al*, 2010). Similar mutations at the bottom of the groove do not significantly affect TA binding; however, they clearly play a role in the targeting process (Mateja *et al*, 2009; Suloway *et al*, 2009). The dimer model requires that a structured groove accommodate a wide variety of TA substrates along with necessitating the binding of an α-helix. This is different from the model for SRP binding of a hydrophobic signal sequence, which uses a flexible loop to form its helical binding pocket (Bernstein *et al*, 1989; Keenan *et al*, 1998).

The tetramer model resolves a number of the issues that arise in the dimer model. Here, the residues in helix 6 that affect nucleotide hydrolysis do not contact the hydrophobic cage directly and would not be expected to directly affect TA binding. Instead, they would relay the binding of substrate and oligomerization state to the NHD. Moreover, in a deuterium exchange assay the binding of TA leads to exposure of the C-terminus of helix 6, as seen in the tetramer; whereas, in the dimer model this should be occluded by TA (Bozkurt *et al*, 2009). The binding of a hydrophobic TA stabilizes the tetrameric state of fungal Get3. This suggests that energy must be input to destabilize the complex. It also suggests why a Get3 complex with cytochrome b_5, a more hydrophilic TA, does not require energy for insertion as it is likely less stable (Stefanovic and Hegde, 2007; Favaloro *et al*, 2010). It is interesting to note that in the conditions tested here the archaeal tetramer is significantly more stable than eukaryotic homologues independent of bound TA.

Both Get3 and ArsA are members of the 'deviant Walker A motif' family of ATPases (Koonin, 1993). ArsA is structurally very similar to Get3 (Suloway *et al*, 2009) and uses ATP hydrolysis to facilitate arsenite export (Walmsley *et al*, 1999). Binding of arsenite stimulates the monomer (effectively a pseudo-dimer) to dimerize (analogous to a Get3 tetramer) (Ching *et al*, 1991). Structures of ArsA exist only as monomers, so there appears to be an analogous oligomerization in this related system. It would be interesting to see if the solution structure of ArsA bound to arsenite is compatible with the tetramer described here. Other members of this ATPase family that have been characterized are soluble dimers that also form higher order functional complexes either as homo-oligomers (e.g., Soj; Leonard *et al*, 2005 and MinD; Hayashi *et al*, 2001) or as hetero-oligomers (e.g., NifH; Schindelin *et al*, 1997); therefore, a role of higher order oligomerization may be a general feature of this family.

A remaining complication in this study is the stoichiometry of the tetramer/TA complex. The current idea is that a single TA protein binds to the Get3 complex. While this is attractive, there is no biochemical data that supports this; indeed, it is difficult to prove. It is clear that our tetramer complex contains minimally a single TA protein; however, the biophysical data suggest that there are more copies bound. The size of the chamber easily accommodates a single TM with ample space for additional TMs. Moreover, the hydrophobic nature of tetramer formation presumably allows for flexibility of the chamber that may expand to bind more substrates. A recent study by Leznicki *et al* (2011) used chemical modification to the TM of the TA to explore the flexibility of the TA binding pocket. Addition of a single large polyethylene glycol (PEG) adduct to the TM did not inhibit binding or insertion; however, modification at two sites prevented binding independent of whether the attachments were on the same or opposite sides of a presumed helix. The single site addition is clearly compatible with a dimer model; however, it is inconsistent with a second binding site on the same side not binding. The tetramer model is also consistent with a single site modification as the two proposed chamber access points could accommodate both protein and extended PEG. A second site would have a harder time being accommodated and would presumably reduce the affinity. Overall, these results point to surprising flexibility in binding to Get3.

