Using VHHs to study the function and control of the E2 Ubiquitin-conjugating enzyme UBC6e

by

Jingjing Ling

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2018

Signature redacted

Author .....
Department of Chemistry
January 22, 2018

Signature redacted

Certified by .....
Hidde Ploegh
Professor, Boston Children's Hospital
Thesis Supervisor

Signature redacted

Accepted by .....
Robert W. Field
Haslam and Dewey Professor of Chemistry,
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a Committee of the Department of Chemistry as follows:

Professor Barbara Imperiali.. Signature redacted

Chairman, Thesis Committee

Class of 1922 Professor of Biology and Chemistry

Professor JoAnne Stubbe ... Signature redacted

Member, Thesis Committee

Novartis Professor of Chemistry

Professor Hidde Ploegh ..... Signature redacted

Thesis Supervisor

Professor, Boston Children's Hospital

Using VHHs to study the function and control of the E2 Ubiquitin-conjugating enzyme UBC6e

by

Jingjing Ling

Submitted to the Department of Chemistry
on January 22, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemistry

Abstract

Endoplasmic reticulum-associated degradation (ERAD) is essential for protein quality control both during stress and at steady state. We found that an E2 ubiquitin-conjugating enzyme UBC6e contributes to a new layer of homeostatic control, in which ERAD activity itself is regulated post-transcriptionally and independently of the unfolded protein response. Ablation of UBC6e causes up-regulation of active ERAD enhancers and increases clearance not only of terminally misfolded substrates, but also of wild-type glycoproteins that fold comparatively slowly. Tuning of ERAD component level involves a mechanism that is likely distinct from the conventional ERAD by Hrd1/SEL1L complex.

To better understand how UBC6e is controlled, we developed VHH05, which associates with UBC6e with a low nanomolar dissociation constant. VHH05 enhances enzymatic activity of UBC6e by binding to a gateway helix on UBC6e (168-186). As the phosphorylation site S184 is on this gateway loop, we propose that stress-induced phosphorylation of UBC6e allows binding of a cytosolic factor that enhances UBC6e activity similarly to VHH05.

We further characterized the binding of VHH05 to UBC6e to identify a 14mer binding epitope QADQEAKELARQIS. This epitope was translated into a 6e-tag recognized by VHH05. VHH05 specifically retrieves 6e-tagged proteins from a complex lysate mixture. VHH05 sortagged with biotin or fluorophores can stain 6e-tagged protein in an immunoblot and in FACS. VHH05 recognizes 6e-tag in the reducing environment of the cytosol and can be used to target tagged proteins to selective locations in the cell.

To study the ubiquitination machinery in live cells, we used CellSqueeze technology to deliver epitope-tagged ubiquitin (Ub) to HeLa cells. The delivered Ub molecules are readily utilized in the cell. Furthermore, kinetics of Ub incorporation is much faster in lysates than in squeezed cells. Further experiments using E2s pre-

loaded with Ub in the presence of E1 inhibition showed that E2-Ubs are also utilized faster in lysates than in squeezed cells. We hope to use this technology to identify the contribution of E2s to substrate selectivity and characterize interactions between E3s and E2-Ubs.

Thesis Supervisor: Hidde Ploegh
Title: Professor, Boston Children's Hospital

Acknowledgments

This thesis could not have been written without the support of many people.

In the first place I wish to thank my advisor Hidde Ploegh for taking me into his lab and giving me the opportunity of being around the best people. As a leader and a professor, his wisdom and insights, as well as his positive personality, have led me through the sometimes rough sea of graduate school and the journey to become a scientist. I respect, admire, and really like him a lot as a scientist and my advisor.

I am very grateful to Professor Bradley L. Pentelute for taking me for the first two years, and for teaching me the mentality of serious research and hard work.

I would also like to thank Professor Barbara Imperiali and Professor JoAnne Stubbe for their continuous support during my time at graduate school. Barbara has always given insightful suggestions for my projects, and JoAnnes critical thinking skills have taught me a lot about what a good scientist should be.

I would like to acknowledge present and past members of the Ploegh lab for being my colleagues, my friends, and my scientific partners. I would like to specially thank Dr. Ross Cheloha for fruitful collaborations, critical reading of my thesis and small chats on politics, football, and bad weather. I would like to thank Djenet Bousbaine and Dr. Florian Schmidt for their friendship and support; Dr. Andy Woodham, who helped with editing the manuscript and is always open to sharing; Dr. Tao Fang and Zeyang Li, from whom unconditional support can always be expected. And my research on UBC6e would not have been possible without Dr. Masatoshi Hagiwara, who not only took me into the world of ERAD and taught me about cell biology, but also is a friend, a mentor, and a big brother.

I would also like to thank all my collaborators, from whom I obtained help on things I am not capable of, and for sharing their expertise so that our current understanding on UBC6e is possible.

Special thanks to Mr. Robert Miller, for being an interesting friend and a very helpful lab manager.

I thank my scholarship A*STAR, for providing me the funding for undergraduate

and graduate education, so that I can enjoy the opportunities of conducting research amidst the best scientists.

I thank my family for bringing me up the way they did, and for making me who I am now.

I thank my husband Haoyue Wang for being here both in times of happiness and stress, and for helping with formatting this thesis.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Endoplasmic reticulum and protein folding

The endoplasmic reticulum (ER) is an organelle that exists in most eukaryotic cells, but not in red blood cells. It is the first compartment in the protein biosynthetic pathway that specializes in the folding and processing of secreted and membrane proteins (Caro and Palade, 1964). The ER is also responsible for lipid synthesis, membrane assembly (Fagone and Jackowski, 2009), and calcium storage (Koch, 1990). As much as 30% of the proteome is made up of proteins that are destined for the ER (Palade, 1975). In the ER, nascent proteins are modified with glycans, folded by chaperones, and oxidized by disulfide bond formation (Braakman and Hebert, 2013). Only correctly folded proteins are allowed to leave the ER for further processing by the Golgi apparatus, whereas misfolded proteins are targeted for degradation (Ellgaard and Helenius, 2003).

Most proteins destined for the ER are imported co-translationally. When a signal sequence with an exposed hydrophobic region exits from ribosome, it is recognized by signal recognition particle (SPR)(Walter and Blobel, 1981; Walter et al., 1981). Binding of the SPR halts protein synthesis until the ribosome carrying the nascent chain is docked onto the Sec61 translocon (Saraogi and Shan, 2011). Recently, Rapoport and co-workers published structures of the prokaryotic SecY translocon in complex with the ribosome during translation initiation (Park et al., 2014), and of the SecY channel in complex with a translocating substrate (Li et al., 2016) to better understand the molecular mechanisms at play during co-translational translocation (Rapoport et al., 2017).

As nascent peptides enter into the ER, many of them are modified with $Glc_3Man_9GlcNAc_2$ (Glc: glucose, Man: mannose, GlcNAc: N-acetylglucosamine) at conserved Asn-Xxx-Ser/Thr motifs (Ruiz-Canada et al., 2009). Glycans derived from $Glc_3Man_9GlcNAc_2$ are known as N-linked high mannose glycans. Attached glycans assist proteins in folding into their native conformation by providing handles for recognition by the lectin-like chaperones calnexin (CNX) and/or calreticulin (CRT) (Helenius, 1997). The presence and time-dependent processing of glycans also serve as a timer mechanism to mark proteins for degradation by the ER-associated degradation (ERAD) system if multiple rounds of folding fail (Fagioli and Sitia, 2001). In Chapter 2, I will discuss how our findings from UBC6e$^{-/-}$ cells support the mannose timer hypothesis, since we observed defects in the folding of hard-to-fold substrates in UBC6e$^{-/-}$ cells that up-regulate ERAD.

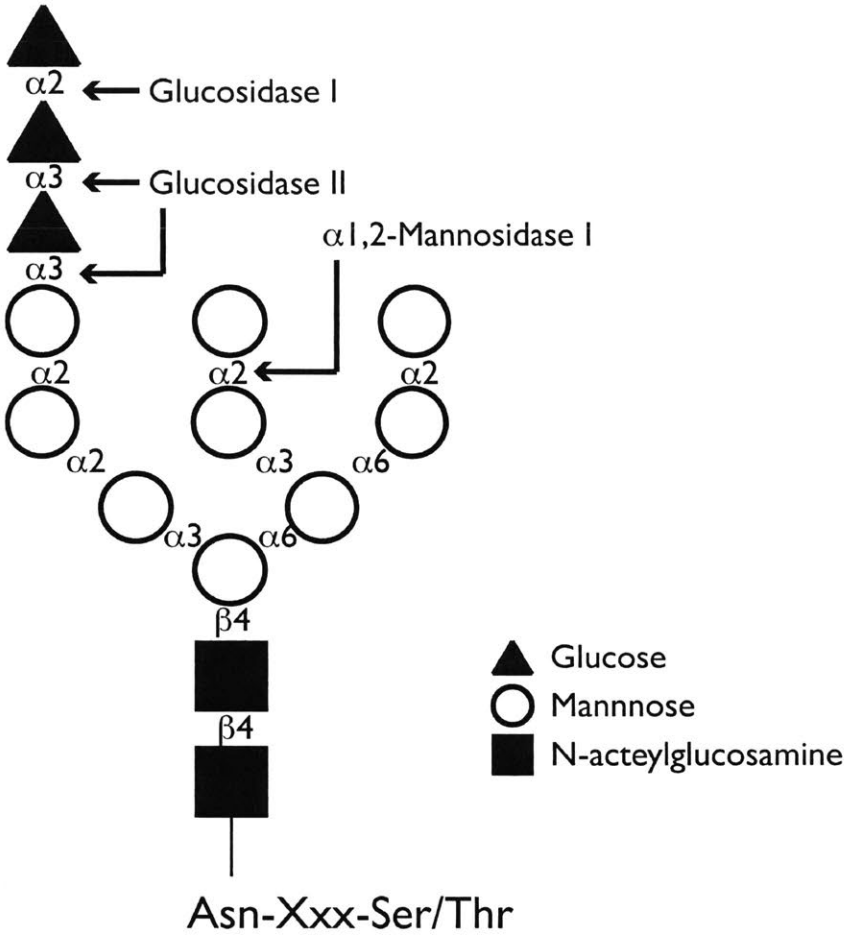


Figure 1.1. Schematic of an N-linked oligosaccharide attached to glycoproteins in the ER. The linkage type and enzymes responsible for trimming at the cleavage sites are indicated. α and β refer to the chirality of the anomeric carbon and the number represents the position of the linkage site on the distal sugar. Adapted from (Ellgaard and Frickel, 2003).

Glycoproteins can undergo multiple rounds of folding in the CNX/CRT cycle. After $Glc_3Man_9GlcNAc_2$ is attached to the N-glycan site, the two distal glucoses on the core glycan are quickly trimmed off. The monoglucosylated substrates are recognized by CNX and CRT. In addition to recognizing glycans, CNX and CRT also bind to hydrophobic regions on the polypeptide to assist in folding by preventing aggregation (Ware et al., 1995; Williams, 2006). ERp57, a glycoprotein-specific thiol-disulfide oxidoreductase that associates with CNX and CRT, helps to establish disulfide bonds in newly synthesized proteins (Ellgaard and Frickel, 2003; Zapun et al., 1998). While glycoprotein substrates are associated with CNX/CRT, glucosidase II (GS-II) trims away the innermost glucose, leading to substrate release from CNX/CRT. At this stage, if the substrates are properly folded, they are exported from the ER to the Golgi apparatus for further processing. If the substrates are not folded, they can be reglycosylated by UDP-glucose glycoprotein glucosyltransferase 1 (UGGT1), which binds to the exposed hydrophobic patches on glycoprotein substrates and adds a glucose molecule to allow re-binding to CNX/CRT (Caramelo and Parodi, 2008; Ritter and Helenius, 2000). Once a protein has exhausted the CNX cycle yet failed to fold

successfully, the terminal mannose can be trimmed by ER-resident mannosidases, such as ER mannosidase I, to yield a $Man_{5-6}GlcNAc_2$ labeled protein, which is then handed off to ER degradation enhancing alpha-mannosidase-like (EDEM) family proteins for degradation by the ERAD system (Avezov et al., 2008; Oda et al., 2003). EDEM family proteins can bind to terminally misfolded substrates on CNX and direct them to ERAD, however, there is no conclusive evidence that EDEM family proteins are enzymatically active.

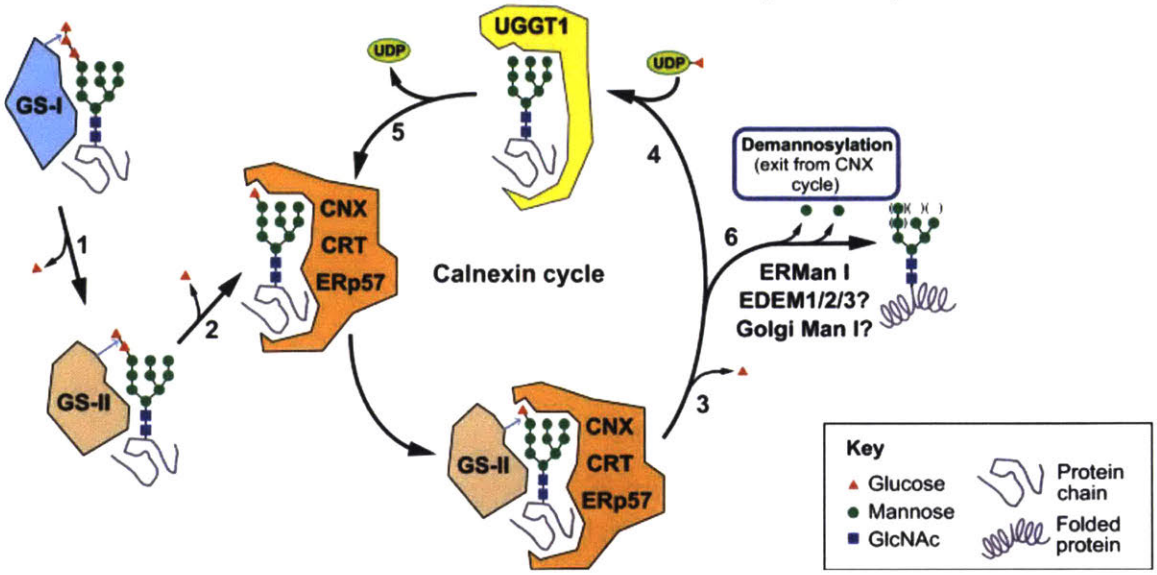


Figure 1.2. Schematic of the CNX/CRT/ERp57 folding cycle. GS-I: glucosidase I. GS-II: glucosidase II. ERMan I: ER mannosidase I. UGGT1: UDP-glucose glycoprotein glucosyltransferase 1. (1) GS-I trims off first the glucose from $Glc_3Man_9GlcNAc_2$. (2) GS-II trims off the second glucose, and the substrate binds to CNX/CRT/ERp57 for folding. (3) GS-II trims off the innermost glucose, leading to substrate release from CNX/CRT. (4) UGGT1 can add a glucose to $Man_9GlcNAc_2$ on unfolded substrates. (5) Substrates with $Glc_3Man_9GlcNAc_2$ bind CNX/CRT/ERp57 for another round of folding. (6) Peptide substrates leave the CNX cycle through mannose trimming. Mannose can be trimmed by Golgi-resident mannosidases, ER-resident mannosidases, and possibly by EDEM family proteins. Adapted from (Ferris et al., 2014).

In addition to the CNX/CRT cycle, immunoglobulin binding protein (Bip), an ER-resident homolog of Hsp70, also assists in protein folding by binding to hydrophobic regions of proteins, but it does so independently of glycosylation status (Okuda-Shimizu and Hendershot, 2007). ER DnaJ-like (ERdj) family proteins 1-7 have also been reported to help with both productive folding and directing terminally misfolded proteins to degradation (Hagiwara et al., 2011; Lai et al., 2012; Meunier et al., 2002; Shen and Hendershot, 2005; Ushioda et al., 2008).

1.2 Endoplasmic reticulum associated degradation

Due to errors in translation and time-dependent accumulation of protein damage, proteins that are terminally misfolded or damaged must be degraded. These proteins are recognized

by the endoplasmic reticulum-associated degradation (ERAD) system and are retrotranslocated to the cytosol to be degraded by the ubiquitin-proteasome system (UPS) or by autophagy. The ERAD system comprises chaperones that recognize proteins destined for degradation, the retrotranslocation complex which transports unfolded substrates to the cytosol, and the UPS, which covalently modifies substrates with ubiquitin and targets them for degradation. Recent studies have also shown that the autophagy pathway can be activated when ERAD fails to remove misfolded and aggregated substrates (Houck et al., 2014; Kario et al., 2011; Kroeger et al., 2009).

During my PhD studies I have focused on the ERAD complex centered on the HMG-CoA reductase degradation 1 protein (HRD1) . This complex includes lectin chaperones (EDEM1, OS-9, XTP3-B, and SEL1L), the retrotranslocon (HRD1 and SEL1L), components of the ubiquitin conjugation machinery (UBC6e and HRD1), proteasome components (p97 AAA ATPase and 26S proteasome), as well as accessory proteins (Figure 1.3). The key components in this complex and their functions are discussed below.

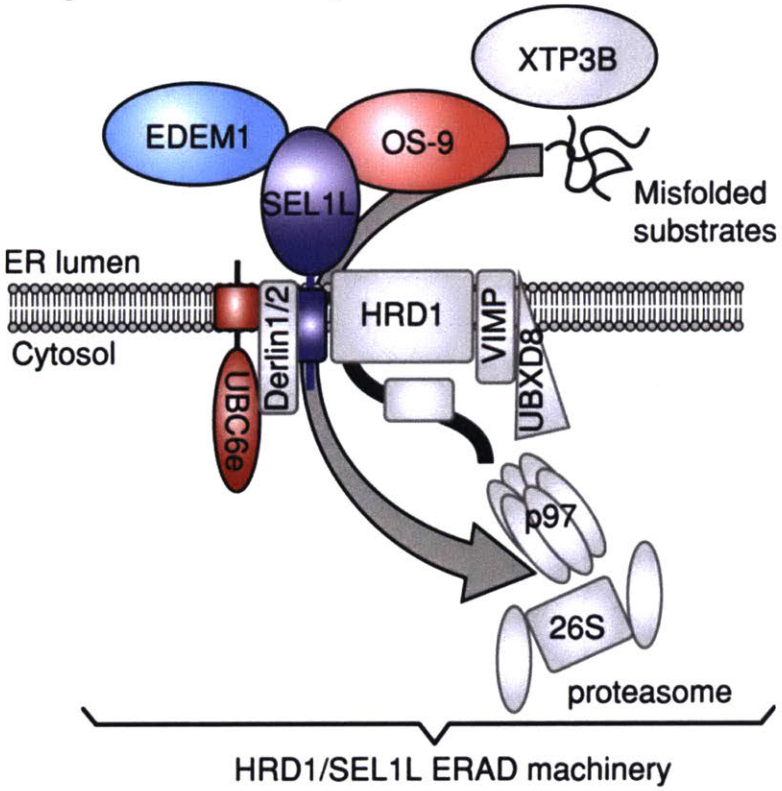


Figure 1.3. Schematic of the role of members of the HRD1/SEL1L ERAD complex. Adapted from (Hagiwara et al., 2016).

Terminally misfolded glycoproteins can be recognized via their glycans. EDEM family proteins 1/2/3 are likely responsible for receiving substrates from CNX. Of these three, EDEM1 is the best studied. Overexpression of EDEM1 increases ERAD activity, and knockdown of EDEM1 delays ERAD, as measured by the rates of degradation of ERAD substrates, such as a mutant form of α 1-antitrypsin (A1AT) and BACE457, which is a

pancreatic isoform of human β-secretase (Hosokawa et al., 2001; Molinari et al., 2003). Overexpressed EDEM1 also interacts with CNX, and accelerates substrate release (Molinari et al., 2003; Oda et al., 2003). EDEM2 and EDEM3 were later reported to enhance ERAD as well (Hirao et al., 2006; Olivari et al., 2005). These EDEM family proteins have mannosidase-like domains, but they do not require enzymatic activity to be functional (Mast et al., 2005; Olivari et al., 2006). Even though overexpression of EDEM family proteins accelerates demannosylation of ERAD substrates *in vivo*, their mannosidase activities have yet to be demonstrated *in vitro.* It is thus not clear whether EDEMs recruit other ER-resident mannosidases to trim off the mannose molecules on glycoprotein substrates, or are enzymatically active themselves.

After the glycans on ERAD substrates are trimmed to $Man_7GlcNAc_2$, they become substrates for the lectins osteosarcoma 9 (OS-9) and XTP3-transactivated gene B protein (XTP3-B)(Groisman et al., 2011; Hosokawa et al., 2009; Satoh et al., 2010; Yamaguchi et al., 2010). OS-9 was identified as an adaptor protein that binds to both EDEM1 and suppressor/enhancer of Lin-12-like (SEL1L), and thereby accelerates ERAD by facilitating substrate transfer from EDEM to SEL1L, which is a core component of an ERAD complex referred to as the HRD1/SEL1L complex (Christianson et al., 2008; Hosokawa et al., 2009; Mueller et al., 2008). On the other hand, XTP3-B also binds glycoprotein substrates carrying a $Man_7GlcNAc_2$ motif, but inhibits ERAD when overexpressed. XTP3-B is therefore considered to be a negative regulator of ERAD (Fujimori et al., 2013; Hosokawa et al., 2008).

EDEM family proteins and OS-9 are thought to relay the terminally misfolded glycan substrates from the CNX cycle to the retrotranslocon, where the substrates are then dislocated to the cytosol and tagged with ubiquitin for degradation. While the translocon for co-translational import into the ER has been defined for about 30 years (Deshaies et al., 1991), much less is known regarding the retrotranslocon(s) responsible for transporting misfolded polypeptides from the ER back to the cytosol. Many studies in this field have thus studied ERAD based on another core component, the E3 ubiquitin ligase, which is responsible for attaching ubiquitins to substrates.

As the name suggested, HRD1 is the central E3 ligase in the HRD1/SEL1L ERAD complex. HRD1 forms a stoichiometric complex with its essential adaptor protein SEL1L (Iida et al., 2011; Sun et al., 2014). The HRD1/SEL1L ERAD complex is involved in the degradation of a variety of substrates, including both endogenous proteins such as 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-R) (Hampton et al., 1996) and MHC class I molecule heavy chain (Burr et al., 2011; Lilley and Ploegh, 2005; Mueller et al., 2006), as well as ERAD substrates that are mutant proteins unable to achieve their native fold (Bordallo et al., 1998). The HRD1 homolog in *Saccharomyces cerevisiae* HRD1p also forms a complex with SEL1L homolog HRD3. This HRD1p/HRD3 ERAD complex is considered to be primarily responsible for degrading substrates from the ER lumen (Bays et al., 2000), with some involvement in the degradation of membrane proteins (Bordallo et al., 1998). Other known ER-associated E3s include Doa10 (yeast)/Teb4 (Foresti et al., 2013), gp78 (Fang et al., 2001), and TEM129 (van den Boomen et al., 2014).

HRD1 contains a N-terminal domain that spans the ER membrane multiple times, and a C-terminal E3 really interesting new gene (RING) domain responsible for ubiquitin ligase activity (Figure 1.4). It was only recently found that HRD1 is likely responsible for forming an aqueous channel to act as the retrotranslocon in the context of this complex. Recent cryo-EM structural analysis of *Saccharomyces cerevisiae* HRD1p in complex with HRD3 showed 8 transmembrane helices (Figure 1.4 and 1.5)(Schoebel et al., 2017). Rapoport *et. al* recently reported that autoubiquitination of HRD1p triggers retrotranslocation of substrates (Baldridge and Rapoport, 2016) and that HRD1p forms an aqueous channel in the ER membrane (Figure 1.6) (Schoebel et al., 2017), providing support for HRD1's role as a translocon in retrotranslocating unfolded substrates from the ER back to the cytosol.

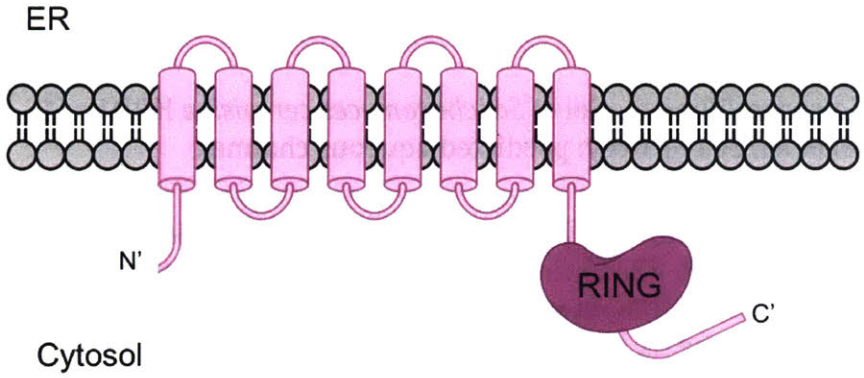


Figure 1.4. Topology of HRD1. The N-terminal domain spans the ER membrane 8 times and HRD1 contains a RING E3 ligase domain. The RING domain is located in the cytoplasm.

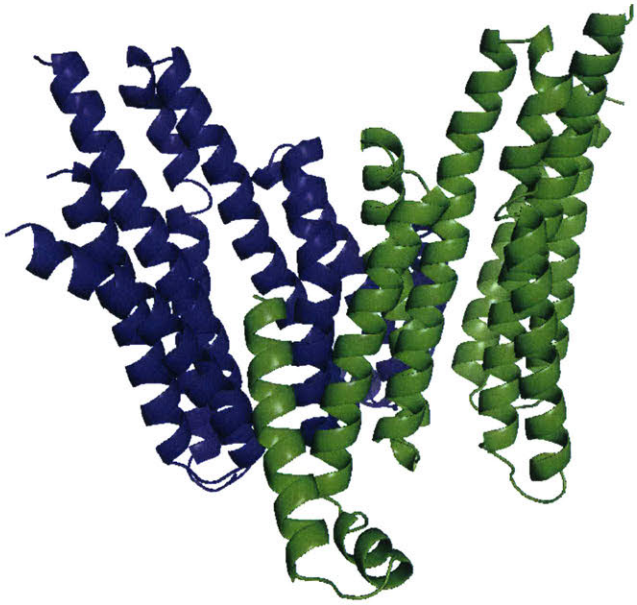


Figure 1.5. Transmembrane domain of *Saccharomyces cerevisiae* HRD1p (1-407) published in (Schoebel et al., 2017). PDB: 5V6P. Green and blue: helices from two different HRD1 molecules.

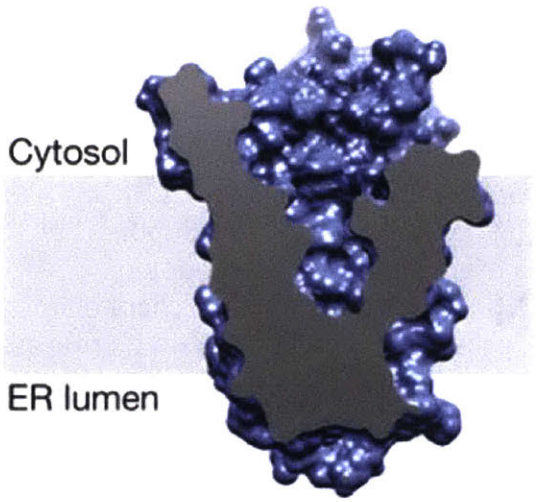


Figure 1.6. Cut-away view of a space-filling model of *Saccharomyces cerevisiae* HRD1p (1-407) as published in (Schoebel et al., 2017). Grey: predicted aqueous channel.

HRD1 is a RING-type E3 (Bordallo and Wolf, 1999). RING-type E3s lack catalytic cysteines, and instead act by forming a complex with E2s to catalyze the ubiquitin (Ub) transfer from E2 directly to substrates. The HRD1 RING domain was shown to be active *in vitro* by enhancing substrate ubiquitination by the E2 ubiquitin conjugating enzyme UBC7 (Kikkert et al., 2004). However, our lab immunoprecipitated overexpressed HA-SEL1L from HeLa cells, and identified only the E2 UBC6e, but not UBC7, in association with the HRD1/SEL1L complex (Mueller et al., 2008). We hypothesize that HRD1 works with multiple E2s with different affinities and for different ubiquitination reactions. UBC7 dimerizes *in vitro* and builds Lys48-linked ubiquitin chains on its active site in the absence of E3 or substrates (Liu et al., 2014). This Lys48-linked ubiquitin chain elongation activity by UBC7 is also reported *in vivo* (Bocik et al., 2011; von Delbruck et al., 2016). Thus, UBC7 might be the E2 that is primarily responsible for elongating ubiquitin chains on substrates for recognition by the proteasome system.

Derlin family proteins are another set of essential components in the HRD1/SEL1L complex. Degradation from the ER protein 1 (Der1p) was identified as an essential component of ERAD in *Saccharomyces cerevisiae* (Knop et al., 1996). The mammalian homologs of Der1p are named Der1-like domain family proteins (Derlins), and there are three homologs in mammalian cells. Genetic deletion of Derlin-1 and Derlin-2 in mice are both embryonic lethal, while Derlin-3-deficient mice appear normal (Dougan et al., 2011; Eura et al., 2012). Derlin-1 was first found to be essential for MHC-I degradation mediated by the human cytomegalovirus-encoded glycoprotein US11 (Lilley and Ploegh, 2004; Ye et al., 2004). Derlin-1 was found to interact with the p97 AAA ATPase, and was thus thought to mediate a step between translocation and proteasomal degradation (Ye et al., 2004). US11 initiates MHC-I degradation by bringing the substrate in proximity to Derlin-1 (Cho et al., 2013; Lilley and Ploegh, 2004). Derlin-1 is also involved in the degradation of various membrane proteins, such as the epithelial Na^+ channel (ENaC), lipidated apolipoprotein B-100 (ApoB), and the cystic fibrosis transmembrane conductance regulator (CFTR) (Sun et

al., 2006; Suzuki et al., 2012; You et al., 2017). Derlin-1 is also thought to recruit the E3 ligase transmembrane protein of ER membrane 129 (TMEM129) to the ER, which functions in addition to the HRD1/SEL1L ERAD complex (van den Boomen et al., 2014). Derlin-2 functions similarly to Derlin-1 in ERAD, where Derlin-2 binds to both HRD1/SEL1L and p97 AAA ATPase (Lilley and Ploegh, 2005). In Derlin-2-deficient mouse tissues, constitutive upregulation of ER chaperones and activation of the unfolded protein response (UPR) was observed (Dougan et al., 2011).

Other components in the HRD1/SEL1L complex include VCP (p97)-interacting membrane protein (VIMP), which harbors reductase activity that reduces the disulfide bonds on substrates as they are translocated back to the cytosol (Christensen et al., 2012). UBX domain containing protein 8 (UBXD8) was also co-immunoprecipitated with SEL1L (Mueller et al., 2008), and UBXD8 was later found to be involved in lipid homeostasis (Loregger et al., 2017; Olzmann et al., 2013; Suzuki et al., 2012). There are also many other proteins thought to be involved in ERAD that were not recovered during the anti-SEL1L immunoprecipitation. One example is homocysteine-induced ER protein (HERP), which is thought to be another adaptor protein that associates with misfolded proteins, especially non-glycosylated substrates, and facilitates their degradation (Leitman et al., 2014; Okuda-Shimizu and Hendershot, 2007). Many of these proteins are important for ERAD , but with the underlying molecular mechanisms still to be elucidated.

1.3 The ubiquitin-conjugating enzyme UBC6e

My thesis focuses on ubiquitin-conjugating enzyme 6e (UBC6e), which is also called ubiquitin E2 J1 (UBE2J1). We focused on this protein because it was the only E2 that co-immunoprecipitated with SEL1L, which is a core member of the SEL1L/HRD1 ERAD complex, indicating that it might be the primary E2 in the SEL1L/HRD1 complex (Mueller et al., 2006). We are interested in the function of this E2 in ERAD. In my thesis, I will discuss our knowledge of the function of UBC6e as ascertained through studying mouse embryonic fibroblasts (MEFs) from UBC6e-deficient mice (Chapter 2) and through perturbation of UBC6e activity with a the variable fragment of a camelid-derived heavy-chain antibody (VHH) that specifically binds to UBC6e (VHH05) (Chapter 3).

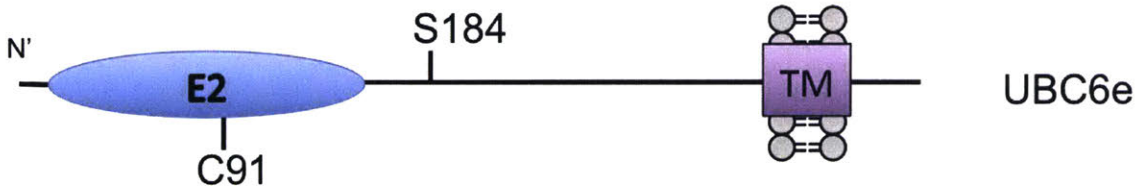


Figure 1.7 Schematic of the linear sequence of UBC6e. S184: phosphorylation site. TM: transmembrane helix for tail-anchor to the ER. C91: active site.

In *Saccharomyces cerevisiae*, there are two E2s associated with the ER: UBC6p and UBC7p. UBC7p and its mammalian homolog UBC7 associate with the ERAD complex through its ER tail-anchored adaptor proteins CUE1p or CUE domain (Chen et al., 2006; Kostova et al.,

2009). UBC6p is an ER tail-anchored E2, where a C-terminal transmembrane domain anchors its E2 enzyme domain to the ER. UBC6 and UBC6e are two mammalian homologs of UBC6p. Both of these homologs are ER tail-anchored and involved in ERAD (Lenk et al., 2002). The ER tail anchor is vital to the proper localization and function of UBC6e (Claessen et al., 2010). Enterovirus 71 cleaves UBC6e from its ER tail-anchor to abolish its activity in ERAD, in order to attenuate ERAD to benefit its viral replication (Wang et al., 2017).

UBC6e co-immunoprecipitates with SEL1L (Mueller et al., 2008). Our lab further established that UBC6e is indeed an E2 that functions in the HRD1/SEL1L ERAD complex by showing its involvement in US11-mediated dislocation of MHC I molecules (Mueller et al., 2008). The involvement of UBC6e in the ERAD of MHC I was confirmed by others (Burr et al., 2011; Burr et al., 2013). UBC6e has also been reported to complex with a different E3 ligase RING finger protein 5 (RNF5) to contribute to the degradation of CFTR (Younger et al., 2006).

UBC6e is phosphorylated at Ser184 in response to ER stress (Oh et al., 2006) or inhibition of protein synthesis (Menon et al., 2013). The function of UBC6e phosphorylation is still not fully understood, but was proposed to be important for the recovery of cells from stress (Elangovan et al., 2017). Our lab generated UBC6e-deficient mice, and found that male UBC6e$^{-/-}$ mice suffer from deficiencies in sperm development (Koenig et al., 2014). The molecular mechanism behind this deficiency is yet to be understood.

Though both are involved in ERAD, the isoelectric points of UBC6 and UBC6e are very different. Based on the primary sequences, UBC6e is predicted to be positively charged at pH 7.4 in almost all species examined, while UBC6 is predicted to be negatively charged at pH 7.4 (Table 1.1). Such charge difference could contribute to their divergence in cellular functions, as the different charges will inevitably lead to different interactions with other cellular proteins. However, the details of how these two UBC6 homologs function in ERAD is still an area where little is known.

Table 1.1. Isoelectric points (pIs) for UBC6e and UBC6 homologs in different organisms.

Protein	Organism	pI	gi number
UBC6e	*Homo sapiens*	6.26	gi|37577122
UBC6e	*Pan troglodytes*	6.26	gi|114608464
UBC6e	*Macaca mulatta*	6.37	gi|109072028
UBC6e	*Canis lupus familiaris*	6.15	gi|545519892
UBC6e	*Bos taurus*	6.15	gi|329664082
UBC6e	*Mus musculus*	6.51	gi|31980960
UBC6e	*Rattus norvegicus*	6.26	gi|157817496
UBC6e	*Gallus gallus*	8.49	gi|45382103
UBC6e	*Danio rerio*	5.34	gi|47550809
UBC6e	*Caenorhabditis elegans*	5.05	gi|115532480
UBC6e	*Xenopus tropicalis*	5.81	gi|58332036
UBC6	*Homo sapiens*	8.6	gi|37577124
UBC6	*Pan troglodytes*	8.91	gi|410032108
UBC6	*Macaca mulatta*	8.8	gi|297279166
UBC6	*Canis lupus familiaris*	8.6	gi|73956479
UBC6	*Bos taurus*	8.6	gi|84000055
UBC6	*Mus musculus*	9.07	gi|85662413
UBC6	*Rattus norvegicus*	8.8	gi|56090425
UBC6	*Gallus gallus*	8.59	gi|50759239
UBC6	*Danio rerio*	8.81	gi|148229951
UBC6	*Drosophila melanogaster*	7.6	gi|21358599
UBC6	*Caenorhabditis elegans*	8.95	gi|32563946
UBC6	*Xenopus tropicalis*	8.6	gi|163915229
UBC6p	*Saccharomyces cerevisiae*	5.81	gi|398364639

Our lab aims to study how UBC6e functions in ERAD, and I will discuss how we have used tools developed in our lab to accomplish this in the next two chapters.

1.4 VHHs and their uses in biology

Antibodies are useful tools to study the functions of proteins of interest. Depending on how an antibody interacts with its target, it can activate or inhibit the activity of the target. Antibodies against cell surface receptors have been used to modulate these receptors to study their functions (Excler et al., 2014; Wajant, 2015; Wilkinson, 2016). We aimed to

generate functional antibodies to perturb the activity of our protein of interest, UBC6e, in order to study its role in the ERAD process. However, conventional antibodies are too large to penetrate the cell membrane, and they rely on intramolecular and intermolecular disulfide bonds to maintain their binding affinities, limiting their intracellular applications and making impossible the assembly of a full immunoglobulin in the reducing environment of the cytosol. Camelids produce a class of antibodies that lack light chains (Hamers-Casterman et al., 1993). The single-domain variable fragments of these heavy chain-only antibodies are termed VHHs or nanobodies. They are the smallest immunoglobulin domains that retain antigen-binding properties (Figure 1.8). VHHs retain the immunoglobulin fold shared by antibodies, using three hypervariable loops, CDR1, CDR2 and CDR3, to bind to their targets (Figure 1.9). Many VHHs bind to their targets with affinities similar to conventional full-size antibodies, but possess other properties superior to them. Therefore, VHHs are attractive tools for use in biological research and therapeutics development.

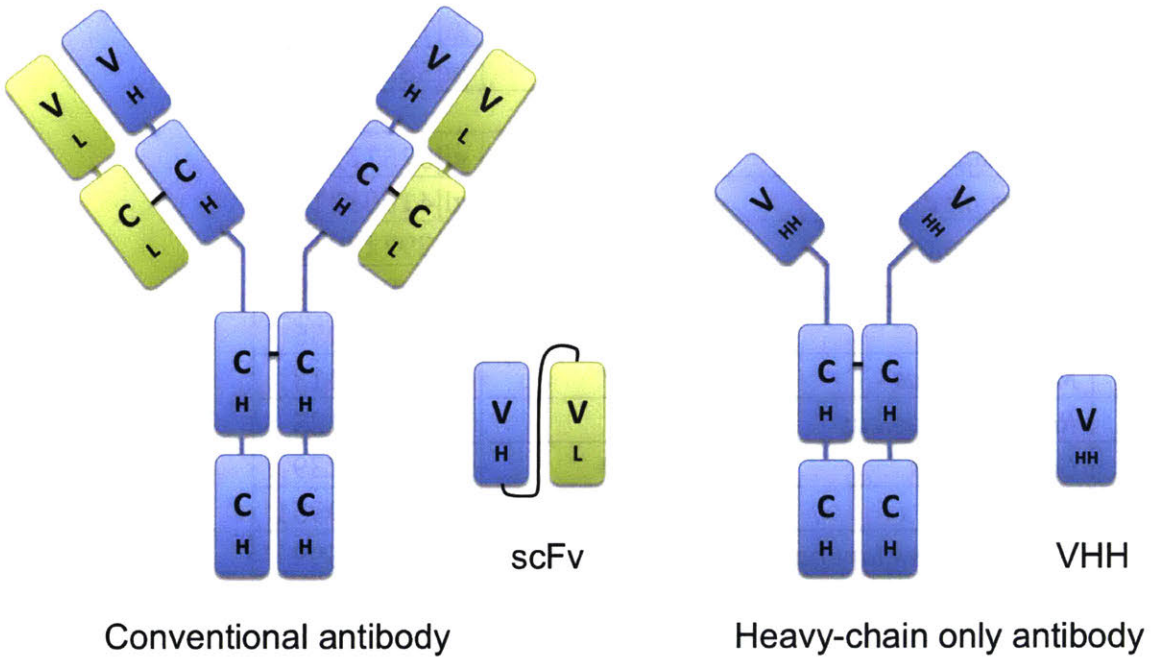


Figure 1.8. Comparison between conventional antibodies and heavy-chain only antibodies. scFvs (single chain variable fragments) are the smallest target-binding fragments obtainable from conventional antibodies, and VHHs are the binding domains of heavy-chain only antibodies. scFvs are typically ~25 kD in size. VHHs are typically 12-15 kD in size.

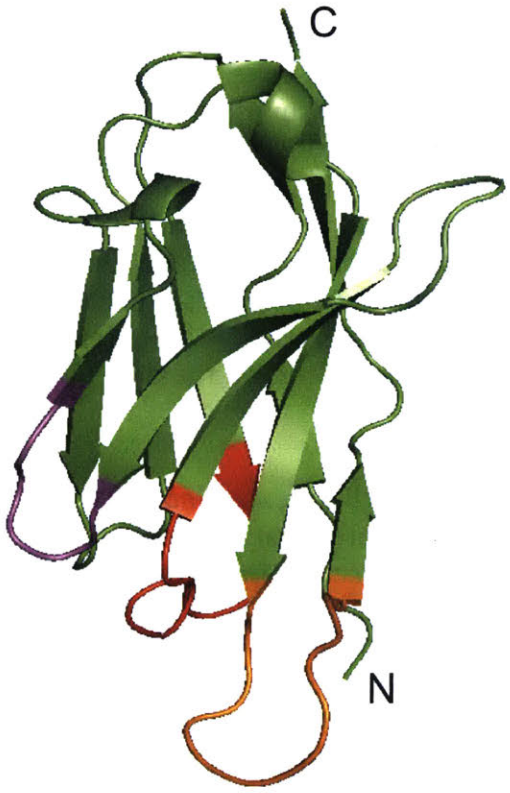


Figure 1.9. Typical structure of a VHH. Green: a typical immunoglobulin fold consists of 4 β -sheet framework regions; red: CDR1; purple: CDR2; orange: CDR3. Structure adapted from a torsin/VHH complex from PDB 5J1S.

One major advantage of VHHs over conventional antibodies is that they can be expressed with yields up to 100-200 mg/L in *E. coli* (Xu et al., 2017; Zarschler et al., 2013). Their small size, high solubility, and lack of a requirement for disulfide bonds all contribute to the observed high yield. The high expression yield in a low-cost organism helps to drive down costs when VHHs are used in applications typically preformed with conventional antibodies, such as immunoprecipitation and immunoblotting.

Additionally, most VHHs retain their target specificity in the reducing environment of the cytosol. While there are cysteine pairs that form disulfide bonds in VHHs, these disulfide bonds are not necessary to maintain VHH structure or target affinity (Harmsen and De Haard, 2007). The extra disulfide bonds marginally increase thermostability (Akazawa-Ogawa et al., 2016; Zabetakis et al., 2014; Zarschler et al., 2013), but labeling these cysteines with maleimides causes neither folding problems nor affects their binding affinities or specificities (Pleiner et al., 2015). A VHH expressed intracellularly in yeast retained its binding properties as it neutralized the toxicity of its target, 15-acetyl-deoxynivalenol (15-AcDON) (Doyle et al., 2009). Researchers rapidly discovered that intracellular VHHs are highly useful tools to mechanistically characterize cellular processes. For example, viral life cycles have been characterized by inhibiting steps in the infection

process using VHHs (Ashour et al., 2015; Hanke et al., 2016; Hanke et al., 2017; Inoue et al., 2007; Serruys et al., 2009). The properties of metastatic cancer cells can also be characterized using similar approaches (Inoue et al., 2007). VHHs have also been used to modulate the enzymatic activity of their targets, for example Huntingtin-associated yeast-interacting protein E (HypE) activity has been modulated with a VHH (Truttmann et al., 2015). Even VHHs that do not directly affect enzymatic activity have proven to be useful tools to study the cellular localization of their targets when the VHHs are expressed as a genetic fusion with green fluorescent protein (GFP), which avoids possible artifacts from fixation or overexpression of the proteins of interest (Schmidt et al., 2016; Truttmann et al., 2015). Therefore, VHHs have great potential in assisting biologists to better understand proteins of interest by either perturbing the normal functions of their targets or allowing for visualization of endogenous targets in living cells.

VHHs are popular also among crystallographers. Many VHHs have extended CDR3s that bind deep into a pocket on their targets and help stabilize protein conformations, which can assist with crystallization (Spiess et al., 2015). This is especially important for the crystallization of enzymes, where different VHHs could help lock enzymes in specific states, and even help with understanding mechanisms of inhibition (Ingram et al., 2015; Rudolph et al., 2016; Schmitz et al., 2013; Spiess et al., 2015). As multiple attempts to crystallize UBC6e did not yield any structural data, we hope to use VHHs that we have generated that recognize Ubc6e to assist in crystallization, and to understand how these VHHs might affect UBC6e activity.

VHHs also have superior properties as imaging and delivery agents compared to conventional antibodies. The small size of VHHs makes them more tissue-permeable. Our lab, among others, has conjugated radioactive labels to VHHs to image specific immune cell types *in vivo* by positron emission tomography (PET), for example during cancer progression, where it was found that VHHs could infiltrate and label immune cells in tumors and in the lymphatics, highlighting their ability to penetrate different tissues (Bannas et al., 2015; Oliveira et al., 2012; Rashidian et al., 2015a; Rashidian et al., 2015b; Rashidian et al., 2016). This deep tissue penetration also makes VHHs useful for targeted delivery of drugs to selected sites. These target sites could be tumors or injured tissues, and the cargo could range from small molecules to immunotoxins, liposomes, and nanoparticles (Cortez-Retamozo et al., 2004; Fang et al., 2016; Hu et al., 2017; Li et al., 2017; Pruszynski et al., 2013; Yu et al., 2017). The targeted delivery of peptide antigens to dendritic cells in mice using VHHs elicits strong immune responses against the covalently attached peptide epitope *in vivo* (Duarte et al., 2016; Fang et al., 2017; Woodham AW et al., 2018). As researchers develop more VHHs that target biomarkers on different cells, we will have more tools for the targeted delivery of selected molecules.

The single-domain nature of VHHs greatly reduces the possibility of complications when multiple antigen-targeting motifs need to be genetically fused for improved therapeutic properties. Mispairing of V_H and V_L domains complicates the development of bispecific antibodies and bispecific scFvs, as the pairing of non-cognate V_Hs and V_Ls could potentially lead to the formation of new binding sites that have off-target specificities (Spiess et al.,

2015). On the other hand, VHHs fused to other targeting domains retain their respective binding properties without generating additional binding sites. Case in point, our lab has fused an anti-GFP VHH (named VHH enhancer as it enhances GFP fluorescence upon binding) to an anti-MHCII VHH, and showed that the bivalent antibody could be used to stain MHCII-expressing cells upon GFP injection into mice (Witte et al., 2012). Other groups have fused an anti-ZnO VHH to an anti-GFP VHH to immobilize GFP onto ZnO surfaces (Hattori et al., 2010).

Most VHHs are identified using phage display libraries encoding the VHH sequences amplified from lymphocyte cDNAs collected from immunized camelids such as alpacas (Figure 1.10) (Pardon et al., 2014). In contrast, the process of identifying scFvs from immunized mice or rabbits is more challenging due to the mispairing between V_H and V_L chains that limits library coverage and reduces the success rate (Harmsen and De Haard, 2007). VHHs identified from these phage library screens are monoclonal and their DNA sequences are ready for subcloning with or without modifications, such as a tag of choice. Our lab prefers to equip VHHs with a C-terminal LPETGG sortase recognition motif, which allows us to site-specifically modify recombinantly produced VHHs with triglycine-equipped nucleophiles, including, but not limited to, biotin, fluorophores, radiolabels, DNA, or peptides using previously developed in-house sortagging methods (Popp et al., 2007).

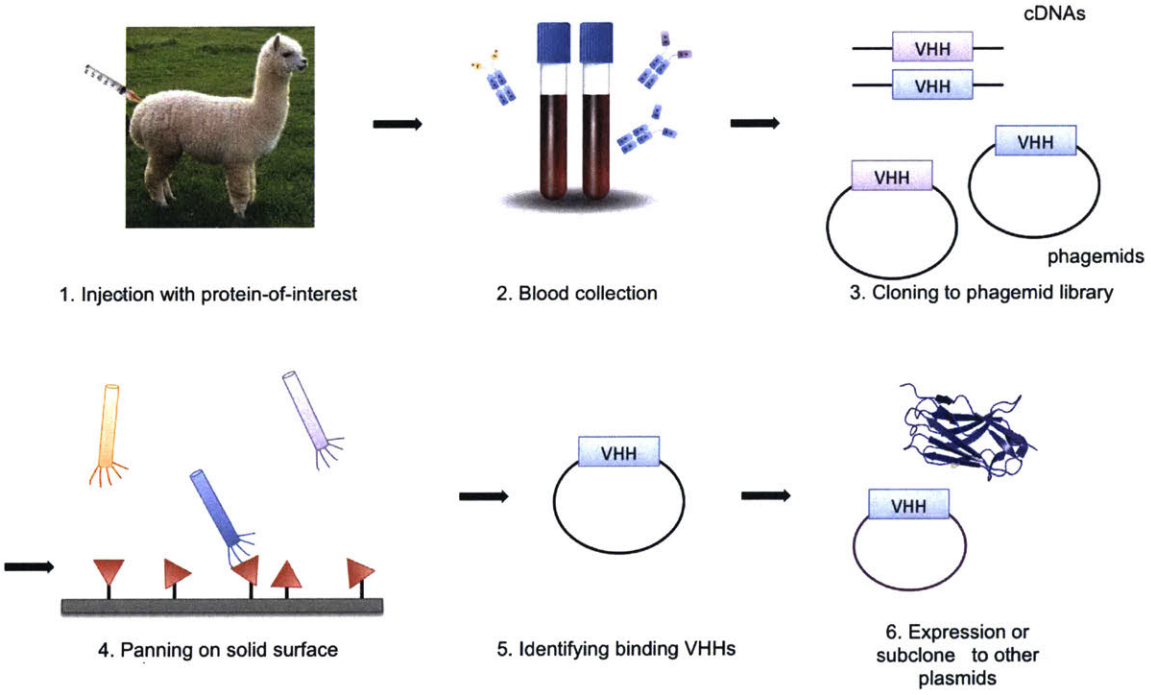


Figure 1.10. Commonly used protocol for identifying VHHs using a phage display library. (1) A camelid is immunized with a purified protein of interest. (2) Blood containing lymphocytes coding for heavy chain-only antibodies are collected. (3) cDNAs are retro-transcribed from lymphocytes, and primers for VHH-coding regions are used to amplify VHH genes and subclone VHHs to a phagemid library. (4) Antigens are immobilized to a solid surface and phage particles are panned. (5) Enriched clones are sequenced for VHH

genes. (6) VHHs are expressed in *E. coli*, subcloned as genetic fusions to other proteins, or expression in mammalian cells.

There have been many efforts to engineer artificial scaffolds to generate protein binders, such as affibodies, fibonectin domains, and designed ankyrin repeat proteins (DARPins) (Frejd and Kim, 2017; Koide and Koide, 2007; Stumpp and Amstutz, 2007). Most of these artificial scaffolds rely on large libraries to ensure identification of potential binders and error-prone polymerases to introduce mutations that improve binding affinities and specificities. On the other hand, VHHs are generated by the immune systems of living animals, where B cells undergo somatic mutations and selection processes in germinal centers (Grace and Papavasiliou, 2007; Victora and Nussenzweig, 2012). In this process, B cells can accumulate point mutations or even add or trim nucleotides from the CDR-coding regions to generate high levels of VHH-domain diversity. B cells also compete for scarce antigens for survival as injected antigens are degraded in the animals over time, which serves as a further selection pressure for increasing antigen affinity. The process of affinity maturation *in vivo*, driven by natural selection, is considered a key factor that leads to the superior performance of VHHs when compared to binders generated from artificial scaffolds.

1.5 References

Akazawa-Ogawa, Y., Uegaki, K., and Hagihara, Y. (2016). The role of intra-domain disulfide bonds in heat-induced irreversible denaturation of camelid single domain VHH antibodies. Journal of Biochemistry *159*, 111-121.

Ashour, J., Schmidt, F.I., Hanke, L., Cragnolini, J., Cavallari, M., Altenburg, A., Brewer, R., Ingram, J., Shoemaker, C., and Ploegh, H.L. (2015). Intracellular expression of camelid single-domain antibodies specific for influenza virus nucleoprotein uncovers distinct features of its nuclear localization. Journal of virology *89*, 2792-2800.

Avezov, E., Frenkel, Z., Ehrlich, M., Herscovics, A., and Lederkremer, G.Z. (2008). Endoplasmic Reticulum (ER) Mannosidase I Is Compartmentalized and Required for N-Glycan Trimming to Man(5–6)GlcNAc(2) in Glycoprotein ER-associated Degradation. Molecular Biology of the Cell *19*, 216-225.

Baldridge, R.D., and Rapoport, T.A. (2016). Autoubiquitination of the Hrd1 Ligase Triggers Protein Retrotranslocation in ERAD. Cell *166*, 394-407.

Bannas, P., Lenz, A., Kunick, V., Fumey, W., Rissiek, B., Schmid, J., Haag, F., Leingärtner, A., Trepel, M., Adam, G., *et al.* (2015). Validation of Nanobody and Antibody Based In Vivo Tumor Xenograft NIRF-imaging Experiments in Mice Using Ex Vivo Flow Cytometry and Microscopy. Journal of Visualized Experiments 52462.

Bays, N.W., Gardner, R.G., Seelig, L.P., Joazeiro, C.A., and Hampton, R.Y. (2000). Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. Nature Cell Biology *3*, 24.

Bocik, W.E., Sircar, A., Gray, J.J., and Tolman, J.R. (2011). Mechanism of polyubiquitin chain recognition by the human ubiquitin conjugating enzyme Ube2g2. The Journal of biological chemistry *286*, 3981-3991.

Bordallo, J., Plemper, R.K., Finger, A., and Wolf, D.H. (1998). Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins. Molecular Biology of the Cell *9*, 209-222.

Bordallo, J., and Wolf, D.H. (1999). A RING-H2 finger motif is essential for the function of Der3/Hrd1 in endoplasmic reticulum associated protein degradation in the yeast Saccharomyces cerevisiae. FEBS Letters *448*, 244-248.

Braakman, I., and Hebert, D.N. (2013). Protein Folding in the Endoplasmic Reticulum. Cold Spring Harbor Perspectives in Biology *5*, a013201.

Burr, M.L., Cano, F., Svobodova, S., Boyle, L.H., Boname, J.M., and Lehner, P.J. (2011). HRD1 and UBE2J1 target misfolded MHC class I heavy chains for endoplasmic reticulum-associated degradation. Proceedings of the National Academy of Sciences of the United States of America *108*, 2034-2039.

Burr, M.L., van den Boomen, D.J., Bye, H., Antrobus, R., Wiertz, E.J., and Lehner, P.J. (2013). MHC class I molecules are preferentially ubiquitinated on endoplasmic reticulum luminal residues during HRD1 ubiquitin E3 ligase-mediated dislocation. Proceedings of the National Academy of Sciences of the United States of America *110*, 14290-14295.

Caramelo, J.J., and Parodi, A.J. (2008). Getting in and out from calnexin/calreticulin cycles. The Journal of biological chemistry *283*, 10221-10225.

Caro, L.G., and Palade, G.E. (1964). Protein synthesis, storage, and discharge in the pancreatic exocrine cell. An Autoradiographic Study. The Journal of cell biology *20*, 473-495.

Chen, B., Mariano, J., Tsai, Y.C., Chan, A.H., Cohen, M., and Weissman, A.M. (2006). The activity of a human endoplasmic reticulum-associated degradation E3, gp78, requires its Cue domain, RING finger, and an E2-binding site. Proceedings of the National Academy of Sciences of the United States of America *103*, 341-346.

Cho, S., Lee, M., and Jun, Y. (2013). Forced interaction of cell surface proteins with Derlin-1 in the endoplasmic reticulum is sufficient to induce their dislocation into the cytosol for degradation. Biochemical and Biophysical Research Communications *430*, 787-792.

Christensen, L.C., Jensen, N.W., Vala, A., Kamarauskaite, J., Johansson, L., Winther, J.R., Hofmann, K., Teilum, K., and Ellgaard, L. (2012). The human selenoprotein VCP-interacting membrane protein (VIMP) is non-globular and harbors a reductase function in an intrinsically disordered region. The Journal of biological chemistry *287*, 26388-26399.

Christianson, J.C., Shaler, T.A., Tyler, R.E., and Kopito, R.R. (2008). OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. Nature Cell Biology *10*, 272-282.

Claessen, J.H., Mueller, B., Spooner, E., Pivorunas, V.L., and Ploegh, H.L. (2010). The transmembrane segment of a tail-anchored protein determines its degradative fate through dislocation from the endoplasmic reticulum. The Journal of biological chemistry *285*, 20732-20739.

Cortez-Retamozo, V., Backmann, N., Senter, P.D., Wernery, U., De Baetselier, P., Muyldermans, S., and Revets, H. (2004). Efficient cancer therapy with a nanobody-based conjugate. Cancer research *64*, 2853-2857.

Deshaies, R.J., Sanders, S.L., Feldheim, D.A., and Schekman, R. (1991). Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. Nature *349*, 806-808.

Dougan, S.K., Hu, C.C., Paquet, M.E., Greenblatt, M.B., Kim, J., Lilley, B.N., Watson, N., and Ploegh, H.L. (2011). Derlin-2-deficient mice reveal an essential role for protein dislocation in chondrocytes. Molecular and cellular biology *31*, 1145-1159.

Doyle, P.J., Saeed, H., Hermans, A., Gleddie, S.C., Hussack, G., Arbabi-Ghahroudi, M., Seguin, C., Savard, M.E., MacKenzie, C.R., and Hall, J.C. (2009). Intracellular expression of a single-domain antibody reduces cytotoxicity of 15-acetyl-deoxynivalenol in yeast. Journal of Biological Chemistry.

Duarte, J.N., Cragnolini, J.J., Swee, L.K., Bilate, A.M., and Bader, J. (2016). Generation of Immunity against Pathogens via Single-Domain Antibody-Antigen Constructs. The Journal of Immunology *197*, 4838-4847.

Elangovan, M., Chong, H.K., Park, J.H., Yeo, E.J., and Yoo, Y.J. (2017). The role of ubiquitin-conjugating enzyme Ube2j1 phosphorylation and its degradation by proteasome during endoplasmic stress recovery. Journal of cell communication and signaling *11*, 265-273.

Ellgaard, L., and Frickel, E.M. (2003). Calnexin, calreticulin, and ERp57: teammates in glycoprotein folding. Cell Biochemistry and Biophysics *39*, 223-247.

Ellgaard, L., and Helenius, A. (2003). Quality control in the endoplasmic reticulum. Nature Reviews Molecular Cell Biology *4*, 181-191.

Eura, Y., Yanamoto, H., Arai, Y., Okuda, T., Miyata, T., and Kokame, K. (2012). Derlin-1 deficiency is embryonic lethal, Derlin-3 deficiency appears normal, and Herp deficiency is intolerant to glucose load and ischemia in mice. PLoS One *7*, e34298.

Excler, J.-L., Ake, J., Robb, M.L., Kim, J.H., and Plotkin, S.A. (2014). Nonneutralizing Functional Antibodies: a New "Old" Paradigm for HIV Vaccines. Clinical and Vaccine Immunology *21*, 1023-1036.

Fagioli, C., and Sitia, R. (2001). Glycoprotein quality control in the endoplasmic reticulum. Mannose trimming by endoplasmic reticulum mannosidase I times the proteasomal degradation of unassembled immunoglobulin subunits. The Journal of biological chemistry *276*, 12885-12892.

Fagone, P., and Jackowski, S. (2009). Membrane phospholipid synthesis and endoplasmic reticulum function. Journal of Lipid Research *50*, S311-S316.

Fang, S., Ferrone, M., Yang, C., Jensen, J.P., Tiwari, S., and Weissman, A.M. (2001). The tumor autocrine motility factor receptor, gp78, is a ubiquitin protein ligase implicated in degradation from the endoplasmic reticulum. Proceedings of the National Academy of Sciences of the United States of America *98*, 14422-14427.

Fang, T., Duarte, J.N., Ling, J., Li, Z., Guzman, J.S., and Ploegh, H.L. (2016). Structurally Defined alphaMHC-II Nanobody-Drug Conjugates: A Therapeutic and Imaging System for B-Cell Lymphoma. Angewandte Chemie International Edition in English *55*, 2416-2420.

Fang, T., Van Elssen, C., Duarte, J.N., Guzman, J.S., Chahal, J.S., Ling, J., and Ploegh, H.L. (2017). Targeted antigen delivery by an anti-class II MHC VHH elicits focused alphaMUC1(Tn) immunity. Chemical science *8*, 5591-5597.

Ferris, S.P., Kodali, V.K., and Kaufman, R.J. (2014). Glycoprotein folding and quality-control mechanisms in protein-folding diseases. Disease Models & Mechanisms *7*, 331-341.

Foresti, O., Ruggiano, A., Hannibal-Bach, H.K., Ejsing, C.S., and Carvalho, P. (2013). Sterol homeostasis requires regulated degradation of squalene monooxygenase by the ubiquitin ligase Doa10/Teb4. eLife *2*, e00953.

Frejd, F.Y., and Kim, K.-T. (2017). Affibody molecules as engineered protein drugs. Experimental & Molecular Medicine *49*, e306.

Fujimori, T., Kamiya, Y., Nagata, K., Kato, K., and Hosokawa, N. (2013). Endoplasmic reticulum lectin XTP3-B inhibits endoplasmic reticulum-associated degradation of a misfolded alpha1-antitrypsin variant. The FEBS journal *280*, 1563-1575.

Grace, T., and Papavasiliou, F.N. (2007). Immunoglobulin Somatic Hypermutation. Annual Review of Genetics *41*, 107-120.

Groisman, B., Shenkman, M., Ron, E., and Lederkremer, G.Z. (2011). Mannose trimming is required for delivery of a glycoprotein from EDEM1 to XTP3-B and to late endoplasmic reticulum-associated degradation steps. The Journal of biological chemistry *286*, 1292-1300.

Hagiwara, M., Ling, J., Koenig, P.A., and Ploegh, H.L. (2016). Posttranscriptional Regulation of Glycoprotein Quality Control in the Endoplasmic Reticulum Is Controlled by the E2 Ub-Conjugating Enzyme UBC6e. Molecular cell *63*, 753-767.

Hagiwara, M., Maegawa, K., Suzuki, M., Ushioda, R., Araki, K., Matsumoto, Y., Hoseki, J., Nagata, K., and Inaba, K. (2011). Structural basis of an ERAD pathway mediated by the ER-resident protein disulfide reductase ERdj5. Molecular Cell *41*, 432-444.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E.B., Bendahman, N., and Hammers, R. (1993). Naturally occurring antibodies devoid of light chains. Nature *363*, 446.

Hampton, R.Y., Gardner, R.G., and Rine, J. (1996). Role of 26S proteasome and HRD genes in the degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum membrane protein. Molecular biology of the cell *7*, 2029-2044.

Hanke, L., Knockenhauer, K.E., Brewer, R.C., van Diest, E., Schmidt, F.I., Schwartz, T.U., and Ploegh, H.L. (2016). The Antiviral Mechanism of an Influenza A Virus Nucleoprotein-Specific Single-Domain Antibody Fragment. mBio *7*.

Hanke, L., Schmidt, F.I., Knockenhauer, K.E., Morin, B., Whelan, S.P., Schwartz, T.U., and Ploegh, H.L. (2017). Vesicular stomatitis virus N protein-specific single-domain antibody fragments inhibit replication. EMBO reports *18*, 1027-1037.

Harmsen, M.M., and De Haard, H.J. (2007). Properties, production, and applications of camelid single-domain antibody fragments. Applied Microbiology and Biotechnology *77*, 13-22.

Hattori, T., Umetsu, M., Nakanishi, T., Togashi, T., Yokoo, N., Abe, H., Ohara, S., Adschiri, T., and Kumagai, I. (2010). High Affinity Anti-inorganic Material Antibody Generation by Integrating Graft and Evolution Technologies: POTENTIAL OF ANTIBODIES AS BIOINTERFACE MOLECULES. The Journal of biological chemistry *285*, 7784-7793.

Helenius, a., Trombetta, E.S., Hebert, D.N., and Simons, J.F. (1997). Calnexin, calreticulin and the folding of glycoproteins. Trends in Cell Biology *7*, 193-200.

Hirao, K., Natsuka, Y., Tamura, T., Wada, I., Morito, D., Natsuka, S., Romero, P., Sleno, B., Tremblay, L.O., Herscovics, A., *et al.* (2006). EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming. The Journal of biological chemistry *281*, 9650-9658.

Hosokawa, N., Kamiya, Y., Kamiya, D., Kato, K., and Nagata, K. (2009). Human OS-9, a lectin required for glycoprotein endoplasmic reticulum-associated degradation, recognizes mannose-trimmed N-glycans. The Journal of biological chemistry *284*, 17061-17068.

Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L.O., Herscovics, A., and Nagata, K. (2001). A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation. EMBO reports *2*, 415-422.

Hosokawa, N., Wada, I., Nagasawa, K., Moriyama, T., Okawa, K., and Nagata, K. (2008). Human XTP3-B forms an endoplasmic reticulum quality control scaffold with the HRD1-SEL1L ubiquitin ligase complex and BiP. The Journal of biological chemistry *283*, 20914-20924.

Houck, Scott A., Ren, Hong Y., Madden, Victoria J., Bonner, Jaclyn N., Conlin, Michael P., Janovick, Jo A., Conn, P.M., and Cyr, Douglas M. (2014). Quality Control Autophagy Degrades Soluble ERAD-Resistant Conformers of the Misfolded Membrane Protein GnRHR. Molecular Cell *54*, 166-179.

Hu, Y., Liu, C., and Muyldermans, S. (2017). Nanobody-Based Delivery Systems for Diagnosis and Targeted Tumor Therapy. Frontiers in Immunology *8*, 1442.

Iida, Y., Fujimori, T., Okawa, K., Nagata, K., Wada, I., and Hosokawa, N. (2011). SEL1L protein critically determines the stability of the HRD1-SEL1L endoplasmic reticulum-associated degradation (ERAD) complex to optimize the degradation kinetics of ERAD substrates. The Journal of biological chemistry *286*, 16929-16939.

Ingram, J.R., Knockenhauer, K.E., Markus, B.M., Mandelbaum, J., Ramek, A., Shan, Y., Shaw, D.E., Schwartz, T.U., Ploegh, H.L., and Lourido, S. (2015). Allosteric activation of apicomplexan calcium-dependent protein kinases. Proceedings of the National Academy of Sciences of the United States of America *112*, E4975-4984.

Inoue, A., Sawata, S.Y., Taira, K., and Wadhwa, R. (2007). Loss-of-function screening by randomized intracellular antibodies: identification of hnRNP-K as a potential target for metastasis. Proceedings of the National Academy of Sciences of the United States of America *104*, 8983-8988.

Kario, E., Amar, N., Elazar, Z., and Navon, A. (2011). A New Autophagy-related Checkpoint in the Degradation of an ERAD-M Target. Journal of Biological Chemistry *286*, 11479-11491.

Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thanedar, S., Roitelman, J., Chau, V., and Wiertz, E. (2004). Human HRD1 Is an E3 Ubiquitin Ligase Involved in Degradation of Proteins from the Endoplasmic Reticulum. The Journal of biological chemistry *279*, 3525-3534.

Knop, M., Finger, A., Braun, T., Hellmuth, K., and Wolf, D.H. (1996). Der1, a novel protein specifically required for endoplasmic reticulum degradation in yeast. The EMBO journal *15*, 753-763.

Koch, G.L. (1990). The endoplasmic reticulum and calcium storage. BioEssays : news and reviews in molecular, cellular and developmental biology *12*, 527-531.

Koenig, P.A., Nicholls, P.K., Schmidt, F.I., Hagiwara, M., Maruyama, T., Frydman, G.H., Watson, N., Page, D.C., and Ploegh, H.L. (2014). The E2 ubiquitin-conjugating enzyme UBE2J1 is required for spermiogenesis in mice. The Journal of biological chemistry *289*, 34490-34502.

Koide, A., and Koide, S. (2007). Monobodies: antibody mimics based on the scaffold of the fibronectin type III domain. Methods in Molecular Biology *352*, 95-109.

Kostova, Z., Mariano, J., Scholz, S., Koenig, C., and Weissman, A.M. (2009). A Ubc7p-binding domain in Cue1p activates ER-associated protein degradation. Journal of cell science *122*, 1374-1381.

Kroeger, H., Miranda, E., MacLeod, I., Pérez, J., Crowther, D.C., Marciniak, S.J., and Lomas, D.A. (2009). Endoplasmic Reticulum-associated Degradation (ERAD) and Autophagy Cooperate to Degrade Polymerogenic Mutant Serpins. The Journal of biological chemistry *284*, 22793-22802.

Lai, C.W., Otero, J.H., Hendershot, L.M., and Snapp, E. (2012). ERdj4 Protein Is a Soluble Endoplasmic Reticulum (ER) DnaJ Family Protein That Interacts with ER-associated Degradation Machinery. The Journal of biological chemistry *287*, 7969-7978.

Leitman, J., Shenkman, M., Gofman, Y., Shtern, N.O., Ben-Tal, N., Hendershot, L.M., and Lederkremer, G.Z. (2014). Herp coordinates compartmentalization and recruitment of HRD1 and misfolded proteins for ERAD. Molecular Biology of the Cell *25*, 1050-1060.

Lenk, U., Yu, H., Walter, J., Gelman, M.S., Hartmann, E., Kopito, R.R., and Sommer, T. (2002). A role for mammalian Ubc6 homologues in ER-associated protein degradation. Journal of cell science *115*, 3007-3014.

Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. Nature *531*, 395-399.

Li, N., Weng, D., Wang, S.M., Zhang, Y., Chen, S.S., Yin, Z.F., Zhai, J., Scoble, J., Williams, C.C., Chen, T., *et al.* (2017). Surfactant protein-A nanobody-conjugated liposomes loaded with methylprednisolone increase lung-targeting specificity and therapeutic effect for acute lung injury. Drug delivery *24*, 1770-1781.

Lilley, B.N., and Ploegh, H.L. (2004). A membrane protein required for dislocation of misfolded proteins from the ER. Nature *429*, 834-840.

Lilley, B.N., and Ploegh, H.L. (2005). Multiprotein complexes that link dislocation, ubiquitination, and extraction of misfolded proteins from the endoplasmic reticulum membrane. Proceedings of the National Academy of Sciences of the United States of America *102*, 14296-14301.

Liu, W., Shang, Y., Zeng, Y., Liu, C., Li, Y., Zhai, L., Wang, P., Lou, J., Xu, P., Ye, Y., *et al.* (2014). Dimeric Ube2g2 simultaneously engages donor and acceptor ubiquitins to form Lys48-linked ubiquitin chains. The EMBO journal *33*, 46-61.

Loregger, A., Raaben, M., Tan, J., Scheij, S., Moeton, M., van den Berg, M., Gelberg-Etel, H., Stickel, E., Roitelman, J., Brummelkamp, T., *et al.* (2017). Haploid Mammalian Genetic Screen Identifies UBXD8 as a Key Determinant of HMGCR Degradation and Cholesterol Biosynthesis. Arteriosclerosis, thrombosis, and vascular biology *37*, 2064-2074.

Mast, S.W., Diekman, K., Karaveg, K., Davis, A., Sifers, R.N., and Moremen, K.W. (2005). Human EDEM2, a novel homolog of family 47 glycosidases, is involved in ER-associated degradation of glycoproteins. Glycobiology *15*, 421-436.

Menon, M.B., Tiedje, C., Lafera, J., Ronkina, N., Konen, T., Kotlyarov, A., and Gaestel, M. (2013). Endoplasmic reticulum-associated ubiquitin-conjugating enzyme Ube2j1 is a novel substrate of MK2 (MAPKAP kinase-2) involved in MK2-mediated TNFalpha production. Biochemical Journal *456*, 163-172.

Meunier, L., Usherwood, Y.K., Chung, K.T., and Hendershot, L.M. (2002). A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. Molecular Biology of the Cell *13*, 4456-4469.

Molinari, M., Calanca, V., Galli, C., Lucca, P., and Paganetti, P. (2003). Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle. Science (New York, NY) *299*, 1397-1400.

Mueller, B., Klemm, E.J., Spooner, E., Claessen, J.H., and Ploegh, H.L. (2008). SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. Proceedings of the National Academy of Sciences of the United States of America *105*, 12325-12330.

Mueller, B., Lilley, B.N., and Ploegh, H.L. (2006). SEL1L, the homologue of yeast Hrd3p, is involved in protein dislocation from the mammalian ER. Journal of Cell Biology *175*, 261-270.

