MONITORING PROTEOLYTIC EFFICIENCY OF ENGINEERED TRYPSIN VIA CHYMOTRYPSINOGEN ACTIVITY ASSAY

A thesis submitted to the faculty of
San Francisco State University
In partial fulfillment of
The requirements for
The Degree

Master of Science
In
Chemistry
Concentration in Biochemistry

By

Rodolfo Villa

San Francisco, California

San Francisco State University

July 2019

CERTIFICATION OF APPROVAL

I certify that I have read *MONITORING PROTEOLYTIC EFFICIENCY OF ENGINEERED TRYPSIN VIA CHYMOTRYPSINOGEN ACTIVITY ASSAY* by Rodolfo Villa, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Science in Chemistry: Biochemistry at San Francisco State University.

Dr. Teaster Baird, Jr.
Principal Investigator

Dr. George Gassner
Professor

Dr. Ray Esquerra
Professor

MONITORING PROTEOLYTIC EFFICIENCY OF ENGINEERED TRYPSIN VIA CHYMOTRYPSINOGEN ACTIVITY ASSAY

Rodolfo Villa
San Francisco, California
2019

Protease therapeutics have been on the rise and have gained recognition for having diverse clinical applications. One major complication faced is the abundance of protease inhibitors that serve to regulate proteolytic activity resulting in a short therapeutic half-life. Previous research with the model serine protease trypsin showed that residues 39 and 60 play a key role in protease: inhibitor binding. Compared to wild type, four trypsin single variants (Y39A, Y39F, K60A, K60V) which were all catalytically similar using a synthetic substrate displayed altered sensitivity towards bovine pancreatic trypsin inhibitor (BPTI) compared to wild type. In order to ascertain the viability of these engineered variants for inhibitor resistance in vivo, the interactions between naturally occurring macro molecular substrates were evaluated in this study. Variant Y39A displayed the highest k_{cat}/K_M ($\mu M^{-1}Min^{-1}$) at 3.34 ± 0.06, while K60V displayed at 0.58 ± 0.01. Variant Y39A activated the most chymotrypsinogen (Cg) with a total 30% activation 1.48 ± 0.05 μM (out of 5 μM) while K60V had the lowest with 14% or 0.70 ± 0.04 μM. With 20 nM BPTI, the K60A variant had an activation drop of 49% while K60V dropped by 16%. At 30 nM, 50 nM and 100nM BPTI, the Y39F variant displayed greater activation than wild-type and the K60 variants in the presence of BPTI. At higher concentrations of BPTI, the K60 variants s have become more sensitive to inhibition, along with having lower k_{cat}/K_M and less total chymotrypsin activation. Overall K60 variants seemed to be less proteolytically efficient towards macromolecular substrates with less resistance towards inhibition, while Y39F displayed promising results in both studies.

I certify that the Abstract is a correct presentation of the content of this thesis.

Chair, Thesis Committee

7.31.19
Date

ACKNOWLEDGEMENTS

I would like to thank the support of the Student Enrichment Opportunities program that allowed me to take part in the Bridges to the Doctorate program which provided me with not only with financial assistance but allowed my development as a scientist.

I extend my sincerest gratitude to my principal investigator of this work Dr. Teaster Baird, who took me into his lab and provided with mentorship support with my studes and research, for his patience, motivation, and immense knowledge that has been shared. I could not have asked for a better mentor and advisor.

I would also like to thank the rest of my thesis committee: Dr. Ray Esquerra, and Dr. George Gassner, for their encouragement and support.

Last, I would like to thank my parents Rodolfo Villa and Candelaria Villa for supporting me on every move and decision I have made throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figures Page

EQUATIONS

Equations Figures

Introduction

Proteases, also known as proteolytic enzymes, are a complex group of enzymes that are responsible for catalyzing the breakdown of proteins and peptides. They can be found in animilia, bacteria, fungi, plantae, viruses, and archaea. Their functions span a wide range of applications that include cell proliferation and differentiation, hemostasis, blood coagulation, inflammation, immunity, and apoptosis.[1] Approximately 2-4% of the typical human genome contains genes encoding for proteases making them one of the most abundant and diverse group of enzymes. [2] Over the decades, their molecular mechanisms and regulation have been widely studied. Along with their physiological importance, proteases are also one of the three largest groups of industrial enzymes and account for about 60% of the total worldwide sale of enzymes, providing an economic benefit.[3] Their physiological roles also make proteases attractive for drug design. The U.S. Food and Drug Administration (FDA) has approved twelve protease therapies, and a number of next generation or completely new protease therapies are in clinical development.[4] Due to the abundance of inhibitors that tightly regulate proteolytic enzymes, proteases tend to have relatively short half-lives causing major problems when designing protease therapies. Engineering proteases for inhibitor resistance has been a focus of the Baird Lab, which studies serine protease activity and inhibitor interaction.

Protease Classification

The protease classification system divides them into clans based on their catalytic mechanism and families based off common ancestry. Over one third of all known proteolytic enzymes are serine proteases grouped into 13 clans and 40 families.[2] The general properties, catalytic residues, clans, and specificities of serine proteases are listed in *Table 1*. The PA protease clan is the largest family of serine proteases and perhaps the best studied group of enzymes.[2] Many of the proteolytic enzymes within the PA protease clan share three-dimensional structural traits. Trypsin is the classical representative member and is used as a structural reference; when referring to structural similarities within the PA clan it is common to refers to the enzyme as having the trypsin-fold. PA clan proteases not only share structural similarities, but also catalytic residues which are responsible for cleaving peptide bonds through a hydrogen-bond network within the active site. The primary specificity of a protease refers to the amino acid residue whose peptide bond gets cleaved (the P1 residue). The identity of this residue is partially determined by specific residues found in the active site. For example, trypsin, whose primary specificity is basic residues, contains a negatively charged asparate residue (Asp^{189}) in its active site that complements the positively charged residues lysine and arginine in substrates. Serine alkaline proteases are produced by several bacteria (like *Arthrobacter*, *Streptomyces*, *Flavobacterium*), molds, yeasts, and fungi.[5] Examination of the amino acid sequences of proteases has allowed the assignment to evolutionary families that can be grouped together to

signs of distant relationship, with as many as 60 evolutionary lines with separate origins.[6] At the highest level, proteases are distinguished by the pH at which their activity is a maximal. Some work best at either acidic, basic, or neutral pH values [1]. The family names stem from the nucleophilic residue in the enzymes active site, which is responsible for nucleophilic attack of the peptide bond. This nucleophilic species can range from serine, cysteine, aspartate, glutamate, and even water activated by metallo-ions such as zinc. In this study, the focus will be on the PA clan serine proteases trypsin and chymotrypsin.

Clan	***Families***	***Representative Member***	***Catalytic Residues***	***Primary Specificity***
PA	12	Trypsin	His, Asp, Ser	A, E, F, G, K,Q, R,W, Y
SE	6	D-A, D-A carboxypeptidase	Ser, Lys	D-S
ST	5	Rhomboid	His, Ser	D, E
SP	3	Nucleoporin	His, Ser	F
SB	2	Subtilisin	Asp, His, Ser	F, W, Y
SC	2	Prolyl oligopeptidase	Ser, Asp, His	G, P

Table 1: *Representative members of serine protease clans along with their catalytic residues according to their representative members.* [2]

Serine Proteases

Serine proteases are named for the nucleophilic serine residue located within the active site. When referring to the residue positioning of the catalytic triad, chymotrypsin residue sequence is used as the reference numbering system. In chymotrypsin, the nucleophilic serine residue is located at position 195, while trypsin is located at residue 183. Although chymotrypsin and trypsin amino acid sequences have slight differences, their tertiary structures both are nearly superimposable where the active and inhibition site are located in the relative three-dimensional locations. Trypsin, a canonical serine protease, has two characteristic regions that are essential for proteolysis in its active site; the catalytic triad and the oxyanion hole. The catalytic triad, three residues that are involved in the cleavage of peptide bonds, consist of a general base (His^{57}), a general acid (Asp^{102}), and a nucleophile (Ser^{195}). The triad spans from the active site cleft with Ser^{195} on one side and Asp^{102} and His^{57} on the other, depicted in green on *Figure 1*. These residues create a hydrogen bond network that allows the stabilization of the nucleophile to attack the peptide. The hydrogen bonding is usually observed between the Nδ1-H of His^{57} and Oδ1 of Asp^{102} and between the OH of Ser^{195} and the Nε2-H of His^{57}.[7] The hydrogen bond created by His^{57}-Asp^{102} has an orientation that is parallel to the P1 carboxylate. This allows the hydrogen bond to form on the more basic electron pair from the His^{57}. A tetrahedral intermediate occurs when the bond arrangement around the P1 carbonyl is converted from a trigonal geometry to tetrahedral allowing the reaction to proceed. The formation of the tetrahedral arrangement forms an oxyanion that is stabilized through hydrogen

interaction with backbone amide hydrogens from Gly^{193} and Ser^{195}, the oxyanion hole. Along with the catalytic triad, the oxyanion hole is also linked to the Ile^{16}-Asp^{194} salt bridge via Ser^{195}. [7] The geometric relation of Asp^{102}, His^{57} and Ser^{195} leads to the role that His^{57} serves for transferring the proton from Ser^{195} to Asp^{102} in a charge relay mechanism, *Figure 2*. [8]

Figure 1: *Structure of wild-type trypsin (PDB ID: 1ANE)[9]. The green residues represent the catalytic triad composed of His^{57}, Asp^{102}, and Ser^{195}. The yellow residues represent the di-sulfide bond formed by Cys^{191} and Cys^{220}. The grey residues show the S1' inhibition pocket, composed of Tyr^{39}, His^{40}, Phe^{41}, and Lys^{60}. (Adapted from Fontecilla-Camps, J.C., 1995)[9]*

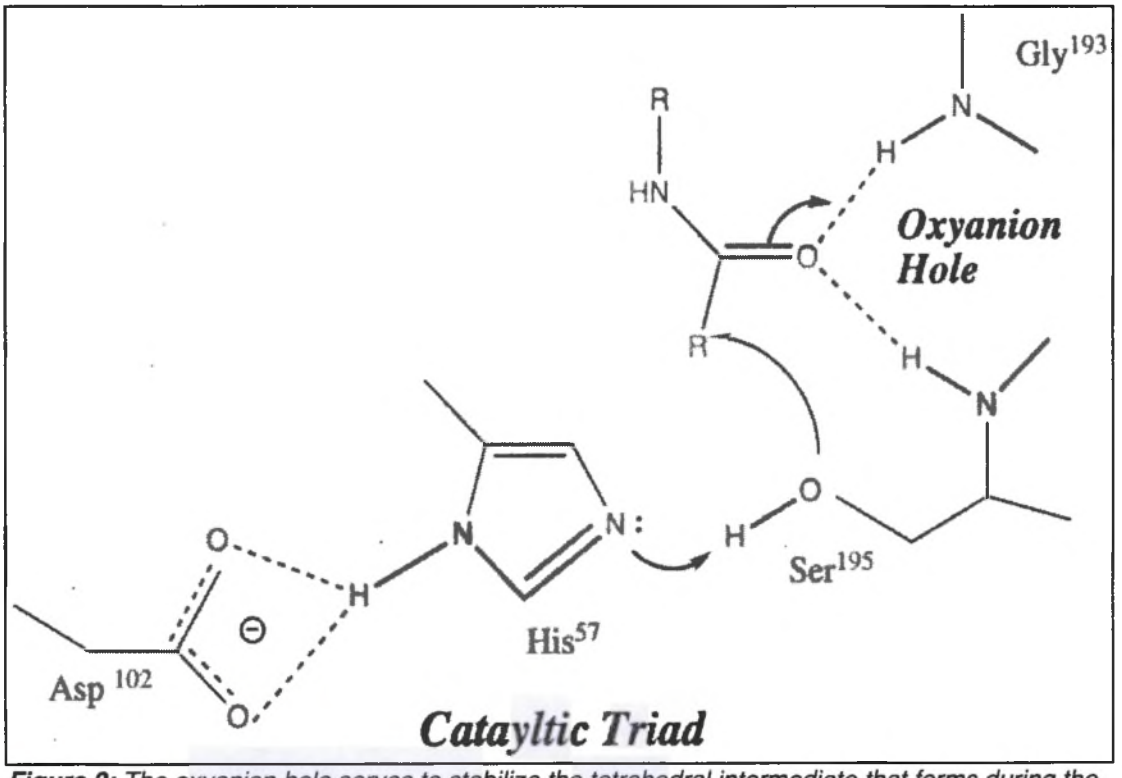


Figure 2: *The oxyanion hole serves to stabilize the tetrahedral intermediate that forms during the catalytic cycle. The amide NH groups from the back bone of Gly[193] and Ser[195] form hydrogen bonds with negatively charged oxygen from the carbonyl of the substrate.*

