CHARACTERIZATION OF THE BINDING SITE OF STE24 DURING THE -AAXING CLEAVAGE

by

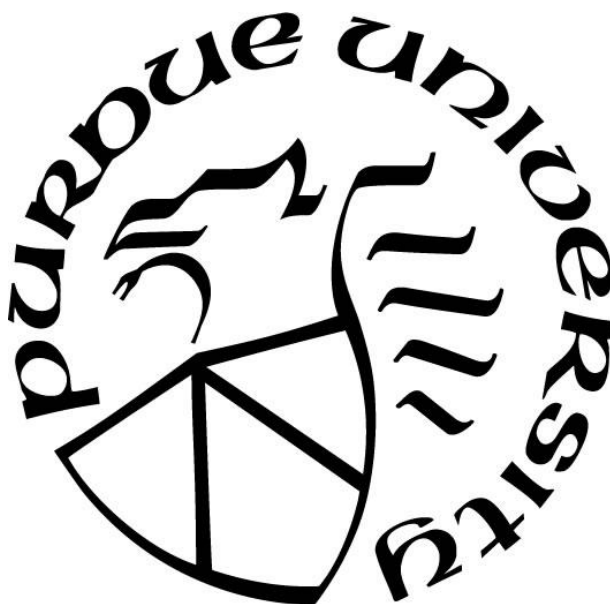**Chelsea Theisen**

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Chemistry

West Lafayette, Indiana

December 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Christine Hrycyna, Chair

Department of Chemistry

Dr. Chittaranjan Das

Department of Chemistry

Dr. Mark Hall

Department of Biochemistry

Dr. Cynthia Stauffacher

Department of Biological Sciences

Approved by:

Dr. Christine Hrycyna

For my friends and family

ACKNOWLEDGMENTS

Finally, I would like to greatly thank my close friends and family. My parents and brother who supported and pushed me through college and gave me the confidence to go to graduate school. Aswathy Chandran, for all our Einstein Bagel breakfasts, all the laughs and fun we had I will never forget. And especially to my amazing, loving husband, who has been my confidant, my rock, my support, and my solace over these past 5 years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure 2.3: Structural analysis of Ste24 residues reveal interesting interactions and possible explanations for changes in activities due to mutations. Images were created in PyMOL Molecular Graphics System using *Sacchromyces mitake* Ste24, PDB: 4IL3, solved in Pryor et al, 2013. A) Electrostatic potential surface representation of Ste24. The zinc ion is shown as an orange sphere. B) Analysis of the TMD 2 and 3 hydrophobic patch. TMD2 is shown in cyan, TMD3 in blue. Residues mutated in the activity studies (Y79, F87 and L139) are shown as green sticks. Additional rings in the TMD2 structure possibly forming pi-interactions with F87 are shown as magenta sticks. C) The proposed binding residues (G255-F266) are shown in green and conserved aspartates (D164 and D280) are shown as sticks. Hydrogen bond interactions between R258/D164 and S256/D280 are displayed. D) Exterior surface representation of the largest Ste24 portal is shown in magenta and the proposed peptide binding region (G255-F266) are shown as green sticks. The zinc ion is shown as an orange sphere. E) Interior cavity surface area shows the opening of the largest portal in the Ste24 structure in magenta. Other internal cavities surrounding the proposed peptide binding region (G255-F266, green sticks) are shown as a gray surface. The zinc ion is shown as an orange sphere.

Figure 2.4: C_{10}-para-a-factor binds in a competitive fashion to the Ste24 a-factor binding site. Samples contained 0.5 µg purified wild-type Ste24, 50 µM C_{10}-para and varying amounts of 15-mer a-factor. The samples were equilibrated on ice and then irradiated with UV light (365 nm) for 30 minutes. 2x SDS was mixed 1:1 with each sample after photolabeling and the samples were run on an SDS-PAGE and then analyzed by immunoblot with Neutravidin®-HRP.

Figure 2.5: Photolabeling analysis of purified wild-type and mutant Ste24 proteins reveal small insights into substrate binding. A) Activity assay results comparing 15-mer a-factor and C10-para substrates. Dots represent the average of duplicate samples in each replicate, bars are at the average of three replicates, error bars represent the standard deviation. The black numbers are the percent activity levels with 15-mer a-factor and blue are for the percent activities for C_{10}-para. (*) indicates mutations that are stasticially significant (p-value < 0.05) between both probes. B) For the photolabeling samples, a pulldown was performed utilizing NeutrAvidin® Agarose resin. Photolabeled protein was eluted from the resin with 50 µL of 2x SDS. Then 10 µL of each sample was used to run on a 10% SDS-PAGE followed by immunoblot analysis with mouse α-HA primary and goat α-mouse-HRP secondary antibody. C) Quantitation of photolabeling signals from (B) was quantified similar to that described at http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf. The bands of both the +UV and -UV samples were quantified with ImageJ and the -UV signal was subtracted from the +UV sample. Then the ratio of the mutant signal was compared to the WT Ste24 signal. The graph was created in Graphpad Prism 9. The bars indicate the average of three replicate photolabeling intensites, dots indicated individual data points and (*) indicates statistical significance (p-value < 0.05). Error bars display the standard deviation.

Figure 2.6: MST and thermostability experiments display direct measurements of K_D and T_m values between Ste24 mutants. A) Purified wild-type and mutant Ste24 proteins were loaded into capillary tubes straight from the concentrated protein stocks, so no protein dilution was performed. The capillaries were then loaded into a Nanotemper Tycho NT 1.6 to measure the T_M of each Ste24 variant. Curves represent average curves of two replicates. B) T_M values of Ste24 mutants obtained from the software of the Nanotemper Tycho NT 1.6, indicative of the inflection point of the curve. Dots represent values of individual replicates. * Indicates statistical significance (p<0.05). C) 20 µL of the concentrated, purified wild-type and mutant Ste24 proteins was used to make a series of

LIST OF ABBREVIATIONS

2µ URA3 P_{PGK}-His₁₀-HA₃-Ste24 – a yeast 2µ replication promotor (high copy number), containing a URA3 gene for uracil production, also containing P_{PGK} promoter for high copy number in *E. coli*, containing the Ste24 gene containing an additional N-terminal His_{10} and HA_3 tag

2µ URA3 P_{PGK}-His₁₀-HA₃-Ste24 – a yeast 2µ replication promotor (high copy number), containing a URA3 gene for uracil production, also containing P_{PGK} promoter for high copy number in *E. coli*, containing the Ste24 gene containing an additional N-terminal His_{10} and HA_3 tag

ABC – Ammonium bicarbonate

ACE-2 – Angiotensin-converting enzyme 2

AEBSF – 4-(2-Aminoethyl)benzenesulfonyl fluoride hydrochloride, serine protease inhibitor

Atm – atmosphere, pressure unit

Axl1 – Axl Receptor Tyrosine Kinase

BSA – Bovine Serum Albumin

CaaX – C-terminal tetra-peptide: cysteine, aliphatic residue, aliphatic residue, any residue

-CC – C-terminal di-peptide: cysteine, cysteine

-CCXX –C-terminal tetra-peptide: cysteine, cystine, any residue, any residue

COVID-19 – Coronavirus 2019, severe acute respiratory syndrome variant

-CXC – C-terminal tripeptide: cysteine, any residue, cysteine

-CXXX – C-terminal tetra-peptide: Cysteine, any residue, any residue, any residue

DDM – N-Dodecyl-beta-Maltoside detergent

DH5α – "Douglas Hanahan" 5 α cells, *E. coli* cells engineered to maximize DNA replication

DNA – Deoxyribonucleic acid

DTT – Dithiothreitol

ER – Endoplasmic reticulum

ESI – Electrospray ionization

FLAG-tag – Protein sequence tag consisting of an Aspartic acid-Tyrosine-Lysine-Aspartic acid- Aspartic acid-Aspartic acid- Aspartic acid-Lysine sequence

FDR – False Discovery Rate

FTase – Farnesyltransferase

FTI – Farnesyl-transferase inhibitor

GGTase – Geranylgeranyltransferase

GTPases – nucleotide guanosine triphosphatase

HA-tag – Human influenza hemagglutinin, protein tag

HExxH – zinc binding motif, Histidine, Glutamic acid, two of any residues, Histidine

HGPS – Hutchinson-Gilford Progeria Syndrome

hICMT – human isoprenylcysteine carboxyl methyltransferase

HIV – human immunodeficiency virus

HRP – horseradish peroxidase

ICMT – isoprenylcysteine carboxyl methyltransferase

IEtOH – iodoethanol

IFITM3 – Interferon Induced Transmembrane Protein 3

INM – Inner Nuclear Membrane

KASH- Klarsicht, ANC-1, Syne homology

KCl – Potassium chloride

K_D – Dissociation constant, concentration of substrate when half the protein target is bound

KDN/YII – Lysine-Aspartic Acid-Asparagine, cut site, Tyrosine-Isoleucine-Isoleucine

K_I – Inhibitory constant, concentration of inhibitor where half the protein target is inhibited

K_M - Michaelis-Menten constant, concentration of substrate when enzyme is at half V_{Max}

Kpsi – kili-pounds per square inch

LAD – Lamina-associated domains

LAP – Lamina-associated protein

LAP2α – Lamina-associated polypeptide 2α

LINC – Linker of Nucleoskeleton and Cytoskeleton complex

LMNA – Lamin A/C gene

MAD – Mandibuloacral dysplasia

MaICMT – *Methanocarcina acetivorans* ICMT

MAPK – Mitogen-activated protein kinase

MATa trp1 leu2 ura3 his4 can1 ste24Δ::LEU2 rce1Δ::TRP1 – yeast strain containing the MATα allele, Tyrosinase-related protein 1, orotidine 5-phosphate decarboxylase (for Uracil production), histidine biosynthesis trifunctional protein, Arginine permease CAN1, Ste24 deletion, beta-isopropylmalate dehydrogenase for leucine production, RCE1 deletion

MDa – Megadalton, atomic mass unit

MgCl₂ – Magnesium chloride

MS – Mass spectrometry

MSMS – Tandem mass spectrometry, mass spectrometry

MST – Microscale thermophoresis

MWCO – Molecular weight cutoff

Myc-tag – C-myc polypeptide used for protein tags

NaCl – Sodium chloride

NaOH – Sodium hydroxide

NE – Nuclear envelope

NPC – Nuclear pore complex

OD### - Optical density, wavelength in nanometers

ONM – Outer nuclear membrane

PAI-1 – Plasminogen Activator Inhibitor-1

PAL-Photoaffinity labeling

PBS – Phosphate buffered saline

PBST – Phosphate buffered saline with Tween-20

pCH#### - plasmid "Christine Hrycyna" plasmid number

PDB – Protein Databank

RAM1/RAM2 – Protein farnesyltransferase subunits

RCE1 – Ras Converting Enzyme 1

RD – Restrictive Dermopathy

RMSD – Root-mean-square deviation

RNA – Ribonucleic acid

SAH – *S*-adenosyl-*L*-homocysteine

SAM – *S*-adenosyl-*L*-methionine

SC-URA – Single dropout media using Synthetic Complete Supplement mixture without uracil

SDS – Sodium Dodecyl Sulfate

SDS-PAGE – Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis

SM3614 – "Susan Michaelis" yeast strain number

Ste14 – "Sterile" 14

Ste24 – "Sterile" 24

Ste3 – "Sterile" 3

Ste6 – "Sterile" 6

SUN – Sad1p, UNC-84 protein domain

TcICMT – *Tribolium castaneum* ICMT

TEP – Triethyl phosphate

TGF-ß – Transforming growth factor ß

TMD – Transmembrane Domain

Tris-HCl – Tris hydrochloride

UV – Ultraviolet

V_{Max} – Maximum product formation rate by a given enzyme concentration

Wnt-ß-catenin – canonical pathway for ß-catenin accumulation

WT - Wild-type

YIIKGVFWDPAC(Fr)-VIA – 15-mer version of **a**-factor, Tyrosine – Isoleucine – Isoleucine – Lysine – Glycine – Valine – Phenylalanine – Tryptophan – Aspartic acid – Proline – Alanine – farnesylated Cysteine – Valine – Isoleucine - Alanine

ZMPSTE24 – Zinc Metalloprotease Sterile 24

α-HA – antibody targeting HA-tag

α-mouse – antibody targeting primary mouse antibodies

Δste24 – Deletion of the Ste24 gene

ABSTRACT

ZMPSTE24 is a seven transmembrane domain zinc metalloprotease that resides in the ER and inner nuclear membranes of mammalian cells. The crystal structures of both the mammalian and yeast homologs, ZMPSTE24 and Ste24, respectively, were solved recently and revealed a common novel structure. Both structures contain a large chamber of mixed hydrophobicity that is capped on both sides. The canonical catalytic HExxH zinc-binding motif lies inside the chamber. Defects in the enzymatic function of human ZMPSTE24 have been shown to cause premature aging disorders. In addition to the well-defined role ZMPSTE24 and Ste24 play in the maturation of prelamin A in mammals and **a**-factor in yeast, both proteins have been proposed to play protective roles in Type 2 diabetes and viral infections by interactions with the cellular translocon. ZMPSTE24 can also be inhibited by several common HIV aspartyl protease inhibitors, possibly causing the frequent and common side-effects of these prescribed drugs. As of now, no precise location for substrate binding has been identified in either ZMPSTE24 or Ste24. Thus, the goal of this project is to localize residues in the enzyme that are important for substrate binding. The yeast homolog Ste24 was used as a model system as it functionally complements the mammalian enzyme and can be reliably cloned, overexpressed, and purified in an active form.

Three approaches were taken to directly determine the K_D values for substrates of Ste24. The ability to perform a direct analysis of K_D values of Ste24 mutations was successfully optimized using microscale thermophoresis. Through K_D analysis, the Ste24 mutation G255A, while completely inactive, does not prevent substrate binding. Alternatively, L441A and L410A mutations showed both an increase in thermal stability and a decrease in binding affinity, that could explain their lower activity levels. A photoaffinity labeling-based proteomics experiment was utilized to precisely locate the site of the prenyl group to a hydrophobic patch lying just under a side portal of Ste24, near K234, during the -aaXing cleavage of **a**-factor maturation. To assess the method of inhibition of HIV protease inhibitors on Ste24 the conserved aspartate mutants were explored. All mutations of these aspartate residues resulted in a severe loss of Ste24 function and instability of the protein.

CHAPTER 1.   INTRODUCTION

1.1   CaaX processing

CaaX processing is a set of three post-translational modifications that occur on proteins that end with a CaaX motif (**Figure 1.1A**). The C in the motif stands for cysteine, a is an aliphatic amino acid, and X is one of several different amino acid residues. The first modification is the prenylation of the CaaX cysteine residue. This prenylation can be either farnesylation or geranylgeranylation, as described below. The second step is a proteolysis reaction that removes the three C-terminal residues and is often called the -aaXing step. Finally, the substrate is carboxylmethylated at the newly exposed C-terminal cysteine. Several proteins that undergo these post-translational processing steps have been identified, including the nuclear lamins, the yeast mating pheromone **a**-factor, and the Ras family of GTPases (1-7). These modifications are believed to promote hydrophobicity and enhance localization/functionality of membrane associating proteins, such as Ras and Lamin proteins (8).

1.1.1   Prenyltransferases

Protein prenylation is the first step in CaaX processing. Prenylation adds either a 15-carbon farnesyl group, or a 20-carbon geranylgeranyl group to the cysteine in the CaaX motif via a thioether bond (7). The significance of this step is believed to be that adding a large hydrophobic group onto the protein allows the protein to be trafficked properly within the cell and associate with the intracellular or plasma membranes. In the majority of cases, the X residue of the motif appears to determine which modification gets added to the CaaX protein (9). Proteins such as Ras, **a**-factor and nuclear lamins end with a serine, cysteine, methionine, or alanine, and are farnesylated. Other proteins, such as heterotrimeric G proteins or Ras-like proteins end with a bulkier or hydrophobic X residue, such as leucine or phenylalanine, and are geranylgeranylated (10).

The identification of modified proteins in yeast allowed for homologs to be discovered in other species (7). In yeast, the two major prenylation enzymes are the farnesyltransferase, RAM1/RAM2 heterodimer, and the geranylgeranyltransferase, Ram2p/Cdc43p heterodimer (9). The yeast mating pheromone, **a**-factor, is farnesylated by Ram1/Ram2 and this modification is believed to aid **a**-factor with the nuclear and ER membrane association specifically (9). Trafficking

tot these intracellular membranes allows the **a**-factor precursor to be proteolyzed in the next step of CaaX processing, catalyzed by either the RCE1 or Ste24 proteases (**Section 1.1.2**) (9).

In mammalian cells, there are three prenyltransferases: protein farnesyltransferase (FTase), protein geranylgeranyltransferase-I (GGTase-I), and Rab geranylgeranyltransferase (GGTase-II or Rab GGTase) (7). All three functional proteins are active as heterodimeric, Zn^{2+}- dependent enzymes (8,11,12). The FTase and GGTase-I enzymes are similar in structure and function with the main difference being in the size of the isoprenoid binding site (10,13). The FTase uses an isoprenoid binding cleft to hold the 15-carbon farnesyl pyrophosphate and the peptide binding site to properly align the CaaX motif of the peptide substrate. It then uses a zinc catalyzed reaction to add the hydrophobic chain onto the cysteine (11). The GGTase-I enzyme uses a larger isoprenoid binding pocket where a threonine is substituted for the tryptophan found in FTase to open the pocket. This larger area allows for the binding of a larger 20 carbon geranylgeranyl pyrophosphate group that is then added to the peptide substrate in a similar manner as the FTase enzyme (10). Commonly farnesylated proteins include those of the Ras superfamily and in the maturation of lamin A. GGTase-I is known mostly for the role of prenylation proteins in the Rho superfamily.

It has been shown that in cases where FTase is inhibited, many CaaX substrates can be alternatively geranylgeranylated by GGTase, that also results in a fully functional mature protein. Thus, while farnesyltransferase inhibitors (FTI's) were initially sought after for treatment of Ras related cancers (8,12), this mode of therapeutics was shown to be ineffective (14). Interestingly though, while Ras can still retain full function by being processed by GGTase-I, lamin A, shows much less efficacy for the GGTase-I protein (15,16). Thus, FTI's have shown to be promising therapeutics for progeria, a deficiency linked to defects in lamin A processing, that is highlighted more in **Section 1.3**.

One additional GGTase, GGTase-II or RabGGTase, is the third type of mammalian geranylgeranyltransferase. It differs from the other prenyltransferases in that it recognizes the -CC, -CXC, -CCXX, or -CXXX motifs. These motifs often result in two subsequent geranylgeranylation sites (8). The main substrates of GGTase-II are the family of Rab proteins (7-9). These mature Rab proteins are mostly found associated with intracellular vesicles or the plasma membrane. There they are known to act as regulators of membrane trafficking pathways (17).

1.1.2 Proteases

1.1.2.1 *Rce1*

The second step of the CaaX posttranslational modifications pathway is the proteolysis of the -aaX residues from the prenylated cysteine. The major proteases in CaaX processing are the Rce1 and ZMPSTE24/Ste24 proteases. These different protease families contain both overlapping and unique substrate specificities (4). Ras appears to depend more heavily on Rce1 function, whereas **a**-factor, while cleavable by Rce1 as well, is the only known substrate of yeast Ste24 (8,18). Rce1 is a unique protease in that it utilizes catalytic glutamate and histidine residues to cleave the substrate (18,19). It can also cleave both farnesylated and geranylgeranylated peptides (18,19).

Like other CaaX modification enzymes, Rce1 proteolysis is critical for the proper localization of Ras (8). In Ras-driven cancers a mutation occurs that results in a constitutively active protein that, upon farnesylation, appears to permanently associate with the plasma membrane (20-22). While Rce1 was an initial target for Ras-related cancers, specifically targeting Rce1 with small molecule inhibitors has proved difficult (18). The only structure to date of Rce1 comes from the archaea organism *Methanococcus maripaludis* (19). Using this structure and the topological analysis of the protein from *S. cerevisiae,* it is believed that this protein functions as a glutamyl protease as two specific glutamate and two histidine residues appear to be highly conserved and are critical for proper function of Rce1 (18). The *M. maripaludis* Rce1 shares only 15% identity with human Rce1, so it is difficult to draw strong conclusions about the human homolog from the crystal structure. Hydropathy and sequence alignments across many eukaryotic species, including *Homo sapiens* and *Saccharomyces cerevisiae* show a possible seven or eight transmembrane helix structure, depending on the species (18,23).

1.1.2.2 *ZMPSTE24/Ste24*

Besides Rce1, there is an additional CaaX protease in eukaryotic cells called ZMPSTE24. This protease was first discovered in yeast and termed Sterile-24 (Ste24), as it aids in the maturation of the yeast mating pheromone, **a**-factor (24,25). The mammalian homolog is called Zinc Metalloprotease Ste24 (ZMPSTE24) and is most widely known for the role it plays in the

maturation of lamin A (26). While the two substrates are distinct to their host organisms, ZMPSTE24 and Ste24 have been found to be functional homologs (25,27).

Residing in both the endoplasmic reticulum and inner nuclear membranes, ZMPSTE24 and Ste24 both have a highly unique 3D structure (28,29). Both contain eight transmembrane helices surrounding a large chamber. The chamber can accommodate a 15 kDa protein, or about 450 water molecules. Several helices cap both ends of the chamber, and the HExxH zinc catalytic motif lies inside the chamber near the cytoplasmic side. Because of the novel and unique structure, the mechanisms of substrate entry and exits are still unclear, but it is believed there are portal regions within the sides of the membrane that may allow the substrate to enter and exit (28-30).

ZMPSTE24 and Ste24 also perform an additional cleavage upstream of the CaaX site on their respective substrates (**Figure 1.1B**) (1,5,8,31,32). The mechanism of substrate selection remains unclear and is the focus of the research chapters herein. Lamin A is cleaved 14 amino acids upstream of the farnesylated cysteine at an RSY/LLG sequence, and this cleavage completely removes the newly modified C-terminus (1,5,26,29,30,33-40). In yeast, the N-terminal cleavage site is 12 amino acids upstream at the sequence KDN/YII. The mature product is on the C-terminal side of the cleavage site, that includes the CaaX modified tail (24). It has also been shown that *S. cerevisiae* Ste24 can cleave mammalian prelamin A, and mammalian ZMPSTE24 can cleave the yeast mating pheromone **a**-factor (25,26). However, the lack of substrate sequence motifs makes it quite difficult to identify how these proteases recognize the correct proteolytic sites. The substrate binding site of Ste24 is the focus of the consecutive chapters.

Figure 1.1: Lamin A and **a**-factor both undergo similar post-translational processing, including CaaX processing. A) Process of typical CaaX processing. B) Similarities and differences between the production of lamin A and **a**-factor. Figure modeled after *Michaelis, S and Barrowman, J. (2012)* with permission from Microbiology and molecular biology reviews: MMBR, license 1142671-1.

1.1.3   Methyltransferases

The third and final step in the CaaX processing pathway is the carboxylmethylation of the isoprenylated cysteine. This reaction adds a methyl group to the exposed carboxylate on the cysteine, following the -aaXing step, creating a neutral methyl ester (1,39,41). The replacement of the negative charge of the carboxylate with the hydrophobic methyl group is believed to aid in membrane association of the mature CaaX proteins, like Ras and **a**-factor (2,42).

Methylation may add some stability to CaaX proteins as well. For example, RhoA appears to have a decreased half-life in macrophage cells, and the **a**-factor peptide degrades quickly in cells without Sterile-14 (Ste14), or the mammalian homolog human isoprenylcysteine carboxyl methyltransferase (hICMT) (43,44). ICMT enzymes bind both the methyl acceptor isoprenylated, cysteine substrate, as well as the methyl donor S-adenosyl methionine (SAM) (41). The proposed mechanism for the methylation reaction is that first the methyl donor SAM binds, followed by the substrate (41). Methylation then occurs and the methylated peptide leaves first, followed by the SAH product (S-adenosyl homocysteine) (41).

Similar to other CaaX processing enzymes, the first homologs were discovered in yeast before their discovery in mammalian cells. The *S. cerevisiae* CaaX methyltransferase is termed Ste14 ("sterile-14") as defects in Ste14 function causes a decreased formation of the mature form of the mating pheromone **a**-factor (3). It was found that the inability to methylate the **a**-factor peptide caused an inability for the export of **a**-factor from the cell through the Ste6 transporter (2). Also the methyl group is important for interaction with cell receptors on α-cells for sexual reproduction in yeast cells (2). Again, this could be due to the negatively charged carboxylate, formed following the -aaXing step, having a low affinity for the hydrophobic plasma membrane (45). Despite this, Ste14 is not considered an essential protein in yeast, as deletion of the *STE14* gene does not appear to cause significant changes in cellular metabolism or detectable defects in the cells other than the lack of mating ability (3). In mammals, ICMT is the only identified CaaX protein methyltransferase. Known substrates are the previously mentioned CaaX proteins (lamin A, Ras, Rho, etc.) and it can methylate both farnesylated and geranylgeranylated proteins.

For quite some time, the only structure of the ICMT family of proteins that was discovered came from a prokaryotic protein (*Ma*-ICMT) (46). This structure only aligns with the C-terminal half of eukaryotic Ste14 but did show a binding pocket for SAM within five transmembrane helices (46). The catalytic domain appeared to be in the C-terminal domain that is connected via a tunnel region to the isoprenylated substrate (46). This substrate was proposed to enter from the N-terminal side of the tunnel (46). The *Ma*-ICMT does share high sequence conservation with the C-terminal regions of eukaryotic ICMT proteins but the N-terminal varies greatly making mechanistic insights of the eukaryotic proteins from the prokaryotic structure challenging (47). In 2018, the eukaryotic beetle *Tribolium castaneum* ICMT structure was resolved to 3.4 Å, revealing eight transmembrane helices, as compared to the five in the *Ma*-ICMT structure (47).

Originally, FTI's were thought to be a great target for treating Ras related cancers. Once in clinical trials, however, it was quickly found that most CaaX proteins, especially Ras, are almost equally geranylgeranylated if FTase is inhibited (14). Because ICMT appears to accept both the farnesyl and geranylgeranyl groups equally (48), Ras was still methylated, and thereby still propagated the cancerous phenotype (42,49,50). An alternative approach has been to try to develop inhibitors for ICMT rather than FTase (8,42,51-54). In rodents and cell lines inhibition of ICMT appeared to strongly decrease cancerogenic phenotypes (20,55-57). Because it was shown that unmethylated Ras does not associate well with the plasma membrane, ICMT has been a novel target proposed for the development of inhibitor drugs to help treat Ras related cancers (42,45).

1.2   Nuclear envelope and lamin proteins

1.2.1   Structure of nuclear envelope

Eukaryotic cells are most distinguishable from prokaryotic cells by the presence of organelles, notably a nucleus. The nucleus, the cellular compartment responsible for containing cellular DNA, is surrounded by a double membrane, termed the nuclear envelope (NE) (**Figure 1.2**). The inner and outer nuclear membranes (INM and ONM, respectively) are continuous with each other at sites called nuclear pore complexes (NPC) (58). The ONM is also continuous with the endoplasmic reticulum (ER), another organelle found in eukaryotic cells responsible for lipid synthesis, protein processing and cellular calcium regulation (59). While the membranes are continuous, the nuclear membrane contains a diverse set of proteins that are not found in the ER membrane (58).

One protein group found in the nuclear membranes is the nucleoporins. These proteins make up the very large NPCs (40-120 MDa) to form large channels across the nuclear membrane (60,61). These pores allow for the passive diffusion of many water-soluble molecules, including proteins, RNA, and ions both into and out of the nucleus (60-63). Some nucleoporins also contain recognition signals for binding of nuclear import receptors, allowing for the nuclear import of larger proteins, like ribosomal or RNA polymerase subunits (63).

The lamin associating polypeptides (LAP proteins) are another group of nuclear proteins that are found only within the INM (64). LAP proteins interact with the lamina and genetic material to aid in proper DNA replication, expression levels of various genes, and in the organization of the

chromatin. The LAP proteins include the lamina-associated polypeptide (LAP2), emerin and the lamin B receptor for interaction with the various lamin proteins (64). LAP proteins associate strongly with heterochromatin and the lamin proteins, playing a strong roll in epigenetic regulation (64). LAP proteins also play a role in regulation of the cell cycle and cellular proliferation (65,66).

A third group of proteins found exclusively in the ONM are those that contain a KASH domain (Klarsicht, ANC-1, Syne Homology). The ONM exclusive proteins are used to properly position the nucleus during various stages of the cell cycle (58). These proteins are anchored to the membrane by a C-terminal hydrophobic region (67). These, along with the SUN proteins, specific LAP proteins in the INM, help transverse mechanical signals from the cytoskeleton to the nucleoskeleton (67).

Lastly, the fourth group of nuclear membrane proteins are the nuclear intermediate filaments. Intermediate filaments lie on the nucleoplasmic side of the membrane and associate greatly with the INM to form the nuclear lamina (39,68-71). These intermediate filaments are highly involved in many regulatory processes, including the cytoskeletal arrangement of the nucleus, mechanosensing and signaling, and gene regulation (66,69,71-76). These proteins consist of a family of lamin proteins that are separated into four main types: lamin A, lamin B1, lamin B2, and lamin C (71). Details of this group will be covered in greater detail in **Section 1.2.2**.
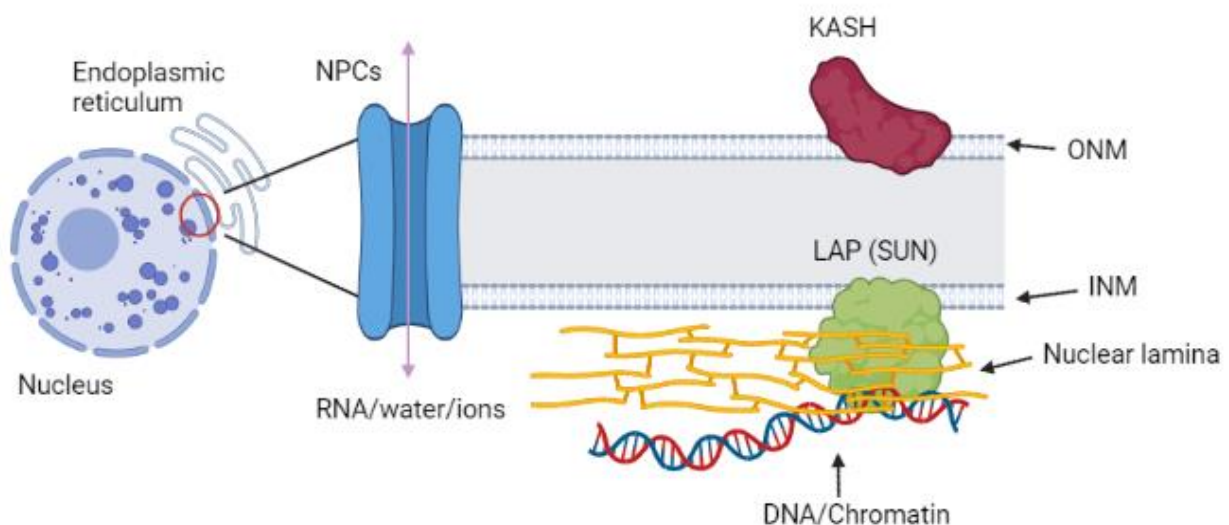
Figure 1.2: Nuclear membrane organization, figure created in biorender.com

1.2.2 Lamin proteins

1.2.2.1 General lamin protein family qualities

Lamin proteins are found exclusively in the animal kingdom where they are highly conserved. Other eukaryotes, including the *S. cerevisiae*, do not express them (71). There are four well known lamin proteins: A, B1, B2, and C (71). These four isoforms are expressed at various levels in different animal species (71). All the lamin proteins share similar structures and sequences in the N-terminal end, that contains a N-terminal head followed by a central rod domain (71). The tail end, consisting of the last 90 amino acids, differ greatly among lamin types (**Figure 1.3A**) (71). Lamin proteins favor the assembly of head-to-tail proto-filament oligomers and give rise to 10-13 nm long filaments (74). Lamin proteins are believed to be the evolutionary ancestor of all intermediate filaments and are the only intermediate filaments in the nucleus (71,74).

The lamin proteins function in several biological regulation pathways. One responsibility of the lamin proteins is that they play critical roles in transcriptional regulation as they associate with lamin-associated domains (LADs) in the DNA and interact with heterochromatin dense regions of the DNA (73,77-80). This role is believed to aid heavily in the repression of many gene sequences (68,73,78,79,81). It has been estimated that 30-40% of the mammalian genome consists of LADs and that these include a high frequency of protein-encoding regions (81). It has been shown in multiple studies that activation of these LAD regions by contact with the nuclear lamina, as well as higher rates of histone methylation, likely results in transcriptional repression of these protein-encoding regions (73,78,79,81). Additionally, as key components to the nuclear cytoskeleton, lamin proteins play a large role in proper reaction and homeostasis in responses to changes in mechanical forces to the cell (82).

Lamin proteins also aid in the movement, disintegration, and reformation of the nuclear membrane during the cell cycle (70,83,84). While the lamin proteins are most commonly localized to the nuclear membrane, studies have shown that both A and C type lamins can also be dispersed in the nucleus (77). These nucleoplasm localized lamins are also believed to aid strongly in epigenetic regulation and mechanical signaling, making the lamins highly important regulatory proteins in the nucleoskeleton (64,77,81,85,86).

1.2.2.2   B-type lamins

The B-type family of lamin, or lamin B, proteins are expressed by *LMNB1* or *LMNB2* at all stages of the cell and organism life cycle (71,74,77). B-type lamins undergo the canonical farnesylation, proteolysis and methylation seen by normal CaaX proteins. The addition of the farnesyl group allows them to associate well with the inner nuclear membrane, where they remain during the cell cycle (77). It has been suggested that lamin B proteins are important for maintaining the integrity of the nuclear envelope, positioning nuclear pore complexes, and may also play a role in DNA replication and transcription (74,87). It has also been shown that lamin B proteins may play a critical role in the proper organization of mitotic spindles during mitosis (74,84). In somatic cells, B-type lamins are critical for cell survival, as their depletion causes the cells to undergo apoptosis, most likely due to mitotic spindles becoming defective and the inability to correctly form spindle poles and improper association with the chromatin (74,87). Similar to lamin A, the lamin B preproprotein undergoes CaaX modifications. However, unlike lamin A, deficiencies in the proteolysis step of CaaX processing appears to be exclusively due to inefficiencies of Rce1, and not dependent on ZMPSTE24 (88).

Several diseases have been proposed to be caused by the overexpression of B-type lamins. The overexpression of lamin B is due to a duplication event in the chromosomes and this can cause many types of cancers, including hepatocellular carcinomas, breast cancer and colorectal cancer (74,89-92). The overexpression of lamin B1 has also been hypothesized to disrupt the production of myelin and proteolipid proteins in oligodendrocytes, causing myelin breakdown in the central nervous system (74,93,94).

1.2.2.3   A/C-type lamins

Expressed in developmentally regulated temporal patterns, the lamin A and C proteins are both encoded by the *LMNA* gene (71,77). They are expressed at low levels in embryonic cells, but become upregulated once cells become differentiated (77). Once transcribed, the mRNA can undergo alternative splicing to form different lamin isoforms, lamin A or lamin C (**Figure 1.3B**). Lamin A and lamin C are nearly identical within the first 566 amino acids, but lamin A contains a CaaX motif and lamin C does not. In fact, lamin C does not contain the last 98 amino acids that are found in the C-terminus of lamin A (68,71).

Both isoforms expressed by the *LMNA* gene appear to play vital roles in many signaling events, including the MAPK, Wnt-ß-catenin, TGF-ß and Notch signaling pathways (68). Mutations in the *LMNA* gene often cause a group of diseases called laminopathies or progeroid diseases, and can display a range of phenotypic effects such as abnormalities in skeletal, cardiac, and adipose tissues (68). At the cellular level, these *LMNA* mutations are usually accompanied by misshapen nuclei resulting in a difficulty responding to changes in mechanical stress, especially in cells surrounding branching points of arteries or skeletal muscles (68). As will be discussed in **Section 1.3**, many of these phenotypes arise due to incomplete modification of the lamin A protein (prelamin A), notably by ZMPSTE24

The A-type lamins, unlike the B-type lamins, can both associate with the nuclear membrane as well as remain soluble within the nucleoplasmic pool (66,77). It is believed that due to the lack of the farnesyl group on the mature protein, the A-type lamins can disassociate from the membrane as needed (66,77). The localization of the Lamin A/C proteins appears to be highly correlated with the binding of the Lamina-associated polypeptide 2α (LAP2α) that appears to help the lamins remain soluble in the nucleoplasm (77).
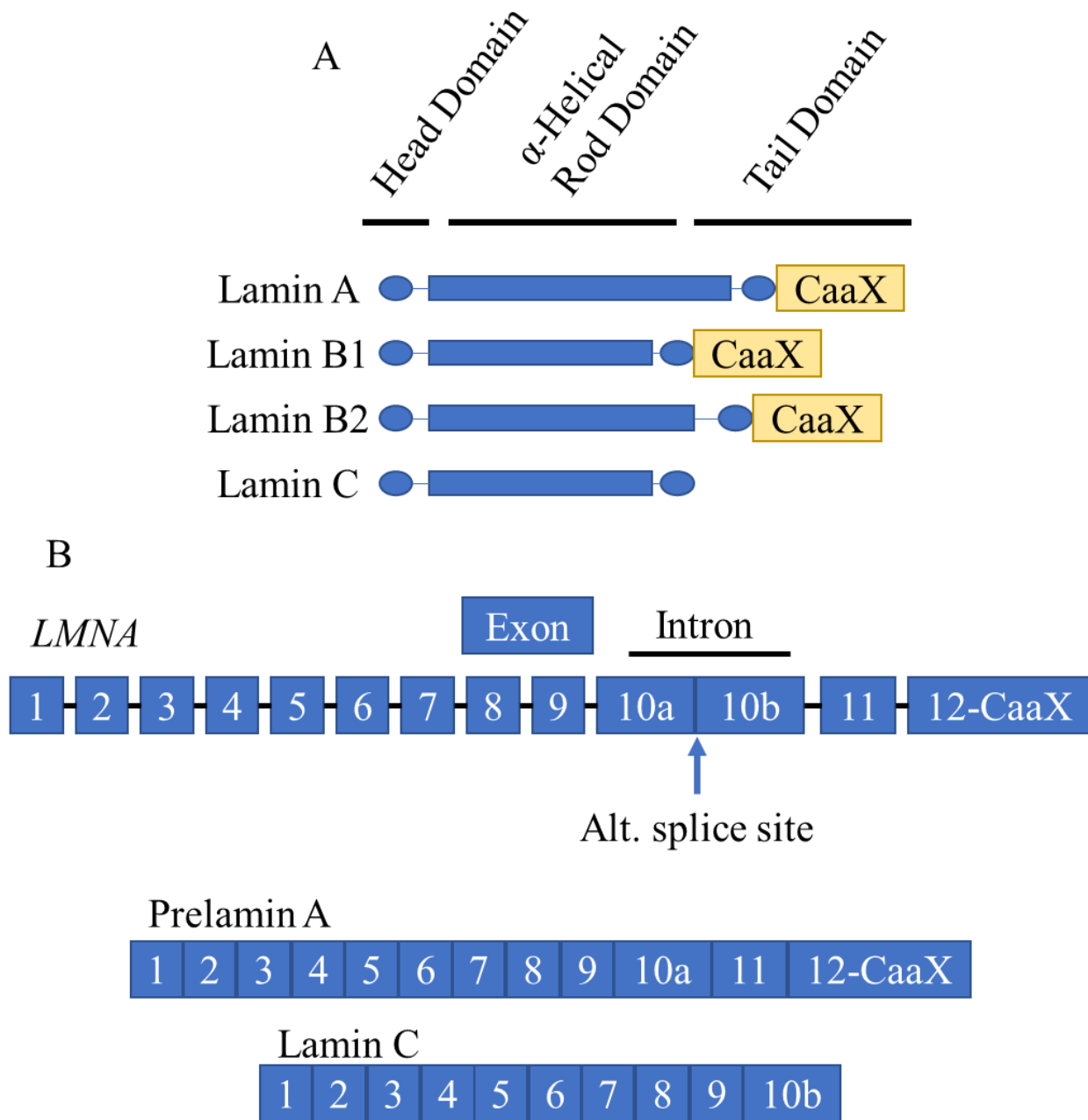
Figure 1.3: Lamin protein structure and alternative splicing. A) Protein structure domains and lengths of each of the lamin proteins (10). B) Alternative splicing variations of the *LMNA* gene result in either the prelamin A or lamin C protein (32).

1.3 Progeroid diseases

Progeroid diseases, or laminopathies, are an umbrella term for a class of syndromes that range greatly in severity. Some, like restrictive dermopathy, result in death either in utero or shortly after birth, while others result in milder metabolic disorders (95). These diseases occur due to mutations either in the *LMNA* gene or the *ZMPSTE24* gene that result in a decreased rate of prelamin A processing. Symptoms usually include some combination of striated muscle defects, skin or bone abnormalities, and changes in fat disposition (1,16,96).

