BIOPHYSICAL CHARACTERIZATION OF UNUSUAL CLPS PROTEOLYTIC ADAPTERS

by

Jennifer Vorn

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Degree in Major with Distinction

Spring Semester 2020

BIOPHYSICAL CHARACTERIZATION OF

UNUSUAL CLPS PROTEOLYTIC ADAPTERS

by

Jennifer Vorn

Approved: _____

Karl Schmitz, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Vijay Parashar, Ph.D.
Committee member from the Department of Medical and Molecular Sciences

Approved: _____

Avron Abraham, Ph.D.
Committee member from the Board of Senior Thesis Readers

Approved: _____

Michael Chajes, Ph.D.
Chair of the University Committee on Student and Faculty Honors

ACKNOWLEDGMENTS

I want to first say a big thank you to Dr. Karl Schmitz for entrusting me to conduct research within his lab and also thank you for supporting me throughout my journey of writing this thesis manuscript and also my career in science. Thank you to my mentors Jialiu Jiang and Monkia Prorok for patiently guiding me throughout lab and always teaching me something new. Thank you to my graduate lab members; Patrick Beardslee, Emmanual Ogbonna, Gaury Dhamdhere, and Chris Presloid for always helping me out in lab and making sure I was not lost during experiments. Thank you to my undergrad lab members for supporting me throughout this project. Thank you to my significant other Evan, for constantly supporting me and being my number one cheerleader. Thank you to my friends and family for believing in me. Thank you to my committee members Dr. Vijay Parashar and Dr. Avron Abraham for taking the time to review my thesis and be a part of my committee. Thank you to my department, Medical and Molecular Science for allowing me to do research with the Schmitz Lab. Thank you to Delaware INBRE for allowing me to start this project and thank you to the University of Delaware for allowing me to continue the research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Bacterial ATP-dependent proteases contribute to cellular health by recognizing, unfolding, and hydrolyzing damaged and misfolded proteins. Within *Escherichia coli*, the ATP-dependent protease ClpAP works with the ClpS proteolytic adaptor to degrade proteins with specific N-terminal amino acids in a process called N-end rule degradation. *E. coli* ClpS has a special binding pocket with specific affinity for N-terminal Leu, Phe, Tyr, and Trp amino acids. ClpS binds to the N-domain of ClpA and delivers substrates to ClpAP for degradation. ClpS shares strong sequence conservation among many species of bacteria, but ClpS from the human pathogen *Helicobacter pylori* has unusual amino acid substitutions that may change the shape of the binding pocket. We hypothesize that *H. pylori* ClpS has altered specificity for N-terminal amino acids compared to *E. coli* ClpS. To test this, I cloned, expressed, and purified *H. pylori* ClpS. I carried out fluorescence anisotropy binding assays to test the preference of *H. pylori* ClpS for N-terminal amino acids. Proteolysis assays tested whether *H. pylori* ClpS functions as an adaptor for N-end rule proteolysis by ClpAP. I also set up crystallization trials of *H. pylori* ClpS to structurally characterize differences in the shape of its binding pocket that may explain its altered binding specificity. My results suggest that *H. pylori* ClpS does not recognize standard N-end rule amino acids. *H. pylori* ClpS can not function as an adaptor for *E. coli* ClpAP protease complex.

Chapter 1

INTRODUCTION

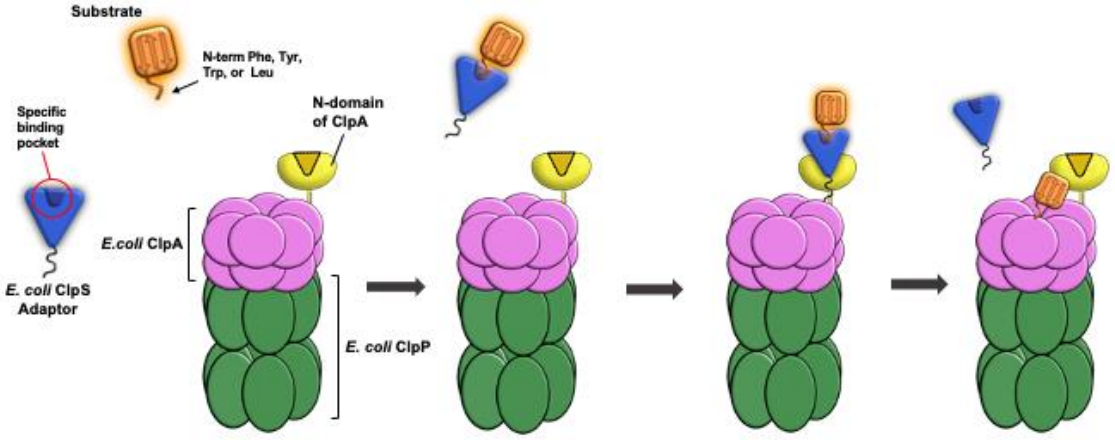
1.1 PROTEOLYTIC REGULATION WITHIN CELLS

Proteins are polymeric macromolecules composed of amino acid building blocks. All organisms produce thousands of distinct proteins that play an array of important roles in the cell. Proteins help define the cellular structure, catalyze chemical reactions, and regulate cellular pathways. The proteome – the composition of proteins within the cell – is dynamic[1]. New proteins are synthesized by ribosomes, and existing proteins are destroyed by ATP-dependent proteolytic complexes. In bacteria, intracellular proteolytic processes regulates different physiological processes, including elimination of misfolded proteins, response to environmental stress, and regulation of virulence phenotypes. [2]

Many different organisms possess a proteolytic program called the N-end rule pathway, which links proteolytic half-life to the identity of a protein's N-terminal amino acid.[3] N-end rule pathways are found in all domains of life, and have been studied in several organisms[2–4]. Chloroplasts have a version of the N-end rule pathway in which charged residues and large hydrophobic residues at the N-terminus of a substrate are avoided by the ClpS adaptor. [4] Eukaryotic cells possess a complex N-end rule pathway in which different N-terminal amino acids are recognized under different conditions, leading to covalent modification with polyubiquitin to mark substrates for degradation.[2–5] However, the N-end rule pathway has been most extensively studied is that of the model bacterium *Escherichia coli*.

1.2 THE N-END RULE PATHWAY AND PROTEIN STRUCURAL COMPONENTS

The N-end rule pathway in *Escherichia coli* involves an ATP-fueled protease complex called ClpAP (**Figure 1.2**). ClpA is an unfoldase enzyme that forms a hexamer consisting of two stacked rings in the presence of ATP (**Figure 1.2**). Proteolytic substrates are gripped by a network of loops in the ClpA axial pore (**Figure 1.2**). Rounds of ATP hydrolysis in ClpA subunits drive conformational changes in the ring that mechanically unfold protein substrate and translocate the unfolded polypeptide through ClpA and into the associated peptidase, ClpP.[5] ClpP is composed of two stacked heptamers that create a barrel-shaped complex (**Figure 1.2**). The solvent-filled interior of the ClpP barrel is lined with 14 peptidase active sites that degrade protein substrates into short peptide products.[3]



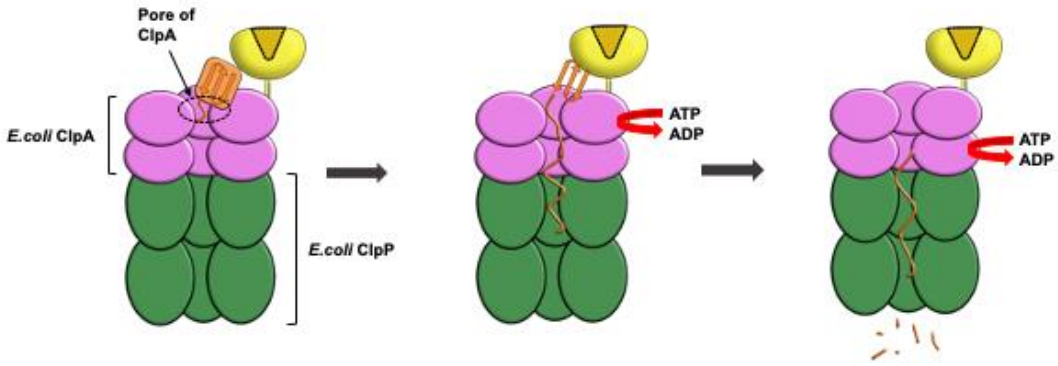


Figure 1.2 The N-end rule pathway and degradation of N-terminal substrates. The N-end rule pathway in *E. coli* dictates the half-life of substrates. The ClpS adaptor has a binding pocket that selectively binds substrates with an exposed N-terminal Phenylalanine, Tyrosine, Tryptophan or Leucine. ClpS binds to the N-domain of ClpA, which allows for the transfer of the substrate to the ClpAP protease. Once the substrate is transferred to pore of ClpA, ClpA uses energy from ATP hydrolysis to unfold the substrate, pull it through the axial pore of the ClpA hexamer, and translocate it into the degradation chamber barrel of ClpP. Serine peptidase actives sites in ClpP hydrolyze the substrate polypeptide into short peptide products.

ClpA has a small N-terminal domain (N-domain) that can bind an adaptor called ClpS. ClpS recognizes protein substrates with specific N-terminal amino acids and delivers these substrates to ClpAP by binding to the ClpA N-domain. N-terminal amino acids that are targeted by the N-end rule pathway are called N-degrons.

The ClpS adaptor comprises three different parts (**Figure. 1.3.1**): a folded core domain that contains the N-end degron binding pocket, a ~20 amino acid unstructured N-terminal extension region, and a functionally important junction portion between these components (**Figure 1.3.2**). ClpS is able to bring N-end rule substrates to ClpAP because it has a conserved hydrophobic binding pocket that selectively recognizes N-terminal Phenylalanine, Tryptophan, Tyrosine, and Leucine amino acids, despite their different side chain shapes.[1,2,6,7] Importantly, the ClpS binding pocket does not bind to β-branched side chains (e.g., Ile) because of a conserved methionine embedded in the

hydrophobic pocket, which sterically clashes with β-branched N-terminal amino acids.[2] The junction region of ClpS consists of four to five residues that are between the core and the N-terminal extension that contributes to the docking of ClpS to ClpA. The junction is substantially conserved among other ClpS orthologs in Proteobacteria. The N-terminal extension of ClpS in comparison to the junction region consists of approximately fourteen random residues. The length of the N-terminal extension is conserved across species, but its amino acid composition is variable. The N-terminal extension and the junction region are required in order for ClpS to efficiently deliver N-end rule substrates to ClpAP, while resisting degradation itself. [5] Additionally, binding of ClpS inhibits degradation of non-N-end rule substrates (e.g., substrates tagged with a C-terminal ssrA tag), thus serving as a specificity switch for the protease.[3,6,8,9]

1.3 EVALUATING UNUSUAL CLPS ORTHOLOG

Despite the N-end rule pathway being conserved throughout the phylum Proteobacteria, in *E. coli* the ClpS adaptor is not an essential protein.[7] Moreover, it has been shown that ClpA is able to weakly recognize and degrade N-end rule proteins even without ClpS.[7] If the protease ClpAP is able to perform the task of degrading N-end rule substrates, why is ClpS highly conserved in other species if the adaptor is not vital for the cell to live? Does ClpS perform different functions in the cell than what was already described in the literature?