Catalytic Mechanism

The serine protease catalytic mechanism is one of the more studied and well-understood mechanisms in biochemistry. To cleave peptide bonds, proteases must overcome several chemical barriers: (a) peptide bonds are very stable due to electron donation from the amide nitrogen to the carbonyl oxygen; (b) water is a poor nucleophile; proteases always activate water, usually via a general base; and (c) amines are poor leaving groups; proteases protonate the amine prior to expulsion.[7] Nucleophilicity from the hydroxyl group of the catalytic Ser^{195} is typically dependent on interactions with Asp^{102} and His^{57}, commonly referred to as the charge relay system.[2] This charge relay system involves multiple simultaneous hydrogen interactions that allow proteolysis to occur. Without a proper nucleophilic residue peptide cleavage cannot occur. Previous studies done in 2006, show the importance of Ser^{195} when threonine, a residue that is chemically similar to serine, was substituted at position 195. The $Ser^{195} \rightarrow Thr^{195}$ substitution caused inactivity of the protease due to the negative steric interactions between the methyl group on the β-carbon of Thr^{195} and the disulfide bridge formed by cysteines 42 and 58.[10] This charge relay mechanism begins by Asp^{102} immobilizing His^{57} through hydrogen bonding. This causes His^{57} to act as a general base and remove a proton from the nucleophilic residue Ser^{195}, facilitating cleavage of the carbonyl carbon of the peptide bond. Next, His^{57} donates a proton to the amide nitrogen of the peptide creating a tetrahedral intermediate. This facilitates bond disassociation into a smaller peptide fragment. Due to the instability of the negative charge from the

oxygen the tetrahedral intermediate breaks down followed by the removal the peptide product, leaving a more stable acyl-intermediate. This is followed by a nucleophilic attack on the carbonyl carbon by water which generates a second tetrahedral intermediate. His^{57} accepts a proton from the water which causes the collapse of the intermediate, assisted by a proton donation from His^{57} to the serine oxygen, regenerating the original catalytic triad. Serine proteases increase the rate of peptide bond hydrolysis by a factor of $\sim 10^{10}$ compared to that of an uncatalyzed reaction. [11] To control cleavage by these proteases, there is an abundance of naturally occurring inhibitors that contribute to regulation of protease activity.

Inhibition

Although serine proteases contribute greatly to maintain homeostasis, they also have the potential to be hazardous to their natural enzymatic environments if they are not properly regulated. The primary level of regulation is by control of expression/secretion. A second level of regulation occurs through inhibition. There are three different types of natural inhibitors that can be distinguished based on the mechanism of action: canonical inhibitors, non-canonical inhibitors, and serpins. Canonical inhibitors bind to the enzyme through an exposed convex binding loop, which is complementary to the active site of the enzyme. Non-canonical inhibitors interact through their N-terminal segments. Serpins interact with their target proteases in a substrate-like manner, which is similar to the canonical inhibitors.[12] A main complication of developing protease therapeutics is

the abundance of these naturally tight binding inhibitors. Serpins, the largest group of protein inhibitors are canonical inhibitors that act according to the standard mechanism of inhibition.[13]. Serpins are a specific class of serine protease inhibitors not only because of their large super family of proteins, but unlike most other families of protein protease inhibitors, the serpins inhibit their target proteases through a unique branched pathway suicide substrate mechanism in which the acyl-intermediate is trapped by large scale conformational changes in both the serpin and protease.[14] Description of serine protease inhibitor features are listed in *Table 2.*

Inhibitor	***Examples***	***Major features of inhibition***	***Size***
conical inhibitors	BPTI, OMTKY3, eglin	Tight, non-covalent interaction resembling enzyme-substrate Michaelis complex, direct blockage of the active site, no conformational changes, antiparallel β-sheet between enzyme and inhibitor. Similar mode of interaction through canonical protease-binding loop despite completely different inhibitor structures.	3-21 kDa per domain
non- canonical inhibitors	hirudin, TAP, ornithodorin	Extremely strong and specific interaction so far known for factor Xa and thrombin only, two-step kinetics, inhibition of the active site through N terminus of the inhibitor, two areas of interaction.	6-8 kDa per domain
serpins	Antitrypsin, antichymotrypsin	Irreversible covalent acyl-enzyme complex, huge conformational changes in inhibitor, disruption of protease active site.	45-55 kDa

Table 2: *Major features of protease inhibitors of serine proteases.* [12] *BPTI, bovine pancreatic trypsin inhibitor; OMTKY3, turkey ovomucoid third domain; TAP, tick anticoagulant peptide.*

The target substrate residue (the P1 residue) generally defines the specificity of the protease. Thus inhibitors with P1 Lys and Arg tend to inhibit trypsin and trypsin-like enzymes. [13] There is great significance in the study of interactions between proteases and their inhibitors. They not only provide a basis for studying protein-protein recognition at the molecular level but minimizing inhibition and preserving catalysis are required for advancing the use of proteases as a safe and efficacious drug class.

In this study we worked with bovine pancreatic trypsin inhibitor (BPTI). BPTI inhibits several types of trypsin fold serine proteases such as trypsin, chymotrypsin, and plasmin. BPTI is a member of the kunitz family of protease inhibitors. Kunitz-type protease inhibitors are proteins of approximately 20 kD that bind tightly to the active site of the protease in a substrate-like manner, acting as a competitive inhibitor. [15] Though the interaction is reversible, some have such strong affinities for their target protease that the reaction can be considered irreversible.[16] The inspection of co-crystal structures of trypsin with various inhibitors revealed a conserved network of hydrogen bond interactions between the protease and the inhibitor, illustrated in *Figure 3.* Fully understanding protease-inhibitor interactions is crucial for the development of new protease therapies.

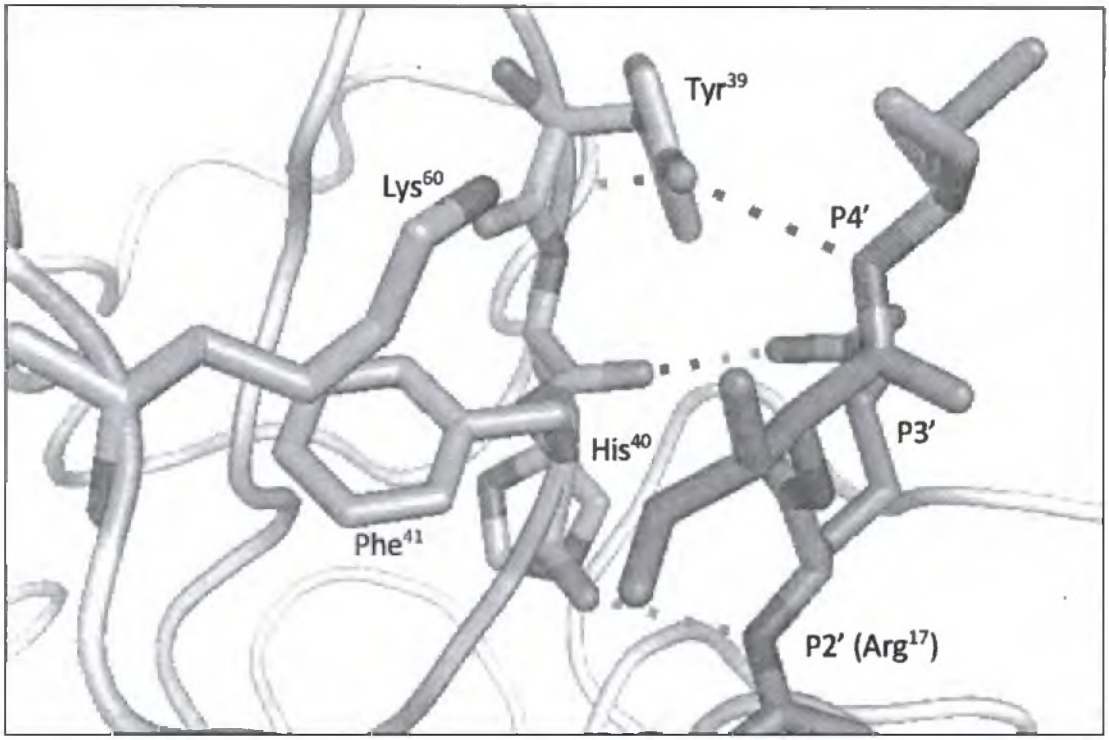


***Figure 3: Wild-type trypsin-inhibitor complex**. The S1' pocket of trypsin residues is shown in green while BPTI residues are shown in purple, and hydrogen bonds shown in yellow dashes. (PDB ID:1TPA).* [17] [9]*The backbone carbonyl oxygen of phenylalanine at position 41 hydrogen bonds with the backbone amide of the inhibitor at P2', along with the hydroxyl group of tyrosine at position 39 makes a hydrogen bond with the backbone amide of inhibitor at P4'. These hydrogen bonds are critical for inhibitor studies and are conserved in almost all serine proteases inhibitor complexes.*

Protease Therapeutics

The development of protease therapeutics have been on the rise in recent years and have a very promising future. For example, Activase (Alteplase), one of the more reliable serine protease therapeutics, was released to the public in 1987. It is the protease tissue plasminogen activator (t-PA) and is used to treat stroke victims. Activase is produced by recombinant DNA technology. When administered, Activaseactivates the protease zymogen plasminogen to plasmin which initiates the breakdown of the fibrins that created the blood clot (thrombus) that caused the stroke.[18]. When introduced into the systemic circulation, Activase acts in the following manner. First, it binds to the fibrin protein threads of a thrombus. Once bound, it then converts the enmeshed plasminogen to plasmin, initiating local fibrinolysis. In the absence of fibrin, Activase produces limited conversion of plasminogen to plasmin, thus causing a limited systemic effect, which is to dissolve clots allowing blood flow to be restored. [19]

Although they are a well-recognized class of *targets* for inhibitors, proteases themselves have not been typically considered as a drug class despite their application in the clinic over the last several decades. [4] Today, twelve protease therapies that have been approved by the United States Food and Drug Administration (USFDA) and a number of potential new therapies are in the clinical development phase. The ability to harness proteolysis for disease treatment will increase our understanding of protease biochemistry and the molecular mechanisms responsible for these diseases. Along with the ability to engineer

proteases, as well as improving delivery options, there will be a great expansion for potential applications of these enzymes. The recognition that proteases are, in fact, an established class of safe and efficacious drugs has already stimulated investigation of additional therapeutic applications for these enzymes. Protease based therapeutics have specific advantages when used as procoagulants and for thrombolysis, as seen in *Table 3*. Out of the 12 approved therapeutics, 10 are serine proteases. One major complication faced when using serine proteases as therapeutic agents is their relatively short half-life due to the abundance of tight binding inhibitors found in the human body that serve to regulate proteolytic activity. Developing an effective protease therapeutic requires resistance towards naturally occurring inhibitors while conserving its catalytic activity. To properly develop inhibitor resistance, one approach is to identify key residues involved in protease-inhibitor interactions. Since the majority of protease therapeutics currently in use are trypsin-fold, trypsin serves as a perfect model for the investigation of protease-inhibitor interactions.

