RECOMBINANT EXPRESSION OF *STREPTOMYCES GRIZEUS* GENES AND THEIR VARIANTS IN *BACILLUS SUBTILIS* AND *ESCHERICHIA COLI*

by

Debbie Carson

B.Sc. (hon.), Simon Fraser University, 1989

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

in the Department
of
Chemistry

© Debbie Carson 1994

SIMON FRASER UNIVERSITY

April 1994

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
Approval

Name: Debbie Carson
Degree: Master of Science
Title of thesis: RECOMBINANT EXPRESSION OF STREPTOMYCES GRISEUS GENES AND THEIR VARIANTS IN BACILLUS SUBTILIS AND ESCHERICHIA COLI

Examinining Committee:

Chair: Dr. F. Einstein

Dr. T. J. Borgford
Senior Supervisor

Dr. W. R. Richards
Professor
Department of Chemistry

Dr. K. Slessor
Professor
Department of Chemistry

Dr. R. B. Cornell
External Examiner
Assistant Professor
Department of Chemistry
Simon Fraser University
PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

RECOMBINANT EXPRESSION OF BACTERIAL
GRISSEUS CYGNS AND THEIR VARIANTS IN BACILLUS SUBTILIS AND E. COLI

Author:

Debra Carson

(signature)

(name)

May 5, 1999

(date)
Abstract

The impetus for this study was a desire to produce inactive variants of *Streptomyces griseus* protease B (SGPB). Studies involving other catalytically altered serine proteases indicated a potential for a proteolytically inactive SGPB to be used in protein semi-synthesis and to be of value due to its specificity and its stability under various conditions. *B. subtilis* was initially chosen as the host for recombinant expression of this *S. griseus* gene due to the similarities in gene structure between SGPB and the extracellular *Bacillus* protease subtilisin and the ability of *Bacillus* to secrete large quantities of protein.

The gene for SGPB had previously been found to code for a signal peptide and pro-peptide joined to the amino terminus of the mature protease. This suggested that SGPB (and the related protease, SGPA) is initially secreted as a precursor which is subsequently self-processed to remove the pro-region in a mechanism common to other bacterial extracellular proteases. Attempts at expressing SGPB without its pro-region in either *Bacillus* or *E. coli* or by substituting the pro-region of *Bacillus amyloliquefaciens* subtilisin BPN' in *Bacillus* did not result in any SGPB protease activity. Properly folded active enzyme was produced only when the protease-specific pro-region was present, although the pro-region is cleaved and does not form part of the active protease.

Expression of inactive variants of SGPB in either *Bacillus* or *E. coli* did not result in overproduction of the mutant protein. Although transcribed in *Bacillus*, the inactive SGPB precursor was not detectable and substitution of the active site serine with glycine resulted in specific targeting for degradation in *E. coli* due to a possible decrease in protease stability. Further study is necessary to determine the fate of the inactive precursor in *Bacillus* and ultimately overexpress the inactive SGPB variant.
Table of Contents

Approval........................................................................................................... ii
Abstract.......................................................................................................... iii
List of Tables .................................................................................................. x
List of Figures ............................................................................................... xi
List of Abbreviations.................................................................................. xiii
Acknowledgments....................................................................................... xiv
Dedication ..................................................................................................... xv
1. Introduction............................................................................................... 1
   1.1. Proteases .............................................................................................. 1
       1.1.1. Serine Proteases ........................................................................ 2
       1.1.2. Cysteine Proteases ................................................................... 6
       1.1.3. Aspartic Proteases ..................................................................... 7
       1.1.4. Metalloproteases ...................................................................... 8
   1.2. Recombinant DNA Technology .......................................................... 9
       1.2.1. Expression .................................................................................. 9
   1.3. Secreted Proteases ............................................................................. 11
   1.4. *Streptomyces griseus* Proteases ........................................................ 15
   1.5. Protein Synthesis ................................................................................ 17
       1.5.1. Chemical Protein Synthesis ....................................................... 17
       1.5.2. Protein Semi-synthesis ............................................................... 18
   1.6. Summary ............................................................................................. 19
2. Materials and Methods ........................................................................... 20
   2.1. Buffers and Other Solutions ............................................................... 20
       2.1.1. Tris Buffers ............................................................................... 20
2.1.2. Transformation Solutions ...................................................... 20
2.1.3. DNA Preparation ............................................................... 20
2.1.4. DNA Manipulation .............................................................. 21
2.1.5. RNA Preparation and Manipulation ...................................... 21
2.1.6. Protein Manipulation .......................................................... 22

2.2. Bacteriological ......................................................................... 23
2.2.1. Media ................................................................................ 23
2.2.2. Bacterial Strains ................................................................ 24
2.2.3. Preparation of Competent E. coli Cells .................................. 25
2.2.4. Transformation of E. coli ..................................................... 26
   2.2.4.1. Transformation of E. coli DH5α™, TB1, and K12c1 .......... 26
   2.2.4.2. Transformation of E. coli QY13 with pLcII-FX constructs ... 26
   2.2.4.3. Transformation of E. coli TG2 by Bacteriophage M13 ...... 27
2.2.5. Preparation of Bacillus Protoplasts ....................................... 27
2.2.6. Transformation of Bacillus Protoplasts .................................. 28
2.2.7. Cryopreservation of Bacterial Strains .................................... 28

2.3. Nucleic Acid Manipulation ....................................................... 29
2.3.1. Plasmids and Bacteriophage Vectors ..................................... 29
   2.3.1.1. E. coli ........................................................................ 29
   2.3.1.2. Bacillus .................................................................... 30
2.3.2. DNA Isolation .................................................................... 30
   2.3.2.1. Preparation of Single-stranded Bacteriophage M13 DNA . 30
   2.3.2.2. Preparation of Double-stranded E. coli Plasmid DNA ...... 31
   2.3.2.3. Preparation of Replicative Form M13 .......................... 32
   2.3.2.4. Preparation of Bacillus Plasmid DNA .......................... 32
2.3.3. Cloning .............................................................................. 33
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.3.1. Restriction Enzyme Digestion</td>
<td>33</td>
</tr>
<tr>
<td>2.3.3.2. Blunt Ending</td>
<td>33</td>
</tr>
<tr>
<td>2.3.3.2.1. T4 DNA Polymerase</td>
<td>33</td>
</tr>
<tr>
<td>2.3.3.2.2. Mung Bean Nuclease</td>
<td>34</td>
</tr>
<tr>
<td>2.3.3.3. Dephosphorylation of Vector</td>
<td>34</td>
</tr>
<tr>
<td>2.3.3.4. Gel Purification of DNA Fragments</td>
<td>35</td>
</tr>
<tr>
<td>2.3.3.4.1. Low Melting Point Agarose</td>
<td>35</td>
</tr>
<tr>
<td>2.3.3.4.2. Electroelution</td>
<td>35</td>
</tr>
<tr>
<td>2.3.3.4.3. Spin Column Purification</td>
<td>36</td>
</tr>
<tr>
<td>2.3.3.5. Ligation</td>
<td>37</td>
</tr>
<tr>
<td>2.3.3.6. DNA Sequencing</td>
<td>37</td>
</tr>
<tr>
<td>2.3.3.6.1. E. coli Plasmid DNA Preparation</td>
<td>37</td>
</tr>
<tr>
<td>2.3.3.6.2. Annealing Template and Primer</td>
<td>38</td>
</tr>
<tr>
<td>2.3.3.6.3. Labeling and Termination Reactions</td>
<td>38</td>
</tr>
<tr>
<td>2.3.4. Site-directed Mutagenesis</td>
<td>38</td>
</tr>
<tr>
<td>2.3.4.1. Preparation of Single-stranded Uracil-containing Bacteriophage M13 DNA</td>
<td>38</td>
</tr>
<tr>
<td>2.3.4.2. Kinasing of Oligonucleotides</td>
<td>40</td>
</tr>
<tr>
<td>2.3.4.3. Site-specific Mutagenesis</td>
<td>40</td>
</tr>
<tr>
<td>2.3.4.4. Transformation of E. coli DH5α™</td>
<td>41</td>
</tr>
<tr>
<td>2.3.5. Northern Blots</td>
<td>41</td>
</tr>
<tr>
<td>2.3.5.1. Isolation of RNA from Bacillus</td>
<td>41</td>
</tr>
<tr>
<td>2.3.5.2. Denaturing Formaldehyde Agarose Gel</td>
<td>42</td>
</tr>
<tr>
<td>2.3.5.3. Transfer of RNA</td>
<td>43</td>
</tr>
<tr>
<td>2.3.5.4. Preparation of Radioactive DNA Probe</td>
<td>44</td>
</tr>
<tr>
<td>2.3.5.5. Hybridization</td>
<td>45</td>
</tr>
<tr>
<td>2.4. Protein Expression</td>
<td>45</td>
</tr>
</tbody>
</table>
2.4.1. Milk Plates ........................................................................45
2.4.2. Activity Assay ..................................................................45
2.4.3. *Bacillus* Expression and Fractionation ..............................46
2.4.4. *E. coli* Expression and Fractionation ...............................47
   2.4.4.1. *E. coli* (pMAL-p/c) Growth and Induction ...............47
   2.4.4.2. *E. coli* (pLcII-FX) Growth and Induction ...............47
   2.4.4.3. Fractionation ............................................................47
2.4.5. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) ......48
   2.4.3.1. SDS-Polyacrylamide Gel ............................................48
2.4.6. Western Blots ...................................................................49
   2.4.6.1. Transfer ......................................................................49
   2.4.6.2. Immunoassay ............................................................49
3. Results ...................................................................................51
   3.1. Expression Plasmids ..........................................................51
   3.2. Mature SGPB ....................................................................51
   3.3. Pre-Pro-Mature Protease ....................................................55
   3.4. Pro-Mature Protease ..........................................................57
      3.4.1. *Bacillus* Expression ....................................................57
      3.4.2. *E. coli* Expression .......................................................60
         3.4.2.1. pMAL Constructs ...................................................60
         3.4.2.2. pLcII-FX Constructs ................................................62
   3.5. Inactive SGPB .................................................................62
   3.6. *Bacillus* Time Course ......................................................65
      3.6.1. Bacterial Growth .........................................................66
      3.6.2. Protease Activity ........................................................66
         3.6.2.1. Spiking of SGPA with SGPB .................................70
   3.7. SDS-Polyacrylamide Gel Electrophoresis and Western Blotting .73
3.7.1. *Bacillus* Expression ................................................. 73
  3.7.1.1. Inactive SGPB Spiking ....................................... 75
3.7.2. *E. coli* Expression ............................................... 78
  3.7.2.1. pLcII-FX-constructs ...................................... 78
  3.7.2.2. pMAL-constructs ......................................... 81

4. Discussion ..................................................................... 90
  4.1. Expression of a Streptomyces Gene in *Bacillus* ........ 90
  4.2. *B. subtilis* Subtilisin BPN' Promoter Pre-pro-region Constructs .... 91
  4.3. Native *S. griseus* Promoter Pre-region Constructs .... 95
  4.4. *B. subtilis* Subtilisin BPN' Promoter Pre-region Constructs .... 96
  4.5. Directional Dependence of Expression in pUB110 .......... 97
  4.6. Effect of Growth Temperature on Protease Activity .... 98
  4.7. Expression of Inactive SGPB .................................. 99
  4.8. Protein Stability .................................................... 100
    4.8.1. *Bacillus* .................................................... 100
    4.8.2. *E. coli* ..................................................... 101
  4.9. Spiking Experiments ............................................. 102
    4.9.1. SGPA ........................................................... 102
    4.9.2. Inactive SGPB .............................................. 103
  4.10. Intramolecular Chaperone ..................................... 103
  4.11. Conclusions and Further Studies ............................ 104

Appendix: Plasmid Construction ........................................ 106
  A.1. *Bacillus* Expression ........................................... 106
    A.1.1. Pre-pro-subtilisin BPN' .................................... 106
    A.1.2. Pre-pro-mature Protease ................................ 109
      A.1.2.1. PCR .................................................... 109
      SGPA ............................................................ 109
List of Tables

Table 1.1. Different Chemical Types of Proteases................................. 1
Table 3.1. SGPA and SGPB Expression Plasmids. ..................................... 52
Table 3.2. Mutant *E. coli* K12 strains.................................................. 83
Table 4.1. Amino acid sequences surrounding the cleavage sites of the pro-mature bond of pre-pro-subtilisin BPN’ fusion constructs................................. 92
Table 4.2. Amino acid sequences surrounding the cleavage sites at the N-terminus of the mature region of subtilisin BPN’ of the three subtilisin BPN’-containing constructs......................................................... 94
Table 4.3. Amino acid sequences surrounding the cleavage sites of the pre-pro bond of pre-subtilisin BPN’ fusion constructs................................. 96
List of Figures

Figure 1.1. Schematic representation of the binding site of proteolytic enzymes .......... 2
Figure 1.2. Schematic representation of the burst and steady-state phases in the hydrolysis of p-nitrophenyl acetate by chymotrypsin .............................................. 3
Figure 1.3. Mechanism of chymotrypsin cleavage of a peptide bond between phenylalanine and glycine .......................................................... 5
Figure 1.4. Structure of bacterial signal peptide ................................................. 10
Figure 1.5. Pre-pro-mature protease precursor .................................................... 11
Figure 1.6. Comparison of SGPB, α-lytic protease and subtilisin E precursors ...... 16
Figure 2.1. Preparation of the spin column ...................................................... 36
Figure 2.2. Preparation of the transfer apparatus .............................................. 43
Figure 3.1. Comparison of plasmids pMSi, pPT30, pKN1, pDC1, and pGK4a ....... 53
Figure 3.2. B. subtilis DB104 cultures harboring the plasmids pUB110, pPT30, pKN1, pDC1, and pGK4a .......................................................... 54
Figure 3.3. Bacillus plasmid pUB110 .............................................................. 56
Figure 3.4. Comparison of plasmids pPA1, pPA2, pPB1, and pPB2 ................. 56
Figure 3.5. B. subtilis DB104 cultures harboring the plasmids pUB110, pPA1, pPA2, pPB1, and pPB2 .............................................................. 57
Figure 3.6. Comparison of plasmids pbDC12+, pbDC12-, pbDC13+, and pbDC13-................................. 58
Figure 3.7. B. subtilis DB104 cultures harboring the plasmids pUB110, pbDC12+, pbDC12-, pbDC13+, and pbDC13-................................. 58
Figure 3.9. Comparison of plasmids pMAL-pA, and pMAL-pB .......................... 60
Figure 3.10. E. coli TB1 cultures harboring various the plasmids pMAL-p, pMAL-c, pMAL-pA, pMAL-pB, pMAL-cA, and pMAL-cB ......................... 61
Figure 3.11. Plasmid pLcII-FX-pro-SGPB .......................................................... 62
Figure 3.12. Inactive SGPB expression plasmids .............................................. 63
Figure 3.13. B. subtilis DB104 cultures harboring various the plasmids pUB110, pbDC13+, pbDC13-, pEB-B8, pbDC14+, pbDC14- and pEB-HDG ...... 64
Figure 3.14. *E. coli* TB1 cultures harboring the plasmids pMAL-p, pMAL-c, pMAL-pB, pMAL-cB, pMAL-pB/SG195, and pMAL-cB/SG195.

Figure 3.15. Growth curves of strains of *B. subtilis* DB104.

Figure 3.16. Protease activity profiles of strains of *B. subtilis* DB104.

Figure 3.17. Protease activity profiles of strains of *B. subtilis* DB104 after addition of SGPB.

Figure 3.18. Protease activity profiles of strains of *B. subtilis* DB104 after addition of various amounts of SGPB.

Figure 3.19. Coomassie Blue-stained SDS-polyacrylamide gel of fractionated *B. subtilis* culture supernatants.

Figure 3.20. Western blot using mouse polyclonal antibodies on fractionated *B. subtilis* DB104 culture supernatant samples.

Figure 3.21. Coomassie Blue-stained SDS-polyacrylamide gel of samples from a protease spiking experiment.

Figure 3.22. Coomassie Blue-stained SDS-polyacrylamide gel of *E. coli* QY13 cell fractions.

Figure 3.23. Coomassie Blue-stained SDS-polyacrylamide gel of *E. coli* QY13 cell fractions.

Figure 3.24. Coomassie Blue-stained SDS-polyacrylamide gel of *E. coli* TB1 cell fractions.

Figure 3.25. Coomassie Blue-stained SDS-polyacrylamide gel of *E. coli* CAG626 cell fractions.

Figure 3.26. Coomassie Blue-stained SDS-polyacrylamide gel of *E. coli* CAG629 cell fractions.

Figure 3.30. Western blot using rabbit anti-MBP monoclonal antibodies on *E. coli* TB1 cell fractions.

Figure 3.31. Western blot using rabbit anti-MBP monoclonal antibodies on *E. coli* CAG626 cell fractions.

Figure 3.32. Western blot using rabbit anti-MBP monoclonal antibodies on *E. coli* CAG629 cell fractions.

Figure 4.1. *Bacillus* plasmid pUB110.
List of Abbreviations

ccc ............ covalently-closed circular
d dH₂O ............ distilled deionized water
DIPF ............ diisopropylphosphofluoridate
DNA ............ deoxyribonucleic acid
DTT ............ dithiothreitol
EDTA ............ ethylenediaminetetraacetic acid, disodium salt
EGTA............ ethylene glycol-bis(β-aminoethyl ether) N,N',N'-tetraacetic acid
IPTG ............ isopropylthio-β-D-galactoside
MBP ............ maltose binding protein
MOPS .......... 3-[N-morpholino]propanesulfonic acid
PAGE ............ polyacrylamide gel electrophoresis
PAP ............ polyacrylamide carrier
PCR ............ polymerase chain reaction
PEG ............ polyethylene glycol
PMSF .......... phenylmethylsulfonyl fluoride
RNA ............ ribonucleic acid
SDS ............ sodium dodecyl sulfate, sodium lauryl sulfate
SGPA ............ *Streptomyces griseus* protease A
SGPB ............ *S. griseus* protease B
TEMED ............ N,N',N'-tetramethylethlenediamine
TCA .......... trichloroacetic acid
Tris ............ Tris[hydroxymethyl]aminomethane
Tween 20 ....... polyoxyethylenesorbitan monolaurate
X-gal ........... 5-bromo-4-chloro-3-indolyl-β-D-galactoside
Acknowledgments

First, I would like to thank Dr. T. J. Borgford and Dr. Gabe Kalmar for their guidance and encouragement throughout this project, and for the many discussions necessary for its completion.

I would also like to thank my lab-mates, past and present, for their helpful advice and humour.

Finally, I would like to give thanks to my friends; to Dale, for his patience and tolerance of my never-ending questions, and especially to Doug, for his understanding and for making life more joyous.
Dedication

This thesis is dedicated to my parents.
1. Introduction

The impetus for this study was a desire to express inactive variants of *Streptomyces griseus* protease B (SGPB). Studies involving other catalytically altered serine proteases indicated a potential for a proteolytically inactive SGPB to be used in semi-synthetic reactions and to be of value due to its specificity and its stability under various conditions. *B. subtilis* was initially chosen as the host for recombinant expression of this *S. griseus* gene due to the similarities in gene structure between SGPB and the extracellular *Bacillus* protease subtilisin and the ability of *Bacillus* to secrete large quantities of protein. Expression systems for SGPB and the related protease SGPA were developed in both *B. subtilis* and *E. coli*.

1.1. Proteases

The term protease describes a large class of enzymes responsible for the cleavage of peptide bonds in proteins. Proteases are divided into two subclasses; peptidases, consisting of exopeptidases; and proteinases, consisting of proteolytic enzymes and endopeptidases. Of the several hundred proteases characterized from a variety of organisms, only four major catalytic mechanisms have been elucidated. The various proteases can be categorized based on their chemical characteristics which are also indicative of their mechanism of action.¹

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>serine proteases</td>
<td>chymotrypsin</td>
<td>inhibited by organophosphorous compounds</td>
</tr>
<tr>
<td>cysteine proteases</td>
<td>papain</td>
<td>sensitive to oxidation</td>
</tr>
<tr>
<td>aspartic proteases</td>
<td>pepsin</td>
<td>active at low pH</td>
</tr>
<tr>
<td>metalloproteases</td>
<td>thermolysin</td>
<td>sensitive to chelators</td>
</tr>
</tbody>
</table>

Proteases differ in their preference or specificity for the amino acids on either side of the susceptible peptide bond. In the nomenclature of Schechter and Berger² (Figure 1.1.), amino acid residues extending from the scissile bond toward the
Figure 1.1. Schematic representation of the binding site of proteolytic enzymes.

amino terminus of the peptide are denoted as $P_1$, $P_2$, ..., $P_n$ and those extending toward the carboxyl terminus are denoted $P'_1$, $P'_2$, ..., $P'_n$. The corresponding binding sites on the enzyme are $S_1$, $S_2$, ..., $S_n$ and $S'_1$, $S'_2$, ..., $S'_n$. Primary specificity usually lies in the $S_1$ site on the enzyme and in the corresponding $P_1$ amino acid residue of the substrate but secondary specificity can include up to three to four residues on either side of the scissile bond.

1.1.1. Serine Proteases

The most abundant and well studied subclass of proteases is the serine proteases. Examples of this subclass of enzymes include the pancreatic digestive enzymes chymotrypsin and trypsin; enzymes responsible for blood clotting, thrombin and coagulation Factor X; as well as the bacterial enzymes, subtilisin (Bacillus spp.) and $\alpha$-lytic protease (Lysobacter). Within this mechanistic group, two distinct families exist, the chymotrypsin and subtilisin families. Chymotrypsin and subtilisin are unrelated in terms of their amino acid sequences and tertiary structures but are nearly identical in the three-dimensional organization of their active sites. It is thought that each family of serine proteases evolved independently, but arrived at the same catalytic mechanism, i.e., convergent evolution.

Catalysis of peptide bond hydrolysis by serine proteases proceeds in two distinct stages as was shown in kinetic studies of the hydrolysis of $p$-nitrophenyl acetate to $p$-nitrophenol:

$$
\text{O}_2\text{N}-\text{OCCH}_3 + \text{H}_2\text{O} \rightarrow \text{O}_2\text{N}-\text{OH} + \text{CH}_3\text{COOH}
$$
When large amounts of enzyme are used, there is an initial rapid burst of \( p \)-nitrophenol production, followed by its formation at a much slower steady-state rate (Figure 1.2.).

![Figure 1.2. Schematic representation of the burst and steady-state phases in the hydrolysis of \( p \)-nitrophenyl acetate by chymotrypsin.]

The simple kinetic scheme of substrate hydrolysis is shown in equation 1:

\[
E + S \rightleftharpoons E\cdot S \quad \xrightarrow{H_2O} \quad E\cdot P_2 \quad \rightarrow \quad E \quad (1)
\]

where
- \( E \) = free enzyme
- \( S \) = free substrate (amide or ester)
- \( E\cdot S \) = enzyme substrate complex
- \( P_1 \) = amine or alcohol component of substrate
- \( E\cdot P_2 \) = covalent intermediate
- \( P_2 \) = acid component of substrate

In the case of the ester, \( p \)-nitrophenyl acetate, the first step in catalysis is the formation of an enzyme-substrate complex followed by the cleavage of the ester bond releasing \( p \)-nitrophenol as the acyl group of the substrate becomes covalently attached to the enzyme. Water then attacks the acyl-enzyme complex to yield carboxylate ion and regenerate the free enzyme. The initial rapid burst of \( p \)-nitrophenol released is due to formation of the acyl-enzyme complex (acylation) and the slower steady-state production of \( p \)-nitrophenol is due to hydrolysis of the covalent complex to regenerate the free enzyme (deacylation). For this substrate, deacylation is much slower than acylation and determines the overall rate of hydrolysis.
The protease residue involved in acylation was identified by isolation of the acyl-enzyme complex (E–P\(_2\)) which was stable at low pH.\(^4\) The acyl group was bound through the oxygen atom of a single, unusually reactive serine residue (Ser–195 of chymotrypsin, one of 27 serine residues in this protein). The same residue was found to react with organic fluorophosphates such as diisopropylphosphofluoridate (DIPF)\(^5\) and phenylmethylsulfonyl fluoride (PMSF).\(^6\) DIPF and PMSF inhibit serine proteases by forming relatively stable covalent complexes with the enzymes. Involvement of histidine–57 in catalysis was implicated by affinity labeling using a molecule that resembles a substrate such as tosyl-L-phenylalanine chloromethyl ketone (TPCK).\(^7\) TPCK specifically binds to the active site due to its phenylalanine side chain and alkylates one of the ring nitrogens of His–57. The final member of the so-called catalytic triad of serine proteases, aspartate–102, was determined through X-ray crystallography studies.\(^8\)

In the free enzyme, His–57, acting as a general base catalyst in its unprotonated form, accepts a proton from the Ser–195 hydroxyl and the oxygen atom carries out a nucleophilic attack on the carbonyl carbon of the scissile peptide bond (Figure 1.3.-A). It was originally thought that Asp–102 would also become protonated from His–57 in a "charge relay network".\(^9\) Results of neutron diffraction studies of trypsin\(^10\) indicate that His–57 remains protonated and that the negatively charged carboxyl group of Asp-102 orientst His–57 and stabilizes the positively charged imidazole. The resulting oxyanion of the tetrahedral intermediate is stabilized by hydrogen bonding to two main chain NH groups as the proton held by His–57 is donated to the nitrogen atom of the susceptible peptide bond (Figure 1.3.-B). Once this bond is cleaved, the amine component of the substrate (P\(_1\)) diffuses away leaving the acid component esterified to Ser–195 (E–P\(_2\)).