1.3.1 Mandibulolacrual Dysplasia (MAD)

Mandibulolacural dysplasia is an autosomal recessive progeroid disorder caused by heterozygous mutations in both copies of either the *LMNA* or *ZMPSTE24* mutations (96-98). The disease typically onsets around 4 years of age (97). As of 2008, there had been only 27 identified patients with MAD, 23 contained a mutation in *LMNA* and 4 were due to mutations in *ZMPSTE24* (97). Mutations found in the ZMPSTE24 proteins of patients with MAD have been studied *in vitro* and have been found to retain partial activity compared to wildtype (27). This partial activity allows for a slower turnover of the prelamin A preproprotein, causing a mild to moderate buildup of the farnesylated and carboxylmethylated form of prelamin A, which is known to be toxic to cells (99). Mutations can also be found in the *LMNA* gene which result in mutated prelamin A and lamin A/C proteins, however the full CaaX processing and upstream ZMPSTE24 cleavage site are intact (99), unlike in HGPS that is discussed in further detail in **Section 1.3.2**.

Symptoms of MAD present in several different tissues in the body. First, there is usually a large number of skeletal abnormalities including osteolysis, or degradation of the bone, clavicular and mandibular hypoplasia, and craniofacial abnormalities (96-98). Often, the patients are described as appearing "bird-like" because they have prominent eyes and a beaked nose (97,98). The skin usually appears with skin atrophy, alopecia, and nail dysplasia, as well as hyperpigmentosa (96-98). Metabolic disorders are very common, and include glucose intolerance, diabetes mellitus, resistance to insulin and hypertriglyceridemia (96,97). Many patients present one of two types of lipodystrophy. Type A, typically presented with *LMNA* mutations, results in partial fat loss from the extremities, with very prominent regions in the face and neck. Type B,

usually more common in the presence of *ZMPSTE24* mutations, is a more generalized loss of the subcutaneous fat (97,98).

1.3.2 Hutchinson-Gilford Progeria Syndrome (HGPS)

One of the most well-studied forms of progeroid diseases is Hutchinson-Gilford Progeria syndrome (HGPS or progeria). Worldwide, at any time, there are about 30 cases of HGPS, indicating how rare this disorder is. HGPS is usually diagnosed within the first few years of life, and symptoms include slowed growth, hypoplasia in the face, lipodystrophy osteolysis, and eventually death around 13 years of age due to coronary artery disease (100,101).

Unlike other progeroid diseases discussed here, HGPS is caused by a specific mutation in exon 11 (C1824T) of the *LMNA* gene. While this does not cause a change in amino acid (G608), it does create a critical splice site. This site causes the removal of 150 base pairs, or 50 amino acids, that includes the N-terminal cleavage site for prelamin A maturation. The CaaX motif remains, so the protein still undergoes CaaX processing. The farnesylated, proteolyzed and methylated product is unable to undergo the necessary second N-terminal cleavage to produce mature lamin A. It is believed that this intermediate builds up and causes the toxic effects observed (102,103). At a cellular level, the indicative sign of progeria is misshapen nuclei (15,37,104). This mutation is also known to cause changes in cellular responses to mechanosensitivitiy and changes in gene regulation (37,76,83,105).

Several clinical trials are ongoing for the treatment of HGPS. The most promising of these use various farnesyl transferase inhibitors (FTI's) (15,16,104,106-109). While FTI's were shown to be ineffective in the treatment of Ras related cancers due to the alternate prenylation pathway, geranylgeranylation, prelamin A appears to show little efficiency with geranylgeranyl transferase (8,15,16). If the first step of the CaaX processing is prohibited, the cleaved, methylated intermediate is not formed, and the toxic effects are ameliorated. A 2012 study found that the treatment of HGPS patients with a FTI, lonafarnib, showed several improvements during the two year study, including: increased rate of weight gain, decreases in heart and arterial defects, increases in skeletal rigidity and slightly increased lifespan (110). A second clinical study, published in 2016, further ameliorated HGPS symptoms by combination of lonafarnib with a statin drug pravastatin and osteoporosis drug zoledronic acid (106).

1.3.3 Restrictive dermopathy (RD)

A more severe syndrome on the progeroid disease scale is called restrictive dermopathy (RD). RD is characterized by thin, tight skin, loss of fat and stiff joints. Death usually occurs in utero or shortly after birth (35,111-114). Unlike the dominant mutation known to cause HGPS, RD results from either recessive homozygous *ZMPSTE24* mutations, or heterozygous mutations in *ZMPSTE24*. These mutations result in a complete knockdown of ZMPSTE24 activity (35). Commonly, these mutations result from insertions or frameshifts in the genetic sequence of the *ZMPSTE24* gene (35). Due to the complete knockdown of ZMPSTE24 function, the concentration of the farnesylated and carboxylmethylated form of the prelamin A molecule increases in the cells and nearly none of the mature lamin A protein is produced. This creates very dramatic shifts in gene regulation and cellular signaling, and results in the death of the child either in utero or shortly after birth (111,115). Rare cases have also been shown to be caused by mutations in *LMNA*. Diagnosis during pregnancy by ultrasound is not very reliable as defects do not appear until well into the 2nd trimester. It is believed that this is due to the slow accumulation of prelamin A, resulting in the expression of the phenotype during the late-stages of pregnancy (35,111-115).

1.4 Recently discovered, novel roles for ZMPSTE24

1.4.1 HIV protease inhibitors

One effect of ZMPSTE24 inhibition has been implicated through the side effects of several medications used to treat patients with human immunodeficiency virus (HIV) (30,36,116,117). These medications, including lopinavir and tipranavir, target the aspartyl protease of HIV (30,118,119). However, several of these HIV aspartyl inhibitors cause various side effects including lipodystrophy, a common symptom in many progeroid disorders. In addition, it has been shown that several of these drugs can inhibit ZMPSTE24 *in vitro* to levels reflecting those found in patients with moderate progeroid diseases (**Figure 1.4**). Because ZMPSTE24 is a known zinc metalloprotease, and there is no apparent structural similarity with the HIV aspartyl protease, it is unclear how these inhibitors have this effect (30,36,116,117). Therefore, a better understanding of the binding site of ZMPSTE24 could provide better insight into how this inhibition occurs.
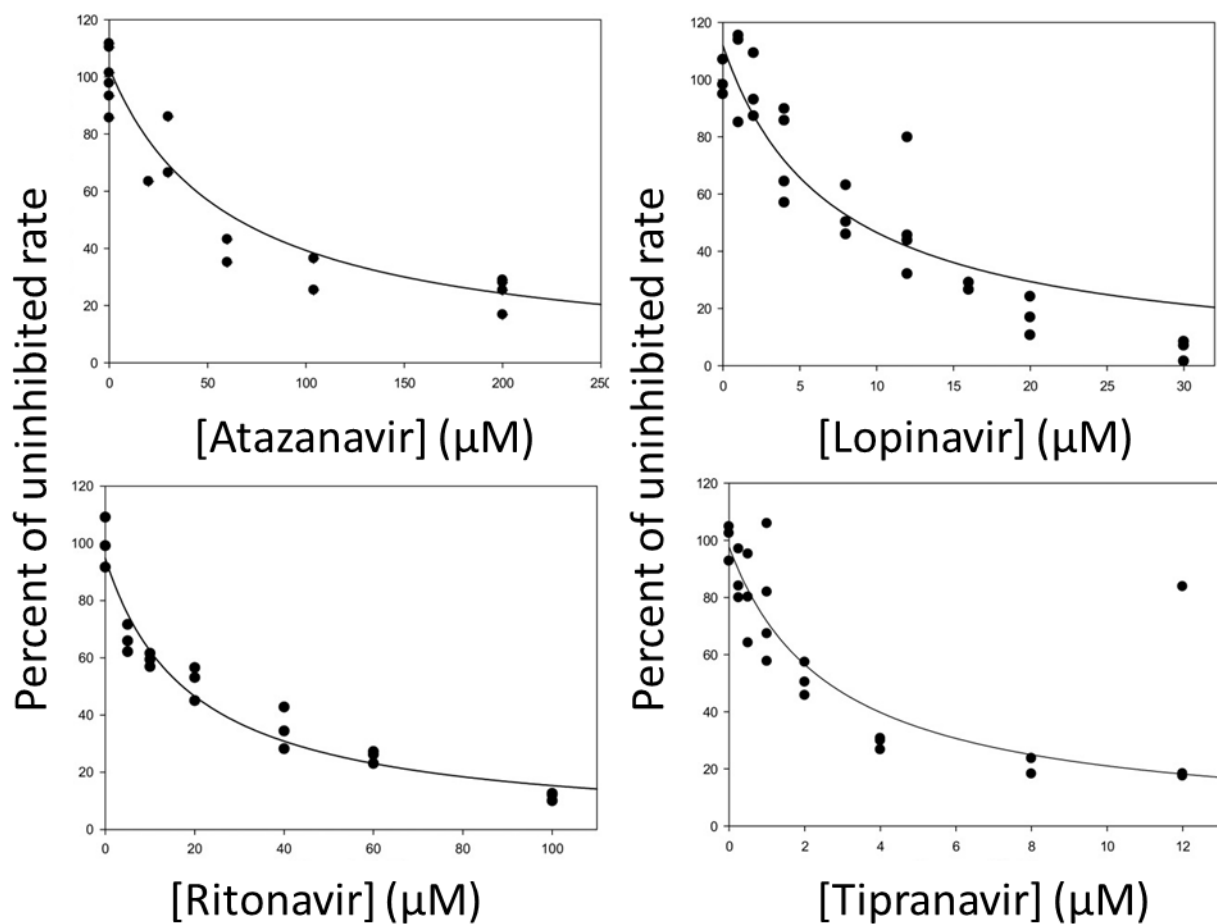
Inhibitor	K_I (μM)
Atazanavir	49 ± 13
Ritonavir	15 ± 2
Lopinavir	5.6 ± 1.1
Tipranovir	2.2 ± 0.6

Figure 1.4: ZMPSTE24 inhibition by various HIV aspartyl protease inhibitors. The activity curves are shown above and the corresponding K_I values are in the table below. Both figure and table obtained with permission from Clark, KM et al, Protein Science, 2017, 26:24.

1.4.2 Type II Diabetes

In 2018, it was discovered that native ZMPSTE24 may play several protective roles within the endoplasmic reticulum membrane (120,121). One of these roles protects against the aggregation of islet amyloid polypeptide (IAPP), a small peptide hormone. IAPP is co-secreted with insulin from ß-cells in the pancreas, and though its native function is unknown, when too much is expressed, IAPP forms aggregates (122-124). These aggregates are believed to be the cause of most ß-cell death in patients with Type II diabetes (120,123-125). It was recently shown that *Δste24* yeast cells showed a highly decreased cell viability when an IAPP hexamer was expressed, suggesting that IAPP aggregation caused a decrease in cell viability (120). When these cells were transformed with wild-type Ste24 cell viability was restored (120). Although this was an indirect experiment, these data suggest that Ste24 somehow plays a role in protection against the effects of IAPP aggregation. In addition, the genetic sequence of patients with Type II diabetes was obtained and many contained mutations within the *ZMPSTE24* gene (120). Several of the identified mutations were previously identified in progeroid disease patients (120). Taken together, these data suggest that ZMPSTE24 may play a role in preventing IAPP aggregation, although a direct mechanism has yet to be discovered.

One proposed explanation is that ZMPSTE24 may be preventing this aggregation by a mechanism of "de-clogging" the translocon (**Figure 1.5A**). A previous study developed a "clogging" protein that would prematurely fold as it was moving through the translocon. This caused the translocon to "clog", and causing a rapid increase in ER stress along with a decrease in the ability of the translocon to transport any other proteins (121). Cells that expressed higher levels of ZMPSTE24 showed a decrease in these effects. In addition, co-immunoprecipitation experiments showed that ZMPSTE24 does associate with some of the proteins within the translocon (121). These studies concluded that ZMPSTE24 may play a protective role against IAPP aggregation by helping to clear the translocon from the aggregated IAPP that would become stuck while moving through the translocon (120,121). While these studies looked at strong correlative effects, the lack of direct interaction of ZMPSTE24 with the clogging protein has yet to be confirmed. Both studies are relatively new findings, so many future experiments are needed to determine the exact role ZMPSTE24 performs with the translocon.

1.4.3 Viral infection

A study published in 2017 linked ZMPSTE24 to a possible protective function against certain viral infections, including influenza virus (126). Fu *et al* found that cells without ZMPSTE24 expression were significantly more susceptible to viral infection, as determined by fluorescence microscopy with a fluorescently tagged viral protein (**Figure 1.5B**). In addition, rates of infection by singix, vaccinia, zika and cowpox were increased in the absence of ZMPSTE24 protein. Viruses that were unaffected by ZMPSTE24 over-expression include murine leukemia virus and adenovirus, indicating a protective role in some, but not all, viral infections (126). Additional experiments showed that ZMPSTE24 may be interacting with interferon-induced transmembrane protein-3 (IFITM3), as determined by co-immunoprecipitation experiments (126). While again the direct role of ZMPSTE24 in the function of viral protection is not yet known, these experiments suggest ZMPSTE24 may be much more relevant to human health than previously described.

Additionally, in mid-2021 it was discovered that ZMPSTE24 may play a strong protective role against the novel SARS-CoV-2 virus (COVID-19) (127). Han et al showed that ZMPSTE24 could be co-purified with IFITMs in COVID-19 infections, and that knockout of ZMPSTE24 prevented the proper cleavage and function of ACE-2 driven immunity (127). This study confirmed that ZMPSTE24 interacts with IFITMs in ACE-2 driven infections, including COVID-19 (127). Also, because ZMPSTE24 activity/expression levels are significantly reduced in patients within the higher risk category, including the elderly, those with obesity or metabolic disorders, this would help explain the higher susceptibility to severe symptoms by these patient groups.
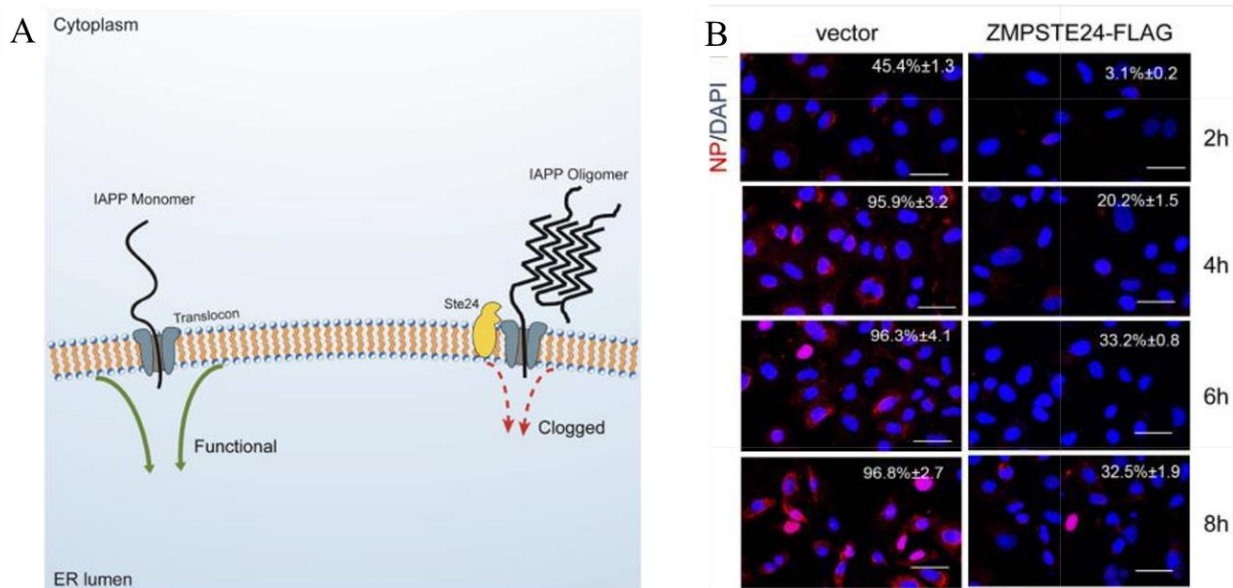
Figure 1.5: Alternative functions were newly discovered for ZMPSTE24/Ste24 proteolysis. A) Ste24 is proposed to "de-clog" the translocon from the IAPP oligomer. Figure obtained and modified with permission from Kayatekin, C. et al Cell. 2018 173: 62-73. B) Cells devoid of ZMPSTE24 show much faster infection rates by influenza virus than those transfected with ZMPSTE24. The cells were stained with a DAPI dye, as well as an antibody targeting the viral nuclear protein (NP) that produces a red fluorescence signal. Figure under Creative Commons License (https://www.ncbinlm.nih.gov/pmc/articles/PMC5379977/) from Fu, BJ Exp Med (2017) 214 (4): 919-929.

1.5   The yeast homolog, Ste24

1.5.1   Maturation of a-factor

The yeast homolog of ZMPSTE24, Sterile-24 (Ste24) plays a crucial role in the maturation of the yeast mating pheromone **a**-factor (25). Yeast can undergo asexual reproduction where they divide similarly to prokaryotes and do not require any interaction with surrounding yeast cells. Alternatively, when they express the **a**-factor and α-factor mating pheromones in nearby yeast cells, yeast can undergo sexual reproduction, allowing for diploid cell formation. If the yeast cell is then unable to express **a**-factor it is unable to undergo sexual reproduction, so is termed "sterile", hence the name "sterile 24" (25).

The **a**-factor peptide is translated as a 36 amino acid, prepropeptide. In a normal cell, the preproprotein **a**-factor is expressed and undergoes CaaX modifications. It is first farnesylated by RAM1/RAM2 at the cysteine in the CaaX box (CVIA) (9,25). Then, the C-terminal -VIA (aaX) is removed by either Rce1 or Ste24 resulting in a 33 amino acid peptide (4). Finally, the resulting

protein is methylated by Ste14 (3,25). After these CaaX modifications, Ste24 performs an additional cleavage to remove the N-terminal seven residues. Following the second Ste24 cleavage, the product is cleaved again by Axl1 to leave the mature, fully processed 12-mer with the farnesylated and carboxylmethylated cysteine at the C-terminus. The mature peptide is then secreted by the cell through Ste6 and binds to the Ste3 receptor on the outside of α-factor producing cells, inducing sexual reproduction (24,25).

1.5.2   Ste24 and ZMPSTE24 are functional homologs

Since the discovery of Ste24 in yeast, homologs have been identified across all eukaryotic species. The yeast Ste24 and mammalian homolog (ZMPSTE24) are both dual processing enzymes that reside uniquely in both the endoplasmic reticulum and inner nuclear membranes (28,29). The yeast homolog is often used as a model system for ZMPSTE24 as the mammalian homolog is more difficult to clone, express and purify as a functional protein. The only verified substrate of ZMPSTE24 is the mammalian protein prelamin A, while for Ste24 the only known substrate is the yeast mating pheromone **a**-factor. While the two cleavage sites in each respective substrate are unique, both proteins have been shown to be able to be functional substitutes for each other (25,27). This means that the mammalian ZMPSTE24 can be used to perform both cleavages of **a**-factor, and yeast Ste24 can be used to perform both cleavages of prelamin A (25,26). Due to the lack of substrate sequence specificity, how these proteins properly select each binding site is unclear.

The crystal structures of both proteins have been solved (**Figure 1.5**). ZMPSTE24 was solved first by Quigley *et al* with 3.4 Å resolution (29). The structure of Ste24 came quickly after by Pryor *et al* who used Ste24 as a model to produce a structure with a 3.1 Å resolution (28). Both structures revealed a novel protein structure that does not match any of the known soluble zinc proteases. These proteins have seven transmembrane helices, all surrounding a very large chamber of mixed hydrophobicity. The chamber is around 14,000 Å^3 and is large enough to accommodate a 10-kD protein or 450 water molecules. The chamber is capped on both ends by various helices, loops and a mixed α-helical/ß-strand domain. Gaps, or portals, are present in the sides of the chamber and are believed to be sites where the substrates can enter or exit. The transmembrane helices (TMD) VI and VII contain the zinc metalloprotease HExxH motif (28,29).

Pryor, *et al* also aligned the active site residues of a water-soluble zinc metalloprotease, thermolysin, to the Ste24 active site (28). This allowed for the determination of residues that are

suspected to aid in proper coordination/orientation of the substrate. Based on the similarities between the active sites, known substrates of thermolysin were mapped onto the active site of Ste24 to determine a proposed region for substrate binding, but other than substrate coordinating residues, specific residues required for binding were not determined.

The resulting conclusion of these structures is that the novel protein shape unfortunately does not provide much insight into the actual mechanism of action. The lack of sequence specificity between substrates also does not provide much insight into how it is properly positioned within the large chamber to undergo two distinct proteolytic cleavages. One proposed schematic is that the substrate (**a**-factor or prelamin A) enters through a portal site in the side of the chamber, then exits through a portal to get methylated by Ste14 or ICMT, respectively, and then re-enter through either the same or a different portal site for the second N-terminal cleavage. Further studies will need to determine the roles of each portal site and how this affects the activity of the protein. In addition, a more detailed analysis of the binding site of Ste24 and ZMPSTE24 needs to be performed.
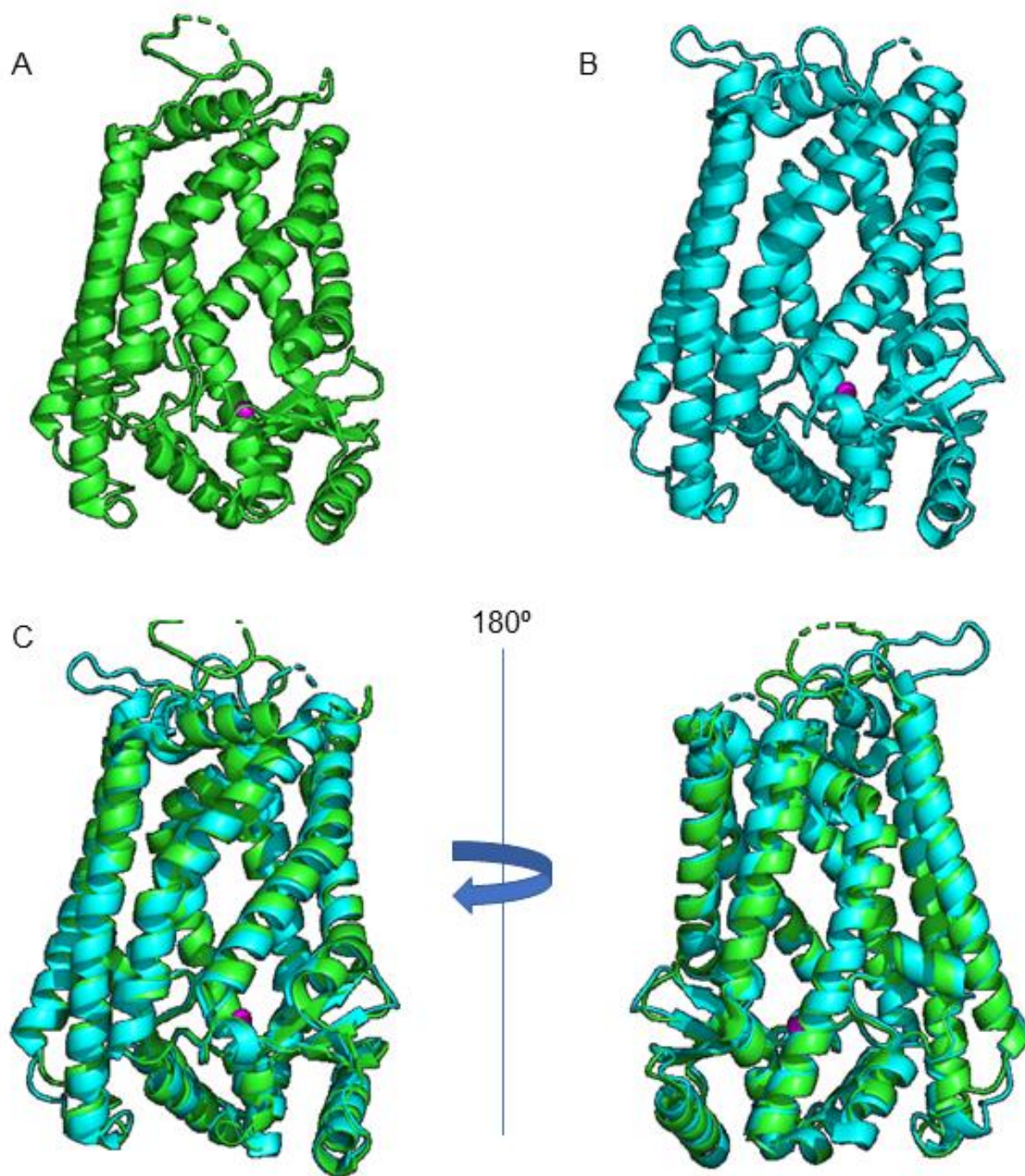
Figure 1.6: ZMPSTE24 and Ste24 both share a novel protein shape. A) The yeast *Saccharomyces mitake* Ste24 protein structure (PDB: 4IL3). B) The human ZMPSTE24 protein structure (PDB: 4AW6). C) Both 4IL3 and 4AW6 were aligned in PyMOL Molecular Graphics System (yeast in green and human in cyan) and the RMSD was calculated at 1.1 angstroms.

1.6   Statement of intent

As discussed above, although the structure of Ste24 has been solved, many questions remain unanswered regarding the mechanism of action. While the only known substrate is **a**-factor and this is cleaved at two distinct sites, how Ste24 recognizes and binds to the substrate is still unclear. Here, the yeast Ste24 protein is being utilized as a model system for ZMPSTE24 as it is easier to express and purify as a functional protein. Based on the solved crystal structure there are several regions of interest that I hypothesize play important roles for substrate binding.

The studies herein aim to identify specific residues or regions within Ste24 that are important for substrate binding during the C-terminal CaaX cleavage. A direct binding assay was developed that utilized microscale thermophoresis to measure direct changes in binding affinity of wild-type and mutant Ste24 proteins. A photoaffinity labeling-based proteomics assay was also developed to precisely locate the site of the prenyl group of a photoactivatable variant of **a**-factor within a hydrophobic patch lying just under a side portal of Ste24. Lastly, the conserved aspartate residues of Ste24 were assessed to determine their role in the proper function of Ste24.

CHAPTER 2.   DEVELOPMENT OF ASSAYS TO IDENTIFY AND INTERROGATE THE STE24 BINDING SITE

2.1   Introduction

Structures of both the human ZMPSTE24 and yeast Ste24 have been solved (28-30). The human ZMPSTE24 has also been co-crystallized in the presence of an unfarnesylated CaaX peptide, CSIM, the CaaX sequence of prelamin A (29). These structures have led to many hypotheses into how ZMPSTE24 and Ste24 select and bind their respective substrates, as discussed below (28,30). In addition, several studies have observed how changes in the amino acid sequences of the substrate effect selectivity (128,129).

Using these crystal structures and biochemical data, several regions within the Ste24 protein have been hypothesized to be important for substrate binding. These hypotheses are based on the co-crystallized ZMPSTE24-CaaX structure (29). Ste24 and ZMPSTE24 contain several portals opening into the chamber of the protein where the substrate is predicted to enter and exit, with the largest portal lying between TMD V and VI (28,29). This portal opens into the chamber near the canonical HExxH catalytic site within the protein. The peptide backbone of the substrate is suspected to bind along the amino acids surrounding the active site and those that start near this portal site (G255-F266). Additionally, the location of the farnesyl group during both the C-terminal and N-terminal cleavage are hypothesized to be in a hydrophobic patch near the C-terminus (L410 and L413) and within TMD 2 and 3 (Y79, F87 and L139), respectively.

One thing to note is, while the co-crystalized structure with CSIM provided strong information regarding the orientation of the CaaX sequence around the catalytic active site, the CaaX sequence did not contain the farnesyl group of the CaaX motif. Although several articles have proposed that the farnesyl group is not required for recognition, these studies were performed *in vitro*, and found the K_M value of farnesylated substrate to be 7-fold lower than that of non-prenylated substrates (16.4 µM vs 112.5 µM, respectively) (130). It has been proposed that the farnesyl group may be required more for substrate trafficking to the intracellular membranes for CaaX processing (130). Regardless of whether it is required for activity, we know the farnesyl group is present during both the CaaX and N-terminal cleavages *in vivo*, meaning there must be a hydrophobic region within the large chamber that can accommodate the farnesyl group during proteolysis.

All functionality of Ste24 amino acids thus far have been based on either structural or activity data. However, to determine exactly how residues play a role in substrate binding, an assay had to be designed that could look directly at binding affinity of Ste24. The goal of this project was to develop a binding assay for the C-terminal -aaXing cleavage of Ste24, and to use the assay to identify binding-oriented residues inside the protein. Alanine-scanning was used to mutate residues identified to lie in highly hydrophobic regions around the active site and in between side portals in the chamber. The aim was to clarify the importance of these residues in substrate binding by characterizing the enzymatic activity of the Ste24 variants and their ability to bind to substrate. The CaaX cleavage is the more well-studied activity of Ste24 and was the focus of this study.

There are several known methods for studying substrate binding. One notable technique is the use of photoaffinity labeling (PAL) that utilizes photolabeling substrates to observe drastic loss or gain in protein-substrate interactions (131,132). This technique utilizes the structure of a known binding partner to act as an affinity tag. The tag is constructed to have a photoactivatable group allowing for a covalent bond to form between the binding partner and the target protein, as well as a handle for enrichment. This technique has been shown to be quite powerful in identifying interactions amongst proteins and their substrates (131).

Additionally, the development of newer, small scale binding assays like microscale thermophoresis (MST) have greatly enhanced the ability to look at protein binding interactions among proteins with low expression rates and in the presence of difficult components like detergents and lipids (133,134). MST can be used either in addition or in place of photoaffinity labeling to provide insights into substrate binding, notably by the determination of specific K_D values. Herein, the first binding assay was used to determine specific K_D values for Ste24 was developed using MST. In addition, by utilizing alanine-screening of the regions of interest, residues could be identified by their role in either catalysis, binding, or both, during the C-terminal CaaX cleavage of **a-**factor by Ste24.

2.2   Methods

2.2.1   Plasmids and yeast strains

All plasmids were created from the starting pCH1283 plasmid, made by prior Hrycyna lab graduates. This plasmid contains a His_{10}-HA_3 tag on the N-terminus of the Ste24 gene (*2μ URA3*

P_{PGK}-His₁₀-HA₃-Ste24). Mutant plasmids were made by traditional cloning methods (135) or by the Q5-site directed mutagenesis kit ® (New England Biolabs) using the primers displayed in **Table 2.1** to make the corresponding plasmids shown in **Table 2.2**. Plasmids were transformed into the yeast strain SM3614, containing a double deletion for endogenous Rce1 and Ste24 (*MATa trp1 leu2 ura3 his4 can1 ste24Δ::LEU2 rce1Δ::TRP1*) (44). For the Ste14 protein used for activity assays, this was expressed in a pCH2733 (*2μ URA3 P_{PGK}-His₁₀-Myc₃-Ste14*) also transformed into the SM3614 strain.

For traditional cloning methods, mutant primers were designed to introduce single point mutations in the Ste24 gene (**Table 2.1**) in the pCH1283 plasmid (135). GoTaq® Green Master Mix (Promega) was used following manufacturer instructions, to perform a two-step PCR procedure to introduce the mutation of interest into the gene. The products of each PCR reaction were purified using the QIAquick PCR Purification Kit (Qiagen). After the final PCR step, both the purified PCR products and the pCH1283 vector were digested with EagI restriction enzyme and separated on a 1% agarose gel. The bands of interest were excised and purified using the QIAquick Gel Extraction Kit (Qiagen). The Rapid DNA Ligation Kit (Invitrogen) was then used to ligate the purified products. The ligated plasmids were then transformed into *E. coli* DH5α competent cells (Invitrogen) according to manufacturer instructions. Single colonies obtained from the transformation were used to grow miniprep cultures by selecting single colonies and growing them in 5 mL of Luria-Bertani (LB) broth containing 100 µg/mL ampicillin overnight at 37 ºC. The plasmids were harvested using QIAprep Spin Miniprep Kit (Qiagen), and the sequences were verified using bi-directional dye-terminator sequencing (BigDye® Terminator v3.1) by Purdue University Genomics Facility, or by GeneWiz sequencing center, using Sanger sequencing. The plasmids were then transformed into SM3614 cells using the Elble protocol (136) and grown at 30 ºC for two days on synthetic complete medium without uracil (SC-URA).

Alternatively, mutant plasmids were created using the Q5-site directed mutagenesis kit ® according to provided instructions and primers listed in **Table 2.1**. Briefly, end to end primers were designed and used to amplify linear versions of the mutated plasmid. Then the KLD enzyme mix (kinase, ligase, DPNase) provided by the kit was used. The kinase adds a phosphate group to the 5'-end of the DNA strand, allowing the ligase to form circular plasmids. The DPNase then chews up the methylated template DNA and leaves only the mutated plasmids that were then transformed into DH5α cells (Invitrogen) where the DNA was collected by the QIAprep Spin

Miniprep Kit (Qiagen). The sequences were verified using bi-directional dye-terminator sequencing (BigDye® Terminator v3.1) by Purdue University Genomics Facility, or by GeneWiz sequencing center, using Sanger sequencing. The plasmids were then transformed into SM3614 cells using the Elble protocol (136) and grown at 30 °C for two days on synthetic complete medium without uracil (SC-URA).

Table 2.1: List of primers used to create the Ste24 mutant library

Primer	Sequence (5' to 3')
SVEC (+)	CAGGGGGTGGTTTAGTTTAG
SVEC (-)	CAACTGTTGGGAAAGGCGATC
E390A (+)-Q5	AGAACTCATGCATATCAAGCTG
E390A (-)-Q5	GGA AAT TAA ACTCATCACG
H434A (+)	GCTATCATTATTCCGCTCCAACTCTAGCT
H434A (-)	CAGCTAGAGTTGGAGCGGAATAATGATA
R440A (+)-Q5	CAACTCTAGCTGAAGCATTGACC
R440A (-)-Q5	GATGGGAATAATGATAGCTAGAATACAGAG
L441A (+)-Q5	CTGAAAGAGCGACCGC
L441A (-)-Q5	CTAGAGTTGGATGGGAATAATG
L410A (+)	CAAAATCTATGT AGG GCT GCA ATT GAT CTA CAA ATC
L410A (-)	GAT TTG TAG ATC AAT TGC AGC CCT ACA TAG ATT TTG
L413A (+)	GTA GGG CTC TAA TTG ATG CAC AAA TCA AAA AC
L413A (-)	GTT TTT GAT TTG TGC ATC AAT TAG AGC CCT AC
Y79A (+)	GGT GAC GTC GCT AAC CTA GCC
Y79A (-)	GGC TAG GTT AGC GAC GTC ACC
F87A (+)	GCC CAA AAG CTA GTT GCC ATC AAA TAC G
F87A (-)	CGT ATT TGA TGG CAA CTA GCT TTT GGG C
L139A (+)	GTT GAT TTG CCA GCC TCT TAC TAT AGC
L139A (-)	GCT ATA GTA AGA GGC TGG CAA ATC AAC
G255A (+)	GAT TTT TGT CAT TGA CGC CTC AAA AAG ATC TTC TC
G255A (-)	GAG AAG ATC TTT TTG AGG CGT CAA TGA CAA AAA TC
S256A (+)	GAC GGC GCA AAA AGA TC
S256A (-)	GAT CTT TTT GCG CCG TC
K257A (+)	GAC GGC TCA GCA AGA TCT TCT C
K257A (-)	GAG AAG ATC TTG CTG AGC CGT C
R258A (+)	GAC GGC TCA AAA GCA TCT TC
R258A (-)	GAA GAT GCT TTT GAG CCG TC
S259A (+)	TCA AAA AGA GCT TCT CAT TCA AAC
S259A (-)	GTT TGA ATG AGA AGC TCT TTT TGA
S260A (+)	CAA AAA GAT CTG CTC ATT CAA ACG C
S260A (-)	GCG TTT GAA TGA GCA GAT CTT TTT
H261A (+)-Q5	AAG ATC TTC TGC TTC AAA CGC ATA TTT CAC
H261A (-)-Q5	TTT GAG CCG TCA ATG ACA
S262A (+)	GAT CTT CTC ATG CAA ACG CAT ATT TC
S262A (-)	GAA ATA TGC GTT TGC ATG AGA AGA TC
A264G (+)	CTC ATT CAA ACG GAT ATT TCA CAG G
A264G (-)	CCT GTG AAA TAT CCG TTT GAA TGA G
Y265A (+)	CAA ACG CAG CTT TCA CAG GTT TGC
Y265A (-)	GCA AAC CTG TGA AAG CTG CGT TTG
F266A (+)-Q5	AAA CGC ATA TGC CCA CAG G
F266A (-)-Q5	GAA TGA GAA GAT CTT TTT GAG AAG

Table 2.2: List of plasmids used in the creation of the mutant Ste24 library

Plasmid	Genotype
pCH1283	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24
pCH1314	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-E390A
pCH1338	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-H434A
pCH1376	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-R440A
pCH1322	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-L441A
pCH1342	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-L410A
pCH1343	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-L413A
pCH1340	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-Y79A
pCH1339	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-F87A
pCH1352	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-L139A
pCH1344	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-G255A
pCH1377	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-S256A
pCH1295	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-K257A
pCH1296	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-R258A
pCH1297	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-S259A
pCH1298	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-S260A
pCH1294	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-H261A
pCH1300	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-S262A
pCH1378	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-A264G
pCH1341	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-Y265A
pCH1379	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-F266A

2.2.2 Crude membrane preparation

Crude membrane preparation was started by first inoculating a small, SC-URA culture with the desired Ste24 (or Ste14) strain and this culture was grown overnight at 30 ºC. Then 15 mL was used to inoculate a larger, 1 L SC-URA flask that was grown to log phase (OD$_{600}$ = 3-5) then harvested at 4000 xg. Lysis of the cells was performed using a yeast sorbitol lysis buffer (0.3 M sorbitol, 0.1 M NaCl, 12 mM MgCl$_2$, 1% aprotinin, 3 mM AEBSF, 1 mM DTT, 10 mM Tris-HCl,

pH 7.5) at a ratio of 1 mL lysis buffer to 800 OD$_{600}$ cells. The lysis buffer was added to the cell pellet in a 50 mL conical tube that was then vortexed and left on ice for 15 minutes to allow the cells to swell. The suspension was then flash frozen in liquid nitrogen and thawed in a 29 °C water bath, twice. Then a french press was used to further lyse and break up the membrane fractions using 18,000 psi twice. The resulting suspension was first centrifuged twice at 500 xg for 10 minutes to remove cellular debris then centrifuged at 100,000 xg for 1 hour at 4 °C. After the final centrifugation step, the supernatant was removed, and the membrane pellet was resuspended in 10 mM Tris-HCl, pH 7.5. Protein concentration was determined using a coomassie protein assay using a known concentration of bovine serum albumin (BSA) as the standard curve (137).

2.2.3 Membrane protein purification

To purify, the crude membrane proteins were first solubilized in buffer A (0.3 M sorbitol, 0.1 M NaCl, 6 mM MgCl$_2$, 10 mM Tris, pH 7.5, 10% glycerol, 1% aprotinin, 2 mM AEBSF) supplemented with 20 mM imidazole and 1% n-Dodecyl-B-D-maltopyranoside (DDM; Anatrace) with rocking at 4 °C for 1 hour. The suspension was then centrifuged at 100,000 xg for 45 minutes to remove the insoluble fractions, and the supernatant was added to Talon® Metal Affinity Resin (Clontech). The resin mixture was rocked at 4 °C for 1 hour, then washed with five column volumes of buffer B (40 mM imidazole and 1% DDM in buffer A) twice, buffer C (40 mM Imidazole, 1% DDM and 0.5 M KCl in buffer A) once, and buffer D (40 mM imidazole, 0.1% DDM and 0.5 M KCl in buffer A) once. The protein was then eluted from the resin with buffer E (250 mM Imidazole, 0.1% DDM in buffer A). The elution was then concentrated using Amicon® Ultra Centrifugal Filter 30,000 MWCO (Millipore) at 4,000 xg for 20-30 minutes at 4 °C until desired volume was reached. Protein concentration was determined using the amido black protein assay using BSA as a standard curve (138).

2.2.4 Radioactive methyltransferase-coupled diffusion assay

The C-terminal, or -aaXing, cleavage activity of Ste24 wild-type and mutant proteins were determined using an *in-vitro* radioactive methyltransferase-coupled diffusion assay, as previously described (100,139). First, 0.24 µg Ste14, 0.12 µg Ste24 and farnesylated 15-mer **a**-factor (final concentration 50 µM, **Figure 2.1A**) were added to 40 µg *E. coli* Polar Lipid Extracts (Avanti).