Usage	Protease	Indications	Source of protein	Target protein or pathway	Type of protease	FDA Year approved
Thrombolysis	-Urokinase (u-PA) -t-PA (alteplase,Activase®) -Reteplase (Retevase) -TNK-tPA(tenectaplase, Metalyse®)	-Thrombus, catheter clearing - AMI, stroke, catheter clearing - AMI - Myocardial Infraction	-Urine or kidney cell -Recomb. CHO cells -Recomb. E. coli cells -Recomb. CHO cells	-Converts Plasminogen into plasmin -Plasminogen Activator -Plasminogen Activator -Plasminogen Activator	-Serine -Serine -Serine -Serine	-1978 -1987 -1996 -2000
Procoagulant	-FIX -FIX(BeneFIX®) -FVIIa (Novosen®) -Thrombin (Recothrom®)	- Haemophilia B - Haemophilia B - Haemophilia A & B - Bleeding	-Human Plasma - Recomb. CHO cells - Recomb. BHK cells - Recomb. CHO cells	-FX Activator -FX Activator -FX and FIX Activator -Fibrinogen Activator	-Serine -Serine -Serine -Serine	-1990 -1997 -1999 -2008
Sepsis	Activated protein C, (drotrecogin alfa, Xigirs®)	-Sepsis, septic shock	-Recomb. Human cell line	-Plasminogen Activator	-Serine	-2001
Neuromuscular	-Botulinum toxin A (Botox®) -Botulinum toxin B (Myobloc)	-Various muscle spasms -Cervical dystonia	-Bacterial (C.botulinum)	-Santaxin and SNAP-25 deactivator -Synaptobrevin deactivator	-Zinc -Zinc	-1989 -2000
Digestion	-Zenpep® (pancrelispase)	-Exocrine Pancreatic Insufficiency	-Porcine pancreatic extract	-Aids digestion of protein	-Serine	-2009

Table 3: *FDA- approved protease drugs. Serine proteases constitute for most of the approved therapeutics in the market.* [4]

Protease Engineering

Trypsin has been extensively studied and is structurally similar to many serine proteases making it a good scaffold enzyme to engineer and allowing for easier application towards other serine proteases. Previously, co-crystal structures of trypsin-inhibitor complexes were inspected with the intent of identifying residues that may be key in protease-inhibitor interactions. A conserved hydrogen bond interaction between the backbone carbonyl of Phe^{41} and the backbone amide of the P2' residue of the inhibitor was observed across complexes. It was also observed that a simultaneous hydrogen-bond network occurred between Lys^{60} and Tyr^{39} of trypsin and the P4' residue of the inhibitor, as illustrated in *Figure 3* with bovine pancreatic trypsin inhibitor.[17] The conservation of this interaction suggested significance with respect to inhibitor binding and molecular interactions. Consequently, it was hypothesized that substitution of residues Lys^{60} and Tyr^{39} in the S1' site of trypsin might allow the development of inhibitor resistance.

The identification of key residues in protease-inhibitor interactions was carried out by *Batt. et. al.* at San Francisco State University.[17] The P1' prime site of trypsin is the location for the most conserved trypsin-inhibitor interactions, making residues 39 and 60 prime targets. A single amino acid substitution for these residues was predicted to increase inhibitor resistance by eliminating certain protease-inhibitor interactions. Four trypsin single variants Y39A, Y39F, K60A, and K60V were created and their sensitivities to inhibition were tested against bovine pancreatic trypsin inhibitor (BPTI) and M84R ecotin. The results showed that, in

the presence of inhibitors, some of the substitutions affected the catalytic activity of the variants against small commercially available substrates. However, differences were observed. The Y39A variant was less sensitive to BPTI and M84R ecotin, while the Y39F variant was more sensitive to both. The experimental results were supported by computational modeling. The BPTI:Y39F complex resulted in the lowest relative binding energy, despite the removal of hydrogen bonding capability with the P4' BPTI residue, while BPTI:Y39A resulted in the highest binding energy. Compared to wild-type, the K60A and K60V variants showed increased sensitivity to BPTI but less sensitivity to ecotin.[17] Additionally, burst inhibition studies showed that Y39A trypsin had a slower association rate (k_{on}) and higher dissociation rate (k_{off}). The Y39A trypsin:inhibitor complex is less stable with a quicker dissociation, showing resistance towards the formation of the complex. The outcomes suggest that residue 39 might be a better candidate for reducing trypsin-inhibitor interactions. Although the results were promising, the concern with these experiments was that the activity was measured with a small tri-peptide substrate Z-Gly-Pro-Arg-*p*-nitroanilide (Z-GPR-pNA, *Figure 4).* In vivo, these proteases do not interact with small substrates such as Z-GPR-pNA, but with proteins. Furthermore, given that BPTI interacts with trypsin in a manner similar to a substrate, there is the possibility that any weakened interaction between trypsin and BPTI may also be observed between trypsin and a protein substrate. The purpose of this study was to test these hypotheses.

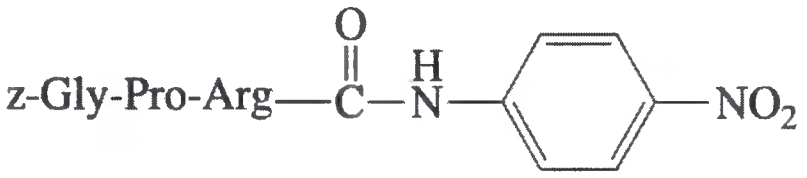


Figure 4. *Tripeptide substrate: Z-Gly-Pro-Arg-p-nitroanilide (Z-GPR-pNA). The peptide bond that is cleaved is shown in atomic detail.*

Experimental Design

An effective protease therapeutic requires resistance towards naturally occurring inhibitors while conserving its catalytic activity. Previously engineered trypsin variants Y39A, Y39F, K60A, K60V retained their catalytic activity against small commercially available substrates while the Y39A variant displayed the most promising results with respect to resistance to inhibition. Since the majority of the protease therapeutics currently in use are of the trypsin-fold, it is logical to use the protease trypsin in this study. The focus of this work is to understand the interactions between the trypsin variants and a macromolecular substrate. In vivo, chymotrypsinogen is activated by trypsin through the cleavage of the peptide bond between Arg^{15} and Ile^{16}, resulting in a liberated NH_3^+ terminus at $Ileu^{16}$ which forms an ion-pair with the carboxylate sidechain of Asp^{194} and triggers the formation of the active enzyme, α-chymotrypsin.[20] Given this relationship, chymotrypsinogen was used as the macromolecular substrate. To accurately determine the concentration of active protease present in a given sample, active site titration is routinely utilized. For this study, α-chymotrypsin was titrated with 4-

methylumbelliferyl *p*-trimethylammoniocinnamate chloride (MUTMAC), allowing quantification of active protease in the presence and absence of inhibitor. *Figure 5* illustrates the scheme of the activation. We hypothesized that the trypsin variants will retain their proteolytic activity toward chymotrypsinogen while displaying a weaker interaction with BPTI. To test this hypothesis, we monitored MUTMAC hydrolysis, which is a specific active-site titrant for α-chymotrypsin. Using a standard curve, we then determined the concentration [μM] of α-chymotrypsin at certain time points. This was then plotted and fitted using the exponential burst equation allowing us to determine k_{cat}/K_M for comparison.

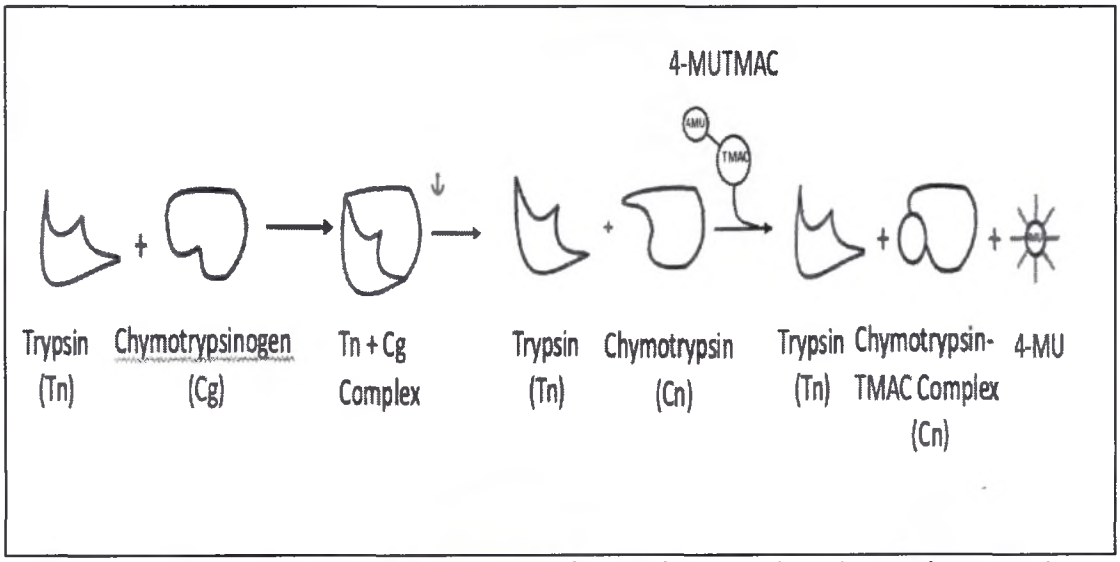


Figure 5. *Displays of kinetic characterization scheme where trypsin activates chymotrypsinogen then followed by titration of 4-MUTMAC.*

Additionally, the question remains how these trypsin variants will fare against inhibitors while in the presence of macromolecular substrates. Since previous work was done with synthetic peptides, we performed inhibition studies of the engineered variants with various inhibitor concentrations of BPTI using chymotrypsinogen as a substrate instead of the tri-peptide Z-GPR-pNA. Through such inhibition studies we will be able to get a better understanding of the effect of the substitutions on inhibition resistance and catalytic activity under condition that are more similar to those that would be found in vivo.

Methods

Generation of Trypsin Variants

Trypsin variants Y39A, Y39F, K60A, and K60V were expressed as described in Baird et. al, 2006.

Expression of Wild Type Trypsin

Pichia pastoris cells that were transformed with a plasmid that contained the genes for rat ionic trypsin were plated on YPDS-Zeocin plates (1% yeast extract, 18% D-sorbitol, 2% peptone 2% agar, 20% dextrose, 100 μg/mL Zeocin™) then incubated at 30°C for two-three days. Following the incubation phase, the healthiest isolated colonies were selected with a sterile pipette tip which was ejected into a sterile culture tube with 4mL of low salt liquid YPDS-Zeocin media and incubated in a shaker over night at 30°C.

The YPDS-Zeocin culture with the highest optical density was added into 1 liter of buffer glycerol complex media (BMGY) and incubated overnight at 30°C. Methanol was added over the next 3-5 days to induce expression and gel samples were prepared. The culture was then centrifuged at 4°C for 10 minutes at 9000 rpm. The crude protein was collected in clean, sterilize bottles and SDS-PAGE (4%-12% gradient) was used to confirm the presence of trypsinogen.

Fast Protein Liquid Chromatography (FPLC)

After the supernatant was collected, NaCl and MES were added to reach a final concentration of 3M and 20mM respectively and the solution was adjusted pH to 6, if necessary. This high-salt supernatant was loaded onto a phenyl sepharose hydrophobic interaction chromatography column (Pharmacia) that was equilibrated with high-salt buffer (3M NaCl, 20mM MES, pH 6). This was followed by washing the column with the high-salt buffer and then eluting into 5mL fractions through the use of a negative linear salt gradient (3M – 0M NaCl, 20mM MES at pH 6). Fractions that contained trypsinogen were identified using 12% SDS-PAGE gels, combined, dialyzed into 4L of 10mM glycine at pH 3 at 4°C with three changes of buffer and stored at -20°C until use.

