

**THE HTRA FAMILY OF PROTEINS IN *SALMONELLA*
TYPHIMURIUM AND *STAPHYLOCOCCUS AUREUS***

Elaine Mo

**A thesis submitted in accordance with the requirements of the
degree of Doctor of Philosophy at the University of London.**

**The Wolfson Institute for Biomedical Research
University College London**

September 2002

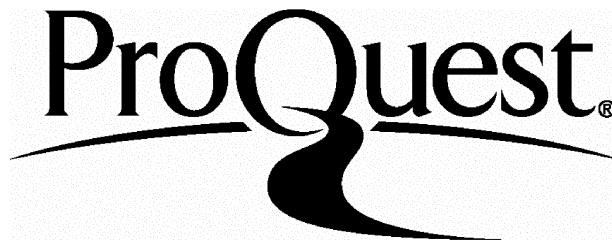
ProQuest Number: U641869

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest U641869

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

This thesis is dedicated to my parents for all their love and support.

ABSTRACT

High temperature requirement A (HtrA) has been identified in many Gram-negative bacteria to be important for bacterial survival during stress conditions, such as heat-shock and oxidative stress. HtrA has been reported to degrade misfolded proteins in the periplasm. Substrate recognition is proposed to be carried out by C-terminal PDZ domains. In *Escherichia coli*, a family of HtrA homologues have been identified. HhoA has a similar domain structure to HtrA, with two PDZ domains, whereas HhoB is a smaller protein and has a single PDZ domain.

This study reports the analysis of the HtrA family in the Gram-negative bacterium, *S. typhimurium*, and identifies a family of HtrA-like homologues in the Gram-positive bacterium, *S. aureus*. Three *S. aureus* HtrA orthologues were identified and named SAHtrA1, SAHtrA2 and SAHtrA3. SAHtrA3 did not have a PDZ domain, thus, this study examines SAHtrA1 and SAHtrA2, both of which have a single PDZ domain. Northern blot analysis showed that the transcription of *sahtrA2* was heat-inducible at 42°C whereas transcription of *sahtrA1* was constitutively expressed. Primer extension experiments predicted that *sahtrA1* had promoter sites similar to that seen in the vegetative σ factor, σ^A . In addition, the PDZ domain of SAHtrA2 was constructed as a recombinant protein and shown to bind the *S. aureus* SsrA-tag, suggesting that this domain was involved in protein:protein interactions.

The *S. typhimurium* HtrA homologues, named STHtrA, STHhoA and STHhoB, exhibited a high level of identity to their *E. coli* counterparts. Mutants of the *htrA* gene family were constructed in *S. typhimurium*. *In vitro* heat-shock studies suggest that *htrA* and *hhoB* are important in bacterial viability at 42°C. The *hhoB* gene also appeared to be important for growth when bacteria were exposed to hydrogen peroxide. Interestingly, under the *in vitro* stress conditions, mutating the *hhoA* gene appeared to have little effect on bacterial viability while mutating both the *htrA* and *hhoB* genes together severely affected bacterial viability and growth. Bacterial virulence was investigated using a mouse model. These *in vivo* experiments showed that in common with the *in vitro* data, *htrA* and *hhoB* mutants were attenuated, whereas mutations in *hhoA* had little affect on bacterial survival. To further characterise the function of the HtrA homologues, recombinant proteins were generated and STHhoA was shown to be proteolytically active.

CONTENTS

<u>ABSTRACT</u>	3
<u>CONTENTS</u>	4
<u>LIST OF FIGURES</u>	12
<u>LIST OF TABLES</u>	16
<u>ABBREVIATIONS</u>	17
<u>ACKNOWLEDGEMENTS</u>	21

CHAPTER 1. INTRODUCTION

1.1 Bacterial stress response	24
1.2 The discovery of HtrA.....	26
1.3 The importance of HtrA.....	28
1.4 The function of HtrA.....	29
1.5 Structural domains of HtrA.....	30
1.5.1 Catalytic domain	31
1.5.2 PDZ domains.....	32
1.6 The crystal structure of HtrA	34
1.7 Substrate recognition by PDZ domains	37
1.7.1 Carboxylate-binding loop	40
1.7.2 Substrate specificity conferred by α B helix	40
1.7.3 Proposed substrate recognition by the PDZ domains of HtrA.....	42
1.7.3.1 Substrate recognition of PDZ1	42
1.7.3.2 Substrate recognition of PDZ2.....	43
1.8 HtrA substrates.....	43
1.8.1 HtrA recognition requirement for unfolded substrates.....	43
1.8.2 HtrA recognition requirement for a hydrophobic core.....	46

1.8.3	<i>In vivo</i> HtrA ligands	47
1.8.3.1	The SsrA-tagging system.....	48
1.9	Regulation of HtrA.....	51
1.9.1	RpoE regulon	52
1.9.1.1	Negative regulation by the anti- σ factors RseABC (regulator of sigma E).....	52
1.9.1.2	The two-component Cpx system.....	54
1.9.1.3	HtrA regulation by the phosphoprotein phosphatases, PrpA and PrpB	55
1.10	Paralogues of HtrA.....	56
1.10.1	HhoA	56
1.10.2	HhoB.....	57
1.11	Identification of HtrA-like homologues in Gram positive bacteria	58
1.11.1	Comparisons of Gram-negative and Gram-positive bacterial HtrA homologues.....	59
1.12	<i>S. aureus</i> , a Gram-positive bacterium	60
1.13	<i>Salmonella</i> , a Gram-negative bacterium	62
1.14	The identification of HtrA-like proteins in <i>S. aureus</i> and <i>S. typhimurium</i>	62
1.15	Aims of this thesis	63

CHAPTER 2. MATERIALS AND METHODS

2.1	DNA manipulation	68
2.1.1	Isolation of <i>S. aureus</i> genomic DNA	68
2.1.1.1	Phenol-chloroform extraction of DNA.....	68
2.1.1.2	Ethanol precipitation	69
2.1.2	Polymerase Chain Reaction (PCR)	69
2.1.2.1	Optimisation of MgCl ₂ concentration for PCR.....	69
2.1.3	Analysis of DNA samples on agarose gel electrophoresis	70
2.1.3.1	Gel preparation.....	70
2.1.3.2	Sample preparation.....	70
2.1.3.3	Electrophoresis.....	71
2.1.4	Cloning of PCR products.....	71

2.1.4.1	Preparation of the plasmid vector.....	71
2.1.4.2	Preparation of insert (PCR product).....	72
2.1.4.3	Ligating inserts into vectors.....	72
2.1.4.4	Transformation of DH5 α competent cells.....	73
2.1.5	Purification of plasmid DNA.....	73
2.1.5.1	Determining presence of insert in plasmid.	74
2.1.6	DNA sequencing	74
2.2	Recombinant protein expression.....	76
2.2.1	Transformation of plasmid into competent cells.....	76
2.2.1.1	Transformation by electroporation.....	76
2.2.1.2	Transformation by heat-shock	78
2.2.2	Expression of recombinant proteins	78
2.2.2.1	Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).....	79
2.2.2.2	Tricine gel electrophoresis.....	80
2.2.2.3	Coomassie staining of SDS-PAGE gels	81
2.2.2.4	Silver staining of SDS-PAGE gels.....	81
2.2.2.5	Western blot analysis of SDS-PAGE separated proteins	81
2.2.3	Verifying solubility of recombinant proteins.....	83
2.2.4	Purification of GST-tagged recombinant proteins	83
2.2.4.1	GSTrap affinity chromatography	84
2.2.4.2	Removal of thrombin using the HiTrap Benzamidine column	85
2.2.4.3	Gel filtration chromatography.....	85
2.2.5	Small scale purification of 6xHis-tagged recombinant proteins.....	86
2.3	RNA manipulation	87
2.3.1	Isolation of <i>S. aureus</i> RNA.....	87
2.3.2	Spectrophotometric quantification of RNA samples.....	88
2.3.3	Separation of RNA by denaturing gel electrophoresis.....	89
2.3.3.1	Preparation of RNA samples	89
2.3.3.2	Electrophoresis of RNA denaturing gel.....	89
2.3.4	Northern blot analysis.....	90
2.3.4.1	Radioactive probe.....	90

2.3.4.2	Hybridisation of probe to membrane.....	92
2.3.4.3	Stripping filters.....	92
2.4	Primer extension.....	92
2.4.1	Construction of template primer for manual sequencing	92
2.4.2	Manual sequencing.....	93
2.4.2.1	Denaturing plasmid DNA	93
2.4.2.2	Sequencing of DNA templates.....	94
2.4.3	Analysis of DNA sequence by polyacrylamide sequencing gels.....	94
2.4.4	Primer extension.....	95
2.4.4.1	Radiolabelling of primer.....	95
2.4.4.2	Primer extension.....	95
2.5	Heat shock of <i>S. aureus</i>	96
2.6	Construction of <i>S. typhimurium</i> mutants	97
2.6.1	Allelic replacement.....	97
2.6.1.1	Template DNA for resistance cassettes.....	97
2.6.1.2	Construction of resistance cassettes	98
2.6.1.3	Transformation of <i>S. typhimurium</i> SL3261 competent cells with pBAD λ RED	98
2.6.1.4	Induction of <i>S. typhimurium</i> SL3261 pBAD λ RED	99
2.6.1.5	Verification of allelic replacement.....	99
2.6.2	Transduction of mutation into <i>S. typhimurium</i> strain SL1344.....	100
2.6.2.1	Plate stock prep	101
2.6.2.2	Transduction	101
2.7	Protease detection using zymograms	102
2.7.1	Zymograms	102
2.8	<i>S. typhimurium</i> stress conditions <i>in vitro</i>	103
2.8.1	Heat shock studies	103
2.8.2	Oxidative stress of <i>S. typhimurium</i> SL3261 mutants	104
2.9	Attenuation of <i>S. typhimurium</i> SL1344 mutants <i>in vivo</i>	105
2.9.1	Quantification of <i>S. typhimurium</i> SL1344 mutants.....	105
2.9.1.1	Inoculation of BALB/c mice.....	105
2.9.1.2	Viable counts.....	106
2.10	Statistical analysis	106

CHAPTER 3. IDENTIFICATION OF HTRA HOMOLOGUES IN *S. aureus* AND *S. typhimurium*

3.1	Introduction.....	107
3.2	Identification of HtrA orthologues in <i>S. typhimurium</i>	108
3.3	Identification of HtrA orthologues in <i>S. aureus</i>	113
3.4	Summary.....	119

CHAPTER 4. TRANSCRIPTIONAL ANALYSIS OF THE *S. aureus htrA* ORTHOLOGUES

4.1	Introduction.....	122
4.2	Isolation of <i>S. aureus</i> RNA.....	122
4.3	Preparation of probes for Northern blot analysis	125
4.4	Northern analysis of RNA transcription by heat shock.....	126
4.5	Determining the start sites for <i>sahtrA1</i> and <i>sahtrA2</i> transcription.....	129
4.5.1	Primer extension of <i>sahtrA1</i> mRNA	130
4.5.2	Primer extension of <i>sahtrA2</i> mRNA	133
4.6	Summary.....	134

CHAPTER 5. FUNCTIONAL ANALYSIS OF THE RECOMBINANT *S. aureus htrA* ORTHOLOGUES

5.1	Introduction.....	136
5.2	Cloning and expression strategies	136
5.2.1	Cloning strategy	136
5.2.2	Expression strategy.....	139
5.3	Cloning, expression and purification of recombinant SAHtrA1.....	145
5.3.1	Cloning and expression of recombinant native SAHtrA1	145
5.3.2	Cloning and expression of recombinant 6xHis-tagged SAHtrA1	148
5.4	Cloning, expression and purification of SAHtrA1 subdomains	154
5.4.1	Cloning and expression of the catalytic domain of SAHtrA1	154
5.4.2	Cloning and expression of the PDZ domain of SAHtrA1	156
5.4.2.1	Cloning of the PDZ of SAHtrA1 as a GST-tagged protein ...	156

5.4.2.2	Purification of GST-PDZ1 by FPLC	158
5.4.2.3	Cloning of the PDZ of SAHtrA1 as a 6xHis-tagged protein .	169
5.5	Cloning, expression and purification of SAHtrA2 and its subdomains	171
5.5.1	Cloning and expression of SAHtrA2.....	171
5.5.2	Cloning, expression and purification of the PDZ domain of SAHtrA2	177
5.5.2.1	Cloning and expression of PDZ2 as a GST-tagged protein...	177
5.5.3	Cloning, expression and purification of 6xHis-tagged PDZ2.....	180
5.6	Protease activity of the full length recombinant proteins	181
5.7	Summary	183

CHAPTER 6. FUNCTIONAL ANALYSIS OF THE RECOMBINANT *S. typhimurium htrA* ORTHOLOGUES

6.1	Introduction.....	186
6.2	Expression of recombinant STHhoA	186
6.2.1	Cloning of STHhoA	186
6.2.2	Expression and purification of STHhoA	190
6.3	Cloning and expression of STHhoB.....	194
6.3.1	Cloning of STHhoB.....	194
6.3.2	Expression and purification of STHhoB	194
6.4	Proteolytic activity of the <i>S. typhimurium</i> HtrA orthologues	196
6.5	Summary	200

CHAPTER 7. VIABILITY OF MUTANTS OF THE *htrA* FAMILY

7.1	Introduction.....	203
7.2	Constructing <i>S. typhimurium</i> mutants	204
7.2.1	Construction of single mutants	204
7.2.1.1	Preparation of a kanamycin resistance cassette for the construction of linear allelic replacement fragments	204
7.2.1.2	Allelic replacement of <i>htrA</i> , <i>hhoA</i> , <i>hhoB</i> or <i>hhoAB</i>	206
7.2.2	Construction of double and triple mutants.....	206

7.2.2.1	Production of <i>htrA::GM</i> and <i>htrA::CAT</i> single mutants	208
7.2.2.2	Production of double and triple mutants.....	210
7.3	Viability of mutants under <i>in vitro</i> stress conditions	210
7.3.1	Survival of mutants under heat shock conditions	212
7.3.2	Survival of mutants under oxidative stress conditions	220
7.4	Viability of mutants in an <i>in vivo</i> system	230
7.4.1	Transduction of <i>S. typhimurium</i> SL3261 mutants into <i>S. typhimurium</i> SL1344.....	230
7.4.2	Bacterial virulence in BALB/c mice	230
7.5	Summary.....	236

CHAPTER 8. DISCUSSION

8.1	Introduction.....	240
8.2	Identification of HtrA homologues in <i>S. aureus</i> and <i>S. typhimurium</i>	241
8.3	HtrA homologues in <i>S. aureus</i>	243
8.3.1	SAHtrA1	243
8.3.2	SAHtrA2	246
8.4	HtrA homologues in <i>S. typhimurium</i>	248
8.5	Conclusions and future work	252

<u>REFERENCES</u>		255
--------------------------	-------	-----

APPENDIX

A.1	Mechanism of serine proteases: the catalytic triad.....	272
A.2	Primers used for the production of recombinant proteins	273
A.3	Bacterial strains used.....	274
A.4	Primers used for primer extension experiments.....	274
A.5	Primers used for sequencing reactions	275
A.6	Primers used for making DNA probes for Northern blot analysis.....	276
A.7	Primers used for constructing linear DNA fragment for allelic replacement .	276
A.8	Primers for colony PCR to verify allelic replacement	277

A.9 Sequenced ORF of <i>sahtrA1</i>	278
A.10 Sequenced ORF of <i>sahtrA2</i>	279
A.11 Sequenced ORF of <i>sthhoA</i>	280
A.12 Sequenced ORF of <i>sthhoB</i>	281

LIST OF FIGURES

CHAPTER 1

Figure 1.1	Diagram of important domains of HtrA in Gram-negative bacteria	31
Figure 1.2	Crystal structure of a monomer of <i>E. coli</i> HtrA	35
Figure 1.3	The crystal structure of the <i>E. coli</i> HtrA oligomer	36
Figure 1.4	Crystal structure of the third PDZ domain of PSD-95 (amino acids 302-402) bound to the peptide KQTSV	39
Figure 1.5	Comparison of the multimeric structures of the chaperone protein GroEL and the periplasmic protease HtrA	45
Figure 1.6	The lateral cavity of <i>E. coli</i> HtrA.....	46
Figure 1.7	Hydrophobic patches within the lateral cavity of <i>E. coli</i> HtrA	47
Figure 1.8	Model for SsrA-mediated tagging of proteins synthesized from messages lacking stop codons.....	49
Figure 1.9	Schematic diagram of transcriptional regulation of <i>htrA</i>	55
Figure 1.10	Schematic diagram of the protein domains of the HtrA paralogues in <i>E. coli</i>	58
Figure 1.11	Comparison of structures of Gram-negative and Gram-positive bacteria.....	60

CHAPTER 3

Figure 3.1	Schematic diagram of the protein domains of the HtrA orthologues in <i>E. coli</i>	108
Figure 3.2	Protein sequence alignment of <i>S. typhimurium</i> HtrA orthologues versus their <i>E. coli</i> counterparts.....	110
Figure 3.3	Protein sequence alignment of the HtrA orthologues in <i>S. typhimurium</i> versus <i>E. coli</i> HtrA	112
Figure 3.4	Protein sequence alignment of <i>S. aureus</i> HtrA orthologues versus the <i>S. typhimurium</i> HtrA	114
Figure 3.5	Protein sequence alignment of the HtrA orthologues in <i>S. aureus</i> versus <i>E. coli</i> HtrA.....	116
Figure 3.6	Hydrophobicity plots of HtrA orthologues identified in Gram-positive bacteria	118

Figure 3.7	Schematic diagram of the protein domains in the HtrA homologues in <i>S. typhimurium</i> and <i>S. aureus</i> in comparison to the <i>E. coli</i> HtrA orthologues.....	120
------------	---	-----

CHAPTER 4

Figure 4.1	Comparison of RNA isolation methods to produce the highest RNA yields.....	124
Figure 4.2	Detection of transcript levels during heat shock of <i>S. aureus</i>	127
Figure 4.3	Primer extension of <i>sahtrA1</i> mRNA	131
Figure 4.4	Predicted domains upstream from the translation start site of <i>sahtrA1</i>	132

CHAPTER 5

Figure 5.1	Effect of increasing MgCl ₂ concentrations on the yield of PCR products.....	138
Figure 5.2	Cloning strategy for native recombinant proteins using pTrc99A.....	140
Figure 5.3	Cloning strategy for GST-tagged recombinant proteins	141
Figure 5.4	Cloning strategy of 6xHis-tagged recombinant proteins using pQE-30	142
Figure 5.5	Cloning strategy of 6xHis-tagged recombinant proteins using pRSET-B.....	143
Figure 5.6	Expression strategy to optimise protein expression	144
Figure 5.7	Optimisation of the expression of soluble SAHtrA1 in <i>E. coli</i> BL21(DE3)Codon Plus and <i>E. coli</i> BL21(DE3)pLysS cells from pTSAHtrA1-41.....	146
Figure 5.8	Expression of soluble SAHtrA1 in JM109, DH5a, BL21(DE3)pLysS and BL21(DE3)CodonPlus.....	149
Figure 5.9	Expression of soluble 6xHis-SAHtrA from pRSAHtrA1-3	151
Figure 5.10	Small-scale purification of 6xHis-SAHtrA1 in <i>E. coli</i> BL21(DE3)CodonPlus.....	153
Figure 5.11	Western blot analysis of the effect of increasing IPTG concentration on protein expression of pGCAT1-13 in <i>E. coli</i> BL21(DE3)pLysS and <i>E. coli</i> BL21(DE3)CodonPlus cells.....	155
Figure 5.12	Expression of soluble GST- PDZ1 from pGPDZ1-1	157

Figure 5.13	Purification of GST-PDZ1 using GSTrap	159
Figure 5.14	Optimisation of thrombin digest of GST-PDZ1	161
Figure 5.15	Separation of species from the digest of GST-PDZ1	163
Figure 5.16	Removal of contaminating GST-species using cation exchange chromatography.....	164
Figure 5.17	Removal of thrombin from the digest reaction by HiTrap Benzamidine FF	166
Figure 5.18	Removal of residual low molecular weight contaminants from PDZ1 by gel filtration chromatography.....	167
Figure 5.19	Expression of 6xHis-PDZ1 from pRPDZ1-3.....	170
Figure 5.20	Alignment of <i>sahtrA2</i> in <i>S. aureus</i> SMITH strain against the NCTC 8325, N315 and Mu50 <i>S. aureus</i> strains.....	173
Figure 5.21	Protein expression from pTSAHtrA2-715.....	174
Figure 5.22	Expression of SAHtrA2 using pRSAHtrA2-2	176
Figure 5.23	Expression of soluble GST-PDZ2 using pGPDZ2-1	178
Figure 5.24	Expression and purification of 6xHis-PDZ2	182

CHAPTER 6

Figure 6.1	Cloning strategy of 6xHis-tagged recombinant proteins using pQE-30	187
Figure 6.2	Cloning strategy of 6xHis-tagged recombinant proteins using pRSET-B.....	188
Figure 6.3	Cloning strategy for native recombinant proteins using pTrc99A.....	189
Figure 6.4	Expression strategy to optimise protein expression	191
Figure 6.5	Expression of soluble 6xHis-tagged STHhoA from pQHhoA-4 and pRHhoA-1	192
Figure 6.6	Expression of soluble 6xHis-STHhoB from pQHhoB-6 and pRHhoB-3	195
Figure 6.7	Protease activity of HtrA homologues	197
Figure 6.8	Effect of protease inhibitors on STHhoA* protease activity	199

CHAPTER 7

Figure 7.1	Preparation of kanamycin cassette as the insert DNA for allelic exchange	205
------------	--	-----

Figure 7.2	Verification of mutants after one round of allelic replacement	207
Figure 7.3	Verification of <i>htrA::GM</i> and <i>htrA::CAT</i> mutants.....	209
Figure 7.4	Verification of allelic replacement in the double and triple mutants .	211
Figure 7.5	Comparison of the growth profiles of the <i>S. typhimurium</i> SL3261 mutants at 42°C and 37°C.....	213
Figure 7.6	Comparison of the growth profiles of the <i>S. typhimurium</i> SL3261 single mutants at 42°C and 37°C	214
Figure 7.7	Comparison of the growth profiles of the <i>S. typhimurium</i> SL3261 double and triple mutants at 42°C and 37°C.....	216
Figure 7.8	Bacterial viability under heat shock conditions	219
Figure 7.9	Growth profile of <i>S. typhimurium</i> SL3261 mutants in the presence of increasing concentrations of hydrogen peroxide	221
Figure 7.10	Effect of increasing concentrations of H ₂ O ₂ on SL3261 mutant growth.....	225
Figure 7.11	Effect of exposure to 1 mM H ₂ O ₂ for 3.5 hr at 37°C	229
Figure 7.12	Verification of <i>S. typhimurium</i> SL1344 mutants by PCR	231
Figure 7.13	Viability of <i>S. typhimurium</i> SL1344, <i>htrA::KAN</i> , <i>hhoA::KAN</i> , <i>hhoB::KAN</i> and <i>hhoAB::KAN</i> mutants in a mouse model of infection	233
Figure 7.14	Viability of <i>S. typhimurium</i> SL1344, <i>htrA::GM</i> , <i>htrA::GM hhoA::KAN</i> , <i>htrA::GM hhoB::KAN</i> and <i>htrA::GM hhoAB::KAN</i> mutants in a mouse model of infection.....	235
Figure 7.15	Viability of <i>S. typhimurium</i> SL1344, <i>htrA::GM</i> , <i>htrA::GM hhoA::KAN</i> , <i>htrA::GM hhoB::KAN</i> and <i>htrA::GM hhoAB::KAN</i> mutants in a mouse model of infection using a higher inoculum dose.....	237

LIST OF TABLES

CHAPTER 1

Table 1.1	Species published with the <i>htrA</i> gene.....	27
Table 1.2	Published crystal structures of PDZ domains	38

CHAPTER 2

Table 2.1	Sequencing conditions for CEQ-2000 sequencing machine.....	75
Table 2.2	Bacterial strains	76
Table 2.3	Length of pre-heat shock incubation	78
Table 2.4	Antibodies used for Western blot analysis	82
Table 2.5	FPLC conditions using GSTrap	84
Table 2.6	FPLC conditions using HiTrap Benzamidine FF	85
Table 2.7	FPLC conditions using HiLoad Superdex 75 26/60 prep grade	86
Table 2.8	Bioscreen conditions for heat shock of <i>S. typhimurium</i> mutants.....	104

ABBREVIATIONS

°C	Degrees Celsius
Å	Angstrom
Acc #	Accession number
ATP	Adenosine triphosphate
BHI	Brain heart infusion
BLAST	Basic local alignment search tools
bp	Base pairs
BSA	Bovine serum albumin
Cal	Colysin A lysis
cDNA	Complementary DNA
Ci	Curies
cm	Centimetres
CV	Column volume
CaCl ₂	Calcium chloride
dATP	2'-Deoxyadenosine-5'-triphosphate
dCTP	2'-Deoxycytosine-5'-triphosphate
ddH ₂ O	Double distilled water
DegP	Degradation P
DegQ	Degradation Q
DegS	Degradation S
DEPC	Diethyl pyrocarbonate
DFP	Di-isopropylfluorophosphate
dGTP	2'-Deoxyguanosine-5'-triphosphate
dH ₂ O	Distilled water
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleoside triphosphates
DTCS	Dye terminator cycle sequencing
DTT	Dithiothreitol
dTTP	2'-Deoxythymidine-5'-triphosphate
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid
fmol	Femtomole

FPLC	Fast protein liquid chromatography
g	Gram
GST	Glutathione-S-transferase
H ₂ O ₂	Hydrogen peroxide
HCl	Hydrochloric acid
HhoA	Homologue of HtrA A
HhoA*	Processed HhoA (signal peptide removed)
HhoB	Homologue of HtrA B
hr	Hour
HtrA	High temperature requirement A
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilo base pair(s)
KCl	Potassium chloride
kDa	Kilo Daltons
K ₂ HPO ₄	Di-potassium phosphate
KH ₂ PO ₄	Potassium di-hydrogen phosphate
L-Broth	Lennox L-Broth
LTP	Long term potentiation
M	Molar
Mb	Mega base pairs
MBP	Maltose binding protein
MCS	Multiple cloning site
Mg ²⁺	Magnesium ion
MgCl ₂	Magnesium chloride
MgSO ₄	Magnesium sulphate
min	Minute
ml	Millilitre
mm	Millimetre
mM	Millimolar
mmol	Millimole
Mn ²⁺	Manganese ion
Na ₂ HPO ₄	Di-sodium phosphate
NaOH	Sodium hydroxide

ng	Nanogram
nm	Nanometre
nM	Nanomolar
NMDA	N-methyl-D-aspartate
nNOS	Neuronal nitric oxide synthase
nt	Nucleotide
OD	Optical density
ORF	Open reading frame
p	Probability
PBP-3	Penicillin-binding protein 3
PCR	Polymerase chain reaction
pmol	Picomole
PNK	Polynucleotide kinase
PDZ	Acronym of PSD-95, Dlg, Zo-1
RNA	Ribonucleotide triphosphate
rpm	Rotations per minute
RT	Reverse transcriptase
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide electrophoresis
sec	Second
TAE	Tris-acetate-EDTA
T-Broth	Terrific Broth
TEMED	N,N,N',N'-tetramethylethylenediamine
tRNA	Transfer RNA
Tris	Tris(hydroxymethyl)aminomethane
TS-Broth	Tryptic Soy Broth
U	Units
μl	Microlitre
μM	Micromolar
UV	Ultraviolet
V	Volts
v/v	Volume to volume
W	Watts

w/v	Weight to volume
X-Gal	5-bromo-4-chloro-3-indoyl- β -D-galactosidase

ACKNOWLEDGMENTS

I would like to thank my supervisor, Professor Ian Charles, for his encouragement, guidance and advice throughout the course of this study. I thank the Biotechnology and Biological Sciences Research Council, Tripos (Missouri, USA) for funding this project and to Arrow Therapeutics Ltd (London, UK) for providing additional financial aid to allow me to complete this study.

Many thanks to Professor Duncan Maskell at the University of Cambridge for his kind gift of the *S. typhimurium* strains and to Dr. Sarah Peters and Ms. Chrissie Willers at Arrow Therapeutics for their help with the *in vivo* work and also for their guidance and friendship. Thanks, also to Professor Alistair Hawkins, Dr. Heather Lamb and Ms. Alison Spiers at the University of Newcastle for their help with the purification of the PDZ domains and the SPR experiments.

Special thanks to Dr. Dagmar Alber for her continual support and advice throughout this study and for taking time to read this thesis. I would also like to thank Mr. Joo Wook Ahn for the many hours of trouble-shooting we have had over RNA. In addition, I would like to thank Professor John Garthwaite and Dr. Giti Garthwaite for their support over this study.

I would like to express my gratitude to Dr. Jo Budworth, Dr. Kerry Wheeler at Arrow Therapeutics for their help and suggestions.

My special thanks are to all my friends who have been patiently waiting for me to exist again, especially Dan for all his love and support (and cooking!)

Finally, my biggest thanks are to my parents and brothers for all the love and support they have provided throughout my life.

CHAPTER 1. INTRODUCTION

1.1	Bacterial stress response.....	24
1.2	The discovery of HtrA.....	26
1.3	The importance of HtrA	28
1.4	The function of HtrA.....	29
1.5	Structural domains of HtrA	30
1.5.1	Catalytic domain	31
1.5.2	PDZ domains.....	32
1.6	The crystal structure of HtrA.....	34
1.7	Substrate recognition by PDZ domains.....	37
1.7.1	Carboxylate-binding loop	40
1.7.2	Substrate specificity conferred by α B helix	40
1.7.3	Proposed substrate recognition by the PDZ domains of HtrA.....	42
1.7.3.1	Substrate recognition of PDZ1	42
1.7.3.2	Substrate recognition of PDZ2.....	43
1.8	HtrA substrates.....	43
1.8.1	HtrA recognition requirement for unfolded substrates.....	43
1.8.2	HtrA recognition requirement for a hydrophobic core.....	46
1.8.3	<i>In vivo</i> HtrA ligands	47
1.8.3.1	The SsrA-tagging system.....	48
1.9	Regulation of HtrA.....	51
1.9.1	RpoE regulon	52
1.9.1.1	Negative regulation by the anti- σ factors RseABC (regulator of sigma E).....	52
1.9.1.2	The two-component Cpx system.....	54
1.9.1.3	HtrA regulation by the phosphoprotein phosphatases, PrpA and PrpB	55
1.10	Paralogues of HtrA.....	56

1.10.1 HhoA	56
1.10.2 HhoB.....	57
1.11 Identification of HtrA-like homologues in Gram positive bacteria	58
1.11.1 Comparisons of Gram-negative and Gram-positive bacterial HtrA homologues	59
1.12 <i>S. aureus</i> , a Gram-positive bacterium	60
1.13 <i>Salmonella</i> , a Gram-negative bacterium	62
1.14 The identification of HtrA-like proteins in <i>S. aureus</i> and <i>S. typhimurium</i>	62
1.15 Aims of this thesis	63

1.1 Bacterial stress response

Bacteria are continuously challenged by environmental changes throughout their lifespan. Such challenges elicit a stress response from the bacteria in order to survive. For example, on infecting a host, bacteria encounter the immune response, a component of which, macrophages, act to remove foreign agents by means of the oxidative burst. In this environment, the bacteria encounter hostile conditions involving reactive oxygen species, such as superoxide radicals, hydroxyl radicals as well as nitric oxide. Another common stress condition, heat shock, can result in denatured proteins that can accumulate and become toxic to the bacteria.

Thus, bacteria must be able to adapt to various environmental conditions in order to enhance their chances of survival. This adaptation is brought about by a variety of mechanisms of stress response whereby a set of proteins are rapidly and transiently expressed in order to limit the damage caused to the bacteria. Such proteins were first observed when observing the cellular response of *E. coli* to heat shock and hence were termed heat shock proteins. In bacteria, heat shock proteins are distributed throughout the cell, within the cytoplasm and periplasm, on the cytoplasmic membrane and on the cell surface of the bacterium (Qoronfleh, M.W. *et al.*, 1998). Heat shock proteins consist of two major categories, chaperone proteins, which assist the renaturation of denatured proteins, and proteases, which degrade misfolded proteins.

The main role of chaperone proteins is to prevent aggregation of proteins and assist in refolding the denatured protein into its native form. Chaperone proteins have a hydrophobic pocket which recognises hydrophobic stretches of amino acids on their

substrates. Hydrophobic regions are normally locked away in the core of the protein, hidden away from the aqueous environment. Thus, exposure of these regions implies that the protein has folded incorrectly or has been denatured. The binding and releasing of a misfolded protein by chaperone proteins is accompanied with ATP hydrolysis, which is probably associated with the release of the protein. (reviewed in Bukau, B. & Horwich, A., 1998, Alberts, B. *et al.*, 1994). The exposure of *E. coli* to stress conditions results in the upregulation of the chaperone proteins DnaJ, DnaK, GrpE, GroEL and GroES (Laskowska, E. *et al.*, 1996).

Should the misfolded protein be beyond repair, chaperone proteins can target the protein to be removed from the cell by degradation by proteases. Some of these proteases also have chaperone subunits, thus, entwining the two methods of aberrant protein removal within the same enzyme. Cytoplasmic proteases are all ATP-dependent (Smith, C.K. *et al.*, 1999). *E. coli* has at least five ATP-dependent proteases, Lon, HflB, ClpAP, ClpXP and ClpYQ (Gottesman, S. *et al.*, 1998) which have distinct substrate preferences. For instance, HflB is important for the degradation of the sigma factor, σ^H , while ClpXP degrades σ^{32} (Gottesman, S. *et al.*, 1998, Smith, C.K. *et al.*, 1999). However, a common feature of these proteases is the ability to recognise tail sequences of their substrates. At least seven *E. coli* periplasmic proteases have been cloned including the genes, *ptr*, *lep*, *lsp*, *prc*, *sspA*, *htrA* and *ompT*. Of these, the *prc* gene product is shown to be important in heat-shock and also in osmotic shock. Like the cytoplasmic proteases, it also recognises its substrates by tail-specific recognition and has been shown to process the penicillin-binding protein 3. The *htrA* gene product is also important in heat-shock response and it is this protein that is investigated in this study.

1.2 The discovery of HtrA

In 1981, Swamy and Goldberg identified eight soluble proteins in *Escherichia coli* with proteolytic abilities. Six of these enzymes, named Do, Re, Mi, Fa, So and La, hydrolysed [¹⁴C]-globin and [³H]-casein (Swamy, K.H.S. & Goldberg, A.L., 1981). The proteolytic activity of protease La (also known as Lon, discussed earlier in 1.1) was shown to be ATP- and Mg²⁺-dependent (Swamy, K.H.S. & Goldberg, A.L., 1981). In 1989, Lipinska *et al.*, identified an *E. coli* gene product which was essential for the bacterial survival at an elevated temperature of 42°C. *E. coli* mutants, obtained from the insertion of mini-Tn10 transposon, were screened for growth at 42°C. Two mutants were unable to grow at this elevated temperature. On closer inspection, these mutants both had insertions in the same gene, which was termed *htrA* (high temperature requirement A). The gene was mapped to approximately 3.7 min on the chromosome and encoded an unstable protein of 52 kDa, which was processed to a mature 48 kDa protein (Lipinska, B. *et al.*, 1989). Meanwhile, also in 1989, Strauch *et al.* characterised *degP* (degradation P). This gene had been previously identified in a mutant that prevented periplasmic degradation of abnormal proteins (Strauch, K.L. & Beckwith, J., 1988). DegP was concluded to be identical to HtrA as the gene was mapped to approximately 3.5 min on the *E. coli* chromosome and the mutant phenotype was also thermosensitive. In addition, cloning of *degP* resulted in a protein with proteolytic properties with a molecular mass of approximately 50 kDa. Finally in 1991, it was concluded that protease Do was identical to HtrA due to the similarities between the proteins. Both were proteases of similar sizes, which were induced under heat-shock conditions and were essential for growth at elevated temperatures. One difference between protease

Do and HtrA, however, was the subcellular location of these proteins. HtrA is a peripheral membrane protein on the periplasmic side of the inner membrane whereas protease Do was thought to be located in the cytoplasm (Swarmy, K.H.S. *et al.*, 1983). This discrepancy in the location of protease Do in the cytoplasm may have occurred because of experimental methods by which the cells were sonicated.

Furthermore, HtrA homologues have been identified in numerous bacterial species, listed in Table 1.1.

Bacterial strain	Accession number	Reference
<i>Rickettsia typhi</i>	D78346	Kim, J.-H. & Hahn, M.-J., 2000
<i>Brucella abortus</i>	Q44597	Phillips, R.W. & Roop, R.M., II, 2001
<i>Buchnera aphidicola</i>	O85291	Thao, M.L. & Baumann, P., 1998
<i>Haemophilus influenzae</i>	O68198	Loosmore, S.M. <i>et al.</i> , 1998
<i>Bradyrhizobium japonicum</i>	U55047	Naberhaus, F. <i>et al.</i> , 1998
<i>Rhizobium meliloti</i>	U31512	Glazebrook, J. <i>et al.</i> , 1996!
<i>Yersinia enterocolitica</i>	Q56885	Yamamoto, T. <i>et al.</i> , 1996
<i>Yersinia pestis</i>	NP668143	Williams, K. <i>et al.</i> , 2000
<i>S. typhimurium</i>	P26982	Johnson, K. <i>et al.</i> , 1991
<i>Chlamydia trachomatis</i>	P18584	Kahane, S. <i>et al.</i> , 1990
<i>Haematococcus pluvialis</i>	O04674	Hershkovits, G. <i>et al.</i> , 1997
<i>Klebsiella pneumoniae</i>	Q8RSS1	Cortés, G. <i>et al.</i> , 2002
<i>Bartonella henselae</i>	L09274	Anderson, B. <i>et al.</i> , 1996

Table 1.1 – Species published with the *htrA* gene

1.3 The importance of HtrA

The identification of HtrA homologues in numerous bacterial species implicates the importance of this gene (discussed earlier 1.2). HtrA has been shown to be important in a number of stress conditions. Mutations in the *htrA* gene in *E. coli* showed increase sensitivity to heat shock indicating its importance in thermoresistance (Seol, J.H. *et al.*, 1991, Strauch, K.L. *et al.*, 1989, Lipinska, B. *et al.*, 1989). This inability of *E. coli htrA* mutant strains to grow at elevated temperatures have also been shown in a number of other bacteria including *L. pneumophila* (Pedersen, L.L. *et al.*, 2001), *B. japonicum* (Naberhaus, F. *et al.*, 1998) and *Y. pestis* (Williams, K. *et al.*, 2000). HtrA also has a role in bacterial viability during oxidative stress conditions (Skorko-Glonek, J. *et al.*, 1999, Smeds, A. *et al.*, 1998, Phillips, R.W. & Roop, R.M., II, 2001). Interestingly, the HtrA homologue in *S. typhimurium* does not appear to be essential for heat shock survival, but instead is important in survival in response to oxidative stress (Johnson, K. *et al.*, 1991).

In vivo studies on mutant HtrA in *S. typhimurium* showed that HtrA was important in bacterial virulence as *S. typhimurium* with mutant HtrA is highly attenuated in mice and macrophages (Johnson, K. *et al.*, 1991, Bäumler, A.J. *et al.*, 1994). HtrA also appears to be important in *Y. enterocolitica* virulence (Li, S.-R. *et al.*, 1996). Infection of the wild-type *Y. enterocolitica* causes death in susceptible mice. However, when dosed with a mutant HtrA strain, all mice survive. Furthermore, the infection is cleared twenty-one days post-infection (Li, S.-R. *et al.*, 1996). The attenuation of mutant HtrA strains in these studies was suggested to be due to an inability to survive or replicate in the host, caused by the mutant HtrA protein (Johnson, K. *et al.*, 1991).

1.4 The function of HtrA

HtrA was initially identified as a protease (see 1.2). The proteolytic activity of HtrA was shown to be important in the degradation of abnormal periplasmic protein (Strauch, K.L. & Beckwith, J., 1988). A native protein, maltose-binding protein (MBP), was engineered with two in-frame deletions in the encoding gene, *malE*, which rendered the protein unstable. When introduced into a *htrA* mutant strain of *E. coli*, the mutant MBP stability was significantly increased suggesting that HtrA was involved in the degradation of this protein. The unstable MBP mutants were then engineered with a signal-sequence mutation which resulted in the proteins being located in the cytoplasm. These cytoplasmic MBP mutants were also unstable but unlike the periplasmic MBP mutants, no stability was conferred when introduced into the same *htrA* mutant (Strauch, K.L. & Beckwith, J., 1988). Thus, HtrA is only involved in the degradation of unstable periplasmic proteins and not involved in cytoplasmic protein degradation. This ability to remove abnormal periplasmic proteins is vital for the bacterial survival in stress conditions, as these environments result in the accumulation of abnormal proteins both in the periplasm and the cytoplasm. HtrA acts to remove the accumulation of the aberrant periplasmic proteins, thus reducing the number of toxic agents and promoting survival.

In 1999, Speiss *et al.* showed that HtrA also has chaperone abilities (Speiss, C. *et al.*, 1999). The function of chaperones, as discussed in 1.1, is to assist in the folding of proteins into their native state. A natural substrate of HtrA is MalS, which belongs to the α -amylase family and hydrolyses the $\alpha(1,4)$ glycosidic bonds in long maltodextrins. MalS has two intramolecular disulphide bonds, which are formed

between C⁴⁰-C⁵⁸ and C¹⁰⁴-C⁵²⁰. The formation of these bonds is dependent on the gene product of *dsbA*, a periplasmic oxidoreductase, which is involved in disulphide bond formation (Bardwell, J.C. *et al.*, 1991). Thus, in a *dsbA* mutant strain, the activity of MalS was reduced by 20-fold at 37°C. In a *dsbA htrA* double mutant strain, a further 18-fold decrease in MalS activity was observed when compared against the single *dsbA* mutant strain. This was reduced to a 11-fold decrease when a non-proteolytic HtrA was expressed in the *dsbA htrA* mutant strain. A temperature-dependent switch between the chaperone and proteolytic activities of HtrA was demonstrated by *in vitro* studies. Proteolytically inactive HtrA could refold MalS at both 28°C and 37°C. Wild-type HtrA also refolded MalS at 28°C but at 37°C MalS was degraded. Furthermore, the proteolytic activity of HtrA was shown to be temperature dependent. The activity at 28°C was shown to be one-eighth of that at 42°C, with a major increase in activity occurring between 32 to 42°C (Speiss, C. *et al.*, 1999). Thus, it was concluded that HtrA could act as a chaperone or a protease in a temperature-dependent manner.

1.5 Structural domains of HtrA

The best studied HtrA is from *E. coli*. The enzyme is a large oligomeric protein of approximately 500 kDa consisting of twelve 48 kDa monomers (Strauch, K.L. *et al.*, 1989, Swamy, K.H.S. *et al.*, 1983). Each monomer is synthesised as an unstable precursor of 51 kDa (Lipinska, B. *et al.*, 1990, Lipinska, B. *et al.*, 1989). This is then processed into the mature 48 kDa form by cleavage at a signal peptidase site which results in the removal of the N-terminal twenty-six amino acids (Lipinska, B. *et al.*, 1990). The resulting protein was proposed to comprised of three main domains (Figure 1.1).

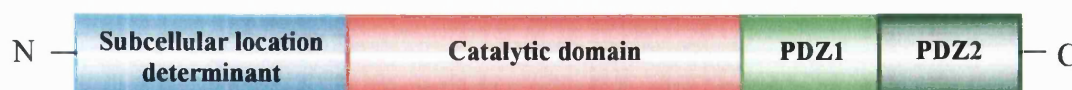


Figure 1.1 - Diagram of important domains of HtrA in Gram-negative bacteria

The N-terminal domain, which differs between homologues, is thought to determine the subcellular location of the enzyme. The catalytic domain contains the serine protease activity that cleaves substrates. The C-terminal PDZ domains (discussed later in 1.5.2) are thought to be important in specific recognition of substrates.

1.5.1 Catalytic domain

HtrA was first identified as a protease by its ability to cleave β -casein (Swarmy, K.H.S. *et al.*, 1983). The enzymatic activity of the enzyme is active over a pH range from 4.8 to 10 (Swarmy, K.H.S. *et al.*, 1983, Lipinska, B. *et al.*, 1990). Degradation of β -casein was independent of ATP and divalent cations, such as Mg^{2+} , Mn^{2+} and Zn^{2+} . The proteolytic activity was further characterised, which resulted in several lines of evidence suggesting that HtrA is a trypsin-like serine protease. Firstly the catalytic domain contains the conserved catalytic triad seen in serine proteases (Smeds, A. *et al.*, 1998). In *E. coli*, these residues are S^{210} , H^{105} and D^{135} , which presumably act in a mechanistic manner known for trypsin cleavage (mechanism, appendix A.1). Secondly, the conserved serine residue is located in the conserved GNS²¹⁰GGAL sequence, similar to the consensus sequence GDSGGPK in trypsin-like serine protease active sites (Skórko-Glonek, J. *et al.*, 1995). Thirdly, inhibition studies using the serine protease inhibitor, di-isopropylfluorophosphate (DFP), resulted in the loss of protease activity (Swarmy, K.H.S. *et al.*, 1983, Skórko-

Glonek, J. *et al.*, 1995). Inhibitors of metalloproteases, aspartic acid proteases and cysteine proteases, however, did not affect proteolytic activity of HtrA (Lipinska, B. *et al.*, 1990).

Site-directed mutagenesis of the residues of the active site resulted in the loss of proteolytic activity by mutating $S^{210} \rightarrow A$ or $H^{105} \rightarrow R$. Analysis of these mutant proteins by Fourier transform infra-red spectroscopy revealed that the proteins were essentially identical in structure to the wild-type HtrA, however, plasmids containing these mutants could not relieve the thermosensitive phenotype of a *htrA* mutant strain (Skórko-Glonek, J. *et al.*, 1995). Thus, the proteolytic activity is directly linked to the thermosensitive phenotype seen in *htrA* mutants.

1.5.2 PDZ domains

PDZ domains are named after the proteins, postsynaptic density-95 (PSD-95), Disc Large (Dlg) and Zonula occludentes-1 (Zo-1) eukaryotic proteins in which the domains were first identified. These domains are involved in protein-protein interactions and receptor clustering (reviewed in Pallen, M.J. & Ponting, C.P., 1997). For instance, in the mammalian central nervous system PSD-95 acts as a scaffold protein via its three PDZ domains, which can recognise and bind to C-terminal sequences. PDZ1 and PDZ2 of PSD-95 can bind N-methyl-D-aspartate (NMDA) receptors (Hu, L.A. *et al.*, 2000). In addition, it is thought that neuronal nitric oxide synthase (nNOS), which is activated by the Ca^{2+} -permeable channels such as NMDA receptors, can also bind to PDZ2 of PSD-95 (Tochio, H. *et al.*, 1999). Thus, PSD-95 colocalises these two receptors in close proximity to each other in order to allow efficient signal transduction. The PDZ domain of nNOS can further recruit other

proteins such as CAPON, a cytosolic protein with a phosphotyrosine-binding domain, or α -syntrophin (Tochio, H. *et al.*, 1999). The PDZ domains of PSD-95 are also important in long-term potentiation (LTP), which is an enhancement of synaptic responses that underlie some forms of memory (Malenka, R.C. & Nicoll, R.A., 1999). NMDA receptors and β_1 -adrenergic receptors are essential for the development of LTP (Malenka, R.C. & Nicoll, R.A., 1999). In 2000, Hu *et al.* demonstrated that the PDZ3 domain of PSD-95 could bind β_1 -adrenergic receptors (Hu, L.A. *et al.*, 2000). As NMDA receptors could bind to PSD-95, these two receptors are also colocalised by their interactions with PDZ domains of one protein. Thus, PDZ domains in the eukaryotic system act primarily as scaffold proteins by recruiting numerous proteins to a localised area and in doing so increases the efficiency of a number of cellular processes such as signal transduction.

In bacteria, the PDZ domains have been shown to be associated with proteolysis (reviewed in Pallen, M.J. & Ponting, C.P., 1997). An example of this is the periplasmic protease, Prc, also known as Tsp. Prc has a role during bacterial exposure to heat-shock and osmotic shock conditions as the *prc* mutant is more sensitive to these conditions compared to the wild-type (Hara, H. *et al.*, 1991). During these stress conditions, Prc can act as a chaperone as it can recognise hydrophobic core residues of a denatured protein by its hydrophobic binding pocket (Bass, S. *et al.*, 1996). In addition, Prc has a protease activity which processes the penicillin-binding protein (PBP-3) by removing eleven amino acids at the C-terminus of the protein. Thus, Prc can behave both as a chaperone and as a protease (Bass, S. *et al.*, 1996). Prc recognises its substrate by means of its PDZ domain,

which is situated in the middle of the protein in close proximity to the active site (reviewed in Pallen, M.J. & Ponting, C.P., 1997).

Due to the ability of PDZ domains to bind to C-terminal protein motifs, the PDZ domains in HtrA are proposed to be important for the specific recognition of substrate. In 1999, Sassoon *et al.* showed that the PDZ domains of HtrA were involved in oligomerisation of HtrA hexamers. Smaller complexes, believed to be dimers or trimers were observed in HtrA mutants without the PDZ domains, whereas a larger complex, suggesting a hexameric form, was observed for the wild-type HtrA. The loss of the PDZ domains also resulted in the loss of proteolytic activity, supporting the change in the arrangement of the oligomeric structure. As Gram-negative bacterial HtrA has been identified with two PDZ domains, it is possible that one domain is involved in oligomerisation while the other is involved in substrate recognition (Sassoon, N. *et al.*, 1999).

1.6 The crystal structure of HtrA

The crystal structure of *E. coli* HtrA was recently resolved (Krojer, T. *et al.*, 2002). Figure 1.2 shows the crystal structure of a monomer of *E. coli* HtrA. The catalytic domain (green) is folded in a manner similar to other trypsin enzymes, where two β -barrel lobes with a carboxy-terminal helix lie perpendicular to each other. The catalytic triad of serine proteases (stick and ball) is located between the two lobes. The structure of the catalytic domain HtrA is most similar to trypsin and thrombin. Although the core of the domain is highly conserved, differences are observed in the surface loops. These differences appear to be important in adjusting the arrangement of the catalytic triad and also in the specificity pocket, S1. Interestingly, what was

originally thought to be the N-terminal subcellular location domain actually forms an α -helix (α A) followed by the β -strand, β 1, which interacts with β 2 and β 3 of the protease domain. The loop connecting β 1 of the N-terminal domain and β 2 of the protease domain is an unusually enlarged loop (LA), and interacts with loop 1 (L1) and loop 2 (L2) in the active site of its partner subunit. This resulting interaction completely removes any access to the active site. To obtain the typical catalytic arrangement of serine protease, the loops must undergo a large conformational change.

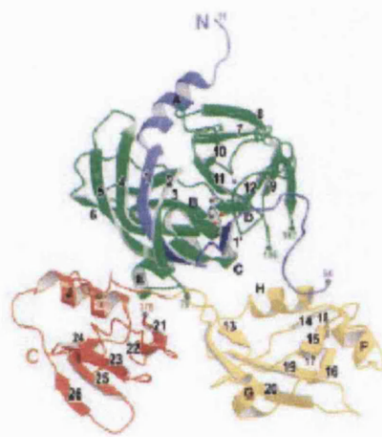


Figure 1.2 – Crystal structure of a monomer of *E. coli* HtrA (taken from Krojer, T. *et al.*, 2002)

The N-terminal domain is coloured blue. The catalytic domain is coloured green. PDZ1 is coloured yellow and PDZ2 is coloured red. Helices are labelled with letters and β -strands are denoted by numbers.

The PDZ domains of HtrA are folded in a similar manner to other PDZ domains. The secondary structure is well conserved in PDZ domain family members even though there is a limited amino acid sequence identity. PDZ domains consist of six β -sheets (β 1 to β 6) and two α -helices (α A and α B). However, PDZ1 of HtrA

contains an extra β -sheet (denoted β_{13} in Figure 1.2) and α -helix (denoted α_F in Figure 1.2), which are important for monomeric interactions in the hexamer.

The crystal structure showed that HtrA monomers associated as trimers, which stacked on top of each other in a staggered manner, with a lateral cavity located between the trimers (Figure 1.3).

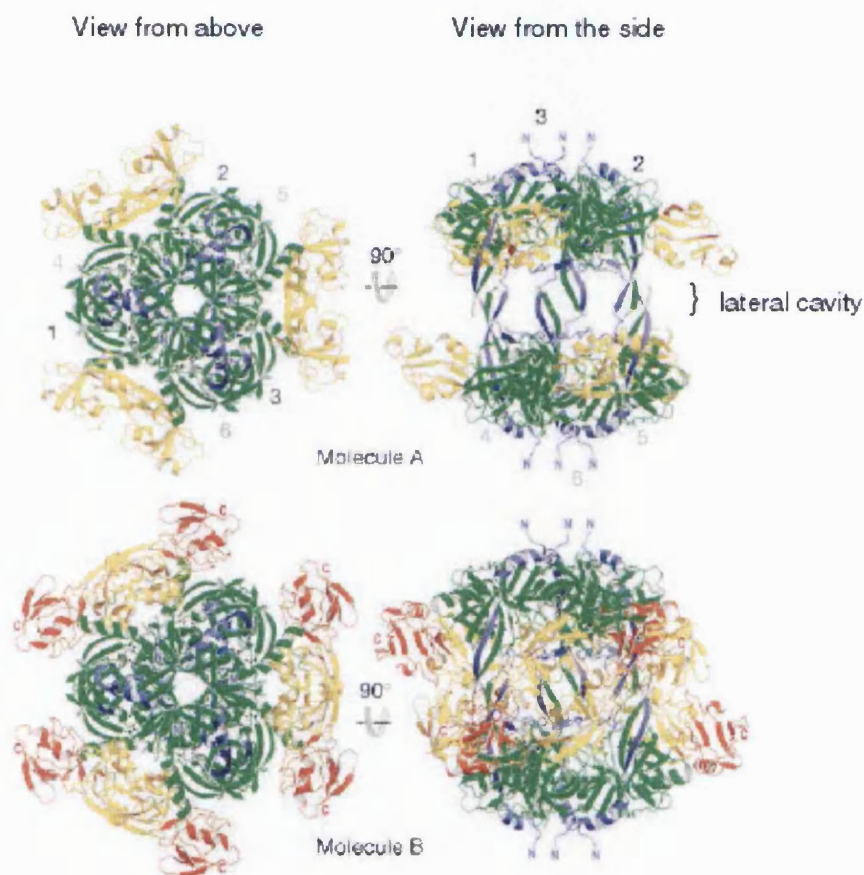


Figure 1.3 – The crystal structure of the *E. coli* HtrA oligomer (taken from Krojer, T. *et al.*, 2002)

The top and side views of molecules A and B are shown. The ‘N-terminal’ subcellular location determinant is coloured blue. The catalytic domain is coloured green. PDZ1 is coloured yellow and PDZ2 is coloured red. Helices are labelled with letters and β -strands are denoted by numbers.

The interactions between the subunits are shown to be via the LA loop or by PDZ interactions. For example, the interaction between subunit 1 and subunit 4 is determined by LA. This is because of the ability of this loop to associate with the active site in its partner subunit. Interactions between subunit 1 and subunit 6 are via the PDZ domains. PDZ1 domains face each other by a α F- β 15- α G interaction while the PDZ2 domains are joined by a β 22- α I interaction.

The LA loop also dictates the height of the lateral cavity, through its interaction with the subunit beneath/above the LA loop. This cavity is lined with the proteolytic sites of the catalytic domains. The proteolytic sites are accessible in the open form of the HtrA, molecule A, where removal of the PDZ1-PDZ1 interactions results in a 70° tilt away from the lateral cavity and a 30Å movement. Restoring the PDZ1-PDZ1 interactions block off access to the lateral cavity resulting in the second HtrA form, molecule B, which is a closed form.

1.7 Substrate recognition by PDZ domains

Proteases often recognise substrates in a distinct manner to their recognition of the cleavage site (reviewed in Pallen, M.J. & Wren, B.W., 1997). HtrA may be able to recognise substrates through PDZ domain interactions. This is an attractive idea as PDZ domains are shown to be involved in protein-protein interactions (reviewed in Pallen, M.J. & Ponting, C.P., 1997) being able to form homodimers as well as being able to recognise both internal and C-terminal sequences (Herman, C. *et al.*, 1998, Gee, S.H. *et al.*, 2000). As discussed earlier in 1.5.2, an example of this is the ability

of the PDZ domain of PSD-95 to bind to the C-terminal sequence of NMDA receptors via the C-terminal S/-X-V sequence (Songyang, Z. *et al.*, 1997). PSD-95 can also bind the PDZ domain of neuronal nitric oxide synthase (nNOS) and thus cluster nNOS close to the NMDA receptors in the mammalian nervous system (Doyle, D.A. *et al.*, 1996).

Understanding the mode of recognition in PDZ domains have been aided by the availability of crystal structures of both native and ligand-bound PDZ domains (Doyle, D.A. *et al.*, 1996). To date, three-dimensional structures have been resolved for several PDZ domains, which are listed below in

Table 1.2.

Protein	PDZ domain	Reference
PSD-95	Second	Tochio, H. <i>et al.</i> , 2000
	Third	Doyle, D.A. <i>et al.</i> , 1996
nNOS		Tochio, H. <i>et al.</i> , 1999,
		Mandal, P.K., 2002,
		Oschkinat, H., 1999
hCASK		Daniels, D.L. <i>et al.</i> , 1998
PTP-BL	Second	Walma, T. <i>et al.</i> , 2002
SAP90	First	Piserchio, A. <i>et al.</i> , 2002
Human Disc Large	Third	Cabral, J.H.M. <i>et al.</i> , 1996
CtpA		Yamamoto, Y. <i>et al.</i> , 2001

Table 1.2 – Published crystal structures of PDZ domains

As discussed in 1.6, the secondary structure of PDZ domains is well conserved consisting of six β -sheets (β A- β F) and two α -helices (α A and α B) (Songyang, Z. *et al.*, 1997), which form a β sandwich (reviewed in Pallen, M.J. & Wren, B.W., 1997). In PSD-95, the peptide ligand of the third PDZ domain lies within a groove made by the β B and α B helix in an antiparallel fashion (Figure 1.4) (Doyle, D.A. *et al.*, 1996). This interaction between the PDZ and ligand explains the slight differences in PDZ specificity for different consensus sequences (discussed later 1.7.2).

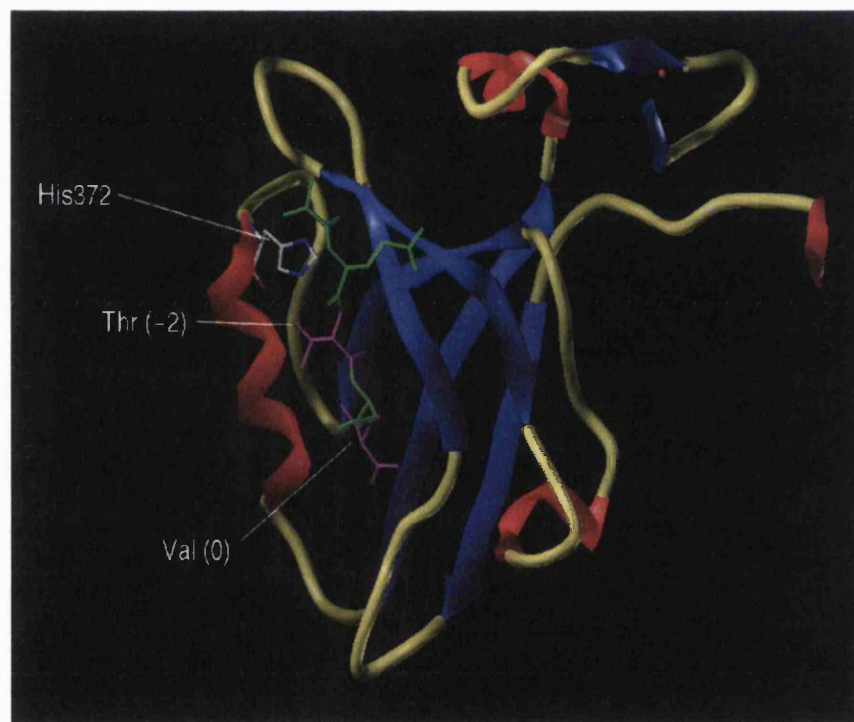


Figure 1.4 - Crystal structure of the third PDZ domain of PSD-95 (amino acids 302-402) bound to the peptide KQTSV

The protein was modelled using Sybyl on the Silicon Graphics computer. (β -strands – blue, α -helices - red, interlinking loops - yellow)

1.7.1 Carboxylate-binding loop

The carboxylate-binding loop consists of the highly conserved GLGF consensus sequence of PDZ domains (Songyang, Z. *et al.*, 1997). However, recently this consensus motif has been refined to $\Phi G \Phi$, where Φ denoted hydrophobic residues (Tochio, H. *et al.*, 2000). For instance, in the carboxyl terminal processing protease (CtpA) of the precursor D1 protein in chloroplast, the carboxylate-binding loop comprises of GVGF (Yamamoto, Y. *et al.*, 2001). The amide groups of the residues which make up the carboxylate-binding loop interact via hydrogen-bonding to the free carboxyl group of the peptide (Doyle, D.A. *et al.*, 1996). In addition, these hydrophobic residues interact with the terminal ligand residue (Doyle, D.A. *et al.*, 1996). Other variations between PDZ domains are the size of this pocket in which the C-terminal residue nestles. For example, in the PDZ3 of PSD-95, has a surface area of $110.69 \text{ \AA}^2$, which can fit a small uncharged residue, such as valine. However, in the PDZ of the $\text{Na}^+ \text{-H}^+$ exchanger regulatory factor (NHERF), specific binding was observed when the C-terminal residue was the larger leucine residue, which would fit better in the larger pocket, calculated to be $183.43 \text{ \AA}^2$ (Karthikeyan, S. *et al.*, 2001).

1.7.2 Substrate specificity conferred by αB helix

Different PDZ domains prefer different consensus sequences. For instance, PSD-95 recognises sequences with the motif S/T-X-V/A/F (Songyang, Z. *et al.*, 1997, Fuh, G. *et al.*, 2000), whilst the PDZ domain of p55 prefers the sequence F/Y-X-A/V/F (Songyang, Z. *et al.*, 1997) This has resulted in separating the PDZ domains into different classes by the differences in the first residue of the αB helix (αB_1), which

forms the peptide-binding groove with $\beta 2$. This residue interacts with the residue at the -2 position on the ligand (where 0 denotes the C-terminal residue).

Class I PDZ domains include PSD-95, the third and fifth PDZ domains of phosphotyrosine phosphatase and the three domains of murine Disc Large (Songyang, Z. *et al.*, 1997). These all selectively recognise the S/T-X-V sequence. Here the -2 residue is hydrophilic (S/T) which in PSD-95 interacts with H³⁷² at αB_1 (Doyle, D.A. *et al.*, 1996).

Class II PDZ domains prefer the F/Y-X-A/V/F sequence and include p55, human LIN-2 and AF-6 (Songyang, Z. *et al.*, 1997). The -2 residue of the ligand is a bulky hydrophobic residue (F/Y) which is selected due to the absence of the basic His residue on the αB helix (Songyang, Z. *et al.*, 1997). In p55 and Lin-2, the αB_1 residue is valine which selects for phenylalanine over tyrosine at the -2 position of the ligand (Songyang, Z. *et al.*, 1997).

Few Class III PDZ domains have been identified to date. The PDZ domain of nNOS is a class III member and recognises a negatively charged residue at the -2 position on the ligand, G-E/D-X-V (Tochio, H. *et al.*, 1999, Bezprozvanny, I. & Maximov, A., 2001).

Thus, the interaction between the αB_1 residue of the PDZ domain and the -2 position residue of the ligand is the important factor in the differing specificities of different PDZ domains.

1.7.3 Proposed substrate recognition by the PDZ domains of HtrA

As discussed in 1.6, the PDZ domains of HtrA are essentially folded in a similar manner to that of other PDZ domains, except for the extra β -strand and α helix in PDZ1 (Krojer, T. *et al.*, 2002). In order to predict the ligand binding sites, the PDZ domains of HtrA were aligned with the crystal structure of the third PDZ domain of PSD-95, which had been resolved in the native form and bound to a short peptide ligand (Doyle, D.A. *et al.*, 1996).

1.7.3.1 Substrate recognition of PDZ1

Mapping PDZ1 to the third PDZ domain of PSD-95 showed that the ligand-binding groove is located between β 14 and α H. The groove is much deeper than that seen in other PDZ domains. The carboxylate-binding loop is composed of the residues ELGI, which agrees with the $\Phi G \Phi$ motif, discussed earlier (1.7.1). Another residue, R²⁶² also interacts and fixes the carboxyl group at the C-terminus of a peptide. This residue is highly conserved as it is involved in the tilt of PDZ1 away from the lateral cavity (discussed earlier in 1.6). Thus, it is proposed that on binding a substrate, R²⁶² mediates the tilt of PDZ1 away from the lateral cavity, hence, promoting proteolysis (Krojer, T. *et al.*, 2002). Substrate specificity is conferred by interactions with the substrate at 0 and -2 positions from the C-terminus (1.7.2). Residues, thought to be involved in such interactions in PDZ1 of HtrA, were all hydrophobic (Krojer, T. *et al.*, 2002). Unlike, other PDZ domains, the ligand binding cleft is far more adaptable for a variety of substrates as it has a flexible nature, namely in β 14 and the carboxylate loop.

1.7.3.2 Substrate recognition of PDZ2

The ligand-binding groove in PDZ2 is comprised of β 21 and α J. The carboxylate-binding consists of the residues GIEGA, which differs from the $\Phi G \Phi$ motif, discussed earlier in 1.7. Unlike PDZ1, residues 355 to 369 form a loop that binds to the groove in an anti-parallel manner, mimicking the substrate. The ability to bind an internal sequence via a loop structure is also seen in nNOS (Oschkinat, H., 1999) where it was shown that an additional C-terminal β -sheet in the PDZ domain could form a β -finger. This structure could then be recognised by the PDZ domain of α -syntrophin (Oschkinat, H., 1999). Like PDZ1, the carboxylate-binding loop of PDZ2 is highly flexible, presumably to allow interactions with residues 355 to 369. Thus, it is suggested that the displacement of this fourteen residue loop may be involved in substrate translocation, and hence the first contact between HtrA and its substrate.

1.8 HtrA substrates

1.8.1 HtrA recognition requirement for unfolded substrates

In vitro, HtrA is capable of cleaving β -casein, the N-terminus of λ repressor and several variants of the arc repressor (reviewed in Pallen, M.J. & Wren, B.W., 1997). In addition, the protease can also cleave mislocalised cytoplasmic proteins such as β -galactosidase (Speiss, C. *et al.*, 1999). However, HtrA cannot cleave serum albumin, ovalbumin, growth hormone, or a number of twenty synthetic protease substrates (Kolmar, H. *et al.*, 1996).

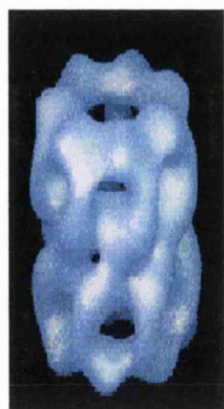
The differences seen here can be explained by the need for the substrate to be in a predominantly unfolded form in order for HtrA to be able to recognise the substrate. In 1999, Kim *et al.* incubated HtrA with citrate synthase, which is properly folded at

35°C and is unfolded at 43°C. At 43°C, citrate synthase was degraded over time whereas at 35°C there was little or no degradation. In a second experiment, bovine α -lactalbumin was used. This protein contains four intramolecular disulphide bonds (Kim, K.I. *et al.*, 1999). Sequential reduction of the disulphide bonds showed that no degradation occurred until all four bonds were reduced and the protein was unfolded.

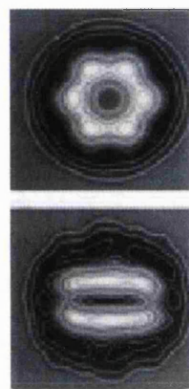
Originally, this requirement for unfolded substrates was proposed to be due to the arrangement of the catalytic sites in the hexameric HtrA complex. Recent studies (Kim, K.I. *et al.*, 1999) showed that the arrangement of the hexamer was reminiscent of that seen in the chaperone protein, GroEL (Figure 1.5a) and also of the cytoplasmic protease, Clp.