Then a rapid dilution was performed with Tris-HCl, pH 7.5 to dilute the fractions 1:8 (final 100 mM Tris-HCl). To start the reaction, S-adenosyl-[^{14}C-methyl]-L-methionine (SAM) (Perkin Elmer) was added to a final concentration of 20 µM and the reaction mixtures were incubated at 30 ⁰C for 30 minutes. The reactions were stopped by addition of a 1 M NaOH/1% SDS mixture (final 450 µM NaOH, 0.5% SDS), then the samples were spread onto a pleated filter paper and placed in the neck of a scintillation vial filled with 10 mL of Biosafe II scintillation fluid (RPI). The vial was capped and left for 3 hours to allow time for the [^{14}C]-methanol, released by the addition of NaOH to stop the reaction, to diffuse into the scintillation fluid. The filter paper was then removed, and the radioactivity was quantified using a Packard TriCarb Scintillation counter. The specific activity was calculated from the counts per minute (CPM) and converted into pmol of -aaX residues cleaved per minute per mg of Ste24 protease. Reactions were set up as duplicates, and each set of duplicates was run in triplicate experiments.

2.2.5   SDS-PAGE and immunoblot analysis

For immunoblot analysis of purified Ste24, 0.05 µg pure protein (diluted in 2x SDS loading dye) was loaded onto a 4% stacking, 10% separating SDS-PAGE gel. The samples were stacked at 85 V, followed by separating at 165 V for 40 minutes. The proteins were then transferred to a 0.45 µm nitrocellulose membrane (Cytiva Amersham ™ Protran ™ NC Nitrocellulose) at 100 V for 90 minutes. The membranes were then blocked overnight in 20% milk in PBST (1x PBS buffer, 0.1% Tween-20) at 4 ⁰C, followed by 2 hours in primary antibody (mouse, α-HA at 1:15,000) in 5% milk in PBST, then 1 hour in secondary antibody (goat-α-mouse-HRP, 1:4,000) in 4% milk in PBST. The bands were then visualized on a GeneGnome XRQ (SynGene) using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific).

To determine purity of the Ste24 proteins, pure proteins were diluted in 2x SDS, then 1 µg total protein was loaded onto a 4% stacking, 10% separating SDS-PAGE gel. The running conditions were identical to the immunoblot SDS-PAGE described above. The gel was then stained overnight in Coomassie Brilliant Blue (0.3 M Coomassie Brilliant Blue, 10% acetic acid, 40% methanol) at room temperature, then destained with 10% acetic acid / 30% methanol to remove background.

2.2.6 Photolabeling assays

First, 0.5 µg of the purified Ste24 protein was added to 125 µg of *E. coli* Polar Lipid Extract (Avanti) and incubated on ice for 10 minutes. Then, a solution of buffer, DTT and photolabeling probe, C_{10}-para-**a**-factor (**Figure 2.1B**) was added to dilute the protein reconstitution by 1:6, resulting in a final concentration of 1 mM DTT, 50 µM probe and 100 mM Tris-HCl, pH 7.5. This was left to incubate on ice another 10 minutes in the dark. Then 65 µL was added to a 96-well plate sitting on ice under a UV lamp (365 nm), serving as the +UV, or photolabeled, sample. Then another 65 µL was added to a fresh tube, kept on ice but left in the dark for a -UV, unlabeled, sample. The samples were crosslinked under UV light (365 nm) for 30 minutes, and then 60 µL of both the -UV and +UV samples were added to 50 µL of a 1:1 slurry of NeutrAvidin® Agarose Resin (Thermo Scientific) in 800 µL RIPA/SDS/I buffer (25 mM tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 1% sodium dodecyl sulfate, 1% aprotinin, 2 mM AEBSF, 50 U/mL micrococcal nuclease, 1 mM $CaCl_2$, and 1 mM DTT). The protein reactions and resin were left rocking for two hours at 4 ⁰C. The resin was then pelleted at 10,000 xg for 2 minutes and washed three times with 800 µL fresh RIPA/SDS/I buffer. The beads were then resuspended in 50 µL 2x SDS loading buffer, heated at 65 ⁰C for 30 minutes, then 10 µL was run on an SDS-PAGE followed by an immunoblot, as described above.

Quantification was performed using ImageJ as described in published methods (http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf). Briefly, rectangular bands were created that fit all bands in an immunoblot and the pixel measurements were counted. The pixel density (X) was inverted (255-X) for each of the -UV and +UV samples. The inverted -UV samples were then subtracted from the +UV samples to express the net values of the photolabeling signal. The net photolabeling signal was reported as the percentage of band intensity as compared to the photolabeling signal of the WT Ste24 sample.

2.2.7 Photolabeling assay for competitive binding

The photolabeling assay was set up similarly to the purified photolabeling assays described in **Section 2.2.6** except with the addition of the 15-mer **a**-factor substrate. First, the wild-type Ste24 protein was incubated in *E. coli* polar lipid extract (Avanti) with the C_{10}-para probe, as described above, incubated for ten minutes, then rapidly diluted 1:6 in buffer (final concentrations: 1 mM

DTT, 100 mM Tris-HCl, 50 µM C_{10}-para, 0.5 µg WT, 5.4 µg lipid) and aliquoted into eight wells of a 96-well plate. Then **a**-factor dilutions were prepared by 1:1 serial dilution in DMSO starting from 7.5 mM through 0.12 mM and these were added to the wells containing the photolabeling samples (final concentrations 1.5 µM to 100 µM, with DMSO alone being added to the final to mimic a 0 µM solution). The samples were incubated in the dark on ice for 10 minutes to allow for equilibration.

Then a UV light was used to irradiate the samples for 10 minutes. Then 60 µL of each reaction was mixed with 60 µL 2x SDS loading buffer, heated at 65 °C, then 10 µL was loaded onto a 10% SDS-PAGE, transferred to a 0.45 µm nitrocellulose membrane. The membrane was blocked overnight in 20% milk in PBST at 4 °C followed by three hours at room temperature with NeutrAvidin-HRP® (ThermoScientific) in 5% BSA (Fisher) in PBST. The bands were then visualized on a GeneGnome XRQ (SynGene) using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific).

2.2.8 Microscale thermophoresis (MST) assay

Purified Ste24 proteins were serially diluted 1:1 from their stock concentrations (ranging from 20 µM to 60 µM) in concentrated buffer E. Then, a 5-FAM linked version of the 15-mer **a**-factor (**Figure 2.1C**) was added to a final concentration of 50 nM in each sample. These samples were loaded into standard capillaries for use in the Nanotemper Monolith NT.115 instrument at room temperature, using a 1 second, 20% initial fluorescence reading for normalization, followed by 15 seconds using the MST laser set at medium intensity, then 1 second after the MST laser was turned off. The K_D values were determined using the Monolith NT.115 software, as well as a verification on GraphPad Prism 9 using the Total Binding, nonlinear fit analysis. Data are represented as the average of duplicate reads for each sample.

Alternatively, the Cy5-labeled isostere form of **a**-factor (**Figure 2.7**) was also used to perform binding analysis utilizing the MST Nanotemper Monolith NT .115 instrument. Samples were prepared as generated by the Monolith NT .115 software. Briefly, 20 µL of the Ste24 stock proteins was used to perform 16, 1:1 serial dilutions of 10 µL each. Then 10 µL of the Cy5-isostere **a**-factor was added to a final concentration of 20 nM, totaling a reaction volume of 20 µL. Two separate MST runs were performed for each set of dilutions.

2.2.9 Tycho NT 1.6 for protein thermal stability

Purified wild-type and mutant Ste24 proteins were loaded from their stock tubes directly into Tycho NT 1.6 capillaries from Nanotemper and loaded into the Tycho NT 1.6 instrument. The fluorescence ratio of 350 nm/330 nm was measured across temperatures from 30 °C up to 95 ° C. The inflection point (T_m) was determined by the Tycho NT1.6 software. All proteins were analyzed twice.

2.2.10 Structural analysis of the Ste24 protein

Structural analysis was performed by utilizing The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. and analyzing the SmSte24 structure (PDB: 4IL3). The zinc ion was depicted as an orange sphere. Residues and modifications were selected as they became of interest. Hydrogen bonds were determined using the "Action→Find→polar contacts→to other atoms in object". Further analysis will be discussed in the legends of the figures.

A) 15-mer a-factor

Farnesyl group

H₂N-YIIKGVFWDPACVIA

15-mer peptide backbone

B) C₁₀-para-a-factor

PEG4-linker

C_{10}-para: Photoactivatable group

YIIKGVFWDPACVIA

15-mer peptide backbone

Biotin handle

C) 5-FAM a-factor

Farnesyl group

YIIKGVFWDPACVIA

15-mer peptide backbone

PEG4-linker

5-FAM: Fluorescent label
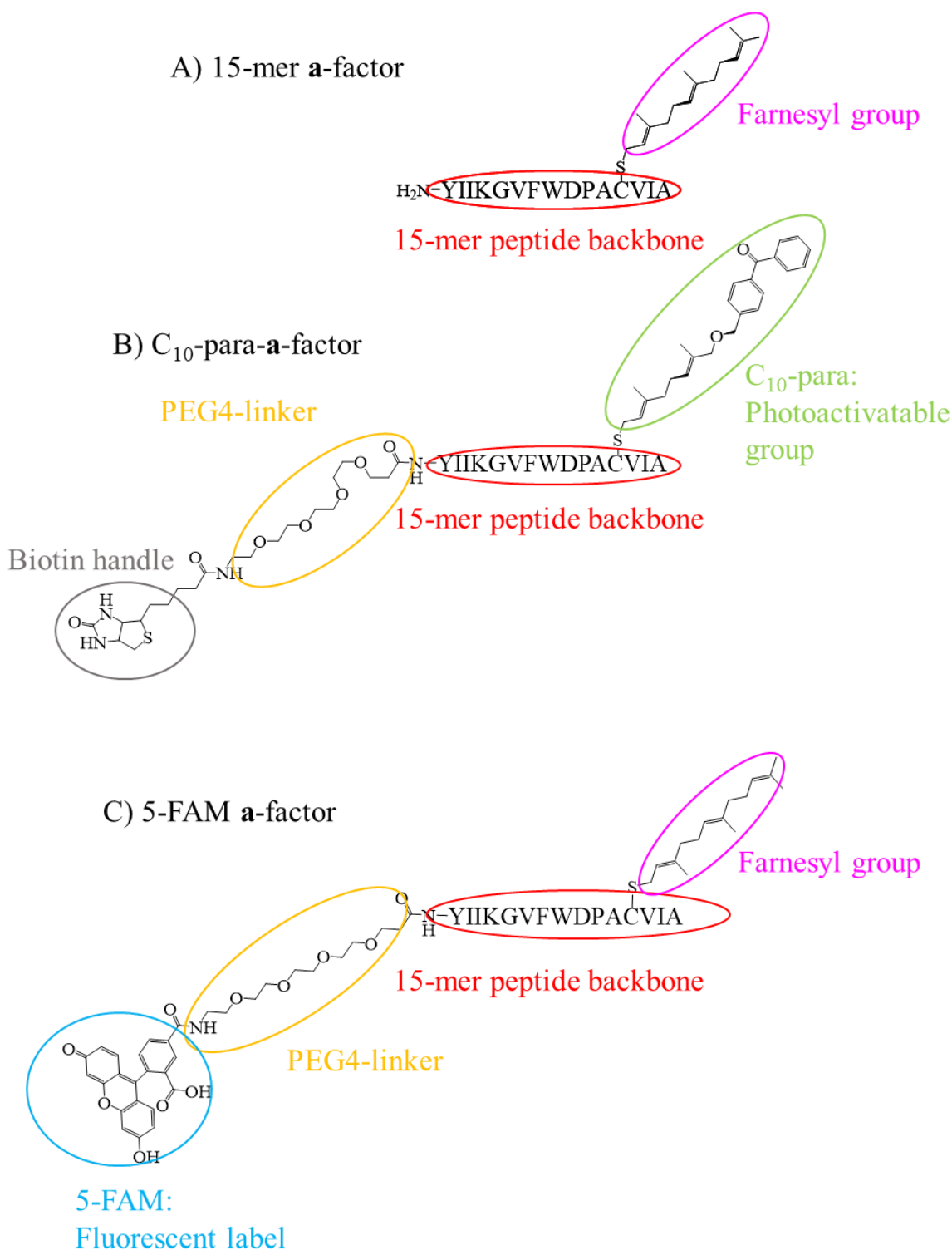
Figure 2.1: Several substrates/binding partners developed by collaborators at the University of Minnesota were used in this study. A) 15-mer **a**-factor substrate used for activity assay analysis. B) C_{10}-para-**a**-factor substrate developed by the Distefano Lab at the University of Minnesota for photolabeling assays. C) 5-FAM **a**-factor was developed by the Distefano Lab at the University of Minnesota for use in MST binding assays and contained an uncleavable isostere moiety.

2.3 Results and Discussion

2.3.1 Activity assay data revealed the importance of several residues required for proper catalysis of the C-terminal -aaX cleavage of a-factor by yeast Ste24

To begin, several residues in yeast Ste24 were selected to undergo alanine screening, or glycine in the case of the native residue being an alanine. These include: one of the zinc-binding residues in the active site (E390), residues mapped to the substrate binding site of another zinc metalloprotease, thermolysin, (H434, R440 and L441) (28,30), amino acids in two separate hydrophobic pockets (the C-terminal hydrophobic patch [L410 and L413] and the TMD 2 and 3 hydrophobic patch [Y79, F87 and L139]), and finally residues predicted to aid in binding of the **a**-factor peptide backbone due to their location between two possible entrance and exit sites, or portal sites, in the protein (G255-F266). Once created, membranes expressing these mutant Ste24 proteins or the wildtype protein were isolated from yeast and purified to perform the presented analyses. All residue functions and data obtained in this chapter is summarized in **Table 2.4**.

To study the activity of the proteins, a well-established methyltransferase-coupled proteolysis assay was utilized (100,139). This assay uses a truncated, 15 amino acid version of **a**-factor that consists of the farnesyl group on the CaaX cysteine, and the active 15 amino acid sequence (**Figure 2.1A**). The 15-mer **a**-factor peptide is incubated for 30 minutes with both Ste24 and Ste14, reconstituted in *E. coli* lipids, along with the Ste14 co-factor *S*-adenosyl-[^{14}C-methyl]-L-methionine (SAM). Ste24 first cleaves the three C-terminal -aaX residues and then Ste14, in excess, adds a radiolabeled ^{14}C-methyl group to the now exposed cysteine. After the 30-minute timepoint a mixture of NaOH and SDS is added to quench the reaction. The SDS denatures the proteins and then the hydroxyl group from NaOH replaces the ^{14}C-methoxy group added by the Ste14 and SAM. The sample was then spread onto a pleated filter paper and rested in the neck of a scintillation vial filled with scintillation fluid for three hours. During the three hours the ^{14}C-methanol released after reaction with the NaOH evaporates and diffuses into the scintillation vial, including into the scintillation fluid. The amount of ^{14}C-methanol is then counted using a scintillation counter and this is used to calculate the amount of substrate turnover of the starting 15-mer **a**-factor substrate.

The results of the methyltransferase-coupled diffusion assay of each mutant are shown in **Figure 2.2**. As expected, the zinc-binding E390A mutation showed a complete knock down of activity, around 0% retained activity compared to the purified wild-type Ste24. It has been

previously shown that Ste24 function is strongly dependent on the presence of the zinc ion (140). Therefore, by removing a zinc-coordinating residue, E390, the catalytic function is abrogated. The putative substrate coordinating residues, predicted based off the structure alignment of the active site of Ste24 with the soluble zinc metalloprotease thermolysin, include H434, R440 and L441 (28). All the putative substrate coordinating residues retained less than 2% activity indicating that all three residues are important for proper function of Ste24.

Interestingly, a strong reduction in activity for the C-terminal hydrophobic patch is observed for the amino acids mutated in this region, L410A and L413A. The methyltransferase-coupled diffusion assay specifically measures the CaaX proteolysis activity of Ste24 proteins. These residues were selected due to their proximity to the co-crystallized ZMPSTE24 and CSIM tetra peptide (29). The CSIM (CaaX sequence) aligns itself around the active site zinc but does not contain a farnesyl group (29). Within an appropriate distance there must be a hydrophobic patch able to accommodate this 15-carbon chain. The branched chain is quite flexible, but as a straight chain measures about 13 Å long, so the hydrophobic patch must be at most about 13 Å from the active site. The location of the L410 and L413 residues are about 8 Å and 10 Å away from the zinc ion, respectively, and are at the beginning of a large hydrophobic patch within the Ste24 protein. Both the L410A and L413A display less than 20% retained function, with the L410A mutation showing only 4.2%.

The other hydrophobic patch that was explored resides in the TMD 2 and 3 region was determined in a similar way as the C-terminal, but now for the N-terminal cleavage that is 15 amino acids upstream, or around 25 Å from the cleavage site. The residues selected in this study all lie on either transmembrane helix 2 or 3 in a hydrophobic region at a distance far enough away from the active site, about 30 Å to 37 Å away from the catalytic zinc ion, to still align the N-terminal cleavage site at the HExxH catalytic motif. The methyltransferase-coupled diffusion assay utilized for this study looks specifically at the CaaX cleavage, not the N-terminal cleavage, so activity was not predicted to drop significantly for the mutations in this region (Y79A, F87A and L139A). The lowest of the three was F87A with a reduction to about 40% activity. Y79A showed a mild reduction of 71.7% retained, whereas L139A was the only mutant to not be significantly different than the wild-type protein, retaining around 100% activity. Using PyMOL Molecular Graphics System to examine the Ste24 structure, PDB: 4IL3, Y79 and L139 residues still lie relatively midway down the chamber and do not appear to be making strong intermolecular

interactions with nearby residues (**Figure 2.4B**). F87, on the other hand, lies near the end of TMD 2 and appears to be participating in base stacking near the end of the helix, possibly adding some strong pi-stacking interactions and helping to stabilize the protein structure (**Figure 2.4B**).

Within the proposed binding region, a high variability of activities was observed. Using the PyMOL Molecular Graphics System, PDB: 4IL3, the reason for these variable activities could be explains. The side chains of S256, K257 and R258, retaining 45.2%, 2.4%, and 3.8% retained activity, respectively, all show hydrogen bonds to other amino acids within the Ste24 structure. S256 hydrogen bonds to the backbone of D280, K257 to the backbone of F221, and R258 performs two hydrogen bonds, one also to the backbone of F221, as well as another to the side chain of D164 (**Figure 2.4C**). Disrupting these hydrogen bonds could have strong effects on proper protein configuration, preventing proper activity. Interestingly, D164 and D280 are shown to be critical to Ste24 activity as well (**Chapter 4**), indicating disruption of either amino acid participating in these hydrogen bonds are required for proper Ste24 function. While other amino acids within the putative peptide binding region do participate in hydrogen bonds, these occur from the backbone of the peptide, so a change to the amino acid side chain would be less likely to disrupt these intramolecular interactions.

Residues A264 through F266 lie within 17 Å of the largest portal in the side of the Ste24 structure and surround a large cavity region around the zinc ion, indicating the active site (**Figure 2.4D** and **Figure 2.4E**). The A264, Y265 and F266 residues retain 30.5%, 9.2% and 18.6% activity, respectively, and likely play special roles in allowing the substrate to correctly enter and align around the zinc ion before catalysis. S260 and S262 are more active, 67.3% and 41.3%, respectively, and are further away from the catalytic zinc ion, more than 20 Å away, and the side chains do not look like they are participating in any intra-protein interactions within this cavity, so may not be as strong of coordinators of substrate orientation.

The other three residues studied also lie within the proposed peptide binding region and include G255, S259, and H261. G255A represents one of the more modest mutations as there are no changes in intramolecular interaction, retains 51.3% activity. G255 also lies the furthest away from the active site, being almost 25 Å away from the catalytic zinc ion. The S259A and H261A both show very low activity levels (15.7%, and 8.1%, respectively). Both S259 and H261 side chains face inwards, towards the zinc ion where they could be responsible for hydrogen bonds or dipole interactions with the substrate. Both S259 and H261 amino acids are further away from the

zinc ion (21.1 Å and 16.6 Å, respectively) and this could provide important substrate interaction partners further down the substrate peptide backbone.

Through alanine screening, I determined residues that, when mutated, have detrimental effects on the activity of Ste24. However, the decrease in activity could be for several reasons: 1) mutation of a catalytic residue, 2) decrease in protein stability, 3) disruption of substrate binding, or 4) a combination of the first three. The major caveat to studying only the activity levels of the mutated Ste24 proteins this is that these experiments were solely looking at the activity of the protein. So, I needed to develop additional assays that could look directly at both protein stability and substrate binding.
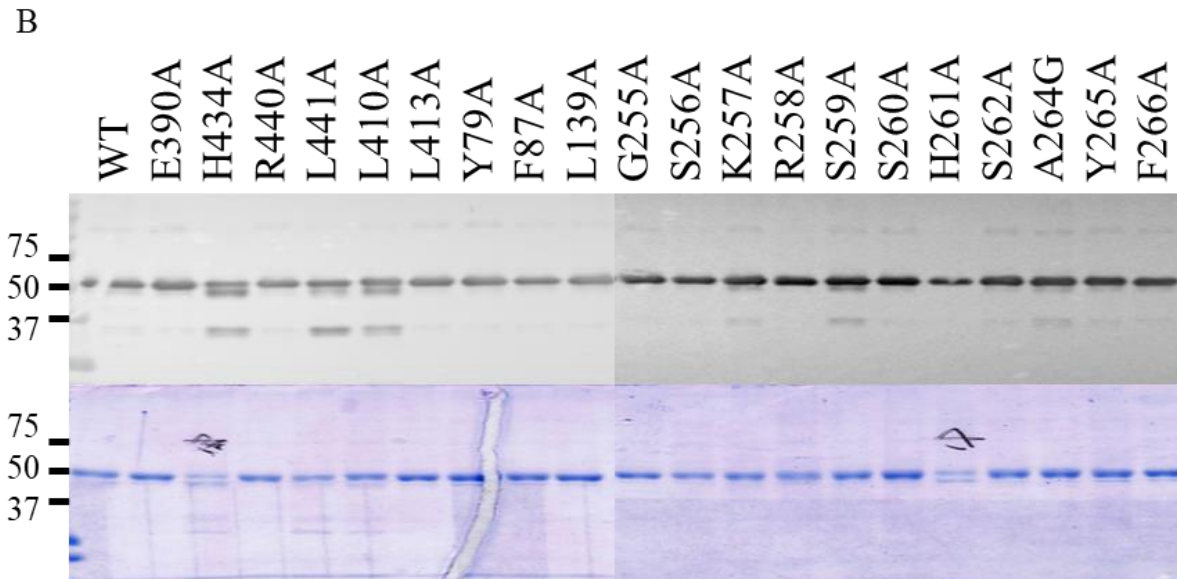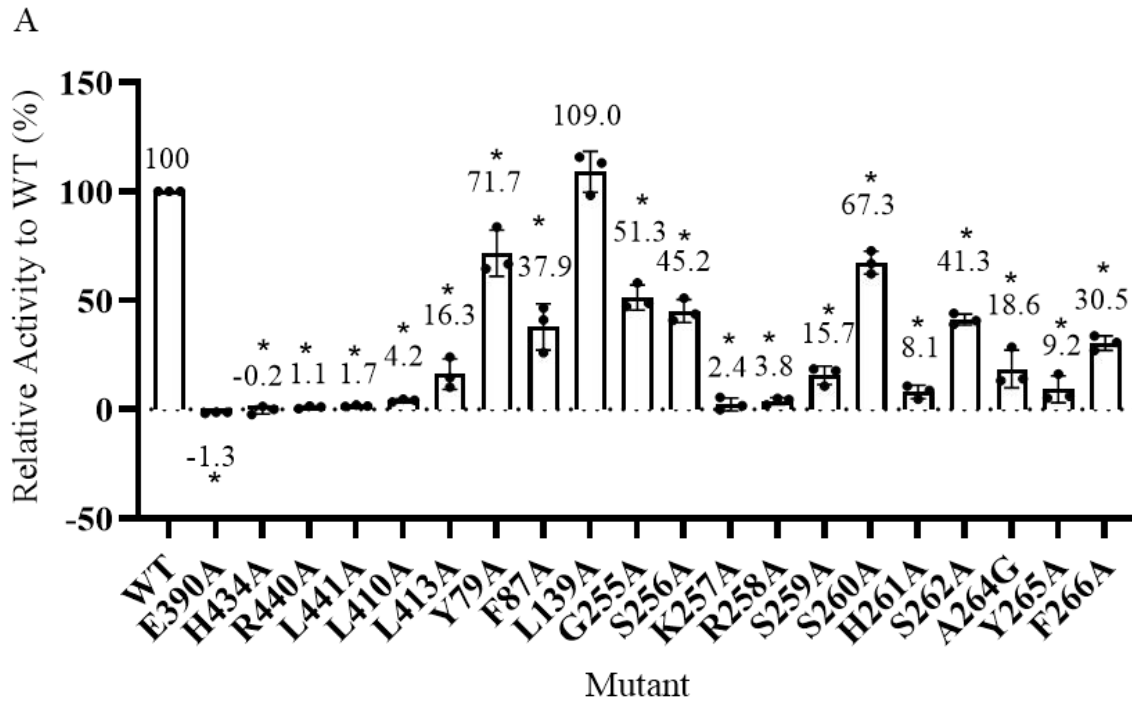
Figure 2.2: Activity levels and expression levels of purified wild-type and mutant Ste24 proteins provide insights into important residues for Ste24 activity. A) Methyltransferase-coupled diffusion assay results of the purified wild-type and mutant Ste24 proteins. Bars indicate the average of triplicate of duplicate experiments. The dots represent the average of duplicate samples in each replicate and the error bars represent the standard deviation amongst the triplicates. Statistical significance ($p < 0.05$) is presented by (*). B) Purified wild-type and mutant Ste24 proteins were run on a 10% SDS-PAGE followed by analysis by immunoblots (0.05 µg total protein, mouse-α-HA primary, and goat-α-mouse secondary) or coomassie stains (1 µg total protein).

Figure 2.3: Structural analysis of Ste24 residues reveal interesting interactions and possible explanations for changes in activities due to mutations. Images were created in PyMOL Molecular Graphics System using *Sacchromyces mitake* Ste24, PDB: 4IL3, solved in Pryor et al, 2013. A) Electrostatic potential surface representation of Ste24. The zinc ion is shown as an orange sphere. B) Analysis of the TMD 2 and 3 hydrophobic patch. TMD2 is shown in cyan, TMD3 in blue. Residues mutated in the activity studies (Y79, F87 and L139) are shown as green sticks. Additional rings in the TMD2 structure possibly forming pi-interactions with F87 are shown as magenta sticks. C) The proposed binding residues (G255-F266) are shown in green and conserved aspartates (D164 and D280) are shown as sticks. Hydrogen bond interactions between R258/D164 and S256/D280 are displayed. D) Exterior surface representation of the largest Ste24 portal is shown in magenta and the proposed peptide binding region (G255-F266) are shown as green sticks. The zinc ion is shown as an orange sphere. E) Interior cavity surface area shows the opening of the largest portal in the Ste24 structure in magenta. Other internal cavities surrounding the proposed peptide binding region (G255-F266, green sticks) are shown as a gray surface. The zinc ion is shown as an orange sphere.

A

B

C

D

E

2.3.2 Photolabeling assays revealed two residues that displayed different binding intensities than wild-type Ste24 protein

To look at substrate binding more directly, many researchers have turned to the use of photoaffinity labeling (PAL), or photolabeling. This process utilizes a probe that contains three parts: 1) a native binding partner mimic, 2) a photoactivatable group, and 3) an enrichment handle. For this study, collaborators from the University of Minnesota in Dr. Mark Distefano's lab created a probe termed C_{10}-para 15-mer **a**-factor (**Figure 2.1B**). This probe contains a similar 15-mer based peptide backbone as our activity assay substrate in **Figure 2.1A** and provides the role of the binding partner mimic. The photoactivatable group is a benzophenone functionality within the hydrophobic farnesyl region, oriented in a C_{10}-para fashion. Finally, the N-terminus of the peptide is linked to a biotin handle by a PEG4 group to allow for enrichment or pulldown of the photolabeled protein. Both the 15-mer **a**-factor and C_{10}-para-**a**-factor substrates have a similar K_M, 8.3 ± 3 µM and 6.4 ± 1.8 µM, respectively. The V_{max} values do differ slightly, with the 15-mer **a**-factor showing a V_{max} of 7814 ± 1043 pmol/min/mg, and the C_{10}-para-**a**-factor substrate having a V_{max} around half of the 15-mer **a**-factor at 3500 ± 334 pmol/min/mg. The similar K_M values indicate they are saturating Ste24 at comparable concentrations, while the large benzophenone group on the C_{10}-para-**a**-factor may make it more difficult for the substrate to enter and properly align in the active site for proteolysis.

After ensuring the probe was a substrate for Ste24, the next important requirement of the C_{10}-para-**a**-factor probe was that this substrate-based probe would bind to a similar site within Ste24 as the 15-mer **a**-factor. A competitive photolabeling assay was performed using 50 µM of C_{10}-para-**a**-factor and varying amounts of the 15-mer **a**-factor substrate (0 µM to 100 µM). The reactions were then quenched with 2x SDS and run on an SDS-PAGE followed by immunoblot analysis with NeutrAvidin®-HRP. The NeutrAvidin®-HRP binds to the biotin tag on the N-terminus of the C_{10}-para-**a**-factor substrate. When the C_{10}-para-**a**-factor photolabels Ste24 it forms a covalent bond with the Ste24 protein that is not affected by the denaturation of 2x SDS. The photolabeled protein should then react with the NeutrAvidin®-HRP to form a band at around 57 kDa.

As shown in **Figure 2.4** as the concentration of 15-mer **a**-factor increases the photolabeling signal decreases, as would be expected if the two substrates showed competitive binding. This is because as the concentration of the native substrate, **a**-factor, increases it would bind within the

binding site and thereby prevent C_{10}-para-**a**-factor from binding and photolabeling the Ste24. These data demonstrate competitive binding between the two substrates, indicating the binding sites between the two molecules are at least overlapping. The peptide sequence between the two are identical, so it is intuitive that the peptide backbone is binding in a similar position for both substrates. However, the larger benzophenone group could cause the lipid group to be oriented differently in Ste24. Both the farnesyl group and the benzophenone groups are similar distances from the cysteine in the CaaX motif, 13 Å and 18 Å, respectively, and are both quite hydrophobic. In addition, both the farnesyl and benzophenone group are at the end of a highly flexible region, allowing for variability in the functional group position. The hope is that even if the precise location of the substrates is not identical, both groups will be binding in a similar region.
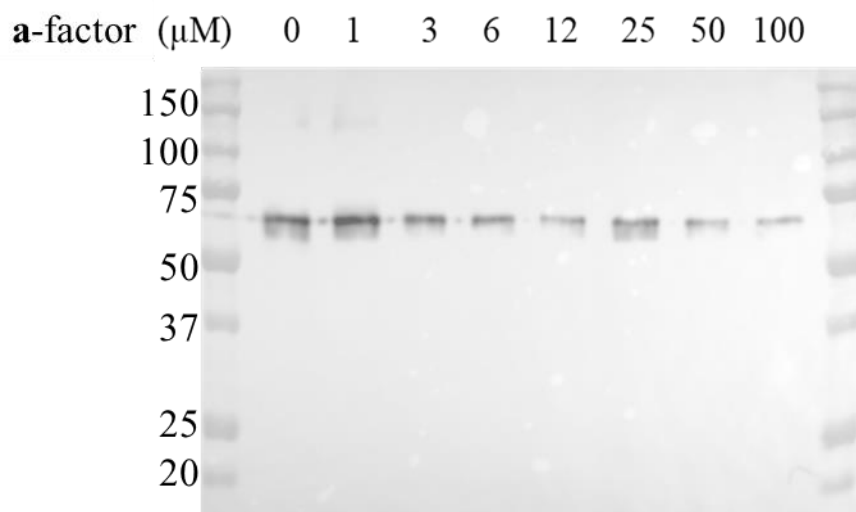
Figure 2.4: C_{10}-para-**a**-factor binds in a competitive fashion to the Ste24 **a**-factor binding site. Samples contained 0.5 µg purified wild-type Ste24, 50 µM C_{10}-para and varying amounts of 15-mer **a**-factor. The samples were equilibrated on ice and then irradiated with UV light (365 nm) for 30 minutes. 2x SDS was mixed 1:1 with each sample after photolabeling and the samples were run on an SDS-PAGE and then analyzed by immunoblot with Neutravidin®-HRP.

At saturating conditions almost all Ste24 variants showed similar relative activity levels with both the C_{10}-para-**a**-factor and **a**-factor substrates (**Figure 2.5A**). A few exceptions are Y79A, F87A, K257A, Y265A and F266A. In the case of Y79A, both substrates still resulted in high activity with the 15-mer **a**-factor retaining 71.7% activity and C_{10}-para-**a**-factor being as active as wild-type (113% compared to wild-type activity). F87A is interesting as the activity with the native 15-mer **a**-factor substrate was moderately low at 37.9% but the large hydrophobic group on the

C$_{10}$-para-**a**-factor probe appeared more favorable in the F87A protein retaining 84.8% activity. The proposed reason for the decreased function of F87A was that the mutation might disrupt existing pi-interactions at the luminal side of the TMD 2 in the native protein (**Figure 2.3B**). The loss of the interaction could be increasing the flexibility of the protein slightly to accommodate the larger benzophenone functional group.

While the K257A mutation did show statistically different activities between substrates, both were still less than 50% active (2.4% with 15-mer **a**-factor and 41.8% with C$_{10}$-para-**a**-factor). The Y265A and F266A were both extremely more active in the presence of C$_{10}$-para-**a**-factor (98.4% and 76.5%, respectively) than with 15-mer **a**-factor (9.2% and 30.5%, respectively). These residues do lie quite close to the largest portal in the Ste24 chamber. The mutation of these large aromatic residues could further open the portal to allow the larger benzophenone group to enter.

Photolabeling reactions were then performed in triplicate for each mutant. In these experiments, the purified Ste24 proteins were reconstituted in *E. coli* lipids with a rapid dilution in the presence of C$_{10}$-para-**a**-factor and incubated on ice for ten minutes to allow the samples to equilibrate and the substrate time to bind to Ste24. As an unlabeled control, half of the sample was kept in the dark (-UV) while the other half was photolabeled by irradiation of a 365 nm light (+UV) for 30 minutes. The samples were loaded onto NeutrAvidin® Agarose Resin to pulldown all photolabeled protein through the biotin handle on the N-terminal end of the C$_{10}$-para-**a**-factor substrate. The protein was then eluted from the NeutrAvidin® Agarose resin with 2x SDS loading dye and run on an SDS-PAGE and analyzed by immunoblot with a mouse α-HA primary antibody followed by a goat α-mouse-HRP secondary antibody.

Figure 2.5B displays representative blots for each sample, and ImageJ quantification values are shown in **Figure 2.5C**. Quantification was performed by measuring the pixel count of each +UV band and subtracting the number of pixels in the -UV sample as described at http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf. Interestingly, the E390A mutation showed that the photolabeling signal was double the intensity of the wild-type (WT) Ste24 signal. The E390 residue is known to be a zinc-binding residue characteristic of zinc metalloproteases (28). It is known that Ste24 is not catalytically active in the absence of zinc (140), and the data presented in **Figures 2.2A** and **2.5A** display the activity levels being reduced to around 0%. This is a clear example that activity data is not the best indicator of substrate binding. The zinc ion in the active site is coordinated to three residues within the Ste24 active site, H297, H301 and E390,

as well as a water molecule. During catalysis, the zinc ion aids in activating the water molecule and forming a stabilizing coordination interaction with the intermediate (141). Removing the interaction of the E390 residue with zinc by mutation to alanine likely prevents the correct coordination state of the zinc for the reaction to proceed. However, this decrease in activity does not provide any direct details into the binding of the substrate. Additionally, because the substrate is not being turned over it could be the probe resides in the binding site longer, allowing more time for the C_{10}-para-**a**-factor to photolabel Ste24. The only other statistically significant point is the H434A mutation showing only a 28% photolabeling signal compared to WT Ste24. This H434 residue corresponds to a substrate-coordinating residue found in thermolysin (28). Loss of the proper substrate orientation likely causes the farnesyl and benzophenone group into a different location or prevents proper binding of the substrate. Either scenario would likely decrease the photolabeling signal observed in the immunoblot analysis.

One big thing to keep in mind with these studies is that photolabeling does not necessarily directly indicate substrate binding and is not quantitative in terms of substrate binding. Changes in the binding site could alter the binding site just enough to alter the orientation of the **a**-factor based sequence. These small changes could cause the C_{10}-para benzophenone to shift in place just enough to be out of range to crosslink nearby amino acids. In this case, the substrate is still binding but is less likely to photolabel. For example, the Y79A mutation maintains over a 100% activity with the C_{10}-para-**a**-factor probe, but only shows a photolabeling signal of 51%. The C_{10}-para-**a**-factor substrate must still be binding to the protein as the substrate turnover rate is very high, but the mutation in this residue could alter the position of the substrate, or slightly alter the Ste24 structure, enough that the photolabeling efficiency of the benzophenone group is reduced.

Figure 2.5: Photolabeling analysis of purified wild-type and mutant Ste24 proteins reveal small insights into substrate binding. A) Activity assay results comparing 15-mer **a**-factor and C10-para substrates. Dots represent the average of duplicate samples in each replicate, bars are at the average of three replicates, error bars represent the standard deviation. The black numbers are the percent activity levels with 15-mer **a-**factor and blue are for the percent activities for C_{10}-para. (*) indicates mutations that are stasticially significant (p-value < 0.05) between both probes. B) For the photolabeling samples, a pulldown was performed utilizing NeutrAvidin® Agarose resin. Photolabeled protein was eluted from the resin with 50 µL of 2x SDS. Then 10 µL of each sample was used to run on a 10% SDS-PAGE followed by immunoblot analysis with mouse α-HA primary and goat α-mouse-HRP secondary antibody. C) Quantitation of photolabeling signals from (B) was quantified similar to that described at http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf. The bands of both the +UV and -UV samples were quantified with ImageJ and the -UV signal was subtracted from the +UV sample. Then the ratio of the mutant signal was compared to the WT Ste24 signal. The graph was created in Graphpad Prism 9. The bars indicate the average of three replicate photolabeling intensites, dots indicated individual data points and (*) indicates statistical significance (p-value < 0.05). Error bars display the standard deviation.

A

B

C

2.3.3 Thermostability analysis and direct substrate binding assay development gave more specific insights into changes in substrate binding in Ste24

Changes in protein activity could be the result of three things: 1) changes in catalytic ability, 2) changes in substrate binding, or 3) changes in protein stability. To determine whether the observed changes in activity and photolabeling were due to protein stability the melting temperature (T_m) of the presented mutations were measured on a Nanotemper Tycho NT 1.6. This instrument utilizes a form of nano-differential scanning fluorimetry that monitors the fluorescence changes of tryptophan and tyrosine in the protein upon protein unfolding. Thermal stability was calculated using the Nanotemper Tycho NT 1.6 instrument and software, that measures the change in the ratio of fluorescence at 350 nm to the fluorescence at 330 nm over a range of temperatures from 30 °C to 95 °C.

The thermal stability of each Ste24 variant was measured twice and the average curve is shown in **Figure 2.6A**. The T_m was quantified at the inflection point in each curve and was graphed in **Figure 2.6B**. The T_m of the wild-type protein was calculated to be 50.7 °C, which to the best of my knowledge is the first time the stability of the wild-type Ste24 protein has ever been reported. Amongst the Ste24 mutants, the range of T_m values contained the maximum, therefore more thermally stable, at L441A (54.9 °C) and the minimum, least stable, at R258A (46.3 °C). A change over 2 °C has been shown to be significant in the thermal stability of protein and protein interactions (142,143). Most of the Ste24 variants were only tested twice, so statistical significance can be misleading. However, the L441A and R258A are the two mutations that do present as statistically significant (p-value < 0.05) by one-way ANOVA, as determined by GraphPad Prism 9. L441A is the only mutation showing a statistical increase in thermal stability even though the activity level is only about 1.7% that of wild-type Ste24. This L441 residue is suspected to aid in coordinating the substrate around the zinc in the active site, due to structural alignments with the active site of other zinc metalloproteases (28). These data indicates that the inactivity of L441A is not due to instability of the protein. R258A is the only mutation that was statistically significant on the lower end of the T_m spectrum. The activity levels of R258A are also very low, 3.8% that of wild-type, so the loss of stability likely contributes to lower activity levels. Other mutants more than 2 °C lower than the wild-type Ste24 T_m were E390A, S256A, S259A, and H261A (48.6 °C, 48.6 °C, 47.7 °C, and 48.3 °C, respectively). Most of these mutations did have a significant loss of activity as well, with E390A displaying an average of less than 0% activity, H261A with 8.1%,

S259A at 15.7% and then S256A still retained 45.2% activity. While S256A does appear to be significantly less thermally stable than wildtype, it does still retain almost half of the activity of the wild-type protein. While the thermal stability curves indicate that half of the protein is denatured at 48.6 °C (T_m), because the activity assay is performed at 30 °C it could be the protein can still withhold enough of the proper shape to turnover substrate during the methyltransferase-coupled diffusion assay. E390 is a zinc-coordinating residue that has been shown to be important for proper Ste24 catalysis (140) and removal of this stabilization of zinc-binding could cause instability in the protein. Both S259A and H261A also show low catalytic activity that could then be explained by the lower stability of the protein. The other mutations studied here (H434A, R440A, L410A, L413A, Y79A, F87A, L139A, G255A, K257A, S260A, A264G, Y265A, and F266A) did not show a drastic change in T_m, so their changes in activity levels are likely not due to changes in protein stability.