We wanted to examine the ClpS binding pocket composition in other species of the phylum Proteobacteria, to see if different amino acids at the sites that interact with the binding pocket would modulate the function and structure of the ClpS adaptor in comparison to *E. coli* ClpS. In order to compare the binding pocket region of ClpS,

we gathered and aligned sequences of ClpS adaptors from different proteobacterial species. We aligned the sequences and focus on a few amino acids that are known to be critical for recognition of N-end rule degrons. The *E. coli* ClpS amino acid residues that are critical for recognition consist of residue Leucine$_{32}$, Methionine$_{40}$, Valine$_{43}$, Leucine$_{47}$, Methionine$_{62}$, Valine$_{65}$, Histidine$_{66}$, Valine$_{88}$, and Leucine$_{99}$. These positions within the sequence impacts the binding pocket structure of the adaptor.

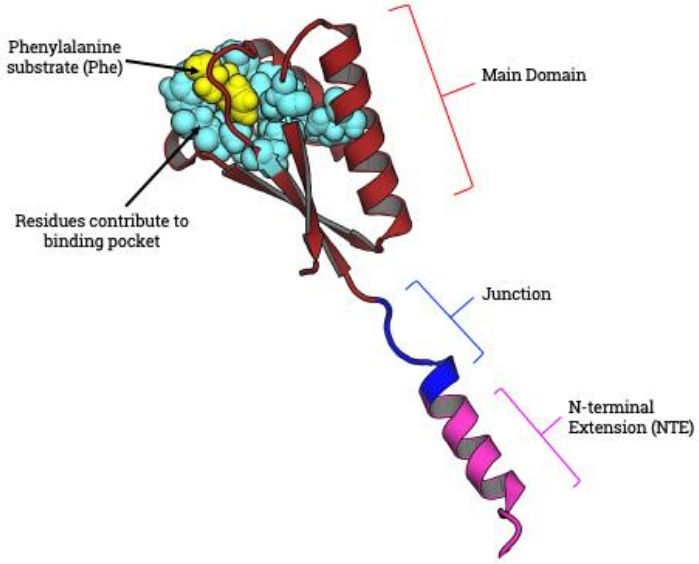


Figure 1.3.1 Important residues that impact the recognition of N-degrons. *E. coli* ClpS structure was obtain from the Protein Data Base (3OH2). Using the sequence of *E. coli* ClpS we are able to highlight the residues that contributes to the binding of N-degrons represented in cyan spheres within the structure. The red folded segment consisting of helices and sheets of the structure is called the main domain. The dark blue strand is called the junction. The magenta helix located at the end of the junction is the N-terminal extension (NTE). The yellow spheres is one of the N-degrons, Phenylalanine binding in the pocket of the main domain.

Within this alignment, we identified ClpS orthologs that possessed unusual amino acid substitutions at these positions. Evaluating the sequences, the ClpS ortholog from human pathogen *Helicobacter pylori* had the highest level of odd amino acid substitutions within the binding pocket (**Figure 1.3.2**).

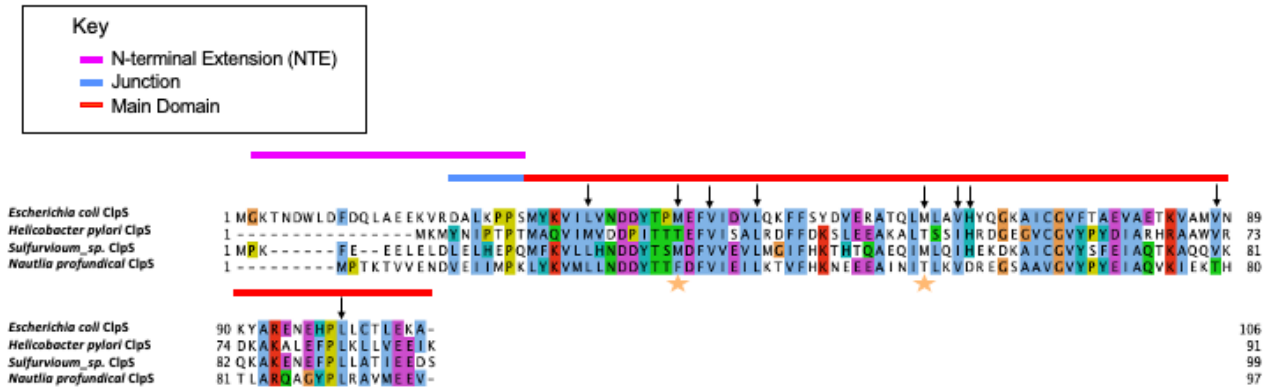


Figure 1.3.2 Conservation comparison among different unusual ClpS adaptors. The ClpS adaptors contains three different components; the N-terminal extension (NTE), junction, and the core domain. The binding pocket is located within the core domain. The black arrows indicated bases location within the sequences that contributes to the overall specificity of the binding pocket. The first sequence on the top is the ClpS adaptor for *Escherichia coli* followed by ClpS orthologs from *Helicobacter pylori*, *Sulfurvioum sp*., and *Nautlia profundical*. The star at the bottom of residue 40 and 62 is where the unusual substitution is prominent. The amino acids were originally a Methionine in *E. coli* to Threonine in *H. pylori*.

After reviewing the alignment sequence and looking at the residues that impacts the binding pocket, the most unusual substitution of amino acid residue in *H. pylori* ClpS is at Methionine$_{32}$, Threonine$_{40}$, and Threonine$_{62}$. Looking at amino acid based on polarity, the Threonine substitution in residue 40 and 62 in *H. pylori* is a drastic change of polarity from the nonpolar amino acid Methionine. Moving forward we want to observe how does these unusual substitution affects the function and structure of $_{Hp}$ClpS adaptor.

Chapter 2

AIMS AND HYPOTHESIS

2.1 *H. PYLORI* CLPS FUNCTION OUTCOMES OF AMINO ACID RESIDUE SUBSITITUIONS

Our first aim of this experiment is to see how the unusual amino acid substitution contributes to the overall function of *H. pylori* ClpS. When evaluating what aspects of function we wanted to observe for this adaptor; we chose to focus on experiments that correlates to binding of N-degrons and its interaction with other proteolytic complex systems.

2.1.1 *H. PYLORI* BINDING WITH N-END RULE DEGRONS

Can *H. pylori* ClpS bind with characterized N-degron substrates Phenylalanine (Phe), Tyrosine (Tyr), Leucine (Leu), and Tryptophan (Trp)? The hypothesis for this first sub aim is that *H. pylori* ClpS can not bind to the N-degrons because of the drastic switch of amino acid residues within the binding pocket of *H. pylori* changes the overall affinity and interaction of the pocket binding towards those N-end degrons.

2.1.2 *H. PYLORI* CLPS, ADAPTOR FOR OTHER PROTEOLYTIC SYSTEMS?

Can *H. pylori* ClpS function as an adaptor for other AAA+ proteolytic complexes like *E. coli* ClpAP? The hypothesis for this second sub aim is that *H. pylori* ClpS is able to function as an adaptor for *E. coli* ClpAP because ClpS is highly conserved within different species. The conservation of the main domain and length of NTE allows *H. pylori* interact with the unfoldase *E. coli* ClpA in a similar manner like *E. coli* ClpS.

2.2 *H. PYLORI* CLPS STRUCTURAL AND BINDING POCKET IMPACT FROM AMINO ACID SUBSITUTIONS

Our second aim of this experiment is to see how the changes within the amino acid sequence affects *H. pylori* secondary structure and its binding pocket shape. The hypothesis for second aim is that the overall secondary structure of *H. pylori* ClpS will be similar to the model ClpS from *E. coli* because the main domain of all ClpS are conserved. The binding pocket of *H. pylori* ClpS will differ in shape compared to *E. coli* because the unusual polar substitution within amino acid residue Theronine$_{40}$ and Theronine$_{62}$ in the *H. pylori* sequence.

Chapter 3

MATERIAL AND METHODS

In order to compare the function and structure of *H. pylori* ClpS and *E. coli* ClpS, I first needed to express and purify *H. pylori* ClpS. I constructed a plasmid that allowed me to recombinantly express *H. pylori* ClpS ($_{Hp}$ClpS) in *E. coli* expression strains. Overexpressed $_{Hp}$ClpS could then be purified from *E. coli* cell lysate. Purified $_{Hp}$ClpS was used to run assays to characterize its function, and for crystallization trials to help determine its three-dimensional structure. A pre-existing stock of purified *E. coli* ClpS ($_{Ec}$ClpS) in the Schmitz lab was used as a control in assays. X-ray crystal structures of $_{Ec}$ClpS has been reported[5] and are available online in the Protein Data Bank (3O2H & 3O2B).

3.1 PLASMID DESIGN FOR *H. PYLORI* CLPS

Plasmids are small circular DNA molecules that are separate from the chromosomal DNA of the cell. *E. coli* plasmids can serve as vectors that carry foreign genetic material. DNA encoding $_{Hp}$ClpS was introduced into an *E. coli* plasmid, which was used to overexpress $_{Hp}$ClpS protein in *E. coli*. The entire *clpS* coding region from *H. pylori*, including the main domain, junction and N-terminal extension, was ordered as a segment of synthetic DNA (Twist Biosciences). This synthetic DNA was cloned into expression vector MBP-SUMO-HTUA (a derivative of pET-21a; EMD Biosciences) using Gibson Assembly[10]. The sequence encoding $_{Hp}$ClpS was cloned in frame with an upstream sequence encoding an His$_6$-MBP-SUMO tag. (The functions of each component of the tag described below.) The resulting plasmid was used to produce full-length $_{Hp}$ClpS.

For crystallization experiments, a second HpClpS construct was cloned that lacked the presumably unstructured N-terminal extension. PCR-based site-directed mutagenesis was used to remove the first 14 codons, corresponding to the N-terminal extension. The resulting plasmid was used to produce N-terminally truncated protein, $_{Hp}ClpS_{\Delta N}$.

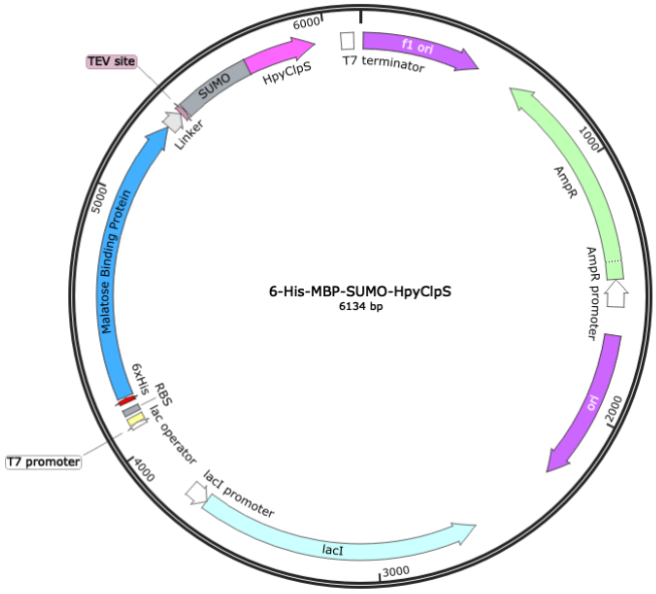


Figure 3.1 Plasmid design of *H. pylori* ClpS. The plasmid is derived from pET-21a and incorporates an *E. coli* p15A origin of replication (ori), an ampicillin resistance cassette (AmpR), a cassette encoding the Lac repressor (LacI). The open reading frame encoding HpClpS was cloned downstream of a fusion tag encoding a hexahistidine sequence (6xHis), maltose binding protein, and SUMO, controlled by a Lac/T7 promoter.

