Tailoring the Catalytic Properties and Specificity of Subtilisin B. lentus via a Combined Site-Directed Mutagenesis and Chemical Modification Approach

by

Grace DeSantis

A Thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy Graduate Department of Chemistry University of Toronto

© Copyright by Grace DeSantis 1999
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
Tailoring the Catalytic Properties and Specificity of Subtilisin B. lentus via a Combined Site-Directed Mutagenesis and Chemical Modification Approach

by

Grace DeSantis

1999

A Thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy Graduate Department of Chemistry University of Toronto

Abstract

Chemically modified mutant enzymes (CMMs) of subtilisin Bacillus lentus (SBL) were designed to alter its specificity and catalytic properties. Modification of S156C, which is located toward the bottom of the S1 pocket and whose side chain is solvated, effected only small changes in its catalytic properties. Conversely, modification of S166C, which is located at the bottom of the S1 pocket and points into the pocket, effected larger changes. The S166C-S-CH2CH2SO3- CMM caused a 16-fold decrease in $k_{cat}/K_M$ with the standard suc-AAPF-pNA substrate, a 5-fold decrease in the binding of the 2,4-dichlorophenylboronic acid inhibitor and a 0.93 unit increase in the $pK_a$ of His64. The restricted S1 pocket of S166C-CMMs was demonstrated by the precluded binding and covalent attachment of a $p$-boronic acid benzophenone photoactivatable active-site directed inhibitor.

SBL prefers large hydrophobic $P_1$ substrate residues and the CMM strategy was exploited to confer a more universal $P_1$ specificity on S1. A large cyclohexyl group was introduced at position S166C to fill up the S1 pocket, causing a 2-fold improvement in $k_{cat}/K_M$ for the small $P_1$ residue suc-AAPA-pNA substrate and a 51-fold improvement in suc-AAPA-pNA/suc-AAPF-pNA selectivity compared to WT.
Negatively and positively charged groups were introduced at position S166C. A monotonic increase in $k_{\text{cat}}/K_m$ with the positively charged $P_1$ residue containing substrate suc-AAPR-pNA was effected by increasing negative charge at position 166. This culminated in a 9-fold improvement in $k_{\text{cat}}/K_m$ for the tri-negatively charged S166C-S-CH$_2$CH$_2$C(COO$^-$)$_3$ CMM and a 61-fold improvement in its suc-AAPR-pNA/suc-AAPF-pNA selectivity. The positively charged S166C-S-CH$_2$CH$_2$NH$_3^+$ CMM showed a 19-fold improvement in $k_{\text{cat}}/K_m$ with the negatively charged $P_1$ residue containing substrate, suc-AAPE-pNA and a 54-fold improvement in its suc-AAPE-pNA/suc-AAPF-pNA selectivity.

Modification of the $S_2$ pocket N62C residue resulted in seven of eleven CMMs, with higher than wild type (WT) levels of activity. A monotonic increase in activity with increasing side-chain size was observed culminating in a 3-fold improvement in $k_{\text{cat}}/K_m$ for N62C-S-c-CH$_2$C$_6$H$_{11}$. The indication from molecular modelling that this was due to increased acidity of the catalytic triad residue His64 was confirmed by pH-activity profile studies which revealed $pK_a$ decreases of up to 0.72 units. A linear correlation between the hydrophobicity of the introduced side chain and the observed $pK_a$ was apparent.
To my parents and Michael
“It is by logic that we prove, but by intuition that we discover.”

J. H. Poincaré
Acknowledgments

I am sincerely grateful to my research supervisor and mentor, Professor J. Bryan Jones, for his guidance, support and inspiration throughout the duration of this work. As well, I am grateful to Professor Marvin Gold for sharing his expertise in protein purification and characterization techniques and for excellent technical guidance.

I am very grateful to Professors Thomas T. Tidwell and Robert A. McClelland for the use of their photoreactors. I am particularly indebted to the members of my doctoral committee Professors Ronald Kluger, Andrew M. MacMillan and G. Andrew Woolley for their advice and encouragement. As well, I am deeply indebted to all of my coworkers in the research group, for their technical assistance as well as many inspiring scientific debates. In particular I would like to thank Drs. Ben Davis and Xiao Shang for the synthesis of methanethiosulfonate reagents, Drs. Per Berglund and Michele Stabile for technical guidance in the techniques of protein purification and modification, Dr. Erika Plettner for advice with the benzophenone work and Dr. Richard Martin for guidance with molecular modelling.

I am appreciative of graduate fellowships from the University of Toronto and the Natural Sciences and Engineering Research Council of Canada.

I am grateful to Genencor International Inc. for providing the WT and cysteine mutants of subtilisin. I am particularly grateful to Drs. Rick Bott, Christian Paech, Thomas Graycar, Colin Mitchinson and Chris Murray of Genencor International Inc. for helpful discussions and technical advice. Particular thanks are extended to Dr. Rick Bott for making available the PDB coordinates for SBL prior to publication, Dr. Christian Paech for his tutelage on tryptic-digestion and HPLC-ESI analysis during my visit to Genencor and for analyzing some of my samples on very short notice and Dr. Sue Middlebrook for peptide sequencing.

Many thanks go to my husband and soul-mate, Michael DiDonato, for his enduring moral support and understanding. Finally, I would like to thank my family for their encouragement, patience and support throughout my education.
**Part Three:**

3.0 Introduction of Unnatural Amino Acids or Side Chains  
3.1 Chemical Modification to Generate Semisynthetic Enzymes  
3.2 Unnatural Amino Acid Mutagenesis  
3.3 Peptide Synthesis and Fragment Ligation  
3.4 Site-Directed Mutagenesis Combined with Chemical Modification  
3.5 Current Study  
References  

Chapter 2: Site-Directed Mutagenesis Combined with Chemical Modification as a Strategy for Altering the Specificity of the S<sub>i</sub> Pocket of Subtilisin *B. lentus*  
Introduction  
Results  
Discussion  
Experimental  
References  

Chapter 3: Probing the Altered Specificity and Catalytic Properties of Mutant Subtilisin Chemically Modified at Position S156C and S166C in the S<sub>i</sub> Pocket  
Introduction  
Results and Discussion  
Experimental  
References  

Chapter 4: Benzophenone Boronic Acid Photoaffinity Labeling of SBL CMMs  
Introduction  
Results and Discussion  
Experimental  
References
Chapter 5: Towards Tailoring the Specificity of the S1 pocket of Subtilisin B. lentus: Chemical Modification of Mutant Enzymes as a Strategy for Removing Specificity Limitations

Introduction 121
Results 122
Discussion 123
Experimental 127
References 133

Chapter 6: Chemical Modification of the N62C Cysteine Mutant of Subtilisin B. lentus Can Create Better Catalysts than the Wild-Type Enzyme

Introduction 140
Results and Discussion 141
Experimental 147
References 149

Chapter 7: Chemical Modification at a Single Site Can Induce Significant Shifts in the pH Profiles of a Serine Protease

Introduction 151
Results and Discussion 152
Experimental 161
References 163

Perspective 166
References 169

Appendix

Representative Calculation of $k_{cat}$ and $K_m$ xvi
Representative Calculation of $K_i$ xvii
# List of Figures

## Chapter 1

| Figure 1.1 | Mechanism of Serine Proteases | 10 |
| Figure 1.2 | Active Site Nomenclature of the Serine Proteases | 16 |
| Figure 1.3 | Crystal Structure of SBL | 17 |
| Figure 1.4 | Lipase Catalyzed Reaction | 20 |

## Chapter 2

| Figure 2.1 | Active Site of SBL showing Ser156 and Ser166 | 67 |
| Figure 2.2 | Specificity Patterns for S156C and S166C CMMs | 70 |
| Figure 2.3 | Molecular Modelling for S166C CMMs | 72 |

## Chapter 3

| Figure 3.1 | pH-Activity Profiles for WT-SBL and S156C-S-CH₂CH₂SO₃⁻ | 86 |
| Figure 3.2 | Binding of Boronic Acid Inhibitors to S156C and S166C CMMs | 89 |
| Figure 3.3 | Modelling for ρ-Carboxyphenyl boronic acid | 90 |

## Chapter 4

| Figure 4.1 | Time Course for Photolysis of WT-SBL and BBP | 104 |
| Figure 4.2 | HPLC Trace of Native and BBP Crosslinked WT-SBL | 106 |
| Figure 4.3 | Proposed Mechanism of BBP Crosslinking | 107 |
| Figure 4.4 | Time Course for Photolysis of SBL CMMs and BBP | 108 |

## Chapter 5

| Figure 5.1 | Screen of S166C CMMs with suc-AAP-F/A/R/E-pNA Substrates | 125 |
| Figure 5.2 | Negative Charge at Position166 versus Activity | 129 |
| Figure 5.3 | Modelling of WT-SBL and S166C-S-CH₂CH₂NH₃⁻ with AAPE | 131 |
Chapter 6

Figure 6.1  Active Site of SBL Indicating Position of N62  141
Figure 6.2  Specificity Patterns for N62C CMMs  145

Chapter 7

Figure 7.1  Molecular Modelling for N62C CMMs  154
Figure 7.2  pH-Activity Profile for WT-SBL and N62C-S-c-CH₂C₆H₆,  155
Figure 7.3  ΔpKₐ of N62C CMMs  157
Figure 7.4  Linear Correlation Between pKₐ and Log P  159
# List of Tables

**Chapter 2**
Table 2.1 Kinetic Parameters for S156C and S166C CMMs (pH 7.5) 69

**Chapter 3**
Table 3.1 Kinetic Constants for S156C and S166C CMMs (pH 8.6) 84
Table 3.2 Inhibition Constants for S156C and S166C CMMs 87

**Chapter 4**
Table 4.1 Kinetic and Inhibition Constants for S166C CMMs with BBP 103
Table 4.2 ES-MS Data for BBP Treated WT-SBL 105
Table 4.3 ES-MS Data for BBP Treated SBL CMMs 110

**Chapter 5**
Table 5.1 Kinetic Evaluation of Altered S1 Pocket Specificity 126

**Chapter 6**
Table 6.1 Kinetic Parameters for N62C CMMs 144

**Chapter 7**
Table 7.1 Kinetic and pK_a Data for N62C CMMs 156

**Appendix**
Table 1 \( k_{\text{cat}}, K_m \) for S166C-S-CH_2-c-C_6H_11 xvi
Table 2 \( K_i \) for WT-SBL and 3-aminophenyl boronic acid xvii
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac</td>
</tr>
<tr>
<td>Ala (A)</td>
</tr>
<tr>
<td>Arg (R)</td>
</tr>
<tr>
<td>Asn (N)</td>
</tr>
<tr>
<td>Asp (D)</td>
</tr>
<tr>
<td>n-BuLi</td>
</tr>
<tr>
<td>CHES</td>
</tr>
<tr>
<td>Cys (C)</td>
</tr>
<tr>
<td>d</td>
</tr>
<tr>
<td>de</td>
</tr>
<tr>
<td>DMF</td>
</tr>
<tr>
<td>DMSO</td>
</tr>
<tr>
<td>DTT</td>
</tr>
<tr>
<td>ee</td>
</tr>
<tr>
<td>ESMS</td>
</tr>
<tr>
<td>EtOH</td>
</tr>
<tr>
<td>Et₂O</td>
</tr>
<tr>
<td>FPLC</td>
</tr>
<tr>
<td>Gln (Q)</td>
</tr>
<tr>
<td>Glu (E)</td>
</tr>
<tr>
<td>Gly (G)</td>
</tr>
<tr>
<td>His (H)</td>
</tr>
<tr>
<td>Ile (I)</td>
</tr>
<tr>
<td>Leu (L)</td>
</tr>
<tr>
<td>Lys (K)</td>
</tr>
<tr>
<td>m</td>
</tr>
<tr>
<td>MeOH</td>
</tr>
<tr>
<td>MES</td>
</tr>
<tr>
<td>Met (M)</td>
</tr>
<tr>
<td>MMTS</td>
</tr>
<tr>
<td>MWCO</td>
</tr>
<tr>
<td>NAD(P)</td>
</tr>
<tr>
<td>PAGE</td>
</tr>
<tr>
<td>PEG</td>
</tr>
<tr>
<td>Phe (F)</td>
</tr>
<tr>
<td>Pro (P)</td>
</tr>
<tr>
<td>s</td>
</tr>
<tr>
<td>SBL</td>
</tr>
<tr>
<td>SBN'</td>
</tr>
<tr>
<td>Ser (S)</td>
</tr>
<tr>
<td>Abbreviation</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>SDS</td>
</tr>
<tr>
<td>suc-AAPF-pNA</td>
</tr>
<tr>
<td>TCA</td>
</tr>
<tr>
<td>TFA</td>
</tr>
<tr>
<td>THF</td>
</tr>
<tr>
<td>Thr</td>
</tr>
<tr>
<td>TMS</td>
</tr>
<tr>
<td>Trp</td>
</tr>
<tr>
<td>Tyr</td>
</tr>
<tr>
<td>Val</td>
</tr>
<tr>
<td>v/v</td>
</tr>
<tr>
<td>w/v</td>
</tr>
</tbody>
</table>
Chapter 1
Introduction

Enzymes are now widely exploited as chiral catalysts in organic synthesis, with their application being driven mainly by the synthetic opportunities provided by the exquisite structural, regio- and stereospecificities of enzyme-catalyzed reactions.\(^1\) As a result of the asymmetric catalytic power offered by enzymes, the availability of information providing guidelines for the prediction of stereospecificity, and the current understanding of the mechanisms of enzyme catalysis, enzymes can now be used to access asymmetric transformations which would otherwise be impractical by any other means.

During the last century, the understanding of the complementary relationship between the enzyme active site and the substrate has developed extensively since it was first recognized and described by Emil Fischer as the “lock and key” model.\(^6\) It was later hypothesized by J. B. S. Haldane that enzymatic action hinges upon an induced fit since “the key does not fit the lock quite perfectly, but exercises a certain strain on it”\(^7\) and further developed to the currently accepted notion of enzyme-transition state complementarity by Linus Pauling.\(^8\) However, the continued development of highly selective enzyme catalysts requires a more thorough understanding of the factors which control enzyme activity, such as steric effects, electrostatic interactions, hydrogen bonding, hydrophobic effects, and \(\pi\)-stacking interactions.\(^9\)

Creating enzymes with new catalytic activities, and tailoring the specificity of existing ones to better accommodate unnatural substrates, is crucial to further increasing the scope of the applicabilities of enzymes in organic synthesis and thus is an area of considerable current research.\(^{10-15}\) Insights into the electrostatic\(^{13-15}\) steric,\(^{16}\) and hydrophobic,\(^{17}\) factors which govern enzyme-substrate interactions have been probed by site-directed mutagenesis studies. However, despite recent advances in understanding the intermolecular interactions which determine enzyme structural-, regio-, and stereo-specificity, truly rational tailoring of enzyme specificity
remains an elusive goal. The current study is aimed at developing a rapid and efficient methodology for tailoring enzyme specificity which is not limited by the 20-natural amino acid constraint of site-directed mutagenesis. The approach outlined entails the introduction of one cysteine residue at a key active site position via site-directed mutagenesis, which is then thioalkylated with an alkyl methanethiosulfonate reagent, \( \text{CH}_3\text{SO}_2\text{-S-R} \), such that: \( \text{WT} \rightarrow \text{Cys}_{\text{mutant}} + \text{H}_3\text{C-SO}_2\text{-S-R} \rightarrow \text{Cys-S-R} \) to give a chemically modified mutant enzyme (CMM) where R is infinitely variable. The serine protease subtilisin \textit{Bacillus lentus} (SBL) was selected as the representative enzyme for this strategy of incorporating unnatural amino acid moieties.
1.0 Enzymes

Enzymes are outstanding biological catalysts, providing rate enhancements of up to $10^{17}$ fold over the corresponding uncatalyzed reactions.\textsuperscript{18} Enzymes operate under mild conditions, and at physiological temperature and pH in aqueous solution.\textsuperscript{5} As a result of the evolutionary pressures imposed by the complexity of biological systems, they are exquisitely selective with respect to the structure and stereochemistry of the substrate, and hence the product. Enzymes can be isolated from mammalian, bacterial, microbial, plant or fungal sources. Of the more than 3000 known enzymes,\textsuperscript{2,19} several hundred are commercially available as pure crystalline products, as whole cell preparations,\textsuperscript{20} or in immobilized\textsuperscript{21} or crosslinked forms.\textsuperscript{22}

Enzymes are composed of unbranched polypeptide chains consisting of the 20 natural L-\(\alpha\)-amino acids linked together by amide bonds between the \(\alpha\)-carboxyl group of one residue and the \(\alpha\)-amino group of the next.\textsuperscript{9} This generates a primary structure defined by the amino acid sequence. The genetically encoded sequence determines the secondary structural components of \(\alpha\)-helix, \(\beta\)-sheet and \(\beta\)-turn, which in turn give rise to the overall fold of the enzyme or the tertiary structure.\textsuperscript{23} In addition to noncovalent chemical interactions, the tertiary fold may be further stabilized by disulfide bridges.\textsuperscript{23} It is the unique tertiary fold of enzymes which controls the spatial arrangement of the amino acid residues and dictates enzyme specificity by defining the active site where substrate binding and reaction occur.\textsuperscript{9}
1.1 Enzymes in Organic Synthesis

In addition to the vital role of enzymes in biological systems, enzymes catalyze a broad spectrum of reactions in vitro which are of interest to organic chemists.\textsuperscript{1-5} The commercial applications of enzymes are for detergent, dairy product, and starch processing, and to a lesser extent for fine chemical synthesis, and they have an annual world market of $900 million dollars.\textsuperscript{24}

Enzymes are employed in organic synthesis applications out of convenience, necessity and opportunity. The milder, more selective, reaction conditions employed for enzymatic reactions minimize problems of isomerization, racemization and epimerization often encountered in traditional chemical synthesis.\textsuperscript{5} Furthermore, since enzymatic reactions are optimal in aqueous media they offer environmental advantages over current organic catalysts, which usually contain toxic transition metal complexes,\textsuperscript{25} and that must be employed in organic solvents which are increasingly becoming more difficult to dispose of. Furthermore, enzymes are intrinsically compatible with one another and consequently a number of enzymes can be used to accomplish multi-step reaction sequences in a single reaction vessel.\textsuperscript{26} In contrast, one-pot use of multiple non-biological catalysts are relatively limited.\textsuperscript{25}
1.2 Enzyme Classes

The International Biochemistry Union has adopted a classification which
divides the enzymes into six groups based upon their function, as described below.¹⁹

1. **The Oxidoreductases** catalyze oxidation-reduction reactions involving
oxygenation, (C-H → C-OH), or the net removal or addition of hydrogen atom
equivalents, as for (CH(OH) ⇌ C=O and CH-CH ⇌ C=C).

2. **The Transferases** mediate the transfer of groups such as acyl, sugar,
phosphoryl, aldehyde and ketone moieties from one molecule to another.

3. **The Hydrolases** hydrolyze a broad range of functional groups, including
glycosides, anhydrides and esters, as well as amides, peptides and other C-N
containing functions.

4. **The Lyases** catalyzes the addition of HX to double bonds such as C=C, C=N and
C=O, as well as the reverse reaction.

5. **The Isomerases** catalyze various isomerizations, including C=C bond migration,
cis-trans isomerization, and racemization.

6. **The Ligases** are responsible for the formation of C-O, C-S, C-N, C-C and
phosphate ester bonds and are also commonly referred to as synthetases.

There is an enzyme-catalyzed counterpart for almost every type of organic
reaction including key synthetic reactions (Scheme 1.1) such as the aldol
condensation, asymmetric epoxidation reactions, Baeyer-Villiger oxidation,
Claisen rearrangement and recently even the Diels-Alder reaction. However, the
class of enzymes most widely applied to organic synthesis applications are the hydrolases. Members of the hydrolase family which have been exploited extensively include, lipases, esterases and proteases whose in vivo function is the hydrolysis of fatty acid esters, esters, and proteins respectively. Amongst these, the serine protease (EC 3.4.21) subclass, of which subtilisin is a member, has been utilized extensively for both preparative applications and for mutagenesis studies. Furthermore, subtilisin can under the appropriate conditions, be employed as an esterification catalyst (Scheme 1.1d).
Scheme 1.1

(a)

\[
\begin{align*}
\text{DHAP} & \quad \text{Aldolase} \\
\text{Rha 1-P} & \quad \text{Aldolase} \\
\text{Fuc 1-P} & \quad \text{Aldolase}
\end{align*}
\]

(b)

\[
\begin{align*}
\text{Alcaligenes sp.} & \quad \text{Cl} & \quad \text{OH} & \quad \text{OH} & \quad \rightarrow & \quad \text{99.4\% ee} \\
\text{Pseudomonas sp.} & \quad \text{Cl} & \quad \text{OH} & \quad \text{OH} & \quad \rightarrow & \quad \text{99.3\% ee}
\end{align*}
\]

(c)

\[
\begin{align*}
\text{cyclohexanone oxygenase} & \quad \text{OH} & \quad \rightarrow & \quad \text{98\% ee}
\end{align*}
\]

(d)

\[
\begin{align*}
\text{Subtilisin} & \quad \text{DMF} \\
\text{OH} & \quad \text{OH} & \quad \rightarrow & \quad \text{OH}
\end{align*}
\]
1.3 Serine Proteases

The sub-class of hydrolases known as the proteases, which catalyze the hydrolysis of amide linkages in proteins in vivo and catalyze the hydrolysis of both amide and ester bonds in vitro have been utilized extensively in organic synthesis.\(^{2-3}\) The serine proteases (EC 3.4.21) are an endopeptidic sub-class of this group requiring no cofactors, and which are characterized by a uniquely reactive serine hydroxyl which is part of the Ser - His - Asp active site catalytic triad, first identified in chymotrypsin.\(^{37}\) The spatial arrangement of the catalytic triad is conserved amongst serine proteases despite the lack of sequence homology and differences in overall fold.\(^{38}\) This reinforces the crucial role of the triad in the catalytic mechanism which has been thoroughly investigated and is well characterized, as noted in Figure 1.1.\(^{9,39-40}\) The imidazole side chain of histidine increases the nucleophilicity of the serine hydroxyl by acting as a general base catalyst. The buried carboxyl side chain of aspartic acid, which differs slightly in position between proteases,\(^{41}\) serves to raise the pK\(_a\) of imidazole through its negatively charged carboxylate side chain, as well as to confine its location. Serine proteases catalyze ester and amide hydrolysis in a pH dependent manner.\(^{42-45}\) Amide and ester substrates are hydrolyzed by the serine proteases via the acyl enzyme mechanism.\(^{39-40,46-48}\)
Figure 1.1 Mechanism of Serine Proteases: The enzyme (E) and substrate (S) first associate to form a noncovalent enzyme-substrate (ES) complex held together by physical interactions. This provides the proper alignment of reactants and catalytic groups so that chemical transformation can occur. A tetrahedral covalent O-acyl intermediate then results from the attack of the serine Oγ hydroxyl on the substrate carbonyl. The negatively charged tetrahedral intermediate formed is stabilized by a series of hydrogen bonds in the "oxyanion hole" binding region of the enzyme. This high-energy tetrahedral intermediate then collapses to yield the acyl-enzyme (AcE) releasing the amine or alcohol product (P₁), which rapidly diffuses away. The acyl-enzyme, which is common to both ester and amide substrates, is subsequently hydrolyzed, with a water molecule acting as the nucleophile, thereby regenerating the native enzyme and releasing the carbonyl component of the peptide (P₂) as a carboxylate anion. The acylation step is usually rate limiting in the hydrolysis of amide bonds. However, deacylation of the acyl enzyme intermediate is usually rate limiting for ester hydrolysis.⁹,³⁹-⁴⁶
1.4 Enzyme Kinetics

For many enzymes, including the serine proteases, a predictable kinetic pattern is observed which corresponds to the kinetic expressions developed initially by Brown and further refined by Michaelis and Menten. The initial formation of the Michaelis complex (ES) is assumed to be rapid and reversible with no chemical changes taking place. The chemical processes then occur in subsequent steps, generating the product P, and releasing the enzyme E as shown in Scheme 1.2.

**Scheme 1.2**

\[
\text{E} + \text{S} \xrightarrow{K_s} \text{ES} \xrightarrow{k_{\text{cat}}} \text{P} + \text{E}
\]

The Michaelis-Menten rate equation (Equation 1.1), where \( K_m = K_s \) is solved based on the following assumptions: (1) The ES complex is in thermodynamic equilibrium with the free enzyme and substrate. This is true only if the formation of product from the ES complex is very fast relative to the rate of the decomposition of the ES complex to give enzyme and substrate. (2) [P] is insignificant compared to [S] and the substrate concentration is constant, which is possible by considering only the initial stage of the reaction. (3) The rate expression is valid only for single substrate system where [S] \( \gg \gg \) [E].

\[
\nu = \frac{[E]_0[S]k_{\text{cat}}}{[S] + K_s}
\]  

Equation 1.1

The true Michaelis constant, \( K_s \), which is the dissociation constant of the ES complex, reflects the strength of enzyme substrate binding, with lower values representing better binding. Under these conditions, \( K_m = K_s \). The turnover number, \( k_{\text{cat}} \), is the apparent first order rate constant for the conversion of the enzyme-substrate complex to product. The hyperbolic Michaelis-Menten equation has two simplified forms. At low substrate concentration, a linear
relationship is observed and can be expressed as \( v = [E]_0[S](k_{cat}/K_M) \), where \( k_{cat}/K_M \) is the apparent second order rate constant which relates the reaction rate to the concentration of free rather than total enzyme. At low substrate concentration the enzyme is largely unbound, with \( [E] \approx [E]_0 \). The specificity constant, \( k_{cat}/K_M \), is used as the best measure of the efficacy and specificity of the enzymatic reaction. At high substrate concentrations, all of the enzyme is in the ES complex and the velocity is independent of substrate. Saturation kinetics is observed and Equation 1.1 simplifies to \( v = k_{cat}[E]_0 \), which is termed the maximal velocity (\( V_{max} \)), as illustrated in Scheme 1.3.