Zymogen Activation

Conversion of trypsinogen to trypsin was achieved by dialyzing 3mL aliquots into 1 L of activation buffer (20mM HEPES, 50mM NaCl, and 2mM $CaCl_2$, pH 8.0) at room temperature. This promotes auto activation of trypsinogen. To measure activation rate, we combined 15μL of 1.5 mM z-GPR-pNA, 4 μL of activating trypsin, and 481 μL of kinetic assay buffer (50 mM HEPES, 100 mM NaCl, 20 mM $CaCl_2$ at pH 8.0), then measured change in absorbance of p-nitroanilide (λ = 410 nm). Readings were taken every 10-20 minutes. Activation was considered to be complete after absorbance no longer increased.

Separation of Trypsin from Trypsinogen

Trypsin was separated from trypsinogen and purified using a p-aminobenzamidine agarose affinity column. The column was prepared by washing with 5 mL 20% ethanol, followed by 5 mL water, and 5 mL kinetic assay buffer (KAB; 50 mM HEPES, 100 mM NaCl, 20 mM $CaCl_2$ at pH 8.0) before and after the addition of our activated trypsinogen sample. The dialyzed solution containing the mature trypsin inhibitor along with residual trypsinogen was cycled through the column twice to maximize binding of the trypsin; trypsinogen does not bind efficiently to the column. The active trypsin was then eluted using 5 mL of 10mM glycine pH 3.0. The presence and purity of the mature enzyme was confirmed by SDS-PAGE.

Active-Site Titration

The concentration of active trypsin was determined by monitoring the cleavage of the fluorogenic active site titrant 4-methylumbelliferyl p-guanidinobenzoate (MUGB). Endpoint fluorescence data was measured at λ_{ex} = 360 and λ_{em} = 450 nm. The titration was performed using 500 μL of kinetic assay buffer in a fluorescence quartz cuvette. To calculate the concentration, difference in fluorescence after the addition of trypsin was used in conjunction with a standard curve constructed with 4-methylumbelliferonem (4-MU).

Kinetic Characterization of Trypsin Variants via Chymotrypsinogen (Cg) Activation

The activities of the wild-type rat anionic trypsin and its variants were measured through the activation of chymotrypsinogen to chymotrypsin. In the assay, 100 nM trypsin and 5μM chymotrypsinogen were incubated in an Eppendorf tube for 0, 2, 4, 6, 8, 10, 20 and 30 minutes at 37°C. This was followed by loading samples in a 1cm x 1cm cuvette into a Spectramax M5^e spectrophotometer (Molecular Devices, San Jose, CA). Settings were set at kinetics with an λ_{ex} = 360 and λ_{em} = 450 nm with no cut off. The scan time was set to 75 seconds, taking readings at 15 second intervals, and adding 4-methylumbelliferyl p-trimethylammoniocinnamate chloride (MUTMAC) after the 2nd reading. Like MUGB, MUTMAC contains fluorophore 4-MU, which allows one to use a 4-MU standard curve for concentration calculations. Kinetic constants were determined at 37°C in kinetic assay buffer (50 m*M* HEPES, 100 m*M* NaCl, 20 m*M* $CaCl_2$ at pH 8.0) and fitted the data with the exponential equation.

Inhibition Characterization Studies

Wild-type trypsin and its variants were characterized by measuring the hydrolytic activity towards chymotrypsinogen in the presence of BPTI. Trypsin (100 nM) was incubated at 37°C for 12 minutes with 10 nM, 20 nM, 30 nM, 50 nM, and 100 nM of BPTI. This was followed by the addition of 5 μM chymotrypsinogen and incubation for an additional 10 minutes. Similar to the kinetic characterization techniques, the samples were loaded in a 1cm x 1cm cuvette into a Spectramax

M5^e spectrophotometer (Molecular Devices, San Jose, CA). Settings were set at kinetics with an λ_{ex} = 360 and λ_{em} = 450 nm with no cut off. The scan time was set to 75 seconds with readings taken at 15-second intervals, adding MUTMAC after the 2nd reading. Kinetic constants were determined at 37°C in kinetic assay buffer (50 m*M* HEPES, 100 m*M* NaCl, 20 m*M* $CaCl_2$ pH 8.0). All Data was acquired using SoftMax Pro (Molecular Devices).

Data Analysis

To determine kinetic constants k_{cat}/K_M, graphing and data analysis was done using Igor Pro (Wave metrics) and the exponential burst equation [Eq. 1]. Value K corresponds to k_{cat}/K_M and F corresponds to total amounts of activated protein.

$$F = F_\infty \bullet \left(1 - e^{(-Kt)}\right) \qquad \text{[Eq. 1]}$$

The amount of active chymotrypsin was determined by active site titration with MUTMAC as described above after incubation of trypsin with BPTI and chymotrypsinogen. The fraction of active trypsin was determined using Eq. 2 where V_i corresponds to the activation in the presence of BPTI while V_o corresponds to the activation without any inhibitor present, with a ten-minute activation period.

$$\frac{V_i}{V_0} = \textit{Fraction of Chymotrypsinogen activated} \qquad \text{[Eq. 2]}$$

Results and Discussion

YPDS-Zeocin plates/ Bacterial growth

Due to previous lab experiments conducted, plenty of *Pichia pastoris* cells transformed with rat-anionic wild type trypsin cells were available. Small amounts of cells were spread into YPDS-Zeocin plates to grow. After incubation at 30°C for 2-4 several days, several plates seemed to have bacterial colonies from some type of contamination. Not knowing whether the contamination occurred during plating or the production of the agar, all the plates were discarded. After the second attempt all the agar plates had healthy colonies. Colonies were chosen from three agar plates and used to inoculate 4 ml YPDS-Zeocin cultures and incubated overnight at 30°C, the cultures with the greatest cell density were chosen for large scale expression. Unused plates where be wrapped in parafilm for storage at 4°C.

Large Scale Expression

Healthy cell colonies were then transferred into 1 liter of BMGY and incubated overnight at 30°C to obtain biomass. For 3-5 additional days, methanol was added to induce expression of wild-type trypsinogen during incubation. During the third day, it was determined that one of the two liters of BMGY was contaminated and was discarded. The previous day while adding methanol the tip fell into the Erlenmeyer flask which was likely the cause of the contamination. Yeast and not bacterial cells are used to express these enzymes due to their non-reducing environment that allows the proper formation of disulfide bridges that

stabilize the enzymes structure. Trypsin contains six disulfide bonds in its structure which are important for the enzyme to properly function. The cytosol of bacteria is a reducing environment that causes denaturation of these disulfide bonds. SDS-PAGE gels confirming trypsin's presence were stored in cellulose but, unfortunately the gels were not treated properly prior to storage causing them to crack.

Fast Protein Liquid Chromatography

After trypsin was expressed it was purified via fast protein liquid chromatography. We successfully purified wild-type trypsin using a phenyl sepharose hydrophobic interaction chromatography column (Pharmacia) and a negative linear salt concentration gradient (3M-0M NaCl) . A total of eighty-four 5 mL fractions were collected at a rate of 5 mL/min. To determine which fractions contained the desired protein, we collected 50 μL aliquots from every fifth fraction and ran a 12% SDS-PAGE gel to confirm which range of fractions contained the protein. The fractions that contained protein were pooled and then dialyzed into 10 mM glycine, pH 3.0. This allows the excess amount of salt surrounding the enzyme to be removed preventing the autoactivation of the enzyme. Approximately 0.3 grams of protein was recovered during the first round of FPLC purification.

Zymogen activation

Trypsin is expressed as an inactive zymogen and must to be converted its active forms through limited proteolytic cleavage. In vivo, enteropeptidase is responsible for activation of trypsinogen. However, trypsinogen has the ability to autoactivate. To promote enzyme maturation, trypsinogen was dialyzed into activation buffer (20mM HEPES, 50mM NaCl, and 2mM $CaCl_2$ with a pH 8.0) at room temperature. This promotes auto activation that allows trypsin to activate trypsinogen. Every ten to twenty minutes, a 4 μL aliquot was collected and tested for activity by measuring absorbance at hydrolysis of Z-GPR-pNA at 410 nm. Once the assay either showed no increase in activity or began to decrease, activation was considered complete. Auto-activation of trypsinogen was unsuccessful the first couple of runs. Even though there was an increase in absorbance after hours of activation, the absorbance change was so miniscule that it was believed to be background noise. Previous trypsinogen activation was noted to happen rather quickly, usually within an hour. That was not the case for our studies. Therefore, in an attempt for a more efficient activation, concentration of calcium chloride was doubled from 2mM to 4mM. Previous studies have shown that trypsinogen autoactivation is accelerated in the presence of Ca^{2+} ions. Increasing the concentration of calcium chloride was successful with trypsinogen fully activating in approximately 50 minutes. As seen in *Figure 6*, Lane 3 displays trypsinogen at time 0' with only one band. At 30 minutes several bands begin to appear which is believed to be trypsinogen, trypsin, and cleaved fragments. At 50 minutes no more

trypsinogen is present leading assumption that all trypsinogen was activate, further activation will only lead to the degradation of trypsin.

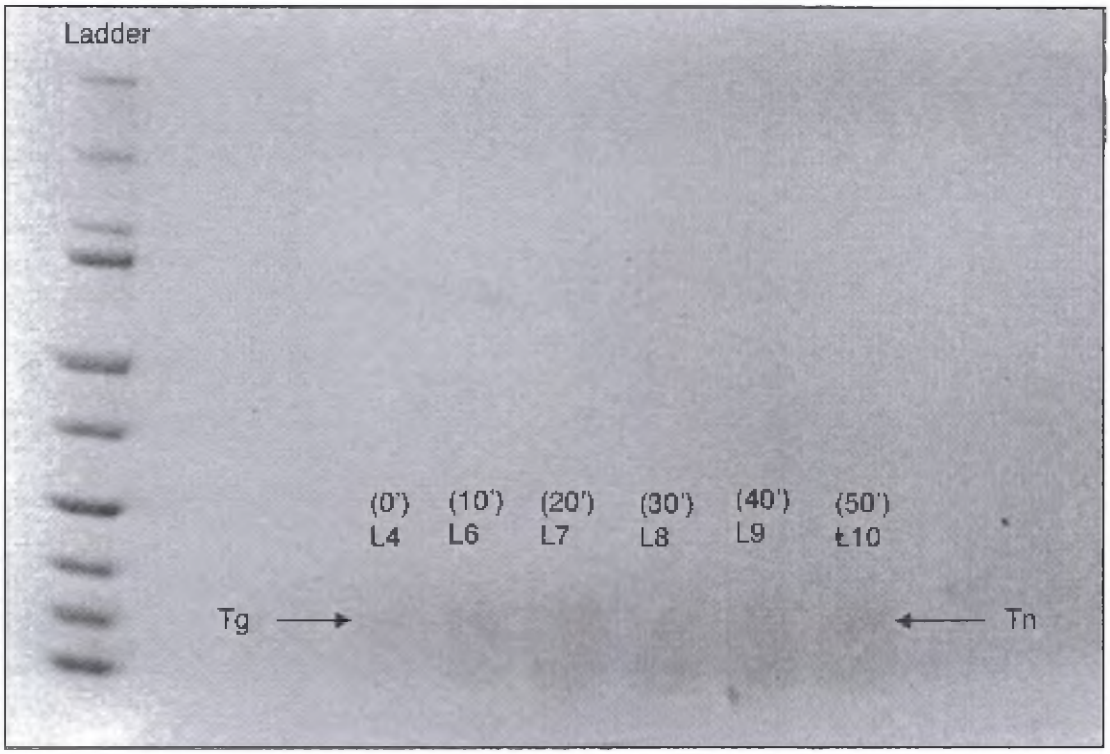


Figure 6. *Activation of wild-type trypsinogen. SDS-PAGE gel was use for the detect the activation of trypsinogen.*

