The Dissertation Committee for Haixin Li Certifies that this is the approved version of the following dissertation:

Engineering and Investigation of Protease Fine Specificity

Committee:

George Georgiou, Supervisor

Brent L. Iverson

David W. Hoffman

Jon D. Robertus

Pengyu Ren

Engineering and Investigation of Protease Fine Specificity

by

Haixin Li, B.Med, M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2010

Dedication

To my family

Acknowledgements

First, I would like to gratefully and sincerely thank my research advisor, Dr. George Georgiou, for his continuous encouragement, tremendous support and guidance throughout my six year's study. I would like to thank Dr. Brent Iverson for his valuable suggestions and advice on my research project and I was much motivated by his strong passion for the scientific research. I appreciate all the thoughtful comments and suggestions from my committee members, Dr. David Hoffman, Dr. Jon Robertus and Dr. Pengyu Ren.

I would like to thank for Dr. Bumyeol Huang for his assistance to get me acquainted with phage display and his kindly help in the human kallikrein project. I would like to thank Dr. Navin Varadarajan for training me the substrate conjugation experiments and FPLC, HPLC analysis and for his valuable pioneer work on the OmpT project. I would like to thank Dr. Stony Lo for his contribution to the Mass spectrum work. I would like to thank Mark Gebhard and Mark Pogson for their help with the FACS sorting. I would like to thank Dr. Tae Hyeon Yoo, Dr. Mridula Rani, Dr. Silvia Arredondo, Dr. Marsha Demers-Lewis, Dr. Sang Taek Jung, Dr. Everett Stone, Dr. Li Zhou and all other BIGG lab members for sharing their experiences, helpful advice and their friendship.

Finally, I would like to thank my whole family: my parents, my husband and my dear daughter Sophie. Without their previous love and endless support, I could not have finished my PhD study.

Engineering and Investigation of Protease Fine Specificity

Publication No._____

Haixin Li, PhD

The University of Texas at Austin, 2010

Supervisor: George Georgiou

The Escherichia coli (*E. coli*) outer membrane protease OmpT is an endopeptidase of the omptin family in gram negative bacteria. OmpT cleave preferentially between two consecutive basic residues, especially Arg-Arg, and it has been classified as an aspartyl protease based on its crystal structure although biochemical confirmation of a catalytic aspartyl residue is lacking (Vandeputte-Rutten, *et al*., 2001). Our lab has successfully engineered the P1 and P1' specificity and selectivity of OmpT by employing novel strategies for the isolation of enzyme variants that cleave desired substrates from large combinatorial libraries screened by flow cytometry. However, the engineering of proteases with altered specificity beyond the P1 and P1' residues of the substrate have not been demonstrated. By applying high throughput screening of large libraries of OmpT constructed by structure-guided saturation mutagenesis of the S2 subsite (which recognizes the P2 residue), as well as random mutagenesis by error prone

PCR and DNA shuffling, we engineered an OmpT variant exhibiting about 56 fold change in the selectivity for the P2 position in peptide substrates. Specifically, this enzyme preferred an acidic residue (Glu) over Tyr which is preferred by the wild type OmpT. Molecular modeling was then employed to provide insights on how mutations in OmpT mediated this change in P2 specificity.

A long term goal of protease engineering is to generate highly specific enzyme variants that can be used for the irreversible inactivation of disease targets. The anaphylatoxin C3a is a key mediator in inflammation and has been implicated with multiple inflammatory diseases. Since the site of anaphylatoxin C3a recognized by cellular receptors lie in its C-terminus, a protease cleaving the C-terminus of C3a could be therapeutically relevant. Using high throughput screening and directed evolution we successful isolated C3a cleaving enzyme variants and have characterized them biochemically.

Finally as part of this dissertation we have employed high throughput screening methods to dissect the substrate specificity of members of the kallikrein family of mammalian proteases which are implicated in a number of physiological and disease functions. The human tissue kallikrein (KLK) family contains 15 secreted serine proteases that are expressed in a wide range of tissues and have been implicated in different physiological functions and disease states. Of these, KLK1 has been shown to be involved in the regulation of multiple physiological processes such as blood pressure, smooth muscle contraction and vascular cell growth. KLK6 is over-expressed in breast and ovarian cancer tissues and has been shown to cleave peptides derived from human

myelin protein and the Aβamyloid peptide *in vitro*. Here we analyzed the substrate specificity of KLK1 and KLK6 by substrate phage-display using a random octapeptide library. Consistent with earlier biochemical data, KLK1 was shown to exhibit both trypsin-and chymotrypsin-like selectivities with Tyr/Arg preferred at the P1 site, Ser/Arg strongly preferred at P1' and Phe/Leu at P2. KLK6 displayed trypsin-like activity, with the P1 position occupied only by Arg and a strong preference for Ser in P1'. Docking simulations of consensus peptide substrates was used to infer possible identities of the enzyme residues that are responsible for substrate binding. Bioinformatic analysis suggested several putative KLK6 protein substrates such as ionotropic glutamate receptor (GluR) and synphilin.

Table of Contents

List of Tables

List of Figures

Chapter 1 Protease Engineering

1.1 Overview of Protease engineering

Proteases are enzymes that cleave the peptide bond between amino acids. These enzymes exist in all organisms. In human proteases are predicted to account for nearly 2% of the genome and represent 5-10% of drug targets (Overall, *et al.*, 2007). Proteases play an important role in protein activation, synthesis and degradation, and are involved in a wide range of physiological processes from food digestion to cell cycle progression, apoptosis, coagulation, wound healing, tissue remodeling, the immune reaction, *etc* (Turk, 2006). Figure1.1 shows the nomenclature of the amino acid positions in substrate peptides cleaved by proteases. Based on the position of the cleavage, proteases are classified into exopepdidases (e.g. aminopeptidases, carboxypeptidases) and endopeptidases (e.g. trypsin, chymotrypsin). Based on their catalytic mechanisms and specifically the active site nucleophile, proteases are divided into serine, threonine, cysteine, aspartic, glutamic and metallopreases. Note that glutamic proteases have not been found in mammals yet. Proteases can be relatively nonspecific, or very specific, for particular substrates. The nonspecific proteases are commonly involved in digestion or protein degradation, while the specific proteases are typically involved in the turning on cellular signals or activating enzyme cascades such as the caspases and blood clotting factors (Walsh, 2001; Turk, 2006).

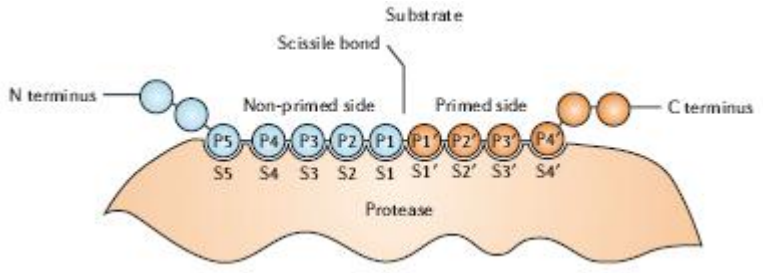


Figure1.1 Schematic representation of protease and substrate interaction; nomenclature of binding sites (Adapted from Turk, 2006).

Biologists and biochemists have studied proteases extensively in order to understand their structures, functions, activities and substrate specificities. However, it is often difficult to fully understand the relationship between a protease's structure and its specificity or activity. Because of the lack of detailed mechanistic models, reprogramming the specificity of proteases has been a challenge to protein engineers. Novel protease variants generated in the laboratory are often found to either display low catalytic rates or broad specificity relative to the corresponding wild type enzymes. In the early days of protease engineering, rational design based on crystal structures was commonly used to identify residues to be mutated. This method has obvious limitations due to its low throughput and the difficulty encountered in trying to predict the consequences of particular mutations.

More recently, directed evolution, has been extensively used for protease engineering. In direct evolution, the Darwinian principles of mutation and selection are mimicked *in vitro*. While applying selective pressure, large numbers of variants are screened for the desired specificity and overall level of activity. Directed evolution requires a functional display system compatible with the normally toxic protease.

Combining rational design with directed evolution has proven to be a particularly powerful approach to engineering enzymes (Reviewed by Johannes and Zhao, 2006; Kaur and Sharma, 2006). Specificities of some proteases have been successfully engineered during the past two decades such as trypsin ((Hedstrom, *et al*., 1992), subtilisin (Ballinger, *et al*., 1996), aqualysin I (Tanaka, *et al.*, 1998), HIV protease (O'Loughlin, *et al*., 2006), rat neurolysin (Kadonosono, *et al*., 2008), HAV 3CP (Sellamuthu, *et al*., 2008), and the *E. coli* OmpT (Varadarajan, *et al*., 2005, 2008) (reviewed by Pogson, *et al*., 2009). However most of these studies did not involve directed evolution. (Some of these studies will be discussed in detail in the following.)

1.2 Engineering the Trypsin-Chymotrypsin-Elastase model system

The pancreatic serine proteases trypsin, chymotrypin and elastase have been extensively studied as model systems for protease engineering. Trypsin, chymotrypsin and elastase have a high degree of structural similarity but distinct substrate specificities. At the P1 position, Trypsin prefers Arg/Lys; chymotrypsin favors hydrophobic residues Phe/Tyr/Trp; and elastase prefers Ala/Val. Residue 189 residing at the bottom of the S1 binding pocket was proposed to determine the specificity of P1 (reviewed by Hedstrom,

2002). However, redesigning the specificities of these enzymes turned out to be quite challenging. The single amino acid substitution of Asp189 with Ser (D189S) failed to generate chymotrypsin specificity (Hedstrom, *et al*., 1992; 1996). The chymotrypsin variant Ser189Asp (S189D) did not have trypsin-like activity either (Venekei, *et al*., 1996). It was proposed that in addition to the S1 binding pocket, distal interactions were important in determining the substrate specificity and trypsin was successfully engineered to have chymotrypsin activity only after replacing the two loops L1 and L2 in addition to the D189S mutation (Hedstrom, 1992). Moreover, the S1' and S2' sites of trypsin have been reprogrammed. The trypsin variant K60E with a single amino acid substitution showed about 70 fold increase in preference for Arg at P1' than wild type trypsin (Kurth, *et al*., 1998). A trypsin variant with specificity for His at P2' was created by introducing a metal binding site at the S2' subsite (Willett, *et al*., 1996). Elastase has not yet been successfully converted to a trypsin-like protease despite extensive efforts (Venekei, *et al*., 1996; Huang and Hedstrom, 1998).

1.3 Engineering subtilisin

Subtilisin has been a model enzyme for protease engineering since the 1980s. It is one of the best characterized of all the enzymes due to its multiple industrial applications, especially its use in laundry detergent. The large amount of structural and functional information about subtilisin made it possible to engineer its stability and specificity via rational design. Plenty of variants have been generated with mutations in over 50% of

the 275 amino acids (reviewed by Bryan, 2000). Subtilisin has been engineered for increased stability, investigations of catalytic mechanism, introduction of novel activity and altered specificity. Successful examples of modulating specificity of the enzyme include: conversion of subtilisin to furilisin for cleaving tri-basic substrates (Ballinger, *et al*., 1996); altering the P1 and P1' specificity of subtilisin by site directed mutagenesis and chemical modification (DeSantis, *et al*., 1998); altering the P2 specificity of aqualysin I (subtilisin related protease) with site directed mutagenesis (Tanaka, *et al*., 1998); a remodeling of subtilisin specificity to cleave phosphotyrosine substrates (Knight, *et al*., 2007); a subtilisin-like yeast Golgi protease Kex2 engineered to exhibit novel P2 specificity by genetic screening (Han, *et al*., 2005); and a subtilisin BPN' variant from *Bacillus amyloliquefaciens* engineered to strongly prefer substrate with Phe or Tyr at P4 position (Ruan, *et al*., 2008).

1.4 Engineering specificity of HIV PR

The human immunodeficiency virus type I protease (HIV PR) is a processing enzyme that cleave the HIV-1 *Gag-Pol* polyprotein and thus activate the proteins required for viral infectivity (Kohl, *et al.*, 1988). The HIV PR is considered as cytotoxic non-specific protease due to its cleavage of essential proteins in E. coli, yeast and mammalian cells (Baum, *et al.*, 1990; Blanco, *et al.*, 2003). The HIV PR has been utilized as a model enzyme to engineer altered specificity (O'Loughlin, *et al*., 2006). In this study, HIV PR library was generated by error-prone PCR and transformed into *E.coli*.

Due to the toxicity of wild type HIV PR, screening for viable cells after induction eliminated the non-specific variants. Wild type HIV PR substrate sequence (decapeptide) was inserted into the surface loop of *E.coli*. β galastosidase (β-gal) and its activity is monitored by the decrease of β-gal activity. The isolated viable variants were further screened for a disrupted β-gal activity using a non-wild type preferred substrate. The variant P9S/I150L was isolated and exhibited novel specificity for substrate NRPDYLLFAE. Moreover, the variant P9S/I150L does not have the cytotoxic effect of the non-specific wild type. In this way, HIV PR specificity was successfully engineered by directed evolution (Khersonsky, *et al*., 2006).

1.5 Engineering substrate specificity of HAV 3CP

The human hepatitis A virus 3C protease has been utilized as a model enzyme to engineer substrate specificity using a genetic assay for site-specific proteolysis (GASP) in yeast (Sellamuthu, *et al*., 2008). The GASP system is shown in Figure 1.2

The triple fusion protein included integral membrane protein, peptide substrate sequence for protease and the transcription factor (TF). The transcription factor LexA-b42 is anchored to plasma membrane due to a linker that connected to a trans-membrane domain. Upon the specific cleavage by protease at the linker region, the LexA-b42 was released from the membrane, trans-located into the cell nucleus and turned on the transcription of the reporter genes. Using this scheme, a four-position saturation library of HAV 3CP was screened for altered P2 specificity (Gln over wild type Thr). An HAV

3CP variant was successfully isolated demonstrating a 160-fold switch of specificity (reviewed by Pogson *et al.* , 2009).

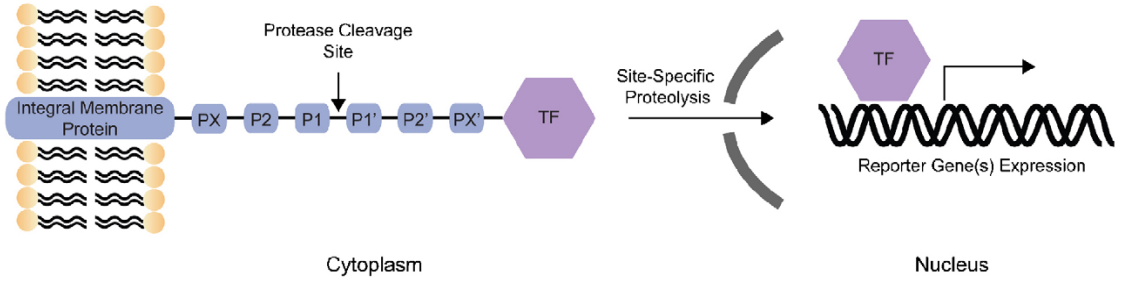


Figure1.2 Screening principle of GASP. (Adapted from Pogson, *et. al*., 2009) The triple fusion protein contains integral membrane protein, peptide substrate sequence for protease and the transcription factor (TF).

1.6 Engineering *E.coli* OmpT

1.6.1 Overview of *E.coli* OmpT

E.coli OmpT (EC 3.4.21.87) belongs to the omptin family which consists of a group of highly homologous outer membrane proteases. This protease family includes the *E. coli* OmpP, the *Yersinia pepsis* plasminogen activator Pla, the *Salmonella enterica* PgtE, the *Shigella flexneri* SopA and the *Erwinia pyrifoliae* PlaA (Kukkonen, *et al*., 2004). OmpT was first identified as an outer membrane protein, and then was named OmpT due to its low expression level at low temperature. (Rupprecht, *et al*., 1983) Mature OmpT has a molecular weight of 33.5 KD. It contains 297 amino acids and folds into a 10-stranded vase-shaped anti-parallel β-barrel (Vandeputte-Rutten, *et al*., 2001; Figure1.3). OmpT was first proposed to be a serine protease with Ser99 and His212

involved in the catalytic mechanism (Rawlings and Barrett, 1994). However, the subsequent crystal structure showed this was not the case and revealed that OmpT has a novel proteolytic mechanism likely involving a His212-Asp210 dyad and an Asp83-Asp85 pair to activate a putative nucleophilic water molecule. Therefore OmpT is tentatively classified as an aspartic protease at this point even though the details of the catalytic mechanism are not known at this time. Interestingly, the active site sequence and therefore presumably the active site structure is fully conserved in the Omptin family (Vandeputte-Rutten, *et al*., 2001). Nonetheless the assignment of OmpT as an aspartyl protease has been controversial and it is not supported by the pH profile of the catalytic activity.

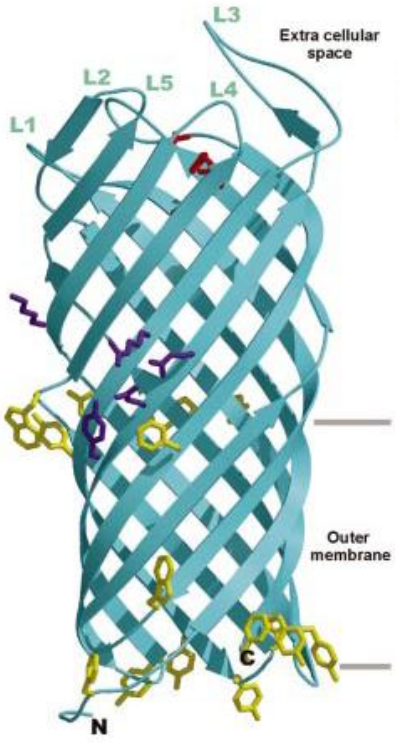


Figure1.3 Overall structure of OmpT (adapted from Vandeputte-Rutten, *et al*., 2001). The extra cellular space is at the top of the figure. There are five extracellular loop labeled as L1 to L5. The horizontal line indicates the boundary of outer membrane bilayer. The putative catalytic residues are labeled in red; the proposed LPS binding sites are in purple; and the aromatic residues at the boundary of hydrophobic and hydrophilic area on the surface are depicted in yellow.

The full physiological function of OmpT may not be understood yet either. So far, OmpT has been suggested to play a role in urinary tract infections by uropathogenic *E. coli* (Webb and Lundrigan, 1996). This was supported by the finding that OmpT hydrolyzes the anti-microbial peptide protamine that is secreted by the epithelial cells in the urinary tract (Stumpe, *et al*., 1997). OmpT is also considered a stress protein in *E. coli* with increased expression in response to heat shock or induction of recombinant protein overexpression (Gill, *et al*., 2000).

Although not a biological role, OmpT has been shown to cleave recombinant proteins during purification such as T7 RNA polymerase, human interferon gamma and cyclin A (Grodberg, *et al*., 1988; Sugimura, *et al*., 1988; Yam, *et al*., 2001); cholesterol esterase/lipase fusions and the cholera toxin B subunit (Hanke, *et al*., 1992; Klauser, *et al*., 1992). OmpT's putative role in E. coli pathogenesis and its potential to be a useful processing enzyme for recombinant fusion protein in biotechnology justifies further research in its enzymatic properties, mechanism of action and specificity.

1.6.2 The specificity of OmpT

The specificity of OmpT has been investigated in studies that span the last 20 years. OmpT prefers to cleave the peptide bond between dibasic residues (Sugimura and Nishibara, 1988) and requires binding to lipopolysaccharide (LPS) to exhibit its catalytic activity (Krammer *et al*., 2000). A synthetic peptide library screen revealed that OmpT has a strict preference for Arg (or sometimes Lys) at the P1 position but a broader

tolerance of Lys, Ile, His and Arg at P1' (Dekker, *et al*., 2001). OmpT exhibits minimal specificity at P2 as evidenced by the acceptance of the Abz fluorophore at this position. At P2', OmpT has been shown to prefer Ile, Val and Ala to a certain degree. However, introducing acidic residues such as Glu or Asp at either P2 or P2' almost abolishes all OmpT activity, suggesting that electrostatic interactions at sites adjacent to P1 and P1' might play an important role in substrate recognition (Dekker, *et al*., 2001).

A phage display study refined these conclusions and further showed that OmpT prefers basic residues at P1 and P1' and Val or Ala are strongly preferred at P2'. At P2, OmpT was found to prefer Ala, Gly, Tyr or Phe, while disfavoring acidic residues (McCarter, *et al*., 2004). OmpT shows a lower K_m rather than higher k_{cat} for substrates containing basic residues at P4 and P6, and this specificity was utilized to design an optimal linker sequence to generate recombinant Atrial Natriuretic Peptide (ANP) that used OmpT as a processing enzyme (Okuno, *et al*., 2002).

The crystal structure study of OmpT shows a large negatively charged groove on the extracellular surface, explaining its substrate specificity for consecutive positively charged residues and thus the cleavage of cationic peptides (Vandeputte-Rutten, *et al*., 2001). The negatively charged groove is composed of 18 residues that are conserved within the omptin family. Well-defined S1 and S2' subsites are proposed according to the crystal structure. In particular, Glu27 and Asp208 within the deep negative pocket likely form the S1 subsite. M81 and I170 locate at the bottom of a shallow hydrophobic groove about 7 Å away from the negative S1 subsite. This area might form the S2'

subsite that is consistent with the high specificity for small hydrophobic residues at P2' (Vandeputte-Rutten, *et al*., 2001).

1.6.3 Engineering of OmpT

OmpT was first engineered via directed evolution in 2000 by our previous lab member Mark Olsen. A mutant of OmpT with preference for cleavage at Arg-Val was isolated using high throughput screening of large scale library by FACS (Olsen, *et al*., 2000) Later OmpT was successfully engineered to exhibit novel and specific substrate selectivity at the P1 and P1' positions, while still maintaining high overall catalytic activity. (Varadarajan, *et al*., 2005, 2008).

In an interesting example of how subtle recognition in a protease active site can be, a single mutation at 223 from Ser to Arg changed the P1 selectivity completely. While the wild-type OmpT prefers to cleave between Arg-Arg, the variant Ser223Arg prefers to cleave between Ala-Arg and shows more than three million-fold selectivity in favor of Ala-Arg over Arg-Arg (Varadarajan, *et al*., 2005).

In the most dramatic change in an engineered protease specificity yet reported, the P1 specificity of OmpT was converted into a negatively charged Glu residue, a hydrophobic residue Tyr and a small residue Thr. All three of these residues are not recognized by wide-type OmpT at P1. The P1' specificity of OmpT has also been engineered to generate variants that cleave specifically between Arg-Val and Glu-Ala (Varadarajan, *et al*., 2008a). The mutation sites of OmpT variants with altered P1

specificity are depicted in Figure1.4. Val29, Tyr221 and Leu265 are close to the previous predicted negatively charged groove thought to play a role in determining the P1 specificity. However, several distal mutations also show up in these P1 variants. At this point it is not clear which of the mutations are essential for the altered specificity. Nevertheless, the distal mutations, if required for activity, suggest there is subsite cooperativity in OmpT specificity determination (Ng, *et al*., 2009).

Importantly, the RV-OmpT variant exhibits higher than wild-type OmpT activity in plasminogen activation *in vitro*. In fact, the catalytic efficiency of RV-OmpT against plasminogen is comparable to human tissue plasminogen activator (tPA), the clinically used treatment for heart attacks. In addition, human β-defensin could be recognized by RV-OmpT and TR2-OmpT (Varadarajan, *et al*., 2008a) with high efficiency. The bottom line is that these studies have verified that engineered versions of OmpT can cleave biologically significant targets, including full-size proteins, as well as synthetic peptide substrates.

OmpT was further engineered to recognize substrates that are post-translationally modified including sulfated tyrosine (Varadarajan, *et al*., 2008b). The best isolated variant displayed a 200-fold higher specificity for sulfotyrosine over phosphotyrosine and a 10-fold higher specificity for sulfotyrosine over unmodified tyrosine.

A novel multicolor assay was established to tailor OmpT specificity for other post-translationally modified Tyrosine residues including 3-nitrotyrosine. Selection of the 3-nitrotyrosine specific variant was based on the fluorescent signal of the 3-nitrotyrosine

containing selection substrate versus multiple counter-selection substrate signals containing unmodified and various other modified Tyrosine residues. The best isolated variant cleaved 3-nitrotyrosine efficiently, exhibiting a 3600-fold discrimination over sulfotyrosine, a 160-fold discrimination over unmodified , an 8000-fold discrimination over phosphotyrosine and an 8000-fold discrimination over phosphoserine (Varadarajan, *et al*., 2009).

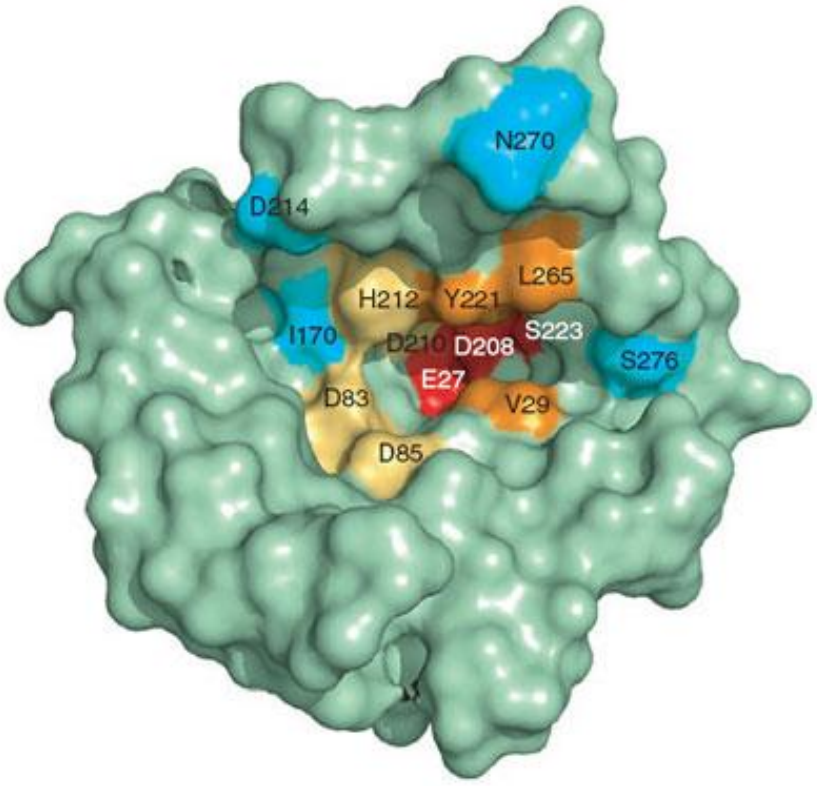


Figure1.4 The surface representation of OmpT. Glu27, Asp208 and Ser223 are shown in red. These red residues are suggested to compose of the S1 subsite. The two pairs of catalytic sites Asp83-Asp85 and His212-Asp210 are shown in yellow. Val29, Tyr221 and Leu265 (in orange) might also account for the P1 specificity. Other distal residues that might account for P1 specificity are shown in cyan. (Di Cera, 2008; Varadarajan, *et al*., 2008).

Despite all of these previous successful engineering efforts using OmpT at P1 and P1', the engineering of positions in the extended active site regions of OmpT has not been reported. In this study, we successfully isolated an engineered OmpT variant with altered P2 specificity.

1.7 Therapeutic applications of protease engineering

1.7.1 Overview of therapeutic proteases

Since proteases play a wide range of roles in different physiological and pathological processes, engineered proteases could have therapeutic application for the treatment of various human diseases. Until now, there have only been a few engineered or recombinant proteases are currently used as therapeutic agents. These include tissue plasminogen activator (tPA, Tenecteplase, Retavase), activated protein C (Xigris), Factor VIIa (NovoSeven) and Factor IXa (Benefix, Mononine).

Human tPA is a serine protease released by endothelial cells. tPA activates the inactive plasminogen into plasmin, and then plasmin, itself a protease, further degrades fibrin clots. Moreover, fibrin binds to tPA and plasminogen, and thus activates its own degradation though a positive feedback mechanism (Hoylaerts, *et al*., 1982). Recombinant tPA has been used in thrombolytic therapy of myocardial infarction, pulmonary embolism, peripheral vessel occlusive disease and stroke (Sarullo, *et al*., 2000; Goldhaber, 2001; Schmulling, *et al*., 2000). Through structure guided genetic engineering, variant TNK-tPA was developed that exhibits a longer elimination half-life than wild type tPA, increased resistance to protease inhibitors and increased fibrin specificity (Stewart, *et al*., 2000). TNK-t-PA (TNK, tenecteplase) has been approved by the US Food and Drug Administration (FDA) (Werner, 2001; www.fda.gov).

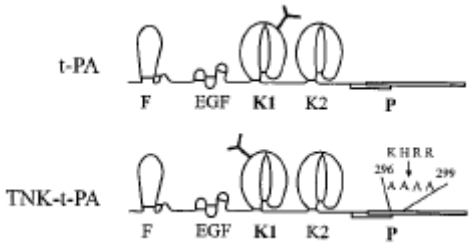


Figure1.5 The secondary structure of tPA and TNK-tPA. (Adapted from Stewart, *et al*., 2000). tPA consists of five domains: F represents a fibronectin finger-like domain; EGF represents a epidermal growth factor domain; K1, K2 represents two kringle domains and P is a protease domain.

Figure1.5 shows a comparison of the secondary structure of tPA and TNK-tPA. There are five domains in tPA: a fibronectin finger-like domain (F), an epidermal growth factor domain (EGF), two kringle domains (K1, K2) and a protease domain (P). Compared to wild-type tPA, TNK tPA contains mutations N117Q (abbreviated N) and T103N (abbreviated T) which shift the glycosylation site in the K1 domain from 117 upstream to 103. This change in glycosylation site accounts for the longer half-life of TNK-tPA. Moreover, the substitution of tetra-alanine to the KHRR (position 296 to 299, abbreviated K) increases the resistance of TNK-tPA to plasminogen activator inhibitor 1, thus strengthening its activity in vivo. The exact reason for increased specificity of TNK-tPA for fibrin remains unclear (Stewart, *et al*., 2000).

Another engineered tPA variant named Reteplase is also widely used as a thrombolytic agent in different clinical situations. Reteplase is a single chain, non-glycosylated polypeptide containing only the K2 and P domains of tPA (presented in

Figure1.5) without the K1, F and EGF domains. Reteplase displays a higher perfusion rate but similar thrombolytic efficacy compared to full-length recombinant tPA. This might be due to the long half-life of Reteplase (Simpson, *et al*., 2006). Moreover, Reteplase lacks fibrin binding so it could penetrate and activate the tissue plasminogen activator inside the clot, which might explain its high in vivo potency (Fisher and Kohnert, 1997).

Human activated protein C (APC) plays a critical role the pathogenesis of sepsis and its related coagulation abnormalities. Protein C exists as an inactivate zymogen in circulation until it is activated by the thrombin-thrombomodulin complex (Yang, *et al.*, 2006). Working together with its cofactor protein S, APC degrades clotting factor V and VIII, thus decreases the generation of factor Xa and thrombin. In this way, APC attenuates the activation of coagulation cascade (Esmon, *et al.*, 1997; Levi, *et al.*, 2007). Recombinant human activated protein C (APC, Xigris) has been used to reduce the mortality of severe sepsis and disseminated intravascular coagulation (DIC) (Bernard, *et al*., 2001; reviewed by Levi, *et al.*, 2007). Recombinant APC has been engineered by structure based approaches to exhibit better pharmacological properties. For example, variant L194S demonstrated higher resistance to protein C inhibitor and slower inactivation rate than wild type. It might be useful in patients with high level of serine protease inhibitors (Berg, *et al*., 2003).

Neither recombinant factor VIIa (Novoseven) nor factor IXa (Benefix, Mononine) has been genetically engineered. However extensive research is going on to search for

better variants of these clotting factors. Both factor VIIa and factor IXa are vitamin K dependent proteases that play a key role in clotting. Novoseven could activate factor X directly in the extrinsic coagulation pathway without the presence of factor VIII and IX. Thus Novoseven is used in hemophilic patients who have the antibodies against factor VIII or IX and it can also be used in the management of some uncontrollable bleeding cases (Lloyd, *et al*., 2003; Franchini, *et al*., 2006). Recombinant factor IXa is routinely used to treat Hemophilia B which is factor IX deficiency (Roth, *et al.*, 2001). Engineered clotting factors with enhanced specificity or pharmaceutical properties are still anticipated.

The Complement system is also a protease cascade which is similar to coagulation system. So far there is no engineered complement interacting protease that is used in clinical therapy. As part of this study we tried to engineer E. coli OmpT to degrade the complement activation products.

