UNIVERSITY OF CALIFORNIA
RIVERSIDE

Quality Control to Mistargeted Endoplasmic-Reticulum Proteins in the Cytosol

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Khanh Kieu Nguyen

June 2020

Dissertation Committee:
Dr. Joseph Genereux, Chairperson
Dr. Min Xue
Dr. Yinsheng Wang

The Dissertation of Khanh K. Nguyen is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

My graduate study at UC Riverside has been valuable experience for me in building up my scientific career. Words can't express how grateful I am to everyone that I receive support and guidance from during my study.

First and foremost, I would like to show my appreciation to my dissertation advisor, Prof. Joseph Genereux. He provided me an opportunity to be part of graduate study at UCR. Furthermore, during my five years of graduate studies, he has been an enthusiastic and motivating advisor to me in guiding my research projects, training me the basic analytical lab techniques, teaching me the concepts of the new chemical and biological background, revising my drafts of my posters, my manuscript and my dissertation, giving me feedbacks on all of my presentations, and supporting me in my job search. Overall, I have learned a lot from him to become a resourceful research scientist. I could not have achieved what I'm currently having without your support.

Additionally, I would like to thank my committee members, Prof. Min Xue and Prof. Yinsheng Wang for serving in my dissertation and giving me advice on my projects as well as feedbacks on my dissertation. Another support system that is crucial during my graduate years is from the other members of the Genereux lab including Dr. Mateo Espinoza, Maureen Montoya, Guy Quanrud, Ziqi Lyu,

Melody Sycks, and Dr. Liangyong Mei in their technical and mental support. I also appreciate help and advice in classes, practice oral exam, and practice talks from other peers in the department, including Dr. Yana Lyon, Michael Trinh, Priyanka Sarkar, and many more chemistry graduate students.

Special thanks to my fiancé, Dr. Mateo Espinoza, who has been constantly there for me when I was having troubles with the experiments or with other issues. Without his great love, care, company, sense of humor, and intense support, I would not have made this far. Last but not least, I would like to thank my own family, my parents and my younger brother, who showed a lot of emotional support during my challenging years of graduate school.

The work in Chapter 2 was done in collaboration with Dr. Mateo Espinoza and guided by Dr. Joseph Genereux. The work in Chapter 4 was done with the help of Dr. Rong Hai and Stephanie Thurmond from Department of Microbiology at UC Riverside and guided by Dr. Joseph Genereux.

The text and figures in Chapter 4, in part or full, are reprint of the following publication:

Nguyen, K.K., Thurmond, S., Hai, R. Genereux, J. Protein profiling and pseudo-parallel reaction monitoring to monitor a fusion-associated conformational change in hemagglutinin. *Anal. Bioanal. Chem.* **411**, 4987-4998 (2019)

DEDICATION

I dedicate this dissertation to my family.

ABSTRACT OF THE DISSERTATION

Quality Control to Mistargeted Endoplasmic-Reticulum Proteins in the Cytosol

by

Khanh Kieu Nguyen

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, June 2020
Dr. Joseph Genereux, Chairperson

Proteins play major roles and participate in multiple pathways in the cell. Thus, cells have to rely on proper folding and localization of the proteome to avoid dysfunction. One of the key steps in maturation for secretory proteins is translocation into the endoplasmic reticulum (ER) rather than mistargeting to and accumulating in the cytosol. However, under ER stress, ER translocation becomes less efficient leading to higher accumulation of proteins in the cytosol. This potentially poses a threat to the cytosol. In our study, we identified an ER chaperone HSPA13 as a regulatory factor of protein translocation into the ER. Overexpression of HSPA13 in HEK293T cells impairs ER and cytosolic proteostasis by inhibiting ER translocation of substrates, disrupting their secretory pathway and promoting their intracellular aggregation. Additionally, we found that mistargeted protein in the cytosol is triaged on the basis of its stability. Highly destabilized mistargeted proteins are rapidly processed and degraded by the proteasome, while more stable proteins are degraded less efficiently.

Furthermore, in the case of transthyretin, the processing is mediated by selective N-terminal proteolysis. Thus, we have found that mistargeted ER-directed proteins are subject to quality control in the cytosol.

Influenza infection requires viral escape from early endosomes into the cytosol, which is enabled by an acid-induced irreversible conformational transformation in the viral protein hemagglutinin. Despite the direct relationship between this conformational change and infectivity, label-free methods for characterizing this and other protein conformational changes in biological mixtures are limited. While the chemical reactivity of the protein backbone and side-chain residues is a proxy for protein conformation, coupling this reactivity to quantitative mass spectrometry is a challenge in complex environments. Herein, we developed electrophilic amidination coupled with reaction monitoring to detect the fusion-associated conformational transformation in recombinant hemagglutinin (rHA). We identified rHA peptides that are differentially amidinated between the pre- and post-fusion states, and validated that this difference relies upon the fusion-associated conformational switch. We further demonstrate that we can distinguish the fusion profile in a matrix of digested cellular lysate.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Protein Folding

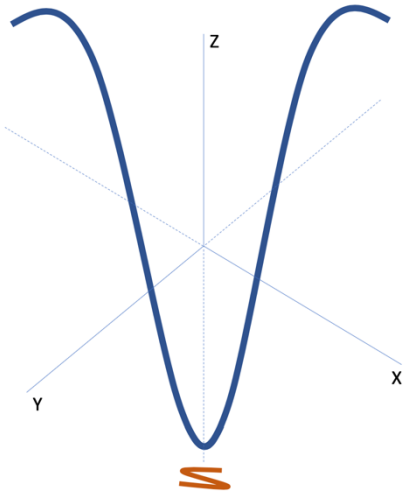
According to the central dogma of molecular biology, DNA contains the genomic information that is ultimately translated to make proteins (Crick 1970). While DNA is a linear sequence of four different nucleotides, proteins are made of up to twenty common amino acids in three-dimensional structures. Protein molecules are typically several hundreds of amino acids long and may contain more than one folding domain. Additionally, they usually have at least one intrinsically disordered domain so that they can interact with other proteins (Dyson & Wright 2005). *In vivo*, proteins can fold spontaneously at low concentrations and at the temperatures in *in vitro* experiments (Anfinsen 1973); however, protein folding is more challenging inside the cells due to a crowded environment (Ellis & Hartl 1999, Ellis & Minton 2006). It was explained that macromolecular crowding can affect folding capacity, and the mobility of the polypeptide chain is hindered inside the cell (Hong & Gierasch 2010, Gershenson & Gierasch 2011). The physicochemical properties of the side chains of amino acids such as hydrophobicity, size, and charge lead to different native states of proteins. For example, in a properly folded protein, hydrophobic residues are shielded while hydrophilic residues are exposed on the surface protein. Moreover, the protein structures are stabilized by intramolecular electrostatic and hydrophobic interactions (Anfinsen 1972). Free energy diagrams can demonstrate the kinetics

of protein folding (Dill & Chan 1997). As shown in Figure 1.1.A, a smooth folding energy funnel represents the possible conformations of a polypeptide chain positioned on the X- and Y-axis, while the energy level is represented by the Z-axis. The higher the intramolecular interactions of the intermediate states are, the closer it is to the native conformation of the protein, which minimizes its free energy. This folding funnel demonstrates that there are several pathways for reaching the native folded states. In Figure 1.1.B, we can observe a rougher folding free energy landscape, which illustrates the kinetic barriers which populate partially folded intermediates. These intermediates are likely to misfold and become aggregates (Dobson 2004). The folded state of a protein is in dynamic equilibrium with its unfolded state during its life (Shea & Brooks 2001, Hartl & Hayer-Hartl 2009).

Misfolded proteins lead to aggregation and neurodegenerative diseases

Proteins in their native conformations are in good shape and become functional. However, at times, the ribosome makes mistakes in translation that leads to proteins that are inherently unstable and even toxic to the cells. The native conformations of proteins are composed of mostly alpha-helixes (Pauling 1951). On the other hand, misfolded proteins acquire a motif known as beta-sheets, which expose hydrophobic amino acids and could promote aggregation and amyloid fibrils (Sunde & Blake 1998). Besides amyloid formation from beta-sheet

A



B

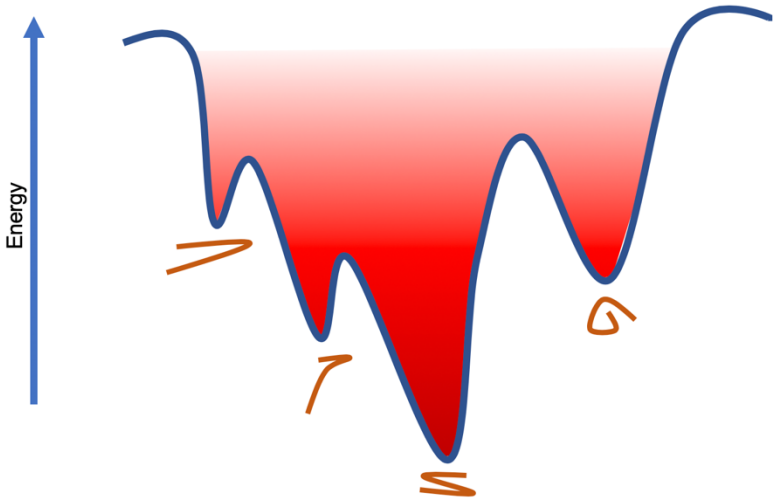


Figure 1.1: Perspective of energy landscape in protein folding. Individual positions along the X- and Y-axis represent different protein conformations and the Z-axis corresponds to the internal free energy of protein folding. A. A smooth folding funnel reveals several pathways that a polypeptide chain can adopt to achieve the native state of a protein. B. A rougher folding free energy landscape also demonstrates numerous possible pathways for a native conformation of a protein; however, occasionally the polypeptide chain can get kinetically trapped at their intermediate conformations.

structures, partially folded intermediates can form amorphous aggregates as well (Khurana 2001). Protein aggregates are involved in several diseases such as Alzheimer's disease, Huntington's disease, prion diseases, amyotrophic lateral schlerosis, etc. Different protein aggregates are associated with different types of neurodegenerative diseases (NDs). For instance, Alzheimer's disease (AD) involves extracellular aggregates of amyloid precursor protein and tau protein while mutations of superoxide dismutase (SOD1) causes amyotrophic lateral schlerosis (ALS) (Ross & Poirier 2004). These diseases are terminal and untreatable, leading to the permanent and fatal damage of the central nervous system (Selkoe 2003).

Noticeably, the soluble intermediate oligomers are the most causative factor of neurodegeneration (Chromy 2003, Sengupta 2016). For example, formation of Aβ oligomers can lead to pathogenesis of AD by interfering with communication between neurons by a mechanism that involves NMDA-type glutamate receptors (NMDARs). NMDAR-calcium influx into the spine head is reduced by Aβ oligomers (Shankar 2007). Generally, ND is observed in aging process, in which elderly people are susceptible to neurodegeneration (Hou 2019). Additionally, there are several rare NDs that are associated with misfolded proteins such as prion disease or familial amyloid polyneuropathy (FAP). Prion disease manifests by the replication of the infectious prion by recruiting the normal prion proteins (PrP^{C}) into the pathological scrapie-causing conformation (PrP^{SC}) (Aguzzi &

Polymenidou 2004). Homotetrameric transthyretin (TTR), a protein model for our study, can form amyloid and oligomers that are toxic to neurons and lead to FAP (Planté-Bordeneuve & Said). Different mutations of TTR can contribute to the progression of FAP differently, which will be mentioned in more details in Chapter 3.

Proteostasis in eukaryotic cells

Protein homeostasis, or proteostasis, is the response of the cells to accumulation of misfolded proteins. Proteostasis network comprises of integrated functions that maintain the stability of the proteome (Balch 2008, Powers 2009). Although the systems are compartment-specific, they can work together or compete to preserve the health of the cells against any insults. Proteostasis is extremely important because of the equilibrium of folded and unfolded state, the weak intramolecular forces stabilizing the folded structure, and the ability to access misfolded and aggregating proteins when mistakes are made. For instance, aging decreases proteostasis network capacity that results in protein misfolding and aggregation (Kaushik & Cuervo 2015). In this chapter, we introduced the factors involved in proteostasis including cellular chaperones, stress responses and degradation systems.

Major Molecular Chaperones

Molecular chaperones are defined as proteins that interact with, stabilize and help other proteins fold correctly (Hartl 1996). Several different classes of chaperones exist in the cell that form cooperative networks. These chaperones are often known as heat-shock proteins (HSPs) since they are upregulated under stress conditions in which the amounts of aggregation-prone folding intermediates increase. HSPs are classified based on their molecular weights (HSP40, HSP70, HSP90 and the small HSPs) (Hartl 2011). They play several roles in maintaining the proteostasis of the cell including de novo protein folding, refolding stress-denatured proteins, dissociating protein aggregates and directing toxic materials to proteolytic degradation. in the manner of ATP- and cofactor-regulated cycles.

The HSP70 system

HSP70 consists of an N-terminal binding domain (NBD) and C-terminal substrate binding domain (SBD) connected by a highly conserved hydrophobic linker region. Chaperones interact with hydrophobic region of the partially folded proteins to prevent aggregation. These substrates were delivered to ATP-bound HSP70 by one of several HSP40 cofactors. Hydrolysis of ATP to ADP triggers the close conformation of HSP70 in which the α-helical lid of the SBD closes over the bound substrate peptide (Figure 1.2). After ATP hydrolysis, a nucleotide-exchange factor (NEF) binds to the ATPase domain on HSP70 to release ADP

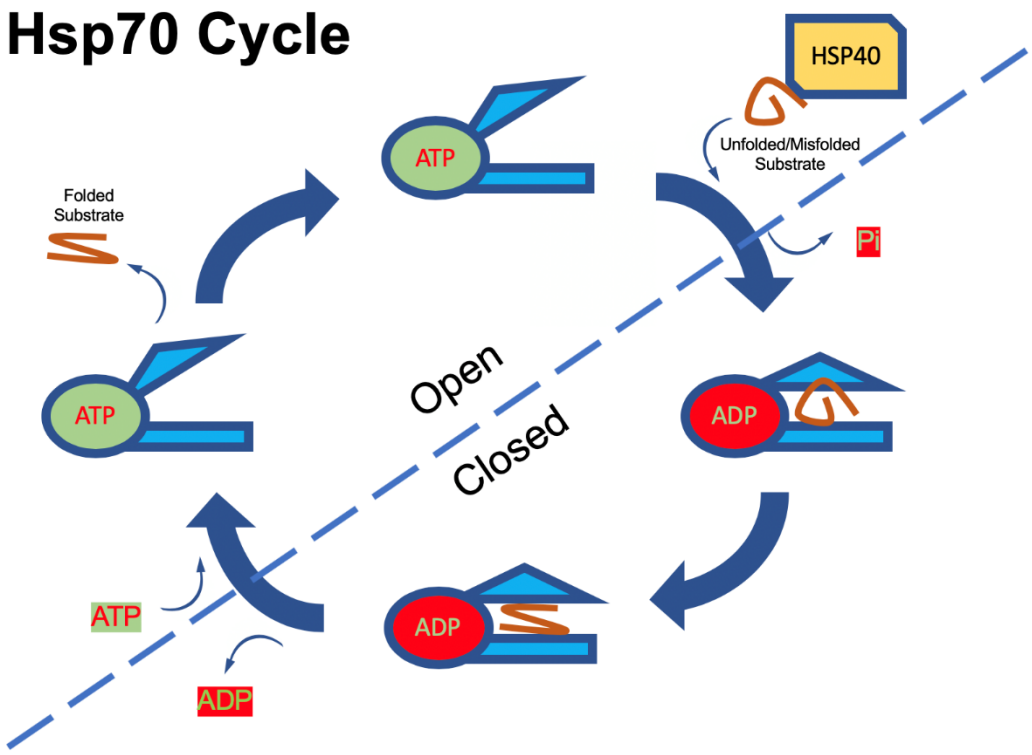


Figure 1.2: The HSP70 chaperone cycle. Misfolded and unfolded protein substrates are delivered to ATP-bound HSP70 by one of HSP40 cofactors. The hydrolysis of ATP results in closed conformation of HSP70. Then, dissociation of ADP catalyzed by one of NEFs leads to recycling of HSP70.

from the NBD. ATP then binds to HSP70 to release the substrate and HSP70 achieves its close conformation again. Substrate release allows the fast-folding proteins to fold and bury hydrophobic regions, while proteins that take longer to fold will rebind to HSP70 to avoid aggregation. If the proteins still fail to fold after dissociating from HSP70, they can be transferred to downstream chaperones or be directed for degradation (Mayer 2010, Kampinga & Craig 2010).

HSP40 (J protein) and NEF cochaperones regulate the HSP70 cycle. The HSP40 proteins contain a J-domain, which binds to the N-terminal ATPase domain on HSP70 and the adjacent linker region. HSP40s function as cochaperones and recruit HSP70 to misfolded proteins (Mayer 2010, Kampinga & Craig 2010). This interaction facilitates the ATP hydrolysis to ADP, resulting in a tight substrate binding by HSP70 in its close conformation. The binding of NEF on the NBD domain of HSP70 catalyzes the exchange of ADP to ATP, promoting the opening of SBD and substrate release. The most abundant eukaryotic NEFs are the HSP110s, whose structures are related to HSP70. These NEFs synergize with HSP70 and HSP40 for protein disaggregation. Proper regulation of the ATPase activity of HSP70s is crucial for their functions. Although HSP40s and NEFs are present at relatively lower levels compared to HSP70, they cooperate to accomplish the aids in protein folding and also establish substrate specificities for HSP70 machinery. There are diseases associated with mutated cochaperones. For example, the inactive mutations on NEF Sil1 disrupt its

interaction with HSPA5 and cause the neurodegenerative disease Marinesco-Sjögren syndrome, which indicated the importance of NEF in maintaining proteostasis (Senderek 2005, Anttonen 2005).

The HSP90 system

The HSP90 system has a major role in cell regulation. Their substrates include misfolded proteins and multiple signaling molecules that are transferred from HSP70 chaperones. HSP90 consists of three domains: the highly conserved N-terminal domain, the middle domain, and the C-terminal domain (Mayer 2010). The N-terminal domain (ND) contains the ATPase binding site, whose mutagenesis and inhibitors can disrupt the binding site for the unfolded proteins. The middle domain (MD), composed of α-β-α motif, plays important role in substrate binding activity and regulation of ATP hydrolysis. The C-terminal domain (CD) of HSP90s contains a pair of helixes that can assemble into a four-helix bundle. Similar to HSP70, the HSP90 dimer interacts with the substrates through the ATP-regulated cycle followed by its structural change. In particular, the open shaped HSP90 receives the inactive substrate and then closes. This activity is driven by the combined effects of ATP binding, ATP hydrolysis, posttranslational modification and the interactions with other cochaperones. Subsequently, the NDs dimerize to form a molecular clamp of the structure with twisted subunits. After hydrolysis, the N termini dissociate to release the client protein (Hartl 2011).

There are various cochaperones that regulate this cycle. The negative regulators of this cycle include CDC37 and HOP. CDC37, which delivers certain kinase substrates to HSP90, inhibits the ATPase activity. HOP inhibits N-terminal dimerization. The factors that adjust kinetic properties of the cycle for HSP90s to obtain their conformational change include AHA1 and p23. AHA1 promotes ATP hydrolysis, whereas p23 stabilizes the ATP-bound state (Hartl 2011).

Along with molecular chaperones that facilitate refolding proteins into their native states, the cells develop their responses to degrade the toxic materials when the proteostasis is unbalanced. The most prominent responses include the integrated stress response, the unfolded protein response and the heat shock response.

The Unfolded Protein Response in the endoplasmic reticulum (ER)

Nearly one-third of the proteome enters the ER secretory compartment (Schröder & Kaufman, 2005; Ron & Walter 2007). Once the proteins enter the ER, they pass through the Golgi or the lysosome to get to the extracellular space by membrane-bound vesicles. Also, the environment of the ER is quite distinct. For instance, most of the secreted proteins from the ER are cotranslationally glycosylated at Asn residues (N-linked glycosylation), which is important for their folding efficiency and functions (Culyba 2011). Additionally, the ER is highly oxidizing, which allows the oxidation of cysteine residues into the stable disulfide

bridges of the protein (Lindquist & Kelly 2011). The key factors that maintain proteostasis in the ER include the ER versions of HSP70 and HSP90 chaperone families, the lectin chaperones calnexin and calreticulin that helps N-linked glycoproteins to fold (Williams 2006), the glycosidases that target misfolded proteins to ER-associated degradation (ERAD) (Williams 2006), and protein disulfide oxidases that catalyze disulfide formation (Schröder & Kaufman, 2005; Ron & Walter 2007).

If the influx of unfolded proteins exceeds the capacity of the ER, the physiological state of the ER will be disrupted. Under this condition, the signaling pathways, termed the unfolded protein response (UPR), are activated to relieve the stress in the ER. The pathway includes three response regulators: PERK, IRE1, and ATF6. These sensors have one domain in the ER and another in the cytosol (Schröder & Kaufman 2005; Ron & Walter 2007). As misfolded proteins populate the ER, GRP78 chaperones are recruited to bind to the misfolded proteins from those three sensors; thus, GRP78 also plays a role as a negative regulator for this pathway.

ATF6 (activating transcription factor 6) is a transmembrane domain protein encoding a basic leucine zipper (bZIP) transcription factor in its cytosolic domain (Haze 2017). Once ATF6 is activated by the accumulation of unfolded proteins, it translocates to the Golgi complex and gets cleaved by metalloprotease site-2

protease (S2P) within the phospholipid bilayer (Chen 2002, Ye 2000). This cleavage results in the release of the cytosolic fragment of ATF6, which then translocates to the nucleus to activate the transcription of genes of the ER chaperones (Kokame 2001, Yoshida 2001).

The activation of PERK (protein kinase-like ER kinase) phosphorylates the eukaryotic translation initiation factor 2 (eIF2α) and the transcription factor Nrf2. Phosphorylated eIF2α results in attenuation of protein translation to reduce the misfolding and aggregation load on the ER proteostasis system. Also, phosphorylation of eIF2α favors the translation of mRNA encoding activating transcription 4 (ATF4), which translocates to the nucleus to induce expression of ER chaperones such as PDI (Halperin 2014). Nrf2 and its homologue Nrf1 are the transducers of the UPR to the nucleus to activate the target genes (Pajares 2017).

IRE1 (inositol requiring kinase 1) is a type I transmembrane kinase endoribonuclease, consisting of an ER-luminal dimerization and cytosolic kinase and endoribonuclease domain (Schroder and Kaufman 2005). When activated, IRE1 cleaves the mRNA encoding a UPR-specific transcription factor XBP1. The spliced mRNA is translated to the active form of the transcription factor (XBP1s). XBP1s appears to have a role in regulating lipid synthesis as well as ERAD components (Walter & Ron 2011).

Cumulatively, to remedy the perturbations of the ER proteostasis due to misfolded proteins, the folding demand has to be decreased and the folding capacity has to be increased. To decrease the folding demand, the transcription of genes and protein translation are downregulated; also, the clearance of misfolded proteins through ERAD is accelerated. To increase the folding capacity, the molecular chaperones in the ER and the folding catalysts are upregulated. The ER also enlarges in order to dilute the concentration of the unfolded proteins by increasing lipid synthesis (Schröder & Kaufman 2005).

The Heat Shock Response

The heat shock response (HSR) signaling pathway can sense the accumulation of misfolded proteins or aggregates in the cytosolic or nuclear compartments. When the pathway gets triggered, the load on proteostasis is alleviated due to the halt of nonessential protein synthesis and degradation of mRNA. The HSR is regulated by HSP90 binding to the transcription factor heat shock factor-1 (HSF1). The HSR is activated by the dissociation of HSP90 from HSF1 to tackle misfolding and/or aggregation, which allows the HSF1 transcription factor to trimerize and be phosphorylated. The active HSF1 initiates the transcription of heat shock element (HSE) (Åkerfelt 2010). Phosphatases and acetyl transferase enzymes act as negative regulators of the HSR.

The Integrated Stress Response

The integrated stress response (ISR) is an elaborate signaling pathway present in eukaryotic cells, in response to various stress stimuli. While UPR senses the stress of misfolded proteins in the ER and HSR operates in cytosol, ISR is a central and conserved signal network that responds to stress in the ER lumen as well as in the cytosol (Costa-Mattioli & Walter 2020). The common downstream factor for all perturbations is phosphorylation of the α-subunit of the eIF2α on serine 51 (Ron 2002). Interestingly, this is catalyzed by four different serine/threonine (S/T) eIF2α kinases that are activated by distinct stress stimuli (Wek 2006, Donnelly 2013). Consequentially, eIF2α phosphorylation attenuates protein synthesis and promotes translation of selected genes including (ATF4) to aid in cell recovery. ATF4 is the core effector of ISR. Activated ATF4 can form homo- and heterodimers and bind to DNA targets to promote expression of genes to adapt to cellular stress (Ameri & Harris 2008).

Several different kinases can phosphorylate eiF2α, with each activated by one or more cellular stress pathways. PKR-like ER kinase (PERK) is located in the endoplasmic reticulum (ER) where its luminal domain binds to 78-kDa glucose-regulated protein (GRP78). PERK is activated during ER stress due to accumulation of misfolded and unfolded proteins in the ER, or disruption in calcium homeostasis, cellular energy, etc. Two models of activating PERK have been proposed. The classical model suggests that misfolded proteins recruit

GRP78 away from PERK, which leads to its autophosphorylation and activation (Korennykh & Walter 2012). The more recent studies propose that PERK is activated by binding of unfolded and misfolded proteins to its luminal domain (Wang & Kaufman 2016). The kinase GCN2 binds to deacylated transfer-RNAs (tRNAs) and is activated in response to amino acid deprivation (Aldana 1994). GCN2 can also be activated by UV light since UV induces crosslinking between GCN2 and tRNAs or rapid consumption of arginine due to UV activation of nitric oxide synthetase (Deng 2002, Lu 2009). The mammalian PKR is activated by double-stranded RNA (dsRNA) during viral infection, which results in inhibition of viral and host protein synthesis. Furthermore, it was noticed that other stresses such as oxidative and ER stress, bacterial infection or growth factor deprivation can activate PKE in a dsRNA-dependent manner. The least uncommon eIF2α is HRI that is mainly expressed in erythroid cells; additionally, HRI is also expressed in several cell types and organs and responds to multiple cellular stress (Guo 2019). HRI requires the availability of heme for the production of hemoglobin to protect erythroid cells against accumulation of toxic globin aggregates (Han 2001, Chefalo 1998). Binding of heme to HRI promotes disulfide bond formation between HRI monomers so that they can stay in inactive dimer conformation and its kinase activity is inhibited (Han 2001). In general, these eIF2α kinases have overlapping functions and can cooperate to regulate cellular responses to multiple stress stimuli.

To restore protein synthesis and cellular functions, ISR has to be terminated by dephosphorylation of eIF2α, which is facilitated by protein phosphatase (Novoa 2001). In mammals, phosphatase activity is regulated by GADD34 or CReP (Novoa 2001, Jousse 2003). GADD34 expression is regulated downstream of phosphorylated eIF2α and ATF4 during the later stages of ISR to increase dephosphorylation of eIF2α (Kojima 2003); on the other hand, CReP is constitutively expressed and cooperates with PP1c in unstressed cells to maintain protein synthesis by regulating phosphorylation of eIF2α (Jousse 2003).

Protein degradation including ubiquitin-proteome system and autophagy

In the cell, from any compartment, the unwanted and toxic proteins need to be degraded to recycle the smaller building blocks for new materials. Two major pathways of degradation in the cell are ubiquitin-proteasome system and the autophagy-lysosome system.

The 26S proteasome

The eukaryotic 26S proteasome complex consists of two subcomplexes: the 20S core particle (CP) and the 19S regulatory particle (RP). The CP is the proteolytic site of the proteasome that is a barrel-shaped cylinder and consists of α subunits and β subunits arranged in four stacked heteroheptameric rings (Lupas 1995, Tanaka 2009, Coux 1996, Kunjappu & Hochstrasser 2014). The RP plays a crucial role in controlling substrate recognition, unfolding the proteins and

translocating them to the CP. The RP consists of two subcomplexes, the base and the lid. The materials that are directed to the proteasomes are the small misfolded and short-lived proteins (Dikic 2017). These substrates that are directed to the proteasome need to be ubiquitinated to be recognized for degradation. Ubiquitin (Ub) is involved in hierachiral and reversible enzymatic reactions to be degradation tags. Firstly, Ub is activated through the hydrolysis of ATP to ADP by E1, followed by a transfer of Ub thioester to a Ub-conjugating enzyme (E2). The formation of the isopeptide bond between the Ub and the substrate is catalyzed by Ub-ligase enzyme (E3). The modification of ubiquitination can be at the N-terminus or at a lysine side chain of Ub (K6, K11, K27, K29, K33, K48, K63). These reactions can be repeated until the substrates are poly-ubiquitinated (Oh 2018). Then, the substrates can be recognized directly by the Ub receptors of the proteasome or by the shuttle factors that contain both a Ub-binding domain and a domain binding to the proteasome. Before being degraded inside the barrel of the proteasome, the substrates are unfolded and deubiquitinated for the Ub to be recycled (Pohl & Dikic 2019). This UPS system is significantly important in degrading misfolded proteins that retrotranslocate to the cytosol from the ER. It was shown that deficiency in UPS pathway caused neurodegeneration in which intracellular aggregates are formed (Bence 2001). Also, injection of epoxomicin or synthetic proteasomal inhibitors into rats resulted in progressive parkinsonism (McNaught 2004). Thus, the UPS plays a vital role in removing protein aggregates that are potentially causative of neurodegeneration.

Autophagy

Unlike the substrates of the proteasomes, the substrates of autophagy are bulky materials such as protein aggregates, bacteria or damaged organelles (Dikic & Elazar 2018). The labeling signals of autophagic substrates are more diverse, including Ub, lipid-based signals (galectins), etc. (Pohl & Dikic 2019, Johannes 2018). There are three types of autophagy: macroautophagy, microautophagy, and chaperone-mediated autophagy (CMA). The most common and best-characterized autophagy is macroautophagy and termed “autophagy” for the rest of the dissertation. The substrates are then recognized by the autophagy receptors, which contain a Ub binding domain and an LC3-interacting region (LIR) (Rogov 2014). A growing phagophore can engulf the whole autophagic cargo via the LIR motif-containing autophagy receptors to form the phagosome, which is eventually sequestered then fuses with the lysosome. The materials are then degraded inside the autolysosome and amino acids are recycled (Pohl & Dikic 2019). Autophagy dysfunction also results in deleterious effects on other systems such as enhancement of apoptosis (Boya 2004, Ravikumar 2006) or nutrient deprivation due to inefficient recycling of macromolecules (Kuma 2004).

In this dissertation, Chapter 2 would discuss the inhibitory effect of ER translocation of HSPA13 on transthyretin. In Chapter 3, we unraveled a quality control system in the cytosol towards mistargeted ER proteins. In Chapter 4, we

presented a new electrophilic profiling technique coupled to reaction monitoring mass spectrometry to probe for protein conformation.

REFERENCES

1. Aguzzi, A. & Polymenidou, M. Mammalian Prion Biology: One Century of Evolving Concepts. *Cell* **116**, 313–327 (2004).

2. Åkerfelt, M., Morimoto, R. I. & Sistonen, L. Heat shock factors: integrators of cell stress, development and lifespan. *Nature Reviews Molecular Cell Biology* **11**, 545–555 (2010).

3. Aldana, C. R. V. de, Wek, R. C., Segundo, P. S., Truesdell, A. G. & Hinnebusch, A. G. Multicopy tRNA genes functionally suppress mutations in yeast eIF-2 alpha kinase GCN2: evidence for separate pathways coupling GCN4 expression to unchanged tRNA. *Molecular and Cellular Biology* **14**, 7920–7932 (1994).

4. Ameri, K. & Harris, A. L. Activating transcription factor 4. *The International Journal of Biochemistry & Cell Biology* **40**, 14–21 (2008).

5. Anfinsen, C. B. The formation and stabilization of protein structure. *Biochem J* **128**, 737–749 (1972).

6. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **181**, 223–230 (1973).

7. Anttonen, A.-K. *et al.* The gene disrupted in Marinesco-Sjögren syndrome encodes SIL1, an HSPA5 cochaperone. *Nature Genetics* **37**, 1309–1311 (2005).

8. Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. Adapting Proteostasis for Disease Intervention. *Science* **319**, 916–919 (2008).

9. Bence, N. F., Sampat, R. M. & Kopito, R. R. Impairment of the Ubiquitin-Proteasome System by Protein Aggregation. *Science* **292**, 1552–1555 (2001).

10. Boya, P. *et al.* Inhibition of Macroautophagy Triggers Apoptosis. *Molecular and Cellular Biology* **25**, 1025–1040 (2005).

11. Chefalo, P. J., Oh, J., Rafie-Kolpin, M., Kan, B. & Chen, J.-J. Heme-regulated eIF-2α kinase purifies as a hemoprotein. *European Journal of Biochemistry* **258**, 820–830 (1998).

12. Chen, X., Shen, J. & Prywes, R. The Luminal Domain of ATF6 Senses Endoplasmic Reticulum (ER) Stress and Causes Translocation of ATF6 from the ER to the Golgi. *J. Biol. Chem.* **277**, 13045–13052 (2002).

13. Chromy, B. A. *et al.* Self-Assembly of Aβ1-42 into Globular Neurotoxins. *Biochemistry* **42**, 12749–12760 (2003).

14. Costa-Mattioli, M. & Walter, P. The integrated stress response: From mechanism to disease. *Science* **368**, (2020).

15. Coux, O., Tanaka, K. & Goldberg, A. L. Structure and Functions of the 20s and 26s Proteasomes. *Annual Review of Biochemistry* **65**, 801–847 (1996).

16. Crick, F. Central Dogma of Molecular Biology. *Nature* **227**, 561–563 (1970).

17. Culyba, E. K. *et al.* Protein Native-State Stabilization by Placing Aromatic Side Chains in N-Glycosylated Reverse Turns. *Science* **331**, 571–575 (2011).

18. Deng, J. *et al.* Activation of GCN2 in UV-Irradiated Cells Inhibits Translation. *Current Biology* **12**, 1279–1286 (2002).

19. Dikic, I. Proteasomal and Autophagic Degradation Systems. *Annual Review of Biochemistry* **86**, 193–224 (2017).

20. Dikic, I. & Elazar, Z. Mechanism and medical implications of mammalian autophagy. *Nature Reviews Molecular Cell Biology* **19**, 349–364 (2018).

21. Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. *Nature Structural Biology* **4**, 10–19 (1997).

22. Dobson, C. M. Principles of protein folding, misfolding and aggregation. *Seminars in Cell & Developmental Biology* **15**, 3–16 (2004).

23. Donnelly, N., Gorman, A. M., Gupta, S. & Samali, A. The eIF2α kinases: their structures and functions. *Cell. Mol. Life Sci.* **70**, 3493–3511 (2013).

24. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **6**, 197–208 (2005).

25. Ellis, R. J. & Hartl, F. U. Principles of protein folding in the cellular environment. *Current Opinion in Structural Biology* **9**, 102–110 (1999).

26. Ellis, R. J. & Minton, A. P. Protein aggregation in crowded environments. *Biological Chemistry* **387**, 485–497 (2006).

27. García, M. A. *et al.* Impact of Protein Kinase PKR in Cell Biology: from Antiviral to Antiproliferative Action. *Microbiol. Mol. Biol. Rev.* **70**, 1032–1060 (2006).

28. Gershenson, A. & Gierasch, L. M. Protein folding in the cell: challenges and progress. *Current Opinion in Structural Biology* **21**, 32–41 (2011).

29. Guo, X. *et al.* Mitochondrial dysfunction is signaled to the integrated stress response by OMA1, DELE1 and HRI. *bioRxiv* 715896 (2019)

30. Halperin, L., Jung, J. & Michalak, M. The many functions of the endoplasmic reticulum chaperones and folding enzymes. *IUBMB Life* **66**, 318–326 (2014).

31. Han, A.-P. *et al.* Heme-regulated eIF2α kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency. *The EMBO Journal* **20**, 6909–6918 (2001).

32. Hartl, F. U. Molecular chaperones in cellular protein folding. *Nature* **381**, 571–580 (1996).

33. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).

34. Hartl, F. U. & Hayer-Hartl, M. Converging concepts of protein folding in vitro and in vivo. *Nature Structural & Molecular Biology* **16**, 574–581 (2009).

35. Haze, K., Yoshida, H., Yanagi, H., Yura, T. & Mori, K. Mammalian Transcription Factor ATF6 Is Synthesized as a Transmembrane Protein and Activated by Proteolysis in Response to Endoplasmic Reticulum Stress. *MBoC* **10**, 3787–3799 (1999).

36. Hoeller, D. & Dikic, I. How the proteasome is degraded. *Proc Natl Acad Sci USA* **113**, 13266–13268 (2016).

37. Hong, J. & Gierasch, L. M. Macromolecular Crowding Remodels the Energy Landscape of a Protein by Favoring a More Compact Unfolded State. *J. Am. Chem. Soc.* **132**, 10445–10452 (2010).

38. Hou, Y. *et al.* Ageing as a risk factor for neurodegenerative disease. *Nature Reviews Neurology* **15**, 565–581 (2019).

39. Johannes, L., Jacob, R. & Leffler, H. Galectins at a glance. *J Cell Sci* **131**, jcs208884 (2018).

40. Jousse, C. *et al.* Inhibition of a constitutive translation initiation factor 2α phosphatase, CReP, promotes survival of stressed cells. *J Cell Biol* **163**, 767–775 (2003).

41. Kampinga, H. H. & Craig, E. A. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology* **11**, 579–592 (2010).

42. Kaushik, S. & Cuervo, A. M. Proteostasis and aging. *Nature Medicine* **21**, 1406–1415 (2015).

43. Khurana, R. *et al.* Partially Folded Intermediates as Critical Precursors of Light Chain Amyloid Fibrils and Amorphous Aggregates. *Biochemistry* **40**, 3525–3535 (2001).

44. Klionsky, D. J. *et al.* Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition). *Autophagy* **12**, 1–222 (2016).

45. Kuma, A. *et al.* The role of autophagy during the early neonatal starvation period. *Nature* **432**, 1032–1036 (2004).

46. Kojima, E. *et al.* The function of GADD34 is a recovery from a shutoff of protein synthesis induced by ER stress: elucidation by GADD34-deficient mice. *The FASEB Journal* **17**, 1573–1575 (2003).

47. Kokame, K., Kato, H. & Miyata, T. Identification of ERSE-II, a New cis-Acting Element Responsible for the ATF6-dependent Mammalian Unfolded Protein Response. *J. Biol. Chem.* **276**, 9199–9205 (2001).

48. Korennykh, A. & Walter, P. Structural Basis of the Unfolded Protein Response. *Annual Review of Cell and Developmental Biology* **28**, 251–277 (2012).

49. Kuma, A. *et al.* The role of autophagy during the early neonatal starvation period. *Nature* **432**, 1032–1036 (2004).

50. Kunjappu, M. J. & Hochstrasser, M. Assembly of the 20S proteasome. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**, 2–12 (2014).

51. Lee, E.-S., Yoon, C.-H., Kim, Y.-S. & Bae, Y.-S. The double-strand RNA-dependent protein kinase PKR plays a significant role in a sustained ER stress-induced apoptosis. *FEBS Letters* **581**, 4325–4332 (2007).

52. Lindquist, S. L. & Kelly, J. W. Chemical and Biological Approaches for Adapting Proteostasis to Ameliorate Protein Misfolding and Aggregation

Diseases–Progress and Prognosis. *Cold Spring Harb Perspect Biol* **3**, a004507 (2011).

53. Lu, W., László, C. F., Miao, Z., Chen, H. & Wu, S. The Role of Nitric-oxide Synthase in the Regulation of UVB Light-induced Phosphorylation of the α Subunit of Eukaryotic Initiation Factor 2. *J. Biol. Chem.* **284**, 24281–24288 (2009).

54. Lupas, A., Zwickl, P., Wenzel, T., Seemüller, E. & Baumeister, W. Structure and Function of the 20S Proteasome and of Its Regulatory Complexes. *Cold Spring Harb Symp Quant Biol* **60**, 515–524 (1995).

55. Mayer, M. P. Gymnastics of Molecular Chaperones. *Molecular Cell* **39**, 321–331 (2010).

56. McNaught, K. S. P., Perl, D. P., Brownell, A.-L. & Olanow, C. W. Systemic exposure to proteasome inhibitors causes a progressive model of Parkinson's disease. *Annals of Neurology* **56**, 149–162 (2004).

57. Needham, P. G., Guerriero, C. J. & Brodsky, J. L. Chaperoning Endoplasmic Reticulum–Associated Degradation (ERAD) and Protein Conformational Diseases. *Cold Spring Harb Perspect Biol* **11**, a033928 (2019).

58. Novoa, I., Zeng, H., Harding, H. P. & Ron, D. Feedback Inhibition of the Unfolded Protein Response by GADD34-Mediated Dephosphorylation of eIF2α. *J Cell Biol* **153**, 1011–1022 (2001).

59. Oh, E., Akopian, D. & Rape, M. Principles of Ubiquitin-Dependent Signaling. *Annual Review of Cell and Developmental Biology* **34**, 137–162 (2018).

60. Pajares, M., Cuadrado, A. & Rojo, A. I. Modulation of proteostasis by transcription factor NRF2 and impact in neurodegenerative diseases. *Redox Biol* **11**, 543–553 (2017).

61. Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *PNAS* **37**, 205–211 (1951).

62. Planté-Bordeneuve, V. & Said, G. Familial amyloid polyneuropathy. *The Lancet Neurology* **10**, 1086–1097 (2011).

63. Pohl, C. & Dikic, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **366**, 818–822 (2019).

64. Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W. & Balch, W. E. Biological and Chemical Approaches to Diseases of Proteostasis Deficiency. *Annual Review of Biochemistry* **78**, 959–991 (2009).

65. Ravikumar, B., Berger, Z., Vacher, C., O'Kane, C. J. & Rubinsztein, D. C. Rapamycin pre-treatment protects against apoptosis. *Hum Mol Genet* **15**, 1209–1216 (2006).

66. Richter, K., Haslbeck, M. & Buchner, J. The Heat Shock Response: Life on the Verge of Death. *Molecular Cell* **40**, 253–266 (2010).

67. Rogov, V., Dötsch, V., Johansen, T. & Kirkin, V. Interactions between Autophagy Receptors and Ubiquitin-like Proteins Form the Molecular Basis for Selective Autophagy. *Molecular Cell* **53**, 167–178 (2014).

68. Ron, D. Translational control in the endoplasmic reticulum stress response. *J Clin Invest* **110**, 1383–1388 (2002).

69. Ron, D. & Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nature Reviews Molecular Cell Biology* **8**, 519–529 (2007).

70. Ross, C. A. & Poirier, M. A. Protein aggregation and neurodegenerative disease. *Nature Medicine* **10**, S10–S17 (2004).

71. Rousseau, A. & Bertolotti, A. Regulation of proteasome assembly and activity in health and disease. *Nature Reviews Molecular Cell Biology* **19**, 697–712 (2018).

72. Schröder, M. & Kaufman, R. J. The Mammalian Unfolded Protein Response. *Annual Review of Biochemistry* **74**, 739–789 (2005).

73. Selkoe, D. J. Folding proteins in fatal ways. *Nature* **426**, 900–904 (2003).

74. Senderek, J. *et al.* Mutations in SIL1 cause Marinesco-Sjögren syndrome, a cerebellar ataxia with cataract and myopathy. *Nature Genetics* **37**, 1312–1314 (2005).

75. Sengupta, U., Nilson, A. N. & Kayed, R. The Role of Amyloid-β Oligomers in Toxicity, Propagation, and Immunotherapy. *EBioMedicine* **6**, 42–49 (2016).

76. Shankar, G. M. *et al.* Natural Oligomers of the Alzheimer Amyloid-β Protein Induce Reversible Synapse Loss by Modulating an NMDA-Type Glutamate Receptor-Dependent Signaling Pathway. *J. Neurosci.* **27**, 2866–2875 (2007).

77. Shea, J.-E. & Brooks III, C. L. FROM FOLDING THEORIES TO FOLDING PROTEINS: A Review and Assessment of Simulation Studies of Protein Folding and Unfolding. *Annual Review of Physical Chemistry* **52**, 499–535 (2001).

78. Shimazawa, M. & Hara, H. Inhibitor of double stranded RNA-dependent protein kinase protects against cell damage induced by ER stress. *Neuroscience Letters* **409**, 192–195 (2006).