Separation of Trypsin from Trypsinogen

In the activation process, the dialyzed solution contains mature trypsin along with residual trypsinogen. To separate the two, we transferred our solution to an agarose column, coupled with *p*-aminobenzamidine (pAB), a competitive trypsin inhibitor. The column was prepared as described in "Materials and Methods." The active trypsin was eluted using 5 mL of 10 mM glycine, pH 3.0

wash. This low pH glycine wash partially denatures trypsin allowing it to release from pAB column. The presence and purity of mature enzyme was confirmed by SDS-PAGE, as seen on *Figure 7.*

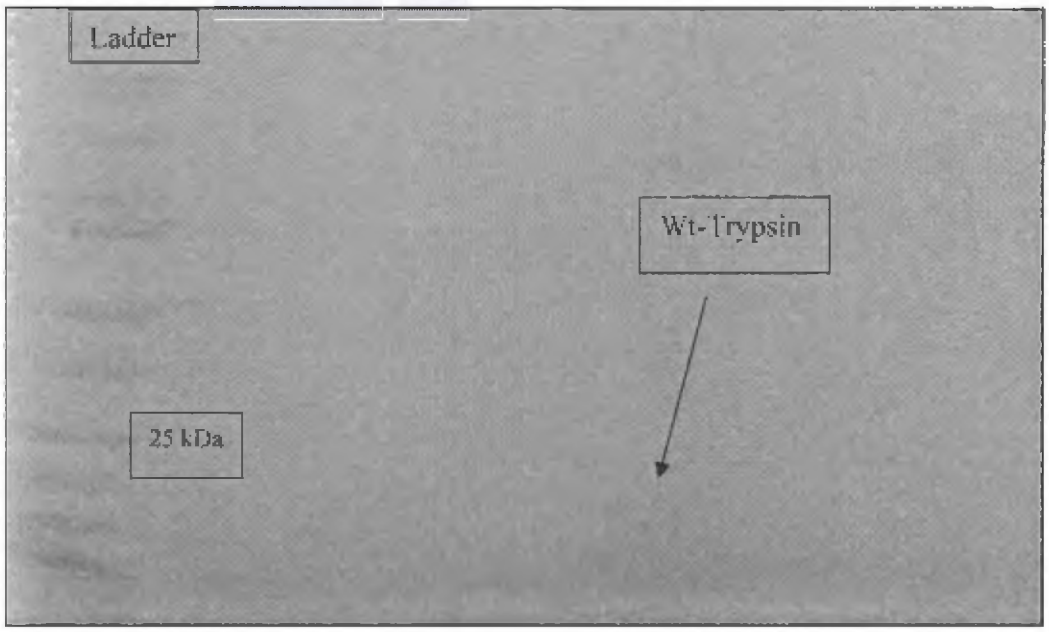


Figure 7. *4-12%Tricine SDS-PAGE confirms out activated and purified wild-type trypsin. Trypsin molecular weight ~ 23.3 kDa.*

Active-Site Titration

To accurately determine the concentration of trypsin, we monitored the cleavage of MUGB at λ_{ex} = 360nm and λ_{em} = 450 nm with no cut off. This titration was performed in kinetic assay buffer (50 m*M* HEPES, 100 m*M* NaCl, 20 m*M* $CaCl_2$ at pH 8.0) and quantified using a standard curve constructed from 4-methylumbelliferyl (4-MU). The standard curve was generated by adding 1 μM (4-MU) to 500 μL of KAB in increments of 2 μL and measuring the increase in

fluorescence on a Spectramax M5^e (Molecular Devices, San Jose, CA). To generate the titration curve, a continuous kinetic run was set up for 75 seconds with a fluorescence reading every 15 seconds. The first two readings measure fluorescence of 4 μL of 1mM MUGB in 486 μL KAB followed by titration of 10 μL of trypsin to initiate the release of 4-MU. The difference between the fluorescence before and after the addition of trypsin is used to calculate the amount of active trypsin. *Figure 8* illustrates a fluorescence standard curve, while *Table 4 shows the fixed concentrations of wild type and variant.*

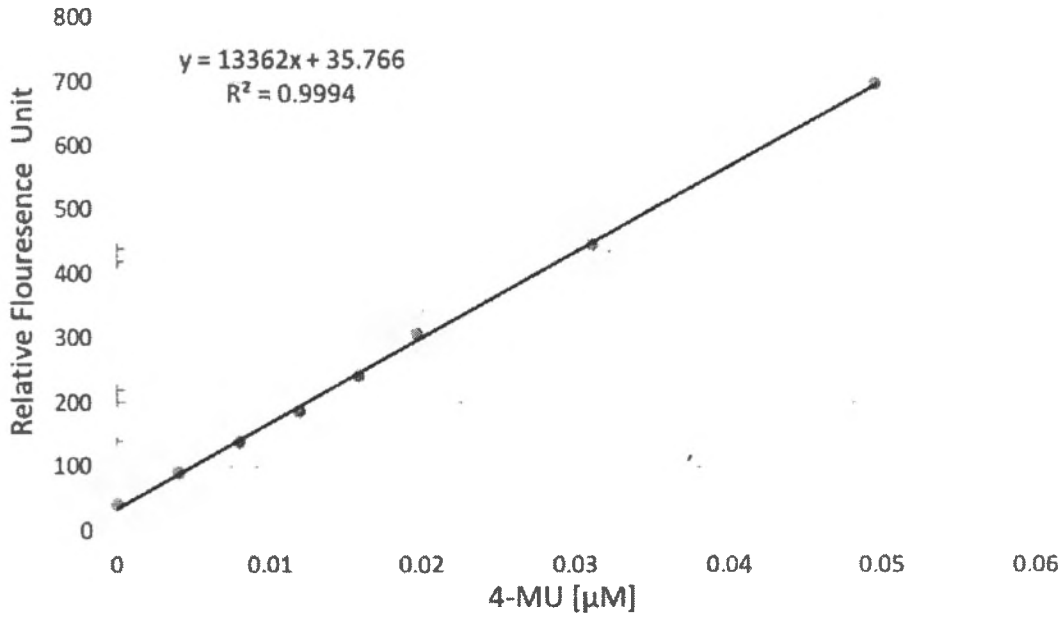


Figure 8. 4MU standard curve was generated to determine enzyme concentration. Linear regression curve displays the best fit for linear regression of RFU vs 4MU to be y = 13362x + 35.766 with a correlation coefficient r^2 = .9994

Enzyme	***Wild-Type***	***Y39A***	***Y39F***	***K60A***	***K60V***
Concentration (µM)	*12.25*	*5.31*	*7.12*	*15.64*	*11.21*

Table 4*. Trypsin variants were previously created. Active site was conducted to assure accurate concentration.*

Sample active site concentration calculation is shown below.

$$[Enzyme] = \frac{RFU(T^{30} - T^{15}) - y\ intercept}{Std\ curve\ slope}\ x\ dilution\ factor$$

$$[wild - type] = \frac{(2000.75) - (35.766)}{13362}\ x\ 83.33$$

[wild -type] = 12.25 μM

T = Time (s)

Kinetic Characterization

A problem that inevitably encountered in the development of protease therapeutic is the abundance of naturally occurring inhibitors. Due to the presence of these tight binding inhibitors, the half-life of protease therapeutics tends to be very short. Previously engineered trypsin variants displayed promising kinetic and inhibition results but were tested with a commercially available peptide substrate. It was important to fully understand the kinetic and inhibition characterization these variants have using a naturally occurring macromolecular substrate such as chymotrypsin. To determine whether Cg autoactivates, it was incubated for the same time points as the activation by trypsin. *Figure 9* confirms that there will be no false positives from autoactivated Cg.

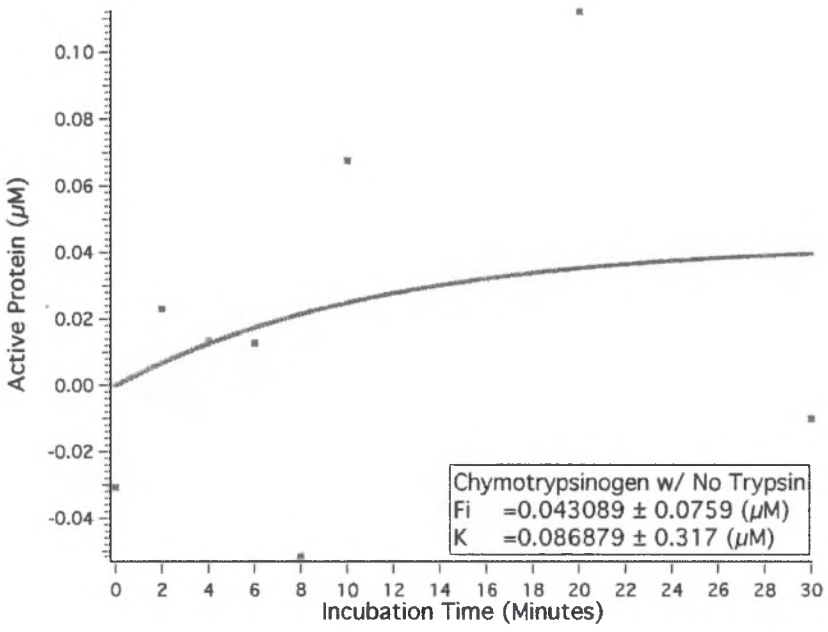


Figure 9. *After 30 minutes of incubation, there was little to no increase in activated Cg. Confirms that Cg will not autoactivate without the presence of trypsin.*

In vivo, trypsin proteolytically activates Cg making the kinetic quantification of Cg activation to be a relevant. Activation of Cg was conducted for 0, 2, 4, 6, 8,10, 20, and 30 minutes in while incubated at 37° C. Y39 and K60 single substitution variant specificity was not affected toward macromolecular substrates. A common trend was observed in all variants at the ten-minute activation mark, where the rate of activation began to slow down. This could be indicative of several things. Since proteases have the ability to degrade themselves, trypsin might be degrading itself causing a drop-in activation. However, control experiments have been done where trypsin was incubated alone, and its activity was followed with no noticeable decrease in activity over the incubation time. Another possibility is

that the turn-over rate slowed down due to the congestion within the sample. Further studies are required to fully comprehend. A drastic drop in efficiency (k_{cat}/K_M) was observed compared to previous work with tripeptide Z-GPR-pNA. Chymotrypsin is around 25 kD while Z-GPR-pNA is approximately 0.6 kD (582.62 g/mol), substantially smaller. Turnover numbers are an important parameter to evaluate the efficiency of designed proteins. A substrate with the highest value is the best substrate for the enzyme. The rate of any reaction is limited by the rate at which reactant molecules collide. [21] As seen in *Figure 10*, the Y39 mutations were more active than wild-type and K60 variants. Variant Y39A displayed the highest k_{cat}/K_M ($\mu M^{-1}Min^{-1}$) at 3.34 ± 0.06, while the value for K60V was 0.58 ± 0.01. Variant Y39A activated the most Cg with a total activation of 1.48 ± 0.047 μM (out of 5 μM Cg) while K60V had the lowest with 0.70 ± 0.04 μM. Overall the K60V mutation was affected the most. Previous studies showed that all 4 variants had similar k_{cat}/K_M towards Z-GPR-pNA. On the other hand, the valine substitution caused the greatest proteolytic rate to drop on macromolecular substrates while both Y39A and Y39F had faster k_{cat}/K_M than wild-type.