Plasmids encoding full-length and truncated proteins both incorporated an N-terminal His_6-MBP-SUMO tag. The His_6 allowed the construct to adhere to Ni-NTA resin for affinity chromatography[11]. The SUMO tag was added immediately upstream of ClpS both to improve solubility, and because it can be readily cleaved by the

SUMO-protease ULP1 without leaving any residual amino acids[12]. (ULP1 can cleave a SUMO tag upstream of any amino acid except proline[13], so the truncation of $_{Hp}ClpS_{\Delta N}$ was constructed to avoid placing a proline after SUMO.) A maltose binding domain (MBP) was included upstream of the SUMO tag because it helps create a large size difference between tag (0.84 kDa 6xHis + 43.82 kDa MBP + 10.9 kDa SUMO = ~56 kDa) and ClpS (~10 kDa) after cleavage with ULP1. This allowed us to more easily follow ClpS by SDS-PAGE during purification, and to separate cleaved tag from ClpS by size-exclusion chromatography

3.2 PROTEIN PURIFICATION

3.2.1 TRANSFORMATION OF PLASMID

To be able to recombinantly express and purify $_{Hp}ClpS/_{Hp}ClpS_{\Delta N}$, we first transformed the plasmid into an *E. coli* expression strain. Transformation is the process of introducing foreign DNA directly into a host cell. There are different ways to transform the *E. coli* cells; I transformed plasmid into cells by electroporation. 10 µL of the miniprepped plasmid DNA drop dialyzed against dH_2O for 3-4 hours on a 0.025 µm VSWP filter (Millipore). 1µL of dialyzed plasmid was mixed with 50 µL dH_2O and 10 µL electrocompetent *E. coli* strain ER2566 (NEB) in a 1 cm gap-width cuvette, and electroporated in a MicroPulser (Bio-Rad) on setting EC1. 1 mL of 1.5XYT medium was added and the transformation solution was transferred to a test tube and incubated shaking for 45 minutes at 37°C. The transformation solution was plated on LB+ampicillin (100 mg/L) plates and incubated at 37°C overnight. Since the vector has a gene conferring ampicillin resistant, this ensures that the colonies that

grow on the plate carry *the plasmid*. Single colonies from these plates were used to inoculate overnight cultures for larger scale overexpressing. (A detailed transformation protocol is included in Appendix A.)

3.2.2 OVERNIGHT CULTURE

To produce large amounts of purified protein, overexpression was carried out with eight liters of culture. Overexpression cultures were inoculated with 10 mL of overnight culture for each liter of culture. Two 50 mL overnight culture were prepared in autoclaved 250 mL flasks, in a sterile environment. Each flask contained 50 mL of 1.5 XYT media + 50 mg/L ampicillin. A single colony from the ampicillin plate was used to inoculate each 250 mL flask. The flasks were covered with aluminum foil and were incubated, shaking, at 37°C overnight.

3.2.3 LARGE SCALE OVEREXPRESSION

Eight cultures were prepared, each with 1 L of 1.5XYT media containing 50 mg/mL ampicillin in a baffled Ultra Yield shaking flask (Thompson). 10 - 12.5 mL of overnight culture was added to each flask. Flasks were incubated, with shaking, at 37°C for two to three hours until the optical density (OD_{600}) was within the range of 0.6-1.0. Once the desired OD was reached, the shaker temperature was set to 30°C and the flasks were allowed to cool down for a few minutes. Overexpression of protein was induced by the addition of 0.5 mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) in each of the eight flasks. IPTG relieves repression of the Lac operon and permits strong transcription of *H. pylori clpS* gene from the T7 RNA polymerase promoter. Overexpression continued for 4 hours at 37°C.

3.2.4 HARVESTING CULTURE

Each 1 L culture was transferred to a centrifuge bottle and spun at 4000 rpm 30 minutes at 4°C. The supernatant was poured off into a collection container and sterilized with 10% bleach. (The collection container is referred to in our lab as the "evil box" because of its pungent odor.) Pelleted bacteria from 1 L of culture were suspended in 25 mL of lysis buffer (25 mM HEPES pH 7.5, 10 mM imidazole pH 7.5, 300 mM NaCl, 10% glycerol) by vortexing. Resuspended pellets were stored at -80°C.

3.2.5 LYSIS AND CLARIFICATION OF PROTEIN

Resuspended cell pellets were lysed by sonication. Samples were initially combined, divided into three 250 mL beakers, and placed on ice. Samples were sonicated with a 4 mm probe tip at 75% amplitude in ten cycles of 30 seconds each. The temperature was monitored to ensure it stayed below 10°C. When necessary, the samples were allowed to cool on ice before starting a new cycle of sonication. Surpassing the temperature of 10°C could denature the protein. Sonication alternated among the three different beakers to give each time to cool before the start of the cycle.

After sonication, samples were divided into four 50 mL conical tubes. The tubes were spun at 15,000 rpm at 4°C for 1 hour. Clarified supernatant was collected and combined in a clear plastic bottle. Supernatant was stored at -80°C.

3.2.6 NI-NTA COLUMN

As noted above, protein constructs incorporate an N-terminal His_6-tag that allows the protein to selectively interact with immobilized nickel ions (e.g., on Ni-NTA resin). High concentrations of imidazole, the same chemical group found on

histidine side chains, competes with histidine for binding and selectively elutes His_6-tagged protein from the column.

(A detailed protocol for Ni-NTA chromatography over a gravity column is included in Appendix C.)

The clarified *E. coli* lysate, containing a His_6-tagged $_{Hp}$ClpS construct, was passed over a column of Ni-NTA agarose pre-equilibrated in lysis buffer by gravity flow. Twice as much of Ni-NTA slurry was used due to the inefficiency of the bead binding to the protein of interest [24 mL of Ni-NTA slurry]. Fractions were collected throughout the Ni-NTA chromatography procedure for later analysis by SDS-PAGE [**Figure 3.2.6**]. The flow-through was collected and passed through the column a second time, to ensure the protein had bind to the nickel beads. The column was washed with ~15 mL lysis buffer to remove nonspecifically bound proteins. Protein was eluted from the column in three fractions of elution buffer (25 mM HEPES pH 7.5, 300 mM NaCl, 300 mM imidazole pH 7.5, 10% glycerol). After the Ni-NTA column, an SDS-PAGE gel was run to identify which fractions contained the $_{Hp}$ClpS construct [**Figure 3.2.6**].

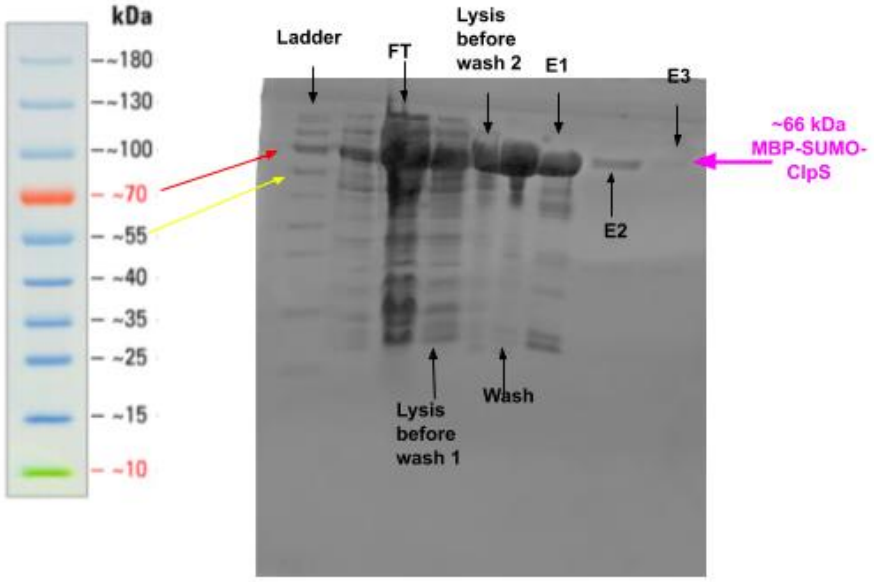


Figure 3.2.6 SDS-PAGE gel after Ni-NTA column. Fractions from Ni-NTA chromatography were run on an SDS-PAGE gel. 10 μL of each sample was mixed with 5 μL of SDS-PAGE loading buffer (450 μL of 3x Sample Buffer, 50 μL of 2-Mercaptoethanol) and added to the designated well. (Samples were not boiled, as boiling was found to cause irreversible precipitation.) 3μL of protein standards ladder was loaded. From left to right, the first well is the ladder followed by FT (flow through), lysis before wash 1, lysis before wash 2, wash, E1 (elution 1), E2 (elution 2), and E3 (elution 3). *H. pylori* ClpS constructs incorporated a His_6-MBP-SUMO and had a mass of ~66 kDa.

3.2.7 SPIN CONCENTRATION

Elution fractions containing $_{Hp}$ClpS constructs were pooled and concentrated in a 10 kDa cutoff Amicon spin concentration device (Millipore) at 4,000 rpm, 4°C, for 30 minutes intervals. Flow-through was discarded and replaced with more of the elution fractions until the total final volume of the concentrate was at or below 1.5 mL. In between spins, a 100 μL pipette was used mix the concentrate, to prevent precipitation at the bottom of the spin concentrator. Concentrated protein was transferred to a 50 mL conical tube and stored at -80°C. Spin concentrators were reused for later preps of the same construct.

3.2.8 BUFFER EXCHANGE

Prior to removal of the His_6-MBP-SUMO tag with ULP1 protease, $_{Hp}$ClpS constructs were buffer exchanged from Ni-NTA elution buffer into a buffer better optimized for ULP1 activity. For full-length $_{Hp}$ClpS, buffer was change to APD buffer (50 mM HEPES pH 7.5, 300 mM Sodium Chloride, 20 mM Magnesium Chloride, 10% glycerol, 0.1 mM EDTA) which is optimized for subsequent assays with *E. coli* ClpAP. For $_{Hp}ClpS_{\Delta N}$, the buffer was changed to a different ratio component of $_{Hp}$APD buffer (10 mM HEPES, 150 mM Sodium Chloride, 10% glycerol) in order to prevent precipitation of the protein.

Buffer exchange was achieved by successive rounds of spin concentration, followed by dilution in the destination buffer in a 10 kDa cutoff spin concentrator. ~1.5 mL of concentrated protein was diluted to ~15 mL with destination buffer, and spin concentrated as described above. The flow-through was discarded. This process was repeated 3 or 4 times, and the protein was again concentrated to 1.5 mL or less. Once the exchange of buffer was completed, concentrated protein was transferred to a 1.7 mL Eppendorf tube.

3.2.9 ULP1 CLEAVAGE

ULP1 (ubiquitin-like protease 1) specifically recognizes SUMO and cleaves the peptide bond immediately downstream of its terminal glycine amino acid[12]. ULP1 was used to remove His_6-MBP-SUMO tag from $_{Hp}$ClpS constructs. The ratio of protein to ULP1 varied empirically depending on the activity of the particular ULP1 prep. A ratio of 1:10 to 1:50 of concentrated ULP1 to concentrated protein was used for cleavage. ULP1 was added to the concentrated, buffer exchanged protein and incubated overnight at 30°C in order to optimize the work efficiency of the enzyme.