The photolabeling data presented in **Section 2.3.2** presented several interesting findings. One was that, even though E390A was completely inactive, this mutant showed a photolabeling signal that was more than double that of wild-type Ste24 (**Figure 2.5C**). This increase in photolabeling signal could be due to an improvement of substrate binding or it could be that the position of the photolabeling benzophenone group of C_{10}-para-**a**-factor was shifted into a position that allowed a more efficient photolabeling signal. It could also be that since there is no turnover of the C_{10}-para-**a**-factor substrate, as there is in WT Ste24, it could be the probe resides in the binding site longer than the wild-type protein, providing more time for the photolabeling reaction to occur. Development of a binding assay for Ste24 was greatly needed to determine whether the change in photolabeling was due directly to a change in the ability of the protein to bind to the substrate.

To do this, I utilized microscale thermophoresis (MST) for Ste24 K_D measurements. MST works by monitoring the movement of fluorescent molecules down a slight temperature gradient. One of the binding partners is fluorescent and is held at a constant concentration while the other partner is not fluorescent and the concentration is varied. The fluorescent binding partner was developed by Dr. Mark Distefano's lab at the University of Minnesota and is displayed in **Figure 2.1C**. This molecule still shares the same 15-mer **a**-factor sequence and farnesyl group shown in **Figure 2.1A**, but also contains an N-terminal, fluorescein (5-FAM) moiety. This molecule showed similar kinetics ($K_M = 2.3 \, \mu M \pm 0.7 \, \mu M$, $V_{max} = 3932 \, pmol/min/mg \pm 410 \, pmol/min/mg$) to the **a**-

factor substrate (K_M = 8.3 μM ± 2.2 μM, V_{max} = 7814 pmol/min/mg ± 867 pmol/min/mg) for WT Ste24 in the methyltransferase-coupled diffusion assay. Initial experiments attempted to fluorescently label the Ste24 protein, but issues arose at high levels of the **a**-factor substrate. At high levels (> 40 μM) the **a**-factor molecule starts to precipitate and causes "aggregation" readings in the MST. The data indicated that the substrate had to be the fluorescent molecule as this allowed it to be held constant at low concentrations.

Preliminary experiments utilized the WT and E390A Ste24 proteins. Due to the increased photolabeling ability seen by E390A (**Figure 2.5B-C**) I hypothesized that the K_D value may also be lower, indicating higher affinity, than the WT protein. As seen in **Figures 2.6A-B**, the WT protein showed a K_D at a similar value to the K_M (3.2 μM ± 1.1 μM) while the E390A mutant showed a moderately lower K_D value at 0.1 μM ± 0.1 μM. This could be a result of two things. The first being that, as discussed before, the E390 residue is a zinc-coordinating residue. While the zinc residue is needed for catalysis it may not actually provide as much structural integrity as was previously believed and the substrate could still be binding well within the active site. The second factor that is likely the predominant reason, is that the 5-FAM-**a**-factor is still a substrate of Ste24. As are the cases with most enzymatic reactions, the product of the reaction has a much lower affinity for the protein to promote substrate turnover. Because E390A can no longer produce product the K_D measured is solely of the substrate, while the WT K_D is a combination of the substrate and product K_D values.

To further determine the role of the Ste24 proposed binding residues, I planned to supplement the mutants described in **Section 2.3.1** with the E390A mutation. I began with the K257A mutation as this residue was also inactive (2.4% activity), thermally stable (T_m = 49.5 °C) and was able to be purified at a high concentration. The K_D of this double-mutant was calculated by MST to be 2.6 μM (**Figure 2.6C**) . This increase in K_D (or decrease in affinity) could result from an unknown interaction being lost between Ste24 and the substrate. The Nanotemper Tycho 1.6 instrument was used to determine the melting temperature (T_m) of each of the WT, E390A and K257A single mutation. The T_m for the WT and K257A mutations were not statistically different (50.7 °C and 49.5 °C, respectively. This indicates that the K257A mutation did not significantly affect the structural stability of Ste24, implicating that these changes in K_D are in fact due to changes directly in substrate binding, independent of activity. The E390A mutation did show a

slight decrease in thermal stability (ΔT_m = -2.1 °C) but the difference was not statistically significant.

A couple of problems did become apparent during these initial MST experiments. One problem was that there was some background fluorescence seen in the reaction buffer and purified protein samples. Further analysis of the components of the buffer indicated that the protease inhibitors used in purification of the Ste24 variants also fluoresced within the same bandwidth of the Nanotemper Monolith .115 blue laser that also excites the 5-FAM fluorophore. The 5-FAM fluorophore has a maximum excitation wavelength of 490 nm and maximum emission wavelength of 514 nm. The blue laser has a bandwidth from about 450 nm to 500 nm for excitation and about 510 nm to 560 nm for the emission. Following these MST experiments, it was found that the protease inhibitors aprotinin and AEBSF were also able to fluoresce within this region and interfered with the data obtained from the MST experiments. These same components did not show any background fluorescence signal in the red laser that has an excitation between 600 nm and 650 nm and an emission spectrum of about 670 nm to 720 nm. The other problem found with the use of the 5-FAM **a**-factor substrate (**Figure 2.1C**) was that because it is a substrate of the CaaX cleavage of Ste24 it is likely that the product is being formed and influencing the K_D value measured for active strains of Ste24. This means that there was a need of a probe that was not a substrate of Ste24 but did still mimic the **a**-factor substrate, and that also would fluoresce in the red laser as opposed to the blue.

Figure 2.6: MST and thermostability experiments display direct measurements of K_D and T_m values between Ste24 mutants. A) Purified wild-type and mutant Ste24 proteins were loaded into capillary tubes straight from the concentrated protein stocks, so no protein dilution was performed. The capillaries were then loaded into a Nanotemper Tycho NT 1.6 to measure the T_M of each Ste24 variant. Curves represent average curves of two replicates. B) T_M values of Ste24 mutants obtained from the software of the Nanotemper Tycho NT 1.6, indicative of the inflection point of the curve. Dots represent values of individual replicates. * Indicates statistical significance ($p<0.05$). C) 20 µL of the concentrated, purified wild-type and mutant Ste24 proteins was used to make a series of 16, 1:1 dilutions of 10 µL volume samples in 10 mM Tris-HCl, pH 7.5. Then 10 µL of the 5-FAM **a**-factor substrate was added to a final concentration of 50 nM. The samples were then loaded into Nanotemper Standard Capillary tubes and loaded into the Nanotemper Monolith .115. Microscale thermophoresis results demonstrate duplicate runs of WT, E390A and K257A-E390A. Calculated K_D values are displayed above the graphs.

A

B

C

WT K$_D$ = 3.2± 1.1 μM

E390A K$_D$ = 0.1+/- 0.07 μM

K257A-E390A K$_D$ = 2.6+/- 0.1 μM
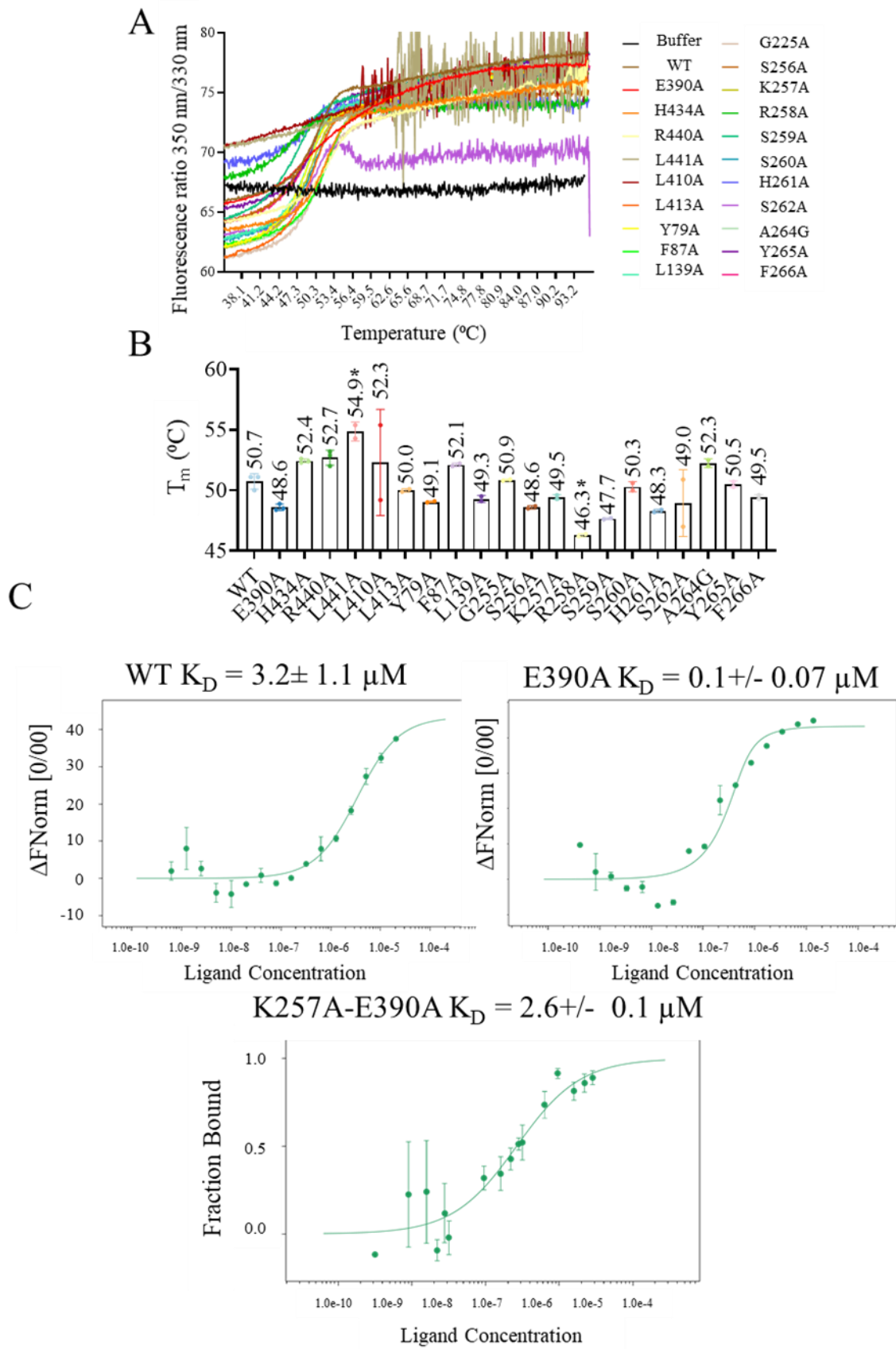
2.3.4 Use of an uncleavable substrate probe allows definition of the K_D values of Ste24 mutant proteins to be determined

I became aware of two problems with the 5-FAM **a**-factor in the MST instrument. One was that because the probe also acted as a substrate of Ste24, the K_D value observed is likely influenced by the K_D of the product being formed during the reaction . The other problem was that the buffer used to purify the Ste24 variants contained protease inhibitors that also fluoresced within the bandwidth of the blue laser of the Nanotemper Monolith .115. Our collaborators at the University of Minnesota then developed a new, uncleavable probe with a red Cy5 fluorophore (**Figure 2.7**). The uncleavable probe contains an isostere group between the cysteine and valine residues of the CaaX sequence, meaning the carbonyl group of the linking peptide bond is reduced, leaving just an -NH- group between the two residues. Because the peptide backbone is lacking the carbonyl, Ste24 is unable to cleave the backbone in the CaaX sequence.
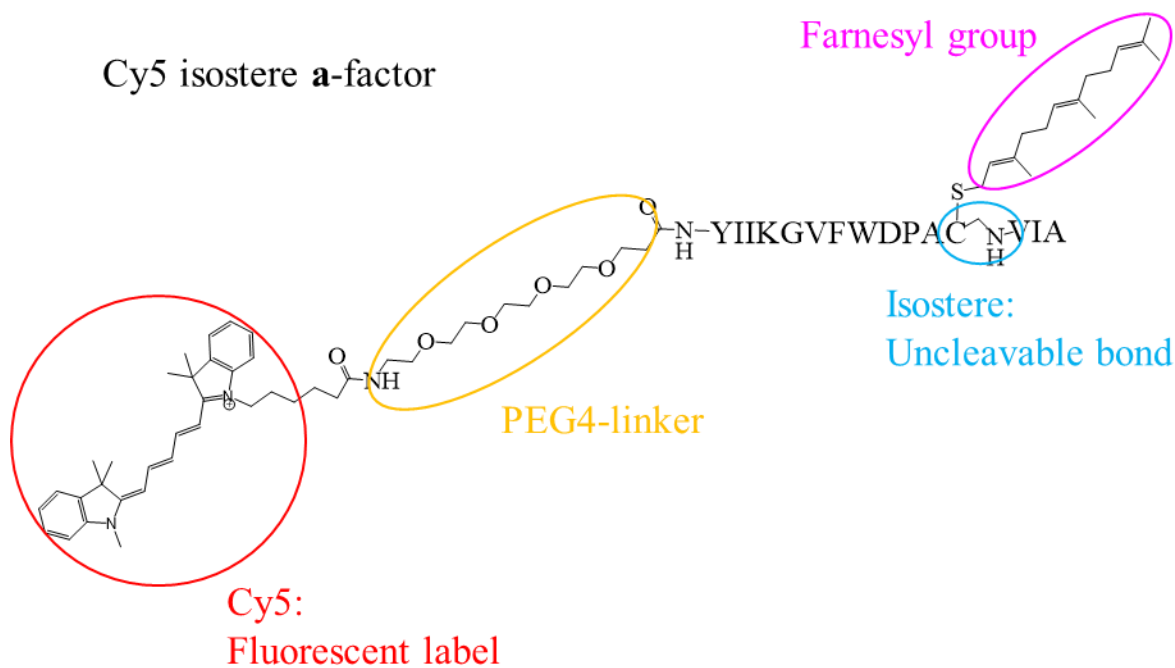
Figure 2.7: Uncleavable Cy5-labeled **a**-factor probe

With this new uncleavable probe it was possible to measure the binding of each single mutation in Ste24 to determine the K_D values (**Figure 2.8**, **Table 2.3**) without the product formation. The binding constant (K_D) value of the Ste24 wild-type protein was determined to be

1.7 µM. The K_D value for the WT protein with the 5-FAM, cleavable substrate was slightly higher at 3.2 µM that was as to be expected. Enzymes are typically designed to have a lower affinity for the product of a reaction, resulting in the release of the product and binding of a new substrate molecule. This value is also consistent with the K_M value of 8.3 µM.

The protein concentrations of the Ste24 mutants were kept as concentrated as possible to try to obtain a full binding curve. However, as the concentrations of the purified proteins vary slightly, some were not concentrated enough to obtain a full saturation curve. Ideally, a binding curve would start with a protein concentration of at least 10x the concentration of the K_D value. In the case of the wild-type protein, the K_D value was determined to be about 1.7 µM, so a concentration of around 20 µM would be sufficient. The K_D value of the other mutant Ste24 proteins is yet unknown. By starting with a concentration of at least 60 µM hopefully at least the start of a binding curve could be observed. If the K_D of these mutants do shift significantly, to even 10 µM or higher, the required Ste24 protein concentration would be much higher than what has been obtained in our lab. Development of a better expression system is currently being designed with baculovirus in SF9 insects cells to attempt to improve yields of purified Ste24 mutants.

The K_D values of the Ste24 mutant proteins are unknown, so to begin testing the samples the Ste24 variants were used at as high of a concentration as was possible from the stock solutions. In **Figure 2.8**, the figures on the left represent the raw F_{norm} data. Those that formed nearly a full binding curve were transformed to fraction bound using the Nanotemper Monolith .115 software associated with the MST instrument and used to obtain a K_D value (**Table 2.3**). Moving forward, mutations that had a concentration that was too low for a full curve will be repurified at a higher concentration. Ideally, the development of a new expression system (baculovirus expression in SF9 insect cells) should allow for a concentration that is considerably higher than in yeast proteins.

Of the catalytic and substrate coordinating residues (**Figure 2.8 A** and **B**), the E390A mutation was the only one to form a full concentration curve and displayed a K_D of 0.5 µM, similar to that seen with the 5-FAM **a**-factor substrate (**Figure 2.7C**). While statistical significance cannot be determined due to this being of only two replicates, it does support the photolabeling data from **Figure 2.4 B** and **C**. The photolabeling signal was almost double the intensity of the WT Ste24 protein and the K_D calculated here was less than half the value shown in WT. However, the standard deviation of the two K_D values for wild-type and E390A do overlap, so no direct conclusions can be made yet. The other two mutations that were also transformed to fraction bound

(H434A and L441A, **Figure 2.8B**) were included to show what the curves look like when the concentration is not high enough to get a full curve. The software did suggest the K_D values of H434A and L441A were 9.8 µM and 7.6 µM, respectively, because there is not an inflection point reached in the curve this is a very rough estimate. If these estimated K_D values are correct, a starting concentration of 100 µM and 80 µM, respectively, would be the ideal minimum to achieve a full binding curve. However, the L441A mutation was the most stable variant of Ste24 measure, with a T_m of 54.9 °C, while being completely inactive (1.7% relative activity to WT). L441 has been proposed to be an important substrate-coordinating residue based on structural alignments between the active site of Ste24 and another well-studied, soluble zinc metalloprotease, Thermolysin (28). Likely, the L441A mutation is then inactive due to changes in substrate binding.

Within the hydrophobic patches there were two mutations that stood out. Within the C-terminal hydrophobic patch, the L410A mutation showed a nearly full binding curve in the fraction bound, as it did make it past the inflection point of the curve. Again, this is a rough estimate as there was not a full curve obtained, but the K_D value, 3.8 µM, was slightly higher than WT, indicating a possible decrease in binding affinity. The L410A mutation did show an increase in thermal stability and also significantly decreased the activity to about 4% compared to WT (**Figure 2.6B** and **Figure 2.2A**). The slight decrease in binding affinity could be the cause for the complete loss of -aaX activity as it was shown to still be a stable protein by nano-DSF. Additional replicates are required, as well as a higher L410A concentration, to obtain a full binding curve to confirm this shift in K_D. The L139A mutation from the TMD 2 and 3 hydrophobic patch showed a much better affinity for the substrate than the wild-type protein, though the error of the data points was higher. This was the only mutation that retained an activity level comparable to the wild-type (**Figure 2.2 A**), indicating it likely is binding at least on par to the wild-type protein.

For the proposed peptide binding region, the only mutation to obtain a full curve was the G255A mutation. The K_D value was slightly lower, but still close to the wild-type protein. The activity of the G255A mutation was around 55% the activity of the wild-type Ste24 protein. The thermostability of the mutation was almost identical to the wild-type (WT at 50.7 °C and G255A at 50.9 °C), so this change in activity is not due to the stability of the protein. This is a clear example of the problem when researchers use activities of proteins to estimate binding. The G255A mutation binds to the substrate just as well as the wild-type protein, but then loses around half the

activity. The glycine residue then must play a role that benefits the activity of the protein in catalysis without an interaction with the binding of the **a**-factor substrate.

While the photolabeling signal of G255A was lower than the WT protein, there is a significant limit with photolabeling assays. As discussed in **Section 2.3.2**, slight changes in substrate orientation can cause drastic changes in the photolabeling signal, especially because the benzophenone group is attached at a large flexible region. A small change in orientation of substrate, that does not affect the binding, could be causing the benzophenone to orient slightly away from a nearby residue to photolabel.

The other two mutations in **Figure 2.8H** (S256A and R258A) were included to show an incomplete binding curve of these mutations to indicate the range of the K_D values that could be obtained by this binding assay. The Nanotemper Monolith .115 software does estimate the K_D values of S256A and R258A to be 20.8 µM and 60.8 µM, respectively. This was very exciting as it indicates that the range of K_D values amongst these Ste24 mutations could differ by almost 10-fold. By increasing the starting concentration of the purified Ste24 proteins this MST assay should be able to explore this large range of K_D values of these mutations. Currently, the expression system in yeast does limit the amount of protein that can be obtained to the 20 µM – 40 µM range. A better expression system is being developed in baculovirus that should allow higher expression levels of the Ste24 protein, allowing for a higher concentration of protein after purification. This will aid in obtaining full binding curves, and thereby K_D values of all the Ste24 mutants.

Figure 2.8: Development of an MST binding assay allows for the determination of the K_D of multiple Ste24 mutants. 20 µL of the concentrated, purified wild-type and mutant Ste24 proteins was used to make a series of 16, 1:1 dilutions of 10 µL volume samples in 10 mM Tris-HCl, pH 7.5. Then 10 µL of the Cy5 isostere **a**-factor substrate was added to a final concentration of 20 nM. The samples are then loaded into Nanotemper Standard Capillary tubes and loaded into the Nanotemper Monolith .115. Microscale thermophoresis results demonstrate duplicate runs of the same reaction samples. A) The raw data of F_{norm} of the zinc-binding mutant E390A as well as the substrate-coordinating residues. B) Fraction bound conversion of the mutations from (A) showing at least a partial binding curve. C) The raw data of F_{norm} of the C-terminal hydrophobic patch mutations. D) Fraction bound conversion of the mutations from (C) showing at least a partial binding curve. E) The raw data of F_{norm} of the TMD 2 and 3 hydrophobic patch mutants. F) Fraction bound conversion of the mutations from (E) showing at least a partial binding curve. G) The raw data of F_{norm} of the predicted peptide binding mutations. H) Fraction bound conversion of the mutations from (G) showing at least a partial binding curve.

Catalytic and Substrate-Coordinating Mutants
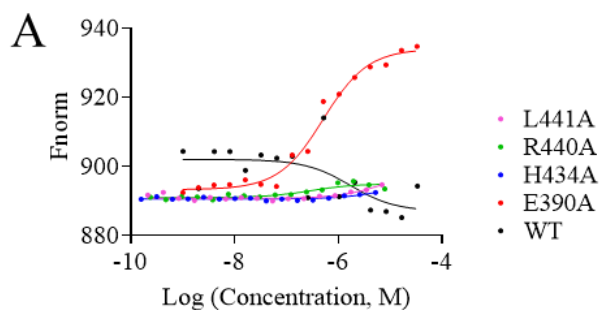
A — Fnorm vs Log (Concentration, M): L441A, R440A, H434A, E390A, WT

B — Catalytic and Substrate-Coordinating Mutants — Fraction Bound vs Log (Concentration, M): L441A, H434A, E390A, WT

C — C-Terminal Hydrophobic Patch — Fnorm vs Log (Concentration, M): L413A, L410A, WT

D — C-Terminal Hydrophobic Patch — Fraction Bound vs Log (Concentration, M): L410A, WT

E — TMD 2 and 3 Hydrophobic Patch — Fnorm vs Log (Concentration, M): L139A, F87A, Y79A, WT

F — TMD 2&3 Hydrophobic Patch — Fraction Bound vs Log (Concentration, M): L139A, WT

G — Binding Mutants — Fnorm vs Log (Concentration, M): S260A, S259A, R258A, K257A, S256A, G255A, WT

H — Binding Mutants — Fraction Bound vs Log (Concentration, M): R258A, S256A, G255A, WT
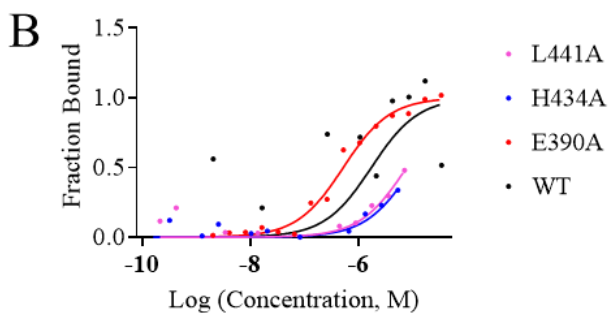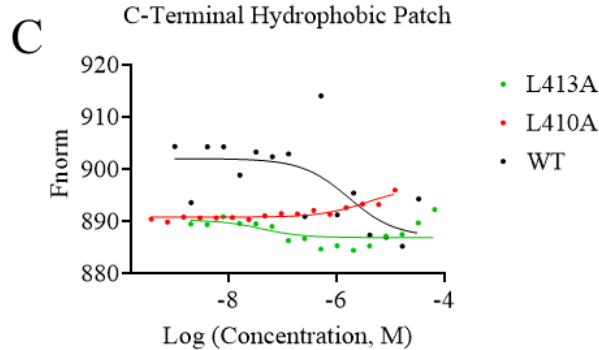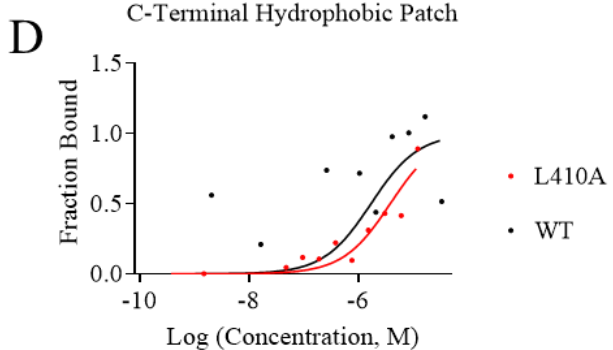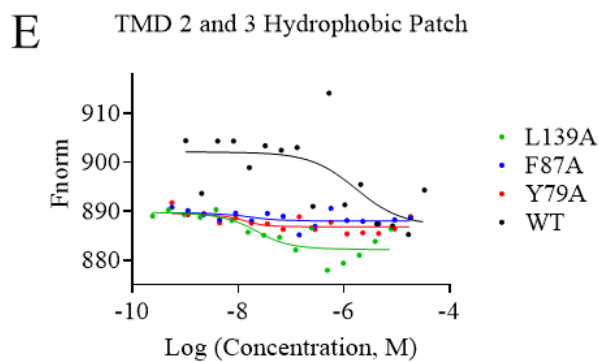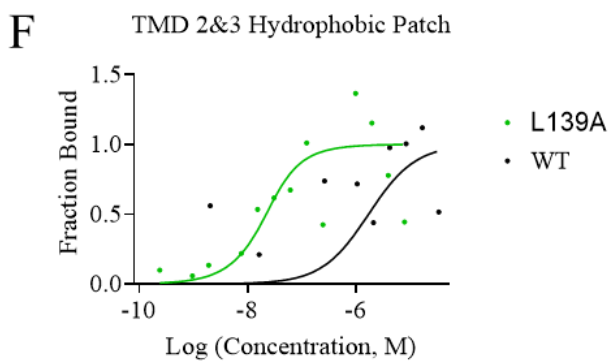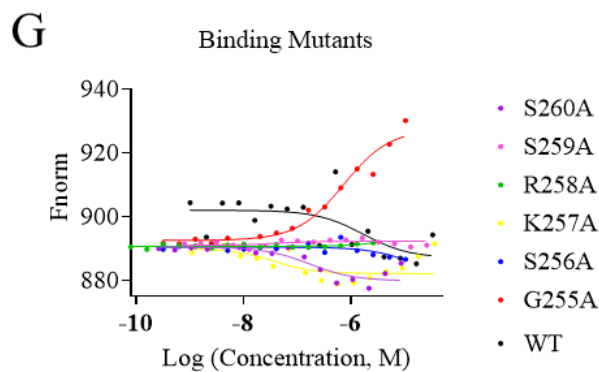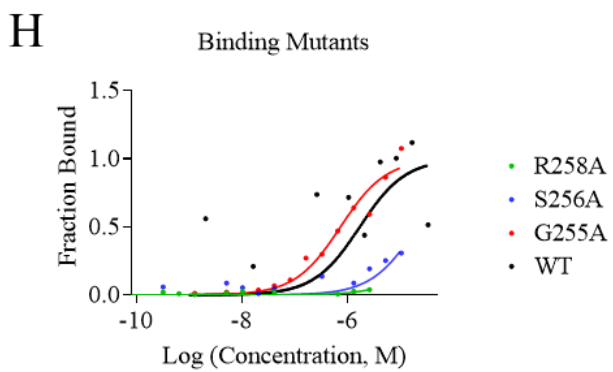
Table 2.3: K_D values of Ste24 mutants that showed at least a nearly full binding curve in the F_{norm} from **Figure 2.8**

Mutant	K_D (µM)
WT	1.7 ± 2.4
E390A	0.5 ± 0.07
L410A	3.8 ± 2.9
G255A	0.7 ± 0.1

2.4   Conclusion

Taken together, the alanine-screening mutagenesis has revealed a lot of information about the Ste24 substrate binding region (**Table 2.4**). Activity assays revealed Ste24 variants that had reduced or abrogated catalytic activity that could possibly be explained by hydrogen bonding interactions, presence of pi-pi interactions, or removal of catalytic/substrate-coordinating interactions. However, this did not present details specifically about substrate binding, and in photolabeling analyses, a lot of these residues (E390 notably) did not match the changes in activity. While the photolabeling studies did provide some information, like H434A, L441A, F87A and L139A showed consistency between activity and photolabeling, indicating changes in binding, other mutations did not (Y79A).

The development of a direct binding assay utilizing MST has allowed for the direct analysis of changes in binding affinity of several Ste24 mutants. While many of these experiments need to be repeated with a higher concentration of Ste24 protein, several mutations did present as full binding curves and the need for a binding assay for Ste24 was successfully satisfied. For example, the G255A mutation showed similar thermal stability to WT but only retained about 50% of the activity. This could have been due to either a change in substrate activity, or a change directly in enzymatic activity. The K_D value of G255A was similar to the WT protein, indicating that this loss of activity must be mostly due to a change in activity directly, and not due to an interaction with the substrate. On the other hand, the L410A and L441A mutations showed a more stable protein than wild-type Ste24 but also a lack of activity. For these two mutations the change in K_D from the

binding assay displays that the change in catalytic ability of these Ste24 variants is due to the change in substrate binding, specifically.

Though many of these mutated proteins will need to be repeated at higher concentrations, there was clear success in the development of a Ste24 binding assay. This MST assay for Ste24 displays the ability to measure a wide range of K_D values, from at least the 0.1 µM to 60 µM range. While it does require a high concentration of Ste24 protein, the development of a better expression system should help to obtain this requirement.

Table 2.4: Summary of all activity, photolabeling, thermostability and binding affinity values obtained in **Chapter 2**. As most of the photolabeling signals were not statistically significant, only those that were significant are indicated. The K_D values presented are only of those that had a nearly complete or saturating binding curve.

Predicted function	Mutation	Retained activity (%)	Statistically different photolabeling signal? (Y/N)	T_M °C	K_D (µM)
Zinc-coordinating	E390A	-1.3	Y	48.6	0.5
Substrate-coordinating	H434A	-0.2	Y	52.4	9.8
	R440A	1.1	N	52.7	0.3
	L441A	1.7	N	54.9	7.6
C-terminal hydrophobic patch for farnesyl binding	L410A	4.2	N	52.3	3.8
	L413A	16.3	N	50.0	--
TMD 2&3 hydrophobic patch for farnesyl binding	Y79A	71.7	N	49.1	--
	F87A	37.9	N	52.1	--
	L139A	109.0	N	49.3	--
Peptide binding	G255A	51.3	N	50.9	0.7
	S256A	45.2	N	48.6	20.8
	K257A	2.4	N	49.5	--
	R258A	3.8	N	46.3	60.8
	S259A	15.7	N	47.7	--
	S260A	67.3	N	50.3	--
	H261A	8.1	N	48.3	--
	S262A	41.3	N	49.0	--
	A264G	18.6	N	52.3	--
	Y265A	9.2	N	50.5	--
	F266A	30.5	N	49.5	--

2.5 Future directions

1) Repeat triplicate experiments with all Ste24 mutant proteins with MST and the Cy5 isostere

2) Develop a new expression system to obtain higher concentrations of the Ste24 mutants to obtain full binding curves for each

3) Create additional mutations that replace the native amino acid with other residues that are not alanine

   a. Y265F, to determine whether it is the aromatic functionality or the hydrogen bond of the tyrosine that affects the activity and substrate binding

   b. F266Y – to determine whether the aromatic group is necessary and whether addition of the hydroxyl group disrupts this effect, especially since it lies next to the Y265 residue near the portal

   c. L410F and L413F – to determine whether a larger hydrophobic group here could enhance the binding of the farnesyl group

   d. Mutations which alter charge states, for example mutating K257D or K257E, R258D or R258E, to determine whether the positive charges at these amino acids are required for proper substrate turnover

4) Triplicate values for the T_m of each mutant

5) Determine whether the Cy5 isostere is cleaved at a different site by Ste24 by incubating the substrate with Ste24 at 30 °C and then running the reaction on a UPLC-MS to determine whether the substrate is altered.

CHAPTER 3.   ELUCIDATION OF THE LIPID BINDING SITE FOR THE FARNESYL GROUP DURING THE -AAXING CLEAVAGE BY USE OF TANDEM MS/MS

3.1   Introduction

CaaX processing is a common post-translational pathway that likely modifies hundreds of proteins in the cell (1,25,39-41,144). Examples of proteins that undergo CaaX processing include the Ras family of proteins, lamin proteins and yeast mating pheromone peptides. CaaX proteins end in a CaaX motif, that is defined as a cysteine, two aliphatic residues, and one of many different residues. The post-translational modification of CaaX proteins follows the same general scheme (**Figure 1.1A**) where the protein or peptide is first lipidated at the C-terminal cysteine with the addition of either a farnesyl or geranylgeranyl group, then cleaved by proteases to remove the three C-terminal -aaX residues, followed by methyl esterification of the now exposed, lipidated cysteine (19,25,144).

The yeast protease Ste24 is known to act upon the mating pheromone **a**-factor in this post-translational pathway. The necessity of the farnesyl group added to the cysteine during the CaaX processing of **a**-factor has been debated, as the mammalian homolog was shown to be able to cleave the native substrate prelamin A without the farnesyl group present *in vitro*, as shown by mass spectrometry and fluorescence-based kinetic analysis (130,145). In this study, while they did show that the non-farnesylated substrates could be acted upon by purified Ste24 the K_M values were greatly increased by around 10-fold between farnesylated and non-farnesylated substrates (130). Regardless of whether farnesylation is required for proteolysis by Ste24, *in vivo* both **a**-factor and prelamin A precursors have been detected as farnesylated substrates (25).

Whether or not it is required specifically for binding, the farnesyl group is a 15-carbon long lipid chain. The interior of the Ste24 protein is a large chamber capable of accommodating a 15 kDa protein (28) . The amino acid residues lining the inside of the chamber are of mixed hydrophobicity with a high propensity for charged residues (28). A large hydrocarbon attached at the residue of the C-terminal CaaX cleavage must have a location within the chamber to accommodate such a large hydrophobic attachment. A co-crystallized structure of ZMPSTE24 and the corresponding, non-farnesylated, prelamin A CaaX sequence (CSIM) has been published and does help to display that the CaaX peptide does appear to orient around the zinc residue, indicating

the location of the active site (29). However, this co-crystallization was performed utilizing the tetra-peptide without the attachment of a farnesyl group. The authors note that the peptide is aligned at the incorrect peptide bond around the zinc ion in the active site, orienting between the S/I site rather than the C/S site (29). This misalignment could be due to the small size of the peptide or could be due to the lack of the farnesyl group. Determining where this lipid group is located during the -aaX cleavage should help to predict the alignment and orientation of the substrate peptide during CaaX proteolysis.

Determining the location of the lipid group can be difficult in typical structural techniques, such as crystallization and electron microscopy, due to the highly flexible characteristics of the farnesyl attachment. Because of this, any crystal structures determined with a substrate in mammalian ZMPSTE24 or yeast Ste24 have lacked the farnesyl group. One technique that could be quite useful, and has quickly become a useful technique in proteomics, is the use photoaffinity labeling (PAL) (131,146,147). The probes for PAL consist of three things: 1) an affinity group based on the native substrate to properly bind to the protein of interest, 2) a photoactivatable group capable of forming radical compounds to create covalent bonds with the protein of interest upon irradiation, and 3) a handle to allow for enrichment of the photolabeled protein-substrate complex. The sample is subsequently digested and run on a tandem LC-MS/MS instrument, followed by software analysis to determine the precise location of the covalent bond formed between the photoactivatable probe and the protein of interest. This technique has become very useful in determining the sites of interaction between proteins and either other proteins, peptides, or small molecules (131,146,147).

Herein, I utilized PAL with a photoactivatable version of **a**-factor, here termed C_{10}-para-**a**-factor designed by collaborators at the University of Minnesota. The affinity group is a 15-amino acid sequence with the photoactivatable group, benzophenone, attached in a C_{10}-para fashion within the farnesyl moiety, and the enrichment handle is a biotin tag attached to the N-terminal end by a PEG4 linker (**Figure 2.1B**). The purpose of this study was to identify the site of the covalent linker formed during the photolabeling reaction with WT Ste24. Locating the attachment site of the photoactivatable, modified farnesyl group would provide insight into the location of the farnesyl group during the -aaXing cleavage reaction.

3.2 Methods and Materials

3.2.1 Plasmids and yeast strains

The pCH1283 plasmid (**Table 2.2**) contains a His_{10}-HA_3 tag on the N-terminus of the *ste24* gene (*2µ URA3 P_{PGK}-His_{10}-HA_3-Ste24*). The plasmid was transformed into the yeast strain SM3614, containing a double deletion for endogenous Rce1 and Ste24 (*MATa trp1 leu2 ura3 his4 can1 ste24Δ::LEU2 rce1Δ::TRP1*) (44).

3.2.2 Crude membrane preparation

Crude membrane preparation was started by first inoculating a small, synthetic complete supplement mixture without uracil (SC-URA) culture with the pCH1283 (WT) Ste24 strain that was grown overnight at 30 ⁰C overnight. Then 15 mL was used to inoculate a larger, 1 L SC-URA culture and was grown to log phase (OD_{600} = 300-500) then harvested at 4000 xg. Lysis of the cells was performed using yeast sorbitol lysis buffer (0.3 M Sorbitol, 0.1 M NaCl, 12 mM $MgCl_2$, 1% aprotinin, 3 mM AEBSF, 1 mM DTT, 10 mM Tris-HCl, pH 7.5) at a ratio of 1 mL lysis buffer to 800 OD_{600} cells. The lysis buffer was added to the cell pellet in a 50 mL conical tube and then vortexed and left on ice for 15 minutes to allow the cells to swell. The suspension was then flash frozen in liquid nitrogen and thawed in a 29 ⁰C water bath, twice. Then a french press was used to further lyse and break up the membrane fractions using 18,000 psi twice. The resulting suspension was first centrifuged twice at 500 xg for 10 minutes to remove cellular debris then centrifuged at 100,000 xg for 1 hour at 4 ⁰C. After the final centrifugation step, the supernatant was removed, and the membrane pellet was resuspended in 10 mM Tris-HCl, pH 7.5. Protein concentration was determined using a coomassie protein assay using a known concentration of bovine serum albumin (BSA) as the standard curve (137).

3.2.3 Membrane protein purification

To purify the wild-type Ste24 from the crude membrane samples, the protein was first solubilized in buffer A (0.3 M sorbitol, 0.1 M NaCl, 6 mM $MgCl_2$, 10 mM Tris, pH 7.5, 10% glycerol, 1% aprotinin, 2 mM AEBSF) supplemented with 20 mM imidazole and 1% n-Dodecyl-B-D-maltopyranoside (DDM, Anatrace) with rocking at 4 ⁰C for 1 hour. The suspension was then centrifuged at 100,000 xg for 45 minutes to remove the insoluble fraction, and the supernatant was

added to Talon® Metal Affinity Resin (Clontech). The resin mixture was rocked at 4 ⁰C for 1 hour, then washed with five column volumes of buffer B (40 mM imidazole and 1% DDM in buffer A) twice, buffer C (40 mM Imidazole, 1% DDM and 0.5 M KCl in buffer A) once, and buffer D (40 mM imidazole, 0.1% DDM and 0.5 M KCl in buffer A) once. The protein was then eluted from the resin with buffer E (250 mM imidazole, 0.1% DDM in buffer A). The elution was then concentrated using Amicon® Ultra Centrifugal Filter 30,000 MWCO (Millipore) at 4,000 xg for 20-30 minutes at 4 ⁰C (until desired volume was reached, about 250 µL). Protein concentration was determined using the amido black protein assay with a known concentration of BSA for a standard curve (138).

3.2.4 Photolabeling of purified Ste24 for use in tandem ESI-MS-MS

Photolabeling of purified Ste24 was performed similar to that in **Section 2.2.6** except without the use of lipids to aid in sample compatibility with downstream ESI-MS-MS experimentation. First, 20 µg of purified wild-type Ste24 enzyme was incubated with 50 µM C_{10}-para-**a**-factor in 100 mM Tris-HCl pH 7.5 and 1 mM DTT. The sample was then irradiated with UV light (365 nm) for 30 minutes on ice. The sample was dried with a vacuum centrifuge and the pellet was washed three times with cold 80% acetone to remove contaminants. Residual acetone was removed at 65 °C until dried, about 10 minutes. The samples were then used for reduction, alkylation and digestion as described in **Section 3.2.5**.