1.7.2 Complement system

Therapeutic applications are typically the driving force for protease engineers. The present therapeutic proteases are mostly coming from the coagulation system and its regulatory proteases. The mammalian complement system contains a similar cascade of different proteases. Therefore, modulation of the complement system is an attractive target with many potential therapeutic applications for engineered proteases.

Complement is composed of a group of proteins circulating in the blood. Complement proteins are key components of the innate immune defenses and also link the innate and adaptive immune systems (Caroll, 2004; Lambris, *et al.*, 2007). There are three complement activation pathways involved in the response of pathogens and other harmful agents: the classic, alternative and lectin pathways (Figure1.6). An antigen-antibody complex is necessary for the activation of the classic pathway. Complement thus serves to link innate and adaptive immunity. In alternative pathway C3 can be automatically activated into C3a on microbial surface by certain microbial components such as bacterial lipopolysaccharide (LPS) or certain immunoglobulin such as nephritic factor and aggregated IgA. The lectin pathway activation is triggered by binding of lectin to mannose residues on the pathogen surface, which activates the MBL-associated serine proteases, MASP-1, and MASP-2. Then C4 is cleaved into C4a, C4b and C2 is cleaved into C2a and C2b. C4b and C2a form the C3-convertase to activate the down stream complements, as in the classical pathway. All three pathways lead to the conversion of C3 to C3a and activate its downstream complement components. Ultimately, C5b and C6-C9 form the membrane attack complex (MAC) that leads to cell lysis (Lambris, *et al.*, 2007).

Deregulation of any step in the complement cascade will inhibit or over-activate the complement system leading to pathological conditions. Not surprisingly, improperly regulated complement is involved in the pathogenesis of a wide range of diseases such as autoimmune diseases, ischemia/reperfusion injury and inflammatory diseases. It has been

stated that the development of molecules that inhibit or inactivate various complement components represents a gold mine for drug discovery (Lambris, *et al.*, 2007).

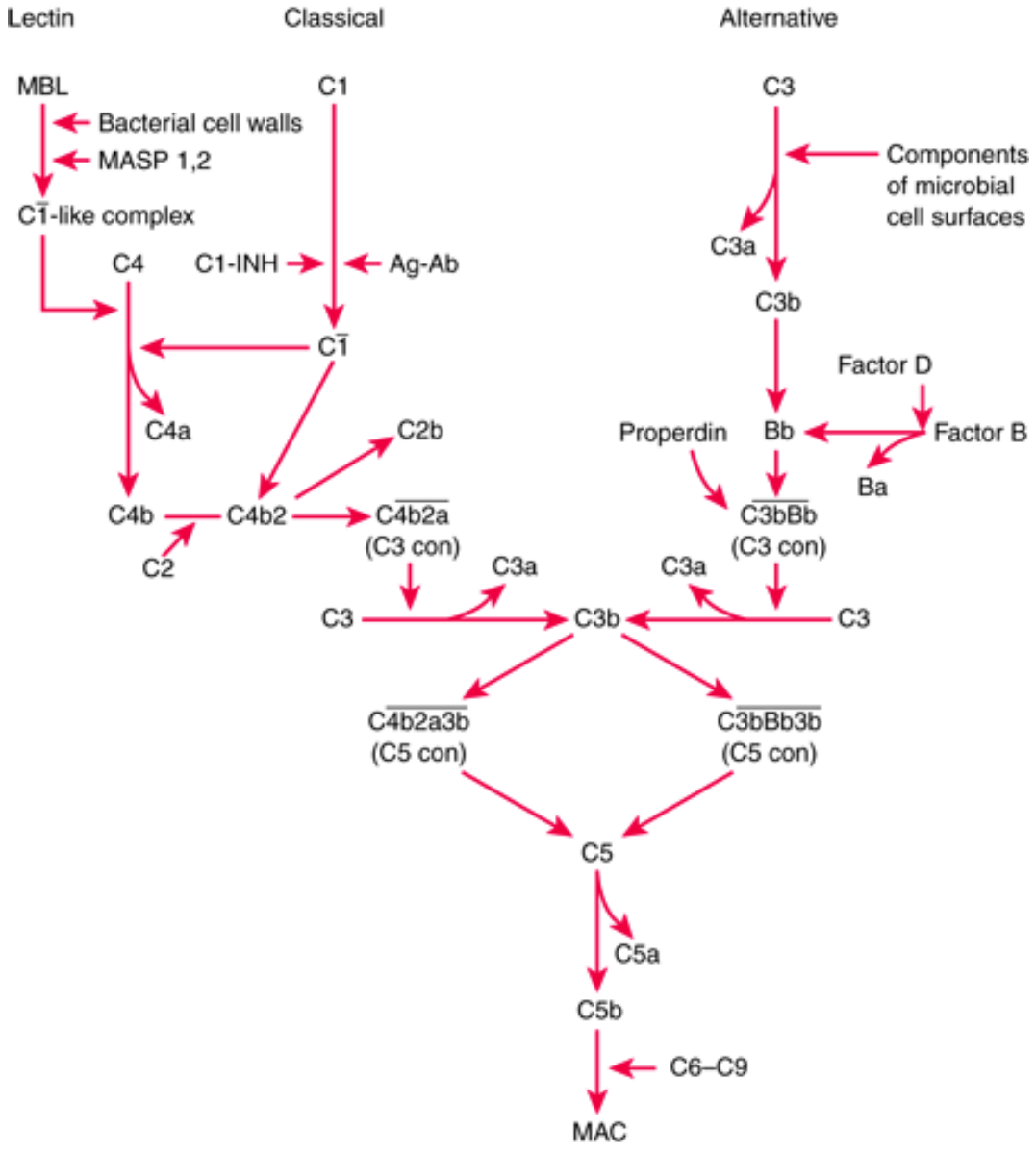


Figure1.6 The complement system: classical, alternative and lectin pathway. C1-INH= C1 inhibitor, MAC= membrane attack complex, MASP= MBL-associated serine protease, MBL= mannose-binding lectin (Adapted from www.merk.com)

1.7.3 Human anaphylatoxin C3a

Anaphylatoxins are protein fragments (C3a, C4a and C5a) generated during the activation of the complement system (Figure1.6) from C3, C4 and C5 respectively. C3a and C5a binding to their receptors leads to mast cell degranulation, an increase of vascular permeability, smooth muscle contraction and release of certain cytokines (Lambris, *et al*., 2005). C3a and C5a are involved in many infectious, autoimmune and inflammatory diseases such as sepsis, asthma, ischemia reperfusion injuries and immune complex diseases (reviewed by Haas and Strijp, 2007).

As part of this study, Complement C3a was selected as a model therapeutic target for inactivation by engineered OmpT. C3a is 77 amino acid peptide (Hugli, 1975), whose level in serum is tightly controlled by carboxypeptidase, generating an inactive desArg-C3a (Ember, *et al*., 1998). Active C3a binds to a G protein coupled receptor (GPCR) to effect its varied (Wetsel, *et al*., 2000). In this way, C3a play an important role in mediating inflammatory diseases such as asthma and sepsis (Ali and Panettieri, 2005; Drouin, *et al*., 2001). The active sites of C3a exist in its C-terminus. It has been reported that the synthetic C3a (70-77) retains only 1-2% of the activity of natural human C3a (Caporale, *et al*., 1980). The disruption of the C terminal peptides of C3a could interfere with its receptor binding, thus decreasing its biological activity. An enzyme cleaving the C-terminus of C3a specifically and efficiently might have anti-inflammatory effects and could have therapeutic applications for the treatment of different

inflammatory diseases such as asthma and sepsis (Humbles, *et al*., 2000; Ali and Panettieri, 2005).

With extensive research studies on the complement system, numerous complement related drugs are under clinical or preclinical trial. However, only two drugs are currently available on the market including complement related protease inhibitor (C1-INH, Cetor/Sanquin, BerinertP/CSL, Behring, Lev Pharma) and monoclonal antibody to complement C5 (Eculizumab/Soliris, Alexion Pharmaceuticals). The purified human C1-INH is used in the treatment of hereditary angio-edema (HAE) which is associated with C1-INH deficiency. The Eculizumab/Soliris is marketed for the treatment of paroxysmal nocturnal hemoglobinuria (PNH) and is still under investigation for other implications (reviewed by Ricklin and Lambris, 2007). Other complement targeted drugs in preclinical or clinical trials include monoclonal antibodies, protein inhibitors, protein regulators and receptor antagonists. None of these is a protease (Ricklin and Lambris, 2007). In this study, we are trying to engineer a protease variant that can degrade complement C3a. This engineered protease variant could be a novel therapeutic agent in the treatment of complement related diseases.

1.8 Substrate specificity investigation

Investigation of substrate specificity plays a central role in characterization of engineered proteases. Substrate phage display provides a means for discerning the extended amino acid specificity of proteases without the need for chemical labeling

(Deperthes, *et al*., 2002; Diamond 2007). It relies on the selective cleavage of specific peptide sequences sandwiched between the PIII minor coat protein of filamentous phage and an affinity tag. The substrate specificity of over 20 bacterial and mammalian proteases have been analyzed by phage display (Cloutier, *et al*., 2002; Felber, *et al*., 2005; Ferrieu-Weisbuch, *et al*., 2006).

As part of this dissertation I used substrate phage to determine the specificity of human tissue kallikrein (KLK) 1 and 6. These two enzymes were evaluated as potential "scaffold proteases' for the engineering of highly selective human peptidase that could be used for therapeutic purposes. The KLK family contains 15 secreted serine proteases that are expressed in a wide range of tissues. For example, KLK3, also called prostate specific antigen (PSA), has been the most effective marker for early detection of prostate cancer (www.cancer.gov). The kallikrein family attracts a lot of attention in biomedical research and drug discovery. KLK1 has been shown to be involved in the regulation of multiple physiological processes such as blood pressure, smooth muscle contraction and vascular cell growth (Bhoola, *et al.*, 1992; Borgono, *et al*., 2004). The physiological function of KLK6 remains unclear. However, KLK6 has been shown to be over-expressed in breast and ovarian cancer tissues. Furthermore, KLK6 been implicated in neurodegenerative diseases such as Alzheimer's disease, multiple sclerosis and Parkinson's disease (Bernett, *et al*., 2002; Iwata, *et al.*, 2003; Magklara, *et al*., 2003). Due to the physiological significance of KLK1 and KLK6, the substrate specificities of these proteases have attracted much attention. As part of the study, we have successfully

profiled the substrate specificities of KLK1 and KLK6 by phage display. A consensus substrate sequence for KLK6 was identified and the identity of the endogenous protein substrates of KLK6 were proposed (Li, *et al*., 2008). The enzyme profiles of KLK1 and KLK6 could provide useful information for engineering human tissue kallikreins in the future.

1.9 Research outline

In this study, chapter 2 represents the engineering of the extended specificity of OmpT at P2. The overall experimental scheme is shown in Figure1.7

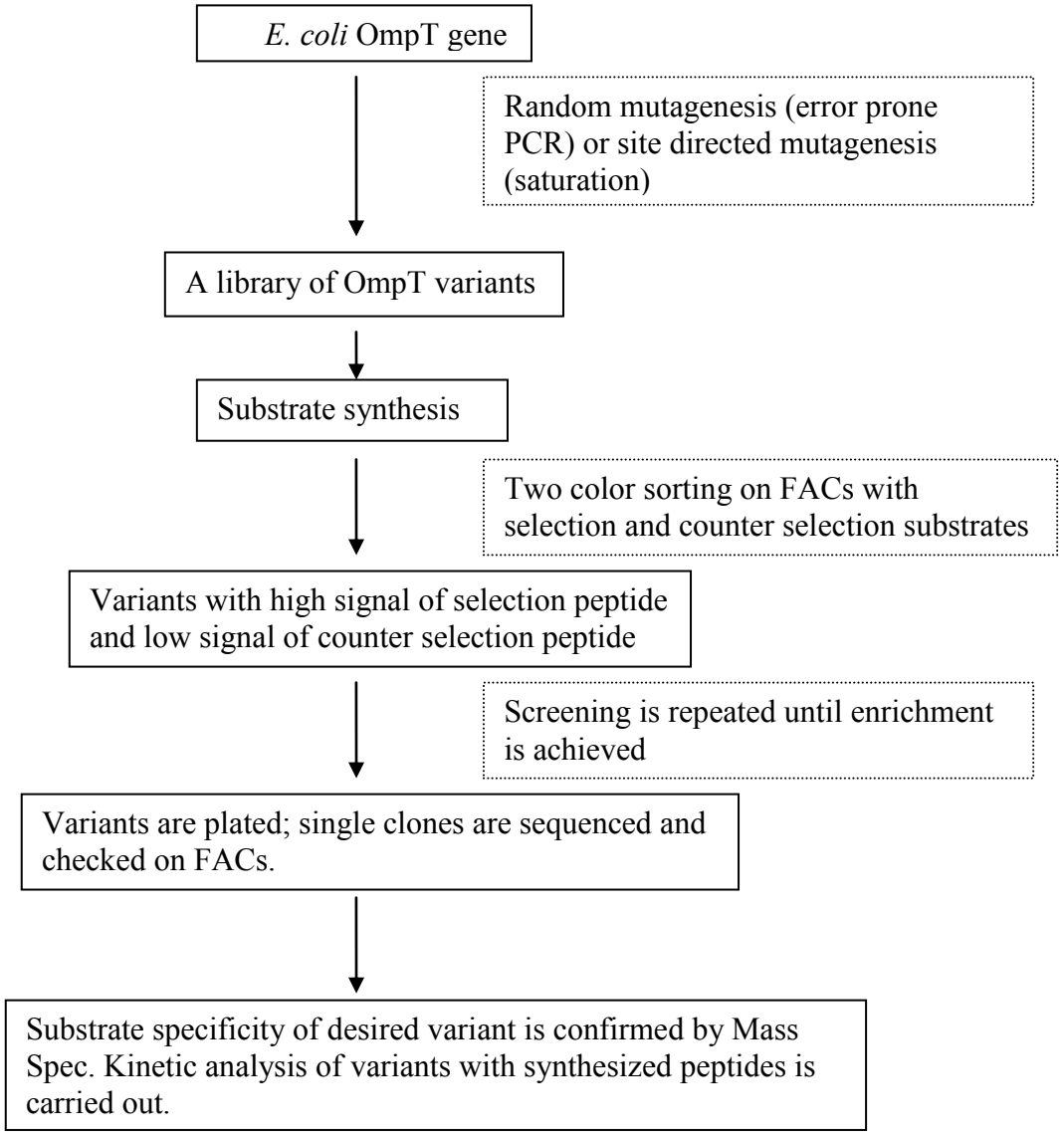


Figure1.7 Overall scheme of engineering *E. coli* OmpT

In chapter3, the engineering of OmpT to specifically cleave C3a is described. The C-terminal residues 67-72, His-Ala-Arg-Ala-Ser-His, are important for C3a activity, so we targeted the cleavage between Arg69 and Ala70 (HAR↓ASH). Although the E. coli outer membrane OmpT and its variants might not be used directly as therapeutic enzyme in humans due to likely allergic reactions in humans, a C3a-specific OmpT could establish a model system for engineering a complement related therapeutic protease.

In chapter 4, the characterization of substrate specificity of Human kallikreins 1 and 6 by phage display is described. The specificities of human serine protease kallikrein 1 and 6 were successfully profiled by screening the substrate peptide library displayed on phage. The putative physiological substrates were proposed. The specificity profile of previously engineered OmpT variants could also be analyzed using this method in the future. The substrate profile might indicate the putative protein targets *in vivo*.

1.10 Reference:

Ali H, Panettieri RA Jr (2005) "Anaphylatoxin C3a receptors in asthma" Respir. Res. 6: 19

Ballinger, M. D., Tom, J. and Wells, J.A. (1996) "Furilisin: a variant of subtilisin BPN engineered for cleaving tribasic substrates." Biochemistry 35: 13579-13585

Baum, E.Z., Bebemitz, G.A. and Gluzman, Y. (1990) "Isolation of mutants of human immunodeficiency virus protease based on the toxicity of the enzyme in Escherichia coli. Proc. Natl. Acad. Sci. USA 87: 5573-5577

Berg, D.T., Gerlitz, B., Shang, J., Smith, T., Santa, P., Richardson, M.A., Kurz, K.D., Grinnell, B.W., Mace, K., Jones, B.E. (2003) "Engineering the proteolytic specificity of activated protein C improves its pharmacological properties." Proc. Natl. Acad. Sci. USA 100: 4423-4428

Bernard, G.R., Vincent, J.L., Laterre, P.F., LaRosa, S.P., Dhainaut, J.F., Lopez-Rodriguez, A., Steingrub, J.S., Garber, G.E., Helterbrand, J.D., Ely, E.W., Fisher Jr., C.J. "Efficacy and safety of recombinant human activated protein C for severe sepsis. N. Engl. J. Med. 344: 699-709

Betnett, M.J., Blaber, S.I., Scarisbrick, I.A., Dhanarajan, P., Thompson, S.M., and Blaber, M. (2002) "Crystal structure and biochemical characterization of human kallikrein 6 reveals that a trypsin-like kallikrein is expressed in the central nervous system." J. Biol. Chem. 277: 24562−24570

Bhoola, K.D., Figueroa, C.D., and Worthy, K. (1992) "Bioregulation of kinins: kallikreins, kininogens, and kininases." Pharmacol. Rev. 44: 1−80

Blanco, R., Carrasco, L. and Ventoso, I. (2003) "Cell killing by HIV-1 protease." J.Biol.Chem. 278: 1086-1093

Borgono, C.A., Michael, L.P., and Diamandis, E.P. (2004) "Human Tissue Kallikreins: Physiologic roles and applications in Cancer." Mol. Cancer Res. 2: 257−280

Bryan, P.N. (2000) "Protein engineering of subtilisin" Biochim. Biophys. Acta. 1543(2): 203-222. Review

.

Caroll, M.C. (2004) "The complement system in regulation of adaptive immunity" Nat. Immuno. 5(10): 981-986

Carters, P. and Wells, J.A. (1987) “Engineering enzyme specificity by "substrate-assisted catalysis". Science 237(4813):394-399

Caporale, L.H., Tippett, P.S. and Erickson, B.W. (1980) “The active site of C3a anaphylatoxin” J. Biol. Chem. 255: 10758-10763

Cloutier, S.M., Chagas, J.R., Mach, J.P., Gygi, C.M., Leisinger, H.J., and Deperthes, D. (2002) “Substrate specificity of human kallikrein 2 (hK2) as determined by phage display technology.” Eur. J. Biochem. 269: 2747−2754

Dekker, N., Cox, R.C., Kramer, R.A. and Egmond, M.R. (2001) “Substrate specificity of the integral membrane protease OmpT determined by spatially addressed peptide libraries.” Biochemistry 40: 1694-1701

Deperthes, D. (2002) “Phage display substrate: a blind method for determining protease specificities.” Biol. Chem. 383: 1107−1112

DeSantis, G., Berglund, P., Stabile, M.R., Gold, M., Jones, J.B. (1998) “Site-directed mutagenesis combined with chemical modification as a strategy for altering the specificity of the S1 and S1’ pockets of subtilisin Bacillus lentus.” Biochemistry 37: 5968-5973

Diamond, S.L. (2007) “Methods for mapping protease specificities.” Curr. Opin. Chem. Biol. 11: 46−51

Di Cera E. (2008) “Engineering protease specificity made simple, but not simpler” Nat. Chem. Bio. 4: 270-271

Drouin, S.M, Kildsgaard, J., Haviland, J., Zabner, J., Jia, H.P., McCray, P.B. Jr, Tack, B.F., Wetsel, R.A. (2001) “Expression of the complement anaphylatoxin C3a and C5a receptors on bronchial epithelial and smooth muscle cells in models of sepsis and asthma” J Immunol. 166(3):2025-32

Ember, J.A., Kildsgaard, J., Haviland, D.L., Lambris J.D. and Holers, M.V. (2000) “Therapeutic interventions in the complement system” Humana press, Totowa, NJ. pp 113-153

Esmon, C.T., Ding, W., Yasuhiro, K., Gu, J.M., Ferrell, G., Regan, L.M., Stearns-Kurosawa, S., Mather, T., Laszik, Z. and Esmon, N.L. (1997) “The protein pathway: new insights” Thromb. Haemostasis 78:70-74

Felber, L.M., Borgono, C.A. Cloutier, S.M., Kundig, C., Kishi, T., Chagas, J.R., Jichlinski, P., Gygi, C.M., Leisinger, H.J., Diamandis, E.P., and Deperthes, D. (2005) "Enzymatic profiling of human kallikrein 14 using phage-display substrate technology." Biol. Chem. 386: 291−298

Ferrieu-Weisbuch, C., Michel, S., Collomb-Clerc, E., Pothion, C., Deleage, G., and Jolivet-Reynaud, C. (2006) "Characterization of prostate-specific antigen binding peptides selected by phage display technology." J. Mol. Recognit. 19: 10−20

Fisher, S. and Kohnert, U. (1997) "Major mechanistic differences explain the higher clot lysis potency of reteplase over alteplase: lack of fibrin binding is an advantage for bolus application of fibrin-specific thrombolytics." Fibrinolysis and Proteolysis 11(3): 129-135

Franchini, M. (2006) "Recombiant factor VIIa: A review on its clinical use" Int. J. Hematol. 83: 126-138

Gill, R.T., DeLisa, M.P., Shiloach, M., Holoman, T.R. and Bentley, W.E. (2000) "OmpT expression and activity increase in response to recombinant chloramphenicol acetyltransferase overexpression and heat shock in E. coli" J. Mol. Microbiol. Biotechnol. 2(3): 283-289

Goldhaber, S.Z. (2001) "Thrombolysis in pulmonary embolism" Circulation 104:2876

Grodberg, J. and Dunn, J.J. (1988) "OmpT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification." J. Bacteriol. 170: 1245-1253

Haas, P.J. and van Strijp J. (2007) "Anaphylatoxins: their role in bacterial infection and inflammation." Immuno. Res. 37(3): 161-175

Han, H.E., Rho, S.H., Lee, Y.J., Park, W.J. (2005) "Engineering of Kex2 variants exhibiting altered substrate specificity" Biochem. Biophys. Res. Commun. 337: 1102-1106

Hanke, C., Hess, J., Schumacher, G., and Goebel, W. (1992) "Processing by OmpT of fusion proteins carrying the HlyA transport signal during secretion by the Escherichia coli hemolysin transport system." Mol. Gen. Genet. 233: 42-48

Hedstrom, L., Szilagyi, L. and Rutter, W.J. (1992) "Converting trypsin to chymotrypsin: the role of surface loops." Science 255(5049):1249-1253

Hedstrom, L. (1996) “Trypsin: a case study in the structural determinants of enzyme specificity.” Biol. Chem. 377(7-8): 465-470

Hedstrom, L. (2002) “Serine protease mechanism and specificity.” Chem. Rev. 102: 4501- 4523

Hoylaerts, M., Rijken, D.C., Lijnen, H.R., and Colleen, D. (1982) “Kinetics of the activation of plasminogen by tissue plasminogen activator” J. Biol. Chem. 257: 2912-2919

Hugli, T.E. (1975) “Human anaphylatoxin (C3a) from the third component of compliment. Primary structure.” J. Biol. Chem. 250: 8293-8301

Humbles, A.A., Lu, B., Nilsson, C.A., Lilly, C., Israel, E., Fujiwara, Y., Gerard, N.P. and Gerard, C. (2000) “A role for the C3a anaphylatoxin receptor in the effector phase of athma.” Nature 406: 998-1001

Hung, S. H. and Hedstrom, L. (1998). “Converting trypsin to elastase: substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity.” Protein Engineering 11(8): 669-673

Iwata, A., Maruyama, M., Akagi, T., Hashikawa, T., Kanazawa, I., Tsuji, S., and Nukima, N. (2003) “Alpha-synuclein degradation by serine protease neurosin: implication for pathogenesis of synucleinopathies.” Hum. Mol. Genet. 12: 2625−2635

Johannes, T., and Zhao H. (2006) “Directed evolution of enzymes and biosynthetic pathways.” Current Opinion in Microbiology 9: 261-267

Kadonosono, T., Kato-Murai M., Ueda M. (2008) “Alteration of substrate specificity of rat neurolysin from matrix metalloproteinase-2/9-type to -3-type specificity by comprehensive mutation.” Protein Eng Des Sel 21: 507-513

Kaur, J. and Sharma R. (2006) “Directed Evolution: An approach to engineer enzymes.” Critical Reviews in Biotechnology 26(3): 165-199

Khersonsky, O., Roodveldt, C., Tawfik, D.S. (2006) “Enzyme promiscuity: evolutionary and mechanistic aspects.” Curr. Opin. Chem. Biol. 10:498-508

Klauser, T., Pohlner, J., and Meyer, T.F. (1992) “Selective extracellular release of cholera toxin B subunit by Escherichia coli: dissection of Neisseria Igαβ-mediated outer membrane transport.” EMBO J. 11: 2327-2335

Knight, Z.A., Garrison, J.L., Chan, K., King, D.S. and Shokat, K.M. (2007) "A remodeled protease that cleaves phosphotyrosine substrates" J. Am. Chem. Soc. 129: 11672-11673

Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M. and Sigal, I.S. (1988) "Active human immunodeficiency virus protease is required for viral infectivity" Proc. Nat. Acad. Sci. 85, 4686-4690

Kramer, R.A., D.Zandwijken, M.R. Egmond, and N. Dekker. (2000) In vitro folding, purification and characterization of Escherichia coli outer membrane protease OmpT. Eur. J. Biochem. 267:885-893

Kukkonen, M., and T. K. Korhonen. (2004) "The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in Escherichia coli to systemic spread of Yersinia pepsis." Int. J. Med. Microbial. 294: 7-14

Kurth, T.; Grahn,S.; Thormann,M.; Ullman, D.; Hofman, H.J.; Jakubke, H.D.; Hedstrom, L. (1998) "Engineering the S1' subsite of trypsin: design of a protease which cleaves between dibasic residues." Biochemistry 37(33): 11434-11440

Lambris, J.D. and Morikis, D. (2005) "Structure biology of the complement system" ISBN: 0-8247-2540-9 page: 161-174

Levi, M. and Poll, T. V. D. (2007) "Recombinant human activated protein C: current insights into its mechanism of action" Critical care 11(suppl. 5): S3

Li, H.X., Hwang, B.Y., Laxmikanthan, G., Blaber, S.I., Blaber, M., Golubkov, P.A., Ren, P., Iverson, B.L. and Georgiou, G. (2008) "Substrate specificity of human kallikreins 1 and 6 determined by phage display." Protein Science 17: 664-672

Lloyd, J.M., Wight, J., Paisley, S., Knight, C. (2003) "Control of bleeding in patients with haemophilia A with inhibitors: a systematic review." Haemophilia 9: 464-520

Magklara, A., Mellati, A.A., Wasney, G.A., Little, S.P., Sotiropoulou, G., Becker, G.W., and Diamandis, E.P. (2003) "Characterization of the enzymatic activity of human kallikrein 6: autoactivation, substrate specificity, and regulation by inhibitors." Biochem. Biophys. Res. Commun. 307: 948−955.

McCarter, J.D., Stephens, D., Shoemaker, K., Rosenberg, S., Kirsch, J.F., Georgiou, G. (2004) "Substrate specificity of the Escherichia coli outer membrane protease OmpT." J. Bacteriol. 186(17): 5919-5925.

McNeil, H.P.; Adachi, R., and Stevens, R.L. (2007) “Mast cell-restricted trypases: structure and function in inflammation and pathogen defense.” J. Bio. Chem. 282: 20786-20789

Ng, N.M., Pike, R.N. and Boyd, S.E. (2009) “Subsite cooperativity in protease specificity” Biol. Chem. 390: 401-407

Olsen, M., Stephens, D.L., Griffiths, D., Daugherty, P.S., Georgiou, G. and Iverson, B.L. (2000). “Function-based isolation of novel enzymes from large libraries.” Nat. Biotechnol. 18: 1071-1076.

O’Loughlin, T.L., Greene, D.N., Matsumua, I. (2006) “Diversification and specialization of HIV protease function during in vitro evolution” Mol. Biol. Evol. 23:764-772

Okuno, K., Yabuta, M., Ohsuye, K., Ooi, T. and Kinoshita, S. (2002) “An analysis of target preferences of Escherichia coli outer membrane endoprotease OmpT for use in therapeutic peptide production: efficient cleavage of substrates with basic amino acids at the P4 and P6 positions” Biotechnol. Appl. Biochem. 36: 77-84

Overall, C.M. and Blobel, C.P. (2007) “In search of partners: linking extracellular proteases to substrates” Mol. Cell Bio. 8: 245-257

Rawlings, N.D. and Barret, A.J. (1994) “Families of serine peptidases.” Methods Enzymol. 244:19-61

Ricklin, D and Lambris, J.D. (2007) “Complement-targeted therapeutics” Nature Biotechnology 25, 1265 - 1275

Roth, D.A., Kessler, C.M., Pasi, K.J., Rup, B., Courter, S.G. and Karen, L. (2001) “Human recombinant factor IX: safety and efficacy studies in hemophilia B patients previously treated with plasma-derived factor IX concentrates” Blood 98(13): 3600-3606

Ruan, B., London, V., Fisher, K.E., Gallagher, T.D. and Bryan, P.N. (2008) “Engineering Substrate Preference in Subtilisin: Structural and kinetic analysis of a specificity mutant” Biochemistry 47: 6628-6636

Rupprecht, K.R., Gordon, G., Lundrigan, M., Gayda, R.C., Markovitz, A., Earhart, C. (1983) “OmpT: Escherichia coli K12 structural gene for protein a (3b).” J. Bacteriol. 153: 1104-1106

Sarullo, F.M., Americo, L., Di Pasquale, P., Castello, A., Mauri, F. (2000) "Efficacy of rescue thrombolysis in patients with acute myocardial infarction: preliminary findings." Cardiovas Drugs Ther 14:83-89

Schmulling, S., Grond, M., Rudolf, J., Heiss, W.D. "One-year follow-up in acute stroke patients treated with rtPA in clinical rountine." Stroke 31: 1552-1554

Sellamuthu, S., Shin, B.H., Lee, E.S., Rho, S.H., Hwang, W., Lee, Y.J., Han, H.E., Kim, J.I., Park, W.J. (2008) "Engineering of protease variants exhibiting altered substrate specificity." Biochem. Biophys. Res. Commun. 371(1): 122-6

Simpon, D., Siddiqui, A.A., Scott, L.J. and Hilleman, D.E. (2006) "Reteplase: A review of its use in the management of thrombotic occlusive disorders" Am. J. Cardiovasc. Drugs 6(4): 265-285

Stewart, R.J., Fredenburgh, J.C., Leslie, B.A., Keyt, B.A. Rischke, J.A. and Weitz, J.I. (2000) "Identification of the mechanism responsible for the increased fibrin specificity of TNK-tissue plasminogen activator relative to tissue plaminogen activator" J. Bio. Chem. 275(14): 10112-10120

Stumpe, S., Schmid, R., Stephens, D..L., Georgiou, G. and Bakker, E.P. (1998). "Identification of OmpT as the protease that hydrolyzes the antimicrobial peptide protamine before it enters growing cells of Escherichia coli." J. Bacteriol. 180(15): 4002-4006.

Sugimura, K. and Higashi, N. (1988). "A novel outer-membrane-associated protease in Escherichia coli." J. Bacteriol. 170(8): 3650-3654.

Sugimura, K. and T. Nishihara. (1988). "Purification, characterization, and primary structure of Escherichia coli protease VII with specificity for paired basic residues: identity of protease VII and OmpT. J. Bacteriol. 170(8): 5625-5632

Tanaka, T., Matsuzawa, H., and Ohta T. (1998) "Engineering of S2 site of aqualysin I; alteration of P2 specificity by excluding side chain." Biochemistry 37: 17402-17407

Turk B. (2006) "Targeting proteases: successes, failures and future prospects." Nature 5: 785-799

Vandeputte-Rutten, L., Kramer, R.A., Kroon, J., Dekker, N., Egmond, M.R. and Gros, P.(2001) "Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site" Embo J. 20: 5033-5039

Varadarajan, N., Gam, J., Olsen M.J., Georgiou, G., and Iverson, B.L. (2005) "Engineering of Protease Variants Exhibiting High Catalytic Activity and Exquisite Substrate Selectivity" Proc. Nat. Acad. Sci. 102, 6855-6860

Varadarajan, N., Rodriguez, S., Georgiou, G., and Iverson, B.L. (2008a) "Highly Active and Selective Endopeptidases with Programmed Substrate Specificities" Nat. Chem. Bio. 4:290-294

Varadarajan, N., Georgiou, G., and Iverson, B.L. (2008b) "An Engineered Protease that Cleaves Specifically after Sulfated Tyrosine" Angew. Chem. Int. Ed. 47:1-4

Varadarajan, N., Pogson, M., Georgiou, G. and Iverson, B.L. (2009) "Proteases that can distinguish among different post-translational forms of tyrosine engineered using multicolor flow cytometry" J. Am. Chem. Soc. 131: 18186-18190

Venekei, I.; Szilagyi, L.; Graf, L. and Rutter, W.J. (1996) "Attempts to convert chymotrypsin to trypsin." FEBS Lett 379: 143-147

Walsh, C. (2001) "Enabling the chemistry of life" Nature 409: 226-231

Webb, R.M. and Lundrigan, M.D. (1996) "OmpT in Escherichia coli correlates with severity of disease in urinary tract infections." Med Microbiol Letters 5: 8-14

Werner, R.G. (2001) "Improving natural principles with genetic engineering: TNK-tissue plasminogen activator." Arzneimittelforschung 51: 516-520

Wetsel, Rick A.; Kildsgaard, Jens; Haviland, David L. (2000) "Complement anaphylatoxins (C3a , C4a, C5a) and their receptors (C3aR, C5aR/CD88) as therapeutic targets in inflammation" Therapeutic Interventions in the Complement System, 113-153.

