The Bag6 Complex:
Biological Complexity through Modularity

Thesis by

Jee-Young Mock

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2017

Defended May 19, 2017

Jee-Young Mock
ORCID: 0000-0002-4656-3357

ACKNOWLEDGEMENTS

First I want to thank professor William (Bil) Clemons for his unrelenting support for my education and growth as a human being. Bil patiently guided me through my struggles, honestly taught me how to deal with my shortcomings, and enthusiastically cheered me on through my triumphs. He fostered an open and safe environment, in which no question was ignored or judged, for me to find my footing as a scientist. I would not be the scientist I am today without his mentorship. I am also grateful to the Clemons family—Jana, Kuba and Ozzie—for accepting me as one of their own. Although I was an ocean apart from my family, I felt at home in Pasadena because of them.

I thank my committee members: Professors Pamela Bjorkman, David Chan and Shu-ou Shan. My conversations with Pamela about her experiences as a woman in science helped me contextualize my struggles and cheered me on through times of self-doubt and frustrations. I thank her for being both an exemplar and champion for women in science. David's class, Bi/Ch113, helped me fall in love with cell biology. I also thank him for taking a sincere interest in my project and career path and advising me with honesty. I thank Shu-ou for her insightful comments that always forced me to think critically about my data and experimental design.

I would also like to acknowledge the mentors throughout my career who believed in me and nurtured my curiosity: my undergraduate thesis advisor Professor Sean J. Elliott, my summer research advisor Professor Han-woong Lee, and the first professor I ever worked for when I was only a high school student, Professor Won-Gi Bang.

I feel grateful to have been part of an inspiring group of scientists in the Clemons lab. I am particularly thankful for Dr. Justin Chartron, the most inspiring and patient mentor I could have asked for. I am a better biochemist largely because of him. I look forward to a lifelong friendship and scientific partnership with Professor Chartron! A special thanks to members of GET/TRC team in the Clemons lab: Dr. Christian Suloway, Dr. Harry Gristick, Dr. Ku-feng Lin and Michelle Fry. Their support and comments were crucial to completing this work. I also thank Dr. Shiho Tanaka, Dr. Sureshkumar Ramasamy, Dr. Kyoung-soon Jang, Ma'ayan Zaslaver, Hyun Gi Yun, Stephen Marshall, Dr. Axel Müller, Shyam Saladi, Gyu-bin Jang, Catherine Day and Romanus Hutchins for their support.

Our collaboration with the Shan lab on campus was indispensable. I thank the GET team, Dr. Michael Rome, Dr. Meera Rao, Dr. Hyunju Cho and Unseng Chio for valuable comments and technical assistance. Off campus, the Dr. Yihong Ye and Dr. Yue Xu from the Ye lab at the National Institute of Health (NIH) were instrumental in addressing questions that required human cell biology tools.

My six years at Caltech would not have been the same without the friends who stood by my side. I thank the 'boys,' Mikhail Hanewich-Hollatz, Dr. Daniel Lin and Dr. Jeffrey Bosco, for being amazing friends, cheerleaders, and at times, critics, throughout my journey at Caltech. I will always cherish our memories of *fiercely* celebrating my passing candidacy, holidays, birthdays and a wedding! And I know that we will continue to celebrate for many years to come. In the Clemons lab, I found not only an amazing

colleague but also a big sister, Dr. Shiho Tanaka. She is one of the most intelligent, kind and caring people I know and I am lucky to have met her and become part of her family. I will always cherish my memories of eating around the country with Melissa Hahn, my best friend from the good old and wild days at Boston University. I thank her for working on our long-distance friendship over the past six years, and look forward to being closer on the other coast! I thank Samantha Adelberg for constantly challenging me and inspiring me to see the world in a different way. I also thank Dr. Katrin Tiemann, Dr. Meera Rao, Hyun Gi Yun, Dr. Helty Adisetiyo, Dr. Christian Suloway, Luke Urban, Lincoln Collins and Anastasios Dimitriadis for their friendship and support. You all have kept me grounded and supported me in different ways over the past 6 years – thank you.

My husband, Jacques Edeline, is kind, loving, intelligent and understanding in my pursuit of science and research. He has been supportive of my career choices, even in sending me off to the other side of the country in pursuit of my dream; I am so grateful for this, on top of all the other ways he enriches my life. I love you, JBE! Marrying into the loving and supportive Edeline family has also been an incredible gift. I feel so lucky to have in-laws who made me feel right at home since day one.

Lastly, I thank my family. My dad, Minsoo Mock, gave my brother and me the opportunity to study abroad at a very young age. My brother, Jeonghoon Mock, has been a quiet cheerleader throughout my journey abroad. I am so proud of all that he has achieved despite difficult circumstances and look forward to many more adventures overseas with my one and only baby brother. No one understands me better than my aunt, Hyekyung

Park, whom I consider to be my second mom. My uncles, Hoyong Park and Sangyong Park, have raised me as if I were their own and never stopped working to put me through school. Without their hard work and persistence, I would not have had the tremendous opportunities in the U.S. My grandfather, Donhee Park, has given me so much love and support. He led by example to teach me how to be bold, ambitious, and loving.

Growing up, I had great examples of strong women. My grandmother, Youngsoo Kwon, was born at the tail end of the Japanese occupation and lived through the Korean War. She rarely spoke about it, but carried her experiences with her as she worked tirelessly to feed and care for her family. Her sacrifice and hard work allowed me to thrive.

The greatest inspiration for my life has been my mother, who had me when she was only 22 years old. She never went to college, but she is one of the most intelligent and hardworking people I know. She instilled in me the value of hard work and sacrificed her own ambitions for my education. I am proud and grateful to have a mother who sees and nurtures my potential in every way she knows how.

This thesis is dedicated to my mother and my grandmother, two intelligent and hardworking women, who, like many Korean women of their generation, never had the opportunity to pursue their own dreams. They instead sacrificed their lives so that we could pursue ours. I love you both with all my heart.

ABSTRACT

Proper synthesis and targeting of membrane proteins that contain hydrophobic transmembrane domains are mediated by chaperones and targeting factors. Tail-anchored (TA) proteins are a special class of membrane proteins that are characterized by a single carboxy (C) terminal helix that anchors them to biological membranes. Fungal *G*uided *E*ntry of *T*ail-anchored protein (GET) pathway components, which include four soluble proteins—Sgt2, Get3, Get4, Get5—and two membrane bound receptors—Get1 and Get2—mediate TA biogenesis. These proteins maintain TA protein solubility in the aqueous cytosol and target TA to the endoplasmic reticulum. While most of the components are conserved in metazoans, one additional protein, Bag6, reorganizes the sorting complex from the heterotetrameric Get4-5 to the heterotrimeric Bag6-TRC35-Ubl4A. To understand the molecular architecture and mechanism of the Bag6 complex, we took a multidisciplinary approach that combines x-ray crystallography, biochemical reconstitution, and cell biology. Our studies demonstrate that the BAG domain of Bag6 is not a canonical BAG domain. Instead, main role of the Bag6 'mock' BAG domain is to dimerize with Ubl4A. Furthermore, the truncated Bag6 complex defined in this study is sufficient to facilitate substrate transfer from SGTA to TRC40. Lastly, our results unequivocally establish TRC35 as a cytoplasmic retention factor for Bag6. These results provide structural, biochemical and cell biological bases for modular Bag6 function and regulation of nucleocytoplasmic distribution of Bag6 by TRC35.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Mock, J.-Y., Chartron, J.W., Zaslaver, M., Xu, Y., Ye, Y., and Clemons, W.M. Jr. (2015) Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain. *Proc Natl Acad Sci USA* 106-111, doi: 10.1073/pnas.1402745112

   J.-Y. Mock participated in designing experiments, conducting experiments, and writing of the manuscript.

2. Mock, J.-Y., Xu, Y., Ye, Y., and Clemons, W.M. Jr. (2017) Structural basis for nucleocytoplasmic distribution of Bag6. *Submitted*.

   J.-Y. Mock participated in designing experiments, conducting experiments, and writing of the manuscript.

TABLE OF CONTENTS

Chapter 5
Bag6: A Modular Multitasker

LIST OF FIGURES AND TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

Chapter 1

INTRODUCTION

Biological Membranes and Cellular Evolution

Formation of the biological membrane was crucial to the emergence of cellular life on earth. While it is likely that the first cellular membranes differed significantly from the phospholipid bilayers we observe in modern cells, the fundamental properties that membranes confer to life do not change. Membranes encapsulate the cell, forming a protective barrier from the changing and often harsh environment. This encapsulation also results in compartmentalization of the biochemical activity of a cell, such as DNA replication and metabolic activity, from the environment. These two conditions were prerequisites to the emergence of the self-replicative cell that comprises life on earth today.

Encapsulation and compartmentalization by biological membranes are also guiding principles for achieving further biological complexity in eukaryotes. Membrane bound compartments within the plasma membrane partitions specific activities within the cell. The encapsulation of DNA by the nuclear membrane, for example, not only ensures greater integrity of the cell's genetic material, but also physically separates transcription from translation. This decoupling allows for mRNA processing prior to translation, which expands the transcriptome and the proteome without altering the genome size. Other subcellular compartments specialize in different functions, such as protein synthesis and modification in the case of endoplasmic reticulum (ER) and metabolic reactions in the case of mitochondria. In complex multicellular organisms, differential distribution of subcellular

organelles in distinct cell types plays a role in determining tissue-specific function. The extensive ER networks in pancreatic cells that secrete digestive enzymes and the abundance of mitochondria in muscle cells exemplify membrane-bound organelle-driven tissue specialization.

Understanding the molecular machineries that have evolved to accommodate and maintain these increasingly complex biological membrane systems is crucial to advancing our understanding of cellular biology and can provide valuable insight into the underlying principles of molecular evolution. This dissertation explores a metazoan molecular machinery involved in proper biosynthesis of proteins that reside in the membrane.

Membrane Protein Biogenesis

Biological membranes do not exist in isolation; rather, membrane proteins that reside in biological membranes are indispensable for cellular homeostasis, membrane function and organelle function (Engel and Gaub, 2008; Tan et al., 2008). Membrane protein synthesis, however, is inherently problematic for the cell. First, membrane proteins are characterized by hydrophobic transmembrane domains that are required for spanning the hydrophobic phospholipid bilayer. Due to these hydrophobic domains, membrane proteins are prone to aggregation in the aqueous and crowded cytosol where they are synthesized. Furthermore, they need to be targeted to appropriate destination membranes. Dedicated cellular machineries work to ensure proper folding and targeting of membrane proteins. One such machinery is the universally conserved and essential *S*ignal *R*ecognition *P*article (SRP) (Akopian et al., 2013). SRP co-translationally recognizes the amino-terminal (N-terminal) signal sequence (Blobel and Dobberstein, 1975), which is usually the first transmembrane

domain encoded by transmembrane proteins. In eukaryotes, SRP binds the ribosome nascent chain complex (RNC) (Walter and Blobel, 1981), and the SRP-RNC complex is recruited to the ER membrane via its interaction with the ER-resident SRP receptor (SR) (Gilmore et al., 1982). Upon arrival, the RNC is transferred to the Sec61 translocon (Simon and Blobel, 1991), a protein channel through which polypeptides are translocated across membranes. In this highly-coordinated process, SRP acts both as a chaperone and a targeting factor (Akopian et al., 2013).

Notably, membrane proteins whose single transmembrane domain is near the carboxy terminus (C-terminus) cannot access the co-translational SRP pathway because the transmembrane domain—the signal sequence—does not emerge from the ribosome until translation has terminated. These proteins, named tail-anchored (TA) proteins, are topologically constrained from accessing the SRP pathway. TA proteins were first described in 1993 (Kutay et al., 1993). Subsequent biochemical studies showed that the post-translational insertion of synaptobrevin, a model TA protein, into the ER membrane is SRP-independent and ATP-dependent (Kutay et al., 1995). For 12 years the TA targeting machinery remained elusive until a 40 kDa cytosolic ATPase, *T*ransmembrane domain *R*ecognition *C*omplex 40 (TRC40), was isolated from an unbiased biochemical screen for a predominant TA binding partner in rabbit reticulocyte lysate (Stefanovic and Hegde, 2007). Thanks to high evolutionary conservation of TRC40, its fungal counterpart Get3 was quickly identified. Get3 was initially identified in a large-scale genetic interaction study of the yeast secretory pathway (Schuldiner et al., 2005). Get1, 2, or 3 knockout yeast displayed phenotypes consistent with defects in retrograde trafficking, and thus were

named *G*olgi *E*R *T*rafficking 1-3 (Schuldiner et al., 2005). After studies in the mammalian TA targeting pathway identified Get3 as the TRC40 homologue (Stefanovic and Hegde, 2007), Get3 and its membrane associated interaction partners, Get1 and Get2, were re-established as central players in fungal TA targeting (Schuldiner et al., 2008; Wang et al., 2011a), which led to the renaming of the Get pathway as *G*uided *E*ntry of *T*ail-anchored protein pathway.

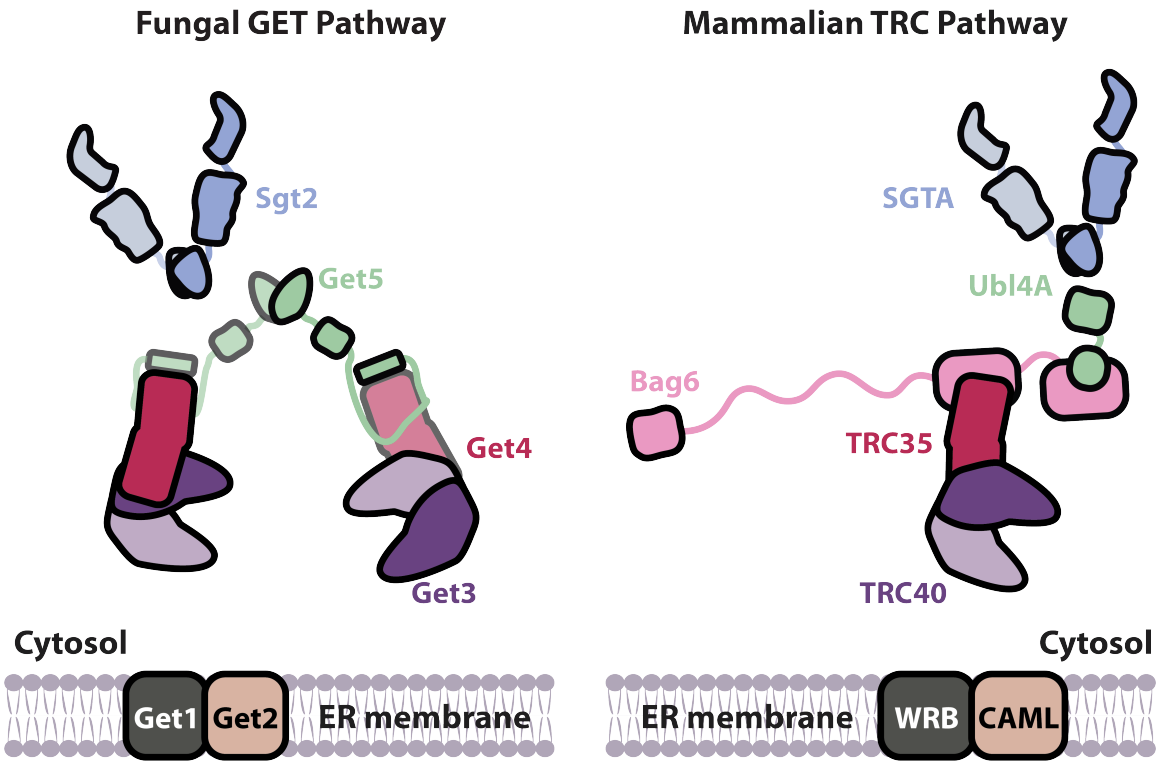


Figure 1. Overview of the TA targeting pathways in fungi and mammals. Upon translation termination and release from the ribosome, TA is captured by Sgt2/SGTA and handed off to Get3/TRC40. This handoff step is facilitated by the Get4-5 heterotetramer in yeast and the Bag6 heterotrimer in mammals. The Get3-TA/TRC40-TA complex that results is ultimately recruited to the ER membrane by membrane receptors, fungal Get1/2 or mammalian WRB/CAML.

Unlike the co-translational SRP pathway, which is comprised of a single cytosolic factor SRP and a single membrane receptor SR, the fungal Get system is composed of four

soluble components—Sgt2, Get3-5—and two membrane bound factors—Get1 and Get2; in metazoans, incorporation of one additional soluble factor results in five soluble components—SGTA, TRC40, TRC35, Ubl4A, Bag6—and two membrane bound factors—WRB and CAML. In yeast, upon termination of translation and release from the ribosome, TA substrate is captured by Sgt2. Sgt2 hands the TA substrate off to the targeting factor Get3 (Wang et al., 2010; Wang et al., 2011a), a process facilitated by the Get4-5 heterotetramer (Gristick et al., 2014). This Get4-5 heterotetramer will be referred to as the fungal sorting complex throughout the text. The Get3-TA complex is recruited to the ER membrane via the ER membrane-associated Get3 receptor, Get1/2 complex, followed by TA release and insertion into the membrane (Stefer et al., 2011; Wang et al., 2014a).

In mammals, the overall TA transfer from the Sgt2 homologue, SGTA, to the Get3 homologue, TRC40, is conserved. SGTA post-translationally captures free TA substrate then hands it off to TRC40, a process facilitated by the mammalian sorting complex, the Bag6 heterotrimer comprised of Bag6, TRC35 (Get4 homologue) and Ubl4A (Get5 homologue) (Mariappan et al., 2011; Mock et al., 2015; Stefanovic and Hegde, 2007). The TRC40-TA complex is targeted to the ER membrane via its ER receptor, the WRB/CAML complex, followed by substrate release and insertion into the membrane (Vilardi et al., 2011; Vilardi et al., 2014; Yamamoto and Sakisaka, 2012). The incorporation of Bag6 into the mammalian TA targeting pathway results in substantial reorganization of the molecular architecture of the sorting complex. In yeast, the N-terminal domain of Get5 forms a direct complex with the C-terminal binding groove of Get4 (Chartron et al., 2010). Get5 also contains a homodimerization domain at its C-terminus (Chartron et al., 2012c).

Bcl-2 associated Athanogene 6 (Bag6, also known as Scythe or BAT3)

Bag6 is a uniquely metazoan protein that appears late in opisthokont evolution (Fig. 2.11). Opisthokonts include a broad range of eukaryotes: the metazoan, choanozoan and fungal lineages. The term opisthokont was first coined in 1987 by Thomas Cavalier-Smith (British Mycological Society. Symposium (1986 : University of Bristol) et al., 1987), who suggested a close evolutionary relationship between fungi and metazoans; analysis of 16S-like rRNA sequences (Wainright et al., 1993) and key protein sequences (Baldauf and Palmer, 1993) confirmed the hypothesis. A commonly isolated isoform of Bag6 is ~130 kDa with an N-terminal ubiquitin-like (UBL) domain, a large predicted-to-be disordered proline-rich domain and a C-terminal putative BAG domain (Fig. 6.2).

Bag6 was initially described as part of the gene cluster that included the human major histocompatibility complex (MHC) class III on chromosome 6, resulting in its first name, HLA (human leukocyte antigen)-*B*-*A*ssociated *T*ranscript 3 (BAT3). The genomic localization suggested a role in immune response, which has been supported by evidence of Bag6 involvement in TH1 cell survival (Rangachari et al., 2012), natural killer cell cytotoxicity (Pogge von Strandmann et al., 2007), and MHC class II molecule presentation (Kamper et al., 2012a; Pai et al., 2002).

Based on limited sequence homology to the *B*cl-2 associated *A*thanogene (BAG) family and apparent Hsc70 regulating activity, BAT3 was renamed Bag6 in 2001 (Thress et al., 2001). BAG family of proteins are multidomain proteins characterized by a conserved C-terminal BAG domain (Qin et al., 2016; Takayama and Reed, 2001). The BAG domain is usually 110 to 130 residues long and folds into a three helix bundle that acts as a nucleotide

exchange factor for Hsp70 family of chaperones (Arakawa et al., 2010; Briknarova et al., 2001). Excess BAG domain has been shown to inhibit Hsc70 refolding activity *in vitro* (Takayama et al., 1997). Initial searches that identified four additional functional homologues of Bag1 (Bag2-5) (Takayama et al., 1999) failed to detect Bag6.

In addition to its roles in immune regulation, there is increasing evidence for nuclear functions of Bag6. Bag6 contains a functional nuclear localization sequence (NLS) (Manchen and Hubberstey, 2001) and endogenously localizes both in the nucleus and the cytoplasm (Kamper et al., 2012b; Manchen and Hubberstey, 2001; Pogge von Strandmann et al., 2007; Tsukahara et al., 2009; Wu et al., 2004). Bag6 NLS is removed by apoptosis inducers (Preta and Fadeel, 2012a, b) or masked by TRC35 (Wang et al., 2011a). In addition, inhibition of Crm1-dependent nuclear export prevents trafficking of Bag6 from the nucleus (Kamper et al., 2012b). Nuclear Bag6 has been shown to play a role in DNA damage response (Krenciute et al., 2013), p53 acetylation (Sasaki et al., 2007), cell cycle regulation (Yong and Wang, 2012), p300-mediated acetylation (Sebti et al., 2014a) and histone methylation (Nguyen et al., 2008; Wakeman et al., 2012). Secreted forms of Bag6 have also been observed, and which appear to modulate natural killer cell cytotoxicity in chronic lymphocytic leukemia patients (Reiners et al., 2013).

Bag6 has been implicated in a variety of additional cellular processes via numerous binding partners such as apoptosis inducing factor (AIF), Reaper, osteopontin, Rpn10c, SGTA, gp78, UbxD8, Rpt5, BORIS, SCP and RNF126 (Akahane et al., 2013; Desmots et al., 2008; Goto et al., 2011; Kikukawa et al., 2005; Long et al., 2012; Nguyen et al., 2008;

Rodrigo-Brenni et al., 2014; Thress et al., 2001; Xu et al., 2013) (Fig. 6.2). Importantly, multiple studies have independently shown the interaction between components of the TA targeting pathway (SGTA, TRC35, Ubl4A) and Bag6, demonstrating that Bag6 exists in complex with TR35 and Ubl4A in the cytoplasm (Hessa et al., 2011; Krenciute et al., 2013; Mariappan et al., 2010; Mariappan et al., 2011; Wang et al., 2011b), with frequent interactions with SGTA (Chartron et al., 2012b; Leznicki and High, 2012; Leznicki et al., 2013; Winnefeld et al., 2006; Xu et al., 2012). Bag6 has also been demonstrated to homo-oligomerize *in vitro*, introducing greater complexity to the system. This builds a picture of Bag6 as a central hub for diverse physiological network of proteins.

Such diverse cellular roles of Bag6 are reflected by its implications in a variety of diseases. Mutations of Bag6 are implicated in Kawasaki Syndrome (Hsieh et al., 2010), rheumatoid arthritis (Harney et al., 2008), lung cancer (Chen et al., 2014; Etokebe et al., 2015b; Wang et al., 2008; Zhao et al., 2014), type 1 diabetes (Degli-Esposti et al., 1992), myasthemia gravis, thymus hyperplasia (Vandiedonck et al., 2004) and osteoarthritis (Etokebe et al., 2015a). Combined, these data establish Bag6 as a critical player in cellular and organismal biology.

In contrast to the complicated plethora of data on the function of Bag6, precise biochemical and structural description had been rare when this work was conceived. This dissertation focuses on structural and biochemical characterization of a trimeric Bag6 complex that includes TRC35 and Ubl4A. Chapter 2 describes the structural characterization of the Bag6-Ubl4A complex and the Bag6-TRC35 complex. TRC35 and Ubl4A bind Bag6 at

distinct binding sites; Ubl4A localizes to the putative BAG domain and TRC35 binds at the region that includes the NLS of Bag6. The Bag6-Ubl4A structure reveals that the Bag6-BAG domain does not resemble a canonical BAG domain, establishing Bag6 as a misnomer. TRC35 masking NLS provides the structural basis for nucleocytoplasmic distribution of Bag6. Chapter 3 describes the biochemical assays used to functionally characterize the BAG domain and the isolated minimal Bag6 complex, which includes a truncated C-terminal Bag6, truncated TRC35 and full length Ubl4A. An Hsc70 refolding assay is used to confirm that the Bag6 "BAG" domain is not a BAG domain. We also developed an *in vitro* TA transfer assay to demonstrate that the minimal Bag6 complex identified in this study is sufficient for facilitating TA handoff, establishing it as a TA targeting module. In Chapter 4, the regulation of nucleocytoplasmic distribution of Bag6 by TRC35 is explored using biochemical and cell biology techniques. Chapter 5 concludes by contextualizing the findings of this dissertation in the body of work on protein targeting and quality control.

C h a p t e r 2

STRUCTURAL CHARACTERIZATION OF THE BAG6-TRC35-UBL4A COMPLEX

Parts of this chapter were first published in

Mock, J.-Y., Chartron, J.W., Zaslaver, M., Xu, Y., Ye, Y., and Clemons, W.M. Jr. (2015) Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain. *Proc Natl Acad Sci USA*. 112(1): 106-11 doi: 10.1073/pnas.1402745112

Mock, J.-Y., Xu, Y., Ye, Y., and Clemons, W.M. Jr. (2017) Structural basis for regulation of nucleocytoplasmic distribution of Bag6. *Submitted*.

Jee-Young Mock solved the crystal structures, carried out yeast 2-hybrid analysis, biophysical experiments with purified proteins, and sequence analysis.

Abstract

Bcl2-associated athanogene cochaperone 6 (Bag6) is a uniquely metazoan protein involved in a diverse array of cellular processes. It is part of the heterotrimeric Bag6 complex, which also includes ubiquitin-like 4A (Ubl4A) and transmembrane domain recognition complex 35 (TRC35). The Bag6 complex plays a central role in the mammalian tail-anchor protein targeting pathway, mislocalized protein degradation pathway and the endoplasmic reticulum-associated degradation pathway. Here we define the architecture of the Bag6 complex, demonstrating that both TRC35 and Ubl4A have distinct C-terminal binding sites on Bag6 defining a minimal Bag6 complex. The crystal structure of the Bag6-Ubl4A dimer demonstrates that Bag6-BAG is not a canonical BAG domain. Instead, its main function is to dimerize with the well-conserved dimerization domain of Ubl4A. The crystal structure of Bag6 and its cytoplasmic retention factor TRC35 reveals remarkable structural conservation of Get4/TRC35 throughout opisthokont lineage except at the C-terminal Bag6-binding groove, which diverged to accommodate Bag6. Together these data advance our molecular understanding of the Bag6-TRC35-Ubl4A complex.

Introduction

Bcl2-associated athanogene cochaperone 6 (Bag6, also known as "BAT3" or "Scythe") plays a central role in membrane protein quality control, with additional links to apoptosis, transcription regulation, and immunoregulation (for reviews, see (Kawahara et al., 2013; Krenciute et al., 2013; Lee and Ye, 2013)). Recent studies demonstrated that Bag6 forms a heterotrimeric Bag6 complex with ubiquitin-like 4A (Ubl4A) and transmembrane domain recognition complex 35 (TRC35) that determines the fate of membrane proteins (Mariappan et al., 2011; Wang et al., 2011b) in tail-anchor (TA) protein targeting (Stefanovic and Hegde, 2007), mislocalized protein degradation (Hessa et al., 2011), and endoplasmic reticulum (ER)-associated protein degradation (Wang et al., 2011b). The many roles of the Bag6 complex likely are centered on its ability to bind exposed hydrophobic regions of proteins, such as transmembrane domains.

In mammals, Bag6 has been shown to be critical in the targeting of TA proteins to the ER by the transmembrane recognition complex (TRC) pathway (Mariappan et al., 2011), a process best understood in the equivalent fungal guided entry of tail-anchored proteins (GET) pathway (Chartron et al., 2012a; Hegde and Keenan, 2011). Although Bag6 is absent in fungi, the analogous yeast complex contains two proteins, Get4 and Get5/Mdy2, which are homologues of the mammalian proteins TRC35 and Ubl4A, respectively. In yeast, these two proteins form a heterotetramer that regulates the handoff of the TA protein from the cochaperone small, glutamine-rich, tetratricopeptide repeat protein 2 (Sgt2) [small glutamine-rich tetratricopeptide repeat-containing protein (SGTA) in mammals] to the delivery factor Get3 (TRC40 in mammals) (Chartron et al., 2010;

Chartron et al., 2012c; Gristick et al., 2014; Wang et al., 2010). It is expected that the mammalian homologs, along with Bag6, play a similar role (Chartron et al., 2012b; Leznicki and High, 2012; Leznicki et al., 2013; Xu et al., 2012).

Bag6 also interacts with other proteins such as apoptosis-inducing factor, glycoprotein 78 (gp78), regulatory particle 5, and brother of regulator of imprinted sites (BORIS) (Akahane et al., 2013; Desmots et al., 2008; Kikukawa et al., 2005; Long et al., 2012; Nguyen et al., 2008; Thress et al., 2001; Xu et al., 2013) and can homo-oligomerize (Xu et al., 2013). These findings build a picture of Bag6 as a central scaffolding factor linking various cellular processes. A variety of diseases, ranging from cancer to autoimmune disorders and diabetes, are linked to Bag6 (Degli-Esposti et al., 1992; Harney et al., 2008; Hsieh et al., 2010; Vandiedonck et al., 2004; Wang et al., 2008).

Despite this demonstrated importance, structural characterization of the Bag6 complex is lacking. The longest and most common isoform of the Bag6 gene encodes an 1,132-aa protein (Banerji et al., 1990) with an N-terminal ubiquitin-like (UBL) domain that has been characterized structurally (PDB ID codes 4EEW, 4DWF, and 1WX9), a large proline-rich central domain that is predicted to be unstructured, and a C-terminal predicted BAG domain (Bag6-BAG).

In this study, we map the TRC35- and Ubl4A-binding regions to the C terminus of Bag6. Based on these results, the structure of the complex between the heterodimerization domains of Bag6 and Ubl4A was solved, revealing unexpected structural homology to Get5 and showing that the Bag6-BAG is a “mock” BAG domain. The structure of the

complex between the heterodimer of Bag6 and TRC35 presented here provides a structural basis for regulation of nucleocytoplasmic distribution of Bag6 by TRC35. Furthermore, the structures reveal that despite the changes in architecture, the overall folds of the TA sorting complex have been remarkably conserved throughout opisthokont evolution.

Results

TRC35 binds the Bag6 nuclear localization sequence, whereas Ubl4A binds the BAG domain

To define the molecular architecture of the heterotrimeric Bag6 complex (Bag6, Ubl4A, and TRC35), a series of yeast two-hybrid assays were performed. Bag6 was divided into five fragments: A (amino acids 1–225), B (amino acids 226–399), C (amino acids 400–659), D (amino acids 660–950), and E (amino acids 951–1,126), with the activating domain attached at the N terminus (Fig. 2.1A); TRC35 or Ubl4A contained N-terminal DNA-binding domains. TRC35 and Ubl4A both showed a positive interaction with the C-terminal Bag6E fragment that contains the nuclear localization sequence (NLS) and the BAG domain (Fig. 2.1B). To refine the interfaces, Bag6E was divided further into an N-terminal domain, an NLS domain, and the putative BAG domain (E_N, E_{NLS}, and E_{BAG} respectively) (Fig. 2.1A). TRC35 showed an interaction with E_{NLS} (Fig. 2.1B) confirming the in vivo result that TRC35 masks the Bag6-NLS, preventing nuclear targeting (Wang et al., 2011b). Ubl4A showed an interaction with E_{BAG} (Fig. 2.1B), a surprising result because none of the five previously characterized BAG domains was known to form stable interactions with other proteins.