Oda, Y., Hosokawa, N., Wada, I., and Nagata, K. (2003). EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin. Science (New York, NY) *299*, 1394-1397.

Oh, R.S., Bai, X., and Rommens, J.M. (2006). Human homologs of Ubc6p ubiquitin-conjugating enzyme and phosphorylation of HsUbc6e in response to endoplasmic reticulum stress. The Journal of biological chemistry *281*, 21480-21490.

Okuda-Shimizu, Y., and Hendershot, L.M. (2007). Characterization of an ERAD pathway for non-glycosylated BiP substrates which requires Herp. Molecular Cell *28*, 544-554.

Olivari, S., Cali, T., Salo, K.E.H., Paganetti, P., Ruddock, L.W., and Molinari, M. (2006). EDEM1 regulates ER-associated degradation by accelerating de-mannosylation of folding-defective polypeptides and by inhibiting their covalent aggregation. Biochemical and Biophysical Research Communications *349*, 1278-1284.

Olivari, S., Galli, C., Alanen, H., Ruddock, L., and Molinari, M. (2005). A Novel Stress-induced EDEM Variant Regulating Endoplasmic Reticulum-associated Glycoprotein Degradation. Journal of Biological Chemistry *280*, 2424-2428.

Oliveira, S., van Dongen, G.A., Stigter-van Walsum, M., Roovers, R.C., Stam, J.C., Mali, W., van Diest, P.J., and van Bergen en Henegouwen, P.M. (2012). Rapid visualization of human tumor xenografts through optical imaging with a near-infrared fluorescent anti-epidermal growth factor receptor nanobody. Molecular imaging *11*, 33-46.

Olzmann, J.A., Richter, C.M., and Kopito, R.R. (2013). Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. Proceedings of the National Academy of Sciences of the United States of America *110*, 1345-1350.

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. Science (New York, NY) *189*, 347-358.

Pardon, E., Laeremans, T., Triest, S., Rasmussen, S.G.F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W.G.J., Kobilka, B.K., and Steyaert, J. (2014). A general protocol for the generation of Nanobodies for structural biology. Nature Protocols *9*, 674.

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

Pleiner, T., Bates, M., Trakhanov, S., Lee, C.-T., Schliep, J.E., Chug, H., Böhning, M., Stark, H., Urlaub, H., and Görlich, D. (2015). Nanobodies: site-specific labeling for super-resolution imaging, rapid epitope-mapping and native protein complex isolation. eLife *4*, e11349.

Popp, M.W., Antos, J.M., Grotenbreg, G.M., Spooner, E., and Ploegh, H.L. (2007). Sortagging: a versatile method for protein labeling. Nature chemical biology *3*, 707-708.

Pruszynski, M., Koumarianou, E., Vaidyanathan, G., Revets, H., Devoogdt, N., Lahoutte, T., and Zalutsky, M.R. (2013). Targeting Breast Carcinoma with Radioiodinated Anti-HER2 Nanobody. Nuclear medicine and biology *40*, 52-59.

Rapoport, T.A., Li, L., and Park, E. (2017). Structural and Mechanistic Insights into Protein Translocation. Annual review of cell and developmental biology *33*, 369-390.

Rashidian, M., Keliher, E., Dougan, M., Juras, P.K., Cavallari, M., Wojtkiewicz, G.R., Jacobsen, J., Edens, J.G., Tas, J.M., Victora, G., *et al.* (2015a). The use of (18)F-2-fluorodeoxyglucose (FDG) to label antibody fragments for immuno-PET of pancreatic cancer. ACS central science *1*, 142-147.

Rashidian, M., Keliher, E.J., Bilate, A.M., Duarte, J.N., Wojtkiewicz, G.R., Jacobsen, J.T., Cragnolini, J., Swee, L.K., Victora, G.D., Weissleder, R., *et al.* (2015b). Noninvasive imaging of immune responses. Proceedings of the National Academy of Sciences of the United States of America *112*, 6146-6151.

Rashidian, M., Wang, L., Edens, J.G., Jacobsen, J.T., Hossain, I., Wang, Q., Victora, G.D., Vasdev, N., Ploegh, H., and Liang, S.H. (2016). Enzyme-Mediated Modification of Single-Domain Antibodies for Imaging Modalities with Different Characteristics. Angewandte Chemie International Edition in English *55*, 528-533.

Ritter, C., and Helenius, A. (2000). Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. Nature Structural & Molecular Biology *7*, 278-280.

Rudolph, M.J., Vance, D.J., Cassidy, M.S., Rong, Y., Shoemaker, C.B., and Mantis, N.J. (2016). Structural analysis of nested neutralizing and non-neutralizing B cell epitopes on ricin toxin's enzymatic subunit. Proteins *84*, 1162-1172.

Ruiz-Canada, C., Kelleher, D.J., and Gilmore, R. (2009). Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. Cell *136*, 272-283.

Saraogi, I., and Shan, S.-o. (2011). Molecular Mechanism of Co-translational Protein Targeting by the Signal Recognition Particle. Traffic *12*, 535-542.

Satoh, T., Chen, Y., Hu, D., Hanashima, S., Yamamoto, K., and Yamaguchi, Y. (2010). Structural basis for oligosaccharide recognition of misfolded glycoproteins by OS-9 in ER-associated degradation. Molecular Cell *40*, 905-916.

Schmidt, F.I., Lu, A., Chen, J.W., Ruan, J., Tang, C., Wu, H., and Ploegh, H.L. (2016). A single domain antibody fragment that recognizes the adaptor ASC defines the role of ASC domains in inflammasome assembly. The Journal of Experimental Medicine *213*, 771-790.

Schmitz, K.R., Bagchi, A., Roovers, R.C., van Bergen en Henegouwen, P.M., and Ferguson, K.M. (2013). Structural evaluation of EGFR inhibition mechanisms for nanobodies/VHH domains. Structure (London, England : 1993) *21*, 1214-1224.

Schoebel, S., Mi, W., Stein, A., Ovchinnikov, S., Pavlovicz, R., DiMaio, F., Baker, D., Chambers, M.G., Su, H., Li, D., *et al.* (2017). Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. Nature *548*, 352.

Serruys, B., Van Houtte, F., Verbrugghe, P., Leroux-Roels, G., and Vanlandschoot, P. (2009). Llama-derived single-domain intrabodies inhibit secretion of hepatitis B virions in mice. Hepatology (Baltimore, Md) *49*, 39-49.

Shen, Y., and Hendershot, L.M. (2005). ERdj3, a stress-inducible endoplasmic reticulum DnaJ homologue, serves as a cofactor for BiP's interactions with unfolded substrates. Molecular Biology of the Cell *16*, 40-50.

Spiess, C., Zhai, Q., and Carter, P.J. (2015). Alternative molecular formats and therapeutic applications for bispecific antibodies. Molecular Immunology *67*, 95-106.

Stumpp, M.T., and Amstutz, P. (2007). DARPins: a true alternative to antibodies. Current Opinion in Drug Discovery and Development *10*, 153-159.

Sun, F., Zhang, R., Gong, X., Geng, X., Drain, P.F., and Frizzell, R.A. (2006). Derlin-1 promotes the efficient degradation of the cystic fibrosis transmembrane conductance regulator (CFTR) and CFTR folding mutants. The Journal of biological chemistry *281*, 36856-36863.

Sun, S., Shi, G., Han, X., Francisco, A.B., Ji, Y., Mendonça, N., Liu, X., Locasale, J.W., Simpson, K.W., Duhamel, G.E., *et al.* (2014). Sel1L is indispensable for mammalian endoplasmic reticulum-associated degradation, endoplasmic reticulum homeostasis, and survival. Proceedings of the National Academy of Sciences of the United States of America *111*, E582-E591.

Suzuki, M., Otsuka, T., Ohsaki, Y., Cheng, J., Taniguchi, T., Hashimoto, H., Taniguchi, H., and Fujimoto, T. (2012). Derlin-1 and UBXD8 are engaged in dislocation and degradation of lipidated ApoB-100 at lipid droplets. Molecular Biology of the Cell *23*, 800-810.

Truttmann, M.C., Wu, Q., Stiegeler, S., Duarte, J.N., Ingram, J., and Ploegh, H.L. (2015). HypE-specific nanobodies as tools to modulate HypE-mediated target AMPylation. The Journal of biological chemistry *290*, 9087-9100.

Ushioda, R., Hoseki, J., Araki, K., Jansen, G., Thomas, D.Y., and Nagata, K. (2008). ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. Science (New York, NY) *321*, 569-572.

van den Boomen, D.J., Timms, R.T., Grice, G.L., Stagg, H.R., Skodt, K., Dougan, G., Nathan, J.A., and Lehner, P.J. (2014). TMEM129 is a Derlin-1 associated ERAD E3 ligase essential for virus-induced degradation of MHC-I. Proceedings of the National Academy of Sciences of the United States of America *111*, 11425-11430.

Victora, G.D., and Nussenzweig, M.C. (2012). Germinal centers. Annual review of immunology *30*, 429-457.

von Delbruck, M., Kniss, A., Rogov, V.V., Pluska, L., Bagola, K., Lohr, F., Guntert, P., Sommer, T., and Dotsch, V. (2016). The CUE Domain of Cue1 Aligns Growing Ubiquitin Chains with Ubc7 for Rapid Elongation. Molecular Cell *62*, 918-928.

Wajant, H. (2015). Principles of antibody-mediated TNF receptor activation. Cell Death and Differentiation *22*, 1727-1741.

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. Journal of Cell Biology *91*, 551-556.

Walter, P., Ibrahimi, I., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum. I. Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein. Journal of Cell Biology *91*, 545-550.

Wang, T., Wang, B., Huang, H., Zhang, C., Zhu, Y., Pei, B., Cheng, C., Sun, L., Wang, J., Jin, Q., *et al.* (2017). Enterovirus 71 protease 2Apro and 3Cpro differentially inhibit the cellular endoplasmic reticulum-associated degradation (ERAD) pathway via distinct mechanisms, and enterovirus 71 hijacks ERAD component p97 to promote its replication. *13*, e1006674.

Ware, F.E., Vassilakos, A., Peterson, P.A., Jackson, M.R., Lehrman, M.A., and Williams, D.B. (1995). The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. The Journal of biological chemistry *270*, 4697-4704.

Wilkinson, T.C. (2016). Discovery of functional monoclonal antibodies targeting G-protein-coupled receptors and ion channels. Biochemical Society transactions *44*, 831-837.

Williams, D.B. (2006). Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. Journal of cell science *119*, 615-623.

Witte, M.D., Cragnolini, J.J., Dougan, S.K., Yoder, N.C., Popp, M.W., and Ploegh, H.L. (2012). Preparation of unnatural N-to-N and C-to-C protein fusions. Proceedings of the National Academy of Sciences of the United States of America *109*, 11993-11998.

Woodham AW, C.R., Ling J,Rashidian M, Kolifrath SC, Mesyngier M,, Duarte JN, B.J., Skeate JG, Da, and Silva DM, K.W., Ploegh HL. (2018). Nanobody-antigen conjugates elicit HPV-specific anti-tumor
immune responses. Pending publication in Cancer Immunology Research.

Xu, L., Song, X., and Jia, L. (2017). A camelid nanobody against EGFR was easily obtained through refolding of inclusion body expressed in Escherichia coli. Biotechnology and Applied Biochemistry *64*, 895-901.

Yamaguchi, D., Hu, D., Matsumoto, N., and Yamamoto, K. (2010). Human XTP3-B binds to alpha1-antitrypsin variant null(Hong Kong) via the C-terminal MRH domain in a glycan-dependent manner. Glycobiology *20*, 348-355.

Ye, Y., Shibata, Y., Yun, C., Ron, D., and Rapoport, T.A. (2004). A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. Nature *429*, 841-847.

You, H., Ge, Y., Zhang, J., Cao, Y., Xing, J., Su, D., Huang, Y., Li, M., Qu, S., Sun, F., *et al.* (2017). Derlin-1 promotes ubiquitylation and degradation of the epithelial Na⁺ channel, ENaC. Journal of cell science *130*, 1027.

Younger, J.M., Chen, L., Ren, H.Y., Rosser, M.F., Turnbull, E.L., Fan, C.Y., Patterson, C., and Cyr, D.M. (2006). Sequential quality-control checkpoints triage misfolded cystic fibrosis transmembrane conductance regulator. Cell *126*, 571-582.

Yu, Y., Li, J., Zhu, X., Tang, X., Bao, Y., Sun, X., Huang, Y., Tian, F., Liu, X., and Yang, L. (2017). Humanized CD7 nanobody-based immunotoxins exhibit promising anti-T-cell acute lymphoblastic leukemia potential. International journal of nanomedicine *12*, 1969-1983.

Zabetakis, D., Olson, M.A., Anderson, G.P., Legler, P.M., and Goldman, E.R. (2014). Evaluation of Disulfide Bond Position to Enhance the Thermal Stability of a Highly Stable Single Domain Antibody. PLoS ONE *9*, e115405.

Zapun, A., Darby, N.J., Tessier, D.C., Michalak, M., Bergeron, J.J.M., and Thomas, D.Y. (1998). Enhanced Catalysis of Ribonuclease B Folding by the Interaction of Calnexin or Calreticulin with ERp57. The Journal of biological chemistry *273*, 6009-6012.

Zarschler, K., Witecy, S., Kapplusch, F., Foerster, C., and Stephan, H. (2013). High-yield production of functional soluble single-domain antibodies in the cytoplasm of Escherichia coli. Microbial Cell Factories *12*, 97-97.

Chapter 2 Post-transcriptional regulation of glycoprotein quality control in the endoplasmic reticulum is controlled by the E2 Ub-conjugating enzyme UBC6e

Masatoshi Hagiwara[1], Jingjing Ling[1], Paul-Albert Koenig[1,2], Hidde L. Ploegh[1,3,*]

1 Whitehead Institute for Biomedical Research, Cambridge, MA 02142
2 Present address: Klinikum rechts der Isar, Technische Universität München, Institut für Klinische Chemie und Pathobiochemie, 81675 München, Germany
3 Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142
* Correspondence: ploegh@wi.mit.edu

This chapter is reprinted from Hagiwara, M., et al. (2016)."Posttranscriptional Regulation of Glycoprotein Quality Control in the Endoplasmic Reticulum Is Controlled by the E2 Ub-Conjugating Enzyme UBC6e." Molecular Cell 63(5): 753-767. **with permissions from all authors**

Most experiments were designed and conducted by Dr Masatoshi Hagiwara and Jingjing Ling. Manuscript was written by the first two authors and the other authors contributed to editing the manuscript.

2.1 Introduction

The endoplasmic reticulum (ER) is the cellular compartment where membrane and secretory proteins are synthesized. Together, these comprise ~30% of the cellular proteome (Palade, 1975). In the ER, these proteins mature and undergo co- and post-translational modifications. Strict quality control is imposed on newly synthesized proteins, such that only correctly folded proteins are released to their final destination. Misfolded or unfolded proteins are distinguished from correctly folded proteins by the ER-resident quality control apparatus. Proteins that fail to fold correctly are then segregated from correctly folded proteins: the latter can traverse the secretory pathway. Misfolded proteins may be directed to the cytosol by means of a protein-conducting complex, also called the dislocon (Smith et al., 2011). Upon transfer to the cytosol, aberrant proteins are degraded by the ubiquitin-proteasome system. The combined action of the sensors that recognize misfolded proteins, and the dislocon with its connections to the ubiquitin-proteasome system, executes what is now known as ER associated degradation (ERAD).

For N-linked glycoproteins, their glycans play an important role in monitoring and controlling folding status (Xu and Ng, 2015) . The ER contains a number of lectins, each of which recognizes and discriminates among specific N-linked glycans. Calnexin (CNX) and calreticulin (CRT), lectin chaperones that recognize the G_1Man_9 (where G is Glucose and Man is Mannose) form of N-linked glycans, promote productive folding in complex with the oxidoreductase ERp57 (Ellgaard and Frickel, 2003)., Interactions of newly synthesized glycoproteins with EDEM-family proteins (EDEM1, EDEM2 and EDEM3), lectins with mannosidase activity, are diagnostic of incompletely folded, or misfolded, N-linked glycoproteins. Overexpression of any of the EDEM family proteins eliminates the lag phase often seen in the degradation of glycoprotein ERAD substrates (Hirao et al., 2006; Hosokawa et al., 2001; Olivari et al., 2005). This lag phase results from attempts at productive folding in the course of the CNX cycle. Elimination of the lag phase by overexpression of EDEM family proteins accelerates ERAD, attributable to premature release of ERAD substrates from CNX, assisted by the mannosidase activity of EDEM-family proteins (Molinari, 2003; Oda, 2003). OS-9 is a lectin with a mannose 6-phosphate receptor homology (MRH) domain that can bind to terminally misfolded glycoproteins in which the N-glycans have been trimmed to the Man_{5-7} form by EDEM-family proteins or by ER mannosidase 1 (Satoh et al., 2010). SEL1L receives

terminally misfolded proteins through a hand-off from lectins that include EDEM1 and OS-9 (Christianson et al., 2008; Cormier et al., 2009) and serves as a scaffold protein for a major ERAD component, the HRD1 complex, which establishes a functional connection with the cytoplasmic ubiquitin-proteasome system (Iida et al., 2011).

The ER resolves stress, as caused by accumulation of aberrant proteins, via induction of the unfolded protein response (UPR) to maintain protein homeostasis. EDEM-family proteins, OS-9 and SEL1L transcripts are upregulated in the course of the UPR, a pathway activated by ER stressors such as tunicamycin, dithiothreitol (DTT) or thapsigargin. Such stress leads to accelerated ERAD of misfolded or unfolded substrate proteins (Yoshida et al., 2003). At steady state, EDEM1, OS-9 and SEL1L have relatively short half-lives (1-2 hours) (Mueller et al., 2006; Reggiori et al., 2010). It was suggested that EDEM1 and OS-9 undergo lysosomal degradation upon exit from the ER and segregation into the secretory pathway for delivery to endosomal compartments (Reggiori et al., 2010), but specific mechanisms of protein recognition and degradation remain unresolved. Likewise, the factor(s) that contribute(s) to the short half-lives of ERAD enhancers are unknown.

UBC6e is an E2 ubiquitin conjugating enzyme that localizes to the ER via its tail-anchor. We identified UBC6e as a protein that interacts with SEL1L, and thus is a component of the larger HRD1 ERAD complex. (Figure 2.1 A) (Mueller et al., 2008). A reduction in the expression of UBC6e impaired degradation of endogenous Class I MHC heavy chains (Burr et al., 2011). UBC6e also forms a complex with RMA1, an ER-resident E3 ligase distinct from the Hrd1 complex, and Derlin1, hypothesized to be a component of the dislocon (Lilley and Ploegh, 2004). This complex has been suggested to regulate degradation of CFTRD508, a commonly used ERAD substrate (Younger et al., 2006). Mouse embryo fibroblasts (MEFs) generated from UBC6e knockout mice show an accumulation of EDEM1, OS-9 and SEL1L, suggesting a possible role for UBC6e in controlling ERAD capacity. UBC6e$^{-/-}$ mice suffer from male sterility and auditory defects. The molecular causes for these defects remain to be identified (Koenig et al., 2014).

Here we demonstrate that UBC6e down-regulates EDEM1, EDEM3, OS-9 and SEL1L (ERAD enhancers) in the absence of ER stress. UBC6e's function depends strictly on its E2 enzymatic activity but not its phosphorylation status. The exact ER membrane localization of UBC6e and the supramolecular complexes in which UBC6e participates

determine its activity, and directly control the levels of components essential for ERAD activity. The absence of UBC6e increases the levels of ERAD enhancers with a corresponding increase in the rate of clearance of misfolded and/or incompletely folded substrates. UBC6e$^{-/-}$ MEFs show accelerated mannose trimming and premature substrate release from CNX, initiated by ER mannosidase-dependent eviction of substrate from the CNX cycle. Finally, by deletion of UBC6e, accelerated degradation is observed in tissue culture and *in vivo,* not only for canonical ERAD substrates, but also for folding intermediates of proteins that fold slowly, such as tyrosinase. ERAD activity is curtailed under normal circumstances to balance productive folding and degradation of native proteins to avoid the loss-of-function due to premature degradation of proteins that fold comparatively slowly.

2.2 Results

EDEM1, OS-9 and Sel1L enhance ERAD and are upregulated post-transcriptionally in UBC6e$^{-/-}$ cells, independently of the UPR

The HRD1 complex is a multi-subunit protein complex involved in retro-translocation of misfolded proteins from the ER to the cytosol (Stein et al., 2014). The complex contains UBC6e, presumed to act as its cognate E2 ubiquitin conjugating enzyme. It further contains Derlin1/2/3, VIMP, UBXD8, EDEM1, OS-9.1/.2 and XTP3B (Figure 2.1A) (Koenig and Ploegh, 2014) but an accurate definition of its composition remains a challenge, because of the possible existence of sub-complexes. In UBC6e$^{-/-}$ mice and their derived embryonic fibroblast (MEF) lines, we observed up-regulation of select components of the HRD1 ERAD complex, including EDEM1, OS-9 and SEL1L (Figure 2.1B). This up-regulation can be reversed by reintroduction of wild-type UBC6e, not only when UBC6e is expressed constitutively (Figure 2.1F and S1) but also when induced for 24 h off a doxycycline-inducible promoter (Figure 2.4B). Deletion of UBC6e does not affect the level of the homologous E2, UBC6 (Figure 2.1B). The observed restoration of ERAD enhancer levels is therefore specific for UBC6e. Notwithstanding the increase in ERAD enhancers over the course of the UPR, RNA-seq data showed no obvious changes in transcript levels for ERAD enhancers. Their protein levels must therefore be regulated post-transcriptionally (Figure 2.1C, Figure 2.S2). We assessed the levels of BiP, Grp170 and P5 as signature indicators of UPR activity and observed no changes (Figure 2.1D). The

levels of ERAD enhancers in UBC6e$^{-/-}$ cells are thus regulated independently of the UPR.

Homeostasis of SEL1L, EDEM1 and OS-9 requires enzymatic activity of UBC6e

UBC6e is class III ubiquitin-conjugating (E2) enzyme, C-terminally tail-anchored in the ER membrane (van Wijk and Timmers, 2010). We investigated the contribution of its catalytic activity and phosphorylation status to the observed increase in ERAD enhancer levels in Ub6e$^{-/-}$ cells. We eliminated the enzymatic activity of UBC6e by mutation of its E2 catalytic cysteine C91S (Lenk et al., 2002). We also mutated the single known phosphorylation site (Menon et al., 2013) in UBC6e by generating a S184A variant. We also created a phospho-mimetic variant, S184E (Figure 2.1E). Both S184A and S184E mutants down-regulated ERAD enhancers as efficiently as did wild-type UBC6e, whereas the C91S mutant did not (Figure 2.1F). E2 activity of UBC6e is thus essential for attenuation of ERAD enhancers.

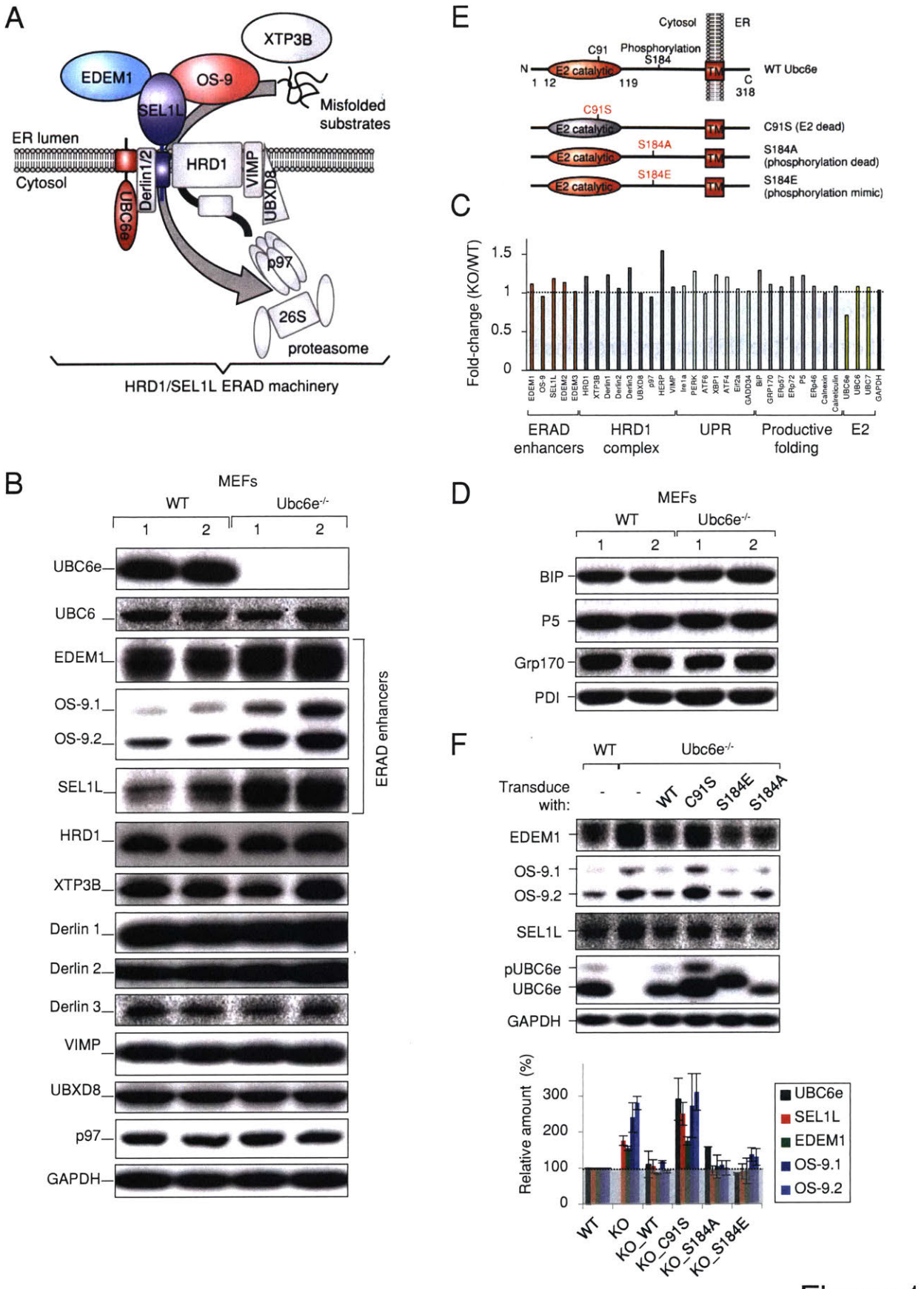

A
XTP3B
EDEM1
OS-9
SEL1L
Misfolded substrates
ER lumen
Cytosol
Derlin1/2
HRD1
VIMP
UBC6e
UBXD8
p97
26S
proteasome
HRD1/SEL1L ERAD machinery
B
MEFs
WT
Ubc6e-/-
1
2
1
2
UBC6e
UBC6
EDEM1
OS-9.1
OS-9.2
SEL1L
ERAD enhancers
HRD1
XTP3B
Derlin 1
Derlin 2
Derlin 3
VIMP
UBXD8
p97
GAPDH
C
Fold-change (KO/WT)
1.5
1
0.5
0
EDEM1 OS-9 SEL1L EDEM2 EDEM3 HRD1 XTP3B Derlin1 Derlin2 Derlin3 UBXD8 p97 HERP VIMP Ire1a PERK ATF6 XBP1 ATF4 Eif2a GADD34 BiP GRP170 ERp57 ERp72 P5 ERp46 Calnexin Calreticulin UBC6e UBC6 UBC7 GAPDH
ERAD enhancers
HRD1 complex
UPR
Productive folding
E2
D
MEFs
WT
Ubc6e-/-
1
2
1
2
BIP
P5
Grp170
PDI
E
Cytosol
ER
C91
Phosphorylation S184
N
E2 catalytic
TM
1 12
119
C
318
WT Ubc6e
C91S
C91S (E2 dead)
S184A
S184A (phosphorylation dead)
S184E
S184E (phosphorylation mimic)
F
WT
Ubc6e-/-
Transduce with:
-
-
WT
C91S
S184E
S184A
EDEM1
OS-9.1
OS-9.2
SEL1L
pUBC6e
UBC6e
GAPDH
Relative amount (%)
300
200
100
0
UBC6e
SEL1L
EDEM1
OS-9.1
OS-9.2
WT
KO
KO_WT
KO_C91S
KO_S184A
KO_S184E

Figure 1

Figure 2.1. Selective ERAD components are regulated by UBC6e

(A) Model showing the components in the canonical HRD1-containing ERAD complex.

(B) ERAD enhancers are selectively up-regulated in two independently derived UBC6e$^{-/-}$ lines of MEFs.

(C) No significant difference is observed for mRNA levels of ER quality control-related genes in wild-type and UBC6e$^{-/-}$ cells. Quantifications of the mRNA expression ratio (KO/WT) of ER-resident E2s and the genes related to ERAD, UPR, productive folding are shown.

(D) Levels of the UPR target proteins BiP, P5 and Grp170 do not differ in wild-type and UBC6e$^{-/-}$ cells.

(E) Scheme for UBC6e structure and the mutants used. Mutation in the active site residue C91 renders the protein enzymatically inactive (Lenk et al., 2002). Phosphorylation occurs at S184 (Menon et al., 2013).

(F) Up-regulation of ERAD enhancers in UBC6e$^{-/-}$ MEF cells is reversed by re-introduction of catalytically active versions of UBC6e. UBC6e$^{-/-}$ cells were transduced with UBC6e variants using pCDH-EF1-CMV-puro vector. Quantifications are normalized to GAPDH. Error bars represent S.D. (n=3). Phosphorylated UBC6e is referred to as pUBC6e.

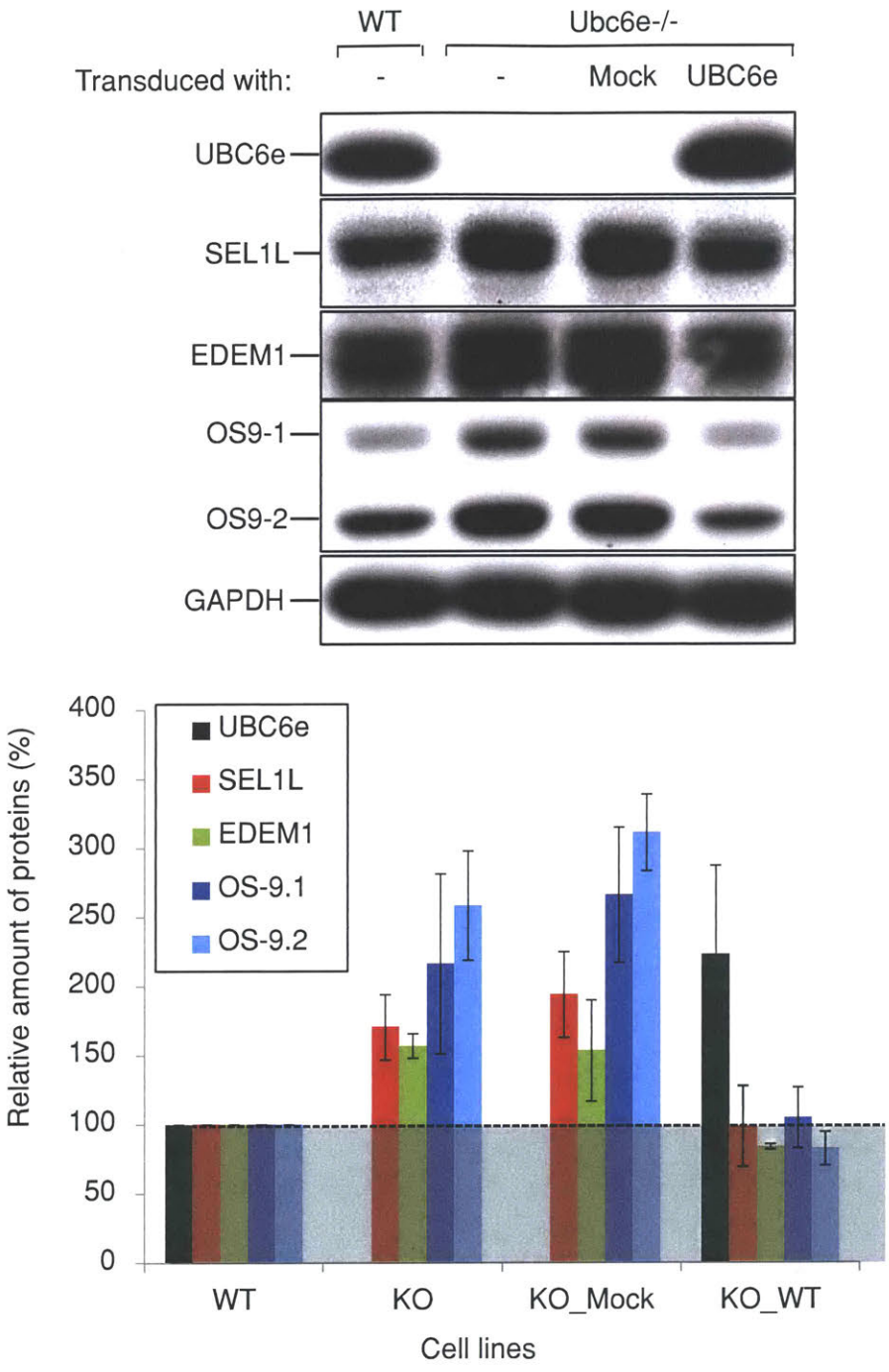


Figure S1.
Re-introduction of UBC6e to Ubc6e$^{-/-}$ MEFs down-regulates ERAD enhancers as wild-type MEFs

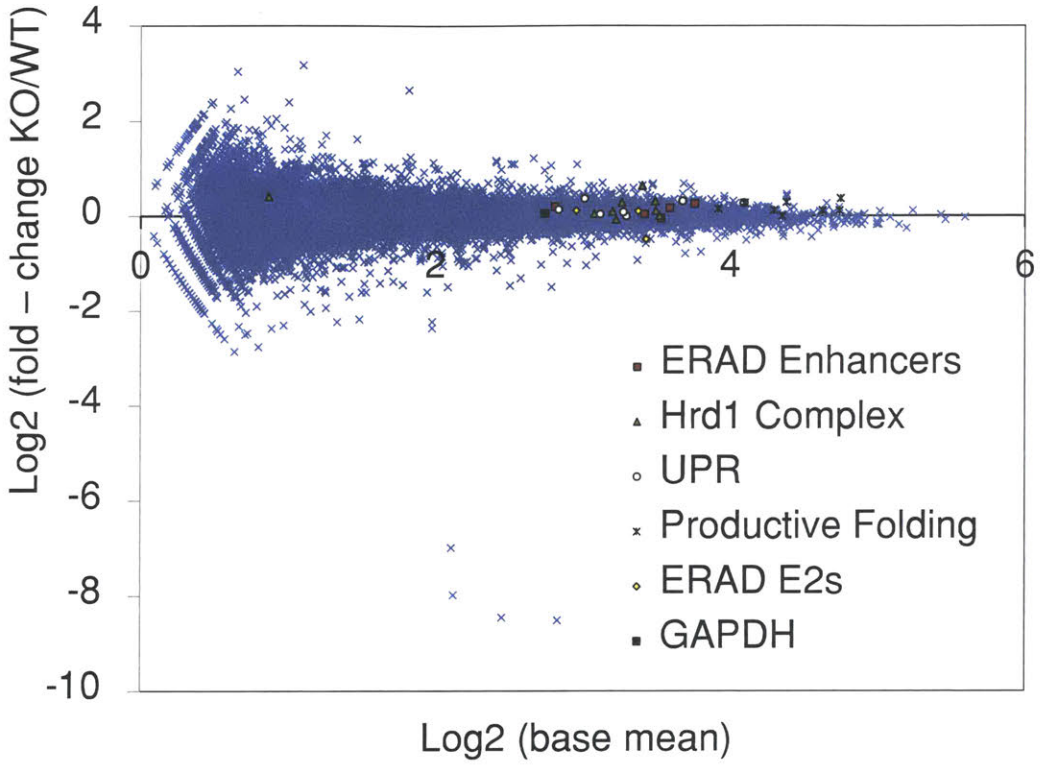


Figure S2.
No significant difference is observed for the mRNA levels for Ubc6e$^{-/-}$ cells in general by RNA sequence analysis

UBC6e$^{-/-}$ cells show elevated ERAD activity

UBC6e$^{-/-}$ cells up-regulate ERAD enhancers selectively. Because UBC6e participates in quality control in the ER, we hypothesized that deletion of UBC6e might attenuate degradation of ERAD substrates, including misfolded EDEM1, OS-9 and SEL1L. However, the observed selective up-regulation of ERAD enhancers in UBC6e$^{-/-}$ cells without UPR led us to consider an alternative possibility, namely that under normal conditions ERAD enhancers are actively degraded in a UBC6e-dependent manner to regulate their participation in ERAD.

We first confirmed by immunofluorescence that excess ERAD enhancers produced in UBC6e$^{-/-}$ cells are localized correctly. We saw no significant changes in their localization and did not note formation of any aggregates (Figure 2.2A). We then tested the activity of ERAD in wild-type and UBC6e$^{-/-}$ cells by transfecting them with a model ERAD substrate, null Hong Kong (NHK), a variant of the glycoprotein a1-antitrypsin bearing a frameshift mutation that causes a C-terminal truncation of 61 amino acids (Sifers et al., 1988). Indeed, degradation of NHK was accelerated in UBC6e$^{-/-}$ cells ($t_{1/2}$ > 120 min in wild-type cells, $t_{1/2}$~80min in UBC6e$^{-/-}$ cells; Figure 2.2B). Addition of the proteasome inhibitor ZL_3VS blocked the degradation of NHK in both cell lines. The UBC6e loss-of-function mutation thus produces a gain-of-function with respect to ERAD activity (Figure 2.2B). We confirmed that the observed increase in ERAD activity applied also to another model ERAD substrate, RI332. RI332 is a truncated variant of the type I ER transmembrane glycoprotein ribophorin (RI), containing only the N-terminal 332 amino acids of its luminal domain (Tsao et al., 1992). We observed acceleration in the rate of degradation of RI332 in UBC6e$^{-/-}$ cells ($t_{1/2}$ > 60 min in wild-type cells, $t_{1/2}$~45min in UBC6e$^{-/-}$ cells; Figure 2.2C). Excess ERAD enhancers produced in UBC6e$^{-/-}$ cells must therefore be folded correctly and functional. Conversely, UBC6e is likely involved in a pathway that -in the absence of obvious stressors- continually degrades otherwise functional ERAD enhancers to control their levels, and therefore should be considered an ERAD rheostat.

Increased ERAD activity is initiated by early mannose trimming of NHK

We focused on identifying the underlying mechanism for increased ERAD activity in UBC6e$^{-/-}$ cells. In pulse-chase experiments in UBC6e$^{-/-}$ cells, we observed a subtle increase

in the mobility of NHK at later chase periods, as monitored by SDS-PAGE, in comparison with NHK retrieved from wild-type cells (Figure 2.2B). We repeated the pulse-chase in UBC6e$^{-/-}$ cells rescued with either wild-type or C91S UBC6e. At the 3-hour time point, NHK showed a noticeable increase in mobility in UBC6e$^{-/-}$ cells compared to wild-type cells. Stable expression of wild-type UBC6e reverts this phenotype, but the catalytically inactive mutant C91S does not. The observed differences in electrophoretic mobility of NHK are therefore UBC6e-dependent (Figure 2.2D). The observed increase in mobility of NHK must be due to increased mannose trimming of N-glycans on NHK. In fact, treatment with an a-mannosidase I selective inhibitor, kifunensine, eliminates both the mobility shift observed for NHK and the difference in rates of degradation of NHK in wild-type and UBC6e$^{-/-}$ cells (Figure 2.2E). Among EDEM family proteins, EDEM1 and EDEM3 but not EDEM2 are up-regulated in an UBC6e-dependent manner, suggesting that they are responsible for the increased mannose trimming in UBC6e$^{-/-}$ cells (Figure 2.2F).

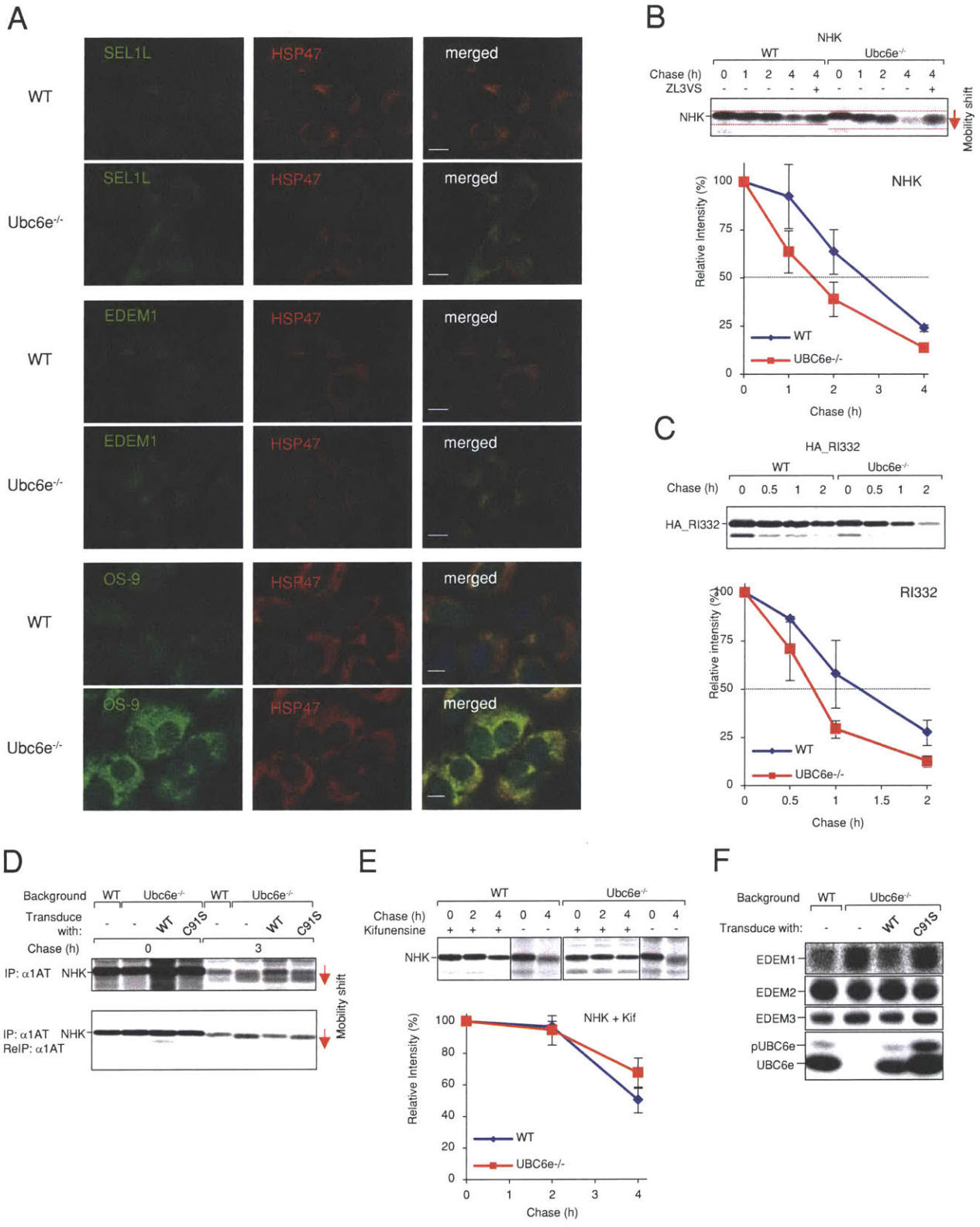

A
SEL1L
HSP47
merged
WT
Ubc6e-/-
EDEM1
OS-9
B
NHK
Chase (h)
ZL3VS
Mobility shift
Relative Intensity (%)
WT
UBC6e-/-
Chase (h)
C
HA_RI332
RI332
Relative intensity (%)
D
Background
Transduce with:
C91S
IP: α1AT
ReIP: α1AT
E
Kifunensine
NHK + Kif
F
EDEM1
EDEM2
EDEM3
pUBC6e
UBC6e

Figure 2

Figure 2.2. UBC6e deletion causes hyperactive ERAD for glycoprotein substrates.

(A) Immunofluorescence of endogenous EDEM1, OS-9 and SEL1L in wild-type and UBC6e$^{-/-}$ cells. Cells are stained with anti-HSP47 as an ER marker.

(B) ERAD is enhanced for the Null Hong Kong (NHK) variant of a1-antitripsin in UBC6e$^{-/-}$ cells. Wild-type and UBC6e$^{-/-}$ cells were transfected with a NHK construct for 24 hours. The cells were starved for 1 hour, labeled with ^{35}S-Met/Cys for 20 min and chased for various time periods. Cells were treated with or without proteasome inhibitor ZL_3VS during the 4 hour chase period. NHK was immunoprecipitated using an a1-antitripsin antibody. Quantifications are normalized to t=0hr. Error bars represent S.D. (n=3).

(C) The glycoprotein ERAD substrate RI332 is degraded more rapidly in UBC6e$^{-/-}$ cells. Wild-type and UBC6e$^{-/-}$ MEFs were transfected with RI332-HA, pulsed for 15 minutes and chased for the indicated time periods. Cell lysates were immunoprecipitated using an anti-HA antibody. Quantifications are normalized to t=0hr. Error bars represent S.D. (n=3).

(D) NHK is deglycosylated more rapidly in cells that lack catalytically active UBC6e. Wild-type, UBC6e$^{-/-}$, and UBC6e$^{-/-}$ MEFs stably expressing wild-type or C91S UBC6e were transfected with NHK. Transfected cells were pulsed for 15 min and chased for 3 hours. NHK was either immunoprecipitated once (upper panel) or re-immunoprecipitated (lower panel) using an a1-antitrypsin antibody.

(E) Kifunensine treatment inhibits the mannose trimming-dependent mobility shift and the degradation of NHK in both wild-type and UBC6e$^{-/-}$ Cells. A pulse-chase experiment as shown in Figure 2B was performed with or without kifunensine treatment. Quantifications are normalized to t=0hr. Error bars represent S.D. (n=3)

(F) EDEM1 and EDEM3, but not EDEM2, are up-regulated in UBC6e-deficient cells.

Early mannose trimming leads to premature release of NHK from Calnexin (CNX)

In a pulse-chase experiment, wild-type cells degrade NHK more slowly during the first hour of chase than do UBC6e$^{-/-}$ cells (Figure 2.2B). This lag phase has been associated with the operation of the calnexin (CNX) / calreticulin (CRT) cycle (Molinari, 2003; Oda, 2003). CNX binds to the G_1Man_9 form of N-glycans on folding intermediates and facilitates productive folding by formation of a complex with the oxidoreductase ERp57. Removal of the innermost glucose from G_1Man_9 by glucosidase II releases the substrate from CNX. (UDP)-glucose: glycoprotein glucosyl transferase (UGGT) senses the folding status of glycoprotein substrates. UGGT monoglucosylates incompletely folded substrates to create the G_1Man_9 form of N-glycans, which then allows the substrate an additional round of the CNX/CRT cycle. If folding is incomplete even after several attempts, terminally misfolded or unfolded substrates undergo further trimming of mannose, which then directs the substrate to the ERAD pathway.

We transfected wild-type and UBC6e$^{-/-}$ cells with NHK and then retrieved CNX by immunoprecipitation. We observed faster release of NHK from CNX in UBC6e$^{-/-}$ cells than in wild-type controls. Time of release corresponded with that observed for the lag phase that precedes degradation in wild-type cells (Figure 2.3A, Figure 2.S3A). In UBC6e$^{-/-}$ and UBC6e$^{-/-}$ cells transduced with C91S UBC6e, we observed faster electrophoretic mobility of NHK, which we attribute to mannose trimming (Figure 2.3A, 3B and Figure 2.S3B). To the best of our knowledge, this is the first time that mannose trimming has been observed on protein substrates bound to CNX.

In UBC6e$^{-/-}$ cells, ERAD substrates are quickly handed over from the CNX cycle to the ERAD pathway

EDEM1, EDEM3, OS-9 and SEL1L are all up-regulated in UBC6e$^{-/-}$ cells, leading to a corresponding increase in functional interactions. More EDEM1, EDEM3 and OS-9 were recovered in association with CNX in UBC6e$^{-/-}$ cells than in wild-type cells (Figure 2.3C). In similar fashion we observed increased interactions between SEL1L and OS-9 (Figure 2.3D, 3E). We did not see direct interactions between EDEM1, EDEM3 and SEL1L (Figure 2.3D). These observations lead us to propose a model for the increase in ERAD activity observed in UBC6e$^{-/-}$ cells. In wild-type cells, glycoproteins carrying G_1M_9 bind to CNX to await folding. However, in UBC6e$^{-/-}$ cells, the increase in EDEM1 and EDEM3 allows for prolonged interactions with CNX, as well as with the protein substrates bound to CNX. These

interactions accelerate mannose trimming on CNX-bound substrates. The more extensively mannose-trimmed substrates no longer bind to CNX and are carried over to the HRD1-SEL1L ERAD complex through interaction with OS-9 (Figure 2.3F).

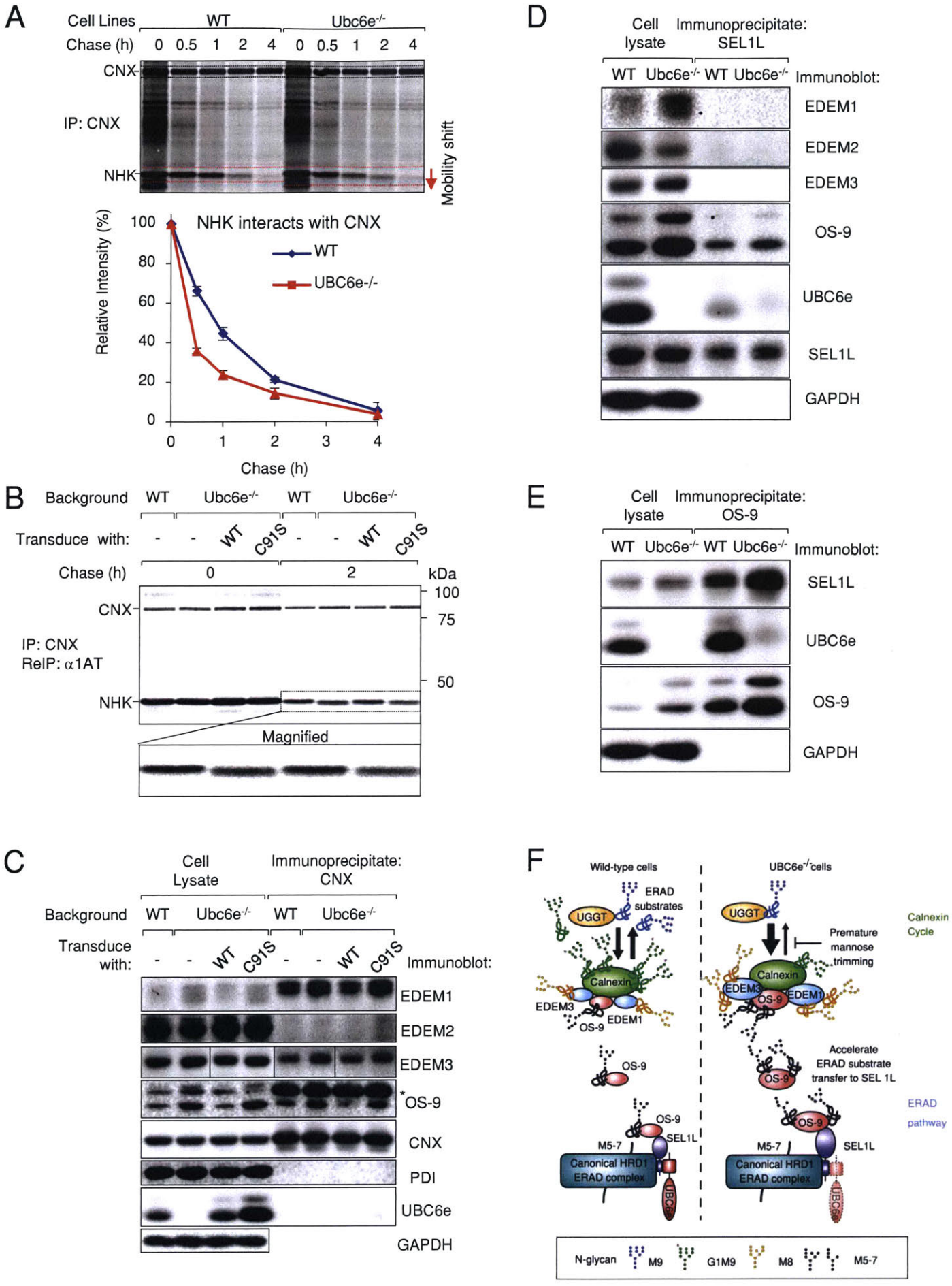

A
Cell Lines
WT
Ubc6e-/-
Chase (h)
0 0.5 1 2 4 0 0.5 1 2 4
CNX
IP: CNX
NHK
Mobility shift
NHK interacts with CNX
WT
UBC6e-/-
Relative Intensity (%)
Chase (h)
B
Background
WT
Ubc6e-/-
Transduce with:
C91S
Chase (h)
kDa
100
75
50
CNX
IP: CNX
ReIP: α1AT
NHK
Magnified
C
Cell Lysate
Immunoprecipitate: CNX
Immunoblot:
EDEM1
EDEM2
EDEM3
*OS-9
CNX
PDI
UBC6e
GAPDH
D
Cell lysate
Immunoprecipitate: SEL1L
Immunoblot:
EDEM1
EDEM2
EDEM3
OS-9
UBC6e
SEL1L
GAPDH
E
Cell lysate
Immunoprecipitate: OS-9
Immunoblot:
SEL1L
UBC6e
OS-9
GAPDH
F
Wild-type cells
UBC6e-/- cells
ERAD substrates
UGGT
Calnexin
EDEM3
EDEM1
OS-9
Premature mannose trimming
Calnexin Cycle
Accelerate ERAD substrate transfer to SEL 1L
ERAD pathway
M5-7
SEL1L
Canonical HRD1 ERAD complex
UBC6e
N-glycan
M9
G1M9
M8
M5-7

Figure 3

Figure 2.3. Deletion of UBC6e leads to pre-mature release of NHK from calnexin (CNX) and enhances interaction of ERAD enhancers

(A) UBC6e deletion promotes the release of NHK from CNX. Wild-type and UBC6e$^{-/-}$ cells were transfected with NHK for 24 hours. Cells were starved for 1 hour, labeled with ^{35}S-Met/Cys for 20 min and chased for various time periods. Lysates were immunoprecipitated with an antibody to CNX. For quantification, the amount of labeled NHK was normalized relative to CNX for each sample. Error bars represent S.D. (n=3)

(B) NHK bound to CNX is deglycosylated in cells that lack functional UBC6e. Wild-type, UBC6e$^{-/-}$, and UBC6e$^{-/-}$ MEFs stably expressing wild-type or C91S UBC6e were transfected with NHK. Transfected cells were labeled for 20 mins and chased for 2 hours. The lysates were immunoprecipitated with a CNX antibody and re-immunoprecipitated using an a1-antitripsin antibody.

(C) More EDEM1, EDEM3 and OS-9 interact with CNX in UBC6e$^{-/-}$ cells. EDEM2 does not co-immunoprecipitate with CNX in either wild-type or UBC6e$^{-/-}$ cells.

(D-E) The amount of interacting SEL1L and OS-9 increases in UBC6e$^{-/-}$ cells. Co-immunoprecipitation of SEL1L and EDEM family proteins is not observed in either wild-type or UBC6e$^{-/-}$ cells.

(F) Model showing the comparison between wild-type and UBC6e$^{-/-}$ cells for a sequential pathway that moves ERAD substrates from productive folding in the CNX cycle to ERAD machineries.

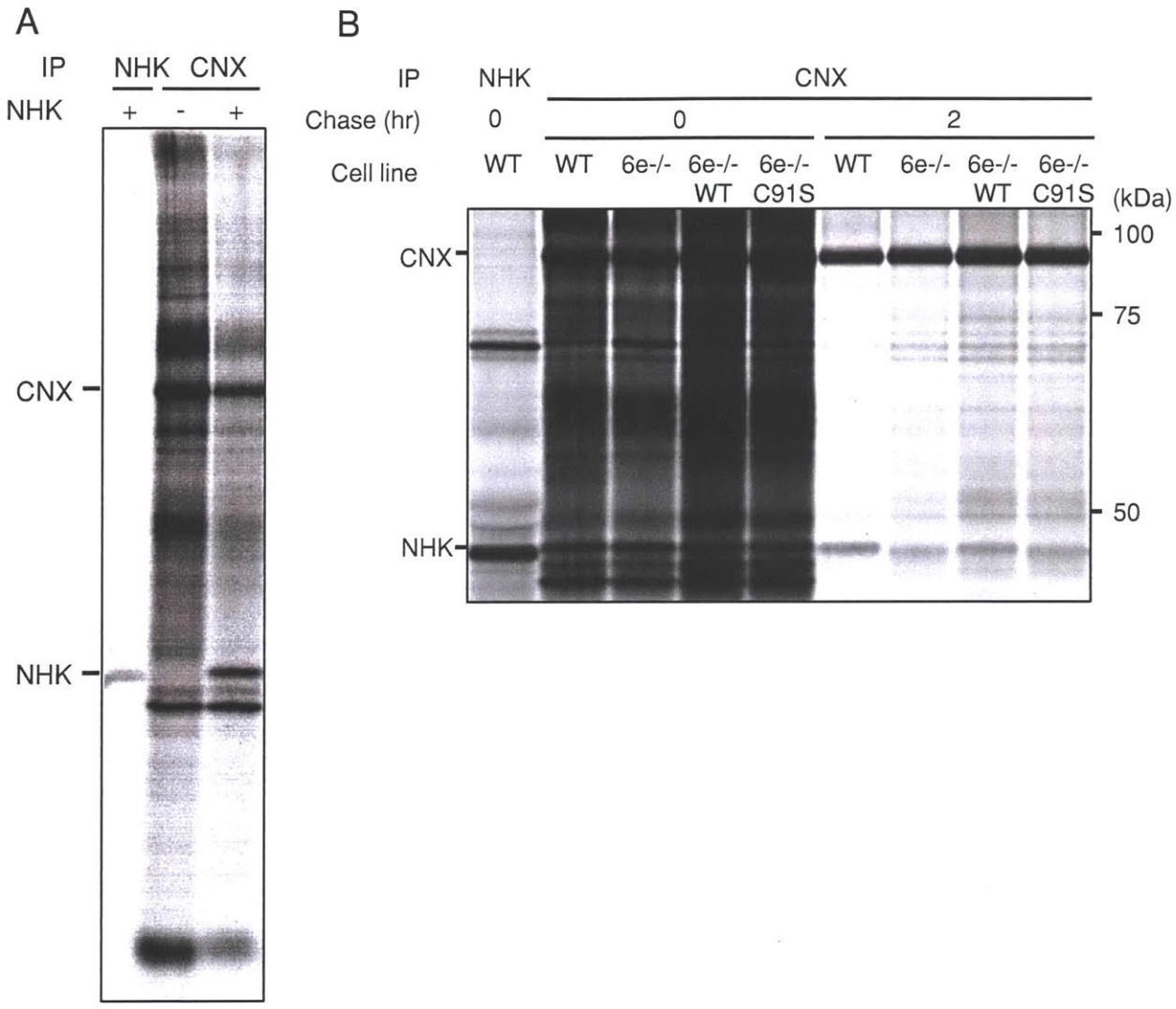


Figure S3.
CNX and NHK interaction in wild-type equivalent cells and in Ubc6e$^{-/-}$ equivalent cells

(A) Detection of co-immunoprecipitated NHK with Calnexin

(B) CNX-bound NHK are de-mannosylated in cells lack functional UBC6e

ER localization and the identity of its transmembrane domain are essential for UBC6e activity

UBC6e regulates the level of ERAD enhancers in order to attenuate ERAD. To assess the importance of proper localization of UBC6e and compare its function with that of mislocalized UBC6e, we generated a cytosolic variant of UBC6e by removal of its C-terminal tail anchor (ΔTM); a version in which the tail anchor is replaced with that of MAVS, a protein targeted to the mitochondria(Seth et al., 2005); and a variant with the transmembrane segment of the type I membrane protein CD4 (Figure 2.4A). The latter construct carries the transmembrane segment in opposite polarity to that of native CD4, yet inserts efficiently into the ER, as shown by glycosylation of an N-linked NXT/S acceptor sequence introduced into the predicted ER luminal portion (Claessen et al., 2010). Finally, we attached the E2 domain of UBC6e to the tail-anchor region of UBC6, the ER-localized and tail-anchored homolog of UBC6e (Figure 2.S6A). Importantly, the E2 domains of all UBC6e mutants are exposed to the cytosol, as is the E2 domain of wild-type UBC6e. This allows an evaluation of the importance of cellular localization. We first confirmed that these variants localize to the respective intended compartments (Figure 2.S5). Upon doxycycline-induced expression in UBC6e$^{-/-}$ cells, only wild-type UBC6e attenuated ERAD enhancers to levels seen in wild-type cells (Figure 2.4B and Figure 2.S6B). We saw only partial down-regulation of ERAD enhancers by UBC6e$_{E2}$-UBC6$_{TM}$ and UBC6e-CD4, even when expressed at high levels (Figure 2.4B, Figure 2.S4 and Figure 2.S6). In the absence of doxycycline, UBC6e$^{-/-}$ cells transduced with wild-type UBC6e still showed almost complete restoration of ERAD enhancer levels, presumably due to leaky promoter activity that allows for expression at about 10% of the wild-type level (data not shown).

We further investigated whether the transmembrane domain of UBC6e is important for its interaction with ERAD enhancers. Only UBC6e with its native transmembrane domain interacts with OS-9 and SEL1L, while interaction with EDEM1 cannot be detected by immunoprecipitation (Figure 2.4C, 4D, S6C and S6D). Even though wild-type UBC6e, UBC6e-CD4 and UBC6e$_{E2}$-UBC6$_{TM}$ all localize to the ER, UBC6e-CD4 and UBC6e$_{E2}$-UBC6$_{TM}$ down-regulate ERAD enhancers only partially, presumably because they were unable to stably interact with them.

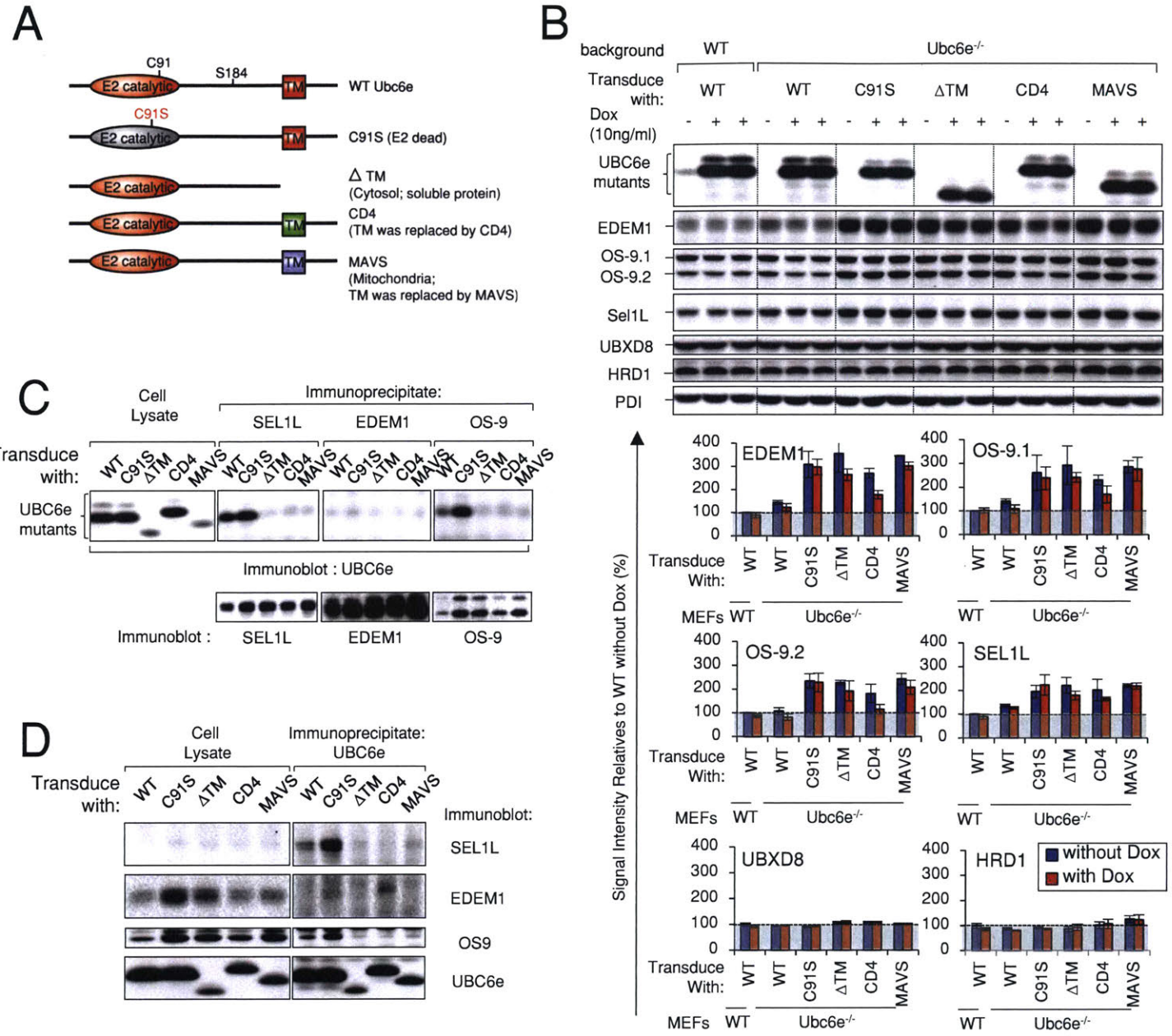

A
C91
S184
E2 catalytic
TM
WT Ubc6e
C91S
C91S (E2 dead)
Δ TM
(Cytosol; soluble protein)
CD4
(TM was replaced by CD4)
MAVS
(Mitochondria;
TM was replaced by MAVS)
B
background
WT
Ubc6e-/-
Transduce with:
Dox (10ng/ml)
UBC6e mutants
EDEM1
OS-9.1
OS-9.2
Sel1L
UBXD8
HRD1
PDI
Signal Intensity Relatives to WT without Dox (%)
without Dox
with Dox
MEFs
C
Cell Lysate
Immunoprecipitate:
SEL1L
EDEM1
OS-9
Immunoblot : UBC6e
Immunoblot :
D
Cell Lysate
Immunoprecipitate: UBC6e
Immunoblot:
SEL1L
EDEM1
OS9
UBC6e

Figure 4

Figure 2.4. Precise ER membrane localization is important for UBC6e function and interaction with EDEM1, OS-9 and SEL1L.

(A) Scheme of the different transmembrane domain (TMD) mutants of UBC6e.

(B) Correct ER localization is essential for the function of UBC6e. Expression of UBC6e-TM mutants was induced using doxycycline for 24 hours. Only UBC6e$^{-/-}$ cells transduced with wild-type UBC6e showed down-regulation of ERAD enhancers to the extent seen in wild-type cells. The other ERAD components, UBXD8 and HRD1, were not affected. Quantification is normalized relative to PDI. Error bars represent S.D. (n=3)

(C-D) The correct transmembrane domain is necessary for UBC6e to interact with the substrates SEL1L and OS-9. UBC6e$^{-/-}$ cells expressing different UBC6e variants were collected in 2% CHAPS and immunoprecipitated with either antibodies against the EDEM1, OS-9 or SEL1L (C) or rabbit anti-UBC6e (D). Interactions between UBC6e and OS-9 or SEL1L were observed only in cells expressing wild-type or C91S UBC6e. We failed to observe interaction between UBC6e and EDEM1.

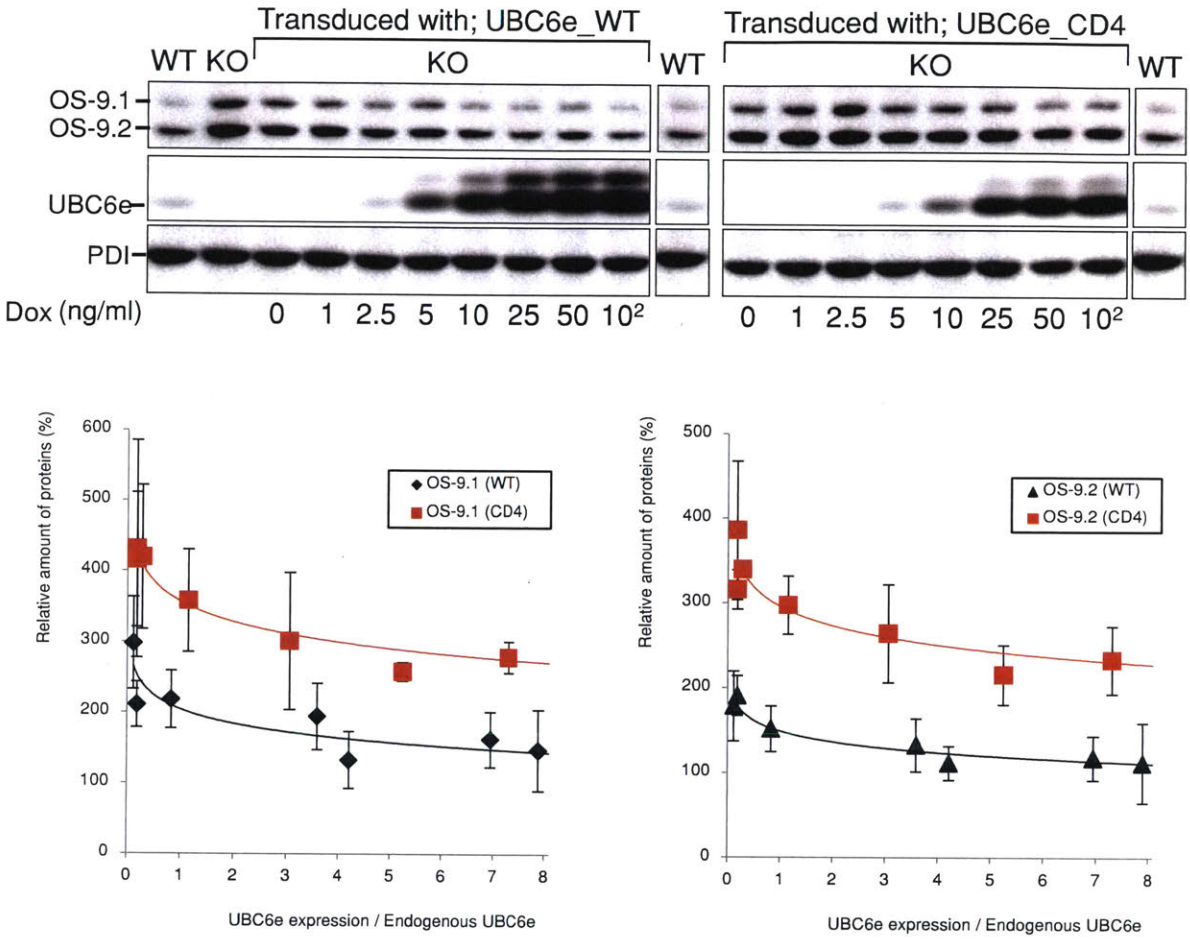


Figure S4.
UBC6e-CD4 mutant cannot down-regulate OS-9 as does wild-type UBC6e even when overexpressed

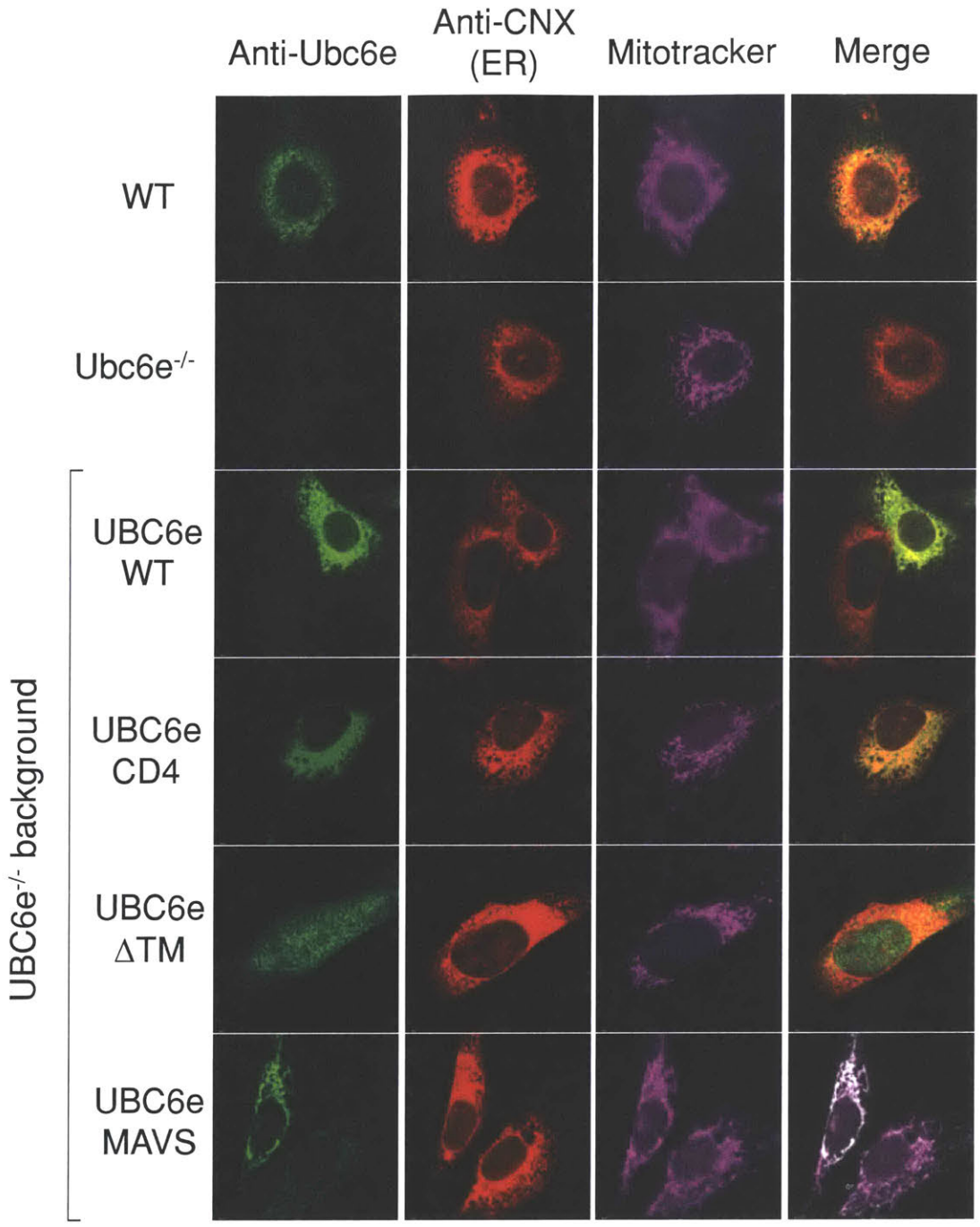


Figure S5.
Immunofluorescence staining showing the localization of the UBC6e-TM mutants.
Ubc6e$^{-/-}$ cells were transduced with UBC6e-TM mutants using the doxycycline-inducible pINDUCER20 system. Cells were stained using Mitotracker (purple), mouse anti-UBC6e (green) and rabbit anti-CNX (red).