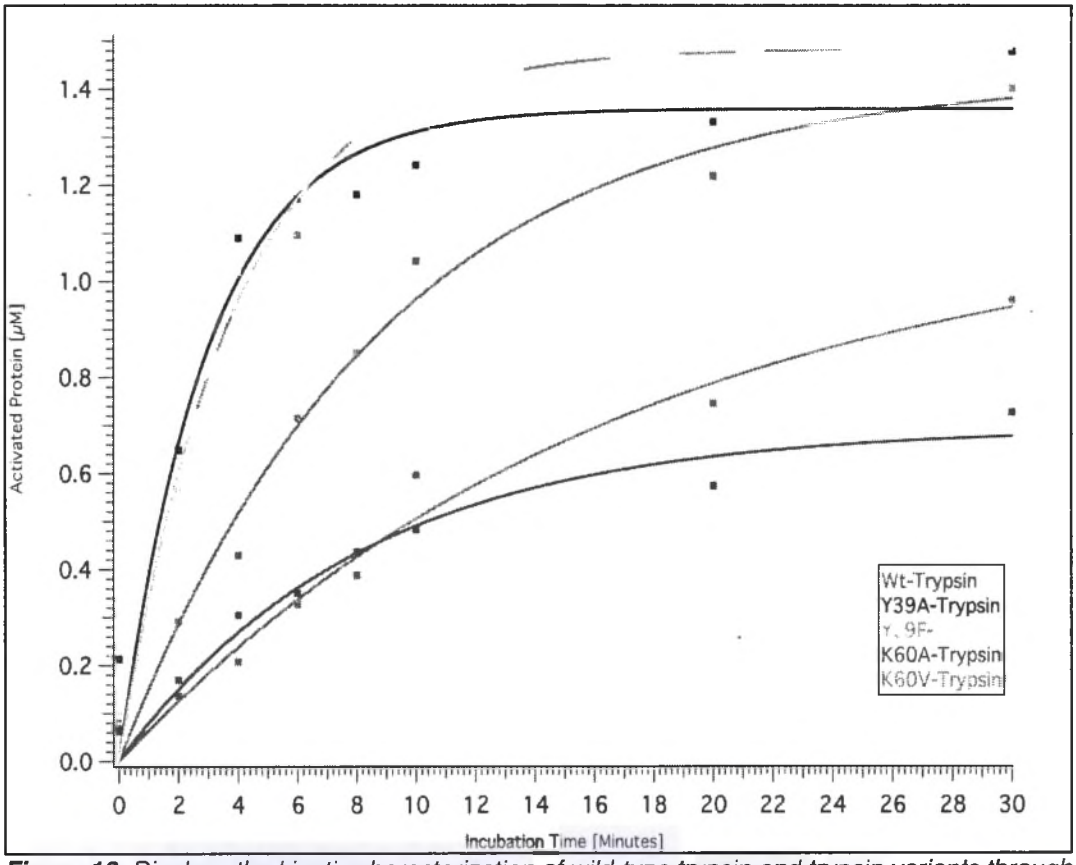


Figure 10. *Displays the kinetic characterization of wild-type trypsin and trypsin variants through the activation of chymotrypsinogen by trypsin.*

Trypsin	***Total Active Chymotrypsin (μM)***	k_{cat}/K_M ***(μM⁻¹Min⁻¹)***
Wt-Tn	1.43 ± 0.072	1.11 ± 0.012
Y39A	1.35 ± 0.069	3.34 ± 0.065
Y39F	1.48 ± 0.047	2.56 ± 0.027
K60V	1.14 ± 0.13	0.579 ± 0.012
K60A	0.70 ± 0.04	1.22 ± 0.018

Table 5. *The table above compares kinetic parameters k_{cat}/K_M and total amount of activated chymotrypsin by trypsin in a 30-minute period.*

BPTI Inhibition Studies

Conducting kinetic characterization studies allowed us to obtain insight on the catalytic efficiency toward macromolecular substrates. All four variants kept their specificity while retaining most of their catalytic efficiency. Our next focus became how these variants will act in the presence of an inhibitor. Wild-type trypsin and variants were incubated with BPTI for twelve minutes until equilibrium was reached.[17] After equilibration, chymotrypsinogen was added. It was observed with kinetic characterizations that maximum activation of Cg by trypsin occurred in 10 minutes; after the 10-minute mark activation began to decrease. For this reason, a 10-minute incubation/activation period was chosen. To obtain a ratio for active Cn, our baseline for activation was specifically chosen from the amount of activated chymotrypsin [μM] at the 10-minute mark from kinetic characterization studies. It was initially hypothesized that the removal of Tyr^{39} or Lys^{60}would cause a decrease in inhibitor sensitivity, and that was indeed the case for the Y39 variants with BPTI using the tri-peptide Z-GPR-pNA as the substrate, but not the K60 variants. After testing inhibition resistance with a macromolecular substrate, similar results were observed. At 10 nM BPTI both K60 variants had a slightly higher percentage of (77-83%) of activated trypsin with respect to the Y39 variants. A drastic activation drop can be observed on the K60 variants at the 20 nM BPTI concentration, with K60A having an activation drop of 49% while K60V dropped by 16%. At 20 nM, Y39F only had a 4% drop compared to Y39A with a 14% drop. At 30 nM, 50 nM and 100 nM the Y39 variants displayed greater activation than wild-

type and K60 variants. With Y39F displaying more than 6% difference in activation in the 30 nM, and ~3% in 100 nM BPTI. It was clear that at the higher concentrations of BPTI the K60 variants seems to have become a lot more sensitive to BPTI and displayed a greater degree of inhibition, as seen on *Figure 9*. Previous studies validated that Y39 substitution corresponded to a decrease in inhibition sensitivity with Y39F displaying the most inhibitor resistance. On the other hand, in the presence of macromolecular substrates Y36A seems to show a decrease in inhibition sensitivity, not as much as both K60 variants. Of all the variants tested in the study, it appears that the Y39 single nucleotide mutation had the most promising results from the rest. Not only did it maintain its specificity, catalytic efficiency but also displayed inhibitor resistance with both micro and macromolecular substrates. In addition, the Y39F showed promising results. It seems that by removing the Tyr^{39} hydroxyl group interactions with Lys^{60} and the P4' of BPTI was the most effective. As seen in previous and current studies a single residue mutation in positions 60 and 39 within the prime side binding region of inhibitors can greatly influence inhibitor interactions in the presence of micro and macromolecular substrates. Reviewing these substitutions in relation to the interactions occurring within the inhibitor complex may provide useful for further development of additional engineered trypsin-fold variants with a final goal of inhibitor resistance. Having the ability to design and control the desired trypsin: inhibitor complex interactions has great up side and will be essential for future developments of protease therapeutics.

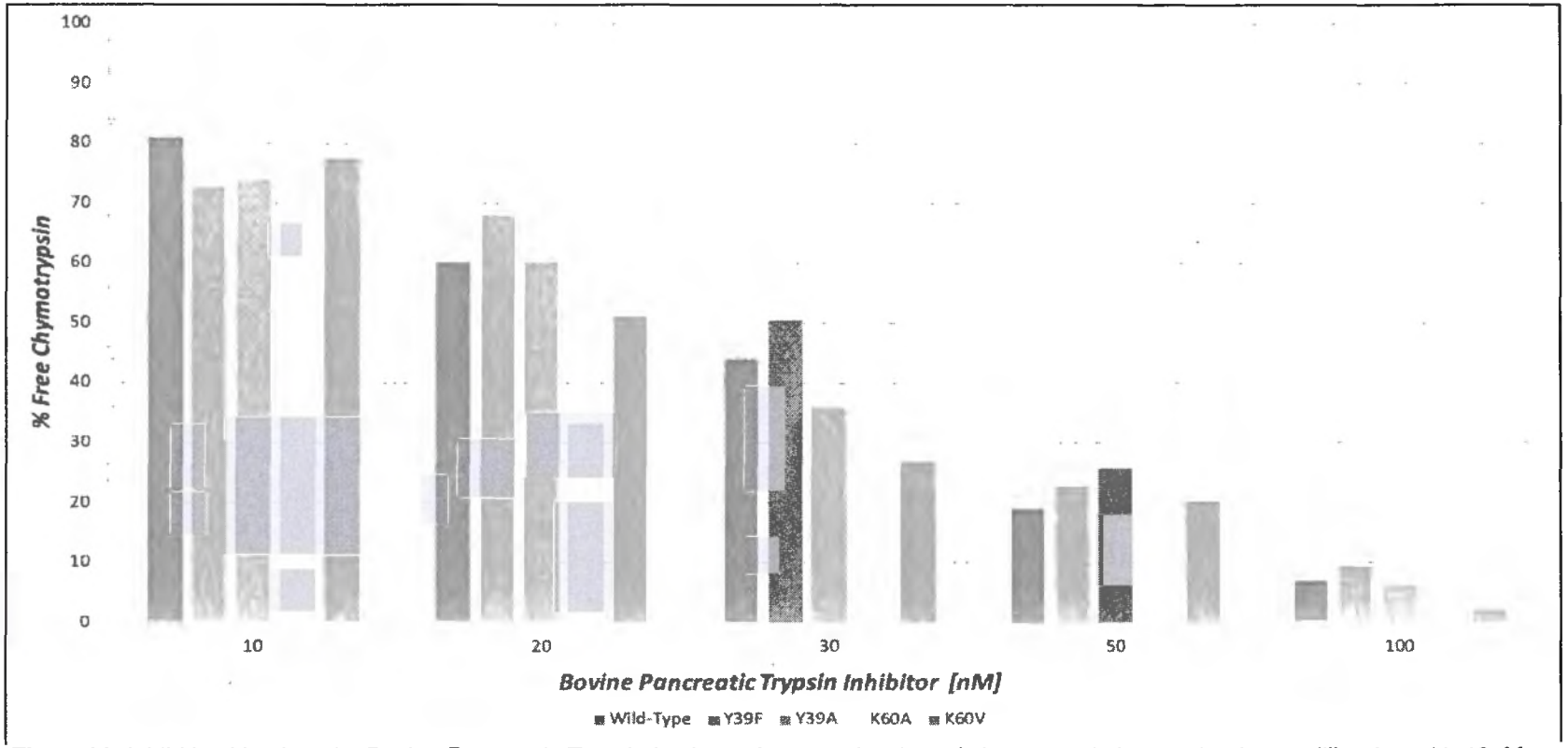


Figure11: *Inhibition kinetics with Bovine Pancreatic Trypsin Inhibitor. Amount of activated chymotrypsin by trypsin after equilibration with 10nM to 100nM BPTI.*

	Wt-Tn	Y39F	Y39A	K60A	K60V
BPTI [nM]	Active Cn (%)	Active Cn (%)	Active Cn (%)	Active Cn (%)	Active Cn (%)
10	80.86	72.60	74.01	82.18	77.39
20	60.08	68.17	60.24	33.18	51.08
30	44.13	50.51	35.90	10.84	26.86
50	18.97	22.93	25.78	8.16	20.47
100	6.79	9.32	6.15	3.51	2.275

Table 6. *The table displays the ratio of active chymotrypsin in the presence of BPTI, after inhibition equilibration and a 10-minute activation period.*