According to sequence alignment, Ubl4A lacks the Get5 N-terminal domain, and TRC35 lacks the β-loop in Get4 (Fig. 2.2A), both of which are involved in the Get4 interface with Get5 (Chartron et al., 2010). This difference suggests there are different interactions in the Bag6 complex. One possibility is that the region around the Bag6 NLS acts structurally like the Get5 N domain by binding the C domain of TRC35. To confirm this hypothesis, a two-hybrid experiment was performed with TRC35 split into either an N

domain (TRC35-N, residues 1–157) or a C domain (TRC35-C, residues 158–327), as had been done previously for Get4 (Fig. 2.2A) (Chang et al., 2010). As predicted, Bag6E showed a clear interaction with TRC35-C and no interaction with TRC35-N (Fig. 2.2B). Surprisingly, the smaller $Bag6E_{NLS}$ did not interact with TRC35-C (Fig. 2.2B) despite the previously seen interaction with full-length TRC35 (Fig. 2.1B). This interaction was restored with the longer $Bag6E_{N,NLS}$, suggesting that additional contacts are required to form a stable interaction. This extended region defines a minimal complex with TRC35 bound to the C terminus of Bag6 in close proximity to Ubl4A, similar to the architecture found in the yeast Get4-5 complex.

The crystal structure of Bag6-BAG/Ubl4AC

For structural characterization, multiple variants of the $Bag6_{min}$ complex were pursued for crystallization. One consisting of the Ubl4A C-terminal dimerization domain (Ubl4A-C) and the Bag6 BAG domain (Bag6-BAG, residues 1054-1107) resulted in well-formed crystals. A complete 2.1 Å native data set was collected in the space group $P2_1$ and phased using an iodide derivative. The final refinement resulted in an R_{free} of 28.0% and good statistics (Table 2.1). Both domains are primarily helical (Fig. 2.3A) with an extensive dimer interface dominated by conserved hydrophobic residues that results in 2485 $Å^2$ of buried surface (Fig. 2.3A). Bag6-BAG contains three helices in an extended conformation making few intramolecular contacts. Ubl4A-C contains three helices with the first two forming an interface and the short third helix wrapping around Bag6-BAG (Fig. 2.3A). All of the conserved hydrophobic residues of Bag6-BAG participate in dimerization (Fig. 2.3B).

While fungal Get5 forms a stable homo-dimer mediated by Get5-C, Ubl4A alone is primarily a monomer (Chartron et al., 2012c). As Ubl4A forms a hetero-dimer with Bag6 one might expect a novel fold for Ubl4A-C; instead, the Ubl4A-C hetero-dimerization domain has an identical structure to the Get5-C homo-dimerization domain, an RMSD of 0.94 Å for equivalent backbone residues (Figs. 2.4A and B). The same hydrophobic residues that form the core in Get5 homo-dimerization—W179, I182, L186, F190, V200, L204, W208—are conserved in Ubl4A—W96, I99, L103, F107, V115, L119, Y123 (Figs. 2.4B and C).

Bag6-BAG is a 'mock' BAG domain

Based on the structural characterization, the Bag6-BAG is shorter (47 residues) than canonical BAG domains (6-112 residues) and the three helices do not form a BAG-like three-helix bundle, with different orientations of the few residues equivalent to those involved in Hsp70 binding of other BAG domains (Figs. 2.5A and B).

Furthermore, circular dichroism of Bag6-BAG alone indicates no stable secondary structure (Fig. 2.6). This implies that a primary role for Bag6-BAG is to hetero-dimerize with Ubl4A, as the two are found in a stoichiometric complex (Hessa et al., 2011; Mariappan et al., 2010).

The crystal structure of Bag6-NLS/TRC35 complex

A complex between TRC35 (residues 23 to 305) and a minimal TRC35-binding domain of Bag6 that contains the NLS (residues 1000-1054) was co-expressed, purified and crystallized. A dataset from a single crystal was collected to 1.8Å resolution and phases

were obtained using single wavelength anomalous dispersion from a rubidium derivative. The final model refined to 1.8 - 40Å had good statistics with an R_{work}=16.5% and an R_{free}=20.1% and no residues in the disallowed region of the Ramachandran. Full crystallographic statistics are provided in Table 2.2. All TRC35 residues in the crystallized construct were visible in the electron density except for S137 and G138 that are disordered in the loop between α6 and α7. For Bag6, residues 1002-1043 were resolved with only a few terminal residues ambiguous in the electron density.

TRC35 has the same overall architecture as the fungal Get4 structures (Fig. 2.8A), revealing that the Get4 fold has been conserved across opisthokonta and is likely conserved across all eukaryotes. Opisthokonta includes metazoan, choanozoan, and fungal lineages (Fig. 2.11), which share common ancestry as determined by analysis of 16S-like rRNA sequences (Wainright et al., 1993) and several protein sequences (Baldauf and Palmer, 1993). The first 15 α-helices form an α-solenoid fold that can be divided into an N- and C-terminal halves (Fig. 2.7). Alignment of the NTD between TRC35 and Get4 (PDB ID: 3LKU) results in an RMSD = 1.380Å for that region, while the equivalent alignment in the CTD results in an RMSD = 3.429Å (Fig. 2.8). As seen in Get4 (Chartron et al., 2010), there is likely some flexibility between these two halves based on differences across crystal forms. As predicted from sequence alignment analysis, the β-hairpin in Get4 is replaced by a shorter loop in TRC35 (Fig. 2.7C, arrows). The C-terminal α16 is flanked by two extended stretches that cover part of Bag6 (Fig. 2.7A).

The Bag6 NLS region wraps around the TRC35-CTD (Fig. 2.7B, light pink) at an interface that resembles the interaction of Get5 with Get4 (Fig. 2.7C). The interaction is stabilized by two hydrophobic interface sites. In interface I, Bag6 α1 and α2 dock into a conserved pocket created by α12 and α13 of TRC35 (Fig. 2.9B) and include W1004, V1008 and W1012 from Bag6 and V254, V257, F242, L258 and Y262 from TRC35 (Fig. 2.7E). In fungal Get4-5, the Get5 N-terminal helix docks into a groove formed by α12, α13, and the β-hairpin of Get4 (Fig. 2.7C). The missing β-hairpin in TRC35 results in the Bag6 helix shifting away from α11 towards α12 and α13 near the bottom of TRC35 (Fig. 2.7C, arrows). The differences in the interface result in changes in the arrangement of α11, α13, and α15 of TRC35 (Fig. 2.7C) relative to those of Get4. Interface II is less extensive involving fewer residues, L1032, Y1036, M1040 of Bag6 and F195, M271 of TRC35 (Fig. 2.7E). While the connecting loops between the two interfaces are well-ordered in both contexts, the Bag6 loop makes fewer contacts to TRC35 than the extensive interactions in the Get5-loop to Get4 (Fig. 2.7B).

The structure reveals how TRC35 masks the Bag6 bipartite NLS (Fig. 2.7B). The conserved first basic cluster of the predicted NLS (K_{1024}RVK) is sequestered between interfaces I and II (Fig. 2.7B, sticks). Only the first lysine residue of the second cluster (K_{1043}RRK) is resolved in our model (Fig. 2.7B, sticks). The truncated C-terminal residues of TRC35 in our construct are predicted to be disordered (Linding et al., 2003) and include conserved multiple negative charges (five Glu and three Asp (Fig. 2.10)) that would be in position to mask the second NLS basic cluster, K_{1043}RRK, through charge-mediated interactions.

TRC35 and TRC40 are Conserved throughout Opisthokont Evolution

Unlike Get4, the CTD of TRC35 does not bind the homologous counterpart Ubl4A, and the change in binding partner is reflected at the sequence level by the lack of the β-hairpin in TRC35 and the lack of the N-terminal extension in Ubl4A. We hypothesized that the loss of the fungal features may coincide with the appearance of Bag6. To address this, we turned to phylogenetic analysis. First, sequence alignments of Bag6, TRC35, Ubl4A, SGTA and TRC40 were used to generate sequence motifs using the Multiple Em for Motif Elicitation suite (MEME) (Bailey et al., 2009). Species were selected based on fully sequenced genomes from representative lineages of Eukaryota. The MEME motifs were used to search each of the genomes shown (Fig. 2.11).

Consistent with our hypothesis, in opisthokonts the split from the fungal lineage coincides with disappearance of the Get4/TRC35 β-hairpin (Figs. 2.10 and 2.11) and the Get5/Ubl4A N-terminal extension. However, species lacking Get4/TRC35 β-hairpin do not immediately gain Bag6. In TA targeting, Bag6 centrally bridges the Ubl4A-mediated TA substrate hand-off from SGTA to TRC40 which itself is mediated by TRC35 (Mock et al., 2015). Curiously, various lineages seem to lack either Sgt2/SGTA or Get5/Ubl4A with Get5/Ubl4A most often missing. In both mammals and yeast, the interaction between the two proteins (Chartron et al., 2011) has been demonstrated to be crucial to TA protein transfer (Mateja et al., 2015; Mock et al., 2015) and ER-associated degradation system (Xu et al., 2013). In species that lack TRC35 β-hairpin, Bag6, and Ubl4A, such as *Demosponge* or *Caenorhabditis*, it is unclear how these processes are

modulated. It is likely that functionally homologous proteins with low sequence homology exist.

Get4/TRC35 and Get3/TRC40 are conserved throughout eukaryotic evolution and seem to occur as a pair in all opisthokonts and in most eukaryotes, suggesting the essentiality of the two proteins in the pathway. Consistent with this notion, residues at the predicted TRC35-TRC40 interface and the TRC35-Bag6 interface in TRC35 are highly conserved (Figs. 2.8A and 2.8C).

Discussion

The Bag6 complex has been implicated in various cellular pathways necessitating a description of its molecular architecture. In this study, we have examined the Bag6 trimeric complex and solved high-resolution crystal structures of the Bag6-Ubl4A and Bag6-TRC35 complexes.

The initial characterization of Bag6 BAG domain relied on limited and inaccurate sequence homology (Thress et al., 2001) without structural information. Comparison of high resolution structures of BAG domains, as well as circular dichroism data that show Bag6-BAG alone is intrinsically disordered, suggests that the Bag6-BAG domain is not a canonical BAG domain. Biochemical data from the previous study that demonstrated functional equivalence between Bag6 BAG and canonical BAG domains will be addressed later in this work.

The first functional annotation of Scythe (the Bag6 Xenopus homolog) was an ability to bind Reaper, an apoptosis-inducing protein in Drosophila, thus inhibiting Reaper-induced apoptosis in Xenopus oocyte extracts (Thress et al., 1999; Thress et al., 1998). Reaper induces apoptosis in a variety of model systems, including Xenopus oocyte extract (Evans et al., 1997), SF-21 insect cells (Vucic et al., 1997), and HeLa human cancer cells (Tait et al., 2004). Sequence alignment of the conserved Ubl4A dimerization domain and Reaper reveals the conservation of most of the residues involved in Bag6-BAG/Ubl4A-C dimerization (Fig. 2.4), suggesting that Reaper may disrupt the Bag6/Ubl4A interaction. This hypothesis would be consistent with Reaper's binding the 312-residue C-terminal

truncation of Scythe (ScytheC312), leading to the release of bound factors (Thress et al., 1998).

The remarkable conservation of TRC35 in both sequence and structure from yeast to human can be attributed to the role of TRC35 as a hub of protein-protein interactions in the TRC pathway. The most important interaction, based on its complete conservation, is between TRC35 and TRC40 (Fig. 2.9) and likely drives evolution of the components in the pathway. This interaction is critical for the regulation of TA protein transfer (Gristick et al., 2014; Mock et al., 2015; Shao et al., 2017) and, as shown in fungal Get4, for release of Get3/TRC40 from the ER membrane to restore the cytosolic pool of complex ready for TA protein transfer (Rome et al., 2014). The importance of linking TRC35 homologs to Ubl4A homologs, while less conserved, is sustained in humans despite the loss of structural features that allow these two proteins to interact with each other. In this case, a new component Bag6 is introduced, which serves as a scaffold to link TRC35 to Ubl4A. The Bag6-TRC35 structure reveals that TRC35 binds this novel binding partner utilizing conserved patches in a distinct fashion (Fig. 2.9). This led to the expansion of the TRC35 protein-protein interaction network while maintaining its crucial interaction with TRC40. Our findings establish support the model in which TRC35 is an important hub of a protein-protein interaction network with its interaction with TRC40 at the core. Higher eukaryotes expanded on this network by the addition of Bag6.

The emergence of Bag6 in metazoan evolution is also an intriguing question. How did such a complex and extended scaffolding protein emerge? The exact genetic events are difficult to pinpoint because the DNA sequences that can provide clues to events such as previous recombination or gene duplication are not under selective pressure, and are therefore quickly removed (Moore et al., 2008). Instead, the protein sequence of Bag6 can be used to speculate the evolutionary origins of Bag6.

There are evolutionary trends that can be applied to deduce the evolutionary origin of Bag6. Bioinformatic survey of multidomain proteins suggest that novel multidomain proteins are often created by single domain insertions and deletions at the N- and C-termini (Bjorklund et al., 2005; Kummerfeld and Teichmann, 2005). These domain rearrangement events seem to occur more frequently in metazoans than in other phylogenetic groups (Ekman et al., 2007). Similarly, the N-terminal UBL domain and C-terminal TRC35 and Ubl4A binding domains are conserved structural elements that could have been added through independent domain insertion and rearrangement events.

A similar phenomenon has been observed in the evolution of the membrane-associated guanylate kinase (MAGUK) family of proteins, which are involved in cell-cell communication in metazoa. Putative ancestral MAGUK-like genes were discovered in unicellular ancestors to metazoans, *Capsaspora owczarzaki* and *Monosiga brevicollis*, as well as the earliest surviving metazoan *Amphimedon queenslandica* (de Mendoza et al., 2010). These ancestral genes contain parts of the domains that have been expanded upon in modern MAGUK proteins by domain insertion and rearrangement events.

These results support a model in which the primary role of the Bag6 C terminus is to bridge TRC35 and Ubl4A. Possible Bag6 dimerization would form a heterohexamer, creating a complex analogous to the Get4-5 heterotetramer found in yeast and providing strong structural parallels in TA targeting (Fig. 2.12). Several unanswered questions about the Bag6 complex remain. Does Bag6-BAG domain behave like a canonical BAG domain? Is the structurally characterized Bag6 complex functionally equivalent to the fungal Get4-5 complex? Can disrupting the Bag6-TRC35 interface lead to changes in the nucleocytoplasmic distribution of Bag6? Biochemical and cell biological methods will be used to answer these questions in the following chapters.

Figures

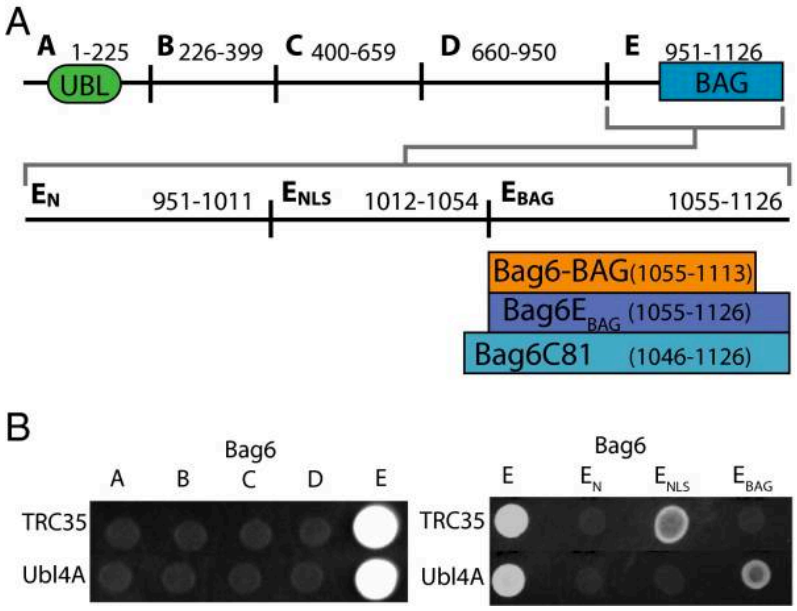


Figure 2.1. Bag6 has distinct binding sites for TRC35 and Ubl4A at its C-terminus

(A) Scheme of the five different Bag6 fragments and of the sub-fragments of Bag6E, which is further divided into N-terminus (E_N), the NLS (E_{NLS}), and a fragment containing only the putative BAG domain (E_{BAG}). (B) Yeast two-hybrid assay between Bag6 fragments and either TRC35 or Ubl4A. The A fragment contains the UBL domain and the E fragment contains the NLS and putative BAG domain. Yeast two-hybrid assay of TRC35 or Ubl4A between the Bag6 E fragments.

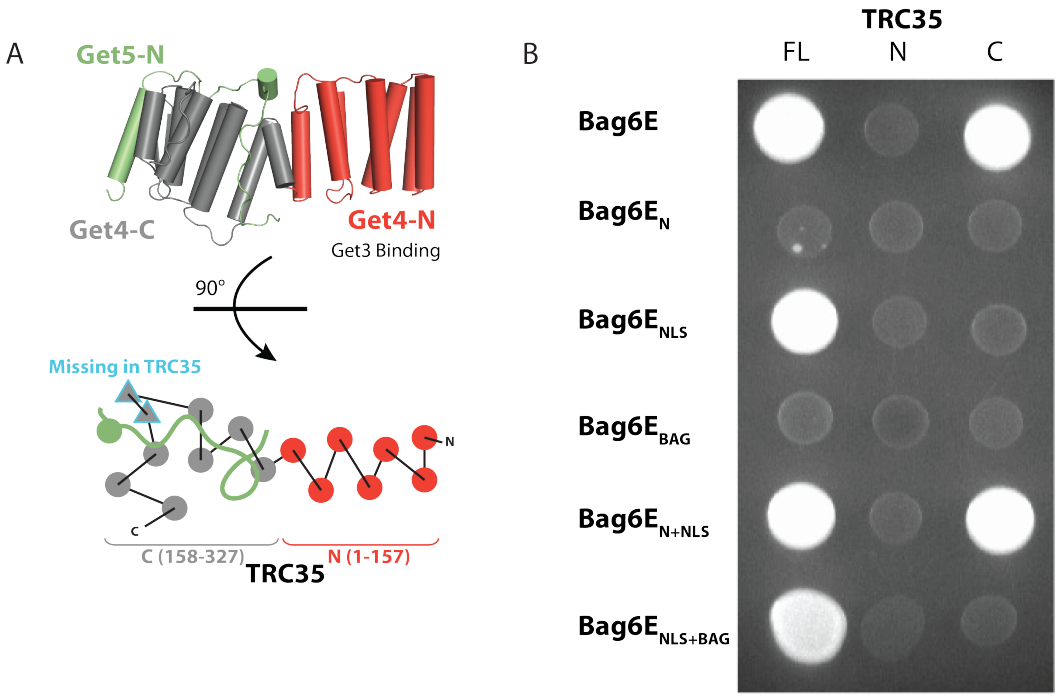


Figure 2.2 Defining the interaction between Bag6 and TRC35 by yeast two-hybrid assay

(A, Upper) Diagram of the Get4/Get5-N complex (PDB ID code 3LKU). Get4 is shown in gray and red, and Get5 is shown in green. (Lower) A top-view schematic representation of the architecture. The two β-strands in Get4 that are missing in TRC35 are outlined in blue. Ubl4A does not contain a Get5-N equivalent. (B) Bag6E fragments containing the activating domain were combined with full-length TRC35 (FL), TRC35-N (residues 1–157), or TRC35-C (residues 158–327) conjugated to the binding domain. TRC35-N and TRC35-C were defined by sequence alignment with Get4.

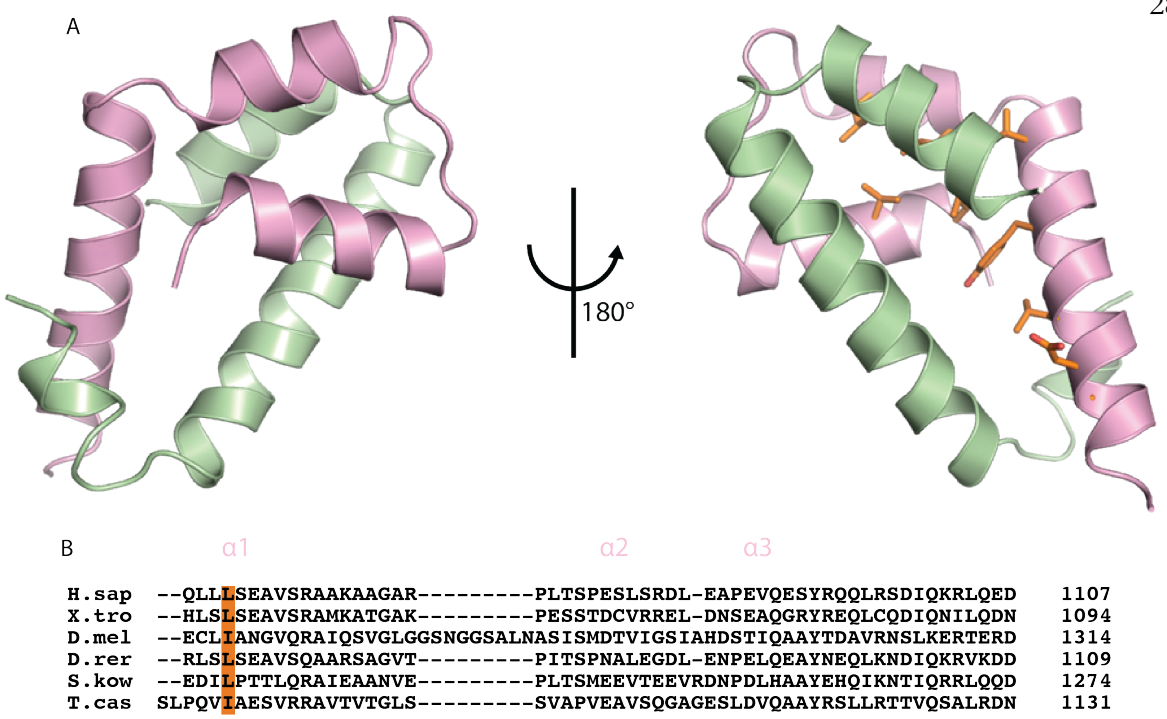


Figure 2.3. Crystal structure of Bag6-BAG/Ubl4A-C heterodimer

(A) The overall structure of Bag6-BAG/Ubl4A-C heterodimer in ribbon representation with Bag6-BAG in cyan and Ubl4A-C in magenta. Hydrophobic residues in Bag6 involved in packing are highlighted as orange sticks. (B) Sequence alignment of Bag6-BAG (H.sap *Homo sapiens*, X.tro *Xenopus tropicalis*, D.mel *Drosophila melanogaster*, D.rer *Danio rerio*, S.kow *Saccoglossus kowalevskii*, and T.cas *Tribolium castaneum*). The secondary structure based on the structure is highlighted above the text. The conserved hydrophobic and aromatic residues involved in the hydrophobic packing interactions between Bag6-BAG and Ubl4A-C are highlighted in orange. The extended Drosophila melanogaster sequence is a predicted protein sequence based on theoretical translation, and may not reflect a physiological isoform.

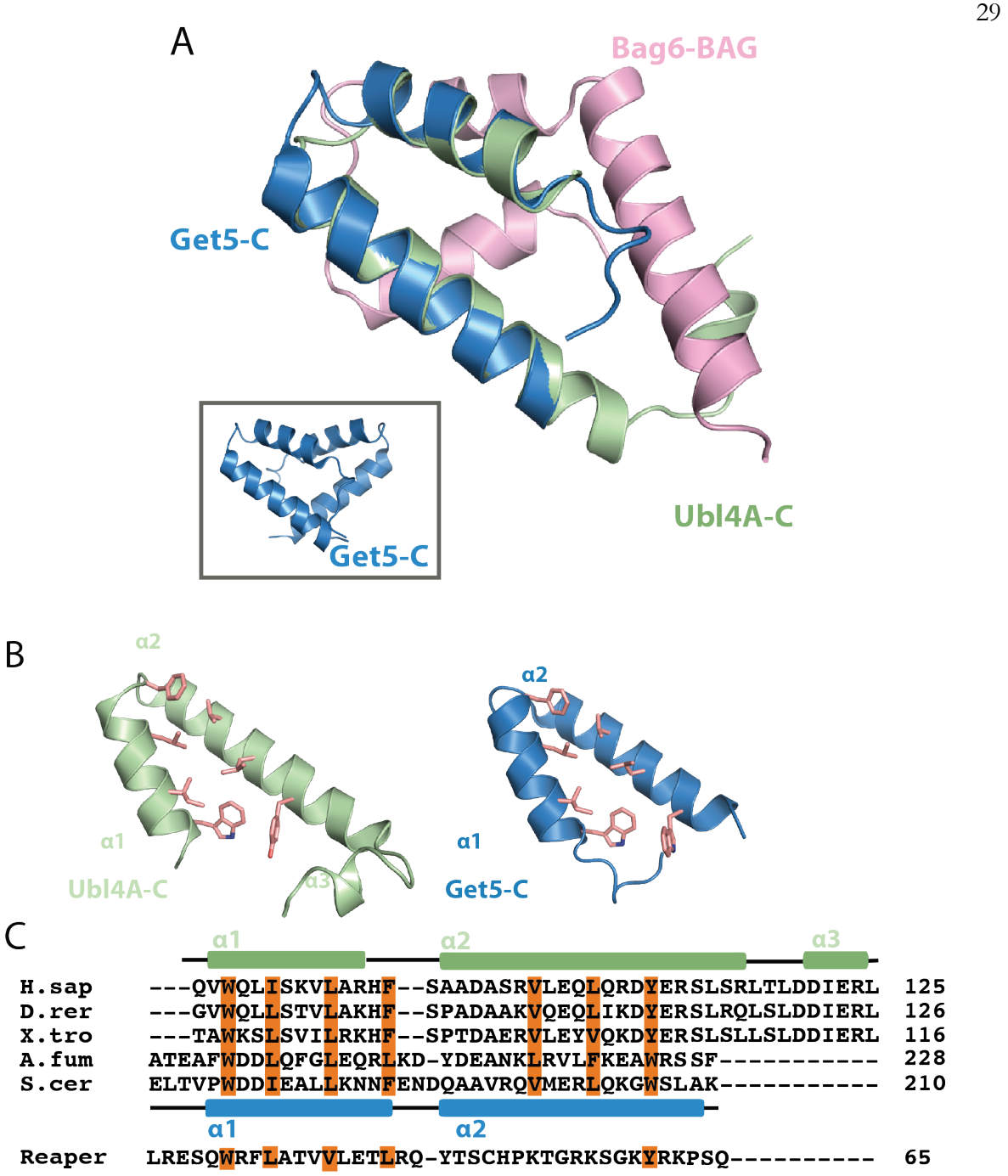


Figure 2.4. Ubl4A-C and Get5-C are structurally homologous

(A) A monomer of Get5-C (green) (PDBID: 3VEJ) and Ubl4A-C (pink) overlaid. Bag6-BAG is included in cyan. The inset shows the Get5-C homo-dimer. (B) Ubl4A-C and

Get5-C are juxtaposed with conserved residues involved in dimerization highlighted as orange sticks. (C) Sequence alignment of Ubl4A-C homologs (H.sap *Homo sapiens*, D.rer *Danio rerio*, X.tro *Xenopus tropicalis*, A.fum *Aspergillus fumigatus*, and S.cer *Saccharomyces cerevisiae*) and the Drosophila apoptosis-inducing protein Reaper. The Ubl4A-C and Get5C secondary structure are shown above (pink) and below (green), respectively. Conserved hydrophobic residues involved in dimerization are highlighted in orange.

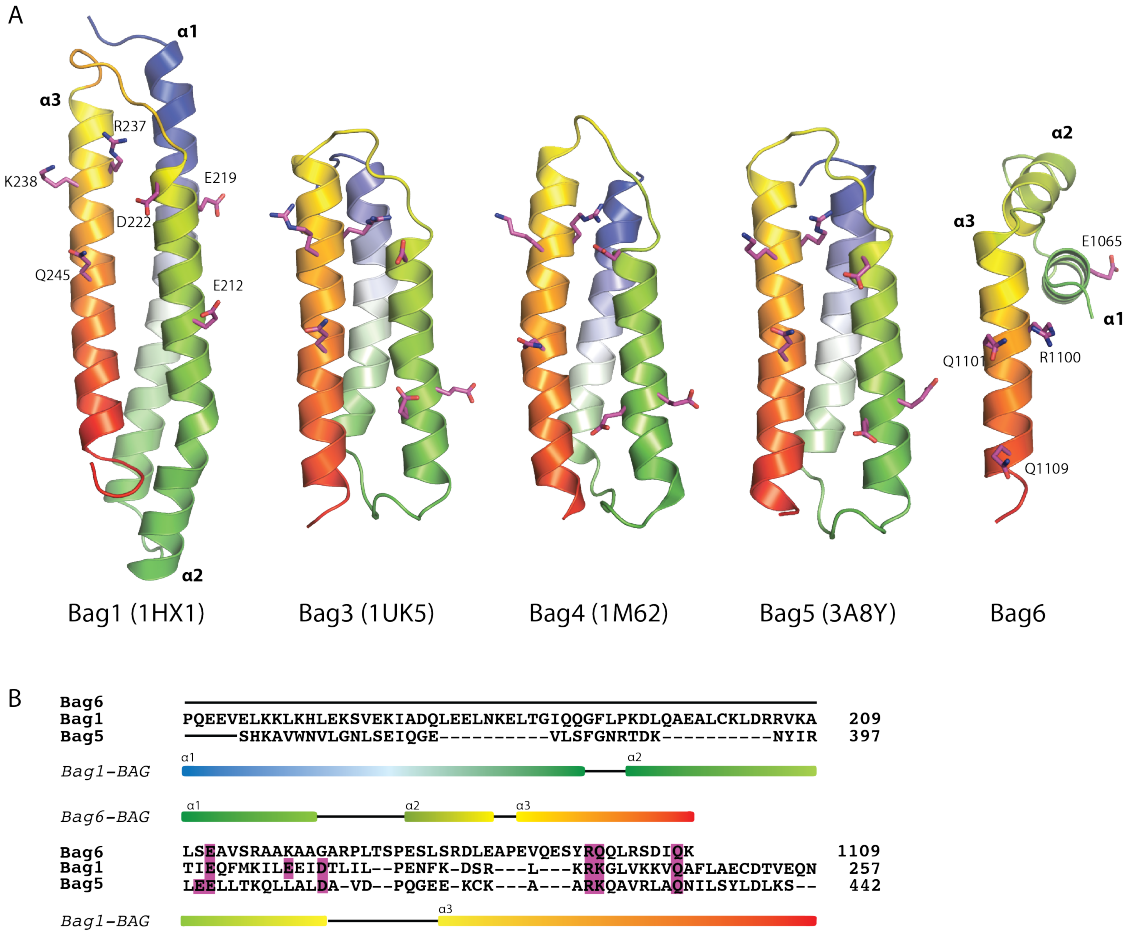


Figure 2.5. A comparison of Bag6-BAG to canonical BAG domains

(A) Published structures of BAG domains are shown as ribbons similar to Bag6-BAG. Residues involved in Hsp/Hsc70 binding are highlighted as magenta sticks. (B) Sequence alignment of human Bag1-BAG, Bag5-BAG5, and Bag6-BAG with known secondary structure indicated.

