CHARACTERISATION OF *STREPTOMYCES GRISEUS* PROTEASE PROPEPTIDE MUTANTS *IN VIVO*

by

Jason Baardsnes

B.Sc. (Honours) Simon Fraser University, 1993

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

in the Department of Chemistry

© Jason Baardsnes 1996

SIMON FRASER UNIVERSITY

April 1996

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

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395 rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

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

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-16780-1

Canada

APPROVAL

Name: Jason Baardsnes

Degree: Master of Science

Title of Thesis: Characterisation of *Streptomyces griseus* Protease Propeptide Mutants *In Vivo*

Examining committee:

Chair: Dr. R. B. Cornell

Dr. T. J. Borgford (Associate Professor)
Senior Supervisor

Dr. A. Bennet (Assistant Professor)
Committee Member

Dr. D. Sen (Assistant Professor)
Committee Member

Dr. C.M. Boone (Assistant Professor)
Internal Examiner

Date Approved: April 11, 1996

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:

Characterisation of Streptomyces griseus Protease Propeptide Mutants In Vivo

Author:

(signature)

Jason Baardsnes

(name)

11 April 1996

(date)

ABSTRACT

Streptomyces griseus is a soil bacterium that secretes several serine proteases, each of which are encoded by a gene that expresses a pre-pro-mature enzyme. The precursor is processed during secretion into a mature, fully folded protease. Maturation involves the proteolytic cleavage of bonds at the junction of the pre and pro domains, and pro and mature domains. The latter cleavage is an autocatalytic process and occurs after the propeptide directs the folding of the precursor polypeptide into active, mature protease. In this study, a chimeric gene consisting of the propeptide encoding region of *Streptomyces griseus* protease A (SGPA) fused directly to the mature encoding domain of *Streptomyces griseus* protease B (SGPB) was constructed and found to generate 6.0 % wild type SGPB activity when expressed in *Bacillus subtilis*. This demonstrates that the propeptide of SGPA can correctly fold mature SGPB. A conserved amino terminal α-helix was found in the propeptides of SGPA and SGPB and targeted for deletion analysis in SGPB. The greater the size of the deletion, the greater was the loss in protease activity indicating the deleted region is important for the folding process. Surprisingly the activity of the secreted enzyme decreased in an exponential manner according to the length of the propeptide deletion. These results found that the propeptide stabilises the folding transition state of the mature enzyme by 1.0 kcal mol^{-1} for every 5 amino acids through many small binding interactions over the entire surface of the propeptide as opposed to interactions involving a few crucial residues. Northern blot analysis demonstrated that all of the mutant genes were transcribed at the same level as the wild type gene for *Streptomyces griseus* protease B. Amino terminal analysis of SGPB generated by the chimeric gene and a 10 amino acid propeptide deletion mutant found that they both underwent wild type processing at the pro-mature junction.

This indicates the observed decrease in protease activity was due to a decrease in the number of correctly folded protease molecules, and not because of misfolded protease with reduced catalytic activity.

For my mom and dad

I would like to thank my senior supervisor Dr. Thor J. Borgford
for giving me this opportunity (and my first airplane trip)

and

Cheers to the other grad students in the lab for all of their
help, friendship and support:
Kathy, Jeff, Dev, Louise, and Grace

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

LIST OF ABBREVIATIONS

Amp	Ampicillin
ATP	Adenosine Triphosphate
BSA	Bovine Serum Albumin
CD	Circular dichroism
CIP	Calf Intestinal Phosphatase
DEPC	Diethyl Pyrocarbonate
ddH_2O	Distilled, deionized water
DHFR	Dihydrofolate reductase
EDTA	Ethyenediaminetetraacetic acid
EtOH	Ethanol
FPLC	Fast protein liquid chromatography
IPTG	Isopropyl-β-galactoside
kan	Kanamycin
kbp	Kilobase pair
kcal	Kilocalorie
mat	Mature
mol	Mole
MOPS	3-(N-morpholino) propanesulfonic acid
MW	Molecular weight
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
pro	Propeptide
RNase A	Ribonuclease A
RNase T_1	Ribonuclease T_1
SDS	Sodium dodecyl sulfate

SGPA	*Streptomyces griseus* protease A
SGPB	*Streptomyces griseus* protease B
SGPC	*Streptomyces griseus* protease C
SGPD	*Streptomyces griseus* protease D
SGPE	*Streptomyces griseus* protease E
SMM	Spizizen's minimal media
SMMP	Spizizen's minimal media containing penassay
TAE	Tris-acetate buffer containing EDTA
TEMED	N,N,N',N'-tetramethylenediamine
Tris	Tris (hydroxymethyl) methylamine
TY	Tryptone yeast
UV	Ultra violet

1. INTRODUCTION

1.1 THE PROTEIN FOLDING PROBLEM

One of the largely unanswered questions in biochemistry is how a protein's linear sequence of amino acids dictates the final three dimensional conformation of the functional protein. A peptide of n +1 residues with m equally probable conformations per residue will give m^n possible final three dimensional conformations per peptide (1). A protein with 100 residues and an average of 8 conformations per residue would have 8^{99} or about 10^{89} possible final conformations. Levinthal (2) calculated that it would require 10^{66} years to sample all of the possible conformations at the diffusion controlled rate limit of 10^{13} s^{-1}. Proteins fold on time scale of seconds, therefore, in order for a protein to reach it's final conformation, it must scan the total conformational space and choose a protein folding pathway that uses only a small subset of the total possible protein configurations. How a protein accomplishes this process is an intriguing question and is yet largely unanswered.

A class of proteins called molecular chaperones that aids in the folding of proteins *in vivo.* has been identified. In the presence of ATP, chaperone proteins such as the Hsp60 class bind to the exposed hydrophobic surfaces of unfolded proteins in order to prevent off-pathway folding events such as aggregation (3, 4) The chaperones assist protein folding by isolating incompletely folded proteins in large multi-subunit complexes, thus producing an optimum environment for protein refolding. The chaperone protein does not directly interact with the folding proteins to increase the forward folding reaction.

Another class of proteins are synthesized as precursor forms with an amino terminal domain called an intramolecular chaperone or propeptide. This domain is common to bacterial proteases (5, 6) and acts as a single turnover catalyst to guide the folding of the mature polypeptide to the native state. In contrast to molecular chaperones, propeptide mediated folding does not require an exogenous energy source such as ATP, and the propeptide directly increases the forward folding reaction instead of preventing off-pathway events. The propeptide is therefore used by many research groups as a model system for research to better understand protein folding.

1.2 PRE-PRO-MATURE POLYPEPTIDE PRECURSORS IN BACTERIAL SERINE PROTEASES

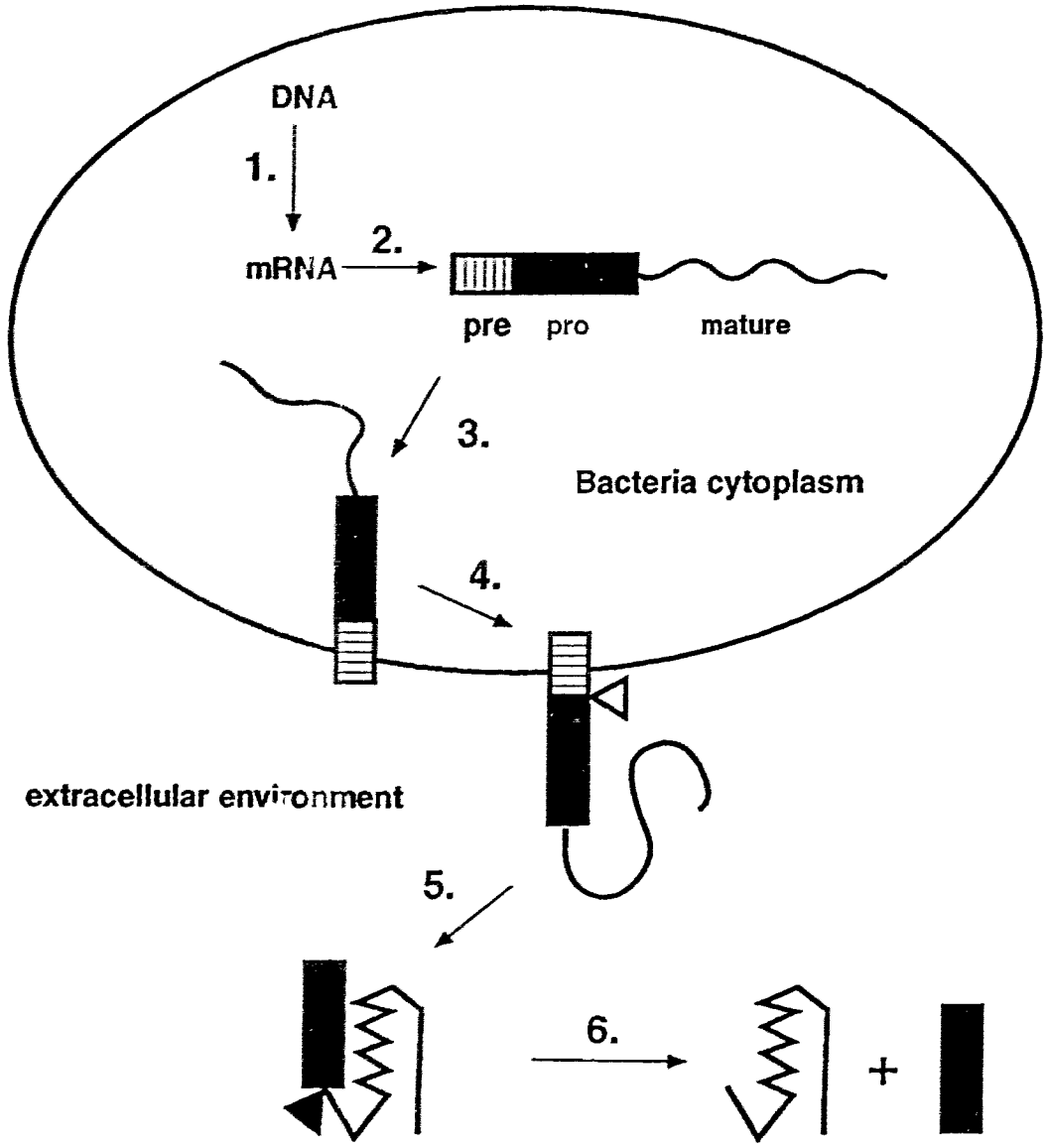
Investigators have used the bacterial serine proteases α-lytic protease from *Lysobacter enzymogenes* (7), subtilisin E from *Bacillus subtilis* (8) and subtilisin BPN' from *Bacillus amyloliquefaciens* (9) as model proteases to study propeptide assisted protein folding. We investigated the family of serine proteases produced by *Streptomyces griseus* (10, 11, 12, 13). All of the proteases mentioned were chosen for investigation because they were well characterised and are single subunit proteins with relatively low molecular weights less than 30,000. The proteases are secreted as pre-pro-mature polypeptide precursors, though only the mature region is detected in culture supernatants. The pre and pro regions are transient amino terminal domains that are hydrolysed at different stages of protease maturation to generate mature protease. The secretion process in gram positive bacteria such as *S. griseus* and *B. subtilis* which have only a single cell membrane is outlined in Figure 1. In gram positive bacteria such as *E. coli* and *L. enzymogenes*, the

process is essentially the same except the protease is secreted into the periplasm, a region between the inner and outer membranes of the bacterium.

1.2.1 The Function of the Signal Sequence

The pre region of the polypeptide is a standard bacterial signal sequence that directs the nascent enzyme to the cell surface for secretion (14). Bacterial signal sequences are typically about 30 amino acids long and comprised of an amino terminal charged region, a membrane-spanning stretch of 10 to 15 hydrophobic amino acids, and a carboxy terminal signal peptidase cleavage site (14). The signal peptidase cleavage site typically contains small neutral residues at positions -1 and -3 (15) where -1 is the final amino acid of the signal sequence before the first amino acid of the propeptide. Once the nascent polypeptide reaches the cell surface, the pre region is hydrolysed by signal peptidase at the pre-pro junction and the rest of the enzyme is released into the extracellular media. There are two types of signal peptidase, one specific for lipoprotein signal peptide (peptidase II), and one that hydrolyses all other signal peptides (peptidase I) (14). All signal peptidases are integral membrane proteins, with the bulk of the protein on the extracellular side of the membrane where the cleavage will occur. Signal peptides have a lack of homology, even between related

Figure 1. Typical Secretion Process of Pre-Pro-Mature Proteases in Gram Positive Bacteria 1) Transcription of the protease gene. **2)** Translation to form pre-pro-mature polypeptide precursor. **3)** Pre-domain directed targeting of the immature polypeptide to the cell membrane. **4)** Translocation of the immature polypeptide across the cell membrane. **5)** Cleavage at the pre-pro junction by signal peptidase (white arrow) and release of the pro-mature enzyme into the extracellular environment; pro-domain directed folding of mature protease. **6)** Cleavage at the pro-mature junction (black arrow) and release of mature protease.


DNA
1.
mRNA
2.
pre
pro
mature
3.
Bacteria cytoplasm
4.
extracellular environment
5.
6.
+

proteins (14). However, the overall structure of the signal sequence is the important factor for correct secretion, therefore, signal sequences can be easily exchanged between different proteins without losing any secretion capabilities (14). *E. coli* signal sequences have been correctly processed in *X. laevis* oocytes. The fact that a bacterial signal sequence can function correctly in a eukaryotic organism indicates that the secretion process has been highly conserved during evolution (14).

1.2.2 The Function of the Propeptide

As compared to the signal sequence, the function of the propeptide is more ambiguous, varying from enzyme to enzyme. Studies of the bacterial enzymes subtilisin and α-lytic protease reveal an essential role for the propeptide in the catalysis of the protein folding during secretion.

Subtilisin has a 77 amino acid propeptide that assists folding of a mature protease of 275 residues. Similarly, α-lytic protease has a 166 amino acid propeptide and a 198 residue mature protease. The creation of a construct containing the *phoA E. coli* signal sequence fused to the first amino acid of the mature α-lytic protease (i.e., missing the propeptide) resulted in no active protease being detected in *E. coli* (16). The mature region of subtilisin E was expressed in *E. coli* using a direct fusion with the ompA signal peptide (17). This construct produced an insoluble protein which when purified generated a mature protease size band on SDS-PAGE. Amino terminal sequencing verified the mature protease sequence with the signal sequence was correctly processed but active protease was not generated due to the absence of the propeptide. Each propeptide is specifically designed to fold it's own mature domain, the propeptides cannot be interchanged and

generate active protease at any significant levels (18). The propeptide of the bacterial proteases are removed by an autocatalytic process, but it is not known if this process is intramolecular (one molecule self-processing), intermolecular (one mature protease processing another zymogen), or if both intramolecular and intermolecular processing takes place (18, 19, 20).

These results stress the importance of the propeptide for folding in α-lytic protease and subtilisin, and contrast results of the propeptides of mammalian chymotrypsin-like enzymes which have no known role in protein folding. Mammalian propeptides hold the enzyme in an inactive conformation until cleaved from the zymogen by the action of enterokinase (21) With few exceptions the bacterial propeptides of the chymotrypsin-like enzymes are longer than their mammalian counterparts (22).

The bacterial proteases such as α-lytic protease and subtilisin that require the propeptide for folding provide a unique opportunity to study the folding process of the mature domain. The propeptide places the residues important for folding within a defined region making the analysis of the folding process an easier task.

1.3 CHARACTERISTICS OF PROPEPTIDE MEDIATED FOLDING IN α-LYTIC PROTEASE AND SUBTILISIN

1.3.1 α-Lytic Protease Propeptide

1.3.1.1 *Trans* Addition of Propeptide to Initiate α-Lytic Protease Folding

Once it was known that the propeptide was crucial for folding the mature region, the task of determining the folding mechanism began. The first *in vitro* studies with α-lytic protease involved the expression of the propeptide and mature protease region *in trans* from the plasmid pCOMP5 in *E. coli*. (23). Expression of the mature domain was driven by the *phoA* promoter and signal sequence which was induced under phosphate starvation and the propeptide was directly fused to the *lpp/lac* promoter and *ompA* signal sequence which was induced in the presence of IPTG (Figure 2). The two different promoters allowed each component to be expressed independently, when *E. coli* cells were grown under low-phosphate conditions without IPTG only the mature domain was expressed and did not generate any activity. After the addition of IPTG, the propeptide was expressed off of the *lpp/lac* promoter and α-lytic protease activity was detected which indicates the mature domain was correctly folded. This experiment was able to eliminate the "kinetic trapping model" of propeptide mediated folding which was dependent on the hydrolysis of a direct linkage between the pro and mature domains. In this model, the propeptide would lower the energy barrier to a folded pro-mature intermediate state, and only after the hydrolysis of the propeptide would the equilibrium shift to favour

the most thermodynamically stable product (23) (Figure 3a). Rather, the ability of the propeptide to fold the mature protease in *trans* demonstrated that a covalently linked propeptide is not required for generating the most thermodynamically stable product (23) (Figure 3b). Once the propeptide is removed, the enzyme is resistant to thermal degradation, autolysis and has a high level of structural rigidity, therefore the propeptide binds a transition state late in the folding pathway (23). This late transition state prevents the formation of inactive intermediates across low energy barriers from the final folded state (Figure 4).

This hypothesis was reinforced by the finding that propeptide will bind and inhibit mature protease *in vitro* which indicates the folding intermediate is very close to the mature structure at the end of the folding pathway (24). In this study, propeptide expressed as a fusion protein with glutathione-S-transferase (GEXPRO) using the pGEX-2T expression system. Denatured α-lytic protease was allowed to refold in the presence of stoichiometric amounts of GEXPRO and 39 % α-lytic activity was recovered, most of the activity lost was due to aggregation of the denatured protease. The inhibitory effects of the propeptide were revealed when GEXPRO was assayed with folded, mature α-lytic protease. A solution of 2.2 μM α-lytic protease retained less than 1 % of its activity when 2.7 μM GEXPRO was added, and 89 % activity was regained after incubation for 20 minutes with 0.1 μg/μL trypsin which degrades the GEXPRO bound to the active site. Kinetic studies of GEXPRO as an inhibitor found it to have a K_i of 0.5 - 2.0 × 10^{-10} M (24) which makes it the highest known affinity inhibitor for α-lytic protease. The protease *Streptomyces griseus* protease B is evolutionarily

Figure 2. Plasmid Map of the pCOMP Plasmid. The propeptide of α-lytic protease is produced as a fusion protein with the OmpA signal peptide and is controlled by the *lpp/lac* promoter. The α-lytic mature region is produced as a fusion with the phoA signal sequence and is controlled by the *phoA* promoter. IPTG induces expression off the lpp/lac promoter, and phosphate starvation induces expression off the *phoA* promoter. This figure is adapted from Silen and Agard (23).

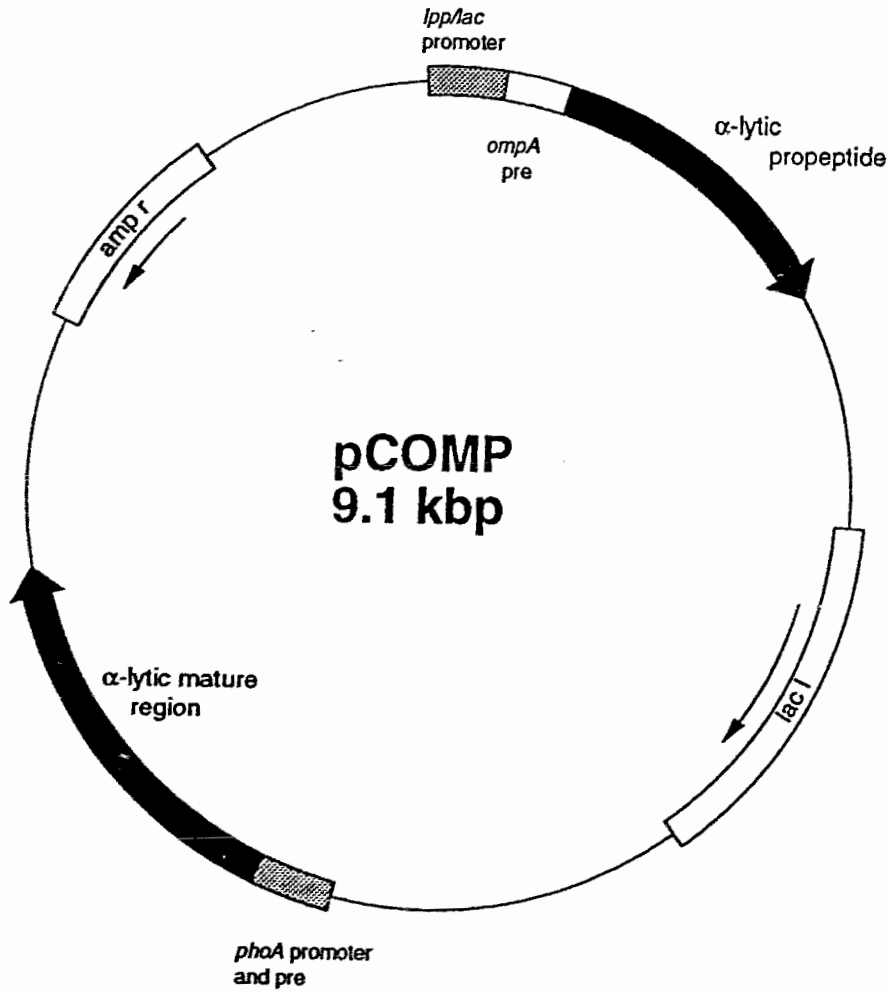
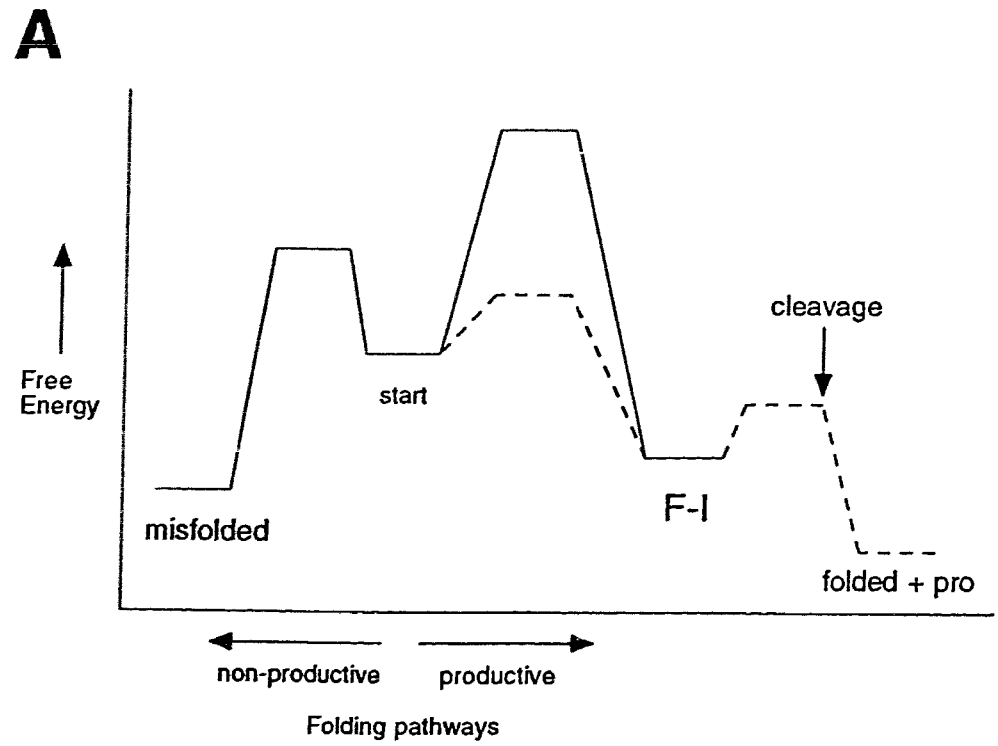

lpp/lac
promoter
ompA
pre
α-lytic
propeptide
amp r
pCOMP
9.1 kbp
α-lytic mature
region
lac I
phoA promoter
and pre

Figure 3. Propeptide Models. **A)** In this model, the presence of the propeptide lowers the energy barrier (dashed lines) to a nearly folded intermediate (F-I) which must have the propeptide hydrolysed from the mature peptide to shift the overall equilibrium to favour the folded native state (F) over the unfolded state (U). **B)** In this case, the propeptide stabilises the folding transition state to lower the energy barrier to favour the formation of the folded native state (F) over the unfolded state (U). The hydrolysis of the propeptide is not necessary for the generation of the most thermodynamically stable state. This figure is adapted from Silen and Agard (23).


A
Free Energy
start
cleavage
misfolded
F-I
folded + pro
non-productive
productive
Folding pathways

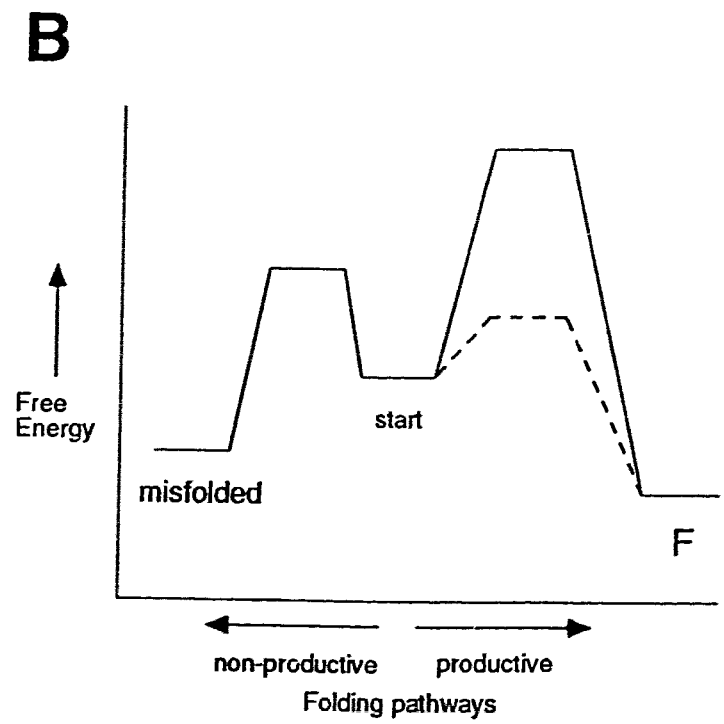
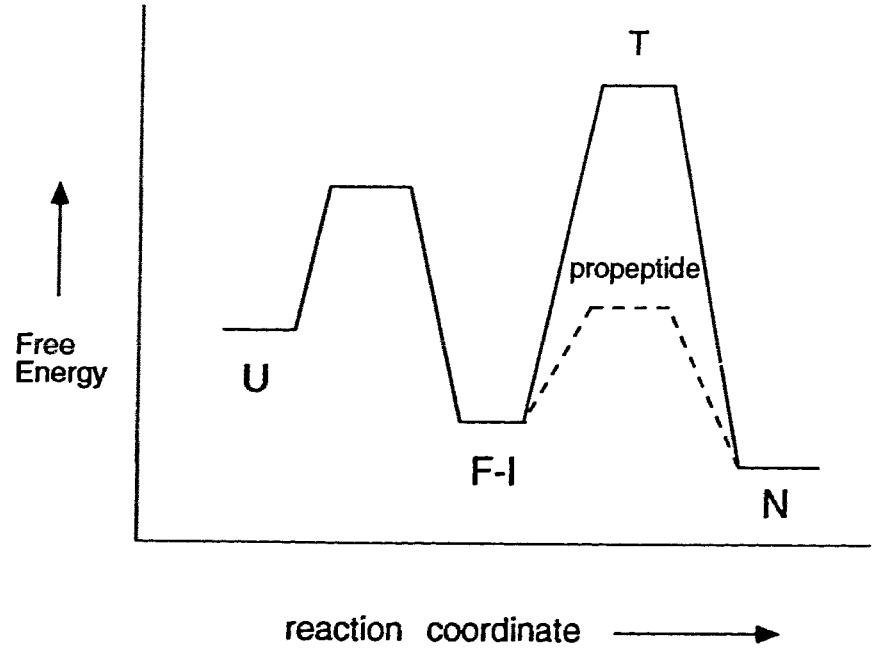

B
Free Energy
start
misfolded
F
non-productive
productive
Folding pathways

Figure 4. Free Energy Diagram for the Folding of α-Lytic Protease and Subtilisin. The propeptide lowers the transition state energy barrier (T) between the folding intermediate (F-I) and native (N) protease (dashed line). Once the propeptide is removed from the mature protease, the transition state energy barrier increases, making the folding intermediate inaccessible from the native state. The folding intermediate is reached from the unfolded (U) state of the protein. Refolding mature protease without the propeptide can trap it in a local minima at the folding intermediate.