Deacylation, essentially the chemical reverse of acylation, proceeds with a water molecule attacking the carbonyl carbon atom of the acyl enzyme and His-57 acting as a general base catalyst (Figure 1.3.-C) to form another transient tetrahedral intermediate (Figure 1.3.-D). His-57 donates a proton to the oxygen atom of Ser-195 which releases
Figure 1.3. Chymotrypsin cleavage of a peptide bond between phenylalanine and glycine. Individual steps in the mechanism are explained in the text.
the acid component of the substrate and regenerates the free enzyme (Figure 1.3.-E).
Thus, catalysis of peptide bond hydrolysis is said to proceed via a repetitive mechanism
where acylation and deacylation require both general base and general acid catalysis
(Equation 2).

\[
E + (S \text{ or } P) \xrightarrow{\text{general base}} \text{THI} \xrightarrow{\text{general acid}} E-P_2
\]

where THI = tetrahedral intermediate

Most of the modification and kinetic studies that elucidated the catalytic mechanism
of serine proteases were carried out on chymotrypsin and trypsin. These results were
found to be applicable to other serine proteases in view of the fact that the catalytic triad
occupies an almost identical geometric arrangement regardless of the folding of the main
chain of the protein.\textsuperscript{11} Thus for simplicity of comparison, the numbering applied to the
catalytic residues of serine proteases are those of chymotrypsin.

\subsection{1.1.2. Cysteine Proteases}

Another widely distributed family of enzymes are the cysteine proteases\textsuperscript{12,13} which
include papain, isolated from papaya, and the mammalian enzymes, cathepsin and calpain.
The catalytic activity of each of these proteases is inhibited by sulphydryl reagents such as
iodoacetate or \textit{p}-chloromercuribenzoate, implying involvement of a sulphydryl group in the
activity of the enzyme.

In the case of papain, the active site cysteine (Cys-25) residue forms an acyl-
thiolenzyme (thiol ester) intermediate with the acyl-moiety of the substrate during catalysis,
following a kinetic scheme similar to that of the serine proteases (Equation 1). X-ray
diffraction data\textsuperscript{14} and chemical studies\textsuperscript{13} point to the formation of a thiolate-imidazolium
ion-pair as a functional unit, implying that the proton is already on the imidazole when the
thiolate ion attacks the substrate. The consequence of the ion-pair formation is that, in
contrast to serine proteases, cysteine proteases follow a non-repetitive mechanism (Equation 3).

\[
E + (S \text{ or } P) \xrightarrow{\text{no catalysis}} \text{THI} \xrightarrow{\text{general acid}} E-P_2 \]

where THI = tetrahedral intermediate

Thus, formation of the tetrahedral intermediate (acyl-enzyme complex) is not promoted by general base catalysis but is a simple nucleophilic attack. Stabilization of the charged intermediate by an oxyanion binding site is not as pronounced as in serine proteases and contributes much less to the overall catalysis. Also in contrast to serine proteases, cysteine proteases do not have a charge stabilizing system involving a negatively charged aspartate residue. Instead, a neutral asparagine is found to be hydrogen bonded to the catalytic imidazole group.

1.1.3. Aspartic Proteases

Aspartic proteases were originally referred to as acid proteases due to their activity in acidic environments. This family includes the human immunodeficiency virus (HIV)-encoded protease; pepsin, the principle protease in gastric juice; and the microbial enzyme, penicillopepsin of Penicillum janthinellum. Studies of pH dependence of substrate hydrolysis\(^{15}\) implicated two catalytically active residues with pKa values of approximately 1 and 4, probably two carboxyl groups, one ionized and the other unionized in the active form of the enzyme. Two aspartic acid residues, Asp-32 and Asp-215 of pepsin, were labeled with active site specific reagents\(^{16}\) and were located by X-ray crystallography\(^{17}\) in the active site binding cleft within hydrogen bonding distance of each other.

Controversy remains regarding the mechanism of the aspartic proteases. Although not excluded as a mechanism, no evidence supporting the formation of a covalent intermediate has been observed; no covalent intermediates have been trapped, no initial rapid release of acyl or amino product have been observed, nor are the two carboxyl groups
accessible to substrate molecules, based on X-ray crystallography data.\textsuperscript{17} The alternative mechanism involving hydrolysis via a noncovalent intermediate has been proposed, utilizing a dissociated Asp-32 as a general base and promoting the nucleophilic attack of a water molecule on the carbonyl carbon of the substrate to form the tetrahedral intermediate. The nondissociated Asp-215 would act as a general acid to enhance polarization of the carbonyl bond. The mechanism of the breakdown of the noncovalent tetrahedral intermediate is not clear although it is known that the protonation of the peptide nitrogen atom is necessarily for the amine component of the substrate to be a good leaving group. It has been proposed that decomposition occurs by a similar concerted general acid-base catalysis with the Asp-32/Asp-215 pair accepting a proton from the protonated oxyanion and transferring a proton to the leaving nitrogen\textsuperscript{13} but it is not known which carboxyl is involved with each proton transfer.

1.1.4. Metalloproteases

The activity of the metalloproteases, such as the digestive enzyme carboxypeptidase A and the bacterial enzyme thermolysin, depends on a tightly bound zinc ion located in a binding cleft near the surface of the molecule where it is coordinated by two imidazole nitrogens, two carboxylate oxygens and a water molecule. Possible mechanisms of action of metalloproteases include; nucleophilic attack of an active site glutamate residue to form an acyl-enzyme (anhydride) intermediate with the carbonyl group of the substrate; or general base catalysis by the active site glutamate to promote the attack of a water molecule on the carbonyl carbon.

Based on X-ray crystal data and a computer graphics study of thermolysin catalysis,\textsuperscript{18} the active site glutamate residue (Glu-143) is hypothesized to promote the attack of the zinc-bound water molecule on the carbonyl carbon of the substrate. At the same time, the carbonyl oxygen is polarized as it becomes a ligand of the zinc which results in the formation of the tetrahedral intermediate bound to the zinc ion and stabilized by hydrogen bonds to Tyr-157 and His-231 sides chains. Decomposition of the tetrahedral
intermediate proceeds by way of proton transfer to the leaving nitrogen by general acid catalysis by Glu-143.

1.2. Recombinant DNA Technology

Recombinant DNA technology or "DNA cloning" evolved out of decades of research on DNA, RNA, and the enzymes involved in their function. It has grown into a powerful means of examining gene architecture, control of gene expression, and protein structure. The study of protein structure-function relationships has especially benefited from the ability to precisely change the DNA sequence of a cloned gene and determine its effects on either activity or function of the expressed protein.

1.2.1. Expression

A major goal of recombinant DNA technology is to over-express proteins for study. A number of general requirements are necessary for the DNA sequence to be competent for RNA and protein synthesis. These include promoter and terminator sequences for transcription of messenger RNA (mRNA) and a ribosome binding site (RBS) and initiation and termination codons for translation of the mRNA into protein. These features have such similar structural organization among the various organisms studied that sequences from one prokaryotic species may function in another, although with varying degrees of success and specificity.

Proteins to be expressed outside of the bacterial cell have further structural requirements due to the nature of the surrounding cell membrane and cell wall. Gram-positive bacteria, e.g. *Bacillus*, have a cell wall composed of a thick layer of peptidoglycan and teichoic or teichuronic acid, whereas Gram-negative bacteria, e.g. *E. coli*, have a thin layer of peptidoglycan and an additional outer membrane layer. To pass through the cell membrane, secreted proteins possess at the amino terminus a conserved region, the signal (leader) peptide, which is necessary for the protein to cross the membrane and is removed by a signal peptidase during or after translocation.19
Signal peptides from Gram-negative and Gram-positive bacteria display little primary sequence homology but have common structural features (Figure 1.4.).

Figure 1.4. Structure of bacterial signal peptide.

These features include an amino terminal basic (positively charged) hydrophilic sequence followed by a stretch of uncharged, hydrophobic residues and a more polar carboxy region. The amino acid sequence at the point of signal peptidase cleavage is that of a general recognition site, Ala-X-Ala * Ala (where * denotes the point of cleavage), found in most prokaryotes.²⁰

Signal peptides from Gram-positive organisms appear to be longer than those in Gram-negative bacteria (30-40 residues in total versus 22) with a tendency toward longer hydrophobic regions (14-23) and more highly charged hydrophilic regions. Despite these differences, signal peptides from Bacillus have been found to function sufficiently in E. coli to affect translocation of proteins across the inner membrane into the periplasmic space, although the reverse is not necessarily true. Also, signal peptides from one gene when fused in frame to the coding region of another is sufficient in many instances to direct secretion of the heterologous protein outside of the cell. Signal peptides from Bacillus subtilisin have been successfully used to secrete a variety of proteins including TEM β-lactamase,²¹ staphylococcal protein A,²²,²³ human atrial natriuretic α-factor,²⁴ and bovine pancreatic ribonuclease A.²⁵ Each of the expressed heterologous proteins are released into the culture medium after cleavage by the signal peptidase.²⁶
1.3. Secreted Proteases

Secreted bacterial proteases have been isolated and characterized from a variety of sources. DNA sequence analysis of the genes for several of these enzymes including various subtilisins from Bacillus spp.;27-30 proteases A and B from Streptomyces griseus;31 and α-lytic protease from Lysobacter enzymogenes;32 revealed that many proteases are initially translated as precursors, consisting of a signal peptide followed by a long polypeptide extension (pro-region) joined to the amino terminus of the mature protease (Figure 1.5.). This suggests that further processing, other than signal peptidase cleavage, is necessary for release of the mature enzyme. The mechanism of this processing was suggested by studying secretion of recombinant subtilisin BPN’ (Bacillus amyloliquefaciens) in a subtilisin E deficient Bacillus subtilis host.33 Mutations of the cloned gene affecting activity of the protease (Asp32Asn) effectively blocked maturation of the precursor. When expressed in a subtilisin E proficient host, processing by endogenous enzyme released the mutant form into the medium. Thus, processing appears to occur by an intramolecular, autocatalytic mechanism. This knowledge was used to produce a variety of inactive subtilisin BPN’ mutants in order to study catalytic contributions of each of the catalytic residues of this enzyme.34 Addition of a wild-type or “helper” subtilisin BPN’ to cultures of the inactive variants catalyzed the necessary maturation.

Figure 1.5. Pre-pro-mature protease precursor.

Many functions of the pro-region have been hypothesized: 1) The pro-region may play a role in the export process;35 2) It may be required for the association of the pro-enzyme with the cell before release into the medium;36 3) It may function to keep the protease inactive;37,38 or 4) It may play an essential role in guiding the folding of protease into the proper conformation necessary for activity. This last function has been clearly
demonstrated for both subtilisin E of *B. subtilis* 168\(^{39,41}\) and **α**-lytic protease of *L. enzymogenes*.\(^{35,42,43}\)

When expressed in *E. coli*, active subtilisin E from *B. subtilis* was secreted into the periplasmic space as directed by its own signal peptide or that of the *E. coli* gene **ompA** when fused to pro-mature subtilisin E. When the **ompA** signal peptide sequence was fused directly to the mature enzyme sequence, protein of the appropriate size was produced although no activity was detected.\(^{39}\) *In vitro* maturation of purified pro-subtilisin E expressed in *E. coli* was also demonstrated by denaturation of pro-subtilisin E in guanidinium chloride and subsequent dialysis which produced processed active enzyme, although active mature subtilisin E alone was not renatured after similar treatment.\(^{40}\) Thus, the pro-region appears to be necessary for subtilisin E activity both *in vivo* and *in vitro*. The pro-region however does not need to be contiguous with the mature sequence to guide proper folding.\(^{41}\) Mutant pro-subtilisin E (Asp32Asn), unable to affect its own processing,\(^{39,40}\) can function in an intermolecular reaction to guide the correct refolding of not only mature subtilisin E but also subtilisin Carlsberg of *B. licheniformis* and subtilisin BPN' of *B. amyloliquefaciens*, accounting for the previously noted pro-sequence similarities of these enzymes in both the predicted secondary structure and hydrophobic residue distribution.\(^{30}\)

Similar studies have also been performed on the **α**-lytic protease gene cloned from *L. enzymogenes*, a Gram-negative bacterium.\(^{32}\) Secretion of the pro-mature portion of the protein from the *E. coli* alkaline phosphatase (**phoA**) promoter and signal peptide resulted in accumulation of active, processed protease whereas constructs containing the conservative active site mutation (Ser143Ala) resulted in accumulation of a higher molecular weight form, indicating an autocatalytic mechanism of processing.\(^{42}\) Similarly, constructs expressed without the pro-region resulted in cell-associated inactive enzyme.\(^{42}\) Independently expressing the pro- and mature regions within the same bacterium (complementation) resulted in active protease thus providing further evidence for the
absolute requirement of the presence of the pro-region for protease activity, although direct covalent linkage was not necessary for production of active protease. In vitro experiments confirmed this result using purified pro- and mature proteins denatured in guanidinium chloride.

Autocatalytic activation of subtilisin E is consistent with its known substrate specificity although the same is not true for α-lytic protease, which is specific for small hydrophobic residues. The amino acid sequence surrounding the pro-mature junction of α-lytic protease is Gln-Thr-Thr * Ala-Asn (where * denotes the bond cleaved) and should be a poor substrate for the mature enzyme. Autocatalytic processing of this protease precursor requires that three hydrophilic residues occupy the S3, S2 and S1 subsites of the enzyme. If the cleavage is autocatalytic, it would require an altered specificity of the proenzyme molecule resulting in cleavage before the alanine residue, which is the preferred P1 residue.

In vivo and in vitro experiments using proteins with deletions of the N-terminal end of the subtilisin E pro-region revealed that the entire pro-region is necessary for producing processed active subtilisin E. In vitro refolding experiments using synthetic pro-peptide fragments of subtilisin E were successful only when using the entire pro-sequence. Similarly, deletions of more than five amino acids at any of six sites spanning the pro-region of α-lytic protease were found to abolish protease activity and disturb secretion, although the mature protease coding region itself was not changed.

Studies involving both subtilisin E and α-lytic protease reveal an inhibitory action of the pro-region on the active enzymes. The synthetic subtilisin E pro-region elicits specific strong binding to subtilisin E and functions as a competitive inhibitor. Smaller fragments had a much weaker interaction, indicating that the entire sequence is essential for specificity of interaction of the pro-region with the mature region. Such interaction with the active site of subtilisin E had been hypothesized previously, based on the amino acid composition of the pro-region and the X-ray crystal structure of the mature protein.
Charged residues were found to be unevenly distributed on the surface of subtilisin E with the active site cleft having only a few charged residues. It had been speculated that the highly charged pro-sequence (23 charged residues out of 77 amino acids) initially binds to this area, resulting in a more even distribution of charges on the pro-subtilisin E surface. This was substantiated by the finding that processing of pro-subtilisin E was not inhibited by PMSF or *Streptomyces* subtilisin inhibitor.\textsuperscript{40} This suggested that if the subtilisin E active site was involved in the autocatalytic processing reaction, the active site was probably occupied by the pro-sequence such that the inhibitors could not interact with the active site. The pro-region of $\alpha$-lytic protease has also been found to be a high affinity inhibitor of the mature protease in *in vitro* refolding experiments.\textsuperscript{38} Other examples of pro-region inhibition include pancreatic carboxypeptidases A\textsuperscript{47} and Y,\textsuperscript{48} and aspartic protease pepsinogen.\textsuperscript{49}

Protein folding involves kinetic competition between reactions off the pathway (e.g. aggregation) and on the pathway leading to the native state. The observation that the pro-region interacts strongly with the native enzyme suggests that the pro-region directly facilitates the on-pathway reaction and that the rate-limiting folding transition state has native-like conformation.

In order to further analyze the role of the pro-region in the folding of these enzymes, the pro-region of subtilisin E was subjected to random mutagenesis to isolate mutations that prevent secretion of active subtilisin E.\textsuperscript{50,51} Using the polymerase chain reaction with *Taq* DNA polymerase (and its inherent spontaneous misincorporation rate), a total of 25 amino acid substitution mutations affecting secretion of active protease were identified and found to occur with a high frequency in the hydrophobic regions of the pro-peptide. This suggests that the pro-peptide consists of a few functional regions which interact with specific regions of the mature region of subtilisin E during the folding process. Assuming that specific interactions between pro-peptide and mature subtilisin E are important during renaturation, it is likely that certain amino acid substitutions in the pro-
peptide would lead to defects in production of active subtilisin E. Further study is necessary to characterize the effects of these changes and to isolate active revertants by performing similar mutagenesis on the coding region of mature subtilisin E. For one of the pro-region mutations Met(-60)Thr, a second-site suppressor mutation Ser(188)Leu was isolated within the mature region, indicating that these two residues may be located very close to one another in the pro-subtilisin E structure. Ser188 is located on the surface of the subtilisin E molecule near the active center cleft. Since the synthetic pro-peptide functions as a competitive inhibitor, the pro-peptide is likely to bind in the region where the active center is exposed. By isolating several second-site suppressor mutations with mature subtilisin E for each pro-peptide mutation, one should be able to map the interactions between specific regions of the pro-peptide and mature subtilisin E.

1.4. *Streptomyces griseus* Proteases

The Gram-positive organism *Streptomyces griseus* secretes a number of extracellular proteases which have a suspected evolutionary relationship with chymotrypsin. *S. griseus* protease B (SGPB) has a high degree of structural homology to bovine chymotrypsin despite the difference in size (18 600 Da versus 21 600 Da) and a very low primary structural homology (<21% identity) between the bacterial and pancreatic enzymes.52 Two of the *S. griseus* proteases, SGPA and SGBP, have been cloned and sequenced and show a high degree of homology with each other (61% amino acid identity)31 as well as to α-lytic protease (37%).38 All three proteases have several common deletions when compared with chymotrypsin and each has a large insertion between residues 186 and 187 (of the chymotrypsinogen A sequence).53 The *S. griseus* proteases show no homology to the other subtilisin-like group of serine proteases although they share a similar gene organization, i.e. that of the pre-pro-mature protease (Figure 1.6.).
The *Bacillus* spp. subtilisins demonstrate high homology amongst themselves both in the pro-region which includes a common hexapeptide sequence (Tyr-Ile-Val-Gly-Phe-Lys) located 10 residues from the signal peptide cleavage site and in the mature region with 62.5% amino acid identity.

The activation of SGPB and the other bacterial protease precursors differ from that of chymotrypsin. Chymotrypsinogen is cleaved by another pancreatic enzyme, trypsin, between Arg15 and Ile16 (a disulfide bond connects two fragments) to give the fully active π-chymotrypsin. π-Chymotrypsin is further modified by chymotrypsin-catalyzed cleavage to remove two dipeptides, Ser14-Arg15 and Thr147-Arg148, and undergoes conformational changes to give α-chymotrypsin. The initial cleavage of chymotrypsinogen allows the rotation of Ile16 in order to form an ionic bond between free α-amino group of Ile16 and carboxylate of Asp194. This structural change allows the formation of the substrate binding pocket and thus activates the enzyme.

The processing of SGPB involves the removal of the pro-region and the release of the mature protease from the bacterial cell. The SGPB pro-region may act as a competitive inhibitor of the protease, as has been demonstrated for the pro-regions of subtilisin E and the related α-lytic protease. The pro-region of α-lytic protease has a high homology to that of SGPB (with the highest homology in the C-terminal 50 residues) and competitively
inhibits mature SGPB in kinetic assays\textsuperscript{38} despite different substrate specificities. This is not the case for the more distantly related yet structurally similar serine protease, elastase, which shows no inhibition by the $\alpha$-lytic protease pro-peptide.

Thus, the mechanism of SGPB activation would appear to differ from that of chymotrypsinogen in that the pro-region is removed from the precursor and that the ionic bond for Asp194 is already provided by the guanidinium group of a buried arginine residue (Arg138)\textsuperscript{52} and does not directly result from cleavage of the pro-region.

1.5. Protein Synthesis

1.5.1. Chemical Protein Synthesis

With the advent of molecular cloning techniques and protein engineering, studies on the structure-function relationship in proteins have increased the understanding of the contributions of individual amino acid residues to the activity of the protein. These techniques are not without limitation; site-directed mutagenesis of cloned genes can only affect substitutions of the twenty naturally occurring amino acids. In some instances, the ability to introduce a residue of novel or unique structure would allow the design of proteins with tailored functional properties.\textsuperscript{54} Proteins possessing artificial analogues of amino acids or nonpeptide regions are potentially capable of many new functions, including physical studies of the altered protein in a reconstituted biological environment using an amino acid analog containing a spectroscopic label at specific locations in the peptide.

Introduction of novel amino acids into peptides is possible through total chemical synthesis. In fact, since the development of the Merrifield step-wise, solid-phases method of peptide synthesis,\textsuperscript{55} together with improvements in peptide purification, chemical methods are often superior for synthesis of peptides of up to 50 amino acids in length. If one desires to study a large number of short peptide analogs (30-40 amino acids), chemical synthesis is the method of choice for its speed and ease of application.
synthesis. Purification is costly and extremely difficult as impurities may only differ by one or two amino acids. For peptides of this length, discrete segments of the peptide are synthesized and the fragments are condensed either in free solution or on solid supports.

1.5.2. **Protein Semi-synthesis**

Semi-synthesis, an alternative to total synthesis, combines fragments derived from natural sources with those from artificial sources. Only the part of the molecule to be changed is chemically synthesized or altered. The resulting fragments are condensed using proteolytic enzymes working in reverse to ensure correct coupling of activated peptides with little or no side chain protection necessary. Enzymes of all mechanistic types have been used in this process.\(^56,57\) The most notable accomplishment was the conversion of porcine insulin into human insulin in marketable quantity and purity.\(^54\)

Recent interest has focused on peptide coupling via the enzyme, thiolsubtilisin,\(^58\) a chemically altered subtilisin with poor protease activity but broad substrate specificity. The difficulty in using proteases with broad specificity include the potential to bind and catalyze transpeptidation reactions at many sites on large substrates. New variants of subtilisin are being constructed using recombinant DNA technology\(^59\) to change substrate specificity in experiments analogous to those performed on the wild-type enzyme. With these new variants, a number of specific protein ligases can be produced.

Alternatively, inactivation of the protease by replacing the active site serine residue with a glycine would result in no competing hydrolytic activity of the protease and would allow the binding of an activated C-terminal methyl ester in the following ligation reaction:

\[
\begin{align*}
R-C-O-O-CH_3 + R'-NH_2 &\rightarrow R-C-O-NH-R' + CH_3OH
\end{align*}
\]

The protease-catalyzed reaction of the donor and recipient peptides with the elimination of methanol would be thermodynamically favourable due to the increased effective concentration of the reactants by virtue of their interaction with the enzyme and the remaining binding determinants. The oxyanion hole will provide electrostatic stabilization
concentration of the reactants by virtue of their interaction with the enzyme and the remaining binding determinants. The oxyanion hole will provide electrostatic stabilization of the transition state complex and His-57 will mediate passage of a proton from the amino group to the leaving group methanol. The equilibrium of this reaction will also be shifted toward amide bond formation by carrying out the reactions in high concentrations of recipient peptide. Peptide bond formation is also favoured by performing reactions in organic solvents, by ensuring that the products of coupling are less soluble than the reactants.

1.6. Summary

A number of proteases have been purified from Pronase, a commercially available product derived from culture filtrates of the thermophilic bacterium *Streptomyces griseus*. One of these proteases, SGPB, has been purified and found to be both thermostable and active in the presence of denaturing chemicals such as guanidinium chloride\(^60,61\) with the conservation of \(P_1\) specificity for leucine, tyrosine or phenylalanine.\(^60\) These characteristics are beneficial for proteases used in protein semi-synthesis as increased temperature and denaturing agents are sometimes necessary to expose the buried ends of the larger peptides. This is the major drawback in the use of subtilisin or chymotrypsin for semi-synthesis due to their susceptibility to chemical denaturation.\(^62\)

The basis of the following thesis has been the development of expression systems for the protease SGPB and the creation of an inactive variant of SGPB for the eventual use in protein semi-synthetic reactions.
2. Materials and Methods

2.1. Buffers and Other Solutions

All solutions were prepared using distilled deionized water (ddH$_2$O, Millipore Milli-Q Water System) and chemicals of the highest grade available. Unless indicated otherwise all percentages (%) were weight per volume (w/v).

2.1.1. Tris Buffers

TE: 10 mM Tris base, 1 mM EDTA. Adjust pH with HCl as required and autoclave.

