# INVESTIGATING THE ROLE OF THE GET3-GET4/GET5 INTERACTION DURING TAIL-ANCHOR PROTEIN TARGETING

Thesis by
Harry Benjamin Gristick

In Partial Fulfillment of the Requirements for the degree
of
Doctor of Philosophy


CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California
2015
(Defended May 29, 2015)

# ACKNOWLEDGEMENTS

First, I need to thank my advisor William "Bil" Clemons for his support and mentorship. We didn't always agree (or never) on scientific matters, but we were still able to drive progress forward in our field. He always forced me to ask the right questions and to think critically about new results, and I am certainly a better scientist because of him. Outside of the lab, I formed a close personal relationship with Bil and his family that were highlighted by many events: birthday parties for his children, baseball games with Kuba, and backstage passes at Bruce Springsteen concerts. I will always have fond memories of this time in my life, and thank Bil for treating me as if I were his own family.

Of course I need to thank the members of my committee: Professor Shu-ou Shan, who taught me theoretical and practical kinetics and our close collaborator; Professor Doug Rees, who has offered me priceless advice in and out of science and who happens to be the best shortstop in the Caltech slow-pitch softball league; Professor David Chan, who also hired me as a research technician before I started graduate school. David also happens to be the scientist I admire most in the world. I'm not aware of any other investigators who can run a lab at the highest level that uses techniques ranging from cell biology to biochemistry to structural biology to mouse models.

I was persuaded to pursue a career in research science by a number of professors at IUP: the late Gary Ciskowski, my advisor who constantly kept me on my toes with his razor-sharp wit; Andrew Browe, my undergraduate research mentor who made me realize my love of human anatomy and physiology; Robert Hinrichsen, whose experiences at the Fred Hutchinsen Cancer Research Center with Dr. Leland H. Hartwell were awe-inspiring;

and Allan Andrew, for his brilliant lectures on immunology and medical microbiology. I also need to thank Professor Michael Root and Dr. Mark Jones for hiring me as a research technician straight from college and introducing me to structural biology.

I thank the members of the Clemons lab, starting with the GET team: Dr. Christian Suloway, Dr. Justin Chartron, Ma'ayan Zaslaver, Jee-Young Mock, and Ku-Feng Lin. Christian and Justin laid the foundation upon which our lab has built our research on protein targeting. They were with me for my first three years and taught me theoretical and practical protein crystallography, among many other things. I'd also like to thank the remaining members of the Clemons lab who were present during this work: Dr. Sureshkumar Ramasamy, Dr. Shiho Tanaka, Dr. Axel Müller, and Dr. Kyoung-Soon Jang, Hyun-gi Yun, Lada Klaic, and Stephen Marshall.

I would like to thank specific members of the Shan lab at Caltech: Mike Rome, Meera Rao, and Un-Seng Chio. Their work was instrumental in generating the model we now have for the Get3-Get4/5 interaction. I am especially grateful to my collaborators, Mike and Meera, whose excellent biochemistry was essential to this thesis and to whom I am indebted.

I would like to acknowledge another graduate student, Rachel Galimidi, my fiancé and all around best friend. Rachel is an extremely intelligent, beautiful, and talented person in and out of science, and makes a positive impression on everyone she meets. Regardless of where our careers take us, I will always be thankful I came to Caltech because I met you.

Finally, I'd like to thank my parents, Paul and Beverly Gristick, to whom I dedicate this thesis. Their love and support throughout my life allowed me to pursue this difficult career path, and for that, I am forever grateful.

# ABSTRACT

The proper targeting of membrane proteins is essential to the viability of all cells. Tail-anchored (TA) proteins, defined as having a single transmembrane helix at their C-terminus, are post-translationally targeted to the endoplasmic reticulum (ER) membrane by the GET pathway (Guided Entry of TA proteins). In the yeast pathway, the handover of TA substrates is mediated by the heterotetrameric Get4/Get5 (Get4/5) complex, which tethers the co-chaperone Sgt2 to the central targeting factor, the Get3 ATPase. Although binding of Get4/5 to Get3 is critical for efficient TA targeting, the mechanisms by which Get4 regulates Get3 are unknown. To understand the molecular basis of Get4 function, we used a combination of structural biology, biochemistry, and cell biology. Get4/5 binds across the Get3 dimer interface, in an orientation only compatible with a closed Get3, providing insight into the role of nucleotide in complex formation. Additionally, this structure reveals two functionally distinct binding interfaces for anchoring and ATPase regulation, and loss of the regulatory interface leads to strong defects *in vitro* and *in vivo*. Additional crystal structures of the Get3-Get4/5 complex give rise to an alternate conformation, which represents an initial binding interaction mediated by electrostatics that facilitates the rate of subsequent inhibited complex formation. This interface is supported by an in-depth kinetic analysis of the Get3-Get4/5 interaction confirming the two-step complex formation. These results allow us to generate a refined model for Get4/5 function in TA targeting.

# TABLE OF CONTENTS

**Acknowledgements** ..... iii
**Abstract** ..... vi
**Table of Contents** ..... vii
**List of Figures and Tables** ..... ix

**Chapter 1**
*Introduction*
Post-translational protein targeting ..... 1
Discovery of the GET pathway ..... 2
Characterizing the GET pathway ..... 3

**Chapter 2**
*Crystal structure of ATP-bound Get3-Get4/Get5 complex reveals regulation of Get3 by Get4*
Abstract ..... 7
Introduction ..... 8
Results ..... 11
Discussion ..... 17
Acknowledgments ..... 19
Methods ..... 20
Figures ..... 26
Tables ..... 33
Supplementary Figures ..... 34
Supplementary Tables ..... 43

**Chapter 3**
*Molecular details of a Get3-Get4/Get5 intermediate complex*
Abstract ..... 44
Introduction ..... 45
Results ..... 47
Discussion ..... 55
Acknowledgments ..... 57
Methods ..... 58
Figures ..... 65
Tables ..... 76

**Chapter 4**
*Conclusions*
Results/Discussion ..... 77
Methods ..... 81

Figures ..... 84

**Appendix A**

*Towards purification of the Get1/Get2 membrane protein complex*

Introduction ..... 86
Results/Discussion ..... 88
Methods ..... 93
Figures ..... 95

**Appendix B**

*Characterization of an Archaeal Get3 knockout*

Introduction ..... 110
Results/Discussion ..... 112
Methods ..... 114
Figures ..... 115

**Bibliography** ..... 117

# LIST OF FIGURES AND TABLES

## Chapter 2

**Figure 2.1.** Get4/5 prefers ATP-bound state of Get3 ..... 26
**Figure 2.2.** 5.4 Å crystal structure of an ATP-bound Get3-Get4/5 complex ..... 27
**Figure 2.3.** Get3-Get4 binding interfaces ..... 28
**Figure 2.4.** Get4/5 regulates Get3 ATPase activity ..... 30
**Figure 2.5.** A working model for Get4/5 ..... 32

**Table 2.1.** Data collection and refinement statistics ..... 33

**Supplementary Figure 2.1.** Representative ITC isotherms for nucleotide-dependent complex formation ..... 34
**Supplementary Figure 2.2.** Surface properties of Get3 and Get4 ..... 36
**Supplementary Figure 2.3.** Stereo views of the Get3D-Get4/5N interface ..... 38
**Supplementary Figure 2.4.** TA targeting assay ..... 39
**Supplementary Figure 2.5.** Tetramers of Get3 ..... 41

**Supplementary Table 2.1.** Summary of ATPase data ..... 43

## Chapter 3

**Figure 3.1.** Get4/5 undergoes a conformational change upon binding Get3 ..... 65
**Figure 3.2.** Modeling of experimental kinetic data ..... 67
**Figure 3.3.** Purification of the Get3-Get4/5N complex ..... 68
**Figure 3.4.** Crystal structures of a Get3-Get4/5 intermediate complex ..... 69
**Figure 3.5.** Surface properties of Get3 and Get4 ..... 71
**Figure 3.6.** Structural overlay comparing the intermediate and inhibited interfaces ..... 72
**Figure 3.7.** Get3 D263A is defective for complex formation ..... 73
**Figure 3.8.** Get3 binds to one-half of the Get4/5 heterotetramer ..... 74

**Table 3.1.** Summary of crystallographic data ..... 76

## Chapter 4

**Figure 4.1.** Purification of full-length Get3-Get4/5 complex ..... 84
**Figure 4.2.** Crystallization of full-length Get3-Get4/5 Complex ..... 85

## Appendix A

**Figure A.1.** Purification of the individual components Get1 and Get2 ..... 95

**Figure A.2.** Purification of the individual components Get1, Get2, and $Get2_{TM}$ used in reconstitution experiments..... 96
**Figure A.3.** Detergent screen for Get1/2 complex formation..... 97
**Figure A.4.** Scaled up detergent screen with LDAO, Fos Choline 12, and Anzergent 3-12 for Get1/2 complex formation..... 99
**Figure A.5.** Co-expression of Get1/2..... 100
**Figure A.6.** Co-expression of MBP-$Get2_{TM}$/10xHis-Get1 in pET-DUET from *Sc*, *Ca*, and *Af*..... 102
**Figure A.7.** Co-expression of 10xHis-Get2/Get1 in pET-DUET from *Sc* and *Af*... 104
**Figure A.8.** Co-expression of 10xHis-Get2/Get1 with untagged Get3 from *Sc*..... 106
**Figure A.9.** Co-expression of 10xHis-Get2/Get1 with MBP-Get3 from *Sc* and *Af*..... 107
**Figure A.10.** Purification of MBP-Get3 from *Sc*..... 109

## Appendix B

**Figure B.1.** Generating knockouts of Get3 homologues in *Haloferax volcanii*..... 115

*Chapter 1*

# INTRODUCTION

In eukaryotes, membrane proteins account for ~30% of the proteome and are involved in many essential functions, including cell signaling, protein transport, and membrane fusion [1,2]. The vast majority of these proteins are initially targeted to the endoplasmic reticulum by the signal recognition particle (SRP) pathway. The SRP recognizes the first hydrophobic transmembrane domain (TMD) as it emerges from the ribosome exit tunnel, and then shuttles the ribosome-nascent chain (RNC) to the ER, where the SRP is recognized by its receptor. Docking of the RNC onto the Sec translocon then ensues, allowing the co-translational insertion of the nascent membrane protein into the ER where it adopts the correct architecture and, if necessary, enters the secretory pathway and is transported to the appropriate destination within the cell. This pathway has been extensively characterized [3,4] and is an example of highly evolved, complex biological macromolecules working in concert to perform a difficult task.

## Post-translational protein targeting

Although the SRP pathway can target an extraordinarily wide range of membrane proteins, there is one class of proteins that cannot access the co-translational targeting machinery. Known as tail-anchored (TA) proteins, they contain a single TMD at their extreme carboxyl-terminus (C-terminus), and are found in nearly all membrane-bound organelles within the cell. As in the SRP pathway, TA proteins are first targeted to the ER, and then enter the secretory pathway where they are trafficked to their final destination. TA proteins account for ~1-2% of the eukaryotic proteome, and are involved in a wide range of

functions including vesicle fusion (SNAREs), apoptosis (Bcl-2), and protein translocation (Sec61ß) [5,6].

Based on experiments using cytochrome $b_5$, it was initially believed that TA proteins did not require any cellular machinery and spontaneously inserted into membranes [7]. However, in 1993, Kutay et al were the first to acknowledge that TA proteins would be unable to access the SRP pathway [8]. They reasoned that because of their topological constraints, the TMD emerges from the ribosome only after protein synthesis is complete, and therefore must be targeted to the ER post-translationally. Shortly after their initial report, Kutay et al demonstrated that the SRP was not essential for TA targeting, but instead was ATP-dependent and required a membrane component at the ER distinct from the Sec translocon [9].

**Discovery of the GET pathway**

Following these initial experiments, the specific factors involved in TA targeting remained unknown for over a decade. Then, in 2007 Stefanovic et al used a cross-linking reagent to an *in vitro* translation system prepared from rabbit reticulocytes and managed to isolate a protein factor from a lysate that specifically bound TA proteins and facilitated their targeting *in vitro*[10]. Importantly, TA targeting by this factor was dependent on both ATP and the presence of a membrane component within the ER. This newly discovered factor, a 40 kDa cytosolic ATPase, was named the TMD Recognition Complex of 40 kDa (TRC40), or Guided Entry of Tail-Anchored proteins (Get3) [10,11].

Originally annotated as ArsA or Arr4 due to its homology to the bacterial arsenite transport protein, homologues of TRC40 were identified in other organisms due to the high level of sequence conservation. Even before a link to TA targeting was known, multiple genetic studies provided evidence that TRC40 homologues were essential for normal

cellular functioning. Initial experiments with *S. cerevisiae* demonstrated knockouts were sensitive to increased temperatures and metal tolerance, much like the bacterial ArsA [12,13]. Growth phenotypes were also observed in higher eukaryotes. A *Caenorhabditis elegans* (*C. elegans*) knockout arrested development at the L1 stage [14], while a mouse knockout was embryonic lethal [15]. Downregulation of TRC40 in human melanoma cells increased their sensitivity to arsenite and cisplatinin, a cancer therapeutic, providing a link to human disease [16].

Prior to the discovery of TRC40, genetic studies in the budding yeast *S. cerevisiae* established a link between *Arr4* and a complex of two resident ER membrane proteins (*MDM39* and *RMD7*) [17]. Furthermore, these factors were implicated in Golgi to ER traffic (GET), and renamed Get1 (*MDM39*), Get2 (*RMD7*), and Get3 (*Arr4*). Subsequent studies demonstrated that Get3 formed a physical interaction with the protein complexes Get1/Get2 (Get1/2) [18,19] and Get4/Get5 [20], and that all five functioned directly in TA targeting. This led to the acronym GET to be redefined as the Guided Entry of Tail-Anchored proteins. Evidence from two separate genetic analyses suggested that an additional factor, the heat shock protein (HSP) co-chaperone Sgt2, also played a role in TA targeting, and most likely functioned upstream of Get4/5 [21,22].

**Characterizing the GET pathway**

Extensive structural characterization of Get3, the central TA targeting factor, demonstrated that it undergoes ATP-dependent conformational changes from an open to closed form required for capturing the TA substrate [23-28]. Once formed, the Get3-TA complex is then localized to the ER by the membrane proteins Get1/2, which stimulate release of the TA protein and subsequent insertion into the ER membrane [29-32]. Crystal structures of Get3 in complex with the soluble domains of Get1 and Get2 revealed

overlapping binding sites on Get3, suggesting that Get1 and Get2 bind independently *in vivo* [29,30]. This was further supported by the observation that the soluble domain of Get2 crystallized with both the open and closed forms of Get3, whereas the soluble domain of Get1 crystallized with only the open form. Rome et al (2014) subsequently demonstrated that Get2 has the highest affinity to ADP-bound Get3, whereas Get1 prefers apo-Get3, suggesting that Get2 binding precedes Get1. In fluorescence microscopy experiments with yeast Get3 or Get1/2 knockout strains, the majority of TA proteins were either mistargeted (Get3), or formed aggregates in the cytosol (Get1/2), underscoring the importance of these factors *in vivo* [19].

Upstream of Get3 is the multi-domain co-chaperone Sgt2 that specifically binds the TA, which is the first committed step in TA targeting [22,33,34]. Sgt2 forms an extended protein in solution and is comprised of three domains: an amino-terminal (N-terminal) dimerization domain, a middle tetratricopeptide repeat (TPR) domain, and C-terminal glutamine-rich (Q) domain. NMR structures of the dimerization domain by itself and in complex with the ubiquitin-like (Ubl) domain of Get5 were recently published [35]. These structures, along with a comprehensive binding analysis, revealed that this interaction is mainly electrostatic and occurs on fast timescales. Disruption of this interaction in yeast leads to similar growth defects seen with the Get3 knockout [36,37]. Structural studies of the TPR domain provided a model for how Sgt2 might interact with upstream chaperones [33]. The Q domain, which binds the TMD of the nascent TA protein, has so far evaded both NMR and protein crystallography due its inherent flexibility.

Efficient delivery of a TA substrate to Get3 requires the hetero-tetrameric Get4/5 complex that provides the link between Sgt2 and Get3 [21,22,34,38]. Structural studies of Get4 and the N-terminal domain of Get5 revealed that Get4 is an alpha-helical repeat protein, with the N-terminus of Get5 wrapping around its C-terminus [36,39]. Size-exclusion

chromatography (SEC) demonstrated that formation of the Get4/5 heterotetramer was dependent on dimerization of the Get5 C-terminus [36]. In SAXS reconstructions, the full-length Get4/5 complex forms an extended structure where Get4 flanks the Get5 Ubl domain and central Get5 C-terminal homodimerization domain [33,36,40].

More recent work has expanded on the role of Get4/5 in TA targeting beyond acting as a simple bridge. In addition to preferentially recognizing a nucleotide bound Get3, Get4 inhibits Get3 ATP hydrolysis [41]. TA binding is the presumptive trigger for hydrolysis [41], thus Get4 helps to stabilize Get3 in a conformation competent for TA binding. Initial biochemical and genetic evidence implicated the N-terminal face of Get4 in Get3 binding [36,42], but additional evidence was lacking until recently. This report showed that Get4 bound to Get3 on near-diffusion limited timescales, and that one phase of the association was salt-dependent, suggesting that electrostatic interactions play a role in complex formation [43].

Although our understanding of the pathway has increased in the last decade, there are many questions that remain unanswered. One important question is how Get4 regulates Get3 activity, i.e., what is the structural basis of Get4 function. The work presented here answers these questions and gives insight into the role of the Get3-Get4/5 interaction during TA protein targeting. Chapter 2 presents the crystal structure of an ATP-bound Get3-Get4/5 complex. This structure provides insight into the role of nucleotide in complex formation, where Get4 binds to both monomers of Get3 in an orientation only compatible with a closed Get3 and revealed two functionally distinct binding interfaces for anchoring and ATPase regulation. Chapter 3 presents two additional crystal structures of the Get3-Get4/5 complex in an alternate conformation. These structures represent an initial binding interaction mediated by electrostatics that facilitates the rate of subsequent inhibited complex formation. This is supported by kinetic analysis of Get3-Get4/5 complex

formation confirming the two-step complex formation. Chapter 4 presents unpublished work towards solving the crystal structure of a full-length Get3-Get4/5 complex, and concludes the thesis with a review of the key findings and insight into possible future experiments that may help further our understanding of the Get3-Get4/5 interaction and the roles it plays in TA targeting.

*C h a p t e r 2*

# CRYSTAL STRUCTURE OF ATP-BOUND GET3-GET4-GET5 COMPLEX REVEALS REGULATION OF GET3 BY GET4

**Abstract**

Correct localization of membrane proteins is essential to all cells. Chaperone cascades coordinate the capture and handover of substrate proteins from the ribosome to their target membrane; yet the mechanistic and structural details of these processes remain unclear. Here we investigate the conserved GET pathway, in which the Get4/Get5 complex mediates the handover of tail-anchor (TA) substrates from the co-chaperone Sgt2 to the Get3 ATPase, the central targeting factor. We present a crystal structure of a yeast Get3-Get4/Get5 complex in an ATP-bound state, and show how Get4 primes Get3 into the optimal configuration for substrate capture. Structure-guided biochemical analyses demonstrate that Get4-mediated regulation of ATP hydrolysis by Get3 is essential to efficient TA protein targeting. Analogous regulation of other chaperones or targeting factors could provide a general mechanism for ensuring effective substrate capture during protein biogenesis.

*Adapted from*

Gristick, H.B., Rao, M., Chartron, J. W., Rome, M. E., Shan, S. O., and Clemons, W.M., Jr. Crystal structure of ATP-bound Get3-Get4-Get5 complex reveals regulation of Get3 by Get4. *Nat Struct Mol Biol* **21**, 437-42 (2014).

## Introduction

In eukaryotes, the proper targeting of membrane proteins is a significant challenge for the cell to overcome [44]. Integral membrane proteins contain hydrophobic transmembrane domains (TMD), which must be protected from the aqueous cytosolic environment prior to integration into the appropriate membrane. For the majority of membrane proteins targeted to the endoplasmic reticulum (ER), this is accomplished by the signal recognition particle (SRP), typically binding the initial TMD, or signal anchor, as it emerges from the ribosome and targets it to the ER for co-translational insertion. Exceptions are the ubiquitous tail-anchor (TA) proteins, defined topologically by a single transmembrane domain near the C-terminus, which are unable to access the SRP pathway and must be targeted to the ER post-translationally [8,9].

Found in most cellular membranes, TA proteins are targeted to either the ER or mitochondria. For the latter, a dedicated pathway has not been identified [45]. Those destined for other organelles are initially targeted to the ER, and then subsequently trafficked to the appropriate membrane [9]. Examples include many essential proteins such as SNAREs (vesicle fusion), Bcl-2 (apoptosis), and Sec61γ (protein translocation machinery) [46]. A need for specific targeting of ER destined TA proteins was first conceptualized nearly two decades ago [8] and the cellular factors responsible have now been identified. The central targeting factor was identified biochemically in a mammalian system as the cytosolic ATPase TRC40, demonstrated to bind the TA and deliver it specifically to the ER [10,11]. Previous genetic experiments involving the yeast homolog, Get3, could then be linked to TA targeting, providing a route for studying this process [17]. This was followed by a series of results that characterized the new pathway, termed GET (Guided Entry of TA proteins) [19,20,22]. In yeast, this pathway consists of six proteins, Get1-5 and Sgt2, all with homologs in higher eukaryotes (Reviewed in Ref. 12).

Structural characterization of Get3, the central TA targeting factor, demonstrates that it undergoes ATP-dependent conformational changes from a apo 'open' to an ATP-bound 'closed' form required for capturing the TA substrate [23-27,47]. Deletion of Get3 in yeast leads to a buildup of mislocalized cytosolic TA proteins [19] and is embryonic lethal in mice [15]. The Get3-TA complex is recruited to the ER by the membrane proteins Get1-Get2 (Get1-2) [29-31]. These stimulate release of the TA protein and subsequent insertion into the ER membrane, although the specifics of this mechanism are, as yet, unknown. Upstream of Get3 is the multi-domain Hsp70 or Hsp90 co-chaperone Sgt2 [22,33,34] that specifically binds the TA, the first committed step in TA targeting, followed by hand-over to Get3 [34].

Efficient delivery of a TA substrate to Get3 requires the hetero-tetrameric Get4-Get5 (Get4/5) complex that provides the link between Sgt2 and Get3 [21,22,34]. Structural studies of Get4 and the N-terminal domain of Get5, also called Mdy2, revealed that Get4 is an alpha-helical repeat protein, with the N-terminus of Get5 wrapping around its C-terminus [36,39]. Biochemical and genetic evidence implicated the N-terminal face of Get4 in Get3 binding at an interface that shared commonalities with the binding sites for Get1 and Get2 [36,39]. In SAXS reconstructions, the full-length Get4/5 complex forms an extended structure where Get4 flanks the Get5 ubiquitin-like domain (Ubl) and central Get5 C-terminal homodimerization domain [33,36,40]. Initial results suggested that binding of Get4 to Get3 required nucleotide [36]; however, a subsequent publication has brought this into question [42]. More recent work has expanded on the role of Get4/5 in TA targeting beyond acting as a simple bridge. In addition to preferentially recognizing a nucleotide bound Get3, Get4 inhibits Get3 ATP hydrolysis [41]. TA binding is the presumptive trigger for hydrolysis [41], and thus Get4 helps to stabilize Get3 in a conformation competent for TA binding. Since a major outstanding question is how Get4 regulates Get3 activity, we set out to understand the structural basis of Get4 function.

This report describes the 5.4 Å crystal structure of an ATP-bound Get3-Get4/5 complex from *Saccharomyces cerevisiae* (*Sc*), a combined ~160 kDa hetero-hexameric structure. Two functionally distinct binding interfaces for anchoring and ATPase regulation were found between Get3 and Get4, and were confirmed biochemically and genetically. Mutations at these interfaces demonstrated that Get4/5-mediated regulation of ATP hydrolysis by Get3 was critical for efficient TA targeting. Finally, crystallographic tetramers of Get3 are compatible with two Get3 dimers bridged by a single Get4/5 hetero-tetramer. In total, this work illustrates how Get4/5 regulates Get3, priming it for TA loading, which is a critical step in this important pathway.

## Results

### Get4/5 binds the ATP-bound state of Get3

It was first important to establish the requirements for forming a stable Get3-Get4/5 complex. Unless noted, Get3 is either wild type or contains an ATPase inactivating mutation (D57V, referred to as Get3D) that prevents ATP hydrolysis while still allowing ATP binding. Get4/5 is either the wild-type hetero-tetramer (Get4/5) or a 22 residue C-terminal truncation of Get4 (1-290) and the 54 residue N-terminal domain of Get5 (Get4/5N), similar to that used in a crystal structure of the heterodimer [36]. As observed previously [42], in low salt (10 mM NaCl) a 1:1 complex of Get3D bound to Get4/5N could be generated that was stable by size-exclusion chromatography (SEC) (**Figure 2.1A,** yellow trace). Increasing the salt concentration resulted in a loss of complex formation such that no complex could be detected at 500 mM NaCl (brown trace), suggesting an electrostatic interaction. At near physiological conditions (175 mM NaCl), a complex between Get3D-Get4/5 was disfavored (dark red trace), suggesting that additional factors were required to stabilize the complex *in vivo*.

Based on previous evidence that suggested a role for nucleotide in complex formation [36], a variety of assays were tested to confirm nucleotide stabilization of the Get3D-Get4/5N complex. SEC was performed in the presence of nucleotide at a salt concentration where the complex was disfavored (250 mM NaCl) (**Figure 2.1A-B**, orange trace). For ADP, a clear stabilization of the complex was seen (**Figure 2.1B**, green trace) while ATP resulted in the most stable complex (blue trace), consistent with previous experiments [36,41]. Additionally, pull-down experiments were quantified where tagged Get4/5N was used to precipitate Get3D with binding presented as a ratio of the two. These experiments confirmed the nucleotide-dependence as observed with SEC demonstrating a preference for ATP (**Figure 2.1C**). Finally, affinity constants ($K_d$) were measured using isothermal titration calorimetry (ITC) at a physiological ionic strength (150 mM KOAc) (**Figure 2.1D** and **Supplementary Figure 2.1**).

For an apo-Get3D, a $K_d$ could not be measured, suggesting that the affinity was less than 10 μM. In the presence of ADP, Get4/5N binds to Get3 with micromolar affinity that increases to ~500 nM with ATP. Collectively, all three assays show that Get4/5 preferentially interacts with ATP-bound Get3 at physiological ionic strength conditions.

**Architecture of the Get3-Get4/5 complex**

While it is clear that Get4 recognizes the ATP-bound state of Get3, the structural details of this interaction were missing. To understand these, a complex of ATP, Get3D and Get4/5N was crystallized and a structure was determined to 5.4 Å resolution (**Table 2.1**). In the structure, there is a 1:1 ratio of Get3D to Get4/5N. Get3 is in a 'closed' conformation, as anticipated based on ATP being bound [24], with the Get4 interaction lying across the dimer interface (**Figure 2.2**). An 'anchoring' (see below) interface (primarily to the brown Get3 in **Figure 2.2**) buries ~920 $Å^2$ surface area, while a 'regulatory' interface (primarily to the purple Get3) contributes ~400 $Å^2$ to this interaction (**Supplementary Figure 2.2**).

At the anchoring interface, Get4 α2 makes extensive contacts roughly parallel to the groove formed by Get3 α10 and α11 (**Figure 2.3A** and **Supplementary Figure 2.3**, we have included side-chains for clarity despite the resolution). This results in an interaction between the invariant residues Phe246 and Tyr250 on Get3 α10 and Tyr30 on Get4 α2 (**Figure 2.3A-B** and **Supplementary Figure 2.2B-C**). In addition to these hydrophobic contacts, a number of highly conserved charged residues are located within this interface (Get3: Glu253, Gln257, Glu258, Glu304, Asp308, and Get4: Arg37, Arg42). In contrast to α2, Get4 α1 is located on the opposite face relative to Get3, and makes fewer contacts. However, the N-terminus of Get4 α1 is tilted toward Get3, placing the conserved charged residues Lys15 and Arg19 in close proximity to Get3 α10. Previously, this group had demonstrated that residues on the N-terminal face of Get4 bound Get3 residues at a similar surface to that demonstrated for the ER receptors

Get1 and Get2 [29,30,36]. A number of these residues map to the anchoring interface (Get3: Tyr250, Glu253 and Get4: Arg19, Tyr30, Arg37, Arg42) (**Figure 2.3A** and **Supplementary Figure 2.2A**).

At the regulatory interface, located on the opposing monomer, the C-terminal end of Get4 α4 packs against the loop following Get3 α3 (**Figure 2.3A-B** and **Supplementary Figure 2.3**). This places a number of highly conserved complementary charged residues within this interface (Get3: Lys69, Lys72, Arg75 and Get4: Asp74, Glu81). Central to this interaction are the invariant residues Lys69 (Get3) and Asp74 (Get4), which are located opposite one another.

**Mutational analysis of the Get3–Get4 interface**

As noted above, the extensive Get3–Get4 interface involves varied contacts to both monomers. ITC and affinity capture assays were performed to determine which residues are essential for binding (**Figure 2.3** and **Supplementary Figure 2.1**). As expected, alanine substitution of the invariant hydrophobic residues (Get3: Phe246, Tyr250 and Get4: Tyr30) dramatically reduced the binding affinity. In addition, a number of the conserved charged residues within the anchoring interface (Get3: Glu253, Glu304 and Get4: Arg19, Arg37, Arg42) produced similar effects following their substitution. A moderate binding defect was seen with Ala substitutions of the remaining residues on Get3 α10 (Gln257, Glu258) and the loop following Get3 α11 (Asp308). Substitution of the remaining residues on Get4 (Get4: Lys15, Lys23, Tyr29, Glu31, His33, Gln34, Arg45), which are located farther away from the core of the anchoring interface and of lower sequence conservation, had little to no effect on binding.

At the regulatory interface, several highly conserved basic residues in Get3 (Lys69, Lys72, Arg75) are in close proximity to oppositely charged residues in Get4 (Asp74, Glu81). Nevertheless, substitution of these Get3 residues by Ala or Asp has little to no effect on Get4

binding (**Figure 2.3C**). Furthermore, substitution of Get4 Asp74 by Lys resulted in a marginal increase in $K_d$ ~1 μM as determined by ITC (**Figure 2.3C**). Collectively, these studies revealed that the anchoring interface is mainly responsible for the specificity of binding between Get3 and Get4.

**Regulation of Get3 nucleotide hydrolysis**

While the regulatory interface is not involved largely in binding, the highly conserved nature of residues at this interface suggests they play an alternative role in TA targeting. Recently published work demonstrated that Get4/5 binding results in inhibition of Get3 ATP hydrolysis [41]. To test if this interface plays a role, $k_{cat}$ was determined for several mutants of Get3 in the absence and presence of Get4–5 (**Supplementary Table 2.1**). The results are presented as a ratio (-Get4/5:+Get4/5) where a larger number indicates inhibition by Get4/5, e.g., wild-type Get4/5 inhibits Get3 ~6-fold (**Figure 2.4A**). Consistent with their binding defects, the Get3 mutants E253K and E304K were not inhibited by Get4/5 (**Figure 2.4A**). Notably, Get3 K69D, situated at the regulatory interface, significantly lost the ability to be inhibited by Get4/5 (**Figure 2.4A**), although it bound Get4/5 with similar affinities to wild type (**Figure 2.3C**). A Get3 K72D mutant also lost the ability to be inhibited by Get4–5 relative to wild-type, albeit to a smaller extent (**Supplementary Table 2.1**). Mutation of the invariant Get4 Asp74, situated opposite Get3 Lys69 (**Figures 2.3B** and **2.4C** and **Supplementary Figures 2.2B-C** and **2.3**), yielded the same phenotype (**Figures 2.3C** and **2.4A**). Importantly, combining both opposing mutants (Get3 K69D/Get4 D74K) restored the ability of Get4–5 to regulate Get3 ATPase activity, demonstrating that these two residues directly interact (**Figure 2.4A**). This is again consistent with the high conservation of residues located on either side of this interface (**Figure 2.3B** and **Supplementary Figure 2.2B-C**). These results demonstrate

that Get4 plays two distinct roles for Get3, recruitment and regulation, which can be biochemically decoupled.