Scheme 1.3

In 1925, Briggs and Haldane\(^5\) introduced the concept of a steady state in the enzymatic reaction and showed that the equilibrium concept of Michaelis and Menten was not essential. In addition, the Michaelis-Menten scheme may be extended to cover a variety of cases in which additional intermediates, covalently or noncovalently bound, occur on the reaction pathway, as outlined in Scheme 1.4.\(^5\)

Scheme 1.4

\[
E + S \xrightarrow{k_1} ES \quad \xrightarrow{k_{-1}} E + S \\
ES \xrightarrow{k_2} EAc \xrightarrow{k_3} E + P_2 \\
\]

In Scheme 1.3, the initial rate \( v \) is plotted against substrate concentration \( [S] \). The concentration \( [S] = K_m \) is a critical point where the rate is half of the maximum, \( \frac{v_{max}}{2} = K_m \).
While the Michaelis-Menten equation, as described above, still applies, $K_M$ and $k_{cat}$ are now combinations of various rate and equilibrium constants, as shown in Equation 1.2. In particular the acylenzyme intermediate (EAc), resulting from serine protease hydrolysis of amide or ester bonds, is incorporated into Scheme 1.4. In the limiting case when $k_1 \gg k_2$, (Scheme 1.4), the Michaelis constant approximates to $k_1/k_2$ and $K_M = K_s$.52

$$v = \frac{[E]_o[S]k_{cat}}{[S] + K_M}$$  \hspace{1cm} \text{Equation 1.2}$$

where, 

$$k_{cat} = k_2k_3/(k_2 + k_3)$$
$$K_M = (k_1/k_2)((k_2)/(k_2 + k_3))$$

A complementary approach to substrate kinetic analysis for probing enzyme specificity is the use of competitive reversible transition state analogue inhibitors. A competitive reversible inhibitor (I) binds to the enzyme active site (E) via noncovalent interactions and forms an EI complex which prevents substrate binding, as shown in Scheme 1.5.

**Scheme 1.5**52-53

![Scheme 1.5](image)
In such situations, the affinity of the enzyme for the inhibitor (I) is defined by the dissociation constant, $K_i$, of the enzyme-inhibitor complex, ([EI]). The kinetic manifestation of competitive inhibition is an apparent increase in the Michaelis constant, $K_M$, by a factor of $(1 + [I]/K_i)$ while the turnover number, $k_{cat}$, remains unaltered, as shown in Equation 1.3. \(^52-53\)

$$v = \frac{[E]_0[S]k_{cat}}{[S] + K_M(\text{apparent})}$$  \hspace{1cm} \text{Equation 1.3} \(^52-53\)

where, $k_{cat} = k_2k_3/(k_2 + k_3)$

$K_M(\text{apparent}) = (1 + [I]/K_i)K_M$
1.5 Subtilisin \textit{Bacillus lentus} (SBL)

The alkaline serine protease, subtilisin \textit{Bacillus lentus} (SBL, EC 3.4.21.14) was chosen as a representative serine protease for the current study since it is a well characterized enzyme and is of synthetic\textsuperscript{12,54} as well as of industrial\textsuperscript{55} interest. Furthermore, SBL's high resolution crystal structure has been solved,\textsuperscript{56-58} it has been cloned, overexpressed and purified,\textsuperscript{59} and its kinetic behaviour well characterized.\textsuperscript{60-63}

The amount of available information on subtilisins, combined with their industrial and synthetic applications, makes them ideal model systems for protein engineering.\textsuperscript{64-66} Subtilisins are a class of serine endopeptidases secreted into the external medium by a wide variety of gram-positive bacteria, such as \textit{Bacillus}.\textsuperscript{39} Subtilisins are synthesized as pre-proenzymes, which are then translocated over a cell membrane via the pre-peptide (or signal peptide), and finally activated by cleavage of the pro-peptide.\textsuperscript{67} The subtilisin proteases are single-domain molecules with no disulfide bridges, and with a hemispherical shape of 40 Å diameter.\textsuperscript{57} Their active site is located on the flat surface of the hemisphere.\textsuperscript{57} The core of the protein is composed of a twisted parallel β-sheet with α-helices running anti-parallel to the β-sheet and parallel to each other.\textsuperscript{38} SBL, which is secreted by an alkalophilic bacterium, has a high isoelectric point, pI = 11.1,\textsuperscript{68} high thermal stability at alkaline pH\textsuperscript{69} and has maximal catalytic activity between pH 8 and pH 12.\textsuperscript{57} SBL has 13 positively charged lysine and arginine residues and 10 negatively charged aspartate and glutamate residues, for a total of 25 charged groups, including the N- and C-termini. Despite having fewer charged residues than subtilisin BPN', SBL has seven salt bridges, with two additional salt bridges being possible at high pH, which is two more than subtilisin BPN' s five.\textsuperscript{57} SBL has two Ca\textsuperscript{2+} binding sites which do not have a catalytic role but rather contribute to its thermal stability\textsuperscript{70} which in turn increases autolytic resistance.\textsuperscript{71} The high affinity Ca\textsuperscript{2+} binding site is seven coordinate and exhibits full occupancy, while the lower affinity site is five coordinate and exhibits an occupancy of 0.6.\textsuperscript{57} The crystal structure of SBL also reveals a linear chain of five
water molecules, linked by a network of hydrogen bonds and extending from the surface of the protein into the active site, and whose function appears to be structural rather than catalytic.\textsuperscript{57}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1_2.png}
\caption{Active Site Nomenclature of the Serine Proteases: The substrate residues are denoted $P_1$, $P_2$ etc. and the active site pockets to which they bind are denoted $S_1$, $S_2$ etc. The residues on the carbonyl side of the scissile bond are denoted $P_1'$, $P_2'$ etc.\textsuperscript{72}}
\end{figure}

The binding region of SBL has been described as a surface channel which can accommodate at least eight residues, $P_5$ to $P_3'$ in the $S_5$ to $S_3'$ pockets,\textsuperscript{60} where the scissile bond is between the residues denoted $P_1$ and $P_1'$, as shown in Figure 1.2.\textsuperscript{72} While SBL, with its 269 residues, is homologous to subtilisin BPN' (62% identical residues) a significant difference between SBL and other subtilisins is a four residue deletion in its $S_1$ pocket, and two single deletions at positions equivalent to residues 36 and 56 in subtilisin BPN'.\textsuperscript{57} The $S_1$ primary specificity determining pocket of SBL is composed of residues 155 to 162, 166 and 126 to 128, and exhibits a preference for large hydrophobic residues such as the benzyl side chain of Phe (Figure 1.3).\textsuperscript{60,62} The catalytic triad residues of subtilisin are Asp32, His64 and Ser221 (BPN' numbering). Residue Asn155 is located in the so called “oxyanion-hole” and helps to stabilize the oxyanion generated in the tetrahedral transition state.\textsuperscript{73}
Figure 1.3 Crystal Structure of SBL: Showing binding pockets $S_2$ to $S_1$ and the catalytic triad residues, Ser221, His64 and Asp32.\textsuperscript{56}
2.0 Altering Enzyme Activity

While the focus of the current study is to expand the specificity and operational parameters of SBL, altering enzyme activity has already been broadly applied to overcome the limitations of biocatalysts. The challenges posed by the narrow substrate specificity, instability and high cost of some enzymes have been, and continue to be, addressed through several approaches, including covalent modification, noncovalent manipulation, and protein engineering of existing catalysts, and through the de novo design of biocatalysts and of biocatalyst mimics. In addition to optimizing biocatalysts for organic synthesis applications, these approaches to altering enzyme activity have increased our collective understanding of protein structure, stability and function and have yielded insights into the factors which govern biocatalysis. The representative examples of these approaches provided below are illustrative of their capacities and of their deficiencies.

2.1 Noncovalent Modification of Enzymes

One of the simplest methods of altering the specificity, reactivity and stability of enzymes is by mechanical manipulation. This involves the use of techniques which do not alter primary structure but nonetheless alter enzyme specificity through changing the properties of the interaction between the biocatalyst and substrate. These include changes in solvent properties, molecular imprinting, substrate engineering, temperature effects, and enzyme immobilization.

For preparative transformations, the use of enzymes in organic solvents has received considerable attention since it addresses the challenges posed by insolubility of some substrates in aqueous solution, and in some cases minimizes side-reactions such as hydrolysis, racemization, and decomposition. An example of increased product optical purity, due to diminished chemical side-reactions, is the preparation of cyanohydrins by the mandelonitrile lyase-catalyzed addition of
hydrogen cyanide to various aldehydes in ethyl acetate. An important advantage of employing enzymes in organic solvents is the possibility to shift thermodynamic equilibria, for example to favour peptide synthesis over hydrolysis. In this way, by using hydrolases in organic solvents, esters, lactones and amides or peptides can be chemo- regio- and enantio-selectively synthesized (Figure 1.1d, page 8). In addition, the use of hydrophobic organic solvents can reverse the usual proteolytic reaction enantiopreference from L-amino acid to D-amino acid substrates. Supercritical fluids have also been exploited as solvents for enzymatic reactions and have been found to increase activity. Furthermore, enzyme specificity may be modulated by altering the pressure of supercritical fluids which in turn alters the solvent dielectric constant, and under higher pressure the solvent becomes increasingly hydrophilic. For example, the stereoselectivity of subtilisin was increased by increasing the pressure in fluoroform. In addition, biocatalyst thermal stability may be enhanced in organic solvents due to increased structural rigidity. The available crystal structures and models of biocatalysts in organic solvents help to provide insights into the molecular basis for the altered specificities.

Another approach toward altering enzyme activity is by immobilization. For example, enzyme adsorption on silica, Celite, potassium phosphate, and on modified controlled-pore glass (CPG) has afforded increased reaction rates for preparative biotransformations in organic solvents. Complementary to physical methods of adsorption is chemical immobilization, in which functional groups on the surface of the protein, such as amino and carboxyl groups, are used for attachment to a support. For example, poly(ethylene glycol) has been exploited extensively for enzyme immobilization and has generated enzymes with increased solubility and stability in organic solvents. In addition, the commercially available cross-linked enzyme crystals (CLECs), which are generated by crosslinking with glutaraldehyde, boast improved biocatalyst stability. Enzymes have also been entrapped within gels or polymeric substances. However, a recent advance in chemical immobilization is to introduce a unique cysteine residue at a position away
from the active site and then use this unique sulfhydryl to link the enzyme onto a solid support, thereby effecting higher activity than immobilization via random residues.\textsuperscript{111}

A conceptually pleasing method to enhance or alter biocatalyst specificity is through molecular imprinting or bioimprinting. This approach exploits the enzyme's flexibility in aqueous solution and its relative inflexibility in organic solvents to first "mold" the enzyme active site around a substrate mimic in aqueous solution and then lock-in the shape by lyophilization. The bioimprinted enzyme is then used in an organic solvent where it "remembers" the structure of the substrate-mimic generating rate enhancements,\textsuperscript{112} altering specificity,\textsuperscript{113-114} and preventing inactivation.\textsuperscript{115} For example, chymotrypsin was made to accept D-amino acids by bioimprinting.\textsuperscript{113-114} A variation on this general theme is the creation of catalytic conformationally modified proteins (CCMP) which are prepared by bioimprinting and then crosslinking a noncatalytic protein such as bovine serum albumin producing a CCMP which is more active than the initial unmodified protein.\textsuperscript{116}

![Chemical structure](image)

**Figure 1.4 Lipase Catalyzed Reaction:** An example of an irreversible Amano Lipase catalyzed transesterification reaction using the vinyl acetate acyl donor.\textsuperscript{117}

Substrate engineering can also be employed to alter the course of a biocatalytic reaction. For example, the use of vinyl or isopropenyl acetates as substrates for transesterification reactions renders the reaction irreversible since the reaction byproducts are unstable alcohols which tautomerize to nonnucleophilic aldehydes (Figure 1.4).\textsuperscript{117,118} Alternatively, highly activated leaving groups such as trichlorethyl, trifluoroethyl, cyanomethyl, thiol, enol and oxime esters can be used to make the reactions kinetically virtually irreversible.\textsuperscript{119,120} Alternate substrate binding modes have also been exploited to change the stereochemical course of reaction.\textsuperscript{121}
2.2 Site-Directed Mutagenesis and Rational Design

The development of site-directed mutagenesis, which permits the routine replacement of any amino acid residue in a protein whose gene has been cloned, with any one of the other 19 natural amino acids, has caused an explosion in attempts to alter enzyme properties.\textsuperscript{122,123} The most intellectually gratifying approach to mutagenesis is that of rational design. However, since it requires a detailed understanding of the enzyme's catalytic mechanism, substrate specificity determinants and tertiary structure, it is only suited to enzymes that are very well understood.\textsuperscript{66}

Mutagenesis studies have not only yielded enzymes with altered specificities, but have contributed significantly to understanding the electrostatic,\textsuperscript{13-15,124-127} steric,\textsuperscript{16,128-133} and hydrophobic\textsuperscript{16-17,134} factors which govern enzyme-substrate interactions. Many of the design strategies are based on structural elements which are exploited in nature by native enzymes with the desired specificity. For example, the P\textsubscript{1} and P\textsubscript{2} substrate specificity of subtilisin was changed from one that prefers large hydrophobic residues to one that prefers positively charged residues by the introduction of negatively charged residues in its S\textsubscript{1} and S\textsubscript{2} pockets. This was based on the structure of kex2 which exhibits the desired substrate specificity.\textsuperscript{13,135-136} The subtle nuances of the rational design approach are apparent from the failure of a similar strategy to change the substrate specificity of trypsin from one that prefers positively charged P\textsubscript{1} residues to one that prefers large hydrophobic ones. Since the S\textsubscript{1} pocket residue, 189, is Asp for trypsin but Ser for chymotrypsin, it was hoped that the designed D189S mutant of trypsin would confer the desired chymotrypsin-like specificity onto trypsin,\textsuperscript{137} however, this approach failed. Instead, several amino acids including some which do not contact the substrate had to be exchanged before the specificity change was achieved.\textsuperscript{138-140} Conversely, attempts to confer trypsin-like S\textsubscript{1} specificity onto chymotrypsin by the S189D mutation were also unsuccessful.\textsuperscript{141}
Rational design has also been applied successfully to the oxidoreductases. For example, the stereoselectivity of a L-lactate dehydrogenase (LDH), which is one of the most stereospecific enzymes known, was significantly reversed. The designed Ile240Lys/Arg171Tyr LDH double mutant yielded up to 2.3% of D-lactate by promoting an altered substrate binding mode, which represents a truly remarkable > 500-fold switch in stereospecificity-preference. In addition, LDH's specificity toward substrates with positively changed side chains was improved by the introduction of negatively changed amino acid residues such as Asn102Asp/Glu in the substrate binding site. Furthermore, LDH's specificity toward substrates with large or branched side chains was improved by replacing Gln102 by smaller amino acid residues such as Asn. Similarly, the A95G mutant of L-lactate oxidase (LOX) was successfully designed to convert LOX to a long chain α-hydroxyacid oxidase. A similar design strategy was applied to *Rhizopus delemar* lipase. Employing a computer generated model of the lipase, the V209W/F11 double mutant was successfully designed to improve the specificity for short and medium chain length fatty acids and preclude binding of long chain fatty acids. This example constitutes the first successful application of rational design to a lipase.

Cofactor tailoring can also be accomplished by rational design. By comparison of the active site residues and structure of isopropylmalate dehydrogenase (IMDH) which exhibits a 100-fold NAD/NADP preference, the cofactor requiring specificity of isocitrate dehydrogenase (IDH) was switched from its natural 7000-fold NAD⁺/NAD⁺ preference to a 200-fold NAD⁺/NAD⁺ preference by mutagenesis of seven amino acids which were selected on the basis of X-ray structures and molecular modelling. This example is particularly important since NAD⁺ is 5-fold more expensive than NAD⁺.
2.3 Random Mutagenesis

An alternate approach to site-directed mutagenesis for altering enzyme specificity is random mutagenesis, which does not require prior understanding of specificity determinants nor knowledge of the structure. The approaches applied to the preparation of libraries of mutants include the use of chemical mutagenesis,\textsuperscript{149-153} random oligonucleotide primers,\textsuperscript{154-156} an error-prone polymerase chain reaction which may be induced by the use of manganese instead of magnesium\textsuperscript{156} or by an engineered error-prone polymerase,\textsuperscript{157} mutagenic nucleotide analogues,\textsuperscript{158-159} DNA shuffling,\textsuperscript{160-161} or the use of an \textit{E. coli} mutator strain.\textsuperscript{162-163} For the purposes of altering enzymatic properties, these randomization techniques are usually coupled to a suitable high throughput screen\textsuperscript{164-168} or to positive genetic selection.\textsuperscript{64} The desired enzyme properties can be optimized through several iterations of mutagenesis and selection and the approach is therefore termed directed evolution. Directed evolution permits rapid sampling of sequence space which is huge since \(20^X\) variants are possible where \(X = \) number of amino acids in the protein. The random mutagenesis approach is particularly well suited to systems that are not well characterized.\textsuperscript{169} However, the success of this strategy to yield a biocatalyst with the desired properties hinges on the design of an appropriate screen.

This methodology has been applied to a subtilisin from \textit{Bacillus subtilis} in order to improve its catalytic activity in organic solvents.\textsuperscript{170} In this way, an enzyme was generated that hydrolyzes the standard suc-AAPF-pNA peptide substrate 256 times more efficiently than does wild-type subtilisin in 60\% DMF.\textsuperscript{170-172} In another illustration, the enantioselectivity of a lipase from \textit{Pseudomonas aeruginosa} which showed an ee of only 2\% in favour of the S-enantiomer of 2-methyldecanoic acid \(p\)-nitrophenol ester was optimized. After only four generations of directed evolution, an enzyme was obtained that gave the S-enantiomer in 81\% ee.\textsuperscript{173} Random mutagenesis has also been used to enhance thermal stability of, for example, cholesterol oxidase from \textit{Streptomyces},\textsuperscript{174} and of a lipase from \textit{Pseudomonas}
Directed evolution has also been employed to expand substrate specificities. After five rounds of selection, a $10^5$-fold increase in the catalytic efficiency of aspartate aminotransferase for $\beta$-branched 2-oxo acids, and a 30-fold decrease for the native substrate, was achieved.\(^\text{176}\)

### 2.4 De Novo Protein Design

The approaches described above recognize that naturally available enzymes represent a valuable starting point for the optimization of biocatalysts. However, a highly desirable goal is the *de novo* design of a biocatalyst with the needed property. *De novo* biocatalyst creation entails the design and generation of a protein scaffold based on the still limited but rapidly growing understanding of the relationship between sequence and structure.\(^\text{23,177}\) Unfortunately, while a few proteins or peptides with secondary structure have been designed, reliable *de novo* design of biocatalysts remains elusive.\(^\text{178}\) Nevertheless, recently a 33-residue polypeptide was designed which was found to catalyze peptide ligation with a rate enhancement of $10^4$ compared to the uncatalyzed reaction, and with a 10-fold diastereoselectivity.\(^\text{179}\) While this approach is still in its infancy, it can clearly be expected to make dramatic progress as improved design algorithms emerge.\(^\text{180}\)

An approach to biocatalyst design that complements the *de novo* approach is to graft an alternate catalytic machinery onto an existing "protein scaffold" possessing the desired specificity, thus generating a hybrid enzyme.\(^\text{136,181}\) For example, the *E. coli* cyclophilin, which binds proline containing peptides was converted into a proline specific protease by mutation of three amino acids in its substrate binding cleft to form a triad resembling that of the serine proteases. The resultant protease enhanced the hydrolysis reaction $10^5$-fold compared to the uncatalyzed reaction.\(^\text{182}\)
2.5 Catalytic Antibodies

In the quest for better biocatalysts, chemists have left no stones unturned. In this regard, the diversity and chemical potential of the immune system has been recognized and exploited. The speed and diversity of the immune response to antigen challenge has been exploited to generate catalytic antibodies or abzymes. Abzymes are prepared by immunizing a host with a chemically stable hapten, coupled to a carrier protein which resembles the transition state of the desired reaction. This elicits the production of isolatable antibodies which bind the transition state analogue hapten. As was predicted early on, some of these possess the capability of promoting catalysis by stabilizing the reactive transition state. Abzymes that catalyze a wide variety of transformations, including the aldolase reaction, the Diels-Alder reaction, the Claisen reaction, peptide hydrolysis and even peptide ligation have been generated. While abzymes are usually plagued by low reaction rates and product inhibition, they are potentially promising.

2.6 Enzyme Mimics

Chemists have also adopted a minimalist approach to catalysis in the development of supramolecular or host-guest chemistry to create small molecules which can mimic the catalysis of naturally occurring enzymes or even attempt to create catalytic abilities not observed in nature. In this search, functionalized and unfunctionalized cyclodextrins, cyclic porphyrin trimers, functionalized Cram ethers, cryptands, and a wide variety of other macrocyclic synthetic hosts have been evaluated as catalysts. An early example by Bender demonstrated the elegance of this approach, grafting a carboxylate group, an imidazoyl group and a hydroxyl group onto the rim of a β-cyclodextrin, the hydrophobic binding pocket, and the catalytic triad of the serine proteases was emulated, yielding an artificial enzyme with approximately the same catalytic activity as chymotrypsin. More recently,
supramolecular structures which could catalyze the Diels-Alder reaction have been sought, such as a porphyrin trimer which can catalyze an exo-selective Diels-Alder reaction.\textsuperscript{198}

2.7 Nucleic Acid Enzymes

Since the initial observation of the autocatalytic property of RNA almost 20 years ago,\textsuperscript{199-200} ribozymes have been recognized as versatile and efficient catalysts.\textsuperscript{201} In addition to their natural nuclease activity, ribozymes have been shown to catalyze the synthetically important Diels-Alder\textsuperscript{202} and peptide ligation\textsuperscript{203} reactions. Furthermore, although DNA enzymes have not yet been observed in nature, deoxyribozymes have recently been developed \textit{in vitro}.\textsuperscript{204-206} For example, a histidine-dependent deoxyribozyme, which catalyzes the cleavage of an RNA phosphodiester bond using the amino acid histidine as a cofactor and producing a rate enhancement of over $10^6$ relative to the uncatalyzed reaction, was identified by \textit{in vitro} selection.\textsuperscript{207} Despite their limited substrate specificity and low reaction rates, nucleic acid enzymes are expected to develop into an interesting alternative to peptide based enzymes since they can be very easily optimized by randomization.
3.0 Introduction of Unnatural Amino Acids or Side Chains

The importance and challenges of altering enzymatic activity are evident from the vastly varied approaches which have been applied toward accomplishing this goal. From optimizing existing enzymes via mutagenesis techniques, noncovalent modifications, exploiting the diversity of the immune response, generating small enzyme mimics, or developing nucleic acid based enzymes, enormous growth has been achieved in the applicability and robustness of enzymes for organic synthesis applications. However, the protein-based approaches described for altering enzyme activity, are limited to the 20 naturally occurring amino acids. In order to more fully understand the factors which control enzyme specificity and to permit the creation of novel selectivity, the need to incorporate unnatural moieties into proteins has been recognized. Nature itself employs modified amino acids for regulatory, signalling and localization functions through the post-translational modifications of amino acid residues generating for example, 5-hydroxylysine, 4-hydroxyproline, O-methylaspartate, 3,5-diiodotyrosine, 4-carboxyglutamate, N,N,N-trimethyllysine, formylglycine as well as glycosylated, phosphorylated, acetylated, and sulfated amino acids. In fact, several approaches have been explored in attempts to overcome the natural amino acid limitation of conventional site-directed mutagenesis.

3.1 Chemical Modification to Generate Semisynthetic Enzymes

The utility of protein chemical modification techniques to incorporate unnatural functionalities and change enzyme properties has been recognized. The first reports of the application of chemical modification to enzymes by the groups of Bender and Koshland in 1966 predated site-directed mutagenesis techniques. Bender and Koshland independently created a thiolsubtilisin by
chemical transformation of the uniquely reactive active site serine of subtilisin BPN' to cysteine (Ser221-CH₂OH → Ser221-CH₂SH) yielding an active enzyme with altered catalytic properties. Subsequently, using an analogous procedure, subtilisin Carlsberg was converted into selenosubtilisin (Ser221CH₂OH → Ser221CH₂SeH). This selenosubtilisin exhibited improved potential as a peptide ligation catalyst and also exhibited peroxidase activity. Chemical modification of active site residues has been exploited extensively to alter the catalytic properties of hydrolases. For example, methylation of the catalytic triad histidine residue of subtilisin and of α-chymotrypsin was found to improve esterase to amidase selectivity. In addition, modification of a bacterial lipase with amino acid specific reagents including, 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, phenylglyoxal, pyridoxal 5'-phosphate, potassium iodide, and tetraniitromethane which modify all of the Asp/Glu, Arg, Lys, Trp, and Tyr residues respectively in the protein, was used to modulate lipase versus esterase activity.

In 1985, interest in chemically produced artificial enzymes, including some with synthetic potential, was renewed. In particular, a semisynthetic flavopapain, which exhibited oxido-reductase activity, was generated by the Kaiser group by covalently linking a flavin to the active site cysteine (Cys25) of the cysteine protease papain through a thioester linkage. This methodology was applied to several other protein templates to generate flavo-glyceraldehyde-3-phosphate, flavo-lysozyme via a flavin ester bond to the active site Asp52 or the surface exposed Asp101 residues, and flavo-hemoglobin via a flavin linkage to the two β-chain Cys93 residues. In a subsequent investigation, papain was converted into a ligase by alkylation of the active site cysteine with 2-bromomethyl-N-methyl/benzyl thiazolium bromide. This thiazolopapain, which had no detectable peptidase activity, catalyzed the dimerization of 6-oxoheptanal at a rate 400-fold better than the free thiazol and constitutes the first example of carbon-carbon bond formation catalyzed by a chemically modified enzyme. In a similar vein, the pyridoxamine cofactor was covalently linked to an interior cysteine residue of the noncatalytic adipocyte lipid binding protein via a disulfide linkage.
generated a protein able to catalyze the reductive amination of \( \alpha \)-keto acids to \( \alpha \)-amino acids with a 1.8 fold rate enhancement compared to the uncoupled pyridoxamine cofactor, and with an enantiomeric excess of 42 – 84%.