The role of nucleotide in TA targeting remains to be determined. All of the current evidence demonstrates that nucleotide hydrolysis is required at the membrane but not for TA binding (Stefanovic and Hegde, 2007; Favaloro *et al*, 2008, 2010). Mutants deficient in nucleotide binding efficiently bind to TA substrate in both *in vitro* (Stefanovic and Hegde, 2007) and heterologous expression assays (Yamagata *et al*, 2010). Structural studies suggest that Get3 undergoes distinct conformational changes from an 'open' to 'closed' form that is stimulated by the nucleotide state. The closed form of Get3 has only been seen in the presence of nucleotide, independent of the γ-phosphate (Bozkurt *et al*, 2009; Mateja *et al*, 2009). The open form has been seen in both the apo and ADP complex (Mateja *et al*, 2009; Suloway *et al*, 2009; Yamagata *et al*, 2010). This suggests that there is conformational flexibility with bound nucleotide favoring the 'closed' state (Chartron *et al*, 2010). Neither our purified *Mj*Get3 nor our *Sc*Get3/TA complex contain bound nucleotide consistent with what has been seen before (Bozkurt *et al*, 2009; Favaloro *et al*, 2010). Purified Get3/TA complexes require nucleotide to stimulate insertion implying that the nucleotide-binding pocket (NBP) is solvent accessible (Bozkurt *et al*, 2009; Favaloro *et al*, 2010). With a bound transition state analogue, the NBP is closed to solvent (Mateja *et al*, 2009); however, in the closed form bound to ADP the switch II loop has moved, exposing the NBP (Bozkurt *et al*, 2009). This conformation cannot hydrolyse ATP and is similar to the position of switch II in our *Mj*Get3 tetramer. Therefore, in the Get3 tetramer the NBP is more accessible allowing diffusion in or out of nucleotide. The high cellular concentration of ATP makes it likely that this is the bound form. Disruption of the tetramer would be required for switch II to occupy a hydrolysis competent conformation. This disruption would result in substrate release and may be facilitated by factors at the membrane.

In the fungal GET pathway, each of the soluble proteins exists minimally as dimers. Sgt2, the first protein in the pathway to specifically bind the TA (Tobaben *et al*, 2003; Liou *et al*, 2007; Wang *et al*, 2010), is a homodimer with an N-terminal dimerization domain (Liou and Wang, 2005). Get4 and Get5, which link Sgt2 to Get3, are hetero-tetramers with a C-terminal dimerization domain in Get5 (Chartron *et al*, 2010). This all suggests a larger functional complex in TA protein recognition and delivery (Chartron *et al*, 2011). The ability of Get4, minimally present in two copies, to bind directly to a dimer of Get3 is consistent with the possible specific recognition of a tetramer or two dimers.

All of this allows us to suggest a modified model for Get3 mediated targeting of a TA protein (Figure 7). (1) Get3 in its apo form is a stable and soluble dimer in equilibrium between an 'open' and 'closed' form. (2) Binding of nucleotide

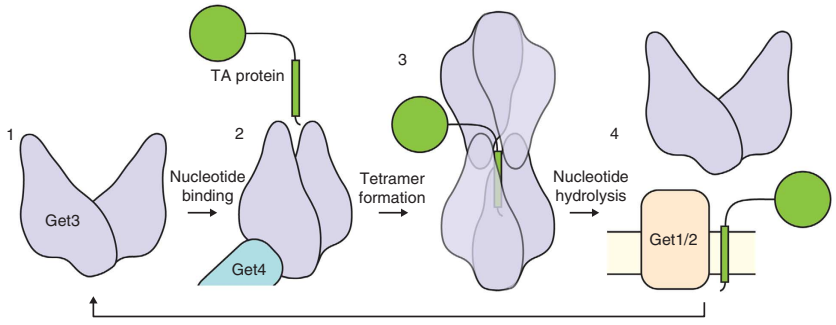

Figure 7 Model for Get3 TA targeting.

shifts the equilibrium towards the 'closed' form that is compatible with binding Get4. This is now competent for TA binding. (3) Binding of the TA results in tetramer formation. Conformational changes to form the tetramer cause release from Get4 and diffusion of the Get3/TA complex to the membrane. (4) Here, the receptor complex stimulates release of TA and insertion into the membrane. Get3 now returns to the dimer state to bind new TA substrates.

A related structural study on archaeal Get3 was published while this manuscript was in review. Their crystal structure of Get3 from the archaea *Methanothermobacter thermautotrophicus* (*Mt*Get3) is consistent with the results presented here in that the archaeal homologue is structurally similar to fungal counterparts (Sherrill *et al*, 2011). In this case, a dimer was specifically purified and crystallized; therefore, they do not investigate the tetramer that we report. Interestingly, the loops that extend to form our tetramer are also extended in the dimer of *Mt*Get3 perhaps explaining why the archaeal tetramer is more stable. Excitingly, they were able to demonstrate that *Mt*Get3 can facilitate TA insertion using a protease protection assay. This difference from our result might be assay specific or could be related to the species of Get3 and TA protein tested.