TAE: 40 mM Tris-acetate, 0.001 M EDTA. Prepare as a 50x stock and dilute as required.

TBE: 90 mM Tris-borate, 0.002 M EDTA. Prepare as a 10x stock and dilute as required.

2.1.2. Transformation Solutions

TfBII: 30 mM potassium acetate, 50 mM MnCl$_2$, 100 mM KCl, 10 mM CaCl$_2$, 15% (v/v) glycerol (BRL, ultrapure). pH should be approximately 7.0. Sterilize by filtration.

TfBII: 10 mM sodium MOPS (3-[N-morpholino]propanesulfonic acid) (pH 7.0), 75 mM CaCl$_2$, 10 mM KCl, 15% (v/v) glycerol (BRL, ultrapure). Sterilize by filtration.

2x SMM: 1.0 M sucrose, 0.04 M MgCl$_2$, 0.04 M maleate (pH 6.5). Autoclave.

4x Penassay: 7.0% antibiotic medium 3 (Difco). Autoclave.

SMMP: Mix equal volumes of 2x SMM and 4x Penassay solutions after autoclaving.

40% PEG solution: 40% polyethylene glycol (approximate molecular weight: 6,000-8,000, Sigma), 1x SMM. Autoclave.

2.1.3. DNA Preparation

Solution I: 25 mM Tris-Cl (pH 8.0), 50 mM glucose, 10 mM EDTA. Autoclave. Add solid lysozyme to a final concentration of 4 mg/mL immediately prior to use.

Solution II (Bacillus): 0.2 M NaCl, 1% SDS. Prepare by mixing equal volumes of 0.4 M NaCl, 2% SDS immediately prior to use.
Solution II (E. coli): 0.2 M NaOH, 1% SDS. Prepare by mixing equal volumes of 0.4 M NaOH, 2% SDS immediately prior to use.

Solution III: 60 mL 5 M potassium acetate, 11.5 mL glacial acetic acid, 28.5 mL ddH2O.

RNase A: Dissolve pancreatic RNase (RNase A, Sigma, Molecular Biology Grade) at a concentration of 10 mg/mL in 10 mM Tris-Cl (pH 7.5)/15 mM NaCl. Boil for 15 min and allow to cool slowly to room temperature. Dispense into aliquots and store at -20°C.

2.1.4. DNA Manipulation

1x Mung bean nuclease buffer: 30 mM sodium acetate (pH 4.6), 50 mM NaCl, 1 mM ZnCl2, 5% glycerol.

10x T4 polynucleotide kinase buffer: 0.5 M Tris-Cl (pH 7.6), 0.1 M MgCl2, 50 mM dithiothreitol (DTT), 1 mM spermidine HCl, 1 mM EDTA.

10x PE1: 200 mM Tris-Cl (pH 7.5), 100 mM MgCl2, 500 mM NaCl, 10 mM DTT.

10x PE2: 200 mM Tris-Cl (pH 7.5), 100 mM MgCl2, 100 mM DTT.

10x Dephosphorylation buffer: 500 mM Tris-Cl (pH 8.5), 1 mM EDTA.

5x Ligation buffer: 250 mM Tris-Cl (pH 7.6), 50 mM MgCl2, 25% polyethylene glycol (8,000), 5 mM ATP, 5 mM DTT.

5x Sequenase buffer: 200 mM Tris-Cl (pH 7.5), 100 mM MgCl2, 250 mM NaCl.

2.1.5. RNA Preparation and Manipulation

RNA Lysis buffer: 30 mM Tris-Cl (pH 7.4), 100 mM NaCl, 5 mM EDTA, 1% SDS. Autoclave.

DNase digestion buffer: 40 mM Tris-Cl (pH 7.9), 10 mM NaCl, 6 mM MgCl2, 0.1 mM CaCl2. Autoclave.

10x MOPS buffer: 200 mM MOPS acid, 50 mM sodium acetate, 10 mM EDTA. Adjust to pH 7.0 with NaOH.

Formaldehyde loading buffer: 0.75 mL formamide (deionized with Amberlite MG-1 mixed bed exchanger, Sigma), 0.15 mL 10x MOPS buffer, 0.24 mL
formaldehyde, 0.1 mL ddH₂O, 0.1 mL glycerol, 0.08 mL 10% bromophenol blue. Store at -20°C in small aliquots.

Pre-swollen Sephadex G-25: Prepare by autoclaving 2-3 g Sephadex G-25 in excess (10-15 mL) TE (pH 8.0). Remove layer of "fines" before use.

Hybridization buffer: 20 mg bovine serum albumin, 5 mL 2 M sodium phosphate/4 mM EDTA (pH 7.2), 5 mL 20% SDS, 10 mL deionized formamide. Prepare by adding components in the order given, warming to dissolve if necessary.

10x SSC: 1.5 M NaCl, 0.15 M sodium citrate (pH 7.0).

2.1.6. Protein Manipulation

Separating buffer: 0.75 M Tris·Cl (pH 8.8), 0.2% SDS.

Stacking buffer: 0.1458 M Tris·Cl (pH 6.8), 0.117% SDS.

Acrylamide stock: 30% acrylamide, 0.8% bis-acrylamide. Make up acrylamide stock to volume, deionize with approximately 2% Amberlite MG-1 mixed bed exchanger (Sigma) for 30 min, and filter using Whatman #1 filter paper.

5x SDS loading buffer: Mix in a 50 mL Falcon tube: 7.5 mL 2.083 M Tris·Cl (pH 6.8), 11.0 mL β-mercaptoethanol (1.12 g/mL), 5 g SDS. Add water up to 25 mL and glycerol up to 50 mL. Add 2.5 mg bromophenol blue. Store in aliquots at -20°C.

SDS-PAGE running buffer: 0.025 M Tris·Cl (pH 8.3), 0.1% SDS, 0.192 M glycine.

Staining solution: 0.3% Coomassie Brilliant Blue G in fast destaining solution.

Fast destaining solution: 50% (v/v) methanol, 10% (v/v) glacial acetic acid.

Slow destaining solution: 10% (v/v) glacial acetic acid.

Transfer buffer: 39 mM glycine, 48 mM Tris·Cl, 0.0375% SDS.

10x TBS: 1.54 M NaCl, 0.20 M Tris·Cl (pH 7.4).

Blocking solution ('blotto'): 1x TBS, 0.05% (v/v) polyoxyethylene sorbitan monolaurate (Tween 20), 5% Carnation skim milk powder. Add milk powder immediately prior to use.
Lysis buffer (protein preparation): 10 mM sodium phosphate (pH 7.0), 30 mM NaCl, 0.25% Tween 20, 10 mM β-mercaptoethanol, 10 mM EDTA, 10 mM EGTA.

Column buffer: 10 mM sodium phosphate (pH 7.0), 0.5 M NaCl, 1 mM sodium azide, 10 mM β-mercaptoethanol, 1 mM EGTA.

2.2. Bacteriological

2.2.1. Media

Culture plates were made by adding 1.5-2.0% agar to the medium prior to autoclaving.

TY broth: 1.0% tryptone, 0.5% yeast extract, 0.5% NaCl.

2x TY broth: 1.6% tryptone, 1.0% yeast extract, 0.5% NaCl.

TY/Mg broth: 2.0% tryptone, 0.5% yeast extract, 0.5% NaCl, 10 mM MgSO4.

5x MM salts: 5.25% K2HPO4, 2.25% KH2PO4, 0.5% (NH4)2SO4, 0.25% sodium citrate.

1x Penassay broth: 1.75% antibiotic medium 3 (Difco).

Minimal agar: Autoclave 7.5 g agar in 400 mL water. Add 100 mL sterile 5x MM salts and cool to 55°C prior to adding 0.5 mL 20% MgSO4, 5.0 mL 20% glucose, and 0.25 mL filter sterilized 1% Vitamin B1.

Milk plates: TY broth, 1.5% agar, 1.5% Carnation skim milk powder (final concentration). Autoclave 350 mL TY-agar (5.0 g tryptone, 2.5 g yeast extract, 2.5 g NaCl, 7.5 g agar) and 150 mL skim milk solution (7.5 g Carnation skim milk powder) separately and mix immediately prior to pouring.

Regeneration medium: Mix the following sterile solutions (per liter); 350 mL nutrient agar (10.0 g agar, 5.0 g casamino acids, 5.0 g yeast extract); 500 mL 1 M sodium succinate (pH 7.3); 100 mL 3.5% K2HPO4 and 1.5% KH2PO4; 25 mL 20% glucose; 20 mL 1 M MgCl2; 5 mL filter-sterilized 2% bovine serum albumin; 1 mL 100 mg/mL kanamycin.

H-Top agar: 1.0% tryptone, 0.8% NaCl, 0.8% agar. Autoclave.
Antibiotics were prepared as stock solutions and were diluted for use as follows:

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Stock Solution (stored at -20°C)</th>
<th>Working Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>200 mg/mL in 50% (v/v) ethanol</td>
<td>200 μg/mL</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>100 mg/mL in water (filter sterilized)</td>
<td>50 μg/mL</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>20 mg/mL in 100% ethanol</td>
<td>20 μg/mL</td>
</tr>
</tbody>
</table>

Note: Magnesium ions inhibit the action of tetracycline, so straight TY was used when selecting bacteria on tetracyline.

2.2.2. Bacterial Strains

*Escherichia coli:*

DH5α™:63 supE44 ΔlacU169 (φ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1.

A patented (Gibco/BRL) recombination-deficient suppressing strain used for plating and growth of plasmids and cosmids. The φ80 lacZΔM15 permits α-complementation with the amino terminus of β-galactosidase encoded in pUC vectors. Grown on rich media (2x TY broth or TY agar) at 37°C.

K12cl: A strain that constitutively expresses cI repressor. Grown on rich media at 37°C.

K12ΔH1:64 F- Δ(bio- uvrB) lacZam λ Nam7 Nam53 cl857 ΔH1(cro-F-A-J-b2)

A defective phage λ lysogenic strain with a temperature-sensitive repressor gene used as a host for plasmids containing the bacteriophage λ pL promoter and lacking a functional N-protein. Grown on rich media at 30°C.

QY13:65 F- lacam trpam B B' bio-256 N+ cl857 ΔH Smr recA.

A defective phage λ lysogenic strain with a temperature-sensitive repressor gene and an active N gene used as a host for plasmids containing the bacteriophage λ pL promoter. Grown on rich media at 30°C.

RZ1032:64 lacZbd-279::Tn10 (Tet') Hfr dut-l ung-1 recA.

A dut- ung- strain that supports growth of M13 phage and can produce phage with uracil in the DNA. Grown on rich media supplemented with tetracycline at 37°C.
TB1:66 F- ara Δ(lac-proAB) rpsL (Str') (φ80 lacZΔM15) hsdR rK−mK+
A strain used for plating and growth of plasmids and cosmids. The φ80 lacZΔM15 permits α-complementation with the amino terminus of β-galactosidase encoded in pUC vectors. Grown on rich media at 37°C.

TG2:63 supE Δ(lac-proAB) hsd5 Δ(srl-recA)306::Tn10 (Tet') F' traD36 proAB+ lacI4 lacZΔM15.
A recombination-deficient EcoK- strain that neither modifies nor restricts transfected DNA. It will support growth of vectors carrying amber mutations. Grown on rich media supplemented with tetracycline or on minimal media at 37°C.

Bacillus subtilis:

DB104:67 his nprR2 nprE18 ΔaprA3
An extracellular protease-deficient strain that carries lesions in the structural genes for extracellular neutral (nprE) and serine (aprA) proteases and has less than 4% of extracellular protease activity of the wild type. Grown on rich media (2x TY broth or TY agar) or 1x Penassay at 37°C.

2.2.3. Preparation of Competent E. coli Cells
One mL TY/Mg broth was inoculated with a fresh colony of E. coli and incubated with shaking at the appropriate temperature (strain-dependent) to mid-log phase measured as optical density of 0.6 at 600 nm (OD600=0.6). This culture was used to inoculate 20 mL TY/Mg broth (prewarmed) in a 250 mL Erlenmeyer flask which was incubated with shaking to mid-log phase. This second culture was used to inoculate 200 mL TY/Mg broth (prewarmed) in a 1 L Erlenmeyer flask which was incubated with shaking to an optical density of 0.6-0.8. The resulting culture was chilled in an ice water bath and centrifuged at 3,500 rpm for 15 min at 2°C (Sorvall centrifuge, GSA rotor). The following procedures were performed at 4°C. The cell pellet was resuspended in 40 mL
cold TfBI buffer and recentrifuged. The resulting cell pellet was resuspended in 8 mL cold TfBI buffer and divided into approximately 300-500 µL aliquots and placed on dry ice. Once frozen, the competent *E. coli* cells were stored at -80°C.

2.2.4. Transformation of *E. coli*

2.2.4.1. Transformation of *E. coli* DH5α™, TB1, and K12cI

A microcentrifuge tube containing frozen competent *E. coli* cells was allowed to thaw on ice for 10 min. Aliquots of 100-200 µL were dispensed into sterile microcentrifuge tubes and DNA [ligation solution or covalently-closed circular (ccc)-plasmid DNA] was added. The mixture was incubated on ice for 15-30 min. Heat-shocking was achieved by swirling the microcentrifuge tube in a 37°C water bath for 1 min, followed by incubation on ice for 1 min. The mixture was added to 2 mL 2x TY broth and incubated at 37°C with shaking for 1-2 h. Dilutions equivalent to 0.1 µL to 500 µL of the 2 mL culture were plated on TY-agar plates containing 200 µg/mL ampicillin and incubated at 37°C for 12-16 h.

If colour selection using α-complementation of β-galactosidase in *E. coli* DH5α™ or TB1 was desired, aliquots of 40 µL X-gal (20 mg/mL in dimethylformamide) and 4 µL IPTG (200 mg/mL, filter-sterilized) were spread onto TY-agar plates containing 200 µg/mL ampicillin and allowed to dry before plating transformed bacteria. After incubation at 37°C, colour was developed further by incubating the plates at 4°C for 1-2 h.

2.2.4.2. Transformation of *E. coli* QY13 with pLcII-FX constructs

A microcentrifuge tube containing frozen competent *E. coli* QY13 was allowed to thaw on ice for 10 min. Aliquots of 100-200 µL were dispensed into sterile microcentrifuge tubes and DNA was added. The mixture was incubated on ice for 40 min. Heat-shocking was achieved by swirling the microcentrifuge tube in a 30°C water bath for 5 min, followed by incubation on ice for 15 min. An equal volume 2x TY
broth was added and this mixture was incubated at 30°C for 1 h with no shaking. Dilutions equivalent to 0.1 μL to 100 μL of the mixture were plated on TY-agar plates containing 200 μg/mL ampicillin and incubated at 30°C for 12-16 h.

2.2.4.3. Transformation of *E. coli* TG2 by Bacteriophage M13

A microcentrifuge tube containing frozen competent *E. coli* TG2 was allowed to thaw on ice for 10 min. Aliquots of 100-200 μL were dispensed into sterile microcentrifuge tubes and DNA (ligation solution or double-stranded bacteriophage M13 DNA) was added. The mixture was incubated on ice for 15-30 min. Heat-shocking was achieved by swirling the microcentrifuge tube in a 37°C water bath for 1 min, followed by incubation on ice for 1 min. Dilutions equivalent to 0.1 μL to 100 μL of the mixture were added to 3 mL molten H-top agar (45°C), followed by 200 μL *E. coli* TG2. The agar solution was mixed and poured onto a TY-agar plate. The plate was gently swirled to evenly distribute the top-agar and allowed to harden for 5 min. Once solidified, the plate was inverted and incubated at 37°C for 8-12 h.

If colour selection using α-complementation of β-galactosidase in *E. coli* TG2 was desired, aliquots of 40 μL X-gal (20 mg/mL) and 4 μL IPTG (200 mg/mL) were added to the top-agar before pouring.

2.2.5. Preparation of Bacillus Protoplasts

Two mL 1x Penassay broth were inoculated with a fresh colony of *Bacillus subtilis* DB104 and incubated with shaking at 37°C to mid-log phase (OD$_{600}$=0.6). This culture was used to inoculate 50 mL 1x Penassay broth in a 250 mL baffled Erlenmeyer flask which was incubated with shaking to an optical density of 0.6. The resulting culture was chilled in an ice water bath and centrifuged at 5,000 rpm for 10 min at 4°C (Sorvall centrifuge, SS-34 rotor). The cell pellet was resuspended gently (by swirling) in 5 mL SMMP. Lysozyme was added to 2 mg/mL (final concentration) and the suspension was incubated at 37°C for 2 h with gentle shaking. The protoplasts were centrifuged at 5,000 rpm for 15 min at room temperature and washed once by resuspending in 5 mL
SMMP and recentrifuging. The final pellet was resuspended in 4 mL SMMP. Although protoplasts are said to be stable at room temperature for at least 5 h, aliquots of 500 μL were dispensed into sterile microcentrifuge tubes and frozen slowly (by placing in styrofoam at -80°C) for future use.

2.2.6. Transformation of *Bacillus* Protoplasts

The following procedure was performed at room temperature. A microcentrifuge tube containing frozen *B. subtilis* DB104 protoplasts was thawed quickly in a 37°C water bath. DNA to be transformed (ligation solution or ccc-plasmid DNA) was prepared by adding an equal volume of 2x SMM. The thawed protoplast solution was added to the DNA solution, mixed and transferred to a sterile 50 mL capped centrifuge tube containing 1.5 mL 40% PEG solution. The tube was vortexed manually to obtain a uniform mix. After incubation for 2 min, 5 mL SMMP was added and the protoplasts were recovered by centrifuging at 5,000 rpm for 10 min. The pellet was resuspended in 1 mL SMMP and incubated at 30°C for 90 min with gentle shaking. Aliquots of 300-500 μL were gently plated onto regeneration medium and incubated for 24-48 h at 37°C. When grown, colonies were restreaked onto TY-agar, 50 μg/mL kanamycin to confirm transformation success.

2.2.7. Cryopreservation of Bacterial Strains

A single bacterial colony was the inoculum in two mL 2x TY broth containing the appropriate antibiotic. After incubation with shaking at the appropriate temperature, aliquots of the mid-log phase (OD$_{600}$=0.6) bacterial culture were dispensed into sterile microcentrifuge tubes and glycerol was added to 15% (v/v). The tubes were placed on dry ice for 5 min and stored at -80°C. Fresh cultures were started from these stocks by removing and streaking small amounts of frozen culture onto the appropriate agar plate followed by incubation at the desired temperature.
2.3. Nucleic Acid Manipulation

2.3.1. Plasmids and Bacteriophage Vectors

2.3.1.1. *E. coli*

M13mp18/19\(^{68}\) Bacteriophage M13 is a filamentous virus that infects *E. coli* strains carrying the F' episome. The viral DNA molecule is a single-stranded circle while the double-stranded replicative form (RF) is an intermediate during DNA replication. The M13mp vectors contain the N-terminal portion of the *E. coli* lacZ gene that expresses the α-peptide of β-galactosidase. The α-peptide is capable of complementing a lac mutation in the host, lacZΔM15, to produce a functional β-galactosidase. Insertion of DNA at the multiple cloning site results in interruption of the α-peptide, producing colourless, rather than blue, plaques on medium containing isopropylthio-β-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyI-β-D-galactoside (X-gal). M13mp18 and M13mp19 differ only in the orientation of the multiple cloning site.

pUC18/19\(^{68}\) Plasmids pUC18 and pUC19 are *E. coli* cloning vectors that carry the β-lactamase gene (ampicillin resistance, Ap\(^{r}\)) and the α-peptide of the lacZ (β-galactosidase) gene. Insertion of DNA at the multiple cloning site results in interruption of the α-peptide, producing colourless, rather than blue, colonies on medium containing IPTG and X-gal. Plasmids pUC18 and pUC19 differ only in the orientation of the multiple cloning site.

pLeII-FX\(^{70}\) Plasmid pLeII-FX is a pUC9-based protein expression vector (carrying the β-lactamase gene) which directs synthesis of a fusion protein transcribed from p\(_L\) promoter. The hybrid protein consists of the 31 amino terminal residues of the λcII protein and the tetrapeptide (Ile-Glu-Gly-Arg) recognition site for a protease isolated from the blood coagulation cascade, Factor Xa, fused to the cloned protein.
Plasmids pMAL-p and pMAL-c are pBR322-based protein expression vectors (carrying the β-lactamase gene) which direct synthesis of fusion proteins transcribed from the p_lac promoter. The vectors express the malE gene (with or without its signal sequence) fused to the lacZα gene. Insertion of DNA at the multiple cloning site results in interruption of the α-peptide, producing colourless colonies on medium containing IPTG and X-gal. The resulting fusion protein is directed to either the periplasm (pMAL-p) or cytoplasm (pMAL-c) and is purified using the affinity of the maltose binding protein (MBP) for maltose. The desired protein is separated from MBP by the tetrapeptide (Ile-Glu-Gly-Arg) recognition site of Factor Xα which is cleaved following purification.

2.3.1.2. Bacillus

Plasmid pUB110 was originally isolated as a Staphylococcus aureus plasmid encoding resistance to kanamycin and neomycin (kanamycin nucleotidyltransferase) and is used for cloning in B. subtilis.

2.3.2. DNA Isolation

2.3.2.1. Preparation of Single-stranded Bacteriophage M13 DNA

Two mL 2x TY broth supplemented with tetracycline was inoculated with a single bacterial colony of E. coli TG2 and incubated with shaking at 37°C for 6-18 h. Bacteriophage M13-infected E. coli TG2 was prepared by dilution of 100 μL of E. coli TG2 with 2 mL 2x TY broth and addition of 5 μL of bacteriophage stock supernatant or a single well-isolated bacteriophage plaque. Infected cultures were incubated with shaking at 37°C for 4-5 h.

All procedures were performed at room temperature unless otherwise stated. Cells were removed by centrifugation in a microcentrifuge at 15,000 rpm for 5 min. The bacteriophage-containing supernatant was transferred to a fresh microcentrifuge tube containing 200 μL of 20% polyethylene glycol (PEG 8,000) in 3.5 M ammonium acetate.
The solution was mixed by inverting the tube several times and incubated for 15 min. Bacteriophage particles were recovered by centrifugation at 15,000 rpm for 10 min and the supernatant was removed by aspiration through a "drawn-out" Pasteur pipette. The bacteriophage pellet was resuspended in 100 μL TE (pH 8.0) by vigorous vortexing and the protein coat was removed by phenol extraction using 50 μL water-saturated phenol. After centrifugation at 15,000 rpm for 1 min, the aqueous (upper) phase was transferred to a fresh microcentrifuge tube containing one-tenth volume of 3 M sodium acetate (pH 5.5), followed by 2.5 volumes of 95% ethanol. The solution was mixed by inverting and incubated for 15 min. The single-stranded bacteriophage DNA was recovered by centrifugation at 15,000 rpm for 10 min. The supernatant was removed and the pellet rinsed with 200 μL 70% (v/v) ice-cold ethanol and recentrifuged. The pellet was dried under vacuum, resuspended in 50 μL TE (pH 8.0) and stored at -20°C.

2.3.2.2. Preparation of Double-stranded E. coli Plasmid DNA

All procedures were performed at room temperature unless otherwise stated. Two mL 2x TY broth containing the appropriate antibiotic were inoculated with a single bacterial colony and incubated with shaking at 37°C for 6-18 h. Cells were harvested by centrifugation in a microcentrifuge at 15,000 rpm for 5 min and the supernatant was removed. The cell pellet was resuspended in 100 μL Solution I and incubated for 10 min. Cells were lysed by addition of 200 μL Solution II (E. coli). The suspension was mixed by inverting and incubated for 5 min. Bacterial debris was precipitated by addition of 150 μL Solution III. The suspension was mixed by inverting and incubated for 5 min. The debris was pelleted by centrifugation at 15,000 rpm for 5 min. The supernatant was transferred to a fresh tube containing 200 μL water-saturated phenol, 100 μL chloroform; mixed thoroughly and centrifuged at 15,000 rpm for 2 min. The aqueous (upper) phase was transferred to a fresh tube containing 1 mL 95% ethanol, inverted to mix, and incubated for 10 min. Total nucleic acids were pelleted by centrifugation at 15,000 rpm for 10 min and the supernatant was removed. The pellet was dried under vacuum,
resuspended in 50 μL TE (pH 8.0) containing 1 μL each of RNase A (Sigma, 10 mg/mL) and RNase T₁ (Boehringer Mannheim, 100 U/mL) and incubated 1 h in a 37°C water bath. The final DNA solution was stored at -20°C.

2.3.2.3. Preparation of Replicative Form M13

Bacteria infected with bacteriophage M13 contain the double-stranded replicative form of bacteriophage M13 DNA which can be isolated in a manner analogous to that of double-stranded plasmid DNA. Following the growth of two mL cultures of M13-infected E. coli TG2 (as given in Section 2.3.2.1.), cells were harvested and treated as described for the preparation of E. coli plasmid DNA (Section 2.3.2.2.).