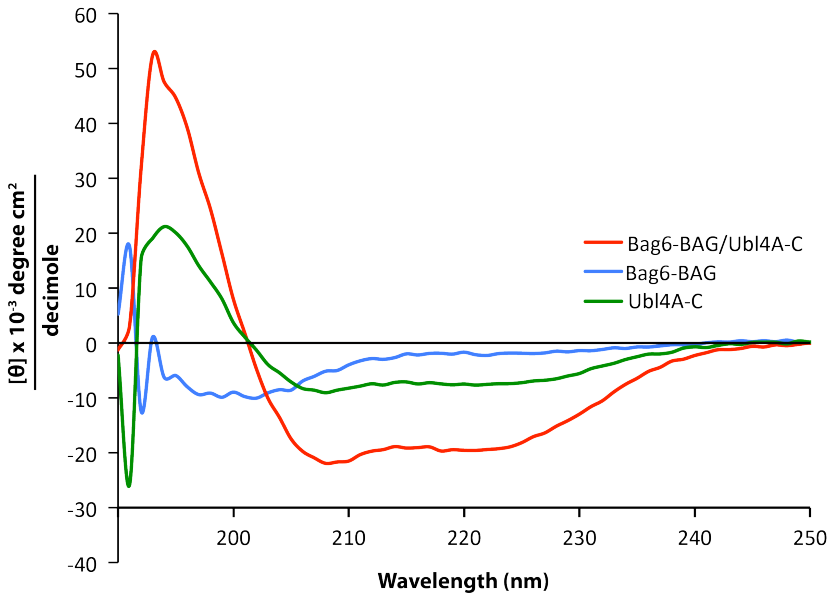


Figure 2.6. Circular dichroism spectra of Bag6-BAG, Ubl4A-C and the complex of the two

Spectra for Bag6-BAG (blue), Ubl4A-C (green), and Bag6-BAG/Ubl4A-C (red).

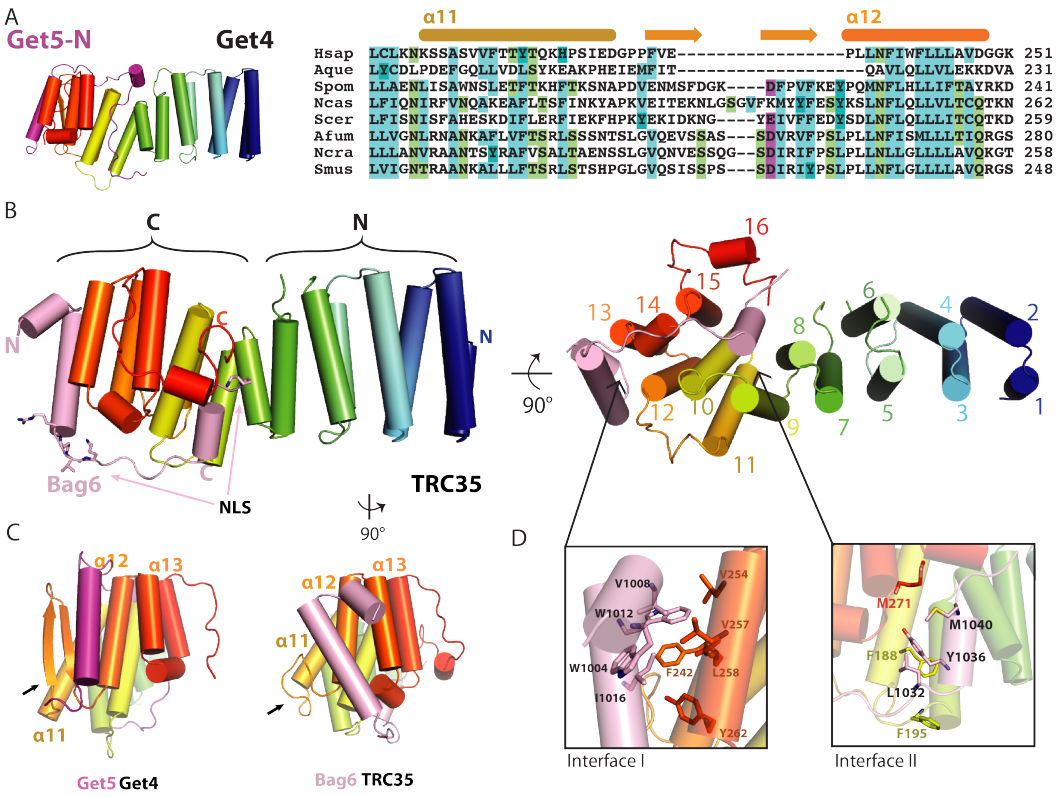


Figure 2.7. The crystal structure of the Bag6-NLS/TRC35 heterodimer

(A) The structure of Saccharomyces cerevisiae Get4-Get5N complex (PDBID: 3LKU), Get4 in rainbow and Get5 in magenta. Sequence alignment of TRC35/Get4 from helix 11 to 12 of two metazoan species, Hsap (Homo sapiens), Aque (*Amphimedon queenslandica*), and six fungal species Spom (*Schizosaccharomyces pombe*), Ncas (*Naumovozyma castelli*), Scer (*Saccharomyces cerevisiae*), Afum (*Aspergillus fumigatus*), Ncra (*Neurospora crassa*), and Smus (*Sphaerulina musiva*). The secondary structure based on fungal Get4s is highlighted above the text. (B) Left, the overall structure of Bag6-TRC35 heterodimer in cylinder representation with Bag6 in light pink and TRC35 in rainbow. The nuclear localization sequence is highlighted in sticks on Bag6. Right, a 90˚ in plane rotated 'bottom' view. The 16 helices of TRC35 are labeled from N to C terminus. (C)

Comparison of the C-terminal faces of TRC35 and Get4 that bind Bag6 and Get5, respectively. The arrows highlight the significant structural difference in the residues between α11 and α12. (D) Zoomed in view of the regions, defined as interface I and II. Hydrophobic residues that are involved in Bag6-TRC35 dimerization are shown as sticks.

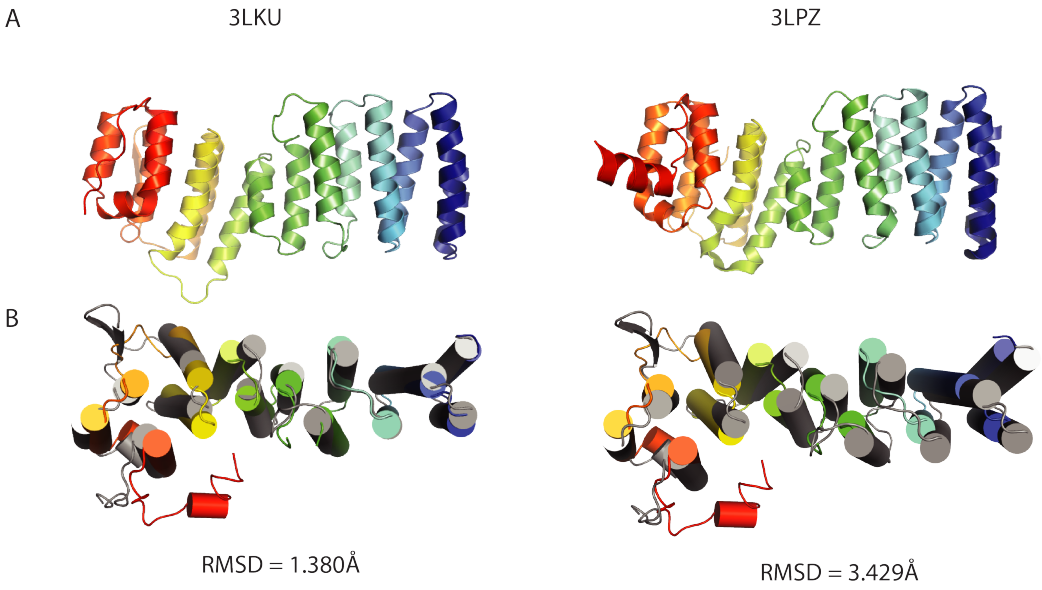


Figure 2.8. Comparison of TRC35 and Get4 structures

(A) Representative structures of fungal Get4 homologs. 3LKU is from *Saccharomyces cerevisiae* (Sc) and 3LPZ is from *Chaetomium thermophilum* (Ct). (B) Aligned human TRC35 (color ramped) and ScGet4 (3LKU, grey) using Pymol (Schrodinger, 2015) super for sequence-independent structural alignment. Left, structures aligned to the six N-terminal helices. Right, structures aligned to the seven C-terminal helices.

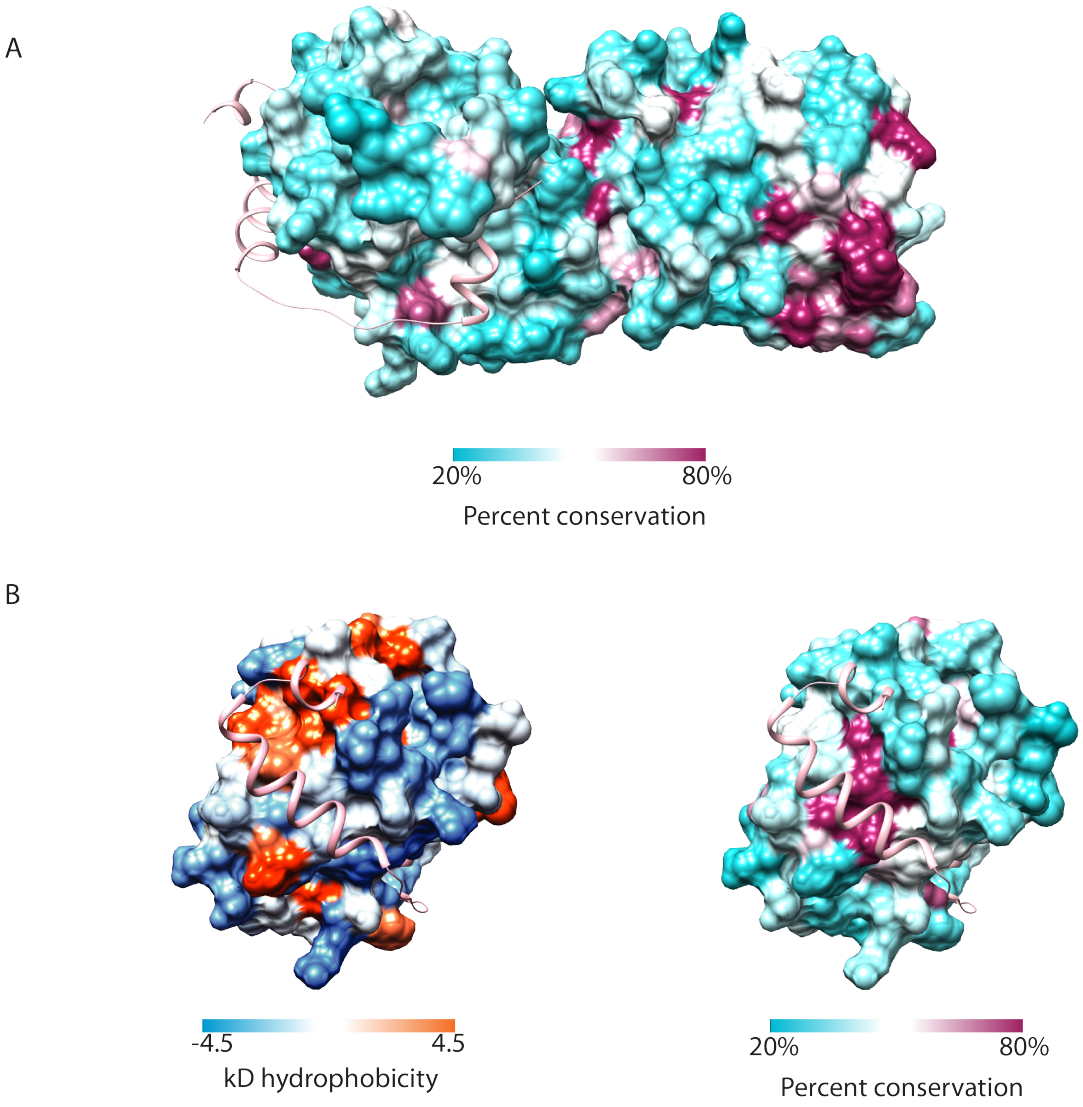


Figure 2.9 Analysis of the surface conservation and hydrophobicity of TRC35

(A) Accessible surface representation of TRC35 colored based on percent conservation as implemented in Chimera (Pettersen et al., 2004). Conservation based on a MAFFT (Katoh and Standley, 2013) alignment of TRC35/Get4 sequences from *Homo sapiens*, *Xenopus laevis*, *Danio rerio*, *Drosophila melanogaster*, *Nematostella vectensis*, *Monosiga brevicollis*, *Schizosaccharomyces pombe*, and *Saccharomyces cerevisiae*. Bag6 is in

ribbons representation in pink. (B) The Bag6-binding surface of TRC35 colored based on hydrophobicity (Kyle-Doolittle scale) and percent conservation as implemented in Chimera (Pettersen et al., 2004).

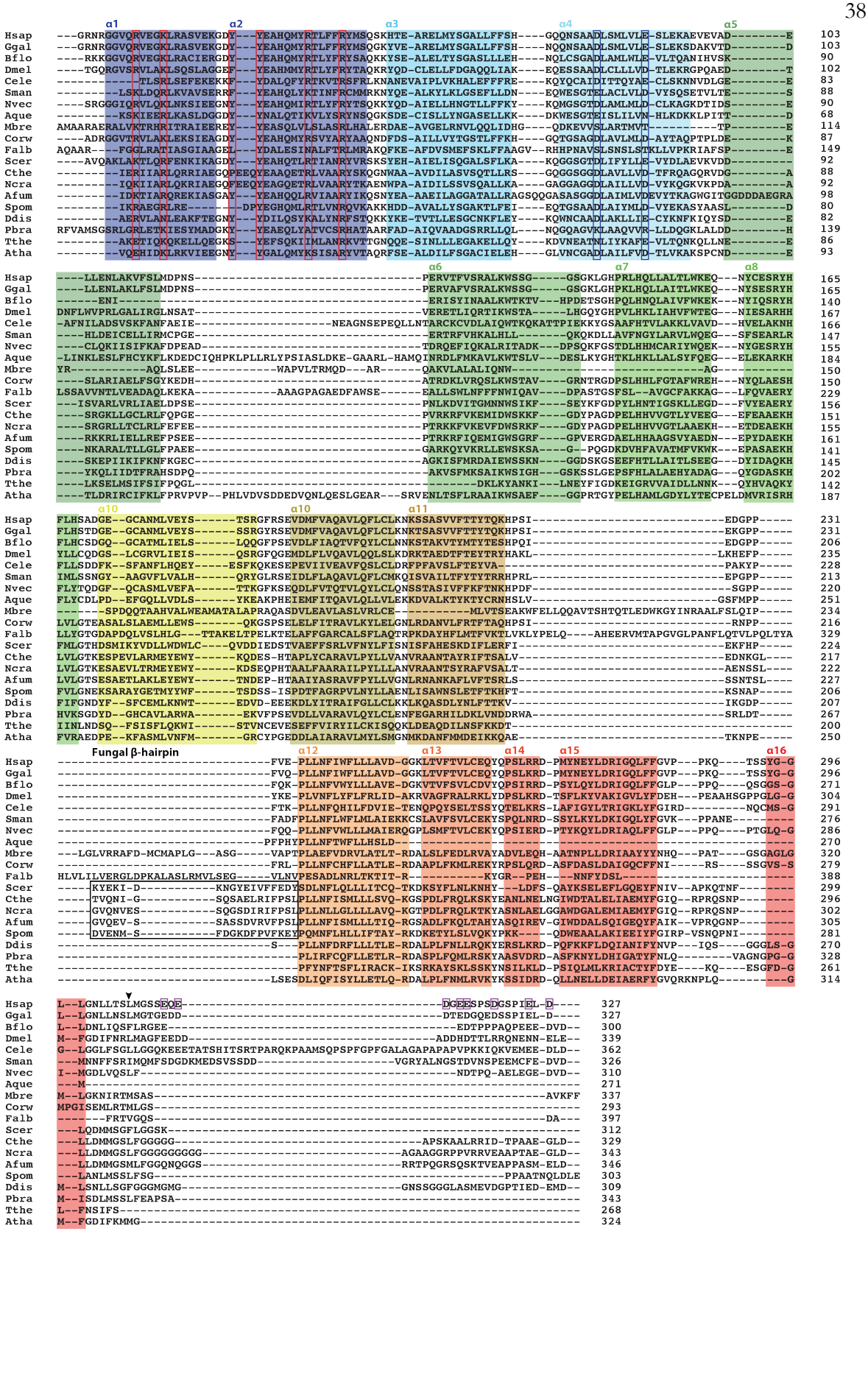

α1
α2
α3
α4
α5
α6
α7
α8
α10
α10
α11
Fungal β-hairpin
α12
α13
α14
α15
α16

Figure 2.10. Representative aligned sequences of eukaryotic TRC35/Get4

Species selected based on the eukaryotic phylogenetic tree by Eme et al (Eme et al., 2009). Sequences were aligned with MAFFT (Katoh and Standley, 2013). α-helices, based on the TRC35 structure, are highlighted in colors that correspond to the crystal structure on Figure 2.1B and numbered on top. Residues highlighted in red boxes were identified as critical to fungal Get4 binding Get3 (Gristick et al., 2014). Residues highlighted in blue boxes are critical to Get4/TRC35 regulating Get3/TRC40 (Gristick et al., 2014). The residues that comprise the fungal β-hairpin are highlighted in a black box. The arrow indicates the end of the crystallization construct. Purple boxes highlight glutamate and aspartate residues in the C-terminus of TRC35.

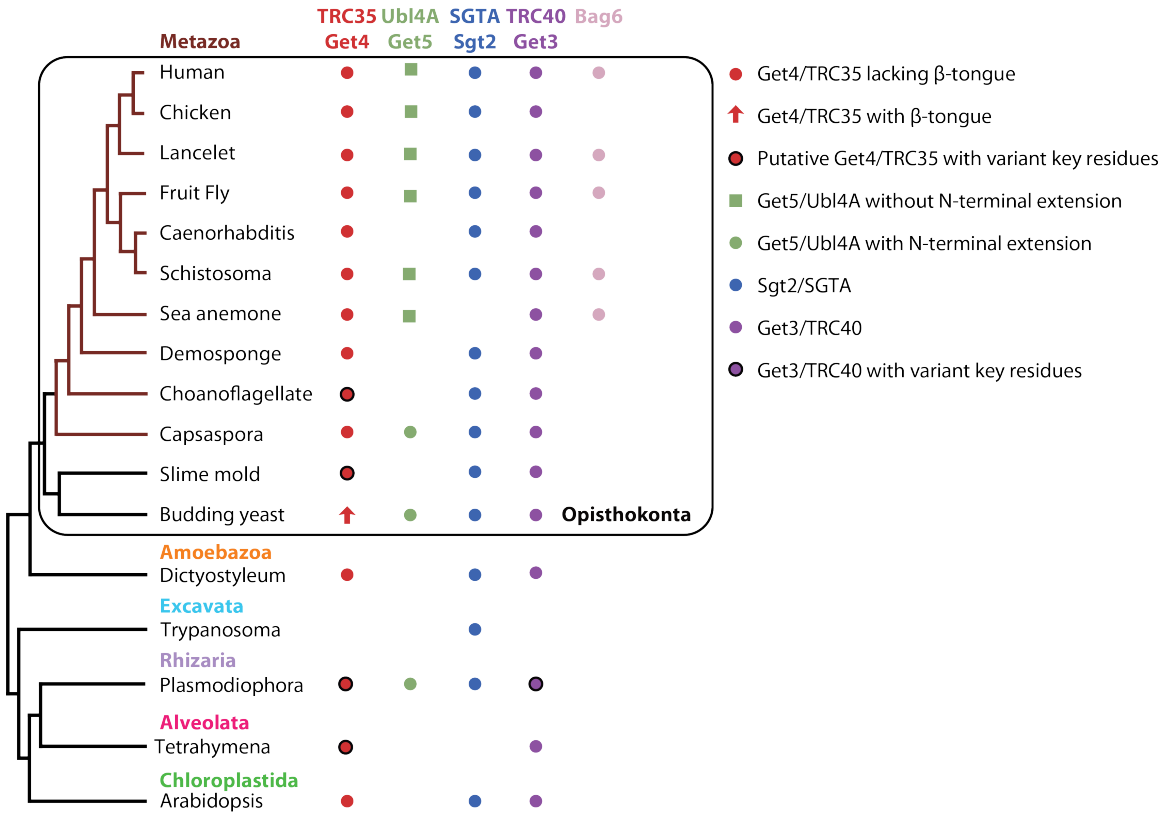


Figure 2.11 Survey of factors involved in the TRC pathway in eukaryotes

A condensed phylogenetic tree of representative eukaryotes was built based on the maximum likelihood phylogenetic tree of eukaryotes by Eme et al (Eme et al., 2014). The genome of each organism was searched for the presence of Get4/TRC35, Ubl4A/Get5, Sgt2/SGTA, Get3/TRC40, and Bag6 using MEME suite motif discovery tool (MEME) (Bailey et al., 2009) and motif scanning tool (MAST) (Bailey and Elkan, 1994) in addition to NCBI protein BLAST (Altschul et al., 1990). Proteins are color coded and sequence elements are highlighted. Black circles indicate homologs that are missing residues demonstrated to be critical in fungal TA-targeting studies.

Table 2.1 Bag6-Ubl4A crystallographic data and model refinement statistics

	Native	Iodide
Data collection		
Protein	Bag6-BAG/Ubl4A-C	Bag6-BAG/Ubl4A-C
Synchrotron	ALS	ALS
Beamline	8.2.1	8.2.1
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, b, c *(Å)*	47.2, 56.4, 75.5	47.4, 57.2, 75.9
α, β, γ (°)	90.0, 96.2, 90.0	90.0, 95.9, 90.0
Wavelength	1.0000	1.7000
Resolution (Å)	28.2 – 2.05	28.6 – 3.5
R_{merge} *(%)*[a]	3.6 (71.1)	14.2 (162.6)
$< I > / < \sigma I >$[a]	18.3 (2.0)	24.8 (3.0)
Completeness (%)[a]	98.7 (98.3)	99.9 (100.0)
No. of observations	78,344	258,426
No. of unique reflections[a]	26,539 (1,950)	11,271 (1,483)
Redundancy[a]	3.0 (2.7)	22.9 (23.3)
Refinement		
Resolution (Å)	28.3 – 2.0	
No. of reflections	26525	
No. of reflections test set	1338	
R_{work} / R_{free}	0.251 (0.335) / 0.280 (0.353)	
No. atoms (non-hydrogen)	3298	
Protein	3159	
Water	139	
B-*factors*		
Protein	43.60	
Water	41.80	
RMSD		
Bond lengths (Å)	0.019	
Bond angles (°)	1.30	
Ramachandran plot		
Favored (%)	95	
Additionally allowed (%)		
Outliers (%)	1.9	

[a]Highest-resolution shell is shown in parentheses

Table 2.2 Bag6-TRC35 crystallographic data and model refinement statistics

	Native	Rubidium
Data collection		
Protein	Bag6-TRC35	Bag6-TRC35
Synchrotron	SSRL	SSRL
Beamline	12-2	12-2
Space group	P21 21 21	P21 21 21
Cell dimensions		
a, b, c *(Å)*	41.7, 84.6, 102.6	42.0, 84.3, 102.5
α, β, γ (°)	90.0 90.0, 90.0	90.0, 90.0, 90.0
Wavelength	0.9200	0.8154
Resolution (Å)	39.1 – 1.80 (1.92 – 1.80)	50.0 – 1.99 (2.11 – 1.99)
R_{meas} *(%)*[a]	9.3 (64.3)	16.9 (118.8)
$I / \sigma(I)$	16.9 (3.8)	14.9 (2.1)
$CC_{1/2}$ *(%)*	99.9 (95)	99.9 (88.0)
Completeness (%)[a]	98.7 (92.7)	99.2 (95.6)
No. of observations	384,047 (60,190)	655,714 (96,962)
No. of unique reflections[a]	40,776 (6,194)	47,793 (7,417)
Redundancy[a]	9.4 (9.7)	13.7 (13.1)
Refinement		
Resolution (Å)	39.1 – 1.80	
No. of reflections	34,334	
No. of reflections test set	1,998	
R_{work} / R_{free}	16.0 / 20.0	
No. atoms (non-hydrogen)	2725	
Protein	2561	
Water	152	
Ligand/Ions	12	
B-*factors*		
Protein	31.7	
Water	39.0	
RMSD		
Bond lengths (Å)	0.010	
Bond angles (°)	1.16	
Ramachandran plot		
Favored (%)	98	
Additionally allowed (%)	2.0	
Outliers (%)	0.0	

[Ta]Highest-resolution shell is shown in parentheses

Table 2.3 Table of genomes surveyed for TRC pathway proteins in figure 2.11

	TRC35	Ubl4A	SGTA	TRC40	Bag6
Homo sapiens	NP_057033.2	NP_055050.1	NP_003012.1	NP_004308.2	NP_004630.3
Gallus gallus	XP_015149638.1	XP_015129423.1	NP_001026550.2	XP_015158190.1	
Branchiostoma floridae	XP_002607071.1	XP_002594764.1	XP_002591391.1	XP_002605952.1	XP_002594881.1
Drosophila melanogaster	NP_649464.1	NP_650680.1	NP_609842.1	NP_610296.2	NP_648132.2
Caenorhabditis elegans	NP_498034.1		NP_494893.1	NP_498965.1	
Schistosoma mansoni	XP_018645653.1	XP_018645468.1	XP_018654909.1	XP_018648676.1	XP_018648249.1
Nematostella vectensis	XP_001640454.1	XP_001638929.1		XP_001638254.1	XP_001635433.1
Amphimedon queenslandica	XP_011404841.1		XP_003387439.1	XP_003385895.1	
Monosiga brevicollis	XP_001750773.1		jgi Monbr1 37257	XP_001743630.1	
Capsaspora orwazaki	XP_004342754.1	XP_004347862.1	XP_004343568.1	XP_004363354.1	
Fonticula alba	KCV72828.1		XP_009498174.1	XP_009492379.1	
Saccharomyces cerevisiae	NP_014807.3	NP_014807.3	NP_014649.1	NP_010183.1	
Dictyostelium discoideum	XP_640617.1	XP_642815.1	XP_641391.1	XP_629069.1	
Trypanosoma brucei			XP_845518.1		
Plasmodiophora brassicae	CEP00951.1	CEP02084.1	CEP01962.1	CEP03329.1	
Tetrahymena thermophila	XP_012656464.1			XP_001033080.2	
Arabidopsis thaliana	NP_201127.2		NP_192572.2	NP_563640.1	

Materials and Methods

Yeast two-hybrid

The PJ69-4α strain was obtained from the Yeast Resource Center at the University of Washington. Bag6 fragments, A (1-225), B (226-399), C (400-659), D (660-950), E (951-1126), Bag6E_N (951-1011), Bag6E_{NLS} (1012-1054), and Bag6E_{BAG} (1055-1126) were cloned into pGAD-C1 vector. TRC35, TRC35-N (1-157), TRC35-C (158-327), and Ubl4A were cloned into pGBDU-C1 vector. Alanine mutants were made using site-directed mutagenesis (Agilent Technologies). pGAD-C1 and pGBDU-C1 containing genes of interest were co-transformed into PJ69-4α using previously described methods (Gietz and Schiestl, 2007) andf then plated on SC-Ura-Leu and incubated at 30°C. The double transformants were then inoculated into 5 mL SC-Ura-Leu and grown over night in a shaking incubator at 200 rpm, 30 °C. From the cultures 2×10^7 cells were re-inoculated into 5 mL total SC-Ura-Leu and grown in a shaker incubator for 6 hours. Cells were harvested by spinning at 3000 ×g at room temperature, washed twice with 5 mL sterile water. 1×10^7 cells were resuspended in 40 µL of sterile water. 4 µL of this resuspended sample were spot plated onto SC-Ura-Leu-Ade and grown for 72 hours at 30°C.

Cloning, expression and purification

For crystallization, 54 residues from Bag6 (Q1054 to D1107) and the C-terminal 53 residues from Ubl4A (P93 to K145) were co-expressed with pET33b plasmid with the N-terminal 6x histidine tag on Ubl4A-C. To facilitate *t*obacco *e*tch *v*irus (TEV) cleavage of the histidine tag, three flexible residues, glycine, alanine, and serine, were inserted between the TEV cut site and P93 of Ubl4A using site-directed mutagenesis. The proteins

were expressed in *E. coli* NiCo21(DE3) (New England Biolabs) for 3 hours at 37°C after induction with 300 μM isopropyl-β-D-thiogalactopyranoside (Affymetrix). Cells were lysed using an S-4000 sonicator (Misonix) in 50 mM NaH_2PO_4, 200 mM NaCl, 20 mM imidazole supplemented with benzamidine, PMSF and 0.5% triton X-100. The complex was purified by nickel-nitrilotriacetic acid-agarose (Ni-NTA) affinity chromatography (Qiagen) then incubated overnight at room temperature with TEV protease in 20 mM NaH_2PO_4, 100 mM NaCl, 20 mM imidazole, and 10 mM β-mercaptoethanol followed by size-exclusion chromatography on Superdex 75 gel-filtration column in 10 mM Tris, 50mM NaCl, pH 8.0 and concentrated to ~10 mg/ml. All truncations of Bag6 were purified using the protocol described above. C-terminal 53 residues from L1055 to D1107 of Bag6 were cloned for Bag6-BAG. Bag6-Ubl4A construct contained the C-terminal 73 residues and full-length Ubl4A. C-terminal 175 residues from V951 to P1126 were cloned in the Bag6 E fragment vector. Bag6-C81 contained the C-terminal 81 residues from K1046 to P1126.

All constructs were derived from human cDNA, specifically the major Bag6 isoform b (NP_542433.1). For crystallization, both Bag6 (1000 to 1054) and TRC35 (23 to 305) were inserted into the multiple cloning site of the pACYCDuet plasmid (Novagen) with an N-terminal hexahistidine tag on Bag6(1000-1054). The plasmid was transformed into *E. coli* NiCo21(DE3) (New England Biolabs). The plate was scraped to inoculate 12 × 2L baffled flasks containing 2xYT media then grown by shaking at 37°C in a shaking incubator (Multitron Standard Infors HT) at 250 rpm. Cell growth was monitored to an OD_{600} = 0.6 then protein expression was induced for 3 hours by the addition of 500 μM

isopropyl-β-D-thiogalactopyranoside (IPTG) (Affymetrix). Cells were harvested by centrifugation in a TLA 8.1 rotor at 4000 ×g for 20 minutes. The pellet was resuspended in 50 mM Mops pH 7.2, 300 mM K•glutamate, 5 mM MgOAc, 20 mM imidazole, 5 mM β-mercaptoethanol (1g cell / 10 mL lysis buffer) and supplemented with 0.1 mM PMSF and 1 mM benzamidine. Cells were lysed using an M-110L microfluidizer (Microfluidics) by two passes at approximately 17,500 psi. The lysate was clarified by centrifugation at 235,000 ×g in a Beckman Ti45 rotor for 30 minutes at 4°C. The clarified lysate was incubated for 1 hour with 3 mL of a 50% (vol/vol) slurry of nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen) by rocking. The mix was poured into a gravity column then washed with 100 mL lysis buffer. The protein was eluted with 12 mL elution buffer (20 mM Mops (pH 7.2), 150 mM K•glutamate, 300 mM Imidazole, 5 mM β-mercaptoethanol). The eluent was mixed with ~0.5 mg of TEV protease in snakeskin dialysis tubing with 10 kDa molecular weight cutoff (Thermo Fisher Scientific) and dialyzed overnight at room temperature in (20 mM Mops (pH 7.2), 100 mM K•glutamate, and 10 mM β-mercaptoethanol). Precipitate was removed by centrifugation with Beckman SX4750A rotor at 3000 ×g for 5 minutes at 4 °C and filtered with a 0.22 μm syringe filter. The sample was concentrated and loaded onto a 5 mL UnoQ ion-exchange column (Biorad) then eluted with a 60 mL 20 mM Mops (pH 7.2), 50 - 500 mM K•glutamate gradient, 5 mM β-mercaptoethanol. The fractions containing the protein (~5-23mL) were dialyzed in snakeskin dialysis bag in 20 mM Mops (pH 7.2), 50 mM K•glutamate, 5 mM β-mercaptoethanol for 2 hours at 4 °C. The sample was concentrated to 2 mL, filtered with a 0.22 μm syringe filter, and then further

purified by size-exclusion chromatography over a 120 mL Superdex 75 column (GE Healthcare) (20 mM Mops (pH 7.2), 50 mM K•glutamate, 5 mM β-mercaptoethanol). The fractions containing the heterodimer (61 mL – 72 mL) and monomeric Bag6 (77 mL – 82 mL) were pooled and concentrated to ~10 mg/mL using centrifugal filter units with 10 kDa molecular weight cutoff (Merck Millipore).