Also while this manuscript was in review, two reports on the interaction of Get3 with the membrane proteins Get1 and Get2 were published (Mariappan *et al*, 2011; Stefer *et al*, 2011). In both studies, structures of Get3 dimers are bound to the soluble domains of either Get1 or Get2. The structures suggest how Get1 and Get2 can recognize Get3 and facilitate release of a TA protein. These structures do not contain TA proteins. We believe the evidence is compatible with a model in which a tetrameric Get3/TA protein complex is captured by Get2 then disrupted by Get1 to release the TA protein for insertion.

Get3 is a dynamic protein that undergoes a complex series of conformational changes in delivery of a TA protein to the ER. Here, we present the first structural information of Get3/TA complex from a heterologously expressed system demonstrating that Get3 in this state is a tetramer. The tetramer model suggests a TA is sequestered within a hydrophobic chamber. Further studies are required to establish the role of the tetramer *in vivo*. An unresolved point is the stoichiometry of the TA proteins to Get3, which probably requires other GET components to determine. Finally, the functional role of the Get3 archaeal homologue is a tantalizing question, particularly with the broader context of the detailed molecular mechanism of TA protein targeting by the GET pathway.

Materials and methods

Cloning, expression and purification

All Get3 homologues were amplified from genomic DNA and cloned into a pET33b vector (Novagen) modified to contain only an N-terminal 6 × His tag. Genomic DNA for *Mj*Get3 (MJ_1142) from *M. jannaschii* DSM 2661 (ATCC), *Tk*Get3 (TK_0994) from *T. kodakaraensis* KW128 (Santangelo *et al*, 2007), *Mm*Get3 (MmarC7_1163) from *M. maripaludis* C7 (ATCC), and *Sc*Get3 from a previous study (Suloway *et al*, 2009). For *Mj*Get3, site-directed mutagenesis was used to generate a truncation encoding amino acids 12–349 of *Mj*Get3 (*Mj*Get3$_{12-349}$) and 12–333 of *Mj*Get3 (*Mj*Get3$_{12-333}$). For co-expression, *Mj*Get3$_{12-349}$, *Tk*Get3 and *Sc*Get3 were cloned into the first multiple-cloning site (MCS) of pACYCDuet (Novagen). TA proteins (Ysy6 YBR162W-A, Sbh1 YER087C-B, Secβ (Kinch *et al*, 2002; van den Berg *et al*, 2004), SecE MJ_0371, MtrA MJ_0851, MtrB MJ_0850) were amplified from genomic DNA and cloned into pMAL-C2 (NEB) modified to contain a thrombin site between MBP and the MCS. Truncations of Sbh1 were generated by site-directed mutagenesis. Gene annotations are from KEGG (http://www.genome.jp/kegg/).

Get3 homologues were expressed in BL21-Gold(DE3) (Stratagene) in 2 × YT at 37°C for 3 h (induced at $A_{600} = 0.6$ with 0.3 mM isopropyl-β-D-thiogalactoside). Cells were pelleted, resuspended in Buffer A (50 mM Tris pH 7.5, 300 mM NaCl, 10 mM β-mercaptoethanol (β-ME)) with protease inhibitors, and lysed through an ML-110 microfluidizer (Microfluidics). Lysate was centrifuged and supernatant was passed over Ni-NTA resin (Qiagen), washed with Buffer A with 10 mM imidazole and eluted in Buffer A with 200 mM imidazole. The eluate was incubated with 2U thrombin per ml (Sigma) at room temperature (RT) while dialysing against Buffer A for 16 h. The reaction was stopped with 1 mM PMSF and passed over Ni-NTA resin to remove uncleaved product and contaminants. Flow through was purified on a Superdex 200 column (GE Healthcare) (10 mM Tris pH 7.5, 100 mM NaCl, 10 mM β-ME). Selenomethione derivatives were expressed by auto-induction media and purified by the same method as native (Studier, 2005). *Sc*Get3 tetramers were analysed for extraneous peptides by tryptic digest followed by LC/MS at the Caltech Protein/Peptide Micro-Analytical Laboratory.