Clp enzymes are located in the cytoplasm and consist of proteolytic and ATPase subunits (Gottesman, S. *et al.*, 1998, Smith, C.K. *et al.*, 1999). Six proteolytic subunits arrange themselves into a ring structure. The ATPase-containing subunits can also form hexameric rings and bind both ends of the 'proteolytic stack', thus forming a channel (Gottesman, S. *et al.*, 1998, Smith, C.K. *et al.*, 1999). It is thought that the catalytic sites are arranged so that they line the channel and thus may access unfolded substrates that enter the channel.

Using electron microscopy on *E. coli* HtrA, ring-shaped structures with a six-fold symmetry were seen from above (Figure 1.5bi). A striated structure was seen when the hexamer was viewed from the side (Figure 1.5bii). A comparison with the three-dimensional structure of GroEL is shown in Figure 1.5.



a) GroEL three dimensional structure (taken from <http://bioc09.uthscsa.edu/~seale/Chap/struc.html>)



i) View from HtrA above the dodecamer
ii) Sideview of the HtrA dodecamer

b) Electron micrograph of HtrA (taken from Kim, K.I. *et al.*, 1999)

Figure 1.5 - Comparison of the multimeric structures of the chaperone protein GroEL and the periplasmic protease HtrA

The diameter of the opening is 3 nm, an intermediate between the size of the chaperone GroEL (4.5 nm) and that of the Clp protease (1.0 nm). This suggests that the channel opening is large enough for partially unfolded substrates such as β -casein (Kim, K.I. *et al.*, 1999).

Although this model was an attractive idea, the crystal structure of *E. coli* HtrA, discussed previously in 1.6, shows that unlike the Clp enzymes, HtrA has a lateral cavity, which is lined with the proteolytic sites (Krojer, T. *et al.*, 2002). The height of this cavity varies between 15Å at the centre of the cavity to 18Å near the outer edges. Due to these constrictions, protein substrates must be partially unfolded in order to reach the active sites as only secondary structures can access the cavity (Figure 1.6).

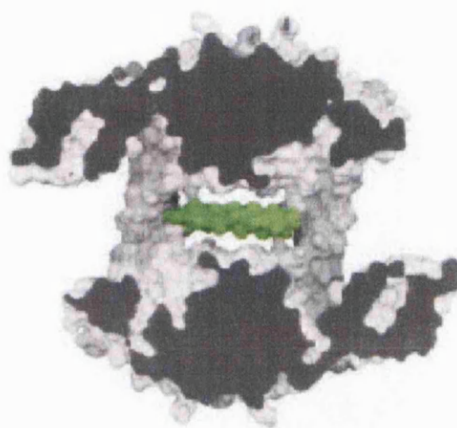


Figure 1.6 – The lateral cavity of *E. coli* HtrA (taken from Krojer, T. *et al.*, 2002)

Cleaved representations of molecule A (discussed in 1.6) were constructed. The cut regions are in dark grey. Surface representation of the lateral cavity with a polyalanine helix modelled to fit within the cavity, seen from the side.

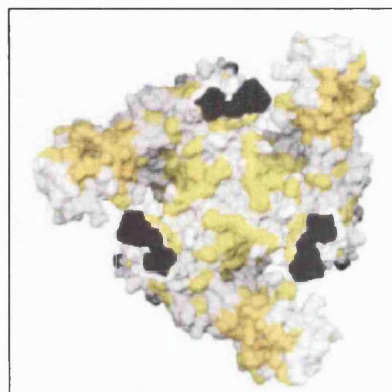
1.8.2 HtrA recognition requirement for a hydrophobic core

As discussed earlier (1.4), HtrA can function to remove misfolded proteins in the periplasm by degradation or renaturation of the protein. The recognition of substrates in both cases are presumed to involve recognising hydrophobic regions of a protein substrate, which are normally locked away from the solvent environment in the core of the protein. In 1999, Kim *et al.* suggested that HtrA recognised a hydrophobic core because it could not degrade small hydrophilic peptides such as glucagons or the insulin A-strand. In addition, it was shown that incubating HtrA with an hydrophobic probe (ANS) abolished β -casein degradation by the enzyme (Kim, K.I. *et al.*, 1999). This hypothesis is supported by the HtrA crystal structure (Krojer, T. *et al.*, 2002). On analysis of the lateral site, it could be seen that there were three hydrophobic grooves surrounding the central pore (Figure 1.7). These grooves

consist of the loops LA and L2 of the protein and extend towards the PDZ1 domain, where the initial substrate interaction is thought to occur (Krojer, T. *et al.*, 2002).

Figure 1.7 – Hydrophobic patches within the lateral cavity of *E. coli* HtrA (taken from Krojer, T. *et al.*, 2002)

Hydrophobic residues of the catalytic domain are coloured yellow, while the uncharged residues of the PDZ1 domain are coloured orange.



1.8.3 *In vivo* HtrA ligands

Although there are a number of substrates used to characterise HtrA *in vitro*, few natural *in vivo* ligands have been identified. The first native ligand of HtrA was shown to be the acylated precursor of colicin A lysis (Cal) protein in 1989 (Cavard, D. *et al.*, 1989). This protein is a small lipoprotein essential for the secretion of colicin A in *E. coli*. Pulse-chase experiments analysing the processing of the Cal precursor were carried out in wild-type *E. coli* and in a *htrA* mutant strain. It was shown that after thirty minutes the acylated precursor form was being processed to the mature form in the wild-type strain. However, in the mutant strain, levels of the Cal precursor increased with time (Cavard, D. *et al.*, 1989). Thus, it was concluded that HtrA was involved in processing the precursor protein. In the following years, it was shown that K88 and K99 pilin subunits (Bakker, D. *et al.*, 1991) and the HMW1 and HMW2 adhesins of *H.influenzae* (St Geme, J.W., II & Grass, S., 1998) were HtrA ligands. Recent studies have shown that the PDZ domains of HtrA can

recognise the Small Stable RNA A (*ssrA*)-encoded peptide tag (discussed later in 1.8.3.1).

1.8.3.1 The *SsrA*-tagging system.

The SsrA-tag is a short C-terminal eleven-residue motif, which marks a protein for degradation (Tu, G.-F. *et al.*, 1995) by cytoplasmic proteases, such as ClpAP, ClpXP (Gottesman, S. *et al.*, 1998) and HflB (Herman, C. *et al.*, 1998). It is also recognised by the PDZ domain of the periplasmic protease, Prc (Keiler, K.C. *et al.*, 1996) and recently been implicated to interact with the PDZ domain of *E. coli* HtrA (Spiers, A. *et al.*, 2002). The importance of the SsrA motif is highlighted by its presence all bacteria and plastid genomes (<http://www.indiana.edu/~tmrna/>) including *B. subtilis* (Muto, A. *et al.*, 2000, Wiegert, T. & Schumann, W., 2001), *Streptococcus pyogenes* (Steiner, K. & Malke, H., 2001) and *Caulobacter crescentus* (Keiler, K.C. *et al.*, 2000). SsrA mutants have a number of phenotypes, including temperature sensitivity (Muto, A. *et al.*, 2000) and reduced motility (Karzai, A.W. *et al.*, 1999). Mutants of *S. typhimurium* displayed a 200-fold decrease in virulence (Julio, S.M. *et al.*, 2000), although the bacteria were still viable. However, in *Neisseria gonorrhoeae*, mutations in *ssrA* are lethal (Huang, C. *et al.*, 2000).

The SsrA-tag is part of an elegant mechanism by which a stalled ribosome is liberated and the resulting partially formed protein is tagged with an eleven-residue motif, which targets the protein for proteolysis (summarised in Figure 1.8).

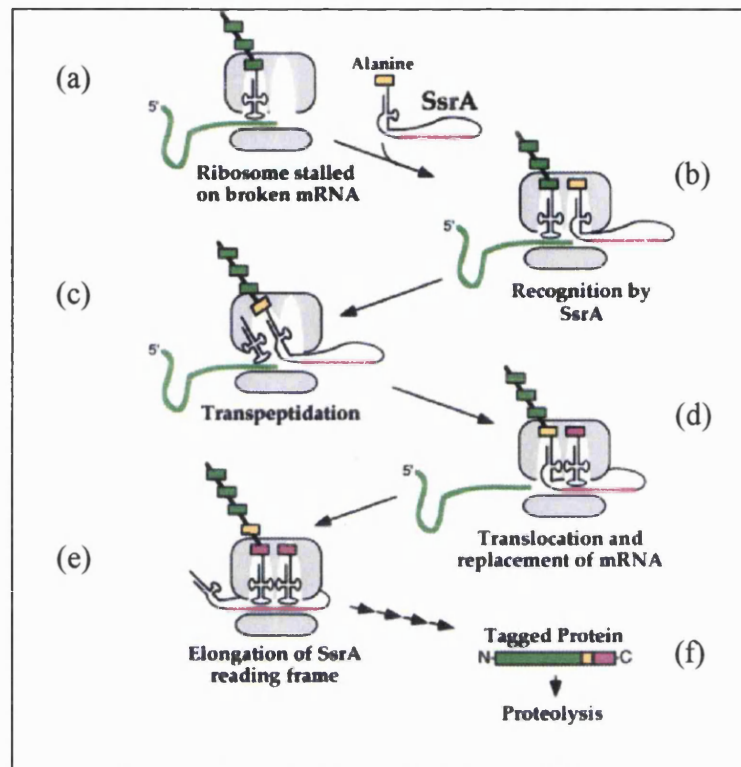


Figure 1.8 – Model for SsrA-mediated tagging of proteins synthesized from messages lacking stop codons (taken from Roche, E.D. & Sauer, R.T., 1999).

The SsrA-tag is encoded by a 363 nucleotide long RNA molecule which has characteristics of both tRNA and mRNA structures (hence called tmRNA) (Gottesman, S. *et al.*, 1998). The tmRNA is formed as a precursor, which is processed by RNase P in order to remove the seven nucleotides at the 5'-end of the structure (Komine, Y. *et al.*, 1994). The RNA is further processed by RNase E to produce the mature tmRNA molecule (Lin-Chao, S. *et al.*, 1999), which is then charged with an alanine residue by alanyl-tRNA synthetase (Komine, Y. *et al.*, 1994).

When a ribosome is stalled (Figure 1.8a), the tmRNA, charged with an alanine residue can enter the ribosome (Figure 1.8b) and polymerise onto the growing peptide chain (Figure 1.8c). The mRNA within the ribosome is released, suggested

to be degraded by RNase R (Karzai, A.W. & Sauer, R.T., 2001, Gillet, R. & Felden, B., 2001), and the tmRNA takes over (Figure 1.8d). The eleven-residue tag is translated from the tmRNA (Figure 1.8e) and the protein is released with the SsrA-tag (Figure 1.8f). This tag can then be recognised and degraded by proteases such as Prc and ClpAP.

The function of SsrA-tagging is to essentially free stalled ribosomes as the proteolysis step has been shown to be unimportant for bacterial viability in *N. gonorrhoeae* (Huang, C. *et al.*, 2000) and in *B. subtilis* (Muto, A. *et al.*, 2000). The stalling of mRNA on a ribosome has detrimental results on the bacteria as it prevents the recycling of ribosomes in production of proteins, which may be required for growth at a certain point in time. In 1999, Roche *et al.* showed that translating proteins with rare codons can stall the ribosome and are tagged with the SsrA motif (Roche, E.D. & Sauer, R.T., 1999). A significant increase for SsrA-tagging was observed in the presence of the rare arginine codon, AGA, which account for 0.26% of the codons used for this amino acid. It was also noted that there are low levels of tRNA^{AGA} in the cell, thus, tRNA^{AGA} was introduced into the bacteria. This resulted in a decrease in SsrA-tagging, whereas reducing the amount of tRNA^{AGA} resulted in an increase in tagging (Roche, E.D. & Sauer, R.T., 1999).

The identification of endogenous protein substrates for SsrA-tagging showed that tagging corresponded to the same positions as stop codons (Roche, E.D. & Sauer, R.T., 2001). Recently, Collier *et al.* suggested that SsrA-tagging of ribokinase occurred because translation was stopped using inefficient stop codons, such as UGA (Collier, J. *et al.*, 2002). Mutating UGA to a strong termination stop codon, UAA,

abolished SsrA-tagging of the protein (Collier, J. *et al.*, 2002). Meanwhile, Hayes *et al.* suggested that the SsrA-tagging of ribokinase was due to a combination of the inefficient stop sequence and two rare codons, AGG, which encoded two arginine residues that preceded the stop codon (Hayes, C. *et al.*, 2002).

Another occasion when ribosomes could be stalled is when bacteria are in the presence of protein synthesis inhibitors, such as chloramphenicol and erythromycin (de la Cruz, J. & Vioque, A., 2001), or in the presence of miscoding antibiotics, such as kanamycin and streptomycin (Abo, T. *et al.*, 2002). In both instances, *ssrA* mutants were more sensitive to the effects of the antibiotics in comparison to the wild-type strains, with less SsrA-tagging occurring. These studies illustrate that the tmRNA is able to recycle the ribosomes and can deal with the translational errors which occur in the presence of low concentrations of antibiotics. Thus, the SsrA-tagging system is important for bacterial viability under a variety of stress conditions.

1.9 Regulation of HtrA

Proteolysis is a powerful property of HtrA, and if uncontrolled would create havoc within a biological system. The presence of a small lateral channel that is only accessed by partially unfolded proteins (discussed in 1.8.1) is one method of controlling this activity of HtrA but other measures also exist. For example, the expression of *E. coli* HtrA is controlled by four systems: the two-component regulatory system CpxRA, the two-phosphoprotein phosphatases, PrpA and PrpB, and the RpoE regulon, which in turn is regulated by the RseA anti- σ factor (reviewed in Pallen, M.J. & Wren, B.W., 1997).

1.9.1 RpoE regulon

Periplasmic stress conditions, such as an accumulation of misfolded proteins in this compartment induce the *rpoE* regulon (Mecsas, J. *et al.*, 1993). The *rpoE* regulon, an extracytoplasmic (ECF) σ factor, is also induced under the heat-shock stress conditions (Rouvière, P.E. *et al.*, 1995) and under hyperosmotic shock conditions (Bianchi, A.A. & Baneyx, F., 1999). The majority of *htrA* transcripts are controlled by the *rpoE* regulon as mutations in σ^E result in a significant decrease in *htrA* transcription (Raina, S. *et al.*, 1995, Resto-Ruiz, S.I. *et al.*, 2000).

The *rpoE* regulon consists of the genes *rpoE*, *rpoH*, *fkpA*, *ompK* and *htrA* (Erickson, J.W. & Gross, C.A., 1989, Wang, R.F. & Kaguni, J.M., 1989, Raina, S. *et al.*, 1995, Danese, P.N. & Silhavy, T.J., 1997). The transcription of its own gene product establishes a positive feedback mechanism (Rouvière, P.E. *et al.*, 1995). In the absence of induction the levels of σ^E are low indicating the presence of a negative feedback mechanism (Raina, S. *et al.*, 1995, Rouvière, P.E. *et al.*, 1995).

1.9.1.1 Negative regulation by the anti- σ factors RseABC (regulator of sigma E)

The σ^E factor is controlled by the anti- σ factors RseA, RseB and RseC (De Las Peñas, A. *et al.*, 1997, Missiakas, D. *et al.*, 1997, Ades, S.E. *et al.*, 1999), which are encoded downstream of *rpoE* (De Las Peñas, A. *et al.*, 1997, Missiakas, D. *et al.*, 1997). RseA and RseB act on *rpoE* in a negative manner as mutations in *rseA* or *rseB* result in an increase in σ^E activity (De Las Peñas, A. *et al.*, 1997, Missiakas, D. *et al.*, 1997). When *rseA* or *rseB* was overproduced separately the basal levels of σ^E activity decreased by 80 to 90% and 30 to 40%, respectively, thus illustrating the

relative effects of the gene products (Missiakas, D. *et al.*, 1997). Mutants of *rseC* could not be obtained, however, overproduction of this gene increased σ^E activity, thus suggesting that RseC is a positive modulator of σ^E (Missiakas, D. *et al.*, 1997).

RseA is an inner membrane protein with a N-terminal cytoplasmic domain and a large coiled-coil periplasmic domain (De Las Peñas, A. *et al.*, 1997, Missiakas, D. *et al.*, 1997). In the absence of induction, the periplasmic domain of RseA binds σ^E , thus sequestering it (De Las Peñas, A. *et al.*, 1997, Missiakas, D. *et al.*, 1997). RseB exerts an lesser negative effect on σ^E activity (De Las Peñas, A. *et al.*, 1997, Missiakas, D. *et al.*, 1997), as it is thought to sense the accumulation of misfolded periplasmic proteins and then transmits a signal to RseA via a conformational change (Missiakas, D. *et al.*, 1997). This induces RseA to release σ^E in order to transcribe the *rpoE* regulon (Missiakas, D. *et al.*, 1997). The role of RseC as a positive modulator could be mediated by its known interactions with RseA and RseB (Missiakas, D. *et al.*, 1997), however, the mechanism of this modulation is unknown.

RseA, itself is also regulated by a serine protease, HhoB, a paralogue of HtrA (discussed later in 1.10.2). HhoB is smaller than HtrA and diverges from HtrA at the N-terminus. It also only contains one PDZ domain and interestingly has been shown not to be important for bacterial survival in heat shock (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). However, HhoB contains the conserved residues found in the serine protease catalytic triad and thus is proposed to have proteolytic activity (Waller, P.R.H. & Sauer, R.T., 1996). A recent study showed that deletion of the *hhoB* gene resulted in a five-fold decrease in basal σ^E activity. In addition, the induction of σ^E activity by the overexpression of the outer membrane protein C

(OmpC) was also impaired. However, the addition of wild-type *hhoB* restored σ^E activity (Ades, S.E. *et al.*, 1999). Thus, the controlled proteolytic activity of HhoB regulates the levels of RseA (Ades, S.E. *et al.*, 1999).

The RseABC: σ^E mechanism is relatively conserved in evolution as homologues have been found in other Gram-negative bacteria with similar mechanisms (Pallen, M.J. & Wren, B.W., 1997). The *Pseudomonas aeruginosa* σ^E homologue, AlgU, is negatively regulated by MucA, a Rse homologue (Ades, S.E. *et al.*, 1999). In addition the protein, MucB, is also involved in negative regulation of AlgU (Schurr, M.J. *et al.*, 1996, Erickson, J.W. & Gross, C.A., 1989).

1.9.1.2 The two-component Cpx system

The Cpx system is a typical two-component system, which responds to cell envelope stresses (Pogliano, J. *et al.*, 1997). CpxA is a sensor kinase/phosphatase embedded in the inner membrane. On activation, CpxA phosphorylates CpxR, the response regulator, which in turn activates a cognate DNA-binding regulating protein by transfer of the phosphate group (Danese, P.N. *et al.*, 1995, Raina, S. *et al.*, 1995).

The *cpxA/cpxR* regulon regulates the expression of genes with σ^{70} , σ^{32} and σ^E promoters and in doing so controls the expression of *htrA* (Raina, S. *et al.*, 1995, Pogliano, J. *et al.*, 1997, Dartigalongue, C. & Raina, S., 1998). Null mutants of either *cpxA* or *cpxR* support this, as they result in complete abolishment of *htrA* transcription (Danese, P.N. *et al.*, 1995, Laskowska, E. *et al.*, 1996).

1.9.1.3 *HtrA* regulation by the phosphoprotein phosphatases, *PrpA* and *PrpB*

The phosphoprotein phosphatases, PrpA and PrpB, are important in modulating the phosphorylation status of the CpxA/R system (reviewed in Pallen, M.J. & Wren, B.W., 1997) and by this regulate the activity of HtrA. Studies mutating *prpA/B* result in a decrease in *htrA* transcription (Missiakas, D. & Raina, S., 1997). However when PrpA/B are overexpressed there is a four- to six-fold increase in *htrA* transcription (Missiakas, D. & Raina, S., 1997).

A summary of the transcriptional regulation of *htrA* is shown in Figure 1.9.

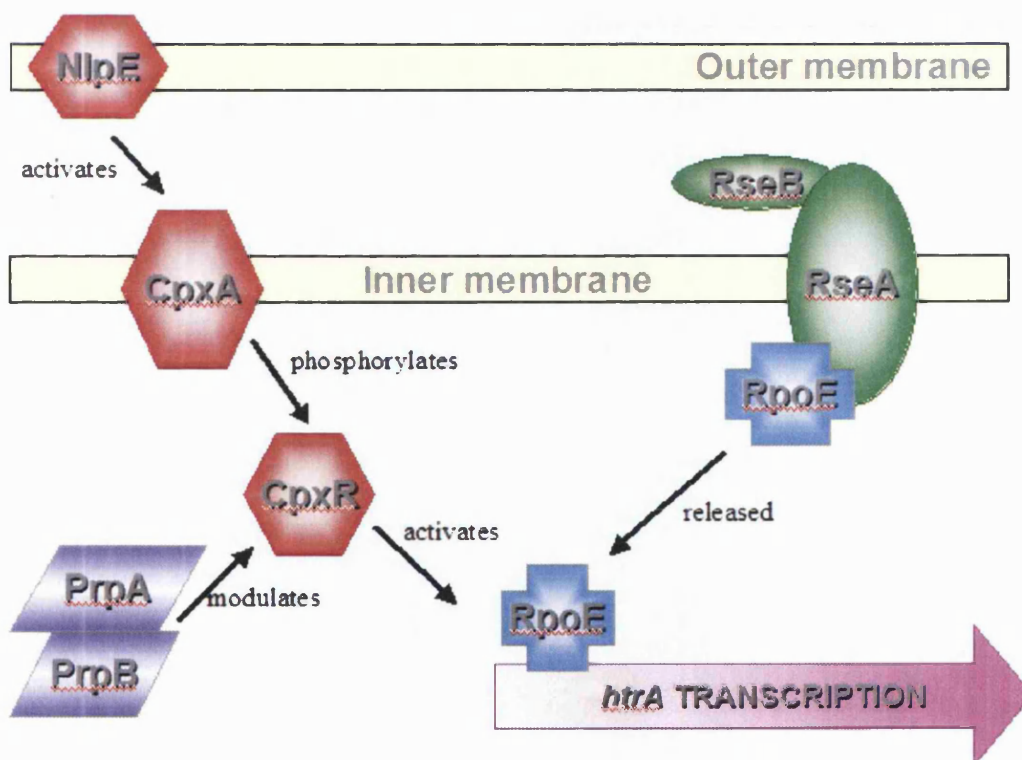


Figure 1.9 – Schematic diagram of transcriptional regulation of *htrA*

1.10 Paralogues of HtrA

Within the *E. coli* genome, there are two paralogues of *htrA*. Paralogues are homologous genes, which evolved from a common ancestor probably by gene duplication, and exist within the same genome. In 1996, two independent groups identified and characterised the HtrA paralogues, HtrA homologue A (HhoA, DegQ) and HhoB, also known as DegS. These two genes lie adjacent on the *E. coli* genome at a large distance to *htrA*. They are separated by 89 base pairs, which can potentially form a stable stem-loop structure (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). The two genes are transcribed in the same direction as separate transcripts of approximately 1.5-kb and 1.1-kb, respectively (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Unlike *htrA*, the two paralogues were not heat-inducible (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Supporting this was the observation that the promoter sequences, located upstream of the two genes, had characteristics with σ^{70} -dependent promoters (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996).

1.10.1 **HhoA**

HhoA consists of 455 amino acids and has 60% identity and 75% similarity to HtrA at the amino acid level (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Unlike HtrA, however, there is no proline/serine/glycine-rich region (positions 80-100 in HtrA) in HhoA (Waller, P.R.H. & Sauer, R.T., 1996). However, there are similarities with HtrA, as both are periplasmic proteins containing an N-terminal signal peptide, a serine protease catalytic domain and two PDZ domains (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). In addition, the substrate profiles of HhoA are indistinguishable (Pallen, M.J. & Wren, B.W., 1997).

It was shown that recombinant HhoA could degrade β -casein and thus was proteolytically active. This activity was inhibited with the serine protease inhibitor, DFP (Waller, P.R.H. & Sauer, R.T., 1996). Mutants of *hhoA* grew *in vitro* in a similar manner to that of the wild-type strain and exhibited no thermosensitivity (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). The thermosensitive phenotype displayed by *htrA* mutants was rescued by overexpression of HhoA as it has proteolytic activity and has similar *in vitro* substrates to HtrA (Waller, P.R.H. & Sauer, R.T., 1996, Kolmar, H. *et al.*, 1996). HhoA could also rescue the thermosensitive phenotype of Prc (Waller, P.R.H. & Sauer, R.T., 1996) indicating that HhoA could also recognise Prc substrates in the periplasm.

1.10.2 HhoB

HhoB is a smaller protein than either HhoA or HtrA. It consists of 355 amino acids and has 27% identity and 58% similarity to HtrA at the amino acid level (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Like HtrA and HhoA, HhoB has a catalytic domain containing the serine protease catalytic triad, however, HhoB has only a single PDZ domain (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). The N-terminus of HhoB, which diverges the most out of the three paralogues, does not appear to contain a consensus signal peptide cleavage site (Waller, P.R.H. & Sauer, R.T., 1996).

Recombinant HhoB was not shown to be proteolytically active (Waller, P.R.H. & Sauer, R.T., 1996), however, as discussed earlier in 1.9.1.1, HhoB is implicated in the degradation of the anti- σ factor, RseA (Ades, S.E. *et al.*, 1999, Alba, B.M. *et al.*,

2001). Mutants of this gene were not thermosensitive, however, they showed a small-colony phenotype (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Thus, HhoB is implicated in normal *E. coli* growth.

A schematic representation of the three *E. coli* HtrA paralogues is show below in Figure 1.10.

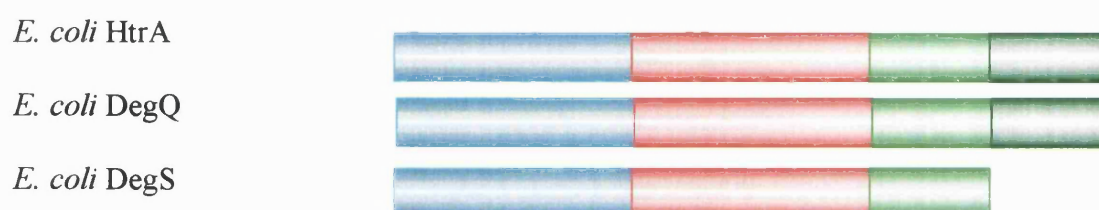


Figure 1.10– Schematic diagram of the protein domains of the HtrA paralogues in *E. coli*

The N-terminal domain is coloured cyan, the catalytic domain is coloured red and the PDZ domains are coloured green (PDZ1 – light green; PDZ2 – dark green).

1.11 Identification of HtrA-like homologues in Gram positive bacteria

Until very recently, HtrA homologues were only identified in Gram-negative bacteria. This has now changed, and to date only six Gram-positive bacteria have been published to contain HtrA homologues. These include *B. subtilis* (Noone, D. *et al.*, 2000, Noone, D. *et al.*, 2001), *Lactobacillus helveticus* (Smeds, A. *et al.*, 1998), *Lactococcus lactis* (Poquet, I. *et al.*, 2000, Poquet, I. *et al.*, 2001), *S. pyogenes* (Jones, C.H. *et al.*, 2001), *Streptococcus mutans* (Diaz-Torres, M.L. & Russell, R.R.B., 2001) and *Streptococcus pneumoniae* (Sebert, M.E. *et al.*, 2002). *B. subtilis* has three HtrA-like proteases encoded in its genome (Noone, D. *et al.*, 2000, Noone, D. *et al.*, 2001) whereas *L. lactis* only contains one. This difference in number is

thought to be due to the differences in the genome size of the two species. *L. lactis* has a much smaller genome (2.4Mb compared with 4.2Mb of *B. subtilis*) and the encoded protease is thought to act as a housekeeping protease involved in the maturation of exported proteins (Smeds, A. *et al.*, 1998).

1.11.1 Comparisons of Gram-negative and Gram-positive bacterial HtrA homologues

Like Gram-negative HtrA, the Gram-positive homologues have been shown to be important in bacterial survival at elevated temperatures in both *B. subtilis* and *L. lactis* (Noone, D. *et al.*, 2000, Poquet, I. *et al.*, 2000). However, the mRNA levels of the *htrA* transcript in *L. helveticus* were only increased two-fold when the bacteria was exposed to 42°C (Smeds, A. *et al.*, 1998). In addition, all the Gram-positive HtrA homologues identified contain the conserved residues of the catalytic triad of serine proteases.

To date, Gram-positive HtrA homologues appear to have a single PDZ-domain, while Gram-negative HtrA homologues often contain two PDZ domains, except in the case of HhoB (reviewed in Pallen, M.J. & Wren, B.W., 1997). In addition the existence of HtrA in Gram-positive bacteria may in itself be unusual as HtrA in Gram-negative bacteria is located in the periplasm, a structure that is missing from Gram-positive bacteria (Figure 1.11).

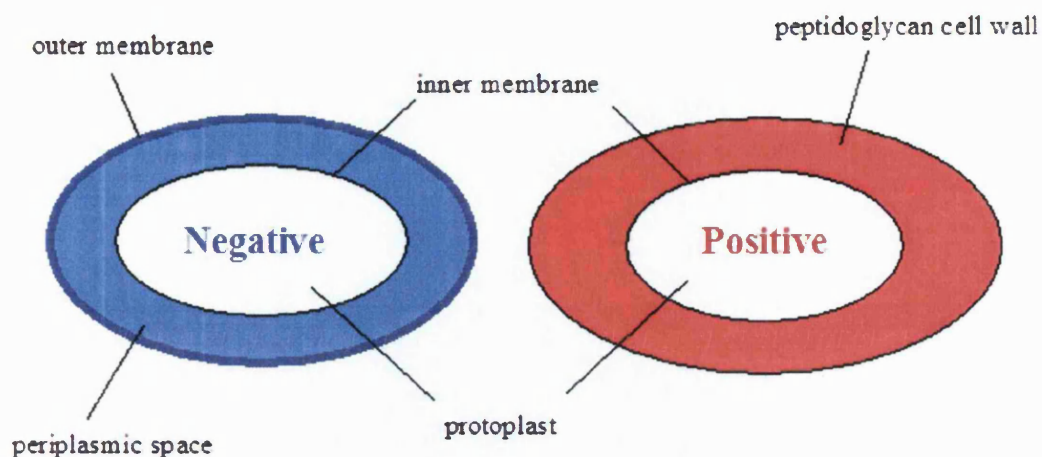


Figure 1.11 – Comparison of structures of Gram-negative and Gram-positive bacteria

As Gram-positive bacteria have no periplasm, the proposed location of HtrA is unknown. It has been suggested that HtrA homologues, which have been identified in Gram-positive bacteria, may be located on the outer surface of the cell membrane on the basis of sequence comparisons (Smeds, A. *et al.*, 1998, Noone, D. *et al.*, 2000, Poquet, I. *et al.*, 2000). The homologues all have a hydrophobic segment at the N-terminus, which could be a membrane anchor, but this in itself is unusual as surface proteins in Gram-positive bacteria are usually associated by a lipoprotein type attachment (Smeds, A. *et al.*, 1998).

1.12 *S. aureus*, a Gram-positive bacterium

S. aureus is a pathogenic Gram-positive bacterium. Alarmingly, it is one of the growing numbers of bacteria that have evolved into a ‘superbug’, which are resistant to many antibiotics. Most strains are resistant to penicillin though in the UK these strains can still be treated with isozacyl penicillins.

Globally, a new strain of *S. aureus* has evolved with resistance to methicillin which until recently could be treated with vancomycin. The emergence of the methicillin-

resistant *S. aureus* (MRSA) has caused much alarm, as this strain is becoming an increasing factor in patient deaths in hospitals. Reports of the emergence of a new strain of MRSA displaying less susceptibility to vancomycin has caused further alarm as this vancomycin-intermediate *S. aureus* (VISA) may give rise to a new strain, vancomycin-resistant *S. aureus* (VRSA), which would not be treatable with current antibiotics.

Although the bacterium is carried asymbiontotically in man on the skin and in the nasopharynx it is capable of causing a range of illnesses. These include food poisoning, wound and line infection, abscesses and toxic shock syndrome. The pathogenicity of *S. aureus* is thought to be facilitated by the production of a number of virulence factors (Chan, P.F. *et al.*, 1998), including both extracellular proteins such as toxins, and also cell-surface-associated proteins (Chan, P.F. & Foster, S.J., 1998). In addition, the bacteria has adapted to survive in a number of environments and stress conditions, including heat shock, reactive oxygen free radicals as well as osmotic shock and starvation conditions (reviewed in Clements, M.O. & Foster, S.J., 1999).

Investigations into the mechanisms by which *S. aureus* survives stress conditions will improve understanding of the bacterium. This understanding may then shed light on possible targets for the future generations of antibiotics in the battle against 'superbugs'.

1.13 *Salmonella*, a Gram-negative bacterium

Salmonella is a facultatively anaerobic Gram-negative bacilli. It is a pathogen, which is common throughout the world, affecting millions of humans and causing a substantial loss of livestock. Clinical manifestations of salmonella infections (salmonellosis) include septicemia, gastroenteritis and enteric fevers.

Gastroenteritis is generally caused by the ingestion of infected meat, and results in diarrhoea, vomiting and nausea. After ingestion of contaminated food, the bacteria colonises the ileum and colon and invades the epithelial lining. The organisms then proliferate and spread into the mesenteric lymph nodes, where it is spread around the body by the circulation system. Eventually, the bacteria are then cleared from the host by reticuloendothelial cells, which contain the infection and remove the bacteria from the system.

A more severe *Salmonella* infection results in enteric fevers such as typhoid fever, which can be fatal if not treated correctly. This form of infection is normally spread by drinking water contaminated with human faeces. Thus, like *S. aureus*, *Salmonella* is a pathogen of clinical importance.

1.14 The identification of HtrA-like proteins in *S. aureus* and *S. typhimurium*

Searching the available information on the NCBI BLAST database revealed three potential HtrA-like proteins encoded in the *S. aureus* genome containing both the conserved catalytic triad and a single PDZ domain. In addition, three *S. typhimurium* proteins were revealed with homology to *E. coli* HtrA and its paralogues.

1.15 Aims of this thesis

This study aims to compare the HtrA homologues in the Gram-negative bacteria, *S. typhimurium* and the Gram-positive bacteria, *S. aureus*.

1. HtrA homologues of *S. aureus* will be cloned and characterised in order to test the hypothesis that these proteins are important in Gram-positive bacteria viability in stress conditions, such as heat shock.
2. Recombinant proteins of HtrA homologues of both *S. aureus* and *S. typhimurium* will also be analysed for protease activity and whether this is its primary function. This is important, as recent studies have shown that *E. coli* HtrA acts mainly as a chaperone at temperatures below 28°C and above this temperature will behave as a protease (Speiss, C. *et al.*, 1999). In addition, the PDZ domains of the *S. aureus* orthologues will be analysed for protein:protein interaction.
3. Mutants of the *htrA* gene family will be constructed in *S. typhimurium* to investigate the importance of this gene family in bacterial viability under stress conditions, such as heat-shock and oxidative stress. In addition, the importance of these genes in bacterial virulence will be analysed.

The findings of this study will increase our understanding of at least one of the mechanisms involved in the maintenance of bacterial viability under stress

conditions and consequently may aid the generation of future antibiotics for pathogenic Gram-positive and Gram-negative bacteria.

CHAPTER 2. MATERIALS AND METHODS

2.1	DNA manipulation.....	68
2.1.1	Isolation of <i>S. aureus</i> genomic DNA	68
2.1.1.1	Phenol-chloroform extraction of DNA	68
2.1.1.2	Ethanol precipitation	69
2.1.2	Polymerase Chain Reaction (PCR)	69
2.1.2.1	Optimisation of MgCl ₂ concentration for PCR	69
2.1.3	Analysis of DNA samples on agarose gel electrophoresis	70
2.1.3.1	Gel preparation	70
2.1.3.2	Sample preparation	70
2.1.3.3	Electrophoresis	71
2.1.4	Cloning of PCR products.....	71
2.1.4.1	Preparation of the plasmid vector	71
2.1.4.2	Preparation of insert (PCR product)	72
2.1.4.3	Ligating inserts into vectors	72
2.1.4.4	Transformation of DH5α competent cells	73
2.1.5	Purification of plasmid DNA.....	73
2.1.5.1	Determining presence of insert in plasmid.....	74
2.1.6	DNA sequencing	74
2.2	Recombinant protein expression	76
2.2.1	Transformation of plasmid into competent cells	76
2.2.1.1	Transformation by electroporation	76
2.2.1.2	Transformation by heat shock	78
2.2.2	Expression of recombinant proteins.....	78
2.2.2.1	Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE).....	79
2.2.2.2	Tricine gel electrophoresis	80
2.2.2.3	Coomassie staining of SDS-PAGE gels.....	81
2.2.2.4	Silver staining of SDS-PAGE gels	81
2.2.2.5	Western blot analysis of SDS-PAGE separated proteins.....	81
2.2.3	Verifying solubility of recombinant proteins	83
2.2.4	Purification of GST-tagged recombinant proteins	83
2.2.4.1	GSTrap affinity chromatography.....	84

2.2.4.2	Removal of thrombin using the HiTrap Benzamidine column.....	85
2.2.4.3	Gel filtration chromatography	85
2.2.5	Small scale purification of 6xHis-tagged recombinant proteins.....	86
2.3	RNA manipulation.....	87
2.3.1	Isolation of <i>S. aureus</i> RNA.....	87
2.3.2	Spectrophotometric quantification of RNA samples	88
2.3.3	Separation of RNA by denaturing gel electrophoresis.....	89
2.3.3.1	Preparation of RNA samples	89
2.3.3.2	Electrophoresis of RNA denaturing gel	89
2.3.4	Northern blot analysis	90
2.3.4.1	Radioactive probe	90
2.3.4.2	Hybridisation of probe to membrane	92
2.3.4.3	Stripping filters	92
2.4	Primer extension	92
2.4.1	Construction of template primer for manual sequencing	92
2.4.2	Manual sequencing.....	93
2.4.2.1	Denaturing plasmid DNA.....	93
2.4.2.2	Sequencing of DNA templates	94
2.4.3	Analysis of DNA sequence by polyacrylamide sequencing gels.....	94
2.4.4	Primer extension	95
2.4.4.1	Radiolabelling of primer	95
2.4.4.2	Primer extension	95
2.5	Heat shock of <i>S. aureus</i>	96
2.6	Construction of <i>S. typhimurium</i> mutants.....	97
2.6.1	Allelic replacement	97
2.6.1.1	Template DNA for resistance cassettes	97
2.6.1.2	Construction of resistance cassettes.....	98
2.6.1.3	Transformation of <i>S. typhimurium</i> SL3261 competent cells with pBAD λ RED.....	98
2.6.1.4	Induction of <i>S. typhimurium</i> SL3261 pBAD λ RED	99
2.6.1.5	Verification of allelic replacement	99
2.6.2	Transduction of mutation into <i>S. typhimurium</i> strain SL1344	100

2.6.2.1	Plate stock prep.....	101
2.6.2.2	Transduction.....	101
2.7	Protease detection using zymograms.....	102
2.7.1	Zymograms.....	102
2.8	<i>S. typhimurium</i> stress conditions <i>in vitro</i>	103
2.8.1	Heat shock studies.....	103
2.8.2	Oxidative stress of <i>S. typhimurium</i> SL3261 mutants.....	104
2.9	Attenuation of <i>S. typhimurium</i> SL1344 mutants <i>in vivo</i>	105
2.9.1	Quantification of <i>S. typhimurium</i> SL1344 mutants.....	105
2.9.1.1	Inoculation of BALB/c mice.....	105
2.9.1.2	Viable counts.....	106
2.10	Statistical analysis.....	106

2.1 DNA manipulation

2.1.1 Isolation of *S. aureus* genomic DNA

Genomic DNA was extracted from *S. aureus* Smith strain kindly grown and donated from Arrow Therapeutics, London. The bacteria was prepared as follows:

A single bacteria colony was inoculated into 2 ml of Lennox L-Broth (L-Broth: 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 5% (w/v) NaCl, Sigma) and grown overnight at 200 rpm, 37°C. The culture was then transferred to 20 ml of L-Broth and grown for a further 6 hr. Cells were pelleted by centrifugation, 3000 x g for 20 min. The pellet was transferred to a microfuge tube and stored at -20°C until required.

The *S. aureus* pellet was resuspended with 1 ml of L-Broth and 200 µl aliquots were transferred to microfuge tubes. Cells were pelleted by centrifugation, 16000 x g for 2 min and the supernatant discarded. Pellets were then stored at -20°C until use.

2.1.1.1 *Phenol-chloroform extraction of DNA*

The *S. aureus* pellet was resuspended in 200 µl of ddH₂O. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1; Invitrogen) was added. The microfuge tube was gently inverted a few times to mix. The organic and aqueous layers were separated by centrifugation at 8900 x g for 10 min at room temperature.

The top aqueous layer containing nucleic acids was carefully transferred to a fresh microfuge tube. The denser organic layer containing proteins was discarded. This extraction procedure was repeated until the top aqueous layer was clear.

2.1.1.2 Ethanol precipitation

DNA was concentrated using ethanol precipitation. To the final aqueous layer obtained from the phenol-chloroform extraction (2.1.1.1), 0.1 volume of 3 M sodium acetate, pH 4.8, and 2.5 volumes of 95% (v/v) ethanol were added. The microfuge tube was gently inverted several times to thoroughly mix the reagents. Samples were then stored at -80°C for 1.5 hr. On removal, DNA was pelleted at 16000 x g for 10 min at room temperature and the supernatant carefully removed and discarded. The DNA pellet was carefully washed with 200 µl of 70% (v/v) ethanol and briefly centrifuged at 16000 x g for 5 min at room temperature. Again the supernatant was removed and discarded and samples dried in a speedvac (Eppendorf Concentrator 5310) for 20 min. The DNA pellet was then resuspended in 50 µl of ddH₂O and stored at -20°C until used.

2.1.2 Polymerase Chain Reaction (PCR)

2.1.2.1 Optimisation of MgCl₂ concentration for PCR

PCR was carried out using the Expand High Fidelity PCR system (Roche) according to the manufacturer's recommendations using PCR nucleotide mix (dNTPs, Promega). In general, 25 µl reactions were set in 200 µl PCR tubes with 0.1 µg of each PCR primer pair (Sigma Genosys), 200 µM dNTPs, 0.3U of High Fidelity Polymerase (Roche), 1X PCR buffer (Roche) and 2 µl of *S. aureus* genomic DNA (purified as in 2.1) or 1 µl of *S. typhimurium* genomic DNA (kindly donated by Arrow Therapeutics, London). Optimum MgCl₂ concentrations for each insert were determined using increasing MgCl₂ concentration from 0 to 4 mM (final concentration). Primers were designed incorporating the two different restriction sites at the 5' and 3' end of the DNA fragment (appendix A.2).

PCR reactions were carried out in a PCR machine (MWG-Biotech Primus 96 plus). Samples were denatured at 97°C for 5 min, followed by 30 cycles of PCR: melting the template at 94°C for 1 min, annealing at 55°C for 1 min and extending the DNA segment at 72°C for 1 min or 3 min, depending on length of product. This was followed by 7 min at 72°C in order to allow extension of DNA fragments to be completed. PCR products were stored at 4°C until further use.

2.1.3 Analysis of DNA samples on agarose gel electrophoresis

2.1.3.1 Gel preparation

DNA fragments were separated by electrophoresis on 1% agarose gels. These were prepared by melting 1 g of agarose (Invitrogen) in 100 ml of Tris-Acetate-EDTA (TAE) buffer (80 μ M Tris, 20 μ M ethylenediaminetetraacetic acid (EDTA), pH 8.0, 1.1% (v/v) glacial acetic acid) in a microwave until the agarose was melted. Ethidium bromide (2.5 μ l/100 ml, Sigma) was added to the molten gels and mixed. Gels were poured into the casting trays of a horizontal electrophoresis tank (Horizon 58, Invitrogen) with the appropriate combs inserted. Gels were allowed to set and TAE buffer was added to cover the gel.

2.1.3.2 Sample preparation

DNA samples (3-5 μ l) were mixed with 1 μ l of gel loading buffer (6X concentrate, Type I, Sigma). The samples were then loaded into the wells using a micro-pipette.

2.1.3.3 Electrophoresis

Agarose gels were electrophoresed at a constant voltage of 100V for 30-45 min. Gels were usually stopped if the marker dye had moved $\frac{1}{2}$ to $\frac{3}{4}$ of the length of the gel. DNA bands were visualised under UV light.

2.1.4 Cloning of PCR products

PCR products were cloned into the multiple cloning sites (MCS) of glutathione-S-transferase (GST) fusion plasmid, pGEX-4T-1 (Amersham), the native plasmid pTrc99A (Amersham) or the 6xHis-fusion plasmids pQE-30 (Qiagen) and pRSET-B (Invitrogen).

2.1.4.1 Preparation of the plasmid vector

Each plasmid was digested with 30U each of the appropriate restriction endonuclease (New England Biosciences (NEB) or Roche) in a final volume of 130 μ l. In addition, 50U of calf intestinal alkaline phosphatase (NEB) was added in order to dephosphorylate the 5'-phosphate group of the plasmid to prevent recircularisation. Incubations were carried out at 37°C for 1.5 hr. After this time, an additional 10U of each restriction endonuclease was added and incubated for a further 1 hr.

To check the reaction was complete with no star products, 5 μ l of the reaction mix on an agarose gel (2.1.3). Linearised plasmid DNA was then purified using the QIAquick PCR purification kit (Qiagen) following the manufacturer's recommendations.

2.1.4.2 Preparation of insert (PCR product)

To clone the PCR products into their respective vectors, PCR products from separate PCR reactions (2 x 50 µl) were checked on an agarose gel for products other than the one of interest. In all cases, the reaction yielded only the desired product and so PCR product purification was carried out using a QIAquick PCR purification kit (Qiagen) following the manufacturer's recommendations.

The PCR product was then digested using 30U of each respective restriction endonuclease in a final reaction volume of 80 µl. Digests were carried out at 37°C for 1.5 hr after which, a further 10U of each enzyme was added. Reactions were continued for a further 1 hr.

To check the reaction was complete with no star products, 5 µl of the reaction mix was loaded on an agarose gel (2.1.3). Linearised plasmid DNA was then purified using the QIAquick PCR purification kit (Qiagen) following the manufacturer's recommendations.

2.1.4.3 Ligating inserts into vectors

Equal volumes (5 µl) of the purified digested insert and the purified digested vector were loaded separately onto an agarose gel (2.1.3) and quantified using Hyperladder I (Bioline). Molar ratios were calculated and ligation reactions were set up using a 7:1 (insert:plasmid) ratio, 2 µl 5X T4 ligase buffer (Invitrogen) and 0.1U of T4 ligase (Invitrogen) in a 10 µl reaction volume. Ligations were carried out at room temperature for 1 hr.

2.1.4.4 Transformation of DH5 α competent cells

MAX EFFICIENCY DH5 α competent cells (Invitrogen) were used to amplify the plasmid DNA. Transformation by heat-shock was carried out following the manufacturer's recommendations. Briefly, 2 μ l of the ligation mixture was added to 20 μ l of competent cells, previously thawed on ice, flicked gently to mix and stored on ice for 30 min. Samples were then heat-shocked at 42°C for 45 sec. Immediately after heat-shock the samples were stored on ice for 2 min before the addition of 180 μ l of SOC media (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose, Invitrogen). Samples were then incubated at 37°C shaking at 200 rpm for 1 hr before plating 100 μ l of cells were plated onto 90 mm petri dishes (Sterilin) containing L-agar (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl and 1.5% (w/v) agar, pH 7.0, Sigma) supplemented with 100 μ g/ml ampicillin (Sigma) and incubated overnight at 37°C.

2.1.5 Purification of plasmid DNA

Single colonies containing a putative plasmid of interest were picked and grown overnight at 37°C at 200 rpm in 5 ml L-Broth supplemented with 100 μ g/ml ampicillin. At the end of the growth period, bacteria were pelleted at 3000 x g for 10 min and the supernatant discarded. Plasmid DNA was purified from these pellets using a commercially available kit, QIAprep miniprep kit (Qiagen) following the manufacturer's recommendations.

2.1.5.1 Determining presence of insert in plasmid.

a) Restriction digests

Purified plasmid DNA (2 µl) was digested in the presence of 10U of each appropriate restriction endonuclease for 1 hr at 37°C in a final volume of 20 µl and 5 µl of the reaction mix was loaded onto an agarose gel (2.1.3) for analysis.

b) Colony PCR

A sterile Gilson pipette tip was used to pick a single colony and then dipped into a 200 µl PCR tube with the appropriate primers (10 µg each) before inoculating the L-Broth. 15 µl 1X PCR mix (Promega) containing 200 µM dNTPs (Promega) and 1U of *Taq Pol* (Promega) was added. Samples were briefly vortexed before carrying out the reactions as in 2.1.2 and analysed using agarose gel electrophoresis (2.1.3).

2.1.6 DNA sequencing

The insert in the plasmid must be sequenced in order to verify that the PCR was correct. DNA sequencing was carried out using a CEQ-2000 sequencing machine (Beckman Coulter). Sequencing reactions were carried out using a commercially available kit, CEQ Dye Terminator Cycle Sequencing (DTCS) quick start kit. All reagents were supplied by the kit, unless otherwise stated. No more than 100 fmol of purified plasmid in a volume of 6 µl was required for each sequencing reaction. This was quantified by comparing single digest reactions of the plasmids with a quantitative DNA molecular weight marker, Hyperladder I (Bioline), on an agarose gel.

The template DNA was made up to a volume of 11 μ l in a 200 μ l PCR tube and denatured at 96°C for 1 min before the addition of 1 μ l of primer (4 pmol/ μ l; listed in appendix A.5) and 8 μ l DTCS quick start master mix. Samples were mixed by vortexing and then underwent 30 cycles of PCR: 96°C for 20 sec, 50°C for 20 sec and 60°C for 4 min, followed by holding at 4°C. In fresh 1.5 ml microfuge tubes, 5 μ l of stop solution was added (2 μ g glycogen, 1.5 M sodium acetate, pH 4.8 (Sigma), 50 mM EDTA (Sigma)). The sequencing reaction was added to this and then 60 μ l of 95% (v/v) ice-cold ethanol was added to precipitate the DNA. Samples were vortexed to mix and then pelleted at 16000 x g for 15 min at 4°C. The supernatant was removed carefully and discarded. Pellets were washed twice with 200 μ l of ice-cold 70% (v/v) ethanol followed by centrifugation at 16000 x g for 2 min, 4°C. The pellets were then dried in a speedvac for 40 min at room temperature. Samples were then resuspended in 35 μ l of sample loading solution and transferred to a sequencing plate. Samples were overlaid with one drop of mineral oil to prevent evaporation. The following conditions (Table 2.1) were used for sequencing:

Table 2.1: Sequencing conditions for CEQ-2000 sequencing machine

Capillary temperature:	Preheated to 50°C	
Denature:	Temperature: 90°C	120 sec
Injection:	Voltage: 2.0 kV	15 sec
Separate:	Voltage: 4.0 kV	100 min

Sequencing data was analysed using the program Bioedit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) to check for mismatches to the genomic sequence.

2.2 Recombinant protein expression

2.2.1 Transformation of plasmid into competent cells

Plasmids with the correct inserts were used for expression studies. Each plasmid was transformed into one or more expression systems by electroporation or heat shock (detailed 2.2.1.1 and 2.2.1.2, respectively). Below in Table 2.2 details the bacterial strains used.

Table 2.2: Bacterial strains (genotypes listed in appendix A.3)

Bacterial strain	Form of transformation	Manufacturer
BL21	Heat-shock	Novagen
BL21(DE3)	Heat shock	Novagen
BL21(DE3)pLysS	Heat-shock	Novagen
BL21(DE3)CodonPlus	Electroporation	Stratagene
DH5 _α	Heat-shock	Novagen
JM109	Heat-shock	Promega
SG1009	Electroporation	Qiagen
M15	Electroporation	Qiagen

2.2.1.1 Transformation by electroporation

The bacterial strains BL21(DE3)CodonPlus, M15, SG13009 were prepared as electrocompetent cells. All other strains were purchased as competent cells for heat-shock uptake of DNA.

a) *Preparation of competent cells*

A single colony was inoculated into 5 ml of L-Broth and grown overnight at 37°C at 200 rpm. 1:100 overnight cultures were used to inoculate 50 ml of L-Broth. Cultures were grown at 37°C at 200 rpm until the OD₆₀₀ was approximately 0.5. Cultures were immediately placed on ice for 15 min. 1.5 ml of culture was aliquoted into sterile microfuge tubes and pelleted briefly at 16000 x g for 30 sec. Supernatant was removed and another 1.5 ml of culture was added and pelleted. The supernatant was removed and pellets were washed in 1 ml of ice-cold sterile ddH₂O. The supernatant was removed and cells were washed once more with 1 ml of ice-cold sterile ddH₂O. Pellets were resuspended in 1 ml of ice-cold sterile 10% (v/v) glycerol followed by centrifugation at 16000 x g for 30 sec. Supernatant was removed and cells resuspended in 80 µl of ice-cold sterile 10% (v/v) glycerol. Cells were used immediately or stored at -80°C until use.

b) *Electroporation*

1 µl of plasmid was added to the electrocompetent cells. Samples were transferred to pre-chilled 1 mm electro-cuvettes (Equibio). Cells were electroporated at 2500V using the Eppendorf electroporator 2510. 500 µl of L-Broth was added immediately to the samples and then transferred to a fresh sterile 1.5 ml microfuge tube and incubated at 37°C for 1 hr at 250 rpm. Cells were pelleted at 16000 x g for 30 sec and the pellet resuspended in 100 µl of L-Broth. 50µl was plated onto L-agar plates supplemented with the appropriate antibiotics.

2.2.1.2 Transformation by heat shock

Plasmid (1 μ l) was added to competent cells and incubated on ice for varying lengths of time (Table 2-3).

Table 2-3: Length of pre-heat shock incubation

Bacterial strain	Length of incubation
BL21	
BL21(DE3)	5 min
BL21(DE3)pLysS	
JM109	20 min
DH5 α	30 min

Samples heat-shocked at 42°C for 45 sec and subsequently placed immediately on ice for 2 min before addition of 180 μ l SOC media. Cells were then incubated at 37°C for 1hr before plating 50 μ l onto L-agar plates supplemented with the appropriate antibiotics.

2.2.2 Expression of recombinant proteins

A single colony was picked to inoculate 5 ml of L-Broth, supplemented with the appropriate antibiotics, and grown overnight at 37°C at 200 rpm. An overnight culture was used to inoculate 10 ml of fresh L-Broth or Terrific broth (T-Broth, 1.2% (w/v) tryptone, 2.4% (w/v) yeast extract, 0.94% (w/v) K₂HPO₄, 0.22% (w/v) KH₂PO₄, 0.8% (v/v) glycerol, Sigma) to give an OD₆₀₀ of 0.2. Cultures were grown at 37°C or 26°C at 200 rpm until an OD₆₀₀ of 1.0 was reached. The cultures were subsequently induced with varying concentrations (0 to 5 mM) of isopropyl- β -D-

thiogalactoside (IPTG, Melford Laboratories Ltd). At varying time points, 100 µl aliquots were removed and the OD₆₀₀ readings were taken. Cells were pelleted and resuspended in 50 µl of dH₂O and prepared for protein gel electrophoresis, described in 2.2.2.1 and 2.2.2.3 and the equivalent of OD₆₀₀ of 0.1 was loaded per well. Western blot analysis was carried out (2.2.2.5) in most cases to verify the presence of the protein of interest.

2.2.2.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Mini-gels were prepared and performed using the Mini-Protean II protein electrophoresis cell (BioRad) according to the manufacturer's instructions. All gels had an acrylamide to bisacrylamide ratio of 37.5:1 (BioRad). Resolving gels were prepared with varying amounts of acrylamide/bisacrylamide ranging from 8 % to 20% in 375 mM Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.1% (v/v) ammonium persulphate (Sigma) and 0.06% (v/v) N, N, N', N'-tetramethylethylenediamine (TEMED, Sigma). The appropriate resolving gel was used with a 4% stacking gel (prepared as resolving gel but with the exception of using 187.5 mM Tris-HCl pH 6.8 and 0.1% TEMED). An equal volume of 2X solubilisation buffer (100 mM Tris-HCl, pH 6.8, 4% (w/v) SDS, 0.2% (w/v) bromophenol blue, 20% (w/v) glycerol, 10% (v/v) mercaptoethanol) was added to each sample. Samples were denatured at 99°C for 5 min, briefly cooled, before loading on the gel. The gels were run in electrode buffer (25 mM Trizma base (Sigma), 192 mM glycine, 0.1% (w/v) SDS) and electrophoresed at constant voltage of 200V for 45 min using the BioRad Protean II gel system. Rainbow molecular weight markers (range 14 to 200 kDa, Amersham) were loaded onto all gels.

After electrophoresis, gels were analysed by either Coomassie or silver staining, and also by western blot analysis according to the methods below.

2.2.2.2 Tricine gel electrophoresis

Tricine gel electrophoresis was carried out to separate smaller protein molecules (5 to 20 kDa). Mini gels were prepared and performed using the Mini Protean II protein electrophoresis cell according to the manufacturer's recommendations. Resolving gels were prepared with 12% acrylamide/bisacrylamide (29:1 ratio, BioRad) in 1 M Tris-HCl pH8.45, 0.1 % (w/v) SDS, 12% (v/v) glycerol, 3.6% (v/v) ammonium persulphate and 0.1% (v/v) TEMED. A 4% stacking gel, in 778 mM Tris-HCl, pH 8.45, 0.078% (v/v) SDS, 0.008% (v/v) ammonium persulphate and 0.0008% TEMED, was immediately poured on top of the stacking gel. An equal volume of 2X solubilisation buffer was added to the samples, which were then denatured by heating to 99°C for 5 min. The samples were briefly cooled before loading onto the gel. The gels were run in electrode buffer (100 mM Trisma, 100 mM tricine (Sigma), 0.1% (w/v) SDS, pH8.3) and electrophoresed at constant voltage of 150V for 60 to 70 min. Protein ladders, the rainbow molecular weight markers (range 14.3 to 200 kDa, Amersham) and the MultiMark® multicoloured standard (Invitrogen), were loaded onto all gels.

After electrophoresis, the gels were analysed by Coomassie-staining (2.2.2.3) and/or Western blot analysis (2.2.2.5).

2.2.2.3 *Coomassie staining of SDS-PAGE gels*

Mini-gels were electrophoresed as described in 2.2.2.1 or 2.2.2.2. The gel apparatus was dismantled and gels placed in Coomassie stain (0.25% (w/v) Coomassie blue, 45% (v/v) ethanol, 10% (w/v) glacial acetic acid) and rocked for 45 min. The stain was replaced with destain (10% (v/v) glacial acetic acid, 30% (v/v) ethanol) and gels destained until protein bands were visible.

2.2.2.4 *Silver staining of SDS-PAGE gels*

Mini-gels were electrophoresed as described in 2.2.2.1. The gel apparatus was dismantled and gels stained using the BioRad silver stain plus kit following the manufacturer's recommendations.

2.2.2.5 *Western blot analysis of SDS-PAGE separated proteins*

Four pieces of 3MM filter paper (Whatman) was cut to the exact size of the gels to be used for western blot analysis and soaked in transfer buffer (25 mM Tris-HCl, 192 mM glycine, 20% (v/v) methanol). A 'sandwich' was made by placing two pieces of filter paper on a pre-soaked sponge in the cassette of a wet electroblotter (BioRad). A pre-soaked sheet of Hybond™ ECL™ nitrocellulose membrane (Amersham), cut to size, was placed on top of the filter paper. On top on the membrane, a protein gel was placed, onto which the remaining two filter papers were placed. Another pre-soaked sponge was placed on top and the cassette closed. The blotting apparatus was assembled and the proteins transferred at a constant voltage of 100V for at least 1 hr at room temperature. At the end of the transfer the filter paper and gel were discarded. The membrane was incubated in blocking buffer (5% (w/v) non-fat milk in PBST: 0.1% (v/v) Tween 20, 80 mM Na₂HPO₄, 20 mM NaH₂PO₄,

100 mM NaCl) for 2 hr at room temperature or overnight at 4°C in order to block non-specific protein binding. Blots were washed three times at room temperature in PBST for 5 min each before addition of primary antibody (see Table 2.4 for antibodies used) diluted in appropriate buffer for 2 hr at room temperature.

Table 2.4: Antibodies used for Western blot analysis

		Dilution	Dilution buffer	Length of incubation	Manufacturer
Goat anti-GST antibody	Primary antibody	1:40,000	5% (w/v) milk in PBST	2 hr	Amersham
Monoclonal anti-polyhistidine clone His-1	Primary antibody	1:2000	1% (w/v) BSA in PBST	2 hr	Sigma
Donkey, anti-goat IgG, HRP conjugate	Secondary antibody	1:20,000	5% (w/v) milk in PBST	2 hr	Promega
Anti-mouse IgG (H + L), HRP conjugate	Secondary antibody	1:5000	1% (w/v) milk in PBST	1 hr	Promega

Again, blots were washed three times at room temperature in PBST for 5 min each. The appropriate secondary antibodies, which were conjugated with horseradish peroxidase diluted in the appropriate buffer, were added and incubated for the correct length of time (Table 2.4). Immunoreactive bands were visualised using ECL substrate for western blotting (Amersham). Equal volumes of each reagent were added and poured onto the blots. Blots were incubated for 1 min and then any excess liquid was drained off prior to wrapping the blots in Saranwrap. Blots were then exposed to HyperfilmTM ECLTM (Amersham) for between 30 sec to 5 min after which the film was developed.

2.2.3 Verifying solubility of recombinant proteins

Cell pellets from 10 ml induction incubations were resuspended in 500 µl of ice-cold sonicating buffer (100 mM Tris-HCl, pH 8.0, 150 mM NaCl) and transferred to a fresh microfuge tube. Samples were sonicated on ice. Sonication was carried out using twenty 1 sec bursts, followed by a 20 sec rest on ice, and then another twenty 1 sec bursts (Status 70, Philip Harris Scientific).

Cell debris was removed by ultracentrifugation at 105000 x g for 30 min at 4°C (TLA.2 rotor, Beckman). The supernatant was transferred to a fresh microfuge tube and 50 µl of supernatant was added to 50 µl of 2X solubilisation buffer and denatured at 99°C for 5 min. A 2 µl sample was loaded onto SDS-PAGE or Tricine gels and electrophoresed as described in 2.2.2.1 and 2.2.2.2, respectively. Gels were stained with Coomassie (2.2.2.3). Western blot analysis (2.2.2.5) was carried out in most cases to verify the presence of the recombinant protein.

2.2.4 Purification of GST-tagged recombinant proteins

A 200 ml induction culture was resuspended into 5 ml of sonicating buffer. Samples were sonicated and cell debris removed as described in 2.2.3. Nucleic acids were sheared by passing through a 22-gauge needle (VWR) and then further clarified by filtering through a 0.22 µm filter. Purification of the recombinant proteins were carried out by fast protein liquid chromatography (FPLC) on the AKTA FPLC system (Amersham).

2.2.4.1 GSTrap affinity chromatography

GST-tagged proteins were isolated by affinity chromatography using a 1 ml GSTrap column (Amersham) on the AKTA FPLC system (Amersham) with the conditions described below in Table 2.5.

Table 2.5: FPLC conditions using GSTrap

Column	GSTrap
Column volume (CV)	1 ml
Buffer A:	PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na ₂ HPO ₄ , 1.8 mM KH ₂ PO ₄ , pH 7.3)
Buffer B:	50 mM Tris-HCl, 10 mM reduced glutathione (Sigma), pH 8.0
Flow rate:	1 ml/min
Equilibration	5 CV buffer A
Load sample	1 ml
Wash	5 CV buffer A
Elution	6 CV buffer B in 1ml fractions
Reequilibration	5 CV buffer A

Fractions were analysed by SDS-PAGE (2.2.2.1) and by Western blot analysis (2.2.2.5) probing with an anti-GST antibodies. Fractions containing the GST-tagged protein were pooled and protein concentration determined (BioRad BCA protein assay). The GST-tag was removed by incubation at room temperature for 16 hr with 0.2 NIH U of thrombin (Amersham) per 100 µg protein. Samples were clarified through a 0.22 µm filter before being loaded onto the next column.

2.2.4.2 Removal of thrombin using the HiTrap Benzamidine column

To prevent ‘overcleaving’, the thrombin was removed by passing the sample through a HiTrap benzamidine FF column (1 ml, Amersham). The conditions used are shown in Table 2.6.

Table 2.6: FPLC conditions using HiTrap Benzamidine FF

Column	HiTrap Benzamidine FF 1ml
Column volume (CV)	1 ml
Buffer C:	50 mM Tris-HCl, pH 8.0
Buffer D:	50 mM Tris-HCl, 1 M NaCl, pH 8.0
Flow rate:	1 ml/min
Equilibration	5 CV buffer C
Load sample	5 ml
Wash	20 CV buffer C in 5 ml fractions
Elution	10 CV buffer D in 2 ml fractions
Reequilibration	5 CV buffer C
Thrombin removal	Buffer D + 20 mM p-aminobenzamidine

Fractions were analysed by SDS-PAGE (2.2.2.1) and by Western blot analysis (2.2.2.5). Fractions containing the protein of interest were pooled, concentrated (Millipore Ultra-Free-15) and clarified by passing through 0.22 µm filter.

2.2.4.3 Gel filtration chromatography

Gel filtration chromatography was used to remove the remaining contaminants. A gel affinity column, HiLoad Superdex 75 26/60 prep grade (Amersham) was used with the conditions outlined in Table 2.7.

Table 2.7: FPLC conditions using HiLoad Superdex 75 26/60 prep grade

Column	HiLoad Superdex 75 26/60 prep grade
Column volume (CV)	318.557 ml
Buffer E:	20 mM MES, 1 mM DTT, pH 6.6
Flow rate:	2 ml/min
Equilibration	3 CV buffer E (at 1 ml/min for removal of 20% ethanol)
Load sample	13 ml
Elution	1.5 CV buffer E in 12 ml fractions
Reequilibration	2 CV buffer E

Fractions were analysed by SDS-PAGE (2.2.2.1) and Western blotting (2.2.2.5).

Fractions containing the protein of interest were pooled, concentrated (Millipore Ultra-Free 15) and stored at -20°C until use.

2.2.5 Small scale purification of 6xHis-tagged recombinant proteins

Induction cultures of 10 ml were pelleted at 3000 x g and supernatant discarded. The pellet was resuspended in 500 µl of PBS (140 mM NaCl, 2.7 mM KCl, 10.1 mM Na₂PO₄, 1.8 mM KH₂PO₄, pH 7.3) and sonicated as described in 2.2.3. The cellular debris was pelleted by centrifugation at 16000 x g at 4 °C for 15 min. The supernatant (400 µl) was added to 100 µl of 50% Ni-NTA slurry (Qiagen), which was pre-washed with 400 µl of PBS. Samples were continually inverted at 4°C for 2 hr to allow binding of the 6xHis-tagged proteins to the column. Beads were pelleted at 750 x g for 10 sec, and the flowthrough removed. The Ni-NTA resin was then washed twice with 400 µl of washing buffer (20 mM phosphate, 500 mM NaCl, 20

mM imidazole pH 7.4). The samples were eluted by the addition of 50 µl of elution buffer (20 mM phosphate, 500 mM NaCl, 400 mM imidazole pH 7.4). Fractions were analysed by SDS-PAGE (2.2.2.1) or tricine gel electrophoresis (2.2.2.2) and Western blotting (2.2.2.5).