Willett, W. S., Brinen, L.S., Fletterick and R.J., Craik, C.S. (1996). "Delocalizing trypsin specificity with metal activation." Biochemistry 35(19): 5992-5998.

Yam, C. H., Siu, W.Y., Kaganovich, D., Ruderman, J.V. and Poon, R.Y. (2001) "Cleavage of cyclin A at R70-R71 by the bacterial protease OmpT." Proc. Nat. Acad. Sci. 98(2): 497-501.

Yang L., Manithody, C. and Rezaie, A.R. (2006) "Activation of protein C by the thrombin–thrombomodulin complex: Cooperative roles of Arg-35 of thrombin and Arg-67 of protein C" Proc. Nat. Acad. Sci. 103(4): 879-884

Chapter 2 Engineering Extended Protease Specificity: the S2 subsite of the E. coli outer membrane protease OmpT

2.1 Introduction

Proteases, enzymes that degrade proteins or peptides, play an important role in a wide range of cellular processes involving protein activation and degradation. Both rational design and directed evolution have been used to develop proteases with altered specificity (reviewed by Johannes and Zhao, 2006; Kaur and Sharma, 2006; Pogson *et al*, 2009). However, reprogramming extended specificity of proteases other than P1 and P1' has proven to be a challenge for protein engineers with only a relatively small number of reported successes. Published examples include alteration of the P2 specificity of aqualysin I using site directed mutagenesis (Tanaka, *et al*., 1998) and a subtilisin-like yeast Golgi protease Kex2 that was engineered to exhibit novel P2 specificity through genetic screening (Han, *et al*., 2005). In addition, Hepatitis A virus 3C protease was engineered to exhibit altered P2 specificity by a yeast-based transcriptional activation screen (Sellemuthu, *et al*., 2008). In the present study we used directed evolution with high throughput fluorescence activated sorting to engineer the S2 subsite of the *E. coli* outer membrane protease OmpT.

Engineering protease specificity has therapeutic implications for the treatment of various human diseases. For example, recombinant tissue plasminogen activator (rt-PA) has been used in thrombolytic therapy of myocardial infarction, pulmonary embolism and stroke (Sarullo, *et al*., 2000; Goldhaber, 2001; Schmulling, *et al*., 2000). The

genetically engineered and FDA approved variant TNK-t-PA exhibits increased fibrin specificity and a longer elimination half-life compared to wild-type t-PA. Similarly, recombinant human activated protein C (APC) has been used to reduce the mortality of severe sepsis (Bernard, *et al*., 2001). Specificity of APC has been engineered and variants exhibiting better pharmacological properties have been generated by structure-based mutagenesis (Berg, *et al*., 2003). The general ability to dramatically alter and fine-tune the specificity of a protease would greatly expand the potential for using proteases as catalytic alternatives to therapeutic antibodies. To be useful in a therapeutic context, however, it will be necessary to engineer extended specificity into a protease active site.

In this study, the *E. coli* protease OmpT was used as a model to establish a high throughput method to fine-tune extended enzyme specificity. OmpT is an *E. coli* outer membrane protease that belongs to the omptin family. OmpT prefers dibasic residues in the P1 and P1' positions of the substrate (Sugimura and Nishibara, 1988) and requires lipopolysaccharide (LPS) to exhibit its catalytic activity (Krammer *et al*., 2000). Wild-type OmpT favors small hydrophobic residues at the P2' position of the substrate, but exhibits little specificity at the P2 position (Dekker, *et al*., 2001). OmpT has previously been engineered to prefer a variety of different residues at the P1 and P1' substrate positions while maintaining exquisite overall substrate specificity and high catalytic activity. (Varadarajan, *et al*., 2005, 2008).

The OmpT crystal structure revealed a relatively well-formed S1 subsite, but the putative S2 subsite does not appear to be well-defined (Vandeputte-Rutten, *et al*., 2001).

Phage display analysis indicated that OmpT prefers Ala, Gly, Tyr or Phe and disfavors acidic residues at the P2 position (McCarter, *et al*., 2004).

In the study reported here, screening of OmpT led to the isolation of variants preferring an acidic amino acid (Glu) at P2 while maintaining relatively high catalytic activity, amounting to a roughly 60-fold switch in P2 specificity compared to wild-type OmpT. These results indicate that directed evolution can be successfully employed to engineer extended specificity distal to the scissile bond even when a binding pocket (i.e. subsite) appears to be poorly defined structurally.

2.2 Material and methods

2.2.1 Conjugation of FRET substrates

The peptides Ac-CERRVGKGRGR-NH_2 and Ac-KYRRVGCGRGR-NH_2 were synthesized by GenScript (NJ, USA). BODIPY^{TM}-FL-SE, Qsy7-maleimide and Qsy21-SE were purchased from Invitrogen. Atto-633-maleimide was purchased from ATTO-TEC (Germany). Qsy7-maleimide and BODIPY^{TM}-FL-SE were conjugated to Ac-CERRVGKGRGR-NH_2 at the Cys and Lys residues, respectively, to generate FRET selection substrate **1**. Qsy21-SE and Atto-633-maleimide were conjugated to Ac-KYRRVGCGRGR-NH_2 at the Lys and Cys residues, respectively, to generate FRET counter selection substrate **2** according to literature procedures (Varadarajan, *et al*., 2005). The substrate purity and identity was confirmed by HPLC and ESI-MS.

2.2.2 Library construction

The OmpT gene used as the starting point for library generation was synthesized using the so-called overlapping PCR method (Stemmer, *et al*., 1995). For library generation, the NNS (G/C) was used at chosen positions. The putative S2 region of the active site was identified by inspection of a substrate-OmpT model that was based on our previous results with P1 and P1' engineering. The synthetic OmpT genes were assembled using 48 primers that were designed as previously described (Varadarajan, *et al*., 2009) with 5' EcoRI and 3' HindIII restriction sites surround the finished gene to facilitate unidirectional cloning. PCR products corresponding to the fully assembled gene were gel purified, digested with EcoR1 and HindIII according to the manufacturer's protocols, then ligated into a similarly digested pMLE19 vector. Ligated products were used to transform the *E. coli* MC1061 strain.

An initial saturation library was constructed using primers with NNS codons at residues 29, 31, 39, 40, 263, 265, 274, 276, 280 as previously described (Varadarajan, *et al*., 2009). Following isolation of variant 3 and variant 6 (sequences in Table 2) from the initial saturation library, a new library was constructed using DNA shuffling of the two isolated clones along with wild-type OmpT. The three genes were amplified using outside primers 5' CCGGGAATTCACCATGCGGGCGAAACTTCTGGGAATAGTC 3' and 5' AACAGCCAAGCTTTTAAAATGTGTACTTAAGACCAGCAGT 3'. 5 μg of PCR product was purified and digested using 0.2-1 unit DNase I (NEB) at room temperature for 9 minutes. The reaction was stopped by heating at 75^{o}C for 10 minutes. The cleavage

products were purified by PAGE, and the 50-100 base pair bands were cut out of the gel and purified using Qiagen gel extraction kit.

The isolated DNA fragments were amplified using two rounds of PCR. In the first round, no primers were used. The double-stranded DNA was denatured at 95^{o}C 90s, followed by 40 cycles of denaturation at 94^{o}C for 30s, annealing at 52^{o}C for 45s and extension at 72^{o}C 4 min. A final extension was performed at 72^{o}C for 7 min. The second round of PCR was carried out using the same outside primers listed above. The DNA was denatured at 95^{o}C for 90s, followed by 25 cycles of heating at 94^{o}C 30s, annealing at 57^{o}C 30s and extension at 72^{o}C 90s. The final primer extension was performed at 72^{o}C for 10 min.

An error-prone PCR library was constructed using the 1st round DNA shuffling PCR product as template. The error-rate was targeted to be 0.5% by using 0.5 mM of Mn^{2+} in the buffer mix. The protocol was revised from previously described (Arnold F.H., and Georgiou G., 2003)

In each case, library DNA was digested, ligated into plasmid pMLE19 (Varadarajan, *et al*., 2009) and transformed into *E. coli* strain MC1061. Copies of each library were stored at -80^{o}C as full plasmids isolated from overnight cultures of the transformed MC1061.

2.2.3 Library screening

The plasmid library was transformed into *E.coli* BL21(DE3) (F^- *ompT* $hsdS_B$(r_B-m_B-) *gal dcm* (DE3)) for expression and sorting. The cell suspension was prepared according to the procedure previously described (Varadarajan, *et al*., 2009). A 1 µl of each FRET substrate was added into 973 µl of 1% sucrose and 25 µl of cell suspension. The reaction was incubated at room temperature in the dark for 9 min. A 900 µl reaction mixed with 100 µl of 1% sucrose and analyzed on BD FACS-Aria. The sorting gate was set according to the fluorescence signal and about 0.5% of the cells were collected in 200 µl 2 x YT. The collected cells were plated on LB plates 200 µg/ml and grown at 37^{o}C for 8 to 10 hours. The cells were scraped, sub-cultured and resorted. After 4-6 rounds of screening, plated colonies were analyzed as single colonies by FACS then analyzed by DNA sequencing. The screening scheme is shown in Figure 2.1

2.2.4 Enzyme purification and peptide cleavage analysis

OmpT and its variants were purified according to published procedures (Mangel, *et al*., 1994). Purity of the enzyme was over 90% in each case as determined by SDS-PAGE. Peptide substrates were synthesized by solid phase synthesis (Genescript, NJ, US) and were >95% pure based on HPLC. The kinetic assay was carried out in 50 mM MES (morpholinoethanesulfonic acid), 50 µM TCEP (tris (2-carboxyethyl) phosphine hydrochloride) containing 10 mM EDTA (pH 6.2). 5-320 µM of substrates were incubated with 0.5-5 nM of purified enzyme at 37^{o}C for 5 to 30 min. The reaction was

stopped by adding 1% TFA in 1:1 ratio and freezing in liquid nitrogen. The cleavage of peptides was determined by high-performance liquid chromatography (HPLC) as described previously (Li, et al., 2008). The reported product amounts were determined as peak integration areas at 280 nm and fitted to nonlinear regression of Michaelis-Menten equation using KaleidaGraph.

2.3 Results and Discussion

The presumed S1 subsite, labeled red in Figure 2.1, is well-defined in the OmpT crystal structure. Our previous successful attempts to engineer novel specificity at the P1 substrate position were largely consistent with this predicted S1 location (Vandeputte-Rutten, *et al*., 2001; Varadarajan, *et al*., 2005; 2008). On the other hand, the S2 area of OmpT has not been conclusively identified. Based on modeling with substrate and proximity to the presumed S1, a putative S2 area was identified and is shown as the circled region in Figure 2.1. This area, comprised of residues V29, L31, V39, S40, T263, L265, D274, S276 and A280, lacks any well-defined features such as a cavity that would be expected to facilitate specific recognition. Nevertheless, wild-type OmpT is known to disfavor acidic residues at P2, while favoring Ala, Gly, Tyr or Phe (McCarter, *et al*., 2004).

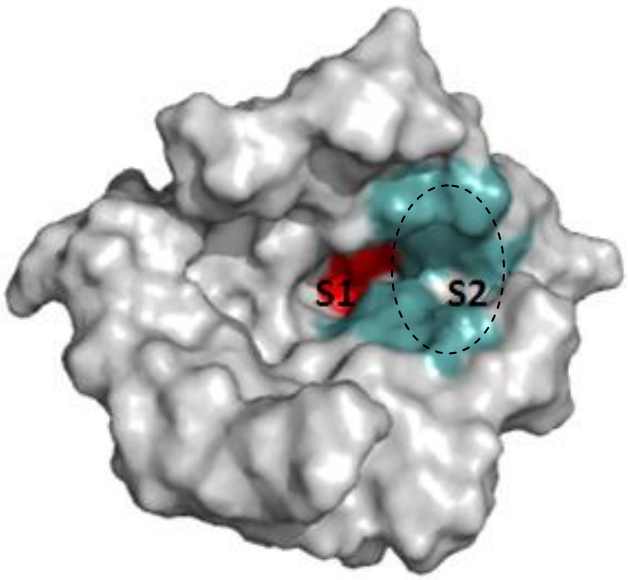


Figure 2.1 Molecular modeling of wild type OmpT. Partial saturation mutagenesis was performed using degenerate codons (NNC/G) at putative S2 subsite shown in circled area in light teal including residues 29, 31, 39, 40, 263, 265, 274, 276, 280. The putative key residues (27, 208 and 223) of the S1 pocket are shown in red for comparison. This figure was produced using PyMOL (www.pymol.org).

In preparation for directed evolution studies, a synthetic saturation library was created using the overlapping PCR method with NNS codons at residues centered around the putative S2 site (residues 29, 31, 39, 40, 263, 265, 274, 276, 280). A two-color FACS screening scheme with selection and counter selection substrates conjugated to different fluorophores and quenchers was used to screen the library, as shown schematically in Figure 2.2.

The selection substrate (**1**) contained a Glu residue at the expected P2 site along with a BODIPY^{TM} and Qsy7 FRET pair on either side of the scissile bond. The counter selection substrate (**2**) contained the wild-type preferred Tyr at P2 along with an Atto633 and Qsy21 FRET pair on either side of the scissile bond. For either substrate, OmpT cleavage results in significant fluorescence enhancement due to separation of the FRET pair. Importantly, the highly fluorescent BODIPY^{TM} and Atto633 containing cleavage products have a +3 overall charge, allowing for electrostatic capture on the negatively charged *E. coli* surface. Using an isotonic sucrose solution provided for retention of surface bound fluorescent product at concentrations that are dependent on enzymatic turnover, allowing for the FACS isolation of the most active (most intense fluorescence) clones.

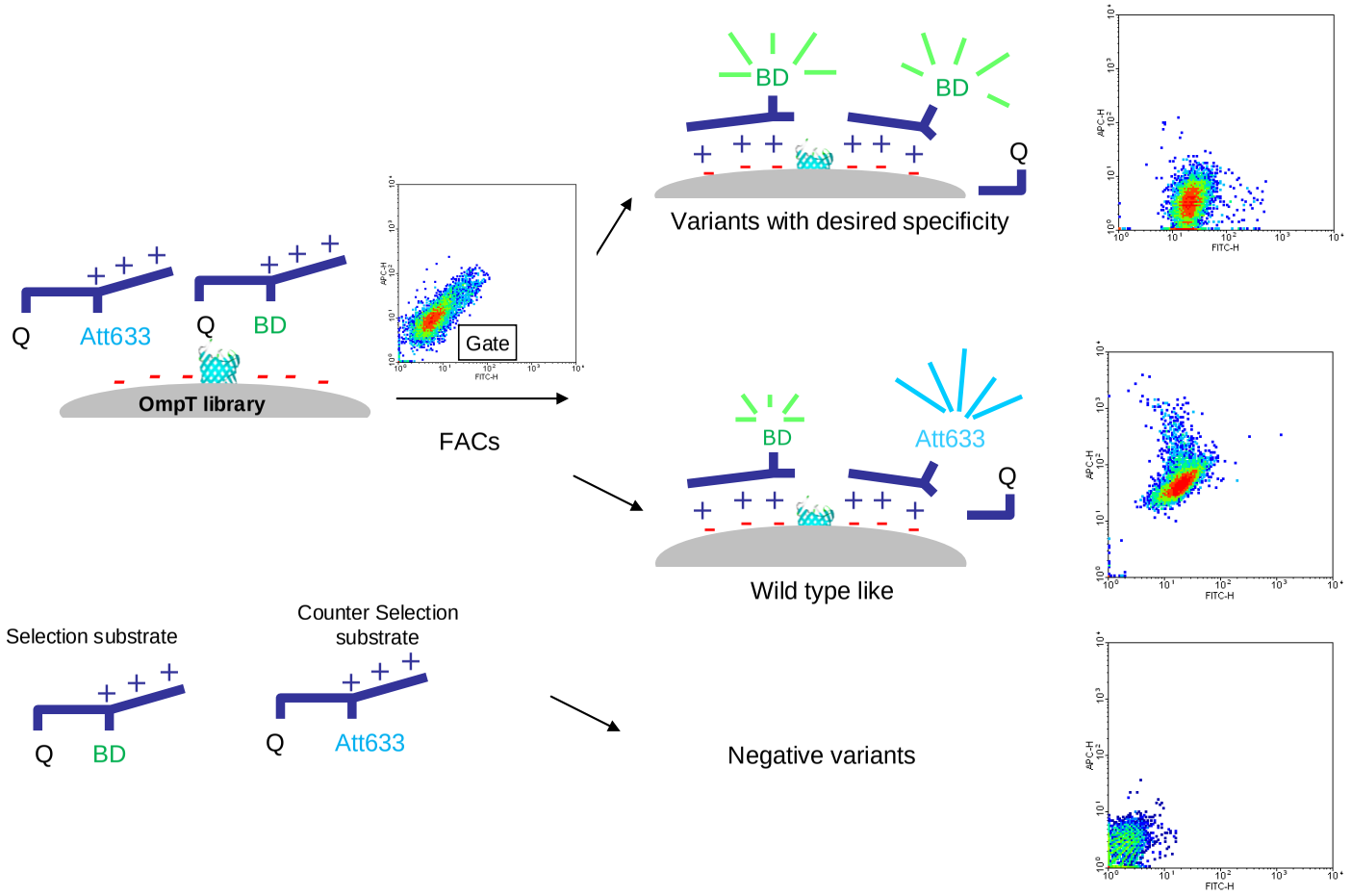
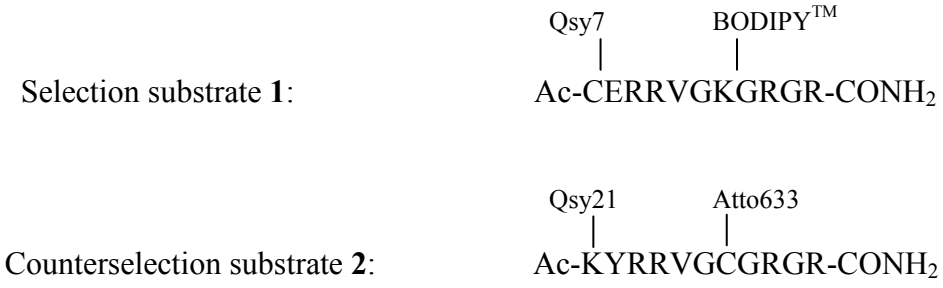


Figure 2.2 Two-color fluorescence activated cell sorting scheme used for isolating OmpT variants exhibiting altered P2 specificity. BD=BODIPY, Att633=ATTO633, Q=Qsy7/Qsy21). Both substrates are FRET substrates. Selection substrate is conjugated to BD and its quencher Qsy7 while counterselection substrate is conjugated to Att633 and its quencher Qsy21.

FACS sort gates were set to collect cells expressing OmpT variants with high BODIPY^{TM} and low Atto633 fluorescence signal. This strategy was intended to maximize cleavage of the Glu containing selection substrate (**1**), while minimizing cleavage of the wild-type preferred Tyr containing counter-selection substrate (**2**). After five rounds of sorting the saturation P2 library, twenty different variants were isolated (sequences shown in Table 2.1). Variant 3 and 17 have the same sequences. A near consensus 20 out of 21 clones contain a Val to a positively charged Arg or Lys mutation at position 39. In addition, 11 out of 21 clones possess a Leu to Arg or Lys mutation at position 265 and 9 contain an Asp to Arg or Lys at position 274. Taken together, the striking preponderance of positively charged residues in selected clones would seem to be consistent with selection for greater activity with the negatively charged Glu residue in P2.



Variants #	L31	V39	L265	D274	S276	other
1		R	I	R		S12R
2		R	K		Y	
3		R		R		
4		R	R	W		
5	V	K	I	R		E27D
6	V	K	V		K	
7		R	R	N		
8		R	R	M		
9		Q	H	W		S99T
10		R	K	W		
11		R	R	R		
12	K		R	R		
13		R	V	K		
14	K	K	V		K	
15	K	R		R		
16		R	I	K		
17		R		R		
18		R	R	L		
19	R	R	R	L		
20		R	R		Y	
21		R	K	W		

Table 2.1 Sequences of variants that were isolated from five rounds sorting of the saturation library.

The isolated OmpT variants were analyzed as single clones with both substrates **1** and **2**. In addition, the variants were individually purified and analyzed on HPLC using synthesized peptides WCYR↓RVGKGR (**3**) (wild type preferred) and WCER↓RVGKGR (**4**) (target specificity with a Glu residue at P2). On these peptides, the *N*-terminal tryptophan was added as an absorbance handle for detection purposes. Based on results of these initial screens (data not shown), variant 3 (V39R, D274R) and variant 6 (L31V,

V39K, L265V, S276K) were selected to undergo further mutagenesis. The variants 3 and 6 were "backcrossed" with wild-type OmpT via DNA shuffling, followed by further error-prone PCR to generate a library of 10^6 different transformants. After four rounds of FACS sorting with substrates **1** and **2**, the mean intensity of BODIPY^{TM} fluorescence was increased (Figure 2.3). Seven clones were isolated (shown in Table 2.2). Single clones were cultured and analyzed on FACS as before. The variant 25 was selected for further analysis because it exhibited the highest preference for the selection substrate **1** over substrate **2** (Figure 2.4).

The enzymatic kinetics of purified samples of wild-type OmpT and the variant 25 were determined using an HPLC assay with peptide substrates **3** and **4** (Table 2.3). The wild-type OmpT cleaved the YR↓R containing substrate (**3**) roughly 14-fold more efficiently than the ER↓R substrate (**4**) based on a comparison of catalytic efficiency (k_{cat}/K_M) values. This wild-type specificity is largely due to differences in K_M values. On the other hand, comparison of catalytic efficiencies (k_{cat}/K_M) for the variant 25 revealed a roughly 4-fold preference for the ER↓R substrate (**4**) compared to the YR↓R containing substrate (**3**). For the variant 25, specificity is largely the result of differences in K_M values. It is possibly due to the positive charged mutations on putative S2 subsite which altered the binding of Glu or Tyr at P2 of the substrate. It is important to note that overall, the catalytic efficiency of the variant 25 reacting with the ER↓R substrate (**4**) was about three fold lower than that of wild type OmpT reacting with the YR↓R containing substrate (**3**).

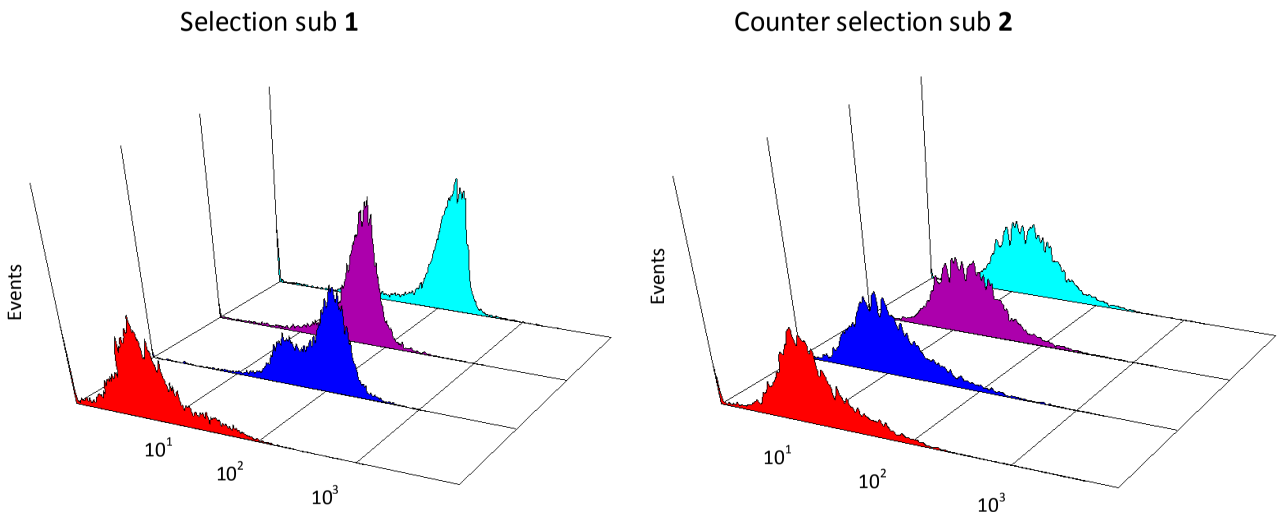
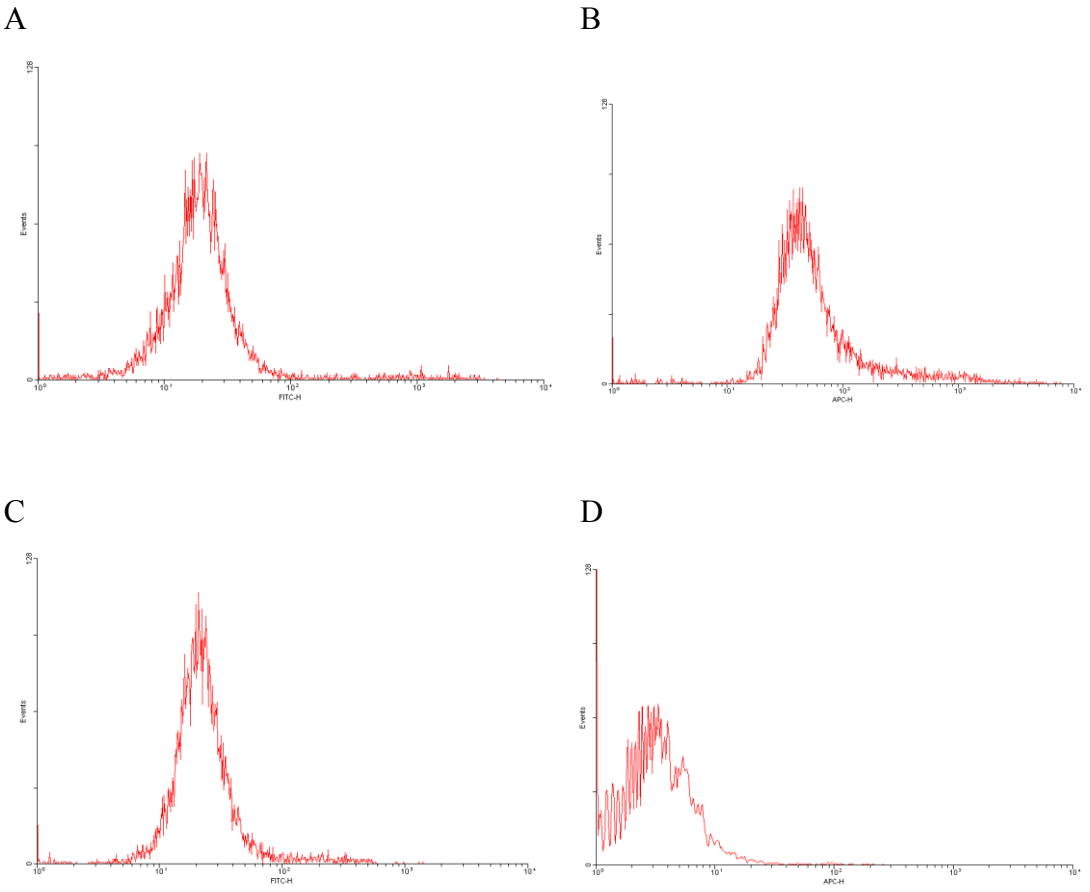


Figure 2.3 Two-color fluorescence activated cell sorting scheme used for isolating OmpT variants exhibiting altered P2 specificity. BD=BODIPY, Att633=ATTO633, Q=Qsy7/Qsy21). Both substrates are FRET substrates. Selection substrate is conjugated to BD and its quencher Qsy7 while counterselection substrate is conjugated to Att633 and its quencher Qsy21.



Selection substrate **1**: Ac-C(Qsy7)ERRVGK(BODIPY™)GRGR-CONH$_2$

Counterselection substrate **2**: Ac-K(Qsy21)YRRVGC(Atto633)GRGR-CONH$_2$

Figure 2.4 FACS analysis of wild type OmpT and the variant 25. A: wild type OmpT with the selection substrate **1**. B: wild type OmpT with the counter selection substrate **2**. C: The variant 25 with the selection substrate **1**. D: The variant 25 with the counter selection substrate **2**.

Table 2.2 Sequences of variants that were isolated from sorting of DNA shuffling and error-prone PCR library. Positive charged mutations were marked in red.

Variants #	L 31	V39	L265	D274	S276	other
3 (parent)		R		R		
6 (parent)	V	K	V		K	
22	V	K	V		K	
23		R		R		R127H
24	V	K	Q	R		
25	V	K		R		E153K
26		R	R	W		
27	V	R		R		
28		R		R		G145S

Closer inspection of the data in Table 2.3 reveals that although there is an approximately 56-fold change in specificity between wild type OmpT and variant 25, the difference comes down to variant 25 having a significantly decreased ability to react with the YR↓R containing substrate (**3**) and an increase in activity with the ER↓R substrate (**4**). Michaelis-Menten parameters of kinetics of wild type OmpT and variant 25 with substrate (**3**) and (**4**) are shown in Figure 2.6 and 2.7.

The variant 25 contains four mutations: L31V, V39K, E153K and D274R (shown in Figure 2.5). Among these, the L31V, V39K, and D274R mutations reside within the putative S2 subsite of OmpT. The observed alteration of efficiency appears to be the result of both favoring ER↓R (**4**) and disfavoring reaction with the YR↓R containing

substrate (**3**). Nevertheless, by the judicious choice of selection and counter-selection substrates along with targeted mutagenesis, we have demonstrated that dramatic changes in specificity at P2 can be achieved through directed evolution/high throughput screening, even in the case of OmpT which apparently lacks a well-defined S2 subsite. These results therefore bode well for the engineering of extended specificity into the active sites of proteases intended for therapeutic or other applications.

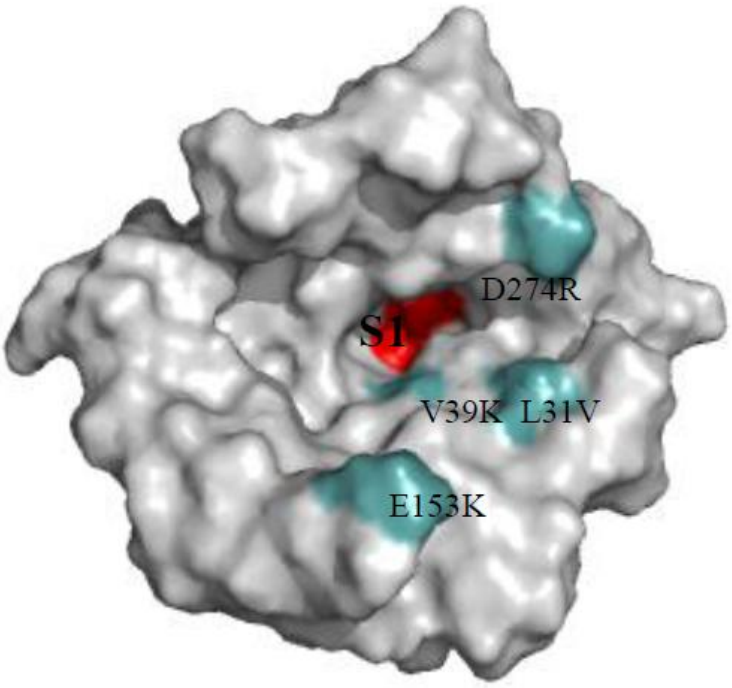
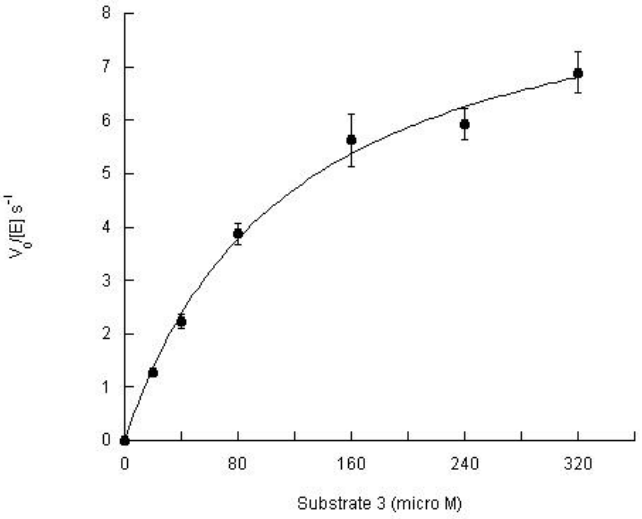


Figure 2.5 Molecular modeling of variant 25 locations of mutations (light teal) found in the variant 25 is indicated on the wild type OmpT surface. This figure was produced using PyMOL (www.pymol.org). Note that multiple positive charged mutations on variant 25 at putative S2 subsite but not on WT OmpT.