79. Sunde, M. & Blake, C. C. F. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quarterly Reviews of Biophysics* **31**, 1–39 (1998).

80. Suragani, R. N. V. S. *et al.* Heme-regulated eIF2α kinase activated Atf4 signaling pathway in oxidative stress and erythropoiesis. *Blood* **119**, 5276–5284 (2012).

81. Tanaka, K. The proteasome: Overview of structure and functions. *Proceedings of the Japan Academy, Series B* **85**, 12–36 (2009).

82. Walter, P. & Ron, D. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. *Science* **334**, 1081–1086 (2011).

83. Wang, M. & Kaufman, R. J. Protein misfolding in the endoplasmic reticulum as a conduit to human disease. *Nature* **529**, 326–335 (2016).

84. Wek, R. C., Jiang, H.-Y. & Anthony, T. G. Coping with stress: eIF2 kinases and translational control. *Biochem Soc Trans* **34**, 7–11 (2006).

85. Williams, D. B. Beyond lectins: the calnexin/calreticulin chaperone system of the endoplasmic reticulum. *J Cell Sci* **119**, 615–623 (2006).

86. Xu, P. *et al.* Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* **137**, 133–145 (2009).
78. Ye, J. *et al.* ER Stress Induces Cleavage of Membrane-Bound ATF6 by the Same Proteases that Process SREBPs. *Molecular Cell* **6**, 1355–1364 (2000).

87. Yoshida, H. *et al.* Endoplasmic Reticulum Stress-Induced Formation of Transcription Factor Complex ERSF Including NF-Y (CBF) and Activating Transcription Factors 6α and 6β That Activates the Mammalian Unfolded Protein Response. *Molecular and Cellular Biology* **21**, 1239–1248 (2001).

CHAPTER 2: REGULATION OF HSPA13 ON ENDOPLASMIC-RETICULUM TRANSLOCATION OF TRANSTHYRETIN

INTRODUCTION

An important step that is involved in protein synthesis and maturation is translocation into endoplasmic reticulum (ER) during or after translation. This step is required for the proteins that get secreted from the cells, membrane proteins or ER-localized proteins. These proteins contain N-terminal and cleavable signal sequences that are 7-12 hydrophobic amino acids in length (Rapoport 2007). The signal sequences direct the proteins to enter the ER, which is termed “ER translocation”. The translocation process is able to take place owing to the protein-conducting channel. In eukaryotes, this channel is Sec61 complex located on the ER membrane while SecY complex is positioned on the cell membrane (Berg 2004, Orborne 2005). This complex by itself is a passive pore; therefore, it has to associate with other components for translocation. In co-translational translocation, the ribosome and the signal-recognition particle (SRP) are the partners. While the proteins are being translated, the signal or transmembrane sequence is recognized by the SRP. Then, the complex of the ribosome, the polypeptide chain and the SRP binds to the membrane, first by an interaction between SRP and its membrane receptor and by the interaction between the ribosome and the translocation channel (Rapoport 2007). Subsequently, the polypeptide chain moves from the ribosome into the associated membrane channel and gets inserted into the ER. Some proteins employ post-translational translocation into the ER. These proteins have moderately hydrophobic residues on their signal sequences that can escape the

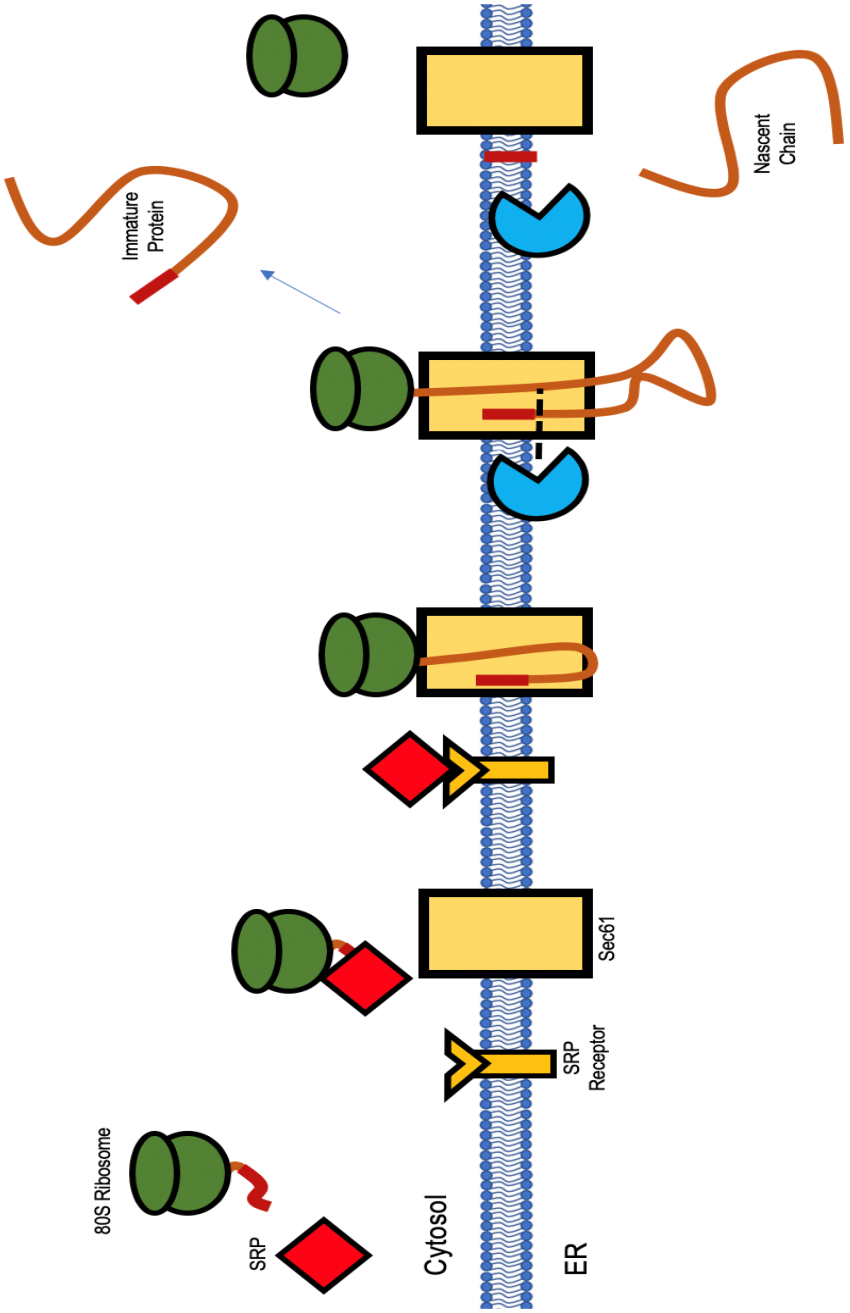

Immature Protein
Nascent Chain
Sec61
SRP Receptor
80S Ribosome
SRP
Cytosol
ER

Figure 2.1: Cotranslational ER Translocation of Proteins. The N terminus of the newly synthesized polypeptide protrudes out of the ribosome. Synthesis is arrested when the signal recognition particle (SRP) binds to the signal sequence of the polypeptide chain and synthesis resumes when SRP docks on the SRP receptor. The polypeptide is then inserted into the ER lumen via the Sec61 translocon. Afterwards, the signal sequence is cleaved and the nascent protein can enter the secretory pathway.

recognition of SRP. Another type of these proteins is tail-anchored (TA) proteins, whose transmembrane domains (TMD) are at the C-terminus. Since C-terminal TMDs cannot be identified until the proteins are completely translated, the proteins are targeted to the ER post-translationally (Borgese & Gaetani 1983). In this mode of translocation, Sec61 channel has to partner with Sec62/Sec63 complex and with the luminal chaperone BiP (Deshaies 1991, Panzner 1995, Meyer 2000, Tyedmers 2000). To initiate translocation, substrate binds to the channel and cytosolic chaperones dissociate from the protein. Multiple BiP molecules are able to bind to the polypeptide chain through ATP hydrolysis in consecutive rounds, routing the polypeptide into the ER (Rapoport 2007). After the translocation, the signal sequence of the secretory proteins is cleaved by the signal peptidase and the protein is properly folded. Other modifications can take place after the proteins are inserted into the ER. The completion of protein folding and modification is made possible owing to the presence of chaperones and molecular catalysts in the ER. As mentioned above, HSP70 family is a major chaperone network in the proteostasis system. GRP78/BiP is a well-characterized chaperone that is upregulated when UPR is activated. Some other chaperones, such as HSPA13 also known as STCH, have not been well-established in their functions (Kampinga 2009).

HSP70 chaperones contain an NBD and SBD domain which are necessary for their functions, as illustrated in Chapter 1. Interestingly, HSPA13 contains a

segmented NBD and possibly the ATPase activity; however, it does not have an SBD and presumably cannot interact with misfolded proteins as the other HSP70 chaperones do (Otterson 1994). Nevertheless, ATPase activity on the NBD regulates the SBD, so HSPA13 can still serve as a regulatory factor of its substrates. HSPA13 is expressed in all human cell types and is conserved across metazoan, which implies its significant physiological role. Additionally, the sequence of HSPA13 is 42% aligned with GRP78/BiP; however, the linker region is not conserved but still present in all species. Like other HSP70s being upregulated by UPR, the expression of HSPA13 increases when the UPR transcription factor X-box binding protein is spliced (Bensellam 2012, Grandjean 2019).

Although its functionality is not intensively studied, HSPA13 is implicated in a few diseases, such as cancer and prion diseases. HSPA13 mutant was found to be associated with gastric cancer in the Japanese population (Aoki 2005). Another disease that HSPA13 plays a role in is prion disease, a fatal neurodegenerative disorder affecting several mammalian species including human. This disease is caused by misfolding of prion protein (PrP). The toxic Scrapie conformation of prion protein (PrP^{SC}) can recruit and convert the normal PrP (PrP^{C}) to PrP^{SC}. Eventually, PrP^{SC} becomes resistant to degradation and form oligomeric aggregates that give rise to the disease. The actual mechanism of the PrP pathogenesis has not been well-defined; however; there are identified factors

that can change the states of PrP. First, the maintenance of ER proteostasis can prevent PrP aggregation. Immediately after ER stress, PrP would get exported from the ER to the Golgi, transiently get access to the cell surface and eventually get degraded by the lysosomes (Satpute-Krishnan 2014). Additionally, PrP can be modulated by its efficiency in ER translocation. Levine et al. found out that switching the signal sequence of PrP can alter its ER translocation (Levine 2005). Notably, Kang et al. figured that during ER stress, translocational attenuation of proteins into the ER serves as pre-emptive quality control (pQC) to maintain proteostasis in the cell. Thus, PrP is less efficient in translocating into the ER during ER stress (Kang 2006). Noticeably, Rane et al. indicated that decreased translocation of PrP to the ER during ER stress contributes to neurodegeneration of the transgenic mice (Rane 2008). Another host susceptibility factor is the expression of HSPA13. Interestingly, HSPA13 overexpression reduced the incubation time of different prion strains in mice, while knockdown of HSPA13 mRNA in chronically prion infected cells significantly reduces PrP^{SC} accumulation (Grizenkova 2012). This observation suggests that changes in HSPA13 expression result in direct effects on the phenotype of the disease. In another study conducted by Estoppey *et al.*, it suggested that HSPA13 associates with signal peptidase. Different concentrations of an inhibitor of signal peptidase SEC11 cavinafungin greatly sensitized the HSPA13 knockout cell line (Estoppey 2017). Regarding the severity of prion disease and the efficiency of translocation of PrP, we

hypothesized HSPA13 plays a role as an inhibitor of protein translocation that may accelerate the disease. We set out to characterize its functionality with respect to its effects on protein translocation and proteostasis.

The model we employed in our study is transthyretin (TTR). TTR is a homotetrameric protein that shuttles thyroxin and retinol. It is predominantly synthesized in the liver as a single polypeptide chain and forms a tetramer of noncovalently bound monomers in the plasma (Johnson 2005). Also, TTR is synthesized from the choroid plexus of the brain. Wild-type TTR or mutants are highly amyloidogenic that can lead to fatal diseases characterized by progressive neuropathy and/or cardiomyopathy (Benson & Roudebush 1989, Ando & Suhr 1998, Ng 2005, Westermark 1990). In molecular aspects, TTR enters the ER during its translation and progresses to the extracellular secretion. Therefore, TTR is likely to bear similarity to PrP which HSPA13 has effects on. PrP is toxic and hard to handle, so we employed TTR as the model for the study of protein of interest HSPA13.

We hypothesize that HSPA13 negatively regulates translocation of proteins from the cytosol to the ER, impairing their maturation as well as their cleavage of signal sequence. To accomplish the purpose of the study, we employed a proximity labeling assay to partition the TTR protein between the ER and the cytosol. To further support the hypothesis of the modulating role of HSPA13 with

ER translocation, we applied quantitative proteomics to identify potential interactors of HSPA13. Furthermore, we also unraveled the consequences of TTR and the cellular proteomics under the expression of HSPA13.

MATERIALS AND METHODS

Plasmid DNA

HSPA13 and eGFP were ligated into pDEST30 backbone. $^{Flag}TTR^{WT}$ (N-terminal Flag) was prepared from pcDNA3.1. Cytosolic peroxidase pcDNA3 APEX-NES was a gift from Alice Ting (Addgene plasmid # 49386 ; http://n2t.net/addgene:49386 ; RRID:Addgene_49386) [64]. pCMV-erHRP(N175S mutant) was a gift from Joshua Sanes (Addgene plasmid # 79909 ; http://n2t.net/addgene:79909 ; RRID:Addgene_79909) [65].

PIPE cloning for $HSPA13^{Flag}$

Primers (Integrated DNA Technologies) for PIPE addition of Flag tag on HSPA13 were designed to bind to the vector backbone (HSPA13) or the insert backbone (Flag) with the overhanging on the 5' end complementary to the opposite amplicon. The amplified vector and insert from PCR were digested with DpnI to remove the template, and mixed in the ratio of 3:17 (vector:insert). The mixture was chemically transformed to DH5α E. coli, which were heat shocked at 42^{o}C for 45 s.

Site-directed Mutagenesis of $HSPA13^{WT}$ to HSPA13 mutants

Primers (Integrated DNA Technologies) incorporating the nucleotide changes needed for amino acid substitution were included for site-directed mutagenesis (SDM) PCR. $HSPA13^{WT}$ was amplified with the designed primers with Q5 polymerase (New England BioLabs s M0491S). The template of $HSPA13^{WT}$ was digested with *DpnI* (New England BioLabs R0176S). Amplicons were ligated with T7 DNA ligase (New England Biotechnologies M0318S) and transformed into DH10β electrocompetent cells using a Bio-Rad MicroPulser electroporator set at 1.4 kV.

Cell culture

Human embryonic kidney cell line 293 with T antigen (HEK293T) were grown in Dulbecco's Modification of Eagle's Medium (DMEM; Corning 10-013-CV) containing 10% fetal bovine serum, 2 mM L-Glutamine and Penicillin (100 IU/ mL) - Streptomycin (100 μg/ mL). Cells were transfected at 20-60% confluence while being attached to the plates. All transfection performed in the study were $Ca_3(PO_4)_2$ transfection. 250 mM $CaCl_2$ containing DNA plasmid was vortexed with drop-wise addition of 2X HEPES-buffered saline (2X HBS). The resultant transfection solution was added drop-wise to the cells with fresh replenishment of media after 12-16 hours. To enhance the adherence of HEK293T cells on the plates, cells were reseeded on poly-D-lysine treated plates. 1 μM MG132 was

added to the media and incubated for 16 hours before harvest. The negative control received the same volume of DMSO vehicle.

Cell lysis

Cells were lysed in RIPA buffer (50 mM Tris, 15 mM NaCl, 1% Triton X100, 0.5% deoxycholate and 0.1% sodium dodecyl sulfate (SDS)) containing 1X protease inhibitor cocktail (PI) (Roche) for at least 15 minutes on ice. The lysate was collected from the supernatant of a 21,100 g spin for 15 minutes. Protein was quantified from Bradford assay by spectrophotometry (Agilent Cary 50 and 60).

Western Blot sample preparation

The proteins from different samples were normalized and brought to the same volume by RIPA buffer. If the samples are cultured media, no normalization would be performed. The 6X Laemli loading dye was added to a final concentration of 1X and 16.7 mM dithothreitol was also included in the samples. To denature and reduce the proteins, the samples were boiled at 100^{o}C for 5 minutes or 20 minutes as indicated. Proteins were loaded on homemade SDS-polyacrylamide gel electrophoresis (SDS-PAGE) gels cast in the Bio-Rad Mini-PROTEAN system. The samples were run through the stacking gel and the resolving gel at 60 V and 170 V, respectively.

The gels, filter paper and nitrocellulose membrane were immersed in Towbin's buffer, consisting of 25 mM Tris, 192 mM glycine, 20% methanol (per Bio-Rad transfer buffer recipes). The semi-dry transfer of the sandwich was conducted in a Bio-Rad Trans-Blot Turbo for 60 minutes at 25 volts and 1 A. The quality of the transfer was confirmed by Ponceau S assay.

The membranes were then blocked with 5% milk in Tris-buffered saline (TBS) for 1 hour. Milk was rinsed well with TBS for several rounds. The proteins were then probed with primary antibody diluted in 5% bovine serum albumin (BSA), 0.1% NaN_3 in TBS for at least two hours. The membranes were then rinsed with TBS buffer with 0.1% Tween (TBST) and probed with secondary antibody targeting the primary host such as rabbit or mouse for 30 minutes. The secondary antibody was diluted in 1:20000 ratio in blocking buffer. The membranes were repeatedly rinsed with TBST, followed by a rinse in TBS and water. After all steps of blotting, the membranes were imaged with a Li-Cor Odyssey Fc, using 700 nm and 800 nm filters and exposure times of 10 minutes each. The images were adjusted with Li-Cor's Image Studio, with quantification also performed, if necessary.

Microsomal Extraction Assay

Transfected HEK293T cells were suspended in HBS/PI (20 mM HEPES, 100 mM NaCl and 1 mM EDTA). Cell suspensions were incubated with 0-0.25% digitonin for 15 minutes, then centrifuged for 10 minutes at 10,000 x g, resulting in

separated pellets and supernatants. Pellets were resuspended in HBS/PI with 1% Triton and spun down for 10 minutes at 10,000 x g. Both pellets and supernatants were reduced and boiled for Western blot.

Proximity labeling of living cells

ER-HRP or APEX-NES plasmids and other plasmids were transfected to the cells. Cells were reseeded on PDK-treated plates and subject to drug treatment as indicated above. Fresh media including 500 μM biotin-phenol was added to the cells 30 minutes before harvest. Cells were then incubated at 37°C, 5% CO_2 as usual. To initiate the labeling reaction of peroxidase to biotinylated proteins, 1 mM H_2O_2 was added to the media and the plates were gently agitated for 1 minute. Immediately after 1 minute, the media was aspirated and the quenching solution (5 mM Trolox, 10 mM sodium ascorbate and 10 mM sodium azide in cold DPBS, pH 7.4). was added to the cells. The cells were rinsed in the quenching solution for 3 times before harvest. Cells were then lysed in RIPA buffer containing 5 mM Trolox, 10 mM sodium ascorbate and 10 mM sodium azide in the same protocol as described above. Equal amounts of proteins were added to streptavidin beads and rotated overnight at 4°C. After rotation, the beads were spun down at 1500 x g for 1 minute and the supernatant was aspirated. The beads were then washed 2X with RIPA buffer, 1X with 1 M KCl, 1X with 0.1 M Na_2CO_3, 1X with 2 M urea, and 2X with RIPA buffer. The beads

were then eluted in buffer containing 12% SDS, 10% glycerol in 10 mM Tris pH 6.8, with 10 mM DTT and 2 mM biotin, and boiled for 10 minutes.

Immunoprecipitation

HEK293T cells were transfected with plasmid of $HSPA13^{Flag}$ or GFP as a control. After 2 days post-transfection, the cells were harvested and crosslinked to 2 mM dithiobis(succinimidyl propionate) (DSP) for 30 minutes at RT. Cells were then lysed, and the lysates were normalized and precleared for 30 minutes at 4°C with Sepharose beads. The lysates were transferred to M2 anti-Flag beads and rotated overnight at 4°C. After incubation, the beads were washed stringently 4X with RIPA buffer, eluted in Leamli buffer and boiled for 5 minutes. 10% of the eluates were reduced and saved for confirmation of immunoprecipitation by silver stain. For immunoprecipitation from media, cultured media after 24-hour incubation was collected and spun down at 700 x g to remove cell debris. The media was incubated with M2 anti-Flag beads by rotating overnight at 4°C. The cleared media was removed and the beads were washed with RIPA buffer for 4 times. The proteins were eluted from the beads in Laemli buffer and the eluates were reduced with 10 mM DTT.

Ultracentrifugation assay

Cell lysates after normalization were balanced in weight with RIPA in Beckman-Coulter polypropylene ultracentrifuge tubes (357448) and spun down in a

Beckman-Coulter Optima Max-XP ultracentrifuge, using a TLA-55 rotor. Samples were spun at 77,000 x g for four hours, at 4°C and under vacuum. Pellets formed from ultracentrifuge were rinsed four times with RIPA and incubated for 96 hours with 8 M urea at 4°C. The solubilized pellets were diluted to 2 M urea with RIPA and prepared for Western blot by addition of reducing Laemmli and boiling for 5 minutes.

PNGase treatment

Eluates from immunoprecipitation of A1AT in the media were denatured in Glycoprotein Denaturing Buffer (NEB) by boiling for 10 minutes. The total reaction of deglycosylation contained GlycoBuffer (NEB), NP-40 (NEB), PNGase F (NEB) and the protein sample. The reaction was incubated at 37^{o}C for 1 hour and analyzed by SDS-PAGE and Western blotting.

Quantitative Proteomic Analysis by LC-MS/MS

The protein samples were precipitated with methanol-chloroform precipitation. The volumes of the samples were brought to 100 μL. 400 μL of methanol was added and the samples were vortexed well. 100 μL of chloroform was added and the samples were vortexed well. 300 μL of water were added and the samples were vortexed well. The mixtures were then spun down at 16,000 x g for 5 minutes. A layer of proteins would appear on the interphase. The supernatant was aspirated, then the proteins were rinsed 3 more times with methanol. The

protein pellets were air-dried and then solubilized in 100 mM HEPES buffer containing 1% Rapigest. The samples were reduced with 10 mM tris(2-carboxyethyl)phosphine (TCEP) for 30 minutes, followed by alkylation with 5 mM iodoacetamide for 15 minutes in the dark. The proteins were digested into peptide fragments with 1:100 trypsin:protein ratio and shaken overnight at 37°C. The peptide samples were labeled with tandem mass tags (TMT) dissolved in acetonitrile as indicated in the table below, up to a final 40% concentration of acetonitrile. Incubation with TMT solution took place over one hour at room temperature, followed by quenching with 0.4% ammonium bicarbonate. The labeled peptides were pooled and dried to 10 μL by speedvac, then brought up to 200 uL with buffer A, which consists of 5% acetonitrile and 0.1% formic acid, and acidified to pH 2. The samples were incubated at 37°C for 1 hour and centrifuged at 21,000 x g to remove Rapigest and debris.

We prepared trapping and analytical columns in house by packing resins to the capillary tubing with pressurized He gas. We employed MudPIT separation technology on our samples (Washburn 2001). The prepared peptide mixtures were preloaded onto MudPIT trapping column, which comprises of three phases of separation types (reverse phase – strong cation exchange – reverse phase). Liquid chromatography mass spectrometry (LC-MS) was performed with a Thermo LTQ Velos mass spectrometer and EASY-nLC 1000 UPLC. Separation was achieved by applying 7-55% gradient of solvent B (80% ACN, 0.1% formic

acid) at a flow rate of 500 nL/min. A voltage of 3.0 kV was applied for electrospray ionization and the inlet capillary was heated to 275oC. Data-dependent acquisition of MS/MS spectra with the LTQ-Velos-Orbitrap was performed with the following settings: MS/MS on the 10 most intense ions per precursor scan, 1 microscan, rejection of unassigned and charge state 1; dynamic exclusion repeat count 1, repeat duration, 30s; exclusion list size 500, and exclusion duration, 60s. The MS/MS spectra were obtained from HCD fragmentation in the Orbitrap of 38.0 with a stepped collision of 12% for 3 times. Protein and peptide identification and peptide quantification were performed on Integrated Proteomics Pipeline IP2. Tandem mass spectra were extracted from the raw files using Rawconverter and were searched on a Uniprot human database with reversed sequences using ProLucid algorithm (Peng 2003, He 2015, Uniprot Consortium 2019). The search parameters included semi-tryptic peptides. Carbamidomethylation (+57.02146) of cysteine and tandem mass labeling (+229.1629) of lysine and N-terminus were set as static modification. Peptide candidate were filtered using DTASelect, with these parameters: -p 2 –y 1 –sfp 0.01 –DM 10.

Table 2.1. Tandem Mass Tag Channels (Immunoprecipitates, Cellular proteome and Insoluble proteome)

IP – Replicate	**Channel**	
GFP-1	126	
GFP-2	127C	
GFP-3	128C	
HSPA13-FLAG-1	129N	
HSPA13-FLAG -2	130C	
HSPA13-FLAG -3	131	
UC – Replicate	**Lysate Channel**	**Pellet Channel**
GFP-1	126	126
GFP-2	127C	127N
GFP-3	128C	128N
HSPA13-1	129N	129N
HSPA13-2	130C	130N
HSPA13-3	131	131

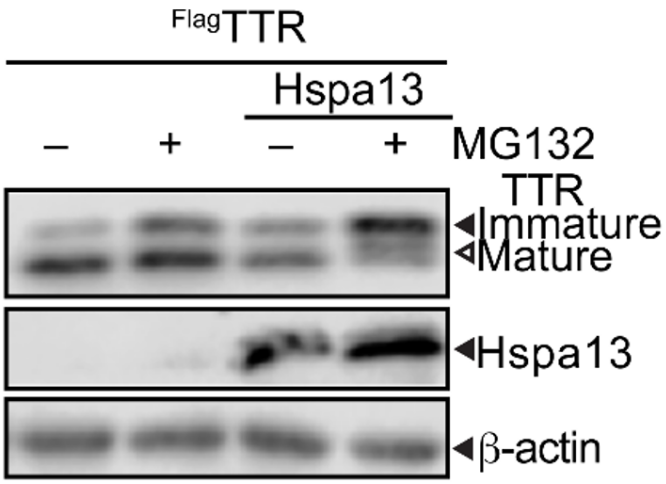
RESULTS

Effects of HSPA13 on ER translocation of $^{Flag}TTR^{WT}$

To evaluate the effect of HSPA13 on protein translocation to the ER, we employed a model of a secretory protein transthyretin with appended N-terminal Flag tag after its signal sequence. We coexpressed the plasmids of $^{Flag}TTR^{WT}$ and HSPA13 in HEK293T cells, and treated with proteasomal inhibitor MG132 at a final concentration of 1 μM for 16 hours. The cells were harvested and lysed and the lysates were separated on reducing SDS-PAGE gel for western blotting. Monomeric $^{Flag}TTR^{WT}$ reveals two distinct bands on the blot: a lighter species at the expected 16 kDa (including the double FLAG tag) and a heavier species at 18 kDa (Figure 2.2.A). The difference of 2 kDa is believed to be the molecular weight of the signal sequence of TTR (Sato 2012). Therefore, we were able to distinguish two different species of monomeric TTR: the immature form that contains the signal sequence and the mature form whose signal sequence is cleaved off. Presumably, the immature TTR has not entered into the ER and the mature TTR has completely entered the ER. The protein has to be in the ER compartment to get its signal sequence cleaved since the signal peptidase is located inside the ER. The results showed that in the absence of HSPA13 overexpression, MG132 slightly increases the formation of the immature band (Figure 2.2.A, lane 1 and 2); nevertheless, when HSPA13 was overexpressed and in the absence of MG132, we observed a three-fold increase in the formation

of the immature band (Figure 2.2.A, lane 3; Figure 2.2.B). Overexpression of HSPA13 and inclusion of MG132 significantly increase the relative amount of immature band and decreases the amount of mature TTR (Figure 2.2.A, lane 4; Figure 2.2.B). We observed reduced amount of mature TTR because MG132 phosphorylates eIF2α to reduce protein synthesis (Jiang 2005). Also, the immature TTR that fails to enter the ER is prone to proteasomal degradation, most likely due to being an unstable species. MG132 inhibits ERAD and thus increases protein load in the ER (Jiang 2005); however, this is not the case because we observed accumulation of immature TTR but not mature TTR. The accumulation of immature TTR upon HSPA13 and MG132 treatment indicated that HSPA13 inhibits the translocation into the ER of TTR and therefore lowers the cleavage of the signal sequence. Nevertheless, we could not base on this result to conclude about the effect of HSPA13 on ER translocation of TTR since it is possible that TTR has entered the ER but the signal sequence has not been cleaved. Therefore, we performed microsomal isolation assay in which a mild detergent digitonin was used to lyse the cells. Digitonin is able to extract the cytosolic materials of the cells but still keep other organelles intact. At different concentrations of digitonin, we were able to distinguish the ER and cytosolic proteomes. As demonstrated, at low concentrations of digitonin, cytosolic proteome is collected in the supernatant and proteome of the ER is present in the pellet after centrifugation. At higher concentrations of digitonin, the membrane of

A


FlagTTR
Hspa13
– + – + MG132
TTR
Immature
Mature
Hspa13
β-actin

B

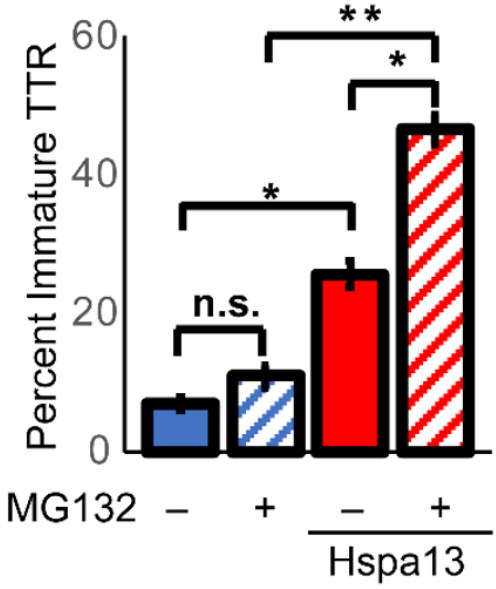
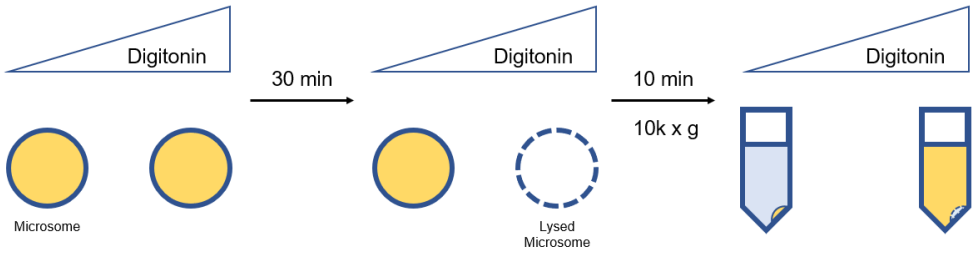

Percent Immature TTR
60
40
20
0
**
*
*
n.s.
MG132 – + – +
Hspa13

Figure 2.2: HSPA13 effects on $^{Flag}TTR^{WT}$ ER translocation and maturation. A) HEK293T cells transfected with $^{Flag}TTR^{WT}$ mammalian expression vectors with or without HSPA13 co-transfection followed by a 16-hour treatment with MG132 proteasomal inhibitor or vehicle. The lysates were reduced and denatured for SDS-PAGE, followed by Western blotting. The indicated antibodies were used for probing the blot, followed by near-IR fluorescent secondary staining. B) Percent immature FLAG-TTR of total near-IR signal (n = 4; **, $p < 0.01$; *, $p < 0.05$).

A


Digitonin
30 min
Digitonin
10 min
10k x g
Digitonin
Microsome
Lysed
Microsome

B

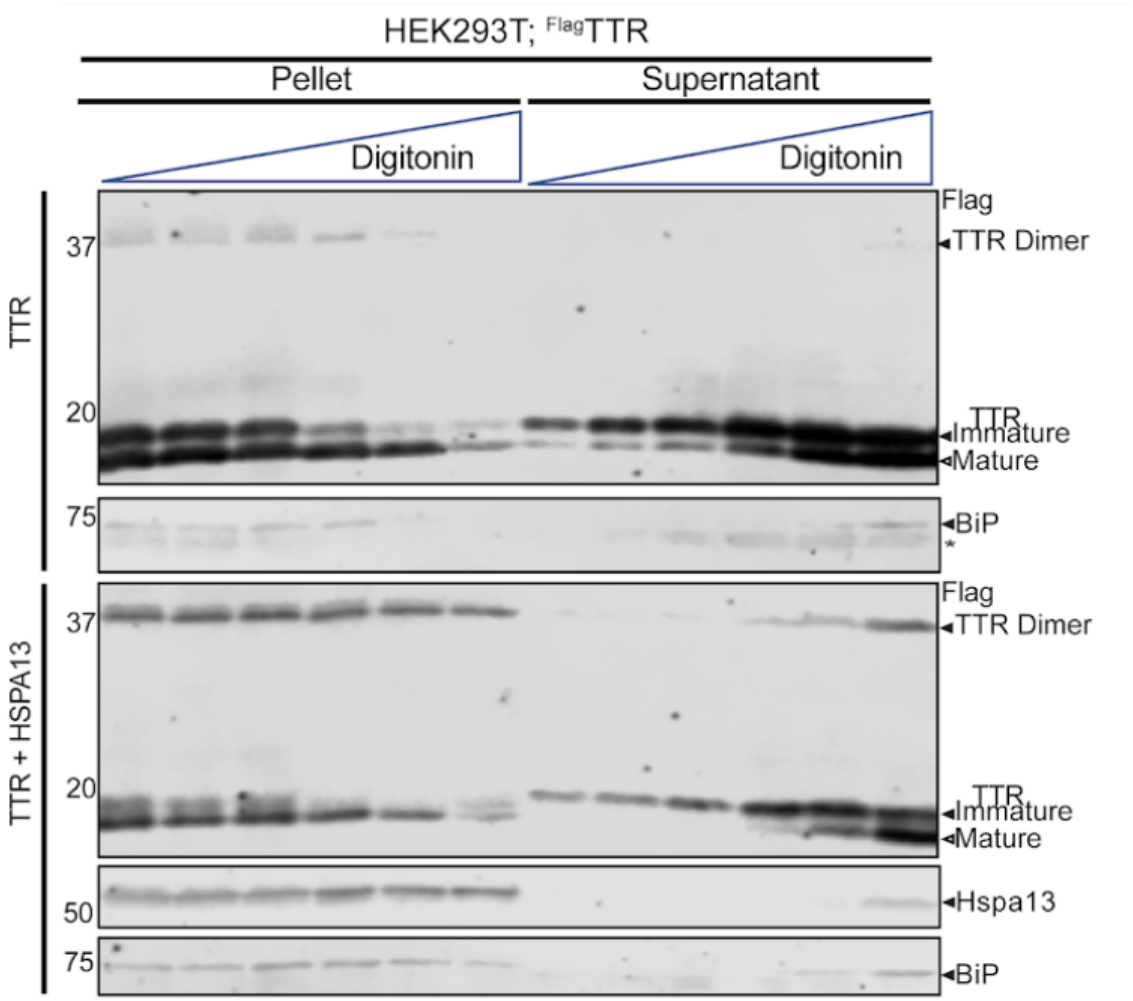

HEK293T; FlagTTR
Pellet
Supernatant
Digitonin
Digitonin
TTR
TTR + HSPA13
37
20
75
37
20
50
75
Flag
TTR Dimer
TTR
Immature
Mature
BiP
*
Flag
TTR Dimer
TTR
Immature
Mature
Hspa13
BiP

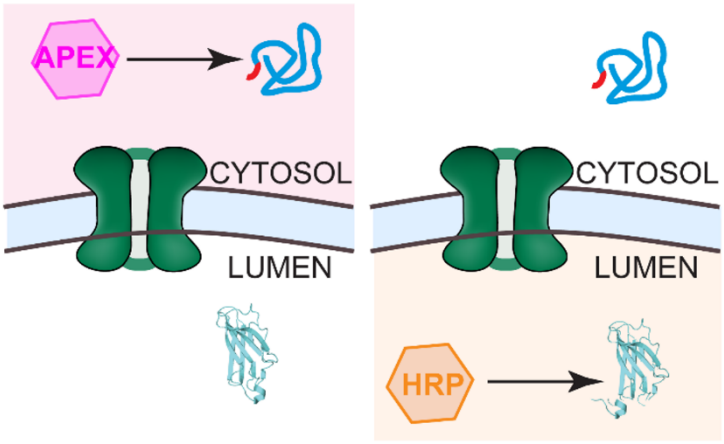
Figure 2.3: Microsomal separation of ER and cytosolic proteins by digitonin. A) Schematic of digitonin-based microsomal separation assay. B) HEK293T cells transfected with ^{Flag}TTR, with or without HSPA13 were incubated with increasing concentrations of digitonin for 30 minutes, then centrifuged for pellet and supernatant fractions.

the ER gets more disrupted so the proteome of the ER will get released. The immature TTR was present in the supernatant at low concentrations of digitonin while the mature form was detected in the supernatant at higher concentrations of digitonin (Figure 2.3.B). HSPA13 was selectively detected in the pellet, which confirms its ER localization. Overexpression of HSPA13 relatively changes the population of the two TTR monomeric species. However, it does not affect the recovery rate of TTR with digitonin, indicating that HSPA13 impacts the localization of TTR rather than the activity of the signal peptidase.

Besides employing the assay of microsomal extraction to differentiate the localization of immature and mature TTR, we developed a proximity labeling assay in our laboratory that enabled characterizing modulation of ER translocation of TTR. Thus, we employed this technique to identify the localization of TTR species under overexpression of HSPA13. In this assay, we transfected two different peroxidases that are functional in the ER and cytosol. ER-targeting horseradish peroxidase (^{ER}HRP) with an ER-retaining sequence KDEL (Joesch 2016) and its signal sequence is targeted in the ER, while an engineered ascorbate peroxidase (^{cyt}APEX2) is localized in the cytosol (Lam 2015) (Figure 2.4.A). The transfected cells were treated with 1 μM MG132 to prevent degradation of mistargeted proteins. Before the labeling process, biotin-phenol (BP) was added to the cells. In the presence of hydrogen peroxide and BP, endogenous proteins within a few nanometers of the peroxidases are

covalently biotinylated. Since labeling is performed within 1 minute while the cells and organelles are still intact, compartmentalized proteins are preserved, resulting in their spatially specific tagging by biotin. Cells are then lysed, and biotinylated proteins are enriched with avidin beads and analyzed by Western blotting. As expected, GRP78/BiP and HSPA13 were preferentially labeled with ^{ER}HRP while beta-actin and HSPA1A were selectively labeled with ^{cyt}APEX2. There was large background of TTR staining in the presence of ER-HRP when the blot was imaged for HSPA1A due to having the same secondary antibodies. ^{cyt}APEX2 labeling and subsequent purification substantially increases ^{Flag}TTR in the cytosolic fraction when HSPA13 was overexpressed (Figure 2.4.B). Additionally, we observed another intermediate band of ^{Flag}TTR biotinylated with ^{cyt}APEX2, implying another species of immature TTR present in the cytosol. This result demonstrated that HSPA13 promotes ^{Flag}TTR mistargeting in the cytosol. Furthermore, ^{Flag}TTR cytosolic fraction is the immature monomer containing the signal sequence, which suggests regulation of ER translocation of HSPA13 before the substrate is exposed to signal peptidase.

A


APEX
CYTOSOL
LUMEN
CYTOSOL
LUMEN
HRP

B

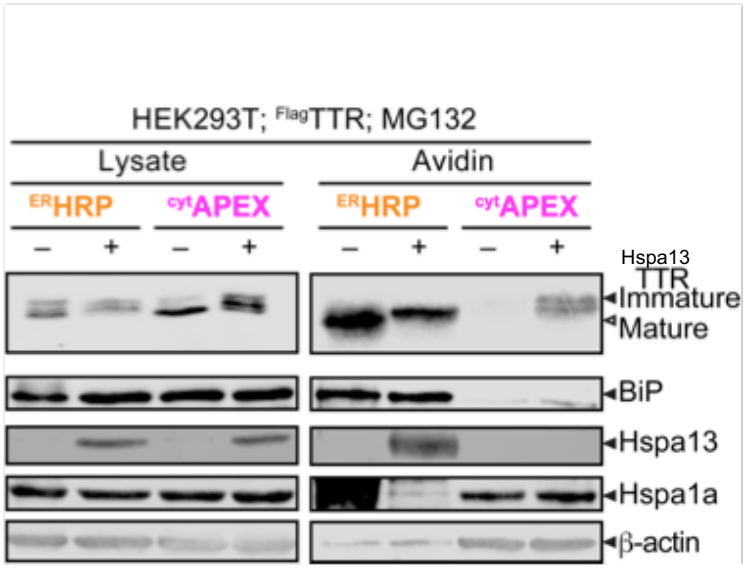

HEK293T; FlagTTR; MG132
Lysate
Avidin
ERHRP
cytAPEX
ERHRP
cytAPEX
– + – + – + – +
Hspa13
TTR
Immature
Mature
BiP
Hspa13
Hspa1a
β-actin

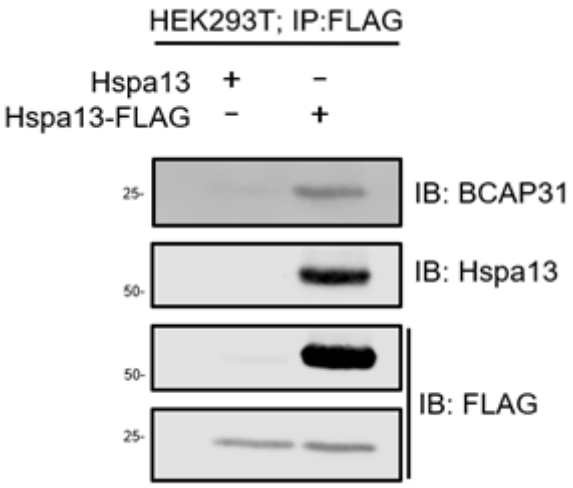
Figure 2.4: Proximity Labeling Assay of ^{Flag}TTR. A) Schematic of labeling strategy. ER-HRP and ^{cyt}APEX2 peroxidases (green hexagons) are expressed in either the ER or cytosol, respectively. Upon addition of hydrogen peroxide (H_2O_2) and biotin-phenol (BP), the targeted locale is labeled (pink and orange background), allowing for selective purification. B) Representative Western blot of HEK293T lysates (left) and avidin-purified eluates (right).

Potential interactors of HSPA13

We observed that HSPA13 inhibits the ER translocation of ^{Flag}TTR leading to mistargeting of ^{Flag}TTR in the cytosol. Additionally, the sensitivity of HSPA13 knockout cells to the inhibitor of signal peptidase suggests that HSPA13 may be associated with Sec11 signal peptidase complex. Therefore, we set out to unravel how HSPA13 might participate in regulatory functions that impact protein translocation. To isolate potential interactors of HSPA13, we employed affinity purification mass spectrometry (AP-MS) approach. To immunoprecipitate HSPA13, we appended the Flag tag to the C-terminus to obtain HSPA13^{Flag}. Since regulatory interactors are transient, we used crosslinking to stabilize the proteins that interact with HSPA13. We transfected HEK293T cells with GFP or HSPA13^{Flag}, crosslinked with DSP, and immunopurified HSPA13^{Flag} from lysates. In a pilot experiment, we discovered a Sec61 interactor that notably interacted with HSPA13, B-cell receptor-associated protein (BCAP31) (Wang 2008) (Figure 2.5.A). This interaction was used to optimize the crosslinking conditions for following experiments (Figure 2.5.B). We found that at 2 mM DSP, we were able to pull down the most BCAP31; thus, we used this concentration for crosslinking before co-immunoprecipitation of HSPA13. Once the interactors of HSPA13 were isolated, they were precipitated by methanol-chloroform precipitation, digested into peptides and subjected to quantitative proteomics (Figure 2.6.A). We determined the fold change of potential interactors to be two-fold with FDR < 0.05. We observed a large number of nonspecific interactors possibly due to high

concentration of DSP used; additionally, overexpression of HSPA13 may subject a portion of the proteome to degradation since we observed lower number of preys. Other identified interactors of HSPA13 associated with Sec61 translocon include subunits of the signal peptidase (SP) such as SP catalytic subunits (SPCS2) and Sec11C (Figure 2.6.C). These interactors are responsible for the inefficiency of signal sequence cleavage of TTR. Other identified components such as dolichyldiphosphooligosaccharide-protein glycosyltransferase subunits (STT3A and STT3b) are part of the oligosaccharyl transferase (OST) complex that is involved in N-linked glycosylation. Noticeably, STT3A and STT3B were identified with unique peptides to each protein. Besides these strong interactors, we also identified a few more potential interactors that provide insights about HSPA13 being the regulator for maturation of its substrates. Those interactors including hypoxia upregulated protein 1 (HYOU1), GRP78/BiP and peptidyl-prolyl cis-trans isomerase B (PPIB) belong to a large ER-residing chaperone complex (Meunier 2002). HYOU1 is a chaperone and a NEF for BiP (Andreasson 2010, Inoue 2015). PPIB also participates in protein folding due to its proline isomerase activity (Gothel 1999). Interaction of HSPA13 with these components suggested HSPA13 associating with signal sequence cleavage and glycosylation events.