Circular dichroism

Circular dichroism spectra were obtained using an Aviv 62A DS circular dichroism spectrometer. The ellipticity of 10 μM of Bag6-BAG, Ubl4A-C, or Bag6-Ubl4A suspended in 10 mM Tris, 50 mM NaCl, pH 8.0 was measured.

Crystallization

For the Bag6-Ubl4A dimer, crystallization screening was performed using the sitting drop vapor-diffusion method with commercially available screens (Hampton) and was set up by a Mosquito robot (TTP Labtech) and incubated at room temperature. The heterodimer crystallized after 4 days as rectangular prisms in 20% PEG-3350, 0.05 M Hepes, 1% tryptone (Hampton). They were soaked in 20% glycerol for 15 seconds and cryopreserved in liquid nitrogen. Iodide derivatives were generated by soaking crystals in 40 μl mother liquor (20% PEG-3350, 0.05 M Hepes, 1% tryptone (Hampton)), 10 μl ethylene glycol, and 3 μl 6 M sodium iodide for 2 to 10 seconds prior to cryopreservation.

For the Bag6-TR35 dimer, crystallization screening was performed using the sitting drop vapor-diffusion method with the commercially available PEG Ion screen (Hampton

Research) and MRC 2 Well Crystallization plate (Hampton Research). Initial screening was performed by a Mosquito robot (TTP Labtech) with 100nL:100nL drops (protein:well solution) and 50 μL wells. No refinement was necessary. Crystals grew to full-size after 5 days as rectangular prisms in 20% PEG-3350 (wt/vol), 0.2 M DL-malic acid (pH 7.0). For cryo-protection, crystals were transferred into 100uL of well solution with 20% (vol/vol) glycerol for 15 seconds and then frozen in liquid nitrogen. Rubidium derivatives were generated by transferring crystals into 40 μl well solution plus 350 mM rubidium iodide for 2 to 10 seconds prior to cryopreservation with glycerol.

Data collection, structure solution and refinement

For the Bag6-Ubl4A structure, x-ray diffraction data were collected on beam line 8.2.1 at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory. A complete dataset was collected from a single crystal to 2.1 Å resolution. Data were integrated, scaled, and merged using MOSFLM (Battye et al., 2011) and SCALA (Collaborative Computational Project, 1994; Winn et al., 2011). Phases were determined by single-wavelength anomalous dispersion data merged from three iodide derivative crystals, which diffracted to 3.5 Å. The model was refined using COOT (Emsley et al., 2010) and PHENIX (Adams et al., 2010) . Statistics are provided in Table 1.

For the Bag6-TRC35 dimer, x-ray diffraction data were collected on beam line BL12-2 at the Stanford Synchrotron Radiation Laboratory (SSRL). A complete native dataset was collected from a single crystal to 1.8 Å resolution and a single rubidium derivative crystal dataset was collected from a single crystal to 2.0 Å with a Dectris Pilatus 6M detector. Data were integrated, scaled, and merged using XDS (Kabsch, 2010) and SCALA

(Collaborative Computational Project, 1994; Winn et al., 2011). Phases were determined by molecular replacement single-wavelength anomalous dispersion (MRSAD) using SHARP (Bricogne et al., 2003) and PHASER-MR (McCoy et al., 2007) on PHENIX (Adams et al., 2010). The initial structure was built by PHASER as implemented by PHENIX (Adams et al., 2010). The model was further built and refined against the native dataset over several rounds using COOT (Emsley et al., 2010) and PHENIX (Adams et al., 2010). Statistics are provided in Table 2.

Phylogenetic Tree of GET/TRC Components

The phylogenetic tree was modified from the maximum likelihood phylogenetic tree of eukaryotes by Eme et al (Eme et al., 2014). The MEME suite (Bailey et al., 2009) was used to determine the presence or absence of the GET/TRC components in the genomes presented in the tree. The MEME motif discovery tool was used to make motifs for Ubl4A, Get5, Get4/TRC35, and Sgt2, SGTA, and Bag6. The motifs were then used to search the genomes using the motif-scanning tool MAST (Bailey and Elkan, 1994). Identified proteins were confirmed using BLAST (Altschul et al., 1990) and the corresponding reference ID is provided on table 3.

Acknowledgements

We thank Daniel Lin and Jens Kaiser for help with data processing. We thank Dr. Meera Rao and Dr. Michael Rome for critical reading of the manuscript. We thank members of the laboratory for support and useful discussions. We are grateful to Gordon and Betty Moore for support of the Molecular Observatory at Caltech. We thank the staff at the Advanced Light Source and Stanford Synchrotron Radiation Laboratory for assistance with synchrotron data collection. W.M.C. is supported by NIH grant R01GM097572.

C h a p t e r 3

BIOCHEMICAL CHARACTERIZATION OF THE BAG DOMAIN AND THE MINIMAL TA TARGETING MODULE

Parts of this chapter were first published in

Mock, J.-Y., Chartron, J.W., Zaslaver, M., Xu, Y., Ye, Y., and Clemons, W.M. Jr. (2015) *Proc Natl Acad Sci USA*. 112(1): 106-11 doi: 10.1073/pnas.1402745112

Mock, J.-Y., Xu, U., Ye, Y., and Clemons, W.M. Jr. (2017) Structural basis for regulation of nucleocytoplasmic distribution of Bag6. *Submitted*.

Jee-Young Mock carried out yeast 2-hybrid analysis and biochemical experiments with purified proteins. Jee-Young Mock participated in designing immunoprecipitation experiments.

Abstract

The metazoan protein BCL2-associated athanogene cochaperone 6 (Bag6) forms a hetero-trimeric complex with ubiquitin-like 4A (Ubl4A) and transmembrane domain recognition complex 35 (TRC35). This Bag6 complex is involved in tail-anchored protein targeting and various protein quality control pathways in the cytosol as well as regulating acetylation and histone methylation in the nucleus. Crystal structure of the Bag6 BAG domain revealed that it does not fold into a canonical BAG domain fold. In this study, we biochemically demonstrate that Bag6-BAG domain does not behave like a canonical BAG domain, establishing it as a "mock" BAG domain. Furthermore, we show that C-terminal 125 residues of Bag6, which can form a stable complex with TRC35 and Ubl4A, are sufficient for TA targeting.

Introduction

BCL2-associated athanogene cochaperone 6 (Bag6, also known as "BAT3" or "Scythe") is a metazoan protein that is involved in membrane protein targeting and quality control in the cytosol. Biochemical (Mariappan et al., 2011; Stefanovic and Hegde, 2007; Wang et al., 2011b) and structural studies (Mock et al., 2015) have demonstrated that Bag6 forms a stable heterotrimer with transmembrane domain recognition complex 35 (TRC35) and ubiquitin-like 4A (Ubl4A).

Crystal structures of the Bag6-Ubl4A and Bag6-TRC35 complexes revealed that the Bag6 BAG domain, previously designated based sequence homology and biochemical characterization (Thress et al., 2001), does not resemble the canonical three helix bundle conformation of other BAG domains (Fig. 2.5). Furthermore, despite organizational changes from the heterotetrameric fungal complex to the heterotrimeric metazoan complex, all the functional domains required for TA targeting are structurally conserved. The fungal heterotetrameric Get4-5 sorting complex binds Sgt2 via the UBL domain of Get5 (Chartron et al., 2012b; Xu et al., 2012). The N-terminal domain of Get4 binds and regulates the ATPase activity of Get3 (Chartron et al., 2010; Gristick et al., 2014; Rome et al., 2013). Structurally, the UBL domain of Get5 and the N-terminal domain of Get4 have both been conserved in the metazoan homologues, Ubl4A and TRC35, respectively. This suggests that the minimal Bag6 complex structurally characterized in chapter 1 is functionally equivalent to the fungal Get4-5 complex.

This chapter seeks to functionally verify structural observations using biochemical methods. *In vitro* Hsc70-mediated refolding assay was used to confirm that the Bag6

BAG domain does not behave like a canonical BAG domain, which binds and inhibits Hsc70 *in vitro*. Furthermore, the minimal Bag6-TRC35-Ubl4A complex identified and structurally characterized in chapter 2 facilitates TA protein transfer from SGTA to TRC40 *in vitro*. These findings establish the minimal Bag6 complex identified in this study as a TA targeting module.

Results

Functional characterization of the Bag6-BAG domain

Previous results have shown that the ability of Bag6 to inhibit Hsc70 refolding of substrates *in vitro* is dependent on the presence of the C-terminal 81 residues (Bag6-C81), which include the BAG domain (Thress et al., 2001). This was suggested to be equivalent to results for the Bag1 BAG domain, a demonstrated nucleotide exchange factor for Hsc70 (Sondermann et al., 2001). If true, Bag6-BAG should inhibit Hsc70-mediated protein folding. To assay this, denatured β-galactosidase was folded in vitro in the presence of human Hsc70 and Hdj1, as done previously (Freeman and Morimoto, 1996; Thress et al., 2001). Folding was measured as a percent of β-galactosidase activity recovered after the folding reaction was quenched. With both Hsc70 and Hdj1 present a maximal refolded activity of ~35% was recovered after 180 min (Fig. 3.1, brown line) while no refolding was seen when only BSA was added (black line). The addition of human Bag1-BAG to Hsc70 and Hdj1 completely inhibited the ability of Hsc70 to fold the protein (purple line) consistent with previous results (Takayama et al., 1997). Conversely, the addition of Bag6-BAG had no effect on refolding by Hsc70 (orange solid line).

The inconsistency with previous results might be explained by co-purification of endogenous Bag6 binding partners. In the earlier study, an affinity-tagged full-length human Bag6 and C-terminal 81 residues of Bag6 (Bag6-C81) were expressed and purified from insect cells over a single affinity resin (Thress et al., 2001). As the proteins are well conserved, it is reasonable to assume that endogenous insect Ubl4A could be a

contaminant and may contribute to the inhibition. The introduction of the Bag6-BAG/Ubl4A complex to the reaction had no effect on folding (blue line). SGTA, a co-chaperone, has recently been shown to form a specific complex with Ubl4A (Xu et al., 2013) and may also have been in the endogenously purified sample. SGTA had no significant effect on activity (Fig. 3.2) with or without the other factors (Fig. 3.1A, dashed lines). Bag6-C81, which was required for inhibition in the previous study, is slightly larger than the BAG domain defined here (Fig. 2.1A). However, using this larger fragment also had no effect on folding (Fig. 3.2). Together, these results suggest that although the C-terminal 81 residues of Bag6 are required for its chaperoning activity, it does not act as a bona fide BAG domain to cooperate with Hsc70.

This conclusion was further supported by binding assays. A 6xHis-tagged Hsc70-NBD (*N*ucleotide *B*inding *D*omain), the expected binding site for Bag proteins, was incubated with purified Bag6-C81, Bag6-BAG, Bag6-BAG/Ubl4A, or Bag1-BAG. After incubation, Hsc70-NBD was captured on Ni-NTA beads along with any associated protein. As expected, Hsc70-NBD could capture the Bag1-BAG (Fig. 3.3). On the other hand, Hsc70-NBD was unable to capture any of the Bag6 C-terminal fragments at levels above background. Thus, one would conclude that Bag6-BAG is unlike canonical BAG domains in its ability to interact with Hsc70 in vitro.

While the evidence makes it clear that the Bag6 BAG domain does not interact with Hsc70 in isolation, the role of other unknown factors could not be ruled out. For instance, SGTA could mediate the interaction between Bag6-BAG and Hsp70 as it binds Hsp70

via its tetratricopeptide repeat (TPR) domain and Ubl4A via its N-terminus (Chartron et al., 2012b). To address the possibility of accessory factors, tagged variants of BAG domains were incubated with 293T whole cell lysate (Fig. 3.1B). The positive control Bag1-BAG again could capture a significant amount of Hsc70. A complex with Bag6-BAG and Ubl4A, either the crystallized fragment or the full-length, was unable to capture Hsc70.

These results contrast with previous experimental results in which Bag6 inhibition of Hsc70 was dependent on the presence of the BAG domain (Thress et al., 2001). This could not rule out the possibility of binding Hsc70 in the context of the full-length Bag6. To test this, 293T cells were transfected with either FLAG-Bag6 or FLAG-Bag6ΔC81 and captured with anti-FLAG resin. Proteins bound to the beads were blotted with both Bag6 (green) and Hsc70 (red) antibodies. Full-length Bag6 captured a small amount of Hsc70 over background; however, the lack of the previously annotated BAG domain, Bag6ΔC81, had no effect on the amount of captured Hsc70 (Fig. 3.1C). Combined, these results suggest that Bag6-BAG, both structurally and biochemically, is not a true BAG domain.

Expression and purification of Bag6$_{min}$ complex

The next goal was to purify the hetero-trimeric complex. Both full-length Bag6 and TRC35 were recalcitrant to recombinant expression in *E. coli.* For TRC35, expression required removal of residues at the N- and C-terminus that are not conserved, TRC35(23-305). This truncated TRC35 behaved as wild-type by yeast 2-hybrid with Bag6E (Fig. 3.4A). Additionally, based on the interaction results, a minimal Bag6 fragment (Bag6$_{min}$)

was constructed, residues 1001 to 1126, that removed the N-terminal 50 residues of Bag6E. Co-expressing $Bag6_{min}$, TRC35(23-305) and Ubl4A resulted in a stable complex that could be purified (Fig. 3.4).

The $Bag6_{min}$ complex is an independent module that facilitates TA handoff

The fungal Get4-5 complex binds ATP bound Get3 and inhibits its ATPase activity priming Get3 for TA substrate capture from Sgt2 (Gristick et al., 2014; Rome et al., 2014). One would expect that the trimeric Bag6 complex, which contains the mammalian Get4-5 orthologues, regulates TRC40 in a similar manner. To test this, an in vitro assay was developed to probe the role of the $Bag6_{min}$ complex in TA handoff from SGTA to TRC40 using recombinantly purified proteins (scheme in Fig. 3.5A). When co-expressed, fungal Sgt2 binds GET dependent TA substrates and this complex can be purified (Wang et al., 2010). Here, histidine-tagged SGTA (hSGTA) was co-expressed in E. coli with the yeast TA protein Sbh1 that contained an N-terminal maltose-binding protein (MBP•Sbh1) resulting, after a two-step purification, in a stable hSGTA/MBP•Sbh1 complex (Fig. 3.4B). The final component, TRC40, was generated as a N-terminal glutathione-S-transferase tag (GST•TRC40) (Fig. 3.4B), as done previously (Vilardi et al., 2011; Yamamoto and Sakisaka, 2012). Transfer was initiated by incubation of hSGTA/MBP•Sbh1 with either GST•TRC40 alone or with the $Bag6_{min}$ complex (Fig. 3.5A). The samples were then precipitated with anti-GST resin, washed in two steps in the absence of ATP, and then probed after Western blotting with both MBP and GST antibodies.

GST•TRC40 was able to capture some TA from SGTA alone as seen before for the yeast system (Wang et al., 2010) (Fig. 3.6). The additional presence of the $Bag6_{min}$ complex resulted in an ATP-dependent increase in TA transfer to TRC40 (compare lanes 4 & 6).

This was not a result of a bridged capture of TRC40 pulling down TA still bound to SGTA. The interaction between SGTA and Ubl4A is predicted to have very fast off-rates (Chartron et al., 2012b) and SGTA would be rapidly removed during the wash steps. Moreover, capture of TA by TRC40 was insensitive to increasing salt concentration (Fig. 3.6) despite the SGTA/Ubl4A and TRC35/TRC40 interactions being dominated by electrostatics, the latter also requiring ATP (Chartron et al., 2012b; Gristick et al., 2014). The $Bag6_{min}$ fragment used in the study does not contain the Bag6 substrate-binding domain (Leznicki et al., 2013); therefore, the Bag6 complex can promote substrate hand-off from SGTA to TRC40 without Bag6 directly engaging the substrate.

SGTA binds the Bag6 complex via the UBL domain of Ubl4A; consequently, the Bag6 dependent hand-off should require this interaction (Chartron et al., 2012b; Xu et al., 2012). To address this, the mutants hSGTA(C38A) and Ubl4A(L43A) were generated where the equivalent mutations in yeast were previously shown to disrupt the homologous interaction (Chartron et al., 2012b). As expected, each resulted in a similar loss of substrate hand-off relative to wild type (Figs. 3.5B and 3.7C). This highlights the importance of this interaction for the bridging by Bag6 during TA transfer.

In yeast, Get4-5 binding regulates Get3 ATPase activity and TA targeting (Rome et al., 2013). Therefore, in addition to the bridging role of the Bag6 complex, it is critical to test if

this larger complex plays a regulatory role in TA targeting. The crystal structure of yeast Get4 bound to Get3 highlighted a regulatory interface separate from the binding interface (Gristick et al., 2014). When mutated, residues on either side of this regulatory interface (the charge swaps K69D on Get3 and D74K on Get4) each resulted in a loss of ATPase inhibition, reduction of TA insertion into microsomes, and a loss of fitness in vivo, despite maintaining a stable complex in vitro (Gristick et al., 2014). Combining the charge swap mutants restored the Get4 regulatory activity (Gristick et al., 2014). For TRC40, the corresponding regulatory mutation, K86D, resulted in a reduction of the Bag6 complex facilitated hand-off (Fig. 3.5C, compare lanes 4 & 8 and Fig. 3.7D). Similarly, the corresponding regulatory mutation in TRC35, D84K, resulted in a reduction of facilitated hand-off (compare lanes 4 & 5 and Fig. 3.7D). Excitingly, as seen for the yeast system (Gristick et al., 2014), the combination of these two charge swap mutants resulted in a rescue of the facilitated hand-off (lane 9 and Fig. 3.7D). These results highlight that this minimal Bag6 complex acts as an independent TA targeting module and performs a similar regulatory role to the fungal Get4-5 complex despite the different architectures.

Discussion

Our structural characterization of the mammalian TA sorting complex in chapter 1 established Bag6 as a scaffolding factor that forms a stable complex with TRC35 and Ubl4A. The structure of the putative BAG domain, however, did not resemble published structures of canonical BAG domains. Meanwhile, although the introduction of Bag6 alters the overall molecular architecture in mammals relative to that of yeast, all the functional elements required to bind and regulate relevant TA components, Sgt2/SGTA and Get3/TRC40, were conserved in the Bag6 complex. In this chapter, we biochemically confirm that Bag6 BAG domain is not a BAG domain. Furthermore, we define and characterize *in vitro* a minimal Bag6 TA targeting module that is sufficient for TA handoff.

The demonstrated inability of Bag6-BAG to influence Hsc70 activity is unsurprising if one considers sequence and structural comparisons with canonical BAG domains. This result brings into question the reported inhibition of Hsp70-mediated β-galactosidase refolding by Bag6 (Thress et al., 2001). A possible explanation is the holdase activity of Bag6. Bag6 prevents the aggregation of unfolded luciferase by forming a stable interaction with the exposed hydrophobic core, preventing Hsp70-mediated refolding (Wang et al., 2011b). Similar Bag6 holdase activity may have prevented refolding of β-galactosidase. Deletion of the 81 C-terminal residues may disrupt this holdase through an unknown mechanism, possibly by occluding the binding site on the truncated Bag6. The observed Hsp70 inhibition by Bag6, therefore, would be a result of sequestration of the unfolded substrate by Bag6 via its hydrophobic substrate-binding region, which also could be recognized by Hsp70, resulting in the interaction between Bag6 and Hsp70.

While full-length Scythe and Bag6 inhibit reaper and ricin-triggered apoptosis, excess ScytheC312 (Thress et al., 1998) or the C-terminal 131 residues (Wu et al., 2004) of Bag6 can induce apoptotic events. As the C-terminal fragment would be consistent with the $Bag6_{min}$ complex defined here, excess Bag6 C-terminus would disconnect the triaging and holdase/degradation roles of the complex. The apoptosis connection could then be linked to tail-anchor targeting. Overexpression of Bag6 in HeLa cells exposed to ricin, an apoptosis inducer, leads to an increase in endogenous Bcl-2 protein levels, whereas Bag6 knockdown causes down-regulation of Bcl-2 proteins (Wu et al., 2004). Several proteins that belong to the Bcl-2 family are tail-anchored, including Bcl-2, MCL1, BAX, and BOK, and reside both at the ER and the mitochondria (Echeverry et al., 2013; Strasser, 2005; Szegezdi et al., 2009). The Bag6 complex would then play an important role in regulating the localization and turnover of these Bcl-2 proteins, which could be altered by Bag6 cleavage.

The results presented in this chapter establish the C-terminal domain of Bag6 as an independent TA targeting module of Bag6. The N-terminal UBL domain of Bag6 connects the proteasome, where it interacts with RP non-ATPase 10c (Kikukawa et al., 2005), with components of quality control pathways, where it interacts with RNF126, gp78 and ubiquitin regulatory X domain-containing protein 8 (Rodrigo-Brenni et al., 2014; Wang et al., 2011b; Xu et al., 2013). Downstream of this connection, the proline-rich domain has been implicated as the holdase domain binding to exposed hydrophobic regions and polyubiquitylated defective ribosomal products (Leznicki et al., 2013; Minami et al., 2010; Wang et al., 2011b). Bag6 then acts as a scaffolding protein,

simultaneously binding ubiquitylation machinery, the proteasome, TA-targeting factors, and proteins to be triaged. Recent biochemical characterization of the triaging process revealed that TA substrate that is not handed off to TRC40 within ~3 cycles is rerouted to the degradation pathway by Bag6 (Shao et al., 2017). The molecular details of how its decision-making process relates to its other functions in apoptosis, gene regulation, and immunoregulation are important questions for future studies.

Figures

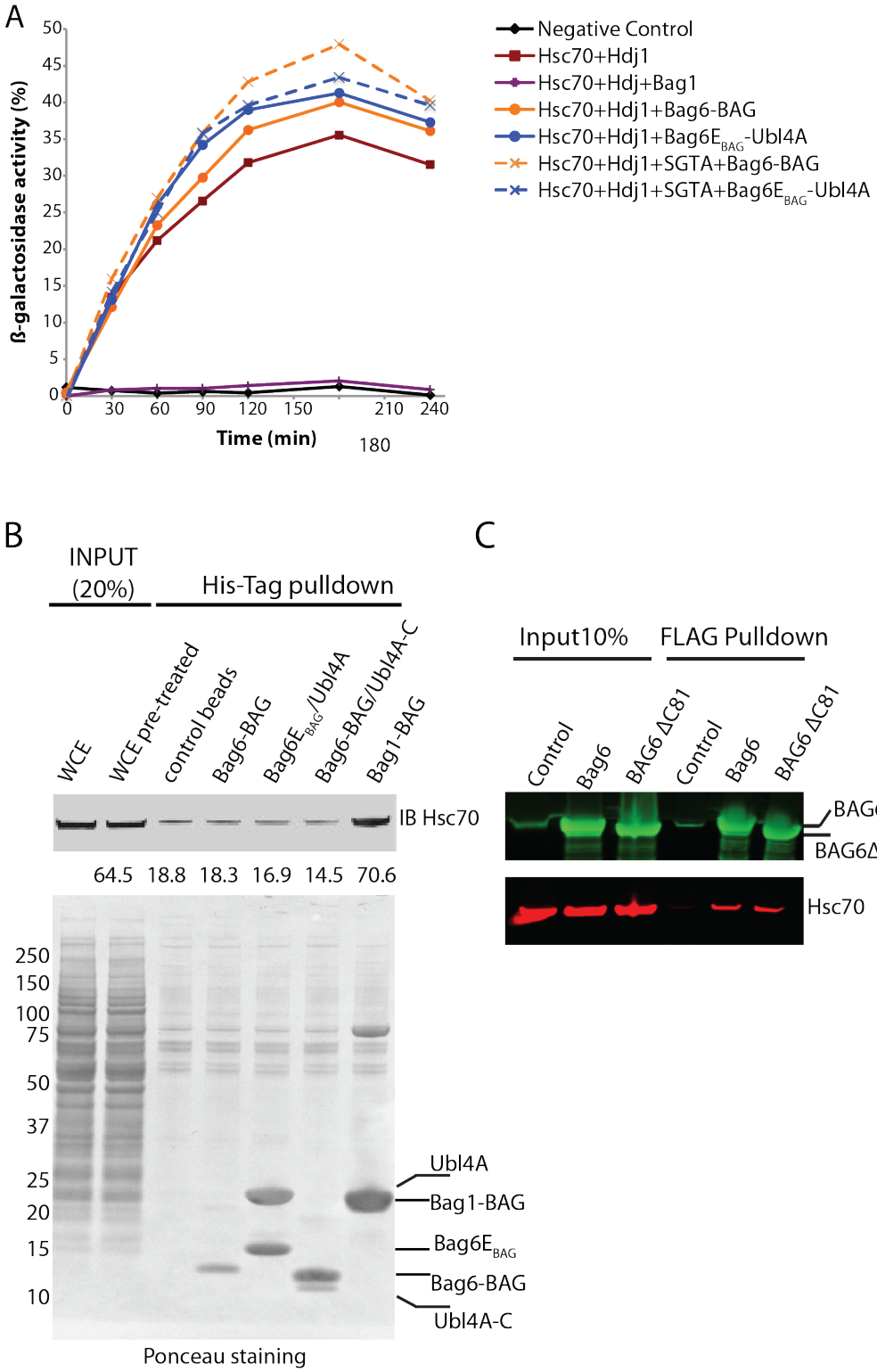

A
ß-galactosidase activity (%)
Time (min)
Negative Control
Hsc70+Hdj1
Hsc70+Hdj+Bag1
Hsc70+Hdj1+Bag6-BAG
Hsc70+Hdj1+Bag6E$_{BAG}$-Ubl4A
Hsc70+Hdj1+SGTA+Bag6-BAG
Hsc70+Hdj1+SGTA+Bag6E$_{BAG}$-Ubl4A
B
INPUT (20%)
His-Tag pulldown
WCE
WCE pre-treated
control beads
Bag6-BAG
Bag6E$_{BAG}$/Ubl4A
Bag6-BAG/Ubl4A-C
Bag1-BAG
IB Hsc70
64.5 18.8 18.3 16.9 14.5 70.6
Ubl4A
Bag1-BAG
Bag6E$_{BAG}$
Bag6-BAG
Ubl4A-C
Ponceau staining
C
Input10%
FLAG Pulldown
Control
Bag6
BAG6 ΔC81
BAG6
BAG6ΔC81
Hsc70

Figure 3.1. Bag6-BAG is not a canonical BAG domain

(A) Hsc70-mediated refolding of β-galactosidase in the presence of Bag1-BAG (purple), Bag6-BAG (orange solid), Bag6E_{BAG}/Ubl4A (blue solid), SGTA & Bag6-BAG (orange dashed), SGTA & Bag6E_{BAG}/Ubl4A (blue dashed) or BSA (black) as a negative control. (B) Affinity tagged Bag6-BAG, Bag6E_{BAG}/Ubl4A, Bag6-BAG/Ubl4A-C or Bag1-BAG was loaded onto cobalt resin beads and incubated with 293T whole cell lysate. Eluted samples were immunoblotted with Hsc70 antibody (top panel) and Ponceau stained (bottom panel). (C) FLAG-Bag6 or FLAG-Bag6ΔC81 was overexpressed in 293T cells and anti-FLAG resin was used to capture them with bound factors. Bag6 antibody on the blot is detected in green, and Hsp70 antibody in red.

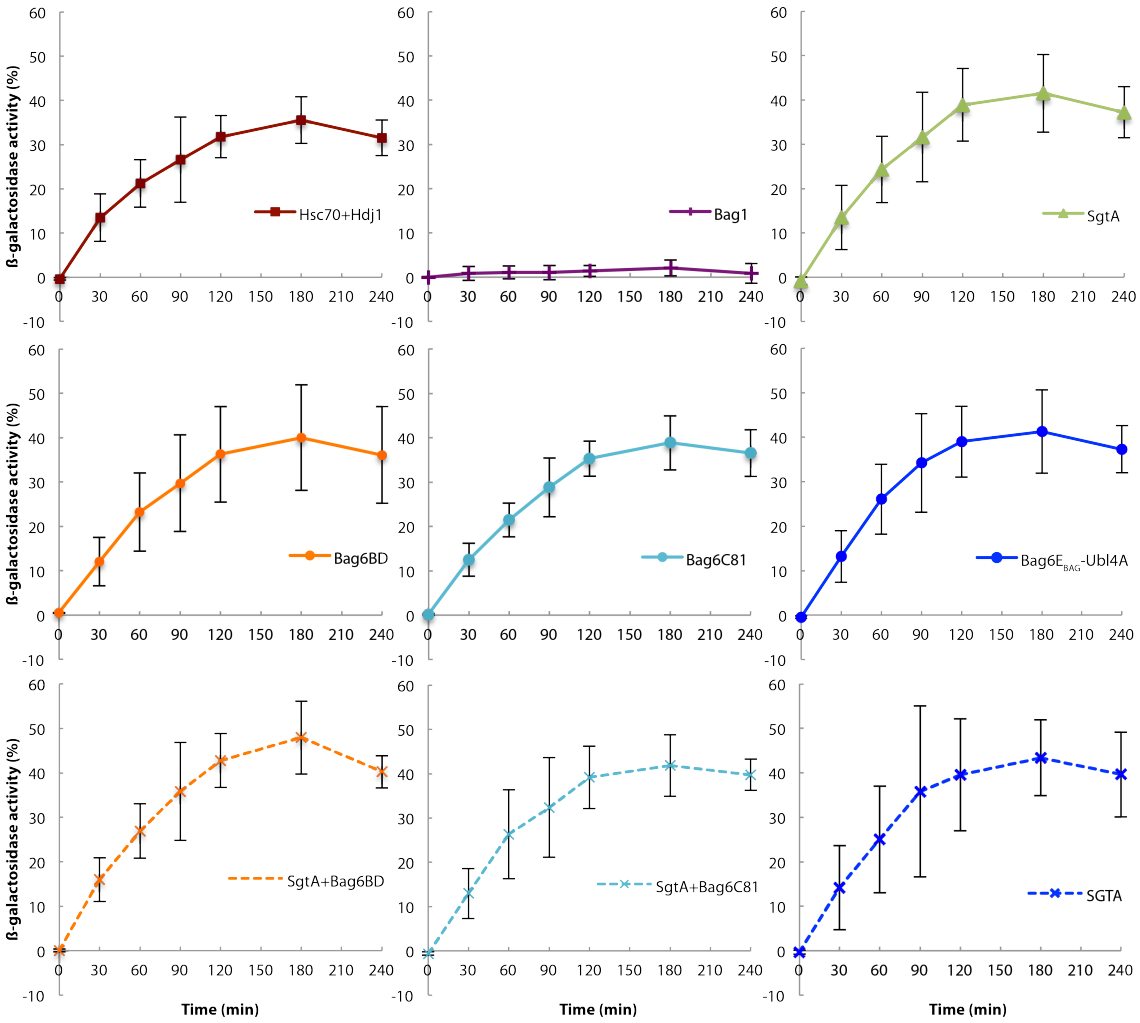


Figure 3.2. Individual results for various *in vitro* refolding assays

β-Galactosidase refolding assays in the presence of Hsc70, Hdj1 and/or other factors (labeled). Colors are based on Figure 3.1 except for assays containing Bag6-C81 that are in cyan. Error bars are from three independent experiments.