Samples before and after incubation with ULP1 were compared by SDS-PAGE to determine if they had been efficiently cleaved off. There was a clear size differentiation associated with cleavage: the ~66 kDa construct was cleaved into ~56 kDa and ~10 kDa fragments.

3.2.10 SIZE EXCLUSION CHROMATOGRAPHY

Protein was further purified by size exclusion chromatography, which effectively separated the small $_{Hp}$ClpS constructs from the larger His$_6$-MBP-SUMO tag. (I found that ClpS does not tolerate low salt concentrations, which is why ion exchange chromatography was not used as a purification step.)

Cleaved protein was filtered through a 0.2 µm syringe filter and loaded into the sample loop of an FPLC instrument (Fast Protein Liquid Chromatography). ($_{Hp}$ClpS was viscous and frequently clogged the syringe filter. Multiple filters were used when necessary.) Sample was injected over a HiLoad 16/600 Superdex® 75 prep-grade column (GE Healthcare), and elution fractions were collected during isocratic elution. During size exclusion chromatography, larger molecules elute first, followed by medium and smaller molecules. $_{Hp}$ClpS constructs are small proteins, and eluted towards the end of the fractionation range, in fractions 24-31. The chromatograph trace was used as a guide for fractions containing protein, and fractions were analyzed by SDS-PAGE. This final purified protein was concentrated to 1.5 mL or less.

3.3 ASSAYS TO FIND FUNCTION OF *H. PYLORI* CLPS

In order to analyze *H. pylori* ClpS function in comparison to *E. coli* ClpS, two assays will be conducted; Anisotropy and Protease Assay. Anisotropy assays analyzes the binding ability of *H. pylori* ClpS binding pocket with the N-end rule peptides (Phe,

Tyr, Trp, and Leu) and serine peptide. Anisotropy is based on the ratio of emission that shows up in the two emission channels. The total light emission is the same in both channels over all samples. In the assay the changes in the polarization of emission is observed, but not the total amount. The changes in polarization of emission of the sample is measured by two channels that are parallel and perpendicular.

The ratio that these two channels is converted to an “anisotropy”. An anisotropy value can be large or small. A large anisotropic signal means that there is a large difference in the ratio. When the fluorescence labeled peptide with N-terminal amino acid binds with the ClpS protein, this protein-peptide complex with create a slow tumble within the sample. The polarized light diffracts off the complex and has a longer time to pass through the grating filter, depicting the large different ratio in anisotropic signal. A low anisotropy value means that there is the same ratio amount in both polarization channels. The low number in dictates the peptide and protein have not formed together in a complex. The alone peptide will tumble faster in the sample, cutting the signal frequently not allowing the diffracted light to spend time in the channels, making both channels have the same ratio.

In-vitro proteolytic assay allows us to see if *H. pylori* ClpS can act as an adaptor for the protease complex *E. coli* ClpAP by interacting with the E. coli ClpA unfoldase. If an adaptor is bound to the N-domain of ClpA, the interaction inhibits -ssrA substrates degradation. The different components that makes up this assay is 0.3μM E. coli ClpA, 0.5 μM E. coli ClpP, 1 μM E. coli ClpS, 1μM H. pylori ClpS, 10 μM GFP-ssrA substrate, and 1x PK regeneration. The substrate has GFP attached, which helps detects the level of substrate present within the wells they are in. *E. coli*

ClpA is the unfoldase that requires ATP hydrolysis in order to unfold the substrates. PK regeneration is able to utilize low ADP concentrations and convert it back to ATP[14], supplying the unfoldase the energy to unfold the substrates. [**Figure 4.3**] (Both of these protocols can be found in the appendix D and E)

3.4 CRYSTALLOGRAPHY TO FIND STRUCTURE OF *H. PYLORI* CLPS

X-ray crystallography is the most commonly used method for determining the three-dimensional structure of folded proteins[15]. A beam of X-rays is directed at a protein crystal, and the resulting diffraction pattern is measured at different angles. X-ray diffraction data are ultimately used to calculate the electron density of the crystallized protein and build a model of the protein structure. This method necessarily requires producing high quality crystals of the protein of interest. Conditions that support crystal growth are generally discovered by trial and error, by mixing concentrated protein with screens of potential crystallization conditions containing different buffers, salts, and organic precipitants.

To find conditions that support crystal growth, crystallization trails of purified $_{Hp}ClpS_{\Delta N}$ were set up using two 96-condition crystallization screens: JCSG-PLUS and Pact Premier (Molecular Dimensions). $_{Hp}ClpS_{\Delta N}$ that was used in these crystallization screenings had a concentration of 5 mg/mL. The volume protein drops, and reservoir drops were 0.5 μL. The total reservoir volume in the crystallization screening was 75 μL. Crystallization screens were incubated at 20°C and checked periodically under a microscope for crystal formation.

Chapter 4

RESULTS

4.1 OPTIMIZATION OF $_{Hp}$CLPS OVEREXPRESSION AND PURIFICATION

To begin to understand the properties of $_{Hp}$ClpS, we first set out to recombinantly overexpress this protein in *E. coli*. I transformed plasmid encoding $_{Hp}$ClpS into the *E. coli* expression strain ER2566. Unexpectedly, cells carrying the expression plasmid grew slower than typical, even before IPTG induction. Inoculated 1 L cultures of $_{Hp}$ClpS took longer than normal to reach the target OD_{600}: 5-6 hours, instead of 2-3 hours at the temperature of 37°C. When prepping for the overnight cultures, I also noticed that the cells grew more slowly. Majority of the times I would extend the amount of time per step to ensure at least enough growth of cell were present before proceeding to the next step. Moreover, the overexpress yield from this strain was much lower than observed for other ClpS orthologs, notably *E. coli* and *M. smegmatis*.

We suspected that $_{Hp}$ClpS was toxic to the ER2566 host cells, and that leaky expression was selecting for cells with low expression of the protein of interest. To circumvent this, we transformed the expression plasmid into a different *E. coli* expression strain, C43(DE3), which is documented as being particularly successful at expressing toxic protein targets[16]. We observed that inoculated cultures of C43(DE3) cells reach the target OD faster, and gave greater overexpression yield. These observations suggest that overexpression of $_{Hp}$ClpS is toxic to *E. coli* under some circumstances, which we have not observed for other orthologs.

Although $_{Hp}$ClpS overexpressed well, its poor behavior during purification resulted in some loss across purification steps. Spin-concentrated protein at high

concentration was viscous and frequently clogged syringe filters while loading the sample on the FPLC. I suspect that a large portion of each prep was lost due to precipitation, interactions with filtration membranes, and adherence to size-exclusion column media. However, there was still sufficient purified HpClpS for downstream assays and crystallography trials.

4.2 EVALUATION OF N-END RULE PEPTIDE BINDING BY FLUORESCENCE ANISTROPY

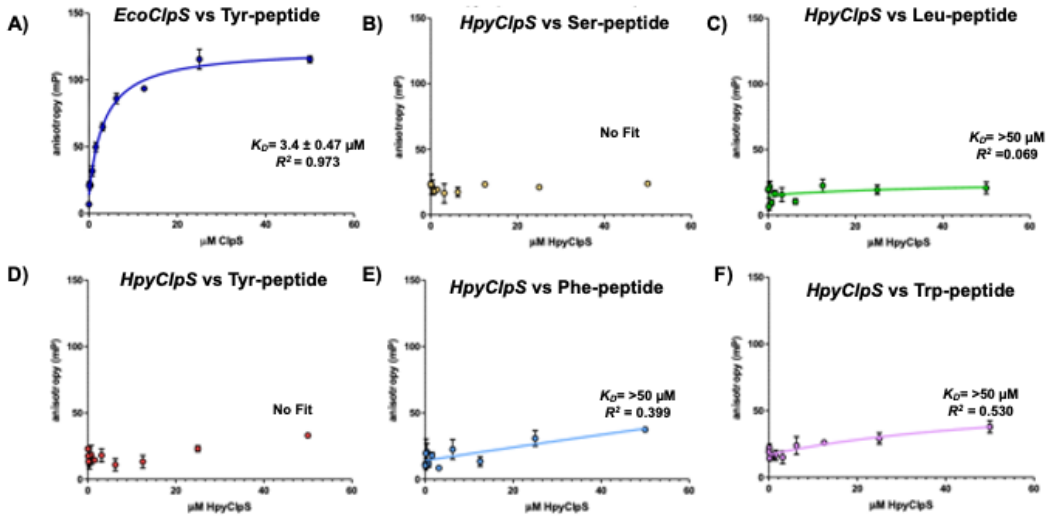


Figure 4.2 *H. pylori* ClpS does not interact with N-end rule peptides. Changes in fluorescence anisotropy were used to monitor the ability of ClpS (0 – 50 µM) to bind fluorescently labeled peptides (0.1 µM) with different N-terminal amino acids. Data are fit to a single site binding equation. Error is reported as K_D ± SD. (A) *E. coli* ClpS bound a Tyr-peptide with a K_D of 3.4 µM, as expected[17].For HpClpS , weak binding >50 µM or no significant binding was observed to peptides bearing N-terminal (C) Leu, (D) Tyr, (E) Phe, or F) Trp amino acids, all of which are canonical N-end degrons. (B) No binding was observed between HpClpS and a peptide bearing an N-terminal Ser, which is not an N-end rule degron.

To test if purified HpClpS can bind canonical N-end rule degrons, we monitored the fluorescence anisotropy of fluorescently labeled peptides with various N-terminal amino acids over a range of ClpS concentrations (**Fig. 4.2**). In these

experiments, the sample was excited with vertically polarized light and the emission fluorescence was measured in vertical and horizontal polarization planes. Peptide alone tumbles rapidly in solution, resulting in similar emission in both polarization planes. The emission anisotropy increases if ClpS binds the peptide, because the ClpS•peptide complex tumbles much more slowly than the peptide alone.

As a positive control, we tested binding of *E. coli* ClpS ($_{Ec}$ClpS) to a known N-end rule degron. $_{Ec}$ClpS binds a Tyr-peptide with K_D of 3.4 μM, similar to reported values which range from 3-9 μM for similar peptides[17] (**Fig. 4.2**). By contrast, we observed weak or no binding of $_{Hp}$ClpS to peptides with N-end rule amino acids (**Fig. 4.2 C-F**). $_{Hp}$ClpS appeared to bind Leu, Phe, and Trp peptides weakly, with poorly determined K_D values >50 μM and large fitting error. No fit was determined for $_{Hp}$ClpS binding to a Tyr peptide. Similarly, saw no binding between $_{Hp}$ClpS and a non-N-end rule control peptide with an N-terminal Ser (**Fig. 4.2 B**). These data indicate that $_{Hp}$ClpS is unable to bind canonical N-end rule degrons.