To test whether the regulation of Get3 ATPase activity is important for TA targeting, a reconstituted *in vitro* targeting assay was used [41]. Specifically, a TA-substrate, Sbh1, was translated in Δ*get3* yeast extracts and targeted to ER microsomes by exogenously added Get3. The efficiency of targeting is then reported by the glycosylation of an engineered opsin tag on Sbh1 upon insertion into microsomes. Mutant Get3 K69D exhibits a ~40% loss of Sbh1 insertion compared to wild-type, which agrees with its loss in Get4/5-induced regulation of ATPase activity (**Figure 2.4A-B** and **Supplementary Figure 2.4A**). Importantly, this effect is only seen in the presence of Get4/5 as both wild-type Get3 and Get3 K69D have the same targeting efficiency using translation extracts from a Δ*get3/*Δ*mdy2* strain (Get4 is depleted in this strain[34]) (**Supplementary Figure 4B-C**). This is distinct from Get1-2 binding mutants as the critical E253K mutant (that cannot bind Get1 or Get2) [29-31] completely abolishes insertion in both Δ*get3* and Δ*get3/*Δ*mdy2* extracts (**Supplementary Figure 4B-C**), which demonstrates that the Get3 K69D mutant does not directly affect the membrane-associated steps. The formation of functional Get3-TA complexes likely follows a mechanism similar to wild-type in these mutants, as the data still fits a Hill coefficient of 2, previously shown to correlate with Get3 tetramer formation [41]. In addition, the targeting by Get3 K69D cannot be rescued by increasing protein concentration (**Figure 4B**), consistent with a model in which premature ATP hydrolysis in this mutant reduces the fraction of productive Get3-Get4/5 complexes that can capture and target the TA substrate. Thus, Get4/5-induced delay of ATP hydrolysis from Get3 is integral for ensuring efficient TA protein targeting.

To examine whether this regulation is important for Get3 function *in vivo*, the ability of Get3 K69D and Get4 D74K to rescue known knockout phenotypes was tested using a yeast growth assay [19,23,36]. As before, neither Δ*get3* nor Δ*get4* strains showed a phenotype when

grown on synthetic complete media at 30$^{o}$C. However, growing these strains at 40$^{o}$C in the presence of 2mM $Cu^{2+}$ produced a strong phenotype that could be rescued by expression of the wild-type protein on a plasmid (**Figure 2.4D**). A Get3 K69D mutant was unable to fully rescue the growth phenotype supporting a role for regulation *in vivo*. A Get4 D74K mutant was also unable to fully rescue, resulting in an even stronger phenotype than the Get3 K69D mutant (**Figure 2.4D**). It is important to note that this surface of Get4 has no other known interacting partners, meaning that this phenotype can only be readout of the regulatory role of Get4. In total, these results provide strong evidence that Get4/5-mediated regulation of Get3 ATPase activity is critical for a functional GET pathway.

## Discussion

The structure of the Get3-Get4/5 complex presented here reveals the molecular basis of Get3 recognition by Get4. In particular, the structure provides insight into the role of nucleotide in complex formation, where Get4 binds to both monomers of Get3 in an orientation only compatible with a closed Get3. The anchoring interface was demonstrated to mediate the interaction between Get3-Get4, while the regulatory interface is critical for inhibition of Get3 ATP activity. This regulation of Get3 is necessary for efficient targeting *in vitro* and loss of regulation leads to growth defects *in vivo*. While it is difficult to speculate at sub-atomic resolution, it is interesting to note that Get3 Lys69 connects through a short helix to the critical Switch I loop that contains the catalytic D57. This would make the interaction between Get4 Asp74 and Get3 Lys69 allosteric, leading to inactivating conformational changes in the catalytic pocket.

There is growing evidence that the soluble Get3-TA complex contains a tetramer of Get3, in which two copies of the dimer form a hydrophobic chamber [28,41]. A tetramer of Get3 is observed in the Get3-Get4/5 crystal packing that is strikingly similar to the tetramer structures seen in the archaeal Get3 homolog [28] and in the crystal packing of the Get2-Get3 complex [30] (**Figure 2.5A-B** and **Supplementary Figure 2.5**). If one considers the orientations of the Get4 monomers across the tetramer, the distances are compatible with the requirement that they be bridged by the rest of the Get5 dimer (**Figure 2.5A**) [33]. In contrast, while a Get3 dimer presents two potential Get4 binding sites, a single Get4/5 hetero-tetramer would be unable to occupy both, as this would require steric clashes to Get3. The tetramer seen here then likely represents a pre-hydrolysis Get3 complex waiting for a TA substrate to trigger hydrolysis and release of the complex from Get4.

Coupling this structural data and biochemistry with the current literature allows us to provide a refined model for TA selection by the Get3-Get4/5 complex (**Figure 2.5C**). (1) In the

absence of nucleotide Get3 is predominantly in the open conformation with low affinity for Get4/5. (2) Binding of ATP would shift Get3 to a closed conformation that would be recognized by Get4/5, and promote inhibition of Get3 ATPase activity. (3) A second Get3 dimer binding to the Get4/5 hetero-tetramer would follow. The complex would now be primed for capture of a TA substrate from the co-chaperone Sgt2 [24] bound to the Ubl-domain of Get5 [40]. (4) The stabilized tetramer-TA complex results in ATP hydrolysis [41] causing release from Get4 and subsequent delivery to the ER.

In all organisms, cascades of protein biogenesis factors mediate the chaperoning and handover of nascent proteins from the ribosome to their final folding state or cellular destination. Active regulation of the conformation and nucleotide state of protein biogenesis factors, as studied here for Get3, has also been observed for the SRP-SRP receptor complex during co-translational protein targeting [48,49]. This likely represents a general mechanism for ensuring efficient and productive biogenesis of nascent proteins.

**Acknowledgements**

We are grateful to V. Denic (Harvard University) for providing the Δ*get3*/Δ*get5* yeast strain. We thank Graeme Card, Ana Gonzalez and Michael Soltis for help with data collection at SSRL BL12-2. We are grateful to Gordon and Betty Moore for support of the Molecular Observatory at Caltech. Operations at SSRL are supported by the US Department of Energy and US National Institutes of Health (NIH). This work was supported by career awards from the David and Lucile Packard Foundation and the Henry Dreyfus Foundation (S.-o.S.), US National Science Foundation Graduate Research Fellowship DGE-1144469 (M.E.R.), NIH training grant 5T32GM007616-33 (H.G. and M.R.) and NIH research grant R01GM097572 (W.M.C.).

## Methods

### Protein cloning, expression, and purification

The sequences of Get4 and Get5 were cloned as previously described [36]. To generate the Get4/5N used in this study, this construct was further modified by truncating the C-terminus of Get4 (residues 291-312), and by the addition of a stop codon after residue 54 within Get5. All *S. cerevisiae* Get4 mutants were generated using the QuikChange mutagenesis method (Stratagene) and verified by DNA sequencing. All Get4/5 proteins were overexpressed in BL21-Gold (DE3) (Novagen) grown in 2xYT media at 37 °C and induced for 3h by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were lysed using a microfluidizer (Microfluidics) and purified as a complex by Ni-affinity chromatography (Qiagen). The affinity tag was removed by an overnight TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 7.5, 30 mM NaCl, and 5 mM β-mercaptoethanol (BME). A second Ni-NTA column was used to remove any remaining his-tagged protein, and the sample was then loaded onto a 6 mL Resource Q anion exchange column (GE Healthcare). The peak containing the Get4/5N complex was collected and concentrated to 15-20 mg/mL. Initial purifications of the Get4/5N complex were verified to be a single monodispersed species over SEC using a Superdex 200 16/60 column (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 100 mM NaCl, and 5 mM BME. Full-length Get4/5 used in ATPase assays and translocation experiments was further purified using a Superdex 200 16/60 column (GE Healthcare) equilibrated with 20 mM K-HEPES pH 7.5, 150 mM KOAc, 10 mM MgOAc, 10% (v/v) Glycerol, and 5 mM BME. Fractions containing Get4/5 were pooled and concentrated to ~5 mg/mL.

The *S. cerevisiae* Get3 coding region was cloned as previously described [23]. A 6xHis-tag followed by a tobacco etch virus (TEV) protease site was fused to the N-terminus, and a stop codon was placed in front of the C-terminal 6xHis-tag. All *S. cerevisiae* Get3 mutants

were generated using the QuikChange method. Get3 mutants used in SEC, ITC, or capture assays were introduced into the Get3D construct, whereas mutants used in ATPase assays or translocation assays were introduced into the wildtype Get3 construct. All Get3 proteins were made in BL21-Gold(DE3), grown in 2xYT media, and induced with 0.5 mM IPTG for 16h at 22$^{o}$C. Cells were lysed using a microfluidizer (Microfluidics) and purified by Ni-affinity chromatography (Qiagen). The affinity tag was removed by an overnight TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 7.5, 100 mM NaCl, and 5 mM BME. A second Ni-NTA column was used to remove any remaining his-tagged protein, and then the sample was run on a Superdex 200 16/60 column (GE Healthcare) equilibrated with the dialysis buffer. Fractions corresponding to a dimer of Get3 were pooled and concentrated to 15-20 mg/mL. Get3 derivatives used for ATPase assays and translocation experiments were further purified over a MonoQ anion exchange column to remove contaminating ATPases.

Get3D-Get4/5N complex was formed by equilibrating 105 μmol Get4/5N with 100 μmol Get3D at room temperature in 500 μL of 20mM Tris pH 7.5, 10mM NaCl, 5mM BME, 1mM $MgCl_2$, and 1mM of either AMP-PNP or ATP. Prior to complex formation, Get3 had been pre-equilibrated with 1mM $MgCl_2$ and either 1mM AMP-PNP or ATP for 5 min at room temperature. Get3-Get4/5N complex was further separated from free Get4/5N using a Superdex 200 10/300 (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 10 mM NaCl, and 5 mM BME or with 20 mM Bis-Tris Propane pH 9.0, 10 mM NaCl, and 5 mM BME. All complexes were concentrated to 10-12 mg/mL before use in crystallization experiments.

**Crystallization**

Purified Get3D-Get4/5N complex was concentrated to 10-12 mg/ml and crystal trials were carried out using the sitting-drop vapor diffusion method at room temperature by equilibrating equal volumes of the protein complex solution and reservoir solution using a TTP

LabTech Mosquito robot and commercially purchased kits (Hampton Research, Qiagen, Molecular Dimensions Limited). Get3D-Get4/5N crystals grew in the presence of 18% PEG 3350, 0.2 M KSCN, 0.1 M Bis-Tris Propane pH 9.0, and 5% DMSO. Crystals were cryoprotected by transferring directly to 10μL of a reservoir solution supplemented with 20% glycerol, 1mM ATP, and 1 mM $MgCl_2$ and incubated for ≥10 minutes before being flash frozen in liquid nitrogen.

**Data collection, structure solution, and refinement**

All structures were solved using datasets collected on Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL) at 1 Å at ~100 K. Each structure was solved from a single dataset that was integrated using MOSFLM [50] or XDS [51], and scaled and merged using SCALA [52,53]. Crystals of Get3D-Get4/5N diffracted to 5.4 Å and was solved by molecular replacement with PHASER [54] as implemented in PHENIX [55], using a monomer of the closed (ADP-$AlF_4$) form of Get3 (PDB ID 2WOJ [24]) and one Get4/5N heterodimer (PDB ID 3LKU [36]) as search models. No solution was found using the open (*apo*) form of Get3 (3A37 [27]). Refinement was performed using REFMAC v6.3 with rigid body restraints and in CNS v1.2 [56] using DEN refinement. Manual rebuilding was performed using COOT [57]. The final model refined to an *R*-factor of 27.0% ($R_{free}$ = 32.8%) with residues in the Ramachandran plot in 92.2% preferred, 6.0% allowed, and 1.8% in the disallowed and restricted regions [57]. Full statistics in Table 1.

**Size exclusion chromatography for complex stability**

250 μL of 25 μM Get3D and 250 μL of 25 μM Get4/5N were combined and dialyzed at room temperature in 20 mM Tris pH 7.5, 10 mM NaCl, 5 mM BME, 1 mM $MgCl_2$, and, where indicated, 1 mM of either ADP or ATP. The total samples were injected onto a Superdex 200

10/300 (GE Healthcare) equilibrated in 20 mM Tris pH 7.5, 10-500 mM NaCl, and 5 mM BME.

**Capture assay**

500 nmol of 6xHis-tagged Get4/5N was incubated with 10 µL Ni-NTA agarose resin for one hour at 4°C in 500 µL binding buffer containing 20 mM K-HEPES pH 7.5, 150 mM KOAc, 10 mM MgOAc, 10% (v/v) Glycerol, 25 mM imidazole, and 1 mM ADP or ATP where indicated. Following the addition of 1 µmol of Get3D, the solution was incubated for an hour at 4°C. After incubation, the reaction was spun for 30 sec at 500xg. The supernatant was removed, and 500 µL binding buffer was added to the solution and gently mixed through inversion. The wash step was repeated twice, and following the final wash, the remaining bound proteins were eluted with 30 µL of 20mM Tris pH 7.5, 100 mM NaCl, 5mM BME, and 300 mM imidazole. The samples were spun for 30 sec at 500 x g, and the supernatant was removed and added to 6 µL of 6xSDS-PAGE buffer. All samples were run on 15% SDS-PAGE gels and stained with Coomasie blue G-250. Gels were then analyzed by infared scanning in the 700 nm channel using a LI-COR Odyssey Infared Imaging System and Odyssey Application Software v3.0.30.

**Isothermal titration calorimetry**

Get3D-Get4/5N binding experiments were carried out using the MicroCal iTC200 system (GE Healthcare). Binding affinities were measured by filling the sample cell with 50 µM Get3D and titrating 350 µM Get4/5N. The buffer conditions were identical for all samples and contained 20 mM K-HEPES pH 7.5, 150 mM KOAc, 10 mM MgOAc, 10% (v/v) Glycerol, and 1mM ATP. For each experiment, 2 µl of Get4/5N was injected into Get3 for 20 intervals spaced 120 sec apart at 25°C. For the first titration, 0.4 µl of Get4/5N was injected.

The stirring speed and reference power were 1000 rpm and 5 μcal/s. Affinity constants were calculated from the raw data using Origin v7.4 software (MicroCal).

**ATPase assay**

Get3 ATPase rates were measured as previously described [41]. Briefly, the $k_{cat}$ for 8μM Get3 was determined in the presence of excess of ATP doped with ($\gamma$-$^{32}$P) ATP, and analyzed by autoradiography. Each Get3 ATPase reaction was conducted in the presence or absence of excess (20μM) full-length Get4/5. For Figure 2.4A, individual ratios were calculated for each of *n* independent trials (**Supplementary Table 2.1**) performed on separate days and then a mean and standard deviation were calculated across *n* ratios. Each independent trial was the average of values from two side-by-side reactions. Values used in Supplementary Table 2.1 are means and standard deviations calculated across *n* experiments.

**Translocation assay**

The coding sequence for yeast Sbh1 was cloned into a transcription plasmid [34] under control of an SP6 promoter and modified as previously described [41]. Sbh1 mRNA was transcribed using the SP6 Megascript kit (Ambion). All translation and translocation assays were performed in yeast as previously described [28] using extracts and microsomes from either a *Δget3* [19] or *Δget3/Δget5* [34] strain. Get3 translocation efficiency was plotted as a function of Get3 concentration and analyzed as previously described [41].

**Yeast growth assay**

Knockout strains BY4741 *YDL100C::KanMX* (Get3) and BY4741 *YOR164C::KanMX* (Get4) were purchased from America Type Culture Collection (ATCC) and used as previously described [23,36]. The Get3 rescue plasmid was constructed by PCR amplifying the open reading

frame with 242 bp upstream and 263 bp downstream flanking regions from BY4741 genomic DNA. The Get4 rescue plasmid was constructed by PCR amplifying the open reading frame with 233 bp upstream and 86 bp downstream flanking regions from BY4741 genomic DNA. Both genes were amplified with SalI and NotI restriction sites and ligated into the pRS316 vector [58]. Mutants were generated by Quikchange site-directed mutagenesis. Yeast strains were transformed using the Li/Ac/single-stranded carrier DNA/PEG method [59]. Phenotypic rescue was determined by growing each transformant in SC-Ura media at 30 °C to an $OD_{600\ nm}$ between 1 and 2, diluting to 3.85 x $10^6$ cells⁄mL and spotting 4 μL of serial dilutions onto SC -Ura agar plates in the presence or absence of 2 mM $CuSO_4$. Plates were then incubated at 30 °C or 40 °C for 24-48 h and photographed. The results were consistent through three trials.

**Structure analysis and Figures**

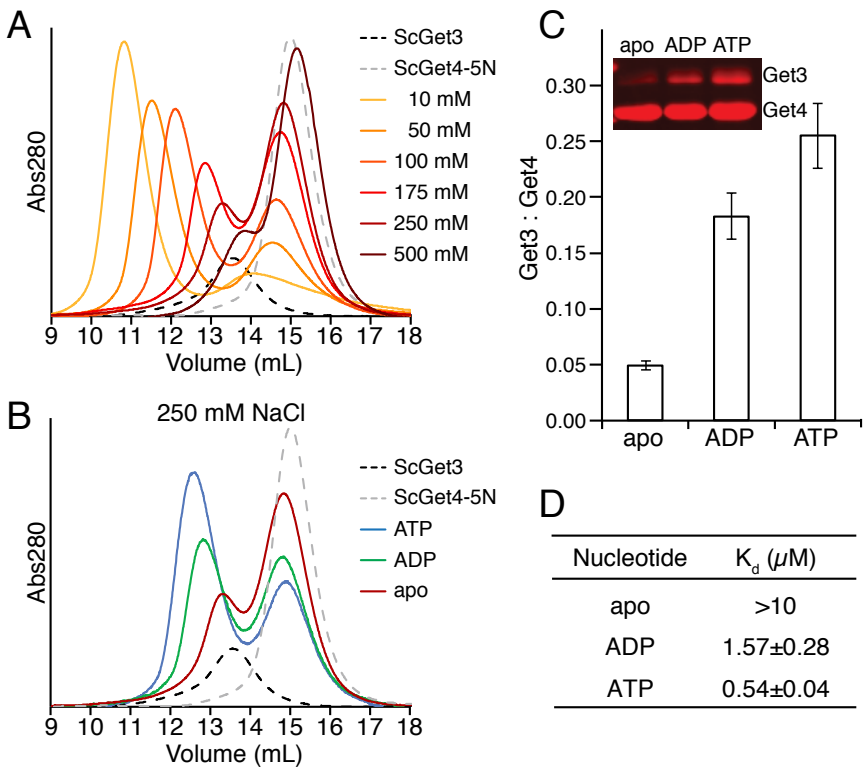
Cartoon representations of protein structures were prepared using PyMol (Schrodinger, LLC), while surface representations were prepared using UCSF Chimera [60]. Surface figures were made in Chimera. Conservation used values for individual residues based on an alignment from ClustalW [61]. Electrostatic surface potentials were calculated using APBS with default values as implemented in the PDB2PQR webserver [62,63].

**Accession codes**

Atomic coordinates and structure factors of the Get3D-Get4/5N complex have been deposited in the Protein Data Bank under accession code 4PWX.

# Figures

## Figure 2.1




| Nucleotide | $K_d$ ($\mu$M) |
|---|---|
| apo | >10 |
| ADP | 1.57±0.28 |
| ATP | 0.54±0.04 |

**Figure 2.1. Get4/5 prefers ATP-bound state of Get3.**

(A) SEC of Get3D-Get4/5N run in various salt concentrations. (B) SEC of Get3D-Get4/5N in 250 mM NaCl. (C) Pulldown experiments using 6xHis-tagged Get4/5N and Get3D. The amount of Get3D retained after elution is plotted as a fraction of Get4, and error bars represent the s.d. from n=3 technical replicates. (D) Summary of the binding affinities of Get4/5N to Get3D obtained by ITC experiments. Data is expressed as $K_d$ (μM) with n = at least 3 technical replicates.

**Figure 2.2**

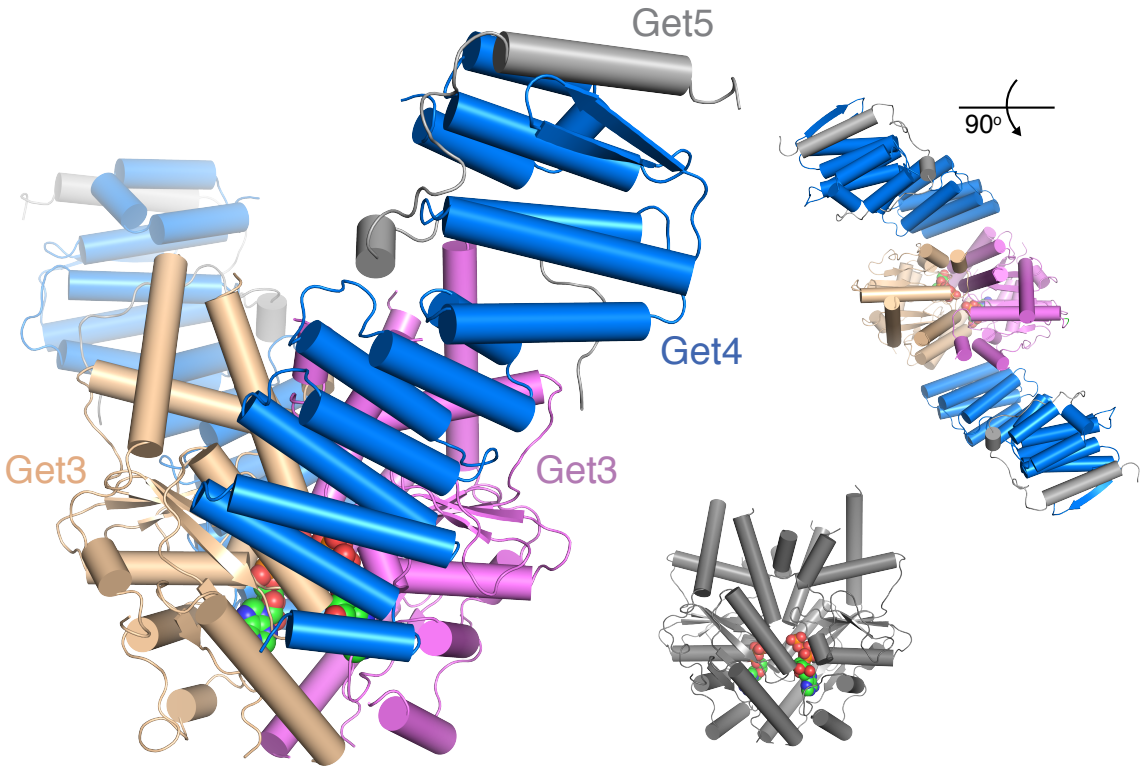
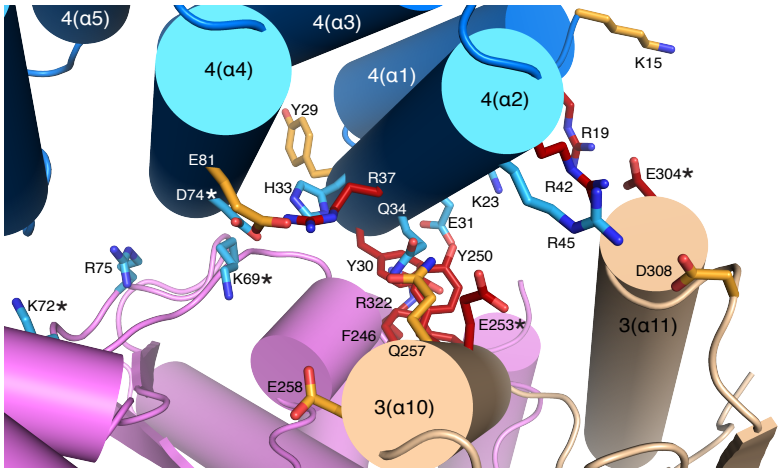



**Figure 2.2. 5.4 Å crystal structure of an ATP-bound Get3-Get4/5 complex.**

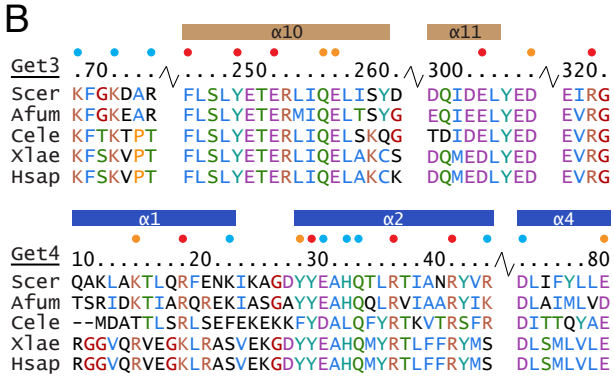
The asymmetric unit of the Get3D-Get4/5N complex in two orientations (Get4 blue and Get5N gray) bound to Get3D (wheat and magenta). Bottom right: Get3D dimer alone in gray to emphasize the 'closed' structure. ATP is represented as spheres.

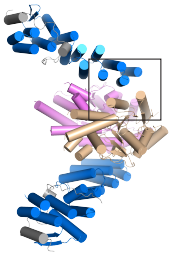
**Figure 2.3**

A




B



C

| | | Mutant | ITC Kd (nM) | Capture (Get3 : Get4) |
|---|---|---|---|---|
| Get3 D57V | | WT | 540 | 1.00 |
| | Anchoring | F246A | >10,000 | 0.33 |
| | | Y250A | >10,000 | 0.06 |
| | | E253K★ | >10,000 | 0.03 |
| | | Q257A | 2,990 | 0.66 |
| | | E258A | 3,120 | 0.36 |
| | | E304A | - | 0.17 |
| | | E304K★ | >10,000 | 0.04 |
| | | D308A | 3,720 | 0.45 |
| | | R322A | >10,000 | 0.10 |
| | Regulatory | K69A★ | 760 | 0.76 |
| | | K69D★ | 660 | 0.86 |
| | | K72A★ | 1,150 | 0.76 |
| | | K72D★ | 840 | 0.87 |
| | | R75A | 990 | 0.80 |
| Get4 | Anchoring | K15A | 1,190 | 0.69 |
| | | R19A | - | 0.25 |
| | | K23A | 340 | 1.12 |
| | | Y29A | 1,930 | 0.74 |
| | | Y30A | 3,730 | 0.34 |
| | | E31A | 170 | 1.00 |
| | | H33A | 300 | 0.93 |
| | | Q34A | 400 | 0.80 |
| | | R37A | >10,000 | - |
| | | R42A | >10,000 | 0.44 |
| | | R45A | 950 | 0.86 |
| | Reg. | D74K★ | 1,010 | 0.51 |
| | | E81K | 3,480 | 0.63 |

**Figure 2.3. Get3-Get4 binding interfaces.**

(A) View of the Get3-Get4 interface showing interactions between Get4 (blue) and both monomers of Get3D (wheat and purple). Residues at the interface tested for interaction are shown as sticks and colored based on phenotype. The positions of side chains cannot be determined at this resolution and are only shown here for reference. Below, overview of Get3D-Get4/5N in same orientation used to show the interface. Area within the box represents the interface shown above. (B) Sequence alignments of regions involved in contacts in the Get3-Get4 interface and colored based on ClustalW output [61]. Sequences are Scer – *S. cerevisiae*, Afum – *Aspergillus fumigatus*, Cele – *Caenorhabditis elegans*, Xlae – *Xenopus laevis* and Hsap – *Homo sapiens*. Helices are indicated above the sequence and labeled. Residues tested are highlighted by spheres and colored based on phenotype (blue, none or minimal; orange, moderate; red, strong). (C) Summary of the data obtained by ITC and pulldown experiments. Mutants are colored based on strength of phenotype as in (A-B). ITC data was generated from a single experiment; pulldown experiments were performed in triplicate, with the mean shown.

**Figure 2.4**

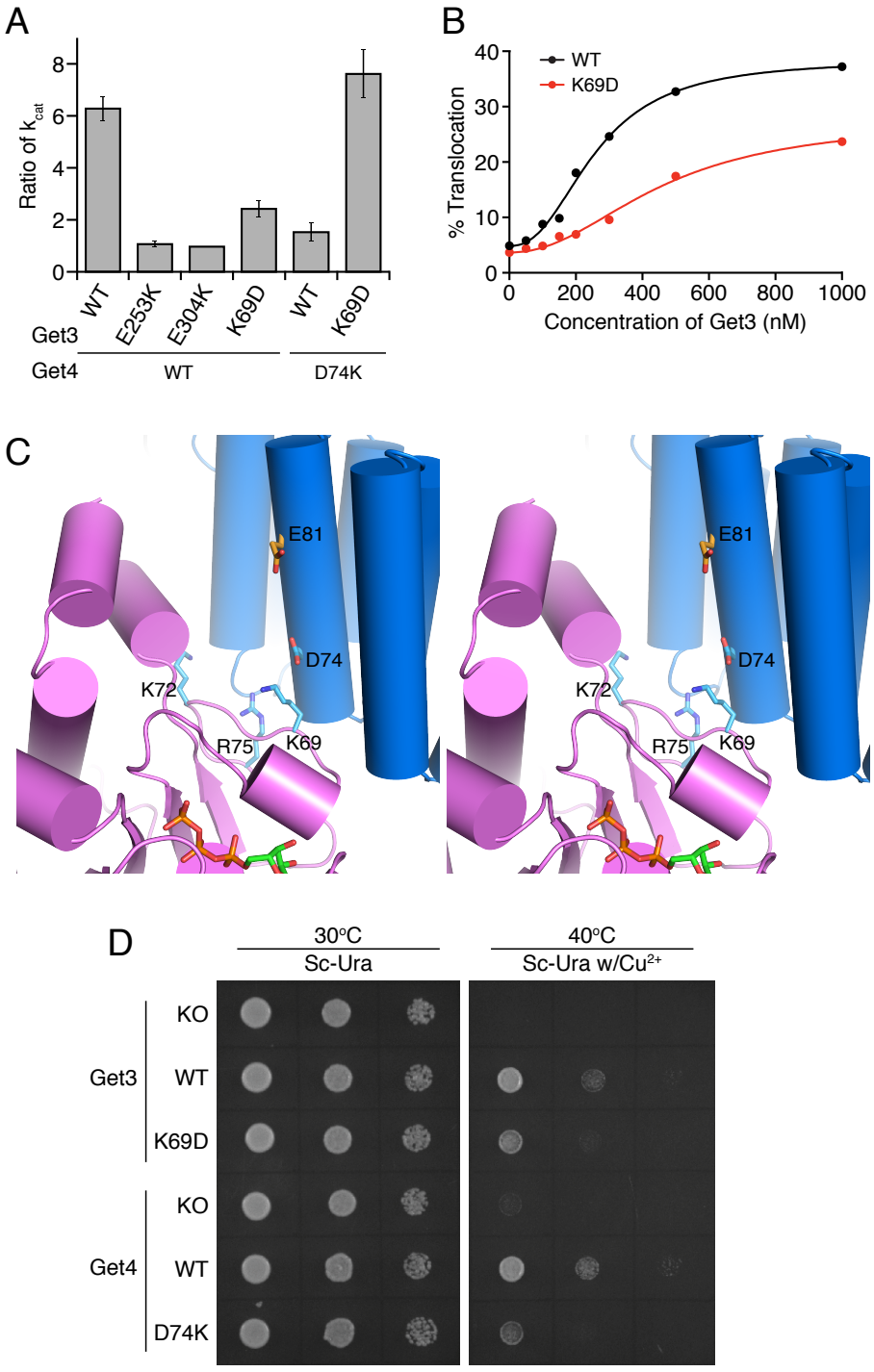

A
Ratio of $k_{cat}$
8
6
4
2
0
WT
E253K
E304K
K69D
WT
K69D
Get3
Get4
WT
D74K
B
% Translocation
40
30
20
10
0
WT
K69D
0
200
400
600
800
1000
Concentration of Get3 (nM)
C
E81
D74
K72
R75
K69
E81
D74
K72
R75
K69
D
30°C
Sc-Ura
40°C
Sc-Ura w/Cu2+
Get3
KO
WT
K69D
Get4
KO
WT
D74K

**Figure 2.4. Get4/5 regulates Get3 ATPase activity.**

(A) Get3 ATPase assay in the presence and absence of Get4/5 and mutants. The Get4/5 effect is represented as a ratio of $k_{cat}$ in the absence and presence of Get4/5, with a value of 1 indicating no inhibition by Get4/5. The values are shown as means and standard variations for ratios calculated from *n* independent trials (Supplementary Table 1). (B) Comparison of Get3 translocation efficiency between *wt* and K69D. (C) Stereo view of the regulatory interface showing interactions between Get4 (blue) and Get3 (purple). (D) Spot plate growth assays of pRS316 derived rescue plasmids under control of genetic promoters in the BY4741 Get3::KanMx or Get4::KanMx background. "KO" represents transformation with empty vector. Plates consisted of Sc-Ura with or without 2 mM $CuSO_4$. Each image was taken from a single plate at either 24 h (30$^{o}$C, Sc-Ura) or 48 h (40$^{o}$C, Sc-Ura w/2 mM $CuSO_4$).