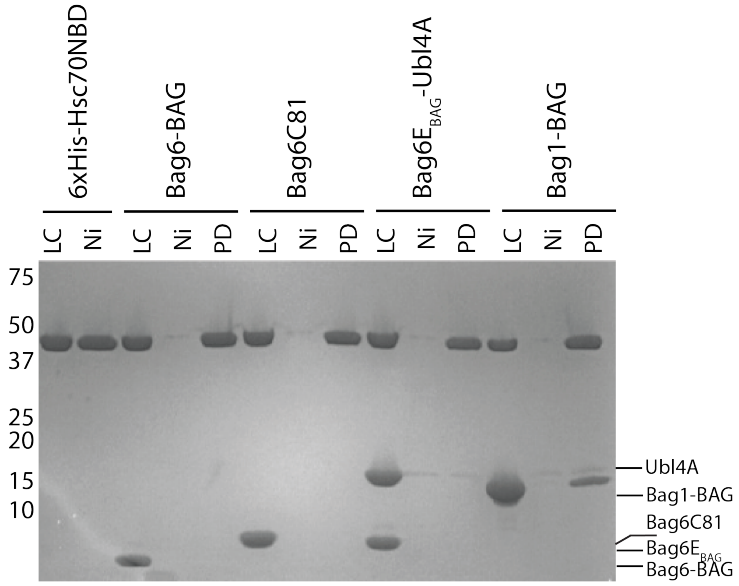


Figure 3.3. Bag6-BAG does not bind Hsc70 nucleotide binding domain (NBD)

In vitro capture by 6xHis-Hsc70-NBD of Bag6-BAG, Bag6-C81, Bag6E$_{BAG}$/Ubl4A, or Bag1-BAG. Protein was pulled down (PD) with Ni-NTA after incubation with 6xHis-Hsc70-NBD. Four percent of total loaded protein is shown as a loading control (LC), and each protein was incubated alone with Ni-NTA (Ni) to assess background binding.

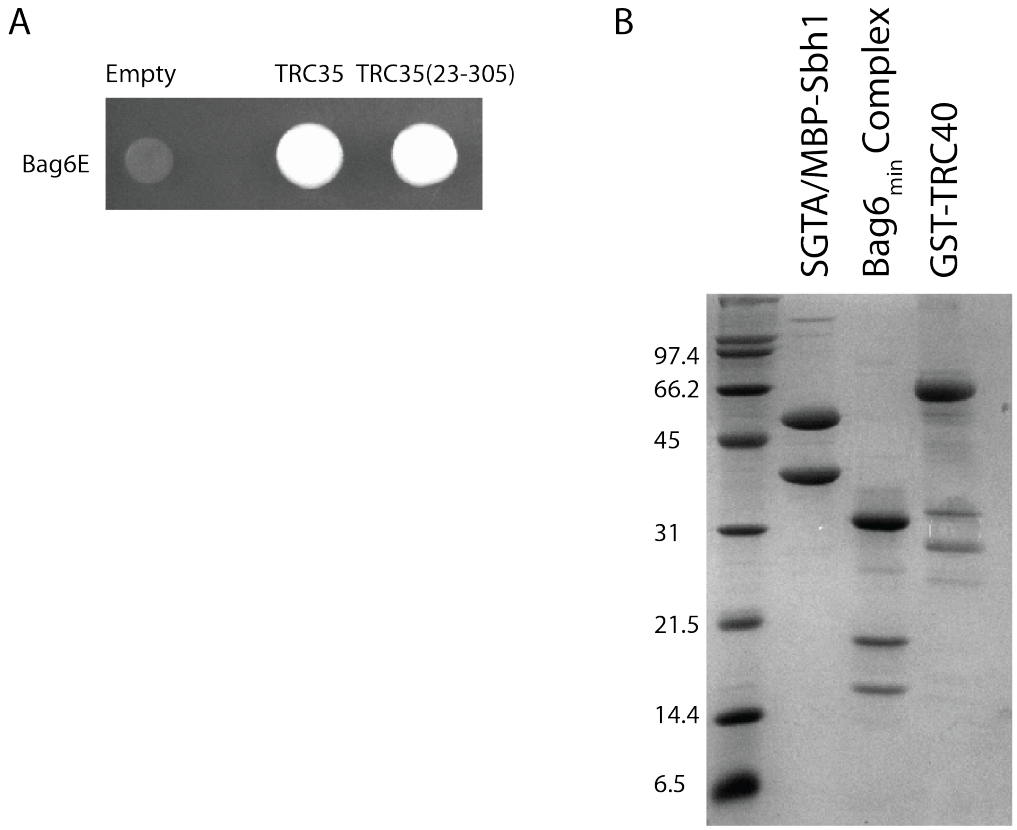
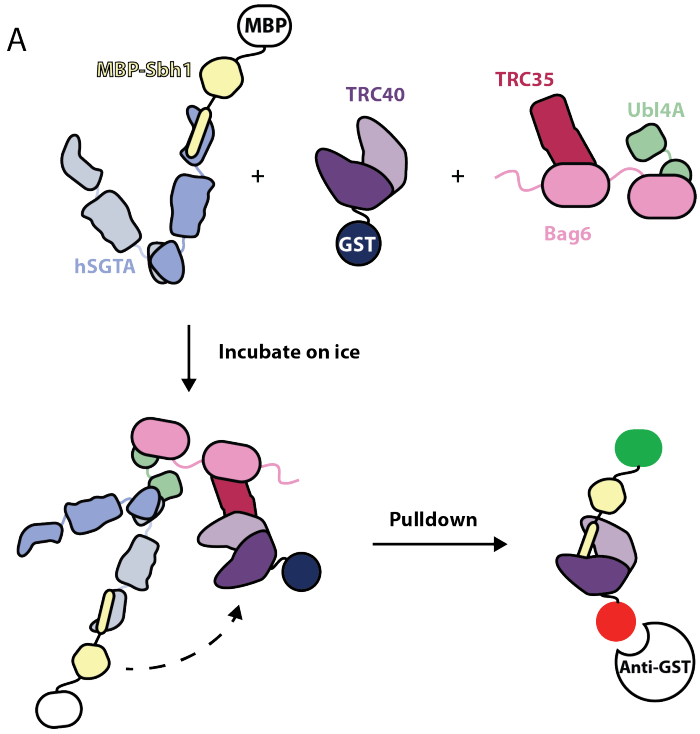
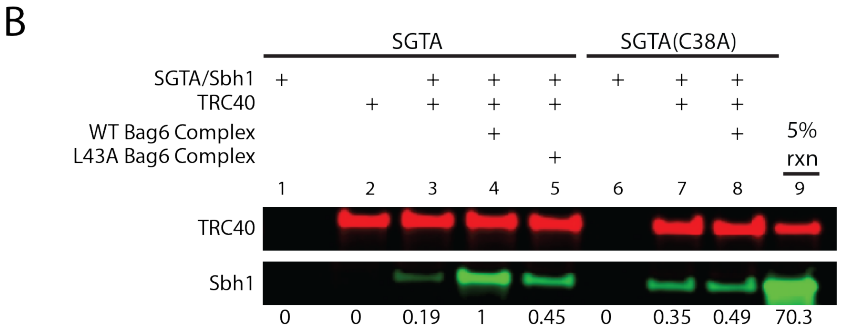


Figure 3.4. Purification of recombinant proteins used in TA transfer assay

(A) Two-hybrid using full-length TRC35 and TRC35(23-305) with Bag6E fragment. Both TRC35 constructs display strong two-hybrid interactions. (B) Representative Coomassie stained 12% SDS-PAGE gel of purified hSGTA/MBP•Sbh1, $Bag6_{min}$ Complex, and GST•TRC40.


A
MBP
MBP-Sbh1
hSGTA
TRC40
GST
TRC35
Ubl4A
Bag6
Incubate on ice
Pulldown
Anti-GST


B
SGTA
SGTA(C38A)
SGTA/Sbh1
TRC40
WT Bag6 Complex
L43A Bag6 Complex
5% rxn
1 2 3 4 5 6 7 8 9
TRC40
Sbh1
0 0 0.19 1 0.45 0 0.35 0.49 70.3

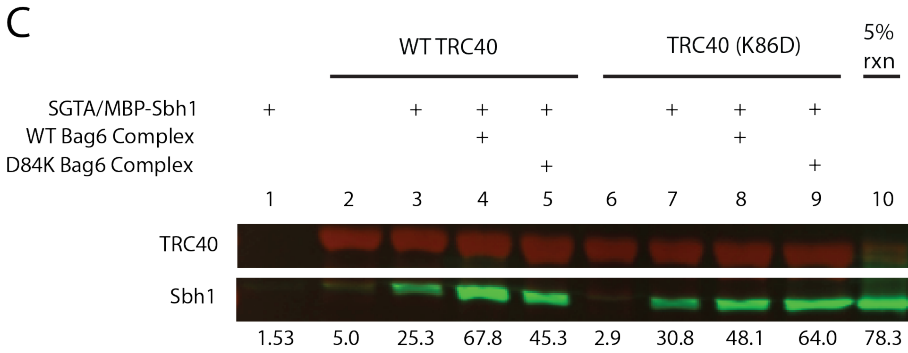

C
WT TRC40
TRC40 (K86D)
5% rxn
SGTA/MBP-Sbh1
WT Bag6 Complex
D84K Bag6 Complex
1 2 3 4 5 6 7 8 9 10
TRC40
Sbh1
1.53 5.0 25.3 67.8 45.3 2.9 30.8 48.1 64.0 78.3

Figure 3.5. The $Bag6_{min}$ complex facilitates TA transfer from SGTA to TRC40

(A) The in vitro TA handoff reaction scheme. Recombinantly purified hSGTA-MBP•Sbh1 complex was incubated with GST•TRC40 and indicated recombinant proteins. After incubation on ice for 10 minutes, GST•TRC40 and bound substrate were precipitated with anti-GST resin followed by three wash steps and Western blotting. (B) Mutants affecting SGTA binding to the $Bag6_{min}$ complex reduce TA transfer to TRC40. GST•TRC40 was captured on anti-GST resin after incubation in the presence of ATP with SGTA/MBP-Sbh1 or SGTA(C38A)/MBP•Sbh1 alone or with the Bag6min or $Bag6_{min}$(Ubl4A(L43A)) complex. Eluted samples were immunoblotted with anti-GST (red) and anti-MBP antibody (green) then quantified by Odyssey Infrared Imaging System analysis software. Relative values of captured Sbh1 underneath each lane with the experiment containing all wild-type components as the reference. Sbh1 fluorescence values were normalized for each trial based GST•TRC40 captured in each lane. Values are averages of six independent experiments. Standard deviations are included in figure 3.7. The 5% rxn lane corresponds, in all cases, to loading 5µL of the wild-type reaction prior to capture. (C) Regulatory mutants GST•TRC40(K86D) and TRC35(D84K)$Bag6_{min}$ complex were incubated with indicated recombinant proteins and ATP then captured on anti-GST resin and analyzed as in B.

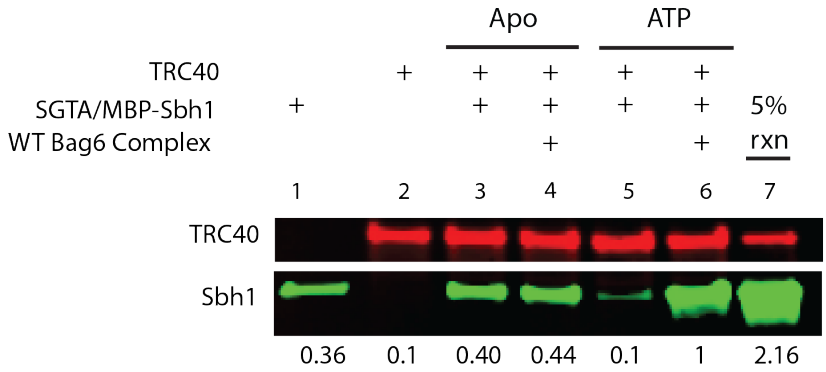


Figure 3.6. The $Bag6_{min}$ complex facilitates TA transfer from SGTA to TRC40 in an ATP dependent manner

Nucleotide-dependent TA handoff facilitated by $Bag6_{min}$ complex. GST•TRC40 was captured on anti-GST resin after incubation with SGTA-MBP•Sbh1 and $Bag6_{min}$ complex with or without ATP. Eluted samples were immunoblotted with anti-GST (red) and anti-MBP antibody (green) then quantified by Odyssey Infrared Imaging System analysis software and the Sbh1 values were normalized based on total GST•TRC40 captured. MBP•Sbh1 signal from WT experiment was designated 1, and the rest represented as a fraction of the WT value. Values are the average of four independent experiments.

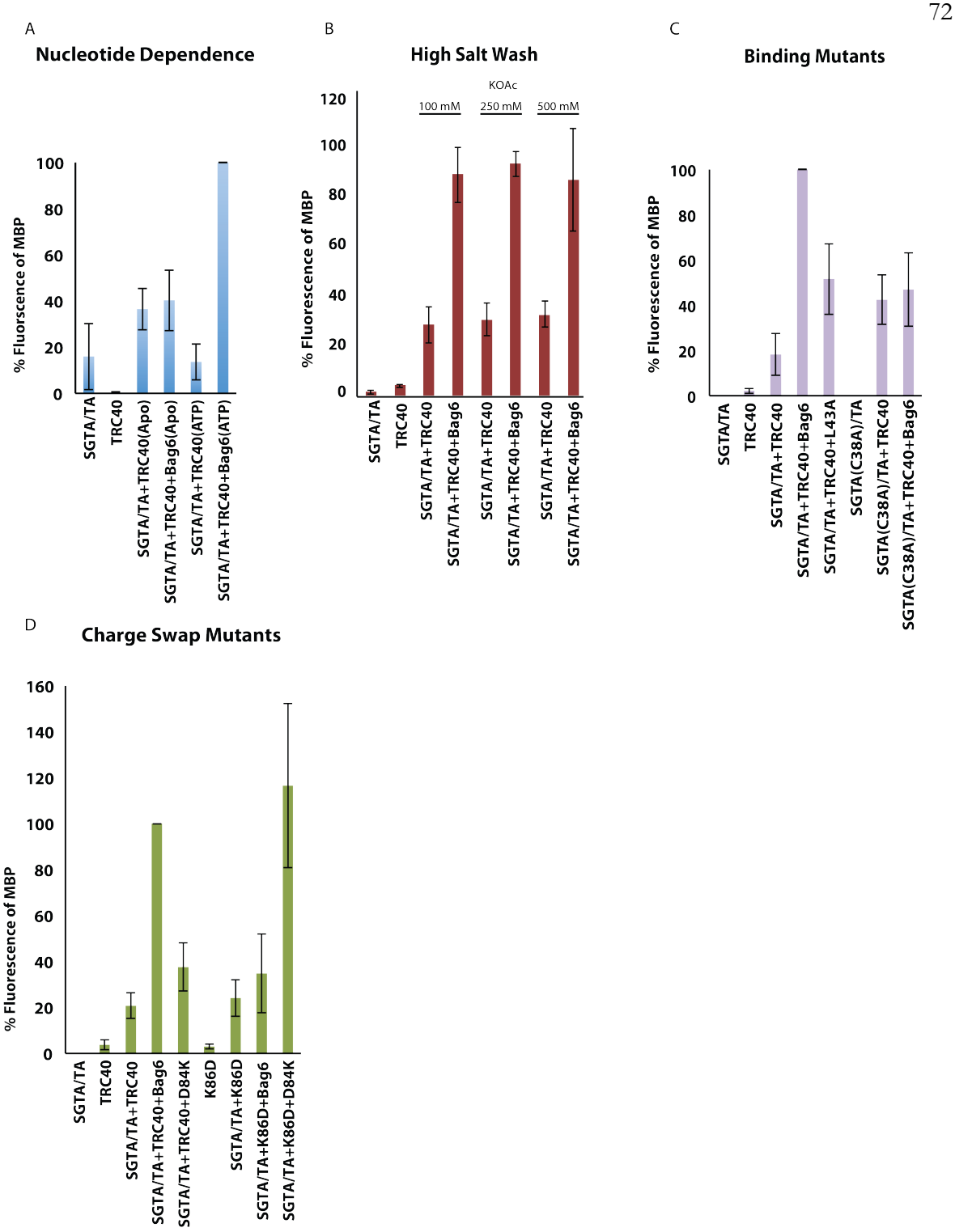

A
Nucleotide Dependence
% Fluorscence of MBP
0
20
40
60
80
100
SGTA/TA
TRC40
SGTA/TA+TRC40(Apo)
SGTA/TA+TRC40+Bag6(Apo)
SGTA/TA+TRC40(ATP)
SGTA/TA+TRC40+Bag6(ATP)
B
High Salt Wash
KOAc
100 mM
250 mM
500 mM
% Fluorescence of MBP
0
20
40
60
80
100
120
SGTA/TA
TRC40
SGTA/TA+TRC40
SGTA/TA+TRC40+Bag6
SGTA/TA+TRC40
SGTA/TA+TRC40+Bag6
SGTA/TA+TRC40
SGTA/TA+TRC40+Bag6
C
Binding Mutants
% Fluorescence of MBP
0
20
40
60
80
100
SGTA/TA
TRC40
SGTA/TA+TRC40
SGTA/TA+TRC40+Bag6
SGTA/TA+TRC40+L43A
SGTA(C38A)/TA
SGTA(C38A)/TA+TRC40
SGTA(C38A)/TA+TRC40+Bag6
D
Charge Swap Mutants
% Fluorescence of MBP
0
20
40
60
80
100
120
140
160
SGTA/TA
TRC40
SGTA/TA+TRC40
SGTA/TA+TRC40+Bag6
SGTA/TA+TRC40+D84K
K86D
SGTA/TA+K86D
SGTA/TA+K86D+Bag6
SGTA/TA+K86D+D84K

Figure 3.7. TA handoff from SGTA to TRC40

(A) Average values of nucleotide-dependent TA handoff from SGTA to TRC40 facilitated by $Bag6_{min}$ complex. Error bars are from four independent experiments. (B) Average values of TA handoff from SGTA to TRC40 facilitated by $Bag6_{min}$ complex in the presence of ATP washed with buffers with varying salt concentrations. Error bars are from three independent experiments. Fluorescence values are represented as a percentage of WT handoff as measured by MBP fluorescence. (C) Average values of TA handoff by binding mutants hSGTA(C38A)/MBP•Sbh1 and $Bag6_{min}$(Ubl4A(L43A)) complex as compared to WT. Error bars are from six independent experiments. (D) Average values of TA handoff of regulatory mutants GST•TRC40(K86D) and $Bag6_{min}$(TRC35(D84K)) complex as compared to WT. Error bars are from four independent experiments.

Materials and Methods

Cloning, Expression, and Purification. cDNA of human Hsc70, Hdj1, and Hsc70-NBD (P5 to S381) were subcloned into pET33b vector, expressed, and purified similar to previously described methods with some modifications (Chartron et al., 2012b). Full-length Hsc70 and Hsc70-NBD were purified over a UnoQ column (Biorad) (50 mM Hepes, 50 - 500 mM KCl gradient, pH 8.0, 5 mM β-mercaptoethanol). Hdj1 was purified over a UnoS column (50 mM Hepes, 20 mM KCl 500 mM KCl, pH 7.0, β-mercaptoethanol). The plasmid vectors containing the cDNA of the chaperone proteins were obtained from the Morimoto group at Northwestern University (Freeman and Morimoto, 1996).

cDNA of human TRC40 was subcloned into pGEX-6P-1 vector and expressed in NiCo21(DE3) cells. Cells were lysed using a M-110L Microfluidizer Processor (Microfluidics) in 50 mM NaH_2PO_4 and 400 mM NaCl supplemented with benzamidine, PMSF, and 5mM β-mercaptoethanol. The protein was purified in a single step Glutathione SuperFlow resin affinity chromatography (Clontech). cDNA of human SGTA was subcloned into pET33b vector, and co-expressed with MBP-Sbh1 in pACYCDuet vector in NiCo21(DE3) cells and lysed using a M-110L Microfluidizer Processor (Microfluidics) in 50 mM NaH_2PO_4, 150 mM NaCl, 20 mM imidazole supplemented with benzamidine, PMSF, and 5mM β-mercaptoethanol. The complex was purified in two steps using Ni-NTA resin (Qiagen) and amylose resin (NEB). Human TRC35(23-305) was subcloned from TRC35 cDNA into pACYCDuet vector in the first multiple cloning site (MCS) with N-terminal 6x histidine tag and a TEV protease cut site.

In the second MCS, cDNA of untagged human Ubl4A was subcloned. TRC35 and was co-expressed with untagged Bag6(1001-1126) in pET33b in NiCo21(DE3). The complex was purified by Ni-NTA affinity chromatography (Qiagen). Contaminants were further removed using chitin affinity chromatography (NEB). TRC40(K86D), TRC35(D84K), SGTA(C38A), and Ubl4A(L43A) mutants were generated using site-directed mutagenesis (Agilent Technologies).

Hsc70 refolding assay. The Hsc70 mediated β-galactosidase refolding assay was carried out as previously reported with modifications. Stock solution of β-galactosidase at 10 mg/ml was prepared by dissolving the enzyme (Sigma-Aldrich) in 50 mM Tris-HCl, 10 mM $MgCl_2$, 5mM β-mercaptoethanol (pH 7.3). For experiments, the stock enzyme was 1:10 diluted in 1 M glycylglycine (pH 7.4). 5 µl of this was diluted into 95 µl of unfolding buffer (25 mM Hepes, 5 mM $MgCl_2$, 50 mM KCl, 5 mM β-mercaptoethanol, 6 M guanidine-HCl, pH 7.4), and 5 µl was diluted into 95 µl 1 M glycylglycine pH 7.4 for the control. Final β-galactosidase concentration was 3.4 nM. The two samples were incubated at 30°C for 30 minutes.

Folding reactions were performed in refolding buffer (1.6 µM Hsc70 and 3.2 µM Hdj1 suspended in 25 mM Hepes, 5 mM $MgCl_2$, 50 mM KCl, 2 mM ATP, 10 mM DTT, pH 7.4). Varying concentrations of Bag6 constructs were tested for their effect on Hsc70 folding activity. After a 30 minute incubation at 30°C, 1 µl of denatured enzyme was added to 124 µl of each refolding reaction tube and incubated at 37 °C. In regular time intervals, 10 µl of each folding reaction was added to 10 µl of 0.8 mg/ml ONPG (ortho-

nitrophenyl-β-galactoside) and incubated at 37°C for 15 minutes. The reaction was stopped by the addition of 80 µl 0.5 M sodium carbonate. β-galactosidase activity was measured as a rate of conversion of ONPG by absorbance at 413 nm.

Hsc70 capture assay

500 pmol of Hsc70-NBD with N-terminal 6x histidine tag was incubated with 2 nmol of untagged Bag6-BAG, Bag6-C81, Bag6E_{BAG}/Ubl4A, and Bag1-BAG in total 100 µl of binding buffer (20 mM Hepes, 100 mM KCl, 20 mM imidazole, and 5 mM β-mercaptoethanol) for 1 hour at room temperature. The samples were then added to 15 µl of Ni-NTA beads (Qiagen). The beads were washed twice with 100 µl binding buffer. Bound proteins were eluted with 15 µl of 20 mM Hepes, 100 mM KCl, 300 mM imidazole, and 5 mM β-mercaptoethanol then run on to 4-20% gradient SDS-PAGE gel (Biorad). 4% of total sample was run as loading control.

For capture assays from 293T whole cell lysate, cells from 10cm dish (90% confluent) were lysed in 1.5 ml NP40 lysis buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 0.5% NP40, 2 mM β-mercaptoethanol, protease inhibitor cocktail). Cell extracts were subject to centrifugation at 20,000 ×g for 5 min to remove insoluble materials. The soluble fractions were pre-treated with 180 ml HisPurTM Coblat resin (Thermo) and then incubated with 30 ml HisPurTM Coblat resin immobilized with His-tagged Bga6-fragments as indicated in the figure at 4 degree for 1 hour. The resins were quickly washed twice with 400 ml a wash buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$), 0.1% NP40, and 2 mM β-mercaptoethanol. The proteins bound to the resin were eluted with 60 ml SDS-PAGE loading buffer and denatured by heating at 65 °C for 10

minutes. The samples were analyzed by SDS-PAGE. Proteins were detected by either Ponceau staining and immunoblotting.

To examine in vivo interaction of Bag6 and Hsp70, cells grown in a 6 well plate were transfected with plasmids expressing FLAG-tagged wild type Bag6 or a mutant Bag6 lacking the C-terminal 81 amino acids using TransIT 293 (Mirus). Cells were lysed in the NP40 lysis buffer 24 hours post transfection. Bag6 was pulled down from the lysate by FLAG M2 beads (Sigma-Aldrich). The precipitated material was analyzed by immunoblotting.

***In vitro* TA handoff assay**

0.625 pmol of hSGTA/MBP•Sbh1 or hSGTA(C38A)/MBP•Sbh1 was incubated with 0.04 μg/μL (6.25 pmol) GST•TRC40 or GST•TRC40(K86D) with or without 0.008 μg/μL of $Bag6_{min}$, $Bag6_{min}$(D84K), or $Bag6_{min}$(L43A) complex in 100 μL of incubation buffer (50 mM Hepes (pH 7.5), 4 mM $Mg(OAc)_2$, 150 mM KOAc, 10% glycerol, and 1 mM DTT) on ice for 10 minutes. After 10 minutes, 10 μL of MagnetGST resin (Promega) was added to each reaction and incubated at room temperature for 15 minutes to pull down GST•TRC40 and bound factors. The resin was washed three times with 500 μL of incubation buffer and eluted with 20 μL of 20 mM Tris, 300 mM NaCl, and 33 mM L-Glutathione pH 7.4. The precipitated material was analyzed by immunoblotting. For the wild-type capture experiment, approximately 12 ± 5.2% of TRC40 and 4.1 ± 1.6% of Sbh1 were eluted from the beads. Assuming one TA per TRC40 dimer, one would expect ~6% of the TA to be captured assuming 100% transfer; therefore, our yield

is reasonable and differences could be attributed to a variety of factors such as differences in stoichiometry assumptions or affinities of the various TRC40 complexes.

Acknowledgements

We thank Yoko Shibata and Richard Morimoto (Northwestern) for plasmids. We thank members of the laboratory for support and useful discussions. W.M.C. is supported by NIH grant R01GM097572.

C h a p t e r 4

BIOCHEMICAL AND CELL BIOLOGICAL INVESTIGATION OF THE MECHANISM FOR NUCLEO-CYTOPLASMIC DISTRIBUTION OF BAG6 BY TRC35

A version of this chapter was first published as

Mock, J.-Y., Xu, U., Ye, Y., and Clemons, W.M. Jr. (2017) Structural basis for regulation of nucleocytoplasmic distribution of Bag6. *Submitted*.

Jee-Young Mock carried out yeast 2-hybrid analysis, biochemical characterization of Bag6 using purified proteins, and exchange assays with purified proteins. Jee-Young Mock participated in designing mammalian cellular experiments.

Abstract

The metazoan protein BCL-2 associated athanogene cochaperone 6 (Bag6) acts as a central hub for several essential cellular processes, including immunoregulation, gene regulation, autophagy, apoptosis, and proteostasis. These roles are in both the nucleus and the cytosol, but the mechanism by which Bag6 trafficking is regulated remains elusive. Here we present biochemical and cell biological characterization of the cytoplasmic retention factor of Bag6, transmembrane domain recognition complex 35 (TRC35). Disrupting the interface between Bag6 and TRC35 results in nuclear localization of Bag6. TRC35 binds Bag6 with higher affinity than karyopherins. Free TRC35 that cannot bind Bag6 at its native binding site is ubiquitylated and degraded. Combined, these results suggest a mechanism for regulation of the nucleo-cytoplasmic distribution of Bag6.

Introduction

The metazoan protein Bag6 is a multidomain protein implicated in various essential cellular processes. Recent efforts have elucidated its extensive cytosolic role as a protein targeting and quality control triaging factor (Mock et al., 2015; Rodrigo-Brenni et al., 2014; Shao et al., 2017; Wang et al., 2011b). It is well known, however, that Bag6 can localize to the nucleus via its nuclear localization sequence (Manchen and Hubberstey, 2001), which has been assumed to be a bipartite nuclear localization sequence.

Several studies have investigated nuclear roles of Bag6. First, Bag6 modulates histone methylation. In U2OS cells, Bag6 constitutively co-localizes in the nucleus with *DOT1 L*ike histone methyltransferase (DOT1L) (Wakeman et al., 2012), a methyltransferase that methylates histone 3 at lysine 79 (H3K79). siRNA-mediated knockdown of Bag6 results in reduced DOT1L-dependent H3K79 methylation in cells treated with ionizing radiation (Wakeman et al., 2012). In HCT116 cells, Bag6 interacts with *B*rother *O*f the *R*egulator of *I*mprinted *S*ites (BORIS), a DNA-binding protein that localizes to the promoter regions of myc and BRCA1 (Nguyen et al., 2008). Bag6 binding facilitates BORIS-mediated histone 3 lysine 4 (H3K4) dimethylation (Nguyen et al., 2008).

Bag6 is also implicated in regulating the acetyltransferase p300. Bag6 promotes the interaction between p300 and p53 (Sasaki et al., 2007), which is crucial for p300-mediated acetylation of p53 in the nucleus. siRNA-mediated knockdown of Bag6 results in reduced p300-mediated acetylation of p53, which cannot be rescued by expressing Bag6ΔNLS (Sasaki et al., 2007). Another study showed that Bag6 promotes p300-mediated p53 acetylation but inhibits p300-mediated acetylation of ATG7 (Sebti et al.,

2014b). These cells also cannot be rescued by expressing Bag6 ΔNLS mutant (Sebti et al., 2014b). Notably, in Bag6$^{-/-}$ mouse embryonic fibroblasts, Bag6 is a nuclear shuttling factor of p300 (Sebti et al., 2014a, b).

Other examples include cell cycle dependent (Yong and Wang, 2012) and ionizing radiation-induced nuclear localization of Bag6 (Krenciute et al., 2013).

While it is clear that some endogenous Bag6 is translocated into to the nucleus, our understanding of the molecular mechanism by which Bag6 distribution is modulated between the nucleus and the cytoplasm is lacking. Importantly, many localization studies did not consider the molecular characterization of the Bag6 complex that includes TRC35, which binds and retains Bag6 in the cytosol (Mock et al., 2015; Wang et al., 2011b). There is strong evidence that Bag6, TRC35 and Ubl4A exist in a trimeric complex in the cytoplasm (Hessa et al., 2011; Liu et al., 2014; Mariappan et al., 2010; Mariappan et al., 2011; Mock et al., 2015; Shao et al., 2017; Wang et al., 2011b; Xu et al., 2013) and influence each other's stability. Knocking down Bag6 leads to reduced cellular levels of TRC35 and Ubl4A (Krenciute et al., 2013), and knocking down both TRC35 and Ubl4A results in reduction of Bag6 levels (Krenciute et al., 2013).

This chapter explores the molecular basis for cytosolic retention of Bag6 by TRC35. Our results reveal that the Bag6-TRC35 interface seen in our crystal structure (Fig. 2.7) is crucial for cytoplasmic retention of Bag6. Furthermore, this interaction is important for maintaining TRC35 stability. Our results, combined with previous human genetics and

qualitative mass spectrometry studies, suggest that nucleo-cytoplasmic distribution of Bag6 is partially mediated by modulation of Bag6 expression.

Results

Probing the Bag6-TRC35 interface

To validate the interfaces observed in the Bag6-TRC35 structure (Fig. 2.7), we generated alanine mutants for analysis by yeast 2-hybrid analysis. A fragment of Bag6 (residues 951-1126) was attached to the GAL4 transcription activating domain and full-length TRC35 was attached to the GAL4 DNA-binding domain. Of the single amino acid substitutions, only one residue, TRC35 (Y262A), disrupted the yeast 2-hybrid interaction (Fig. 4.1A). The Bag6 mutations W1004A and W1012A are localized at interface I and Y1036A is localized at interface II. The combination of the two mutations synthetically disrupted the interaction (W1004A/Y1036A or W1012A/Y1036A) (Fig. 4.1D) confirming that both interfaces are critical for forming a stable complex between Bag6 and TRC35. Expression of Bag6 and TRC35 in the yeast used from two-hybrid experiments was confirmed by immunoblotting (Fig. 4.1C).