2.3 RNA manipulation

2.3.1 Isolation of *S. aureus* RNA

Total RNA was isolated using the TRIZOL reagent (Invitrogen), a commercially available guanidine isothiocyanate-phenol solution. *S. aureus* pellets were resuspended in 1 ml of TRIZOL reagent to which 500 µl of zirconia beads (0.1 mm diameter, Biospec Products Inc) were added. Samples were then vortexed for 90 sec to disrupt the bacterial cell wall. Samples were incubated for 5 min at room temperature to permit the complete dissociation of nucleoprotein complexes and also to allow the beads to settle. The solution was transferred to a fresh microfuge tube containing 200 µl of chloroform (Sigma). Samples were vortexed for 30 sec followed by incubation for 3 min at room temperature. The aqueous and organic layers were separated by centrifugation at 12000 x g for 15 min at 4°C. The upper aqueous layer containing RNA was transferred to a fresh microfuge tube.

To precipitate the RNA, 0.5 ml isopropanol (Sigma) was gently mixed with the sample. The samples were briefly vortexed before pelleting the RNA at 12000 x g for 15 min at 4°C. Pellets were washed with 1 ml of 80% (v/v) ethanol (in diethyl pyrocarbonate (DEPC, Sigma) treated ddH₂O) followed by centrifugation at 12000 x g for 5 min at 4°C. Pellets were air-dried and then resuspended in 87 µl of DEPC ddH₂O.

The isolated RNA was treated with DNase to remove any residual DNA. The RNA sample was incubated with 20U of RNase-free DNase (Roche) in 1X DNase buffer (Roche) in a final volume of 100 μ l. Samples were incubated for 30-60 min at 37°C and the reaction was stopped with 10 μ l of 100 mM EDTA, pH 8.0.

Samples were then passed through a RNeasy mini kit column (Qiagen) to remove RNA species smaller than 200 nucleotides (nt). Samples were eluted in 3 x 50 μ l of DEPC ddH₂O and pooled. RNA was then precipitated with 0.1 volume of 3 M sodium acetate, pH 5.2 (Sigma), and 2.5 volumes of ice-cold ethanol (> 99% (v/v)). Samples were precipitated at -80°C for at least 1 hr or on dry ice for 30 min. RNA was pelleted by centrifugation at 16000 x g for 30 min at 4°C. Pellets were then washed with 1 ml of 80% (v/v) ethanol (in DEPC ddH₂O) and centrifuged at 16000 x g for 10 min at 4°C. The supernatant was discarded and samples air-dried for approximately 10 min. Finally, the samples were dissolved in 7 μ l of DEPC ddH₂O and incubated at 55°C for 5 min.

2.3.2 Spectrophotometric quantification of RNA samples

A 40 μ g/ml solution of RNA has an absorbance of OD = 1.0 at a wavelength of 260 nm. Hence,

$$\text{Concentration } (\mu\text{g/ml}) = \text{absorbance} \times 4 \times \text{dilution factor}$$

RNA samples were diluted in 1:100 in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). RNA was judged to be of sufficient quality for subsequent experiments if the OD₂₆₀/OD₂₈₀ was greater than 1.8.

2.3.3 Separation of RNA by denaturing gel electrophoresis

A 1% (w/v) agarose gel was prepared by dissolving 0.5 g agarose (Invitrogen) in 37 ml of DEPC ddH₂O and 5 ml of 10X MOPS buffer (200 mM MOPS, 50 mM sodium acetate, 1 mM EDTA). Molten agarose was allowed to cool to room temperature before the addition of 8 ml of formaldehyde (37%, Sigma). The gel was poured immediately into the casting tray of a horizontal electrophoresis tank (Horizon 58, Invitrogen). Once set, the gel was covered with running buffer (1X MOPS buffer (20 mM MOPS, 5 mM sodium acetate, 100 µM EDTA).

2.3.3.1 Preparation of RNA samples

To each RNA sample, 2 volumes of 1.5X sample loading buffer (0.006 % (w/v) bromophenol blue, 0.06% (w/v) xylene cyanol, 1.5X MOPS buffer, 9% (v/v) formaldehyde, 60% (v/v) deionised formamine, 4.7% (v/v) ethidium bromide) was added. Samples were linearised by heating at 65°C for 15 min and then were briefly cooled before loading onto the gel. The 0.24-9.5-kb RNA ladder (Invitrogen) was also loaded onto the gel.

2.3.3.2 Electrophoresis of RNA denaturing gel

Gels were run at 100V (constant voltage) until the bromophenol blue band was approximately $\frac{2}{3}$ down the length of the gel. Gels were visualised under a UV light in order to analyse the quality of the RNA.

2.3.4 Northern blot analysis

RNA was transferred onto a nylon membrane (Hybond N+, Amersham Biosciences) by capillary action. A long strip of 3MM filter paper was presoaked in 20X SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0) and used as a wick on a raised plastic platform. The gel was placed on to top and covered with the nylon membrane. Five strips of 3MM filter paper presoaked in 20X SSC was placed on top of the membrane. A stack of paper towels (2 inches deep) was placed on top and the stack weighed down. RNA was allowed to transfer overnight to the membrane by capillary action. The RNA was then crosslinked to the membrane by UV.

2.3.4.1 Radioactive probe

The required DNA sequence to be used as a probe was amplified by PCR (2.1.2) using the primers listed in appendix A.5. PCR products (3 µl) were then ligated with 1 µl of the pGEM-T Easy vector (Promega) as recommended by the manufacturer's instructions. JM109 competent cells were transformed with 2 µl of the ligation mix as described in 2.2.1.2. A 100 µl aliquot of the transformation mix was plated onto L-agar plates supplemented with 100 µg/ml ampicillin, 0.5 M IPTG and 40 µg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactosidase (X-Gal, Melford Laboratories Ltd) in order to carry out blue/white screening. Positive single white colonies were picked and grown overnight in 5 ml of L-Broth supplemented with 100 µg/ml ampicillin at 200 rpm at 37°C. Plasmids were isolated using the QIAprep miniprep kit (Qiagen) according to the manufacturer's recommendations and 2 µl of the plasmid was used to check for the correct sequence by DNA sequencing (2.1.6).

The required DNA sequence was excised from pGEM-T Easy vector using 30U of each of the appropriate restriction endonucleases and incubating for 1.5 hr at 37°C. The reaction was run on an agarose gel (2.1.3) and the required band excised from the gel using a scalpel. DNA extraction was performed using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's recommendations. Aliquots (5µl) were run on an agarose gel (2.1.3). Quantitative ladders (Hyperladder I or IV, Bioline) were loaded onto the gel to quantify the amount of DNA fragment.

The High Prime DNA Labelling kit (Roche) was used to radiolabel 25 ng of probe DNA with [$\alpha^{32}\text{P}$] using a mixture of random hexanucleotide primers. Probe DNA (25 ng in a volume of 8 µl) was denatured at 100°C for 10 min. Sample was chilled briefly on ice. The reaction mix below was prepared and added to the sample.

1 µl dATP	(0.5 mM in Tris buffer)
1 µl dTTP	(0.5 mM in Tris buffer)
1 µl dGTP	(0.5 mM in Tris buffer)
4 µl Reaction mix (5X)	
2 µl [$\alpha^{32}\text{P}$] dCTP (3000 Ci/mmol).	

The reaction mix was then incubated at 37°C for 30 min.

Unreacted [$\alpha^{32}\text{P}$] dCTP and incomplete fragments were removed from the samples by passing through a Nick column (Amersham). The cleaned probe was then denatured at 100°C for 10 min.

2.3.4.2 Hybridisation of probe to membrane

The denatured probe was chilled on ice for 1 min before adding to 10 ml of ExpressHyb (Clontech), which was pre-warmed to 68°C. This was then poured onto the membrane, which had been pre-incubated at 68°C for 30 min with ExpressHyb containing 100 µg/ml denatured salmon sperm DNA. The probe was allowed to hybridise for 1 hr at 68°C, with continual rotation, before washing three times with 2X SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0, 0.05% (w/v) SDS) for 15 min at room temperature. The membrane was then subjected to three washes in 0.1X SSC (15 mM NaCl, 1.5 mM sodium citrate, pH 7.0, 0.1% (w/v) SDS) for 15 min at 50°C. Membranes were then wrapped in Clingfilm and exposed to phosphorimager screens (Packard Instruments Co.) overnight and then visualised by using Optiquant (Packard Instruments Co.).

2.3.4.3 Stripping filters

In order to probe for another mRNA transcript, the membrane had to be stripped of any hybridised probes. This was performed by heating 0.5% (w/v) SDS to 100°C and immediately pouring onto the membrane. Membranes were subsequently incubated at 68°C for 2 hr on a rotating platform. Excess liquid was drained off, and membranes wrapped in Clingfilm and exposed to phosphorimager screens for 1 hr to check for residual radioactivity.

2.4 Primer extension

2.4.1 Construction of template primer for manual sequencing

The DNA sequence upstream of the sequence of interest was cloned into pGEM-T Easy vector (Promega). Firstly, the sequence of interest was amplified by PCR

(2.1.2) and then 3 μ l was ligated into pGEM-T vector as recommended by manufacturer's instructions. JM109 competent cells (Promega) were transformed with 2 μ l of the ligation mix as described in 2.2.1.2. L-agar plates supplemented with 100 μ g/ml ampicillin, 0.5 M IPTG and 40 μ g/ml X-Gal were plated with 100 μ l transformation mix in order to carry out blue/white screening. Positive single white colonies were picked and grown overnight in 5 ml of L-Broth supplemented with 100 μ g/ml ampicillin at 200 rpm at 37°C. Plasmids were isolated using QIAprep miniprep kits (Qiagen) as manufacturer's recommendations and 2 μ l used to check for the correct sequence by DNA sequencing (2.1.6).

2.4.2 Manual sequencing

2.4.2.1 Denaturing plasmid DNA

JM109 cells, containing the correct plasmid, were inoculated into 5 ml of L-Broth supplemented with 100 μ g/ml ampicillin. Plasmids was isolated using the QIAprep miniprep kits (Qiagen) according to the manufacturer's recommendations and eluted with 50 μ l ddH₂O. Plasmid DNA (32 μ l) was denatured with 0.1 volume of 2 M NaOH and 0.1 volume of 2 mM EDTA and incubated at room temperature for 5 min. DNA was precipitated by the addition of 0.1 volume of 3 M sodium acetate, pH 5.2 and 2.5 volumes of ice-cold ethanol (> 99%) and incubating at -80°C for at least 30 min. DNA was pelleted by centrifugation at 16000 x g for 25 min at 4°C. The supernatant was discarded and pellets were then washed with 100 μ l of 70% (v/v) ethanol. Samples were centrifuged for 5 min at 16000 x g at 4°C, and the supernatant was discarded. DNA pellets were then dried in a speedvac for approximately 10 min or until dry.

2.4.2.2 Sequencing of DNA templates

Nucleotide sequence was analysed by a modification of the dideoxy chain termination method using a commercially available kit (T7 Sequenase v 2.0, Amersham) from which all reagents were taken unless otherwise stated.

The denatured plasmid was resuspended in 7 µl of ddH₂O and combined with 2 µl of T7 sequenase reaction buffer and 1 µl of primer (0.5 pmol/µl, listed in appendix A.4). The DNA:plasmid mixture was heated to 65°C for 2 min and cooled slowly to room temperature to anneal the two species.

DNA was labelled by the addition of 1 µl of 0.1 M DTT, 2 µl of labelling mix (1:10 dilution), 0.5 µl of [α^{35} S] dATP (10 µCi/µl, Amersham) and 2 µl of T7 polymerase (diluted 1:10) and incubated for 5 min at room temperature. The reaction was terminated by transferring 3.5 µl of labelling reaction to a nucleotide mix containing either dideoxy dATP, dTTP, dGTP or dCTP. This reaction mix was then incubated for 5 min at 37°C before 4 µl of stop solution was added. Samples were heated to 90°C for 2 min immediately before loading 3 µl onto a sequencing gel.

2.4.3 Analysis of DNA sequence by polyacrylamide sequencing gels

A 8% polyacrylamide sequencing gel was prepared using GEL-MIX[®] 8 (Invitrogen) and the Sequi-Gen[®] GT nucleic acid electrophoresis cell (38 x 50 cm, BioRad). Samples were loaded carefully and the gel electrophoresed at 75W (constant power) for approximately 2 hr or until the frontline of the bromophenol blue, a component of the stop solution, was approximately 5-7 cm from the bottom of the gel. The gel was then stopped and transferred to 3MM filter paper, covered with Clingfilm and

any excess was trimmed away. The gel was then dried for 1hr under vacuum at 80°C. The gel was analysed either by autoradiography, using film (Kodak BioMax film, Amersham), or by phosphoimaging screens.

2.4.4 Primer extension

2.4.4.1 Radiolabelling of primer

Ten units of T4 polynucleotide kinase (T4 PNK, Promega) was used to label 2 µl of primer (0.5 pmol/µl) with 3µl of [γ -³²P] ATP (3000 Ci/mmol, Amersham) and 1 µl of 10X T4 PNK buffer in a final volume of 10 µl. The reaction mix was incubated at 37°C for 10 min and then terminated by heating to 90°C for 2 min. The resulting labelled primer was diluted 1:10 with DEPC ddH₂O.

2.4.4.2 Primer extension

RNA (10-100 µg/3.3 µl) was isolated as described in 2.3.1. [γ -³²P]-labelled primer (1 µl) and 1 µl of 5X first strand buffer (Invitrogen) were added to the RNA sample. The primer and RNA mixture was heated to 90°C for 2 min in order to linearise the RNA. The primer was then annealed to the RNA by slowly cooling to 42°C over a period of 1 hr. Samples were placed briefly on ice. Primer extension was carried out by the addition of 1 µl of 5X first strand buffer (Invitrogen), 0.7 µl of 40 mM sodium pyrophosphatase (Promega) and 0.5 µl of Superscript II reverse transcriptase (RT, Invitrogen) and incubated for 90 min at 42°C. The RT was inactivated by heating the samples to 90°C for 10 min. The RNA was removed by incubating the primer extension reaction with 1.5U of RNase H (Invitrogen) for 30 min at 37°C. T7 sequenase loading dye (6.7 µl) was added to each sample, which were then heated to

90°C for 2 min. The samples were allowed to cool briefly before loading 4 µl of each sample was loaded onto a 8% polyacrylamide gel (2.4.3).

Gels were electrophoresed, dried and subjected to autoradiography or exposed to phosphoimager screens as described in 2.4.3.

2.5 Heat shock of *S. aureus*

The *S. aureus* strain 8325 PS47 (36) 8409 * 47 (Public Health Laboratory Service, London, appendix A.3) was streaked onto Brain Heart Infusion (BHI) agar (20% (w/v) infusion from calf brains, 25% (w/v) infusion from beef heart, 1% (w/v) proteose peptone, 0.2% (w/v) dextrose, 18 mM Na₂HPO₄, 85.6 mM NaCl, 1.5% (w/v) agar). Single *S. aureus* colonies were inoculated into 15 ml of Tryptic Soy Broth (TS-Broth: 1.7% (w/v) pancreatic digest of casein, 0.3% (w/v) papaic digest of soybean meal, 0.25% (w/v) dextrose, 85.6 mM NaCl, 0.15 mM K₂HPO₄, pH 7.3) and incubated overnight at 37°C at 200 rpm. A 1:50 dilution of overnight culture was inoculated into 400 ml of fresh TS-Broth and grown at 30°C at 200 rpm, until the OD₆₀₀ reached 0.5. The culture was then split into seven flasks containing 50 ml. At time points, 0, 15, 30, 45, 60, 90, 120 min, a flask was removed and the bacteria recovered by centrifugation at 3000 x g for 10 min. Pellets were immediately frozen on dry ice and stored at -80°C until required.

RNA was isolated as described in 2.3.1, which were subsequently separated on a RNA denaturing gel (2.3.3) and analysed by Northern blotting (2.3.4).

2.6 Construction of *S. typhimurium* mutants

Mutants were constructed in the attenuated *S. typhimurium* SL3261 strain (genotype listed in appendix A.3) by replacing the gene of interest with a resistance marker. The mutation was then transduced into a virulent strain for *in vivo* studies.

2.6.1 Allelic replacement

S. typhimurium mutants were constructed by a modification of the ET-cloning procedures (Muyrers, J.P.P. *et al.*, 1999, Yu, D. *et al.*, 2000, Zhang, Y. *et al.*, 2000) patented by Arrow Therapeutics Ltd. Single mutants were obtained using a kanamycin resistance cassette. Double mutants were created by two rounds of replacement, firstly with gentamycin and then with kanamycin. Two of the genes of interest (*hhoA* and *hhoB*) lie adjacent on the *S. typhimurium* chromosome and this fact was exploited by reducing the number of rounds required for the construction of a triple mutant.

2.6.1.1 Template DNA for resistance cassettes

The kanamycin resistance cassette was excised from plasmid pUC18:KAN (Arrow Therapeutics) with 10U of *Xba*I (Roche) in buffer H (Roche) in a final volume of 100 µl. Digests were performed at 37°C for 1hr followed by purification of the resistance cassette by gel extraction (QIAquick gel extraction kit, Qiagen). Purified DNA (5 µl) was analysed on a 1% agarose gel (2.1.3) to confirm the presence of the DNA of interest.

The gentamycin resistance cassette was obtained from the plasmid pBSL182 (Netherlands Culture Collection of Bacteria) and the chloramphenicol resistance

cassette was obtained from the plasmid pACYC184 (Northumberland Biologicals Ltd).

2.6.1.2 Construction of resistance cassettes

Resistance cassettes (kanamycin or gentamycin) were amplified with 5' and 3' arms (60-mers) complementary to the flanking regions to the gene of interest. Amplification was performed by PCR using a proofreading DNA polymerase (2.1.2). MgCl₂ concentrations (0 to 4 mM) were first optimised to obtain the maximum yield of products using 10 pmol of each primer pair (listed in appendix A.7) and 1 µl template DNA. PCR was performed by denaturing the DNA at 94°C for 6 min, followed by 30 cycles of PCR: denaturation at 94°C for 1 min, hybridisation at 60°C for 1 min and extension at 72°C for 2 min. With the optimised MgCl₂ concentration, 2 x 50 µl PCR reactions were carried out.

PCR products were purified using QIAquick PCR purification kits (Qiagen) for the kanamycin cassette and gel extraction (QIAquick gel extraction kit, Qiagen) for the gentamycin cassette.

*2.6.1.3 Transformation of *S. typhimurium* SL3261 competent cells with pBADλRED*

S. typhimurium SL3261 (kindly Professor Duncan Maskell, University of Cambridge) was prepared as electrocompetent cells (prepared as 2.2.1.1) and was transformed with 1 µl pBADλRED (kindly donated from Arrow Therapeutics). Colonies containing pBADλRED were selected on L-agar supplemented with 100 µg/ml ampicillin.

2.6.1.4 Induction of *S. typhimurium* SL3261 pBAD λ RED

A single mutant colony was inoculated into 5 ml of L-Broth supplemented with 100 μ g/ml ampicillin and grown overnight at 37°C, 250 rpm. Overnight cultures of each mutant were diluted 1:100 into 50 ml of fresh L-Broth supplemented with 100 μ g/ml ampicillin and grown until the OD₆₀₀ reached 0.25. Cultures were treated with 0.2% (w/v, final concentration) of arabinose (Sigma) in order to induce the bacteria. Cultures were then grown until the OD₆₀₀ reached 0.5. The cultures were then prepared as electrocompetent cells (2.2.1.1a) except that the electrocompetent cells were resuspended in 50 μ l of 10% (v/v) glycerol. Varying volumes of insert DNA (5 or 10 μ l) were transformed into the bacteria as described in 2.2.1.1b. Following transformation, the samples were incubated for 2 hr at 37°C at 250 rpm. The cells were pelleted and then were resuspended into 100 μ l of L-Broth, which was then plated onto L-agar or BHI agar plates supplemented with 25 μ g/ml kanamycin (Sigma), 20 μ g/ml gentamycin (Sigma) or 25 μ g/ml chloramphenicol (Sigma). Plates were then incubated overnight at 37°C.

2.6.1.5 Verification of allelic replacement

Several methods were performed to verify that the resistance marker had undergone recombination into the genomic DNA, including colony PCR and DNA sequencing.

a) Colony PCR

This was performed as described in 2.1.5.1 using 10 pmol of each primer (listed in appendix A.8), which amplified segments from within the kanamycin cassette to the

flanking chromosomal DNA. For the gentamycin and chloramphenicol resistance cassettes, primers corresponding to the chromosomal DNA flanking the cassette were used, as size difference was sufficient to distinguish mutant from wild-type PCR product.

b) Sequencing amplified DNA

To confirm that the resistance gene was inserted correctly into the chromosomal DNA to create the desired mutant, the PCR products obtained from the colony PCR (2.6.1.4a) were TA-cloned (Invitrogen) following manufacturer's recommendations. This method ligated the TA-overhangs of the PCR product into the *lacZ* gene in the pTOPO plasmid. The ligation mix was then transformed into One Shot TOP10 competent cells (Invitrogen) by heat shock at 42°C for 30 sec and then plated onto L-agar plates supplemented with 100 µg/ml of ampicillin and 40 µl of 40 mg/ml X-Gal. Single, positive white colonies were inoculated into 5 ml of L-Broth and grown overnight at 37°C at 200 rpm. Cultures were pelleted at 3000 x g for 10 min and supernatant discarded. Plasmids were purified from these pellets using the QIAprep miniprep kit (Qiagen) according to manufacturer's recommendations. The plasmids (2 µl each) were then sequenced as described in 2.1.6 using primers for the M13 forward and reverse primers (Invitrogen).

2.6.2 Transduction of mutation into *S. typhimurium* strain SL1344

The mutations created in the attenuated *S. typhimurium* strain SL3261 were transduced into the mouse virulent strain SL1344 (kindly donated by Professor Duncan Maskell, University of Cambridge) for *in vivo* studies. This was performed by Dr. Sarah Peters (Arrow Therapeutics Ltd, London).

2.6.2.1 Plate stock prep

A single mutant colony derived from the SL3261 strain (donor colony) was inoculated into 5 ml of L-Broth supplemented with the appropriate antibiotics and grown overnight at 37°C. P22 phage was diluted to a variety of dilutions (1:20 to 1:100000) in SM buffer (100 mM NaCl, 8 µM MgSO₄·7H₂O, 0.01% (w/v) gelatin, 50 mM Tris-HCl, pH 7.5). To a volume of 200 µl overnight donor culture, 100 µl of diluted P22 phage was added and the samples were mixed briefly. Phage was allowed to adsorb to the bacteria by incubating at 37°C for 30-40 min. Top agar (137 mM NaCl, 1% (w/v) tryptone, 0.7% agar) was pre-warmed to 42°C and 3-5 ml was added to the samples, mixed briefly and poured onto L-agar plates supplemented with the appropriate antibiotics. Plates were incubated for 3-6 hr until the plaques became visible. SM buffer (2 ml) was added to semi-confluent plates, which were rocked overnight at 4°C to elute the phage from the agar. The buffer was collected and the entire top agar was removed and added to it. Chloroform (1% v/v final concentration) was added to lyse any remaining bacteria. The agar was pelleted by centrifugation at 10000 x g for 15 min. The supernatant was removed and clarified through a 0.22 µm filter.

2.6.2.2 Transduction

A single colony of the recipient strain (SL1344) was inoculated into L-Broth and grown overnight at 37°C. Increasing amounts of plate stocks (10 to 300 µl), which had been prepared in 2.6.2.1, were added to 200 µl of overnight recipient culture and mixed briefly. Phage was allowed to adsorb to the bacteria by standing at 37°C for 1 hr. Culture was then plated onto L-agar supplemented with the appropriate

antibiotics and 5 mM ethylene glycol-bis(2-aminoethylether)-N,N,N',N'-tetraacetic acid (EGTA, Sigma), which was added to inhibit further phage adsorption. Samples were then incubated overnight at 37°C. Single colonies were streaked onto a fresh plate of L-agar supplemented with the appropriate antibiotics and 5 mM EGTA. This was repeated once more before finally streaking onto L-agar supplemented with the appropriate antibiotics.

Colonies were screened for agglutination with *Salmonella*-specific antisera by Dr. Sarah Peters (Arrow Therapeutics).

Colonies were also screened for the appropriate mutation by PCR with 1 µl of heat-inactivated overnight culture (diluted 1:100 in PBS) using the same conditions and primers as 2.6.1.5a.

2.7 Protease detection using zymograms

2.7.1 Zymograms

Zymograms are essentially protein gels which contain a protease substrate, in this case, β -casein. Zymograms was prepared as described for 12.5 % SDS-PAGE protein gels (2.2.2.1) except no SDS was added to the gels. In addition, β -casein was added to a final concentration of 0.05% (w/v). A 4% stacking gel (2.2.2.1) was set on top of the resolving gel. Samples were diluted in 1.5X solubilisation buffer (62.5 mM Tris-HCl pH 6.8, 25% (v/v) glycerol, 0.04% (w/v) SDS, 0.1% (w/v) bromophenol blue) and incubated for 10 min at room temperature. The proteins were separated by electrophoresis at a constant voltage of 100V for 90 min in electrode buffer (25 mM Trizma base (Sigma), 192 mM glycine, 0.1% (w/v) SDS).

After electrophoresis, the gel apparatus was disassembled and the gels were incubated in renaturation buffer (2.5% (v/v) Triton-X100) for 30 min at room temperature with gentle agitation. The renaturation buffer was discarded and the gels were briefly washed in dH₂O before the developing buffer (50 mM Tris-HCl pH7.5, 200 mM NaCl, 5 mM CaCl₂·2H₂O) was added. The gels were then incubated for 16 hr at 37°C with gentle agitation. After the incubation, the developing buffer was discarded and the gels were stained for at least 1 hr with 0.5% (w/v) Coomassie in 40% methanol and 10% acetic acid. The gels were then destained (20% methanol, 10% acetic acid) until the protease bands could be visualised.

2.8 *S. typhimurium* stress conditions in vitro

The viability of the *S. typhimurium* SL3261 mutants under stress conditions was investigated using heat shock and oxidative stress conditions.

2.8.1 Heat shock studies

A single mutant colony was inoculated into 1 ml of L-Broth supplemented with the appropriate antibiotics and grown overnight at 37°C, 250 rpm. A 1:200 dilution of overnight culture was used to inoculate 2 ml of fresh L-Broth and grown for 8 hr at 37°C at 250 rpm. Cultures were diluted to give an OD₆₀₀ of 0.015 and 200 µl was aliquoted per well in a 100-well honeycomb plate (Thermo Life sciences Ltd). This was incubated for 16 hr in a Labsystems Bioscreen C machine (Labsystems) using the conditions outlined in Table 2.8.

Filter	Single
Wavelength	600 nm
Temperature	42°C
Preheat	6 min
Length of incubation	16 hr
OD ₆₀₀ reading at	15 min intervals
Shaking	Medium for 10 sec before reading

Table 2.8: Bioscreen conditions for heat shock of *S. typhimurium* mutants

After 16 hr, 10 µl aliquots were removed and serially diluted before plating onto L-agar plates to determine the viability of the mutants after 16 hr of heat shock.

Controls for each mutant were carried out in a similar manner except the incubation temperature was 37°C.

2.8.2 Oxidative stress of *S. typhimurium* SL3261 mutants

The ability of the *S. typhimurium* SL3261 mutants to withstand oxidative stress caused by reactive oxygen species (ROS) was investigated using H₂O₂, a source of hydroxyl radicals. Investigations into the effect of increasing concentrations of H₂O₂ were performed in a similar manner to the heat shock experiments (2.8.1) except increasing concentrations (0 to 10 mM) of H₂O₂ (10 µl) was added. In addition the incubation temperature was 37°C.

2.9 Attenuation of *S. typhimurium* SL1344 mutants *in vivo*

The viability of the *S. typhimurium* SL1344 mutants *in vivo* was investigated in a mouse model of infection.

2.9.1 Quantification of *S. typhimurium* SL1344 mutants

The number of bacteria needed to be quantified so that the same number of mutants is introduced into each mouse. This was done by viable counts on overnight culture. Briefly, a single colony was inoculated into 10 ml of L-Broth supplemented with the appropriate antibiotics and incubated overnight at 37°C. The OD₆₀₀ was measured and then the culture was diluted and plated onto L-agar to determine the number of viable colonies present. This procedure gives a good indication of the number of colonies at a specific OD₆₀₀, and therefore was subsequently used as a comparison in future growth experiments.

2.9.1.1 *Inoculation of BALB/c mice*

A single colony was inoculated into 10 ml of L-Broth supplemented with the appropriate antibiotics and stood overnight at 37°C. Bacteria were pelleted by centrifugation at 3000 x g for 10 min followed by resuspension in 10 ml of PBS. Samples were then diluted to approximately 10³ bacteria per 200 µl.

Four BALB/c mice (Harlan) were used per mutant per time point. A volume of 200 µl of bacteria culture was injected into the tail vein of each mouse. Four mice per group were culled over a period of seven days post-infection. The spleen and livers were removed and placed into 10 ml of PBS and disrupted using a Stomacher® 80 Lab System (Seward).

Procedures involving animals were carried out by Dr. Sarah Peters and Ms Chrissie Willers (Arrow Therapeutics Ltd, London).

2.9.1.2 Viable counts

The number of viable *S. typhimurium* mutants was quantified by plating disrupted spleen and liver samples (2.9.1.1) onto L-agar. A volume of 1 ml or 5 ml disrupted liver or spleen samples was added to 20 ml molten L-agar, mixed briefly and poured. In addition, agar dilutions were performed whereby, serial dilutions were carried out in molten L-agar and 500 µl plated onto petri dishes. Plates were incubated overnight at 37°C and the number of colonies counted.

2.10 Statistical analysis

Statistical analysis was carried out by comparing the raw data using the Student t-test in Excel. P-values less than 0.05 were deemed significant.

CHAPTER 3. IDENTIFICATION OF HTRA
HOMOLOGUES IN *S. aureus* AND *S. typhimurium*

3.1	Introduction.....	108
3.2	Identification of HtrA orthologues in <i>S. typhimurium</i>	109
3.3	Identification of HtrA orthologues in <i>S. aureus</i>	113
3.4	Summary.....	119

3.1 Introduction

HtrA is a periplasmic serine protease identified to be an important factor in the maintenance of viability when bacteria are exposed to stress conditions. During these conditions, misfolded proteins may accumulate and consequently become toxic to the bacteria. HtrA is reported to remove misfolded proteins by proteolysis and thus promoting bacteria survival (Pallen, M.J. & Wren, B.W., 1997). This is achieved by the serine protease activity of the enzyme, however, at the lower temperature of 28°C, HtrA has chaperone re-folding activity (Speiss, C. *et al.*, 1999).

HtrA is constructed of three domains, a N-terminal subcellular localisation domain, a catalytic serine protease domain, containing the catalytic triad, and C-terminal PDZ domains. In *E. coli*, a family of HtrA paralogues were identified, and named HhoA and HhoB (Figure 3.1). HhoA has a similar domain structure to HtrA, with two PDZ domains, whereas HhoB is a smaller protein and has a single PDZ domain (reviewed in Pallen, M.J. & Wren, B.W., 1997).

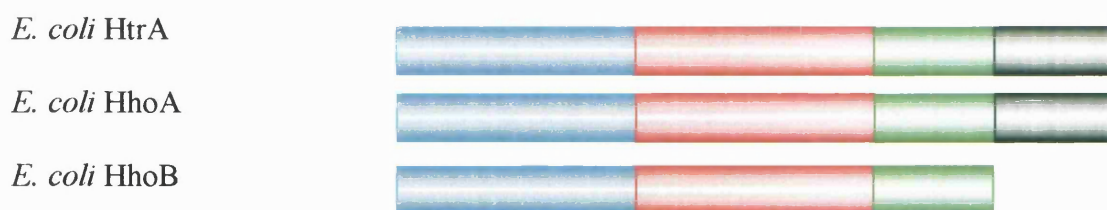


Figure 3.1 – Schematic diagram of the protein domains of the HtrA orthologues in *E. coli*

The N-terminal subcellular domain is coloured cyan, while the catalytic domain is coloured red. The PDZ domains are coloured light green (PDZ1) and dark green (PDZ2).

The majority of information about HtrA and its orthologues has been obtained from Gram-negative bacterial species. In recent years, six Gram-positive bacteria species have been published to possess a gene that has high homology to HtrA. This chapter describes the identification of HtrA homologues in the pathogenic Gram-negative bacterium, *S. typhimurium*, and the pathogenic Gram-positive bacterium, *S. aureus*.

3.2 Identification of HtrA orthologues in *S. typhimurium*

E. coli HtrA orthologues were identified in *S. typhimurium* LT2 strain using the HtrA paralogue sequences of *E. coli* as the query sequences. Three hits (STHtrA, STHhoA, STHhoB) were obtained using the BLAST algorithm (<http://www.ebi.ac.uk>), which corresponded to the three HtrA homologues in *E. coli*.

Figure 3.2 illustrates the protein alignments of each pair of homologues using pairwise alignment (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>). All three *S. typhimurium* homologues show a high degree of identity (88% STHtrA; 90% STHhoA; 92% STHhoB) and similarity (94% STHtrA; 98% STHhoA; 97% STHhoB) at the protein level when compared with their *E. coli* counterparts.

STHtrA has a predicted open reading frame (ORF) of 1,428-bp which encodes a gene product of 475 amino acids. The gene is located in the genome at positions 244,484 – 245,917 on the plus strand. Further downstream, STHhoA and STHhoB lie adjacent on the genome (STHhoA: 3,515,488 - 3,516,756; STHhoB: 3,516,852 – 3,517,919) separated by 136-bp. STHhoA comprises a predicted ORF of 1,368-bp encoding a gene product of 455 amino acids whereas STHhoB is smaller (ORF

a) Protein sequence alignment of STHtrA versus *E. coli* HtrA

E.coli HtrA	MKKTLLA	SR	LALSGLALS	PLSATAAETS	SAT	TAQQMP	SLAPMLEKVM	PSVVSINVEG	STTVNTPRMP	69
STHtrA	MKKTLLA	SA	LALSGLALS	PLSATAAETS	SSAM	TAQQMP	SLAPMLEKVM	PSVVSINVEG	STTVNTPRMP	70
E.coli HtrA	RNFQQFFGDD	SPFCQEGSPF	QSSPFCQGGQ	GGNGGGQQQK	FMALGSGVII	DADKGYVVTN	NHVVDNATVI	139		
STHtrA	RNFQQFFGDD	SPFCQDGSPI	QNSPFCQGGG	NGNGGGQQQK	FMALGSGVII	DAAKGYVVTN	NHVVDNASVI	140		
E.coli HtrA	KVQLSDGRKF	DAKMYGKDPR	SDIALIQIQN	PKNLTAIKMA	DSDALRVGDY	TVGIGNPFGL	GETVTSGIVS	209		
STHtrA	KVQLSDGRKF	DAKVVYKDPK	SDIALIQIQN	PKNLTAIKLA	DSDALRVGDY	IYVIGNPFGL	GETVTSGIVS	210		
E.coli HtrA	ALGRSGLNAE	NYENFIQTDA	AINRGNSSGA	LVNLNGELIG	INTAILAPDG	GNIGIGFAIP	SNMVKNLTSQ	279		
STHtrA	ALGRSGLNVE	NYENFIQTDA	AINRGNSSGA	LVNLNGELIG	INTAILAPDG	GNIGIGFAIP	SNMVKNLTSQ	280		
E.coli HtrA	MVEYGQVKRG	ELGIMGTELN	SELAKAMKVD	AQRGAFFVSQV	LPNSSAAKAG	IKAGDVITSL	NGKPISSFAA	349		
STHtrA	MVEYGQVKRG	ELGIMGTELN	SELAKAMKVD	AQRGAFFVSQV	MPNSSAAKAG	IKAGDVITSL	NGKPISSFAA	350		
E.coli HtrA	LRAQVGTPMV	GSKLTGLLR	DGKQVNVNLE	LQSSQNQVD	SSSIFNGIEG	AEMSNKCKDQ	GVVVNNVKTG	419		
STHtrA	LRAQVGTPMV	GSKISGLLR	EGKAITVNLE	LQSSSQVD	SSTIFSGIEG	AEMSNKCKDK	GVVVSSVKAN	420		
E.coli HtrA	TAAQIGLKK	GDVVIIGANQQ	AVKNIAELRK	VLDKSPSVLA	LNIQRGDRHL	PVNAVISLNP	FLKTGRGSPY	489		
STHtrA	SPAAQIGLKK	GDVVIIGANQQ	PVKNIARELR	ILDSKPSVLA	LNIQRGD	---S--S---	YL-----L	473		
E.coli HtrA	NL	491								
STHtrA	MQ	475								

b) Protein sequence alignment of STHhoA versus *E. coli* HhoA

E.coli HhoA	MKKQTQLLSA	LALSGLTSL	ASFCAYASIP	GQVADQAPLP	SLAPMLEKVL	PAVVSVRVEG	TASQGGKIPE	70
STHhoA	MKKHTQLLSA	LALSGLTSL	APFALASIP	GQVPGQATLP	SLAPMLEKVL	PAVVSVKVEG	TAAQSQKVPE	70
E.coli HhoA	EFKKFFGDDL	PDQPAQPFEG	LGSGVIINAS	KGYVLTNNHV	INQAQKISIQ	LNDGREFDAK	LIGSDDQSDI	140
STHhoA	EFKKFFGEDI	PDQPSQPFEG	LGSGVIDAA	KGYVLTNNHV	INQAQKISIQ	LNDGREFDAK	QIGGDDQSDI	140
E.coli HhoA	ALLQIQNPSSK	LTQIAIADSD	KLRVGDFAVA	VGNPFLGQT	ATSGIVSALG	RSGLNLEGLE	NFIQTASIN	210
STHhoA	ALLQIQNPSSK	LTQIAIADSD	KLRVGDFAVA	VGNPFLGQT	ATSGIISALG	RSGLNLEGLE	NFIQTASIN	210
E.coli HhoA	RGNSGGALLN	LNGLIGINT	AIALPGGGSV	GIGFAIPSNM	ARTLAQQLID	FGEIKRGLLG	IKGTEMASDI	280
STHhoA	RGNSGGALLN	LNGLIGINT	AIALPGGGSV	GIGFAIPSNM	AQTLAQQLIQ	FGEIKRGLLG	IKGTEMATDI	280
E.coli HhoA	AKAFNLDVQR	GAFFSEVLPN	SGSAKAGVKA	GDVITSLNGK	PLNSFAELRS	RIATTEPGTK	VKLGLLRNGK	350
STHhoA	AKAFKLNVR	GAFFSEVLPN	SGSAKAGVKS	GDVITSLNGK	QLNSFAELRS	RIATTEPGTK	VKLGLLRDQK	350
E.coli HhoA	PLEVEYTLDT	STSSASAEM	ITPALEGATL	SDGQLKDGK	GKVIDEVYKG	SPAAQAGLQK	DDVIGVNRD	420
STHhoA	PLEVEYTLDS	NTSSASAEM	IAPALQATL	SDGQLKDGK	GKVIDSVYKS	SPAAQAGLQK	DDVIGVNRD	420
E.coli HhoA	RVNSIAEMRK	VLAAPKATIA	LQIVRGNESI	YLLLR	455			
STHhoA	RISIAEMRK	VMAAKPSIA	LQIVRGNENI	YLLLR	455			

c) Protein sequence alignment of STHhoB versus *E. coli* HhoB

E.coli HhoB	MFVKLLRSVA	IGLIVGAILL	VAMPSLRSLN	PLSTPQFDST	DETPASYNLA	VRRAPAVVN	VYNRGLNTNS	70
STHhoB	MFVKLLRSVA	IGLIVGAILL	AVMPSLRKIN	PIAVPQFDST	DETPASYNFA	VRRAPAVVN	VYNRSANSTA	70
E.coli HhoB	HNQLEIRTLG	SGVIMDQRGY	IITNKHVIND	ADQIIVALQD	GRVFEALLVG	SDSLTDLAVL	KINATGGLPT	140
STHhoB	HNQLEIRTLG	SGVIMDQRGY	IITNKHVIND	ADQIIVALQD	GRVFEALLVG	SDSLTDLAVL	KINATGGLPT	140
E.coli HhoB	IPINARRVPH	IGDVVLAIGN	PYNLGQTITQ	GIISATGRIG	LNPTGRQNFL	QTASINHG	SGGALVNSLG	210
STHhoB	IPINTKRTPH	IGDVVLAIGN	PYNLGQTITQ	GIISATGRIG	LNPTGRQNFL	QTASINHG	SGGALVNSLG	210
E.coli HhoB	ELMGINTLSF	DKSNDGETPE	GIGFAIPFQL	ATKIMDKLIR	DGRVIRGYIG	IGGREIAPLH	AQGQGGIDQL	279
STHhoB	ELMGINTLSF	DKSNDGETPE	GLGFAIPFQL	ATKIMDKLIR	DGRVIRGYIG	IGGREIAPLH	AQGGSGMDPI	280
E.coli HhoB	QGIVVNEVSP	DGPAANAGIQ	VNDLIISVYN	KPAISALETM	DQVAEIRPGS	VIPVVVMRDD	KQLTLQVTIQ	349
STHhoB	QGIVVNEVTP	DGPAANAGIQ	VNDLIISVYN	KPAVSALETM	DQVAEIRPGS	VIPVVVMRDD	KQLTLQVTIQ	350
E.coli HhoB	EYPATN	355						
STHhoB	EYPASN	356						

Figure 3.2 – Protein sequence alignment of *S. typhimurium* HtrA orthologues versus their *E. coli* counterparts (amino acid identity – blue; amino acid similarity – yellow)

1,071-bp) encoding a 356 amino acid protein. Although STHtrA has been identified in *S. typhimurium* LT2 strain (Accession number (Acc#) P26982) and characterised in *S. typhimurium* C5 strain (Johnson, K. *et al.*, 1991), STHhoA (Acc # NP 462258, AAL22217) and STHhoB (Acc# NP 462259) are only been annotated in the protein databases to be serine proteinases.

The high level of identity between the *S. typhimurium* and *E. coli* counterparts indicates that the arrangement of protein domains is likely to be identical (Figure 3.3). All *S. typhimurium* HtrA orthologues have a catalytic domain (residues 114-272 STHtrA; 91-249 STHhoA; 79-239 STHhoB, boxed in red) with the typical catalytic triad of serine proteases. The catalytic triad are followed by two PDZ domains (boxed in green) in HtrA (residues 273-381: PDZ1; 382-475: PDZ2) and HhoA (residues 250-358: PDZ1; 359-455) homologues, whereas the HhoB homologues have only the one PDZ domain (residues 240-356). The N-terminal subcellular location determinant domain is very similar between the *E. coli* HtrA, STHtrA and HhoA, suggesting that in common with the *E. coli* HtrA, they exist within the periplasm (reviewed in Pallen, M.J. & Wren, B.W., 1997). However, the N-terminal subcellular domain of STHhoB differs from STHtrA and HhoA suggesting that it could exist within another cellular compartment, although to date, this is unknown. Analysis of the sequences also shows that there are a segment of approximately twenty amino acids (residues 80-100) in *E. coli* HtrA and STHtrA which is missing in STHhoA and STHhoB.

The three *S. typhimurium* orthologues were analysed for signal peptide cleavage sites using SignalP (<http://www.cbs.dtu.dk/services/SignalP>). STHtrA was predicted to

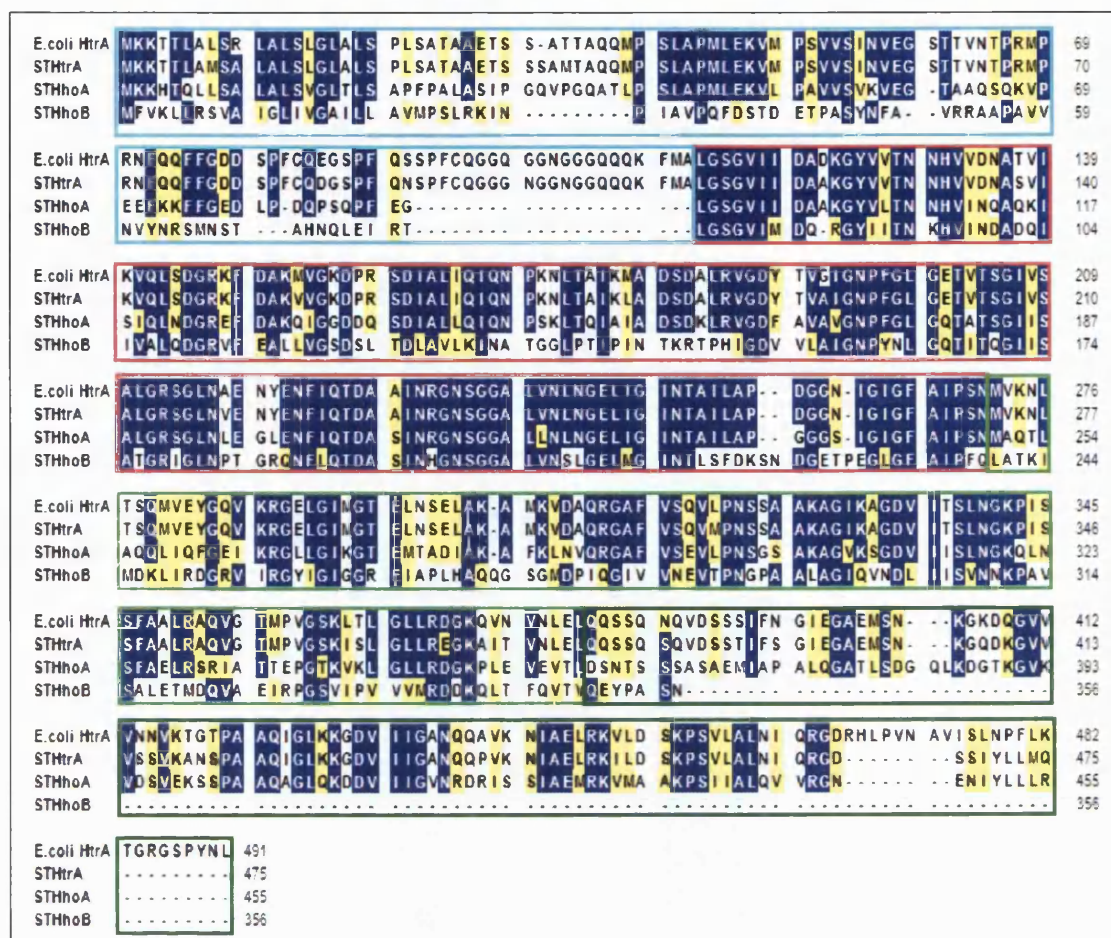


Figure 3.3 – Protein sequence alignment of the HtrA orthologues in *S. typhimurium* versus *E. coli* HtrA

Protein sequences were aligned using ClustalW (within Bioedit) and identical residues (blue background) and similar residues (yellow background) were highlighted. Protein domains have been outlined – the N-terminal domain (cyan), the catalytic domain (red), PDZ1 (light green) and PDZ2 (dark green).

have a signal peptide cleavage site between positions A²⁶ and A²⁷ (p = 0.992). STHhoA was also predicted to have a signal peptide cleavage site between positions A²⁷-S²⁸ (p = 0.996). In *E. coli*, both HtrA (Pallen, M.J. & Wren, B.W., 1997) and HhoA (Waller, P.R.H. & Sauer, R.T., 1996) are reported to have signal peptide cleavage sites in similar positions and so these predictions are likely to function *in vivo*. However, HhoB has not been shown to possess a signal peptide (Waller, P.R.H. & Sauer, R.T., 1996) and in agreement to this STHhoB has a lesser probability of 0.657 for a cleavage site at position 28.

3.3 Identification of HtrA orthologues in *S. aureus*

Three DNA sequences (SAHtrA1, SAHtrA2 and SAHtrA3) were identified to be HtrA-like homologues in the unfinished genome sequence of *S. aureus* NCTC 8325 strain, using the HtrA homologue of *S. typhimurium* as the query sequence. This study concentrates on SAHtrA1 and SAHtrA2, although SAHtrA3 is discussed briefly later. Figure 3.4 shows the protein alignments of the *S. aureus* homologues in comparison to the HtrA homologue of *S. typhimurium*.

SAHtrAs shows a high degree of identity (42% SAHtrA1; 35% SAHtrA2) and similarity (64% SAHtrA1; 57% SAHtrA2) at the protein level when compared with the catalytic and PDZ domains of STHtrA. SAHtrA1 comprises a predicted ORF of 1,275-bp encoding a gene product of 425 amino acids agreeing with a previous observation (Jones, C.H. *et al.*, 2001). SAHtrA2 has a predicted ORF of 2,325-bp with a gene product of 775 amino acids. Recently, a second HtrA-like homologue was reported to be encoded in the *S. aureus* genome (Jones, C.H. *et al.*, 2001, but this sequence had 413 amino acids missing at the N-terminus compared to the

a) Protein sequence alignment of SAHtrA1 versus *S. typhimurium* HtrA

STHtrA	---MKKTTLA	MSALA-LS-L	GL-ALSPLSA	TAAETSSSAM	TQQMPSLAP	MLEKVMPVV	SINVEGSTTV	54
SAHtrA1	MSDFNHIDHS	TTNHSQTPRY	RRPKFPWFKT	VIVALIAGII	GALLVLGIGK	VLNSTILNKD	GSTVQTNNK	70
STHtrA	NTPRM---PR	NF---QQFFG	DDSPFCQDGS	PFNSPFCQG	GGNGNGGQQ	QKFMALGSGV	I--IDAAGKY	126
SAHtrA1	GGNQLDQSK	KFGTVHEMIK	SWSPITVGV	NMQKAS-SVD	DLKSKSSKP	SE-AGVGSGV	IYQINNNSAY	138
STHtrA	VVTNNHVVDN	ASVIVVQLSD	GRKFDKVVVG	KDPRSDIALI	QIQNPKNLTA	IKLADSDALR	VGDYTVAIQN	196
SAHtrA1	VVTNNHVIDG	ANEIRVQLHN	KKQVKAHLVG	KDAVTRIVVL	KIENTKGIKA	IQFANSSKVQ	TGDSVFRMGN	208
STHtrA	PFGL--GETV	TSGIVSALGR	---SGLNVEN	YE-NFIQTDA	AINRGNSSGA	LVNLNGELIG	INTAILAPDG	260
SAHtrA1	PLGLQFANSV	TSGIISASER	TIDAETTGGN	TKVSVLQTD	AINFGNSGGA	LVDINGNLVG	INSMKIAATQ	278
STHtrA	GNIGIGFAIP	SNMVKNLTSQ	QVEYGVQVKG	ELGIMGTEN	SELAKAM-KV	DACSGAFVSQ	VMPNSSAAKA	329
SAHtrA1	VE-GIGFAIP	SNEVKNLTSQ	LKHKGIKIDP	SIGISLINLK	DIPEEREQL	HTDREDGIY-	W-AK-ADSDI	344
STHtrA	GIKAGGVITS	LNGPISSFA	ALRAQVGTMP	VGSKISLGL	REGKAITVNL	ELQQSSQSQV	DSTIFSGIE	399
SAHtrA1	DLKKGDIITE	IDGKKI----	---KDDVD-LR	-----SYL	YENK-----	-----KPG	ESVTV--TV-	384
STHtrA	GAEMSNKQGD	KGVVSSVDA	NSPAAQIGLK	KGDVIGANQ	QPVKNIAELR	KILDSKPSVL	ALNIQRGDS	469
SAHtrA1	---IRD-Q--	K-----T-KE	V-----KVALK	Q-----QKE	QP-----	KR-QFRSERQ	SPG--QGDRD	421
STHtrA	IYLLMQ	475						
SAHtrA1	-FF--R	424						

b) Protein sequence alignment of SAHtrA2 versus *S. typhimurium* HtrA

STHtrA	-----	-----	-----	-----	-----	-----	-----	1
SAHtrA2	MEWTLVDIGK	KHVIPKSQYR	RKRREFFHNE	DREENLNQHQ	DKQNIDNTTS	KKADKQIHKD	SIDKHERFKN	70
STHtrA	-----	-----	-----	-----	-----	-----	-----	1
SAHtrA2	SLSSHLEQRN	RDVNEKAE	SKSNQDSKSA	YNRDHYLTDD	VSKKQNSLDS	VDQDTEKSKY	YEQNSEATLS	140
STHtrA	-----	-----	-----	-----	-----	-----	-----	1
SAHtrA2	TKSTDKVEST	EMRKLSSDKN	KVGHEEQHVL	SKPSEHOKET	RIDSESSPTD	SDSSMQTEKI	KKDSSDGNKS	210
STHtrA	-----	-----	-----	-----	-----	-----	-----	1
SAHtrA2	SNLKSEVISD	KSNTPVKLSE	SDDEVNNQKP	LTLPEEQKLK	RQSQNEQTK	TYTYGDSEQN	DKSNHENDLS	280
STHtrA	-----	-----	-----	-----	-----	-----	-----	1
SAHtrA2	HHIPISDDK	DNVRENHIV	DDNPONDINT	PSLSKTDDDR	KLDEKIHVED	KHKQNADSS	TVGYQSQSTA	350
STHtrA	-----	-----	-----	-----	-----	-----	-----	1
SAHtrA2	SHRSTEKRN	SINDHDKLNG	QKNTKTSAN	NNQKKATSKL	SLDLALSPLS	ATAETSSSA	MTAQQMPSLA	43
SAHtrA2	SHRSTEKRN	SINDHDKLNG	QKNTKTSAN	NNQKKATSKL	NKGRATNNNY	SDILKKFWMM	YWPCLVILMG	420
STHtrA	PM-LEKVMPS	VV-SINVEGS	TTVNTPRMPR	NFQQFFGDDS	PFCQDGSPPQ	NSPFCQGGGN	GGNGGQQQKF	111
SAHtrA2	IILIVILNA	IFNNVKNDR	MNDNDADAQ	KYTTTMMKAN	NTVKSVVTV	NET-SKDSSL	PKDKASQDE	428
STHtrA	MALGSGVVI	DAAGK-VVVT	NNHVVDNASV	IKVQLSDGRK	FDKVVVGKDP	RSDIALIQIQ	NP-KNLTAIK	178
SAHtrA2	--VSGSVVYK	KSGDTLYIVT	NAHVVDKEN	QKITFSNNKS	VVGKVLGKDK	WSQLAVVKAT	SSDSSVKEIA	556
STHtrA	LADSDALRVG	DYTVAIQNP	GLG--ETVTS	GIVSALGR--	S-GLNVEN-Y	ENFI--QTD	AAINFGNSGG	239
SAHtrA2	IGDSNNLVVG	EPILVGNPL	GVDFKGTVTE	GIVSGLNRNV	PIDFDKDNKY	DMLMKAFQID	ASVNFNSGG	625
STHtrA	ALVNLNGELI	GINTAILAPD	GGNIGTGFAI	PSNMVKNLTS	QMVETQVQR	GELGI--MGT	E-LNSELAKA	306
SAHtrA2	AVVNRGKLI	GVVAAKISMP	NVE-NMSFAI	PVNEVQKIVK	DLETGKIDY	PDVGVKMKNI	VSLNSFERQA	695
STHtrA	MKV--DAQRG	AFVSQVMPNS	SAKAGIKAG	DVITSNGKP	ISSFAALRAQ	VGTMVVGSK	ISLGLIFREGK	373
SAHtrA2	VKLPGKVKNG	VVQDVQVNG	LADQSQLKNG	DVITELGKLG	LEDDLRFQRI	IFSHKDDLMS	ITAKIYRQK	765
STHtrA	AITMNLLELQ	SSQSQVDSST	IFSGIEGAEM	SNKGQDKGVV	VSSVKANSFA	AQIGLKKGDV	IIGANQQPVK	443
SAHtrA2	EKEINIKLK-	-----	-----	-----	-----	-----	-----	774
STHtrA	NIAELRKILD	SKPSVLALNI	QRGDSIIYLL	MQ	475			
SAHtrA2	-----	-----	-----	---	774			

Figure 3.4 - Protein sequence alignment of *S. aureus* HtrA orthologues versus the *S. typhimurium* HtrA (amino acid identity – blue; amino acid similarity – yellow)

SAHtrA2 identified in this study. Two homologues have also been identified in *S. aureus* MW2 strain. One homologue (Acc # BAB95535, NP 646487) is identical to SAHtrA1. The second homologue (Acc # BAB94768, NP 645720) is identical to SAHtrA2 except for some extra residues at the N-terminal sequence (highlighted in bold) of M¹**EWTLVDIGL** in SAHtrA2 instead of M¹DIGL. Interestingly the sequences in the databases are annotated such that the homologue corresponding to SAHtrA2 is recognised as the serine protease HtrA, even though the sequence identity and similarity of SAHtrA1, with respect to STHtrA, is higher. This latter protein, however, is annotated to be the hypothetical protein similar to HtrA.

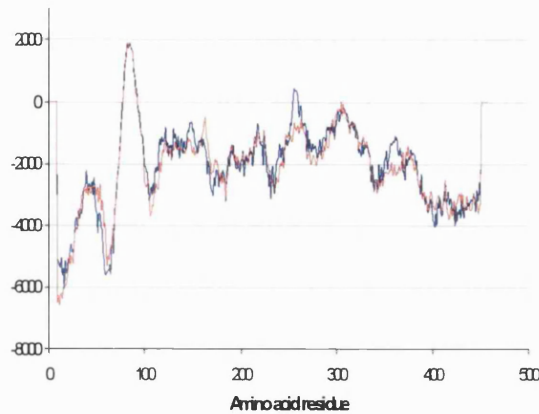
Analysis of the protein sequences (Figure 3.5) shows the presence of a catalytic triad typical of serine protease in both SAHtrAs (G²⁵³ NSGGAL in SAHtrA1; G⁶²²NSGGAV in SAHtrA2). In addition, there is a single PDZ domain following the catalytic domain in both proteins in agreement with observations that identified HtrA homologues in Gram-positive bacteria contain a single PDZ domain (reviewed in Pallen, M.J. & Wren, B.W., 1997, Jones, C.H. *et al.*, 2001). The PDZ domains in the *S. aureus* homologues, however, do not have the typical GLGF consensus sequence but instead have GIGL, agreeing with the observation of a Φ G Φ consensus sequence where two hydrophobic residues (Φ) flank both sides of the second glycine residue (Tochio, H. *et al.*, 2000). The proteins differ significantly from each other and from STHtrA at the N-terminus. In particular, SAHtrA2 has an extra 622 residues at the N-terminus. The N-terminal domains of both SAHtrAs, however, are not homologous to any other known domain or any sequence in the bacterial database. HtrA homologues seen in other Gram-positive bacteria are thought to be located on the cell surface although the cell wall sorting signal, LTSPG, seen in Gram-positive

bacterium is absent. If the HtrA-like proteins are located on the cell surface, then the sorting of the protein to the cell surface is likely to be via a novel mechanism.

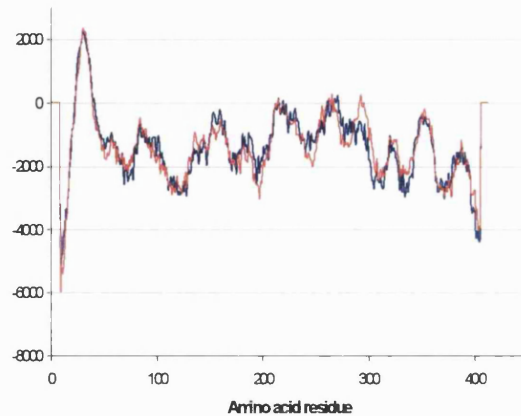
Analysis for signal peptide cleavage sites in the SAHtrA homologues using SignalP (<http://www.cbs.dtu.dk/services/SignalP>) did not return scores likely to suggest a signal peptide. However, predictions for transmembrane sequences indicate a possible transmembrane segment near the N-terminal in SAHtrA1. The hydrophobicity plot for SAHtrA1 (Figure 3.6) using TMPRED (http://www.ch.embnet.org/software/TMPRED_form.html) show a similar profile to that seen in HtrA homologues identified in other Gram-positive bacterium. In all cases, there is a positively charged region before a hydrophobic region. In SAHtrA2, a transmembrane region is predicted to be located in the middle of the protein (residues 407 - 432) preceding the catalytic domain.

A third sequence (SAHtrA3) with a predicted ORF of 706-bp and a gene product of 235 amino acids was found in the *S. aureus* genome. The gene product has a similar N-terminus to the SAHtrA1 and SAHtrA2 but is missing the PDZ domain (Figure 3.5). SAHtrA3 contains a serine protease catalytic domain, however, there is a difference between the catalytic triad of SAHtrA1 (G²⁵³NSGGAL) and STHtrA2 (G⁶²²NSGGAV) and that of SAHtrA3 (G¹⁸⁷NSGSPV). The extent of the GA→SP mutation is unknown at this time. Like the other SAHtrAs, SAHtrA3 has no signal peptide cleavage site although a transmembrane region may be present at the N-terminus of the protein.

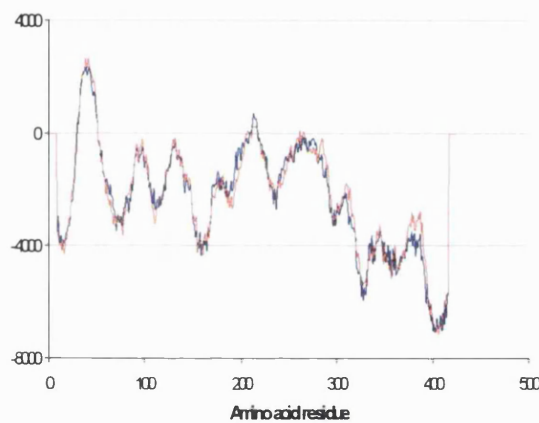
a) *B. subtilis* YvtA (Noone, D. *et al.*, 2000)



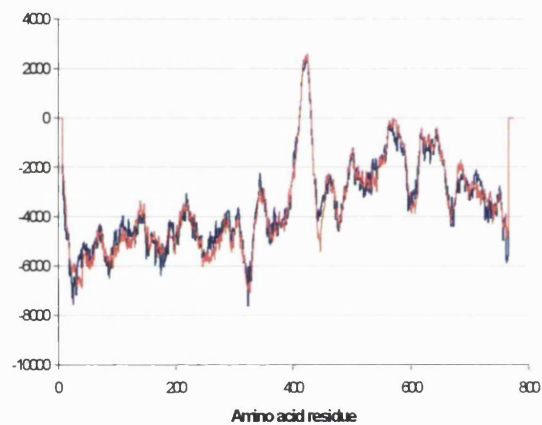
b) *L. helveticus* HtrA (Smeds, A. *et al.*, 1998)



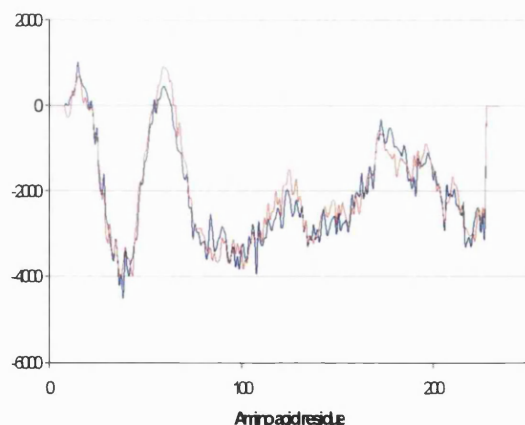
c) SAHtrA1



d) SAHtrA2



e) SAHtrA3



Hydrophobicity plots using TMPRED algorithm which predicts the presence of transmembrane regions. Peaks above the value of 0 indicate potential transmembrane domains within the query sequence

Figure 3.6 – Hydrophobicity plots of HtrA orthologues identified in Gram-positive bacteria

Helix orientation: inside->outside helices (blue), outside->inside helices (red).

3.4 Summary

A family of HtrA proteins are found in both the Gram-negative *S. typhimurium* and the Gram-positive *S. aureus* bacteria. *S. typhimurium* orthologues are highly similar to those identified in *E. coli* containing the catalytic and both PDZ domains in STHtrA and STHhoA, while STHhoB only contains one PDZ domain. The *S. aureus* orthologues, however, are similar to other Gram-positive bacteria where only one PDZ domain is present, and in the case of SAHtrA3 is completely absent. Major differences between the proteins lies in the N-terminal domains, particularly in SAHtrA2, where there is a 622 amino acid extension for which the purpose is unknown. A summary of the protein domains is shown in Figure 3.7.

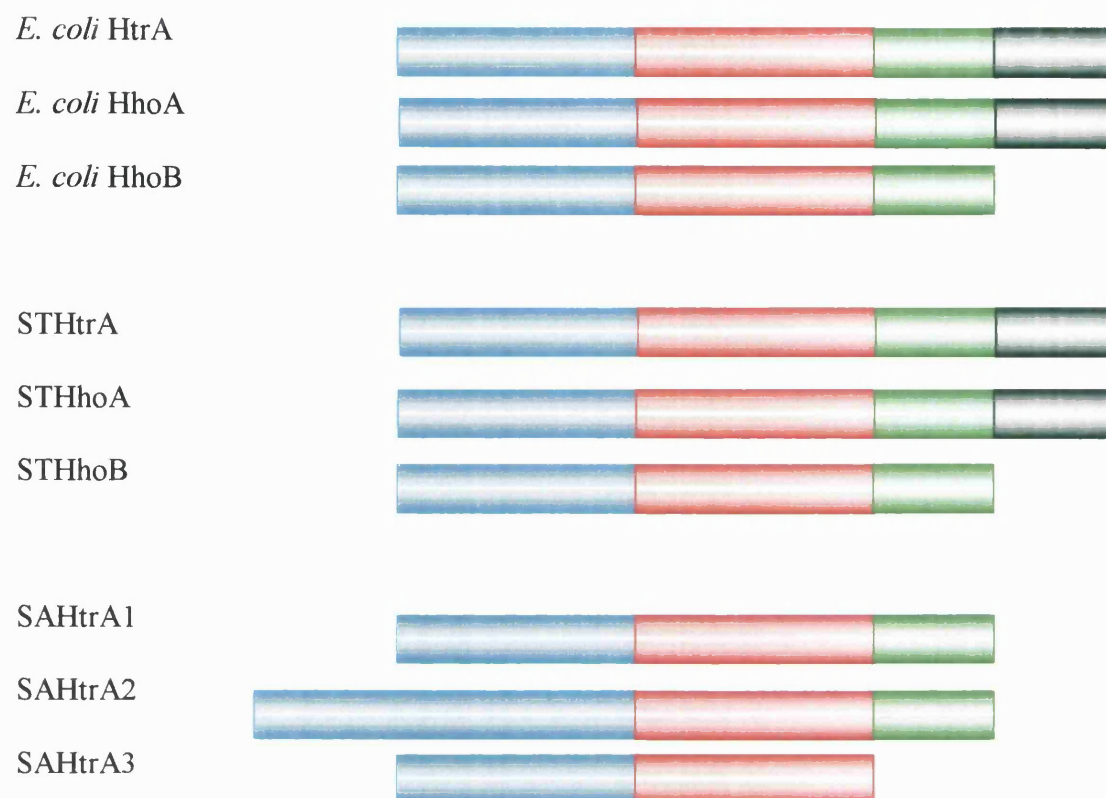


Figure 3.7 – Schematic diagram of the protein domains in the HtrA homologues in *S. typhimurium* and *S. aureus* in comparison to the *E. coli* HtrA orthologues

The N-terminal subcellular location determinant domains are coloured cyan and catalytic domains are coloured red. PDZ domains are coloured green, PDZ1 is light green whilst PDZ2 is dark green.

CHAPTER 4. TRANSCRIPTIONAL ANALYSIS OF THE ***S. aureus htrA* ORTHOLOGUES**

4.1	Introduction.....	122
4.2	Isolation of <i>S. aureus</i> RNA.....	122
4.3	Preparation of probes for Northern blot analysis	125
4.4	Northern analysis of RNA transcription by heat-shock	126
4.5	Determining the start sites for <i>sahtrA1</i> and <i>sahtrA2</i> transcription.....	129
4.5.1	Primer extension of <i>sahtrA1</i> mRNA	130
4.5.2	Primer extension of <i>sahtrA2</i> mRNA	133
4.6	Summary.....	134

4.1 Introduction

E. coli HtrA has been shown to be induced in stress conditions such as heat-shock and during oxidative stress. However, it is not induced during ethanol, and osmotic stress suggesting that it is not required under these conditions. In *E. coli*, the gene is under the control of the σ^E regulon, which in turn is regulated by periplasmic stress caused by the accumulation of misfolded proteins and also by the two-component CpxRA system (discussed earlier in 1.9).