T
propeptide
Free
Energy
U
F-I
N
reaction coordinate

related to α-lytic protease (37 % sequence homology) and was shown to be inhibited by GEXPRO with a K_i of 1.1 nM (24).

1.3.1.2 Isolation of an α-Lytic Protease Folding Intermediate

An inactive α-lytic protease folding intermediate was isolated and found to be as stable as native protease under identical conditions (25). The intermediate state was isolated by rapidly diluting guanidine denatured mature α-lytic protease without the propeptide 200 to 500 fold in buffer lacking guanidine but containing 100 μM AcProBoroVal α-lytic protease inhibitor (K_i = 3 μM). This protease solution contains the folding-competent intermediate state as seen by a 74 % recovery of activity with the addition of GEXPRO fusion protein (25) (Figure 5). Gel filtration chromatography using G-75 Sephadex revealed the intermediate state has an apparent Stokes radius of 28 Å which lies between the denatured state (≥ 41 Å) and native protease (20 Å) (25). Circular dichroism measurements on native, intermediate, and denatured states of the protease showed that the intermediate state contains almost as much β sheet as the native state (70 % compared to 75 % for wild type), and the denatured state contains no β sheet (25). Aromatic region CD and fluorescence spectroscopy indicated few tertiary interactions. The intermediate state was extremely stable, even after 800 hours of incubation at 4 °C no α-lytic protease activity could be detected and about 20 % of activity could be recovered on the addition of GEXPRO compared to addition of GEXPRO at time zero(25). From this data, it was approximated that the folding intermediate will convert to the native state in the absence of propeptide with a rate constant of $<10^{-9}$ s^{-1} which corresponds to an energy barrier in the order of 27 kcal mol^{-1} (25). This data suggests that the folding

intermediate is a long-lived state that has different enzymatic properties compared to the native state. The increased Stokes radius, near native like secondary structure, and lack of tertiary structure are characteristics of a molten globule state (26). In the presence of the propeptide, the folding reaction occurs with single turnover kinetics and an observed rate constant of 0.016 s^{-1}, a rate enhancement of 1.6×10^7. This data demonstrates that there are two thermodynamically stable states; the folding intermediate molten globule-like state and native state. The propeptide acts to lower the energy barrier to the final folded state, making it kinetically accessible from the intermediate. Once the propeptide is hydrolysed, the energy barrier back to the folding intermediate state will be raised making it inaccessible. The stability of the isolated folding intermediate *in vitro* highlights the fact that the folding pathway is under kinetic control, and not thermodynamic control. Kinetically controlled folding is pathway dependent, therefore intermediates can be isolated that lie on the folding path unlike thermodynamically controlled folding where the folding process is pathway independent and no intermediates can be isolated (27) (Figure 6).

Further elucidation of the propeptide of α-lytic protease has been thwarted by a lack of a pro-mature crystal structure and by the inability of the propeptide to accommodate mutations for analysis. Deletions of more than 5 amino acids to the propeptide abolished all protease activity within *E. coli* and in the extracellular medium (28).

Figure 5. *In Vitro* Refolding Experiments with Pro-Mature Proteases. *In vitro* refolding experiments with α-lytic protease and subtilisin used purified mature protease or pro-mature protease from inclusion bodies in *E. coli.* The proteins were solubilised in denaturant such as guanidinium-HCl, therefore the pro-mature mutants could not process the propeptide. Refolding was initiated by removing the denaturant by dilution or dialysis. Refolding of denatured mature protease was initiated in the presence of purified propeptide. Protease mutants were created that have active site mutations that inactivate the catalytic triad. This was done to ensure refolded protease does not degrade any unfolded protease. Inactive pro-mature protein can refold, however, the propeptide cannot be processed.

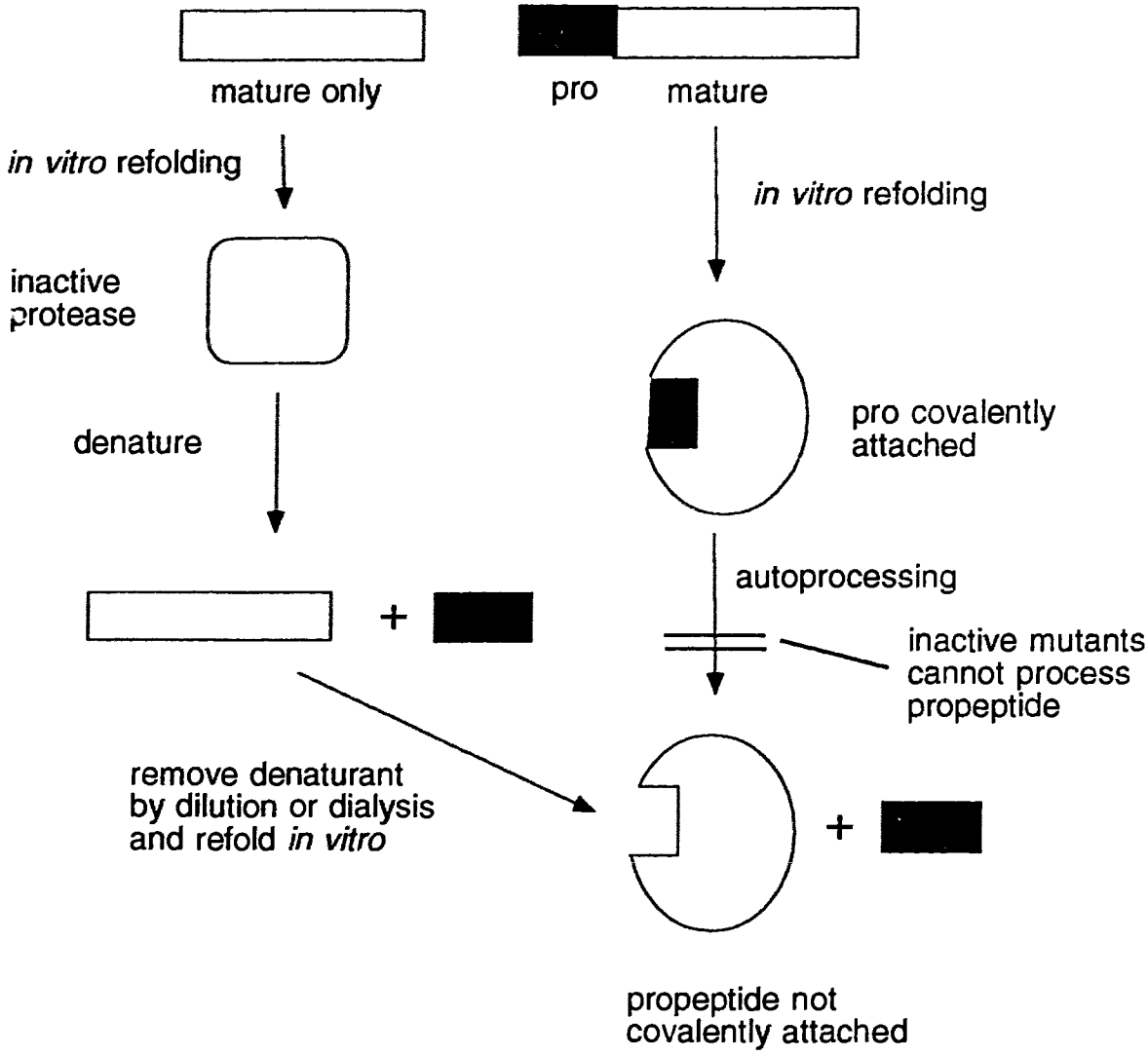
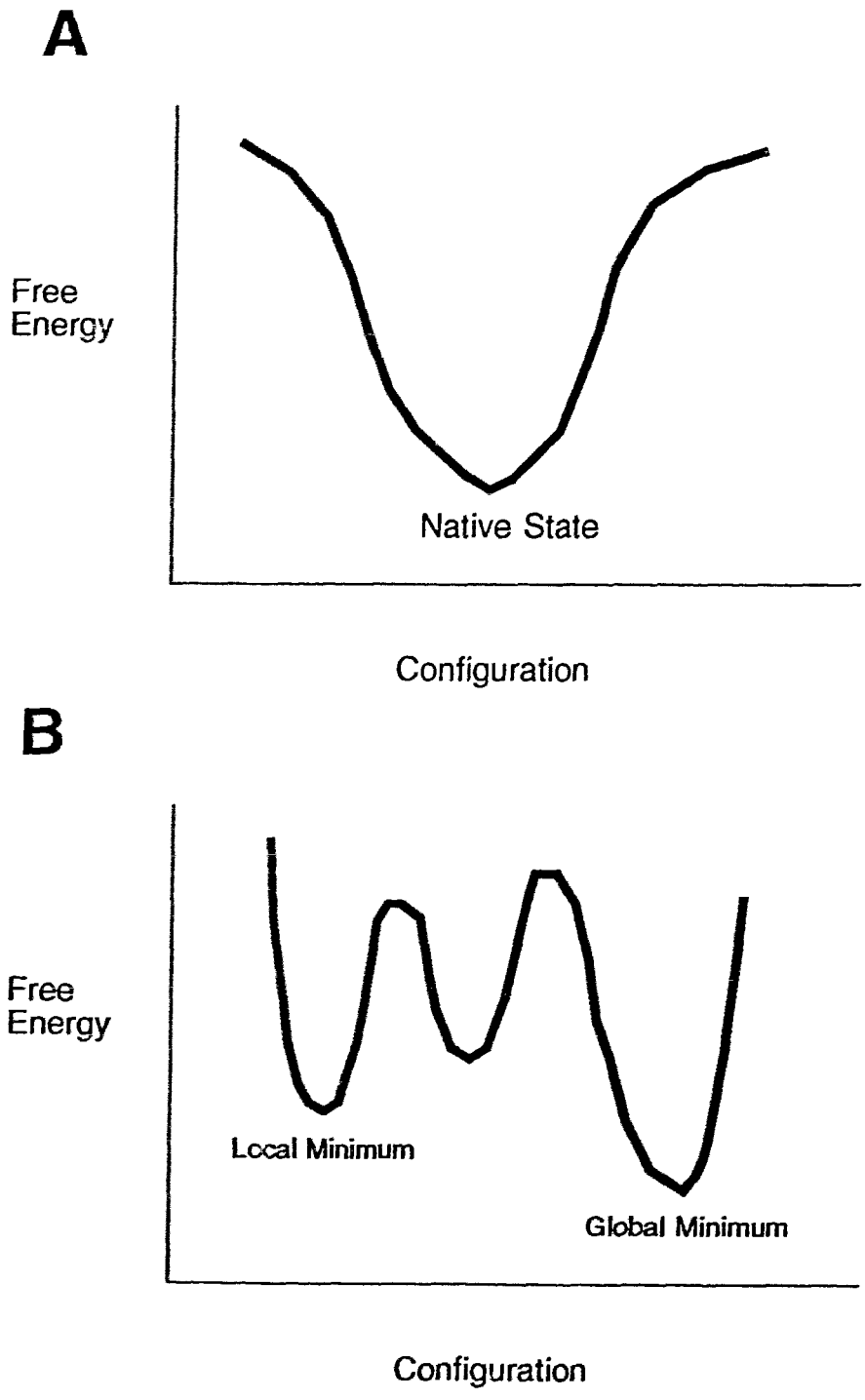

mature only
pro
mature
in vitro refolding
in vitro refolding
inactive protease
denature
pro covalently attached
autoprocessing
+
inactive mutants cannot process propeptide
remove denaturant by dilution or dialysis and refold in vitro
+
propeptide not covalently attached

Figure 6. Kinetic Versus Thermodynamic Control of Folding. **A)** A folding protein under thermodynamic control can start with any conformation and end at the most thermodynamically stable state. **B)** A folding protein under kinetic control is pathway dependent. For example, if the folding polypeptide started on the left, it could end up trapped in a local energy minimum if the energy barriers were high enough. Such is the case with the folding intermediate. The propeptide is needed to lower the free energy of the folding transition state in order for the native state to be reached. This figure was adapted from Baker and Agard (27).


A
Free
Energy
Native State
Configuration
B
Free
Energy
Local Minimum
Global Minimum
Configuration

1.3.2 Subtilisin Propeptide

1.3.2.1 *Trans* Addition of Propeptide to Initiate Subtilisin Folding

The propeptide of subtilisin was found to behave in a similar manner as the propeptide from α-lytic protease. The research into the subtilisin has paralleled that of α-lytic protease and similar results have been found even though the subtilisin class of enzymes and α-lytic protease are evolutionarily unrelated (22). Like α-lytic protease, the requirement for the propeptide for folding was established with the expression studies where a *E. coli* ompA signal sequence fused directly to the subtilisin E mature region failed to produce any detectable activity *in vivo* (17).

In order to analyse the ability of the propeptide to fold the mature domain in *trans*, an inactive pro-mature mutant of subtilisin E was created that contains the active site aspartic acid to alanine mutation (29). The inactive mutant is required to ensure that refolded protease does not degrade unfolded mature polypeptide. This mutant and the mature region of subtilisin E were produced as direct fusions to the ompA signal sequence in the *E. coli* pIN-ompA$_3$ expression vector that is controlled by IPTG induction off the *lpp* promoter. Both proteins were correctly processed by signal peptidase and purified from the periplasmic space using a Tris buffer containing guanidine-HCl as a denaturant. Various concentrations of the pro-mature inactive mutant were mixed with the mature domain, allowed to incubate at 4 °C, and then dialysed against renaturing buffer lacking denaturant (Figure 5). It was found that the pro-mature protein was able to refold denatured mature protease, thus demonstrating that the propeptide

can refold the mature domain *in trans*. The maximum specific activity recovered was about 20 % of wild type subtilisin E. In an attempt to determine if the subtilisin E inactive pro-mature mutant could refold denatured subtilisin from other strains of *Bacillus*, denatured pro-mature mutant of subtilisin E was added to denatured mature subtilisin BPN' and subtilisin Carlsberg (2.4 moles pro-mature to 1 mole mature) and dialysed against renaturing buffer. Only 5.5 % and 2.3 % of wild type specific activity was recovered for BPN' and Carlsberg respectively even though there is a very high homology between the enzymes (29). Subtilisin E has the greatest amino acid homology with subtilisin BPN' (30) (88.5 % in the propeptide and 86.2 % in the mature region) which is probably directly related to the greater recovery of specific activity. The low specific activities highlight the specificity that each propeptide has with it's own mature region; with only 9 out of 77 residues changed, the propeptide from subtilisin E still unable to recover appreciable activity from denatured subtilisin BPN'.

This study demonstrated that inactive pro-mature mutant of subtilisin E could refold denatured mature protease. However, it failed to isolate the propeptide from the covalently attached mature domain, thus any effects the mature domain had on refolding could not be ruled out. Ohata *et al* chemically synthesized the complete propeptide for subtilisin E (31). It was found to be able to convert denatured subtilisin E to the native, active state using the same methodology as previous studies (29). The optimal molar ratio of propeptide : denatured mature protease for refolding was 1 : 1. At higher molar ratios of propeptide, it acts as a competitive inhibitor of the mature protease with a $K_i = 10^{-7}$. A synthetic propeptide that had the first 12 amino acids deleted from the amino terminus (64-mer) could not refold denatured subtilisin E or inhibit the native enzyme which highlights the

importance of this region for folding and verifies the *in vivo* results previously obtained (17). Two other propeptides were synthesised that comprised the amino terminal 34 amino acids (34-mer) and the carboxy terminal 43 amino acids (43-mer) of the subtilisin E propeptide. Both were found to be incapable of refolding denatured protease. However, the 43-mer and the 64-mer were found to act as competitive inhibitors and have K_i values of about 10^{-5}, although still much weaker than the full length synthetic propeptide. The 34-mer and 64-mer were mixed in a 1:1 molar ratio and it was found that the two peptides couldn't complement each other and were unable to refold denatured protease even though all of the residues of the entire propeptide were accounted for. This suggests that the amino and carboxy terminal ends of the propeptide must be covalently linked in order to direct the folding of the protease.

1.3.2.2 Isolation of a Subtilisin Folding Intermediate

In another experiment, the propeptide of subtilisin BPN' was generated *in vivo* as a carboxy terminal fusion to dihydrofolate reductase (DHFR) (32). DHFR activity was used to detect fusion protein during purification. Once the DHFR-propeptide fusion protein was purified, an engineered factor X site was used to release the propeptide from the DHFR domain. Much like the results seen with α-lytic protease, the BPN' propeptide was able to direct the folding of isolated folding intermediate to correctly folded mature protease. Wild type subtilisin BPN' that had been denatured in guanidinium-HCl and refolded in the presence of propeptide regained activity, indicating the enzyme is in it's native conformation. Other refolding studies of subtilisin BPN' were carried out with serine 221 to alanine (S221A) active site mutants

to reduce the amount of autocatalytic degradation of the folding protease (32). The folding intermediate state was isolated by dialysing the guanidinium-HCl denatured mature region against potassium phosphate buffer. The resulting folding intermediate has a relative hydrodynamic volume between that of native subtilisin and guanidinium-HCl denatured subtilisin. Spectroscopic analysis of the protein by CD and fluorescence spectroscopy revealed that the folding intermediate state has secondary structure similar to the native enzyme, little tertiary structure, and that the tryptophan residues in the intermediate are slightly more protected from solvent, much like the observations gained from the α-lytic folding intermediate (25). Again, all of these observations suggest that the folding intermediate exists in a molten globule state (26) (Figure 4). The S221A subtilisin BPN' mutant was refolded with various concentrations of propeptide and the amount of refolding measured by near UV CD. After incubation in the presence of propeptide for 8 days at 4 °C, about 50 % of the wild type signal at 278 nm was recovered (32). Therefore, the folding intermediate must be on the folding pathway or in equilibrium with a conformation on the folding pathway because it is able to retain a folding competent state for extended periods. A large fraction of the signal was probably lost to protein aggregation over the long incubation period (32). Refolded S221A subtilisin was shown to bind chymotrypsin inhibitor 2 ($K_i = 2.9 \times 10^{-12}$ against subtilisin BPN'), a further indication demonstrating that the protein was correctly folded (32).

The rate constant for the refolding of subtilisin BPN' was determined by observing the increase in tryptophan fluorescence of the S221A pro-mature mutant at 340 nm (33). Unlike the rate constant determined with α-lytic protease which was based on the intermolecular interaction of added propeptide (25), this rate constant was determined from the intramolecular

interaction of propeptide covalently linked to the mature region. It will more accurately reflect the actual rate enhancement for folding produced by the propeptide. In the absence of the pro sequence, based on the detection limit of the protease assay, the rate constant for folding of the S221A mutant lacking the propeptide was $<1.4 \times 10^{-8}$ s^{-1}. The S221A pro-mature mutant had a rate constant of 4.7×10^{-3} s^{-1} which corresponds to a reduction of the folding transition state energy barrier by more than 7.5 kcal/mol (33).

The folded S221A pro-mature mutant, has similar spectrophotometric characteristics to native subtilisin BPN' (33). The folded S221A pro-mature mutant had a far UV CD spectra with ellipticity minima at 220 nm and 208 nm suggesting well defined α-helical structure and the near UV CD has a strong negative signal at 278 nm which is typical of well developed tertiary structure. The fluorescence emission spectra is quenched by about 30 % in native subtilisin BPN' compared to the S221A pro-mature mutant. This suggests that the propeptide is partially shielding 1 of the 3 tryptophan residues found in the mature protease and increasing the fluorescence. There are no tryptophan residues present in the propeptide to affect the fluorescence emission. The propeptide alone has no defined structure as seen by near and far UV CD, however, the difference spectra between folded S221A pro-mature mutant and native subtilisin has a negative ellipticity between 230 and 210 nm which suggests that the propeptide adopts a defined secondary structure in the context of the mature region only (33). Propeptide independent of the mature region could be induced to fold in the presence of 50 % ethylene glycol (34) and almost match the difference spectra. Based on the transition of unfolded to folded propeptide in ethylene glycol, the equilibrium of free propeptide was calculated to favor the unfolded state, with a ΔG of folding equal to about 2 kcal/mol at 25 ^{o}C (34). However, even

though the equilibrium is unfavorable, a few percent of the total population will be correctly folded which will have enough binding energy to catalyse folding of intermediate to the native state. Wild type subtilisin will have the propeptide covalently attached, therefore the local concentration of propeptide will be much higher and collisions of correctly folded propeptide and the folding intermediate with the correct orientation will be more frequent (34).

Further verification that the folding intermediate lies on the folding pathway came from analysing the unfolding of isolated subtilisin S221A folding intermediate and the S221A pro-mature folded protease in the presence of increasing guanidinium-HCl (33). Fluorescence (excitation at 280 nm and emission at 340 nm) and ellipticity at 222 nm of the unfolding proteins was measured and both the S221A folding intermediate and pro-mature mutant followed the same unfolding pathway after an initial unfolding transition for the pro-mature mutant at guanidinium-HCl concentrations below 1 M. Below 1 M guanidinium-HCl, the pro-mature and folding intermediates diverge due the tertiary structure found only in the pro-mature mutant. Near UV-CD of the S221A pro-mature mutant has shown that it lacks tertiary structure above 1 M guanidinium-HCl, much like the intermediate state lacks tertiary structure (33).

Analysis of the propeptide using α-lytic protease and subtilisin as model systems have shown that the propeptide in the two classes of proteases has the same general function even though the enzymes are structurally unrelated. The propeptide of these enzyme systems functions to lower the energy barrier of the transition state of folding between the folding intermediate and the final native state, most likely by stabilizing the folding transition state (18, 22, 35). This folding process is under kinetic control since

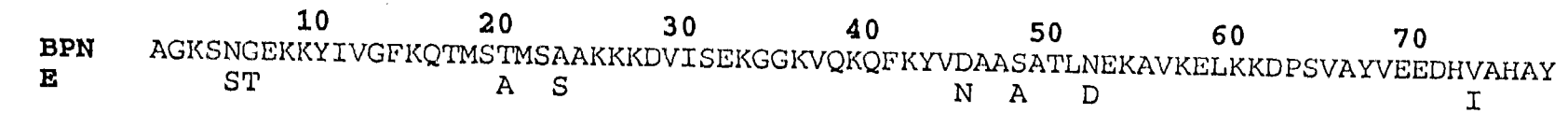
the intermediate to folding can be isolated *in vitro* and shown to be stable indefinitely if the propeptide is not present (25, 32). The folding intermediate is late in the folding pathway as observed by the large amount of secondary structure seen by CD spectroscopy (25, 32, 33) the intermediate hydrodynamic radius in gel filtration (25, 32), and the ability of the propeptide to inhibit native protease which indicates that the folding intermediate likely has native like structure (25, 31, 33).

1.3.2.3 Structural Analysis of Specific Residues in the Subtilisin Propeptide

Compared to α-lytic protease, the propeptides of subtilisin E and subtilisin BPN' have had a more thorough analysis for specific amino acids or sequences that directly interact with the mature protease during folding. Specific interactions between the propeptide and mature domain have been difficult to elucidate due to the lack of a pro-mature x-ray crystal structure until recently (34).

Figure 7. Amino Acid Sequence of the Subtilisin BPN' and Subtilisin E Propeptide. The amino acid sequence of subtilisin BPN' and subtilisin E are propeptide are shown (18). Only the divergent amino acids of subtilisin E are shown, all other residues are conserved. The propeptides have a 88.5 % amino acid identity.

23b


10
20
30
40
50
60
70
BPN
E
AGKSNGEKKYIVGFKQTMSTMSAAKKKDVISEKGGKVQKQFKYVDAASATLNEKAVKELKKDPSVAYVEEDHVAHAY
ST
A S
N A D
I

Sequence analysis of the propeptide of subtilisin E revealed that it contains 36 % charged residues (28 of 77) compared to only 12 % in the mature region (36) and these charges are unevenly distributed throughout the propeptide (Figure 7). For example, the amino terminal 27 amino acids of the propeptide contains 7 lysine and only one glutamic acid, while the carboxy terminus 16 amino acids has the reverse situation with 5 acidic residues, two histidines, and no lysine residues (36). Another noteworthy feature is that the propeptide contains three hydrophobic stretches of amino acids of 9, 5 and 4 residues (36). The substrate binding site of subtilisin is predominantly hydrophobic, and the charge distribution is uneven on the surface of subtilisin (37). Therefore, it is speculated that the hydrophobic residues interact with the substrate binding site which in turn causes the charged residues on the surface to even out the distribution of surface charge (37). The charged residues may also interact with the membrane *in vivo* (33).

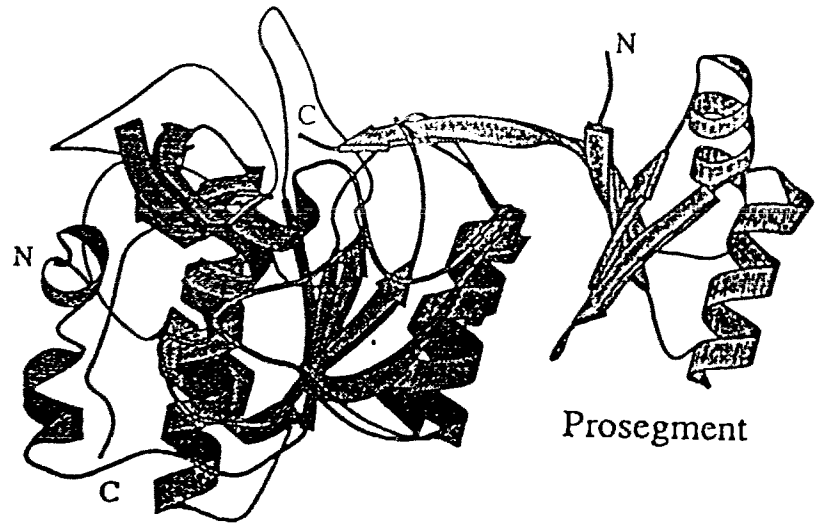
Error-prone *Taq* DNA polymerase was used to generate random mutations within the region of the subtilisin E gene encoding the propeptide (37, 38). Mutants were screened by looking for transformed colonies that lack subtilisin activity. Almost half of the mutations fell within three short stretches of hydrophobic residues which comprise only 18 of the 77 possible amino acids in the propeptide (37). Therefore it is suspected that these regions are important for the folding process. Six of the twenty five propeptide mutants that were unable to produce active subtilisin *in vivo* (37, 38) were analysed for their ability to refold denatured subtilisin and inhibit native subtilisin *in vitro* (30). Six of the mutants were found to generate active protease as seen by zones of clearing around *E. coli* colonies incubated for 6 days at room temperature, longer than the previous screening assay (37, 38). The six mutant pro sequences were subcloned and expressed in *E. coli* using a

T_7 expression system induced by IPTG and purified to greater than 99 % homogeneity. The mutant propeptides were found to refold denatured subtilisin in a range from 80 % wild type efficiency down to only 13 % efficiency (30). A direct relationship between the ability of a propeptide to correctly fold and inhibit subtilisin activity was established (30); the ability of a propeptide to refold denatured protease is directly related to how well it can bind (inhibit) the native protease. However, a direct correlation between the location of mutation and "foldase" activity could not be established because the loss of activity does not follow a predictable pattern.