2.3.2.4. Preparation of Bacillus Plasmid DNA

Two mL 2x TY broth containing 50 μg/mL kanamycin were inoculated with a single bacterial colony and incubated with shaking at 37°C for 6-18 h. Cells were harvested by centrifugation and the supernatant was removed. The cell pellet was resuspended in 100 μL Solution I and incubated in a 37°C water bath for 20-30 min. Cells were lysed by addition of 200 μL Solution II (Bacillus). The suspension was mixed by inverting and incubated on ice for 5 min. Bacterial debris was precipitated by addition of 150 μL Solution III. The suspension was mixed by inverting and kept on dry ice for 2 min. The debris was pelleted by centrifugation at 15,000 rpm for 20 min. The supernatant was transferred to a fresh tube containing 400 μL isopropanol, mixed and kept on dry ice for 5 min. Total nucleic acids were pelleted by centrifugation at 15,000 rpm for 20 min and the supernatant was removed. The pellet was dried under vacuum, resuspended in 100 μL TE (pH 8.0) containing 1 μL each of RNase A (Sigma, 10 mg/mL) and RNase T₁ (Boehringer Mannheim, 100 U/mL) and incubated 1 h in a 37°C water bath. The final DNA solution was stored at -20°C.
2.3.3. Cloning

2.3.3.1. Restriction Enzyme Digestion

Diagnostic digests were typically performed in a volume of 15 μL using approximately 100 ng of double-stranded DNA. DNA (typically 2-3 μL of a standard *E. coli* or *Bacillus* DNA preparation), 1.5 μL 10x digestion buffer (supplied by manufacturer) and sterile ddH₂O were combined to a final volume of 14 μL. Digestion was initiated with the addition of 2-4 units of restriction enzyme and the solution was incubated at the appropriate temperature for 1 h. The resulting fragments were analyzed by agarose gel electrophoresis. Test digests were scaled up for preparing DNA fragments for subsequent isolation and/or cloning. DNA (typically 15-25 μL), 10x digestion buffer and water were combined to a final volume of 50-100 μL. Restriction enzyme (5-10 units) was added and the mixture was incubated at the appropriate temperature for 2-4 h. Small aliquots were analyzed by agarose gel electrophoresis to confirm complete digestion. If necessary, additional enzyme was added and the mixture incubated further.

2.3.3.2. Blunt Ending

2.3.3.2.1. T4 DNA Polymerase

Treatment with T4 DNA polymerase was used to fill in 3' recessed ends or remove 3' overhangs of DNA fragments. After restriction enzyme digestion the DNA was phenol extracted and ethanol precipitated as follows. The DNA solution was mixed with an equal volume of 1:1 phenol:chloroform and centrifuged at 15,000 rpm for 2 min at room temperature. The aqueous (upper) layer was transferred to a fresh microcentrifuge tube containing an equal volume of chloroform. The tube was mixed thoroughly and centrifuged for 2 min. The aqueous layer was transferred to a fresh microcentrifuge tube containing one-tenth volume 3 M sodium acetate (pH 5.5) and 2.5 volumes 95% ethanol was added. The solution was incubated at room temperature
for 30 min and the DNA was recovered by centrifugation for 15 min. The supernatant was removed and the pellet was dried under vacuum. For removing 3’ overhangs, the DNA was resuspended in 89 μL ddH₂O and 10 μL 10x T4 DNA polymerase buffer. T4 DNA polymerase (New England Biolabs, 3 U/μL, 1 μL) was added and the solution incubated at 37°C for 10 min. For filling in 3’ recessed ends, 2 μL of 25 mM deoxynucleotide triphosphates were included in the above reaction.

2.3.3.2.2. Mung Bean Nuclease

Treatment with mung bean nuclease was used to remove 5’ overhangs of DNA fragments. After restriction enzyme digestion, phenol extraction and ethanol precipitation, the DNA (approximately 1 μg) was resuspended in:

- ddH₂O: 50 μL
- 150 mM sodium acetate (pH 4.6), 250 mM NaCl: 20 μL
- 5 mM ZnCl₂: 20 μL
- 50% (v/v) glycerol: 10 μL
- mung bean nuclease (Pharmacia, 169 U/μL): 5 units

The nuclease solution was mixed and incubated at 37°C for 10 min.

2.3.3.3. Dephosphorylation of Vector

Restriction enzyme-digested vectors were dephosphorylated by treatment with calf intestinal alkaline phosphatase (CIP). After digestion and/or subsequent blunt-ending, the DNA was phenol extracted and ethanol precipitated. The resulting DNA pellet was resuspended in 44 μL ddH₂O and 5 μL 10x dephosphorylation buffer. The dephosphorylation reaction was initiated with the addition of 1 μL of calf intestinal alkaline phosphatase (Boehringer Mannheim, 1 U/μL) and incubated for 60 min at 50°C. A second 1 μL aliquot of phosphatase was added and the reaction was incubated for a further 30-60 min.
2.3.3.4. Gel Purification of DNA Fragments

Restriction enzyme-digested DNA fragments used for cloning were isolated by agarose gel electrophoresis. DNA was electrophoresed on an agarose gel (1.0-1.5%) to separate fragments and the desired band(s) were identified using low energy ultraviolet illumination (366 nm) and excised from the gel. The DNA fragment of interest was purified by use of low melting point agarose, by electroelution, or by spin column purification.

2.3.3.4.1. Low Melting Point Agarose

DNA was electrophoresed on a low melting point agarose gel (1.0-1.5%) and desired band(s) were excised from the gel. The gel slice was centrifuged at 15,000 rpm for 2 min and the volume of the agarose was estimated (typically 50 µL). Five volumes of TE (pH 8.0) were added and the solution was incubated at 65°C for 5 min. When melted, the hot agarose solution was extracted once each with equal volumes of phenol, phenol:chloroform, and chloroform. The DNA fragment was ethanol precipitated and collected by centrifugation at 15,000 rpm for 30 min. The supernatant was removed and the pellet was dried under vacuum. The recovered DNA was resuspended in 6-10 µL ddH₂O and the yield was estimated by agarose gel electrophoresis of a 1 µL aliquot.

2.3.3.4.2. Electroelution

The agarose gel slice containing the DNA fragment of interest was placed in rinsed, prepared dialysis tubing (described below) with a small volume (200-400 µL) of TE (pH 8.0). The dialysis tubing (Spectrum Spectra/Por®, Mol. Wt. Cut-off 12,000-14,000) was immersed in 0.5x TAE and current was applied (constant voltage of 200 V) for 15 min. The current was reversed for 30 sec and the TE transferred to a microcentrifuge tube. The dialysis tubing and gel slice were rinsed with approximately 100 µL buffer and the buffer solutions pooled. The DNA was ethanol precipitated at room temperature overnight after addition of 1 µL polyacrylamide carrier (PAP).77 The
recovered DNA was resuspended in 6-10 µL ddH₂O and the yield was estimated by agarose gel electrophoresis of a 1 µL aliquot.

Dialysis tubing for electroelution was prepared by cutting the tubing into 3-4 cm lengths and boiling in a large volume of 2% sodium bicarbonate and 1 mM EDTA for 10 min. Once cooled, the tubing was rinsed thoroughly in distilled water and boiled in 1 mM EDTA and 0.1% SDS for 10 min. Once cooled, the prepared dialysis tubing was stored at 4°C in the same solution.

2.3.3.4.3. Spin Column Purification

The spin column for purifying DNA from an agarose gel was prepared by packing sterile silanized glass wool into the lower 2.5 cm of a P1000 tip, cut 3.0 cm from the end. The tip was placed in a 0.5 mL microcentrifuge tube (lid removed, hole pierced in the bottom) which was nested into a 1.5 mL microcentrifuge tube (lid removed).

Figure 2.1. Preparation of the spin column.

The agarose gel slice containing the DNA fragment of interest was placed inside the P1000 tip (on top of the glass wool) and the spin column was centrifuged at 6,000 rpm for 10 min. The recovered DNA solution (in the 1.5 mL microcentrifuge tube) was extracted once with an equal volume of phenol:chloroform and once with chloroform. The DNA was ethanol precipitated and resuspended in 6-10 µL ddH₂O. Yield was estimated by agarose gel electrophoresis of a 1 µL aliquot.
2.3.3.5. Ligation

DNA fragments were cloned by combining approximately 20 ng of vector with a suitable amount of insert (in a insert:vector molar ratio of 1-2:1) in a final volume of 3 μL. The ligation reaction was initiated with the addition of 1 μL 5x ligation buffer and 1 μL of T4 DNA ligase (BRL, 10 U/μL) and incubated for 2-16 h at 16°C or room temperature (≤22°C). Transformation of *E. coli* was routinely performed with one-half of the total ligation mixture. Transformation of *Bacillus* protoplasts was performed with the entire ligation mixture. Successful ligations were identified primarily by restriction enzyme digestion or, if necessary, by DNA sequencing of the resulting constructs.

2.3.3.6. DNA Sequencing

DNA templates were sequenced by the dideoxy chain-termination method using Sequenase Version 2.0 (United States Biochemical Corporation) or T7 DNA polymerase (Pharmacia).

2.3.3.6.1. *E. coli* Plasmid DNA Preparation

Double-stranded plasmid DNA from *E. coli* (Section 2.3.2.2.) was purified further prior to sequencing. Total DNA from the standard 2 mL-culture preparation (50 μL) was diluted to 120 μL with ddH2O and mixed with one-third volume 10 M ammonium acetate. The DNA was precipitated by addition of 2 volumes 95% ethanol and incubation at room temperature for 10 min followed by centrifugation at 15,000 rpm for 10 min. The pellet was resuspended in 90 μL ddH2O and the DNA was reprecipitated by addition of 10 μL 3 M sodium acetate (pH 5.5) and 150 μL 95% ethanol. After incubation at room temperature for 10 min followed by centrifugation for 10 min, the pellet was resuspended in a final volume of 10 μL ddH2O.

Double-stranded plasmid DNA was denatured for sequencing by addition of 1 μg of DNA template (in 1-4 μL) to 200 mM NaOH/0.2 mM EDTA for a final volume of 17 μL. The denaturation solution was incubated at room temperature for 3 min before addition of 3 μL 3 M sodium acetate (pH 5.5) and 50 μL 95% ethanol. After incubation
at room temperature for 5 min, the DNA solution was centrifuged at 15,000 rpm for 5 min. The pellet was dried under vacuum and used for the annealing reaction immediately.

2.3.3.6.2. Annealing Template and Primer

For each set of four sequencing lanes, a single annealing (and subsequent labeling) reaction was used. The denatured double-stranded plasmid DNA was resuspended in the following cocktail: 1 μL primer (0.5 pmole/mL), 2 μL 5x Sequenase buffer, and 7 μL ddH2O. The annealing solution was incubated in a 37°C water bath for 15 min, then at room temperature for 5 min.

When using single-stranded M13 DNA template, 1 μg DNA (in 7 μL) was combined with 1 μL sequencing primer (0.5 pmole/μL) and 2 μL 5x Sequenase buffer. The mixture was heated in a heating block for 2 min at 65°C after which the heating block was turned off and allowed to cool slowly to ≤30°C. The annealed template was placed on ice.

2.3.3.6.3. Labeling and Termination Reactions

Labeling and termination reactions were performed as described in the Sequenase Version 2.0 sequencing manual using materials and buffers provided. The resulting termination reactions were analyzed on a denaturing polyacrylamide sequencing gel.63

2.3.4. Site-directed Mutagenesis

2.3.4.1. Preparation of Single-stranded Uracil-containing Bacteriophage M13 DNA

Using a sterile toothpick, a single plaque produced by the desired bacteriophage M13 recombinant was transferred to a microcentrifuge tube containing 1 mL 2x TY broth. The tube was incubated at 60°C for 5 min to kill the host bacteria and vortexed vigorously for 30 sec to release bacteriophage trapped in the top agar. Debris was removed by centrifuging at 15,000 rpm for 2 min and the supernatant was transferred to a fresh tube. The bacteriophage suspension was used to infect E. coli RZ1032 by
combining 50 µL of the bacteriophage supernatant with 5 mL of a mid-log phase *E. coli* RZ1032 culture in a 500 mL flask containing 50 mL 2x TY broth supplemented with 0.25 µg/mL uridine. The culture was incubated with vigorous shaking for 6 h at 37°C. The resulting culture was chilled in an ice water bath and centrifuged at 5,500 rpm for 30 min at 4°C (Sorvall centrifuge, GSA rotor). The supernatant was transferred to a 250 mL centrifuge bottle and a small aliquot (500 µL) was reserved for titer determination (described below). The bacteriophage particles were precipitated by adding 0.25 volume of 20% polyethylene glycol (PEG 8,000) in 3.5 M ammonium acetate. The solution was mixed by swirling and incubated on ice for 1 h. The bacteriophage were recovered by centrifugation at 5,500 rpm for 20 min at 4°C (Sorvall centrifuge, GSA rotor) and all traces of supernatant were removed by aspiration. The bacteriophage pellet was resuspended in 4 mL TE (pH 8.0) and transferred to a 15 mL Corex centrifuge tube. The walls of the centrifuge bottle were washed with another 2 mL TE (pH 8.0) and transferred to the Corex tube. The suspension was vortexed vigorously for 30 sec and incubated on ice for 1 h. The suspension was vortexed vigorously for 30 sec and centrifuged at 6,500 rpm for 20 min at 4°C (Sorvall centrifuge, SS-34 rotor). The supernatant was transferred to a clean Corex tube and extracted twice with phenol and once with phenol:chloroform, separating the phases by centrifugation at 6,500 rpm for 5 min at room temperature. The final aqueous phase was mixed thoroughly in a glass centrifuge tube with 0.1 volume 3 M sodium acetate (pH 5.5), followed by 2 volumes of ice-cold 95% ethanol and incubated on ice for 30 min. The DNA was recovered by centrifugation at 6,500 rpm for 20 min at 4°C and washed with ice-cold 70% (v/v) ethanol. Once resuspended in 200 µL TE (pH 8.0), the DNA was quantitated spectrophotometrically at 260 nm (1 OD<sub>260</sub>=40 µg/mL) and analyzed by gel electrophoresis, using as marker single-stranded DNA of the original bacteriophage M13 recombinant.
To determine the relative titer of viable bacteriophage, dilutions of $10^{-4}$ to $10^{-10}$ of the bacteriophage suspension were prepared and 100 µL of the dilutions were plated using either *E. coli* TG2 or RZ1032. The titer on *E. coli* RZ1032 should be four to five orders of magnitude greater than on *E. coli* TG2 for successful selection after mutagenesis to occur.

2.3.4.2. Kinasing of Oligonucleotides

Oligonucleotides used in mutagenesis were phosphorylated using T4 polynucleotide kinase in the following reaction:

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>oligonucleotide</td>
<td>100-200 pmoles</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>to 16.5 µL</td>
</tr>
<tr>
<td>10x bacteriophage T4 polynucleotide kinase buffer</td>
<td>2 µL</td>
</tr>
<tr>
<td>10 mM ATP</td>
<td>1 µL</td>
</tr>
<tr>
<td>bacteriophage T4 polynucleotide kinase (BRL, 10 U/µL)</td>
<td>4 units</td>
</tr>
</tbody>
</table>

The reaction was incubated for 1 h at 37°C and the kinase was inactivated by incubation for 10 min at 68°C.

2.3.4.3. Site-specific Mutagenesis

The phosphorylated mutagenic oligonucleotide and phosphorylated universal sequencing primer were annealed to the single-stranded uracil-containing bacteriophage M13 DNA containing the target sequence in the following solution:

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>single-stranded template DNA (1 µg)</td>
<td>0.5 pmole</td>
</tr>
<tr>
<td>phosphorylated mutagenic oligonucleotide</td>
<td>10 pmole</td>
</tr>
<tr>
<td>phosphorylated universal sequencing primer</td>
<td>10 pmole</td>
</tr>
<tr>
<td>10x PE1 buffer</td>
<td>1 µL</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>to 10 µL</td>
</tr>
</tbody>
</table>

The mixture was heated in a heating block for 5 min to 20°C above the theoretical $T_m$ of a perfect hybrid formed by the mutagenic oligonucleotide, calculated from the formula $T_m=4(G+C) + 2(A+T)$, after which the heating block was turned off and allowed to cool.
slowly to ≤30°C. While the reaction cooled, the PE3 mixture was prepared as follows and stored on ice until needed:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PE2</td>
<td>1.0 µL</td>
</tr>
<tr>
<td>2 mM deoxynucleotide triphosphates</td>
<td>1.0 µL</td>
</tr>
<tr>
<td>10 mM ATP</td>
<td>1.0 µL</td>
</tr>
<tr>
<td>ddH$_2$O</td>
<td>to 8.0 µL</td>
</tr>
<tr>
<td>bacteriophage T4 DNA ligase (BRL, 10 U/µL)</td>
<td>1.0 µL</td>
</tr>
<tr>
<td>DNA polymerase I-Klenow fragment (New England Biolabs, 5 U/µL)</td>
<td>2.5 units</td>
</tr>
</tbody>
</table>

The tube containing the annealing reaction was centrifuged briefly and 10 µL of the ice-cold PE3 mixture was added. The extension-ligation reaction was incubated for 6-15 h (overnight) at 16°C.

2.3.4.4. Transformation of *E. coli* DH5α™

The extension-ligation reaction was used to transform competent *E. coli* DH5α™ as described for the transformation of *E. coli* TG2 by bacteriophage M13 (Section 2.2.4.3.) with the following modifications. Aliquots of 200 µL competent *E. coli* DH5α™ were combined with 1 µL and 5 µL aliquots of undiluted extension-ligation reaction and the entire transformation mix was plated using *E. coli* TG2.

Confirmation of mutagenesis was achieved by DNA sequencing of single-stranded DNA isolated from the resulting bacteriophage.

2.3.5. Northern Blots

2.3.5.1. Isolation of RNA from *Bacillus*$_{^{80}}$

Cells from a 5 mL bacterial culture were harvested by centrifugation at 10,000 rpm for 10 min at 4°C (Sorvall centrifuge, SS-34 rotor). The cell pellet was resuspended in 0.5 mL lysis buffer and the suspension was transferred to a microcentrifuge tube placed on dry ice. Once frozen, the suspension was thawed and sonicated on ice with a microtip sonicator (Fisher Sonic Dismembrator, Model 300) three times for 15 sec with 20-30 sec intervals on ice. Proteinase K (Sigma) was added to a
final concentration of 100 µg/mL and the solution was incubated for 60 min in a 37°C water bath. The suspension was then extracted twice with an equal volume of 1:1 phenol/chloroform and once with an equal volume of chloroform. The aqueous phase was transferred to a fresh tube containing 15 µL 5 M NaCl and the tube was filled with ice-cold 95% ethanol. The mixture was incubated 15-30 min on ice or overnight at -20°C and the precipitated nucleic acids were collected by centrifugation at 15,000 rpm for 15 min at 4°C. The pellet was rinsed with 500 µL ice-cold 70% (v/v) ethanol and resuspended in 96 µL DNase digestion buffer to which 4 µL RQ1 RNase-free DNase (Promega, 1 U/µL) was added. After incubation for 60 min in a 37°C water bath, the RNA solution was extracted once with 1:1 phenol/chloroform at room temperature and the organic phase was extracted once with 100 µL TE (pH 8.0). The pooled aqueous phase was extracted once with chloroform and the RNA was precipitated by the addition of 10 µL 5 M NaCl and 600 µL ice-cold 95% ethanol. After incubation overnight at -20°C or 15 min on dry ice, the precipitate was collected by centrifugation at 15,000 rpm for 15–30 min at 4°C. The pellet was rinsed with 500 µL ice-cold 70% (v/v) ethanol, dried under vacuum and resuspended in 100 µL ddH2O. Absorbance at 260 nm was used to quantify RNA yield. (An $A_{260}$ of 1.0 = 40 µg/mL single-stranded RNA.)

2.3.5.2. Denaturing Formaldehyde Agarose Gel

A denaturing formaldehyde agarose gel (1.2%) was prepared by boiling a solution of 21 mL autoclaved ddH2O, 2.5 mL 10x MOPS buffer and 0.3 g agarose, allowing it to cool to 55°C and adding 1.1 mL deionized formaldehyde. The solution was mixed by swirling and poured into a mini-slab apparatus in the fumehood. RNA samples were prepared by mixing 10 µL RNA and/or sterile ddH2O with 5 µL formaldehyde loading buffer and heating at 60°C for 15 min. Ethidium bromide (0.5 µL of 1 mg/mL stock) was added prior to loading samples on the formaldehyde gel. The gel was electrophoresed in the fumehood using 1x MOPS as the running buffer at a constant voltage of approximately 5 V/cm until the bromophenol blue dye band had migrated two-thirds of
the way down the gel. The gel was photographed using ultraviolet illumination using a plastic ruler as an external reference.

2.3.5.3. Transfer of RNA

After electrophoresis the formaldehyde gel was soaked in water for 10 min followed by 10x SSC for 20 min. Whatman #1 filter paper was cut into two wicks (2 cm wider than the gel and approximately 30 cm long) and four squares (the size of the gel), and soaked in 10x SSC. Zeta-Probe (Bio-rad) cut to the size of the gel was wet carefully in 10x SSC. The transfer apparatus was arranged as shown in Figure 2.2.

Figure 2.2. Preparation of the transfer apparatus.

(Note: Parafilm or Saran Wrap was placed on the wick up against the gel to block fluid transfer directly from the wick to the paper towel.). The transfer was allowed to proceed overnight after which the apparatus was dismantled and the position of the agarose gel wells marked on the membrane in pencil. The membrane was washed briefly in 10x SSC and the RNA was UV-crosslinked to the membrane (Stratagene UV Stratalinker 2400). After fixing the RNA, the membrane was soaked in 1% SDS for 15 min.
2.3.5.4. Preparation of Radioactive DNA Probe

The radioactive DNA probe was prepared with the Random Priming System I (New England Biolabs). Gel-purified DNA fragment (25 ng in a volume of 34 µL) was denatured by boiling in a water bath for 5 min and then placing on ice for 5 min. The tube containing DNA fragment was centrifuged briefly in the cold and the following reagents were added in the order listed:

- 10x labeling buffer (includes random hexadeoxyribonucleotides) 5 µL
- dNTP mixture (.167 mM each of dGTP, dCTP, and dTTP) 5 µL
- α-32P-dATP (3,000 Ci/mmol, 50 µCi) 5 µL
- DNA polymerase I-Klenow fragment (5 U/µL) 1 µL

The priming reaction was incubated at 37°C for 1 h and was terminated by addition of 50 µL TE (pH 8.0). The synthesized probes were separated from unincorporated nucleotides by gel filtration on a Sephadex G-25 spin column.

The Sephadex spin column was prepared by removing the needle and plunger of a 1-cc syringe and packing a few mm of silanized glass wool in the bottom of the syringe. Pre-swollen Sephadex G-25 was added to the syringe with a Pasteur pipette, being careful not to introduce any bubbles. The aqueous solution was allowed to drain from the syringe and more Sephadex was added until the Sephadex formed a hump at the top of the syringe. The syringe was placed in a plastic test tube and centrifuged at approximately 225x g for 2 min to drain the buffer. The Sephadex was washed twice by adding 100 µL TE (pH 8.0) to the top of the column and recentrifuging. The radioactive probe sample (100 µL) was added slowly to the center of the packed column and the syringe was placed in a plastic test tube containing a 1.5 mL microcentrifuge tube (lid removed) and centrifuged. The purified probe was recovered from the microcentrifuge tube and denatured by boiling in a water bath for 5 min and then placing on ice for 5 min. The tube containing the DNA probe was centrifuged briefly in the cold and the probe was used directly.
2.3.5.5. Hybridization

The membrane was prehybridized in 10 mL hybridization buffer at 42°C for 1 h. The purified radioactive DNA probe (30-50 μL) was added and allowed to hybridize at 42°C for 16 h. The temperature was increased to 55°C and the membrane was washed twice for 20 min with 10-15 mL of 2x SSC/0.3% SDS, twice for 20 min with 10-15 mL of 1x SSC/0.5% SDS and twice for 20 min with 10-15 mL of 0.3x SSC/1.0% SDS. The washed membrane was sealed in plastic and exposed to Kodak XAR film. The probes were stripped from the membrane by incubating the membrane in 200 mL 1% SDS at 95°C for 3 h with two changes of SDS solution.

2.4. Protein Expression

2.4.1. Milk Plates

Extracellular protease activity was detected by observing zones of clearing surrounding bacteria plated on TY-agar plates containing 1.5% Carnation skim milk powder. Bacillus strains were streaked onto milk plates containing 50 μg/mL kanamycin and grown for 18-24 h at 37°C. E. coli strains harboring pMAL-p constructs were streaked onto milk plates containing 100 μg/mL ampicillin and 0.1 mM IPTG and grown for 48 h at room temperature.