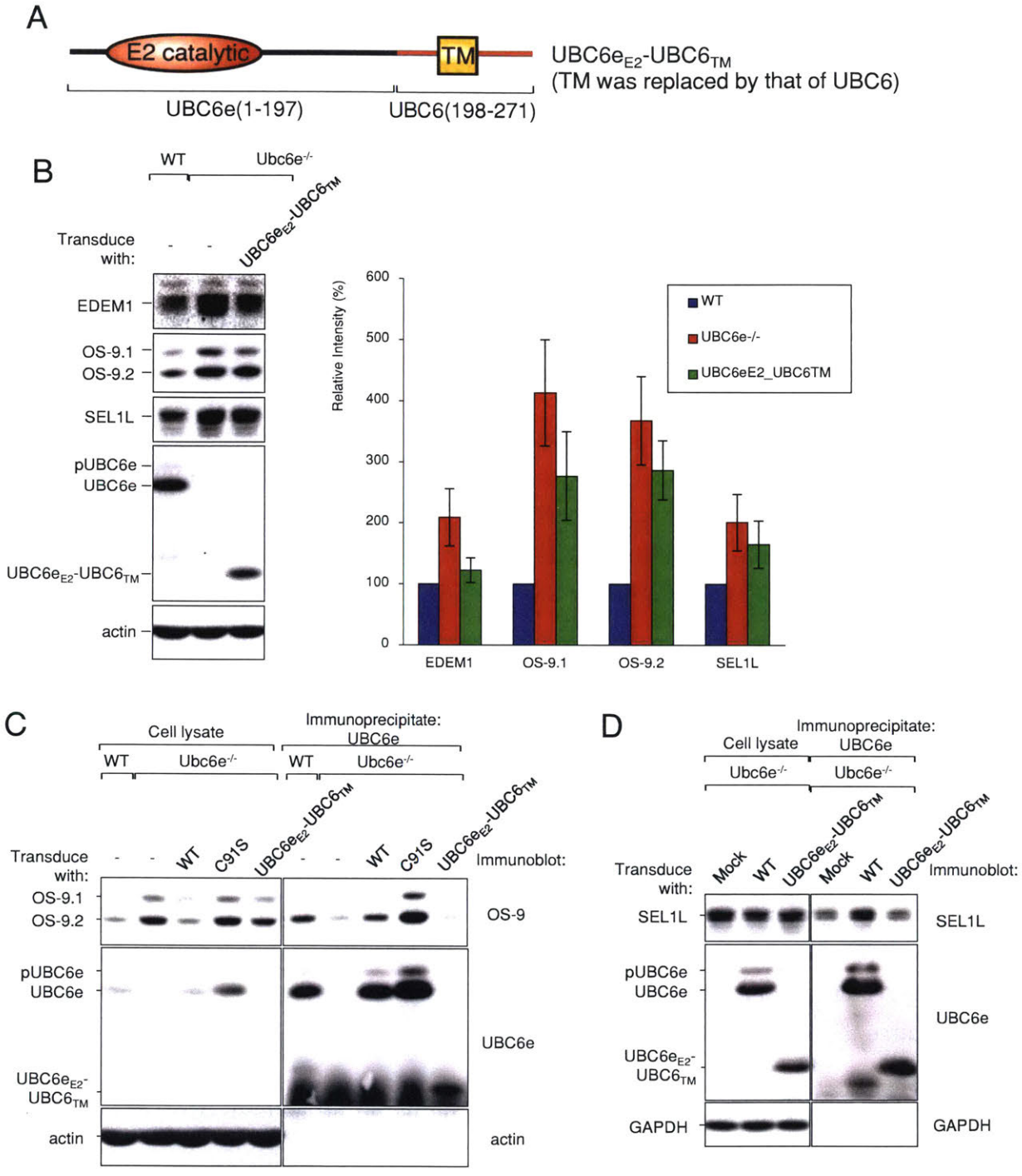


Figure S6.

The native C-terminal tail-anchored region of UBC6e is important for its function

A. Schematic representation of the tail anchor mutant UBC6e$_{E2}$-UBC6$_{TM}$

B. UBC6e$_{E2}$-UBC6$_{TM}$ only partially downregulates ERAD enhancers

C-D. UBC6e$_{E2}$-UBC6$_{TM}$ does not co-immunoprecipitate with OS-9(C) and/or SEL1L(D)

Proper complex formation is essential for UBC6e function

We fractionated various MEF cell lysates on 10-60% linear sucrose density gradients to study complex formation of UBC6e with its partners. For wild-type UBC6e, ~60% of UBC6e is recovered as monomers and ~40% appears in larger complexes (fractions 6-8). However, we observed a different UBC6e distribution profile in UBC6e-CD4 cells, indicating the inability of UBC6e-CD4 to fully engage in proper complex formation (Figure 2.5A). This result is in line with the observation that UBC6e-CD4, despite being localized to the ER, does not down-regulate ERAD enhancers to the same extent as does wild-type UBC6e (Figure 2.4B and S4).

When wild-type cell lysates are fractionated, we observe a signal corresponding to a larger complex (fraction 6-8) that contains UBC6e and ERAD enhancers (Figure 2.5B). This must be the complex that is essential for UBC6e's function in maintaining proper ERAD enhancer levels, since these are the only fractions that contain all members of the complex. We confirmed that UBC6e interacts with OS-9 and SEL1L in this complex, as well as the interaction between OS-9 and SEL1L in the same complex (Figure 2.5C).

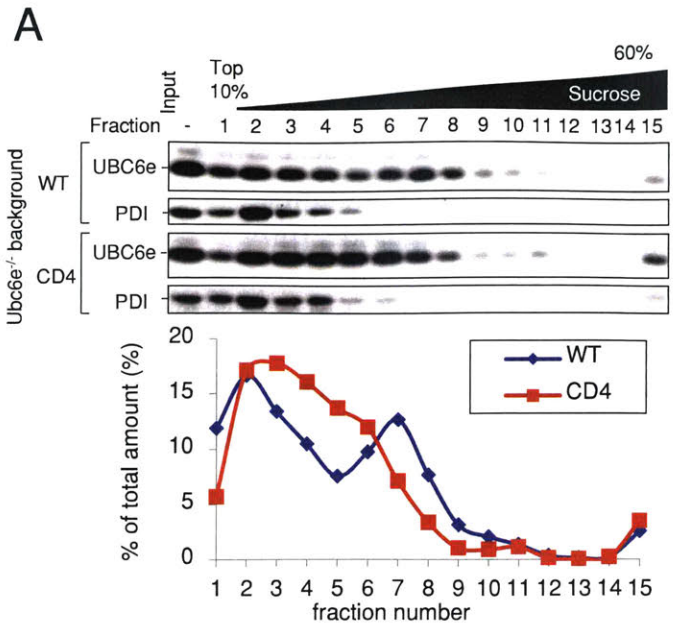
The enzymatically inactive C91S mutant of UBC6e, when over-expressed, might displace wild-type UBC6e from the complex and inhibit normal UBC6e function. We transduced wild-type MEF cells with wild-type or the C91S mutant of UBC6e under the control of a doxycycline-inducible promoter and titrated the expression of UBC6e. As expected, when the C91S mutant was overexpressed to levels about 10-fold over the endogenous level, ERAD enhancers accumulated in a manner similar to that seen in UBC6e$^{-/-}$ cells (Figure 2.5D). Overexpression of the C91S mutant replaces active UBC6e in the complex and prevents degradation of its usual substrates. Overexpression of wild-type UBC6e does not further reduce the levels of ERAD enhancers.

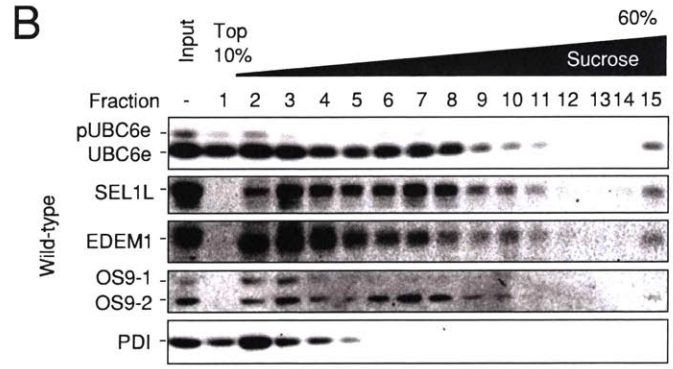
Derlin2 interacts with UBC6e in the complex involved in the regulation of ERAD enhancers

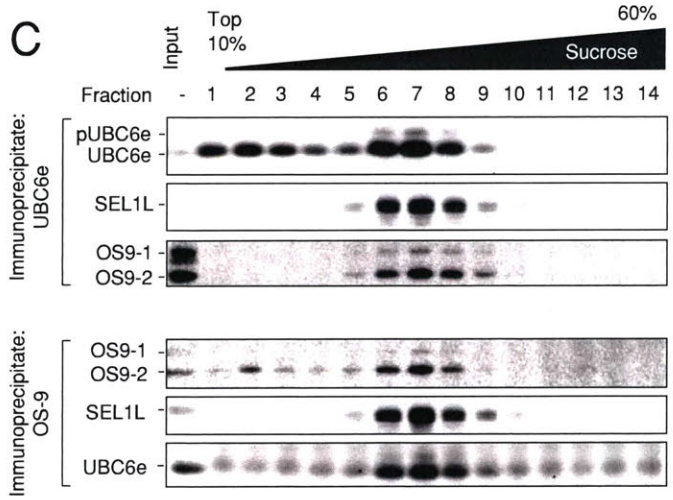
Having found at least two multi-molecular complexes in which UBC6e participates, we considered the presence of yet other complexes. We generated Derlin2$^{-/-}$ mice and MEF cell lines (Dougan et al., 2011), and showed that not only ERAD enhancers but also UBC6e itself was upregulated (Figure 2.S7). Derlin2 might thus be part of the same or a similar complex. Immunoprecipitation using anti-Derlin2 antiserum confirmed its binding to UBC6e in a

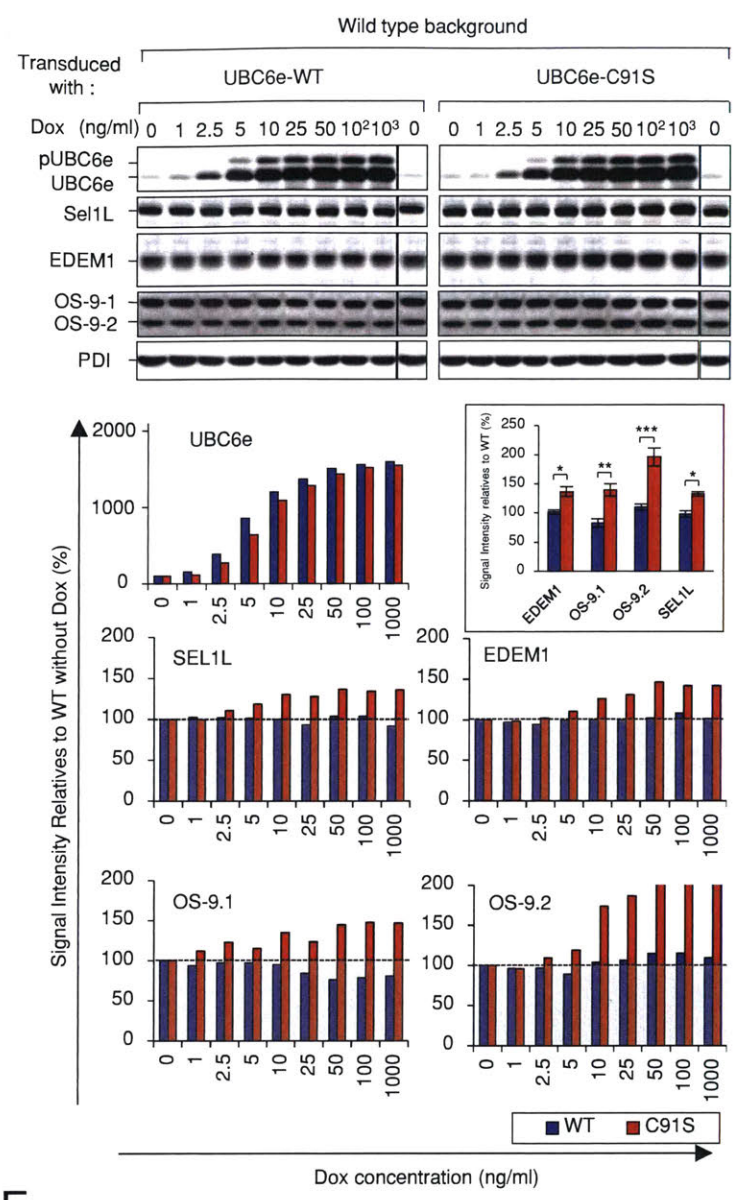
manner that depends on the identity of the transmembrane domain of UBC6e. Derlin2 levels do not change in UBC6e$^{-/-}$ cells transduced with any of the UBC6e mutants, suggesting that it is not a substrate regulated by UBC6e, but is a member of a UBC6e complex (Figure 2.5E).

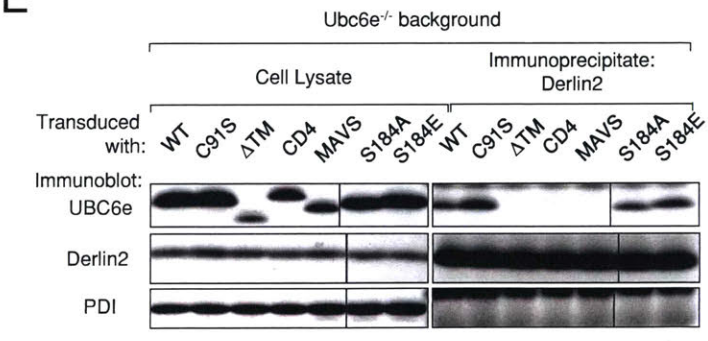
UBC6e exists in at least two different configurations: as an apparent monomer and as part of a complex. Only UBC6e that is part of such complexes controls the levels of ERAD enhancers (Figure 2.5F).


A
Input
Top 10%
60%
Sucrose
Fraction - 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Ubc6e-/- background
WT
UBC6e
PDI
CD4
UBC6e
PDI
% of total amount (%)
WT
CD4
fraction number


B
Input
Top 10%
60%
Sucrose
Fraction - 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Wild-type
pUBC6e
UBC6e
SEL1L
EDEM1
OS9-1
OS9-2
PDI


C
Input
Top 10%
60%
Sucrose
Fraction - 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Immunoprecipitate: UBC6e
pUBC6e
UBC6e
SEL1L
OS9-1
OS9-2
Immunoprecipitate: OS-9
OS9-1
OS9-2
SEL1L
UBC6e


D
Wild type background
Transduced with :
UBC6e-WT
UBC6e-C91S
Dox (ng/ml) 0 1 2.5 5 10 25 50 10² 10³ 0
pUBC6e
UBC6e
Sel1L
EDEM1
OS-9-1
OS-9-2
PDI
Signal Intensity Relatives to WT without Dox (%)
UBC6e
SEL1L
EDEM1
OS-9.1
OS-9.2
Signal Intensity relatives to WT (%)
WT
C91S
Dox concentration (ng/ml)


E
Ubc6e-/- background
Cell Lysate
Immunoprecipitate: Derlin2
Transduced with:
WT C91S ΔTM CD4 MAVS S184A S184E
Immunoblot:
UBC6e
Derlin2
PDI

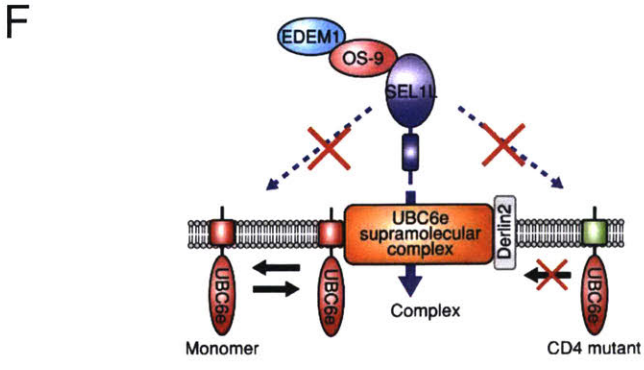

F
EDEM1
OS-9
SEL1L
UBC6e supramolecular complex
Derlin2
UBC6e
Monomer
Complex
CD4 mutant

Figure 5

Figure 2.5. Participation in complex formation is essential for the function of UBC6e.

(A) Wild-type UBC6e and UBC6e-CD4 fractionate differently in a sucrose density gradient. UBC6e$^{-/-}$ cells transduced with wild-type UBC6e and UBC6e-CD4 were lysed in 2% CHAPS and fractionated by centrifugation on a 10-60% linear sucrose gradient in the absence of detergent.

(B) UBC6e and its substrates exist both in monomeric form and in fractions that contain larger complexes.

(C) UBC6e interacts with its substrates only in fractions containing large complexes. Lysate fractions in (B) are immunoprecipitated with anti-UBC6e or anti-OS-9 followed by immunoblot with anti-UBC6e, anti-OS-9 or anti-SEL1L sera.

(D) UBC6e C91S acts as a dominant negative mutant that competes with wild-type UBC6e for complex formation and antagonizes UBC6e regulation of ERAD enhancers. Wild-type MEF cells were transduced with wild-type or C91S UBC6e using doxycycline-inducible pINDUCER20 system and titrated for UBC6e expression. Statistical analysis was conducted on cells that reach induction saturation for UBC6e (>10 fold compared to endogenous). Statistical analysis is presented in the inset as means ± SEM, n = 8. *$p < 0.05$, **$p < 0.001$, ***$p < 0.0001$

(E) The UBC6e transmembrane domain is essential for its interaction with Derlin2. Cells were induced for 24 hours and lysed in 2% CHAPS, immunoprecipitated with anti-Derlin2 and blotted for the interaction with UBC6e.

(F) Model diagram showing that only UBC6e involved in complex are able to regulate SEL1L, OS-9 and EDEM1.

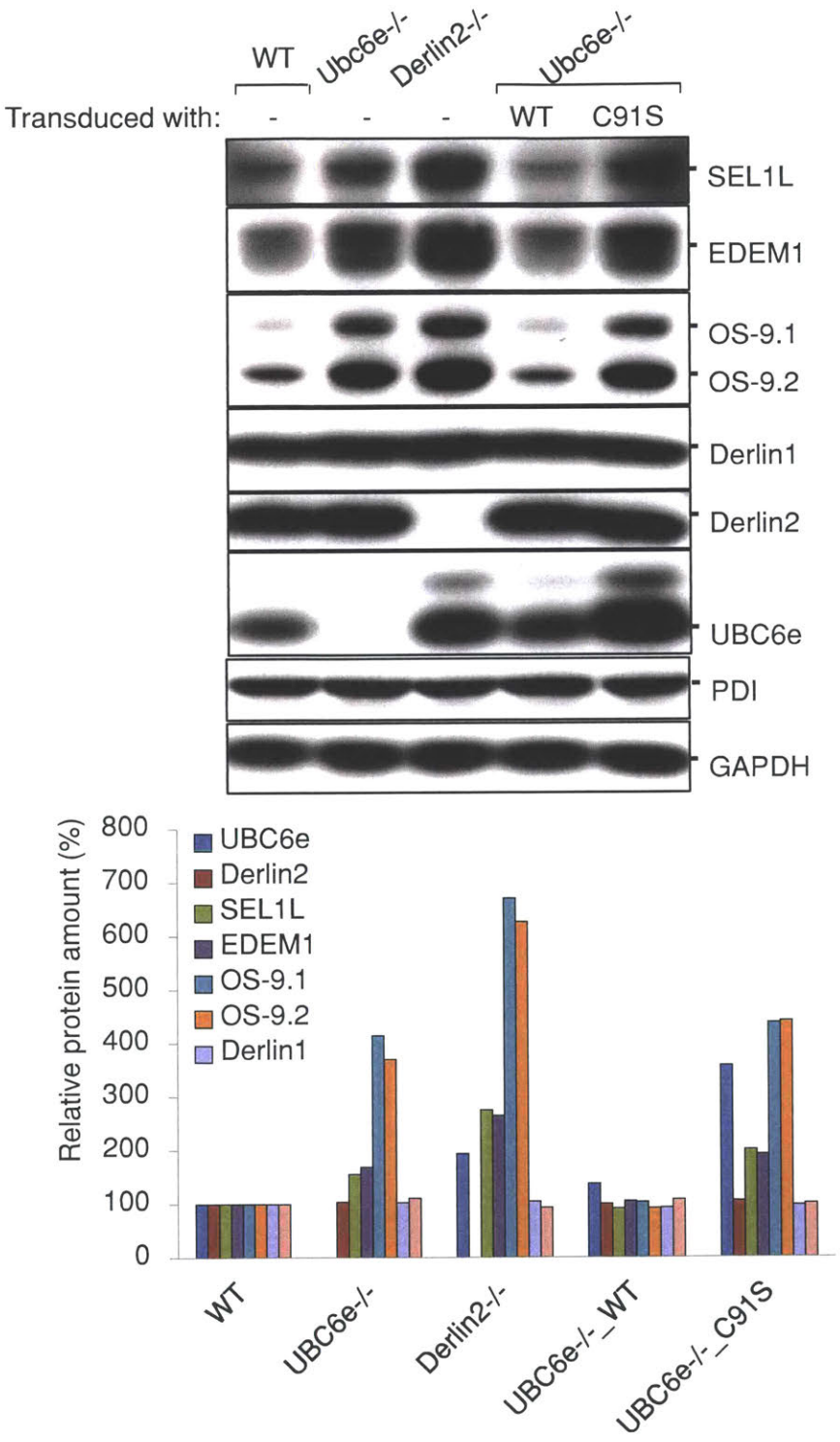


Figure S7.
Derlin2$^{-/-}$ cells showed similar up-regulation of specific ERAD factors compared to UBC6e$^{-/-}$ cells

Overexpression of EDEM1 accelerates degradation of wild-type tyrosinase.

UBC6e controls the amount of ERAD enhancers in order to attenuate ERAD activity under normal circumstances. Increased ERAD activity would not be considered harmful to the cell if only misfolded substrates were subject to accelerated degradation. However, how ER quality control sensors would distinguish between folding intermediates and unfolded proteins is not immediately apparent, because at some point in time both carry the same G_1Man_9 form of N-glycans. We hypothesized that an increase in ERAD activity could be harmful by eliminating proteins that fold comparatively slowly, before they have had the chance to attain their native fold.

To investigate this possibility, we chose tyrosinase as a substrate that folds comparatively slowly. Tyrosinase is a melanocyte-specific type I membrane glycoprotein that catalyzes the rate-limiting step in the production of melanin (Wang and Hebert, 2006). It folds slowly, with 6 N-glycosylation sites and 15 cysteine residues poised to make intramolecular disulfide bonds. Only one-third of tyrosinase attains its properly folded state and is then exported to melanocytes (Wang and Hebert, 2006). Furthermore, a single mutation, C89R or A206T, renders the protein unable to fold and results in a complete diversion to the ERAD pathway (Svedine et al., 2004; Toyofuku et al., 2001). To impose increased ERAD activity, we expressed EDEM1-HA together with wild-type tyrosinase or the C89R mutant fused to YFP in HEK293T cells. Both tyrosinase mutants showed decreased accumulation when co-expressed with EDEM1. The C89R mutant appeared to be more sensitive to increased EDEM1 levels, underscoring a role for EDEM1 in the recognition of misfolded glycoproteins (Figure 2.6A). EDEM1 interacted with all tyrosinase-YFP constructs, as did endogenous EDEM1 (Figure 2.6B and Figure 2.S8). Even in wild type cells, the EDEM1-mediated ERAD pathway removes a fraction of wild-type tyrosinase.

Increased degradation of tyrosinase in UBC6e$^{-/-}$ cells

UBC6e$^{-/-}$ cells have increased ERAD activity for two model ERAD substrates, NHK and RI332 (Figure 2.2B and 2C). Would ERAD also affect proteins known to fold comparatively slowly, such as wild-type tyrosinase? We observed less accumulation of tyrosinase-YFP in UBC6e$^{-/-}$ cells. This phenotype is reversed by expression of wild-type UBC6e but not the C91S mutant (Figure 2.6C). Similar to the canonical ERAD substrates, in pulse chase experiments, tyrosinase-YFP showed accelerated degradation in UBC6e$^{-/-}$ cells, most

clearly during the lag phase ($t_{1/2}$ ~180 min in wild-type cells, $t_{1/2}$ ~90min in UBC6e$^{-/-}$ cells; Figure 2.6D). Inclusion of ZL_3VS stabilizes tyrosinase-YFP, indicating that degradation of tyrosinase is indeed proteasomal and mediated by the ERAD pathway (Figure 2.6D). Retrieval of the transfected wild-type tyrosinase-YFP fusion by immunoprecipitation showed strong interactions with EDEM1, EDEM3 and OS-9 in UBC6e$^{-/-}$-equivalent cells (Figure 2.6E), consistent with the model proposed in Figure 2.3F.

We tracked maturation of tyrosinase by observing its acquisition of EndoH-resistance. Less endoH-resistant tyrosinase appeared in UBC6e$^{-/-}$ cells than in wild-type cells (Figure 2.6F and Figure 2.S9). Addition of ZL_3VS increased the amount of EndoH-resistant tyrosinase in both cell lines, presumably by allowing more tyrosinase molecules the opportunity to attain an export-competent state (Figure 2.S9). Dislocation of misfolded proteins is kinetically coupled to their degradation, and therefore stalling proteasomal degradation is expected to lead to the accumulation of all intermediates upstream of the block, including those that have yet to leave the ER (Wiertz et al., 1996).
We quantified the fraction of EndoH-resistant tyrosinase for each chase timepoint (Figure 2.6F). There is a slight decrease in the fraction of EndoH-resistant tyrosinase in UBC6e$^{-/-}$ cells. While this trend is clear, the difference observed nonetheless fails to reach statistical significance. This difference is much smaller than the faster degradation rate observed (Figure 2.6D).

When analyzed under non-reducing conditions, tyrosinase-YFP from UBC6e$^{-/-}$ cells had a more diffuse appearance at the 4-hour time point (Figure 2.S10A). To distinguish between differences in glycosylation versus redox state of tyrosinase as contributing factors, we analyzed samples with and without EndoH treatment, followed by separation by non-reducing SDS-PAGE (Figure 2.S10B). Treatment with EndoH should remove glycan modifications on ER-resident tyrosinase. Any remaining differences in mobility might thus provide an indication of different redox states of tyrosinase. We observed similar redox states of tyrosinase in wild-type and UBC6e$^{-/-}$ MEFs (Figure 2.S10B). Thus the more diffuse tyrosinase band in UBC6e$^{-/-}$ cells must result from differences in N-linked glycans. Deletion of UBC6e does not affect the redox state of tyrosinase but rather its N-glycan modifications, possibly via up-regulation of ERAD enhancers.. Increased turnover concerns the pool of folding intermediates and is responsible for the reduction of the mature form of tyrosinase.

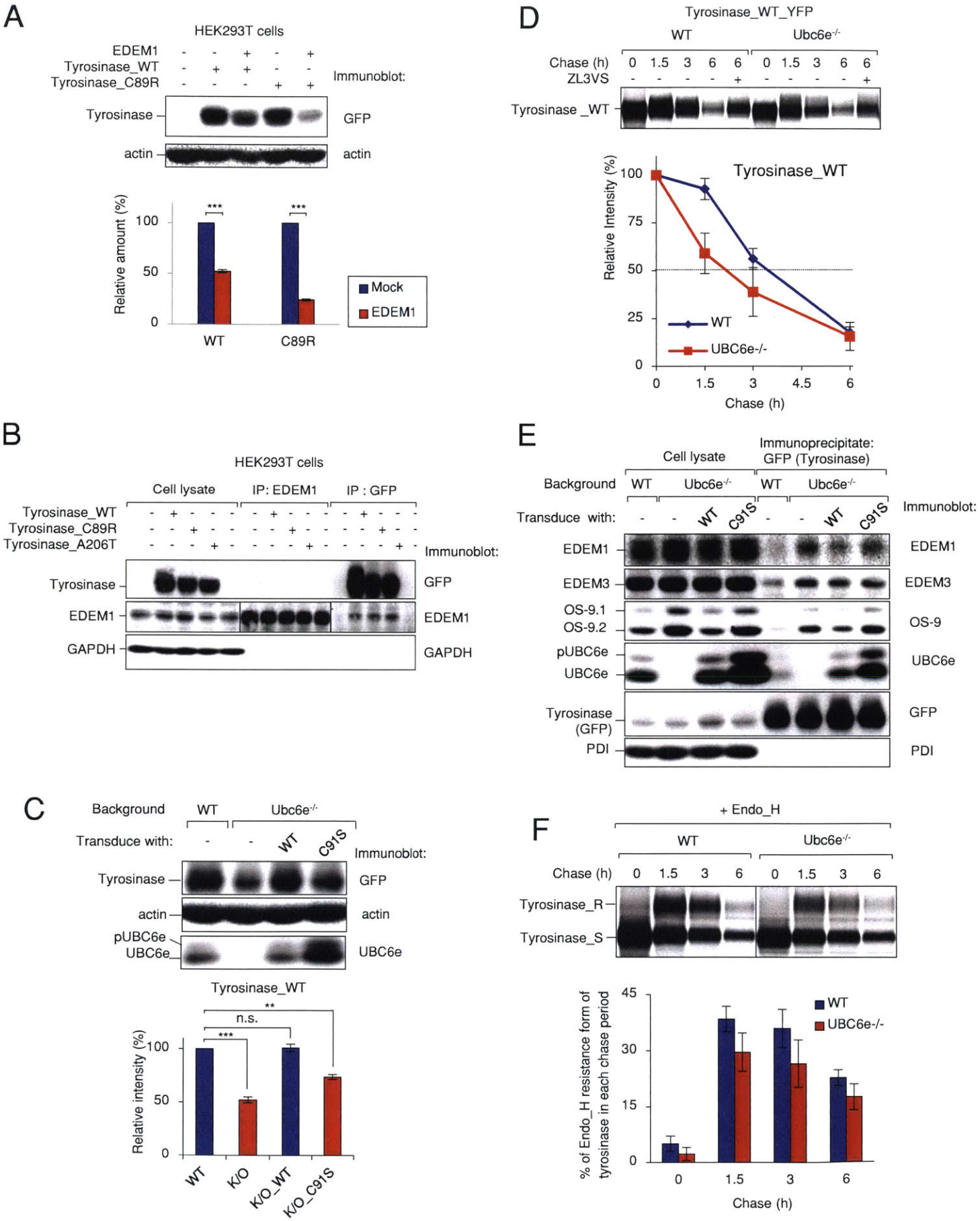

A
HEK293T cells
EDEM1
Tyrosinase_WT
Tyrosinase_C89R
Immunoblot:
Tyrosinase
GFP
actin
actin
Relative amount (%)
100
50
0
***
***
Mock
EDEM1
WT
C89R
B
HEK293T cells
Cell lysate
IP: EDEM1
IP : GFP
Tyrosinase_WT
Tyrosinase_C89R
Tyrosinase_A206T
Immunoblot:
Tyrosinase
GFP
EDEM1
EDEM1
GAPDH
GAPDH
C
Background
WT
Ubc6e-/-
Transduce with:
WT
C91S
Immunoblot:
Tyrosinase
GFP
actin
actin
pUBC6e
UBC6e
UBC6e
Tyrosinase_WT
Relative intensity (%)
100
50
0
**
n.s.
***
WT
K/O
K/O_WT
K/O_C91S
D
Tyrosinase_WT_YFP
WT
Ubc6e-/-
Chase (h)
0 1.5 3 6 6 0 1.5 3 6 6
ZL3VS
Tyrosinase _WT
Tyrosinase_WT
Relative Intensity (%)
100
75
50
25
0
WT
UBC6e-/-
0 1.5 3 4.5 6
Chase (h)
E
Cell lysate
Immunoprecipitate:
GFP (Tyrosinase)
Background
WT
Ubc6e-/-
WT
Ubc6e-/-
Transduce with:
WT
C91S
WT
C91S
Immunoblot:
EDEM1
EDEM1
EDEM3
EDEM3
OS-9.1
OS-9.2
OS-9
pUBC6e
UBC6e
UBC6e
Tyrosinase
(GFP)
GFP
PDI
PDI
F
+ Endo_H
WT
Ubc6e-/-
Chase (h)
0 1.5 3 6 0 1.5 3 6
Tyrosinase_R
Tyrosinase_S
% of Endo_H resistance form of tyrosinase in each chase period
45
30
15
0
WT
UBC6e-/-
0
1.5
3
6
Chase (h)

Figure 6

Figure 2.6. Hyperactive ERAD in UBC6e$^{-/-}$ cells accelerates degradation of the slow-folding substrate tyrosinase and inhibits maturation of tyrosinase

(A) Overexpression of EDEM1 accelerates degradation of both wild-type tyrosinase and C89R mutant. HEK293T cells were transfected with plasmids encoding EDEM1-HA and tyrosinase-YFP for 24 hours. The cells were lysed and immunoblotted for accumulation of tyrosinase-YFP. Error bars represent S.D. (n=3)

(B) Wild-type and mutant tyrosinases interact with endogenous EDEM1. HEK293T cells were transfected with wild-type, C89R or A206T tyrosinase-YFP fusion. Cells were lysed and immunoprecipitation with anti-GFP or anti-EDEM1.

(C) Wild-type tyrosinase accumulates less in cells that lack functional UBC6e. Wild-type, UBC6e$^{-/-}$ and UBC6e$^{-/-}$ stably expressing wild-type or C91S UBC6e are transfected with tyrosinase-YFP for 24 hours and analyzed for tyrosinase accumulation. Error bars represent S.D. (n=3)

(D) Degradation of wild-type tyrosinase is accelerated in UBC6e$^{-/-}$ MEFs. Cells transfected with tyrosinase-YFP were pulsed for 30 min with ^{35}S-Met/Cys and chased for the indicated periods. Cell lysates were immunoprecipitated with anti-GFP. ZL_3VS treatment blocks degradation of the fusion protein. Quantifications are normalized to t=0hr. Error bars represent S.D. (n=3)

(E) More EDEM1, EDEM3 and OS-9 co-immunoprecipitated with transfected tyrosinase-YFP in UBC6e$^{-/-}$ cells.

(F) Tyrosinase maturation and export to the post-ER compartments is inhibited by deletion of UBC6e. Cells were treated as in (D) and tyrosinase-YFP was immunoprecipitated with anti-GFP. The recovered samples were treated with Endoglycosidase H (EndoH). The fraction of EndoH-resistant tyrosinase-YFP in each chase time point was quantified. Error bars represent S.D. (n=3).

The statistical results are presented as means ± SEM, n = 4. **$p < 0.001$, ***$p < 0.0001$

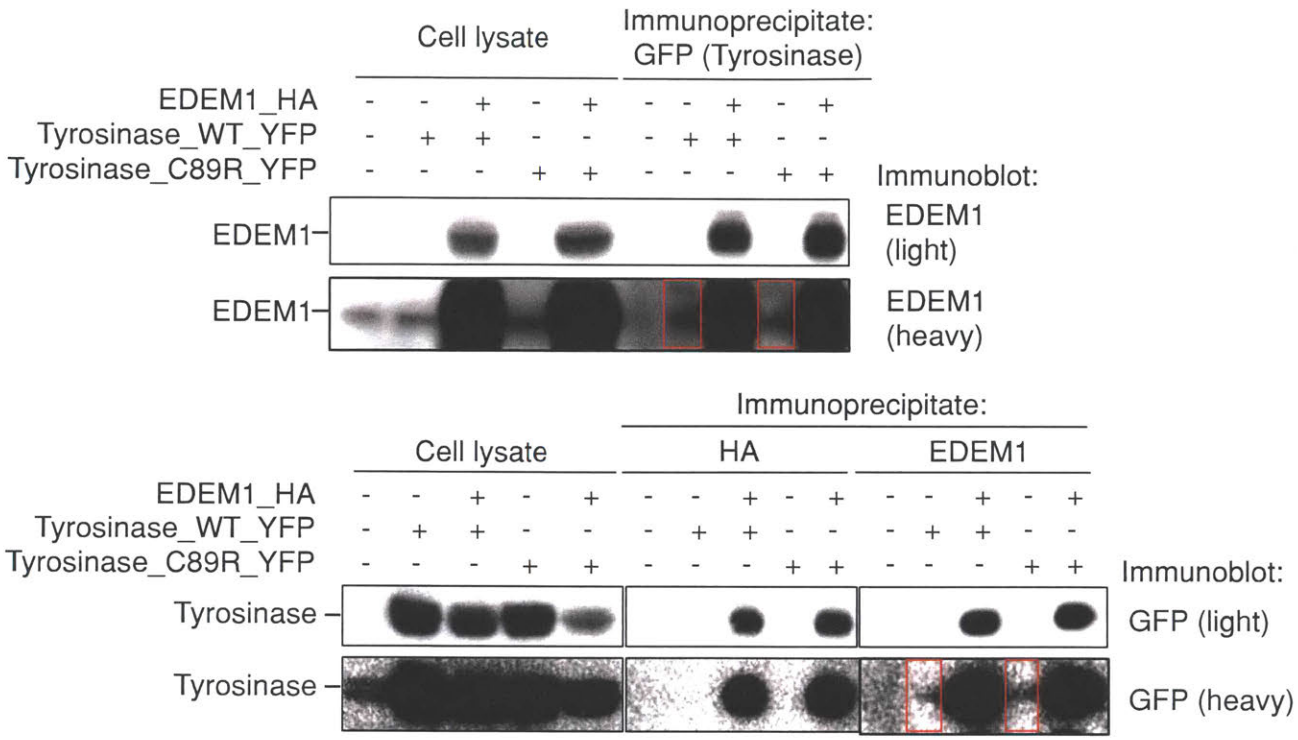


Figure S8.
Overexpressed and endogenous of EDEM1 interact with both wild-type and the C89R Tyrosinase mutant
Cells were lysed in 2% CHAPS and immunoprecipitated with anti-GFP (upper), HA and EDEM1 (lower) antibodies. The eluates were immunoblotted for EDEM1 (upper) or GFP (lower).

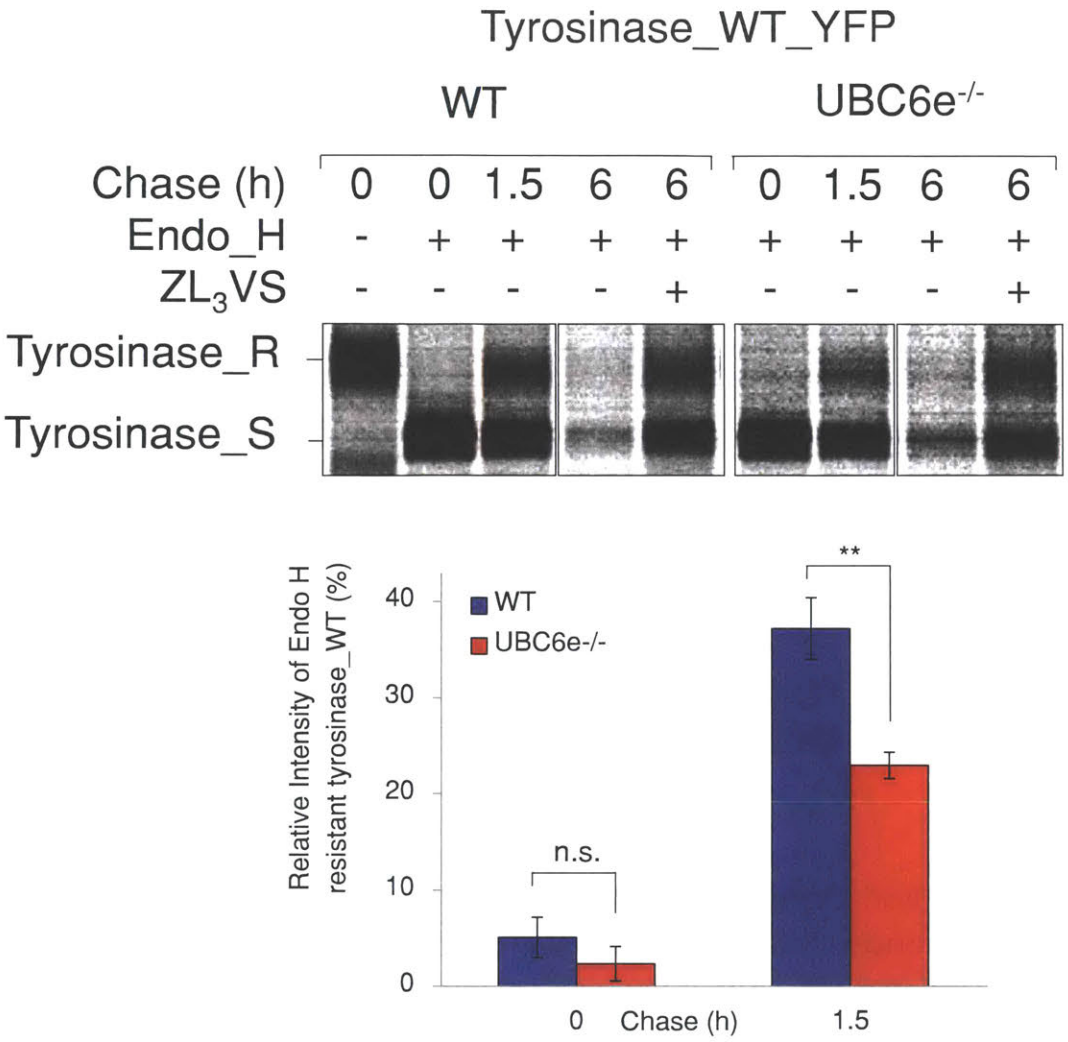


Figure S9.

Deletion of UBC6e affects the maturation of newly synthesized tyrosinase_WT

Cells expressing tyrosinase_WT were pulse labeled with ^{35}S-Met/Cys for 30 min and chased for the indicated time. Cell lysates were immunoprecipitated with anti-GFP for tyrosinase. The immunoprecipitated tyrosinase_WT was treated with Endo_H. ZL_3VS treatment increased the levels of Endo_H resistant tyrosinase in both cell lines. Quantifications are normalized by the total signal of tyrosinase_WT at t=0hr. Error bars represent S.D. (n=3) . **p < 0.001

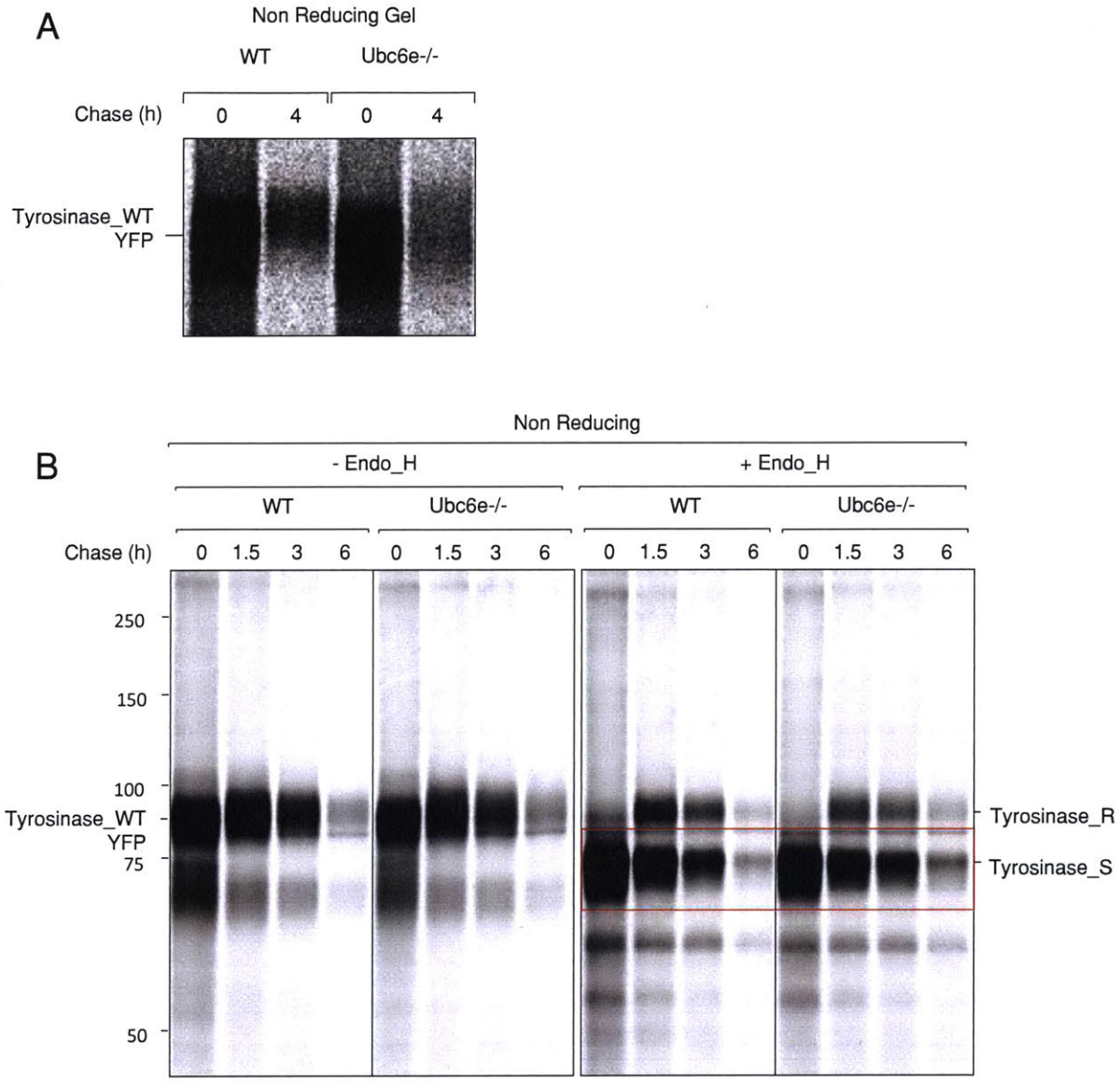


Figure S10.

Deletion of UBC6e does not affect the redox states of tyrosinase_WT

A. Cells expressing tyrosinase_WT were pulse labeled for the indicated periods, chased and analyzed by non reducing SDS-PAGE.

B. Pulse chase analysis for tyrosinase_WT as Fig. 6D. The immunoprecipitated tyrosinase_WT was treated with/without Endo_H and analyzed by non reducing SDS-PAGE.

UBC6e$^{-/-}$ mice have reduced skin tyrosinase levels

ERAD activity in an UBC6e$^{-/-}$ environment affects not only artificial substrate degradation, but also the maturation of wild-type tyrosinase in MEFs, as shown above. To examine whether this finding applied also *in vivo*, we assessed the level of endogenous tyrosinase in the mouse epidermis, as this organ is the site of highest tyrosinase expression. We collected skin lysates from a wild-type mouse, a heterozygous UBC6e$^{+/-}$ mouse and two UBC6e$^{-/-}$ mice. Endogenous tyrosinase levels in UBC6e$^{-/-}$ mice are quite low as assessed by immunoblotting (Figure7A). We therefore used skin lysates in a first round of immunoprecipitation to enrich for tyrosinase, followed by immunoblotting to detect the recovered tyrosinase. We observed a clear reduction in tyrosinase levels only in samples from UBC6e$^{-/-}$ mice (Figure 2.7B). The reduction in total tyrosinase levels correlate with our observations made in cell lines (Figure 2.6C) and correlated with the up-regulation of ERAD enhancers in UBC6e$^{-/-}$ skin lysates. UBC6e$^{+/-}$ and wild-type samples showed comparable results, even with a significant reduction in UBC6e levels in the UBC6e$^{+/-}$ sample (Figure 2.7C).

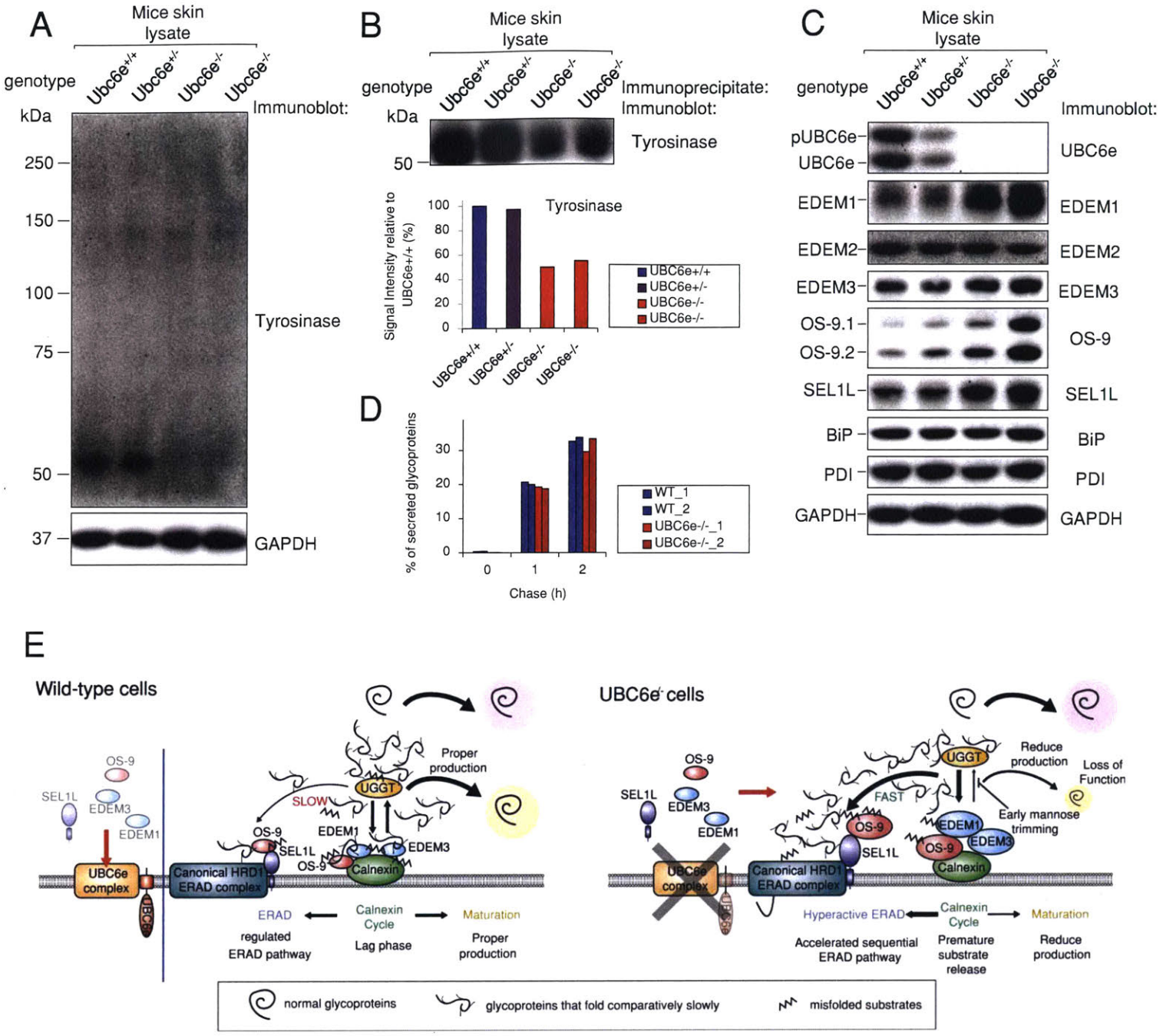

A
Mice skin lysate
genotype
Ubc6e+/+
Ubc6e+/-
Ubc6e-/-
Ubc6e-/-
kDa
Immunoblot:
250
150
100
75
50
37
Tyrosinase
GAPDH
B
Mice skin lysate
genotype
Immunoprecipitate:
Immunoblot:
kDa
50
Tyrosinase
Signal Intensity relative to UBC6e+/+ (%)
100
80
60
40
20
0
Tyrosinase
UBC6e+/+
UBC6e+/-
UBC6e-/-
UBC6e-/-
D
% of secreted glycoproteins
30
20
10
0
0
1
2
Chase (h)
WT_1
WT_2
UBC6e-/-_1
UBC6e-/-_2
C
Mice skin lysate
genotype
Immunoblot:
pUBC6e
UBC6e
UBC6e
EDEM1
EDEM1
EDEM2
EDEM2
EDEM3
EDEM3
OS-9.1
OS-9.2
OS-9
SEL1L
SEL1L
BiP
BiP
PDI
PDI
GAPDH
GAPDH
E
Wild-type cells
OS-9
SEL1L
EDEM3
EDEM1
UBC6e complex
Canonical HRD1 ERAD complex
Proper production
UGGT
SLOW
OS-9
EDEM1
SEL1L
EDEM3
OS-9
Calnexin
ERAD
regulated ERAD pathway
Calnexin Cycle
Lag phase
Maturation
Proper production
UBC6e-/- cells
OS-9
SEL1L
EDEM3
EDEM1
UBC6e complex
Canonical HRD1 ERAD complex
UGGT
Reduce production
Loss of Function
FAST
Early mannose trimming
OS-9
EDEM1
EDEM3
OS-9
SEL1L
Calnexin
Hyperactive ERAD
Accelerated sequential ERAD pathway
Calnexin Cycle
Premature substrate release
Maturation
Reduce production
normal glycoproteins
glycoproteins that fold comparatively slowly
misfolded substrates

Figure 7

Figure 2.7. UBC6e regulates ERAD activity through ERAD enhancers to curtail premature destruction of proteins that fold comparatively slowly.

(A) Immunoblot of endogenous tyrosinase in skin samples from UBC6e$^{+/+}$, UBC6e$^{+/-}$ and UBC6e$^{-/-}$ mice failed to yield detectable signal. Epidermis of individual mice was homogenized and lysed in 1% NP40. Post-nuclear supernatants were used as the source of skin lysates. 15 μg total protein were separated on SDS-PAGE and immunoblotted for tyrosinase and GAPDH.

(B) UBC6e$^{-/-}$ mice show reduced level of endogenous tyrosinase. 500 μg skin lysate in (A) were immunoprecipitated with antibody to tyrosinase. The recovered tyrosinase was then detected by immunoblotting using the same antibody.

(C) ERAD enhancers are selectively up-regulated in skin lysates from UBC6e$^{-/-}$ mice.

(D) UBC6e deletion does not affect the overall secretion of glycoproteins. Two independently derived wild-type and UBC6e$^{-/-}$ MEFs were pulse-labeled with ^{35}S-Met/Cys for 10 min and chased for the indicated periods. Cell lysates and media were precipitated with Concanavalin A sepharose beads to recover the bulk of glycoproteins. The recovered samples were examined by SDS-PAGE/fluorography. The percentage of secreted glycoproteins was calculated by dividing the signal intensity in the media by the total signal in both the cells and media at that time point.

(E) Model on how UBC6e fine-tunes ERAD activity to achieve a balance between productive folding and degradation. In UBC6e$^{-/-}$ cells, the UBC6e-dependent pathway that constantly downregulates ERAD enhancers is shut off and ERAD enhancers accumulate. The high-mannose glycans on substrates captured by CNX/CRT are more rapidly trimmed off by the excess ER mannosidases, EDEM1 and EDEM3. After formation of MAN_{5-7}-carrying glycans, the substrates are released from the productive folding cycle on CNX, transferred to OS-9, and then SEL1L in the HRD1 complex for ubiquitination and degradation. The decreased time spent in a productive folding cycle may not affect the proteins that fold regularly, but could be detrimental to those that fold comparatively slowly, leading to their loss-of-function.

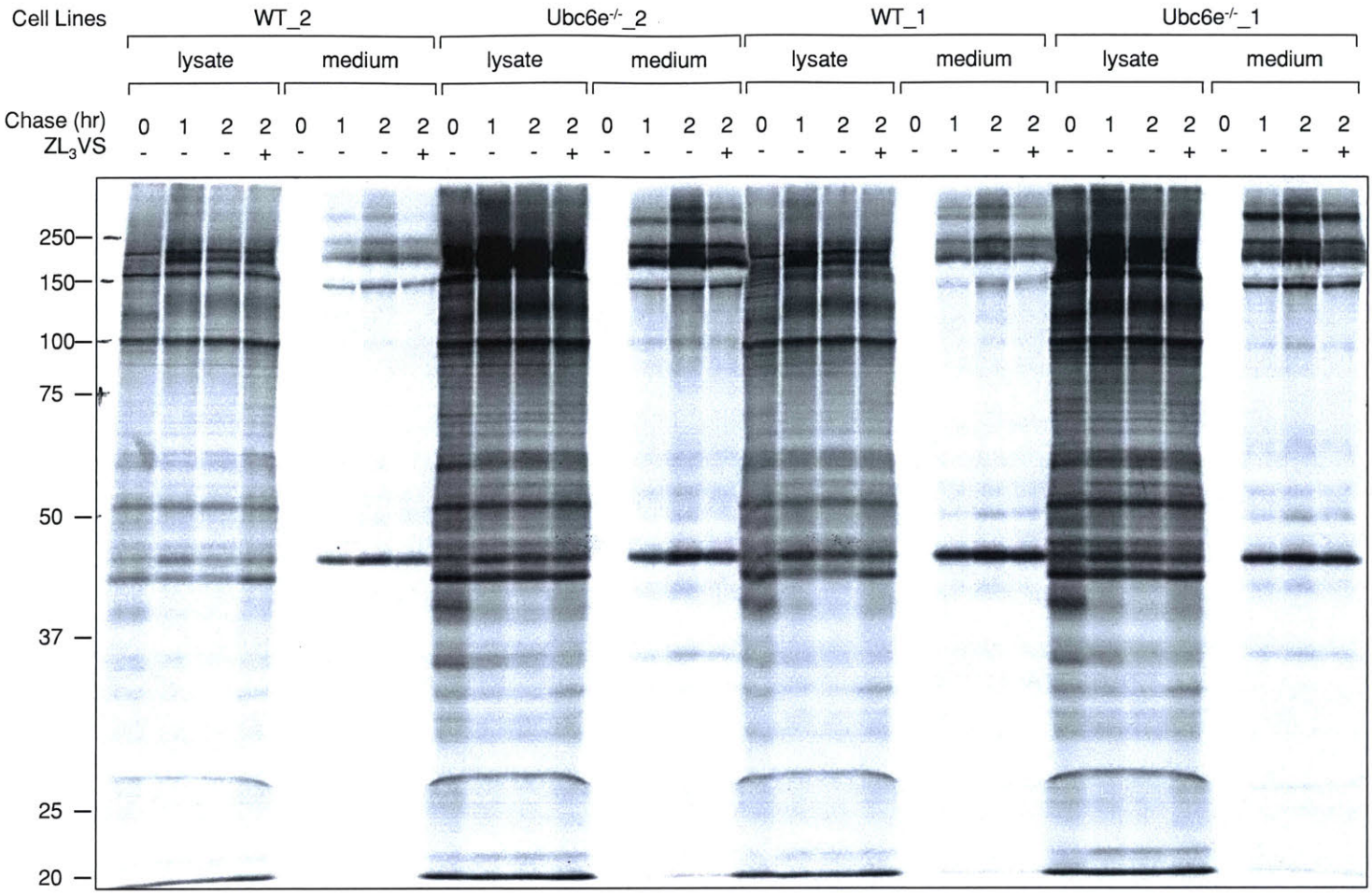


Figure S11.
Detection of total glycoproteins in cell lysates and in conditioned medium

2.3 Discussion

ER quality control serves to remove misfolded proteins and distinguishes between fully functional, folded proteins, which are spared from degradation, and terminally misfolded proteins, which are destroyed (Koenig and Ploegh, 2014; Smith et al., 2011). Only few cases are known in which an otherwise fully functional protein is degraded in the ER. A physiological example is the regulation of HMG-CoA-reductase (HMGR) by the intracellular levels of lanosterol. HMGR and lanosterol levels are inversely correlated, with excess HMGR being degraded by the ERAD pathway (Hampton and Garza, 2009). Pathogen-specific events can also control protein levels in the ER: the US2, US3 and US11 glycoproteins encoded by human cytomegalovirus (HCMV) co-opt the ERAD pathway to degrade Class I MHC heavy chains (Morito and Nagata, 2015). Our findings suggest a new layer of homeostatic control, in which ERAD activity itself is regulated by controlling the levels of select components through the action of UBC6e. The targets of this regulation are EDEM1, EDEM3, OS-9 and SEL1L (ERAD enhancers). Ablation of UBC6e causes up-regulation of active ERAD enhancers and so enhances clearance not only of terminally misfolded substrates but also of wild-type glycoproteins that fold comparatively slowly, such as tyrosinase.

ERAD enhancers are relatively short-lived proteins at steady state. Their mode of degradation and the molecular features that help explain their rather short half-life remained to be characterized. The present work on UBC6e provides insight into the underlying mechanisms. Deletion of UBC6e results in accumulation of ERAD enhancers without overt activation of the UPR (Figure 2.1B and 1D). This increase depends on the E2 enzymatic activity of UBC6e, its exact localization and on proper engagement in complex formation in its ER environment (Figure 2.1F, 4B and S6B). SEL1L is a possible candidate as the primary target for UBC6e-dependent degradation. However, UBC6e is an E2 ubiquitin-conjugating enzyme and E2s are not commonly considered to directly recognize their targets. The associated E3s, or the E2/E3 complexes are usually viewed as the elements responsible for substrate recognition. The identity of the E3 (or E3s) responsible for the observed down-regulation of EDEM1, OS-9 and Sel1L by UBC6e remains to be established.

Overexpression of catalytically inactive UBC6e in wild-type MEFs causes a dominant negative effect, whereas overexpression of wild-type UBC6e did not affect ERAD enhancer

levels (Figure 2.5D). Deletion of UBC6e (Figure 2.2F) does not change EDEM2 levels, suggesting that EDEM2 activity is presumably independent of ERAD activation. Careful regulation of the degradative capacity of these ERAD factors therefore operates at steady state in the absence of obvious stressors. When we replaced the C-terminal segment of UBC6e with that of the integral membrane protein CD4 or that of the tail-anchored protein UBC6, these chimeras were less active in normalizing ERAD enhancer levels in UBC6e$^{-/-}$ cells. The CD4 and UBC6 chimeras did not interact with OS-9 or SEL1L. The activity of such UBC6e complexes, which help establish the ERAD capacity of the ER, may thus have been compromised.

Even expression at ~10% of the endogenous level of UBC6e in UBC6e$^{-/-}$ cells re-establishes ERAD enhancer levels to those seen in wild-type cells: apparently very little UBC6e suffices to properly regulate them. Knock-down of UBC6e by means of siRNA delayed degradation of CFTRΔF508 and Class I MHC heavy chains (Burr et al., 2011; Younger et al., 2006). Knock-down mediated by siRNA or shRNA is rarely complete, yet may have been sufficient to examine the role of UBC6e in canonical ERAD, but would not have achieved the reduction required to observe control of ERAD enhancer levels by UBC6e. UBC6e$^{-/-}$ cells up-regulate selective ERAD-related factors with a concomitant increase in ERAD activity. The HRD1 ERAD complex may accept multiple E2s for degradation of terminally misfolded substrates, whereas degradation of folding-competent ERAD enhancers apparently requires only UBC6e as the functional E2 responsible.

Even if the ER engages in normal rates of cell type-appropriate protein synthesis and if overloading with misfolded proteins can be avoided, mechanisms of quality control should still apply at steady state and under normal conditions of growth, because errors in protein synthesis and protein folding cannot be avoided completely. Dysregulated ERAD activity as the result of deletion of key ERAD components such as HRD1 or SEL1L showed embryonic lethality in mice (Sun et al., 2014; Yagishita et al., 2005). MEFs deficient in HRD1 or SEL1L tend to accumulate canonical ERAD substrates and show activation of the UPR (Kaneko et al., 2010). In contradistinction, exaggerated ERAD activity may lead to loss-of-function through premature degradation of potentially functional proteins that have yet to attain their final and stable conformation (Noack et al., 2014). It has been difficult to test this notion experimentally due to the inability to initiate a more active ERAD pathway in the absence of triggering the UPR. The precise sequence of steps that control the mammalian ERAD pathway remains to be established and further adds to the challenge.

Here we addressed this question by exploring the mechanism of increased ERAD activity mediated by increased levels of ERAD components, as brought about by the ablation of UBC6e (Figure 2.7E). Our observations support the 'mannose timer hypothesis' (Fagioli, 2001; Helenius et al., 1997). In the course of the CNX/CRT cycle, terminally misfolded substrates are not distinguishable by any presently known criterion from folding intermediates, since both can carry the same G_1Man_9 N-glycan(s). EDEM2 and the other EDEM family proteins are functionally distinct, as inferred from cells that lack EDEM2 (Ninagawa et al., 2014). Glycoproteins whose N-linked glycans showed advanced mannose trimming could no longer attain their native structure, and were targeted for degradation. They could no longer productively engage the CNX/CRT cycle, nor could they serve as substrates for the folding sensor UGGT. Up-regulation of OS-9 and SEL1L facilitates the transfer of ERAD substrates from CNX to the HRD1-centered ERAD machinery and so enhances overall ERAD activity. We have shown that ERAD of glycoproteins is indeed regulated in a UBC6e-dependent manner. The non-glycoprotein ERAD pathway is far less well understood, presumably differs from the pathway we explored here and merits a separate and detailed analysis.

Under normal conditions and in the absence of obvious stressors, ERAD also promotes the degradation of proteins that fold comparatively slowly, such as proteins with many disulfide bonds and/or with many transmembrane segments (Guerriero and Brodsky, 2012). The folding intermediates of proteins that fall into these categories require prolonged engagement of the CNX/CRT cycle to allow completion of correct folding. We hypothesized that the increases in ERAD activity seen upon ablation of UBC6e might interfere with normal maturation of these slow-folding proteins and would thus compromise their function by reducing their levels. Indeed, tyrosinase showed accelerated degradation and failed to mature in both $UBC6e^{-/-}$ cells in tissue culture and in the skin of mice. Deletion of UBC6e does not affect the overall secretion of glycoproteins as assayed in vitro (Figure 2.7D and S11). Therefore, deletion of UBC6e showed only a minor effect in mice, compared with the ablation of other key ERAD components.

Transcription is usually considered the main contributor to the regulation of the ER quality control machinery. Regulation of ER homeostasis under the agency of ATF6, PERK and Ire1a in the course of the unfolded protein response (UPR) are well-established examples (Hetz et al., 2015), the manipulation of which usually requires the use of chemical stressors

such as thapsigargin, tunicamycin or DTT. How the activity of ERAD is controlled at steady state and in the absence of any ER stressors therefore remains an important question. We show that ERAD accelerators are members subject to quantity control in the ER by a novel UBC6e-dependent degradation pathway. The male sterility and auditory defects seen in UBC6e$^{-/-}$ mice might result from premature degradation of slow-folding proteins that participate in these processes. Post-transcriptional control of the ERAD machinery uncovered here is yet another means of homeostatic ER quality control.

2.4 Experimental Procedures

Plasmids, Antibodies, and Reagents

Plasmids, antibodies, and reagents are described in the Supplemental Experimental Procedures.

Cell Culture and Transfection

Cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% inactivated fetal bovine serum with or without 0.0007% v/v β-mercaptoethanol. Transient transfections were performed using Lipofectamine 2000 for HEK293T cells, Lipofectamine 3000 for MEFs according to the manufacturer's instructions (Invitrogen). Further experiments were performed 24 h after transfection.

Lentiviral Infection

Lentiviruses were generated in HEK293T cells. MEFs stably expressing various UBC6e mutants were created by lentiviral infection, followed by selection with geneticin or puromycin. See the Supplemental Experimental Procedures for details.

Mouse Experiments

All mouse experiments were performed in compliance with the protocols approved by the Massachusetts Institute of Technology Committee on Animal Care.

Fluorescence Microscopy

Images captured by confocal microscopes were viewed and processed by a PerkinElmer confocal system and Velocity software. Images were captured using a confocal microscope with a 63°- 1.40 N.A. of the Carl Zeiss Plan Apo oil objective. See the Supplemental Experimental Procedures for detail.

Immunoprecipitation and Western Blot

Cells were lysed in lysis buffer and subjected to immunoprecipitations with primary antibody. Immuno complexes were subjected to SDS-PAGE and analyzed by immunoblot. For details of detergents, conditions and antibodies used, see the Supplemental Experimental Procedures.

Pulse-Chase Analysis and Glucosidase Digestions

For pulse-labeling experiments, cells were starved for 45–60 min in methionine/cysteine-free DMEM at 37 °C, and labeled for 15-30 min at a concentration of 0.1mCi/ml of [^{35}S] methionine/cysteine, followed by incubation in complete MEF medium during the chase periods. Cell lysates were subjected to immunoprecipitation, followed by SDS-PAGE and examined by fluorography analysis. An EndoH digestion of radiolabeled substrates was performed after immunoprecipitation according to the manufacturer's instructions (New England Biolabs). See the Supplemental Experimental Procedures for details.

Sucrose Density Gradient Fractionation

Post-nuclear supernatants were applied to a 10%-60% linear sucrose gradient and centrifuged at 160,000 g for 16 hr. Fractions were collected from the top and used for further analysis.

Statistical Analysis

Welch's t-test was used for statistical analysis. P value less than 0.05 were considered statistically significant.

2.5 References

Burr, M.L., Cano, F., Svobodova, S., Boyle, L.H., Boname, J.M., and Lehner, P.J. (2011). HRD1 and UBE2J1 target misfolded MHC class I heavy chains for endoplasmic reticulum-associated degradation. Proc. Natl. Acad. Sci. U. S. A. *108*, 2034–2039.

Christianson, J.C., Shaler, T.A., Tyler, R.E., and Kopito, R.R. (2008). OS-9 and GRP94 deliver mutant α1-antitrypsin to the Hrd1?SEL1L ubiquitin ligase complex for ERAD. Nat. Cell Biol. *10*, 272–282.

Claessen, J.H.L., Mueller, B., Spooner, E., Pivorunas, V.L., and Ploegh, H.L. (2010). The transmembrane segment of a tail-anchored protein determines its degradative fate through dislocation from the endoplasmic reticulum. J. Biol. Chem. *285*, 20732–20739.

Cormier, J.H., Tamura, T., Sunryd, J.C., and Hebert, D.N. (2009). EDEM1 Recognition and Delivery of Misfolded Proteins to the SEL1L-Containing ERAD Complex. Mol. Cell *34*, 627–

633.

Dougan, S.K., Hu, C.-C.A., Paquet, M.-E., Greenblatt, M.B., Kim, J., Lilley, B.N., Watson, N., and Ploegh, H.L. (2011). Derlin-2-deficient mice reveal an essential role for protein dislocation in chondrocytes. Mol. Cell. Biol. *31*, 1145–1159.

Ellgaard, L., and Frickel, E.-M. (2003). Calnexin, calreticulin, and ERp57: teammates in glycoprotein folding. Cell Biochem. Biophys. *39*, 223–247.

Fagioli, C. (2001). Glycoprotein Quality Control in the Endoplasmic Reticulum. Mannose trimming by endoplasmic reticulum mannosidase i times the proteasomal degradation of unassembled immunoglobulin subunits. J. Biol. Chem. *276*, 12885–12892.

Guerriero, C.J., and Brodsky, J.L. (2012). The Delicate Balance Between Secreted Protein Folding and Endoplasmic Reticulum-Associated Degradation in Human Physiology. Physiol. Rev. *92*, 537–576.