4.3 INFLUENCE OF HPCLPS ON PROTEASE ACTIVITY OF ECCLPAP IN VIRTO

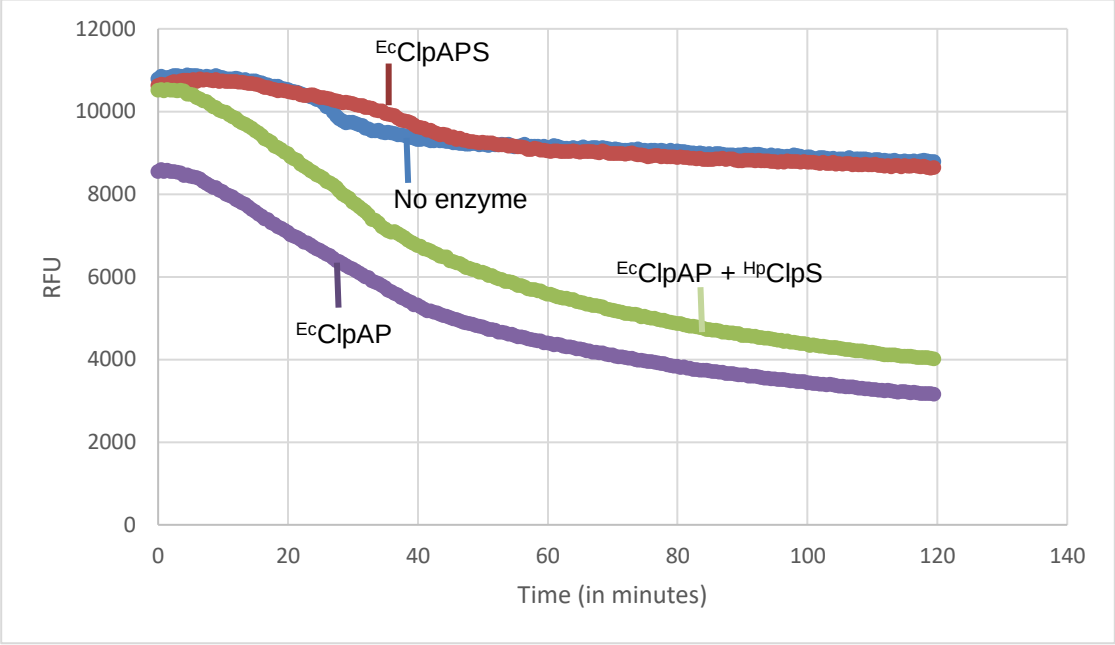


Figure 4.3 Effect of HpClpS on proteolysis by *E. coli* ClpAP. Proteolysis of a CP7GFP-ssrA model substrate (10 μM) by *E. coli* ClpAP (0.3 μM + 0.5 μM) was monitored in the absence of ClpS (purple), the presence of EcClpS (1 μM; ruby), or the presence of HpClpS (1 μM; green). A control reaction containing substrate, but no enzyme is also shown (blue). This in-vitro proteolytic assay was done in triplicates. This graph displays one of the triplicates.

The anisotropy data suggest that suggest that HpClpS cannot interact with typical N-end rule substrates and deliver them to ClpA. However, it is still possible that HpClpS interacts with ClpA and influences its activity. EcClpAP within the cell can degrade -ssrA tagged substrates without the aid of an adaptor. When EcClpS interacts with EcClpA a switch in degradation of substrate occurs. EcClpS binds to the N-domain of EcClpA (**Figure 1.2**), suppresses proteolysis of non-N-end rule substrates including -ssrA tagged substrates and forces the EcClpAP complex to now degrade substrates that are characterized as N-degrons.

In order to see if HpClpS can alter the substrate specificity of *E.coli* ClpAP, we tested proteolysis of a simple model substrate, circularly-permuted GFP-ssrA (CP7-GFP-ssrA[5], by *E. coli* ClpAP in the presence and absence of a molar excess of either *E. coli* or *H. pylori* ClpS . **Fig. 4.3** shows the kinetic trajectories of GFP fluorescence over the course of the assay. A negative control reaction with substrate but lacking enzyme is show in blue and showed little change in fluorescence signal over time. When ClpAP was also present (**Fig. 4.3,** purple curve), proteolysis of substrate resulted in a loss of fluorescence over the time. As expected, the presence of EcClpS inhibited substrate degradation (**Fig. 4.3**, ruby curve), resulting in a trajectory similar to the GFP-only control. This is consistent with the reported function of EcClpS as a specificity switch for *E. coli* ClpAP[3,18].However, the presence of HpClpS did not suppress CP7GFP-ssrA degradation by *E. coli* ClpAP (**Fig. 4.3**, green curve), and resulted in a proteolysis trajectory similar to *E. coli* ClpAP in the absence of ClpS. There is no evidence that HpClpS alters the ability of *E. coli* ClpAP to recognize and degrade a ssrA-tagged model substrate. This may reflect its behavior in *H. pylori*. Alternatively, it is possible that HpClpS is simply unable to interact with *E. coli* ClpA and interacts differently with its cognate ClpA ortholog in *H. pylori*.

4.4 CRYSTALIZATION TRIALS

The anisotropy data indicated that HpClpS is unable to bind peptides with canonical N-end rule amino acids. Our hypothesis predicts that this is due to the noted substitutions in key positions that form the ClpS substrate pocket. To test this directly, we sought to determine the three-dimensional structure of HpClpS by X-ray crystallography.

To generate crystals for X-ray crystallography, I expressed and purified an N-terminally truncated $_{Hp}$ClpS construct ($_{Hp}$ClpS$_{\Delta N}$), lacking the presumably unstructured N-terminal extension. As stated in the previous paragraph (pg. 20), the protein yield of $_{Hp}$ClpS was low and behaved poorly during overexpressing when the protein was transformed in ER2566 cells. This led me to change the cell line C43 to ultimately deal with the protein of interest and obtain a higher yield. Throughout the protein purification process, C43 cells was less difficult to overexpress, and concentrate compared to the ER2566 cells. The main issue that was observed when using C43 cell line was when ULP1 cleaved the 6His-MBP-SUMO tag off ClpS. ULP1 was able to cleave the SUMO tag efficiently but now that the MBP tag was not attached to the protein, the protein precipitated. In order to bring the protein back into solution, I had to unfold the protein by placing the precipitated protein in 3M Guanidine Hydrochloride. After unfolding the protein, I had to find a buffer that the protein can refold and remain soluble in solution.

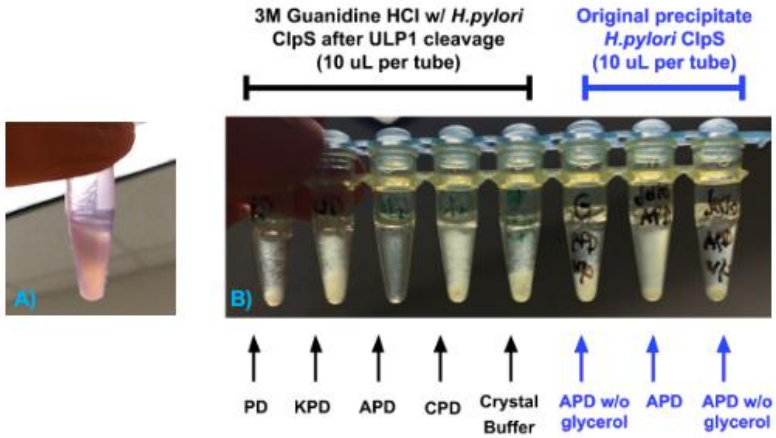
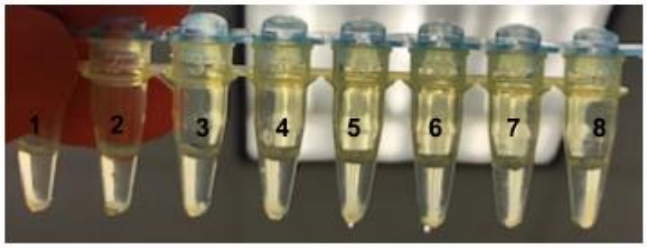


Figure 4.4.1 Attempting to bring $_{Hp}$ClpS back into solution. This figure displays the aftermath of ULP1 cleavage of HpyClpS, precipitated by MBP tag (**Fig. 4.4.1 A**). **Fig. 4.4.1 B** displays the 3M Guanidine Hydrochloride with $_{Hp}$ClpS (10 μL) in different buffers (PD, KPD, APD, CPD, crystal buffer: 100 μL) that contain different

ratios of salt and glycerol percentage. The last three tubes have the precipitated ClpS (10 μL) that was test against APD buffer (100 μL) in the present and absence of glycerol.

With the testing of different buffers to see which will aid unfolded HpClpS back into solution, APD buffer with glycerol seems to be the best candidate out of all the buffers. In **figure 4.4.1 B**, APD buffer is transparent in color; a sign that precipitation is not present in the solution. With this initial APD buffer condition, I wanted to optimize the buffer to not only bring ClpS back into solution but to maintain suitable conditions for crystallography in order for the protein to crystalize properly.



Tube Number	HEPES Concentration pH 7.5	Sodium Chloride (NaCl) Concentration	Glycerol/Sucrose Percentages
1	10 mM	100 mM	10% glycerol
2	10 mM	150 mM	10% glycerol
3	10 mM	200 mM	10% glycerol
4	10 mM	100 mM	10% sucrose
5	10 mM	150 mM	10% sucrose
6	10 mM	200 mM	10% sucrose
7	10 mM	250 mM	10% sucrose
8	10 mM	300 mM	10% sucrose

Figure 4.4.2 Optimizing buffer for HpClpS for crystallography. This figure displays the different ratios of HEPES, sodium chloride (NaCl), and glycerol/sucrose amount with 5 μL of 3M Guanidine Hydrochloride with HpClpS. Tube 2 ratio of components was able to keep the protein in solution the best out of all the ratios. The total volume of the different buffers and unfolded protein is 50 μL.

I proceeded to exchange the 3M Guanidine Hydrochloride that HpClpS was unfolded in and replace that solution with the HpAPD buffer (10 mM HEPES pH 7.5,

150 mM NaCl, and 10% glycerol). Buffer exchange was achieved by successive rounds of spin concentration, followed by dilution in the HpAPD buffer in a 3 kDa cutoff spin concentrator. ~1.5 mL of concentrated protein was diluted to ~15 mL with HpAPD buffer, and spin concentrated as described above (pg. 15). The flow-through was discarded. This process was repeated 3 or 4 times, and the protein was again concentrated to 1.5 mL or less. Once the exchange of buffer was completed, concentrated protein was transferred to a 1.7 mL Eppendorf tube. Despite successfully refolding and bring HpyClpS back into solution, the project could not be continued. Although I was not able to do crystallization screening with the prep of HpyClpS, I did do initial screenings with $_{Hp}ClpS_{\Delta N}$ that I was able to obtain from the previous rounds of overexpression.

Concentrated $_{Hp}ClpS_{\Delta N}$ (5 mg/mL) in our optimized buffer was used to set up two 96-well crystallization screens – JSCG-PLUS and PACT Premier – in sitting drop format. Several conditions produced small crystals (**Table 4.4, Figure 4.4.3**). In the JCSG-PLUS (MDI-37) well, A1 and I10 had a sign of small crystals growth. In Pact Premier (MDI-29) well, C3 had present of small crystal.