2.4.2. Activity Assay

Production of extracellular protease in B. subtilis was also detected by growing bacteria harboring plasmid derivatives in liquid culture media. A volume of 100 mL of 2x TY supplemented with 50 μg/mL kanamycin was inoculated with 1 mL of a mid-log B. subtilis DB104 and samples were removed to monitor growth (optical density at 600 nm) and protease activity in the culture supernatant as the culture was incubated with shaking at 30°C. Protease activity in aqueous samples was determined by cleavage of the substrate, N-succinyl-al a-ala-pro-phe-p-nitroanilide, and measurement of released chromophore, p-nitroaniline. Samples of protease (up to 40 μL) were mixed with 1 mL
of prepared substrate (1 mg/mL in 50 mM Tris·Cl, pH 8.0) and the increase in absorbance at 412 nm was recorded per unit time at 23°C.

2.4.3. Bacillus Expression and Fractionation

Cultures (12.5 mL) were treated with 1 mM phenylmethylsulfonyl fluoride (PMSF) and the cells were harvested by centrifugation at 10,000 rpm for 10 min at 4°C (Sorvall centrifuge, SS-34 rotor). The cell pellet was resuspended with 1 mL 10 mM Tris·Cl (pH 7.4) and transferred to a microcentrifuge tube. The suspension was treated with PMSF, and recentrifuged. The cells were resuspended in 100 μL 30 mM Tris·Cl (pH 8.0) and treated with T4 lysozyme (20 μg/mL) at 37°C for 20 min. After centrifugation at 40,000× g (Beckman TL-100 Ultracentrifuge) for 20 min, the supernatant (cytosolic fraction) was saved. The pellet (crude membrane fraction) was resuspended in 100 μL 50 mM sodium phosphate (pH 6.6) and treated with 5 mM EDTA. The crude membrane fraction was treated with DNase (10 μg/mL)/10 mM MgCl₂, and incubated at 37°C for 10 min. The membranes were recovered by centrifugation at 40,000× g for 20 min, washed once in 100 μL 50 mM sodium phosphate buffer, and resuspended in 100 μL 50 mM sodium phosphate buffer.

Certain cell pellets were treated with exogenous protease in vitro prior to analysis. Cells from 12.5 mL of culture were resuspended in 1.0 mL of osmotically supported Tris buffer (30 mM Tris·Cl, 25% sucrose, pH 8.0) and treated with exogenous protease. After overnight incubation at room temperature with stirring, the cells and medium were separated by centrifugation and fractionation as described above.

The pH of culture supernatant samples was adjusted to approximately pH 5.8 with acetic acid and diluted with an equal volume of 100 mM sodium acetate (pH 5.8). To a minimum of 4 mL solution, ice-cold acetone was added to 40% (v/v, assuming additive volumes). The solution was mixed and incubated on ice for 1 h. The 40% acetone fraction was collected by centrifugation at 10,000 rpm for 20 min at 4°C (Sorvall centrifuge, SS-34 rotor). The supernatant was transferred to a fresh centrifuge tube and
0.5 volume of ice-cold acetone was added. The solution was mixed and incubated on ice for 1 h. The 40-60% acetone fraction was collected by centrifugation. Both pellets were dried under vacuum and resuspended in 0.5 mL acetate buffer. The protein suspension was transferred to a microcentrifuge tube and precipitated using 50 μL 100% trichloroacetic acid (TCA) and incubating 30 min on ice. The TCA precipitate was collected by centrifugation at 15,000 rpm for 15 min at room temperature. The supernatant was decanted and the pellet washed by adding 0.5 mL acetone and vortexing. The precipitate was pelleted by centrifugation for 5 min at room temperature. The supernatant was discarded and the TCA pellet was resuspended in acetate buffer.

2.4.4. E. coli Expression and Fractionation

2.4.4.1. E. coli (pMAL-p/c) Growth and Induction

Two mL 2x TY broth containing 100 μg/mL ampicillin were inoculated with a fresh colony of E. coli and incubated with shaking at 30°C to mid-log phase (OD_{600}=0.6). The culture was divided into two aliquots and IPTG was added to one sample to a final concentration of 0.3 mM. Both samples were incubated with shaking at 37°C for a further 2 h.

2.4.4.2. E. coli (pLcII-FX) Growth and Induction

Two mL 2x TY broth containing 100 μg/mL ampicillin were inoculated with a fresh colony of E. coli and incubated with shaking at 30°C to mid-log phase (OD_{600}=0.6). The culture was divided into two aliquots and one sample was returned to the 30°C incubator. The second sample was incubated at 42°C for 5 min and placed in a 37°C incubator. Both samples were incubated with shaking for a further 2 h.

2.4.4.3. Fractionation

Cells were harvested by centrifugation at 15,000 rpm for 5 min. The supernatant was removed and the cell pellet was resuspended in 100 μL Solution I (without lysozyme). The suspension was subjected to three cycles of freeze/thawing followed by
bath sonication for 10 min. Crude pellet was separated from the cytosolic fraction by centrifugation for 5 min.

2.4.5. SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

2.4.3.1. SDS-Polyacrylamide Gel

SDS-polyacrylamide gels were cast and electrophoresed using the Bio-rad Mini-PROTEAN II Cell apparatus. The separating gel was prepared by mixing 5.0 mL separating buffer with 5.0 mL acrylamide stock and degassing by aspiration for 15 min. Polymerization was initiated by addition of 50 µL of freshly prepared 10% ammonium persulfate and 5 µL N,N,N',N'-tetramethylethylenediamine (TEMED). The gel was cast in the apparatus (leaving approximately 5 mm between the top of the separating gel and the bottom of the comb) and 50% ethanol was layered over the gel solution. After allowing 45 min for complete polymerization, the top of the separating gel was rinsed with water and the apparatus dried with Whatman #1 filter paper. The stacking gel was prepared by mixing 7.0 mL stacking buffer with 1.6 mL acrylamide stock and degassing by aspiration for 15 min. Polymerization was initiated by addition of 37.5 µL 10% ammonium persulfate and 7.5 µL TEMED. The stacking gel was layered over the separating gel and the comb was inserted. The gel was allowed to polymerize for a minimum of 30 min before removing the comb and rinsing the wells first with water and then with SDS-PAGE running buffer. The Mini-PROTEAN II Cell apparatus was assembled and the samples were electrophoresed at 200 V (constant voltage) until the bromophenol blue dye band ran off the gel. The apparatus was disassembled and the gel stained in staining solution for 10 min followed by destaining in fast destaining solution for 10 min. The gel was destained in slow destaining solution until all bands were visible (usually 2-16 h).

Samples were prepared by adding one-fifth volume 5x SDS loading buffer and heating at 95°C for 3 min. Prestained SDS-PAGE standards (Bio-rad, Low Range) were used as molecular weight markers.
2.4.6. Western Blots

2.4.6.1. Transfer

After gel electrophoresis, the proteins were transferred to Immobilon-P transfer membrane (Millipore) using the LKB 2117 Multiphor II electrophoresis apparatus. For each gel, four pieces of 3M Whatman paper (cut to the size of the gel) were soaked in transfer buffer and one piece of Immobilon-P membrane was immersed briefly in methanol, rinsed with water and soaked in transfer buffer. The SDS-PAGE apparatus was disassembled and the gel rinsed with transfer buffer. The transfer apparatus was assembled by wetting both the anodal and cathodal plates with transfer buffer and stacking two sheets of Whatman paper, the membrane, the gel, and two sheets of Whatman paper on the cathodal plate. Each layer was carefully laid using sufficient quantities of buffer to avoid trapping air bubbles. The transfer stack was covered with the anodal plate and the apparatus was assembled. The blot was electrophoresed for 2 h at a constant current of $1.2 \times [\text{total area (cm}^2\text{) of the gel(s)}] \text{ mA}$. Typically a current of 56 mA was used for one gel and a current of 112 mA was used for two gels.

2.4.6.2. Immunoassay

After electrophoresis, the transfer apparatus was disassembled and the gel stained to verify protein transfer. Unoccupied sites on the membrane were blocked in 30 mL of blocking solution ('blotto') for 2 h at room temperature on a rotatory shaker. The membrane was then incubated with the appropriate primary antibody diluted in a small volume of blotto for 2 h at room temperature with gentle shaking. After washing the membrane three times (10 min each) in 30 mL blotto, the membrane was incubated with the appropriate horseradish peroxidase-coupled secondary antibody for 2 h at room temperature with gentle shaking. The membrane was then washed twice (5 min) in blotto, twice (5 min) in 1x TBS with 0.05% Tween 20, and twice (10 min) in 1x TBS. To detect the target protein, the membrane was carefully immersed in a fresh solution containing 6 mg 3,3'-diaminobenzidine in 9 mL 50 mM Tris-Cl (pH 7.4), 1 mL
0.03% NiCl₂, 10 μL H₂O₂. When the bands were of the desired intensity, the reaction was quenched by washing the membrane in water.
3. Results

3.1. Expression Plasmids

Numerous plasmids were constructed in attempts to express *S. griseus* proteases A and B (SGPA and SGPB) and their variants. The first series of plasmids were constructed in order to express the mature portion of SGPB from the pre-pro-portion of subtilisin BPN' in *B. subtilis*. The second series of plasmids were constructed in order to express the *S. griseus* genes using their own pre-pro-region in *B. subtilis*. The third series involved expressing the pro-mature proteases coupled to the pre-region of subtilisin BPN' in *B. subtilis* as well as fused to the maltose binding protein (pMAL plasmids) or to a portion of the cII protein (pLcII-FX plasmids) in *E. coli*. The final series of plasmids were constructed in order to express the SGPB inactive variant. The various plasmids are summarized in Table 3.1. and a thorough description of the construction of each is found in the Appendix.

3.2. Mature SGPB

The plasmid pMSi, as provided by Cangene Corporation (Mississauga, Ontario), contained the sequences for mature SGPB\(^1\) missing the first 21 nucleotides. Without prior knowledge of the exact function of the SGPB pro-region and the possibility of it having only a general involvement in the secretion process, initial attempts at expression of SGPB in *B. subtilis* consisted of using oligonucleotides to reconstruct the missing segment and to link the coding sequence for the mature region of SGPB to the coding region for the subtilisin BPN' pre-pro-region. The resulting construct would utilize the subtilisin BPN' pre-region as a genus specific signal peptide to transport the fusion protein to the cell surface.

Because subtilisin BPN' is responsible for cleavage of its pro-mature bond during secretion,\(^3\) it was thought that due to its P\(_1\) specificity this enzyme might still recognize the pro-subtilisin BPN'-mature-SGPB junction. Thus, subtilisin BPN' could be used to
Table 3.1. SGPA and SGPB Expression Plasmids.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Vector (Orientation)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Promoter</th>
<th>Pre-Region&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Pro-Region</th>
<th>Mature Protease</th>
<th>Description (Pages)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacillus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pPT30</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>53-54, 106</td>
</tr>
<tr>
<td>pKN1</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGPB</td>
<td>53-54, 108</td>
</tr>
<tr>
<td>pDC1</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGPB/subtilisin</td>
<td>53-55, 108</td>
</tr>
<tr>
<td>pPK4a</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGPB/subtilisin</td>
<td>53-55, 109</td>
</tr>
<tr>
<td>pPA1</td>
<td>pUB110 (+)</td>
<td>SGBP</td>
<td>SGBP</td>
<td>SGBP</td>
<td>56-57, 111</td>
<td></td>
</tr>
<tr>
<td>pPA2</td>
<td>pUB110 (-)</td>
<td>SGBP</td>
<td>SGBP</td>
<td>SGBP</td>
<td>56-57, 111</td>
<td></td>
</tr>
<tr>
<td>pPB1</td>
<td>pUB110 (+)</td>
<td>SGBP</td>
<td>SGBP</td>
<td>SGBP</td>
<td>56-57, 112</td>
<td></td>
</tr>
<tr>
<td>pPB2</td>
<td>pUB110 (-)</td>
<td>SGBP</td>
<td>SGBP</td>
<td>SGBP</td>
<td>56-57, 112</td>
<td></td>
</tr>
<tr>
<td>pbDC8-</td>
<td>pUB110 (-)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>70-72, 116</td>
<td></td>
</tr>
<tr>
<td>pbDC12+</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>57-59, 117</td>
<td></td>
</tr>
<tr>
<td>pbDC12-</td>
<td>pUB110 (-)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>57-59, 117</td>
<td></td>
</tr>
<tr>
<td>pbDC13+</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>57-59, 117</td>
<td></td>
</tr>
<tr>
<td>pbDC13-</td>
<td>pUB110 (-)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>57-59, 117</td>
<td></td>
</tr>
<tr>
<td>pbDC14+</td>
<td>pUB110 (+)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>62-64, 124</td>
<td></td>
</tr>
<tr>
<td>pbDC14-</td>
<td>pUB110 (-)</td>
<td>subtilisin</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>62-64, 124</td>
<td></td>
</tr>
<tr>
<td>pEB-B8</td>
<td>pUB110 (-)/pUC18</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>SGBP</td>
<td>58-59</td>
<td></td>
</tr>
<tr>
<td>pEB-HDG</td>
<td>pUB110 (-)/pUC18</td>
<td>subtilisin</td>
<td>SGBP</td>
<td>SGBP</td>
<td>63-64</td>
<td></td>
</tr>
<tr>
<td>E. coli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pMSi</td>
<td>pUC19</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>SGBP</td>
<td>51, 53, 106</td>
</tr>
<tr>
<td>pLcII-FX-pro-SGBP</td>
<td>pLcII-FX</td>
<td>PL</td>
<td>cII-FX</td>
<td>SGBP</td>
<td>62, 120</td>
<td></td>
</tr>
<tr>
<td>pLcII-FX-pro-SGBP/SG195</td>
<td>pLcII-FX</td>
<td>PL</td>
<td>cII-FX</td>
<td>SGBP</td>
<td>62, 126</td>
<td></td>
</tr>
<tr>
<td>pMAL-cA</td>
<td>pMAL-pA</td>
<td>Ptac</td>
<td>pMAL</td>
<td>SGBP</td>
<td>60-61, 118</td>
<td></td>
</tr>
<tr>
<td>pMAL-pA</td>
<td>pMAL-cB</td>
<td>Ptac</td>
<td>pMAL</td>
<td>SGBP</td>
<td>60-61, 118</td>
<td></td>
</tr>
<tr>
<td>pMAL-pB</td>
<td>pMAL-pB</td>
<td>Ptac</td>
<td>pMAL</td>
<td>SGBP</td>
<td>60-61, 118</td>
<td></td>
</tr>
<tr>
<td>pMAL-pB/SG195</td>
<td>pMAL-pB</td>
<td>Ptac</td>
<td>pMAL</td>
<td>SGBP</td>
<td>62, 125</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> orientation of the fragment in the Bacillus vector pUB110 where (+) indicates a clockwise direction and (-) indicates a counter-clockwise orientation.

<sup>b</sup> for E. coli constructs, the 'pre-region' refers to the expression plasmid-encoded fusion sequences.

catalyze the release of mature SGBP (or SGPB mutants) in cultures expressing the pre-pro-subtilisin BPN'-pro-mature-SGBP fusion construct. To facilitate the cleavage of the fusion protein an additional amino acid (glycine, alanine or valine) was added to the amino
terminus of mature SGPB due to the $P_1'$ specificity of subtilisin BPN' for small hydrophobic residues.\textsuperscript{81,82} The clone that was chosen for further study contained an alanine which is also the native $P_1'$ amino acid for self-maturation in pre-pro-mature subtilisin BPN'.

The coding region for the reconstructed mature portion of SGPB was placed behind the pre-pro-region of subtilisin BPN' in pPT30. Due to the method of construction two plasmids resulted: pKN1, containing the entire coding sequence for the mature region of SGPB, including the 3'-non-coding region; and pDC1, containing the coding sequence for the mature region of SGPB missing the last 18 nucleotides of the coding sequence. In pDC1 the sequence of mature SGPB was in-frame with the mature subtilisin BPN' sequence and allowed for the translation of mature subtilisin BPN' following the mature SGPB (Figure 3.1.).

![Figure 3.1. Comparison of plasmids pMSi, pPT30, pKN1, pDC1, and pGK4a.](attachment:image.png)

Production of extracellular protease was detected by growing bacteria harboring plasmid derivatives on agar plates containing 1.5\% Carnation skim milk powder.\textsuperscript{28} Secretion of active protease created a clear halo surrounding the bacterial colony on an otherwise opaque agar plate due to the proteolysis of insoluble milk proteins. The extent of
the clearing varied with the bacterial strain, the plasmid construct and the temperature of incubation.

Figure 3.2. shows milk plates on which *B. subtilis* DB104 cultures harboring the plasmids PUB110, pPT30, pKN1, pDC1, and pGK4a were grown. *B. subtilis* DB104/pUB110 was used as a negative control for all experiments using *B. subtilis*. This strain, containing the plasmid pUB110 without any insert, had a small zone of clearing as a result of a low amount of endogenous protease activity and cell lysis. *B. subtilis* DB104/pPT30 secretes subtilisin BPN', a serine protease with broad proteolytic specificity, and demonstrated a large zone of clearing. *B. subtilis* DB104/pKN1, containing the mature portion of SGPB linked to the pre-pro-region of subtilisin BPN', showed no active protease secretion. The activity that resulted from *B. subtilis* DB104/pDC1 was shown by examination of the culture supernatant by SDS-polyacrylamide gel electrophoresis to be due to secretion of subtilisin BPN' with no indication of polypeptides corresponding to mature SGPB.

![Figure 3.2](Image)

Figure 3.2. *B. subtilis* DB104 cultures harboring the plasmids PUB110, pPT30, pKN1, pDC1, and pGK4a grown on milk plates for 48 h at (A) 30°C and (B) 37°C.
To rule out the possibility that the last six amino acids of mature SGPB were necessary for stability, oligonucleotide linkers were used to reconstruct the last 18 nucleotides of the coding sequence for the mature portion of SGPB in pDC1. The resulting plasmid, pGK4a, contained the entire coding sequence for the mature region of SGPB inserted in frame after the pre-pro region of subtilisin BPN' (Figure 3.1.). Again, the activity was due to the secretion of subtilisin BPN' (Figure 3.2.).

In studies with two other proteases, subtilisin E from *B. subtilis* and α-lytic protease of *Lysobacter enzymogenes*, it was revealed that there is an absolute requirement of the protease-specific pro-region for production of active protease. However, from the above results, it was found that the pro-region of subtilisin BPN' need not be contiguous with its mature region on the same polypeptide to produce active protease. In the construct pGK4a, the pro and mature regions of subtilisin BPN' are separated by the 185 amino acids of the mature SGPB.

### 3.3. Pre-Pro-Mature Protease

If there is a common mechanism of folding and activation among extracellular proteases in different bacterial species and the protease-specific pro-region is required for the production of active SGPB, it was then necessary to express SGPB from constructs containing its own pro-region. Further attempts at production of SGPB and the related *S. griseus* protease A (SGPA) in *B. subtilis* were made by expressing the two proteases under the control of their own promoters and with their own pre-pro-regions. This was accomplished by using the polymerase chain reaction (PCR) to isolate the necessary gene sequences from *S. griseus* IMRU3499 (work done in this lab by Sachdev Sidhu with assistance from Kevin Neden). The pre-pro-mature gene sequences were provided as *E. coli* clones.

For expression in *B. subtilis*, the coding regions for SGPA and SGPB were ligated into the unique *PvuII* site of the *Bacillus* plasmid pUB110. Due to the nature of the subcloning (blunt-ended ligation), the genes were oriented in either direction with
respect to the plasmid pUB110. Constructs pPA1 and pPB1 contain the SGPA and SGPB fragments in a counterclockwise orientation relative to pUB110 (as it is conventionally drawn with the plasmid encoded genes transcribed in a counterclockwise direction, Figure 3.3.) and constructs pPA2 and pPB2 contain the fragments in a clockwise orientation (Figure 3.4.).

![Plasmid diagram](image)

**Figure 3.3.** *Racillus* plasmid pUB110.

**Figure 3.4.** Comparison of plasmids pPA1, pPA2, pPB1, and pPB2.

Figure 3.5. shows milk plates on which *B. subtilis* DB104 cultures harboring the plasmids pUB110, pPA1, pPA2, pPB1, and pPB2 were grown. Those strains harboring the constructs containing pre-pro-mature-SGPA (pPA1 and pPA2) showed no activity on milk plates compared to the strain harboring the construct pPB1 containing pre-pro-mature-SGBP. Comparing strains harboring the SGBP constructs, the level of activity depended on the orientation of the insert in the plasmid. *B. subtilis* DB104/pPB1 resulted in secretion of protease activity and contained the plasmid pPB1 with the SGBP fragment in a counterclockwise orientation relative to pUB110 (as it is conventionally drawn) and in the same orientation as all four open reading frames on pUB110.
3.4. Pro-Mature Protease

3.4.1. *Bacillus* Expression

The construct pPB1 demonstrated that the *S. griseus* protease SGPB could be expressed successfully from its endogenous promoter and pre-region and that the pro-region is necessary for the production of active, mature SGPB. To produce greater quantities of protease the pre-pro-mature constructs were modified to use a genus specific signal that already demonstrated production of great quantities of protein in the form of subtilisin BPN'. This was accomplished by using the PCR to isolate the necessary fragments; pre-subtilisin BPN', pro-mature SGPA, and pro-mature SGPB. To be expressed in *B. subtilis*, the fragments containing pre-subtilisin BPN' and the pro-mature-proteases were ligated into pUB110 to give the constructs pbDC12+, pbDC12-, pbDC13+, and pbDC13-. pbDC12+ and pbDC13+ contain the pre-subtilisin BPN'-pro-mature-protease fragments in a clockwise orientation relative to
pUB110 and pbDC12- and pbDC13- contain the pre-subtilisin BPN'-pro-mature-protease fragments in a counterclockwise orientation (Figure 3.6.).

![Comparison of plasmids pbDC12+, pbDC12-, pbDC13+, and pbDC13-](image)

Figure 3.6.  Comparison of plasmids pbDC12+, pbDC12-, pbDC13+, and pbDC13-.

Figure 3.7. shows milk plates on which *B. subtilis* DB104 cultures harboring the plasmids pUB110, pbDC12+, pbDC12-, pbDC13+, and pbDC13- were grown. Also shown is *B. subtilis* DB104/pEB-B8 harboring a similar construct for SGPB (pre-subtilisin BPN'-pro-mature-SGPB) contained on an *E. coli*/*B. subtilis* shuttle vector constructed by Sachdev Sidhu of this lab (Figure 3.8.).

![B. subtilis DB104 cultures harboring the plasmids pUB110, pbDC12+, pbDC12-, pbDC13+, and pbDC13- grown on milk plates for 48 h at (A) 30°C and (B) 37°C.](image)

Figure 3.7.  *B. subtilis* DB104 cultures harboring the plasmids pUB110, pbDC12+, pbDC12-, pbDC13+, and pbDC13- grown on milk plates for 48 h at (A) 30°C and (B) 37°C.
B. subtilis harboring constructs containing pre-subtilisin BPN'-pro-mature-SGPA (pbDC12+ and pbDC12-) showed less activity than those harboring the constructs containing pre-subtilisin BPN'-pro-mature-SGPB (pbDC13+ and pbDC13-). For both proteases, those constructs that contain the fusion gene in a counterclockwise orientation relative to pUB110 and in the same orientation as all four open reading frames on pUB110 (pbDC12- and pbDC13-) produced greater amounts of protease activity. Similar amounts of active protease resulted from the SGPB constructs in both the Bacillus-specific plasmid (pbDC13-) and the E. coli/Bacillus shuttle vector (pEB-B8). Differences in secreted activity were also observed when the bacteria were grown at different temperatures. As a general rule, those cultures grown at lower temperature secreted more activity, which was especially noticeable with the SGPA-containing constructs.

Northern blots were performed to determine whether this orientational dependence on expression was due to RNA transcriptional read through from promoters located on the plasmid. For both pbDC12- and pbDC13-, bands corresponding to the full length pre-subtilisin BPN'-pro-mature-S. griseus protease were evident as well as a second band approximately 285 nucleotides longer, indicating transcription originating from an upstream promoter, namely from the mob gene. Thus, increases in the levels of protease activity are due in part to transcriptional read through of the insert from a plasmid promoter.
3.4.2. *E. coli* Expression

3.4.2.1. pMAL Constructs

For expression in *E. coli* the coding sequences for the pro-mature regions of SGPA and SGPB were ligated behind the maltose binding protein (MBP) sequence on each of the plasmids pMAL-c and pMAL-p (Figure 3.9.71,72) When induced, bacteria harboring pMAL derivatives express fusion protein consisting of the MBP linked to the pro-mature protease by the tetrapeptide recognition site for Factor Xa. pMAL-p derivatives directed the fusion protein to the periplasm of the cell due to the presence of the MBP signal peptide sequence. pMAL-c derivatives which lack the leader sequence produced a cytoplasmic protein.

![Diagram of pMAL Constructs](image)

Figure 3.9. Comparison of plasmids pMAL-pA, and pMAL-pB.