Hampton, R.Y., and Garza, R.M. (2009). Protein Quality Control as a Strategy for Cellular Regulation: Lessons from Ubiquitin-Mediated Regulation of the Sterol Pathway. Chem. Rev. *109*, 1561–1574.

Helenius, a., Trombetta, E.S., Hebert, D.N., and Simons, J.F. (1997). Calnexin, calreticulin and the folding of glycoproteins. Trends Cell Biol. *7*, 193–200.

Hetz, C., Chevet, E., and Oakes, S.A. (2015). Proteostasis control by the unfolded protein response. Nat. Cell Biol. *17*, 829–838.

Hirao, K., Natsuka, Y., Tamura, T., Wada, I., Morito, D., Natsuka, S., Romero, P., Sleno, B., Tremblay, L.O., Herscovics, A., et al. (2006). EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming. J. Biol. Chem. *281*, 9650–9658.

Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L.O., Herscovics, a, and Nagata, K. (2001). A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation. EMBO Rep. *2*, 415–422.

Iida, Y., Fujimori, T., Okawa, K., Nagata, K., Wada, I., and Hosokawa, N. (2011). SEL1L Protein Critically Determines the Stability of the HRD1-SEL1L Endoplasmic Reticulum-associated Degradation (ERAD) Complex to Optimize the Degradation Kinetics of ERAD Substrates. J. Biol. Chem. *286*, 16929–16939.

Kaneko, M., Koike, H., Saito, R., Kitamura, Y., Okuma, Y., and Nomura, Y. (2010). Loss of HRD1-Mediated Protein Degradation Causes Amyloid Precursor Protein Accumulation and Amyloid- Generation. J. Neurosci. *30*, 3924–3932.

Koenig, P.A., and Ploegh, H.L. (2014). Protein quality control in the endoplasmic reticulum. In F1000Prime Rep,.

Koenig, P.-A., Nicholls, P.K., Schmidt, F.I., Hagiwara, M., Maruyama, T., Frydman, G.H., Watson, N., Page, D.C., and Ploegh, H.L. (2014). The E2 Ubiquitin-conjugating Enzyme UBE2J1 Is Required for Spermiogenesis in Mice. J. Biol. Chem. *289*, 34490–34502.

Lenk, U., Yu, H., Walter, J., Gelman, M.S., Hartmann, E., Kopito, R.R., and Sommer, T. (2002). A role for mammalian Ubc6 homologues in ER-associated protein degradation. Journal of Cell Science *115*, 3007-3014.

Stein, A., Ruggiano, A., Carvalho, P., and Rapoport, Tom A. (2014). Key Steps in ERAD of Luminal ER Proteins Reconstituted with Purified Components. Cell *158*, 1375-1388.

Menon, M.B., Tiedje, C., Lafera, J., Ronkina, N., Konen, T., Kotlyarov, A., and Gaestel, M. (2013). Endoplasmic reticulum-associated ubiquitin-conjugating enzyme Ube2j1 is a novel substrate of MK2 (MAPKAP kinase-2) involved in MK2-mediated TNFα production. Biochem. J. *456*, 163–172.

Molinari, M. (2003). Role of EDEM in the Release of Misfolded Glycoproteins from the Calnexin Cycle. Science (80-.). *299*, 1397–1400.

Morito, D., and Nagata, K. (2015). Pathogenic Hijacking of ER-Associated Degradation: Is ERAD Flexible? Mol. Cell *59*, 335–344.

Mueller, B., Lilley, B.N., and Ploegh, H.L. (2006). SEL1L, the homologue of yeast Hrd3p, is involved in protein dislocation from the mammalian ER. J. Cell Biol. *175*, 261–270.

Mueller, B., Klemm, E.J., Spooner, E., Claessen, J.H., and Ploegh, H.L. (2008). SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. Proc. Natl. Acad. Sci. U. S. A. *105*, 12325–12330.

Ninagawa, S., Okada, T., Sumitomo, Y., Kamiya, Y., Kato, K., Horimoto, S., Ishikawa, T., Takeda, S., Sakuma, T., Yamamoto, T., et al. (2014). EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. J. Cell Biol. *206*, 347–356.

Noack, J., Bernasconi, R., and Molinari, M. (2014). How Viruses Hijack the ERAD Tuning Machinery. J. Virol. *88*, 10272–10275.

Oda, Y. (2003). EDEM As an Acceptor of Terminally Misfolded Glycoproteins Released from Calnexin. Science (80-.). *299*, 1394–1397.

Olivari, S., Galli, C., Alanen, H., Ruddock, L., and Molinari, M. (2005). A novel stress-induced EDEM Variant regulating endoplasmic reticulum-associated glycoprotein degradation. J. Biol. Chem. *280*, 2424–2428.

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. Science *189*, 867.

Reggiori, F., Monastyrska, I., Verheije, M.H., Calì, T., Ulasli, M., Bianchi, S., Bernasconi, R., de Haan, C. a M., and Molinari, M. (2010). Coronaviruses Hijack the LC3-I-positive EDEMosomes, ER-derived vesicles exporting short-lived ERAD regulators, for replication. Cell Host Microbe *7*, 500–508.

Satoh, T., Chen, Y., Hu, D., Hanashima, S., Yamamoto, K., and Yamaguchi, Y. (2010). Structural Basis for Oligosaccharide Recognition of Misfolded Glycoproteins by OS-9 in ER-Associated Degradation. Mol. Cell *40*, 905–916.

Seth, R.B., Sun, L., Ea, C.-K., and Chen, Z.J. (2005). Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF-κB and IRF3. Cell *122*, 669–682.

Sifers, R.N., Brashears-Macatee, S., Kidd, V.J., Muensch, H., and Woo, S.L.C. (1988). A

frameshift mutation results in a truncated ??1-antitrypsin that is retained within the rough endoplasmic reticulum. J. Biol. Chem. *263*, 7330–7335.

Smith, M.H., Ploegh, H.L., and Weissman, J.S. (2011). Road to Ruin: Targeting Proteins for Degradation in the Endoplasmic Reticulum. Science (80-.). *334*, 1086–1090.
Stein, A., Ruggiano, A., Carvalho, P., and Rapoport, Tom A. (2014). Key Steps in ERAD of Luminal ER Proteins Reconstituted with Purified Components. Cell *158*, 1375-1388.
Sun, S., Shi, G., Han, X., Francisco, A.B., Ji, Y., Mendonça, N., Liu, X., Locasale, J.W., Simpson, K.W., Duhamel, G.E., et al. (2014). Sel1L is indispensable for mammalian endoplasmic reticulum-associated degradation, endoplasmic reticulum homeostasis, and survival. Proc. Natl. Acad. Sci. U. S. A. *111*, E582–E591.

Svedine, S., Wang, T., Halaban, R., and Hebert, D.N. (2004). Carbohydrates act as sorting determinants in ER-associated degradation of tyrosinase. J. Cell Sci. *117*, 2937–2949.

Toyofuku, K., Wada, I., Spritz, R. a, and Hearing, V.J. (2001). The molecular basis of oculocutaneous albinism type 1 (OCA1): sorting failure and degradation of mutant tyrosinases results in a lack of pigmentation. Biochem. J. *355*, 259–269.

Tsao, Y.S., Ivessa, N.E., Adesnik, M., Sabatini, D.D., and Kreibich, G. (1992). Carboxy terminally truncated forms of ribophorin I are degraded in pre-Golgi compartments by a calcium-dependent process. J. Cell Biol. *116*, 57–67.

Wang, N., and Hebert, D.N. (2006). Tyrosinase maturation through the mammalian secretory pathway: Bringing color to life. Pigment Cell Res. *19*, 3–18.

Wiertz, E.J.H.J., Jones, T.R., Sun, L., Bogyo, M., Geuze, H.J., and Ploegh, H.L. (1996). The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. Cell *84*, 769–779.

van Wijk, S.J.L., and Timmers, H.T.M. (2010). The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. FASEB J. *24*, 981–993.

Xu, C., and Ng, D.T.W. (2015). Glycosylation-directed quality control of protein folding. Nat. Rev. Mol. Cell Biol. needs page numbers etc

Yagishita, N., Ohneda, K., Amano, T., Yamasaki, S., Sugiura, A., Tsuchimochi, K., Shin, H., Kawahara, K. -i., Ohneda, O., Ohta, T., et al. (2005). Essential Role of Synoviolin in Embryogenesis. J. Biol. Chem. *280*, 7909–7916.

Yoshida, H., Matsui, T., Hosokawa, N., Kaufman, R.J., Nagata, K., and Mori, K. (2003). A time-dependent phase shift in the mammalian unfolded protein response. Dev. Cell *4*, 265–271.

Younger, J.M., Chen, L., Ren, H.-Y., Rosser, M.F.N., Turnbull, E.L., Fan, C.-Y., Patterson, C., and Cyr, D.M. (2006). Sequential Quality-Control Checkpoints Triage Misfolded Cystic Fibrosis Transmembrane Conductance Regulator. Cell *126*, 571–582.

Chapter 3: Using VHHs to study the function and control of the E2 Ubiquitin-conjugating enzyme UBC6e

Jingjing Ling[1], Ross Cheloha[1], Nicholas MCCaul[1], Zhen-Yu Sun[2], Tao Fang[1], Gerhard Wagner[2] and Hidde Ploegh[1,*]

1 Boston Children's Hospital, Boston, MA 02115

2 Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA

*Correspondence: Hidde.Ploegh@childrens.harvard.edu
To be submitted

All experiments were designed by Jingjing Ling. Experiments were assisted by other authors listed.

3.1 Introduction

Most secretory proteins in mammalian cells are folded and modified in the Endoplasmic Reticulum (ER) before export and secretion. However, many proteins are unable to reach their natively folded state despite multiple rounds of chaperone-assisted folding cycles, and need to be degraded. Misfolded proteins and/or unfolded proteins may be transported back into the cytosol to be degraded by the proteasome. This process is termed ER-associated degradation (ERAD). There are many components involved in the ERAD process, including the sensor proteins that recognize the unfolded substrates, the dislocon that transports substrates from the ER back to the cytosol, the ubiquitination machinery that poly-ubiquitinates the substrates for degradation, as well as the proteasome that finally degrades the substrates.

HRD1 is a well-studied E3 ubiquitin ligase that is considered central to ERAD. It has been implicated in the degradation of both endogenous proteins such as 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-R) (Hampton et al., 1996) and MHC class I heavy chains (Burr et al., 2011a), as well as ERAD substrates such as mutants that are unable to achieve their native fold (Bordallo et al., 1998). HRD1 acts as a part of the machinery that transports substrates from the ER lumen to the cytosol. Rapoport *et. al* have reported that auto-ubiquitination of HRD1 triggers retrotranslocation of substrates(Baldridge and Rapoport, 2016) and recently reported a cryo-EM structure of *Saccharomyces cerevisiae* HRD1 in complex with Hrd3, in which HRD1 formed an membrane-embedded aqueous channel (Schoebel et al., 2017). HRD1 also contains a RING-H2 domain, which allows it to act as an E3(Bordallo and Wolf, 1999). E3 enzymes require collaboration with E2 ubiquitin conjugating enzymes to ubiquitinate substrates. Human HRD1 has *in vitro* ubiquitination activity in conjunction with the E2 UBC7(Kikkert et al., 2004). In addition, E2 UBC6e, also known as UBE2J1, has been implicated -together with HRD1- in the degradation of MHC class I molecules (Burr et al., 2011b; Mueller et al., 2008; Mueller et al., 2006).

In mammalian cells HRD1 requires stoichiometric association with its adaptor protein Hrd3, also known as SEL1L, to be stable and functional(Iida et al., 2011; Sun et al., 2014). The HRD1-SEL1L complex has been implicated in the immune response, where it is required for B cell development through its involvement in the degradation of pre-BCR complex (Ji et al., 2016). Deletion of HRD1 also inhibits T cell proliferation (Xu et al., 2016) and diminishes expression of MHC II molecules in dendritic cells. (Yang et al., 2014) The importance of ERAD for operation of the immune system underscores a role not only in quality control, but also for targeted degradation of select molecules when needed.

Immunopurification of HA-TEV-SEL1L was performed to identify the members of HRD1-SEL1L complex, and showed UBC6e to be the predominant co-immunoprecipitated E2 (Mueller et al., 2008). This led to the proposal that UBC6e was the E2 acting in concert with HRD1. However, knockout of UBC6e led to an accumulation of functional chaperones involved in targeting folding intermediates for ERAD, and therefore an increase in ERAD activity (Hagiwara et al., 2016). UBC6e may thus be involved in fine-tuning of ERAD activity. We therefore sought to better understand how this E2 is functionally controlled.

UBC6e is one of the two ER tail-anchored E2 that are mammalian homologs of the *Saccharomyces cerevisiae* E2 UBC6p (Lenk et al., 2002; Lester et al., 2000). Both homologs, UBC6 and UBC6e, are implicated in ERAD, as expression of catalytically inactive, dominant-negative mutants delays the degradation of ERAD substrates (Lenk et al., 2002). The ER membrane tail anchor is essential for the proper function of UBC6e, as overexpression of ΔTM mutant does not restore the function of UBC6e in UBC6e$^{-/-}$ cells (Hagiwara et al., 2016). UBC6e is cleaved by Enterovirus 71 3C^{pro} in its C-terminal region to be released from ER, and therefore perturbs normal ERAD during viral infection (Wang et al., 2017). Genetic knock-out of UBC6e is not lethal in mice, but causes sterility in UBC6e-/- homozygous males (Koenig et al., 2014).

There is a known stress-sensitive phosphorylation site Ser184 on UBC6e. During ER stress, protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK) is thought to phosphorylate UBC6e(Oh et al., 2006); during inhibition of protein synthesis, Mitogen-activated protein kinase 2 (MK-2) has been reported to be the kinase responsible. (Menon et al., 2013) (Menon et al., 2013) However, the functional role of phosphorylation remains unclear. Phosphorylation may affect the rate of ubiquitin (Ub) loading onto UBC6e as well as the rate at which it is degraded by the proteasome (Elangovan et al., 2017; Oh et al., 2006).

VHHs, also known as nanobodies, are the binding domain of heavy-chain only antibodies found in camelids. They retain the binding affinities and specificities of conventional antibodies, but do not require the formation of disulfide bonds or glycosylation to be active. They are usually between 10 to 15 kDa in size, and can be recombinantly expressed in high yields, both in the cytosol and in the periplasm in *E. coli*. VHHs can bind to their targets in mammalian cytosol. We have used VHHs to block nuclear import of nuclear protein (NP) during influenza infection(Hanke et al., 2016), protect cells from VSV infection(Hanke and Schmidt, 2017), as well as to track the formation of formation of Asc filaments during inflammasome activation(Schmidt et al., 2016). (Schmidt et al., 2016). In the current study, we aimed to generate VHHs that can target UBC6e so as to allow us to study the function of UBC6e.

We generated 7 VHHs that can bind to UBC6e, with one VHH (VHH05) exhibiting nanomolar binding affinity. VHH05 enhances the enzymatic activity of UBC6e *in vitro*, and blocks stress-induced phosphorylation of UBC6e when expressed in cells. VHH05 may serve as a tool to study the function of phosphorylation and allosteric control of E2 enzyme activity.

3.2 Results

Identifying VHHs that specifically recognize UBC6e

We immunized an alpaca with purified recombinant UBC6e (1-234) protein to elicit an antibody response in the animal. UBC6e (1-234) was produced using the same construct used to generate a polyclonal rabbit serum (Mueller et al., 2008). A strong antibody response was elicited in the animal after immunization (Figure 3.1a). The VHH-coding regions were selectively amplified from lymphocyte cDNA library and inserted into a pD phagemid plasmid. After two rounds of panning, 7 VHHs with distinct CDR regions were identified. Among all VHHs identified, only VHH05 exhibits low nanomolar affinity for UBC6e, and therefore we chose to focus on this VHH for subsequent studies (Table 3.1). VHH05 immunoprecipitated recombinant UBC6e, but was inefficient in the immunoprecipitation of the UBC6e homolog UBC6, the other ER-related E2 UBC7 and the other E2s tested (Figure 3.1b). We then modified VHHs equipped at the C-terminus with the sortase LPETGG motif and a hexahistadine tag with G_3-biotin, using previously developed methods and used them for immunoblots (Popp and Ploegh, 2011). While many previously reported VHHs, like VHH "enhancer", specific for GFP (Kirchhofer et al., 2010), only recognize their targets in their native conformations, VHH05 can be used to detect UBC6e after SDS-denaturation (Figure 3.1c).

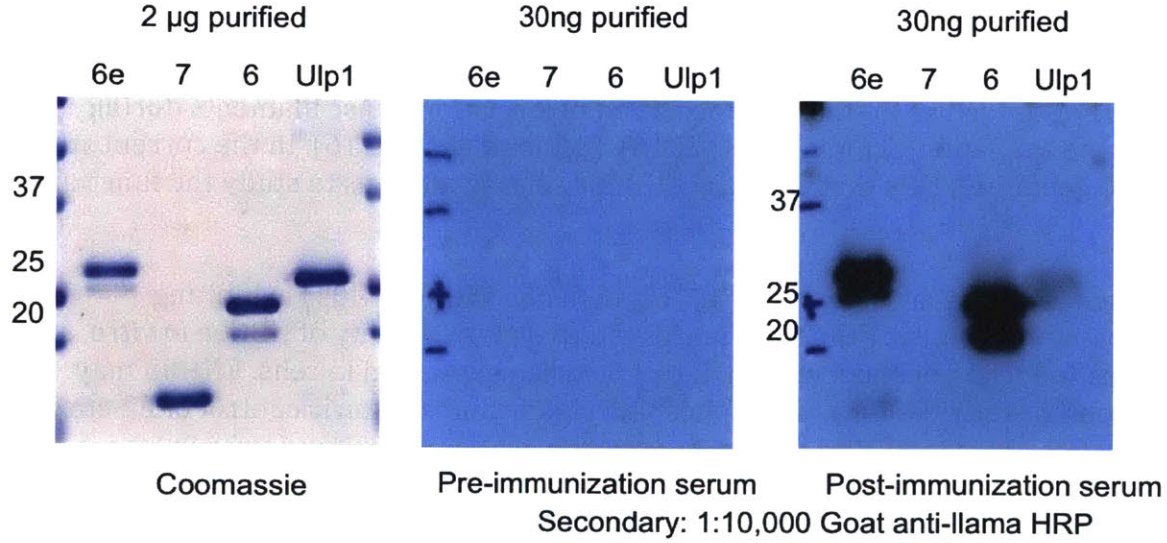


Figure 3.1a. An antibody response was elicited in the alpaca immunized with purified recombinant UBC6e(1-234). The pre- and post-immunization serum were used in a 1:1000 dilution to blot for purified UBC6e, UBC6 and UBC7 and SUMO protease Ulp1. UBC6(1-197) and UBC7 were injected together with UBC6e during alpaca immunization. SUMO protease Ulp1 was not injected. The alpaca developed antibody response against UBC6e(1-234) and UBC6(1-197), but not for UBC7.

Table 3.1. Dissociation constants for the 7 unique VHHs identified.

Kd(nM) / VHH	UBC6e (1-234)	UBC6e (1-197)
UBC6e 02	80 ± 20	60 ± 30
UBC6e 03	160 ± 40	280 ± 60
UBC6e 05	5 ± 2	6 ± 2
UBC6e 08	900 ± 200	600 ± 300
UBC6e 11	2200 ± 1500	4600 ± 1300
UBC6e 16	800 ± 200	No Binding
UBC6e 18	1000 ± 700	1500 ± 600

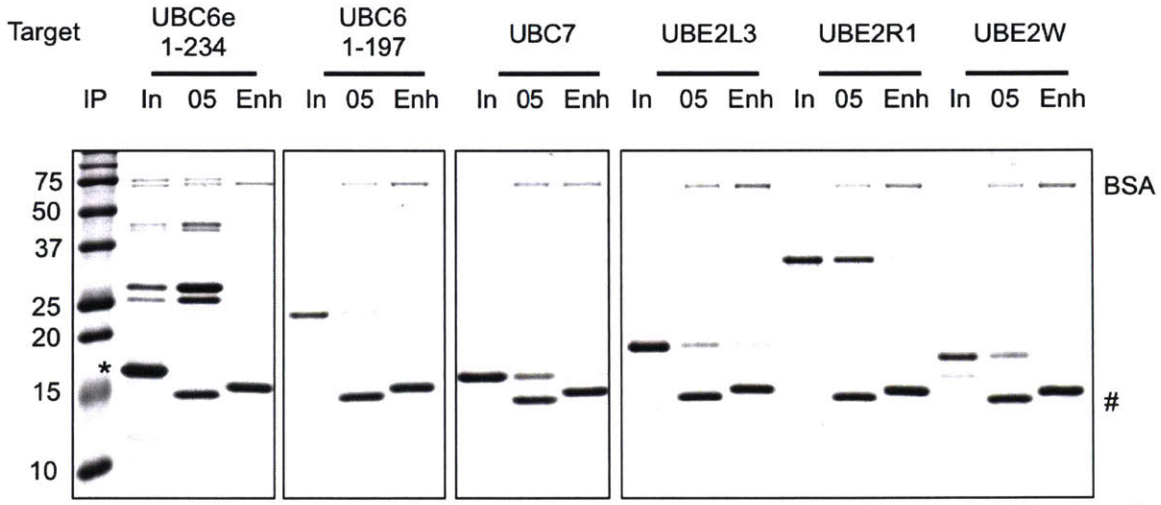


Figure 3.1b. VHH 6E05 specifically immunoprecipitates recombinant UBC6e. Purified proteins were incubated with sepharose beads conjugated with VHH05 or anti-GFP VHH enhancer (Kirchhofer et al., 2010). In: Input. * : cleaved SUMO from expression as a SUMO fusion. #: VHH 6E05 and VHH enhancer.

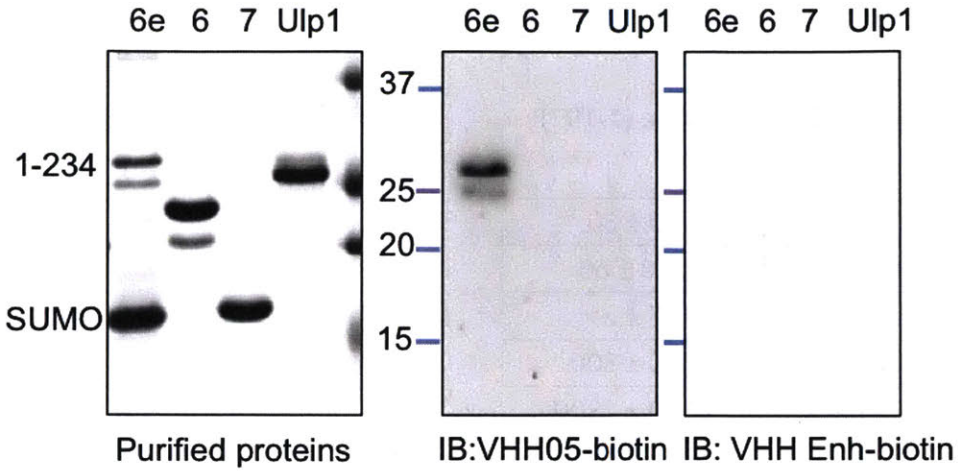


Figure 3.1c. VHH 6E05 immunoblots for UBC6e after SDS denaturation. Purified proteins were blotted with VHH05-biotin or VHH enhancer-biotin. Strepatavidin-HRP was used as the secondary antibody for detection.

In contrast to conventional antibodies, VHHs do not require the formation of stable disulfide bonds between heavy chain and light chain to fold and function. They can be expressed in good yield in the reducing environment of the cytosol of *E. coli*, and when purified, retain their binding properties. Therefore, VHHs should also be able to bind to their targets in the reducing environment in the cytosol of mammalian cells. To test the function of these VHHs in the cytosol, we expressed the indicated VHHs equipped with HA tags in mouse embryonic fibroblasts (MEFs), and immunoprecipitated with anti-HA beads. Endogenous UBC6e was co-immunoprecipitated with expressed VHH05-HA (Figure 3.1d). The other VHHs were not able to enrich for UBC6e in this manner, possibly due to their lower affinities.

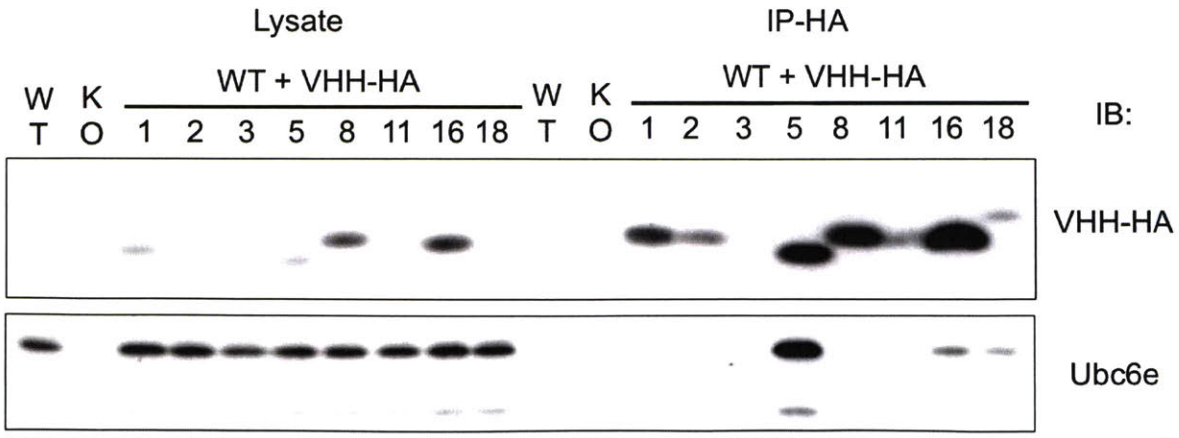


Figure 3.1d. When expressed in cytosol, HA-tagged VHH 6E05 can co-immunoprecipitates endogenous UBC6e. Mouse embryonic fibroblasts (MEFs) were transduced with VHHs listed in Table 3.1. Lysates were collected in 1% NP-40 in HBS and immunoprecipitated with HA antibody-conjugated sepharose beads. WT: wild type. KO: UBC6e$^{-/-}$ cells.

VHH05 accelerates the E2 loading of UBC6e and E3-catalyzed UBC6e-Ub hydrolysis

E2 ubiquitin-conjugating enzymes, interposed beween E1 and E3-type activities, execute a key step of ubiquitination pathway. Ubiquitin-activating enzyme E1 activates the C-terminus of free ubiquitin and forms a thioester between ubiquitin and the active site cysteine on E1. This ubiquitin is then transferred onto a ubiquitin-conjugating enzyme, E2, to form a second thioester. The loaded E2 can either transfer its ubiquitin onto a HECT type E3 ubiquitin ligase in thioester linkage and subsequently onto the substrates, or be recruited by a RING-type E3 to the substrate and directly transfer the ubiquitin from the E2 onto the substrates (Figure 3.2a).

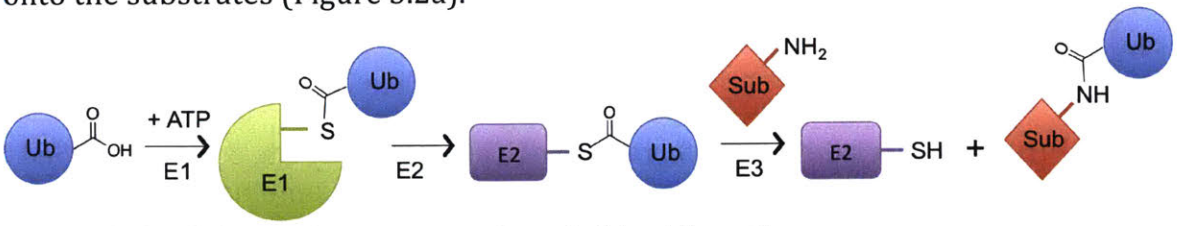
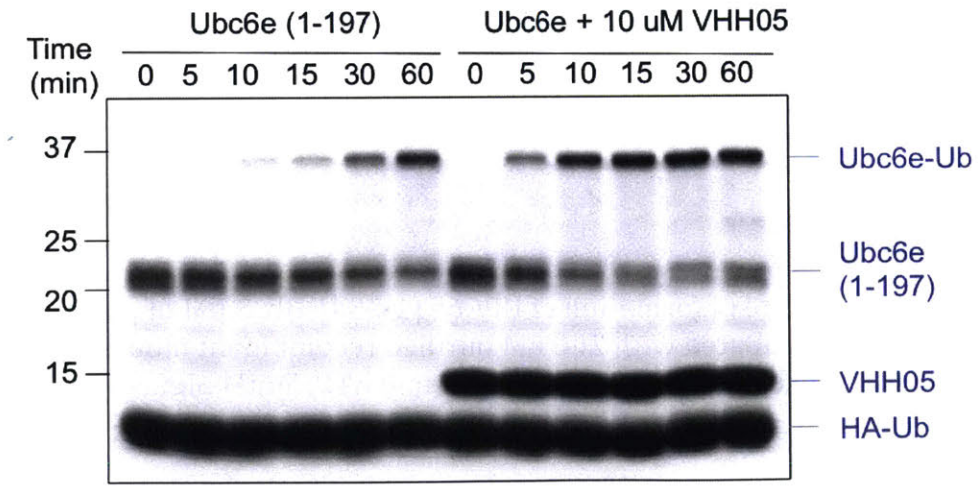


Figure 3.2a. Schematic representation of ubiquitin pathway.

The loading of ubiquitin onto UBC6e can be achieved using purified recombinant proteins, and the UBC6e-Ub thioester is stable upon FPLC purification and non-reducing SDS-PAGE analysis. Because UBC6e (1-234) migrates as two distinct bands on SDS-PAGE, a new truncation (1-197) was designed based on the construct used for the crystallization of its homolog UBC6 (pdb: 2F4W) (Sheng et al., 2012). Both the UBC6e (1-197) and (1-234) constructs can be loaded with ubiquitin by recombinant E1 (Figure 3.2b and 3.2c). Upon addition of VHH05, the rate of loading is accelerated by approximately 3-fold. The acceleration is achieved with UBC6e-specific VHH05, but not with VHH enhancer, which is specific for GFP. VHH05 has no effect on the loading of UBC6e homolog UBC6 or UBC7 with Ub (Figure 3.2d).



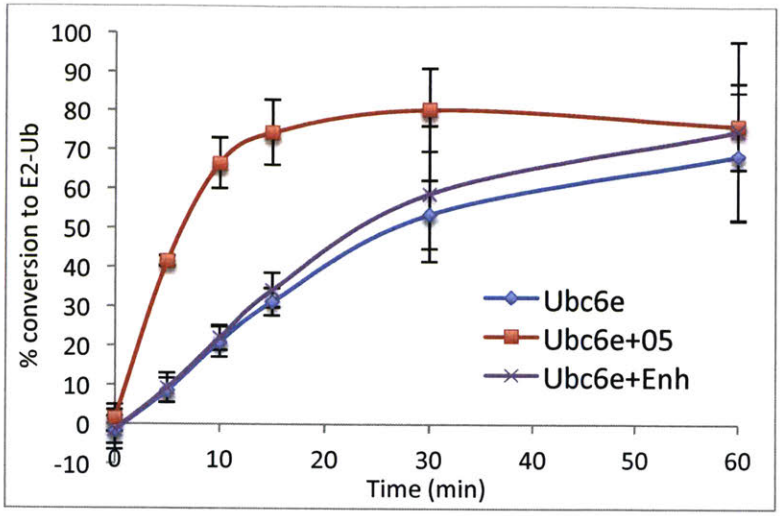


Figure 3.2b. VHH05 enhances the rate of formation of UBC6e (1-197)-Ub. 3 μM of purified recombinant UBC6e (1-197) was incubated with 200 nM purified recombinant mE1, 10 μM HA-Ub in the presence of 50 mM Tris pH 7.4, 2 mM ATP, 5 mM $MgCl_2$ and 200 μM DTT at 37°C. VHHs were added at 10 μM concentration. Samples were quenched with 2x SDS sample buffer with no reducing agents. Quantification was conducted by normalizing the band intensity of UBC6e-Ub band to that of UBC6e (1-197) at t=0 minutes. Error Bars represent standard deviation (n=3).

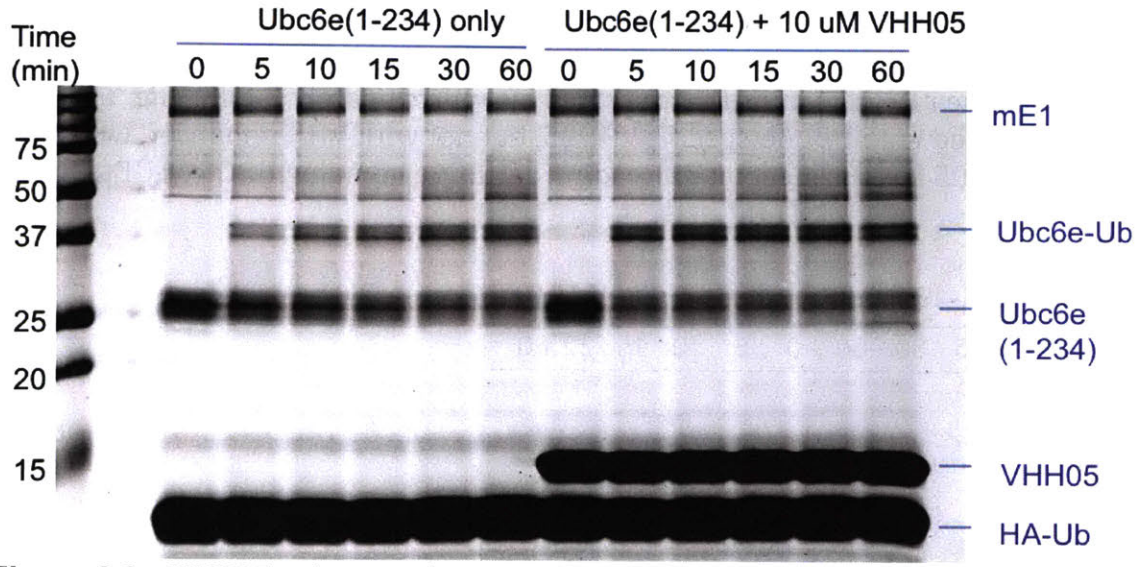


Figure 3.2c. VHH05 enhances the rate of formation of UBC6e (1-234)-Ub. 3 μM of purified recombinant UBC6e(1-234) was incubated with 200 nM purified recombinant mouse E1 (mE1), 10 μM HA-Ub in the presence of 50 mM Tris pH 7.4, 2 mM ATP, 5 mM $MgCl_2$ and 200 μM DTT at 37°C. VHHs were added at 10 μM concentration.

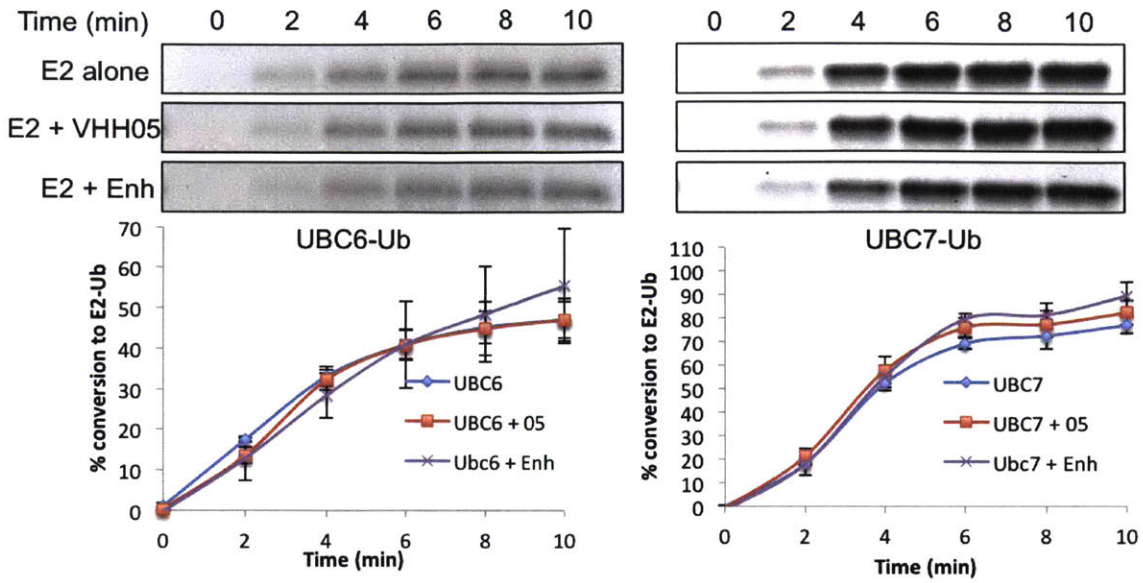


Figure 3.2d. VHH05 does not affect the rate of formation of UBC6-Ub or UBC7-Ub. 3 μM of purified recombinant UBC6(1-197) or UBC7 was loaded and characterized as described in Figure 3.2c. Only the E2-Ub bands from SDS-PAGE are shown in this panel.

We then investigated whether the hydrolysis of UBC6e-Ub is likewise accelerated in the presence of VHH05. UBC6e-Ub is sufficiently stable in pH 7.0 phosphate buffer to be purified using size exclusion FPLC. However, addition of RING domains from E3 enzymes has been reported to significantly increase the hydrolysis of many E2-Ub thioesters (Stewart et al., 2016). We therefore used recombinant GST-HRD1 RING domain (272-343) to explore thioester hydrolysis. The molecular weight of GST-HRD1 RING is very close to that of UBC6e-Ub, leading to an overlap in bands. To observe the behavior of UBC6e-Ub without such interference, we expressed UBC6e (1-197) with a sortase-compatible LPSTGG motif and labeled it with G_3-TAMRA. We could successfully load this TAMRA-labeled UBC6e with ubiquitin at rates and with yields similar to the unlabeled counterpart (data not shown). In the absence of GST-HRD1 RING, addition of VHH05 has little effect on the rate of hydrolysis of UBC6e-Ub (Figure 3.3a). There is a slight increase in the formation of UBC6e with 2 ubiquitins attached, and a slight increase in the formation of DTT-resistant UBC6e-Ub, products which we attribute to auto-ubiquitination of Ubc6e on a lysine residue. On the other hand, addition of GST-HRD1 RING significantly enhanced the rate of UBC6e-Ub hydrolysis. Addition of VHH05 further increased the rate of disappearance of UBC6e-Ub (Figure 3.3b). Formation of UBC6e with 2 ubiquitins was abolished upon addition of GST-HRD1 RING (data not shown). We suggest that VHH05 binding to UBC6e may allow opening of its active site, thus accelerating both E2 loading and unloading. Addition of 4 different control VHHs (anti-CD38 VHH, VHH enhancer, anti-MHCII VHH7 and anti-Flu NP VHH68) all delayed the hydrolysis of UBC6e-Ub in the presence of HRD1 RING (Data not shown). Since E2/E3 interactions are usually too weak to be purified (Stewart et al., 2016), inclusion of irrelevant proteins might interfere with their binding.

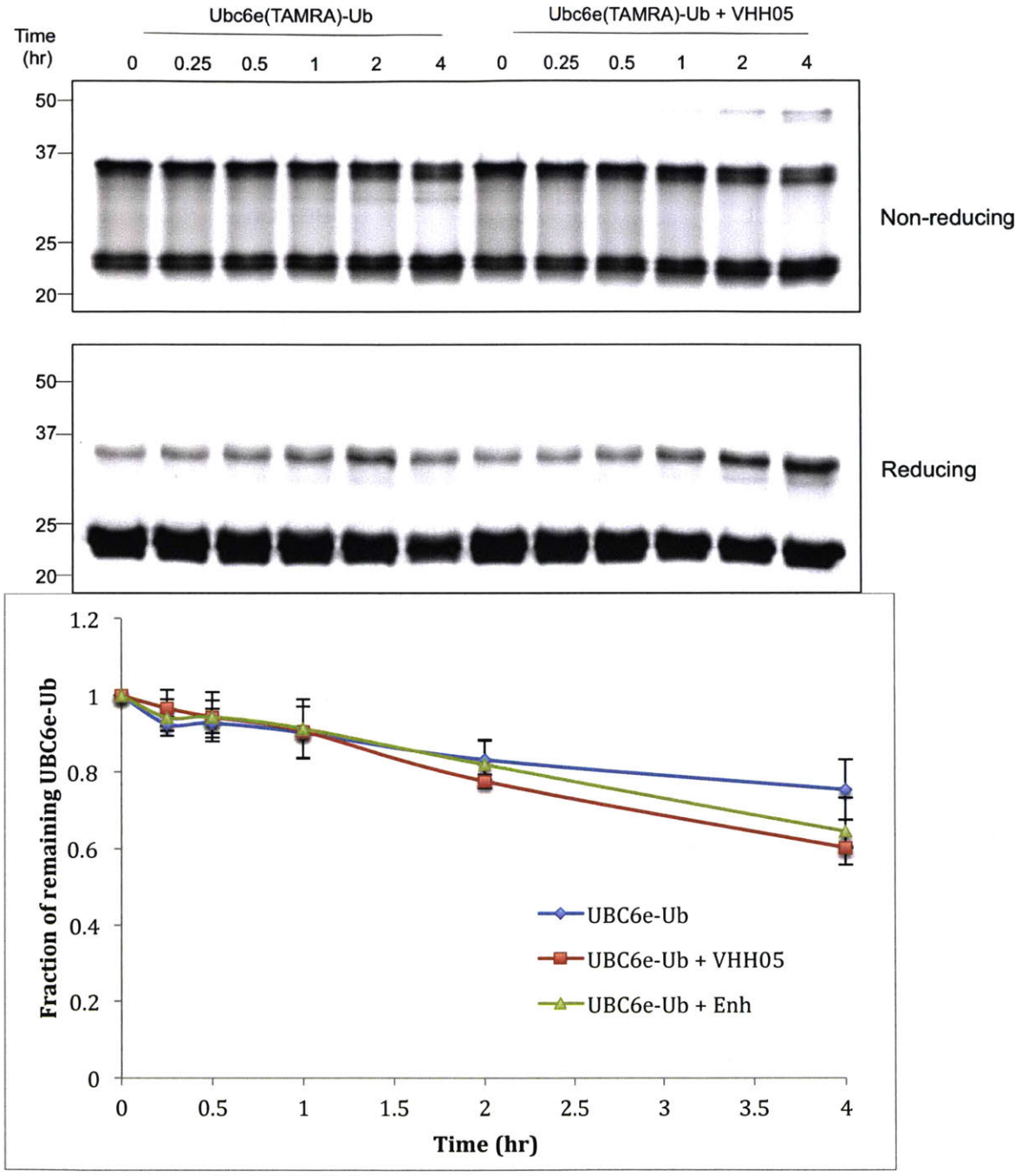


Figure 3.3a. VHH05 does not affect the rate of hydrolysis of UBC6e-Ub in the absence of GST-HRD1 RING. Approximately 3 μM of purified UBC6e(1-197)-Ub was incubated in TBS at 37°C. VHHs were added at 10 μM concentration. Samples were quenched in 2x SDS sample buffer with no reducing agent and analyzed with 15% SDS-PAGE. Gels were imaged for absorbance at 555 nm. Quantification was conducted by normalizing the band intensity of E2-Ub band to that at t=0 minutes. Error Bars represent standard deviation (n=3).

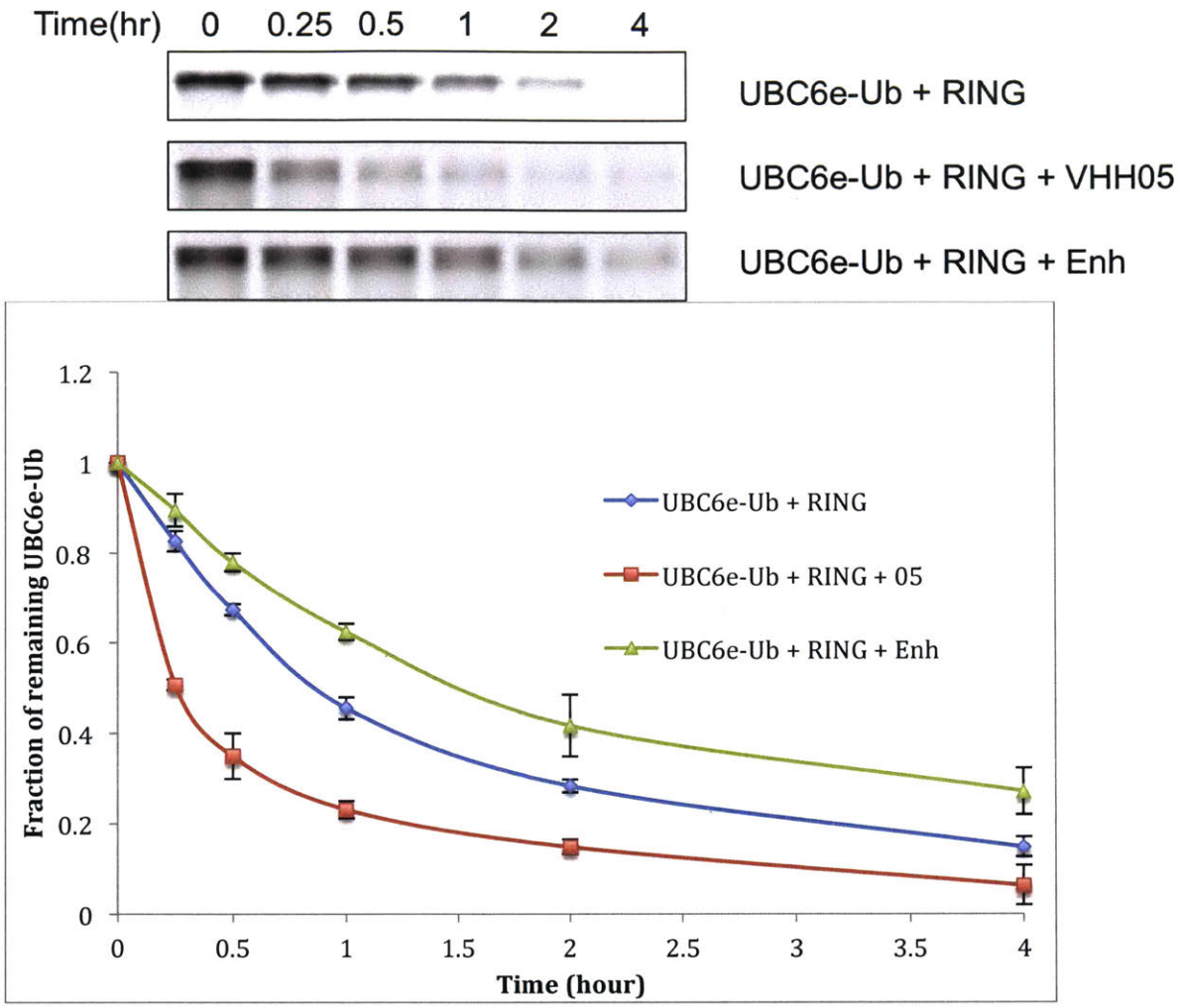


Figure 3.3b. VHH05 enhances the rate of hydrolysis of UBC6e-Ub in the presence of GST-HRD1 RING. Approximately 3 μM of purified UBC6e(1-197)-Ub was incubated in TBS at 37°C. VHHs were added at 10 μM concentration. Samples were quenched in 2x SDS sample buffer with no reducing agent and analyzed by 15% SDS-PAGE. Gels were imaged for absorbance at 555 nm. Quantification was conducted by normalizing the band intensity of the E2-Ub band to that at t=0 minutes. Error Bars represent standard deviation (n=3).

VHH05 binds to both UBC6e and E2-loaded UBC6e-Ub

For VHH05 to have an effect on the hydrolysis of UBC6e-Ub, it must bind to it. We have confirmed this by immunoprecipitation. UBC6e(1-197) and UBC6e(1-197)-Ub both bind to VHH05 but not VHH enhancer, while neither UBC7-Ub nor UBC7 binds to VHH05 or VHH enhancer (Figure 3.4a).

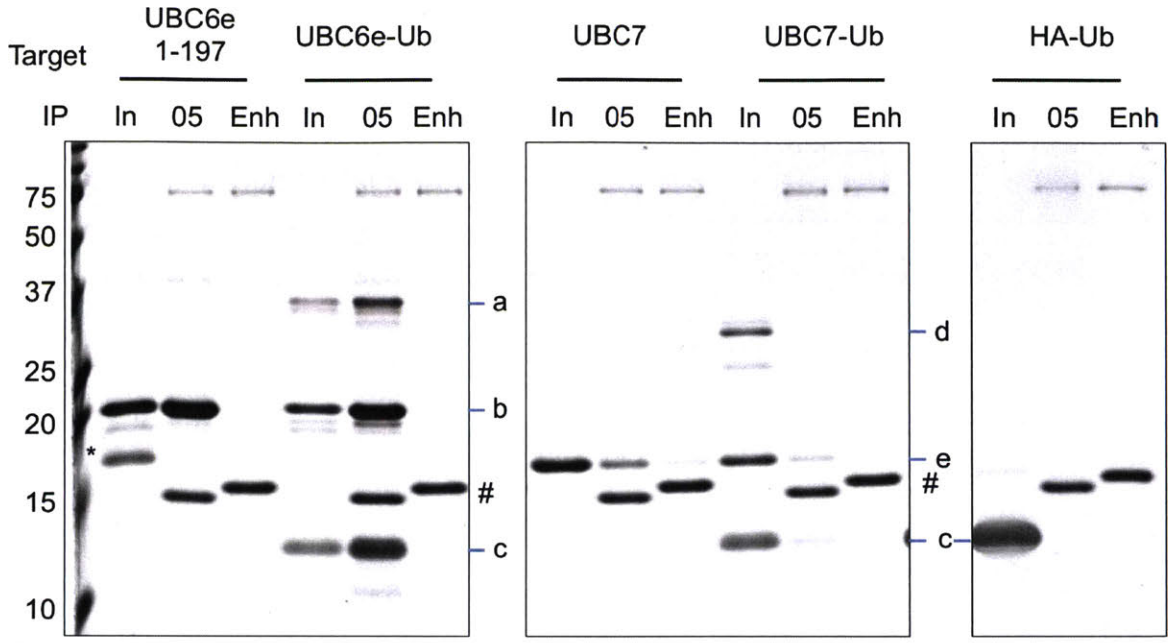


Figure 3.4a. VHH05 binds to both UBC6e and UBC6e-Ub. a: UBC6e(1-197)-Ub. b: UBC6e(1-197). c: HA-Ub. d: UBC7-UB. e: UBC7. *: SUMO. #: VHH05 and VHH enhancer. Samples were eluted in 100 mM DTT, therefore the thioester has been partially hydrolyzed.

VHH05 binding induces structural changes in UBC6e

Multiple attempts to crystallize UBC6e (1-197), both in the presence and absence of VHH05, failed to yield crystals. We therefore turned to NMR analysis of ^{15}N-labeled UBC6e. UBC6e (1-197) is not stable in solution for prolonged periods, forming aggregates that lead to signal loss over time. Therefore, we were unable to determine the complete structure of UBC6e. However, when non-labeled VHH05 was added to ^{15}N-labeled UBC6e in solution, a significant change in the distribution of N-H peaks was observed, indicating that there is a structural change upon binding of VHH05 (Figure 3.4b).

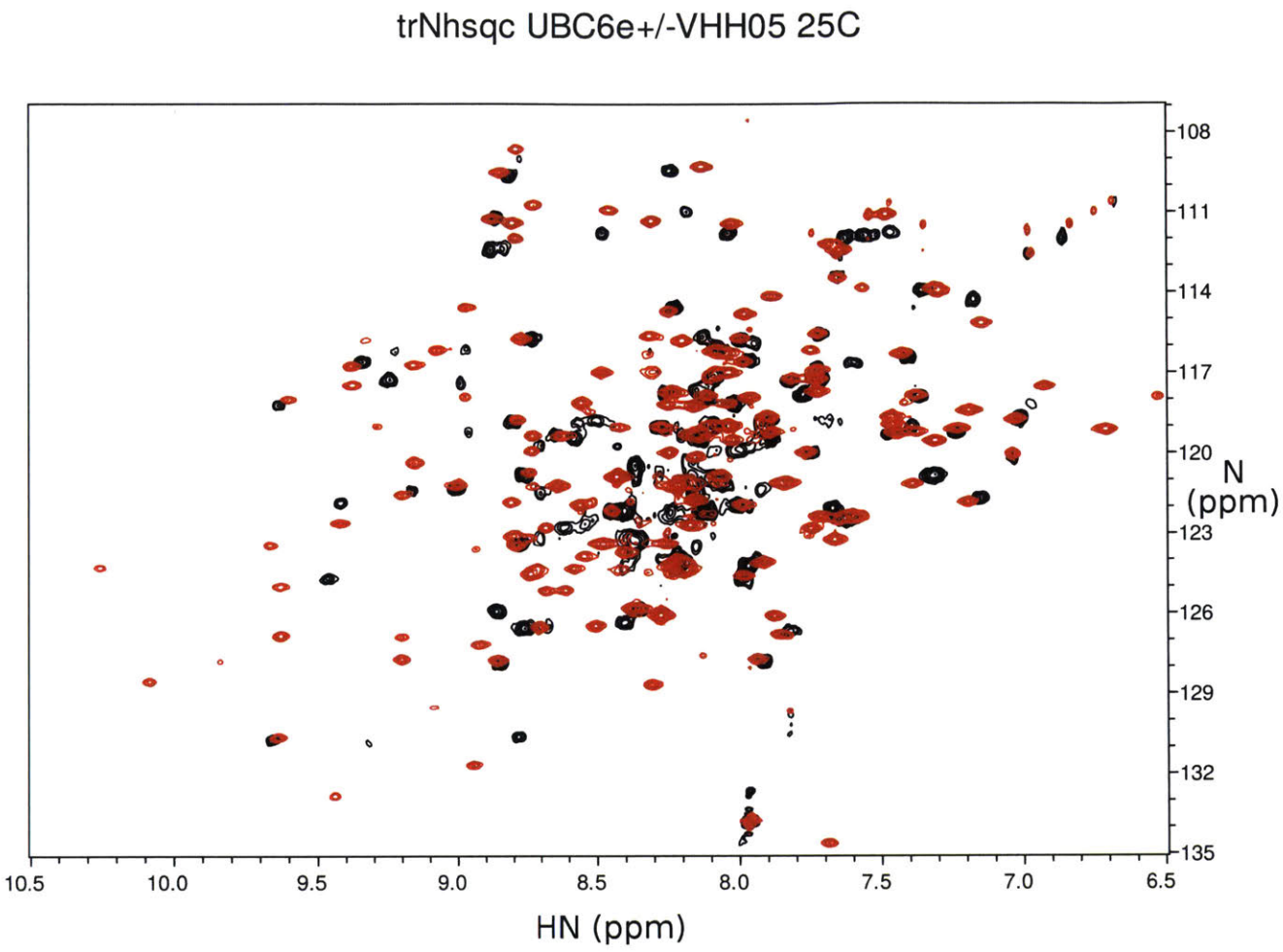


Figure 3.4b. Addition of VHH05 induces structural changes in UBC6e. HN-TROSY-HSQC spectrum of 0.2mM ^{15}N-labeled UBC6e (1-197), 20mM $NaPO_4$, pH7.5, with 50mM NaCl and 10mM DTT, 10% D_2O in the absence (shown in black) or presence of 0.22mM unlabeled VHH05 (shown in red). Data acquired on Bruker Avance III 800Mhz spectrometer at 25ºC. Data were collected with 16 scans per FID, 1024 complex points in direct ^{1}H dimension, and 50 complex points in indirect ^{15}N dimension for UBC6e alone, and 256 scans per FID, 1024 complex points in direct ^{1}H dimension, and 116 complex points in indirect ^{15}N dimension for UBC6e with VHH05.

Identification of the VHH05 epitope on UBC6e

VHH05 can be used to detect SDS-denatured UBC6e by immunoblot, indicating that it might recognize a peptide epitope rather than a 3-dimentional conformational epitope in the target protein (Figure 3.1c). Furthermore, VHH05 can recognize construct UBC6e (1-197) but not UBC6e 1-165 both by immunoprecipitation and immunoblot (Figure 3.5a and 3.5b). The construct UBC6e (1-165) was produced in an attempt to identify a minimal construct

for crystallization by running homology analysis to other crystallized E2s. In addition, VHH05 no longer affects the rate of E2-Ub loading for UBC6e (1-165) (Figure 3.5c). We therefore inferred that the epitope probably resides between residues 166 to 197 in UBC6e. The amino acid sequence for this region is GSGSSQADQEAKELARQISFKAEVNSSGKTIA, where the underlined serine was mutated to glycine to create a GGG sortag motif. This peptide was sortagged to EGFP-LPSTGG and the chimeric product was tested for its binding to VHH05. EGFP-UBC6e (168-197), but not EGFP-LPSTGG-His_6, was recovered on VHH05-conjugated Sepharose beads (Figure 3.5d). Biotinylated EGFP-UBC6e (168-197) exhibited a binding affinity (K_d at 7 ± 2 nM) similar to that of UBC6e (1-197) as measured using Octet (Table 3.2), showing that the epitope is indeed within the region grafted.

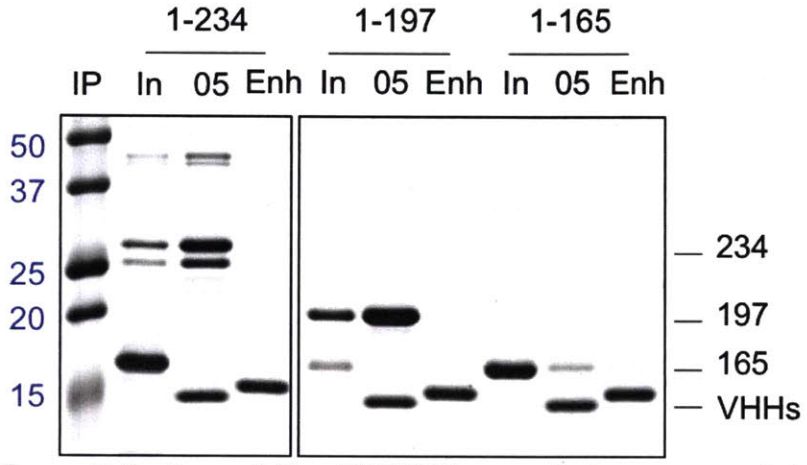


Figure 3.5a. Immobilized VHH05 can only immunoprecipitate UBC6e(1-234) and UBC6e (1-197), but not UBC6e (1-165). Band identities are indicated by the labels on the right.

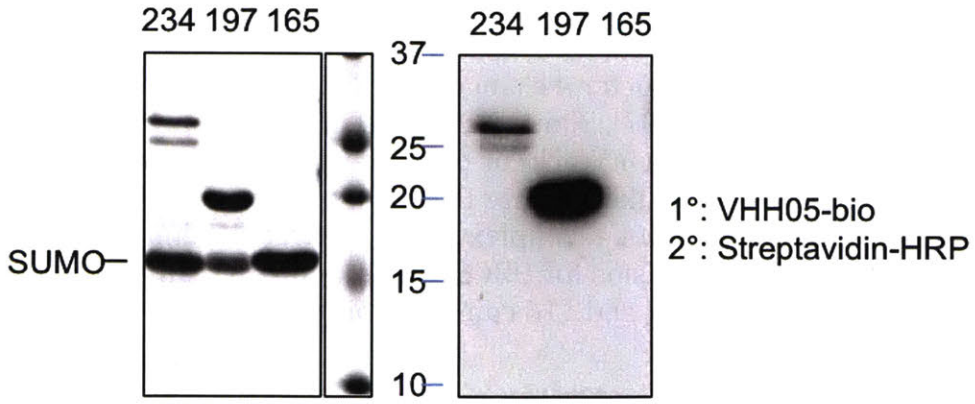


Figure 3.5b. VHH05-biotin recognizes UBC6e (1-234) and UBC6e (1-197), but not UBC6e (1-165). 2 μg of each protein was loaded (Coomassie-stained gel on the left), and 30 ng of each protein was loaded for the blot on the right. UBC6e (1-165) co-migrates with SUMO on 15% SDS-PAGE. The identity of UBC6e (1-165) was confirmed by LC-MS (data not shown).

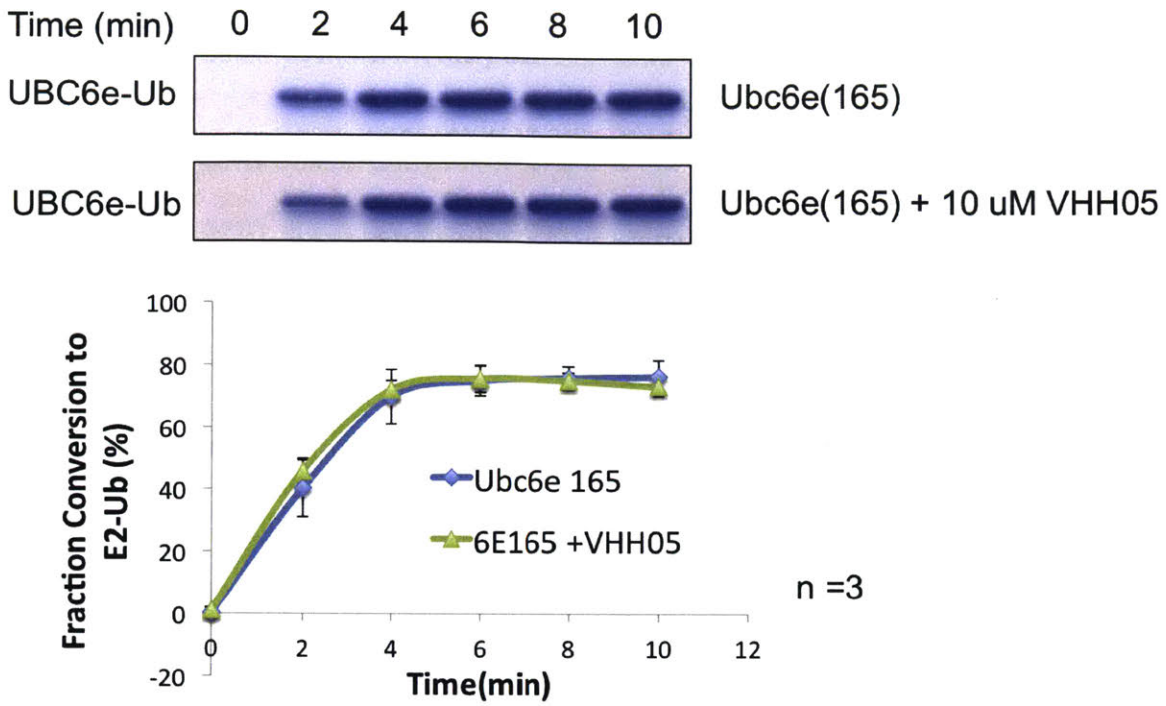


Figure 3.5c. VHH05 does not affect the rate of formation of UBC6e(1-165)-Ub. The experiment was performed as described in Figure 3.2b.

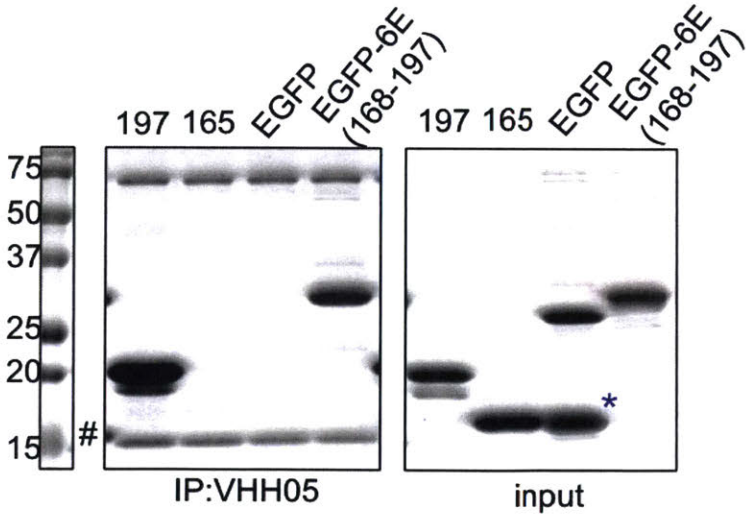


Figure 3.5d. Immobilized VHH05 immunoprecipitates EGFP protein sortagged with UBC6e fragment (168-197). *: cleaved SUMO protein from expression of SUMO-EGFP-LPSTGG-His_6. #: VHH05-LPETGG-His_6.

Table 3.2. Binding constants of constructs carrying various lengths of UBC6e fragments.

Immobilized construct	**Kd (nM)**
UBC6e (1-197)-LPSTGGG-biotin	16 ± 4
EGFP(biotin)-LPSTGGG-(168-197)	7 ± 2
Biotin-LPET-GGGSSQADQEAKELARQISFK (168-186)	9 ± 2

Further analysis of the UBC6e(168-197) sequence, using online secondary structure prediction software Jpred4, suggested that the region between residues 172 and 183 might be helical (Figure 3.5e)(Drozdetskiy et al., 2015). A shorter truncated peptide containing GGG-UBC6e(168-186) was synthesized and sortagged onto EGFP-LPSTGG or Arl3-LPSTGG to see whether it still conferred binding to VHH05. Arl3 (ADP-ribosylation factor-like protein 3) is an irrelevant protein with no homology to E2s, and is used as a second carrier protein. While neither EGFP nor Arl3 bind to VHH05, the chaemeric proteins sortagged with peptide (168-186) are recovered on immobilized VHH05 (Figure 3.5f). GGG-UBC6e(168-186) was biotinylated by sortagging to biotin-LPETGG, and VHH05 exhibited an affinity for the peptide measured at 9 ± 2 nM, comparable to the measurement for UBC6e (1-197) and EGFP(biotin)-LPSTGGG-(168-197)(Table 3.2). We thus concluded that the epitope of VHH 6E05 is within the predicted helical region UBC6e(168-186).

```
                 165       170       175       180       185       190       195
                  |         |         |         |         |         |         |
               :  GSGSSQADQEAKELARQISFKAEVNSSGKTIA  :

Jnet           :  -------HHHHHHHHHHH--------------  :
jhmm           :  --------HHHHHHHHHH--------------  :
jpssm          :  ------HHHHHHHHHHHH---EE---------  :

Jnet Rel       :  99986221888888741121333677788 9  :
```

Figure 3.5e. Secondary structure prediction on UBC6e (165-197). H: helical. E: sheets. Jnet: Jnet overall prediction. Jhmm: Jnet hmm profile prediction. Jpssm: Jnet PSIBLAST pssm profile prediction. Jnet Rel: confidence level range from 0 to 9. Higher number means higher confidence.

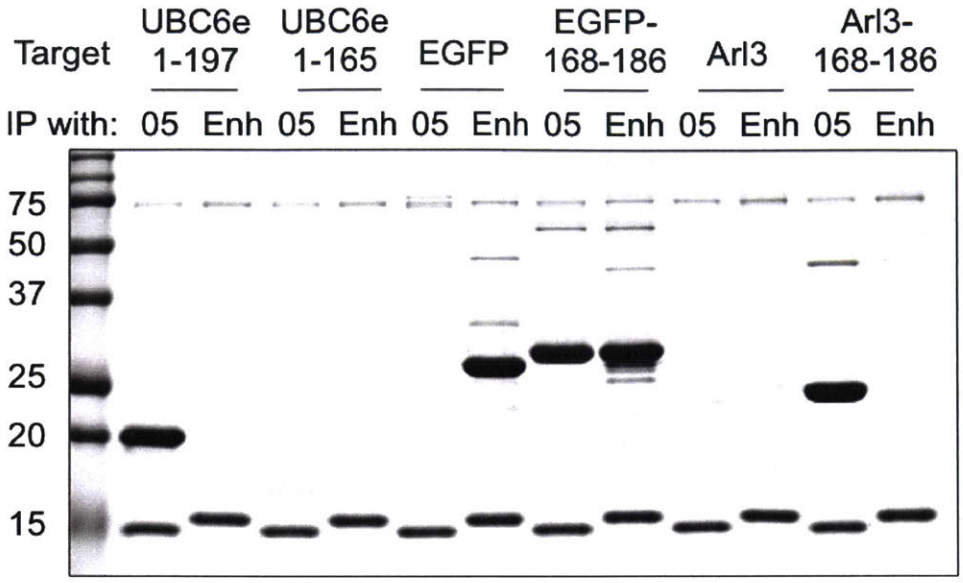


Figure 3.5f. Immobilized VHH05 immunoprecipitates EGFP and Arl3 protein sortagged with UBC6e fragment (168-186) as evaluated by SDS-PAGE.

Recognition by VHH05 is not affected by phosphorylation state.
UBC6e can be phosphorylated on Ser184, which occurs both during ER stress and protein synthesis inhibition. PERK has been reported to be responsible for ER stress-induced UBC6e phosphorylation, and MK2 has been reported to phosphorylate UBC6e during inhibition of protein synthesis.(Menon et al., 2013; Oh et al., 2006) In addition, we observed high levels of UBC6e phosphorylation within 40 minutes of B cell activation (Müller, 2009). Based on the epitope identified for VHH05, the site of phosphorylation (Serine 184) lies very close to, if not within the binding site of VHH05.

We first tested whether VHH05 distinguished between phosphorylated and non-phosphorylated UBC6e. We recombinantly expressed a phosphomimetic mutant, S184E, of UBC6e and tested whether VHH05 still binds to the phosphomimetic mutant. Immobilized VHH05 recovered UBC6e (S184E) to a similar extent as observed for UBC6e(1-197). We then non-specifically biotinylated UBC6e(S184E) using biotin-NHS and measured its affinity to VHH05 using Octet. The K_d for the interaction between VHH05 and UBC6e (S184E) was 15 ± 10 nM, similar to that measured for UBC6e(1-197)-biotin produced with biotin-NHS (14 ± 8 nM) and UBC6e(1-197)-LPETGGG-biotin (16 ± 4 nM) produced by sortagging. In addition, the loading of UBC6e(S184E) with Ub is also similar to that of UBC6e(1-197), both in the presence or absence of VHH05.

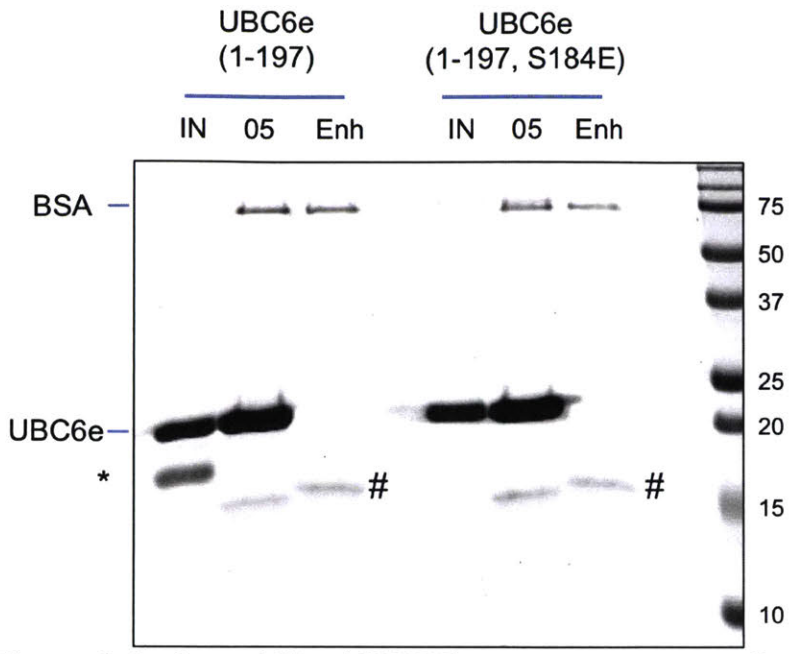


Figure 3.6a. Immobilized VHH05 immunoprecipitates both UBC6e (1-197) and UBC6e (1-197, S184E). *: cleaved SUMO protein from expression as a SUMO fusion. #: VHH05-LPETGG-His_6 and VHH enhancer-LPETGG-His_6. BSA: Bovine Serum Albumin used for blocking.

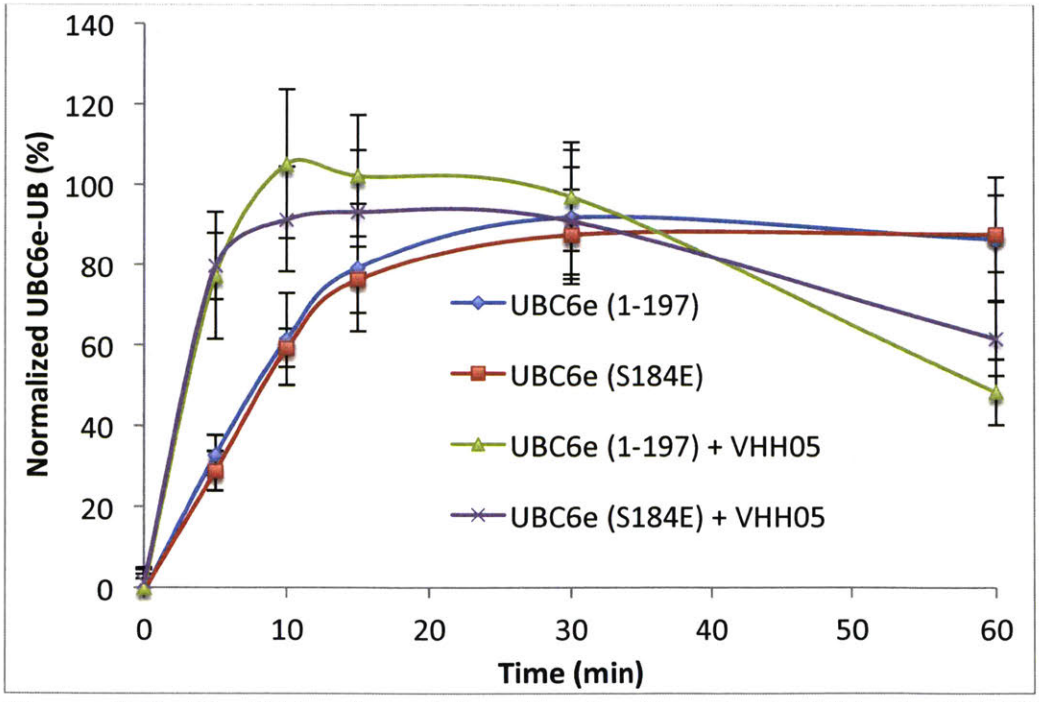


Figure 3.6b. The E2 loading of phosphomimetic mutant UBC6e (1-197, S184E) is similar to that of wildtype UBC6e(1-197). The experiment and quantification was performed as described in Figure 3.2b. Error Bars represent standard deviation (n=3).