Table 4.4 Conditions with Crystal Growth of $_{Hp}ClpS_{\Delta N}$

Screening Tray Name	Well Number in Tray	Component in Well
JCSG-PLUS (MDI-37)	A1 (1-1)	0.2 M Lithium Sulfate 0.1 M Sodium acetate pH 4.5 50% v/v PEG 400
JCSG-PLUS (MDI-37)	I10 (2-46)	0.2 M Ammonium Sulfate 0.1 M Bis Tris pH 5.5 25% w/v PEG 3350
Pact Premier (MDI-29)	C3 (1-27)	1.1 M PCTP pH 6.0 25 % w/v PEG 1500

Seeing the presence of crystal growth allowed us to move forward to optimize these conditions and in 24-well hanging drop trays. However, due to the current climate of social distancing and pause on non-essential work, we were not able to progress with the optimization of $_{Hp}$ClpS$_{\Delta N}$ crystals. (As of writing, the 24 well trays are dried up, so crystals are no longer able to form.)

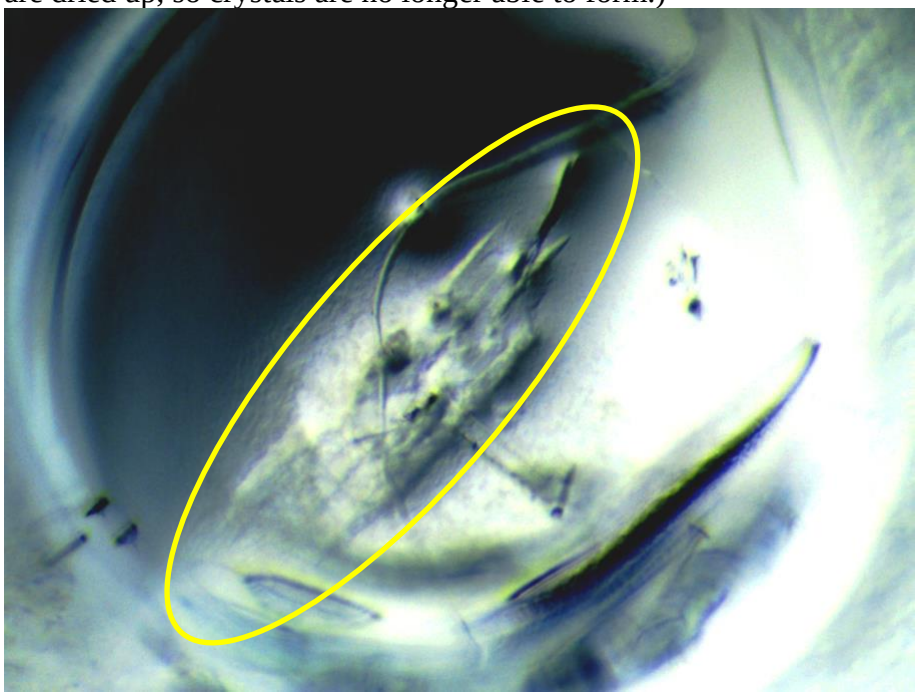


Figure 4.4.3 Remain of crystal outline from screening trays. *H. pylori* ClpS was able to form crystal in JCSG-PLUS (MDI-37), A1 although this well is dried.

4.5 COMPUTER MODELING OF *H. PYLORI* CLPS STRUCTURE

Unfortunately, I was not able to successfully crystalize $_{Hp}$ClpS. To be able to view a prediction structure of *H. pylori*, I utilized the primary sequence of *H. pylori* ClpS and put it through a threading server called Phyre2. Threading allows you to use the primary sequence of the protein and put it into this computer server. The program will then make secondary structure predictions using homology data base it has access

to. Once the server generates the structure, I was able to do comparison of this predicted HpClpS in comparison to the published *E. coli* ClpS5 in a molecular structural program called Pymol.

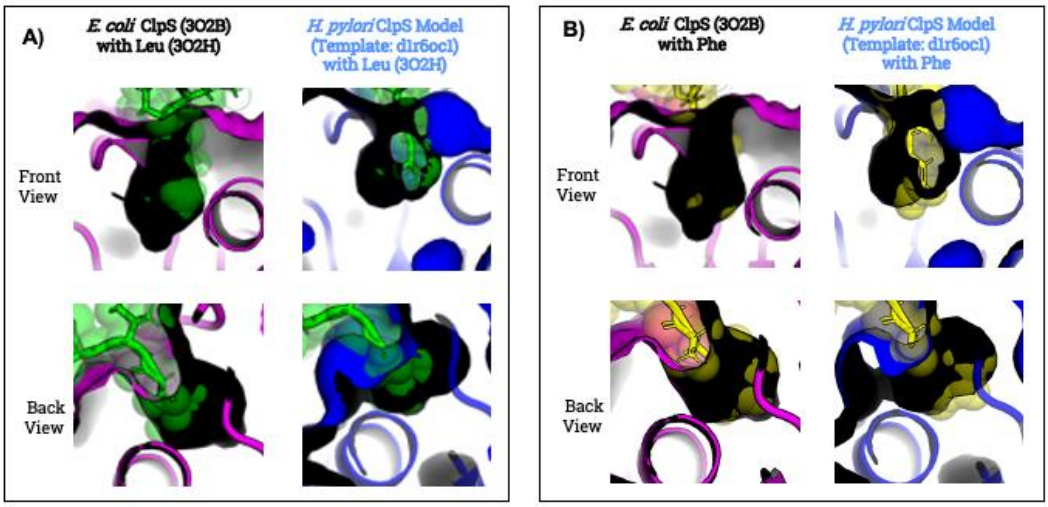


Figure 4.5 HpClpS binding pocket lacks depth and width to bind with N-degrons. Using the threading server Phyre2, it generated the prediction secondary structure of HpClpS (template: d1r6oc1). **Fig. 4.5 A)** compares *E. coli* ClpS and HpClpS binding against one of the N-degrons, Leucine (Leu). The first row of photo provides a different front view perspective of the interaction of the binding pocket and Leu substrate. The second row of photos shows a back view of the binding pocket with the Leu substrate. **Fig. 4.5 B)** compares both ClpS with different views, but shows binding with a different N-degron, Phenylalanine (Phe).

Chapter 5

DISSCUSION AND FUTURE DIRECTIONS

5.1 DISSCUSION

We hypothesized that unusual substitutions in amino acids that form the binding pocket in $_{Hp}$ClpS would alter its ability to bind N-end rule degrons. Indeed, we observed that $_{Hp}$ClpS at best weakly interacts with canonical N-end rule degrons. The unusual substitution may have contributed to the change in the binding pocket of $_{Hp}$ClpS. By observing the threading model of $_{Hp}$ClpS in comparison to *E. coli* ClpS (**Figure 4.5**), the binding pocket is shallower and wider. The change of the binding pocket can contribute to why *H. pylori* does not strongly bind with majority of the N-terminal substrates that is known to bind with *E. coli* ClpS. The change does not help $_{Hp}$ClpS accommodate the N-degrons.

In **figure 1.3.2.**, residue 40 and 62 that contributed to binding was completely changed a Methionine to a Threonine residue from *E. coli* ClpS to *H. pylori* ClpS. In *E. coli* ClpS, this methionine residue is responsible for the specificity ClpS has for particular N-degrons. This methionine residue blocks clashes with β-branched N-terminal amino acids.[2] The substitution of threonine in *H. pylori* ClpS adaptor potentially contributed to why this adaptor did not have the same binding affinity with the N-degrons.

We hypothesized that *H. pylori* ClpS could behave as an adaptor for *E. coli* ClpAP because of the highly conserved main domain similar to *E. coli* ClpS. Unfortunately, we did not observe $_{Hp}$ClpS inhibiting *E. coli* ClpAP from degrading the substrate CP7- GFP-ssrA, suggesting that *H. pylori* ClpS is not a compatible adaptor for *E. coli* ClpAP. The protease assay we conducted (**Figure 4.3**) only displays the

amount of substrate present by measuring the GFP level that are attached to the substrate. We are not able to visually see the interaction HpClpS with *E. coli* ClpA.

From the literature we are able to come up with explanations why this phenomenon occurred within the protease assay. Within the *E. coli* ClpAP proteolytic complex, non-N-degrons substrates like -ssrA tagged substrates can be degraded automatically without the aid of an adaptor. The interaction of *E. coli* ClpS adaptor and *E. coli* ClpA unfoldase causes a "specificity switch", only allowing N-degrons substrates to be degraded. Since HpClpS was unable to inhibit ssrA substrate degradation, this adaptor may not perform the specificity switch that *E. coli* ClpS can do. Another explanation is that HpClpS is unable to interact with *E. coli* ClpA which can contribute to the lack of inhibition of ssrA substrates in the assay.

Regarding *H. pylori* structure, we were unable to obtain a successful protein structure although *H. pylori* grew in conditions within the screening trays. Using the threading model as seen in **figure 4.5,** we are able to see the noticeable changes in depth and width in HpClpS bind pocket in comparison to the model *E. coli* ClpS.

5.2 FUTURE DIRECTIONS

For future directions, I plan to successfully overexpress *H. pylori* ClpS using C43 competent cells since this preparation of protein had the highest yield. After purification of *H. pylori* ClpS, we plan to redo the two 96-screening trays and optimize the growth conditions of *H. pylori* ClpS that we observe in the wells. With the protein crystal, the structure can be determined by x-ray crystallography. Figuring out the structure will also help us decide which N-terminal amino acids would likely bind to *H. pylori* ClpS binding pocket oppose to testing all of the remaining 15 amino acids. I would like to redo the protease assay to obtain better positive and negative

controls. I would like to purify *H. pylori* ClpA to see if *H. pylori* ClpS interacts with its own unfoldase.

REFERENCES

(1) Sauer, R. T.; Baker, T. A. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. *Annu. Rev. Biochem.* **2011**, *80* (1), 587–612. https://doi.org/10.1146/annurev-biochem-060408-172623.

(2) Román-Hernández, G.; Grant, R. A.; Sauer, R. T.; Baker, T. A. Molecular Basis of Substrate Selection by the N-End Rule Adaptor Protein ClpS. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (22), 8888–8893. https://doi.org/10.1073/pnas.0903614106.

(3) Dougan, D. A.; Truscott, K. N.; Zeth, K. The Bacterial N-End Rule Pathway: Expect the Unexpected. *Mol. Microbiol.* **2010**, *76* (3), 545–558. https://doi.org/10.1111/j.1365-2958.2010.07120.x.

(4) Bouchnak, I.; van Wijk, K. J. N-Degron Pathways in Plastids. *Trends Plant Sci.* **2019**, *24* (10), 917–926. https://doi.org/10.1016/j.tplants.2019.06.013.

(5) Román-Hernández, G.; Hou, J. Y.; Grant, R. A.; Sauer, R. T.; Baker, T. A. The ClpS Adaptor Mediates Staged Delivery of N-End Rule Substrates to the AAA+ ClpAP Protease. *Mol. Cell* **2011**, *43* (2), 217–228. https://doi.org/10.1016/j.molcel.2011.06.009.

(6) A. Erbse, R. Schmidt, T. Bornemann, J. Schneider-Mergener, A. Mogk, R. Zahn, D. A. Dougan, & B. B. ClpS Is an Essential Component of the N-End Rule Pathway in Escherichia Coli. *Nature*. 2006, pp 753–756. https://doi.org/10.1038/nature04412.