A


HEK293T; IP:FLAG
Hspa13 + −
Hspa13-FLAG − +
25-
IB: BCAP31
50-
IB: Hspa13
50-
25-
IB: FLAG

B

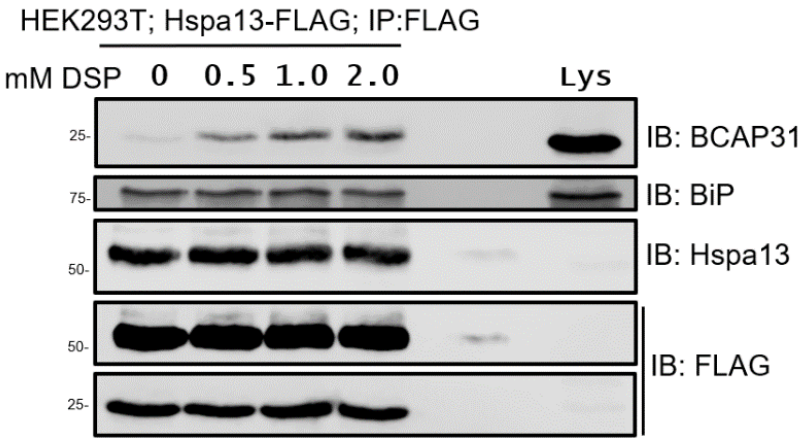

HEK293T; Hspa13-FLAG; IP:FLAG
mM DSP 0 0.5 1.0 2.0 Lys
25-
IB: BCAP31
75-
IB: BiP
50-
IB: Hspa13
50-
25-
IB: FLAG

Figure 2.5. IP optimization of HSPA13^{Flag}. A) Western blot of immunoprecipitations (IP) of lysates from HEK293T cells transfected with either HSPA13 or HSPA13^{Flag} and crosslinked with 1 mM dithiobis(succinimidyl propionate) (DSP). B) Western blot of eluates of IP from cells expressing HSPA13^{Flag} resolved on a 10% acrylamide SDS-PAGE gel. Cells were treated with variable concentrations of DSP prior to IP. Lysate (Lys) was run on the rightmost lane for comparison. Blot was probed with the indicated antibodies. Upper FLAG slice shows HSPA13^{Flag}; lower slice is the anti-FLAG light chain.

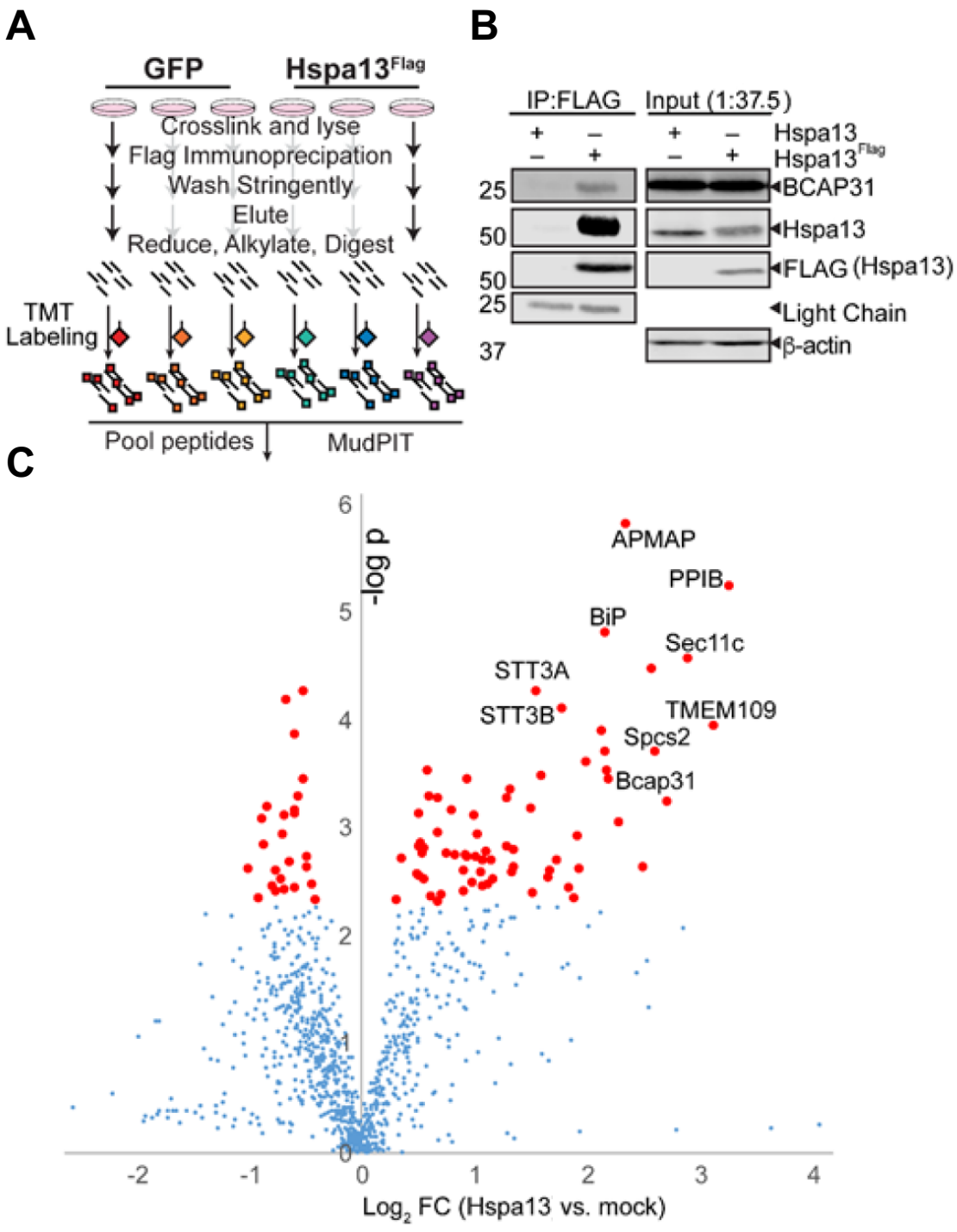
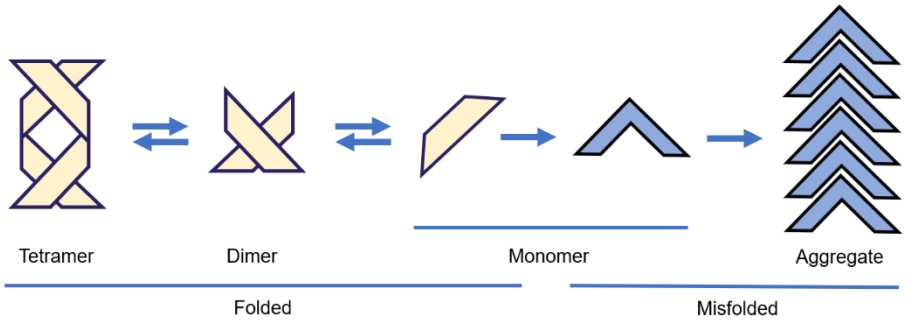

A
GFP
Hspa13Flag
Crosslink and lyse
Flag Immunoprecipation
Wash Stringently
Elute
Reduce, Alkylate, Digest
TMT Labeling
Pool peptides
MudPIT
B
IP:FLAG
Input (1:37.5)
Hspa13
Hspa13Flag
25
50
50
25
37
BCAP31
Hspa13
FLAG (Hspa13)
Light Chain
β-actin
C
-log p
APMAP
PPIB
BiP
Sec11c
STT3A
STT3B
TMEM109
Spcs2
Bcap31
Log2 FC (Hspa13 vs. mock)

Figure 2.6: Quantitative Proteomic Analysis of HSPA13 interactors. A) Schematic of immunoprecipitation-based identification of HSPA13 interactors. HEK293T cells transfected with GFP mock or $HSPA13^{Flag}$ expression vectors were crosslinked and lysed for IP. The eluates were labeled with tandem mass tag (TMT), pooled and prepared for LC-MS utilizing multidimensional protein identification technology (MudPIT). B) Western blot of eluates acquired during procedure shown in A. Immunoprecipitations (IP) are shown on the left slices; input lysates on the right. C) Volcano plot of identified proteins co-immunoprecipitated with $HSPA13^{Flag}$.

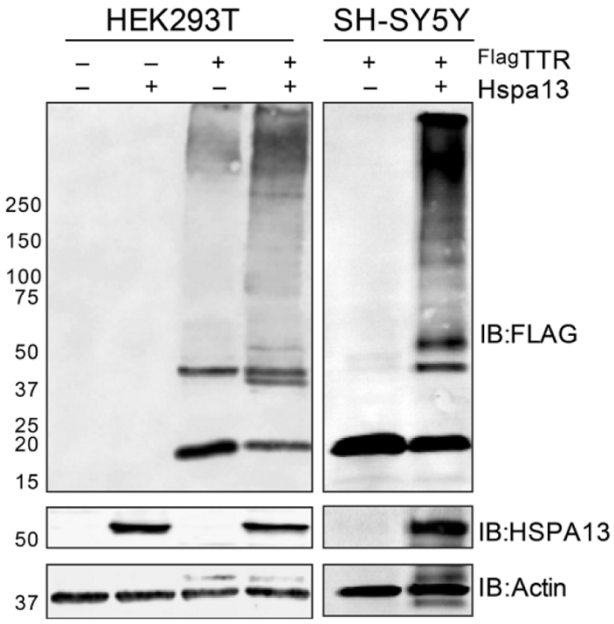
Mistargeted $^{Flag}TTR^{WT}$ in cytosol is aggregation prone

Transthyretin is an amyloidogenic protein present in the plasma or CSF (Bulawa 2012, Gasperini 2012). The aggregation is likely due to the instability of the tetramer caused by mutation or accumulation of extracellular TTR^{WT}. In this study, we used TTR^{WT} as the model since it is an ER substrate. TTR^{WT} is more kinetically and thermodynamically stable than most of the mutants (Sekijima 2005); thus, we would not expect to observe intracellular aggregates of TTR over the course of 48 hours of the experiments. However, the failure of ER translocation of TTR upon HSPA13 overexpression accumulates TTR in the cytosol. The reducing environment of the cytosol and the lack of chaperones might not be ideal for proper protein folding, which can increase the amyloidogenicity of TTR. Additionally, HSPA13 was observed to enhance the pathogenicity of another amyloidogenic protein PrP. To evaluate our hypothesis, we transfected $^{Flag}TTR^{WT}$ to HEK293T cells with or without cotransfection of HSPA13. Cell lysates were denatured and reduced as described, and were separated on a 4-20% gradient gel. Inherently, TTR showed its aggregates as a high molecular smear above 50 kDa (Figure 2.7.B, lane 3). HSPA13 overexpression increases the intensity of the smear and lowers the intensity of the monomeric TTR, indicating that the aggregation is further enhanced in the presence of HSPA13 overexpression (Figure 2.7.B, lane 4). Noticeably, the aggregates of TTR are extremely stable after being reduced and boiled for up to 20 minutes.

A


Tetramer
Dimer
Monomer
Aggregate
Folded
Misfolded

B


HEK293T
SH-SY5Y
FlagTTR
Hspa13
250
150
100
75
50
37
25
20
15
IB:FLAG
50
IB:HSPA13
37
IB:Actin

C

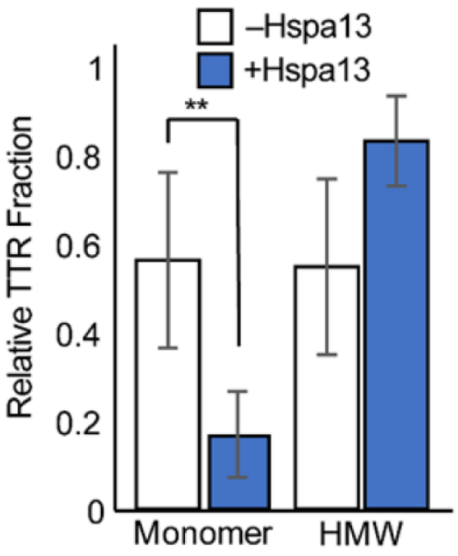
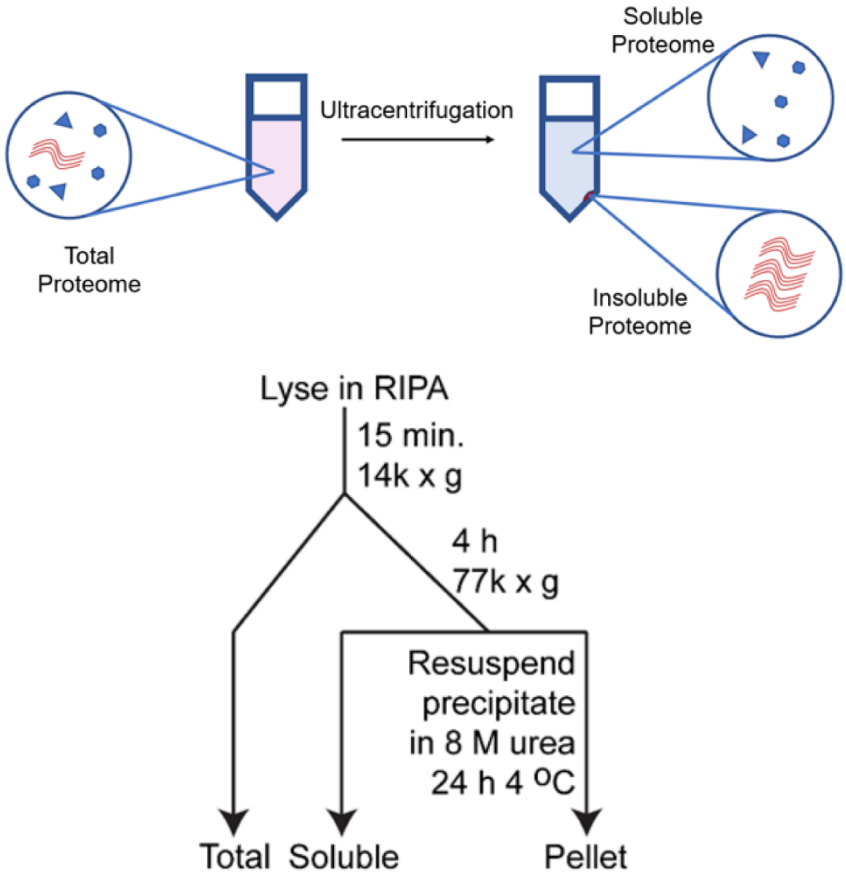

–Hspa13
+Hspa13
**
Relative TTR Fraction
1
0.8
0.6
0.4
0.2
0
Monomer
HMW

Figure 2.7: Aggregation of TTR under HSPA13 Coexpression. A) Graphic of TTR aggregation for familial and systemic diseases. TTR in normal conformation (tan trapezoids) form tetramers to shuttle hormones. Tetramers dissociate into stable dimers, then monomeric TTR. The monomers can misfold into an aggregation-prone conformation (blue arrowheads) and form aggregates. B) Western blot of lysates from cells expressing ^{Flag}TTR with or without HSPA13 co-transfections. Lysates of HEK293T and SH-SY5Y cells were resolved on a 4-20% acrylamide gradient gel and processed for Western blot. C) Densitometry quantification of TTR monomers and high molecular weight (HMW) smears from lysates of HEK293T cells transfected with ^{Flag}TTR and HSPA13 expression vectors (n=8; **, $p < 0.005$). Error bars are standard error of mean by two-way Student's t-test.

We repeated the same experiment on another cell line, SH-SY5Y neuroblastoma cells, and observed the same result. The aggregation of TTR with HSPA13 overexpression is not specific in HEK293T but also applicable to neuronal cell lines. The appearance of aggregates as smears can be visualized on gradient gels; however, it is not robust for quantification. Therefore, we employed ultracentrifugation to partition the soluble and insoluble proteomes from total lysates. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ with or without HSPA13. The total lysates of the cells were normalized and subjected to ultracentrifugation for 77,000 x g for 4 hours at 4^{o}C. The soluble supernatant was saved and the insoluble pellet was rinsed 3X with RIPA and incubated in 8 M urea for 24 hours to solubilize aggregation-prone TTR. Subsequently, the total lysates prior to ultracentrifugation, the soluble and insoluble fractions were analyzed on SDS-PAGE and Western blots. In the absence of HSPA13 overexpression, most of the TTR partitioned into soluble fraction and barely TTR was present in the pellet (Figure 2.8.B, lane 1-3). However, in the presence of HSPA13 overexpression, we observed similar amounts of TTR in the total lysate and the soluble fraction but there was an increase in the insoluble TTR in the pellet (Figure 2.8.B, lane 4-6). Importantly, the soluble TTR is mostly the mature band while insoluble TTR appears to be the immature band, indicative of immature TTR being aggregation-prone. This result is associated with previous results about the effect of HSPA13 on impairment of ER translocation of TTR, which eventually leads to aggregation of mistargeted TTR and potentially can be harmful for the cells.

A


Soluble
Proteome
Ultracentrifugation
Total
Proteome
Insoluble
Proteome
Lyse in RIPA
15 min.
14k x g
4 h
77k x g
Resuspend
precipitate
in 8 M urea
24 h 4 °C
Total
Soluble
Pellet

B

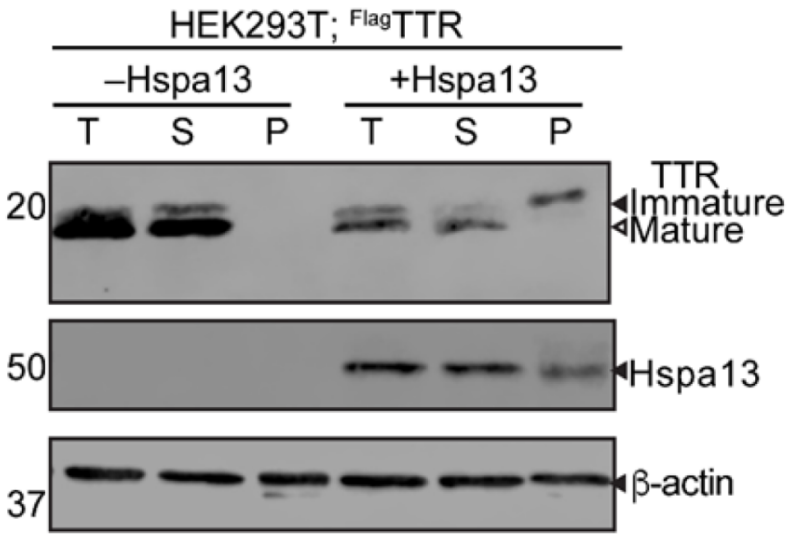

HEK293T; FlagTTR
–Hspa13
+Hspa13
T
S
P
T
S
P
20
TTR
Immature
Mature
50
Hspa13
37
β-actin

Figure 2.8: Ultracentrifugation assay to isolate insoluble proteins. A) Schematic of ultracentrifugation strategy. Lysates subjected to 77,000 x g for 4 hours are separated into supernatant and pellet. Protein is solubilized and extracted from the pellet with 8 M urea. B) Western blot of fractions resulting from ultracentrifugation of HEK293T cell lysates, with 2.7% of total volume loaded for total lysate and supernatant; 33% of the insoluble pellet. Cells were transfected with ^{Flag}TTR, with or without HSPA13 co-transfection. Total lysate (T), supernatant (S) and pellet (P) of each treatment were resolved on a 15% acrylamide gel for SDS-PAGE, followed by near-IR fluorescent Western blot with the indicated antibodies.

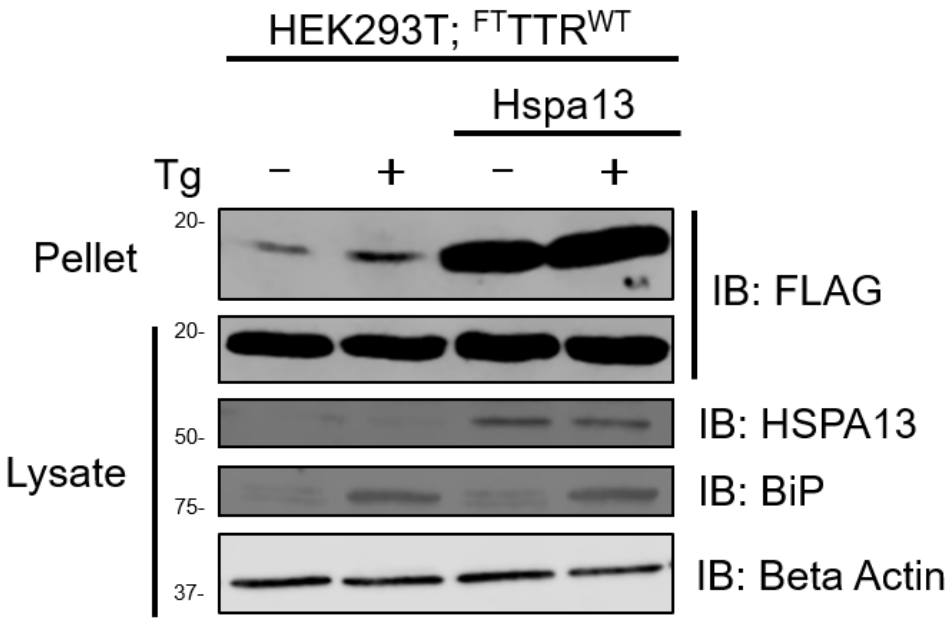
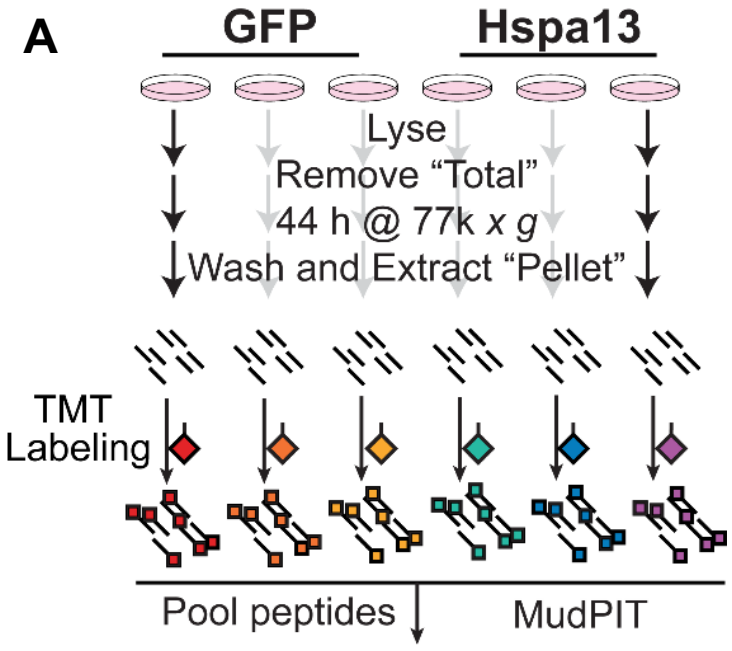


Figure 2.9: Insolubility of TTR is specific to its maturation defect. Ultracentrifugation of proteins under ER-PQC. HEK293T cells were transfected for ^{Flag}TTR expression with or without HSPA13. Thapsigargin (Tg) was treated for 16 hours before harvest. Cell lysates were subjected to ultracentrifugation and processed for Western blot, with 66% of the solubilized pellet loaded versus 2.6% of the total lysate.

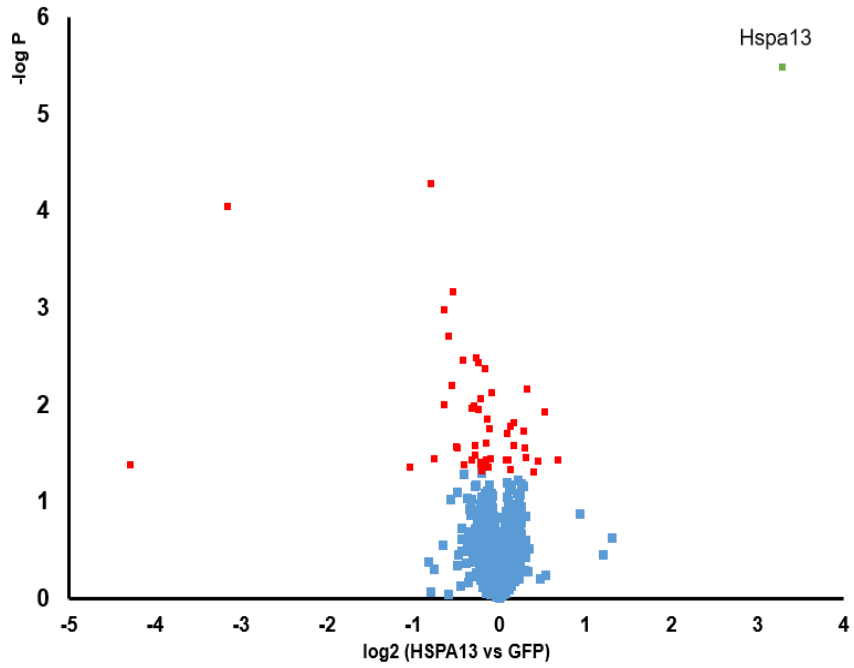
HSPA13 is regulated by UPR activation upon ER stress. It was observed in earlier study that during ER stress, ER translocation gets attenuated as a pQC activity to lower the burden of protein loads in the ER (Kang 2006). To evaluate whether our results are due to UPR activation, we overexpressed $^{Flag}TTR^{WT}$ in HEK293T cells and treated with 1 μM of thapsigargin (Tg) to induce ER stress (Sehgal 2017). Tg impairs the sarcoendoplasmic reticulum calcium transport ATPase (SERCA) pump, depleting calcium inside the ER. Two of the major ER chaperones, calnexin and calreticulin, require calcium binding for their function. Therefore, Tg treatment increases protein misfolding in the ER, activating the UPR. We observed that Tg does not produce TTR aggregates at the same intensity as HSPA13 did (Figure 2.9). We concluded that overexpression of HSPA13 induces intracellular aggregation of TTR in the absence of stress.

Effects of HSPA13 on total cellular proteome and insoluble proteome

Our previous results showed that HSPA13 inhibits ER translocation of TTR and leads to aggregation of mistargeted TTR in the cytosol. We investigated whether overexpression of HSPA13 has any effects on the cellular proteome. HEK293T cells were transfected with the plasmid expressing HSPA13 or GFP control. The total lysates and the insoluble pellets dissolved in 8 M urea for 96 hours were subjected to quantitative proteomics as described earlier. There were possible hits whose FDR was lower than 0.05, but the FC was not high enough to be


A
GFP
Hspa13
Lyse
Remove "Total"
44 h @ 77k x g
Wash and Extract "Pellet"
TMT
Labeling
Pool peptides
MudPIT

B


-log P
6
5
4
3
2
1
0
-5
-4
-3
-2
-1
0
1
2
3
4
log2 (HSPA13 vs GFP)
Hspa13

C

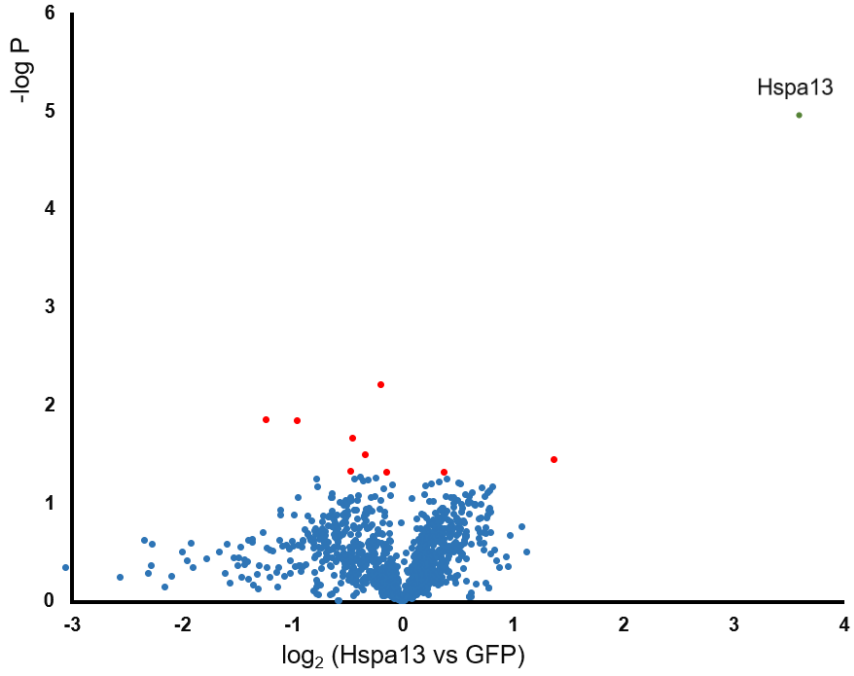

-log P
6
5
4
3
2
1
0
-3
-2
-1
0
1
2
3
4
log2 (Hspa13 vs GFP)
Hspa13

Figure 2.10: Quantitative Proteomic Analysis of Total Cellular Lysates and Insoluble Fractions. A) Schematic of liquid chromatography mass spectrometry (LC-MS) coupled with multidimensional protein identification technology (MudPIT). HEK293T cells transfected with either GFP or HSPA13 expression vectors were lysed and the protein solutions were subjected to ultracentrifugation at 77,000 x g for 4 hours. The replicate protein pellets were extracted in 8 M urea then labeled with tandem mass tags (TMT) and pooled for LC-MS and MudPIT. B,C) Volcano plots showing the abundance ratios of identified proteins in total lysates (B) and insoluble pellets (C). Proteins with significant change ($p < 0.05$) are labeled in red, while proteins having FDR < 0.05 are labeled in green.

considered true results of the assay. Thus, we did not observe any proteins that are upregulated or downregulated except the overexpression of HSPA13 (Figure 2.10.B). Also, the ultracentrifugation assay did not reveal any proteins that demonstrated differential solubility with HSPA13 overexpression except insoluble HSPA13 itself due to overexpression (Figure 2.10.C). Overall, there was no noticeable difference in the cellular proteome as well as the insoluble proteome with HSPA13 overexpression. Perhaps our cell line and the experimental setup were not appropriate to observe any regulation in the proteome of the cells. We did not inhibit protein degradation in the experiment after transfecting the cells with HSPA13; therefore, the proteins that were regulated by HSPA13 might have been degraded. Furthermore, HSPA13 overexpressing cells were cultured for 48 hours, which was not long enough for intracellular aggregates to accumulate.

ATPase mutation of HSPA13 enhances its effect on ER translocation

As mentioned earlier, HSPA13 contains an NBD domain but does not have the SBD necessary for binding substrates as other HSP70s. We assume that the function of HSPA13 largely depends on its ATPase activity on the NBD, including the effect on ER translocation. In *E.coli* DnaK, lysine 70 mutated to alanine (K70A) removed its hydrolytic activity and threonine 199 mutated to alanine (T199A) abolished its autophosphorylation (Buchberger 1995, Barthel 2001). These residues are conserved in HSPA13; therefore, we mutated lysine 100 to

alanine (K100A) and threonine 230 to alanine (T230A) on HSPA13. These mutants would lack the ATPase activity observed in HSP70.

We hypothesized that these mutants would not have any effects on the $^{Flag}TTR^{WT}$ as observed in wild-type HSPA13. We cotransfected $^{Flag}TTR^{WT}$ with wild-type HSPA13, $HSPA13^{K100A}$ and $HSPA13^{T230A}$ to evaluate whether TTR is aggregation-prone with the overexpression of HSPA13 mutants. Surprisingly, on the gradient gel, we still observed high molecular smears of $^{Flag}TTR^{WT}$ with $HSPA13^{K100A}$ and $HSPA13^{T230A}$ coexpression (Figure 2.11.B, top slice). Additionally, the inhibitory effect of ER translocation on TTR was more severe with the mutants of HSPA13 (Figure 2.11.B, second slice). In particular, $HSPA13^{K100A}$ overexpression resulted in more immature TTR compared to wild-type HSPA13, and the monomeric TTR appears to be mostly immature under the expression of $HSPA13^{T230A}$. This observation completely opposed to what we expected, which showed that the ATPase deficient mutants of HSPA13 exacerbated the inhibitory effect of ER translocation on ER substrates. We also performed ultracentrifugation assay to further validate our results on the gradient gels, and the results were consistent. In the absence of HSPA13 overexpression, TTR is mostly soluble (Figure 2.11.C, lane 1-3). The overexpression of wild-type HSPA13 promoted an amount of insoluble TTR which is mostly comprised of immature TTR (Figure 2.11.C, lane 4-6). The promotion of aggregating TTR is more noticeable with $HSPA13^{K100A}$ since we observed more immature TTR in the insoluble pellet. Especially, $HSPA13^{T230A}$ produces mostly immature TTR

A

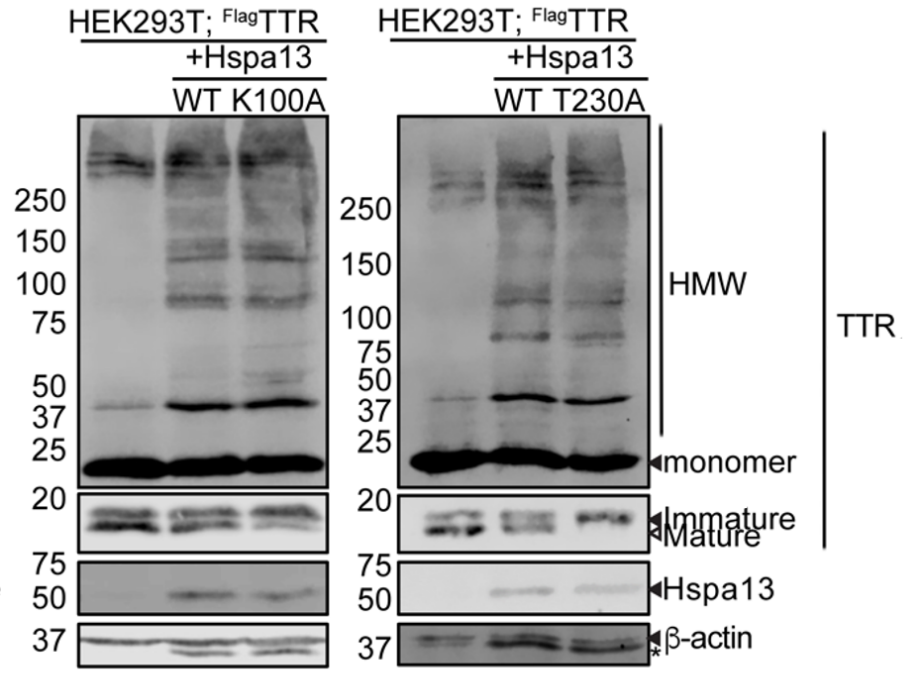
K100

DnaK	64-NTLFAIKRLIGR-75
BiP	90-NTVFDAKRLIGR-101
HSPA13	64-NTIYDAKRFIGK-105

T230

DnaK	104-DLGGGTFDISII-205
BiP	224-DLGGGTFDVSLL-235
HSPA13	225-DLGGGTLDVSLL-236

B



C

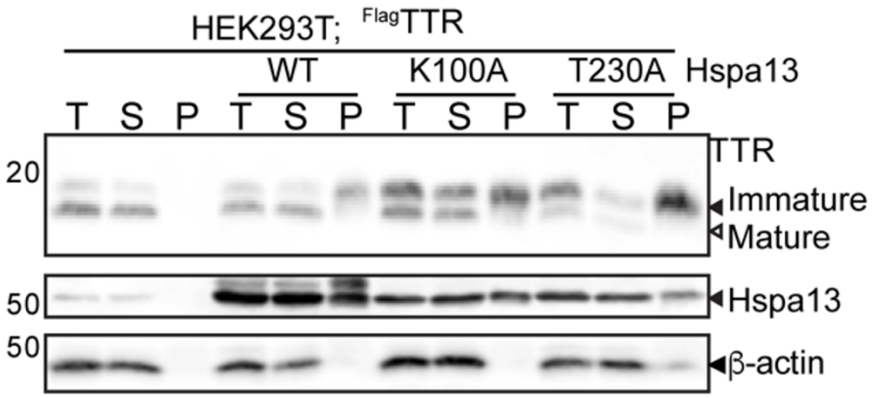

HEK293T; FlagTTR
WT
K100A
T230A
Hspa13
T S P T S P T S P T S P
20
TTR
Immature
Mature
50
Hspa13
50
β-actin

Figure 2.11: ATPase mutation of HSPA13. A) Protein alignments of HSPA13 (bottom rows) to Hsp70s including human BiP/GRP78 (middle rows) and *E. coli* DnaK (top rows). Top panel shows conserved lysine necessary for ATP hydrolysis, bottom panel shows conserved threonine needed for autophosphorylation. B) Western blot of homogenous 12% (lower three slices) and 4-20% acrylamide gradient gels (top slices) and from HEK293T cells transfected with ^{Flag}TTR and HSPA13 variants consisting of wild-type (WT), K100A (left) and T230A (right). High molecular weight ^{Flag}TTR (HMW) ladders are shown in the gradient gel slice, with monomeric ^{Flag}TTR marked with the black arrowhead. Monomeric ^{Flag}TTR was separated on the 12% gel into the immature (black arrowhead) and mature (white arrowhead) forms. Asterisks points to the dimer of TTR. C) Western blot of different fractions pre- and post-ultracentrifugation from HEK293T cells transfected with ^{Flag}TTR and one of the HSPA13 wild-type and variants. Each treatment provided a total (T) lysate, a soluble supernatant (S), and an insoluble pellet (P).

partitioning into the insoluble pellet and a significant drop in the amount of TTR present in the soluble fraction. Thus, we concluded that ATPase mutants of HSPA13 increases the inhibitory effects of ER translocation for TTR and the mechanism of this effect remains unknown. Possibly, the conformation induced by the ATPase depletion can enhance the interaction with the translocon to further inhibits ER translocation of TTR.

Effects of HSPA13 on Secretion and Maturation of ER Substrates

We observed that HSPA13 lowers ER translocation and maturation of $^{Flag}TTR^{WT}$, which resulted in mistargeted TTR in cytosol. Due to being in a different cellular compartment, TTR tend to be misfolded and becomes aggregates. We also interrogated other downstream effects such as secretory capacity of $^{Flag}TTR^{WT}$ and glycosylation and secretion of ^{Flag}A1AT.

Besides being the hub for a third of the proteome in the cell, the ER also serves as an important entry of the canonical secretory pathway. Once in the ER, the proteins are properly folded, packaged into vesicles and transported to the Golgi apparatus. Presumably, the inhibitory effect of HSPA13 on protein translocation also affects secretion. We measured the amount of extracellular $^{Flag}TTR^{WT}$ from the cells with and without overexpression of wild-type HSPA13 or HSPA13^{T230A} and found that HSPA13 significantly decreases the secreted TTR (Figure 2.12.A). Additionally, due to inhibited TTR maturation in the cells, we also looked

at its maturity after TTR is secreted. To obtain adequate signal, we immunoprecipitated the $^{Flag}TTR^{WT}$ from the conditioned media, followed by elution in Laemli buffer and denaturation in 16.7 mM DTT, we noticed the immature TTR also being secreted from the cells (Figure 2.12.A). We assume the immature TTR that is mistargeted to the cytosol being exported out of the cells without bypassing the Golgi (Rabouille 2017, Volkmar 2016, Lee 2016). Additionally, there is a possibility that the cells were lysed during the incubation, releasing the immature TTR.

Because HSPA13 affects TTR secretion, we considered looking at alpha-1 antitrypsin (A1AT). Unlike TTR, A1AT is glycosylated, which enabled us to investigate another ER-specific post-translational modification besides the cleavage of signal sequence. Impaired ER proteostasis could also lead to impairment of protein modification in the ER. We also considered that the signal sequence affects glycosylation efficiency (Chen 2001, Rutkowski 2003). Also, different signal sequences can vary ER translocation efficiency (Levine 2004). Therefore, we constructed two different signal sequences on the plasmids expressing ^{Flag}A1AT, which are the native signal sequence of A1AT (ssA1AT-^{Flag}A1AT) or the preprotrypsin signal sequence (ssPPT-^{Flag}A1AT) (Figure 2.12.B). We overexpressed ssA1AT-^{Flag}A1AT or ssPPT-^{Flag}A1AT in HEK293T cells, along with HSPA13, HSPA13^{K100A} or HSPA13^{T230A} cotransfection. The secreted A1AT was immunopurified in the same manner as $^{Flag}TTR^{WT}$. We also included the

PNGase F to remove modified glycans, resulting in smaller proteins. When untreated, both types of ^{Flag}A1AT appeared as a single band at greater than 50 kDa (Figure 2.12.C; top, lane 1). However, treatment with PNGase F shifts the ^{Flag}A1AT band to less than 50 kDa (Figure 2.12.C; bottom, lane 2), allowing for distinguishing between glycosylated and unglycosylated ^{Flag}A1AT. We observed that secreted unglycosylated ssA1AT-^{Flag}A1AT appears with HSPA13^{K100A} and HSPA13^{T230A} overexpression (Figure 2.12.C, lane 5-8). In comparison, unglycosylated ssPPT-^{Flag}A1AT is secreted with overexpression with all HSPA13 variants (Figure 2.12.C, lane 3-8). Interestingly, HSPA13^{T230A} also lowers ssA1AT-^{Flag}A1AT expression (Figure 2.12.D; lane 4), indicating ER proteostatic imbalance. Thus, HSPA13 impacts ER proteostasis and has inhibitory effects on signal peptide cleavage and glycosylation.