Chemical modification of enzymes has also been employed to alter surface properties. For example, succinylation of the 14 lysine residues of chymotrypsin (CT) decreased its positive surface charge and conversely, reaction of its 13 surface carboxylates with ethylenediamine increased its positive surface charge. These changes effected alterations in the pH-activity profile of CT. Similarly, the 59 carboxy groups of *Arthrobacter* D-xylose isomerase were chemically coupled to glycinamide in order to decrease the negative surface charge of the enzyme and lower its pH-optimum. Enzyme solubility and stability have also been improved by surface chemical modifications.

In addition there are countless examples of chemical modification as a technique to identify catalytically important residues. For example, the irreversible inhibitors, 5-azo-1-H-tetrazole and diisopropylphosphorylfluoridate were employed to identify the active site serine and histidine residues of the serine proteases. Iodoacetamide was used to identify the presence of sulphydryls essential for activity in rubredoxin and to determine the pH-activity dependence of papain. Chemical modification with thiol specific reagents has also been applied to probe receptor ligand binding such as for the glucocorticoid-receptor. While the chemical modification technique has proven to be very successful in altering biocatalyst properties, it suffers from the problems of nonspecific reactions. In addition, it is generally limited to either surface exposed residues or specially activated residues such as catalytic residues.
3.2 Unnatural Amino Acid Mutagenesis

Recently, the 20 amino-acid limitation of conventional site-directed mutagenesis has been recognized and addressed by biosynthetic methods.\textsuperscript{249-252} For this approach, the codon for the amino acid of interest is replaced with one of the stop codons (UAA, UAG, and UGA) by conventional oligonucleotide-directed mutagenesis. These codons are not recognized by any of the common tRNAs involved in protein biosynthesis and thus can be viewed as blanks. This methodology requires construction of a suppresser tRNA which recognizes the stop codon of choice, and its \textit{in vitro} aminoacylation with the unnatural amino acid to be introduced.\textsuperscript{253} A suppresser tRNA that recognizes this codon is then chemically acylated with the unnatural amino acid of interest. Addition of the mutated gene or mRNA and acylated tRNA to an \textit{in vitro} transcription-translation system results in the specific incorporation of the unnatural amino acid at the position of the stop codon. The \textit{in vitro} transcription-translation reactions are typically performed on a 30 \textmu l to 5 mL scale. Thus, a limitation of this approach is the relatively small quantity of protein which can be obtained.\textsuperscript{249} The scope and efficiency of amino acids incorporation has been investigated and, in general, large hydrophobic amino acids are inserted more efficiently than small or charged ones.\textsuperscript{254} In addition, D-amino acids are not accommodated by the translation machinery, although some \(\alpha,\alpha\)-disubstituted amino acids are.\textsuperscript{254} To address the limitations on the amino acid which can be incorporated, unnatural amino acids with side chains that can be elaborated by chemical modification have been incorporated and subsequently modified.\textsuperscript{255} Furthermore, thus far only a single unnatural amino acid has been incorporated into a protein at a time.

Despite the small amount of proteins expressed, this approach represents a significant advance in protein engineering since it permits the site-specific incorporation of unnatural amino acids, and it has already been applied to the incorporation of biophysical probes such as spin-labeled amino acids, fluorescent amino acids, photolabile benzophenone derivatized amino acids and photolabile
protected amino acids.\textsuperscript{256-257} The approach has also permitted detailed studies of the energetics of hydrogen bonding,\textsuperscript{258} protein stability and structure\textsuperscript{259-260} and has been applied to mechanistic investigations.\textsuperscript{261-264} In addition, the methodology has been expanded through the use of frame-shift suppression mutagenesis by utilizing synthetically prepared tRNA with a four base anticodon. This results in a truncation due to an out-of-frame reading if the four base codon is not correctly translated.\textsuperscript{265-266} While the unnatural amino acid mutagenesis approach has great potential, major improvements are required, including: increasing the quantities of protein obtainable, expanding the specificity of the translation factors to permit incorporation of more diverse amino acids, altering the selectivity of aminoacyl tRNA synthases to allow enzymatic charging of tRNA with unnatural amino acids, and developing the methodology to permit incorporation of more than one unnatural amino acid at a time.

### 3.3 Peptide Synthesis and Fragment Ligation

Unnatural amino acids have been introduced into peptides by incorporation of the new side chains directly onto the growing peptide chain during solid phase peptide synthesis,\textsuperscript{267} by simply coupling the individually prepared unnatural amino acid residues on solid support\textsuperscript{268-269} or in solution,\textsuperscript{270} or by enzyme catalyzed couplings.\textsuperscript{271} These chemically synthesized peptides, which are typically up to 40 residues long, can then be ligated on solid support,\textsuperscript{272-274} enzymatically,\textsuperscript{11,275-287} or chemically in solution\textsuperscript{288-289} to generate full length proteins. A related approach is protein semisynthesis in which a synthetic peptide which contains an unnatural amino acid is ligated to a natural protein fragment to produce a full-length protein.\textsuperscript{290} These approaches are feasible only for small proteins of less than 12 kDa which are able to refold or reassociate spontaneously to give functionally active protein, as for ribonuclease\textsuperscript{277} or cytochrome C.\textsuperscript{289} Furthermore, the viability of the incorporation of unnatural amino acids into self splicing protein motifs via synthetic inteins, has recently been demonstrated.\textsuperscript{291}
3.4 Site-Directed Mutagenesis Combined with Chemical Modification

In response to the deficiencies of the chemical modification, unnatural amino acid mutagenesis and solid phase peptide synthesis approaches for unnatural amino acid incorporation into proteins, the combined site-directed mutagenesis and chemical modification approach was recognized by Kaiser as a viable alternative. He recognized that it would be much more convenient if one did not have to rely on the natural availability of suitable residues for chemical modification but rather could introduce appropriate "handles" by genetic engineering which could subsequently be modified site-specifically. For this purpose, the most commonly introduced "handle" is the cysteine residue due to its unique reactivity.

This combined approach for the creation of new active-site environments and altered specificity was first applied to the hydrolytic enzyme carboxypeptidase Y. The S1' pocket M398C mutation was made by conventional site-directed mutagenesis and the new cysteine side chain was then chemically modified with phenylacetyl bromide and with 1-bromo-2-butanone. However, careful control of reaction conditions was required to prevent the reaction of these alkylating agents with other nucleophilic residues in the enzyme such as methionine and lysine. Carboxypeptidase Y M398C, was also modified with the thiol-specific alkyl methanethiosulfonate (CH₃SO₂S-R, MTS) reagents, where R = -CH₃, -CH₂CH₂CH₃, -CH₂C₆H₅, and -CH₂CH₂NH₃⁺, which reacted with the introduced cysteine to form a disulfide modified enzyme. Fortunately, residue C341 which was present both in the WT and M398C mutant enzyme was not modified due to its inaccessibility.

This combined site directed mutagenesis chemical modification approach has also aided mechanistic investigations. For example, two lysine residues in the active site of ribulosebisphosphate carboxylase were established as being critical to catalysis by their mutation to cysteine. This resulted in a complete loss of activity which was partially restored on aminoethylation of the cysteine residues with 2-bromoethylamine. However, the WT enzyme, which itself contains five
cysteines, was also alkylated by this procedure. In another example, the essential catalytic Lys258 residue of aspartate aminotransferase was mutated to Cys (K258C), thereby inactivating the enzyme. The mutant protein was then aminoethylated with 2-bromoethyl amine after reversible protection of the untargeted sulphydryl groups with Ellman’s reagent. The chemically elaborated enzyme, K258C-CH₂CH₂NH₃⁺, regained much of the activity lost after mutation. In another example, modification of the R292K mutant of aspartate aminotransferase with the guanidinating reagent O-methylisourea (MIU) was used to convert the WT active site residue Arg292 into homoarginine, however this was accompanied by modification of other reactive lysine side chains. The combined site-directed mutagenesis chemical modification approach was recently applied to glucoamylase effecting a 2-fold increase in activity. This combined approach has also yielded insights into the protein packing of thioredoxin and staphylococcal nuclease.

### 3.5 Current Study

The goal of the current study is to better understand the factors which control enzyme specificity, and to create novel specificities that will further expand the synthetic applicabilities of the serine proteases. To do this, we have adopted the strategy of chemical modification and site directed mutagenesis to alter the catalytic properties of the enzyme subtilisin *Bacillus lenta*us (SBL).

For this study, the sulphydryl group of cysteine was chosen as a modification handle since it is the most reactive of all amino acid side-chain functional groups under physiological conditions. In addition, the chemistry of sulphydryls is very robust and they are easily alkylated, acylated, arylated and oxidized. The cysteine sulphydryl, which has a pKₐ of 8.37, reacts via its thiolate anion with both reversible and irreversible sulphydryl blocking reagents. For example, N-ethylmaleimide, O-methylisourea, and iodoacetamide form irreversible adducts with cysteine thiols but 4,4'-dithiodipyridine, and alkyl methanethiosulfonate (MTS) reagents form readily reversible adducts with sulphydryls.
modification the use of a readily reversible modifying reagent is desirable since this permits restoration of the native activity. Furthermore, the chosen modification reagent must be devoid of the cross reactivity with other nucleophiles such as lysine and histidine which plagues many sulfhydryl blocking reagents. The alkyl methanethiosulfonate (MTS) reagents were chosen for the current study since they fulfill all of these requirements. While there is one report of the reaction of MTS reagents with lysine, this occurred under more vigorous conditions and this is consistent the 10-fold higher rate constant for the reaction of sulfhydryls versus amines with methanethiosulfonate reagents.

**Scheme 1.6**

The strategy involves the introduction of one cysteine residue at a key active site position via site-directed mutagenesis which is then thioalkylated with an alkyl methanethiosulfonate reagent (CH$_3$SO$_2$S-R) to give chemically modified mutant enzymes (CMMs) as illustrated in Scheme 1.6. Alkyl methanethiosulfonate reagents react specifically and quantitatively with sulfhydryls and are routinely used for chemical modification of protein thiols. The modification can be performed under mild reaction conditions, on a large scale, and is independent of the nature of the R group. Furthermore, the reaction is readily reversible by treatment with β-mercaptoethanol or other reducing agents.

The relatively low abundance of cysteine in proteins and in extracellular enzymes of gram-positive bacteria in particular makes this approach widely applicable. In addition to permitting the site-specific incorporation of unnatural amino acid side chains, the CMM approach avoids the difficulties which can result from nonconservative mutations that produce proteins which are unable to fold
properly or that do not undergo the required auto-processing. This potential difficulty is particularly relevant to the subtilisins, which are initially expressed as pre-pro-enzymes that must undergo autoproteolysis.\textsuperscript{67} SBL is especially well suited for this strategy since it contains no natural cysteine residues and therefore the introduced cysteine provides a unique sulfhydryl for modification. The particular utility of the combined site-directed mutagenesis chemical modification approach to subtilisins has been recognized. In particular, in order to generate an oxidation resistant subtilisin, the M222C mutant of a subtilisin from \textit{Bacillus lentus} was prepared, and then thioalkylated with the methyl methanethiosulfonate (CH\textsubscript{3}SO\textsubscript{2}SCH\textsubscript{3}) reagent, generating an enzyme which exhibited 65\% of WT activity.\textsuperscript{314} In addition to the specificity and mechanistic investigations, mutagenic introduction of cysteine combined with chemical modification has also been used for the site-directed incorporation of thiol spin-labels\textsuperscript{315-318} and fluorescence labels.\textsuperscript{319} This methodology has also facilitated detailed investigations of membrane spanning proteins,\textsuperscript{320-324} ion-channel properties,\textsuperscript{325-327} and ligand-receptor associations.\textsuperscript{328-329} As well, aminoethylation of cysteine residues has been long utilized as a method to introduce trypsin-susceptible cleavage sites into proteins to facilitate sequence studies.\textsuperscript{330}
References


37. Blow DM, Birkoff JJ, Hartley BS. Role of a Buried Acid Group in the

38. Wright CS, Alden RA, Kraut J. Structure of Subtilisin BPN' at 2.5 Angstrom


40. Walsh C. Enzymatic Reaction Mechanisms. San Francisco: Freeman and
Co., 1979: Chapter 3.

41. McPhalen CA, James MNG. Structural Comparison of Two Serine
Proteinase-Protein Inhibitor Complexes Eglin-C-Subtilisin Carlsberg and Cl-2-

42. Alberty RA, Massey V. On the Interpretation of the pH Variation of the
Maximum Initial Velocity of an Enzyme-Catalyzed Reaction. Biochim.

43. Bender ML, Clement GE, Kézdy FJ, Heck HD. The Correlation of the pH (pD)
Dependence and the Stepwise Mechanism of α-Chymotrypsin-Catalyzed

44. Fersht AR, Renard M. pH Dependence of Chymotrypsin Catalysis.


46. Zerner B, Bond RPM, Bender ML. Kinetic Evidence for the Formation of Acyl-
Enzyme Intermediates in the α-Chymotrypsin Catalyzed Hydrolyses of

47. Zerner B, Bender ML. The Kinetic Consequences of the Acyl-Enzyme
Mechanism for the Reaction of Specific Substrates with Chymotrypsin. J. Am.

Science 1957; 125:1145-6.


56. Knapp M, Daubermann J, Bott RR. Brookhaven Database Entry 1JEA.


109. Lalond JJ, Govardhan C, Khalaf N, Martinez AG, Visuri K, Margolin AL.


132. Takagi H, Ohtsu I, Nakamori S. Construction of Novel Subtilisin E with High Specificity, Activity and Productivity Through Multiple Amino Acid


155. Black ME, Loeb LA. Random Sequence Mutagenesis for Generation of Active


244. Welches WR, Baldwin TO. Active Center Studies on Bacterial Luciferase: Modification of the Enzyme with 2,4-Dinitrofluorobenzene. Biochemistry 1981; 20:512-7.


252. Bain JB, Diala ES, Glabe CG et al. Site-Specific Incorporation of Nonnatural


263. Parsons JF, Xiao G, Gilliland G, Armstrong RN. Enzymes Harboring


297. Planas A, Kirsch JF. Reengineering the Catalytic Lysine of Aspartate


329. Heinonen P, Koskua K, Pihlavisto M et al. A Series of 6-(Methanesulfonylethioalkoxy)-2-N-methyl-1,2,3,4-tetrahydroquinolines:

Chapter 2

Reproduced in part with permission from Biochemistry, 1998, 37, 5968-5973.
Copyright 1998, American Chemical Society (Reference 19)

-65-
Site-Directed Mutagenesis Combined with Chemical Modification as a Strategy for Altering the Specificity of the $S_1$ Pocket of Subtilisin *B. lentus*

**Introduction**

Employing the strategy of combined site-directed mutagenesis and chemical modification (Scheme 2.1), the goal of the current strategy was to alter the specificity of the primary-specificity determining pocket of SBL, $S_1$. Employing the crystal structure of SBL as our guide two residues at the bottom of the $S_1$ pocket, Ser156, one which is surface exposed, and Ser166, one whose side chain points into the pocket, were chosen for mutagenesis and chemical modification (Figure 2.1). The choice of methane thiosulfonate reagents is a representative one used to explore the viability of the CMM approach to alter enzyme specificity and activity. These include alkyl, (1a-1c), aromatic (1d), negatively charged (1e), and positively charged (1f) methanethiosulfonate reagents.

**Scheme 2.1**

![Scheme 2.1](image)
Figure 2.1 Active Site of SBL Showing Ser156 and Ser166. The catalytic triad residues are Asp32, His64, and Ser221. The S, pocket residues chosen for mutation and modification are Ser156 which is located at the bottom of the S, pocket and whose side chain points out toward solvent, and Ser166 which is located at the bottom of the S, pocket and whose side chain points in toward the active site.

Results

Each of the S156C and S166C SBL mutants was treated with the methanethiosulfonate reagents (1a-f) as shown in Scheme 2.1, and modification reactions were monitored by specific activity measurements. In all cases, reactions at the 156 and 166 sites with reagents 1a-f were complete within 30 min, consistent with the surface exposure of the 156 residue and the ready accessibility of the 166 residue.
Free thiol titration of the S156C and S166C CMMs with Ellman's reagent established that reactions were quantitative. In all cases, the free thiol content of the CMMs was less than 2%. Mass analysis of the CMMs by electrospray mass spectrometry was consistent (± 6 Da) with the calculated mass. The purity of the modified enzymes was assessed by native-PAGE and in all cases only one band was visible. CMMs S156C-S-a to -d, and S166C-S-a to -d, could not be distinguished from the parent cysteine mutant nor the WT enzyme on native-PAGE. However, the negatively charged CMMs derived from reactions with 1e, S156C-S-CH₂CH₂SO₃⁻, and S166C-S-CH₂CH₂SO₃⁻, displayed retarded mobility in the direction of the cathode, while the positively charged S156C-S-CH₂CH₂NH₃⁺, and S166C-S-CH₂CH₂NH₃⁺ CMMs derived from reactions with 1f displayed greater mobility relative to wild type.

That modification of cysteine is wholly responsible for altered activity was established by demonstrating that treatment of SBL-WT with each of 1a-f resulted in no change in activity, nor of molecular weight. Furthermore, treatment of S166C-S-a to-f, and S156C-S-a to-f, with β-mercaptoethanol restored activity to that of the parent cysteine mutant, verifying that chemical modification at cysteine was solely responsible for the observed changes in activity and is fully reversible by this treatment.

Kinetic constants for each of the CMMs, were determined with the suc-AAPF-pNA substrate and are shown in Table 2.1. Kᵣₛ's are virtually unaltered and consequently, kₒᵣ/Kᵣₛ changes are mainly reflective of kₒᵣ variations.
Table 2.1: Kinetic Parameters\textsuperscript{a} for S156C and S166C CMMs (pH 7.5)

<table>
<thead>
<tr>
<th></th>
<th>$k_{cat}$ (s\textsuperscript{-1})</th>
<th>$K_M$ (mM)</th>
<th>$k_{cat}/K_M$ (s\textsuperscript{-1} mM\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>48 $\pm$ 2</td>
<td>0.55 $\pm$ 0.06</td>
<td>87 $\pm$ 10</td>
</tr>
<tr>
<td>S156C</td>
<td>43 $\pm$ 2</td>
<td>0.65 $\pm$ 0.08</td>
<td>66 $\pm$ 9</td>
</tr>
<tr>
<td>S156C-S-a</td>
<td>30 $\pm$ 1</td>
<td>0.79 $\pm$ 0.08</td>
<td>38 $\pm$ 4</td>
</tr>
<tr>
<td>S156C-S-b</td>
<td>23.4 $\pm$ 0.5</td>
<td>0.68 $\pm$ 0.04</td>
<td>34 $\pm$ 2</td>
</tr>
<tr>
<td>S156C-S-c</td>
<td>24.2 $\pm$ 0.6</td>
<td>0.60 $\pm$ 0.04</td>
<td>40 $\pm$ 3</td>
</tr>
<tr>
<td>S156C-S-d</td>
<td>21.8 $\pm$ 0.7</td>
<td>0.54 $\pm$ 0.05</td>
<td>40 $\pm$ 4</td>
</tr>
<tr>
<td>S156C-S-e</td>
<td>31.6 $\pm$ 0.9</td>
<td>0.78 $\pm$ 0.06</td>
<td>40 $\pm$ 3</td>
</tr>
<tr>
<td>S156C-S-f</td>
<td>39 $\pm$ 1</td>
<td>0.86 $\pm$ 0.08</td>
<td>45 $\pm$ 4</td>
</tr>
<tr>
<td>S166C</td>
<td>14.2 $\pm$ 0.5</td>
<td>0.51 $\pm$ 0.05</td>
<td>28 $\pm$ 3</td>
</tr>
<tr>
<td>S166C-S-a</td>
<td>17.4 $\pm$ 0.7</td>
<td>0.39 $\pm$ 0.05</td>
<td>45 $\pm$ 6</td>
</tr>
<tr>
<td>S166C-S-b</td>
<td>11.8 $\pm$ 0.3</td>
<td>0.92 $\pm$ 0.05</td>
<td>12.8 $\pm$ 0.5</td>
</tr>
<tr>
<td>S166C-S-c</td>
<td>31 $\pm$ 1</td>
<td>0.72 $\pm$ 0.07</td>
<td>47 $\pm$ 4</td>
</tr>
<tr>
<td>S166C-S-d</td>
<td>6.9 $\pm$ 0.4</td>
<td>0.74 $\pm$ 0.09</td>
<td>9 $\pm$ 1</td>
</tr>
<tr>
<td>S166C-S-e</td>
<td>3.8 $\pm$ 0.1</td>
<td>0.70 $\pm$ 0.06</td>
<td>5.4 $\pm$ 0.5</td>
</tr>
<tr>
<td>S166C-S-f</td>
<td>16.3 $\pm$ 0.5</td>
<td>0.60 $\pm$ 0.06</td>
<td>27 $\pm$ 3</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Michaelis-Menten constants were measured at 25 °C according to the initial rates method in 0.1 M phosphate buffer containing 0.5 M NaCl at pH 7.5, and 1% DMSO, with succinyl-AAPF-pNA as the substrate.

Discussion

Once characterized, the specificity of each of the CMMs was evaluated with succinyl-AAPF-pNA as the standard reference substrate. Chemical modifications at each of the three sites resulted in very different activity patterns, as indicated in Table 2.1. Chemical modification of the S156C mutant effected only very small changes in $k_{cat}/K_M$. In contrast, chemical modification of the S166C mutant resulted in more dramatic changes in activity.
Figure 2.2 Specificity Patterns for S156C and S166C CMMs: Determined at 25 °C in 0.1 M phosphate buffer containing 0.5 M NaCl at pH 7.5, and 1 % DMSO, with succinyl-AAPF-pNA as the substrate.

Mutation of Ser156 to cysteine and modification with 1a-1f results in somewhat decreased $k_{cat}/K_M$. As illustrated in Figure 2.2 chemical modification of S156C resulted in small changes in activity relative to WT as indicated by the ratio of $\ln \left( \frac{k_{cat}/K_M}{k_{cat}/K_M_{WT}} \right)$. They are in accord with introduced groups being able to take up positions away from the active site as a result of the surface exposed nature of the S156C residue. Even when the enzyme's surface charge was altered, as for S156C-S-CH$_2$CH$_2$SO$_3^-$ (e) and S156C-S-CH$_2$CH$_2$NH$_3^+$ (f), activity remained
unaffected. This result contrasts previous observations for the related enzyme subtilisin BPN', for which increasing positive charge on the enzyme surface destabilizes the imidazolium form of the active site histidine and stabilizes the negatively charged oxyanion resulting in altered activity.\textsuperscript{4} The absence of such an electrostatic effect for SBL indicates a significant difference in specificity between the S\textsubscript{1} pockets of SBL compared to subtilisin Carlsberg or subtilisin BPN'.\textsuperscript{5} This may be due to the 4 amino acid deletion in SBL relative to subtilisin BPN' for which mutation of residue 156 significantly affects P\textsubscript{1} binding\textsuperscript{6} and enzyme activity.\textsuperscript{7} While these data were determined under conditions of high ionic strength, which may mask electrostatic effects\textsuperscript{8} $k_{cat}/K_M$ determinations at lower ionic strength did not significantly affect $k_{cat}/K_M$ either (Table 3.1, page 84). This suggests that the -S-CH$_2$CH$_2$NH$_3^+$ group of S156C-S-f and -S-CH$_2$CH$_2$SO$_3^-$ group of S156C-S-e are extensively solvated.

The most interesting and unexpected results are manifest for the S166C CMMs. An intriguing structure-activity relationship (SAR) of alternating increases and decreases in $k_{cat}/K_M$ upon modification of S166C with methanethiosulfonate reagents of increasing steric volume, specifically 1a (-S-CH$_3$), 1b (-S-CH$_2$CH$_3$) 1c (-S-CH$_2$CH(CH$_3$)$_2$) and 1d (-S-CH$_2$C$_6$H$_5$) is observed. The SAR illustrated in Figure 2.1b is unprecedented and quite different from what would have been predicted from previous site-directed mutagenesis studies of residue 166 in the related enzyme, subtilisin BPN'.\textsuperscript{9,10} The molecular basis for this novel activity pattern was analyzed by molecular modelling of the peptidyl product inhibitor AAPF bound to SBL-CMMs.
Figure 2.3 Molecular Modelling for S166C CMMs: Top panel shows the active site of WT-SBL (—) with the product inhibitor Ala-Ala-Pro-Phe (—) bound, highlighting the residues which comprise the S, pocket: 155-156,166 and 126-129, and the catalytic triad residues, 32, 64 and 221. Oxyanion stabilization is provided by the amide backbone hydrogens of Met222 and Asn155. Lower panels show the position of the S166C-S-R (—) side chain and the altered binding of the Phe residue for each of the S166C-S-a to -f CMMs (—) with respect to WT (——).
As shown in Figure 2.3, molecular modeling analysis of the S166C CMMs reveals that the observed $k_{cat}/K_m$ changes correlate with altered binding of the P₁ benzyl side-chain of the AAPF product inhibitor. While the phenyl ring of the P₁ Phe residue is positioned well into the S₁ pocket of WT-SBL, it is slightly distorted out of the pocket for S166C-S-CH₃ (-S-a). For S166C-S-CH₂CH₃ (-S-b) the phenyl binding is further distorted, but is less so for S166C-S-CH₂CH(CH₃)₂ (-S-c). Molecular modelling also revealed a very significant repositioning of the phenyl ring out of the S₁ pocket of S166C-S-CH₂C₆H₅ (-S-d), consistent with the lowered $k_{cat}/K_m$ for this CMM. Excitingly, in all cases the extent of binding distortion of the P₁ benzyl side chain in the S₁ pocket of the CMMs modeled correlates with the changes in $k_{cat}/K_m$. The molecular modelling results also suggest that for the S166C CMMs, disrupted phenyl binding may be due to changes in the preferred position of the β-carbon ($^{β}$-CH₂-R) of Cys166. As illustrated in Figure 2.3, for S166C-S-CH₂CH₃ (-S-b) the β-carbon of Cys166 points further into the S₁ pocket than it does for the S166C-SCH₃ (-S-a), S166C-S-CH₂CH(CH₃)₂ (-S-c), or WT enzymes, and thus pushes the P₁ Phe residue out of the S₁ pocket. However, for S166C-S-CH₂C₆H₅ (-S-d), which also displays lower activity, and for which molecular modelling indicates significantly distorted Phe binding, the position of the β-carbon of Cys166 is not altered. It is not immediately evident how the local re-orientation of the modified side chain of S166C-S-CH₂CH₃ (-S-b) may serve to relieve strain induced elsewhere in the S₁ pocket.