3.2.5 Reduction, alkylation, and enzymatic digestion of photolabeled Ste24

Following photolabeling, the sample was prepared for ESI-MS-MS by reduction, alkylation, and digestion. First, the sample pellet was incubated with 10 µL of 10 mM DTT in 8 M urea at 700 rpm and 37 °C for 1 hour to reduce all disulfide bonds. Then the free thiols were alkylated by addition of 10 µL of the alkylation solution (0.5% TEP and 2% IEtOH in acetonitrile) and at 37 °C and 700 rpm. The sample was then split into two tubes, one for each digestion enzyme, and dried in a vacuum centrifuge to remove reagents from the alkylated sample.

The pellet was dissolved in 20 µL of 25 mM ammonium bicarbonate (ABC) followed by addition of 50 µL of a Trypsin/Lys-C digestion mixture or Chymotrypsin (2.5 µg total digestion enzyme). The samples were digested first in a barocylcer (50 sec at 20 kpsi, then 10 sec at atm, 60

rounds, 55 °C). Then another 5 µL of each digestion enzyme was added and the samples were left to digest overnight at 37 °C and 700 rpm. Digested peptides were isolated by use of UltraMicroSpin C18 columns (3-30 µg capacity, BIOZ) according to manufacturer protocols. The samples were dried again in a vacuum centrifuge and submitted to the Purdue Proteomics Facility.

3.2.6 Tandem MSMS of digested samples

Samples were analyzed in the Dionex UltiMate 3000 RSLC nano System (Thermo Fisher Scientific, Waltham, MA, USA) coupled to the Q-Enactive High-Field (HF) Hybrid Quadrupole Orbitrap MS (Thermo Fisher Scientific, Waltham, MA, USA) as described previously (148-150). Peptides were re-suspended in 3% ACN/0.1% Formic Acid (FA)/96.9% MilliQ, and about 1µg (equivalent volume) was used for LC-MS/MS analysis. Peptides were loaded into the trap column (300 µm ID × 5 mm) packed with 5 µm 100 Å PepMap C18 medium using a flow rate of 5 µL/min with 98% purified water/ 2% ACN/0.1% FA and washed for 5 min. Peptides were then separated using a reverse-phase Acclaim PepMap analytical column (Thermo Fisher Scientific, Waltham, MA, USA) packed with 2 µm of 100 Å C18 medium (75 µm × 50 cm) using a 130-min method at a flow rate of 150 nL/min. The mobile phase A consisted of 0.1% FA in water and mobile phase B consisted of 0.1% FA in 80% ACN. The linear gradient started at 2% B, switched to 8% B in 5.1 min, and reached to 27% B in 80 min, 45% B in 100 min, and 100% B in 105 min. At this point, concentration of B was held constant at 100% B for 7 min before reverting to 2% B in 112 min and maintained at 2% B until the end of the run. The column temperature was maintained at 50°C. MS was operated at positive ion mode with default charge of 2+ and data were collected with a Top 20 data-dependent acquisition method with a maximum injection time of 100 ms, a resolution of 120 000 at 200 m/z, AGC of 3×10^6, scan range of 350-1600 m/z for MS1. Fragmentation of precursor ions was performed by high-energy C-trap dissociation (HCD) with the normalized collision energy of 27 eV. MS/MS scans were acquired at a resolution of 15 000 at m/z 200, AGC of 1×10^5, maximum injection time of 20 ms. The dynamic exclusion was set at 15 s to avoid repeated scanning of identical peptides, and chromatographic peak width was maintained at 15 s.

3.2.7 Analysis of MSMS .RAW data for photolabeled Ste24

The .RAW files for the photolabeled and unlabeled samples were loaded into PEAKS Studio X Pro and were set for a search allowing for: 1) a mass tolerance of 0.02 Da, 2) a specific enzyme digestion with trypsin or chymotrypsin allowing for up to 2 missed cleavages. Different PTM conditions were explored, depending on the sample prep: 1) Raney nickel addition of C_{10}-para (+348.21), 2) full C_{10}-para probe (+2512.3), 3) trypsinized C_{10}-para (+1522.76), 4) -aaXed C_{10}- para (+2244.13), 5) trypsinized and -aaXed C_{10}-para (+1239.57), 6) chymotrypsinized C_{10}-para (+1033.52), and 7) chymotrypsinized and aaXed C_{10}-para (+750.33) and all of these also with the loss of water (+ 330.19, +2494.28, +1504.74, +2226.11, +1221.55, +1015.5 and +732.31 respectively). A fixed modification of carbamidomethylation on cysteine residues (+57.02) was also set due to the alkylation step with iodoethanol. The database used was a .fasta file containing the N-terminal His_{10}-HA_3-Ste24 protein, 15-mer **a**-factor peptide sequence, trypsin, and Lys-C protein sequences. Label-free quantification was performed between the two samples.

3.2.8 Structural analysis of Ste24

Structural analysis was performed by utilizing The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. and analyzing the *smSte24* (PDB: 4IL3). The zinc ion was colored in orange. Residues were selected as they became of interest and colors were modified from there. Further analysis will be discussed in the legends of the figures.

3.2.9 Photolabeling efficiency assay

To determine the efficiency of the photolabeling reaction, the samples were set up in a similar fashion as **Section 2.2.6**. The wild-type Ste24 protein was reconstituted in *E. coli* polar lipid extract (Avanti) to a final reaction concentration of 100 mM Tris-HCl pH 7.5, 1 mM DTT, 50 µM C_{10}-para, 0.5 µg WT Ste24 and 50 µg lipid. Then 60 µL was initially taken to serve as the time = 0 min control and loaded onto 50 µL of a 1:1 slurry of NeutrAvidin® Agarose Resin (ThermoScientific) in RIPA/SDS buffer (1% SDS, 1% Sodium deoxycholate, 1% Triton X, 150 mM NaCl, 25 mM Tris-HCl pH 7.5) supplemented with 1% aprotinin, 2 mM AEBSF, 50 U/mL micrococcal nuclease, 1 mM $CaCl_2$ and 1 mM DTT (RIPA/SDS/I). Then the stock reaction was aliquoted into seven different wells in a 96-well plate and irradiated with UV light (365 nm) and

aliquots were taken at different time points (0 min, 1 min, 5 min, 10 min, 20 min, 30 min, and 40 min) and loaded directly onto NeutrAvidin® Agarose Resin in RIPA/SDS/I buffer. Additional RIPA/SDS/I buffer was added to improve turnover and the samples were rocked for 2 hours at 4 °C. Then the supernatant was collected from each time point and precipitated with TCA (final concentration 20% TCA). The precipitated protein was then redissolved into 200 µL 2x SDS. The resin was washed once with fresh RIPA/SDS/I and then resuspended in 50 µL 2x SDS. All SDS samples were then heated at 65 °C for 30 minutes to fully denature the protein.

Then 10 µL of supernatant and resin samples were run on separate 10% SDS-PAGE gels followed by transfer to a 0.2 µm nitrocellulose membrane. The membranes were blocked overnight in 20% milk in PBST (1x PBS buffer, 0.1% Tween-20) at 4 ⁰C, followed by 2 hours in primary antibody (mouse, α-HA at 1:15,000) in 4 % milk in PBST, then 1 hour in secondary antibody (goat-α-mouse-HRP, 1:4,000) in 4 % milk in PBST. The bands were then visualized on a GeneGnome XRQ (SynGene) using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific).

3.2.10 Enrichment of photolabeled Ste24 for tandem MSMS

Multiple aliquots of WT Ste24 were photolabeled simultaneously (20 µg each, 260 µg total) according to the procedure in **Section 3.2.3**. The samples were crosslinked under UV light (365 nm) for 30 minutes, and then 60 µL to 50 µL of a 1:1 slurry of NeutrAvidin® Agarose Resin (ThermoScientific) in 800 µL 1x PBS. The samples were left rocking for two hours at 4 ⁰C. The supernatant for all aliquots were combined and added to one reaction flask to serve as a flowthrough, or unlabeled, control. The resin was then pelleted at 10,000 xg for 2 minutes and washed three times with 800 µL fresh PBS buffer. Residual supernatant was removed and all the resin aliquots were combined and added to a second reaction flask.

Then to each reaction flask, about 200 mg of the thioether-hydrogenation catalyst Raney®-Nickel slurry (Sigma Aldrich) was added along with a stir bar. The reaction was set up under a H_2 atmosphere overnight at room temperature. In the morning, imidazole was added to a final concentration ~1 M and incubated for 30 minutes to release the N-terminal His-tag of the Ste24 protein from the nickel in the catalyst. The sample was then run through a gravity column, along with addition of another 1 mL of 1 M imidazole. The sample was concentrated slightly through a Amicon® Ultra Centrifugal Filter 30,000 MWCO (Millipore) at 4,000 xg to remove the imidazole,

then dried in a vacuum centrifuge followed by reduction, alkylation, and digestion, as described in **Section 3.2.5** with just a trypsin/Lys-C digestion.

3.3   Results and Discussion

3.3.1   Near full sequence coverage of Ste24 was obtained using parallel trypsin/Lys-C and chymotrypsin digestions

Obtaining nearly full sequence coverage of proteins, especially with transmembrane proteins, can be quite difficult. To be detected by the ESI-MS/MS instrument, the digested peptides must be charged and at a sufficient mass to charge ratio to be detected within the range of the instrument. Transmembrane proteins tend to have very long peptide regions that are highly hydrophobic. The lack of charged residues makes it difficult for these peptides to ionize properly and can make obtaining multiple charge states quite problematic. As these sections are usually quite long and cannot obtain a high charge state, many of the peptides tend to be outside the m/z range of detection of the instrument.

Two of the most utilized proteases for the digestion of proteins for tandem MS/MS are trypsin and chymotrypsin. Trypsin digests proteins on the C-terminal side of arginine and lysine residues (except when next to a proline). Trypsin is very common as a digestion enzyme as it ensures each peptide has at least one positive charge. The only problem seen with using this protease in the Ste24 digest is that the transmembrane domains have few lysine and arginine residues, making these hydrophobic peptides very large and likely outside of the detection limit unless a multiple charge state is obtained. Lys-C is a protease often used in tandem with trypsin as it also cleaves on the C-terminal side of lysine but is not restricted by the presence of proline.

Chymotrypsin is also very common, especially for transmembrane proteins. The cut specificity is not as restricted as trypsin, as it cuts preferentially on the C-terminal side of large hydrophobic residues (tryptophan, tyrosine, and phenylalanine) and can also cleave on the C-terminal side of leucine and methionine at a slower rate. Ste24 does contain many more cut sites for chymotrypsin than trypsin, but there were a couple cons for this protease selection. The first is that the specificity of the protease does not ensure a positively charged residue on the peptide. The other is that the transmembrane regions have so many chymotrypsin cut sites that the problem is on the opposite end of the spectrum as trypsin in that the peptides are too short and below the size detection limit of the instrument.

Often, the best way to obtain the highest sequence coverage of a protein is to utilize several digestion enzymes. Initial experiments began with 10 µg of Ste24 protein run on a 10% SDS-PAGE gel. The gel was stained with coomassie stain and the pure Ste24 bands were cut out of the gel and destained using a 1:1 solution of 25 mM ammonium bicarbonate (ABC) : 100% acetonitrile (ACN). The gel pieces were then dried in 100% ACN followed by running the tubes in a vacuum centrifuge. Reduction of the cysteine residues was performed at 55 °C for one hour by adding enough 10 mM DTT in 25 mM ABC to fully hydrate and cover the gel slices. Then the DTT solution was removed and 55 mM iodoacetamide in 25 mM ABC was added to alkylate the free thiol groups for 45 minutes at 55 °C in the dark. The gel pieces were then washed with 100 µL of 25 mM ABC then dried again with a 1:1 mixture of 25 mM ABC : 100% ACN followed by addition of 100% ACN and then vacuum centrifugation to remove all the liquid. Then 1.5 µg of either trypsin/Lys-C enzyme cocktail or chymotrypsin was added and allowed to absorb into the gel slices. Then additional 25 mM ABC was added to cover the gel and the digestions were run in a barocylcer at 50 °C for 60 cycles of 20 kpsi for 50 seconds and 10 seconds at atmospheric pressure. The peptides were then extracted by sonicating the gel pieces in 60% ACN/5% TFA twice. The supernatant from both sonication steps was combined and then dried in a vacuum centrifuge and submitted to the Purdue Proteomics Facility to run the sample. The raw data was used to check the sequence coverage using Mascot MS/MS Ion Query (matrixscience.com). The in-gel digestion resulted in a sequence coverage of 44% from the trypsin/Lys-C digestion and 53% coverage by the chymotrypsin digestion. The majority of the TMD 2 and 3 sequence, as well as much of the other transmembrane regions were not identified in the sample. It is likely the extraction of the peptides out of the SDS-PAGE gel was performed with a solution that was too polar to sufficiently dissolve the hydrophobic sections of Ste24, as the majority of the identified peptides were the charged, water exposed sections of Ste24.

Then, an overnight in-solution digestion for both chymotrypsin and trypsin/Lys-C digestion was attempted. For initial in-solution digestion, 10 µg of purified wild-type Ste24 was precipitated using 4x volume of -20 °C acetone overnight. In the morning the supernatant was removed by centrifugation followed by heating at 65 °C for 10 minutes to remove any residual acetone. The protein pellet was dissolved in a solution of 10 mM DTT in 8 M urea. The DTT was used to reduce the cysteine residues while the 8 M urea aids in protein denaturation. The reduction reaction was incubated for one hour in the dark at 700 rpm and 37 °C. Then 10 µL of an alkylating

solution (2% Iodoethanol, IEtOH, 0.5% triethylphosphine, TEP, in ACN) and incubated for another hour at 700 rpm and 37 °C. The samples were then dried in a vacuum centrifuge. Then 4 µg of either a trypsin/Lys-C enzyme cocktail or chymotrypsin enzyme dissolved in 25 mM ABC were added to the protein sample and the digests were first run in a barocylcer at 50 °C for 60 cycles at 20 kpsi for 50 sec and atmospheric pressure for 10 seconds to aid in protein denaturation. An additional 0.5 µg of each digestion enzyme was then added to the respective samples and the digestion continued overnight at 37 °C. The peptides were extracted using an UltraMicroSpin C18 Column (3-30 µg capacity) according to manufacturer instructions. The trypsin/Lys-C digestion resulted in ~75% coverage (**Figure 3.1A**) while the chymotrypsin digestion resulted in ~80% (**Figure 3.1B**). Taken together, both digestions resulted in about 93% sequence coverage of Ste24 (**Figure 3.1C**). The topology diagram in **Figure 3.2** displays the regions that were not identified by either digestion sample. Very small regions of the hydrophobic transmembrane regions are still unidentified. The short sequence in TMD 2 may provide digested peptides that are too small to be separated on the Q-Exactive C-trap instrument. The sections in the TMD 3, 4, and 5 that are unidentified are a part of long very hydrophobic sequences. Digestion with chymotrypsin would not allow for the presence of many charged residues while digestion with trypsin/Lys-C would likely create peptides that are too large for the Q-Exactive C-trap instrument. The two residues at the end of the Ste24 protein follow three lysine residues that were likely cleaved by the trypsin enzyme and resulted in a very short dipeptide. While full sequence coverage of Ste24 was not obtained, 93% coverage is very high and it was very likely it would be sufficient for this project. The goal of this project was to identify the direct site of photolabeling by C_{10}-para-**a**-factor and with a nearly complete sequence coverage it was expected that this interaction would be within the identified region.

A **Trypsin/Lys-C 75% coverage**

MHHHHHHHHHHHYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAAQC
GRDLKTILDHPNIPWKLIISGFSIAQFSFESYLTYRQYQKLSETKLPPV
LEDEIDDETFHKSRNYSRAKAKFSIFGDVYNLAQKLVFIKYDLFPKI
WHMAVSLLNAVLPVRFHMVSTVAQSLCFLGLLSSLSTLVDLPLSYYS
HFVLEEKFGFNKLTVQLWITDMIKSLTLAYAIGGPILYLFLKIFDKFPT
DFLWYIMVFLFVVQILAMTIIPVFIMPMFNKFTPLEDGELKKSIESLA
DRVGFPLDKIFVIDGSKRSSHSNAYFTGLPFTSKRIVLFDTLVNSNST
DEITAVLAHEIGHWQKNHIVNMVIFSQLHTFLIFSLFTSIYRNTSFYNT
FGFFLEKSTGSFVDPVITKEFPIIIGFMLFNDLLTPLECAMQFVMSLIS
RTHEYQADAYAKKLGYKQNLCRALIDLQIKNLSTMNVDPLYSSYHY
SHPTLAERLTALDYVSEKKKN

B **Chymotrypsin 80% coverage**

MHHHHHHHHHHHYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAAQC
GRDLKTILDHPNIPWKLIISGFSIAQFSFESYLTYRQYQKLSETKLPPV
LEDEIDDETFHKSRNYSRAKAKFSIFGDVYNLAQKLVFIKYDLFPKI
WHMAVSLLNAVLPVRFHMVSTVAQSLCFLGLLSSLSTLVDLPLSYYS
HFVLEEKFGFNKLTVQLWITDMIKSLTLAYAIGGPILYLFLKIFDKFPT
DFLWYIMVFLFVVQILAMTIIPVFIMPMFNKFTPLEDGELKKSIESLA
DRVGFPLDKIFVIDGSKRSSHSNAYFTGLPFTSKRIVLFDTLVNSNST
DEITAVLAHEIGHWQKNHIVNMVIFSQLHTFLIFSLFTSIYRNTSFYNT
FGFFLEKSTGSFVDPVITKEFPIIIGFMLFNDLLTPLECAMQFVMSLIS
RTHEYQADAYAKKLGYKQNLCRALIDLQIKNLSTMNVDPLYSSYHY
SHPTLAERLTALDYVSEKKKN

C **Combined 93.1% coverage**

MHHHHHHHHHHHYPYDVPDYAGYPYDVPDYAGSYPYDVPDYAAQC
GRDLKTILDHPNIPWKLIISGFSIAQFSFESYLTYRQYQKLSETKLPPV
LEDEIDDETFHKSRNYSRAKAKFSIFGDVYNLAQKLVFIKYDLFPKI
WHMAVSLLNAVLPVRFHMVSTVAQSLCFLGLLSSLSTLVDLPLSYYS
HFVLEEKFGFNKLTVQLWITDMIKSLTLAYAIGGPILYLFLKIFDKFPT
DFLWYIMVFLFVVQILAMTIIPVFIMPMFNKFTPLEDGELKKSIESLA
DRVGFPLDKIFVIDGSKRSSHSNAYFTGLPFTSKRIVLFDTLVNSNST
DEITAVLAHEIGHWQKNHIVNMVIFSQLHTFLIFSLFTSIYRNTSFYNT
FGFFLEKSTGSFVDPVITKEFPIIIGFMLFNDLLTPLECAMQFVMSLIS
RTHEYQADAYAKKLGYKQNLCRALIDLQIKNLSTMNVDPLYSSYHY
SHPTLAERLTALDYVSEKKKN

Figure 3.1: A 93% sequence coverage of Ste24 was obtained by in-solution digest with dual digestion reactions with trypsin/Lys-C and chymotrypsin. Purified wild-type Ste24 (20 µg) was reduced DTT, alkylated with iodoethanol, and then digested with a 5:1 ratio of Ste24 to either trypsin/Lys-C or chymotrypsin in a barocylcer followed by an overnight digestion at 37 °C. N-terminal tags are displayed in blue and all identified amino acid sequences are in red. A) trypsin/Lys-C digestion resulted in a sequence coverage of 75% by using MS/MS Ion Search from Mascot Server. B) Chymotrypsin digestion resulted in a sequence coverage of 80% by using MS/MS Ion Search from Mascot Server. C) Combined results from both digestions resulted in a sequence coverage of 93% by using MS/MS Ion Search from Mascot Server.
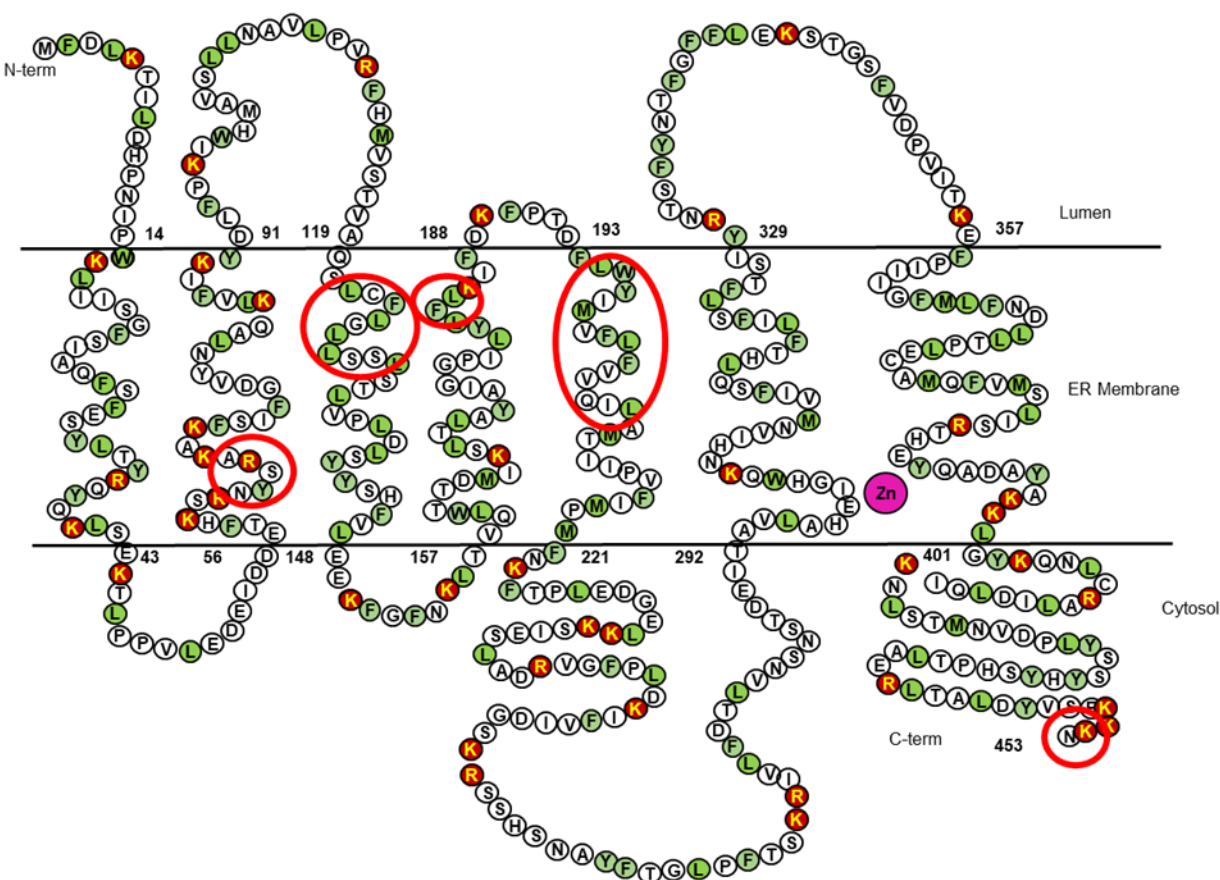
Figure 3.2 Topology map of Ste24 displaying digestion sites and holes in sequence coverage from dual digestion with trypsin/Lys-C and chymotrypsin. 20 µg of purified wild-type Ste24 was reduced with DTT, alkylated with iodoethanol, and then digested with either trypsin/Lys-C or chymotrypsin. Combined results covered 93% of the Ste24 sequence. Red circles highlight areas that were not detected in either digestion sample. The red K and R residues display the location of trypsin/Lys-C digestion sites while the green F, L, W,Y, and M residues display the location of chymotrypsin digestion sites.

3.3.2 Peptides were identified that contained the site of photolabeling using the C_{10}-para-**a**-factor probe

Our next goal was to determine the site of photolabeling by the C_{10}-para-**a**-factor probe (**Figure 2.1B**). Wild-type Ste24 (20 µg) was incubated with the C_{10}-para-**a**-factor substrate and then split into two tubes. One group was kept in the dark as an unlabeled sample and the other was irradiated with UV light (365 nm) for 30 minutes to photolabel the protein. Both samples were reduced with DTT, alkylated with iodoethanol, then split in half. One half was digested with trypsin/Lys-C and the other with chymotrypsin. Then the digested peptides were combined and

isolated with C_{18} UltraSpin columns (3-30 µg) according to manufacturer instructions and run on a nanoESI- Q-Exactive High-Field (HF) Hybrid Quadrupole Orbitrap MS.

The mass chromatogram for both the photolabeled and unlabeled samples are displayed in the combined graph in **Figure 3.3A**. As shown, there is very little variability between the two groups. There are slight differences at time points around 31 minutes and 43 minutes, but nothing initially stands out as a different peptide in the photolabeled sample. Because there were no changes easily identified by eye, PEAKS Studio X Pro was used to search for PTM modifications from the various C_{10}-para additions (full probe, digestion proteolysis only, -aaXed only, both digestion proteolysis and -aaXed). An additional search for the same modifications with a -H_2O loss was also used, as the benzophenone reaction can often result in the loss of water. The photolabeled sample resulted in 61% sequence coverage and the unlabeled sample had 71% sequence coverage. Of all the Ste24 peptides identified, 113 were found in both samples, 22 in the unlabeled alone and 33 in the photolabeled alone (**Figure 3.3B**).

The PTM search revealed several sites that could potentially contain the photolabeling site of the trypsinized and -aaXed, minus water, addition of the C_{10}-para-**a**-factor probe. As shown in **Figure 3.4A**, the only modification identified with confidence from PEAKS Studio was the addition of the C_{10}-para at the K278 location. This 278 position is the position on the N-terminal His_{10}-HA_3-tagged Ste24 protein and corresponds to the lysine 234 (K234) residue in the native protein. This modification was identified exclusively in the photolabeled sample, and the annotated MS2 spectrum is shown in **Figure 3.5A**. It is important to note that the spectra shown is the data refined spectra and the analysis is based off an average of surrounding ions in the spectra, not just of the one scan. Therefore, ions such as the labeled Y15+ ion was not visible in the one spectra but was located in surrounding MS2 spectra. This modification was identified when the default FDR setting was used (all peptides included).

A

Annotated Chromatogram of Protein Ste24

Intensity

F 1

37.71 50.11 59.27

28.60   43.03   55.50

Photolabeled

1E9

63.55

49.26   64.44

5E8

67.23

83.02   93.15

70.83

0

67.30   83.04   96.34

F 2

85.83

5E8

63.51

31.62

50.12   59.55

Unlabeled

37.88

43.11

1E9

28.60

RT

10   20   30   40   50   60   70   80   90   100   110   120   130

B

Photolabeled   Unreacted

33   113   22
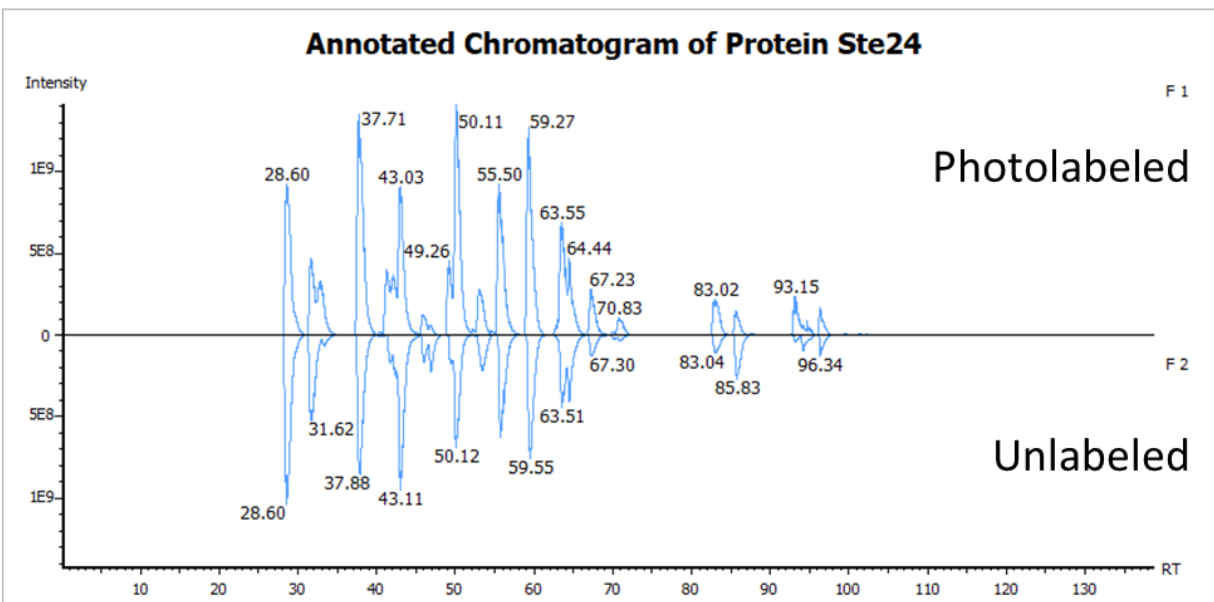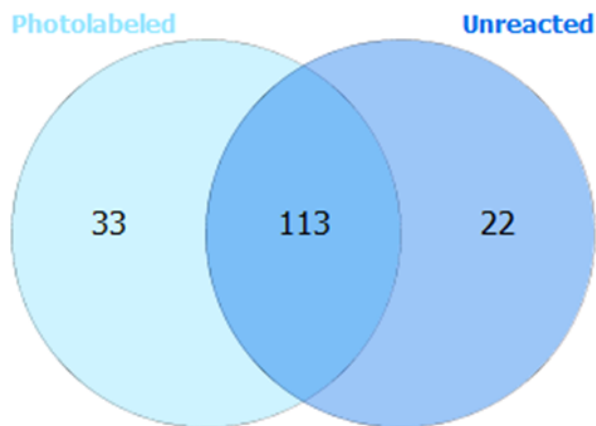
Figure 3.3: Comparison of the MS chromatogram between the photolabeled and unlabeled samples offers little into insight of newly labeled species. A) The refined chromatograms of the photolabeled sample (top) and unlabeled sample (bottom) graphed as retention time versus ion intensity. B) Venn Diagram prepared by PEAKS Studio X Pro showing the number of peptides found in either or both samples.

Figure 3.4: Statistical determination of the location of the photolabeling site by PEAKS Studio analysis. A) By utilizing PEAKS Studio and searching for possible PTM modifications of C_{10}-para-addition, and by comparing between both a photolabeled and sample, the PEAKs studio software obtained 71% sequence coverage in the photolabeled sample and 74% sequence coverage in the unlabeled sample (default FDR settings used, utilizing all identified peptides). Blue bands indicate spectra that were statistically matched to the Ste24 protein while gray bars indicate De Novo sequencing peptides that also obtained the same sequence as the Ste24 protein. The carbamidomethylation was a fixed modification of cysteine residues resulting from sample workup, adding an alkyl group with iodoethanol, and the orange C boxes indicate possible sites of C_{10}-para addition where the probe had been trypsinized and cleaved of the –aaX residues. Location K278 here was the only location scored with a "high confidence" score, where the modification site confidence cutoff value was set at 5% minimal ion intensity (default). It should be noted that K278 in this Fasta sequence correlates to K234 in the native protein, that does not contain the His_{10} and HA_3 N-terminal tags and this labeling site was identified only in the photolabeled sample. B) By reducing the FDR to 1%, the sequence coverage by the photolabeled sample was reduced to 61% and the unlabeled to 71%. Only one possible site of photolabeling remains at position E401 (E357 in the native protein). This modification was found only in the unlabeled protein sample.

A

```
1    MHHHHHHHHH HYPYDVPDYA GYPYDVPDYA GSYPYDVPDY AAQCGRDLKT ILDHPNIPWK LIISGFSIAQ FSFESYLTYR

81   QYQKLSETKL PPVLEDEIDD ETFHKSRNYS RAKAKFSIFG DVYNLAQKLV FIKYDLFPKI WHMAVSLLNA VLPVRFHMVS

161  TVAQSLCFLG LLSSLSTLVD LPLSYYSHFV LEEKFGFNKL TVQLWITDMI KSLTLAYAIG GPILYLFLKI FDKFPTDFLW

             K278                                    278
                                                      C
241  YIMVFLFVVQ ILAMTIIPVF IMPMFNKFTP LEDGELKKSI ESLADRVGFP LDKIFVIDGS KRSSHSNAYF TGLPFTSKRI
                                              C

321  VLFDTLVNSN STDEITAVLA HEIGHWQKNH IVNMVIFSQL HTFLIFSLFT SIYRNTSFYN TFGFFLEKST GSFVDPVITK
                                                                  C

401  EFPIIIGFML FNDLLTPLEC AMQFVMSLIS RTHEYQADAY AKKLGYKQNL CRALIDLQIK NLSTMNVDPL YSSYHYSHPT
     C                   C                              C

481  LAERLTALDY VSEKKKN
```

C Carbamidomethylation (+57.02)
C C10 tryp and aaX -H2O (+1221.54)

B

```
1    MHHHHHHHHH HYPYDVPDYA GYPYDVPDYA GSYPYDVPDY AAQCGRDLKT ILDHPNIPWK LIISGFSIAQ FSFESYLTYR

81   QYQKLSETKL PPVLEDEIDD ETFHKSRNYS RAKAKFSIFG DVYNLAQKLV FIKYDLFPKI WHMAVSLLNA VLPVRFHMVS

161  TVAQSLCFLG LLSSLSTLVD LPLSYYSHFV LEEKFGFNKL TVQLWITDMI KSLTLAYAIG GPILYLFLKI FDKFPTDFLW

241  YIMVFLFVVQ ILAMTIIPVF IMPMFNKFTP LEDGELKKSI ESLADRVGFP LDKIFVIDGS KRSSHSNAYF TGLPFTSKRI

321  VLFDTLVNSN STDEITAVLA HEIGHWQKNH IVNMVIFSQL HTFLIFSLFT SIYRNTSFYN TFGFFLEKST GSFVDPVITK

401  EFPIIIGFML FNDLLTPLEC AMQFVMSLIS RTHEYQADAY AKKLGYKQNL CRALIDLQIK NLSTMNVDPL YSSYHYSHPT
     C                   C
      E401

481  LAERLTALDY VSEKKKN
```

A more well accepted FDR rate is 1% and when this was applied the only C_{10}-para addition was identified at the glutamate 401 (E401) residue in the N-terminally tagged Ste24, corresponding to the glutamate 357 (E357) residue in the native protein that does not contain the N-terminal tags. This modification was found with a lack of confidence in exclusively the unlabeled sample (**Figure 3.4B**). The total sequence coverage was also reduced slightly to 61% in the photolabeled sample and 71% in the unlabeled sample. The annotated MS2 spectra found with this peptide is shown in **Figure 3.5B** and contains more of the predicted y ion series of the Ste24 peptide. It is possible that there is some background photolabeling in the absence of UV light and that this happened to appear in this unlabeled sample.

The MS2 spectra of both identified photolabeled spectra shown in **Figure 3.5** do lack any b ions of the C_{10}-para probe. It would be expected that one could see the b ion series of the probe to identify that it is also present. Most ions in both spectra are low in intensity, so it could also be possible that the intensity is just too low to be detected. Also, the data in the MS2 scan had undergone data refinement under default settings in the mass spectrometry search program and displays a spectrum averaging all nearby MS2 scans, causing some of the peaks in the spectrum to be omitted. Likely, the majority of the positive charge lies on the Ste24 peptide due to the higher frequency of chargeable amino acids. Most fragment ions in the MS2 spectrum are detectable due to the transferable protons down the peptide backbone. However, because of the large hydrophobic group connecting the Ste24 peptide and probe peptide, through the photoactivatable benzophenone group, it is likely the protons cannot be transferred to the probe peptide sequence. Another problem with the spectra in **Figure 3.5** was that most MS2 spectra display a y ion series on the right-hand side of the precursor ion, towards the higher m/z axis. In the spectra in **Figure 3.5** all the y ions identified were on the lower m/z scale than the precursor ion. While this does not disprove the Ste24 peptide is present, it does provide a cause for concern in the data interpretation.
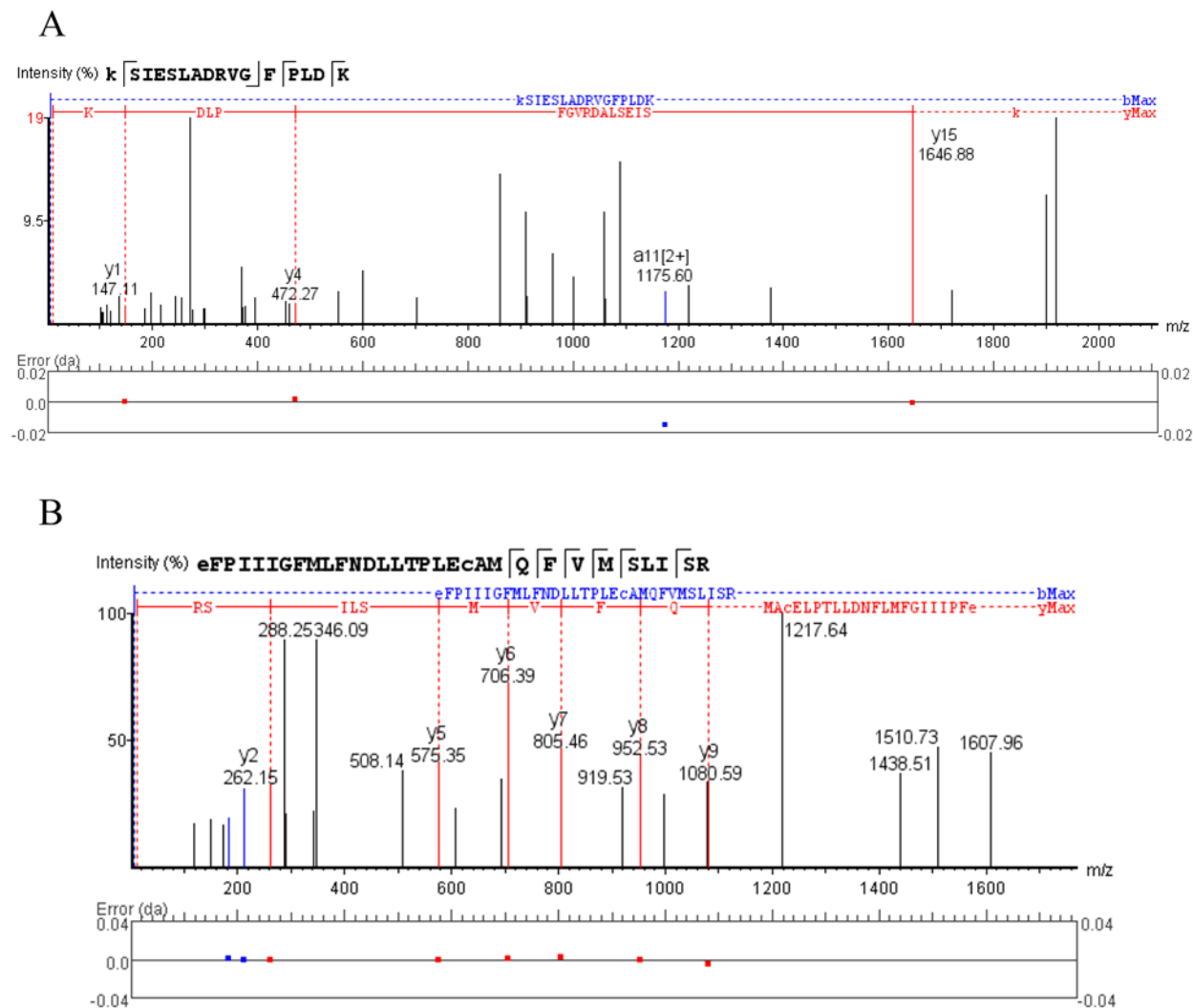
Figure 3.5: Fragmentation ion assignment by PEAKS Studio recognizes a few ions within the assigned MS2 spectra. Diagram created by PEAKS Studio and assignment of peaks was performed by default settings with the addition of a-ion annotations and after data refinement (spectral averaging). A) MS2 spectrum annotated for the modification found in the photolabeled sample at position K278 (K234 in native protein). The mass error for each assigned peak is in the graph below the spectrum. B) MS2 spectrum annotated for the modification found in the unlabeled sample at position E401 (E357 in native protein). The mass error for each assigned peak is in the graph below the spectrum.

The K234 residue is of particular interest. As shown in **Figure 3.6A** and **Figure 3.6B**, the K234 residue lies at the bottom of a very large, hydrophobic cavity inside the Ste24 protein. C_{10}-para-**a**-factor typically labels the α-carbons of amino acids, so with the K234 side chain facing outward, this puts the α-carbon facing into this hydrophobic pocket. The α-carbon is approximately 21 Å from the catalytic zinc enzyme, making it close enough to be labeled by the benzophenone functional group that was about 18 Å from the cysteine. This hydrophobic pocket also lies just below the portal in the side of the protein. The portals within the Ste24 chamber were discovered when the crystal structure was solved and present a possible entry and exit site for the substrate (28,29). There were several possible portals identified, but the largest lies between TMD V and VI. This portal site is positioned near the catalytic site and opens into the proposed peptide binding site explored in **Chapter 2**. Based on the position, it is possible that the **a**-factor substrate enters through the portal site and positions the farnesyl group in this large hydrophobic area close to the K234 residue, thus positioning the peptide near the catalytic site. This would also allow a close exit site, being close to the portal, to leave and be methylated before the secondary cleavage.