Table 2.3 Kinetics of wild type OmpT and isolated variant 25. . The kinetic assay was carried out in 50 mM MES, 50 μM TCEP containing 10 mM EDTA (pH 6.2). 5-320 μM of substrates were incubated with 0.5-5 nM of purified enzyme at 37°C for 5 to 30 min.

Enzyme	Peptide substrate	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})
WT OmpT	WCYR↓RVGKGR (**3**)	7.4± 0.2	9 ± 1	$8 \pm 1.2 \times 10^5$
	WCER↓RVGKGR (**4**)	4.0 ± 0.2	70±8	$6 \pm 1 \times 10^4$
Variant 25	WCYR↓RVGKGR (**3**)	9.3± 0.4	120 ± 12	$8 \pm 0.9 \times 10^4$
	WCER↓RVGKGR (**4**)	8.9± 0.3	27 ± 3	$3 \pm 0.5 \times 10^5$



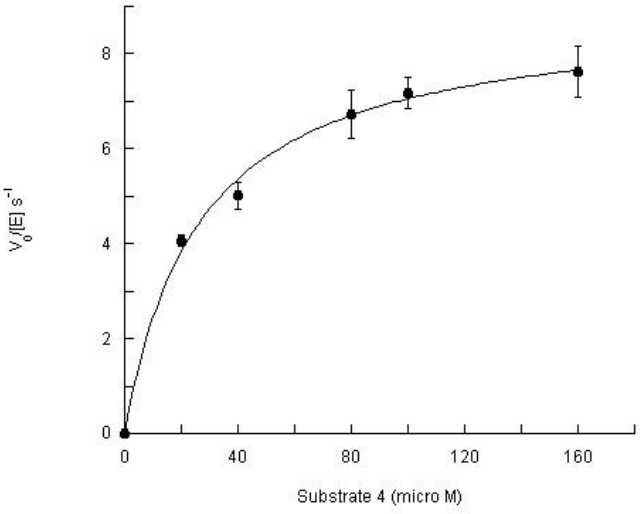
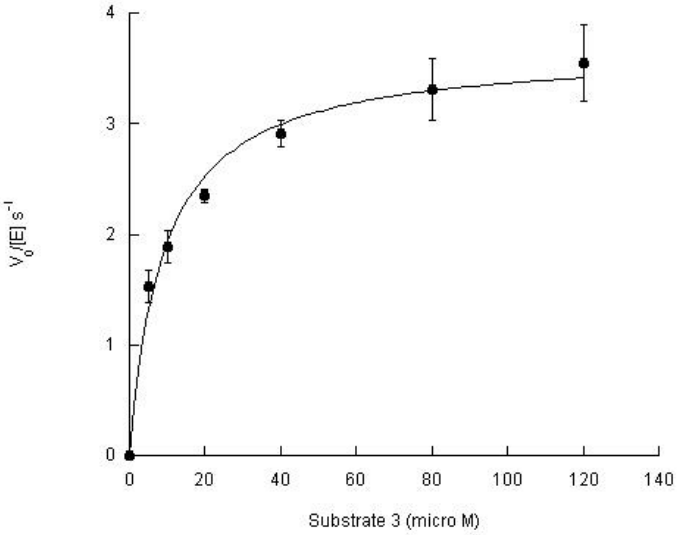


Figure 2.6: Michaelis-Menten parameters for the hydrolysis of substrate 3 (WCYRRVGKGR) and 4 (WCERRVGKGR) by variant 25. The kinetic assay was carried out in 50 mM MES (morpholinoethanesulfonic acid), 50 μM TCEP (tris (2-carboxyethyl) phosphine hydrochloride) containing 10 mM EDTA (pH 6.2).



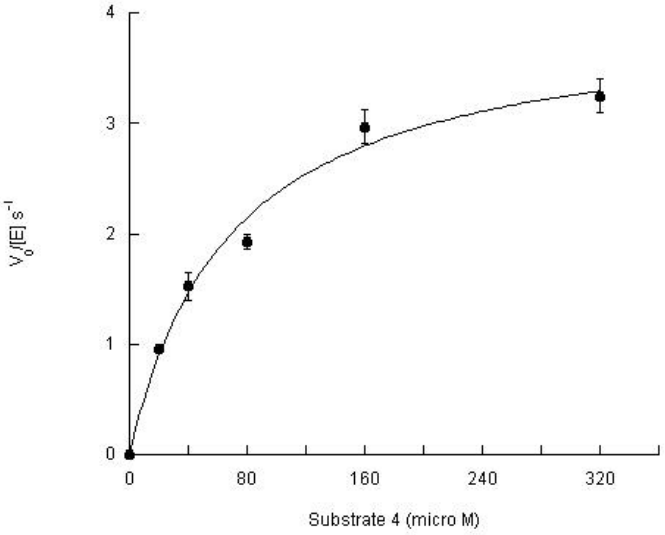


Figure 2.7: Michaelis-Menten parameters for the hydrolysis of substrate 3 (WCYRRVGKGR) and 4 (WCERRVGKGR) by wild type OmpT. The kinetic assay was carried out in 50 mM MES (morpholinoethanesulfonic acid), 50 μM TCEP (tris (2-carboxyethyl) phosphine hydrochloride) containing 10 mM EDTA (pH 6.2).

Acknowledgements

We thank Dr. Navin Varadarajan for his thoughtful input in this project. This work was supported by National Institute of Health Grants R01 GM065551 (to Dr. Brent Iverson and George Georgiou).

2.4 References

Arnold F.H., and Georgiou G. (2003) “Directed Evolution Library Creation. Methods and protocols” 75-90

Berg, D.T., Gerlitz, B., Shang, J., Smith, T., Santa, P., Richardson, M.A., Kurz, K.D., Grinnell, B.W., Mace, K., Jones, B.E. (2003) “Engineering the proteolytic specificity of activated protein C improves its pharmacological properties.” Proc. Natl. Acad. Sci. USA 100: 4423-4428

Bernard, G.R., Vincent, J.L., Laterre, P.F., LaRosa, S.P., Dhainaut, J.F., Lopez-Rodriguez, A., Steingrub, J.S., Garber, G.E., Helterbrand, J.D., Ely, E.W., Fisher Jr., C.J. “Efficacy and safety of recombinant human activated protein C for severe sepsis. N. Engl. J. Med. 344:699-709

Carters, P. and Wells, J.A. (1987) Engineering enzyme specificity by "substrate-assisted catalysis". Science 237(4813):394-399

Dekker, N., Cox, R.C., Kramer, R.A. and Egmond, M.R. (2001) “Substrate specificity of the integral membrane protease OmpT determined by spatially addressed peptide libraries.” Biochemistry 40: 1694-1701

Diamond, S.L. (2007) “Methods for mapping protease specificities.” Curr. Opin. Chem. Biol. 11: 46−51

Di Cera E. (2008) “Engineering protease specificity made simple, but not simpler” Nat. Chem. Bio. 4: 270-271

Ekici O.D., Karia A., Paetzel M, Lively M.O., Pei D, Dalbey R.E. (2007) “ Altered P3 substrate specificity of Escherichia coli signal peptidase 1 mutants as revealed by screening a combinatorial peptide library. J Biol. Chem. 282:417-425

Goldhaber, S.Z. (2001) “Thrombolysis in pulmonary embolism” Circulation 104:2876

Grodberg, J. and Dunn, J.J. (1988) “OmpT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification.” J. Bacteriol. 170: 1245-1253

Han, H.E., Rho, S.H., Lee, Y.J., Park, W.J. (2005) “Engineering of Kex2 variants exhibiting altered substrate specificity” Biochem. Biophys. Res. Commun. 337: 1102-1106

Hwang, B.Y., Varadarajan, N., Li, H.X., Rodriguez, S., Iverson, B. L., Georgiou, G. (2007) "Substrate specificity of the Escherichia coli outer membrane protease OmpP" Journal of Bacteriology, 189(2), 522-530

Johannes, T., and Zhao H. (2006) "Directed evolution of enzymes and biosynthetic pathways." Current Opinion in Microbiology 9: 261-267

Kaur, J. and Sharma R. (2006) "Directed Evolution: An approach to engineer enzymes." Critical Reviews in Biotechnology 26(3): 165-199

Kerr, F.K., O'Brien G., Quinsey N.S., Whisstock J.C., Boyd S., de la Banda M.G., Kaiserman D., Matthews A.Y., Bird P.I., and Pike R.N. (2005) "Elucidation of the substrate specificity of the C1s protease of the classical complement pathway." J. Biol. Chem. 280: 39510–39514.

Kramer, R.A., D.Zandwijken, M.R. Egmond, and N. Dekker. (2000) In vitro folding, purification and characterization of Escherichia coli outer membrane protease OmpT. Eur. J. Biochem. 267:885-893

Kukkonen, M., and T. K. Korhonen. (2004) "The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in Escherichia coli to systemic spread of Yersinia pepsis." Int. J. Med. Microbial. 294: 7-14

Li, H.X., Hwang, B.Y., Laxmikanthan, G., Blaber, S.I., Blaber, M., Golubkov, P.A., Ren, P., Iverson, B.L. and Georgiou, G. (2008) "Substrate specificity of human kallikreins 1 and 6 determined by phage display." Protein Science 17: 664-672

McCarter, John D.; Stephens, Daren; Shoemaker, Kevin; Rosenberg, Steve; Kirsch, Jack F.; Georgiou, George. (2004) "Substrate specificity of the Escherichia coli outer membrane protease OmpT." Journal of Bacteriology 186(17): 5919-5925.

Mangel, w., d.L. Toledo, M.T. Brown, K. Worzalla, Ml Lee, and J.J. Dunn. (1994). Omptin: an Escherichia coli outer membrane proteinase that activates plasminogen. Methods enzymol. 244: 384-399.

Olsen, M., Stephens, D.L., Griffiths, D., Daugherty, P.S., Georgiou, G. and Iverson, B.L. (2000). "Function-based isolation of novel enzymes from large libraries." Nature Biotechnology 18: 1071-1076.

Pogson, M., Georgiou G. and Iverson, B.L. (2009) "Engineering next generation proteases." Current Opinion in Biotechnology 20: 390-397

Ruan, B., London, V., Fisher, K.E., Gallagher, T.D. and Bryan, P.N. (2008) "Engineering Substrate Preference in Subtilisin: Structural and kinetic analysis of a specificity mutant" Biochemistry 47: 6628-6636

Sarullo, F.M., Americo, L., Di Pasquale, P., Castello, A., Mauri, F. (2000) "Efficacy of rescue thrombolysis in patients with acute myocardial infarction: preliminary findings." Cardiovasc. Drugs. Ther. 14: 83-89

Schmulling, S., Grond, M., Rudolf, J., Heiss, W.D. "One-year follow-up in acute stroke patients treated with rtPA in clinical rountine." Stroke 31: 1552-1554

Sellamuthu, S., Shin, B.H., Lee, E.S., Rho, S.H., Hwang, W., Lee, Y.J., Han, H.E., Kim, J.I., Park, W.J. (2008) "Engineering of protease variants exhibiting altered substrate specificity." Biochem. Biophys. Res. Commun. 371(1): 122-6

Stemmer, Willem P. C.; Crameri, Andreas; Ha, Kim D.; Brennan, Thomas M.; Heyneker, Herbert L. (1995) "Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides" Gene 164(1): 49-53.

Stumpe, S., Schmid, R., Stephens, D..L., Georgiou, G. and Bakker, E.P. (1998). "Identification of OmpT as the protease that hydrolyzes the antimicrobial peptide protamine before it enters growing cells of Escherichia coli." J. Bacteriol. 180(15): 4002-4006.

Sugimura, K. and Higashi, N. (1988). "A novel outer-membrane-associated protease in Escherichia coli." J. Bacteriol. 170(8): 3650-3654.

Sugimura, K. and T. Nishihara. (1988). "Purification, characterization, and primary structure of Escherichia coli protease VII with specificity for paired basic residues: identity of protease VII and OmpT. J. Bacteriol. 170(8): 5625-5632

Tanaka, T., Matsuzawa, H., and Ohta T. (1998) "Engineering of S2 site of aqualysin I; alteration of P2 specificity by excluding side chain." Biochemistry 37: 17402-17407

Turk B. (2006) "Tarkgeting proteases: successes, failures and future prospects." Nature 5: 785-799

Varadarajan, N., Gam, J., Olsen M.J., Georgiou, G., and Iverson, B.L. (2005) "Engineering of Protease Variants Exhibiting High Catalytic Activity and Exquisite Substrate Selectivity" Proc. Nat. Acad. Sci. 102, 6855-6860

Varadarajan, N., Rodriguez, S., Georgiou, G., and Iverson, B.L. (2008) "Highly Active and Selective Endopeptidases with Programmed Substrate Specificities" Nat. Chem. Bio. 4:290-294
Varadarajan, N., Cantor, J.R., Georgiou, G., and Iverson, B.L. (2009) "Construction and flow cytometric screening of targeted enzyme libraries." Nat. Protoc. 4(6):893-901

Werner, R.G. (2001) "Improving natural principles with genetic engineering: TNK-tissue plasminogen activator." Arzneimittelforschung 51: 516-520

Chapter 3 Engineering OmpT Variants that can Degrade Human Anaphylatoxin C3a

3.1 Introduction

Complement factors is a network of plasma and cell surface proteins that participate in the response to disease agents and play an important role in immunity and inflammation. Complement factors were first discovered by Bordet in 1896 as "heat liable" serum component that "complements" antibodies to kill bacteria. The complement system is known to be a complex network comprising more than 30 proteins in the blood and on cell surfaces (reviewed by Dunkelberger and Song, 2010). It constitutes a principal component of the innate immune system and also links the innate and adaptive immune responses (Caroll, 2004; Lambris, *et al.*, 2007). There are three complement activation pathways: the classic, the alternative and the lectin pathway (Figure1.5). The presence of antigen antibody complexes is necessary for the activation of the classic pathway. The alternative and lectin pathway can be spontaneously activated by microbial components such as bacterial surfaces. All these pathways lead to the conversion of C3 to C3a which in turn activates downstream processes mostly via a series of proteolytic cleavage reactions Eventually, the C5b and C6-C9 proteins form the membrane attack complex (MAC) which leads to cell lysis (Lambris, *et al.*, 2007). However, the whole picture of the function and control of the complement system is still under investigation.

The complement pathway plays an essential role in immune responses. Disregulation of any step in this cascade will disrupt the balance of complement activation and inhibition leading to pathological circumstances. Similarly, deficiency of components of the complement system leads to diseases. For example, deficiency of the complement C1 esterase inhibitor causes autosomal dominant disorder: hereditary angioedema (Bernstein, 2008). Another inhibitory component of the complement system, the decay accelerating factor (DAF, a GPI-anchored enzyme) residing on the surface of red blood cells protects RBCs from lysis by inadvertent complement activation on the RBC surface. Deficiency of DAF leads to complement mediated hemolysis and causes paroxysmal nocturnal hemoglobinuria (PNH) (Jonhson and Hillmen, 2002). Deficiency of C5 to C8 causes inability to form the MAC complex and leads to increased risk of *Neisseria* infections (Hellerud, *et al.*, 2010). Complement C3 deficiency is associated with recurrent pyogenic sinus and respiratory infection (Tedesco, *et al.*, 1993). Moreover, deficiency of C3 has been shown recently to increase β-amyloid deposition and neurodegeneration in amyloid precursor protein (APP) transgenic mice (Maier, *et al.*, 2008). It is proposed that C3 might play a beneficial role in protection from Alzheimer's disease by clearing plaque and by maintaining neuronal health (Maier, *et al.*, 2008). Furthermore, disregulation of the complement system is associated with the pathogenesis of a wide range of diseases such as autoimmune diseases, immunodeficiency, ischemia/reperfusion injury and inflammatory diseases (Lambris, *et al.*, 2007).

Anaphylatoxins are pro-inflammatory mediators (C3a, C4a and C5a) generated during the activation of complement system (Figure1.5). C3a and C5a bind to their receptors on mast cells and lead to mast cell degranulation. Then, mast cells release inflammatory molecules such as histamine, which increases of vascular permeability and cause smooth muscle contraction (Lambris, *et al*., 2007). C3a and C5a are involved in different inflammatory diseases such as sepsis, asthma and ischemia-reperfusion (I/R) injury following stroke (reviewed by Haas and Strijp, 2007). It has been reported that C3a level increased in asthma patients but not in healthy individuals after allergen challenge (Ali and Panettieri, 2005). It has been suggested that C3a regulates the interaction between mast cells and airway smooth muscle cells. Therefore, it has been proposed that therapeutic intervention to decrease the C3a level could be useful in the treatment of asthma (Ali and Panettieri, 2005).

I/R injury is a common clinical situation in stroke. The disruption of blood flow causes tissue ischemia and the deprivation of oxygen causes severe tissue damage. However the reperfusion after ischemia elicits intense inflammatory responses causing further tissue damage. It has been shown that in multiple organ I/R injury triggers the activation of the complement system and the release of the anaphylatoxin C3a, C5a as well as the membrane attack complex (Arumugam *et al.,* 2004). No effective treatment for I/R injury is available

We hypothesized that a protease that cleaves selectively C3a would be of therapeutic value. Therefore, as part of this dissertation C3a was used as a model disease target substrate for engineered OmpT. C3a is a 77 amino acid peptide (Hugli, 1975). In serum, the level of C3a is tightly controlled by the non-specific protease carboxypeptidase N which removes the C terminal Arg and generates the inactive product desArg-C3a (Ember, *et al*., 2000). After binding to its G protein coupled receptor (GPCR) on mast cells and basophils, C3a causes degranulation of these cells releasing the inflammatory molecules such as histamine, which leads to smooth muscle contraction, increase of vascular permeability and other inflammatory responses (Wetsel, *et al*., 2000). Thus C3a play an important role in mediating inflammatory diseases such as asthma and sepsis (Ali and Panettieri, 2005; Drouin, *et al*., 2001).

C3a receptor binds to the C-terminus of C3a. It has been reported that a synthetic C3a (70-77) peptide retains 1-2% of the activity of the natural human C3a (Caporale, *et al*., 1980). Thus, the disruption of the C-terminal active site of C3a could decrease its biological activity by interfering with receptor binding. (Humbles, *et al*., 2000; Ali and Panettieri, 2005). An engineered protease which specifically degrades C3a, could be useful in attenuating the inflammatory response in I/R injury thus minimize the tissue damage. Thus, the engineered C3a specific proteases could have wide therapeutic potentials in the treatment of shock, stroke, myocardial infarction, vascular diseases and trauma (Arumugam *et al.*, 2004).

The complement system has been extensively investigated recent years. There are many complement-related drugs are under clinical or preclinical investigation. For example the complement related protease inhibitor (C1-INH) is currently used in the treatment of hereditary angioedema (HAE) which is associated with C1-INH deficiency (Ricklin and Lambris, 2007). Another complement related therapeutic agent is Eculizumab (Solaris, Alexion Pharmaceuticals). Solaris is monoclonal antibody of complement protein C5 and it prevents C5 from breaking down in to C5a and C5b, thus prevents the activation of downstream complement cascade. It is marked for the treatment of paroxysmal nocturnal hemoglobinuria (PNH).

As part of this study, we tried to engineer a novel OmpT variant which degrades complement C3a specifically. Though *E. coli* OmpT might not be a direct therapeutic agent used in clinical applications, it could set up a platform for engineering of proteases degrading and regulating components of complement system.

3.2 Material and Method

3.2.1 Preparation of FRET substrates

The peptides Ac-CHARASHKGRGR-NH_2 and Ac-KYRRVGCGRGR-NH_2 were synthesized by GenScript (Piscataway, NJ). BODIPY^{TM}-FL-SE, Qsy7-maleimide and Qsy21-SE were purchased from Invitrogen (Carlsbad, CA). Atto-633-maleimide was purchased from ATTO-TEC (Siegen, Germany). Qsy7-maleimide and BODIPY^{TM}-FL-SE were conjugated to Ac-CHARASHKGRGR-NH_2 at the Cys and Lys residues,

respectively, to generate FRET selection C3a substrate **1**. Qsy21-SE and Atto-633-maleimide were conjugated to Ac-KYRRVGCGRGR-NH_2 at the Lys and Cys residues, respectively, to generate FRET counter selection substrate **2** according to literature procedures (Varadarajan, *et al*., 2005). The substrate purity and identity was confirmed by HPLC and ESI-MS.

Selection C3a substrate **1**: Ac-CHARASHKGRGR-$CONH_2$ (Qsy7, BODIPY™)

Counterselection substrate **2**: Ac-KYRRVGCGRGR-$CONH_2$ (Qsy21, Atto633)

Figure 3.1 The sequence of conjugated FRET substrates for C3a library screening and single clone activity assay.

3.2.2 Library construction

Preliminary results showed that the RV-OmpT variant (Varadarajan, *et al*., 2008) could hydrolyze C3a peptide at two positions both C terminal and N terminal. Therefore the RV-OmpT variant was subjected to random mutagenesis by error prone PCR. The error-rate was targeted to be 1% by using 0.5 mM of Mn^{2+} in the buffer mix. The genes were amplified using outside primers 5' CCGGGAATTCACCATGCGGGCGAAACTTCTGGGAATAGTC 3' and 5' AACAGCCAAGCTTTTAAAATGTGTACTTAAGACCAGCAGT 3' using described protocol (Chapter, Arnold F.H., and Georgiou G., 2003)

5’ EcoRI and 3’ HindIII restriction sites flanking the RV-OmpT gene were used for unidirectional cloning. The error-prone PCR products corresponding to the full size of RV-OmpT gene were gel-purified using the Qiagen gel extraction kit (CA, USA), digested with EcoR1 and HindIII, then ligated into a similarly digested pMLE19 vector. Ligated products were used to transform *E. coli* MC1061 (F^- Δ(ara-leu)7697 gal E15 galK16 Δ(lac)X74 rpsl(strR) hsdR2(r^-m^+) mcrA mcr B1; Casabadan and Cohen, 1980) and the cells were incubated in SOC at 37^{o}C for 1 hr. Subsequently the cells were spread on LB plates plus 1% glucose containing 200 μg/ml of ampicillin for 8 to 10 hours. Then, the plated cells were collected and sub-cultured in LB medium with 1% glucose containing 200 μg/ml of ampicillin at 37^{o}C either for 8 hours or overnight. The cells were pelleted by centrifugation and plasmid DNA was purified using the Qiagen miniprep kit. The plasmid DNA library was stored -80^{o}C After four rounds of sorting of the error-prone PCR library, clones 1-8 (Table 3.1) were isolated. A secondary error-prone PCR library was constructed using variants 1-5 as templates. The error-rate was targeted to be 1% by using 0.5 mM of Mn^{2+}.

3.2.3 Library screening

The plasmid library was transformed into *E.coli* BL21(DE3) (F^- *ompT* $hsdS_B$(r_B-m_B-) *gal dcm* (DE3)) for expression and sorting. The cell suspension was prepared as previously described (Varadarajan, *et al*., 2009). A 1 μl of C3a selection substrate **1** and counter-selection substrate **2** was added into 973 μl of 1% sucrose and 25 μl of cell

suspension. The reaction was incubated at room temperature in the dark for 9 min. A 900 µl reaction mixed with 100 µl of 1% sucrose and analyzed on FACS-Aria. The sorting gate was set according to the fluorescence signal and about 0.5% of the cells were collected in 200 µl 2 x YT. The collected cells were plated on LB plates 200 µg/ml and grown at 37^{o}C for 8 to 10 hours. The cells were scraped, sub-cultured and resorted. After 3-4 rounds of screening, plated colonies were analyzed as single colonies by FACS then analyzed by DNA sequencing. The screening scheme is the same as previously described in Chapter 2 (shown in Figure 2.1).

Single colonies were inoculated in 3 ml of 2xYT medium containing 200 µg/ml of ampicillin. The cells were cultured at 37^{o}C for 8 to 10 hours until an OD_{600} approximately 2 units. Then 1ml of the cells was centrifuged at 10,000 rpm for 2 minutes and resuspended in 1ml of 1% sucrose. The centrifugation step was repeated and the pellet was finally resuspended in 1ml of 1% sucrose. For activity assays, 25 µl of cells were labeled with 10 nM of either one FRET substrate or with both **1** and **2** and 1% sucrose was added to make a 1 ml reaction mixture (shown in Figure 4.1). 900 µl of the reaction was diluted with another 100 µl of 1% sucrose. The 1ml of the reaction mixture was analyzed on BD FACS-Aria.

3.2.4 Enzyme purification and C3a protein cleavage analysis

RV-OmpT and its variants were purified according to published procedures (Mangel, *et al*., 1994). The purity of the enzyme was over 90% in each case as

determined by SDS-PAGE. C3a protein substrates were purchased from CALBIOCHEM (Darmstadt, Germany) and the purity was >98% pure based on SDS-PAGE. For the C3a cleavage assay, 3 to 5 μg of C3a protein was incubated with 200 nM to 1 μM of the purified enzyme at 37^{o}C for 2 hours to overnight. The cleavage assay was carried out in 50 mM MES (morpholinoethanesulfonic acid), pH 6.2. , containing 10 mM EDTA. The reaction was stopped by adding 1% Trifluoroacetic acid (TFA) at a 1:1 ratio and was frozen in liquid nitrogen. Upon thawing the reaction products were analyzed by gel electrophoresis on 16% SDS PAGE (Invitrogen). LC-MS peptide or protein mass analysis was performed by Dr. Herng-Hsiang Stony Lo in the University of Texas, ICMB/CRED Protein and Metabolite Analysis Facility. An electrospray ion trap mass spectrometer (LCQ, ThermoFisher, San Jose, CA) coupled with a microbore HPLC (Magic 2002, Michrom BioResources, Auburn, CA) was used to acquire spectra of peptides or proteins. Automated acquisition of full scan MS spectra was executed by Finnigan Excalibur™ software (ThermoFisher, San Jose, CA). The full scan range for MS was 400-2000 Da. The acquired convoluted protein spectra from LCQ were deconvoluted by the ProMass for Xcalibur version 2.5.0 software (Novatia LLC, Princeton, NJ) to afford the MH+ m/z value(s) of the protein sample. For the peptide samples, zoom scans of the ions of interest were also performed to determine the charge state of the ions, consequently the molecular weight of the peptides.

3.3 Results and discussion

In Chapter 2, we successfully engineered OmpT P2 specificity using a high-throughput screening scheme (Figure 2.2). In this chapter attempted to use a similar strategy to engineer OmpT into a specific protease for the cleavage of C3a. To determine the C3a cleavage sites generated by WT OmpT or RV-OmpT the reaction products were analyzed by LC-MS (ESI). Figure 3.10 showed the intact C3a protein with MW of 9089. Figure 3.11 and 3.12 showed that the major cleavage product of both WT OmpT and RV-OmpT were around 8284 and 824. These results confirmed the primary cleavage site of WT and RV-OmpT was between C terminal residue 69 and 70 (R ↓ A). This is consistent with our screening scheme. Figure 3.11A and B also showed that WT OmpT might have a secondary cleavage site between N-terminal residue 7 and 8 (K ↓ R). The cleavage sites on C3a were labeled with arrow in Figure 3.2.

Based on these preliminary results, we chose the RV-OmpT as the model enzyme to be engineered to hydrolyze C3a specifically. The gene for the RV-OmpT protease was subjected to random mutagenesis by error prone PCR with an error rate of 0.6%. A library of 10^7 transformants was obtained.

Figure 3.1 shows the sequence of FRET substrates that were used for library screening. The C3a sequence 67-72 was included in the selection substrate **1**. The RR and RV sequence which are preferred by WT OmpT and the RV-OmpT variant respectively were included in the counter-selection substrate **2**.

After 4 rounds of sorting, we isolated five different variants (shown in Table 3.1). Variants 6-8 are exactly the same sequences as RV-OmpT. All 8 variants contain the RV mutation Q63R, D97H, S223D. However, we did not find consensus mutation on other positions. Selected single clone activity was shown in Figure 3.3 to 3.6. Among the novel variants that we isolated, none of them showed significantly increase specificity for C3a than the parent enzyme-RV variant.

In order to increase the specificity for C3a sequence, a secondary error PCR library was constructed using clone 1-5 as templates. The final library size is around 10^7 with a 0.5% error rate estimated from sequencing results. This 2^{nd} library was screened in the same way as the 1^{st} library above. After 3 rounds of sorting, clones 9 and 10 (shown in Table 3.1) were isolated. Single clone activity on FACS is shown in Figures 3.7 and 3.8. As can be seen from the fluorescence histograms in Figure 3.7, clone showed similar activity as RV-OmpT for the C3a selection substrate **1** but lower activity for counter selection substrate **2**. In Figure 3.8, clone10 showed decreased activity for both substrates. Because of the high cost of purified C3a protein, we selected clone 9 for a detailed analysis. WT OmpT, RV-OmpT and clone #9 proteins were purified and incubated with C3a at 37^{o}C for 4 hour or overnight and the reaction products were by SDS PAGE. As can be seen from Figure 3.9 RV-OmpT cleaved C3a protein most efficiently. 200nM of RV-OmpT incubated with 4 μg C3a protein for 4 hour, gave more than 30% cleavage of C3a was cleaved. Unfortunately #9 showed a low cleavage efficiency that was similar to that of the WT OmpT. .

Totally we isolated 10 novel variants by directed evolution (shown in Table 3.1). However we could not find consensus sequence among these mutations. Single clone of these variants was analyzed on FACs (Figure 3.3-3.8). Comparing to RV variant, the isolated clones 1-5, 9, 10 displayed relatively higher preference for selection substrate **1** over counter-selection substrate **2**. Unfortunately, our best clone #9 showed significantly lower efficiency in C3a protein cleavage comparing parent enzyme RV-OmpT. It might be due to the structural difference between selection FRET peptide and the C3a protein. In the future, if we could discover novel screening system utilizing protein substrates, we might isolate better variants by directed evolution with higher efficiency in protein cleavage.

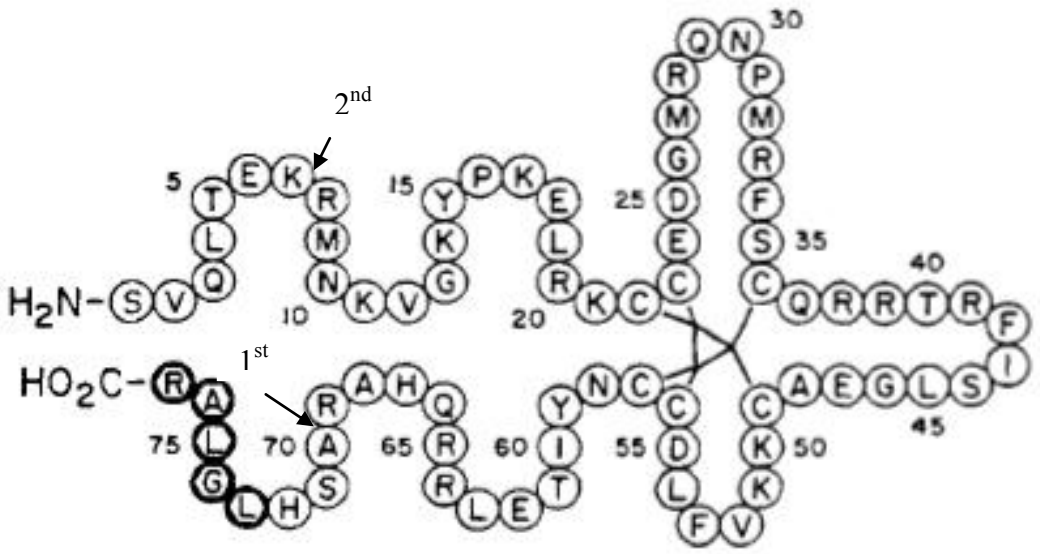


Figure 3.2 C3a sequence (adapted from Caporale *et al.,* 1980). The 1st arrow indicated the primary cleavage site of both WT and RV-OmpT. The 2nd arrow indicated the secondary cleavage site of WT OmpT.