In order to study whether or not the transcription of the HtrA gene family was induced by stress, *S. aureus* was cultured under heat-shock conditions for a period of time. This time course was chosen to determine if the genes were induced under prolonged exposure to environmental stress. RNA was isolated and Northern blot analysis was carried out probing for *sahtrA1* and *sahtrA2*.

Primer extension experiments were carried out to determine the 5' end of the mRNA for both *sahtrA1* and *sahtrA2* in order to determine with more accuracy the promoter sites for each gene and to understand the transcriptional regulation of the gene better.

4.2 Isolation of *S. aureus* RNA

To isolate *S. aureus* RNA, the bacterial cell must first be disrupted in order to obtain the cellular contents. *S. aureus* is a Gram-positive bacteria with a thick peptidoglycan cell wall and so must undergo harsher disruptive methods than those commonly used for Gram-negative bacteria or mammalian cells, which have much less rigid outer membranes. Thus, methods (Cheung, A.L. *et al.*, 1994) using small beads to break up the cells were performed and compared with phenol-chloroform

extraction methods to investigate the best method for the highest RNA yield from *S. aureus*.

Method 1 used phenol-chloroform to directly extract the RNA from bacterial pellets. Pellets were resuspended in TE buffer to which phenol-chloroform extraction was carried out. The resulting aqueous layer was precipitated by ethanol either overnight or for 5 min. The resulting pellet was resuspended in DEPC ddH₂O and loaded onto a RNA denaturing gel.

Method 2 investigated the ability for TRIZOL to isolate RNA in the presence or absence of zirconia beads. Bacterial pellets were resuspended in TRIZOL and vortexed in the presence and absence of zirconia beads. RNA in the aqueous layer was extracted by isopropanol precipitation. Samples were treated with DNase to remove any remaining DNA and then cleaned up by phenol-chloroform extraction. RNA in the resulting aqueous layer was precipitated with ethanol. The precipitate was resuspended in DEPC ddH₂O and loaded onto an RNA denaturing gel.

Figure 4.1a shows that the phenol-chloroform extraction method yield less RNA with respect to the TRIZOL method. However, the overnight precipitated method does yield a lot of RNA but the smearing seen on the gel suggests that far more RNA has been degraded when compared to the clean profile from the TRIZOL method in the presence of beads. A large band ran lower in the gel and is probably tRNA, which accounts for a large proportion of the total RNA and may influence the quantification of the RNA. As the contaminating tRNAs are small molecules, they

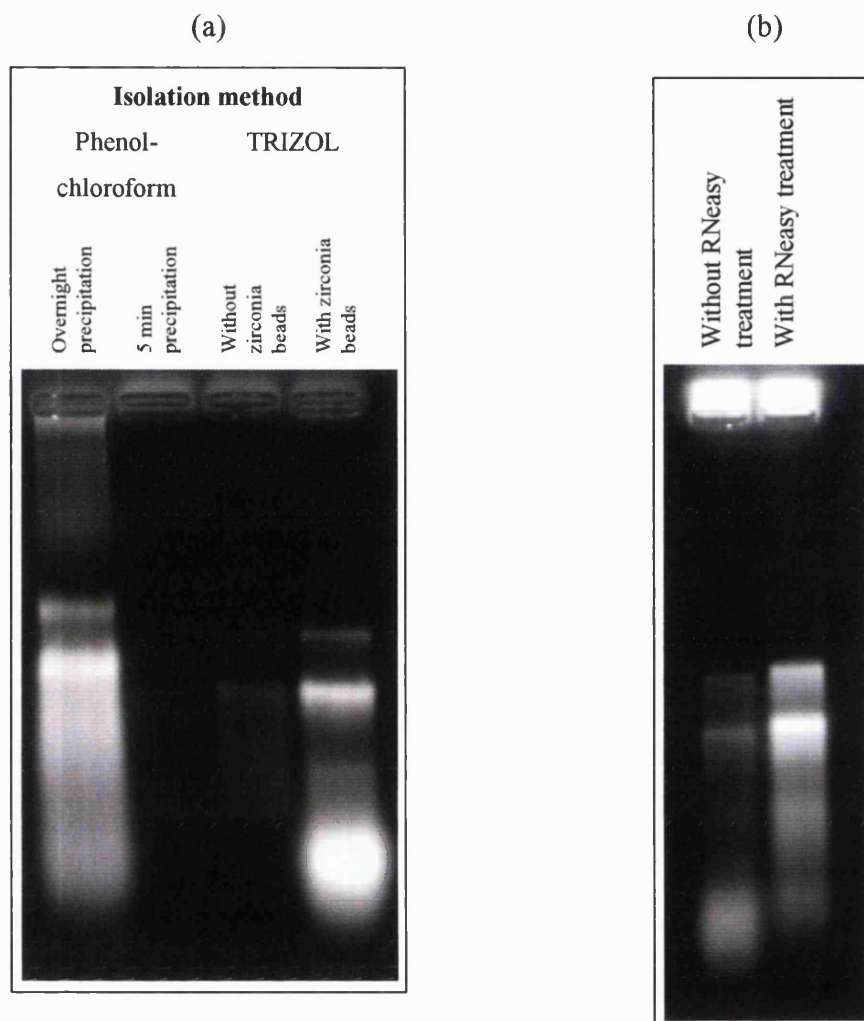


Figure 4.1 – Comparison of RNA isolation methods to produce the highest RNA yields.

Equal volumes from RNA extraction in the presence or absence of zirconia bead were loaded onto a RNA denaturing gel. Ribosomal RNA bands were visualised under UV light. Two methods were investigated: method 1 isolation used phenol chloroform extraction and precipitating the RNA overnight or for 5 min. Method 2 isolation used TRIZOL in the presence and the absence of the zirconia beads. (b) RNA extracted using zirconia beads were treated with a RNeasy mini column to remove the lower contaminating band.

were removed by putting the DNase treated samples down a RNeasy column, which removes species smaller than 200-bp (Figure 4.1 b). Treatment of RNA samples by this method resulted in the highest yield of high quality RNA (the OD_{260/280} reading gave values of approximately 1.8 or above implying very few protein contaminants of protein).

4.3 Preparation of probes for Northern blot analysis

Probes were designed to specially detect *sahtrA1* and *sahtrA2*. The DNA sequences for the PDZ domains of each SAHtrA was amplified by PCR using primers, EMR6 + EMR7, and EMR16 + EMR17 (sequences are listed in appendix A.2). The resulting PCR product was ligated into the *lacZ* gene of pGEM-T Easy vector. The resulting pGPDZ1 and pGPDZ2 plasmids were transformed into JM109 and positive clones determined by blue/white screening. Positive white colonies were picked and used to amplify the plasmid containing the DNA sequence of interest. The DNA of the plasmids were then sequenced using the primer for the T7 promoter to obtain the 5'→3' sequence and the primer, pGEM-REV, (appendix A.5) was used to obtain the 3'→5' sequence. A positive plasmid for each probe was used as the template for PCR. DNA was isolated by gel extraction and DNA was quantified by running on an agarose gel with a quantitative ladder, Hyperladder IV.

In addition, in order to demonstrate heat-shock was occurring, a probe (positive control) was designed to follow the transcription of *dnaK*, a gene known to be produced in heat-shock (Ohta, T. *et al.*, 1994), under exposure to the same environmental conditions. A 447-bp probe was designed to correspond to positions 233 to 680-bp in the *dnaK* gene (Acc # BAB45502, BAB42672). The 23S rRNA

gene (Acc # 14349228) was used in order to determine that equal amounts of RNA were loaded onto the gel. Thus, a 503-bp probe designed to correspond to positions 263 to 756-bp. Both probes were created in the same manner as the PDZ domains, using primers listed in appendix A.6.

4.4 Northern analysis of RNA transcription by heat-shock

The effect of heat-shock on the transcription of *sahtrA1* and *sahtrA2* was examined. *S. aureus* strain NCTC 8325 was grown at 30°C before transferring to 42°C. Samples were removed at different time points, pelleted and stored on dry ice. RNA was isolated by TRIZOL using zirconia beads to disrupt the bacterial cell wall. Samples were treated with DNase and then processed through an RNeasy column. RNA was quantified and equal amounts of RNA were loaded onto a denaturing RNA gel. The quality of the RNA was checked by analysing the gel under UV (Figure 4.2a) and the RNA was then transferred to Hybond N+ by Northern transfer by a manner of capillary action. RNA was then fixed to the membrane by UV. The membrane was then probed for *sahtrA1*, *sahtrA2*, *dnaK* and 23S rRNA (using the probes prepared as discussed in 4.3) stripping the membrane between each hybridisation. Probing for 23S rRNA (Figure 4.2e) implies that there is equal amount of RNA as levels of this RNA is not expected to vary during heat-shock.

Heat-shock conditions were shown by the induction of *dnaK* transcripts after 15 min of heat-shock (Figure 4.2d). This correlates with the study of the translation levels of *dnaK* in *S. aureus* (Laport, M.S. *et al.*, 2001), which showed that the levels of DnaK protein were at maximal at between 5 to 20 min. The amount of translation is then reduced by 25 to 30 min to levels lower than that seen during 37°C. Even though the

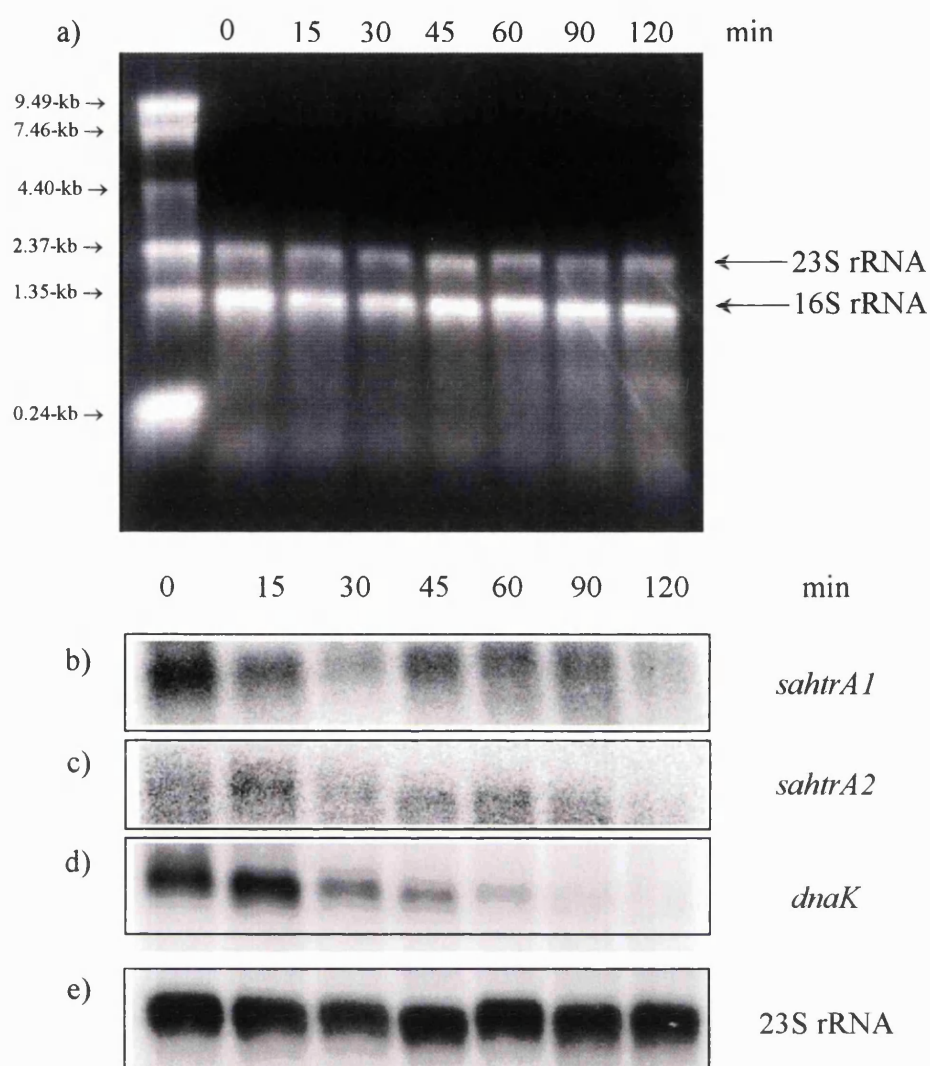


Figure 4.2 – Detection of transcript levels during heat-shock of *S. aureus*

S. aureus was grown at 30°C before transferring to 42°C for heat-shock. Samples were removed at various time points over 120 min after transferral to 42°C, and pelleted. RNA was isolated from these pellets and 2 µg total RNA was loaded per well on a denaturing RNA gel. RNA was separated and the quality of the RNA was checked by UV (a). RNA was then transferred to Hybond N+ by Northern transfer and fixed using UV. The membranes were then sequentially probed for *sahtrA1* (b) and *sahtrA2* (c). The membranes were probed for *dnaK* to check that heat-shock was occurring (d) and for 23S rRNA to check that equal amounts of RNA were loaded (e). (n = 3)

time points are not as frequent in this experiment, this is also seen at the transcriptional level where levels of *dnaK* transcripts are at maximal at 15 min and then are decreased by 30 min to levels which are even lower than that seen at 0 min.

The probe designed to detect *sahtrA1* gave a signal around 1-kb in size, consistent with the predicted ORF of 1.275-kb. Although the size of the RNA detected appears smaller, this can be explained by the inaccuracy of the RNA marker. However, given that the size of the gene and the transcript are approximately the same size, this shows that *sahtrA1* is transcribed as a monocistronic mRNA and not as a polycistronic mRNA. Under the conditions used here, *sahtrA1* does not seem to be induced (Figure 4.2b), as there is no significant difference in the transcript levels throughout the time course. This is interesting, as although this HtrA homologue has a higher homology to *E. coli* HtrA, it is not expressed under heat-shock conditions. From this it could be argued that SAHtrA1 is actually a homologue of HhoA or even HhoB as these two *E. coli* homologues are constitutively expressed (Waller, P.R.H. & Sauer, R.T., 1996) unlike *E. coli* HtrA, which is expressed during heat-shock. Protein sequence alignments show that SAHtrA1 has 34% identity, 53% similarity with HhoA and 35% identity, 52% similarity with HhoB at the amino acid level. However, without further characterisation of HhoA, HhoB and the two SAHtrA homologues, this cannot be concluded as yet.

The probe designed against *sahtrA2* detected a transcript of 2.65-kb. This is in good agreement with the predicted ORF of 2.325-kb, and the difference in size could be due to the relative inaccuracies consistent with measurements against the RNA markers. Again, due to the similar sizes of the gene and transcript, *sahtrA2* is

predicted to be transcribed as a monocistronic mRNA. Unlike *sahtrA1*, *sahtrA2* is induced under heat-shock conditions at 42°C (Figure 4.2c). Though the relative amount of mRNA is much lower, in comparison to *sahtrA1*, there is a significant induction of transcription and this is seen clearly by 15 min. The levels then return to pre-heat-shock conditions by 30 min. Induction during heat-shock occurs in *E. coli* HtrA and this similar characteristic in SAHtrA2 suggests that SAHtrA2 is the HtrA homologue in *S. aureus* even though at the protein level, it is not as similar as SAHtrA1 (discussed in 3.3).

4.5 Determining the start sites for *sahtrA1* and *sahtrA2* transcription

In order to analyse the promoter sequences involved in the transcription of *sahtrA1* and *sahtrA2*, primer extension was performed. This technique combines manual sequencing techniques with essentially one round of reverse transcription. Due to the single round of reverse transcription, a small quantity of cDNA is made, which can only be detected by a sensitive detection method. The method with the best sensitivity is the usage of radioisotopes.

Primer extension uses a primer designed to run 5'→3' complementary to the RNA sequence. This primer is radiolabelled and then annealed to the RNA before allowing one round of reverse transcription to occur in a primer extension assay. The RNA is then removed using RNase H, which removes RNA from DNA:RNA duplexes. In conjunction with this, the same primer (unlabelled) is used in a manual sequencing reaction using a plasmid containing DNA sequence –1000 to +100 (where A of ATG is designated +1) for the gene of interest. The sequencing ladder is then separated on a DNA sequencing gel adjacent to the primer extension product. The position on the

sequencing gel where the primer extension product (or products) run to, indicates the start site of the candidate mRNA. This sequence can then be read off the sequencing gel and can be used to map the 5' end of the mRNA of interest.

4.5.1 Primer extension of *sahtrA1* mRNA

For *sahtrA1*, a 1305-bp sequence (-975 to +330 where A of the predicted ATG is designated +1) was cloned in pGEM-T Easy vector. The resulting pGEMHit1-5F plasmid was used in manual sequencing reactions. A variety of primers were designed to anneal to the mRNA from +30 to +113 on the gene. These primers were first checked to see if they produced good sequencing data before being used in primer extension experiments. Two primers, PEXT-13 and PEXT-15, both gave good sequencing data and so were used for the primer extension experiments.

For the primer extension experiments, attempts were made to maximise the amount of RNA that could be used in each experiment due to the small amount of cDNA which would be generated. Optimisation of the annealing conditions included optimising both the initial temperature to linearise the RNA and the annealing temperature of the probe and the mRNA. In addition, the time for each reaction was optimised. The final conditions are described in the Methods and Materials section (2.4.4.2).

Primer extension experiments using two different primers yielded the same result (Figure 4.3). The 5' end of *sahtrA1* mRNA begins 26-bp upstream from the predicted ATG start site. Figure 4.4 shows predicted sequences important for transcriptional and translational start signals. The gene is preceded by a putative

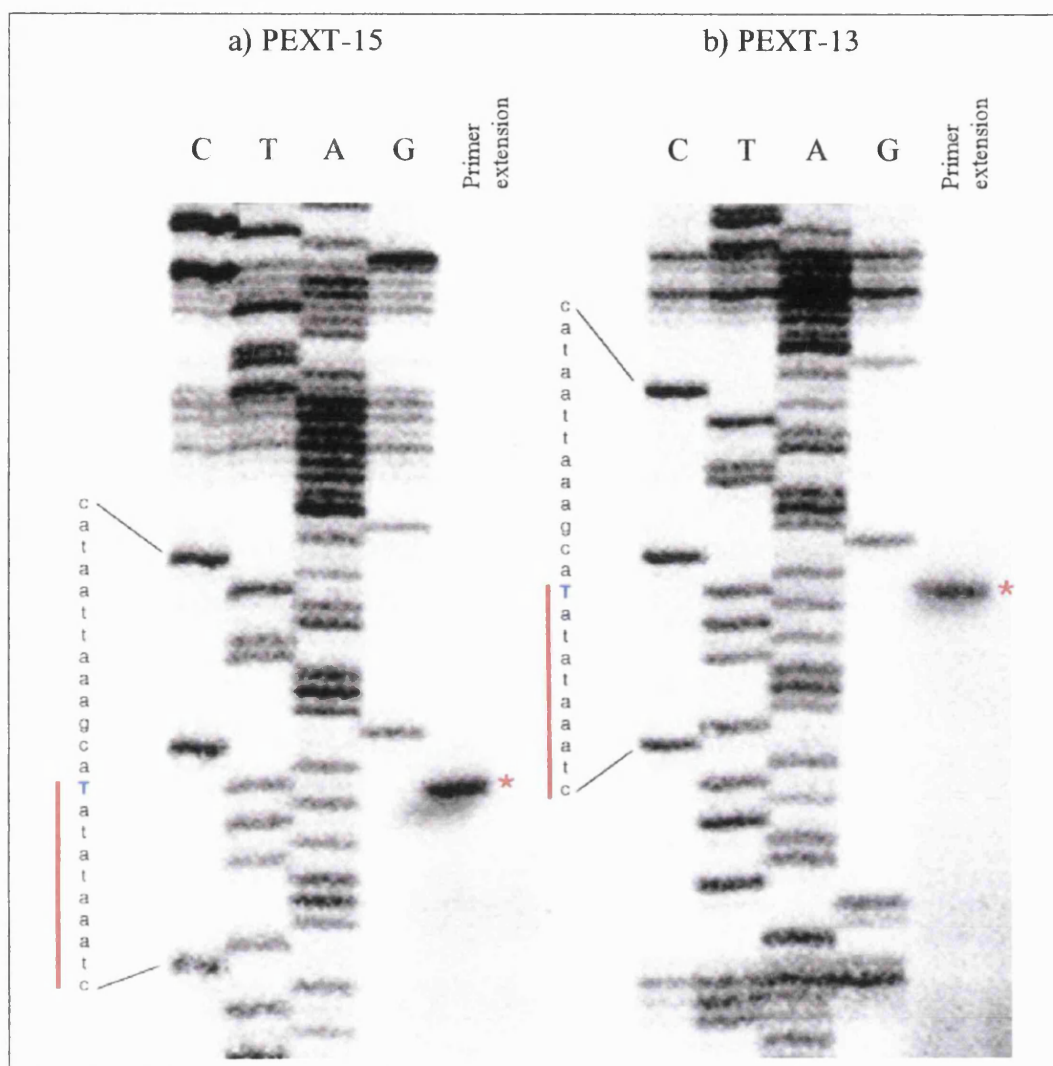


Figure 4.3 – Primer extension of *sahtA1* mRNA

The 5' end of *sahtA1* mRNA was extended by annealing a radiolabelled primer to the mRNA. One round of reverse transcription was performed to produce radiolabelled cDNA. RNA was removed by treatment with RNase H and samples then ran alongside a sequencing reaction, which used the same primer (unlabelled) and pGEMHit-5F as the template DNA. Two primers (PEXT-15 (a), and PEXT-13 (b)) which annealed to different areas gave the same transcriptional start site. The 5' end of the mRNA is highlighted with a red line and transcription start site is in uppercase (blue). The primer extension product is marked with a red asterisk.

	-35	Spacer	-10
<i>sea</i>	TAGACA	N ₁₈	TAATAT
<i>sar-P2</i>	TAGCAA	N ₁₈	TAATAT
<i>blaZ</i>	TTGACA	N ₁₈	TATTAT
<i>sar-P1</i>	TTTACT	N ₁₈	TATAAT

AGTGTATTAC**TAGCAA**TAATTATAAAAAAAGAC**ATAATT**TAAAGCAT**T**ATATAAATCATATAA**TGGAGGT**TAAGT**ATG**

-35 -10 +1 RBS

<i>spr</i>	TGGAGTT
<i>EII^{mtl}</i>	TGGAGTG
<i>lacG</i>	AGGAGTT
<i>tetM</i>	TGGAGGA
Gene	RBS

The nucleotide (blue) corresponding to the primer extension product is marked by an arrow and designated +1. The potential -10 and -35 sequences for the *sahtrA1* transcript are marked in a bold red font. A suggested ribosome binding site (RBS) is highlighted in a bold green font. The translation start site, ATG, is highlighted in a bold dark red font. Below the sequence are other promoter regions (a) and RBS sequences (b), which have been determined for other *S. aureus* genes (Novick, R.P., 1991, Deora, R. & Misra, T.K., 1996, Manna, A.C. *et al.*, 1998).

promoter sequence with sequences TAGCAA and TAATTA for the -10 and -35 regions, respectively. These sequences are similar to other σ^A -like promoter sequences identified (Deora, R. & Misra, T.K., 1996, Novick, R.P., 1991, Manna, A.C. *et al.*, 1998). σ^A is a vegetative σ factor (Deora, R. & Misra, T.K., 1996) and thus is involved in the transcription of 'housekeeping' genes. This predicted promoter site for *sahtrA1* is therefore supportive of the Northern data, which indicated that *sahtrA1* mRNA is not induced under heat-shock conditions. These data support the suggestion that SAHtrA1 could indeed be a homologue of either HhoA or HhoB.

Interestingly, the ribosome binding site (RBS) sequence, TGGAGGT, is identical to that seen in *spr*, a *S. aureus* gene encoding the V8 serine protease (Novick, R.P., 1991).

4.5.2 Primer extension of *sahtrA2* mRNA

For *sahtrA2*, a 1407-bp sequence (-903 to +505 where A of the predicted ATG is designated +1) was cloned into the pGEM-T Easy vector. The resulting pGEMHit2-5B plasmid was used for manual sequencing reactions. A variety of primers were designed to anneal to the mRNA from +26 to +84 on the gene. These primers were first checked to see if they produced good sequencing data before using in primer extension experiments. One primer, PEXT-21A gave good sequencing data and so was used for the primer extension experiments. However, attempts to map the start site of *sahtrA2* mRNA were not successful. This could be due to the small amount of starting material as Northern (Figure 4.2) indicate that the *sahtrA2* mRNA is much

less abundant than *sahtrA1* mRNA. Thus, with less mRNA in the sample, the signal produced in primer extension experiments may be too weak to be detected.

4.6 Summary

sahtrA1 is expressed as a monocistronic mRNA but is not induced by heat-shock at 42°C. This finding suggests that *sahtrA1* may actually be a homologue of *E. coli* *hhoA* or *hhoB* as both these homologues are constitutively expressed in *E. coli* with only *htrA* transcription being regulated under stress conditions, such as heat-shock. This constitutive transcription of *sahtrA1* is further supported by the predicted promoter sites as it contains sequence motifs at the -10 and -35 regions that are similar to sequences identified for σ^A , the vegetative promoter in *S. aureus*.

Unlike *sahtrA1*, *sahtrA2* is induced under heat-shock conditions suggesting that *sahtrA2* is more likely to be the HtrA homologue in *S. aureus*. The mRNA is transcribed with an approximate length of 2.7-kb, which is in agreement with a monocistronic mRNA as the predicted ORF is 2.325-kb. However, primer extension experiments did not produce any detectable product. This was likely to be due to the low abundance of *sahtrA2* mRNA under the conditions used.

Northern blot analysis has shown the presence of the mRNA of *sahtrA1* and *sahtrA2*. Thus, in order to investigate the function of these genes, recombinant proteins of these genes were generated.

CHAPTER 5. FUNCTIONAL ANALYSIS OF RECOMBINANT *S. aureus* HTRA ORTHOLOGUES

5.1	Introduction.....	136
5.2	Cloning and expression strategies.....	136
5.2.1	Cloning strategy	136
5.2.2	Expression strategy.....	139
5.3	Cloning, expression and purification of recombinant SAHtrA1	145
5.3.1	Cloning and expression of recombinant native SAHtrA1	145
5.3.2	Cloning and expression of recombinant 6xHis-tagged SAHtrA1	148
5.4	Cloning, expression and purification of SAHtrA1 subdomains	154
5.4.1	Cloning and expression of the catalytic domain of SAHtrA1	154
5.4.2	Cloning and expression of the PDZ domain of SAHtrA1	156
5.4.2.1	Cloning of the PDZ of SAHtrA1 as a GST-tagged protein	156
5.4.2.2	Purification of GST-PDZ1 by FPLC.....	158
5.4.2.3	Cloning of the PDZ of SAHtrA1 as a 6xHis-tagged protein.....	169
5.5	Cloning, expression and purification of SAHtrA2 and its subdomains	171
5.5.1	Cloning and expression of SAHtrA2.....	171
5.5.2	Cloning, expression and purification of the PDZ domain of SAHtrA2	177
5.5.2.1	Cloning and expression of PDZ2 as a GST-tagged protein	177
5.5.3	Cloning, expression and purification of 6xHis-tagged PDZ2.....	180
5.6	Protease activity of the full length recombinant proteins	181
5.7	Summary.....	183

5.1 Introduction

For the purpose of characterisation, recombinant proteins of the HtrA homologues in *S. aureus* were generated. The *S. aureus* homologues, SAHtrA1 and SAHtrA2 were expressed in their native form. Fusion proteins were also generated with a 6xHis-tag so that purification of the protein would be facilitated.

In order to investigate the ligand binding characteristics of the *S. aureus* orthologues, the PDZ domains of SAHtrA1 and SAHtrA2 were generated. Construction of the catalytic domain of SAHtrA1 was attempted in order to characterise the catalytic activity of this homologue. The recombinant proteins were produced with a N-terminal GST-tag or 6xHis-tag in order to aid purification.

5.2 Cloning and expression strategies

Several strategies were explored in order to express high levels of soluble proteins using a variety of plasmids and expression systems.

5.2.1 Cloning strategy

All the insert sequences were cloned using similar strategies with a variety of expression plasmids. Target sequences were amplified by PCR using the proofreading DNA polymerase, High Fidelity Expand (Roche). Preliminary studies were performed to maximise the yield of PCR products for cloning procedures by optimising the concentration of MgCl₂. Generally, for these experiments, increasing the concentration of MgCl₂ increased the amount of PCR product generated. However, under these conditions there is more non-specific binding of primers and so more PCR species are produced. In this study, optimisation resulted in a

significant increase in the amount of PCR products with no generation of unwanted PCR products. An example of this is seen in Figure 5.1 where the ORF of SAHtrA1 was amplified with increasing concentrations of MgCl₂. It can be seen that the level of amplification of *sahtrA1* is significantly higher at 4 mM than at any other concentration and so these conditions were used for further PCR experiments.

Using the optimal MgCl₂ concentrations, target insert DNA was amplified from genomic DNA using primers which were designed to incorporate different restriction sites at the 5' and 3' ends of the PCR product (appendix A.2). Integrating different restriction sites into the insert DNA simplifies the cloning of the gene into an expression plasmid as the DNA can only ligate in one direction, thus reducing the number of steps required to obtain the correct clone. Purified PCR products were digested with the appropriate restriction endonucleases and purified using the QIAquick PCR purification kits. The plasmid, into which the target DNA was to be inserted, was linearised with the appropriate restriction endonucleases and dephosphorylated to prevent recircularisation. The target DNA was then ligated into the plasmid DNA and transformed into *E. coli* DH5 α cells in order to amplify the levels of plasmids. Single colonies were cultured and the plasmids isolated for subsequent DNA sequencing. A plasmid with the correct DNA sequence was chosen for use in further expression studies.

For expression studies, four different expression plasmids were used; pTrc99A was used for production of native recombinant proteins, pGEX-4T-1 was used for expression of N-terminal GST-tagged proteins, pQE-30 and pRSET-B were used for expression of N-terminal 6xHis-tagged proteins. The plasmid pRSET-B differs from

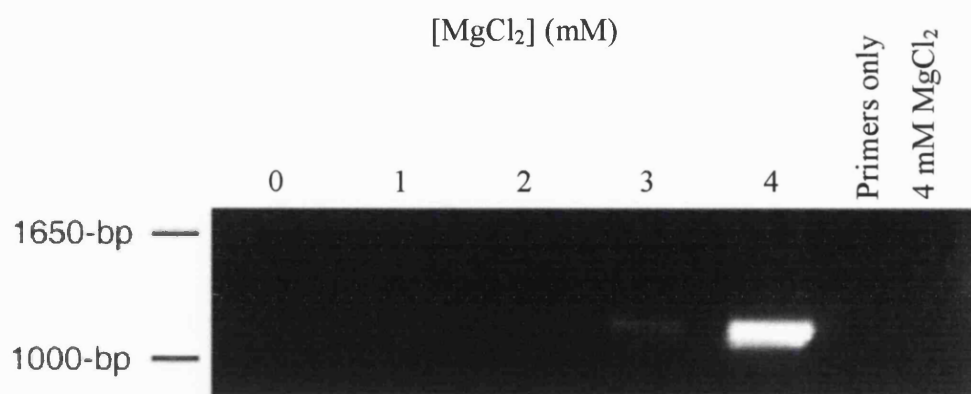


Figure 5.1 – Effect of increasing MgCl₂ concentrations on the yield of PCR products

PCR reactions (25 µl) were set up with 1 µl *S. aureus* genomic DNA and 1 µl each of primers EMR1 and EMR2. MgCl₂ concentrations ranged from 0 mM to 4 mM. Samples were then subjected to 30 cycles of PCR as described in 2.1.2. A 1% agarose gel was loaded with 5 µl PCR products and electrophoresed (2.1.3). PCR products were visualised under UV light.

pQE-30 as it has an enterokinase site that facilitates removal of the 6xHis-tag. The cloning strategies for these four plasmids and the target sequences used for each are illustrated in Figure 5.2 (pTrc99A), Figure 5.3 (pGEX-4T-1), Figure 5.4 (pQE-30) and Figure 5.5 (pRSET-B).

5.2.2 Expression strategy

To generate the target proteins, plasmids containing the correct sequence were transformed into a variety of expression strains. Single colonies containing the correct plasmids were cultured and protein expression induced by IPTG. An expression strategy (Figure 5.6) was employed in order to determine the optimal conditions for protein expression. The optimal time for protein induction was examined by inducing cultures with 1 mM IPTG over a time period of 4 hr. Subsequently, the optimal IPTG concentration was determined by inducing cultures with increasing concentrations of IPTG (0 to 5 mM) using the pre-determined optimal time for induction. These optimisation steps were carried out in L-Broth and T-Broth to investigate the effect of the two media. The effect of temperature was also investigated, as lower temperatures are likely to increase the generation of soluble protein. Thus, inductions were carried out at 37°C and 26°C, for which the latter experiment 16 hr time points were also taken. After optimising the conditions for protein expression, the solubility of the recombinant proteins was examined. Bacterial pellets from a 10 ml induction, using the optimal conditions, was sonicated and cell debris cleared by centrifugation. The supernatant was analysed to determine whether the protein was soluble. If detected on a Coomassie-stained SDS-PAGE gel, the protein was deemed feasible for purification for future work.

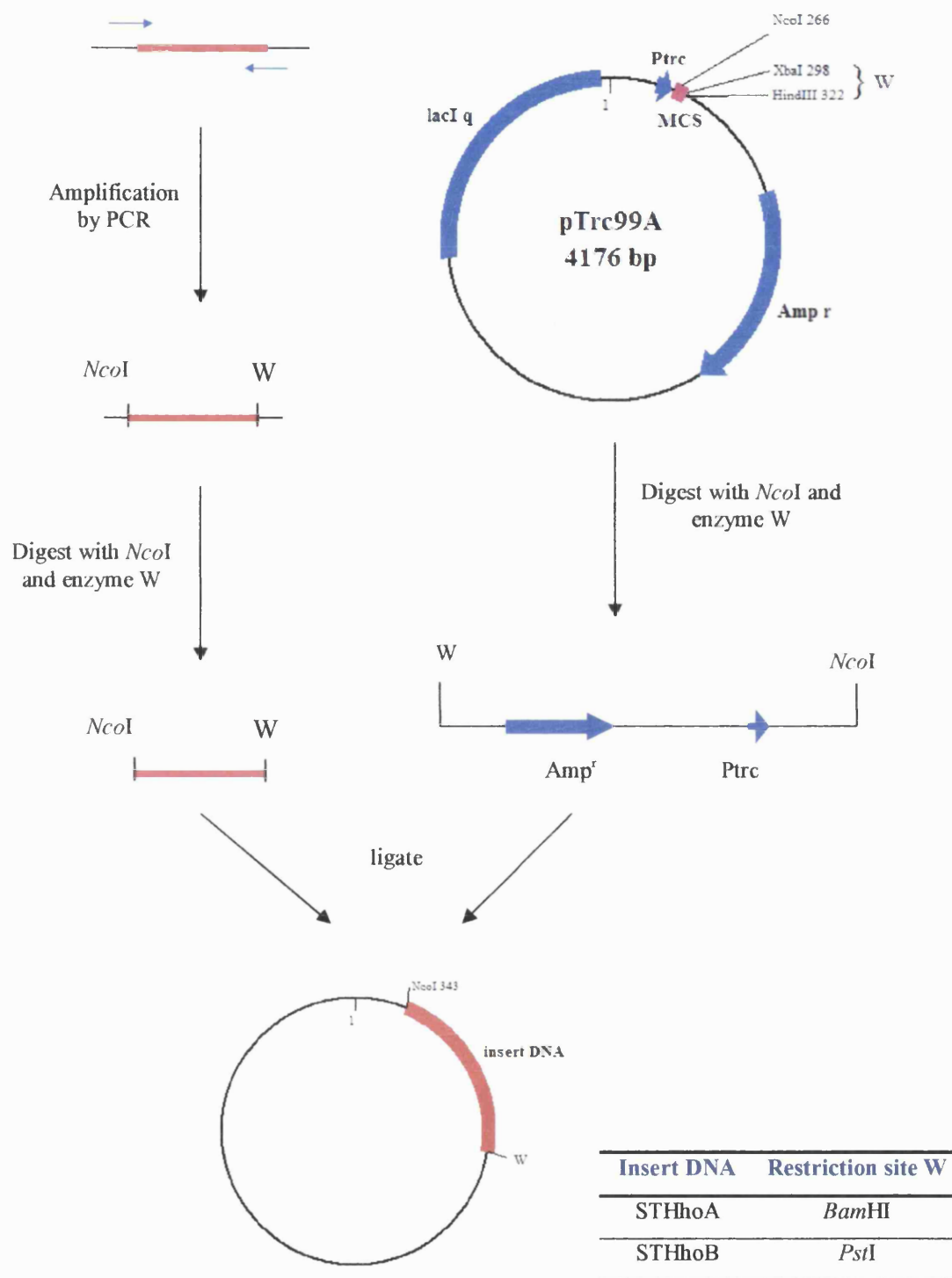


Figure 5.2 – Cloning strategy for native recombinant proteins using pTrc99A

Insert DNA cloned into pTrc99A were engineered with a *NcoI* (or *Afl*III for SAHtrA1) site at the 5' end and site W (see table) at the 3' end of the insert.

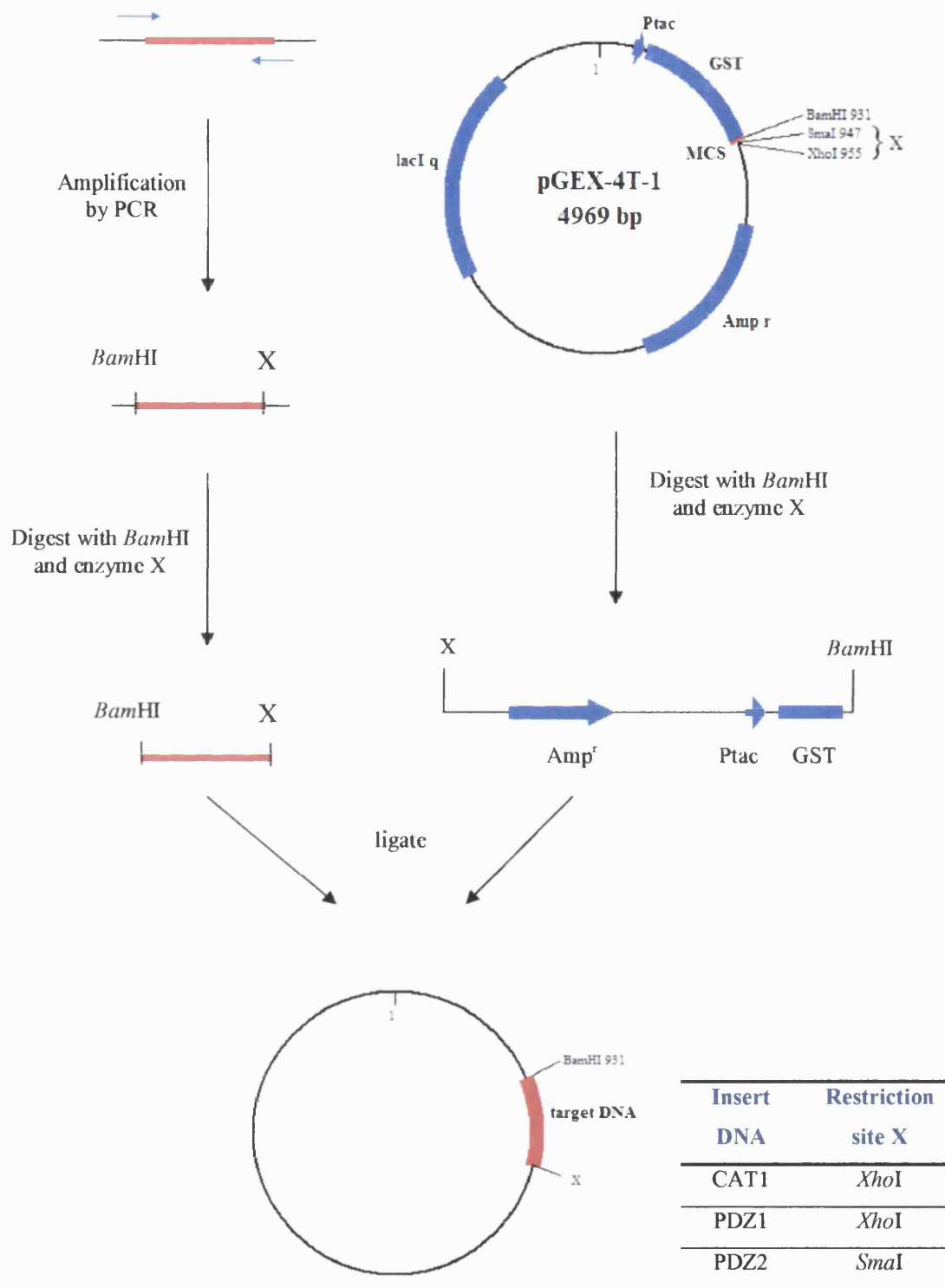


Figure 5.3 – Cloning strategy for GST-tagged recombinant proteins

All insert DNA cloned into pGEX-4T-1 was engineered with a *Bam*HI site at the 5' end and site X (see table) at the 3' end of the insert.

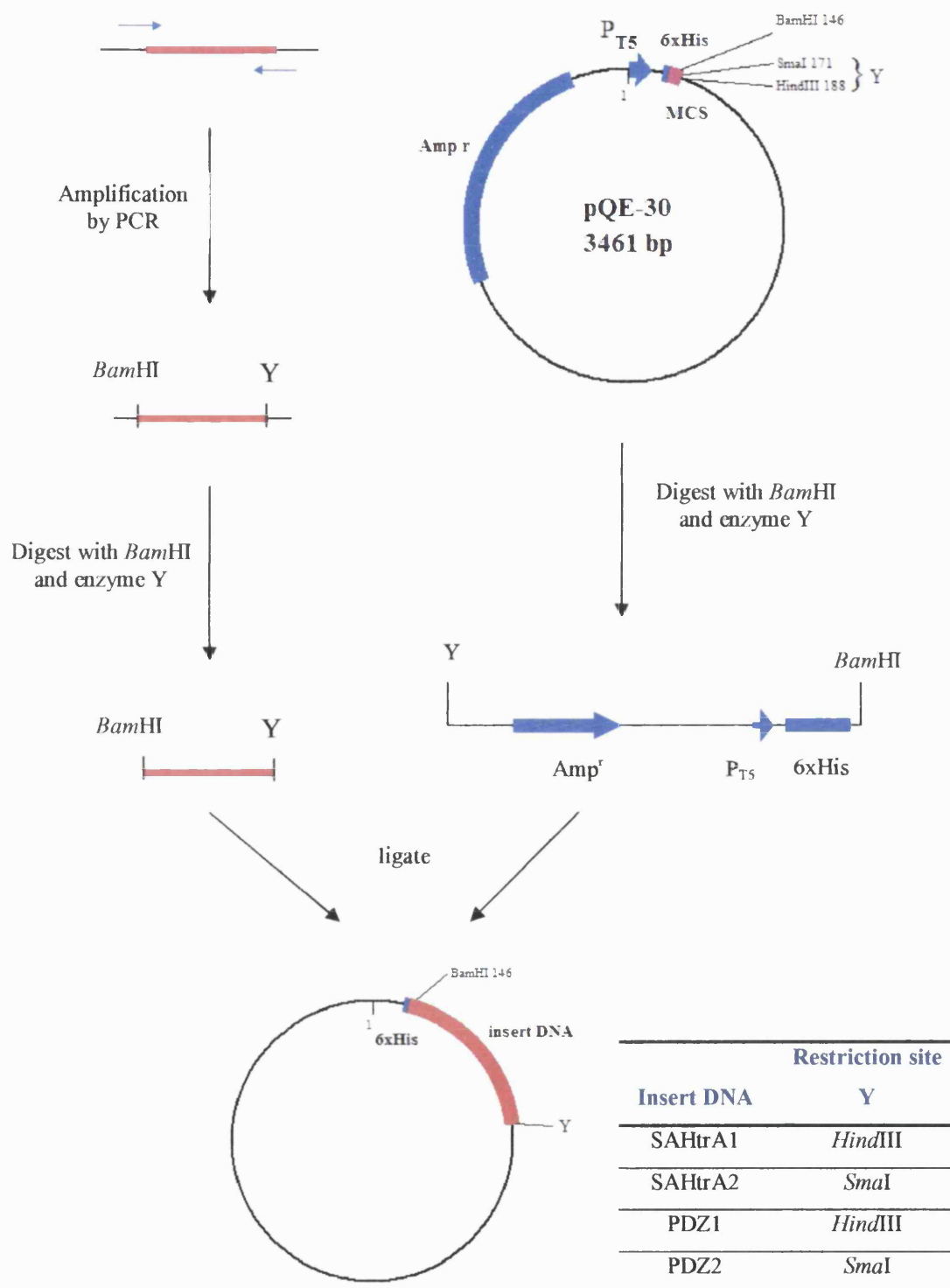


Figure 5.4 – Cloning strategy of 6xHis-tagged recombinant proteins using pQE-30

All insert DNA cloned into pQE-30 was engineered with a *Bam*HI site at the 5' end and site Y (see table) at the 3' end of the insert.

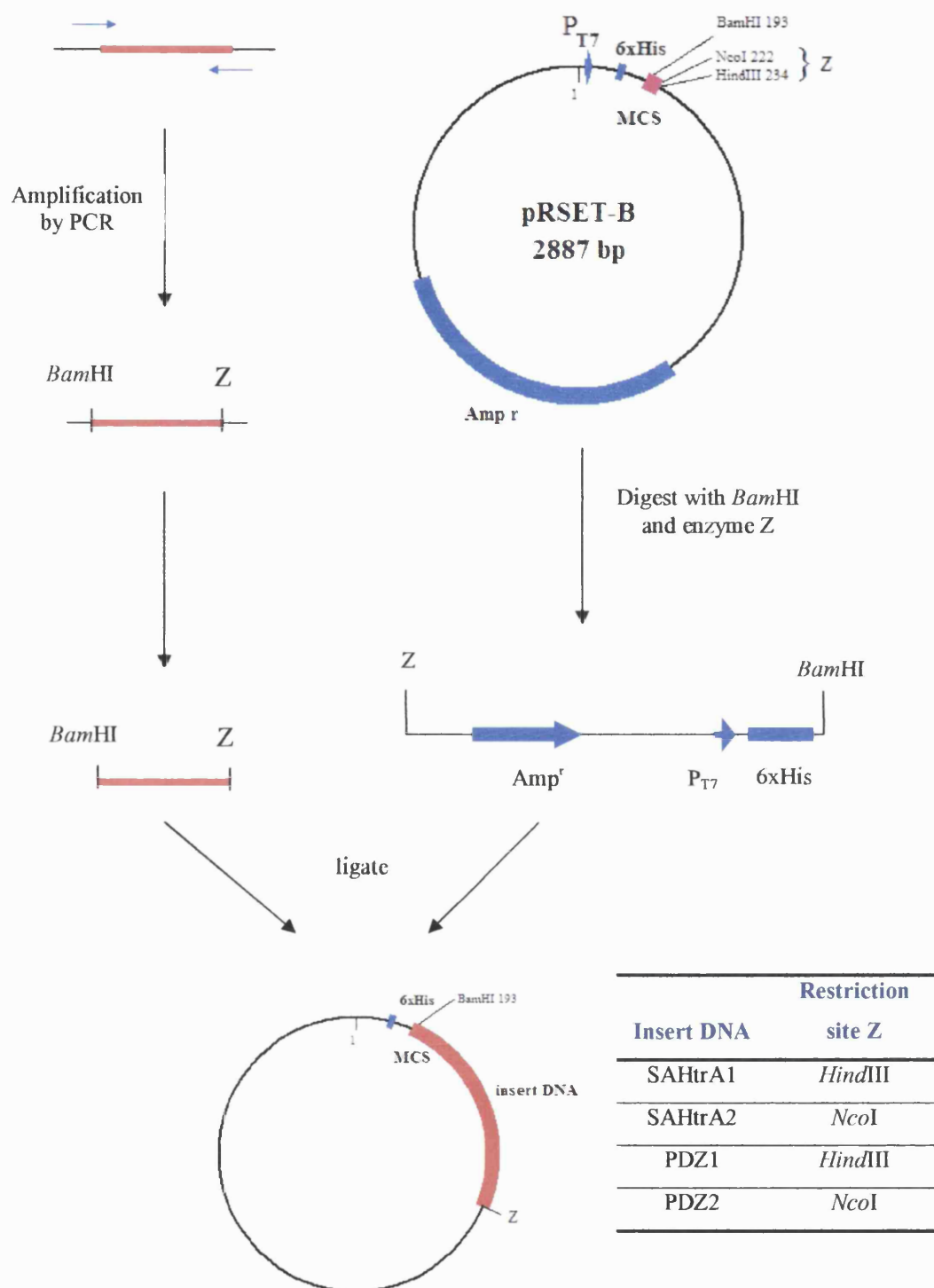


Figure 5.5 – Cloning strategy of 6xHis-tagged recombinant proteins using pRSET-B

All insert DNA cloned into pRSET-B was engineered with a *Bam*HI site at the 5' end and site Z (see table) at the 3' end of the insert

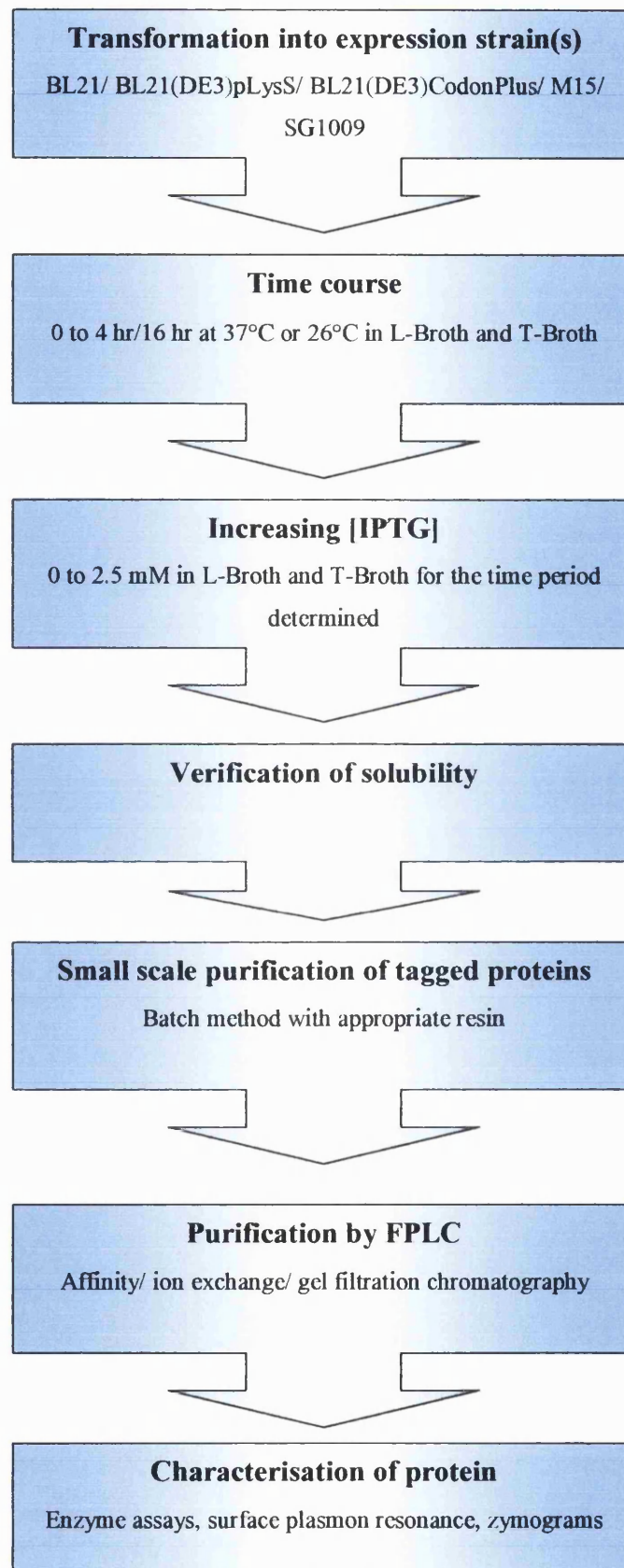


Figure 5.6 – Expression strategy to optimise protein expression

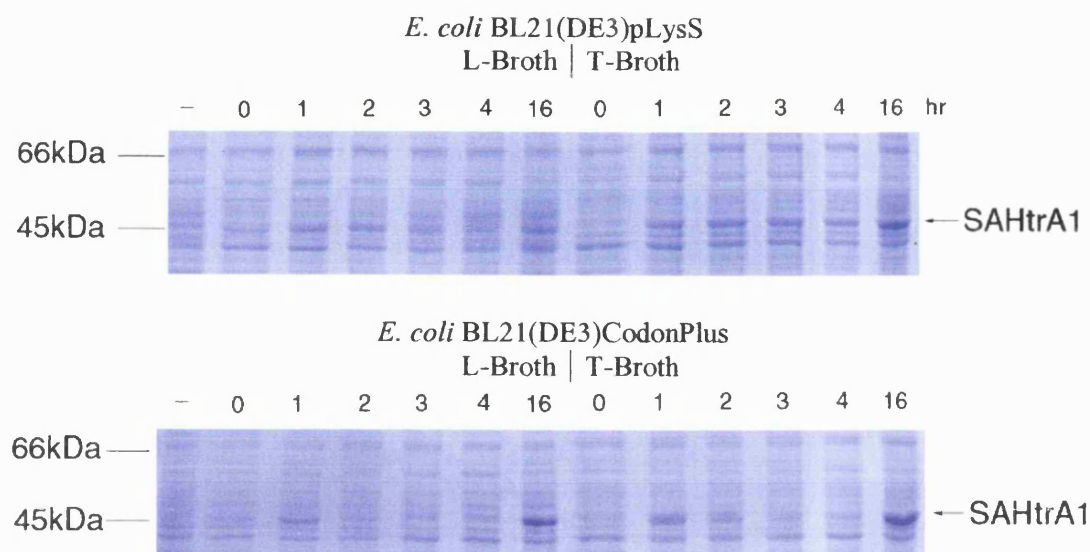
5.3 Cloning, expression and purification of recombinant SAHtrA1

5.3.1 Cloning and expression of recombinant native SAHtrA1

The ORF of *sahtrA1* was cloned into the *Nco*I and *Hind*III sites of pTrc99A as illustrated in Figure 5.2. However, *sahtrA1* has an internal *Nco*I site, so an *Afl*III site was incorporated at the 5' end of the sequence by the primer EMR1. This is because the resulting overhang from the digestion with *Afl*III (5' A[▼]CATGT 3') is the same as for *Nco*I (5' C[▼]CATGG) and so was exploited for cloning *sahtrA1* into pTrc99A. A *Hind*III site was incorporated at 3' end by EMR2.

The coding region for *sahtrA1* was amplified using the proof reading DNA polymerase, High Fidelity Expand, using primers EMR1 and EMR2 (section 2.1.2 Table 2.1) using an optimal MgCl₂ concentration of 4 mM (Figure 5.1). The resulting 1275-bp product was purified before digestion with *Afl*III and *Hind*III. The expression plasmid, pTrc99A, was digested with *Nco*I and *Hind*III and ligated to the insert DNA. The ligation mix was transformed into *E. coli* DH5 α cells. Single colonies were cultured and plasmids isolated from these for sequencing. The plasmid, pTSAHtrA1-41, contained the correct sequence and was verified by a second round of sequencing. This plasmid was transformed into *E. coli* BL21(DE3)CodonPlus cells and *E. coli* BL21(DE3)pLysS cells for expression. The optimisation of protein expression was performed by investigating various parameters in two different broths. Figure 5.7a shows that over a period of 4 hr a protein band is induced by IPTG and is approximately the correct size of SAHtrA1 (calculated molecular mass 45.8 kDa). The time course shows that expression of this band is maximal after 1 to 2 hr in *E. coli* BL21(DE3)CodonPlus and after 3 hr in *E. coli* BL21(DE3)pLysS. In both cases, the overnight culture at 26°C resulted with in a

a) Time course of recombinant native SAHtrA1



b) Increasing IPTG concentrations

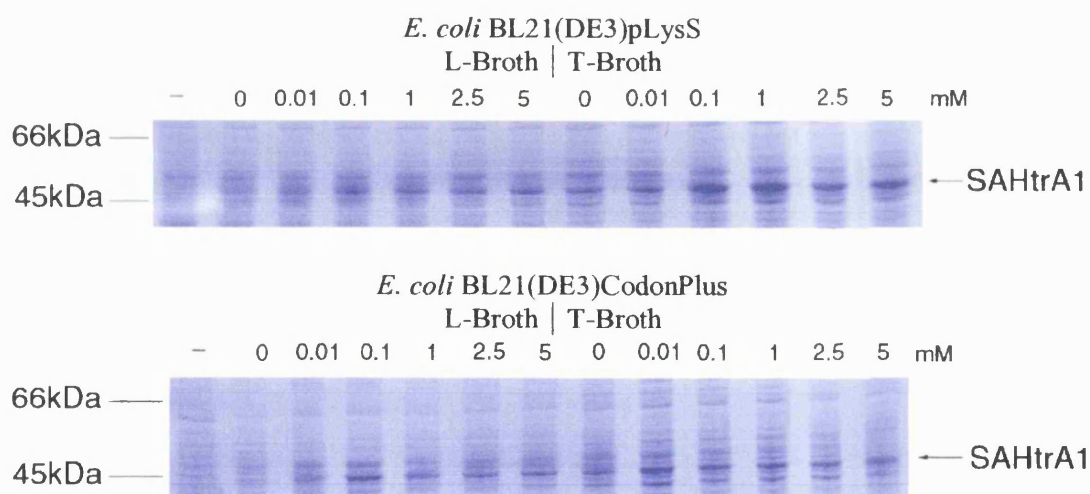


Figure 5.7 – Optimisation of the expression of soluble SAHtrA1 in *E. coli* BL21(DE3)Codon Plus and *E. coli* BL21(DE3)pLysS cells from pTSAHtrA1-41

The expression of native SAHtrA1 was optimised by investigating the yield of recombinant protein over time (a - 0 to 4 hr at 37°C and 16 hr at 26°C) and the effect of increasing IPTG (b - 0 to 2.5 mM). The effect of two culture media (L-Broth and T-Broth) was also investigated to maximise yield. Negative controls, containing pTrc99A only, are denoted by ‘-’.

heavier staining protein band in the gel suggesting that more SAHtrA1 was produced under these conditions than at 37°C for up to 4 hr. The overnight culture is further favoured for the induction conditions as the density readings (at OD₆₀₀) of the cell culture were also much higher (generally twice that of 4 hr at 37°C) suggesting that there is also a higher population of bacteria in the culture. Furthermore, expression of SAHtrA1 was more evident in T-Broth when compared to L-Broth. To investigate the effect of IPTG concentration, the *E. coli* BL21(DE3)CodonPlus cells transformed with pTSAHtrA were induced overnight at 26°C with increasing concentrations of IPTG (Figure 5.7b). Induction in L-Broth produced similar yields of SAHtrA1 at all IPTG concentrations used, although levels are higher in *E. coli* BL21(DE3)pLysS. Expression in T-Broth again was much more pronounced and expression in *E. coli* BL21(DE3)CodonPlus cells was maximal at 0.01 mM IPTG whereas in *E. coli* BL21(DE3)pLysS maximal expression was between 0.1-1.0 mM. At higher concentrations of IPTG, SAHtrA1 yield is reduced perhaps suggesting that the protein is unstable under these conditions, possibly due to host proteases. With the expression of SAHtrA1 optimised, the next step was to determine whether the protein was soluble. Bacterial pellets were sonicated and cell debris separated by centrifugation. The supernatant was then loaded onto SDS-PAGE gels (Figure 5.7c). Unfortunately, the protein did not appear to be in the supernatant and hence insoluble.

Due to the insoluble nature of the protein, induction conditions were altered to increase the solubility of the recombinant protein. Alterations included reducing IPTG concentration to 0.001 mM, incubating for a shorter time of 4 hr, incubating at a lower temperature of 20°C and also allow more aeration of the culture. In addition,

two other bacterial expression strains, DH5 α and JM109, were used. Time courses over 4 hr were carried out at 26°C and 20°C using 0.001 mM IPTG for all four strains (data not shown). Due to the slow growth at 20°C no SAHtrA1 could be detected (data not shown). However at 26°C, SAHtrA1 could be clearly seen in all four strains used (Figure 5.8). Under these conditions, SAHtrA1 was soluble albeit at very low levels.

5.3.2 Cloning and expression of recombinant 6xHis-tagged SAHtrA1

Although native recombinant SAHtrA1 was soluble, it was expressed at low levels. Thus, the protein was generated with a N-terminal 6xHis-tag as this would aid purification procedures. In order to clone SAHtrA1 with a 6xHis-tag two different expression plasmids were chosen. One plasmid, pRSET-B, had an enterokinase recognition site following the 6xHis-tag, facilitating the removal of the tag. The second plasmid, pQE-30, encoded the 6xHis-tag directly before the multiple cloning site (MCS) and so a smaller tag (0.84 kDa compare to 3.5 kDa of pRSET-B) would be added, however, the tag cannot be removed.

The cloning strategies are illustrated in Figure 5.4 and Figure 5.5 and pTSAHtrA1 was used as the template DNA for amplification by PCR. The primers EMR3 and EMR2 were used for amplifying *sahtA1* for pQE-30, while EMR4 and EMR2 were used for pRSET-B. Optimal MgCl₂ concentrations (2 mM for pQE-30, 3 mM for pRSET-B) were determined as described in 2.1.2.1 (data not shown). The two PCR products were purified and then digested with *Bam*HI and *Hind*III. The DNA fragments were subsequently purified before ligating to the respective plasmids, which had been digested with *Bam*HI and *Hind*III and subjected to

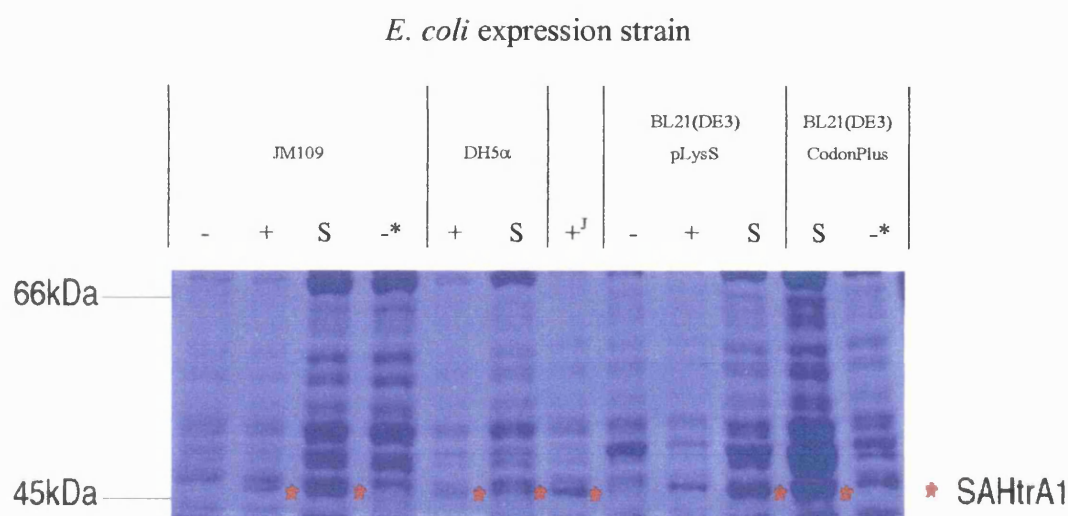


Figure 5.8 – Expression of soluble SAHtrA1 in JM109, DH5α, BL21(DE3)pLysS and BL21(DE3)CodonPlus

Four bacterial strains were transformed with pTSAHtrA1-41. Cultures were induced with 0.001 mM IPTG for 4 hr and were subsequently pelleted by centrifugation. The supernatants were discarded and pellets resuspended in sonicating buffer (2.2.3), Cell debris was removed by centrifugation and 1 µl supernatant (S) was loaded onto a 10% SDS-PAGE gel. Negative controls were performed by inducing the same bacterial strains containing the empty pTrc99A. . Negative controls (-) (equivalent of OD₆₀₀ = 0.01) were obtained from the 4 hr time point (T-Broth) in the time course. In addition, 1 µl of the sonicate from a 10 ml negative control culture was loaded in lanes denoted ‘-*’. Positive controls (+) (equivalent of OD₆₀₀ = 0.01) were obtained from the 4 hr time point in the time course.

dephosphorylation. The ligation mix was then transformed into *E. coli* DH5 α cells and single colonies were picked and cultured before extracting the resulting pQSAHtrA1 and pRSHtrA1 plasmids for sequencing. Correct plasmids, pQSAHtrA1-2 and pRSAHtrA-3, were then transformed into a series of bacterial expression strains in order to determine the best host.

Western blot analysis showed that there was no protein expression from pQSAHtrA1 in any of the strains used and so this strategy was discarded. Expression of 6xHis-tagged SAHtrA1 (6xHis-SAHtrA1) of the expected size (calculated molecular mass 49.1 kDa) from pRSAHtrA1 occurred in all four strains used (Figure 5.9a). Western blot analysis detected a second smaller protein in the M15 and SG1009 strains, which could be a result of degradation or premature termination during translation. This contaminant is not seen in *E. coli* BL21(DE3)CodonPlus as this strain is constructed to contain extra copies tRNA with codons that are rare in *E. coli*. Levels of expression in *E. coli* BL21(DE3)CodonPlus cells are slightly less than that of *E. coli* M15 and *E. coli* SG1009, which express similar levels to each other. However, the manufacturers recommend that *E. coli* M15 is the better strain for proteins, which may be toxic to the cell. Thus, these two strains were chosen for further optimisation; *E. coli* M15 was chosen for the higher yield and *E. coli* BL21(DE3)CodonPlus was chosen for the cleaner product.