The positively charged $-S$-CH₂CH₂NH₃⁺ side chain of S166C-S-f effected no change in $k_{cat}/K_m$ relative to S166C, while the negatively charged $-S$-CH₂CH₂SO₃⁻ moiety of S166C-S-e caused a 5-fold decrease. A positive charge is expected to stabilize the unprotonated form of His64 as well as provide oxyanion stabilization, while a negative charge destabilizes these interactions.⁴ Thus factors such as non-optimal positioning of the positive charge within the enzyme active site, or masking of a favourable electrostatic interaction by an unfavourable steric one, are clearly operating to negate the beneficial influence of the positive charge of the $-S$-CH₂CH₂NH₃⁺ side chain of S166C-S-f. A dramatic structural change is seen for
S166C-S-f, in which the -S-CH₂CH₂NH₃⁺ side chain becomes directed into the S₁ pocket, and causes the phenyl ring of phenylalanine to be displaced into the upper region of the pocket. This unexpected orientation of the ethylammonium moiety into the S₁' pocket is attributed to a potential hydrogen bond between its ammonium hydrogen and the α-carbonyl of Gly 127, (N-O distance 3.9 Å in the energy minimized structure). While the conformation of the -S-CH₂CH₂NH₃⁺ side chain of S166C-S-f is unusual, it parallels the conformational flexibility of the side chain of Lys166 evident from the disordered electron density map observed in the X-ray structure of the G166K mutant of subtilisin BPN'. Furthermore, molecular modeling analysis shows that, while the -S-CH₂CH₂SO₃⁻ side chain of S166C-S-e does not orient itself directly into the S₁ pocket, it does cause significant disruption of inhibitor binding.

The fact that AAPF binding is correlated to the trends in $k_{cat}/K_m$ while $K_m$ is virtually unaltered suggests that the product inhibitor mimics transition state binding and not ground state binding of the succinyl-AAPF-pNA substrate. These modifications change enzyme turnover to a greater extent than substrate binding.

These results demonstrate that the combination of site-directed mutagenesis and site-specific chemical modification can alter the specificity of the S₁ and S₁', pockets of subtilisin B. lentus in unusual ways. Also, molecular modeling has again proven to be a powerful tool for analyzing the molecular basis for the activity changes induced by chemical modification and suggest that activity changes are correlated to altered binding of the P₁ moiety of AAPF in the S₁ pocket. It has also become apparent that the degree of activity changes which may be engendered upon modification is correlated with the extent of surface exposure of the residue modified.
**Experimental**

**Enzyme Purification.** Wild type subtilisin *Bacillus lentus*, and its S156C and S166C mutants were purified by the general method of Stabile *et al.* The crude protein concentrates containing PEG (50%) as a stabilizer, which were obtained as previously described, were purified on a Sephadex G-25 desalting matrix with a pH 5.2 buffer (20 mM sodium acetate, 5 mM CaCl₂), to remove small molecular weight contaminants. Pooled fractions from the desalting column were then applied to a strong cation exchange column (SP Sepharose FF) in the HEPES buffer (20 mM HEPES, 2 mM CaCl₂, pH 7.8), and SBL was eluted with a one-step gradient of 1-200 mM NaCl-HEPES buffer, (20 mM HEPES, 2 mM CaCl₂, 10 mM DTT, 200 mM NaCl, pH 7.8). Enzyme powder was obtained by dialysis (MWCO 12-14,000) of the eluent against 1 mM CaCl₂ and subsequent lyophylization. The purity of the WT and mutant enzymes, denatured by incubation with 0.1 M HCl at 0 °C for 30 min., was determined by SDS-PAGE on 20% homogeneous gels using the Phast System from Pharmacia. For all enzymes used, only one band was visible.

**Methanethiosulfonate reagents.** Reagent 1a was purchased from Aldrich Chemical Co. Inc., 1e and 1f from Toronto Research Chemical (2 Brisbane Rd. Toronto, ON), and all were used as received. Reagents 1b, 1c, and 1d were prepared as previously described.

**Site-Specific Chemical Modification.** To 25 mg of a SBL mutant in CHES buffer (2.5 mL; 70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5) at 20 °C was added one of the methanethiosulfonate reagents (1a-f) (100 μL of a 1 M solution: 1a in MeOH, 1b in EtOH, 1c in EtOH, 1d in CH₃CN, 1e in CHES buffer, 1f in MeOH), in a PEG (10,000) coated polypropylene test tube, and the mixture agitated in an end-over-end rotator. Blank reactions containing 100 μL of solvent instead of the reagent solution were run in parallel. Each of the modification reactions was monitored spectrophotometrically (ε₄₁₀ = 8800 M⁻¹ cm⁻¹) on a Perkin Elmer Lambda 2 spectrophotometer, by specific activity measurements. After the reaction was quenched by dilution in MES buffer (5 mM MES, 2 mM CaCl₂, pH 6.5) at 0 °C, the
specific activity of the CMM (10 µL), was determined in buffer containing: 0.1 M TRIS pH 8.6, 0.005 % Tween 80, and 1% DMSO, with the succinyl-AAPF-pNA substrate (1 mg/mL) (purchased from Bachem Bioscience Inc.) at 25 °C. The reaction was terminated when the addition of a further 100 µL of methane thiosulfonate solution effected no further change in specific activity, generally within 30 min. The reaction solution was purified on a disposable desalting column (Pharmacia Biotech PD-10, Sephadex G-25 M) pre-equilibrated with MES buffer. The CMM was eluted with MES-buffer (3.5 mL), dialyzed against 1 mM CaCl₂ (3 x 1 L) at 4 °C and subsequently lyophilized. Modified enzymes were analyzed by nondenaturing gradient (8-25%) gels at pH 4.2, run towards the cathode on the Pharmacia Phast-System, and appeared as one single band. Each of the CMMs was analyzed in parallel with its parent cysteine mutant and the WT enzyme.


**Regeneration of Unmodified Enzyme by Treatment with β-Mercaptoethanol.** To a solution of CMM (2.0 mg) in 250 µL of CHES-buffer (70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5) was added 10 µL of a solution of β-mercaptoethanol (1 M in 95% EtOH ). The reaction was monitored by specific activity measurements and in all cases the activity of the cysteine parent was restored.
**Free Thiol Titration:** The free thiol content of WT, S156C, S166C and their CMMs, was determined spectrophotometrically by titration with Ellman's reagent \( \epsilon_{\lambda=1} = 13600 \text{ M}^{-1} \text{ cm}^{-1} \) in phosphate buffer 0.25 M, pH 8.0 as was found to be < 2% in all cases.

**Active Site Titrations:** The enzyme concentration was determined\(^{16}\) by monitoring fluoride release upon enzyme reaction with \( \alpha \)-toluenesulfonyl fluoride (Aldrich Chemical Co. Inc.) as measured by a fluoride ion sensitive electrode (Orion Research 96-09). The enzyme concentration determined in this way was used to calculate kinetic parameters for each CMM.

**Kinetic Measurements:** Michaelis-Menten constants were measured at 25 °C by curve fitting (GraFit\(^{\circledR} \text{3.03} \)) of the initial rate data determined at eight concentrations (0.125 mM–4.0 mM) of the succinyl-AAPF-pNA substrate in 0.1 M phosphate buffer containing 0.5 M NaCl, 0.005% Tween 80, 1% DMSO, pH 7.5 \( (\epsilon_{410} = 8800 \text{ M}^{-1} \text{ cm}^{-1}) \).

**Molecular Modelling:** The X-ray structure of subtilisin *Bacillus lichen* with the peptide inhibitor AAPF bound\(^2\) was used as the starting point for calculations on wild type and CMMs. The enzyme setup was done with Insight II, version 2.3.0.\(^{17}\) To create initial coordinates for the minimization, hydrogens were added at the pH 7.5 used for kinetic measurements. This protonated all Lys and Arg residues and the N-terminus and deprotonated all Glu and Asp residues and the C-terminal carboxyl group. The protonated form of His 64 was used in all calculations. The model system was solvated with a 5 Å layer of water molecules. The total number of water molecules in the system was 1143. The overall charge of the enzyme-inhibitor complex resulting from this setup was +4 for the WT enzyme. Energy simulations were performed with the Discover program, Version 2.9.5\(^{18}\) on a Silicon Graphics Indigo computer, using the consistent valence force field function (CVFF). A non-bonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. The energy of the WT enzyme was minimized in stages, with initially only the water molecules being allowed to move, followed by water molecules and the amino acid side chains, and then finally the entire enzyme. The
mutated and chemically modified enzymes were generated by modifying the relevant amino acid using the Builder module of Insight. These structures were then energy minimized in a similar manner. Initially the side-chain of the mutated residue and the water molecules were energy minimized. Then the amino acid side chains within a 10 Å radius of the α-carbon of the mutated residue were energy minimized while all other residues were constrained, then all of the atoms within a 10 Å shell were energy minimized. The AAPF inhibitor was free to move throughout all stages of the minimization.
References


2. Knapp M, Daubermann J, Bott RR. Brookhaven Database Entry 1JEA.

3. All of the non-commercially available methanethiosulfonate reagents used for the CMM approach were prepared by Drs. Xiao Shang and Ben Davis.


Chapter 3

Reproduced in part with permission from Bioorg. Med. Chem. in press.
Unpublished work copyright 1999, Elsevier. (Reference 32)

-81-
Probing the Altered Specificity and Catalytic Properties of Mutant Subtilisin Chemically Modified at Position S156C and S166C in the S₁ Pocket

Introduction

The objective of the current study was to further explore the S₁ pocket specificity changes induced for the S156C-a to -d and S166C-a to -d CMMs (Scheme 3.1).¹ It is often desirable to utilize enzymes under conditions that differ from their pH optima particularly for synthetic applications.² Since, the strategy of chemical modification of site-directed mutant enzymes offers a new opportunity for controlling pH optima, pH-activity profiles for these CMMs were determined. In addition, an approach which is complementary to substrate kinetic analysis for probing enzyme specificity is the use of competitive reversible transition state analogue inhibitors. In this regard, boronic acid inhibitors have emerged as particularly effective probes for the serine proteases³⁹ and therefore a screen of a representative structural range of aromatic boronic acid inhibitors was applied to the S156C and S166C CMMs.

Scheme 3.1

\[
\text{R} = \begin{array}{llll}
a) \text{CH}_3 & b) \text{CH}_2 & c) \text{CH}_2\text{CH}_2\text{NH}_3^+ & d) \text{CH}_2\text{CH}_2\text{SO}_3^- \\
\end{array}
\]

\[1= H_3C-SO\text{SR} \xrightarrow{\text{pH 9.5}} SBL-SR \]

S156C S166C
S156C S166C
Results and Discussion

Each of the chemically modified mutant enzymes (CMMs) was prepared by the general method described previously,\textsuperscript{1} which entails reaction of the S166C and S156C mutants of SBL with each of the MTS reagents 1a-1d, yielding S156C-S-a to -d and S166C-S-a to -d as outlined in Scheme 3.1. Each of the CMMs was fully characterized, with electrospray mass spectrometry (ES-MS), the absence of residual free thiol was established by titration with Ellman's reagent,\textsuperscript{10} and by native-polyacrylamide gel electrophoresis (PAGE), all of which demonstrate that the modifications were quantitative and specific for the introduced cysteine.\textsuperscript{1}

The kinetic constants were evaluated at the pH 8.6 optimum of WT-SBL in Tris-HCl buffer with suc-AAPF-pNA as the standard substrate. These results are summarized in Table 3.1, which for comparative purposes also includes data determined at pH 7.5 (0.1 NaHPO\textsubscript{4}, 0.5 M NaCl) under high salt conditions (Table 2.1, page 69).\textsuperscript{1} In all cases the CMMs were more active under the pH 8.6 conditions. However, the trends in the $k_{cat}/K_m$ variations at both pH 8.6 and 7.5 parallel each other very closely.
Table 3.1 Kinetic Constants for S156C and S166C CMMs (pH 8.6)

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}$ s$^{-1}$ (a)</th>
<th>$K_m$ mM (a)</th>
<th>$k_{cat}/K_m$ s$^{-1}$ mM$^{-1}$ (a)</th>
<th>$pK_a$ (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>153 ± 4</td>
<td>0.73 ± 0.05</td>
<td>209 ± 15</td>
<td>7.01 ± 0.02</td>
</tr>
<tr>
<td>156C</td>
<td>125 ± 4</td>
<td>0.85 ± 0.06</td>
<td>147 ± 11</td>
<td>(66 ± 9)$^b$</td>
</tr>
<tr>
<td>156C-S-a</td>
<td>84 ± 3</td>
<td>0.84 ± 0.06</td>
<td>100 ± 8</td>
<td>(38 ± 4)$^b$</td>
</tr>
<tr>
<td>156C-S-b</td>
<td>72 ± 2</td>
<td>0.59 ± 0.05</td>
<td>122 ± 11</td>
<td>(40 ± 4)$^b$</td>
</tr>
<tr>
<td>156C-S-c</td>
<td>90 ± 2</td>
<td>0.73 ± 0.04</td>
<td>123 ± 7</td>
<td>(45 ± 4)$^b$</td>
</tr>
<tr>
<td>156C-S-d</td>
<td>87 ± 2</td>
<td>1.2 ± 0.07</td>
<td>74 ± 4</td>
<td>(45 ± 4)$^b$</td>
</tr>
<tr>
<td>166C</td>
<td>42 ± 1</td>
<td>0.5 ± 0.05</td>
<td>84 ± 7</td>
<td>(28 ± 3)$^b$</td>
</tr>
<tr>
<td>166C-S-a</td>
<td>46 ± 2</td>
<td>0.34 ± 0.05</td>
<td>135 ± 21</td>
<td>(45 ± 6)$^b$</td>
</tr>
<tr>
<td>166C-S-b</td>
<td>23.1 ± 0.5</td>
<td>1.17 ± 0.06</td>
<td>20 ± 1</td>
<td>(9 ± 1)$^b$</td>
</tr>
<tr>
<td>166C-S-c</td>
<td>50 ± 1</td>
<td>0.68 ± 0.04</td>
<td>74 ± 5</td>
<td>(27 ± 3)$^b$</td>
</tr>
<tr>
<td>166C-S-d</td>
<td>25.0 ± 0.7</td>
<td>1.34 ± 0.08</td>
<td>19 ± 1</td>
<td>(5.4 ± 0.5)$^b$</td>
</tr>
</tbody>
</table>

(a) Michaelis-Menten constants were measured by the initial rates method in pH 8.6 Tris-HCl buffer at 25 °C with suc-AAPF-pNA as the substrate. (b) Measured in pH 7.5 phosphate buffer (0.1 NaHPO$_4$, 0.5 M NaCl) at 25 °C with suc-AAPF-pNA as the substrate. Taken from reference 1. (c) The observed $pK_a$ was calculated from pH-activity profiles of $k_{cat}/K_m$", measured in 0.02 M ethylenediamine buffer, ionic strength 0.05 M adjusted with KCl 25 °C. n.d. = not determined.

While all CMMs exhibited lower than WT $k_{cat}/K_m$'s, all remain viable enzymes (Table 3.1). The modest activity changes caused by chemical modification of S156C are consistent with the surface exposed nature of this residue. Notably, a 2.9-fold decrease and a 1.7-fold decrease in $k_{cat}/K_m$ was effected for S156C-S-CH$_2$CH$_2$SO$_3^-$ (-d) and for S156C-S-CH$_2$CH$_2$NH$_3^+$ (-c) respectively, relative to WT. More dramatic changes in specificity were induced by modification of the 166 site. In particular, a 10.5-fold decrease in $k_{cat}/K_m$ compared to WT, resulting from a synergistic decrease in $k_{cat}$ and increase in $K_m$, was observed for the hydrophobic
side chain of S166C-S-CH₂C₆H₅ (-b) and is attributed to the unfavourable steric contacts revealed previously by molecular modelling analysis. However, for the negatively charged side chain CMM, S166C-S-CH₂CH₂SO₃⁻ (-d) the \( k_{cat}/K_m \) was equally (11-fold) reduced. In this case we conclude that the unfavourable steric interactions revealed by molecular modelling⁴ are augmented by the destabilizing effect of the repulsion between the negative charge of the sulfonato group with that of the incipient oxyanion of the transition state. Conversely, the introduction of a positive charge in the S, pocket, as for S166C-S-CH₂CH₂NH₃⁺ (-c), ameliorated the rate reducing effects, with only a 3-fold sterically induced decrease in \( k_{cat}/K_m \) relative to WT being manifest, and no change at all relative to the unmodified cysteine parent S166C.

It is often desirable to use enzymes under conditions that differ from their optima². Accordingly, the effect of chemical modifications on the pH-activity profiles of SBL-CMMs was examined. This approach to altering pH-activity profiles turned out to be quite general, in that pKₐ changes of up to 0.92 are also observed in the current study. A representative pH-activity profile is illustrated in Figure 3.1, and the overall pKₐ data are summarized in Table 3.1. In the absence of a change in rate determining step, and with no group in the substrate being ionizable within the pH range of the study, the pH dependence of \( (k_{cat}/K_m)_{obs} \) for the serine proteases follows the ionization of a catalytically important residue in the free enzyme.¹³⁻¹⁵ The ΔpKₐ values are observed to vary from -0.16 to +0.92 and are attributed to differences in the ionization of the catalytic His64 in the various CMMs.¹⁶⁻²⁰ For the uncharged modifications of S156C-S-CH₂C₆H₅ (-b) and S166C-SCH₂C₆H₅ (-b) the pKₐs of 7.07 and 7.00 respectively are virtually unchanged relative to WT (pKₐ = 7.01). In contrast, introduction of a negative charge, as for S156C-S-CH₂CH₂SO₃⁻ (-d) and S166C-SCH₂CH₂SO₃⁻ (-d), elicited a 0.48 and a 0.92 unit increase in the pKₐ of His64 respectively, while introduction of a positive charge as for S156C-S-CH₂CH₂NH₃⁺ (-e) and S166C-SCH₂CH₂NH₃⁺ (-e) caused a 0.09 unit and 0.16 unit decrease in the pKₐ respectively. The observed increases in pKₐ with increasing negative surface charge, and decreases with increasing positive surface charge, are attributed to
stabilization of the imidazolium form of His64 by a negative charge and its destabilization by a positive charge. These results are in accord with previous reports on charge effects\textsuperscript{16-20} and also demonstrate that electrostatic interactions are capable of more far reaching influences than hydrophobic or steric ones.\textsuperscript{15} The electrostatic effect of modification is more pronounced for the 166 than for the 156 site due to the different environments of the side chains. Since the 156 residue is more surface exposed and thus more readily solvated, the impact of the positively and negatively charged side chains on the active site are reduced. In contrast, the more buried nature of the S166C CMM side chains provides a protected environment whose influence on the enzyme active site residues is not diluted by solvation.

Figure 3.1 pH-Activity Profiles for WT-SBL (\(\bullet\), \(pK_a = 7.01 \pm 0.02\)) and S156C-S-\(\text{CH}_2\text{CH}_2\text{SO}_3^-\) (\(\circ\), \(pK_a = 7.49 \pm 0.06\)).
As a further probe of the S₁ pocket specificity changes induced by SBL-CMMs the altered binding of the representative structural range of transition state analog boronic acid inhibitors³⁹ (Formulae 2-6) was evaluated. The observed $K_s$ revealed significant specificity changes from WT, as summarized in Table 3.2.

**Formulae 2-6:**

![Image of chemical structures](image)

**Table 3.2 Inhibition Constants[^a] for S156C and S166C CMMs**

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Inhibitor 2 ($K_i$ mM)</th>
<th>Inhibitor 3 ($K_i$ mM)</th>
<th>Inhibitor 4 ($K_i$ mM)</th>
<th>Inhibitor 5 ($K_i$ mM)</th>
<th>Inhibitor 6 ($K_i$ mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>1.3 ± 0.1</td>
<td>1.3 ± 0.1</td>
<td>0.15 ± 0.01</td>
<td>8.4 ± 0.6</td>
<td>1.07 ± 0.08</td>
</tr>
<tr>
<td>156C</td>
<td>0.94 ± 0.07</td>
<td>1.09 ± 0.08</td>
<td>0.18 ± 0.01</td>
<td>7.1 ± 0.6</td>
<td>1.08 ± 0.08</td>
</tr>
<tr>
<td>156C-S-a</td>
<td>0.79 ± 0.06</td>
<td>1.4 ± 0.1</td>
<td>0.18 ± 0.01</td>
<td>5.4 ± 0.4</td>
<td>1.13 ± 0.08</td>
</tr>
<tr>
<td>156C-S-b</td>
<td>1.16 ± 0.09</td>
<td>1.19 ± 0.09</td>
<td>0.22 ± 0.02</td>
<td>2.4 ± 0.2</td>
<td>1.3 ± 0.1</td>
</tr>
<tr>
<td>156C-S-c</td>
<td>2.3 ± 0.1</td>
<td>1.8 ± 0.1</td>
<td>0.15 ± 0.01</td>
<td>8.1 ± 0.5</td>
<td>1.26 ± 0.07</td>
</tr>
<tr>
<td>156C-S-d</td>
<td>1.6 ± 0.1</td>
<td>2 ± 0.1</td>
<td>0.26 ± 0.02</td>
<td>6 ± 0.4</td>
<td>1.60 ± 0.1</td>
</tr>
<tr>
<td>166C</td>
<td>0.38 ± 0.04</td>
<td>0.42 ± 0.04</td>
<td>0.18 ± 0.2</td>
<td>9.0 ± 0.9</td>
<td>1.9 ± 0.2</td>
</tr>
<tr>
<td>166C-S-a</td>
<td>3.8 ± 0.6</td>
<td>0.42 ± 0.06</td>
<td>0.22 ± 0.03</td>
<td>4.6 ± 0.7</td>
<td>2.6 ± 0.5</td>
</tr>
<tr>
<td>166C-S-b</td>
<td>2.7 ± 0.2</td>
<td>4.3 ± 0.3</td>
<td>1.8 ± 0.1</td>
<td>65 ± 4</td>
<td>9.7 ± 0.5</td>
</tr>
<tr>
<td>166C-S-c</td>
<td>0.57 ± 0.04</td>
<td>0.78 ± 0.05</td>
<td>0.3 ± 0.02</td>
<td>4.52 ± 0.03</td>
<td>1.4 ± 0.1</td>
</tr>
<tr>
<td>166C-S-d</td>
<td>0.96 ± 0.06</td>
<td>2.3 ± 0.2</td>
<td>0.72 ± 0.05</td>
<td>18 ± 1</td>
<td>3.7 ± 0.2</td>
</tr>
</tbody>
</table>

[^a]: Inhibition constants were determined in pH 8.6 Tris-HCl buffer at 25 °C with suc-AAPF-pNA as the substrate by the method of Waley.⁲¹
The influence of the surface exposed 156 site is further reinforced since, as illustrated in Figure 3.2a, despite the widely varying structures of the boronic acids (2-6) evaluated only small changes in $K_s$ were observed. The greatest of these was a 3.5-fold improvement in the binding of 4-carboxyphenyl boronic acid (5) to S156C-S-CH$_2$C$_6$H$_5$ (-b). This indicated that some electrostatic interaction between the carboxylate group of 5 and the $S_1$ pocket of S156C-S-CH$_2$C$_6$H$_5$ was operating.

As illustrated in Figure 3.2b, more dramatic changes in boronic acid binding were observed for the S166C CMMs, most notably for S166C-S-CH$_2$C$_6$H$_5$ (-b). For this CMM, poorer binding of each of the boronic acids 2-6 was observed, with the largest effect being a 12-fold increase over WT in the $K_i$ of 2,4-dichlorophenyl boronic acid (4) with S166-S-CH$_2$C$_6$H$_5$ (-b). As expected, phenethyl boronic acid (6), the inhibitor with the largest $P_1$ group of the series, binds best to WT-SBL, which has the biggest $S_1$ pocket. That the $K_i$ of 4-carboxyphenyl boronic acid (5) with S166C-S-CH$_2$CH$_2$NH$_3^+$ (-c) is 2-fold lower than for WT is attributed to the improved interaction between the positively charged ammonium side chain of S166C-S-c and the negatively charged carboxylate moiety of 5. In further support of this interpretation, 4-carboxyphenyl boronic acid (5), with its negatively charged carboxylate group, binds more poorly to S166C-S-CH$_2$CH$_2$SO$_3^-$ (-c) than to WT, due to electrostatic repulsion.
Figure 3.2 Binding of Boronic Acid Inhibitors (2-6) to S156C and S166C CMMs: Stronger binding of the boronic acids to the WT is shown as a bar below the X-axis and stronger binding of the boronic acids to the CMM as a bar above the X-axis.
Figure 3.3 Molecular Modelling for $p$-Carboxyphenyl boronic acid:
(a) Energy minimized structure of the covalent complex of $para$-carboxyphenyl boronic acid (5) covalently linked to the Ser221 Oγ oxygen of WT-SBL showing H-bonds between:
- Asp32 carboxylate and His64 HN₂ (2.68 Å), His64 HN₃ and boronic acid hydroxyl oxygens (2.95 and 2.88 Å), Ser125 carbonyl oxygen and boronic acid hydroxyl oxygen (2.57 Å), Ser166 Oγ hydroxyl and carboxylate of 5 (2.57 Å).
- The H-bond between the Am155 backbone NH and carboxylate of 5 (2.98 Å) is a new one which was not present in the WT complex shown in (a).