Table 3.1 Sequences of variants that were isolated from sorting of error-prone PCR library.

Variants	Mutations
RV	Q63R, D97H, S223D
1	Q63R, D97H, S223D, N47S, A69T, N91Y, S140G, F151L, G196S, A245T
2	Q63R, D97H, S223D, E34K, G261S, S267N
3	Q63R, D97H, S223D, T20A, T25I, N80Y, S90C, D158E
4	Q63R, D97H, S223D, E3K, D113G
5	Q63R, D97H, S223D, G35R
6-8	Q63R, D97H, S223D
9	Q63R, D97H, S223D, I16V, G161D, F192Y, S207P, H268Y
10	Q63R, D97H, S223D, L31V, E34K, E153K, S222K, S267N, D274R

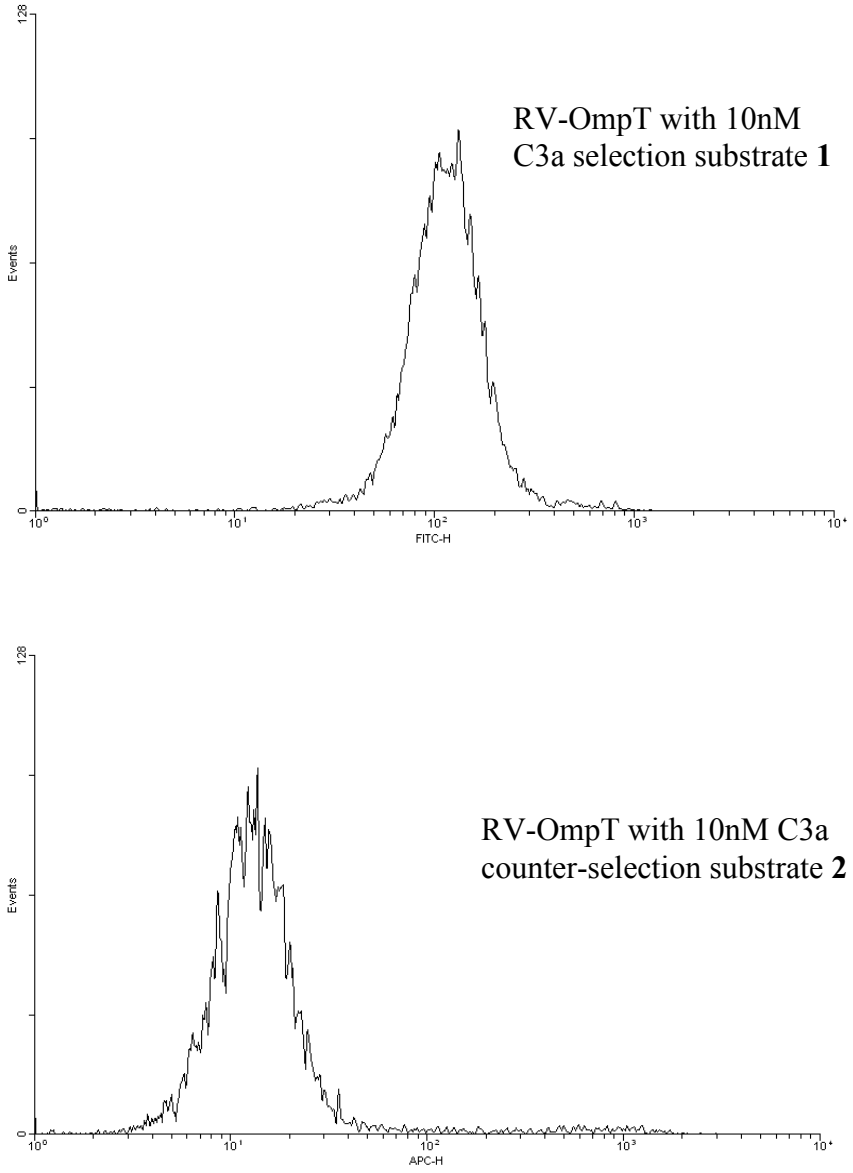


Figure 3.3 Fluorescence histogram of single clone of RV-OmpT incubated with 10nM of C3a selection substrate **1** and counter-selection substrate **2**.

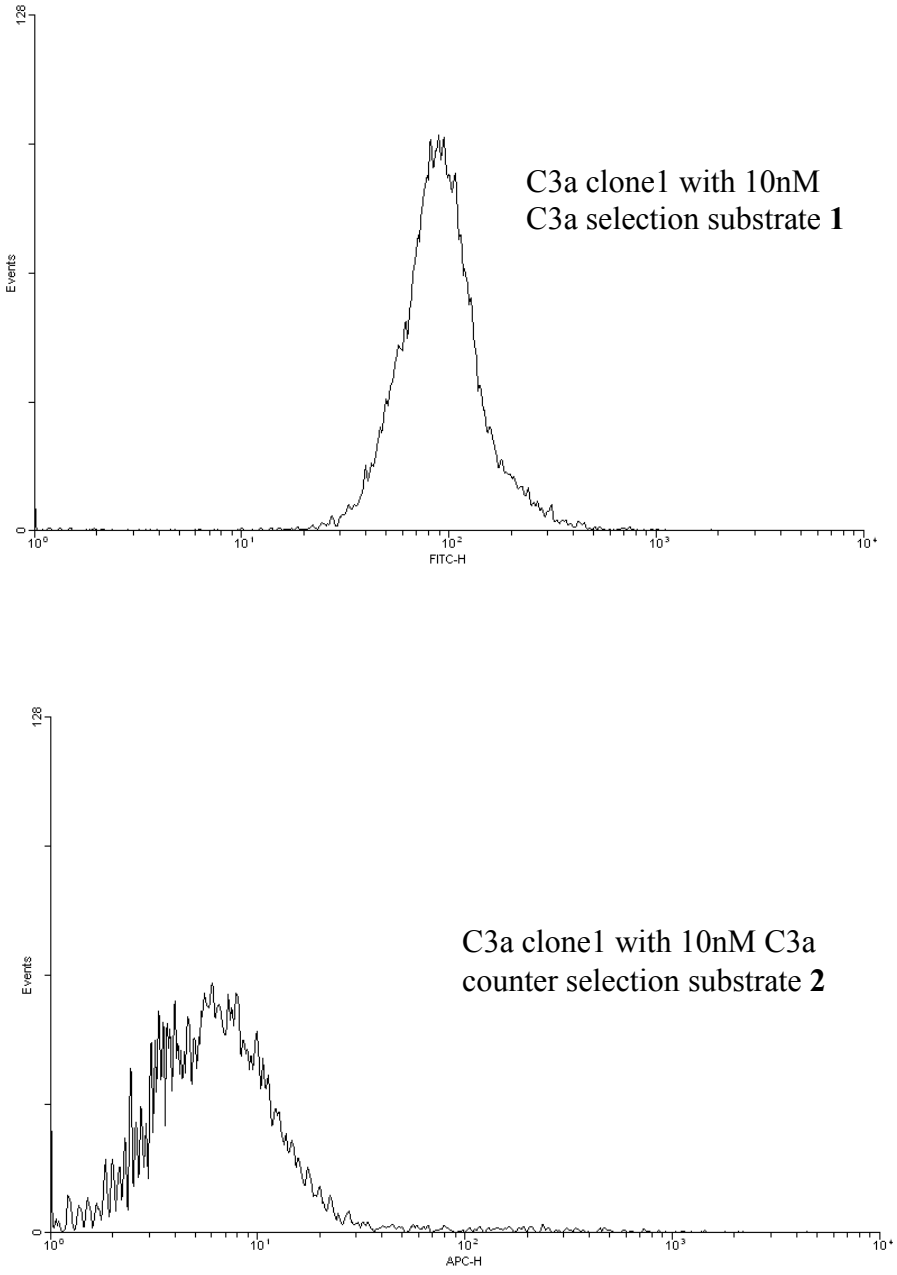
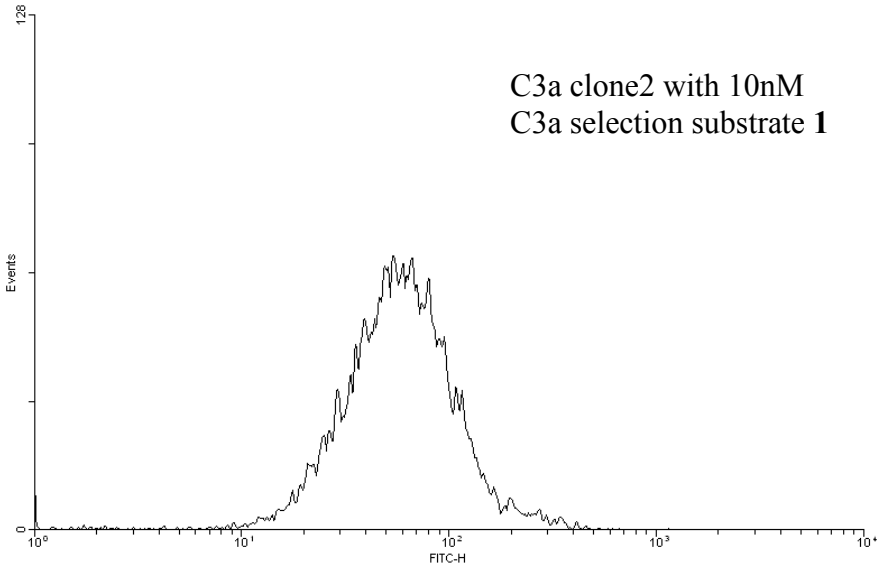


Figure 3.4 Fluorescence histogram of single clone of C3a clone1incubated with 10nM of C3a selection substrate **1** and counter-selection substrate **2**. The fluorescence of hydrolysis substrate **1** is detected under FITC and substrate **2** is detected under APC.



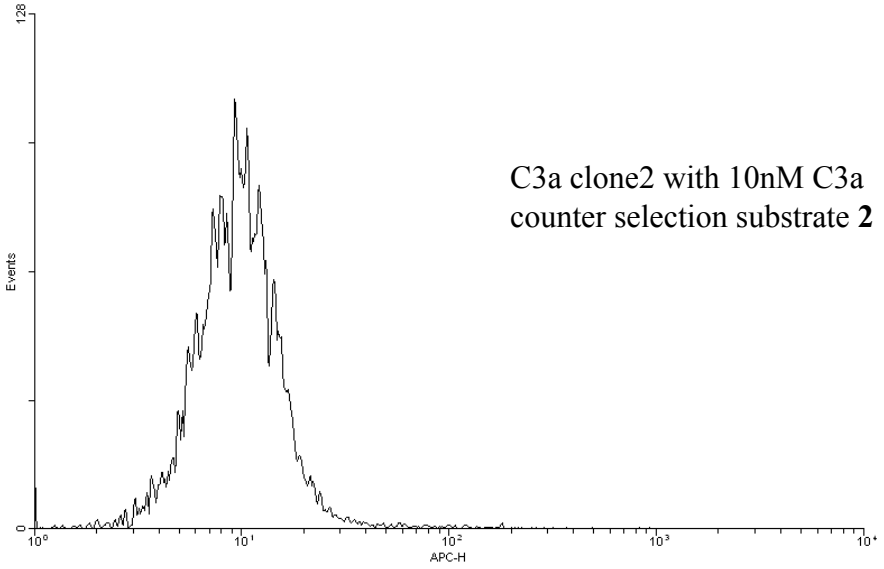


Figure 3.5 Fluorescence histogram of single clone of C3a clone2incubated with 10nM of C3a selection substrate **1** and counter-selection substrate **2**.

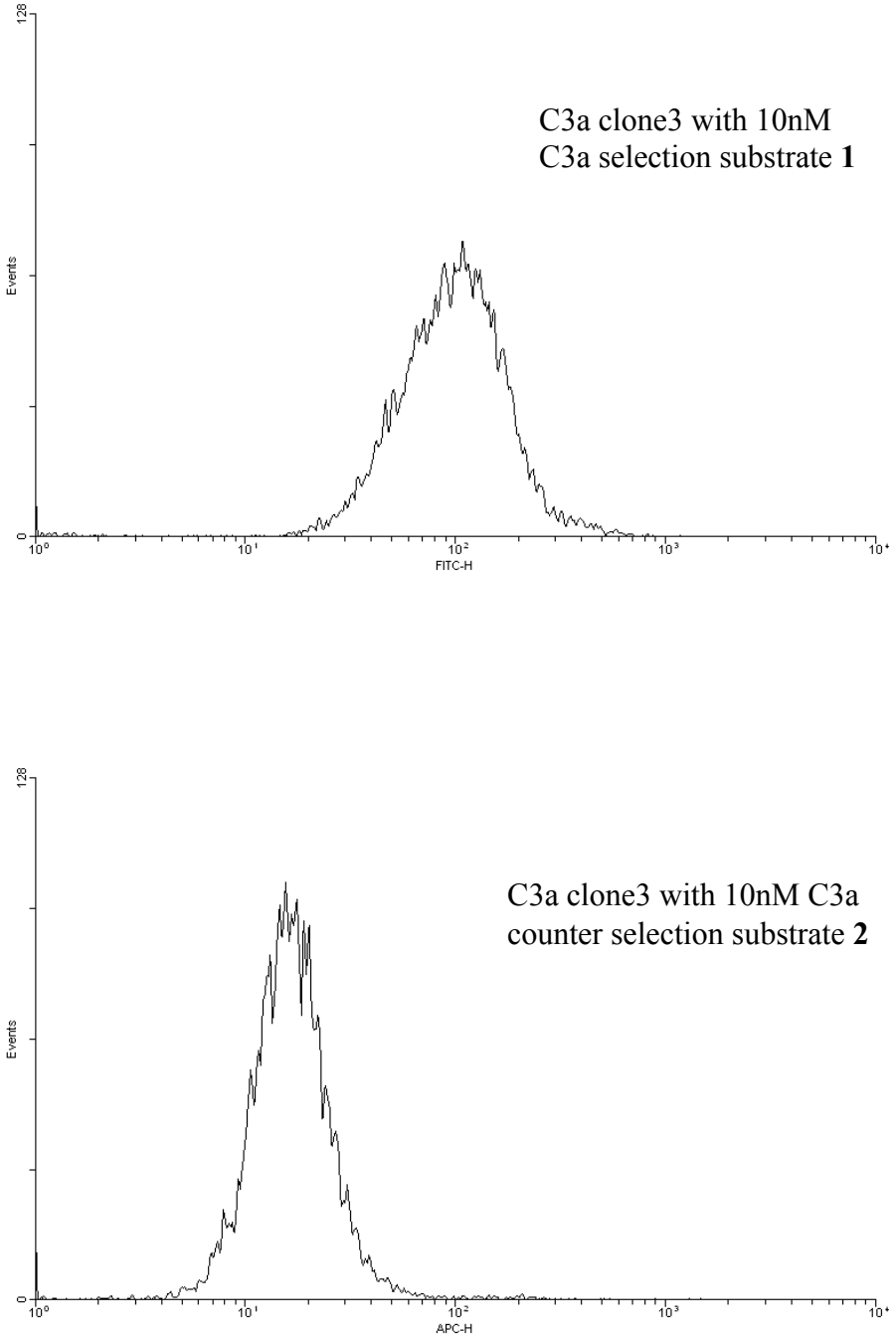


Figure 3.6 Fluorescence histogram of single clone of C3a clone3 incubated with 10nM of C3a selection substrate **1** and counter-selection substrate **2**.

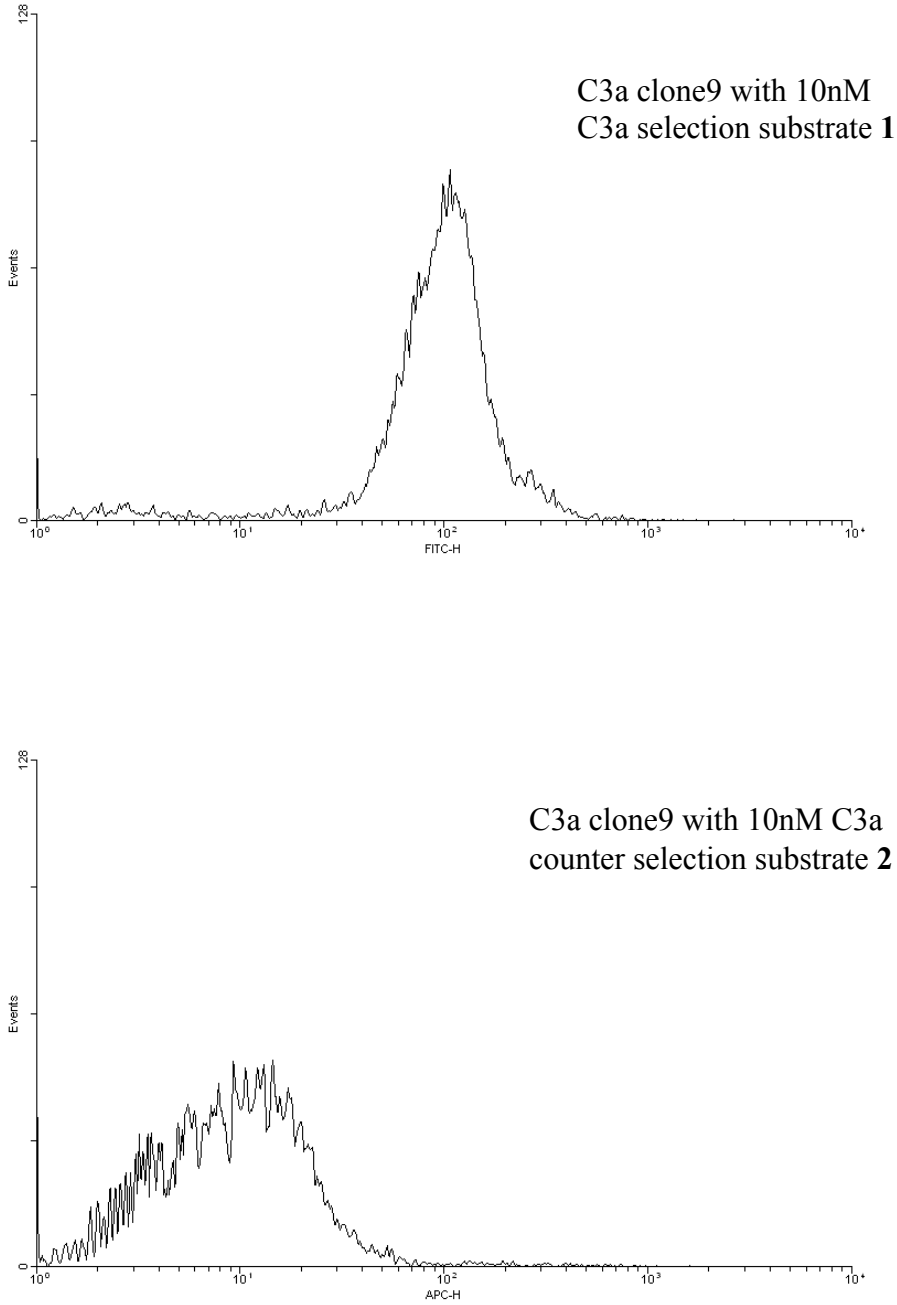
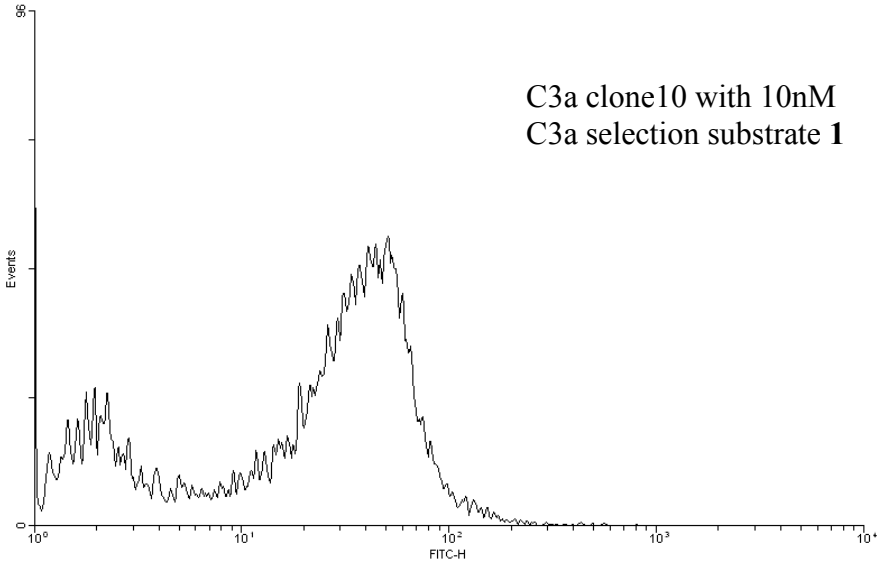
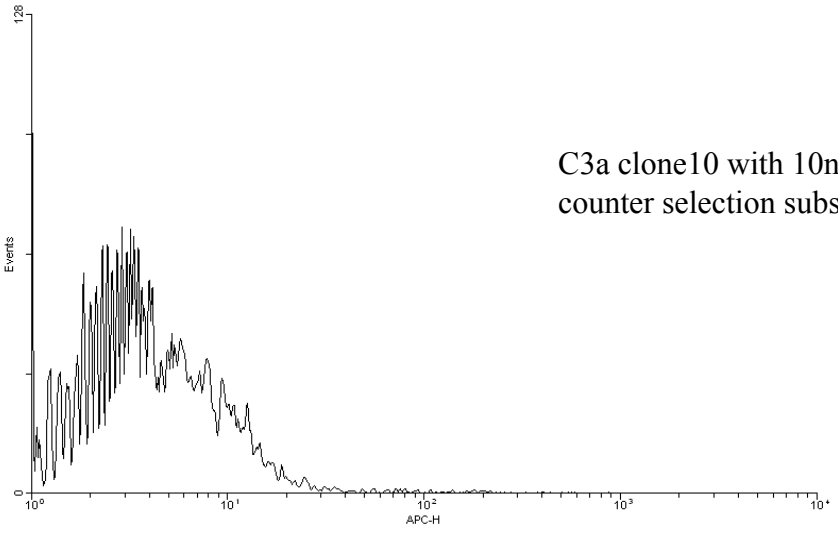


Figure 3.7 Fluorescence histogram of single clone of C3a clone9 incubated with 10nM of C3a selection substrate **1** and counter-selection substrate **2**.



C3a clone10 with 10nM
C3a selection substrate **1**



C3a clone10 with 10nM C3a
counter selection substrate **2**

Figure 3.8 Fluorescence histogram of single clone of C3a clone10 incubated with 10nM of C3a selection substrate **1** and counter-selection substrate **2**.

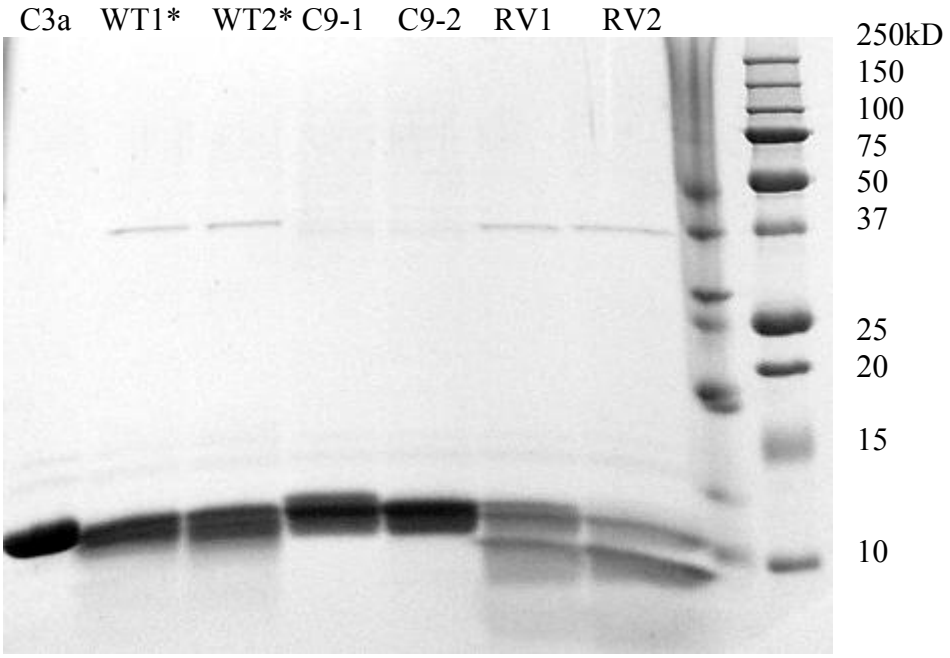


Figure 3.9 C3a protein cleavage assay. * 1 in the above figure indicates reaction incubation time as 4 hours. *2 indicates reaction incubation time as overnight. C9 refers to C3a variant 9 shown in Table 3.1. In all reactions 4 μg of C3a protein incubated with 200 nM of purified enzyme for either 4 hour or overnight at 37^{o}C.

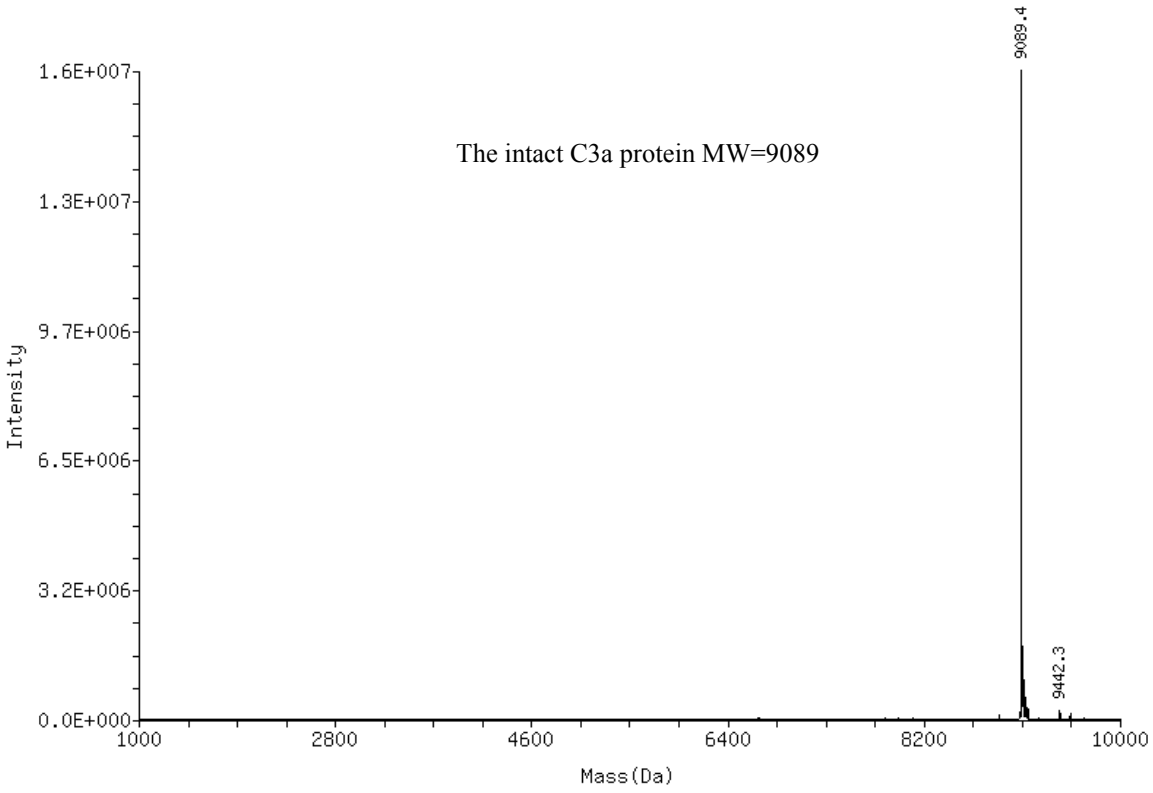


Figure 3.10 Mass spectrum of C3a protein. The molecular weight of C3a protein is detected as 9089.

A.

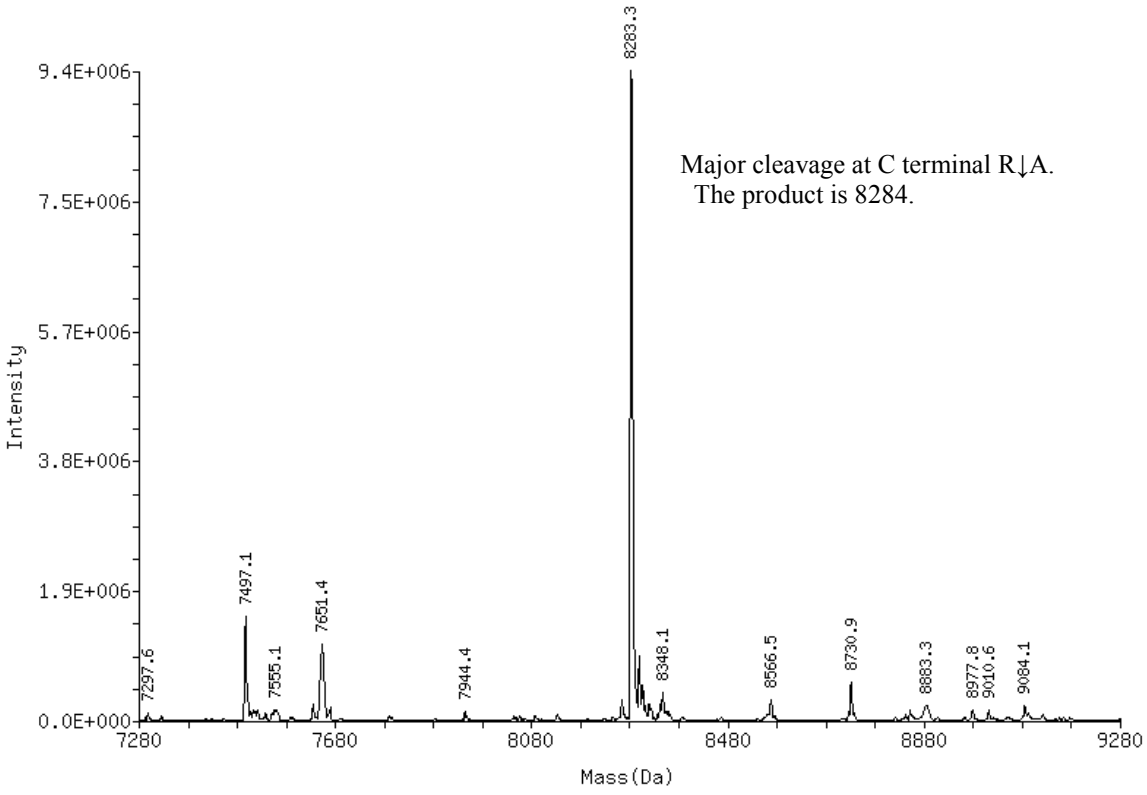


Figure 3.11A Mass spectrum of cleavage product of C3a by WT OmpT. The major cleavage product is detected as 8284 which is consistent with cleavage site between C terminal residue 69 and 70 (R↓A).

B.