Wild-Type Trypsin S1' Pocket

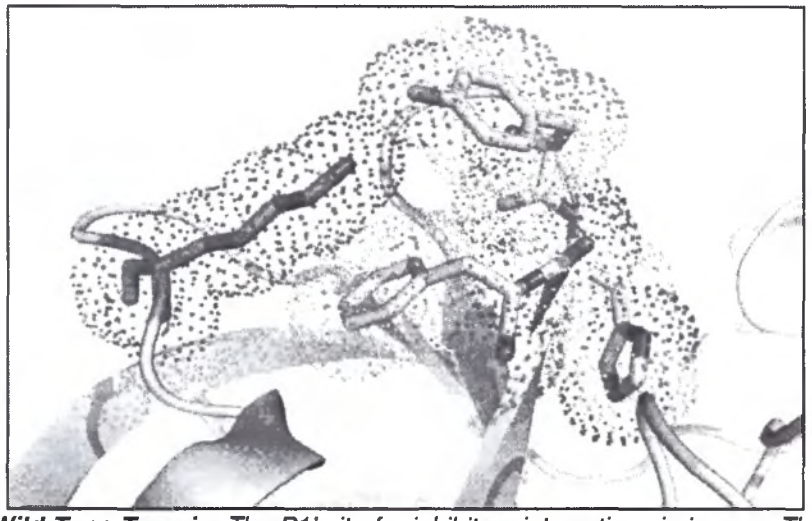
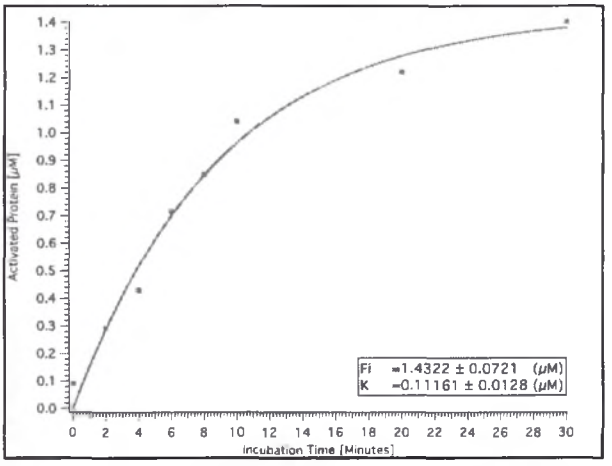


Figure 12: Wild-Type Trypsin. *The P1' site for inhibitors interactions is in grey. These residues consist of Tyr[39], His[40], Phe[41], Lys[60].*

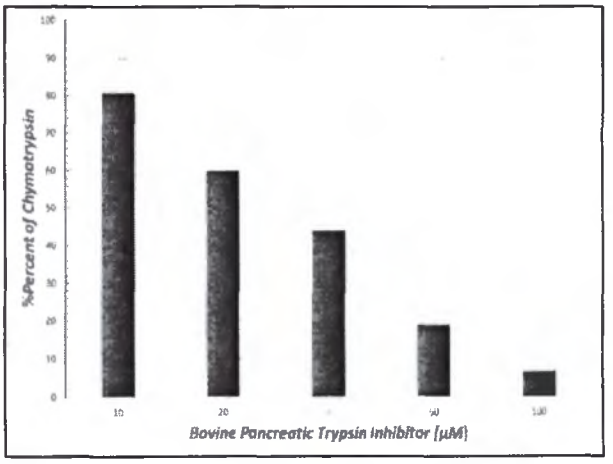
Trypsin	*Active Cn (μM)*	k_{cat}/K_M *(μM⁻¹Min⁻¹)*
WT-Tn	1.43 ± 0.072	1.11 ± 0.012

Figure 12b. *Wt-Trypsin. The plot measures the amount of activated chymotrypsin over a time period of 30 minutes.*



BPTI [nM]	*Active Cn in 10 Min. [μM]*	*Ratio of Active Cg (%*
10	1.04	80.97
20		60.08
30		44.13
50		18.96
100		6.79

Figure 12c: *WT-Trypsin. Amount of active chymotrypsin after equilibration with 10nM to 100nM BPTI.*



Y39A-Trypsin S1' Pocket

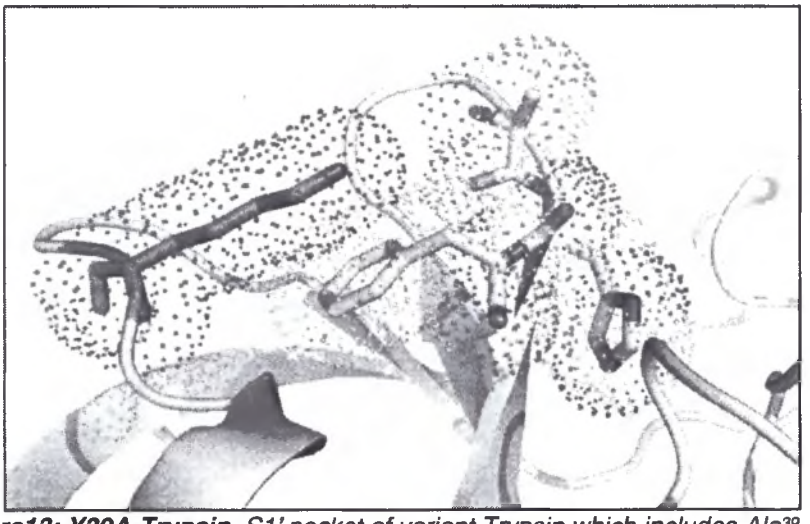
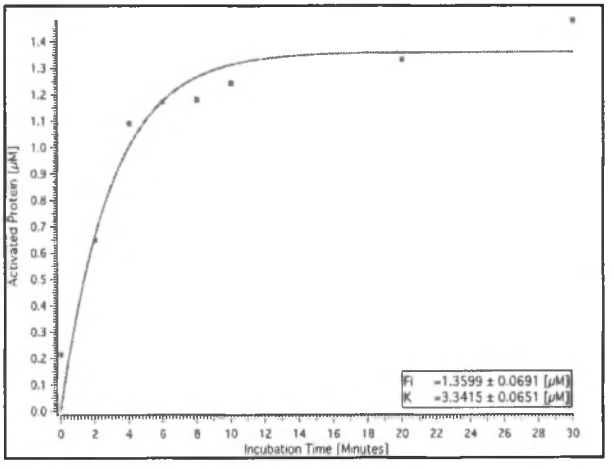


Figure13: Y39A-Trypsin. *S1' pocket of variant Trypsin which includes Ala[39], His[40], Phe[41], and Lys[60].*

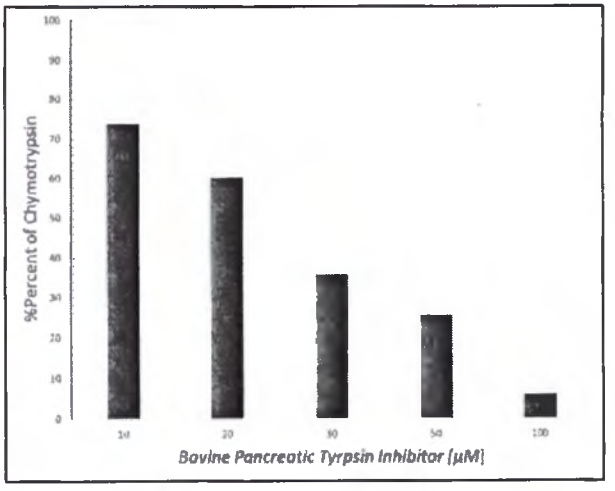
Trypsin	*Active Cn (μM)*	k_{cat}/K_M *(μM⁻¹Min⁻¹)*
Y39A	1.35 ± 0.069	3.34 ± 0.065

Figure 13b*: Y39A Trypsin. The plot measures the amount of activated chymotrypsin over a time period of 30 minutes.*



BPTI [nM]	Active Cn in 10 Min. [μM]	Ratio of Active Cg (%)
10	1.24	74.01
20		60.24
30		35.90
50		25.78
100		6.14

Figure 13c: *Y39A-Trypsin. Amount of active chymotrypsin after equilibration with 10nM to 100nM BPTI.*



Y39F-Trypsin S1' Pocket

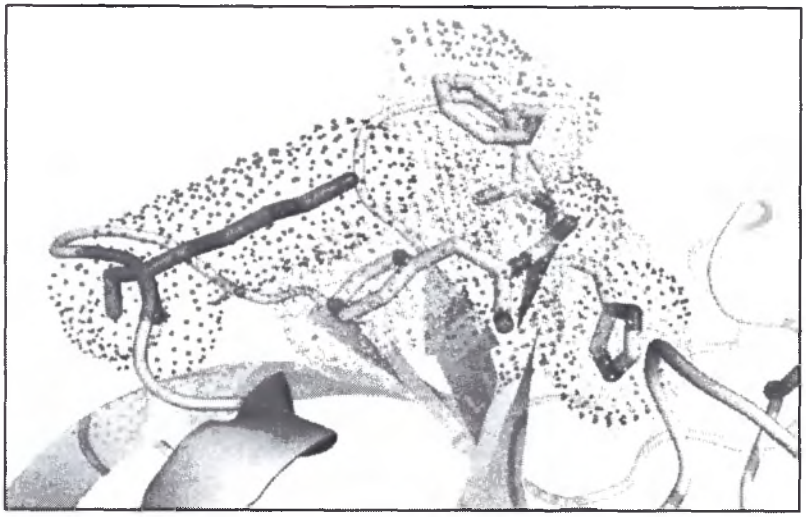
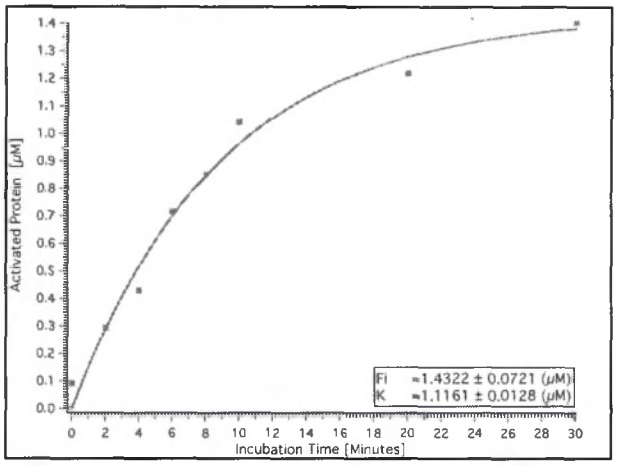


Figure14*: *Y39F-Trypsin. *S1' pocket of variant Trypsin which includes Phe[39], His[40], Phe[41], and Lys[60].*

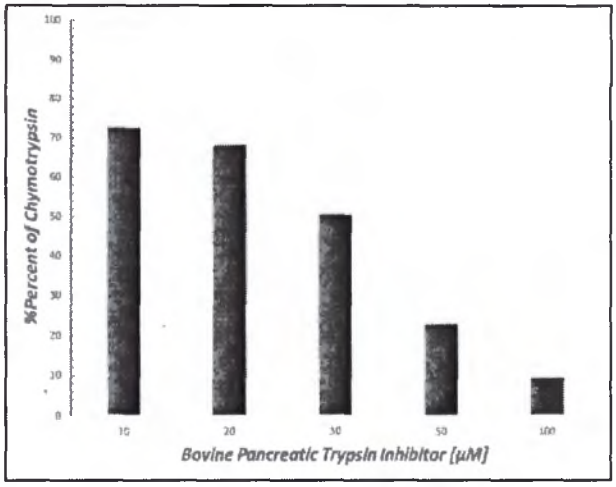
Trypsin	*Active Cn (μM)*	*k_{cat}/K_M ($\mu M^{-1}Min^{-1}$)*
Y39F	1.48± 0.047	2.56 ± 0.027

Figure 14b: *Y39F-Trypsin. The plot measures the amount of activated chymotrypsin over a time period of 30 minutes.*



BPTI [nM]	Active Cn in 10 Min. [μM]	Ratio of Active Cn (%)
10	1.34	72.59
20		68.17
30		50.51
50		22.92
100		9.32

Figure 14c: *Y39F-Trypsin. Amount of active chymotrypsin after equilibration with 10nM to 100nM BPTI.*