Secretion of active protease from the *E. coli* constructs was also tested by growth on milk plates. Figure 3.10. shows milk plates on which *E. coli* TB1 cultures harboring the plasmids pMAL-p, pMAL-c, pMAL-pA, pMAL-pB, pMAL-cA, and pMAL-cB were grown. *E. coli* TB1 harboring pMAL-c derivatives did not demonstrate any extracellular protease activity due to the nature of the expression plasmid. In these constructs, the fusion proteins were directed to the bacterial cytoplasm. The *E. coli* TB1 strain directing expression of MBP-pro-mature-SGPA to the periplasm (pMAL-pA) demonstrated less activity than the strain expressing a similar construct for MBP-pro-mature-SGPB (pMAL-pB). Differences in secreted activity were also observed when the bacteria were grown at different temperatures. At 37°C, no activity is apparent but as the growth temperature is reduced from 37°C to 30°C to ~23°C (room temperature), there is an increase in secreted activity. Those cultures grown at lower temperature secreted more
activity, although they had a slower growth rate and produced smaller colonies than those grown at high temperatures.

Figure 3.10. *E. coli* TB1 cultures harboring various the plasmids pMAL-p, pMAL-c, pMAL-pA, pMAL-pB, pMAL-cA, and pMAL-cB grown on milk plates for 48 h at (A) room temperature (approximately 23°C); (B) 30°C, and (C) 37°C.
3.4.2.2. pLcII-FX Constructs

The coding sequence for the pro-mature region of SGPB was also ligated behind the coding region for the cII protein on the *E. coli* plasmid pLcII-FX\(^70\) to create pLcII-FX-pro-SGBP (Figure 3.11.). When induced, bacteria harboring pLcII-FX-pro-SGBP express fusion protein consisting of the first 31 amino acids of the lambda (\(\lambda\)) cII protein linked to the pro-mature SGBP by the tetrapeptide recognition site for a protease isolated from the blood coagulation cascade, Factor Xa.

![Figure 3.11](image)

Figure 3.11. Plasmid pLcII-FX-pro-SGBP.

*E. coli* QY13 harboring pLcII derivatives do not demonstrate any extracellular protease activity due to the cytosolic nature of expression from this plasmid and therefore assaying for protease activity by growth on milk plates is not possible.

3.5. Inactive SGBP

Once production of active SGBP was demonstrated, site-directed mutagenesis was performed by the Kunkel method\(^84,85\) in order to substitute a glycine residue for the active site serine. The oligonucleotide SG195 hybridized to the sequence surrounding the codon for Ser-195 of SGBP and directed the change of the DNA sequence from TCC (coding for serine) to GGT (coding for glycine) at position 195. To express the inactive SGBP protein, the coding region containing the SG195 mutation was ligated into the various expression plasmids to create the inactive counterparts: pbDC14+, pbDC14-, pMAL-pB/SG195, pMAL-cB/SG195, and pLcII-FX-pro-mature-SGBP/SG195 (Figure 3.12.). Changes in the secretion of protease activity were tested for by plating the bacteria containing these constructs on milk plates.
Figure 3.12. Inactive SGPB expression plasmids.

Figure 3.13. shows milk plates on which *B. subtilis* DB104 cultures harboring the plasmids pUB110, pbDC13+, pbDC13-, pEB-B8, pbDC14+, and pbDC14- were grown. Also shown is the *B. subtilis* DB104/pEB-HDG (constructed by Ian Reilly and Sachdev Sidhu) harboring a construct similar to pEB-B8 that codes for an inactive SGPB on an *E. coli/B. subtilis* shuttle vector.

Mutation of the codon for the active site serine to a glycine (SG195) in SGPB resulted in no secretion of active protease when expressed in *B. subtilis* (pbDC14+, pbDC14-) compared to the secretion in constructs containing the active protease (pbDC13+, pbDC13-). There was no difference in protease secretion when comparing the two inactive constructs in the *Bacillus* specific plasmid (pbDC14-) and the *E. coli/Bacillus* shuttle vector (pEB-HDG) although the codon for the glycine was different (GGT for pbDC14- and GGC for pEB-HDG).
Figure 3.13. *B. subtilis* DB104 cultures harboring various the plasmids pUB110, pbDC13+, pbDC13-, pEB-B8, pbDC14+, pbDC14- and pEB-HDG grown on milk plates for 48 h at 37°C.

Figure 3.14. shows milk plates on which *E. coli* TB1 cultures harboring the plasmids pMAL-p, pMAL-c, pMAL-pB, pMAL-cB, pMAL-pB/SG195, and pMAL-cB/SG195 were grown. Inactivation of SGPB by mutation of the codon for the active site serine to glycine (pMAL-pB/SG195) resulted in no secretion of active protease when expressed in *E. coli* TB1 compared to the secretion in constructs containing the active protease (pMAL-pB).
Figure 3.14. *E. coli* TB1 cultures harboring the plasmids pMAL-p, pMAL-c, pMAL-pB, pMAL-cB, pMAL-pB/SG195, and pMAL-cB/SG195 grown on milk plates for 48 h at (A) room temperature (approximately 23°C).

3.6. *Bacillus* Time Course

Production of extracellular protease in *B. subtilis* was also detected by growing bacteria harboring plasmid derivatives in liquid culture media. A volume of 100 mL of 2x TY supplemented with 50 μg/mL kanamycin was inoculated with 1 mL of a mid-log *B. subtilis* DB104 and samples were removed to monitor growth (optical density at 600 nm) and protease activity in the culture supernatant (cleavage of the substrate N-succinyl-ala-ala-pro-phe-p-nitroanilide) as the culture was incubated with shaking at 30°C. Due to variations in growth conditions (temperature fluctuations and changes in speed of the environmental shaker) in each separate experiment, the following graphical results represent only one of several trials and each include the necessary control samples.
3.6.1. Bacterial Growth

Figure 3.15. shows the growth curves (optical density at 600 nm versus time) of strains of *B. subtilis* DB104 monitored during growth. In general, each culture had a “lag” period of about 4 h where little increase in optical density was noted, followed by a period of exponential growth. The optical density of each culture peaked after approximately 30 h of growth and then declined as the cultures entered stationary phase. Comparing the shape and magnitude of each of the growth curves indicated that harboring a pUB110 plasmid derivative did not adversely affect the growth of any of the *B. subtilis* strains studied and that any plasmid-encoded proteins were not toxic to the bacterial cell.

3.6.2. Protease Activity

Figure 3.16. shows protease activity in culture supernatants as a function of time for strains of *B. subtilis* DB104 monitored during growth.

**Subtilisin BPN'-containing Constructs**

Figure 3.16. A shows the protease activity profile of *B. subtilis* DB104 strains harboring subtilisin BPN'-containing constructs and confirms the results shown on the milk plates. Harboring the plasmid pPT30 resulted in the greatest amount of protease activity due to the secretion of subtilisin BPN'. This activity level was ten-fold the amount of activity secreted by strains harboring either pGK4a and pDC1 in which the subtilisin BPN' transcript was interrupted by the mature-region of SGPB. Both strains harboring pUB110 and pKN1 (pre-pro-subtilisin BPN'-mature-SGPB) had baseline protease activity.

**SGPA-containing Constructs**

Figure 3.16. B shows the protease activity profile of strains harboring SGPA-containing constructs and confirms the results shown on the milk plates. Harboring the plasmid pbDC12- (pre-subtilisin BPN'-pro-mature-SGPA in the counterclockwise orientation) resulted in the greatest amount of secreted activity compared to the same fragment in the opposite orientation (pbDC12+) or either of the pre-pro-mature-SGPA
constructs (pPA1 or pPA2). For pbDC12- the peak activity level is approximately 100-fold less than that of subtilisin BPN' (pPT30) and unlike subtilisin BPN' and SGPB proteases, the activity due to SGPA is not stable under these conditions and declines after peaking at 30 h.

SGPB-containing Constructs

Figure 3.16. C shows the protease activity profile of strains harboring SGPB-containing constructs and confirms the results shown on the milk plates. Harboring the plasmid pbDC13- (pre-subtilisin BPN'-pro-mature-SGPB in the counterclockwise orientation) resulted in the greatest amount of protease activity. The strain *B. subtilis* DB104/pbDC13+ with the pre-subtilisin BPN'-pro-mature-SGPB fragment in the opposite orientation had twice the activity of *B. subtilis* DB104/pPB1 (pre-pro-mature-SGPB in the counterclockwise orientation) and the strains harboring PUB110 and pPB2 have baseline amounts of protease activity (Figure 3.16. D).
Figure 3.15. Growth curves (optical density at 600 nm versus time) of strains of *B. subtilis* DB104 monitored during growth with shaking at 30°C.
Figure 3.16. Protease activity profiles (protease activity versus time) of strains of *B. subtilis* DB104 monitored during growth with shaking at 30°C.
3.6.2.1. Spiking of SGPA with SGPB

When expressed under the control of the subtilisin BPN' promoter and pre region, both SGPA and SGPB were successfully secreted from *B. subtilis* although to different degrees. The peak activity of each strain grown at 30°C for ~30 h (Figure 3.16. B, 3.16. C) differed 50-fold and there was a significant difference in stabilities of the protease activity under these conditions. Since both proteases are secreted simultaneously in the endogenous organism, *S. griseus*, it was thought that one protease (namely SGPB) might assist the maturation of the other (SGPA). To test this, a small amount of SGPB activity or an equivalent volume of buffer was added to several *B. subtilis* DB104 cultures harboring SGPA constructs, and the total protease activity in the culture supernatant was monitored over time. As shown in Figure 3.17., total activity of the SGPB-spiked *B. subtilis* DB104/pbDC8- culture supernatant (pre-subtilisin BPN'-pro-mature-SGPA) increased over time while the total activity of the buffer-spiked culture supernatant did not. Spiking with SGPB did not lead to increases in total activity for any of the other cultures containing SGPA constructs. (Note: Plasmid pbDC8- was constructed in a manner similar to pbDC12- except that a fragment of 324 nucleotides was removed from the *mob* gene on pUB110 during construction and secretes similar levels of protease activities.)

To determine whether or not the increases in activity were dependent on the amount of protease added or on the growth of the bacteria, various amounts of SGPB activity were added to *B. subtilis* DB104 cultures harboring pUB110 or pbDC8- and the total activity was determine after 7 h (Figure 3.18.). No further increases in activity were observed in the *B. subtilis* DB104/pbDC8- culture supernatant when more than approximately 30 nmol/min/mL of SGPB activity was added.
Figure 3.17. Protease activity profiles (total protease activity versus time) of strains of *B. subtilis* DB104 monitored during growth with shaking at 37°C. To several *B. subtilis* DB104 cultures harboring SGPA constructs (optical density at 600 nm -3.0-4.0), a small amount of SGBP activity (20 nmol/min/mL) or an equivalent volume of buffer (50 mM Tris-Cl, pH 8.0) was added and the total protease activity in the culture supernatant was monitored over time. (Note: All samples with buffer additions had baseline activity levels.)
Protease activity profiles (total protease activity versus time) of strains of *B. subtilis* DB104 monitored during growth with shaking at 37°C. To *B. subtilis* DB104 cultures harboring pUB110 or pDC8- (optical density at 600 nm -3.0-4.0), various amounts of SGPB activity or an equivalent volume of buffer (50 mM Tris Cl, pH 8.0) were added and the total protease activity in the culture supernatant was monitored over time. (Note: All samples with buffer additions had baseline activity levels.)
3.7. SDS-Polyacrylamide Gel Electrophoresis and Western Blotting

SDS-polyacrylamide gel electrophoresis and western blotting was used to demonstrate the presence of the SGPB polypeptides, whether as processed mature proteases or as fusion proteins.

3.7.1. Bacillus Expression

Culture Supernatant

Figure 3.19. shows a Coomassie Blue-stained SDS-polyacrylamide gel of fractionated culture supernatants. The B. subtilis DB104 strain harboring the construct pbDC13- had a unique band (lane 6) compared to the negative control, B. subtilis DB104/pUB110 (lane 5), at an approximate molecular weight of 17 800 Da in the 40-60% acetone fraction. This value corresponded with the theoretical molecular weight of 18 645 Da for mature SGPB, indicating that this strain successfully secretes an active processed SGPB. Neither of the strains harboring constructs containing the inactive genes (pbDC14- and pEB-HDG; lanes 7-8) secreted a protein of a similar size (detectable by Coomassie staining). This result should be anticipated if the assumption of autoprocessing of the SGPB precursor is correct. If the signal peptide of the precursor is cleaved and the pro-mature SGPB (26 673 Da) released to the supernatant, any evidence of its presence would be obscured due to the heavily staining band in all samples at 27 000 Da.
Mouse polyclonal antibodies raised against mature SGPB (work done by Sachdev Sidhu and Dr. G. Kalmar) were used in a western blot to detect SGPB or SGPB precursors in the culture supernatant of a number of *B. subtilis* cultures (Figure 3.20.). The polyclonal antibodies did not react strongly to any proteins in the 0-40% acetone fraction of the culture supernatants (lanes 1-4) but reacted with a number of proteins in all samples of the 40-60% acetone fraction (lanes 5-8), including the strain harboring pUB110 which acted as a negative control for SGPB expression (lane 5). The only unique reactive protein band occurred in the strain harboring an active construct of SGPB (pbDC13-) at a molecular weight comparable to that of mature SGPB (lane 9). No unique reactive bands occurred in either strain harboring an inactive construct although any evidence of a pro-mature form of SGPB (26 673 Da) released into the supernatant would be obscured by the highly reactive band at 27 000 Da in all samples.
3.7.1.1. Inactive SGPB Spiking

As no evidence of processed or unprocessed inactive SGPB was observed in the culture supernatant of the *B. subtilis* strains pbDC14- or pEB-HDG, it was thought that if the SGPB precursor was successfully transported across the cytoplasmic membrane, it might still be attached to the outer surface of the bacterial cell, unable to affect auto-maturation due to the inactivation of the protease. Addition of active SGPB might be able to process the inactive form in an intermolecular reaction but the active and inactive proteins would be indistinguishable. Thus, processing of the inactive form of SGPB was attempted using the related *S. griseus* proteases C and D. All *S. griseus* proteases have similar precursor structure (pre-pro-mature protease) and proteases C and D have similar sequences surrounding the pro-mature junctions similar to SGPB.
<table>
<thead>
<tr>
<th>Protease</th>
<th>Sequence surrounding pro-mature junction</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGPB</td>
<td>ERTPGKFTKL -ISGGDAIYS</td>
</tr>
<tr>
<td>SGPC</td>
<td>ARSAEQPRAL -ADIRGGDAYYM</td>
</tr>
<tr>
<td>SGPD</td>
<td>NRTAGEFTPL -IAGGDAIWG</td>
</tr>
</tbody>
</table>

Strains of *B. subtilis* DB104 harboring the plasmids pUB110, pbDC13-, pbDC14- and pEB-HDG were incubated with equal amounts of exogenous SGPC and SGPD activities (Section 2.4.3.). After treatment, each culture pellet was fractionated and samples applied to duplicate SDS-polyacrylamide gels. One gel was stained with Coomassie Blue while the other was used for a western blot. The results of the Coomassie Blue stained gels are shown in Figure 3.21. No unique protein bands were observed in any of the three fractions, independent of the experimental conditions.

The western blots performed on the duplicate set of gels did not reveal any highly reactive bands. Three faint unique bands were observed in the incubation supernatant in the samples containing *B. subtilis* DB104/pbDC13- (active SGPB). These bands were the size of mature SGPB and were presumably due to the release of either nascent protease or protease temporarily restricted by the cell wall. No unique bands were observed in the cell supernatant fraction while three very faint bands were observed in the membrane fraction of pEB-HDG. These bands were of the size of pre-pro-mature-SGPB (30 526 Da) but were not evident in the *B. subtilis* DB104 strain expressing active SGPB (pbDC13-). Similar results were demonstrated in the expression of an inactive variant of subtilisin BPN' (Asparagine-32) in *B. subtilis*. After extended periods of incubation the unprocessed pre-pro-mature precursor accumulated in the membrane whereas the active subtilisin precursor was rapidly processed and was not detected in the membrane fractions.
Figure 3.21. Coomassie Blue-stained SDS-polyacrylamide gel (15%) of (A) incubation supernatant, (B) cytosolic fraction, and (C) membrane fractions of *B. subtilis* DB104 cultures treated with buffer (50 mM Tris-Cl, pH 8.0), SGPC, or SGPD. Each gel was loaded as follows: pUB1110 treated with buffer (Lane 1), SGPC (Lane 2) or SGPD (Lane 3); pbDC13- treated with buffer (Lane 4), SGPC (Lane 5) or SGPD (Lane 6); pbDC14- treated with buffer (Lane 7), SGPC (Lane 8) or SGPD (Lane 9); pEB-HDG treated with buffer (Lane 10), SGPC (Lane 11) or SGPD (Lane 12); and purified SGBP (Lane 13).
3.7.2.  *E. coli* Expression

3.7.2.1.  pLcII-FX-constructs

Expression of both active and inactive forms of SGPB was attempted in *E. coli* QY13 using the pLcII-FX expression system. Figure 3.22. shows the results of a small scale test expression. No unique proteins were observed in the cell supernatant of induced samples (Figure 3.22. A, lanes 2, 4-7) while a single protein was observed in the cell pellet of the induced sample for the active SGPB construct (Figure 3.22. B, lane 2). This unique protein had an approximate molecular weight of 32 000 Da which corresponded with the theoretical weight of 30 802 Da for the cII-FX-pro-mature-SGPB fusion protein. No unique proteins were observed in the cell pellet of the induced inactive samples (lanes 4-7).

As expression using the pLcII-FX expression system was known to be dependent on the conditions of the transformation, bacterial growth and induction, another test expression was conducted using several colonies from a separate transformation. Figure 3.23. shows the results of this test expression. Again no unique proteins were observed in the cell supernatant (Figure 3.23. A) while a single protein of an approximate molecular weight of 32 000 Da was observed in the cell pellet of the induced samples for the active SGPB construct (Figure 3.23. B, lanes 2-3). No unique proteins were observed in the cell pellet of the induced inactive samples (lanes 5-9).
Figure 3.22. Coomassie Blue-stained SDS-polyacrylamide gel (15%) of (A) cell supernatant and (B) cell pellet fractions of bacterial pellets. Lanes 1-2 contain samples from *E. coli* QY13 harboring pLcII-FX-pro-mature-SGPB (active). Lanes 3-7 contain separate sample from a strain containing the inactive SGPB construct (pLcII-FX-pro-mature-SGPB/SG195). Samples in lanes 1 and 3 were from uninduced cultures while the remaining were from induced cultures.
Figure 3.23. Coomassie Blue-stained SDS-polyacrylamide gel (15%) of (A) cell supernatant and (B) cell pellet fractions of bacterial pellets. Lanes 1-3 contain samples from *E. coli* QY13 harboring pLcII-FX-pro-mature-SGPB (active). Lanes 4-9 contain separate samples from a strain containing the inactive SGPB construct (pLcII-FX-pro-mature-SGPB/SG195). Samples in lanes 1 and 4 were from uninduced cultures while the remaining were from induced cultures.
3.7.2.2.  **pMAL-constructs**

Expression of both active and inactive forms of SGBP were attempted in *E. coli* TB1 using the pMAL expression system. Figure 3.24. shows the results of a small scale test expression. In the cell supernatant of pMAL-p (Figure 3.24. A; lane 1), a protein band appeared corresponding in size to the MBP-β-galactosidase-α fusion (52 000 Da). No other overexpressed proteins were observed in the cell supernatant of induced samples.

In the pellet fraction of pMAL-p (Figure 3.24. B; lane 1), one major protein band appeared at a position corresponding in size to the MBP-β-galactosidase-α fusion protein. In the pellet fraction of pMAL-pB (lane 2) a single unique protein band appeared at an approximate molecular weight of 67 700 Da which corresponds with a theoretical molecular weight of 68 673 Da for an MBP-pro-mature-SGBP fusion. No unique proteins were observed in the cell pellet of the induced inactive samples (lanes 3-6).

As conceded by the supplier (New England BioLabs), not all gene fusions in the pMAL system produce full length protein fusions. In some cases the protein fusion is subject to degradation by *E. coli* proteases, which is a potential fate for the inactive variants. In effort to express inactive SGBP, both active and inactive constructs were transformed into a total of six *E. coli* K12 mutants (acquired from New England BioLabs) possessing a variety of mutations affecting heterologous protein stability (Table 3.2.). Figures 3.25. and 3.26. show the results using the mutant *E. coli* strains K12 CAG626 and K12 CAG629. No overexpressed proteins were observed in the cell supernatant of either strain (Figures 3.25. A, 3.26. A) and only the active constructs (lanes 1-4) demonstrated an overexpressed protein of the appropriate size in the pellet fraction (Figure 3.25. B, 3.26. B). It was also noted that control of protein expression in the strain K12 CAG626 was less stringent than for the others attempted as MBP-pro-SGBP (active) fusion protein was observed in both the uninduced and induced pellet fraction samples.
Figure 3.24. Coomassie Blue-stained SDS-polyacrylamide gel (15%) of (A) cell supernatant and (B) cell pellet fractions of bacterial pellets. Lane 1 contains an induced sample from *E. coli* TB1 harboring pMAL-p. Lane 2 contains an induced sample from a strain containing pMal-pB (active SGPB). Lanes 3-4 contain separate samples from a strain containing pMal-pB/SG195 (inactive SGPB). Lanes 5-6 contain separate samples from a strain containing pMal-cB/SG195 (inactive SGPB). The two center lanes contain the molecular weight markers and a purified sample of the maltose binding protein (MBP).
Table 3.2. Mutant *E. coli* K12 strains.

<table>
<thead>
<tr>
<th><em>E. coli</em> strain</th>
<th>Relevant Genotype</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>K12 CAG597&lt;sup&gt;88&lt;/sup&gt;</td>
<td>rpoHam&lt;sup&gt;a&lt;/sup&gt;165 lacZam supCts</td>
<td>-The <em>rpoHam</em> allele encodes an amber mutant form of the sigma factor for the heat shock response. The strain has a temperature sensitive amber suppressor (<em>supCts</em>), so upon incubation at 37°C or 42°C the strain becomes deficient in expression of the heat shock proteins.</td>
</tr>
<tr>
<td>K12 CAG626&lt;sup&gt;89&lt;/sup&gt;</td>
<td>lon supCts</td>
<td>-strain is <em>lon</em>&lt;sup&gt;−&lt;/sup&gt; which knocks out the major ATP-dependent protease in the <em>E. coli</em> cytoplasm.</td>
</tr>
<tr>
<td>K12 CAG629&lt;sup&gt;88,89&lt;/sup&gt; (CAG597 lon)</td>
<td>lon rpoHam165 supCts</td>
<td>-strain contains the <em>rpoHam</em> allele and has a temperature sensitive amber suppressor (<em>supCts</em>), and is <em>lon</em>&lt;sup&gt;−&lt;/sup&gt;.</td>
</tr>
<tr>
<td>K12 CAG748&lt;sup&gt;90&lt;/sup&gt;</td>
<td>lacY Δ(lac)&lt;sup&gt;x&lt;/sup&gt;90 dnaJ259</td>
<td>-strain is deficient in the DnaJ &quot;chaperonin&quot;</td>
</tr>
<tr>
<td>K12 PR745&lt;sup&gt;89&lt;/sup&gt;</td>
<td>lon Δ(malB) Δ(lac)U169</td>
<td>-strain is <em>lon</em>&lt;sup&gt;−&lt;/sup&gt; which knocks out the major ATP-dependent protease in the <em>E. coli</em> cytoplasm. The <em>malE</em> gene, which encodes maltose binding protein, is deleted as part of the <em>malB</em> deletion.</td>
</tr>
<tr>
<td>K12 UT5600&lt;sup&gt;91&lt;/sup&gt;</td>
<td>lacY1 Δ(ompT-fepC)266</td>
<td>-strain is deficient in the OmpT protease, a periplasmic protease which can degrade cytoplasmically expressed proteins when the cells are lysed to make a crude extract, as well as proteins expressed in the periplasm.</td>
</tr>
</tbody>
</table>
Figure 3.25. Coomassie Blue-stained SDS-polyacrylamide gel (15%) of (A) cell supernatant and (B) cell pellet fractions of bacterial pellets. Lanes 1-2 contain uninduced and induced samples (respectively) from *E. coli* K12 CAG626 harboring pMAL-pB; Lanes 3-4 contain uninduced and induced samples of strains containing pMAL-cB; Lanes 5-6 contain uninduced and induced samples of strains containing pMAL-pB/SG195 (inactive); and Lanes 7-8 contain uninduced and induced samples of strains containing pMAL-cB/SG195 (inactive).
Figure 3.26. Coomassie Blue-stained SDS-polyacrylamide gel (15%) of (A) cell supernatant and (B) cell pellet fractions of bacterial pellets. Lanes 1-2 contain uninduced and induced samples (respectively) from *E. coli* K12 CAG629 harboring pMAL-pB; Lanes 3-4 contain uninduced and induced samples of strains containing pMAL-cB; Lanes 5-6 contain uninduced and induced samples of strains containing pMAL-pB/SG195 (inactive); and Lanes 7-8 contain uninduced and induced samples of strains containing pMAL-cB/SG195 (inactive).
In an effort to determine whether or not any inactive SGPB was being translated in *E. coli*, rabbit monoclonal antibodies specific to the MBP (supplied by New England BioLabs) were used to detect the MBP and the MBP-pro-mature-SGPB fusion proteins in both cell supernatant and pellet fractions of small scale test expressions of *E. coli* strains harboring pMAL constructs.