We further established that the binding of VHH05 does not depend on the phosphorylation state of the epitope by synthesizing the phosphorylated peptide (GGGQADQEAKELARQIS*FKAEVNS, where S* is phosphoserine). We site-specifically modified this peptide with biotin at the N-terminus via sortagging using biotin-LPETGG and measured the affinity of VHH05 for immobilized, biotinylated peptide using Octet to be 2 ± 0.2 nM, representing slightly tighter binding than that measured for Biotin-LPET-GGGSSQADQEAKELARQISFK (168-186) (9 ± 2 nM).

Inhibition of protein synthesis strongly induces UBC6e phosphorylation (Menon et al., 2013). We tested whether VHH05 could immunoprecipitate phosphorylated UBC6e induced by cycloheximide (CHX) treatment. VHH05 indeed immunoprecipitates both non-phosphorylated UBC6e and phosphorylated UBC6e, supporting the hypothesis that binding of VHH05is independent of the phosphorylation state of its epitope.

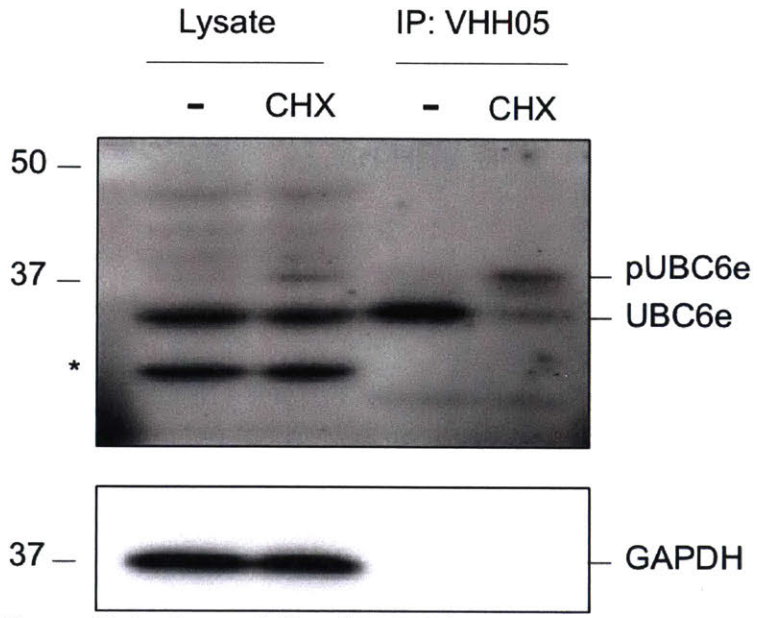


Figure 3.6c. Immobilized VHH05 immunoprecipitates both UBC6e and phosphorylated UBC6e. HeLa cells were treated with 50 μM cycloheximide for 30 minutes before lysates were collected in 1% NP-40 in HBS. The lysates were incubated with sepharose beads functionalized with VHH05 overnight at 4 °C, and the beads were washed 3 times in 1% NP-40 in TBS before elution in 2x SDS sample buffer with 100 mM DTT. Lysate and IP samples were immunoblotted with 1:3000 polyclonal rabbit anti-UBC6e serum. pUBC6e: phosphorylated UBC6e.

VHH05 does not affect intracellular UBC6e or OS-9 levels when expressed in mammalian cells

We showed that VHH05 forms a complex with endogenous UBC6e when expressed in MEF cells (Figure 3.1d). We also wanted to check whether expression of VHH05 leads to changes in endogenous UBC6e levels. MEFs and HeLa cells were transduced with VHH05 or control

VHH anti-Flu NP VHH68 using the pINDUCER20 system (Meerbrey et al., 2011). TheVHH05 epitope UBC6e (165-186) is conserved between mouse and human. We observed no consistent and obvious differences for UBC6e levels in either MEF cells or HeLa cells that express VHH05. There is also no significant change in OS-9 levels, which is up-regulated in UBC6e$^{-/-}$ cells (Hagiwara et al., 2016; Koenig et al., 2014).

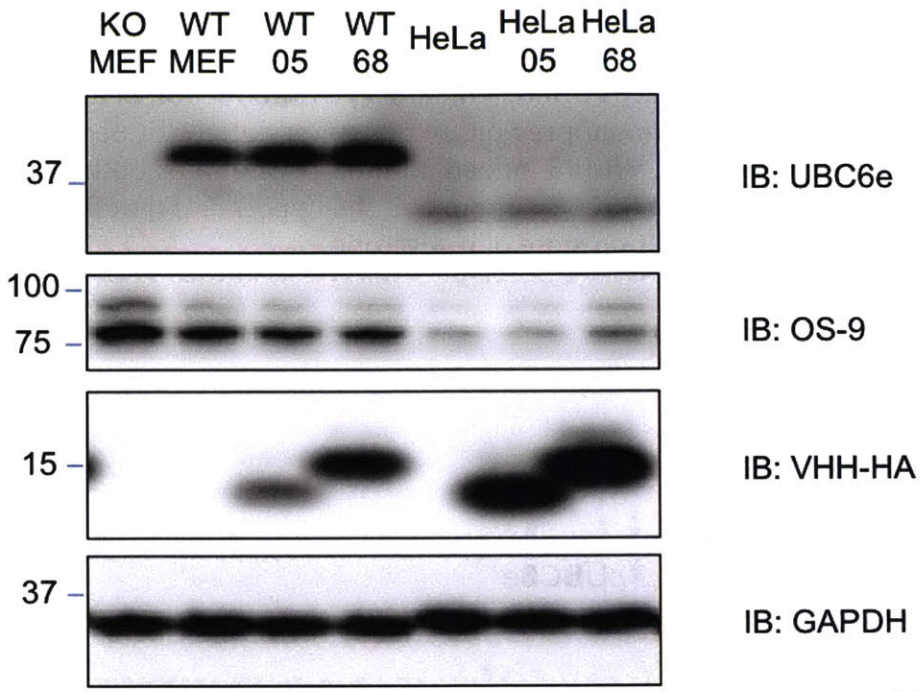


Figure 3.7a. VHH05 expression does not result in changes on UBC6e levels or OS-9 levels. Human UBC6e migrates slightly faster than mouse UBC6e on SDS-PAGE. Experiments were repeated 3 times. Numbers listed correspond to the identity of the VHH expressed intracellularly.

VHH05 inhibits UBC6e phosphorylation in cells

We then tested whether intracellularly expressed VHH05 affects phosphorylation of UBC6e in response to stress. Cytosolic expression of VHH05 inhibits phosphorylation of UBC6e induced by cycloheximide (CHX, protein synthesis inhibitor), dithiothreitol (DTT, reducing agent and ER stressor), thapsigargin (Tg, Ca^{2+} channel inhibitor and ER stressor) or MN7243 (E1 inhibitor), while the control VHH68 that targets the hemagglutinin of influenza virus has no effect (Figure 3.7b). Consistent with previous observations, protein synthesis inhibitor cycloheximide induced the strongest UBC6e phosphorylation (Menon et al., 2013).

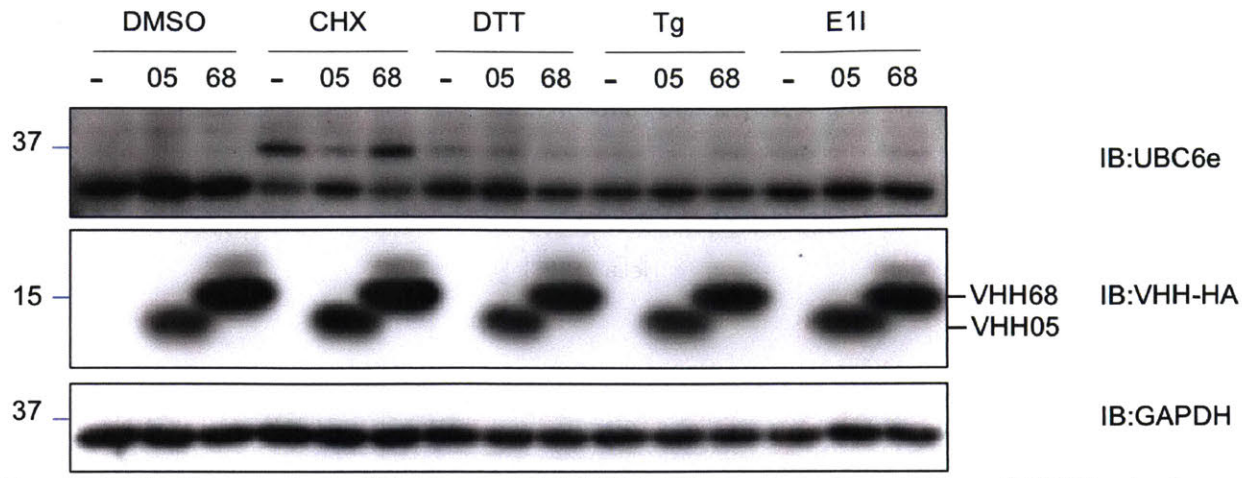
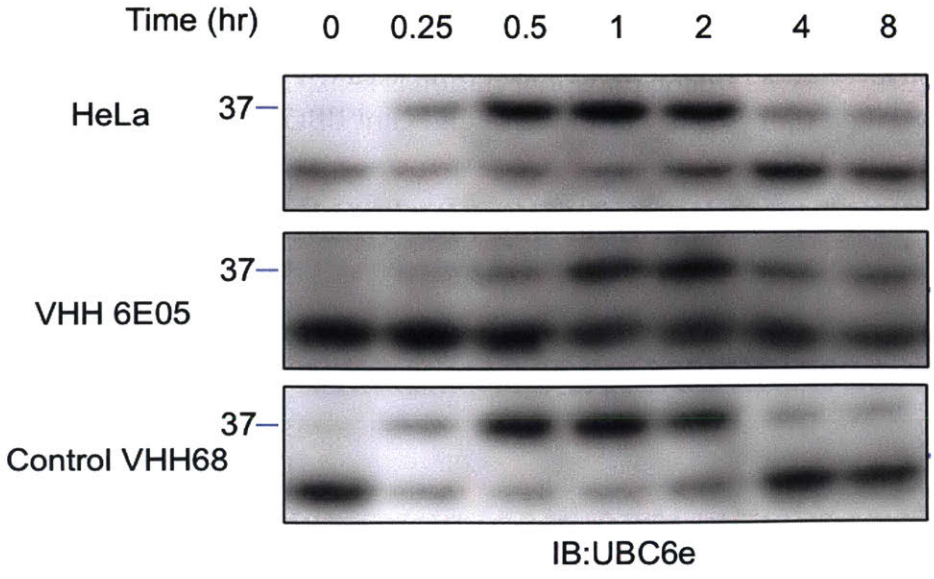


Figure 3.7b. Expression of HA-tagged VHH05 inhibits phosphorylation of UBC6e during protein synthesis inhibition and induction of ER stress. Treatment conditions: DMSO (1%, 30 minutes), CHX (50 μM, 30 minutes), DTT (10 mM, 1 hour), Tg (100 μM, 2 hours) and E1 inhibitor (25 μM, 1 hour). 05: HeLa cells expressing VHH05. 68: HeLa cells expressing control VHH68.

We then analyzed the changes in the fraction of UBC6e that is phosphorylated at different time points following cycloheximide treatment. Cycloheximide treatment leads of phosphorylation of ~80% UBC6e within an hour, and this modification is then slowly reversed. Cytosolic expression of VHH05 delays phosphorylation of UBC6e to peak at 2 hours, and decreases maximum phosphorylation to ~60% (Figure 3.7c). These results indicate that VHH05 not only was able to bind to its target UBC6e in the reducing environment of cytosol, but also that it interferes with normal protein-protein interactions involving UBC6e and the kinases responsible for stress-induced phosphorylation (in the case of protein synthesis inhibition: MK2).



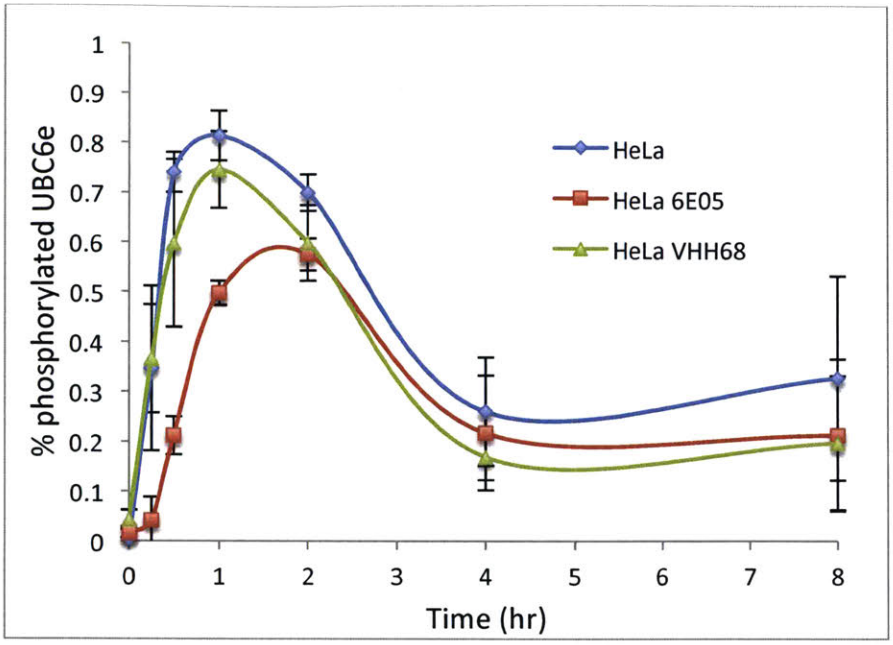


Figure 3.7c. Expression of VHH05 both delays and dampens UBC6e phosphorylation induced by inhibition of protein synthesis. Quantification was calculated for the fraction of phosphorylated UBC6e at each time point. Error bars represent standard deviation (n=3).

Expression of VHH05 accelerates degradation of ERAD substrate NHK

We have shown that binding by VHH05 allosterically enhances UBC6e activity (Figure 3.2b and 3.3b). To investigate whether this enhanced UBC6e activity accelerates ERAD, we expressed VHH05 and VHH68 in WT MEF cells stably expressing an ERAD substrate: the misfolded null Hong Kong mutant of α1-antitrypsin (NHK)(Sifers et al., 1988). Cells were starved for 10 minutes, pulsed with ^{35}S-labeled Met/Cys for 30 minutes and chased with complete media for the indicated times. In MEF cells transduced with VHH05, we observed a faster clearance of NHK compared to untransduced cells, or cells transduced with the control VHH, VHH68 (Figure 3.8a). As Ubc6e phosphorylation is sensitive to protein synthesis inhibition, we verified that we did not induce significant UBC6e phosphorylation during our pulse chase (Figure 3.8b).

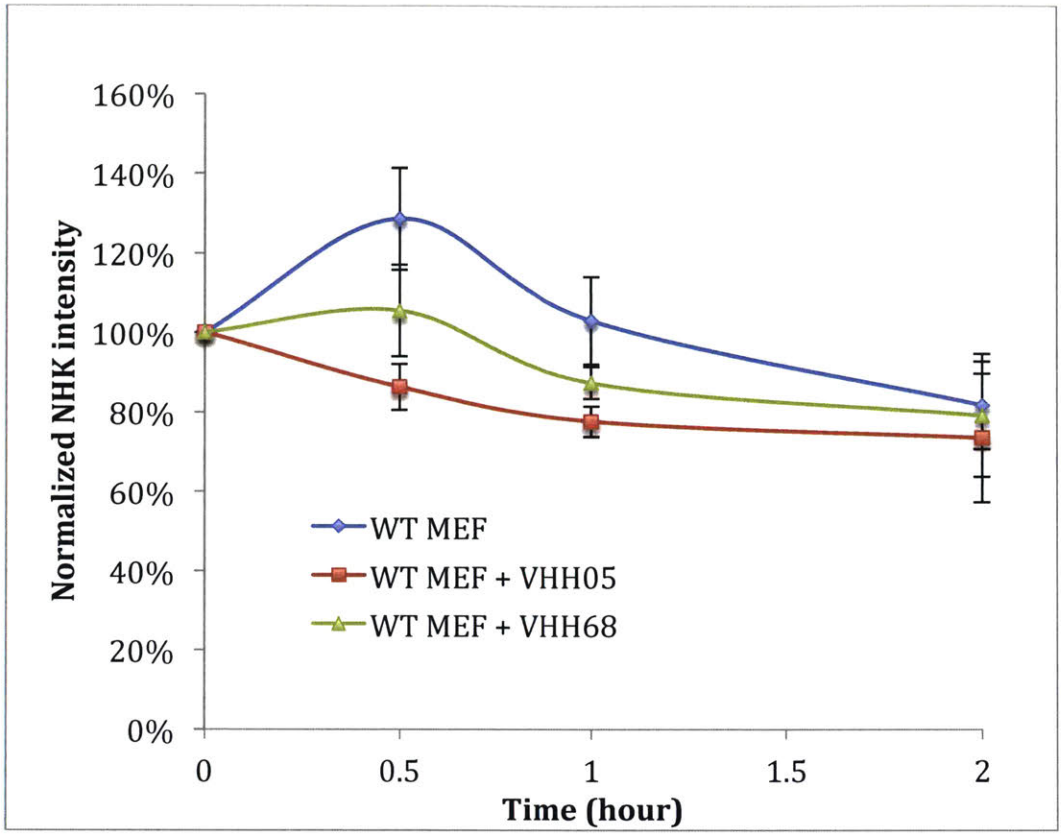


Figure 3.8a. Cytosolic expression of VHH05 accelerates degradation of NHK-HA Signals were normalized to t = 0. WT MEF cells untransduced or transduced with VHH05 or VHH68 were starved for 10 minutes, labeled with ^{35}S-labeled Met/Cys for 30 minutes and chased for the indicated times. Detergent lysates were immunoprecipitated with sepharose beads functionalized with an antibody against the HA tag and analyzed by SDS-PAGE. Error bars represent standard deviation (n=3).

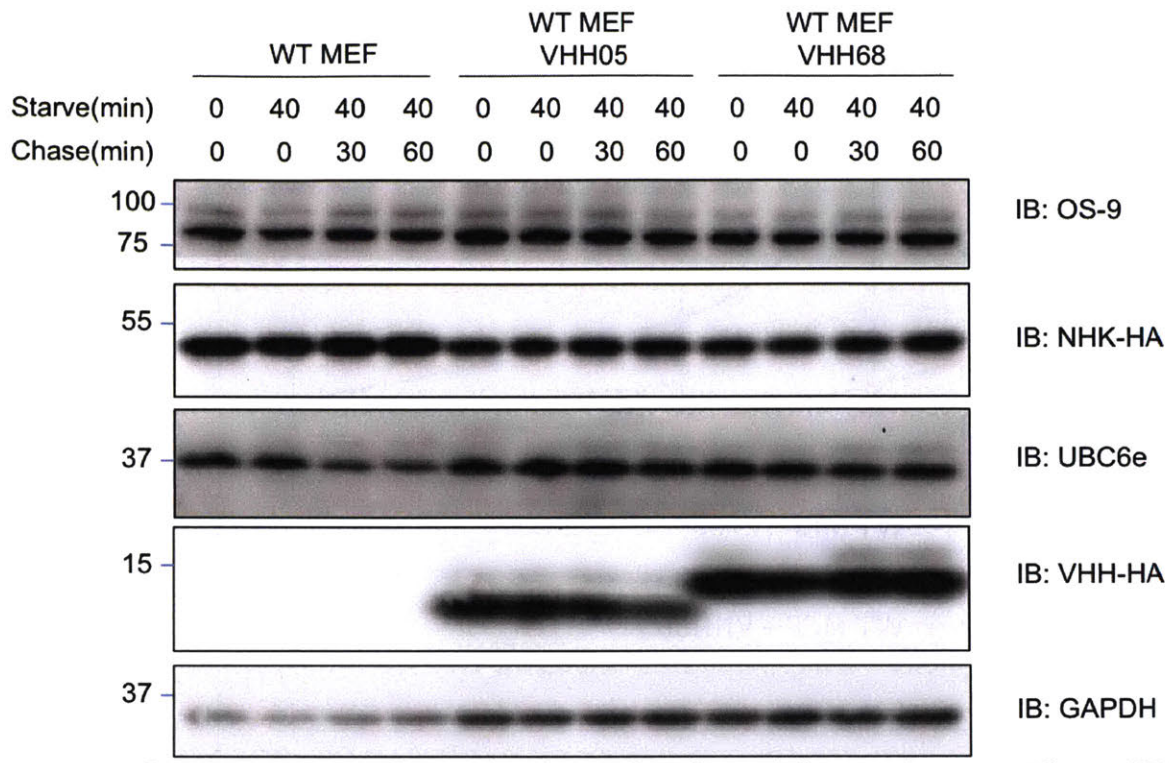


Figure 3.8b. Starvation conditions used in pulse-chase did not induce significant UBC6e phosphorylation. Cells were starved in DMEM without Met/Cys for 40 minutes and chased in complete MEF media for the indicated times to mimic the conditions during ^{35}S-Met/Cys pulse-chase.

3.3 Discussion

ERAD machinery is important for keeping the ER functional by removing misfolded proteins before they form aggregates that may be toxic. During ER stress, in which an excess of unfolded protein accumulates in the ER, an extensive unfolded protein response (UPR) is activated through three different branches: PKR-like ER kinase (PERK), Inositol Requiring Enzyme 1 (IRE1), and Activating Transcription Factor 6 (ATF6) (Gong et al., 2017; Oslowski and Urano, 2011; Walter and Ron, 2011). Activation of PERK results in phosphorylation of eIF2α and subsequent shut down of protein synthesis (Shi et al., 1998). Stress-activated IRE1 splices X-box Binding Protein 1 (XBP1) mRNA, and XBP1 protein transcribed from the spliced mRNA, XBP1s stimulates expression of UPR related genes (Lee et al., 2003; Walter and Ron, 2011). During ER stress, ATF6 is released into the Golgi compartment, cleaved from the membrane and enters the nucleus to activate transcription of XBP1, ER resident chaperones, as well as ERAD components (Bravo et al., 2013). HRD1 and SEL1L are overexpressed during ER stress, which is mediated by the IRE1 and ATF6 pathways respectively (Kaneko et al., 2002; Kaneko et al., 2007; Yamamoto et al., 2008). Each of the ER-associated E2s, UBC6, UBC6e and UBC7, is upregulated during the UPR (Hampton, 2000; Oh et al., 2006).

Such transcription-mediated responses take hours, yet there is substantial evidence that cells mount responses to misfolded proteins more rapidly. It is only logical that the cell would activate ERAD as quickly as possible to meet an increased demand for protein disposal. Our study confirms that UBC6e has the capacity to be activated during stress, and we propose that this activation is likely mediated by phosphorylation on Ser184 during stress.

We have used a newly discovered VHH as a tool to identify a region that allosterically regulates the enzymatic activity of UBC6e. VHH05, which binds to a region of UBC6e predicted to be helical (168-186), with a dissociation constant in the low nanomolar range, enhances the rate of Ub loading on UBC6e by approximately 3-fold (Figure 3.2b). Inclusion of VHH05 also increases the rate of UBC6e-Ub hydrolysis in the presence of HRD1 RING domain (Figure 3.3b). While we were unable to obtain high-resolution structural information for UBC6e with or without VHH05, NMR analysis of ^{15}N-labeled UBC6e shows a significant change in the chemical environments of N and H atoms upon VHH05 binding (Figure 3.4b). This data supports our hypothesis that binding to UBC6e (168-186) by VHH05 induces a significant change in UBC6e structure, likely promoting a more accessible active site. In addition, cytosolic expression of VHH05 leads to faster clearance of ERAD substrate NHK (Figure 3.8a). Thus, the observed increase in UBC6e enzymatic activity *in vitro* applies to intact mammalian cells, and enhances ERAD.

Allosteric regulation of enzymatic activity has been reported for many E2s. Binding of E3 RING domains allosterically enhances E2 reactivity for transferring their Ubs to substrates (Pruneda et al., 2012). For example, UBCH5b activity is enhanced upon binding to multiple E3s (Das et al., 2009; Özkan et al., 2005). In line with these precedents, we also observed an enhanced rate of UBC6e-Ub hydrolysis upon inclusion of HRD1 RING domain (Figure 3.3a and 3.3b). Precedent for non-RING domain-mediated allosteric regulation of E2s has also been reported. UBC7 is allosterically activated by a non-RING domain on its cognate E3 ligase gp78 (Das et al., 2013; Das et al., 2009). An E2 (UBCH5B)-Ub-RING E3 complex can be activated by a non-covalently associated ubiquitin (Buetow et al., 2015). Aside from binding on allosteric sites, phosphorylation of "gateway residues" in E2s has also been reported to open the active site and enhance enzymatic activity for CDC34 (Ub-loading), UBC3B (substrate degradation) and UBE2A (substrate ubiquitination) (Papaleo et al., 2011; Semplici et al., 2002; Shchebet et al., 2012; Valimberti et al., 2015). Since phosphomimetic S184E UBC6e is loaded with kinetics similar to wildtype *in vitro*, Ser184 on UBC6e is unlikely to be a gateway residue. However, we propose that UBC6e (168-186) is a gateway helix that -when engaged- leads to opening of the active site. VHH05 binds to this helix and enhances UBC6e activity. Ser184 phosphorylation could induce binding by phosphoserine-binding proteins, thus similarly enabling allosteric enhancement of UBC6e activity.

The phosphorylation site Ser184 falls within the epitope recognized by VHH05. VHH05 binds to both phosphorylated and non-phosphorylated epitopes with similar affinities, showing that Ser184 is not essential for VHH05 binding. However, cytosolic expression of VHH05 inhibited phosphorylation of UBC6e induced by protein synthesis inhibition. This suggests that VHH05 binding competes with intracellular binding of kinases to the target site, implying that the VHH epitope is adjacent to this Ser184 phosphorylation site. We

propose that phosphorylation of Ser184 could facilitate interaction with intracellular phosphoserine binding domains, which could in turn activate UBC6e as observed with VHH05.

Phosphorylation of UBC6e is induced both by ER stress and protein synthesis inhibition(Menon et al., 2013; Oh et al., 2006). Therefore, UBC6e might not only be activated by unfolded proteins in the ER, but could also be activated to enhance ERAD in response to severe cellular stress that accompanies protein synthesis inhibition. Interestingly, UBC6e phosphorylation is induced by TLR agonists such as LPS (TLR4), CpG (TLR9), Imiquimod (TLR7), Malp2 (TLR2/6), and Pam3CSK4 (TLR1/2) within 40 minutes (Müller, 2009). Since the kinase MK2 is also activated in TLR-activated dendritic cells (McGuire et al., 2013; Zaru et al., 2015), and activated MK2 induces strong phosphorylation of UBC6e, it is likely that MK2 is responsible for TLR-induced UBC6e phosphorylation. Since MK2 is activated during many cellular processes, phosphorylation of UBC6e may provide a mechanism by which the condition of cellular stress can be conveyed to the ER, where ERAD can be pre-emptively activated in preparation for the impact of cytosolic stress on the ER environment.

We previously reported that UBC6e is involved in the post-translational control of levels of ERAD enhancers, such as OS-9, EDEM1 and SEL1L (Hagiwara et al., 2016). In the current study, we observed that OS-9 levels, which are up-regulated in UBC6e deficient cells, are not significantly changed when VHH05 was expressed (Figure 3.7b). OS-9 was restored to similar levels by re-introduction of wildtype, phosphomimetic S184E or phosphorylation-deficient S184A UBC6e into UBC6e$^{-/-}$ cells, and is not affected by overexpression of UBC6e (Hagiwara et al., 2016). This finding suggests that the level of OS-9 expression is not controlled by Ser184 phosphorylation or the interaction of binding partners with phosphorylated UBC6e (168-186). As we have previously proposed, the mechanism by which UBC6e down-regulates OS-9, EDEM1 and SEL1L is likely distinct from conventional ERAD mediated by HRD1. This difference must account for the observation that upon cytoplasmic expression of VHH05, NHK degradation is accelerated, whereas OS-9 levels are unchanged. Future experiments will determine how UBC6e is involved in both mechanisms.

3.4 Materials and Methods

Antibodies and reagents

OS-9 (monoclonal rabbit ; Abcam), GAPDH (Horseradish peroxidase (HRP)-conjugated monoclonal rabbit ;Cell Signaling Technology), Rabbit anti-UBC6e(1:3000) (Mueller et al., 2008) was produced in our laboratory).
Secondary antibodies for immunoblotting: Horseradish peroxidase (HRP)-conjugated goat anti-mouse, -rabbit antibody (Southern Biotech), HRP-conjugated rabbit anti-goat antibody (Southern Biotech). High sensitivity streptavidin-HRP (Pierce, MA) were all used at 1:10,000 dilutions.

Cycloheximide and Doxycycline (Clontech) were purchased. Protein G Sepharose-4 Fast Flow, rProtein A Sepharose 4 Fast Flow were purchased from GE Healthcare.

Plasmids, Protein expression and purification

pCMV-CDH-EF1-puro UBC6e constructs were previously described (Hagiwara et al., 2016). VHHs were amplified from cDNA library generated from alpaca lymphocytes and subcloned into pD vector. VHHs enriched after 2 rounds of panning were then subcloned to pHEN6 vector for expression in *E. coli* and to pINDUCER20 vector for expression in mammalian cells(Meerbrey et al., 2011).

UBC6e constructs were expressed using Champion pET SUMO system (Thermo Scientific). Plasmids coding for SUMO-UBC6e(1-234), (1-197) or (1-165) were transformed into LOBSTER BL21(DE3) *E. coli.* (Andersen et al., 2013) An overnight starter culture was prepared in LB media with 70 μg/ml kanamycin and 50 μg/ml chloramphenicol. Bacteria were grown in TB media at 37 °C until OD_{600} reached 0.5, and protein expression was induced with 1mM IPTG at 30 °C overnight. Bacterial pellets were collected by spinning down at 6000 rpm for 15 minutes. The pellet from a 1L culture was resuspended in 50 ml PBS, cooled on ice for 30 minutes, and lysed by sonication for 60 seconds 3 times. Lysed bacteria were spun down at 17000 rpm for 1 hour to remove cell debris. His-tagged SUMO-UBC6e was purified by binding to Ni-NTA beads, washed with 10 mM imidazole in PBS, and eluted with 250 mM imidazole in 50mM Tris, 500 mM NaCl, pH 8.0. 50 μg of SUMO protease was added to the eluent to cleave off SUMO, and the mixture was dialyzed into 50 mM Tris, 150 mM NaCl, pH 7.4 overnight. SUMO and uncleaved SUMO-UBC6e was then removed by passing through fresh Ni-NTA beads, and the beads were washed with 20 mM imidazole in 50 mM Tris, 500 mM NaCl. UBC6e protein in flow through and 20 mM wash was pooled and further purified with an S75 column (GE Healthcare, UK). A typical yield is 5-10 mg per 1L culture.

GST-hmHRD1 RING (272-343) was expressed using pGEX-6P-1 plasmid. (Kikkert et al., 2004) Bacterial lysates were collected as described above and the lysates were incubated with 1 ml Glutathione Sepharose 4B (GE Healthcare, UK) at 4°C for 1 hour. The beads were washed 3 times with 10 ml PBS, and GST-HRD1 RING was eluted with 10 mM reduced glutathione in TBS. The eluted protein was further purified by S75 FPLC column.

VHHs were expressed using pHEN6 vector. Plasmids coding for PelB-VHH-LPETGG-His_6 were transformed into WK6 *E.coli* bacteria. Bacteria were grown in TB media with 100 μg/ml ampicillin at 37 °C until OD_{600} reached 0.5, and protein expression was induced with 1mM IPTG at 30 °C overnight. Bacteria pellets were collected by spinning down at 6000 rpm for 15 minutes, and pellet from a 1L culture was resuspened in 15 ml TES buffer (50 mM Tris, 650 μM EDTA, 2 M sucrose) to prepare for osmotic shock. After incubating for 2 hours at 4 °C, 75 ml distilled H_2O was added and the bacteria was further incubated overnight at 4 °C. His-tagged VHHs were purified by binding to Ni-NTA beads, washed with PBS, and eluted with 250 mM imidazole in 50mM Tris, 500 mM NaCl, pH 8.0. The eluent was further purified with an S75 column. A typical yield is 20-40 mg per 1L culture.

Cell Culture

Wildtype(WT) and UBC6e$^{-/-}$ MEF cells were previously described(Hagiwara et al., 2016). MEF cells were maintained in Dulbecco's modified Eagle's medium (DME) supplemented with 10% IFS and 0.0007% v/v b-mercaptoethanol (MEF medium) in humidified air containing 5% CO_2 at 37°C. HeLa cells were maintained in Dulbecco's modified Eagle's medium (DME) supplemented with 10% IFS.

Immobilization of VHHs to CNBr-activated sepharose beads

200 mg of CNBr-activate sepharose beads (Sigma-aldrich) were washed twice with 20 ml of 1 mM HCl, and neutralized with 100 mM $NaHCO_3$, 500 mM NaCl, pH 8.3. 2 mg of each VHH (VHH6E05, VHH enhancer or VHH MHCII07) in PBS was added to beads in 20 ml of 100 mM $NaHCO_3$, 500 mM NaCl, pH 8.3 and reacted overnight at 4 °C. The beads were blocked with 50 mM Tris, 500 mM NaCl, pH8.0 for 2 hours at room temperature, and then TBS containing 1% BSA overnight before use.

Immunoblotting and immunoprecipitation (purified protein)

Unless otherwise stated, 30 ng (as measured by nanodrop) of each protein was loaded for immunoblot. VHH-biotin (VHH05-biotin and VHH enhancer-biotin) was used at 1 μg/ml in 5% milk in TBST. High sensitivity HRP-Conjugated Streptavidin (Pierce, MA) was used as 1:10,000 dilution.

For immunoprecipitation using immobilized VHH05 and VHH enhancer, 50 μl of beads (approximately 100 μg VHHs) and 50 μg of target protein was used. Immunoprecipitations were conducted in TBS containing 1% BSA. Beads were washed with TBS for 3 times before boiled in 2x sample buffer with 100 mM DTT for elution.

Generation of stable cells using lentiviral transduction

Lentivirus was produced with HEK293 cells. HEK cells were plated in 6-well plate and grown until 70% confluency, and transfected with 1.5 μg of plasmids containing gene of interest (backbones in pINDUCER20 or pCDH-CMV-MCS-EF1-Puro), 0.65 μg psPAX2 and 0.35 μg pMD2.G. Media containing lipofectamine 2000/plasmids complex were removed after 6 hours of tranfection, and HEK cells were grown in DMEM containing 10% heat-inactivated FBS for 2 days. Virus particles were harvested by spinning down the HEK cell pellet, and passing the media through a 0. 45 μm filter. The media containing virus particles were mixed 1:1 with fresh media and used to transduce MEF or HeLa by incubating for 6 hours. The cells were grown in regular media overnight and selected with 500 μg/ml geneticin for pINDUCER20-transduced cells and 2 μg/ml puromycin for pCDH-CMV-MCS-EF1-Puro-transduced cells.

Immunoblotting and immunoprecipitation (lysates)

Cultured cells were lysed in HEPES-buffered saline (HBS) with 1% NP-40 supplemented with protease inhibitor cocktail (complete protease inhibitor cocktail tablets; Roche Applied Science) on ice for 20 minutes. Cell debris were removed by spinning at 13,000 rpm for 20 minutes at 4°C. Protein content was determined using a BCA protein assay (Thermo Fisher Scientific, Inc.). Proteins were denatured in SDS sample buffer containing 0.1 M DTT, resolved by SDS-PAGE, transferred to PVDF membranes and immunoblotted using standard procedures.

All immunoprecipitate procedures were performed at 4°C or on ice. Cell lysates were collected in HBS with 1% NP-40 and incubated with CNBr-activated sepharose beads conjugated with VHH05 (50 μg bead volume, 30 μg VHH-LPETGG-His_6) or anti-HA agrose beads (Thermo Scientific) overnight. The beads were washed for 3 times with TBS with 1% NP-40 and resuspended in SDS-PAGE sample buffer with 0.1M DTT.

Pulse chase

Cells transduced with pCDH-EF1-CMV-NHK-HA and pINDUCER-VHH05-HA or pINDUCER VHH68-HA were starved in DMEM lacking methionine and cysteine for 10 minutes and pulse labeled with Expre35S35S protein labeling mixture (PerkinElmer Life Sciences) for 30 minutes at a concentration of 0.18 mCi/ml. After pulse labeling, pulse medium was diluted with chase medium containing 5 mM met/cys was added and aspirated immediately to prevent further incorporation of label. The cells were then incubated in complete MEF medium supplemented with 5 mM Met/Cys and 10 mM HEPES for the indicated chase periods. Pulse chase was stopped by transferring dishes to ice and replacing chase media with stop media (ice cold PBS + 20 mM NEM). Cells were lysed with 1% Triton X-100 in MNT buffer (20 mM Mes, 30 mM Tris, 100 mM NaCl pH 7.8). Cell lysates were centrifuged at 18,000 g for 10 minutes at 4°C to remove nuclei. NHK-HA was immunoprecipitated with 50 μl of anti-HA sepharose beads, ~0.4mg/ml loading, overnight at 4 °C. Beads were washed twice in a buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 0.1% SDS and 0.05% Triton X-100. Samples were analyzed by reducing 10% SDS-PAGE. Dried gels were then exposed either to Kodak BioMax MR films or phosphor screens (GE Life Sciences) for visualization. Phosphor screens were scanned using a typhoon FLA-9000 imager and quantifications performed using ImageQuant TL 8.1 (GE Life Sciences).

Peptide Synthesis and purification

Peptides were synthesized using flow-based solid-phase peptide synthesizer using Fmoc chemistry as previously described(Simon et al., 2014). Peptides were conjugated to Rink-amide linker resin to produce C-terminal amides. Phospho-serine was incorporated on solid phase using Fmoc-L-Ser(HPO3Bzl)-OH (Chem Impex). Following synthesis beads were air-dried and peptides were cleaved from resin and deprotected in 92.5% trifluoroacetic acid, 5% H_2O, and 2.5% TIPS. Peptides were precipitated into cold diethyl ether, air-dried, and purified using reverse phase C18 HPLC using 10-50% Acetonitrile/H_2O gradient.

Sortase reactions

VHH-LPETGG-His_6, eGFP-LPSTGG-His_6 or biotin-LPETGG were incubated with 500 μM of GGG-peptides in the presence of 5 μM HeptMutSrt-His_6 at room temperature for 1 hour or 4°C overnight. Unreacted substrates and sortase were removed by incubating with Ni-NTA beads, and excess peptides were removed using P10 column for VHHs and eGFP, and HPLC for biotin. A typical yield is 50% after purification.

E2 loading reactions and UBC6e-Ub hydrolysis reactions

For analytical E2-loading reactions, 3 μM of purified E2 (UBC6e, UBC6 or UBC7) was incubated with 10 μM HA-UB and 100 nM purified His_6-mouse E1 in 100 mM Tris with 2

mM ATP, 5 mM $MgCl_2$ and 200 μM DTT at 37°C, in the presence or absence of 10 μM VHHs. Reactions were quenched at different time points with equal volume of 2x SDS sample buffer with no reducing agents. Samples were analyzed with non-reducing 15% SDS-PAGE gels.

For the preparation of UBC6e(TAMRA)-Ub, 4 mg of SUMO-UBC6e-LPSTGG in 2 ml TBS was SUMO cleaved with 100 μg of SUMO protease, followed by exposure to 5 μM HeptMutSrt-His_6 and 500 μM of G_3-TAMRA. The sortase reaction mixture was directly used for E2-loading reaction without purification. To the reaction mixture, 10 μM HA-UB and 100 nM purified His_6-mouse E1 in 100 mM Tris with 2 mM ATP, 5 mM $MgCl_2$ and 200 μM DTT was added, and the E2 loading reaction was allowed to proceed at 37°C for 45 minutes. The reaction mixture was spun twice at 13000 rpm for 3 minutes at 4°C to remove insoluble debris, and directly loaded to S75 FPLC column (GE healthcare, UK) for purification. The peak corresponding to UBC6e(TMR)-Ub was pooled and concentrated. The product contains about 50% of unloaded UBC6e (TAMRA), both from incomplete purification and subsequent hydrolysis during purification. The yield was about 20%.

To analyze the hydrolysis of UBC6e(TAMRA)-Ub, 2 μM UBC6e(TAMRA)-Ub was incubated with 10 μM GST-HRD1 RING(272-343) and 10 μM various VHHs in TBS. Reactions were quenched with equal volume of 2x SDS sample buffer with no reducing agents at indicated time points, and analyzed using 15% non-reducing SDS-PAGE. The gel was imaged for absorbance at 546 nm for visualization of TAMRA-labeled UBC6e-Ub and UBC6e.

Octet measurements

All Octet experiments were conducted at Biophysical Instrumentation Facility at MIT using the Octet RED96 Bio-Layer Interferometry. The target of interest is biotinylated either through sortagging or labeled with (+)-Biotin N-hydroxysuccinimide ester (Sigma-aldrich). For NHS labeling, 2-fold molar excess of biotin-NHS was used for labeling at room temperature for 1 hour in PBS. Excess biotin-NHS was removed using ZeBa columns (7K MWCO, Thermo Fisher). All octet measurements were taken in 1% BSA in PBS with 1% Tween. 10 ng/ml of biotin-labeled target were loaded on Dip and Read™ Streptavidin (SA) Biosensors (Pall Fortebio) until the reading excess 1 nm. These sensors were then dipped into wells containing different concentrations of VHH05, ranging from 50 nM to 8 μM. Average and standard deviations were taken for at least 3 measurements.

3.5 References

Andersen, K.R., Leksa, N.C., and Schwartz, T.U. (2013). Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins *81*, 1857-1861.

Baldridge, R.D., and Rapoport, T.A. (2016). Autoubiquitination of the Hrd1 Ligase Triggers Protein Retrotranslocation in ERAD. Cell *166*, 394-407.

Bordallo, J., Plemper, R.K., Finger, A., and Wolf, D.H. (1998). Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins. Molecular biology of the cell *9*, 209-222.

Bordallo, J., and Wolf, D.H. (1999). A RING-H2 finger motif is essential for the function of Der3/Hrd1 in endoplasmic reticulum associated protein degradation in the yeast Saccharomyces cerevisiae. FEBS Letters *448*, 244-248.

Bravo, R., Parra, V., Gatica, D., Rodriguez, A.E., Torrealba, N., Paredes, F., Wang, Z.V., Zorzano, A., Hill, J.A., Jaimovich, E., *et al.* (2013). Endoplasmic Reticulum and the Unfolded Protein Response: Dynamics and Metabolic Integration. International review of cell and molecular biology *301*, 215-290.

Buetow, L., Gabrielsen, M., Anthony, N.G., Dou, H., Patel, A., Aitkenhead, H., Sibbet, G.J., Smith, B.O., and Huang, D.T. (2015). Activation of a primed RING E3-E2-ubiquitin complex by non-covalent ubiquitin. Molecular cell *58*, 297-310.

Burr, M.L., Cano, F., Svobodova, S., Boyle, L.H., Boname, J.M., and Lehner, P.J. (2011a). HRD1 and UBE2J1 target misfolded MHC class I heavy chains for endoplasmic reticulum-associated degradation. Proceedings of the National Academy of Sciences of the United States of America *108*, 2034-2039.

Burr, M.L., Cano, F., Svobodova, S., Boyle, L.H., Boname, J.M., and Lehner, P.J. (2011b). HRD1 and UBE2J1 target misfolded MHC class I heavy chains for endoplasmic reticulum-associated degradation. Proceedings of the National Academy of Sciences of the United States of America *108*, 2034-2039.

Das, R., Liang, Y.-H., Mariano, J., Li, J., Huang, T., King, A., Tarasov, S.G., Weissman, A.M., Ji, X., and Byrd, R.A. (2013). Allosteric regulation of E2:E3 interactions promote a processive ubiquitination machine. The EMBO Journal *32*, 2504-2516.

Das, R., Mariano, J., Tsai, Y.C., Kalathur, R.C., Kostova, Z., Li, J., Tarasov, S.G., McFeeters, R.L., Altieri, A.S., Ji, X., *et al.* (2009). Allosteric Activation of E2-RING Finger-Mediated Ubiquitylation by a Structurally Defined Specific E2-Binding Region of gp78. Molecular cell *34*, 674-685.

Drozdetskiy, A., Cole, C., Procter, J., and Barton, G.J. (2015). JPred4: a protein secondary structure prediction server. Nucleic Acids Research *43*, W389-W394.

Elangovan, M., Chong, H.K., Park, J.H., Yeo, E.J., and Yoo, Y.J. (2017). The role of ubiquitin-conjugating enzyme Ube2j1 phosphorylation and its degradation by proteasome during endoplasmic stress recovery. Journal of cell communication and signaling *11*, 265-273.

Gong, J., Wang, X.-z., Wang, T., Chen, J.-j., Xie, X.-y., Hu, H., Yu, F., Liu, H.-l., Jiang, X.-y., and Fan, H.-d. (2017). Molecular signal networks and regulating mechanisms of the unfolded protein response. Journal of Zhejiang University Science B *18*, 1-14.

Hagiwara, M., Ling, J., Koenig, P.A., and Ploegh, H.L. (2016). Posttranscriptional Regulation of Glycoprotein Quality Control in the Endoplasmic Reticulum Is Controlled by the E2 Ub-Conjugating Enzyme UBC6e. Molecular cell *63*, 753-767.

Hampton, R.Y. (2000). ER stress response: Getting the UPR hand on misfolded proteins. Current Biology *10*, R518-R521.

Hampton, R.Y., Gardner, R.G., and Rine, J. (1996). Role of 26S proteasome and HRD genes in the degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum membrane protein. Molecular biology of the cell *7*, 2029-2044.

Hanke, L., Knockenhauer, K.E., Brewer, R.C., van Diest, E., Schmidt, F.I., Schwartz, T.U., and Ploegh, H.L. (2016). The Antiviral Mechanism of an Influenza A Virus Nucleoprotein-Specific Single-Domain Antibody Fragment. mBio *7*.

Hanke, L., and Schmidt, F.I. (2017). Vesicular stomatitis virus N protein-specific single-domain antibody fragments inhibit replication. EMBO reports *18*, 1027-1037.

Iida, Y., Fujimori, T., Okawa, K., Nagata, K., Wada, I., and Hosokawa, N. (2011). SEL1L protein critically determines the stability of the HRD1-SEL1L endoplasmic reticulum-associated degradation (ERAD) complex to optimize the degradation kinetics of ERAD substrates. The Journal of biological chemistry *286*, 16929-16939.

Ji, Y., Kim, H., Yang, L., Sha, H., Roman, C.A., Long, Q., and Qi, L. (2016). The Sel1L-Hrd1 Endoplasmic Reticulum-Associated Degradation Complex Manages a Key Checkpoint in B Cell Development. Cell reports *16*, 2630-2640.

Kaneko, M., Ishiguro, M., Niinuma, Y., Uesugi, M., and Nomura, Y. (2002). Human HRD1 protects against ER stress-induced apoptosis through ER-associated degradation11The nucleotide sequence data reported are available in the DDBJ/EMBL/GenBank databases under the accession number AB085847. FEBS Letters *532*, 147-152.

Kaneko, M., Yasui, S., Niinuma, Y., Arai, K., Omura, T., Okuma, Y., and Nomura, Y. (2007). A different pathway in the endoplasmic reticulum stress-induced expression of human HRD1 and SEL1 genes. FEBS Letters *581*, 5355-5360.

Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thanedar, S., Roitelman, J., Chau, V., and Wiertz, E. (2004). Human HRD1 Is an E3 Ubiquitin Ligase

Involved in Degradation of Proteins from the Endoplasmic Reticulum. The Journal of biological chemistry *279*, 3525-3534.

Kirchhofer, A., Helma, J., Schmidthals, K., Frauer, C., Cui, S., Karcher, A., Pellis, M., Muyldermans, S., Casas-Delucchi, C.S., Cardoso, M.C., *et al.* (2010). Modulation of protein properties in living cells using nanobodies. Nature structural & molecular biology *17*, 133-138.

Koenig, P.A., Nicholls, P.K., Schmidt, F.I., Hagiwara, M., Maruyama, T., Frydman, G.H., Watson, N., Page, D.C., and Ploegh, H.L. (2014). The E2 ubiquitin-conjugating enzyme UBE2J1 is required for spermiogenesis in mice. The Journal of biological chemistry *289*, 34490-34502.

Lee, A.H., Iwakoshi, N.N., and Glimcher, L.H. (2003). XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. Molecular Cell Biology *23*, 7448-7459.

Lenk, U., Yu, H., Walter, J., Gelman, M.S., Hartmann, E., Kopito, R.R., and Sommer, T. (2002). A role for mammalian Ubc6 homologues in ER-associated protein degradation. Journal of Cell Science *115*, 3007-3014.

Lester, D., Farquharson, C., Russell, G., and Houston, B. (2000). Identification of a family of noncanonical ubiquitin-conjugating enzymes structurally related to yeast UBC6. Biochemical and biophysical research communications *269*, 474-480.

McGuire, V.A., Gray, A., Monk, C.E., Santos, S.G., Lee, K., Aubareda, A., Crowe, J., Ronkina, N., Schwermann, J., Batty, I.H., *et al.* (2013). Cross talk between the Akt and p38alpha pathways in macrophages downstream of Toll-like receptor signaling. Molecular and cellular biology *33*, 4152-4165.

Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T., *et al.* (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America *108*, 3665-3670.

Menon, M.B., Tiedje, C., Lafera, J., Ronkina, N., Konen, T., Kotlyarov, A., and Gaestel, M. (2013). Endoplasmic reticulum-associated ubiquitin-conjugating enzyme Ube2j1 is a novel substrate of MK2 (MAPKAP kinase-2) involved in MK2-mediated TNFalpha production. Biochemical Journal *456*, 163-172.

Mueller, B., Klemm, E.J., Spooner, E., Claessen, J.H., and Ploegh, H.L. (2008). SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. Proceedings of the National Academy of Sciences of the United States of America *105*, 12325-12330.

Mueller, B., Lilley, B.N., and Ploegh, H.L. (2006). SEL1L, the homologue of yeast Hrd3p, is involved in protein dislocation from the mammalian ER. Journal of Cell Biology *175*, 261-270.

Müller, B. (2009). Protein degradation from the mammalian endoplasmic reticulum. In Faculty of Chemistry (OPUS Publication Server of the University of Stuttgart: University of Stuttgart), pp. 163.

Oh, R.S., Bai, X., and Rommens, J.M. (2006). Human homologs of Ubc6p ubiquitin-conjugating enzyme and phosphorylation of HsUbc6e in response to endoplasmic reticulum stress. The Journal of biological chemistry *281*, 21480-21490.

Oslowski, C.M., and Urano, F. (2011). Measuring ER stress and the unfolded protein response using mammalian tissue culture system. Methods in enzymology *490*, 71-92.

Özkan, E., Yu, H., and Deisenhofer, J. (2005). Mechanistic insight into the allosteric activation of a ubiquitin-conjugating enzyme by RING-type ubiquitin ligases. Proceedings of the National Academy of Sciences of the United States of America *102*, 18890-18895.

Papaleo, E., Ranzani, V., Tripodi, F., Vitriolo, A., Cirulli, C., Fantucci, P., Alberghina, L., Vanoni, M., De Gioia, L., and Coccetti, P. (2011). An acidic loop and cognate phosphorylation sites define a molecular switch that modulates ubiquitin charging activity in Cdc34-like enzymes. PLoS computational biology *7*, e1002056.

Popp, M.W., and Ploegh, H.L. (2011). Making and breaking peptide bonds: protein engineering using sortase. Angewandte Chemie International Edition in English *50*, 5024-5032.

Pruneda, J.N., Littlefield, P.J., Soss, S.E., Nordquist, K.A., Chazin, W.J., Brzovic, P.S., and Klevit, R.E. (2012). Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases. Molecular cell *47*, 933-942.

Schmidt, F.I., Lu, A., Chen, J.W., Ruan, J., Tang, C., Wu, H., and Ploegh, H.L. (2016). A single domain antibody fragment that recognizes the adaptor ASC defines the role of ASC domains in inflammasome assembly. The Journal of Experimental Medicine *213*, 771-790.

Schoebel, S., Mi, W., Stein, A., Ovchinnikov, S., Pavlovicz, R., DiMaio, F., Baker, D., Chambers, M.G., Su, H., Li, D., *et al.* (2017). Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. Nature *548*, 352.

Semplici, F., Meggio, F., Pinna, L.A., and Oliviero, S. (2002). CK2-dependent phosphorylation of the E2 ubiquitin conjugating enzyme UBC3B induces its interaction with beta-TrCP and enhances beta-catenin degradation. Oncogene *21*, 3978-3987.

Shchebet, A., Karpiuk, O., Kremmer, E., Eick, D., and Johnsen, S.A. (2012). Phosphorylation by cyclin-dependent kinase-9 controls ubiquitin-conjugating enzyme-2A function. Cell cycle (Georgetown, Tex) *11*, 2122-2127.

Sheng, Y., Hong, J.H., Doherty, R., Srikumar, T., Shloush, J., Avvakumov, G.V., Walker, J.R., Xue, S., Neculai, D., Wan, J.W., *et al.* (2012). A human ubiquitin conjugating enzyme (E2)-HECT E3 ligase structure-function screen. Molecular & Cellular Proteomics *11*, 329-341.

Shi, Y., Vattem, K.M., Sood, R., An, J., Liang, J., Stramm, L., and Wek, R.C. (1998). Identification and characterization of pancreatic eukaryotic initiation factor 2 alpha-subunit kinase, PEK, involved in translational control. Molecular and cellular biology *18*, 7499-7509.

Sifers, R.N., Brashears-Macatee, S., Kidd, V.J., Muensch, H., and Woo, S.L. (1988). A frameshift mutation results in a truncated alpha 1-antitrypsin that is retained within the rough endoplasmic reticulum. The Journal of biological chemistry *263*, 7330-7335.

Simon, M.D., Heider, P.L., Adamo, A., Vinogradov, A.A., Mong, S.K., Li, X., Berger, T., Policarpo, R.L., Zhang, C., Zou, Y., *et al.* (2014). Rapid flow-based peptide synthesis. Chembiochem : a European journal of chemical biology *15*, 713-720.

Stewart, M.D., Ritterhoff, T., Klevit, R.E., and Brzovic, P.S. (2016). E2 enzymes: more than just middle men. Cell Research *26*, 423.

Sun, S., Shi, G., Han, X., Francisco, A.B., Ji, Y., Mendonça, N., Liu, X., Locasale, J.W., Simpson, K.W., Duhamel, G.E., *et al.* (2014). Sel1L is indispensable for mammalian endoplasmic reticulum-associated degradation, endoplasmic reticulum homeostasis, and survival. Proceedings of the National Academy of Sciences of the United States of America *111*, E582-E591.

Valimberti, I., Tiberti, M., Lambrughi, M., Sarcevic, B., and Papaleo, E. (2015). E2 superfamily of ubiquitin-conjugating enzymes: constitutively active or activated through phosphorylation in the catalytic cleft. Scientific Reports *5*, 14849.

Walter, P., and Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. Science (New York, NY) *334*, 1081-1086.

Wang, T., Wang, B., Huang, H., Zhang, C., Zhu, Y., Pei, B., Cheng, C., Sun, L., Wang, J., Jin, Q., *et al.* (2017). Enterovirus 71 protease 2Apro and 3Cpro differentially inhibit the cellular endoplasmic reticulum-associated degradation (ERAD) pathway via distinct mechanisms, and enterovirus 71 hijacks ERAD component p97 to promote its replication. PLos Pathogen *13*, e1006674.

Xu, Y., Zhao, F., Qiu, Q., Chen, K., Wei, J., Kong, Q., Gao, B., Melo-Cardenas, J., Zhang, B., Zhang, J., *et al.* (2016). The ER membrane-anchored ubiquitin ligase Hrd1 is a positive regulator of T-cell immunity. Nature Communication *7*, 12073.

Yamamoto, K., Suzuki, N., Wada, T., Okada, T., Yoshida, H., Kaufman, R.J., and Mori, K. (2008). Human HRD1 promoter carries a functional unfolded protein response element to which XBP1 but not ATF6 directly binds. The Journal of Biochemistry *144*, 477-486.

Yang, H., Qiu, Q., Gao, B., Kong, S., Lin, Z., and Fang, D. (2014). Hrd1-mediated BLIMP-1 ubiquitination promotes dendritic cell MHCII expression for CD4 T cell priming during inflammation. The Journal of Experimental Medicine *211*, 2467-2479.

Zaru, R., Edgar, A.J., Hanauer, A., and Watts, C. (2015). Structural and Functional Basis for p38-MK2-Activated Rsk Signaling in Toll-Like Receptor-Stimulated Dendritic Cells. Molecular and cellular biology *35*, 132-140.

Chapter 4. Using VHH05 as a peptide-binding nanobody

Jingjing Ling[1], Ross Cheloha[1], Joy Xie[1], Sarah A. Nordeen[2], Thomas Schwartz[2] and Hidde Ploegh[1,*]

1 Boston Children's Hospital, Boston, MA 02115

2 Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Correspondence: Hidde.Ploegh@childrens.harvard.edu

To be submitted

All experiments were designed by Jingjing Ling.

4.1 Introduction

Epitope tags are useful tools in molecular biology. Proteins of interest are often tracked with an antibody that specifically recognizes them. Antibodies can be used to quantify a target from a complex mixture (i.e. immunoblotting or enzyme-linked immunosorbent assay ELISA), to pull down proteins from a mix of proteins (immunoprecipitation; IP), and to stain cells or tissues carrying the protein of interest (immunofluorescence or immunohistochemistry). However, it takes at least a few months for immunization, and it is not always straightforward to generate a specific antibody with good affinity for every protein of interest. Biologists therefore often tag their proteins of interest with epitope tags where a good antibody has already been developed. Examples of commonly used epitope tags include the hemagglutinin (HA) tag derived from influenza hemagglutinin amino acid 98-106, and the Myc-tag derived from c-Myc transcription factor (amino acids 410-419) (Kimple et al., 2004). Epitope tags allow researchers to study their proteins of interest without having an applicable antibody to the endogenous protein. Thus, they have been shown to be important tools for biology.

Nanobodies, also known as VHHs, are the target-recognizing domain of heavy-chain only antibodies. They have attracted attention in the last decade because of their small size, high expression in both *E. coli* and mammalian cells, and ability to act as crystallization chaperones. While many VHHs have been identified and used to stabilize proteins for crystallization, few have been shown to recognize peptide epitopes, and even fewer have been used as tool for corresponding epitope tags. As of 2016, a nanobody that recognizes a 12-amino acid peptide on β-catenin has been published, and this peptide sequence can be used as an epitope tag (Braun et al., 2016). While a nanobody against the 4 amino acid tag EPEA has been published (Pardon et al., 2013), many proteins contain the EPEA motif, thus making it difficult for *in vivo* studies without the risk of broad cross reactivity. Chromotek, a company specialized in alpaca antibodies, recently made a Myc-specific nanobody available. While scientists in nanobody research have been working to identify VHHs that recognize simple linear epitopes, only a few have been published thus far.

We have previously reported identification of a VHH against UBC6e, called VHH05 or VHH 6E05, which recognizes amino acid 168-168 in UBC6e at low nanomolar affinity. UBC6e (168-186) can also be sortagged to irrelevant proteins and result in recognition by VHH05. In the current study, we would like to investigate the potential of using the described epitope as a tag appended to expressed proteins of interest. Beyond the use this VHH05-epitope pair as a mere substitute for conventional antibodies, we would also like to explore the added benefit of VHHs —specifically their lack of need for stable disulfide bonds and compatibility with intracellular environment — and as tagged proteins could be targeted to specific membranes by VHH05s carrying signal sequences.

4.2 Results

Identifying the minimal epitope tag

In the previous chapter, where VHH05 was used to help the study the functions of its target UBC6e, we showed that the epitope of VHH05 was within an 18-amino acid region that is predicted to be helical in structure. To find the minimal epitope, we synthesized various truncations of this 18-amino acid region with an N-terminal G_3 motif, and conjugated them onto EGFP-LPSTGG or Bio-LPETGG via a sortase reaction for analysis. We immunoprecipitated EGFP-peptide constructs with VHH05 or VHH "enhancer" specific for GFP. Peptides as short as 10 mer can still bind to VHH05, even though we observed that sequences shorter than the 14mer QADQEAKELARQIS showed reduced binding (Figure 4.1). We compared the recovery using VHH05 to that of VHH enhancer. For EGFP carrying a 14mer or 16mer peptide, VHH05 can recover more protein than VHH enhancer (Table 4.1). On the other hand, VHH05 can only recover 80% of total EGFP-peptide as compared to VHH enhancer. We used Octet to measure the affinity of VHH05 to biotin-LPETGG-peptides. Surprisingly, the 14mer peptide has the highest affinity, even higher than the 16mer construct. However, trimming beyond the 14mer reduced the affinity, by significantly increasing the observed k_{off}s. These increased k_{off}s are consistent with the observed lower recovery of the EGFPs carrying peptides shorter than 14mer.

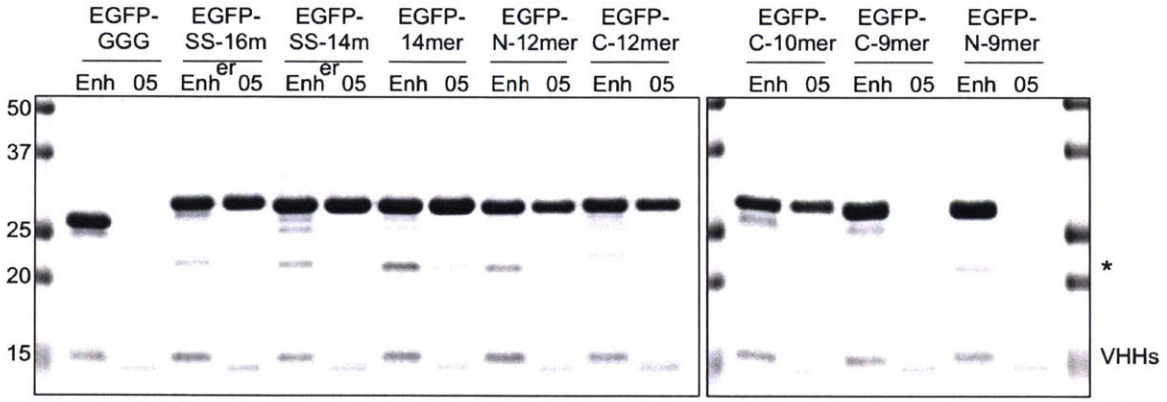


Figure 4.1. Identification of the minimal peptide epitope. Different fragments of the 18 amino acid region identified in the previous chapter were synthesized with a sortase recognition motif and conjugated onto EGFP. These EGFP constructs were immunoprecipitated with immobilized VHH enhancer or VHH05 in TBS (50 mM Tris, 150 mM NaCl, pH 7.4). Samples were eluted with 2x SDS sample buffer to ensure complete recovery.

Table 4.1. Affinities of peptide fragments to VHH05.

Immobilized Construct		K_d (nM)	k_{on}(M^{-1} s^{-1})	k_{off} (s^{-1})	% recovery by VHH05 compared to VHH enhancer
(SS-16mer)	Biotin-LPETGGG-SSQADQEAKELARQISFK	9 ± 1	9.4 ± 0.6 x 10^4	8.13 ± 0.5 x10^{-4}	110%
(SS-14mer)	Biotin-LPETGGG-SSQADQEAKELARQIS	0.6 ± 0.1	1.03 ± 0.06 x 10^6	5.6 ± 0.9 x10^{-4}	107%
(14mer)	Biotin-LPETGGG- - - QADQEAKELARQIS	3.3 ± 0.6	1.7 ± 0.2 x 10^5	5.7 ± 0.5 x10^{-4}	114%
(N-12mer)	Biotin-LPETGGG- - - - -DQEAKELARQIS	20 ± 5	2.2 ± 0.4 x10^5	4 ± 2 x10^{-3}	81%
(C-12mer)	Biotin-LPETGGG- - - QADQEAKELARQ	10.0 ± 0.5	2.8 ± 0.4 x10^5	2.8 ± 0.3 x10^{-3}	77%
(C-10mer)	Biotin-LPETGGG- - - QADQEAKELA	8.9 ± 0.5	4.6 ± 0.3 x10^5	4.1 ± 0.5 x10^{-3}	80%
(C-9mer)	Biotin-LPETGGG- SSQADQEA	N.A	N.A	N.A	1%
(N-9mer)	Biotin-LPETGGG- - - - - -QEAKELAR	N.A	N.A	N.A	1%

The 6e tag can be used for affinity purification

Based on all peptide fragments tested, the 14mer tag QADQEAKELARQIS currently serves as our best tag with the highest affinity. The 16mer tag QADQEAKELARQISFK was also analyzed to examine how the difference in binding affinities contributes to the property of this tag.

EGFP was expressed with an N-terminal 14mer or 16mer tag in *E. coli*. Transformed *E. coli* cells were grown in Terrific Broth (TB) overnight at 37°C without induction with IPTG. The EGFP constructs expresses very well when induced, with more than 80% of cytosolic protein being the target protein (Data not shown). To show that VHH05 can be used to purify tagged proteins from a more complex mixture, I decided to use uninduced *E. coli* for a lower expression level. From uninduced *E. coli* lysates, 14mer-EGFP and 16mer-EGFP are both recovered on CNBr-activated beads modified with VHH enhancer or VHH05, but not with the non-specific anti-MHCII VHH07 (Figure 4.2). With the same amount of lysate and similar amounts of VHHs used, VHH05 can immunoprecipitate much more 14mer-EGFP and 16mer-EGFP than VHH enhancer. Untagged EGFP is only immunoprecipitated by VHH enhancer, not VHH05, showing that its specificity persists even in a complex environment such as a bacterial lysate.

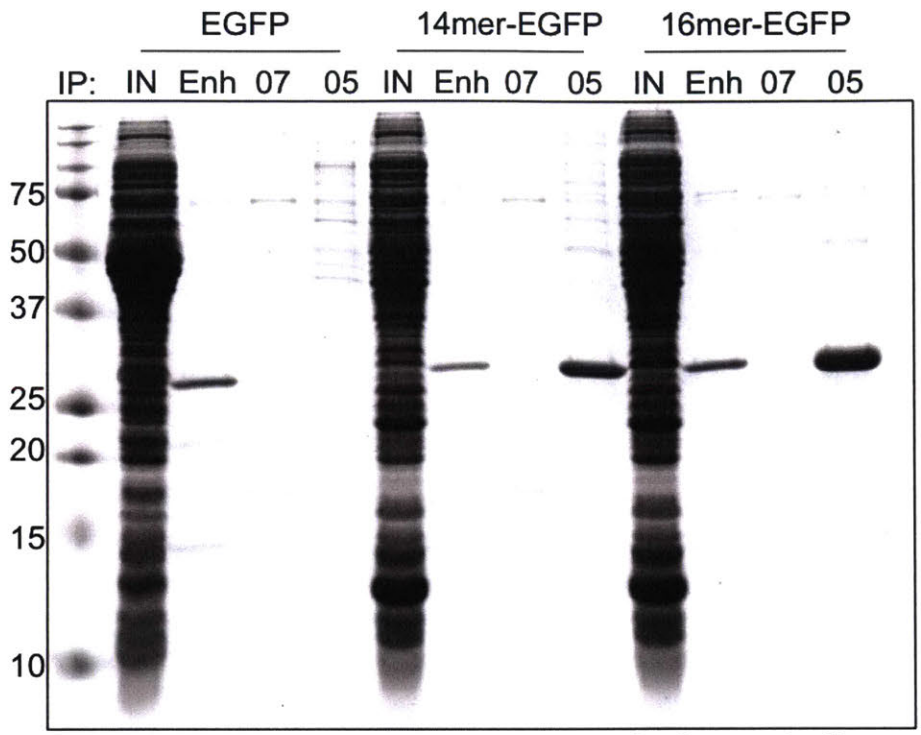


Figure 4.2. 14mer-EGFP and 16mer-EGFP are both specifically recovered by immobilized VHH05. *E. coli* lysates were collected by sonication. Beads were washed 3 times before samples were eluted using 2x SDS sample buffer. Enh: anti-GFP VHH enhancer. 07: anti-MHCII VHH07. 05: anti-UBC6e VHH05.

For affinity purification, it is very important to be able to elute the samples collected under mild conditions. The 6e tagged protein is readily eluted by incubating the beads with a small volume of 1mM peptide YGQADQEAKELARQIS in TBS for 30 minutes at room temperature. 83 ± 6 % of 16mer-EGFP was eluted using peptide, while only 60 ± 10 % of 14mer-EGFP was eluted (Figure 4.3, n=3). This discrepancy can be partially explained by the differences in the observed in K_d values, where the 14mer tag has a higher binding affinity than that of the 16mer tag. Though giving a slightly lower yield was observed, the proteins eluted with 1 mM peptide are markedly cleaner with fewer contaminants than those eluted with 2% SDS sample buffer.

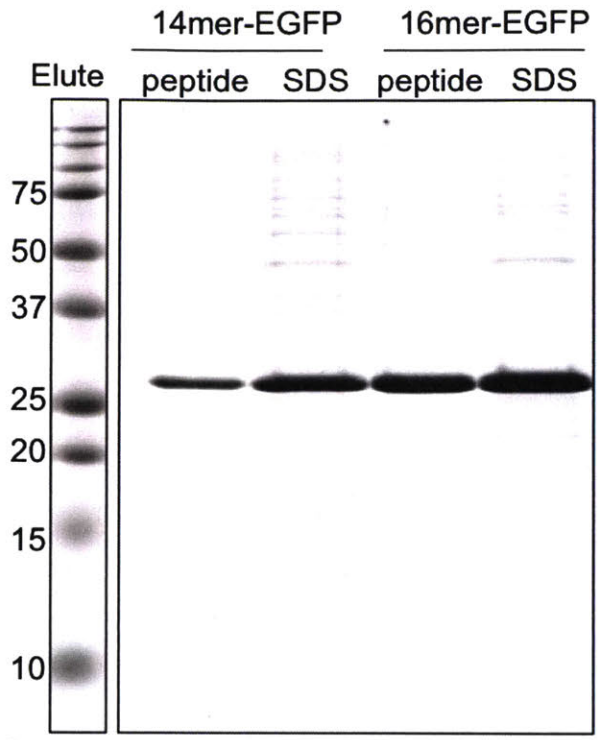


Figure 4.3. 14mer-EGFP and 16mer-EGFP can be eluted from VHH05 beads using 1 mM YGQADQEAKELARQIS peptide in TBS.