Optimisation experiments compared the expression of 6xHis-SAHtrA1 in L-Broth and T-Broth using 26°C and 37°C following expression over time and with increasing concentrations of IPTG, as carried out for pTSAHtrA1. These experiments showed that the levels of 6xHis-SAHtrA1 were higher in T-Broth in all

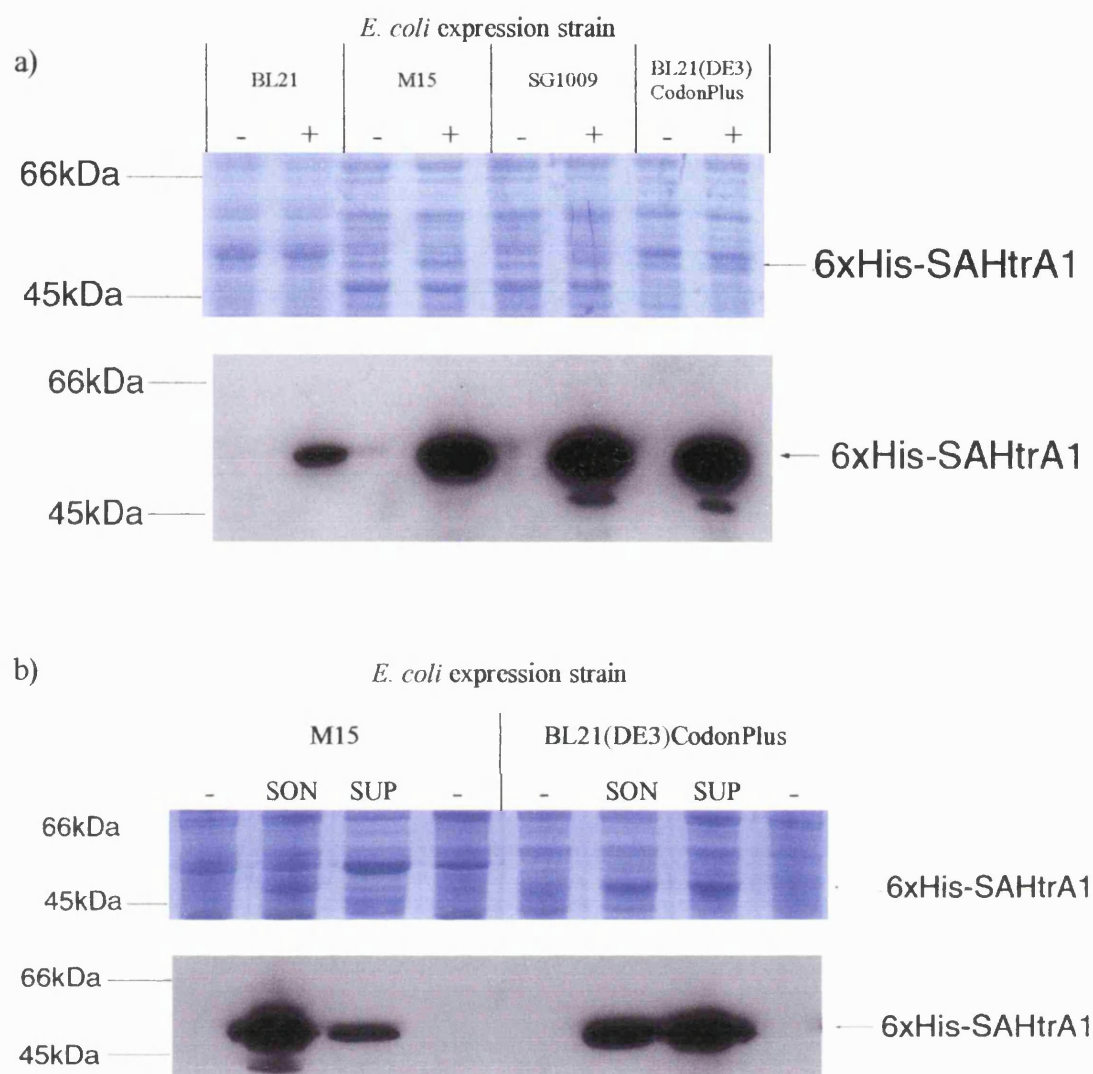


Figure 5.9 –Expression of soluble 6xHis-SAHtrA from pRSAHtrA1-3

Preliminary studies were carried out to investigate the best expression strain to use for expression studies of SAHtrA1 (a). *E. coli* M15 and *E. coli* BL21(DE3)CodonPlus were selected for further experiments, which included time courses and analysis of the effect of increasing IPTG concentrations on 6xHis-SAHtrA1 expression. Using the optimal parameters, a 10 ml culture was induced with IPTG and the pellet was sonicated to investigate the solubility of the protein. The sonicate (SON) and supernatant (SUP) from these samples were analysed by 10% SDS-PAGE gels and Western blot. Negative controls were carried out using the same strains transformed with pRSET-B only.

conditions used. Expression was maximal at 4 hr at both 26°C and 37°C and was expressed to similar levels when using increasing levels of IPTG (0.01 to 2.5 mM).

To investigate solubility of the 6xHis-SAHtrA1 inductions were carried out for 4 hr using 0.01 mM IPTG at 37°C. Bacterial pellets from a 10 ml induction were sonicated and cell debris separated by centrifugation. Figure 5.9b shows the SDS-PAGE and Western blot analysis of both sonicate and supernatant for each strain. The SDS-PAGE protein gel is inconclusive as 6xHis-SAHtrA1 runs closely to another protein band that is in the negative control (sonicate of cells containing pTrc99A only). However, the Western blot analysis shows that the protein is soluble at 37°C.

Small-scale purification of 6xHis-SAHtrA1 was performed in order to visualise the protein band on an SDS-PAGE gel. Purification was carried out using a commercially available Ni^{2+} -resin, Ni-NTA, which interacts with the 6xHis-tag and so aids in purification of the recombinant protein. The supernatant containing 6xHis-SAHtrA1 was allowed to bind to the resin and then washed with buffer. The 6xHis-SAHtrA1 was then eluted from the resin using imidazole, which competes for the divalent cation, Ni^{2+} . Figure 5.10 shows that although 6xHis-SAHtrA1 is purified to some extent by this method (Figure 5.10b), it is still indistinguishable by SDS-PAGE analysis (Figure 5.10a). Due to this, it was not feasible to carry out large-scale purification of this protein for further biochemical analysis.

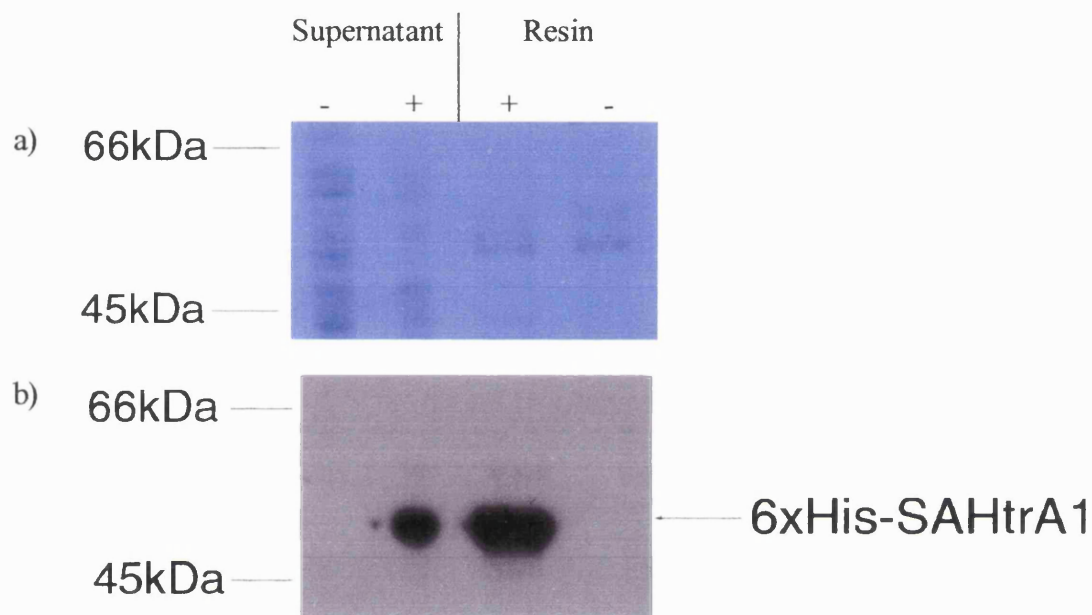


Figure 5.10 – Small-scale purification of 6xHis-SAHtrA1 in BL21(DE3)CodonPlus

Bacterial pellets were sonicated and cell debris removed by centrifugation. The supernatant was then allowed to bind to Ni-NTA resin for 2 hr. The resin was then washed with PBS buffer. An aliquot of resin (1 μ l) was loaded onto a 10% SDS-PAGE gel and Western blot analysis was also performed. Samples are denoted '+' and negative controls, containing empty pRSET-B, are denoted '-'.

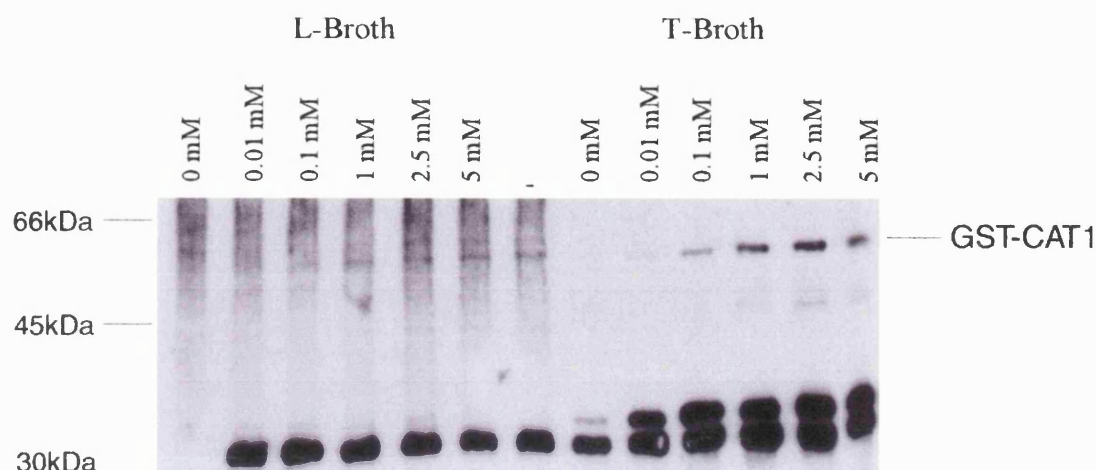
5.4 Cloning, expression and purification of SAHtrA1 subdomains

5.4.1 Cloning and expression of the catalytic domain of SAHtrA1

The catalytic domain of SAHtrA1 (CAT1) was determined using the ProfileScan algorithm (http://www.isrec.isb-sib.ch/software/PFSCAN_form.html). In order to examine the properties of the domain, the 867-bp ORF of CAT1 was cloned into the GST plasmid, pGEX-4T-1, using the cloning strategy described in Figure 5.3. The CAT1 domain was amplified by PCR with primers EMR3 and EMR5, which incorporate *Bam*HI and *Xho*I sites respectively. A stop codon (TAA) was also incorporated at the 3' end. The optimal MgCl₂ concentration for the PCR was determined as described (2.1.2.1) and found to be 2 mM. The PCR products and pGEX-4T-1 were digested with *Bam*HI and *Xho*I before ligation. The DNA sequence of the resulting pGCAT1 plasmids were determined as described in 2.1.6. The plasmid, pGCAT1-13, contained the correct DNA sequence and was transformed into *E. coli* BL21(DE3)pLysS and *E. coli* BL21(DE3)CodonPlus for expression.

The resultant GST-tagged CAT1 (GST-CAT1) protein could not be detected following electrophoresis on SDS-PAGE and Coomassie-staining and so detection of the protein had to be carried out by Western blot analysis. Overnight inductions at 26°C produced the maximal level of protein (data not shown) and so experiments increasing IPTG concentrations were carried out for this length of time. Figure 5.11 shows that the *E. coli* BL21(DE3)CodonPlus strain produces GST-CAT1 with fewer extra bands than *E. coli* BL21(DE3)pLysS. Western blot analysis also shows that there is degradation of the CAT as two smaller fragments are detected on the blots, which do not occur in the negative controls containing the empty pGEX-4T-1 plasmid. These degradation products may be typical of proteases with an

a) *E. coli* BL21(DE3)pLysS



b) *E. coli* BL21(DE3)CodonPlus

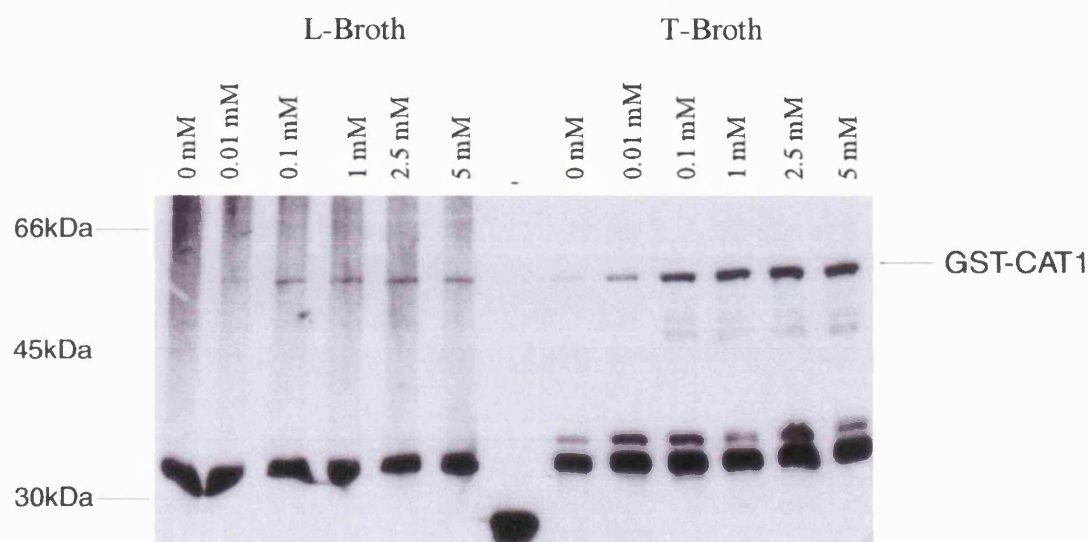


Figure 5.11 – Western blot analysis of the effect of increasing IPTG concentration on protein expression of pGCAT1-13 in BL21(DE3)pLysS and BL21(DE3)CodonPlus cells

The plasmid, pGCAT1-13 was induced for 16 hr at 26°C with a range of IPTG concentrations in *E. coli* BL21(DE3)pLysS and *E. coli* BL21(DE3)CodonPlus cells. The equivalent of OD600 of 0.01 was loaded onto 10% SDS-PAGE and expression of CAT1 checked by Western blot analysis probing with an anti-GST antibody. The negative control (-) was the expression of pGEX-4T-1 in the same strains using the same expression procedures.

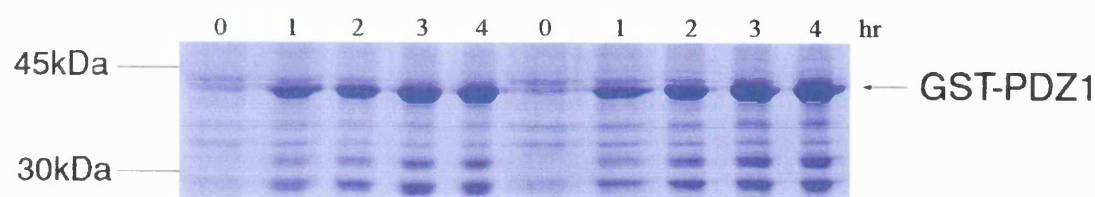
autoproteolytic activity, which the *E. coli* HtrA has been described to possess (Lipinska, B. *et al.*, 1990, Skórko-Glonek, J. *et al.*, 1995). Thus, due to the unstable nature of the recombinant protein, no further work was carried out on this protein domain.

5.4.2 Cloning and expression of the PDZ domain of SAHtrA1

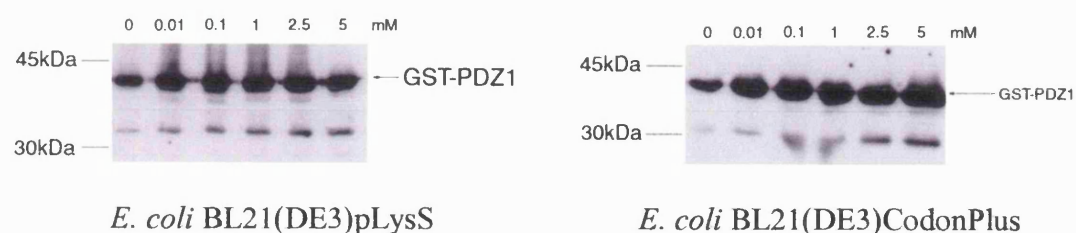
5.4.2.1 Cloning of the PDZ of SAHtrA1 as a GST-tagged protein

The PDZ domain of SAHtrA1 (PDZ1) was determined using the ProfileScan algorithm (http://www.isrec.isb-sib.ch/software/PFSCAN_form.html). In order to study to the properties of the domain, the 408-bp ORF of PDZ1 was cloned into the GST plasmid, pGEX-4T-1, using the cloning strategy in Figure 5.3. PDZ1 was amplified by PCR with primers EMR6 and EMR7, which incorporate *Bam*HI and *Xho*I sites respectively. The optimal MgCl₂ concentration for the PCR was 4 mM. The PCR products and pGEX-4T-1 were digested with *Bam*HI and *Xho*I before subsequently ligation. The DNA of the resulting pGPDZ1 plasmids were sequenced. The plasmid, pGPDZ1-1, contained the correct sequence and so was transformed into *E. coli* BL21(DE3)pLysS and *E. coli* BL21(DE3)CodonPlus for expression. The resultant GST-tagged PDZ1 (GST-PDZ1) was expressed to maximal levels in T-Broth induction experiments when compared to L-Broth (Figure 5.12a). Levels of expressed GST-PDZ1 was comparable in the two expression strains Figure 5.12a. However, from Western blot analysis (Figure 5.12b), the product from *E. coli* BL21(DE3)CodonPlus cells contained fewer GST-tagged products which did not correspond to GST-PDZ1 and so this strain was used for purification studies. Optimal conditions for induction of the protein were at 26°C for 16 hr with 0.01 mM IPTG determined from time courses and increasing IPTG concentration, performed

a) Time course of GST-PDZ1 in *E. coli* BL21(DE3)CodonPlus over 4 hr at 37°C



b) Western blot analysis of the effect of increasing concentrations of IPTG



c) Verification of soluble GST-PDZ1

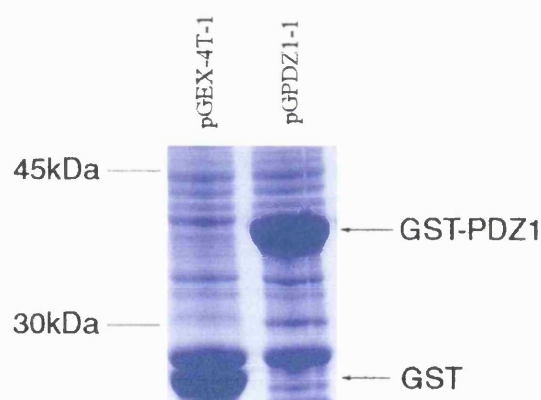


Figure 5.12 – Expression of soluble GST- PDZ1 from pGPDZ1-1

GST-PDZ1 was expressed from pGPDZ1-1 in *E. coli* BL21(DE3)pLysS and *E. coli* BL21(DE3)CodonPlus expression strains. Induction of GST-PDZ1 in BL21 cells was analysed over time (0 to 4 hr) at 37°C in L-Broth and T-Broth (a). Expression under increasing concentrations of IPTG (0 to 5 mM) in T-Broth at 26°C for 16 hr was analysed by Western blot analysis (b) for both expression strains. Supernatant from sonicated *E. coli* BL21(DE3)CodonPlus cells show that GST-PDZ1 is soluble (c).

as for previous recombinant protein expression experiments. Solubility of the recombinant protein was verified by loading the supernatant of sonicated cells onto an SDS-PAGE gel (Figure 5.12c).

5.4.2.2 *Purification of GST-PDZ1 by FPLC*

a) Affinity chromatography (GSTrap)

E. coli BL21(DE3)CodonPlus containing pGPDZ1-1 was inoculated into 10 ml of T-Broth and induced using the optimal induction conditions. In this case, 0.01 mM IPTG was used to induce production of GST-PDZ1 protein for 16 hr at 26°C. The cells were recovered by centrifugation and the bacterial pellet used for a small-scale purification. The pellet was disrupted by sonication as described in 2.2.3 and the cell debris separated by centrifugation. The supernatant was then incubated with GSTrap resin, which is conjugated with glutathione, a ligand for the GST-tag (data not shown). Binding of the GST-tag to the resin was allowed to proceed for 2 hr at 4°C. The resin was pelleted by centrifugation and the supernatant removed. The resin was then washed three times with PBS, and the GST-PDZ1 protein was eluted with 10 mM glutathione. GST-PDZ1 could be purified by this method and thus, large-scale purification by FPLC (2.2.4) was performed.

Cells from a 200 ml induction experiment were disrupted via sonication (2.1.3), and subsequently clarified by ultracentrifugation (2.1.3) and followed by filtering with a 0.22 µm filter. The supernatant was loaded at a slow flow rate (0.2 ml/min) due to the slow binding kinetics of GST to the column. The column was then washed with PBS and then the GST-PDZ1 was eluted with glutathione (2.2.4.1). A typical chromatogram is shown in Figure 5.13a. It can be seen in Figure 5.13b that GST-

a) Elution profile of GST-PDZ1 using affinity chromatography

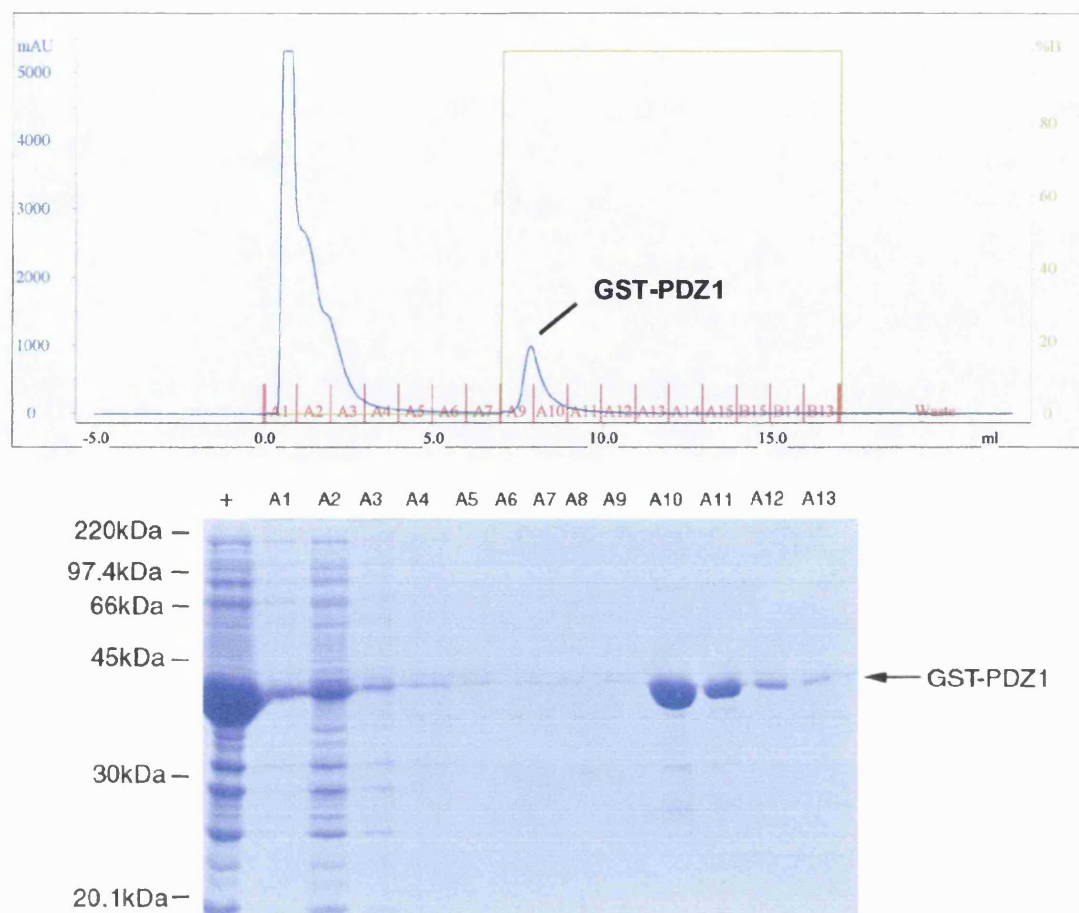


Figure 5.13 – Purification of GST-PDZ1 using GSTrap

GST-PDZ1 was isolated by affinity chromatography using GSTrap. A typical chromatogram (a) is shown where the blue line is the absorbance at 280 nm, a crude way of detecting protein. The green line corresponds to the percentage of the elution buffer, and each fraction collected is shown in red. The corresponding SDS-PAGE analysis of each fraction (2.5 μ l) is shown (b) where + is 1 μ l of the supernatant applied to the column.

PDZ1 was isolated with few contaminants in fractions A10 to A13. The typical yield from a pellet from a 200 ml induction experiment was 60 mg of protein.

As GST-PDZ1 has been purified from the majority of contaminating bacterial proteins, the next step was to remove the GST-tag to yield PDZ1. The plasmid, pGEX-4T-1, which had been used to clone the protein, has a thrombin cleavage site encoded after the GST-coding domain. This was utilised to obtain the pure PDZ1 domain. Firstly, conditions were examined to facilitate optimisation. This process was carried out because the thrombin may remove extra amino acids if incubated too long or if too high a concentration of enzyme was used. This may result in a variety of unwanted PDZ1 species with smaller molecular weights. Figure 5.14 shows a time-dependent digestion of GST-PDZ1 (calculated molecular mass 41.5 kDa) into its fragments of GST (26 kDa) and PDZ1 (calculated molecular mass 15.5 kDa). The addition of increasing concentrations of thrombin resulted in increased amounts of GST and PDZ1. However, at concentrations higher than 1U, another protein band becomes apparent, which is smaller than the full-length PDZ1. This product may result from thrombin digesting within PDZ1 during the incubation period. Thus, for the removal of the GST-tag from GST-PDZ1, 1 NIH U of thrombin was incubated per 200 µg GST-PDZ1 for 16 hr, as these conditions would favour an incomplete digestion and so reduce the likelihood of extra residues of GST-PDZ1 being removed.

b) Gel filtration chromatography (HiLoad Superdex 75 26/60 prep grade)

The next stage of purification utilised gel filtration chromatography in order to separate the different species yielded from the thrombin digest. Samples were loaded

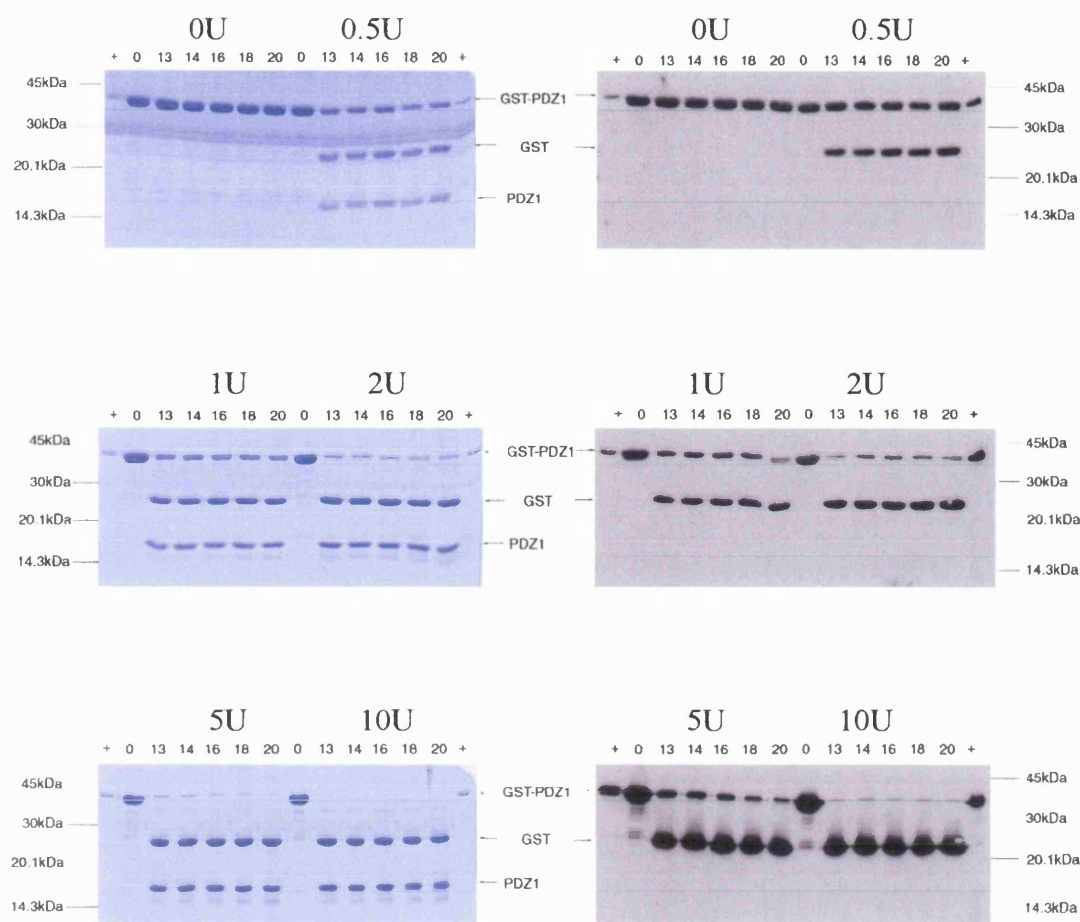


Figure 5.14 – Optimisation of thrombin digest of GST-PDZ1

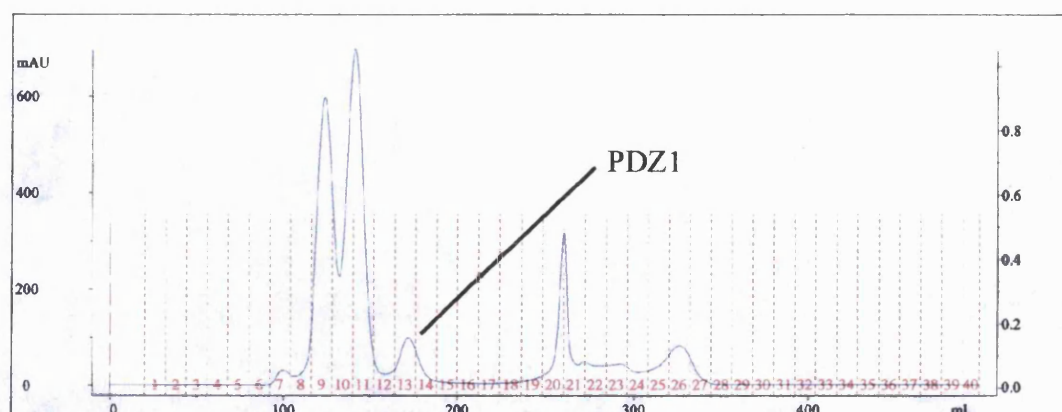
GST-PDZ1 (200 μ g) was incubated at room temperature with varying concentrations of thrombin (0 to 10U) over 20 hr. At each time point 20 μ g protein was removed and prepared for loading onto a 15% SDS-PAGE gel. 2.5 μ g protein was loaded per well and Western blot analysis was carried out to verify the presence of GST-tagged proteins.

onto a Superdex 75 26/60 prep grade column equilibrated with 20 mM MES pH6.6, 1mM DTT. A typical chromatogram is shown in Figure 5.15a, with the corresponding SDS-PAGE gel (Figure 5.15b). Analysis of the gel suggests that PDZ1 has been purified from the GST-tagged species, however, on examining the Western blot data (Figure 5.15c), PDZ1 appears to be contaminated with small amounts of both GST-PDZ1 and GST. Thus, another purification step was carried out.

c) *Cation exchange chromatography (Mono S 5/5 HR)*

In order to remove the contaminating GST species, the fractions of interest were pooled and clarified before loading onto a cation exchange column. The theoretical pI of PDZ1 is approximately 8, whereas the pI of the undigested fusion proteins and the GST-tag are approximately 6. Thus, at the correct pH, cation exchange chromatography can be used to purify the PDZ domains from the contaminating fusion protein. A typical chromatogram of the purification of PDZ1 is shown in Figure 5.16a, which shows that instead of one peak for PDZ1, there is a doublet. Improved separation of these peaks, however, could not be achieved. From the corresponding SDS-PAGE (Figure 5.16b), the two peaks seen in the chromatogram appear to be the same protein. However, as there is a difference as indicated by the chromatogram, it is possible that these one of these peaks is an overdigested species of PDZ1 as a result from thrombin cleavage, which may still be in the protein sample.

a) Elution profile of GST-PDZ1 digest reaction using gel filtration chromatography



b)



c)

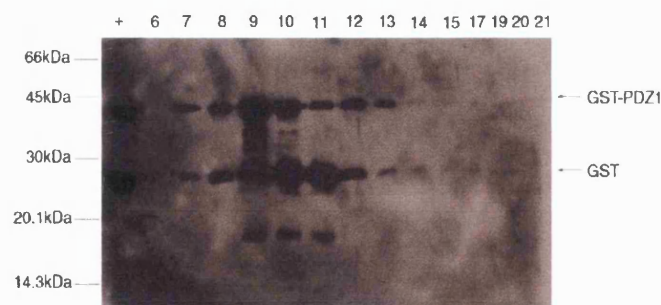
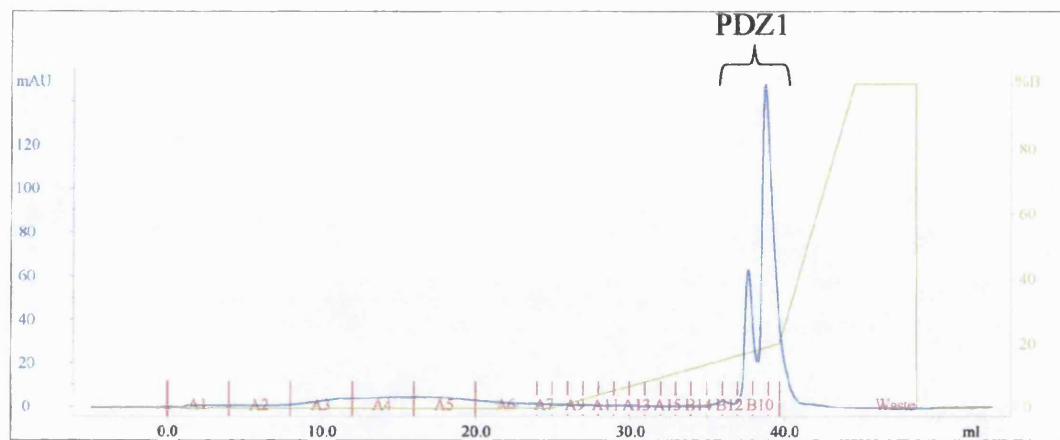


Figure 5.15 – Separation of species from the digest of GST-PDZ1

GST-PDZ1 digested with thrombin was loaded onto a Superdex 75 26/60 prep grade column. A typical chromatogram (a) is shown where the blue line is the absorbance at 280 nm and each fraction collected is shown in red. The corresponding SDS-PAGE analysis, where 2.5 μ l of each fraction was loaded (b), and Western blot analysis (c) of each fraction is shown where + is the digest reaction applied to the column.

a) Elution profile of cation exchange chromatography



b)

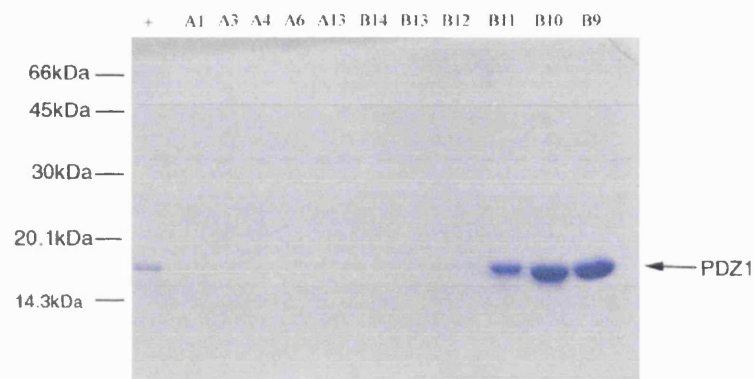


Figure 5.16 – Removal of contaminating GST-species using cation exchange chromatography

Fractions from gel filtration experiments containing PDZ1 were pooled and subjected to cation chromatography using MonoS 5/5 HR. A typical chromatogram (a) is shown where the blue line is the absorbance at 280 nm. The green line corresponds to the percentage of the elution buffer, and each fraction collected is shown in red. The corresponding SDS-PAGE analysis of each fraction (2.5 μ l loaded) is shown (b) where + is the pooled sample applied to the column.

d) Removal of thrombin by affinity chromatography

To remove any residual thrombin, the sample was loaded onto an affinity column, HiTrap Benzamidine FF column, an affinity column. This column consists of resin conjugated with benzamidine, an inhibitor of serine proteases, and thus will bind thrombin and allow the rest of the sample to pass through the column. This column was introduced in the purification procedure immediately after the digest step to reduce the possibility of overdigestion by thrombin. The elution profile analysed by SDS-PAGE is shown in Figure 5.17. Interestingly, the column had additional anion exchange properties, as it was also capable of removing undigested GST-PDZ1 and GST, which were eluted in a high salt step (1 M NaCl). Thus, the flowthrough fractions from this column, containing PDZ1, can be used in the gel filtration step in order to purify PDZ1. The removal of GST-PDZF1A and GST with this column removed the need for cation exchange chromatography.

e) Removal of residual low molecular weight contaminants

The gel filtration step used in 5.4.2.2b was used to remove any residual low molecular weight contaminants from the sample. One example would be the glutathione used in the first step of purification. The pooled flowthrough fractions from the HiTrap Benzamidine column were loaded onto the Superdex 75 26/60 prep grade column and fractions collected. A typical elution profile is shown in Figure 5.18a. To check that the PDZ1 protein was pure, fractions were loaded onto a 20% SDS-PAGE gel, which was subsequently silver-stained as this is a much more sensitive protein detection method than Coomassie staining (Figure 5.18b).

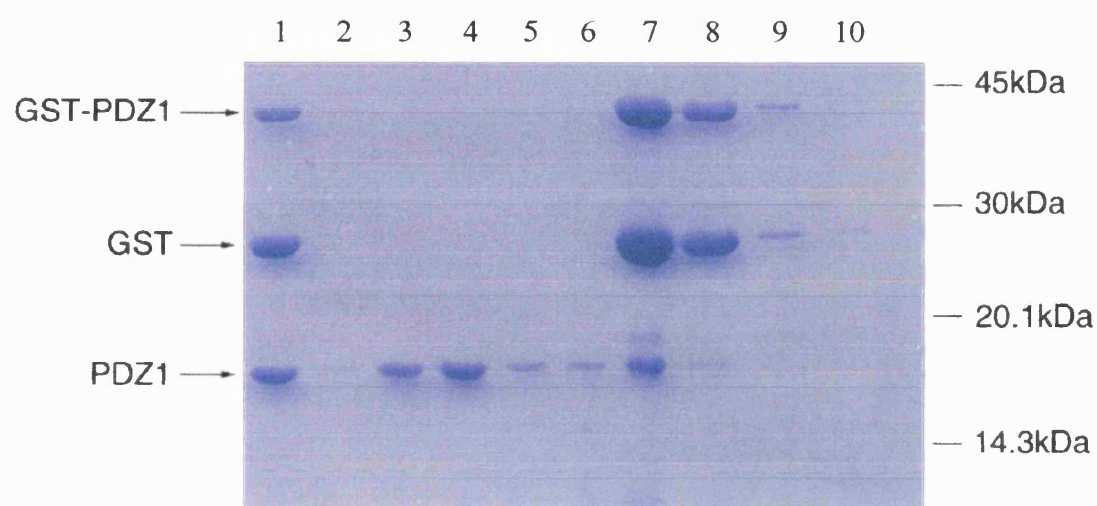
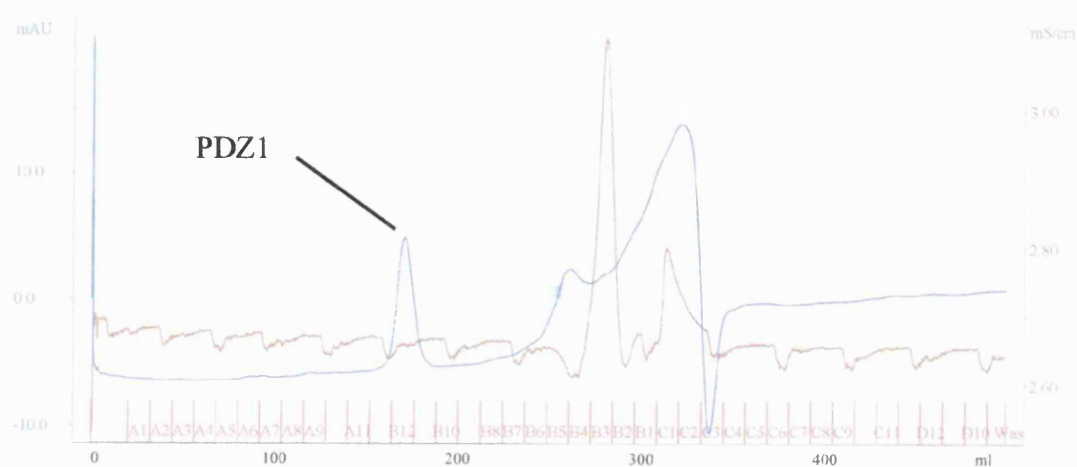


Figure 5.17 – Removal of thrombin from the digest reaction by HiTrap Benzamidine FF

GST-PDZ1 isolated from the GSTrap column was digested with thrombin and then loaded onto a HiTrap Benzamidine FF column. Flowthrough fractions (lanes 2-6) were collected. A high salt wash (1 M NaCl) was used to remove any other proteins, which had adhered to the column via electrostatic interactions (lanes 7-10). Lane 1 contains the digest mix loaded onto the column. 2.5 μ l was loaded per fraction was loaded onto the gel.

a) Elution profile of gel filtration chromatography



b)

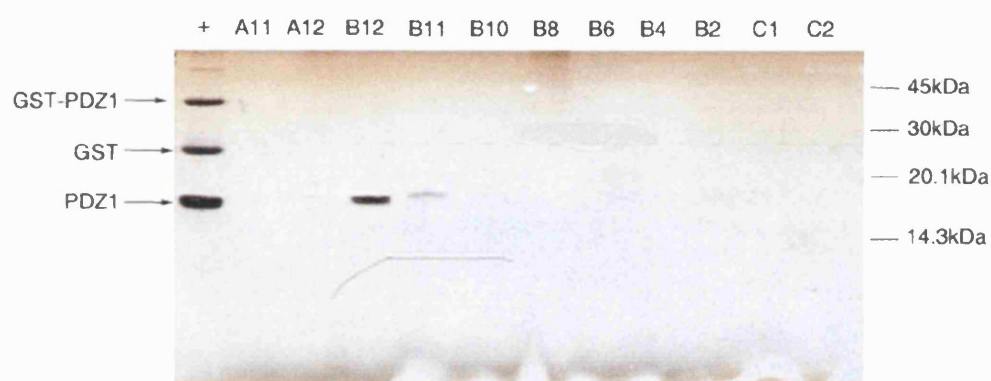


Figure 5.18 – Removal of residual low molecular weight contaminants from PDZ1 by gel filtration chromatography

Flowthrough fractions from HiTrap Benzamidine experiments containing PDZ1 were pooled and subjected to gel filtration using the Superdex 75 26/60 prep grade column. A typical chromatogram (a) is shown where the blue line is the absorbance at 280 nm. The brown line corresponds to the conductivity of the buffer, and each fraction collected is shown in red. The corresponding SDS-PAGE analysis of each fraction (2.5 μ l loaded) is shown (b) where + is the pooled sample applied to the column.

The PDZ1 domain was generated using recombinant protein procedures. Though these procedures can produce a protein of the correct sequence, which can then be purified, the protein may not have folded correctly and therefore be unusable for characterisation procedures. Checking that a recombinant protein is folded correctly is often overlooked. Circular dichroism (CD) is a spectrometric technique which can provide information on the secondary structure of proteins. Circular dichroism experiments were carried out by Dr. Heather Lamb (University of Newcastle) and showed that PDZ1 had secondary structure. Further work is planned to investigate the binding properties of this domain utilising a method called surface plasmon resonance (SPR), using the commercial BIAcore 2000 biosensor (Pharmacia Biosensor AB). BIAcore 2000 allows the detection and monitoring of binding events between two or more molecules on the sensor surface. One interactant, the ligand, is immobilised on the sensor surface and the other interactant, the analyte, is injected over the immobilised ligand in a controlled flow. Any interactions between the molecules are detected by a change in mass at the sensor surface and measure in Resonance Units (RU). This change is measured continuously to produce a binding-progress curve (sensorgram) and allows the estimation of binding rate constants. However, the PDZ1 domain has a N-terminal GST-tag, and it has been argued that the use of GST fusion proteins in SPR studies can overestimate the affinity of a measured interaction (Ladbury, J.E. *et al.*, 1995). This is proposed to be due to the GST-fusion dimers binding simultaneously to ligand molecules on the sensor surface. Thus, to clarify any interactions that may be shown with this technique, the PDZ1 domain was also cloned with a 6xHis-tag.

5.4.2.3 Cloning of the PDZ of *SAHtrA1* as a 6xHis-tagged protein

The use of GST fusion proteins may be unsuitable for biological evaluation for accurate SPR studies (Ladbury, J.E. *et al.*, 1995). To generate an alternative fusion protein, the PDZ1 domain was cloned as a 6xHis-tagged protein. PDZ1 was cloned with a *Bam*HI at the 5' terminus and a *Hind*III site at the 3' terminus by incorporating these sites into the PCR primers used. Primers EMR6 and EMR2 were used to clone PDZ1 into pQE-30, while EMR8 and EMR2 were used for pRSET-B. Optimal MgCl₂ concentrations for PCR were determined to be 4 mM in both cases (2.1.2.1). PDZ1 was then cloned into pQE-30 and pRSET-B by the cloning strategies shown in Figure 5.4 and Figure 5.5, respectively. The resultant plasmids were pQPDZ1 and pRPDZ1, respectively. Single *E. coli* DH5 α colonies, which had been transformed with ligation mix, were cultured to obtain plasmids for sequencing. Plasmids containing the correct DNA sequence were named pQPDZ1-10 and pRPDZ1-3.

Expressed proteins were separated on a Tricine gel due to the small molecular weights of the 6xHis-tagged PDZ1 (6xHis-PDZ1; calculated molecular mass 19.2 kDa). The 6xHis-tagged PDZ1 (6xHis-PDZ1) were induced with 1 mM IPTG in a variety of *E. coli* expression strains. Expression from pQPDZ1 in several bacterial expression strains failed to yield detectable levels of 6xHis-PDZ1 by Western blot analysis and so was discarded. However, four expression strains produced 6xHis-PDZ1 on expression from pRPDZ1 (Figure 5.19a). Induction conditions were determined as for previous recombinant proteins (data not shown). Maximal expression of 6xHis-PDZ1 was induced by 0.01 mM IPTG for 16 hr at 26°C. The protein was judged to be soluble, thus, a small-scale purification using Ni-NTA resin

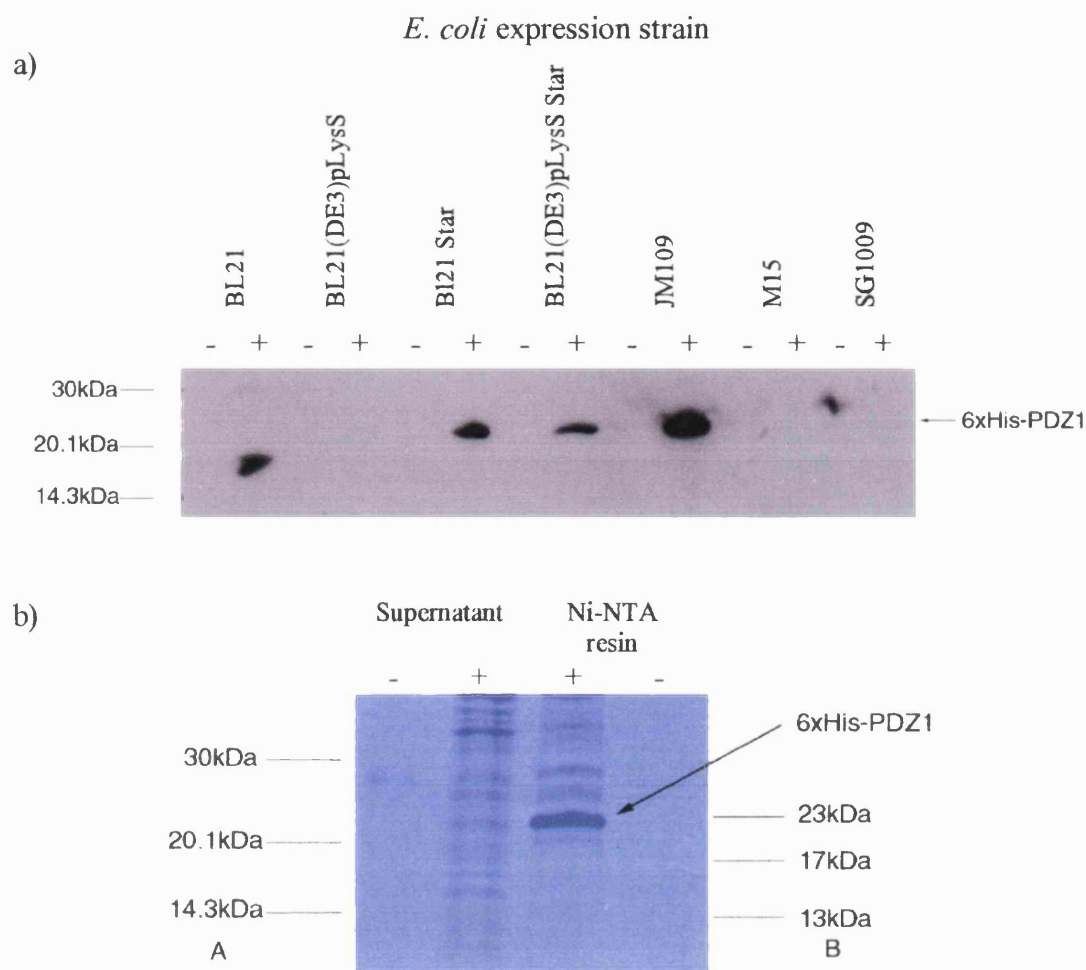


Figure 5.19 – Expression of 6xHis-PDZ1 from pRPDZ1-3

Seven bacterial expression strains were transformed with pRSET-B (-) or pRPDZ1-3 (+) and induced for 4 hr at 37°C with 1 mM IPTG. Proteins from 5 µl culture separated on a 12% Tricine gel by electrophoresis before transferral onto nitrocellulose for Western blot analysis using anti-His antibodies (a). Induction conditions were optimised and a 10 ml induction culture was pelleted. The supernatant was separated from cell debris and subjected to a small-scale purification using Ni-NTA resin. Supernatant (1 µl) and resin bound to protein (8 µl) was loaded per well (b). Negative controls, containing pRSET-B only, were treated in the same manner as the samples. Ladder A is the rainbow marker (Amersham) and ladder B is MultiMark® multicoloured standard (Invitrogen).

was performed (Figure 5.19b). Previous gels, (Figure 5.19a) indicated that 6xHis-PDZ1 was larger than the expected 19kDa. However, this was thought to be due to the protein markers used as the DNA sequence of the cloned fragment was correct and Western blot analysis gave a strong single band. To test this hypothesis, a second marker (MultiMark® multi-coloured marker (Invitrogen) was run in conjunction, and this showed that 6xHis-PDZ1 purified under the conditions used (2.2.5) was approximately the correct size.

Thus, 6xHis-PDZ1 can be expressed as a soluble protein and partially purified affinity chromatography.

5.5 Cloning, expression and purification of SAHtrA2 and its subdomains

5.5.1 Cloning and expression of SAHtrA2

SAHtrA2 was initially constructed as a native recombinant protein. In order to do this, the gene was amplified with *Nco*I and *Xba*I sites at the 5' and 3' termini, respectively. The optimal MgCl₂ concentration was determined to be 4 mM for PCR reactions. The gene was cloned as shown in Figure 5.2, using primers EMR9 and EMR10, which incorporated the restriction sites mentioned earlier. However, initial experiments showed that no plasmid was obtained with the correct sequence. To obtain a plasmid with the correct sequence, *sahtrA2* was cloned in two segments utilising the single *Cla*I site at position 856 of the *sahtrA2* gene. The 5' segment was amplified by PCR using primers EMR9 and EMR11, whereas the 3' segment was amplified by PCR using primers EMR12 and EMR10 and TA-cloned into the *lacZ* gene in pGEM-T-Easy plasmid. DNA sequence of the plasmids was determined and

the plasmids containing the correct DNA sequenced, pSAHtrA25 and pSAHtrA23 were transformed into *E. coli* DH5 α cells. Single colonies containing pSAHtrA25 or pSAHtrA23 were cultured and the plasmids were isolated using the QIAprep miniprep kit. The plasmid containing the 5' segment, pSAHtrA25 was excised with *Sac*II and *Cla*I and ligated into the plasmid containing the 3' segment, pSAHtrA23, treated in the same manner. Conversely, the 3' segment was excised from pSAHtrA23 by digestion with *Sac*I and *Cla*I, and ligated into pSAHtrA25, which had been treated with the same enzymes. The two ligation mixes were transformed separately into JM109 cells. Colonies were picked to amplify the plasmids, which were then isolated and sequenced to check that the segments had ligated correctly. It was observed from separate cloning experiments using different PCR reactions that there were several differences in the *sahtrA2* gene in the *S. aureus* SMITH strain compared to *S. aureus* NCTC 8325. However, the differences in the SMITH strain were identical to those recently published data for the MRSA strain (N315) and the VRSA strain (Mu50) (Kuroda, M. *et al.*, 2001). It should also be noted that the NCTC 8325 strain is an unfinished genome at the moment. Figure 5.20 illustrates the changes in amino acid sequence caused by these sequence changes and compares the sequences of the four strains discussed.

The *sahtrA2* gene was excised from the pGEM-T Easy plasmid with *Nco*I and *Xba*I and ligated into pTrc99A. Again, the resultant pTSAHtrA2 plasmids were sequenced to verify the insert had ligated correctly into the plasmid. The plasmid pTSAHtrA2-715 contained the correct sequence and was used in expression studies. However, on induction, the expected protein band (calculated molecular mass 87.1 kDa) was not seen, although a band of approximately 66 kDa was seen when the protein gels were

NCTC 8325	MEWTLVDIGK	KHVIPKSQYR	RKRREFFHNE	DREENLNQH	DKQNIDNTTS	KKADKQIHKD	SIDKHERFKN	70
SMITH	MEWTLVDIGK	KHVIPKSQYR	RKRREFFHNE	DREENLNQH	DKQNIDNTTS	KKADKQIHKD	SIDKHERFKN	70
N315	MEWTLVDIGK	KHVIPKSQYR	RKRREFFHNE	DREENLNQH	DKQNIDNTTS	KKADKQIHKD	SIDKHERFKN	70
Mu50	MEWTLVDIGK	KHVIPKSQYR	RKRREFFHNE	DREENLNQH	DKQNIDNTTS	KKADKQIHKD	SIDKHERFKN	70
NCTC 8325	SLSSHLEQRN	RDVVENKAAE	SKSNQDSKSA	YNRDHYLTDD	VSKKQNSLDS	VDQDTEKSKY	YEQNSEATLS	140
SMITH	SLSSHLEQRN	RDVVENKAAE	SKSNQDSKSA	YNRDHYLTDD	VSKKQNSLDS	VDQDTEKSKY	YEQNSEATLS	140
N315	SLSSHLEQRN	RDVVENKAAE	SKSNQDSKSA	YNRDHYLTDD	VSKKQNSLDS	VDQDTEKSKY	YEQNSEATLS	140
Mu50	SLSSHLEQRN	RDVVENKAAE	SKSNQDSKSA	YNRDHYLTDD	VSKKQNSLDS	VDQDTEKSKY	YEQNSEATLS	140
NCTC 8325	TKSTDKVEST	EMRKLSDDKN	KVGHEEQHVL	SKPSEHDKET	RIDSESSRTD	SDSSMQTEKI	KKDSSDGNKS	210
SMITH	TKSTDKVEST	EMRKLSDDKN	KVGHEEQHVL	SKPSEHDKET	RIDSESSRTD	SDSSMQTEKI	KKDSSDGNKS	210
N315	TKSTDKVEST	EMRKLSDDKN	KVGHEEQHVL	SKPSEHDKET	RIDSESSRTD	SDSSMQTEKI	KKDSSDGNKS	210
Mu50	TKSTDKVEST	EMRKLSDDKN	KVGHEEQHVL	SKPSEHDKET	RIDSESSRTD	SDSSMQTEKI	KKDSSDGNKS	210
NCTC 8325	SNLKSEVSD	KSNTVPKLSE	SDDEVNNQKP	LTLPEEQKLK	RQSQSNEQTK	TYTYGDSEQN	DKSNHENDLS	280
SMITH	SNLKSEVSD	KSNTVPKLSE	SDDEVNNQKP	LTLPEEQKLK	RQSQSNEQTK	TYTYGDSEQN	DKSNHENDLS	280
N315	SNLKSEVSD	KSNTVPKLSE	SDDEVNNQKP	LTLPEEQKLK	RQSQSNEQTK	TYTYGDSEQN	DKSNHENDLS	280
Mu50	SNLKSEVSD	KSNTVPKLSE	SDDEVNNQKP	LTLPEEQKLK	RQSQSNEQTK	TYTYGDSEQN	DKSNHENDLS	280
NCTC 8325	HHTPSI SDDK	DNVMRENHIV	DDNPNDINT	LSLSKIDDDR	KLDEKIHVED	KHKQNADSSE	TVGYQSQSTA	350
SMITH	HHTPSI SDDK	DNVMRENHIV	DDNPNDINT	LSLSKIDDDR	KLDEKIHVED	KHKQNADSSE	TVGYQSQSTA	350
N315	HHTPSI SDDK	DNVMRENHIV	DDNPNDINT	LSLSKIDDDR	KLDEKIHVED	KHKQNADSSE	TVGYQSQSTA	350
Mu50	HHTPSI SDDK	DNVMRENHIV	DDNPNDINT	LSLSKIDDDR	KLDEKIHVED	KHKQNADSSE	TVGYQSQSTA	350
NCTC 8325	SHRSTEKRN	SINDHDKLNG	QKTNTKTSAN	NNQKKATSKL	NKGRATNNNY	SDILKKFWM	YWPKLVLIMG	420
SMITH	SHRSTEKRN	SINDHDKLNG	QKTNTKTSAN	NNQKKATSKL	NKGRATNNNY	SDILKKFWM	YWPKLVLIMG	420
N315	SHRSTEKRN	SINDHDKLNG	QKTNTKTSAN	NNQKKATSKL	NKGRATNNNY	SDILKKFWM	YWPKLVLIMG	420
Mu50	SHRSTEKRN	SINDHDKLNG	QKTNTKTSAN	NNQKKATSKL	NKGRATNNNY	SDILKKFWM	YWPKLVLIMG	420
NCTC 8325	IIILIVILNA	IFNNVNKNDR	MNDNNDADAQ	KYTTTMMKAN	NTVKS SVTVE	NETSKDSSLP	KDKASQDEVG	490
SMITH	IIILIVILNA	IFNNVNKNDR	MNDNNDADAQ	KYTTTMMKAN	NTVKS SVTVE	NETSKDSSLP	KDKASQDEVG	490
N315	IIILIVILNA	IFNNVNKNDR	MNDNNDADAQ	KYTTTMMKAN	NTVKS SVTVE	NETSKDSSLP	KDKASQDEVG	490
Mu50	IIILIVILNA	IFNNVNKNDR	MNDNNDADAQ	KYTTTMMKAN	NTVKS SVTVE	NETSKDSSLP	KDKASQDEVG	490
NCTC 8325	SGVVYKSGD	TLYIVTNAHV	VGDKENQKIT	FSNNKSVVGK	VLGKDKWSDL	AVVKATSSDS	SVKEIAIGDS	560
SMITH	SGVVYKSGD	TLYIVTNAHV	VGDKENQKIT	FSNNKSVVGK	VLGKDKWSDL	AVVKATSSDS	SVKEIAIGDS	560
N315	SGVVYKSGD	TLYIVTNAHV	VGDKENQKIT	FSNNKSVVGK	VLGKDKWSDL	AVVKATSSDS	SVKEIAIGDS	560
Mu50	SGVVYKSGD	TLYIVTNAHV	VGDKENQKIT	FSNNKSVVGK	VLGKDKWSDL	AVVKATSSDS	SVKEIAIGDS	560
NCTC 8325	NNLVLGEPIL	VVGNPLGVDF	KGTVTEGIIS	GLNRRNPIDF	DKDNKYDMLM	KAFQIDASVN	PGNSGGAVVN	630
SMITH	NNLVLGEPIL	VVGNPLGVDF	KGTVTEGIIS	GLNRRNPIDF	DKDNKYDMLM	KAFQIDASVN	PGNSGGAVVN	630
N315	NNLVLGEPIL	VVGNPLGVDF	KGTVTEGIIS	GLNRRNPIDF	DKDNKYDMLM	KAFQIDASVN	PGNSGGAVVN	630
Mu50	NNLVLGEPIL	VVGNPLGVDF	KGTVTEGIIS	GLNRRNPIDF	DKDNKYDMLM	KAFQIDASVN	PGNSGGAVVN	630
NCTC 8325	REGKLIGVVA	AKISMPNVEN	MSFAIPVNEV	QKIVKDLETK	GKIDYPDVGV	KMKNIASLNS	FERQAVKLP	700
SMITH	REGKLIGVVA	AKISMPNVEN	MSFAIPVNEV	QKIVKDLETK	GKIDYPDVGV	KMKNIASLNS	FERQAVKLP	700
N315	REGKLIGVVA	AKISMPNVEN	MSFAIPVNEV	QKIVKDLETK	GKIDYPDVGV	KMKNIASLNS	FERQAVKLP	700
Mu50	REGKLIGVVA	AKISMPNVEN	MSFAIPVNEV	QKIVKDLETK	GKIDYPDVGV	KMKNIASLNS	FERQAVKLP	700
NCTC 8325	KVKNGVVVDQ	VDNNGLADQS	GLKKGDVITE	LDGKLEDDL	RFRQIIFSHK	DDLKSITAKI	YRDGKEKEIN	770
SMITH	KVKNGVVVDQ	VDNNGLADQS	GLKKGDVITE	LDGKLEDDL	RFRQIIFSHK	DDLKSITAKI	YRDGKEKEIN	770
N315	KVKNGVVVDQ	VDNNGLADQS	GLKKGDVITE	LDGKLEDDL	RFRQIIFSHK	DDLKSITAKI	YRDGKEKEIN	770
Mu50	KVKNGVVVDQ	VDNNGLADQS	GLKKGDVITE	LDGKLEDDL	RFRQIIFSHK	DDLKSITAKI	YRDGKEKEIN	770
NCTC 8325	IKLK	774						
SMITH	IKLK	774						
N315	IKLK	774						
Mu50	IKLK	774						

Figure 5.20 – Alignment of *sahtrA2* in *S. aureus* SMITH strain against the NCTC 8325, N315 and Mu50 *S. aureus* strains

The *sahtrA2* gene from *S. aureus* SMITH strain is aligned against *S. aureus* strains NCTC 8325, N315 and Mu50. Identical residues are highlighted in pale yellow while differences in amino acid sequence are highlighted in blue.

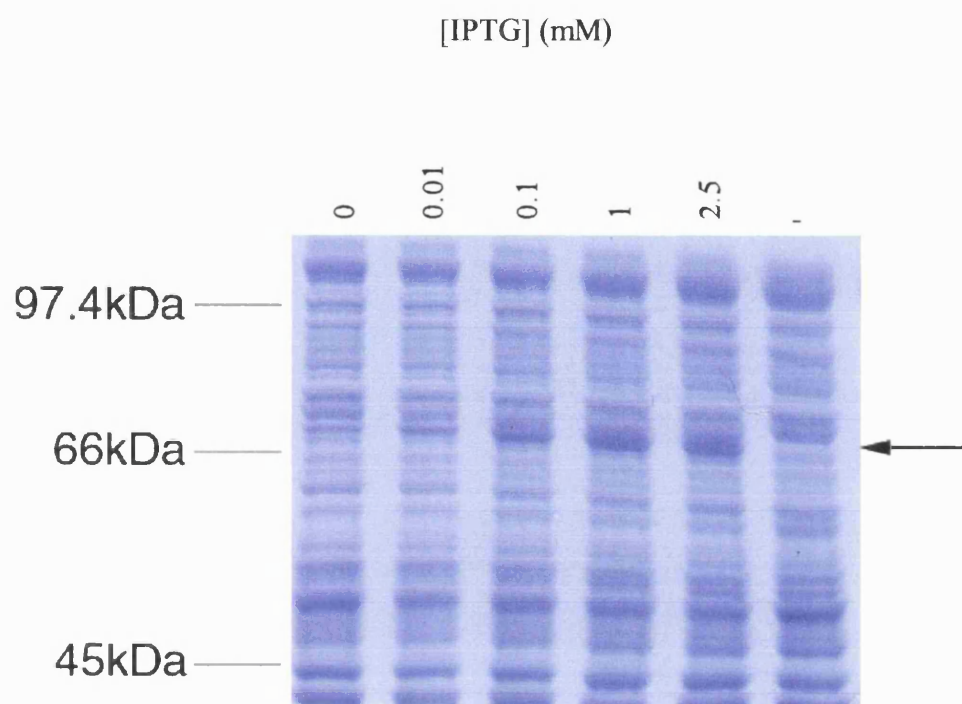


Figure 5.21 – Protein expression from pTSAHtrA2-715

E. coli BL21(DE3)CodonPlus cells were transformed with pTSAHtrA2-715. Protein expression was induced with increasing concentrations of IPTG for 16 hr at 26°C. The equivalent of OD₆₀₀ = 0.1 was loaded per well. A negative control (-) containing the empty pTrc99A was carried out, with protein induction with 1 mM IPTG under the same conditions. The arrow (←) denotes the a protein of approximately 68 kDa which increases with time.

stained with Coomassie blue (Figure 5.21). Because the SAHtrA2 was expressed as a native protein, it was not possible to determine if the 66 kDa protein band corresponded to any component of SAHtrA2 protein. As a consequence of this, SAHtrA2 was cloned and expressed as a N-terminal 6xHis-tagged protein using the fusion plasmids, pQE-30 and pRSET-B. Cloning was carried in a similar manner as before except that pTSAHtrSA2-715 served as the template DNA in the PCR reactions. Primers, EMR13 and EMR14 were used to clone *sahtrA2* into pQE-30 as shown in Figure 5.4, whereas primers EMR15 and EMR14 were used to clone *sahtrA2* into pRSET-B as shown in Figure 5.5. The resultant pQSAHtrA2 and pRSAHtrA2 plasmids were sequenced and the correct plasmids, pQSAHtrA2-2 and pRSAHtrA2-2 were transformed into the *E. coli* expression strains, BL21, M15, SG1009 and BL21(DE3)CodonPlus expression strains. Unfortunately, pQSAHtrA2-2 did not express well and so was discarded. Figure 5.22 shows that expression of the 6xHis-tagged SAHtrA2 (6xHis-SAHtrA2) occurs in all strains except *E. coli* BL21 cells. A band corresponding SAHtrA2 (calculated molecular mass of 90.6 kDa) is strongest in the M15 cells. However, there is a heavier band, which is expressed in the positive lanes, which is not present in the negative lanes. This band cannot be explained, as the plasmid DNA sequence is correct and the protein samples are electrophoresed under reducing conditions, so it cannot be dimerisation. Due to the unidentified larger band, no further work was carried out on SAHtrA2.

E. coli expression strains

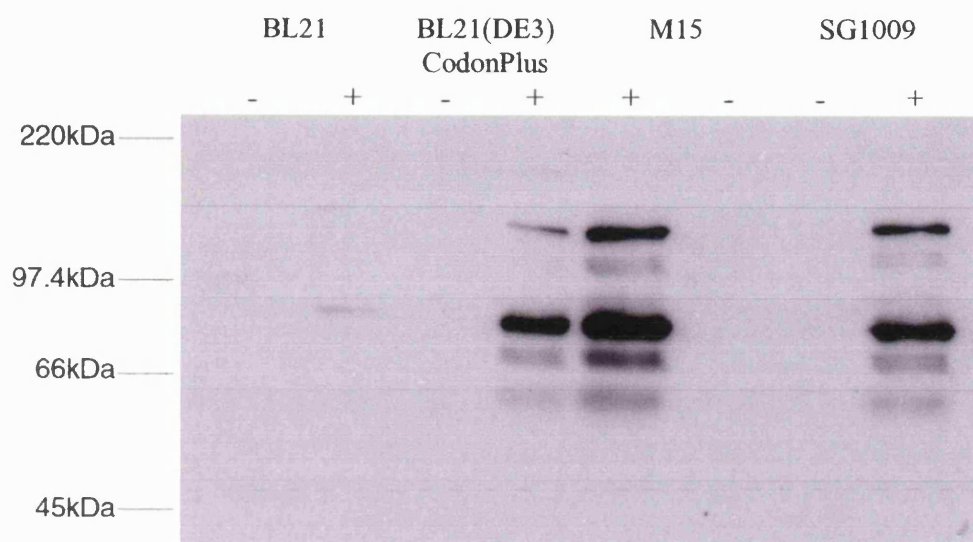


Figure 5.22 – Expression of SAHtrA2 using pRSAHtrA2-2

E. coli expression strains, BL21, BL21(DE3)CodonPlus, M15 and SG1009, were transformed with pRSAHtrA2-2. Single colonies were cultured to an OD600 of 1.0 and then induced with 1 mM IPTG for 4 hr. 100 µl culture was pelleted and resuspended in 100 µl dH₂O. Solubilisation buffer (2X) was added and samples (+, 5 µl) ran on an 8% SDS-PAGE protein gel. Protein was then transferred onto nitrocellulose by Western blotting and probed with anti-His antibodies. Negative controls (-, 5 µl) were carried out using the same conditions except that pRSET-B was transformed into the expression strains.