**Figure 2.5**

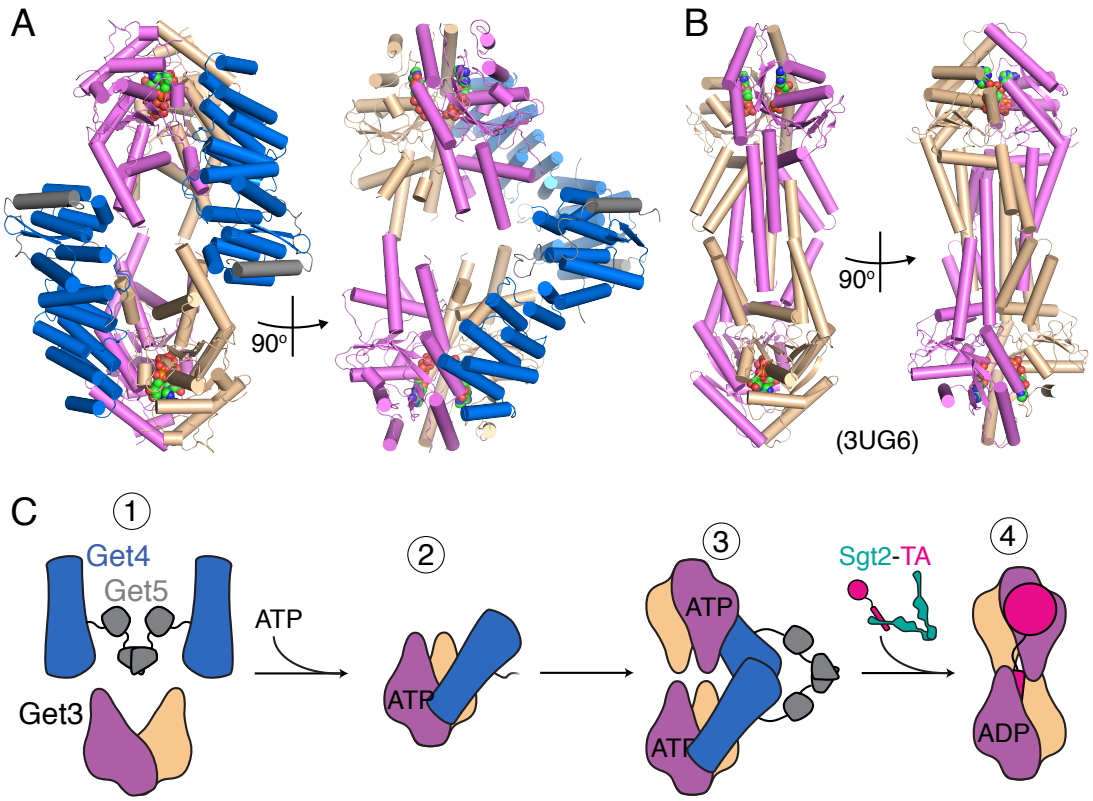



**Figure 2.5. A working model for Get4/5.**

(A) The tetramer of Get3 from the Get3D-Get4/5N crystal lattice in two orientations. Colored as in Figure 2B. (B) The tetramer of an archaeal Get3 (PDB ID 3UG6 [28]) oriented similar to (A). (C) A model for the assembly of the Get3-Get4/Get5 tail-anchor binding complex. Colors correspond to those in Figure 2 and Figure 5.

## Tables

**Table 2.1. Data collection and refinement statistics.**

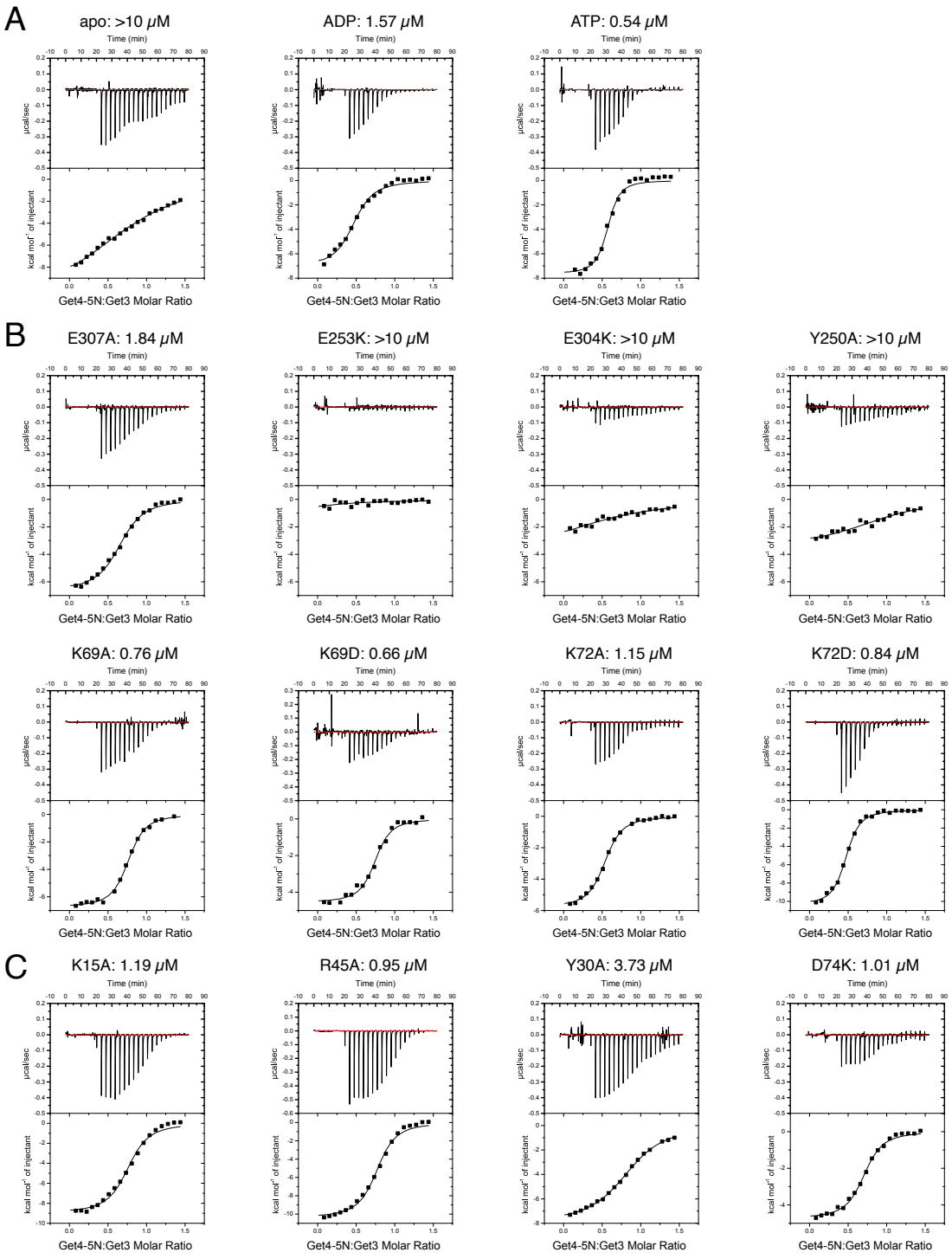
| | SSRL BL12-2[1] |
|---|---|
| **Data collection** | |
| Space group | C 2 |
| Cell dimensions | |
| *a*, *b*, *c* (Å) | 166.3, 134.5, 84.1 |
| a, b, g (°) | 90, 113.4, 90 |
| Resolution (Å) | 30.0-5.4 (6.0-5.4)[2] |
| $R_{sym}$ or $R_{merge}$ | 0.05 (0.46) |
| $I$ / s$I$ | 7.4 (2.0) |
| Completeness (%) | 93.8 (95.9) |
| Redundancy | 0.05 (0.46) |
| | |
| **Refinement** | |
| Resolution (Å) | 30.0-5.4 (6.0-5.4) |
| No. reflections | 5,529 |
| $R_{work}$ / $R_{free}$ | 0.270/0.328 |
| No. atoms | 10,177 |
| Protein | 10,112 |
| Ligand/ion | 65 |
| Water | n/a |
| *B*-factors | |
| Protein | 348.5 |
| Ligand/ion | 317.4 |
| Water | n/a |
| R.m.s. deviations | |
| Bond lengths (Å) | 0.0032 |
| Bond angles (°) | 0.94 |

[1]A single native crystal was used to determine the structure of Get3D–Get4–Get5

[2]Values in parentheses are for highest-resolution shell.

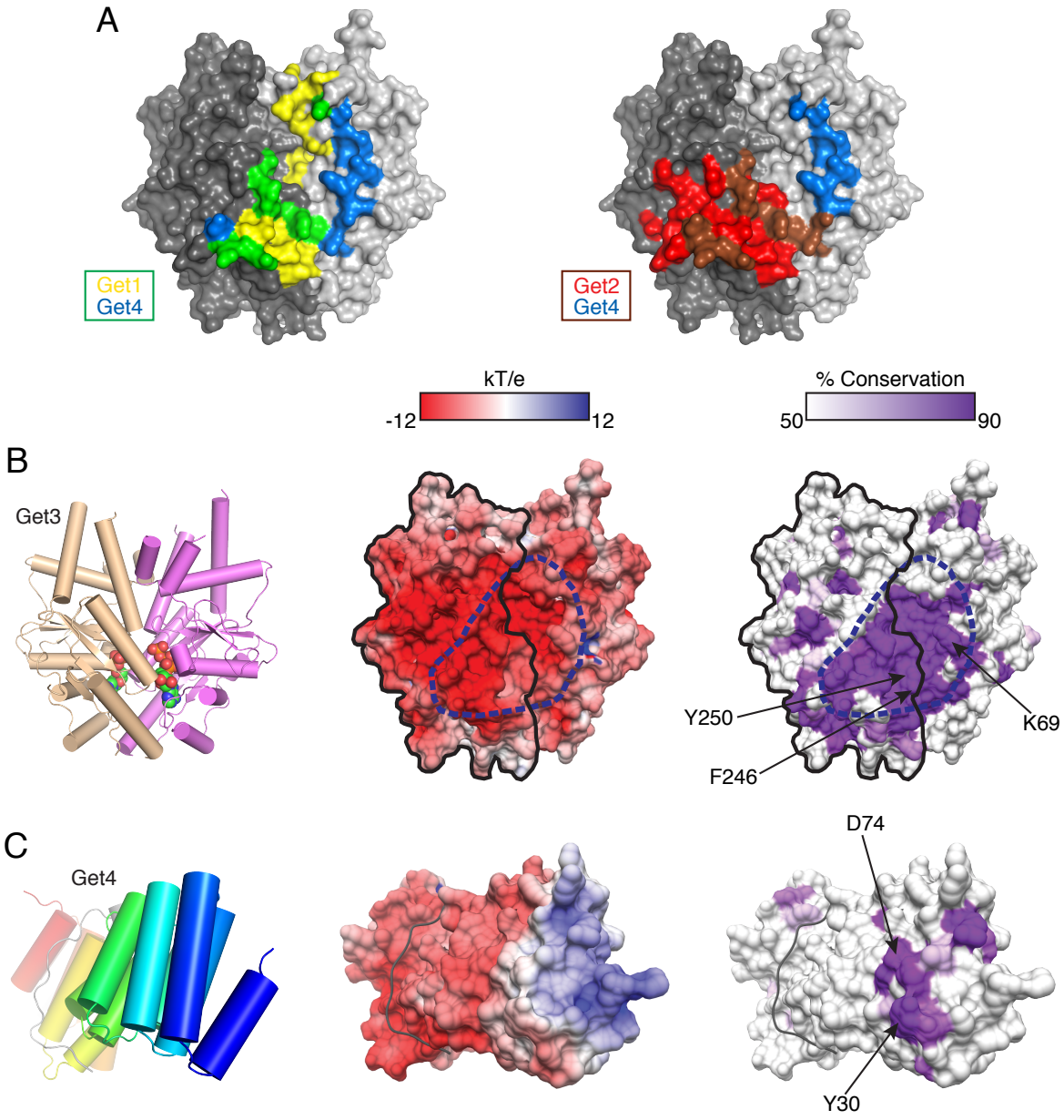
# Supplementary Figures

## Supplementary Figure 2.1



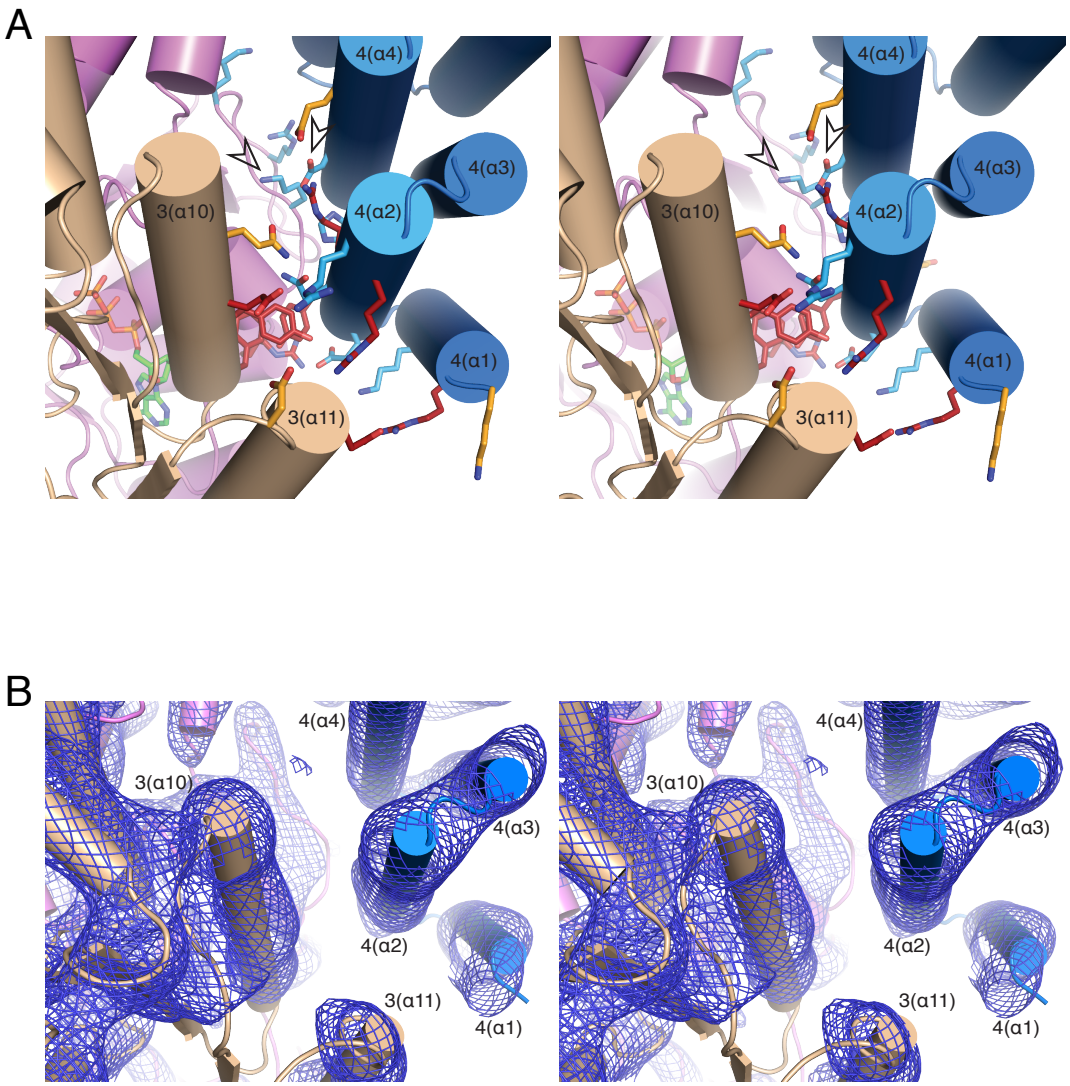
**Supplementary Figure 2.1. Representative ITC isotherms for nucleotide-dependent complex formation.** (A) Representative measurements for Get3 in various nucleotide states. Raw data are shown in the top panel of each trial represented as the power input into the sample cell over time. Integrated data are shown in the bottom panels in terms of the total energy required for equilibration as a function of the molar ratio of Get4/5N:Get3. The solid line represents the best-fit model used to calculate affinities. The affinities indicated for each nucleotide state are the average from at least three experiments. In all experiments Get4/5N was titrated into the sample cell containing Get3. (B) Representative measurements for Get3 mutants from the Anchoring interface, top, or the Regulatory interface, bottom. The indicated affinity for each mutant represents the value obtained from a single experiment carried out in the presence of ATP. (C) Representative measurements for Get4 mutants from the various interfaces as in (B).

**Supplementary Figure 2.2**



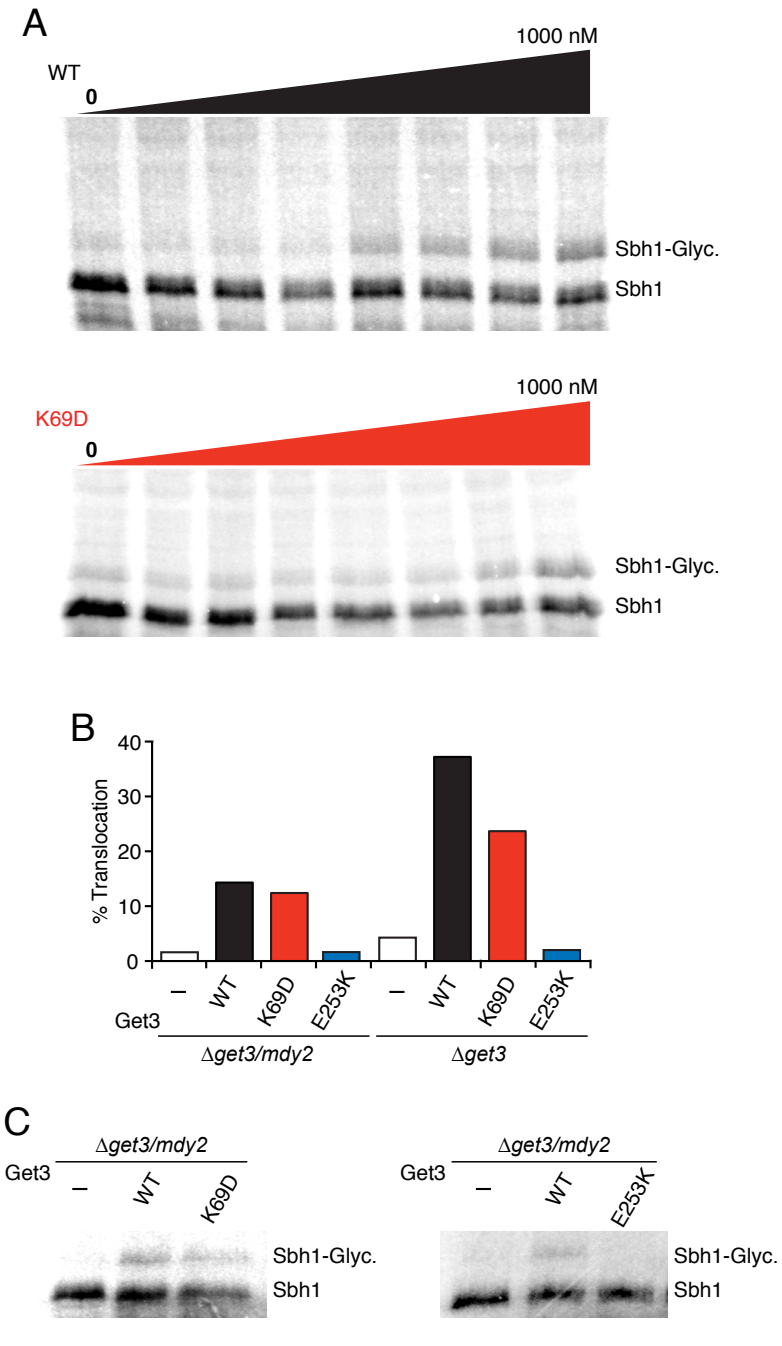
**Supplementary Figure 2.2. Surface properties of Get3 and Get4**. (A) Left, Surface representation of closed Get3 dimer (gray) from (B) showing Get4 binding site (blue), Get1 cytoplasmic domain binding site (yellow) and the overlap (green). Right: Similar surface representation highlighting the Get4 binding site (blue), Get2 cytoplasmic domain binding site (red), and the overlap (brown). (B) Left: orientation of Get3 used in subsequent panels to highlight the binding interfaces. Monomers are colored as in Figure 2. For surface representations one monomer is outlined in black and dashed lines (blue) highlight the interaction surface of Get4 on Get3. Center, accessible surface color ramped based on electrostatic potential from -12 kT/e (red) to +12 kT/e (blue). Get5N colored gray. Right: accessible surface color ramped based on conservation from ≤50% (white) to ≥90% (purple). Get5N colored gray. Coloring is based on a ClustalW [34] alignment using the following sequences: *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Aspergillus fumigatus, Candida albicans, Pichia pastoris, Nematostella vectensis, Caenorhabditis elegans, Drosophila melanogaster, Danio rerio, Xenopus laevis, Mus musculus, Homo sapiens, Arabidopsis thaliana, Neurospora crassa, Anolis carolinensis*. (C) Surface properties of Get4 as in (B), except Get4 is color-ramped from N-terminus (blue) to C-terminus (red).

**Supplementary Figure 2.3**



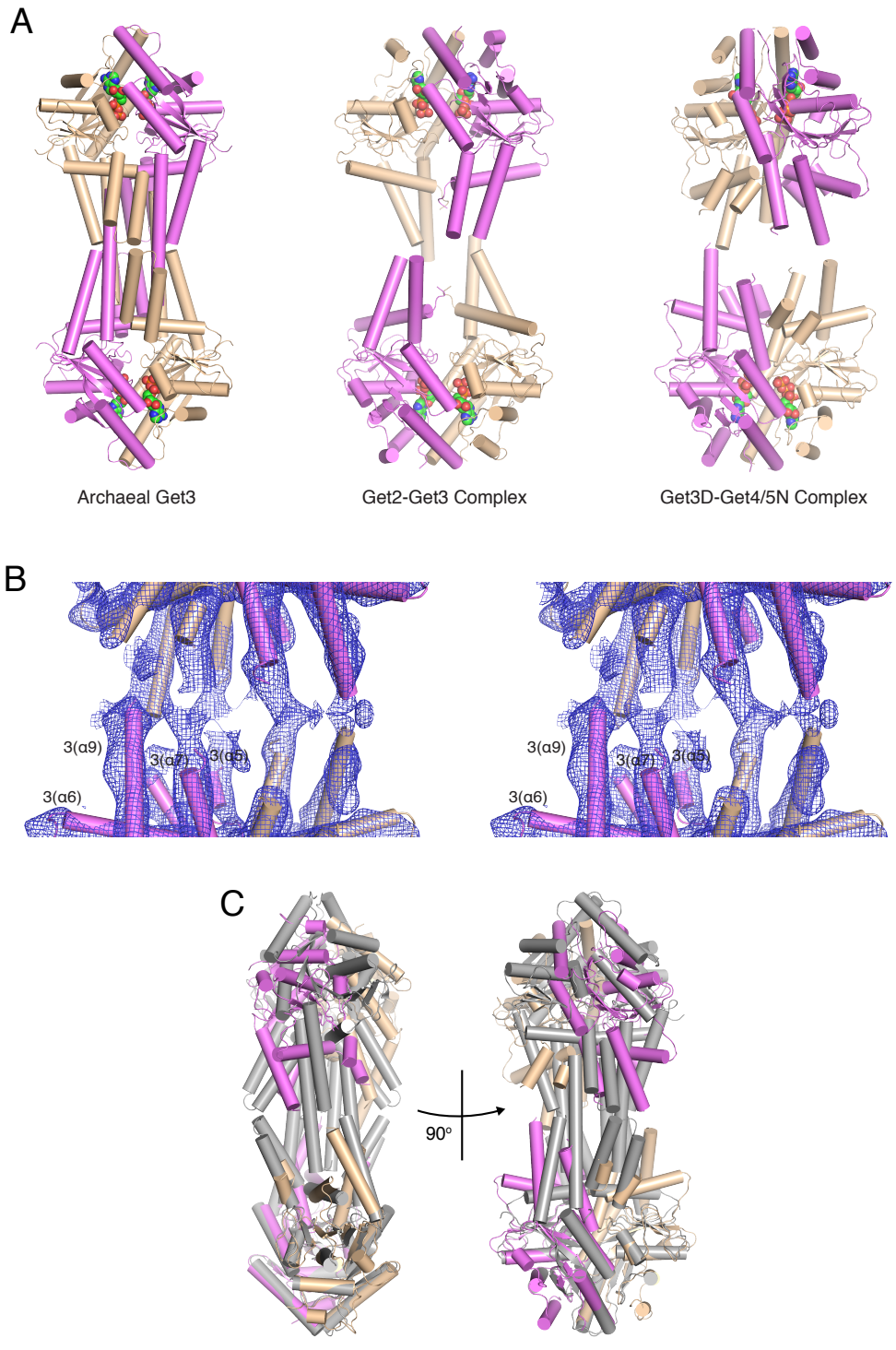

**Supplementary Figure 2.3. Stereo views of the Get3D-Get4/5N interface.** (A) The Get3D-Get4 interface. Coloring is the same as in Figure 3a. Get3 Lys69 and Get4 Asp74 are highlighted by open arrowheads. (B) $2F_O$-$F_C$ Electron density contoured at 1.5σ highlighting the Get3D-Get4 interface.

**Supplementary Figure 2.4**



**Supplementary Figure 2.4. TA targeting assay.** (A) Representative autoradiograph images used to calculate Get3 translocation efficiency corresponding to Figure 4b for wildtype (Top) and K69D (Bottom). (B) Comparison of Get3 translocation efficiencies in Δ*get3* and Δ*get3*/Δ*mdy2* lysates. (C) Representative autoradiograph images used to calculate Get3 translocation efficiencies for Δ*get3*/Δ*mdy2* lysate in (B).

**Supplementary Figure 2.5**



**Supplementary Figure 2.5. Tetramers of Get3.** A) Tetramers of Get3 seen in various crystal forms with Get3 dimers similar to Figure 5A,B. Each is highlighted with the archaeal Get3 (*Methanocaldococcus jannaschii*) from PDB ID 3UG6, the Get2-Get3 complex from PDB ID 3SJD, and the Get3D-Get4/5N crystal structure. (B) Stereo view of the $2F_o$-$F_c$ density contoured at 1.5σ showing the loops involved in formation of the Get3D-Get4/5N complex tetramer. (C) Comparison of the Archaeal MjGet3 dimer aligned to one dimer of the Get3D-Get4/5N complex dimer.

## Supplementary Tables

**Supplementary Table 2.1. Summary of ATPase data**

| Get3 ± Get4/5 | $k_{cat}$ ($min^{-1}$) Mean | SD[1] | *n* |
|---|---|---|---|
| WT | 0.870 | 0.254 | 7 |
| + Get4/5 | 0.139 | 0.043 | |
| E253K | 0.566 | 0.034 | 2 |
| + Get4/5 | 0.536 | 0.030 | |
| E304K | 0.718 | 0.020 | 2 |
| + Get4/5 | 0.739 | 0.041 | |
| K69A | 0.847 | | 1 |
| + Get4/5 | 0.132 | | |
| K69D | 1.213 | 0.298 | 4 |
| + Get4/5 | 0.508 | 0.145 | |
| K72A | 1.417 | | 1 |
| + Get4/5 | 0.203 | | |
| K72D | 1.519 | 0.078 | 2 |
| + Get4/5 | 0.314 | 0.029 | |
| WT | 0.676 | 0.065 | 2 |
| +D74K | 0.462 | 0.107 | |
| K69D | 1.213 | 0.298 | 4 |
| +D74K | 0.135 | 0.004 | 2 |

[1]SD = standard deviation; *n* = number of trials

*C h a p t e r 3*

# MOLECULAR DETAILS OF A GET3-GET4/GET5 INTERMEDIATE COMPLEX

## Abstract

Tail-anchored (TA) proteins, defined as having a single transmembrane helix at their C-terminus, are post-translationally targeted to the endoplasmic reticulum (ER) membrane by the GET (Guided Entry of TA proteins) pathway. In the yeast pathway, the handover of TA substrates is mediated by the heterotetrameric Get4/Get5 (Get4/5) complex, which tethers the co-chaperone Sgt2 to the targeting factor, the Get3 ATPase. Binding of Get4/5 to Get3 is critical for efficient TA targeting; however, questions remain about the formation of the Get3-Get4/5 complex. Here we report crystal structures of a Get3-Get4/5 complex from *Saccharomyces cerevisiae* (*Sc*) at 2.8 Å and 6.0 Å, which reveal a novel interface between Get3 and Get4 dominated by electrostatic interactions. Kinetic and mutational analyses strongly suggest that these structures represent an on-pathway intermediate that rapidly assembles and then rearranges to the final Get3-Get4/5 complex. Furthermore, we provide evidence that the Get3-Get4/5 complex is dominated by a single Get4/5 heterotetramer bound to one monomer of a Get3 dimer, uncovering an intriguing asymmetry in the Get4/5 heterotetramer upon Get3 binding. Ultrafast electrostatically driven Get3-Get4/5 association enables Get4/5 to rapidly sample and capture Get3 at different stages of the GET pathway.

*Adapted from*

Gristick, H.B., Rome, M. E., Chartron, J. W., Rao, M., Hess, S., Shan, S. O., and Clemons, W.M., Jr. Molecular details of a Get3-Get4/Get5 intermediate complex. *In Submission*.

## Introduction

The targeting of membrane proteins to their correct location in the cell is a highly regulated process [44]. The majority of membrane proteins are targeted via the signal recognition particle (SRP), which typically recognizes the initial hydrophobic transmembrane domain (TMD) as it emerges from the ribosome [3]. However, the ubiquitous TA proteins, defined topologically by a single TMD near the C-terminus, are unable to access the SRP pathway as their targeting signal, the single TMD, emerges from the ribosome only after protein synthesis is complete and must be targeted to the ER post-translationally [8,9]. In eukaryotes, TA proteins account for at least 1% of the proteome, and are involved in many essential cellular processes such as apoptosis, vesicle fusion, and protein trafficking [45,46].