Crystallization

*Mj*Get3$_{12-349}$ crystallized in the P2$_1$ form in 2 days at RT by sitting-drop vapour diffusion by mixing 1 μl of 10 mg/ml *Mj*Get3 (10 mM Tris pH 7.5, 100 mM NaCl, 10 mM β-ME, 2 mM $MgCl_2$, 2 mM ADP, 0.5 mM $AlCl_3$, 8 mM NaF) with 1 μl of reservoir solution (0.1 M Na_2SO_4 and 9% (w/v) PEG 3350). Crystals were cryoprotected with artificial mother liquor containing 20% glycerol or 17.5% sucrose and 17.5% xylitol before flash freezing in liquid N_2. Selenomethione crystals were obtained in the same way. *Mj*Get3$_{12-333}$ crystallized in the P2 form after 1 day in drops of 1 μl of 10 mg/ml *Mj*Get3 (10 mM Tris pH 7.5, 100 mM NaCl, 10 mM β-ME, 2 mM $MgCl_2$, 2 mM ADP) and 1 μl 0.2 M Na_2SO_4 and 10% (w/v) PEG 3350 and cryoprotected with 20% ethylene glycol.

Data collection, structure solution and refinement

All native data collection was done at SSRL BL12-2 at a wavelength of 1.000 Å at 100 K. Selenomethionine derivative data were collected at the APS GM/CA-CAT BL23ID-D at a wavelength of 0.9795 Å at

100 K. Diffraction data were integrated with iMosflm (Leslie, 1992) and scaled with CCP4/Scala for $P2_1$ data (CCP4, 1994) and XDS for P2 data (Kabsch, 2010). MR with native $P2_1$ data was performed with CCP4/Phaser (McCoy, 2007) with an NHD from 3IBG. MR for native P2 data was by phenix.automr used a starting model of an NHD monomer from *Mj*Get3 $P2_1$. The $P2_1$ form consisted of a single *Mj*Get3 tetramer in the asymmetric unit. The P2 form contained four monomers assembled in two crystallographic tetramers aligned along the long axis of symmetry (Supplementary Figure S2). Rounds of model building and refinement were done with Coot (Emsley and Cowtan, 2004) and phenix.refine (Adams *et al*, 2002). Global NCS was used in the $P2_1$ refinement with weights calculated in phenix.refine. Residues that we were not able to resolve in both crystal forms varied between subunits but included the N-terminus (2–23/25), the loop between helices 8 and 9 (202–209) and the C-terminus (333/334–349). In P2, we could see density in chain A that is consistent with helix 1 of chain A in $P2_1$ but were unable to build into it with confidence. Additionally, we were unable to convincingly model density in the NBP near the magnesium and aluminium fluoride-binding site. As we see it in both forms it could be a sulphate ion; however, we do not have direct evidence for this. TLS (translation/libration/screw) vibrational motions were calculated using the TLSMD server (Painter and Merritt, 2006a, b) and used in the refinement. After initial modelling and refinement the P2 model then refined against data corrected by the Diffraction Anisotropy Server (Strong *et al*, 2006) limiting the resolution in directions a* to 3.3 Å, b* to 2.9 Å and c* to 3.4 Å. The final $P2_1$ model had an R_{work} of 25.1% and an R_{free} of 28.6% with residues in the Ramachandran plot in 97.5% preferred, 2.5% allowed and 0.0% in the disallowed and restricted regions. The final P2 model had an R_{work} of 27.0% and an R_{free} of 29.6% with residues in the Ramachandran plot in 96.2% preferred, 3.8% allowed and 0.0% in the disallowed and restricted regions. Ramachandran statistics are taken from PHENIX. All structure figures were made using Pymol (Delano, 1998) except for Figures 2A and 5 and Supplementary Figure S5, which were made using UCSF Chimera (Pettersen *et al*, 2004).

Pull downs

*Sc*Get3, *Mj*Get3 or *Tk*Get3 and TA proteins were co-expressed in BL21-Gold(DE3) *E. coli*. Soluble complexes were purified in two steps using amylose resin (NEB) and Ni-NTA resin. *Sc*Get3 complexes with Sbh1 truncations were purified by the same method followed by incubation with 2 U of thrombin per ml at RT overnight.

Microsome insertion assay

The *Sc*Get3 complex with MBP–Sbh1-op was co-expressed and purified as for the pull downs followed by SEC on a Superdex 200 10/300. MBP–Sbh1-op was purified using amylose resin. Microsomes from WT and Δ*get3* strains were prepared as in Schuldiner *et al* (2008). Purified complex or MBP–Sbh1-op and Δ*get3* microsomes were used for the insertion assay using the conditions reported in Bozkurt *et al* (2009).