5.5.2 Cloning, expression and purification of the PDZ domain of SAHtrA2

5.5.2.1 Cloning and expression of PDZ2 as a GST-tagged protein

To investigate the function of the PDZ domain of SAHtrA2, the coding sequence (357-bp) was cloned into pGEX-4T-1. The coding sequence was amplified by PCR using the optimal concentration of MgCl₂ (4 mM) with the primers EMR18 and EMR19, which incorporated the *Bam*HI and *Sma*I sites, respectively. The resultant pGPDZ2 plasmid was generated in a manner shown in Figure 5.3 and plasmids obtained from single DH5α colonies were sequenced. The correctly sequenced pGPDZ2-1 was transformed into the *E. coli* expression strains BL21 and BL21(DE3)CodonPlus. Expression of GST-PDZ2 occurred in both strains used and although the levels of GST-PDZ2 was much smaller in *E. coli* BL21(DE3)CodonPlus than *E. coli* BL21, the product was much cleaner (Figure 5.23a). Thus, for further expression studies, *E. coli* BL21(DE3)CodonPlus was used. The optimal expression conditions were determined to be at 26°C for 16 hr using the 1.0 mM IPTG, using time course experiments and investigating the effect of IPTG. GST-PDZ2 was verified to be soluble (Figure 5.23b).

GST-PDZ2 was purified by the same techniques used to purify GST-PDZ1. Affinity chromatography was performed to isolate the GST-PDZ2 by interaction of the GST-tag with glutathione conjugated onto the resin of a GSTrap column. The protein was eluted from this column using glutathione and was subsequently treated with thrombin for 16 hr. Thrombin was then removed by the HiTrap Benzamidine FF column in which the thrombin binds to the column and the cleaved untagged PDZ2 passes through. Due to the anion exchange properties of this column, the uncleaved

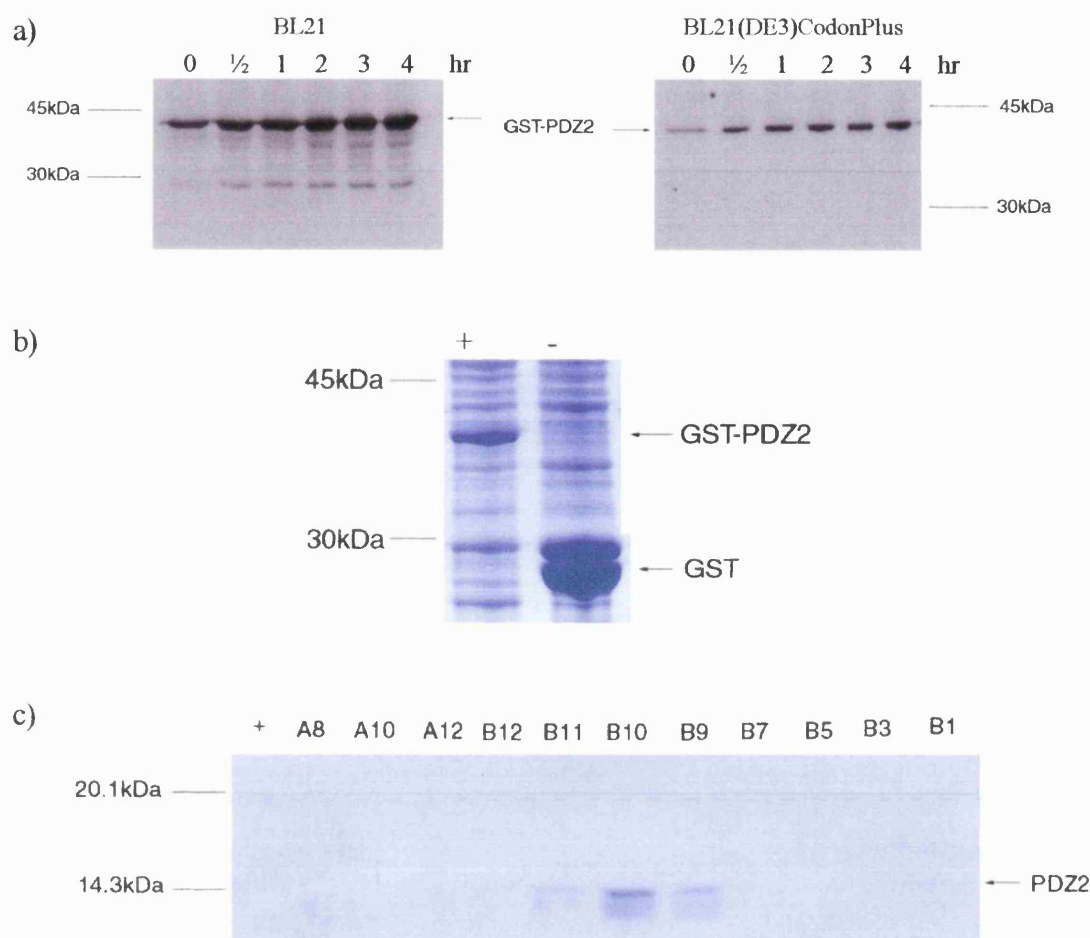


Figure 5.23 – Expression of soluble GST-PDZ2 using pGPDZ2-1

E. coli BL21 and *E. coli* BL21(DE3)CodonPlus, transformed with pGPDZ2-1, were induced with 1 mM IPTG over a period of 4 hr. Expression of the protein was followed by Western blot analysis (a). A bacterial pellet obtained from optimal expression conditions for *E. coli* BL21(DE3)CodonPlus was sonicated and cellular debris was pelleted. The supernatant (+, 2 μ l) was analysed by SDS-PAGE (b). A negative control using the same strain transformed with pGEX-4T-1. PDZ2 was then purified using various FPLC methods and selected fractions from the final purification procedure analysed by SDS-PAGE (c).

GST-PDZ2 and GST-tag are also retained on the column. The PDZ2 sample was then further purified by gel filtration to remove small molecules, such as glutathione, from the solution. To check protein purity, the samples from the fractions collected during gel filtration were loaded onto a 20% SDS-PAGE gel and stained with Coomassie. Figure 5.23c shows that although GST-PDZ2 is pure, there is a smear underneath the protein band, which could indicate that the protein is not stable in the buffer used. Circular dichroism experiments carried out by Dr. Heather Lamb at the University of Newcastle showed that the GST-PDZ2 fusion protein had secondary structure.

It is thought that the PDZ domains of *S. aureus* HtrA orthologues are important in ligand binding and recently it was shown that the PDZ domains of *E. coli* and *S. typhimurium* HtrA could bind the SsrA-tag (discussed earlier in 1.8.3.1). The ability of PDZ2 to bind the wild-type *S. aureus* SsrA-tag, GKSNNNF₁₀AVAA, was investigated using SPR (discussed earlier in 5.4.2.2e). SPR experiments were carried out by Miss Alison Spiers (University of Newcastle) and showed that the GST-PDZ2 could bind to the wildtype SsrA-tag with an apparent equilibrium dissociation constant, K_D , of 26×10^{-9} . No interaction was observed with a SsrA variant with the same amino acid composition but in randomised order (FANVNAGANSK). In addition, there was no interaction between the PDZ domain with a SsrA variant where the C-terminal alanine residue is replaced with a glutamate residue (GKSNNNF₁₀AVAE). In both cases, the C-terminal residue of the SsrA-tag was a polar residue (lysine or glutamate, respectively). This suggests that a non-polar C-terminal amino acid is important in the interaction between PDZ2 and the *S. aureus* SsrA-tag and this interaction is abolished by replacing the non-polar C-terminal

amino acid with a polar amino acid. Thus, the PDZ domains of the SAHtrAs appear to be involved in protein:protein interactions suggesting that it could recruit substrates for subsequent proteolysis.

As discussed earlier (5.4.2.2e), although the SPR experiments can show that an interaction is occurring between PDZ2 and the SsrA-tag, the affinity of this interaction may be overestimated due to the GST-moiety. Therefore, PDZ2 was also cloned as a 6xHis-tagged protein.

5.5.3 Cloning, expression and purification of 6xHis-tagged PDZ2

In order to clarify the PDZ2:SsrA interaction, PDZ2 was cloned as a N-terminal 6xHis-tagged protein. The coding sequence of PDZ2 was cloned into pQE-30 and pRSET-B as shown in Figure 5.4 and Figure 5.5, respectively. The insert DNA for pQE-30 was cloned by PCR using primers EMR16 and EMR17, which incorporated *Bam*HI and *Sma*I sites, respectively. The insert DNA for pRSET-B was cloned by PCR using primers EMR18 and EMR17, which incorporated *Bam*HI and *Nco*I sites, respectively. The optimal MgCl₂ concentrations for both PCR reactions were determined to be 4 mM (2.1.2.1) and the template DNA used was pGPDZ2-2. The resultant plasmids, pQPDZ2 and pRPDZ2, were obtained and transformed into DH5α cells. Single colonies were cultured to amplify the plasmid of interest. The isolated plasmids were then sequenced to identify clones with the correct sequence.

The plasmids, pQPDZ2-2 and pRPDZ2-3, were judged to contain the correct sequence and so were transformed in several expression strains for induction of the protein. Neither *E. coli* BL21 and *E. coli* BL21(DE3)CodonPlus strains containing

pQPDZ2-2 or pRPDZ2-3 grew very well. In addition, the *E. coli* SG1009 strain containing pQPDZ2-2 also did not grow. From Figure 5.24a, it can be seen that 6xHis-tagged PDZ2 (6xHis-PDZ2) expressed from pQPDZ2-2 was not detected by Western blot analysis and so this plasmid was discarded. Expression of 6xHis-PDZ2 from pRPDZ2-3, however, was detected in the *E. coli* M15 and *E. coli* SG1009 strains, though levels were higher in the former strain. Induction conditions were optimised in a similar manner for all other proteins discussed and were determined to be at 26°C for 16 hr using 0.1 mM IPTG. The protein was verified to be soluble and thus, subjected to a small-scale purification using Ni-NTA resin in a similar manner described for 6xHis-PDZ1. Figure 5.24b shows samples of the supernatants loaded onto the resin and the resin, bound to its ligand, 6xHis-PDZ2, which was separated on a Tricine gel and stained with Coomassie. Negative samples were carried out in conjunction using the empty pRSET-B plasmid transformed into *E. coli* M15 and induced and purified under the same conditions. Due to inaccuracies in the rainbow marker used, a second protein marker was loaded onto the gel. The gel shows a clear band corresponding to the calculated molecular mass of 6xHis-PDZ2 (16.6kDa). Further work will be carried out to clarify the SsrA:PDZ2 interaction using the soluble 6xHis-PDZ2.

5.6 Protease activity of the full length recombinant proteins

HtrA and its homologues all contain the catalytic triad seen in serine proteases. Thus, to investigate if the recombinant proteins had any protease activity zymograms were performed. Zymograms are essentially SDS-PAGE protein gels with a substrate added to the gel, in this case, β -casein. The proteins were electrophoresed under non-reducing conditions and are subsequently renatured in a mild detergent buffer. The

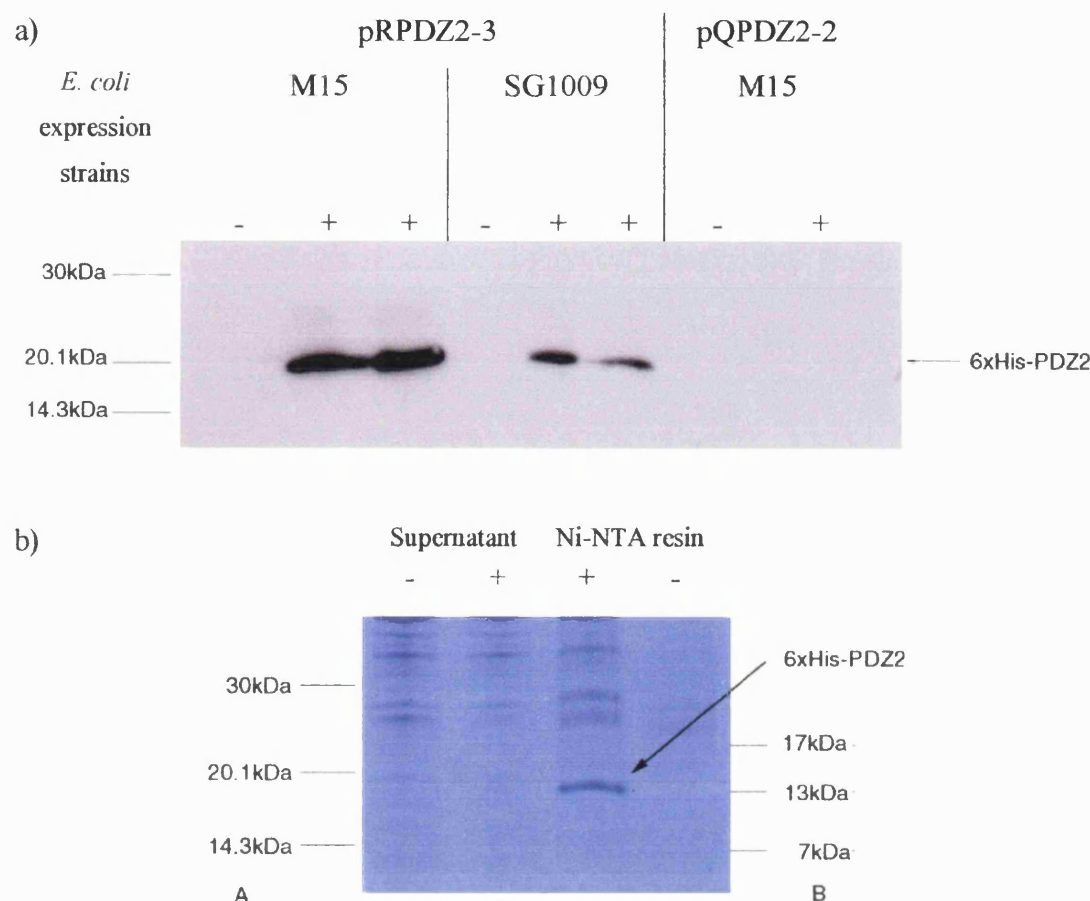


Figure 5.24 – Expression and purification of 6xHis-PDZ2

Two colonies of *E. coli* M15 and *E. coli* SG1009 strains, transformed with pRSET-B (-) or pRPDZ2-3 (+) or pQPDZ2-2 (+), were induced for 4 hr at 37°C with 1 mM IPTG. Culture (5 µl) was loaded per well onto a 12% Tricine gel and separated by electrophoresis. Protein was then transferred onto nitrocellulose for Western blot analysis using anti-His antibodies (a). Induction conditions were optimised and a 10 ml induction culture was pelleted. The supernatant was separated from cell debris and subjected to a small-scale purification using Ni-NTA resin. Supernatant (1 µl) and resin bound to protein (8 µl) was loaded per well (b). Ladder A is the rainbow marker (Amersham) and ladder B is MultiMark® multicoloured standard (Invitrogen). Samples are denoted '+', while the negative controls are denoted '-'.

gels are then incubated in a Tris buffer overnight to allow digestion of β -casein where proteases have separated on the gel. The gels are then stained with Coomassie, which stains the gel blue leaving clear bands where β -casein has been removed due to the presence of a protease. A negative control was run in conjunction to the samples used. This was derived from inductions with *E. coli* M15 expression strain containing the empty pRSET-B plasmid and subjected to the same purification procedures as positive samples. However, SAHtrA1 showed no protease activity under the conditions used (data not shown). This could be due to experimental conditions, as the proteins are denatured during electrophoresis by the high concentrations of SDS in the buffers. Although there is a renaturing step, it is possible that SAHtrA1 had not refolded properly and thus there is no detectable activity. Alternatively, the recombinant proteins may not be expressed with the correct tertiary structure.

5.7 Summary

Recombinant proteins were constructed the HtrA homologues in *S. aureus*. The construction of native SAHtrA1 and SAHtrA2 was not successful as the expression of the proteins could not be detected with confidence and so 6xHis-tagged versions were constructed. These proteins did not express well, even though a number of expression strains and conditions were used. SAHtrA1 did not appear to contain protease activity. However, this could have been a result of the experimental conditions, as the protein may not have refolded correctly, if it did have the correct tertiary structure to being with.

The expressed catalytic domain of SAHtrA1 proved to be unstable as it showed characteristics of autoproteolysis and so investigations in this domain were discarded.

The expression of the PDZ domains, PDZ1 and PDZ2, proved successful when cloned as a GST-tagged protein. Proteins were successfully purified by FPLC and contained secondary structure. SPR experiments showed that PDZ2 could interact with the *S. aureus* SsrA-tag and thus, is involved in protein:protein interactions. However, the GST-moiety may influence the accuracy of the SPR experiments, thus, the PDZ domains were cloned with a N-terminal 6xHis-tag. Expression of the 6xHis-tagged PDZ1 and PDZ2 were much reduced in comparison to the GST-tagged proteins, but a small-scale purification using Ni-NTA resin showed that the proteins could be partially purified and subsequently, large-scale purification was carried out.

CHAPTER 6. FUNCTIONAL ANALYSIS OF RECOMBINANT *S. typhimurium* HTRA ORTHOLOGUES

6.1	Introduction.....	186
6.2	Expression of recombinant STHhoA	186
6.2.1	Cloning of STHhoA	186
6.2.2	Expression and purification of STHhoA	190
6.3	Cloning and expression of STHhoB.....	194
6.3.1	Cloning of STHhoB.....	194
6.3.2	Expression and purification of STHhoB	194
6.4	Proteolytic activity of the <i>S. typhimurium</i> HtrA orthologues	196
6.5	Summary.....	200

6.1 Introduction

The expression of the *S. aureus* HtrA orthologues resulted in insufficient quantities of protein for further biochemical analysis (Chapter 5). Thus, the *S. typhimurium* HtrA orthologues were expressed to investigate this family of HtrA homologues. Recombinant *S. typhimurium* HtrA has been previously expressed and shown to be proteolytically active (Spiers, A. *et al.*, 2002). Thus, this chapter investigates the proteolytic activity of the STHhoA and STHhoB.

6.2 Expression of recombinant STHhoA

6.2.1 Cloning of STHhoA

To study the function of STHhoA, the gene was cloned as a native and a 6xHis-tagged protein. The coding sequence (1,368-bp) was cloned into, pQE-30, pRSET-B and pTrc99A. The ORF was amplified by PCR using 4 mM MgCl₂, pre-determined to be the optimal concentration in all cases. PCR reactions for cloning into the fusion plasmids were performed with the primers EMR19 and EMR20 for pQE-30, and EMR21 and EMR20 for pRSET-B. Primers EMR19 and EMR21 incorporated the 5' *Bam*HI restriction site while EMR20 incorporated the 3' *Hind*III site. PCR reactions for cloning into pTrc99A were carried out with the primers EMR22 and EMR23, which incorporated the 5' *Nco*I site and the 3' *Bam*HI site, respectively. The cloning strategies were carried out as described for previous recombinant proteins and illustrated in Figure 6.1 for pQE-30, Figure 6.2 for pRSET-B and Figure 6.3 for pTrc99A. *E. coli* DH5 α cells were transformed with the ligation mix and the resultant plasmids, pQHhoA, pRHhoA and pTHhoA, were amplified by culturing single colonies the plasmids of interest isolated for sequencing. Plasmids containing the correct sequence were pQHhoA-4, pRHhoA-1 and pTHhoA-1.

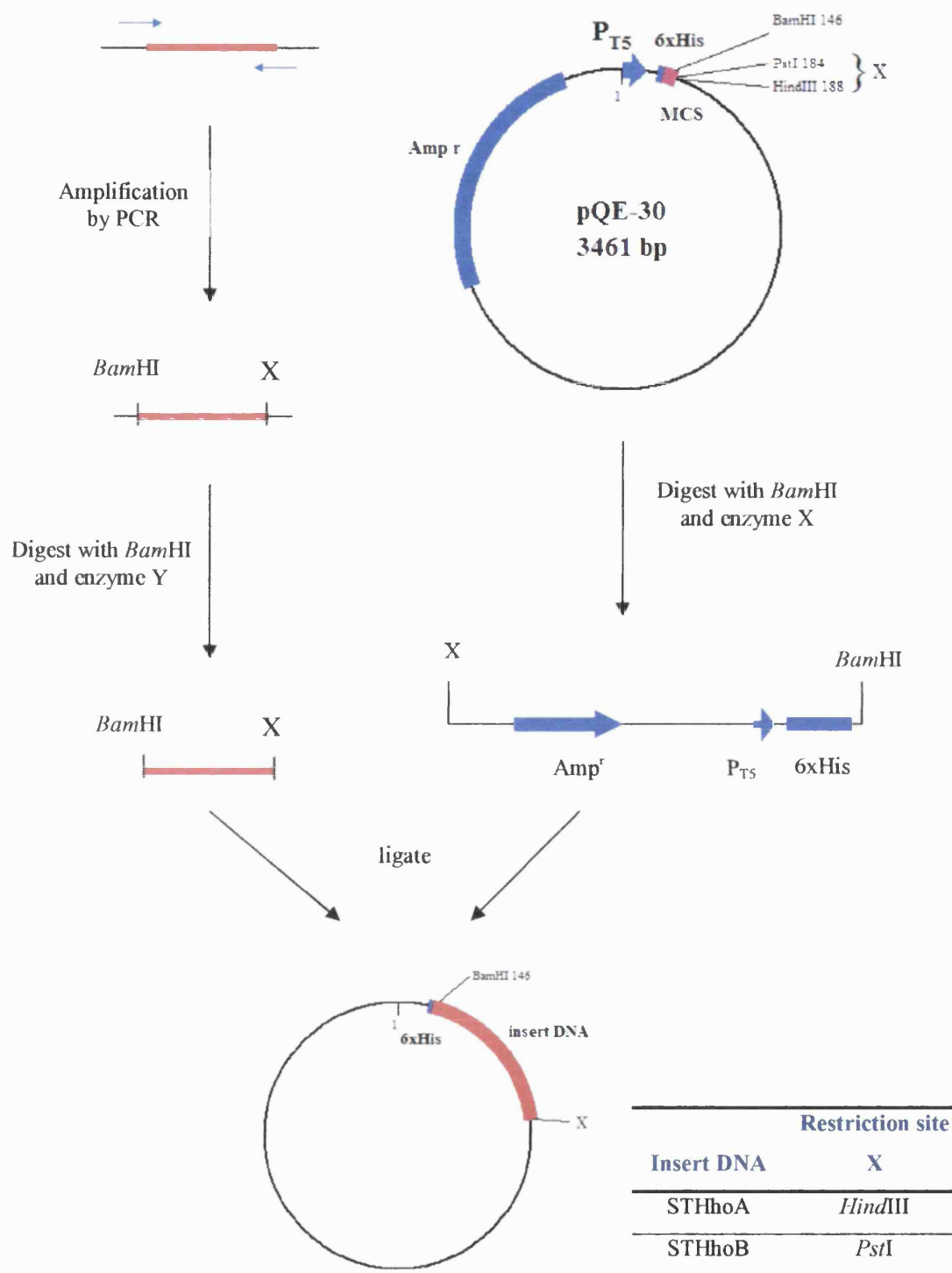


Figure 6.1 – Cloning strategy of 6xHis-tagged recombinant proteins using pQE-30

All insert DNA cloned into pQE-30 was engineered with a *Bam*HI site at the 5' end and site X (see table) at the 3' end of the insert.

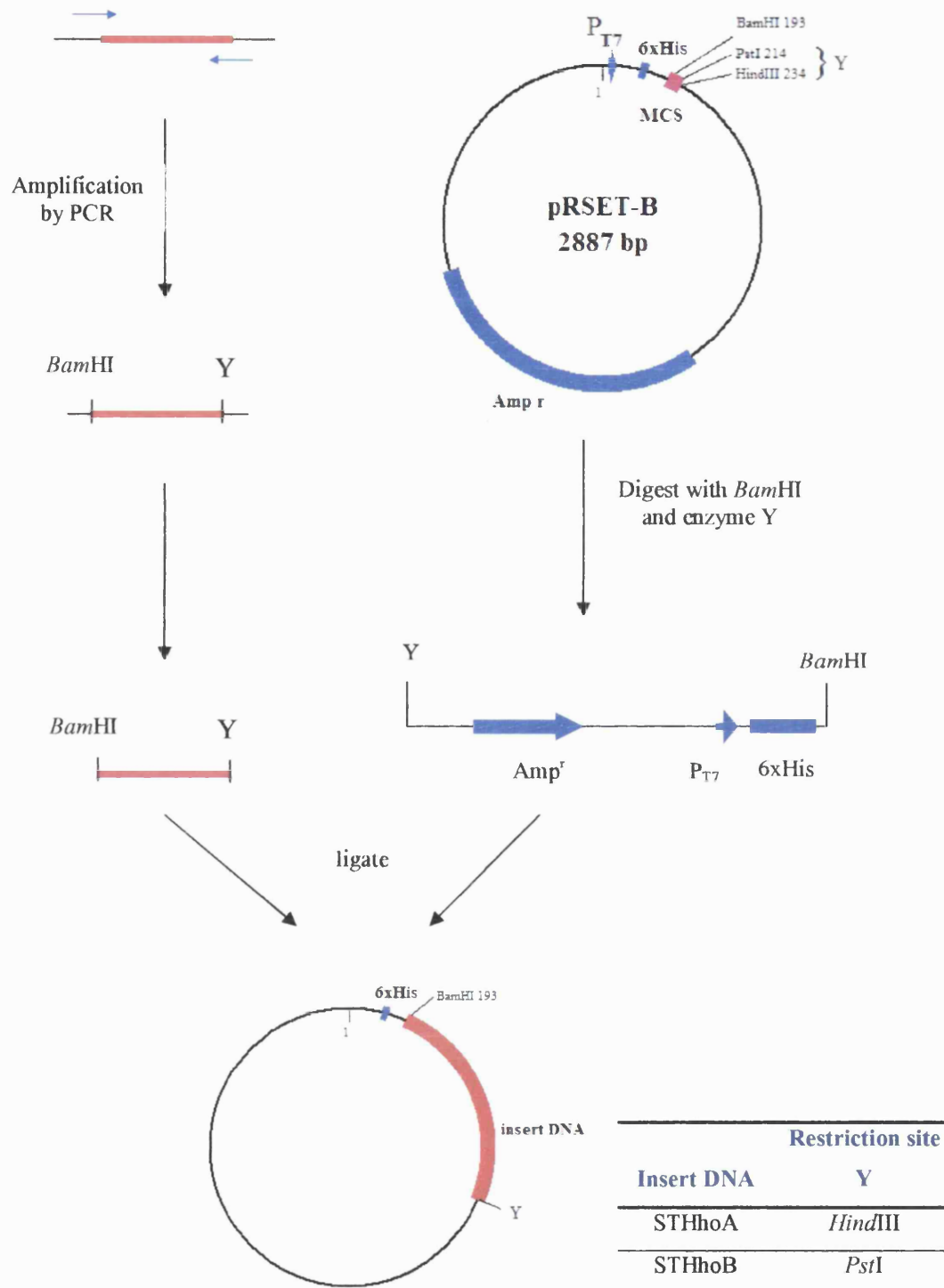


Figure 6.2 – Cloning strategy of 6xHis-tagged recombinant proteins using pRSET-B

All insert DNA cloned into pRSET-B was engineered with a *Bam*HI site at the 5' end and site Y (see table) at the 3' end of the insert

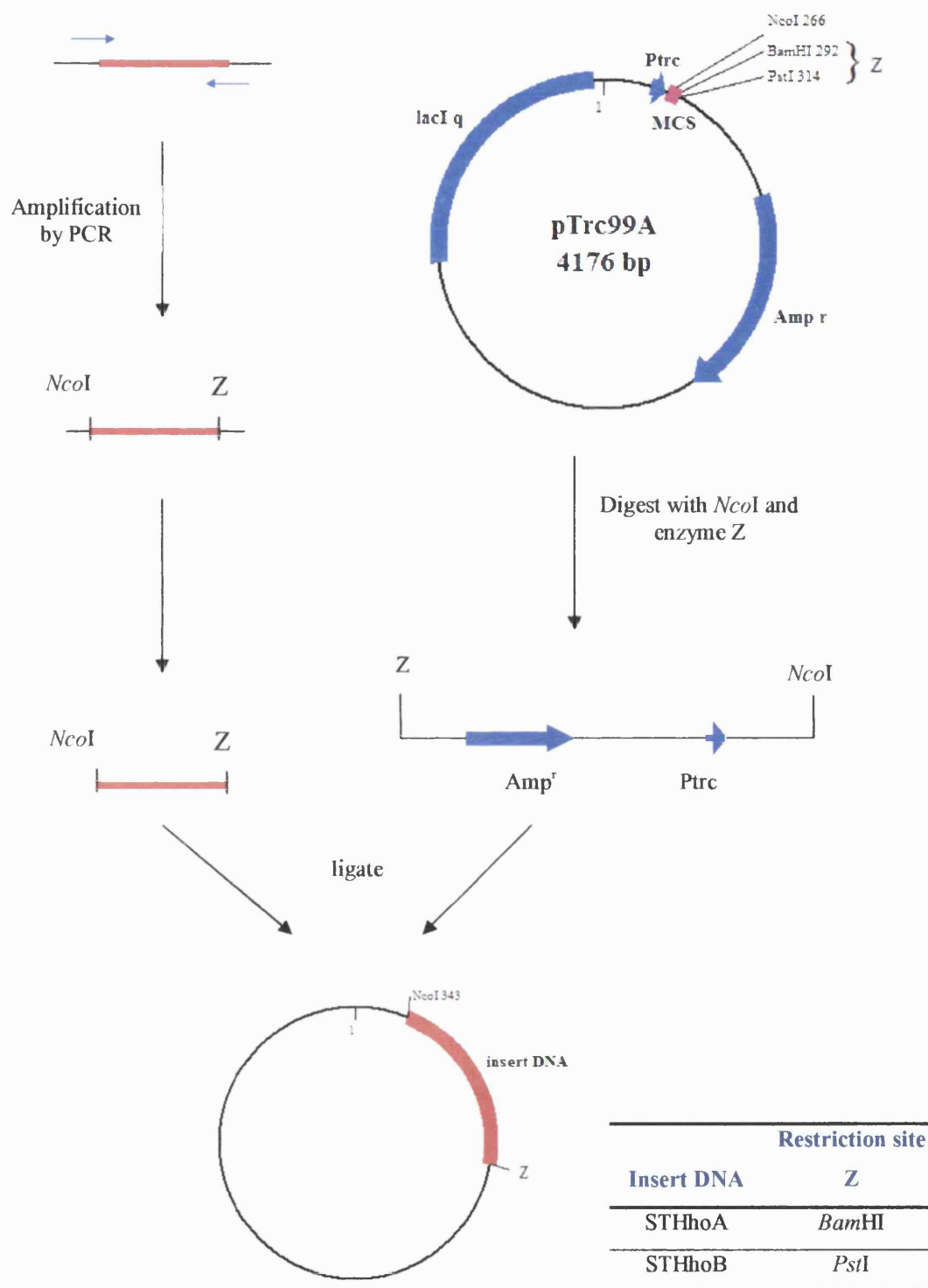


Figure 6.3 – Cloning strategy for native recombinant proteins using pTrc99A

All insert DNA cloned into pTrc99A were engineered with a *NcoI* site at the 5' end and site Z (see table) at the 3' end of the insert.

6.2.2 Expression and purification of STHhoA

The expression strategy is shown in Figure 6.4. The fusion plasmids were expressed in several expression strains by induction at 1 mM IPTG for 4 hr at 37°C. Expression was compared to that of the empty plasmid, pQE-30 and pRSET-B, which were used as negative controls. Expression of the 6xHis-tagged STHhoA (6xHis-STHhoA; calculated molecular mass, 48.1 kDa in pQE-30, and 50.8 kDa in pRSET-B) was not detected in any of the *E. coli* expression strains used on Western blot analysis. However, from the protein gel (Figure 6.5a), expression of a smaller protein (~ 47 kDa) was detected in all samples for expression from both pQHhoA-4 and pRHhoA-1. *E. coli* HhoA is known to undergo signal peptide cleavage at S²⁸ (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996) and this region is virtually identical in STHhoA (Chapter 3, Figure 3.1b), thus it is suggested that signal peptide cleavage has occurred. If this were to occur at S²⁸ in STHhoA, the cleaved STHhoA (denoted STHhoA*) would be 47.3 kDa. This is in agreement with the data from the gel, as although the 6xHis-tags produced from pQE-30 and pRSET-B differ by 2.7 kDa, removal of this along with the signal peptide would result in a protein of 47.3 kDa in all cases, which would not be detected by Western blot analysis.

All the *E. coli* expression strains used grew and expressed STHhoA* when transformed with pQHhoA-4. However, levels of STHhoA* were lower than that observed when the expression strains were transformed with pRHhoA-1. Of the strains used, *E. coli* M15 and *E. coli* SG1009 transformed with pRHhoA-1 did not grow and so were discarded. The *E. coli* BL21(DE3)Codon Plus expression strain was transformed with pTHhoA-1 and expression studies were carried out to optimise

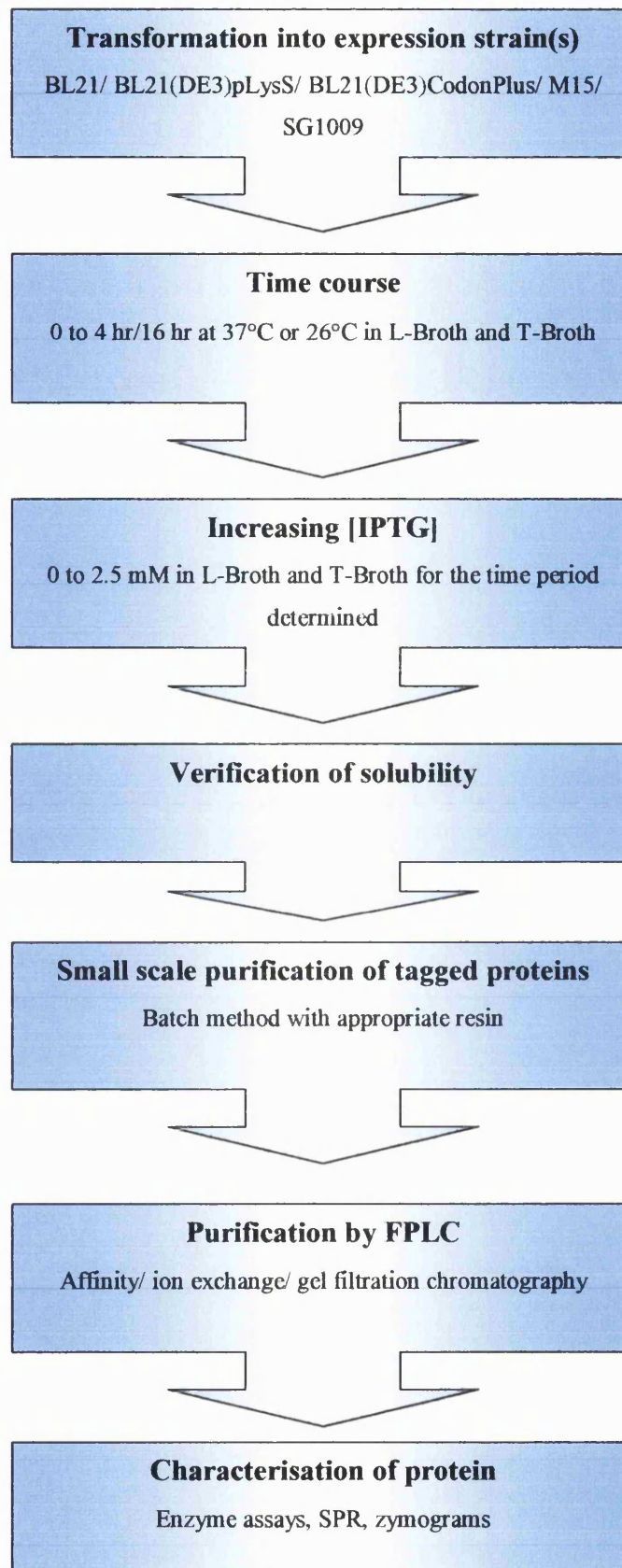


Figure 6.4 – Expression strategy to optimise protein expression

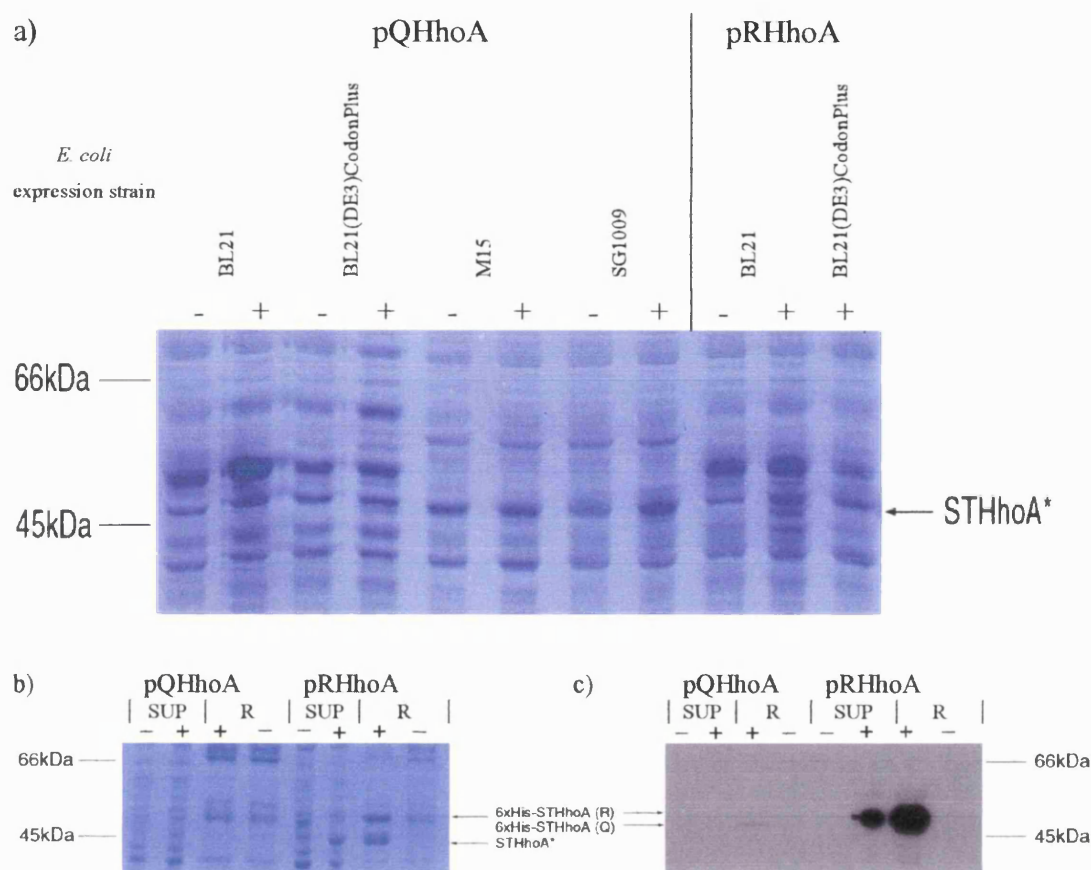


Figure 6.5 – Expression of soluble 6xHis-tagged STHhoA from pQHhoA-4 and pRHhoA-1

E. coli expression strains were transformed with either pQHhoA-4 or pRHhoA-1 and induced with 1 mM IPTG over 4 hr. Culture (5 μ l) was analysed by SDS-PAGE (a). Strains M15 and BL21 were utilised to express pQHhoA-1 and pRHhoA-1, respectively. The optimal induction conditions were determined and the protein was verified to be soluble. Supernatant from a 10 ml induction was subjected to a small-scale purification (b). The supernatant (SUP, 0.5 μ l) and resin bound to 6xHis-STHhoA (R, 2 μ l) was analysed by SDS-PAGE and Western blot analysis. Negative controls (-) were carried out in parallel using either pQE-30 or pRSET-B transformed into *E. coli* M15 and *E. coli* BL21 strains, respectively.

the optimal induction conditions along with *E. coli* BL21 containing pRHhoA-1 and the *E. coli* BL21(DE3)CodonPlus strain containing pQHhoA-4.

Optimal expression was achieved in T-Broth for all samples at 37°C for 4 hr. The optimal concentration of IPTG used differed for the three proteins, 0.01 mM (pQHhoA-4), 0.1 mM (pRHhoA-1) and 0.01 mM (pTHhoA-1). The recombinant STHhoA* was verified to be soluble and the 6xHis-tagged versions were subjected to purification by Ni-NTA resin in order to purify any uncleaved products. Negative controls were carried out using the same conditions except that the plasmid transformed into the expression strains was the empty plasmid, pQE-30 or pRSET-B. Figure 6.5b shows the Coomassie stained SDS-PAGE of the supernatant and bound resin of samples ran against the negative controls. Recombinant protein expressed from pQHhoA-4 was not observed on the gel and from the Western blot analysis (Figure 6.5c), very little full length 6xHis-STHhoA is detected. Recombinant protein expressed from pRHhoA-1, however, was detected on the protein gel as there were two proteins expressed that are not seen in the negative controls. The heavier protein band corresponded to the full length 6xHis-STHhoA was detected on the protein gel and corresponded to a signal detected by Western blot analysis. The smaller protein band corresponded to the calculated molecular mass of STHhoA* and was not detected by Western blot analysis indicating that the N-terminal 6xHis-tag had been removed. Interestingly, STHhoA* could be isolated by the Ni-NTA resin, even though there was no 6xHis-tag present. It is unclear how the protein interacts with the resin, though it could be due to another characteristic of the resin or that STHhoA* can bind to Ni^{2+} .

6.3 Cloning and expression of STHhoB

6.3.1 Cloning of STHhoB

In order to characterise STHhoB, a similar approach to that for STHhoA was performed. The ORF (1,071-bp) was cloned into pQE-30 and pRSET-B using primers EMR24 (pQE-30) or EMR25 (pRSET-B), which incorporated a 5' *Bam*HI site, and EMR26, which incorporated a 3' *Pst*I site. The cloning strategy shown in Figure 6.1 was used for pQE-30 and Figure 6.2 shows the strategy for pRSET-B. In addition, the ORF was cloned into pTrc99A (Figure 6.3) using EMR27, which incorporate a 5' *Nco*I site, and EMR26. For all PCR reactions, an optimal MgCl₂ concentration of 4 mM was determined as for previous experiments. Colonies obtained from transforming the ligation reactions into DH5α cells were grown and plasmids isolated. The DNA of the resulting pQHhoB, pRHhoB and pTHhoB plasmids were then sequenced. Plasmids pQHhoB-6, pRHhoB-3 and pTHhoB-1 were determined to have the correct DNA sequence.

6.3.2 Expression and purification of STHhoB

The pQHhoB-6 and pRHhoB-3 plasmids were transformed into several expression strains. The *E. coli* M15 expression strain containing pRHhoB-3 did not grow when cultured and so was discarded. All the expression strains, which did grow, expressed the 6xHis-STHhoB, detectable by Western blot analysis (Figure 6.6a). There was no indication of any occurrence of signal peptide cleavage on examination of the protein gel (data not shown). This is in agreement with previous observations (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). The expression strain *E. coli* M15 was used for further optimisation of the expression from the pQHhoB-6

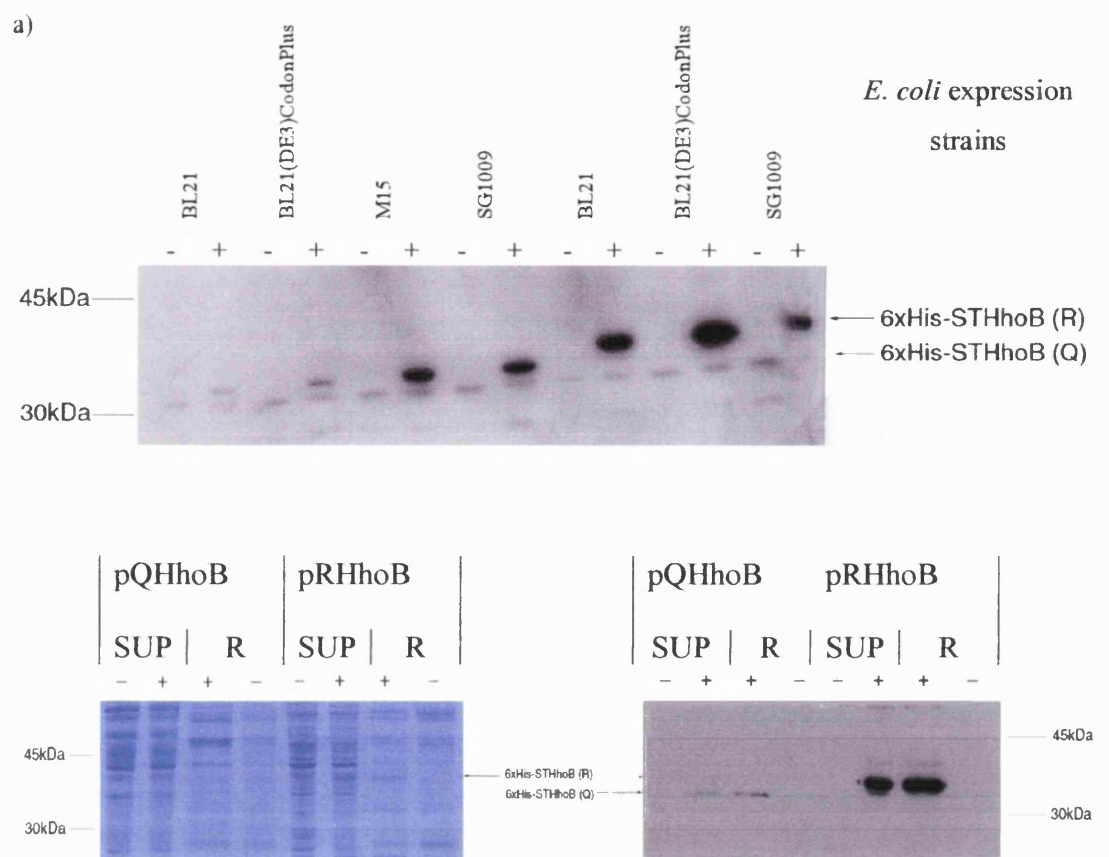


Figure 6.6 – Expression of soluble 6xHis-STHhoB from pQHhoB-6 and pRHhoB-3

E. coli expression strains were transformed with either pQHhoB-6 or pRHhoB-3 and induced with 1 mM IPTG over 4 hr. Culture (5 μ l) was analysed by SDS-PAGE (a). *E. coli* strains M15 and BL21(DE3)CodonPlus were utilised to express pQHhoB-6 and pRHhoB-3, respectively. The optimal induction conditions were determined and the protein was verified to be soluble. Supernatant from a 10 ml induction was subjected to a small-scale purification (b). The supernatant (SUP, 0.5 μ l) and resin bound to 6xHis-STHhoA (R, 2 μ l) was analysed by SDS-PAGE and Western blot analysis. Negative controls (-) were carried out in parallel using either pQE-30 or pRSET-B transformed into *E. coli* M15 and *E. coli* BL21(DE3)CodonPlus strains, respectively.

plasmid, whilst the expression strain, *E. coli* BL21(DE3)CodonPlus, was used for expression from the pRHhoB-3 plasmid. The native pTHhoB-1 was transformed into *E. coli* BL21(DE3)CodonPlus strain for expression studies.

Optimal expression, once again, was maximal in T-Broth for all the expression studies. Maximal expression occurred at 37°C for 4 hr inducing with 0.01 mM IPTG in all cases. The proteins were then verified to be soluble and then subjected to a small-scale purification step with Ni-NTA. Figure 6.6b shows the resin can purify 6xHis-STHhoB produced from the different plasmids, although there is a much higher yield of protein from pRHhoB-3 with respect to pQHhoB-6. Due to the small yield of purified protein from pQHhoB-6, this plasmid was discarded.

6.4 Proteolytic activity of the *S. typhimurium* HtrA orthologues

As discussed earlier, HtrA and its homologues all contain the catalytic triad seen in serine proteases. Thus, as with SAHtrA1, recombinant proteins of the *S. typhimurium* HtrA orthologues were examined for proteolytic activity by performing zymograms (discussed earlier in 5.6). In all cases, negative controls were run in conjunction to the samples. These were derived from inductions with the same *E. coli* expression strain containing the empty plasmid, pRSET-B or pTrc99A. For the 6xHis-tagged proteins, the negative samples were subjected to the same purification procedures as the positive samples.

Figure 6.7 shows the zymograms of STHhoA (R), derived from expression from pRHhoA (Figure 6.7a), STHhoB (T), derived from expression from pTHhoB (Figure 6.7b) and STHhoB (R), derived from expression from pRHhoB. Protein expression

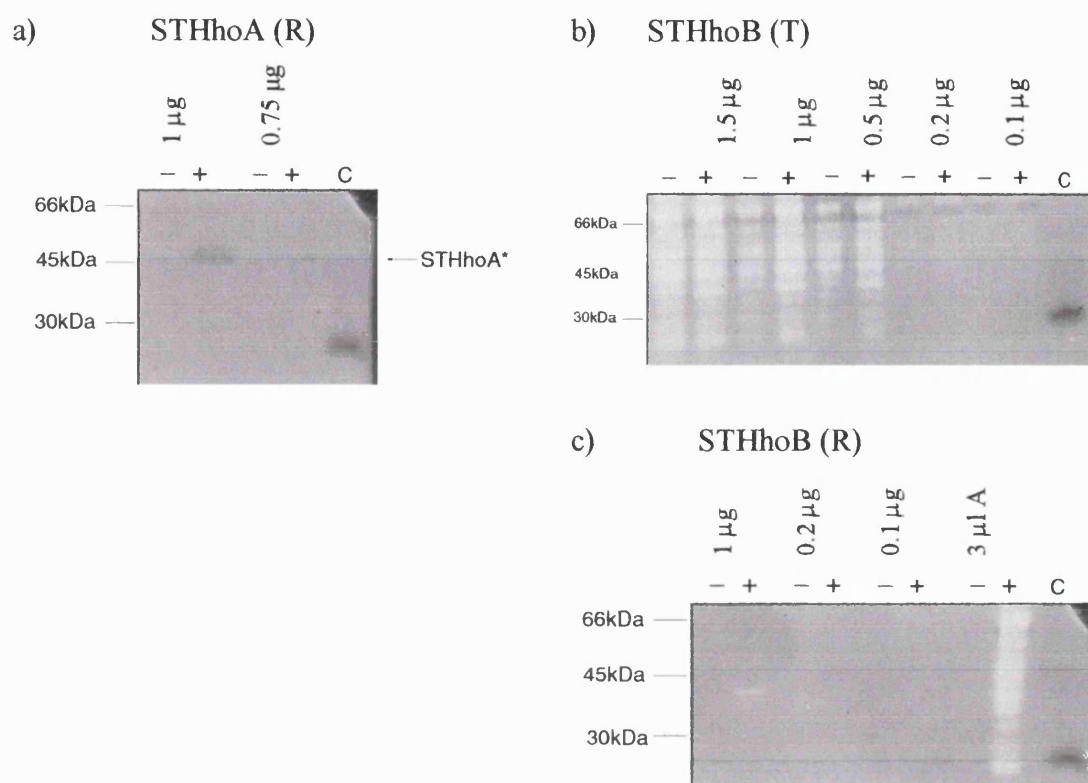


Figure 6.7 – Protease activity of HtrA homologues

Varying amounts of protein (+) (partially purified for 6xHis-tagged proteins or crude supernatant, in the case of native proteins and samples A and B) were loaded onto a zymogram. Proteins were separated by electrophoresis and proteins allowed to renature before incubation for 16 hr at 37°C. The gels were stained with Coomassie and the presence of protease is indicated by the removal of β -casein (clear bands) on the gel after destaining. Sample A (c) was the supernatant from the induction study of pRHhoA, which had not been partially purified. Sample C on all gels is the positive control chymotrypsin (2.5 μ g). Negative controls (-) were performed in parallel using the respective empty plasmid, pRSET-B or pTrc99A, in the appropriate *E. coli* expression strain. Gel images were manipulated in Photoshop in order to visualise the protease bands (dark grey) clearly (pictures were converted to greyscale and then inverted).

of the STHhoA (R) was performed in *E. coli* BL21(DE3)CodonPlus cells. This is because although the BL21 strain produced more recombinant protein, the protein was inactive under the conditions used. There was no protease activity seen for either of the STHhoB recombinant proteins. This could be due to experimental conditions, as the proteins are denatured during electrophoresis by the high concentration of SDS in the buffers. Although there is a renaturing step, it is possible that STHhoB did not refold properly and thus there was no detectable activity. Alternatively, the recombinant protein may not have been expressed with the correct tertiary structure.

As STHhoA was active, the effects of a variety of protease inhibitors were investigated. Figure 6.8a shows a zymogram of STHhoA derived from pTHhoA and pRHhoA. The clear band, not detected in the negative controls, is more pronounced as the amount of protein increases. In addition, this band is approximately the same size in both samples indicating that it is STHhoA*, as the protein expressed from pRHhoA would be 3.5 kDa heavier than the native protein. The cysteine protease inhibitor E64 (Figure 6.8c), the aspartic acid protease inhibitor pepstatin A (Figure 6.8e) and the metalloprotease inhibitor EDTA (Figure 6.8f) do not inhibit the protease activity. In addition, the serine protease inhibitors, aprotonin (Figure 6.8b) and AEBSF (Figure 6.8d) do not abolish protease activity of STHhoA* or chymotrypsin, although there does seem to be a reduction in the degradation of β -casein. Efforts to optimise the conditions, such as the removal of the detergent Brij 35, and increasing the concentrations of the inhibitors did not abolish the activity of STHhoA* or chymotrypsin. This could be also due to incorrect refolding of the protein after electrophoresis. Alternatively, there may be other factors involved in the catalytic mechanism of STHhoA*. However, further work must be done to

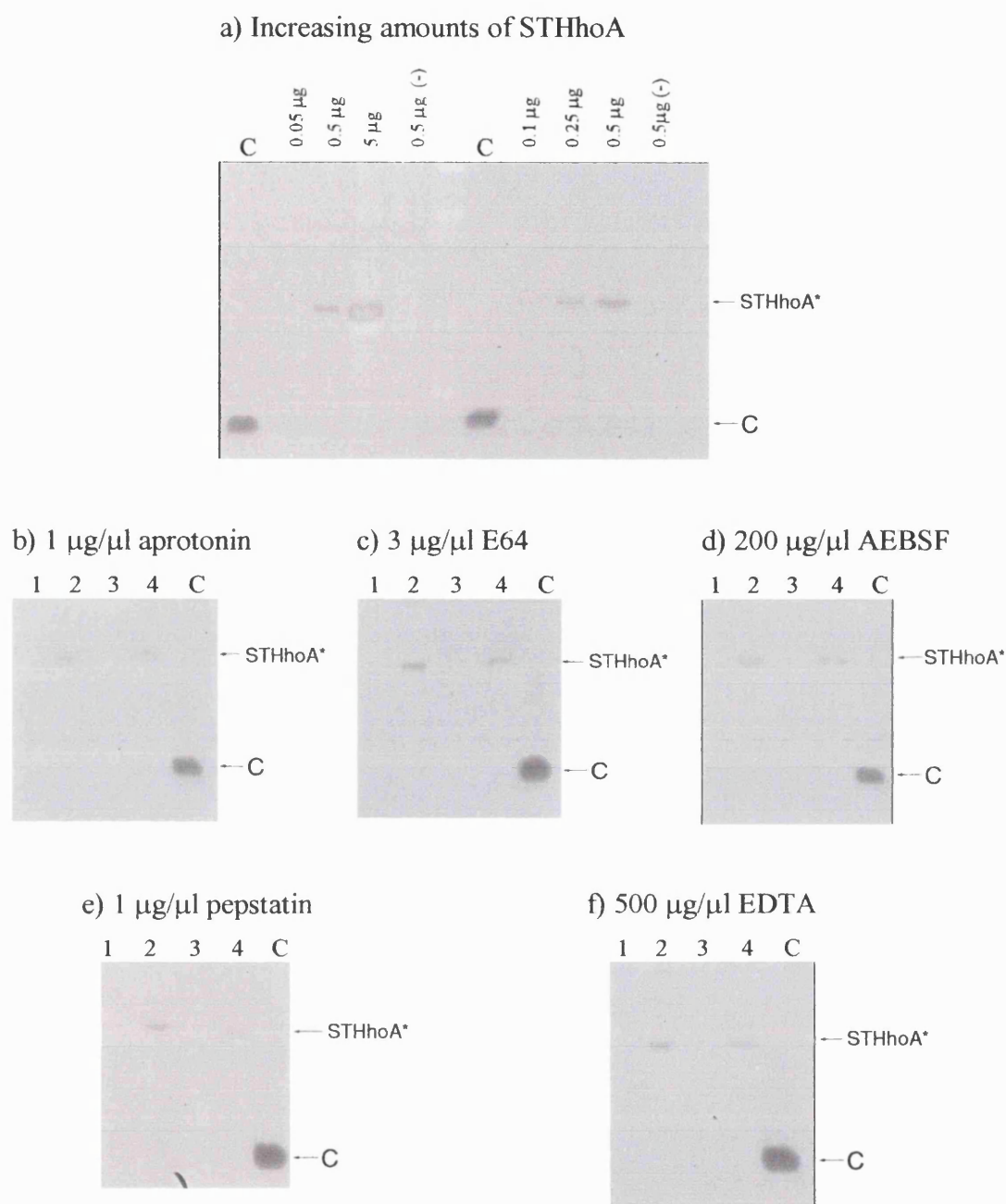


Figure 6.8 – Effect of protease inhibitors on STHhoA* protease activity

Zymograms were performed using protein partially purified from induction studies with pTHhoA-1 and pRHhoA-1 (a). Gels were incubated in the presence of either 1 µg/µl aprotinin (b), 3 µg/µl E64(c), 200 µg/µl AEBSF (d), 1 µg/µl pepstatin (e), 500 µg/µl EDTA (f) or with no inhibitors (a). Gel images were manipulated in Photoshop in order to visualise the protease bands (dark grey) clearly (pictures were converted to greyscale and then inverted).

elucidate any differences. Purification of the proteins would undoubtedly be invaluable for future studies.

6.5 Summary

Recombinant proteins were constructed of the HtrA orthologues in *S. typhimurium*. *S. typhimurium* HtrA had been cloned and expressed previously in the lab and was shown to be proteolytically active (Spiers, A. *et al.*, 2002). Thus, STHhoA and STHhoB were cloned and expressed as native and 6xHis-tagged protein in order to further characterise these HtrA orthologues in *S. typhimurium*. Expression using the pQE-30 plasmid produced a much poorer yield than that seen with the pRSET-B versions. STHhoA appeared to undergo signal cleavage, as the N-terminal 6xHis-tag was not detected by Western blot analysis. The processed STHhoA* could be partially purified by Ni-NTA resin with the uncleaved STHhoA protein. STHhoB, however, did not undergo signal peptide cleavage, in agreement with observations discussed in 3.2. This protein could also be partially purified by Ni-NTA resin. Large-scale purification of these proteins are being carried out at the moment.

Partially purified STHhoA and STHhoB was examined for protease activity. In addition, as the native proteins expressed well, they were also tested and compared to their respective negative controls. STHhoB did not exhibit any protease activity on the zymograms. This could be a result of experimental conditions, as the protein may not have refolded correctly, if indeed it had the correct tertiary structure to begin with.

These experiments showed that the *S. typhimurium* protein, STHhoA*, was proteolytically active. This is in agreement with work published on the *E. coli* HhoA (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Cysteine protease inhibitors, aspartic acid protease inhibitors or metalloprotease inhibitors do not affect this protease activity. However, serine protease inhibitors were effective but could not completely abolish this activity, resulting in a reduced amount of β -casein degradation. Further optimisation of the experimental conditions may be required to characterise this enzyme activity.

CHAPTER 7. VIABILITY OF MUTANTS OF THE *htrA* FAMILY

7.1	Introduction.....	203
7.2	Constructing <i>S. typhimurium</i> mutants	204
7.2.1	Construction of single mutants	204
7.2.1.1	Preparation of a kanamycin resistance cassette for the construction of linear allelic replacement fragments	204
7.2.1.2	Allelic replacement of <i>htrA</i> , <i>hhoA</i> , <i>hhoB</i> or <i>hhoAB</i>	206
7.2.2	Construction of double and triple mutants.....	206
7.2.2.1	Production of <i>htrA::GM</i> and <i>htrA::CAT</i> single mutants	208
7.2.2.2	Production of double and triple mutants.....	210
7.3	Viability of mutants under <i>in vitro</i> stress conditions	210
7.3.1	Survival of mutants under heat shock conditions	212
7.3.2	Survival of mutants under oxidative stress conditions	220
7.4	Viability of mutants in an <i>in vivo</i> system	230
7.4.1	Transduction of <i>S. typhimurium</i> SL3261 mutants into <i>S.</i> <i>typhimurium</i> SL1344	230
7.4.2	Bacterial virulence in BALB/c mice	230
7.5	Summary.....	236

7.1 Introduction

In order to understand the biological role of *htrA* and its homologues, knockout mutants were constructed in *S. typhimurium*. Mutants of *htrA* in *E. coli* (Strauch, K.L. & Beckwith, J., 1988, Lipinska, B. *et al.*, 1989, Lipinska, B. *et al.*, 1990, Seol, J.H. *et al.*, 1991) and other species, such as *K. pneumoniae* (Cortés, G. *et al.*, 2002) and *Y. pestis* (Williams, K. *et al.*, 2000) have been constructed and have shown increased sensitivity to heat-shock, indicating the importance *htrA* in thermoresistance. *S. typhimurium htrA* mutants demonstrated an inferior ability to survive oxidative stress conditions such as H₂O₂. *In vivo* studies of *htrA* mutants illustrated the importance of this gene in bacterial virulence. *S. typhimurium* and *Y. enterocolitica* mutants were highly attenuated in mice (reviewed in Pallen, M.J. & Ponting, C.P., 1997).

The genetic systems available for the construction of targeted mutants in *S. aureus* are less well characterised than those in *S. typhimurium*, therefore, mutants of the *htrA* gene family were made in *S. typhimurium* and analysed *in vitro* and *in vivo*.

For the purpose of examining the function of the *htrA* gene family, mutants of *htrA* and its paralogues, *hhoA* and *hhoB*, were constructed in *S. typhimurium* SL3261, an attenuated *aroA* mutant of the wild-type strain, SL1344. *In vitro* studies examining the ability of bacterial growth and viability in stress conditions were investigated using these mutants. In order to investigate bacterial survival *in vivo*, the mutations constructed in SL3251 were transduced into the mouse virulent strain SL1344.

To date, no comprehensive *in vivo* study has been published which examines the properties of mutants of the *htrA* paralogues.

7.2 Constructing *S. typhimurium* mutants

7.2.1 Construction of single mutants

7.2.1.1 Preparation of a kanamycin resistance cassette for the construction of linear allelic replacement fragments

For the generation of single mutants, the target gene is replaced with an antibiotic resistance marker, in order to certify that the gene has been replaced. Single mutants for *htrA*, *hhoA* and *hhoB* were created using a novel technique, which exploits ET-cloning (Muyrers, J.P.P. *et al.*, 1999, Zhang, Y. *et al.*, 2000) and recombination via the red genes, *exo*, *bet* and *gam* (Yu, D. *et al.*, 2000). In addition, the double mutant *hhoAB*, was constructed by exploiting the fact that these two gene are adjacent to each other on the chromosome and so can be deleted together in a single allelic replacement experiment.

The resistance marker used for the construction of the single mutants was the kanamycin resistance gene (*kan*), which was excised from pUC18:KAN using *Xba*I sites either side of the gene (Figure 7.1a). Gel extraction (Figure 7.1b) was performed using the QIAquick gel extraction kit to remove the *kan* cassette from the plasmid DNA. The *kan* gene was then used as the template DNA in PCR (Figure 7.1c) with 60-mer primers (appendix A.7) which had been designed to anneal with 20-bp at each end of *kan* with the remaining forty bases complementary to the DNA sequence flanking the gene of interest. PCR was carried out with a proofreading

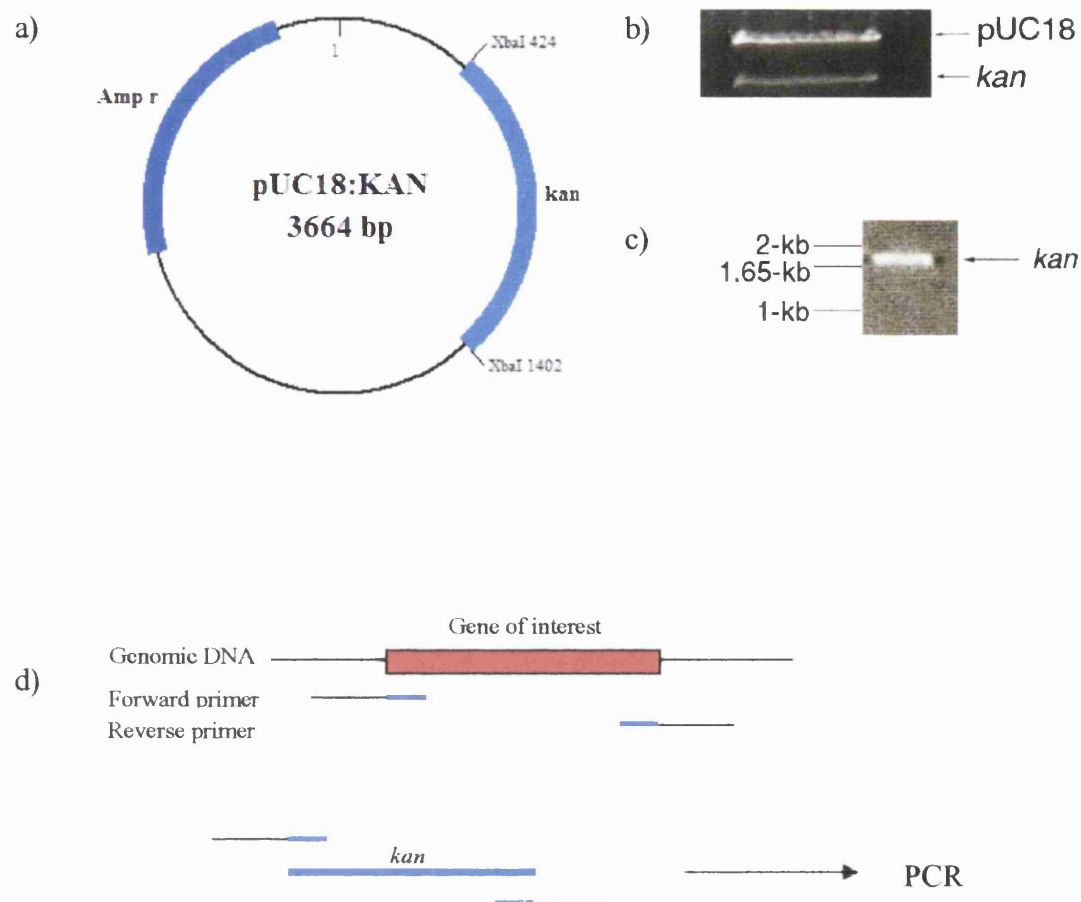


Figure 7.1 – Preparation of kanamycin cassette as the insert DNA for allelic exchange

Schematic diagram of pUC18:KAN (a), the source of the template DNA, *kan*. The resistance marker *kan* was excised with 10U *Xba*I (b) and isolated from the plasmid DNA by gel extraction (c). This DNA was then used for PCR (d) with 60-mer primers designed to anneal to 20-bp (blue) at either end of the *kan* gene and the remaining 40-mers complementary to 40-bp of the DNA flanking the gene of interest (black).