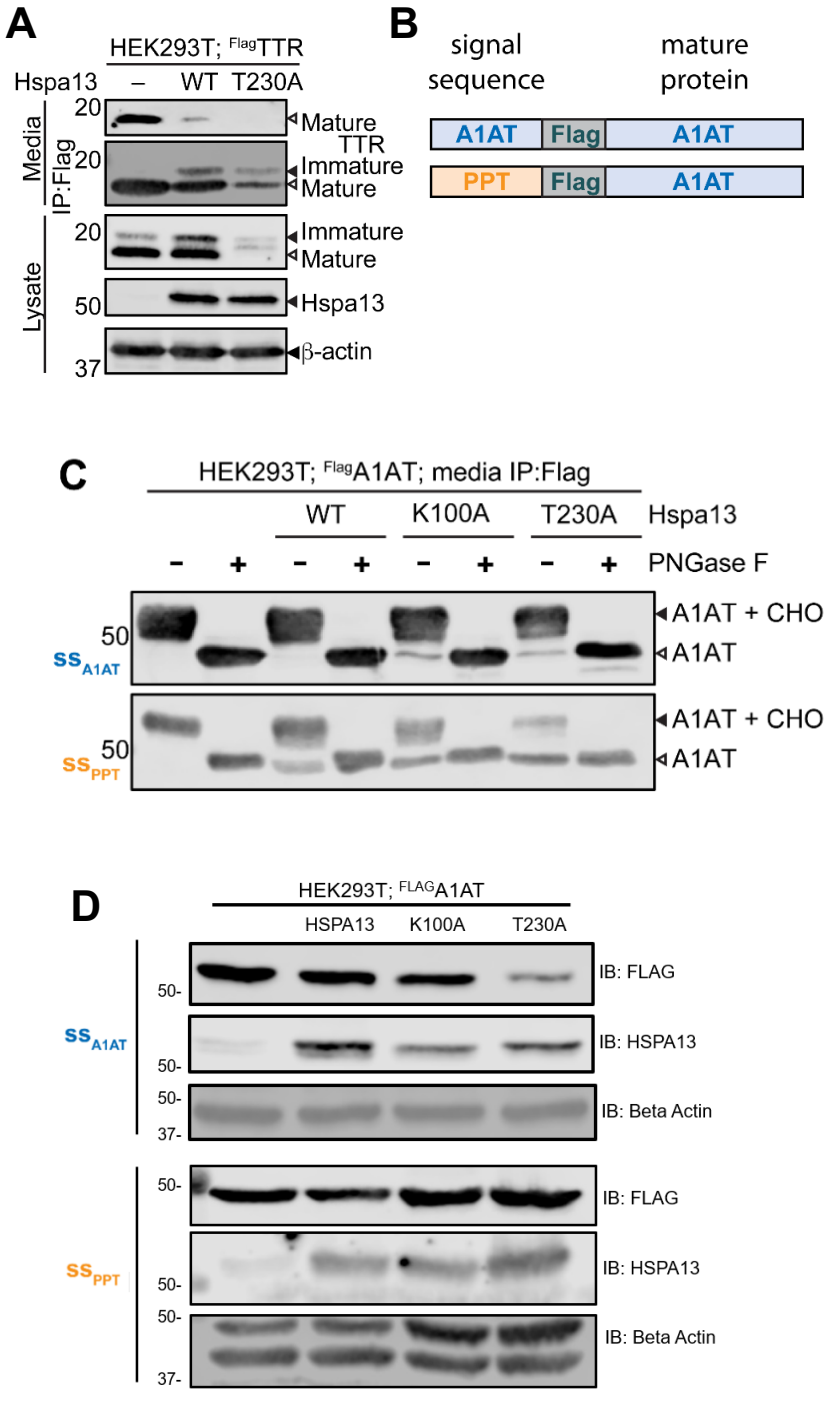

A
HEK293T; FlagTTR
Hspa13 – WT T230A
Media
IP:Flag
Lysate
20
20
20
50
37
Mature TTR
Immature
Mature
Immature
Mature
Hspa13
β-actin
B
signal sequence
mature protein
A1AT
Flag
A1AT
PPT
Flag
A1AT
C
HEK293T; FlagA1AT; media IP:Flag
WT K100A T230A Hspa13
- + - + - + - + PNGase F
50
SSA1AT
A1AT + CHO
A1AT
50
SSPPT
A1AT + CHO
A1AT
D
HEK293T; FLAGA1AT
HSPA13 K100A T230A
SSA1AT
50-
50-
50-
37-
IB: FLAG
IB: HSPA13
IB: Beta Actin
SSPPT
50-
50-
50-
37-
IB: FLAG
IB: HSPA13
IB: Beta Actin

Figure 2.12: Effects of HSPA13 on secretion and glycosylation of ER substrates. A) HEK293T cells expressing ^{Flag}TTR with or without HSPA13 were cultured for both lysate and media collection. Media input (top slice), media immunoprecipitated (IP) eluate (second from top) and cellular lysate were resolved on 12% acrylamide SDS-PAGE gel and processed for Western blot. Antibodies used for probing indicated. White arrowheads denote mature ^{Flag}TTR; black arrowheads for immature ^{Flag}TTR. B) Diagram of FLAG-tagged alpha 1 antitrypsin (^{Flag}A1AT) variants with top containing native A1AT signal peptide and bottom containing preprotrypsin (PPT) signal peptide. C) Western blot of IP material from media incubated with HEK293T cells transfected with either A1AT, with or without HSPA13. Media was treated with PNGase F or mock after IP. D) Western blot of intracellular ssA1AT-^{Flag}A1AT and ssPPT-^{Flag}A1AT from the same HEK293T cells that the cultured media were collected from. Samples were resolved on a 10% acrylamide SDS-PAGE gel for Western blotting with anti-FLAG primary antibodies. N-linked glycosylated (+ CHO) forms of A1AT indicated by black arrowheads; unglycosylated A1AT with white arrowheads.

DISCUSSION

In this study, we aimed to investigate the role as well as the functionality of a truncated HSP70. HSPA13 (STCH) is conserved across Eumetazoa, implying an essential role that has not been identified. We used an ER substrate TTR with appended N-terminal flag tag as a model to interrogate its ER translocation under the effect of HSPA13. Overexpression of HSPA13 inhibits maturation of ^{Flag}TTR into the mature form, in which the signal sequence is cleaved off. The immature TTR containing signal sequence is, therefore, shifted to localize in the cytosol, which was confirmed in out proximity labeling assay. Apparently, the inhibitory effect of HSPA13 on ER translocation is likely to lead to the worsening of prion disease (Rane 2010, Grizenkova 2012). The mistargeted TTR is sensitive to proteasomal degradation, which was also observed in mistargeted protein from the mitochondria (Wrobel 2015, Topf 2016). Since cytosol is a reducing environment, mistargeted proteins are likely to misfold and become toxic. As expected, immature TTR aggregates under overexpression of HSPA13. Perhaps, HSPA13 also lowers the translocation of PrP, rendering its aggregation and toxicity. Consequently, we also observed significantly lowered secreted TTR due to its failure to enter the ER to begin with and also, the ER proteostasis might be disrupted.

The observed effects indicated that HSPA13 is a translocation inhibitor; therefore, it has to associate with other translocation complexes. In earlier

studies, HSPA13 is shown to be membrane-associated by BioID and association with signal peptidase (Go 2019, Estoppey 2017). Our AP-MS results revealed a few interactors that are part of the translocon and the signal peptidase system. We identified the translocon alpha subunit Sec61A1 as an interactor. However, we did not identify the beta and gamma subunits of the Sec61 translocon, probably due to their localization on the cytosolic side of the ER membrane (Mavylutov 2018). Other identified interactors that associate with the membrane and are proximal to Sec61 include BCAP31 (Wang 2008), the signal peptidase complex subunits, oligosaccharyl transferase. This result further supported the localization of HSPA13 in proximity to the translocon; thus, it is partially influential in modulation of ER translocation either by itself or by direct interaction with Sec61. Furthermore, HSPA13 interacting with the signal peptidase complex indicated its effect on the maturation of TTR. The secreted immature TTR can also result from its bypass of ER proteostasis or the activity of the signal peptidase gets inhibited by HSPA13. Another possibility is that the mistargeted TTR in the cytosol got recruited to the surface of the ER for deubiquitylation. These deubiquitylation cargos are encapsulated into ER-associated late endosomes to be exported extracellularly (Volkmar 2016, Lee 2016).

We also observed unglycosylated A1AT secreted from the cells transfected with HSPA13 wild-type and mutants. This result can be related to HSPA13 interacting with STT3A and STT3B subunits of the OST complex, as well as the OST-

associated magnesium transporter 1 (MAGT1). The OST complex located nearby signal peptidase is largely responsible for post-translational N-linked glycosylation. This modification serves as a marker for processing the proteins in the ER. The glycans added to the proteins can affect the fate of the proteins to be secreted or degraded. Our result indicated that HSPA13 has an inhibitory effect on glycosylation of A1AT and this pattern also depended on the signal sequences of A1AT. The signal peptidase complex and the OST complex cooperate to properly process glycoproteins (Chen 2001), and since HSPA13 appears to interact with both complexes, we observed the inhibitory effects on signal sequence cleavage and post-translational glycosylation of different ER substrates. Unfortunately, we did not observe any difference in the expression as well as solubility of endogenous proteins in HEK293T cells with the overexpression of HSPA13, implying the effect of HSPA13 on a limited number of proteins. The effects of HSPA13 identified in this study are mostly on the secretory proteins, perhaps in some other cell lines, we can observe its effects on the proteome. For example, HepG2 expressing endogenous TTR and other secretory protein can be a good model for the investigation of the aggregation of the proteome. Also, the neuronal cell lines that grow more slowly than HEK293T can accumulate the aggregates over a period of time that can be identified with HSPA13 overexpression.

The mutations on the ATPase activity of HSPA13 increased inhibition of all three observed processed: ER translocation, cleavage of signal sequence, and glycosylation. Therefore, it is likely that ATP-bound HSPA13 acts as a negative regulator. This conformation might tightly bind to complexes such as the translocon, the OST and the signal peptidase complex. However, further investigation is needed to confirm this hypothesis such as comparing the interactome of $HSPA13^{WT}$ and HSPA13 mutants.

REFERENCES

1. Ando, Y. & Suhr, O. B. Autonomic dysfunction in familial amyloidotic polyneuropathy (FAP). *Amyloid* **5**, 288–300 (1998).

2. Andréasson, C., Rampelt, H., Fiaux, J., Druffel-Augustin, S. & Bukau, B. The Endoplasmic Reticulum Grp170 Acts as a Nucleotide Exchange Factor of Hsp70 via a Mechanism Similar to That of the Cytosolic Hsp110. *J. Biol. Chem.* **285**, 12445–12453 (2010).

3. Aoki, M. *et al.* A genetic variant in the gene encoding the stress70 protein chaperone family member STCH is associated with gastric cancer in the Japanese population. *Biochemical and Biophysical Research Communications* **335**, 566–574 (2005).

4. Aviram, N. & Schuldiner, M. Targeting and translocation of proteins to the endoplasmic reticulum at a glance. *J Cell Sci* **130**, 4079–4085 (2017).

5. Barthel, T. K., Zhang, J. & Walker, G. C. ATPase-Defective Derivatives of Escherichia coliDnaK That Behave Differently with Respect to ATP-Induced Conformational Change and Peptide Release. *Journal of Bacteriology* **183**, 5482–5490 (2001).

6. Bensellam, M., Gilon, P. & Jonas, J.-C. Physiological ER Stress: The Model of Insulin-Secreting Pancreatic b-Cells. in *Endoplasmic Reticulum Stress in Health and Disease* (eds. Agostinis, P. & Afshin, S.) 185–211 (Springer Netherlands, 2012).

7. Benson, M. D. & Roudebush, R. L. Familial amyloidotic polyneuropathy. *Trends in Neurosciences* **12**, 88–92 (1989).

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

9. Borgese, N. & Gaetani, S. In vitro synthesis and post-translational insertion into microsomes of the integral membrane protein, NADH-cytochrome b5 oxidoreductase. *The EMBO Journal* **2**, 1263–1269 (1983).

10. Buchberger, A. *et al.* Nucleotide-induced Conformational Changes in the ATPase and Substrate Binding Domains of the DnaK Chaperone Provide Evidence for Interdomain Communication. *J. Biol. Chem.* **270**, 16903–16910 (1995).

11. Bulawa, C. E. *et al.* Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *PNAS* **109**, 9629–9634 (2012).

12. Chen, X., VanValkenburgh, C., Liang, H., Fang, H. & Green, N. Signal Peptidase and Oligosaccharyltransferase Interact in a Sequential and Dependent Manner within the Endoplasmic Reticulum. *J. Biol. Chem.* **276**, 2411–2416 (2001).

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

14. Dreier, L., Hartmann, E. & Rapoport', T. A. Posttranslational Protein Transport in Yeast Reconstituted with a Purified Complex of Set Proteins and Kar2p. 10.

15. Estoppey, D. *et al.* The Natural Product Cavinafungin Selectively Interferes with Zika and Dengue Virus Replication by Inhibition of the Host Signal Peptidase. *Cell Reports* **19**, 451–460 (2017).

16. Gasperini, R. J., Klaver, D. W., Hou, X., Aguilar, M.-I. & Small, D. H. Mechanisms of Transthyretin Aggregation and Toxicity. in *Protein Aggregation and Fibrillogenesis in Cerebral and Systemic Amyloid Disease* (ed. Harris, J. R.) 211–224 (Springer Netherlands, 2012).

17. Go, C. D. *et al.* A proximity biotinylation map of a human cell. *bioRxiv* 796391 (2019)

18. Göthel, S. F. & Marahiel, M. A. Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *CMLS, Cell. Mol. Life Sci.* **55**, 423–436 (1999).

19. Grandjean, J. M. D. *et al.* Deconvoluting Stress-Responsive Proteostasis Signaling Pathways for Pharmacologic Activation Using Targeted RNA Sequencing. *ACS Chem. Biol.* **14**, 784–795 (2019).

20. Grizenkova, J. *et al.* Overexpression of the Hspa13 (Stch) gene reduces prion disease incubation time in mice. *PNAS* **109**, 13722–13727 (2012).

21. Grudnik, P., Bange, G. & Sinning, I. Protein targeting by the signal recognition particle. *Biological Chemistry* **390**, 775–782 (2009).

22. Halic, M. & Beckmann, R. The signal recognition particle and its interactions during protein targeting. *Current Opinion in Structural Biology* **15**, 116–125 (2005).

23. He, L., Diedrich, J., Chu, Y.-Y. & Yates, J. R. Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. *Anal. Chem.* **87**, 11361–11367 (2015).

24. Heinrich, S. U., Mothes, W., Brunner, J. & Rapoport, T. A. The Sec61p Complex Mediates the Integration of a Membrane Protein by Allowing Lipid Partitioning of the Transmembrane Domain. *Cell* **102**, 233–244 (2000).

25. Imai, Y., Ito, A. & Sato, R. Evidence for Biochemically Different Types of Vesicles in the Hepatic Microsomal Fraction. *The Journal of Biochemistry* **60**, 417–428 (1966).

26. Inoue, T. & Tsai, B. A Nucleotide Exchange Factor Promotes Endoplasmic Reticulum-to-Cytosol Membrane Penetration of the Nonenveloped Virus Simian Virus 40. *Journal of Virology* **89**, 4069–4079 (2015).

27. Janda, C. Y. *et al.* Recognition of a signal peptide by the signal recognition particle. *Nature* **465**, 507–510 (2010).

28. Jiang, H.-Y. & Wek, R. C. Phosphorylation of the α-Subunit of the Eukaryotic Initiation Factor-2 (eIF2α) Reduces Protein Synthesis and Enhances Apoptosis in Response to Proteasome Inhibition. *J. Biol. Chem.* **280**, 14189–14202 (2005).

29. Joesch, M. *et al.* Reconstruction of genetically identified neurons imaged by serial-section electron microscopy. *eLife* **5**, e15015 (2016).

30. Johnson, S. M. *et al.* Native State Kinetic Stabilization as a Strategy To Ameliorate Protein Misfolding Diseases: A Focus on the Transthyretin Amyloidoses. *Acc. Chem. Res.* **38**, 911–921 (2005).

31. Kadowaki, H. *et al.* Pre-emptive Quality Control Protects the ER from Protein Overload via the Proximity of ERAD Components and SRP. *Cell Reports* **13**, 944–956 (2015).

32. Kampinga, H. H. *et al.* Guidelines for the nomenclature of the human heat shock proteins. *Cell Stress and Chaperones* **14**, 105–111 (2009).

33. Kang, S.-W. *et al.* Substrate-Specific Translocational Attenuation during ER Stress Defines a Pre-Emptive Quality Control Pathway. *Cell* **127**, 999–1013 (2006).

34. Kowarik, M., Küng, S., Martoglio, B. & Helenius, A. Protein Folding during Cotranslational Translocation in the Endoplasmic Reticulum. *Molecular Cell* **10**, 769–778 (2002).

35. Lam, S. S. *et al.* Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2015).
36. Lee, J.-G., Takahama, S., Zhang, G., Tomarev, S. I. & Ye, Y. Unconventional secretion of misfolded proteins promotes adaptation to proteasome dysfunction in mammalian cells. *Nature Cell Biology* **18**, 765–776 (2016).

37. Levine, C. G., Mitra, D., Sharma, A., Smith, C. L. & Hegde, R. S. The Efficiency of Protein Compartmentalization into the Secretory Pathway. *MBoC* **16**, 279–291 (2004).

38. Lindner, A. B., Madden, R., Demarez, A., Stewart, E. J. & Taddei, F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proceedings of the National Academy of Sciences* **105**, 3076–3081 (2008).

39. Luirink, J. & Sinning, I. SRP-mediated protein targeting: structure and function revisited. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1694**, 17–35 (2004).

40. Mavylutov, T., Chen, X., Guo, L. & Yang, J. APEX2- tagging of Sigma 1-receptor indicates subcellular protein topology with cytosolic N-terminus and ER luminal C-terminus. *Protein Cell* **9**, 733–737 (2018).

41. Meunier, L., Usherwood, Y.-K., Chung, K. T. & Hendershot, L. M. A Subset of Chaperones and Folding Enzymes Form Multiprotein Complexes in Endoplasmic Reticulum to Bind Nascent Proteins. *MBoC* **13**, 4456–4469 (2002).

42. Meyer, H.-A. *et al.* Mammalian Sec61 Is Associated with Sec62 and Sec63. *J. Biol. Chem.* **275**, 14550–14557 (2000).

43. Moir, D. T. & Dumais, D. R. Glycosylation and secretion of human alpha-1-antitrypsin by yeast. *Gene* **56**, 209–217 (1987).

44. Ng, B., Connors, L. H., Davidoff, R., Skinner, M. & Falk, R. H. Senile Systemic Amyloidosis Presenting With Heart Failure: A Comparison With Light Chain–Associated Amyloidosis. *Arch Intern Med* **165**, 1425–1429 (2005).

45. Osborne, A. R., Rapoport, T. A. & van den Berg, B. Protein Translocation by the Sec61/Secy Channel. *Annual Review of Cell and Developmental Biology* **21**, 529–550 (2005).

46. Otterson, G. A. *et al.* Stch encodes the 'ATPase core' of a microsomal stress 70 protein. *EMBO J* **13**, 1216–1225 (1994).

47. Panzner, S., Dreier, L., Hartmann, E., Kostka, S. & Rapoport, T. A. Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* **81**, 561–570 (1995).

48. Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J. & Gygi, S. P. Evaluation of Multidimensional Chromatography Coupled with Tandem Mass Spectrometry (LC/LC−MS/MS) for Large-Scale Protein Analysis: The Yeast Proteome. *J. Proteome Res.* **2**, 43–50 (2003).

49. Rabouille, C. Pathways of Unconventional Protein Secretion. *Trends in Cell Biology* **27**, 230–240 (2017).

50. Rane, N. S., Chakrabarti, O., Feigenbaum, L. & Hegde, R. S. Signal sequence insufficiency contributes to neurodegeneration caused by transmembrane prion protein. *J Cell Biol* **188**, 515–526 (2010).

51. Rane, N. S., Kang, S.-W., Chakrabarti, O., Feigenbaum, L. & Hegde, R. S. Reduced Translocation of Nascent Prion Protein During ER Stress Contributes to Neurodegeneration. *Developmental Cell* **15**, 359–370 (2008).

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

53. Rapoport, T. A., Li, L. & Park, E. Structural and Mechanistic Insights into Protein Translocation. *Annual Review of Cell and Developmental Biology* **33**, 369–390 (2017).

54. Rutkowski, D. T., Ott, C. M., Polansky, J. R. & Lingappa, V. R. Signal Sequences Initiate the Pathway of Maturation in the Endoplasmic Reticulum Lumen. *J. Biol. Chem.* **278**, 30365–30372 (2003).

55. Sato, T. *et al.* STT3B-Dependent Posttranslational N-Glycosylation as a Surveillance System for Secretory Protein. *Molecular Cell* **47**, 99–110 (2012).

56. Satpute-Krishnan, P. *et al.* ER Stress-Induced Clearance of Misfolded GPI-Anchored Proteins via the Secretory Pathway. *Cell* **158**, 522–533 (2014).

57. Sehgal, P. *et al.* Inhibition of the sarco/endoplasmic reticulum (ER) Ca2+-ATPase by thapsigargin analogs induces cell death via ER Ca2+ depletion and the unfolded protein response. *J. Biol. Chem.* jbc.M117.796920 (2017).

58. Sekijima, Y. *et al.* The Biological and Chemical Basis for Tissue-Selective Amyloid Disease. *Cell* **121**, 73–85 (2005).

59. Topf, U., Wrobel, L. & Chacinska, A. Chatty Mitochondria: Keeping Balance in Cellular Protein Homeostasis. *Trends in Cell Biology* **26**, 577–586 (2016).

60. Tyedmers, J. *et al.* Homologs of the yeast Sec complex subunits Sec62p and Sec63p are abundant proteins in dog pancreas microsomes. *PNAS* **97**, 7214–7219 (2000).

61. Volkmar, N., Fenech, E. & Christianson, J. C. New MAPS for misfolded proteins. *Nature Cell Biology* **18**, 724–726 (2016).

62. Wang, B. *et al.* BAP31 Interacts with Sec61 Translocons and Promotes Retrotranslocation of CFTRΔF508 via the Derlin-1 Complex. *Cell* **133**, 1080–1092 (2008).

63. Washburn, M. P., Wolters, D. & Yates, J. R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* **19**, 242–247 (2001).

64. Westermark, P., Sletten, K., Johansson, B. & Cornwell, G. G. Fibril in senile systemic amyloidosis is derived from normal transthyretin. *PNAS* **87**, 2843–2845 (1990).

65. Wrobel, L. *et al.* Mistargeted mitochondrial proteins activate a proteostatic response in the cytosol. *Nature* **524**, 485–488 (2015).

66. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**, D506–D515 (2019).

CHAPTER 3: CYTOSOLIC QUALITY CONTROL OF MISTARGETED PROTEINS

INTRODUCTION

In general, cells contain multiple pathways to perform quality control (QC) to identify any errors and damages in protein synthesis. To maintain proteostasis, QC has to be active in every cellular compartment to rapidly recognize and correct or degrade the erroneous substrates such as misfolded proteins, mislocalized proteins and damaged proteins. Failure to perform corrections can lead to imbalance of proteostasis that affects the health of the cell overall. On the other hand, if QC systems are too promiscuous, then they can degrade or segregate folded proteins, depleting the cell of useful resources (Shao & Hegde 2016). For example, QC target loss-of-function includes a mutated protein with intact biochemical functions being prematurely recognized and degraded by the QC machinery, which leads to the loss of normal functions in the cell. This defect can cause degradation of important enzymes or signal transducers; for instance, the mutations of CFTR (cystic fibrosis transmembrane conductance regulator) result in cystic fibrosis disorder (Cheng 1990). In contrast, QC target gain-of-function occurs when a mutated protein that should be recognized and degraded happen to escape at a low rate; thus, these mutants accumulate over time and cause pathologic consequences. This occurrence is common is late-onset neurodegenerative disorders such as Huntington's and poly-glutamine expansion disease (Fan 2014) or other amyloidosis linked to cancers (Chiti & Dobson 2006). QC itself faces several challenges in balancing between accumulation of misfolded protein, and depletion of functional proteins. Firstly, there is a wide

range of degrees of misfolded proteome in the cell, making it hard for the QC system to identify the right substrates to degrade. Secondly, newly synthesized peptides have to undergo proper folding, modification, localization and assembly to functional units. The folding intermediates, although non-functional, should not be degraded. Therefore, the QC machinery has to be able to set time for the folding process to be complete and also prioritizes the degradation of potential substrates. To overcome those challenges, QC system has to rely on properties of the proteins such as biochemical features, cellular localization, and time to selectively degrade the aberrant proteins in several discriminatory steps. In this chapter, we aimed to unravel the cytosolic QC in handling the mistargeted proteins.

In the previous chapter, we identified the inhibitory effects on ER translocation and signal peptide cleavage to ER substrates of HSPA13. In particular for TTR as a substrate, this inhibition redirected it to be mistargeted in the cytosol. As observed in Chapter 2, the mistargeted TTR containing signal sequence is prone to proteasomal degradation. However, it still remains inconclusive since we did not intensively investigate their fate after being mistargeted. As introduced in Chapter 1, there are two major degradation pathways in the cell: UPS and lysosomal autophagy. Short-lived, misfolded proteins are directed to UPS, whereas bulky materials such as protein aggregates, bacteria or damaged organelles (Pohl & Dikic 2019). There are a few common characteristics between

the two systems. First, Ub serves as degradation signals for UPS and autophagy (Lu 2017); second, the proteasome itself can also be ubiquitinated for autophagic degradation (Besche 2014). Lastly, UPS and autophagy are also linked by properties of p62 and Ub. For example, p62 can trigger proteasomal degradation of N-end rule degrons, but the binding of p62 to the substrates also form aggregates among the molecules themselves leading to autophagy (Cha-Molstad 2017). Therefore, the two systems of degradation can be mediated by each other and act as an integrated QC network in response to accumulation of misfolded or damaged substrates in the cell.

As described earlier in Chapter 2, protein entering the ER is regulated by transmembrane domain or N-terminal signal sequence. These targeting sequences are recognized by SRP, enabling their co-translational translocation (Shan & Walter 2005, Halic & Beckmann 2005). If the TMD is located near the C-terminus, it cannot be recognized by SRP, but instead uses a post-translational targeting factor termed TRC40 (TMD Recognition Complex 40 kDa Subunit) (Hegde & Keenan 2011, Stefanovic & Hegde 2007). Targeting can sometimes fail due to the variance of signal sequence or during ER stress (Levine 2005, Kang 2006), rerouting the proteins in the cytosol that are degraded by the proteasome; otherwise, they can form aggregates and become toxic to the cell. These mislocalized proteins can be recognized by BAG6 (BCL2-associated

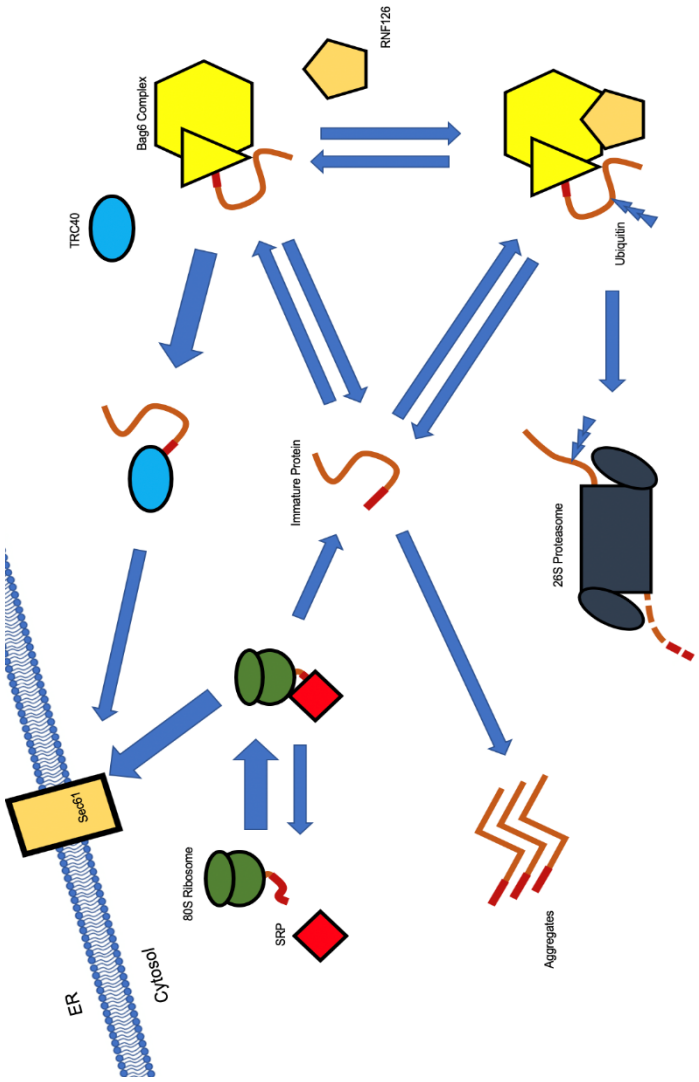

RNF126
Bag6 Complex
TRC40
Ubiquitin
Immature Protein
26S Proteasome
Sec61
80S Ribosome
SRP
Aggregates
ER
Cytosol

Figure 3.1: Quality control of mislocalized ER proteins. The signal recognition particle (SRP) recognizes the signal peptide of an elongating polypeptide. SRP binds to the signal peptide and the ribosome to deliver clients to the Sec61 translocon. If the protein fails to enter the ER, it will be released into the cytosol and will aggregate if left unattended. BAG6 complex binds the mistargeted protein and hands over to TRC40 if the protein is a TA-protein or it has a C-terminal signal sequence. Otherwise, BAG6 will recruit E3 ligase RNF126 and the clients will be ubiquitinated and proteasomally degraded.

athanogene 6), which recruits the E3 ligase RNF126 for ubiquitination. The ubiquitination event commits the substrates to proteasomal degradation (Hessa 2011, Rodrigo-Brenni 2014). This current model of proteasomal degradation was proposed for mistargeted proteins in the cytosol. However, we assumed this mechanism would be more complicated than the simple action of Bag6 recognition and ubiquitination. Different proteins with various energetics as well as stability are mistargeted to the cytosol. Hypothetically, cytosolic QC is mostly involved in degradation of mistargeted proteins. In the ER, folded proteins are sorted from folding intermediates and misfolded proteins to enter the secretory pathway by ER quality control (Sato 2007). This mechanism was also observed in correlation of protein stability to efficiency of secretion (Kowalski 1998). In particular, the lower the stability of each variant is, the less efficient its secretion from the cell is (Sekijima 2005). The energetic stability can be defined by kinetic stability of the TTR tetramer, which is the rate of dissociation of tetramer; and thermodynamic stability, which is the transition from folded tetramer to unfolded monomer. These two stabilities can predict the severity of amyloidogenicity of TTR (Sekijima 2005). Furthermore, the aggregation propensity of extracellular TTR variants also depends on their energetics, in which more destabilized variants are more prone to aggregation (Chen 2014).

The amyloidosis of TTR that is associated with neurodegeneration and cardiomyopathies is caused by its rate-determining tetramer dissociation,

followed by partial monomer unfolding (Colon & Kelly 1992, Lai 1996). The unfolded monomer of more than 100 TTR variants is competent for aggregation that can lead to familial amyloidosis. The disease can be more severe with variants that exhibit tissue-selective deposition and pathology. For instance, TTR^{D18G} is present in CNS involvement in the fifth decade of the patient. This variant is highly prone to aggregation under physiological condition. Its low stability protects the carrier from early onset by disrupting the secretion of TTR^{D18G} into the plasma and CSF (Hammarström 2003). More importantly, TTR^{D18G} is efficiently degraded by ERAD under regulation of GRP78/BiP (Sörgjerd 2006, Susuki 2009). Another unstable TTR variant TTR^{A25T} with CNS amyloidosis at the fourth decade of the patient has low concentration in serum that also explains the late onset of peripheral disease (Sekijima 2003). On the other hand, variant TTR^{L55P} is associated with early onset and relatively progressive amyloidosis (Jacobson 1992). TTR^{D18G} and TTR^{A25T} secretion is inefficient, making it hard to be detected in serum. However, they are still present in the cell in destabilized conformations, which can be prone to aggregation and become toxic to the neuronal cells. Meanwhile, TTR^{L55P} is more readily secreted from the cell. However, this variant is less stable than TTR^{WT} and it can become more aggregation-prone in the extracellular space than TTR^{WT}. Given the scope of the energetics of TTR variants regulating their secretion efficiency and thus mediating their progression of amyloidosis, we included a few TTR variants in our study to investigate how different mislocalized variants are handled in the cell.

Another ER substrate that we include in our study is alpha-1 antitrypsin (A1AT). A1AT is a major model for secretory protein trafficking. A1AT belongs to the serpin family that protects the connective tissue of the lungs from elastase. A1AT is predominantly produced from the liver and secreted to blood. A1AT serves as a serine protease inhibitor against neutrophil elastase. During cellular response to pathogens in the lung, proteases are oxidants released into the lung tissues. Neutrophil elastase is an important protease but it can destroy lung matrix components. Therefore, the lung is protected by a variety of antiproteases including A1AT. Thus, the imbalance between levels of A1AT and this elastase can lead to emphysema (Silverman & Sandhaus 2009). A1AT is present at the highest concentration and has the highest affinity for neutrophil elastase (Köhnlein & Welte 2008). On a molecular basis, A1AT is an ER substrate and is glycosylated after entering the ER. In some rare occasions patients can possess null A1AT alleles, which are characterized by the absence of A1AT in serum (Lee 2000). One variant of these null alleles that was included in our study is null Hong Kong A1AT (NHK.A1AT), whose TC deletion at Leu 318 results in a frameshift to obtain the premature stop codon at 334 (Lee & Brantly 2000). In the cell, NHK.A1AT is retained shortly in the ER and efficiently degraded by ERAD, indicating that this variant is highly destabilized (Wu 2003, Hosokawa 2003, Hosokawa 2007). Thus, NHK.A1AT is not efficiently secreted from the cells, leading to deficiency in A1AT (Sifers 1988). In the absence of A1AT, neutrophil elastase accumulates in the lung to cause severe emphysema (Fregonese

2008). We found that A1AT is also an appropriate model for investigation of how $A1AT^{WT}$ and NHK.A1AT are regulated for degradation when they are mistargeted to the cytosol.

MATERIALS AND METHODS

Plasmid DNA

Wild-type ^{Flag}TTR and its variants (N-terminal Flag) and were prepared from pcDNA3.1. Cytosolic peroxidase pcDNA3 APEX-NES was a gift from Alice Ting (Addgene plasmid # 49386 ; http://n2t.net/addgene:49386 ; RRID:Addgene_49386) [64]. pCMV-erHRP(N175S mutant) was a gift from Joshua Sanes (Addgene plasmid # 79909 ; http://n2t.net/addgene:79909 ; RRID:Addgene_79909) [65].

Cell culture

Human embryonic kidney cell line 293 with T antigen (HEK293T) were grown in Dulbecco's Modification of Eagle's Medium (DMEM; Corning 10-013-CV) containing 10% fetal bovine serum, XX L-Glutamine and XX Penicillin-Streptomycin. Cells were transfected at 20-60% confluence while being attached to the plates. All transfection performed in the study were $Ca_3(PO_4)_2$ transfection. 250 mM $CaCl_2$ containing DNA plasmid was vortexed with drop-wise addition of 2X HEPES-buffered saline (2X HBS). The resultant transfection solution was

added drop-wise to the cells with fresh replenishment of media after 12-16 hours. To enhance the adherence of HEK293T cells on the plates, cells were reseeded on poly-D-lysine treated plates. To inhibit ER translocation, we treated the transfected cells with 25 nM mycolactone for 12 hours. To prevent degradation of mistargeted proteins in the cytosol, 1 μM MG132 and/or 100 μM chloroquine (CQ) was added to the cells for 12 hours. The negative control received the same volume of DMSO vehicle.

Cell lysis

Cells were lysed in RIPA buffer (50 mM Tris, 15 mM NaCl, 1% Triton X100, 0.5% deoxycholate and 0.1% sodium dodecyl sulfate (SDS)) containing 1X protease inhibitor cocktail (PI) (Roche) for at least 15 minutes on ice. The lysate was collected from the supernatant of a 21,100 g spin for 15 minutes. Protein was quantified from Bradford assay by spectrophotometry (Agilent Cary 50 and 60).

Western Blot sample preparation

The proteins from different samples were normalized and brought to the same volume by RIPA buffer. If the samples are cultured media, no normalization would be performed. The 6X Laemli loading dye was added to a final concentration of 1X and 16.7 mM dithothreitol was also included in the samples. To denature and reduce the proteins, the samples were boiled at 100^{o}C for 5 minutes or 20 minutes as indicated. Proteins were loaded on homemade SDS-

polyacrylamide gel electrophoresis (SDS-PAGE) gels cast in the Bio-Rad Mini-PROTEAN system. The samples were run through the stacking gel and the resolving gel at 60 V and 170 V, respectively.

The gels, filter paper and nitrocellulose membrane were immersed in Towbin's buffer, consisting of 25 mM Tris, 192 mM glycine, 20% methanol (per Bio-Rad transfer buffer recipes). The semi-dry transfer of the sandwich was conducted in a Bio-Rad Trans-Blot Turbo for 60 minutes at 25 volts and 1 A. The quality of the transfer was confirmed by Ponceau S assay. The membranes were then blocked with 5% milk in Tris-buffered saline (TBS) for 1 hour. Milk was rinsed well with TBS for several rounds. The proteins were then probed with primary antibody diluted in 5% bovine serum albumin (BSA), 0.1% NaN_3 in TBS for at least two hours. The membranes were then rinsed with TBS buffer with 0.1% Tween (TBST) and probed with secondary antibody targeting the primary host such as rabbit or mouse for 30 minutes. The secondary antibody was diluted in 1:20000 ratio in blocking buffer. The membranes were repeatedly rinsed with TBST, followed by a rinse in TBS and water. After all steps of blotting, the membranes were imaged with a Li-Cor Odyssey Fc, using 700 nm and 800 nm filters and exposure times of 10 minutes each. The images were adjusted with Li-Cor's Image Studio, with quantification also performed, if necessary.

Proximity labeling of living cells

ER-HRP or APEX-NES plasmids and other plasmids were transfected to the cells. Cells were reseeded on PDK-treated plates and subject to drug treatment as indicated above. Fresh media including 500 μM biotin-phenol was added to the cells 30 minutes before harvest. Cells were then incubated at 37°C, 5% CO_2 as usual. To initiate the labeling reaction of peroxidase to biotinylated proteins, 1 mM H_2O_2 was added to the media and the plates were gently agitated for 1 minute. Immediately after 1 minute, the media was aspirated and the quenching solution (5 mM Trolox, 10 mM sodium ascorbate and 10 mM sodium azide in cold DPBS, pH 7.4). was added to the cells. The cells were rinsed in the quenching solution for 3 times before harvest. Cells were then lysed in RIPA buffer containing 5 mM Trolox, 10 mM sodium ascorbate and 10 mM sodium azide in the same protocol as described above. Equal amounts of proteins were added to streptavidin beads and rotated overnight at 4°C. After rotation, the beads were spun down at 1500 x g for 1 minute and the supernatant was aspirated. The beads were then washed 2X with RIPA buffer, 1X with 1 M KCl, 1X with 0.1 M Na_2CO_3, 1X with 2 M urea, and 2X with RIPA buffer. The beads were then eluted in buffer containing 12% SDS, 10% glycerol in 10 mM Tris pH 6.8, with 10 mM DTT and 2 mM biotin, and boiled for 10 minutes.

Ultracentrifugation assay

Cell lysates after normalization were balanced in weight with RIPA in Beckman-Coulter polypropylene ultracentrifuge tubes (357448) and spun down in a Beckman-Coulter Optima Max-XP ultracentrifuge, using a TLA-55 rotor. Samples were spun at 77,000 x g for four hours, at 4°C and under vacuum. Pellets formed from ultracentrifuge were rinsed four times with RIPA and incubated for 96 hours with 8 M urea at 4°C. The solubilized pellets were diluted to 2 M urea with RIPA and prepared for Western blot by addition of reducing Laemmli and boiling for 5 minutes.

Ubiquitination Purification of Mistargeted TTR

HEK293T cells were transfected with TTR^{WT} or TTR^{D18G} and Ubiquitin-HA followed by 25 nM mycolactone, 1 μM MG132 and 100 μM CQ as indicated for 16 hours. Cells were harvested and lysed in RIPA/PI buffer containing 12.5 mM N-ethylmaleimide to inhibit endogenous deubiquitinylases during lysis. The lysates were precleared with Sepharose beads for 30 minutes at 4^{o}C and incubated overnight at 4^{o}C with anti-Flag M2 beads to purify TTR. The immunoprecipitates were run on 12% SDS-PAGE gel for western blotting and immunoblotted for HA to probe for ubiquitin.

Peptide Mapping of Proteolytic Product of mistargeted TTR

HEK293T cells were transfected with $^{Flag}TTR^{D18G}$ and treated with 25 nM mycolactone and 1 μM MG132 to obtain the majority of the intermediate band. Cells were then harvested and lysed in RIPA/PI. The lysates were precleared with Sepharose for 30 minutes at 4°C, and incubated with M2 anti-Flag beads overnight at 4°C. The beads were then washed stringently with RIPA buffer for 4 times and eluted in Laemli buffer by boiling for 5 minutes. The eluate was precipitated by methanol-chloroform precipitation as described above. The air-dried protein pellet was resuspended in 9 M urea, reduced with 10 mM tris(2-carboxyethyl)phosphine for 30 minutes at room temperature, and alkylated with 5 mM iodoacetamide for 15 minutes at room temperature in the dark. The solution was diluted to 2 M urea by 50 mM Tris pH 8.0. The sample was digested with 0.5 ug trypsin overnight at 37°C overnight. The digestion was quenched by 0.5% acetic acid and spun down at 21,100 x g for at least 30 minutes to remove debris.

The samples were preloaded on injected to a 2-cm trapping column containing C18 resin. Separation was achieved by applying 7-55% gradient of solvent B (80% ACN, 0.1% formic acid) at a flow rate of 500 nL/min. A voltage of 3.0 kV was applied for electrospray ionization and the inlet capillary was heated to 275]C. Data-dependent acquisition of MS/MS spectra with the LTQ-Velos-Orbitrap was performed with the following settings: MS/MS on the 10 most intense ions per precursor scan, 1 microscan, rejection of unassigned and

charge state 1; dynamic exclusion repeat count 1, repeat duration, 30s; exclusion list size 500, and exclusion duration, 60s. The MS/MS spectra were obtained from CID fragmentation in the Ion trap of 38.0. Peptide identification was performed on Integrated Proteomics Pipeline IP2. Tandem mass spectra were extracted from the raw files using Rawconverter and were searched on a Uniprot human database with reversed sequences using ProLucid algorithm (Peng 2003, He 2015, Uniprot Consortium 2019). The search parameters included nontryptic peptides. Carbamidomethylation (+57.02146) of cysteine and tandem mass labeling (+229.1629) of lysine and N-terminus were set as static modification. Peptide candidate were filtered using DTASelect, with these parameters: -p 2 –y 0 –sfp 0.01 –DM 10. We require at least 2 identified peptides for a protein to be accepted. The peptides can be accepted regardless of enzyme specificity. The peptide level false positive rate was set to 0.01. We accepted peptides that have delta masses no more than 5 ppm.

RESULTS

The immature $^{Flag}TTR^{D18G}$ was not identified with overexpression of HSPA13

In Chapter 2, we always observed two distinct bands of $^{Flag}TTR^{WT}$ which are immature TTR with signal sequence in cytosol and mature TTR with signal sequence cleaved in the ER (Figure 3.2, lane 1-2). Overexpression of HSPA13 increased the relative amount of immature TTR and introduces a third,

intermediate band (Figure 3.2, lane 3-4). However, $^{Flag}TTR^{D18G}$ only showed up with the mature TTR^{D18G} band, even with HSPA13 overexpression. When MG132 was included as an inhibitor of proteasomal degradation, we were able to observe the immature TTR^{D18G} band. This result implied that the mistargeted TTR^{D18G} is not inherently present in cytosol and can only be recovered in the presence of MG132. We also observed the mature TTR^{D18G} band getting more intense with MG132, which explains the retention of ER-associated degradation of TTR^{D18G}. We inferred that mistargeted TTR^{D18G} is more rapidly degraded by the proteasome in cytosol compared to TTR^{WT}.

Inhibition of ER translocation leads to accumulation of stable proteins in the cytosol while destabilized variants are degraded efficiently

The major band of TTR^{WT} detected on the Western blot is the mature TTR, but a small amount of immature TTR is present as well. The immature TTR is representative of mistargeted TTR in the cytosol. To observe the degradation kinetics of the steady state population of mistargeted TTR, we performed cycloheximide chase of mistargeted TTR^{WT} in the presence or absence of proteasomal inhibitor MG132, which is a peptide aldehyde that effectively blocks the activity of the 26S proteasome complex (Goldberg 2012). HEK293T cells transfected with $^{Flag}TTR^{WT}$ were incubated in the presence of translational inhibitor cycloheximide (CHX). Aliquots of cells were harvested immediately after and at specific time points of cycloheximide treatments (Figure 3.3.A). We

observed that regardless of the presence of MG132, the mistargeted TTR^{WT} in cytosol is persistent while the mature TTR^{WT} decreases when translation is halted (Figure 3.3.B). This decrease is due to the attenuation of protein synthesis. This result indicated mistargeted TTR^{WT} is stable in the cytosol for at least 4 hours even when the proteasomal degradation is active.