A series of genetic and biochemical experiments in yeast and mammalian cells identified members of a dedicated pathway for delivery of TA-proteins to the ER. In yeast, these factors compose the Guided Entry of TA proteins (GET) pathway that consists of six proteins, Get1-5 and Sgt2, all with homologs in higher eukaryotes [47,64]. The first committed step in TA targeting is the formation of an Sgt2/TA complex [34], which may be assisted by chaperones. TA substrate is transferred to Get3 in a Get4/5-dependent manner [34]. Extensive structural characterization of Get3, the central TA targeting factor, demonstrated that it undergoes ATP-dependent conformational changes from an open to closed form required for capturing the TA substrate [23-28]. Once formed, this Get3-TA complex is then localized to the ER by the membrane proteins Get1/Get2 (Get1/2), which stimulate release of the TA protein and subsequent insertion into the ER membrane [29-32].

Efficient delivery of a TA substrate to Get3 requires the hetero-tetrameric Get4/5 complex that provides the link between Sgt2 and Get3 [21,22,34]. Get4/5 stabilizes the ATP-bound state of Get3 delaying ATP release and inhibiting ATP hydrolysis, thereby locking Get3 in a conformation competent for TA substrate binding [41]. Recently, the first crystal structure of a

Get3-Get4/5 complex provided insight into the role of nucleotide in complex formation, where Get4 binds to two functionally distinct binding interfaces, anchoring and ATPase regulation, on a closed Get3 [65]. Importantly, mutations introduced at these interfaces demonstrated that Get4-mediated regulation of ATP hydrolysis by Get3 were critical for efficient TA targeting. A recent study has now shown that Get4/5 binding to Get3 occurs in very rapid diffusion-limited timescales, suggesting an electrostatic interaction [43] However, the nature of this interaction cannot be explained by the aforementioned Get3-Get4/5 structure.

This report describes the crystal structure of a Get3-Get4/5N intermediate complex from *Saccharomyces cerevisiae* (*Sc*) in two crystal forms, at 2.8 Å and 6.0 Å respectively. The structure represents an initial binding interaction mediated by electrostatics that facilitates the rate of subsequent inhibited complex formation. This is supported by kinetic analysis of Get3-Get4/5 complex formation confirming the two-step complex formation. Finally, mass spectrometry and multi-angle light scattering are used to demonstrate that, under physiological steady-state conditions, a single Get4/5 heterotetramer is bound to one monomer of an empty Get3 dimer. This work allows us to generate a refined model for Get3-Get4/5 complex formation.

## Results

### Formation of the Get3-Get4/5 complex follows an electrostatically-driven intermediate

To understand the mechanism of assembly for the Get3-Get4/5 complex, we characterized the assembly pathway using pre-steady-state kinetics. Complex formation was monitored using acrylodan labeled Get4/5, whose fluorescence is enhanced upon binding Get3 [43]. The Get3-Get4/5 complex has been demonstrated to be highly stable yet dynamic, and its formation follows rapid biphasic association at near diffusion-limited timescales [43]. The observed association rate constants in both kinetic phases are linearly dependent on Get3 concentration [43], indicating that both phases represent bi-molecular association of Get3 with Get4/5. While the initial report used full-length Get4/5, the same biphasic association was observed with a truncated Get4/5N that lacks the homodimerization domain of Get5 and hence forms a heterodimer (**Figure 3.1A-B**). This rules out the possibility that the second kinetic phase arises from binding of Get3 to the second Get4 molecule in a Get4/5 heterotetramer.

Additional characterizations strongly suggest the presence of an additional rearrangement step in the assembly pathway. First, the rate of the initial "burst" phase is strongly dependent on ionic strength consistent with electrostatically-driven association [43] (**Figure 3.1C**), whereas the second association phase is independent of ionic strength [43] (**Figure 3.1D**). This suggests a structural rearrangement to adopt a more stable interface prior to the second association phase. Secondly, despite the large changes in the relative association rates of the salt-sensitive and -insensitive phases, the amplitudes of these two phases are unaltered by changes in ionic strength (**Figure 3.1E**). This argues against the possibility that the biphasic association arises from parallel pathways, i.e., that the second association phase arises from a sub-population of Get3 (or Get4/5) that binds the interaction partner with slower kinetics. Thirdly, the dissociation rate constant extrapolated from the fast association phase (**Figure 3.1B**) is 8–10 $s^{-1}$ ($k_{obsd} = k_{on}$[Get3] + $k_{off}$); this $k_{off}$ value is 10–100 fold faster than those

of the final Get3-Get4/5 complex determined directly in pulse-chase experiments [43] (**Figure 3.1F**), suggesting that the initial association gives rise to an unstable intermediate that must subsequently rearrange to adopt more stable interfaces. Further in support of this model are results from the dissociation rate measurements. While dissociation of the Get3-Get4/5 complex also exhibits biphasic kinetics, neither the rates nor the magnitude of the phases are dependent on ionic strength (**Figure 3.1F**). Finally, the magnitude of the slow dissociation phase increases successively when apo-Get3 is compared with ADP- and ATP-bound Get3 (**Figure 3.1G**), consistent with previous data showing that ATP strengthens binding between Get3 and Get4/5 by inducing Get3 into a more closed conformation [36,41,43,55,65].

These new data, coupled with kinetic simulation, support a multi-step mechanism for the Get3-Get4/5 interaction as the simplest model to explain all the available observations (**Figure 3.1H**) [43]. In this model, association of Get3 with Get4/5 involves the following steps: (i) resting Get3 bound to ATP is biased towards a 'closed' conformation; (ii) initial binding of Get4/5 to Get3 is rapid and generates an intermediate dominated by electrostatic interactions; (iii) this is followed by a conformational change to the final stable structure dominated by hydrophobic interactions; (iv) under the experimental conditions, the second subunit in the Get3 dimer can bind another Get4/5 complex to give rise to the second phase in association kinetics Using the experimentally observed $k_{on}$ and $k_{off}$ values from both the salt-sensitive and salt-insensitive phases, the equilibrium binding data and kinetic modeling, we were able to completely assign the rate constant for individual microscopic steps in this model. Analytical simulations based on this model reproduced the experimentally observed association/dissociation kinetics and equilibrium titrations (**Figure 3.2A**). Furthermore, the equilibrium titration data with full-length Get4/5 [43] could only be fit to a 1:2 stoichiometry in which a Get3 dimer is bound to two different Get4/5 heterotetramers (**Figure 3.2B**), consistent with the biphasic nature of complex formation.

Although binding of a second Get4/5 complex to Get3 was observed (**Figure 3.1H**), multiple observations suggest that the stable Get3-Get4/5 complex (species (ii)) dominates under most conditions and accumulation of the saturated Get3-Get4/5 complex is modest. First, previous assays of Get3 activity show that binding of one full-length Get4/5 complex is sufficient to inhibit Get3's ATPase reaction [41]. Second, kinetic simulations show that during complex assembly with relatively stoichiometric amounts of proteins, the saturated Get3-Get4/5 complex accumulates to <20% of all the complexes formed (**Figure 3.2C**). Coupled with the different Get3-Get4/5N complexes in various crystal structures [65], it appears that binding of Get4/5 to Get3 has a preferred, but not obligatory, stoichiometry.

**The structure of a Get3-Get4/5 intermediate complex**

The recent crystal structure of an ATP-bound Get3-Get4/5N complex demonstrated a stable interface that included conserved hydrophobic interactions [65]. The structure allowed rationalization of how Get4 stabilized an ATP-bound Get3; however, it was unclear how the structure could account for the multi-step Get3-Get4/5 assembly seen experimentally. To address this question, we set out to obtain a structure of Get4/5N bound to a Get3 dimer under conditions that would stabilize an initial intermediate complex. As before [65], Get3 and Get4/5 were expressed and purified separately, and then combined and purified as a 1:1 complex in low ionic strength conditions over size-exclusion chromatography (SEC) (**Figure 3.3**). Initial crystals were obtained using a Get4/5N construct similar to that used for solving the structure of the Get4/5N heterodimer [36]. This resulted in a crystal that diffracted to 6.0 Å in space group $P2_1$. Phases for this crystal form were obtained by molecular replacement using an "open" form of Get3 (3A37) [27] and the Get4/5N heterodimer (3LKU) as search models. The final model was refined to an R/R-free of 27.4/30.3% (Table 1). To improve the diffraction limit, purification and crystallization trials were performed in the presence of either ADP or the non-hydrolyzable

ATP-analogue AMP-PNP. In addition, 22 residues from the C-terminus of Get4 were truncated, as they were disordered in previous structures [25,36]. This resulted in a new crystal form that diffracted to 2.8 Å in space group $P2_12_12$. The final model was refined to an R/R-free of 22.4/26.1% (Table 1).

**The Get3-Get4/5 intermediate complex is dominated by an electrostatic interface**

In both crystal forms, Get4 binds in a similar orientation on Get3 (**Figure 3.4A-B**). The 2.8 Å structure contains a 2:1 complex of a Get3 dimer bound to a single Get4/5N heterodimer (**Figure 3.4A**). This crystal form is related to another Get3 crystal form [27] in which limited space in the crystal lattice allows for only one Get4/5N molecule (**Figure 3.4C**) despite the 1:1 stoichiometry of the complex put into crystallization trays as measured by SEC (**Figure 3.3**). While nucleotide was present throughout purification and crystallization, the nucleotide-binding pocket appears empty. For the 6.0 Å structure (**Figure 3.4B**), the asymmetric unit contained a 1:1 complex with 8 copies of Get3 bound to 8 Get4/5N heterodimers (**Figure 3.4B**).

In both cases, Get3 is in an 'open' conformation, and Get4/5N binds in the same interface and orientation, defined here as the intermediate complex. Despite the differences in crystallization and space group, there is little difference in the interface between the two structures, with only a slight 10° rotation at the furthest point (**Figure 3.4D**). The higher resolution structure will be used for reference for the rest of the text. In this intermediate complex, Get4 interacts with a single monomer of Get3 in an orientation compatible with both the 'open' and 'closed' structures, burying ~970 $Å^2$ of surface area (**Figure 3.5A-B**). The interactions in this interface are electrostatic and involve the positive face of Get4 binding to the negative surface of Get3 (**Figure 3.5A-B**).

In this orientation, Get4 α1 sits on top of a groove formed by the loop before Get3 α1 and the loop following Get3 α11, such that the N-terminus of Get4 α1 packs against the loops following Get3 α9 and α10, resulting in an interaction between K23 on Get4 α1 and E307 on Get3 (**Figure 3.5C**). In addition, Get4 α2 packs roughly perpendicular to Get3 α10 and α11, such that the N-terminus contacts Get3 α11 and the C-terminus contacts Get3 α10. This results in interactions between the invariant residues R42 on Get4 α2 and D308 on Get3, and R45 on Get4 α2 with both E253 and Q257 on Get3 α10 (**Figure 3.5C**). Finally, the N-terminus of Get4 α3, which is oriented between α1 and α2, contacts the top of Get3 α10, allowing H51 on Get4 α3 to contact D263 on Get3 α10 (**Figure 3.5C**).

**Comparison of the two distinct Get3-Get4/5 complexes**

In the previously reported ATP-bound Get3-Get4/5N complex Get4/5N is bound across the dimer interface of a "closed" Get3, forming interactions with both monomers, which we define as the inhibited complex (**Figure 3.6**). This is in contrast to the current structure, which is bound to an "open" Get3 and interacts with only one monomer (**Figure 3.6**). Since the Get4-binding surfaces do not undergo a conformational change between the "open" and "closed" Get3 structures, we aligned the helices (α10 and α11) involved in Get4 binding from our intermediate complex onto the inhibited complex. Both structures have overlapping binding sites and predominantly utilize the N-terminus of Get4 for the interaction (**Figure 3.6**). The difference in orientation involves a rotation centered on the 2$^{nd}$ helix of Get4. This rotation allows Get4 α2 to form the majority of the interactions in the inhibited complex, whereas Get4 α1 is responsible for the majority of the interactions in the intermediate complex. Based on the multi-step assembly observed from kinetic analysis, we believe that the new structure dominated by electrostatics represents the intermediate Get3-Get4/5 complex (**Figure 3.1H**, species (ii)) prior to rearrangement to the final structure stabilized by hydrophobic interactions.

**Mutational analysis of the intermediate complex**

The intermediate complex interface involves numerous electrostatic residues on both Get3 and Get4. Of these interactions, only Get3 D263-Get4 H51 and Get3 E307-Get4 K23 are unique to the intermediate interface and are not found in the stable complex (**Figure 3.7A**). The first interaction (D263-H51) is not conserved; however, in higher eukaryotes there is conservation of a likely salt bridge (**Figure 3.7A**). The conserved Get3 E307 also interacts with the cytoplasmic domain of Get2 [29,30], whereas Get4 K23 is not conserved (**Figure 3.7A**). Substitution of Get3 D263 leads to a loss of observable binding to Get4/5N, as measured by ITC (**Figure 3.7B**), whereas substitution of Get3 E307 had a more modest loss in affinity (**Figure 3.7B**).

If the intermediate complex is on-pathway to the final Get3-Get4/5 complex, one would predict that disrupting the interactions unique to this intermediate would slow down Get3-Get4/5 association. Disrupting the putative Get3 E307-Get4 K23 interaction (E307A) resulted in similar association rates compared to wildtype Get3 and, together with the ITC results, suggests that the interaction is not essential for complex formation (**Figure 3.7B-C**). In contrast, significantly reduced rates of Get3-Get4/5 association were observed with Get3 D263A (**Figure 3.7B-C**). While the association rates are slightly higher than previously reported, the differences are small (~2-fold) and do not change the interpretation. Both mutants had virtually identical dissociation rates compared to wildtype Get3; therefore, the difference in affinity for D263A is specific for complex formation kinetics (**Figure 3.7B-C**).

If Get3 D263A follows the same assembly pathway as wildtype Get3, one would expect that once the Get3(D263A)-Get4/5 complex is formed, it has the same equilibrium stability as the wildtype complex since D263 does not form any interactions in the final complex (**Figure 3.7D**). However, the weakened equilibrium stability of the Get3(D263A)-

Get4/5 complex indicates that the changes induced by D263A are more complex, and this mutant likely follows a different assembly pathway leading to a non-native complex. Since a reaction follows the fastest pathway, this observation in turn suggests that the D263A mutation slows the native assembly pathway more extensively than was experimentally measured. Collectively, mutational analysis of the intermediate complex provides independent evidence that the structure observed here represents an on-pathway intermediate during Get3-Get4/5 assembly.

**Stoichiometry between Get3 and Get4/5**

The stoichiometry of the Get3-Get4/5 interaction is not fully understood. A recent SEC-MALLS analysis using an ATPase deficient Get3 (D57N) suggested that the complex contains one Get4/5 heterotetramer bound to one Get3 dimer [66]. Using wildtype Get3, SEC-MALLS analysis using equimolar concentrations of Get3 and Get4/5 confirmed that the size of the complex is consistent with a single Get4/5 heterotetramer bound to a single Get3 dimer (**Figure 3.8A**).

One puzzling observation arises from this data: despite the presence of two Get4 molecules in full-length Get4/5, there has been no evidence for binding of Get3 to the second Get4 in the Get4/5 heterotetramer. To directly test this asymmetry of Get4/5 during complex formation, we developed an alkylation-protection assay. An engineered cysteine at the Get3-Get4/5 interface (Get4 S48C [43]) is allowed to react with N-ethyl-maleimide (NEM) in the absence and presence of various factors (**Figure 3.8B**). In the free Get4/5 complex, Get4 C48 is solvent exposed [36] and rapidly alkylates to completion (**Figure 3.8B-C**). If both copies of Get4 in the Get4/5 complex can bind Get3, S48C should be rendered solvent inaccessible and be completely protected from alkylation (**Figure 3.8B**). Consistent with binding at a single Get4 interface, only 50% of Get4 C48 was protected from alkylation by Get3 (**Figure 3.8C**).

This protection pattern was observed at Get3 concentrations nearly 1000-fold above the dissociation constant for the Get3-Get4/5 complex, indicating that the 50% protection did not arise from incomplete Get3-Get4/5 binding. These results strongly suggest that only one site of a Get4/5 heterotetramer is able to bind Get3.

To further support these results, accessibility of Get4/5 was probed using a 10kDa PEG-maleimide label. A second site was chosen to further validate the interface (Get4 Q34C) that is also occluded in the Get3-Get4/5N structure (**Figure 3.8D**). Importantly, substitution of Q34 to alanine does not impair binding to Get3 [65]. In this reaction, Get4 Q34C with PEG-maleimide forms a covalent adduct that gives a 10kDa increase in mass, which can be detected by SDS-PAGE and Coomassie staining (**Figure 3.8E**). In the absence of Get3, this reaction resulted in 100% pegylation of Get4/5 over a two-minute time course, whereas addition of saturating Get3 resulted in 50% pegylation (**Figure 3.8E**). This provides corroborating evidence that only one Get4 molecule in a heterotetramer binds Get3. Together, these data indicate that once Get3 binds to Get4/5, the other Get4 molecule in the Get4/5 heterotetramer is inhibited from further interaction with another Get3 dimer.

## Discussion

### A model for Get3-Get4/5 complex formation

Multiple lines of evidence support a model in which Get3 and Get4/5 rapidly form an electrostatic intermediate complex, and then undergo a structural rearrangement to a more stable complex [43] (**Figure 3.2A**). This conformational change is consistent with previous results showing that Get4/5 binding induces Get3 into an "occluded" state, leading to Get3 ATPase inhibition and delayed ATP dissociation kinetics [41]. Based on this model, once Get3 has transitioned to the occluded conformation, complex dissociation would be insensitive to buffer ionic strength (**Figure 3.1F**) but nucleotide-dependent (**Figure 3.1G**).

The new crystal structures of the Get3-Get4/5 complex presented here (**Figure 3.2A-B**) reveal a novel-binding interface composed of electrostatic interactions (**Figure 3.5**). These structures provide the molecular basis for the initial, salt-dependent association seen in our kinetic description of the Get3-Get4/5 interaction (**Figure 3.1**). This is corroborated by the D263A mutation introduced within this interface, which drastically slows down complex formation (**Figure 3.7**). These data coupled with the kinetic experiments argue that these structures represent the initial intermediate complex between Get3 and Get4/5.

As we previously reported, Get4/5 is able to precisely discriminate between nucleotide states of Get3, enabling Get4/5 to regulate Get3 activity and prime it for efficient capture of the TA substrate [41,43,65]. The structure described here provides evidence for an additional interaction dominated by electrostatic interactions and characterized by fast association rates. This ultrafast diffusion-limited association of Get4/5 to Get3 [43] may function as an additional mechanism for Get4/5 to select for the appropriate conformational state of Get3, or inversely, enables Get3 to constantly sample Get4/5 complexes until transfer of TA substrate occurs. Interestingly, this may also function to ensure Get3 is quickly recruited following dissociation from the membrane to prevent re-binding to Get1 or Get2 [43]. This type of interaction has been

seen in other systems including barnase-barstar [67] and ribosome interacting proteins [68], and likely represents a recurring theme in protein-protein interactions. This work provides further evidence for the importance of Get4/5 in TA targeting.

**Acknowledgements**

We thank Graeme Card, Ana Gonzalez and Michael Soltis for help with data collection at SSRL BL12-2. We are grateful to Gordon and Betty Moore for support of the Molecular Observatory at Caltech. The PEL (S.H.) was supported by the Gordon and Betty Moore Foundation, through Grant GBMF775 and the Beckman Institute. Operations at SSRL are supported by the US DOE and NIH. This work was supported by career awards from the David and Lucile Packard Foundation and the Henry Dreyfus Foundation (S.-o.S.), NSF Graduate Research Fellowship DGE-1144469 (M.E.R.), NIH training grant 5T32GM007616-33 (H.G. and M.R.) and NIH research grant R01GM097572 (W.M.C.).

## Methods

### Protein cloning, expression, and purification

The sequences of Get4 and Get5 were cloned as previously described [36]. To generate the Get4/5N used in this study, this construct was further modified by truncating the C-terminus of Get4 (residues 291-312), and by the addition of a stop codon after residue 54 within Get5, and verified by DNA sequencing. All Get4/5 proteins were overexpressed in BL21-Gold (DE3) (Novagen) grown in 2xYT media at 37 °C and induced for 3h by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were lysed using a microfluidizer (Microfluidics) and purified as a complex by Ni-affinity chromatography (Qiagen). The affinity tag was removed by an overnight TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 7.5, 30 mM NaCl, and 5 mM β-mercaptoethanol (BME). A second Ni-NTA column was used to remove any remaining his-tagged protein, and the sample was then loaded onto a 6 mL Resource Q anion exchange column (GE Healthcare). The peak containing the Get4/5N complex was collected and concentrated to 15-20 mg/mL. Initial purifications of the Get4/5N complex were verified to be a single monodispersed species over SEC using a Superdex 200 16/60 column (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 100 mM NaCl, and 5 mM BME. Full-length Get4/5 used in ATPase assays and translocation experiments was further purified using a Superdex 200 16/60 column (GE Healthcare) equilibrated with 20 mM K-HEPES pH 7.5, 150 mM KOAc, 10 mM MgOAc, 10% (v/v) Glycerol, and 5 mM BME. Fractions containing Get4/5 were pooled and concentrated to ~5 mg/mL.

The *S. cerevisiae* Get3 coding region was cloned as previously described [23]. A 6xHis-tag followed by a tobacco etch virus (TEV) protease site was fused to the N-terminus, and a stop codon was placed in front of the C-terminal 6xHis-tag. All *S. cerevisiae* Get3 mutants were generated using the QuikChange method. All Get3 proteins were made in BL21-

Gold(DE3), grown in 2xYT media, and induced with 0.5 mM IPTG for 16h at 22$^{o}$C. Cells were lysed using a microfluidizer (Microfluidics) and purified by Ni-affinity chromatography (Qiagen). The affinity tag was removed by an overnight TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 7.5, 100 mM NaCl, and 5 mM BME. A second Ni-NTA column was used to remove any remaining his-tagged protein, and the sample was run on a Superdex 200 16/60 column (GE Healthcare) equilibrated with the dialysis buffer. Fractions corresponding to a dimer of Get3 were pooled and concentrated to 15-20 mg/mL. Get3 derivatives used for ATPase assays and translocation experiments were further purified over a MonoQ anion exchange column to remove contaminating ATPases.

Get3-Get4/5N complex was formed by equilibrating 105 μmol Get4/5N with 100 μmol Get3 at room temperature in 500 μL of 20mM Tris pH 7.5, 10mM NaCl, 5mM BME, 1mM $MgCl_2$, and 1mM of either ADP, AMP-PNP, or ATP. Prior to complex formation, Get3 had been pre-equilibrated with 1mM $MgCl_2$ and either 1mM ADP or 1mM AMP-PNP for 5min at room temperature. Get3-Get4/5N complex was further separated from free Get4/5N using a Superdex 200 10/300 (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 10 mM NaCl, and 5 mM BME. Get3-Get4/5N complex using full-length Get4 was formed as above but in the absence of nucleotide and $MgCl_2$. This complex was further separated from free Get4/5N using a Superdex 200 16/60 (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 10 mM NaCl and 5 mM BME. All complexes were concentrated to 10-12 mg/mL before use in crystallization experiments.

**Crystallization**

Purified Get3-Get4/5N complex was concentrated to 10-12 mg/ml and crystal trials were carried out using the sitting-drop vapor diffusion method at room temperature by equilibrating equal volumes of the protein complex solution and reservoir solution using a TTP

LabTech Mosquito robot and commercially purchased kits (Hampton Research, Qiagen, Molecular Dimensions Limited). The 2.8 Å Get3-Get4/5N OC crystals grew in the presence of 17% PEG 3350, 0.24 M Na citrate, and 30 mM TCEP. Crystals were cryoprotected by transferring directly to 10μL of a reservoir solution supplemented with 20% glycerol, 1mM ADP or AMP-PNP, and 1 mM $MgCl_2$ and incubated for ≥10 min before being flash frozen in liquid nitrogen. The 6.0 Å Get3-Get4/5N OC crystals grew in the presence of 12% PEG 3350 and 0.1 M Na malonate (pH 5.0), and were cryoprotected by transferring directly to 10μL of a reservoir solution supplemented with 20% glycerol before being flash frozen in liquid nitrogen.

**Data collection, structure solution, and refinement**

All structures were solved using datasets collected on Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL). Each structure was solved from a single dataset that was integrated using MOSFLM [50] or XDS [51], and scaled and merged using SCALA [52,53]. Crystal of Get3-Get4/5N in the OC conformation diffracted to 2.8 Å and was solved by molecular replacement with PHASER [54] as implemented in PHENIX [55], using a monomer of the open (*apo*) form of Get3 (PDB ID 3A37) and one Get4/5N heterodimer (PDB ID 3LKU) as search models. No solution was found using the closed (ADP-$AlF_4$) form of Get3 (2WOJ). Coordinates were refined using REFMAC v6.3 [52,53] with NCS and B-factor restraints. Manual rebuilding was performed using COOT [57]. TLS groups were included in the later refinement stages and were determined using the TLSMD web server [69]. These associated regions within the structure were allowed a certain degree of variation based on predicted motions they might have during data collection. The final model refined to an *R*-factor of 22.4% ($R_{free}$ = 26.1%). The second structure in the OC conformation was solved to 6.0 Å using methods similar to those described above. This structure lacked the Get4 C-terminal truncation engineered for the other structures and refined to an *R*-factor of 27.4% ($R_{free}$ = 30.3%). Full statistics in Table 3.1.

**Structure analysis and Figures**

Cartoon representations of protein structures were prepared using PyMol (Schrodinger, LLC), while surface representations were prepared using UCSF Chimera [60]. Surface figures were made in Chimera. Conservation used values for individual residues based on an alignment from ClustalW [61]. Electrostatic surface potentials were calculated using APBS with default values as implemented in the PDB2PQR webserver [62,63].

**Isothermal titration calorimetry**

Get3(D57V)-Get4/5N binding experiments were carried out using the MicroCal iTC200 system (GE Healthcare). Binding affinities were measured by filling the sample cell with 50 μM Get3D and titrating 350 μM Get4/5N. The buffer conditions were identical for all samples and contained 50 mM K-HEPES pH 7.5, 150 mM KOAc, 10 mM MgOAc, 10% (v/v) Glycerol, and 1mM ATP. For each experiment, 2 μl of Get4/5N was injected into Get3 for 20 intervals spaced 120 sec apart at 25°C. For the first titration, 0.4 μl of Get4/5N was injected. The stirring speed and reference power were 1000 rpm and 5 μcal/s. Affinity constants were calculated from the raw data using Origin v7.4 software (MicroCal).

**Multi Angle Light Scattering**

100 μL of 35 μM Get3 and 100 μL of 35 μM Get4/5 diluted in 50 mM K-HEPES pH 7.5, 150 mM KOAc, 10 mM MgOAc, 10% (v/v) Glycerol, 5 mM BME, and 2 mM ATP were either loaded separately or together onto a Superdex 200 10/300 (GE Healthcare) column equilibrated in the above buffer. All protein concentrations were 35 μM. Samples were analyzed using a Dawn HELEOS II multi-angle light scattering unit (Wyatt Technology).

**Alkylation and Mass Spectrometry**

Cysteine mutants of Get4/5 were reduced with 2.5 mM TCEP at RT for 2 h, followed by the addition of 100 μM NEM. The reaction was quenched with 50 mM DTT at various time points, concentrated under vacuum, redissolved in 0.2% formic acid, and 25 pmol protein was analyzed on an LC-MSD SL 1100 series (Agilent). The samples were chromatographed on a 2.1 x 150 mm Zorbax 300SB-C3 column (Agilent) using a gradient consistent of 0.2% formic acid and 0.2% formic acid in acetonitrile (89.8%) and methanol (10%). The m/z of the intact proteins were measured in the single quadrupole, deconvoluted, and quantified using the ChemStation software (Agilent). Control experiments, where different ratios of un-alkylated and alkylated proteins were mixed and subjected to MS analysis, showed the quantification of ratio of alkylated species to be reliable [70].

**Get4/5 PEGylation Assay**

Get4/5 Q34C/C177T containing a 6xHIS tag was reduced with 2.5 mM TCEP in Get buffer (without DTT) at RT for 2 hours. For each PEGylation reaction, 0.45 μM Get4/5 alone or in the presence of 2 μM Get3 was pre-incubated in 2mM ATP for 10 minutes to allow for complex formation, followed by the addition of 60 μM PEG-maleimide (10,000 Da Mw conjugates, SIGMA). The reaction was quenched with 9.0 mM DTT at the indicated time points, and the extent of Get4/5 PEGylation was followed by Western Blot with an anti-HIS antibody (Qiagen). The ratio of un-modified Get4/5 to PEGylated Get4/5 was determined using ImageQuant software (GE Healthcare).

**Fluorescent labeling**

Get4/5 S48C/C177T and Get4/5N S48C/C177T were labeled with thiol-reactive acrylodan. Protein was dialyzed in labeling buffer (50 mM KHepes, pH 7.0, 300 mM NaCl)

and treated with 2 mM TCEP to reduce the disulfide bonds. The labeling reaction was carried out using a 10-30 fold excess of dye over protein. The reaction was incubated overnight at 4 °C and stopped by adding 2 mM DTT. Excess dye was removed by gel filtration using Sephadex G-25 (Sigma) equilibrated with GET buffer.

**Association and Dissociation rate measurements**

All rate measurements were performed on a Kintek stopped-flow apparatus. For association rate measurements, acrylodan-labeled Get4/5 was held constant at 0.2 μM, Get3 concentration was varied as indicated, and ATP was present at 2 mM. Observed rate constants ($k_{obsd}$) were plotted as a function of Get3 concentration and fit to the following equation (Eq. 1),

$$k_{obsd} = k_{on}[\text{Get3}] + k_{off} \quad (1)$$

in which $k_{on}$ is the association rate constant, and $k_{off}$ is the dissociation rate constant.

For dissociation rate measurements, a pulse-chase experiment was used. A complex between acrylodan-labeled Get4/5 (at 0.15 μM) and Get3 (at 0.3 μM) was preformed by incubation in 2 mM ATP for 10 minutes, followed by addition of unlabeled Get4/5 at 6 μM as the chase to initiate Get3-Get4/5 dissociation. The time course for change in fluorescence ($F_{obsd}$) was fit to the following double exponential function (Eq. 2),

$$F_{obsd} = F_e + \Delta F_1 \times e^{-k_{fast}t} + \Delta F_2 \times e^{-k_{slow}t} \quad (2)$$

in which $F_e$ is the fluorescence when the reaction reaches equilibrium, $\Delta F_1$ and $k_{fast}$ are the magnitude and rate constant of the fluorescence change in the fast phase, and $\Delta F_2$ and $k_{slow}$ are the magnitude and rate constant of the fluorescence change in the slow phase.

**Kinetic Modeling**

The kinetic model was derived using KinTek Explorer Professional Software (Kintek Corporation) [71,72]. Kinetic and equilibrium simulations were carried out using the rate constants and theoretical protein concentrations listed in Figure 1.

**Accession codes**

Atomic coordinates and structure factors of the Get3-Get4/5N intermediate complexes have been deposited in the Protein Data Bank under accession codes 5BW8 (2.8 Å) and 5BWK (6.0 Å).