(b) Energy minimized structure of the covalent complex of $para$-carboxyphenyl boronic acid (5) covalently linked to the Ser221 Oγ oxygen of S156C-S-b showing H-bonds between:
- Asp32 carboxylate and His64 HN₂ (2.70 Å), His64 HN₃ and boronic acid hydroxyl oxygen (2.86 Å), Ser125 carbonyl oxygen and boronic acid hydroxyl oxygen (2.58 Å), Ser166 Oγ hydroxyl and carboxylate of 5 (2.70 Å).
Support for the hypothesis of an electrostatic interaction between the carboxylate group of 5 and the S₁ pocket of S156C-S-CH₂C₆H₅ was sought by molecular modelling. Molecular modelling revealed an additional hydrogen bond between the carboxylate of *para*-carboxyphenyl boronic acid (5) and the Asn155 amide backbone N-H of S156C-S-b, which is not present in the WT (Figure 3.3). This additional hydrogen bond appears to have been gained at the expense of a His64 HNₐₙ hydrogen bond with a boronic acid hydroxyl oxygen which is present in the energy minimized structure of the WT.

These results demonstrate that using a screen of different specificity and catalytic property approaches in combination, such as kinetic, pH-activity profiles etc., provide complementary insights into the effect of site-directed chemical modifications of SBL, which taken together are more powerful than independently.
**Experimental**

**Preparation of Methanethiosulfonate Regents:** Reagent 1a was purchased from Aldrich Chemical Co. Inc., 1c and 1d from Toronto Research Chemicals (2 Brisbane Rd. Toronto, ON), and all were used as received. 1b was prepared via the nucleophilic displacement of benzyl bromide by sodium methanethiosulfonate.²²⁻²⁴

**Boronic Acids:** Boronic acids, 2 and 3 were purchased from Aldrich Chemical Co. Inc., 4 from Lancaster Synthesis Inc. (Windham, NH) and used as received. Boronic Acids 5-6 were obtained as previously described.³⁶

**Preparation and Characterization of CMMs:** The chemically modified mutants of SBL were prepared and purified as described previously.¹ Briefly each of S156C or S166C (25 mg) was exposed to each of the methanethiosulfonate reagents 1a-1d in CHES buffer (2.5 mL; 70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5) at 20 °C. These were purified by ion exchange chromatography (Pharmacia Biotech PD-10, Sephadex G-25 M). The free thiol content of each of the CMMs was determined by titration with Ellman's¹⁰ reagent and established that the free thiol content of all CMMs was less than 2 % demonstrating completeness of reaction. The purity of each of the CMMs was analyzed by nondenaturing gradient (8-25%) gels run at pH 4.2 run towards the cathode, on the Pharmacia Phast-System, and each appeared as a single band.¹ In addition, ES-MS spectral analysis of the CMMs revealed that their determined masses were found to be in agreement (± 6 Da) of the calculated.¹ The active enzyme concentration was determined as previously described²⁵ by monitoring fluoride release upon enzyme reaction with α-toluenesulfonfyl fluoride (Aldrich Chemical Co. Inc.) as measured by a fluoride ion sensitive electrode (Orion Research 96-09). The active enzyme concentration determined in this way was used to calculate kinetic parameters for each CMM.

**Determination of Kinetic Constants:** Michaelis-Menten constants were measured at 25 °C by curve fitting (GraFit® 3.03) of the initial rate data determined at eight concentrations (0.125 mM—4.0 mM) of the suc-AAPF-pNA (Bachem Inc., Torrance, California) in pH 8.6, 0.1 M Tris-HCl containing 0.005% Tween 80, and 1% DMSO, (ε₄₁₀ = 8800 M⁻¹ cm⁻¹)²⁶, [E] = 6.33 x 10⁻⁹ to 1.7 x 10⁻⁷ M, as described previously.¹
**pH-Activity Profiles:** pH-activity profiles for the subtilisin CMMs were constructed by monitoring product release via its colourometric absorbance at 410 versus time at 25 °C ε410 = 8800 cm⁻¹ M⁻¹)²⁸ as described previously.¹¹ $k_{cat}/K_M$ values were determined in duplicate in 0.02 M ethylenediamine buffer, ionic strength 0.05 M adjusted with KCl. [S] = 4.2 - 12.5 x 10⁻⁵ M of suc-AAPF-pNA, [E] = 2.53 to 21.7 x 10⁻³ M at 25 °C. pKₐ's were calculated using GraFit® version 3.03 curve fit, single pKₐ or bell-shaped double pKₐ, with the minimum set to zero. Specifically, into a 1.5 mL polystyrene cuvette was added 980 μL of buffer (0.02 M ethylenediamine) and 10 μL of substrate (4.2 -12.5 x 10⁻³ M in DMSO). The solution was incubated in a cell holder at 25 °C, before the absorbance reading was set to zero. Then 10 μL of enzyme solution (2.53 to 21.7 x 10⁻⁶ M in pH 5.8, 20 mM MES, 1 mM CaCl₂) was added to initiate the reaction. After an 8 sec delay, absorbance versus time measurements were recorded on a Perkin Elmer lambda 2 spectrophotometer. $k_{cat}/K_M$ values were calculated employing the low substrate approximation, where the Michaelis-Menten Equation reduces to: $v = k_{cat}/K_M [E]_o [S]$ when [S]<<$K_M$²⁷

**Determination of Boronic Acid Inhibition Constants:** Boronic acid $K_i$'s were determined in duplicate by the method of Waley.²¹ The progress curve without inhibitor was determined, from an assay mixture containing 980 μL of buffer (0.1 M Tris-HCl, 0.005% Tween 80) and 10 μL of suc-AAPF-pNA substrate (25 mM in DMSO). This mixture was incubated in a water jacketed cell for 5 min. at 25 °C. The absorbance reading was set to zero prior to initiating the reaction by addition of 10 μL of enzyme solution (3.2 x 10⁻⁶ to 1.1 x 10⁻⁴ M in pH 5.8, 20 mM MES, 1 mM CaCl₂). The final volume of the assay mixture was 1 mL. The progress curve with inhibitor was determined, similarly but with 10 to 980 μL of the boronic acid inhibitor solution (1 x 10⁻³ to 0.1 M, in 0.1 M Tris-HCl, 0.005% Tween 80) added to a final volume of 1 mL. Absorbance versus time measurements were recorded on a Perkin Elmer lambda 2 spectrophotometer and points for calculation were taken at 15, 18, 21, 24, 27, 30, 33 and 36% substrate conversion.²⁸

**Molecular Modelling Analysis:** Energy simulations were performed with the Discover program version 2.9.5²⁹ on a Silicon Graphics Indigo computer, using the
consistent valence force field function. The X-ray structure of subtilisin \textit{Bacillus lentus} with the peptide inhibitor AAPF bound was used as the starting point for calculations.\textsuperscript{30} To create initial coordinates for the minimization, hydrogens were added at the pH 8.6 used for kinetic measurements. This protonated all Lys and Arg residues and the N-terminus and deprotonated all Glu and Asp residues and the C-terminal carboxyl group. The protonated form of His 64 was used in all calculations. The model system was solvated with a 5 Å layer of water molecules. The total number of water molecules in the system was 1143. The overall charge of the enzyme-inhibitor complex resulting from this setup was +4 for the WT enzyme. A non-bonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. The energy of the WT enzyme was minimized in stages, with initially only the water molecules being allowed to move, followed by water molecules and the amino acid side chains, and then finally the entire enzyme, until the maximum derivative of 0.1 kcal mol\(^{-1}\) Å\(^{-1}\) was reached. The S156C-S-b enzyme was generated by modifying the relevant amino acid using the Builder module of Insight.\textsuperscript{31} The energy of this structure was then minimized in a similar manner. Initially the side-chain of the mutated residue and the energy of the water molecules was minimized, followed by water molecules and the amino acid side chains, and then finally the entire enzyme. The AAPF inhibitor was free to move throughout all stages of the minimization. The AAPF inhibitor was deleted and then the energy minimized enzyme structure was used as the starting points for calculation of the boronic acid-enzyme complex. In the calculation of the boronic acid-enzyme complex, a tetrahedral carbon atom was used to mimic the boron atom since, no force field parameters are available for boron. The boron atom equivalent of the inhibitor was covalently bound to the O\(\gamma\) of Ser221. The inhibitor was docked such that its phenyl moiety was directed into the S\(_1\) pocket, one of the hydroxyl groups was directed to Asn155 of the oxyanion hole and the other toward His64. The energy of this structure was then minimized as above.
References


30. Knapp M, Daubermann J, Bott RR. Brookhaven Database Entry 1JEA.


Chapter 4
Benzophenone Boronic Acid Photoaffinity Labeling of SBL CMMs

Introduction

Previously, (Table 3.2, page 87) the $S_1$ binding pocket properties of the $S_{166C}$ and $S_{156C}$ chemically modified mutant enzymes (CMMs) of SBL were probed using a series of boronic acid transition state analogue competitive inhibitors.\cite{1} The inhibitor probing approach revealed substantial changes in $S_1$ pocket binding and also yielded insights into the molecular basis for these changes.\cite{1} In order to further investigate the altered binding modes of boronic acid inhibitors to the $S_1$ pocket, and since X-ray crystal structures of the boronic acid inhibitors bound to the CMMs are not available, we considered the strategy of employing an active site directed photolabile $p$-boronic acid benzophenone (BBP, 1) inhibitor to map SBL’s binding site.

Active site directed photoaffinity labels\cite{2,3} have been employed to probe enzyme active sites,\cite{7,8} and provide structural information which is complementary to that which may be available from X-ray crystallography\cite{9} or NMR\cite{10,11} approaches. The photolysis of benzophenone derivatives at $\lambda$ 320-360 nm effects an $n$ to $\pi^*$ transition which induces formation of a covalent link to amino acid residues, or other functionalities, by H-abstraction and radical recombination,\cite{3,12} as shown in Scheme 4.3. Benzophenone (BP) inhibitor\cite{7,8} and substrate\cite{13} derivatives have been employed extensively in active-site directed photoaffinity labeling studies.\cite{3,12} In addition, the benzophenone moiety has also been introduced into peptides or proteins.\cite{3,12} The advantages of the BP photophore have been documented and include its chemical stability, its stability to ambient light and its 330-360 nm photoactivation range, which is not damaging to proteins.\cite{4} In addition, BP is particularly well suited to the current probing study since SBL’s active site is quite hydrophobic and BP exhibits preferential labeling of hydrophobic binding regions in proteins.\cite{4}
In the current study, the photolabile \( p \)-boronic acid benzophenone (BBP, 1) was designed to incorporate both the boronic acid transition state analogue inhibitor structure\(^{14-20} \) and the photoactivatable benzophenone functionality\(^3-12 \). The \( p \)-benzophenone boronic acid (BBP) photoprobe, has the active site-directing phenyl boronic acid inhibitor incorporated into the structure of BP rather than tethered to the end of the photoprobe, which should in principle reduce nonspecific labeling\(^4 \).

![BBP, 1](image)

Since more substantial changes were effected by modifications at the S166C site compared to modifications at the S156C site as revealed by substrate (Table 3.1, page 84)\(^1,21 \) and inhibitor (Table 3.2, page 87)\(^1 \) kinetics, only the S166C CMMs were subjected to the current \( p \)-boronic acid benzophenone probing studies. In addition to WT-SBL as a control, four representative S166C CMMs were investigated. Two of these, S166C-S-CH\(_2\)C\(_6\)H\(_5\) (-a) and S166C-S-CH\(_2\)-c-C\(_6\)H\(_{11}\) (-b) have a sterically constricted S\(_1\) pocket while S166C-S-CH\(_2\)CH\(_2\)SO\(_3^\-\) (-c) and S166C-S-CH\(_2\)CH\(_2\)NH\(_3^\+\) (-d) have a negatively and a positively charged S\(_1\) group respectively (Scheme 4.1). In this study we report the synthesis of the BBP reagent, its evaluation as a competitive reversible inhibitor of WT-SBL and of the four representative S166C CMMs (Scheme 4.1) and photolysis of the EI complexes.
Results and Discussion

Each of the chemically modified mutant enzymes (CMMs) were prepared \(^{21}\) and characterized \(^{21-22}\) by the general method described previously, yielding S166C-S-a to -d as outlined in Scheme 4.1.

The \(\rho\)-boronic acid benzophenone (BBP, 1) photoprobe designed for the current study was prepared by standard procedures as outlined in Scheme 4.2. Briefly, 4-bromobenzophenone (3) was protected as the glycal in refluxing benzene. \(^{23,24}\) The glycal (4) was then subjected to halogen-lithium exchange by reaction with \(n\)-BuLi, followed by treatment with trimethyl borate to generate the corresponding dimethyl borate. \(^{17,24}\) In order to facilitate purification, \(^{17}\) this dimethyl borate was subsequently deprotected in sulfuric acid and reprotected with ethylene glycol, generating the diglycal (5) which was recrystallized from benzene. The purified diglycal (5) was then fully deprotected in refluxing 6 M HCl, \(^{24}\) yielding the \(\rho\)-boronic acid benzophenone (BBP, 1) in 77 % overall yield (Scheme 4.2).
Initially, the efficacy of binding of the designed photolabile competitive inhibitor, BBP, to each of the WT, S166C-S-CH₂C₆H₅, S166C-S-CH₂-c-C₆H₁₁, S166C-S-CH₂CH₂NH₃⁻ and S166C-S-CH₂CH₂SO₃⁻ enzymes was determined using the standard suc-AAPF-pNA substrate. The kinetic data are summarized in Table 4.1, and as expected, in all cases BBP binds best to WT. BBP binding to S166C-SCH₂C₆H₅ (-a) and S166C-CH₂-c-C₆H₁₁ (-b), with their large hydrophobic side chains, is reduced by 86- and 9-fold, respectively, compared to WT. Previously, all five of the boronic acids evaluated (Figure 3.2b, page 89) exhibited poorer binding with the S166C-S-CH₂C₆H₅ CMM compared to WT, and therefore the increased $K_i$ of BBP with this CMM is not unexpected. That BBP binding to S166C-S-CH₂-c-C₆H₁₁ is better than to S166C-SCH₂C₆H₅ may be due to the greater hydrophobicity of the side chain or to differences in side chain orientation. BBP binding to the charged CMMs, S166C-S-CH₂CH₂SO₃⁻ (-c) and S166C-S-CH₂CH₂NH₃⁺ (-d), is reduced 170- and 4-fold respectively compared to
to WT. This inhibition pattern is comparable to that revealed by the previous boronic acid inhibitor screens, for which, all expect for the smallest inhibitor, phenyl boronic acid, exhibited poorer binding to S166C-CH$_2$CH$_2$SO$_3^-$ compared to WT (Figure 3.2b, page 89). In contrast, the apparently more accommodating nature of the S$_1$ pocket of S166C-CH$_2$CH$_2$NH$_3^+$ is in accord with the fact that three of the five boronic acid inhibitors evaluated exhibited improved binding compared to WT (Figure 3.2b, page 89). Thus, BBP binding to SBL exhibits the same activity patterns as did the previously evaluated boronic acid transition state analogue inhibitors.

Table 4.1 Kinetic and Inhibition Constants for S166C CMMs with BBP

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_M$ mM</th>
<th>$k_{cat}$ s$^{-1}$</th>
<th>$k_{cat}/K_M$ mM$^{-1}$s$^{-1}$</th>
<th>$K_i$ mM$^{(a)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.55 ± 0.06</td>
<td>48 ± 2</td>
<td>87 ± 10</td>
<td>0.037 ± 0.004</td>
</tr>
<tr>
<td>S166C-S-a</td>
<td>0.74 ± 0.09</td>
<td>6.9 ± 0.4</td>
<td>9 ± 1</td>
<td>3.2 ± 0.04</td>
</tr>
<tr>
<td>S166C-S-b</td>
<td>0.61 ± 0.04</td>
<td>8.7 ± 0.2</td>
<td>14.3 ± 0.9</td>
<td>0.32 ± 0.02</td>
</tr>
<tr>
<td>S166C-S-c</td>
<td>0.70 ± 0.06</td>
<td>3.8 ± 0.1</td>
<td>5.4 ± 0.5</td>
<td>6.3 ± 0.6</td>
</tr>
<tr>
<td>S166C-S-d</td>
<td>0.60 ± 0.06</td>
<td>16.3 ± 0.5</td>
<td>27 ± 3</td>
<td>0.15 ± 0.02</td>
</tr>
</tbody>
</table>

(a) Michaelis-Menten constants were measured by the initial rates method in pH 7.5 phosphate buffer (0.1 NaHPO$_4$, 0.5 M NaCl) at 25 °C with suc-AAPF-pNA as the substrate. (b) Inhibition constants were determined in pH 7.5 phosphate buffer at 25 °C with suc-AAPF-pNA as the substrate by the method of Waley.

The initial photolysis time-course experiments were done with WT-SBL. The photolysis reactions were conducted in pH 5.8 storage buffer (10 MES, 1 mM CaCl$_2$) in order to reduce autolysis. The photolysis reactions were conducted at 4 °C in a Rayonet photoreactor equipped with 350 nm lamps. The BBP photoprobe was used at both a 2-fold and a 10-fold molar excess concentration with respect to enzyme. In addition, WT was photolyzed in the absence of BBP and as a dark control, a sample of WT + BBP in a foil wrapped cuvette was also
photolyzed alongside the other reaction vessels. Aliquots were withdrawn at intervals and specific activity measurements with the substrate were determined.

![Figure 4.1 Time Course for Photolysis (350 nm) of SBL-WT and BBP.](image)

The photolysis time course experiment revealed a time dependent decrease in specific activity falling off to zero after 60 min, as shown in Figure 4.1. Both the 2- and 10-fold molar excess BBP reactions resulted in the same final activity with the more concentrated reaction resulting in a slightly faster activity loss. For both the WT photolyzed control (WT + hv) and the dark control (WT + BBP) reactions, activity was unaltered. This established that the activity losses observed were due exclusively to reaction between the enzyme and BBP. After photolysis, each sample was purified and analyzed by ES-MS. Both the WT photolyzed (WT + hv) control and the dark (WT + BBP + dark) control retained the original WT mass, suggesting that no chemical changes were induced by photolysis of the enzyme in the absence of BBP nor by treatment of the enzyme with BBP in the dark. However, for the WT enzyme photolyzed in the presence of either a 2-fold or a 10-fold molar excess of BBP a mass increase consistent with
covalent linkage of BBP and the loss of two moles of water was observed as summarized in Table 4.2.

**Table 4.2 ES-MS Data for BBP Treated WT-SBL**

<table>
<thead>
<tr>
<th>Enzyme Preparation</th>
<th>Mass Calcd.</th>
<th>Mass Found</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT + hν</td>
<td>26698</td>
<td>26702</td>
<td>native enzyme</td>
</tr>
<tr>
<td>WT + 2X BBP + hν</td>
<td>26888</td>
<td>26888</td>
<td>+ BBP -2 H₂O</td>
</tr>
<tr>
<td>WT + 10X BBP + hν</td>
<td>26888</td>
<td>26891</td>
<td>+ BBP -2 H₂O</td>
</tr>
<tr>
<td>WT + 2X BBP + dark</td>
<td>26698</td>
<td>26702</td>
<td>native enzyme</td>
</tr>
<tr>
<td>WT + 10X BBP + dark</td>
<td>26698</td>
<td>26706</td>
<td>native enzyme</td>
</tr>
</tbody>
</table>

Each of the enzyme samples was denatured by treatment with urea and then degraded by trypsin into 13 fragments designated 1T to 13T, which correspond to the tryptic fragments numbered sequentially from the amino terminus of the enzyme. These fragments were then separated by reverse-phase HPLC which, coupled to mass spectrophotometric analysis, permits identification of each of the tryptic fragments on a peptide map. Comparison of the peptide map of the BBP treated and photolyzed enzyme to that of the control was used to identify the BBP modified peptide fragment. It revealed a modification on tryptic fragment 6T, which is comprised of amino acids 93 to 143 in the linear sequence. Modification resulted in the disappearance of fragment 6T and yielded two new peaks. These were a minor peak designated as 6T' and a major peak designated as 6T". Both of these are more hydrophobic than 6T itself, as evidenced by their longer retention times by reverse-phase HPLC, and showed mass increases over 6T of 197 and 189 Da for 6T' and 6T" respectively (Figure 4.2).
The peptides $6T'$ and $6T''$ were thought to correspond to the original tryptic fragment $6T$, in the 269 amino acid linear sequence of SBL, covalently modified by BBP crosslinking. The $6T'$ and $6T''$ fragments were isolated by HPLC. The specific location of the covalent bond between BBP and the $6T'$ and $6T''$ tryptic fragments of the enzyme were identified by amino acid sequencing of the isolated peptides.\textsuperscript{27-29} For the minor peptide $6T'$ the Edman sequencing signal became weak after cycle 33 revealing the sequence
\[\text{VLGASGSGSVSSIAQGLEWAGNGMHVANLS}^{126}\text{L}^{127}\text{G}^{128}\text{xxx (BPN' numbering).}\]
however, it was not clear if the cycle stopped at this point, or just weakened.\textsuperscript{30} In contrast, for the major peptide $6T''$ an abrupt termination after cycle 32 was noted, revealing the sequence $\text{...VANLS}^{126}\text{L}^{127}\text{xxx (BPN' numbering).}$ The
sequence information obtained for the major peptide 6T" identifies a modification at residue Gly127 (BPN' numbering). Gly127 is located on the back wall of the S, pocket. A proposed crosslinking mechanism which accounts for this modification and for the loss of two water molecules, and which is consistent with the observed mass increase is illustrated in Figure 4.3. The modification site of the minor peptide fragment 6T' is not conclusive but may be at residue Ser128.

![Diagram](image)

**Figure 4.3** Proposed Mechanism of BBP Crosslinking: The Photoactivated Crosslinking of BBP to Gly127 and Dehydration. The last dehydration step forming a boronate ester between BBP and the enzyme may precede crosslinking.

The above results demonstrate the validity of the approach for probing the active site of WT-SBL. Accordingly, the same protocol was applied to the CMMs S166C-S-a to -d, using a 10-fold molar excess of BBP, in order to probe BBP binding changes. However, despite the 12 hour photolysis time, specific activity was not decreased to zero for any of the CMM samples (Figure 4.4).
Figure 4.4 Time Course for Photolysis (350 nm) of SBL CMMs and BBP
Nonetheless each of the photolyzed CMMs and the dark controls were purified and analyzed by ES-MS. For each of S166C-S-a, -c, -d, no modification is indicated by ES-MS (Table 4.3). A mass increase of 220 Da is observed for S166C-S-CH$_2$-c-C$_6$H$_{11}$ (-b) 10X BBP + hv, which is consistent (± 6 Da) with the formation of a covalent bond between S166C-S-b and BBP (calculated mass of BBP 226 Da).

For the S166C-S-c-C$_6$H$_{11}$ CMM, both the dark control (S166C-S-b 10X BBP + dark) and photolyzed sample (S166C-S-b 10X BBP + hv), were subjected to tryptic digestion and peptide mapping as described above for the WT. For the photolyzed S166C-S-b (10X BBP + hv) sample a reduction in intensity of the tryptic fragment 6T (50%) was observed which was accompanied by two new peaks, a minor peak designated as 6T' and a major peak designated as 6T". Both 6T' and 6T" are more hydrophobic than 6T itself and showed mass increases over 6T of 197 and 188 Da respectively. The tryptic fragments 6T' and 6T" were isolated and subjected to amino acid sequencing. The sequences for 6T' and 6T" were identical. The observed ...VANLS$_{126}^{127}$xxx (BPN' numbering) sequence is identical to that derived from the 6T" tryptic fragment of the WT-SBL + BBP photolyzed sample. Consistent with the observed mass increase of 188 Da (BBP-2H$_2$O:190 Da Calcd.) for the tryptic fragment, this implicates the formation of a crosslink between the S$_1$ pocket Gly127 residue of S166C-S-b and BBP with concomitant loss of two water molecules. The identity of the minor tryptic fragment 6T' is not apparent. Both the WT and S166C-S-CH$_2$-c-C$_6$H$_{11}$ enzymes show a crosslink to BBP at residue Gly127 despite having very different active site environments. This suggests that BBP forms a boronate ester adduct with the enzyme, which although reversible is sufficiently long-lived that it acts as a tether restricting the spatial span of the BBP reagent in the enzyme active site. Such adducts between boronic acid inhibitors and the Ser221 O$_7$ hydroxyl of serine proteases have been observed previously by X-ray crystallography and by NMR.
Table 4.3 ES-MS Data for BBP Treated SBL CMMs

<table>
<thead>
<tr>
<th>Enzyme Preparation</th>
<th>Mass Calc.</th>
<th>Mass Found</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>S166C-S-a</td>
<td>10X BBP + dark</td>
<td>26836</td>
<td>26847 native enzyme</td>
</tr>
<tr>
<td>S166C-S-a</td>
<td>10X BBP + hv</td>
<td>27062</td>
<td>26839 native enzyme</td>
</tr>
<tr>
<td>S166C-S-b</td>
<td>10X BBP + dark</td>
<td>26842</td>
<td>26844 native enzyme</td>
</tr>
<tr>
<td>S166C-S-b</td>
<td>10X BBP + hv</td>
<td>27068</td>
<td>27062&lt;sup&gt;(a)&lt;/sup&gt; + BBP</td>
</tr>
<tr>
<td>S166C-S-c</td>
<td>10X BBP + dark</td>
<td>26853</td>
<td>26868 native enzyme</td>
</tr>
<tr>
<td>S166C-S-c</td>
<td>10X BBP + hv</td>
<td>27079</td>
<td>26867 native enzyme</td>
</tr>
<tr>
<td>S166C-S-d</td>
<td>10X BBP + dark</td>
<td>26790</td>
<td>26818 native enzyme</td>
</tr>
<tr>
<td>S166C-S-d</td>
<td>10X BBP + hv</td>
<td>27016</td>
<td>26818 native enzyme</td>
</tr>
</tbody>
</table>

<sup>(a)</sup> Poorly resolved spectrum.