Secretory proteins first enter the ER through the Sec61 translocon and its signal sequence is cleaved by the signal peptidase to be secreted. However, different protein signal sequences of the proteins vary in their ER translocation efficiency (Levine 2005). This leads to mislocalization of ER proteins in the cytosol. Additionally, pre-emptive quality control attenuates protein translocation during ER stress (Kang 2006). To investigate how mistargeted proteins are handled in the cytosol, we employed the secretory proteins transthyretin (^{Flag}TTR) and alpha-1 anti-trypsin (^{Flag}A1AT) as the protein models for our study. Their destabilized variants $^{Flag}TTR^{D18G}$ and Null Hong Kong A1AT (NHK.$A1AT^{HA}$) were transfected into HEK293T cells in parallel with the wild-type. To inhibit protein translocation, 25 nM mycolactone was added to the cells for 16-hour prior to harvest (McKenna 2016, Demangel & High 2018). Similar to the TTR bands we observed in the experiments of Chapter 2, TTR is present as a mature form once its signal sequence is cleaved upon entering the ER. Also, its immature form with the uncleaved signal sequence is located in the cytosol before entering the ER (Figure 3.4, lane 1). Both cotranslational and post-translational translocation of

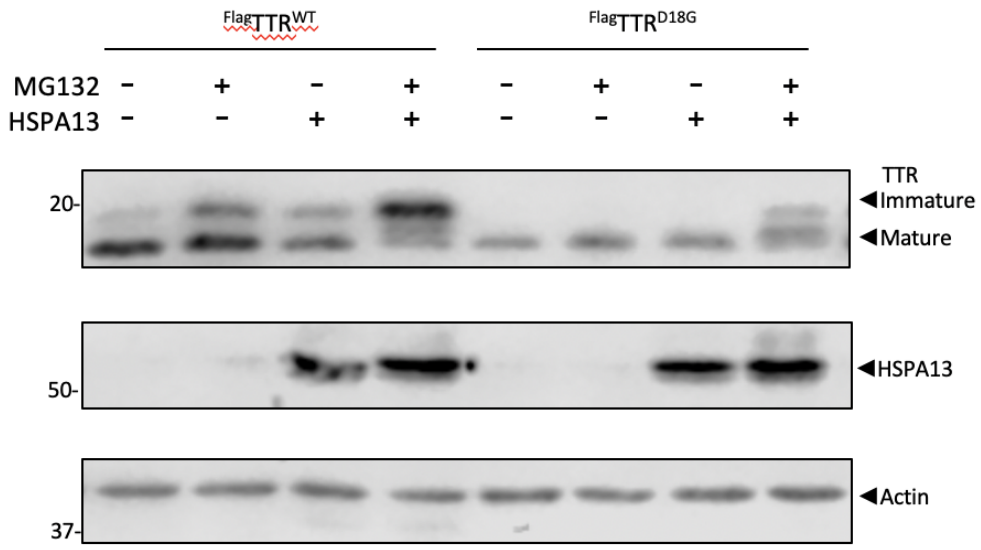


Figure 3.2: Inhibitory effect of HSPA13 did not reveal the immature TTR band in TTR^{D18G} variant. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{D18G}$ and HSPA13 cotransfection followed by 16-hour treatment of 1 μM MG132 or vehicle. The lysates were reduced and denatured for SDS-PAGE and Western blotting. The indicated antibodies were used to probing the proteins, followed by near-IR fluorescent secondary staining.

newly synthesized proteins into the ER is made possible by Sec61 translocon; thus, impairing this gateway significantly inhibits the translocation process. The nontranslocated proteins are normally degraded via proteasomal degradation (Demangel & High 2018). Our result indicated that upon treatment of mycolactone, in the absence of inhibitors of degradation, TTR^{WT} accumulates as the immature and intermediate bands (Figure 3.4, lane 2), while no TTR^{D18G} was observed (Figure 3.4, lane 8). The presence of the intermediate band of mistargeted TTR implicates a proteolytic process happening in the cytosol. Previous studies reported that TTR^{D18G} is highly destabilized, implying that the destabilized variant would be efficiently degraded to maintain homeostasis. Since these mislocalized proteins are destined for degradation, we applied inhibitors of proteasomal and lysosomal degradation to observe those mislocalized TTR. Besides MG132 as proteasomal inhibitor, we also included chloroquine that inhibits autophagosome acidification and prevents later steps of autophagy from taking place (Mauthe 2018). Under proteasomal inhibitor MG132 treatment, we were able to observe two different proteoforms of the immature $^{Flag}TTR^{WT}$ at higher intensity (Figure 3.4, lane 4). On the other hand, $^{Flag}TTR^{D18G}$ is mostly present as the proteolytic product in the cytosol (Figure 3.4, lane 10). Once both pathways of degradation were blocked, the immature and intermediate bands of $^{Flag}TTR^{WT}$ (Figure 3.4, lane 6) and $^{Flag}TTR^{D18G}$ (Figure 3.4, lane 12) were observed to a higher extent. Importantly, we did not observe this proteolytic

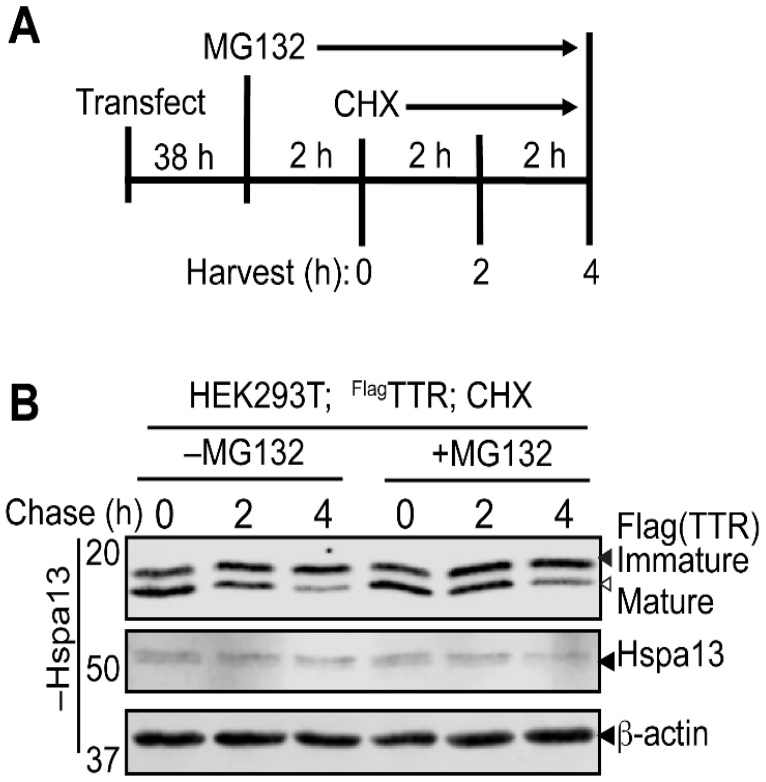
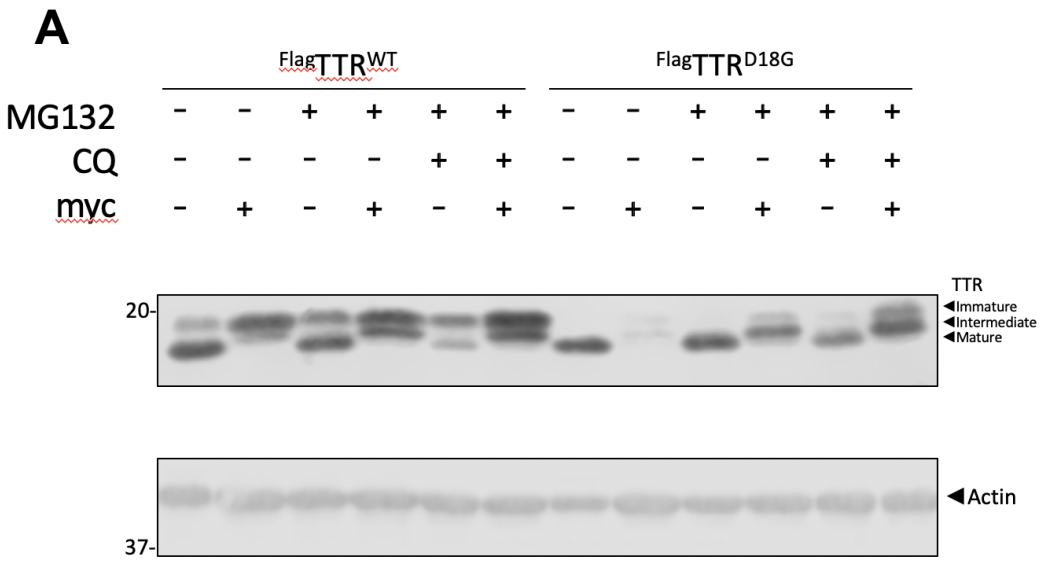


Figure 3.3: Cycloheximide chase of ^{Flag}TTR. A) Timeline of chase treatment and cell harvests. B) HEK293T cells transfected with ^{Flag}TTR^{WT} were treated with cycloheximide (CHX), followed by harvests at 0, 2 and 4 hours post-treatment with or without MG132 treatment.


A
FlagTTRWT
FlagTTRD18G
MG132
CQ
myc
20-
TTR
Immature
Intermediate
Mature
37-
Actin

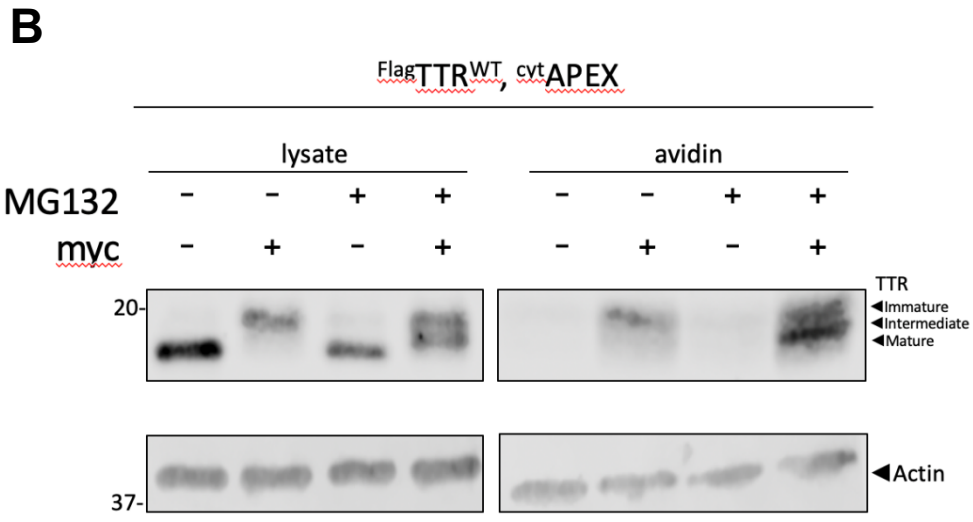

B
FlagTTRWT, cytAPEX
lysate
avidin
MG132
myc
20-
TTR
Immature
Intermediate
Mature
37-
Actin

Figure 3.4: Highly destabilized TTR variant is rapidly degraded in the cytosol. A. Representative Western blot of mistargeted $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{D18G}$. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{D18G}$ followed by 16-hour treatment of inhibitor of ER translocation and inhibitors of degradation. The lysates were reduced and denatured for SDS-PAGE and Western blotting. B. Representative Western blot of lysates and avidin enrichment of mistargeted $^{Flag}TTR^{WT}$ in the cytosol. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ followed by 16-hour treatment of inhibitor of ER translocation with and without MG132. Cells were labeled with 1 mM H_2O_2, quenched and harvested. The lysates were enriched by avidin purification for SDS-PAGE and Western blotting.

species in the absence of mycolactone, indicating that it is not produced by ERAD.

In our laboratory, we developed a proximity labeling technique that employs compartments-specific peroxidase as described earlier. To further confirm the localization of the species we observed, we coexpressed $^{Flag}TTR^{WT}$ and the cytosolic peroxidase APEX-NES to enable the enrichment of the cytosolic fraction of TTR. We confirmed both of the immature and intermediate species of $^{Flag}TTR^{WT}$ to be located in the cytosol (Figure 3.4.B). Additionally, mistargeted TTR is not completely degraded through UPS pathway since we observed higher intensity of the immature bands when we inhibited autophagy as well. However, the proteolytic TTR product is efficiently degraded proteasomally (Figure 3.4.A, lane 2, 4 and 10) but the full-length TTR product is slowly handled by autophagy (Figure 3.4.A, lane 4,10 and 12).

In addition to TTR, we would like to evaluate whether another ER substrate A1AT would also behave in a similar manner when being mistargeted to the cytosol. The NHK mutant variant of A1AT has a TC deletion mutation at 318 amino acid that leads to a C-terminal truncation (Lee 2000). This variant is also highly misfolded to be an ERAD substrate. Therefore, we hypothesized that their destabilization can be prone to degradation upon being mistargeted. A1AT is also a glycosylated protein in the ER; thus, once entering the ER, its

glycosylation leads to a band with a higher molecular weight. If A1AT in unable to translocate into the ER, there would not be post-translational glycosylation. In the same manner, we treated the cells that were transfected with ^{Flag}A1AT^{WT} or NHK.A1AT^{HA} with 25 nM mycolactone and observed that in the absence of inhibitors, ^{Flag}A1AT^{WT} is present at a higher molecular weight compared to the truncated NHK.A1AT^{HA} (Figure 3.5, lane 1 and 5). When Sec61 translocation is inhibited, in the absence of proteasomal inhibition, ^{Flag}A1AT^{WT} accumulates in the cytosol but NHK.A1AT does not (Figure 3.5, lane 3 and 7). When ER translocation and proteasomal degradation are inhibited, we observed a slight increase of A1AT^{WT} and NHK.A1AT signals in the cytosol (Figure 3.5, lane 4 and 8). This observation is consistent with what we obtained from mistargeted TTR that destabilized mutants are efficiently degraded when they are mistargeted to the cytosol.

Progressing and degradation of mistargeted proteins in cytosol depends on its stability

As observed earlier, mistargeted TTR^{D18G} is handled at higher rate and more efficiently than mistargeted TTR^{WT}. Thus, we hypothesize that the extent at which mistargeted proteins are proteolytically processed into the intermediate form, and/or the extent of how rapid those proteins are proteasomally degraded, could be influenced by the stability of those proteins. Using different variants of a single protein, rather than many different proteins with distinct structures and

folding/unfolding mechanisms, enables isolated evaluation of the relationship between those energetic parameters and the biological mechanisms under investigation.

Physiologically, there are several disease-associated mutants of TTR identified with different energetics (Sekijima 2005). Since we observed different proteasomal degradation between the wild-type and the most destabilized TTR variant, we also set out to investigate the rate of degradation of other variants (TTR^{V122I}, TTR^{L55P}, TTR^{A25T}, and TTR^{D18G}) when they fail to enter the ER. While TTR^{V122I} and TTR^{L55P} are slightly less stable than wild-type, TTR^{A25T} and TTR^{D18G} have high rates of dissociation and are highly destabilized. The selection of these variants allows us to differentiate the extent of processing and degradation of the cytosolic QC to variants of the same protein with various degrees of stability. Therefore, their rate of degradation is expected to be correlated to their stability. In Figure 3.6.A and 3.6.B, inhibition of ER translocation accumulated the immature TTR^{WT}, TTR^{V122I} and TTR^{L55P} variants; however, there was little TTR^{A25T} and hardly TTR^{D18G} accumulated in the cytosol under this treatment. When we applied the inhibitors of degradation to observe the retained mistargeted TTR, we observed the proteolytic product along with the immature band of TTR. We calculated the steady state yield of degradation of individual variants being the remaining fraction of the ratio of the amount of TTR in the cytosol in the absence of inhibitors over the amount of TTR being retained in the

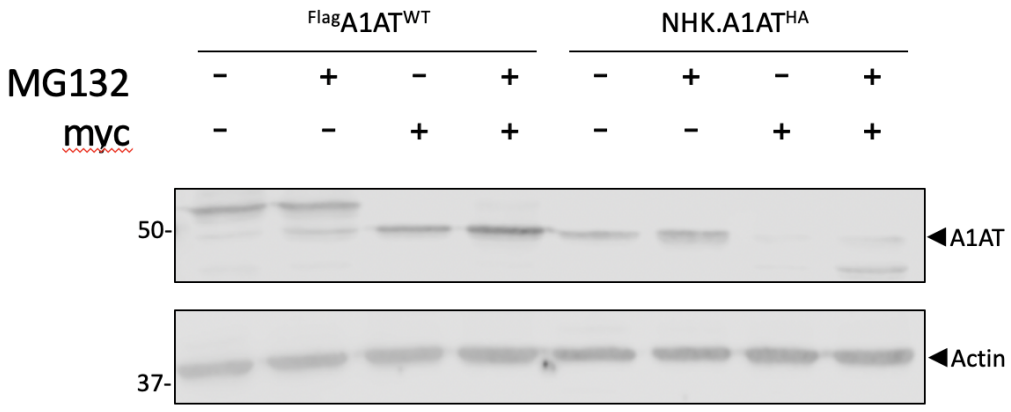
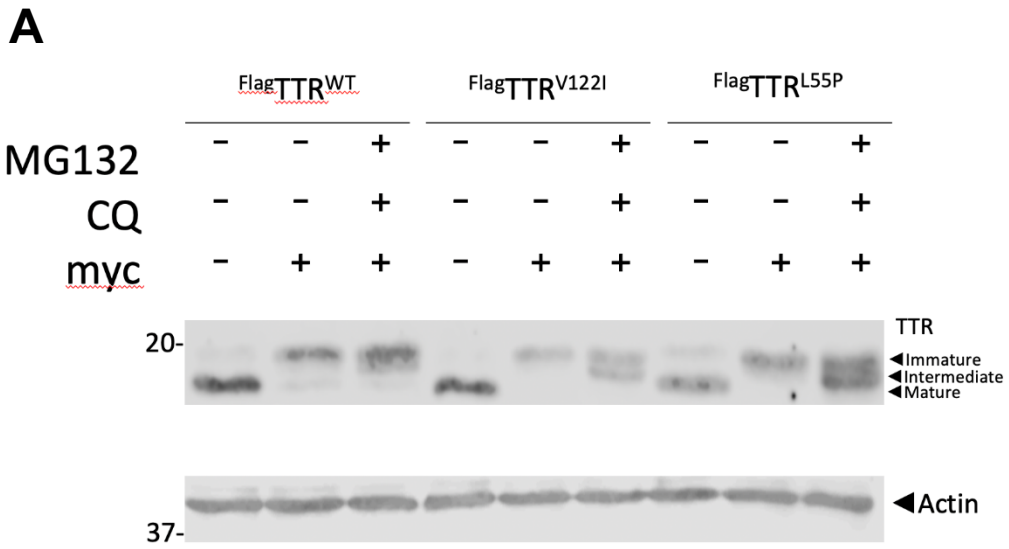
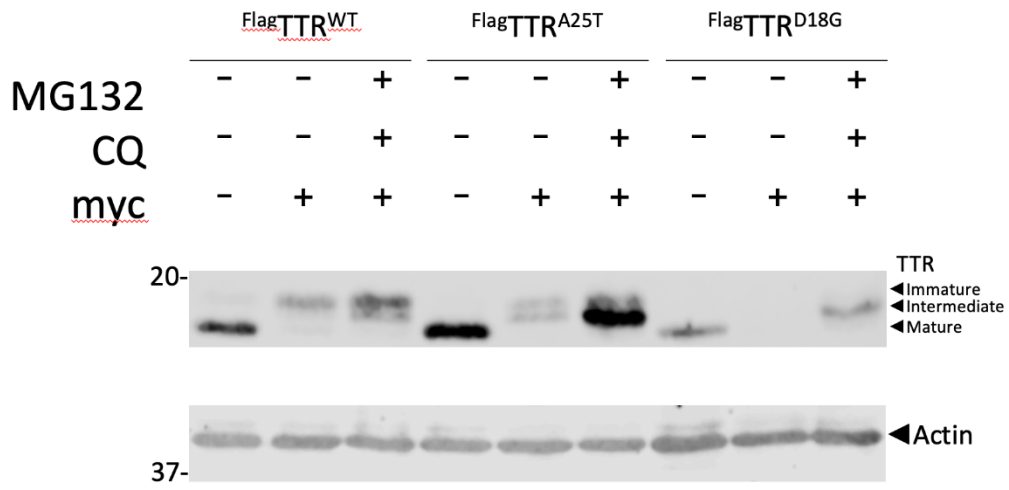


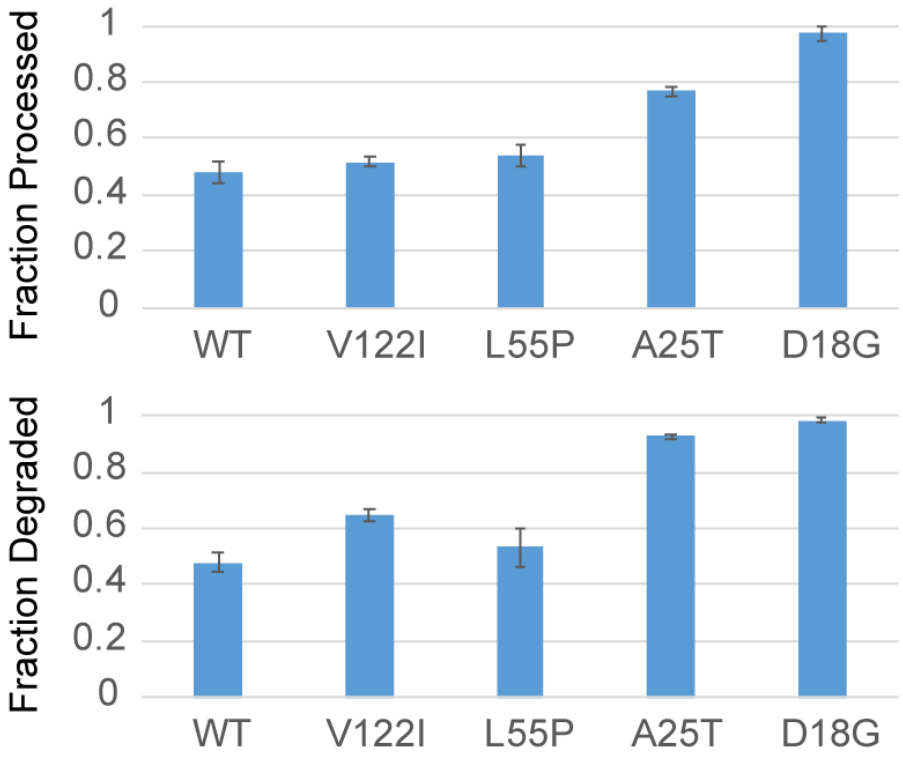
Figure 3.5: Destabilized NHK.A1AT is rapidly degraded in the cytosol compared to A1AT^{WT}. Representative Western blot of mistargeted A1AT^{WT} and NHK.A1AT. HEK293T cells were transfected with ^{Flag}A1AT^{WT} or NHK.A1AT^{HA} followed by 16-hour treatment of 25 nM mycolactone and/or 1 μM MG132. Lysates were reduced and denatured for SDS-PAGE and Western blotting.

presence of inhibitors. Additionally, the ratio of the proteolytic product over the total amount of mislocalized TTR in the presence of proteasomal degradation was also calculated. Noticeably, the proteolytic product is the most dominant band of TTR^{A25T} and TTR^{D18G}, indicating an efficient proteolytical process for highly destabilized variants. As we can observe from Figure 3.6.C, the efficiency of the proteolytic process and the degradation of mistargeted TTR variants are higher in the more destabilized variants. Additionally, we also plotted the efficiency of the proteolytic process against the combined stability score of TTR variants (Figure 3.6.D). The combined stability score was calculated based on the values of the kinetic and thermodynamic stabilities of TTT^{WT} and other variants, as shown in Sekijima *et al.* (Sekijima 2005). We plotted the fraction of the proteolytic process against the stability score of the variants and found a correlation between them. Additionally, we obtained the kinetic and thermodynamic scores from Sekijima *et al.* to plot against the amount of proteolytically processed fraction (Figure 3.6.E) and observed a monotonic trend in both graphs, in which destabilized variants (TTR^{A25T} and TTR^{D18G}) are more efficiently handled than other more stabilized variants. This observation can imply a quality control mechanism that the cytosol handles the mislocalized proteins differently depending on their stabilities. In the ER, highly destabilized variant TTR^{D18G} is degraded via ERAD; when it is mistargeted to the cytosol, a similar manner of degradation was observed in which more destabilized protein structures are degraded more rapidly.

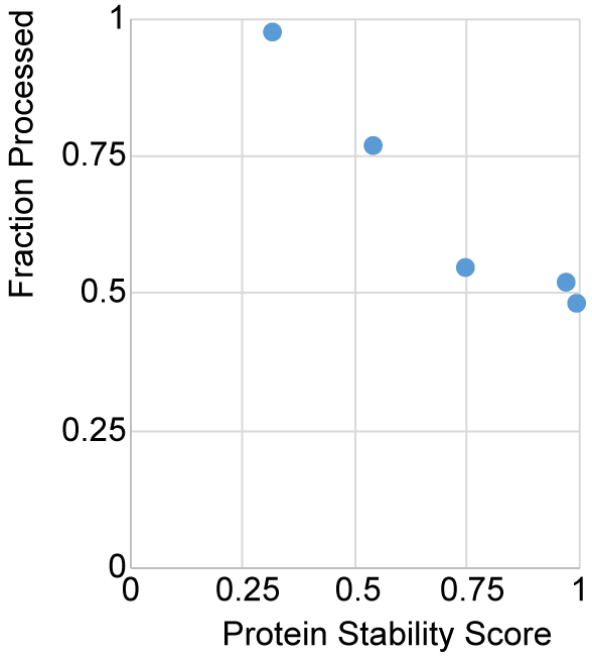

A
FlagTTRWT
FlagTTRV122I
FlagTTRL55P
MG132
CQ
myc
20-
TTR
Immature
Intermediate
Mature
37-
Actin


B
FlagTTRWT
FlagTTRA25T
FlagTTRD18G
MG132
CQ
myc
20-
TTR
Immature
Intermediate
Mature
37-
Actin

C


Fraction Processed
1
0.8
0.6
0.4
0.2
0
WT
V122I
L55P
A25T
D18G
Fraction Degraded
1
0.8
0.6
0.4
0.2
0
WT
V122I
L55P
A25T
D18G

D


Fraction Processed
1
0.75
0.5
0.25
0
0
0.25
0.5
0.75
1
Protein Stability Score

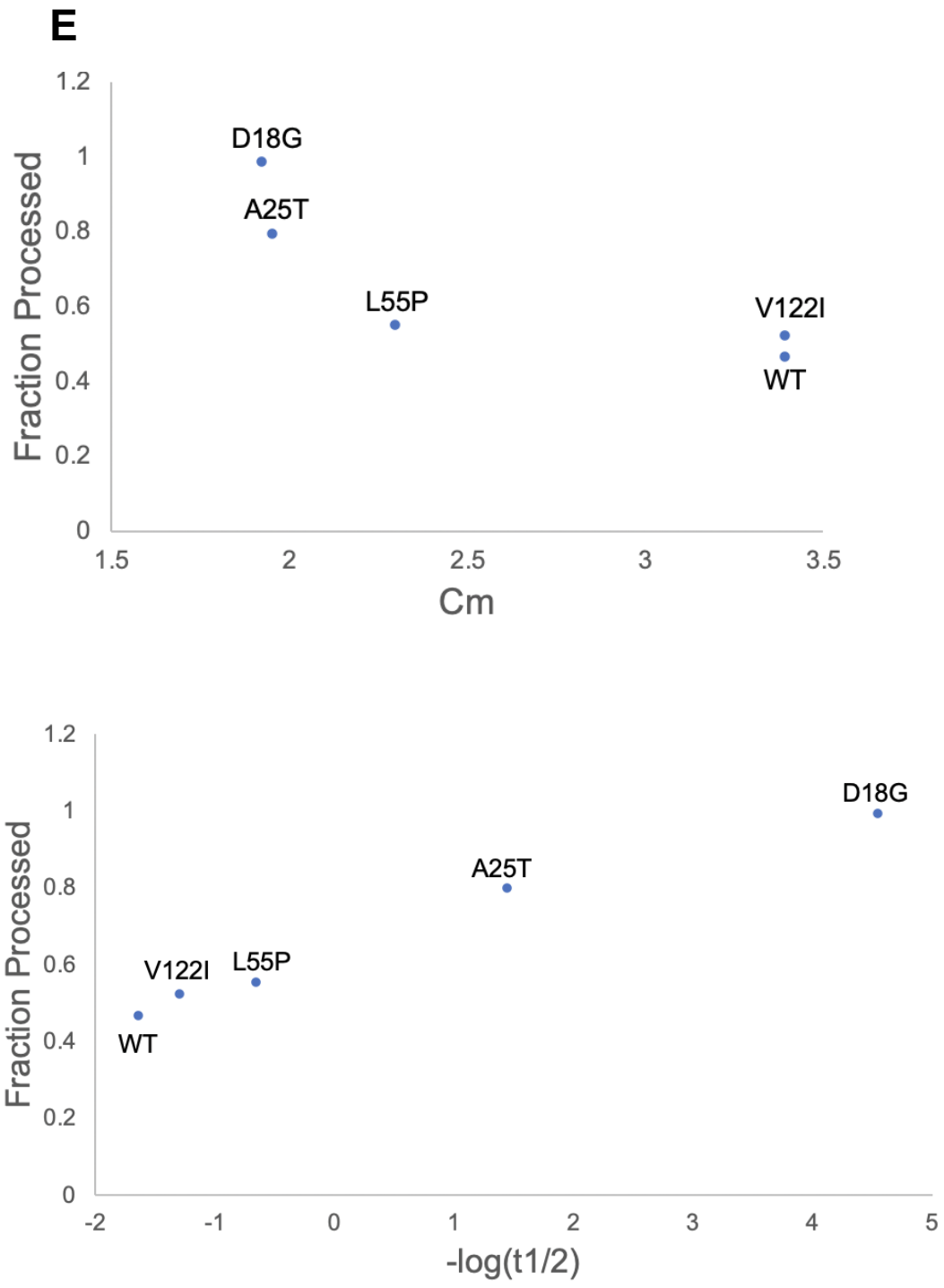
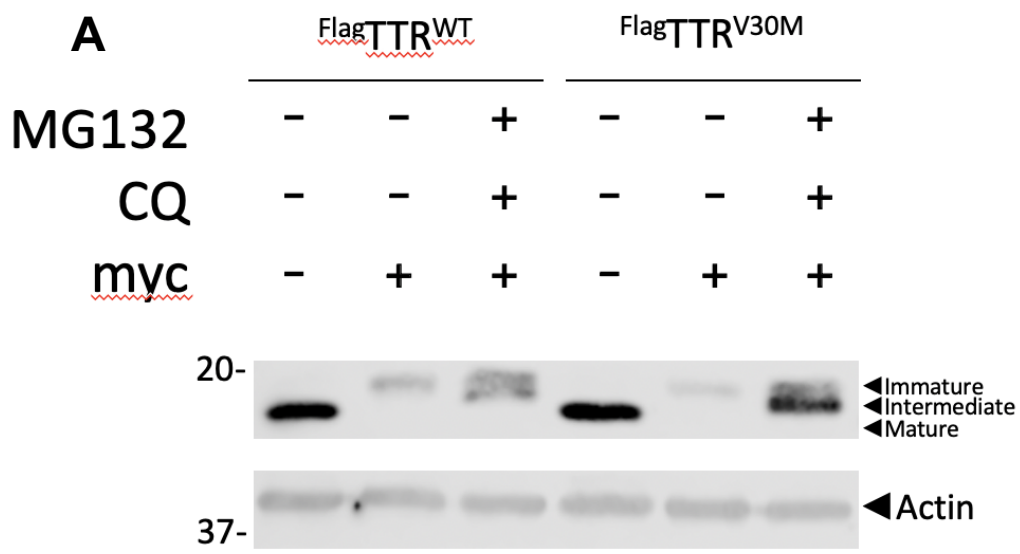

E
Fraction Processed
1.2
1
0.8
0.6
0.4
0.2
0
D18G
A25T
L55P
V122I
WT
1.5
2
2.5
3
3.5
Cm
Fraction Processed
1.2
1
0.8
0.6
0.4
0.2
0
D18G
A25T
V122I
L55P
WT
-2
-1
0
1
2
3
4
5
-log(t1/2)

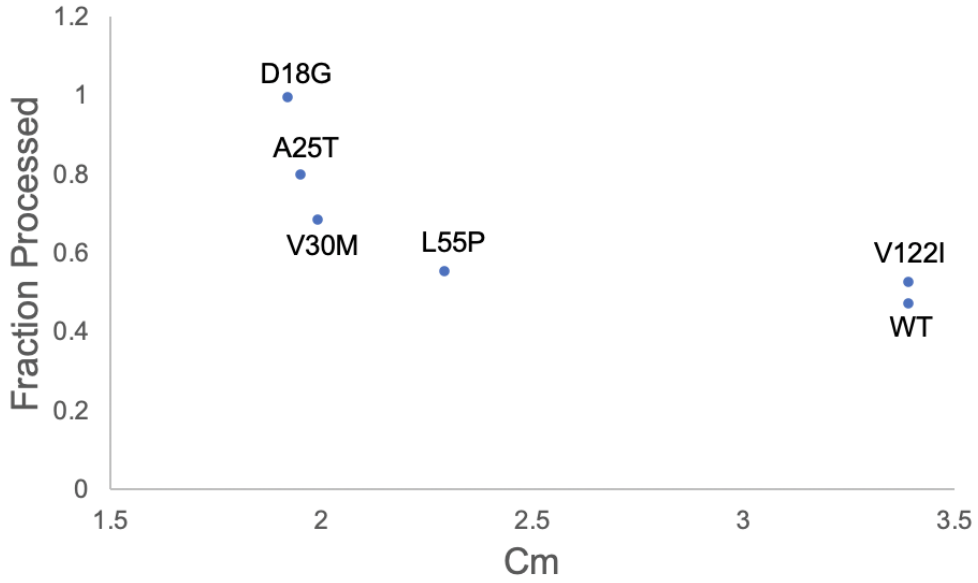
Figure 3.6: Degradation and proteolytic processing of mistargeted TTR in the cytosol depends on stability. A. Representative Western blot of mistargeted TTR^{WT}, TTR^{V122I} and TTR^{L55P} with or without inhibitors of degradation. B. Representative Western blot of mistargeted TTR^{WT}, TTR^{A25T} and TTR^{D18G} with or without inhibitors of degradation. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or ^{Flag}TTR variants followed by 12-hour treatment of myc with or without MG132 and CQ. The lysates were reduced and denatured for SDS-PAGE and Western blotting. C. Quantification of the relative degradation (upper graph) and relative amounts of processed fraction of mistargeted TTR^{WT} and variants (lower graph) (n=6). Error bars represent standard deviations. D. The correlation between stability of TTR and the efficiency of proteolytic processing. E. The correlation between thermodynamic stability (upper graph) and kinetic stability (lower graph) and the efficiency of proteolytic processing.

Proteolytic processing of mistargeted TTR in cytosol depends on its thermodynamic stability

The stability of TTR is defined in two parameters as mentioned earlier: thermodynamic and kinetic stability. We found out that the degradation and proteolytic processing of mistargeted TTR variants in the cytosol are dependent on their stability and we set out to identify whether thermodynamic or kinetic stability plays a role. Among the variants investigated, TTR^{V122I} is as kinetically stable but less thermodynamically stable as TTR^{WT}. The thermodynamic and kinetic stability of TTR^{L55P} variant is intermediate between being stable and destabilized. The two variants TTR^{A25T} and TTR^{D18G} are highly thermodynamically and kinetically destabilized. To determine which energetics of TTR determines their fate if being mistargeted to the cytosol, we made an additional TTR^{V30M} construct by site-directed mutagenesis. This variant is shown to be thermodynamically destabilized with similar C_m as TTR^{A25T} and TTR^{D18G} (Sekijima 2005). The C_m value of TTR^{V30M} is close to that of TTR^{A25T} and TTR^{D18G} while its rate of tetramer dissociation is similar to the rate of TTR^{WT}. TTR^{V30M} was identified in cardiac and leptomeningeal amyloidosis (Yoshinaga 2004). Energetically, this variant is thermodynamically unstable but kinetically stable, which is useful in the study to identify the related specific kinetics of TTR. Mistargeted TTR^{V30M} is also readily degraded in the absence of inhibitors of degradation and its immature form is efficiently processed to the proteolytic


A
FlagTTRWT
FlagTTRV30M
MG132 - - + - - +
CQ - - + - - +
myc - + + - + +
20-
37-
Immature
Intermediate
Mature
Actin

B


1.2
1
0.8
0.6
0.4
0.2
0
Fraction Processed
D18G
A25T
V30M
L55P
V122I
WT
1.5
2
2.5
3
3.5
Cm

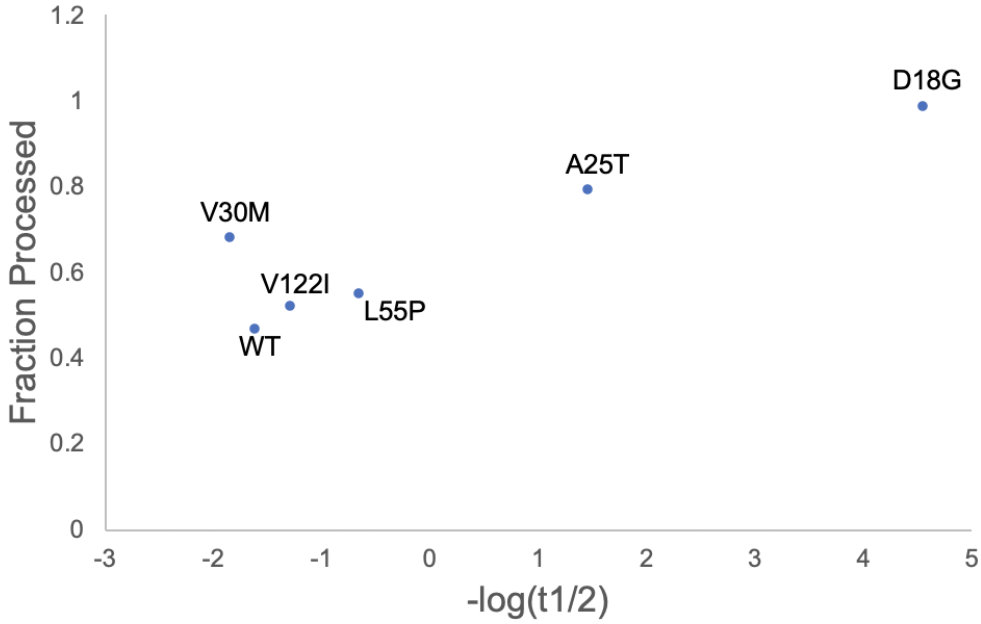
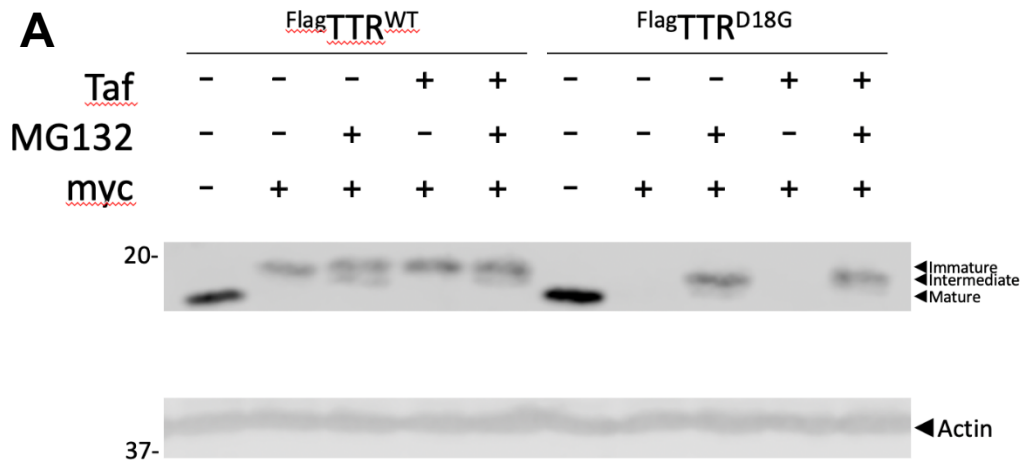

1.2
1
0.8
0.6
0.4
0.2
0
Fraction Processed
D18G
A25T
V30M
V122I
L55P
WT
-3
-2
-1
0
1
2
3
4
5
-log(t1/2)

Figure 3.7: Degradation and proteolytic processing of mistargeted TTR in the cytosol depends on thermodynamic stability. A. Representative Western blot of mistargeted TTR^{WT} and mistargeted TTR^{V30M} with or without inhibitors of degradation. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{V30M}$ followed by 12-hour treatment of inhibitor of ER translocation and inhibitors of degradation. The lysates were reduced and denatured for SDS-PAGE and Western blotting. B. The correlation between thermodynamic stability (upper graph) and kinetic stability (lower graph) and the efficiency of proteolytic processing (n=6).


A
FlagTTRWT
FlagTTRD18G
Taf
MG132
myc
20-
Immature
Intermediate
Mature
37-
Actin

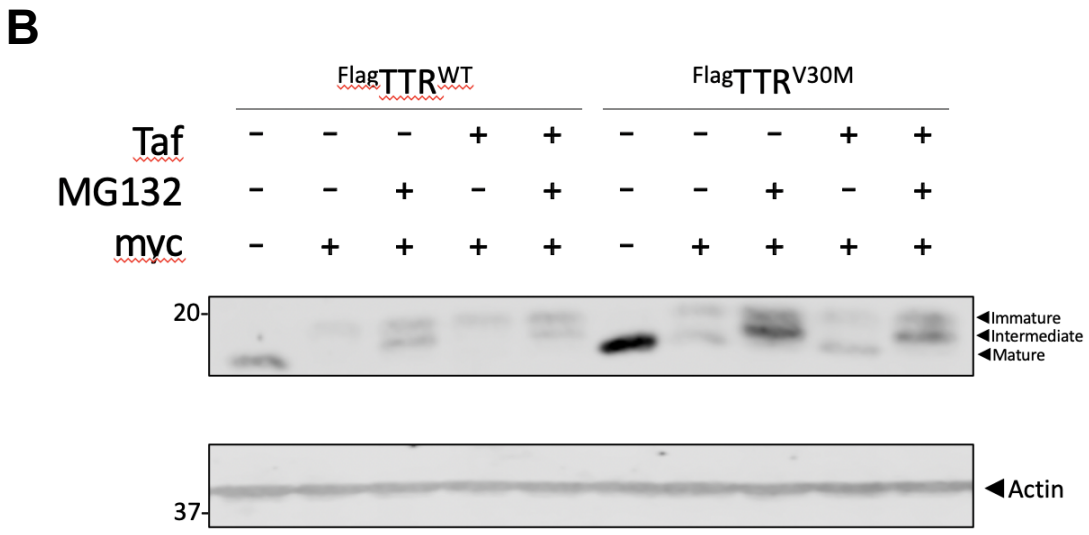

B
FlagTTRWT
FlagTTRV30M
Taf
MG132
myc
20-
Immature
Intermediate
Mature
37-
Actin

Figure 3.8: Tafamidis treatment did not recover mistargeted TTR^{D18G} and TTR^{V30M}. A. Representative Western blot of mistargeted TTR^{WT} and TTR^{D18G} with or without Taf and/or MG132 treatment. B. Representative Western blot of mistargeted TTR^{WT} and TTR^{V30M} with or without Taf and/or MG132 treatment. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{D18G}$ or $^{Flag}TTR^{V30M}$ followed by 16-hour treatment of inhibitor of ER translocation, Tafamidis, and/or MG132. The lysates were reduced and denatured for SDS-PAGE and Western blotting.

product as observed in TTR^{A25T} and TTR^{D18G} (Figure 3.7.A; lane 6,7). When we plotted the fraction of proteolytic product against the score of C_m value or $t_{1/2}$ of tetrameric dissociation, we found that the rate of proteolytic processing is correlated with thermodynamic stability of TTR (Figure 3.7.B).