One of the mutations, methionine to leucine at position -60 of the propeptide was used to find a supressor mutation in the mature region (37). Again, *Taq* DNA polymerase was used to generate random mutations in the mature region, and a serine to leucine 188 (S188L) mutation was found to have activity when expressed in *E. coli*. The S188L mutation is located on the surface of the active site binding cleft (37), and could interact with residue -60 in the propeptide and suppress it's effects.

A co-crystal structure of subtilisin BPN' and propeptide has been obtained (34). It was generated from refolding denatured subtilisin BPN' with purified propeptide *in vitro*, and the structure of subtilisin was found to be identical to the crystal structure of subtilisin folded *in vivo*. As expected, the C-terminus of the propeptide binds onto the active site much like a substrate, with a tyrosine residue of the propeptide occupying the S1 subsite. As a whole, the propeptide appears to stabilise the αβα substructure of subtilisin BPN' to allow folding to the native state and prevent the generation of unproductive misfolded or unfolded conformations (Figure 8).

Figure 8. Co-crystal Structure of Refolded Subtilisin and Propeptide. A ribbon drawing showing the backbone of mature subtilisin E and the propeptide. Mature subtilisin BPN' was refolded in the presence of exogenously added propeptide *in vitro*. The C-terminus of the propeptide is in the active site of subtilisin. This figure was adapted from Bryan *et al* (34).


N
C
N
C
Prosegment

Although the subtilisin propeptide has no homology to the propeptide of α-lytic protease and other pro-proteins (22), the subtilisin results will aid in understanding the more general principles of propeptide-mediated protein folding.

1.4 *Streptomyces griseus* PROTEASES AS A MODEL SYSTEM TO ELUCIDATE PROPEPTIDE MEDIATED PROTEIN FOLDING

1.4.1 *S. griseus* Proteases

Our research involves the analysis of propeptide mediated folding of the *Streptomyces griseus* serine proteases. *S. griseus* is a gram positive soil bacterium that secretes a number of serine proteases (39). These are enzymes that catalyse the hydrolysis of amides and esters with the serine, histidine, and aspartate catalytic triad mechanism (40). The *S. griseus* proteases under study include *Streptomyces griseus* protease A (SGPA), SGPB (10), SGPC (12), SGPD (11), and SGPE (13) which are encoded by *sprA*, *sprB*, *sprC*, *sprD* and *sprE* respectively (Table 1). Four of the five proteases are small monomeric proteins. SGPD is a homodimer with an amino terminal 24 residues of the signal sequence that has high homology to the mitochondrial import signal sequence of hsp60 (41, 42) All of the proteases except SGPE have chymotrypsin-like specificity which means that they will catalyse the hydrolysis of peptide bonds next to large, bulky aromatic or aliphatic amino acids (40). SGPE, unlike the other proteases, has specificity for glutamic acid. SGPC has an additional 50 amino acid domain fused to the C-terminus of the mature domain via a 20 amino acid linker region. This 50 amino acid

Table 1. Protease Data. Listed are data for the family of *S. griseus* proteases under investigation, also included are data for the proteases that have been the primary focus of propeptide research to date.

Protease	Source Organism	Gene	Amino Acids			Mature
			pre	pro	mature	Protein (kDa)
SGPA	*S. griseus*	*sprA*	38	78	181	18
SGPB	*S. griseus*	*sprB*	38	76	185	18
SGPC	*S. griseus*	*sprC*	40	162	190 + 20 + 50 [1]	26
SGPD	*S. griseus*	*sprD*	64	140	188	18/36 [2]
SGPE	*S. griseus*	*sprE*	29	139	187	18
α-lytic protease	*L. enzymogenes*	*ALPHA-LP*	33	166	198	20
subtilisin BPN'	*B. amyloliquefaciens*	*apr*	29	77	275	28
subtilisin E	*B. subtilis*	*aprE*	29	77	275	28

[1] SGPC contains an additional 50 amino acid carboxy terminus domain that has homology to the chitin binding domain of chitinases A1 and D of *B. circulans* joined to the mature region by a 20 amino acid linker region (12).

[2] *sprD* encodes a 18 kDa mature protein that exists as a stable 36 kD homodimer (11).

domain has 57 % amino acid homology to the chitin binding domain of chitinase A1 from *B. circulans* (43), however, the domain in SGPC has not been found to bind chitin (unpublished results). Like α-lytic protease and the subtilisin proteases, the *S. griseus* proteases are secreted as pre-pro-mature polypeptide precursors, though only the mature region is detected in the culture supernatant. Unlike α-lytic protease and the subtilisin proteases, all five of the proteases come from the same organism. The *S. griseus* proteases have a high degree of homology, yet still have structural and functional differences such as the various propeptide lengths, the additional chitin binding domain of SGPC, and the glutamic acid substrate specificity of SGPE give a framework for structure and function analysis of the proteases. The *S. griseus* proteases and α-lytic protease are members of the chymotrypsin superfamily of enzymes and are structurally unrelated to the subtilisin family of serine proteases, although the same catalytic mechanism was developed through convergent evolution (40).

The proteases can be divided into 3 groups based on the length of the propeptide: SGPA and SGPB with 78 and 76 residue propeptides respectively; SGPD and SGPE with 140 and 139 residue propeptides respectively; and SGPC with a 162 residue propeptide. The propeptide of α-lytic protease is 166 residues in length, close to that of SGPC.

1.4.2 Expression in *Bacillus subtilis*

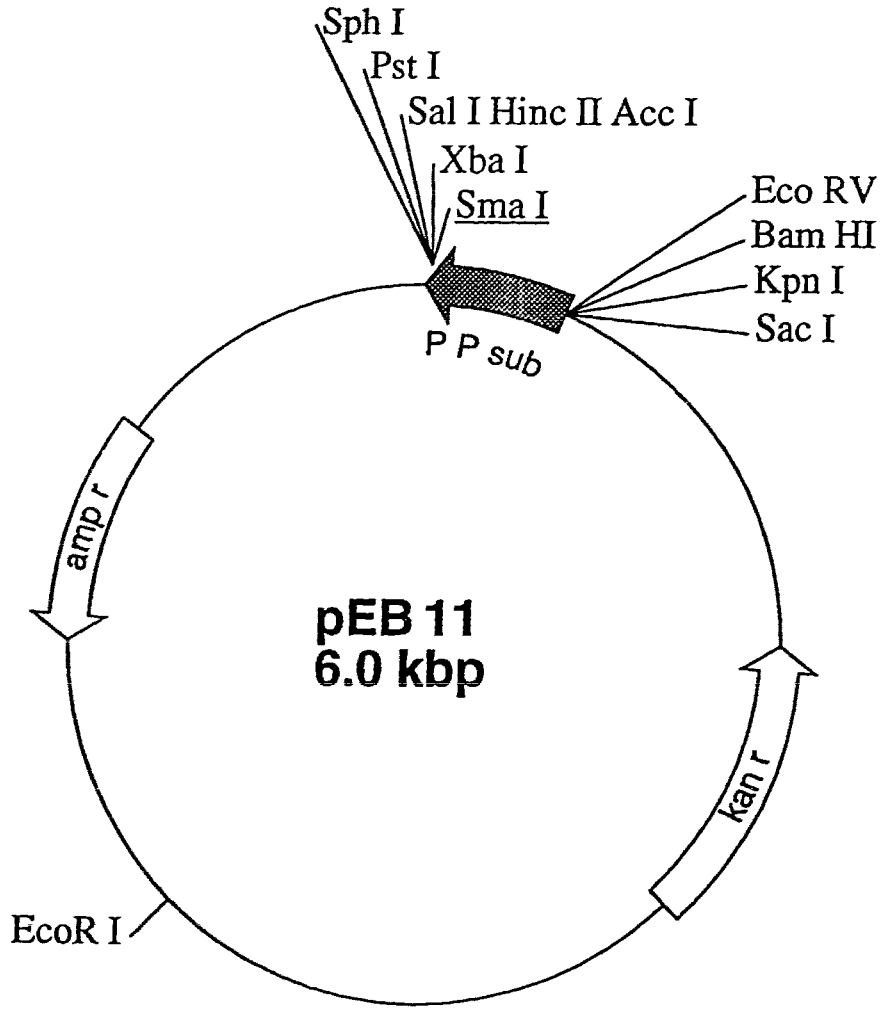
All of the *S. griseus* proteases have been cloned and expressed in *B. subtilis* for purification and analysis. *B. subtilis* like *S. griseus* is a gram positive bacterium that naturally secretes proteins (including subtilisin) across a single cell membrane into the extracellular environment (44). The

commonly used gram negative *E. co'i* bacterium secretes protein into an inter-membrane space called the periplasm. This region is located between the inner and outer membranes, and proteins are only slowly transported across the outer cell membrane (45) which makes purification of the secreted protein problematic. The advantages of using *B. subtilis* over *S. griseus* are that it is a more well characterised organism, easier to manipulate, plasmids have been produced specifically for *B. subtilis* expression, and *B. subtilis* will secrete proteins directly into the cellular media where they can be easily purified.

The pEB (*E. coli - B. subtilis*) expression vector was used for the expression of the *S. griseus* proteases (13) (Figure 9). It is a shuttle vector that can be transformed into *E. coli* for genetic manipulation and into *B. subtilis* for expression. The plasmid contains resistance genes for selection with either ampicillin and kanamycin for selection, but more importantly, it contains the ribosome binding site, promoter region and signal sequence from subtilisin BPN'. A *Sma*I restriction site has been engineered into the 3' end of the signal sequence to allow the generation of direct in-frame fusions with the propeptide of the cloned protease gene. The subtilisin BPN' signal sequence allows for efficient processing of the signal sequence and secretion of the foreign pro-mature protease in *B. subtilis*.

The *B. subtilis* strain DB104 used to express the *S. griseus* proteases is a protease deficient strain, lacking it's endogenous secreted proteases subtilisin and neutral protease (46). Therefore, any protease activity detected *in vivo* will be produced from a cloned pro-mature protease in the pEB expression vector.

Figure 9. Plasmid Map of the pEB-11 *B. subtilis* - *E. coli* Shuttle Vector. Genes are represented by the boxes and the arrowheads indicate the direction of transcription. The vector contains the ampicillin resistance gene (amp r) from pUC-18, the kanamycin resistance gene (kan r) from pUB110, and a 330 bp fragment of the subtilisin BPN' gene containing the promoter and pre region. Genes for expression in *B. subtilis* are blunt-cloned into the *Sma*I site to create in-frame fusion proteins with the pre region.


Sph I
Pst I
Sal I Hinc II Acc I
Xba I
Sma I
Eco RV
Bam HI
Kpn I
Sac I
P P sub
amp r
pEB 11
6.0 kbp
kan r
EcoR I

1.4.3 *S. griseus* Protease Gene Mutants

In order to determine the function of the propeptides, mutagenesis of the genes encoding the proteases must be done to generate structural alterations of the propeptides. Since no analysis of the *S. griseus* proteases' propeptides has been done to date, the initial focus of their structural analysis will involve SGPA and SGPB. These two proteases have the shortest propeptides, and have the highest amino acid identity (43%) (10, 11), therefore isolation of important residues for folding will be easier to determine than with longer or more divergent propeptides. A chimeric gene containing the region encoding the propeptide of *sprA* and the region encoding mature *sprB* will be expressed and used to determine if the propeptide of SGPA can generate mature, active SGPB. Previously, *in vitro* experiments with the *trans* addition of propeptide from a protease from another organism failed to appreciably fold denatured protease (29). In this experiment, both the propeptide and mature region will be expressed as a single polypeptide *in vivo*. In addition, while both the propeptide and mature regions of the chimera will be from two separate genes, both are from the same organism unlike previous experiments. These two factors may contribute to the successful folding of the mature protease beyond what has been previously observed. Also, a series of genes encoding deletions to the amino terminus of the propeptides of SGPB, SGPC, SGPD and SGPE will be expressed in *B. subtilis* and also analysed for their ability to correctly fold the mature protease *in vivo*. How well these mutants propeptides can generate active protease will give some insights about the nature of the folding process with the *S. griseus* proteases. For example, if the deletions have little or no

effect on folding, than the deleted amino acids do not play a major role in the folding process.

In vivo analysis is a good starting point for the study of the perviously uncharacterised *S. griseus* propeptides. Compared to *in vitro* analysis, a large number of mutant propeptides can be easily analysed for their importance to the folding process by the detection of active protease directly from the *B. subtilis* culture supernatants. While the propeptide is known to direct the folding of the mature protease based on *in vitro* results, care must be taken with the interpretation of *in vivo* results. In addition to propeptide mediated folding of the mature protease, there are many factors such as protein translation, secretion across the cell membrane and cell growth that may also be affected by the propeptide mutants and alter the amount of active protease in the cell supernatant. Evidence suggests that the function of the propeptide is independent of these other processes in bacterial proteases and any changes in activity of the propeptide mutants can be directly attributed to the folding process (17, 28). However, *in vitro* analysis of purified propeptide mutants will eventually have to be carried out for verification of the predicted propeptide function based on *in vivo* results.

Protease activity will be used as an indication of how well the mutant propeptide can fold the mature protease *in vivo*; protease activity will be directly correlated to the ability of the propeptide to correctly fold mature protease. A reduced rate of activity of a mutant propeptide compared to wild type protease will indicate the folding transition state energy barrier of the mutant is increased compared to wild type. Fewer folding intermediates will be able to traverse the increased energy barrier to native protease before they are degraded *in vivo*. Mutant propeptides will not generate protease molecules that have a reduced ability to hydrolyse substrate based on the

observations with other pre-pro-mature bacterial proteases; either correctly folded protease is generated by the reduction of the folding transition state barrier by the propeptide, or the pro-mature protease will be unable to overcome the transition state energy barrier and remain in the folding intermediate state. Propeptide mutants with reduced activity *in vivo* will highlight important residues required for the folding process. The ability of both the chimera and deletion mutants to correctly fold mature protease may help in gaining a better understanding of how the propeptides function *in vivo*.

2. MATERIALS AND METHODS

2.1 MATERIALS

Calf intestinal alkaline phosphatase (CIP) was purchased from Boehringer Mannheim. Vent DNA polymerase (New England Biolabs) was used for all polymerase chain reactions (PCR). All other enzymes for digesting or modifying DNA except ribonuclease A were purchased from New England Biolabs or Life Technologies, Inc. Enzymes were used in accordance with the recommendations of the supplier. [α-^{32}P]dATP (~3,000 Ci/mmol) was from Amersham Corp. The Random Priming System I kit was purchased from New England Biolabs. Oligonucleotides were purchased from the Institute of Molecular Biology and Biochemistry Services Laboratory at Simon Fraser University, Burnaby, BC. Amino terminal sequencing was provided by the Protein Microchemistry Centre, Department of Biochemistry and Molecular Biology, University of Victoria, Victoria, BC. Zeta-Probe membrane was purchased from Bio-Rad and x-ray film from Eastman Kodak Co. Immobilon-PVDF membrane was purchased from Millipore. The Centricon-10 tubes were purchased from Amicon. The QIAEX Gel Extraction Kit and QIAGEN Plasmid Maxi Kit was from Qiagen. The Protein-Pak SP FPLC column was purchased from Waters. Acrylamide, agarose, ampicillin, boric acid, glycine, kanamycin, lysozyme, MOPS, polyethylene glycol 8000, ribonuclease A, *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide, and sodium dodecyl sulphate were from Sigma. Agar, glacial acetic acid, glucose potassium acetate, sodium acetate, sucrose, and yeast extract were purchased from BDH. Trypticase peptone was purchased from Becton Dickinson. Penassay and casamino acids were purchased from Difco. *N*-succinyl-Ala-

Ala-Pro-Glu-*p*-nitroanilide was purchased from Bachem Biosciences. All chemicals and reagents were of the highest grade commercially available.

2.2 BUFFERS AND OTHER SOLUTIONS

1. *E. coli* plasmid Preparation Solutions

Solution I: 25 mM Tris-HCl (pH 8.0), 10 mM EDTA, and 50 mM glucose. Autoclave, then add 4 mg/mL lysozyme just prior to use.

Solution II: 0.2 N NaOH and 1 % (w/v) SDS. Prepare fresh just prior to use.

Solution III: 3 M potassium acetate and 2 M acetic acid.

TE8: 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA (pH 8.0).

RNase A1: Dissolve 10 mg/mL RNase I in 10 M Tris-HCl (pH 7.5) containing 15 mM NaCl. Add 1 mL aliquots into 1.5 mL eppendorf tubes and place in a boiling water bath for 10 minutes. Allow the tubes to cool slowly to room temperature and store at -20 ^{o}C.

2. *E. coli* Competent Cell Preparation Buffers

TfBI: 30 mM CH_3COOK, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$ and 15 % (v/v) glycerol. Adjust to pH 7.0 and filter sterilize.

TfBII: 10 mM Na-MOPS (pH 7.5), 75 mM $CaCl_2$, 10 mM KCl and 15 % (v/v) glycerol. Adjust to pH 7.0 and filter sterilize.

3. *B. subtilis* Plasmid Preparation Solutions

The solutions for *B. subtilis* plasmid preparations are the same as those for *E. coli* plasmid preparations except for the following:

Solution IIb: 0.2 N NaCl and 1 % (w/v) SDS. Prepare fresh just prior to use.

4. *B. subtilis* Transformation and Protoplast Solutions

2 × SMM Buffer: Dissolve 17.115 g sucrose, 232.1 mg maleic acid, 406.6 mg $MgCl_2$ in ddH_2O, adjust to pH 6.5 with NaOH. Dilute to 50 mL total volume with ddH_2O and autoclave.

4 × Penassay: Dissolve 3.5 g Penassay (DIFCO bacto antibiotic medium 3) in 50 mL ddH_2O and autoclave.

40 % PEG: Dissolve 8.0 g PEG8000 in 20 mL 1 × SMM and autoclave.

2 × SMMP: Mix together equal volumes of 2 × SMM and 4 × Penassay.

Lysozyme Stock Solution: Dissolve 100 mg/mL lysozyme in ddH_2O and filter sterilize.

5. Agarose Gel Solutions

50 × TAE Buffer Solution: 242 g Tris base, 57.1 mL glacial acetic acid, 100 mL 0.5 M EDTA (pH 8.0). Dilute to 1 L with ddH_20.

1000 × Ethidium Bromide Stock Solution: 50 mg ethidium bromide and 100 mL ddH_2O.

36 × Agarose Gel Loading Dye: 10 mL 30 % glycerol, 0.025 g bromophenol blue, and 0.025 g xylene cyanol FF. Dilute × 6 with 30 % glycerol before use.

6. SDS-PAGE Gel Solutions

Separating Buffer: 0.75 M Tris-HCl (pH 8.8) and 0.2 % (w/v) SDS.

Stacking Buffer: 0.146 M Tris-HCl (pH 6.8) and 0.117 % (w/v) SDS.

SDS-PAGE Acrylamide Stock Solution: 30 % (w/v) acrylamide and 2 % (w/v) N,N - bis-acrylamide, deionized with Amberlite MB-1.

5 × SDS-Page Loading Buffer: 7.5 mL 2.083 M Tris-HCl (pH 6.8), 11.0 mL β-mercaptoethanol, 5 g SDS was diluted to 25 mL total volume with ddH_20.

Then glycerol was added to 50 mL total volume, followed by 2.5 mg of bromophenol blue.

7. SDS-Page Staining and Destaining Solutions

Fast Destain Solution: 50 % (v/v) methanol and 10 % (v/v) glacial acetic in ddH_2O.

Slow Destain Solution: 10 % (v/v) glacial acetic acid in ddH_2O.

Coomassie Stain: 0.3 % (w/v) in fast destain solution.

8. Sequencing Gel Solutions

10 × TBE: Dissolve 10.8 g-base, 5.5 g boric acid in 80 mL ddH_2O. Add 4.0 mL of 0.5 M EDTA and dilute to 100 mL total volume.

Acrylamide Stock Solution: 38 % (w/v) acrylamide and 2 % (w/v) N,N-bis-acrylamide, deionized with Amberlite MB-1.

9. Northern Blot Solutions

10 × MOPS Buffer: 200 mM MOPS acid, 50 mM sodium acetate, 10 mM EDTA, adjust to pH 7.0 with NaOH.

3 × Loading Buffer: 60 µL 10 × MOPS, 210 µL formaldehyde, 600 µL formamide, 200 µL glycerol, 5 mg bromophenol blue. Deionize the formaldehyde and formamide with Amberlite MB-1 before use.

Hybridization Solution: 10 mL deionized formamide, 5 mL 20 % SDS, 3.42 mL 2 M Na_2HPO_4, 1.58 mL 2 M NaH_2PO_4, 40 µL 0.5 M EDTA, and 20 mg BSA.

20 × SSC Buffer: Dissolve 173.3 g NaCl and 88.2 g trisodium citrate·$2H_2O$ in 800 mL ddH_2O and adjust to pH 7.0 with NaOH. Dilute to 1 L total volume with ddH_2O and autoclave.

10. Amino Terminal Sequencing Solutions

Transfer Buffer: 39 mM glycine, 48 mM Tris-base, 0.0375% SDS.

PVDF Stain: 0.025 % (w/v) in 40 % methanol

11. Protein Purification and Column Chromatography Buffers

Buffer A: 50 mM sodium acetate (pH 4.65).

Buffer B: 50 mM sodium acetate (pH 4.65) and 1 M NaCl.

12. Substrate Solutions

Phe Substrate Solution: 0.1 mg/mL *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide in 50 mM Tris pH 8.0.

Glu Substrate Solution: 0.1 mg/mL *N*-succinyl-Ala-Ala-Pro-Glu-*p*-nitroanilide in 50 mM / 5 % methanol (pH 8.0).

2.3 BACTERIOLOGICAL MEDIA, STRAINS AND METHODS

2.3.1 Media

TY/Mg Broth: 20.0 g tryptone, 5.0 g yeast extract, 5 g NaCl, 10 mM $MgSO_4$ per 1 L ddH_2O. The $MgSO_4$ is added from a 1 M filter sterilized stock.

2 × TY Broth: 16 g tryptone, 10 g yeast extract, and 5 g NaCl per 1 L ddH_2O. Autoclave at 121 ^oC to sterilize.

2 × TY/Amp Broth: 2 × TY broth supplemented with 200 µg/mL filter-sterilized ampicillin after autoclaving.

2 × TY/Kan Broth: 2 × TY broth supplemented with 50 µg/mL filter-sterilized kanamycin after autoclaving.

2 × TY Plates: 2 × TY broth containing 1.5 % (w/v) agar. Autoclave at 121 ^oC to sterilize.

2× TY/Amp Plates: 2 × TY/Amp plates supplemented with 200 µg/mL filter-sterilized ampicillin after autoclaving.

TY/Milk Plates: 20 g skim milk powder was dissolved in 200 mL ddH_2O; and 10 g yeast extract, 16 g tryptone, 5 g NaCl and 15 g agar were dissolved in 800 mL ddH_2O. After autoclaving at 121 ^oC, both solutions are mixed for 1 L total volume.

TY/Milk/Kan Plates: TY/Milk plates supplemented with 50 μg/mL filter-sterilized kanamycin after autoclaving.

TY/Milk/Amp Plates: TY/Milk plates supplemented with 200 μg/mL filter-sterilized ampicillin after autoclaving.

S-Broth: 35 g tryptone, 2 g yeast extract, 5 g NaCl per 1 L ddH_2O. After autoclaving at 121 ^oC to sterilize, add 50 μg/mL filter sterilized kanamycin and 10 mM filter sterilized $CaCl_2$.

1 × Penassay Broth: 17.5 g Penassay dissolved per 1 L ddH_2O. Autoclave at 121 ^oC to sterilize.

Protoplast regeneration plates: Autoclave each of the following separately at 121 °C:

Sodium succinate - 33.77 g in 125 mL ddH_2O adjusted to pH 7.3 with 1 N HCl.

Agar nutrient- 4 g agar, 1.25 g casamino acids and 1.25 g yeast extract in 125 mL ddH_2O.

Phosphates - 1.15 g K_2HPO_4 and 0.375 g KH_2PO_4 in 25 mL ddH_2O.

20 % glucose - 1.25 g glucose in 6.25 mL ddH_2O.

1 M $MgCl_2$ - 1.02 g $MgCl_2 \cdot 6H_2O$ in 5 mL ddH_2O.

Mix all of the components after they have cooled to about 60 °C, then add 1.25 mL of 2 % filter sterilized BSA and 1 mL of 50 mg/mL filter sterilized kanamycin. These plates will last for about 2 weeks at 4 °C.

2.3.2 Bacterial Strains

2.3.2.1 *Escherichia coli* strain DH5α:

Genotype: *sup*E44 Δ*lac*U169 (ϕ80 *lac*ZΔM15) *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1.

Growth conditions and Relevant Characteristics: *E. coli* DH5α (47) was grown and maintained on 2 × TY plates and broth at 37 °C. This strain was used for routine transformations of plasmids and ligation mixtures.

2.3.2.2 *Escherichia coli* strain TB1:

Genotype: *ara* Δ (*lacproAB*) *rpsL* (ϕ80 *lacZ*ΔM15) *hsdR* .

Growth Conditions and Relevant Characteristics: *E. coli* TB1 (48) was grown and maintained on 2 × TY plates and broth at 37 °C. This strain was used to transform the construct pMP-BPΔ10 and for routine transformations, interchangeably, with DH5α.

2.3.2.3 *Bacillus subtilis* strain DB104:

Genotype: *his nprR2 nprE18* Δ*aprA3* (pUBHSΔ)3.

Growth Conditions and Relevant Characteristics: *B. subtilis* DB104 (46) was grown and maintained on TY/Milk plates and 2 × TY broth at 30 °C. This strain was used for transforming pEB vector constructs for expression studies.

2.3.3 Preparation of Competent Cells

2.3.3.1 Preparation of Competent *E. coli* Cells

A single *E. coli* colony from a plate grown overnight at 37 °C was used to inoculate a culture tube containing 2 mL of TY/Mg broth. The culture was allowed to grow overnight at 37 °C with shaking and then used to inoculate 20 mL of TY/Mg prewarmed to 37 °C. The 20 mL culture was grown with vigorous shaking at 37 °C until OD_{600} = 0.5 was reached. The cells were then transferred to 200 mL of pre-warmed TY/Mg broth in a 1 L flask and grown with shaking at 37 °C until the cell density reached OD_{600} = 0.6. The culture was then chilled on ice, transferred to cooled, sterile Oakridge tubes, and

centrifuged at 4000 RPM in a Sorvall SS-34 rotor for 5 minutes at 4 ^{o}C. The following steps were all performed at 4 ^{o}C. The supernatant was decanted and the cell pellets were resuspended in 40 mL TfBI buffer and centrifuged again as before. The TfBI buffer was decanted and the cell pellet was resuspended in 10 mL TfBII buffer. Aliquots of cells were placed in 0.7 mL microcentrifuge tubes, frozen on dry ice, and stored at -80 ^{o}C.

2.3.3.2 Preparation of Competent *B. subtilis* DB104 Protoplasts

The production of *B. subtilis* protoplasts is based on the method of Lovett and Ambulos (49). A single colony of *B. subtilis* was selected off a TY plate grown overnight at 37 ^{o}C and used to inoculate 2 mL of 1 × Penassay broth. The culture was grown for 8 hours with 200 RPM shaking at 37 ^{o}C and used to inoculate 1 L of 1 × Penassay in a 2 L Erlenmeyer flask which was grown under the same conditions until OD_{600} = 0.6 (mid-log phase). The cells were transferred to four sterile GSA tubes, iced for 5 minutes, and then centrifuged at 5000 RPM in a GSA rotor for 10 minutes at 4 ^{o}C. The supernatant was decanted and the cells were resuspended in 100 mL SMMP (25 mL per tube). To the resuspended cells, 2.0 mL of 100 mg/mL lysozyme solution was added (0.5 mL per tube) and the tubes were incubated at 30 ^{o}C for 2 hours with gentle shaking just sufficient to move the cells. The cells were then transferred to sterile Oakridge tubes and centrifuged at 5000 RPM in a SS-34 rotor for 15 minutes at 4 ^{o}C. The supernatant was discarded and the cell pellets were gently resuspended by swirling in 100 mL 2 × SMMP and centrifuged as in the previous step. The supernatant was then discarded and the cell pellets resuspended in 80 mL 2 × SMMP. The cell suspension was

divided into 500 µL aliquots in 1.5 mL eppendorf tubes and placed directly into the -80 °C freezer for storage.