K60A-Trypsin S1' Pocket

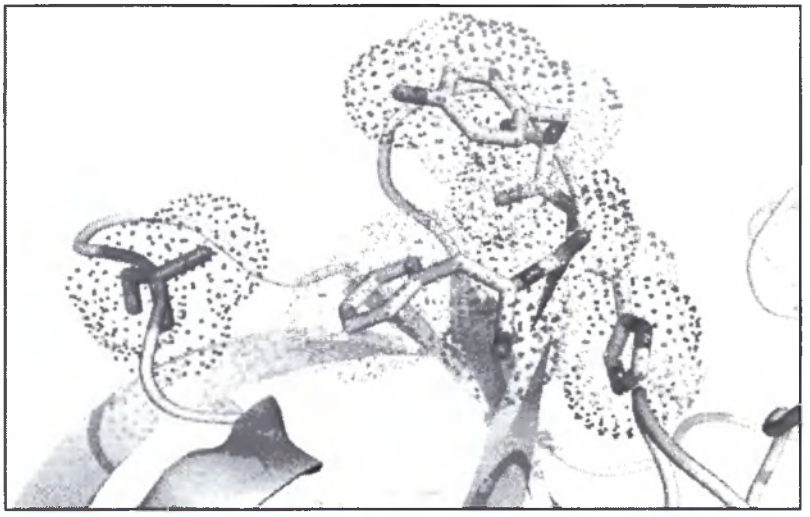
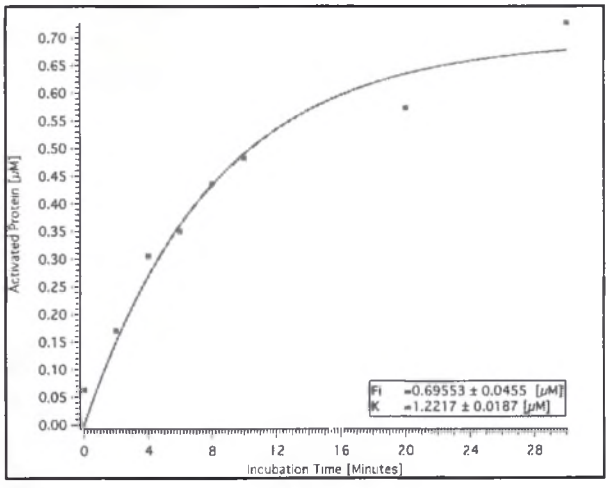


Figure 15: K60A-Trypsin. *S1' pocket of variant Trypsin which includes Tyr[39], His[40], Phe[41], and Ala[60].*

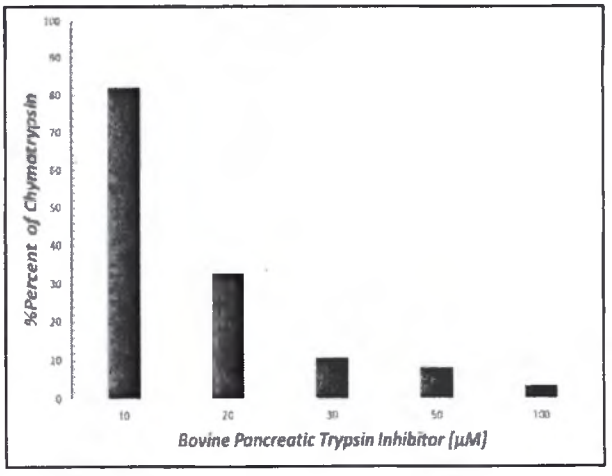
Trypsin	*Active Cn (μM)*	*Kcat/Km (μM⁻¹Min⁻¹)*
K60A	0.695 ± 0.045	1.22 ± 0.018

Figure 15b: *K60A-Trypsin: The plot measures the amount of activated chymotrypsin over a time period of 30 minutes.*



BPTI [nM]	Active Cn in 10 Min. [μM]	Ratio of Active Cn (%)
10	0.482	82.18
20		33.18
30		10.83
50		8.15
100		3.50

Figure 15c: *K60A-Trypsin. Amount of active chymotrypsin after equilibration with 10nM to 100nM BPTI .*



K60V-Trypsin S1' Pocket

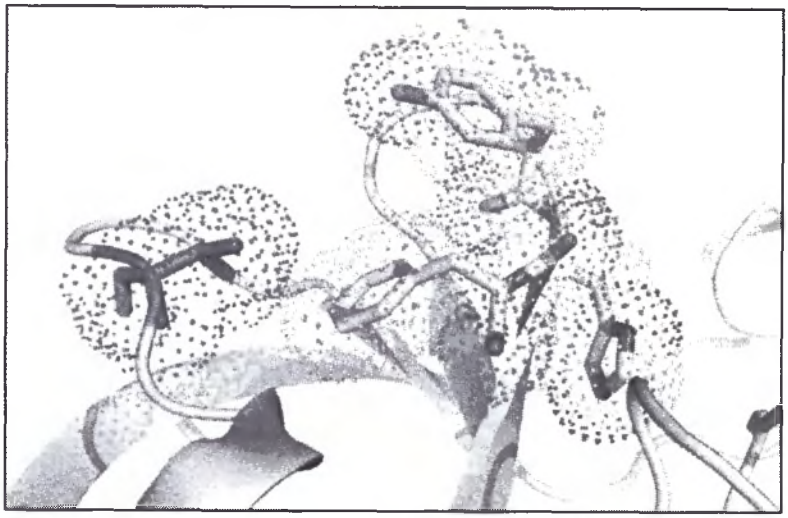
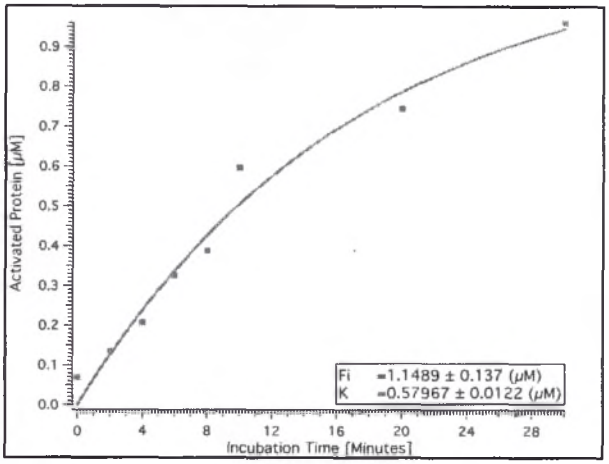


Figure 16: K60V-Trypsin: *S1' pocket of variant Trypsin which includes Tyr[39], His[40], Phe[41], and Val[60].*

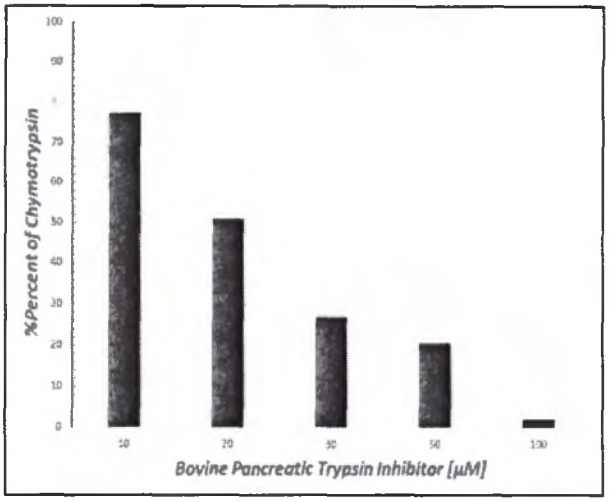
Trypsin	*Active Cn (μM)*	*Kcat/Km ($\mu M^{-1}Min^{-1}$)*
K60V	1.14 ± 0.13	.579 ± 0.012

Figure 16b: *K60V-Trypsin. The plot measures the amount of activated chymotrypsin over a time period of 30 minutes.*



BPTI [nM]	Active Cn in 10 Min. [μM]	Ratio of Active Cn (%)
10	0.59	77.4
20		51.1
30		26.86
50		20.47
100		2.26

Figure 16c: *K6VA-Trypsin. Amount of active chymotrypsin after equilibration with 10nM to 100nM BPTI.*



Future Directions

Kinetic Characterization

Using the activation of chymotrypsinogen to chymotrypsin to monitor proteolytic rate and efficiency brings many outside factors. Since chymotrypsin is a serine protease, a small degree of the enzyme concentration will be inhibited by many of the inhibitors that also inhibit trypsin. Due to this conflict, using a different substrate will such as azocasein, a macromolecular protease substrate, can provide better understanding by eliminating unwanted interactions.

Inhibition Studies

In this project, BPTI was the only inhibitor used. Although BPTI inhibits trypsin and can inhibit chymotrypsin to a certain extent, the studies displayed that chymotrypsin was not inhibited to a noticeable degree. Using a broader range of inhibitors would give more insight on inhibition kinetics using a macromolecular substrate. Studying inhibitors such as M84R ecotin with a broad selectivity or soy bean trypsin inhibitor that offers more selective may provide additional information.

Citations

1. Turk, B. Targeting proteases: successes, failures and future prospects. *Nature Reviews Drug Discovery* **5**, 785–799 (2006).
2. Di Cera, E. Serine proteases. *IUBMB Life* **61**, 510–515 (2009).
3. Mala, R., Aparna, T., Mohini, G. & Vasanti, D. *MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS.* **63(3)**, (1998).
4. Craik, C. S., Page, M. J. & Madison, E. L. Proteases as therapeutics. *Biochemical Journal* **435**, 1–16 (2011).
5. Rani, K., Rana, R. & Datt, S. REVIEW ON LATEST OVERVIEW OF PROTEASES. 8 (2012).
6. Rawlings, N. D. & Barrett, A. J. Evolutionary families of peptidases. *Biochemical Journal* **290**, 205–218 (1993).
7. Hedstrom, L. Serine Protease Mechanism and Specificity. *Chemical Reviews* **102**, 4501–4524 (2002).
8. Polgár, L. The catalytic triad of serine peptidases. *Cellular and Molecular Life Sciences* **62**, 2161–2172 (2005).
9. *PyMOL. San Carlos, CA, 700.* (Delano Scientific).
10. Baird, T. T., Wright, W. D. & Craik, C. S. Conversion of trypsin to a functional threonine protease. *Protein Sci.* **15**, 1229–1238 (2006).
11. Brandt, M. Enzyme_mech_examples.pdf. (2016).
12. Krowarsch, D., Cierpicki, T., Jelen, F. & Otlewski, J. Canonical protein inhibitors of serine proteases. *Cellular and Molecular Life Sciences (CMLS)* **60**, 2427–2444 (2003).

13. Laskowski and Kato - 1980 - Protein Inhibitors of Proteinases.pdf.

14. Maddur, A. A., Swanson, R., Izaguirre, G., Gettins, P. G. W. & Olson, S. T. Kinetic Intermediates en Route to the Final Serpin-Protease Complex: STUDIES OF COMPLEXES OF α_1-PROTEASE INHIBITOR WITH TRYPSIN. *J. Biol. Chem.* **288**, 32020–32035 (2013).

15. Major, I. T. & Constabel, C. P. Functional Analysis of the Kunitz Trypsin Inhibitor Family in Poplar Reveals Biochemical Diversity and Multiplicity in Defense against Herbivores. *Plant Physiol.* **146**, 888–903 (2008).

16. Bannister, W. Medicinal Chemistry. A Biochemical Approach (Second Edition). *Biochemical Education* **18**, 55 (1990).

17. Batt, A. R., St. Germain, C. P., Gokey, T., Guliaev, A. B. & Baird, T. Engineering trypsin for inhibitor resistance: Engineering Trypsin for Inhibitor Resistance. *Protein Science* **24**, 1463–1474 (2015).

18. Genentech: Press Releases. Available at: https://www.gene.com/media/press-releases/4271/1987-11-13/licensing-of-activase-marks-new-era-in-t. (Accessed: 5th April 2018)

19. Frequently asked questions about Activase (Alteplase), 2010, Genentech.pdf.

20. Khan, A. R. & James, M. N. G. Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Science* **7**, 815–836 (2008).

21. *Peptide, protein and enzyme design*. (Elsevier, Academic Press, 2016).

22. Bannister, W. Medicinal Chemistry. A Biochemical Approach (Second Edition). *Biochemical Education* **18**, 55 (1990).