Tafamidis treatment of mislocalized TTR^{WT} and TTR^{D18G}

To further elucidate whether thermodynamic or kinetic stability of TTR determines the triage of its mislocalized fraction in the cytosol, we employed a potent and selective kinetic stabilizer, Tafamidis (Taf), and evaluated whether mistargeted $^{Flag}TTR^{D18G}$ can be rescued from degradation due to tetrameric stabilization (Bulawa 2012). Taf molecule binds to empty thyroxine-binding sites of the tetramer to kinetically stabilize TTR. Taf was shown to facilitate secretion of TTR^{A25T} since Taf can bind and stabilize TTR^{A25T} tetramers in the ER (Chen 2014, Chen 2016). Our previous results indicated that destabilized TTR variants are rapidly degraded in the cytosol; thus, stabilization of tetramers may lower the degradation of TTR^{D18G}. Mistargeted TTR^{WT} and TTR^{D18G} were subjected to 10 μM Taf in the presence or absence of MG132. Overall, we observed a slight increase of accumulated mistargeted TTR^{WT} under Taf treatment (Figure 3.8.A and B; lane 4,5); however, there was no change to the proteolytic product. Taf molecules did not prevent TTR^{D18G} from degradation (Figure 3.8.A, lane 7,9), which implied that TTR cannot form tetramers in the cytosol. The mistargeted TTR in the cytosol is under the less stable state and misfolded; therefore, the

assembly to form tetrameric TTR is not favored. TTR^{V30M} is kinetically stable but thermodynamically unstable. To further illustrate the quality control of mistargeted TTR depends on its thermodynamic stability, we also performed Taf treatment on TTR^{V30M} and assumed that Taf can improve kinetic stability of TTR^{V30M} and thus, the variant would not be rapidly degraded. However, we found that there was no significant increase of mistargeted TTR^{V30M} in the inclusion of Taf molecules as the increase observed when MG132 was included (Figure 3.8.B, lane 7-10). This result further indicated that thermodynamically unstable proteins are degraded rapidly in the cytosol.

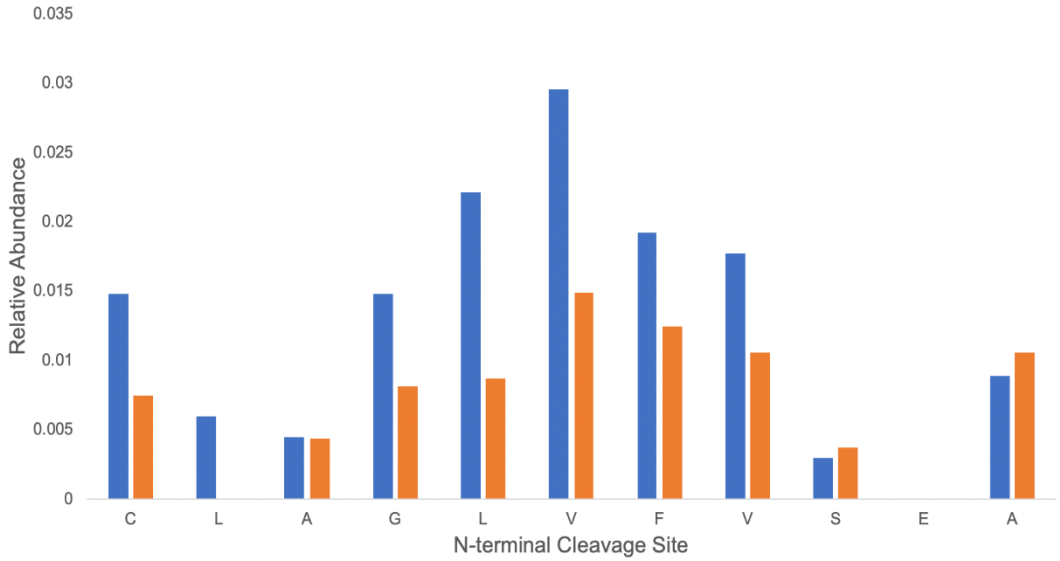
Peptide mapping of proteolytic product in the cytosol of TTR^{D18G}

To further understand the proteolytic product of TTR mislocalized in the cytosol, we performed peptide mapping of this product by data-dependent analysis of mass spectrometry. Peptide mapping is a proteomic method which analyzes the digested peptides to provide insight into the full sequence of a protein. Using this method, we would be able to figure out the cleavage site of the proteolysis of immature TTR. As observed in Figure 3.4.A, under treatment of mycolactone and MG132 on HEK293T cells, we obtained predominantly the intermediate immature band, which can enable analysis of this product only. HEK293T cells were transfected with TTR^{D18G} followed by mycolactone and MG132 treatment. The lysate was immunoprecipitated with anti-Flag M2 beads overnight and eluted in

A


TTR SS
FT
FT
TTR
MASHRLLLLCLAGLVFVSEADYKDDDDKDYKDDDDK

B


0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
Relative Abundance
C
L
A
G
L
V
F
V
S
E
A
N-terminal Cleavage Site

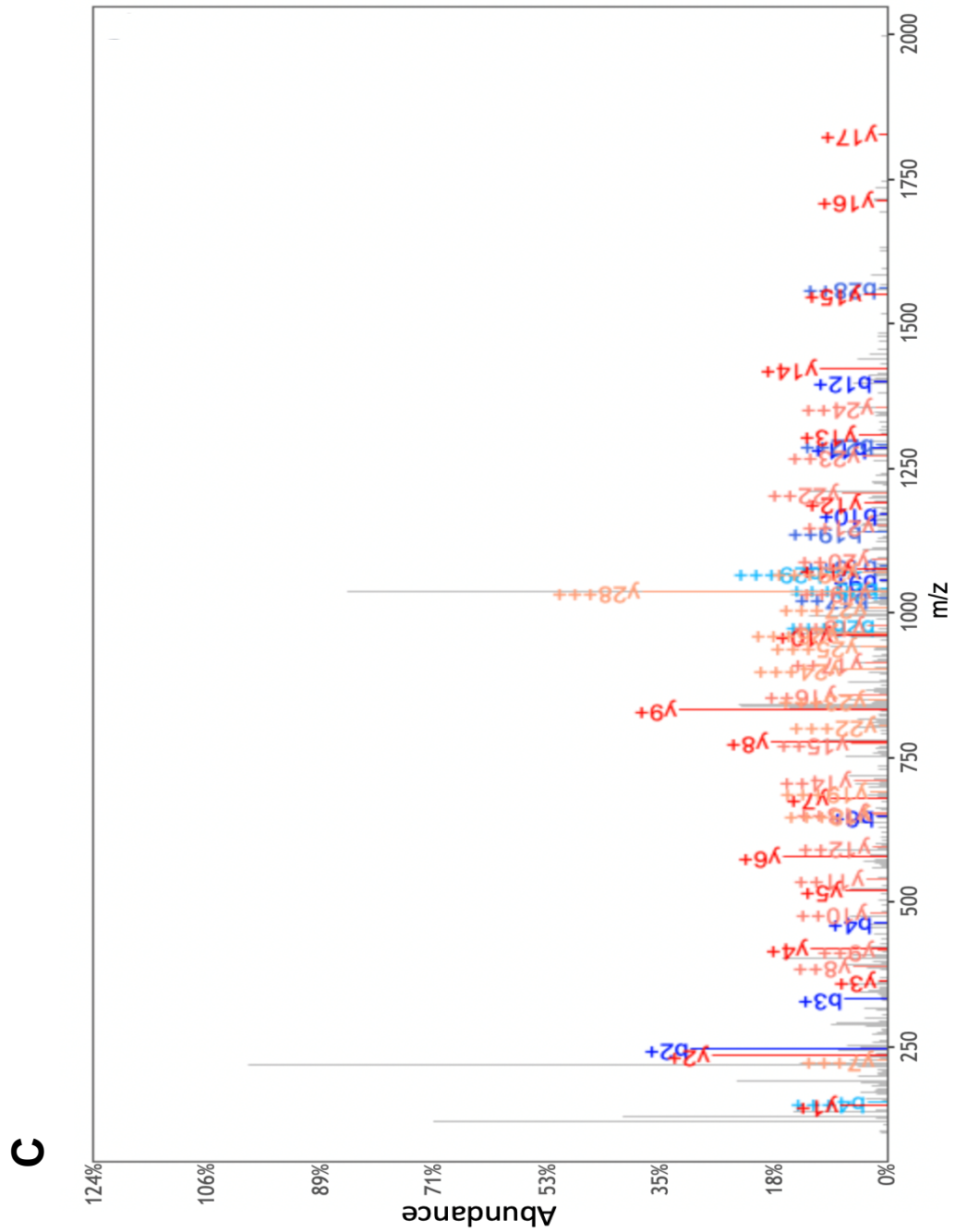
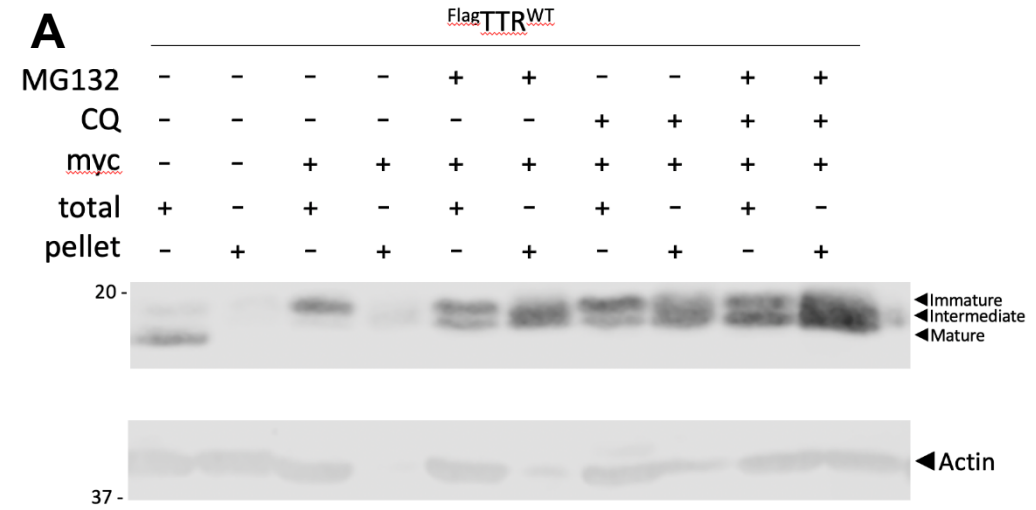

C
Abundance
124%
106%
89%
71%
53%
35%
18%
0%
m/z
250
500
750
1000
1250
1500
1750
2000
y1+
b2+
y2+
b3+
y3+
y4+
b4+
y5+
y6+
b8+
y7+
y8+
y9+
y10+
b10+
y12+
b12+
y13+
y14+
y15+
b15+
y16+
y17+

Figure 3.9: Peptide mapping of proteolytic TTR product in the cytosol. A. Construct of TTR protein consisting of the signal sequence, N-terminal Flag tag and TTR protein sequence. B. The relative frequency of identified cleavage site of proteolytic TTR product (blue: replicate 1, orange: replicate 2). HEK293T cells were transfected with $^{Flag}TTR^{D18G}$, followed by 16-hour treatment of myc and MG132. The lysates were immunoprecipitated with anti-Flag M2 beads and eluted in Leamli buffer. The eluate was prepared for bottom-up proteomics to obtain the sequence of TTR. C. MS/MS Spectrum of peptide containing the cleavage site with the highest number of spectral counts. The sequence of this peptide is FVSEADYKDDDDKDYKDDDDKGPTGTGESK, whose charge state is 4+. The theoretical MH+ is calculated as 3355.4244 Da, and the measured m/z was 839.8604.

Leamli buffer. The immunoprecipitate was extracted by methanol-chloroform precipitation, denatured in 9 M urea, reduced, alkylated and digested overnight with trypsin. The peptides were separated on a reversed-phase column and ionized by ESI to the mass spectrometer. The precursor ions were identified and fragmented by CID in the Orbitrap at normal resolution. We searched the MS/MS spectra with Prolucid algorithm on human database. We observed the coverage of ~ 85% of the protein. There was no cleavage site identified at the C-terminus; however, we identified cleavage at the N-terminus on the signal sequence. There was one noticeable peptide which was cleaved at phenylalanine or valine residue within the signal sequence with the highest relative frequency compared to other cleavage sites, implying the cleavage site of the proteolytic product (Figure 3.9.B). The relative frequency is calculated as the ratio of the number of spectral counts of each peptide and the total number of spectral counts obtained from the entire sequence. The obtained MS2 spectrum of this peptide has a high number of spectral counts and a good coverage of fragments (Figure 3.9.C).

Aggregation-assay of mistargeted TTR in the cytosol

In the ER, proteins are properly folded but they are misfolded and unstable when being mistargeted. The inhibition of ER translocation due to overexpression of HSPA13 led to the accumulation of higher molecular weight of TTR, indicating that TTR forms aggregates when they fail to enter the ER. Thus, we hypothesized that mistargeted TTR would be highly aggregating if they are not efficiently degraded. We also performed ultracentrifugation assay and


A
FlagTTRWT
MG132 - - - - + + - - + +
CQ - - - - - - + + + +
myc - - + + + + + + + +
total + - + - + - + - + -
pellet - + - + - + - + - +
20 -
Immature
Intermediate
Mature
37 -
Actin

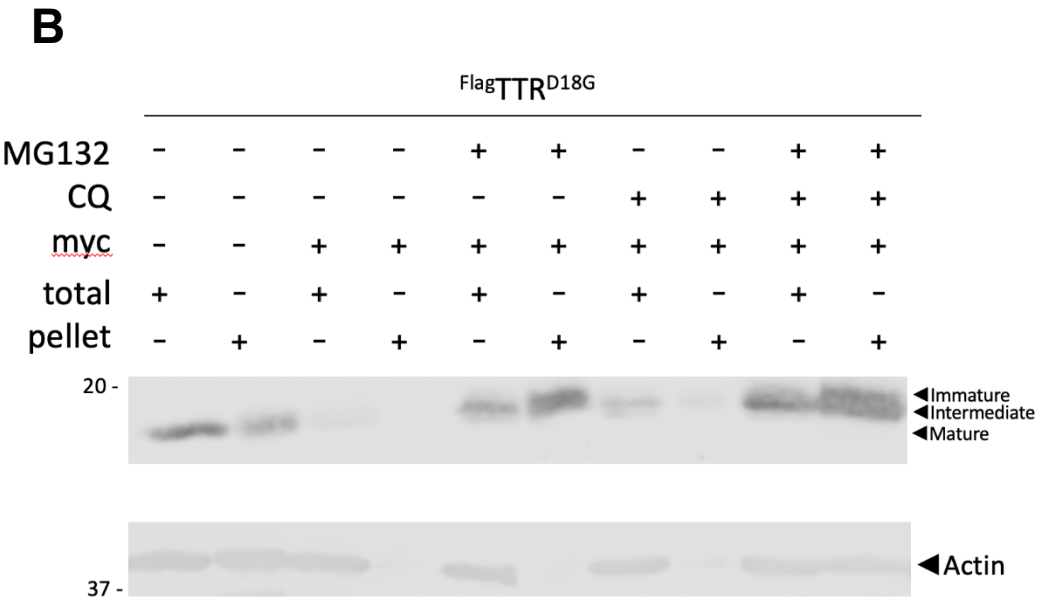

B
FlagTTRD18G
MG132 - - - - + + - - + +
CQ - - - - - - + + + +
myc - - + + + + + + + +
total + - + - + - + - + -
pellet - + - + - + - + - +
20 -
Immature
Intermediate
Mature
37 -
Actin

Figure 3.10: Mistargeted TTR forms aggregates when proteasomal and lysosomal degradation are inhibited. Representative Western blot of total lysates and insoluble pellets of mistargeted TTR^{WT} (A) and TTR^{D18G} (B) with or without inhibitors of degradation. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{D18G}$ followed by 16-hour treatment of myc in the presence of MG132 and/or CQ as indicated. Lysates were subjected to 77,000 g ultracentrifugation and resuspended in 8 M urea for 94 h. Total lysates and insoluble pellets were denatured and reduced for SDS-PAGE and Western blotting.

resuspended the insoluble fraction in 8 M urea for 4 nights to assess whether failure to be translocated into the ER under the impairment of Sec61 results in formation of aggregates of TTR. In comparison between TTR^{WT} and TTR^{D18G}, at basal condition, TTR^{D18G} is extremely prone to aggregation while TTR^{WT} is not (Figure 3.10.A and B, lane 4). While mistargeting happens, in the absence of inhibitors, TTR^{WT} is present while TTR^{D18G} is efficiently cleared out from the cell, resulting in no aggregates of TTR^{D18G}. In the presence of MG132, proteasomal degradation is inhibited, so aggregation of TTR^{WT} and TTR^{D18G} increased (Figure 3.10.A and B, lane 6). In the presence of CQ, lysosomal degradation is inhibited, and TTR^{WT} is aggregation-prone while TTR^{D18G} is not present (Figure 3.10.A and B, lane 8). The insoluble TTR was present on the blot as the intermediate band. Both of MG132 and CQ increased the aggregates of TTR^{WT} and TTR^{D18G}, present in the two immature TTR bands (Figure 3.10.A and B, lane 10). From this result, we confirmed that mistargeted TTR^{D18G} and its proteolytic product are susceptible to proteasomal degradation and form aggregates if it is not degraded. When lysosomal degradation is inhibited by CQ, we did not detect much of mistargeted TTR^{D18G} but TTR^{WT} accumulated and resulted in aggregates. Both of degradation pathways are synergistic in accumulating mistargeted TTR and hence increasing its aggregates in the cytosol.

Physiologically, TTR is highly amyloidogenic due to the dissociation of tetramers into monomers. In our experiment, we observed that TTR is not stable and

misfolded in the cytosol; thus, it is impossible for tetrameric assembly to take place. Misfolded and mistargeted TTR can become aggregates if they are not degraded properly, which was observed in our results regarding TTR^{WT}. Though at the basal conditions, TTR^{D18G} is more aggregation-prone, there was no accumulation of aggregates of mistargeted TTR^{D18G} due to its efficient degradation. We also hypothesized that the proteasome is more active in degrading destabilized TTR^{D18G} to prevent aggregation, while autophagy may mostly degrade aggregates.

Ubiquitination of mistargeted TTR^{WT} and TTR^{D18G} under different inhibitors of degradation

The ubiquitin-proteasome system is strongly dependent on ubiquitin as a degradation signal of proteins (Amm 2014). Before a protein gets degraded by the proteasome, it is polyubiquitinated in an inducible and reversible manner (Pohl & Dikic 2019). The polyubiquinated substrates can be recognized by the shuttle factors or by the proteasome itself to be degraded. Therefore, we investigated the ubiquitination of mistargeted TTR^{WT} and TTR^{D18G} to further confirm their proteasomal degradation. To better visualize the signals of ubiquitin on Western blots, we coexpressed ubiquitin with TTR^{WT} or TTR^{D18G} in HEK293T cells and treated with indicated drugs. Then, they were immunoprecipitated with anti-Flag M2 beads to enable observing the specific degree of ubiquitination of TTR^{WT} or TTR^{D18G} under each pathway, which would indicate the intensity

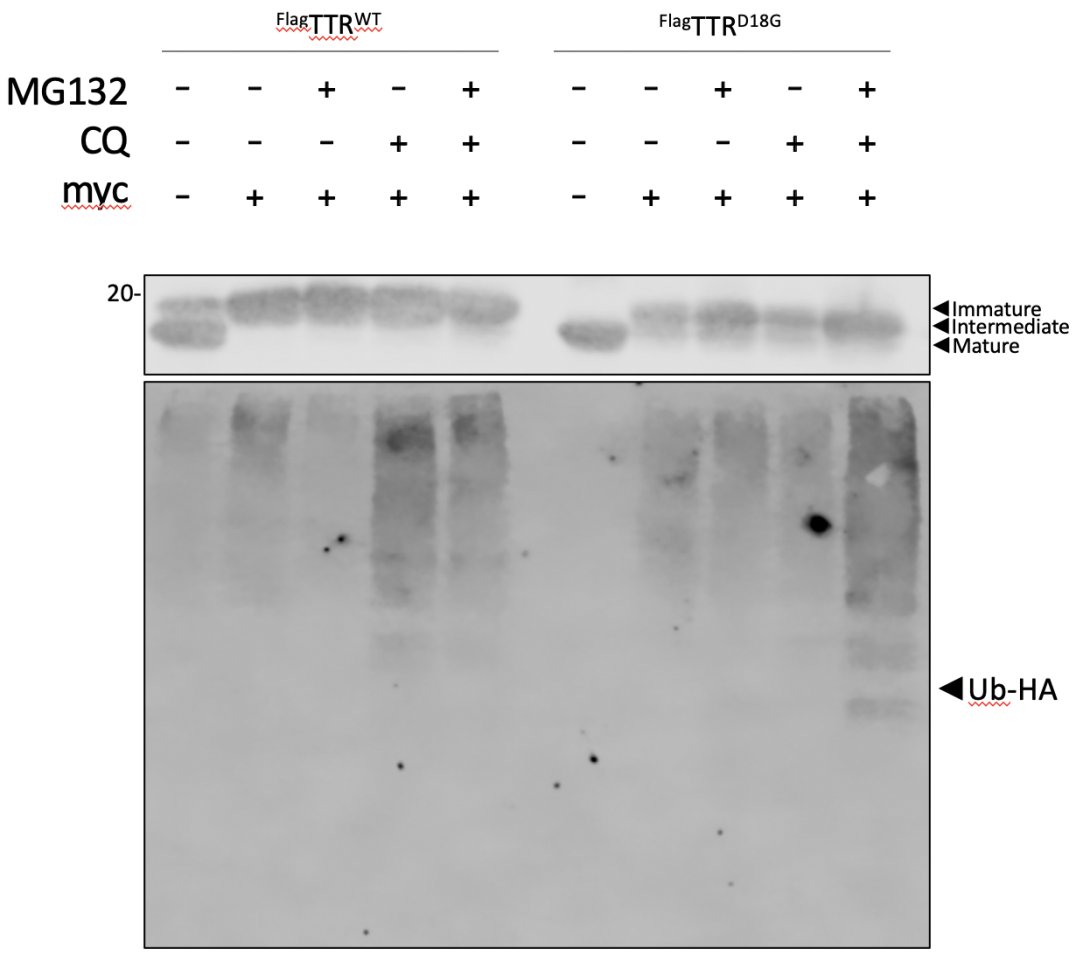

FlagTTRWT
FlagTTRD18G
MG132 - - + - + - - + - +
CQ - - - + + - - - + +
myc - + + + + - + + + +
20-
Immature
Intermediate
Mature
Ub-HA

Figure 3.11: Ubiquitin pull-down from mistargeted TTR^{WT} and TTR^{D18G} with or without inhibitors of proteasomal and lysosomal degradation. HEK293T cells were transfected with $^{Flag}TTR^{WT}$ or $^{Flag}TTR^{D18G}$ followed by 16-hour treatment of myc in the presence of MG132 and/or CQ as indicated. TTR was immunoprecipitated with anti-Flag M2 beads then eluted in Leamli and reduced. The eluates were analyzed by SDS-PAGE and Western blotting.

of degradation in each variant. Since there are deubiquitinylases in the cells which become unregulated during lysis, we included 12.5 mM N-ethylmaleimide in the lysis buffer to inhibit the activity of those deubiquitinylases. The ubiquitin conjugates in the eluate of immunoprecipitated TTR were separated by 12% SDS-PAGE and observed as smears on the blot. The higher intensity of the smears indicated higher ubiquitination of TTR. In the absence of mycolactone, there was no ubiquitination observed in TTR^{D18G} compared to TTR^{WT}, indicating a rapid degradation of this misfolded variant (Figure 3.11; lane 1 and 6). In the presence of mycolactone, we observed a higher intensity of ubiquitination in TTR^{WT} due to its accumulation in the cytosol compared to TTR^{D18G} which is mostly degraded (Figure 3.11; lane 2 and 7). When we inhibited the proteasomal degradation by MG132, surprisingly, the amount of ubiquitination on TTR did not increase in mistargeted TTR^{D18G} and even decreased in mistargeted TTR^{WT} (Figure 3.11; lane 3 and 8). This observation was not expected since ubiquitinated TTR should accumulate when the proteasome is non-functional. When CQ was included as an inhibitor for autophagy, the intensity of mistargeted TTR^{WT} increased leading to a higher ubiquitination, while it stayed the same for mistargeted TTR^{D18G} (Figure 3.11; lane 4 and 9). When both degradation pathways are inhibited, ubiquitination significantly increased in TTR^{D18G} but only slightly increased in TTR^{WT} (Figure 3.11; lane 5 and 10). In the absence of the inhibitors, mistargeted or retrotranslocated TTR^{D18G} was degraded efficiently that resulted in no visualization of ubiquitination compared to TTR^{WT}. The decrease in

ubiquitination in TTR^{WT} and TTR^{D18G} in the presence of MG132 can be attributed to the depletion of the ubiquitin pool. In the presence of CQ, we recovered the ubiquitinated TTR^{WT}, but not TTR^{D18G} because TTR^{D18G} can be efficiently degraded by the proteasome. Thus, we observe the highest signal of ubiquitin when both degradation pathways are inhibited.

DISCUSSION

TTR is amyloidogenic and can be transformed into fibrils that can lead to familial amyloid polyneuropathy (FAP). TTR variants can accelerate TTR amyloidosis due to their thermodynamic or kinetic destabilization. In our study, we found out that the stability of TTR is also a mechanistic factor when it comes to the rate of cellular degradation of misfolded proteins. We inhibited the ER translocation pathway pharmacologically to redirect TTR to the cytosol. The reducing environment of the cytosol can disrupt the disulfide bonds of proteins, leading to unstable structures. Furthermore, the lack of molecular chaperones in the cytosol makes it more challenging for the proteins to be refolded. In fact, the efficiency of ER translocation of proteins is not always 100%. The difference in signal sequences with respect to their properties as well as their requirements for translocation machinery varies their efficiencies of proteins to translocate to the ER (Kim 2002, Levine 2005, Fons 2003, Ng 1996, Voigt 1996). Besides the failure of proteins to enter the ER under normal conditions, acute ER stress was

reported to lower ER translocation of proteins as pre-emptive quality control (Kang 2006). Earlier studies indicated that mistargeted proteins in the cytosol are degraded (Drisaldi 2003, Rane 2004). Another study elaborated this result by stating that Bag6 complex acts as a triage factor in an early step of regulating mistargeted proteins. Bag6 complex distinguishes tail-anchored (TA) proteins from other misfolded ones and hands those TA-proteins over to TRC40 for post-translational ER translocation (Hessa 2011). We extended the triage mechanism of QC in cytosol in further steps involving variants of the same proteins with different energetics. As we did not observe the immature TTR^{D18G} when HSPA13 was overexpressed (Figure 3.2) as observed with TTR^{WT}, we assumed that it is degraded rapidly in the cytosol. Therefore, we set out to investigate how the mistargeted different TTR variants are handled in the cytosol. Our experimental results showed that the more destabilized the variants, the more efficient the degradation is. This observation does not entirely agree with current model of QC for mistargeted proteins which stated that all mistargeted proteins besides TA-proteins are directly degraded by the proteasome. Additionally, we identified another TTR species present in the cytosol that is related to the mistargeting process. This species is specifically degraded by the proteasome and is possibly cleaved by a protease in the cytosol. The formation of this species is also correlated with the stability of TTR, further indicating that the destabilized variants are prioritized to be processed and degraded.

Similar observation was obtained from $A1AT^{WT}$ and NHK.A1AT, in which NHK.A1AT is a highly destabilized variant compared to $A1AT^{WT}$. Both TTR^{D18G} and NHK.A1AT are ERAD substrates that get removed from the secretory pathway to be degraded by the proteasome. When they are mistargeted to the cytosol, these destabilized variants are also rapidly degraded compared to their wild-type structures. Efficient degradation of these proteins regardless of their localization is a protective mechanism of the cell to prevent accumulation of misfolded and aggregation-prone proteins. Madhinavan *et al.* demonstrated that clearance of TTR in *C. elegans* by RNAi or by kinetically stabilizing natively folded TTR lowers TTR aggregation and neuronal dysfunction (Madhinavan 2018). In our aggregation assay, we also observed the absence of mistargeted TTR^{D18G} aggregates in the cytosol due to the efficient degradation, compared to the presence of TTR^{D18G} at basal conditions. When the proteasomal degradation is inhibited, we observed the aggregates of TTR^{D18G}; thus, mistargeted TTR^{D18G} is rapidly processed for degradation. This implicates that toxic aggregates have to be degraded efficiently as a protective control to maintain homeostasis (Chakrabarti 2011).

Ubiquitination is involved in proteasomal or lysosomal degradation and we performed the ubiquitin pull-down to quantify the amount of ubiquitinated TTR^{WT} or TTR^{D18G} for degradation. We observed increase in ubiquitination of TTR^{WT} and TTR^{D18G} in the presence of mycolactone, indicating upregulation of the

degradation machinery towards mistargeted proteins. Surprisingly, in the presence of MG132, we observed ubiquitinated TTR decreasing in both TTR^{WT} and TTR^{D18G}. There is a possibility that ubiquitin gets depleted due to proteasomal inhibition leading to lower ubiquitinated TTR (Wang 2008, Guerra & Callis 2012, Kim 2015).

The most notable result in our study was the proteolytic processing of the immature TTR in the cytosol. Our preliminary peptide mapping data indicated the N-terminal cleavage site of the immature TTR at a phenylalanine or valine residue. One of the established degradation pathways in the cytosol is N-end rule degradation, which involves the degrons containing destabilizing N-terminal residues to be ubiquitinated and degraded by the proteasomes (Varshavsky 2011). These residues would be exposed at the N-terminus by proteolytic cleavage. Our peptide mapping data to obtain the sequence of the proteolytic process was not quantitative enough; thus, we need intact protein analysis to find out the molecular weight of the protein to confirm the identity of the proteolytic product. Further experiments on TTR and A1AT are required to investigate the full mechanism of cytosolic quality control to mistargeted proteins.

REFERENCES

1. Amm, I., Sommer, T. & Wolf, D. H. Protein quality control and elimination of protein waste: The role of the ubiquitin–proteasome system. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**, 182–196 (2014).

2. Besche, H. C. *et al.* Autoubiquitination of the 26S Proteasome on Rpn13 Regulates Breakdown of Ubiquitin Conjugates. *The EMBO Journal* **33**, 1159–1176 (2014).

3. Bulawa, C. E. *et al.* Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *PNAS* **109**, 9629–9634 (2012).

4. Cha-Molstad, H. *et al.* p62/SQSTM1/Sequestosome-1 is an N-recognin of the N-end rule pathway which modulates autophagosome biogenesis. *Nature Communications* **8**, 1–17 (2017).

5. Chakrabarti, O., Rane, N. S. & Hegde, R. S. Cytosolic aggregates perturb the degradation of nontranslocated secretory and membrane proteins. *MBoC* **22**, 1625–1637 (2011).

6. Chen, J. J. *et al.* ATF6 Activation Reduces the Secretion and Extracellular Aggregation of Destabilized Variants of an Amyloidogenic Protein. *Chemistry & Biology* **21**, 1564–1574 (2014).

7. Chen, J. J. *et al.* Endoplasmic Reticulum Proteostasis Influences the Oligomeric State of an Amyloidogenic Protein Secreted from Mammalian Cells. *Cell Chemical Biology* **23**, 1282–1293 (2016).

8. Cheng, S. H. *et al.* Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* **63**, 827–834 (1990).

9. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annual Review of Biochemistry* **75**, 333–366 (2006).

10. Colon, W. & Kelly, J. W. Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry* **31**, 8654–8660 (1992).

11. Demangel, C. & High, S. Sec61 blockade by mycolactone: A central mechanism in Buruli ulcer disease. *Biology of the Cell* **110**, 237–248 (2018).

12. Drisaldi, B. *et al.* Mutant PrP Is Delayed in Its Exit from the Endoplasmic Reticulum, but Neither Wild-type nor Mutant PrP Undergoes Retrotranslocation Prior to Proteasomal Degradation. *J. Biol. Chem.* **278**, 21732–21743 (2003).

13. Fan, H.-C. *et al.* Polyglutamine (PolyQ) Diseases: Genetics to Treatments. *Cell Transplant* **23**, 441–458 (2014).

14. Fons, R. D., Bogert, B. A. & Hegde, R. S. Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *J Cell Biol* **160**, 529–539 (2003).

15. Fregonese, L., Stolk, J., Frants, R. R. & Veldhuisen, B. Alpha-1 antitrypsin Null mutations and severity of emphysema. *Respiratory Medicine* **102**, 876–884 (2008).

16. Goldberg, A. L. Development of proteasome inhibitors as research tools and cancer drugs. *J Cell Biol* **199**, 583–588 (2012).

17. Guerra, D. D. & Callis, J. Ubiquitin on the Move: The Ubiquitin Modification System Plays Diverse Roles in the Regulation of Endoplasmic Reticulum- and Plasma Membrane-Localized Proteins. *Plant Physiology* **160**, 56–64 (2012).

18. Halic, M. & Beckmann, R. The signal recognition particle and its interactions during protein targeting. *Current Opinion in Structural Biology* **15**, 116–125 (2005).

19. Hammarström, P. *et al.* D18G Transthyretin Is Monomeric, Aggregation Prone, and Not Detectable in Plasma and Cerebrospinal Fluid: A Prescription for Central Nervous System Amyloidosis? *Biochemistry* **42**, 6656–6663 (2003).

20. Hegde, R. S. & Keenan, R. J. Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology* **12**, 787–798 (2011).

21. Hessa, T. *et al.* Protein targeting and degradation are coupled for elimination of mislocalized proteins. *Nature* **475**, 394–397 (2011).

22. Hosokawa, N. *et al.* Enhancement of Endoplasmic Reticulum (ER) Degradation of Misfolded Null Hong Kong α1-Antitrypsin by Human ER Mannosidase I. *J. Biol. Chem.* **278**, 26287–26294 (2003).

23. Hosokawa, N., You, Z., Tremblay, L. O., Nagata, K. & Herscovics, A. Stimulation of ERAD of misfolded null Hong Kong α1-antitrypsin by Golgi α1,2-

mannosidases. *Biochemical and Biophysical Research Communications* **362**, 626–632 (2007).

24. Jacobson, D. R., McFarlin, D. E., Kane, I. & Buxbaum, J. N. Transthyretin Pro55, a variant associated with early-onset, aggressive, diffuse amyloidosis with cardiac and neurologic involvement. *Hum Genet* **89**, 353–356 (1992).

25. Kang, S.-W. *et al.* Substrate-Specific Translocational Attenuation during ER Stress Defines a Pre-Emptive Quality Control Pathway. *Cell* **127**, 999–1013 (2006).

26. Kim, J., So, D., Shin, H.-W., Chun, Y.-S. & Park, J.-W. HIF-1α Upregulation due to Depletion of the Free Ubiquitin Pool. *Journal of Korean Medical Science* **30**, 1388–1395 (2015).

27. Kim, S. J., Mitra, D., Salerno, J. R. & Hegde, R. S. Signal Sequences Control Gating of the Protein Translocation Channel in a Substrate-Specific Manner. *Developmental Cell* **2**, 207–217 (2002).

28. Köhnlein, T. & Welte, T. Alpha-1 Antitrypsin Deficiency: Pathogenesis, Clinical Presentation, Diagnosis, and Treatment. *The American Journal of Medicine* **121**, 3–9 (2008).

29. Kowalski, J. M., Parekh, R. N., Mao, J. & Wittrup, K. D. Protein Folding Stability Can Determine the Efficiency of Escape from Endoplasmic Reticulum Quality Control. *J. Biol. Chem.* **273**, 19453–19458 (1998).

30. Lai, Z., Colón, W. & Kelly, J. W. The Acid-Mediated Denaturation Pathway of Transthyretin Yields a Conformational Intermediate That Can Self-Assemble into Amyloid. *Biochemistry* **35**, 6470–6482 (1996).

31. Lee, J. H. & Brantly, M. Molecular mechanisms of alpha1-antitrypsin null alleles. *Respiratory Medicine* **94**, S7–S11 (2000).

32. Levine, C. G., Mitra, D., Sharma, A., Smith, C. L. & Hegde, R. S. The Efficiency of Protein Compartmentalization into the Secretory Pathway. *MBoC* **16**, 279–291 (2004).

33. Lu, K., den Brave, F. & Jentsch, S. Receptor oligomerization guides pathway choice between proteasomal and autophagic degradation. *Nature Cell Biology* **19**, 732–739 (2017).

34. Madhivanan, K. *et al.* Cellular clearance of circulating transthyretin decreases cell-nonautonomous proteotoxicity in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* **115**, E7710–E7719 (2018).

35. Mauthe, M. *et al.* Chloroquine inhibits autophagic flux by decreasing autophagosome-lysosome fusion. *Autophagy* **14**, 1435–1455 (2018).

36. McKenna, M., Simmonds, R. E. & High, S. Mechanistic insights into the inhibition of Sec61-dependent co- and post-translational translocation by mycolactone. *J Cell Sci* **129**, 1404–1415 (2016).

37. Ng, D. T., Brown, J. D. & Walter, P. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* **134**, 269–278 (1996).

38. Pohl, C. & Dikic, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **366**, 818–822 (2019).

39. Rane, N. S., Yonkovich, J. L. & Hegde, R. S. Protection from cytosolic prion protein toxicity by modulation of protein translocation. *The EMBO Journal* **23**, 4550–4559 (2004).

40. Rodrigo-Brenni, M. C., Gutierrez, E. & Hegde, R. S. Cytosolic Quality Control of Mislocalized Proteins Requires RNF126 Recruitment to Bag6. *Molecular Cell* **55**, 227–237 (2014).

41. Sato, T. *et al.* Endoplasmic reticulum quality control regulates the fate of transthyretin variants in the cell. *The EMBO Journal* **26**, 2501–2512 (2007).

42. Schulze, P. C. & Maurer, M. S. Transthyretin Val30Met Mutation in an African American with Cardiac Amyloidosis. *Congest Heart Fail* **16**, 73–76 (2010).

43. Sekijima, Y. *et al.* Energetic Characteristics of the New Transthyretin Variant A25T May Explain Its Atypical Central Nervous System Pathology. *Laboratory Investigation* **83**, 409–417 (2003).

44. Sekijima, Y. *et al.* The Biological and Chemical Basis for Tissue-Selective Amyloid Disease. *Cell* **121**, 73–85 (2005).

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

46. Shao, S. & Hegde, R. S. Target Selection during Protein Quality Control. *Trends in Biochemical Sciences* **41**, 124–137 (2016).

47. Sifers, R. N., Brashears-Macatee, S., Kidd, V. J., Muensch, H. & Woo, S. L. A frameshift mutation results in a truncated alpha 1-antitrypsin that is retained within the rough endoplasmic reticulum. *J. Biol. Chem.* **263**, 7330–7335 (1988).

48. Silverman, E. K. & Sandhaus, R. A. Alpha1-Antitrypsin Deficiency. *New England Journal of Medicine* **360**, 2749–2757 (2009).

49. Sörgjerd, K. *et al.* Retention of Misfolded Mutant Transthyretin by the Chaperone BiP/GRP78 Mitigates Amyloidogenesis. *Journal of Molecular Biology* **356**, 469–482 (2006).

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

51. Susuki, S. *et al.* The Endoplasmic Reticulum-associated Degradation of Transthyretin Variants Is Negatively Regulated by BiP in Mammalian Cells. *J. Biol. Chem.* **284**, 8312–8321 (2009).

52. Varshavsky, A. The N-end rule pathway and regulation by proteolysis. *Protein Sci.* **20**, 1298–1345 (2011).

53. Voigt, S., Jungnickel, B., Hartmann, E. & Rapoport, T. A. Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *J Cell Biol* **134**, 25–35 (1996).

54. Wang, Q., Li, L. & Ye, Y. Inhibition of p97-dependent Protein Degradation by Eeyarestatin I. *J. Biol. Chem.* **283**, 7445–7454 (2008).

55. Wu, Y., Swulius, M. T., Moremen, K. W. & Sifers, R. N. Elucidation of the molecular logic by which misfolded α1-antitrypsin is preferentially selected for degradation. *PNAS* **100**, 8229–8234 (2003).

56. Yoshinaga, T., Takei, Y., Katayanagi, K. & Ikeda, S. Postmortem findings in a familial amyloid polyneuropathy patient with homozygosity of the mutant Val30Met transthyretin gene. *Amyloid* **11**, 56–60 (2004).

CHAPTER 4: PROTEIN PROFILING AND PSEUDO-PARALLEL REACTION MONITORING TO MONITOR A FUSION-ASSOCIATED CONFORMATIONAL CHANGE IN HEMAGGLUTININ

INTRODUCTION

Protein conformation reports on its stability, functionality, and interactions with other proteins or ligands. Therefore, it is necessary for continuing development of technologies that are informative about protein structures. Furthermore, those techniques have to be applicable for in vivo and in vitro experiments, as well as being able to detect endogenous proteins in native conditions. Mass spectrometry (MS) has been known for identifying and quantifying the proteome. Over the past decade, a number of techniques have been added to the toolbox of MS to probe for protein conformations in their complex environments. These include analysis of protein structures at their native condition by chemical modifications or utilizing protease digestion and protein denaturation strategies with quantitative MS-based proteomics. These techniques have constantly evolved to be applicable to a whole-cell approach to study individual protein structures, which facilitates the investigation of the conformational distributions or protein interactions in physiological environments. Traditional biophysical methodologies such as crystallography (Ilari & Savino 2008), nuclear magnetic resonance (Cavalli 2007) or circular dichroism (Kelly 2005) were inherently applied on purified proteins. However, these methods are lower throughput and more limited in probing protein structures. Some of the methods in MS-based structural analysis discussed below, but not all, are more advantageous than those biophysical ones.

Denaturation-based approaches are widely applied to measure protein stability and its kinetics. These approaches employ a gradient of chemicals or temperature to denature the proteins. Consequently, in pulse proteolysis (PP), aliquots of samples are incubated with nonspecific protease such as thermolysin to selectively digest unfolded regions in different concentrations of denaturants (Park & Marqusee 2005). Similarly, in the method of the Stability of Proteins from Rates of Oxidation (SPROX), protein samples denatured in different rates are treated with hydrogen peroxide to report on thermodynamic properties of proteins (West 2010). In Thermal Protein Profiling, protein samples are incubated in a series of different temperature in a given time to denature and aggregate the proteins. The samples are then subjected to ultracentrifugation to separate the soluble (folded) proteins (Molina 2013, Savitski 2014). The quantification for these denaturation-based approaches generally includes quantitative bottom-up proteomics using isobaric mass tags or SILAC (West 2010, Tran 2014).

Hydrogen-deuterium exchange (HDX) is widely applied in vitro to characterize protein folding and stability. This method depends on different proton exchange kinetics based on the local environment of the backbone and the solvent accessibility of the hydrogen bonds (Masson 2019). The kinetics can range from milliseconds for exposed sites to days for buried sites. In particular, a method termed Stability of Unpurified Proteins from Rates of H/D Exchange (SUPREX) involves incubation of protein samples with different concentrations of denaturant

followed by HDX. This enabled measuring $\Delta G_{unfolding}$ of maltose binding protein in complex matrix (Ghaemmaghami 2000). Major challenges of HDX are prevention of back exchange to hydrogen and low identification of proteins in complex samples due to pepsin digestion.

Alternatively, structural analysis of proteins in their native states involves chemical crosslinker or covalent labeling methods followed by bottom-up proteomic methods. Treatment of samples with a chemical cross-linking reagent can detect the inter- and intraprotein sites, providing spatial information about protein structures and protein interaction (Huang 2004, Schulz 2004, Sinz 2006, Bruce 2012). Recently, there have been successful applications in employing crosslinking methods to study protein structures and protein interactions in cellular environment (Sinz 2010, Sinz 2017, Schweppe 2017). The most well-established covalent method is hydroxyl-radical footprinting (HRF), which utilizes hydroxyl radicals generated by UV irradiation (Sharp 2004) or laser photolysis of hydrogen peroxide (Hambly & Gross 2005) to oxidatively modify solvent-accessible residues in an irreversible manner. Different solvent accessibility can report structural changes of proteins resulting from interaction with ligands or other proteins. Notably, HRF has proven to be applicable to proteome in intact cells in nanosecond laser photolysis (Zhu 2017) and fast photochemical oxidation of proteins (FPOP) (Espino 2015, Rinas 2016). Especially, FPOP was recently

developed to probe protein structures *in vivo* with *Caenorhabditis elegans* (Espino & Jones 2019).