We also sought to validate the interaction of TRC35 with full length Bag6 in the context of a mammalian cellular environment by co-immunoprecipitation. For this, we co-expressed N-terminally GFP-tagged wild-type (wt) Bag6 or the mutants Bag6(W1004A), Bag6(W1012), Bag6(Y1036A), Bag6(W1004A/Y1036A), and Bag6(W1012A/Y1036A) with FLAG-tagged TRC35 in a Bag6 knock-out 293T cell ($Bag6^{-/-}$). Bag6 and associated proteins were captured from detergent-derived cell extracts by immunoprecipitating using a GFP-antibody. Immunoblotting analysis showed that these mutations on full length Bag6 protein only partially disrupted the interaction: compared to wtBag6 that efficiently captured TRC35 (Fig. 4.2 lane 9), the single mutation W1004A reduced but not completely

abolished the amount of TRC35 captured by Bag6 (lane 10). This was also true for the double mutants (lane 13 &14).

TRC35 Masks the Nuclear Localization Sequence of Bag6 and Retains Bag6 in the Cytosol

To investigate the Bag6-TRC35 interface in the context of Bag6 localization, the mutants identified from yeast 2-hybrid and immunoprecipitation were used for localization studies. Overexpression of TRC35 has been shown to retain wtBag6 in the cytosol (Wang et al., 2011b), suggesting that TRC35 binding is required for cytosolic localization of Bag6. Because the mutations identified in our yeast 2-hybrid experiments specifically prevent TRC35 binding, we postulated that exogenously expressed Bag6 mutants defective in TRC35 binding would localize primarily in the nucleus regardless of TRC35 expression. To test this hypothesis, wt and mutant Bag6 were expressed in Cos7 cells with or without TRC35•FLAG and the localization of Bag6 and TRC35 was examined by immunofluorescence.

As expected (Wang et al., 2011b), given the NLS, overexpressed wtBag6 and various Bag6 mutants were localized to the nucleus (Fig. 4.3A). This is likely due to excess Bag6 that cannot be retained in the cytosol by endogenous TRC35. Indeed, when TRC35 is co-expressed with wtBag6, the increased cytosolic pool of TRC35 captures wtBag6 and both proteins stain primarily in the cytosol (Fig. 4.3A). In accordance with the immunoprecipitation results (Fig. 4.2), introduction of a single mutation—W1004A, W1012A, or Y1036A—results in some Bag6 localization to the nucleus (Figs. 4.3B, C, and D). Mutations at both interface I (W1004A or W1012A) and interface II (Y1036A) further

reduce binding between Bag6 and TRC35 (Fig. 4.2) and Bag6 localizes primarily to the nucleus (Fig. 4.3E and F). These results unequivocally establish TRC35 as a cytoplasmic retention factor of Bag6.

Free TRC35 is Ubiquitylated

Careful examination of the immunoprecipitation results (Fig. 4.2) revealed two intriguing observations: (1) an unexpected increase in TRC35 binding to Bag6 double mutants relative to single mutants even though these double mutants contain exposed NLS due to lack of TRC35 shielding (Fig. 4.2, compare lanes 11-12 vs. lanes 13-14) and (2) the appearance of higher molecular weight products for Bag6 double mutants (Fig. 4.2, asterisk). As there is evidence that the stability of TRC35 requires forming a proper complex with Bag6 (Krenciute et al., 2013; Mariappan et al., 2010) and given the implication of Bag6 as a chaperone holdase in protein quality control processes such as mis-localized protein degradation (Hessa et al., 2011) and ER-associated protein degradation pathways (Payapilly and High, 2014; Wang et al., 2011b), we postulated that TRC35 mutants that fail to form a complex with Bag6 at its physiological binding site are unstable and become a target for degradation pathways. This would result in TRC35 becoming a target for Bag6 dependent degradation leading to TRC35 binding at the Bag6 substrate-binding site. In this case, the higher molecular weight bands observed in figure 4.2 would probably be ubiquitylated TRC35 bound to Bag6.

To verify this hypothesis, $Bag6^{-/-}$ 293T cells were used to co-express TRC35, Bag6 variants and ubiquitin with the expectation that destabilized TRC35 would show increased ubiquitin complexes. Cells were transfected with TRC35•FLAG, Bag6•GFP, and HA•ubiquitin.

Proteins bound to Bag6 were first immunoprecipitated with GFP antibody. To remove other ubiquitinated Bag6 substrates (Minami et al., 2010; Wang et al., 2011b; Xu et al., 2013), the samples obtained from the GFP immunoprecipitation were subject to a second round of immunoprecipition using FLAG beads under denaturing conditions. Immunoblotting analysis of the samples from the first round of immunoprecipitation showed that all Bag6 variants pulled down ubiquitinated proteins, suggesting that these mutations did not affect its substrate-binding activity (Fig. 4.4, lane 2-7). Re-immunoprecipitation with FLAG antibody showed that TRC35 associated with wtBag6 carried a small amount of ubiquitin conjugates, but those associated with Bag6 variants that disrupted physiological association with TRC35 carried significantly more ubiquitin conjugates. Compared to single Bag6 mutations, TRC35 bound to the double mutants had the highest ubiquitin to TRC35 ratio (Figs. 4.4 compare lanes 10-12 to 13-14) supporting the idea that the subset of TRC35 molecules unable to associate with Bag6 via the NLS domain are unstable and become targets for ubiquitin-dependent degradation through a client-chaperone interaction with Bag6. This suggests that the ubiquitylated TRC35 associates with the quality control module (QC) of Bag6 (Shao et al., 2017).

We sought to ensure that TRC35 is a Bag6-QC substrate by investigating the effect of the E3 ligase RNF126 on TRC35 ubiquitylation. RNF126 is a Bag6-associated E3 ligase utilized by the Bag6-QC for ubiquitylation of Bag6-associated clients in the cytosol (Rodrigo-Brenni et al., 2014). Ubiquitylated Bag6-QC substrates are proteasomally degraded (Shao et al., 2017). If TRC35 ubiquitylation is mediated by Bag6-QC, knocking down RNF126 in cells expressing Bag6 mutants would result in reduced ubiquitylation of

TRC35. Furthermore, if ubiquitylated TRC35 is degraded by the ubiquitin proteasome system, RNF126 reduction would also stabilize TRC35. To test this hypothesis, the effect of RNF126 and proteasome inhibition on TRC35 ubiquitylation and stability were examined.

We first examined the effect of siRNA-mediated RNF126 knockdown on TRC35 ubiquitylation. 293T cells co-expressing HA•ubiquitin, TRC35•FLAG and either wt or mutant Bag6•GFP (W1004A/Y1036A or W1012A/Y1036A) were treated with siRNA against RNF126. Changes in ubiquitylation were compared by taking the relative ratio between ubiquitylated TRC35 and unmodified TRC35 (Ub-TRC35/TRC35). The ratio in cells expressing both wtTRC35 and wtBag6 without RNF126 was defined as 1. In cells expressing Bag6 mutants, TRC35 bound to Bag6 is mostly ubiquitylated (Fig. 4.5A lanes 13 and 14); the relative ratio between ubiquitylated TRC35 (Ub-TRC35) and unmodified TRC35 in cells expressing Bag6(W1004A/Y1036A) and Bag6(W1012A/Y1036A) are 193 and 111, respectively (Fig. 4.5B). Knockdown of RNF126 reduces TRC35 ubiquitylation in cells expressing Bag6 mutants, resulting in ~10-20-fold reduction in Ub-TRC35/TRC35 (Fig. 4.5B). In contrast the Ub-TRC35 to TRC35 ratio in cells expressing wtBag6 remains constant (Figs. 4.5A lane 12 & 16 and 4.5B) regardless of RNF126 treatment, suggesting that TRC35 that is bound to Bag6 at the native binding site is stable. Both unmodified and polyubiquitylated TRC35 is stabilized by proteasome inhibition by MG132 (Fig. 4.5C lanes 9 & 10 vs. 11 & 12), confirming that Ub-TRC35 is degraded by the ubiquitin proteasome system. The accumulation of unmodified TRC35 (Fig. 4.5C lanes 9 & 10 vs 11 & 12) is probably due to depletion of free ubiquitin in the cell (Melikova et al., 2006). To

better observe the changes in accumulation of Ub-TRC35 upon RNF126 knockdown, cells were simultaneously treated with RNF126 siRNA and MG132 to prevent the degradation of Ub-TRC35. For both cells expressing wtBag6 and Bag6(W1004A/Y1036A), accumulation of ubiquitylated TRC35 was reduced (Fig. 4.6 compare lanes 9 & 10 vs. 11 & 12). Together these results demonstrate that ubiquitylation of free TRC35 is modulated by the quality control role of Bag6 and RNF126.

TRC35 has higher affinity for Bag6 than Karyopherin-α 2 (KPNA2)

All molecules destined for the nucleus must move through the nuclear pore complex (NPC), a large assembly that spans the nuclear envelope and facilitates nucleo-cytoplasmic traffic (Hoelz et al., 2011). Macromolecules larger than ~40 kDa cannot freely diffuse through the NPC and require carrier proteins, such as the karyopherin-α (KPNA) and -β families of transport receptors (Lange et al., 2007; Pumroy and Cingolani, 2015). The two basic clusters, R_{1024}KVK and K_{1043}RRK (Fig. 4.7A), in Bag6 are thought to act as a bipartite NLS by specifically recognizing the acidic substrate-binding surface of karyopherins (Manchen and Hubberstey, 2001). In HeLa cells, the K_{1045}R to S_{1045}L mutation has been shown to abrogate nuclear localization of Bag6 (Manchen and Hubberstey, 2001). Therefore, TRC35 and KPNA both bind Bag6 at the fragment that contains the NLS although likely distinct residues mediate the respective interactions.

To confirm that the Bag6-NLS is a KPNA binding site and to define the residues involved in KPNA binding, a Bag6 C-terminal 131 residues (Bag6C131) (Wu et al., 2004) in complex with hexahistidine-tagged full-length Ubl4A (Bag6C131-6xHis•Ubl4A) was purified from E. coli. KPNA2 was chosen specifically as it had been seen to interact with

Bag6 (Rouillard et al., 2016) and could be stably purified. Mutations previously shown to abrogate Bag6 nuclear localization (Manchen and Hubberstey, 2001) were introduced at either the first basic cluster (1024SL) or the second basic cluster (1045SL) (Fig. 4.7A). The purified Bag6C131-6xHis•Ubl4A variants were incubated with either GST•TRC35 or MBP•KPNA2 (58-529) and the resulting complexes were isolated using Ni-NTA beads. GST•TRC35 pulled down wild-type, 1024SL, and 1045SL Bag6C131-6xHis•Ubl4A with similar efficiency (Fig. 4.7B lanes 5-7), demonstrating that these residues are not involved in binding TRC35. MBP•KPNA2 formed a sufficiently stable complex with Bag6 that can be visualized in a pull-down where MBP•KPNA2 captures wtBag6 (Fig. 4.7B lane 9). Surprisingly, only mutating of the second basic cluster led to significant disruption of the interaction between Bag6C131-6xHis•Ubl4A and MBP•KPNA2 (Fig. 4.7B compare lane 10 and 11). These results show that the residues required for binding KPNA2 are distinct from those required for binding TRC35, and TRC35 acts as a cytosolic retention factor by occluding the Bag6 NLS. Moreover, the second basic cluster is necessary and sufficient for binding KPNA2.

If TRC35 binding to Bag6 prevents KPNA-mediated nuclear translocation, only TRC35-free Bag6 should be able to bind KPNA and be translocated to the nucleus. There are several ways in which this could be achieved. One such is that if KPNAs have a higher affinity for Bag6 than TRC35, upregulation of KPNA expression would lead to displacement of TRC35 from Bag6 and formation of a Bag6-KPNA complex. To test this, we sought to compare the binding affinities of the TRC35 and KPNA2 to Bag6 using an exchange assay. Bag6C131-6xHis•Ubl4A complexes with GST•TRC35 were generated

and bound to glutathione affinity resin beads via the GST-tag. After washing, varying amounts of MBP•KPNA2 were added to the bound beads. The ability of MBP•KPNA2 to displace Bag6C131-6xHis•Ubl4A from GST•TRC35 was determined by the amount of Bag6C131-6xHis•Ubl4A that was eluted from the resin after incubation. In this case, even at the highest concentration tested (2x molar excess), there was no significant displacement of Bag6 from TRC35 by KPNA2 (Fig. 4.8 lanes 4-7). Performing the opposite experiment, starting with MBP•KPNA2-Bag6C131-6xHis•Ubl4A on amylose beads, adding excess GST•TRC35 resulted in the dissociation of the MBP•KPNA2-Bag6C131-6xHis•Ubl4A complex (Fig. 4.8 lanes 11-14). These results highlight the stability of the TRC35-Bag6 complex and argue against the ability of KPNA regulation as a means for modulating the nuclear pool of Bag6.

Discussion

Bag6 is a critical scaffolding factor that has important nuclear and cytosolic roles. It is unclear how the localization of Bag6 is regulated. Here we report biochemical and cell biological characterization of the Bag6-TRC35 complex and suggest a mechanism for regulation of Bag6 localization.

Our results demonstrate that TRC35 acts as an intermolecular mask to the monopartite Bag6 NLS in a role similarly performed in other pathways. Examples of other cytosolic retention factor pairs include IκB and NF-κB (Beg et al., 1992), HIC and Rev (Gu et al., 2011) and BRAP2 that retains HMG20A (Davies et al., 2013). Unlike other cytoplasmic retention factors, which have only been shown to bind their target NLS-containing proteins for occlusion of the NLS, TRC35 also plays a distinct role in the cytoplasmic TA targeting and protein quality control when in complex with Bag6. This dual functionality seems to have been evolutionarily conserved. One study showed that Ubl4A also has a nuclear role of promoting STAT3 dephosphorylation (Wang et al., 2014b). This study did not explore the localization of Bag6 relative to Ubl4A, but Bag6 that translocates to the nucleus brings Ubl4A with it (Krenciute et al., 2013). In yeast, fungal Get4 binds the N-terminal domain of Get5, which appears to contain a functional NLS that directs Get5 to the nucleus during a 'Get5-mediated stress response' (Arhzaouy and Ramezani-Rad, 2012).

These results allow speculation of possible regulatory mechanism for Bag6 nuclear localization. First, disrupting the Bag6-TRC35 interface by introducing alanine mutations results in Bag6 and RNF126-dependent ubiquitylation and degradation of TRC35 (Fig. 4.4). Similarly, knocking down Bag6 has been shown to reduce levels of TRC35 in HeLa

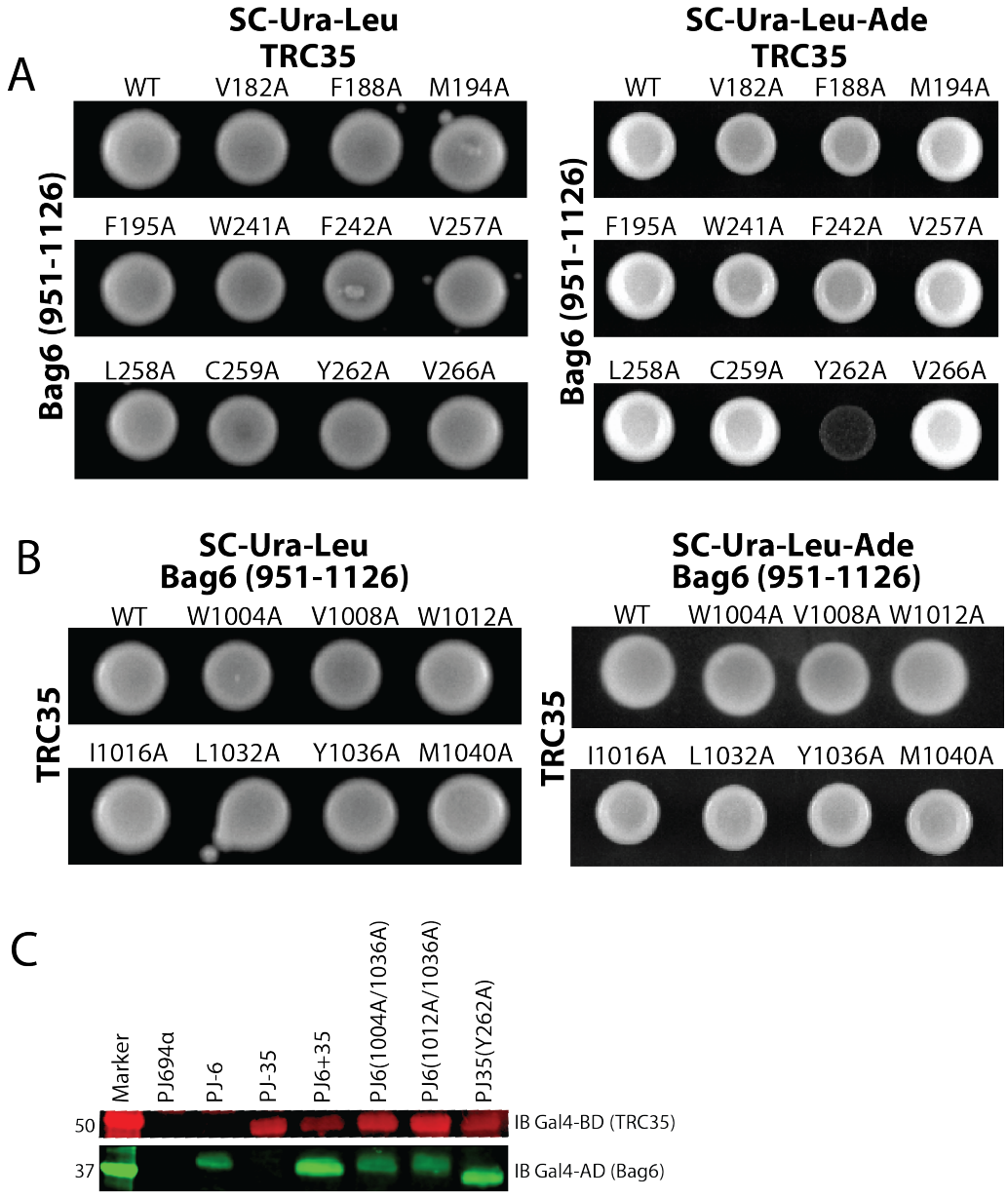
cells (Krenciute et al., 2013), demonstrating that Bag6 is required for TRC35 stability. We also show that KPNA2 has a lower affinity for Bag6 than TRC35 (Fig. 4.8B), which suggests that for Bag6 to bind KPNA and translocate into the nucleus, it needs to be free of TRC35. The most likely explanation is that cells regulate Bag6 localization by modulating Bag6 levels or decreasing TRC35 levels. In humans, Bag6 rs3117582 single nucleotide polymorphism at the promoter region of Bag6, likely affecting expression levels, is associated with higher incidence of lung cancer (Chen et al., 2014; Etokebe et al., 2015b; Zhao et al., 2014) and osteoarthritis (Etokebe et al., 2015a). The localization could also be pre-translationally regulated with differential splicing. In brain and breast tissue, for instance, Bag6 isoforms that lack the NLS are expressed at higher levels than isoforms with the NLS (Luce et al., 2016).

TRC35 utilization of the Bag6 quality control module has important implications on previous Bag6 knockdown studies. When Bag6 is knocked down TRC35 protein levels, not mRNA levels, decrease (Krenciute et al., 2013). Thus, when Bag6 is knocked down, TRC35 translation probably continues. Knocking down Bag6 then not only eliminates Bag6 localization to the nucleus due to higher TRC35 to Bag6 ratio, but also affects the quality control module of Bag6 because excess TRC35 would become Bag6 substrates. The combination of these effects perhaps explains the pleiotropic effects of Bag6 reduction or removal from cells.

Our biochemical studies unequivocally establish TRC35 as a cytosolic retention factor of Bag6 and suggest that Bag6 needs to be in excess for nuclear localization. However, how

different cell types regulate Bag6 localization, whether the TRC35-dependent regulation can be mediated by specific stress, and the biological implications of differential distribution of Bag6 are important questions for future studies.

Figures


A
SC-Ura-Leu
TRC35
SC-Ura-Leu-Ade
TRC35
WT V182A F188A M194A
F195A W241A F242A V257A
L258A C259A Y262A V266A
Bag6 (951-1126)
B
SC-Ura-Leu
Bag6 (951-1126)
SC-Ura-Leu-Ade
Bag6 (951-1126)
WT W1004A V1008A W1012A
I1016A L1032A Y1036A M1040A
TRC35
C
Marker
PJ694α
PJ-6
PJ-35
PJ6+35
PJ6(1004A/1036A)
PJ6(1012A/1036A)
PJ35(Y262A)
50
37
IB Gal4-BD (TRC35)
IB Gal4-AD (Bag6)

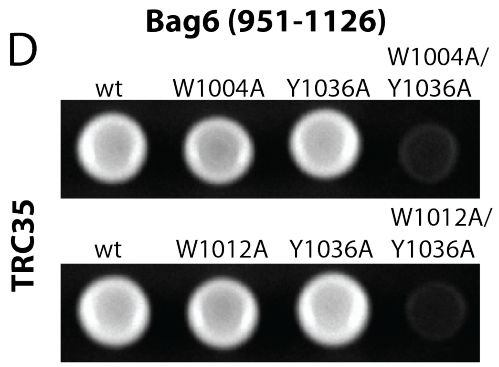


Figure 4.1. Validation of the Bag6-TRC35 interface

(A) Yeast 2-hybrid assay to validate the interface identified in the crystal structure. Wild-type or mutant full length TRC35 conjugated to the DNA binding domain was expressed with wild-type Bag6(951-1126) conjugated to the transcription activating domain. Transformation was confirmed by ability to grow on SC-Ura-Leu media. Interaction was determined by ability to grow on SC-Ura-Leu-Ade media. (B) Wild-type full-length TRC35 conjugated to the DNA binding domain was expressed with wild-type or mutant Bag6(951-1126) conjugated to the transcription activating domain. (C) Expression of TRC35 and Bag6 in yeast cells used from yeast 2-hybrid was examined by Western blot. Antibodies against Gal4 DNA binding domain or trans-activating domain were used to detect expression of Gal4BD-TRC35 and Gal4AD-Bag6. (D) Combination of mutations at interface I (W1004A and W1012A) and interface II (Y1036A) is sufficient for disrupting yeast 2-hybrid interaction.

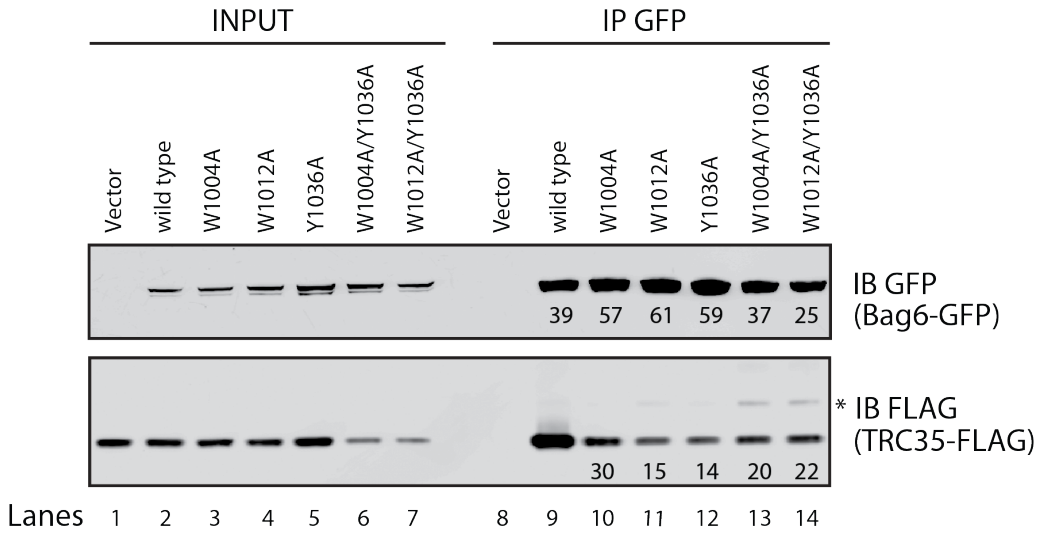


Figure 4.2. Validation of Bag6-TRC35 interface in mammalian cells

Wild-type or mutant Bag6•GFP was co-expressed in Bag6$^{-/-}$ 293T cells with TRC35•FLAG and immunoprecipitated using anti-GFP antibody. Amount of TRC35 retrieved by Bag6 was assessed by blotting with anti-FLAG antibody. The position of the higher molecular weight TRC35•FLAG is highlighted with an asterisk.

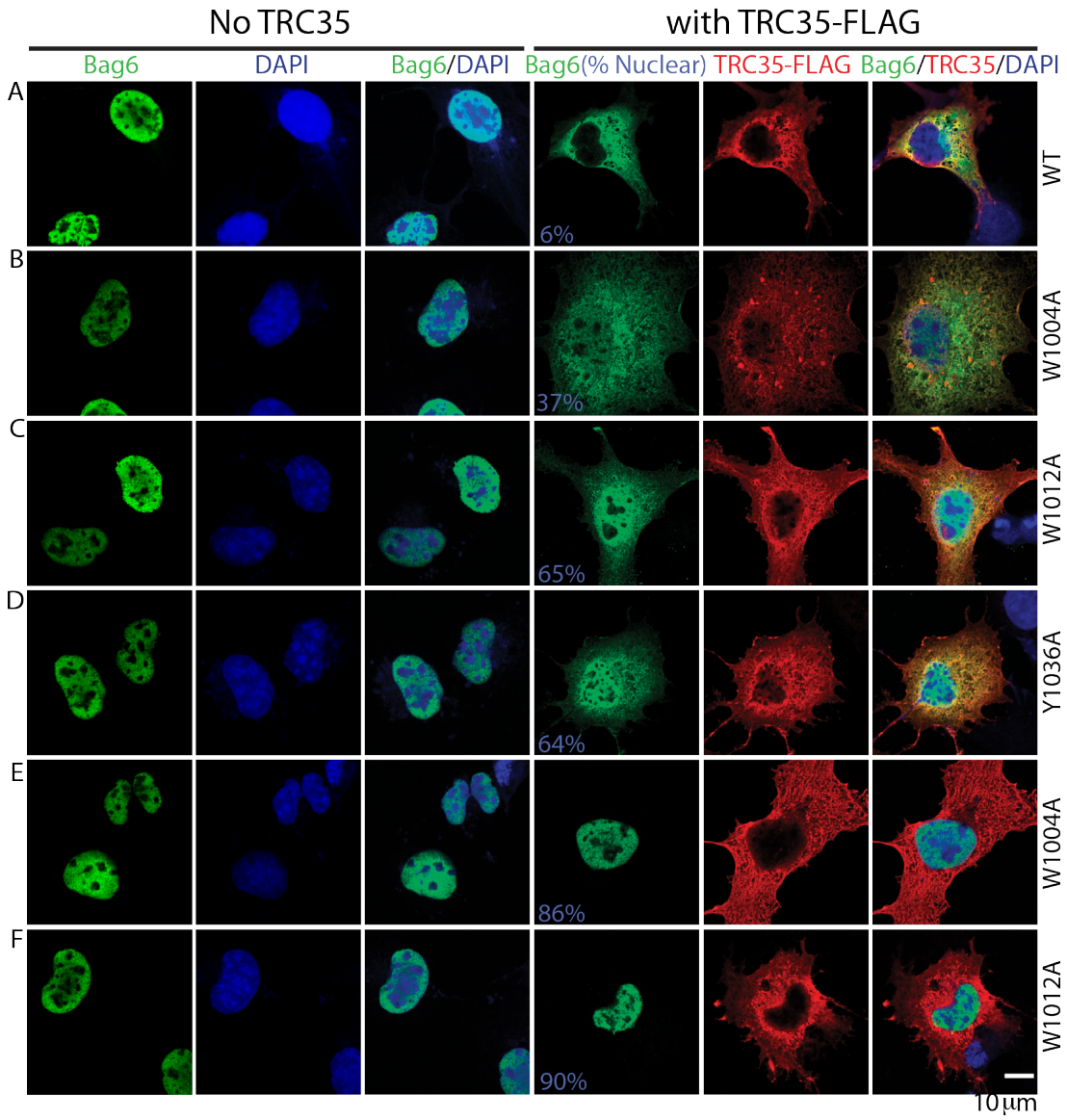


Figure 4.3. Bag6 mutations at the TRC35 binding site results in nuclear localization of Bag6

(A-F) Cos7 cells were transfected either with Bag6•GFP (wt or mutant) expressing plasmid alone or co-transfected with TRC35•FLAG (wt) expressing plasmid. Cells were stained with anti-GFP (green) and/or anti-FLAG (red) antibodies. DNA was stained with

DAPI where indicated (blue). In the Bag6 column, the percent of Bag6 calculated to be in the nucleus is noted. Shown are representative cells imaged by a confocal microscope.

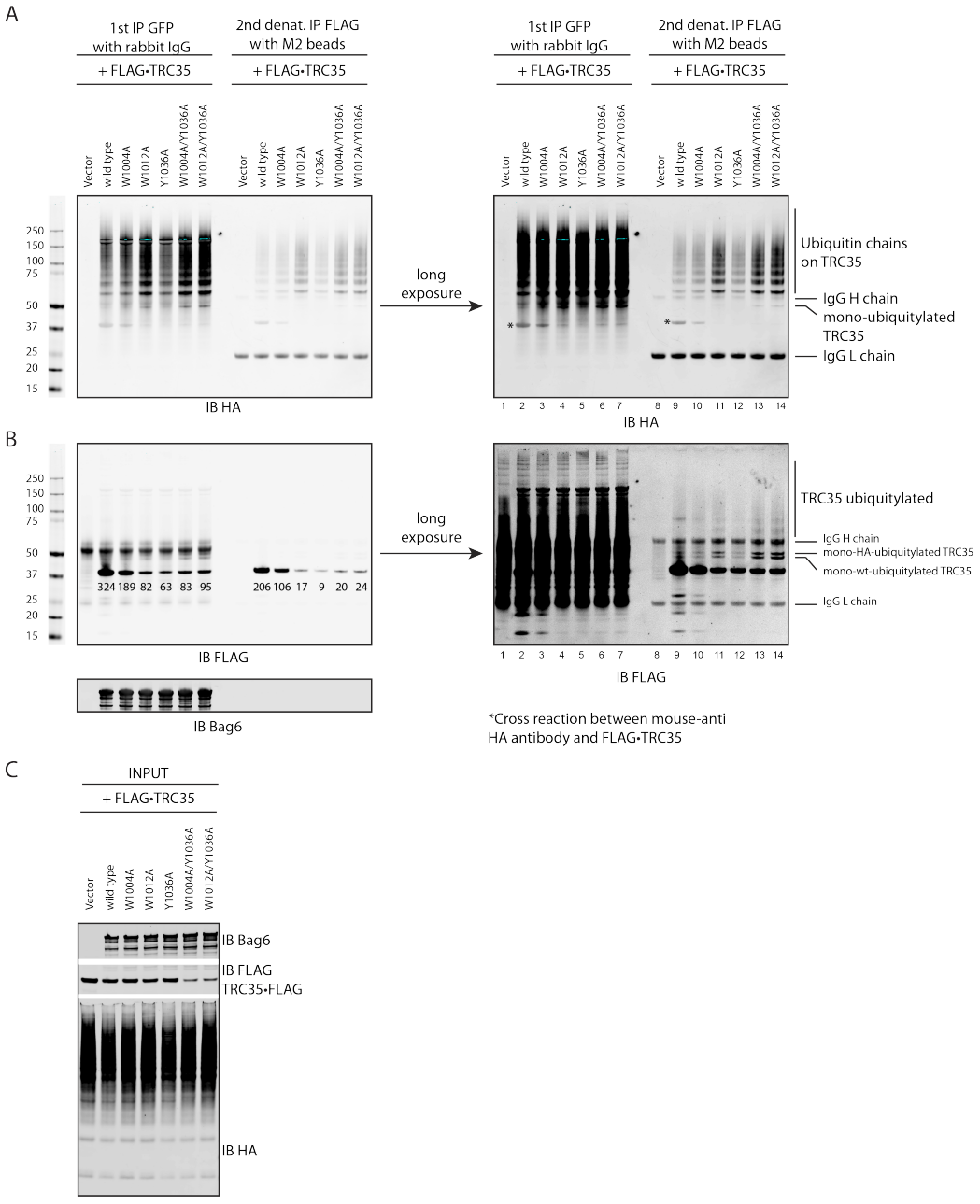


Figure 4.4 Ubiquitylation of TRC35 upon mutant Bag6 expression

(A) Immunoprecipitation (IP) of TRC35 in Bag6$^{-/-}$ 293T cells co-transfected with plasmids encoding TRC35•FLAG (wt), Bag6•GFP (wt or mutants), and HA•ubiquitin. Anti-GFP antibody was used for the first IP. Anti-FLAG antibody was used for the second IP in

denaturing conditions. TRC35 ubiquitination was assessed by immunoblotting with anti-HA antibody. The exposure times were adjusted to improve visibility of the reactive bands. (B) IP was carried out as in (A) in Bag6$^{-/-}$ 293T cells co-transfected with plasmids encoding TRC35•FLAG (wt) and Bag6•GFP (wt or mutants). TRC35 ubiquitination was assessed by immunoblotting with anti-FLAG antibody. The exposure times were adjusted to improve visibility of the reactive bands. (C) The cell extract used for immunoprecipitation was immunoblotted for Bag6, TRC35, and ubiquitin with Bag6 antibody, FLAG antibody and HA antibody, respectively.