# Figures

## Figure 3.1

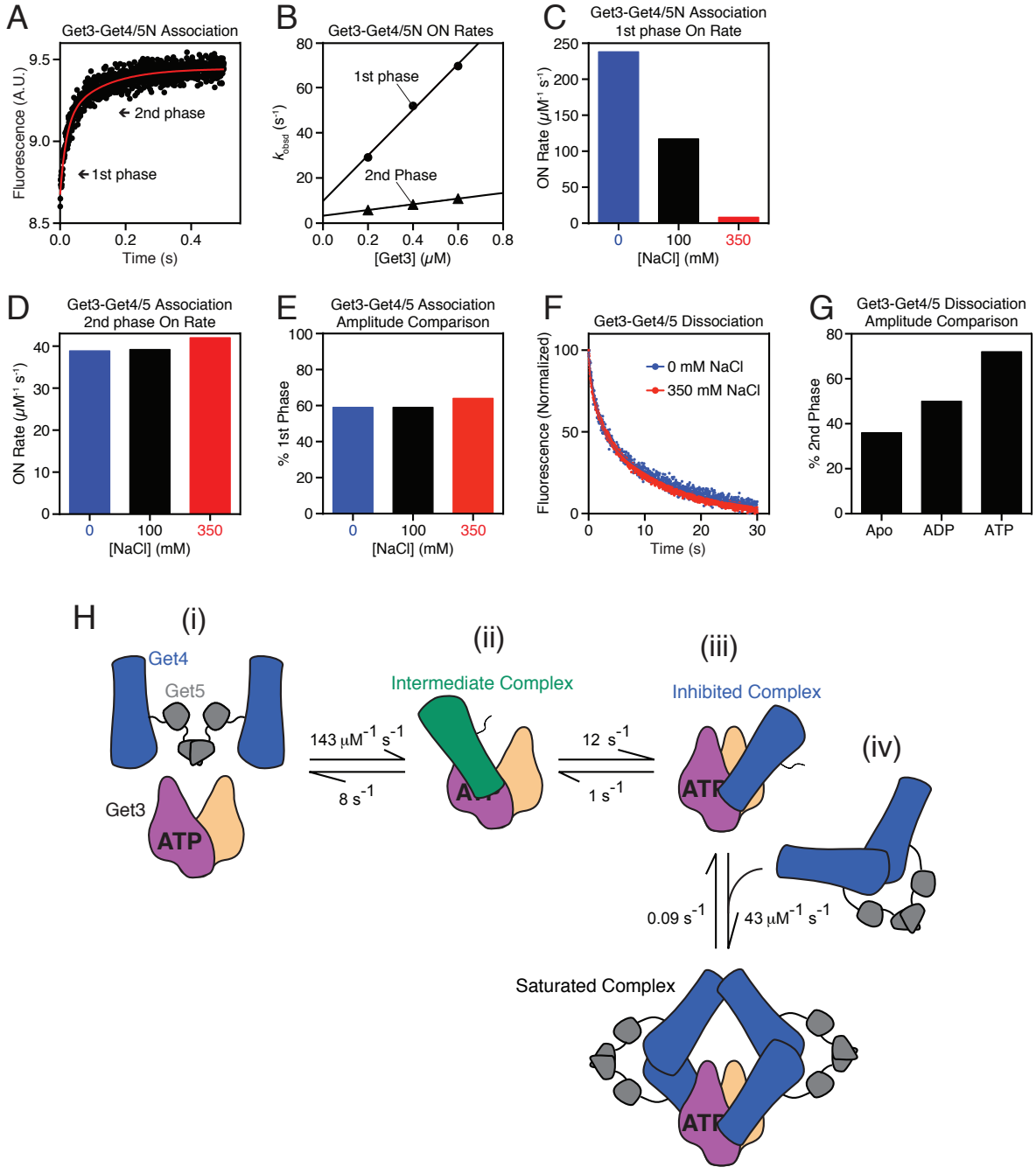


**Figure 3.1. Get4/5 undergoes a conformational change upon binding Get3.**

(A) Time course of Get4/5N binding to ATP-bound Get3. Arrows indicate the two kinetic phases. (B) Observed association rate constants ($k_{obsd}$) are analyzed as a function of Get3 concentration to determine the association rate constant $k_{on}$ for both the first (circles) and second (triangles) kinetic phases. (C) Plots Get3-Get4/5N association rates are highly salt-sensitive. (D) Summary of association rate constants in different salt concentrations for the 2$^{nd}$ kinetic phase in Get3-Get4/5 ON rate measurements. (E) The amplitude of the 1$^{st}$ kinetic phase for Get3-Get4/5 association in different salt concentrations is invariant to buffer ionic strength. (F) Dissociation rate measurement of Get3 from Get4/5 in 350mM NaCl (red) and no salt (blue). (G) Percent amplitude of the 2$^{nd}$ kinetic phase for Get3 dissociation from Get4/5N in different nucleotide states. (H) Proposed Model for Get3-Get4/5 complex association based on kinetic data.

**Figure 3.2**

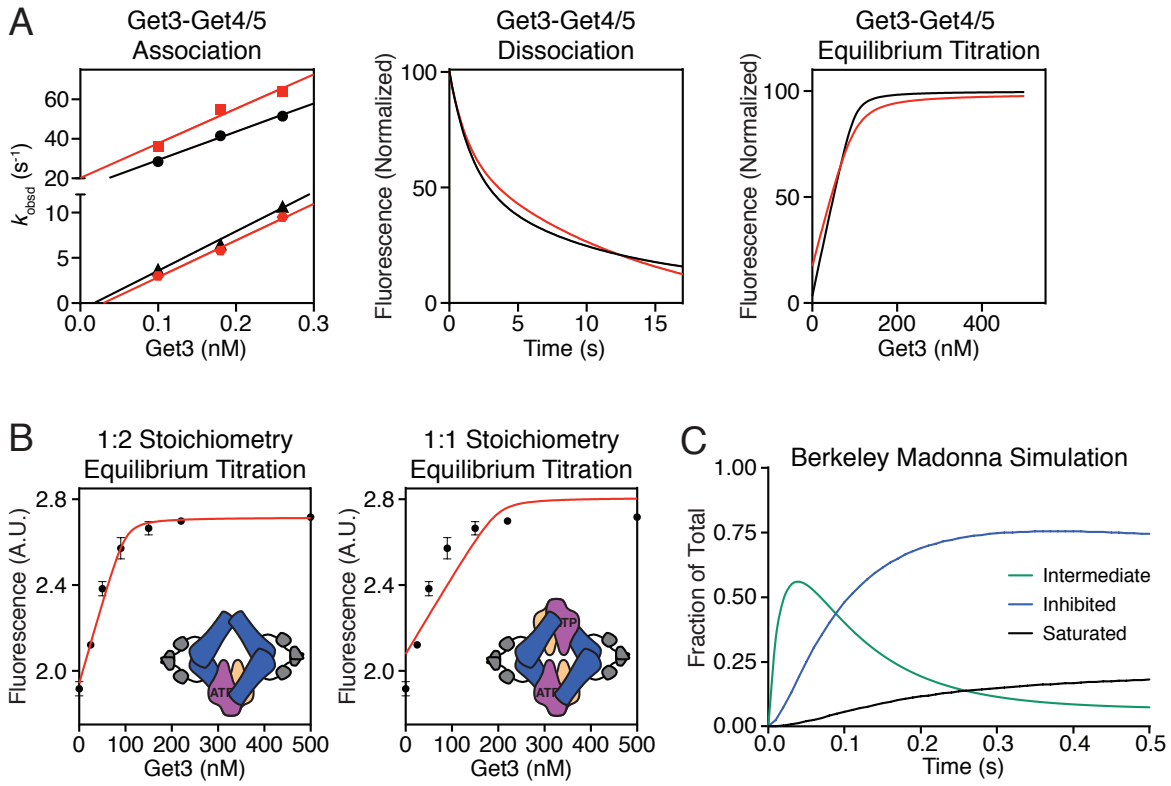



**Figure 3.2. Modeling of experimental kinetic data.**

(A) Comparison of experimental (black) and theoretical (red) binding data for the Get3-Get4/5 complex. Association rate measurements (left), dissociation rate measurements (center), equilibrium titrations (right). All values reported in Table 3.2. (B) Left, model depicting a 1:2 binding stoichiometry of Get3 with Get4/5. This configuration was used to fit an equilibrium titration of ACR-labeled Get4/5-FL (100 nM) with Get3 in 2mM ATP. Right, same as on left but with a 1:1 binding stoichiometry. (C) Berkeley Madonna simulation of Get3-Get4/5 complex formation.

**Figure 3.3**

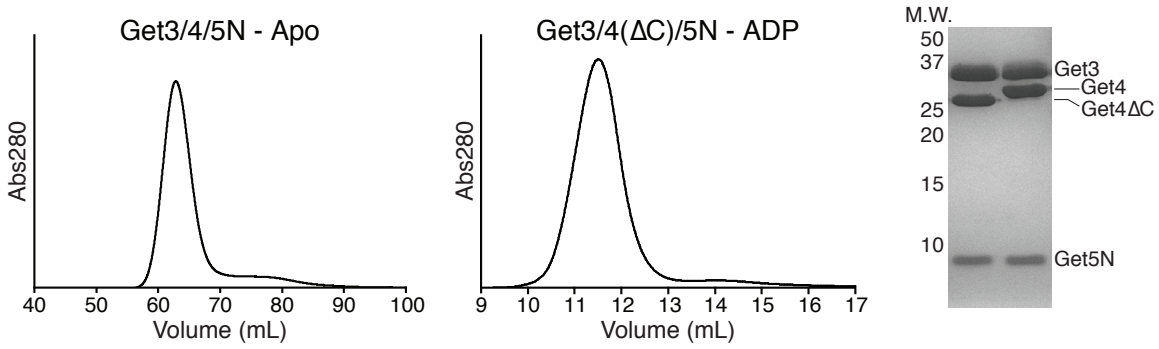



**Figure 3.3. Purification of the Get3-Get4/5N complex.**

Left, SEC of Get3/Get4/5N over a Superdex 16/60 column in the absence of nucleotide. Center, SEC of Get3/Get4ΔC/5N over a Superdex 10/300 column in the presence of ADP. Right, SDS-PAGE of purified complexes.

**Figure 3.4**

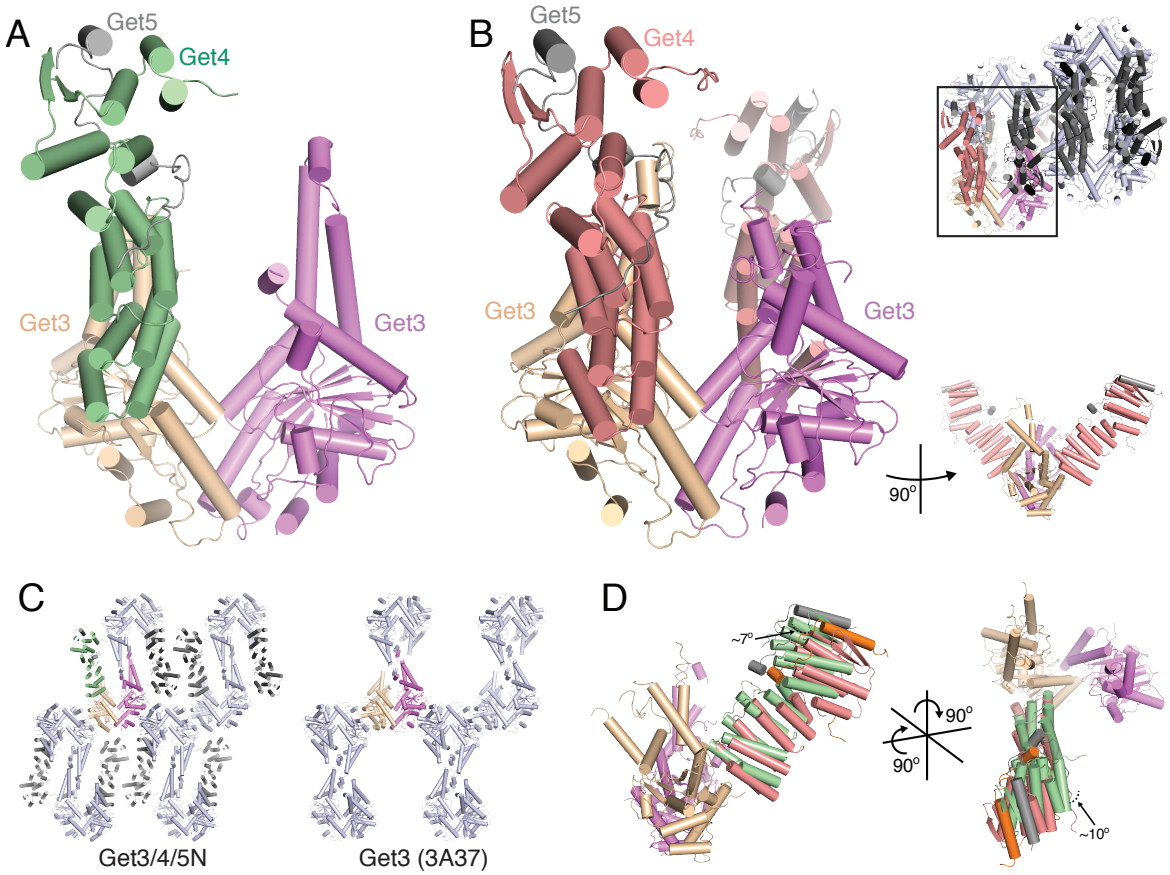

A
Get5
Get4
Get3
Get3
B
Get5
Get4
Get3
Get3
90°
C
Get3/4/5N
Get3 (3A37)
D
~7°
90°
90°
~10°

**Figure 3.4. Crystal structures of a Get3-Get4/5 intermediate complex.**

(A) The asymmetric unit of the 2.8 Å structure of Get4/5N (Get4, green and Get5N, gray) bound to Get3 (wheat and magenta). (B) Left; one subunit from the 6.0 Å structure containing two Get4/5N molecules (Get4, green and Get5N, gray) bound to a Get3 dimer (wheat and magenta). Top right; the asymmetric unit of the 6.0 Å structure of Get4/5N (Get4, red and Get5N, gray) bound to Get3 (wheat and magenta). The black box outlines one Get3 dimer (wheat and magenta) bound to two Get4/5N heterodimers (Get4, red and Get5N, gray). The rest of the asymmetric unit is colored as follows: Get3, light gray; Get4, dark gray; Get5N, black. Bottom right; Additional view of a subunit from the asymmetric unit of the 6Å Get3-Get4/5N complex. (C) Comparison of the crystal packing between 2.8 Å Get3/4/5N structure (left) and the apo Get3 structure (right, PDB ID 3A37). (D) Two views showing the relative orientation of Get4 monomers from the 2.8Å (green) and 6.0Å (red) crystal forms. Alignment was based on helices α10 and α11 from the contacting Get3 monomer.

**Figure 3.5**

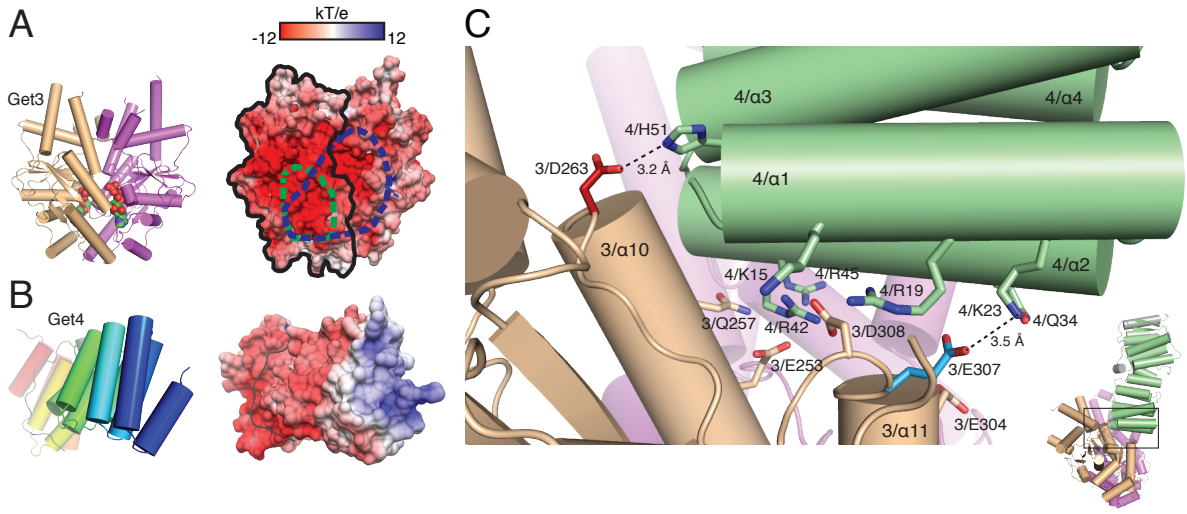



**Figure 3.5. Surface properties of Get3 and Get4.** (A) Left, orientation of Get3 used in subsequent panels to highlight binding interface. For surface representations one monomer is outlined in black and the interaction surface of Get4 on Get3 is highlighted by dashed lines for intermediate complex (green) and inhibited complex (blue). Right, accessible surface color ramped based on electrostatic potential from -12 kT/e (red) to +12 kT/e (blue). (B) Surface properties of Get4 as in (A), except Get4 is color-ramped from N-terminus (blue) to C-terminus (red), and Get5N is colored gray. (C) View of the interface from the 2.8 Å structure showing interactions between Get4 (green) and Get3 (wheat and purple). Residues colored D263 (red) and E307 (blue) were tested for interaction and colored based on phenotype. These residues are unique to this interface and are not found in the ATP-bound structure [65]. Dashed lines represent potential interactions and are labeled with their atomic distances.

**Figure 3.6**

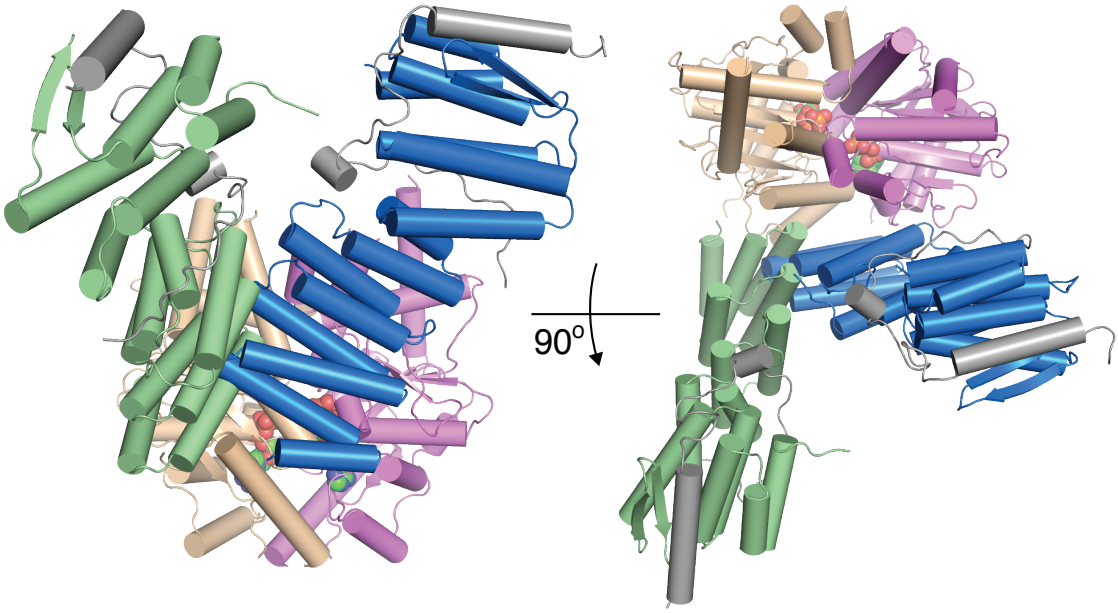
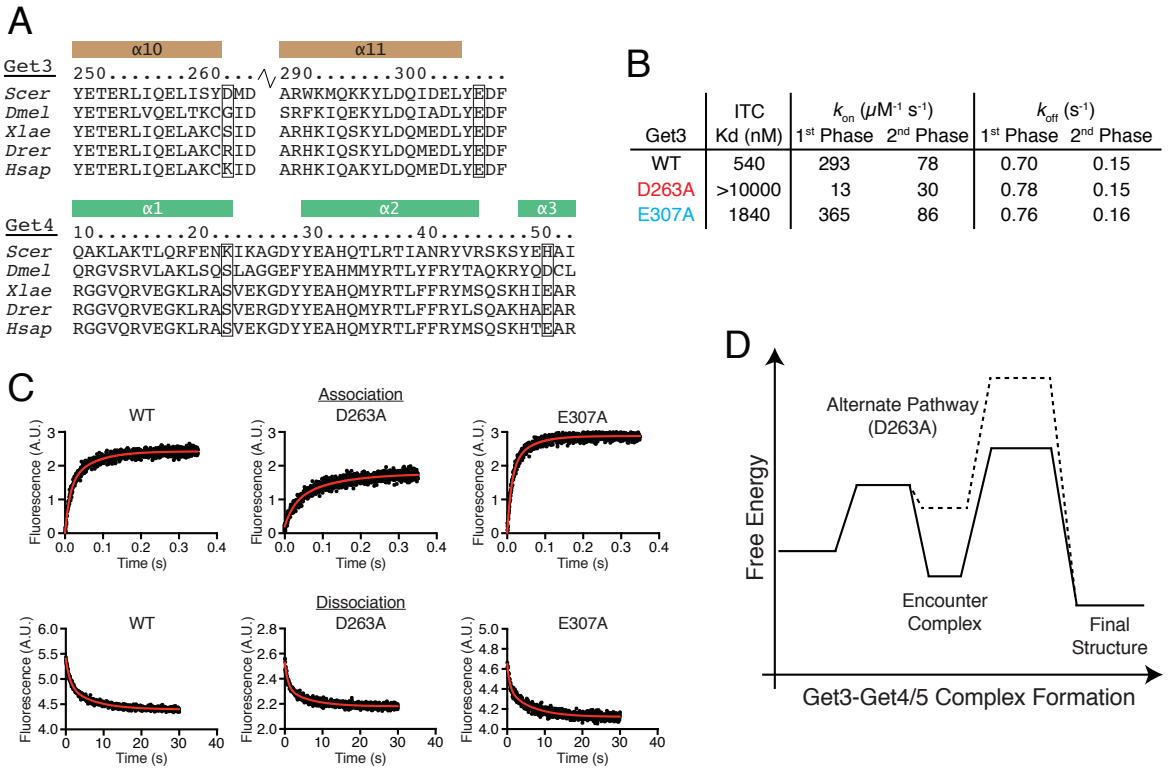



**Figure 3.6. Structural overlay comparing the intermediate and inhibited interfaces.**

Two views of the overlay of Get4 from the intermediate complex (green) and inhibited complex (blue) on the closed Get3 dimer from CC. The Get4 from OC was orientated by aligning helices α10 and α11 on one monomer of Get3.

**Figure 3.7**

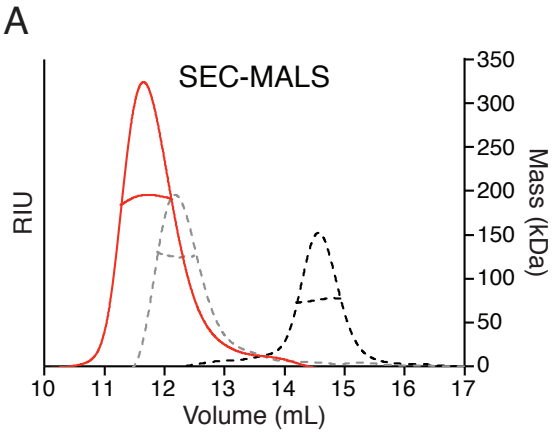



| Get3 | ITC Kd (nM) | $k_{on}$ ($\mu M^{-1}$ $s^{-1}$) 1st Phase | $k_{on}$ ($\mu M^{-1}$ $s^{-1}$) 2nd Phase | $k_{off}$ ($s^{-1}$) 1st Phase | $k_{off}$ ($s^{-1}$) 2nd Phase |
|---|---|---|---|---|---|
| WT | 540 | 293 | 78 | 0.70 | 0.15 |
| D263A | >10000 | 13 | 30 | 0.78 | 0.15 |
| E307A | 1840 | 365 | 86 | 0.76 | 0.16 |

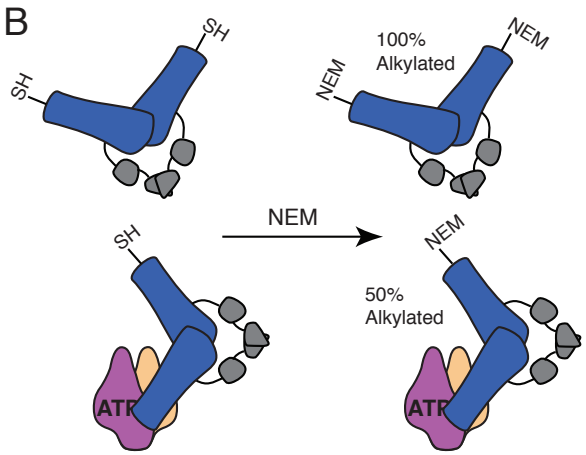
**Figure 3.7. Get3 D263A is defective for complex formation.**

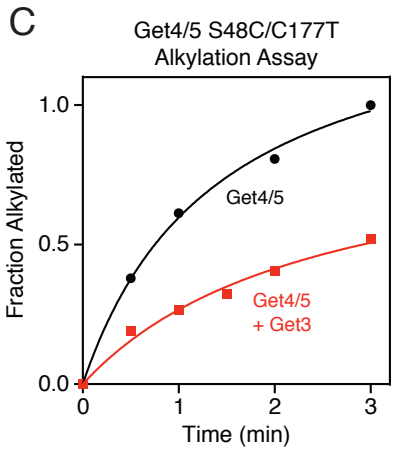
(A) Sequence alignments of regions involved in contacts in the Get3-Get4 interface using ClustalW [61]. Sequences are: Scer – *S. cerevisiae*, Dmel – *Drosophila melanogaster*, Xlae – *Xenopus laevis*, Drer – *Danio rerio* and Hsap – *Homo sapiens*. Helices are indicated above the sequence and labeled. Residues enclosed in boxes tested are unique to the OC interface and tested . (B) Summary of the data obtained by ITC, pulldown assays, and kinetic experiments. ITC data was generated from a single experiment; pulldown experiments were performed in triplicate. (C) Top, Representative association rate measurements of Get3 at 260 nM (wildtype, E307A) and 500 nM (D263A). Each plot represents the average of ten experiments. Bottom, Representative dissociation rate measurements of Get3 (wildtype, E307A, and D263A). Each plot represents the average of 3 experiments. (D) Model for effects of Get3 D263A mutation on Get3-Get4/5 complex formation.

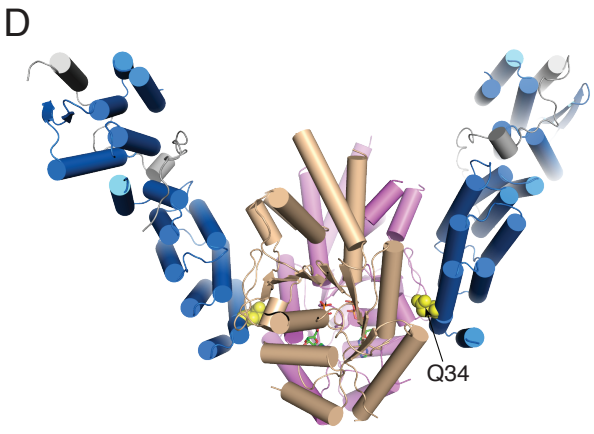
**Figure 3.8**




| Sample | M.W. (kDa) Calculated | Experimental |
|---|---|---|
| ScGet3 | 79.04 | 76.50 |
| ScGet4/5 | 119.56 | 122.61 |
| ScGet3/4/5 | 198.60 | 196.43 |



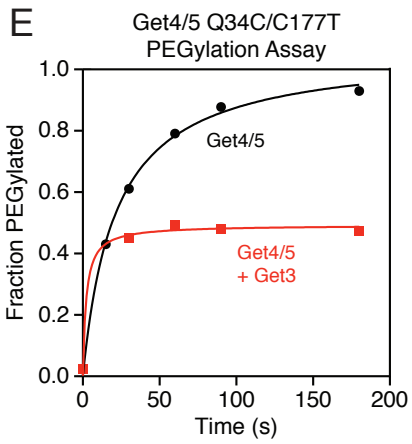



**Figure 3.8. Get3 binds to one-half of the Get4/5 heterotetramer.**

(A) SEC-MALS analysis of the Get3/4/5 complex. (B) Model depicting the NEM accessibility of a solvent exposed Cys residue (S48C) on Get4/5 alone and in complex with Get3. (C) Results of the Alkylation assay shown in A. (D) ATP-bound structure of Get3/4/5N (PDB ID 4PWX) demonstrating the location of Q34C on the Get3-Get4/5 structure. (E) Results of a PEGylation assay with Get4/5 alone or in presence of Get3.

## Tables

**Table 3.1. Summary of crystallographic data**

| Dataset | Intermediate Complexes | |
|---|---|---|
| | 2.8Å | 6Å |
| Wavelength (Å) | 1.0000 | 0.9795 |
| Resolution range (Å) | 29.8–2.8 (2.9–2.8)[a] | 30.0–6.0 (6.6–6.0) |
| Space group | P $2_12_12$ | P $2_1$ |
| Cell parameters: | | |
| a, b, c (Å) | 112.0, 238.0, 52.2 | 152.4, 127.3, 210.3 |
| *α, β, γ* (°) | 90, 90, 90 | 90, 110.2, 90 |
| Unique reflections | 34114 (4406) | 18388 (4452) |
| Completeness (%) | 97.0 (95.7) | 96.3 (98.7) |
| Redundancy | 3.3 (3.1) | 3.2 (3.4) |
| $R_{merge}$[b] | 0.06 (0.58) | 0.06 (0.78) |
| Mean I/σ(I) | 8.3 (1.6) | 8.9 (1.4) |
| **Refinement** | | |
| Reflections: work/free | 32332/1710 | 17462/917 |
| $R_{work}/R_{free}$ | 0.224/0.261 | 0.274/0.303 |
| No. protein atoms | 7750 | 41984 |
| No. ligand atoms | 1 | 4 |
| Protein B-factors (Å$^2$) | 111.0 | 361.8 |
| Water/ligand B-fact (Å$^2$) | 111.1 | 357.5 |
| rmsd of bond lengths (Å) | 0.0148 | 0.0031 |
| rmsd of bond angles (°) | 1.70 | 0.74 |
| Ramachandran preferred | 96.0% | 91.0% |
| Ramachandran allowed | 3.3% | 6.7% |
| Ramachandran outliers | 0.7% | 2.3% |

[a]Values in parentheses are for the highest resolution shell

[b]$R_{merge} = \Sigma hkl\ \Sigma_i\ |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma hkl\ \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is *i*th observation of reflection hkl and $\langle I(hkl)$ is the weighted average intensity for all observations *i* of reflection hkl.

*Chapter 4*

# CONCLUSIONS

In this chapter I will begin by describing our attempts at solving the structure of a full-length Get3-Get4/5 complex. I will then summarize key findings from our research, and highlight areas of study that will help advance our understanding of the role of the Get3-Get4/5 interaction.

## Crystallization of the full-length Get3-Get4/5 complex

Possibly the most surprising result from our studies of the Get3-Get4/5 interaction was that only one binding site on the Get4/5 heterotetramer was occupied by Get3. This asymmetrical binding was unexpected because a free Get4/5 heterotetramer forms an extended particle in solution, with two available Get3 binding sites on opposite ends of the molecule [36]. This suggests that Get3 binding to one Get4 forces the Get4/5 heterotetramer to adopt a conformation incompatible with binding to a second Get3. However, given our current understanding of the full-length Get4/5 molecule, and the lack of a full-length Get3/4/5 structure, it is unknown how this could occur. To answer this question, we attempted to determine a high-resolution crystal structure of the full-length Get3/4/5 complex.

Using a similar protocol as for the Get3-Get4/5N intermediate complex [65], I was able to purify the wildtype Get3-Get4/5 complex in the presence of ADP (**Figure 4.1**). Crystallization trials with this complex produced one hit: 0.02 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, and 22% (w/v) Poly (acrylic acid sodium salt) 5100 (**Figure 4.2A**). These crystals fluoresced under UV (**Figure 4.2B**), and grew larger when the crystallization conditions were optimized (**Figure 4.2C**). These crystals were very reproducible, but on

average only diffracted to ~20 Å. However, a single hexagonal crystal formed that diffracted to ~8.5 Å (**Figure 4.2D-E**). Using this dataset, and known high-resolution structures of Get3 and Get4/5N as search models, I attempted to solve the structure using molecular replacement methods [55], but was unsuccessful.