The E357 residue is a little less intuitive. Although it was identified with stricter search requirements, with a False Discovery Rate of 1%, it was not identified with confidence by PEAKS Studio X Pro. This program defines "confidence" to be evidence that there are y or b ions found on both the N-terminal and C-terminal side of the modification side of the modified amino acid. In this case, the modification is on the N-terminal amino acid so there can only be ions identified from the C-terminal side of the modification. The α-carbon of the E357 residue is also facing into a very hydrophobic region, not too far from the hydrophobic patch observed in **Chapter 2** (**Figure 3.6C** and **Figure 3.6D**), about 45 Å from the zinc ion. This hydrophobic patch is where the farnesyl group is suspected to bind during the N-terminal cleavage. Since this sample was not photolabeled, it is possible this was just a background signal that came through under the search criteria. On the other hand, it could be that the substrate is also capable of entering in such a way as to orient itself for the N-terminal cleavage. It has been proposed in previous studies that the CaaX processing is more-or-less a by-product of the **a**-factor maturation and is not actually required by Ste24 to perform this secondary cleavage (130,145). Without identification of a strong signal from the C_{10}-para probe, this may all be speculation, and the aid of an enrichment step may improve the support for these photolabeling results.
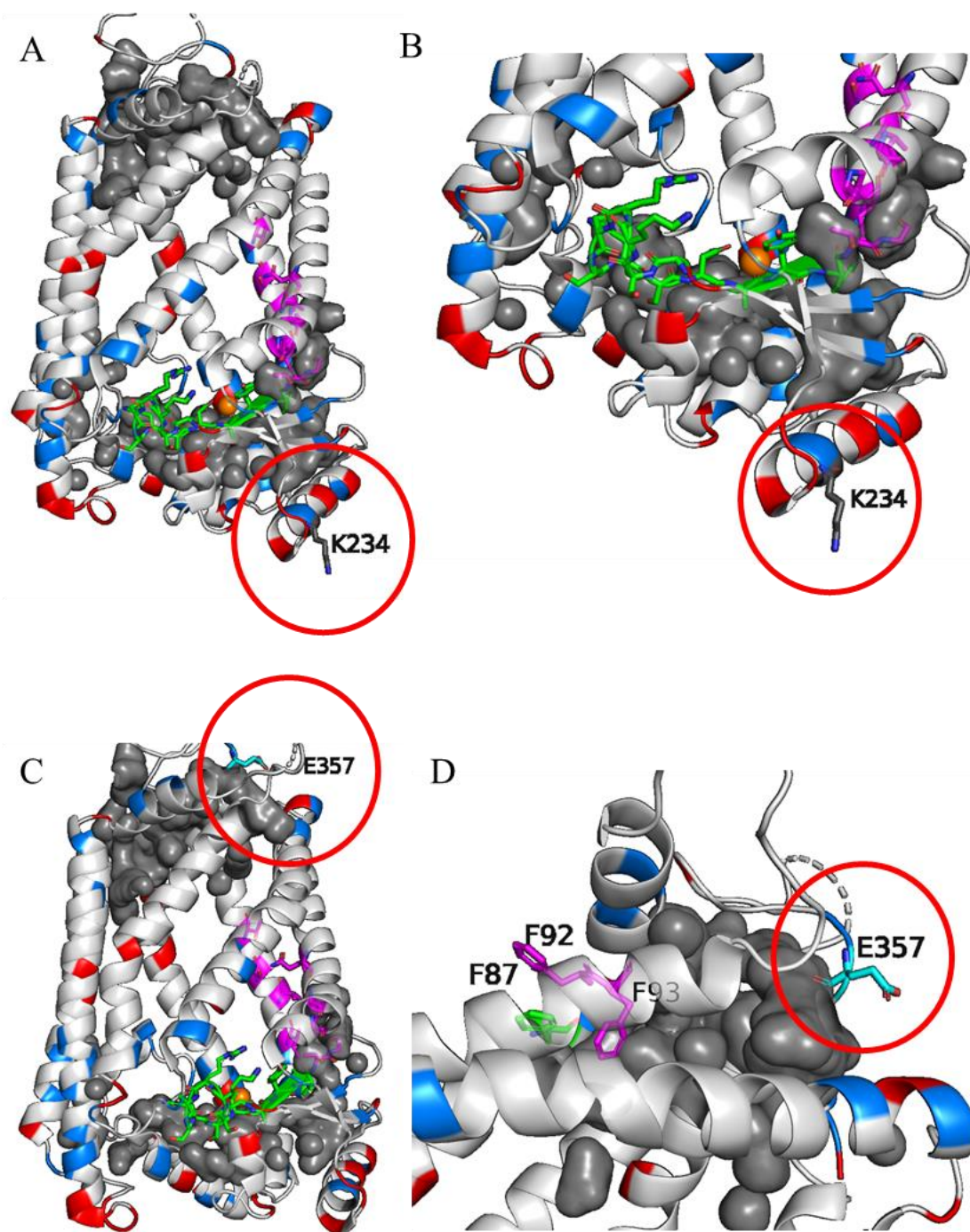
Figure 3.6: Structural analysis of the identified photolabeling sites. A) The full Ste24 protein is colored by electrostatics (negative is red, blue is positive and all hydrophobic residues are shown in gray. The portal site is highlighted with magenta sticks and the peptide binding site in green sticks. K234 is labeled on the protein. Internal cavities or pockets are shown as a dark gray surface. B) Close up of the location of K234, coloring as described in A. C) E357 is highlighted as a cyan stick on the top part of the protein, otherwise coloring is the same as A. D) Close up of E357 residue and including residues F87, F92 and F93, that indicates part of the hydrophobic pocket discussed in **Chapter 2**.

3.3.3 Raney nickel reduced the C₁₀-para probe, allowing for the introduction of an enrichment step for PAL proteomics techniques

The problem identified in **Section 3.3.3** was that the photolabeling site of Ste24 was not identified with high confidence likely due to the low efficiency of photolabeling. Because most of the protein, and thereby most of the digested peptides, are not photolabeled, it will be difficult to detect the photolabeled Ste24 fragment. By use of an enrichment step, only the photolabeled protein could be isolated from the photolabeled sample and thereby increase the relative abundance of the photolabeled fragment. The higher relative abundance of the photolabeled fragment would then increase the chance that it would be detected and identified by proteomics search software.

First, I wanted to have an approximation of the photolabeling efficiency of C_{10}-para. Separate reactions of 20 µg of wild-type Ste24 were photolabeled for different amounts of time and then a pulldown was performed with NeutrAvidin® Resin to isolate any photolabeled protein while any unlabeled protein was expected to remain in the supernatant. The first time point was at 0-min and was used as a starting point of the reaction. Then samples were taken at 1-, 5-, 10-, 20-, 30- and 40-minute time points. The unlabeled protein in the supernatant was isolated using a TCA precipitation, then re-dissolved in 2x SDS. The photolabeled protein in the resin was eluted in 2x SDS. Both the photolabeled and unlabeled samples were run on separate 10% SDS-PAGE gels followed by analysis by immunoblots using mouse α-HA primary antibody and goat α-mouse-HRP secondary antibody. The band intensities were determined using ImageJ.

As shown in **Figure 3.7A**, as more protein becomes photolabeled the amount of protein remaining unlabeled decreases. This results in more protein pulled down with the NeutrAvidin® Resin. While quantification by ImageJ showed that the amount of retained protein at the 30- and 40-minute mark was 67.6% and 21.5% respectively, visually there was not a large change in protein intensity at any of the time points. The band intensities across most of the timepoints were only slightly lighter than that seen at the 0-minute timepoint. The amount of photolabeled protein isolated by the pulldown appeared to be maximized around 30 minutes (**Figure 3.7B**). The difference in the supernatant sample between the 30 minute and 40-minute timepoints in **Figure 3.7A** may be due to variability of the TCA precipitation and aspiration performed to isolate the protein. The amount on unlabeled protein, determined by ImageJ, is highest at the 10-minute timepoint appears to be higher than the other timepoints even though the band intensity from the resin in **Figure 3.7B** continues to increase. These data indicate that any fluctuation in the

supernatant samples is more likely due to variability in the workup. Even if there is 80% of the protein left, this would mean that the photolabeling efficiency is 20%, that is quite high for the photoaffinity labeling of proteins. As high as 40% is considered very efficient labeling, depending on the probe and photoactivatable group (151). Due to low efficiency of labeling, there is a strong need for an enrichment technique to identify the photolabeled peptide in the sample.

A

Supernatant

B
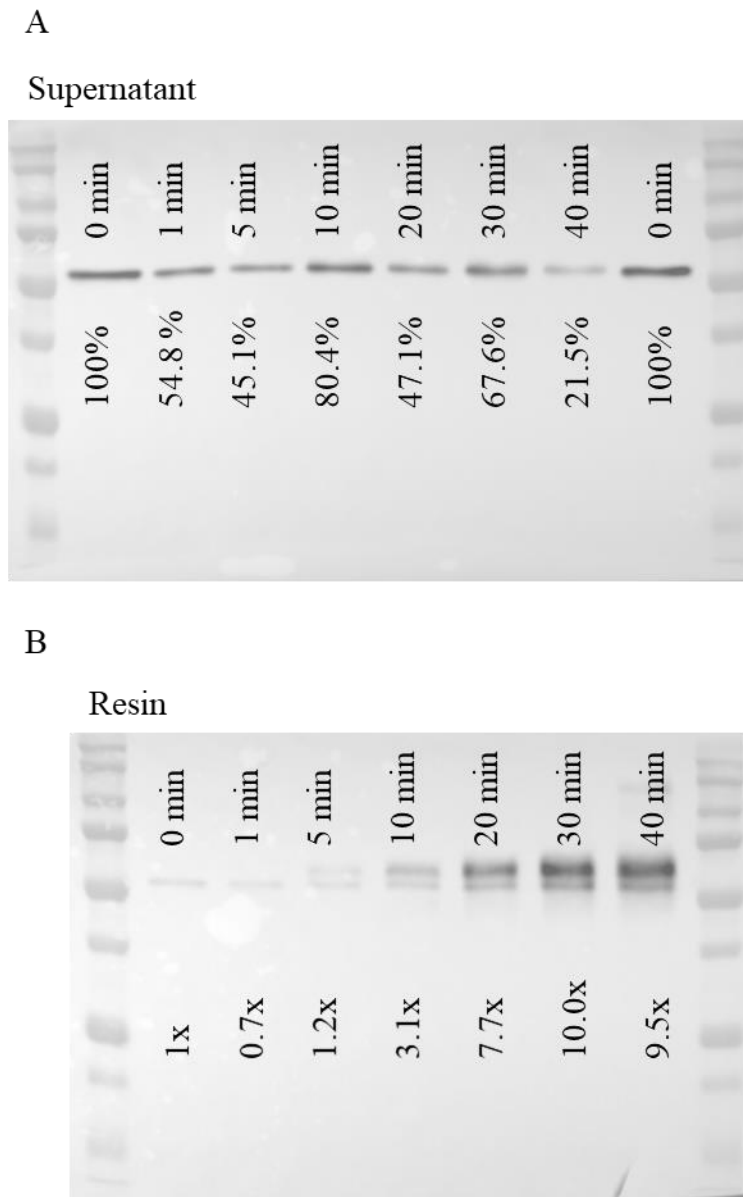
Resin

Figure 3.7: Photolabeling efficiency of Ste24 does not get significantly reduced within about 40 minutes. Photolabeled samples were harvested at the given timepoints and pulled down with 50 µL of a 1:1 slurry of Neutravidin®-Agarose Resin. A) The supernatant was used as a control for unlabeled protein. The protein was precipitated with TCA and re-dissolved in 2x SDS. B) The resin was washed three times with RIPA/SDS/I and then suspended in 50 µL SDS.

The other problem that arises in the proteomics search is that most proteomics software aims to identify the y and b ion series of an individual peptide, but when the C_{10}-para-**a**-factor probe photolabels Ste24 you are left with two peptide sequences that are linked together through the photolabeling reaction. Most proteomics software programs are not capable of determining the presence of two different peptide sequences in one MS2 spectra. To use the proteomics software, I needed to separate both peptides in between the photolabeling step and the ESI-MS-MS analysis.

Ideally, the enrichment step would be able to aid in both problems, the enrichment of the photolabeled peptide and the separation of the probe and protein sequences. One option for this enrichment would be to pulldown the photolabeled Ste24 protein with NeutrAvidin® Agarose resin followed by incubation with raney nickel. Raney nickel is a catalyst capable of hydrogenating thioether bonds. The C_{10}-para probe consists of two thioether bonds, one in the biotin handle and one connecting the photoactivatable benzophenone to the peptide sequence via the cysteine in the CaaX motif. After the pulldown with NeutrAvidin® Agarose Resin, incubation with raney nickel would be predicted to hydrogenate both sulfur functional groups, thereby releasing the isolated samples from the resin and separating the C_{10}-para peptide sequence from the photoactivatable group, and in doing so releasing the covalently bound Ste24 sequence.

Preliminary experiments showed that the probe is modified by the raney nickel reaction. For enrichment experiments, a total of 260 µg of WT Ste24 was photolabeled in 20 µL aliquots. The samples were then combined and pulled down with NeutrAvidin® Agarose resin for two hours at 4 °C. The supernatant of the pulldown was considered unlabeled as it was not bound to the resin and was added to a reaction flask containing 200 mg of raney nickel in phosphate buffered saline (PBS) pH 7.5. The resin was washed twice and then added to a separate reaction flask with 200 mg of raney nickel in PBS. Both reactions were incubated under a H_2 atmosphere overnight. Then the His_{10}-tag of Ste24 was eluted from the nickel in the catalyst using 1 M imidazole followed by concentration steps to remove the added imidazole. The sample was dried in a vacuum centrifuge, then digest, alkylated and digested with trypsin/Lys-C as discussed in **Section 3.3.2**. No chymotrypsin digest was performed here as I wanted to see if this enrichment step would provide enough of a protein signal to pursue further with an additional digestion reaction.

Without the enrichment step, the mass chromatograms show very large peptide peaks in both the unlabeled and photolabeled samples from only a 20 µg sample (**Figure 3.8A** and **Figure 3.8B**). When the enrichment step was incorporated, however, the signal of peptides were greatly

diminished (**Figure 3.8C** and **Figure 3.8D**) despite the large increase in amount of protein used. There are a few things that could have happened here. One is that the nickel in the raney nickel could be binding to the His_{10}-tag in the Ste24 protein, however imidazole had been added after the reaction to try to diminish this effect. Another problem could be that in typically pulldown assays are performed in the presence of detergent, in the case of methods herein RIPA/SDS/I buffer. The detergents help keep the protein in solution and properly denatured. But in preparation for the ESI-MS/MS detergents are normally avoided for downstream compatibility with the proteomics instruments. Preliminary experiments showed that performing a pulldown with PBS buffer worked sufficiently when compared with RIPA/SDS/I buffer (data not shown), but when it comes to harvesting the protein at the end of the raney nickel experiments, this lack of detergent could have caused the very hydrophobic Ste24 protein to aggregate and precipitate, resulting in a large loss of protein. Future studies will be performed to optimize the transition of solubilizing the Ste24 protein after enrichment for the ESI-MS/MS experimentation.

Despite the lack of protein in the enriched samples, PEAKS Studio X Pro recognized a few peptides in the sample, indicating some of the protein was still present. The pulled down, photolabeled sample showed a sequence coverage of 9% and the unlabeled (supernatant) contained 13% coverage. Four peptides appeared in both samples, 15 exclusively in the unlabeled and only one exclusively in the photolabeled sample. Unfortunately, no confident assignments were made for the addition of the C_{10}-para substrate.
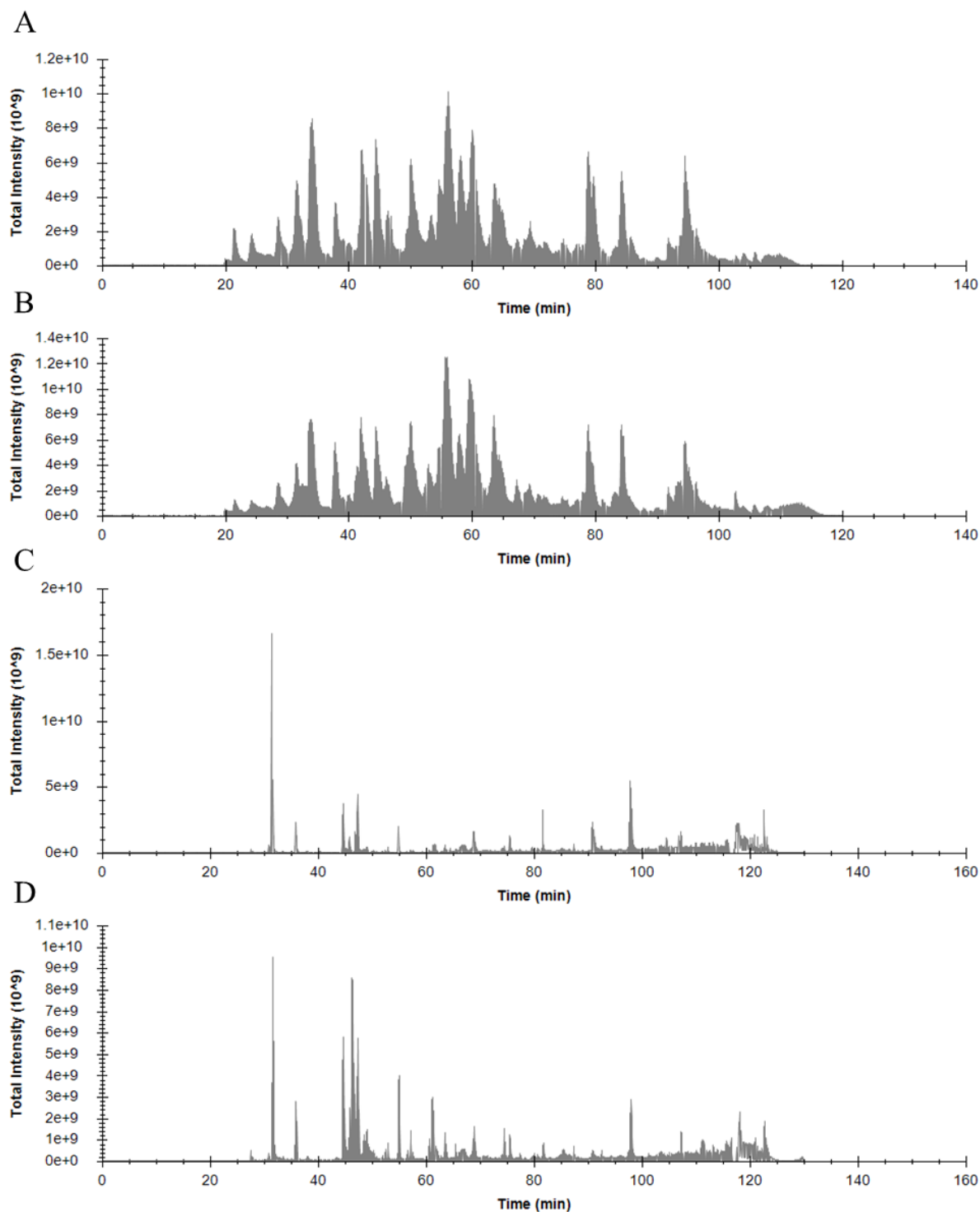
Figure 3.8: Raney nickel reactions greatly reduce the available protein for LC-MS/MS experimentations. The mass chromatograms for: A) Unlabeled sample utilized in **Section 3.3.4**, B) photolabeled sample utilized in **Section 3.3.4**, C) unlabeled, or supernatant Ste24 after NeutrAvidin resin pulldown and treatment with raney nickel according to **Section 3.2.10**, D) photolabeled Ste24 pulled down by NeutrAvidin Resin prepared as described in **Section 3.2.10**

3.4 Conclusion

Near full sequence coverage was obtained for the Ste24 protein using ESI-MS-MS and dual protease digests. In-gel digestion with only a trypsin/Lys-C cocktail digestion only obtained a sequence coverage of 44%. By incorporating a chymotrypsin digestion as well as switching to an overnight in-solution digestion the sequence coverage was increased to 64%. Finally, by utilizing a barocylcer digestion followed by an overnight digestion with both trypsin/Lys-C and chymotrypsin the sequence coverage of Ste24 by ESI-MS-MS was obtained at 93%.

The enrichment of the photolabeled Ste24 followed by hydrogenation with raney nickel was partially successful but is still in need of optimization. The photolabeled sample was pulled down by the NeutrAvidin® Agarose resin and hydrogenated by raney nickel. This sample did show a few identified peptides in the Peaks Studio X Pro, but the signals were low. The unlabeled and photolabeled samples also did not show a strong signal intensity in the mass chromatogram even though 260 µg of Ste24 protein was used. This was likely due to issues with solubility during the raney nickel reaction causing a loss of a significant amount of protein.

Despite the low signal intensity, there were a couple sites identified that potentially could contain the photolabeling site of C_{10}-para. K234 was identified with high confidence as being one of these sites. The location of this photolabeling site does intuitively make sense as the α-carbon is positioned towards a large hydrophobic cavity near the active site and portal in the side chamber. Unfortunately, no ions were identified by de novo sequencing that match the peptide sequence of the C_{10}-para sequence.

3.5 Future directions

1) Perform the pulldown and raney nickel isolation steps in the presence of Sodium deoxycholate, then perform a TCA precipitation followed by subsequent acetone washes to remove the detergent

2) Perform western blot analysis on the supernatant and resin of the raney nickel reaction to determine where the Ste24 protein was being lost during the reaction

3) Determine specific ions of benzophenone fragmentation which could be used to isolate the MS2 spectra that are likely to contain photolabeled fragments

CHAPTER 4.   STE24 SUBSTRATE-BINDING ANALYSES PROVIDED INSIGHTS INTO THE MECHANISM OF INHIBITION OF STE24 BY HIV PROTEASE INHIBITORS

4.1   Introduction

ZMPSTE24 inhibition has been implicated as a possible contributing cause of the side effects of several medications used to treat patients with Human Immunodeficiency Virus (HIV, **Figure 1.4A**) (30,36,116,117). These medications, such as lopinavir and ritonavir, are designed to target the aspartyl protease of HIV and it is currently unclear why these drugs have an inhibitory effect on the zinc metalloprotease ZMPSTE24. Unlike mammalian aspartyl proteases, the HIV enzyme can cleave between F/P and Y/P bonds (118,119,152). Because of this unique specificity, most HIV drugs target the aspartyl protease by mimicking the structure of a phenylalanine-proline substrate (118,119).

The HIV aspartyl protease inhibitors were designed to be highly specific towards the viral protease, as they tend to mimic dipeptide substrates that are not cleavable by mammalian aspartyl proteases. However, strong evidence has arisen that these drugs are capable of competitively inhibiting ZMPSTE24/Ste24 proteins to various extents (30). Patients who are taking these drugs typically start to display the hallmark progeroid disease symptom lipodystrophy (153,154). Furthermore, cell lines treated with HIV protease inhibitors start to accumulate high levels of prelamin A (117,154,155). These effects arose at drug concentrations comparable to those seen in patients prescribed the drugs daily (30,117). While many studies have shown that the HIV protease inhibitors bind in a competitive fashion to ZMPSTE24/Ste24, no precise location of the drug binding interaction has been obtained(30).

HIV protease inhibitors are designed to interact with aspartate residues within the active site of the HIV aspartyl protease, not to zinc metalloprotease motifs including histidine and glutamic acids. Upon inspecting the structures, it remains unclear how these drugs interact and inhibit ZMPSTE24/Ste24. The work presented in this chapter contains preliminary experiments that aim to explore the effects of the highly conserved aspartic acid residues found in the ZMPSTE24/Ste24 structures to gain insight into how the HIV drugs bind to these zinc metalloproteases.

4.2 Methods and Materials

4.2.1 Plasmids and yeast strains

All plasmids were created from the starting pCH1283 plasmid, made by prior Hrycyna lab graduates. This plasmid contains a His_{10}-HA_3 tag on the N-terminus of the Ste24 gene (*2µ URA3 P_{PGK}-His_{10}-HA_3-Ste24*). Mutant plasmids were made by traditional cloning methods (135) or by Q5-site directed mutagenesis kit ® (New England Biolabs) using the primers displayed in **Table 4.1** to make the corresponding plasmids shown in **Table 4.2**. Plasmids were transformed into the yeast strain SM3614, containing a double deletion for endogenous Rce1 and Ste24 (*MATa trp1 leu2 ura3 his4 can1 ste24Δ::LEU2 rce1Δ::TRP1*) (44). For the Ste14 protein used for activity assays, this was expressed in a pCH2733 (*2µ URA3 P_{PGK}-His_{10}-Myc_3-Ste14*) also transformed in the SM3614 strain.

For traditional cloning methods, mutant primers were designed to introduce single point mutations in the Ste24 gene (**Table 4.1**) in the pCH1283 plasmid (135). GoTaq® Green Master Mix (Promega) was used, by kit instructions, to perform a two-step PCR procedure to introduce the mutation of interest into the gene. The products of each PCR reaction were purified using the QIAquick PCR Purification Kit (Qiagen). After the final PCR step, both the purified PCR products and the pCH1283 vector were digested with EagI restriction enzyme and separated on a 1% agarose gel. The bands of interest were excised and purified using the QIAquick Gel Extraction Kit (Qiagen). The Rapid DNA Ligation Kit (Invitrogen) was then used to ligate the purified products. The ligated plasmids were then transformed into *E. coli* DH5α competent cells (Invitrogen) according to manufacturer instructions. Single colonies obtained from the transformation were used to grow miniprep cultures by selecting single colonies and growing them in 5 mL of Luria-Bertani (LB) broth containing 100 µg/mL ampicillin overnight at 37 ⁰C. The plasmids were harvested using QIAprep Spin Miniprep Kit (Qiagen), and the sequences were verified using bi-directional dye-terminator sequencing (BigDye® Terminator v3.1) by Purdue University Genomics Facility, or by GeneWiz sequencing center, using Sanger sequencing. The plasmids were then transformed into SM3614 cells using the Elble protocol (136) and grown at 30 ⁰C for two days on synthetic complete medium without uracil (SC-URA).

Alternatively, mutant plasmids were created using the Q5-site directed mutagenesis kit ® according to provided instructions and primers listed in **Table 4.1**. Briefly, end to end primers were designed and used to amplify linear versions of the mutated plasmid. Then the KLD enzyme

mix (kinase, ligase, DPNase) provided by the kit was used. The kinase adds a phosphate group to the 5'-end of the DNA strand, allowing the ligase to form circular plasmids. The DPNase then chews up the methylated template DNA and leaves only the mutated plasmids that were then transformed into DH5α cells (Invitrogen) where the DNA was collected by the QIAprep Spin Miniprep Kit (Qiagen) as described in the traditional cloning techniques. The sequences were verified using bi-directional dye-terminator sequencing (BigDye® Terminator v3.1) by Purdue University Genomics Facility, or by GeneWiz sequencing center, using Sanger sequencing. The plasmids were then transformed into SM3614 cells using the Elble protocol (136) and grown at 30 ⁰C for two days on synthetic complete medium without uracil (SC-URA).

Table 4.1: List of primers used to create mutations in pCH1283 WT Ste24 plasmid

Primer	Sequence (5' to 3')
SVEC (+)	CAGGGGGTGGTTTAGTTTAG
SVEC (-)	CAACTGTTGGGAAAGGCGATC
D55A (+)-Q5	GAAGACGAAATTGCTGATGAAAC
D55A (-)-Q5	CAGCACAGGTGGCAAC
D164A (+)	CTATGGATCACCGCTATGATCAAGAG
D164A (-)	CTCTTGATCATAGCGGTGATCCATAG
D280A (+)-Q5	CAAGAGAATTGTTTTGTTCGCCACTTTAG
D280A (-)-Q5	GAGGTGAATGGCAAACC
D394A (+)	GAATATCAAGCTGCTGCATAT GC
D394A (-)	GCATTGCAGCAGCTTGATATTC
D424A (+)	CCATGAATGTAGCTCCTCTG
D424A (-)	CAGAGGAGCTACATTCATGG

Table 4.2: List of plasmids used from the Ste24 plasmid library

Plasmid	Genotype
pCH1283	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24
pCH1353	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-D55A
pCH1345	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-D164A
pCH1354	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-D280A
pCH1318	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-D394A
pCH1321	2μ URA3 P_{PGK}-His$_{10}$-HA$_3$-Ste24-D424A

4.2.2   Crude membrane preparation

Crude membrane preparation was started by first inoculating a small, SC-URA culture with the desired Ste24 (or Ste14) strain that was grown overnight at 30 ⁰C. Then 15 mL was used to inoculate a larger, 1 L SC-URA flask and was grown to log phase (OD_{600} = 300-500) then harvested at 4000 xg. Lysis of the cells was performed using a yeast sorbitol lysis buffer (0.3 M sorbitol, 0.1 M NaCl, 12 mM $MgCl_2$, 1% aprotinin, 3 mM AEBSF, 1 mM DTT, 10 mM Tris-HCl, pH 7.5) at a ratio of 1 mL lysis buffer to 800 OD_{600} cells. The lysis buffer was added to the cell pellet in a 50 mL conical tube that was then vortexed and left on ice for 15 minutes to allow the cells to swell. The suspension was then flash frozen in liquid nitrogen and thawed in a 29 ⁰C water bath, twice. Then a french press was used to further lyse and break up the membrane fractions using 18,000 psi twice. The resulting suspension was first centrifuged twice at 500 xg for 10 minutes to remove cellular debris then centrifuged at 100,000 xg for 1 hour at 4 ⁰C. After the final centrifugation step, the supernatant was removed, and the membrane pellet was resuspended in 10 mM Tris-HCl, pH 7.5. Protein concentration was determined using a coomassie protein assay using a known concentration of bovine serum albumin (BSA) as the standard curve (137).

4.2.3   Membrane protein purification

To purify, the crude membrane proteins were first solubilized in buffer A (0.3 M sorbitol, 0.1 M NaCl, 6 mM $MgCl_2$, 10 mM Tris, pH 7.5, 10% glycerol, 1% aprotinin, 2 mM AEBSF) supplemented with 20 mM imidazole and 1% n-Dodecyl-B-D-maltopyranoside (DDM; Anatrace) with rocking at 4 ⁰C for 1 hour. The suspension was then centrifuged at 100,000 xg for 45 minutes

to remove the insoluble fractions, and the supernatant was added to Talon® Metal Affinity Resin (Clontech). The resin mixture was rocked at 4 °C for 1 hour, then washed with five column volumes of buffer B (40 mM imidazole and 1% DDM in buffer A) twice, buffer C (40 mM Imidazole, 1% DDM and 0.5 M KCl in buffer A) once, and buffer D (40 mM imidazole, 0.1% DDM and 0.5 M KCl in buffer A) once. The protein was then eluted from the resin with buffer E (250 mM Imidazole, 0.1% DDM in buffer A). The elution was then concentrated using Amicon® Ultra Centrifugal Filter 30,000 MWCO (Millipore) at 4,000 xg for 20-30 minutes at 4 °C until desired volume was reached. Protein concentration was determined using the amido black protein assay using BSA as a standard curve (138).

4.2.4 Radioactive methyltransferase-coupled diffusion assay

The C-terminal, or -aaXing, cleavage activity of Ste24 wild-type and mutant proteins were determined using an *in-vitro* radioactive methyltransferase-coupled diffusion assay, as previously described (100,139). First, 0.24 µg Ste14, 0.12 µg Ste24 and farnesylated 15-mer **a**-factor (final concentration 50 µM, **Figure 2.1A**) were added to 40 µg *E. coli* Polar Lipid Extracts. Then a rapid dilution was performed with Tris-HCl, pH 7.5 to dilute the fractions 1:8 (final 100 mM Tris-HCl). To start the reaction, *S*-adenosyl-[^{14}C-methyl]-L-methionine (SAM) (Perkin Elmer) was added to a final concentration of 20 µM and the reaction mixtures were incubated at 30 °C for 30 minutes. The reactions were stopped by addition of a 1 M NaOH/1% SDS mixture (final 450 µM NaOH, 0.5% SDS), then the samples were spread onto a pleated filter paper and placed in the neck of a scintillation vial filled with 10 mL of Biosafe II scintillation fluid (RPI). The vial was capped and left for 3 hours to allow time for the [^{14}C]-methanol, released by the addition of NaOH to stop the reaction, to diffuse into the scintillation fluid. The filter paper was then removed, and the radioactivity was quantified using a Packard TriCarb Scintillation counter. The specific activity was calculated from the counts per minute (CPM) and converted into pmol of -aaX residues cleaved per minute per mg of Ste24 protease. Reactions were set up as duplicates, and each set of duplicates was run in triplicate experiments.

4.2.5   SDS-PAGE and immunoblot assays

For immunoblot analysis of purified Ste24, 0.05 µg pure protein (diluted in 2x SDS loading dye) was loaded onto a 4% stacking, 10% separating SDS-PAGE gel. The samples were stacked at 85 V, followed by separating at 165 V for 40 minutes. The proteins were then transferred to a 0.45 µm nitrocellulose membrane (Cytiva Amersham ™ Protran ™ NC Nitrocellulose) at 100 V for 90 minutes. The membranes were then blocked overnight in 20% milk in PBST (1x PBS buffer, 0.1% Tween-20) at 4 ºC, followed by 2 hours in primary antibody (mouse, α-HA at 1:15,000) in 5% milk in PBST, then 1 hour in secondary antibody (goat-α-mouse-HRP, 1:4,000) in 4% milk in PBST. The bands were then visualized on a GeneGnome XRQ (SynGene) using SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific).

To determine purity of the Ste24 proteins, pure proteins were diluted in 2x SDS, then 1 µg total protein was loaded onto a 4% stacking, 10% separating SDS-PAGE gel. The running conditions were identical to the immunoblot SDS-PAGE described above. The gel was then stained overnight in Coomassie Brilliant Blue (0.3 M Coomassie Brilliant Blue, 10% acetic acid, 40% methanol) at room temperature, then destained with 10% acetic acid / 30% methanol to remove background.

4.2.6   Photolabeling assays

First, 0.5 µg of the purified Ste24 protein was added to 125 µg of *E. coli* Polar Lipid Extract (Avanti) and incubated on ice for 10 minutes. Then, a solution of buffer, DTT and photolabeling probe, C_{10}-para (**Figure 2.1B**) was added to dilute the protein reconstitution by 1:6, resulting in a final concentration of 1 mM DTT, 50 µM probe and 100 mM Tris-HCl, pH 7.5. This was left to incubate on ice another 10 minutes, in the dark. Then 65 µL was added to a 96-well plate sitting on ice under a UV lamp (365 nm), serving as the +UV, or photolabeled, sample. Then another 65 µL was added to a fresh tube, kept on ice but left in the dark for a -UV, unlabeled, sample. The samples were crosslinked under UV light (365 nm) for 30 minutes, and then 60 µL of both the -UV and +UV samples were added to 50 µL of a 1:1 slurry of NeutrAvidin® Agarose Resin (Thermo Scientific) in 800 µL RIPA/SDS/I buffer (25 mM tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton X-100, 1% sodium deoxycholate, 1% sodium dodecyl sulfate, 1% aprotinin, 2 mM AEBSF, 50 U/mL micrococcal nuclease, 1 mM $CaCl_2$, and 1 mM DTT). The protein reactions and resin

were left rocking for two hours at 4 ºC. The resin was then pelleted at 10,000 xg for 2 minutes and washed three times with 800 µL fresh RIPA/SDS/I buffer. The beads were then resuspended in 50 µL 2x SDS loading buffer, heated at 65 ºC for 30 minutes, then 10 µL was run on an SDS-PAGE followed by an immunoblot, as described above.

Quantification was performed using ImageJ as described in published methods (http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf). Briefly, rectangular bands were created that fit all bands in an immunoblot and the pixel measurements were counted. The pixel density (X) was inverted (255-X) for each of the -UV and +UV samples. The inverted -UV samples were then subtracted from the +UV samples to express the net values of the photolabeling signal. The net photolabeling signal was normalized to the photolabeling signal of the WT Ste24 sample.

4.2.7   MST assay for substrate binding

The Cy5-labeled isostere form of **a**-factor (**Figure 2.7**) was used to perform binding analysis utilizing the MST Nanotemper Monolith NT .115 instrument. Samples were prepared as generated by the Monolith NT .115 software. Briefly, 20 µL of the Ste24 stock solutions was used to perform 16 1:1 serial dilutions of 10 µL each. Then 10 µL of the Cy5-isostere **a**-factor was used added to a final concentration of 20 nM totaling a reaction volume of 20 µL. Two separate MST runs were performed for each set of dilutions. When using lopinavir, this was added along with the Cy5-isostere to a final concentration of 40 nM.

4.2.8   Tycho NT.6 for protein stability

Purified Ste24 mutants were loaded from their stock tubes directly into Tycho NT.6 capillaries from Nanotemper and loaded into the Tycho NT.6 instrument. The fluorescence ratio of 350 nm/330 nm was measured across temperatures from 30 °C up to 95 ° C. The inflection point (T_m) was determined by the Tycho NT.6 software.

4.2.9   PyMOL analysis

Structural analysis was performed by utilizing The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC. and analyzing the smSte24 (PDB: 4IL3). The zinc ion was colored in orange. Residues were selected as they became of interest and colors were modified from there.

Hydrogen bonds were determined using the "Action→Find→polar contacts→to other atoms in object". Further analysis will be discussed in the legends of the figures.

4.3 Results and Discussion

4.3.1 Activity levels of conserved aspartate residues in Ste24 revealed the importance of these residues in proper enzymatic function

The Ste24 family of zinc metalloproteases all share a conserved set of aspartate residues. In the yeast protein, the residues are D55, D164, D280 and D394. These residues all lie close to the Ste24 active site where they could potentially be involved in Ste24 binding to **a**-factor (**Figure 4.2A**). D164 and D280 residues are thought to hydrogen bond with other residues in the substrate binding site (**Figure 2.3C**), but the direct roles these aspartate residues play in Ste24 function is still unclear.

The conserved aspartate residues were mutated separately to alanine to determine the effect they had on Ste24 activity and substrate binding. The C-terminal cleavage activity levels retained by the conserved aspartate mutants was first assessed by using the well-established methyltransferase-coupled diffusion assay. Mutation of each of these residues resulted in Ste24 mutants with severely decreased catalytic activity (**Figure 4.2B**). D55A retained 50% activity and the other three demonstrated a complete loss of activity. As shown in **Figure 4.2C**, this was not due to a decreased expression level in most, although D394 did degrade upon purification. The retention of partial function of the D55A mutant could be because this residue lies further from the active site and may not have as strong of an interaction with the substrate as the others. As was shown in **Figure 2.4C**, the D164 and D280 residues appear to be important for forming hydrogen bonds with the side chains of a several of the residues in the substrate binding region, including S256 and R258. The disruption of these hydrogen bonds may interrupt proper Ste24 activity. The D394A mutation did show some degradation in both the immunoblot and coomassie stain, and this lack of stability may be at least partially responsible for the lack of activity.

It is interesting to note that previous studies have crystallized ZMPSTE24 in the presence of various HIV inhibitors (lopinavir and tipranavir) (30). In these crystal structures, an exact location of the drugs was not defined. It was suggested that the drugs may be binding to multiple sites (multiple aspartates) to inhibit the protein. The HIV protease drugs are designed to bind to the active site of the HIV aspartyl protease consisting of a DTG motif at the dimer interface (119).