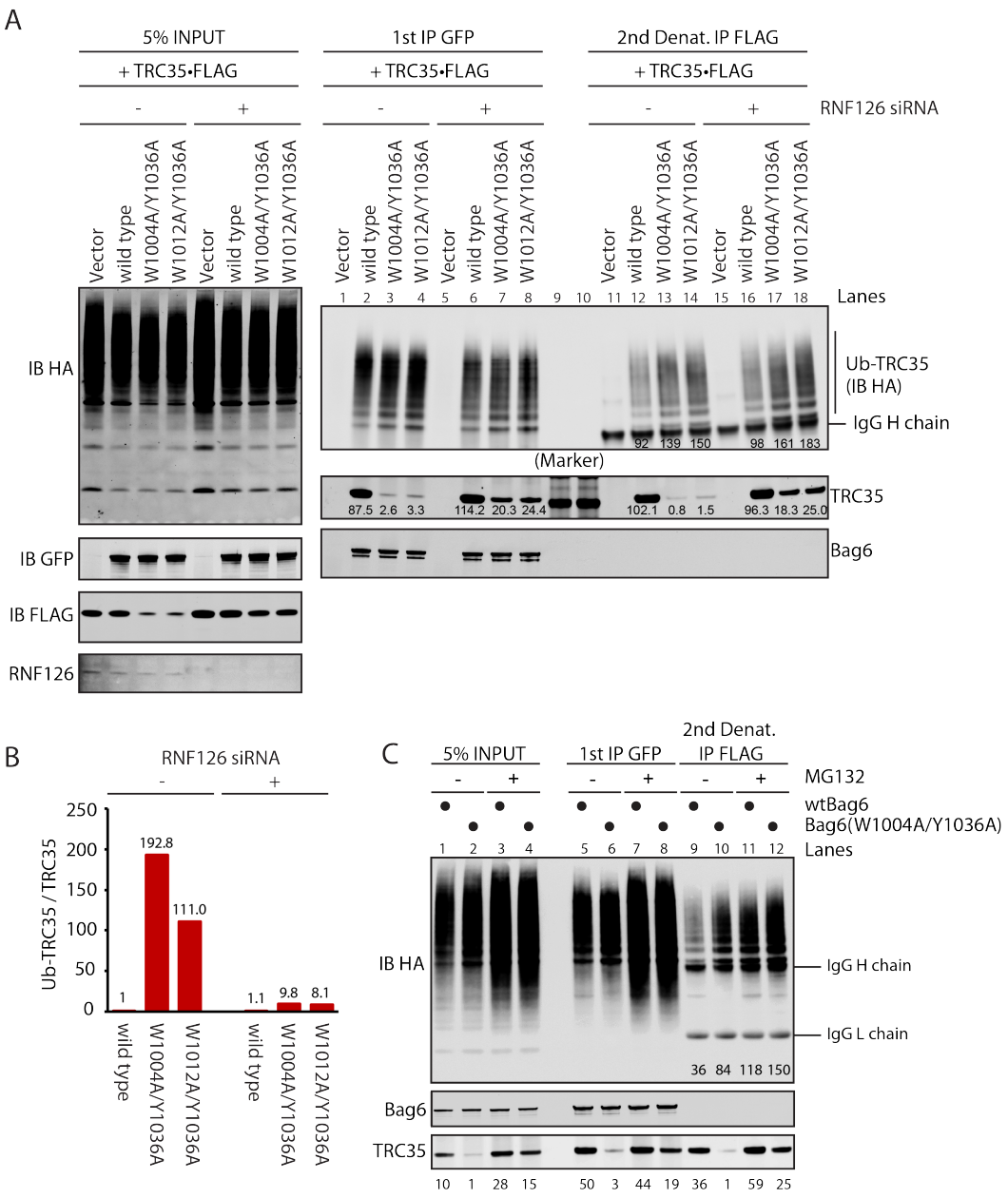


Figure 4.5 RNF126 knockdown stabilizes TRC35 in cells expressing mutant Bag6

(A) wt293T cells expressing TRC35•FLAG, Bag6•GFP and HA•ubiquitin were treated with siRNA against RNF126. Cells were lysed and TRC35 bound to Bag6 were immunoprecipitated first with GFP antibody then with FLAG antibody in denaturing condition. The amount of TRC35 and ubiquitylated TRC35 was assessed by

immunoblotting with anti-TRC35 antibody and anti-HA antibody, respectively. (B) The relative amounts of ubiquitylated TRC35 and unmodified TRC35 from figure 4.5 are calculated as a ratio. (C) 293T cells expressing wtTRC35•FLAG, Bag6•GFP (wt or W1004A/Y1036A) and HA•ubiquitin were treated 10 μM MG132. Bag6 and TRC35 were immunoprecipitated and immunoblotted as in A.

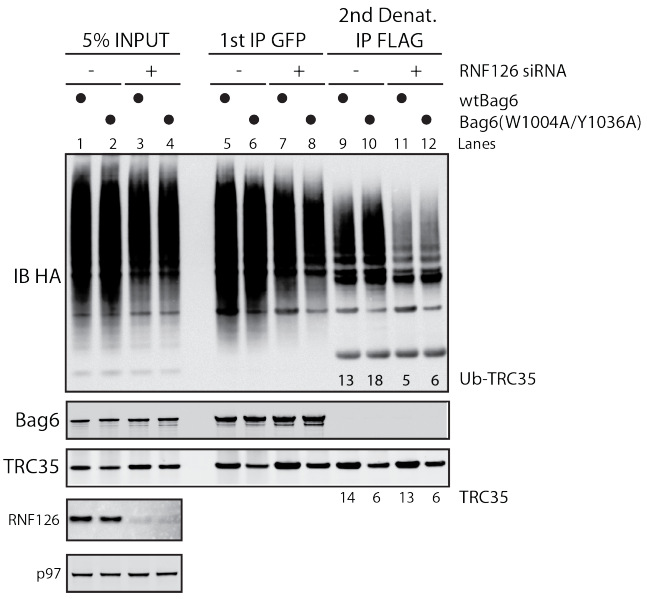


Figure 4.6 TRC35 ubiquitylation is dependent on RNF126.

wt293T cells co-transfected with plasmids encoding TRC35•FLAG (wt), Bag6•GFP (wt or mutants), and HA•ubiquitin. The cells were simultaneously treated with MG132 and siRNA against RNF126. The cell extracts were subject to two rounds of denaturing immunoprecipitation with anti-GFP antibody then with anti-FLAG antibody. TRC35 ubiquitylation was assessed by immunoblotting with HA antibody.

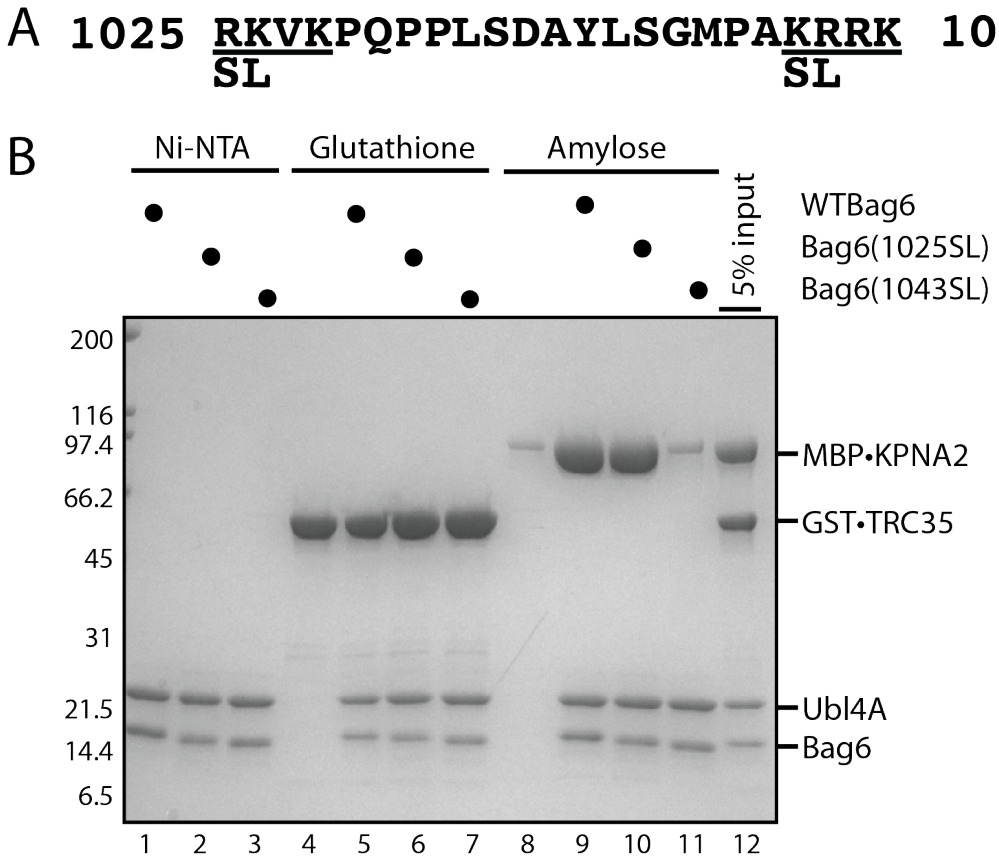


Figure 4.7 Biochemical identification of the monopartite Bag6 NLS

(A) The putative bipartite nuclear localization sequence of Bag6. The serine and leucine mutations introduced in this study are highlighted. (B) Recombinantly purified Bag6C131-6xHis•Ubl4A was incubated with excess GST•TRC35 or MBP•KPNA2 for 20 minutes at room temperature. Ni-NTA beads were used to capture purified Bag6C131-6xHis•Ubl4A and bound factors.

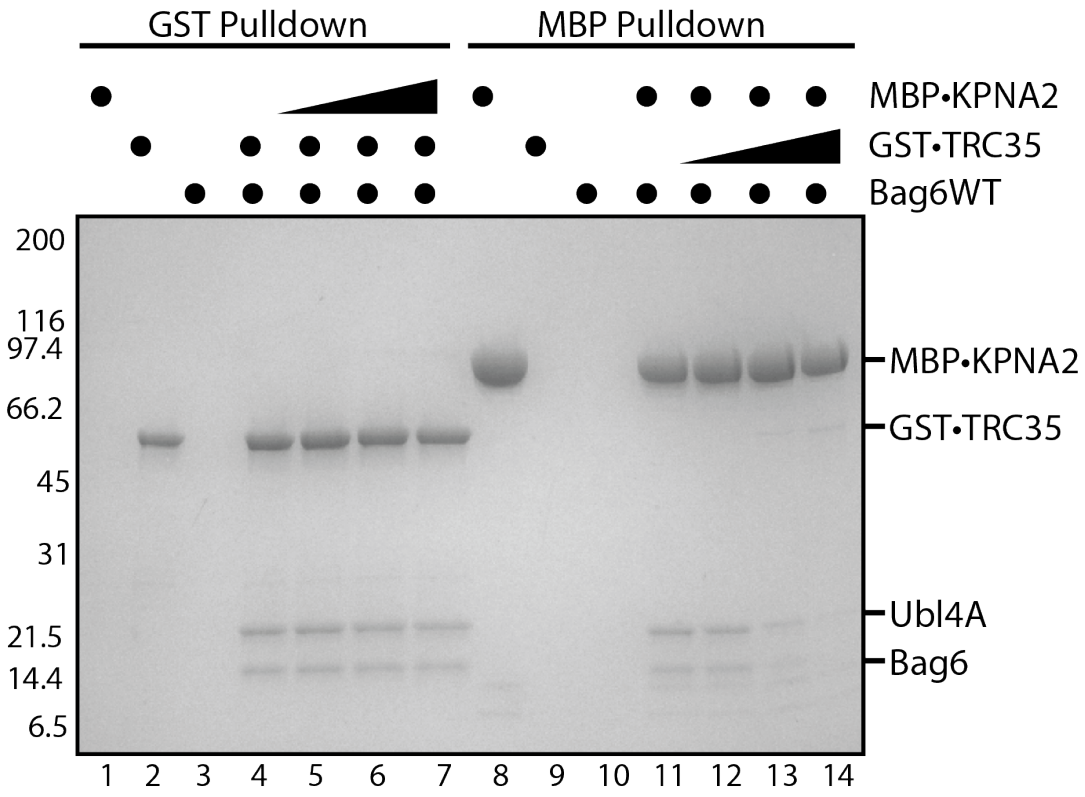


Figure 4.8 TRC35 binding precludes karyopherin α binding to Bag6

Recombinantly purified Bag6C131-6xHis•Ubl4A was first incubated with either GST•TRC35 or MBP•KPNA2. The resulting GST•TRC35-Bag6C131-6xHis•Ubl4A complex was incubated with glutathione resin beads, and increasing amounts of MBP•KPNA2 was added. The ability of MBP•KPNA2 to displace Bag6C131-6xHis•Ubl4A from GST•TRC35 was examined by eluting the GST•TRC35 and bound Bag6C131-6xHis•Ubl4A from the glutathione resin. The opposite experiment was also carried out using amylose resin beads.

Materials and Methods

Expression and purification

The plasmid containing human KPNA2 was obtained from Addgene (#26677). A truncated KPNA2 (58-529) was subcloned into pMAL-C2 vector (New England Biolabs) and transformed into E. coli NiCo21(DE3). The cells were grown in 2×YT media and induced at 37 °C until OD600 = 0.1 then cooled on ice for 1 hour. Expression was induced with 500 μM IPTG for 18 hours at 16°C in a shaking incubator at 200 rpm. The cells were harvested as above and resuspended in 50 mM Mops (pH 7.2), 300 mM K•glutamate, 5 mM β-mercaptoethanol, supplemented with cOmplete EDTA-free protease inhibitor cocktail (Roche) and lysed using the M-110L microfluidizer by two passes at approximately 17,500 psi. The lysate was clarified by centrifugation at 235,000 ×g in a Beckman Ti45 rotor for 30 minutes at 4°C. The lysate was incubated for 1 hour with 3 mL of a 50% (vol/vol) slurry of amylose resin (New England Biolabs) by rocking. The mix was poured into a gravity column then washed with 100 mL lysis buffer. The protein was eluted with 12 mL elution buffer (20 mM Mops (pH 7.2), 150 mM K•glutamate, 10 mM maltose, 5 mM β-mercaptoethanol). The sample was placed in snakeskin dialysis bag (10 kDa cutoff, Thermo Fisher) and dialyzed overnight at 4 °C in 20 mM Mops (pH 7.2), 50 mM - 800 mM K•glutamate gradient, 5 mM β-mercaptoethanol. The fractions containing the protein (~36 - 70 mL) were pooled and dialyzed in snakeskin dialysis bag in 20 mM Mops (pH 7.2), 50 mM K•glutamate, 5 mM β-mercaptoethanol for 2 hours at 4 °C. The sample was concentrated to 2 mL, filtered with a 0.22 μm syringe filter, and purified by size-exclusion chromatography over a 120 mL Superdex 200 column (GE Healthcare) (20 mM Mops (pH 7.2), 50 mM K•glutamate,

5 mM β-mercaptoethanol). Fractions containing the sample (60 mL – 90 mL) were pooled and concentrated with centrifugal filtration units with 50 kDa molecular weight cutoff.

For expression and purification of GST•TRC35, full-length TRC35 was subcloned into pGEX6P-1 (GE Healthcare) and transformed into NiCo21(DE3). The cells were grown at 37 °C until OD_{600} = 0.1, chilled on ice for 1 hour, then induced with 500 μM IPTG at 16 °C for 18 hours in a shaking incubator at 200 rpm. Cells were lysed using 10 mL per g cell pellet lysis buffer (50 mM Hepes (pH 7.2), 400 mM KCl, 5 mM mercaptoethanol) and incubated with 3 mL 50% (vol/vol) slurry of glutathione resin (GE healthcare) for 2 hours at 4 °C by rocking. The resin was washed with 100 mL lysis buffer and the protein was eluted with 12 mL of freshly prepared 20 mM Hepes (pH 7.2), 150 mM KCl, 33 mM glutathione, 5 mM β-mercaptoethanol. The sample was placed in snakeskin dialysis bag (10 kDa cutoff, Thermo Fisher) and dialyzed overnight at 4 °C in 20 mM Hepes (pH 7.2), 100 mM K•glutamate, 5 mM β-mercaptoethanol, before concentration with centrifugal filter unit (50 kDa molecular weight cutoff, Millipore).

Yeast Two-Hybrid

The PJ69-4α strain was obtained from the Yeast Resource Center at the University of Washington. Bag6 isoform b residues 951 to 1126 were cloned into pGAD-C1 vector and TRC35 was cloned into pGBDU-C1 vector. Alanine mutations were made using Q5 site-directed mutagenesis (New England Biolabs). pGAD-C1-TRC35 (wt or mutant) and pGBDU-C1-Bag6 (wt or mutant) were co-transformed into PJ69-4α using previously described methods (12) and then plated onto SC-Ura-Leu plates and incubated at 30 °C.

Single colonies from the transformants were cultured in 5 mL SC-Ura-Leu liquid media and grown overnight in a shaking incubator (Multitron Standard Infors HT) at 200 rpm at 30 °C. 2×10^7 cells were transferred into total 5 mL SC-Ura-Leu media and grown in a shaking incubator at 30 °C at 200 rpm for 6 hours. Cells were harvested by centrifugation with a Beckman SX4750A rotor at 3000 ×g at 25 °C. The cells were washed twice by resuspension in 5 mL sterile water followed by centrifugation at 3000 ×g at 25 °C. After the second wash, cells were resuspended in 1 mL sterile water. Concentration was measured and 1×10^7 cells were resuspended in total 40 µL of sterile water. 4 µL of this resuspended sample were spot plated onto SC-Ura-Leu-Ade plate and incubated for 72 hours at 30 °C.

Yeast Two-hybrid Expression Controls

Cells were grown overnight in SC-Ura-Leu medium. 2.5 x 10^7 cells were transferred to a total 5 mL media and grown for ~6-8 hours until OD_{600} = 2.0. Cells were harvested by centrifugation at 4000 ×g for 5 minutes at 4 °C. Cells were resuspended in 2 mL 10 mM TE buffer (Tris-HCl, pH 8.0, 0.1 mM EDTA). Cells were centrifuged in a microcentrifuge at 13,500 ×g at 4 °C for 2 minutes. ~37.5 µL glass beads (415-600 µm size) for each OD unit were added to the sample. Same volume of 1x SDS-PAGE loading buffer was added. The cells were vortexed and immediately plunged into a boiling water bath. After boiling for 3 minutes, tubes were immersed into an ice bucket. The tubes were placed into FastPrep (MP biomedicals) homogenizer at 6 m/s for 45 seconds. Cells were boiled for 3 minutes and centrifuged in a microcentrifuge for 2 minutes at 13,500 g at room temperature. Supernatant was transferred to a clean microfuge tube and loaded onto 5-20% gradient gel

(Biorad). The protein was transferred onto PVDF membrane. Membrane was blocked with Odyssey TBS blocking buffer (Li-Cor) for 1 hour. The membrane was incubated with primary antibody against Gal4 DNA binding domain (for TRC35 detection) or Gal4 trans-activating domain (for Bag6 detection). After washing with TBST (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.2% Tween-20) for 10 minutes 3 times, membrane was incubated with anti-mouse secondary antibody (IRDye 680RD or 800CW, Li-Cor) for 1 hour at room temperature. Membrane was washed 3 times with TBST for 10 minutes each. Proteins were detected using an Odyssey Imaging System (Li-Cor).

Immunoprecipitation from 293T cells

On day 0, 5 x 10^5 $Bag6^{-/-}$ cells were seeded in plates. On day 1, TRC35 with a C-terminal FLAG tag (TRC35•FLAG), Bag6 with a C-terminal GFP tag (wt or mutant Bag6•GFP), and HA•ubiquitin were co-transfected. On day 3, cells were collected and washed with 1 mL of ice cold phosphate-buffered saline. The cells were lysed with NP40 lysis buffer containing 1 mM N-ethylmaleimide to inhibit the ubiquitin-proteasome system. The detergent soluble fraction was used for immunoprecipitation (IP) using GFP antibody. After IP, the beads were divided into two fractions. Half was analyzed directly with SDS-PAGE and immunoblotting with indicated antibodies to assess the interaction between TRC35 and Bag6 and Bag6 mutants. The other half was used for denaturing immunoprecipitation to detect TRC35 ubiquitylation. The beads were resuspended in 150 µL denaturing buffer (1× PBS, 1% SDS, 5 mM DTT) and heated at 95 °C for 10 minutes. Then 1.35 mL NP40 lysis buffer was added into the tube and incubated at 4 °C for 30 minutes. The supernatant was used for immunoprecipitation with FLAG antibody

conjugated M2 beads (Sigma-Aldrich). The eluate was analyzed with SDS-PAGE and immunoblotting.

For RNF126 knockdown, 5×10^5 wild type 293T cells were seeded in 6 well plates. On day 1, cells were transfected with 60 pmol RNF126 siRNA (Thermo Fisher) or control siRNA using Lipofectamin RNAiMAX (Thermo Fisher). The sequence of the RNF126 siRNAs are GCAUCUUCGAUGACAGCUU (catalog number S31185), GAUUAUAUCUGUCCAAGAU (catalog number S31186), and GCAGGGCUACGGACAGUUU (catalog number S13387). Cells were passaged in 1:2 ratio onto new plates on day 2. Cells were transfected with TRC35•FLAG (wt), Bag6•GFP (wt or mut) and HA•ubiquitin on day 3. On day 4, the cells were collected and cell extracts were assayed with sequential immunoprecipitation and immunoblotting as above. To assay TRC35 ubiquitylation after proteasome inhibition, 5×10^5 wild type 293T cells were seeded in plates on day 0. On day 1, cells were transfected as above. 24-hours post transfection, proteasome inhibitor, MG132, was added to the concentration of 10 μM and incubated overnight. Cells were collected and analyzed with sequential immunoprecipitation and immunoblotting as above.

To investigate the effect of simultaneous RNF126 down-regulation and proteasomal inhibition on TRC35 ubiquitylation, 5×10^5 Bag6 293T cells were seeded in plates on day 0. On day 1, cells were transfected with RNF126 siRNA or control siRNA. On day 2, cells were split in half into new plates. After ~9 hours, plasmids encoding TRC35•FLAG (wt), Bag6•GFP (wt or mutant) and HA•ubiquitin were transfected. On day 3, medium was

replaced with fresh medium containing either 10 μM MG132 or DMSO. On day 4, cells were collected and sequential IP was performed as described above.

Localization assay

On day 0, 1×10^5 Cos7 cells were seeded onto a 12-well plate with a poly-D-lysine coated cover glass. After approximately 8 hours, the cells were co-transfected with TRC35•FLAG and Bag6•GFP (wt or mutant). 20 hours after transfection, the cells were washed with 1× PBS, fixed with 4% (vol/vol) paraformaldehyde for 15 minutes, then washed with 1× PBS before permeabilization with staining solution (1× PBS, 5% fetal bovine serum (vol/vol), 0.1% (vol/vol) NP40 with primary and secondary antibodies). The staining solution was replaced with staining solution with FLAG antibody and incubated at room temperature for 1 hour. After washing the cells with 1× PBS, the cells were incubated with staining solution with secondary antibody then washed with 1× PBS. Cells were counterstained with a mounting medium containing DAPI to illuminate the nucleus. Cover glass was mounted for visualization with Axiovert 200M microscope (Zeiss).

To assay the localization of Bag6, 3 view fields were evaluated suing an Axiovert 200 inverted microscope equipped with a 40x oil immersion objective. Confocal analyses were performed with a Zeiss LSM 780 system. Approximately 100 cells were counted for each constructed tested. The localization of Bag6 in the nucleus was assessed visually and cells were categorized as either Bag6 out of nucleus or Bag6 inside the nucleus.

***In vitro* binding assay with purified proteins**

nmol hexahistidine-tagged Ubl4A-Bag6C131 (wt, 1025SL or 1043SL) was incubated with 1 nmol GST•TRC35 or 1 nmol MBP•KPNA2 in 100 μL total volume of binding buffer (20 mM Mops (pH 7.2), 100 mM K•glutamate 20 mM imidazole, 5 mM β-mercaptoethanol) at room temperature for 30 minutes. 30 μL 50% slurry of Ni-NTA beads (Qiagen) equilibrated with binding buffer were added to the reaction and incubated at room temperature for 30 minutes. The beads were resuspended using a pipet every 10 minutes. The beads were then washed twice with 100 μL room temperature binding buffer. Samples bound to the resin were eluted with 25 μL elution buffer (20 mM Mops (pH 7.2), 100 mM K•glutamate, 300 mM imidazole, 5 mM β-mercaptoethanol) and evaluated with Coommassie-stained SDS-PAGE gel.

Exchange assay

1 nmol hexahisidine-tagged Ubl4A-Bag6C131 (wt, 1024SL, or 1043SL) was incubated with 1 nmol GST•TRC35 in 100 uL total volume of binding buffer (20 mM Mops (pH 7.2), 100 mM K•glutamate, 20 mM imidazole, 5 mM β-mercaptoethanol). 30 μL 50% (vol/vol) slurry of glutathione beads (GE Healthcare) equilibrated with binding buffer were added to the reaction and incubated at room temperature for 30 minutes. The beads were resuspended every 10 minutes. The beads were washed twice with 100 μL room-temperature binding buffer (20 mM Mops (pH 7.2), 100 mM K•glutamate, 5 mM β-mercaptoethanol), and then 0.5, 1 or 2 nmol MBP•KPNA in 100 μL was added to the resin and incubated for 30 minutes at room temperature. The beads were washed twice with 100 μL binding buffer and then eluted with 25 μL elution buffer (20 mM Mops (pH

7.2), 100 mM K•glutamate, 33 mM glutathione, 5 mM β-mercaptoethanol). The reverse experiment starting with hexahistidine-tagged Ubl4A-Bag6C131 and MBP•KPNA2 was carried out as above but using 30 μL 50% (vol/vol) slurry of amylose beads (New England Biolabs) in 20 mM Mops (pH 7.2), 100 mM K•glutamate, 5 mM β-mercaptoethanol and eluted with 25 μL of 20 mM Mops (pH 7.2), 100 mM K•glutamate, 10 mM maltose, and 5 mM β-mercaptoethanol.

Acknowledgements

We thank members of the laboratory for support and useful discussions. We thank Yihong Ye and Yue Xu for useful discussions and technical assistance. W.M.C. is supported by NIH grant R01GM097572.

Chapter 6

BAG6: A MODULAR MULTITASKER

Concluding Remarks

Initial appreciation of the unique nature of Bag6 as a component of the TA targeting machinery led to two major questions that this dissertation seeks to address. The first question pertains to the difference in the molecular architecture between the fungal Get4-5 and the mammalian Bag6-TRC35-Ubl4A sorting complexes. Atomic resolution crystal structures presented in this study reveal surprising degree of structural conservation. Although the incorporation of Bag6 results in overall reorganization—from a heterotetramer to a heterotrimer—the structural elements crucial for protein-protein interactions between the TA targeting components have all been conserved (Fig. 5.1). Accordingly, the trimeric Bag6 complex for TA sorting is functional equivalent to the Get4-5 sorting complex.

Importantly, the architectural reorganization leads to changes in stoichiometry. In the fungal system, the heterotetramer of Get4-5, which contains two copies of both Get4 and two Get5, could bring together two dimers of Get3, ultimately forming a Get3 tetramer. In the mammalian system, the sorting complex includes one TRC35 and one Ubl4A, which bind a single dimer of TRC40 (Fig. 5.1). Homo-oligomerization of Bag6 has been observed *in* vitro (Xu et al., 2013), but it is unclear whether this oligomeric complex is physiologically relevant. Similarly, physiological relevance of the Get3 tetramer is under debate in the field. Several lines of evidence support the existence of a tetrameric Get3. First, a crystal structure of a tetrameric archaeal homologue of Get3 has been solved (Suloway et al., 2012). Second, recombinantly expressed Get3-TA complex purifies as a tetramer (Bozkurt et al., 2009; Suloway et al., 2012) and is competent for *in vitro* TA

insertion into purified microsomes (Suloway et al., 2012). Third, Get3 tetramerization stimulates its ATPase activity by ~100 fold (Rome et al., 2013). However, the only atomic resolution structure of a Get3-TA complex solved thus far suggests that dimeric Get3 (Mateja et al., 2015) is sufficient for binding and targeting TA proteins. It should be noted that the crystallized Get3-TA complex was artificially stabilized by (1) introducing a mutation that rendered Get3 incompetent for ATP hydrolysis and (2) adding high-affinity synthetic antibody fragments. As a result, the crystal structure could have trapped and captured a specific intermediate or an artificial state of the Get3-TA complex. Whether tetrameric Get3/TRC40 plays a physiological role in TA targeting remains to be seen.

Ultimately, the observation that the C-terminal TA targeting module—the minimal Bag6 complex—characterized in this study is structurally and functionally equivalent to the fungal Get4-5 complex begged the question: what different purpose, if any, does Bag6 serve?

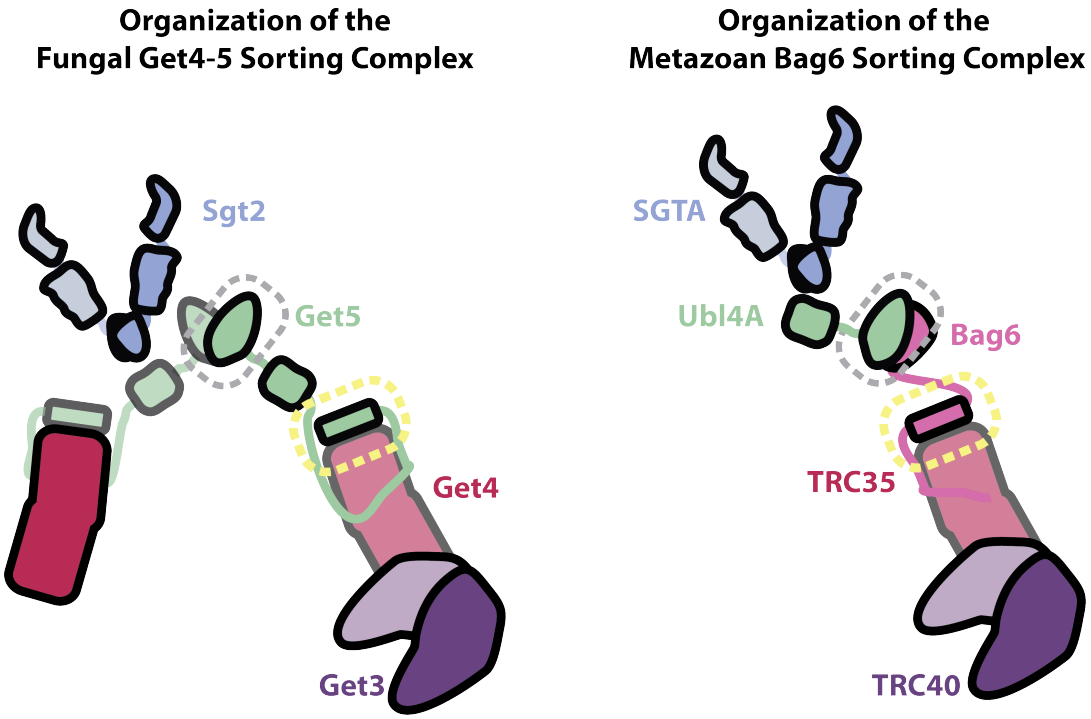


Figure 5.1 Cartoon summary of the fungal and metazoan sorting complexes. Left: fungal Get4-5 TA sorting complex and its binding partners, Sgt2 and Get3, are illustrated. Right: metazoan Bag6 TA sorting complex and its binding partners, SGTA and TRC40, are illustrated. The structurally conserved dimerization domains, whose structures have been solved in this study, are highlighted with grey and yellow dotted boxes.