DNA polymerase and MgCl₂ concentrations were optimised for maximal yield of the PCR product as described in 4.2.1. In all cases the optimum MgCl₂ concentration was 2 mM and this was used to set up 2 x 50 µl PCR reactions. PCR products were purified by gel extraction using the QIAquick gel extraction kit.

7.2.1.2 Allelic replacement of *htrA*, *hhoA*, *hhoB* or *hhoAB*

The plasmid pBADλRED was electroporated into *S. typhimurium* SL3261. This plasmid contains the λ red genes under the control of the P_{araB} promoter. The *gam* gene encodes a protein, which inhibits the host RecBCD exonuclease V while *exo* and *bet* encode proteins which bind DNA-ends to promote recombination. After a short period of induction with arabinose to allow to the production of the λ red genes, the bacteria culture was prepared for electroporation as described in 2.2.1.1. The allelic replacement fragments were electroporated separately into the electrocompetent bacteria. Single *kan*^r colonies were verified by colony PCR (Figure 7.2), using internal *kan* primers in combination with primers which anneal to the DNA flanking the replaced gene. Positive PCR products were TA-cloned and the plasmid DNA from the resulting transformants were sequenced to confirm the insertion of the resistance marker into the target gene.

7.2.2 Construction of double and triple mutants

The *hhoA* and *hhoB* genes are adjacent to each other on the chromosome, and this fact was exploited in constructing double and triple mutants. Thus, the construction of the remaining double and triple mutants was created from two rounds of mutations using a second antibiotic resistance gene. The two markers used were the

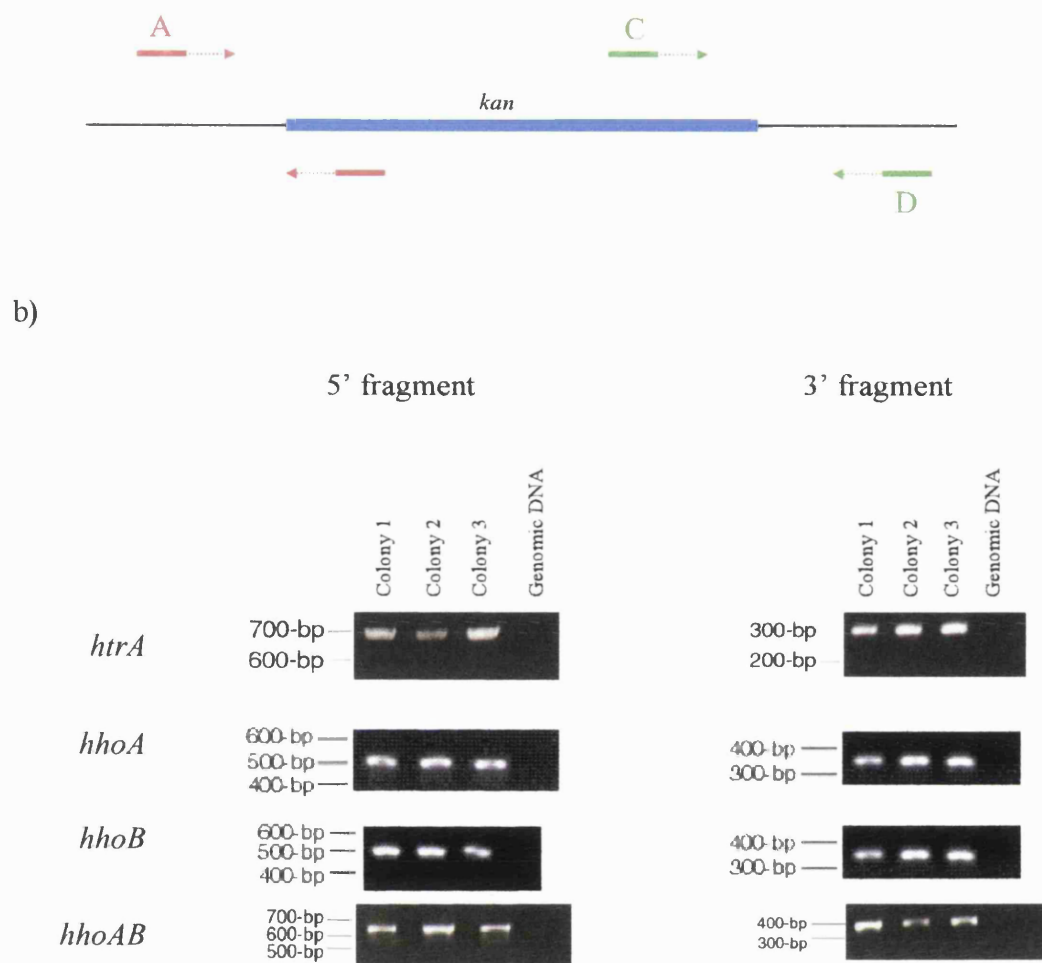


Figure 7.2 – Verification of mutants after one round of allelic replacement

Mutants were verified using internal *kan* primers (B and C) and primers which anneal to DNA flanking the replaced gene (A and D). The PCR products (b) from colony PCRs from three single colonies compared with the negative control (genomic *S. typhimurium* DNA as the template DNA) in the last lane.

gentamycin resistance gene (*gen*) and the chloramphenicol resistance gene (*cat*). Replacing genes with *kan* proved to be more efficient (generally one hundred to two hundred colonies per transformation) than with *gen* (ten to fifteen colonies per transformation) or *cat* (one colony per transformation). Thus, *htrA* was replaced with *gen* or *cat* in the first round of mutations.

7.2.2.1 Production of *htrA::GM* and *htrA::CAT* single mutants

Plasmids, pBSL182 and pACYC184, containing *gen* and *cat*, respectively, were used as the template DNA in PCR reactions with 60-mer primers designed to anneal to the 20-bp flanking the ends of the drug resistance determinant and the remaining forty bases complementary to the DNA flanking the *htrA* gene. Again, MgCl₂ conditions were optimised to obtain the maximum yield of products. Using the optimal concentration of MgCl₂ (2 mM for both genes), *gen* and *cat* were amplified by PCR and purified by gel extraction resulting in the linear DNA fragments for allelic replacement. Electroporation was carried out to modulate the linear fragments into SL3261. Single *gen*^r and *cat*^r colonies containing the deletion of interest were verified by colony PCR, using primers flanking the *htrA* gene, as the size difference between wild-type *htrA* (1.428-kb) and the resistance markers (*gen*: 0.816-kb, *cat*: 1.059-kb) was distinguishable on an agarose gel (Figure 7.3). Positive PCR products were TA-cloned and the plasmid DNA of the resulting transformants were sequenced to confirm the insertion of the resistance marker into the correct site. The resultant strains, *htrA::GM-1* and *htrA::CAT-1* were selected for the construction of double and triple mutants.

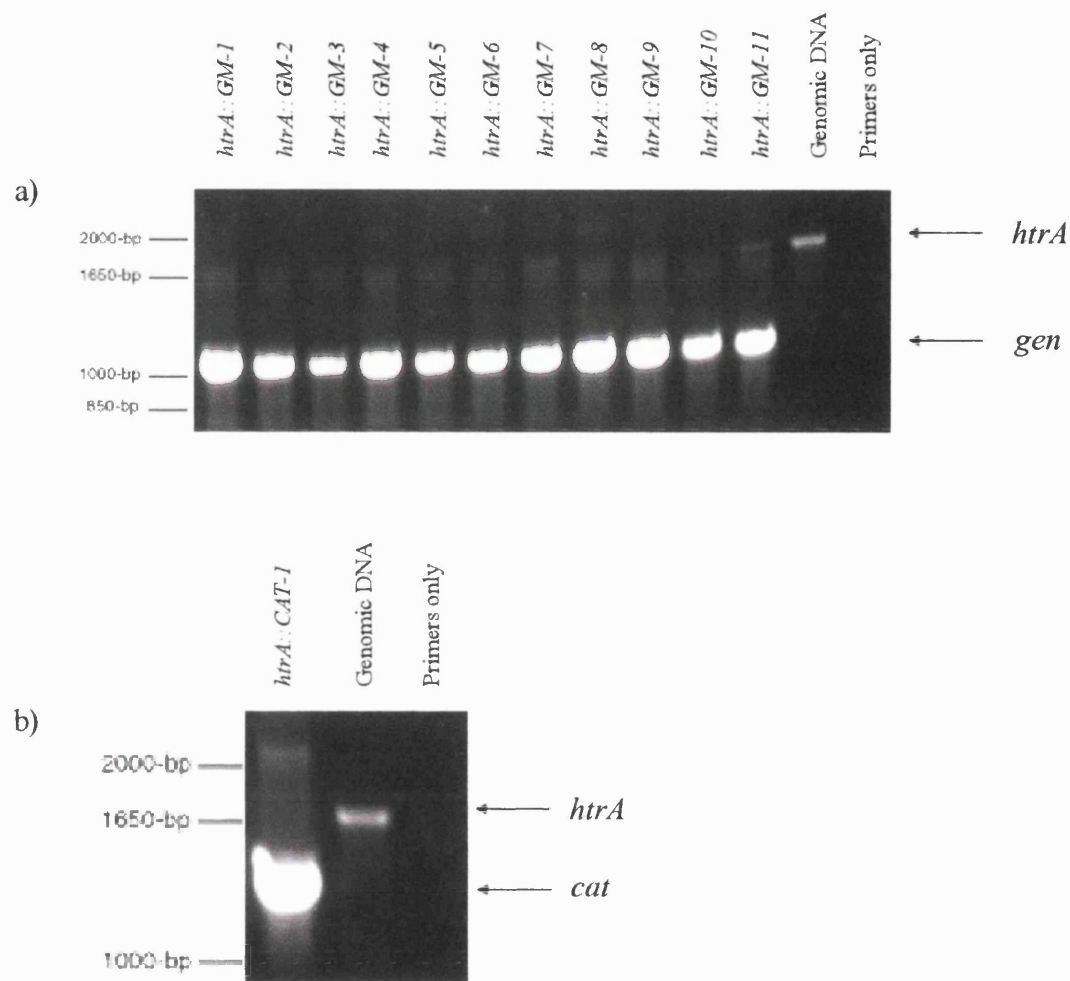


Figure 7.3 – Verification of *htrA::GM* and *htrA::CAT* mutants

Single mutant colonies of *htrA::GM* (a) and *htrA::CAT* (b) were checked by colony PCR using primers that flank the *htrA* gene. The PCR products from three single colonies were separated on a 1% agarose gel with the negative control (genomic *S. typhimurium* DNA as the template DNA).

7.2.2.2 Production of double and triple mutants

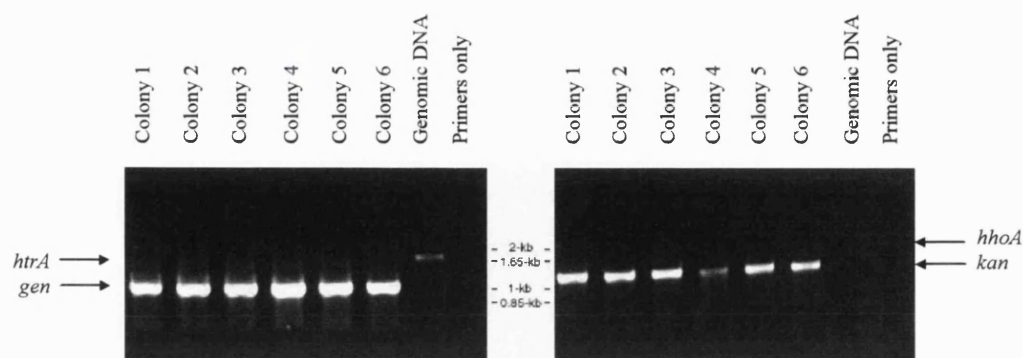
In order to construct the double and triple mutants of the *htrA* gene family, a second round of allelic replacement was carried out. As the plasmid pBAD λ RED is required to facilitate allelic replacement it was necessary to check if it was still present in the selected strains, *htrA::GM-1* and *htrA::CAT-1*. This was achieved by culturing the strains in L-Broth supplemented with ampicillin, as pBAD λ RED has an ampicillin resistance cassette encoded in it. Unfortunately, the *htrA::CAT-1* bacteria did not grow in these conditions suggesting that the plasmid has been lost and so further construction of double and triple mutants was performed using *htrA::GM-1* only.

The second round of allelic replacement was identical to that outlined for the construction the single *kan^r* mutants and so the appropriate linear PCR fragments were electroporated into the *htrA::GM-1* strain after induction using arabinose and preparation into electrocompetent cells. Colonies were picked and the presence of *htrA::GM* and the presence of the *kan* gene was verified using colony PCR (Figure 7.4) and confirmed by DNA sequencing.

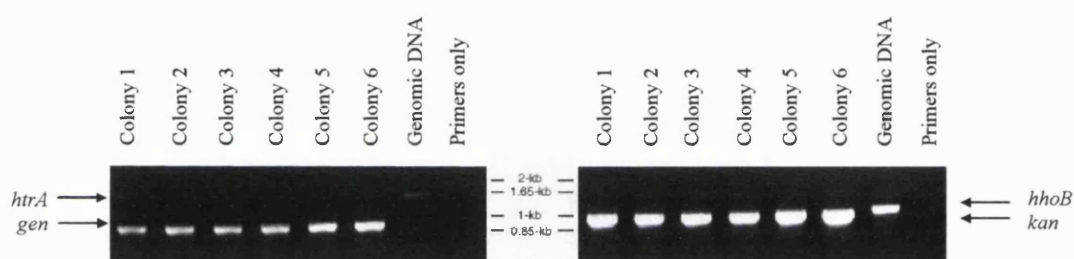
7.3 Viability of mutants under *in vitro* stress conditions

A freshly streaked colony for each mutant was picked and grown overnight. An aliquot of overnight culture was inoculated into fresh broth and grown at 37°C. Cultures were then diluted to give the same OD₆₀₀ and 200 μ l aliquots were plated into a 100-well honeycomb plate. Plates were then incubated overnight in a Bioscreen machine. This machine cultures the bacteria in the honeycomb plate at a

a) *htrA::GM;hhoA::KAN*



b) *htrA::GM;hhoB::KAN*



c) *htrA::GM;hhoAB::KAN*

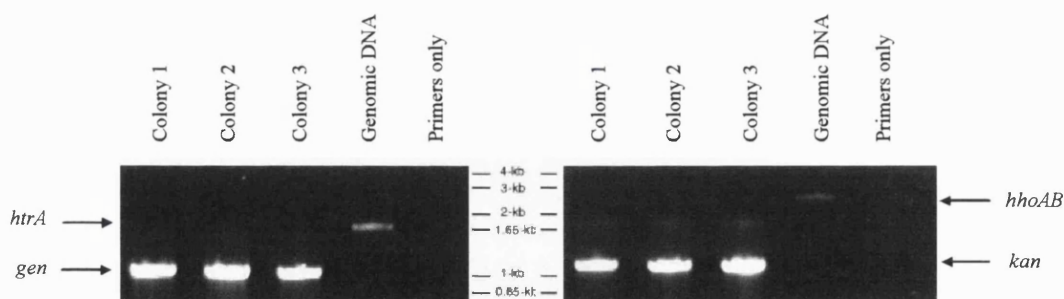


Figure 7.4 – Verification of allelic replacement in the double and triple mutants

Single mutant colonies of *htrA::GM;hhoA::KAN* (a), *htrA::GM;hhoB::KAN* (b) and *htrA::GM;hhoAB::KAN* (c) were checked by colony PCR using primers that flank the *htrA* gene. The PCR products of three single colonies were separated on a 1% agarose gel with the negative control (genomic *S. typhimurium* DNA as the template DNA).

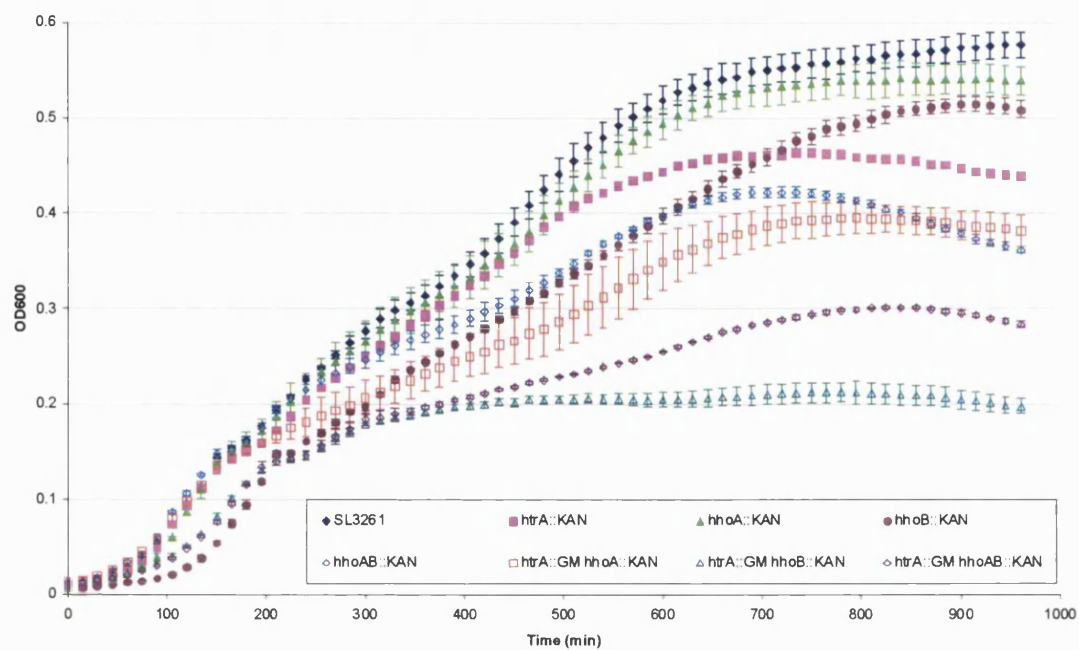
constant temperature and measures the density of culture in each well at 600 nm at specified intervals over a set period of time. At the end of the set time period bacteria underwent a series of dilutions and plated to obtain viable bacterial counts.

7.3.1 Survival of mutants under heat shock conditions

Sublethal heat shock conditions were used in experiments investigating heat shock. The bacteria was grown at 42°C or at 37°C for 16 hr and OD₆₀₀ readings were taken every 15 min, which were plotted to obtain a growth profile. Figure 7.5 shows the comparison of the growth profiles of the parental strain SL3261 and the *htrA* family of mutants when grown at 42°C or at the control temperature 37°C.

Comparisons of the single mutants are illustrated more clearly in (Figure 7.6). Growth profiles for all the bacterial strains are very similar at 42°C and 37°C. However, the *htrA::KAN* mutant appears to be affected under the heat-shock conditions used. At the higher temperature of 42°C, the *htrA::KAN* mutant reached a plateau at an OD₆₀₀ value significantly lower than that observed at 37°C ($p < 0.001$). This value was also significantly lower than that of SL3261 at 42°C ($p < 0.001$). Mutations in *htrA* in other species show that this gene is important for survival under heat shock conditions (Strauch, K.L. & Beckwith, J., 1988, Lipinska, B. *et al.*, 1989, Lipinska, B. *et al.*, 1990, Seol, J.H. *et al.*, 1991), however, mutating *htrA* in *S. typhimurium* did not affect bacterial growth (Johnson, K. *et al.*, 1991). The growth profile of the *htrA::KAN* mutant reported here does not support the findings by Johnson *et al* who showed that *htrA* was not important for bacterial viability under heat-shock conditions. The work reported here shows that in *S. typhimurium* *htrA* is

a) 42°C



b) 37°C

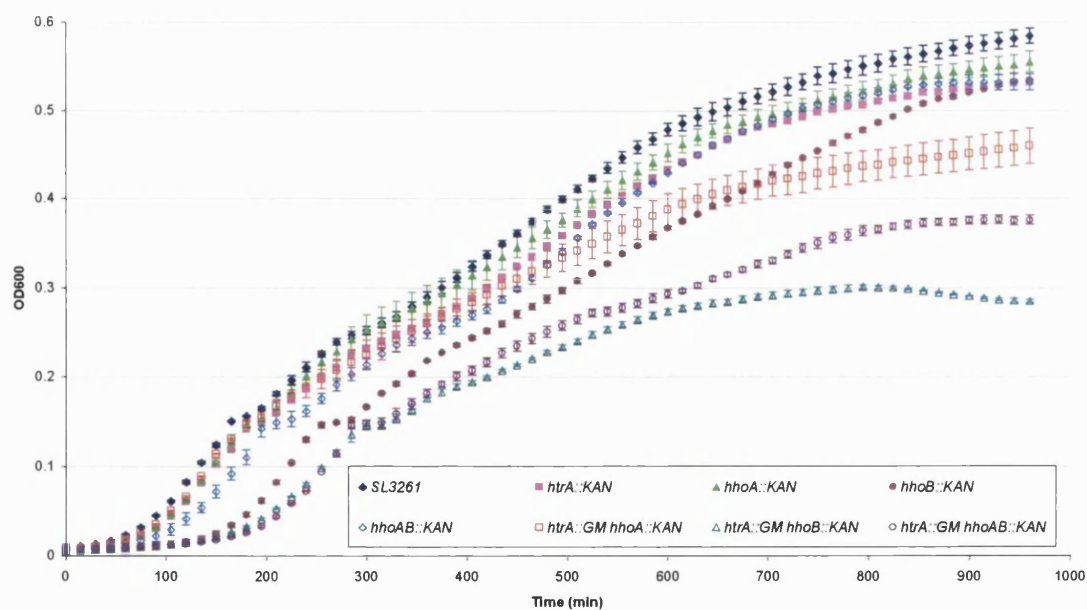
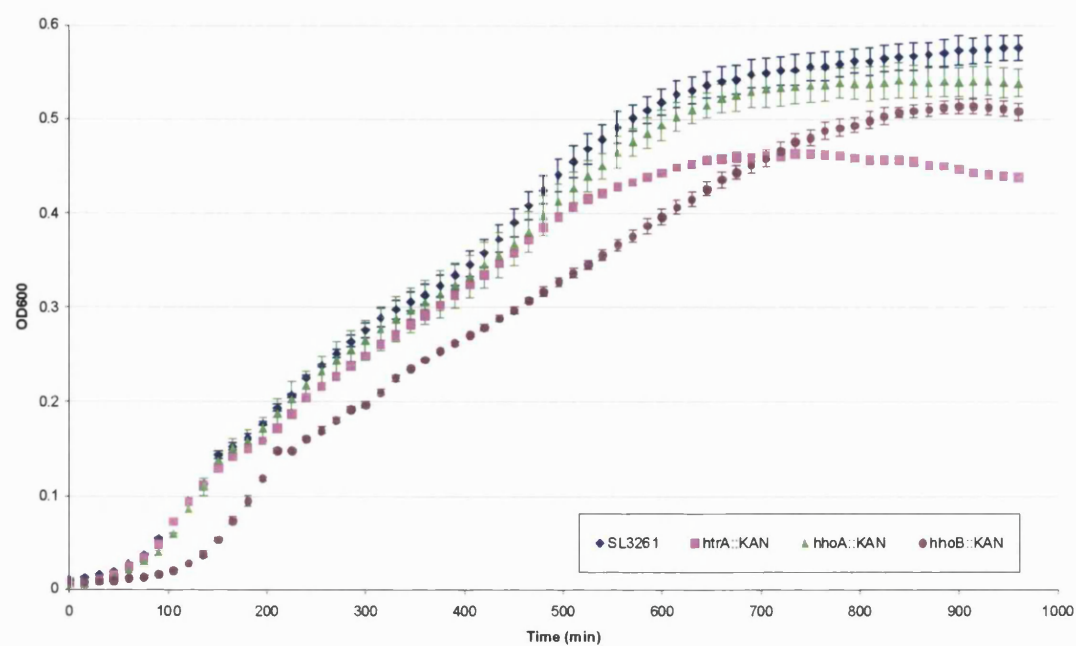


Figure 7.5 – Comparison of the growth profiles of the *S. typhimurium* SL3261 mutants at 42°C and 37°C

Typical growth profiles were measured at OD₆₀₀ over 16 hr at 42°C (a) or 37°C (b).

(n = 3, repeated 4 times)

a) 42°C



b) 37°C

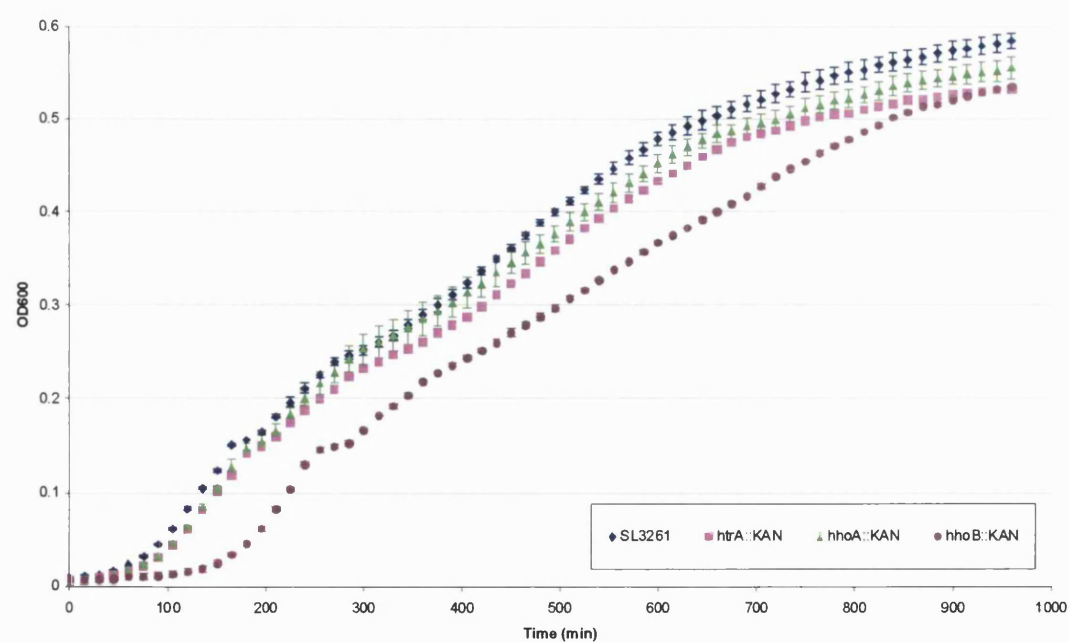


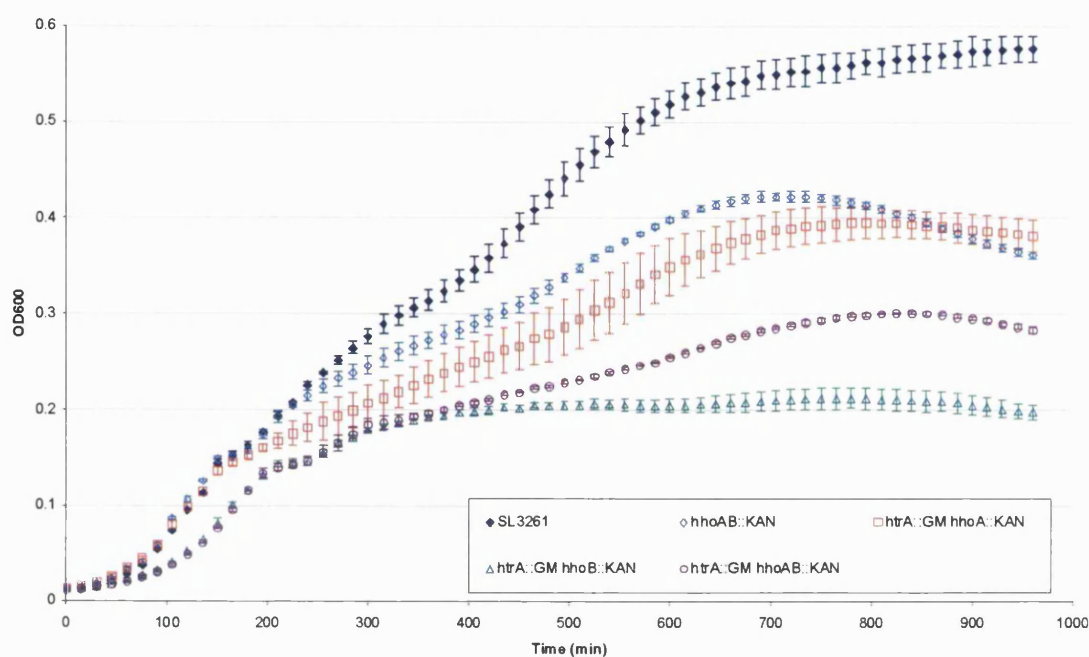
Figure 7.6 – Comparison of the growth profiles of the *S. typhimurium* SL3261 single mutants at 42°C and 37°C

Growth profiles were measured at OD₆₀₀ over 16 hr at 42°C (a) or 37°C (b). (n = 3, repeated 4 times)

required for survival under heat-stress conditions and supports the findings from the *E. coli htrA* mutants (Strauch, K.L. & Beckwith, J., 1988, Lipinska, B. *et al.*, 1989, Lipinska, B. *et al.*, 1990, Seol, J.H. *et al.*, 1991). At both temperatures, the *hhoA::KAN* mutant had a growth profile similar to that seen in the parental SL3261. This finding supports previous reports in *E. coli* (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996) showing that mutations in *hhoA* do not affect bacterial growth during heat-shock. The *hhoB::KAN* mutant reached a plateau at an OD₆₀₀ value significantly lower to that seen at 37°C ($p < 0.01$). Surprisingly, the mutant reached a higher density than that seen in the *htrA::KAN* mutant ($p < 0.001$). However, between 0 to 800 min, there is a lag in *hhoB::KAN* mutant growth, which occurs at both 37°C and 42°C. This suggests that *hhoB* has a role in bacterial growth under non-heat-shock conditions, which is exacerbated at higher temperatures.

Comparisons of the growth profiles of the double and triple SL3261 mutants (Figure 7.7) show that at 37°C, the growth of all the mutants is higher than that seen at 42°C. The increase in temperature appears to significantly affect the growth of the *hhoAB::KAN* double mutant ($p < 0.0001$). At 42°C, the *hhoAB::KAN* mutant reached a plateau at an OD₆₀₀ significantly lower than both the *hhoB::KAN* ($p < 0.0001$) and *hhoA::KAN* ($p < 0.0001$) single mutants. This strain did not have the observed lag time seen for the *hhoB::KAN* mutant between 0 to 200 min, however, the growth slowed down after 200 min incubation at 42°C. At 37°C, the *hhoAB::KAN* mutant did not significantly differ from the single mutants, *hhoA::KAN* and *hhoB::KAN*. This suggests that the loss of both these genes is additive in terms of disadvantageous for bacteria growth. This is especially interesting, as after 16 hr, the

a) 42°C



b) 37°C

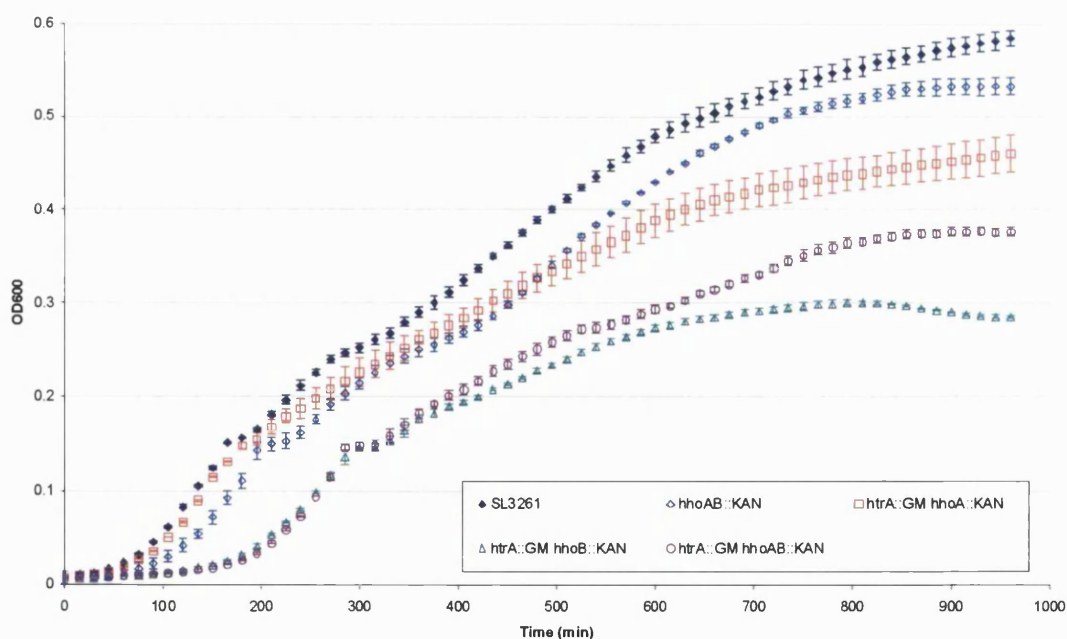


Figure 7.7 – Comparison of the growth profiles of the *S. typhimurium* SL3261 double and triple mutants at 42°C and 37°C

Growth profiles were measured at OD₆₀₀ over 16 hr at 42°C (a) or 37°C (b). (n = 3, repeated 4 times)

growth for the *hhoA* single mutation was not significantly affected at the higher temperature.

The double mutant *htrA::GM hhoA::KAN* is also affected during heat shock in comparison to the parent strain ($p < 0.0001$) and also the two single mutants ($p < 0.01$). This is supported by previous studies in *E. coli* (Waller, P.R.H. & Sauer, R.T., 1996) that have suggested that *hhoA* can compensate the loss of *htrA* function in *htrA* mutants. These studies showed that overexpression of *hhoA* in these strains relieved the thermosensitive phenotype of the *htrA* mutant, probably because like HtrA, HhoA has serine protease activity and is located in the periplasm. This finding is further supported by the amino acid similarity between the two.

At both 42°C and 37°C, the *htrA::GM hhoB::KAN* double mutant and the *htrA::GM hhoAB::KAN* triple mutants had very similar growth profiles to each other. Both strains reached a plateau at a significantly lower OD₆₀₀ value than the wild-type ($p < 0.0001$). The lag in growth, observed with the single *hhoB::KAN* mutant, was also seen albeit more pronounced. This was seen particularly at 42°C where the lag in growth lasted approximately 100 min longer than at 37°C. One main difference between the *htrA::GM hhoB::GM* double mutant and the *htrA::GM hhoAB::KAN* triple mutant is that the triple mutant grows slightly better than the double mutant as the OD₆₀₀ values were consistently higher at each time point. Overall, it can be seen that mutating the *hhoB* and *htrA* genes together severely affects the ability of the bacteria to grow under these conditions. Interestingly, mutating the *hhoA* gene alone

appears to have little effect on bacterial growth, and if anything may facilitate bacterial growth when compared to the *htrA::GM hhoB::KAN* mutant.

The results discussed so far have used the culture density as a measure of bacterial growth. However, density readings at OD₆₀₀ can be misleading because although optical density does correlate with cell numbers, it will also correlate with the cell volume. This is important in these studies as it was noted that the colonies formed by the *hhoB* mutants were much smaller than the *htrA* and *hhoA* mutants when plated on L-agar plates. In addition, density readings do not finely discriminate between live and dead bacteria in the culture, and thus may produce misleading results. In order to address these variables, viable bacterial counts were performed on the samples after 16 hr incubation (Figure 7.8). Interestingly, viable count experiments demonstrate that all the mutants have a similar survival count following growth at 37°C although there is a slight decrease in viability in the *hhoB::KAN*, *htrA::GM hhoB::KAN* and *htrA::GM hhoAB::KAN* mutants. At 42°C, there is a very slight decrease in viability in the parent SL3261 strain, whereas there are significant decreases in viability for all the mutants notably, *htrA* ($p < 0.01$), *htrA hhoB* ($p < 0.0001$) and *htrA hhoAB* ($p < 0.005$). The viable counts support the results from the growth profiles which suggest that *htrA* is important in heat shock as there is a significant decrease in viability. The *htrA::GM hhoB::KAN* double mutant has even less ability to survive when compared with all the other mutant strains. Growth at the elevated temperature of 42°C had a lesser affect on the number of *hhoA::KAN* colonies obtained after 16 hr ($p < 0.05$). This supports the growth profiles obtained and suggests that the *hhoA* gene is not significant in maintaining bacterial viability at elevated temperatures in the presence of the other two family members. It is

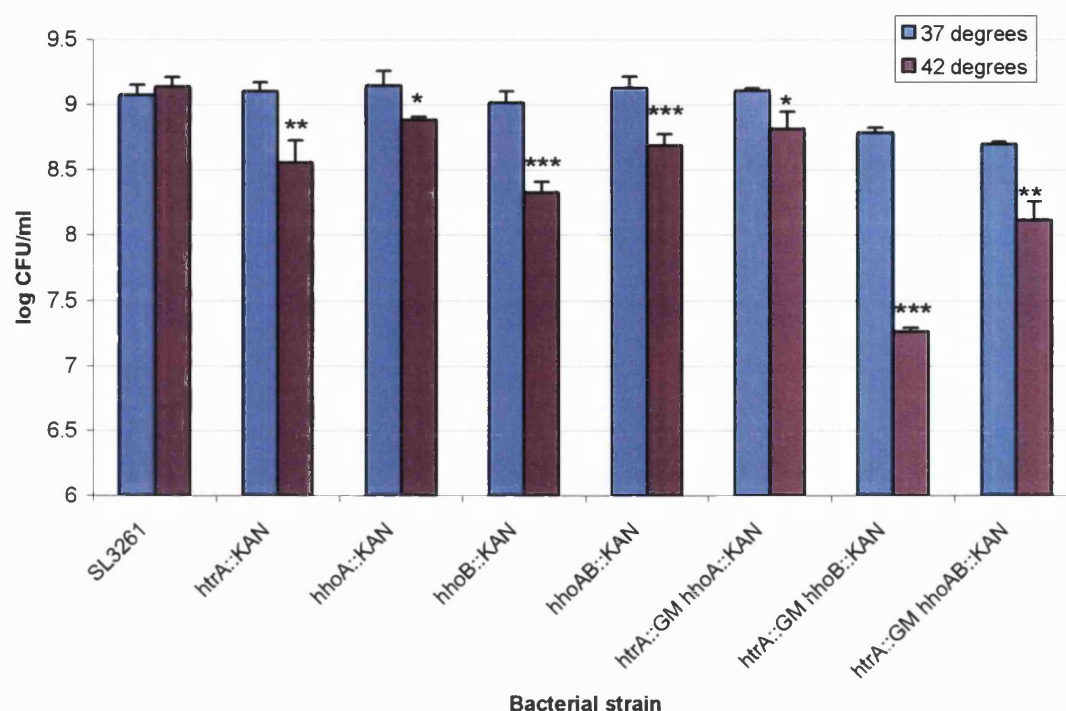


Figure 7.8 – Bacterial viability under heat shock conditions

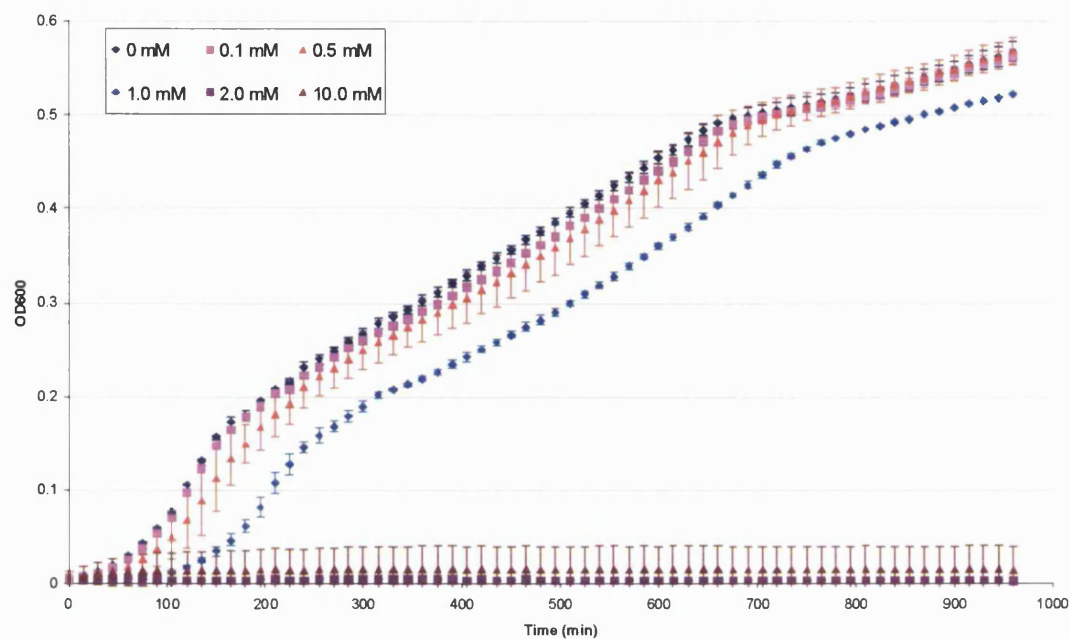
Viability of mutants were investigated after 16 hr growth in sub-lethal heat shock conditions (42°C) and under control conditions (37°C). Statistical analysis was performed using the Student's t-test comparing the heat-shock conditions against the control conditions; $p < 0.05$ (*); $p < 0.01$ (**), $p < 0.005$ (***). (n = 3, repeated 3 times)

interesting to note that there appears to be a positive effect when mutating *hhoA*. All mutants with the *hhoA* mutation had a higher viability count, under heat shock conditions, than the respective single and double mutants. However, further work must be carried out to determine the role of *hhoA* and *hhoB* in order to understand the mechanisms involved in heat shock survival.

7.3.2 Survival of mutants under oxidative stress conditions

Another common stress condition which *S. typhimurium* encounter is oxidative stress. On invasion, the host immune system engulfs the bacteria within macrophages and subject it to the oxidative burst. In order to test whether the members of the *htrA* gene family were important to bacterial survival during oxidative stress, the *S. typhimurium* mutants were subjected to increasing concentration of hydrogen peroxide (H_2O_2), a source of hydroxy radicals. Figure 7.9 shows that increasing concentrations of H_2O_2 results in an increasing lag-time before bacterial growth begins. For all the mutants, the density of bacteria after 16 hr reaches the wild-type OD_{600} value when subjected to concentrations up to 1 mM H_2O_2 . Concentrations higher than 2 mM H_2O_2 , reduced the ability of any of the bacteria strains, including the parent SL3261, to grow. In Figure 7.10, the mutants have been plotted to show the effect of increasing H_2O_2 concentrations against the growth profile of parental SL3261 strain in the absence of H_2O_2 . It can be seen that the growth profiles are similar between 0 mM and 1 mM (Figure 7.10a-d) while at 2 mM and 10 mM H_2O_2 (Figure 7.10e-f) there appears to be no growth. To compare the differences between the bacterial strains, the culture density (OD_{600}) after 3.5 hr exposure to 1 mM H_2O_2 were compared and plotted as a percentage of the culture density observed when the bacteria strains were incubated in the absence of H_2O_2

a) *S. typhimurium* SL3261



b) *htrA::KAN*

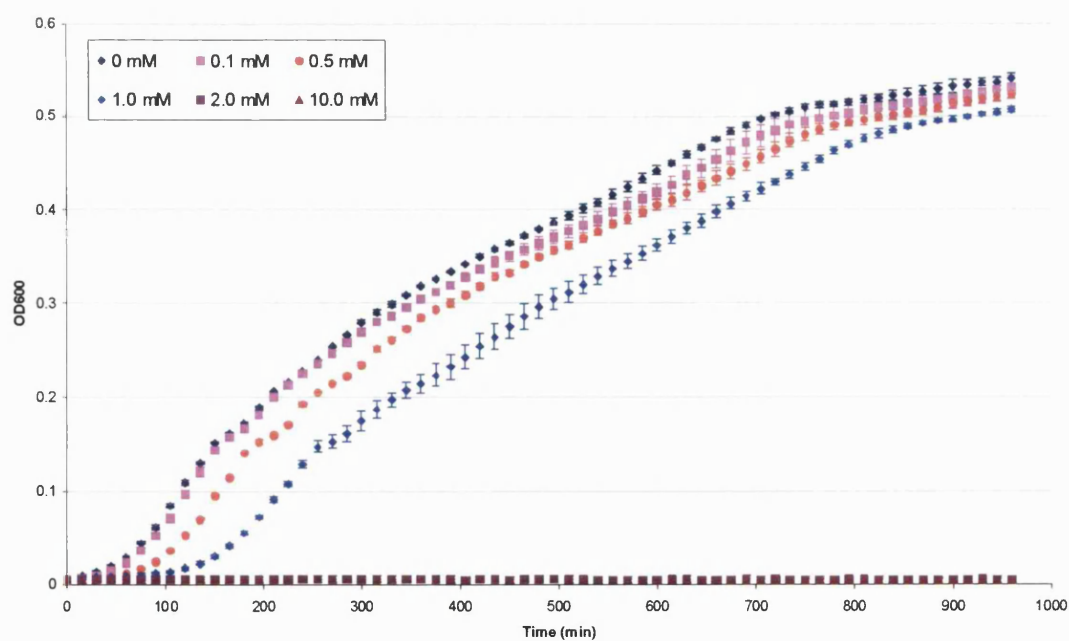
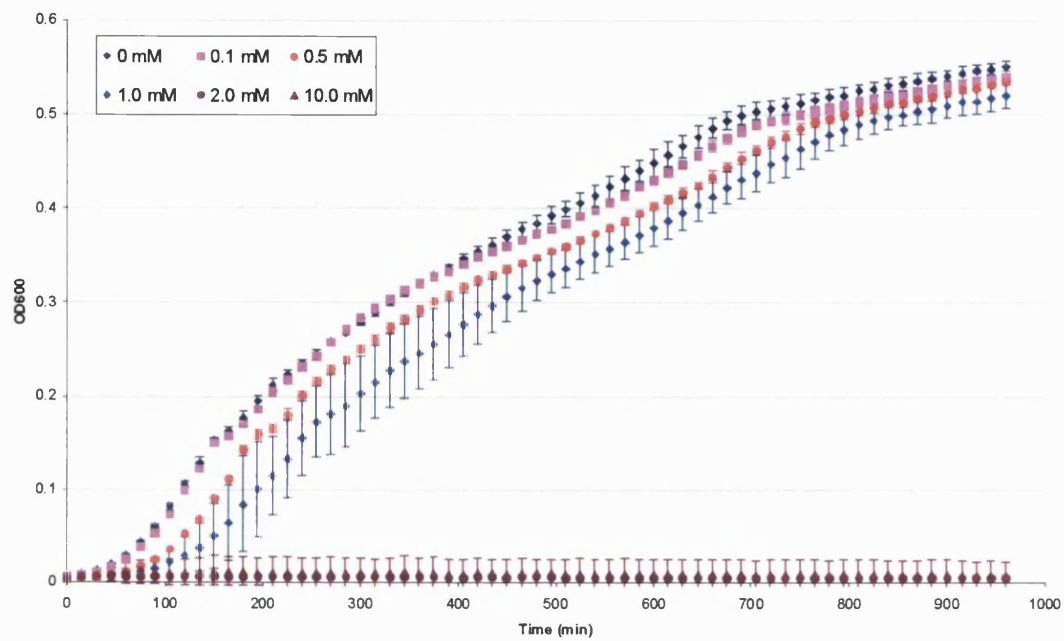


Figure 7.9 – Growth profile of *S. typhimurium* SL3261 mutants in the presence of increasing concentrations of hydrogen peroxide

Growth profiles were measured at OD₆₀₀ over 16 hr at 37°C in the presence of increasing concentrations of hydrogen peroxide. (n = 3, repeated 2 times)

c) *hhoA::KAN*



d) *hhoB::KAN*

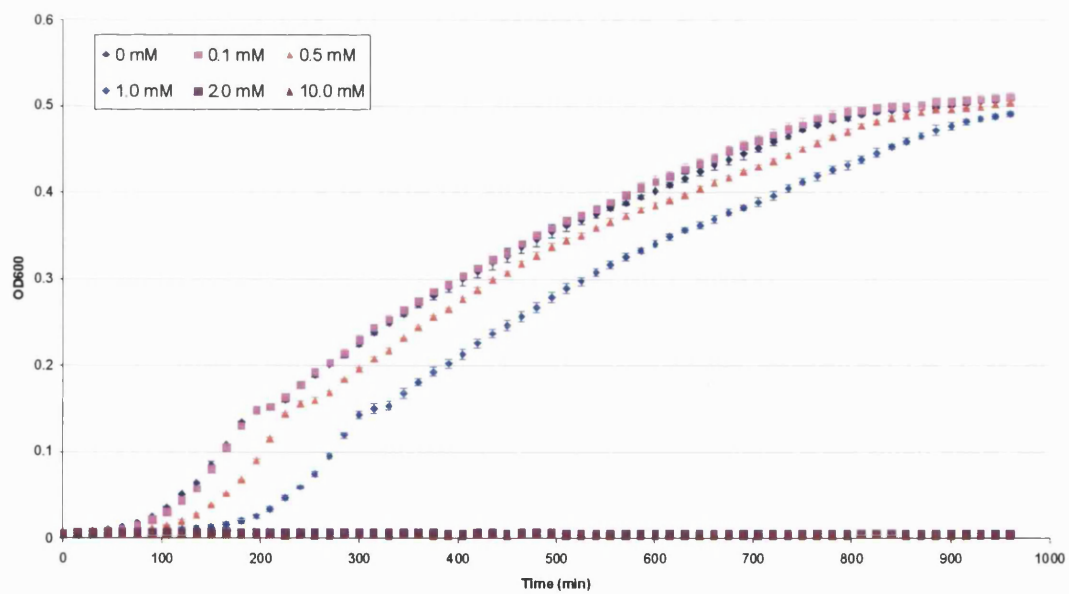
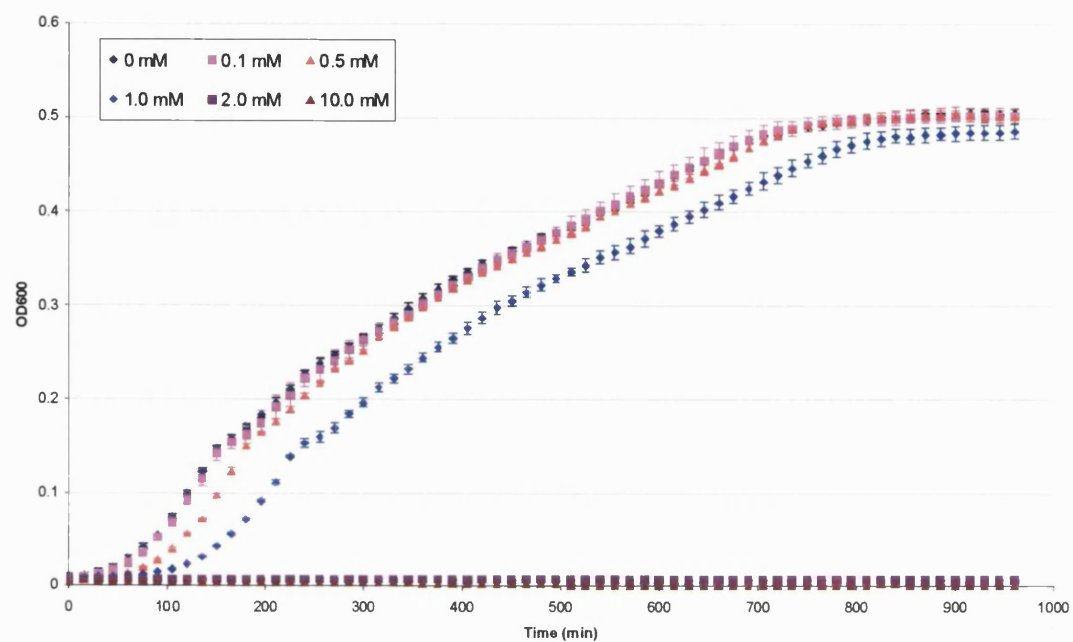


Figure 7.9 continued

e) *hhoAB::KAN*



f) *htrA::GM hhoA::KAN*

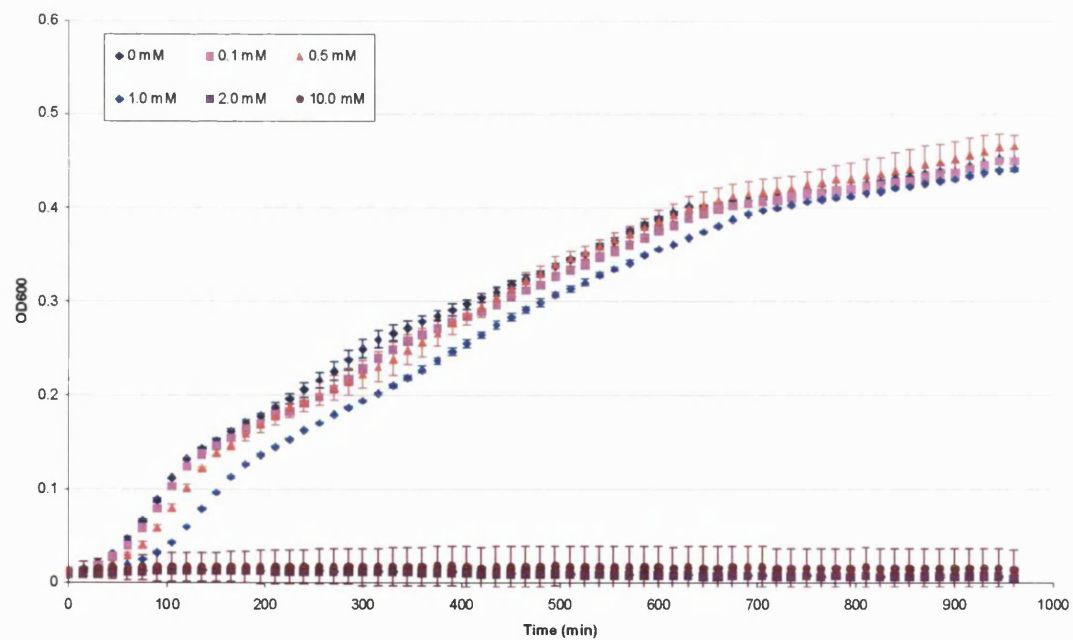
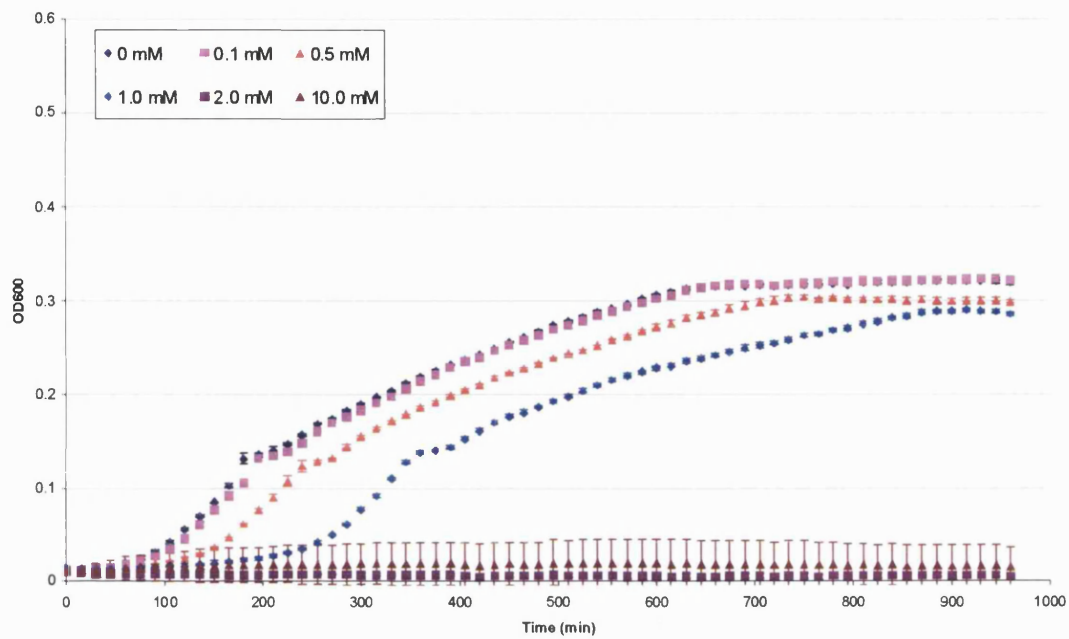


Figure 7.9 continued

g) *htrA::GM hhoB::KAN*



h) *htrA::GM hhoAB::KAN*

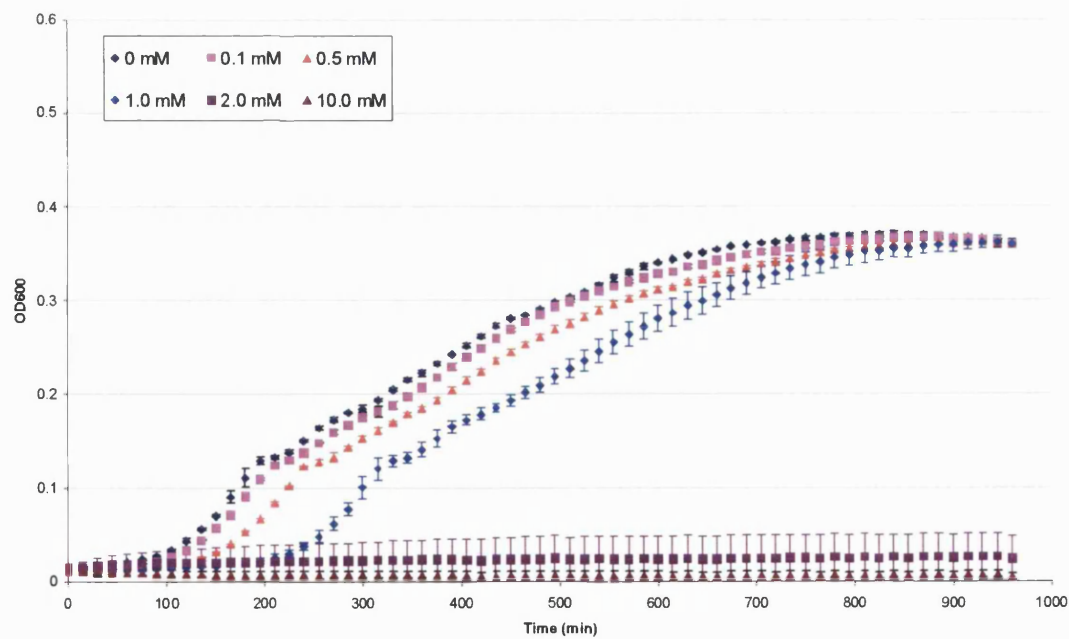
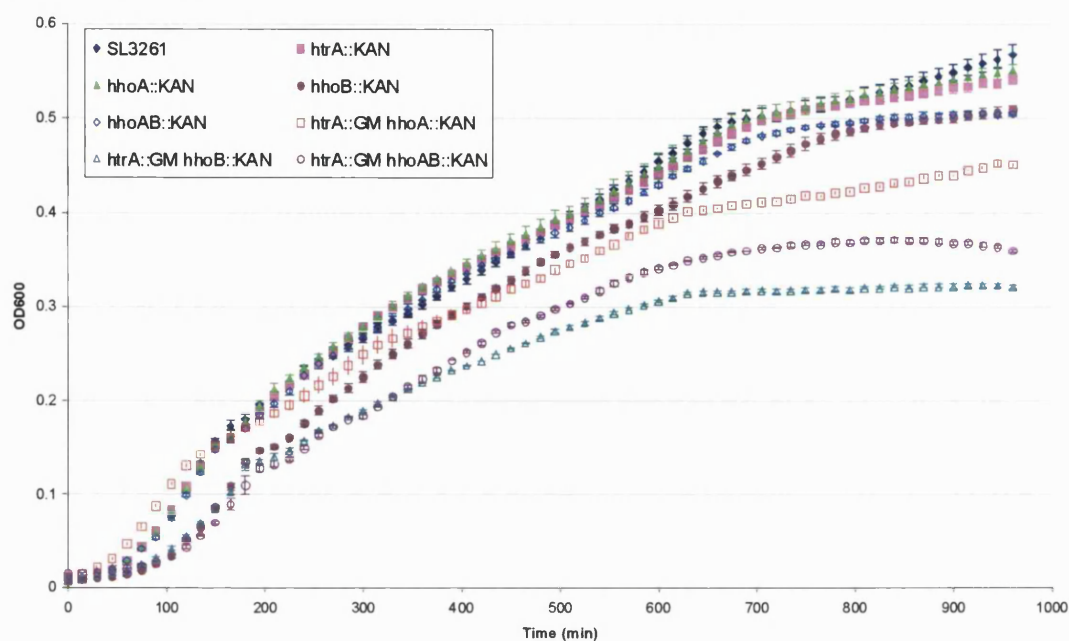


Figure 7.9 continued

a) 0 mM H_2O_2



b) 0.1 mM H_2O_2

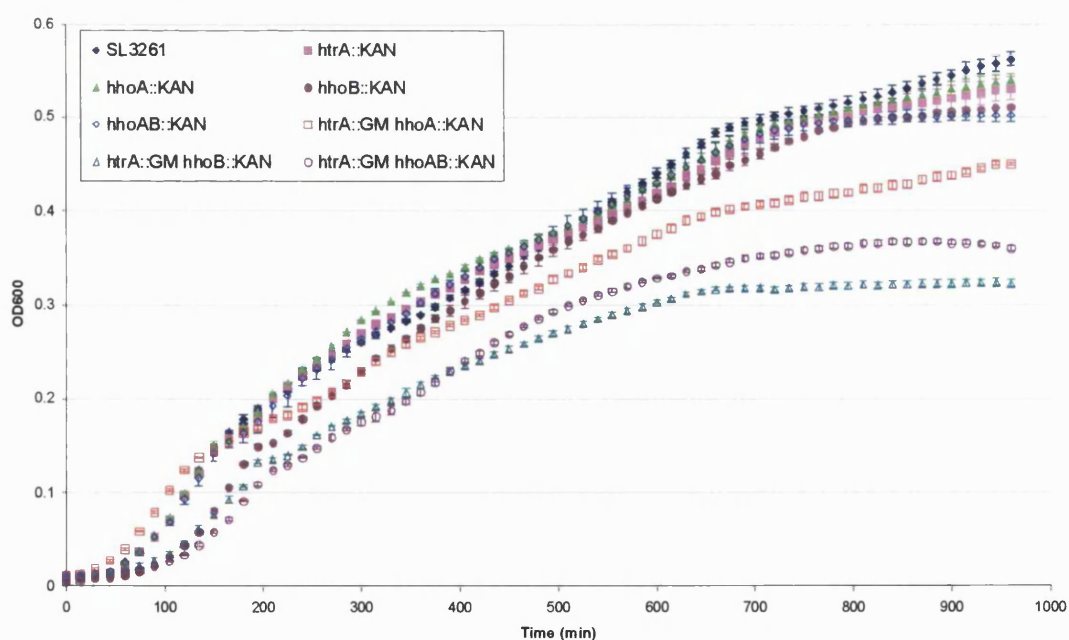
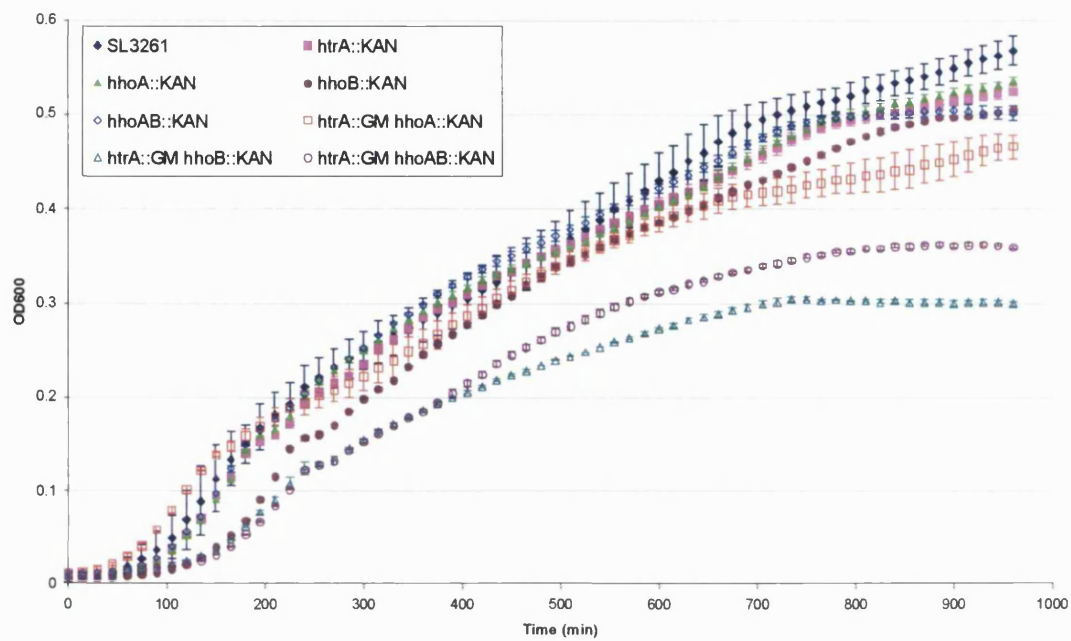


Figure 7.10 – Effect of increasing concentrations of H_2O_2 on SL3261 mutant growth

For each concentration of H_2O_2 used, the relative growth profile of each mutant has been plotted. In addition, the growth profile of SL3261 in the absence of H_2O_2 was plotted to indicate the effect of the presence of H_2O_2 . (n = 3, repeated 2 times)

c) 0.5 mM H₂O₂



d) 1.0 mM H₂O₂

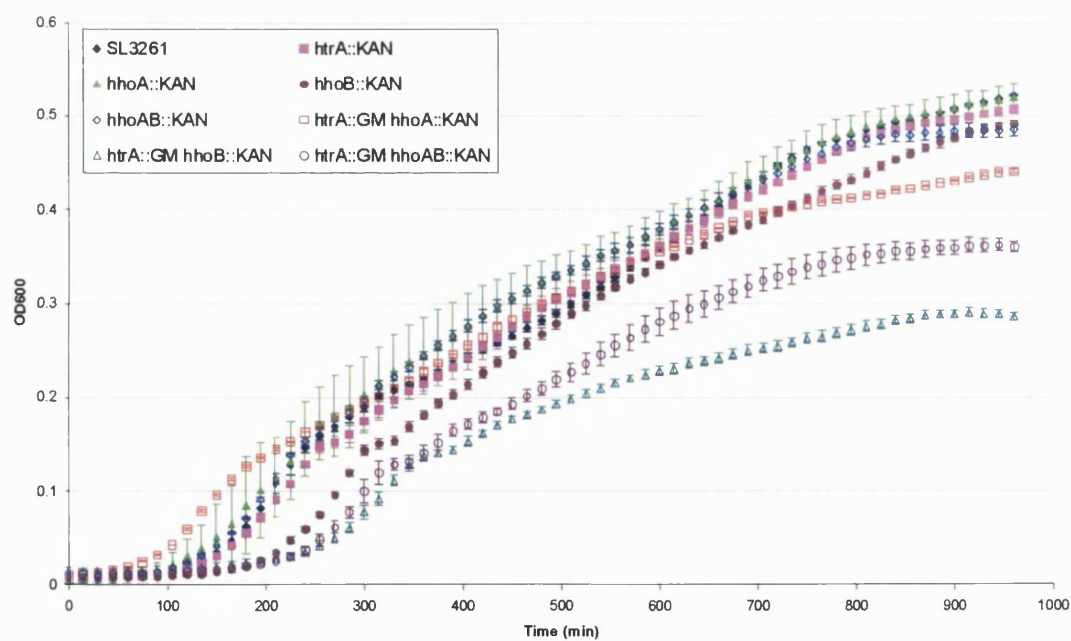
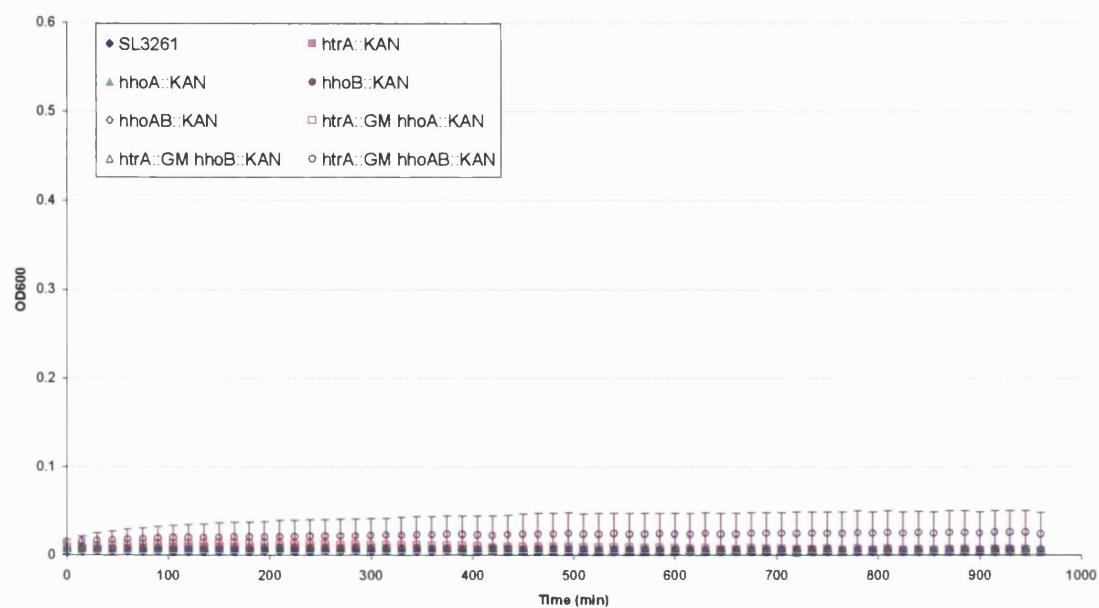


Figure 7.10 continued

f) 2.0 mM H₂O₂



g) 10.0 mM H₂O₂

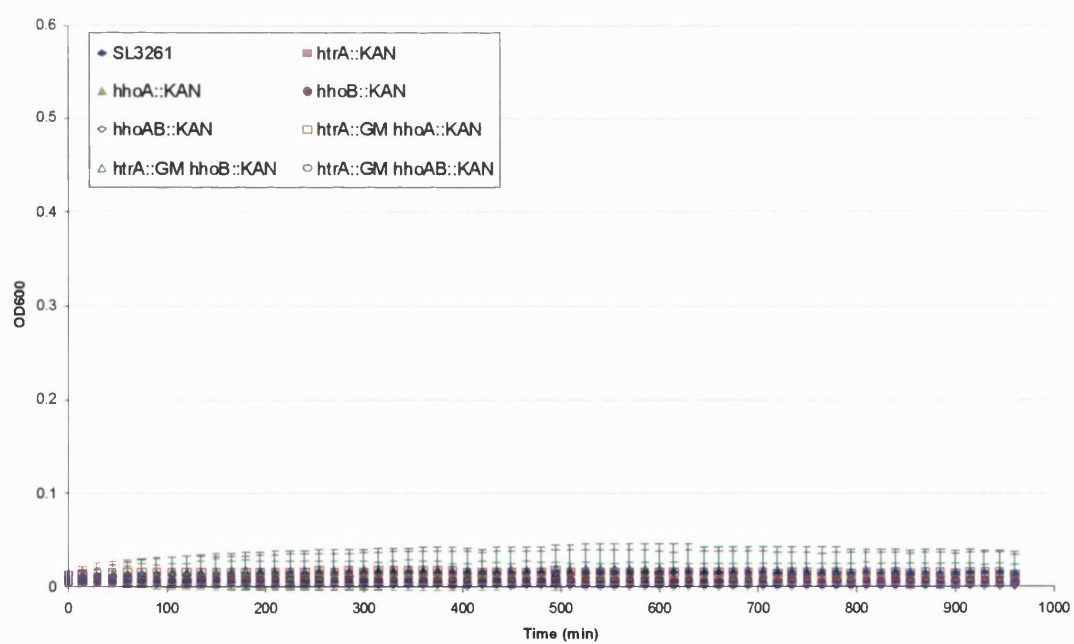


Figure 7.10 continued

(Figure 7.11). It is interesting to note that under these conditions, the culture density of single and double mutants containing the *hhoA* mutation appear to be as similar or higher than the parent strain. Contrastingly, the culture density of mutants containing the *hhoB* mutation (except *hhoAB::KAN*) is significantly lower than the rest ($p < 0.05$ with respect to the *hhoA::KAN* strain and $p < 0.001$ in comparison to all the other non-*hhoB*-containing strains). However, they do not significantly differ between themselves. This data suggests that *hhoB* is important for bacteria growth under these conditions. The *hhoA* gene does not appear to be important for the maintenance of bacterial growth under the conditions used. Interestingly, the *hhoAB::KAN* and the *htrA::GM hhoA* double mutants appear to be able to grow better than their respective single mutants. This suggests that mutating *hhoA* results in a positive effect on bacterial growth under these conditions.