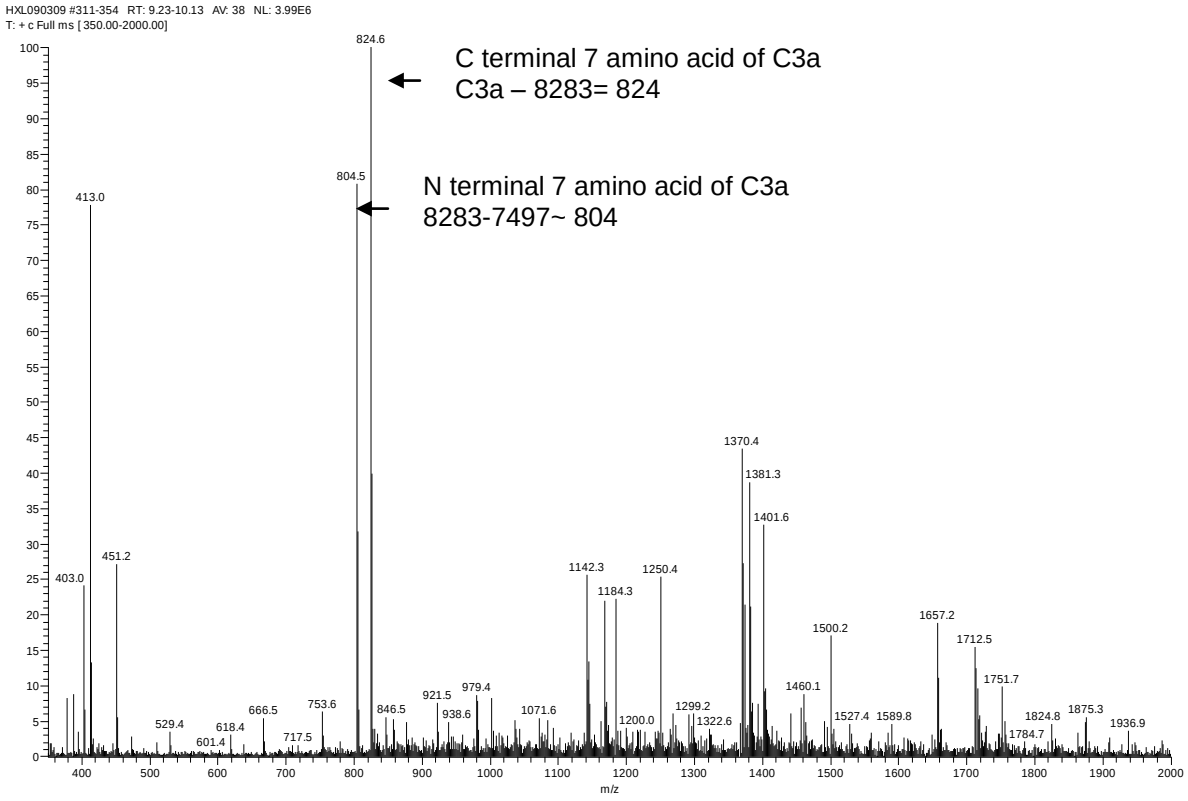


Figure 3.11 B Mass spectrum of cleavage product of C3a by WT OmpT. The cleavage product is detected as 824 which is consistent with the MW of C terminal 7 amino acids. Another minor piece of product is detected as 804 which is consistent with MW of N terminal 7 amino acid.

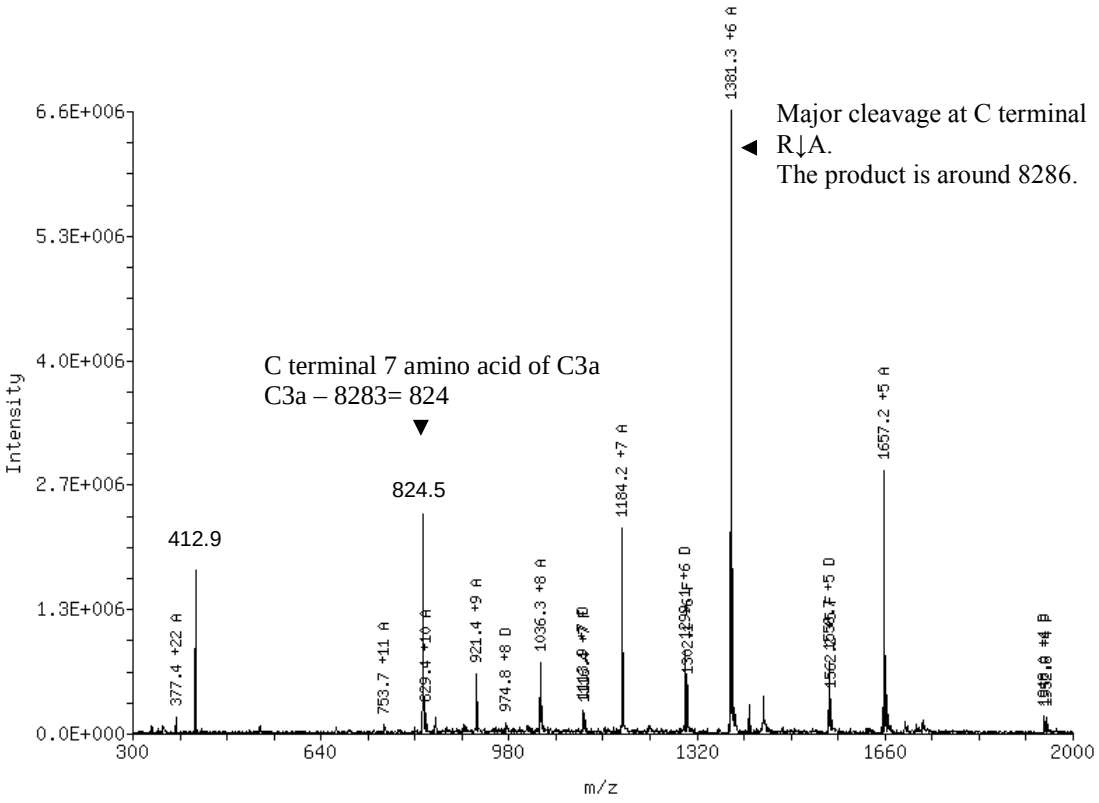


Figure 3.12 Mass spectrum of cleavage product of C3a by RV-OmpT. The major cleavage product is detected as 8286 which is consistent with cleavage site between C terminal residue 69 and 70 (R↓A). The small piece of cleavage product is detected as 824 which is consistent with the MW of C terminal 7 amino acids.

3.4 Reference

Ali, H. and Panettieri, R.A. (2005) “Anaphylatoxin C3a receptors in asthma”. Respir. Res. 21: 6-19

Arnold F.H., and Georgiou G. (2003) “Directed Evolution Library Creation. Methods and protocols” 75-90

Arumugam, T.V., Shiels, I.A, Woodruff, T.M., Granger, D.N., Taylor, S.M. (2004). "The role of the complement system in ischemia-reperfusion injury". Shock 21 (5): 401.

Bernstein J.A. (2008) "Hereditary angioedema: a current state-of-the-art review, VIII: current status of emerging therapies". Ann. Allergy Asthma Immunol. 100 (1 Suppl 2): S41–6.

Caroll, M.C. (2004) “The complement system in regulation of adaptive immunity” Nat. Immuno. 5(10): 981-986

Caporale, L.H., Tippett, P.S. and Erickson, B.W. (1980) “The active site of C3a anaphylatoxin” J. Biol. Chem. 255: 10758-10763

Casabadan, M.J. and Cohen, S.N. (1980) “Analysis gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138; 179-207

Drouin SM, Kildsgaard J, Haviland J, Zabner J, Jia HP, McCray PB Jr, Tack BF, Wetsel RA. (2001) “Expression of the complement anaphylatoxin C3a and C5a receptors on bronchial epithelial and smooth muscle cells in models of sepsis and asthma”, J Immunol. 166(3):2025-32

Dunkelberger, J.R. and Song, W.C. (2010). “Complement and its role in innate and adaptive immune responses” Cell Res. 20:34-50

Ember, J.A., Kildsgaard, J., Haviland, D.L., Lambris J.D. and Holers, M.V. (2000) “Therapeutic interventions in the complement system” Humana press, Totowa, NJ. pp 113-153

Haas, P.J. and Strijp, J.V. (2007) “Anaphylatoxins: their role in bacterial infection and inflammation” Immun. Res. 37/3:161-175

Hellerud, B.C. et al (2010) “Critical roles of complement and antibodies in host defense mechanisms against Neisseria meningitidis as revealed by human complement genetic deficiencies” Infection and Immunity 78(2):802-809

Humbles, A.A., Lu, B., Nilsson, C.A., Lilly, C., Israel, E., Fujiwara, Y., Gerard, N.P. and Gerard, C. (2000) "A role for the C3a anaphylatoxin receptor in the effector phase of athma." Nature 406: 998-1001

Juliano, L. 1997. "Hydrolysis of somatostatin by human tissue kallikrein after the amino acid pair Phe-Phe." Biochem. J. 327: 27−30.

Johnson, R.J. and Hillmen, P.(2002) "Paroxysmal nocturnal haemoglobinuria: Nature's gene therapy?" Mol. Path. 55: 145-152

Lambris, J.D. and Morikis, D. (2005) "Structure biology of the complement system" ISBN: 0-8247-2540-9 page: 161-174

Maier, M., Peng, Y., Jiang, L., Seabrook, T.J., Carroll, M.C. and Lemere, C.A. (2008) "Complement C3 deficiency leads to accelerated amyloid β plaque deposition and neurodegeneration and modulation of the microglia/macrophage phenotype in amyloid precursor protein transgenic mice" J. Neurosci. 28(25): 6333-6341

Ricklin, D and Lambris, J.D. (2007) "Complement-targeted therapeutics" Nature Biotechnology 25, 1265 – 1275

Scarisbrick, I.A., Towner, M.D., and Isackson P.J. (1997) "Nervous system-specific expression of a novel serine protease: regulation in the adult rat spinal cord by excitotoxic injury." J. Neurosci. 17: 8156−8168.

Tedesco, F., Nurnberger, W., Perssutti, S. (1993) "Inherited deficiencies of the terminal complement components." Int. Rev. Immunol. 10: 51-64

Varadarajan, N., Gam, J., Olsen M.J., Georgiou, G., and Iverson, B.L. (2005) "Engineering of Protease Variants Exhibiting High Catalytic Activity and Exquisite Substrate Selectivity" Proc. Nat. Acad. Sci. 102, 6855-6860.

Varadarajan, N., Rodriguez, S., Georgiou, G., and Iverson, B.L. (2008) "Highly Active and Selective Endopeptidases with Programmed Substrate Specificities" Nat. Chem. Bio. 4:290-294

Varadarajan, N., Cantor, J.R., Georgiou, G., and Iverson, B.L. (2009) "Construction and flow cytometric screening of targeted enzyme libraries." Nat. Protoc. 4(6):893-901

Wetsel, Rick A.; Kildsgaard, Jens; Haviland, David L. (2000) “Complement anaphylatoxins (C3a , C4a, C5a) and their receptors (C3aR, C5aR/CD88) as therapeutic targets in inflammation” Therapeutic Interventions in the Complement System, 113-153.

Chapter 4 Investigation of the Substrate Specificity of Human Kallikrein 1 and 6

4.1 Introduction

The human tissue kallikrein locus contains 15 structurally similar genes. They represent the largest contiguous cluster of serine protease genes within the human genome. Human tissue kallikreins (KLKs) are expressed in a wide range of tissues and have been implicated in many physiological and pathological functions (Diamandis et al. 2000; Yousef and Diamandis. 2001; Paliouras and Diamandis 2006). KLK1 was the first kallikrein to be discovered and was detected at high levels in human pancreas. KLK1 together with KLK2 and KLK3 constitute the classical kallikreins. The primary physiological function of KLK1 is the cleavage of low molecular weight kininogen at Arg-Ser and Met-Lys bonds to release lysyl-bradykinin (kallidin) (Sueiras et al. 1994). Kallidin binding to its receptors B1 and B2 leads to a signaling cascade that regulates multiple physiological processes such as blood pressure, smooth muscle contraction and vascular cell growth, etc (Bhoola et al. 1992; Borgono et al. 2004). In addition to its primary function in the release of kallidin, KLK1 has also been shown to cleave kallistatin, somatostatin, pro-insulin, low density lipoprotein, the precursor of atrial natriuretic factor, prorenin, vasoactive intestinal peptide, procollagenase and

Substantial portions of this chapter were previously published as "Substrate specificity of human kallikreins 1 and 6 determined by phage display." Protein Science **17**: 664-672 Li, H.X., Hwang, B.Y., Laxmikanthan G., Blaber, S.I., Blaber, M., Golubkov, P.A., Ren, P., Iverson, B.L. and Georgiou, G. (2008)

angiotensinogen (Bhoola et al. 1992; Yousef and Diamandis 2001). It has been suggested that KLK1 exhibits different physiological functions in different tissues but it is not known how its activity is regulated. The enzyme was shown to exhibit both trypsin- and chymotrypsin-like activities. Its trypsin-like activity is manifest in the cleavage of low molecular weight kininogen at Arg-Ser (Fiedler et al. 1979) while the chymotrypsin-like activity is evident in the cleavage of kallistatin (human kallikrein binding protein) and somatostatin, after a Phe-Phe dipeptide (Zhou et al. 1992; Pimenta et al. 1997).

KLK6 was called zyme, protease M, neurosin or PRSS9 prior to its current nomenclature (Little et al. 1997; Yamashiro et al. 1997; Yousef et al. 1999). Though the physiological function of KLK6 is not clear, KLK6 was reported as myelencephalon-specific protease which may involve in the regulation of myelin turnover and in demyelinating disease such as multiple sclerosis (Scarisbrick et al. 1997). Many proteins have been shown to be cleaved by KLK6 *in vitro* including human myelin basic protein (MBP), Aβ amyloid peptide, plasminogen, myelin and alpha-synuclein. Notably, KLK6 has been implicated in neurodegenerative diseases such as Alzheimer's, multiple sclerosis and Parkinson's (Bernett et al. 2002; Iwata et al. 2003; Magklara et al. 2003). Furthermore, the KLK6 gene has been shown to be up-regulated in ovarian cancer and therefore KLK6 constitutes a potential serum biomarker for diagnosis and prognosis of ovarian cancer (Diamandis et al. 2003). Also, its expression is up-regulated in primary breast tumors and down-regulated at metastatic breast cancer sites. Since KLK6 has been shown to cleave extracellular matrix peptides derived from laminin and fibronectin, it

may be involved in tissue remodeling (Anisowicz et al. 1996; Yousef et al. 2003; Ghosh et al. 2004). Finally, in a recent study, Oikonomopoulou et al. reported that KLK6 cleaves protease-activated receptors (PARs), a family of G coupled protein receptors that are activated via an N-terminal cleavage proteolytic event (Oikonomopoulou et al. 2006).

The biochemical properties and substrate specificity of KLK6 have attracted significant attention. Crystal structures for both proteins KLK1 and 6 are available (Bernett et al. 2002; Laxmikanthan et al. 2005). Even though the physiological role of KLK1 has been studied in detail, relatively little is known about its extended substrate specificity. Fluorogenic peptides derived from a preferred KLK1 substrate indicated that Arg, Lys. Phe, Tyr and Met are accepted in the S1 subsite, Ser is strongly preferred in S1' which can also accommodate Arg to a lesser extent, and finally, Phe and Leu are preferred in S2 (Pimenta et al. 1999). The substrate specificity of KLK6 has been investigated by positional scanning using two different approaches: Debela et al. employed peptide substrates of the general structure acetyl-P4-P3-P2-P1-7-aminomethylcoumarin and reported that KLK6 exhibits strong preference for Arg or Lys in S2 whereas the S1 subsite accepts not only Arg but also Lys, Ala or Met (Debela et al. 2006). These results were in conflict with the data of Angelo et al. who examined the substrate preference of KLK6 using the same set of fluorescence resonance energy transfer (FRET) peptides that had been used for the analysis of the specificity of KLK1 mentioned above. Specially Angelo et al. reported that only Arg is accepted in the S1

subsite of KLK6 whereas S2 strongly prefers Phe and, to a lesser extent Leu, over all other amino acids (Angelo et al. 2006).

The determination of protease specificity using positional libraries of FRET or fluorogenic peptide substrates suffers from two drawbacks. First, interactions between the fluorescent dyes and the enzymes can bias substrate preference. This can be particularly true when the dye occupies the P1' position as is the case the with 7-aminomethylcoumarin substrates used by Debela et al. Second, positional scanning methods cannot take into account neighboring effects whereby the occupancy of one subsite by a particular amino acid affects the preference of the adjacent subsites.

Substrate phage display provides a means for discerning the extended amino acid specificity of proteases without the need for chemical labeling (Deperthes 2002; Diamond 2007). It relies on the selective cleavage of specific peptide sequences sandwiched between the gene pIII minor coat protein of fd bacteriophage and an affinity tag. The phage is immobilized on a solid support via the affinity tag and, following treatment with a purified protease of interest, clones containing susceptible peptide sequences are cleaved, which in turn releases them from the support, allowing amplification. Consensus substrate sequences are isolated after several rounds of panning. The substrate specificity of over 20 bacterial and mammalian proteases including KLK2, KLK3 and KLK14 have been analyzed by substrate phage (Cloutier et al. 2002; Felber et al. 2005; Ferrieu-Weisbuch et al. 2006). In an effort to address discrepancies in earlier studies of KLK1 and KLK6 fine specificity, here we used substrate phage with a random

octapeptide library. Cleavage and enrichment were performed in buffers with a physiological pH and salt concentration. Selected peptides were synthesized, their respective cleavage sites were determined by MS and the kinetics of hydrolysis was determined. KLK1 showed its dual activity both trypsin-like and chymotrysin-like while KLK6 only showed trypsin-like activity with strong preference of Arg at P1 site. Both KLK1 and KLK6 cleaved the synthesized peptides efficiently which derived from phage panning sequences. Docking studies with selected peptides were employed to provide information on the KLK1 and KLK6 residues that form the S1 and S1' subsites and interact with side chains on the substrate.

4.2 Material and methods

4.2.1 Bacterial strains, plasmids

E. coli strain MC1061 was used for cloning (Casabadan and Cohen 1980) and K91BK was used for phage infection and amplification. The polyvalent phage display vector fUSE55 which is derived from fUSE5 was used for the construction of peptide libraries (Scott and Smith 1990). Unless otherwise stated, *E. coli* cells were grown in NZY medium for phage infection and propagation containing the appropriate antibiotics at 37^{o}C in a shaking incubator.

4.2.2 Protein purification

The mature recombinant KLK1 and KLK6 were expressed in a baculovirus/insect cell line system and purified as previous described (Laxmikanthan et al. 2005).

4.2.3 Construction of phage library

First, sequences encoding the FLAG tag (DYKDDDDK) and XhoI restriction site were inserted to the 5' of gene gIII by ligating a synthetic DNA fragment into the SfiI site of fUSE55, giving rise to fUSE55-FLAG. Subsequently, the 8-mer library was synthesized by PCR using fUSE55-FLAG DNA (100 pg) as template, the 5' primer, 5'-GATAAAGGACTCGAGGCTNNKNNKNNKNNKNNKNNKNNKGGGGCCGA AACTGTTGAAAG-3' (where N represents any nucleotide and K represents T/G), and the 3' primer, 5'-CAAACGAATGGATCCTCATT AAAGCCAG-3', followed by digestion with XhoI/BamHI and ligation into fUSE55-FLAG (restriction sites are underlined). The ligated DNA (~ 1 μg) was used to transform electrocompetent *E. coli* MC1061 (~ 100 μl) via electroporation. The library was sequenced to confirm the randomization. No bias was found within the library.

4.2.4 Library screening and phage substrate purification

10^{11} cfu (colony forming unit) of amplified substrate phage library was first immobilized onto 3 ml of anti-FLAG M2 Affinity Gel (Sigma, MO, US) at 4 ^{o}C for 4 h. The phage-bounded resin was washed extensively with TBS buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.5) to remove unbound phages, then incubated with 100 to 200 nM of

purified enzyme in TBS buffer containing 1 mM EDTA at 37 ^{o}C for 45 min. The eluted phage particles were subsequently amplified in *E. coli* K91BK, purified with polyethylene glycol (PEG) 8000 (Sigma, MO, US) and used as input in subsequent rounds of screening. From third to the last round of screening, phage clones were randomly selected and plasmid DNA were isolated and sequenced to determine the displayed substrate peptide sequences. A total of six rounds of enrichment were performed for KLK1 and five rounds for KLK6.

4.2.5 Western blotting

Selected phage clones were purified and incubated with 0 - 1 μM of KLK6 or 4 μM of KLK1 in 50 mM Tris-HCl, 150 mM NaCl buffer (pH 7.5) containing 1 mM EDTA at 37 ^{o}C for 45 min. The reaction mixture was separated with 4-20% gradient SDS-Polyacrylamide gel electrophoresis (PAGE) (Pierce, IL, US), and then analyzed with anti-FLAG antibody by western blotting.

4.2.6 Kinetic measurements

Peptides were synthesized by solid phase synthesis. (EZbiolab InC., Westfield, IN). Kinetic assays were carried out in 20 mM Tris-HCl buffer, containing 1 mM EDTA at pH 9.0 for KLK1 and 50 mM Tris-HCl, 150 mM NaCl containing 1 mM EDTA at pH 7.5 for KLK6, respectively. 1 to 200 μM of substrates were incubated with 0.125-100 nM of purified enzyme at 37 ^{o}C for 5 to 30 min. The reactions were

quenched by liquid nitrogen and were analyzed on a C18 reverse-phase column (Phenomenex, Torrance, CA) by high-performance liquid chromatography (HPLC) using the following gradient: 5% acetonitrile / 95% H_2O for 1 min, increasing to 95% acetonitrile / 5% H_2O for 29 min and returning to 5% acetonitrile / 95% H_2O for 5 min. The product amount was calculated upon the integration area at 280 nm and fitted to non-linear regression of Michaelis-Menten equation by Prism program (Graphpad, San Diego, CA).

4.2.7 Molecular modeling

Docking was performed using GOLD v3.1.1 (www.ccdc.cam.ac.uk) on TI-3D cluster with coordinates taken from human kallikrein 1 structure (PDB code 1SPJ) and human kallikrein 6 structure (PDB code 1L2E). For KLK 6, cavity was defined with a radius of 20 Å centered around the OH group of Ser195. Five peptides were tested, FRSA, FRSV, VRSA, VRSV, and FRFSQ, all capped with acetyl and n-methylamide groups. A total of 200 genetic algorithm (GA) simulations were performed on each peptide, with population size of 150 and $1x10^6$ operations. The following residues on target protein were defined with flexible side chains: His57, His99, Asp189, Ser190, Gln192, Ser195, Ser214, Trp215, Asn217, and Ile218, using standard rotamer libraries (Lovell et al. 2000). All figures were rendered in PyMOL (DeLano scientific, Palo Alto, CA; http://www.pymol.org). For KLK1, seven peptides were tested: FYRR, LYSR, ARSA, SRSA, KRSY, FYSQ, and FYSR. The following residues on the protein were

defined with flexible side chains: Gln41, His57, Tyr99, Asp102, Asp189, Thr190, Ser195, Ser214, Trp215, and Ser226. The rest of docking/GA parameters were identical to KLK6 docking simulations.

4.3 Results

4.3.1 Analysis of the specificity of KLK6

We first examined whether phage encoding a putative KLK6 sequence inserted between the pIII protein and the FLAG peptide epitope can be recognized and cleaved following incubation with protease. Phages containing the octapeptide sequence TAFRSAYG from the panning and SSYISNYG which contains no putative cleavage sites, were constructed, purified and were incubated with 0, 0.1 or 1 μM of KLK6 at 37^{o}C for 45min. Cleavage of the phage was then analyzed by western blotting using anti-FLAG and anti-pIII antibody. The TAFRSAYG sequence was efficiently cleaved under these conditions resulting in loss of the FLAG tag. In contrast, phage encoding SSYISNYG remained intact (Fig. 4.1). Thus, KLK6 does not cleave pIII nor does it affect the integrity of the phage but can recognize substrate sequences inserted between pIII and the FLAG epitope. Similarly KLK1 does not cleave either pIII or the intact phage (data not shown).

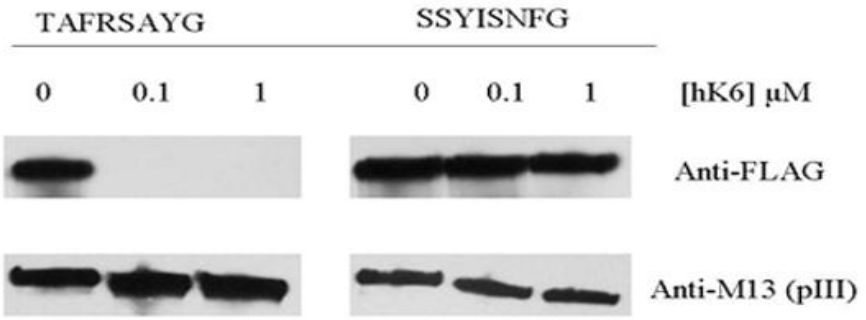


Figure 4.1 Cleavage of putative substrate phage and non-substrate phage by KLK6 by western blotting with anti-FLAG and anti-pIII. Purified phage was incubated with 0 μM, 0.1 μM, or 1 μM of KLK6 in 50 mM Tris-HCl, 150 mM NaCl containing 1 mM EDTA (pH 7.5) at 37 ^{o}C for 45 min and analyzed by western blotting. Phage containing the TAFRSAYG octapeptide is a substrate for KLK6 whereas phage containing the SSYISNFG sequence is not.

A random octapeptide library was constructed as described earlier (Hwang et al. 2007). The diversity of the library was $1.1x10^{8}$ cfus and therefore it encoded only a small fraction of the possible 8-mer amino acid sequences ($2.56x10^{10}$), nonetheless it should be sufficient to capture the key properties of amino acids preferred by the protease (Kerr et al. 2005). The phage library was immobilized on anti-FLAG conjugated beads and subjected to panning with purified recombinant KLK1 or KLK6. Cleaved phage clones were amplified in *E. coli* K91BK and used as input in subsequent panning experiments. After each round of panning, phage clones were picked at random and sequenced. The sequencing results of KLK6-panning revealed that 6 out of 12 clones in the third round contained sequences encoding an Arg-Ser dipeptide sequence while 10 out of 12 clones

in the fourth round contained Arg-Ser dipeptide consistent with the expected cleavage of this sequence. Individual isolated clones from the third to the fifth round were amplified; phage was purified, subjected to digestion by KLK6 and the loss of the FLAG epitope was examined by western blotting. For comparison we also constructed phage in which the random octapeptide in pfUSE55 was replaced with the sequence encoding AAFRFSQA. This peptide had been reported to exhibit the maximum rate of cleavage by KLK6 *in vitro* (Angelo et al. 2006). Six phage clones isolated after five rounds showed a greater degree of cleavage relative to phage containing the AAFRFSQA peptide (Fig. 4.2). All 20 isolated phage clones examined showed KLK6-specific cleavage. The selected sequences that were cleaved by KLK6 are shown in Table 4.1. Consistent with the expected cleavage preference of the enzyme, the sequences were aligned so that Arg is positioned in the P1 site. Fig. 4.3 shows a histogram of the observed amino acid frequencies between the P4 to P4' normalized relative to the codon usage in the respective NNK library position. Only four different amino acids were found in the P1' position with a strong preference for Ser (8-fold greater than the frequency of Ser that would be expected based on the NNK randomization scheme). We also observed an increased frequency of Val or Phe at the P2 position (4-fold and 9.6 fold, respectively) whereas Ala is strongly preferred at the P2' position (9.6-fold enrichment over background). KLK6 appears to accept a variety of amino acids at P3, P4, P3', and P4' positions. Residues derived from the linker sequence in the octapeptide library showed up

at P3, P4 or P3' and P4' and thus do not affect the significance of the analysis described above. All Gly on P4' are from linker which is not considered as significant.

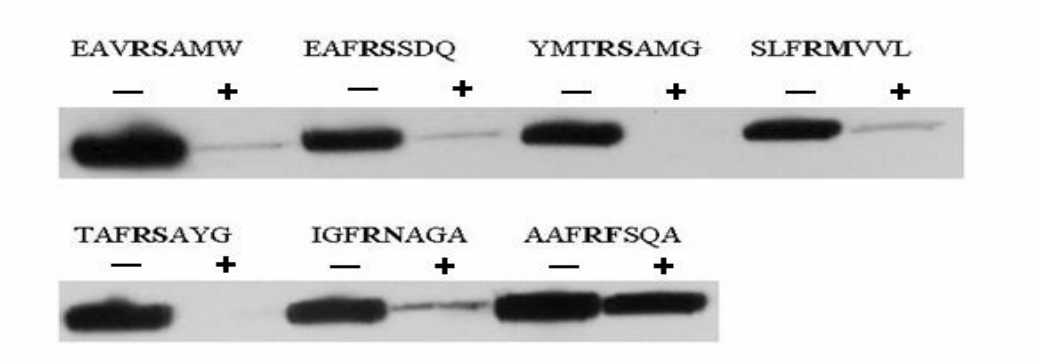
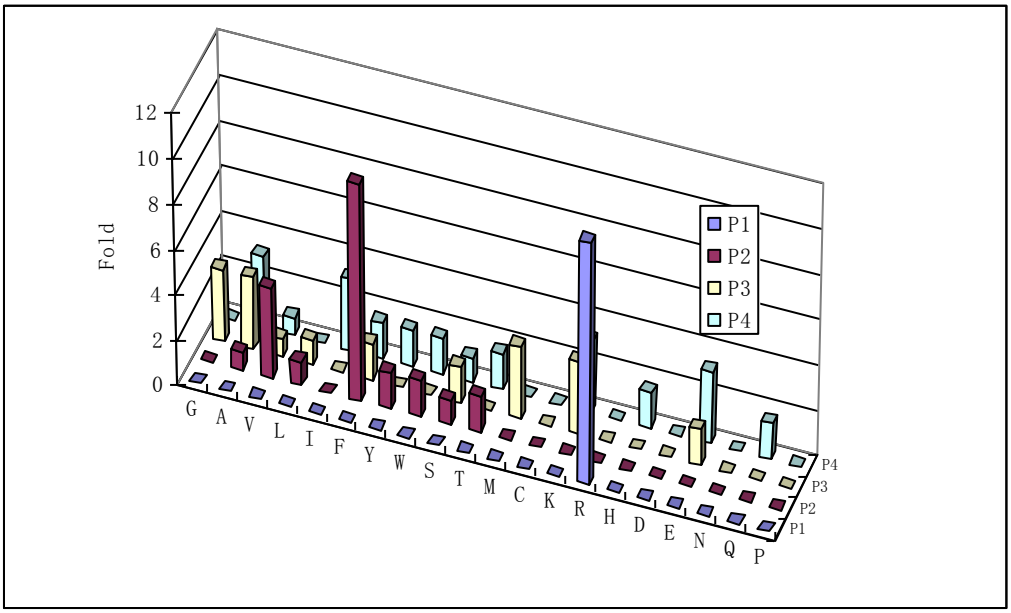


Figure 4.2 Cleavage of selected phage clones by KLK6. Purified phages were incubated with 0 or 0.1 μM KLK6 in 50 mM Tris-HCl, 150 mM NaCl containing 1 mM EDTA (pH 7.5) at 37 ^{o}C for 45 min. The reaction mixture was separated with 4-20% gradient SDS-PAGE and then analyzed with anti-FLAG antibody by western blotting. The label + or – indicates with or without enzyme. The octapeptide sequences were shown above the according band. AAFRFSQA has been reported to exhibit the maximum rate of cleavage by KLK6 *in vitro* (Angelo et al. 2006).

(A)



(B)

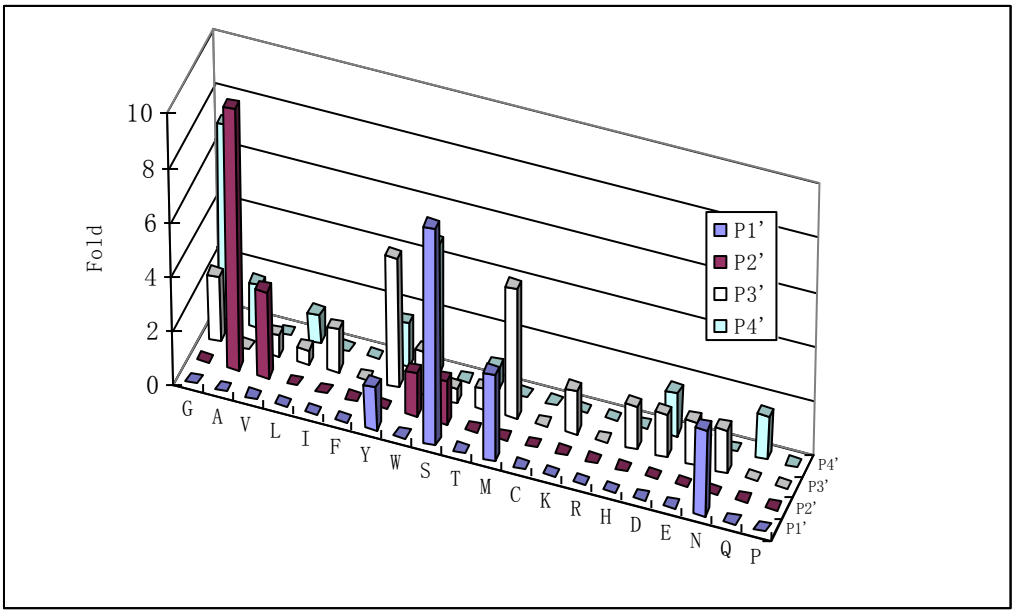


Figure 4.3. Normalized frequency of occurrence of amino acids at P1-P4 (A) and P1'-P4' (B) of KLK6 substrates. The data represent the observed amino acid frequencies between the P4 to P4' normalized relative to the codon usage in the respective NNK library position. Y-axis shows the ratio of the observed frequency to the theoretical frequency of the respective amino acids in NNK library.