The 6e tag can be used for ELISA and sandwich ELISA

The enzyme-linked immunosorbent assay (ELISA) has been widely used for specific detection and quantification of proteins of interest. Since VHH05 was identified using ELISA-based high-throughput screening, we know it can recognize purified UBC6e immobilized on ELISA plates. We first tested the detection limit of immobilized UBC6e, EGFP-LPSTGG-14mer, and EGFP-LPSTGG-16mer using VHH05-biotin. Various amounts of the protein constructs were directly immobilized to ELISA plates, and 1 μg/mL of VHH05-biotin was used to detect the immobilized constructs. The response between 3 ng to 50 ng of immobilized antigen is close to a linear response. Under these conditions, we can detect as little as 3 ng of UBC6e and 6 ng of EGFP-14mer/16mer constructs

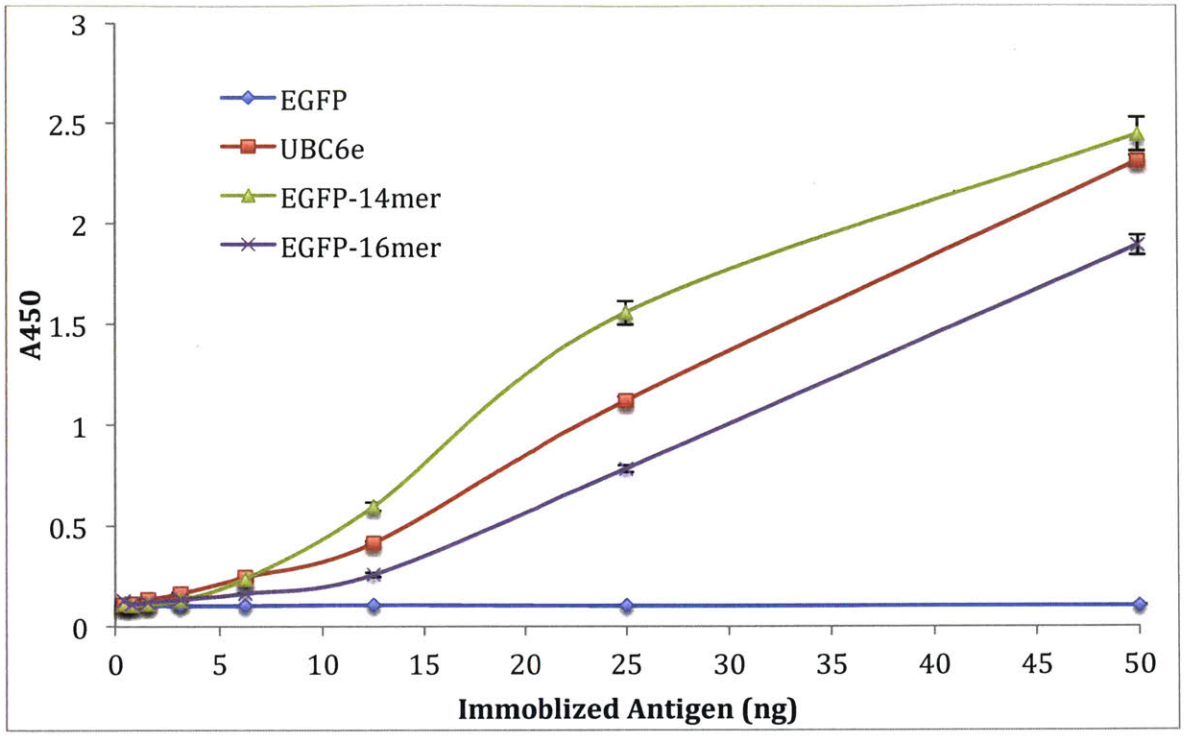
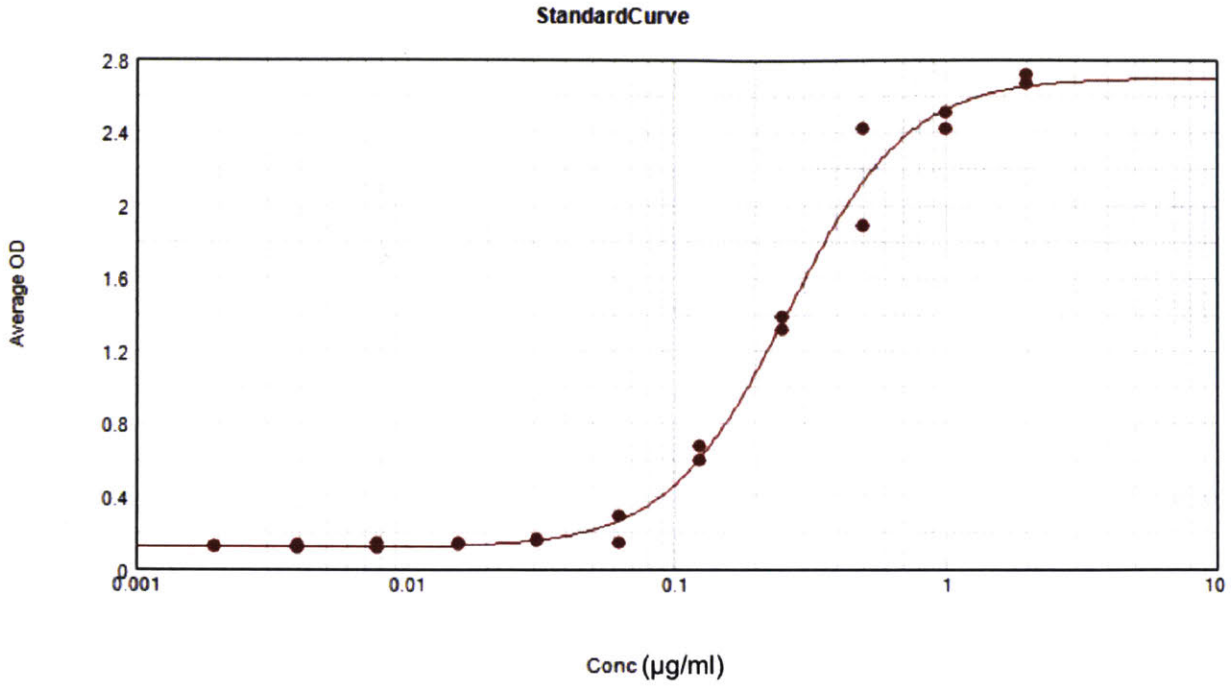


Figure 4.4a. VHH05-biotin can be used to detect purified UBC6e, EGFP-14mer, and EGFP-16mer via ELISA. Streptavidin-HRP and TMB were used for the detection of bound VHH05-biotin.

Sandwich ELISA is an economic, fast, and quantifiable method to detect small amounts of specific analytes in a complex mixture. It is a method that has been widely used in diagnostics such as in pregnancy and HIV infection tests. For sandwich ELISA, an antibody is first immobilized to a surface. When samples are incubated on this surface, the first antibody selectively captures the analyte. A second antibody that binds to a different epitope can then be used to detect and quantify the analyte of interest. We immobilized VHH05 to ELISA plates to test whether it can specifically capture tagged-EGFPs from solution. We could indeed capture purified EGFP-LPSTGG-14mer and generate a clean and reproducible standard curve (Figure 4.4b, n=2). We also aimed at quantifying the amount of tagged EGFP in crude *E. coli* lysate through ELISA. We made serial dilutions of the *E. coli* lysate in PBS, and using the results within the linear range of 0.5-2 absorbance units, we found the concentration of 14mer-EGFP in *E. coli* lysate to be 130 ± 60 μg/mL (n=4) and the concentration of 16mer-EGFP in *E. coli* lysate to be 200 ± 60 μg/mL (n=4).



Curve Fit Results ▲

Curve Fit : 4-Parameter $y = D + \frac{A - D}{1 + \left(\frac{x}{C}\right)^B}$

	Parameter	Estimated Value	Std. Error	Confidence Interval
STD#1 $R^2 = 0.992$ EC50 = 0.261	A	0.118 (Abs)	0.033	[0.049, 0.188]
	B	1.956 (Abs)	0.198	[1.540, 2.371]
	C	0.261 (µg/ml)	0.016	[0.228, 0.294]
	D	2.700 (Hills coefficient)	0.075	[2.543, 2.858]

Figure 4.4b. VHH05 can be immobilized to ELISA plates to capture tagged protein from a complex mixture for sandwich ELISA quantification. Sortagged EGFP-14mer was used as a standard. ELISA was conducted in 5% milk in TBST (TBS with 0.1% Tween-20). EGFP was detected using polyclonal rabbit anti-GFP (Abcam Ab290) and HRP-conjugated anti-Rb secondary antibody.

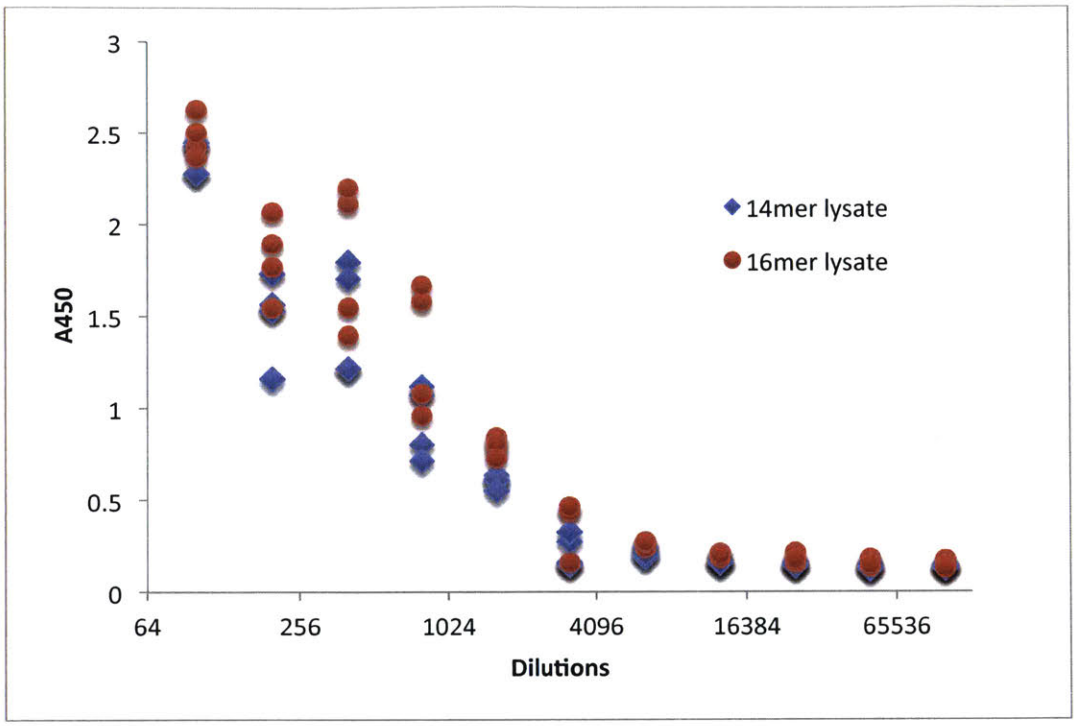


Figure 4.4c. 14mer-EGFP and 16mer-EGFP can be captured from *E. coli* lysates and quantified. Dilutions with readings between 0.5-2 absorbance units were used to calculate for the concentrations of tagged EGFPs in the lysates.

The 6e tag can be used to tag proteins in mammalian cells

While epitope tags have been used for purification of recombinant proteins in *E. coli*, they are much more commonly used for expression in mammalian cells for overexpression of proteins of interest. We transduced 6e-tagged proteins to mammalian cells to test whether VHH05 could be used to detect their expression in immunoblots, FACS and immunoprecipitation. We expressed EGFP, N-terminal tagged 14mer-EGFP, or 16mer-EGFP in HeLa and UBC6e$^{-/-}$ MEF cells using lentiviral transduction using pCDH-CMV-EF1-puro plasmid for constitutive expression (Figure 4.5a). We collected the lysates and performed immunoprecipitations with VHH05 or anti-MHCII VHH07. When blotting with 5 µg/mL VHH05-biotin, both 14mer-EGFP and 16mer-EGFP are visible on the blot. There is a prominent non-specific band at ~37kD (labeled with *) in the lysates that are also visible on the immunoblot. These non-specific targets are not immunoprecipitated by immobilized VHH05, and they are not UBC6e since they are present in UBC6e$^{-/-}$ MEF cells, and are not recognized by polyclonal rabbit anti-UBC6e (Figure 4.5b). The immunoprecipitated tagged-EGFP can be blotted with VHH05-biotin, anti-GFP, and anti-UBC6e. In UBC6e$^{-/-}$ MEF cells, only tagged EGFP can be detected while the endogenous protein can also be observed in HeLa cells. The tagged protein is specifically immunoprecipitated only by VHH05, but not

by anti-MHCII VHH07 used at similar concentration.

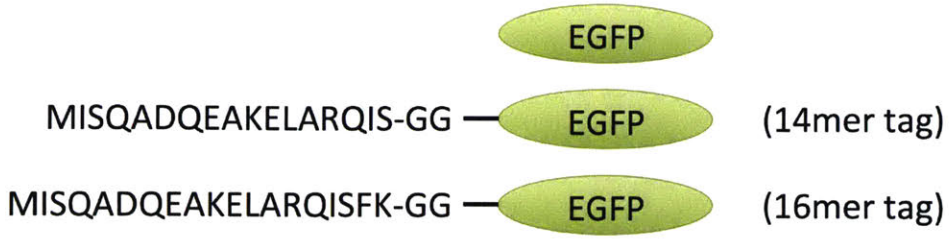


Figure 4.5a. Constructs used for expressing 6e tagged EGFPs in mammalian cells. These constructs were inserted into pCDH-CMV-EF1-puro plasmids using NheI and BamHI sites with a GCCACC Kozak sequence immediately in front of the ATG start codon.

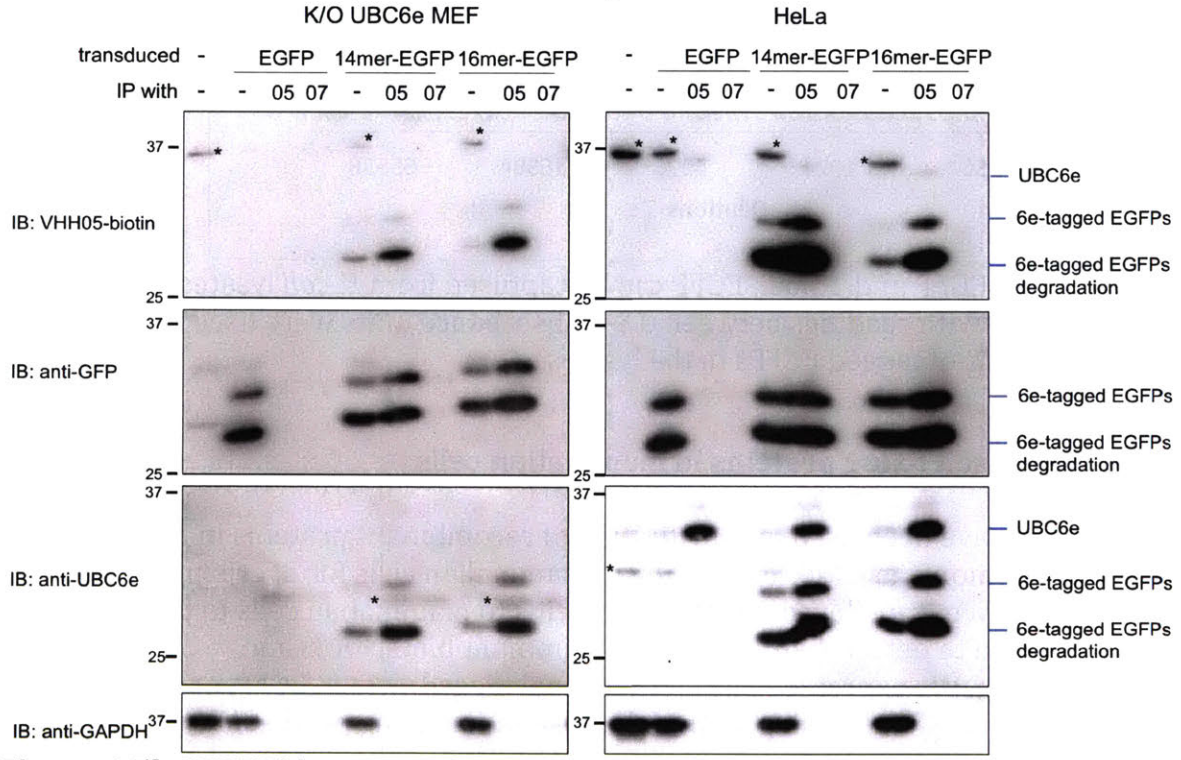


Figure 4.5b. VHH05-biotin can detect tagged proteins in an immunoblot in mammalian cell lysates. VHH05 immunoprecipitates 6e-tagged proteins and endogenous UBC6e from cell lysates. *: non-specific bands

In order the test whether VHH05 labeled with Alexa647 can be used to detect expressed 6e-tagged EGFP, we trypsinized, fixed, and permeablized cells, and used 5 μg/mL VHH05-LPETGG-Alexa647 to stain these cells. The stained cells were then analyzed by FACS. Since we observed strong non-specific signals when mammalian cell lysates were immunoblotted, we first wanted to examine the intensity of non-specific signals. Even for UBC6e$^{-/-}$ MEF cells, there was a 10^3-10^4 increase in fluorescence when cells were stained with VHH05 (Figure 4.5c). A similar increase in fluorescence was also observed for HeLa cells, where endogenous UBC6e is present for specific staining. However, when we compared cells that express the 6e-tagged EGFP proteins, we still observed a further shift

that we attribute to the tagged protein (Figure 4.5d). The Alexa647 signal also correlates well with GFP signal observed, indicating that we can use VHH05-Alexa647 to monitor protein expression in mammalian cells.

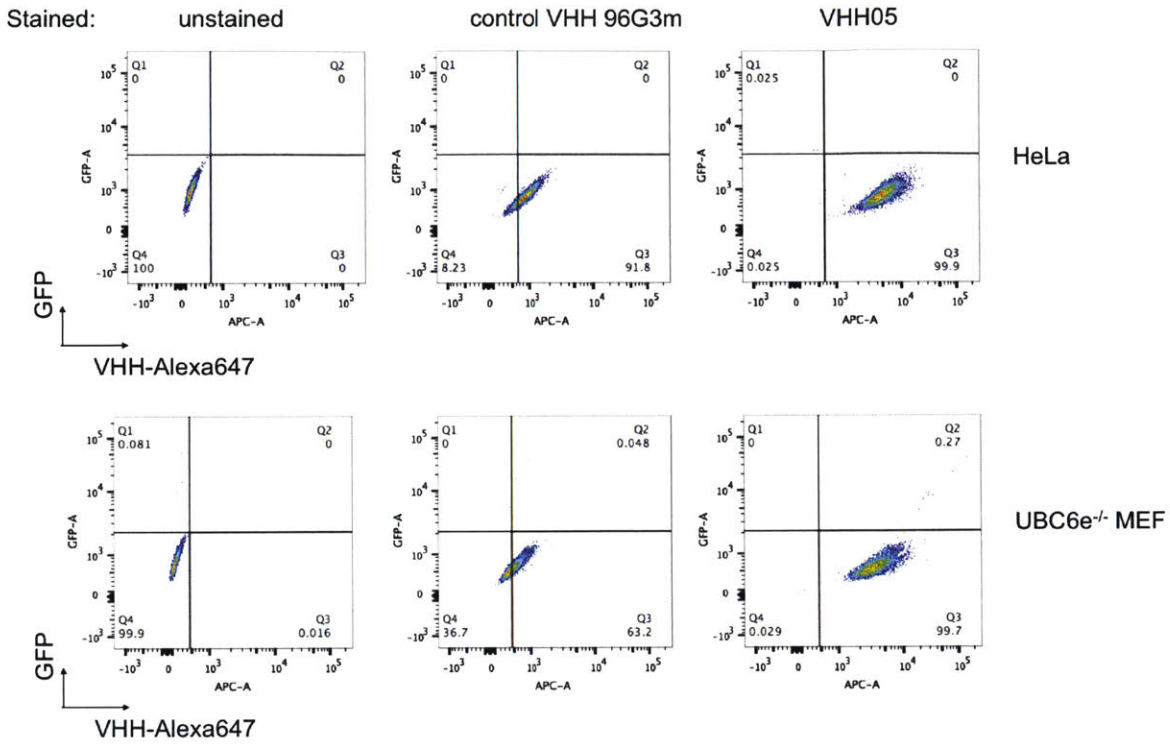


Figure 4.5c. VHH05 non-specifically stains UBC6e$^{-/-}$ MEF and HeLa cells.

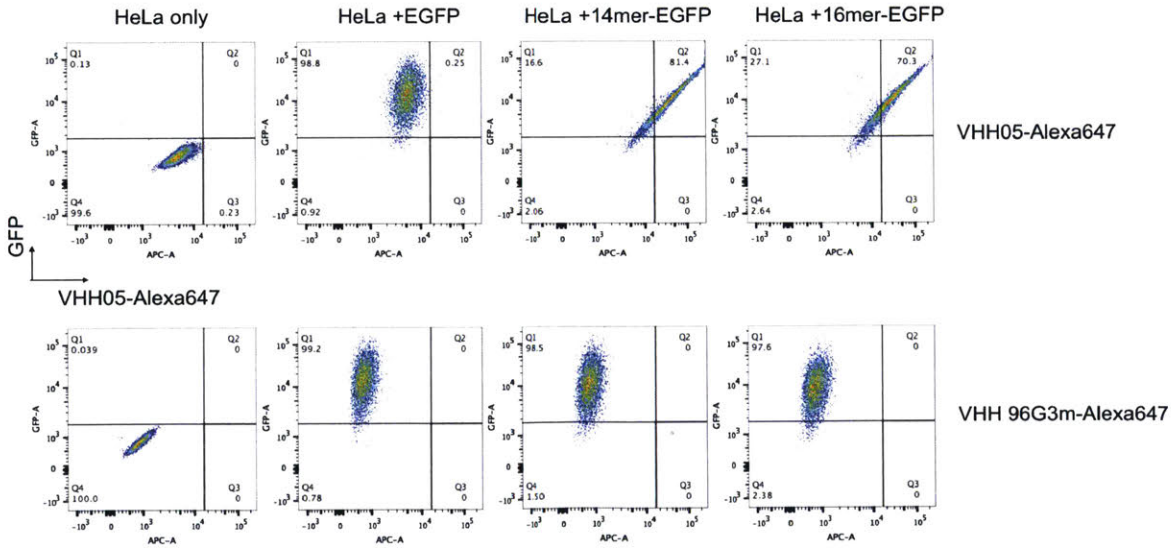


Figure 4.5d. VHH05-Alexa647 can stain tagged EGFP constructs, and can be used to monitor protein expression in HeLa cells.

Work in Progress: using VHH05 to target 6e-tagged proteins to desired locations in the cell

A significant benefit of VHHs over traditional antibodies is that they do not require disulfide bonds to maintain their binding properties. We have also previously shown that cytoplasmically expressed VHH05-HA co-immunoprecipitates with endogenous UBC6e. This antibody-antigen interaction is both specific and tight, thus we should be able to engineer new protein-protein interactions. To test our hypothesis, we first asked whether VHH05 could be used to target tagged constructs to selective intracellular compartments. VHH05 is tagged with different tail-anchor targeting motifs, and we express them using a deoxycycline-inducible system to target the tagged protein-of-interest to selective membranes. We tagged VHH05 with the transmembrane domain of UBC6e for targeting to the ER membrane, as well as the MAVS motif, which should target VHH05 to the outer membrane of mitochondria (Seth et al.) (Figure 4.6a).

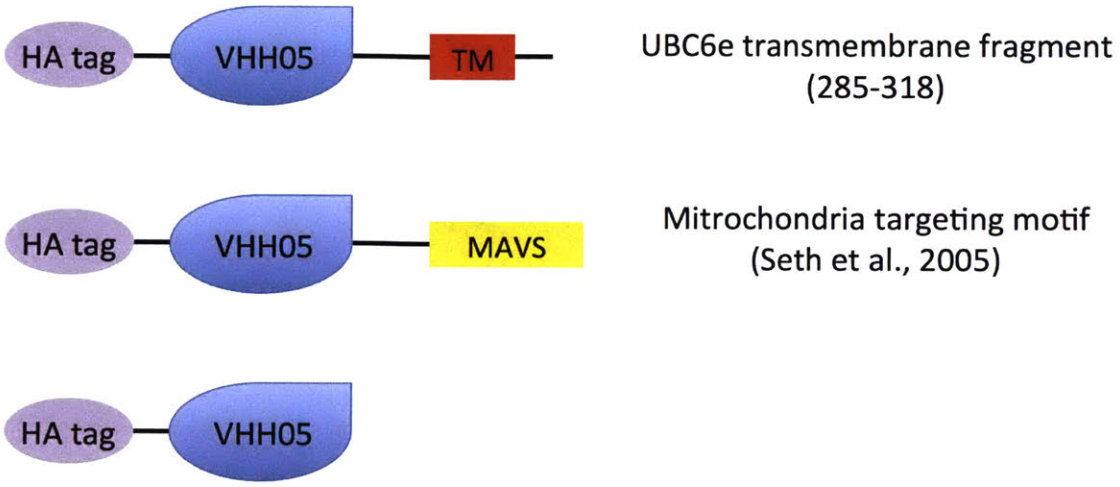


Figure 4.6a. Constructs used for targeting VHH05 to the ER (UBC6e transmembrane domain) or mitochondria (MAVS).

Work in Progress: obtaining a crystal structure of the VHH05-peptide complex

A crystal structure should allow us to observe the molecular interactions between VHH05 and its target. While crystallization of UBC6e has been difficult even in the presence of VHH05, we turned to crystallization of VHH05 with only the peptide epitope it recognizes. The peptide we chose was YGQADQEAKELARQIS, where an A280-sensitive Tyr was appended to the 14mer that measured the lowest K_d. The first crystal was obtained by incubating expressed tag-free VHH05 with the HPLC purified synthetic peptide, followed by purification of the presumptive complex by size exclusion. However, the structure obtained contained two VHH05 molecules domain swapped to form a stable dimer and no peptide. Currently we are co-expressing VHH05 with SUMO-tagged peptide epitope, where the tag was removed to allow for collection of only the VHH05-peptide complex.

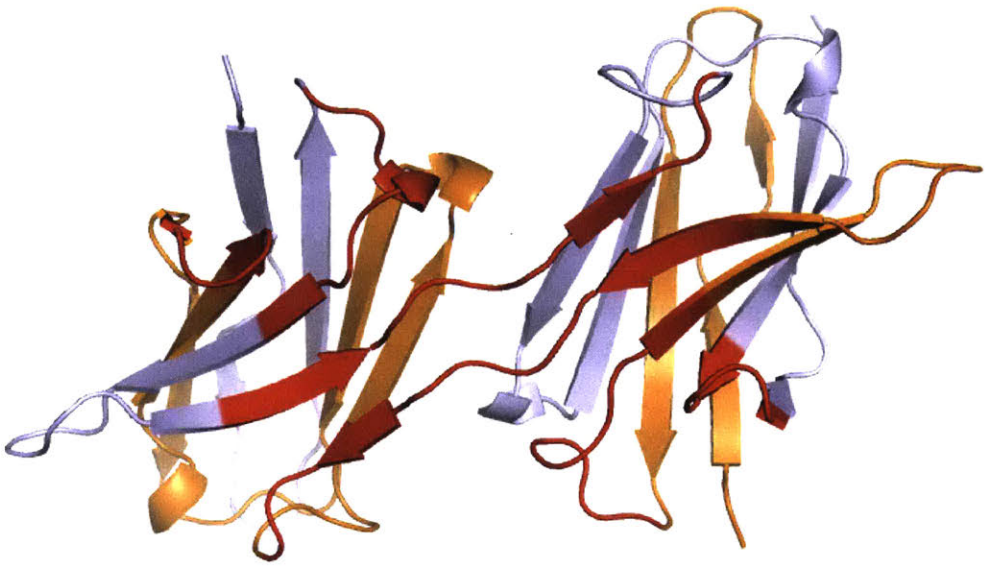


Figure 4.7a. Crystal structure of apo VHH05 dimers that showed extensive domain swap. The map has 1.3 Å resolution. Purple: residues from chain A. Orange: residues from chain B. Red: residues involved in CDR1, CDR2 and CDR3, the three loops that usually contribute most to binding to target.

4.3 Discussion

We have shown that the UBC6e-specific VHH05 recognizes its 14mer epitope in the absence of the remainder of UBC6e. This 14mer epitope can be used to tag proteins-of-interest and thus can be used as a nanobody-peptide tag system.

While peptide epitope tags have been widely used with conventional antibodies, only a few tag-specific nanobodies have been described. Nanobodies present a range of advantages over conventional antibodies. First of all, VHHs can be easily expressed in *E. coli* in good yield, which drastically reduces their production costs. In addition, we routinely express our VHHs with a LPETGG sortag motif, with which we can easily site-specifically install biotin molecules to be used with Streptavidin-HRP or magnetic Streptavidin beads, fluorescent molecules for FACS analysis, and even radioactive isotopes for PET imaging (Rashidian et al., 2015).

When conventional antibodies are used for immunoprecipitation, we usually cannot use the same antibodies for detection. The excess antibody heavy chain (~50kD), light chain (~20kD) and heavy-light chain dimer (~75kD) react with secondary antibodies and affect the signal-to-noise ratio. Conformation-specific secondary antibodies that only recognize the native antibodies have been used to solve this, but they are usually not as sensitive and specific as the conventional ones. Since VHHs are usually much smaller than their targets,

and are directly conjugated to CNBr-activated sepharose beads without biotin, the immunoprecipitated samples can be directly analyzed by the same VHH that is labeled with biotin. They can also be analyzed with conventional antibodies for co-immunoprecipitated targets with no interference, making them perfect alternatives for IP-immunoblot analysis.

In conclusion, we have shown that the 14mer peptide fragment recognized by the UBC6e-specific VHH05 can be used as an epitope tag that can be grafted onto other proteins of interest and retain binding affinity. This epitope tag can be used similarly to a conventional antibody in immunoblots, immunoprecipitations, FACS and ELISA. In addition, expressed VHH05 can be used to target tagged constructs to different locations within the cell. We have shown that VHHs that recognize short peptide fragments have the potential to be developed into useful tools to study protein functions by introducing new stable protein-protein interactions.

4.4 Material and Methods

Antibodies and reagents

GAPDH (Horseradish peroxidase (HRP)-conjugated monoclonal rabbit; Cell Signaling Technology), GFP (polyclonal rabbit; Abcam) were purchased. Rabbit anti-UBC6e(1:3000) (Mueller et al., 2008) were produced in our laboratory. Secondary antibodies for immunoblotting: HRP-conjugated goat anti-rabbit antibody (Southern Biotech) was used at 1:10,000, High sensitivity streptavidin-HRP (Pierce, MA) was used at 1:5,000 dilutions. Doxycycline was purchased from Clonetech. CNBr-activated Sepharose 4B were purchased from Sigma-Aldrich.

Plasmids, Protein expression and purification

pCMV-CDH-EF1-puro UBC6eΔTM and UBC6e(1-197) constructs were cloned from previously described pINDUCER UBC6e plasmids (Hagiwara et al., 2016). pET 30b(+) and pCMV-CDH-EF1-puro 14mer- and 16mer-EGFP plasmids were cloned from pET SUMO EGFP-LPSTGG-His_6 previously described (Ling et al., 2012) using extending primers coding for the 14mer and 16mer region. pINDUCER VHH05 with different C-terminal targeting motifs were cloned from geneblocks (IDT technologies).

Plasmids coding for pET30b(+) 14mer, 16mer EGFP, or EGFP only were transformed into Rosetta BL21(DE3) *E. coli.* An overnight starter culture was prepared in LB media with 70 μg/mL kanamycin and 50 μg/mL chloramphenicol. Bacteria were grown in TB media at 37°C overnight without induction. Bacteria pellets were collected by spinning down at 6000 rpm for 15 minutes. Pellet from a 100 mL culture was resuspended in 5 mL PBS, cooled on ice for 30 minutes, and lysed by sonication for 60 seconds 3 times. Lysates were collected as supernatants after spinning down at 13,000 rpm for 20 minutes.

VHHs were expressed using the pHEN6 vector. Plasmids coding for PelB-VHH-LPETGG-His_6 were transformed into WK6 *E.coli* bacteria. Bacteria were grown in TB media with 100 μg/mL ampicillin at 37°C until the OD_{600} reached 0.5, and protein expression was induced with 1 mM IPTG at 30°C overnight. Bacteria were collected by spinning down at 6000 rpm

for 15 minutes, and pellet from a 1 L culture was resuspened in 15 mL TES buffer (50 mM Tris, 650 μM EDTA, 2M sucrose) to prepare for osmotic shock. After incubating for 2 hours at 4°C, 75 mL distilled H_2O was added and the bacteria was further incubated overnight at 4°C. His-tagged VHHs were purified by binding to Ni-NTA beads, washed with PBS, and eluted with 250 mM imidazole in 50 mM Tris, 500 mM NaCl, pH 8.0. The eluent was further purified with an S75 column. A typical yield is 20-40 mg per 1 L culture.

Cell Culture

Wildtype(WT) and UBC6e$^{-/-}$ MEF cells were maintained as previously described (Hagiwara et al., 2016). Briefly, MEF cells were maintained in Dulbecco's modified Eagle's medium (DME) supplemented with 10% IFS and 0.0007% v/v b-mercaptoethanol (MEF medium) in humidified air containing 5% CO_2 at 37°C. HeLa cells were maintained in DME supplemented with 10% IFS.

Immobilization of VHHs to CNBr-activated sepharose beads

200 mg of CNBr-activate sepharose beads (Sigma-Aldrich) were washed twice with 20 mL of 1 mM HCl, and neutralized with 100 mM $NaHCO_3$, 500 mM NaCl, pH 8.3. 2 mg of each VHH (VHH 6E05, VHH enhancer, or VHH MHCII07) in PBS was added to beads in 20 mL of 100 mM $NaHCO_3$, 500 mM NaCl, pH 8.3 and reacted overnight at 4°C. The beads were blocked with 50 mM Tris, 500 mM NaCl, pH8.0 for 2 hours at room temperature, and then TBS containing 1% BSA overnight before use.

Generation of stable cells using lentiviral transduction

Lentivirus was produced in HEK293 cells. HEK cells were plated in 6-well plates and grown until 70% confluency, and transfected with 1.5 μg of plasmids containing the gene of interest (backbones in pINDUCER20 or pCDH-CMV-MCS-EF1-Puro), 0.65 μg psPAX2 and 0.35 μg pMD2.G. Media containing lipofectamine 2000/plasmid complexes were removed after 6 hours of transfection, and HEK cells were grown in DMEM containing 10% heat-inactivated FBS for 2 days. Virus particles were harvested by spinning down the HEK cell pellet, and passing the media through a 0. 45 μm filter. The media containing virus particles were mixed 1:1 with fresh media and used to transduce MEF or HeLa by incubating for 6 hours. The cells were grown in regular media overnight and selected with 500 μg/mL geneticin for pINDUCER20-transduced cells and 2 μg/mL puromycin for pCDH-CMV-MCS-EF1-Puro-transduced cells.

Immunoblotting and immunoprecipitation

Cultured cells were lysed in HEPES-buffered saline (HBS) with 1% NP-40 supplemented with protease inhibitor cocktail (complete protease inhibitor cocktail tablets; Roche Applied Science) on ice for 20 minutes. Cell debris was removed by spinning at 13,000 rpm for 20 minutes at 4°C. Protein concentration was determined using a BCA protein assay (Thermo Fisher Scientific, Inc.). Proteins were denatured in SDS sample buffer containing 0.1M DTT, resolved by SDS-PAGE, transferred to PVDF membranes, and immunoblotted using standard procedures in 5% milk in TBST. Specifically, GFP and UBC6e were detected with respective rabbit primary antibodies, followed by anti-rabbit IgG HRP. 6e-tagged constructs were detected with 5 μg/ml VHH05-biotin, and then Streptavidin-HRP.

All immunoprecipitation procedures were performed at 4°C. Mammalian cell lysates collected in HBS with 1% NP-40 and incubated with CNBr-activated sepharose beads conjugated with VHH05 (50 μg bead volume, 30 μg VHH-LPETGG-His_6) or anti-HA agarose beads (Thermo Scientific) overnight. The beads were washed 3 times with TBS with 1% NP-40 and resuspended in SDS-PAGE sample buffer with 0.1 M DTT. For *E. coli* lysates, incubation and washing was conducted without detergent.

Peptide Synthesis and purification

Peptides were synthesized using flow-based solid-phase peptide synthesizer using Fmoc chemistry as previously described(Simon et al., 2014). Peptides were conjugated to Rink-amide linker resin to produce C-terminal amides. Phospho-serine was incorporated on solid phase using Fmoc-L-Ser(HPO3Bzl)-OH (Chem Impex). Following synthesis beads were air-dried and peptides were cleaved from resin and deprotected in 92.5% trifluoroacetic acid, 5% H_2O, and 2.5% TIPS. Peptides were precipitated into cold diethyl ether, air-dried, and purified using reverse phase C18 HPLC using 10-50% Acetonitrile/H_2O gradient.

Sortase reactions

VHH-LPETGG-His_6, eGFP-LPSTGG-His_6, or biotin-LPETGG were incubated with 500 μM of GGG-peptides in the presence of 5 μM HeptMutSrt-His_6 at room temperature for 1 hour or 4°C overnight. Unreacted substrates and sortase were removed by incubating with Ni-NTA beads, and excess peptides were removed using P10 column for VHHs and eGFP, and HPLC for biotin. A typical yield is 50% after purification.

Octet measurements

All Octet experiments were conducted at Biophysical Instrumentation Facility at MIT using the Octet RED96 Bio-Layer Interferometry. The target of interest is biotinylated through sortagging with Biotin-LPETGG. All Octet measurements were taken in 1% BSA in PBS with 1% Tween. 10 ng/mL of biotin-labeled target were loaded on Dip and Read™ Streptavidin (SA) Biosensors (Pall Fortebio) until the reading exceeds 1 nm. These sensors were then dipped into wells containing different concentrations of VHH05, ranging from 50 nM to 8 μM. Average and standard deviations were taken for at least 3 measurements.

4.5 References

Braun, M.B., Traenkle, B., Koch, P.A., Emele, F., Weiss, F., Poetz, O., Stehle, T., and Rothbauer, U. (2016). Peptides in headlock--a novel high-affinity and versatile peptide-binding nanobody for proteomics and microscopy. Scientific reports *6*, 19211.

Hagiwara, M., Ling, J., Koenig, P.A., and Ploegh, H.L. (2016). Posttranscriptional Regulation of Glycoprotein Quality Control in the Endoplasmic Reticulum Is Controlled by the E2 Ub-Conjugating Enzyme UBC6e. Molecular cell *63*, 753-767.

Kimple, M.E., Brill, A.L., and Pasker, R.L. (2004). Overview of Affinity Tags for Protein Purification. In Current Protocols in Protein Science, p. Unit 9.9.

Ling, J.J., Policarpo, R.L., Rabideau, A.E., Liao, X., and Pentelute, B.L. (2012). Protein thioester synthesis enabled by sortase. Journal of the American Chemical Society *134*, 10749-10752.

Mueller, B., Klemm, E.J., Spooner, E., Claessen, J.H., and Ploegh, H.L. (2008). SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. Proceedings of the National Academy of Sciences of the United States of America *105*, 12325-12330.

Pardon, E., Steyaert, J., and Wyns, L. (2013). Epitope tag for affinity-based applications (Google Patents).

Rashidian, M., Keliher, E.J., Bilate, A.M., Duarte, J.N., Wojtkiewicz, G.R., Jacobsen, J.T., Cragnolini, J., Swee, L.K., Victora, G.D., Weissleder, R., *et al.* (2015). Noninvasive imaging of immune responses. Proceedings of the National Academy of Sciences of the United States of America *112*, 6146-6151.

Seth, R.B., Sun, L., Ea, C.-K., and Chen, Z.J. Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF-kB and IRF3. Cell *122*, 669-682.

Simon, M.D., Heider, P.L., Adamo, A., Vinogradov, A.A., Mong, S.K., Li, X., Berger, T., Policarpo, R.L., Zhang, C., Zou, Y., *et al.* (2014). Rapid flow-based peptide synthesis. Chembiochem : a European journal of chemical biology *15*, 713-720.

Chapter 5 Work in Progress: Delivery of recombinantly expressed proteins into mammalian cell cytosol

5.1 Introduction

While macromolecules like antibodies have shown good properties as therapeutic agents, their uses have been limited to extracelullar targets. Cell surface membranes are impermeable to macromolecules, including proteins, and delivery of proteins from extracellular environment into the cytosol has always been a challenge in the field.

During my PhD, I have attempted using two different methods to deliver recombinant protein molecules into the cell: (1) utilizing a bacterial toxin that naturally transfer one of its domains into cytosol as a carrier for our proteins of interest; (2) utilizing a platform termed CellSqueeze developed in the Langer and Jensen lab, where cells were squeezed through a microfluidic device to allow temporary disruption of their membranes (Sharei et al., 2013). We are focusing on delivering HA-tagged ubiquitin (HA-Ub) and HA-Ub loaded E2s into the cells. We aim to understand the substrate control by E2s, the kinetics where HA-Ub are utilized by different ubiquitin pathways, as well as tools to understand E2-E3 co-operativity.

5.2 Utilization of diphtheria toxin(DT) as a tool for protein delivery

Advisor: Professor Bradley L. Pentelute

5.2.1 Introduction

Diphtheria, the infectious disease caused by a gram-positive pathogenic bacteria *Corynebacterium diphtheriae*, had been a wide-spread disease that caused many childhood deaths. While mass immunization program has eradicated the disease in most developed countries, there still have been some small outbreaks since the 1990s. The bacteria usually only colonize in surface epithelial cells in human body, such as skin and the respiratory tract. The direct cause of disease and death is usually from diphtheria toxin (DT), which is an exotoxin secreted by the bacteria, circulated in bloodstream and causes multiple organ failures (Collier, 1975). Wild-type DT is very toxic to sensitive species, such as human, primates, guinea pigs, rabbits, and many birds, with killing doses of 50-100 ng/kg.(Pappenheimer, 1977) However, rodents are resistant to the damage of DT. Given the high cytotoxicity of DT, I mostly work with an E148S mutant that is about 800-fold less toxic than the wild-type under federal regulation.(Barbieri and Collier, 1987)

Diphtheria toxin is synthesized by *Corynebacterium diphtheriae* as one single polypeptide with an N-terminal signal sequence that targets the protein for secretion via the Sec system.(Greenfield et al., 1983) The signal sequence is cleaved after secretion, leaving the mature DT with 535 amino acids that is composed of three domains: catalytic domain (1-

193), translocation domain (194-386), and receptor-binding domain (387-535). An alternative system describes the catalytic domain as the A fragment (DTA) while the translocation domain and receptor-binding domain are collectively called the B fragment (DTB) (Figure 5.2.1a). Eisenberg's group has solved the crystal structure of dimeric DT in 1994 (Figure 5.2.1b).(Bennett et al., 1994) Though many active site residues are conserved and the overall fold of the DT catalytic domain is similar to other ADP-ribosylating bacterial toxins, such as exotoxin A, there is little sequence homology among the members of the ADP-ribosylation enzyme family.(Wilson and Collier, 1992; Yates et al., 2006)

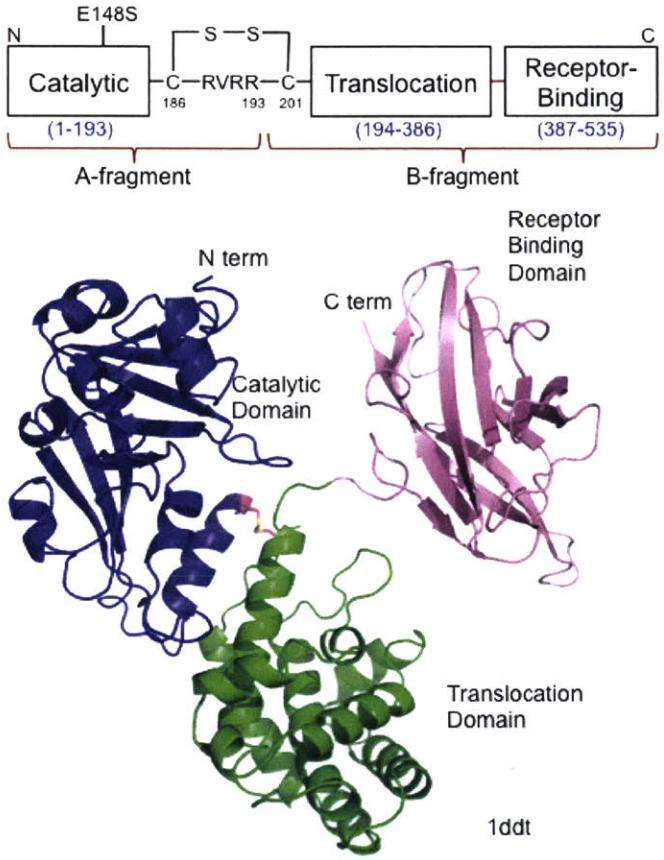


Figure 5.2.1. (a) Linear construct of diphtheria toxin (DT). **(b)** Crystal structure of DT (pdb 1ddt). Blue: catalytic domain. Green: translocation domain. Pink: Receptor-binding domain.

DTA is the catalytic domain that inhibits protein synthesis on ribosome by enzymatically ADP-ribosylating elongation factor-2 (EF-2).(Collier, 1975) The site of ADP-ribosylation is on diphthamide, a conserved post-tranlationally modified histidine (H715 in mammalian and H699 in yeast) across eukaryotic EF-2s. The reaction catalyzed by DTA is shown in Figure 5.2.2. The equilibrium for this reaction lies very much to the right with the equilibrium constant K estimated to be 6.3×10^{-4} M, meaning more than 99.99% EF-2s will be ADP-ribosylated in cytosolic concentrations.(Collier, 1975) Yamaizumi *et al.* shows that one molecule of DTA introduced into the cell through erythrocyte ghost fusion is enough to kill a mouse L-cell when monitored after 7 days.(Yamaizumi et al., 1978) For wild-type DTA, the measured K_m for NAD^+ is about 10 μM and the K_m for EF-2 is about 1 μM. The k_{cat} is about 60-70 min^{-1}. For the E184S mutant I am working with, the measured K_m for NAD^+

is about 9 μM and the K_m for EF-2 is about 2 μM. The k_{cat} is about 0.2-0.5 min^{-1}.(Wilson et al., 1990)

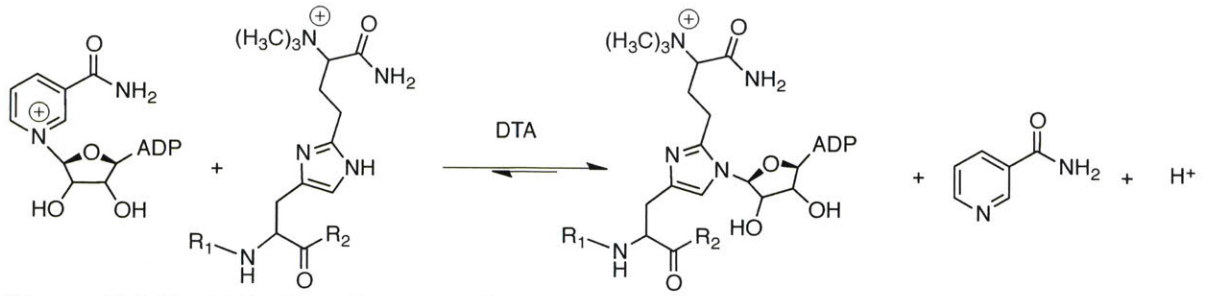


Figure 5.2.2. ADP-ribosylation of diphthamide in EF-2 catalyzed by DTA.

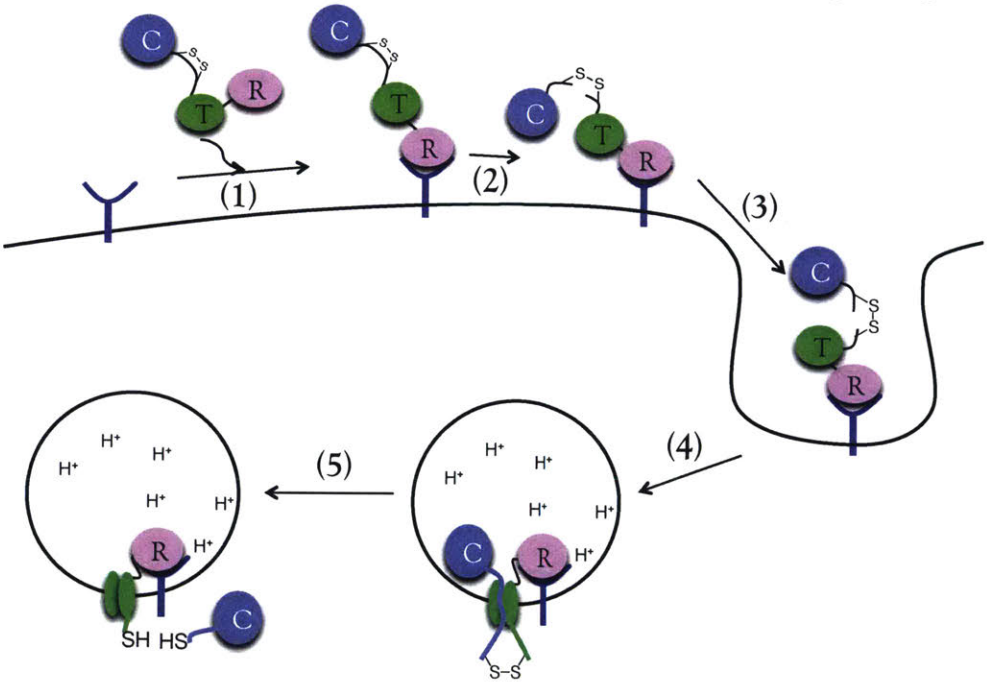


Figure 5.2.3. Molecular mechanism of DTA entry into cells. **(1)** Binding to receptors (hb-EGF) on target cells. **(2)** Furin cleavage at R193. **(3)** Clathrin-mediated endocytosis. **(4)** Acidification of endosome and insertion of DT T domain into membrane. In the process, the N-terminus of DT T domain has been observed to translocate across membrane, possibly bringing the C-terminus of DTA covalently attached by disulfide bond. **(5)** Translocation of catalytic domain across endosome membrane into cytosol. The translocation could occur by DTA passing through the pore formed T domain from C terminus to N terminus, or by chaperone activity of T domain. The process has been observed to be mediated by T domain alone and reported to require cytosolic factors. Once DTA accesses the cytosol, the disulfide bond between DTA and DTB is reduced by cytosolic glutathione and DTA is released from the endosome. C: catalytic domain (DTA). T: translocation domain. R: receptor-binding domain.

DTA needs to access the cytosol for ADP-ribosylation of EF-2, causing inhibition of protein synthesis and eventually killing the cell. DTA enters into eukaryotic cells with assistance of receptor-binding domain (R domain) and translocation domain (T domain) (Figure 5.2.3).

First, the receptor-binding domain binds to the receptor on target cells. In 1978, Leppla's group measured the number of receptors on the very sensitive vero cells by measuring the amount of radiolabelled diphtheria toxin bound to cells.(Middlebrook et al., 1978) They determined that there were about 200,000 binding sites per cell. They also showed that the more radiolabelled DT bound to the cells, the more sensitive the cell-line is to DT cytotoxicity. The resistance comes from defect in toxin entry, and is attributed to defect in receptor binding. The receptor was finally identified in 1992 by cloning a monkey cDNA library into non-sensitive mouse L-M cells and looking for the gene that conferred sensitivity.(Naglich et al., 1992) The receptor is a membrane-bound precursor of heparin-binding EGF-like growth factor (hb-EGF). This precursor can be processed into mature heparin-binding EGF-like growth factor, which is a mitogen that encourages cell division. The binding affinity between DT and the receptor has been reported to be 10^{-8}-10^{-9} M with a stoichiometry of 1:1.(Mitamura et al., 1995)[16] Crystal structure of DT bound to it's the extracellular fragment of its receptor has been reported, which shows mostly hydrophobic interactions between the two, with a burial area of ~1100 $Å^2$ (Figure 5.2.4).(Louie et al., 1997) It was later found that rodent hb-EGF, though 80% similar to human and monkey homolog, does not bind to DT. The >16,000-fold difference in toxin sensitivity has been attributed to differences in 3 amino acids.(Cha et al., 1998)

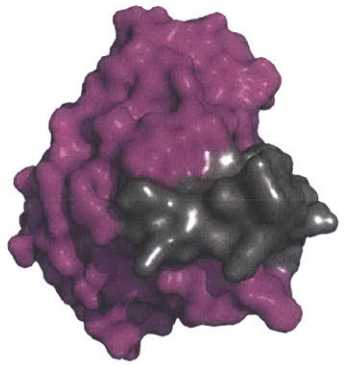


Figure 5.2.4. Crystal structure of DT R domain (pink) bound to a fragment of its receptor hb-EGF (gray) (pdb 1xdt).

DT needs to be enzymatically cleaved between DTA and DTB to allow release of DTA into the cytosol after translocation. There is a flexible loop 186-201 that does not show up in crystal structures. A RVRR (190-193) furin site is contained within this loop, which can be nicked by furin on cell surface, or by trypsin (Figure 5.2.1a). This nicking breaks the amide bond between DTA and DT-B, leaving these two fragments bonded together only by a disulfide bond between C186 and C201. This disulfide bond can be reduced by cellular glutathione once DTA is translocated across phospholipid membrane of endosome and allows for release of DTA into the cytosol. It should be noted that nicking could happen before, during and after endocytosis.

DT enters the cell by clathrin-mediated endocytosis after receptor binding. Electron microscope studies with DT-gold clearly show gold particles in clathrin-coated endocytic vesicles and their movements into cells.(Morris et al., 1985)

Early experiment suggest lysosomal involvement because DT toxicity can be inhibited by addition of NH_4Cl, which inhibits lysosome acidification.(Draper and Simon, 1980) Later Olsnes' group found that after binding of DT to cells, brief exposure of cells to pH 4.5-5.5 media allows for cell-surface translocation of DTA into the cytosol.(Moskaug et al., 1988; Sandvig and Olsnes, 1980) DT has been found to be able to insert into membrane at acidic pHs and form an ion-conducting pore.(Donovan et al., 1981) The pore-forming activity has been attributed to T domain. Among the 9 helices in T domain, TH8 and TH9 have been identified by deletion screening to be the minimal construct for pore formation (Figure 5.2.5).(Silverman et al., 1994) When fluorescent and EPR-sensitive labels are incorporated into this region, spectrometric changes indicated that they were buried into membrane.(Ladokhin et al., 2004; Zhan et al., 1995) Olsnes' group, through cell surface translocation and formation of DT pore in radioactive $^{22}Na^+$ labeled media, showed that monovalent cations, such as Na^+ and K^+, passes through the pore readily.(Sandvig and Olsnes, 1988) Another bilayer study shows that PEG1500 can pass through the pore, suggesting that the pore size is 18-20 Å.(Kagan et al., 1981)

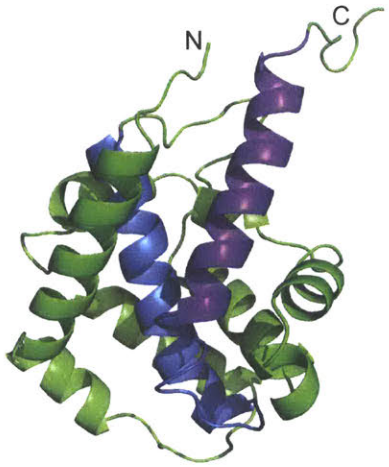


Figure 5.2.5. Crystal structure of DT T domain showing TH8 (blue) and TH9 (purple) (pdb 1ddt).

Though the domain responsible for translocating DTA across the membrane is known, there is still much debate about how the translocation happens.(Oh et al., 1999) It remains unclear whether the T domain alone is enough to translocate the catalytic domain across the membrane (Oh et al., 1999), or other host molecules, such as Hsp90 and COPI complex, are required.(Ratts et al., 2003) Additionally, though the cation-selective pore is observed after the catalytic domain has been translocated to the cytosol, there has not been any account showing DTA blocks the channel formed by DT-T. It could be conceived as the formation of the pore is concerted with initiation of translocation, therefore no pore is observed until the translocation is complete, or that the pore is formed after translocation but is not directly involved in the process of translocation. In the case that DTA passes through a pore, it appears that translocation is initiated from the C-terminus of DTA, since Finkelstein's group observed that the N-terminus of DTB is translocated across the membrane in bilayer studies and the translocation of DTA possibly initiates from the loop region connecting DTA and T domain.(Finkelstein et al., 2000) In terms of DTA's folding state during translocation, London's group have shown that T domain can interact with molten-globule-like proteins when it is inserted into membrane, suggesting that it could have some chaperone activity to help DTA unfold into a molten-globule like state during translocation.(Ren et al., 1999)

Olsnes and Stenmark's group have shown that cargos on the N-terminus of diphtheria toxin can be translocated into the cytosol. The cargos they tested includes various peptides of 12-30 amino acids and proteins such as an extra DTA domain or a 124 amino acid fragment apolipoprotein that forms amphipathic helices.(Madshus et al., 1992; Stenmark et al., 1991) However, they had not been able to establish a structure-function relationship on the cargo requirements for translocation. Another report by *Wu et al.* claims that, with a fusion between myoglobin and DT T domain, they see myoglobin translocation through pore

formed by DT T domain in bilayer studies.(Wu et al., 2006) They claim that the myoglobin is folded during translocation because it has been incubated with hemin that stabilizes myoglobin fold. I personally do not think the translocation of hemin-bound myoglobin is enough evidence to claim myoglobin was folded when it was translocated. Most people in the field think there is at least some degree of unfolding of the cargo during translocation. Diphtheria toxin was one of the first bacteria toxins used to develop immunotoxins. In 1980s, many scientists started developing immunotoxins based on naturally occurring bacterial toxins to specifically kill target cells. When the receptor domain of DT is either mutated to abolish its normal function or completely removed, and a desired receptor-binding domain is appended, the receptor specificity of DT can be changed and the new molecule can be used to specifically kill cells bearing the corresponding receptor. ONTAK (DAB389IL-2) developed by Murphy is one of those immunotoxins that has been approved by FDA to treat CD25 positive refractory cutaneous T cell lymphoma. Through these immunotoxins, we can also learn some properties of DT. Though it appears that the catalytic, translocation and receptor-binding domains are independent from each other in crystal structures, it has been found that immunotoxins with R domain deletion is about 100-fold less active than the ones made with the whole DT, suggesting that the R-domain does facilitate translocation in addition to binding to DT receptors.(Greenfield et al., 1987) Additionally, they found that certain mutations in R-domain reduce receptor binding but not translocation efficiency, indicating that the receptor-binding and translocation function of R-domain can be separated.

To study whether DT can be systematically used to deliver macromolecule cargos, specifically proteins of interest across cell membranes into cytosol, I have first developed a sortagging strategy where cargos were expressed with a C-terminal LPSTGG motif to be sortagged onto the N-terminus of G_5-DT. I have tested that the cargo-DTA(E148S) were successfully translocate into the cytosol through protein synthesis inhibition assays and western blots. However, the translocation efficiency is significantly reduced upon protein cargo attachment. DT is thus not an appropriate tool to deliver significant amounts of protein cargos into the cytosol.

5.2.2 Results

Expression of cargo proteins

To conjugate cargo proteins onto DT and study the effect of cargos on intoxication of DT to vero cells, I expressed several cargo molecules with LPSTGG motif at their C-termini. The protein cargos I have chosen are shown in Figure 5.2.6. Among them, affibody (AB), immunoglobulin G-binding domain 1 (pGB1), and tenth fibronectin type 3 domain (FN3) are small protein scaffolds that can be engineered to bind to protein targets and can interfere with protein-protein binding. LF_N is the N-terminal, protective antigen binding domain of anthrax toxin lethal factor that does not have any enzymatic activity of lethal factor. Previously Madshus *et al.* have conjugated wild-type DTA to E148S DT and showed that they can get similar activity as wild-type DT, and therefore concluded that DT can carry an extra DTA as a cargo protein during translocation.(Madshus et al., 1992) I repeated this experiment to show that fusion proteins conjugated by sortagging behaves

similarly to recombinantly expressed ones in DT cell assay. All cargo proteins are expressed as SUMO fusions and purified with more than 20 mg yield per liter LB.

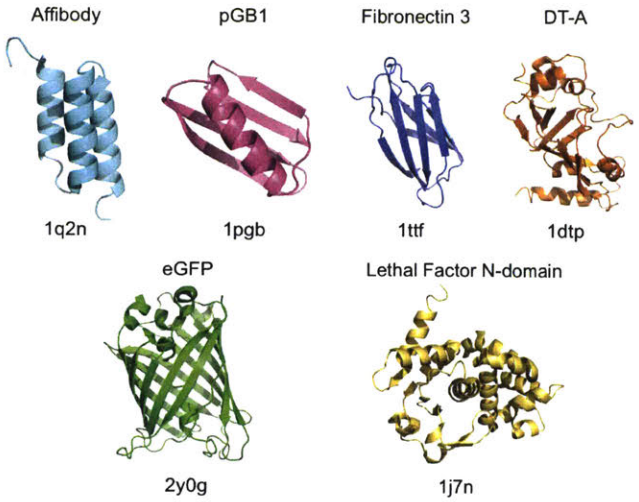


Figure 5.2.6. Crystal structures of cargo proteins conjugated to DT(E148S).

Synthesis of analogs

To study the effect of cargo molecules on intoxication of vero cells by diphtheria toxin, I synthesized DT (E148S) variants with different cargo proteins through sortagging reaction (Figure 5.2.7a). Sortagging reaction was chosen because it allows us to modularly change the cargo that is appended to diphtheria toxin, and is a reaction that can be carried out in native condition without using reducing agents.

The sortagging reaction is catalyzed by sortase A, an enzyme made by *Staphylococcus aureus* that is responsible for ligating surface proteins to the cell wall. Sortase A recognizes an LPXTG motif and cleaves between the threonine-glycine bond, forming a thioester intermediate on C184 in the active site. The thioester intermediate is primed to react with the N-terminus of an oligoglycine moiety and forms back an amide bond, thus covalently linking the two reactants. In 2011, Chen *et al.* reported an evolved sortase A (P94S/D160N/K196T, SrtA*) with improved enzymatic kinetics.(Chen et al., 2011) All my experiments are conducted with SrtA*.

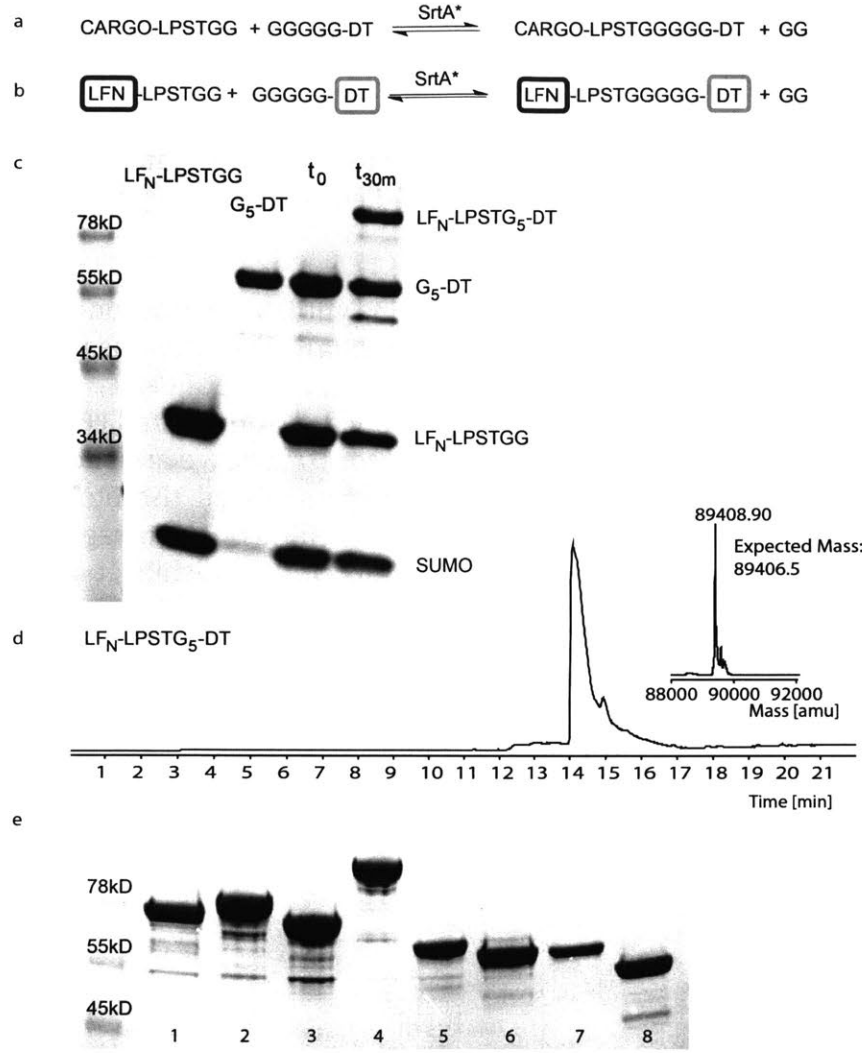


Figure 5.2.7. Synthesis of cargo-DT(E148S) conjugates through sortagging. **(a)** Synthetic scheme for cargo-DT. **(b)** Synthetic strategy for LF_N-DT. **(c)** SDS-PAGE analysis for sortagging reaction between LF_N-LPSTGG and G_5-DT. **(d)** LC-MS analysis of purified LF_N-$LPSTG_5$-DT. **(e)** SDS-PAGE analysis of purified cargo-DTs.
1: eGFP-DT
2: LF_N-DT
3: DTA(wt)-DT
4: LF_N-DTA(wt)-DT
5: FN3-DT
6: pGB1-DT
7: AB-DT
8: G_5-DT

This reaction has been previously used in our lab to conjugate cargo molecules to the C-terminus of anthrax toxin lethal factor N-domain (LF_N) and study the translocation through protective antigen pore. The sortagging protocols have been re-optimized for sortagging on the N-terminus of G_5-DT.

For most proteins, reaction proceeds to about 50% yield within 30 minutes with 75 or 100 μM G_5-DT and 150 or 200μM cargo-LPSTGG in the presence of 5 μM SrtA*, with the exception of SUMO-pGB1-LPSTGG, which needs 100 μM SrtA* for effective sortagging within an hour. The conjugate product can be separated from the reactants on size exclusion column (SEC). For small cargo proteins, SUMO fusion (13kD) is used to help with separation on SEC, and the SUMO fusion is cleaved after separation. I am able to obtain pure cargo-DT conjugates with 5-15% isolated yield. A sample reaction between LF_N-LPSTGG and G_5-DT is shown (5.2.7b-d). Other cargo-DT constructs can be prepared to a similar purity (Figure 5.2.7e).

Vero Cell Protein Synthesis Inhibition Assay

The intoxication effect of DT is studied by incubating vero cells with serial dilutions of cargo-DT conjugates and measuring the extent of ^{3}H-Leu incorporation after treatment. When a critical amount of cargo-DTAs are translocated into cells to ADP-ribosylate all the EF-2s in the cells, protein synthesis will be stopped and little ^{3}H-Leu will be incorporated into cells.

For AB-DT, LF_N-DT, DTA-DT and LF_N-DTA-DT, significant protein synthesis inhibition is observed, where the observed EC_{50s} are all below 1 nM (Figure 5.2.8a); while there is no significant protein synthesis inhibition observed for pGB1-DT, FN3-DT and eGFP-DT, where EC_{50s} appears to be greater than 100 nM, the maximum concentration tested (Figure 5.2.8b). It is important to note that for LF_N-DTA-DT and DTA-DT constructs, we attached a wild-type DTA that is about 800-fold more active than DTA(E148S) expressed, therefore the apparent EC_{50} is decreased. The DTA(wt)-DT(E148S) I synthesized through sortagging shows similar protein inhibition curve to that reported by Madshus *et al.*(Madshus et al., 1992)

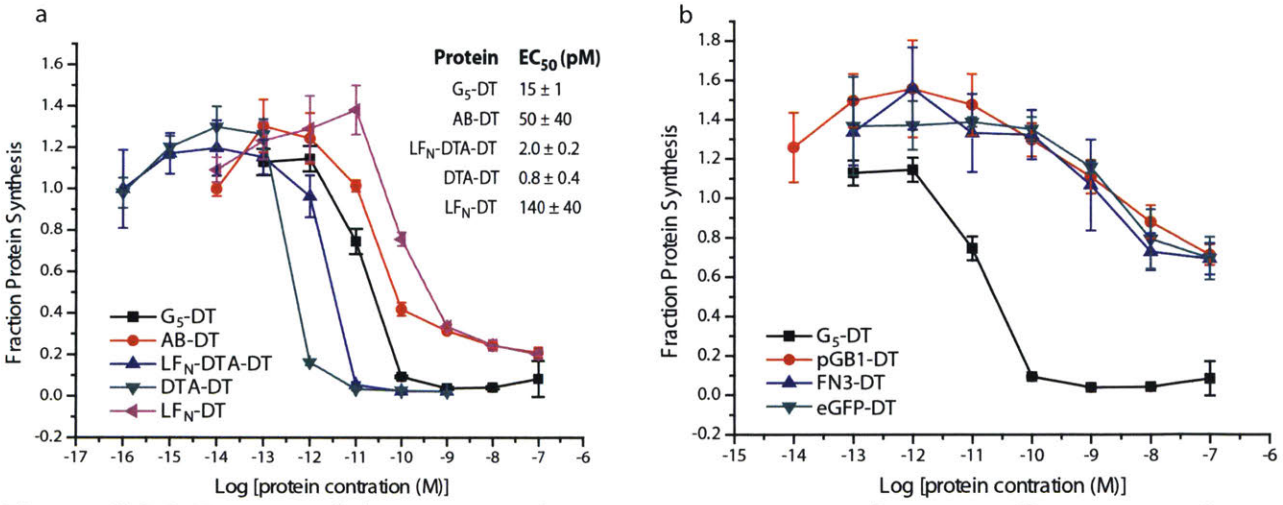


Figure 5.2.8. Protein inhibition assay for cargo-DTs in vero cells. Vero cells are treated with various concentrations of cargo-DT for 24 hours and then incubated with ^{3}H-Leu media for 1 hour. The amount of radioactivity the cells incorporated are measured and used to quantify amount of protein synthesis. **(a)** Protein synthesis inhibition curve for cargo-DT variants with significant intoxication of vero cells. **(b)** Protein synthesis inhibition curve for cargo-DT variants that does not show significant intoxication with 100 nM variants. Each data point represents the average of triplicated experiments at the specific concentration. Data is normalized to cells that are not treated with DT.

From this cell assay, it can be observed that the cargo proteins that translocate effectively are rich in α-helices, while the ones that do not show significant translocation all have significant β-sheet structures. However, the difference in translocation efficiency can also be attributed to the differences in the melting temperatures (T_m), as the latter proteins generally have higher T_m (Table 5.2.1).

Table 5.2.1. Observed EC_{50} for cargo-DT in vero cell and T_ms of isolated cargo proteins

Protein	Observed EC_{50}	T_m (°C)
AB	50 ± 40 pM	75 (Grimm et al., 2012)
FN3	> 100 nM	88
pGB1	> 100 nM	85 (Barakat et al., 2007)
eGFP	> 100 nM	>80 (Stepanenko et al., 2004)
LF_N	140 ± 40 pM	Not reported
DTA(wt)	0.8 ± 0.4 pM	44 (Ramsay et al., 1989)
LF_N-DTA(wt)	2.0 ± 0.2 pM	Not reported

Western blots for translocated constructs After vero cells have been treated with 50 nM of cargo-DTs for 24 hours, the cells are collected and lysed with 0.5% digitonin to avoid extracting microsomes. Cell lysates are run on SDS-PAGE, transferred to nitrocellulose membranes and then blotted with antibody against DTA. During translocation, DTA translocates together with the cargo, leaving DTB and untranslocated DT molecules to be degraded by the lysosome. When cargo-DTA is translocated into the cytosol, we expect to see a band at a different molecular weight than G_5-DTA. Indeed, we observed bands at expected molecular weight for DTA(wt)-DT, LF_N-DT and LF_N-DTA(wt)-DT into vero cells (Figure 5.2.9). However, bands for LF_N-DT and LF_N-DTA(wt)-DT are very weak comparing that that of DTA(wt)-DT, indicating that the translocation of DT cargo might be optimal for DTA through natural selection.

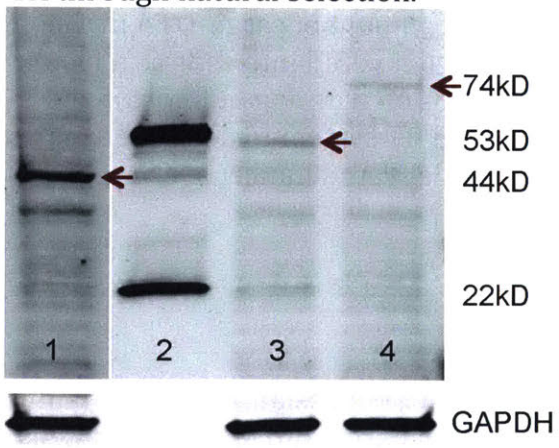


Figure 5.2.9. Western Blot for translocated cargo-DTA(E148S) into vero cells. Vero cells are treated for 24 hours before washed and lysed. The lysates are run on SDS-PAGE and transferred to nitrocellulose membrane. The membrane is blotted with antibody to DTA and then a secondary antibody to image.
1: DTA(wt)-DTA(E148S).
2: 4 ng of LF_N-DTA-ZF(55kD, top) and 4 ng of G_5-DTA(E148S) (22kD, bottom).
3: LF_N-DTA(E148S).
4: LF_N-DTA(wt)-DTA(E148S).

From the first few experiments with

5.1.3. Discussion

I have shown that I can successfully covalently attach protein cargos to the N-terminus of G_5-DT through sortagging. These cargo-DTs can be purified from unreacted G_5-DT using size exclusion chromatography to more than 90% purity. However, not all the cargos can be effectively translocated by DT. DTA(wt) was translocated well into the cytosol, cargos rich in α-helical structure with lower T_m can be translocated less well, while cargos rich in β-sheet structure with very high T_m are translocated very poorly.

It is possible that attachment of non-DTA cargos adversely affects the receptor-binding of DT. Madshus *et al.* previously reported that a number of their DT fusion proteins are unable to bind to receptors on vero cells. Though we attempted to minimize the effect by attaching the cargos to recombinantly expressed, folded DT on the N-terminus, which has significant distance from the receptor binding domain located at the C-termnius both in primary sequence and crystal structure, we could still affected the property of receptor binding domain in solution.