Since no crosslinks were observed for S166C-S-a, -c, -d after photolysis for 12 h in the presence of 10-fold molar excess of BBP, each of these CMMs were photolyzed for up to 72 hours in the presence of an 80-fold molar excess of BBP, the solubility limit. However, although activity decreases were effected under these conditions, purification of the enzymes followed by their ES-MS analysis did not provide evidence of crosslinking. Therefore, S166C-S-a, -c and -d CMMs preclude proper orientation or alignment of BBP in the S₁ pocket for crosslinking with the enzyme. The reduced binding of BBP to S166C-S-a to -d, as manifested by higher $K_s$, and the lack of an observed crosslink between BBP and S166C-S-a, -c and -d indicate that modification of residue C166 makes the S₁ pocket less conducive to binding large P₁ groups.
**Experimental**

**Preparation of p-Boronic acid benzophenone (1):** A stirred solution of 4-bromobenzophenone (3.12 g, 12.0 mmol), ethylene glycol (4.7 g, 75 mmol), and p-toluenesulfonic acid (0.2 g, 1.2 mmol) in benzene (30 mL) was heated under reflux for 48 h and the resulting water was removed by azeotropic distillation in a Dean-Stark trap. The reaction was allowed to cool to 22 °C and then aqueous 1 M NaOH (40 mL) was added and the mixture was extracted with Et₂O (3 × 30 mL). The organic phases were combined, dried (MgSO₄), filtered, and then evaporated *in vacuo* yielding 4-bromobenzophenone ethylene glycol (2) as a white solid which was recrystallized from EtOH (3.7 g, 12 mmol, quantitative yield).

\[ {^1}H \text{ NMR (200 MHz, CDCl}_3) \delta 4.03 (4H, s, CH₂), \delta 7.27-7.50 (9H, m, Ar-H) \] ppm.

To a stirred solution of 4-bromo benzophenone ethylene glycol (3), (1.0 g, 3.28 mmol) dissolved in dry THF (10 mL) under N₂ at -78 °C was added dropwise via syringe n-butyl lithium (1.7 mL of a 2.5 M solution in hexane, 4.3 mmol). This mixture was stirred for 1 h then transferred by cannula to a solution of trimethylborate (0.51 g, 4.9 mmol) in dry THF (10 mL) at -78 °C. This mixture was stirred at -78 °C for 1 h and then allowed to warm to room temperature overnight. The reaction mixture was slowly added to aqueous 10% H₂SO₄ (20 mL) at 0 °C and stirred for 30 min. The reaction mixture was then extracted with Et₂O (3 × 30 mL). The combined organic phases were dried (MgSO₄), filtered and then evaporated *in vacuo* to yield the boronic acid as a yellow oil which was added directly to a flask containing a solution of ethylene glycol (0.2 g, 3.3 mmol) in diethyl ether (25 mL) and stirred vigorously for 1 h at 20 °C. To this reaction mixture were added hexanes (50 mL) and water (1 mL) and the aqueous layer was separated. The organic layer was dried (MgSO₄), filtered, and then evaporated *in vacuo* yielding the diglycal (4) as a white solid which was recrystallized from benzene (0.87 g, 2.9 mmol, 88% yield).
(200 MHz, CDCl₃) δ 4.06 (4H, s, CH₂), 4.36 (4H, s, CH₂), 7.27-7.50 (9H, m Ar-H) ppm.

The diglycal 4. (270 mg, 0.91 mmol) was then dissolved in a stirred solution of 6 M HCl (15 mL) and acetone (5 mL) and heated to 60 °C for 2 h with stirring and then stirred for a further 48 h at room temperature. The reaction mixture was extracted into diethyl ether (2 x 20 mL). The combined organic phases were then washed with brine (20 mL), poured into H₂O (5 mL) and concentrated in vacuo, with the benzophenone boronic acid precipitating as white crystals (180 mg, 0.8 mmol, 87% yield). ¹H NMR (200 MHz, DMSO-d₆) δ 8.35 (2H, s, OH), δ 7.58-7.98 (9H, m, Ar-H) ppm. ¹³C NMR (400 MHz, DMSO-d₆) δ 196.2, 138.2, 137.1, 134.1, 132.8, 129.7, 128.6, 128.5 ppm. IR (KBr) 3500-3200 (B(OH)₂), 2924 (aromatic CH) 1654 (carbonyl), 1318, 1284, 1113, 875, 704 cm⁻¹. mp 103 °C. HRMS for trimeric anhydride C₃₉H₂₇BO₃ Calcd: 624.2086. Found: 624.2095.

**Site-Specific Chemical Modification** To 25 mg of the S166C mutant, purified as previously described and stored flash frozen in CHES buffer (2.5 mL; 70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5), at 20 °C was added in turn one of the methanethiosulfonate reagents (2a-d) (100 µL of a 0.2 M solution) in a PEG (10,000) coated polypropylene test tube, and the mixture agitated in an end-over-end rotator. Blank reactions containing 100 µL of solvent instead of the reagent solution were run in parallel. Each of the modification reactions was monitored spectrophotometrically (ε₄₁₀ = 8800 M⁻¹ cm⁻¹) on a Perkin Elmer Lambda 2 spectrophotometer, by specific activity measurements. After the reaction was quenched by dilution with MES buffer (5 mM MES, 2 mM CaCl₂, pH 6.5) at 0 °C, the specific activity of the CMM (10 µL) was determined in buffer (0.1 M TRIS pH 8.6, 0.005 % Tween 80 and 1% DMSO, with the suc-AAPF-pNA substrate (1mg/mL, purchased from Bachem Bioscience Inc.) at 25 °C. The reaction was terminated when the addition of a further 100 µL of methanethiosulfonate solution effected no further change in specific activity, generally within 30 min. The reaction solution was purified on a disposable
desalting column (Pharmacia Biotech PD-10, Sephadex G-25 M) pre-equilibrated with MES buffer (5 mM MES, 2 mM CaCl₂, pH 6.5) then dialyzed against 20 mM MES, 1 mM CaCl₂, pH 5.8 (3 x 1 L) at 4 °C and aliquoted into 0.5 - 1.5 mL volumes, flash frozen in liquid nitrogen and then stored at -20 °C. Modified enzymes were analyzed by nondenaturing gradient (8-25%) gels at pH 4.2, run towards the cathode on the Pharmacia Phast-System,™ and appeared as one single band.

**Electrospray Mass Spectrometry:** The CMMs were purified for ES-MS analysis, by FPLC (BioRad, Biologic System) on a Source 15 RPC matrix (17-0727-20 from Pharmacia) with 5% acetonitrile, 0.01% TFA as the running buffer and eluted with 80% acetonitrile, 0.01% TFA in a one step gradient. Electrospray mass spectra were recorded on a PE SCIEX API III Biomolecular Mass Analyzer. Mass: WT: Calcd: 26698. Found: 26694. S166C-S-a: Calcd: 26836. Found: 26832. S166C-S-b: Calcd: 26842. Found: 26844. S166C-S-c: Calcd: 26853. Found: 26851. S166C-S-d: Calcd: 26714. Found: 26708

**Regeneration of Unmodified Enzyme by Treatment with β-Mercaptoethanol:** To a solution of CMM (2.0 mg) in 250 μL of CHES-buffer (70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5) was added 10 μL of a solution of β-mercaptoethanol (1 M in 95% EtOH ). The reaction was monitored by specific activity measurements and in all cases the activity of the S166C cysteine parent was restored.

**Free Thiol Titration:** The residual free thiol content of the S166C CMMs, was determined spectrophotometrically by titration with Ellman’s reagent (ε₄₁₂ = 13600 M⁻¹ cm⁻¹) in phosphate buffer (0.25 M, pH 8.0).

**Active Site Titration:** The active enzyme concentration was determined as previously described³⁴ by monitoring fluoride release upon enzyme reaction with α-toluenesulfonyl fluoride (Aldrich Chemical Co. Inc.) as measured by a fluoride ion sensitive electrode (Orion Research 96-09). The active enzyme concentration determined in this way was used to calculate kinetic parameters for each CMM.
**Photolysis:** Enzyme solutions for photolysis were prepared in 3.5 mL quartz cuvettes and contained 3 mL of enzyme dissolved (2 mg/mL) in storage buffer (20 mM MES, 1 mM CaCl₂, pH 5.8). To each cuvette was added 10 μL of p-boronic acid benzophenone (for 10X samples, 0.02 M BBP in DMSO). For the blank reactions 10 μL of DMSO was added instead of BBP. The solutions were photolyzed at 4 °C in a Rayonet photoreactor equipped with 350 nm lamps. The dark controls were treated in the same way except the cuvettes were wrapped in aluminum-foil. At the time intervals indicated, 10 μL aliquots were withdrawn, diluted into 190 μL of storage buffer and specific activity measurements made. After photolysis, the enzymes were purified on a disposable desalting column (Pharmacia Biotech PD-10, Sephadex G-25 M) pre-equilibrated with MES buffer (3.5 mL, 5 mM MES, 2 mM CaCl₂, pH 6.5). The enzyme was eluted with MES buffer, dialyzed against storage buffer (20 mM MES, 1 mM CaCl₂, pH 5.8, (3 x 1 L) at 4 °C aliquoted, flash frozen and stored at -20 °C.

**Specific Activity Measurements for Time Course Monitoring:** For specific activity determinations, absorbance versus time measurements were made on a Perkin Elmer lambda 2 spectrophotometer. The spectrophotometer was zeroed against a 1 mL disposable cuvette containing 980 μL Tris buffer (0.1 M, 0.005% Tween80, pH 8.6) and 10 μL of the suc-AAPF-pNA substrate (100.0 mg/mL in DMSO, 160 mM). The enzyme aliquot (10 μL) was added and, after a 10 second delay, the absorbance reading was started and sampled with an interval of 0.2 s during 60 s at λ = 410 nm. The slope (in A s⁻¹) was determined by linear curve fit of the data below 5% conversion of the substrate and the rate of product formation, dP/dt determined from Beer's Law (A=εcι, where ε₄₁₀ = 8800 M⁻¹ cm⁻¹, ι = 1 cm). The enzyme activity was then calculated: specific activity (μmol min⁻¹ mg⁻¹) = (dP/dt M s⁻¹)(60 s min⁻¹)(10⁶ μmol mol⁻¹)(volume in cuvette L)/{[E] mg mL⁻¹}.
Tryptic Digests and HPLC-MS Analysis:26
From a stock solution of the SBL samples in storage buffer (20 mM MES, 1 mM CaCl₂, pH 5.8) was withdrawn 0.1 mg (≤100 µL) of enzyme which was mixed with water (≥200 µL) and 1N HCl (30 µL) to give a final concentration of 0.03 mg/mL protein in 0.09 N HCl. This mixture was chilled for 10 min on ice and any debris was removed by centrifugation for 2 min at 14,000 rpm. The supernatant was transferred to a clean eppendorf tube to which was added 35 µL TCA (50% w/v) to give a final concentration of 5% (w/v) TCA. This mixture was chilled for 15 min on ice and then the protein was collected by centrifugation (2 min at 14,000 rpm). The pellet was washed once with 1 mL of 90% (v/v) acetone (-20 °C) and briefly dried in vacuo. The pellet was resuspended in 12 µL of 0.4 M ammonium bicarbonate, containing 8 M urea and then incubated at 37 °C for 4 min. To this mixture was added 48 µL of 1% (w/v in H₂O) 1-O-octyl-β-D-glucopyranoside and 1 µL of trypsin (Sigma Chemical Co. TPCK-treated, 2.5 mg/mL in 1 mM HCl) and then the mixture was at 37 °C for 30 min after which time the reaction was terminated by adding 5 µL of TFA (10% v/v).

An aliquot of the tryptic digest was then analyzed on an Hewlett Packard HPLC Model 1090, which is equipped with a Vydac C-18 Reverse Phase Column Model 218T952 (2.1 x 250 mm), with a gradient of water + 0.1% TFA and CH₃CN + 0.1% TFA as the mobile phase, coupled to a ESI Ionization Mass Spectrophometer Model 9889B.

Peptide Sequencing: The isolated tryptic fragments were sequenced on the Applied Biosystems Model 473A Protein Sequencer.

Kinetic Measurements: Michaelis-Menten constants were determined at 25 °C by curve fitting (GraFit®, 3.03) of the initial rate data determined at eight concentrations (0.125 mM—4.0 mM) of the suc-AAPF-pNA substrate (ε₄₁₀ = 8800 M⁻¹ cm⁻¹) in 0.1 M phosphate buffer containing 0.5 M NaCl, 0.005% Tween 80, 1% DMSO, pH 7.5.
Determination of p-Boronic Acid Benzophenone Inhibition Constants: $K_i$s of BBP were determined in duplicate by the method of Waley. The progress curve without inhibitor was determined from an assay mixture containing 980 μL of buffer (0.1 M phosphate buffer containing 0.5 M NaCl, 0.005% Tween 80, pH 7.5), 5 μL DMSO and 5 μL of suc-AAPF-pNA substrate (50 mM in DMSO). This mixture was incubated in a water jacketed cell for 5 min. at 25 °C and the absorbance reading was set to zero prior to initiating the reaction by addition of 10 μL of enzyme solution (3.2 $\times$ 10$^{-6}$ to 1.1 $\times$ 10$^{-4}$ M in pH 5.8, 20 mM MES, 1 mM CaCl$_2$). The final volume of the assay mixture was 1 mL. The progress curve with inhibitor was determined similarly but with 5 μL of the BBP inhibitor solution (7.24 $\times$ 10$^{-3}$ to 1.4 $\times$ 10$^{-1}$ M in DMSO) being added instead of 5 μL DMSO, to a final volume of 1 mL. Absorbance versus time measurements were recorded on a Perkin Elmer lambda 2 spectrophotometer and points for calculation were taken at 15, 18, 21, 24, 27, 30, 33 and 36% substrate conversion.
References


30. Amino acid sequencing was done by Sue Middlebrook at Genencor International Inc.

31. Tryptic Digest and HPLC ES-MS Analysis for this sample was performed by Dr. Christian Paech at Genencor International Inc.


Chapter 5

Reproduced in part with permission from J. Am. Chem. Soc. submitted.
Unpublished work copyright 1999, American Chemical Society
Towards Tailoring The Specificity of the S, Pocket of Subtilisin
B. lentus: Chemical Modification of Mutant Enzymes as a
Strategy for Removing Specificity Limitations

Introduction

For both protein chemistry\textsuperscript{1-3} and organic synthesis applications,\textsuperscript{4-7} it is desirable to have available a diverse toolbox of inexpensive proteases with high selectivity and diverse substrate preferences. Since WT enzymes will never accept all substrate structures, it is attractive to contemplate the tailoring of a readily available protease in order to expand its substrate specificity with the ultimate goal of creating any desired specificity at will.

WT-SBL has a marked preference for substrates with large hydrophobic uncharged P\textsubscript{1} residues. In this study, we explore tailoring the S\textsubscript{1} pocket of subtilisin Bacillus lentus (SBL) to also accept small hydrophobic and positively or negatively charged P\textsubscript{1} residues using the combined site-directed mutagenesis chemical modification approach (CMM), as illustrated in Scheme 5.1.\textsuperscript{8-12} In order to achieve this broadened P\textsubscript{1} tolerance a simplistic strategy of steric and electrostatic complementarity was applied.\textsuperscript{13} Employing the crystal structure of SBL as our guide,\textsuperscript{14} the Ser166 residue, which is located at the bottom of the S, pocket and whose side chain points inward toward the pocket, was chosen for mutagenesis to cysteine and subsequent chemical modification. Firstly, to expand SBL’s specificity toward small uncharged P\textsubscript{1} residues, such as the small P\textsubscript{1} Ala residue of the suc-AAPA-pNA substrate, large groups were introduced at position 166 in S\textsubscript{1} such as benzyl (-c) and cyclohexyl (-f) with a view to reducing the volume of S\textsubscript{1} and inducing a better fit of small P\textsubscript{1} groups, thereby conferring elastase-like\textsuperscript{15} substrate specificity on SBL. Then, to expand SBL’s specificity toward positively charged P\textsubscript{1} residues, such as the P\textsubscript{1} Arg residue of the suc-AAPR-pNA substrate, we introduced negatively charged groups at position S166C in S\textsubscript{1} such as the ethylsulfonato (-b) moiety to elicit complementary
electrostatic attractions, thus making SBL trypsin-like in its specificity. Similarly, to expand SBL's specificity toward negatively charged P residues, such as the negatively charged P Glu residue of the suc-AAPE-pNA substrate, we introduced the positively charged ethylamino (-a) group at position S166C in S.

**Scheme 5.1**

\[
\text{SBL-SH} + \overset{\text{O}}{\overset{\text{S}}{\overset{\text{R}}{\text{S166C}}}}} \rightarrow \text{SBL-S-S-R}
\]

\[
\text{pH 9.5}
\]

\[
\text{S166C} \quad \text{1a-1g} \quad \text{S166C-S-R}
\]

\[
R = \begin{cases} 
(a) & \text{NH}_2^+ \\
(b) & \text{SO}_3^- \\
(c) & \text{COOH} \\
(d) & \text{COOH} \\
(e) & \text{COOH} \\
(f) & \text{COOH}
\end{cases}
\]

**Results**

In order to expand the substrate specificity of SBL towards small hydrophobic P residues such as Ala, the S166C mutant was modified with the large hydrophobic MTS reagents 1c, 1e-g by the general method described previously with the goal of restricting the volume of the S, pocket. Broadening the substrate specificity of SBL to accept positively charged P residues such as Arg was addressed by modifying S166C with the negatively charged MTS reagents 1b, 1d, 1h-1j. Conversely, for conferring acceptance of negatively charged P residues such as Glu, S166C was modified with the complementary.
positively charged MTS reagent 1a. The preparations of the requisite MTS reagents 1c, 1e-f, and 1d, have been reported previously and the steroidyl MTS reagent 1g was prepared from cholic acid by the same methodology. Each of the CMMs obtained was characterized in order to establish its purity and integrity. Titration of the CMMs with Ellman’s reagent showed a residual thiol content of less than 2% in all cases, demonstrating that the MTS reactions were virtually quantitative. Mass analyses of the CMMs by electrospray mass spectrometry were consistent (± 6 Da) with the calculated mass. The purities of the modified enzymes was assessed by native-PAGE and in all cases only one band was visible. Furthermore, relative to WT, the negatively charged CMMs S166C-S-b-d, and -i to -j displayed retarded mobility in the direction of the cathode, while the positively charged S166C-S-a CMM displayed greater mobility, as expected. That modification of cysteine is wholly responsible for altered activity was established by the absence of reaction of WT-SBL with the MTS reagents. Also, the modifications are fully reversible by treatment of each of the CMMs with β-mercaptoethanol, further verifying that chemical modification at cysteine was solely responsible for the observed changes in activity. The total amount of active enzyme was determined by titration with phenylmethylsulfonyl fluoride.
Figure 5.1 $k_{cat}/K_M$ Screen of S166C CMMs with suc-AAPF/A/R/E-pNA Substrates.

Initially, three CMMs S166C-S-a, -b and -c, with a positive, a negative, and with a large hydrophobic side chain respectively, were subjected to a $k_{cat}/K_M$ screen with each of the test substrates, suc-AAPF/A/R/E-pNA in order to identify any induced complementary electrostatic or improved hydrophobic interactions. While as expected $k_{cat}/K_M$s with the standard suc-AAPF-pNA were lowered, the $k_{cat}/K_M$s of the CMMs whose $S_1$ sites were tailored toward the Ala, Arg or Glu $P_1$ residues, improved with the appropriate substrate. This is illustrated in Figure 5.1 in the higher activity of S166C-S-c with suc-AAAPA-pNA, of S166C-S-b with suc-AAPR-pNA, and of S166C-S-a with suc-AAAPE-pNA compared to WT.

Following the validation of the general design strategy from this initial screen, a more complete kinetic analysis was undertaken. The substrate specificity of each of the CMMs was evaluated kinetically with the standard large hydrophobic $P_1$ residue containing substrate, suc-AAPF-pNA. In addition, the S166C CMMs modified with the large hydrophobic MTS reagents 1c, 1e-g, were evaluated with the targeted small hydrophobic $P_1$ residue containing substrate, suc-AAAPA-pNA. The S166C CMMs modified with the negatively charged MTS reagents 1b, 1d, 1h-j were evaluated with the positively charged $P_1$ residue containing substrate, suc-AAPR-pNA. The S166C CMMs modified with the
positively charged MTS reagent 1a was evaluated with the negatively charged P, residue containing substrate, suc-AAPE-pNA. The results are summarized in Table 5.1.

Table 5.1 Kinetic Evaluation\(^{(a)}\) of Altered S1 Pocket Specificity

<table>
<thead>
<tr>
<th>Entry</th>
<th>Enzyme</th>
<th>Substrate</th>
<th>(K_m) mM</th>
<th>(k_{cat}) s(^{-1})</th>
<th>(k_{cat}/K_m) s(^{-1})mM(^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WT</td>
<td>suc-AAPF-pNA</td>
<td>0.73 ± 0.08</td>
<td>153 ± 4</td>
<td>209 ± 15</td>
</tr>
<tr>
<td>2</td>
<td>S166C-S-a</td>
<td>suc-AAPF-pNA</td>
<td>0.68 ± 0.04</td>
<td>50 ± 1</td>
<td>74 ± 5</td>
</tr>
<tr>
<td>3</td>
<td>S166C-S-b</td>
<td>suc-AAPF-pNA</td>
<td>1.34 ± 0.08</td>
<td>25.0 ± 0.7</td>
<td>19 ± 1</td>
</tr>
<tr>
<td>4</td>
<td>S166C-S-c</td>
<td>suc-AAPF-pNA</td>
<td>1.17 ± 0.06</td>
<td>23.1 ± 0.5</td>
<td>20 ± 1</td>
</tr>
<tr>
<td>5</td>
<td>S166C-S-d</td>
<td>suc-AAPF-pNA</td>
<td>1.6 ± 0.2</td>
<td>47 ± 3</td>
<td>29 ± 4</td>
</tr>
<tr>
<td>6</td>
<td>S166C-S-e</td>
<td>suc-AAPF-pNA</td>
<td>1.09 ± 0.07</td>
<td>82 ± 2</td>
<td>75 ± 5</td>
</tr>
<tr>
<td>7</td>
<td>S166C-S-f</td>
<td>suc-AAPF-pNA</td>
<td>0.70 ± 0.05</td>
<td>4.8 ± 0.1</td>
<td>6.90 ± 0.05</td>
</tr>
<tr>
<td>8</td>
<td>S166C-S-g</td>
<td>suc-AAPF-pNA</td>
<td>0.74 ± 0.07</td>
<td>29 ± 1</td>
<td>41 ± 4</td>
</tr>
<tr>
<td>9(^{(b)})</td>
<td>S166C-S-h</td>
<td>suc-AAPF-pNA</td>
<td>1.52 ± 0.06</td>
<td>48 ± 1</td>
<td>31 ± 1</td>
</tr>
<tr>
<td>10(^{(b)})</td>
<td>S166C-S-i</td>
<td>suc-AAPF-pNA</td>
<td>2.26 ± 0.10</td>
<td>67 ± 2</td>
<td>30 ± 2</td>
</tr>
<tr>
<td>11(^{(b)})</td>
<td>S166C-S-j</td>
<td>suc-AAPF-pNA</td>
<td>2.46 ± 0.11</td>
<td>76 ± 2</td>
<td>31 ± 2</td>
</tr>
<tr>
<td>12</td>
<td>WT</td>
<td>suc-AAPA-pNA</td>
<td>2.0 ± 0.1</td>
<td>17.7 ± 0.3</td>
<td>8.8 ± 0.4</td>
</tr>
<tr>
<td>13</td>
<td>S166C-S-c</td>
<td>suc-AAPA-pNA</td>
<td>0.8 ± 0.1</td>
<td>6.8 ± 0.3</td>
<td>9 ± 1</td>
</tr>
<tr>
<td>14</td>
<td>S166C-S-e</td>
<td>suc-AAPA-pNA</td>
<td>1.90 ± 0.03</td>
<td>6.8 ± 0.4</td>
<td>3.6 ± 0.6</td>
</tr>
<tr>
<td>15</td>
<td>S166C-S-f</td>
<td>suc-AAPA-pNA</td>
<td>1.90 ± 0.07</td>
<td>28.2 ± 0.4</td>
<td>14.8 ± 0.6</td>
</tr>
<tr>
<td>16</td>
<td>S166C-S-g</td>
<td>suc-AAPA-pNA</td>
<td>1.74 ± 0.04</td>
<td>9.65 ± 0.07</td>
<td>5.54 ± 0.3</td>
</tr>
<tr>
<td>17</td>
<td>WT</td>
<td>suc-AAPR-pNA</td>
<td>7.2 ± 0.7</td>
<td>0.16 ± 0.01</td>
<td>0.022 ± 0.002</td>
</tr>
<tr>
<td>18</td>
<td>S166C-S-b</td>
<td>suc-AAPR-pNA</td>
<td>3.4 ± 0.3</td>
<td>0.17 ± 0.01</td>
<td>0.050 ± 0.005</td>
</tr>
<tr>
<td>19</td>
<td>S166C-S-d</td>
<td>suc-AAPR-pNA</td>
<td>5.5 ± 1.1</td>
<td>0.68 ± 0.08</td>
<td>0.12 ± 0.03</td>
</tr>
<tr>
<td>20</td>
<td>S166C-S-h</td>
<td>suc-AAPR-pNA</td>
<td>8.2 ± 0.9</td>
<td>0.35 ± 0.02</td>
<td>0.041 ± 0.005</td>
</tr>
<tr>
<td>21</td>
<td>S166C-S-i</td>
<td>suc-AAPR-pNA</td>
<td>5.3 ± 0.5</td>
<td>0.43 ± 0.02</td>
<td>0.080 ± 0.008</td>
</tr>
<tr>
<td>22</td>
<td>S166C-S-j</td>
<td>suc-AAPR-pNA</td>
<td>5.2 ± 0.6</td>
<td>1.06 ± 0.07</td>
<td>0.20 ± 0.03</td>
</tr>
<tr>
<td>23</td>
<td>WT</td>
<td>suc-AAPE-pNA</td>
<td>4.4 ± 0.4</td>
<td>1.75 ± 0.08</td>
<td>0.40 ± 0.04</td>
</tr>
<tr>
<td>24</td>
<td>S166C-S-a</td>
<td>suc-AAPE-pNA</td>
<td>1.9 ± 0.1</td>
<td>14.5 ± 0.3</td>
<td>7.6 ± 0.4</td>
</tr>
</tbody>
</table>

(a) Michaelis-Menten constants were measured by the initial rates method in pH 8.6 Tris-HCl buffer at 25°C with the suc-AAPF-pNA substrate \(\)\(^{(b)}\) Determined by Dr. Ben Davis, Reference 17.
**Discussion**

The significant substrate preference of WT-SBL for large hydrophobic P₁ residues is apparent from its preference for the Phe P₁ residue of the standard suc-AAPF-pNA substrate by a factor of 9500-fold over the small P₁ residue of suc-AAPA-pNA, by a factor of 24-fold compared to the positively charged P₁ residue of suc-AAPR-pNA, and by a factor of 522-fold compared to the negatively charged P₁ residue of suc-AAPE-pNA (Table 5.1: entries 1,12,17 and 23). These substrate preferences are attributed to changes both in binding as reflected by $K_M$ and in turnover number, $k_{cat}$. In addition, as expected, with suc-AAPF-pNA, the WT enzyme is the best catalyst, and its conversion to the CMMs resulted in $k_{cat}/K_M$ decreases of up to 34-fold (Table 5.1: entries 2-11).