Table 4.1 Alignment of sequences selected by substrate phage display with KLK6

Clone	P4	P3	P2	P1	P1’	P2’	P3’	P4’
4R-1	*E*	*A*	V	R	S	A	M	W
4R-11	H	L	V	R	S	W	N	*G*
5R-6	V	G	V	R	S	V	Y	*G*
5R-8	*A*	S	V	R	S	A	M	Y
4R-9	S	K	V	R	S	A	*G*	*A*
4R-8	Q	M	Y	R	S	S	W	*G*
Re5R-6	F	G	F	R	S	V	H	*G*
4R-22	T	A	F	R	S	A	Y	*G*
5R-17	T	A	F	R	N	S	L	*G*
5R-16	I	G	F	R	N	A	*G*	*A*
4R-13	S	L	F	R	M	V	V	L
4R-3	*E*	*A*	F	R	S	S	D	Q
4R-7	*A*	S	S	R	S	V	K	W
Re5R-14	*A*	K	S	R	S	A	G	D
5R-3	*A*	F	L	R	M	A	S	L
4R-17	K	V	L	R	S	A	T	*G*
4R-6	Y	M	T	R	S	A	M	*G*
4R-24	I	S	T	R	S	A	I	W
4R-20	W	G	W	R	Y	A	*E*	*T*
5R-11	K	E	A	R	S	A	Y	*G*

Amino acids from the linker region are shown in italic.

Table 4.2 Kinetic measurements of KLK6 catalyzed hydrolysis of substrate

Peptide no.	Amino acid sequence of peptide substrate	KLK6		
		k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
1	WYMTR↓SAMG	1.4 ±0.1	69± 7	$1.9 \pm 0.5 \times 10^4$
2	WIGFR↓NAGA	1.7 ± 0.7	182.6 ±87.4	$1.5 \pm 1.1 \times 10^4$
3	WTAFR↓SAYG	3.8 ± 1.1	336.3± 148	$1.6 \pm 1 \times 10^4$
4	WEAVR↓SAMW	2 ± 1.3	92.1 ± 85.1	$2.2 \pm 1.4 \times 10^4$
5	WEAFR↓SSDQ	1.2 ± 0.4	248.8 ± 86.4	$5 \pm 2.6 \times 10^3$
6	WAAFR↓FSQA*	0.9 ± 0.2	62.3 ± 25.7	$1.8 \pm 1.1 \times 10^4$
7	WSLFRMVVL	N.D.†	N.D.†	N.D.†

* WAAFRFSQA is from Ref. Angelo et al., 2006.
† N.D. means not determined.

Six peptides derived from selected phages were synthesized by solid-phase synthesis. Each peptide possessed an N-terminal Trp to facilitate detection of the intact peptide and the hydrolysis product by monitoring absorbance at 280 nm. Initial rates of hydrolysis determined by measuring product formation relative to the initial substrate concentration via HPLC, and the kinetic parameters k_{cat} and K_m were evaluated and fitted by non-linear regression to the Michaelis-Menten equation (Table 4.2). 6/7 peptides were cleaved by KLK6 with k_{cat}/K_m values between $5 \pm 2.6 \times 10^3$ $M^{-1}s^{-1}$ and $2.2 \pm 1.4 \times 10^4$ $M^{-1}s^{-1}$. ESI-MS analysis revealed that all peptides were cleaved at a single site following the Arg residue. No secondary cleavage products could be detected. Even though phage encoding SLFRMVVL was efficiently cleaved by KLK6, the corresponding synthetic peptide showed very little cleavage even after four hours of incubation with 100 μM

KLK6. The lack of cleavage of WSLFRMVVL was at least partly due to the lower solubility of this peptide relative to the other sequences tested. The FLAG sequence in the phage encoding SLFRMVVL may have affected the solubility so that it was cleaved by the enzyme efficiently.

4.3.2 Analysis of the specificity of KLK1

Phage containing octapeptide sequences that are cleaved by KLK1 were isolated after six rounds of panning. Individual phage clones were amplified and examined by western blotting for loss of the FLAG tag following incubation with 4 μM KLK1 for 45 minutes. 16 clones that showed extensive cleavage in this assay were examined further. Peptides were aligned such that either Arg or Phe and other hydrophobic amino acids occupied the P1 position, in accord with the published information on the substrate specificity of KLK1 (for a review, see Paliouras and Diamandis 2006). As expected, KLK1 showed both trypsin-like and chymotrypsin-like specificity (Table 4.3). 12 of 16 phage clones encoded chymotrypsin-like substrate sequences, having large hydrophobic amino acids N-terminal at P1. 4 out 16 phage clones encoded trypsin-like substrates containing an Arg at the P1 position. The P2' position favored polar or charged amino acids. No specific bias for particular amino acids could be discerned for the P2-P4 and P3', P4' positions.

Seven peptides corresponding to selected sequences were synthesized. ESI/MS of the KLK1 hydolysis products revealed in all cases a single cleavage occurred N-terminal

to Arg or Ser, consistent with the assignment of these amino acids to the P1' position. WAHRTTFYRRGA was found to be the best substrate with a k_{cat}/K_m of $1.3 \pm 0.5 \times 10^7$ (Table 4.4).

Table 4.3 Alignment of sequences selected by substrate phage display with KLK1

(A) Chymotrypsin-like group

Clone	P4	P3	P2	P1	P1'	P2'	P3'	P4'
6R-29	T	T	F	Y	R	R	*G*	*A*
6R-28	*E*	*A*	S	Y	R	R	K	Q
6R-5	*A*	S	S	Y	R	T	S	R
6R-30	*A*	A	W	Y	R	T	S	R
6R-26	A	R	L	Y	S	R	*G*	*A*
4R-25	*E*	*A*	F	Y	S	Q	R	F
6R-8	T	R	F	Y	S	R	G	R
6R-44	S	F	H	Y	R	M	V	*G*
6R-13	G	T	L	F	R	S	G	N
3R-7	P	N	R	W	S	T	*G*	*A*
6R-11	S	S	E	W	S	M	P	Y
5R-18	S	S	Y	I	S	N	F	G

(B) Trypsin-like group.

Clone	P4	P3	P2	P1	P1'	P2'	P3'	P4'
3R-1	*L*	*E*	*A*	R	S	A	Y	H
5R-4	N	A	A	R	S	T	*G*	*A*
3R-13	*E*	*A*	K	R	S	Y	H	S
6R-27	*E*	*A*	S	R	S	A	T	L

Amino acids from the linker region are shown in italic.

Table 4.4 Kinetic measurements of KLK1 catalyzed hydrolysis of substrate peptides

Peptide no.	Amino acid sequence of peptide substrate	KLK1		
		k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
1	WAHRTTFY↓RRGA	35.5 ± 1.5	3.2 ± 1.2	$1.3 \pm 0.5 \times 10^7$
2	WGYARLY↓SRGA	13.1±1.8	1.6±0.5	$9.4 \pm 4.1 \times 10^6$
3	WAFY↓SQRFRK	1.5 ± 0.5	0.3 ±0.1	$5.6 \pm 2.8 \times 10^6$
4	WGTLF↓RSGN	9.9 ±1.3	83±17.9	$1.3 \pm 0.4 \times 10^5$
5	WGTLR↓SSGN	4.2 ± 0.9	452.8± 118.3	$1.1 \pm 0.5 \times 10^4$
6	WTRFY↓SRGRG	16.5± 1	3± 1.3	$7.3 \pm 3.5 \times 10^6$
7	WKLR↓SSKQ*	9.9 ± 0.5	2.3 ± 1	$5.4 \pm 2.6 \times 10^6$

* WKLRSSKQ is from Ref. Laxmikanthan et al., 2005.

4.3.3 Molecular modeling of peptide substrates bound to KLK6 or KLK1

Molecular docking of the consensus peptide substrates onto the crystal structure of KLK1 or KLK6 was performed using GOLD v3.1.1 For KLK 6, cavity was defined with a radius of 20 Å centered around the OH group of Ser195. Five acetylated peptides were tested (FRSA, FRSV, VRSA, VRSV, and FRFSQ) and a total of 200 genetic algorithm (GA) simulations were performed on each peptide. Among the top five candidates from the docking calculations, the Arg is shown to predominantly occupy S1 pocket (Fig. 4.4), consistent with the biochemical data. The Arg guanidinium group is capable of making a number of hydrogen bonds with protein residues Asp189, Ser190, and Asn217 in the S1 pocket and with Ser214, Thr229, Asp102 and Tyr94 in the S2 pocket. The catalytic Ser195 residue makes hydrogen bond contacts to the backbone of the peptide at P1-P1'. C-terminal residues show consistent binding to the S2'

hydrophobic pocket comprising of residues Gly193, Leu40, Leu41, Gly142, Trp141, and Phe151. When Phe precedes Arg, as in FRSA and FRSV, the docking also predict among the top candidates binding modes where Phe reside in S1 and Arg in S2. This is likely an artifact of the scoring function of GOLD, which utilizes approximated functions describing hydrophobic forces.

For KLK1, seven peptides were tested: FYRR, LYSR, ARSA, SRSA, KRSY, FYSQ, and FYSR. The predominant binding mode for peptides with Arg at P1 (trypsin-like substrate) into KLK1 cavity is with Arg in S1 (Fig 4.5A). When Arg is not present at P1 (chymotrypsin-like substrate), Tyr is found in S1 almost exclusively, making a hydrogen bond with ASP189 (Fig 4.5B). For FYRR, the docking calculations show that either Tyr or the first Arg can occupy in the S1 pocket indicating that the benzene ring competes with guanidinium for the S1 pocket. As the KLK1 possess both trypsin and chymotrypsin characteristics, the S1 pocket is able to recognize both Tyr and Arg.

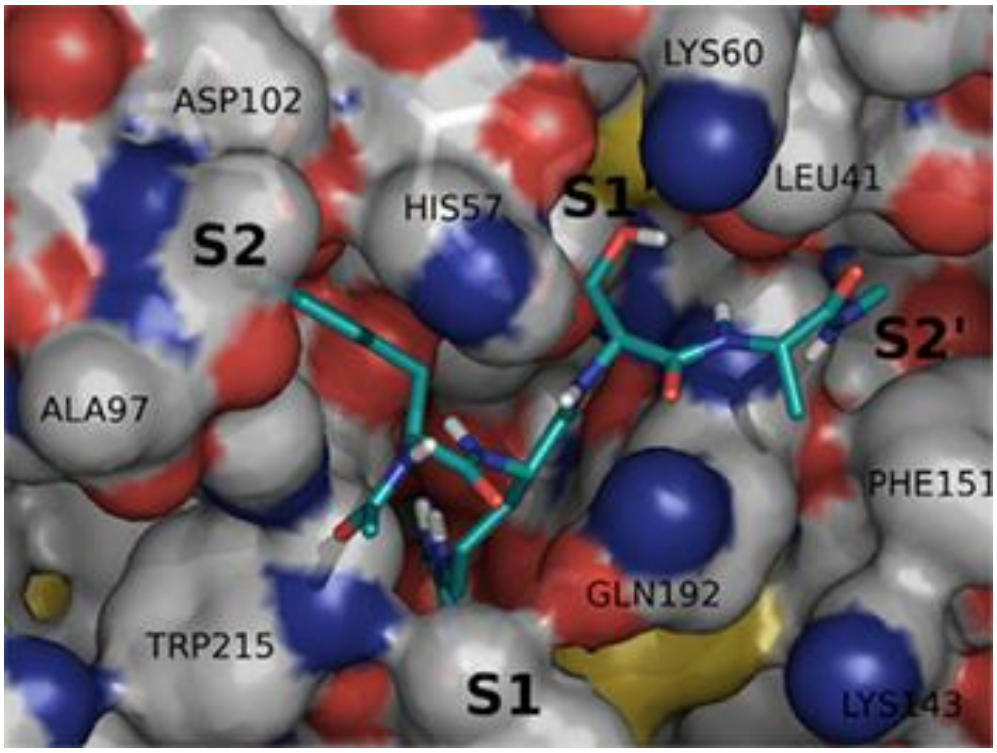
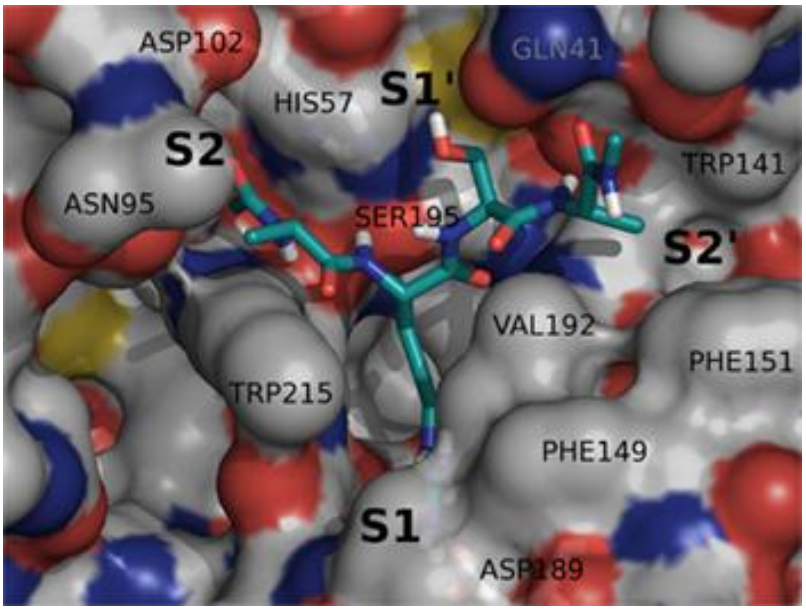
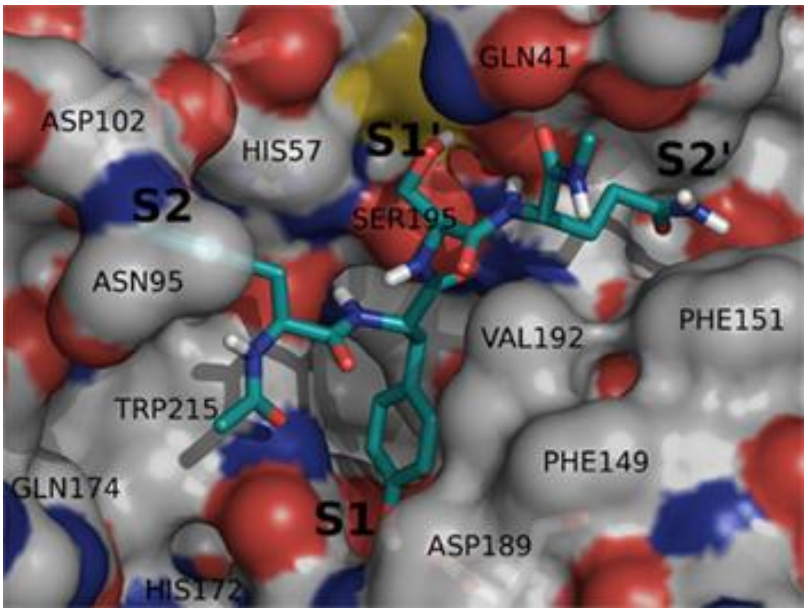


Figure 4.4 The proposed substrate binding around the active site binding pocket of KLK6 by docking simulation using GOLD v3.1.1. The peptide FRSA exists in the active site binding pocket of KLK6. Arg is capable of making hydrogen bonds with protein residues Asp189, Ser190, and Asn217 which compose S1 binding pocket.



(A)



(B)

Figure 4.5 The proposed substrate binding around the active site binding pocket of KLK1 by docking simulation using GOLD v3.1.1. (A) The peptide ARSA (trypsin-like

substrate) and (B) the peptide FYSQ (chymotrypsin-like substrate) bind the active site binding pocket of KLK1.

4.4 Discussion

In this study we investigated the substrate specificity of kallikrein 1 and 6 using substrate phage display. The advantages of substrate phage over methods that rely on the cleavage of fluorescently labeled synthetic substrate libraries have been reviewed recently (Sedlacek and Chen 2005). However, one potential problem with this technique is that clones that are cleaved by endogenous proteases either during phage biogenesis or upon phage purification become enriched together with sequences that are cleaved following incubation by the target protease. To avoid such artifacts, the cleavage of all isolated phage clones by either KLK1 or KLK6 as appropriate, were examined by western blotting following incubation with or without target protease. In this manner we identified a total of 16 and 20 unique sequences cleaved by KLK1 and KLK6, respectively. Peptides corresponding to sequences isolated by substrate phage were synthesized. N-terminal tryptophan was added to provide a chromophore which allowed the kinetics of cleavage to be determined by HPLC. The products of cleavage were analyzed by ESI-MS. Synthetic peptides based on isolated sequences were found to be cleaved efficiently, further confirming that they constitute true substrates. Meanwhile the high efficiency cleavage of peptide substrates by KLK1 and KLK6 indicated that the N-terminal tryptophan did not interfere with enzyme activity. The one exception was the sequence WSLFRMVVL which however was less soluble than the other synthetic substrates and could only be tested at lower concentrations.

Analysis of the sequences from the KLK6 selection revealed that the presence of an Arg at the P1 position of all 20 phage clones analyzed. P1' showed a very strong preference for Ser although clones with Asn, Met and Tyr at that position were also enriched. A significant preference for Phe or Val in P2 and for Val or Ala in P2' was also discerned (Table 4.1). Our results are consistent with the analysis of Angelo et al. except for the amino acid preferences at P2' where positional scanning of peptide substrates did not detect a preference for Ala or Val. Debela et al. examined the specificity of KLK6 using positional libraries of the general structure acetyl-P4-P3-P2-P1-ACC (ACC = 7-amino-4-carbamoylmethylcoumarin) and reported that in addition to Arg, the S1 subsite can accommodate Ala, His, Met, Lys and that Arg and Lys preferentially occupy the S2 subsite. These results are in conflict with both the data from Angelo et al. and from the present study and are probably a consequence of the ACC moiety occupying the P1' position. It should be noted that human pro-KLK6 is self-autoactivated to an active by self-cleavage between Lys21 and Leu22 (Yoon et al. 2007). Self-autoactivation is detected after 24 hr incubation at 37 ^{o}C but not after 1 hr incubation at 37 ^{o}C indicating that the rate of hydrolysis is slow. In this work we carried out incubation of the phage with protease 30 min, to avoid the selection of slowly hydrolyzed peptides containing non-optimal cleavage sites.

The tetrapetide Phe/Val-Arg-Ser-Ala/Val from P2 to P2' constitutes the consensus recognition site for KLK6. A blast search (www.ncbi.nlm.nih.gov) was performed based on sequence homology to phage display selected substrates and the consensus sequence

Phe/Val-Arg-Ser-Ala/Val of KLK6. The blast search gave us hundreds of hits. However several putative substrate candidates of KLK6 were found based on additional information from published biological studies. For example, ionotropic glutamate receptor (GluR) has been proposed to be possible substrate of KLK6 so that KLK6 may be involved in the modulation of glutamate-mediated neuronal activity in CNS. (Angelo et al. 2006). Our blast search showed that GluR N-methyl D-aspartate 2C (GRIN2C) contains the sequence VRSV that matches the consensus. Similarly, GluR N-methyl D-aspartate 2D (GRIN2D) contains the consensus sequence VRSA. This finding further supports the notion that the inonotropic glutamate receptor is a potential target of KLK6 and might interact with GRIN mediating further GluR excitation in CNS.

The major character of Parkinson syndrome is the aggregation of insoluble alpha synuclein and the formation of Lewy bodies in the patients' brain (Spillantini et al. 1997). KLK6 was shown to degrade alpha synuclein and prevent its polymerization (Iwata et al. 2003). We found that the synuclein alpha binding protein-synphilin contains the sequence FRSI which exhibits strong homology to the consensus cleavage motif determined here. Synphilin has been shown to interact with alpha synuclein, promote aggregation of synuclein and the formation of inclusion bodies in neurons (Engelender et al. 1999). Our results suggest that KLK6 might play a role in the degradation of synphilin thus decreasing the formation of inclusions in neurons in parkinson's disease. However, further experiments to check the KLK6 digestion of these proteins will be suggested to confirm these results.

In addition to kininogen, KLK1 appears to have many other physiological protein substrates including pro-insulin, prorenin and low density lipoprotein etc (for a review, see Yousef and Diamandis 2002; Paliouras and Diamandis 2006). Therefore KLK1 has been proposed to exhibit different enzyme functions in specific tissues or cell types. KLK1 cleaves human low molecular weight kininogen at Met-Lys and Arg-Ser bonds (Sueiras et al. 1994) while it can also cleave after the Phe-Phe pair of kallistatin (Zhou et al. 1992) and somatostatin (Pimenta et al. 1997). Here, phage display confirmed that KLK1 is able to accept two distinct classes of substrates and exhibits trypsin-like and chymotrypsin-like activities (Table 4.3A and 4.3B). For the latter class of substrates, Tyr was the predominant residue in the P1 position. Further, our analysis indicates a strict preference for Ser or Arg in the P1' position of chymotrypsin-like substrates of KLK1 (Table 4.3A). As expected based on the known substrates of KLK1 a variety of amino acids can be accommodated in the P2 position of the substrate, although there is a preference for aromatic residues (5 out of 12 phage clones).

Substrates containing trypsin-like sites were enriched at a lower frequency relative to chymotrypsin-like substrates (Table 4.3B). Among the four such sequences obtained there was a perfect consensus for Arg at P1 and Ser at P1'. Table 4.3B also indicates a preference for Ala in the P3 position (3/4 sequences). However, given the small number of trypsin-like substrates obtained, this data should be interpreted with caution.

Five peptides corresponding to sequences isolated by substrate phage were synthesized, the kinetics of cleavage determined by HPLC and the products of cleavage were analyzed by ESI-MS. All five peptides were cleaved with high catalytic efficiency under the conditions used (k_{cat}/K_m $>10^5$ M^{-1} s^{-1}). The best substrate was WAHRTTFYRRGA which was hydrolyzed very rapidly, with k_{cat}/K_m of $1.3 \pm 0.5 \times 10^7$ (Table 4.4). Interestingly, the higher rate of cleavage of this peptide was primarily due to an increased k_{cat}. Although, peptide #4 (WGTLFRSGN) contained both a putative chymotrypsin cleavage site (between Phe and Arg) and a trypsin cleavage site (between Arg and Ser) incubation with KLK1 resulted exclusively in the formation of WGTLF and RSGN. To examine whether the amino acids distal to the cleavage site determine whether KLK1 exhibits trypsin-like or chymotrypsin like activity, we synthesized a derivative of the substrate phage-derived peptide #4 (WGTLF↓RSGN) in which the P1 and P1' were changed from Phe-Arg to Arg-Ser, respectively. This substitution resulted in a 10-fold reduction in the k_{cat}/K_m. Keeping the P2-P2' sequence in peptide #5 fixed but changing the P3-P4 and P3'P4' residues increased the catalytic rate with peptide #7 by 500 fold.

Superposition of KLK1 and KLK6 reveals the two enzymes are very similar structurally, with RMSD of 0.88 Å for C_α. A closer examination shows that S1 cavity in KLK6 is generally tighter and is surrounded by more polar groups. Specifically, Tyr288, Tyr172, Thr190, Gln192, Asn217, Ile218, Gly226, and Ile138 form the S1 cavity of KLK6 compared to Ala228, His172, Ser190, Val192, Try217, Val218, Ser226, and

Ala138 in KLK1. Also, loop 216-226 is brought closer to the S1 cavity in KLK6. The presence of Asp189 in both KLK1 and KLK6 indicates that, unlike chymotrypsin, KLK1 enzyme employs more subtle differences around the S1 cavity to allow for chymotrypsin-like activity, namely a larger cavity and smaller hydrophobic groups to accommodate binding of aromatic groups. The docking simulations show that Arg or Tyr fit into the S1 pocket of KLK1 equally well, which places the peptide in the binding mode consistent with the dual specificities of the enzyme. A major factor in determining the binding mode of KLK1 appears to be the residues at P1' and P2', as S1' site accepts small polar residues and S2' prefers large hydrophobic residues almost exclusively. The S2 site, according to GOLD, appears to be largely promiscuous, capable of accepting a range of polar and aromatic residues. For KLK6, the above preferences are also true; however, since the S1 pocket of KLK6 is less tolerant to large aromatic residues, GOLD predicts that Arg binds to S1 pocket of KLK6, facilitating trypsin-like reaction. Comparing different binding modes we observe that the peptide backbone stays close to catalytic Ser195, indicating that a nucleophilic attack is likely to take place between P1-P1'.

Acknowledgements

This work was supported by National Institute of Health Grants R01 GM065551, RO1 GM073089 and 1R15NS057771-01. The authors PG and PR would like to acknowledge the support by grant number R01GM079686 from the National Institute of General Medical Sciences. This chapter is revised from previously published work: Li, H.X., *et al.*

(2008) “Substrate specificity of human kallikreins 1 and 6 determined by phage display.” Protein Science **17**: 664-672

4.5 References

Angelo, P.F., Lima, A.R., Alves, F.M., Blaber, S.I., Scarisbrick, I.A., Blaber, M., Juliano, L., and Juliano, M.A. 2006. Substrate specificity of human kallikrein 6: salt and glycosaminoglycan activation effects. J. Biol. Chem. 281: 3116−3126.

Anisowicz, A., Sotiropoulou, G., Stenman, G., Mok, S.C., and Sager, R. 1996. A novel protease homolog differentially expressed in breast and ovarian cancer. Mol. Med. 2: 624−636.

Bayes, A., Tsetsenis, T., Ventura, S., Vendrell, J., Aviles, F.X., and Sotiropoulou, G. 2004. Human kallikrein 6 activity is regulated via an autoproteolytic mechanism of activation/inactivation. Biol. Chem. 385: 517−524.

Betnett, M.J., Blaber, S.I., Scarisbrick, I.A., Dhanarajan, P., Thompson, S.M., and Blaber, M. 2002. Crystal structure and biochemical characterization of human kallikrein 6 reveals that a trypsin-like kallikrein is expressed in the central nervous system. J. Biol. Chem. 277: 24562−24570.

Bhoola, K.D., Figueroa, C.D., and Worthy, K. 1992. Bioregulation of kinins: kallikreins, kininogens, and kininases. Pharmacol. Rev. 44: 1−80.

Borgono, C.A., Michael, L.P., and Diamandis, E.P. 2004. Human Tissue Kallikreins: Physiologic roles and applications in Cancer. Mol. Cancer Res. 2: 257−280.

Casabadan, M.J., and Cohen, S.N. 1980. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138: 179–207.

Cloutier, S.M., Chagas, J.R., Mach, J.P., Gygi, C.M., Leisinger, H.J., and Deperthes, D. 2002. Substrate specificity of human kallikrein 2 (hK2) as determined by phage display technology. Eur. J. Biochem. 269: 2747−2754.

Debela, M., Magdolen, V., Schechter, N., Valachova, M., Lottspeich, F., Craik, C.S., Choe, Y., Bode, W., and Goetting, P. 2006. Specificity profiling of seven human tissue kallikreins reveals individual subsite preferences. J. Biol. Chem. 281: 25678−25688.

Deperthes, D. 2002. Phage display substrate: a blind method for determining protease specificities. Biol. Chem. 383: 1107−1112.

Diamandis, E.P., Yousef, G.M., Clements, J., Ashworth, L.K., Yoshida, S., Egelrud, T., Nelson, P.S., Shiosaka, S., Little, S., Lilja, H., Stenman, U.H., Rittenhouse, H.G., and

Wain, H. 2000. New nomenclature for the human tissue kallikrein gene family. Clin. Chem. 46: 1855−1858.

Diamandis, E.P. Scorilas, A., Fracchioli, S., Gramberen, M., Bruijin, H., Henrik, A., Soosaipillai, A., Grass, L., Yousef, G.M., Stenman, U.H., Massobrio, M., Zee, A.G.J., Vergote, I., and Katsaros, D. 2003. Human kallikrein 6 (hK6): a new potential serum biomarker for diagnosis and prognosis of ovarian carcinoma. J. Clin. Oncol. 21: 1035−1043.

Diamond, S.L. 2007. Methods for mapping protease specificities. Curr. Opin. Chem. Biol. 11: 46−51.

Engelender, S., Kaminsky, Z., Guo, X., Sharp, A.H., Amaravi, R.K., Kleiderlein, J.J., Margolis, R.L., Troncoso, J.C., Lanahan, A.A., Worley, P.F., and Dawson, V.L. 1999. Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. Nat. Genet. 22: 110−114.

Felber, L.M., Borgono, C.A. Cloutier, S.M., Kundig, C., Kishi, T., Chagas, J.R., Jichlinski, P., Gygi, C.M., Leisinger, H.J., Diamandis, E.P., and Deperthes, D. 2005. Enzymatic profiling of human kallikrein 14 using phage-display substrate technology. Biol. Chem. 386: 291−298.

Ferrieu-Weisbuch, C., Michel, S., Collomb-Clerc, E., Pothion, C., Deleage, G., and Jolivet-Reynaud, C. 2006. Characterization of prostate-specific antigen binding peptides selected by phage display technology. J. Mol. Recognit. 19: 10−20.

Fiedler, F., and Leysath, G. 1979. Substrate specificity of porcine pancreatic kallikrein. Adv. Exp. Med. Biol. 120A: 261−271.

Ghosh, M.C., Grass, L., Soosaipillai, A., Sotiropoulou, G., and Diamandis, E.P. 2004. Human kallikrein 6 degrades extra cellular matrix proteins and may enhance the metastatic potential of tumor cells. Tumor Biol. 25: 193−199.

Hwang, B.-Y., Varadarajan, N., Li, H., Rodriquez, S., Iverson, B.L., and Georgiou, G. 2007. Substrate specificity of the Escherichia coli outer membrane protease OmpP. J. Bacteriol. 189: 522−530.

Iwata, A., Maruyama, M., Akagi, T., Hashikawa, T., Kanazawa, I., Tsuji, S., and Nukima, N. 2003. Alpha-synuclein degradation by serine protease neurosin: implication for pathogenesis of synucleinopathies. Hum. Mol. Genet. 12: 2625−2635.

Kerr, F.K., O'Brien G., Quinsey N.S., Whisstock J.C., Boyd S., de la Banda M.G., Kaiserman D., Matthews A.Y., Bird P.I., and Pike R.N. 2005. Elucidation of the substrate specificity of the C1s protease of the classical complement pathway. J. Biol. Chem. 280: 39510–39514.

Laxmikanthan, G., Blaber, S.I., Bernett, M.J., Scarisbrick, I.A., Juliano, M.A., and Blaber, M. 2005. 1.70 Å x-ray structure of human apo kallikrein 1: structural changes upon peptide inhibitor/substrate binding. Proteins 58: 802−814.

Little, S.P., Dixon, E.P., Norris, F., Buckley, W., Becker, G.W., Johnson, M., Dobbins, J.R., Wyrick, T., Miller, J.R., MacKellar, W., Hepburn, D., Corvalan, J., McClure, D., Liu, X., Stephenson, D., Clemens, J., and Johnstone, E.M. 1997. Zyme, a novel and potentially amyloidogenic enzyme cDNA isolated from Alzheimer's disease brain. J. Biol. Chem. 272: 25135−25142.

Lovell, J.M., Richardson, J.S., and Richardson, D.C. 2000. The penultimate rotamer library. Proteins 40: 389−408.

Magklara, A., Mellati, A.A., Wasney, G.A., Little, S.P., Sotiropoulou, G., Becker, G.W., and Diamandis, E.P. 2003. Characterization of the enzymatic activity of human kallikrein 6: autoactivation, substrate specificity, and regulation by inhibitors. Biochem. Biophys. Res. Commun. 307: 948−955.

Oikonomopoulou, K., Hansen, K.K., Saifeddine, M., Tea, I., Blaber, M., Blaber, S.I., Scarisbrick, I., Andrade-Gordou, P., Cottrell, G.S., Bunnett, N.W., Diamandis, E.P., and Hollenberg, M.D. 2006. Proteinase-activated receptors, targets for kallikrein signaling. J. Biol. Chem. 281: 32095−32112.

Paliouras, M., and Diamandis E.P. 2006. The kallikrein world: an update on the human tissue kallikreins. Biol. Chem. 387: 643−652.

Pimenta D.C., Juliano M.A., and Juliano L. 1997. Hydrolysis of somatostatin by human tissue kallikrein after the amino acid pair Phe-Phe. Biochem. J. 327: 27−30.

Scarisbrick, I.A., Towner, M.D., and Isackson P.J. 1997. Nervous system-specific expression of a novel serine protease: regulation in the adult rat spinal cord by excitotoxic injury. J. Neurosci. 17: 8156−8168.