2.3.4 Transformation of Cells

2.3.4.1 Transformation of *E. coli* Cells

For each transformation, 1 µL of plasmid or a ligation mixture was diluted to 50 µL total volume with ddH_2O. To this mixture, 100 µL of competent cells thawed on ice was added and the resulting mixture incubated on ice for 30 minutes. The cells were then heat shocked at 37 °C for 1 minute, incubated on ice for 1 minute, and then transferred to 2 mL 2 × TY broth. The 2 mL culture was incubated at 37 oC with vigorous shaking for 2 hours. After incubation, 1 mL, 100 µL, and 10 µL aliquots were spread plated on 2 × TY plates and incubated overnight at 37 °C.

2.3.4.2 Transformation of *B. subtilis* DB104 Protoplasts

This procedure is based on the method of Chang and Cohen (50). In a 1.5 mL Eppendorf tube, 10 µL of plasmid DNA isolated from a standard alkaline lysis plasmid preparation (section 3.4.2.1) was mixed with 10 mL 2 × SMM. The protoplast suspension was quick thawed in a 37 °C water bath, and 100 µL was added to the plasmid solution followed quickly by 300 µL 40 % PEG. The solution was gently mixed with a P1000 pipet and after exactly 2 minutes, 1 mL of 2 × SMMP was added. The resulting solution was centrifuged for 10 minutes at 6000 RPM in a benchtop microfuge to pellet the protoplasts. After centrifugation, the supernatant was decanted, tubes spun

again for 1 minute, and most of the remaining supernatant was removed with a P200 pipet except for a trace amount at the bottom of the tube to ensure the protoplast pellet remains intact. The pellet was resuspended in 1 ml 2 × SMMP and incubated with very gentle shaking (approximately 30 RPM) for 1.5 hours at 30 °C. After the incubation, 1 mL of the protoplast solution was spread on a dry regeneration plate and then allowed to dry in a laminar flow hood. The plates were incubated at 37 °C, and colonies were streaked onto TY/Milk/Kan plates as they appeared on the regeneration plates in order to screen out contaminants.

2.4 PLASMIDS

2.4.1 Cloning Plasmids Used

2.4.1.1 pMAL-p

This plasmid pMAL-p (51) is an *E. coli* expression vector that is under control of the *tac* promoter and generates a fusion protein than contains a maltose binding protein and signal sequence on the amino terminus of the expressed protein (Figure 10). The signal sequence allows the expressed protein to be secreted into the periplasm of *E. coli*, after which the expressed protein is non-specifically transported through the outer cell membrane.

2.4.1.2 pDS-A3

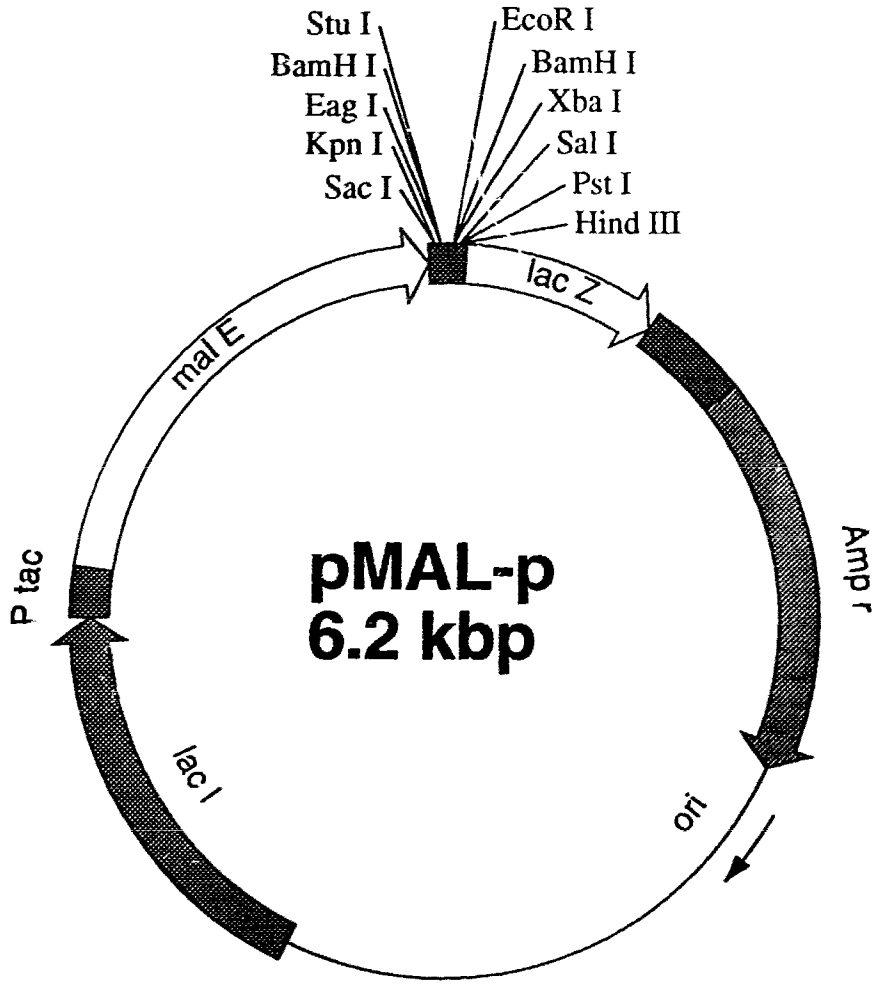
The plasmid pDS-A3 was constructed by Sidhu (52) and contains the region of *sprA* encoding for the entire propeptide and mature region in pUC-

18. A *Pst*I site was engineered into the 5' end of the propeptide encoding portion of *sprA* . Digestion of the vector with *Pst*I followed by removal of the 3' overhang with T_4 polymerase produces a blunt end where the first codon at the 5' end encodes the first amino acid of the propeptide. This allows propeptide and mature region of *sprA* to be cloned in-frame into pEB-11 for expression.

2.4.1.3 pEB-11

This plasmid was constructed by Sidhu (13) and functions as a *E. coli* - *B. subtilis* shuttle vector (Figure 9). The vector contains components of pUC-18 for maintenance in *E. coli*, and components of pUB110 for maintenance in *B. subtilis*. In addition, it contains the promoter region and the pre region of the subtilisin *BPN* gene which is the signal sequence for secretion out of *B. subtilis*. An engineered *Sma*I site downstream of the pre region allows for the cloning of genes in-frame for expression of fusion proteins with the signal sequence on the amino terminal end of the protein.

Figure 10. Plasmid Map of the *E. coli* Expression Vector pMAL-p. Genes are represented by the boxes and the arrowheads indicate the direction of transcription. The cloned gene is inserted downstream of the *malE* gene which encodes the maltose binding protein and pre domain. The tac promoter (P tac) is induced by IPTG. The *lacI* encodes the Lac repressor which prevents transcription from the tac promoter in the absence of IPTG. The plasmid also contains the ampicillin resistance gene (amp r) for selection. The *Stu*I site was engineered for blunt in-frame cloning of insert with the *malE* signal sequence (51).


Stu I
BamH I
Eag I
Kpn I
Sac I
EcoR I
BamH I
Xba I
Sal I
Pst I
Hind III
lac Z
mal E
P tac
pMAL-p
6.2 kbp
Amp r
ori
lac I

2.4.1.4 pEB-13

The vector pEB-13 is identical to pEB-11 except that is has a single cytidine nucleotide inserted 3' of the pre region to introduce a +1 reading frame shift. This is for the expression of cloned inserts that require the addition of a nucleotide on the 5' end of the insert for in-frame transcription with the subtilisin BPN' signal sequence.

2.4.2 Expression Plasmids Used

2.4.2.1 pEB-B8

This plasmid was constructed by Sidhu (12) and consists of region of *sprB* encoding the entire propeptide and mature region cloned into pEB-11 for expression.

2.4.2.2 pEB-C8

This plasmid was constructed by Sidhu (12) and consists of the region of *sprC* encoding entire propeptide and mature region cloned into pEB-11 for expression.

2.4.2.3 pEB-D8

This plasmid was constructed by Sidhu (11) and consists of the region of *sprD* encoding the entire propeptide and mature region cloned into pEB-11 for expression.

2.4.2.4 pEB-E

This plasmid was constructed by Sidhu (13) and consists of the entire propeptide and mature region of *sprE* cloned into pEB-11.

2.4.2.5 pEB-BPΔ10

This plasmid was constructed by Sidhu (52) and consists of the region of *sprB* encoding from the eleventh amino acid of the propeptide to the carboxy terminus of SGPB cloned into pEB-11 for expression.

2.4.2.6 pEB-BPΔ15

This plasmid was constructed by Sidhu (52) and consists of the region of *sprB* encoding from the sixteenth amino acid of the propeptide to the carboxy terminus of SGPB cloned into pEB-11 for expression.

2.4.2.7 pEB-BPΔ20

This plasmid was constructed by Sidhu (52) and consists of the region of *sprB* encoding from the twenty-first amino acid of the propeptide to the carboxy terminus of SGPB cloned into pEB-11 for expression.

2.4.3 Preparation of Plasmids

2.4.3.1 Small Scale *E. coli* Plasmid Preparation

A single bacterial colony was used to inoculate 2 mL of 2 X TY/Amp media and grown for about 8 hours or overnight at 37 ^{o}C with shaking. The culture was then placed in a 1.5 mL eppendorf tube and centrifuged for 2 minutes at 15000 RPM in a microfuge tube. The supernatant was discarded and the pellet resuspended in 100 µL of Solution I and incubated at room temperature for 5 minutes, followed by the addition of 200 µL of Solution II. The tube was inverted to mix, and incubated at room temperature until the solution clears (about 5 minutes). At this point, 150 µL of solution III was added, the tube inverted to mix, and incubated for 5 minutes at room temperature. The solution was centrifuged for 10 minutes at 15000 RPM in a microfuge and the supernatant decanted into a fresh 1.5 mL eppendorf tube. To the supernatant, 150 µL each of phenol and chloroform was added, the mixture vortexed, and then centrifuged as before for 5 minutes. The aqueous phase was carefully removed with a P1000 pipet set on 450 µL and added to a 1.5 mL eppendorf tube containing 900 µL 95 % EtOH to precipitate the DNA. The DNA was pelleted by centrifuging the solution for 10 minutes at 15000 RPM in a microfuge, and the supernatant was removed with an aspirated

drawn out pipet. The DNA pellet was dried under vacuum, and resuspended in 100 µL of TE8 containing 0.5 µL of RNase A_1 10 mg/mL stock and 0.5 µL of 100 units/µL RNase T_1 from Boehringer Mannheim. The plasmid solution was incubated at 37 °C for 1 hour to digest the contaminating RNA. This plasmid DNA was used for restriction digests and transformation procedures, further purification of the plasmid was necessary for sequencing (section 2.4.3.4).

2.4.3.2 Large Scale *E. coli* Plasmid Preparation

A 2 mL *E. coli* 2 × TY/Amp culture containing a plasmid of interest was grown overnight at 37 °C with shaking and used to inoculate 200 mL 2 × TY/Amp which was grown with the same conditions until OD_{600} = 1.5. The plasmid was purified using the standard protocol from the QIAGEN Plasmid Maxi Kit. The final plasmid pellet was resuspended in 200 µL of TE8. Up to 500 µg of plasmid can be isolated with the QIAGEN-tip 500. No RNasing is required with this method, and the plasmid is sufficiently pure for restriction digests and transformations.

2.4.3.3 *B. subtilis* Plasmid Preparation

This is based on the method of Lovett *et al* (49). A single bacterial colony was used to inoculate 2 mL of 2 × TY/Kan media and grown overnight at 37 °C with shaking. The culture was then placed in a 1.5 mL eppendorf tube and centrifuged for 2 minutes at 15000 RPM in a microfuge tube. The supernatant was discarded and the pellet resuspended in 100 µL of Solution I and incubated at room temperature for 5 minutes, followed by the

addition of 200 µL of Solution IIb. The tube was inverted to mix, and incubated at room temperature until the solution clears (about 5 minutes). At this point, 150 µL of solution III was added, the tube inverted to mix, and incubated for 2 minutes on dry ice. The solution was centrifuged for 20 minutes at 15000 RPM in a microfuge and the supernatant decanted into a fresh 1.5 mL eppendorf tube containing 400 µL isopropanol. The solution was incubated on dry ice as before and centrifuged for 10 minutes at 15000 RPM in a microfuge. The supernatant was discarded and the pellet washed with 0.5 mL cold 70 % ethanol. The solution was centrifuged as in the previous step and the supernatant removed with an aspirated drawn out pipet. The DNA pellet was dried under vacuum, resuspended in 100 µL of TE8, and incubated at 37 °C for 10 minutes. The solution was recentrifuged as before to remove the insoluble precipitate and the supernatant was transferred to a fresh tube. To the supernatant, 0.5 µL of RNase A_1 10 mg/mL stock and 0.5 µL of 100 units/µL RNase T_1 from Boehringer Mannheim were added. The plasmid solution was incubated at 37 °C for 1 hour to digest the contaminating RNA. This plasmid DNA was only used for test restriction digests to ensure the correct plasmid was transformed into *B. subtilis*.

2.4.3.4 Plasmid Preparation for Sequencing

Plasmid from small scale *E. coli* preparations (section 2.4.3.1) was further purified for the dideoxy chain termination method (53). In a 0.7 mL eppendorf tube, 100 µL ddH_2O, 40 µL ammonium acetate, 250 µL 95 % ethanol, and 20 µL plasmid were combined. The mixture was incubated for 5 minutes at room temperature, and centrifuged for 5 minutes at 15000 RPM in a microfuge to pellet the plasmid. The supernatant was removed by

aspiration and the pellet dried under vacuum. The plasmid pellet was resuspended in 25 µL 0.4 M and allowed to incubate at room temperature for exactly 10 minutes. After 10 minutes, 7.5 µL 3 M sodium acetate and 160 µL 95 % ethanol were added and the solution placed on dry ice for 10 minutes. The solution was the centrifuged for 10 minutes at 15000 RPM in a microfuge to pellet the DNA. The supernatant was removed using a drawn out pipet with aspiration and the pellet dried under vacuum. The plasmid pellet was resuspended in 10 µL of annealing primer solution consisting of 2 µL 4 ng/µL primer, 2 µL 5 × sequencing buffer, and 6 µL of ddH_2O. The solution was incubated for 20 minutes at 37 °C for annealing, spun down, and placed on ice prior to sequencing. Sequencing was carried out with the Sequenase T7 polymerase (version 2.0) and standard protocols with [α-^{32}P]dATP as supplied by the United States Biochemical Corporation.

2.4.3.5 Separation of Nucleotides by Electrophoresis

Prior to electrophoresis, sequencing samples were heated to 85 °C for 2 minutes. Large 6 % acrylamide sequencing gels (30 × 40 cm) were poured and sequence samples run at 60 watts in 1 × TBE using standard protocols (47). After electrophoresis, the gels were dried using a Bio-Rad gel drier (Model 583) at 80 °C for 1 hour and exposed on Kodak CSB/1 (blue film) or XAR film until the desired band intensity has been reached (typically overnight). The film was developed in a Kodak M35A X-OMAT automatic film developer.

2.5.2 DNA Modifications

All restriction digest were performed according to the manufacturers recommendations. When more than one digest was required with incompatible reaction buffers, or the restriction enzymes had to used in a specific order, the first restriction enzyme was heat inactivated (if possible) and the DNA extracted from solution using the QIAEX Desalting and Concentration Protocol (QIAGEN). The purified plasmid was then resuspended in the next reaction buffer for digestion with the new restriction enzyme. Other DNA modification protocols such as blunt-ending with T_4 DNA polymerase, filling in 5' overhangs with T4 DNA polymerase, dephosphorylation of linearized plasmid with calf intestinal alkaline phosphatase, and ligation of DNA fragments were carried out according to the manufacturers recommendations.

2.5.3 Size Estimation of DNA Fragments by Agarose Gel Electrophoresis

Agarose gel electrophoresis was carried out using standard protocols (47). In general, 1 % (w/v) agarose in 1 x TAE containing 0.5 mg/mL ethidium bromide was used to separate DNA fragments. A maximum 15 µL plasmid and 3 µL loading dye could be easily loaded per lane on a gel. For resolution of small DNA fragments less than 300 bp, 1.5 % (w/v) agarose gels were used. Electrophoresis was performed at 70 volts for 30 mL gels, or 120

volts for 80 mL gels. DNA fragments were compared to 1 kbp ladder standard (BRL), and visualized under ultraviolet light.

2.5.4 Agarose Gel Extraction of DNA Fragments

Plasmid digested with the desired restriction enzymes was precipitated with 2 volumes 95 % ethanol and 0.1 volumes of 3 M sodium acetate. The solution was placed at -80 ^{o}C for about 20 minutes to ensure total precipitation and then centrifuged for 20 minutes at 15000 RPM in a microfuge. The supernatant was removed and the DNA pellet dried under vacuum. The DNA was resuspended in 1 × agarose gel loading dye and run out on a 30 mL 1 % (w/v) agarose gel at 70 V. The DNA fragments were compared to 1 kbp ladder on the gel under low energy UV light (350 nm), and the desired bands were cut out of the gel with a sterile scalpel. The agarose slabs containing the desired bands were placed in a spin column to elute the DNA from the agarose. These columns consist of the lower half of a blue P1000 tip plugged with a small amount of glass wool placed inside a small 0.7 mL Eppendorf tube with a hole punched in the bottom. The small Eppendorf tube sits in a large 1.5 mL Eppendorf tube. The agarose slab was placed on top of the glass wool and the column was spun at 6000 RPM for 10 minutes. The eluent collected in the bottom of the large eppendorf tube was centrifuged for 10 minutes at 15000 RPM to pellet any remaining agarose and the supernatant was transferred to a fresh eppendorf tube. A small amount of the purified plasmid (usually about 3 µL) was run out on an agarose gel and compared to 1 kbp standard in order to verify that the correct fragment was isolated. If the DNA fragment was low in concentration, it was precipitated

in 2 volumes of 95 % ethanol and resuspended in a smaller volume of TE8. No further purification to the fragments was required.

2.6 PLASMID CONSTRUCTION

2.6.1 Oligonucleotide Primers

Oligonucleotide primers required for the construction of several of the mutant gene constructs. All of the oligonucleotides were synthesized on an Applied Biosystems 329 DNA/RNA Synthesizer at the Institute of Molecular Biology and Biochemistry Services Laboratory, Simon Fraser University, Burnaby, BC. Primers complementary to the 5' end of the gene are complementary to the anti-sense strand, and primers complementary to the 3' end of the gene are complementary to the sense strand. In the following figures, mutant residues are shown in **bold**, and restriction sites are underlined and labeled with the specific restriction enzyme. Amino acids coded for by the sense strand are noted above the primer, and mutant amino acids are in bold.

APR: 3' *sprA* primer used to create the gene construct encoding the propeptide of SGPA only.

```
                    L   P   T
5' - AAAGGTACCAGCGGGGTGAACTTGCCC - 3'
          Kpn I
```

BM: 5' *sprB* primer used to create the gene construct encoding the mature region of SGPB only.

```
                        I   S   G
5' - GGATCC GATATCTCCGGCGGCGACGCG - 3'
     BamH I  EcoRV
```

BPΔ4: 5' *sprB* primer used to create the gene construct encoding from the fifth amino acid to the carboxy terminus of SGPB. The third base in the threonine codon is a silent mutation where the wild type guanine was changed to a thymine in order to generate a *Sca* I site.

```
            T   F   S
5' - AGTACTTTCAGTGCCAACCAGCTG - 3'
     Sca I
```

CPΔ11: 5' *sprC* primer used to create the gene construct encoding from the twelfth amino acid to the carboxy terminus of SGPC. In order to generate the *Fsp* I site for cloning, the wild type glycine was replaced with alanine. A silent mutation was created where the wild type guanine was changed to a cytosine in order to generate a *Xho* I site. This site was then used to generate a construct with a 16 amino acid deletion to the propeptide.

```
          A   M   L
5' - TCGGCAATGCTGGCCGCCCTCGAGCG - 3'
     Fsp I                  Xho I
```

CR: 3' *sprC* primer that is complementary to the region of sprC immediately downstream of the stop codon.

```
5' - GAATTCGGGACCGGGGCGG - 3'
     EcoRI
```

2.6.2 Amplification of DNA by PCR

The polymerase chain reaction was used to generate specific deletion mutants and chimeric gene products. The PCR reaction was carried out in 50 µL of 1 × Vent DNA polymerase buffer that contained 1 ng of linearized template DNA, 600 ng of each of the 5' and 3' primers, 25 mM of each dNTP, and 0.5 µL of 2 U/µL Vent DNA polymerase (New England Biolabs). Thirty PCR cycles were used for each PCR reaction and each cycle had a denaturation, annealing and extension step. Denaturation was at 97 ^{o}C for 30

seconds, annealing was at 48 °C to 56 °C (depending on GC content of the DNA) for 1 minute, and extension was at 72 °C for 1 minute per kbp of DNA in the region amplified. Before the first cycle, the DNA was denatured for 2 minutes at 97 °C, and after the cycles the DNA was extended for another 5 minutes at 72 °C. After PCR was complete, 4 µL of a PCR reaction was separated by electrophoresis in a 1 % agarose gel along with 1 Kb ladder standard and visualized under UV light to ensure that a PCR fragment of the correct size was present.

2.6.3 Purification of PCR Fragments

A 10 µL aliquot of the desired PCR reaction was separated by electrophoresis on a 1 % agarose gel. The band corresponding to the desired PCR fragment was gel extracted as previously described (section 2.5.4).

2.7 DELETION MUTANTS FOR *B. Subtilis* EXPRESSION

The propeptide deletion mutants were constructed by using PCR or unique restriction sites in combination with T_4 polymerase to generate blunt ended in-frame deletions. When PCR was used, the PCR fragments were cloned into pUC-18 followed by subcloning into the pEB-11 expression vector. The deletion mutants created by restriction digests were cloned directly into pEB-11. The pEB constructs were initially transformed into *E. coli* TB1 or DH5α for verification by restriction digests and small scale plasmid preparations of the correct constructs. The pEB constructs were then transformed into *B. subtilis* DB104 for expression analysis.

2.7.1 pJB-BPΔ4 (SGPB 4 amino acid propeptide deletion cloning vector)

A four amino acid deletion was made to the amino terminal end of the propeptide as follows. The universal forward primer and BPΔ4 were used as PCR primers with pDS-B8 (12) as a template to amplify the region of *sprB* encoding from the fifth amino acid of the propeptide to the carboxyl terminus of SGPB. The amplified product was digested with *Pst* I and ligated into pUC-18 which had been digested with *Pst* I and treated with CIP. *E. coli* TB1 was transformed with the ligation mix and vectors containing the correct insert were isolated based on restriction enzyme analysis and designated pJB-BPΔ4.

2.7.2 pEB-BPΔ4 (SGPB 4 amino acid propeptide deletion expression vector)

The plasmid pJB-BPΔ4 was digested with *Sca* I, *Pst* I and *Ssp* I. The 0.9 kbp *Sca*I /*Pst*I fragment was gel purified and ligated into pEB-11 (13)that had been digested with *Sma*I and *Pst*I and treated with CIP. The ligation reaction was transformed into *E. coli* TB1, and plasmids containing the correct insert were isolated by restriction analysis and designated pEB-BPΔ4.

2.7.3 pJB-CPΔ11 (SGPC 11 amino acid propeptide deletion cloning vector)

The primers CPΔ11 and CR were used as PCR primers with pDS-C (12)as a template to amplify the region of *sprC* encoding from the twelfth amino acid of the propeptide to the carboxyl terminus of SGPC. The

amplified product was ligated into pUC-18 digested with *Sma*I and treated with CIP. *E. coli* TB1 was transformed with the ligation mix and vectors containing the correct insert in the correct orientation were isolated based on restriction enzyme analysis and designated pJB-CPΔ11.

2.7.4 pEB-CPΔ11 (SGPC 11 amino acid propeptide deletion expression vector)

The plasmid pJB-CPΔ11 was digested with *Pst*I and *Fsp*I. The 1.3 kbp fragment was gel purified and ligated into pEB-11 digested with *Sma*I and *Pst*I and treated with CIP. The ligation reaction was transformed into *E. coli* TB1, and plasmids containing the correct insert were isolated by restriction analysis and designated pEB-CPΔ11.

2.7.5 pEB-CPΔ16 (SGPC 16 amino acid propeptide deletion expression vector)

The plasmid pJB-CPΔ11 was digested with *Xho*I, the 5' overhangs filled in with T4 polymerase, and then digested with *Pst*I. The 13 kbp *spr*C fragment encoding from the seventeenth amino acid of the propeptide to the carboxyl terminus of SGPC was gel purified and ligated into pEB-13 digested with *Sma*I and *Pst*I and treated with CIP. The ligation reaction was transformed into *E. coli* TB1, and plasmids containing the correct insert were isolated by restriction analysis and designated pEB-CPΔ16.

2.7.6 pEB-DPΔ20 (SGPD 20 amino acid propeptide deletion expression vector)

The plasmid pEB-D8 (11)was digested with *Msc*I and *Pst*I. The 1.0 kbp *sprD* fragment encoding from the twenty-first amino acid from the propeptide to the carboxyl terminus of SGPD was gel purified and ligated into pEB-13 digested with *Sma*I and *Pst*I and treated with CIP. The ligation reaction was transformed into *E. coli* TB1, and plasmids containing the correct insert were isolated by restriction analysis and designated pEB-DPΔ20.

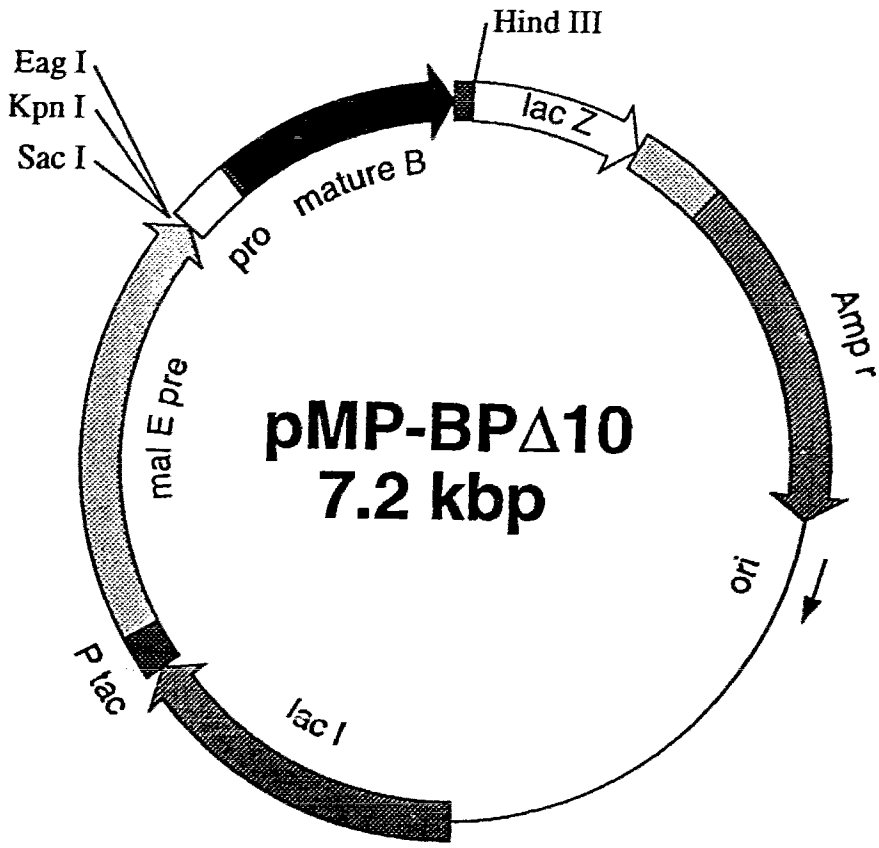
2.7.7 pEB-EPΔ26 (SGPE 26 amino acid propeptide deletion expression vector)

The plasmid pEB-E (13) was partially digested with *Ava*I followed by the addition of T4 polymerase to fill in the 5' overhangs. Plasmid cut only once was gel purified and digested with *Pst*I. The 0.9 kbp *sprE* fragment encoding from the twenty-seventh amino acid of the propeptide to the carboxyl terminus of SGPE was ligated into pEB-13 digested with *Pst*I and *Sma*I and treated with CIP. The ligation reaction was transformed into *E. coli* TB1, and plasmids containing the correct insert were identified by restriction analysis and designated pEB-BPΔ4.