Figure 3.30. shows a western blot of crude protein samples from *E. coli* TB1 harboring various pMAL constructs. A band corresponding to the MBP-β-gal-α fusion (52 000 Da) was found in both the cell supernatant and pellet fraction of a strain harboring pMAL-p (Figures 3.30.A, 3.30.B; lane 1). A band corresponding to the MBP-pro-mature-SGPB (68 673 Da) of the active construct (pMAL-pB) was found in both the cell supernatant and pellet fractions (lane 2) but no single major band was evident for the inactive constructs (lanes 3-6). In all samples, degradation products were observed as a smear of antigenic proteins and in some samples (lane 3) a band corresponding to MBP (42 000 Da) was apparent.

Western blots were performed on samples obtained from mutant *E. coli* K12 strains described in Section 3.2.3.2.2. In each case although no protein band for the inactive SGPB fusion protein was apparent on the Coomassie-stained SDS-polyacrylamide gel, an immuno-reactive protein was found at the same apparent molecular weight as the MBP-pro-mature SGPB fusion as well as a few thousand Daltons (~3 000 Da) heavier. Figures 3.31. and 3.32. are western blots on samples from *E. coli* strains K12 CAG626 and K12 CAG629.
Figure 3.30. Western blot using rabbit anti-MBP monoclonal antibodies on (A) cell supernatant and (B) cell pellet fractions of bacterial pellets. Lane 1 contains an induced sample from *E. coli* TB1 harboring pMAL-p. Lane 2 contains an induced sample from a strain containing pMal-pB (active SGPB). Lanes 3-4 contain separate samples from a strain containing pMal-pB/SG195 (inactive SGPB). Lanes 5-6 contain separate samples from a strain containing pMal-cB/SG195 (inactive SGPB). Note: Lanes 3-6 were loaded with 5x the sample volume as lanes 1-2.
Figure 3.31. Western blot using rabbit anti-MBP monoclonal antibodies on cell supernatant and cell pellet fractions of bacterial pellets. Lanes 1-2 contain pellet and cell supernatant fractions (respectively) from an uninduced culture of *E. coli* K12 CAG626 harboring pMAL-pB; Lanes 3-4 contain pellet and cell supernatant fractions from an induced culture of the strain harboring pMAL-pB (active); Lanes 5-6 contain pellet and cell supernatant fractions from induced samples of the strain containing pMAL-pB/SG195 (inactive); and Lanes 7-8 contain pellet and cell supernatant fractions from induced samples of the strain containing pMAL-cB/SG195 (inactive). Note: Lanes 5-8 were loaded with 5x the sample volume as lanes 1-4.
Figure 3.32. Western blot using rabbit anti-MBP monoclonal antibodies on cell supernatant and cell pellet fractions of bacterial pellets. Lanes 1-2 contain pellet and cell supernatant fractions (respectively) from an uninduced culture of E. coli K12 CAG629 harboring pMAL-pB; Lanes 3-4 contain pellet and cell supernatant fractions from an induced culture of the strain harboring pMAL-pB (active); Lanes 5-6 contain pellet and cell supernatant fractions from induced samples of the strain containing pMAL-pB/SG195 (inactive); and Lanes 7-8 contain pellet and cell supernatant fractions from induced samples of the strain containing pMAL-cB/SG195 (inactive). Note: Lanes 5-8 were loaded with 5x the sample volume as lanes 1-4.
4. Discussion

4.1. Expression of a *Streptomyces* Gene in *Bacillus*

Successful expression of foreign genes in bacterial systems depends on many factors, including DNA composition, signals directing RNA transcription, protein translation (codon usage) and secretion mechanisms. Each can hamper the synthesis of the protein product and comparisons must be made between the organism that is the source of the foreign DNA and the host bacterium when attempting recombinant expression, in this case between *Streptomyces* and *Bacillus* species.

The genome of *Streptomyces* species has a high content of guanidine-cytidine base pairs (70-75%)\(^\text{92}\) compared to *E. coli* (~50%) and *Bacillus* (~43%). This high G-C content results in a highly biased codon usage with ~90% of G or C in the third and most degenerate position of the translated codons.\(^\text{92}\) (For the proteases SGPA and SGPB respectively, 92 and 96% of the bases in the third position of the coding region are either G or C.)\(^\text{31}\)

Although many organisms used to express foreign genes demonstrate highly biased codon usage, comparison of codon frequency in endogenous sequences shows that *B. subtilis* is less biased than *E. coli*\(^\text{93}\) or *Saccharomyces cerevisiae*.\(^\text{94}\) Thus, *S. griseus* proteins have a higher potential of being translated in *B. subtilis* although expression in other bacterial systems can not be excluded.

Recombinant expression of the mature portion of SGPB was attempted in *B. subtilis* using portions of the *B. amyloliquefaciens* subtilisin BPN' gene due to the similarities in expression of the two proteases. Both SGPB and subtilisin BPN' are synthesized as precursors with a signal peptide (pre-region) separated from the mature enzyme by a long pro-region and both are secreted as the processed mature protease. Both *S. griseus* and *B. subtilis* are Gram-positive bacteria with a similar cell structure consisting of a single cell membrane surrounded by a thick peptidoglycan cell wall. Secretion in these bacteria results in the release of protein directly into the culture
medium and as a general consequence, formation of disulfide bonds is favored in the more oxidative environment of the culture media. Expression of proteins in the more reductive cytosol does not support formation of disulfide bonds and has an important consequence for SGPB which contains two disulfide bonds. Release of the desired protein into the culture supernatant of Gram-positive bacteria leads to easier purification compared to the Gram-negative organism *E. coli* which has a second lipid membrane surrounding the cell wall. Secretion in *E. coli* results in accumulation of the exported protein within the periplasmic space between the inner and outer membranes, thus complicating purification.

*Bacillus* species are also well-known for the synthesis of large quantities of secreted proteins produced on an industrial scale. Most are endogenous proteins such as proteases (serine and metalloproteases) and starch hydrolyzing enzymes (amylases) secreted from strains selected for high extracellular enzyme levels or strains that have had the enzyme gene copy number increased through the introduction of recombinant plasmids. As such, these proteins are produced in grams per liter quantities in large scale fermentations.

Numerous heterologous proteins have been secreted by *Bacillus* species on a much smaller scale. Promoters and translation start signals from other Gram-positive bacteria are generally functional in *Bacillus* and signal peptides are recognized by the secretion apparatus (e.g. staphylococcal nuclease from *Staphylococcus aureus*) and cleaved by the signal peptidase. In cases where the heterologous sequences do not function, coupling of *Bacillus* sequences and signal peptide to the amino terminus of the protein can result in secretion of the foreign protein.

4.2. *B. subtilis* Subtilisin BPN' Promoter Pre-pro-region Constructs

Without prior knowledge of the specific function of the pro-region and the possibility of it having only a general involvement in the secretion process, an attempt was made to secrete SGPB as a fusion protein consisting of the pre-pro-region of
subtilisin BPN' followed by the mature portion of SGPB (plasmid pKN1; Figure 3.1.). This construct would utilize the subtilisin BPN' pre-region as a Bacillus specific signal peptide to transport the fusion protein to the cell surface. Because subtilisin BPN' is responsible for cleavage of its pro-mature bond after secretion, it was thought that due to its P$_1$ specificity this enzyme might still recognize the pro-subtilisin BPN'-mature-SGPB junction. Thus, subtilisin BPN' could be used to catalyze the release of mature SGPB (or SGPB mutants) in cultures expressing the pre-pro-subtilisin BPN'-pro-mature-SGPB fusion construct. To facilitate the cleavage of the fusion protein with subtilisin BPN', an additional amino acid (alanine) was added to the amino terminus of mature SGPB due to the P$_1$' specificity of subtilisin BPN' for small hydrophobic residues. This residue is also the native P$_1$' amino acid for self-maturation in pre-pro-mature subtilisin BPN' and would replace the isoleucine residue that is located at the N-terminus of SGPB.

The cleavage sites of the pro-mature bond of the native and fusion proteins are listed in Table 4.1. where * denotes the point of cleavage. The pro-mature junction of the fusion construct could also potentially be a good substrate for SGPB. SGPB itself has extended specificity requirements with a P$_2$ specificity for hydrophobic, non-aromatic residues; a P$_1$ specificity for the residues Phe, Tyr, Leu, and Trp; and a P$_1$' specificity for aromatic or long aliphatic residues. Although the hybrid junction would not be ideal, one might expect cleavage to some degree if SGPB was properly folded and active in the fusion protein.

<table>
<thead>
<tr>
<th>subtilisin BPN'</th>
<th>-1</th>
<th>+1</th>
<th>Gly-Lys-Phe-Thr-Lys-Leu</th>
<th>*Ile-Ser-Gly-Gly-Asp-Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGPB</td>
<td></td>
<td></td>
<td>His-Val-Ala-His-Ala-Tyr</td>
<td>*Ala-Gln-Ser-Val-Pro-Tyr</td>
</tr>
<tr>
<td>pro-subtilisin BPN'-mature SGPB fusion</td>
<td></td>
<td></td>
<td>subtilisin BPN' -1</td>
<td>SGPB +1</td>
</tr>
</tbody>
</table>

Table 4.1. Amino acid sequences surrounding the cleavage sites of the pro-mature bond of pre-pro-subtilisin BPN' fusion constructs.
Expression of SGPB as a pre-pro-subtilisin BPN'-mature-SGPB fusion protein did not result in the production of active protease as shown by growth of *B. subtilis* DB104/pKN1 on milk plates or in liquid culture (Figures 3.2 and 3.16). Attempts at expression of subtilisin E (*B. subtilis*) without its native pro-region resulted in inactive protease \(^{39}\) and only expression of a construct containing the intact pro-region resulted in active protease. By analogy, if the pro-region is necessary for the production of active SGPB, then attempts at expression without it would result in accumulation of inactive protease.

Cleavage of only the signal peptide of the pKN1 fusion would result in the release of the pro-subtilisin BPN'-mature-SGPB fusion protein into the supernatant. Signal peptide cleavage could occur in theory due to the existence of *Bacillus* signal peptidase I that has recently been cloned \(^{98}\) and *in vitro* experiments that demonstrate cleavage of the pre-pro junction of subtilisin E in *B. subtilis*. \(^{26}\) Even so, no pro-subtilisin BPN'-mature-SGPB fusion protein was released to the supernatant as demonstrated by SDS-PAGE. Lack of secretion of pro-mature precursor was also found in the study of expression of an active site mutant of subtilisin BPN' in *B. subtilis*. \(^{33}\) Cleavage of either the pre-pro or the pro-mature junctions in the inactive form did not occur and the full length pre-pro-mature-subtilisin BPN' remained on the cell surface. The possibility exists that the presence of the pro- or pro-mature region prevents the signal peptidase from cleaving the recognition sequence or changes its conformation such that it is not recognized and cleaved. Thus, there might be an order of cleavage reactions with the pro-mature bond being cleaved prior to or concurrent to the pre-pro bond. On the other hand, heterologous expression of the mature sequence of staphylococcal protein A fused to the pro-region of subtilisin BPN' resulted in the release of pro-subtilisin BPN'-protein A into the culture media after cleavage of the signal peptide. \(^{22}\) These two examples demonstrate the variability of the bacterial response to sequence changes in the secretion of heterologous fusion proteins.
Given the importance of the pro-region to the production of active protease, it was found that the pro-region need not be contiguous with the mature region of subtilisin BPN' during secretion to produce active protease. Secretion of active subtilisin BPN' was demonstrated in *B. subtilis* DB104 strains harboring the plasmids pDC1 and pGK4a in which the pro and mature regions of the subtilisin BPN' gene are separated by the mature region of the SGPB gene (Figures 3.2. and 3.16.). In each case active subtilisin BPN' is released into the media although the mechanism of activation and cleavage, whether by an inter- or an intramolecular reaction, is open to speculation.

The plasmid pDC1 was constructed fortuitously during the construction of pKN1 and was missing the last 18 nucleotides on the 3'-end of the SGPB gene. Since no active SGPB was evident in the culture supernatant along with subtilisin BPN', the full length mature SGPB was reconstructed (pGK4a) to investigate whether SGPB would be released, having subtilisin BPN' present to cleave the hybrid pro-mature junction. Again no SGPB was detected in the culture supernatant. The amino acid sequences surrounding the N-terminus of the mature region of subtilisin BPN' in the three subtilisin BPN'-containing constructs are listed in Table 4.2, where * denotes the point (pPT30) or potential point (pDC1 and pGK4a) of cleavage. In each construct, the site of subtilisin BPN' autocatalytic cleavage (Tyr * Ala) has been conserved and mature subtilisin BPN' is released into the media.

Table 4.2. Amino acid sequences surrounding the cleavage sites at the N-terminus of the mature region of subtilisin BPN' of the three subtilisin BPN'-containing constructs.

<table>
<thead>
<tr>
<th></th>
<th>subtilisin BPN' –1 +1 subtilisin BPN'</th>
</tr>
</thead>
<tbody>
<tr>
<td>pPT30</td>
<td>His-Val-Ala-His-Ala-Tyr&lt;sup&gt;*&lt;/sup&gt; Ala-Gln-Ser-Val-Pro-Tyr</td>
</tr>
<tr>
<td>pDC1</td>
<td>SGPB +180 +1 subtilisin BPN'</td>
</tr>
<tr>
<td></td>
<td>Glu-Ala-Leu-Ser-Ala-Tyr&lt;sup&gt;*&lt;/sup&gt; Ala-Gln-Ser-Val-Pro-Tyr</td>
</tr>
<tr>
<td>pGK4a</td>
<td>SGPB +185 +1 subtilisin BPN'</td>
</tr>
<tr>
<td></td>
<td>Tyr-Gly-Val-Ser-Val-Tyr&lt;sup&gt;*&lt;/sup&gt; Ala-Gln-Ser-Val-Pro-Tyr</td>
</tr>
</tbody>
</table>
Due to the high degree of expression of subtilisin BPN' from the plasmid constructs pDC1 and pGK4a, it follows that the mature portion of SGPB must be translated also, although the fate of the protein is not known. If the mature SGPB polypeptide is released to the media in an incorrectly folded state, it would be subject to proteolytic digestion by the abundant and active subtilisin BPN'. However, transcription and translation of mature SGPB from plasmids pDC1 and pGK4a would confirm that this *S. griseus* gene is not subjected to a *B. subtilis* codon bias.

4.3. Native *S. griseus* Promoter Pre-region Constructs

Due to the necessity of the pro-region to produce active enzyme, the complete gene sequences for SGPB and the related protease, SGPA, were isolated from *S. griseus* by use of the polymerase chain reaction (work done in this lab by Sachdev Sidhu with assistance from Kevin Neden). Expression of these genes was first attempted by using the native promoters, ribosome binding sites and signal sequences. Some measure of success was anticipated due to a number of factors. Promoters from one Gram-positive bacteria are generally recognized and function in other Gram-positive organisms. The transport of the protein product across the cell surface was anticipated due to the "universal design" of signal peptide structure among organisms. Release of the protease was possible due to the species-wide consensus sequence of signal peptidase cleavage.

The resulting low levels of protease activities secreted from the native constructs in *B. subtilis* (Figures 3.5. and 3.16.) could result from difficulties with ill-defined promoters in the two bacterial species (*Streptomyces* and *Bacillus*), with the stringency of the ribosome binding sequence in *Bacillus*, or RNA transcript stability. Although the levels of protease activity were not high, the presence of small amounts of activity in the culture supernatant of *B. subtilis*/pPB1 confirmed the requirement of the pro-region for the production of active SGBP.
4.4. *B. subtilis* Subtilisin BPN’ Promoter Pre-region Constructs

The demonstration of activity both on milk plates and in liquid culture indicated that secretion of the *S. griseus* proteases in *B. subtilis* was possible using their endogenous promoters and pre-regions (Figures 3.5. and 3.16.). To produce greater quantities of protease the method of expression was modified to use a genus specific signal that already demonstrated production of vast quantities of protein in the form of subtilisin BPN’. Other examples exist of heterologous expression, such as the secretion of the mature bovine pancreatic ribonuclease A25 and *Staphylococcus aureus* protein A.23 In both examples, the precursor was cleaved precisely after the subtilisin BPN’ signal peptide as demonstrated by microsequencing of the N-terminus of the secreted product.

The signal peptide cleavage sites of the subtilisin BPN’ pre-region fusion proteins are listed in Table 4.3, where * denotes the point of cleavage. Each retains the Ala-Gln-Ala * recognition sequence although it is not known whether this bond is cleaved. No pro-mature SGPB was detected in the culture supernatant and the mature protease could possibly cleave itself from the cell-associated hybrid pre-pro-protein.

Table 4.3. Amino acid sequences surrounding the cleavage sites of the pre-pro bond of pre-subtilisin BPN’ fusion constructs.

<table>
<thead>
<tr>
<th></th>
<th>subtilisin BPN’ -78</th>
<th>-77 subtilisin BPN’</th>
</tr>
</thead>
<tbody>
<tr>
<td>pPT30</td>
<td>Thr-Ser-Ser-Ala-Gln-Ala * Ala-Gly-Lys-Ser-Asn-Gly</td>
<td></td>
</tr>
<tr>
<td>pbDC12-</td>
<td>subtilisin BPN’ -78</td>
<td>-78 SGPA</td>
</tr>
<tr>
<td></td>
<td>Thr-Ser-Ala-Gln-Ala * Ala-Pro-Glu-Ala-Glu-Ser</td>
<td></td>
</tr>
<tr>
<td>pbDC13-</td>
<td>subtilisin BPN’ -78</td>
<td>-76 SGPB</td>
</tr>
<tr>
<td></td>
<td>Thr-Ser-Ala-Gln-Ala * Glu-Thr-Pro-Arg-Thr-Phe</td>
<td></td>
</tr>
</tbody>
</table>

Each pre-subtilisin BPN’ fusion construct produced active SGPB and SGPA in greater quantities compared to the constructs containing *S. griseus* sequences (Figures 3.7. and 3.16.). Similar amounts of SGPB and subtilisin BPN’ were secreted as demonstrated by protein bands of similar intensities on SDS-polyacrylamide gels but the
proteases differ by more than 20-fold in their catalytic efficiency for the enzyme substrate chosen for the activity assays. SGPB has a $k_{cat}/K_m$ value of $10,900 \text{ (M}^{-1}\text{s}^{-1})$ (100 mM Tris pH 8.0) while subtilisin BPN' has a value of $250,000 \text{ (M}^{-1}\text{s}^{-1})$ (100 mM Tris pH 8.6).

4.5. Directional Dependence of Expression in pUB110

Expression of the proteases SGPA and SGPB in B. subtilis was found to depend on the direction of the cloned gene fragment in the plasmid pUB110 and occurred when the expression was under the control of both the Streptomyces and Bacillus promoters (Figures 3.5. and 3.7.). Higher expression resulted when the gene was in the same orientation as the existing four open reading frames on the plasmid itself (Figure 4.2.) and appeared to be the result of transcriptional read-through. These open reading frames include $km$ and $blm$, which are responsible for kanamycin and bleomycin resistance; $rep$, which codes for a replication initiation protein having nicking-closing activity; and $mob$, which is required for conjugative mobilization of the plasmid.

![Figure 4.1. Bacillus plasmid pUB110.](image)

This directional phenomenon had been demonstrated previously in the expression of chloramphenicol acetyltransferase (both with and without promoter). In vitro experimental results gave evidence of transcriptional read-through from the promoters of the open reading frames later identified as $mob$ and $km$. Only the introduction of a transcription termination signal 3’ to the plasmid-encoded genes led to expression which

97
was unaffected by orientation effects. Thus, transcriptional read-through from promoters on pUB110 affects the rate of expression of the cloned genes, either by supplying more polymerase molecules and producing longer transcripts when the gene is in the same orientation or by retarding transcribing ones when the gene is in the opposite orientation.

Another factor that possibly affects the overall level of expression in pUB110 is the strengths of the plasmid promoters and the position of the promoter relative to the cloned gene during transcriptional read-through. This aspect has not been fully explored due to the lack of convenient cloning sites within the other plasmid-encoded genes compared to the three unique restriction sites within the mob gene (Figure 4.2.).

4.6. Effect of Growth Temperature on Protease Activity

Expression of active *S. griseus* proteases was also dependent on the growth temperature of the bacteria. In *B. subtilis*, strains harboring SGPA and SGPB under the control of the subtilisin BPN' promoter/pre-region secreted a greater amount of protease activity when grown at 30°C than at 37°C (Figure 3.7.). Since subtilisin BPN' expression was not dependent on bacterial growth temperature (Figure 3.2.), these effects are specific for the *S. griseus* proteases, especially SGPA. A number of factors may be involved. There may be a particular step in the folding pathway for the pro-mature proteases that is temperature sensitive during secretion in *B. subtilis* but not in the native host. Alternatively, the rate of secretion (which may increase with higher growth rate at the higher temperature) may negatively affect proper folding to produce active protease.

Similar temperature-dependence was observed when expressing the *S. griseus* proteases in *E. coli* under the control of the tac promoter and fused to the maltose binding protein. Protease activity was almost negligible on milk plates grown at 37°C compared to those grown at 23°C (Figure 3.10.). These results mimic those found when expressing pro-subtilisin E (*B. subtilis*) in *E. coli* from the OmpA signal peptide in the expression plasmid pIN-III-ompA. A 16-fold increase in production of active subtilisin E was
achieved by using a low concentration of inducer (IPTG) and a low culture temperature (23°C). The rate of protein synthesis and the growth temperature were found to be important factors for the folding of pro-subtilisin BPN’ and the production of active mature protease. At high temperature and high inducer concentrations the precursor formed insoluble inclusion bodies, a general finding when proteins are overexpressed at 37°C. Over expression of SGPB at 37°C in *E. coli* also resulted in the formation of inclusion bodies as demonstrated by the isolation of insoluble MBP fusion protein in the cell pellet of crude test expressions (Figure 3.24.). Only with the low growth temperature did *E. coli* secrete an active SGPB (Figure 3.10.).

4.7. Expression of Inactive SGPB

Attempts at expression of inactive SGPB confirmed the statement made by Schein (1989) that "simply cloning a gene and making mRNA is not enough to produce soluble and active protein in bacteria." For SGPB, changing the codon for the active site serine (TCC) to that for glycine (GGT or GGC) negatively affected the expression of the protein in both *B. subtilis* and *E. coli*.

As indicated by northern blot analysis, mRNA specific for the inactive variants of SGPB was transcribed at the same levels as the active constructs in *B. subtilis* although overexpressed inactive protein was not detected (Figures 3.19. and 3.20.). Low level expression of the inactive SGPB could be explained by a possible negative feedback mechanism between synthesis of the precursor and export. Expression of an inactive subtilisin BPN’ in *B. subtilis* was found to result in membrane-associated precursor that was not released as a mature protease due to the inability of the precursor to process itself. The precursor reached a steady-state level of ~1000 sites per cell which was less than 0.2% of the total amount of protein produce by the wild-type construct. Power (1986) concluded that when processing of the protease precursor is blocked, further translation of the precursor may be inhibited. For subtilisin BPN’, this negative feedback was partially lifted and protein synthesis increased up to 5-fold when processing was
accomplished by the addition of endogenous protease. It is thought that the translocated but unprocessed subtilisin BPN’ precursor could affect its own synthesis by occupying the secretion machinery in the cell membrane.\textsuperscript{99} Similar examples of feedback control have been reported\textsuperscript{108,109} where heterologous genes were efficiently transcribed but not translated, giving further indication that the feedback control functions at the level of translation.

The question remains as to why inactive SGPB should affect the cleavage of the subtilisin BPN’ signal peptide when other similarly expressed heterologous proteins are released. Like the inactive subtilisin BPN’ construct, inactive pro-mature SGPB did not appear in the culture supernatant. Although the final mature protein structures are dissimilar, perhaps the folding of the mature protease prior to autolytic cleavage of the pro-mature bond precludes cleavage of the signal peptide bond. If the protease is inactive and unable to be released, the signal peptide would not be removed and the entire precursor would remain cell-bound.

4.8. Protein Stability

A major difficulty with heterologous expression of proteins is protein stability. Newly synthesized proteins are subjected to degradation by endogenous proteases found both within the bacterium and those secreted outside the cell. Much effort has been made to reduce the effects of this degradation, mostly through elimination of protease activity.