One purpose Bag6 serves is to physically couple protein targeting and quality control. A recent study demonstrated that purified N-terminal domain of Bag6 (Fig. 5.2), which excludes the TA targeting module, is sufficient for substrate ubiquitylation (Shao et al., 2017). Bag6 seems to decide the fate of TA substrate as it is handed off from SGTA by utilizing either its C-terminal targeting module for productive synthesis of well-folded proteins or its N-terminal quality control module for degradation of misfolded proteins. Such coordination, enabled by modularity, likely minimizes the risk of aggregation for nascent TA proteins.

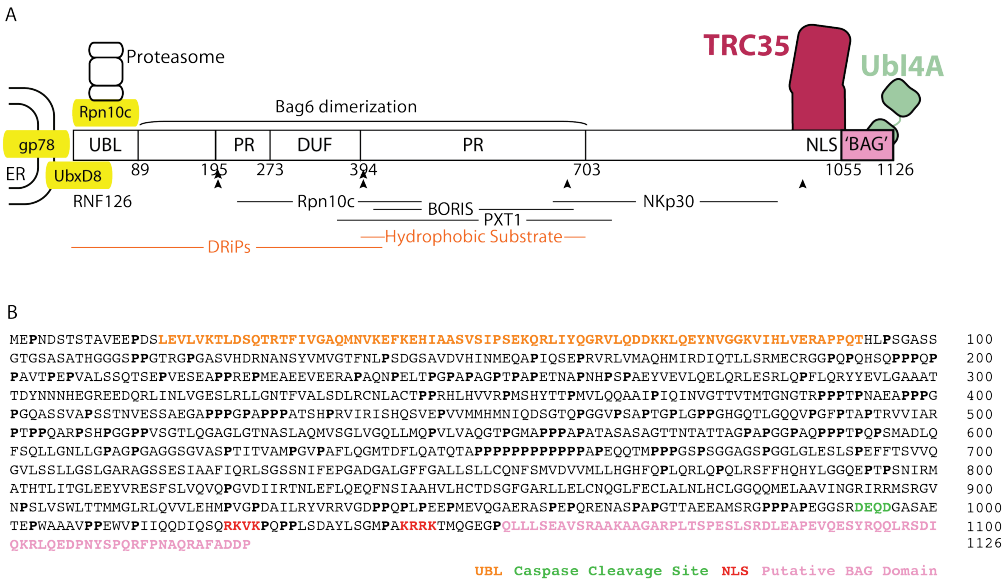


Figure 5.2 Summary of Bag6 domains and binding partners. (a) The Bag6 complex is a scaffold for a broad range of activities. (a) A model of the Bag6 complex and its defined interactions. For Bag6, the locations of the ubiquitin-like domain (UBL), proline-rich domain (PR), domain of unknown function (DUF), nuclear localization sequence (NLS), and the BAG domain (pink) are shown. Arrows indicate the five fragments of Bag6 (A, B, C, D, E) used in this study. Regions required for interaction with Rpn10c, BORIS, PXT1, and NKp30 are indicated in black lines. The regions required for binding of hydrophobic substrates and polyubiquitylated defective ribosomal products (DRiPs) are displayed in orange. The region required for Bag6 dimerization is indicated on top. Proteins with defined interactions are shown as colored boxes. Bag6-UBL binding proteins are highlighted in yellow. The membrane-embedded UbxD8 and transmembrane protein gp78 are thought to anchor the Bag6 complex to the ER. Rpn10c is a component of the proteasome. For the Bag6 complex, TRC35 (dark red) and Ubl4A (green) are shown with their domains indicated. (b) Amino acid sequence of the Bag6 isoform used in this study. Key structural elements are highlighted in unique colors.

The demonstrated nuclear localization of Bag6 may also be an important functionality conferred by Bag6 incorporation. Our structural and biochemical characterization of the Bag6-TRC35 complex unequivocally establishes TRC35 as the cytoplasmic retention factor for Bag6. While the exact function of Bag6 in the nucleus remains to be seen, cell-cycle dependence of the nucleo-cytoplasmic distribution of Bag6 (Yong and Wang, 2012) and observations of endogenous Bag6 localization in the nucleus (Krenciute et al., 2013; Wang et al., 2011b) suggest that nuclear localization of Bag6 is physiologically relevant. Furthermore, our exchange assay demonstrated that excess Bag6 relative to TRC35 is required for nuclear translocation. In fact, estimated physiological ratio of Bag6 to

TRC35 in HeLa cells as determined by qualitative mass spectrometry is 1.2 (Kulak et al., 2014). We postulate that the relative protein abundance of Bag6 and TRC35 is fine-tuned, in which Bag6 acts as a stoichiometry sensor. TRC35 levels would be interpreted as a signal for Bag6 to localize to the nucleus. As such, results presented here necessitate re-interpretation of a large body of work in the field that investigated the localization of Bag6 independent of TRC35 and Ubl4A.

Our interest in Bag6 began in the context of membrane protein biosynthesis. Bag6, however, is a multifaceted protein that regulates diverse cellular processes in complex multicellular animals. Its modularity demonstrates the benefits of modular protein structure in coordinating cellular signals.

Juxtaposition of functional modules leads to physical proximity and coupling, resulting in efficiency. In the case of Bag6, it reduces the amount of time that a failed nascent substrate would spend in the cytosol. The incorporation of Bag6 in metazoans suggests that in higher eukaryotes, the cost of misfolded protein and potential aggregation is higher than it is in fungi.

Modularity also allows for additional fine-tuning of protein function via alternative splicing. Due to the modular nature of Bag6 protein function, the Bag6 gene can generate multiple proteins with varying combinations of functional modules. Of the ~20 predicted Bag6 isoforms documented on the NCBI database, 7 are missing the BAG domain identified in this study. Bag6 that lacks Ubl4A would not be competent for TA handoff, but its quality control module remains intact (Shao et al., 2017) (Fig. 3.7). Its shortest

predicted isoform is 903 residues long and is missing a large portion of the second proline rich domain (Fig. 5.2). Analysis of RNA-seq data from the Illumina Human Body Map Project revealed that Bag6 transcripts without NLS are abundant in breast and brain, while transcripts with NLS are abundant in liver, lung, testes, prostate, kidney and lymph nodes (Luce et al., 2016). In rats, different Bag6 isoforms are expressed in distinct developmental stages (Kwak et al., 2008). Thus, in addition to the modulation of TRC35 expression, alternative splicing would affect Bag6 localization and function. The functional and physiological consequence of different Bag6 isoforms, however, is unclear and is a question for future studies.

Recent structural and biochemical breakthroughs have led to rapid leaps in our understanding of TA targeting and Bag6, but challenging questions remain. Is the Get3/TRC40 tetramer physiologically relevant? Does Bag6 change the stoichiometry of TRC components in metazoans? How is TRC35 expression, and Bag6 localization, modulated? Could some of the pleiotropic effects of knocking down Bag6 be explained through the resulting destabilization of TRC35? A multifaceted approach that combines structural and mechanistic characterization of Bag6 with cell biological and organismal studies is required for a complete understanding of this fascinating protein.

Bibliography

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr *66*, 213-221.

Akahane, T., Sahara, K., Yashiroda, H., Tanaka, K., and Murata, S. (2013). Involvement of Bag6 and the TRC pathway in proteasome assembly. Nat Commun *4*, 2234.

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

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Arakawa, A., Handa, N., Ohsawa, N., Shida, M., Kigawa, T., Hayashi, F., Shirouzu, M., and Yokoyama, S. (2010). The C-terminal BAG domain of BAG5 induces conformational changes of the Hsp70 nucleotide-binding domain for ADP-ATP exchange. Structure *18*, 309-319.

Arhzaouy, K., and Ramezani-Rad, M. (2012). Nuclear import of UBL-domain protein Mdy2 is required for heat-induced stress response in Saccharomyces cerevisiae. PLoS One *7*, e52956.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res *37*, W202-208.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol *2*, 28-36.

Baldauf, S.L., and Palmer, J.D. (1993). Animals and fungi are each other's closest relatives: congruent evidence from multiple proteins. Proc Natl Acad Sci U S A *90*, 11558-11562.

Banerji, J., Sands, J., Strominger, J.L., and Spies, T. (1990). A gene pair from the human major histocompatibility complex encodes large proline-rich proteins with multiple repeated motifs and a single ubiquitin-like domain. Proc Natl Acad Sci U S A *87*, 2374-2378.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr *67*, 271-281.

Beg, A.A., Ruben, S.M., Scheinman, R.I., Haskill, S., Rosen, C.A., and Baldwin, A.S., Jr. (1992). I kappa B interacts with the nuclear localization sequences of the subunits of NF-kappa B: a mechanism for cytoplasmic retention. Genes Dev *6*, 1899-1913.

Bjorklund, A.K., Ekman, D., Light, S., Frey-Skott, J., and Elofsson, A. (2005). Domain rearrangements in protein evolution. J Mol Biol *353*, 911-923.

Blobel, G., and Dobberstein, B. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol *67*, 835-851.

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

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M., and Paciorek, W. (2003). Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr D Biol Crystallogr *59*, 2023-2030.

Briknarova, K., Takayama, S., Brive, L., Havert, M.L., Knee, D.A., Velasco, J., Homma, S., Cabezas, E., Stuart, J., Hoyt, D.W.*, et al.* (2001). Structural analysis of BAG1 cochaperone and its interactions with Hsc70 heat shock protein. Nat Struct Biol *8*, 349-352.

British Mycological Society. Symposium (1986 : University of Bristol), Rayner, A.D.M., Brasier, C.M., and Moore, D. (1987). Evolutionary biology of the fungi : Symposium of the British Mycological Society, held at the University of Bristol, April 1986 (Cambridge ; New York: Cambridge University Press).

Chang, Y.W., Chuang, Y.C., Ho, Y.C., Cheng, M.Y., Sun, Y.J., Hsiao, C.D., and Wang, C. (2010). Crystal structure of Get4-Get5 complex and its interactions with Sgt2, Get3, and Ydj1. J Biol Chem *285*, 9962-9970.

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

Chartron, J.W., Gonzalez, G.M., and Clemons, W.M., Jr. (2011). A structural model of the Sgt2 protein and its interactions with chaperones and the Get4/Get5 complex. J Biol Chem *286*, 34325-34334.

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

Chartron, J.W., VanderVelde, D.G., and Clemons, W.M., Jr. (2012b). 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*, 1620-1632.

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

Chen, J., Zang, Y.S., and Xiu, Q. (2014). BAT3 rs1052486 and rs3117582 polymorphisms are associated with lung cancer risk: a meta-analysis. Tumour Biol *35*, 9855-9858.

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Davies, R.G., Wagstaff, K.M., McLaughlin, E.A., Loveland, K.L., and Jans, D.A. (2013). The BRCA1-binding protein BRAP2 can act as a cytoplasmic retention factor for nuclear and nuclear envelope-localizing testicular proteins. Biochim Biophys Acta *1833*, 3436-3444.

de Mendoza, A., Suga, H., and Ruiz-Trillo, I. (2010). Evolution of the MAGUK protein gene family in premetazoan lineages. BMC Evol Biol *10*, 93.

Degli-Esposti, M.A., Abraham, L.J., McCann, V., Spies, T., Christiansen, F.T., and Dawkins, R.L. (1992). Ancestral haplotypes reveal the role of the central MHC in the immunogenetics of IDDM. Immunogenetics *36*, 345-356.

Desmots, F., Russell, H.R., Michel, D., and McKinnon, P.J. (2008). Scythe regulates apoptosis-inducing factor stability during endoplasmic reticulum stress-induced apoptosis. J Biol Chem *283*, 3264-3271.

Echeverry, N., Bachmann, D., Ke, F., Strasser, A., Simon, H.U., and Kaufmann, T. (2013). Intracellular localization of the BCL-2 family member BOK and functional implications. Cell Death Differ *20*, 785-799.

Ekman, D., Bjorklund, A.K., and Elofsson, A. (2007). Quantification of the elevated rate of domain rearrangements in metazoa. J Mol Biol *372*, 1337-1348.

Eme, L., Sharpe, S.C., Brown, M.W., and Roger, A.J. (2014). On the age of eukaryotes: evaluating evidence from fossils and molecular clocks. Cold Spring Harb Perspect Biol *6*.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Engel, A., and Gaub, H.E. (2008). Structure and mechanics of membrane proteins. Annu Rev Biochem *77*, 127-148.

Etokebe, G.E., Jotanovic, Z., Mihelic, R., Mulac-Jericevic, B., Nikolic, T., Balen, S., Sestan, B., and Dembic, Z. (2015a). Susceptibility to large-joint osteoarthritis (hip and knee) is associated with BAG6 rs3117582 SNP and the VNTR polymorphism in the second exon of the FAM46A gene on chromosome 6. J Orthop Res *33*, 56-62.

Etokebe, G.E., Zienolddiny, S., Kupanovac, Z., Enersen, M., Balen, S., Flego, V., Bulat-Kardum, L., Radojcic-Badovinac, A., Skaug, V., Bakke, P.*, et al.* (2015b). Association of the FAM46A gene VNTRs and BAG6 rs3117582 SNP with non small cell lung cancer (NSCLC) in Croatian and Norwegian populations. PLoS One *10*, e0122651.

Evans, E.K., Kuwana, T., Strum, S.L., Smith, J.J., Newmeyer, D.D., and Kornbluth, S. (1997). Reaper-induced apoptosis in a vertebrate system. EMBO J *16*, 7372-7381.

Freeman, B.C., and Morimoto, R.I. (1996). The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. EMBO J *15*, 2969-2979.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc *2*, 31-34.

Gilmore, R., Blobel, G., and Walter, P. (1982). Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. J Cell Biol *95*, 463-469.

Goto, K., Tong, K.I., Ikura, J., and Okada, H. (2011). HLA-B-associated transcript 3 (Bat3/Scythe) negatively regulates Smad phosphorylation in BMP signaling. Cell Death Dis *2*, e236.

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

Gu, L., Tsuji, T., Jarboui, M.A., Yeo, G.P., Sheehy, N., Hall, W.W., and Gautier, V.W. (2011). Intermolecular masking of the HIV-1 Rev NLS by the cellular protein HIC: novel insights into the regulation of Rev nuclear import. Retrovirology *8*, 17.

Harney, S.M., Vilarino-Guell, C., Adamopoulos, I.E., Sims, A.M., Lawrence, R.W., Cardon, L.R., Newton, J.L., Meisel, C., Pointon, J.J., Darke, C.*, et al.* (2008). Fine mapping of the MHC Class III region demonstrates association of AIF1 and rheumatoid arthritis. Rheumatology (Oxford) *47*, 1761-1767.

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

Hessa, T., Sharma, A., Mariappan, M., Eshleman, H.D., Gutierrez, E., and Hegde, R.S. (2011). Protein targeting and degradation are coupled for elimination of mislocalized proteins. Nature *475*, 394-397.

Hoelz, A., Debler, E.W., and Blobel, G. (2011). The structure of the nuclear pore complex. Annu Rev Biochem *80*, 613-643.

Hsieh, Y.Y., Lin, Y.J., Chang, C.C., Chen, D.Y., Hsu, C.M., Wang, Y.K., Hsu, K.H., and Tsai, F.J. (2010). Human lymphocyte antigen B-associated transcript 2, 3, and 5 polymorphisms and haplotypes are associated with susceptibility of Kawasaki disease and coronary artery aneurysm. J Clin Lab Anal *24*, 262-268.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Kamper, N., Franken, S., Temme, S., Koch, S., Bieber, T., and Koch, N. (2012a). gamma-Interferon-regulated chaperone governs human lymphocyte antigen class II expression. FASEB J *26*, 104-116.

Kamper, N., Kessler, J., Temme, S., Wegscheid, C., Winkler, J., and Koch, N. (2012b). A novel BAT3 sequence generated by alternative RNA splicing of exon 11B displays cell type-specific expression and impacts on subcellular localization. PLoS One *7*, e35972.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772-780.

Kawahara, H., Minami, R., and Yokota, N. (2013). BAG6/BAT3: emerging roles in quality control for nascent polypeptides. J Biochem *153*, 147-160.

Kikukawa, Y., Minami, R., Shimada, M., Kobayashi, M., Tanaka, K., Yokosawa, H., and Kawahara, H. (2005). Unique proteasome subunit Xrpn10c is a specific receptor for the antiapoptotic ubiquitin-like protein Scythe. FEBS J *272*, 6373-6386.

Krenciute, G., Liu, S., Yucer, N., Shi, Y., Ortiz, P., Liu, Q., Kim, B.J., Odejimi, A.O., Leng, M., Qin, J.*, et al.* (2013). Nuclear BAG6-UBL4A-GET4 complex mediates DNA damage signaling and cell death. J Biol Chem *288*, 20547-20557.

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

Kummerfeld, S.K., and Teichmann, S.A. (2005). Relative rates of gene fusion and fission in multi-domain proteins. Trends Genet *21*, 25-30.

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

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

Kwak, J.H., Kim, S.I., Kim, J.K., and Choi, M.E. (2008). BAT3 interacts with transforming growth factor-beta (TGF-beta) receptors and enhances TGF-beta1-induced type I collagen expression in mesangial cells. J Biol Chem *283*, 19816-19825.

Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E., and Corbett, A.H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. J Biol Chem *282*, 5101-5105.

Lee, J.G., and Ye, Y. (2013). Bag6/Bat3/Scythe: a novel chaperone activity with diverse regulatory functions in protein biogenesis and degradation. Bioessays *35*, 377-385.

Leznicki, P., and High, S. (2012). SGTA antagonizes BAG6-mediated protein triage. Proc Natl Acad Sci U S A *109*, 19214-19219.

Leznicki, P., Roebuck, Q.P., Wunderley, L., Clancy, A., Krysztofinska, E.M., Isaacson, R.L., Warwicker, J., Schwappach, B., and High, S. (2013). The association of BAG6 with SGTA and tail-anchored proteins. PLoS One *8*, e59590.

Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., and Russell, R.B. (2003). Protein disorder prediction: implications for structural proteomics. Structure *11*, 1453-1459.

Liu, Y., Soetandyo, N., Lee, J.G., Liu, L., Xu, Y., Clemons, W.M., Jr., and Ye, Y. (2014). USP13 antagonizes gp78 to maintain functionality of a chaperone in ER-associated degradation. Elife *3*, e01369.

Long, P., Samnakay, P., Jenner, P., and Rose, S. (2012). A yeast two-hybrid screen reveals that osteopontin associates with MAP1A and MAP1B in addition to other proteins linked to microtubule stability, apoptosis and protein degradation in the human brain. Eur J Neurosci *36*, 2733-2742.

Luce, M.J., Akpawu, A.A., Tucunduva, D.C., Mason, S., and Scott, M.S. (2016). Extent of pre-translational regulation for the control of nucleocytoplasmic protein localization. BMC Genomics *17*, 472.

Manchen, S.T., and Hubberstey, A.V. (2001). Human Scythe contains a functional nuclear localization sequence and remains in the nucleus during staurosporine-induced apoptosis. Biochem Biophys Res Commun *287*, 1075-1082.

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

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

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

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

Melikova, M.S., Kondratov, K.A., and Kornilova, E.S. (2006). Two different stages of epidermal growth factor (EGF) receptor endocytosis are sensitive to free ubiquitin depletion produced by proteasome inhibitor MG132. Cell Biol Int *30*, 31-43.

Minami, R., Hayakawa, A., Kagawa, H., Yanagi, Y., Yokosawa, H., and Kawahara, H. (2010). BAG-6 is essential for selective elimination of defective proteasomal substrates. J Cell Biol *190*, 637-650.

Mock, J.Y., Chartron, J.W., Zaslaver, M., Xu, Y., Ye, Y., and Clemons, W.M., Jr. (2015). Bag6 complex contains a minimal tail-anchor-targeting module and a mock BAG domain. Proc Natl Acad Sci U S A *112*, 106-111.

Moore, A.D., Bjorklund, A.K., Ekman, D., Bornberg-Bauer, E., and Elofsson, A. (2008). Arrangements in the modular evolution of proteins. Trends Biochem Sci *33*, 444-451.

Nguyen, P., Bar-Sela, G., Sun, L., Bisht, K.S., Cui, H., Kohn, E., Feinberg, A.P., and Gius, D. (2008). BAT3 and SET1A form a complex with CTCFL/BORIS to modulate H3K4 histone dimethylation and gene expression. Mol Cell Biol *28*, 6720-6729.

Pai, R.K., Askew, D., Boom, W.H., and Harding, C.V. (2002). Regulation of class II MHC expression in APCs: roles of types I, III, and IV class II transactivator. J Immunol *169*, 1326-1333.

Payapilly, A., and High, S. (2014). BAG6 regulates the quality control of a polytopic ERAD substrate. J Cell Sci *127*, 2898-2909.

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

Pogge von Strandmann, E., Simhadri, V.R., von Tresckow, B., Sasse, S., Reiners, K.S., Hansen, H.P., Rothe, A., Boll, B., Simhadri, V.L., Borchmann, P.*, et al.* (2007). Human leukocyte antigen-B-associated transcript 3 is released from tumor cells and engages the NKp30 receptor on natural killer cells. Immunity *27*, 965-974.

Preta, G., and Fadeel, B. (2012a). AIF and Scythe (Bat3) regulate phosphatidylserine exposure and macrophage clearance of cells undergoing Fas (APO-1)-mediated apoptosis. PLoS One *7*, e47328.

Preta, G., and Fadeel, B. (2012b). Scythe cleavage during Fas (APO-1)-and staurosporine-mediated apoptosis. FEBS Lett *586*, 747-752.

Pumroy, R.A., and Cingolani, G. (2015). Diversification of importin-alpha isoforms in cellular trafficking and disease states. Biochem J *466*, 13-28.

Qin, L., Guo, J., Zheng, Q., and Zhang, H. (2016). BAG2 structure, function and involvement in disease. Cellular & Molecular Biology Letters *21*, 18.

Rangachari, M., Zhu, C., Sakuishi, K., Xiao, S., Karman, J., Chen, A., Angin, M., Wakeham, A., Greenfield, E.A., Sobel, R.A.*, et al.* (2012). Bat3 promotes T cell responses and autoimmunity by repressing Tim-3-mediated cell death and exhaustion. Nat Med *18*, 1394-1400.

Reiners, K.S., Topolar, D., Henke, A., Simhadri, V.R., Kessler, J., Sauer, M., Bessler, M., Hansen, H.P., Tawadros, S., Herling, M.*, et al.* (2013). Soluble ligands for NK cell receptors promote evasion of chronic lymphocytic leukemia cells from NK cell anti-tumor activity. Blood *121*, 3658-3665.

Rodrigo-Brenni, M.C., Gutierrez, E., and Hegde, R.S. (2014). Cytosolic quality control of mislocalized proteins requires RNF126 recruitment to Bag6. Mol Cell *55*, 227-237.

Rome, M.E., Chio, U.S., Rao, M., Gristick, H., and Shan, S.O. (2014). Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway. Proc Natl Acad Sci U S A *111*, E4929-4935.

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

Rouillard, A.D., Gundersen, G.W., Fernandez, N.F., Wang, Z., Monteiro, C.D., McDermott, M.G., and Ma'ayan, A. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford) *2016*.

Sasaki, T., Gan, E.C., Wakeham, A., Kornbluth, S., Mak, T.W., and Okada, H. (2007). HLA-B-associated transcript 3 (Bat3)/Scythe is essential for p300-mediated acetylation of p53. Genes Dev *21*, 848-861.

Schrodinger, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F*., et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. Cell *123*, 507-519.

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

Sebti, S., Prebois, C., Perez-Gracia, E., Bauvy, C., Desmots, F., Pirot, N., Gongora, C., Bach, A.S., Hubberstey, A.V., Palissot, V*., et al.* (2014a). BAG6/BAT3 modulates autophagy by affecting EP300/p300 intracellular localization. Autophagy *10*, 1341-1342.

Sebti, S., Prebois, C., Perez-Gracia, E., Bauvy, C., Desmots, F., Pirot, N., Gongora, C., Bach, A.S., Hubberstey, A.V., Palissot, V*., et al.* (2014b). BAT3 modulates p300-dependent acetylation of p53 and autophagy-related protein 7 (ATG7) during autophagy. Proc Natl Acad Sci U S A *111*, 4115-4120.

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

Simon, S.M., and Blobel, G. (1991). A protein-conducting channel in the endoplasmic reticulum. Cell *65*, 371-380.

Sondermann, H., Scheufler, C., Schneider, C., Hohfeld, J., Hartl, F.U., and Moarefi, I. (2001). Structure of a Bag/Hsc70 complex: convergent functional evolution of Hsp70 nucleotide exchange factors. Science *291*.

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

Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y.Y., Schwarz, D., Bomke, J., Hein, C., Lohr, F., Bernhard, F*., et al.* (2011). Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. Science *333*, 758-762.

Strasser, A. (2005). The role of BH3-only proteins in the immune system. Nat Rev Immunol *5*, 189-200.

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

Szegezdi, E., Macdonald, D.C., Ni Chonghaile, T., Gupta, S., and Samali, A. (2009). Bcl-2 family on guard at the ER. Am J Physiol Cell Physiol *296*, C941-953.

Tait, S.W., Werner, A.B., de Vries, E., and Borst, J. (2004). Mechanism of action of Drosophila Reaper in mammalian cells: Reaper globally inhibits protein synthesis and induces apoptosis independent of mitochondrial permeability. Cell Death Differ *11*, 800-811.

Takayama, S., Bimston, D.N., Matsuzawa, S., Freeman, B.C., Aime-Sempe, C., and Xie, Z. (1997). BAG-1 modulates the chaperone activity of Hsp70Hsc70. EMBO J *16*.

Takayama, S., and Reed, J.C. (2001). Molecular chaperone targeting and regulation by BAG family proteins. Nat Cell Biol *3*.

Takayama, S., Xie, Z., and Reed, J.C. (1999). An evolutionarily conserved family of Hsp70/Hsc70 molecular chaperone regulators. J Biol Chem *274*.

Tan, S., Tan, H.T., and Chung, M.C. (2008). Membrane proteins and membrane proteomics. Proteomics *8*, 3924-3932.

Thress, K., Evans, E.K., and Kornbluth, S. (1999). Reaper-induced dissociation of a Scythe-sequestered cytochrome c-releasing activity. EMBO J *18*, 5486-5493.

Thress, K., Henzel, W., Shillinglaw, W., and Kornbluth, S. (1998). Scythe: a novel reaper-binding apoptotic regulator. EMBO J *17*, 6135-6143.

Thress, K., Song, J., Morimoto, R.I., and Kornbluth, S. (2001). Reversible inhibition of Hsp70 chaperone function by Scythe and Reaper. EMBO J *20*, 1033-1041.

Tsukahara, T., Kimura, S., Ichimiya, S., Torigoe, T., Kawaguchi, S., Wada, T., Yamashita, T., and Sato, N. (2009). Scythe/BAT3 regulates apoptotic cell death induced by papillomavirus binding factor in human osteosarcoma. Cancer Sci *100*, 47-53.

Vandiedonck, C., Beaurain, G., Giraud, M., Hue-Beauvais, C., Eymard, B., Tranchant, C., Gajdos, P., Dausset, J., and Garchon, H.J. (2004). Pleiotropic effects of the 8.1 HLA haplotype in patients with autoimmune myasthenia gravis and thymus hyperplasia. Proc Natl Acad Sci U S A *101*, 15464-15469.

Vilardi, F., Lorenz, H., and Dobberstein, B. (2011). WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. J Cell Sci *124*, 1301-1307.

Vilardi, F., Stephan, M., Clancy, A., Janshoff, A., and Schwappach, B. (2014). WRB and CAML are necessary and sufficient to mediate tail-anchored protein targeting to the ER membrane. PLoS One *9*, e85033.

Vucic, D., Seshagiri, S., and Miller, L.K. (1997). Characterization of reaper- and FADD-induced apoptosis in a lepidopteran cell line. Mol Cell Biol *17*, 667-676.

Wainright, P.O., Hinkle, G., Sogin, M.L., and Stickel, S.K. (1993). Monophyletic origins of the metazoa: an evolutionary link with fungi. Science *260*, 340-342.

Wakeman, T.P., Wang, Q., Feng, J., and Wang, X.F. (2012). Bat3 facilitates H3K79 dimethylation by DOT1L and promotes DNA damage-induced 53BP1 foci at G1/G2 cell-cycle phases. EMBO J *31*, 2169-2181.

Walter, P., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein. J Cell Biol *91*, 551-556.

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

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

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

Wang, Q., Liu, Y., Soetandyo, N., Baek, K., Hegde, R., and Ye, Y. (2011b). A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. Mol Cell *42*, 758-770.

Wang, Y., Broderick, P., Webb, E., Wu, X., Vijayakrishnan, J., Matakidou, A., Qureshi, M., Dong, Q., Gu, X., Chen, W.V.*, et al.* (2008). Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet *40*, 1407-1409.

Wang, Y., Ning, H., Ren, F., Zhang, Y., Rong, Y., Wang, Y., Su, F., Cai, C., Jin, Z., Li, Z.*, et al.* (2014b). GdX/UBL4A specifically stabilizes the TC45/STAT3 association and promotes dephosphorylation of STAT3 to repress tumorigenesis. Mol Cell *53*, 752-765.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.

Winnefeld, M., Grewenig, A., Schnolzer, M., Spring, H., Knoch, T.A., Gan, E.C., Rommelaere, J., and Cziepluch, C. (2006). Human SGT interacts with Bag-6/Bat-3/Scythe and cells with reduced levels of either protein display persistence of few misaligned chromosomes and mitotic arrest. Exp Cell Res *312*, 2500-2514.

Wu, Y.H., Shih, S.F., and Lin, J.Y. (2004). Ricin triggers apoptotic morphological changes through caspase-3 cleavage of BAT3. J Biol Chem *279*, 19264-19275.

Xu, Y., Cai, M., Yang, Y., Huang, L., and Ye, Y. (2012). SGTA recognizes a noncanonical ubiquitin-like domain in the Bag6-Ubl4A-Trc35 complex to promote endoplasmic reticulum-associated degradation. Cell Rep *2*, 1633-1644.

Xu, Y., Liu, Y., Lee, J.G., and Ye, Y. (2013). A ubiquitin-like domain recruits an oligomeric chaperone to a retrotranslocation complex in endoplasmic reticulum-associated degradation. J Biol Chem *288*, 18068-18076.

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

Yong, S.T., and Wang, X.F. (2012). A novel, non-apoptotic role for Scythe/BAT3: a functional switch between the pro- and anti-proliferative roles of p21 during the cell cycle. PLoS One *7*, e38085.

Zhao, J., Wang, H., Hu, W., and Jin, Y. (2014). Effect of HLA-B-associated transcript 3 polymorphisms on lung cancer risk: a meta-analysis. Med Sci Monit *20*, 2461-2465.