To limit the flexibility within the crystal lattice, 22-residues were truncated from the C-terminus of Get4, which was previously used to produce better-diffracting crystals of Get3-Get4/5N [65]. This led to larger, more three-dimensional crystals (**Figure 4.1F**), but did not improve diffraction. To further stabilize the complex during crystallization trials and improve the diffraction limit of these crystals, I purified the Get3-Get4/5 complex in the presence of ATP using the catalytically inactive ScGet3 D57V (ScGet3D) [65]. Although this complex did not crystallize in the initial condition, crystals did form in an entirely different condition: 8% (v/v) Tacsimate pH 8.0 and 20% (w/v) PEG 3350. These crystals did fluoresce under UV light (**Figure 4.1G**); however they did not diffract x-rays. Due to the large linker regions and inherent flexibility between domains on Get5 [36,40], obtaining a high-resolution structure of the full-length Get3-Get4/5 complex will be difficult using protein crystallography.

**Confirming asymmetry of Get3-Get4/5 binding**

Although crystallizing the full-length Get3-Get4/5 complex may prove too difficult, more analytical biochemical methods may provide an alternate route to study the asymmetry in Get3-Get4/5 binding. It was previously shown that two different constructs of Get4/5 exchange their C-domains to form mixed Get4/5 dimers [36]. Thus, one can use this method to generate an asymmetric Get4/5 molecule, containing one wildtype Get4 and one mutant Get4 defective for binding Get3. This chimeric Get4/5 can then be used in the same kinetic experiments as wildtype Get4/5. If the same biphasic association and dissociation are observed with the

chimeric Get4/5, then the asymmetry in Get3-Get4/5 binding observed in our initial experiments is accurate.

**Molecular details of Get3 ATPase inhibition**

The 5.4 Å ATP-bound Get3-Get4/5 structure revealed two functionally distinct binding interfaces for anchoring and ATPase regulation [65]. The anchoring interface was demonstrated to mediate the interaction between Get3-Get4, while the regulatory interface is critical for inhibition of Get3 ATP activity. The proximity of the interaction between Get3 Lys69 and Get4 Asp74 to the catalytic Get3 D57 suggests that allosteric effects lead to inhibition of Get3 ATPase activity. However, the sub-atomic resolution prevents direct visualization of these changes in our electron density reconstruction [65].

Unfortunately, a careful examination of the available Get3 structures in different nucleotide states suggests that crystallography might not provide us with an answer. Get3 has been solved in complex with AMP-PNP [25], ADP-AlF4 [24], and ADP [23,25]. Although obvious conformational changes occur when compared to apo Get3 [23,27], there is virtually no difference in the active sites of the nucleotide bound structures. The problem is further complicated if these changes are due to altered side chain conformations, and not changes in the backbone. First, this would require high-resolution data to accurately model the side chains within the active site. Second, this would require wildtype Get3 to be used for crystallization, as mutants would bias the interpretation of the active site sidechains. Both of these requirements suggest that an alternate route to crystallography is required to understand the molecular details of Get4-mediated Get3 ATPase inhibition.

**Alternate pathways for complex formation**

In addition to the inhibited complex described in chapter 2, we solved crystal structures of the Get3-Get4/5 complex in an alternate conformation. These structures represent an initial binding interaction mediated by electrostatics that facilitates the rate of subsequent inhibited complex formation. This so-called intermediate complex is supported by a thorough kinetic analysis of Get3-Get4/5 complex formation confirming the two-step complex formation. Similar mechanisms have been seen in other systems such as Barnase-Barstar [67], and the co-translational targeting machinery [73]. In a role analogous to Get4/5, the SRP RNA not only prevents the premature dissociation of SRP and FtsY (Get3 and ATP), but also plays an active role in ensuring the formation of productive assembly intermediates (intermediate complex), thus guiding the SRP and FtsY through the most efficient pathway of assembly (inhibited complex and Get3-TA complex).

While the intermediate complex definitely forms in yeast, it remains to be seen whether the same biphasic Get3-Get4/5 association occurs in other organisms. Sequence analysis of Get3 and Get4 homologues suggest that the yeast Get3 D263-Get4 H51 interaction is lost in flies and frogs, but then acquired again in fish and mammals (**Figure 3.7A**). This raises multiple questions: Is the intermediate complex essential for efficient TA targeting? Do other organisms form an initial Get3-Get4/5 complex different from that in yeast? Only a more comprehensive kinetic analysis of Get3 homologues from higher eukaryotes can answer these questions.

## Methods

### Protein cloning, expression, and purification

The sequences of Get4 and Get5 were cloned as previously described [36]. In this study, Get4/5 was either full-length or modified by truncating the C-terminus of Get4 (residues 291-312) [65]. All Get4/5 proteins were overexpressed in BL21-Gold (DE3) (Novagen) grown in 2xYT media at 37 °C and induced for 3h by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were lysed using a microfluidizer (Microfluidics) and purified as a complex by Ni-affinity chromatography (Qiagen). The affinity tag was removed by an overnight TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 7.5, 30 mM NaCl, and 5 mM β-mercaptoethanol (BME). A second Ni-NTA column was used to remove any remaining his-tagged protein, and the sample was then loaded onto a 6 mL Resource Q anion exchange column (GE Healthcare). The peak containing the full-length Get4/5 complex was collected and concentrated to a final volume of ~5 mL. This sample was further purified using a Superdex 200 16/60 column (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 100 mM NaCl, and 5 mM BME. Fractions containing Get4/5 were pooled and concentrated to ~5 mg/mL.

The *S. cerevisiae* Get3 coding region was cloned as previously described [23]. A 6xHis-tag followed by a tobacco etch virus (TEV) protease site was fused to the N-terminus, and a stop codon was placed in front of the C-terminal 6xHis-tag. All *S. cerevisiae* Get3 mutants were generated using the QuikChange method. All Get3 proteins were made in BL21-Gold(DE3), grown in 2xYT media, and induced with 0.5 mM IPTG for 16h at 22$^{o}$C. Cells were lysed using a microfluidizer (Microfluidics) and purified by Ni-affinity chromatography (Qiagen). The affinity tag was removed by an overnight TEV protease digest at room temperature while dialyzing against 20 mM Tris pH 7.5, 100 mM NaCl, and 5 mM BME. A second Ni-NTA column was used to remove any remaining his-tagged protein, and the sample

was run on a Superdex 200 16/60 column (GE Healthcare) equilibrated with the dialysis buffer. Fractions corresponding to a dimer of Get3 were pooled and concentrated to 15-20 mg/mL.

Get3-Get4/5 and Get3-Get4ΔC/5 complex was formed by equilibrating 105 μmol Get4/5 with 100 μmol Get3 at room temperature in ~5 mL of 20mM Tris pH 7.5, 10mM NaCl, 5mM BME, 1mM $MgCl_2$, and 1mM ADP. Prior to complex formation, Get3 had been pre-equilibrated with 1mM $MgCl_2$ and 1mM for 5min at room temperature. Get3-Get4/5 complex was further separated from free Get4/5 using a Superdex 200 16/60 (GE Healthcare) equilibrated with 20 mM Tris pH 7.5, 10 mM NaCl, and 5 mM BME. Get3D-Get4ΔC/5 complex was formed as above with ATP substituted for ADP. All complexes were concentrated to 10-12 mg/mL before use in crystallization experiments.

**Crystallization and data collection**

Purified Get3-Get4/5 complex was concentrated to 10-12 mg/ml and crystal trials were carried out using the sitting-drop vapor diffusion method at room temperature by equilibrating equal volumes of the protein complex solution and reservoir solution using a TTP LabTech Mosquito robot and commercially purchased kits (Hampton Research, Qiagen, Molecular Dimensions Limited). The Get3-Get4/5 and Get3-Get4ΔC/5 crystals grew in the presence of 0.02 M Magnesium chloride hexahydrate, 0.1 M HEPES pH 7.5, and 22% (w/v) Poly (acrylic acid sodium salt) 5100. The Get3D-Get4ΔC/5 crystals grew in the presence of 8% (v/v) Tacsimate pH 8.0 and 20% (w/v) PEG 3350. All crystals were cryoprotected by transferring directly to 10μL of a reservoir solution supplemented with 20% glycerol before being flash frozen in liquid nitrogen.

All diffraction data was collected on Beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL). A single Get3-Get4/5 crystal diffracted to ~8.5 Å, and multiple datasets were collected from this crystal. Integration of the data using MOSFLM [50] produced a

9 Å dataset in spacegroup C $222_1$ with the unit cell dimensions: a = 274 Å, b = 390 Å, c = 111 Å. Using this dataset, a structure of the full-length Get3-Get4/5 complex could not be determined by molecular replacement methods [55]. However, a protein crystal with ~50% solvent content with these unit cell dimensions, would have enough space for ~2.5 full-length Get3-Get4/5 complexes in the assymetric unit [52].

# Figures

**Figure 4.1**

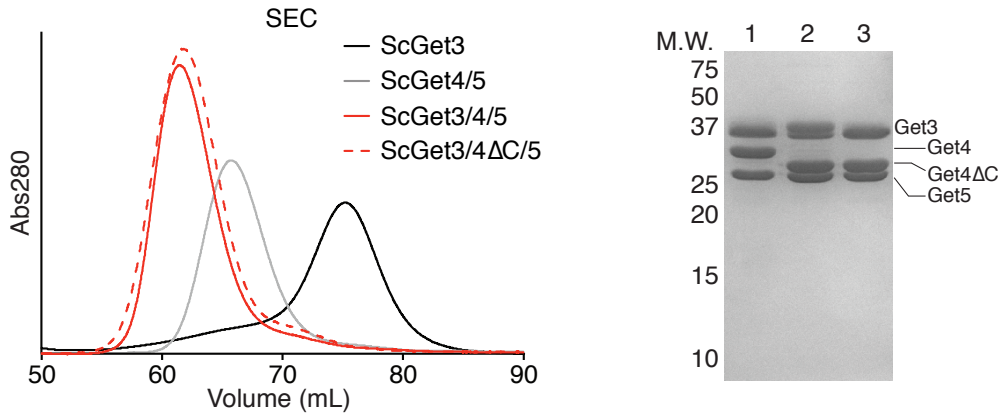



**Figure 4.1. Purification of full-length Get3-Get4/5 complex.** Left; SEC of Get3, Get4/5, Get3-Get4/5, or Get3-Get4ΔC/5. Right; SDS-PAGE corresponding to purified complexes on left. Lane 1 – ScGet3/4/5, Lane 2 – ScGet3/4ΔC/5, Lane 3 – ScGet3(D57V)/4ΔC/5.

**Figure 4.2**

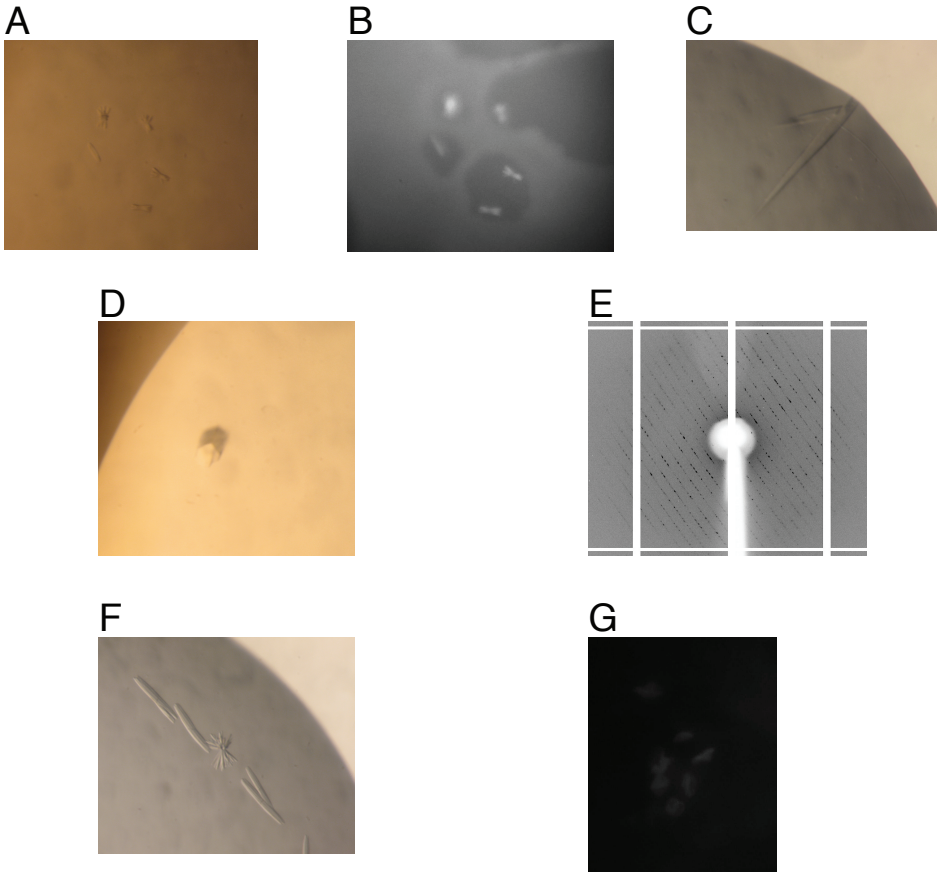



**Figure 4.2. Crystallization of full-length Get3-Get4/5 Complex.** (A) Crystals of full-length Get3-Get4/5 complex from yeast. (B) UV image of same drop in (A). (C) Optimized crystals using similar condition as in C. (D) Hexagonal crystals using Get3-Get4/5. (E) Diffraction image of crystal in (D). The shortest dimension of the image is ~8.5 Å. (G) Image of crystals using Get3-Get4ΔC/5. (G) UV image of crystals using Get3D-Get4ΔC/5.

*Appendix A*

# TOWARDS PURIFICATION OF THE GET1/GET2 MEMBRANE PROTEIN COMPLEX

## Introduction

The targeting of membrane proteins to their correct location in the cell is a highly regulated process [44]. The majority of membrane proteins are targeted via the signal recognition particle (SRP), which typically recognizes the initial hydrophobic transmembrane domain (TMD) as it emerges from the ribosome. However, the ubiquitous TA proteins, defined topologically by a single TMD near the C-terminus, are unable to access the SRP pathway. Their targeting information, contained solely within their TMD, emerges from the ribosome only after protein synthesis is complete and must therefore be targeted to the ER post-translationally [8,9]. In eukaryotes, TA proteins account for at least 1% of the proteome, and are involved in many essential cellular processes such as apoptosis, vesicle fusion, and protein trafficking [8].

A dedicated TA targeting pathway was first described almost twenty years ago [8]. A series of genetic and biochemical experiments in eukaryotes identified the components involved in this pathway, and is now recognized as the Guided Entry of TA proteins (GET) pathway [19,20,22]. In yeast, this pathway consists of at least six proteins, Get1-5 and Sgt2, all with homologs in higher eukaryotes [47]. The first committed step in TA targeting is the formation of the Sgt2/TA complex. TA substrate is then transferred to Get3 in a Get4/5-dependent manner [34]. Extensive structural characterization of Get3, the central TA targeting factor, demonstrated that it undergoes ATP-dependent conformational changes from an open to closed form required for capturing the TA substrate [23-28,47].

Once formed, this Get3-TA complex is then localized to the ER by the membrane proteins Get1/Get2 (Get1/2), which stimulate release of the TA protein and subsequent

insertion into the ER membrane [29-32]. Recent cross-linking studies have demonstrated a physical link between the TMD of a TA protein and Get1/2, suggesting this complex actively assists the TA during insertion, and not only in Get3 recruitment [74]. However, the exact mechanism for Get1/2 function within this pathway is still largely unknown. This section describes the expression and purification of Get1 and Get2 using an *E.coli* expression system in an attempt to obtain a high-resolution crystal structure of the Get1/2 complex.

## Results/Discussion

### Expression and purification of the individual components Get1 and Get2 from *Saccharomyces cerevisiae*

Multiple reports have demonstrated that Get1 and Get2 from *S. cerevisiae* can be recombinantly expressed and purified from *E.coli*, and then reconstituted in proteoliposomes to form a functional complex [29,31,74]. For this reason, we used *E.coli* as an expression system to generate Get1/2 for crystallization trials. For initial experiments, the full-length Get1 or Get2 was cloned into pET33b containing an N-terminal 6xHis-tag and TEV cleavage site and expressed in Rosetta2 pLysS cells using the autoinduction system [75]. Although this system had been previously used to express Get1 and Get2 to milligram quantities, our constructs did not express (**Figure A.1A-B**). Comparing our constructs to the ones used previously [29] revealed that the sequence encoding our 6xHis-tag was poly-CAC, whereas the constructs in previously used in pET28 contained a poly-CAT sequence. Mutating our sequence to a poly-CAT lead to a drastic increase in expression and yields of 1-2 mg/L of total protein (**Figure A.1C**). These fractions were then combined and run over SEC, with both Get1 and Get2 purifying over multiple elution volumes suggesting that these proteins exist as a heterogeneous population (**Figure A.1D**). In addition, greater than 75% of the sample precipitated out of solution during concentration prior to loading onto SEC. The apparent instability and presence of multiple species for both Get1 and Get2 could be the reasons that SEC with Get1/2 has never been reported. To increase both purity and yield, a 10xHis-tag was introduced in place of the 6xHis-tag (**Figure A.2A-B**).

One reason for Get1/2 crashing out of solution during concentration is that they are inherently instable on their own, but form a stable complex together. To test this, individually purified full-length Get1 and Get2 were incubated together prior to concentration. Once again both proteins precipitated during concentration meaning they were unable to form a complex

under these conditions or the complex is also unstable. To further stabilize Get2, the initial 150 residues were truncated from the N-terminus and replaced with either a 10xHis-tag or an MBP-fusion. These constructs, expressed to similar levels as full-length Get2 (**Figure A.2C-D**), but were still unstable during concentration. Using more tightly controlled conditions I repeated the reconstitution experiments with the purified components (**Figure A.2**) by varying temperature and incubation times. However, the samples were unstable and crashed out under all of these conditions.

One disadvantage in starting with purified components is that Get1 or Get2 is in a state that is unfavorable for complex formation. To test this, cells expressing 10xHis-Get1 and MBP-Get2 were mixed prior to lysis and extracted and purified together. In this assay, the majority of MBP-Get2$_{TM}$ did not form complex (**Figure A.3A-B**). Until this point, all of the above purifications were carried in the presence of LDAO. Therefore, a detergent screen was performed to find a different detergent that stabilized complex formation. Using the same assay described above with 10xHis-Get1 and MBP-Get2$_{TM}$, the detergents Fos-Choline 12 (FC12) and Anzergent 3-12 increased extraction efficiency and final protein yields, however there was only a minor increase in complex formation (**Figure A.3C-D**). This assay was then scaled up and repeated at varying temperatures using the detergents FC12, Anzergent 3-12, and LDAO, all which produced similar results (**Figure A.4**). Anzergent 3-12 contains a very charged head group making it somewhat unsuitable for protein purification, whereas FC12 has been increasing in popularity due to its structural similarity to lipids found in the membrane bilayer and increasing success in solving membrane protein crystal structures [76]. Based on these results, FC12 appeared to be the best detergent for future experiments.

**Co-expression and purification of the fungal Get1/2 complex**

Since reconstituting the Get1/2 complex from the individual components was

unsuccessful, co-expression trials were carried out. Initially, Get1/2 was expressed on the same transcript with 10xHis-Get1 followed by $Get2_{TM}$ under control of one promoter (**Figure A.5A-B**), which did not express (**Figure A.5C**). Interestingly, swapping the orientation of Get1 and $Get2_{TM}$ lead to increase in expression (**Figure A.5D-E**), however the yields were still very low (~50 μg/L). To further increase yields, co-expression was carried out using the Duet system (Novagen). In this system, 10xHis-Get1 was cloned into pRSF-Duet, while MBP-$Get2_{TM}$ was cloned into pET-Duet (**Figure A.5F-G**). MBP-$Get2_{TM}$ expressed as before, however His-Get1 did not express at all (**Figure A.5H**). This is most likely due to the high copy number of pRSF-Duet, which could be detrimental to membrane protein production as the kinetics of insertion are slower than transcription. Based on copy number pACYC-Duet would be a more logical choice instead of pRSF-Duet, unfortunately the Rosetta2 pLysS cells contain a Chloramphenicol marker ruling out the possibility of using pACYC-Duet.

Next, co-expression was carried out using the two-promoter system on the same plasmid. Here, MBP-$Get2_{TM}$ was inserted into MCS-1 in pET-Duet, while 10xHis-Get1 was inserted into MCS-2. In addition to *Sc*, Get1/2 was cloned from *Candida albicans* (*Ca*) and *Aspergillus fumigatus* (*Af*). Although both 10xHis-Get1 and MBP-$Get2_{TM}$ expressed very well for all species, the majority of MBP-$Get2_{TM}$ was present in the Ni-NTA flow through suggesting a stable complex could not be purified (**Figure A.6**). For *Sc* and *Ca*, the Ni-NTA elutions were concentrated and then purified over amylose resin. However, the majority of Get1/2 complex was found in the amylose flow through (**Figure A.6C-D**) suggesting that complex formation was not favored and further validating the results from Figure 6A-B. To determine whether the MBP fusion disrupted Get1/2 complex formation, a 10xHis-tag was used in place of MBP, while Get1 remained untagged (**Figure A.7A-B**). This construct gave better expression of the 10xHis-$Get2_{TM}$, but the final yield was not stoichiometric as Get1 was present in roughly half the amount compared to $Get2_{TM}$ (**Figure A.7C**). To determine whether

this contained stable Get1/2 complex, the elution fractions were concentrated and loaded onto SEC. While the chromatogram contained multiple peaks, the major peak eluted at ~12.8 mL, corresponding to a size of ~80 kDa (**Figure A.7D**). Running the fractions on SDS-PAGE revealed that all of the peaks contained $Get2_{TM}$, but only the major one also contained Get1, suggesting that complex formation occurred, although not stoichiometrically (**Figure A.7E**). Strangely, the *Af* proteins used in the same construct design did not express to a significant amount (**Figure A.7F**). To understand whether the truncated version of Get2 containing only the TMDs was enough for complex formation, we performed the same experiment using full-length Get2. In this assay, Get2 expressed to a lesser amount than $Get2_{TM}$, but the complex appeared to be more stoichiometric (**Figure A.7G**). However, concentration of this sample again lead to precipitation suggesting that complex formation did not occur.

One reason for the inability to purify a stable Get1/2 complex from *E.coli* is that an additional factor only found in eukaryotes in required for complex formation or stabilization. Indeed, this factor may be Get3, whose recruitment to the ER membrane may either stimulate complex formation or stabilize an otherwise weak interaction between Get1 and Get2. To test this, full-length Get1/2 were expressed as in Figure 7E, and untagged Get3 was expressed on pRSF-Duet. This lead to decreased Get1/2 expression (**Figure A.8A**) when compared to the same construct without Get3. In this experiment, both Get1 and Get2 contain Get3-binding domains, which might prevent formation of a stable Get1/2/3 trimer complex and instead would cause formation of a heterogeneous population. To stabilize the Get1/2/3 complex, $Get2_{TM}$ was expressed instead of full-length Get2, but produced similar results as with full-length Get2 (**Figure A.8B**). Overall, this strategy leads to decreased Get1/2 expression (**Figure A.8**) when compared to the same construct without Get3 (**Figure A.7A,E**).

Based on the results from Figure 8, it is unclear whether Get3 is not expressing or simply not forming a stable complex with Get1/2. Although no bands corresponding to Get3

were detected in SDS-PAGE, Get3 expression could not be ruled out because the lack of a tag prevents detection in a western blot. To test whether Get3 expressed in this system, an MBP-Get3 fusion was expressed on pRSF-Duet, along with the Get1/2 constructs in Figure 7. In this experiment, Get3 expressed very well, but did not form a stable complex with Get1/2 (**Figure A.9**). Surprisingly, a large fraction of Get3 co-purified with the membrane, but did not form a stable complex with Get1/2. To test whether Get3 non-specifically interacted with the membrane, the MBP-Get3 fusion was expressed on pRSF-Duet by itself and purified under the same conditions. In this experiment, Get3 again interacted with the membrane to the same extent as with Get1/2, suggesting that this interaction is non-specific (**Figure A.10**). Overall, this strategy leads to a decrease in Get1/2 expression (**Figs. A.8-9**) when compared to the same constructs without Get3 (**Figure A.7**). When taken together, the expression and purification studies described above suggest that Get1/2 expression in *E.coli* is not advantageous and that a eukaryotic expression system should be used.

## Methods

### Protein cloning, expression, and purification

The sequences of Get1 (residues 1-235), Get2 (1-285), and $Get2_{TM}$ (151-285) were cloned into pET33 or pET28 as previously described [29]. These constructs were further modified to contain a 10x-His tag using the QuikChange method. In addition to the his-tagged version, $Get2_{TM}$ was fused to an N-terminal Maltose Binding Protein (MBP) followed by a tobacco etch virus (TEV) protease site in the pMAL vector. Co-expression vectors were generated using the Duet system (Novagen). All constructs were verified by DNA sequencing. All Get1 and Get2 proteins were overexpressed in Rosetta2 pLysS (DE3) (Novagen) using the autoinduction system [75]. A single colony was used to inoculate 50 mL Terrific Broth (TB), and grown at 37 °C to an $OD_{600}$ ~0.4. 10 mL of this starter culture was used to inoculate 1 L autoinduction media and grown at 37 °C for ~16 h. Cells were harvested by centrifugation at 5000 rpm for 20min at 4 °C, resuspended in lysis buffer (50 mM Na-HEPES pH 7.5, 500 mM NaCl, 10 mM β-mercaptoethanol (BME), 20 mM imidazole, and 0.1 mM PMSF), and lysed using a microfluidizer (Microfluidics). Following an initial centrifugation of 14,000 rpm for 30 min at 4 °C, the lysate was collected and centrifuged at 45,000 rpm for 60 min at 4 °C. This membrane pellet was then resuspended in lysis buffer and the membrane fraction solubilized by addition of 1% (w/v) detergent (2% for CHAPS, CHAPSO, and Deoxy-Big CHAPS). Room temperature solubilization was performed for 1 hr, while solubilization at 4 °C was performed for 1-16 h. These samples were then centrifuged for 30 min at 4 °C to remove any precipitate, and the supernatant was incubated with 0.5 mL Ni-NTA resin (Qiagen) for ~2 h at 4 °C. The Ni-NTA resin was collected using a 15 mL disposable column (Biorad), and washed with lysis buffer 0.1% detergent and 30 mM imidazole. The sample was eluted by addition of 5 column volumesof lysis buffer containing 0.1% detergent and 300 mM imidazole. TEV protease treatments were performed overnight at room temperature while dialyzing against 20

mM Na-HEPES pH 7.5, 150 mM NaCl, 5 mM BME, and 0.1 % detergent. A second Ni-NTA column was used to remove any remaining his-tagged protein. Size exclusion chromatography (SEC) analysis of the affinity purified proteins was performed by loading .5 mL sample onto a Superdex 200 10/300 column (GE Healthcare) equilibrated with 20 mM Na-HEPES pH 7.5, 150 mM NaCl, 5 mM BME, and 0.1 % detergent. Protein samples were visualized by coomasie staining with SDS-PAGE or by western blot against the His-tag (Qiagen anti-PentaHis) or MBP-tag (NEB anti-MBP).

# Figures

**Figure A.1**

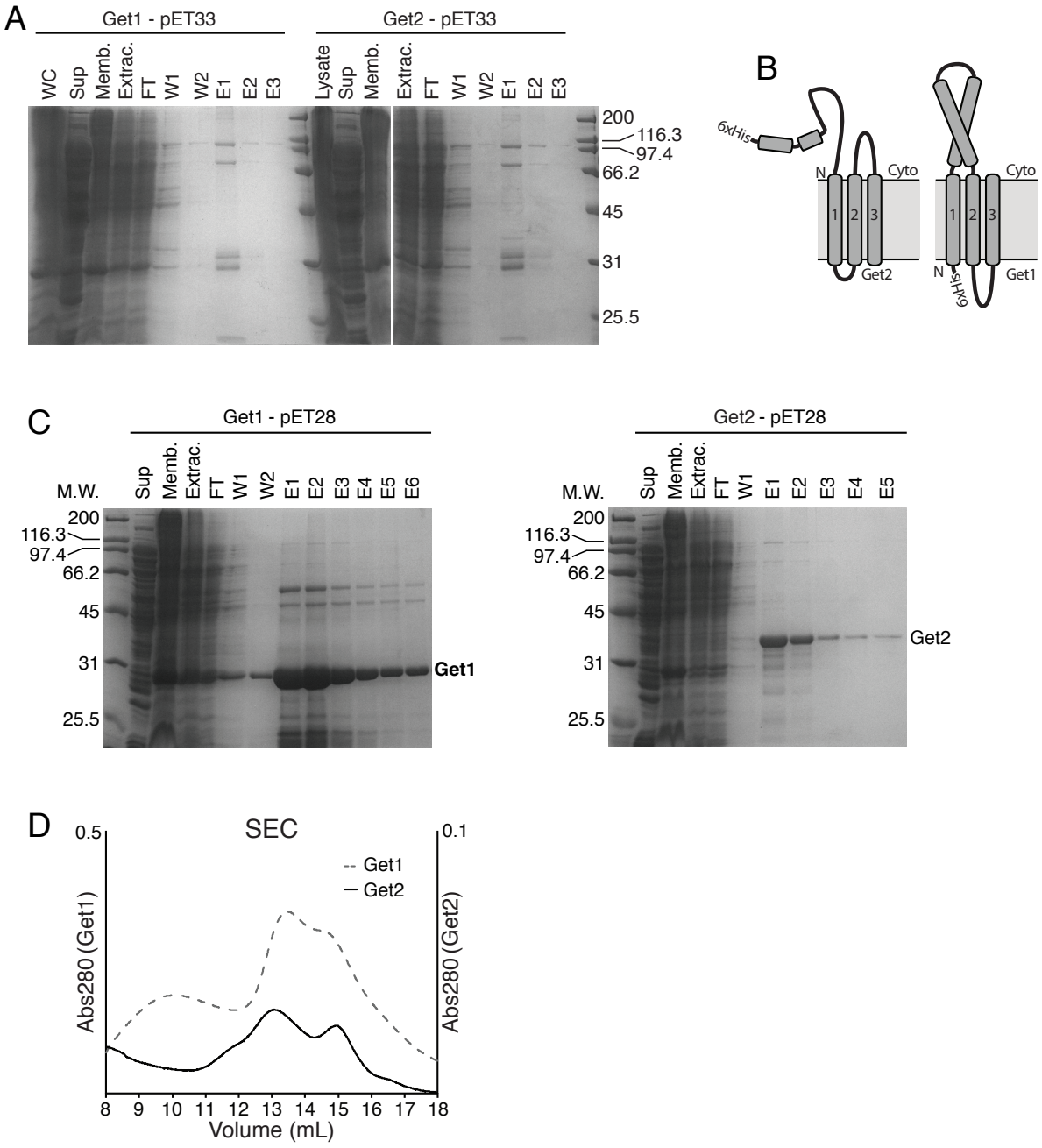



**Figure A.1. Purification of the individual components Get1 and Get2.** (A) Get1 and Get2 expressed from pET33 containing a poly-CAC 6xHis-tag. (B) Cartoon representation of proteins and tags used in the experiment. (C) Get1 and Get2 expressed from pET28 containing a poly-CAT 6xHis-tag. (D) SEC of the individual proteins from (C).