2.7.8 Transformation of pEB Constructs Into *B. subtilis* DB104

All of the pEB vectors were transformed into *B. subtilis* DB104 (section 2.3.4.2). Transformants with the correct plasmid were identified using restriction analysis, and secretion of SGPB, SGPC, SGPD, and SGPE was verified using a skim milk clearing assay (9).

Figure 11. Plasmid Map of the pMP-BPΔ10 *E. coli* Expression Vector. The construction pMP-BPΔ10 is described in section 2.6.6. The pro-mature deletion mutant is expressed as a fusion protein with the maltose binding protein and maltose binding protein signal sequence (mal E). The pMAL-p vector is described in section 2.4.1.1.


Hind III
Eag I
Kpn I
Sac I
lac Z
pro
mature B
Amp r
mal E pre
pMP-BPΔ10
7.2 kbp
ori
P tac
lac I

2.8 *sprB* DELETION MUTANT FOR EXPRESSION IN *E. coli*

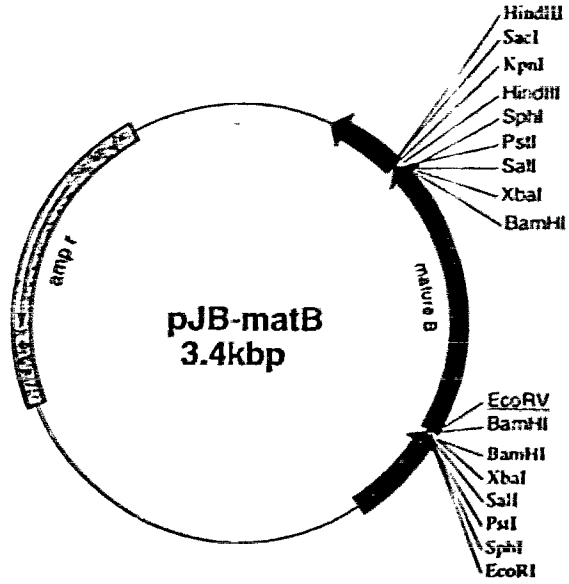
This plasmid was created in order to compare results with other *in vivo* studies that express pre-pro-mature proteases in *E. coli* (17, 28) The plasmid pNCB (12) was digested with *Pvu*II and *Pst*I. The 0.9 kbp fragment encoding from the eleventh amino acid to the carboxy terminus of SGPB was gel purified and ligated into pMAL-p (Figure 10) digested with *Stu*I and *Pst*I and treated with CIP. The ligation mix was used to transform *E. coli* TB1 and vectors containing inserts of the correct size were identified using restriction enzyme analysis and designated pMP-BPΔ10 (Figure 11).

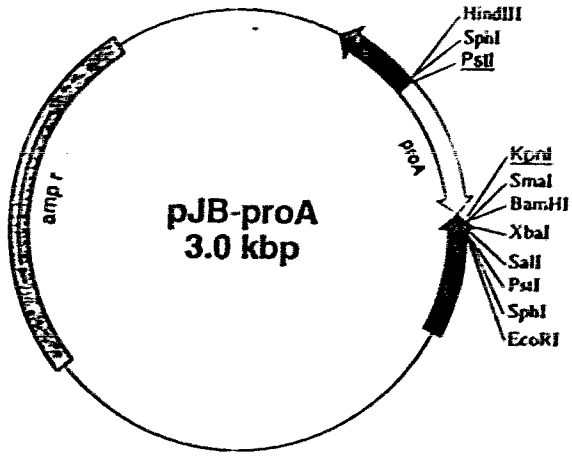
2.9 PRO A - MATURE B CHIMERIC GENE CONSTRUCT

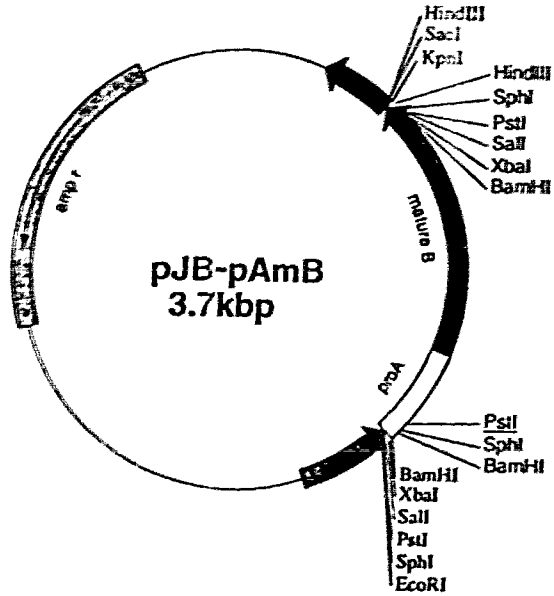
A chimeric gene construct was made that consisted of the propeptide encoding region of *sprA* fused in-frame to the mature encoding region of *sprB* (Figure 12). The universal forward primer and APR were used as PCR primers with pDS-A3 as template to amplify the region of *sprA* encoding the propeptide of SGPA. The amplified product was digested with *Hin*dIII and *Kpn*I and ligated into pUC-18 digested with *Hin*dIII and *Kpn*I and treated with CIP. The ligation mix was used to transform *E. coli* TB1. Vectors containing the correct insert were identified using restriction enzyme analysis and designated pJB-proA. The universal reverse primer and BM primer were used as PCR primers with pNCB (12) a template to amplify the region of *sprB* encoding the mature domain of SGPB. The amplified product was ligated into pUC-18 cut with *Sma*I and treated with CIP. The ligation mix was used to transform *E. coli* DH5α and vectors containing the correct insert in the correct orientation (5' end of *sprB* at the *Hin*dIII side of the polylinker) were selected

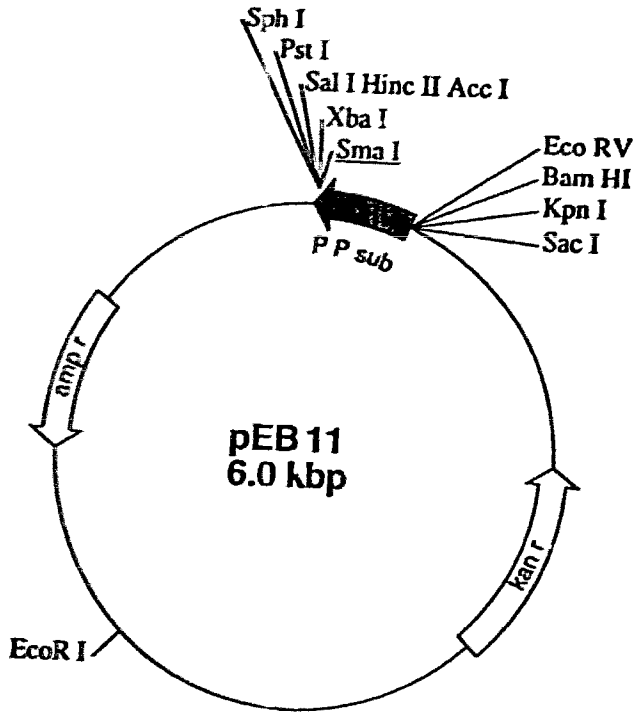
and designated pJB-matB. The plasmid pJB -proA was digested with *Hin*dIII and *Kpn*I and treated with T4 polymerase to produce blunt ends. The resulting 0.2 kbp fragment was gel purified and ligated into pJB-matB that had been digested with *Eco*RV and treated with CIP. The ligation mix was used to transform *E. coli* DH5α. A plasmid containing the pro-*sprA* in the same orientation as the mature-*sprB* was isolated using restriction enzyme analysis and designated pJB-pAmB. The plasmid pJB-pAmB was digested with *Pst*I, treated with T4 polymerase to produce blunt end and then digested with *Xba*I. The 1.0 kbp band was gel purified and ligated into pEB-11 digested with *Xba*I and *Sma*I and treated with CIP. The ligation reaction was transformed into *E. coli* TB1, and plasmids containing the correct insert were identified by restriction analysis and designated pEB-pAmB. The vector was then transformed into *B. subtilis* DB104. Transformants with the correct plasmid were identified using restriction analysis, and secretion of SGPB was verified using a skim milk clearing assay (9).

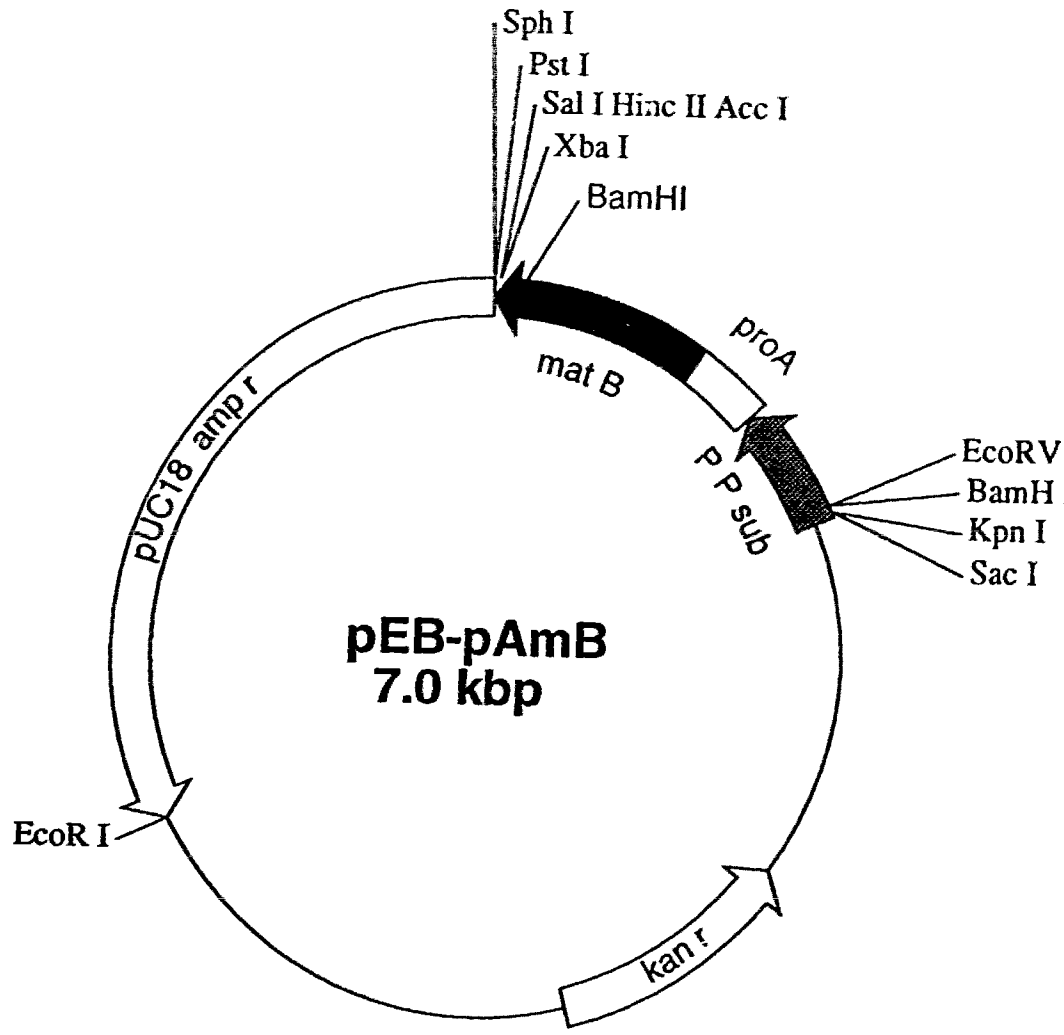
Figure 12. Cloning Summary of pEB-pAmB. The generation of pEB-pAmB is outlined starting from the constructs pJB-proA and pJB-matB as described in section 2.9. Plain text designate restriction sites subcloned into the vector along with the gene insert. Underlined text designate restriction sites engineered to generate blunt DNA for in-frame ligations with the pre sequence in pEB-11.


HindIII
SacI
KpnI
HindIII
SphI
PstI
SalI
XbaI
BamHI
amp r
mature B
pJB-matB
3.4kbp
EcoRV
BamHI
BamHI
XbaI
SalI
PstI
SphI
EcoRI


HindIII
SphI
PstI
proA
KpnI
SmaI
BamHI
XbaI
SalI
PstI
SphI
EcoRI
amp r
pJB-proA
3.0 kbp


HindIII
SacI
KpnI
HindIII
SphI
PstI
SalI
XbaI
BamHI
amp r
mature B
pJB-pAmB
3.7kbp
proA
PstI
SphI
BamHI
BamHI
XbaI
SalI
PstI
SphI
EcoRI


Sph I
Pst I
Sal I Hinc II Acc I
Xba I
Sma I
Eco RV
Bam HI
Kpn I
Sac I
P P sub
amp r
pEB 11
6.0 kbp
kan r
EcoR I


Sph I
Pst I
Sal I Hinc II Acc I
Xba I
BamHI
proA
mat B
P P sub
EcoRV
BamH I
Kpn I
Sac I
pUC18 amp r
pEB-pAmB
7.0 kbp
EcoR I
kan r

2.10 PROTEASE EXPRESSION IN *B. subtilis* DB104

2.10.1 Expression for Time Course Analysis

B. subtilis DB104 transformants were streaked onto TY media containing 1.5% skim milk powder and 50 µg/mL kanamycin. Single colonies were used to inoculate 2 mL TY/Kan and grown overnight at 30 °C with vigorous shaking. From the 2 mL cultures, 50 µL aliquots were removed to inoculate 50 mL TY/Kan in 125 mL Erlenmeyer flasks. The 50 mL cultures were grown in duplicate at 30 °C with shaking at 200 RPM in a Lab-Line environmental shaker. At various time points, 1 mL aliquots were removed for measuring cell growth and activity. A 100 µL sample from the aliquot was diluted with 900 µL TY broth and cell growth was measured spectrophotometrically at 600 nm. The remaining cells were centrifuged at 15000 RPM for 3 minutes in a microfuge and protease activity was measured directly from 20 µL supernatant aliquots (section 2.10.4).

2.10.2 Partial Purification of SGPB

A single *B. subtilis* colony harbouring a pEB vector for either mutant or wild type SGPB expression was used to inoculate 2 mL TY/Kan and grown overnight at 30 °C with vigorous shaking. From the 2 mL culture, 1 mL aliquots were used to inoculate duplicate 1 L of S-broth in a 2 L Erlenmeyer flask. The cultures were grown for 72 hours at 30 °C with shaking at 180 RPM. The cultures were cooled on ice and then centrifuged at 5000 RPM for 10 minutes in a GS-3 rotor at 4 °C. The supernatant containing the secreted protease was saved, and the protease concentrated and partially purified

using a modified version of the protocol developed by Sidhu (11). To 1980 mL of supernatant chilled on ice, 523.26 g of ammonium sulphate was gradually added with stirring and mixed for an additional 10 minutes after addition of the ammonium sulphate. The supernatant solution was then divided into GS-A centrifuge tubes and centrifuged at 8000 RPM for 60 minutes at 4 ^{o}C. To supernatant, an additional 261.9 g ammonium sulphate was dissolved and the mixture was centrifuged as before. The supernatant was discarded and the pellet, containing protease, resuspended in 30 mL of Buffer A. Next, acetone was added to 30 % (v/v) and allowed to mix for 10 minutes while on ice. The solution was divided into SS-34 tubes and centrifuged at 15000 RPM for 15 minutes at 4 ^{o}C. The supernatant was saved and acetone added to 70 % (v/v) and mixed on ice as before. The solution was centrifuged at 10000 RPM for 20 minutes in a GS-A rotor. The pellet was saved and resuspended in 4 mL Buffer A. The resuspended solution was divided into 1.5 mL eppendorf tubes and centrifuged at 15000 RPM for 10 minutes to remove any insoluble precipitate. The supernatant was removed and the tubes stored at -80 ^{o}C.

2.10.3 FPLC of Partially Purified Protease

Further purification of the partially purified protease (section 2.10.2) was required for amino terminal analysis. A 3.0 mL aliquot of the partially purified protein solution was filtered through a 0.22 µm filter and diluted to 25 mL in Buffer A. The protein solution was degassed and loaded onto a Waters Protein-Pak SP FPLC column previously equilibrated with Buffer A. The protein was eluted at 1 mL/minute using a gradient of 10 % to 30 % (v/v) Buffer B in 80 minutes. Peaks were collected manually and assayed for

protease activity (section 2.10.4). The most active fractions were retained for amino terminal analysis.

2.10.4 Activity Assays

Enzyme activity for all of the proteases except SGPE was determined spectrophotometrically at 412 nm using the following assay mixture: 1.00 mL of 0.1 mg/mL *N*-succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (54)in 50 mM buffer (pH 8.0) + 20 µL of culture supernatant or 10 µL of a FPLC fraction. For SGPE, the following assay mixture was used: 0.1 mg/mL *N*-succinyl-Ala-Ala-Pro-Glu-*p*-nitroanilide in 50 mM / 5% methanol (pH 8.0) + 20 µL of culture supernatant. A unit of activity is defined as the amount of enzyme required to produce 1 mmol of *p*-nitroaniline in 1 hour at 20 ^{o}C.

2.10.5 SDS-Polyacrylamide Gel Electrophoresis

Proteins were separated by electrophoresis on 15 % polyacrylamide gels using the Laemmeli buffer system (55). Protein samples of up to 25 µL total volume in 1 × loading dye in 0.7 mL eppendorf tubes were denatured by heating at 98 ^{o}C for 10 minutes. The samples and 5 µL of prestained low molecular weight marker (Bio-Rad) as a standard were then loaded onto the polyacrylamide gel that was assembled using the Mini-PROTEAN II system (Bio-Rad). The gel was run at 200 volts (constant) until the bromophenol dye front or the desired molecular weight marker reached the bottom of the gel. After electrophoresis, the gel was stained in Coomassie Stain for 20 minutes with gentle shaking, followed by destaining in Fast Destain for 20 minutes and the Slow Destain until all of the background colour disappears. The

destained gel was then permanently recorded by scanning it using an Apple OneScanner desktop scanner (Model M1381) and the Ofoto program (Light Source Computer Images). The gel image could then printed out or manipulated further using the Adobe Photoshop program (version 2.0.1).

2.10.6 Amino Terminal Analysis of Secreted Protease

A 2 mL aliquot of the most active FPLC fraction (section 2.10.3) was concentrated to 75 µL in a Centricon-10 microconcentrator by centrifugation in a SS-34 rotor at 5500 RPM for 2 hours. Samples were analysed on 15 % SDS-PAGE gels with SGPB standard to determine the amount required to produce an intense band after coomassie staining for amino terminal analysis (about 8 µL). A large amount of protein was required for a successful transfer to Immobilon-PVDF membrane for amino terminal sequencing. A 15 % SDS-PAGE that was allowed to polymerize at 4 ^{o}C for at least 24 hours was loaded with the concentrated FPLC fraction, SGPB standard, and prestained low molecular weight standard (Bio-Rad) in duplicate so one half of the gel could be stained with Coomassie Stain for a permanent record (section 2.10.5) and the other half transferred to Immobilon-PVDF membrane. Electrophoresis was carried out at 200 volts until the lysozyme molecular weight marker (18500 MW) reached the bottom of the gel. This allowed the maximal separation between protein bands. The gel half used for transfer was rinsed in ddH_2O and allowed to soak in Transfer Buffer. Four sheets of filter paper and the PVDF membrane were cut to the size of the gel. Two of the filter paper sheets were soaked in filter paper, the PVDF membrane was soaked in methanol for 3 seconds, washed in ddH_2O for 1 minute, then soaked in Transfer Buffer. The transfer was carried out by electroblotting with a LKB

2117-250 Novablot Electrophoretic Transfer Kit. The blot was assembled in layers from the anode to the cathode as follows: 2 sheets of filter paper, membrane, SDS-PAGE gel (soaked in transfer buffer), and topped with 2 sheets of filter paper soaked in transfer buffer. The apparatus was run at 1 mA / cm^2 for 2 hours. After electroblotting, the membrane was rinsed for 5 minutes with ddH_2O three times, then stained with PVDF Stain for 5 minutes. The membrane was destained for 10 minutes in 50 % methanol or until the background disappears, and then blotted dry on filter paper. The band corresponding to the SGPB standard was cut out and sent to the Protein Microchemistry Centre, Department of Biochemistry and Molecular Biology, University of Victoria, Victoria, BC for amino terminal sequencing of the first 5 residues.

2.11 TRANSCRIPTIONAL ANALYSIS OF PEB CONSTRUCTS IN *B. subtilis*

2.11.1 RNA Isolation

Total RNA was isolated from *B. subtilis* cultures harbouring a pEB vector. The cultures were grown with shaking in 50 mL TY/Kan media at 30 °C. Aliquots of 1 mL were taken at various times and added to 1.5 mL eppendorf tubes. The cells were pelleted by centrifugation at 15000 RPM for 3 minutes in a microfuge and the supernatant removed. The cell pellets were frozen at -80 °C if not immediately used. The RNA was extracted using a modified *B. subtilis* plasmid preparation where DEPC was added to the solutions to inhibit RNases and Solution III was omitted to prevent the loss of RNA. The cells were resuspended in 200 µL of Solution I containing 0.09 % (w/v) DEPC. Next, the cells were lysed with 200 µL of Solution II containing 0.05 % (w/v) DEPC and extracted twice with 0.5 mL phenol/chloroform (1:1). Since solution III was omitted, there was a large amount of denatured protein at the interphase after the addition of phenol and chloroform. It was sometimes necessary to remove the protein with a pipet followed by recentrifugation in order to prevent protein contamination in the aqueous phase containing the RNA. Two volumes of 95 % ethanol was added to the aqueous phase and stored at -80 °C for 20 minutes. The precipitated RNA was centrifuged for 20 minutes at 15000 RPM and the supernatant removed with a drawn out pipet. The RNA pellet was dried under vacuum and the pellet resuspended in 50 µL ddH_2O containing 0.1 % (w/v) DEPC. The RNA solution was stored at -80 °C if not used immediately.

2.11.2 RNA Agarose Gel Electrophoresis

The RNA prepared in section 2.11.2 was electrophoresed using a 1.5 % agarose gels containing formaldehyde. For a 70 mL gel, 1.05 g of agarose was dissolved in a solution of 59.5 mL ddH_20 and 7.0 mL 10 × MOPS Buffer by heating in a 600 W microwave oven on high power for 90 seconds. After cooling the solution to approximately 60 °C, 3.5 mL of deionized formaldehyde was added and the gel poured. To 10 µL of isolated RNA, 5 µL of 3 X Loading Buffer was added and the solution heated at 60 °C for 15 minutes. Before the sample was loaded onto the gel, 0.5 µL of 1 mg/mL ethidium was added to the RNA sample. The RNA was electrophoresed at 120 volts until the rRNA bands were well resolved as determined by visualisation under UV light.

2.11.3 RNA Northern Transfer

Total RNA that was electrophoresed on a 1.5% formaldehyde gel was centrifugally transferred to Zeta-probe membrane using the spin-blot procedure of Wilkins and Snell (56). After electrophoresis the gel was soaked in ddH_20 for 20 minutes followed by 10 × SSC for 45 minutes. Filter paper and membrane were cut to gel size. The membrane was soaked in ddH_20 and then 10 × SSC. The spin blot was assembled in the following order from the bottom to the top: an inverted microtitre plate lid (8 × 12 cm), 4 mm paper towel, 2 dry pieces of filter paper, 1 piece of filter paper soaked in 10 × SSC, Zeta-probe cut to the formaldehyde gel size, and 2 pieces of filter paper soaked in 10 × SSC on top. The blot was centrifuged at 1000 RPM for 45 minutes in an International Equipment Co. HN-S centrifuge equipped with a

microtiter rotor. After centrifugation, the gel wells were marked directly onto the Zeta-probe membrane with a pencil before the gel was removed. The membrane was then washed in 20 × SSC and then the RNA was covalently cross-linked to the membrane using a UV Stratalinker 2400 according to the manufacturer's instructions. After crosslinking, the membrane was washed in 1 % (w/v) SDS for 15 minutes and then sealed in a plastic bag containing 0.3 × SSC for storage at 4 °C.

2.11.4 DNA Probe Creation

The *Nae*I / *Pst*I fragment of mature *sprB* DNA from pJB-BPΔ4 was labeled with [α-^{32}P]dATP using New England Biolabs' NEBlot kit according to the manufacturers instructions. The probe was purified using a Sephadex G-50 column in a pasteur pipet plugged with glass wool. The column was equilibrated with TE8 buffer, and the crude probe sample added to the resin surface. The probe was eluted with TE8 and 15 drop fractions were collected in 1.5 mL Eppendorf tubes. After the fractions were collected, 5 µL aliquots were removed from each fraction and added to 0.7 mL eppendorf tubes that were placed in small scintillation tubes. A Beckman LS 6000 SC scintillation counter was used to determine the activity in counts per minute from the ^{32}P. The most active fraction was used as the RNA probe.

2.11.5 RNA Northern Hybridization

The Zeta-probe membrane with covalently linked RNA was placed in a short hybridization tube and incubated for 1 hour with 7 mL Hybridization Solution at 42 °C in a hybridization oven. The entire probe fraction that was

purified on the G-50 column (section 2.11.4) was added to the membrane and incubated overnight at 42 °C. In the morning, the membrane was washed twice with 2 × SSC / 0.3 % SDS for 20 minutes at room temperature. Then it was washed once with 1 × SSC / 0.5 % SDS , followed by two 0.3 X SSC / 1.0 % SDS washes, all for 20 minutes at 65 °C. The activity in counts per minute was monitored between washes using a hand held Geiger counter, if they dropped significantly, the washes were stopped. After the last wash, the membrane was washed with ddH_2O and sealed in a plastic bag using a vacuum sealer for storage at 4 °C. The blots were developed on Kodak CSB/1 (blue film) at -80 °C until the desired band intensity was reached.

3. RESULTS

3.1 CLONE SUMMARY

To analyse the effects of the propeptide on the generation of active protease, deletions were made to the propeptide-encoding region of *sprB* and a chimeric gene was constructed containing the propeptide encoding region of *sprA* and the mature encoding region of *sprB* as described in Materials and Methods (section 2.6). The *S. griseus* protease propeptide deletion mutants used for expression are summarised in Table 2. All of the constructs in Table 2 were for expression in *B. subtilis* DB104 except pMP-BPΔ10 which was for expression in *E. coli* TB1. The *B. subtilis* expression construct containing the chimeric gene was designated pEB-pAmB.

The constructs generated using PCR were sequenced by the dideoxy method (section 2.4.2.4 and 2.4.2.5). All of the constructs were verified by restriction digest analysis.