It is also possible that attachment of cargo affect DTA activity. DTA has been reported to be at least partially unfolded during the process of translocation, and the cargo-DTA probably needs to refold after it accesses the cytosol. While DTA has been designed by nature to refold promptly, the cargo appended could impede refolding, resulting in insoluble DTA or misfolded DTA that does not have enzymatic activity. Alternatively, it is possible that the cargo molecule inhibits DTA activity, rendering the fusion protein inactive. Indeed, Auger *et al.* recently reported N-terminal fusions to DTA significantly reduces DTA activity (Auger et al., 2015). They have shown successful delivery of ubiquitin molecules and amylases into the cells and detected intracellular enzymatic activity.

The cargo-DTA translocated could also have a shorter half-life than that of DTA molecules on their own. Due to the long treatment time needed for adequate DTA (E148S) activity, if the cargo-DTA is degraded before they have enough time to ADP-ribosylate all the EF-2s, we would not observe either protein synthesis inhibition or bands in lysate.

In general, while I have not been able to detect a significant amount of cargo-DTA in the cytosol of cells, other labs have shown that enzymes can be delivered into cells as a fusion to DT and be active in the cytosol. Further work should focus on working out how much protein can be delivered using the DT system and to use the DT system to deliver enzymes or protein binders with high affinity so that the amount of delivered molecule are appropriate for the phenotype expected.

5.2.4 References

Auger, A., Park, M., Nitschke, F., Minassian, L.M., Beilhartz, G.L., Minassian, B.A., and Melnyk, R.A. (2015). Efficient Delivery of Structurally Diverse Protein Cargo into Mammalian Cells by a Bacterial Toxin. Molecular Pharmaceutics *12*, 2962-2971.

Barakat, N.H., Barakat, N.H., Carmody, L.J., and Love, J.J. (2007). Exploiting elements of transcriptional machinery to enhance protein stability. Journal of Molecular Biology *366*, 103-116.

Barbieri, J.T., and Collier, R.J. (1987). Expression of a mutant, full-length form of diphtheria toxin in Escherichia coli. Infection and immunity *55*, 1647-1651.

Bennett, M.J., Choe, S., and Eisenberg, D. (1994). Refined structure of dimeric diphtheria-toxin at 2.0-Angstrom resolution. Protein Science *3*, 1444-1463.

Cha, J.H., Brooke, J.S., and Eidels, L. (1998). Toxin binding site of the diphtheria toxin receptor: loss and gain of diphtheria toxin binding of monkey and mouse heparin-binding, epidermal growth factor-like growth factor precursors by reciprocal site-directed mutagenesis. Molecular microbiology *29*, 1275-1284.

Chen, I., Dorr, B.M., and Liu, D.R. (2011). A general strategy for the evolution of bond-forming enzymes using yeast display. Proceedings of the National Academy of Sciences of the United States of America *108*, 11399-11404.

Collier, R.J. (1975). Diphtheria-toxin - Mode of action and structure. Bacteriological Reviews *39*, 54-85.

Donovan, J.J., Simon, M.I., Draper, R.K., and Montal, M. (1981). Diphtheria toxin forms transmembrane channels in planar lipid bilayers. Proceedings of the National Academy of Sciences of the United States of America *78*, 172-176.

Draper, R.K., and Simon, M.I. (1980). The entry of diphtheria-toxin into the mammalian-cell cytoplasm - evidence for lysosomal involvement. Journal of Cell Biology *87*, 849-854.

Finkelstein, A., Oh, K.J., Senzel, L., Gordon, M., Blaustein, R.O., and Collier, R.J. (2000). The diphtheria toxin channel-forming T-domain translocates its own NH2-terminal region and the catalytic domain across planar phospholipid bilayers. International journal of medical microbiology *290*, 435-440.

Greenfield, L., Bjorn, M.J., Horn, G., Fong, D., Buck, G.A., Collier, R.J., and Kaplan, D.A. (1983). Nucleotide-sequence of the structural gene for diphtheria-toxin carried by corynebacteriophage-beta. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences *80*, 6853-6857.

Greenfield, L., Johnson, V.G., and Youle, R.J. (1987). Mutations in diphtheria toxin separate binding from entry and amplify immunotoxin selectivity. Science *238*, 536-539.

Grimm, S., Salahshour, S., and Nygren, P.A. (2012). Monitored whole gene in vitro evolution of an anti-hRaf-1 affibody molecule towards increased binding affinity. New Biotechnology *29*, 534-542.

Kagan, B.L., Finkelstein, A., and Colombini, M. (1981). Diphtheria toxin fragment forms large pores in phospholipid bilayer membranes. Proc Natl Acad Sci U S A *78*, 4950-4954.

Ladokhin, A.S., Legmann, R., Collier, R.J., and White, S.H. (2004). Reversible refolding of the diphtheria toxin T-domain on lipid membranes. Biochemistry *43*, 7451-7458.

Louie, G.V., Yang, W., Bowman, M.E., and Choe, S. (1997). Crystal structure of the complex of diphtheria toxin with an extracellular fragment of its receptor. Molecular Cell *1*, 67-78.

Madshus, I.H., Olsnes, S., and Stenmark, H. (1992). Membrane translocation of diphtheria toxin carrying passenger protein domains. Infect Immun *60*, 3296-3302.

Middlebrook, J.L., Dorland, R.B., and Leppla, S.H. (1978). Association of diphtheria-toxin with vero cells - demonstration of a receptor. Journal of Biological Chemistry *253*, 7325-7330.

Mitamura, T., Higashiyama, S., Taniguchi, N., Klagsbrun, M., and Mekada, E. (1995). Diphtheria toxin binds to the epidermal growth factor (EGF)-like domain of human heparin-binding EGF-like growth factor/diphtheria toxin receptor and inhibits specifically its mitogenic activity. J Biol Chem *270*, 1015-1019.

Morris, R.E., Gerstein, A.S., Bonventre, P.F., and Saelinger, C.B. (1985). Receptor-mediated entry of diphtheria-toxin into monkey kidney (vero) cells - electron-microscopic evaluation. Infection and Immunity *50*, 721-727.

Moskaug, J.O., Sandvig, K., and Olsnes, S. (1988). Low pH-induced release of diphtheria-toxin A-fragment in vero cells - biochemical evidence for transfer to the cytosol. Journal of Biological Chemistry *263*, 2518-2525.

Naglich, J.G., Metherall, J.E., Russell, D.W., and Eidels, L. (1992). Expression cloning of a diphtheria toxin receptor: identity with a heparin-binding EGF-like growth factor precursor. Cell *69*, 1051-1061.

Oh, K.J., Senzel, L., Collier, R.J., and Finkelstein, A. (1999). Translocation of the catalytic domain of diphtheria toxin across planar phospholipid bilayers by its own T domain. Proc Natl Acad Sci U S A *96*, 8467-8470.

Pappenheimer, A.M. (1977). Diphtheria toxin. Annual Review of Biochemistry *46*, 69-94.

Ramsay, G., Montgomery, D., Berger, D., and Freire, E. (1989). Energetics of diphtheria-toxin membrane insertion and translocation - calorimetric characterization of the acid pH induced transition. Biochemistry *28*, 529-533.

Ratts, R., Zeng, H., Berg, E.A., Blue, C., McComb, M.E., Costello, C.E., vanderSpek, J.C., and Murphy, J.R. (2003). The cytosolic entry of diphtheria toxin catalytic domain requires a host cell cytosolic translocation factor complex. The Journal of cell biology *160*, 1139-1150.

Ren, J., Kachel, K., Kim, H., Malenbaum, S.E., Collier, R.J., and London, E. (1999). Interaction of diphtheria toxin T domain with molten globule-like proteins and its implications for translocation. Science *284*, 955-957.

Sandvig, K., and Olsnes, S. (1980). Diphtheria-toxin entry into cells is facilitated by low pH. Journal of Cell Biology *87*, 828-832.

Sandvig, K., and Olsnes, S. (1988). Diphtheria toxin-induced channels in Vero cells selective for monovalent cations. The Journal of biological chemistry *263*, 12352-12359.

Sharei, A., Zoldan, J., Adamo, A., Sim, W.Y., Cho, N., Jackson, E., Mao, S., Schneider, S., Han, M.-J., Lytton-Jean, A., *et al.* (2013). A vector-free microfluidic platform for intracellular delivery. Proceedings of the National Academy of Sciences *110*, 2082-2087.

Silverman, J.A., Mindell, J.A., Zhan, H., Finkelstein, A., and Collier, R.J. (1994). Structure-function relationships in diphtheria toxin channels: I. Determining a minimal channel-forming domain. The Journal of membrane biology *137*, 17-28.

Stenmark, H., Moskaug, J.O., Madshus, I.H., Sandvig, K., and Olsnes, S. (1991). Peptides fused to the amino-terminal end of diphtheria toxin are translocated to the cytosol. The Journal of cell biology *113*, 1025-1032.

Stepanenko, O.V., Verkhusha, V.V., Kazakov, V.I., Shavlovsky, M.M., Kuznetsova, I.M., Uversky, V.N., and Turoverov, K.K. (2004). Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, "dimer2", and DsRed1. Biochemistry *43*, 14913-14923.

Wilson, B.A., and Collier, R.J. (1992). Diphtheria-toxin and pseudomonas-aeruginosa exotoxin-A - active-site structure and enzymatic mechanism. Current Topics in Microbiology and Immunology *175*, 27-41.

Wilson, B.A., Reich, K.A., Weinstein, B.R., and Collier, R.J. (1990). Active-site mutations of diphtheria-toxin - effects of replacing glutamic acid-148 with aspartic-acid, glutamine, or serine. Biochemistry *29*, 8643-8651.

Wu, Z.Y., Jakes, K.S., Samelson-Jones, B.S., Lai, B., Zhao, G., London, E., and Finkelstein, A. (2006). Protein translocation by bacterial toxin channels: A comparison of diphtheria toxin and colicin Ia. Biophysical Journal *91*, 3249-3256.

Yamaizumi, M., Mekada, E., Uchida, T., and Okada, Y. (1978). One molecule of diphtheria-toxin fragment A introduced into a cell can kill cell. Cell *15*, 245-250.

Yates, S.P., Jorgensen, R., Andersen, G.R., and Merrill, A.R. (2006). Stealth and mimicry by deadly bacterial toxins. Trends in Biochemical Sciences *31*, 123-133.

Zhan, H., Oh, K.J., Shin, Y.K., Hubbell, W.L., and Collier, R.J. (1995). Interaction of the isolated transmembrane domain of diphtheria toxin with membranes. Biochemistry *34*, 4856-4863.

5.3 Work in Progress: Using CellSqueeze technology (SQZ) to deliver epitope-tagged Ub and tagged-Ub loaded E2s into cells

in collaboration with Dr. Ross Cheloha with equal contribution

5.3.1 Introduction

Ubiquitin (Ub) is a small protein that is appended to cellular protein targets either as a single copy or in the form of Ub chains. Ubiquitination of proteins affects many aspects of protein function. The best-characterized function of ubiquitin conjugation is to tag substrates and direct them to the proteasome for degradation. Ubiquitin is conjugated to proteins through the sequential action of three classes of enzymes known as ubiquitin-activating enzyme or E1s, ubiquitin-conjugating enzyme or E2s, and ubiquitin ligase or E3s, while deubiquitylating enzymes (DUbs) can remove attached ubiquitin tags (Figure 5.3.1). Dysfunction of the ubiquitin conjugation machinery has been linked to a variety of disease states. There is therefore much interest in understanding how E2s and E3s selectively target their substrates (Corn, 2007; Lim and Tan, 2007; Tsuchida et al., 2017; Upadhya and Hegde, 2007). Ub can be covalently linked using different lysine residues, and these different linkages determine how the substrates are processed (Jacobson et al., 2009; Miranda and Sorkin, 2007). The Ub chains with different linkage types can be probed through use of mutant versions of Ub that lack one or more Lys residues used in chain building.

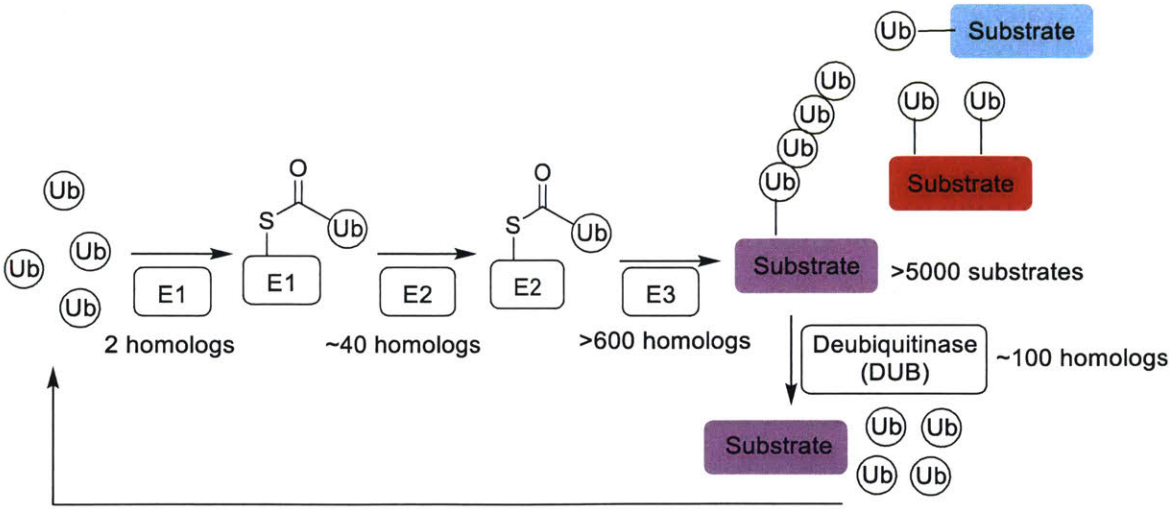


Figure 5.3.1. Scheme of ubiquitin(Ub) pathway in cells. Free ubiquitin is activated by E1 ubiquitin conjugating enzyme at the Ub C-terminus, converting the carboxylic acid into a thioester with the active site Cys on E1. The ubiquitin thioester is then transferred to an E2 conjugating enzyme. Finally, E2 and E3 ligases work together to transfer ubiquitin onto the substrate, to target it for degradation, translocation *etc.* Deubiquitylases remove ubiquitin from the substrates, to recycle Ub and to allow the attenuation of ubiquitin-induced responses.

While there are only 2 known E1s, there are ~40 E2 enzymes, >600 E3 enzymes and ~100 DUbs that regulate the process. Even though most of the ubiquitination pathway can be reconstituted with purified proteins *in vitro*, the reconstituted reactions do not necessarily represent what happens inside a cell and only give limited information on the specificity of the ubiquitination reactions. We hope to better understand the ubiquitination process in the complex environment of the cytosol. How the specificity of the two E1s for downstream E2s is determined remains to be clarified. Specifically, how do the E1s decide which E2 to load during in times of need? In addition, the 40 E2s serve more than 600 E3s. E2s contribute to both substrate specificity and the identity of Ub chains formed (Stewart et al., 2016). We have little knowledge about the factors that contribute to the selection of particular E3s by Ub-loaded E2s. The binding between E2-Ub and E3 is apparently too weak and transient for co-immunoprecipitation to be successful. Overexpression, knockouts and siRNA-mediated knockdowns of selected E2s can give us information on how E2 levels might affect cellular functions. However, there is still much to be learnt about how each step in the pathway is controlled. In addition, E2-Ub conjugates comprise a thioester bond, which cannot be specified by genetic encoding and requires completion of several enzymatic reactions instead. It has not been possible to directly produce an E2-Ub adduct of interest *in vivo* and characterize its substrates. Furthermore, efforts to characterize the use of ubiquitin via inducible genetic expression of epitope-tagged ubiquitin does not allow resolution of ubiquitination events on timescales likely operative *in vivo*, due to the delay inherent in transcription and translation to arrive at the product(s) of interest.

Some of these limitations can be addressed by CellSqueeze, (hereafter: SQZ) technology, which enables transient permeabilization of the cell membrane and concomitant delivery of protein-sized cargo into the cytoplasm following brief mechanical constriction of cells in a microfluidic device (Figure 5.3.2) (Sharei et al., 2013). Cell membrane repair involves the ESCRT III complex and starts within 1.5 minutes post-squeeze, enabling highly efficient delivery and maintenance of cell viability (Ding et al., 2017; Jimenez et al., 2014). SQZ technology allows rapid delivery of chemically modified macromolecules into living cells, thus enabling the study of biological processes at adequate kinetic resolution in the context of an intact cell.

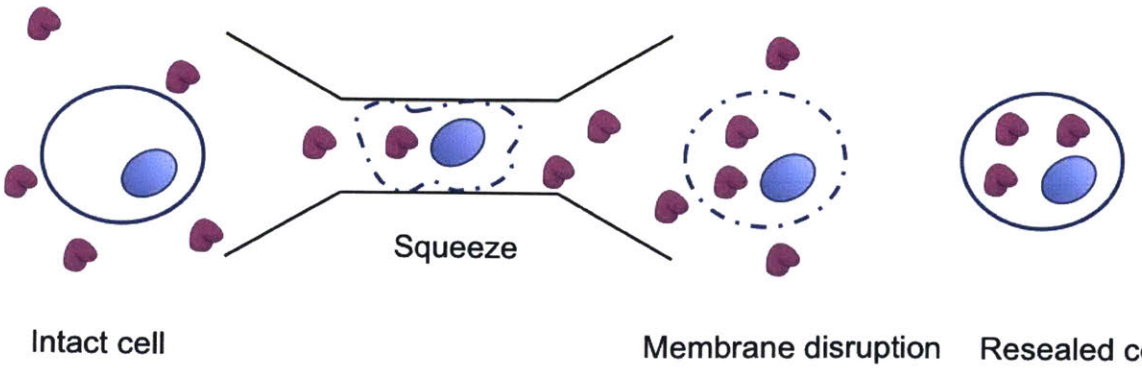


Figure 5.3.2. SQZ technology allows delivery of macromolecules into cells. Intact, living cells are squeezed through microfluidic channels with a diameter smaller than that of the cells. This causes mechanical deformation and transient disruption of the cell's plasma

membrane. Macromolecules can thus diffuse from the environment into the cells prior to membrane repair..

We have begun to use the SQZ platform to deliver HA epitope-tagged ubiquitin (HA-Ub) and recombinantly expressed E2s loaded with HA-Ub into the cytosol of mammalian cells. By doing so, we hope to better understand the kinetics and specificity of HA-Ub utilization when it is delivered into the ubiquitin pool or when it is delivered as a thioester adduct with a particular E2.

5.3.2 Results

HA-Ub can be delivered to squeezed cells and incorporated into the ubiquitin pathway

To assess whether we could deliver proteins into mammalian cells, we squeezed recombinantly expressed HA-Ub into HeLa cells. If HA-Ub is delivered into the cytosol, it should be utilized by the endogenous ubiquitination pathway and incorporated covalently into protein substrates. We optimized the squeeze conditions for HeLa cells and found that when cells are forced through the microfluidic device at 80 psi N_2 pressure using a 10/7 microfluidic chip (channels are 10 μm long by 7 μm wide and 20 μm deep), more than 80% of HeLa cells survive, with more than 80% live cells labeled with the fluorescent dextran in solution (Data not shown). Cells were suspended in Opti-MEM media containing 10 μM HA-Ub and squeezed under the conditions mentioned above. We tracked the formation of substrates labeled with HA-Ub over time in order to explore the kinetics of ubiquitin utilization in the cell.

We observed strong signals corresponding to HA-Ub-conjugated proteins at molecular weights above 37 kD in immunoblotting using an anti-HA antbody (Figure 5.3.2). Since we included 100 mM DTT in the lysis mixture to reduce and therefore eliminated thioesters, such as E1-Ub thioesters as well as E2-Ub thioesters, these observed high molecular weight bands must be HA-Ub-tagged substrates. The high molecular weight HA-containing bands are strongest at 5 and 15 minutes after squeezing, and slowly decrease in intensity thereafter. This result shows that HA-Ub is quickly used by the ubiquitin pathway upon delivery via cell squeezing. HA-Ub can be transferred from E1 to substrates in as short as 5 minutes.

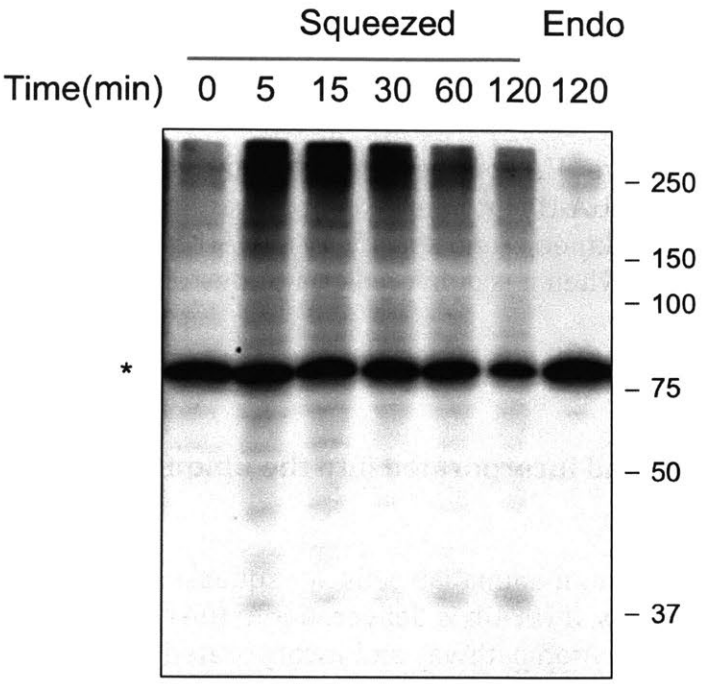


Figure 5.3.3. Recombinantly expressed HA-Ub can be delivered to squeezed cells and incorporated into substrates. Numbers represents minutes post-squeeze. 5 x 10^5 cells in 100 µL were sequeezed in each experiment, transferred into DMEM containing 10% IFS after 90s, and incubated at 37 °C until they were pelleted, washed, and lysed in 2x SDS sample buffer with 100 mM DTT. Endo: cells are exposed to 10 µM HA-Ub for the indicated time but not squeezed. Time 0: Cells were exposed to 10 µM HA-Ub and collected at the beginning of the experiment. *:Irrelevant polypeptide recognized by the HA antibody.

30% of 10 µM HA-Ub is delivered into HeLa cells during squeeze

We sought to quantify the efficiency of HA-Ub delivery by SQZ, and to determine the corresponding cytoplasmic HA-Ub concentration in HeLa cells. To achieve this, we needed to prevent HA-Ub from being used once introduced into the cytoplasm of the cell. We treated HeLa cells with an E1 inhibitor, MLN7243, prior to squeezing to prevent HA-Ub from entering the ubiquitin pathway (Misra et al., 2017). It is noteworthy that high molecular weight HA-containing adducts are not formed following pretreatment with E1 inhibitor, indicating that the adducts formed following introduction of HA-Ub in untreated cells result from Ub attachment to cellular substrates by the endogenous conjugation machinery (Figure 5.3.2). Based on the signal from dilution of a purified HA-Ub solution used as a standard, we estimated the amount of HA-Ub in the cells to be about 4 ± 1 ng HA-Ub in 10^5 squeezed HeLa cells (Figure 5.3.3). Assuming that 1 million HeLa cells take up 1 µL (Fujioka et al., 2006) and that all HA-Ub recovered was delivered to the cytosol, we estimated the cytosolic concentration of HA-Ub to be 3.2 ± 0.8 µM following squeezing in the presence of 10 µM HA-Ub. This number represents a delivery efficiency of 30% for HA-Ub, which is comparable to the efficiency observed using other kinds of cargo (Kollmannsperger et al., 2016; Sharei et al., 2013). Each HeLa cell contains about 1.5 x 10^8 ubiquitin molecules, which corresponds to approximately 250 µM in molar concentration (Carlson and Rechsteiner, 1987). The amount of HA-Ub introduced is much less than that

naturally found in the cell, and we therefore do not anticipate major disruptions in the ubiquitin pathway due to the transient spike in ubiquitin concentration.

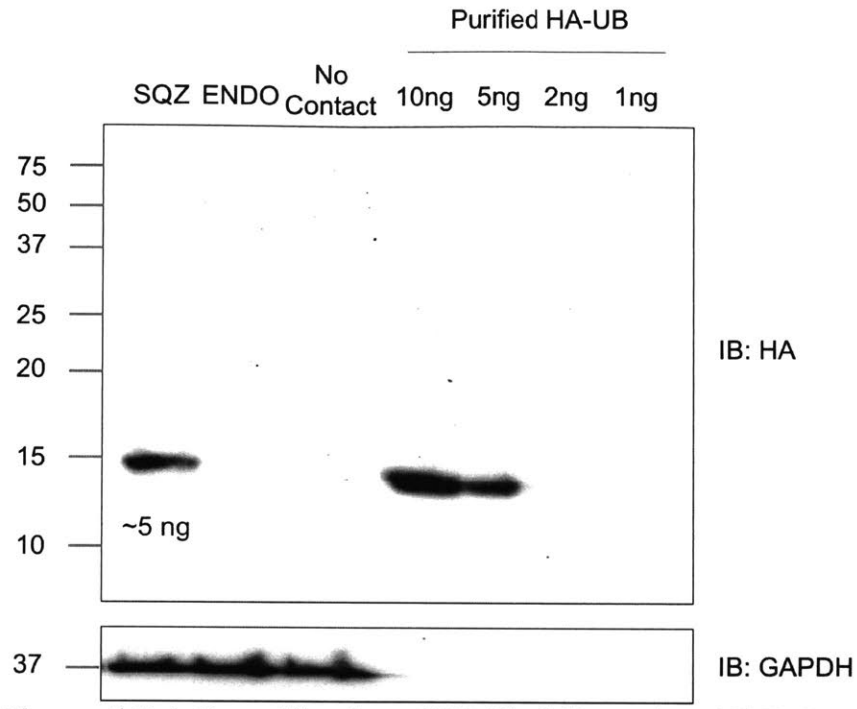


Figure 5.3.4. Quantification of HA-Ub delivered to 10^5 HeLa cells. HeLa cells were treated with 20 μM E1 inhibitor MLN7243 for 1 hour before being trypsinized for squeezing. 10^6 cells were squeezed in 100 μl OptiMEM containing 10 μM HA-Ub. Cells were washed 3 times with PBS and lysed in 2x SDS sample buffer with 100 mM DTT. 10% of lysates were loaded for analysis.

Ub utilization efficiency and kinetics differ for cell lysates and squeezed cells

We compared the kinetics of HA-Ub utilization following addition of HA-Ub into HeLa cell lysates or in squeezed HeLa cells. We prepared cell lysates by sonicating 5 x 10^6 HeLa cells in PBS with Ca^{2+}/Mg^{2+} on ice and collecting the soluble fraction. We added 10 μM of HA-Ub into the lysates and incubated the mixture at 37 °C. The samples were quenched using 2x SDS sample buffer with 100 mM DTT for analysis, as performed for squeezed cells. For comparison, an equivalent number of HeLa cells was used per sample for lysates or squeezed cells.

We compared the intensity and pattern of HA-tagged protein signals for the two experiments and found the outcome to differ substantially when HA-Ub was added to lysates instead of delivered into HeLa cells by SQZ (Figure 5.3.4). The signals for high MW HA-Ub conjugates are much stronger in lysate samples than in squeezed cell samples. This observation can possibly be explained in part by the lower concentration of HA-Ub present in HeLa cells during SQZ (Figure 5.3.3). However, the lysate volume is >200x of that of the volume occupied by cells, and therefore all cellular components are substantially more dilute in lysates than in living cells, which we expect should delay incorporation of HA-Ub.

We therefore conclude that disruption of HeLa cell integrity in the course of lysate preparation leads to loss of control in the ubiquitination pathway, causing higher levels of ubiquitination than in intact cells.

The more striking difference is seen in the kinetics of HA-Ub incorporation and removal. At 5 minutes, there are already significant amounts of HA-Ub-labeled substrates in squeezed cells, equivalent or greater to levels seen at that timepoint in lysates. These HA-Ub conjugates are quickly resolved, either by deubiquitylating enzymes (DUbs) or by components of the proteasome or autophagy system. On the other hand, in lysates, HA-Ub-labeled high MW conjugates accumulate over time without obvious signs of degradation.

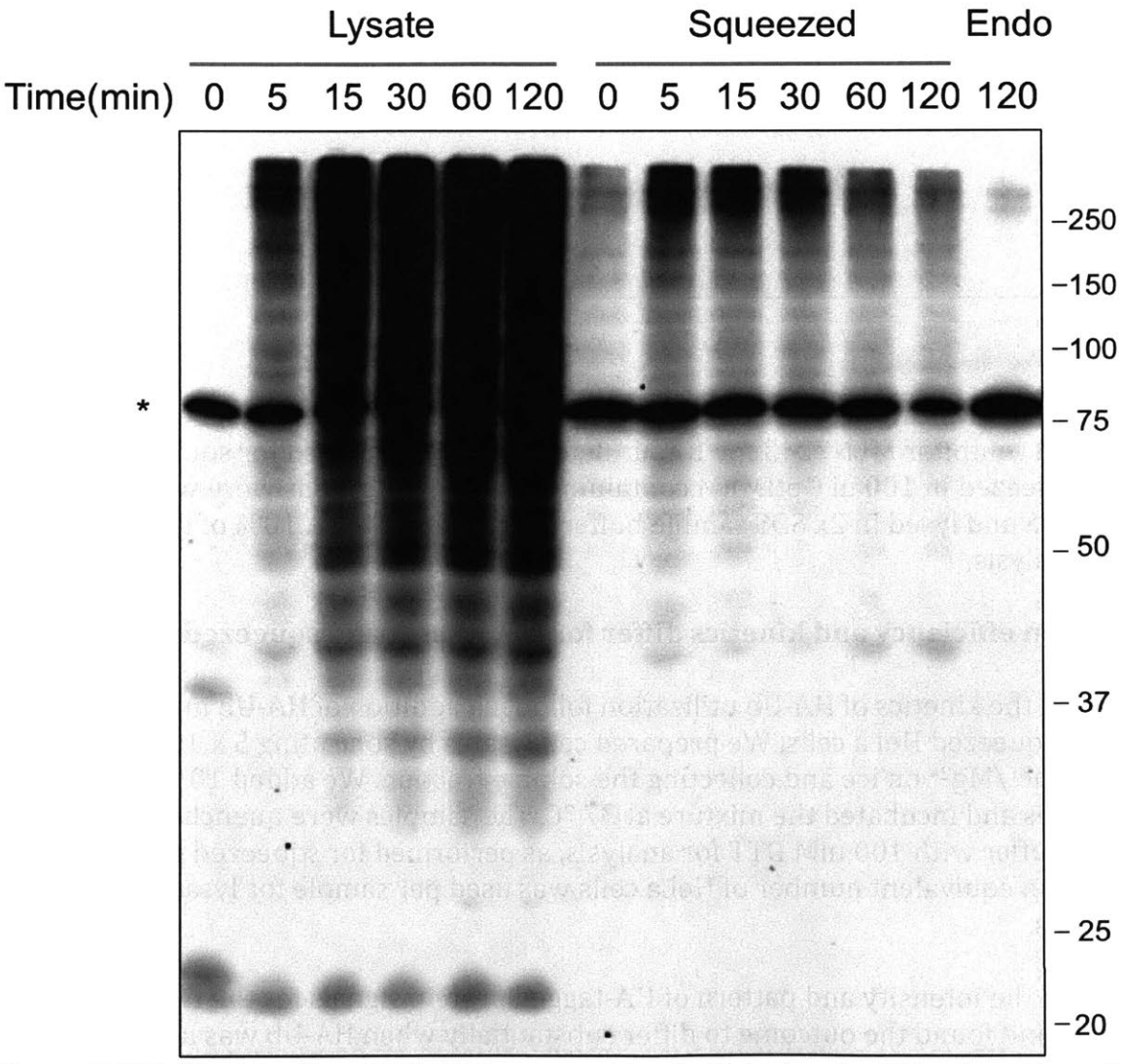


Figure 5.3.5. The kinetics of HA-Ub incorporation in HeLa lysate differ from those in squeezed cells. Identical numbers and concentrations of HeLa cells were loaded for lysates and for squeezed cells. Lysates were prepared by sonicating in PBS with Ca^{2+}/Mg^{2+} followed by centrifugation to remove insoluble material and subsequent addition of HA-Ub. SQZ experiments were performed in HA-Ub containing OptiMEM. For 0 minute timepoint in

the squeeze experiments, HeLa cells were exposed to HA-Ub, but not squeezed and collected at the beginning of experiments. *: Irrelevant polypeptide recognized by the HA antibody

Proteasome inhibition and autophagy inhibition delay degradation of HA-Ub labeled adducts

We treated HeLa cells with either proteasome inhibitor or autophagy inhibitor prior to cell squeezing to explore the reasons for decay of HA-Ub tagged material. Proteasomal degradation and autophagy are two major pathways via which ubiquitinylated substrates are degraded. We observed stronger high MW HA signals for both proteasome inhibitor-treated cells and autophagy inhibitor-treated cells (Figure 5.3.5). However, neither inhibitor was able to completely abolish the disappearance of high molecular weight adducts. This is not surprising, as both mechanisms are responsible for clearing proteins tagged for degradation, and shutting down only one branch will not result in a complete shutdown of protein degradation. The autophagy inhibitor appears to have a stronger effect, a finding that may be related to a connection to the stress experienced during squeezing/membrane repair and possible induction of autophagy. In contrast, HeLa lysates were centrifuged to collect only the supernatant, removing membrane-bound organelles and therefore minimizing the impact of autophagy on Ub consumption and removal in lysate experiments.

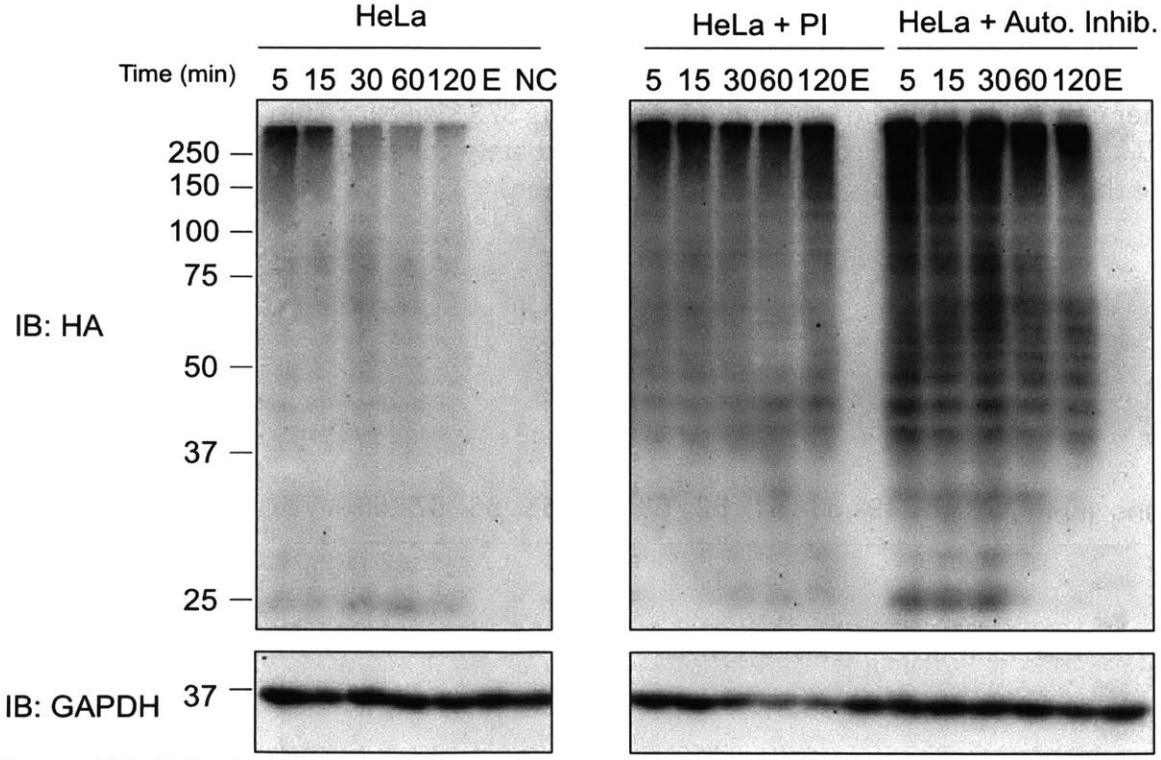


Figure 5.3.6. Both the proteasome inhibitor and the autophagy inhibitor delayed the disappearance of high molecular weight HA-Ub adducts. HeLa cells were treated with proteasome inhibitor (PI) bortezomib (25 μM) or autophagy inhibitor wortmannin (5 μM) for 3 hours prior to trypsinization and squeezing in OptiMEM medium containing 10 μM

HA-Ub. E: endocytosis control. NC: no contact control. All samples were analyzed on the same blot.

HA-Ub E2-Ub thioesters can be delivered to HeLa cells

After squeezing HA-Ub into HeLa cells, we investigated the use of a closely related group of molecules in living cells: Ub-loaded E2 thioesters. E2 enzymes are important participants in the ubiquitination pathway. These enzymes are not only responsible for delivering Ub to E3s, but are also involved in determining substrate specificity as well as the type of ubiquitin linkages installed (Stewart et al., 2016). However, E2s are active only after being loaded with Ub, and the E2-Ub thioester cannot be genetically encoded, as it is produced post-translationally by action of the E1 enzyme. We therefore sought to study how E2 identity contributes to substrate specificity in living cells. We chose two E2s for initial studies: Ubiquitin conjugating enzyme E2 L3 (UBE2L3) and Ubiquitin conjugating enzyme UBC7. UBE2L3 is an E2 linked to many autoimmune diseases and functions with the HECT and RING-in-between-RING (RBR) classes of E3 ligases(Eldridge et al.; Fu et al., 2013; Lewis et al., 2015). UBC7 is an E2 that is recruited to the ER membrane by interaction with an ER-membrane localized anchor protein, CUE1, and is reported to work with HRD1 RING type E3 and gp78 involved in ER associated degradation (Bazirgan and Hampton, 2008; Cohen et al., 2015; Kikkert et al., 2004).

Both UBE2L3 and UBC7 can be expressed in active form in *E. coli*, purified using standard Ni purification protocols, and can be loaded with Ub *in vitro*. Loading is performed with HA-Ub *in vitro* using recombinantly expressed E1. UBE2L3-Ub and UBC7-Ub were purified from other components in the reaction mixture using S75 size exclusion chromatography (Figure 5.3.6a). The purified constructs were immediately concentrated and flash frozen to minimize thioester hydrolysis and auto-ubiquitination.

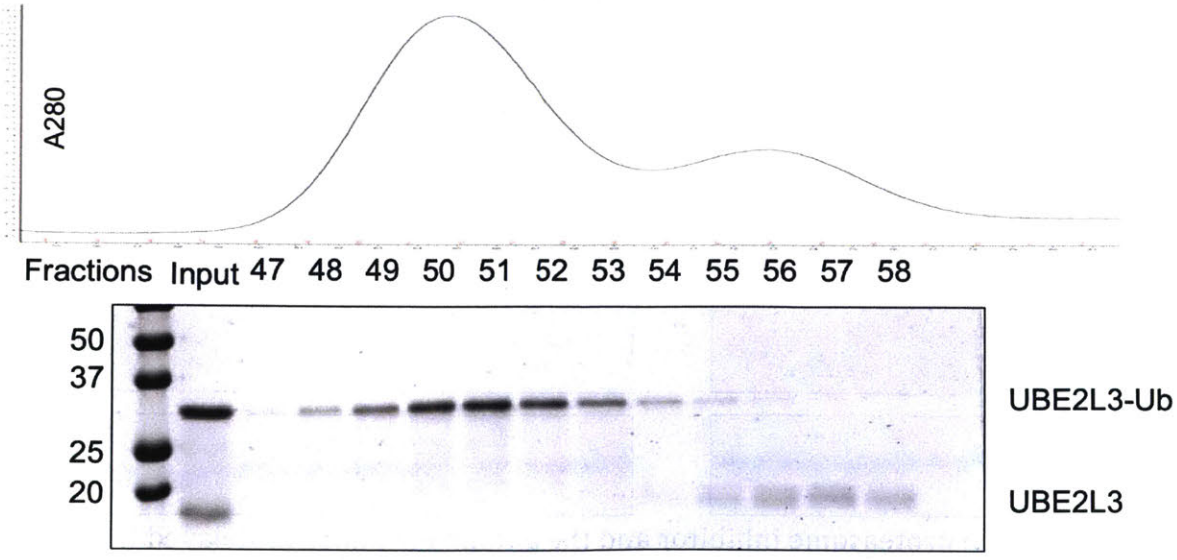


Figure 5.3.7a. Purification of UBE2L3-Ub by S75 16/600 size exclusion column. 45 μM UBE2L3 was loaded with 50 μM HA-Ub and 200 nM mE1 in 100 mM Tris pH7.4, 2 mM ATP,

5 mM $MgCl_2$ and 200 μM DTT at 37°C for 30 minutes. Flow rate: 1 ml/min. Fraction size: 1.3 ml.

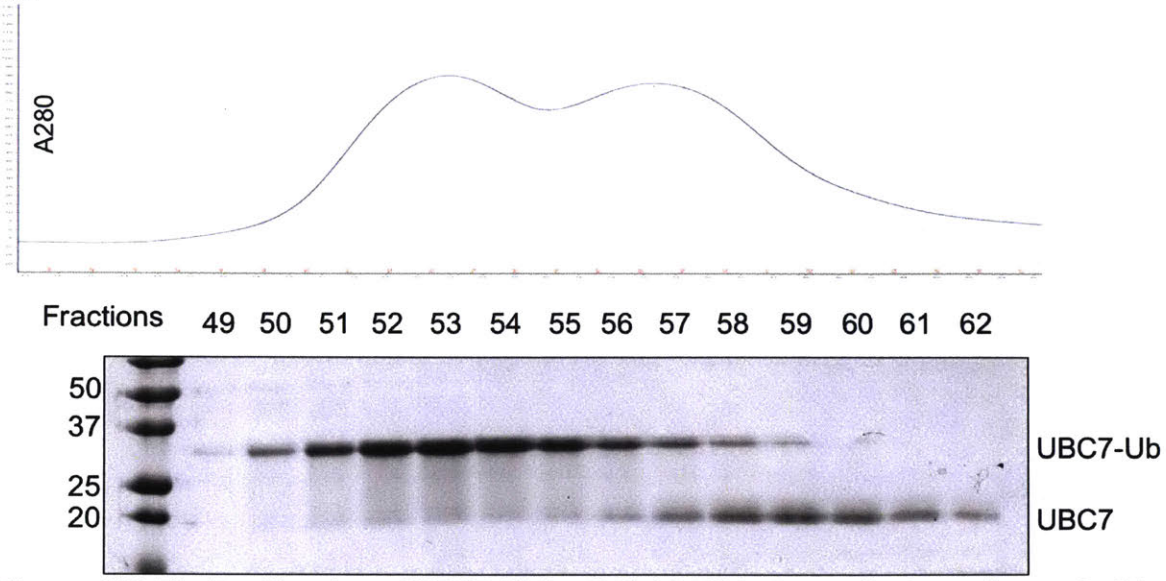
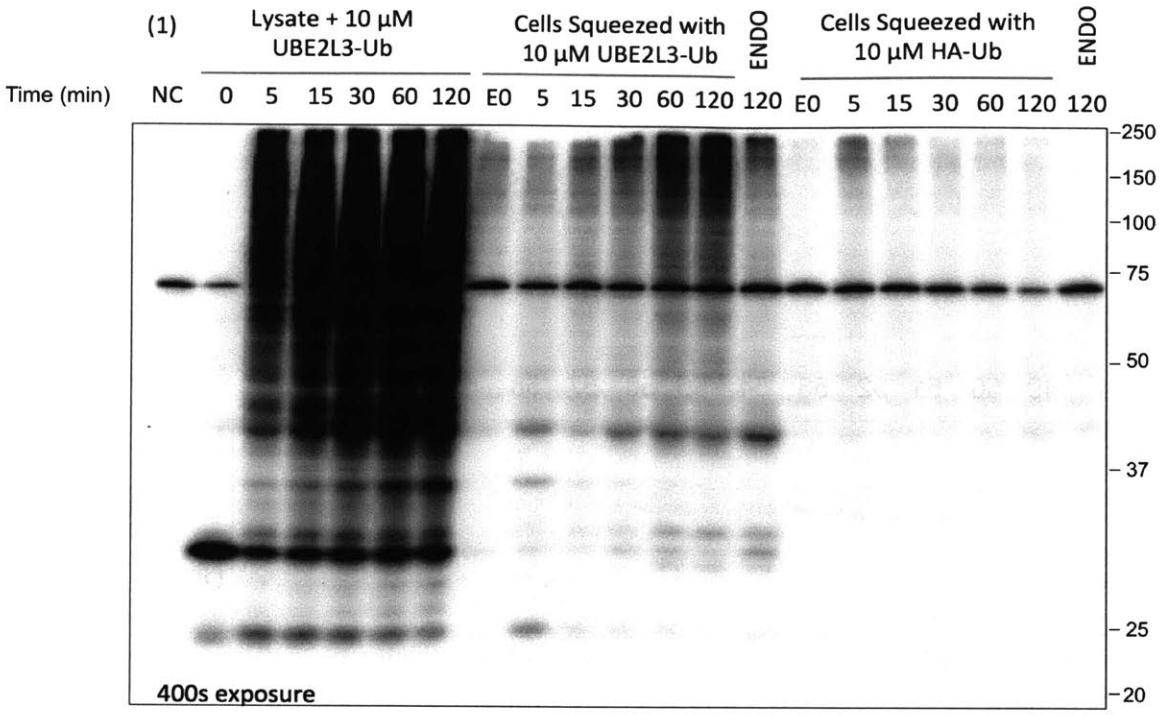


Figure 5.3.7b. Purification of UBC7-Ub by S75 16/600 size exclusion column. 8 μM UBE2L3 was loaded with 10 μM HA-Ub and 200 nM mE1 in 100 mM Tris pH7.4, 2 mM ATP, 5 mM $MgCl_2$ and 200 μM DTT at 37°C for 30 minutes. Flow rate: 1 ml/min. Fraction size: 1.3 ml.

We first squeezed UBE2L3-Ub into HeLa cells and analyzed the formation of HA-tagged adducts over time using immunoblotting. We treated cells with E1 inhibitor prior to introduction of the E2-Ub conjugates by SQZ to prevent any recycling of HA-Ub following conjugation and deconjugation to substrates. Unlike HA-Ub, for which an initial peak was observed shortly after squeezing, HA adducts accumulated over a 2-hour period in cells exposed to UBE2L3-Ub (Figure 5.3.6c). Again, we observed much stronger signals in lysates than in squeezed cells, possibly due to a loss of the machinery responsible for destruction of ubiquitinated proteins during cell lysis. We compared blots from a shorter exposure of lysate to a longer exposure for squeezed cells, and observed that the banding patterns of HA adducts differ depending on the method of UBE2L3-Ub introduction. In lysate experiments, most of the signal is observed in the range of 50-100 kD, whereas in squeezed cells, more signal accumulates >150 kD. We propose that in cells, UBE2L3-Ub labels substrates that have accumulated longer ubiquitin chains than the substrates labeled in lysate. This difference may be a consequence of the loss of regulation in ubiquitinated protein production and destruction caused cell lysis, as discussed above.


(1)
Lysate + 10 μM UBE2L3-Ub
Cells Squeezed with 10 μM UBE2L3-Ub
ENDO
Cells Squeezed with 10 μM HA-Ub
ENDO
Time (min)
NC 0 5 15 30 60 120 E0 5 15 30 60 120 120 E0 5 15 30 60 120 120
250
150
100
75
50
37
25
20
400s exposure

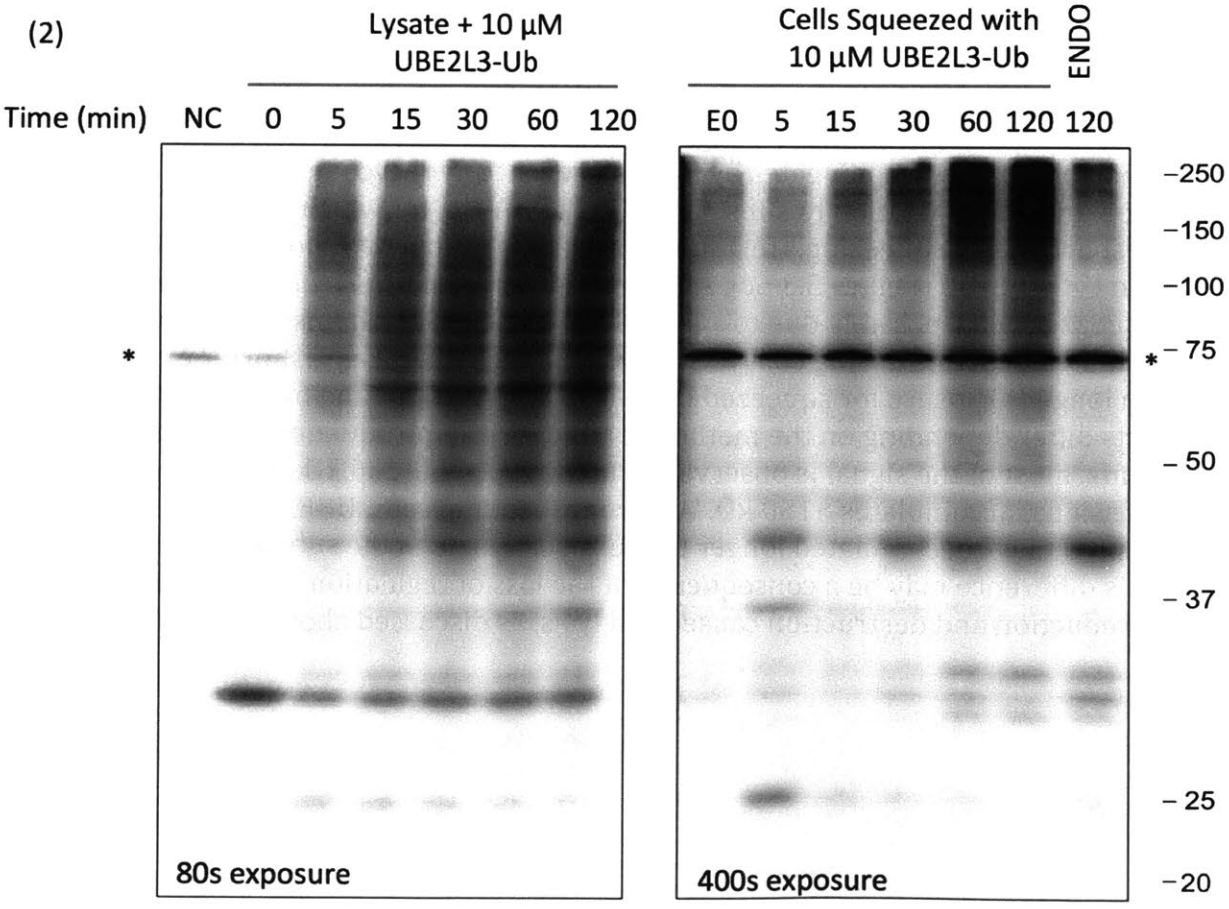

(2)
Lysate + 10 μM UBE2L3-Ub
Cells Squeezed with 10 μM UBE2L3-Ub
ENDO
Time (min)
NC 0 5 15 30 60 120
E0 5 15 30 60 120 120
*
*
250
150
100
75
50
37
25
20
80s exposure
400s exposure

Figure 5.3.7c. Pre-loaded UBE2L3-Ub can be delivered into HeLa cells and transfer HA-Ub onto substrates. Identical numbers and concentrations of HeLa cells were used in lysate and squeezing conditions. HeLa cells were treated with 20 μM E1 inhibitor MLN7243 for 1 hour before being trypsinized for sonication or for squeezing. Lysates were prepared by sonicating in PBS with Ca^{2+}/Mg^{2+} followed by centrifugation to remove insoluble material and subsequent addition of UBE2L3-HA-Ub. Squeeze experiments were performed in Opti-MEM containing 10 μM UBE2-L3-HA-Ub. (1) Comparison at the same exposure time. (2) Comparing lysate signal at a briefer exposure time to SQZ signal. *: non-specific band. E0: endocytosis control where cells were exposed but not squeezed. NC: no contact control.

We then squeezed UBC7-HA-Ub into HeLa cells pre-treated with E1 inhibitor and compared it by immunoblotting to addition of UBC7-HA-Ub directly to the lysate. UBC7 has been reported to form dimers and build ubiquitin chains on its active site *in vitro*, even in the absence of E3 ligases or substrates (Liu et al., 2014). We observed that the purified UBC7-Ubs had multiple Ubs covalently attached (multiple bands at time = 0 and when UBC7-Ub was run separately). The UBC7-nUb adducts disappear after UBC7-Ub are added to lysates or squeezed into cells, indicating that these UBC7-Ubs are utilized or degraded in both lysate and squeezed cells over time (Figure 5.3.6d). For UBC7-Ub, there is still a significant difference in the rate of HA-tagged adduct formation for squeezed cells and cell lysates, but differences in the banding pattern in lysate and squeezed cells are less pronounced than observed for UBC2L3.

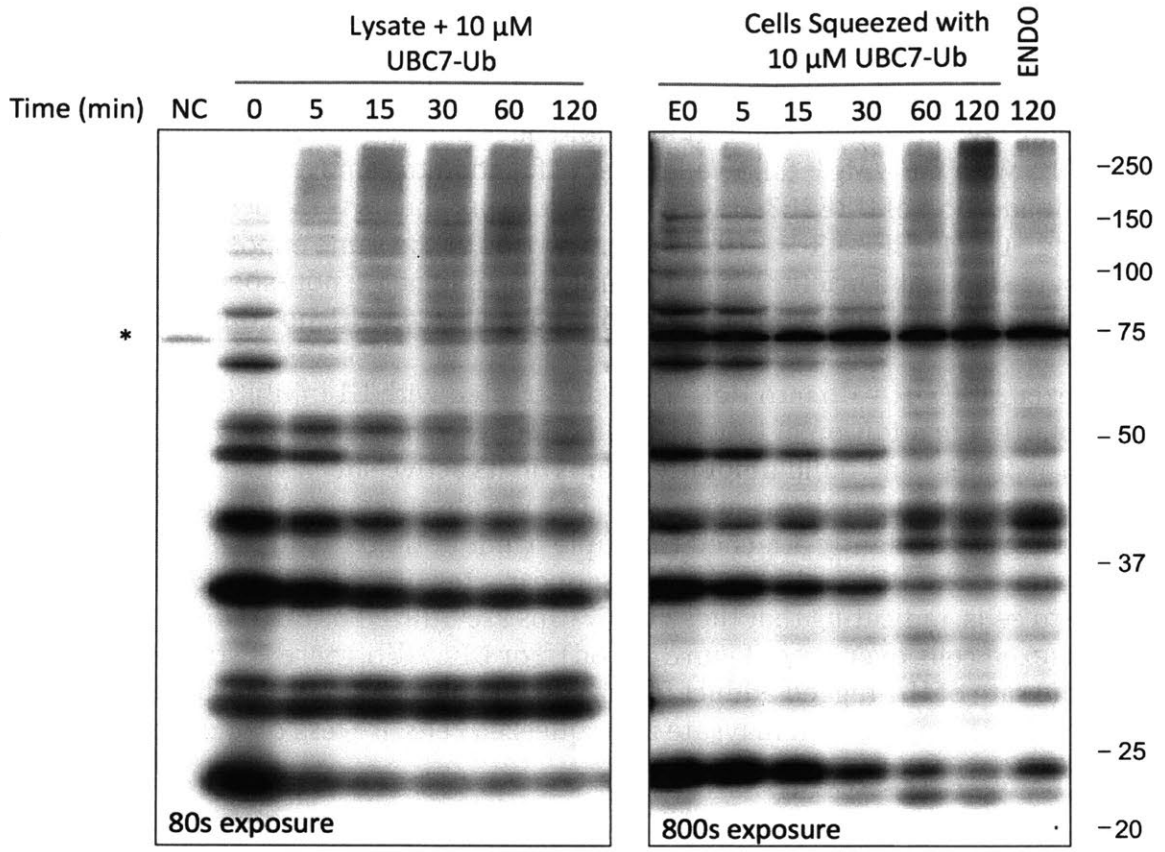


Figure 5.3.7d. Pre-loaded UBC7-Ub can be delivered into HeLa cells and transfer its attached HA-Ub onto substrates. Lysate collection and SQZ were conducted as described in Figure 5.3.6c. SQZ experiments were performed in OptiMEM containing 10 μM UBC7-HA-Ub. (*: non-specific band. E0: endocytosis control where cells were not exposed but not squeezed. NC: no contact control.

5.3.3 Discussion and Future Directions

We have shown that SQZ technology can be used to deliver both HA-Ub and HA-Ub-loaded E2s into the cytosol of HeLa cells. The delivered HA-Ub molecules are used by the endogenous ubiquitin conjugation machinery and attached to substrates within 5 minutes of squeezing. HA-Ub delivered as thioesters carried by select E2s are also attached to substrates in an E1-independent manner. The identity of the E2s that carry the Ub moiety has a substantial impact on the kinetics of Ub use and the pattern of HA-tagged adducts formed.

While epitope-tagged ubiquitins have been expressed by transfection in cells and even in animals to study the ubiquitination process, we have not been able to resolve the kinetics of Ub conjugation in the context of an intact cell. Even with an inducible tagged-Ub expression system, the timescale for ubiquitin production following transcription and translation is on the order of hours. On the other hand, it is possible to obtain kinetic

information from experiments conducted in HeLa lysates, but these results do not represent what happens in an intact cell. The SQZ platform allows us to deliver a suitable amount of material, such as tagged-Ub, into the cytosol in the timeframe of 1-2 minutes. Since we can remove most HA-Ub in the media and thereby stop delivery by simple centrifugation, we can expose cells to HA-Ub for as short as 3 minutes and then chase them as long as is required. This method has thus allowed us to examine ubiquitination events that occur as rapidly as within 5 minutes in the context of a semi-intact cell, which to the best of our knowledge has not been possible with previous methods.

In addition, the SQZ platform allowed us to introduce E2 thioesters into the cytosol. Since E2-Ub thioesters are produced post-translationally by the action of E1s, simple genetic expression of such adducts is impossible. While overexpression of E2s has been used, E2 loading with Ub is regulated by factors within the cell. We have not observed significant levels of E2s loaded with Ub in cells prepared under non-reducing conditions in SDS sample buffer (data not shown). Delivery of pre-loaded E2 allows us to overcome possible regulatory and feedback mechanisms that would complicate interpretation of E2 overexpression and directly observe the contribution of E2 identity to the collection of cellular proteins targeted.

In addition to the naturally occurring E2 thioesters, there are alternative non-natural chemically modified Ub analogs that we can now deliver into cells. Various chemical modifications on Ub can be applied to label proteins involved in the ubiquitin/deubiquitination cascade. A recently published Ub-based probe is the E1-E2-E3 cascading probe Ub-dehydroalanine (Dha), which can be used to label components of the ubiquitin pathway (Mulder et al., 2016). When loaded by an E1, about 50% of Ub-Dha reacts with the E1 active site cysteine as a thioester, while the remainder forms a thioether bond that cannot be transferred to an E2. In similar fashion, 50% of E2-Ub-Dha exists as thioesters, while the remainder yield the non-hydrolysable thioether. When HA-Ub-Dha is loaded onto a selected E2 as a thioester *in vitro* and subsequently delivered to cells, it should label any interacting E3s with an active site cysteine using HA-Ub-Dha thioether, which would allow us to immunoprecipitate for the HECT or RBR E3 ligases that interact with the target E2. Our lab has also developed Ub analogs that selectively label deubiquitinases (DUbs). These probes are comprised of a C-terminally truncated ubiquitin protein functionalized at the C-terminus with electrophiles: ubiquitin vinyl sulfone (UbVS) and Ub vinylmethylester (UbVME) have been used in lysates to label such DUBs (Borodovsky et al., 2001; Love et al., 2009). As an extension of these results we can now deliver these DUb probes into cells to identify the active DUbs in the context of intact cells

There is still a limited understanding of the extent to which cells are stressed during SQZ, and the duration and mechanism of recovery. Since the process of cell squeezing must cause mechanical stress, with might include a corresponding activation of the ubiquitin conjugation machinery, our results must be interpreted with this possibility in mind. Future work should include a characterization of the cellular response to membrane disruption.

5.3.4 Material and Methods

HA (Horseradish peroxidase (HRP)-conjugated monoclonal rat 3F10, Sigma Aldrich; Horseradish peroxidase (HRP)-conjugated monoclonal mouse sc-7392, Santa Cruz Biotechnology; HA rabbit polyclonal sc-805, Santa Cruz Biotechnology), GAPDH (Horseradish peroxidase (HRP)-conjugated monoclonal rabbit; Cell Signaling Technology), Horseradish peroxidase (HRP)-conjugated goat-anti-rabbit secondary antibody (Cell Signaling Technology)

Plasmids, Protein expression and purification

HA-Ub, UBC7 and UBE2L3 were expressed using a pET SUMO plasmid cleaved at the NheI and HindIII site to remove the SUMO fusion but retaining the N-terminal His tag and the thrombin cleavage site. Murine E1 is expressed using a plasmid provided by Addgene (#32534). Proteins were expressed using BL21(DE3) *E. coli* in terrific broth (TB), with overnight 1 mM IPTG induction at 30 °C. Bacteria were collected by centrifugation, re-suspended in 50 ml PBS with 10 mg lysozyme and 1 protease inhibitor pellet (Roche), sonicated for 60 seconds three times and spun at 17000 rpm for 1 hour at 4 °C to collect lysate. His_6-tagged proteins were purified by Ni-NTA column, washed with 10 mM imidazole, 500 mM NaCl, 50 mM Tris pH 7.5, and eluted in 250 mM imidazole, 500 mM NaCl, 50 mM Tris pH 7.5. All proteins were further purified with S75 size exclusion column (GE healthcare) into TBS.

Cell Culture

HeLa cells were maintained in Dulbecco's modified Eagle's medium (DME) supplemented with 10% IFS in humidified air containing 5% CO_2 at 37°C.

5.3.5 References

Bazirgan, O.A., and Hampton, R.Y. (2008). Cue1p is an activator of Ubc7p E2 activity in vitro and in vivo. The Journal of biological chemistry *283*, 12797-12810.

Borodovsky, A., Kessler, B.M., Casagrande, R., Overkleeft, H.S., Wilkinson, K.D., and Ploegh, H.L. (2001). A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. The EMBO journal *20*, 5187-5196.

Carlson, N., and Rechsteiner, M. (1987). Microinjection of ubiquitin: intracellular distribution and metabolism in HeLa cells maintained under normal physiological conditions. The Journal of Cell Biology *104*, 537-546.

Cohen, I., Wiener, R., Reiss, Y., and Ravid, T. (2015). Distinct activation of an E2 ubiquitin-conjugating enzyme by its cognate E3 ligases. Proceedings of the National Academy of Sciences of the United States of America *112*, E625-E632.

Corn, P.G. (2007). Role of the ubiquitin proteasome system in renal cell carcinoma. BMC Biochemistry *8 Suppl 1*, S4.

Ding, X., Stewart, M., Sharei, A., Weaver, J.C., Langer, R.S., and Jensen, K.F. (2017). High-throughput Nuclear Delivery and Rapid Expression of DNA via Mechanical and Electrical Cell-Membrane Disruption. Nature Biomedical Engineering *1*, 0039.

Eldridge, M.J.G., Sanchez-Garrido, J., Hoben, G.F., Goddard, P.J., and Shenoy, A.R. The Atypical Ubiquitin E2 Conjugase UBE2L3 Is an Indirect Caspase-1 Target and Controls IL-1β Secretion by Inflammasomes. Cell Reports *18*, 1285-1297.

Fu, B., Li, S., Wang, L., Berman, M.A., and Dorf, M.E. (2013). The ubiquitin conjugating enzyme UBE2L3 regulates TNFα-induced linear ubiquitination. Cell Research *24*, 376.

Fujioka, A., Terai, K., Itoh, R.E., Aoki, K., Nakamura, T., Kuroda, S., Nishida, E., and Matsuda, M. (2006). Dynamics of the Ras/ERK MAPK Cascade as Monitored by Fluorescent Probes. The Journal of biological chemistry *281*, 8917-8926.

Jacobson, A.D., Zhang, N.-Y., Xu, P., Han, K.-J., Noone, S., Peng, J., and Liu, C.-W. (2009). The lysine 48 and lysine 63 ubiquitin conjugates are processed differently by the 26S proteasome. The Journal of biological chemistry.

Jimenez, A.J., Maiuri, P., Lafaurie-Janvore, J., Divoux, S., Piel, M., and Perez, F. (2014). ESCRT machinery is required for plasma membrane repair. Science (New York, NY) *343*, 1247136.

Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thanedar, S., Roitelman, J., Chau, V., and Wiertz, E. (2004). Human HRD1 is an E3 ubiquitin ligase involved in degradation of proteins from the endoplasmic reticulum. The Journal of biological chemistry *279*, 3525-3534.

Kollmannsperger, A., Sharei, A., Raulf, A., Heilemann, M., Langer, R., Jensen, K.F., Wieneke, R., and Tampé, R. (2016). Live-cell protein labelling with nanometre precision by cell squeezing. Nature Communications *7*, 10372.

Lewis, M., Vyse, S., M Shields, A., Boeltz, S., Gordon, P., Spector, T., J Lehner, P., Walczak, H., and Vyse, T. (2015). UBE2L3 Polymorphism Amplifies NF-κB Activation and Promotes Plasma Cell Development, Linking Linear Ubiquitination to Multiple Autoimmune Diseases. The American Journal of Human Genetics *96*, 221-234.

Lim, K.-L., and Tan, J.M.M. (2007). Role of the ubiquitin proteasome system in Parkinson's disease. BMC Biochemistry *8*, S13-S13.

Liu, W., Shang, Y., Zeng, Y., Liu, C., Li, Y., Zhai, L., Wang, P., Lou, J., Xu, P., Ye, Y., *et al.* (2014). Dimeric Ube2g2 simultaneously engages donor and acceptor ubiquitins to form Lys48-linked ubiquitin chains. The EMBO journal *33*, 46-61.

Love, K.R., Pandya, R.K., Spooner, E., and Ploegh, H.L. (2009). Ubiquitin C-terminal electrophiles are activity-based probes for identification and mechanistic study of ubiquitin conjugating machinery. ACS Chemical Biology *4*, 275-287.

Miranda, M., and Sorkin, A. (2007). Regulation of receptors and transporters by ubiquitination: new insights into surprisingly similar mechanisms. Molecular interventions *7*, 157-167.

Misra, M., Kuhn, M., Löbel, M., An, H., Statsyuk, A.V., Sotriffer, C., and Schindelin, H. (2017). Dissecting the Specificity of Adenosyl Sulfamate Inhibitors Targeting the Ubiquitin-Activating Enzyme. Structure *25*, 1120-1129.e1123.

Mulder, M.P.C., Witting, K., Berlin, I., Pruneda, J.N., Wu, K.-P., Chang, J.-G., Merkx, R., Bialas, J., Groettrup, M., Vertegaal, A.C.O., *et al.* (2016). A cascading activity-based probe sequentially targets E1–E2–E3 ubiquitin enzymes. Nature Chemical Biology *12*, 523-530.

Sharei, A., Zoldan, J., Adamo, A., Sim, W.Y., Cho, N., Jackson, E., Mao, S., Schneider, S., Han, M.-J., Lytton-Jean, A., *et al.* (2013). A vector-free microfluidic platform for intracellular delivery. Proceedings of the National Academy of Sciences of the United States of America *110*, 2082-2087.

Stewart, M.D., Ritterhoff, T., Klevit, R.E., and Brzovic, P.S. (2016). E2 enzymes: more than just middle men. Cell Research *26*, 423.

Tsuchida, S., Satoh, M., Takiwaki, M., and Nomura, F. (2017). Ubiquitination in Periodontal Disease: A Review. International journal of molecular sciences *18*.

Upadhya, S.C., and Hegde, A.N. (2007). Role of the ubiquitin proteasome system in Alzheimer's disease. BMC Biochemistry *8 Suppl 1*, S12.

Chapter 6 Conclusions and Outlook

During my PhD, I have used UBC6e-deficient cells and VHHs to study the function of UBC6e in ERAD. We have found that UBC6e is involved in at least two processes: general ERAD for the degradation of misfolded substrates; and a process in which ERAD components EDEM1, OS-9 and SEL1L are down regulated. In addition, we found that the epitope of VHH05, located within residues 168-186 of UBC6e, might be a gateway helix that regulates activity. Stress-induced phosphorylation of Ser184 possibly allows binding by cytosolic factors to enhance UBC6e activity.

The study of ERAD complexes has frequently centered on the E3 ubiquitin ligases, but much less is known about the roles of E2 ubiquitin conjugating enzymes in the process. There are about 10 known E3s that localize to the ER, while there are only three known ER-localized E2s. These three E2s need to co-operate with different E3s for different cellular processes. Our studies in UBC6e-deficient cells indicate that, in a similar manner, E3 ligases can work with multiple E2s. How is E2/E3 complex formation controlled? How do E2s contribute to substrate specificity? These are some of the questions that, if answered, would contribute to our understanding of the ERAD process.

During ER stress, the unfolded protein response (UPR) is activated to increase production of ER chaperones and ERAD components to resolve ER stress. However, transcription and translation can take hours, so there must be mechanisms that take less time to enact. Our data supports the hypothesis that UBC6e plays a role in enhancing ERAD without transcriptional activation. Stress-induced UBC6e phosphorylation occurs within 30 minutes, and the phosphorylation site is on a gateway helix, which when bound, enhances UBC6e enzymatic activity. We would like to identify the cellular factors that bind to phosphorylated UBC6e to induce activation, and further understand how this pre-UPR response contributes to resolving ER stress.

E2s need to be loaded with ubiquitin by E1 activating enzyme to become active. Since E2-Ub thioesters cannot be genetically coded, scientists lack the tools to study functional E2-Ubs in the cell. We have recently employed SQZ technology developed by the Langer and Jensen Labs to deliver purified E2s loaded with epitope-tagged ubiquitins *in vitro*. We hope to better understand how E2-Ub complexes contribute to substrate selection, and to understand how E2-Ubs interacts with different E3s by using ubiquitin analogs that allow for E2/E3 cross-linking.

We have also developed VHH05, which specifically targets UBC6e (170-184). We have used this VHH to perturb UBC6e function both *in vitro* and *in vivo*. Because the epitope of VHH05 is a linear peptide in UBC6e, we further explored the use of this epitope as a tag for proteins of interest. We have shown that VHH05 sortagged with different labels can specifically immunoprecipitate, immunblot and immunofluorescently label tagged proteins. We hope to use this VHH to intracellularly target tagged proteins to selected compartments.

Appendix A

Protein Thioester Synthesis Enabled by Sortase

Reprinted From

Jingjing J. Ling, Rocco L. Policarpo, Amy E. Rabideau, Xiaoli Liao, and Bradley L. Pentelute

Journal of the American Chemical Society 134(26): 10749-10752

Protein Thioester Synthesis Enabled by Sortase

Jingjing Ling, Rocco L. Policarpo, Amy E. Rabideau, Xiaoli Liao, Bradley L. Pentelute*

Department of Chemistry, Massachusetts Institute of Technology, 16-573a, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139.

Supporting Information Placeholder

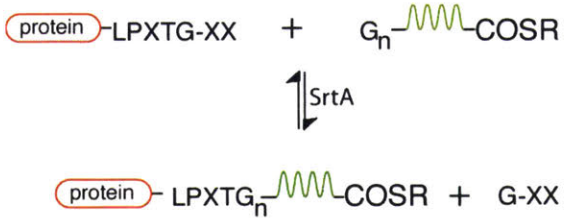
ABSTRACT: Proteins containing a C-terminal thioester are important intermediates in semi-synthesis. Currently there is one main method for the synthesis of protein thioesters that relies upon the use of engineered inteins. Here we report a simple strategy, utilizing Sortase A, for routine preparation of recombinant proteins containing a C-terminal $^{\alpha}$thioester. We used our method to prepare two different anthrax toxin cargo proteins: one containing an $^{\alpha}$thioester and another containing a D-polypeptide segment situated between two protein domains. We show that both variants can translocate through protective antigen pore. This new method to synthesize a protein thioester allows for interfacing of sortase-mediated ligation and native chemical ligation.

Chemical tailoring of proteins is a powerful approach to investigate structure function relationships and the role of post-translational modifications.[1-6] Protein semi-synthesis[4,7] and total synthesis[8] are commonly used to introduce novel functionalities into proteins. Both approaches rely on native chemical ligation (NCL)–a chemoselective amide forming reaction between an $^{\alpha}$thioester and an N-terminal cysteine moiety.[9] When protein $^{\alpha}$thioesters are generated from engineered inteins and then modified with NCL the process is referred to as expressed protein ligation.[10-12] Inteins are protein self-splicing elements that can be engineered to generate protein $^{\alpha}$thioesters after self-cleavage in the presence of small molecule thiol.[10-12] Despite a number of new methods for the chemical synthesis of peptide $^{\alpha}$thioesters, intein mediated synthesis of protein $^{\alpha}$thioesters is the only direct route to generate this important functionality needed for semi-synthesis.

Here we report a sortase-mediated approach for the facile synthesis of protein $^{\alpha}$thioesters (Scheme 1). We show that the calcium dependent enzyme sortase A (SrtA) from *Staphylococcus aureus* can be used to attach synthetic oligoglycine $^{\alpha}$thioesters to a number of different proteins with good yield and efficiency. In addition, this approach allowed us to prepare two different model cargo proteins and probe their translocation through anthrax toxin protective antigen. We found that the anthrax toxin pore can translocate cargo proteins into the cell that either contained an $^{\alpha}$thioester or a D-polypeptide segment linking two protein domains.