To tailor the substrate specificity of SBL toward small hydrophobic P₁ residues such as Ala, the simplistic approach of trying to partially fill up the S₁ binding cleft was followed by preparing the S166C-S-CH₂C₆H₅ (-c), S166C-S-CH₂(CH₂)₈CH₃ (-e), S166C-S-CH₂-c-C₆H₁₁ (-f), and S166C-S-steroidyl (-g) CMMs. This design strategy attempted to mimic the function of the bulky S₁-pocket side chains of α-lytic protease²⁰-²² and of elastase,¹⁵ for which both enzymes exhibit substantial preference for the suc-AAPA-pNA substrate over suc-AAPF-pNA. These CMMs (S166C-S-c,-e,-f, -g. Table 5.1:entries 12-16) were then evaluated with the suc-AAPA-pNA substrate. All revealed slightly improved binding compared to WT, with the greatest improvement in $K_M$ being 2-fold for the S166C-S-CH₂C₆H₅ (-c) CMM. Of the four CMMs, only S166C-S-CH₂C₆H₁₁ (-f) showed both an improved $k_{cat}$ and an improved $k_{cat}/K_M$ compared to WT. While this design strategy yielded only one CMM with an increased preference for the small Ala P₁ residue, all of these modifications effectively excluded the larger Phe P₁ residue preferred by WT-SBL (Table 5.1:entries 4, 6-8). Overall, the selectivities, with respect to $k_{cat}/K_M$, for the suc-AAPA-pNA substrate compared to the suc-AAPF-pNA substrate were improved by 11-fold for S166C-S-CH₂C₆H₅ (-c), 1.1-fold for S166C-S-CH₂(CH₂)₈CH₃ (-e), 51-fold for S166C-S-CH₂-c-C₆H₁₁,
and 3.2-fold for S166C-S-steroidyl (-g), compared to WT. The extent of improvement in P₁ Ala selectivity may be a reflection of the orientation of the R side chain of the CMMs, with the side chains of S166C-S-CH₂C₆H₅ (-c), and -CH₂-c-C₆H₅, (-f), behaving as though directed into the pocket and the side chains of S166C-S-CH₂(CH₂)₆CH₃ (-e) and S166C-S-steroidyl (-g) outward.

Engineering complementary steric fit in the S₁ pocket has already proven to be challenging and previous attempts to alter the steric complementarity of subtilisins have resulted in modest but inconsistent successes with maximal improvements of up to 10-fold. Starting with subtilisin Yab, whose specificity was already elastase-like, the G127A mutation which was designed to fill up its S₁ pocket and thus increase the selectivity for small hydrophobic P₁ residues such as Ala, effected a 10-fold improvement in $k_{cat}/K_M$ with the suc-AAPA-pNA substrate. Similarly, the G166I mutant of subtilisin BPN' caused a 10-fold improvement in $k_{cat}/K_M$ with the suc-AAPA-pNA substrate. However, the G127V mutant of subtilisin E caused a decrease in $k_{cat}/K_M$ with the suc-AAPA-pNA substrate, compared to WT.

Improving the substrate specificity of SBL toward positively charged P₁ residues such as Arg, making it more trypsin-like, was based on mimicking the common motif of the high negative charge density of acidic residues inducing a substrate preference for positively charged substrates. This was achieved with the S166C-S-CH₂CH₂SO₃⁻ (-b), S166C-S-CH₂(CH₂)₂CH₂COO⁻ (-h), S166C-S-CH₂C₆H₄-3,5-(COO⁻)₂ (-d), S166C-S-CH₂CH₂C(CH₃)(COO⁻)₂ (-i) and S166C-S-CH₂CH₂C(COO⁻)₃ (-j) CMMs. Evaluation of each of these CMMs with the suc-AAPR-pNA substrate revealed improved binding of up to 2-fold compared to WT, with the exception of S166C-S-CH₂(CH₂)₂CH₂COO⁻ (-h) for which $K_M$ was slightly increased (Table 5.1: entries 17-22). The general success of this approach is evident since all of the CMMs with a negatively charged R side chain showed an improved $k_{cat}$ of up to 7-fold and an improved $k_{cat}/K_M$ of up to 9-fold with suc-AAPR-pNA compared to WT (Table 5.1: entries 17-22). Furthermore, each additional negative charge introduced at position 166 resulted in an approximate
doubling in $k_{cat}/K_M$ with the complementary suc-AAPR-pNA substrate (Figure 5.2). This additive effect contrasts the previous observation that while noninteracting remote mutations are additive, multiple adjacent mutations are not.\textsuperscript{30} Overall, the selectivities, with respect to $k_{cat}/K_M$, for the suc-AAPR-pNA substrate compared to the suc-AAPF-pNA substrate were improved 25-fold for S166C-S-CH\textsubscript{2}CH\textsubscript{2}SO\textsubscript{3}\textsuperscript{-} (-b), 13-fold for S166C-S-CH\textsubscript{2}(CH\textsubscript{2})\textsubscript{2}CH\textsubscript{2}COO\textsuperscript{-} (-h), 39-fold for S166C-S-CH\textsubscript{2}C\textsubscript{6}H\textsubscript{4}-3,5-(COO\textsuperscript{-})\textsubscript{2} (-d), 25-fold for S166C-S-CH\textsubscript{2}CH\textsubscript{2}C(\textsubscript{2}H\textsubscript{3})(COO\textsuperscript{-})\textsubscript{2} (-i) and 61-fold for S166C-S-CH\textsubscript{2}CH\textsubscript{2}C(COO\textsuperscript{-})\textsubscript{3} (-j), compared to WT.

![Figure 5.2 Negative Charge at Position 166C versus Activity: Plot of negative charge on R at position 166 versus $k_{cat}/K_M$ with the positively charged P\textsubscript{i} residue containing substrate, suc-AAPR-pNA. The points on the line are for the aliphatic series of -R groups.](image)

These results compare favourably to previous attempts to induce charge complementarity by site-directed mutagenesis of subtilisins and to make its specificity trypsin-like.\textsuperscript{31-33} For example, the G166D mutation of subtilisin BPN' was designed to introduce a negative charge in its S\textsubscript{i} pocket in order to engender trypsin-like specificity. However, rather than the anticipated increase in specificity for a complementary positively charged P\textsubscript{i} residue containing substrate, a 2.5-fold decrease in $k_{cat}/K_M$ with the suc-AAPR-pNA substrate was
observed.\textsuperscript{32} Furthermore, despite the introduction of two negatively charged amino acids in the S\textsubscript{1} pocket of subtilisin BPN', one at position 156 and one at position 166, only a maximal 2-fold improvement in $k_{\text{cat}}/K_M$ with the suc-AAPR-pNA substrate was observed compared to WT.$^{31}$

Conversely, to tailor the specificity of SBL toward negatively charged P\textsubscript{1} residues such as Glu, a design strategy of introducing a complementary positive charge in the S\textsubscript{1} binding cleft by the CMM approach was adopted (S166C-S-a). This strategy was based on mimicking the specificity determinants of the serine proteases pronase\textsuperscript{34-35} and granzyme B,\textsuperscript{36-38} which exhibit a substrate preference for negatively charged P\textsubscript{1} residues, and whose S\textsubscript{1} pockets are lined with positively charged residues. In fact, S166C-S-CH\textsubscript{2}CH\textsubscript{2}NH\textsubscript{3}\textsuperscript{+} (-a) displayed a remarkable 19-fold increase in $k_{\text{cat}}/K_M$ with the suc-AAPE-pNA substrate compared to WT. This enhancement is due to a combination of better binding, evident from the 2-fold lower $K_M$, and 8-fold higher $k_{\text{cat}}$ (Table 5.1: Entries 23, 24).

The induction of electrostatic complementarity was most unequivocally demonstrated by the 54-fold improvement in suc-AAPE-pNA to suc-AAPF-pNA substrate selectivity, with respect to $k_{\text{cat}}/K_M$, for S166C-S-CH\textsubscript{2}CH\textsubscript{2}NH\textsubscript{3}\textsuperscript{+} (-a) compared to WT.

Moreover, these results compare favourably with previous site-directed mutagenesis approaches toward tailoring the preference of subtilisin BPN' for substrates with negatively charged P\textsubscript{1} residues.$^{31,39}$ While, $k_{\text{cat}}/K_M$ for the suc-AAPE-pNA substrate was improved, the reported mutants did not exhibit selectivity for negatively charged P\textsubscript{1} residues compared to hydrophobic ones.$^{31,39}$

For example, the G166R and G166K mutants of subtilisin BPN', which although they caused 23- and 340-fold improvements respectively in $k_{\text{cat}}/K_M$ with the suc-AAPE-pNA substrate, compared to WT, displayed even higher $k_{\text{cat}}/K_M$s with the suc-AAPN/M/K-pNA substrates, which lack the complementary charged P\textsubscript{1} residue.$^{31}$ The broad specificity of the G166K mutant has been attributed to the flexible conformation of the K166 side chain.$^{39}$
Figure 5.3 Modelling of WT-SBL and S166C-S-CH₂CH₂NH₃⁺ with AAPE: (a) Active site of WT-SBL showing position of AAPE, as determined by molecular modelling. (b) Active site of S166C-S-a showing position of AAPE, as determined by molecular modelling. In both cases, the P₁ residue of AAPE binds well into the S₁ pocket and the γ-carboxylate of glutamic acid forms a hydrogen bond with the backbone amide nitrogen of Asn155. The terminal carboxylate of AAPF is hydrogen bonded to the HN₂ of His64 and to the HO₇ of Ser221 in the WT enzyme (a) but only to the HN₂ of His64 in the S166C-S-CH₂CH₂NH₃⁺ CMM (b).
Since the S166C-S-CH₂CH₂NH₃⁺ (-a) and suc-AAPE-pNA CMM-substrate pair exhibited the greatest $k_{cat}/K_M$ improvement relative to WT, at 19-fold (Table 5.1), more detailed insights into the molecular basis of their interaction was sought using molecular modelling. Using the modelling strategy reported previously,⁹ the product inhibitor, AAPE was bound to WT-SBL and to the S166C-S-CH₂CH₂NH₃⁺ (-a) CMM. The energies of these structures were minimized. As shown in Figure 5.3, the binding conformations for both the WT and S166C-S-CH₂CH₂NH₃⁺ structures are remarkably similar. However, the ammonium moiety of the S166C-S-CH₂CH₂NH₃⁺ CMM side chain is oriented toward the carboxylate of the glutamic acid P₁ residue, and although it is not within salt-bridge distance (Figure 5.3b), an additional favourable coulombic interaction between the ammonium side chain of S166C-S-CH₂CH₂NH₃⁺ and the carboxylate of the glutamic acid P₁ residue causes the observed 19-fold improvement in $k_{cat}/K_M$ for this CMM substrate pair, compared to WT.

The effectiveness of the general design strategy is validated, since for each of the Ala, Arg, and Glu P₁ residues at least one and up to five of the designed CMMs exhibit improved $k_{cat}/K_М$ compared to WT.
**Experimental**

**Methanethiosulfonate Reagents:** Sulfonatoethyl methanethiosulfonate (1a) and ethylammonium methanethiosulfonate (1b) were purchased from Toronto Research Chemicals (2 Brisbane Rd., Toronto, ON, Canada). Reagents 1c-1f and 1h-1j were prepared as previously described.

**Site-Specific Chemical Modification:** To 25 mg of a S166C mutant, purified as previously described8 and stored flash frozen in CHES buffer (2.5 mL; 70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5) was added at 20 °C one of the methanethiosulfonate reagents (1a-g) (100 μL of a 0.2 M solution), in a PEG (10,000) coated polypropylene test tube, and the mixture agitated in an end-over-end rotator. Blank reactions containing 100 μL of solvent instead of the reagent solution were run in parallel. Each of the modification reactions was monitored spectrophotometrically (ε₄₁₀ = 8800 M⁻¹ cm⁻¹) on a Perkin Elmer Lambda 2 spectrophotometer, by specific activity measurements. After the reaction was quenched by dilution in MES buffer (5 mM MES, 2 mM CaCl₂, pH 6.5) at 0 °C, the specific activity of the CMM (10 μL), was determined in buffer (0.1 M TRIS pH 8.6, 0.005 % Tween 80, and 1% DMSO) with the succ-AAPF-pNA substrate (1mg/mL) (purchased from Bachem Bioscience Inc.) at 25 °C. The reaction was terminated when the addition of a further 100 μL of methanethiosulfonate solution effected no further change in specific activity, generally within 30 min. The reaction solution was purified on a disposable desalting column (Pharmacia Biotech PD-10, Sephadex G-25 M) pre-equilibrated with MES buffer (5 mM MES, 2 mM CaCl₂, pH 6.5) then dialyzed against 20 mM MES, 1 mM CaCl₂, pH 5.8 (3 x 1 L) at 4 °C aliquoted into 0.5 - 1.5 mL volumes, flash frozen in liquid nitrogen and then stored at -20 °C. Modified enzymes were analyzed by non-denaturing gradient (8-25%) gels at pH 4.2, run towards the cathode on the Pharmacia Phast-System,™ and appeared as one single band.

**Electrospray Mass Spectrometry:** Prior to ES-MS analysis, CMMs were purified by FPLC (BioRad, Biologic System) on a Source 15 RPC matrix (17-
0727-20 from Pharmacia) with 5% acetonitrile, 0.01% TFA as the running buffer and eluted with 80% acetonitrile, 0.01% TFA in a one step gradient. ES-MS data were acquired using a PE SCIEX API III Biomolecular mass spectrometer. Mass:


**Regeneration of Unmodified Enzyme by Treatment with β-Mercaptoethanol:**

To a solution of CMM (2.0 mg) in 250 μL of CHES-buffer (70 mM CHES, 5 mM MES, 2 mM CaCl₂, pH 9.5) was added 10 μL of a solution of β-mercaptoprothanol (1 M in 95% EtOH). The reaction was monitored by specific activity measurements and in all cases the activity of the cysteine parent enzyme was restored.

**Free Thiol Titration:** The free thiol content of S166C CMMs was determined spectrophotometrically by titration with Ellman's reagent ($\varepsilon_{412} = 13600$ M⁻¹ cm⁻¹) in phosphate buffer 0.25 M, pH 8.0.

**Active Site Titrations:** The active enzyme concentration was determined as previously described by monitoring fluoride release upon enzyme reaction with α-toluenesulfonyl fluoride (Aldrich Chemical Co. Inc.) as measured by a fluoride ion sensitive electrode (Orion Research 96-09). The active enzyme concentration determined in this way was used to calculate kinetic parameters for each CMM.

**Kinetic Measurements:** Michaelis-Menten constants were measured at 25 °C by curve fitting (GraFit® 3.03) of the initial rate data determined at eight concentrations (0.125 mM–8.0 mM) of the suc-AAPX-pNA substrate (Bachem California Inc. Torrance, CA) in buffer (pH 8.6, 0.1 M Tris-HCl, 0.005% Tween 80, 1% DMSO), ($\varepsilon_{410} = 8800$ M⁻¹ cm⁻¹).

**Molecular Modelling:** The X-ray structure of subtilisin *Bacillus lento* was used as the starting point for calculations on the wild type and chemically modified
mutant enzymes. The enzyme setup was performed with Insight II. To create initial coordinates for the minimization, hydrogens were added at the pH used for kinetic measurements. This protonated all Lys and Arg residues and the N-terminus and deprotonated all Glu and Asp residues and the C-terminus. In addition, the active site His64 was protonated. The model system with the Ala-Ala-Pro-Phe (from crystal structure) product inhibitor bound in the S_{1} - S_{4} pocket was solvated with a 5 Å layer of water molecules giving a total number of water molecules of 1143 in this system. The overall charge of the enzyme-inhibitor complex resulting from this setup was +4 for the WT enzyme. Energy simulations were performed with the Discover program, on a Silicon Graphics Iris Indigo computer, using the consistent valence force field function (CVFF). A non-bonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. The energy of the structure of the WT enzyme was minimized in stages, with initially only the water molecules being allowed to move, then the water molecules and the amino acid side chains, and then the entire enzyme. The mutated and chemically modified enzymes were generated using the Builder module of Insight. Then the amino acid side chains within a 10 Å radius of the α-carbon of the mutated residue were energy minimized while all other residues were constrained, then all of the atoms within a 10 Å shell were energy minimized, followed by minimization of the whole system. To examine the effect of a different P_{1} residue (Glu), the Phe to Glu mutation of the product inhibitor was constructed using Insight II, and then this structure was energy minimized as above.
References


18. Synthesized by Dr. Xiao Shang.


42. Insight II [Biosym Technologies, Inc. San Diego, CA, USA]. Version 2.3.0.

Chemical Modification of the N62C Cysteine Mutant of Subtilisin
*B. lentus* Can Create Better Catalysts than the Wild-type Enzyme

Introduction

All too often, site-directed mutagenesis creates mutant enzymes with lower-than-WT activities. Therefore, one of the goals of the current study was to create chemically modified mutant enzymes (CMMs) that could at least match the activity of the WT parent. In addition, the specificity changes which could be induced in the $S_2$ pocket were of interest. With these objectives in mind, using SBL's X-ray structure as a guide, residue Asn62 which is located in the $S_2$ pocket and is adjacent to the catalytic His64 residue (Figure 6.1) was selected for mutagenesis to cysteine and subsequent reaction with eleven 1a-k methanethiosulfonate reagents (Scheme 6.1).

![Image of active site of SBL indicating position of N62:Catalytic triad residues Asp32, His64, and Ser221, and the S2 pocket residue Asn62 chosen for mutation and modification are shown.](image)

Figure 6.1  Active Site of SBL Indicating Position of N62: The catalytic triad residues Asp32, His64, and Ser221, and the $S_2$ pocket residue Asn62 chosen for mutation and modification are shown.
Results and Discussion

Each of the CMMs N62C-S-a to-k were prepared and characterized as described previously. Briefly, the N62C-SBL mutant was treated with the methanethiosulfonate reagents (1a-k) as shown in Scheme 6.1, with the modification reactions being complete within 30 min. Titration of the CMMs with Ellman's reagent established that reactions were quantitative with the free thiol content of the CMMs being less than 2% in all cases. The CMMs were purified on a disposable Sephadex G-25 desalting column and analysis of each CMM by electrospray mass spectrometry was consistent (± 6 Da) with the calculated mass. The purities of the modified enzymes were assessed by native-PAGE and in all cases only one band was visible. Furthermore, treatment of SBL-WT, in which no cysteine is present, with each of 1a-k resulted in no change in activity. Additionally, β-mercaptoethanol treatment of the CMMs restored activity to that of the parent cysteine mutant, verifying that chemical modification at cysteine is solely responsible for the observed changes in activity and that the modification is fully reversible. The active enzyme concentration was determined by
monitoring fluoride release upon reaction of each CMM with phenylmethylsulfonyl fluoride (PMSF).4

The kinetic data obtained, with the standard suc-AAPF-pNA reference substrate, are summarized in Table 6.1. Despite the breadth of the structural range of the R-groups introduced, all of the chemical modifications resulted in CMMs with higher $k_{cat}/K_m$ than its cysteine mutant enzyme predecessor. However, introductions of the charged sulfonatoethyl ($R = j$) and aminoethyl ($R = k$) groups are clearly only marginally beneficial relative to those of any of the hydrophobic moieties of 1a-i. It is evident that the more hydrophobic the introduced modifying group, the higher the $k_{cat}$ values become. The only exception is for the decyl modified CMM, N62C-S-g, for which $k_{cat}$ is only slightly increased compared to WT. Whereas the $K_m$ for the SBL-N62C mutant itself is almost 3-fold higher than for SBL-WT, the $K_m$ values for the CMMs with small hydrophobic modifications (-a to -e) are only increased slightly (Table 6.1). However, for the larger more hydrophobic (-f to -i) modified enzymes, binding is unaltered or improved slightly. For both the negatively and positively charged N62C-S-j and N62C-S-k respectively, an increase in $K_m$ is observed. This general trend of improved $K_m$ for hydrophobic -R groups and poorer $K_m$ for hydrophilic modifications reinforce the trends in $k_{cat}$. 
Table 6.1. Kinetic Parameters for N62C CMMs

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_M$ (mM)</th>
<th>$k_{\text{cat}}/K_M$ (s$^{-1}$·mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>48 ± 2</td>
<td>0.55 ± 0.06</td>
<td>87 ± 10</td>
</tr>
<tr>
<td>N62C</td>
<td>61 ± 2</td>
<td>1.49 ± 0.1</td>
<td>41 ± 3</td>
</tr>
<tr>
<td>N62C-S-a</td>
<td>45.5 ± 0.4</td>
<td>0.69 ± 0.02</td>
<td>66 ± 2</td>
</tr>
<tr>
<td>N62C-S-b</td>
<td>66 ± 2</td>
<td>0.77 ± 0.05</td>
<td>86 ± 6</td>
</tr>
<tr>
<td>N62C-S-c</td>
<td>76 ± 2</td>
<td>0.84 ± 0.09</td>
<td>90 ± 10</td>
</tr>
<tr>
<td>N62C-S-d(b)</td>
<td>96 ± 2</td>
<td>0.78 ± 0.04</td>
<td>123 ± 6</td>
</tr>
<tr>
<td>N62C-S-e(b)</td>
<td>138 ± 4</td>
<td>0.81 ± 0.06</td>
<td>170 ± 13</td>
</tr>
<tr>
<td>N62C-S-f(b)</td>
<td>136 ± 6</td>
<td>0.54 ± 0.06</td>
<td>252 ± 30</td>
</tr>
<tr>
<td>N62C-S-g(b)</td>
<td>69 ± 3</td>
<td>0.35 ± 0.06</td>
<td>197 ± 35</td>
</tr>
<tr>
<td>N62C-S-h</td>
<td>135 ± 8</td>
<td>0.48 ± 0.09</td>
<td>281 ± 55</td>
</tr>
<tr>
<td>N62C-S-i</td>
<td>70 ± 2</td>
<td>0.37 ± 0.03</td>
<td>189 ± 16</td>
</tr>
<tr>
<td>N62C-S-j</td>
<td>66 ± 2</td>
<td>0.83 ± 0.05</td>
<td>80 ± 5</td>
</tr>
<tr>
<td>N62C-S-k</td>
<td>63 ± 3</td>
<td>1.2 ± 0.1</td>
<td>52 ± 5</td>
</tr>
</tbody>
</table>

(a) Kinetic constants were obtained by nonlinear regression of initial rates determined at eight substrate concentrations (0.25 mM–4.0 mM, suc-AAPF-pNA) at 25 °C in 0.1 M phosphate buffer, pH 7.5, containing 0.5 M NaCl, 1% DMSO. (b) Determined by Dr. Michele Stabile

Upon modification of N62C with 1a-f, $k_{\text{cat}}/K_M$ increases monotonically as the length of the R-side chain grows. The greater importance of chain length over steric volume in this regard is demonstrated by the higher $k_{\text{cat}}/K_M$ benefit conferred by n-pentyl (N62C-S-e) than neo-pentyl (N62C-S-d) modification. For acyclic alkyl modifications, the maximum activity augmentation is seen for N62C-S-f, for which $R = n$-hexyl and whose $k_{\text{cat}}/K_M$ is 6-fold higher than for N62C itself. The kinetic parameters deteriorate slightly with an n-decyl group (N62C-S-g). The largest catalytic improvement was observed for the cyclohexylmethyl modified CMM, N62C-S-h, for which the $k_{\text{cat}}/K_M$ was almost 7-fold higher than that of its N62C parent. Benzylation was also highly beneficial, with $k_{\text{cat}}/K_M$ for N62C-S-i being 4.6-fold elevated.
The overall effects of chemical modification are more effectively displayed by the $k_{cat}/K_M$ ratios shown in Figure 6.2. This plot clearly reveals the divergent effects of the various chemical modifications, firstly to ameliorate the loss of activity engendered by the initial introduction of cysteine by site directed mutagenesis ($R = a-c$), and then to surpass ($R = d-i$) the WT levels by up to 3.2-fold ($R = h$). The N62C-modified enzymes are consistently activated with large, hydrophobic, alkyl groups conferring the most favourable effects. It is remarkable that of eleven CMMs evaluated, seven of them not only match the wild-type $k_{cat}/K_M$ level, but in fact exceed it by factors of up to $>3$. The deleterious, relative to WT, effects of both negatively and positively charged modifications ($j, k$) are also evident from this plot. It is interesting to note that chemical modification effects can mimic those achieved by traditional protein engineering. For
example, the modification of SBL-N62C with the polar side chain j, mirror those induced by introduction of amino acids with the charged (aspartic acid) amino acid side chains by site-directed mutagenesis at position 62, in the closely related enzyme subtilisin BPN'.

To our knowledge, the attainment of these levels of augmented-WT activity for such a broad range of modified enzymes is unprecedented, and has not been matched so far in its breadth by protein engineering methods alone. These results demonstrate the considerable potential that the combined site-directed mutagenesis-chemical modification approach offers for creating novel enzymes with better-than-WT activity.
Experimental

**Enzyme Purification:** WT-SBL and its N62C mutant were prepared as previously described \(^7\) and purified by desalting on a Sephadex G-25 matrix followed by strong cation exchange chromatography on SP Sepharose FF. The purity of WT- and N62C-SBL, denatured by incubation with 0.1 M HCl at 0 °C for 30 min., was determined by SDS-PAGE and in each case, only one band was visible.