Table 2. Deletion Mutant Construct Summary. This table lists the constructs, the protease that each construct expresses, and the size of deletion to the amino terminal end of the propeptide each mutant gene produces.

Construct	Protease	Propeptide Deletion
pEB-B8	SGPB	0
pEB-BPΔ4	SGPB	4
pEB-BPΔ10	SGPB	10
pEB-BPΔ15	SGPB	15
pEB-BPΔ20	SGPB	20
pMP-BPΔ10	SGPB	10
pEB-C8	SGPC	0
pEB-CPΔ11	SGPC	11
pEB-CPΔ16	SGPC	16
pEB-D8	SGPD	0
pEB-DPΔ20	SGPD	20
pEB-E	SGPE	0
pEB-EPΔ26	SGPE	26

3.2 CHIMERA MUTANT RESULTS

3.2.1 Growth and Activity of *B. subtilis* Cultures Harbouring pEB-pAmB

Growth of *B. subtilis* cultures harbouring pEB-pAmB did not significantly differ from cells harbouring pEB-B8 over 58 hours (Figure 13). Colonies harbouring pEB-pAmB produced zones of clearing on skim milk plates, indicating the secretion of active protease (Figure 14). After 74 hours of growth at 30 ^{o}C, the chimeric gene produced 6.0 % wild type *sprB* activity normalised for cell density.

3.2.2 Northern Blot Analysis of *B. subtilis* Cultures Harbouring pEB-pAmB

Transcription of the chimeric gene in *B. subtilis* was analysed by Northern blot analysis (Figure 15). The total RNA was electrophoresed and probed with the mature encoding region of *sprB* . Significant hybridization occurred with pEB-B8 (the positive control, lane 2), and with the chimeric gene (lane 3). This indicates that the reduced activity of the chimera is not due to reduced transcriptional activity. No hybridization could be seen with the expression vector pEB-11 (the negative control, lane 1).

Figure 13. Growth of *B. subtilis* Harbouring pEB-pAmB. Duplicate 50 mL cultures of *B. subtilis* harbouring pEB-11, pEB-B8, or pEB-pAmB were grown at 30 °C and monitored for cell growth spectrophotometrically at 600 nm as described in section 2.7.1. The average cell density of the duplicate cultures was reported at each time point.

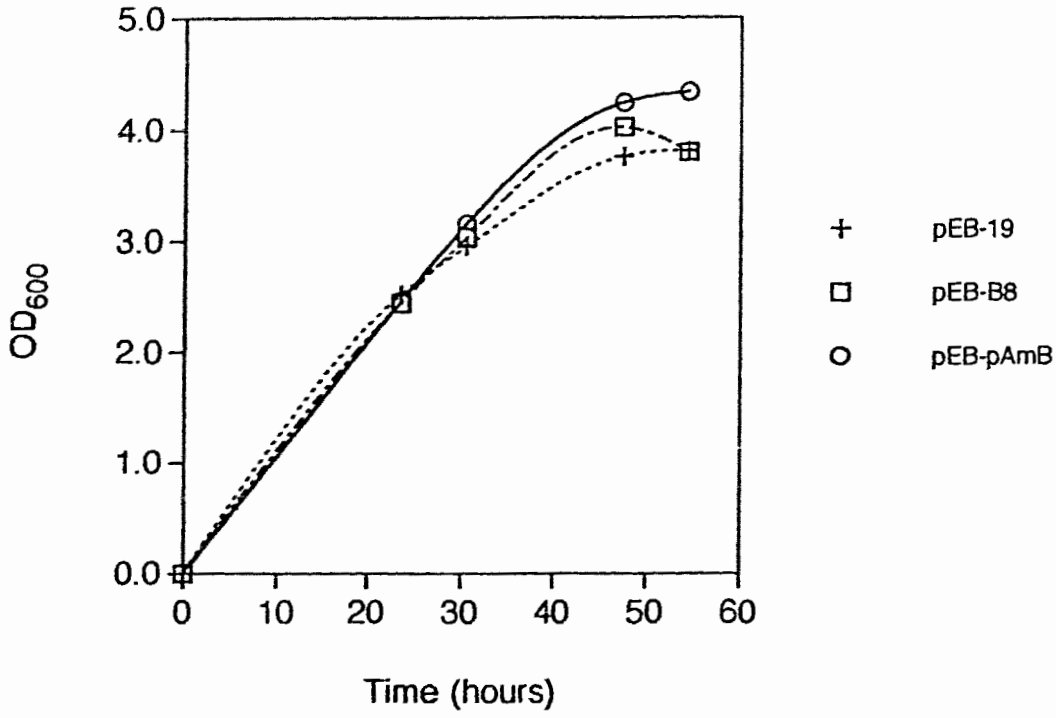

5.0
4.0
3.0
2.0
1.0
0.0
OD_{600}
0
10
20
30
40
50
60
Time (hours)
pEB-19
pEB-B8
pEB-pAmB

Figure 14. Recombinant Expression of SGPB and Chimera in *B. subtilis*. *B. subtilis* DB104 transformants harbouring 1) pEB-pAmB, 2) pEB-B8 or 3) pEB-11 were plated on YT/Milk/Kan plates and photographed after 18 hours of incubation at 30 °C. Zones of clearing indicating the secretion of active protease are visible with both the pEB-B8 and pEB-pAmB transformants; the diameter of each zone of clearing is proportional to proteolytic activity.

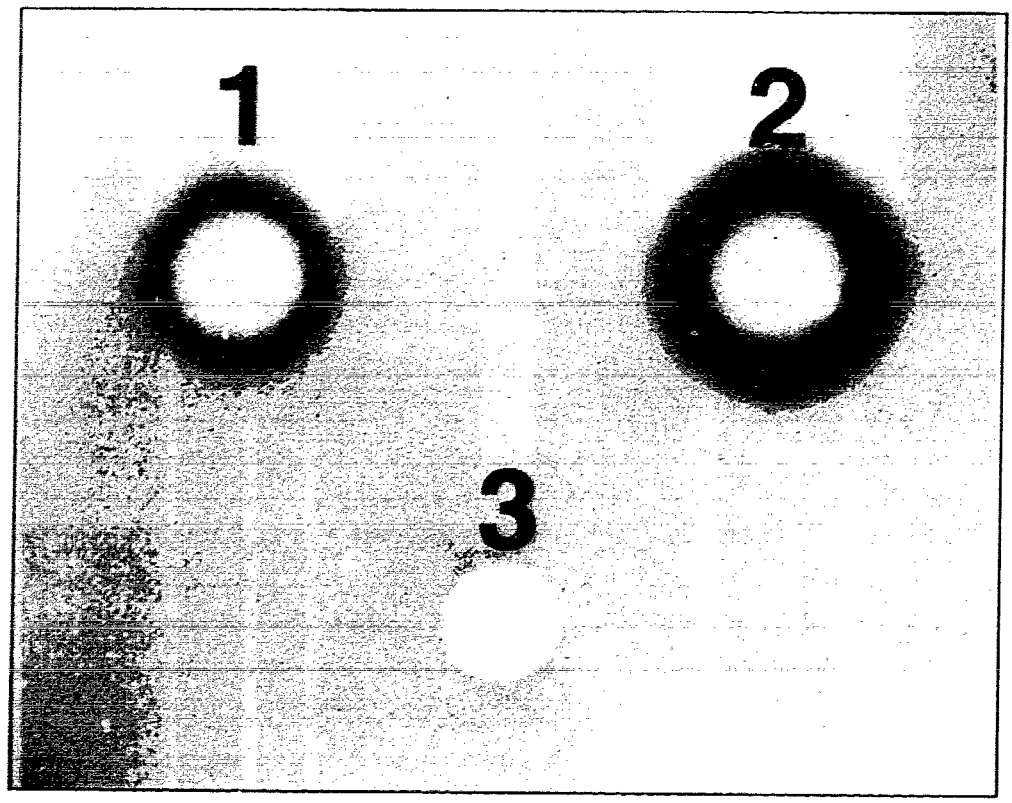

1
2
3

Figure 15. Northern Blot Analysis of the Chimeric Gene. After 36 hours of growth at 30 ^{o}C, total RNA was isolated from *B. subtilis* transformants harbouring **1)** pEB-19, **2)** pEB-B8, **3)** pEB-pAmB and hybridized with a radiolabeled segment of mature *sprB* as described in section 2.8. Significant hybridization could be seen in all lanes except lane 1 (the negative control) indicating that the chimeric gene is being transcribed at approximately the same levels as wild type *sprB*.

1 2 3

3.3 DELETION MUTANT RESULTS

3.3.1 Growth of *B. subtilis* Cultures Harbouring Deletion Mutants

B. subtilis cultures harbouring the deletion mutants were grown in 50 mL 1 × TY/Kan and monitored for growth over time. Figure 16 shows a typical growth time course for *B. subtilis* cultures harbouring pEB-B8, pEB-BPΔ4, pEB-BPΔ10, pEB-BPΔ15, and pEB-BPΔ20 at 30 °C. The growth of cells harbouring pEB-B8 (secreting wild type SGPB) plateaued after 30 hours. Cells harbouring pEB-BPΔ10, pEB-BPΔ15 and pEB-BPΔ20 had their growth plateau after 50 hours, while pEB-BPΔ4 had it's growth plateau about 7 hours earlier. The 10, 15 and 20 amino acid deletion mutants had their growth remain at the highest OD_{600} level, cells harbouring wild type *sprB* (pEB-B8) had the lowest overall OD_{600} values ending at 2.0 after 98 hours of growth, and cells harbouring the 4 amino acid deletion mutant had intermediate OD_{600} values. Similar results were observed with wild type and mutant SGPD, cultures harbouring the wild type protease completed their log phase of growth much earlier and had a lower cell density than the deletion mutant culture (Figure 17A). Cells harbouring wild type SGPC had a drop in growth at 50 hours but recovered and ended with a cell density like the deletion mutants (Figure 17B).

Figure 16. Growth of *B. subtilis* Harbouring the *sprB* Deletion Mutants. Duplicate 50 mL cultures of *B. subtilis* harbouring pEB-11, pEB-B8, pEB-BPΔ4, pEB-BPΔ10, pEB-BPΔ15 or pEB-BPΔ20 were grown at 30 °C and monitored for cell growth spectrophotometrically at 600 nm as described in section 2.7.1. The average cell density of the duplicate cultures was reported at each time point.

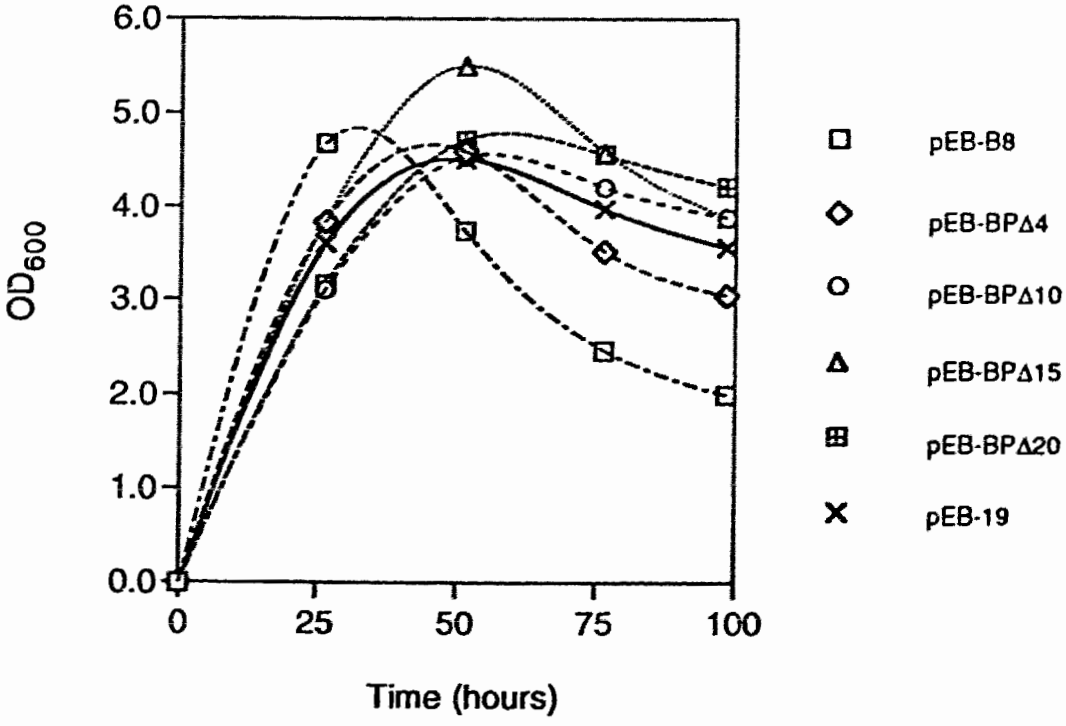
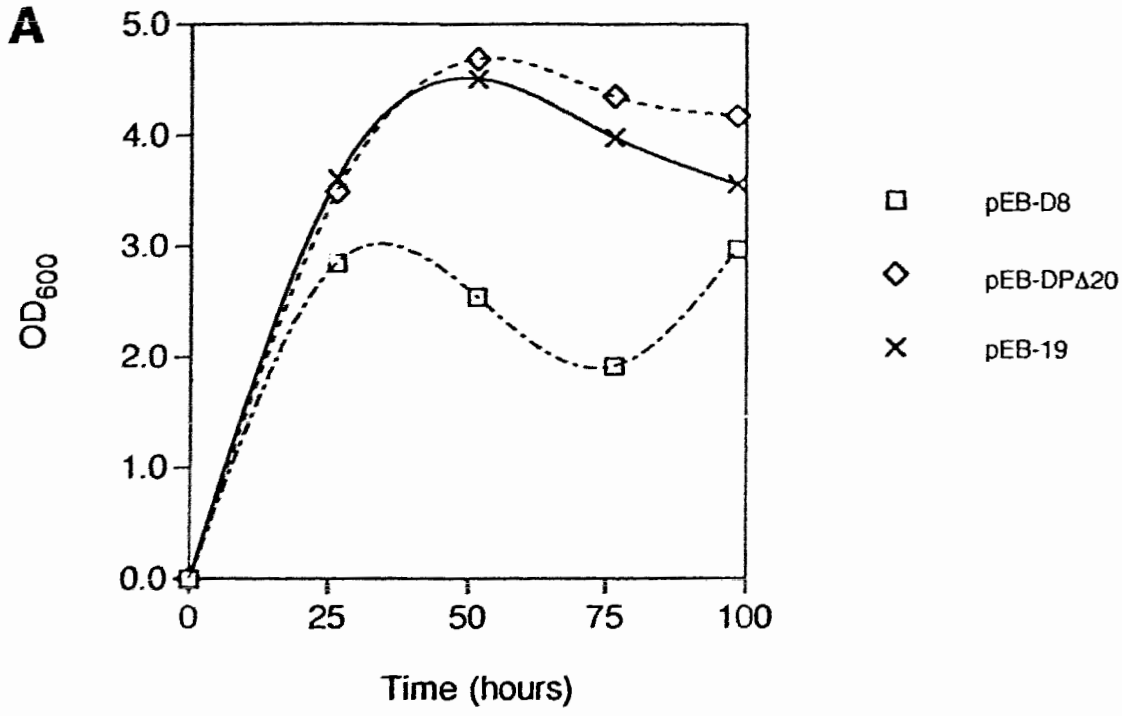
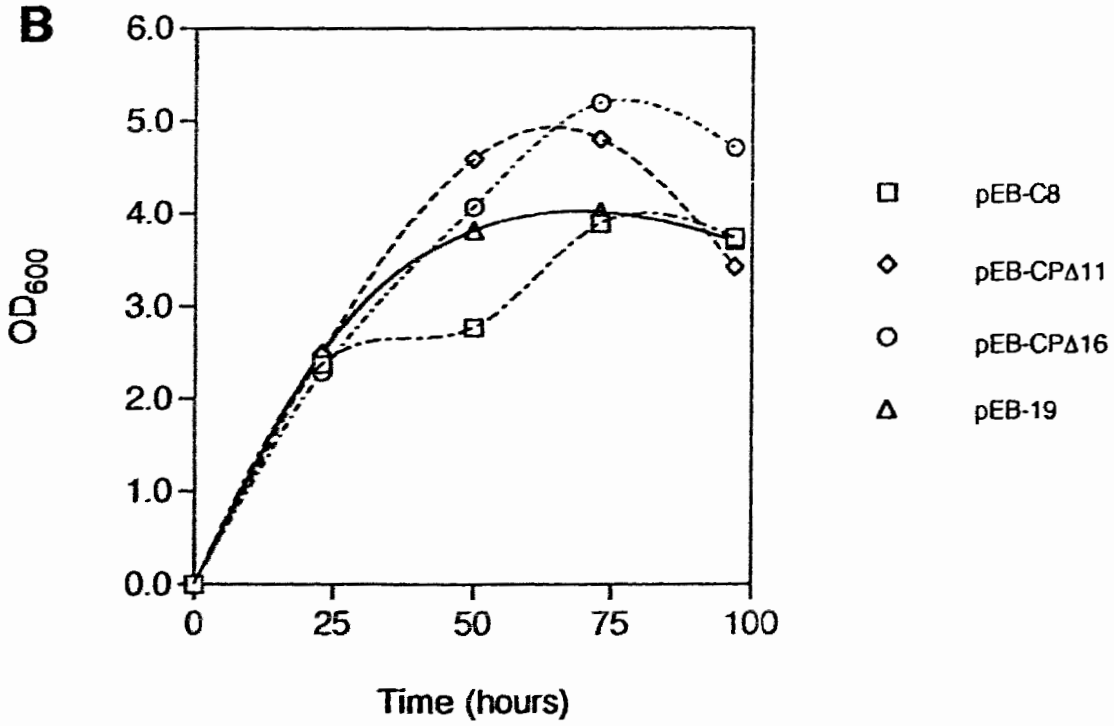

6.0
5.0
4.0
3.0
2.0
1.0
0.0
OD_{600}
0
25
50
75
100
Time (hours)
pEB-B8
pEB-BPΔ4
pEB-BPΔ10
pEB-BPΔ15
pEB-BPΔ20
pEB-19

Figure 17. Growth of *B. subtilis* Harbouring the *sprC* and *sprD* Deletion Mutants. Duplicate 50 mL cultures of *B. subtilis* harbouring **A)** pEB-11, pEB-D8, or pEB-DPΔ20; **B)** pEB-11, pEB-C8, pEB-CPΔ11, or pEB-CPΔ16 were grown at 30 °C and monitored for cell growth spectrophotometrically at 600 nm as described in section 2.7.1. The average cell density of the duplicate cultures was reported at each time point.


A
5.0
4.0
3.0
2.0
1.0
0.0
OD600
0
25
50
75
100
Time (hours)
pEB-D8
pEB-DPΔ20
pEB-19


B
6.0
5.0
4.0
3.0
2.0
1.0
0.0
OD600
0
25
50
75
100
Time (hours)
pEB-C8
pEB-CPΔ11
pEB-CPΔ16
pEB-19

3.3.2 Activity of *B. subtilis* Cultures Harbouring Deletion Mutants

While the *B. subtilis* cultures were monitored for growth, the cell culture supernatants were monitored for secretion of active protease. Figure 18 shows the time course of activities (normalised for cell density) of pEB-B8, pEB-BPΔ4, pEB-BPΔ10, pEB-BPΔ15, and pEB-BPΔ20 cultures. The negative control, pEB-11, did not produce any detectable protease activity at any time, therefore protease activity depends on SGPB production. The detection limit for the assay was found to be 0.2 ng SGPB. The overproduction of active protease in the cultures harbouring wild type protease genes appears to contribute to the lower observed cell density like that seen with the pEB-B8 culture in Figure 16. Figure 19 shows the maximum activities from the various deletion mutant genes normalised for cell density in comparison to wild type *sprB* after the log phase of growth at 76.5 hours. The activities correspond to 39.5 %, 8.6 %, 1.1 %, and 0.16 % wild type activity for the 4, 10, 15 and 20 amino acid propeptide deletions respectively. With increasing size in deletions to the propeptide of SGPB, there was a corresponding exponential decrease in activity normalised for cell density. The log of the normalised activity was plotted as a function of amino acid deletion from the propeptide amino terminus, and the data generated a linear slope (Figure 20). Data from 51.5 and 98.5 hours also produced slopes of -0.141 indicating the slope at 76.5 hours is not time dependent (data not shown). In order to determine if the linear relationship was temperature dependent, pEB-BPΔ10, pEB-BPΔ15 and pEB-BPΔ20 were grown at 21 ^{o}C, 30 ^{o}C, and 37 ^{o}C. The log of normalised activity plotted as a function of amino acid deletion produced parallel lines at each temperature (Figure 21).

Deletions to the propeptide encoding regions of *sprC*, *sprD*, and *sprE* also produced reduced, but detectable levels of activity similar to the *sprB* deletion mutants. Again, the activities were normalised for cell density and compared to wild type activity (Table 3). The plasmids pEB-CPΔ11 and pEB-CPΔ16 produced 3.0 % and 0.05 % maximum wild type pEB-C8 activity respectively. The plasmids pEB-DPΔ20 and pEB-EPΔ26 produce 3.4 % and 1.7 % of the maximal wild type activity of pEB-D8 and pEB-E respectively.

Figure 18. Normalised Activity of the Deletion Mutants. *B. subtilis* cultures harbouring pEB-B8, pEB-BPΔ4, pEB-BPΔ10, pEB-BPΔ15, or pEB-BPΔ20 were grown at 30 °C and assayed for protease activity as described in sections 2.7.1 and 2.7.4. At each time point, activity was normalised for cell growth by dividing the activity by the cell density. The average normalised activity of the duplicate cultures is reported at each time point.

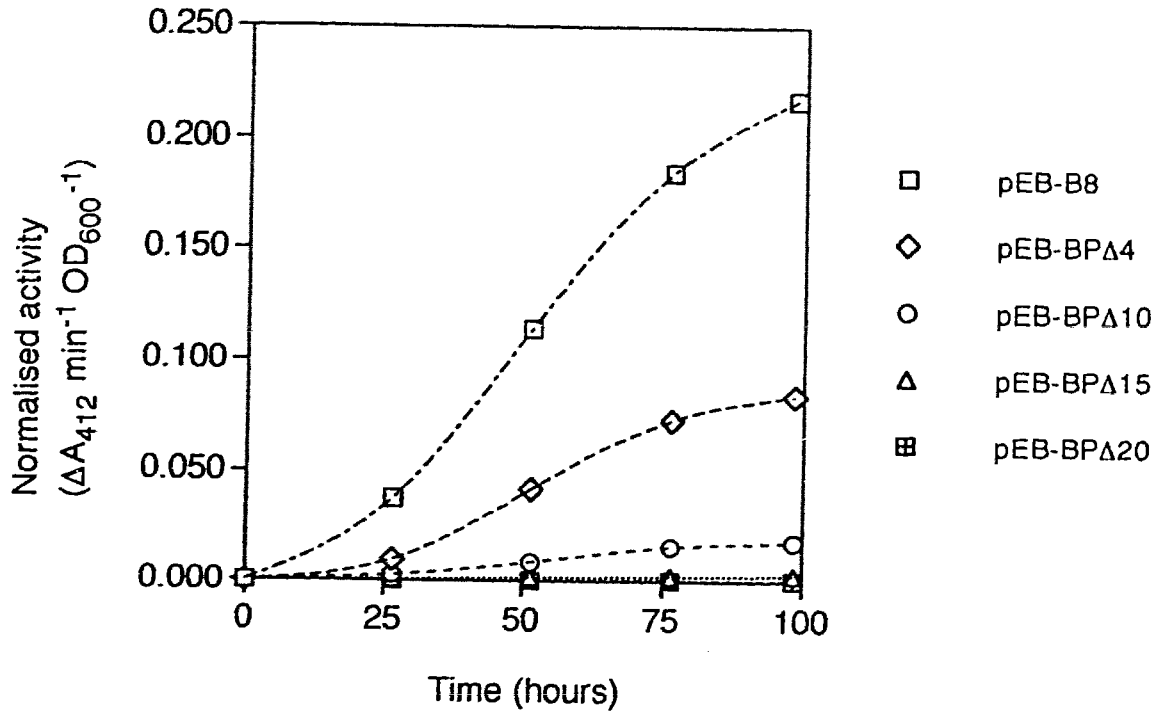
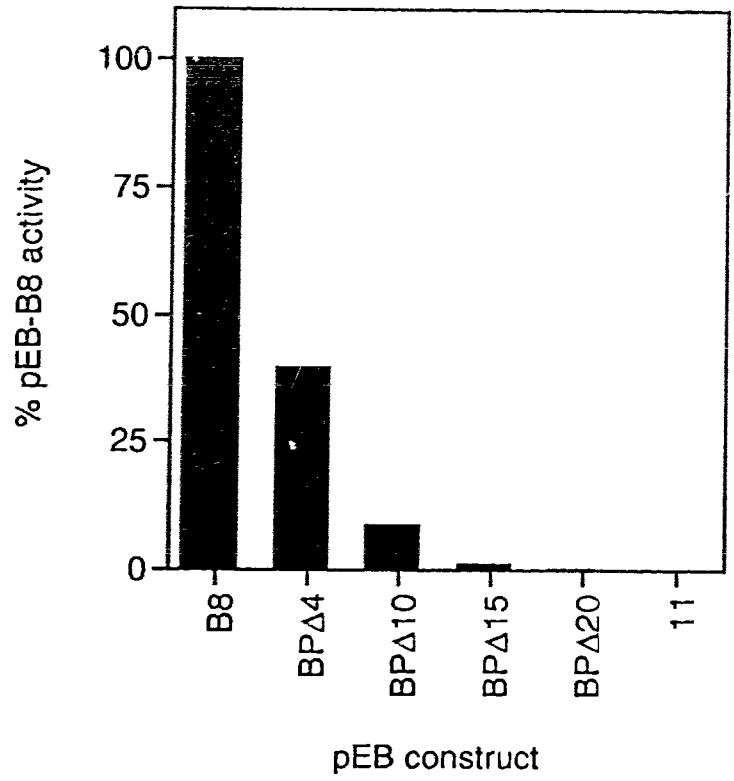

0.250
0.200
0.150
0.100
0.050
0.000
0
25
50
75
100
Normalised activity
(ΔA_{412} min^{-1} OD_{600}^{-1})
Time (hours)
pEB-B8
pEB-BPΔ4
pEB-BPΔ10
pEB-BPΔ15
pEB-BPΔ20

Figure 19. Relative Activity of the Deletion Mutants. *B. subtilis* cultures harbouring pEB-B8, pEB-BPΔ4, pEB-BPΔ10, pEB-BPΔ15, or pEB-BPΔ20 were grown at 30 ^{o}C as described in section 2.7.1. After 76 hours of growth, the activity was assayed (section 2.7.4) and normalised for cell growth by dividing the activity by the cell density. The average normalised activity of duplicate cultures listed are relative to that of pEB-B8, and note the exponential decrease in activity.



Construct	Normalised Activity (ΔA_{412} min^{-1} OD_{600}^{-1})	% Wild Type Activity
pEB-B8	0.1842	100.0
pEB-BPΔ4	0.0727	39.5
pEB-BPΔ10	0.0158	8.58
pEB-BPΔ15	0.0021	1.1
pEB-BPΔ20	0.0003	0.2
pEB-11	0.0000	0.0

Figure 20. Log of Normalised Deletion Mutant Activity as a Function of Amino Acid Deletion. The normalised activity data shown in Figure 3 for the pEB-B8, pEB-BPΔ4, pEB-BPΔ10, pEB-BPΔ15, or pEB-BPΔ20 deletion mutants was converted into log values and plotted against amino acid deletion from the propeptide.