To lower the promiscuity of generic labeling reagents, we can choose to profile protein conformation by electrophilic footprinting. Cysteines and lysines are more commonly targeted for labeling due to their nucleophilic activity. Because cysteines are prone to multiple post-translational modifications (PTMs), it is more challenging to analyze the data as the PTMs can be convoluted with the labeling information. On the other hand, lysines can be selectively modified independently of cysteines by amidination reactions because cysteines have to be reduced by 2-bromoethylamine beforehand (Degraan-Weber 2018). NHS-esters are primarily employed in popular isobaric labeling reagents such as TMT and iTRAQ to label solvent-exposed lysines (Zhou & Vachet 2013). Thioimidates were developed and characterized to be efficient labeling reagent on lysines (Beardsley & Reilly 2003, Chang 2011). The advantage of this reagent includes retention of basicity of lysines that the protein conformations are preserved and ionization during electrospray is promoted. Additionally, thioimidates selectively modify lysines that result in homogenous products and mass spectra. Notably, this type of reagent is cell-permeable that allows labeling proteins in their native environment. In an established study comparing the labeling capacity of ribosomal proteins in *E. coli* cells and cellular lysates, they observed more than 90% of exposed lysines are labeled, and the labeling in lysates and in cells are identical (Jaffee 2012). Also,

Misal *et al.* employed amidination to investigate the processing sites of secretory proteins from *Staphylococcus aureus* (Misal 2019). Thus, we set out to develop electrophilic profiling by S-methylthioimidate (SMTA) followed by quantification by selected reaction monitoring to distinguish the conformational change of hemagglutinin upon viral infection.

Influenza virus is a major cause of respiratory disease, causing >300,000 deaths in a typical year, millions to tens of millions of deaths in pandemic years, and extensive loss of economic activity (Iuliano 2018). An obligate step in viral infection is fusion of the viral membrane with the endosomal membrane, which allows for release of the viral RNA-polymerase complex (RNPs) into the host cytosol. This fusion is driven by an acid-induced conformational change in the viral coat protein hemagglutinin (HA). Because fusion is essential for infection, fusion inhibition has been an intensively pursued approach to block influenza infection using small molecules. Although recent studies have elucidated the biochemical mechanism of HA-mediated fusion (Garcia 2015, Das 2018), the sequence dependence, and particularly intracellular factors, that govern HA fusion competence are still unclear. Largely, this is due to a lack of tools to readily characterize HA conformational changes inside the cell.

Mature HA (cleaved to two domains: HA1 and HA2) mediates viral endocytosis into the cell through engagement of the sialic acid receptor by the HA1 domain.

Under the acidic conditions of the early endosome, HA undergoes a global conformational change (Webster & Rott 1987), including extensive remodeling of both HA1 and HA2 domains. As a result, an N-terminal region of HA2, termed the fusion peptide, is exposed. The fusion peptide embeds into the endosomal membrane, promoting fusion with the viral membrane and subsequent release of the viral genome into the cytosol. The most dramatic structural change takes place over the HA2 region in the stalk domain near the C-terminus; acidification and conversion to the post-fusion state leads to a substantial increase in solvent exposure in this region. Given the large change in solvent engagement over the entirety of HA, we hypothesized that profiling reactions would be sensitive to the fusion event.

MATERIALS AND METHODS

Preparation and Purification of Recombinant Hemagglutinins

HEK293T was obtained from ATCC (Manassas, VA, USA). The cells were cultured in complete media containing DMEM (Corning) supplemented with 10% FBS (Seradigm, a division of VWR), 1% L-Glutamine (Corning) and 1% Penicillin Streptomycin (Corning). HA was prepared as previously described (Hai 2012). Briefly, PR8HA-6xHistag (strain A/Puerto Rico/8/1934 H1N1) expressing plasmid was transfected to HEK293T cells by the PEI transfection method. After 72 hours, the cultured media were cleared by low speed centrifugation collected

(3500 rpm, 20 min, 4°C) and the HA was purified by His6 Ni Superflow resin (Clontech) overnight at 4°C. The resin-media were passed through 10 mL polypropylene columns (Qiagen, Germantown, MD, USA). The retained resin was washed 4 times with wash buffer (50mM NaH_2PO_4, 30 mM NaCl, 20 mM imidazole, pH 8.0). The protein was eluted with elution buffer (50mM NaH_2PO_4, 30 mM NaCl, 250 mM imidazole, pH 8.0). The eluate was concentrated using an Amicon 10 kDa MWCO spin column (Merck Millipore, Burlington, MA, USA) and dialyzed 4 times with PBS and quantified by Thermo Nanodrop 3.0 (Thermo, Waltham, MA, USA). Purity was evaluated by Silver Stain (Chevallet 2016).

Acid-induced Activation of Recombinant HA

PR8 was chosen as the HA wild-type model for the study. Purified recombinant PR8 HA was diluted in PBS buffer. To precleave the precursor HA0 to HA1 and HA2, TPCK-treated trypsin was added to the protein at a ratio of 10:1 (protein-to-protease) and incubated with 600 rpm shaking at 37 °C for 5 min. Pre-cleavage was quenched with 1 mM PMSF (Sigma), which remains in the sample solution through later steps. The pH of PR8 HA solution was brought to pH 4.9 by adding citrate pH 3.0 at the final concentration of 7.0 mM. The acidified HA was incubated at ambient temperature for 3 hours, then brought to pH 8.0 by 20 mM phosphate pH 11.00. The neutral control sample was treated in the same manner with citrate pH 7.4 at the final concentration of 7.0 mM. Both acidified and neutral samples were brought to equal volumes by PBS. The fusion inhibitor (Kadam

2016) arbidol-HCl (Fisher) was added as indicated in DMSO following pre-cleavage at a concentration of 400 μM and incubated for 30 min at ambient temperature prior to acidification. Control samples received 4% DMSO vehicle alone.

Trypsin Susceptibility of Acid-induced HA

The treated HA above was incubated at 37 °C for 3 hours to deactivate PMSF, then was digested with TPCK-trypsin (Sigma) at 5:1 (protein:protease ratio) for 30 min at 37 °C, shaken at 600 rpm (White 2015). Immediately after being digested, the samples were denatured and reduced in 1X Laemmli buffer (12% SDS (w/v), 0.06% bromophenol blue (w/v), 47% glycerol (v/v) and 60mM Tris pH 6.8) and 20mM DTT. The denatured protein was boiled at 100 °C for 5 min. The digest was separated on 12% reducing SDS-PAGE gel for silver staining.

SMTA Labeling of Recombinant HA

SMTA was prepared as the hydrogen iodide salt from thioacetamide (Sigma) and methyl iodide (Sigma), as previously described (Beardsley 2003), and aliquots stored under Ar at -80 °C. Aliquots were freshly dissolved prior to labeling in 25 mM HEPES, 50 mM NaCl pH 7.4 at 200 mM SMTA, which was then neutralized with sodium hydroxide. SMTA solution was added to PR8 HA samples, mixed well and let sit for 30 min at ambient temperature. The reaction was quenched by methanol-chloroform precipitation, by adding 4 equivalents methanol, 1

equivalent chloroform, and 3 equivalents water with vortexing after each addition, followed by sedimentation of protein to the interphase at 10,000 *x g* for 5 min. The aqueous layer was removed and the protein pellet washed well with methanol and air-dried.

Sample Preparation for Mass Spectrometry

Air-dried protein pellets were solubilized in freshly prepared 9 M urea in 50 mM Tris pH 8.0. 10 mM aqueous TCEP was added to reduce the disulfide bonds for 30 min in the dark at room temperature. 5 mM aqueous iodoacetamide was added to alkylate the sulfhydryl groups for 15 min in the dark at room temperature. The solutions were diluted with 50 mM Tris pH 8.0 to 2 M urea and digested with sequencing grade porcine trypsin (Thermo) or Glu-C (Thermo) at a ratio of 20:1 protein-to-protease for 18 to 24 hours at 37 °C, shaken at 600 rpm. After digestion, the samples were acidified with 5% formic acid and debris eliminated at 21,100 *x g* for 30 min.

Preparation of complex mixtures containing digested HA

HEK293T cells were harvested by scrapping in PBS and lysed in RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% Triton X100, 0.1% deoxycholate, 0.1% SDS) or by freeze-thaw in PBS. The lysate concentration was quantified by Bradford assay (Bio-rad, Hercules, CA, USA) with the measurement by Varian Cary 60 UV/Vis spectrophotometer (Agilent, Santa Clara, CA, USA). The protein was

precipitated by methanol-chloroform precipitation and prepared as digested peptides similar to the above description. Digested peptides from either acidified or neutral as described above was spiked in lysate at 10% (w/w). 1μg lysate was analyzed by pseudo-PRM in the LTQ Orbitrap Velos Pro (Thermo).

Data-dependent analysis of SMTA-labeled PR8 HA

Digest was injected from the autosampler of an Easy-nLC 1000 to the trapping column containing a Kasil frit and 2-cm 5 μm C18 resin (Phenomenex). High salt concentration was removed by washing with buffer A through the waste line. The analytical column is a pulled 5 μm tip packed with 15- to 20-cm 3 μm C18 resin (Phenomenex, Torrance, CA, USA). The peptides were injected directly onto the trapping column at a flow rate of 5.0 μL/min and separated on the analytical column at a flow rate of 500 nL/min. Buffer A was 5% acetonitrile, 0.1% formic acid, and buffer B was 80% acetonitrile, 0.1% formic acid. The gradient conditions were as follows: 0-5 min: 1-7%B; 5-95 min: 7-55%B; 95-115 min: 55-100%B; 115-120 min: 100%B; 120-125 min: 100-1%B; 125-145 min: 1%B. The eluate was sprayed at 3.0 kV and the ion transfer tube maintained at 275 °C. The mass spectrometer (LTQ Orbitrap Velos Pro) precursor scan was acquired with the AGC set to 10^5 and 50 ms maximum injection time. In all data acquisition strategies, MS2 scans were acquired by CID at a normalized collision energy set to 35 kV.

Targeted Library Generation

Data-dependent acquisition of MS/MS spectra with the linear ion trap in the LTQ Orbitrap Velos Pro were performed with the following settings: MS/MS on the most 10 intense ions per precursor scan, dynamic exclusion repeat count, 1, repeat duration, 30s; exclusion list size, 500; and exclusion duration, 5 s. The raw data was extracted using to MS1 and MS2 spectral files. These files were searched against a database of the PR8 HA sequence and its reversed sequence by the ProLucid algorithm in the IP2 software suite (Integrated Proteomics, San Diego, CA) (Xu 2015). The database was supplemented with forward and reverse (decoy) sequences for 400 common contaminants (e.g. keratins, immunoglobulins, albumen, etc.). The precursor mass tolerance was set to 3000.0 milli-amu, and the fragment mass tolerance was set to 600.0 ppm. The search space included up to 3 missed cleaved tryptic peptides. Carbamidomethylation (+57.02146) of cysteine was considered as a static modification. Amidination of lysine (+41.02655) was permitted as a differential modification. Peptide candidates were filtered using DTASelect for less than 1% false positive peptides (Cociorva 2007).

Pseudo-PRM Analysis

From data-dependent analysis, the precursors were chosen such that their spectral counts are > 3 and they are fully tryptic peptides (shown in Table 4.1). From each precursor, the 3 most intense fragments were chosen for monitoring.

The peptides were injected directly to the trapping column at a flow rate of 5.0 μL/min and separated on the analytical column at a flow rate of 350 nL/min. Buffer A was 5% acetonitrile, 0.1% formic acid, and buffer B was 80% acetonitrile, 0.1% formic acid. The gradient conditions were as follows: 0-1min: 1-7%B; 1-59min: 7-35%B; 59-64min: 35-100%B; 64-69min: 100%B; 69-74min: 100-1%B; 74-94min: 1%B. The monitoring time for each peptide was 10 min. In the linear ion trap, the precursor scan was acquired at normal resolution, with the AGC set to 10^4 and 50 ms maximum injection time. In the FT Orbitrap, the precursor scan was acquired at a resolution of 15000, with the AGC set to $5x10^4$ and 25 ms maximum injection time. The three most intense product ions were chosen for quantification. The integration of each transition was measured from the extracted ion chromatogram from Xcalibur and summed to obtain the total integration of each peptide. The m/z windows used to quantify the transition in the linear ion trap and Orbitrap are 4 m/z and 0.2 m/z, respectively. For each peptide, the evaluation of differential amidination between acidified and neutral conditions of HA was calculated as: Integration of modified peptide/ Integration of unmodified peptide.

Multiple Reaction Monitoring

MRM was performed on a Thermo TSQ Vantage with Easy-nLC II. The digest was injected (Easy-nLC II) onto a C18 column and washed using a similar trapping set-up as described above. Buffer A was 0.1% formic acid, and buffer B

was 100% acetonitrile, 0.1% formic acid. The peptides were injected onto the trapping column at a flow rate of 2.50 μL/min and separated on the analytical column at a flow rate of 230 nL/min. The gradient conditions were as follows: 0-1min: 0-7%B, 1-46min: 7-33%B, 46-47min: 33-80%B, 47-77min: 80-90%B. Trapping and analytical columns were equilibrated before every analysis. The elution time window set for each peptide was 6 min. The spray voltage was set at 1.8 kV and the discharge current was 4.0 μA for all 3 quadrupoles. Collision energies were calculated with the following formula: CE = 0.03*(precursor m/z) + 2.905 for 2+ charge ions and CE = 0.038*(precursor m/z) + 2.281 for 3+ charge ions. The Q1 and Q3 peak width (FWHM) were set to 0.70 and 0.80, respectively. The cycle time was set to 800ms. Each peptide was integrated directly from its extracted ion chromatogram on Xcalibur.

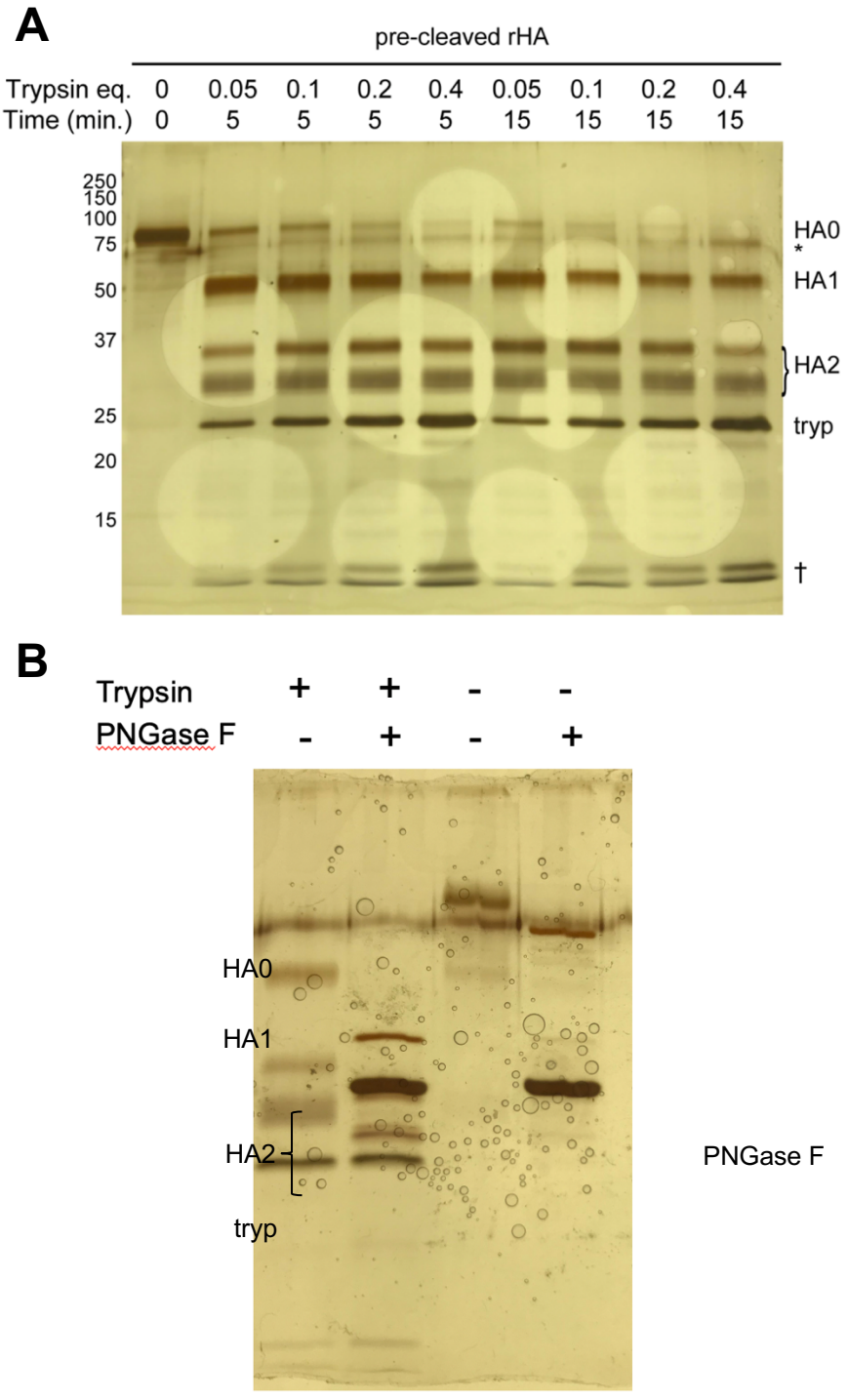
Statistical Analysis

Statistical analysis was performed in Excel. To compare the amidination profile between acidified and neutral conditions, we employed the Student's t-test with the p-value set to 0.05. For the comparison of multiple treatments, one-way ANOVA was employed with the alpha factor set to 0.05, followed by Tukey's test. The sample size is indicated in the figure captions for each analysis. All data are expressed as means ± standard deviations.

RESULTS

Generation of a Library of Amidinated HA Peptides

We chose PR8 HA (strain A/Puerto Rico/8/1934 H1N1) as our nominal "wild-type" HA, as this subtype's structure and fusion has been widely studied (Brunner 1991). To increase recombinant PR8 yield by allowing secretion, the C-terminal trimerization and membrane pass regions were replaced with a bacterial foldon trimerization domain (Wiley 1981, Wilson 1981). We optimized cleavage of full-length HA (HA0) to mature HA1/HA2 by brief TPCK trypsinolysis (Figure 4.1.A). The two bands for HA2 represent glycoforms, as confirmed by coalescence upon PNGase treatment (Figure 4.1.B). Furthermore, citrate acidification of mature HA1/HA2 renders the protein more sensitive to further tryptic degradation (Figure 4.1.B; lanes 3,4) as opposed to the mature protein in PBS (lanes 1,2) or treated with an isotonic neutral citrate buffer (Figure 4.1.B; lanes 5,6). The tryptic sensitivity validates that acidification promotes the conformational change to the post-fusion state. PR8 HA contains 37 lysine residues. To determine their amidination profile, we prepared PR8 HA in both the pre- and post-fusion states and treated each with 20 mM SMTA for 30 min. at pH 8.0. (Figure 4.2.A). SMTA selectively amidinates lysine residues at this pH (Thumm 1987) (Figure 4.2.B). The reaction was quenched by chloroform-methanol precipitation, and the protein reduced, alkylated, and digested by either Glu-C or trypsin. Peptide sequences were determined by data-dependent LC-MS/MS. Qualitatively similar results in peptide identifications were observed for


A
pre-cleaved rHA
Trypsin eq. 0 0.05 0.1 0.2 0.4 0.05 0.1 0.2 0.4
Time (min.) 0 5 5 5 5 15 15 15 15
250
150
100
75
50
37
25
20
15
HA0
*
HA1
HA2
tryp
†
B
Trypsin + + - -
PNGase F - + - +
HA0
HA1
HA2
tryp
PNGase F

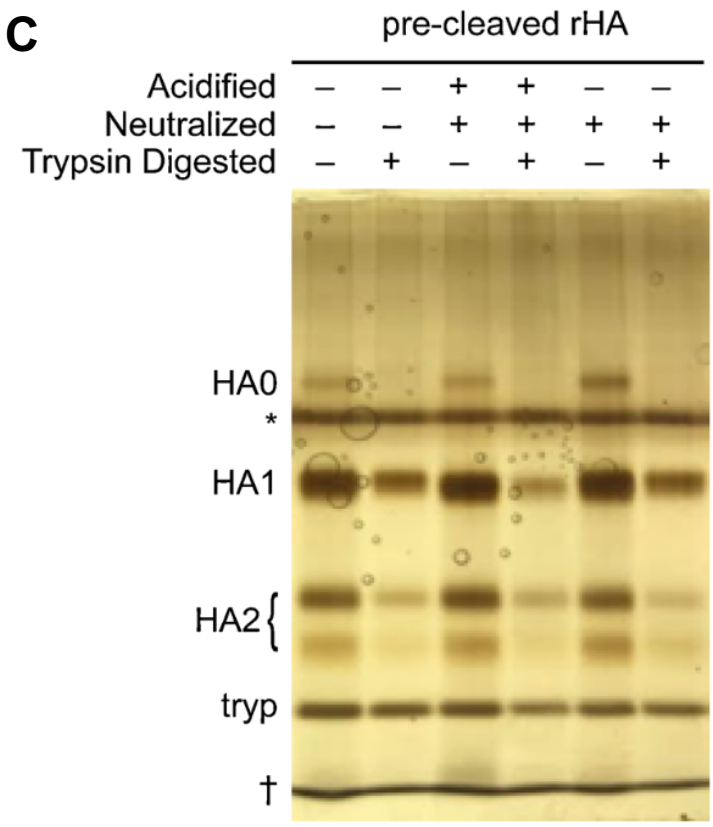
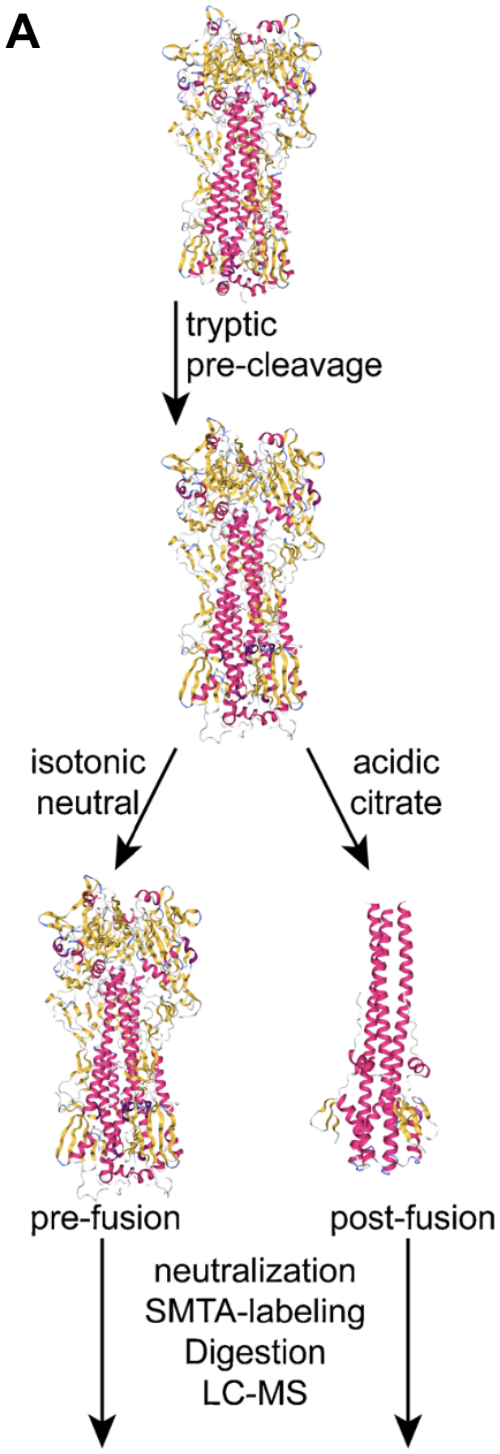
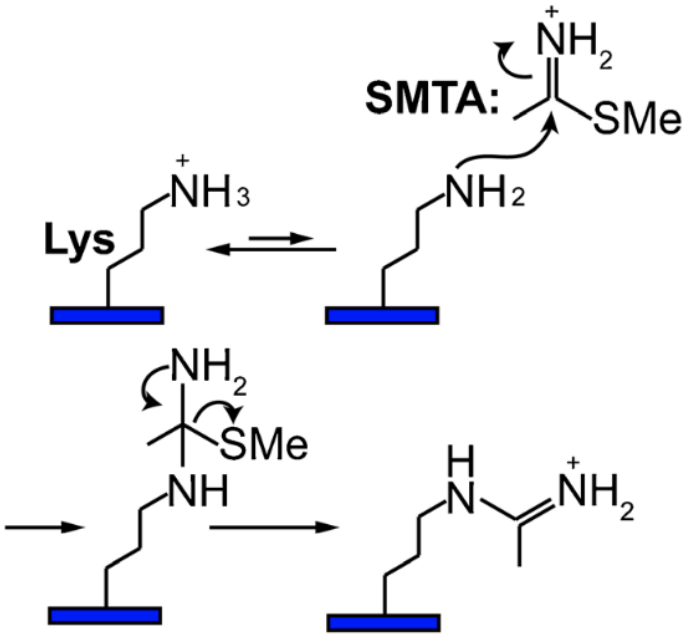

C
pre-cleaved rHA
Acidified – – + + – –
Neutralized – – + + + +
Trypsin Digested – + – + – +
HA0
*
HA1
HA2
tryp
†

Figure 4.1: Optimization of preparation of rHA for SMTA labeling. A. Silver stain of SDS-PAGE separated rHA pre-cleaved by TPCK-treated trypsinization. rHA was incubated for the indicated time at the indicated equivalencies of trypsin (w/w). B. Silver stain of SDS-PAGE separated recombinant PR8 HA (rHA), demonstrating trypsin susceptibility following acid-induced fusion. rHA subjected to brief trypsin precleavage (1:10 w/w TPCK-treated trypsin in PBS for 5 min., 37 °C, 600 rpm) was incubated in either acidic or neutral citrate buffer. Afterwards all samples were brought to pH 8.0, followed by trypsin digestion at 37 °C for 30 min. as indicated. † indicates small degradation products. * indicates residual protein from the media purification of HA; presumably bovine serum albumen. C. Silver stain of SDS-PAGE of PNGase treatment of rHA with and without trypsin digestion.

both pre- and post-fusion HA; hence the following summarizes what was observed for the post-fusion HA. With trypsin modification, we observed 75-85% coverage of the protein (theoretical digest yields 86% coverage by peptides with between 5 and 30 residues long), with 0 lysines observed only with the modification, 9 lysines observed solely without the modification, and 25 lysines observed both with and without amidination. By contrast, Glu-C digestion yielded 60-65% coverage of the protein (theoretical digest yields 68% coverage by peptides between 5 and 30 residues long), with 0 lysines observed solely as the with the modification, 14 lysines observed solely without the modification, and 12 lysines observed both with and without amidination. Hence, we moved forward with trypsin, despite its inability to cleave amidinated lysine (Liu 2007). Interestingly, we found lysine amidination to be distributed throughout the protein sequence, indicating that the PR8 HA is accessible to SMTA across all domains (Figure 4.2.C). Labeling does not appear to be saturating, as although some lysines were never observed to be modified, and other lysines are identified from Peptide-Spectral Matches in either the modified or unmodified state, there are no lysines that we only observed to be amidinated.


A
tryptic
pre-cleavage
isotonic
neutral
acidic
citrate
pre-fusion
post-fusion
neutralization
SMTA-labeling
Digestion
LC-MS

B



C

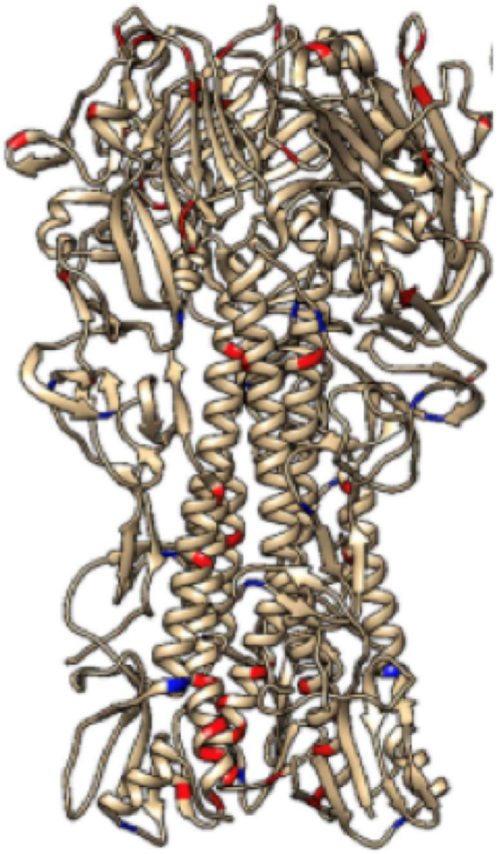


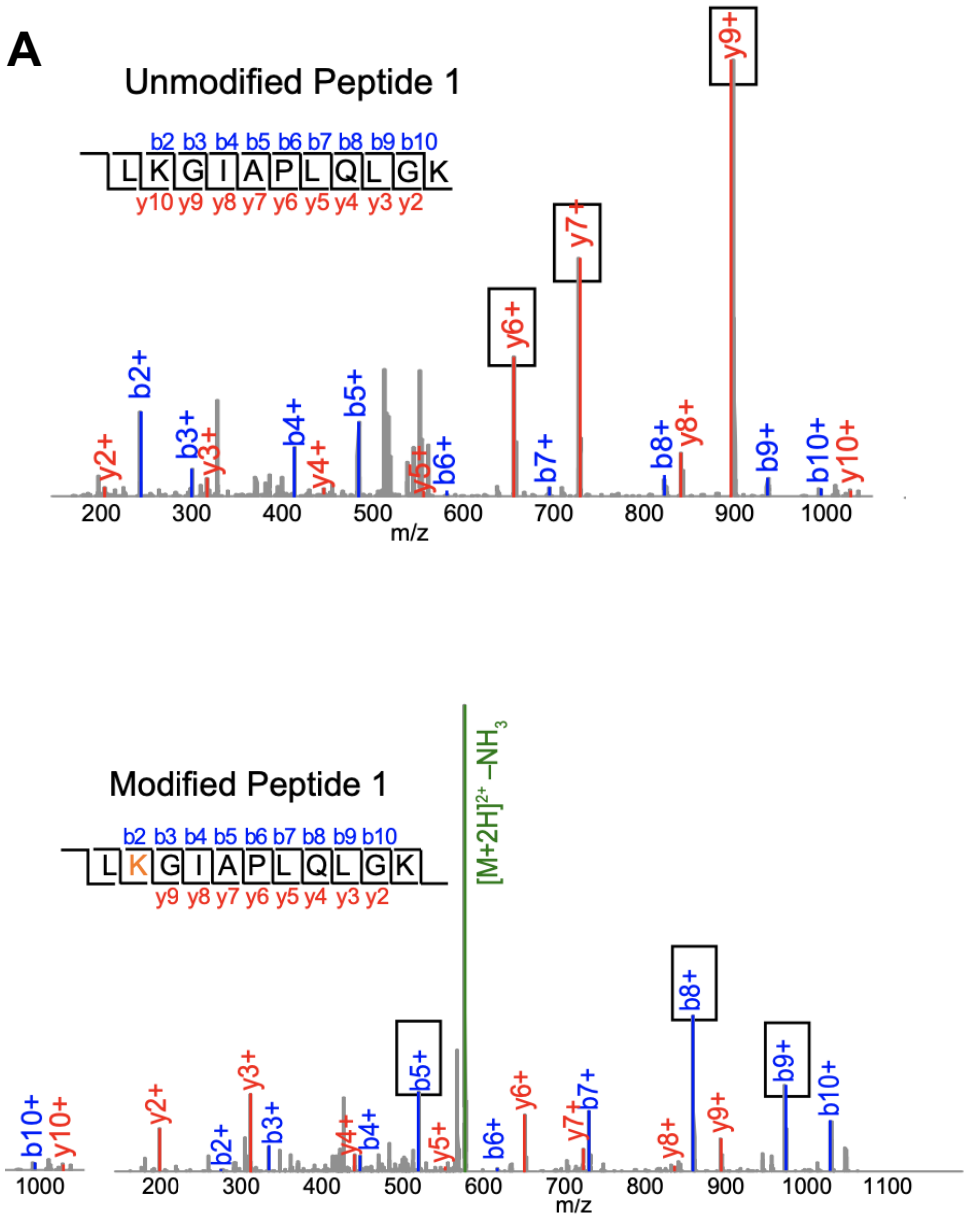
Figure 4.2: SMTA labeling of pre- and post-fusion rHA: A. Flow chart for HA pre-cleavage, fusion, neutralization, and SMTA-labeling. Structures are of H3N2 from PDB 1HA0, 3HMG, and 1HTM and visualized in NGL viewer. B. Reaction mechanism for SMTA modification of the lysine amine. At neutral or greater pH, deamination to the imide is disfavored, driving amidination as shown. C. Map for modification of acidified (post-fusion) SMTA-labeled rHA. The lysines that either do (**red**) or do not (**blue**) react with SMTA (200 μM, pH 7.0, 30 min., ambient temperature), mapped onto a crystal structure for PR8/HA PDB 1RU7 (Gamblin 2004, Petterson 2004)

Pseudo-PRM Quantification of Fusion-dependent HA Amidination

To identify peptides that could serve as markers of lysine amidination, we evaluated the tryptic library (3 runs) for peptides that met the following criteria (Liebler & Zimmerman 2013): 1) robustness of identification, with ≥ 3 spectral counts in each run, 2) fully tryptic, 3) no missed cleavage *except* for at the amidination site, 4) containing lysines that were identified as both modified and unmodified, indicating that there is a dynamic range of modification, 5) lacking methionine, 6) no K/R within 2 amino acids of the cleavage site (except for -1 of course), and 7) charge state of either 2 or 3. We identified 4 amidinated peptides that were worth further consideration based on these criteria (Table 4.1). The three most intense transitions from the library were used for each (Figure 4.3). For normalization, we considered three internal reference peptides (Sherrod 2012): the equivalent unlabeled peptide with missed cleavage at the central lysine, the predicted cleavage product for the unlabeled peptide with the lysine now available for tryptic cleavage, and a peptide elsewhere in the protein generated through arginine cleavage (and hence presumably unaffected by SMTA labeling).

Table 4.1: Peptide Sequences for Reaction Monitoring

Peptide	Lysine	Sequence	Charge state
1	K 62	R.LKGIAPLQLGK	2+
2	K 170	R.NLLWLTEKEGSYPK	2+
3	K 394	K.STQNAINGITNKVNTVIEK	2+
4	K 459	R.TLDFHDSNVKNLYEK	3+


A
Unmodified Peptide 1
b2 b3 b4 b5 b6 b7 b8 b9 b10
L K G I A P L Q L G K
y10 y9 y8 y7 y6 y5 y4 y3 y2
y9+
y7+
y6+
b2+
b3+
y3+
b4+
y2+
b5+
y4+
y5+
b6+
b7+
b8+
y8+
b9+
b10+
y10+
200 300 400 500 600 700 800 900 1000
m/z
Modified Peptide 1
b2 b3 b4 b5 b6 b7 b8 b9 b10
L K G I A P L Q L G K
y9 y8 y7 y6 y5 y4 y3 y2
[M+2H]2+ –NH3
b8+
b5+
b9+
y3+
y2+
b10+
y10+
b2+
b3+
y4+
b4+
y5+
b6+
y6+
y7+
b7+
y8+
y9+
1000 200 300 400 500 600 700 800 900 1000 1100
m/z

B

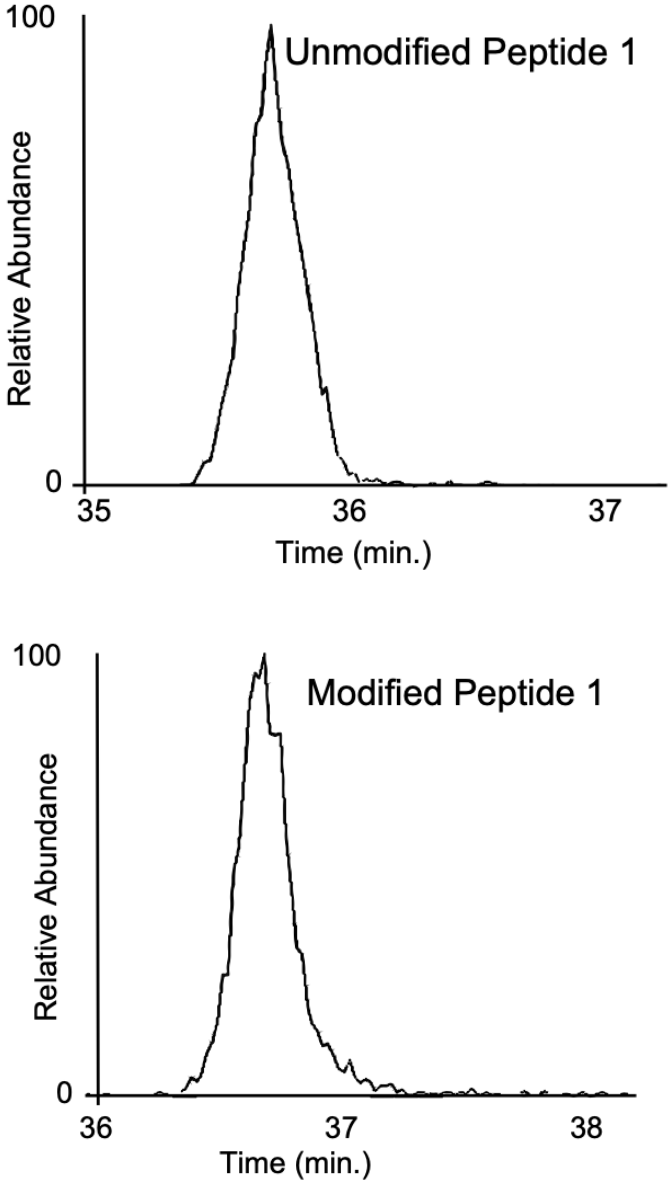

100
Unmodified Peptide 1
Relative Abundance
0
35
36
37
Time (min.)
100
Modified Peptide 1
Relative Abundance
0
36
37
38
Time (min.)

Figure 4.3: Identification and quantification of amidination of rHA by MS. A. CID Fragmentation spectra collected on the LTQ for unmodified and amidinated **Peptide 1** during DDA analysis of acidified, SMTA-labeled rHA digests prepared as in Figure 4.2.A. Identified b/y ions are indicated above and below the peptide sequence. Scan ranges were determined by the instrument. The transitions chosen for MRM, pseudo-PRM and PRM analysis are indicated by boxes. B. Extracted ion chromatograms for unmodified and amidinated **Peptide 1** analyzed during pseudo-PRM in a LTQ Velos ion trap, summed from transitions indicated in Figure 4.3.A.

Surprisingly, despite the inherently increased variance at missed cleavages, normalizing amidinated peptides by the fully cleaved, non-amidinated peptides yielded higher CV values. Therefore, we quantified the modification ratio at each lysine by taking the ratio of the intensity of modified peptide over the intensity of unmodified peptide. SMTA concentration was set to 20 mM, the minimum value that maximized this modification ratio for any of the analyzed peptides. Because we are using an ion trap to characterize three transitions in each peptide, as opposed to the Orbitrap traditionally used for parallel reaction monitoring, we describe this approach as pseudo PRM, analogous to pseudo-SRM, when single transitions are monitored using an ion trap (Sherrod 2012, Abbatiello 2013). Using this approach, we found that amidination at **Peptide 1** (K62) on PR8 HA is significantly higher in the post-fusion (acid-induced) conformational state. More modest effects were observed at **Peptide 2** (K170; increased amidination) and **Peptide 4** (K459; decreased amidination), with no change observed at **Peptide 3** (K394) (Figure 4.4.A-D).

To validate that the conformational switch itself drove the change in SMTA labeling, we considered two orthogonal methods of blocking viral fusion. HA requires pre-cleavage to free the fusion peptide and enable the conformational change (Garten 1981). This precleavage occurs physiologically by endogenous host proteases, but can be accomplished in vitro through brief tryptic digest (Klenk 1975). In the absence of tryptic pre-cleavage, no difference on **Peptide 1**

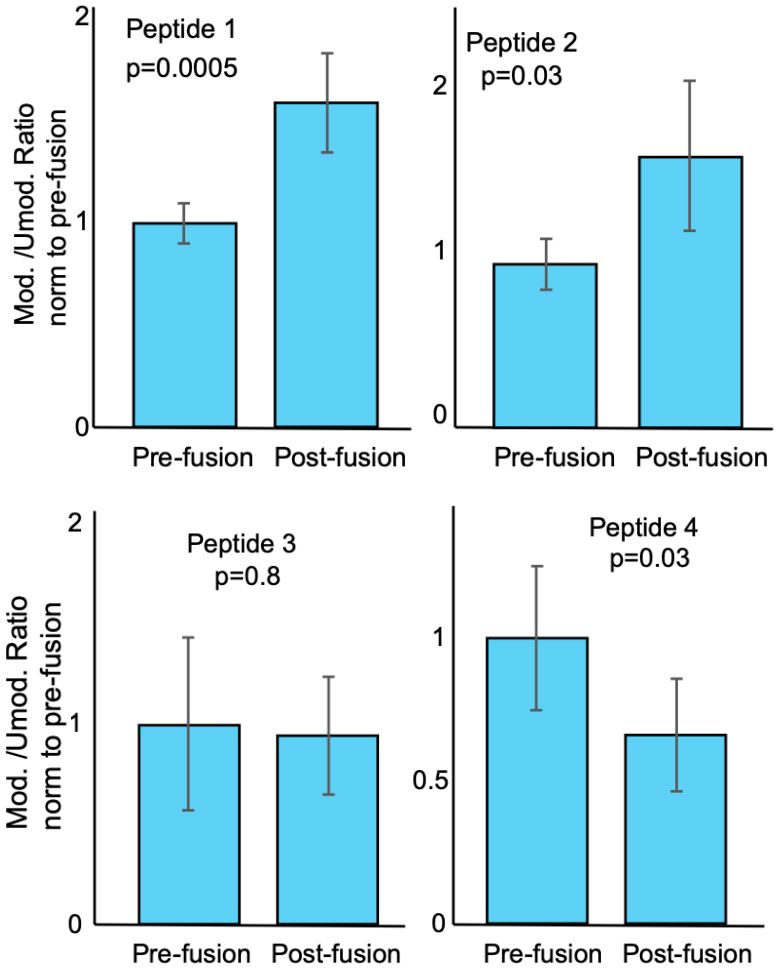
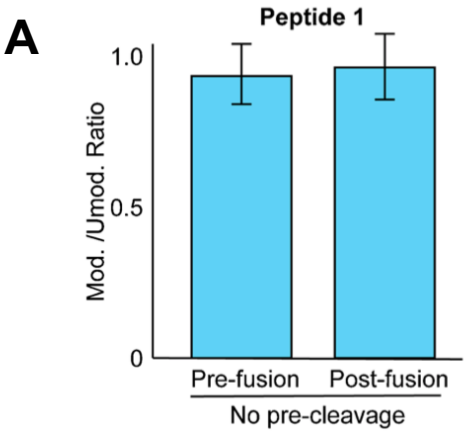
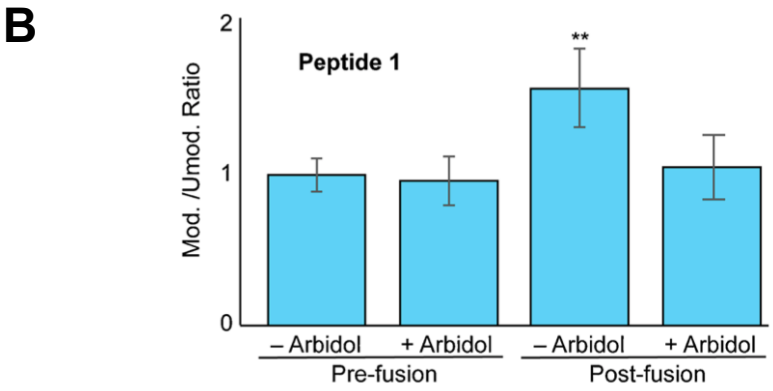

Peptide 1
p=0.0005
Peptide 2
p=0.03
Peptide 3
p=0.8
Peptide 4
p=0.03
Mod. /Umod. Ratio
norm to pre-fusion
Pre-fusion
Post-fusion
0
0.5
1
2

Figure 4.4: Differential amidination of Peptides 1-4 by SMTA. pseudo-MRM quantification of the amidination ratio for **Peptides 1-4** in digests of pre- or post-fusion rHA as indicated. The ratios of amidination are determined from the AUC, as measured by pseudo-MRM, for each target modified and unmodified peptide as described in Table 1. “Pre-fusion” and “post-fusion” preparations are as in Figure 4.2.A. SMTA amidination is performed with 20 mM SMTA, pH 7.0, 30 min. at ambient temperature). Error bars represent standard deviation. p-values were determined by two-tailed unpaired Student’s t test, n = 6 replicates prepared independently


A
Peptide 1
1.0
0.5
0
Mod. /Umod. Ratio
Pre-fusion
Post-fusion
No pre-cleavage


B
Peptide 1
2
1
0
**
Mod. /Umod. Ratio
– Arbidol
+ Arbidol
– Arbidol
+ Arbidol
Pre-fusion
Post-fusion

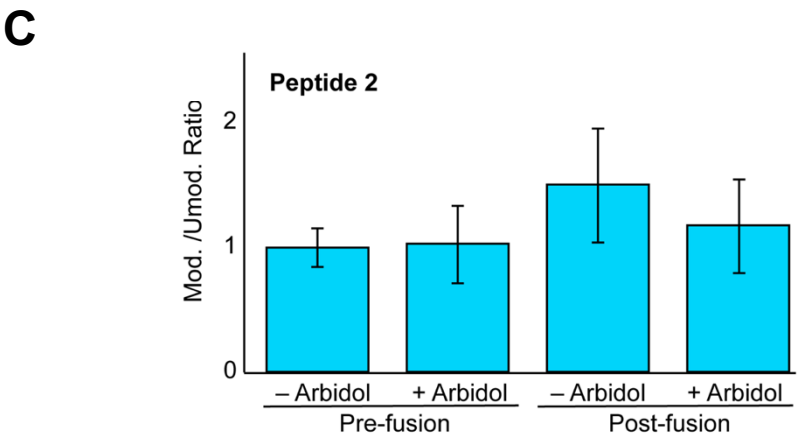

C
Peptide 2
2
1
0
Mod. /Umod. Ratio
– Arbidol
+ Arbidol
– Arbidol
+ Arbidol
Pre-fusion
Post-fusion

Figure 4.5: The amidination assay is sensitive to conditions that interfere with viral fusion. A. Amidination ratios for rHA that has not been tryptically pre-cleaved, subject to acidification and labeling as in Figure 4.2.A. B. Amidination ratios at **Peptide 1** in rHA prepared in either the pre- and post-fusion states following pre-incubation with Arbidol (400 μM, 30 min.), a fusion inhibitor, as indicated. "Pre-fusion" and "post-fusion" preparations are as in Figure 4.2.A. Error bars represent standard deviation. Significance was first assessed by single factor ANOVA, followed by two-tailed Tukey's post-hoc test. ANOVA $F=12$, $p<10^{-4}$, ** Tukey $p < 0.002$. $n = 6$ replicates prepared independently. C. Amidination ratios at **Peptide 2 and 3** in rHA prepared in either the pre- and post-fusion states following pre-incubation with Arbidol (400 μM, 30 min.), a fusion inhibitor, as indicated.

amidination of PR8 HA was observed between acidic and neutral incubation (Figure 4.5.A). Because of the central role of fusion in viral entry and infection, fusion inhibitors have been intensively investigated as antiviral candidates. Arbidol (umifenavir) has been reported to inhibit fusion of H1 serotype HA, including PR8, at 400 μM (Kadam & Wilson 2017). We compared the amidination of target peptides in the presence or absence of arbidol, as well as under neutral or acid incubation conditions. Consistent with the amidination fingerprint serving as a reporter of fusion competence, acidification in the presence of arbidol yielded an amidination profile similar to that of HA that was never exposed to low pH in **Peptide 1** (Figure 4.5.B). Similar, though not statistically significant (ANOVA $p = 0.07$), effects were observed for **Peptide 2** (Figure 4.5.C). Furthermore, in support of this effect being mediated by a direct binding event, **Peptides 3** and **4**, which are present near the arbidol binding pocket (Figure 4.6.A), demonstrated decreased amidination across conditions (Figure 4.6.B), consistent with arbidol's presence in the pocket lowering the accessibility of this residue. In summary, both biochemical and chemical inhibition of fusion yield an electrophilic footprint resembling the pre-fusion state.