Both monomers of the protease contain this DTG within the active site, bringing the two aspartate residues of the separate monomers close together to interact with the substrate. In the Ste24 structure, as shown in **Figures 2.3C** and **4.2A**, the D164 and D280 residues are close together and form hydrogen bonds with important residues in the peptide binding region. Disruption of the interactions between these highly conserved aspartate mutants with their hydrogen binding partners appears to be detrimental in function and binding of an inhibitor to either residue would additionally disrupt these important interactions.
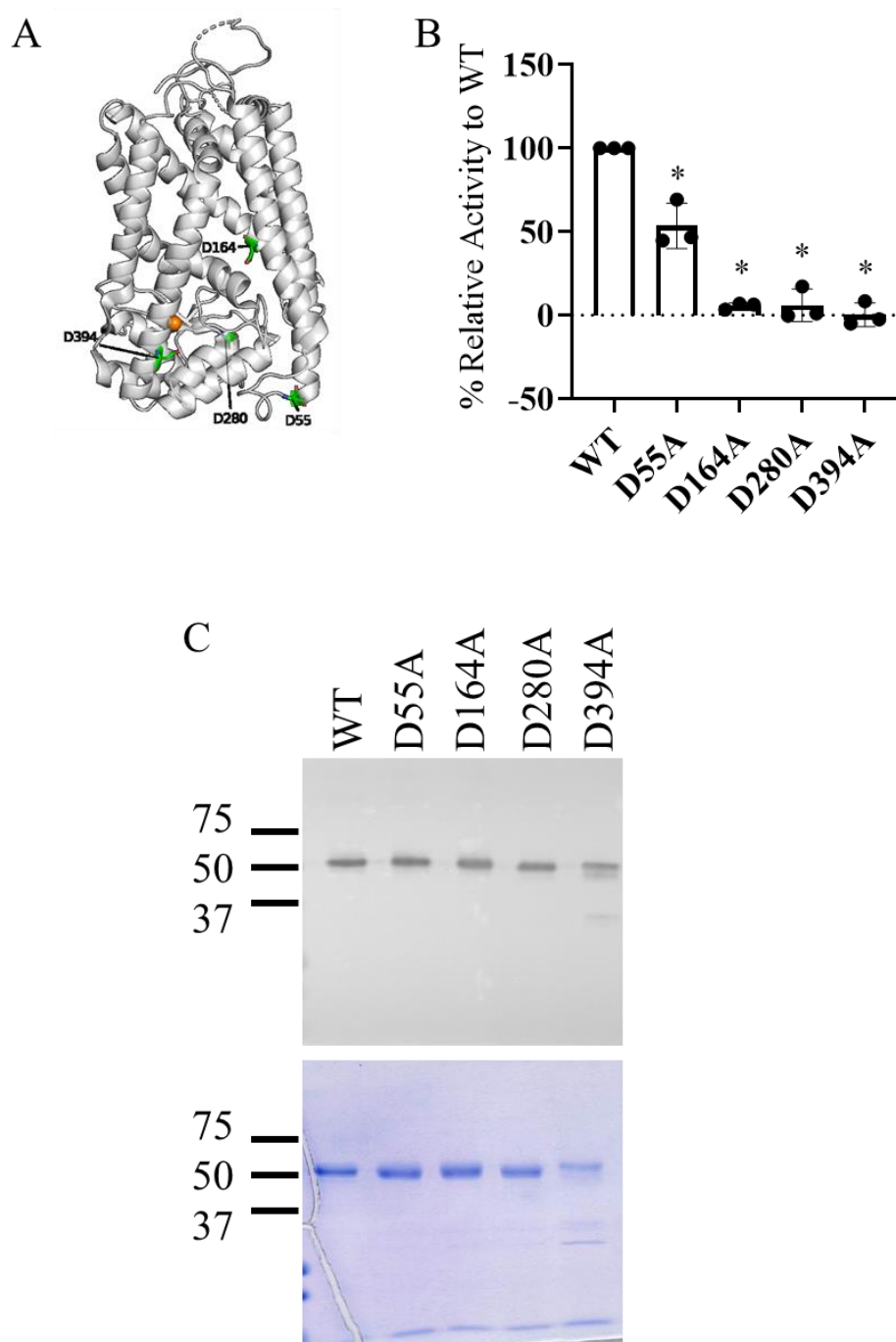
Figure 4.1: Mutating any of the conserved aspartate residues resulted in a significance decrease in activity relative to the wild-type Ste24 protein. A) Location of the conserved aspartates are depicted as green sticks. B) Methyltransferase-coupled diffusion assay results of purified conserved aspartate mutants. Bars indicate the mean of triplicate experiments; error bars depict the standard deviation and dots represent the relative activity levels of each individual experiment. (*) indicates statistical significance ($p<0.05$). C) Immunoblot (0.05 µg protein) and coomassie gels (1 µg protein) of purified conserved aspartate mutants. Immunoblot was incubated with the primary mouse α-HA (1:15,000) and the secondary goat α-mouse (1:4000).

4.3.2   Protein stability of conserved aspartate mutants

To determine whether the changes in activity were due to protein stability, a Nanotemper Tycho NT1.6 was used to measure changes in the fluorescence ratio between 350/330 nm over a range of temperatures. The 330 and 350 nm wavelengths correspond to the fluorescence of tryptophan and tyrosine in a non-hydrated (folded protein) or hydrated (protein unfolded) state, respectively. As proteins denature, more tryptophan residues become hydrated due to the higher exposure to the water in the solvent, so the ratio increases in a sigmoidal fashion. The inflection point of this sigmoidal curve can then be used to determine the melting temperature, or T_m, of the protein. More stable proteins tend to have a higher T_m as it requires more energy to denature them.

The aspartate mutants of Ste24 showed a statistically significant decrease in T_m. The strongest decrease is seen with the D55A, D164A and D280A mutants (45.4 °C, 45.3 °C, and 47.8 °C, respectively, compared to wild-type at 50.7 °C, **Figure 4.3**). The change in T_m, while moderate, does indicate a loss in protein stability with a ΔT_m of over 5 °C indicates a significant change in protein stability. Often, changes over 2 °C is considered a significant change in terms of protein stability (142,143). The hydrogen bonds formed between the proposed binding site (**Chapter 2**) with the aspartate residues D164 and D280 may be a stabilizing interaction, and loss of these aspartate side chains could cause denaturation or instability. The D394A trace (yellow line) does not follow the normal sigmoidal curve found by thermal instability changes, similar to that seen in the buffer (black line), indicating that variant is not as stable as the wildtype. This is further supported by the degradation seen in the immunoblot and coomassie gels (**Figure 4.2B**). Before purification, the D394A mutant still showed a complete knockdown of activity in crude samples, despite there being no degradation present (data not shown). These data suggest that the degradation is not the cause of inactivity. In all, the effects of the conserved aspartate mutants do not appear to be due to the stability of the protein.
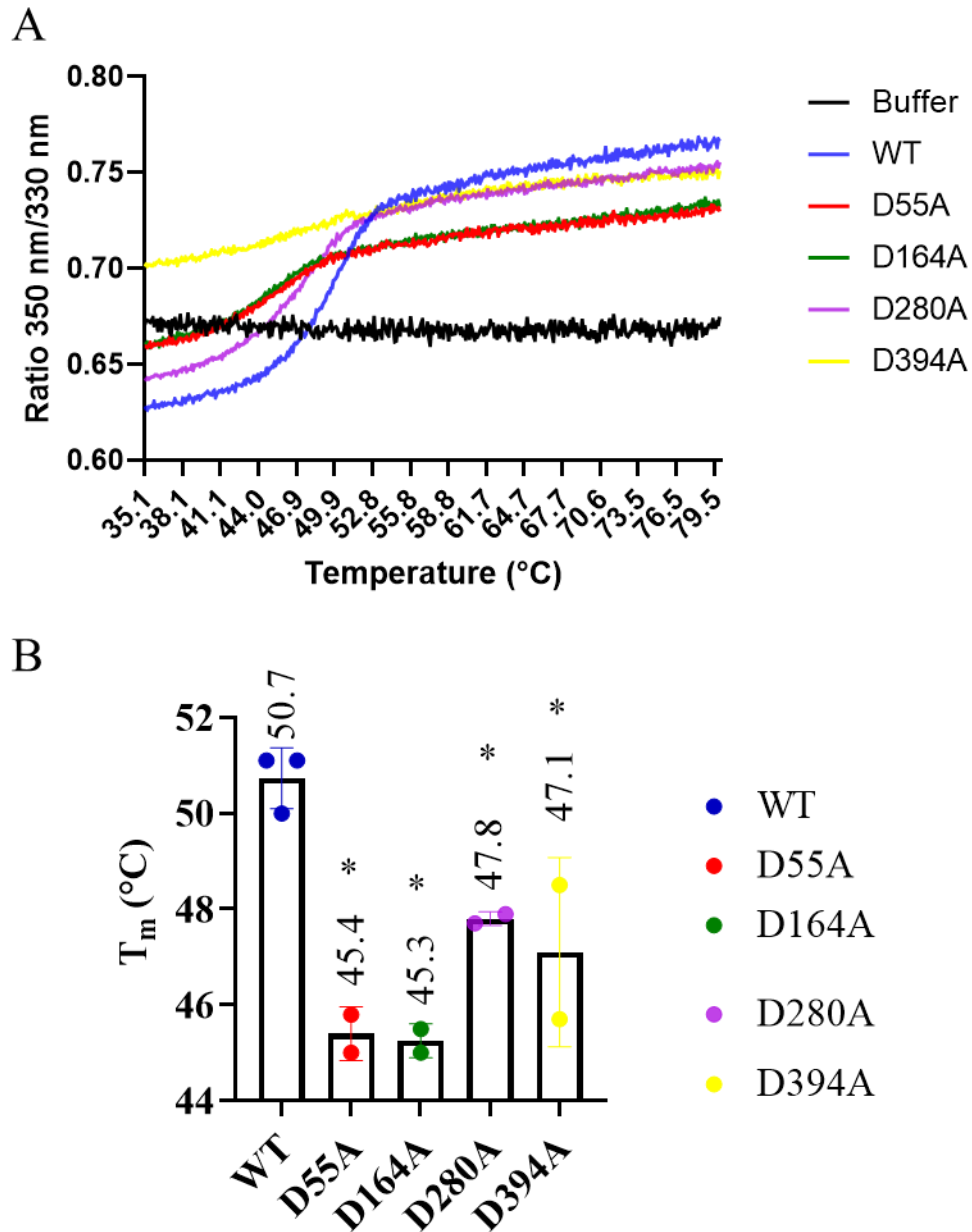
Figure 4.2: Purified aspartate mutations showed statistically significant, mild stability changes compared to the purified WT Ste24 protein. A) Thermostability measurements of purified conserved aspartate Ste24 mutants obtained from the Nanotemper Tycho NT 1.6. Curves represent average curves of two. B) T_m values of purified conserved aspartate Ste24 mutants obtained from the software of the Nanotemper Tycho N1.6, indicative of the inflection point of the curve. Dots represent values of individual replicates. * Indicates statistical significance ($p<0.05$).

4.3.3 All conserved aspartate mutants showed similar levels of photolabeling with the C_{10}-para-a-factor substrate compared to wild-type Ste24

To determine whether the Ste24 aspartate variants disrupted substrate binding, a photolabeling assay was performed, similar to that discussed in **Section 2.3.2**. Photolabeling assay analysis with C_{10}-para of the conserved aspartate residues was performed to determine if these residues affected substrate binding. As shown in **Figure 4.4**, the photolabeling signal intensities were all quite similar to the WT protein. Statistical analyses were not performed as the experiment was performed only once. The D55A and D280A mutations were relatively similar in the photolabeling signal to that of WT. The D55A mutation did retain partial activity in the activity assays so it is possible that it would still photolabel in a similar fashion. The D280 residue was shown in the crystal structure in **Figure 2.3C** to form a hydrogen bond with other residues in the Ste24 structure (**Chapter 2**). The hydrogen bond formed could be crucial in protein activity, but the loss of such may not be enough to fully prevent substrate binding. The D394A mutation was the only one that showed a decrease in photolabeling signal compared to the WT Ste24 protein. Due to the degradation seen in the expression blots, this result was not necessarily surprising as this means there was not as much intact Ste24 protein to begin. Finally, the D164A mutation showed almost a doubling of the photolabeling signal as compared to wildtype Ste24. As D164 lies near the proposed binding site of Ste24, mutation of the aspartate residue to a smaller alanine could open the binding site and allow more room for the bulky, photolabeling benzophenone group on the C_{10}-para probe. It could also be that a slight change in substrate orientation that allowed the benzophenone to photolabel more residues.

One important thing to keep in mind with photoaffinity labeling is that it is a great tool when monitoring large changes in substrate binding. Small changes are less detectable as the benzophenone group is highly reactive towards any nearby C-H bonds next to heteroatoms. This means that slight changes in the substrate orientation may not fully prevent photolabeling. It could be the photolabeling bond is labeling a different residue than the WT protein. Especially since the C_{10}-para substrate contains the benzophenone at the end of a highly flexible region, small changes in substrate binding likely are not preventing the photolabeling reaction.

To really observe the effects of the HIV aspartyl protease inhibitors with the Ste24 protein requires a direct measurement of substrate binding. As discussed in **Chapter 2**, a binding assay would allow the detection of smaller binding adjustments than the photolabeling assays presented

in this chapter. Then the addition of the HIV aspartyl protease inhibitors to the binding assay could reveal the site or sites of drug binding. If the drug specifically binds to one of the aspartate residues, then mutation of this residue would thereby prevent the drug from binding. Activity assays are only able to show protein inhibition, but when a mutation already causes the protein to be inactive, then adding an inhibitor will not show any effect. While these data are very preliminary, the development of a proper binding assay for Ste24 will significantly enhance the knowledge of how the Ste24 protein binds to substrates as well as small molecule inhibitors like the HIV protease drugs.
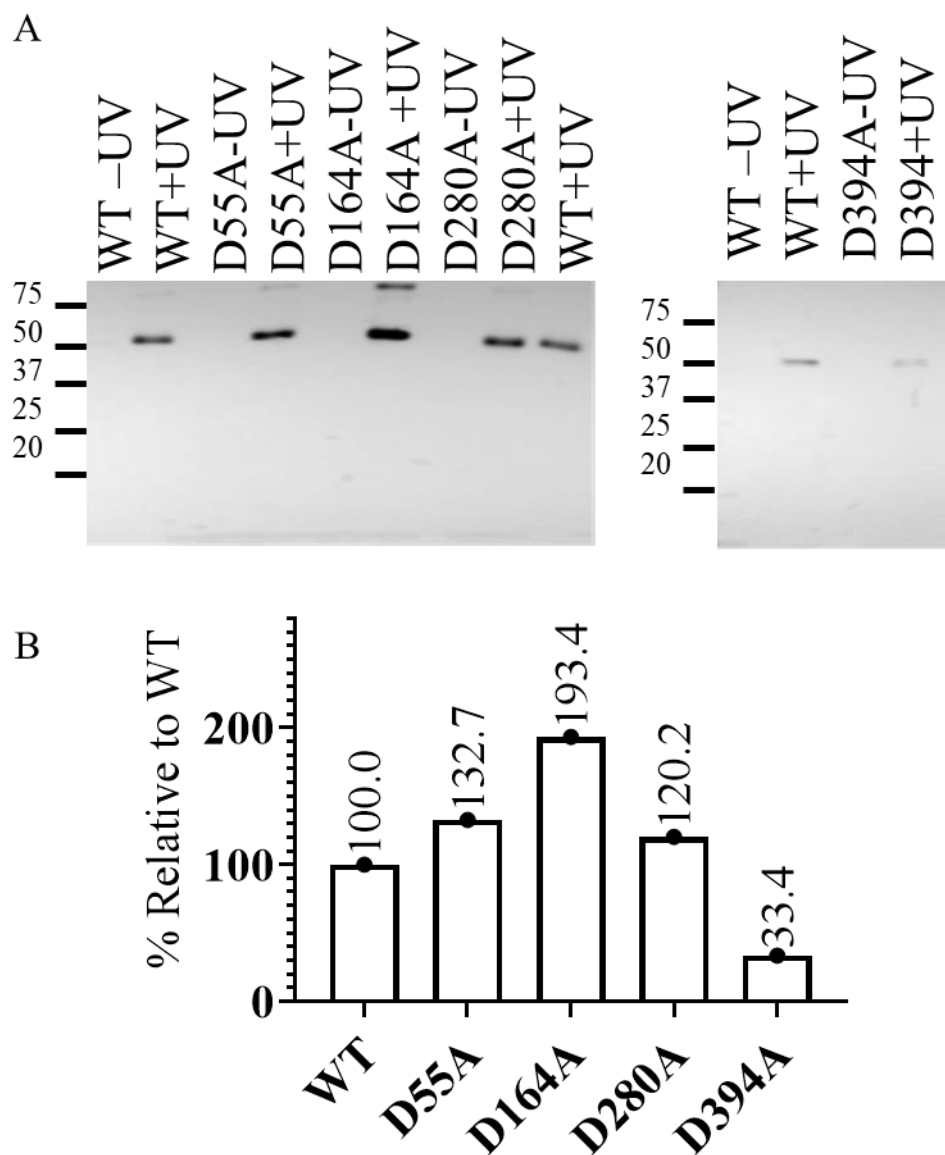
Figure 4.3: Photolabeling of the conserved aspartate mutants shows little variability from the wild-type. A) Photolabeling immunoblots of conserved aspartate mutants. B) Quantification of photolabeling signal from a single photolabeling analysis of conserved aspartate residues.

4.3.4 Binding analysis by MST allowed for the observation of changes in binding affinity, K_D, of wild-type Ste24 in the presence of the HIV aspartate inhibitor, lopinavir

With the development of a Ste24 binding assay (**Chapter 2**), the binding analysis of the conserved aspartate mutants could be performed. Using microscale thermophoresis (MST) with the Nanotemper Monolith 1.15 instrument and software, the K_D values of the aspartate mutants were calculated. The D55A, D280A and D394A concentrations, unfortunately, were not large enough to obtain a sufficient binding curve for the K_D values (**Figure 4.5**). But the D164A displayed a K_D value of 0.06 µM (**Table 4.3**). These were the result of two subsequent experiments. While the K_D values do differ between the WT and D164A, with the standard deviation given, they are not significantly different. These data suggest that the D164A mutation is still capable of binding to the substrate, that is also supported by the photolabeling results in **Figure 4.4**.

A Conserved Aspartates Mutants

D394A
D280A
D164A
D55A
WT

Conserved Aspartates Mutants

B

D164A
WT
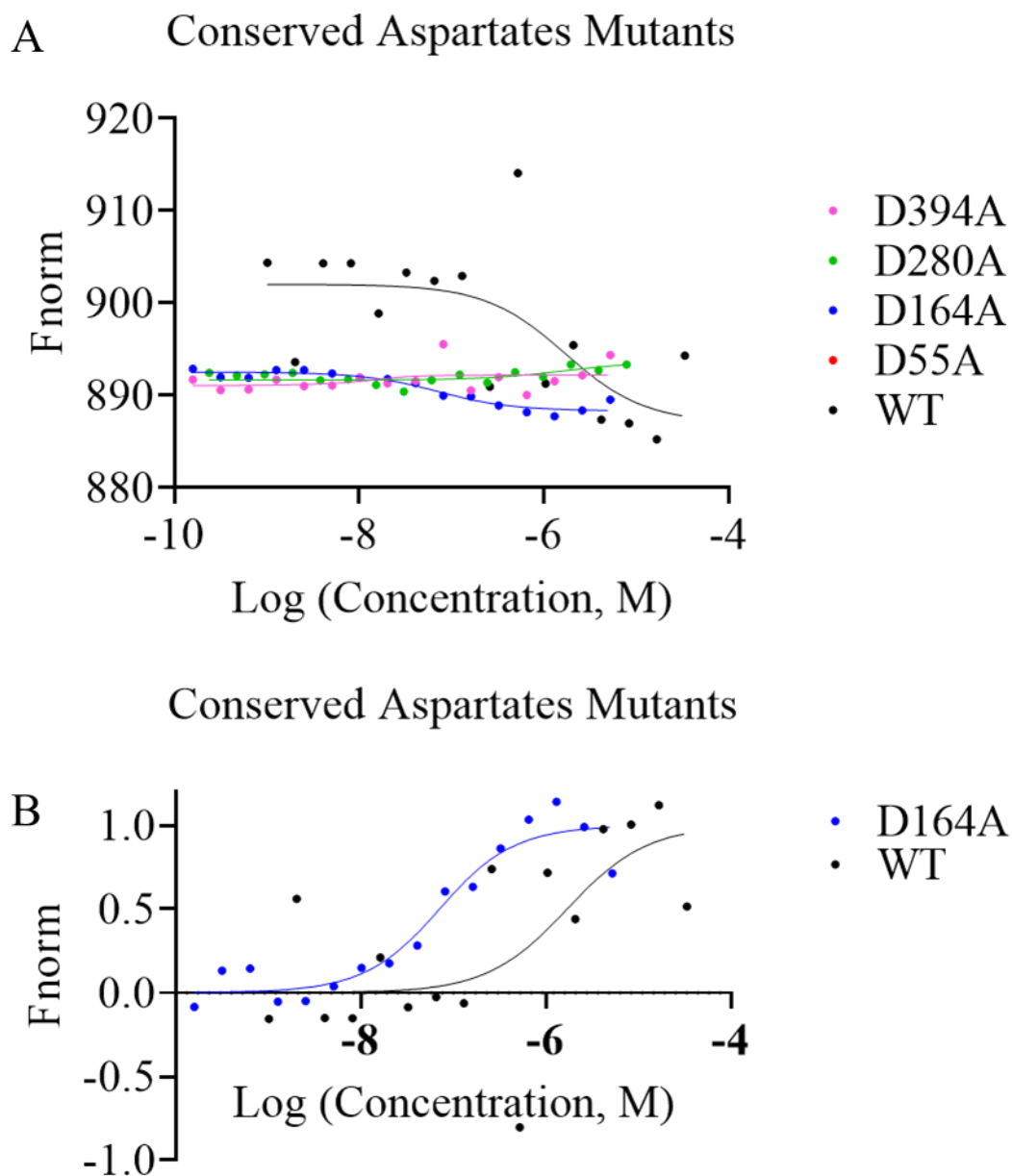
Figure 4.4: Binding analysis of the conserved aspartates revealed likely a higher K_D value for most of these mutations. Both best-fit curves were derived by the MST instrument, Nanotemper Monolith NT 1.16 software. The dots represent the average value of duplicate samples. A) F_{norm} of the conserved aspartate mutants. B) Fraction bound transformation of the binding curves of WT and D164A from (A).

Table 4.3: K_D values of WT and D164A from **Figure 4.5**

Strain	K_D (µM)
WT	1.7 ± 2.5
D164A	0.1 ± 0.03

The HIV aspartate inhibitor lopinavir is capable of inhibiting Ste24 (36,116,117). In 2016, Mehmood *et al* showed, by mass spectrometry analysis, that the HIV protease inhibitors lopinavir and ritonavir were able to bind to the human ZMPSTE24 protein (145). This binding of the inhibitor also showed a decrease in the binding and catalysis of ZMPSTE24 in the maturation of the prelamin A substrate (145). While Mehmood *et al* did show that binding of the inhibitor lopinavir can stabilize the ZMPSTE24 structure and prevent binding of the prelamin A substrate, it remains unclear whether the inhibition models a competitive or non-competitive inhibitor.

By utilizing MST, the addition of lopinavir to the WT Ste24 protein caused a drastic change in the binding curve (**Figure 4.6A**). While there was a full binding curve for the WT without the lopinavir, the addition of this drug flattened out the curve substantially and did not show any detectable binding to the Cy5-**a**-factor substrate. This discovery is in support of the 2016 study that showed that lopinavir binding to mammalian ZMPSTE24 prevented binding of the prelamin A substrate (145). The drug was added at 40 nM and the Cy5-isostere was only at 20 nM. Due to the higher concentration of the drug, it was likely that the binding site of Ste24 was fully saturated and may have required a much higher concentration of protein to overcome the inhibitory effect. The addition of lopinavir to the D164A mutant Ste24 also increased the apparent K_D of the D164A mutation (from 0.06 µM without and 0.2 µM with lopinavir). The curve had not been as drastically changed as with the WT protein, however.

It has been suggested that the HIV aspartyl protease inhibitors could bind to more than one location within Ste24, as there have been attempts to crystalize ZMPSTE24 in the presence of the drugs without a specific binding site being discovered (30). While they were able to obtain crystal structures of the protein they could not determine the location of the HIV protease inhibitors (29). Based on the K_D values presented here, while the K_D values are different, the effect of lopinavir addition does not change the binding curve as drastically as it does with the WT protein. If the

binding site of the drug was altered, thereby preventing the drug from binding, it would be hypothesized that there would not be a difference between the K_D in the presence or absence of the HIV drug. As this was the case with D164A it could be this is at least one binding site of lopinavir.
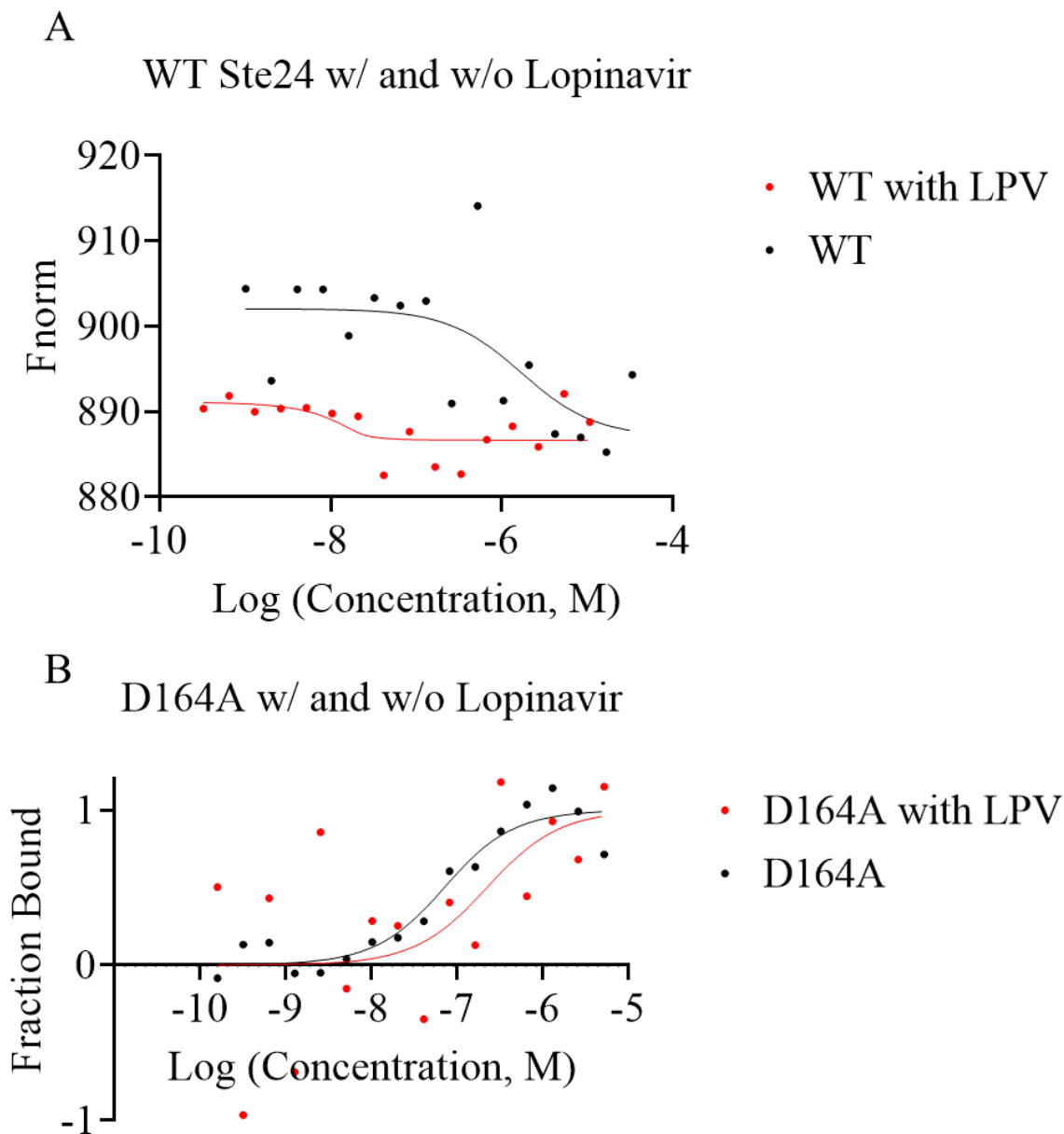
Figure 4.5: Binding assays in the presence of the inhibitor lopinavir (LPV) alters the binding curve of Ste24 WT and D164A, indicating possible competitive binding. Best-fit curves were derived by the MST instrument, Nanotemper Monolith NT 1.16 software. The dots represent the average value of duplicate samples. A) F_{norm} of the WT Ste24 protein with and without 40 nM LPV. B) Fraction bound curves of D164A with and without 40 nM LPV.

Table 4.4: K_D values for the D164A mutant with and without lopinavir from **Figure 4.6B**

Sample	Apparent K_D (μM)
D164A	0.06 ± 0.03
D164A with Lopinavir	0.2 ± 0.3

The K_D values of the other conserved aspartate mutations, and the WT protein with lopinavir, were unable to be determined at these concentrations. Future studies will follow up these experiments with higher concentrations of Ste24 proteins. In addition, triplicate samples will be obtained. From these future experiments it should be possible to determine which residues are being bound by lopinavir, thereby allowing future drug design to be more selective against this site within Ste24.

4.4   Conclusion

These preliminary experiments provide some insight into the importance of conserved aspartate residues in proper Ste24 function. Mutations of these residues greatly reduced activity without changes in thermal stability. The photolabeling ability by the C_{10}-para probe was also not greatly affected by the presence of the mutations. The addition of HIV protease inhibitors to this analysis would be very interesting in determination of the direct role of these drugs on the inhibition of Ste24 and ZMPSTE24. Performing competitive photolabeling experiments within the presence of HIV aspartyl protease inhibitors would be interesting to see if the removal of any of the conserved aspartates could prevent the competitive binding/inhibition effect of these drugs.

Also, the development of a direct binding assay provides some insight into the mechanism of inhibition by the HIV protease inhibitors. Photolabeling analysis is a great technique when determining large changes in substrate binding. When the binding changes are small, differences in binding of the photolabeling probe can be too subtle to be detected. On the other end of the spectrum, because photolabeling is dependent on proximity of the photoactivatable group to

nearby C-H bonds, small changes in one area of the substrate can lead to large changes in the photolabeling efficiency of the probe. So, a substrate could still be binding just as well as WT, but the mutation may move the photoactivatable group just outside of the proximity required to form the covalent bond.

4.5 Future directions

1) Increase the concentration of the purified aspartate mutant proteins through use of a better expression system, likely a baculovirus expression system
2) Repeat the MST and T_m experiments in triplicate to determine changes that are statistically significant
3) Run the MST experiments in the presence of different HIV aspartyl protease inhibitors, including ritonavir as it was shown to be a strong inhibitor of Ste24, lopinavir and darunavir. The darunavir was not shown to be an inhibitor of Ste24, so may act as a negative control
4) Perform the MST experiments with the aspartate mutants and the various HIV protease inhibitors to determine how the replacement of each aspartate affects the binding of the protease inhibitors, allowing a possible drug binding location to be obtained

REFERENCES

1. Barrowman, J., Hamblet, C., George, C. M., and Michaelis, S. (2008) Analysis of prelamin A biogenesis reveals the nucleus to be a CaaX processing compartment. *Mol Biol Cell* **19**, 5398-5408

2. Dai, Q., Choy, E., Chiu, V., Romano, J., Slivka, S. R., Steitz, S. A., Michaelis, S., and Philips, M. R. (1998) Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. *J Biol Chem* **273**, 15030-15034

3. Romano, J. D., Schmidt, W. K., and Michaelis, S. (1998) The Saccharomyces cerevisiae prenylcysteine carboxyl methyltransferase Ste14p is in the endoplasmic reticulum membrane. *Mol Biol Cell* **9**, 2231-2247

4. Schmidt, W. K., Tam, A., Fujimura-Kamada, K., and Michaelis, S. (1998) Endoplasmic reticulum membrane localization of Rce1p and Ste24p, yeast proteases involved in carboxyl-terminal CAAX protein processing and amino-terminal a-factor cleavage. *Proc Natl Acad Sci U S A* **95**, 11175-11180

5. Corrigan, D. P., Kuszczak, D., Rusinol, A. E., Thewke, D. P., Hrycyna, C. A., Michaelis, S., and Sinensky, M. S. (2005) Prelamin A endoproteolytic processing in vitro by recombinant Zmpste24. *Biochem J* **387**, 129-138

6. Wright, L. P., and Philips, M. R. (2006) Thematic review series: lipid posttranslational modifications. CAAX modification and membrane targeting of Ras. *J Lipid Res* **47**, 883-891

7. Zhang, F. L., and Casey, P. J. (1996) Protein prenylation: molecular mechanisms and functional consequences. *Annu Rev Biochem* **65**, 241-269

8. Gelb, M. H., Brunsveld, L., Hrycyna, C. A., Michaelis, S., Tamanoi, F., Van Voorhis, W. C., and Waldmann, H. (2006) Therapeutic intervention based on protein prenylation and associated modifications. *Nat Chem Biol* **2**, 518-528

9. He, B., Chen, P., Chen, S. Y., Vancura, K. L., Michaelis, S., and Powers, S. (1991) RAM2, an essential gene of yeast, and RAM1 encode the two polypeptide components of the farnesyltransferase that prenylates a-factor and Ras proteins. *Proc Natl Acad Sci U S A* **88**, 11373-11377

10. Taylor, J. S., Reid, T. S., Terry, K. L., Casey, P. J., and Beese, L. S. (2003) Structure of mammalian protein geranylgeranyltransferase type-I. *EMBO J* **22**, 5963-5974

11. Long, S. B., Casey, P. J., and Beese, L. S. (1998) Cocrystal structure of protein farnesyltransferase complexed with a farnesyl diphosphate substrate. *Biochemistry* **37**, 9612-9618

12. Park, H. W., Boduluri, S. R., Moomaw, J. F., Casey, P. J., and Beese, L. S. (1997) Crystal structure of protein farnesyltransferase at 2.25 angstrom resolution. *Science* **275**, 1800-1804

13. Strickland, C. L., Windsor, W. T., Syto, R., Wang, L., Bond, R., Wu, Z., Schwartz, J., Le, H. V., Beese, L. S., and Weber, P. C. (1998) Crystal structure of farnesyl protein transferase complexed with a CaaX peptide and farnesyl diphosphate analogue. *Biochemistry* **37**, 16601-16611

14. Fiordalisi, J. J., Johnson, R. L., Weinbaum, C. A., Sakabe, K., Chen, Z., Casey, P. J., and Cox, A. D. (2003) High affinity for farnesyltransferase and alternative prenylation contribute individually to K-Ras4B resistance to farnesyltransferase inhibitors. *J Biol Chem* **278**, 41718-41727

15. Toth, J. I., Yang, S. H., Qiao, X., Beigneux, A. P., Gelb, M. H., Moulson, C. L., Miner, J. H., Young, S. G., and Fong, L. G. (2005) Blocking protein farnesyltransferase improves nuclear shape in fibroblasts from humans with progeroid syndromes. *Proc Natl Acad Sci U S A* **102**, 12873-12878

16. Capell, B. C., Erdos, M. R., Madigan, J. P., Fiordalisi, J. J., Varga, R., Conneely, K. N., Gordon, L. B., Der, C. J., Cox, A. D., and Collins, F. S. (2005) Inhibiting farnesylation of progerin prevents the characteristic nuclear blebbing of Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A* **102**, 12879-12884

17. Zhang, H., Seabra, M. C., and Deisenhofer, J. (2000) Crystal structure of Rab geranylgeranyltransferase at 2.0 A resolution. *Structure* **8**, 241-251

18. Hampton, S. E., Dore, T. M., and Schmidt, W. K. (2018) Rce1: mechanism and inhibition. *Crit Rev Biochem Mol Biol* **53**, 157-174

19. Manolaridis, I., Kulkarni, K., Dodd, R. B., Ogasawara, S., Zhang, Z., Bineva, G., Reilly, N. O., Hanrahan, S. J., Thompson, A. J., Cronin, N., Iwata, S., and Barford, D. (2013) Mechanism of farnesylated CAAX protein processing by the intramembrane protease Rce1. *Nature* **504**, 301-305

20. Bergo, M. O., Gavino, B. J., Hong, C., Beigneux, A. P., McMahon, M., Casey, P. J., and Young, S. G. (2004) Inactivation of Icmt inhibits transformation by oncogenic K-Ras and B-Raf. *J Clin Invest* **113**, 539-550

21. Bos, J. L. (1988) The ras gene family and human carcinogenesis. *Mutat Res* **195**, 255-271

22. Bos, J. L. (1989) ras oncogenes in human cancer: a review. *Cancer Res* **49**, 4682-4689

23. Hildebrandt, E. R., Davis, D. M., Deaton, J., Krishnankutty, R. K., Lilla, E., and Schmidt, W. K. (2013) Topology of the yeast Ras converting enzyme as inferred from cysteine accessibility studies. *Biochemistry* **52**, 6601-6614

24. Michaelis, S., Chen, P., Berkower, C., Sapperstein, S., and Kistler, A. (1992) Biogenesis of yeast a-factor involves prenylation, methylation and a novel export mechanism. *Antonie Van Leeuwenhoek* **61**, 115-117

25. Michaelis, S., and Barrowman, J. (2012) Biogenesis of the Saccharomyces cerevisiae pheromone a-factor, from yeast mating to human disease. *Microbiol Mol Biol Rev* **76**, 626-651

26. Spear, E. D., Alford, R. F., Babatz, T. D., Wood, K. M., Mossberg, O. W., Odinammadu, K., Shilagardi, K., Gray, J. J., and Michaelis, S. (2019) A humanized yeast system to analyze cleavage of prelamin A by ZMPSTE24. *Methods* **157**, 47-55

27. Spear, E. D., Hsu, E. T., Nie, L., Carpenter, E. P., Hrycyna, C. A., and Michaelis, S. (2018) missense mutations that cause progeroid diseases decrease prelamin A cleavage activity and/or protein stability. *Dis Model Mech*

28. Pryor, E. E., Horanyi, P. S., Clark, K. M., Fedoriw, N., Connelly, S. M., Koszelak-Rosenblum, M., Zhu, G., Malkowski, M. G., Wiener, M. C., and Dumont, M. E. (2013) Structure of the integral membrane protein CAAX protease Ste24p. *Science* **339**, 1600-1604

29. Quigley, A., Dong, Y. Y., Pike, A. C., Dong, L., Shrestha, L., Berridge, G., Stansfeld, P. J., Sansom, M. S., Edwards, A. M., Bountra, C., von Delft, F., Bullock, A. N., Burgess-Brown, N. A., and Carpenter, E. P. (2013) The structural basis of ZMPSTE24-dependent laminopathies. *Science* **339**, 1604-1607

30. Clark, K. M., Jenkins, J. L., Fedoriw, N., and Dumont, M. E. (2017) Human CaaX protease ZMPSTE24 expressed in yeast: Structure and inhibition by HIV protease inhibitors. *Protein Sci* **26**, 242-257

31. Pendás, A. M., Zhou, Z., Cadiñanos, J., Freije, J. M., Wang, J., Hultenby, K., Astudillo, A., Wernerson, A., Rodríguez, F., Tryggvason, K., and López-Otín, C. (2002) Defective prelamin A processing and muscular and adipocyte alterations in Zmpste24 metalloproteinase-deficient mice. *Nat Genet* **31**, 94-99

32. Bergo, M. O., Gavino, B., Ross, J., Schmidt, W. K., Hong, C., Kendall, L. V., Mohr, A., Meta, M., Genant, H., Jiang, Y., Wisner, E. R., Van Bruggen, N., Carano, R. A., Michaelis, S., Griffey, S. M., and Young, S. G. (2002) Zmpste24 deficiency in mice causes spontaneous bone fractures, muscle weakness, and a prelamin A processing defect. *Proc Natl Acad Sci U S A* **99**, 13049-13054

33. Wiley, P. A., and Purdue University. Chemistry. *Molecular and biochemical analyses of prelamin A and a-factor proteolytic processing by the novel zinc metalloprotease ZMPSTE24 and its yeast homolog, Ste24p*,

34. Shackleton, S., Smallwood, D. T., Clayton, P., Wilson, L. C., Agarwal, A. K., Garg, A., and Trembath, R. C. (2005) Compound heterozygous ZMPSTE24 mutations reduce prelamin A processing and result in a severe progeroid phenotype. *J Med Genet* **42**, e36

35. Navarro, C. L., Cadiñanos, J., De Sandre-Giovannoli, A., Bernard, R., Courrier, S., Boccaccio, I., Boyer, A., Kleijer, W. J., Wagner, A., Giuliano, F., Beemer, F. A., Freije, J. M., Cau, P., Hennekam, R. C., López-Otín, C., Badens, C., and Lévy, N. (2005) Loss of ZMPSTE24 (FACE-1) causes autosomal recessive restrictive dermopathy and accumulation of Lamin A precursors. *Hum Mol Genet* **14**, 1503-1513

36. Coffinier, C., Hudon, S. E., Lee, R., Farber, E. A., Nobumori, C., Miner, J. H., Andres, D. A., Spielmann, H. P., Hrycyna, C. A., Fong, L. G., and Young, S. G. (2008) A potent HIV protease inhibitor, darunavir, does not inhibit ZMPSTE24 or lead to an accumulation of farnesyl-prelamin A in cells. *J Biol Chem* **283**, 9797-9804

37. Yang, S. H., Qiao, X., Farber, E., Chang, S. Y., Fong, L. G., and Young, S. G. (2008) Eliminating the synthesis of mature lamin A reduces disease phenotypes in mice carrying a Hutchinson-Gilford progeria syndrome allele. *J Biol Chem* **283**, 7094-7099

38. Tamanoi, F., Hrycyna, C. A., and Bergo, M. O. (2011) Protein Prenylation. Part A. in *The enzymes,* , 1st Ed., Elsevier,, Amsterdam

39. Davies, B. S. J., Coffinier, C., Yang, S. H., Jung, H.-J., Fong, L. G., and Young, S. G. (2011) Posttranslational processing of Nuclear Lamins. in *Protein Prenylation: Part A* (Inc, E. ed.), 1 Ed., Elsevier Inc., London, UK. Amsterdam, The Netherlands. Oxford, UK. Waltham, Massachusetts, USA. San Diego, California, USA pp 21-41

40. Barrowman, J., Hamblet, C., Kane, M. S., and Michaelis, S. (2012) Requirements for efficient proteolytic cleavage of prelamin A by ZMPSTE24. *PLoS One* **7**, e32120