(7) Wang, K. H.; Sauer, R. T.; Baker, T. A. ClpS Modulates but Is Not Essential

for Bacterial N-End Rule Degradation. *Genes Dev.* **2007**, *21* (4), 403–408. https://doi.org/10.1101/gad.1511907.

(8) Dougan, D. A.; Reid, B. G.; Horwich, A. L.; Bukau, B. ClpS, a Substrate Modulator of the ClpAP Machine. *Mol. Cell* **2002**, *9* (3), 673–683. https://doi.org/10.1016/S1097-2765(02)00485-9.

(9) Hou, J. Y.; Sauer, R. T.; Baker, T. A. Distinct Structural Elements of the Adaptor ClpS Are Required for Regulating Degradation by ClpAP. *Nat. Struct. Mol. Biol.* **2008**, *15* (3), 288–294. https://doi.org/10.1038/nsmb.1392.

(10) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6* (5), 343–345. https://doi.org/10.1038/nmeth.1318.

(11) Bornhorst, Joshua, Falke, J. Purification of Proteins Using Polyhistidine Affinity Tags. *Methods Enzym.* **2000**, *326*, 245–254.

(12) Malakhov, M. P.; Mattern, M. R.; Malakhova, O. A.; Drinker, M.; Weeks, S. D.; Butt, T. R. SUMO Fusions and SUMO-Specific Protease for Efficient Expression and Purification of Proteins. *J. Struct. Funct. Genomics* **2004**, *5* (1), 75–86. https://doi.org/10.1023/B:JSFG.0000029237.70316.52.

(13) Lau, Y. T. K.; Baytshtok, V.; Howard, T. A.; Fiala, B. M.; Johnson, J. L. M.; Carter, L. P.; Baker, D.; Lima, C. D.; Bahl, C. D. Discovery and Engineering of Enhanced Sumo Protease Enzymes. *J. Biol. Chem.* **2018**, *293* (34), 13224–13233. https://doi.org/10.1074/jbc.RA118.004146.

(14) Resnick, S. M.; Zehnder, A. J. B. In Vitro ATP Regeneration from Polyphosphate and AMP by Polyphosphate:AMP Phosphotransferase and Adenylate Kinase from Acinetobacter Johnsonii 210A. *Appl. Environ. Microbiol.* **2000**, *66* (5), 2045–2051. https://doi.org/10.1128/AEM.66.5.2045-2051.2000.

(15) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(16) Dumon-Seignovert, L.; Cariot, G.; Vuillard, L. The Toxicity of Recombinant Proteins in Escherichia Coli: A Comparison of Overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* **2004**, *37* (1), 203–206. https://doi.org/10.1016/j.pep.2004.04.025.

(17) Stein, B. J.; Grant, R. A.; Sauer, R. T.; Baker, T. A. Structural Basis of an N-Degron Adaptor with More Stringent Specificity. *Structure* **2016**, *24* (2), 232–242. https://doi.org/10.1016/j.str.2015.12.008.

(18) Schmidt, R.; Zahn, R.; Bukau, B.; Mogk, A. ClpS Is the Recognition Component for Escherichia Coli Substrates of the N-End Rule Degradation Pathway. *Mol. Microbiol.* **2009**, *72* (2), 506–517. https://doi.org/10.1111/j.1365-2958.2009.06666.x.

(19) PACT Premier. *Mol. Dimens.* 1–5.

(20) JCSG- plus TM Screen. 1–4.

Appendix A

TRANSFORMATION PROTOCOL

Ratio of liquids within cuvette:

- 40 μL of distilled water
- 1 μL of species plasmid (from PCR w/ Phusion Polymerase)
    - Completed product of the PCR
- 10 μL competent cells (weak cells) *E. coli* based
    - NEB10β for PLASMID
    - ER2566 for PROTEIN

Other materials:

- Cuvette
- Electroporator
- Ice bucket
- Culture Broth (1000 μL =1 mL)
- Plastic tube w/ cap
- Bunsen burner
- Ampicillin plate

Steps:

- Obtain special cuvette and add liquids in THIS ORDER
    - Water
    - Plasmid
- Mix thoroughly as you add each liquid to combine with one another
- (4-5 pipette pumps)
- Competent cells should be added before putting cuvette in electroporator
    - These cells should always be on ice bucket
    - Mix all component together
- Put cuvette on the correct side in the electroporator
    - Slide in the cuvette to lock in place
    - Turn on the electroporator
    - The setting should be Bacteria: *E. coil*
    - Zap one time
    - Switch k/v to m/s to see the amount
    - The good amount should be around ~5
- Turn on the flame to kill contaminants
- Add 1000 μL (1 mL) of culture broth to the cuvette and mix thoroughly
- Pour the liquid into the plastic tube
    - Cover plastic tube w/ cap and put into the shaker machine
    - 37 C for 45 min (for single colony growth)

- Transfer the liquid to a 1.7 mL Eppendorf tube
    - Spin down at 5,000 rpm for 5 mins
    - Make sure you have a balance for centrifuged
- Remove extra liquid
    - Save a tad bit liquid to break up the pellet in
        - Break up the pellet
- Turn on the flame and add 10-20 μL of the product to ampicillin plate
    - Use a spreader to obtain single colonies
    - Incubate overnight at 37 C

Appendix B

BUFFERS/SOLUTIONS USED FOR EXPERIMENTS

Lysis Buffer

- 25 mM HEPES pH 7.5
- 300 mM NaCl
- 10 mM imidazole pH 7.5
- 10% glycerol

Loading Dye: Protein Electrophoresis

- 450 μL of 3x Sample Buffer
- 50 μL of 2-Mercaptoethanol

Elution Buffer

- 25 mM HEPES pH 7.5
- 300 mM NaCl
- 300 mM imidazole pH 7.5?
- 10% glycerol

APD Buffer

- 50 mM HEPES pH 7.5
- 300 mM NaCl
- 20 mM $MgCl_2$
- 10% glycerol
- 0.1 mM EDTA

HpAPD Buffer

- 10 mM HEPES pH 7.5
- 150 mM NaCl
- 10% glycerol

Appendix C

NI-NTA COLUMN GRAVITY BINDING METHOD

Before using the column, make sure there are no contaminants within the column

- The amount of nickel beads used depends on the amount of protein you plan to purify
    - Rule of thumb
        - 3 mL of new Ni-beads liquid (composed of 1.5 ml of beads, 1.5 ml ethanol/water liquid) per 1 L of protein

Running Column

1. Equilibrate Ni-NTA beads with Lysis buffer
    a. Let all Lysis buffer flow out
2. Add lysate to the gravity column
    a. Collect fraction
        i. Pour collected fraction back to the top of the gravity column and recollect lysate
3. Add approximately ~15 mL of Lysis buffer to make sure no impurities are not present in the next fraction
    a. Collect this fraction
4. Wash the column with 2 column volumes of Wash buffer
    a. Collect the fractions
5. Add approximately 3 x 2 column volumes of Elution buffer and collect the fractions
    a. To check if protein is present in each elution use 45 μL of Coomassie Blue and 5 μL of collected elution
        i. If the color changes to blue, protein is detected
    b. You can collect elution until the blue color is no longer present
        i. Usually about the third elution there is no color change

Cleaning up the column

- After you finish collecting the elution, let the remaining amount of elution buffer run out the column
- Ni-NTA beads need to be stored in a 50:50 ratio of ethanol and diH2O
- Pour # volume (mL) of 100% ethanol in the column
    - Close the top and bottom of the column
    - Shake the column up and down in order to get the beads
    - Pour the beads out into the used Ni-NTA bead bottle
- Pour the same # volume (mL) of diH2O in the same column
    - Close the top and bottom of the column
    - Shake the column up and down in order to remove the rest of the beads
    - Pour the beads into the used Ni-NTA bead bottle
- Repeat the previous two steps in order to get most of the beads out the column

- Rinse the empty column with diH2O
    - Use soap and cleaning brush to clean out the column
    - Rinse with diH2O and ensure the column is clean without soap residues
    - Let the column to dry

Appendix D

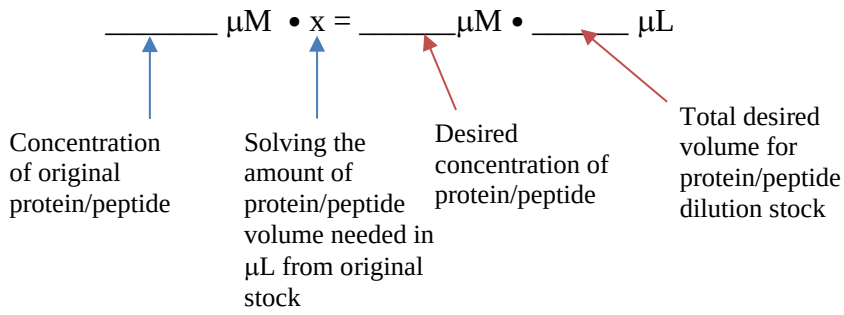
ANISOTROPY ASSAY

Dilution Protein/Peptide Stock

Depending on the desired concentration of the peptide and protein that is needed for the assay, you need to dilute the protein and peptide to make the dilution you will use for the assay.

Components	Desired Concentrations
Protein	100 μM
Peptide	0.1 μM

- Math equation to dilute the protein and peptide stock

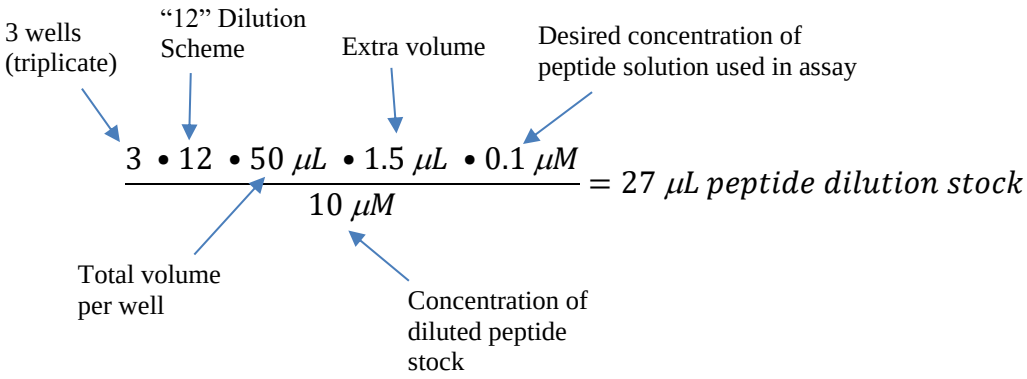


To find the amount of buffer you need for your protein/peptide dilution stock:

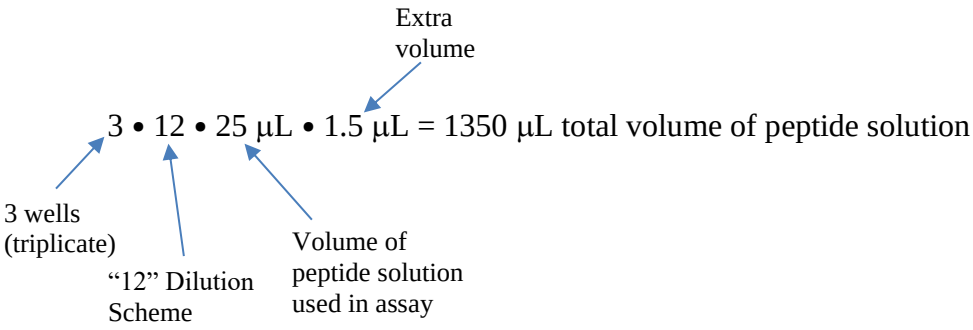
Total volume desired for protein/peptide dilution stock – amount of protein/peptide volume needed from original stock = Buffer volume needed

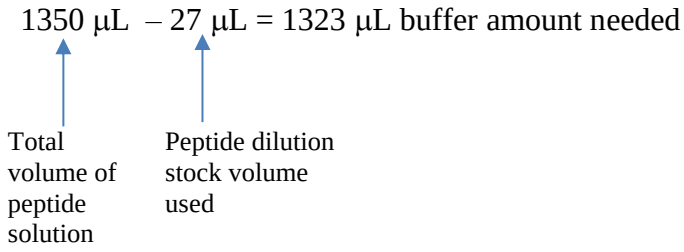
Combine the buffer volume with the protein/peptide volume calculated from the original stock.