Sortases are a class of thiol-containing transpeptidases that anchor proteins to the bacterial cell wall.[13] SrtA recognizes a C-terminal LPXTG sequence and cleaves the threonine-glycine bond to form a thioacyl-linked intermediate.[14] This intermediate is primed to react with the N-terminal amino group of an oligoglycine motif. Recent effort has shown that sortases can be co-opted for the site-specific modification of proteins at the N or C-terminus.[15-25] This sortagging, transpeptidation reaction has been used extensively to attach virtually any water-soluble molecule to a protein of interest. Sortagging reactions are executed in calcium containing aqueous buffer (pH 7-8.5) at nanomolar to micromolar concentration of SrtA.[24] To carry out the sortagging reactions in water the N-terminal membrane spanning region of SrtA is removed.

Scheme 1. Protein $^{\alpha}$thioester synthesis using SrtA



We prepared variants of an oligoglycine $^{\alpha}$thioester and carried out model studies with peptide KLPETGG. During the initial stages of our studies Chen et al. reported an evolved SrtA (SrtA*) that had improved enzymatic kinetics.[26] Once we confirmed SrtA* and SrtA gave similar product yield in model peptide studies, we chose to work exclusively with SrtA* because it substantially shortened reaction time. Reactions with SrtA* were complete in 30 minutes as opposed to 2 hours with SrtA (Figure S1-S2). The analytical RP-HPLC traces for a model reaction between KLPETGG and G_5F-COSR in the presence of SrtA* are shown in Figure 1. We used RP-HPLC to purify the sortagged reaction product (6.9 mg, 54% yield) and then performed an NCL reaction under standard conditions. After purification, the NCL product was isolated (83 % yield, 2 mg) and characterized by high resolution

LCMS (Figure 1b-d). We conducted additional sortagging studies to investigate two aspects of the oligoglycine $^{\alpha}$thioester: varying the C-terminal amino acid (Xaa) or varying the number of glycine residues. It has been demonstrated that the reactivity of a peptide $^{\alpha}$thioester is dependent upon the identity of the C-terminal amino acid; Gly is more reactive than Leu, for instance.[27,28] Product yields determined from analytical RP-HPLC ranged from 54-68 % with SrtA (Table S1) and 74-84 % with SrtA* (Table 1) suggesting the C-terminal amino acid Xaa does not significantly affect the reaction yield. We next investigated the relationship between reaction yield and the number of N-terminal glycine residues. The best yields were obtained when three or more glycines were used (Table 2).

Table 1. SrtA*-mediated ligation reaction yields with different G_5-Xaa-COSR to model peptide

Xaa	Gly	Phe	Ser	Leu
Yield (%)	76(2)	84(3)	74(3)	75(2)

Reagents and conditions: 500 μM KLPETGG was reacted with 1 mM G_5-Xaa-COSR for 30 min in the presence of 5 μM SrtA* and SrtA* buffer (10 mM $CaCl_2$, 50 mM Tris, 150 mM NaCl) pH 7.5. The thioester R group was -CH_2-CH_2-L-R-$CONH_2$. Standard deviations are shown in parentheses.

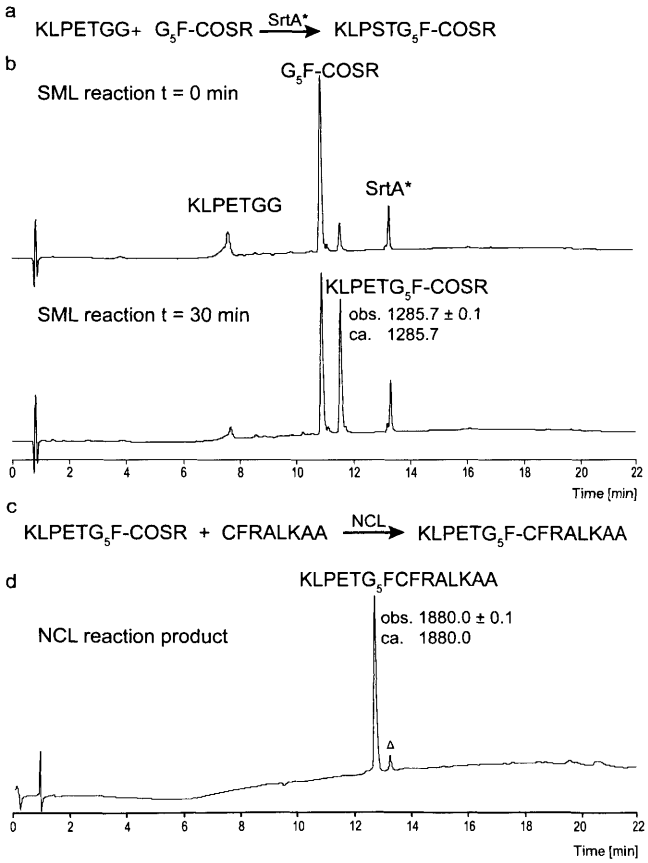


Figure 1. LC data for peptide $^{\alpha}$thioester synthesis using SrtA* followed by NCL. **(a)** SrtA*-mediated ligation (SML) of model peptide. **(b)** Analytical RP-HPLC of crude SML reaction for t = 0 min and t = 30 min. Reaction conditions: 1 mM G_5F-COSR, 500 μM KLPETGG, 5 μM SrtA*, SrtA* buffer, pH 7.5. **(c)** Native chemical ligation (NCL) with sortagged thioester reaction product. **(d)** Analytical RP-HPLC of purified NCL reaction product. Reaction conditions are reported in the supporting information. All analytical RP-HPLC signals are measured at 214 nm. Δ: impurity peak.

Table 2. SrtA*-mediated ligation reaction yields with increasing number of glycines in G_n-COSR to model peptide.

Number of Gly	1	2	3	4	5	6
Yield(%)	62(1)❖	46(8)	68(6)	69(3)	70(1)	59(8)

Reagents and conditions: 500 μM KLPETGG was reacted with 1 mM G_n-COSR for 30 min in the presence of 5 μM SrtA* and SrtA* buffer, pH 7.5. ❖: Both KLPETGG-COSR and KLPETG-COSR were formed when G-COSR was used as the nucleophile. The yield represents the sum of both reactions.

With the ability to sortag oligopeptide $^{\alpha}$thioesters to peptides bearing a C-terminal acceptor sequence, we carried out investigations with three different model proteins. We prepared variants of eGFP, lethal factor N-terminal domain (LF_N) from anthrax toxin,[29] and lethal factor N-terminal domain fused to diphtheria toxin A-chain (LF_N-DTA).[29,30] DTA is the catalytic domain from diphtheria toxin and catalyzes the ADP-ribosylation of elongation factor-2 within the cytosol thereby halting protein synthesis and causing cell death.[31,32] LF_N and LF_N-DTA are moieties used to probe the molecular basis of anthrax toxin protein translocation.[33] We prefer to express proteins as SUMO-protein fusions because expression yields are enhanced and the native N-terminus is generated after removal of SUMO.

We first investigated, by use of high resolution LCMS, whether SrtA* could tag thioesters onto our three different model proteins and found that protein thioester products are readily formed in 30 minutes (Figure S4). We then proceeded to prepare protein thioesters on milligram scale using the approach shown in Figure 2a. In particular, a one-pot method was employed whereby we first removed the N-terminal SUMO tag with SUMO protease and subsequently added SrtA*, Ni-NTA agarose beads, and oligoglycine $^{\alpha}$thioester peptide. After completion of the SrtA*-mediated ligation (SML) reaction, pure protein thioester was isolated by simple filtration and concentration because all unreacted material remained bound to the Ni-NTA agarose beads. We obtained good yields of pure protein $^{\alpha}$thioester and observed minimal amounts of SrtA-mediated hydrolysis when analyzed by high resolution LCMS (Figure 2b-d). Our isolated yields for the three different model protein thioesters ranged from 40-80% (Table S3), which is consistent with prior reports for SML. When we monitored the sortagging reaction by SDS-PAGE and LCMS, we found some product still bound to the Ni-NTA agarose beads for LF_N and LF_N-DTA but not eGFP (Figure S5) suggesting the reaction

yields are in part determined by the intrinsic properties of the protein. For eGFP-COSR we carried out NCL with the model peptide CFRALKAA under standard conditions (pH 7, TCEP, MPAA catalyst) and obtained 0.8 mg product (98% yield) (Figure S6).[34]

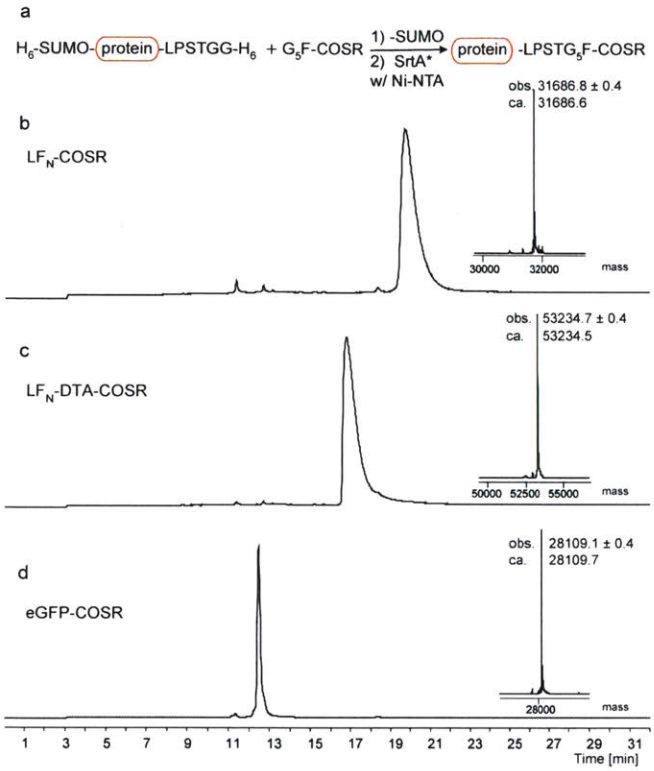


Figure 2. LCMS data for protein $^{\alpha}$thioester synthesis using SrtA*. **(a)** SML of model proteins. **(b)-(d)** LCMS and deconvoluted MS (inset) of the main component for LF_N, LF_N-DTA and eGFP thioesters. Reaction conditions: **(1)** SUMO cleavage: 1 μg SUMO protease per 1 mg protein for 90 min at room temperature. **(2)** SML: 500 μM G_5F-COSR, 100 μM protein-LPSTGG, 5 μM SrtA*, 75 μl Ni-NTA agrose slurry per mg protein, SrtA* buffer, pH 7.0, 30 min. LCMS traces are shown as total ion current.

One of the proteins we have chosen to work with in these studies is LF_N-DTA. This chimera serves as an excellent model cargo protein to probe translocation through the anthrax toxin protective antigen (PA) pore. Having facile chemical access to this molecule allows for the incorporation of non-natural moieties to further elucidate the mechanisms in which this protein enters the cell through the PA pore. To begin these investigations, we aimed to study the translocation of LF_N-DTA-COSR and LF_N-D-linker-DTA, where "D-linker" refers to a small D-peptide tether between LF_N and DTA. The preparation of these synthetic constructs is simplified by our new approach to generate protein thioesters. We prepared LF_N-D-linker-DTA using the approach shown in Figure 3a. The synthetic approach involves first sortagging an oligoglycine thioester containing D-amino acids to LF_N followed by NCL to ligate on the C-terminal DTA domain. We also prepared LF_N-L-linker-DTA using the same approach to serve as a control in our translocation assays (Figure S7). By comparing the translocation efficiency of LF_N-D-linker-DTA against the L-amino acid variant, we were able to evaluate the stereochemical requirement for successful translocation through the pore.

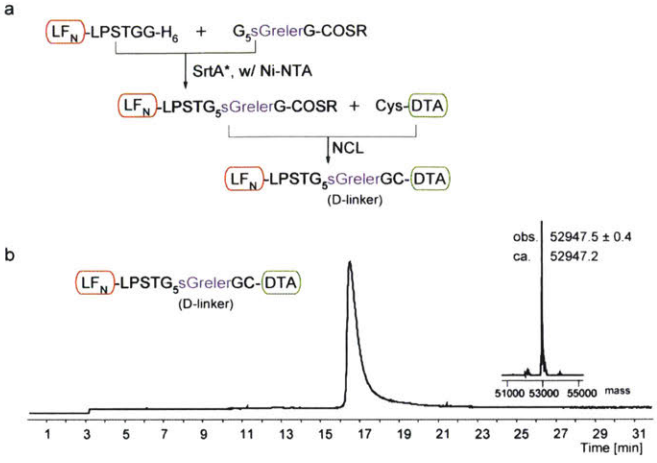


Figure 3. Installing a D-peptide segment between LF_N and DTA. **(a)** Synthetic strategy used to modify the linker region (lower case = D-amino acid). **(b)** LCMS and deconvoluted MS (inset) for LF_N-D-linker-DTA. The reaction conditions are reported in the supporting information. LCMS traces are shown as total ion current.

With our modified LF_N-DTA variants in hand, we tested if they could translocate through PA.[29] In this protein translocation assay, anthrax toxin PA and LF_N-DTA were added to CHO-K1 cells and the amount of LF_N-DTA delivered to the cytosol was inferred by measuring protein synthesis via ^{3}H-Leu incorporation into the cellular proteome.[30,33] Once LF_N-DTA accesses the cytosol it inhibits protein synthesis. The protein synthesis levels for the variants are shown in Figure 4. We found all variants to translocate at levels similar to wild-type LF_N-DTA.

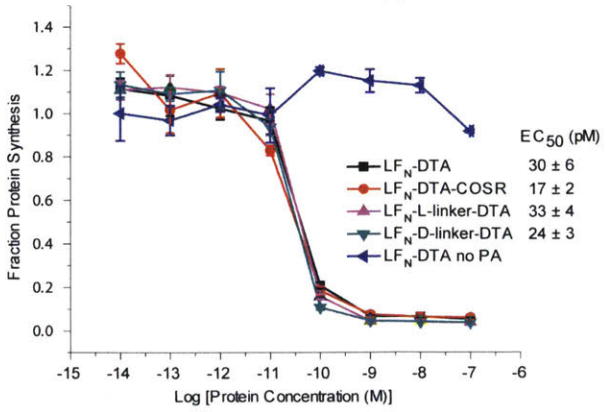


Figure 4. Translocation of LF_N-DTA variants into CHO-K1 cells. Cells were incubated with LF_N-DTA, LF_N-DTA-COSR, LF_N-L-linker-DTA, or LF_N-D-linker-DTA at different concentrations in the presence or absence of 10 nM PA for 30 minutes. The media was then replaced with leucine-

free medium supplemented with 1 μCi/ mL ^{3}H-Leu, and incubated for 1 hour. After incubation, the cells were washed three times with cold PBS, scintillation fluid was added, and incorporated ^{3}H-Leu was determined by scintillation counting. Each data point represents the average of three trials at the specified concentration.

The translocation of LF_N-DTA-COSR suggests that the amide bond can be replaced with a thioester without affecting passage through the pore. We plan to explore the possibility of using the thioester functionality to capture translocated proteins and their possible binding partners in the cytosol.

Our findings indicate that a model cargo protein containing non-natural modifications in the linker region is capable of translocation through the PA pore. Successful translocation of LF_N-D-linker-DTA demonstrates that the stereochemical constraints on PA-mediated translocation are minimal, provided that requirements for prepore binding and translocation initiation are met. This finding is inline with a prior report investigating replacement of the N-terminus of LF_N with a D-peptide segment.[35]

This method to generate an $^{\alpha}$thioester provides flexibility for covalent modification of proteins by interfacing key ligation approaches. The termini or linker regions of proteins can be site-specifically modified in a modular manner by the use of sortase-mediated ligation and native chemical ligation. Linker regions between two protein domains can be easily modified with natural or unnatural moieties by synthesizing various peptide thioesters. We demonstrated the utility of this approach by preparing a variant of LF_N-DTA in containing a D-peptide segment in the linker region. To the best of our knowledge, this is the first time that a D-peptide fragment has been installed between two recombinantly expressed protein domains.

Our method of protein thioester generation could be used to overcome solubility limitations of sortagging reactions. We have found that sortagging efficiency directly parallels oligoglycine nucleophile concentration. Hydrolysis of LPXTG is a competing side-reaction for sortagging endeavors.[36,37] We found that optimal concentrations of oligogylcine needed to be 300 μM and above; at lower concentrations, we observed significant LPSTGG hydrolysis (Figure S8). However, some oligoglycine peptides and proteins are insoluble at these concentrations. Our approach could be used to overcome this solubility limitation because short oligoglycine peptide thioesters tend to be highly soluble in aqueous solution. Moreover, if a given peptide thioester is still found to be insoluble it can be further modified to increase solubility by installing an Arg tag on the thioester leaving group.[38] After sortagging with a peptide $^{\alpha}$thioester, native chemical ligation can be carried out in solvents and buffers that denature sortase but solubilize the coupling partners. Common solubilizing agents that can be used in native chemical ligation reactions include denaturants (urea or guanidinium), detergents, and organic solvents.

However, it should be noted that the utility of sortagging is limited to ligations in which introduction of the $LPXTG_n$ moiety does not significantly alter protein structure. In cases when the $LPXTG_n$ segment may alter the properties of the protein, other ligation methods should be considered. Despite this limitation, sortagging has been used in numerous instances to modify proteins for biological study.[15-25] In our case, the $LPSTG_5$ linker did not affect the translocation of our LF_N-DTA constructs. One way to overcome this limitation would be to evolve sortase to recognize different and possibly shorter sequences. Recently, Piotukh et al. evolved SrtA to recognize FPXTG or APXTG motif, suggesting this may be possible.[39]

In summary, we have developed a SrtA-mediated ligation approach for the synthesis of recombinant protein thioesters. Protein thioesters are generated in 30 minutes with good yields, and pure products are isolated without elaborate purification steps.

Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org."

Corresponding Author

* blp@mit.edu

ACKNOWLEDGMENT

This research was supported the MIT startup funds for B.L.P., Agency for Science, Technology and Research National Science Scholarship (PhD) for J.L., a National Science Foundation Graduate Research Fellowship for A.E.R. We also thank R. John Collier for providing some of the laboratory equipment used to carry out the experiments. Some of the recombinant proteins employed in the study were prepared in the Biomolecule Production Core for the New England Regional Center of Excellence, supported by NIH grant number AI057159. We thank R. Ross and Ben T. Seiler for these services. We thank Mark Simon and Xiuyuan Li for for valuable comments and discussions. We thank Hidde L. Ploegh and Carla Guimaraes for the SrtA plasmid.

Supporting Information Available: Details of the peptide synthesis and purification, model SrtA and SrtA* reactions, protein construct preparation, and anthrax toxin cells assay. This material is available free of charge via the internet at http://pubs.acs.org.

REFERENCES

(1) Murakami, M.; Okamoto, R.; Izumi, M.; Kajihara, Y. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 3494.

(2) Nagorny, P.; Sane, N.; Fasching, B.; Aussedat, B.; Danishefsky, S. J. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 975.

(3) Bavikar, S. N.; Spasser, L.; Haj-Yahya, M.; Karthikeyan, S. V.; Moyal, T.; Kumar, K. S.; Brik, A. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 758.

(4) Vila-Perello, M.; Muir, T. W. *Cell* **2010**, *143*, 191.

(5) Scheuermann, J. C.; de Ayala Alonso, A. G.; Oktaba, K.; Ly-Hartig, N.; McGinty, R. K.; Fraterman, S.; Wilm, M.; Muir, T. W.; Muller, J. *Nature* **2010**, *465*, 243.

(6) Chatterjee, C.; McGinty, R. K.; Fierz, B.; Muir, T. W. *Nat. Chem. Biol.* **2010**, *6*, 267.
(7) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249.
(8) Kent, S. B. *Chem. Soc. Rev.* **2009**, *38*, 338.
(9) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776.
(10) Severinov, K.; Muir, T. W. *J. Biol. Chem.* **1998**, *273*, 16205.
(11) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705.
(12) Evans, T. C., Jr.; Benner, J.; Xu, M. Q. *Protein. Sci.* **1998**, *7*, 2256.
(13) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. *Science* **1999**, *285*, 760.
(14) Navarre, W. W.; Schneewind, O. *Mol. Microbiol.* **1994**, *14*, 115.
(15) Wu, Z.; Guo, X.; Guo, Z. *Chem. Commun. (Camb)* **2011**, *47*, 9218.
(16) Popp, M. W.; Ploegh, H. L. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 5024.
(17) Wu, Z.; Guo, X.; Wang, Q.; Swarts, B. M.; Guo, Z. *J. Am. Chem. Soc.* **2010**, *132*, 1567.
(18) Proft, T. *Biotechnol. Lett.* **2010**, *32*, 1.
(19) Clancy, K. W.; Melvin, J. A.; McCafferty, D. G. *Biopolymers* **2010**, *94*, 385.
(20) Yamamoto, T.; Nagamune, T. *Chem. Commun. (Camb)* **2009**, 1022.
(21) Tsukiji, S.; Nagamune, T. *Chembiochem* **2009**, *10*, 787.
(22) Guo, X.; Wang, Q.; Swarts, B. M.; Guo, Z. *J. Am. Chem. Soc.* **2009**, *131*, 9878.
(23) Samantaray, S.; Marathe, U.; Dasgupta, S.; Nandicoori, V. K.; Roy, R. P. *J. Am. Chem. Soc.* **2008**, *130*, 2132.
(24) Popp, M. W.; Antos, J. M.; Grotenbreg, G. M.; Spooner, E.; Ploegh, H. L. *Nat. Chem. Biol.* **2007**, *3*, 707.
(25) Mao, H.; Hart, S. A.; Schink, A.; Pollok, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2670.
(26) Chen, I.; Dorr, B. M.; Liu, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11399.
(27) Bang, D.; Pentelute, B. L.; Kent, S. B. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 3985.
(28) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068.
(29) Young, J. A.; Collier, R. J. *Annu. Rev. Biochem.* **2007**, *76*, 243.
(30) Milne, J. C.; Blanke, S. R.; Hanna, P. C.; Collier, R. J. *Mol. Microbiol.* **1995**, *15*, 661.
(31) Collier, R. J.; Cole, H. A. *Science* **1969**, *164*, 1179.
(32) Collier, R. J. *J. Mol. Biol.* **1967**, *25*, 83.
(33) Pentelute, B. L.; Sharma, O.; Collier, R. J. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2294.
(34) Johnson, E. C. B.; Kent, S. B. H. *J. Am. Chem. Soc.* **2006**, *128*, 6640.
(35) Pentelute, B. L.; Sharma, O.; Collier, R. J. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2294.
(36) Ton-That, H.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. *J. Biol. Chem.* **2000**, *275*, 9876.
(37) Möhlmann, S.; Mahlert, C.; Greven, S.; Scholz, P.; Harrenga, A. *Chembiochem* **2011**, *12*, 1774.
(38) Johnson, E. C.; Kent, S. B. *Tetrahedron. Lett.* **2007**, *48*, 1795.
(39) Piotukh, K.; Geltinger, B.; Heinrich, N.; Gerth, F.; Beyermann, M.; Freund, C.; Schwarzer, D. *J. Am. Chem. Soc.* **2011**, *133*, 17536.

Supporting Information for

Protein Thioester Synthesis Enabled by Sortase

Jingjing Ling, Rocco L. Policarpo, Amy E. Rabideau, Xiaoli Liao, Bradley L. Pentelute*

Department of Chemistry, Massachusetts Institute of Technology, 16-573a, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139.

Model Peptide Studies

Materials

2-(1H-Benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and α-Boc protected L-amino acids (Chem-Impex International, IL and Peptide Institute, Japan). MBHA resin was obtained from Anaspec, CA. N,N-Dimethylformamide (DMF), dichloromethane (DCM), diethyl ether, HPLC-grade acetonitrile, and guanidine hydrochloride were from VWR, PA. Trifluoroacetic acid (TFA) was purchased from NuGenTec, CA and Halocarbon, NJ. All other reagents were purchased from Sigma-Aldrich, MO and Invitrogen, CA, or as otherwise indicated.

Solid phase peptide synthesis

All model peptides were synthesized on a 0.2 mmol scale on 4-methylbenzhydrylamine (MBHA) resin using manual SPPS in situ neutralization Boc chemistry protocols[1]. Peptide thioesters were prepared as described above and thioester formation relied upon a mercaptopropionic acid (MPA) strategy[2]. Side-chain protection for L-amino acids were as follows: Arg(Tos), Cys(4-MeBzl), Glu(OcHex), Lys(2-ClZ), and Thr(Bzl). The resins were washed with DCM and air dried after completion of stepwise SPPS. The peptides were simultaneously cleaved from the resin and side-chain deprotected by treatment with 10% (v/v) *p*-thiocresol and 10%(v/v) *p*-cresol in anhydrous HF for 1 hr at 0 °C. Peptides were then triturated with cold diethyl ether, dissolved in 50% H_2O: 50% acetonitrile containing 0.1% TFA and lyophilized. ***These same solvent compositions were used in most experiments and will be referred to as A: 0.1 % TFA in H_2O and B: 0.1 % TFA in acetonitrile.***

Peptide purification

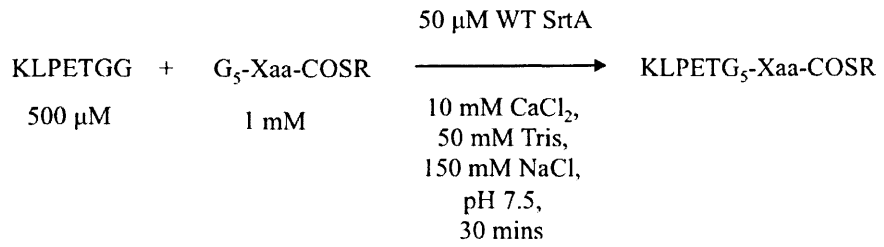
The crude peptides were dissolved in 95 % A: 5 % B and purified by preparative RP-HPLC (Agilent Zorbax SB C_{18} column: 21.2 x 250 mm, 7 μm, linear gradient: 1-41% B over 80 min, flow rate: 10 mL/min). HPLC fractions were spotted with MALDI matrix alpha-cyano-4-hydroxycinnamic acid (CHCA) in 50 % A: 50% B and checked for correct molecular masses. The purity of fractions was confirmed by analytical RP-HPLC (Agilent C_{18} Zorbax SB column: 2.1 x 150 mm, 5 μm, gradient: 1-51% B over 12 min, flow rate: 0.5 mL/min). Analytical HPLC UV absorbance traces in this work were measured at 214 nm. Model peptides used in this investigation are as follows:

Sequence	Observed (Da)	Calculated (Da; Mono.)
GGGGG-G-MPA-LR-$CONH_2$	716.4 ± 0.5	716.33
GGGGG-F-MPA-LR-$CONH_2$	806.6 ± 0.5	806.38
GGGGG-S-MPA-LR-$CONH_2$	746.8 ± 0.5	746.35
GGGGG-L-MPA-LR-$CONH_2$	773.0 ± 0.5	772.40
GGGGG-MPA-LR-$CONH_2$	659.3 ± 0.5	659.31
GGGG-MPA-LR-$CONH_2$	602.8 ± 0.5	602.29
GGG-MPA-LR-$CONH_2$	546.0 ± 0.5	545.27
GG-MPA-LR-$CONH_2$	489.2 ± 0.5	488.25
G-MPA-LR-$CONH_2$	432.3 ± 0.5	431.23
GGGGG-SGRELERG-MPA-LL-$CONH_2$	1500.8 ± 0.1	1500.74
GGGGG-sGrelerG-MPA-LL-$CONH_2$	1500.8 ± 0.1	1500.74
KLPETGG-$CONH_2$	700.1 ± 0.5	699.40
CFRALKAA	878.8 ± 0.5	878.48
GGGGGLRL-$CONH_2$	684.1 ± 0.5	684.80

MPA: 3-mercaptopropionic acid. Lower case letters represents d-amino acids.

Model sortase reaction with wildtype (WT) SrtA

Yield analysis with different G_5-Xaa-COSR



Xaa= Gly, Phe, Ser or Leu

For wild type (WT) sortase-mediated ligations, the substrate KLPETGG (500 μM) was incubated with G_n-Xaa-COSR (1 mM) in the presence of 50 μM WT sortase A (SrtA) in sortase buffer (10 mM $CaCl_2$, 50 mM Tris, 150 mM NaCl), pH 7.5 for 2 hours at room temperature. The reaction mixture was then quenched with equal volume of 50 % A: 50% B. The crude product yield was analyzed using RP-HPLC (**Agilent Zorbax SB C_{18} column: 2.1 x 150 mm, 5 μm, linear gradient: 1-61% B over 15 min, flow rate: 0.5 mL/min. The same analytical HPLC method has been used to analyze peptide sortagging reaction unless otherwise stated**). The yield was calculated by manual integration of area under the curve for KLPETGG (t_R: 7.2min). All experiments were performed in triplicate.

Table S1. Percentage yield for WT SrtA mediated ligations for G_5-Xaa-COSR thioesters (Xaa = Gly, Phe, Ser or Leu) with 500 μM KLPETGG.

Xaa	Gly	Phe	Ser	Leu
Yield (%)	54(4)	69(3)	67(3)	68(1)

Kinetic analysis of WT SrtA

The substrate KLPETGG (500 μM) was incubated with G_5F-COSR (500 μM) in the presence of 50 μM WT SrtA at room temperature in sortase buffer, pH 7.5. The reaction mixture was quenched with 20 μl of 50 % A: 50% B at time (in minutes) = 1, 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, and 120. The formation of product was analyzed using RP-HPLC via manual integration to obtain area under curve for the product peak (t_R: 12.6 min).

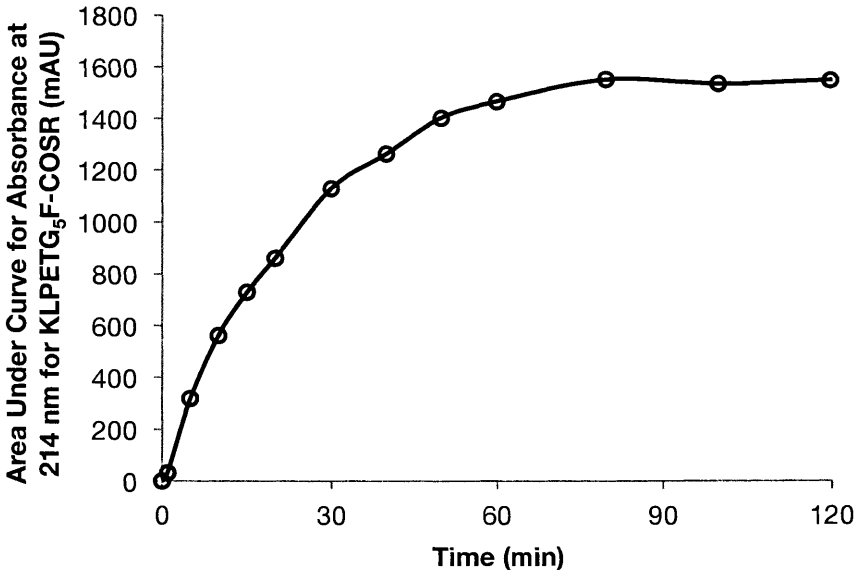


Figure S1. Rate of product formation in the presence of WT SrtA.

Kinetic analysis of SrtA*

The substrate KLPETGG (500 μM) was incubated with G_5F-COSR (500 μM) in the presence of 5 μM SrtA* at room temperature in sortase buffer, pH 7.5. The reaction was quenched with 20 μl of 50 % A: 50% B at time (in minutes) = 1, 5, 10, 15, 20, 30, 40, 50, and 60. The product yield was analyzed using RP-HPLC and then manually integrated to obtain area under curve for the product peak (t_R: 12.6 min).

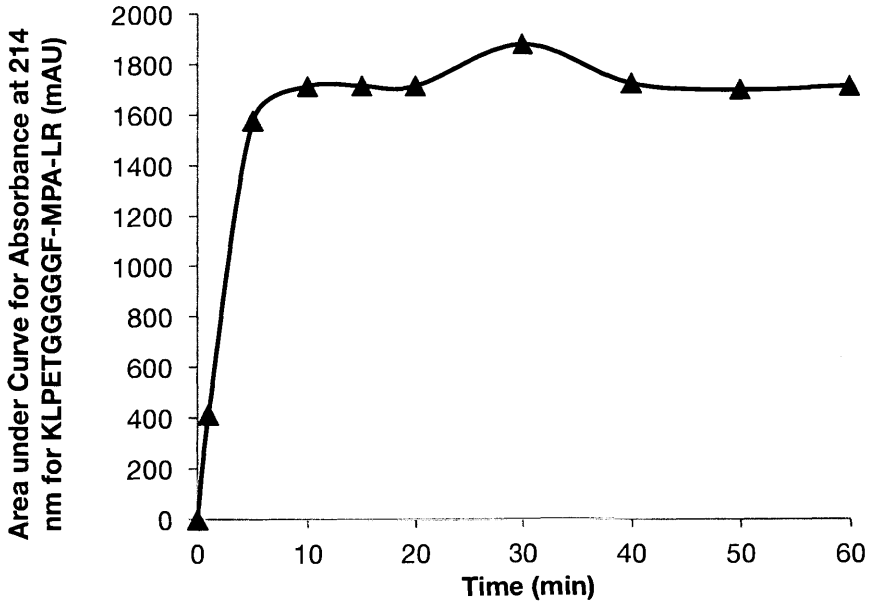
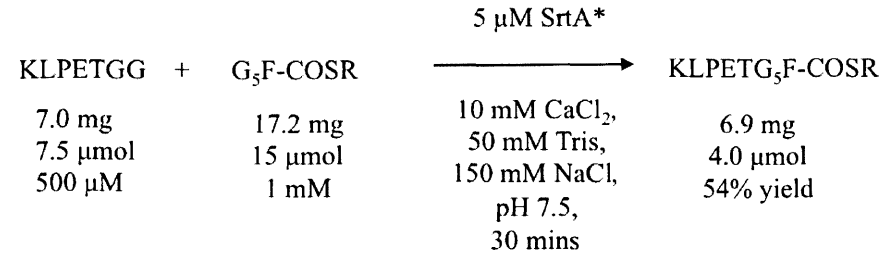


Figure S2. Rate of product formation in the presence of SrtA*.

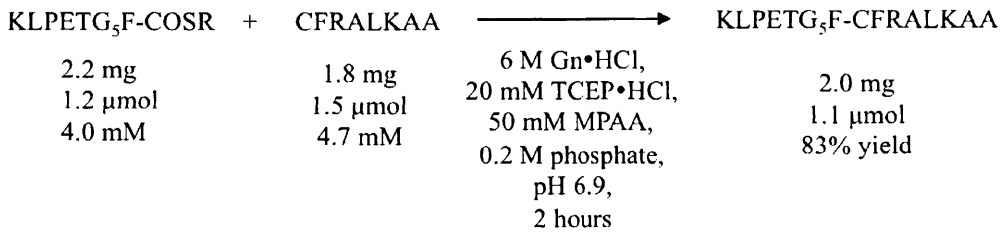
Sortase reaction followed by NCL with model peptide

SrtA* mediated ligation followed by purification



KLPETGG (7.0 mg, 7.5 μmol, 500 μM) was reacted with G_5F-COSR (17.2 mg, 15 μmol, 1 mM) with 5 μM SrtA* at room temperature for 30 minutes in sortase buffer, pH 7.5. The reaction was quenched with an equal volume of 95 % A: 5% B. The product KLPETG_5F-COSR was purified by preparative RP-HPLC (Agilent Zorbax SB C_{18} column: 21.2 x 250 mm, 7 μm, gradient: 1-41% B in 80 min, flow rate: 5 mL/min). Preparative RP-HPLC fractions were analyzed using MALDI mass spectrometry, analytical RP-HPLC (Agilent Zorbax SB C_{18} column: 2.1 x 150 mm, 5 μm, linear gradient: 1-51% B over 12 min, flow rate: 0.5 mL/min) and LCMS (Agilent Zorbax 300SB C_3 column: 2.1 x 150 mm, 5 μm, linear gradient: 1-61% B' over 15 min, flow rate: 0.4 mL/min. **For LCMS we used solvent A': 0.1% formic acid in H_2O and solvent B': 0.1% formic acid in acetonitrile. This same solvent system was used for all LCMS performed during the course of this work.** Fractions containing the purified target peptide were combined and lyophilized to give 6.9 mg, 4.0 μmol of product (54% yield). The observed mass was 1374.7 ± 0.1 Da (calculated monoisotopic mass: 1374.71 Da).

Native chemical ligation



KLPETG_5F-COSR (2.2 mg, 1.2 μmol, 4.0 mM) was reacted with CFRALKAA (1.8 mg, 1.5 μmol, 4.7 mM) for 2 hours at room temperature in 6 M guanidine hydrochloride (Gn•HCl), 20 mM tris(carboxyethyl)phosphine hydrochloride (TCEP•HCl), 50 mM 4-mercaptophenylacetic acid (MPAA), 0.2 M potassium phosphate, at pH 6.9. The reaction mixture was quenched with an equal volume of 95 % A: 5 % B. The product KLPETG_5F-CFRALKAA was purified using preparative RP-HPLC (Agilent Zorbax SB C_{18} column: 21.2 x 250 mm, 7 μm, linear gradient: 1-61% B over 120 min, flow rate: 5 mL/min). Preparative RP-HPLC fractions were then analyzed using MALDI mass spectrometry, analytical RP-HPLC (Agilent Zorbax SB C_{18} column: 2.1 x 150 mm, 5 μm, linear gradient: 1-51% B over 12 min, flow rate: 0.5 mL/min) and LCMS (Agilent Zorbax 300SB C_3 column: 2.1 x 150 mm, 5 μm, linear gradient: 1-61% B' over 15 min, flow rate: 0.4 mL/min). Fractions containing the purified target peptide were combined and lyophilized to give 2.0 mg, 1.1 μmol (83% yield). The observed mass for product is 1879.0 ± 0.1 Da (calculated monoisotopic mass: 1879.04 Da).

Protein Studies

Construction of plasmids

pET SUMO His_6-SUMO-eGFP-LPSTGG-His_6 was prepared by use of the Champion pET SUMO protein expression system (Invitrogen, CA). AccuPrime Taq DNA polymerase (Invitrogen, CA) was used to PCR amplify eGFP and the LPSTGG tag was added by using 5'-ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGG-3' (forward) and 5'-CTAATGGTGGTGGTGGTGGTGGCCGCCGGTGCTCGGCAGCTTGTACAGCTCGTCCATGC-3' (reverse) primers. The product was confirmed by 1% (w/v) agarose gel electrophoresis and subsequently PCR purified by the QIAquick PCR purification kit (Qiagen, Netherlands). The PCR product was then cloned into pET SUMO by an overnight ligation at 15 °C with 6 ng PCR product, 25 ng pET SUMO vector, and 0.5 µl T4 DNA ligase. 2 µl of the ligation product was transformed into One Shot Mach1-T1 competent cells and plated on 30 µg/mL kanamycin plates and incubated overnight at 37 °C. Colonies were grown in LB media containing 30 µg/mL kanamycin. The plasmid DNA was isolated using the Qiaprep spin miniprep kit (Qiagen, Netherlands). Glycine at position 2 in SUMO vector was mutated to isoleucine in all model proteins in order to prevent SrtA*-mediated intramolecular cyclization. These mutations were introduced using QuikChange Lightening Multi Site-Directed Mutagenesis Kit (Agilent, CA) with a 5'-GAAGGAGATATACATATGatCAGCAGCCATCATCATCATC-3' primer. The $SrtA_{59\text{-}206}$ construct employed in this study is a truncated version of naturally occurring SrtA in which the N-terminal membrane-spanning domain has been removed. The remaining portion of SrtA exhibits greater solubility in aqueous solution as consequence of membrane-spanning domain excision yet retains full catalytic activity.[3] SrtA* (P94S/D160N/K196T), evolved by Chen et al[4]., was prepared from WT SrtA in a pET 21 vector using a QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent, CA). The following primers were used for the QuikChange: **P94S:** 5'-CCAGGACCAGCAACAtccGAACAATTAAATAGAGGTG-3', **D160N:** 5'-GTATAAAATGACAAGTATAAGAaacGTTAAGCCAACAGATGTAGAAG-3' **K196T:** 5'-GACAGGCGTTTGGGAAaccCGTAAAATCTTTGTAG-3'

Protein Expression

Recombinant His_6-SUMO-LF_N-LPSTGG-His_6, His_6-SUMO-LF_N-DTA(C186S)-LPSTGG-His_5 and His_6-SUMO-Cys-DTA(C186S) were expressed at New England Regional Center of Excellence/Biodefense and Emerging Infectious Diseases (NERCE). His_6-SUMO-eGFP-LPSTGG-His_6 was overexpressed in *E. coli* BL21 (DE3) cells. Approximately 10 g of cell pellet was lysed by sonication in 50 ml of 50 mM Tris, 150 mM NaCl, pH 7.5 buffer containing 30 mg lysozyme, 2 mg DNAase I, and 1 tablet of Roche protease inhibitor cocktail. The suspension was centrifuged at 17,000 rpm for one hour to remove cell debris. The supernatant was loaded onto a 5 ml HisTrap FF crude Ni-NTA column (GE Healthcare, UK) and washed with 50 mL of 40 mM imidazole in 20 mM Tris, 150 mM NaCl, at pH 8.5. The protein was eluted from the column with buffer containing 500 mM imidazole in 20 mM Tris, 150 mM NaCl, pH 8.5. Imidazole was removed from proteins using a HiPrep 26/10 Desalting column (GE Healthcare, UK) into 20 mM Tris, 150 mM NaCl, pH 8.5. Purified proteins were analyzed using an Any kD Mini-PROTEAN TGX Precast Gel (Bio-Rad, CA).

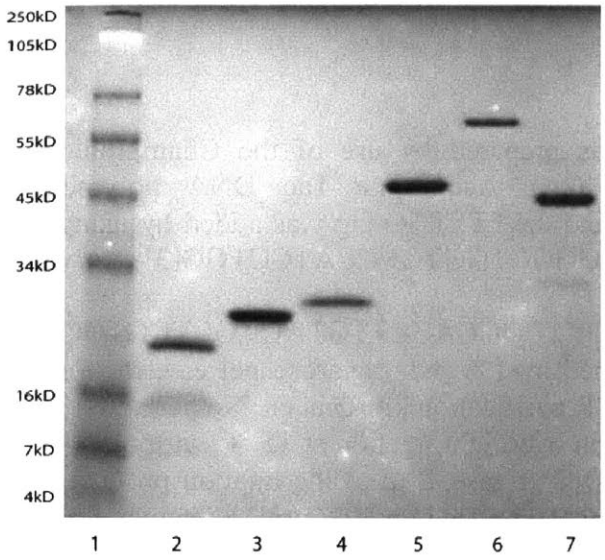


Figure S3. Coomassie stained SDS-PAGE gel of protein constructs. **Lane 1**: Invitrogen Seeblue@plus2 protein standard. **Lane 2**: SrtA*-His_6. **Lane 3**: GB1-His_6-SrtA. **Lane 4**: His_6-SUMO protease. **Lane 5**: His_6-SUMO-LF_N-LPSTGG-His_6. **Lane 6**: His_6-SUMO-LF_N-DTA-LPSTGG-His_5. **Lane 7**: His_6-SUMO-eGFP -LPSTGG-His_6.

In addition, the proteins were analyzed by LCMS to confirm their purity and molecular weight analyzed via high-resolution ESI-QTOF MS (Agilent 6520) (Table S2). The charge-state series of the species were deconvoluted using Agilent MassHunter Bioconfirm using maximum entropy setting (Agilent Zorbax 300SB C_3 column: 2.1 x 150 mm, 5 μm, linear gradient: 5-35% B' over 6 min, 35-45% B' over 20 min, 45-65% B' over 3 min, flow rate: 0.4 mL/min).

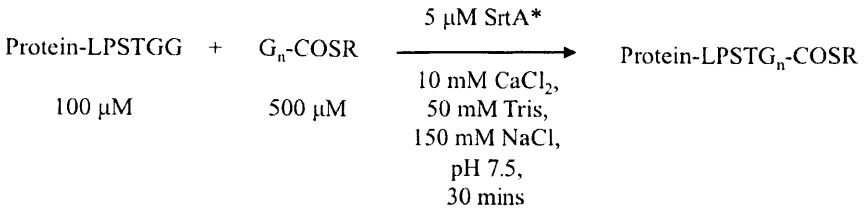
Table S2. Observed molecular masses of protein constructs when analyzed by LCMS.

Protein	Observed MW (Da)	Calculated MW (Da; average)
SrtA*-His_6	19214.6 ± 0.4	19214.5
GB1-His_6-SrtA[3]	25218.5 ± 0.4	25218.2
His_6-SUMO-LF_N-LPSTGG-His_6	45289.3 ± 0.4	45289.8
His_6-SUMO-LF_N-DTA-LPSTGG-His_5	66700.0 ± 0.4	66700.1
Cys-DTA	20840.1 ± 0.4	20840.1
His_6-SUMO-eGFP-LPSTGG-His_6❖	41711.7 ± 0.4	41710.9

❖: The eGFP mass was first calculated from the native primary sequence and then 20 Da was subtracted for peptide backbone cyclization and oxidation to form the chromophore.

Sort-tagging of peptide thioesters to proteins containing C-terminal LPSTGG

All the reactants and products were analyzed via high-resolution ESI-QTOF MS (Agilent 6520) following protocol outlined in the previous section.



His_6-SUMO-Protein-LPSTGG-$His_{6/5}$ proteins (100 μM) were incubated with G_n-COSR (500 μM) with 5 μM SrtA* for 30 min at room temperature (10 mM $CaCl_2$, 50 mM Tris, 150 mM NaCl, pH 7.5). The reaction mixture was then quenched with an equal volume of 50 % A: 50% B. The reaction was analyzed using LCMS (linear gradient: 5-35% B' over 6 min, 35-45% B' over 20 min, 45-65% B' over 3 min, flow rate = 0.4 mL/min). All experiments were performed in triplicate. For all experiments, product formation was observed, but we were unable to determine the yield because the starting material and product co-eluted on LCMS. In addition, similar to peptide studies, for G-COSR, both protein-G-COSR and protein-GG-COSR were observed in products formed.

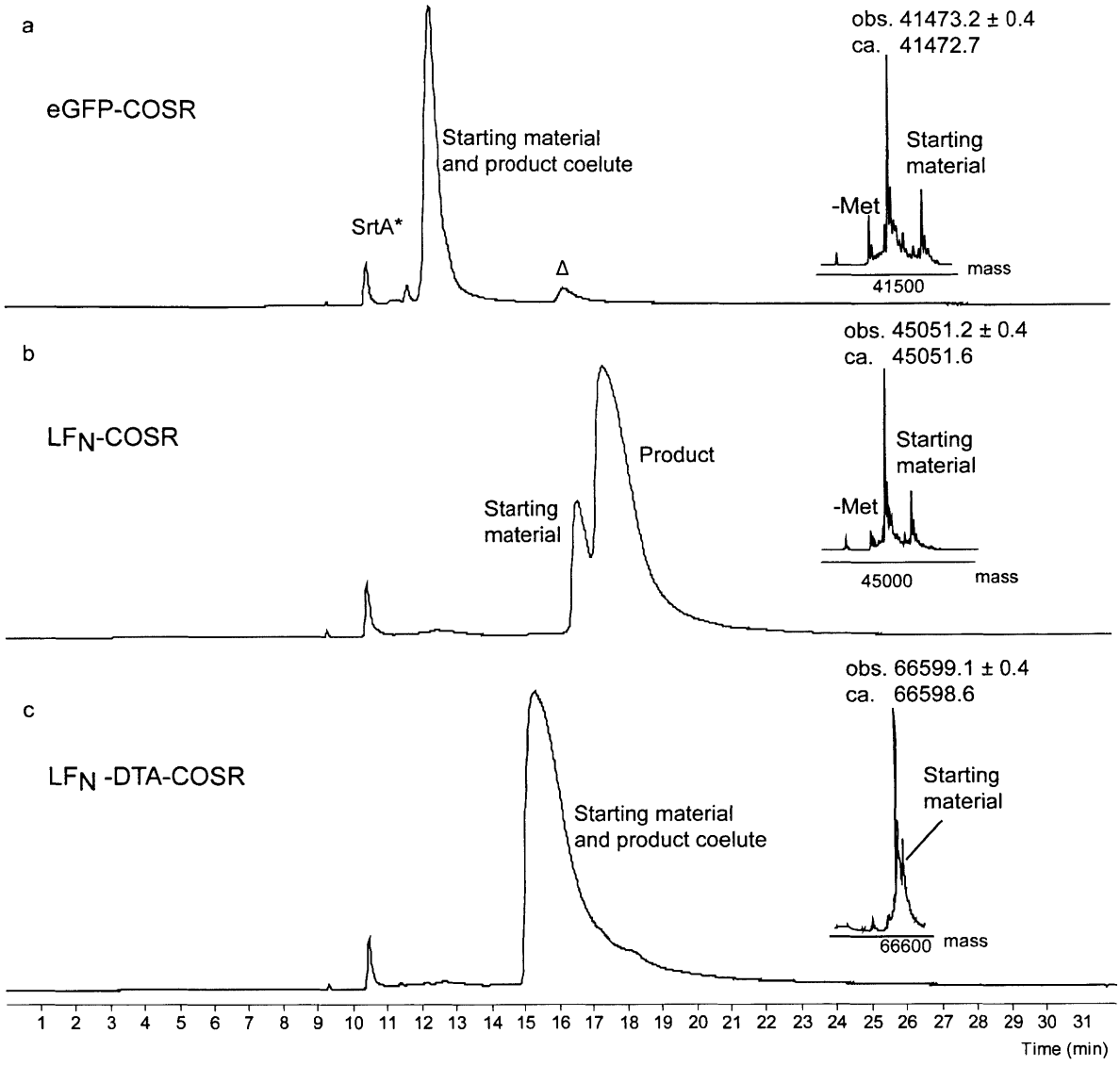
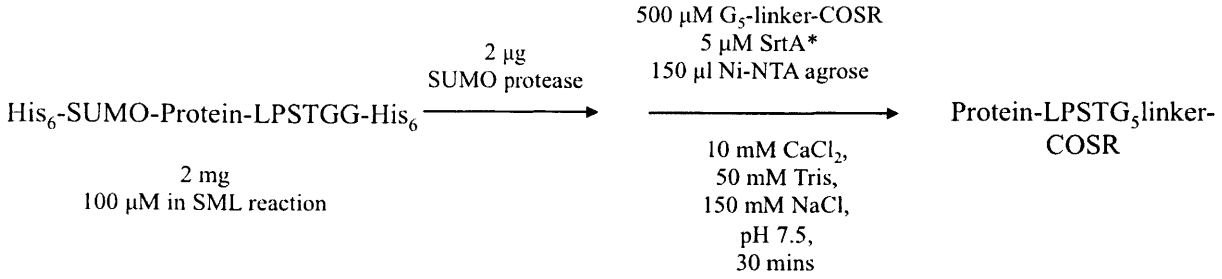


Figure S4. Sample LCMS data for crude SrtA*-mediated ligation (SML) using G_6-COSR with the following model proteins: a) EGFP-COSR, b) LF_N-COSR and c) LF_N-DTA-COSR. The reactions were done in the absence of SUMO protease and Ni-NTA agarose. The total ion current chromatogram is displayed and the inset is the protein mass from the major peak after deconvolution. Δ: impurity in eGFP starting material.

One-pot SrtA* mediated preparation of protein thioesters



SUMO cleavage and SrtA* mediated ligation were conducted using a one-pot method. In particular, His_6-SUMO group on His_6-SUMO-protein-LPSTGG-$His_{6/5}$ was cleaved first by incubating with 1 μg SUMO protease per mg protein at room temperature for 90 mins. The resultant mixture containing protein-LPSTGG-$His_{6/5}$ (100 μM) was reacted with G_5F-COSR, G_5-SGRELER-COSR or G_5-sGreler-COSR (500 μM) in the presence of 5 μM SrtA* and 150 μl of Ni-NTA slurry (Qiagen, Netherlands) for 30 min in buffer containing 10 mM $CaCl_2$, 50 mM Tris, 150 mM NaCl, pH 7.0. The His-tag free, ligated product (protein-LPSTG$_5$-linker-COSR) does not bind to Ni-NTA beads and was obtained by isolation of supernatant after centrifuging Ni-NTA agarose beads at 13,000 rpm for 1 min. Ni-NTA agarose beads were washed 3 times with 20 mM Tris, 150 mM NaCl, pH 7.0 buffer to capture residual product. The product was then concentrated using a Millipore centrifugal filter unit (10K) and washed 3 times with 20 mM Tris, 150 mM NaCl, pH 7.0 buffer to remove unreacted peptide thioesters. Yields are summarized in Table 3. We detected a small amount of thioester hydrolysis when the reactions were conducted at pH 7.5. The hydrolysis was eliminated when the experiments were performed at pH 7.0.

Table S3. Yield for SrtA* mediated synthesis of protein thioesters.

Protein	Thioester used	Starting Material (mg)	Protein-thioester (mg)	Yield (%)
eGFP	G_5F-COSR	4.2	2.3	81
LF_N-DTA	G_5F-COSR	2.0	0.7	40
LF_N	G_5F-COSR	2.3	0.6	50
LF_N	G_5-L-linker-COSR	2.3	0.7	50
LF_N	G_5-d-linker-COSR	4.6	1.4	40

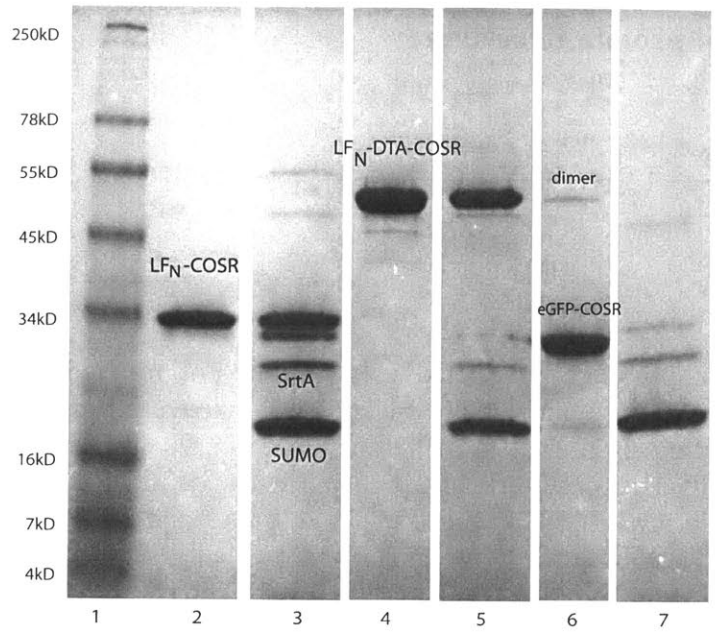


Figure S5. Coomassie stained SDS-PAGE analysis of SrtA* mediated ligation in the presence of Ni-NTA agarose beads and SUMO protease. **Lane 1**: Invitrogen Seeblue@plus2 protein standard. **Lane 2, 4, 6**: LF_N-G_5F-COSR, LF_N-DTA-G_5F COSR and eGFP- G_5F-COSR. **Lane 3, 5, 7**: elute using 500 mM imidazole in 20 mM Tris, 150 mM NaCl, pH 8.5.

Native chemical ligation with eGFP thioester

eGFP-COSR	+	CFRALKAA	⟶	eGFP-CFRALKAA
0.81 mg 0.028 μmol 28 μM		0.5 mg 0.4 μmol 4 mM	20 mM TCEP•HCl, 10 mM MPAA, 0.2 M phosphate, pH 7.0, 3 hours	0.80 mg 0.028 μmol 98% yield

eGFP-LPSTG$_5$F-COSR (0.81 mg, 0.028 μmol, 28 μM) was reacted with CFRALKAA (0.5 mg, 0.4 μmol, 4 mM) for 3 hours in 20 mM TCEP•HCl, 10 mM MPAA, 0.2 M phosphate, pH 7.0. After 3 hours, the reaction mixture was concentrated with a Millipore centrifugal filter unit (10K) and washed three times with buffer (20 mM Tris, 150 mM NaCl, pH 7.0) to remove small molecule reactants (0.80 mg, 0.028 μmol, 98% yield).

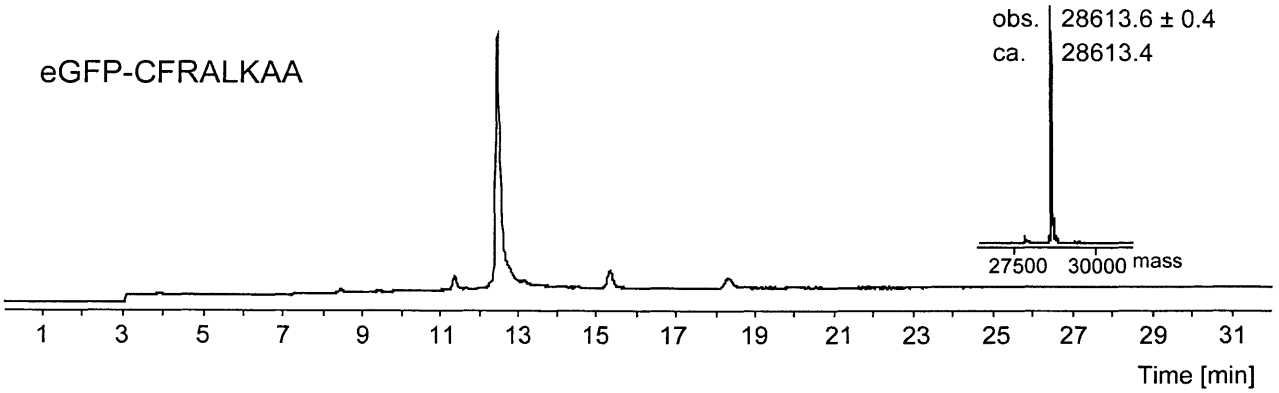
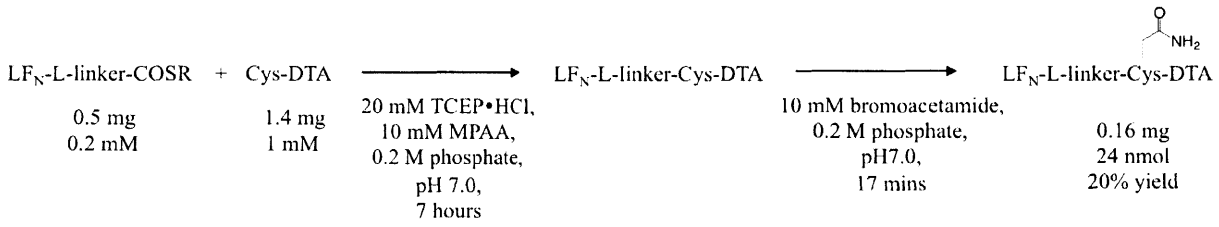


Figure S6. LCMS data for native chemical ligation products of EGFP-CFRALKAA. The total ion current chromatogram is displayed and the inset is the protein mass from the major peak after deconvolution.

Native chemical ligation with LF_N-L-linker-COSR and LF_N-D-linker-COSR to Cys-DTA



LF_N-SGRELER-COSR (0.5 mg, 15 nmol, 0.2 mM) was reacted with Cys-DTA (1.4 mg, 66 nmol, 1 mM) for 7 hours in 20 mM TCEP•HCl, 10 mM MPAA, 0.2 M phosphate, pH 7.0. After 7 hours, the reaction mixture was diluted 5 times with 0.2 M phosphate buffer and 10 mM bromoacetamide was added to alkylate Cys on LF_N-L-linker-Cys-DTA. The alkylation reaction was quenched with 100 mM sodium 2-mercaptoethanesulfonate after 17 mins. LF_N-L-linker-DTA was isolated from the reaction mixture using HiLoad 26/600 Superdex 200 prep grade size exclusion chromatography column (GE Healthcare, UK) in 20 mM Tris, 150 mM NaCl, pH 7.5 buffer. Fractions containing pure product were concentrated with a Millipore centrifugal filter unit (10K). 0.16 mg of LF_N-L-linker-DTA (24 nmol, 20% yield) was obtained.

The D-peptide analogue LF_N-sGreler-COSR (0.6 mg, 18 nmol, 0.2 mM) was reacted with Cys-DTA (1.4 mg, 66 nmol, 1 mM) under the same conditions as the L-peptide variant presented above. 0.065 mg of LF_N-D-linker-DTA (15.8 nmol, 7% yield) was obtained after size exclusion chromatography purification.

Through monitoring in LCMS, we observed near quantitative NCL reaction (Figure S7). The unexpected low isolated yield has been attributed to loading a very small amount of protein on size exclusion column with 320 ml capacity, which lead to loss of product during purification.

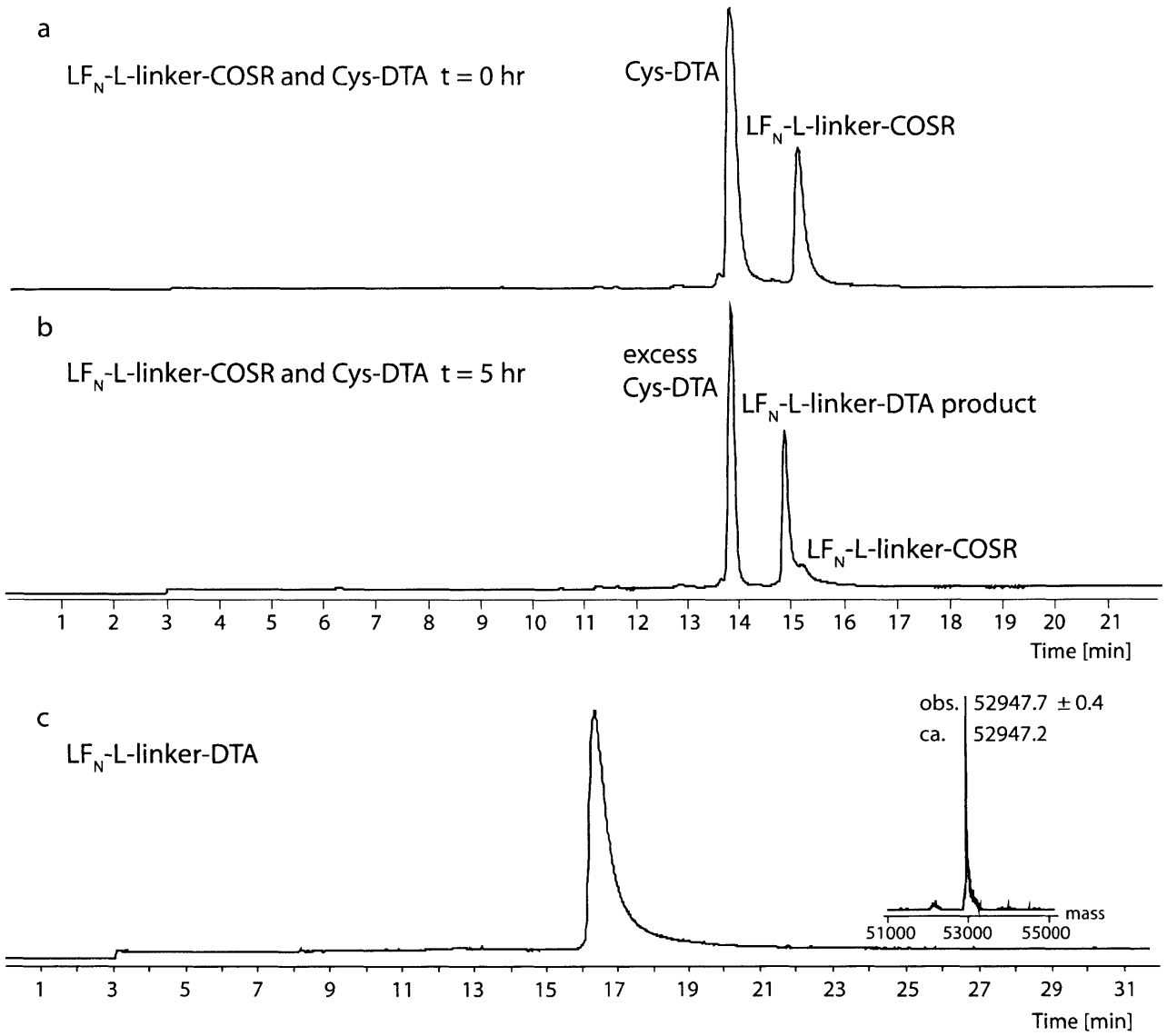


Figure S7. LCMS data for native chemical ligation reaction between LF_N-L-linker and Cys-DTA. **(a)** at t = 0 hr. **(b)** at t = 5 hr. **(c)** purified alkylated LF_N-L-linker-DTA. The total ion current chromatogram is displayed. LCMS conditions: for (a) and (b), linear gradient: 5-65% B' over 15 min, flow rate = 0.4 mL/min. for (c), linear gradient: 5-35% B' over 6 min, 35-45% B' over 20 min, 45-65% B' over 3 min, flow rate: 0.4 mL/min. The NCL reaction between LF_N-L-COSR and Cys-DTA is 80% complete after 5 hours. The NCL reaction for LF_N-D-linker-DTA precedes the same as LF_N-F-linker-DTA.

Evaluation of LPSTGG hydrolysis at varying concentrations of G_5 nucleophile

To investigate the effect of G_5 nucleophile concentration on the extent of hydrolysis byproduct formation during the sortagging reaction, eGFP-LPSTGGHis$_6$ (50 μM) was incubated with G_5LRL (0, 10, 30, 50, 100, 300 or 500 μM) and 7.35 μM SrtA* for 30 min at room temperature (10 mM $CaCl_2$, 50 mM Tris, 150 mM NaCl, pH 7.5). The reaction mixture was then quenched with an equal volume of 50 % A: 50% B and analyzed via LCMS (linear gradient: 1-61% B'

over 15 min, flow rate = 0.4 mL/min). Formation of a hydrolysis byproduct was noticeable at G_5 concentrations lower than 300 μM. Selected LCMS traces are shown in Figure S8.

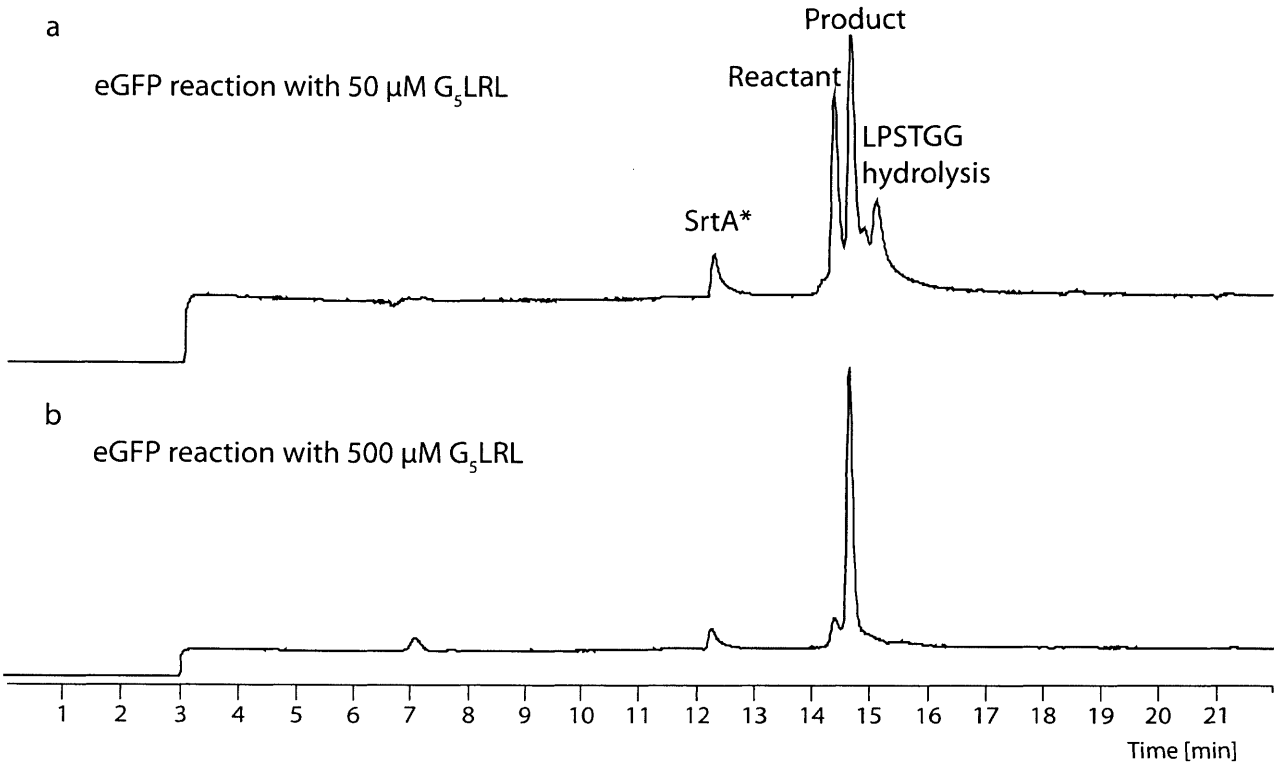


Figure S8. LCMS data for hydrolysis analysis at varying G_5 nucleophile concentrations. **(a)** 50 μM G_5LRL. **(b)** 500 μM G_5LRL. Both 300 and 500 μM G_5 nucleophile showed minimal hydrolysis byproduct and proceeded with high yield. The traces above show total ion current chromatograms.

Translocation of LF_N-DTA-COSR, LF_N-L-linker-DTA and LF_N-D-linker-DTA into cells using anthrax protective antigen

LF_N-DTA, LF_N-L-linker-DTA, LF_N-D-linker-DTA and LF_N-DTA-LPSTG$_5$F-COSR were prepared at varying concentrations in 50 μL of F-12 media supplemented with 10 % (v/v) fetal bovine serum, then 50 μL of F-12 media containing 10 nM protective antigen (PA_{83}) was added and mixed. Then 100 μl samples were added to CHO-K1 cells in 96-well plate and incubated for 30 minutes at 37 °C and 5 % CO_2. The medium was then removed, washed 3 times with PBS, replaced with leucine-free F-12 medium supplemented with 1 μCi/ mL ^{3}H-Leu, and incubated for 1 hr at 37 °C 5% CO_2. The cells were washed 3 times with PBS and then 150 μL of scintillation fluid was added. The inhibition of protein synthesis by LF_N-DTA was determined by measuring ^{3}H-Leu incorporation into cellular proteins. The scintillation counts were normalized to control value with PA only in the media. Each experiment was done in triplicate. The data were then plotted using Origin8 using Sigmoidal Boltzmann Fit using equation $y = A2 + \frac{A1-A2}{1+e^{\frac{x-x_0}{dx}}}$ where x_0 represents the $\log EC_{50}$ values.

References

(1) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. *Int. J. Pept. Protein Res.* **1992**, *40*, 180.

(2) Hackeng, T. M.; Griffin, J. H.; Dawson, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10068.

(3) Kobashigawa, Y.; Kumeta, H.; Ogura, K.; Inagaki, F. *J. Biomol. NMR* **2009**, *43*, 145.

(4) Chen, I.; Dorr, B. M.; Liu, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 11399.

Appendix 2 Publication List

Woodham AW, Cheloha RW, Ling J, Rashidian M, Kolifrath SC, Mesyngier M, Duarte JN, Bader JM, Skeate JG, Da Silva DM, Kast WM, Ploegh HL. Nanobody-antigen conjugates elicit HPV-specific anti-tumor immune responses. Pending publication in Cancer Immunology Research.

Fang, T., Van Elssen, C., Duarte, J.N., Guzman, J.S., Chahal, J.S., Ling, J., and Ploegh, H.L. (2017). Targeted antigen delivery by an anti-class II MHC VHH elicits focused alphaMUC1(Tn) immunity. Chemical science 8, 5591-5597.

Chahal, J.S., Fang, T., Woodham, A.W., Khan, O.F., Ling, J., Anderson, D.G., and Ploegh, H.L. (2017). An RNA nanoparticle vaccine against Zika virus elicits antibody and CD8+ T cell responses in a mouse model. Scientific reports 7, 252.

Hagiwara, M., Ling, J., Koenig, P.A., and Ploegh, H.L. (2016). Posttranscriptional Regulation of Glycoprotein Quality Control in the Endoplasmic Reticulum Is Controlled by the E2 Ub-Conjugating Enzyme UBC6e. Molecular cell 63, 753-767.

Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. Nature 531, 395-399.

Fang, T., Duarte, J.N., Ling, J., Li, Z., Guzman, J.S., and Ploegh, H.L. (2016). Structurally Defined alphaMHC-II Nanobody-Drug Conjugates: A Therapeutic and Imaging System for B-Cell Lymphoma. Angewandte Chemie (International ed in English) 55, 2416-2420.

Spokoyny, A.M., Zou, Y., Ling, J.J., Yu, H., Lin, Y.S., and Pentelute, B.L. (2013). A perfluoroaryl-cysteine S(N)Ar chemistry approach to unprotected peptide stapling. J Am Chem Soc 135, 5946-5949.

Ling, J.J., Policarpo, R.L., Rabideau, A.E., Liao, X., and Pentelute, B.L. (2012). Protein thioester synthesis enabled by sortase. J Am Chem Soc 134, 10749-10752.