Scott, J.K., and Smith, G.P. 1990. Searching for peptide ligands with an epitope library. Science 249: 386–390.

Sedlacek, R., and Chen, E. 2005. Screening for protease substrate by polyvalent phage display. Comb. Chem. High Throughput Screen. 8: 197−203.

Spillantini, M.G., Schmidt, M.L., Lee, V.M., Trojanowski, J.Q., Jakes, R., and Goedert, M. 1997. Alpha-synuclein in Lewy bodies. Nature 388: 839−840.

Sueiras, D.J., Jones, D.M., Ashworth, D., Horton, J., Evans, D.M., and Szelke, M. 1994. Cleavage of human kininogen fragments at Met-lys by human tissue kallikrein. Braz. J. Med. Biol. Res. 27:1935−1942.

Yamashiro, K., Tsuruoka, N., Kodama, S., Tsujimoto, M., Yamamura, Y., Tanaka, T., Nakazato, H., and Yamaguchi, N. 1997. Molecular cloning of a novel trypsin-like serine protease (neurosin) preferentially expressed in brain. Biochim. Biophys. Acta. 1350: 11−14.

Yoon, H., Laxmikanthan, G., Lee, J., Blaber, S.I., Rodriguez, A., Kogot. J.M., Scarisbrick I.A., and Blaber, M. 2007. Activation profiles and regulatory cascades of the human kallikrein-related peptidases. Biochemistry 46: 5209−5217.

Yousef, G.M., Luo, L.Y., Scherer, S.W., Sotiropoulou, G., and Diamandis, E.P. 1999. Molecular characterization of Zyme/Protease M/Neurosin (PRSS9), a hormonally regulated kallikrein-like serine protease. Genomics 62: 251−259.

Yousef, G.M., and Diamandis, E.P. 2001. The new human tissue kallikrein gene family: structure, function, and association to disease. Endocr. Rev. 22: 184−204.

Yousef, G.M., Plymeris, M.E., Yacoub, G.M., Scorilas, A., Soosaipillai, A., Popalis, C., Fracchioli, S., Katsaros, D., and Diamandis, E.P. 2003. Parallel overexpression of seven kallikrein genes in ovarian cancer. Cancer Res. 63: 2223−2227.

Zhou. G. X, Chao L., and Chao J. 1992. Kallistatin: a novel human tissue kallikrein inhibitor. Purification, characterization, and reactive center sequence. J. Biol. chem. 267: 25873−25880.

Bibliography

Ali H, Panettieri RA Jr (2005) "Anaphylatoxin C3a receptors in asthma" Respir. Res. 6: 19

Angelo, P.F., Lima, A.R., Alves, F.M., Blaber, S.I., Scarisbrick, I.A., Blaber, M., Juliano, L., and Juliano, M.A. 2006. Substrate specificity of human kallikrein 6: salt and glycosaminoglycan activation effects. J. Biol. Chem. 281: 3116−3126.

Anisowicz, A., Sotiropoulou, G., Stenman, G., Mok, S.C., and Sager, R. 1996. A novel protease homolog differentially expressed in breast and ovarian cancer. Mol. Med. 2: 624−636.

Arnold F.H., and Georgiou G. (2003) "Directed Evolution Library Creation. Methods and protocols" 75-90

Arumugam, T.V., Shiels, I.A, Woodruff, T.M., Granger, D.N., Taylor, S.M. (2004). "The role of the complement system in ischemia-reperfusion injury". Shock 21 (5): 401.

Ballinger, M. D., Tom, J. and Wells, J.A. (1996) "Furilisin: a variant of subtilisin BPN engineered for cleaving tribasic substrates." Biochemistry 35: 13579-13585

Baum, E.Z., Bebemitz, G.A. and Gluzman, Y. (1990) "Isolation of mutants of human immunodeficiency virus protease based on the toxicity of the enzyme in Escherichia coli. Proc. Natl. Acad. Sci. USA 87: 5573-5577

Bayes, A., Tsetsenis, T., Ventura, S., Vendrell, J., Aviles, F.X., and Sotiropoulou, G. 2004. Human kallikrein 6 activity is regulated via an autoproteolytic mechanism of activation/inactivation. Biol. Chem. 385: 517−524.

Berg, D.T., Gerlitz, B., Shang, J., Smith, T., Santa, P., Richardson, M.A., Kurz, K.D., Grinnell, B.W., Mace, K., Jones, B.E. (2003) "Engineering the proteolytic specificity of activated protein C improves its pharmacological properties." Proc. Natl. Acad. Sci. USA 100: 4423-4428

Bernard, G.R., Vincent, J.L., Laterre, P.F., LaRosa, S.P., Dhainaut, J.F., Lopez-Rodriguez, A., Steingrub, J.S., Garber, G.E., Helterbrand, J.D., Ely, E.W., Fisher Jr., C.J. "Efficacy and safety of recombinant human activated protein C for severe sepsis. N. Engl. J. Med. 344: 699-709

Bernstein J.A. (2008) "Hereditary angioedema: a current state-of-the-art review, VIII: current status of emerging therapies". Ann. Allergy Asthma Immunol. 100 (1 Suppl 2): S41–6.

Betnett, M.J., Blaber, S.I., Scarisbrick, I.A., Dhanarajan, P., Thompson, S.M., and Blaber, M. (2002) "Crystal structure and biochemical characterization of human kallikrein 6 reveals that a trypsin-like kallikrein is expressed in the central nervous system." J. Biol. Chem. 277: 24562−24570

Bhoola, K.D., Figueroa, C.D., and Worthy, K. (1992) "Bioregulation of kinins: kallikreins, kininogens, and kininases." Pharmacol. Rev. 44: 1−80

Blanco, R., Carrasco, L. and Ventoso, I. (2003) "Cell killing by HIV-1 protease." J.Biol.Chem. 278: 1086-1093

Borgono, C.A., Michael, L.P., and Diamandis, E.P. (2004) "Human Tissue Kallikreins: Physiologic roles and applications in Cancer." Mol. Cancer Res. 2: 257−280

Bryan, P.N. (2000) "Protein engineering of subtilisin" Biochim. Biophys. Acta. 1543(2): 203-222. Review

.

Caroll, M.C. (2004) "The complement system in regulation of adaptive immunity" Nat. Immuno. 5(10): 981-986

Carters, P. and Wells, J.A. (1987) "Engineering enzyme specificity by "substrate-assisted catalysis". Science 237(4813):394-399

Caporale, L.H., Tippett, P.S. and Erickson, B.W. (1980) "The active site of C3a anaphylatoxin" J. Biol. Chem. 255: 10758-10763

Casabadan, M.J. and Cohen, S.N. (1980) "Analysis gene control signals by DNA fusion and cloning in Escherichia coli. J. Mol. Biol. 138; 179-207

Cloutier, S.M., Chagas, J.R., Mach, J.P., Gygi, C.M., Leisinger, H.J., and Deperthes, D. (2002) "Substrate specificity of human kallikrein 2 (hK2) as determined by phage display technology." Eur. J. Biochem. 269: 2747−2754

Debela, M., Magdolen, V., Schechter, N., Valachova, M., Lottspeich, F., Craik, C.S., Choe, Y., Bode, W., and Goetting, P. 2006. Specificity profiling of seven human tissue kallikreins reveals individual subsite preferences. J. Biol. Chem. 281: 25678−25688.

Dekker, N., Cox, R.C., Kramer, R.A. and Egmond, M.R. (2001) "Substrate specificity of the integral membrane protease OmpT determined by spatially addressed peptide libraries." Biochemistry 40: 1694-1701

Deperthes, D. (2002) "Phage display substrate: a blind method for determining protease specificities." Biol. Chem. 383: 1107−1112

DeSantis, G., Berglund, P., Stabile, M.R., Gold, M., Jones, J.B. (1998) "Site-directed mutagenesis combined with chemical modification as a strategy for altering the specificity of the S1 and S1' pockets of subtilisin Bacillus lentus." Biochemistry 37: 5968-5973

Diamandis, E.P., Yousef, G.M., Clements, J., Ashworth, L.K., Yoshida, S., Egelrud, T., Nelson, P.S., Shiosaka, S., Little, S., Lilja, H., Stenman, U.H., Rittenhouse, H.G., and Wain, H. 2000. New nomenclature for the human tissue kallikrein gene family. Clin. Chem. 46: 1855−1858.

Diamandis, E.P. Scorilas, A., Fracchioli, S., Gramberen, M., Bruijin, H., Henrik, A., Soosaipillai, A., Grass, L., Yousef, G.M., Stenman, U.H., Massobrio, M., Zee, A.G.J., Vergote, I., and Katsaros, D. 2003. Human kallikrein 6 (hK6): a new potential serum biomarker for diagnosis and prognosis of ovarian carcinoma. J. Clin. Oncol. 21: 1035−1043.

Diamond, S.L. (2007) "Methods for mapping protease specificities." Curr. Opin. Chem. Biol. 11: 46−51

Di Cera E. (2008) "Engineering protease specificity made simple, but not simpler" Nat. Chem. Bio. 4: 270-271

Drouin, S.M, Kildsgaard, J., Haviland, J., Zabner, J., Jia, H.P., McCray, P.B. Jr, Tack, B.F., Wetsel, R.A. (2001) "Expression of the complement anaphylatoxin C3a and C5a receptors on bronchial epithelial and smooth muscle cells in models of sepsis and asthma" J Immunol. 166(3):2025-32

Dunkelberger, J.R. and Song, W.C. (2010). "Complement and its role in innate and adaptive immune responses" Cell Res. 20:34-50

Ekici O.D., Karia A., Paetzel M, Lively M.O., Pei D, Dalbey R.E. (2007) " Altered P3 substrate specificity of Escherichia coli signal peptidase 1 mutants as revealed by screening a combinatorial peptide library. J Biol. Chem. 282:417-425

Ember, J.A., Kildsgaard, J., Haviland, D.L., Lambris J.D. and Holers, M.V. (2000) "Therapeutic interventions in the complement system" Humana press, Totowa, NJ. pp 113-153

Engelender, S., Kaminsky, Z., Guo, X., Sharp, A.H., Amaravi, R.K., Kleiderlein, J.J., Margolis, R.L., Troncoso, J.C., Lanahan, A.A., Worley, P.F., and Dawson, V.L. 1999. Synphilin-1 associates with alpha-synuclein and promotes the formation of cytosolic inclusions. Nat. Genet. 22: 110−114.

Esmon, C.T., Ding, W., Yasuhiro, K., Gu, J.M., Ferrell, G., Regan, L.M., Stearns-Kurosawa, S., Mather, T., Laszik, Z. and Esmon, N.L. (1997) "The protein pathway: new insights" Thromb. Haemostasis 78:70-74

Felber, L.M., Borgono, C.A. Cloutier, S.M., Kundig, C., Kishi, T., Chagas, J.R., Jichlinski, P., Gygi, C.M., Leisinger, H.J., Diamandis, E.P., and Deperthes, D. (2005) "Enzymatic profiling of human kallikrein 14 using phage-display substrate technology." Biol. Chem. 386: 291−298

Ferrieu-Weisbuch, C., Michel, S., Collomb-Clerc, E., Pothion, C., Deleage, G., and Jolivet-Reynaud, C. (2006) "Characterization of prostate-specific antigen binding peptides selected by phage display technology." J. Mol. Recognit. 19: 10−20

Fiedler, F., and Leysath, G. 1979. Substrate specificity of porcine pancreatic kallikrein. Adv. Exp. Med. Biol. 120A: 261−271.

Fisher, S. and Kohnert, U. (1997) "Major mechanistic differences explain the higher clot lysis potency of reteplase over alteplase: lack of fibrin binding is an advantage for bolus application of fibrin-specific thrombolytics." Fibrinolysis and Proteolysis 11(3): 129-135

Franchini, M. (2006) "Recombiant factor VIIa: A review on its clinical use" Int. J. Hematol. 83: 126-138

Ghosh, M.C., Grass, L., Soosaipillai, A., Sotiropoulou, G., and Diamandis, E.P. 2004. Human kallikrein 6 degrades extra cellular matrix proteins and may enhance the metastatic potential of tumor cells. Tumor Biol. 25: 193−199.

Gill, R.T., DeLisa, M.P., Shiloach, M., Holoman, T.R. and Bentley, W.E. (2000) "OmpT expression and activity increase in response to recombinant chloramphenicol acetyltransferase overexpression and heat shock in E. coli" J. Mol. Microbiol. Biotechnol. 2(3): 283-289

Goldhaber, S.Z. (2001) “Thrombolysis in pulmonary embolism” Circulation 104:2876

Grodberg, J. and Dunn, J.J. (1988) “OmpT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification.” J. Bacteriol. 170: 1245-1253

Haas, P.J. and van Strijp J. (2007) “Anaphylatoxins: their role in bacterial infection and inflammation.” Immuno. Res. 37(3): 161-175

Han, H.E., Rho, S.H., Lee, Y.J., Park, W.J. (2005) “Engineering of Kex2 variants exhibiting altered substrate specificity” Biochem. Biophys. Res. Commun. 337: 1102-1106

Hanke, C., Hess, J., Schumacher, G., and Goebel, W. (1992) “Processing by OmpT of fusion proteins carrying the HlyA transport signal during secretion by the Escherichia coli hemolysin transport system.” Mol. Gen. Genet. 233: 42-48

Hedstrom, L., Szilagyi, L. and Rutter, W.J. (1992) “Converting trypsin to chymotrypsin: the role of surface loops.” Science 255(5049):1249-1253

Hedstrom, L. (1996) “Trypsin: a case study in the structural determinants of enzyme specificity.” Biol. Chem. 377(7-8): 465-470

Hedstrom, L. (2002) “Serine protease mechanism and specificity.” Chem. Rev. 102: 4501- 4523

Hellerud, B.C. et al (2010) “Critical roles of complement and antibodies in host defense mechanisms against Neisseria meningitidis as revealed by human complement genetic deficiencies” Infection and Immunity 78(2):802-809

Hoylaerts, M., Rijken, D.C., Lijnen, H.R., and Colleen, D. (1982) “Kinetics of the activation of plasminogen by tissue plasminogen activator” J. Biol. Chem. 257: 2912-2919

Hugli, T.E. (1975) “Human anaphylatoxin (C3a) from the third component of compliment. Primary structure.” J. Biol. Chem. 250: 8293-8301

Humbles, A.A., Lu, B., Nilsson, C.A., Lilly, C., Israel, E., Fujiwara, Y., Gerard, N.P. and Gerard, C. (2000) “A role for the C3a anaphylatoxin receptor in the effector phase of athma.” Nature 406: 998-1001

Hung, S. H. and Hedstrom, L. (1998). “Converting trypsin to elastase: substitution of the S1 site and adjacent loops reconstitutes esterase specificity but not amidase activity.” Protein Engineering 11(8): 669-673

Hwang, B.Y., Varadarajan, N., Li, H.X., Rodriguez, S., Iverson, B. L., Georgiou, G. (2007) “Substrate specificity of the Escherichia coli outer membrane protease OmpP” Journal of Bacteriology, 189(2), 522-530

Iwata, A., Maruyama, M., Akagi, T., Hashikawa, T., Kanazawa, I., Tsuji, S., and Nukima, N. (2003) “Alpha-synuclein degradation by serine protease neurosin: implication for pathogenesis of synucleinopathies.” Hum. Mol. Genet. 12: 2625−2635

Johannes, T., and Zhao H. (2006) “Directed evolution of enzymes and biosynthetic pathways.” Current Opinion in Microbiology 9: 261-267

Johnson, R.J. and Hillmen, P.(2002) “Paroxysmal nocturnal haemoglobinuria: Nature’s gene therapy?” Mol. Path. 55: 145-152

Juliano, L. 1997. “Hydrolysis of somatostatin by human tissue kallikrein after the amino acid pair Phe-Phe.” Biochem. J. 327: 27−30.

Kadonosono, T., Kato-Murai M., Ueda M. (2008) “Alteration of substrate specificity of rat neurolysin from matrix metalloproteinase-2/9-type to -3-type specificity by comprehensive mutation.” Protein Eng Des Sel 21: 507-513

Kaur, J. and Sharma R. (2006) “Directed Evolution: An approach to engineer enzymes.” Critical Reviews in Biotechnology 26(3): 165-199

Kerr, F.K., O'Brien G., Quinsey N.S., Whisstock J.C., Boyd S., de la Banda M.G., Kaiserman D., Matthews A.Y., Bird P.I., and Pike R.N. (2005) “Elucidation of the substrate specificity of the C1s protease of the classical complement pathway.” J. Biol. Chem. 280: 39510–39514.

Khersonsky, O., Roodveldt, C., Tawfik, D.S. (2006) “Enzyme promiscuity: evolutionary and mechanistic aspects.” Curr. Opin. Chem. Biol. 10:498-508

Klauser, T., Pohlner, J., and Meyer, T.F. (1992) “Selective extracellular release of cholera toxin B subunit by Escherichia coli: dissection of Neisseria Igαβ-mediated outer membrane transport.” EMBO J. 11: 2327-2335

Knight, Z.A., Garrison, J.L., Chan, K., King, D.S. and Shokat, K.M. (2007) "A remodeled protease that cleaves phosphotyrosine substrates" J. Am. Chem. Soc. 129: 11672-11673

Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M. and Sigal, I.S. (1988) "Active human immunodeficiency virus protease is required for viral infectivity" Proc. Nat. Acad. Sci. 85, 4686-4690

Kramer, R.A., D.Zandwijken, M.R. Egmond, and N. Dekker. (2000) In vitro folding, purification and characterization of Escherichia coli outer membrane protease OmpT. Eur. J. Biochem. 267:885-893

Kukkonen, M., and T. K. Korhonen. (2004) "The omptin family of enterobacterial surface proteases/adhesins: from housekeeping in Escherichia coli to systemic spread of Yersinia pepsis." Int. J. Med. Microbial. 294: 7-14

Kurth, T.; Grahn,S.; Thormann,M.; Ullman, D.; Hofman, H.J.; Jakubke, H.D.; Hedstrom, L. (1998) "Engineering the S1' subsite of trypsin: design of a protease which cleaves between dibasic residues." Biochemistry 37(33): 11434-11440

Lambris, J.D. and Morikis, D. (2005) "Structure biology of the complement system" ISBN: 0-8247-2540-9 page: 161-174

Laxmikanthan, G., Blaber, S.I., Bernett, M.J., Scarisbrick, I.A., Juliano, M.A., and Blaber, M. 2005. 1.70 Å x-ray structure of human apo kallikrein 1: structural changes upon peptide inhibitor/substrate binding. Proteins 58: 802−814.

Levi, M. and Poll, T. V. D. (2007) "Recombinant human activated protein C: current insights into its mechanism of action" Critical care 11(suppl. 5): S3

Li, H.X., Hwang, B.Y., Laxmikanthan, G., Blaber, S.I., Blaber, M., Golubkov, P.A., Ren, P., Iverson, B.L. and Georgiou, G. (2008) "Substrate specificity of human kallikreins 1 and 6 determined by phage display." Protein Science 17: 664-672

Little, S.P., Dixon, E.P., Norris, F., Buckley, W., Becker, G.W., Johnson, M., Dobbins, J.R., Wyrick, T., Miller, J.R., MacKellar, W., Hepburn, D., Corvalan, J., McClure, D., Liu, X., Stephenson, D., Clemens, J., and Johnstone, E.M. 1997. Zyme, a novel and potentially amyloidogenic enzyme cDNA isolated from Alzheimer's disease brain. J. Biol. Chem. 272: 25135−25142.

Lloyd, J.M., Wight, J., Paisley, S., Knight, C. (2003) "Control of bleeding in patients with haemophilia A with inhibitors: a systematic review." Haemophilia 9: 464-520

Lovell, J.M., Richardson, J.S., and Richardson, D.C. 2000. The penultimate rotamer library. Proteins 40: 389−408.

Magklara, A., Mellati, A.A., Wasney, G.A., Little, S.P., Sotiropoulou, G., Becker, G.W., and Diamandis, E.P. (2003) "Characterization of the enzymatic activity of human kallikrein 6: autoactivation, substrate specificity, and regulation by inhibitors." Biochem. Biophys. Res. Commun. 307: 948−955.

Maier, M., Peng, Y., Jiang, L., Seabrook, T.J., Carroll, M.C. and Lemere, C.A. (2008) "Complement C3 deficiency leads to accelerated amyloid β plaque deposition and neurodegeneration and modulation of the microglia/macrophage phenotype in amyloid precursor protein transgenic mice" J. Neurosci. 28(25): 6333-6341

Mangel, w., d.L. Toledo, M.T. Brown, K. Worzalla, Ml Lee, and J.J. Dunn. (1994). Omptin: an Escherichia coli outer membrane proteinase that activates plasminogen. Methods enzymol. 244: 384-399.

McCarter, J.D., Stephens, D., Shoemaker, K., Rosenberg, S., Kirsch, J.F., Georgiou, G. (2004) "Substrate specificity of the Escherichia coli outer membrane protease OmpT." J. Bacteriol. 186(17): 5919-5925.

McNeil, H.P.; Adachi, R., and Stevens, R.L. (2007) "Mast cell-restricted trypases: structure and function in inflammation and pathogen defense." J. Bio. Chem. 282: 20786-20789

Ng, N.M., Pike, R.N. and Boyd, S.E. (2009) "Subsite cooperativity in protease specificity" Biol. Chem. 390: 401-407

Oikonomopoulou, K., Hansen, K.K., Saifeddine, M., Tea, I., Blaber, M., Blaber, S.I., Scarisbrick, I., Andrade-Gordou, P., Cottrell, G.S., Bunnett, N.W., Diamandis, E.P., and Hollenberg, M.D. 2006. Proteinase-activated receptors, targets for kallikrein signaling. J. Biol. Chem. 281: 32095−32112.

Olsen, M., Stephens, D.L., Griffiths, D., Daugherty, P.S., Georgiou, G. and Iverson, B.L. (2000). "Function-based isolation of novel enzymes from large libraries." Nat. Biotechnol. 18: 1071-1076.

O'Loughlin, T.L., Greene, D.N., Matsumua, I. (2006) "Diversification and specialization of HIV protease function during in vitro evolution" Mol. Biol. Evol. 23:764-772

Okuno, K., Yabuta, M., Ohsuye, K., Ooi, T. and Kinoshita, S. (2002) "An analysis of target preferences of Escherichia coli outer membrane endoprotease OmpT for use in therapeutic peptide production: efficient cleavage of substrates with basic amino acids at the P4 and P6 positions" Biotechnol. Appl. Biochem. 36: 77-84

Overall, C.M. and Blobel, C.P. (2007) "In search of partners: linking extracellular proteases to substrates" Mol. Cell Bio. 8: 245-257

Paliouras, M., and Diamandis E.P. 2006. The kallikrein world: an update on the human tissue kallikreins. Biol. Chem. 387: 643−652.

Pimenta D.C., Juliano M.A., and Juliano L. 1997. Hydrolysis of somatostatin by human tissue kallikrein after the amino acid pair Phe-Phe. Biochem. J. 327: 27−30.

Pogson, M., Georgiou G. and Iverson, B.L. (2009) "Engineering next generation proteases." Current Opinion in Biotechnology 20: 390-397

Rawlings, N.D. and Barret, A.J. (1994) "Families of serine peptidases." Methods Enzymol. 244:19-61

Ricklin, D and Lambris, J.D. (2007) "Complement-targeted therapeutics" Nature Biotechnology 25, 1265 - 1275

Roth, D.A., Kessler, C.M., Pasi, K.J., Rup, B., Courter, S.G. and Karen, L. (2001) "Human recombinant factor IX: safety and efficacy studies in hemophilia B patients previously treated with plasma-derived factor IX concentrates" Blood 98(13): 3600-3606

Ruan, B., London, V., Fisher, K.E., Gallagher, T.D. and Bryan, P.N. (2008) "Engineering Substrate Preference in Subtilisin: Structural and kinetic analysis of a specificity mutant" Biochemistry 47: 6628-6636

Rupprecht, K.R., Gordon, G., Lundrigan, M., Gayda, R.C., Markovitz, A., Earhart, C. (1983) "OmpT: Escherichia coli K12 structural gene for protein a (3b)." J. Bacteriol. 153: 1104-1106

Sarullo, F.M., Americo, L., Di Pasquale, P., Castello, A., Mauri, F. (2000) "Efficacy of rescue thrombolysis in patients with acute myocardial infarction: preliminary findings." Cardiovas Drugs Ther 14:83-89

Scarisbrick, I.A., Towner, M.D., and Isackson P.J. (1997) "Nervous system-specific expression of a novel serine protease: regulation in the adult rat spinal cord by excitotoxic injury." J. Neurosci. 17: 8156−8168.

Schmulling, S., Grond, M., Rudolf, J., Heiss, W.D. “One-year follow-up in acute stroke patients treated with rtPA in clinical rountine.” Stroke 31: 1552-1554

Scott, J.K., and Smith, G.P. 1990. Searching for peptide ligands with an epitope library. Science 249: 386–390.

Sellamuthu, S., Shin, B.H., Lee, E.S., Rho, S.H., Hwang, W., Lee, Y.J., Han, H.E., Kim, J.I., Park, W.J. (2008) “Engineering of protease variants exhibiting altered substrate specificity.” Biochem. Biophys. Res. Commun. 371(1): 122-6

Simpon, D., Siddiqui, A.A., Scott, L.J. and Hilleman, D.E. (2006) “Reteplase: A review of its use in the management of thrombotic occlusive disorders” Am. J. Cardiovasc. Drugs 6(4): 265-285

Spillantini, M.G., Schmidt, M.L., Lee, V.M., Trojanowski, J.Q., Jakes, R., and Goedert, M. 1997. Alpha-synuclein in Lewy bodies. Nature 388: 839−840.

Stemmer, Willem P. C.; Crameri, Andreas; Ha, Kim D.; Brennan, Thomas M.; Heyneker, Herbert L. (1995) “Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides” Gene 164(1): 49-53.

Stewart, R.J., Fredenburgh, J.C., Leslie, B.A., Keyt, B.A. Rischke, J.A. and Weitz, J.I. (2000) “Identification of the mechanism responsible for the increased fibrin specificity of TNK-tissue plasminogen activator relative to tissue plaminogen activator” J. Bio. Chem. 275(14): 10112-10120

Stumpe, S., Schmid, R., Stephens, D..L., Georgiou, G. and Bakker, E.P. (1998). “Identification of OmpT as the protease that hydrolyzes the antimicrobial peptide protamine before it enters growing cells of Escherichia coli.” J. Bacteriol. 180(15): 4002-4006.

Sueiras, D.J., Jones, D.M., Ashworth, D., Horton, J., Evans, D.M., and Szelke, M. 1994. Cleavage of human kininogen fragments at Met-lys by human tissue kallikrein. Braz. J. Med. Biol. Res. 27:1935−1942.

Sugimura, K. and Higashi, N. (1988). “A novel outer-membrane-associated protease in Escherichia coli.” J. Bacteriol. 170(8): 3650-3654.

Sugimura, K. and T. Nishihara. (1988). “Purification, characterization, and primary structure of Escherichia coli protease VII with specificity for paired basic residues: identity of protease VII and OmpT. J. Bacteriol. 170(8): 5625-5632

Tanaka, T., Matsuzawa, H., and Ohta T. (1998) "Engineering of S2 site of aqualysin I; alteration of P2 specificity by excluding side chain." Biochemistry 37: 17402-17407

Tedesco, F., Nurnberger, W., Perssutti, S. (1993) "Inherited deficiencies of the terminal complement components." Int. Rev. Immunol. 10: 51-64

Turk B. (2006) "Targeting proteases: successes, failures and future prospects." Nature 5: 785-799

Vandeputte-Rutten, L., Kramer, R.A., Kroon, J., Dekker, N., Egmond, M.R. and Gros, P.(2001) "Crystal structure of the outer membrane protease OmpT from Escherichia coli suggests a novel catalytic site" Embo J. 20: 5033-5039

Varadarajan, N., Cantor, J.R., Georgiou, G., and Iverson, B.L. (2009) "Construction and flow cytometric screening of targeted enzyme libraries." Nat. Protoc. 4(6):893-901

Varadarajan, N., Gam, J., Olsen M.J., Georgiou, G., and Iverson, B.L. (2005) "Engineering of Protease Variants Exhibiting High Catalytic Activity and Exquisite Substrate Selectivity" Proc. Nat. Acad. Sci. 102, 6855-6860

Varadarajan, N., Rodriguez, S., Georgiou, G., and Iverson, B.L. (2008a) "Highly Active and Selective Endopeptidases with Programmed Substrate Specificities" Nat. Chem. Bio. 4:290-294

Varadarajan, N., Georgiou, G., and Iverson, B.L. (2008b) "An Engineered Protease that Cleaves Specifically after Sulfated Tyrosine" Angew. Chem. Int. Ed. 47:1-4

Varadarajan, N., Pogson, M., Georgiou, G. and Iverson, B.L. (2009) "Proteases that can distinguish among different post-translational forms of tyrosine engineered using multicolor flow cytometry" J. Am. Chem. Soc. 131: 18186-18190

Venekei, I.; Szilagyi, L.; Graf, L. and Rutter, W.J. (1996) "Attempts to convert chymotrypsin to trypsin." FEBS Lett 379: 143-147

Walsh, C. (2001) "Enabling the chemistry of life" Nature 409: 226-231

Webb, R.M. and Lundrigan, M.D. (1996) "OmpT in Escherichia coli correlates with severity of disease in urinary tract infections." Med Microbiol Letters 5: 8-14

Werner, R.G. (2001) "Improving natural principles with genetic engineering: TNK-tissue plasminogen activator." Arzneimittelforschung 51: 516-520

Wetsel, Rick A.; Kildsgaard, Jens; Haviland, David L. (2000) "Complement anaphylatoxins (C3a , C4a, C5a) and their receptors (C3aR, C5aR/CD88) as therapeutic targets in inflammation" Therapeutic Interventions in the Complement System, 113-153.

Willett, W. S., Brinen, L.S., Fletterick and R.J., Craik, C.S. (1996). "Delocalizing trypsin specificity with metal activation." Biochemistry 35(19): 5992-5998.

Yam, C. H., Siu, W.Y., Kaganovich, D., Ruderman, J.V. and Poon, R.Y. (2001) "Cleavage of cyclin A at R70-R71 by the bacterial protease OmpT." Proc. Nat. Acad. Sci. 98(2): 497-501.

Yamashiro, K., Tsuruoka, N., Kodama, S., Tsujimoto, M., Yamamura, Y., Tanaka, T., Nakazato, H., and Yamaguchi, N. 1997. Molecular cloning of a novel trypsin-like serine protease (neurosin) preferentially expressed in brain. Biochim. Biophys. Acta. 1350: 11−14.

Yang L., Manithody, C. and Rezaie, A.R. (2006) "Activation of protein C by the thrombin–thrombomodulin complex: Cooperative roles of Arg-35 of thrombin and Arg-67 of protein C" Proc. Nat. Acad. Sci. 103(4): 879-884

Yoon, H., Laxmikanthan, G., Lee, J., Blaber, S.I., Rodriguez, A., Kogot. J.M., Scarisbrick I.A., and Blaber, M. 2007. Activation profiles and regulatory cascades of the human kallikrein-related peptidases. Biochemistry 46: 5209−5217.

Yousef, G.M., Luo, L.Y., Scherer, S.W., Sotiropoulou, G., and Diamandis, E.P. 1999. Molecular characterization of Zyme/Protease M/Neurosin (PRSS9), a hormonally regulated kallikrein-like serine protease. Genomics 62: 251−259.

Yousef, G.M., and Diamandis, E.P. 2001. The new human tissue kallikrein gene family: structure, function, and association to disease. Endocr. Rev. 22: 184−204.

Yousef, G.M., Plymeris, M.E., Yacoub, G.M., Scorilas, A., Soosaipillai, A., Popalis, C., Fracchioli, S., Katsaros, D., and Diamandis, E.P. 2003. Parallel overexpression of seven kallikrein genes in ovarian cancer. Cancer Res. 63: 2223−2227.

Zhou. G. X, Chao L., and Chao J. 1992. Kallistatin: a novel human tissue kallikrein inhibitor. Purification, characterization, and reactive center sequence. J. Biol. chem. 267: 25873−25880.

Vita

Haixin Li was born in Tianjin, China, the daughter of Disheng Li and Hong Guan. She received a Bachelor of Medicine in Medicine (MD equivalent) from Tianjin Medical University, China, in 1997. The she continued her master study in National university of Singapore in 1999. In the autumn of 2004 she enrolled in the Graduate of School at the University of Texas at Austin.

Permanent address: #2-409, Zhishan Li, Ningxia Road, Heping District, Tianjin, China, 300020

This dissertation was typed by Haixin Li.