**Figure A.2**

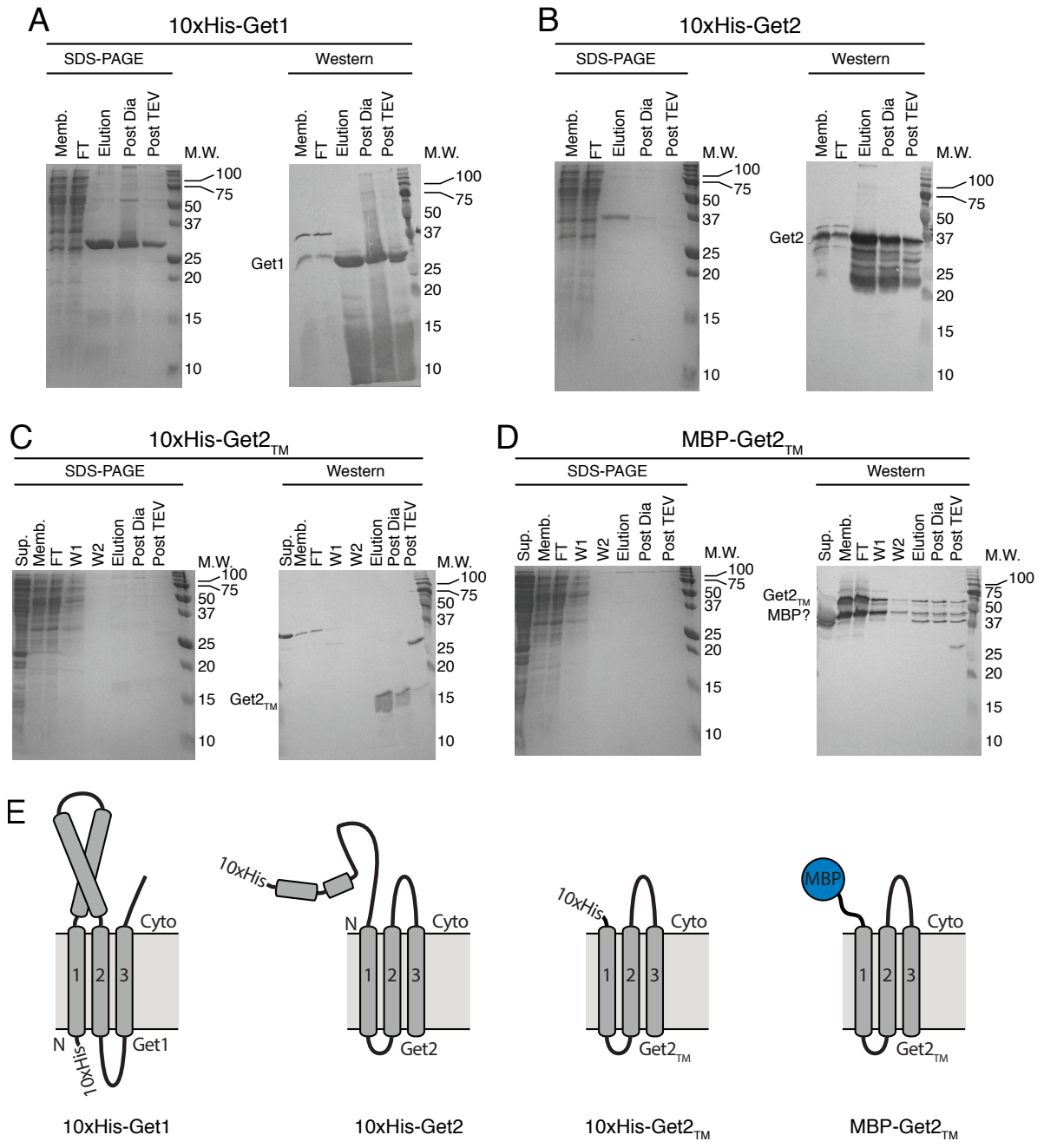



**Figure A.2. Purification of the individual components Get1, Get2, and $Get2_{TM}$ used in reconstitution experiments.** SDS-PAGE and Western blots (Anti-His and Anti-MBP) of (A) 10xHis-Get1 (B) 10xHis-Get2 (C) 10xHis-$Get2_{TM}$ (D) MBP-$Get2_{TM}$. (E) Cartoon representation of proteins and tags used in the experiments.

**Figure A.3**

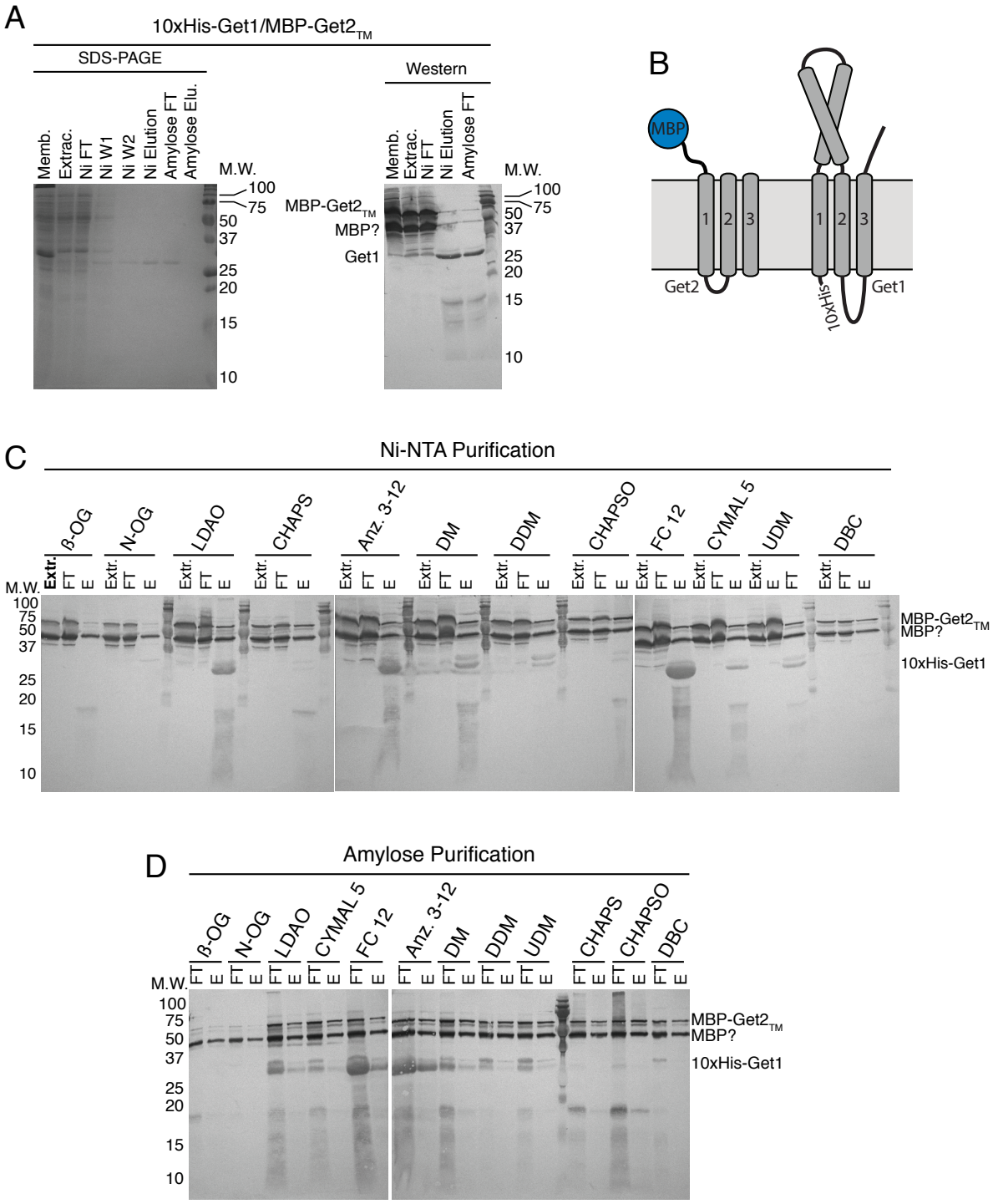

A
10xHis-Get1/MBP-Get2TM
SDS-PAGE
Western
Memb.
Extrac.
Ni FT
Ni W1
Ni W2
Ni Elution
Amylose FT
Amylose Elu.
M.W.
100
75
50
37
25
20
15
10
MBP-Get2TM
MBP?
Get1
B
MBP
1
2
3
Get2
Get1
10xHis
C
Ni-NTA Purification
β-OG
N-OG
LDAO
CHAPS
Anz. 3-12
DM
DDM
CHAPSO
FC 12
CYMAL 5
UDM
DBC
Extr.
FT
E
MBP-Get2TM
MBP?
10xHis-Get1
D
Amylose Purification

**Figure A.3. Detergent screen for Get1/2 complex formation.** (A) Initial purification of 10xHis-Get1 and MBP-Get2$_{TM}$. The proteins were individually expressed and the cell pellets were combined prior to lysis and extraction in LDAO. (B) Cartoon representation of proteins and tags used in the experiment. (C) Ni-NTA purification from detergent screen of 10xHis-Get1 and MBP-Get2$_{TM}$. (D) Amylose purification from detergent screen of 10xHis-Get1 and MBP-Get2$_{TM}$.

**Figure A.4**

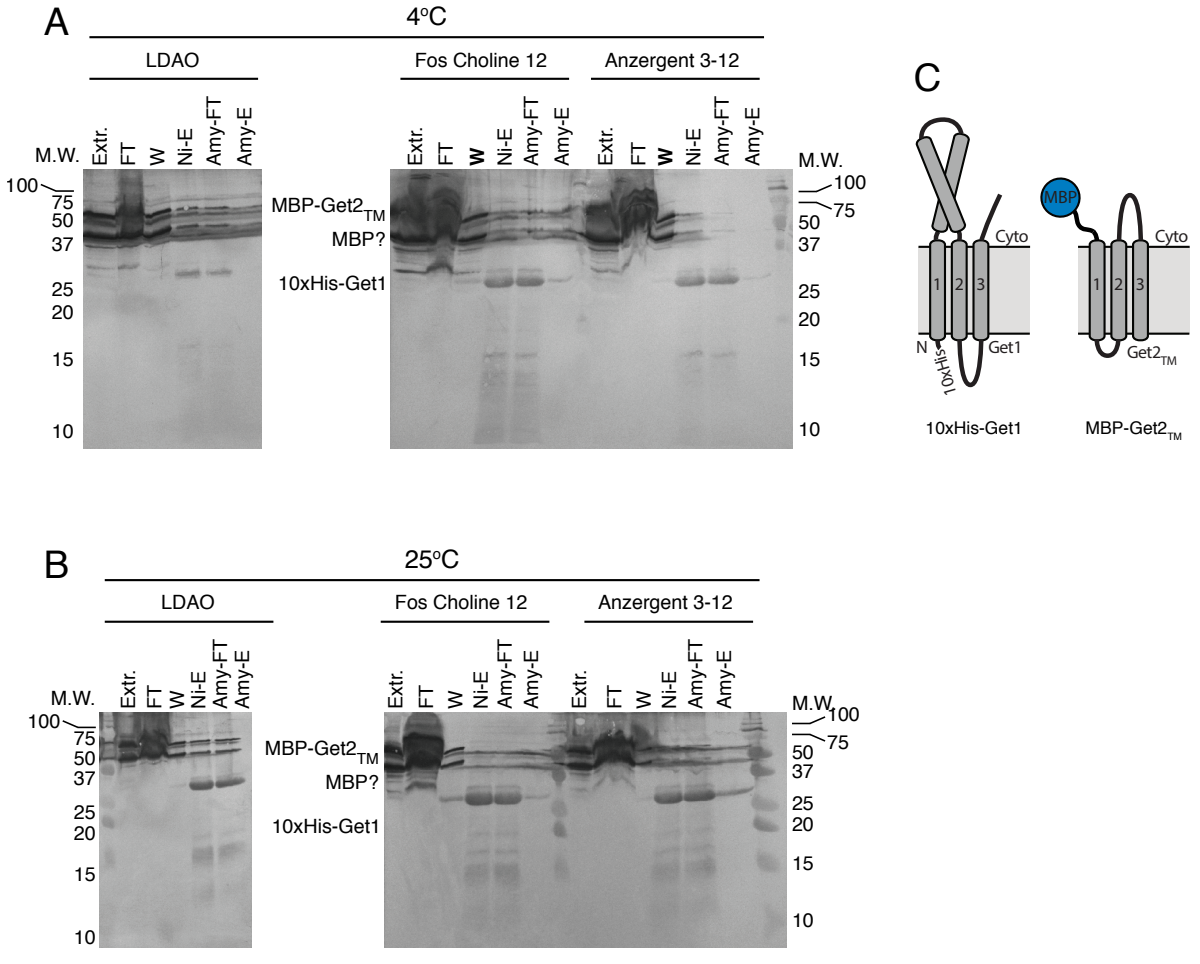



**Figure A.4. Scaled up detergent screen with LDAO, Fos Choline 12, and Anzergent 3-12 for Get1/2 complex formation.** Ni-NTA and amylose purification from detergent screen of 10xHis-Get1 and MBP-Get2$_{TM}$ at (A) 4°C and (B) 25°C. (C) Cartoon representation of proteins and tags used in the experiment.

**Figure A.5**

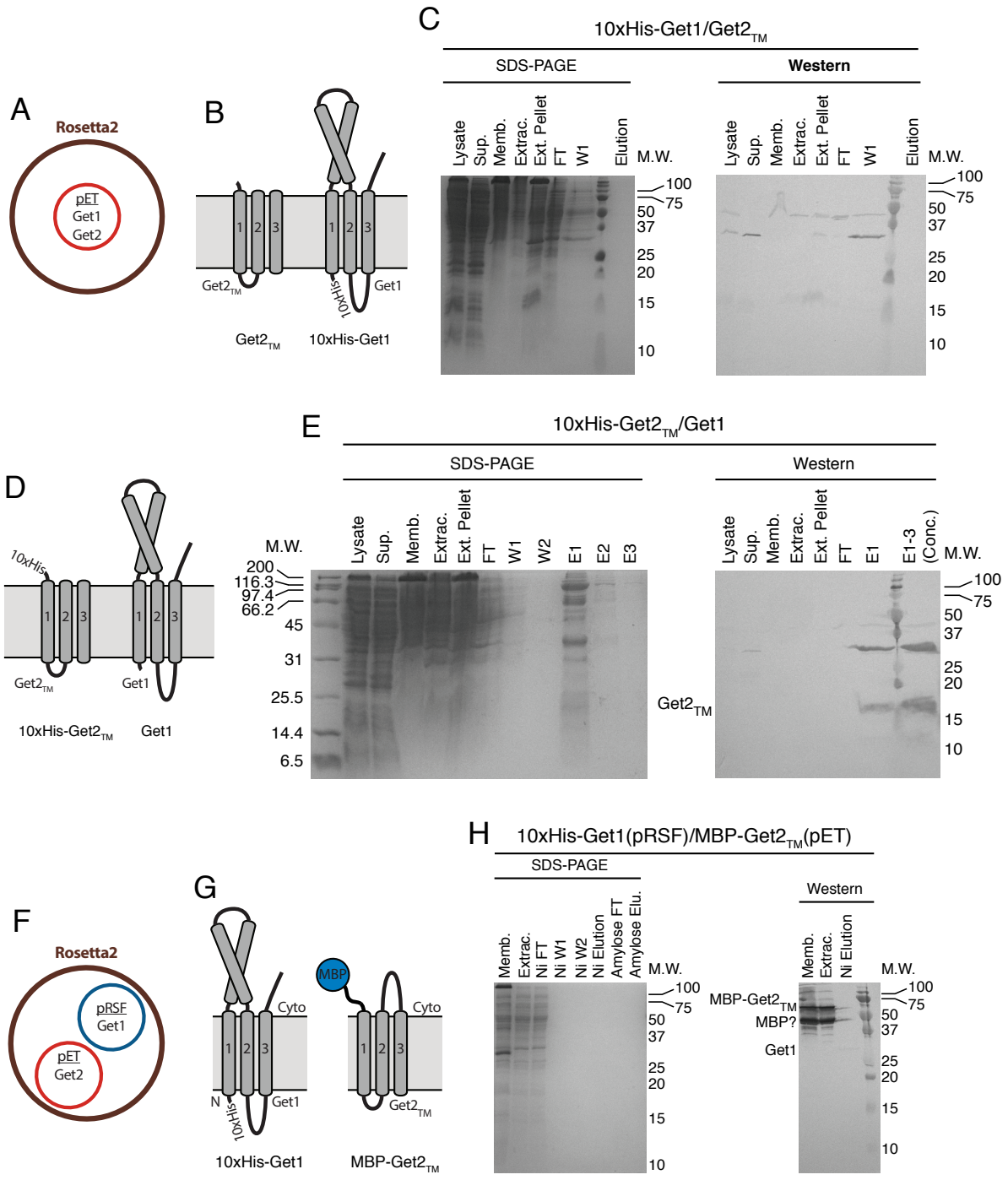

A
Rosetta2
pET
Get1
Get2
B
Get2TM
10xHis-Get1
C
10xHis-Get1/Get2TM
SDS-PAGE
Western
Lysate
Sup.
Memb.
Extrac.
Ext. Pellet
FT
W1
Elution
M.W.
100
75
50
37
25
20
15
10
D
10xHis-Get2TM
Get1
E
10xHis-Get2TM/Get1
SDS-PAGE
Western
Lysate
Sup.
Memb.
Extrac.
Ext. Pellet
FT
W1
W2
E1
E2
E3
E1-3 (Conc.)
M.W.
200
116.3
97.4
66.2
45
31
25.5
14.4
6.5
F
Rosetta2
pRSF
Get1
pET
Get2
G
Cyto
N
10xHis
Get1
MBP
Get2TM
10xHis-Get1
MBP-Get2TM
H
10xHis-Get1(pRSF)/MBP-Get2TM(pET)
SDS-PAGE
Western
Memb.
Extrac.
Ni FT
Ni W1
Ni W2
Ni Elution
Amylose FT
Amylose Elu.
M.W.
MBP-Get2TM
MBP?
Get1

**Figure A.5. Co-expression of Get1/2.** (A-B) Cartoon representation of expression system (A) and the proteins and tags (B) used in the experiment in (C). (C) SDS-PAGE and Western blot of 10xHis-Get1-Get2$_{TM}$. (D) Cartoon representation of proteins and tags used in (E). (E) SDS-PAGE and Western blots of 10xHis-Get2$_{TM}$-Get1. (F) Cartoon representation of the Duet expression system. (G) Cartoon representation of proteins and tags used in (H). (H) SDS-PAGE and Western blots of 10xHis-Get1 (pRSF) and MBP-Get2$_{TM}$ (pET).

**Figure A.6**

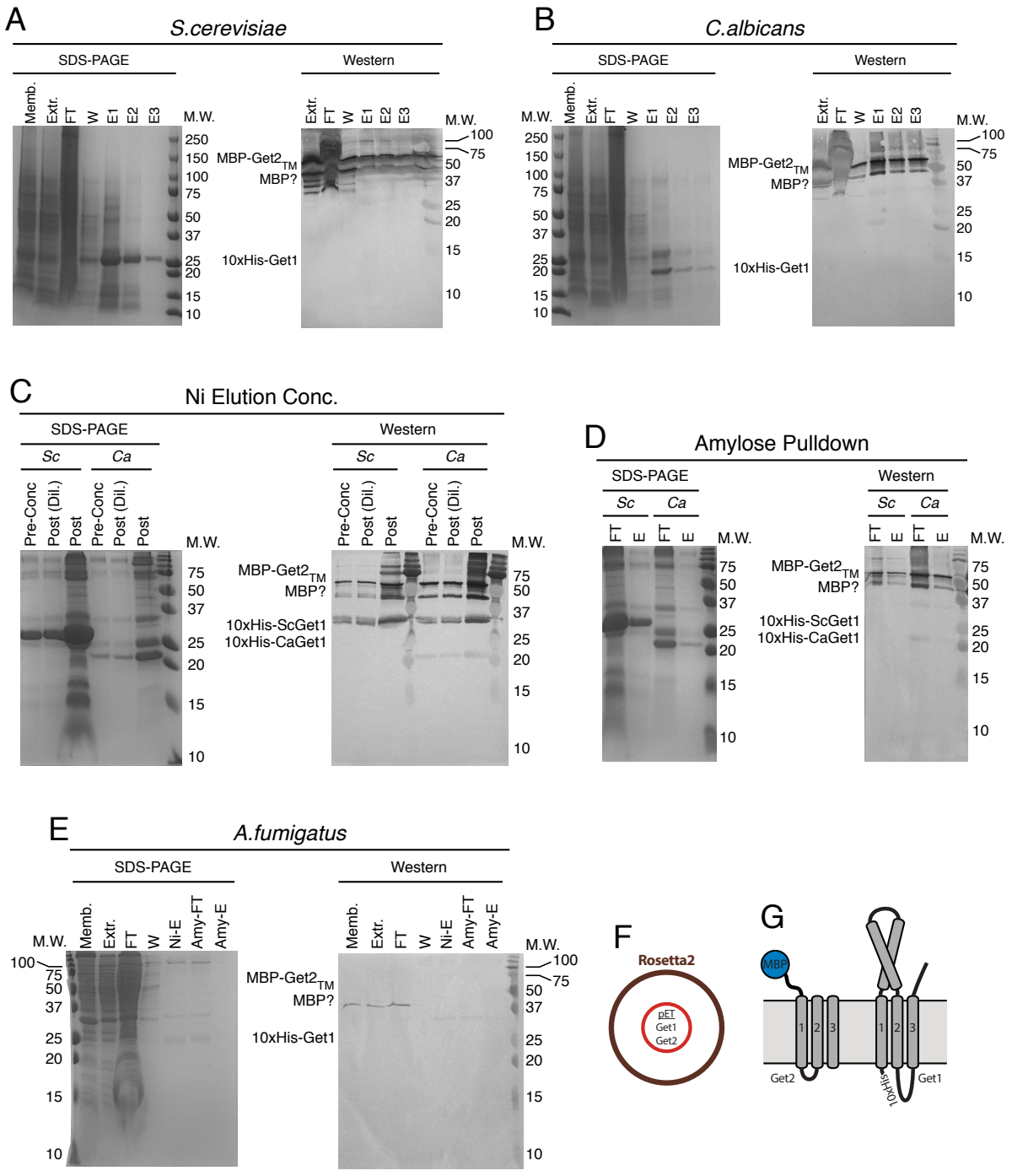

A
S.cerevisiae
SDS-PAGE
Western
Memb.
Extr.
FT
W
E1
E2
E3
M.W.
250
150
100
75
50
37
25
20
15
10
MBP-Get2TM
MBP?
10xHis-Get1
100
75
B
C.albicans
C
Ni Elution Conc.
Sc
Ca
Pre-Conc
Post (Dil.)
Post
10xHis-ScGet1
10xHis-CaGet1
D
Amylose Pulldown
E
A.fumigatus
Ni-E
Amy-FT
Amy-E
F
Rosetta2
pET
Get1
Get2
G
MBP
Get2
Get1
10xHis

**Figure A.6. Co-expression of MBP-Get2$_{TM}$/10xHis-Get1 in pET-DUET from *Sc*, *Ca*, and *Af*.** (A-B) SDS-PAGE and Western blots of Ni-NTA purifications from (A) *Sc* and (B) *Ca*. (C) SDS-PAGE and Western blots of concentrated smaples from (A) and (B). (D) SDS-PAGE and Western blots of amylose purifications from *Sc* and *Ca*. (E) SDS-PAGE and Western blots of Ni-NTA and amylose purifications from *Af*. (F-G) Cartoon representation of expression system (F) and the proteins and tags (G) used in the experiments.

**Figure A.7**

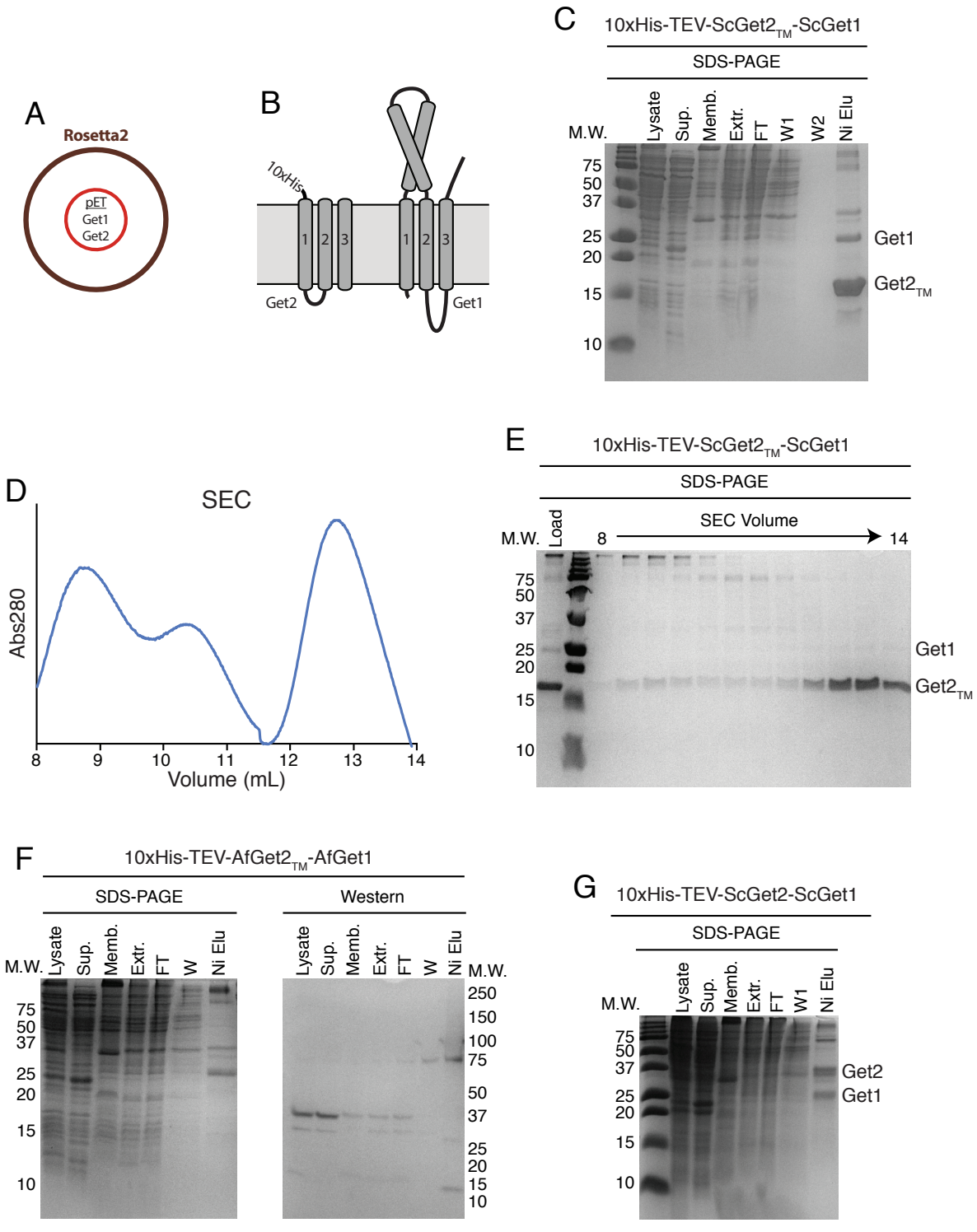

A
Rosetta2
pET
Get1
Get2
B
10xHis
1
2
3
1
2
3
Get2
Get1
C
10xHis-TEV-ScGet2TM-ScGet1
SDS-PAGE
M.W.
Lysate
Sup.
Memb.
Extr.
FT
W1
W2
Ni Elu
75
50
37
25
20
15
10
Get1
Get2TM
D
SEC
Abs280
8
9
10
11
12
13
14
Volume (mL)
E
10xHis-TEV-ScGet2TM-ScGet1
SDS-PAGE
M.W.
Load
SEC Volume
8
14
75
50
37
25
20
15
10
Get1
Get2TM
F
10xHis-TEV-AfGet2TM-AfGet1
SDS-PAGE
Western
M.W.
Lysate
Sup.
Memb.
Extr.
FT
W
Ni Elu
75
50
37
25
20
15
10
250
150
100
75
50
37
25
20
15
10
G
10xHis-TEV-ScGet2-ScGet1
SDS-PAGE
M.W.
Lysate
Sup.
Memb.
Extr.
FT
W1
Ni Elu
75
50
37
25
20
15
10
Get2
Get1

**Figure A.7. Co-expression of 10xHis-Get2/Get1 in pET-DUET from *Sc* and *Af*.** (A-B) Cartoon representation of expression system (A) and the proteins and tags (B) used in (C,E-F). (C) SDS-PAGE of 10xHis-Get2$_{TM}$/Get1 Ni-NTA purification from *Sc*. (D) SEC of sample from (C). (E) SDS-PAGE of fractions from (D). (F) SDS-PAGE and Western blot of 10xHis-Get2$_{TM}$/Get1 Ni-NTA purification from *Af*. (G) SDS-PAGE of 10xHis-Get2/Get1 Ni-NTA purification from *Sc*.

**Figure A.8**

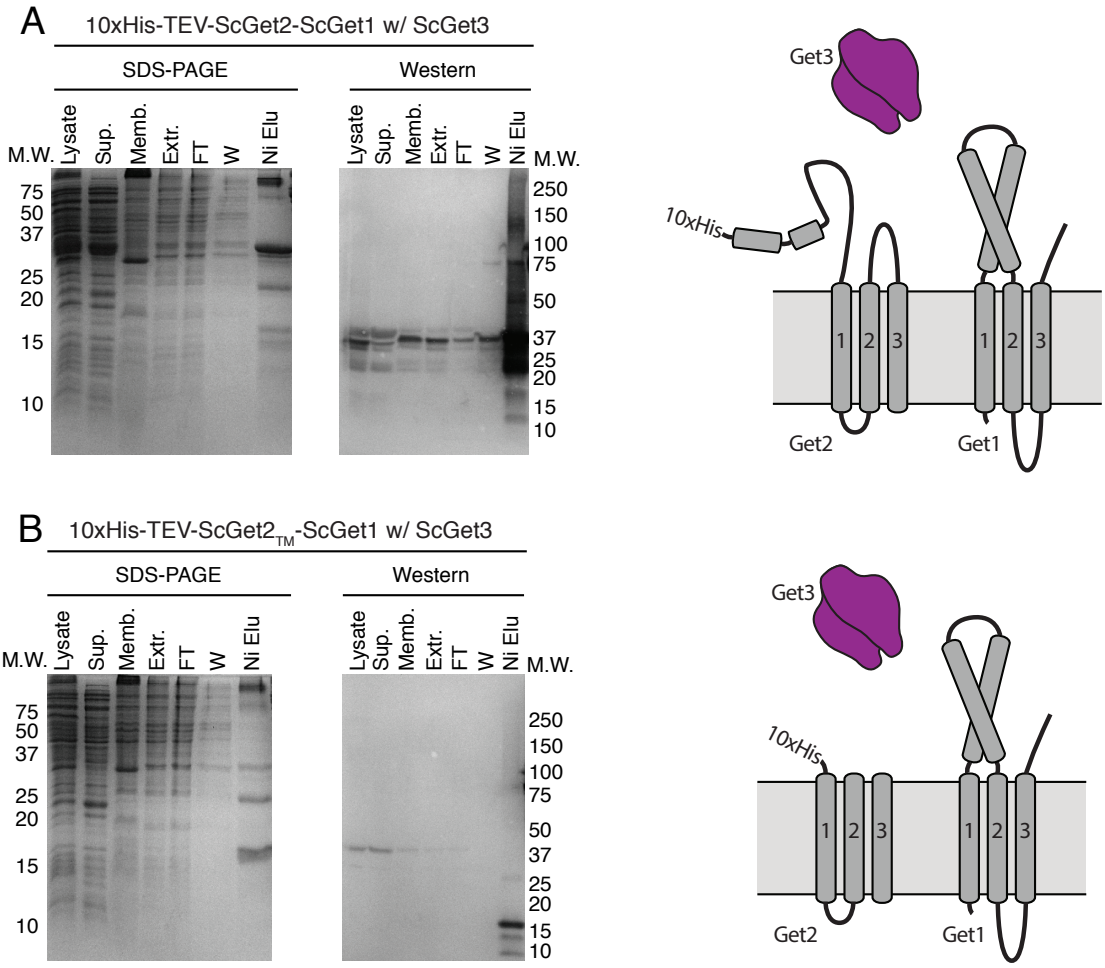



**Figure A.8. Co-expression of 10xHis-Get2/Get1 with untagged Get3 from *Sc*.** (A) Left, SDS-PAGE and Western blot of 10xHis-Get2/Get1/Get3 Ni-NTA and amylose purification from *Sc*. Right, cartoon representation of proteins and tags used in the experiment. (B) Left, SDS-PAGE and Western blot of 10xHis-Get2$_{TM}$/Get1/Get3 Ni-NTA and amylose purification from *Sc*. Right, cartoon representation of proteins and tags used in the experiment.