**Preparation of Methanethiosulfonate reagents:** Reagent 1a was purchased from Aldrich Chemical Co. Inc., 1j and 1k from Toronto Research Chemical (2 Brisbane Rd. Toronto, ON), and all were used as received. Reagents 1b-i were prepared as previously described.\(^8\)

**Site-specific Chemical Modification:** To N62C (25 mg) in CHES buffer (2.5 mL; 70 mM CHES, 5 mM MES, 2 mM CaCl\(_2\), pH 9.5) at 20 °C was added a 100-fold molar excess of each of the methanethiosulfonate reagents 1a-k (100 μL of a 1.0 M solution, 100 μM: 1a-d in EtOH, 1e-i in CH\(_3\)CN, 1j in CHES buffer, 1k in MeOH) in a PEG (10,000) coated polypropylene test tube, and the mixture agitated in an end-over-end rotator. Blank reactions were run in parallel. Each of the modification reactions was monitored by specific activity measurements on a Perkin Elmer Lambda 2 spectrophotometer (ε\(_{410}\) = 8800 M\(^{-1}\) cm\(^{-1}\)).\(^9\) The specific activity of the CMM was determined in buffer containing: 0.1 M TRIS pH 8.6, 0.005 % Tween 80, and 1% DMSO, with the suc-AAPF-pNA substrate (1 mg/mL) at 25 °C. The reaction was terminated when the addition of a further 100 μL of methanethiosulfonate solution effected no further change in specific activity, generally within 30 min. The reaction was quenched by dilution in MES buffer (5 mM MES, 2 mM CaCl\(_2\), pH 6.5) at 0 °C and then purified on a disposable desalting column (Pharmacia Biotech PD-10, Sephadex G-25 M) pre-equilibrated with MES buffer. The CMM was eluted with MES-buffer (3.5 mL), dialyzed against 1 mM CaCl\(_2\) (3 x 1 L) at 4 °C and subsequently lyophilized. Modified enzymes were analyzed by native gradient PAGE (8-25%) at pH 4.2, run
towards the cathode on the Pharmacia Phast-System, and appeared as one single band. It was also established that SBL-WT was unchanged on exposure to 1a-k under the modifying conditions.\textsuperscript{10}


\textbf{Free Thiol Titration:} The residual free thiol content of N62C CMMs, was determined spectrophotometrically by titration with Ellman’s reagent ($\varepsilon_{412} = 13600$ M$^{-1}$ cm$^{-1}$)\textsuperscript{11} in phosphate buffer 0.25 M, pH 8.0 and was less than 2 % in all cases.

\textbf{Active Site Titration:} The active enzyme concentration was determined as previously described\textsuperscript{4} by monitoring fluoride ion release upon enzyme reaction with the irreversible inhibitor, phenylmethylsulfonyl fluoride (Aldrich Chemical Co. Inc.) as measured by a fluoride ion sensitive electrode (Orion Research 96-09). The active enzyme concentrations determined in this way was used to calculate kinetic parameters for each CMM.
References


2. Knapp M, Daubermann J, Bott RR. Brookhaven Database Entry 1JEA.


Chemical Modifications at a Single Site Can Induce Significant Shifts in the pH Profiles of a Serine Protease

**Introduction**

The ability to vary the pH-activity profile of an enzyme in a controlled manner has been a long sought-after goal.\(^1\) Such tailoring can help to elucidate important insights into mechanism and permits optimization of enzyme performance for use in organic synthesis.\(^10,11\) Thus far, the most successful approaches to altering pH-activity profiles have used either site-directed mutagenesis\(^2\) \(^6,8\) or chemical modification\(^1,7\) to alter enzyme surface charge. These studies have established that increasing positive surface charge of the serine proteases destabilizes the protonated imidazolium group of the active site histidine and thus lowers its pK\(_a\) while increasing negative surface charge stabilizes the histidine imidazolium, thereby raising its pK\(_a\).\(^1\)\(^8\) While predictive models to determine the effect of electrostatic modification on His64 pK\(_a\) have recently been established,\(^8\) the introduction of uncharged hydrophobic mutations to alter enzyme pH-activity profiles has not been investigated. In addition, while it has been recognized that the pK\(_a\)s of amino acid side chains in proteins, particularly those essential for catalysis, are affected by their environment,\(^12\) these values cannot always be predicted.

The strategy of chemical modification of site-directed mutant enzymes offers a new opportunity for controlling pH optima. Previously, the strategy illustrated in Scheme 7.1 was exploited to increase the activity of SBL effecting an up to 3-fold higher than wild type (WT) activity enhancement for N62C-SCH\(_2\)-c-C\(_6\)H\(_{11}\) (-h) (Table 6.1, page 144).\(^13\) Furthermore, it was apparent that increased alkyl side chain length caused a monotonic increase in \(k_{cat}/K_M\). Due to the proximity of Cys62 to His64 it is hypothesized that these activity enhancements were due in part to the altered pK\(_a\) of the catalytic triad residue His64. The
current study was undertaken to investigate the validity of this hypothesis and to uncover altered pH-activity profiles of SBL which can expand its synthetic utility.

**Results and Discussion**

Initially, molecular mechanics calculations were performed to probe the molecular basis for the 3-fold activity enhancement previously observed for N62C-S-CH₂-c-C₆H₁₁ (-h) (Table 6.1, page 144). These results, which are depicted in Figure 7.1a, revealed that in the energy minimized structure the cyclohexyl moiety is positioned directly over the His64 imidazole of the catalytic triad. In contrast, for N62C-k, for which a 1.7-fold decrease in activity was observed the -S-CH₂CH₂NH₃⁺ side chain is solvated and oriented away from the catalytic triad region in the energy minimized structure, as shown in Figure 7.1b. In light of these molecular modelling results, and considering the proximity of Cys62 to His64, we postulate that the more hydrophobic environment of the catalytic triad in N62C-S-CH₂-c-C₆H₁₁ induced by the cyclohexyl group would cause a decrease in the observed pKₐ of His64. Accordingly, pH-activity profiles for WT-SBL, the parent cysteine mutant N62C, and each of the eleven CMMs N62C-a to -k were evaluated.
Figure 7.1 Molecular Modelling for N62C CMMs: (a) The active site of the N62C-S-h CMM showing the position of the CH₂-cyclohexyl side chain (in gray) over the top of the His64 imidazole (heavy black). (b) Active site of the N62C-S-k CMM showing the position of the ethylamino side chain (in gray) for which the ammonium group points toward the aqueous solvent. (c) Active site of the N62C-S-g CMM showing the disulfide-linked decyl side chain (in gray) which is coiled over the top of the imidazole side chain of His64 (heavy black).
Gratifyingly, a decrease in pKₐ for N62C-S-CH₂-c-C₆H₁₁ was observed, as shown in the representative pH-activity profile (Figure 7.2), indicating that the His64 pKₐ is shifted toward the more acidic side by 0.56 units. This pKₐ change turned out to be general and other, very significant, pKₐ changes were also observed for N62C-S-a to -k (Table 7.1). The most marked pKₐ shifts reported in Table 7.1 are for the -S-CH₂-c-C₆H₁₁ (N62C-S-h), -S-CH₂C₆H₅ (N62C-S-i) and -S-CH₂(CH₂)₆CH₃ (N62C-S-g) CMMs. In fact, the ΔpKₐ of 0.72 observed for N62C-S-CH₂(CH₂)₈CH₃ (-g) is unprecedented for a single uncharged modification.
Table 7.1: Kinetic and pK$_a$ data for N62C CMMs.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>pK$_a$ (a)</th>
<th>(k$_{cat}$/K$<em>M$)$</em>{max}$ (b) s$^{-1}$ mmol$^{-1}$</th>
<th>log $P_{calc}$ (c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>7.01 ± 0.02</td>
<td>157 ± 2</td>
<td>-1.23 ± 0.22</td>
</tr>
<tr>
<td>N62C</td>
<td>6.80 ± 0.03</td>
<td>91 ± 2</td>
<td>1.25 ± 0.24</td>
</tr>
<tr>
<td>N62C-S-a</td>
<td>6.70 ± 0.04</td>
<td>105 ± 2</td>
<td>1.77 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-b</td>
<td>6.79 ± 0.06</td>
<td>200 ± 5</td>
<td>2.30 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-c</td>
<td>6.52 ± 0.04</td>
<td>196 ± 3</td>
<td>3.18 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-d</td>
<td>6.59 ± 0.07</td>
<td>228 ± 6</td>
<td>3.53 ± 0.29</td>
</tr>
<tr>
<td>N62C-S-e</td>
<td>6.61 ± 0.06</td>
<td>271 ± 6</td>
<td>3.90 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-f</td>
<td>6.53 ± 0.07</td>
<td>228 ± 7</td>
<td>4.42 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-g</td>
<td>6.29 ± 0.04</td>
<td>256 ± 7</td>
<td>6.55 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-h</td>
<td>6.45 ± 0.05</td>
<td>238 ± 5</td>
<td>4.36 ± 0.28</td>
</tr>
<tr>
<td>N62C-S-i</td>
<td>6.56 ± 0.08</td>
<td>344 ± 15</td>
<td>3.55 ± 0.31</td>
</tr>
<tr>
<td>N62C-S-j</td>
<td>7.03 ± 0.05</td>
<td>125 ± 3</td>
<td>N. A.</td>
</tr>
<tr>
<td>N62C-S-k</td>
<td>6.59 ± 0.05</td>
<td>111 ± 2</td>
<td>N. A.</td>
</tr>
</tbody>
</table>

(a) Determined in duplicate by the method of initial rates at low substrate concentration,$^{15}$ with succinyl-AAPF-pNA as the substrate, in 0.02 M ethylenediamine buffer, ionic strength 0.05 M adjusted with KCl at 25 °C. (b) (k$_{cat}$/K$_M$)$_{max}$ represents the maximal rate at the pH optimum and was calculated from pH-activity profiles. (c) Calculated using ACD/LogP.$^{16}$

To evaluate the stabilities of WT-SBL and the CMMs, and to ensure that the enzymes were not irreversibly altered under the pH range of this investigation, each enzyme was incubated in the buffers used in this evaluation and then assayed under the standard conditions. All enzymes were stable under these conditions and their activities were not irreversibly altered. Since, the observed pH dependence of k$_{cat}$/K$_M$ follows the ionization of the free enzyme or free substrate$^{17}$ unless there is a pH-dependent change of rate-determining step, and since the suc-AAPF-pNA substrate used in the present study does not have any groups which ionize within the range of pHs investigated, the pK$_a$ observed in the pH-activity profiles is that of a catalytically important residue in the free
enzyme. Previous pH-activity profiles of subtilisins have also indicated an apparent pKₐ of ~ 7, 24.5.8 which was ascribed by 'H NMR studies to the ionization of His64.18

Furthermore, there appears to be a correlation between the magnitude of the change in pKₐ [ΔpKₐ(WT-CMM)] and increasing alkyl side chain length. This is illustrated in Figure 7.3 by comparison of the monotonic increase of ΔpKₐ for the series N62C-S-CH₃, N62C-S-CH₂CH₃, N62C-S-CH₂(CH₂)₃CH₃, N62C-S-CH₂(CH₂)₄CH₃ and N62C-S-CH₂(CH₂)₈CH₃. While, a Brønsted plot19 of pKₐ versus log(kₐ/K Murraymax) s⁻¹ mM⁻¹ for a serine protease has not been reported, since to a first approximation it appeared that increased activities for these CMMs corresponded to decreased His64 pKₐ values, such a plot was constructed to see if it would be informative. However, no correlation was evident despite the fact that general acid-base catalysis is exploited by the serine proteases.20,21 This result is consistent with the overall catalytic rate being affected by factors other than His64 acidity, such as local conformational changes, or electrostatic, hydrophobic and steric interactions.

Figure 7.3 ΔpKₐₐ of N62C CMMs
That the decyl group was the largest hydrophobic group of the current series prompted us to ascertain if there was a linear relationship between pKₐ and hydrophobicity. The Figure 7.4 plot of pKₐ vs the calculated hydrophobicity index log \( P_{\text{calculated}} \)\(^{16} \) where log \( P = \log \text{(solubility in n-octanol/solubility in H}_2\text{O)} \), confirms this relationship and validates the hypothesis that the hydrophobicity of the environment around His64 modulates its pKₐ. That such dramatic pKₐ shifts are induced by the single, uncharged, groups of the current study is truly remarkable, since shifts of these magnitudes have previously been achievable only by multiple electrostatic modifications of serine proteases. For example, succinylation of 14 lysine side chains of chymotrypsin to decrease the positive charge on the enzyme surface by 28 units was required to effect a 1.0 unit increase in the pKₐ of the active site histidine.\(^1 \) On the other hand, increasing chymotrypsin's positive surface charge by the reaction of 13 surface carboxylates with ethylenediamine was needed to elicit a 0.9 unit pKₐ decrease.\(^1 \) Also, by the site-directed mutagenesis approach, a 4 unit increase of net positive surface charge of the related protease, subtilisin BPN', was needed to induce a 1.0 unit His64 pKₐ decrease.\(^3 \) The slight increase in pKₐ effected in this study by the introduction of a negative charge, as for N62C-S-CH₂CH₂SO₃⁻ (N62C-S-j), and the decrease effected by the addition of a positive surface charge, as for N62C-S-CH₂CH₃NH₃⁺ (N62C-S-k), are consistent with previously established models of the effects of altered surface charges.
All of the above results support the hypothesis that the lower dielectric
constant environment induced by hydrophobic modifications stabilize the
unprotonated form of His64, and thus renders its conjugate acid a more effective
general acid.\textsuperscript{22} Since the decyl modified enzyme N62C-S-g caused the greatest
pK$_a$ change, the modified crystal structure of this enzyme was subjected to
molecular modelling. As shown in Figure 7.1c, the modeled decyl moiety adopts
a coiled structure such that it is positioned over the His64 imidazole in a manner
similar to that of the cyclohexyl modified N62C-S-h, apparently generating an
even more hydrophobic environment. From these additional data, the basis of
the rate enhancements observed previously for CMMs N62C- S-b to -i (Table
6.1, page 144),\textsuperscript{13} which reflect the monotonic ΔpK$_a$ changes, can now be
attributed, at least in part, to shifts in their pH-activity profiles and in particular to
lowered His64 pK$_a$'s.

This is the first report of significant changes in the pH-activity profile of a
serine protease induced by chemical modifications or mutations which do not
alter enzyme surface charge. It is clear that inducing pK$_a$ changes by controlled
hydrophobic modifications represents a new strategy with significant potential.
Furthermore, the CMM approach described should be generally applicable to tailoring the pH-activity profiles of other synthetically useful hydrolases.
Experimental

Kinetics: $k_{cat}/K_M$ values were determined in duplicate in 0.02 M ethylenediamine buffer, ionic strength 0.05 M adjusted with KCl, $[S] = 2.083 - 12.5 \times 10^{-5}$ M of succinyl-AAPF-pNA ($\varepsilon_{410} = 8800$ cm$^{-1}$ M$^{-1}$).$^{23}$ $[E] = 1 - 50 \times 10^{-7}$ M at 25 °C. pK$_a$'s were calculated using GraFit version 3.0 curve fit, single pK$_a$ or bell-shaped double pK$_a$, with the minimum set to zero. Specifically, into a 1.5 mL polystyrene cuvette was added 980 μL of buffer (0.02 M ethylenediamine) and 10 μL of substrate (2.083 -12.5 $\times 10^{-3}$ M in DMSO). The solution was incubated in a cell holder at 25 °C, before the absorbance reading was set to zero. Then 10 μL of enzyme solution (1 x 10$^{-5}$ to 5 x 10$^{-6}$ M) was added to initiate the reaction. After an 8 sec delay, absorbance versus time measurements were recorded on a Perkin Elmer lambda 2 spectrophotometer. $k_{cat}/K_M$ values were calculated employing the low substrate approximation, where the Michaelis-Menten Equation reduces to: $^{15,24}$ $v = k_{cat}/K_M [E]_0[S]$ when $[S]<K_M$.

Evaluation of Enzyme Stability in pH Range: To ensure that the enzymes were not irreversibly altered by pH changes, the CMMs were incubated in ethylenediamine buffers at the same pH as the assay conditions at room temperature (20 °C) for 1 min. These were then diluted (100-fold) into 0.1 M Tris, pH 8.6 containing 0.005% Tween 80 and enzyme specific activity was determined. All CMMs were found to be stable under these conditions.

Calculation of Log P: Calculated using ACD/Log P$^{16}$ based on side chain structure of the amino acid or modified amino acid after replacing the α-carbon atom with a hydrogen atom. For example, the model structure for WT (Asn62) was CH$_3$C(O)NH$_2$ whereas model structure for N62C-S-b is CH$_3$-S-S-CH$_2$CH$_3$. N62C-S-j and -k were omitted since the software package is unable to calculate log P for these enzymes. The plot of pK$_a$ versus log P gave a slope of -0.093 ± 0.009, an intercept of 6.91 ± 0.03 and a correlation coefficient of -0.96.

Molecular Modelling: The X-ray structure of subtilisin Bacillus lentus (Brookhaven PDB entry 1JEA)$^{25}$ was used as the starting point for calculations.
on the wild type and chemically modified mutant enzymes. The enzyme setup was performed with Insight II, version 2.3.0. To create initial coordinates for the minimization, hydrogens were added at the pH used for kinetic measurements. This protonated all Lys and Arg residues and the N-terminus and deprotonated all Glu and Asp residues and the C-terminus. The model system with the phenylmethylsulfonate ester of Ser221 bound in the S1 pocket was solvated with a 5 Å layer of water molecules giving a total number of water molecules of 1174 in this system. The overall charge of the enzyme-inhibitor complex resulting from this setup was +3 for the WT enzyme. Energy simulations were performed with the DISCOVER program, Version 2.9.5 on a Silicon Graphics Iris Indigo computer, using the consistent valence force field function (CVFF). A non-bonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. The energy of the structure of the WT enzyme was minimized in stages, with initially only the water molecules being allowed to move, then the water molecules and the amino acid side chains, and then the entire enzyme. The mutated and chemically modified enzymes were generated using the Builder module of Insight. These structures were then energy minimized in a similar manner. Specifically, the side-chain of the mutated residue and the water molecules were energy minimized. The amino acid side chains within a 10 Å radius of the α-carbon of the mutated residue were energy minimized while all other enzyme residues were constrained, then all of the atoms were energy minimized.
References


16. ACD/LogP Advanced Chemistry Development, Version 0.9 Beta.


25. Knapp M, Daubermann J, Bott RR. Brookhaven Database Entry 1JE.A.


Perspective
The goals of this study were to contribute to the understanding of the factors that control enzyme specificity, to develop a method to overcome the 20-amino acid limitation of conventional site-directed mutagenesis, and to create novel enzyme specificities to further expand the synthetic applicability of the serine proteases. The strategy of chemical modification and site directed mutagenesis of subtilisin *Bacillus lentus* was shown to be a rapid and convenient one in this regard. The choice of residues S156, S166 and N62 for mutagenesis to cysteine and subsequent modification proved to be appropriate, and effected significant changes in the activity, specificity, binding properties, and pH-activity profiles of SBL. The most dramatic changes in the catalytic properties of SBL-CMMs were rationalizable by molecular modelling analysis.

Subtilisin has proven to be an ideal template for this approach. However, the CMM approach is not necessarily limited to cysteine-free enzymes since cysteine residues are usually buried in proteins¹ and therefore will not react readily with the MTS modifying agents used. Thus, introduced cysteine residues on the active site surface can be selectively modified in the presence of buried ones. Alternatively, site-directed mutagenesis could be used to remove potentially reactive natural cysteine residues in order to ensure the desired regio-selectivity of modification by the methanethiosulfonate reagents.

This method is amenable to a combinatorial² approach to generating libraries of CMM enzymes, which may be subjected to selection screens to identify a desirable property, such as improved esterase-to-amidase selectivity.³ In addition, the CMM approach can address the challenges of generating precisely glycosylated proteins *in vitro*.⁴ A further important application of CMMs is for novel organic synthesis. The evaluation of the preparative viability of the CMMs generated in this study has already begun and the preliminary results are very positive.⁵

A further benefit of the current approach may be to provide a rapid and convenient screen identifying the most desirable types of amino acid
replacements for a given enzyme location by traditional protein engineering methodology.
References


Appendix
Representative Calculation of $k_{\text{cat}}$ and $K_M$

Table 1 $k_{\text{cat}}$, $K_M$ for S166C-S-CH$_2$-c-C$_6$H$_{11}$ (pH 7.5)

<table>
<thead>
<tr>
<th>[S] mM</th>
<th>Absorbance at 5% Conversion</th>
<th>Slope, A/s to 5% Conversion</th>
<th>Initial Velocity $v_i$ M s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125</td>
<td>0.056</td>
<td>0.002349</td>
<td>$2.269 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.25</td>
<td>0.111</td>
<td>0.004794</td>
<td>$5.448 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.375</td>
<td>0.165</td>
<td>0.005917</td>
<td>$6.723 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.5</td>
<td>0.220</td>
<td>0.007453</td>
<td>$8.469 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.75</td>
<td>0.330</td>
<td>0.008605</td>
<td>$9.778 \times 10^{-7}$</td>
</tr>
<tr>
<td>1.0</td>
<td>0.440</td>
<td>0.011161</td>
<td>$1.268 \times 10^{-7}$</td>
</tr>
<tr>
<td>1.5</td>
<td>0.660</td>
<td>0.011994</td>
<td>$1.363 \times 10^{-7}$</td>
</tr>
<tr>
<td>2.0</td>
<td>0.880</td>
<td>0.012746</td>
<td>$1.448 \times 10^{-7}$</td>
</tr>
<tr>
<td>3.0</td>
<td>1.320</td>
<td>0.013146</td>
<td>$1.498 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

where, $v_i = \frac{(\text{slope A s}^{-1})}{(8800 \text{ A M}^{-1})}$

$k_{\text{cat}}$ and $K_M$ calculated using GraFit 3.0 Enzyme Kinetics Simple Weighting

![Graph showing enzymatic activity vs substrate concentration](image)

$[E] = 2.14 \times 10^{-7}$ M

$k_{\text{cat}} = 8.7 \pm 0.2$ s$^{-1}$

$k_M = 0.61 \pm 0.04$ mM
Representative Calculation of $K_i$

### Table 2 $K_i$ for WT-SBL and 3-aminophenyl boronic acid

<table>
<thead>
<tr>
<th>% Conversion</th>
<th>$A_{410\text{ nm}}$</th>
<th>$t$ sec</th>
<th>$t_c$ sec</th>
<th>$\ln(S_o/S)$</th>
<th>$t-t_c$ sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.15</td>
<td>0.331</td>
<td>42.4</td>
<td>9</td>
<td>0.1625</td>
<td>33.4</td>
</tr>
<tr>
<td>0.18</td>
<td>0.397</td>
<td>56.4</td>
<td>15</td>
<td>0.1985</td>
<td>41.4</td>
</tr>
<tr>
<td>0.21</td>
<td>0.463</td>
<td>71.2</td>
<td>21.4</td>
<td>0.2357</td>
<td>49.8</td>
</tr>
<tr>
<td>0.24</td>
<td>0.528</td>
<td>86.4</td>
<td>28.1</td>
<td>0.2744</td>
<td>58.3</td>
</tr>
<tr>
<td>0.27</td>
<td>0.595</td>
<td>102.2</td>
<td>34.8</td>
<td>0.3147</td>
<td>67.4</td>
</tr>
<tr>
<td>0.30</td>
<td>0.661</td>
<td>119</td>
<td>42.2</td>
<td>0.3567</td>
<td>76.8</td>
</tr>
<tr>
<td>0.33</td>
<td>0.727</td>
<td>136.4</td>
<td>49.6</td>
<td>0.4005</td>
<td>86.8</td>
</tr>
<tr>
<td>0.36</td>
<td>0.793</td>
<td>155.2</td>
<td>58</td>
<td>0.4463</td>
<td>97.2</td>
</tr>
<tr>
<td>0.39</td>
<td>0.859</td>
<td>174.8</td>
<td>62.2</td>
<td>0.4943</td>
<td>108.6</td>
</tr>
</tbody>
</table>

where,

$A_{410\text{ nm}} = (A)(% \text{ Conversion})$

Beer's Law $A = \varepsilon c \alpha$

$= (8800 \text{ M}^{-1} \text{ cm}^{-1}) (2.5 \times 10^{-3} \text{ M})(1 \text{ cm})$

$= 2.2$

$t =$ time for inhibited enzyme reaction to reach absorbance indicated

$t_c =$ time for the uninhibited enzyme reaction to reach absorbance

$\ln(S_o/S) = \ln(100/85)=0.1625$

$S_o =$ 100% [substrate]

$S =$ [substrate] at % conversion

*GraFit* 3.0 is used to plot $\ln(S_o/S)$ against $t-t_c$. Simple weighting linear regression is applied to slope and error.

$$
[E] = 3.2 \times 10^{-6} \quad [l] = 0.001913 \text{ M} \\
K_{cat} = 153 \pm 4 \text{ s}^{-1} \quad K_M = 0.00073 \pm 0.00005 \text{ M} \\
\text{slope} = 225.0 \pm 0.7 \text{ s}^{-1}
$$

$$
K_i = \frac{K_M \cdot [l]}{V_{max} \cdot \text{slope}} \pm K_i \sqrt{(\Delta K_m/K_m)^2 + (\Delta V_{max}/V_{max})^2 + (\Delta \text{slope}/\text{slope})^2}
$$

$$
K_i = \frac{7.3 \times 10^{-4} \text{ M} \cdot 1.9 \times 10^{-3} \text{ M}}{4.89 \times 10^{-6} \text{ M s}^{-1} \cdot 225.0 \text{ s}} \pm K_i \sqrt{(5E-5/7.3E-4)^2 + (2E-7/4.89E-6)^2 + (0.7/225)^2}
$$

$$
K_i = 1.3 \pm 0.1 \text{ mM}
$$