41. Baron, R. A., and Casey, P. J. (2004) Analysis of the kinetic mechanism of recombinant human isoprenylcysteine carboxylmethyltransferase (Icmt). *BMC Biochem* **5**, 19

42. Anderson, J. L., Henriksen, B. S., Gibbs, R. A., and Hrycyna, C. A. (2005) The isoprenoid substrate specificity of isoprenylcysteine carboxylmethyltransferase: development of novel inhibitors. *J Biol Chem* **280**, 29454-29461

43. Cushman, I., and Casey, P. J. (2011) RHO methylation matters: a role for isoprenylcysteine carboxylmethyltransferase in cell migration and adhesion. *Cell Adh Migr* **5**, 11-15

44. Tam, A., Nouvet, F. J., Fujimura-Kamada, K., Slunt, H., Sisodia, S. S., and Michaelis, S. (1998) Dual roles for Ste24p in yeast a-factor maturation: NH2-terminal proteolysis and COOH-terminal CAAX processing. *J Cell Biol* **142**, 635-649

45. Chiu, V. K., Silletti, J., Dinsell, V., Wiener, H., Loukeris, K., Ou, G., Philips, M. R., and Pillinger, M. H. (2004) Carboxyl methylation of Ras regulates membrane targeting and effector engagement. *J Biol Chem* **279**, 7346-7352

46. Yang, J., Kulkarni, K., Manolaridis, I., Zhang, Z., Dodd, R. B., Mas-Droux, C., and Barford, D. (2011) Mechanism of isoprenylcysteine carboxyl methylation from the crystal structure of the integral membrane methyltransferase ICMT. *Mol Cell* **44**, 997-1004

47. Diver, M. M., Pedi, L., Koide, A., Koide, S., and Long, S. B. (2018) Atomic structure of the eukaryotic intramembrane RAS methyltransferase ICMT. *Nature* **553**, 526-529

48. Svensson, A. W., Casey, P. J., Young, S. G., and Bergo, M. O. (2006) Genetic and pharmacologic analyses of the role of Icmt in Ras membrane association and function. *Methods Enzymol* **407**, 144-159

49. Zhang, F. L., Kirschmeier, P., Carr, D., James, L., Bond, R. W., Wang, L., Patton, R., Windsor, W. T., Syto, R., Zhang, R., and Bishop, W. R. (1997) Characterization of Ha-ras, N-ras, Ki-Ras4A, and Ki-Ras4B as in vitro substrates for farnesyl protein transferase and geranylgeranyl protein transferase type I. *J Biol Chem* **272**, 10232-10239

50. Whyte, D. B., Kirschmeier, P., Hockenberry, T. N., Nunez-Oliva, I., James, L., Catino, J. J., Bishop, W. R., and Pai, J. K. (1997) K- and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors. *J Biol Chem* **272**, 14459-14464

51. Henriksen, B. S., Anderson, J. L., Hrycyna, C. A., and Gibbs, R. A. (2005) Synthesis of desthio prenylcysteine analogs: sulfur is important for biological activity. *Bioorg Med Chem Lett* **15**, 5080-5083

52. Majmudar, J. D., Hahne, K., Hrycyna, C. A., and Gibbs, R. A. (2011) Probing the isoprenylcysteine carboxyl methyltransferase (Icmt) binding pocket: sulfonamide modified farnesyl cysteine (SMFC) analogs as Icmt inhibitors. *Bioorg Med Chem Lett* **21**, 2616-2620

53. Butler, K. V., Bohn, K., Hrycyna, C. A., and Jin, J. (2016) Non-Substrate Based, Small Molecule Inhibitors of the Human Isoprenylcysteine Carboxyl Methyltransferase. *Medchemcomm* **7**, 1016-1021

54. Bergman, J. A., Hahne, K., Hrycyna, C. A., and Gibbs, R. A. (2011) Lipid and sulfur substituted prenylcysteine analogs as human Icmt inhibitors. *Bioorg Med Chem Lett* **21**, 5616-5619

55. Lau, H. Y., Ramanujulu, P. M., Guo, D., Yang, T., Wirawan, M., Casey, P. J., Go, M. L., and Wang, M. (2014) An improved isoprenylcysteine carboxylmethyltransferase inhibitor induces cancer cell death and attenuates tumor growth in vivo. *Cancer Biol Ther* **15**, 1280-1291

56.   Marín-Ramos, N. I., Balabasquer, M., Ortega-Nogales, F. J., Torrecillas, I. R., Gil-Ordóñez, A., Marcos-Ramiro, B., Aguilar-Garrido, P., Cushman, I., Romero, A., Medrano, F. J., Gajate, C., Mollinedo, F., Philips, M. R., Campillo, M., Gallardo, M., Martín-Fontecha, M., López-Rodríguez, M. L., and Ortega-Gutiérrez, S. (2019) A Potent Isoprenylcysteine Carboxylmethyltransferase (ICMT) Inhibitor Improves Survival in Ras-Driven Acute Myeloid Leukemia. *J Med Chem* **62**, 6035-6046

57.   Lau, H. Y., Tang, J., Casey, P. J., and Wang, M. (2017) Isoprenylcysteine carboxylmethyltransferase is critical for malignant transformation and tumor maintenance by all RAS isoforms. *Oncogene* **36**, 3934-3942

58.   Hetzer, M. W. (2010) The nuclear envelope. *Cold Spring Harb Perspect Biol* **2**, a000539

59.   Phillips, M. J., and Voeltz, G. K. (2016) Structure and function of ER membrane contact sites with other organelles. *Nat Rev Mol Cell Biol* **17**, 69-82

60.   Nordeen, S. A., Andersen, K. R., Knockenhauer, K. E., Ingram, J. R., Ploegh, H. L., and Schwartz, T. U. (2020) A nanobody suite for yeast scaffold nucleoporins provides details of the nuclear pore complex structure. *Nat Commun* **11**, 6179

61.   Doye, V., and Hurt, E. (1997) From nucleoporins to nuclear pore complexes. *Curr Opin Cell Biol* **9**, 401-411

62.   Panté, N., and Kann, M. (2002) Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Mol Biol Cell* **13**, 425-434

63.   Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002) The Transport of Molecules between the Nucleus and the Cytosol. in *Molecular Biology of the Cell*, 4th Edition Ed., Garland Science, New York, NY. pp

64.   Vlcek, S., and Foisner, R. (2007) Lamins and lamin-associated proteins in aging and disease. *Curr Opin Cell Biol* **19**, 298-304

65.   Dorner, D., Vlcek, S., Foeger, N., Gajewski, A., Makolm, C., Gotzmann, J., Hutchison, C. J., and Foisner, R. (2006) Lamina-associated polypeptide 2alpha regulates cell cycle progression and differentiation via the retinoblastoma-E2F pathway. *J Cell Biol* **173**, 83-93

66.   Pekovic, V., Harborth, J., Broers, J. L., Ramaekers, F. C., van Engelen, B., Lammens, M., von Zglinicki, T., Foisner, R., Hutchison, C., and Markiewicz, E. (2007) Nucleoplasmic LAP2alpha-lamin A complexes are required to maintain a proliferative state in human fibroblasts. *J Cell Biol* **176**, 163-172

67.   Starr, D. A. (2011) KASH and SUN proteins. *Curr Biol* **21**, R414-415

68.   Andrés, V., and González, J. M. (2009) Role of A-type lamins in signaling, transcription, and chromatin organization. *J Cell Biol* **187**, 945-957

69.   Butin-Israeli, V., Adam, S. A., Goldman, A. E., and Goldman, R. D. (2012) Nuclear lamin functions and disease. *Trends Genet* **28**, 464-471

70.   Capo-chichi, C. D., Cai, K. Q., Smedberg, J., Ganjei-Azar, P., Godwin, A. K., and Xu, X. X. (2011) Loss of A-type lamin expression compromises nuclear envelope integrity in breast cancer. *Chin J Cancer* **30**, 415-425

71.   Dittmer, T. A., and Misteli, T. (2011) The lamin protein family. *Genome Biol* **12**, 222

72.   Barboro, P., Repaci, E., D'Arrigo, C., and Balbi, C. (2012) The role of nuclear matrix proteins binding to matrix attachment regions (Mars) in prostate cancer cell differentiation. *PLoS One* **7**, e40617

73.   Harr, J. C., Luperchio, T. R., Wong, X., Cohen, E., Wheelan, S. J., and Reddy, K. L. (2015) Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biol* **208**, 33-52

74.   Hutchison, C. J. (2014) B-type lamins in health and disease. *Semin Cell Dev Biol* **29**, 158-163

75.   Sakthivel, K. M., and Sehgal, P. (2016) A Novel Role of Lamins from Genetic Disease to Cancer Biomarkers. *Oncol Rev* **10**, 309

76.   Verstraeten, V. L., Ji, J. Y., Cummings, K. S., Lee, R. T., and Lammerding, J. (2008) Increased mechanosensitivity and nuclear stiffness in Hutchinson-Gilford progeria cells: effects of farnesyltransferase inhibitors. *Aging Cell* **7**, 383-393

77.   Naetar, N., Ferraioli, S., and Foisner, R. (2017) Lamins in the nuclear interior - life outside the lamina. *J Cell Sci* **130**, 2087-2096

78.   Brueckner, L., Zhao, P. A., van Schaik, T., Leemans, C., Sima, J., Peric-Hupkes, D., Gilbert, D. M., and van Steensel, B. (2020) Local rewiring of genome-nuclear lamina interactions by transcription. *EMBO J* **39**, e103159

79.   Leemans, C., van der Zwalm, M. C. H., Brueckner, L., Comoglio, F., van Schaik, T., Pagie, L., van Arensbergen, J., and van Steensel, B. (2019) Promoter-Intrinsic and Local Chromatin Features Determine Gene Repression in LADs. *Cell* **177**, 852-864.e814

80.   Meuleman, W., Peric-Hupkes, D., Kind, J., Beaudry, J. B., Pagie, L., Kellis, M., Reinders, M., Wessels, L., and van Steensel, B. (2013) Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **23**, 270-280

81. Wu, F., and Yao, J. (2017) Identifying Novel Transcriptional and Epigenetic Features of Nuclear Lamina-associated Genes. *Sci Rep* **7**, 100

82. Wang, N., Tytell, J. D., and Ingber, D. E. (2009) Mechanotransduction at a distance: mechanically coupling the extracellular matrix with the nucleus. *Nat Rev Mol Cell Biol* **10**, 75-82

83. Goldman, R. D., Shumaker, D. K., Erdos, M. R., Eriksson, M., Goldman, A. E., Gordon, L. B., Gruenbaum, Y., Khuon, S., Mendez, M., Varga, R., and Collins, F. S. (2004) Accumulation of mutant lamin A causes progressive changes in nuclear architecture in Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A* **101**, 8963-8968

84. Tsai, M. Y., Wang, S., Heidinger, J. M., Shumaker, D. K., Adam, S. A., Goldman, R. D., and Zheng, Y. (2006) A mitotic lamin B matrix induced by RanGTP required for spindle assembly. *Science* **311**, 1887-1893

85. Karoutas, A., and Akhtar, A. (2021) Functional mechanisms and abnormalities of the nuclear lamina. *Nat Cell Biol* **23**, 116-126

86. Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S. S., Janssen, H., Amendola, M., Nolen, L. D., Bickmore, W. A., and van Steensel, B. (2013) Single-cell dynamics of genome-nuclear lamina interactions. *Cell* **153**, 178-192

87. Harborth, J., Elbashir, S. M., Bechert, K., Tuschl, T., and Weber, K. (2001) Identification of essential genes in cultured mammalian cells using small interfering RNAs. *J Cell Sci* **114**, 4557-4565

88. Maske, C. P., Hollinshead, M. S., Higbee, N. C., Bergo, M. O., Young, S. G., and Vaux, D. J. (2003) A carboxyl-terminal interaction of lamin B1 is dependent on the CAAX endoprotease Rce1 and carboxymethylation. *J Cell Biol* **162**, 1223-1232

89. Sun, S., Xu, M. Z., Poon, R. T., Day, P. J., and Luk, J. M. (2010) Circulating Lamin B1 (LMNB1) biomarker detects early stages of liver cancer in patients. *J Proteome Res* **9**, 70-78

90. Wong, K. F., and Luk, J. M. (2012) Discovery of lamin B1 and vimentin as circulating biomarkers for early hepatocellular carcinoma. *Methods Mol Biol* **909**, 295-310

91. Wazir, U., Ahmed, M. H., Bridger, J. M., Harvey, A., Jiang, W. G., Sharma, A. K., and Mokbel, K. (2013) The clinicopathological significance of lamin A/C, lamin B1 and lamin B receptor mRNA expression in human breast cancer. *Cell Mol Biol Lett* **18**, 595-611

92. Liu, L., Wang, J., Shi, L., Zhang, W., Du, X., Wang, Z., and Zhang, Y. (2013) β-Asarone induces senescence in colorectal cancer cells by inducing lamin B1 expression. *Phytomedicine* **20**, 512-520

93.   Lin, S. T., and Fu, Y. H. (2009) miR-23 regulation of lamin B1 is crucial for oligodendrocyte development and myelination. *Dis Model Mech* **2**, 178-188

94.   Lin, S. T., Ptácek, L. J., and Fu, Y. H. (2011) Adult-onset autosomal dominant leukodystrophy: linking nuclear envelope to myelin. *J Neurosci* **31**, 1163-1166

95.   Rieckher, M., Garinis, G. A., and Schumacher, B. (2021) Molecular pathology of rare progeroid diseases. *Trends Mol Med*

96.   Ben Yaou, R., Navarro, C., Quijano-Roy, S., Bertrand, A. T., Massart, C., De Sandre-Giovannoli, A., Cadiñanos, J., Mamchaoui, K., Butler-Browne, G., Estournet, B., Richard, P., Barois, A., Lévy, N., and Bonne, G. (2011) Type B mandibuloacral dysplasia with congenital myopathy due to homozygous ZMPSTE24 missense mutation. *Eur J Hum Genet* **19**, 647-654

97.   Agarwal, A. K., Kazachkova, I., Ten, S., and Garg, A. (2008) Severe mandibuloacral dysplasia-associated lipodystrophy and progeria in a young girl with a novel homozygous Arg527Cys LMNA mutation. *J Clin Endocrinol Metab* **93**, 4617-4623

98.   Miyoshi, Y., Akagi, M., Agarwal, A. K., Namba, N., Kato-Nishimura, K., Mohri, I., Yamagata, M., Nakajima, S., Mushiake, S., Shima, M., Auchus, R. J., Taniike, M., Garg, A., and Ozono, K. (2008) Severe mandibuloacral dysplasia caused by novel compound heterozygous ZMPSTE24 mutations in two Japanese siblings. *Clin Genet* **73**, 535-544

99.   Cenni, V., D'Apice, M. R., Garagnani, P., Columbaro, M., Novelli, G., Franceschi, C., and Lattanzi, G. (2018) Mandibuloacral dysplasia: A premature ageing disease with aspects of physiological ageing. *Ageing Res Rev* **42**, 1-13

100.   Barrowman, J., Wiley, P. A., Hudon-Miller, S. E., Hrycyna, C. A., and Michaelis, S. (2012) Human ZMPSTE24 disease mutations: residual proteolytic activity correlates with disease severity. *Hum Mol Genet* **21**, 4084-4093

101.   Dreesen, O., and Stewart, C. L. (2011) Accelerated aging syndromes, are they relevant to normal human aging? *Aging (Albany NY)* **3**, 889-895

102.   McClintock, D., Ratner, D., Lokuge, M., Owens, D. M., Gordon, L. B., Collins, F. S., and Djabali, K. (2007) The mutant form of lamin A that causes Hutchinson-Gilford progeria is a biomarker of cellular aging in human skin. *PLoS One* **2**, e1269

103.   Eriksson, M., Brown, W. T., Gordon, L. B., Glynn, M. W., Singer, J., Scott, L., Erdos, M. R., Robbins, C. M., Moses, T. Y., Berglund, P., Dutra, A., Pak, E., Durkin, S., Csoka, A. B., Boehnke, M., Glover, T. W., and Collins, F. S. (2003) Recurrent de novo point mutations in lamin A cause Hutchinson-Gilford progeria syndrome. *Nature* **423**, 293-298

104.   Young, S. G., Fong, L. G., and Michaelis, S. (2005) Prelamin A, Zmpste24, misshapen cell nuclei, and progeria--new evidence suggesting that protein farnesylation could be important for disease pathogenesis. *J Lipid Res* **46**, 2531-2558

105.   Yang, S. H., Andres, D. A., Spielmann, H. P., Young, S. G., and Fong, L. G. (2008) Progerin elicits disease phenotypes of progeria in mice whether or not it is farnesylated. *J Clin Invest* **118**, 3291-3300

106.   Gordon, L. B., Kleinman, M. E., Massaro, J., D'Agostino, R. B., Shappell, H., Gerhard-Herman, M., Smoot, L. B., Gordon, C. M., Cleveland, R. H., Nazarian, A., Snyder, B. D., Ullrich, N. J., Silvera, V. M., Liang, M. G., Quinn, N., Miller, D. T., Huh, S. Y., Dowton, A. A., Littlefield, K., Greer, M. M., and Kieran, M. W. (2016) Clinical Trial of the Protein Farnesylation Inhibitors Lonafarnib, Pravastatin, and Zoledronic Acid in Children With Hutchinson-Gilford Progeria Syndrome. *Circulation* **134**, 114-125

107.   Gordon, L. B., Massaro, J., D'Agostino, R. B., Campbell, S. E., Brazier, J., Brown, W. T., Kleinman, M. E., Kieran, M. W., and Collaborative, P. C. T. (2014) Impact of farnesylation inhibitors on survival in Hutchinson-Gilford progeria syndrome. *Circulation* **130**, 27-34

108.   Davies, B. S., Barnes, R. H., Tu, Y., Ren, S., Andres, D. A., Spielmann, H. P., Lammerding, J., Wang, Y., Young, S. G., and Fong, L. G. (2010) An accumulation of non-farnesylated prelamin A causes cardiomyopathy but not progeria. *Hum Mol Genet* **19**, 2682-2694

109.   Varela, I., Pereira, S., Ugalde, A. P., Navarro, C. L., Suárez, M. F., Cau, P., Cadiñanos, J., Osorio, F. G., Foray, N., Cobo, J., de Carlos, F., Lévy, N., Freije, J. M., and López-Otín, C. (2008) Combined treatment with statins and aminobisphosphonates extends longevity in a mouse model of human premature aging. *Nat Med* **14**, 767-772

110.   Gordon, L. B., Kleinman, M. E., Miller, D. T., Neuberg, D. S., Giobbie-Hurder, A., Gerhard-Herman, M., Smoot, L. B., Gordon, C. M., Cleveland, R., Snyder, B. D., Fligor, B., Bishop, W. R., Statkevich, P., Regen, A., Sonis, A., Riley, S., Ploski, C., Correia, A., Quinn, N., Ullrich, N. J., Nazarian, A., Liang, M. G., Huh, S. Y., Schwartzman, A., and Kieran, M. W. (2012) Clinical trial of a farnesyltransferase inhibitor in children with Hutchinson-Gilford progeria syndrome. *Proc Natl Acad Sci U S A* **109**, 16666-16671

111.   Guida, V., Giovannetti, A., Piceci-Sparascio, F., Mazza, T., Marchionni, E., Menale, L., Fatigante, G., Hozhabri, H., Traversa, A., Caputo, V., Pizzuti, A., Ferraris, A., and De Luca, A. (2019) Restrictive Dermopathy: Novel *ZMPSTE24* mutation and clues for prenatal diagnosis. *Annals of Biotechnology* **2**

112.   Morais, P., Magina, S., Ribeiro, M. o. C., Rodrigues, M., Lopes, J. M., Thanh, H. e. T., Wehnert, M., and Guimarães, H. (2009) Restrictive dermopathy--a lethal congenital laminopathy. Case report and review of the literature. *Eur J Pediatr* **168**, 1007-1012

113.   Thill, M., Nguyen, T. D., Wehnert, M., Fischer, D., Hausser, I., Braun, S., and Jackisch, C. (2008) Restrictive dermopathy: a rare laminopathy. *Arch Gynecol Obstet* **278**, 201-208

114.	Moulson, C. L., Go, G., Gardner, J. M., van der Wal, A. C., Smitt, J. H., van Hagen, J. M., and Miner, J. H. (2005) Homozygous and compound heterozygous mutations in ZMPSTE24 cause the laminopathy restrictive dermopathy. *J Invest Dermatol* **125**, 913-919

115.	Paige, D. G. (2020) Hyalinoses, Stiff Skin Syndrom and Restrictive Dermopathy. in *Harper's Textbook of Pediatric Dermatology* (Hoeger, P., Kinsler, V., and Yan, A. eds.), 4 Ed., John Wiley & Sons LTD. pp 1164-1171

116.	Hudon, S. E., Coffinier, C., Michaelis, S., Fong, L. G., Young, S. G., and Hrycyna, C. A. (2008) HIV-protease inhibitors block the enzymatic activity of purified Ste24p. *Biochem Biophys Res Commun* **374**, 365-368

117.	Coffinier, C., Hudon, S. E., Farber, E. A., Chang, S. Y., Hrycyna, C. A., Young, S. G., and Fong, L. G. (2007) HIV protease inhibitors block the zinc metalloproteinase ZMPSTE24 and lead to an accumulation of prelamin A in cells. *Proc Natl Acad Sci U S A* **104**, 13432-13437

118.	Debouck, C. (1992) The HIV-1 protease as a therapeutic target for AIDS. *AIDS Res Hum Retroviruses* **8**, 153-164

119.	Flexner, C. (1998) HIV-protease inhibitors. *N Engl J Med* **338**, 1281-1292

120.	Kayatekin, C., Amasino, A., Gaglia, G., Flannick, J., Bonner, J. M., Fanning, S., Narayan, P., Barrasa, M. I., Pincus, D., Landgraf, D., Nelson, J., Hesse, W. R., Costanzo, M., Myers, C. L., Boone, C., Florez, J. C., Lindquist, S., and Consortium, A. T. D.-G. (2018) Translocon Declogger Ste24 Protects against IAPP Oligomer-Induced Proteotoxicity. *Cell* **173**, 62-73.e69

121.	Ast, T., Michaelis, S., and Schuldiner, M. (2016) The Protease Ste24 Clears Clogged Translocons. *Cell* **164**, 103-114

122.	Brender, J. R., Salamekh, S., and Ramamoorthy, A. (2012) Membrane disruption and early events in the aggregation of the diabetes related peptide IAPP from a molecular perspective. *Acc Chem Res* **45**, 454-462

123.	Narita, R., Toshimori, H., Nakazato, M., Kuribayashi, T., Toshimori, T., Kawabata, K., Takahashi, K., and Masukura, S. (1992) Islet amyloid polypeptide (IAPP) and pancreatic islet amyloid deposition in diabetic and non-diabetic patients. *Diabetes Res Clin Pract* **15**, 3-14

124.	Permert, J., Larsson, J., Westermark, G. T., Herrington, M. K., Christmanson, L., Pour, P. M., Westermark, P., and Adrian, T. E. (1994) Islet amyloid polypeptide in patients with pancreatic cancer and diabetes. *N Engl J Med* **330**, 313-318

125.   Maloy, A. L., Longnecker, D. S., and Greenberg, E. R. (1981) The relation of islet amyloid to the clinical type of diabetes. *Hum Pathol* **12**, 917-922

126.   Fu, B., Wang, L., Li, S., and Dorf, M. E. (2017) ZMPSTE24 defends against influenza and other pathogenic viruses. *J Exp Med* **214**, 919-929

127.   Han, M., and Pandey, D. (2021) ZMPSTE24 Regulates SARS-CoV-2 Spike Protein-enhanced Expression of Endothelial Plasminogen Activator Inhibitor-1. *Am J Respir Cell Mol Biol*

128.   Mokry, D. Z., Manandhar, S. P., Chicola, K. A., Santangelo, G. M., and Schmidt, W. K. (2009) Heterologous expression studies of Saccharomyces cerevisiae reveal two distinct trypanosomatid CaaX protease activities and identify their potential targets. *Eukaryot Cell* **8**, 1891-1900

129.   Babatz, T. D., Spear, E. D., Xu, W., Sun, O. L., Nie, L., Carpenter, E. P., and Michaelis, S. (2021) Site specificity determinants for prelamin A cleavage by the zinc metalloprotease ZMPSTE24. *Journal of Biological Chemistry* **296**, 1-12

130.   Hildebrandt, E. R., Arachea, B. T., Wiener, M. C., and Schmidt, W. K. (2016) Ste24p Mediates Proteolysis of Both Isoprenylated and Non-prenylated Oligopeptides. *J Biol Chem* **291**, 14185-14198

131.   Murale, D. P., Hong, S. C., Haque, M. M., and Lee, J. S. (2016) Photo-affinity labeling (PAL) in chemical proteomics: a handy tool to investigate protein-protein interactions (PPIs). *Proteome Sci* **15**, 14

132.   Vervacke, J. S., Funk, A. L., Wang, Y. C., Strom, M., Hrycyna, C. A., and Distefano, M. D. (2014) Diazirine-containing photoactivatable isoprenoid: synthesis and application in studies with isoprenylcysteine carboxyl methyltransferase. *J Org Chem* **79**, 1971-1978

133.   Seidel, S. A., Dijkman, P. M., Lea, W. A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J. S., Srinivasan, P., Baaske, P., Simeonov, A., Katritch, I., Melo, F. A., Ladbury, J. E., Schreiber, G., Watts, A., Braun, D., and Duhr, S. (2013) Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* **59**, 301-315

134.   Seidel, S. A., Wienken, C. J., Geissler, S., Jerabek-Willemsen, M., Duhr, S., Reiter, A., Trauner, D., Braun, D., and Baaske, P. (2012) Label-free microscale thermophoresis discriminates sites and affinity of protein-ligand binding. *Angew Chem Int Ed Engl* **51**, 10656-10659

135.   Ling, M. M., and Robinson, B. H. (1997) Approaches to DNA mutagenesis: an overview. *Anal Biochem* **254**, 157-178

136. Elble, R. (1992) A simple and efficient procedure for transformation of yeasts. *Biotechniques* **13**, 18-20

137. Sedmak, J. J., and Grossberg, S. E. (1977) A rapid, sensitive, and versatile assay for protein using Coomassie brilliant blue G250. *Anal Biochem* **79**, 544-552

138. Schaffner, W., and Weissmann, C. (1973) A rapid, sensitive, and specific method for the determination of protein in dilute solution. *Anal Biochem* **56**, 502-514

139. Tam, A., Schmidt, W. K., and Michaelis, S. (2001) The multispanning membrane protein Ste24p catalyzes CAAX proteolysis and NH2-terminal processing of the yeast a-factor precursor. *J Biol Chem* **276**, 46798-46806

140. Fujimura-Kamada, K., Nouvet, F. J., and Michaelis, S. (1997) A novel membrane-associated metalloprotease, Ste24p, is required for the first step of NH2-terminal processing of the yeast a-factor precursor. *J Cell Biol* **136**, 271-285

141. Holland, D. R., Hausrath, A. C., Juers, D., and Matthews, B. W. (1995) Structural analysis of zinc substitutions in the active site of thermolysin. *Protein Sci* **4**, 1955-1965

142. Tamaki, F. K., Textor, L. C., Polikaropv, I., and Marana, S. R. (2014) Sets of Covariant Residues Modulate the Activity and Thermal Stability of GH1 ß-Glucosidases.

143. Mezzasalma, T. M., Kranz, J. K., Chan, W., Struble, G. T., Schalk-Hihi, C., Deckman, I. C., Springer, B. A., and Todd, M. J. (2007) Enhancing recombinant protein quality and yield by protein stability profiling. *J Biomol Screen* **12**, 418-428

144. Kitten, G. T., and Nigg, E. A. (1991) The CaaX motif is required for isoprenylation, carboxyl methylation, and nuclear membrane association of lamin B2. *J Cell Biol* **113**, 13-23

145. Mehmood, S., Marcoux, J., Gault, J., Quigley, A., Michaelis, S., Young, S. G., Carpenter, E. P., and Robinson, C. V. (2016) Mass spectrometry captures off-target drug binding and provides mechanistic insights into the human metalloprotease ZMPSTE24. *Nat Chem* **8**, 1152-1158

146. Robinette, D., Neamati, N., Tomer, K. B., and Borchers, C. H. (2006) Photoaffinity labeling combined with mass spectrometric approaches as a tool for structural proteomics. *Expert Rev Proteomics* **3**, 399-408

147. Cusick, M. E., Klitgord, N., Vidal, M., and Hill, D. E. (2005) Interactome: gateway into systems biology. *Hum Mol Genet* **14 Spec No. 2**, R171-181

148. Connelly, K. E., Hedrick, V., Paschoal Sobreira, T. J., Dykhuizen, E. C., and Aryal, U. K. (2018) Analysis of Human Nuclear Protein Complexes by Quantitative Mass Spectrometry Profiling. *Proteomics* **18**, e1700427

149.   Zembroski, A. S., Buhman, K. K., and Aryal, U. K. (2021) Proteome and phosphoproteome characterization of liver in the postprandial state from diet-induced obese and lean mice. *J Proteomics* **232**, 104072

150.   Mohallem, R., and Aryal, U. K. (2020) Regulators of TNFα mediated insulin resistance elucidated by quantitative proteomics. *Sci Rep* **10**, 20878

151.   Guo, L. W., Hajipour, A. R., Gavala, M. L., Arbabian, M., Martemyanov, K. A., Arshavsky, V. Y., and Ruoho, A. E. (2005) Sulfhydryl-reactive, cleavable, and radioiodinatable benzophenone photoprobes for study of protein-protein interaction. *Bioconjug Chem* **16**, 685-693

152.   Flexner, C., Broyles, S. S., Earl, P., Chakrabarti, S., and Moss, B. (1988) Characterization of human immunodeficiency virus gag/pol gene products expressed by recombinant vaccinia viruses. *Virology* **166**, 339-349

153.   Garg, A. (2011) Clinical review#: Lipodystrophies: genetic and acquired body fat disorders. *J Clin Endocrinol Metab* **96**, 3313-3325

154.   Caron, M., Auclair, M., Donadille, B., Béréziat, V., Guerci, B., Laville, M., Narbonne, H., Bodemer, C., Lascols, O., Capeau, J., and Vigouroux, C. (2007) Human lipodystrophies linked to mutations in A-type lamins and to HIV protease inhibitor therapy are both associated with prelamin A accumulation, oxidative stress and premature cellular senescence. *Cell Death Differ* **14**, 1759-1767

155.   Caron, M., Auclair, M., Sterlingot, H., Kornprobst, M., and Capeau, J. (2003) Some HIV protease inhibitors alter lamin A/C maturation and stability, SREBP-1 nuclear localization and adipocyte differentiation. *AIDS* **17**, 2437-2444

VITA

Education:
Purdue University, West Lafayette IN
Ph.D., Chemistry, Biochemistry Division, May 2021
> <u>Dissertation Topic:</u> "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site" <u>Advisor</u>: Dr. Christine Hrycyna, Professor of Chemistry, Department Head of Chemistry Administration <u>Relevant Courses</u>: Macromolecules, Scientific Communications, Principle Chemical Action on Biological Systems, Neural Systems, Biophysical Chemistry, Responsible Conduct of Research, Membrane: Structure/Function, Teaching and Research in Chemistry, Seminar in Biochemistry, GrantWriting, Membrane Proteins
> <u>GPA 4.0 / 4.0</u>

Lake Superior State University, Sault Ste. Marie, MI
B.S., Biochemistry, April 2015, Summa Cum Laude
> <u>Dean's List:</u> 2011-2015, all semesters<u>GPA 3.99 / 4.0</u>
> <u>Undergraduate Thesis Topic:</u> "Synthesis of N-(2-Haloethyl)amides and 1,2-Disubstituted 2-Imidazolines"
> <u>Advisor:</u> Dr. R. Adam Mosey, School of Physical Sciences, Assistant Professor

Work and Research Experience:
Purdue University, West Lafayette, IN
Biochemistry Graduate Research Assistant   Fall 2015
Present Advisor: Dr. Christine Hrycyna
> *Project*: Mapping the Binding Site of the *S. cerevisiae* Transmembrane Zinc Metalloprotease, Ste24
> - Performed numerous PCR mediated, site-directed mutagenesis to prepare a library of mutantproteins
> - Characterized protein activity by utilizing an *in vitro* radioactive assay
> - Determined substrate binding ability by use of photolabeling probes and microscale thermophoresis
> - Developed a protocol for identifying substrate binding location by tandem MS/MS experiments

Lake Superior State University, Sault Ste. Marie, MI
Undergraduate Researcher   Summer 2014
Advisor: Dr. R. Adam Mosey
> *Project:* Synthesis of N-(2-Haloethyl)amides and 1,2-Disubstituted 2-Imidazolines
> - Performed organic synthesis to create novel compounds
> - Completed flash chromatography and TLC to isolate compounds of interest
> - Utilized ^{1}H- and ^{13}C-NMR, FT-IR, and high-resolution mass spectrometry to characterize newlysynthesized compounds

Michigan State University, East Lansing, MI
REU Plant Genomics   Summer 2013
Advisor: Dr. Kevin D. Walker

> *Project:* Use of Bacteria Engineered with an Aminomutase to Produce Non-natural β-Amino Acids *in vivo*
>
> - Optimized conditions to maximize biochemical reaction efficiency *in vivo*
> - Utilized GC/EI-MS to recognize and quantify levels of various non-natural ß-aryl amino acids
> - Participated in oral and poster presentation at the end of the REU program

Lake Superior State University, Sault Ste. Marie, MI
Supplemental Instructor/Tutor/Manager   Fall 2011 – Spring 2015
Learning Center

> *Overview*: Worked in conjunction with Lake Superior State University to provide additional informationand tutoring for students
>
> - Provided students with opportunities to ask questions and seek additional practice in topics fortheir biology and chemistry courses
> - Maintained records of meetings and events including tutoring appointments, training schedulesand additional student learning opportunities
> - Collaborated with professors to ensure students were learning the appropriate level of detail fortheir assignments

Publications:

Ellsworth, A., Magyar, C., Hubbell, G., **Theisen, C.**, Holmes, D., Mosey, R. "One-pot triflic anhydride-mediated synthesis of 1,2-disubstituted 2-imidazolines from N-(2-haloethyl)amides and amines." *Tetrahedron* 2016, pg. 6380-6389.

Ratnayake, N., **Theisen, C.**, Walter, T., Walker, K. "Whole-cell biocatalytic production of variouslysubstituted ß-aryl- and ß-heteroaryl-ß-amino acids." *Journal of Biotechnology* 2016, pg. 12-21.

Presentations:

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (May 2021) "Characterization of the Binding Site of the Zinc Metalloprotease" <u>Oral</u> at the Hitchhiker's Guide to the Biomolecular Galaxy Symposium, West Lafayette, IN. <u>Best speaker award</u>.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (November 2020) "Characterization of the Binding Site of the Zinc Metalloprotease" <u>Oral</u> at the Biochemistry Seminar in the Chemistry Department at Purdue University, West Lafayette, IN. <u>Joseph F. Foster award for best presentation</u>.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (July 2019) "Mapping of the Proposed Substrate Binding Site of the Zinc Metalloprotease, Ste24" Poster presentation at the Experimental Biology Conference, Orlando, FL.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (January 2019) "Mapping of the Proposed Substrate Binding Site of the Zinc Metalloprotease, Ste24" Poster presentation at the Experimental Biology Conference, Orlando, FL.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (January 2019) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site" <u>Oral</u> presentation at the PULSE Welcome Session week 2, West Lafayette, IN.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (January 2019) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site" <u>Oral</u> presentation at the PULSe Welcome Session Week 1 5-minute Thesis competition, West Lafayette, IN.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (May 2018) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site" <u>Oral</u> presentation at the Hitchhiker's Guide to the Biomolecular Galaxy symposium, West Lafayette, IN.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (February 2018) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site" Poster presentation at the Health and Disease poster session, West Lafayette, IN.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C.A. (July 2017) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site". Poster presentation at the FASEBResearch Conference, Saxton River, VT.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C. A. (May 2017) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site". Poster presentation at the Hitchhiker's Guide to the Biomolecular Galaxy Conference, West Lafayette, IN.

Theisen, C., Hsu, E., Wiley, P., Vervacke, J., Distefano, M., Hrycyna, C. A. (April 2017) "Mutational Analysis and Characterization of the Ste24p Substrate Binding Site". Poster presentation at the Office of Interdisciplinary Graduate Programs Conference, West Lafayette, IN.

Theisen, C., Heintz, V., Wang, L., LaCount, D. (May 2016) "Preparing Marburg Virus Protein Fragments for FutureHost-Pathogen Interactome Studies." Poster presentation at PULSe Spring Reception, West Lafayette,IN.

Theisen, C., Mosey, R. A. (April 2015) "Synthesis of *N*-(2-haloethyl)amides and 1,2-Disubstituted 2- Imidazolines". Poster and <u>oral</u> at Lake Superior State University Senior Thesis Symposium, Sault Ste.Marie, MI.

Theisen, C., Ratanayake, R. M. D., Walker, K. D. (July 2013) "Use of Bacteria Engineered with an Aminomutase to Produce Non-Natural ß-Amino Acids *In-Vivo*." Poster and <u>oral</u> presentation at MidSURE Poster Symposium and Plant Genomics at MSU Research Symposium, East Lansing, MI.

Teaching Experience:

Purdue University, West Lafayette IN
Lecture Supervisor, General Chemistry with a Biological Focus (CHM 129)   Fall 2018
- Supervised 16 other graduate teaching assistants to help organize grading schedules and communicate important information
- Worked alongside other supervisors to distribute workload and answer questions from teachingassistants and students
- Helped write and review exams, as well as provide alternative testing conditions, as needed bystudents

Teaching Assistant, General Chemistry with a Biological Focus (CHM 129)   Fall 2016 -2019
- Supervised approximately 50 students as they carried out prewritten laboratory assignments
- Analyzed and provided feedback on group assignments
- Held weekly office hours, laboratory periods and recitations

Teaching Assistant, Biochemistry, Laboratory (CHM 339)   Spring 2017-2019
- Supervised approximately 40 students as they performed experiments designed to mimic a biochemicalresearch project, including cloning, bacterial protein expression, purification, and enzyme activity
- Held weekly office hours, along with the weekly laboratory time
- Provided constructive criticism on a group final lab report
- Aided in grading of exams for the CHM 339 lecture course
- Attended weekly staff meetings to discuss problems discovered during the semester

Honors & Awards:
- Joseph F. Foster Award, Best Research Talk, Purdue University Department of Chemistry, 2020-2021
- Hitchhiker's Guide to the Biomolecular Galaxy Symposium Best Talk (3rd place), 2021
- PULSe 2020 Outstanding Graduate Student in Teaching Award, 2020
- Emerson Kampen Fellowship, Purdue University, 2019-2020

- Women in Science Travel Grant Award, Purdue University, 2018
- Outstanding Biochemistry Graduate, Lake Superior State University, 2015
- Margaret Haag Memorial Endowed Scholarship, Lake Superior State University, 2014-2015

Leadership and Communication Experience:

United Way of Greater Lafayette Event Organizer, 2019

An organization that helps raise funds for various groups around the Greater Lafayette area in Indiana. Groups that United Way funds include the American Red Cross, Big Brothers Big Sisters, Bauer Family Resources, and many more. Event organizers help advertise and recruit both attendeesand sponsorships to put on a range of money-raising events.

Hitchhiker's Guide to the Biomolecular Galaxy Symposium Event Planner, 2018-2019

A symposium at Purdue University that brings together students and professors from schools throughout the Midwest to present their work at a biomolecular conference. The event planners are a small group of graduate students that help to bring in funding for the symposium, organize catering, poster sessions and oral presentation schedules. They help to organize and schedule all details of theconference while staying within an appropriate monetary budget.

Purdue University Life Sciences Program Mentor, 2016-2017

A group of volunteers who help first year graduate students in the Purdue University Life Sciences program while they acclimate to the new environment. Volunteers provide advice and contacts from peers to aid the new student while he/she determines which classes to take, which programs to become involved in, which laboratories they should rotate in, and possibly provide aid in deciding which laboratory to join at the end of the first year

Skills and Techniques:

Biophysical and Analytical Techniques

Microscale Thermophoresis, targeted MSMS

Molecular Biology Techniques

PCR & oligonucleotide design, site-directed mutagenesis, DNA isolation and quantification, DNA recombination,dye-termination sequencing, bacterial and yeast transformation, mammalian, and insect cell culture training

Protein Biochemistry Techniques

Cell culturing of yeast and bacterial cell lines, expression of recombinant proteins, purification of membrane proteins, SDS-PAGE, Western blotting, gel staining, *In vitro* enzyme kinetic assay, *In vitro* photolabeling assay,trypsin digestion

Instrumentation

Microscale Thermophoresis (MST), High performance liquid chromatography (HPLC), ^{1}H- and ^{13}C-nuclear magnetic resonance spectroscopy (NMR),Gas Chromatography – Mass spectrometry (GC - MS), flash chromatography, Fourier-transform infrared spectroscopy (FT-IR), UV-Vis spectroscopy, Matrix-assisted laser desorption ionization-time of flight (MALDI- TOF/TOF), tandem MS/MS, microscale thermophoresis