**Figure A.9**

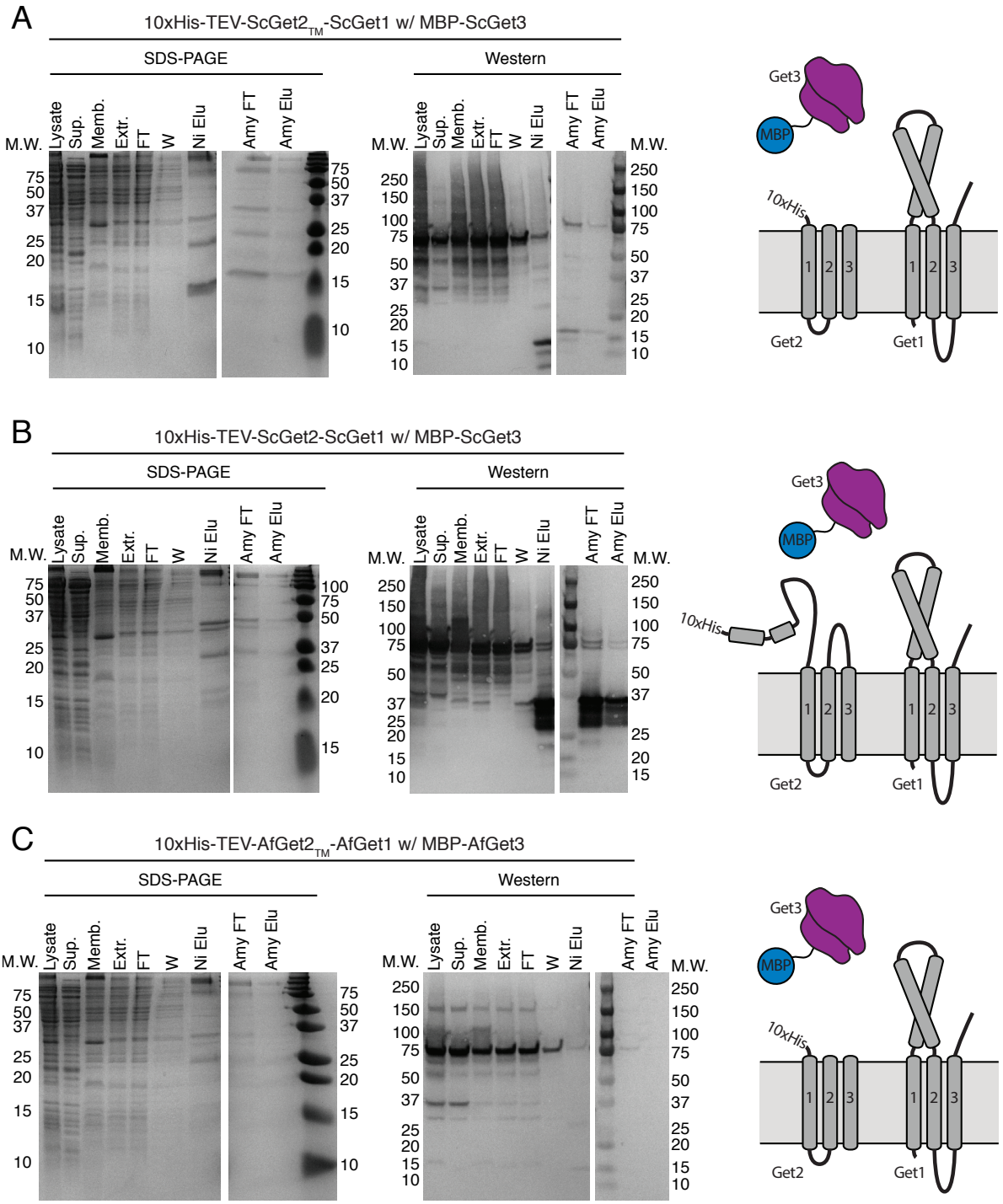

A
10xHis-TEV-ScGet2TM-ScGet1 w/ MBP-ScGet3
SDS-PAGE
Western
Lysate
Sup.
Memb.
Extr.
FT
W
Ni Elu
Amy FT
Amy Elu
M.W.
Get3
MBP
10xHis
Get2
Get1
B
10xHis-TEV-ScGet2-ScGet1 w/ MBP-ScGet3
C
10xHis-TEV-AfGet2TM-AfGet1 w/ MBP-AfGet3

**Figure A.9. Co-expression of 10xHis-Get2/Get1 with MBP-Get3 from *Sc* and *Af*.** (A) Left, SDS-PAGE and Western blot of 10xHis-Get2$_{TM}$/Get1/MBP-Get3 Ni-NTA and amylose purification from *Sc*. Right, cartoon representation of proteins and tags used in the experiment. (B) Left, SDS-PAGE and Western blot of 10xHis-Get2/Get1/MBP-Get3 Ni-NTA and amylose purification from *Sc*. Right, cartoon representation of proteins and tags used in the experiment. (C) Left, SDS-PAGE and Western blot of 10xHis-Get2$_{TM}$/Get1/MBP-Get3 Ni-NTA and amylose purification from *Af*. Right, cartoon representation of proteins and tags used in the experiment.

**Figure A.10**

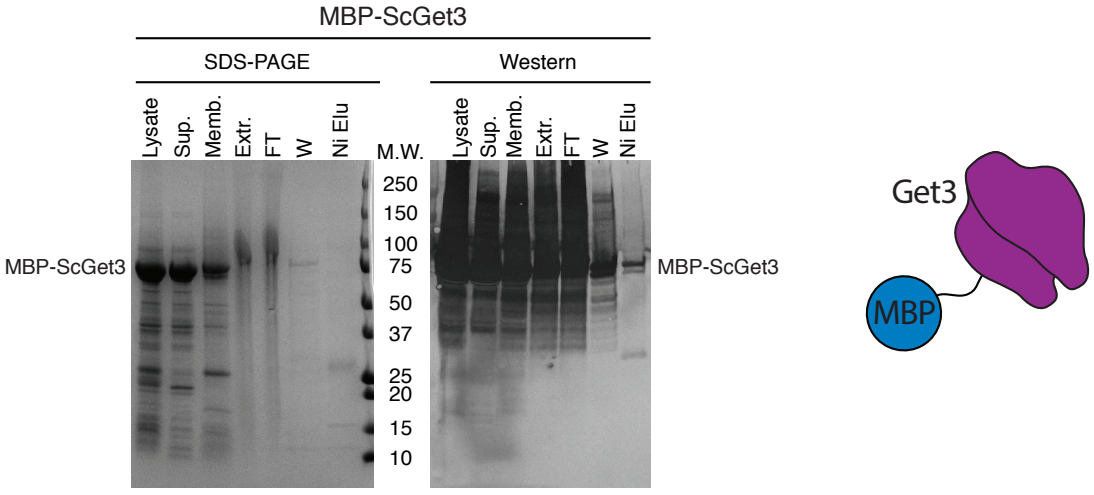



**Figure A.10. Purification of MBP-Get3 from *Sc*.** Left, SDS-PAGE and western blot of Ni-NTA and amylose purification. Right, cartoon representation of proteins and tags used in the experiment.

*A p p e n d i x  B*

# CHARACTERIZATION OF AN ARCHAEAL GET3 KNOCKOUT

## Introduction

In addition to eukaryotes, tail-anchored proteins are found in both bacteria and archaea [77]. However, given the lack of multiple membrane organelles and the spontaneous membrane insertion of TA proteins [7], it is unclear whether these organisms would require a dedicated TA targeting pathway. Surprisingly, homologues of the central TA targeting factor Get3 have been found in numerous archaeal species [77]. Recent studies of these proteins demonstrated they are structurally identical to the fungal Get3, and are able to rescue Get3 function *in vitro* [28,78].

Archaeal species have been used as model systems for over thirty years, and have been useful in studying both the protein translocation channel from *Methanocaldocaucus jannaschi* (*M. jannaschii)* [79] as well as the proton pump bacteriorhodopsin from *Halobacterium salinarum* (*H. salinarum)* [80]. To determine whether Get3 homologues were essential in Archaea, we generated knockout strains of *Haloferax volcanii*, a model halophilic archaea. Using these strains, we set out to address whether a conserved Get pathway exists in Archaea.

*H. volcanii* contains two apparent Get3 homologs (annotated Hvo_0724 and Hvo_2977), with ~30% sequence identity between the two proteins and ~25% sequence identity to *Sc*Get3 [28]. Although they contain much of the same sequence motifs, Hvo_0724 contains two distinct regions essential for Get3 function that are absent in Hvo_2977. First, Hvo_0724 contains the conserved "CxxC motif" [23-25,28], which tetrahedrally coordinates a zinc ion forming the "hinge" that stabilizes the dimer interface [23-25,28]. Second, Hvo_0724 contains the putative "Get3 motif", a hydrophobic stretch that is required for binding the TM of a TA

protein [24]. For these reasons, we believe Hvo_0724 is a Get3 homologue that may function as a TA targeting factor in an Archaeal TA targeting pathway.

*This work was performed in collaboration with Manuela Tripepi from the lab of Mechthild Pohlschroder at the University of Pennsylvania.*

## Results/Discussion

### *Haloferax volcanii* Get3 is not essential

To gain insight into the physiological roles of these proteins, we generated knockout strains in *H. volcanii*, done through collaboration with Mecky Poehlshroeder's lab at the University of Pennsylvania (**Figure B.1A**) [81]. Growing these strains in a variety of conditions did not reveal any differences in growth between the wildtype and knockout strains. To test whether there was a functional redundancy between the two genes, we then generated a double knockout strain in which both Hvo_0724 and Hvo_2977 were deleted. This also proved to be a non-lethal knockout, and showed no phenotypic differences in growth when compared to the wildtype strain. One theory for the presence of these genes in Archaea is that they initially functioned in arsenite/arsenate transport and evolved TA-binding capabilities as organisms evolved multiple membrane-bound organelles. However, the double knockout was not more sensitive than the wildtype when grown in the presence of either arsenate or arsenite. This result suggests an alternate function for these genes *in vivo* than simply arsenite/arsenate transport.

To test a wider array of growth conditions, the phenotype microarray system (Biolog) was used. This assay can test over five hundred conditions in a 96-well high-throughput format [82]. Preliminary experiments have given rise to conditions that show a difference in growth between the wildtype strain and the double knockout strain (**Figure B.1B**). While none of the compounds are directly involved in TA protein targeting, these results imply that the function of either Hvo_0724 or Hvo_2977 becomes essential during stress conditions, similar to growth assays in *S. cerevisiae* [19,23,65].

One hypothesis for the lack of an obvious phenotype in the initial growth assays is that

a chaperone becomes upregulated, which replaces the function of Hvo_0724 and Hvo_2977, and rescues growth in the knockout strains. A previous study analyzed the entire transcriptome from *H. volcanii* under different growth conditions [83]. Two genes of interest that were up regulated during stress conditions were DnaK (Hsp70 homolog), and Prefoldin Beta, two Archaeal chaperones [83,84]. Although Archaeal Get3 homologues were demonstrated to bind TA proteins *in vitro* [28], and replace Get3 function in an *in vitro* TA targeting assay [78], it is unknown whether Archaeal Get3 homologues function in TA targeting *in vivo*. The knockouts described here can be used to determine whether TA targeting is defective in these cells, and provide insight into the function of Get3 homologues in Archaea.

## Methods

### Generating *H. volcanii* knockout strains

*H. volcanii* strains were grown at 45 °C in complex medium (CX) or defined medium (CA) [81]. Plasmid constructs for use in knockout experiments were generated as previously described [85]. In brief, approximately 700 nucleotides upstream and downstream of the genes of interest (Hvo_0724 and Hvo_2977) were cloned into the haloarchaeal suicide vector pTA131 [86] using a modified version of the overlap PCR method [87].

Knockout strains were generated using the homologous recombination (pop-in/pop-out) method as previously described [85]. The knockout constructs were transformed into *E. coli* DL739 to first obtain non-methylated plasmid DNA. Using the standard polyethylene glycol method, the non-methylated plasmid DNA was used to transform *H. volcanii* H98 cells [86]. A single homologous recombination event between one of the flanking regions cloned into the plasmid and the chromosome (pop in) was selected for by growth on CA agar lacking uracil. Recombinants were then grown for 48 h in liquid CA medium supplemented with thymidine, hypoxanthine, and uracil to allow a second recombination event, which results in excision of the plasmid from the chromosome (pop out). After 48 h, liquid cultures were transferred to CA agar plates supplemented with thymidine, hypoxanthine, uracil, and 5-FOA. These conditions only permit the growth of cells in which the plasmid has been excised from the chromosome. Finally, colonies were screened by PCR to confirm the knockout.

# Figures

**Figure B.1**

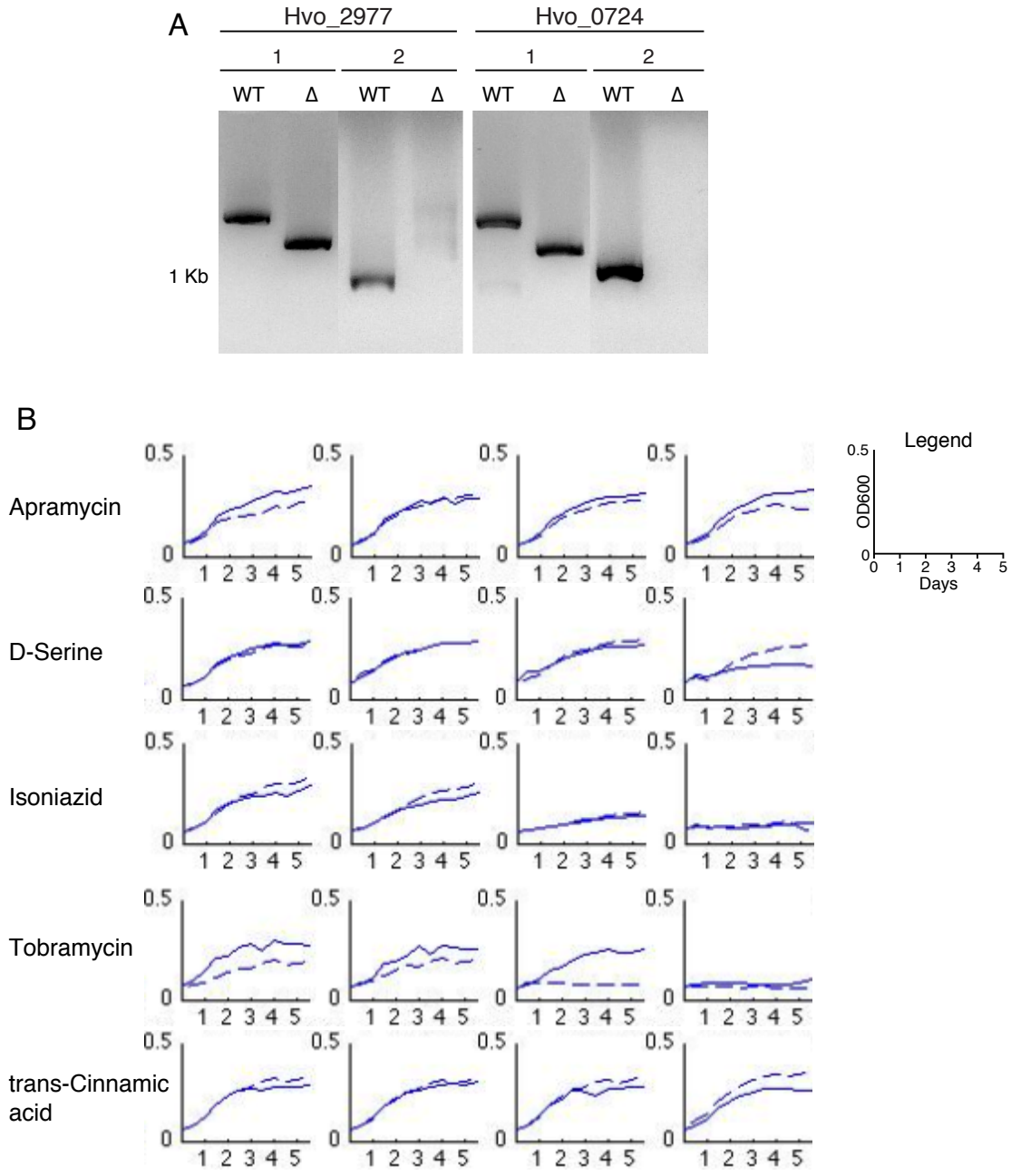


**Figure B.1. Generating knockouts of Get3 homologues in *Haloferax volcanii*.** (A) PCR confirmation of knockouts using flanking primers (1) or gene specific primers (2) for Hvo_2799 and Hvo_0724. (B) Conditions from BIOLOG experiments that gave rise to difference in growth between *H. volcanii* wildtype (dashed line) and ΔHvo_2799/ΔHvo_0724 (solid line) double knockout strains. The concentration of each compound from left to right is: Apramycin (70 μM, 209 μM, 627 μM, 1880 μM); D-Serine (3.1 mM, 9.2 mM, 28 mM, 83 mM); Isoniazid (0.33 mM, 1 mM, 3 mM, 9 mM); Tobramycin (44.3 μM, 133 μM, 400 μM, 1200 μM); trans-Cinnamic acid (8.13 μM, 24.4 μM, 73.3 μM, 220 μM).

## Bibliography


1. Wallin, E. & von Heijne, G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7**, 1029-38 (1998).

2. Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M. & Berglund, L. Prediction of the human membrane proteome. *Proteomics* **10**, 1141-9 (2010).

3. Shan, S.O. & Walter, P. Co-translational protein targeting by the signal recognition particle. *FEBS Lett* **579**, 921-6 (2005).

4. Akopian, D., Shen, K., Zhang, X. & Shan, S.O. Signal recognition particle: an essential protein-targeting machine. *Annu Rev Biochem* **82**, 693-721 (2013).

5. Beilharz, T., Egan, B., Silver, P.A., Hofmann, K. & Lithgow, T. Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in Saccharomyces cerevisiae. *J Biol Chem* **278**, 8219-23 (2003).

6. Kalbfleisch, T., Cambon, A. & Wattenberg, B.W. A bioinformatics approach to identifying tail-anchored proteins in the human genome. *Traffic* **8**, 1687-94 (2007).

7. Dailey, H.A. & Strittmatter, P. Structural and functional properties of the membrane binding segment of cytochrome b5. *J Biol Chem* **253**, 8203-9 (1978).

8. Kutay, U., Hartmann, E. & Rapoport, T.A. A class of membrane proteins with a C-terminal anchor. *Trends in cell biology* **3**, 72-5 (1993).

9. Kutay, U., Ahnert-Hilger, G., Hartmann, E., Wiedenmann, B. & Rapoport, T.A. Transport route for synaptobrevin via a novel pathway of insertion into the endoplasmic reticulum membrane. *The EMBO journal* **14**, 217-23 (1995).

10. Stefanovic, S. & Hegde, R.S. Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell* **128**, 1147-59 (2007).

11. Favaloro, V., Spasic, M., Schwappach, B. & Dobberstein, B. Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *Journal of cell science* **121**, 1832-40 (2008).

12. Zuniga, S., Boskovic, J., Jimenez, A., Ballesta, J.P. & Remacha, M. Disruption of six Saccharomyces cerevisiae novel genes and phenotypic analysis of the deletants. *Yeast* **15**, 945-53 (1999).

13. Shen, J., Hsu, C.M., Kang, B.K., Rosen, B.P. & Bhattacharjee, H. The Saccharomyces cerevisiae Arr4p is involved in metal and heat tolerance. *Biometals* **16**, 369-78 (2003).

14. Tseng, Y.Y., Yu, C.W. & Liao, V.H. Caenorhabditis elegans expresses a functional ArsA. *FEBS J* **274**, 2566-72 (2007).

15. Mukhopadhyay, R., Ho, Y.S., Swiatek, P.J., Rosen, B.P. & Bhattacharjee, H. Targeted disruption of the mouse Asna1 gene results in embryonic lethality. *FEBS Lett* **580**, 3889-94 (2006).

16. Hemmingsson, O., Zhang, Y., Still, M. & Naredi, P. ASNA1, an ATPase targeting tail-anchored proteins, regulates melanoma cell growth and sensitivity to cisplatin and arsenite. *Cancer Chemother Pharmacol* **63**, 491-9 (2009).

17. Schuldiner, M. et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* **123**, 507-19 (2005).

18. Auld, K.L. et al. The conserved ATPase Get3/Arr4 modulates the activity of membrane-associated proteins in Saccharomyces cerevisiae. *Genetics* **174**, 215-27 (2006).

19. Schuldiner, M. et al. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* **134**, 634-45 (2008).

20. Jonikas, M.C. et al. Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* **323**, 1693-7 (2009).

21. Costanzo, M. et al. The genetic landscape of a cell. *Science* **327**, 425-31 (2010).

22. Battle, A., Jonikas, M.C., Walter, P., Weissman, J.S. & Koller, D. Automated identification of pathways from quantitative genetic interaction data. *Molecular systems biology* **6**, 379 (2010).

23. Suloway, C.J., Chartron, J.W., Zaslaver, M. & Clemons, W.M., Jr. Model for eukaryotic tail-anchored protein binding based on the structure of Get3. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 14849-54 (2009).

24. Mateja, A. et al. The structural basis of tail-anchored membrane protein recognition by Get3. *Nature* **461**, 361-6 (2009).

25. Bozkurt, G. et al. Structural insights into tail-anchored protein binding and membrane insertion by Get3. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 21131-6 (2009).

26. Hu, J., Li, J., Qian, X., Denic, V. & Sha, B. The crystal structures of yeast Get3 suggest a mechanism for tail-anchored protein membrane insertion. *PloS one* **4**, e8061 (2009).

27. Yamagata, A. et al. Structural insight into the membrane insertion of tail-anchored proteins by Get3. *Genes to Cells* **15**, 29-41 (2010).

28. Suloway, C.J., Rome, M.E. & Clemons, W.M., Jr. Tail-anchor targeting by a Get3 tetramer: the structure of an archaeal homologue. *The EMBO journal* **31**, 707-19 (2012).

29. Mariappan, M. et al. The mechanism of membrane-associated steps in tail-anchored protein insertion. *Nature* **477**, 61-6 (2011).

30. Stefer, S. et al. Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. *Science* **333**, 758-62 (2011).

31. Wang, F., Whynot, A., Tung, M. & Denic, V. The mechanism of tail-anchored protein insertion into the ER membrane. *Mol Cell* **43**, 738-50 (2011).

32. Yamamoto, Y. & Sakisaka, T. Molecular machinery for insertion of tail-anchored membrane proteins into the endoplasmic reticulum membrane in mammalian cells. *Mol Cell* **48**, 387-97 (2012).

33. Chartron, J.W., Gonzalez, G.M. & Clemons, W.M., Jr. A Structural Model of the Sgt2 Protein and Its Interactions with Chaperones and the Get4/Get5 Complex. *The Journal of Biological Chemistry* **286**, 34325-34334 (2011).

34. Wang, F., Brown, E.C., Mak, G., Zhuang, J. & Denic, V. A chaperone cascade sorts proteins for posttranslational membrane insertion into the endoplasmic reticulum. *Molecular Cell* **40**, 159-71 (2010).

35. Chartron, J.W., VanderVelde, D.G. & Clemons, W.M., Jr. Structures of the Sgt2/SGTA dimerization domain with the Get5/UBL4A UBL domain reveal an interaction that forms a conserved dynamic interface. *Cell Rep* **2**, 1620-32 (2012).

36. Chartron, J.W., Suloway, C.J., Zaslaver, M. & Clemons, W.M., Jr. Structural characterization of the Get4/Get5 complex and its interaction with Get3. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 12127-32 (2010).

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

38. Copic, A. et al. Genomewide analysis reveals novel pathways affecting endoplasmic reticulum homeostasis, protein modification and quality control. *Genetics* **182**, 757-69 (2009).

39. Chang, Y.W. et al. Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1. *The Journal of biological chemistry* **285**, 9962-70 (2010).

40. Chartron, J.W., Vandervelde, D.G., Rao, M. & Clemons, W.M., Jr. The Get5 carboxyl terminal domain is a novel dimerization motif that tethers an extended Get4/Get5 complex. *The Journal of Biological Chemistry* **287**, 8310-7 (2012).

41. Rome, M.E., Rao, M., Clemons, W.M., Jr. & Shan, S.O. Precise timing of ATPase activation drives targeting of tail-anchored proteins. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 7666-71 (2013).

42. Chang, Y.W. et al. Interaction Surface and Topology of Get3-Get4-Get5 Protein Complex, Involved in Targeting Tail-anchored Proteins to Endoplasmic Reticulum. *The Journal of Biological Chemistry* **287**, 4783-9 (2012).

43. Rome, M.E., Chio, U.S., Rao, M., Gristick, H. & Shan, S.O. Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway. *Proc Natl Acad Sci U S A* **111**, E4929-35 (2014).

44. Shao, S. & Hegde, R.S. Membrane protein insertion at the endoplasmic reticulum. *Annual review of cell and developmental biology* **27**, 25-56 (2011).

45. Borgese, N., Gazzoni, I., Barberi, M., Colombo, S. & Pedrazzini, E. Targeting of a tail-anchored protein to endoplasmic reticulum and mitochondrial outer membrane by independent but competing pathways. *Molecular biology of the cell* **12**, 2482-96 (2001).

46. Borgese, N., Brambillasca, S. & Colombo, S. How tails guide tail-anchored proteins to their destinations. *Current Opinion in Cell Biology* **19**, 368-75 (2007).

47. Chartron, J.W., Clemons, W.M., Jr. & Suloway, C.J. The complex process of GETting tail-anchored membrane proteins to the ER. *Current Opinion in Structural Biology* **22**, 217-24 (2012).

48. Zhang, X., Rashid, R., Wang, K. & Shan, S.O. Sequential checkpoints govern substrate selection during cotranslational protein targeting. *Science* **328**, 757-60 (2010).

49. Ataide, S.F. et al. The crystal structure of the signal recognition particle in complex with its receptor. *Science* **331**, 881-6 (2011).

50. Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R. & Leslie, A.G. iMOSFLM: a new graphical interface for diffraction-image processing with

MOSFLM. *Acta crystallographica. Section D, Biological crystallography* **67**, 271-81 (2011).

51. Kabsch, W. XDS. *Acta crystallographica. Section D, Biological crystallography* **66**, 125-32 (2010).

52. CCP4, C.C.P.N.-. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760-3 (1994).

53. Winn, M.D. et al. Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography* **67**, 235-42 (2011).

54. McCoy, A.J. et al. Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).

55. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-21 (2010).

56. Brunger, A. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728-33 (2007).

57. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).

58. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122**, 19-27 (1989).

59. Gietz, R.D. & Schiestl, R.H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols* **2**, 31-4 (2007).

60. Pettersen, E. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605-12 (2004).

61. Larkin, M.A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-8 (2007).

62. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. & McCammon, J.A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10037-41 (2001).

63. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A. & Baker, N.A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research* **32**, W665-7 (2004).

64. Hegde, R.S. & Keenan, R.J. Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nat Rev Mol Cell Biol* **12**, 787-98 (2011).

65. Gristick, H.B. et al. Crystal structure of ATP-bound Get3-Get4-Get5 complex reveals regulation of Get3 by Get4. *Nat Struct Mol Biol* **21**, 437-42 (2014).

66. Mateja, A. et al. Protein targeting. Structure of the Get3 targeting factor in complex with its membrane protein cargo. *Science* **347**, 1152-5 (2015).

67. Schreiber, G. & Fersht, A.R. Rapid, electrostatically assisted association of proteins. *Nat Struct Biol* **3**, 427-31 (1996).

68. Sandikci, A. et al. Dynamic enzyme docking to the ribosome coordinates N-terminal processing with polypeptide folding. *Nat Struct Mol Biol* **20**, 843-50 (2013).

69. Painter, J. & Merritt, E.A. TLSMD web server for the generation of multi-group TLS models. *Journal of Applied Crystallography* **39**, 109-111 (2006).

70. Nguyen, T.X. et al. Mechanism of an ATP-independent protein disaggregase: I. structure of a membrane protein aggregate reveals a mechanism of recognition by its chaperone. *J Biol Chem* **288**, 13420-30 (2013).

71. Johnson, K.A. Fitting enzyme kinetic data with KinTek Global Kinetic Explorer. *Methods Enzymol* **467**, 601-26 (2009).

72. Johnson, K.A., Simpson, Z.B. & Blom, T. Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal Biochem* **387**, 20-9 (2009).

73. Shen, K., Zhang, X. & Shan, S.O. Synergistic actions between the SRP RNA and translating ribosome allow efficient delivery of the correct cargos during cotranslational protein targeting. *RNA* **17**, 892-902 (2011).

74. Wang, F., Chan, C., Weir, N.R. & Denic, V. The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature* **512**, 441-4 (2014).

75. Studier, F.W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-34 (2005).

76. Kefala, G. et al. Structures of the OmpF porin crystallized in the presence of foscholine-12. *Protein Sci* **19**, 1117-25 (2010).

77. Borgese, N. & Righi, M. Remote origins of tail-anchored proteins. *Traffic* **11**, 877-85 (2010).

78. Sherrill, J., Mariappan, M., Dominik, P., Hegde, R.S. & Keenan, R.J. A conserved archaeal pathway for tail-anchored membrane protein insertion. *Traffic* **12**, 1119-23 (2011).

79. Henderson, R. et al. Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol* **213**, 899-929 (1990).

80. Van den Berg, B. et al. X-ray structure of a protein-conducting channel. *Nature* **427**, 36-44 (2004).

81. Dyall-Smith, M. The Halohandbook: Protocols for Halobacterial Genetics, vol. 6.01 (2006).

82. Tohsato, Y. & Mori, H. Phenotype profiling of single gene deletion mutants of E. coli using Biolog technology. *Genome Inform* **21**, 42-52 (2008).

83. Dambeck, M. & Soppa, J. Characterization of a Haloferax volcanii member of the enolase superfamily: deletion mutant construction, expression analysis, and transcriptome comparison. *Arch Microbiol* **190**, 341-53 (2008).

84. Large, A.T., Goldberg, M.D. & Lund, P.A. Chaperones and protein folding in the archaea. *Biochem Soc Trans* **37**, 46-51 (2009).

85. Allers, T. & Ngo, H.P. Genetic analysis of homologous recombination in Archaea: Haloferax volcanii as a model organism. *Biochem Soc Trans* **31**, 706-10 (2003).

86. Allers, T., Ngo, H.P., Mevarech, M. & Lloyd, R.G. Development of additional selectable markers for the halophilic archaeon Haloferax volcanii based on the leuB and trpA genes. *Appl Environ Microbiol* **70**, 943-53 (2004).

87. Hammelmann, M. & Soppa, J. Optimized generation of vectors for the construction of Haloferax volcanii deletion mutants. *J Microbiol Methods* **75**, 201-4 (2008).