S. cerevisiae translation extracts were prepared essentially as in Wu *et al* (2007), and included an additional centrifugation step at 49 000 r.p.m. in an Sw55Ti for 30 min after the low speed centrifugation step (following cell lysis). *In-vitro* translations were carried out as in Wu *et al* (2007). Translation reactions were performed with Δ*get3* extracts in a 10-μl scale with 10 μCi [35S] methionine in the presence of recombinant Get3 (concentrations indicated on gel). To assay post-translational TA protein insertion, the following was added after 30 min: 1 mM cycloheximide, 1 μl energy mix (8.3 mM ATP, 1.7 mM GTP, 200 mM creatine phosphate, 600 mM KOAc, 10 mM MgOAC), and 0.006 U/μl YRMs (WT or Δ*get3*).

SEC-MALLS

Purified proteins were run on a Shodex KW-804 column (10 mM Tris pH 7.5, 100 mM NaCl, 10 mM β-ME) with MALLS data collected on a DAWN HELEOS and Optilab rEX detector (Wyatt). Data were processed using ASTRA (Wyatt) software.

SAXS

Purified *Mj*$Get3_{12–349}$ and *Sc*Get3 complexes with TA substrates were dialysed against 20 mM Tris pH 7.5, 250 mM NaCl and 10 mM β-ME. *Mj*$Get3_{12–349}$ samples were also prepared by dialysis against the same buffer containing 2 mM $MgCl_2$, 2 mM ADP, 8 mM NaF, 0.5 mM $AlCl_3$ and 1 mM $ZnSO_4$. Solution SAXS experiments were done at SSRL BL4-2 at RT. SAXS diffraction images were processed using SASTool and PRIMUS (Konarev *et al*, 2003), data were analysed with PRIMUS/autorg/autoporod, particle distance functions were generated with GNOM/autognom (Svergun, 1992) and *ab initio* shape determination was done with DAMMIN (Svergun, 1999).

ICP-MS measurement

The zinc and magnesium occupancy was quantified by ICP-MS similarly to Yamagata *et al* (2010). Samples were measured using an HP-4500 ICP-MS (Agilent Technologies). The concentration of the Get3/TA complex was measured at between 3.24 and 3.48 μM assuming a Get3 tetramer bound to either one or four substrates. The zinc concentration was measured at 5.66 μM consistent with two ions per tetramer. The magnesium concentration was measured at 1.9 μM indicating less than one ion per tetramer, consistent with no detectable amount of magnesium binding, which is required for nucleotide binding.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Grace Wu for help generating the TA protein constructs and Shu-ou Shan and Raymond Liu for help in microsome preparation. We thank DC Rees, JW Chartron, S Tanaka, HB Gristick and SO Shan for critical reading of the manuscript. We thank members of the laboratory for support and useful discussions. *T. kodakaraensis* genomic DNA was kindly provided by Tom Santangelo (Ohio State). We thank Graeme Card, Ana Gonzalez and Michael Soltice for help with data collection at SSRL BL12-2, Justin Chartron, Michael Becker and Craig Ogata for help with data collection at APS GM/CA CAT BL23ID-D, and Tsutomu Matsui and Hiro Tsuruta for help with data collection and processing at the bioSAXS SSRL BL4-2. ICP-MS instrumentation was provided by the Environmental Analysis Center at the Caltech and we were aided by Nathan Dalleska. We are grateful to Gordon and Betty Moore for support of the Molecular Observatory at Caltech. Operations at SSRL and APS are supported by the US DOE and NIH. This work is funded by grants to WMC including, the Searle Scholar program, a Burroughs-Wellcome Fund Career Award for the Biological Sciences and a National Institutes of Health Grant R01 GM097572. The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank, http://www.pdb.org (PDB ID codes 3UG6 and 3UG7 for the $P2_1$ and P2 forms, respectively).