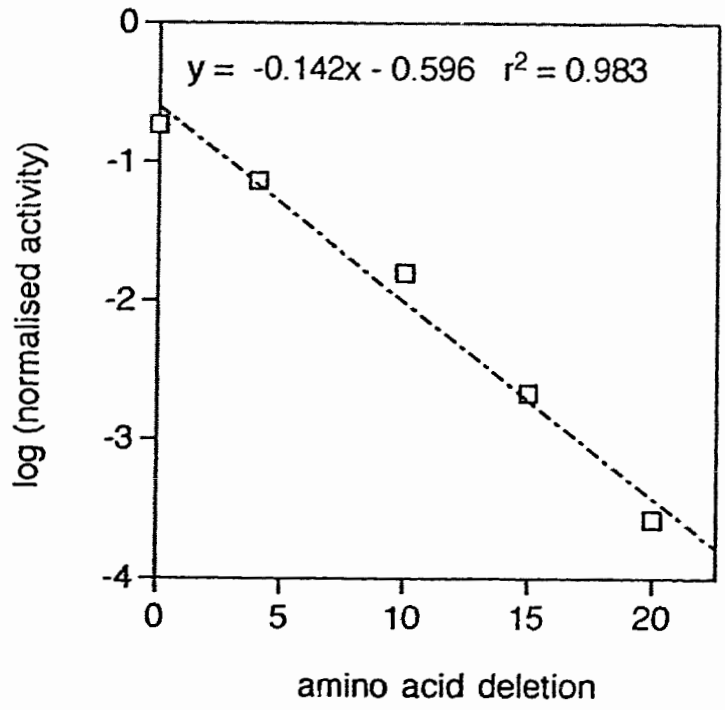

y = -0.142x - 0.596 r2 = 0.983
log (normalised activity)
0
-1
-2
-3
-4
0
5
10
15
20
amino acid deletion

Figure 21. Effect of Temperature on Deletion Mutant SGPB Activity. *B. subtilis* cultures harbouring pEB-BPΔ10, pEB-BPΔ15, pEB-BPΔ20, or pEB-11 were grown at 21, 30 and 37 °C in addition to 30 °C as described in section 2.71. After 59 hours, activity in each culture was assayed, normalised for cell growth, and the log of the normalised activity was plotted as a function of residues deleted from the amino terminus of the propeptide.

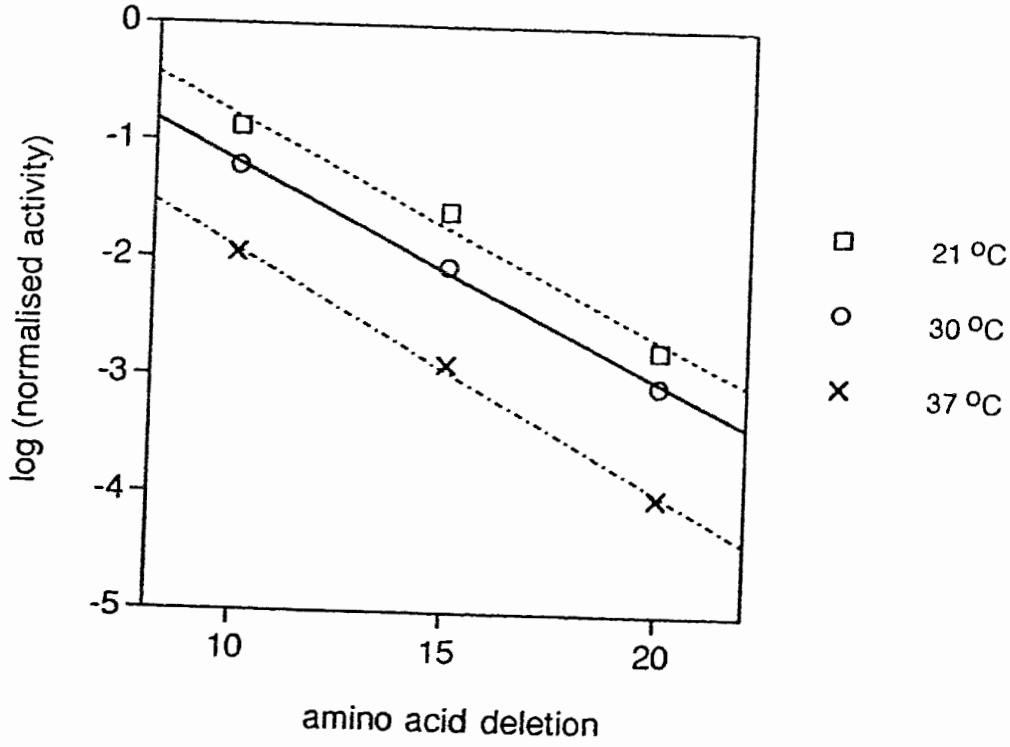

0
-1
-2
-3
-4
-5
log (normalised activity)
10
15
20
amino acid deletion
21 °C
30 °C
37 °C

Table 3. Protease Expression Generated by the *sprC, sprD* and *sprE* Deletion Mutants. Duplicate 50 mL cultures of *B. subtilis* harbouring the deletion mutants or corresponding wild type protease gene were grown at 30 ^{o}C and assayed for protease activity as described in sections 2.7.1 and 2.7.4. The maximum normalised activity each of the deletion mutants obtained relative to it's wild type gene are listed along with the conversion to percent wild type activity for each mutant.

Construct	Normalised Activity (ΔA_{412} min^{-1} OD_{600}^{-1})	% Wild Type Activity
pEB-C8	0.1829	100.0
pEB-CPΔ11	0.005445	2.98
pEB-CPΔ16	8.498×10^{-5}	0.046
pEB-D8	0.1666	100.0
pEB-DPΔ20	0.002484	1.49
pEB-E	0.007365	100.0
pEB-EΔ26	1.894×10^{-4}	0.16

3.3.3 Northern Blot Analysis of *B. subtilis* Cultures Harbouring Deletion Mutants

Transcription of deletion mutant genes in *B. subtilis* was analysed by Northern blot analysis (Figure 22). The total RNA was electrophoresed and probed with the mature encoding region of *sprB*. Significant hybridization occurred with all of the deletion mutant genes (lanes 2-6) at approximately the same levels. This indicates that the observed decrease in activity with increasing size of deletion was not a transcriptional artifact. No hybridization could be seen with the expression vector pEB-11, the negative control.

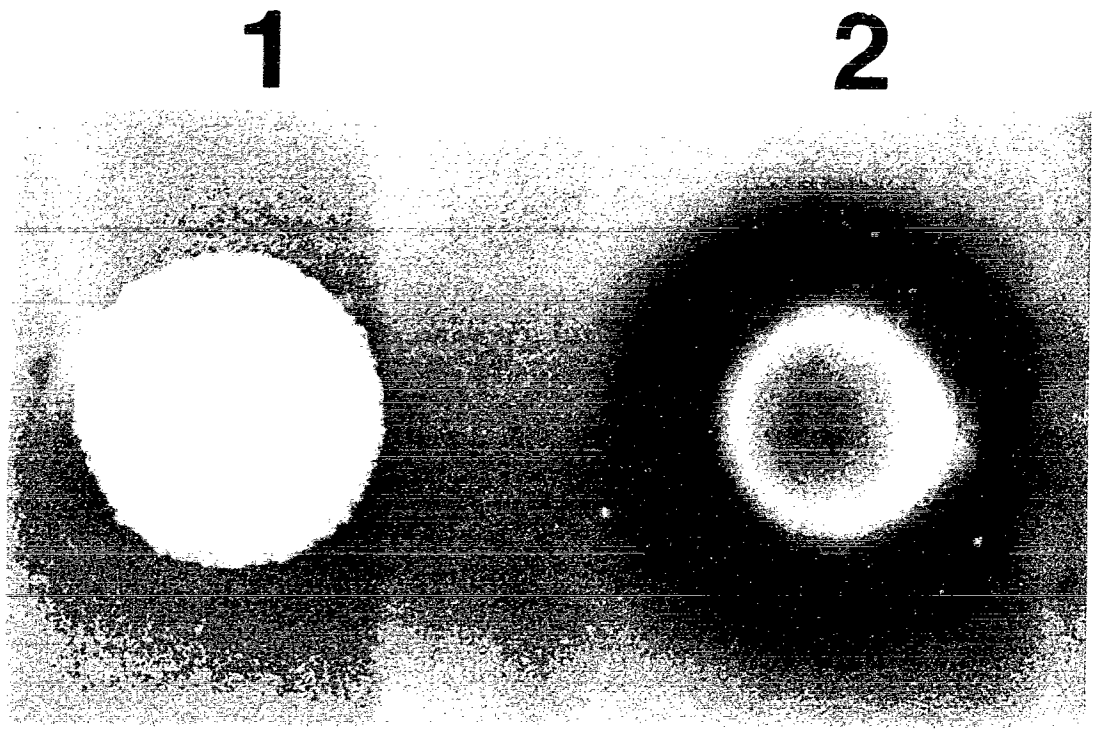
3.3.4 Expression of pMP-BPΔ10 in *E. coli*

In order to compare the SGPB propeptide deletion mutant to previous *in vivo* studies of α-lytic protease and subtilisin propeptide deletion mutants in *E. coli* expression systems (17, 28), a 10 amino acid propeptide deletion was produced in *E. coli* using the pMAL-P secretion vector. Colonies harbouring the plasmid pMP-BPΔ10 secreted active protease that could be visualized as a zone of clearing on TY/Milk/Amp plates containing 80 μM IPTG to induce transcription (Figure 23).

Figure 22. Northern Blot Analysis of *sprB* Genes Encoding Propeptide Deletion Mutants. After 36 hours of growth at 30 °C, total RNA was isolated from *B. subtilis* transformants harbouring **1)** pEB-19, **2)** pEB-B8, **3)** pEB-BPΔ4, **4)** pEB-BPΔ10, **5)** pEB-BPΔ15, **6)** pEB-BPΔ20 and hybridized with a radiolabeled segment of mature *sprB* as described in section 2.8. Significant hybridization could be seen in all lanes except lane 1 (the negative control) indicating that the propeptide deletion mutants are being transcribed at approximately the same levels as wild type *sprB*.

1 2 3 4 5 6

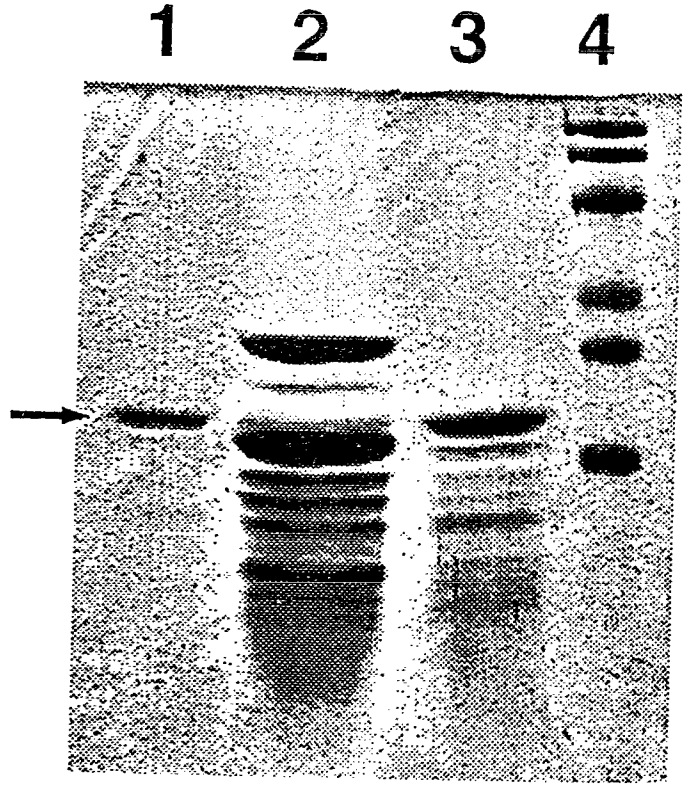
Figure 23. Recombinant Expression of 10 Amino Acid Deletion Mutant in *E. coli* TB1. *E. coli* TB1 transformants harbouring 1) pMAL-p or 2) pMP-BPΔ10 were plated on YT/Milk/Amp plates containing 80 μM IPTG and photographed after overnight incubation at 37 °C followed by 72 hours of incubation at room temperature. A zone of clearing indicating the secretion of active protease is visible with the pMP-BPΔ10 transformants.


1
2

3.4 AMINO TERMINAL ANALYSIS OF pEB-BPΔ10 AND pEB-pAmB

In order to ensure the mutant genes were being processed with wild type specificity, amino terminal analysis of protease generated by the 10 amino acid deletion mutant gene and chimeric gene was carried out. The 10 amino acid deletion was chosen because it was the longest deletion that had readily detectable protease activity in culture supernatant. The protease generated by pEB-BPΔ10 and pEB-pAmB was partially purified as described (section 2.7.2) and further purified by FPLC (section 2.7.3). Active FPLC fractions were concentrated and electrophoresed on 12 % SDS-PAGE. Protein bands that co-migrated with wild type SGPB were generated by the pEB-BPΔ10 and pEB-pAmB cultures (Figure 24). Amino terminal sequencing of these bands revealed that they both had mature SGPB sequence at the amino terminus (Ile-Ser-Gly-Gly-Asp) indicating that both of the mutant genes generated correctly processed mature SGPB.

Figure 24. SDS-PAGE of SGPB Generated by pEB-B8, pEB-pAmB, and pEB-BPΔ10. SGPB generated by pEB-pAmB and pEB-BPΔ10 was purified and concentrated for amino terminal analysis as described in section 2.7.6. Aliquoits of FPLC purified SGPB produced by **1)** pEB-B8, **2)** pEB-pAmB, **3)** pEB-BPΔ10 were run out on 12 % SDS-PAGE gels along with **4)** low molecular weight standard (BioRad) and stained with coomassie blue. The arrow indicates the position of mature SGPB.


1
2
3
4

4. DISCUSSION

Several examples of propeptide catalyzed protein folding have now been reported. Moreover, the examples within the serine proteases come from evolutionarily convergent groups of enzymes. The proteases subtilisin E from *B. subtilis* (8), subtilisin BPN' from *B. amylofaciens*, (9) and α-lytic protease from *Lysobacter enzymogenes* (7) represent two enzyme families which are evolutionarily unrelated but have converged on a common mechanism involving the Asp-His-Ser catalytic triad (40) Remarkably, the nascent proteases in each case possess propeptides implicated in the folding and maturation of the enzymes; thus the different enzymes may have also converged on a similar propeptide-dependent folding pathway. Some authors have suggested that propeptides are the intramolecular analogs of chaperonins, catalyzing the rate-determining step in protein folding. However, propeptides are not able to function with same broad specificity of some chaperonins, rather they appear to have become specifically adapted to one enzyme through the course of evolution (5, 18). By analysing the sequences of propeptides of closely related proteases, we hoped to discover important conserved regions that may be involved in the folding process.

4.1 CHIMERIC GENE CONSTRUCT

The propeptides of subtilisin E (31) and α-lytic protease (24) are reported to be potent inhibitors of the activity of the mature enzymes. This enzyme inhibition may be a major biological role for pro regions; ensuring that the enzyme is inactive until it has been secreted from the cell. Alternatively, enzyme inhibition may be an artifact of the primary function of

the propeptide; to reduce the activation energy between the folded and unfolded states of the protein (23) In either case, propeptides appear to possess both folding and inhibitory functions. Notably, it has been reported that the propeptide of the α-lytic protease is an effective inhibitor of its distantly related cousin SGPB (24). The implication of this observation is that the propeptide of one protease is able to catalyze the folding of another related enzyme; the question is how well? We addressed this question through the creation of a chimera between the two closely related enzymes SGPA and SGPB.

Extensive propeptides are present in the six related proteases SGPA, SGPB (10), SGPC (12), SGPD (11), and SGPE (13) of *S. griseus* and the α-lytic protease of *L. enzymogenes* (7) Although the mature enzymes are similar in size, sequence and three dimensional structure (where known), the propeptides are extremely diverse (11). For example, the propeptides of SGPC and the α-lytic protease are 162 and 166 residues in length whereas the propeptides of SGPA and SGPB are 78 and 76 amino acids (Table 1). A gene construct was prepared in which the propeptide-encoding region of *sprA* was fused in frame to the mature encoding region of *sprB* using PCR cloning techniques. The *Bacillus* expression vector with the chimeric gene was designated pEB-pAmB and it directs the synthesis of a chimeric polypeptide with the propeptide of SGPA covalently attached to the mature region of SGPB (Figure 12). This particular combination of enzymes was chosen because they are the two most closely related in terms of the sequences (43 % amino acid homology) and dimensions of their propeptides (76 and 78 amino acids for SGPB and SGPA respectively) (Figure 25). The mature

Figure 25. Amino Acid Alignment Between the Propeptides of SGPA (proA) and SGPB (proB). The complete amino acid sequence of the propeptides of SGPA (residues 1 - 78) and SGPB (residues 1 - 76). Conserved amino acids are **bold**. There is a 43 % amino acid homology between the two domains.

proA APEAESKATVSQLADASSAILAADVAGTAWYTEASTGKIVLTADSTVSKAELAKVSNALAGSKAKLTVKRAEGKFTPL
proB ETPRTFSANQLTAASDAVLGADIAGTAWNIDPQSKRLVVTVDSTVSKAEINQIKKSAGANADALRIERTPGKFTKL

regions of SGPA and SGPB enzymes are 181 and 185 residues long respectively and are 61% identical in sequence. Furthermore, the α-carbon positions of the SGPB and SGPA crystal structures are 85% topologically equivalent (57). Therefore, it would be expected that the folding pathway of SGPA and SGPB would be quite similar, and the propeptides from both proteases would have cognate functions.

When protease expression was measured, the chimeric gene was found to generate 6.0 % of wild type SGPB activity (normalised for cell density). Therefore, despite evolutionary divergence, the propeptides retain a certain ability to fold related enzymes; at least this was true of the particular combination tested. This is the first example of the propeptide from one bacterial protease gene correctly folding the mature region of another protease gene *in vivo*. It was evident from Northern blotting experiments that the chimera gene was as efficiently transcribed as the gene encoding wild type SGPB (Figure 15), therefore the reduced activity observed with the chimera was not due to a transcriptional artifact.

Zhu *et al* examined the ability of the propeptide of subtilisin E to refold denatured subtilisin BPN' and subtilisin Carlsberg *in vitro* (29). Although there is over 80 % amino acid homology between the propeptides, the *trans* addition of subtilisin E propeptide could only recover a maximum of 5.5 % and 2.3 % specific activity from denatured subtilisin BPN' and subtilisin Carlsberg respectively. Given the high homology between both the propeptides and mature region of subtilisin E and subtilisin BPN' (88.5 %) compared to the propeptides of SGPA and SGPB (43 % homology), it is interesting that both experiments resulted in a similar recovery of wild type activity. However, one cannot overlook the fact that the trans addition of propeptide will have an inherently reduced "foldase" activity due to the

propeptide-mature folding intermediate complex having to be generated from a bimolecular collision, whereas the chimera will have the propeptide covalently linked to the mature region *in vivo* which will increase the effective propeptide concentration. A covalently attached propeptide of 100 amino acids in extended conformation will have a local concentration of about 10 μM relative to the mature domain (34). Typically, exogenous propeptide is added in concentrations around 5 to 10 μM in refolding experiments (25, 32, 34). These results suggest that the success of a particular propeptide in folding other enzymes may be a criterion for determining the degree to which the two enzymes are evolutionarily related.

Since the propeptide of SGPA can direct the folding of SGPB, we decided to compare the propeptides for predicted structural similarities. Secondary structure predictions using the Chou, Fasman, and Rose method (58, 59) suggested that the propeptides of SGPA and SGPB are predominantly helical (Figure 26). This is also a characteristic of the propeptide of carboxypeptidase Y (60) and other heat shock chaperonin proteins (61) Helical wheel analysis of the α-helix in SGPB (residues 5 - 22) and the corresponding α-helix in SGPA (residues 7 - 24) revealed a conserved face (Figure 27). The fact that the α-helices are conserved on one face suggest that this face has an important function in the folding process of the proteases and selective pressures have retained the conserved sequence through evolution. The non-conserved face may have diverged because it has no important function in the folding process. In order to explore the importance of the amino terminal α-helix of the propeptide, it was targeted for mutational analysis in SGPB.

Figure 26. Secondary Structure Predictions of the Propeptides of SGPA (proA) and SGPB (proB). The secondary structure prediction was generated by the Chou, Faseman, and Rose method with the MacDNASIS Pro computer program (Hitachi). 'H', 'S', 'T', and 'C' indicates α-helix, β-sheet, turn and random coil respectively. Capitalized letters indicate the favoured conformation for each amino acid. Helical wheel analysis of the boxed region is shown in Figure 27.

```
                 10        20        30        40        50        60        70
proA     APEAESKATVSQLADASSAILAADVAGTAWYTEASTGKIVLTADSTVSKAELAKVSNALAGSKAKLTVKRAEGKFTPL
HELIX    HHHHHHHHhhhhhHHHHHHHHHHHHHHHHHHHHHHhhhhhhHHHHHHHHHHHHHHHHHHHHHHHHhhhHHHH
SHEET             SSSSs        sssssssssss   SSSSSSs                       SSs      SSSS
TURN       t   t           t                  t       t           t    t          t   t
COIL                                                                              CC

                   10        20        30        40        50        60        70
proB       ETPRTFSANQLTAASDAVLGADIAGTAWNIDPQSKRLVVTVDSTVSKAEINQIKKSAGANADALRIERTPGKFTKL
HELIX           HHHHHHHHHHHHHHHHH              hhhhhhhhhhhhhhhhHHHHHHHHHHHHHHH
SHEET       SSSSsssssss       ssSSSSSSSSS      SSSSSSSSSSSSSSSSs                   SSSS
TURN          t           t    t          TTTT       t           t            t  t   t
COIL       C                             C    C                              CCCCCC
```

Figure 27. Helical Wheel Representations of A) Amino Acids 7 to 24 of the Propeptide of SGPA and B) the Homologous Amino Acids 5 to 22 of the Propeptide of SGPB. The helical wheel was generated using a computer program (MacDNASIS Pro, Hitachi), with 3.6 amino acids per turn. Chemically identical amino acids are dark gray, while chemically similar amino acids (one methyl group difference) are light gray. Notably, the conserved amino acids all occupy one face in each α-helix.

A
Ala 8
Asp 15
Ala 22
Ser 11
Ser 18
Lys 7
Ala 14
Leu 21
Val 10
Ser 17
Asp 24
Leu 13
Ile 20
Thr 9
Ala 16
Ala 23
Gln 12
Ala 19
B
Phe 6
Ala 13
Gly 20
Asn 9
Asp 16
Thr 5
Thr 12
Leu 19
Ala 8
Ser 15
Asp 22
Leu 11
Val 18
Ser 7
Ala 14
Ala 21
Gln 10
Ala 17

4.2 DELETION MUTANT CONSTRUCTS

Considering the vast differences in sizes and sequences of propeptides in bacterial proteases, it is reasonable to question the necessity for the entire propeptide in folding. The *S. griseus* proteases are a good example of how propeptides differ in length. The propeptides range in size from 76 residues for SGPB to 162 residues for SGPC, over two times the length. With deletion analysis of the propeptide of SGPB, we hoped to resolve whether the folding function of the propeptide involves; 1) interactions over the entire length of the amino terminal α-helix (gradual reduction in activity), 2) a few key residues involved in the folding activity (stepwise reduction in activity), 3) or none of the residues in the amino terminal α-helix are important for folding (no reduction in activity with the removal of the α-helix). In a previous study Fujishige *et al* (28) reported that the deletion of as few as 5 residues from the amino terminal end of the propeptide of α-lytic protease completely abolished the secretion of active enzyme. Similarly, a 14 amino acid deletion in the propeptide of subtilisin fails to generate active protease *in vivo* (17), and a series of deletions in carboxypeptidase Y resulted in a reduction of intracellular enzyme activity (62). In the present study we created a set of nested deletions in the amino terminal α-helix of the propeptide of the enzyme SGPB.

Propeptide deletions were created using either restriction enzymes or PCR and the mutant genes were expressed in *B. subtilis* DB104. From the histogram shown in Figure 19, it can be seen that the amount of enzyme secreted diminishes incrementally with each deletion. Remarkably, when the data was normalized for cell growth and replotted, we were able to demonstrate that SGPB activity diminished exponentially according to the

length of the deletion (Figure 20). This indicates that propeptide residues involved in the folding process are spread homogeneously throughout the entire propeptide, and not localised to just a few key residues.

We made the assumption in studies of the variant proteases that the activity measured in culture supernatants corresponds to only the concentration of native protease, as opposed to incompletely processed or partially folded forms of the enzyme. SDS-PAGE of proteins from the partially purified 10 amino acid deletion mutant culture revealed a protein band that co-migrated with wild type SGPB (Figure 24). Amino terminal analysis of the protein in this band verified that a correctly processed protease was secreted into the culture supernatant (section 3.4). In any case, abherently folded forms of the enzyme are unlikely to persist in culture supernatants as they would be susceptible to degradation by the mature native enzyme. Intermediate folding state of α-lytic protease or subtilisin has never been detected *in vivo*, it is only stable *in vitro* (25, 32, 33, 62).

Once outside the cell, the propeptide is cleaved from the enzyme and degraded, thereby trapping the enzyme in a stable, native conformation (23). The half life of *S. griseus* proteases at 30 ^{o}C in culture supernatants is on the order of several days (unpublished results). Consequently, differences in the rate of degradation of SGPB do not explain the differences in activity in the culture supernatant.

It is important to emphasize that in every case the mature enzyme is the same, and the gene promoter and signal sequences are unchanged. Northern blot analyses of *B. subtilis* harbouring the mutant genes were conducted simultaneously with activity assays. The Northern blot shown in Figure 22 indicates that the variant genes are all transcribed to the same degree as the wild type *sprB*. Also, since the deletions occur well inside the

open reading frame, it is unlikely that secondary structures will be created in the RNA that will interfere with translation, though this possibility cannot be completely discounted. Because the deletions change the residue context around the junction of the pre and pro domains they may possibly effect recognition by the leader peptidase. However, the primary specificity for signal peptidase lies in the residues amino terminal to the pre-pro junction, especially those at positions -1 and -3 (15)

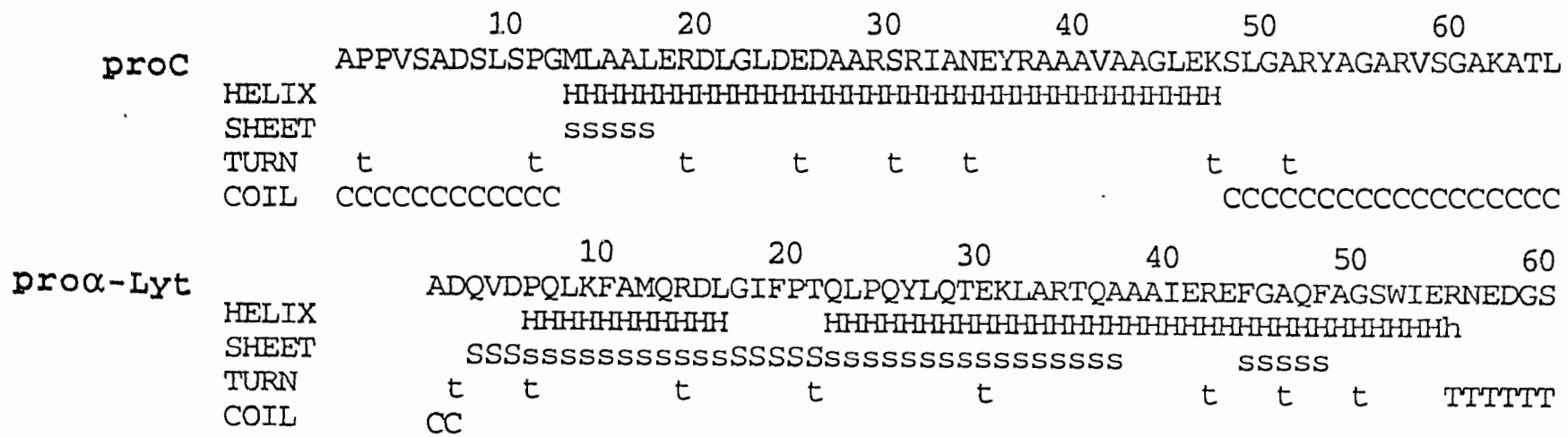
From the plot of the data in Figure 20, it can be seen that on average, each residue removed from the amino terminal end of the propeptide resulted in 0.142 log loss in the secreted activity. If we are correct in assuming that the loss in activity is directly attributable to a reduction in binding energy at the transition state for folding, the Arrhenius equation ($\Delta\Delta G^{\ddagger} = -2.303\ RT \log K$) can be used to calculate the binding energy each amino acid contributes to stabilising the folding transition state since the differences in reaction rate can be directly related to differences in reaction activation energy. On average each amino acid within the deleted region contributes 200 cal mol^{-1}, therefore, every 5 amino acids contributes 1.0 kcal mol^{-1} to the ΔG for folding at 30 ^{o}C. This result assumes that the decrease in activity is due to the reduced ability of the propeptide mutants to correctly fold protease independent of the secretion process as discussed in sections 4.3 and 4.4. For every ten-fold difference in reaction rate, there is a difference of 1.39 kcal mol^{-1} at 30 ^{o}C. The 20 amino acid deletion will have a transition state energy barrier that is 4 kcal mol^{-1} higher than wild type which corresponds to a 750 fold reduction in folding activity at 30 ^{o}C. By comparison, single hydrogen bonds in proteins are known to contribute 1 to 3 kcal mol^{-1} to binding (63).