4.8.1. Bacillus

\textit{Bacillus} species are known for the large amounts of degradative enzymes, including proteases, that are secreted into the culture medium.\textsuperscript{99} The most significant of these are the alkaline and neutral proteases that are responsible for the majority (96%) of the protease activity secreted by the bacterium\textsuperscript{67} and many attempts at production of heterologous proteins have been unsuccessful due to proteolytic degradation of the secreted product. To this end, strains of \textit{Bacillus} have been isolated or engineered to produce low levels of endogenous protease activity\textsuperscript{27,67,110,111} and the search continues.
for other minor extracellular and intracellular proteases. Intracellular protease activity may be significant as it has been found that strains deficient in extracellular proteases are more apt to lyse, thus releasing these proteases into the culture medium.

When overexpressed in *B. subtilis*, the processed mature SGPB remained stable in the culture supernatant in excess of 72 h of growth (Figure 3.16.). However, the related protease, SGPA, had a significantly lower level of expression in *Bacillus*, both in the total amount of activity achieved and in protein released. The lack of SGPA stability could be due to a lack of proper protein folding in *Bacillus*, a lack of sufficient processing of the pro-mature protease or a lack of stability of the mature enzyme following release. Direct evidence of a decreased amount of stability of the mature SGPA is found in the time course measurements of the *Bacillus* strains harboring SGPA constructs (Figure 3.16.). Protease activity peaks following exponential growth, but is followed by rapid decline. This activity profile for SGPA differs from the sustained levels of activity in the cultures containing SGPB and subtilisin BPN' constructs and more closely resembles the expression of TEM-β-lactamase expression from the *B. amyloliquefaciens* α-amylase promoter and signal sequence. This construct, grown in a protease-proficient *Bacillus* strain, produces a heterologous protein that rapidly loses activity in the culture media as compared to the endogenously produced α-amylase. Thus for SGPA, the mature protein appears to be degraded *in vivo* following proper folding and release from the bacterium.

### 4.8.2. *E. coli*

Two separate SGPB fusion proteins, cII-FX-pro-SGPB and MBP-pro-SGPB, were constructed and expressed in *E. coli*. Each protein was found to be stable and demonstrated little degradation during bacterial cell lysis and crude purification (Figures 3.22. and 3.24.). Upon substituting the active site serine of SGPB with a glycine to produce an inactive variant, neither fusion protein was evident on a Coomassie Blue stained gel. Only when probed with a MBP-specific antibody could the MBP-pro-SGPB protein band be detected and then as a doublet consisting of a band of the expected size
along with one of comparable intensity that migrated slower (Figure 3.30.). Also evident on the western blot was a small amount of the intact MBP protein that was in a native conformation and able to bind to the amylose affinity column.

Expression of the inactive MBP-pro-SGPB construct was also attempted in a number of *E. coli* strains that demonstrate increased protein stability as a result of a variety of mutations (Table 3.2.). None of the mutant strains overexpressed the inactive SGPB fusion protein (Figures 3.25. and 3.26.) but all samples had the band doublet when probed with the anti-MBP antibody (Figures 3.31. and 3.32.). Since abnormal and/or denatured proteins are degraded more rapidly than native ones, the inactive SGPB portion appears to destabilize the MBP fusion protein and target it for a degradation pathway(s) common to all *E. coli* strains used. Covalent modification reactions are known to be involved in specifying proteins for some degradation pathways. These modifications include oxidation of amino acid residues, esterification of hydroxyl groups, acylation of NE-lysine residues, and the deamidation of glutamine and asparagine residues.\textsuperscript{117} The protein doublet demonstrated by the inactive SGPB construct may be the result of an undefined targeting mechanism that either increases the protein’s size or changes its electrophoretic mobility.

### 4.9. Spiking Experiments

#### 4.9.1. SGPA

Possible reasons for the overall lack of activity in SGPA-containing cultures involve the proper folding and release of the mature protease. Based on the subtilisin E model, once the precursor protease folds, it autocatalytically cleaves the pro-mature bond and releases itself from the bacterial cell. Since some activity appeared in SGPA-containing cultures, it is assumed that automaturation also occurred with SGPA. But since all *S. griseus* proteases are expressed at the same time within the native bacterium, the possibility exists for the interaction between proteases to assist in the maturation of one another. Experiments involving the addition of purified active SGPB to growing
cultures of SGPA found that overall activity of the culture supernatant increased several fold (Figure 3.17.), presumably through the release of mature SGPA (although not enough protein was present to purify). Intermolecular processing may not be a major method for activation of the SGPA precursor in S. griseus but within the context of the *Bacillus* expression system, SGPB appears to assist in the maturation of SGPA.

**4.9.2. Inactive SGPB**

In the experiments performed on subtilisin BPN’ that originally determined that this protease was autocatalytically processed, inactive constructs were released from the bacterial cell surface through the addition of exogenous protease (subtilisin Carlsberg) or by expressing the inactive construct in a protease proficient *B. subtilis* strain (subtilisin E). Addition of “helper” cultures of subtilisin E was subsequently employed to release a variety of inactive subtilisin BPN’ mutants in a study of the catalytic triad of subtilisin BPN’.

Similar tactics were used to attempt processing and release of inactive SGPB (Figure 3.21.). Since the only detection of the release of the inactive SGPB would be through western blot analysis on cell fractions, the related *S. griseus* proteases SGPC and SGPD were used to attempt processing of the inactive SGPB in a cross-reaction similar to the processing of pro-SGPA by active SGPB. Due to the low specificity of SGPB-specific polyclonal antibodies and to the low concentration of inactive SGPB in the various fractions, no conclusions could be made from these attempts.

**4.10. Intramolecular Chaperone**

A molecular chaperone is a protein required for the post-translational folding, targeting or assembly of other proteins but which does not itself form part of the final assembled structure. Inouye coined the term “intramolecular” chaperone to describe the function of the subtilisin E pro-peptide in the production of active subtilisin E. The pro-peptide, covalently linked to the amino terminal end of the subtilisin E, intramolecularly guides the folding of subtilisin E into the active enzyme.
The pro-peptide is not required for proteolytic activity and is removed by autoprocessing when protein folding is complete. Subtilisin E was found to be unable to properly fold itself in the absence of its pro-region and the pro-regions from the related subtilisins were only partially able to function as intramolecular chaperone.

The evidence presented for SGPB indicates that this protease also requires the presence of its pro-region for proper folding and activity. Expression of SGPB constructs without the pro-region in either *B. subtilis* or *E. coli* resulted in an inactive polypeptide. Only when the pro-region was present was SGPB folded properly and able to cleave peptide bonds.

The concept of the intramolecular chaperone is an important one, especially when noting that its function is retained even when separate from the protein it helps fold, suggesting a possible evolutionary relationship between the pro-region and the more general class of chaperones found in all organisms.

4.11. Conclusions and Further Studies

Several points have been determined as a result of the attempts at expressing *S. griseus* proteases and their variants:

1. Certain *S. griseus* genes can be expressed in *B. subtilis*.
2. Expression of the *S. griseus* proteases in *B. subtilis* is more effective through the use of genus specific signals, including a *Bacillus* promoter and pre-region.
3. There exists a directional dependence of expression in the *B. subtilis* plasmid pUB110, even when using genus specific signals.
4. The pro-region of the *S. griseus* proteases is important in the production of active enzyme.
5. The pro-region of *B. amyloliquefaciens* subtilisin BPN' need not be contiguous with the mature region of the protease to assist in correct folding of the mature enzyme.
6. During expression in *B. subtilis*, the presence of active SGPB results in the increase in protease activity in cultures containing SGPA-constructs.
To clarify some points, further work is required, including the following:

1. It is necessary to determine what growth conditions affect secretion of the proteases in an effort to optimize expression. Temperature has shown to be a factor in the secretion of active proteases (especially for SGPA), as well as the amount of oxygen available to the growing cultures.\textsuperscript{121} Growth of the bacteria in the controlled conditions of the fermentor will assist in these studies.

2. As a result of the discovery of the importance of oxygen levels in the expression of SGPA,\textsuperscript{121} the spiking experiments involving SGPA need to be repeated under controlled conditions.

3. Further experiments involving the use of polyclonal antibodies (of better antigenic specificity) or monoclonal antibodies are necessary to determine the fate of the SGPB-containing polypeptides, especially in the expression of the inactive variants.

4. Once successful expression of the inactive variants has been determined, experiments can be performed to enhance this process, either by use of another protease to assist in the maturation of the pre-pro-mature polypeptide or by expression of the pro- and mature-regions of the protease separately, within the same bacterium. Success in this endeavor has been achieved in the expression of inactive variants of \(\alpha\)-lytic protease.
Appendix: Plasmid Construction

A.1. Bacillus Expression

A.1.1. Pre-pro-subtilisin BPN′

Initial attempts at expression of *Streptomyces griseus* protease B (SGPB) in *Bacillus* consisted of linking the coding sequence for the mature region of SGPB to the coding region for the subtilisin BPN′ pre-pro-region. To facilitate DNA manipulation, the entire subtilisin BPN′ gene (including 5′- and 3′-non-coding sequences) was isolated as a 3.4 kb *Eco*RI fragment of the *Bacillus* plasmid pPT30 and ligated into the unique *Eco*RI site of the *E. coli* plasmid pUC19 to create the plasmid pBSS.

The plasmid pMSi, as provided by Cangene Corporation (Mississauga, Ontario), contained the mature SGPB missing the first 21 nucleotides. The mature region was reconstructed by isolating the *BglII/Bam*HI fragment of pMSi and attaching the linkers.
(oligonucleotides) REC1 and REC2, designed to link the coding region of the mature SGBP to the pre-pro-region of subtilisin BPN' at the SplI site of subtilisin BPN'. The linkers were degenerate in one codon allowing for translation of glycine, alanine or valine at the beginning of the mature region of SGBP.

\[
\text{REC1} \quad 5'\text{-GTACGXGATCTCCGGCGGCGACGC}\quad \text{where } X=G,C,T \\
\text{REC2} \quad \text{CXCTAGAGGCCGCGCTGCCTAG-5'} \quad \text{where } X=C,G,A
\]

Once the linkers were attached, the mature region of SGBP was ligated into the SplI site of pBSS to create the plasmid pTB5.

Upon sequencing of pTB5, the degenerate nucleotide was found to be deleted.

Plasmid pTB5 was used as a template for the polymerase chain reaction (PCR) with the
oligonucleotide REC1 and the product was blunt-ended with T4 DNA polymerase and ligated into the *SmaI* site of pUC19 to create the plasmid pKN7.

Sequencing of the repaired synthetic linker region of SGBP determined the first codon to be GCG which coded for the amino acid alanine.

The coding region for the mature portion of SGBP was placed behind the pre-pro-region of subtilisin BPN' by digesting pKN7 with *Sphi* and ligating the resulting fragments into the *Sphi* site of pPT30. Two plasmids resulted: pKN1, containing the entire coding sequence for the mature region of SGBP, including the 3'-non-coding region; and pDC1 containing the coding sequence for the mature region of SGBP missing the last 18 nucleotides of the coding sequence. In pDC1 the sequence of mature SGBP was in-frame with the mature subtilisin BPN' sequence and allowed for the translation of mature subtilisin BPN' following the mature SGBP.
The oligonucleotide linkers FOR-END-SGBP and REV-END-SGBP were used to reconstruct the last 21 nucleotides of the coding sequence for the mature portion of SGBP in pDC1. The resulting plasmid, pGK4a, contained the entire coding sequence for the mature region of SGBP inserted in frame at the SplI site of subtilisin BPN'.

FOR-END-SGBP 5'-GTACGGCGTCAGCGT
REV-END-SGBP CCGCAGTCGCACATG-5'
SplI SplI

A.1.2. Pre-pro-mature Protease

Further attempts at production of SGBP and the related S. griseus protease A\(^3\) (SGPA) in Bacillus were made by expressing the two proteases under the control of their own promoters and with their own pre-pro-regions. This was accomplished by using the PCR to isolate the necessary gene sequences (work done in this lab by Sachdev Sidhu with assistance form Kevin Neden).

A.1.2.1. PCR

SGPA

The entire coding sequence (including the 5’- and 3’-non-coding regions) for SGPA was isolated from S. griseus IMRU3499\(^3\) using the PCR and the following oligonucleotides:
SGPA-5’-end
5' ACCCCCATCTCATCATTCCGG
SGPA-3’-end
5' ACTCTCGGCAACTTCGCG

SGPB

The coding sequence (including the 5’-non-coding region) for the pre-pro-region and for part of the mature region (101 nucleotides 3’ to the NaeI site) for SGPB was isolated from *S. griseus* IMRU3499\(^1\) using the PCR and the following oligonucleotides:

SGPB-5’-end
5' CCGAAAGGCATTCTTGCG
SGPB-3’-end (NaeI)
5' CCGTAGTCGTTGGTTCG

A.1.2.2. Plasmid Construction

The products of the above PCR reactions were blunt-ended using T4 DNA polymerase and ligated separately into the *SmaI* site of pUC19 to create the plasmids pKN3A (containing pre-pro-mature-SGPA) and pKN3B (containing pre-pro-mature-SGBP-partial).

![Plasmid diagrams](image)

To reconstruct the complete mature coding region of SGBP, the *NaeI/BamHI* of pMSi was substituted for the *NaeI/BamHI* fragment of pKN3B to create the plasmid pKN3B-R.
For expression in *Bacillus*, the coding regions for SGPA and SGPB were ligated into the *Bacillus* plasmid pUB110. The T4 DNA polymerase blunt-ended *SstI/HindIII* fragment of pKN3A, containing the entire coding sequence for SGPA, was ligated into the *PvuII* site of pUB110 to create the plasmids pPA1 and pPA2. pPA1 contains the SGPA fragment in a counterclockwise orientation relative to pUB110 and pPA2 contains the SGPA fragment in a clockwise orientation relative to pUB110.

The T4 DNA polymerase blunt-ended *KpnI/HindIII* fragment of pKN3B-R, containing the entire coding sequence for SGPB, was ligated into the *PvuII* site of
pUB110 to create the plasmids pPB1 and pPB2. pPB1 contains the SGBP fragment in a counterclockwise orientation relative to pUB110 and pPB2 contains the SGBP fragment in a clockwise orientation relative to pUB110.

A.1.3. Pre-subtilisin BPN'-pro-mature Protease

Further attempts at production of SGPA and SGBP in Bacillus were made by expressing the two proteases under the control of the subtilisin BPN' promoter with the subtilisin BPN' pre-region fused to the pro-mature regions of the S. griseus proteases. This was accomplished by using the PCR to isolate the necessary fragments.

A.1.3.1. PCR

Pre-subtilisin BPN'

The coding sequence (including the 5'-non-coding region) for the pre-region of subtilisin BPN' was isolated from the plasmid pPT30 using the PCR and the following oligonucleotides:
The first oligonucleotide was designed to place \textit{BamHI} and \textit{EcoRV} restriction sites at the 5' end of the fragment. The second oligonucleotide was designed to place \textit{NaeI} and \textit{XbaI} restriction sites at the 3' end of the fragment. Cleavage at the \textit{NaeI} site would leave a blunt end at the 3' end of the subtilisin BPN'-pre-coding sequence without changing the amino acid sequence of the pre-region.

**Pro-mature-SGPA**

The coding sequence (including the 3'-non-coding region) for the pro-mature region of SGPA was isolated from pKN3A using the PCR and the following oligonucleotides:

$$\text{pro-mature-SGPA: 5'-end} \quad 5'-\text{GGCTGCAGCTCCCGAGGCG}$$

$$\text{SmaI} \quad \text{PstI}$$

$$(\text{CCCGGG})$$

The first oligonucleotide was designed to place \textit{SmaI} and \textit{PstI} restriction sites at the 5' end of the fragment. Cleavage at the \textit{PstI} site followed by blunt-ending with T4 DNA polymerase would leave a blunt end at the 5' end of the pro-mature-SGPA coding sequence without changing the amino acid sequence of the pro-region.

**Pro-mature-SGBP**

The coding sequence for the pro-region and for part of the mature region (101 nucleotides 3' to the \textit{NaeI} site) for SGBP was isolated from pKN3B using the PCR and the following oligonucleotides:
The first oligonucleotide was designed to place KpnI and FspI restriction sites at the 5' end of the fragment. Cleavage at the FspI would leave a blunt end at the 5' end of the pro-mature-SGPB coding sequence without changing the amino acid sequence of the pro-region.

### A.1.3.2. Plasmid Construction

The products of the above three PCR reactions were blunt-ended with T4 DNA polymerase and ligated separately into the SmaI site of pUC19 to create the plasmids pDC3+ (containing pre-subtilisin BPN'), pDC4+ (containing pro-mature-SGPA) and pDC5+ (containing pro-mature-SGPB-partial).

To reconstruct the complete mature coding region of SGPB, the NaeI/BamHI of pMSi was substituted for the NaeI/BamHI fragment of pDC5+ to create the plasmid pDC5R+.
To construct a general secretion plasmid in *Bacillus*, the *Bam*HI fragment of pDC3+, containing the coding region for the subtilisin BPN' pre-region, was ligated into the *Bam*HI site of PUB110 to create the plasmids pbDC6+ and pbDC6-. pbDC6+ contains the subtilisin BPN' fragment in a clockwise orientation relative to PUB110 and pbDC6- contains the subtilisin BPN' fragment in a counterclockwise orientation relative to PUB110.

The coding sequences for the pro-mature regions of SGPA and SGPB were ligated behind the subtilisin BPN' pre-region in the plasmid pbDC6-. The *Pst*I(blunt-
ended with T4 DNA polymerase)/XbaI fragment of pDC4+, containing the coding sequence for pro-mature-SGPA, was ligated into pbDC6- digested with NaeI and XbaI to create the plasmid pbDC8-. The FspI/XbaI fragment of pDC5R+, containing the coding sequence for pro-mature-SGPB, was ligated into pbDC6- digested with NaeI and XbaI to create the plasmid pbDC7-. In both constructs, a fragment of 324 nucleotides was removed from the Mob gene on pUB110.

The coding sequences for the pro-mature regions of SGPA and SGPB were also ligated behind the subtilisin BPN' pre-region in E. coli constructs prior to transfer to Bacillus plasmids. The PstI(blunt-ended with T4 DNA polymerase)/XbaI fragment of pDC4+, containing the coding sequence for pro-mature-SGPA, was ligated into pDC3+ digested with NaeI and XbaI to create the plasmid pDC9+.

The FspI/XbaI fragment of pDC5R+, containing the coding sequence for pro-mature-SGPB, was ligated into pDC3+ digested with NaeI and XbaI to create the plasmid pDC10+.
To be expressed in *Bacillus*, the fragments containing the pre-subtilisin BPN'-pro-mature-proteases were ligated into pUB110. The *BamHI* fragment of pDC9+, containing the entire coding sequence for pre-subtilisin BPN'-pro-mature-SGPA, was ligated into the *BamHI* site of pUB110 to create the plasmids pbDC12+ and pbDC12-. pbDC12+ contains the SGPA fragment in a clockwise orientation relative to pUB110 and pbDC12- contains the SGPA fragment in a counterclockwise orientation relative to pUB110.

The *BamHI* fragment of pDC10+, containing the entire coding sequence for pre-subtilisin BPN'-pro-mature-SGPB, was ligated into the *BamHI* site of pUB110 to create the plasmids pbDC13+ and pbDC13-. pbDC13+ contains the SGPB fragment in a clockwise orientation relative to pUB110 and pbDC13- contains the SGPB fragment in a counterclockwise orientation relative to pUB110.
A.2. *E. coli* Expression

A.2.1. Pro-mature Protease

pMAL-p and pMAL-c

The coding sequences for the pro-mature regions of SGPA and SGPB were ligated behind the maltose binding protein (MBP) sequence on each of the plasmids pMAL-c and pMAL-p.\(^ {71,72} \) When induced, bacteria harboring pMAL derivatives express fusion protein consisting of the MBP linked to the desired protein by the tetrapeptide recognition site for Factor Xa. pMAL-p derivatives directed the fusion protein to the periplasm of the cell due to the presence of the MBP signal peptide sequence. pMAL-c derivatives which lack the leader sequence produced a cytoplasmic protein.

pMAL-pA and pMAL-cA

The *Pst*I (blunt-ended with T4 DNA polymerase)/*Xba*I fragment of pDC4+, containing the coding sequence for pro-mature-SGPA, was ligated into pMAL-p and pMAL-c digested with *Stu*I and *Xba*I to create the plasmids pMAL-pA and pMAL-cA.
pMAL-pB and pMAL-cB

The *FspI/XbaI* fragment of pDC5R+, containing the coding sequence for pro-
mature-SGBP, was ligated into pMAL-p and pMAL-c digested with *StuI* and *XbaI* to
create the plasmids pMAL-pB and pMAL-cB.
The coding sequence for the pro-mature region of SGPB was ligated behind the coding region for the cII protein on the *E. coli* plasmid pLcII-FX. When induced, bacteria harboring pLcII-FX derivatives express fusion protein consisting of the first 31 amino acids of the lambda (λ) cII protein linked to the desired protein by the tetrapeptide recognition site for a protease isolated from the blood coagulation cascade, Factor Xa. The *BamHI* fragment of pMAL-pB (containing FX-pro-mature-SGPB) was ligated into the *BamHI* site of pLcII-FX to create the plasmid pLcII-FX-pro-SGPB.
A.3. SGPB Inactivation

A.3.1. DNA Template

The DNA sequence coding for the mature portion of SGPB was ligated into an M13 plasmid such that single-stranded template DNA for mutagenesis could be isolated. Initially the FX-SGPB-containing BamHI fragment of pLcII-FX-SGPB was ligated into the BamHI site of M13mp19 to create the plasmid M13mp19-FX-SGPB.
As mutagenesis using this plasmid was unsuccessful, a second template was created, M13mp19-FX-SGPB(HincII) which contained the HincII fragment (3' portion of mature-SGPB) of M13mp19-FX-SGPB ligated into the HincII site of M13mp19.

**A.3.2. Mutagenesis**

To inactivate SGPB, mutagenesis was performed by the Kunkel method using the forward universal sequencing primer and the following mutagenic oligonucleotide:

SG195 5' CCCGCGACGGTGGCGGCCCG

The oligonucleotide SG195 hybridized to the sequence surrounding the codon for Ser-195 of SGPB and directed the change of the DNA sequence from TCC (coding for serine) to GGT (coding for glycine) at position 195. The change on the resulting plasmid, M13mp19/SG195, was confirmed by DNA sequencing of the single-stranded template. To complete sequencing of the entire fragment, the HincII fragment of M13mp19/SG195 was ligated into the HincII site of pUC18 to create the plasmid pSG2.
A.3.3. Inactive SGPB Reconstruction

To express the inactive SGPB protein, the coding region for the mature protease was reconstructed. The *HincII* fragment of pSG2 was ligated into the T4 DNA polymerase blunt-ended *HincII/KpnI* fragment of pMSi to create the plasmid pSG-R.

A.3.4. Bacillus Expression

To express the inactive SGPB protein in *Bacillus*, the *DraIII/XbaI* fragment of pSG-R containing the SG195 mutation was ligated into pDC10+ digested with *DraIII* and *XbaI* to create the plasmid pDC11+.
The BamHI fragment of pDC11+, containing the entire coding sequence for pre-subtilisin BPN'-pro-mature-SGPB/SG195, was ligated into the BamHI site of pUB110 to create the plasmids pbDC14+ and pbDC14-. pbDC14+ contains the SGPB fragment in a clockwise orientation relative to pUB110 and pbDC14- contains the SGPB fragment in a counterclockwise orientation relative to pUB110.
A.3.5. *E. coli* Expression

**pMAL-pB and pMAL-cB**

To express the inactive SGPB protein in *E. coli* using the pMAL plasmids, the *DraIII/XbaI* fragment of pSG-R containing the SG195 mutation was ligated into the pMAL-pB and pMAL-cB plasmids digested with *DraIII* and *XbaI* to create the plasmids pMAL-pB/SG195 and pMAL-cB/SG195.

**pLcII-FX-pro-SGPB**

To express the inactive SGPB protein in *E. coli* using the pLcII-FX plasmid, the *DraIII/XbaI* fragment of pSG-R containing the SG195 mutation was ligated into the pLcII-FX-pro-SGPB plasmid digested with *DraIII* and *XbaI* to create the plasmid pLcII-FX-pro-SGPB/SG195.
DraIII digest

DraIII/XbaI digest

DraIII/XbaI digest
List of References

   antisignals in proteolysis, pp 137-149. In Biogenetics of Neurohormonal Peptides,


   reaction of chymotrypsin and diisopropylphosphorofluoridate. II. The structure of
   two DP-substituted peptides from chymotrypsin-DP. Biochim. Biophys. Acta
   27:556-563.

   esterases. I. Rates of reaction with acetylcholinesterase, α-chymotrypsin, and

7. Schoellman, G., and Shaw, E. E. (1963) Direct evidence for the presence of


    protonation states of aspartic acid-102 and histidine-57 in the tetrahedral
    intermediate of the serine proteases: neutron structure of trypsin. Biochemistry
    20:6462-6474.

    Re-examination of the charge relay system in subtilisin and comparison with other

    8:171-176.

    Fla.


87. Trigo Gonzalez, G., personal communication.


121. Sidhu, S., personal communication.