Calculating amount of diluted peptide (µL) needed for assay



- Find total volume of peptide solution to find buffer (µL)





- Combine the buffer amount needed and the peptide dilution stock volume used together in a tube

Calculating amount of diluted protein (µL) needed for assay

Using the same math that was used to calculate the volume needed from the diluted stock of peptide and buffer volume, repeat the same calculations for the diluted protein and the buffer that the solution needs.

- The desired concentration for the protein solution used in this assay is 100 µM

    - The diluted protein stock has a concentration of 700 μM
- To find the total volume for the protein solution, repeat the same calculation done for the peptide
    - Both protein solution and peptide solution volumes within each well will be 25 μL, making the total volume of the well 50 μL
- Combine the buffer volume needed with the solved diluted protein stock in a tube

Making the dilution scheme for the assay

Using the protein solution and peptide solution for the dilution scheme

- Number twelve tubes
    - In 11 tubes, put 100 μL of buffer of choice
    - The 12th tube is for reference
        - 100 μL of peptide solution
        - 100 μL of buffer
- Transfer 100 μL of protein solution to the FIRST tube
    - Mix the buffer and protein in tube
    - Pipette 100 μL of the liquid from tube one and transfer the volume to the second tube
- Repeat transferring the 100 μL to the next tube
    - Mix the dilution as you go
- When you reach the 11th tube, place the 100 μL from the 10th tube and mix the dilution
    - Discard 100 μL from tube 11 afterwards

The protein has been diluted and these are the concentrations within the tubes

Tube Number	Concentration of Protein
1	50 μM
2	25 μM
3	12.5 μM
4	6.25 μM
5	3.13 μM
6	1.56 μM
7	0.78 μM
8	0.39 μM
9	0.19 μM
10	0.10 μM
11	0.05 μM

12	0 μM (peptide and buffer only)

Adding peptide solution to dilution scheme

- Using a repeat pipette, load 100 μL of peptide stock into each tube BESIDES tube 12
    - Mix peptide and protein dilution as you go
- Create diagram to keep track of dilution with the plate well
    - Each dilution and reference (tube 12) is tested in triplicates
- Use an extra well on the plate to put the highest concentration of protein
    - 50 μM is the highest concentration
        - 25 μL of protein solution
        - 25 μL of buffer

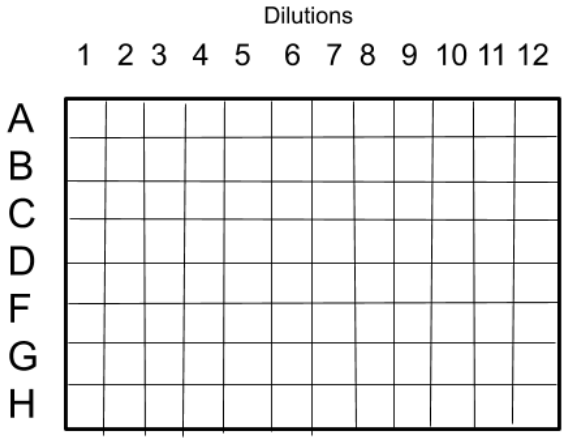


Image of the plate with the wells.

- In each well there should be a total of 50 μL in each well
    - Avoid the bubbles
- After all the samples have been placed in the plate in triplicates, use the plate reader and conduct Anisotropy Assay method

Appendix E

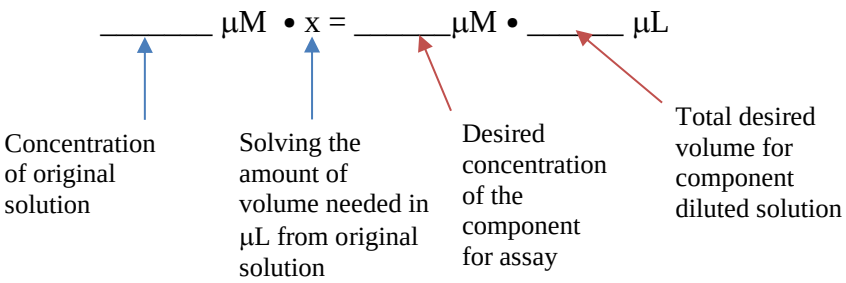
PROTEASE ASSAY

Concentrations used in this Protease Assay

Components	Concentrations
$_{Ec}$ClpA	0.3 μM
$_{Ec}$ClpP	0.5 μM
$_{Ec}$ClpS	1 μM
$_{Hp}$ClpS	1 μM
$_{CP7}$GFP-ssrA	10 μM
PK regeneration	1 x

Depending on the current stock of these components, you may have to make a dilution version in order to obtain the desired concentration for the assay.

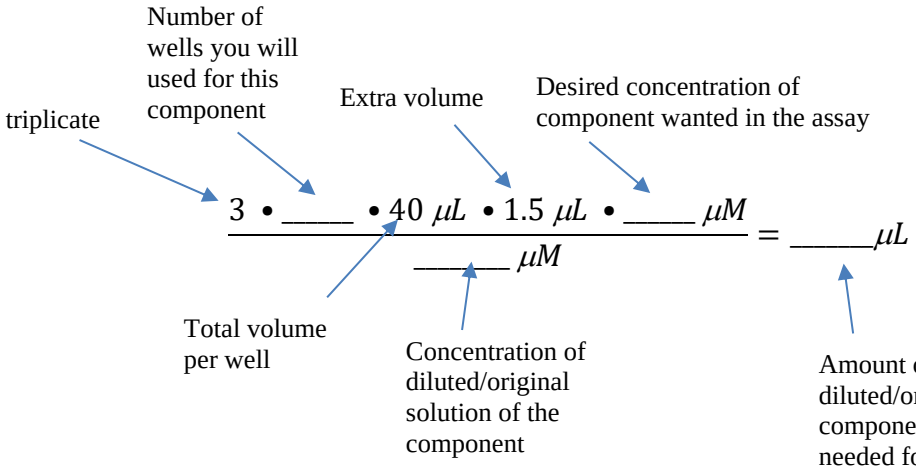
Dilution Math Equation



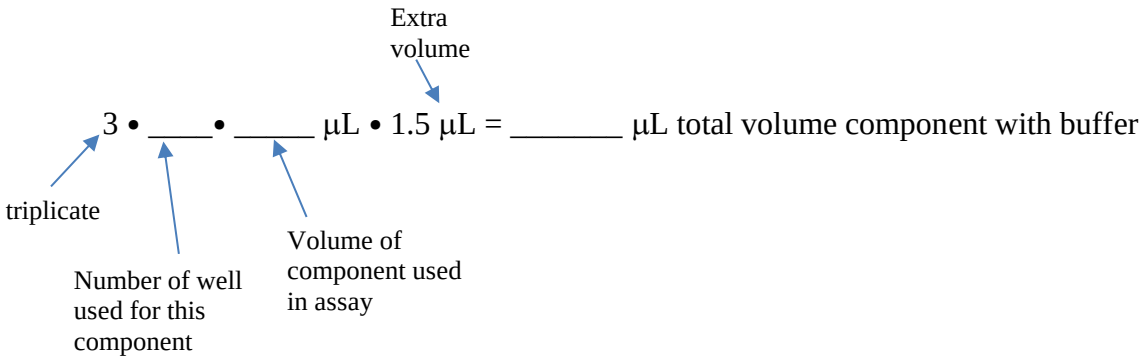
- Use this equation to make a dilution solution for the component if needed

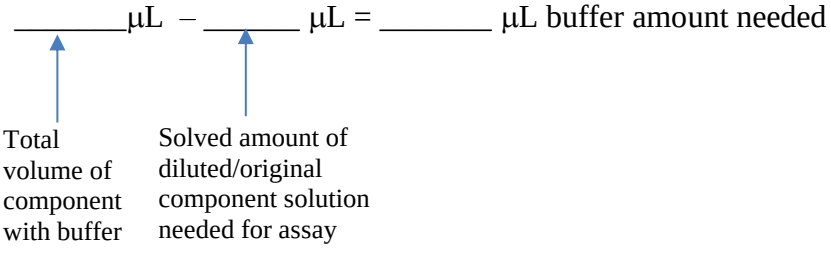
Math to set up assay

- **For each component of the assay**, you need to calculate how much of the diluted/original solution you need in order to obtain the desired concentration for this assay (concentration listed in table above).



- You need to solve for the total volume of this component, in order to find the amount of buffer you need to obtain the concentration
    - For EcClpA and EcClpP, I combined them together in the tube. So, when the total volume is solved, subtract ClpA and ClpP from that volume to obtain the buffer needed.





Volume of component used in Assay

Per well, the total volume should be 40 µL.
The table below showcases the volumes of each components in order to meet the total volume and concentration.
Depending on what you are testing, you have to adjust the volume of the components to meet the total 40 µL per well

- Make a chart and the volumes used to keep track

Sample Component	Volume
Buffer Blank	40 µL
GFP Blank	15 µL GFP + 25 µL buffer
$_{Ec}$ClpAP alone	15 µL $_{Ec}$ClpAP + 25 µL buffer
$_{Ec}$ClpAP and GFP	15 µL $_{Ec}$ClpAP + 15 µL GFP + 10 µL buffer
$_{Hp}$ClpS and Buffer	5 µL $_{Hp}$ClpS + 35 µL buffer
$_{Ec}$ClpAP, GFP substrate + PK regen.	15 µL $_{Ec}$ClpAP + 15 µL GFP + 10 µL PK regen.
$_{Eco}$ClpAPS, GFP substrate + PK regen.	10 µL $_{Ec}$ClpAP + 5 µL $_{Ec}$ClpS + 15 µL GFP + 10 µL PK regen.
$_{Ec}$ClpAP, $_{Hp}$ClpS, GFP substrate, PK regen.	10 µL $_{Ec}$ClpAP + 5 µL $_{Hp}$ClpS + 15 µL GFP + 10 µL PK regen.

- The sample components should be in triplicates
- For this assay I used a 96-well flat bottom black microplate
    - I used APD buffer for my assay
- Combine the sample components within the wells
    - Make sure they are mixed
- For wells that have ClpAPS components, let the wells sit for 5 minutes before adding PK regeneration
- The PK regeneration is the last component that should be added at the very end
    - Before running the method in the plate reader, quickly add the PK regeneration to the correct wells
    - Then start running the protease assay method in the plate reader program