Sensitivity and Robustness

To determine the sensitivity of the approach, we measured the LOD for modified and amidinated **Peptide 1** with respect to rHA, finding 190 and 180 fmol respectively (Table 4.2, Figure 4.7.A), with linear response up to 3 pmol. Similar

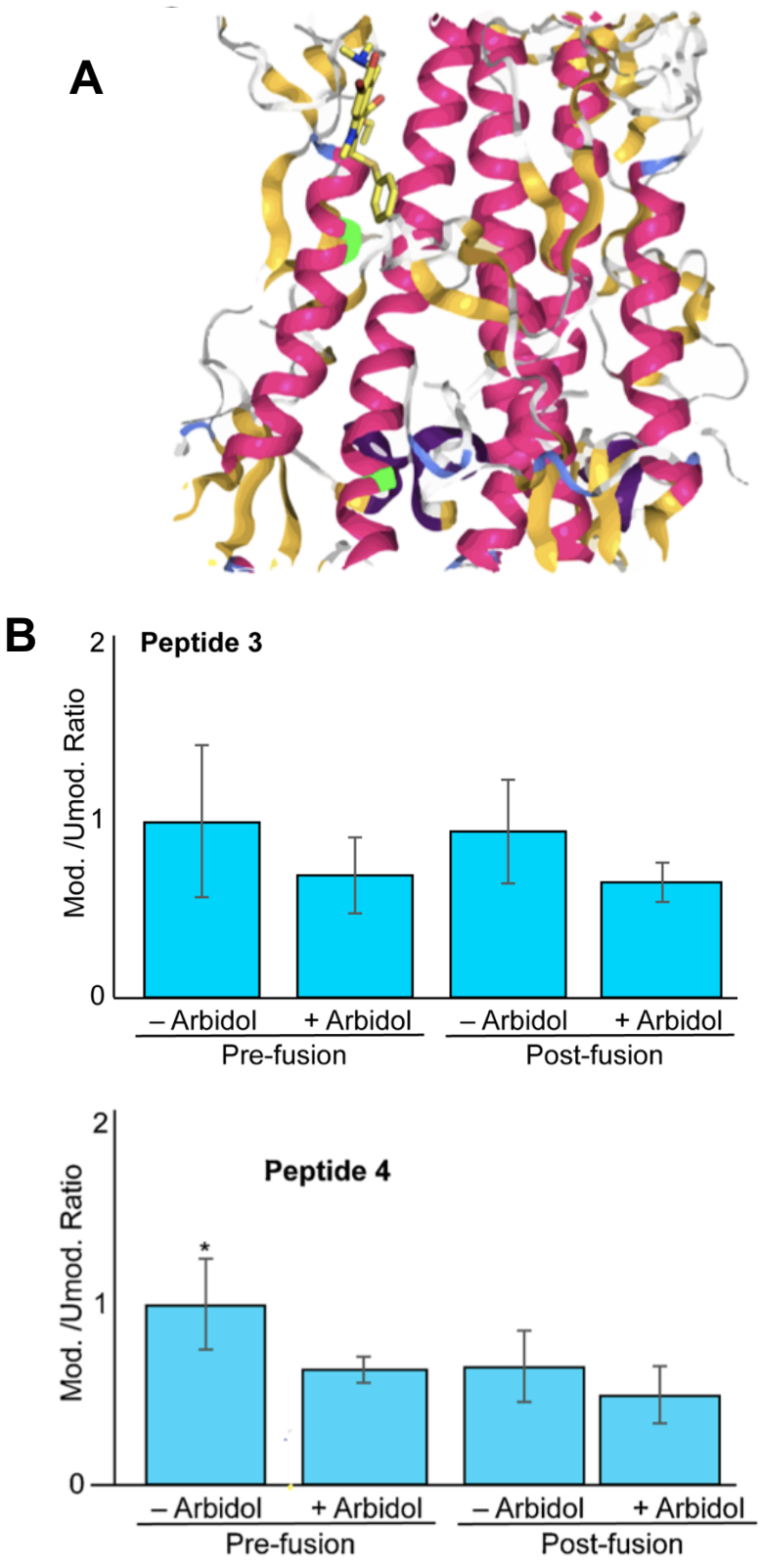
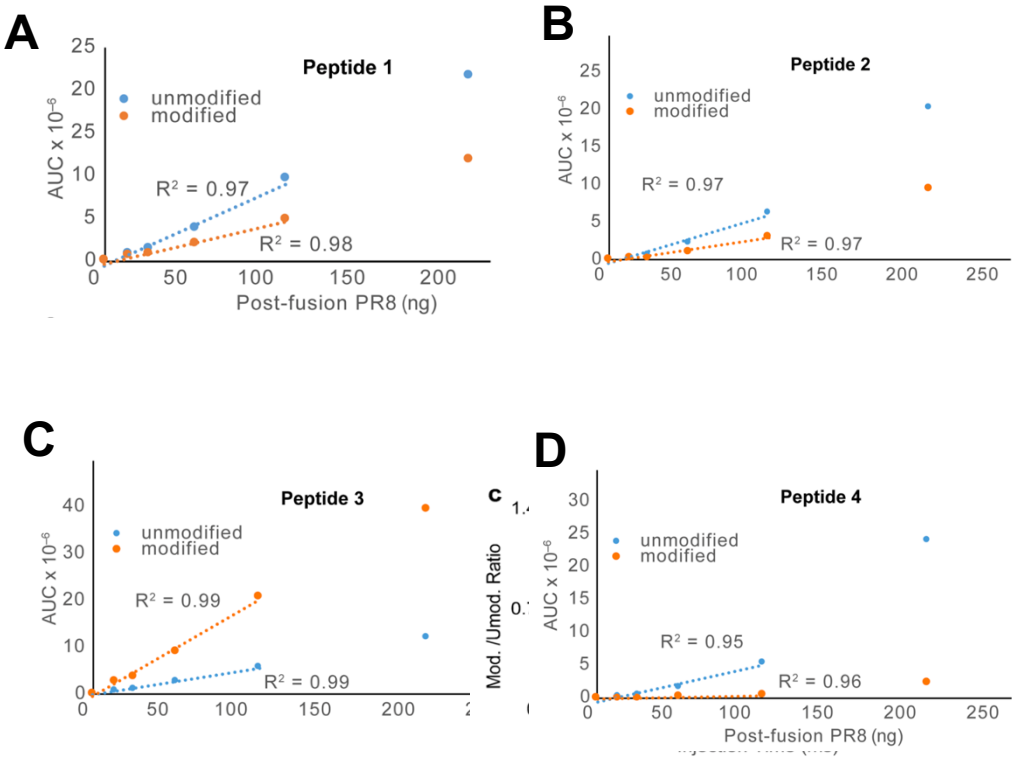
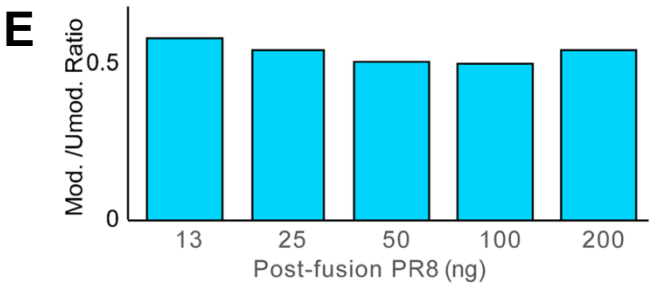

A
B
Peptide 3
Mod. /Umod. Ratio
2
1
0
– Arbidol
+ Arbidol
– Arbidol
+ Arbidol
Pre-fusion
Post-fusion
Peptide 4
*

Figure 4.6: The presence of Arbidol prevents accessibility of SMTA to nearby lysines. A. Arbidol structure superimposed in the binding pocket on a crystal structure for *unbound* PR8/HA, and with the amidination-prone lysines of Peptides 3 and 4 noted in green. Note that the PR8/HA structure would undoubtedly be substantially distorted, as the binding pocket on this isotype is too narrow to accommodate the inhibitor. B. Amidination ratio at **Peptide 4**, as described in Figure 4.5. ANOVA $F=8.3$, $p<10^{-3}$, * Tukey $p < 0.02$

Table 4.2: LODs (from standard deviation of the response) for HA peptides

Peptide	Amidination state	fmol HA
1	Unmodified	188
	Modified	178
2	Unmodified	228
	Modified	235
3	Unmodified	290
	Modified	252
4	Unmodified	78
	Modified	109


A
Peptide 1
unmodified
modified
AUC x 10⁻⁶
R² = 0.97
R² = 0.98
Post-fusion PR8 (ng)
B
Peptide 2
unmodified
modified
AUC x 10⁻⁶
R² = 0.97
R² = 0.97
C
Peptide 3
unmodified
modified
AUC x 10⁻⁶
R² = 0.99
R² = 0.99
D
Peptide 4
unmodified
modified
AUC x 10⁻⁶
R² = 0.95
R² = 0.96
Post-fusion PR8 (ng)
Mod. /Umod. Ratio


E
Mod. /Umod. Ratio
0.5
0
13
25
50
100
200
Post-fusion PR8 (ng)

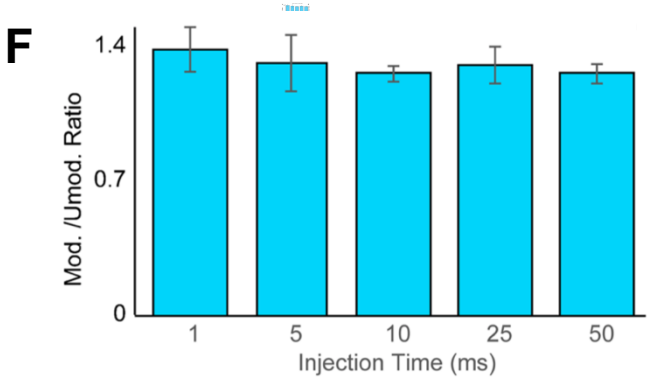

F
Mod. /Umod. Ratio
1.4
0.7
0
1
5
10
25
50
Injection Time (ms)

Figure 4.7: Sensitivity of pseudo-PRM measurement of the peptides. A-D. Linear relationships between AUCs for modified and amidinated Peptide 1-4 and the amount of PR8 injected. E. Minimal dependence of the modification ratio of **Peptide 1** on the amount of PR8 HA loaded between 13 and 200 ng. F. Minimal dependence of the modification ratio of **Peptide 1** on the injection time into the ion trap. Error bars represent standard deviation.

LOD were identified for the other profiled peptides (Table 4.2, Figure 4.7.B-D). Although these LODs are high with respect to typical MRM and PRM thresholds, it should be stressed that these are with respect to the total HA protein concentration, rather than the levels of peptide. The amidination ratio is robust to concentration over this range (Figure 4.7.E). We also considered that the injection time could bias the observed ratio (Sherrod 2012); we maintained an AGC of 10^4 while varying injection time (Figure 4.7.F). Longer injection times allow variance in cycle time when the AGC target is reached, while weaker signals with lower injection times can lead to measurement artifacts due to the Xcalibur noise filter (Savitski 2011). Interestingly, the ratio does not change with injection time, indicating that the AGC target is not getting achieved; the cycle time is consistent across each peak. However, at lower injection times, the variance does increase. To compare to a more traditional MRM approach, we further evaluated amidination of **Peptide 1** on a TSQ Quantiva triple quadrupole mass spectrometer and found similar results, but with greater variation (Figure 4.8).

Evaluation of a mutant form of rHA

A chimeric rHA (cHA) is a mutant form of rHA, which is composed of the head from one HA subtype and the stalk domain from the other (Hai 2012). These play an important role in emerging strategies for universal influenza vaccine development (Krammer 2013). The lack of a fusion assay has limited our

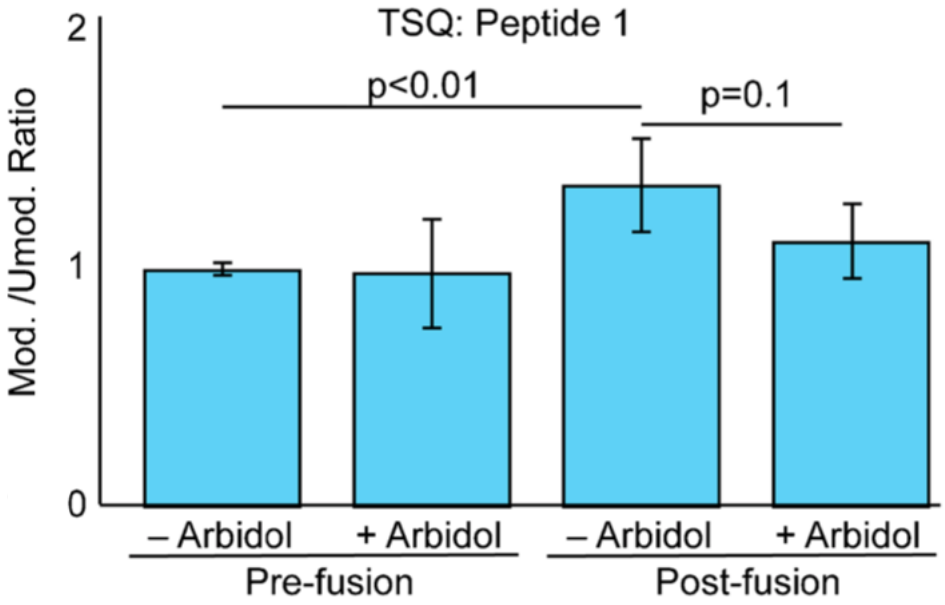
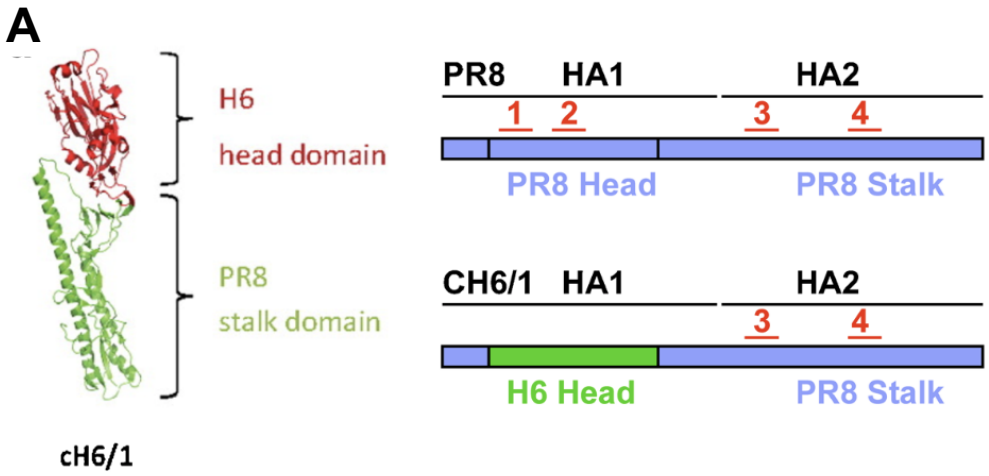

TSQ: Peptide 1
p<0.01
p=0.1
Mod. /Umod. Ratio
2
1
0
– Arbidol
+ Arbidol
– Arbidol
+ Arbidol
Pre-fusion
Post-fusion

Figure 4.8: MRM quantification of **Peptide 1** in rHA digests on a triple-quadrupole mass spectrometer. Significance was first assessed by single factor ANOVA, followed by two-tailed Tukey's post-hoc test. ANOVA $F=6.0$, $p<0.005$. $n = 6$ replicates prepared independently.


A
H6
head domain
PR8
stalk domain
cH6/1
PR8 HA1 HA2
1 2 3 4
PR8 Head PR8 Stalk
CH6/1 HA1 HA2
3 4
H6 Head PR8 Stalk

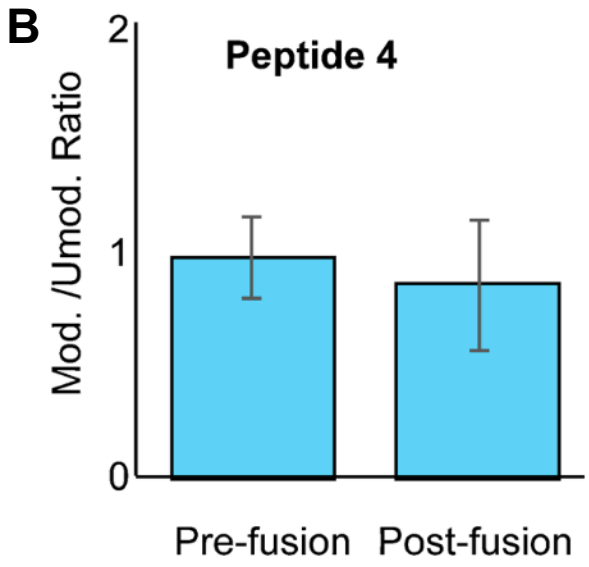
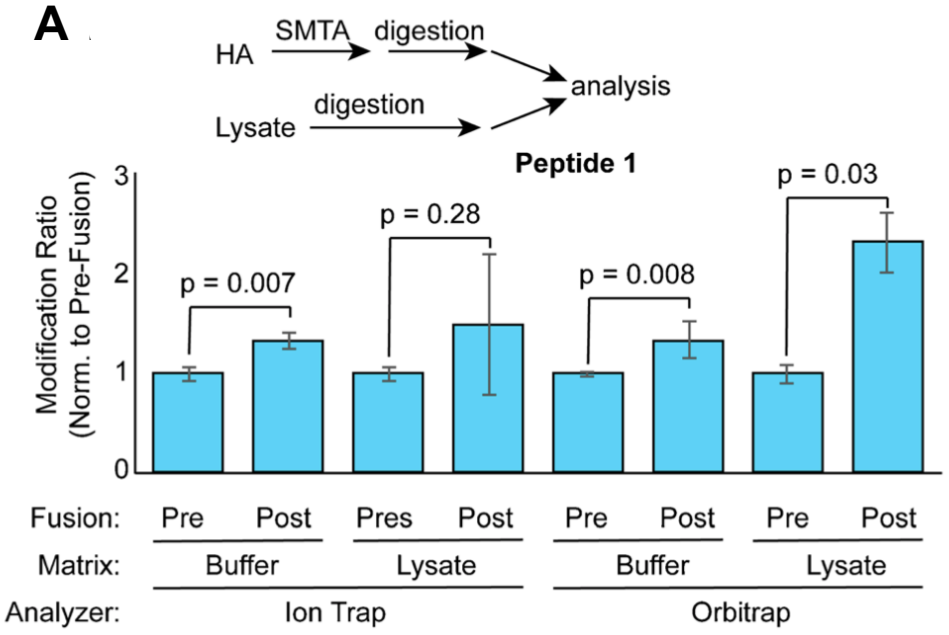

B
Peptide 4
Mod. /Umod. Ratio
2
1
0
Pre-fusion
Post-fusion

Figure 4.9: Evaluation of conformational change of a chimeric rHA. A. Schematic illustrating the design of chimeric c6/1 HA, and comparison to the PR8 HA sequence. Locations of **Peptides 1**-**4** are indicated. **Peptides 1** and **2** are located in the head region of PR8 HA, and hence are absent in the c6/1 HA chimera. B. Differential amidination of **Peptide 4** between pre- and post-fusion states of c6/1 HA.

understanding of the function of the cHA. To this end, we evaluated amidination of one such chimera, cH6/1 (Figure 4.9.A; head from A/mallard/Sweden/81/02 and stalk from A/Puerto Rico/08/34), which contains **Peptide 4** in its PR8 HA stalk domain and whose virus is challenging to generate de novo, implying impaired HA function. Unlike in native PR8 HA, amidination of **Peptide 4** in cH6/1 does not decrease following acidification (Figure 4.9.B), indicating that its conformation is less influenced by acidification, and perhaps less fusion-competent, than that of native PR8 HA.

PRM for Detection in Complex Mixtures

Ultimately, the benefit of PRM profiling of HA fusion, as compared to HDX, would be to identify the fusion-associated conformational switch in complex biological milieu. To evaluate our methods' performance in complex mixtures, we spiked rHA digest into a HEK293T lysate, and compared the measurement of amidination to rHA digest in buffer, as well as between the use of the LTQ ion trap and the Orbitrap (Figure 4.10.A). Both traps perform well, but the high resolution Orbitrap does decrease variance as opposed to the ion trap. Lysate added to labeled PR8 HA prior to tryptic digestion does not affect the Peptide 1 modification ratio. We also considered adding PR8 HA to the lysate prior to labeling. Interestingly, lysate prepared with high detergent RIPA buffer does affect the observed SMTA labeling, but lysate prepared by freeze-thaw of cells in


A
HA
SMTA
digestion
analysis
Lysate
digestion
Peptide 1
Modification Ratio
(Norm. to Pre-Fusion)
3
2
1
0
p = 0.007
p = 0.28
p = 0.008
p = 0.03
Fusion:
Pre
Post
Pres
Post
Pre
Post
Pre
Post
Matrix:
Buffer
Lysate
Buffer
Lysate
Analyzer:
Ion Trap
Orbitrap

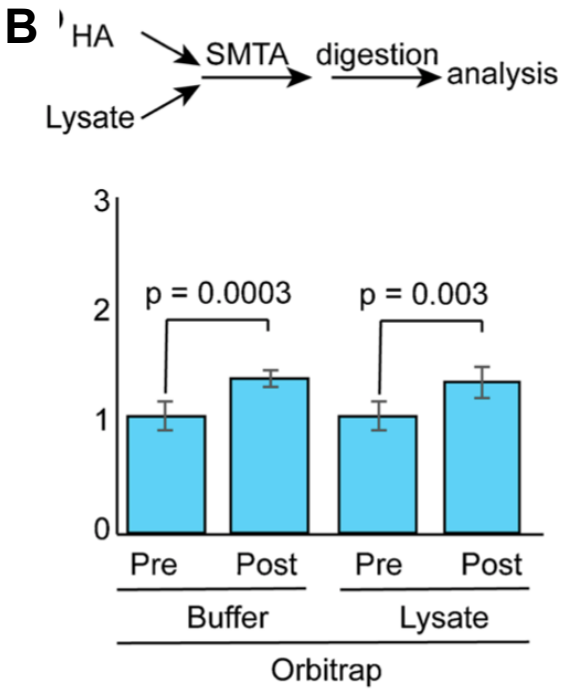

B
HA
SMTA
digestion
analysis
Lysate
3
2
1
0
p = 0.0003
p = 0.003
Pre
Post
Pre
Post
Buffer
Lysate
Orbitrap

Figure 4.10: SMTA labeling and PRM in complex mixtures. A. Pseudo-PRM (ion trap) and PRM (orbitrap) in lysate. Differential amidination at peptide 1 on rHA in the presence or absence of tryptically digested HEK293T lysate. Error bars represent standard deviation. Significance determined from 2-tailed unpaired Students' t test (n = 3 independently prepared samples). B. Differential amidination of 1 ug HA spiked into 5 ug HEK293T lysate prepared by the freeze-thaw method, then SMTA-labeled, digested and analyzed by PRM in the Orbitrap. Error bars represent standard deviation. Significance determined from 2-tailed unpaired Students' t test (n = 6 independently prepared samples).

PBS does not affect the PR8 HA modification ratios measured by PRM (Figure 4.10.C).

DISCUSSION

In this study, we employed a protein model of recombinant wild-type hemagglutinin (PR8 strain) to perform the electrophilic profiling on conformational change of proteins. This form has been validated to form PR8 HA trimers that still undergo acid-dependent fusion, as demonstrated by protease susceptibility following acidic incubation (White 2015). We also confirmed the conformational change of PR8 upon acidification by trypsinolysis. Our protein digestion of PR8 to peptide by trypsin would not be ideal for amidinated lysines; however, we obtained lower coverage of PR8 by digestion with Glu-C. These results are consistent with previous findings that Glu-C yields peptides with poorer chromatographic profile, greater variability in digestion, and up to a three-fold drop in unique peptide identifications (Giansanti 2016). Furthermore, the high density of Glu sites in the HA2 region inherently limits the coverage that can be obtained with this protease. We employed pseudo-PRM as described to quantify the intensity of the selected peptides with and without amidination. With the current labeling conditions, normalization of the modified peptides against their unmodified forms would produce lower variation. We observed a significant increase of amidination in **Peptide 1** in the post-fusion conformation of PR8. Peptide 1 is located in the Fusion (F') subdomain of HA1, far from the receptor

binding site but in a region that undergoes substantial variation with fusion and that stabilizes HA2 in the pre-fusion state (Garcia 2015, Chen 1995). The modest change observed in **Peptide 4** is partly in contrary to a previous HDX study on viral H3 hemagglutinin (2), wherein **Peptide 4** was found to become far more solvent accessible upon fusion. The same study found a modest increase in solvent accessibility post-fusion for the **Peptide 1** region, with little change in the vicinity of the lysines in **Peptides 3** and **4**. These differences could reflect differences between serotypes, or the differences between D_2O accessibility and local lysine reactivity. Also, it is worthwhile to note that these peptides are chosen as reporters of the conformational change that are best suited to reaction monitoring, rather than being regions that themselves drive the fusion-associated conformational change.

The reliance of differential amidination on fusion is validated by its dependence on HA pre-cleavage and the loss of differential amidination upon pre-incubation with the fusion inhibitor Arbidol. Furthermore, the homogeneity of the labeling product allowed straightforward mass spectrometric quantification through MRM on a triple quadrupole instrument, pseudo PRM in a linear ion trap, and PRM in an Orbitrap mass analyzer. Surprisingly, we find that normalization to non-amidinated peptides harboring a missed cleavage as the internal reference peptides provides robust quantification. The ability to discriminate between pre- and post-fusion states, and the robustness of the assay in the presence of

complex matrix interferences, suggest that this approach will allow quantification of intracellular hemagglutinin fusion competence.

One potential application of this approach is to characterize how sequence influences HA fusion competence. Another long-term goal of this methodology is to identify HA fusion yield inside the cell. Under this scheme, we find limits of detection in the 100 to 300 fmol range with respect to HA. At 1000 molecules of HA per virion (Vahey 2019), profiling cells at a multiplicity of infection of 10 would yield about 2 fmol HA in a well of a 96-well plate, in the 0.1 to 1 fmol range often achieved for peptide detection using isotope dilution MRM (Abbatiello 2013, Addona 2009). This method could similarly benefit from isotopically labeled standards. One inherent limitation is the constriction of peptide choice as compared to total protein quantification, wherein the peptides with the best signal, robustness, and freedom from interferences can be chosen. It is possible that in this respect, the product heterogeneity associated with more promiscuous labeling methods such as FPOP might be superior, as the diversity of products increases the number of potential precursor ions that report on individual conformational changes. On the other hand, greater diversity of products might lower the sensitivity to any individual product. Further investigations will compare amidination to other protein labeling strategies as a method to identify HA conformational changes by reaction monitoring of covalent labeling products.

REFERENCES

1. Abbatiello, S. E. *et al.* Design, Implementation and Multisite Evaluation of a System Suitability Protocol for the Quantitative Assessment of Instrument Performance in Liquid Chromatography-Multiple Reaction Monitoring-MS (LC-MRM-MS). *Molecular & Cellular Proteomics* **12**, 2623–2639 (2013).

2. Addona, T. A. *et al.* Multi-site assessment of the precision and reproducibility of multiple reaction monitoring–based measurements of proteins in plasma. *Nature Biotechnology* **27**, 633–641 (2009).

3. Beardsley, R. L. & Reilly, J. P. Quantitation Using Enhanced Signal Tags: A Technique for Comparative Proteomics. *J. Proteome Res.* **2**, 15–21 (2003).

4. Bruce, J. E. In vivo protein complex topologies: Sights through a cross-linking lens. *PROTEOMICS* **12**, 1565–1575 (2012).

5. Brunner, J., Zugliani, C. & Mischler, R. Fusion activity of influenza virus PR8/34 correlates with a temperature-induced conformational change within the hemagglutinin ectodomain detected by photochemical labeling. *Biochemistry* **30**, 2432–2438 (1991).

6. Cavalli, A., Salvatella, X., Dobson, C. M. & Vendruscolo, M. Protein structure determination from NMR chemical shifts. *PNAS* **104**, 9615–9620 (2007).

7. Chang, F.-M. J., Lauber, M. A., Running, W. E., Reilly, J. P. & Giedroc, D. P. Ratiometric Pulse–Chase Amidination Mass Spectrometry as a Probe of Biomolecular Complex Formation. *Anal. Chem.* **83**, 9092–9099 (2011).

8. Chen, J. *et al.* A soluble domain of the membrane-anchoring chain of influenza virus hemagglutinin (HA2) folds in Escherichia coli into the low-pH-induced conformation. *PNAS* **92**, 12205–12209 (1995).

9. Chevallet, M., Luche, S. & Rabilloud, T. Silver staining of proteins in polyacrylamide gels. *Nature Protocols* **1**, 1852–1858 (2006).

10. Cociorva, D., Tabb, D. L. & Yates, J. R. Validation of Tandem Mass Spectrometry Database Search Results Using DTASelect. *Current Protocols in Bioinformatics* **16**, 13.4.1-13.4.14 (2006).

11. Das, D. K. *et al.* Direct Visualization of the Conformational Dynamics of Single Influenza Hemagglutinin Trimers. *Cell* **174**, 926-937.e12 (2018).

12. Degraan-Weber, N., Zhao, B. & Reilly, J. P. Unusual fragmentation of derivatized cysteine-containing peptides. *Rapid Communications in Mass Spectrometry* **32**, 1491–1496 (2018).

13. Espino, J. A. & Jones, L. M. Illuminating Biological Interactions with in Vivo Protein Footprinting. *Anal. Chem.* **91**, 6577–6584 (2019).

14. Espino, J. A., Mali, V. S. & Jones, L. M. In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Anal. Chem.* **87**, 7971–7978 (2015).

15. Gamblin, S. J. *et al.* The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin. *Science* **303**, 1838–1842 (2004).

16. Garcia, N. K., Guttman, M., Ebner, J. L. & Lee, K. K. Dynamic Changes during Acid-Induced Activation of Influenza Hemagglutinin. *Structure* **23**, 665–676 (2015).

17. Garten, W., Bosch, F. X., Linder, D., Rott, R. & Klenk, H.-D. Proteolytic activation of the influenza virus hemagglutinin: The structure of the cleavage site and the enzymes involved in cleavage. *Virology* **115**, 361–374 (1981).

18. Ghaemmaghami, S., Fitzgerald, M. C. & Oas, T. G. A quantitative, high-throughput screen for protein stability. *PNAS* **97**, 8296–8301 (2000).

19. Giansanti, P., Tsiatsiani, L., Low, T. Y. & Heck, A. J. R. Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. *Nature Protocols* **11**, 993–1006 (2016).

20. Hai, R. *et al.* Influenza Viruses Expressing Chimeric Hemagglutinins: Globular Head and Stalk Domains Derived from Different Subtypes. *Journal of Virology* **86**, 5774–5781 (2012).

21. Hambly, D. M. & Gross, M. L. Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J Am Soc Mass Spectrom* **16**, 2057–2063 (2005).

22. Huang, B. X., Kim, H.-Y. & Dass, C. Probing three-dimensional structure of bovine serum albumin by chemical cross-linking and mass spectrometry. *J. Am. Soc. Mass Spectrom.* **15**, 1237–1247 (2004).

23. Ilari, A. & Savino, C. Protein Structure Determination by X-Ray Crystallography. in *Bioinformatics: Data, Sequence Analysis and Evolution* (ed. Keith, J. M.) 63–87 (Humana Press, 2008).

24. Iuliano, A. D. *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *The Lancet* **391**, 1285–1300 (2018).

25. Jaffee, E. G., Lauber, M. A., Running, W. E. & Reilly, J. P. In Vitro and In Vivo Chemical Labeling of Ribosomal Proteins: A Quantitative Comparison. *Anal. Chem.* **84**, 9355–9361 (2012).

26. Kadam, R. U. & Wilson, I. A. Structural basis of influenza virus fusion inhibition by the antiviral drug Arbidol. *Proceedings of the National Academy of Sciences of the United States of America* **114**, (2016).

27. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1751**, 119–139 (2005).

28. Klenk, H.-D., Rott, R., Orlich, M. & Blödorn, J. Activation of influenza A viruses by trypsin treatment. *Virology* **68**, 426–439 (1975).

29. Krammer, F., Pica, N., Hai, R., Margine, I. & Palese, P. Chimeric Hemagglutinin Influenza Virus Vaccine Constructs Elicit Broadly Protective Stalk-Specific Antibodies. *Journal of Virology* **87**, 6542–6550 (2013).

30. Liebler, D. C. & Zimmerman, L. J. Targeted Quantitation of Proteins by Mass Spectrometry. *Biochemistry* **52**, 3797–3806 (2013).

31. Liu, X., Broshears, W. C. & Reilly, J. P. Probing the structure and activity of trypsin with amidination. *Analytical Biochemistry* **367**, 13–19 (2007).

32. Masson, G. R. *et al.* Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods* **16**, 595–602 (2019).

33. Misal, S. A., Li, S., Tang, H., Radivojac, P. & Reilly, J. P. Identification of N-terminal protein processing sites by chemical labeling mass spectrometry. *Rapid Communications in Mass Spectrometry* **33**, 1015–1023 (2019).

34. Molina, D. M. *et al.* Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science* **341**, 84–87 (2013).

35. Park, C. & Marqusee, S. Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. *Nature Methods* **2**, 207–212 (2005).

36. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605–1612 (2004).

37. Rinas, A., Mali, V. S., Espino, J. A. & Jones, L. M. Development of a Microflow System for In-Cell Footprinting Coupled with Mass Spectrometry. *Anal. Chem.* **88**, 10052–10058 (2016).

38. Savitski, M. M. *et al.* Tracking cancer drugs in living cells by thermal profiling of the proteome. *Science* **346**, (2014).

39. Savitski, M. M. *et al.* Delayed Fragmentation and Optimized Isolation Width Settings for Improvement of Protein Identification and Accuracy of Isobaric Mass Tag Quantification on Orbitrap-Type Mass Spectrometers. *Anal. Chem.* **83**, 8959–8967 (2011).

40. Schulz, D. M., Ihling, C., Clore, G. M. & Sinz, A. Mapping the Topology and Determination of a Low-Resolution Three-Dimensional Structure of the Calmodulin−Melittin Complex by Chemical Cross-Linking and High-Resolution FTICRMS: Direct Demonstration of Multiple Binding Modes. *Biochemistry* **43**, 4703–4715 (2004).

41. Schweppe, D. K. *et al.* Mitochondrial protein interactome elucidated by chemical cross-linking mass spectrometry. *PNAS* **114**, 1732–1737 (2017).

42. Sharp, J. S., Becker, J. M. & Hettich, R. L. Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Anal. Chem.* **76**, 672–683 (2004).

43. Sherrod, S. D. *et al.* Label-Free Quantitation of Protein Modifications by Pseudo Selected Reaction Monitoring with Internal Reference Peptides. *J. Proteome Res.* **11**, 3467–3479 (2012).

44. Sinz, A. Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions. *Mass Spectrometry Reviews* **25**, 663–682 (2006).

45. Sinz, A. Investigation of protein–protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal Bioanal Chem* **397**, 3433–3440 (2010).

46. Sinz, A. Divide and conquer: cleavable cross-linkers to study protein conformation and protein–protein interactions. *Anal Bioanal Chem* **409**, 33–44 (2017).

47. Thumm, M., Hoenes, J. & Pfleiderer, G. S-methylthioacetimidate is a new reagent for the amidination of proteins at low pH. *Biochimica et Biophysica Acta (BBA) - General Subjects* **923**, 263–267 (1987).

48. Tran, D. T., Adhikari, J. & Fitzgerald, M. C. Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC)-Based Strategy for Proteome-Wide Thermodynamic Analysis of Protein-Ligand Binding Interactions. *Molecular & Cellular Proteomics* **13**, 1800–1813 (2014).

49. Vahey, M. D. & Fletcher, D. A. Low-Fidelity Assembly of Influenza A Virus Promotes Escape from Host Cells. *Cell* **176**, 281-294.e19 (2019).

50. Webster, R. G. & Rott, R. Influenza virus a pathogenicity: The pivotal role of hemagglutinin. *Cell* **50**, 665–666 (1987).

51. West, G. M. *et al.* Quantitative proteomics approach for identifying protein–drug interactions in complex mixtures using protein stability measurements. *PNAS* **107**, 9078–9082 (2010).

52. White, K. M. *et al.* A Potent Anti-influenza Compound Blocks Fusion through Stabilization of the Prefusion Conformation of the Hemagglutinin Protein. *ACS Infect. Dis.* **1**, 98–109 (2015).

53. Wiley, D. C., Wilson, I. A. & Skehel, J. J. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* **289**, 373–378 (1981).

54. Wilson, I. A., Skehel, J. J. & Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature* **289**, 366–373 (1981).

55. Xu, T. *et al.* ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. *Journal of Proteomics* **129**, 16–24 (2015).

56. Zhou, Y. & Vachet, R. W. Covalent Labeling with Isotopically Encoded Reagents for Faster Structural Analysis of Proteins by Mass Spectrometry. *Anal. Chem.* **85**, 9664–9670 (2013).

57. Zhu, Y. *et al.* Application of Nanosecond Laser Photolysis Protein Footprinting to Study EGFR Activation by EGF in Cells. *J. Proteome Res.* **16**, 2282–2293 (2017).

CHAPTER 4: CONCLUDING REMARKS

In Chapter 2, we investigated an HSPA13 chaperone that lowers ER translocation of TTR, impairing its secretory pathway and promoting intracellular aggregates of TTR. We found that HSPA13 interacts with translocon-associated elements such as Sec61, Sec11, BCAP31, SPCS. This interaction may interfere with the translocation pathway of ER substrates. The ATPase depletion in $HSPA13^{K100A}$ and $HSPA13^{T230A}$ mutants further exacerbate the observed effects of $HSPA13^{WT}$. Overexpression of one of these two mutants substantially increases the immature TTR; hence, the accumulation of TTR aggregates is even more intense due to more products of immature TTR. These results imply that upregulation of HSPA13 can mistarget secretory proteins to the cytosol, which results in intracellular aggregation and degradation. We can overexpress HSPA13 and employ proximity labeling to discover more endogenous substrates of HSPA13 in terms of inhibitory ER translocation. Those substrates can give us some insights about physiological effects of HSPA13. With regards to the molecular mechanism of HSPA13, we observed that removing the ATPase function of HSPA13 worsens that inhibitory effect and enhances higher TTR aggregation. Thus, we assumed that the conformation of HSPA13 induced by the absence of ATPase activity may improve the interaction with the translocon as well as other complementary elements. The strengthened interaction can further inhibit the translocation of other proteins.

In Chapter 3, we proposed a triage mechanism of the cytosolic quality control to mistargeted proteins. We prevented ER translocation of the whole proteome by treating the cells with mycolactone, which impairs Sec61 and plays as a blockade of ER entry. We found that destabilized TTR variants are more efficiently processed and degraded than more stabilized variants. Additionally, immature TTR is processed via an unknown proteolysis before its degradation. Intact analysis of a protein would be a quantitative method of analysis to quantify its size; thus, this method would allow us to confirm the cleavage site obtained from peptide mapping of TTR. Peptide mapping and intact analysis will be conducted to A1AT and NHK.A1AT as well to evaluate if a proteolytic product would be obtained as it was for TTR. To better unravel this quality control mechanism, we have to identify the factors that involve in discriminating between stabilized and destabilized variants to promote the triage mechanism. We have attempted to co-immunoprecipitate mistargeted TTR^{WT} and TTR^{D18G} and subjected the eluates to quantitative MS; however, we were not able to acquire promising interactors for either of the variants. Since several ER chaperones bind to TTR^{D18G}, digitonin extraction would be required to obtain the cytosolic proteome by itself. If we can identify the factors that select the destabilized protein structures to be degraded, we can knock out that factor so that the destabilized proteins cannot be efficiently identified and degraded and we can evaluate the viability of the cells without proper quality control mechanism.

In Chapter 4, we developed a label-free technique followed by pseudo-reaction monitoring (PRM) to probe for conformational change of HA upon activation of viral fusion. SMTA was used to label recombinant HA and the amidination of HA was quantified by PRM. We identified differential amidination of 3 out of 4 selected peptides, which was confirmed to be from the conformational change of viral fusion. We were also able to conduct the labeling reaction of HA in complex lysates, which is promising as an MS-based approach to study protein conformation. There are still spaces to improve this technique. The sensitivity of the quantification by MS is at fmol, which is considered high for other comparable assays; thus, the labeling conditions (concentration of SMTA, duration of incubation) should be optimized to avoid oversaturation of lysine residues. Once we achieve a better sensitivity of this assay, it is worthwhile applying this technique in reporting fusion competence of influenza virus on the platform of drug development towards prevention of influenza infection.