Author contributions: CJMS performed all of the experiments except the *in-vitro* translation insertion assay. MER provided microsome membranes and performed the *in-vitro* translation insertion assay. CJMS and WMC designed experiments and wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Cryst D* **58:** 1948–1954

Battle A, Jonikas MC, Walter P, Weissman JS, Koller D (2010) Automated identification of pathways from quantitative genetic interaction data. *Mol Syst Biol* **6:** 379

Beilharz T, Egan B, Silver PA, Hofmann K, Lithgow T (2003) Bipartite signals mediate subcellular targeting of tail-anchored

membrane proteins in Saccharomyces cerevisiae. *J Biol Chem* **278:** 8219–8223

Bernstein HD, Poritz MA, Strub K, Hoben PJ, Brenner S, Walter P (1989) Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. *Nature* **340:** 482–486

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

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

Boskovic J, Soler-Mira A, García-Cantalejo JM, Ballesta JP, Jiménez A, Remacha M (1996) The sequence of a 16 691 bp segment of Saccharomyces cerevisiae chromosome IV identifies the DUN1, PMT1, PMT5, SRP14 and DPR1 genes, and five new open reading frames. *Yeast* **12:** 1377–1384

Bozkurt G, Stjepanovic G, Vilardi F, Amlacher S, Wild K, Bange G, Favaloro V, Rippe K, Hurt E, Dobberstein B, Sinning I (2009) Structural insights into tail-anchored protein binding and membrane insertion by Get3. *Proc Natl Acad Sci USA* **106:** 21131–21136

CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Cryst D* **50:** 760–763

Chartron JW, Gonzalez GM, Clemons Jr WM (2011) A structural model of SGT2 and its interactions with chaperones and Get4/Get5. *J Biol Chem* **286:** 34325–34334

Chartron JW, Suloway CJM, Zaslaver M, Clemons Jr WM (2010) Structural characterization of the Get4/Get5 complex and its interaction with Get3. *Proc Natl Acad Sci USA* **107:** 12127–12132

Ching MH, Kaur P, Karkaria CE, Steiner RF, Rosen BP (1991) Substrate-induced dimerization of the ArsA protein, the catalytic component of an anion-translocating ATPase. *J Biol Chem* **266:** 2327–2332

Delano WL (1998) *The PyMOL Molecular Graphics System*, Vol. 2008, Palo Alto, California, USA: DeLano Scientific

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Cryst D* **60:** 2126–2132

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

Favaloro V, Vilardi F, Schlecht R, Mayer MP, Dobberstein B (2010) Asna1/TRC40-mediated membrane insertion of tail-anchored proteins. *J Cell Sci* **123:** 1522–1530

Hayashi I, Oyama T, Morikawa K (2001) Structural and functional studies of MinD ATPase: implications for the molecular recognition of the bacterial cell division apparatus. *EMBO J* **20:** 1819–1828

Hu J, Li J, Qian X, Denic V, Sha B (2009) The crystal structures of yeast get3 suggest a mechanism for tail-anchored protein membrane insertion. *PLoS ONE* **4:** e8061

Kabsch W (2010) XDS. *Acta Cryst D* **66:** 125–132

Kalbfleisch T, Cambon A, Wattenberg BW (2007) A bioinformatics approach to identifying tail-anchored proteins in the human genome. *Traffic* **8:** 1687–1694

Keenan RJ, Freymann DM, Walter P, Stroud RM (1998) Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell* **94:** 181–191

Kinch LN, Saier MH, Grishin NV (2002) Sec61beta—a component of the archaeal protein secretory system. *Trends Biochem Sci* **27:** 170–171

Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst* **36:** 1277–1282

Koonin EV (1993) A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J Mol Biol* **229:** 1165–1174

Kriechbaumer V, Shaw R, Mukherjee J, Bowsher CG, Harrison A-M, Abell BM (2009) Subcellular distribution of tail-anchored proteins in Arabidopsis. *Traffic* **10:** 1753–1764

Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372:** 774–797

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

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

Leipe DD, Wolf YI, Koonin EV, Aravind L (2002) Classification and evolution of P-loop GTPases and related ATPases. *J Mol Biol* **317:** 41–72

Leonard T, Butler P, Löwe J (2005) Bacterial chromosome segregation: structure and DNA binding of the Soj dimer—a conserved biological switch. *EMBO J* **24:** 270–282

Leslie AG (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**

Leznicki P, Warwicker J, High S (2011) A biochemical analysis of the constraints of tail-anchored protein biogenesis. *Biochem J* **436:** 719–727

Liou S-T, Cheng M-Y, Wang C (2007) SGT2 and MDY2 interact with molecular chaperone YDJ1 in Saccharomyces cerevisiae. *Cell Stress Chaperones* **12:** 59–70

Liou S-T, Wang C (2005) Small glutamine-rich tetratricopeptide repeat-containing protein is composed of three structural units with distinct functions. *Arch Biochem Biophys* **435:** 253–263