Notably, the relationship between activity of the enzyme and the propeptide deletion exists at each of the temperatures 21, 30 and 37 ^{o}C as

evidenced by the parallel lines of activity (Figure 21). This suggests that the phenomenon is not merely an artifact of the growth conditions of *B. subtilis*. The detection of active enzyme secreted in *E. coli* (Figure 23) demonstrates that the ability of SGPB to fold correctly is not limited by expression in gram positive organisms such as *S. griseus* and *B. subtilis*.

We wanted to demonstrate that the reduction of activity observed with the propeptide deletion mutants was not unique to SGPB rather that it is a general feature of enzymes with pro regions. Therefore, deletions similar to those made in SGPB were made in the propeptides of SGPC, SGPD and SGPE. The fact that the mutants of SGPC, SGPD, and SGPE all have detectable activities (Table 3) indicates that the correlation between length of deletion and activity is not a unique case, but a general phenomena of this family of enzymes. The result seen with the two SGPC deletions is somewhat surprising. SGPC has the longest propeptide but is the least able to accommodate deletions; the 16 amino acid deletion mutant generates extremely low levels of mature protease (Table 3). The propeptide of α-lytic protease has 26 % amino acid homology with the propeptide of SGPC, and behaves in a similar manner to the SGPC propeptide in that it cannot accommodate long deletions. Secondary structure prediction of the propeptides of SGPC and α-lytic protease predict an amino terminal α-helicies for both proteins (Figure 28). The propeptides share 33 % amino acid homology from the eleventh amino acid to the twenty-eighth amino acid of SGPC and the corresponding region in α-lytic protease, although they do not exhibit a conserved face like the propeptides of SGPA and SGPB. The

Figure 28. Secondary Structure Predictions of the Amino Terminal End of the Propeptides of SGPC (proC) and α–Lytic Protease (proα-Lyt). The secondary structure prediction of amino acids 1 to 65 of the propeptide of SGPC and the homologous amino acids 1 to 60 of the α-lytic protease propeptide. The secondary structure prediction was generated by the Chou, Faseman, and Rose method with the MacDNASIS Pro computer program (Hitachi). 'H', 'S', 'T', and 'C' indicates α-helix, β-sheet, turn and random coil respectively. Capitalised letters indicate the favoured conformation for each amino acid.


proC
10 20 30 40 50 60
APPVSADSLSPGMLAALERDLGLDEDAARSRIANEYRAAAVAAGLEKSLGARYAGARVSGAKATL
HELIX HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
SHEET sssss
TURN t t t t t t t t
COIL CCCCCCCCCCCC CCCCCCCCCCCCCCCCCCC
proα-Lyt
10 20 30 40 50 60
ADQVDPQLKFAMQRDLGIFPTQLPQYLQTEKLARTQAAAIEREFGAQFAGSWIERNEDGS
HELIX HHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHh
SHEET SSSsssssssssssSSSSsssssssssssssss sssss
TURN t t t t t t t t TTTTTT
COIL CC

10 amino acid deletion to the propeptide of SGPC which readily generates protease activity does not fall in this region; the 16 amino acid deletion which does fall within the predicted α-helix loses almost all of it's protease activity (Table 3). Similarly, deletions greater than 5 amino acids which do not generate any active protease in α-lytic protease fall within the predicted amino terminal α-helix. This data suggests that the residues before the first α-helix are not as important to the folding process as those within the α-helix.

4.3 PROPEPTIDES AND SECRETION

In many cases, the propeptide appears to be involved in secretion in addition to folding. In the gram negative *E. coli*, propeptide mutants of α-lytic protease which do not produce active protease have been found tightly associated with the outer membrane (28). The interaction between the outer membrane and propeptide mutants is nonspecific since mutants lacking the entire propeptide are also found tightly associated with the outer membrane. Only correctly folded protease will be transported across the outer membrane. Mature protease is folded and processed in the periplasm as propeptide is found to accumulate in this region. Therefore, it seems that a tertiary signal is required for the successful transport of α-lytic across the outer membrane, in contrast to translocation across the inner membrane which is directed by the signal sequence. Translocation across the inner membrane, unlike transport across the outer membrane, is characterised by the lack of tertiary structure and appears to be propeptide independent (28). In another experiment, the propeptide and mature regions are expressed *in trans* in *E. coli* (23). Each component is secreted into the periplasm where the

propeptide directs the folding of the mature region. Again, secretion of the mature protease into the periplasm is propeptide independent.

The gram negative bacterium *Neisseria* gonorrhoeae produces IgA protease which crosses the inner membrane with a standard bacterial signal sequence and traverses the outer membrane by a mechanism different from the transport of α-lytic protease across the outer membrane in *E. coli*. The propeptide of this protease is predicted to interact with the outer cell membrane and form a pore through which the mature region passes and is released into the extracellular medium by autoproteolysis (64).

In eukaryotic systems, the propeptide directs proteins such as carboxypeptidase Y in *Sacchromyces cerevisiae* (65) and prohormones in mouse AtT20 cells (66) from the endoplasmic reticulum lumen through the golgi apparatus to the correct secretory vesicles for secretion.

In the above examples, the signal peptide directs the nascent polypeptide across the first membrane and the signal sequence is then released by cleavage by signal peptidase. With only one membrane to traverse in the gram positive bacterium *B. subtilis*, it is unlikely that the propeptide will have any secretion function. Once the signal sequence directs the propeptide across the cell membrane into the extracellular matrix (44), the target location of the pro-mature protein has been reached. For subtilisin E expression in *E. coli*, constructs were created that produce the OmpA signal peptide fused to 3 different locations within the subtilisin E signal sequence along with a construct that produces the OmpA signal sequence fused directly to the propeptide (19). No difference in activity could be found between all of the fusion proteins including the direct fusion of OmpA signal sequence to the mature domain (19). This suggests the OmpA signal sequence can be interchanged with the subtilisin E signal sequence without a

loss of function. A direct fusion of OmpA signal sequence to mature subtilisin E resulted in the mature protein being secreted, but the protein did not generate any activity. The OmpA signal sequence was correctly processed from the direct fusion to the mature domain as shown by amino terminal sequencing. This suggests the propeptide is required for folding, but not required for the secretion process in *E. coli*. We have shown that *E. coli* harbouring the plasmid pMP-BPΔ10 encoding a 10 amino acid propeptide deletion can secrete active protease *in vivo* (Figure 23). The link between folding and secretion has not been separated in the gram positive bacteria *B. subtilis* and *S. griseus*. However, the replacement of the subtilisin E signal sequence by the OmpA signal sequence suggest that the secretion process in *B. subtilis* behaves similarly to the secretion process in *E. coli*. Also, the pre-pro-mature extracellular basic protease from the gram negative bacterium *Dichelobacter nodosus* was shown to be correctly secreted and processed in *B. subtilis*, demontrating the differences between the secretion in gram negative and gram positive bacteria may not be that large (67). Ideally, all of the effects on activity in *B. subtilis* will be due to the decreased ability of the mutants to fold, and not due to mature enzyme that cannot be correctly secreted into the extracellular environment.

4.4 INTERPRETATION OF *IN VIVO* RESULTS

The chimera and deletion mutant results have been interpreted assuming that transcription, translation and secretion of the protease all function independently of the propeptide.

As stated in section 4.2, the transcription and translation should be equivalent for both the wild type and mutant gene constructs. All of the

constructs use the same promoter and Northern blot analysis demonstrated that all of the genes were transcribed to approximately the same level. All of the constructs use the same ribosome binding site and deletions are well within the open reading frame of the gene, therefore, it is unlikely that any secondary structure would be produced to interfere with translation.

Secretion is more difficult to separate from the folding function of the propeptide. As discussed in section 4.3, the propeptide does not appear to be required for the secretion into the periplasm of gram negative bacteria like *E. coli*. The propeptide is not likely required for secretion in *B. subtilis* assuming secretion into the extracellular environment is similar to secretion into the periplasm of gram negative bacteria like *E. coli*. Also, signal peptidase has specificity for residues at -1 and -3 to the propeptide (15). Therefore, differential hydrolysis of the propeptide mutants by signal peptidase is unlikely.

In order to verify that secretion is not affected by the *S. griseus* protease propeptide mutants in *B. subtilis*, further experiments are required. Western blot analysis should be carried out on both intracellular protein and secreted protease in the supernatant. One of several different results could be obtained: 1) Ideally a decreasing amount of secreted wild type protease with increasing propeptide deletion will be observed along with constant levels of intracellular protein for both wild type and mutant constructs. This will indicate that the propeptide mutants are not changing the rate of secretion from the cell, and once secreted from the cell only the pro-mature protein will be involved in the generation of active protease. This is the assumption made for calculating the increase in the folding transition state energy barrier with increasing size of deletion from the propeptide (section 4.2). 2) There is an increase in intracellular protein and a decrease in extracellular protease with

increasing deletion. This would indicate that the propeptide is required for both secretion and the folding process. 3) Both the intracellular protease and extracellular levels of protease are the same. This would indicate the propeptide folds the deletion mutants to variant forms of the protease with reduced catalytic ability. It would not be possible to tell if the propeptide is linked to secretion in this case with a Western blot. This case is not likely since SDS-PAGE of cell culture supernatant stained with Coomassie Stain shows a decrease in secreted protease between the wild type SGPB and 4 amino acid propeptide deletion mutant (data not shown). 4) There may be a case where the intracellular levels of protein accumulate to very high levels due to a secretion bottle-neck where again it would be difficult to ascertain whether or not the propeptide is involved in the secretion process. A reduction in extracellular protease could be due to the reduced ability of the mutants to be secreted properly in addition to reduced folding ability of the mutants.

The best experiment for determining if secretion is propeptide independent in *B. subtilis* would be to create a construct that expresses the propeptide and mature regions *in trans* as done by Silen for α-lytic protease in *E. coli* (23). Since each component would have it's own signal sequence, active protease in the *B. subtilis* culture supernatant would indicate secretion is independent of folding.

Again, although the effect of the propeptide on secretion cannot be discounted in *B. subtilis* without further experimental evidence, the results in *E. coli* demonstrating that pre-mature protease can be secreted and correctly processed in an inactive state (17, 28, 67), or folded in the presence of propeptide (23) suggest that the propeptide has little or no influence on secretion into the periplasm. Until further experiments are performed in *B.*

subtilis, the parallels drawn between secretion into the periplasm of *E. coli* and extracellular environment in *B. subtilis* will have to be used as a base for the calculation of the folding transition state free energy (section 4.2).

4.5 HYPOTHETICAL FOLDING MECHANISM

As already indicated, SGPA and SGPB have a high degree of structural homology. We have shown that the propeptides of the two enzymes can be interchanged without sacrificing all of their function. Therefore, it follows that the folding pathways of both proteases would be very close. It is difficult to speculate on the mechanism of propeptide mediated folding since no crystal structure of a chymotrypsin like pro-mature enzyme is present, only the pro-mature crystal structure of the structurally unrelated enzyme subtilisin exists. If the propeptide of SGPB behaves like the propeptides of α-lytic protease and subtilisin, then it will interact with a folding transition state late in the folding pathway (23, 33). The stability of the *S. griseus* proteases (unpublished results), like α-lytic protease and subtilisin, demonstrates that there is a high energy barrier that the mature protease must overcome to unfold to an inactive state (23). This energy barrier cannot be easily traversed once the propeptide has been removed by mature protease and subsequently degraded.

The mature regions of SGPA and SGPB seen in the crystal structure analysis of the mature proteases (57)and predicted by the Chou, Fasman and Rose method are predominantly β-sheet. Most of the deletions made to the propeptide of SGPB fall within the first putative amino terminal α-helix (residues 5 to 22) (Figure 26). We predict that the propeptide is aligned so that the conserved face of the amino terminal helix is able to interact with the

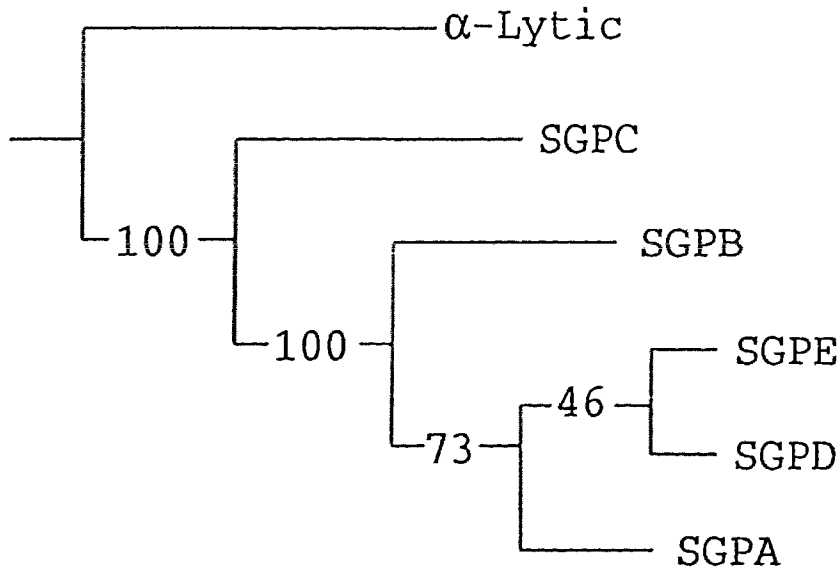
mature domain during folding. In essence, the propeptide acts as a scaffold to stabilize through many small non-covalent interactions the formation of β-sheet and then tertiary structure. We predict that the propeptide of SGPB reduces the activation energy barrier to the native protease much the way the propeptide of subtilisin BPN' stabilises the αβα substructure in the mature domain to allow the completion of folding (34). As shown by the deletions to the pro domain of SGPB, the activation energy barrier would increase by 1.0 kcal mol^{-1} for each 5 amino acids removed. Notably, it has been calculated that the average binding interaction for a single hydrogen bond between uncharged groups will vary in energy between 0.5 to 1.5 kcal mol^{-1}(63), therefore, every 5 amino acids in the analyzed region of the SGPB propeptide contributes the approximately the same amount of energy as one hydrogen bond. This is consistent with an interaction involving three of four hydrogen bonds between the amino terminal α-helix and mature region.

4.6 PROPEPTIDES AS A SCALE OF EVOLUTION

It is interesting to note that the ability of a propeptide to accommodate deletions reflects its position on the phylogenetic tree (11) (Figure 29). The most ancient protease, SGPC, barely has detectable activity with a 16 amino acid deletion. SGPB, which is evolutionarily intermediate to SGPC and the most recent proteases SGPD and SGPE, also is intermediate in activity when comparing the longest deletions (Table 3). SGPD and SGPE appear to retain the most activity with the longest deletions. The propeptide of SGPC is able to accept longer deletions than the α-lytic protease propeptide and still correctly fold mature protease. This correctly places SGPC after α-lytic protease on the phylogenetic tree. This data suggests that as the proteases

evolve, more of the folding function is removed from the propeptide and placed in the mature region. The ultimate goal may be to reduce the size of the propeptide once the folding function has been removed from the propeptide. *S. griseus* trypsin which is evolutionarily related to mammalian trypsin has a shortened propeptide of only 4 amino acids, much like the 6 amino acid propeptide of mammalian trypsin. These propeptides have no folding function, they only serve to keep the mature domain inactive until they are secreted from the cell (68).

Figure 29. Phylogenetic Tree of the Bacterial Proteases SGPA, B, C, D, and E and α-Lytic Protease. The tree was constructed using the nucleotide sequences of the mature proteases. The number at each fork represents the number of times that a particular grouping (consisting of the species to the right of the fork) was generated during the 100 bootstrap replicates. The pre and propeptide encoding sequences and the sequence of the carboxyl-terminal domain were not included in the analysis. This figure is adapted from Sidhu *et al* (11).


α-Lytic
SGPC
100
SGPB
100
SGPE
46
73
SGPD
SGPA

5. CONCLUSION

We have been able to show that propeptide mediated folding of mature SGPB is not specific for only the SGPB propeptide. We constructed a chimeric gene consisting of the propeptide encoding region of *sprA* and mature encoding region of *sprB*. designated pEB-pAmB. The chimeric gene produced correctly processed mature SGPB with detectable activity when expressed in *B. subtilis*. This suggests that propeptides may have evolved from a non-specific propeptide precursor that was able to catalyse the folding of many proteins much like the chaperonin class of proteins found today (69). Sequence analysis of the propeptides of SGPA and SGPB revealed an amino terminal α-helix which has a conserved face. Deletions from the amino terminal end of this α-helix in SGPB results in a logarithmic decrease in activity with increasing deletions. From the activity data, it was calculated that for every 5 amino acids within the deleted region, the propeptide contributes 1.0 kcal mol^{-1} to the stabilization of the folding transition state, assuming that any incompletely folded or miss-folded protease is quickly degraded *in vivo*. Northern analysis of the transcripts indicated the mutant protease genes are being transcribed normally and amino terminal analysis of protease produced by the mutant genes found that the propeptide was processed from the chimeras with wild type specificity. Therefore, the observed decrease in activity could be directly attributed to the reduction of propeptide mediated folding. The conserved α-helix and activity data indicates that the conserved face makes multiple interactions with the folding transition state that are spread evenly through the deleted region; the removal of amino acids from the propeptide reduces it's ability to stabilize and reduce the energy barrier to correctly folded mature protease.

5.1 FUTURE DIRECTIONS

Future work will analyse internal deletions to see if activity is lost at the same rate as the amino terminal deletions. This will also help to define whether or not the interactions within the propeptide are specific or non-specific in nature. Any internal deletions will shift the residues amino terminal to the deletion, therefore any specific interactions will be disrupted. Of particular interest would be the construction of a chimeric gene that produces a chimera consisting of the propeptide of α-lytic protease and the mature region of SGPB. Baker *et al* (24)were able to demonstrate that the α-lytic propeptide can inhibit SGPB, it would be interesting to see if this inhibitory function can translate into folding of mature SGPB.

The ultimate goal would be the generation of a pro-mature crystal structure like the one for subtilisin BPN' (34). From this, the rational selection of point mutations can be carried out for structure and function analysis.

6. REFERENCES

1. Creighton, T. E. (1994) in *Mechanisms of protein folding* Pain, R. H., ed) pp. 1-22, Oxford University Press, New York, N.Y.

2. Zwangig, R., Szabo, A., and Bagchi, B. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 20-22.

3. Agard, D. A. (1993) *Science* **260**, 1903-1904.

4. Craig, E. A. (1993) *Science* **260**, 1902-1903.

5. Shinde, U., and Inouye, M. (1993) *TIBS* **18**, 442-446.

6. Jorg, E., and Fersht, A. R. (1995) *Mol. Microbiol.* **16**, 609-614.

7. Silen, J. L., McGrath, C. N., Smith, K. R., and Agard, D. A. (1988) *Gene* **69**, 237-244.

8. Stahl, M. L., and Ferrari, E. (1984) *J. Bacteriol.* **158**, 411-418.

9. Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., and Chen, E. Y. (1983) *Nucleic Acids Res.* **11**, 7911-7925.

10. Henderson, G., Krygsman, P., Liu, C. J., Davey, C. C., and Malek, L. T. (1987) *J. Bacteriol.* **169**, 3778-3784.

11. Sidhu, S. S., Kalmar, G. B., Willis, L., and Borgford, T. J. (1995) *J. Biol. Chem.* **270**, 7594-7600.

12. Sidhu, S. S., Kalmar, G. B., Willis, L., and Borgford, T. J. (1994) *J. Biol. Chem.* **269**, 20167-20171.

13. Sidhu, S., Kalmar, G., and Borgford, T. (1993) *Biochem. Cell Biol.* **71**, 454-461.

14. Briggs, M. S., and Gierasch, L. M. (1986) *Adv. Prot. Chem.* **38**, 109-180.

15. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683-4690.

16. Silen, J. L., Frank, D., Fujishige, A., Bone, R., and Agard, D. A. (1989) *J. Bacteriol.* **171**, 1320-1325.

17. Ikemura, H., Takagi, H., and Inouye, M. (1987) *J. Biol. Chem.* **262**, 7859-7864.

18. Shinde, U., and Inouye, M. (1994) *J. Bacteriol.* **115**, 629-636.

19. Ikemura, H., and Inouye, M. (1988) *J. Biol. Chem.* **263**, 12959-12963.

20. Ohata, Y., and Inouye, M. (1990) *Mol. Microbiol.* **4**, 295-304.

21. Stroud, R. M., Kossiakoff, A. A., and Chambers, J. L. (1977) *Ann. Rev. Biophys. Bioeng.* **6**, 177-193.

22. Baker, D., Shiau, A. K., and Agard, D. A. (1993) *Curr. Opin. Cell Biol.* **5**, 966-970.

23. Silen, J. L., and Agard, D. A. (1989) *Nature* **341**, 462-464.

24. Baker, D., Silen, J. L., and Agard, D. A. (1992) *Proteins* **12**, 339-344.

25. Baker, D., Sohl, J. L., and Agard, D. A. (1992) *Nature* **356**, 263-264.

26. Bychkova, V. E., and Ptitsyn, O. B. (1993) *Chemtracts-Biochem. Molec. Biol.* **4**, 133-163.

27. Baker, D., and Agard, D. A. (1994) *Biochemistry* **33**, 7505-7509.

28. Fujishige, A., Smith, K. R., Silen, J. L., and Agard, D. A. (1992) *J. Cell Biol.* **118**, 33-42.

29. Zhu, X., Ohata, Y., Jordan, F., and Inouye, M. (1989) *Nature* **339**, 483-484.

30. Li, Y., Hu, Z., Jordan, F., and Inouye, M. (1995) *J. Biol. Chem.* **270**, 25127-25132.

31. Ohata, Y., Hojo, H., Aimoto, S., Kobayashi, T., Zhu, X., Jordan, F., and Inouye, M. (1991) *Mol. Microbiol.* **5**, 1507-1510.

32. Eder, J., Rheinnecker, M., and Fersht, A. R. (1993) *Biochemistry* **32**, 18-26.

33. Eder, J., Rheinnecker, M., and Fersht, A. R. (1993) *J. Mol. Biol.* **233**, 293-304.

34. Bryan, P., Wang, L., Hoskins, J., Ruvinov, S., Strausberg, S., Alexander, P., Almog, O., Gilliland, G., and Gallagher, T. (1995) *Biochemisrty* **34**, 10310-10318.

35. Eder, J., and Fersht, A. R. (1995) *Mol. Microbiol.* **16**, 609-614.

36. Inouye, M. (1991) *Enzyme* **45**, 314-321.

37. Kobayashi, T., and Inouye, M. (1992) *J. Mol. Biol.* **226**, 931-933.

38. Lerner, C. G., Kobayashi, T., and Inouye, M. (1990) *J. Biol. Chem.* **265**, 20085-20086.

39. Jurasek, L., Johnson, P., Olafson, R. W., and Smillie, L. B. (1971) *Can. J. Biochem.* **49**, 1195-1201.

40. Rawlings, N. D., and Barrett, A. J. (1994) *Methods Enzymol.* **244**, 19-61.

41. Jindal, S., Dudani, A. K., Singh, B., Harley, C. B., and Gupta, R. S. (1989) *Mol. Cell Biol.* **9**, 2279-2283.

42. Venner, T. J., and Gupta, R. S. (1990) *Biochim. Biophys. Acta* **1087**, 336-338.

43. Watanabe, T., Oyanagi, W., Suzuki, K., Ohnishi, K., and Tanaka, H. (1992) *J. Bacteriol.* **174**, 408-414.

44. Simonen, M., and Palva, I. (1993) *Microbiol. Rev.* **57**, 109-137.

45. Pugsley, A. P. (1993) *Microbiol. Rev.* **57**, 50-108.

46. Kawamura, F., and Doi, R. H. (1984) *J. Bacteriol.* **160**, 442-444.

47. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* Second Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

48. (1984) *Focus* **6**, 4-7.

49. Lovett, P., and Amulos, N. P. J. (1989) in *Bacillus* Hardwood, C. R., ed) pp. 355, Plenum, New York and London

50. Chang, S., and Cohen, S. N. (1979) *Mol. Gen. Genet.* **168**, 111-115.

51. Maina, C. V., Riggs, P. D., III, A. G. G., Slatko, B. E., Moran, L. S., Tagliamonte, J. A., McReynolds, L. A., and Guan, C. d. (1988) *Gene* **74**, 365-373.

52. Baardsnes, J., Sidhu, S. S., and Borgford, T. J. (1996) *J. Bacteriol.* submitted.

53. Sanger, F., Nicklen, S., and Coulson, R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.

54. Nakajima, K., Powers, J. C., Ashe, B. M., and Zimmerman, M. (1979) *J. Biol. Chem.* **254**, 4027-4032.

55. Laemmli, U. K. (1970) *Nature* **227**, 680-685.

56. Wilkins, R. J., and Snell, R. G. (1987) *Nucleic Acids Res.* **15**, 7200.

57. Delbaere, L. T. J., Brayer, G. D., and James, M. N. G. (1979) *Can. J. Biochem.* **57**, 135-144.

58. Chou, P., and Fasman, G. D. (1978) *Adv. Enz.* **47**, 45-148.

59. Rose, G. D. (1978) *Nature* **272**, 586-590.

60. Winther, J. R., and Sorensen, P. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9330-9334.

61. Sadis, S., Raghavendra, K., and Hightower, L. E. (1990) *Biochemistry* **29**, 8199-8206.

62. Ramos, C., Winther, J. R., and Kielland-Brandt, M. C. (1994) *J. Biol. Chem.* **269**, 7006-7012.

63. Fersht, A. R., et al. (1985) *Nature* **314**, 235-238.

64. Pohlner, J., Halter, R., Beyreuther, K., and Meyer, T. F. (1987) *Nature* **325**, 458-462.

65. Valls, L. A., Winther, J. R., and Stevens, T. H. (1990) *J. Cell Biol.* **111**, 361-368.

66. Kizer, J. S., and Tropsha, A. (1991) *Biochem. Biophys. Res. Commun.* **174**, 586-592.

67. Wang, L.-F., Kortt, A. A., and Stewart, D. J. (1993) *Gene* **131**, 97-102.

68. Huber, R., and Bode, W. (1978) *Acc. Chem. Res.* **11**, 114-122.

69. Ellis, R. J., and van der Vies, S. M. (1991) *Ann. Rev. Biochem.* **60**, 321-347.