Mariappan M, Li X, Stefanovic S, Sharma A, Mateja A, Keenan RJ, Hegde RS (2010) A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* **466:** 1120–1124

Mariappan M, Mateja A, Dobosz M, Bove E, Hegde RS, Keenan RJ (2011) The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature* **477:** 61–66

Mateja A, Szlachcic A, Downing ME, Dobosz M, Mariappan M, Hegde RS, Keenan RJ (2009) The structural basis of tail-anchored membrane protein recognition by Get3. *Nature* **461:** 361–366

McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Cryst D* **63:** 32–41

Metz J, Wächter A, Schmidt B, Bujnicki JM, Schwappach B (2006) The yeast Arr4p ATPase binds the chloride transporter Gef1p when copper is available in the cytosol. *J Biol Chem* **281:** 410–417

Painter J, Merritt EA (2006a) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Cryst D* **62:** 439–450

Painter J, Merritt EA (2006b) TLSMD web server for the generation of multi-group TLS models. *J Appl Cryst* **39:** 109–111

Pedrazzini E (2009) Tail-anchored proteins in plants. *J Plant Biol* **52:** 88–101

Pettersen E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, Ferrin T (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605–1612

Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40:** 191–285

Rabu C, Schmid V, Schwappach B, High S (2009) Biogenesis of tail-anchored proteins: the beginning for the end? *J Cell Sci* **122:** 3605–3612

Santangelo TJ, Cubonová L, James CL, Reeve JN (2007) TFB1 or TFB2 is sufficient for Thermococcus kodakaraensis viability and for basal transcription *in vitro*. *J Mol Biol* **367:** 344–357

Schindelin H, Kisker C, Schlessman JL, Howard JB, Rees DC (1997) Structure of ADP × AIF4(−)-stabilized nitrogenase complex and its implications for signal transduction. *Nature* **387:** 370–376

Schuldiner M, Metz J, Schmid V, Denic V, Rakwalska M, Schmitt HD, Schwappach B, Weissman JS (2008) The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* **134:** 634–645

Shan S-O, Walter P (2005) Co-translational protein targeting by the signal recognition particle. *FEBS Lett* **579:** 921–926

Sherrill J, Mariappan M, Dominik P, Hegde RS, Keenan RS (2011) A conserved archaeal pathway for tail-anchored membrane protein insertion. *Traffic* **12:** 1119–1123

Simpson PJ, Schwappach B, Dohlman HG, Isaacson RL (2010) Structures of Get3, Get4, and Get5 provide new models for TA membrane protein targeting. *Structure* **18:** 897–902

Sprang SR (1997) G protein mechanisms: insights from structural analysis. *Annu Rev Biochem* **66:** 639–678

Stefanovic S, Hegde RS (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, Löhr F, Bernhard F, Denic V, Dötsch V, Sinning I (2011)

Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. *Science* **333:** 758–762

Strong M, Sawaya MR, Wang S, Phillips M, Cascio D, Eisenberg D (2006) Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* **103:** 8060–8065

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41:** 207–234

Suloway CJM, Chartron JW, Zaslaver M, Clemons WM (2009) Model for eukaryotic tail-anchored protein binding based on the structure of Get3. *Proc Natl Acad Sci USA* **106:** 14849–14854

Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst* **25:** 495–503

Svergun DI (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* **76:** 2879–2886

Tobaben S, Varoqueaux F, Brose N, Stahl B, Meyer G (2003) A brain-specific isoform of small glutamine-rich tetratricopeptide repeat-containing protein binds to Hsc70 and the cysteine string protein. *J Biol Chem* **278:** 38376–38383

Van den Berg B, Clemons WM, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA (2004) X-ray structure of a protein-conducting channel. *Nature* **427:** 36–44

Walmsley AR, Zhou T, Borges-Walmsley MI, Rosen BP (1999) The ATPase mechanism of ArsA, the catalytic subunit of the arsenite pump. *J Biol Chem* **274:** 16153–16161

Wang F, Brown EC, Mak G, Zhuang J, Denic V (2010) A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Mol Cell* **40:** 159–171

Wu C, Amrani N, Jacobson A, Sachs MS (2007) The use of fungal *in vitro* systems for studying translational regulation. *Methods Enzymol* **429:** 203–225

Wu CC, Yates JR (2003) The application of mass spectrometry to membrane proteomics. *Nat Biotech* **21:** 262–267

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