Johnson *et al* (1991) reported that HtrA was important for bacterial viability when exposed to H_2O_2 . The data reported in this thesis, however, does not show any significant difference between the parental strain and the mutant *htrA::KAN* mutant. However, it is important to note that this deduction is obtained by comparing culture density, which as discussed previously (7.3.1), is relative to the number and volume of the bacterial cells but does not distinguish between live and dead bacteria in the culture. Attempts were made to investigate viability under oxidative stress conditions using a range of techniques and conditions, but reproducible data could not be obtained. Further work, must be done in order to clarify the effect of H_2O_2 on l viability of these mutants.

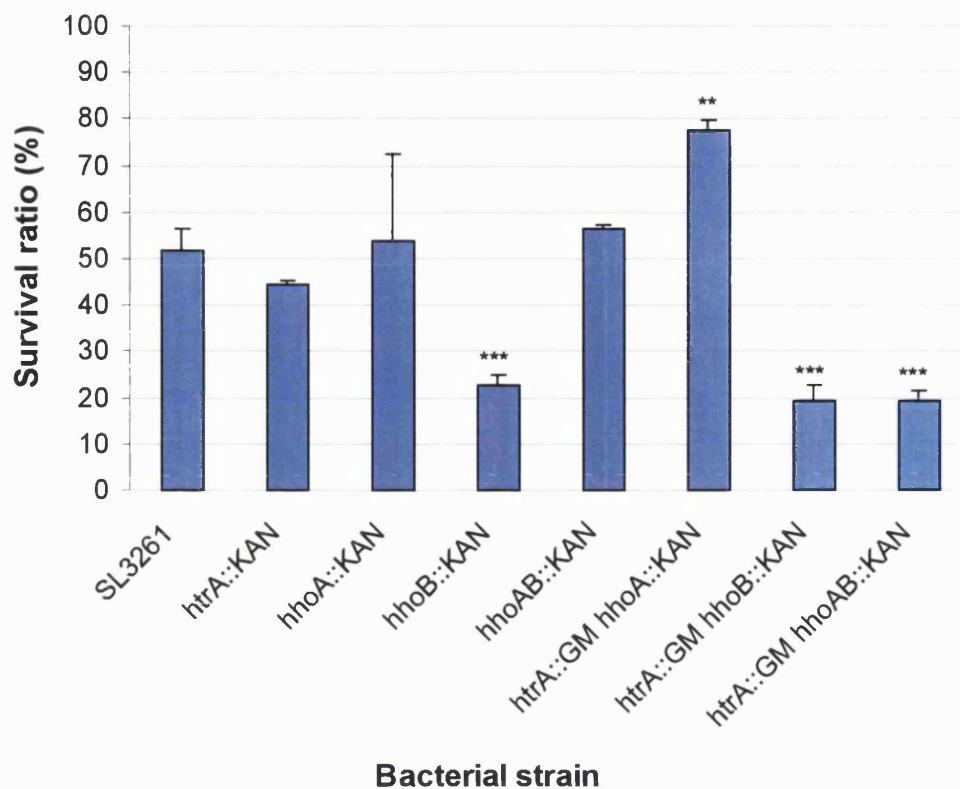


Figure 7.11 – Effect of exposure to 1 mM H₂O₂ for 3.5 hr at 37°C

Percentage bacterial growth in the presence of 1 mM H₂O₂ as compared with control growth (0 mM H₂O₂). Statistical analysis was carried out using the unpaired Student's t-test comparing the different mutant strains against the parent strain SL3261; $p < 0.001$ (***), $p < 0.005$ (**) ($n = 3$, repeated 2 times)

7.4 Viability of mutants in an *in vivo* system

7.4.1 Transduction of *S. typhimurium* SL3261 mutants into *S. typhimurium* SL1344

The studies performed have examined two stress conditions encountered by *S. typhimurium*. Although these *in vitro* studies can lend an understanding of the mechanisms of bacterial viability under stress conditions, the ability of the mutants to survive and propagate within a host must be carried out in *in vivo* conditions. In order to examine bacterial survival *in vivo*, mutants were obtained by transducing the mutations constructed in the attenuated SL3261 strain, into the mouse virulent strain, SL1344. The single mutants, *htrA::KAN*, *hhoA::KAN* and *hhoB::KAN*, and the double mutant *hhoAB::KAN* were transduced into SL1344 using one round of transduction. The mutants carrying the *htrA::GM* genotype were obtained by transducing into the first set of mutants to acquire the full set of mutants. The resulting strains were then verified by PCR using internal primers to confirm the presence of the *kan* gene (Figure 7.12a) and external primers to confirm the presence of the *gen* gene (Figure 7.12b) were at the correct locations in the disrupted target gene.

7.4.2 Bacterial virulence in BALB/c mice

Mutants and wild-type bacteria were grown overnight and then washed in PBS, in order to inject into BALB/c mice. Approximately 10^3 bacteria were used to infect each mouse intravenously, of which four per group were culled and the livers and spleens removed at various time-points over four to seven days. The removed organs

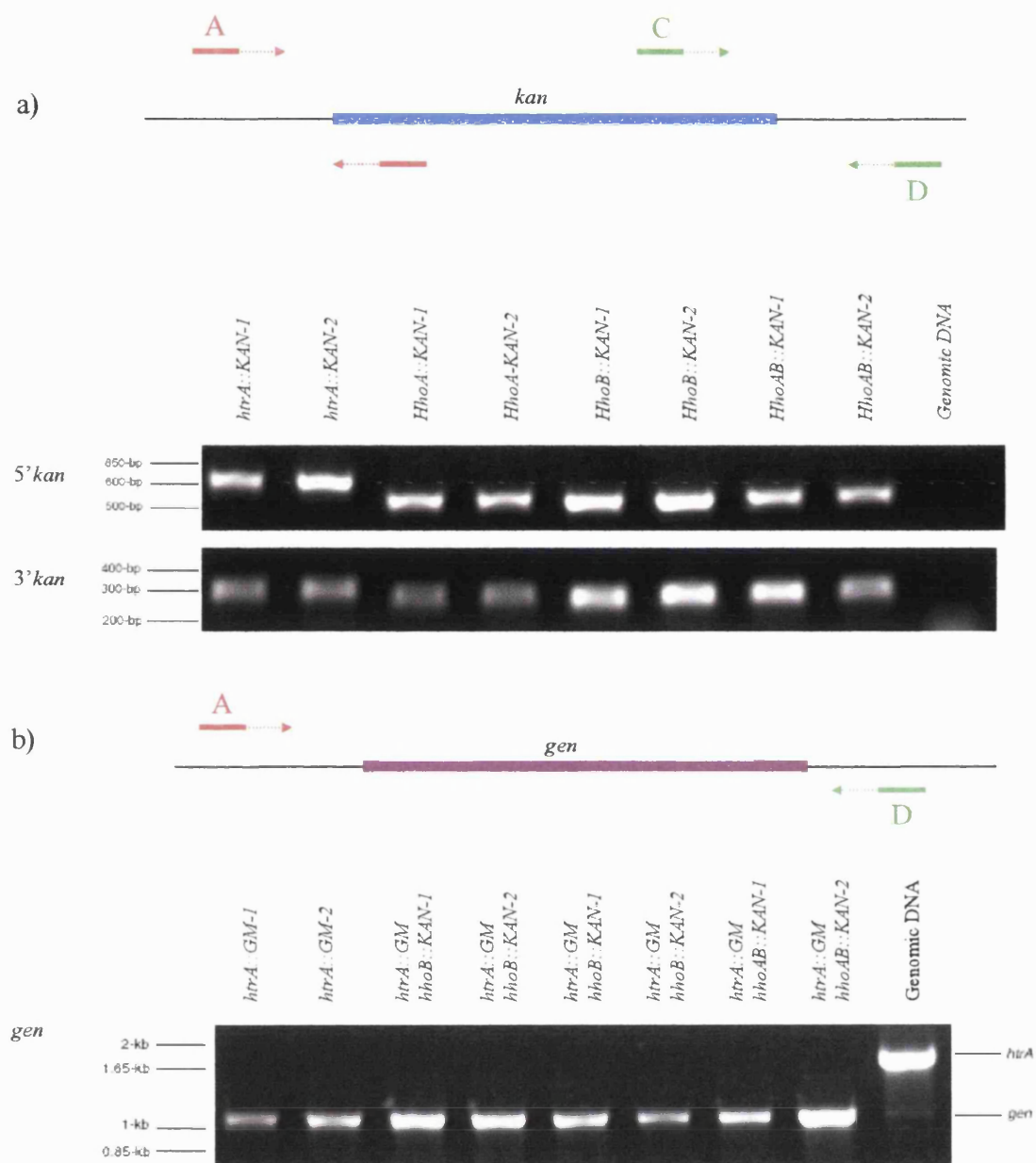


Figure 7.12 – Verification of *S. typhimurium* SL1344 mutants by PCR

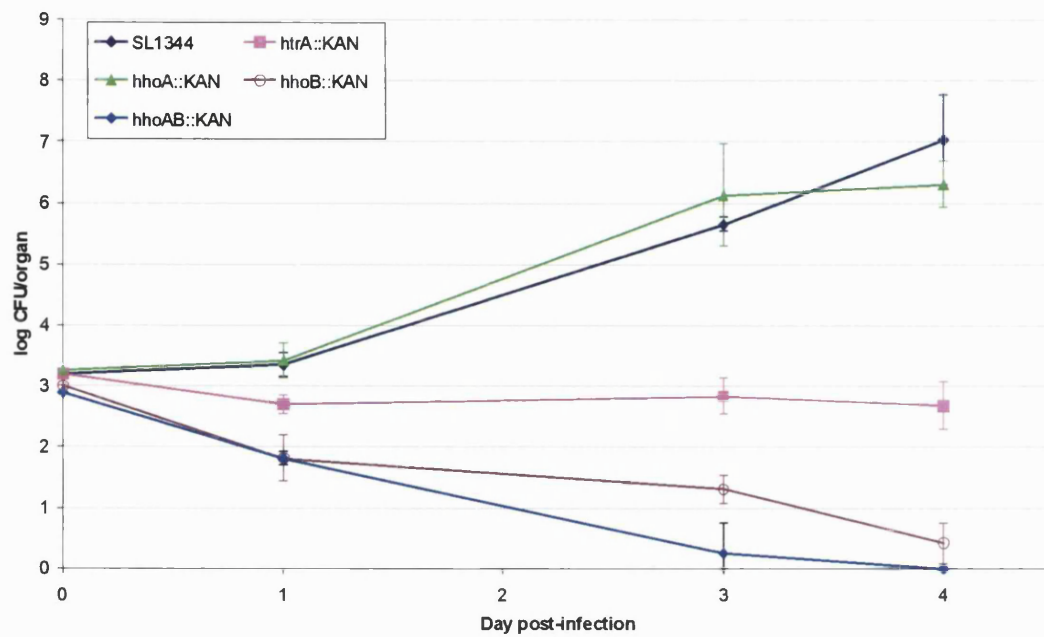
The resistance gene *kan* was detected using internal *kan* primers (B and C) and primers which anneal to DNA flanking the replaced gene (A and D). The *gen* gene was detected using primers A and D. Two colonies for each mutant was tested and ran against a control (genomic *S. typhimurium* DNA as the template DNA) in the last lane

were homogenised and dilutions were plated in order to count the number of viable bacteria at each time point.

At day 1 post-infection, all the mice appeared to be alert. On dissections, some splenomegaly was observed in mice infected with the wild-type SL1344 and with the *hhoA::KAN* mutant. At days 3 and 4 post-infection, all mice were alert except for mice inoculated with wild-type *S. typhimurium* SL1344 or with the *hhoA::KAN* mutant. These mice were much less active and were hunched and reluctant to move. They also exhibited piloerection. In addition, the livers of these mice were larger and much more fragile than mice infected with any of the other mutants. Examination of the spleens showed that they were darker and enlarged in the sick mice in comparison with the alert mice.

Figure 7.13 shows the viability of the kanamycin resistant mutants in the spleen and liver of infected mice. Over a period of four days, *hhoA::KAN* mutants do not significantly differ to wild-type SL1344 and grew 2-3 logs over the course of the experiment. The *htrA::KAN* mutant is attenuated and has significantly lower viability than either the wild-type strain or *hhoA::KAN* mutant ($p < 0.001$ at all time points). It can be seen in Figure 7.13 that in the spleen, the *hhoB::KAN* and *hhoAB::KAN* mutants are more attenuated than the *htrA::KAN* single mutant ($p < 0.01$ for all time points) and have almost cleared by day 4. At day one post-infection, the viable counts from the livers of all livers do not differ significantly, whereas at day three post-infection, there are significantly less *hhoB::KAN* bacteria recovered when compared with the *htrA::KAN* strain ($p < 0.0001$). At day four, counts for the *htrA::KAN* mutant does not significantly differ from *hhoB::KAN* although it is

a) Spleen



b) Liver

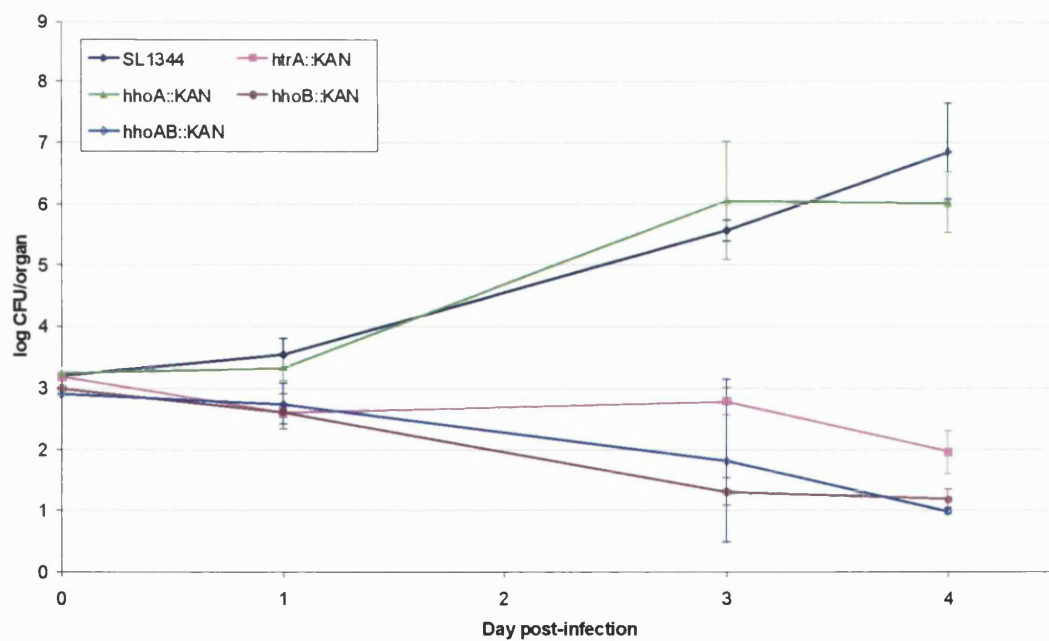


Figure 7.13 – Viability of *S. typhimurium* SL1344, *htrA::KAN*, *hhoA::KAN*, *hhoB::KAN* and *hhoAB::KAN* mutants in a mouse model of infection

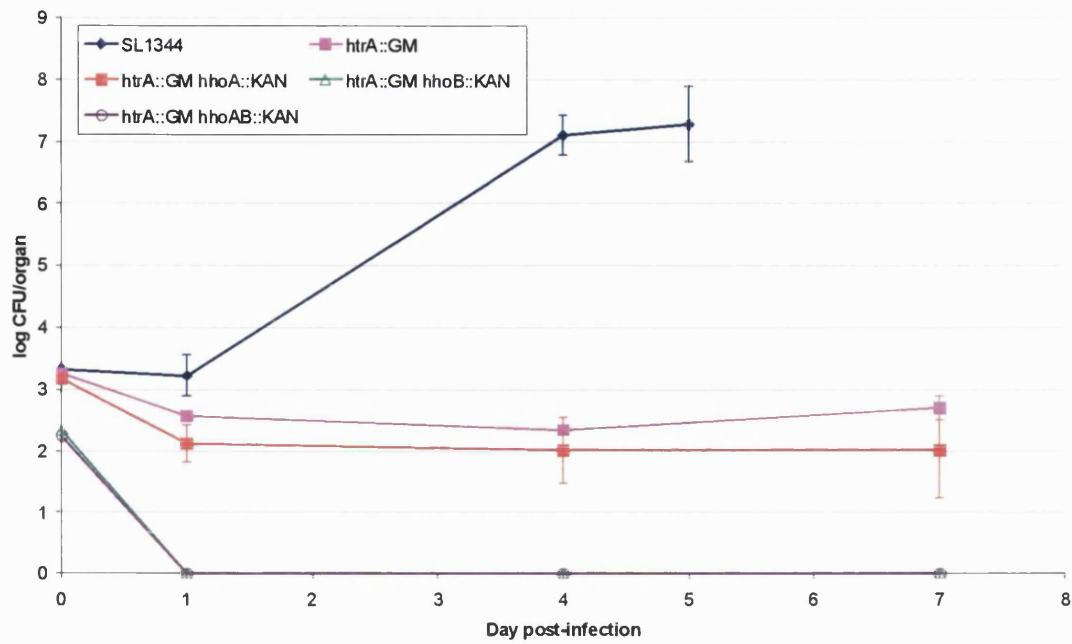
Spleens and livers were recovered from intravenously infected mice at various time points. Disrupted organs were plated to obtain viable counts. (n = 4, repeated 2 times)

significantly higher than the *hhoAB::KAN* mutant ($p < 0.05$). It appears that the *hhoB* is more important in bacterial survival than the *htrA* mutation, supporting observations seen in the *in vitro* studies discussed earlier (7.3). It should be noted that the viable counts of *hhoB::KAN* and *hhoAB::KAN* mutants are not significantly different to each other at all time points taken from the spleen samples. Thus, the mutation of *hhoA* does not appear to affect the bacterial virulence.

Figure 7.14 shows the viability of the *htrA* double and triple mutants in comparison to the wild-type strain SL1344. The mutant strains are all significantly attenuated with respect to the wild-type strain over a period of seven days. Comparison of the viability of the *htrA::GM* mutant and the *htrA::GM hhoA::KAN* mutant shows the two mutants do not differ significantly from each other except at day one post-infection in the spleen ($p < 0.05$). The *htrA::GM hhoB::KAN* and *htrA::GM hhoAB::KAN* mutants are significantly more attenuated in the spleen when compared with the parent strain and other *htrA* mutants ($p < 0.05$). Interestingly, in the liver, these two mutant strains do not appear to differ from the *htrA::GM hhoA::KAN* mutant after day four post-infection. The mutants containing both the *htrA* and *hhoB* mutations do not significantly differ from each other suggesting that the extra mutation in *hhoA* in the triple mutant does not substantially affect bacterial survival *in vivo*.

The *htrA::GM hhoA::KAN* and *htrA::GM hhoAB::KAN* mutants were highly attenuated and appear to be cleared by day one to four. To investigate further, a higher inoculum dose of 10^6 was used to infect the mice (the previous dose was 10^3) and bacteria were recovered at day one and day three post-infection. Due to the high

a) Spleen



b) Liver

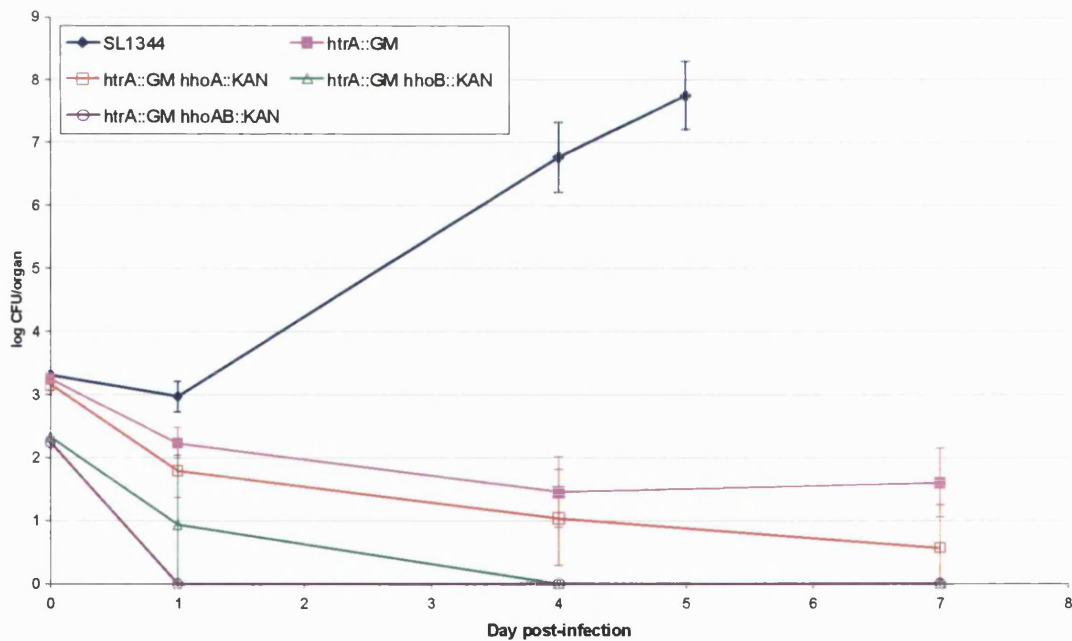


Figure 7.14 – Viability of *S. typhimurium* SL1344, *htrA::GM*, *htrA::GM hhoA::KAN*, *htrA::GM hhoB::KAN* and *htrA::GM hhoAB::KAN* mutants in a mouse model of infection

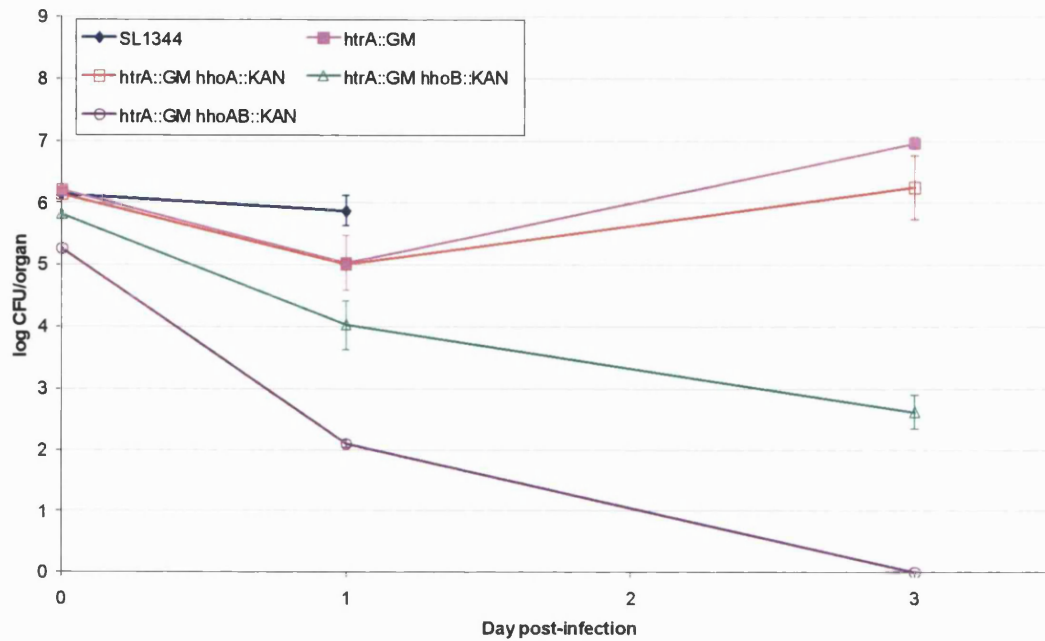
Spleens and livers were recovered from intravenously infected mice at various time points. Disrupted organs were plated to obtain viable counts. (n = 4)

inoculum dose used, experiments with the wild-type strain were not carried on beyond day one. The results of this experiment can be seen in Figure 7.15. The *htrA::GM* mutant and the *htrA::GM hhoA::KAN* mutant appear to proliferate between day 1 and day 3 post-infection. In the spleen, the viable counts do not significantly differ from viable counts recovered for the wild-type strain after day one post-infection. However, in the liver the viability of the *htrA::GM hhoA::KAN* strain is significantly lower than that of the wild-type ($p < 0.001$). At day three post-infection, it can be seen that the viability of the *htrA::GM* and *htrA::GM hhoA::KAN* mutants do not differ significantly in both the organ and the liver. This supports the previous observations (Figure 7.13 and Figure 7.14) that mutating the *hhoA* gene alone does not appear to be significant in bacterial survival. In comparison, the viability of the *htrA::GM hhoB::KAN* and *htrA::GM hhoAB::KAN* mutants are significantly lower than the wild-type, *htrA::GM* and *htrA::GM hhoA::KAN* ($p < 0.05$). At the higher inoculum, it can be seen that the viability of the triple mutant is significantly lower than the double mutant *htrA::GM hhoB::KAN* ($p < 0.05$). This suggests that mutating *hhoA* may have some significance to bacterial viability in combination with the other mutations when a higher inoculum is used.

7.5 Summary

Mutants of the *htrA* family were constructed in the attenuated *S. typhimurium* strain SL3261 by allelic replacement. Exposure of the mutants to elevated temperatures of 42°C suggests that mutating *htrA*, *hhoA* and *hhoB* in any combination reduced the viability of the mutants when compared to control conditions of 37°C. However, exposure of the mutants to H₂O₂ suggests that the *hhoB* gene is important for bacterial survival during oxidative stress, whereas the mutating the *hhoA* gene

a) Spleen



b) Liver

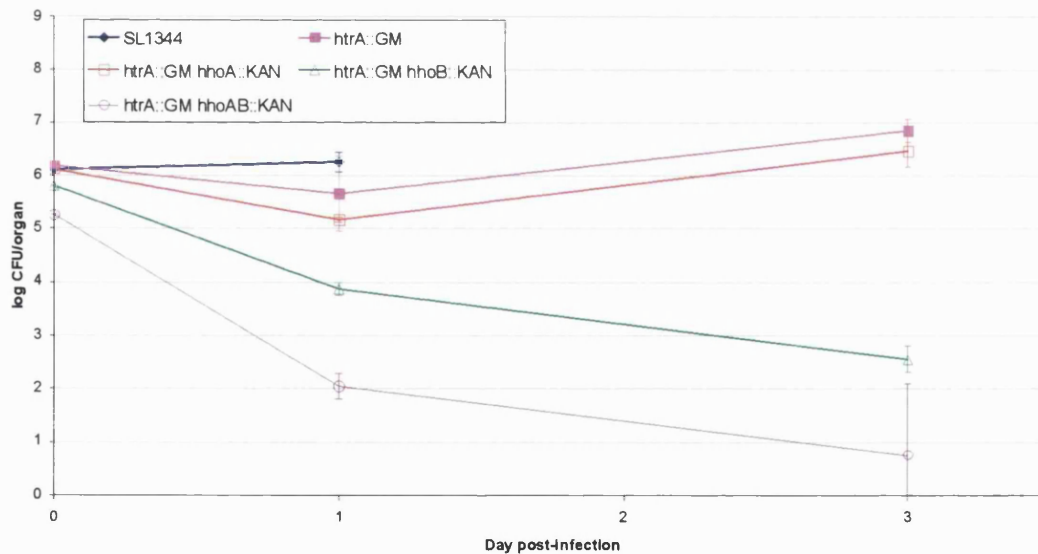


Figure 7.15 – Viability of *S. typhimurium* SL1344, *htrA::GM*, *htrA::GM hhoA::KAN*, *htrA::GM hhoB::KAN* and *htrA::GM hhoAB::KAN* mutants in a mouse model of infection using a higher inoculum dose

Spleens and livers were recovered from intravenously infected mice at various time points. Disrupted organs were plated to obtain viability counts. (n = 4, repeated 2 times)

appeared to increase bacterial growth and viability. The ability of mutants containing the *htrA* and *hhoB* mutation to grow was severely reduced under control conditions suggesting that these genes may work together in normal bacterial growth. However, the *in vitro* stress conditions used appear to exacerbate this reduced ability to grow. Thus, *htrA* and *hhoB* appear to be important for bacterial growth and viability in stress conditions *in vitro*.

In order to investigate the contributions these mutants make to *in vivo* survival, the family of *htrA* mutants (*htrA*, *hhoA* and *hhoB*) were transduced into the virulent wild-type *S. typhimurium* SL1344 and examined in an *in vivo* model of infection. Viable counts were performed on the spleens and livers recovered from mice infected with the different strains. These experiments showed that except for the *hhoA::KAN* mutant, all the mutant strains were attenuated in the infected mice. Interestingly, the presence of the *hhoA* mutation appeared to have little affect on bacterial survival, except in the triple mutant when a higher inoculum dose was used. In contrast, the survival of mutants containing the *hhoB* mutation was significantly reduced. As in the *in vitro* data, this was more pronounced when the *hhoB* mutation was in the presence of the *htrA* mutation. Thus, *htrA* and *hhoB* appear to be important in bacterial survival *in vitro* and *in vivo*.

CHAPTER 8. DISCUSSION

8.1	Introduction.....	240
8.2	Identification of HtrA homologues in <i>S. aureus</i> and <i>S. typhimurium</i>	241
8.3	HtrA homologues in <i>S. aureus</i>	243
8.3.1	SAHtrA1	243
8.3.2	SAHtrA2	246
8.4	HtrA homologues in <i>S. typhimurium</i>	248
8.5	Conclusions and future work	252

8.1 Introduction

HtrA was identified in *E. coli* as a periplasmic serine protease (Swarmy, K.H.S. *et al.*, 1983, Lipinska, B. *et al.*, 1989, Strauch, K.L. *et al.*, 1989). It is thought to remove the misfolded proteins that result from exposure of bacteria to stress conditions. In doing this, HtrA removes aberrant proteins that may otherwise accumulate and become toxic to the cell. Thus, HtrA is an important factor for bacterial survival in certain stress conditions, such as heat-shock and oxidative stress, and has also been shown to be important in bacterial virulence as *htrA* mutants are attenuated *in vivo* (reviewed in Pallen, M.J. & Wren, B.W., 1997).

HtrA has been identified in many Gram-negative bacterial species (Table 1.1). Over the last four years, HtrA have been described in six Gram-positive bacteria (Smeds, A. *et al.*, 1998, Noone, D. *et al.*, 2000, Poquet, I. *et al.*, 2000, Diaz-Torres, M.L. & Russell, R.R.B., 2001, Jones, C.H. *et al.*, 2001, Noone, D. *et al.*, 2001, Poquet, I. *et al.*, 2001, Sebert, M.E. *et al.*, 2002). However, due to the lack of a periplasmic compartment in Gram-positive bacteria, many questions have been raised about the location and function of HtrA in these species.

The aims of the present study were to identify and characterise HtrA homologues in the Gram-negative bacteria, *S. typhimurium*, and the Gram-positive bacteria, *S. aureus*. Both species are pathogens and are of clinical importance. In particular, species of *S. aureus* have now been identified that are resistant to all forms of antibiotics, thus, identifying novel targets in this organism for the development of future antibiotics is of clinical importance.

8.2 Identification of HtrA homologues in *S. aureus* and *S. typhimurium*

Using the BLAST algorithm, HtrA homologues were identified in *S. typhimurium* LT2 strain and in the unfinished genome of *S. aureus* NCTC 8325 strain. The three proteins identified in *S. typhimurium* were homologues of HtrA, HhoB and HhoA, thus, these homologues were named STHtrA, STHhoA and STHhoB, respectively. All three homologues have a high degree of identity (ranging from 82% to 98%) and similarity (ranging from 90% to 97%) with their *E. coli* counterparts at the amino acid level (Figure 3.2). This high degree of similarity indicates that the protein domains of the *S. typhimurium* homologues are likely to be very similar to those seen in *E. coli*. Thus, they all have catalytic domains, containing the catalytic triad of serine proteases. Like HtrA and HhoA in *E. coli*, STHtrA and STHhoA contain two PDZ domains, whereas STHhoB contains a single PDZ domain (as seen in *E. coli* HhoB, Figure 3.3). With such a high degree of homology, it is possible that the *S. typhimurium* homologues function in a similar manner as in *E. coli*.

Three HtrA homologues were identified in the unfinished genome of *S. aureus*, and were consequently named SAHtrA1, SAHtrA2 and SAHtrA3 (Figure 3.5). Like the Gram-negative HtrA homologues, these also contained a catalytic domain containing the conserved catalytic triad of serine proteases. However, SAHtrA1 and SAHtrA2 only contained a single PDZ domain, whereas in SAHtrA3 this was completely absent. The presence of a single PDZ domain has also been observed in all HtrA homologues identified in Gram-positive bacteria.

In *E. coli* HtrA, the PDZ domains are involved in the oligomerisation of the HtrA multimer (Sassoon, N. *et al.*, 1999). This finding was supported in the recently

published crystal structure of *E. coli* HtrA, which showed that the interactions between the two staggered trimeric rings were predominantly due to the PDZ domains (Krojer, T. *et al.*, 2002). The PDZ domains were also suggested to be involved in recognition of HtrA substrates as it is well documented that these domains can bind to C-terminal sequences and also are capable of binding to other PDZ domains (discussed in 1.7). To date, the oligomeric nature of the Gram-positive HtrA homologues have not been investigated. It is possible that Gram-positive HtrA monomers associate as trimers, thus, removing the need for the PDZ domains to stabilise the staggered ring formation seen in *E. coli* HtrA. If this were the case, the primary role of the single PDZ domain in Gram-positive bacteria would be to recognise substrates for proteolysis.

Comparisons between the *S. aureus* HtrA homologues and those of Gram-negative bacteria, such as *E. coli* and *S. typhimurium*, show that the N-terminal domains differ significantly between the two species. The crystal structure of *E. coli* HtrA has shown that the N-terminal domain, previously thought to be a subcellular location determinant, was actually part of the catalytic domain (Krojer, T. *et al.*, 2002). It was shown that the N-terminal region has a role in determining the distance between the trimeric rings and thus, determines the size of the lateral cavity of HtrA, which is lined with the proteolytic sites. Thus, it may be the N-terminal domain and not the PDZ domains that have the primary role in oligomerisation in Gram-positive bacteria. If this were the case, then in SAHtrA2 this cavity would be much larger, as the N-terminal of SAHtrA2 is significantly longer than that seen in STHtrA and SAHtrA1. However, the oligomeric form of the SAHtrAs are not known and there

are no crystal structures available. Therefore, further work is required to be able to determine the possible roles of the N-terminal domain and the PDZ domains.

8.3 HtrA homologues in *S. aureus*

8.3.1 SAHtrA1

HtrA has been shown to be important in stress conditions, such as heat-shock and oxidative stress. In *E. coli*, an extracytoplasmic σ factor, σ^E , controls the transcription of *htrA* (1.9.1). This σ factor is induced under environmental stresses, which result in an accumulation of aberrant proteins in the periplasmic compartment. In Northern blot analysis, *sahtrA1* was found to be transcribed as a monocistronic mRNA, which was not induced on exposure to heat-shock at 42°C (Figure 4.2b). The 5' end of the *sahtrA1* transcript was mapped by primer extension experiments and was found to begin 26-bp upstream of the predicted ATG site (Figure 4.3). Figure 4.4 shows that upstream of the transcription start site were sequence motifs at the –35 and –10 regions that were similar to sequences identified for σ^A , a vegetative σ factor in *S. aureus*. Taken together, this data support the hypothesis that *sahtrA1* is not induced under heat-shock conditions, but is instead expressed constitutively under these conditions. The *E. coli* HtrA homologues, HhoA and HhoB, are also expressed constitutively, thus, SAHtrA1 could be a homologue of HhoA or HhoB, and not HtrA. It is possible that although transcription is not induced under these conditions, translation of the *sahtrA1* transcript may occur under stress conditions. To investigate this further, pulse-chase experiments would need to be carried out.

To further characterise SAHtrA1, the gene was cloned and attempts were made to express recombinant proteins. Initial attempts to clone and express SAHtrA1 as a

native protein were partially successful as they only yielded a small amount of soluble protein (Figure 5.7 and Figure 5.8). SAHtrA1 was then expressed with a N-terminal 6xHis-tag. This protein was soluble and could be partially purified using a Ni^{2+} -column (Figure 5.9 and Figure 5.10). In order to assess if SAHtrA1 was indeed a serine protease, the partially purified protein was subjected to zymography. This technique allows a partially purified protein sample (when ran against a negative control sample) to be assessed for proteolytic activity. Zymography essentially separates protein samples by electrophoresis on a protein gel containing β -casein. The samples are electrophoresed under denaturing conditions due to the presence of SDS. The samples are then allowed to refold by washing the gels in Tween 20 and then incubated overnight in a Tris-buffer. Due to the presence of β -casein, the gels stained blue, when stained with Coomassie. However, where β -casein has been degraded by proteolysis, a clear band is seen in the gel. However, under the conditions used, SAHtrA1 was not found to possess proteolytic activity. As SAHtrA1 has been expressed as a recombinant protein, it is possible that the generated protein is not folded correctly and hence has no proteolytic activity. Alternatively, if SAHtrA1 was expressed in the correct tertiary structure, it is possible that the protein did not refold correctly in the renaturation step in this experiment. Finally, although *E. coli* HtrA is ATP-independent and cation-independent (Swarmy, K.H.S. *et al.*, 1983), SAHtrA1 activity may require ATP, divalent cations or other cofactors.

To investigate the properties of the protein domains of SAHtrA, the subdomains of SAHtrA1 were cloned and expressed as GST-fusion proteins. Expression of the catalytic domain (CAT1) was followed by Western blot analysis because the protein

could not be detected by SDS-PAGE and Coomassie-staining. Western blot analysis using anti-GST antibodies detected an immunoreactive band of the approximate correct size for GST-CAT1 and two smaller bands (Figure 5.11). As the smaller bands were also GST-tagged, it was possible that these were autoproteolytic products of the catalytic domain, as *E. coli* is known to be autoproteolytic (Skórko-Glonek, J. *et al.*, 1995, Lipinska, B. *et al.*, 1990). Thus, due to the unstable nature of the domain, further studies were abandoned.

In order to investigate the binding properties of the PDZ domain of SAHtrA1 (PDZ1), the domain was initially cloned and expressed with a N-terminal GST-tag. GST-PDZ1 was overexpressed in BL21(DE3)CodonPlus cells as a soluble protein (Figure 5.12). The recombinant protein was then isolated by affinity chromatography using GSTrap, which had an affinity to the GST-tag (Figure 5.13). The GST-tag was removed from the eluted GST-tagged species by thrombin (Figure 5.14) and PDZ1 was obtained with subsequent purification with FPLC (Figures 5.15 to 5.18). Circular dichroism experiments showed that PDZ1 had secondary structure, thus, the binding characteristics of this domain will be carried out using SPR. However, GST-fusion proteins do not give accurate SPR binding data and so the domain was also cloned and expressed with a N-terminal 6xHis-tag. This strategy produced soluble 6xHis-PDZ1 but in much lower yields than that seen with GST-PDZ1. The protein could be partially purified by Ni-NTA resin (Figure 5.19b), thus further work must be done for large-scale purification of this protein. Should the purified 6xHis-PDZ1 have a secondary structure, further work would be required to investigate what this domain can bind *in vivo*.

8.3.2 SAHtrA2

To test the hypothesis that SAHtrA2 was induced under heat-shock conditions, *S. aureus* was exposed to heat-shock at 42°C over a period of time. RNA from various time points were analysed by Northern blotting. Figure 4.2c shows that there is an increase in *sahtrA2* mRNA at fifteen minutes after transferral of the bacteria to heat-shock conditions. This transient increase is lost by thirty minutes although it can still be observed after ninety minutes. Thus, unlike *sahtrA1*, transcription of the *sahtrA2* mRNA is induced under heat-shock conditions suggesting that SAHtrA2 is the HtrA homologue in *S. aureus*. Unfortunately, attempts to map the 5' end of the transcript by primer extension were unsuccessful. This was likely to be due to the low abundance of *sahtrA2* mRNA used in the primer extension experiments.

In order to characterise *sahtrA2* further, the gene was cloned and expressed as a native protein. To do this, the gene was cloned in two fragments exploiting a unique *Cla*I site at position 856 in the gene to reconstitute the full-length ORF. These fragments were then ligated into the expression plasmid, pTrc99A. Interestingly, the *sahtrA2* gene in *S. aureus* SMITH strain differed from the unfinished genome of *S. aureus* NCTC 8325 strain (Figure 5.20) at four amino acids positions (I²⁸²→T, P³²¹→L, T³²⁶→I in the N-terminal domain and V⁶⁹⁶→A in the PDZ domain, where the first residue is from the NCTC strain residue). However, the sequence was identical to that seen in the recently published genomes of *S. aureus* N315 and Mu50 (Kuroda, M. *et al.*, 2001). The differences in the N-terminal domain are quite significant, as the properties of the amino acids have changed, for example the uncharged isoleucine residue has been changed to a polar threonine. The change of a valine to an alanine residue in the PDZ domain, however, is not significant as both

residues have a relatively small, uncharged sidechain. However, the genome of the NCTC strain is yet to be finished so these differences may be due to inaccuracies in the databases.

Expression of the SAHtrA2, however, did not yield a protein product of the correct size on SDS-PAGE protein gels. A smaller protein band of approximately 66 kDa was observed (Figure 5.21) but could not be verified to be part of SAHtrA2 as there were no antibodies available to detect this protein by Western blot analysis. Thus, the gene was recloned with a N-terminal 6xHis-tag and expressed in several bacterial expression strains. Although, Western blot analysis detected an immunoreactive band of the approximate size of SAHtrA2, a second heavier immunoreactive band was observed. The identity of this 6xHis-tagged species is unclear as the plasmid DNA sequence was correct and separation of the protein sample was under denaturing and reducing conditions, hence, there would not be any dimerisation occurring. Thus, further characterisation of SAHtrA2 was not undertaken.

Like SAHtrA1, SAHtrA2 also has a single PDZ domain. In order to understand the function of this domain in SAHtrA2, the protein domain was cloned and expressed as a GST-tagged protein in a similar manner to PDZ1. Like PDZ1, PDZ2 was expressed as a soluble protein and was subsequently purified using FPLC (Figure 5.23). Circular dichroism experiments indicated that PDZ2 also had secondary structure, and this protein was shown to bind the wild-type *S. aureus* SsrA-tag. Thus, the PDZ domain of SAHtrA2 appears to be involved in protein:protein interactions and therefore may be involved in substrate binding. As the GST-moiety is thought to influence data obtained from SPR experiments, the protein was expressed with a N-

terminal 6xHis-tagged. Expression of 6xHis-PDZ2 was achieved and the protein was partially purified using Ni-NTA resin (Figure 5.24). For future biochemical characterisation of the PDZ2, large-scale purification studies are underway.

8.4 HtrA homologues in *S. typhimurium*

In order to evaluate the function of the *htrA* homologues in bacterial viability and virulence, it was necessary to construct mutant strains. It is much harder to generate *S. aureus* mutants, therefore, mutants of the *htrA* family were constructed in *S. typhimurium* for analysis. The mutants were generated by allelic replacement using a modification of ET-cloning techniques (Yu, D. *et al.*, 2000, Zhang, Y. *et al.*, 2000). Mutants were obtained by replacing the gene of interest with an antibiotic resistance cassette and the resulting mutant colonies were verified by colony PCR (Figures 7.2 to 7.4) and subsequent DNA sequencing. In this way, *htrA::KAN*, *hhoA::KAN* and *hhoB::KAN* were obtained by allelic replacement using the kanamycin resistance cassette. In addition, *hhoA* and *hhoB* lie adjacent on the genome and this was exploited to obtain a *hhoAB::KAN* mutant in a single round of allelic replacement. To obtain the remaining mutants, two rounds of allelic replacement were required. For this, the *htrA* gene was first replaced by a gentamycin resistance cassette. A second round of allelic replacement was carried out to replace *hhoA*, *hhoB*, and *hhoAB* with the kanamycin cassette.

In vitro studies were carried out to examine the effect of stress conditions, such as heat-shock and oxidative stress, on the viability of these mutants. Growth profiles were obtained by measuring the culture density at OD₆₀₀ throughout a 16 hr time course (Figures 7.5 to 7.7). At 37°C, it was found that mutants containing the *hhoB*

mutation had a longer lag-time than that seen in the other mutant strains and the parent strain. In addition, mutants with the *hhoB* mutation also grew at a slower rate suggesting that under normal growth conditions *hhoB* is important in bacterial growth. The growth of mutants missing both *hhoB* and *htrA* was significantly slower than that seen in the single *htrA* and *hhoB* mutants. This suggests that these two genes function together in a manner which has yet to be identified.

When exposed to sublethal heat-shock conditions mutants containing the *hhoB* mutation had a more pronounced lag-time. Again, this was seen particularly in mutants containing both the *htrA* and *hhoB* mutations. In addition, the *htrA::KAN*, and *htrA::GM hhoA::KAN* mutants reached a plateau at an earlier time point of 650 minutes compared to 800 minutes at 37°C. Viable counts of the mutants after 16 hr incubations found that at 42°C there was a significant decrease in the viability of all the mutant strains when compared with the control conditions of 37°C. Incubation at 42°C had little effect on the viability of the parent SL3261 strain. Johnson *et al* (1991) reported that HtrA was not involved in heat-shock survival in *S. typhimurium*. However, this study suggests that HtrA is involved in heat-shock as both the growth profiles and the viable counts of the *htrA* mutant were significantly lower than the parent strain. In addition, the viability of the *htrA::GM hhoB::KAN* mutant was significantly lower than any of the other mutant strains examined. This supports the observations seen in the growth profiles which suggests that these two genes may function together in some important role in bacterial physiology. Interestingly, mutants with the *hhoA* mutation had higher viability counts than the respective single and double mutants. This is illustrated by the ability of the triple mutant having a higher viability than the *htrA::GM hhoB::KAN* double mutant. This

suggests that mutating the *hhoA* gene has a positive effect on bacterial growth and viability.

The effect of oxidative stress on bacterial growth was examined by exposing the mutant strains to H₂O₂, a source of hydroxyl radicals. Growth profiles were produced by incubating the bacterial strains with a range of H₂O₂ concentrations and measuring the culture density at OD₆₀₀ over a period of 16 hr (Figures 7.9 to 7.10). Figure 7.11 shows that in the presence of 1 mM H₂O₂, mutants with the *hhoB* mutation were affected more than the other mutant strains. Conversely, mutants with the *hhoA* mutation affected less than the other strains. This supports the observations from the heat-shock studies, which also suggest that mutating the *hhoA* gene has a positive effect on bacterial growth and viability. Although a decrease in culture density was observed for the *htrA::KAN* single mutant this was not deemed significant when compared to the parent strain. This data addresses bacterial growth on exposure to H₂O₂ by measuring the OD₆₀₀, and not bacterial viability. Unfortunately, repeated attempts to investigate the viability of the mutants under exposure to H₂O₂ were unsuccessful. Further investigations are needed to order to determine the effect of oxidative stress on bacterial viability of the mutant strains.

In vivo studies were carried out to examine the role of the *htrA* gene family and found that all the mutant strains, except the *hhoA::KAN* mutant, were highly attenuated in a mouse model. The *hhoA::KAN* strain was not attenuated and instead proliferated in a similar manner to that seen by the parent strain, SL1344. The *htrA* mutants did not proliferate but clearance from the liver and spleen were much slower than double and triple mutants strains containing the *htrA* mutation. The *hhoA::KAN*

and *hhoAB::KAN* mutants were cleared from the liver and spleen from infected mice faster than *htrA::KAN*. In addition, the double *htrA::GM hhoB::KAN* mutant and the triple mutant were cleared much faster from the liver and spleen than all the other mutant strains. The reduced virulence observed in the mutants with the *hhoB* mutation may be explained by the slower growth rates observed *in vitro*. The inability to proliferate and infect the mice would increase the chances of the murine immune system to clear these strains. Contrary to observations seen in *in vitro* stress experiments, double and triple mutants with the *hhoA* mutation did not appear to have a positive affect on bacterial virulence. In fact, although the virulence of the single *hhoA::KAN* mutant did not seem affected, a slight decrease in virulence was observed in the *hhoA* double mutants. At a higher inoculum dose, the triple mutant was more attenuated than the *htrA::GM hhoB::KAN* mutant suggesting that mutating the *hhoA* gene did have an effect on bacterial virulence using these conditions. The more complex environment in which the bacteria will be subjected to in an *in vivo* system can explain differences between the *in vitro* and *in vivo* observations. It is possible that *hhoA* has a function in bacterial virulence, which has, as yet, not been determine under *in vitro* conditions. Thus, further analysis is needed to characterise the function of the HtrA homologues in *S. typhimurium*.

In order to biochemically characterise the HtrA homologues of *S. typhimurium*, *hhoA* and *hhoB* were cloned and expressed as native and 6xHis-tagged proteins. The 6xHis-STHhoA was expressed as a smaller protein than expected. However, it was noted that the 6xHis moiety was lost indicating the removal of the N-terminus of the protein. *E. coli* HhoA is known to undergo signal peptide cleavage at the N-terminus (Bass, S. *et al.*, 1996, Waller, P.R.H. & Sauer, R.T., 1996). Analysis of the STHhoA

protein sequence indicated that there was a peptide cleavage site at a similar position to that of *E. coli* HhoA, and removal of this predicted signal peptide resulted in a protein with the same molecular mass as seen on the protein gels. Both STHhoA and STHhoB were expressed as soluble proteins, which could be partially purified by Ni-NTA resin (Figures 6.5 and 6.6). The partially purified homologues were then assessed for protease activity by zymography. Under the experimental conditions used, protease activity was observed for STHhoA (Figure 6.7a) but not STHhoB (Figure 6.7b and 6.7c). The lack of protease activity could be due to the experimental conditions, as these may not have been optimal to detect protease activity. The use of protease inhibitors (Figure 6.8) suggests that STHhoA is a serine protease as metalloprotease, aspartate protease and cysteine protease inhibitors have no effect on the degradation of β -casein. The serine proteases, aprotonin and AEBSF, however did reduce the amount of β -casein degradation, although this was not completely abolished. Abolishment of the signal may require higher concentrations of serine protease inhibitors, thus further work is needed to experimentally verify that STHhoA is a serine protease, although it does contain the catalytic triad typical of serine proteases.

8.5 Conclusions and future work

A family of HtrA homologues have been identified in the Gram-positive bacteria, *S. aureus* and the Gram-negative bacteria, *S. typhimurium*. All these homologues contained the catalytic triad of serine proteases in addition to one or two PDZ domains. Three HtrA homologues were identified in *S. aureus* and two of them, SAHtrA1 and SAHtrA2, were studied in this project. Transcription of *sahtrA2* was found to be induced under heat-shock conditions, whereas, transcription of *sahtrA1*

was constitutively expressed and possess promoter sites similar to those seen for the vegetative σ factor, σ^A . The HtrA homologues in *S. typhimurium* were homologues of *E. coli* HtrA, HhoA and HhoB. In order to investigate the role of STHtrA, STHhoA and STHhoB in bacterial viability and virulence, mutants were constructed. This study suggests that STHhoB has a role in bacterial growth at 37°C and may interact with STHtrA in an unknown manner. *In vitro* stress conditions also suggest that *hhoB* and *htrA* may interact in an unknown manner as strains with mutations in both genes have a significant reduction in bacterial viability. The importance of these genes in bacterial virulence was illustrated by the attenuation of all mutant strains, except for *hhoA::KAN*, which had a similar profile to the wild-type strain SL1344. In *E. coli*, there is another periplasmic protease, Prc, which is involved in the heat-shock response. Thus, it would be interesting to investigate the effect in bacterial viability and virulence when mutating this gene alongside the *htrA* gene family.

Further analysis of the HtrA homologues in both bacterial species will be aided with the recombinant proteins produced during this study. Large-scale purification of the recombinant proteins will greatly assist in further characterisation of this family of proteins. The protease activity of these proteins can also be assessed as only STHhoA was found to have protease activity in this study. This study shows that the PDZ domain of SAHtrA2 is involved in protein:protein interactions, in this case with the SsrA-tag of *S. aureus* and thus, may be involved in substrate binding interactions. However, further characterisation of the binding properties of the HtrA homologues can be performed using purified PDZ domains and the full-length proteins. These experiments will help to understand the function of these homologues in the bacterial system. Crystallisation studies of all these homologues

will illustrate any differences in structure between the homologues and thus, bring an understanding of the purpose of each protein.

The identification of HtrA homologues in *S. aureus* has raised many questions. The subcellular location of the *S. aureus* homologues is intriguing due to the lack of a periplasmic compartment in Gram-positive bacteria. Determining the location of these proteins would shed light on the function of HtrA in Gram-positive species. Finally, the construction of HtrA mutants in *S. aureus* will be useful to study the role of these homologues in the biology of *S. aureus*.

REFERENCES

- Abo, T., K. Ueda, T. Sunohara, K. Ogawa and H. Aiba (2002). "SsrA-mediated protein tagging in the presence of miscoding drugs and its physiological role in *Escherichia coli*." *Genes to Cells* 7: 629-638.
- Ades, S. E., L. E. Connolly, B. M. Alba and C. A. Gross (1999). "The *Escherichia coli* σ^E -dependent extracytoplasmic stress response is controlled by the regulated proteolysis of an anti- σ factor." *Genes & Development* 13: 2449-2461.
- Alba, B. M., H. J. Zhong, J. C. Pelayo and C. A. Gross (2001). "*degS* (*hhoB*) is an essential *Escherichia coli* gene whose indispensable function is to provide σ^E activity." *Molecular Microbiology* 40: 1323-1333.
- Alberts, B., D. Bray, L. Julian, M. Raff, K. Roberts and J. D. Watson (1994). *Molecular Biology of the Cell*, Garland Publishing, Inc. New York & London.
- Anderson, B., D. Jones and A. Burgess (1996). "Cloning, expression and sequence analysis of the *Bartonella henselae* gene encoding the HtrA stress-response protein." *Gene* 178: 35-38.
- Bakker, D., C. E. Vader, B. Roosendaal, F. R. Mooi, B. Oudega and F. K. de Graaf (1991). "Structure and function of periplasmic chaperone-like proteins involved in the biosynthesis of K88 and K89 fimbriae in enterotoxigenic *Escherichia coli*." *Molecular Microbiology* 5: 875-886.
- Bardwell, J. C., K. McGovern and J. Beckwith (1991). "Identification of a protein required for disulfide bond formation *in vivo*." *Cell* 67.
- Bass, S., Q. Gu and A. Christen (1996). "Multicopy suppressors of Prc mutant *Escherichia coli* include two HtrA (Deg P) protease homologs (HhoAB), DksA, and a truncated RlpA." *Journal of Bacteriology* 178: 1154-1161.

Bäumler, A. J., J. G. Kusters, I. Stojiljkovic and F. Heffron (1994). "*Salmonella typhimurium* loci involved in survival withing macrophages." Infection and Immunity **62**: 1623-1630.

Bezprozvanny, I. and A. Maximov (2001). "Classification of PDZ domains." FEBS Letters **509**: 457-462.

Bianchi, A. A. and F. Baneyx (1999). "Hyperosmotic shock induces the σ_{32} and σ_E stress regulons of *Escherichia coli*." Molecular Microbiology **34**: 1029-1038.

Bukau, B. and A. Horwich (1998). "The Hsp70 and Hsp60 chaperone machines." Cell **92**: 351-366.

Cabral, J. H. M., C. Petosa, M. J. Sutcliffe, S. Raza, O. Byron, F. Poy, S. M. Marfatia, A. H. Chishti and R. C. Liddinton (1996). "Crystal structure of a PDZ domain." Nature **382**: 649-652.

Cavard, D., C. Lazdunski and S. P. Howard (1989). "The acylated precursor of the colicin A lysis protein is a natural substrate of the DegP protease." Journal of Bacteriology **171**: 6316-6322.

Chan, P. F. and S. J. Foster (1998). "The role of environmental factors in the regulation of virulence-determinant expression in *Staphylococcus aureus* 8325-4." Microbiology **144**: 2469-2479.

Chan, P. F., S. J. Foster, E. Ingham and M. O. Clements (1998). "The *Staphylococcus aureus* alternative sigma factor σ^B controls the environmental stress response but not starvation survival or pathogenicity in a mouse abscess model." Journal of Bacteriology **180**: 6082-6089.

Cheung, A. L., K. J. Eberhardt and V. A. Fischetti (1994). "A method to isolate RNA from Gram-positive bacteria and mycobacteria." Analytical Biochemistry **222**: 511-514.

Clements, M. O. and S. J. Foster (1999). "Stress resistance in *Staphylococcus aureus*." *Trends in Microbiology* 7: 458-462.

Collier, J., E. Binet and P. Boulloc (2002). "Competition between SsrA tagging and translational termination at weak stop codons in *Escherichia coli*." *Molecular Microbiology* 45: 745-754.

Cortés, G., B. de Astorza, V. J. Benedí and S. Albertí (2002). "Role of the *htrA* gene in *Klebsiella pneumoniae* virulence." *Infection and Immunity* 70: 4772-4776.

Danese, P. N. and T. J. Silhavy (1997). "The σ^E and the Cpx signal transduction systems control the synthesis of periplasmic protein-folding enzymes in *Escherichia coli*." *Genes & Development* 11: 1183-1193.

Danese, P. N., W. B. Synder, C. L. Cosma, L. J. Davis and T. J. Silhavy (1995). "The Cpx two-component signal transduction pathway of *Escherichia coli* regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP." *Genes & Development* 9: 387-398.

Daniels, D. L., A. R. Cohen, J. M. Anderson and A. T. Brünger (1998). "Crystal structure of the hCASK PDZ domain reveals the structural basis of class II domain target recognition." *Nature Structural Biology* 5: 317-325.

Dartigalongue, C. and S. Raina (1998). "A new heat-shock gene, *ppiD*, encodes a peptidyl-isomerase required for folding of outer membrane proteins in *Escherichia coli*." *EMBO Journal* 17: 3968-3980.

de la Cruz, J. and A. Vioque (2001). "Increased sensitivity to protein synthesis inhibitors in cells lacking tmRNA." *RNA* 7: 1708-1716.

De Las Peñas, A., L. Connolly and C. A. Gross (1997). "The σ^E -mediated response to extracytoplasmic stress in *Escherichia coli* is transduced by RseA and RseB, two negative regulators of σ^E ." *Molecular Microbiology* **24**: 373-385.

Deora, R. and T. K. Misra (1996). "Characterization of the primary σ factor of *Staphylococcus aureus*." *Journal of Biological Chemistry* **271**: 21828-21834.

Diaz-Torres, M. L. and R. R. B. Russell (2001). "HtrA protease and processing of extracellular proteins of *Streptococcus mutans*." *FEMS Microbiology Letters* **204**: 23-28.

Doyle, D. A., A. Lee, J. Lewis, E. Kim, M. Sheng and R. MacKinnon (1996). "Crystal structure of a complexed and peptide-free membrane protein-binding domain: Molecular basis of a peptide recognition by PDZ." *Cell* **85**: 1067-1076.

Erickson, J. W. and C. A. Gross (1989). "Identification of the σ^E subunit of RNA polymerase: a second alternate σ factor involved in high-temperature gene expression." *Genes and Development* **3**: 1462-1471.

Fuh, G., M. T. Pisabarro, Y. Li, C. Quan, L. A. Lasky and S. S. Sidhu (2000). "Analysis of PDZ domain-ligand interactions using carboxyl-terminal phage display." *Journal of Biological Chemistry* **275**: 21486-21491.

Gee, S. H., S. Quenneville, C. R. Lombardo and J. Chabot (2000). "Single-amino acid substitutions alter the specificity and affinity of PDZ domains for their ligands." *Biochemistry* **39**: 14638-14646.

Gillet, R. and B. Felden (2001). "Emerging views on the tmRNA-mediated protein tagging and ribosome rescue." *Molecular Microbiology* **42**: 879-885.

Glazebrook, J., A. Ichige and G. C. Walker (1996). "Genetic analysis of *Rhizobium meliloti* *bacA-phoA* fusion results in identification of *degP*: two loci required for symbiosis are closely linked to *degP*." *Journal of Bacteriology* **178**: 745-752.

Gottesman, S., E. Roche, Y. Zhou and R. T. Sauer (1998). "The ClpXP and ClpAP proteases degrade proteins with the carboxy-terminal peptide tails added by the SsrA-tagging system." Genes & Development **12**: 1338-1347.

Hara, H., Y. Yamamoto, A. Higashitani, H. Suzuki and Y. Nishimura (1991). "Cloning, mapping and characterisation of the *Escherichia coli* *prc* gene, which is involved in C-terminal processing of penicillin-binding protein 3." Journal of Bacteriology **173**: 4799-4813.

Hayes, C., B. Bose and R. T. Sauer (2002). "Stop codons preceded by rare arginine codons are efficient determinants of SsrA tagging in *Escherichia coli*." Proceedings of the National Academy of Sciences of the United States of America **99**: 3440-3445.

Herman, C., D. Thévenet, P. Boulloc and G. C. Walker (1998). "Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH)." Genes & Development **12**: 1348-1355.

HersHKovits, G., Z. Dubinsky and D. J. Katcoff (1997). "A novel homologue of the prokaryotic *htrA* gene is differentially expressed in the alga *Haematococcus pluvialis* following stress." Molecular and General Genetics **254**: 345-350.

Hu, L. A., Y. Tang, W. E. Miller, M. Cong, A. G. Lau, R. J. Lefkowitz and R. A. Hall (2000). " β 1-Adrenergic receptor association with PSD-95." Journal of Biological Chemistry **275**: 38659-38666.

Huang, C., M. C. Wolfgang, J. Withey, M. Koomey and D. I. Freidman (2000). "Charged tmRNA but not tmRNA-mediated proteolysis is essential for *Neisseria gonorrhoeae* viability." EMBO Journal **19**: 1098-1107.

Johnson, K., I. Charles, G. Dougan, D. Pickard, P. O'Gaora, G. Costa, T. Ali and C. Hormaeche (1991). "The role of a stress-response protein in *Salmonella typhimurium* virulence." Molecular Microbiology **5**: 401-407.

Jones, C. H., T. C. Bolken, K. F. Jones, G. O. Zeller and D. E. Hruby (2001). "Conserved DegP protease in Gram-positive bacteria is essential for thermal and oxidative tolerance and full virulence in *Streptococcus pyogenes*." *Infection and Immunity* **69**: 5538-5545.

Julio, S. M., D. M. Heithoff and M. J. Mahan (2000). "*ssrA* (tmRNA) plays a role in *Salmonella enterica* serovar Typhimurium pathogenesis." *Journal of Bacteriology* **182**: 1558-1563.

Kahane, S., Y. Weinstein and I. Sarov (1990). "Cloning, characterization and sequence of a novel 59-kDa protein of *Chlamydia trachomatis*." *Gene* **90**: 61-67.

Karthikeyan, S., T. Leung, G. Birrane, G. Webster and J. A. A. Ladas (2001). "Crystal structure of the PDZ1 domain of human Na⁺/H⁺ exchanger regulatory factor provides insights into the mechanism of carboxyl-terminal leucine recognition by Class I PDZ domains." *Journal of Molecular Biology* **308**: 963-973.

Karzai, A. W. and R. T. Sauer (2001). "Protein factors associated with the SsrA-SmpB tagging and ribosome rescue complex." *Proceedings of the National Academy of Sciences of the United States of America* **98**: 3040-3044.

Karzai, A. W., M. M. Susskind and R. T. Sauer (1999). "SmpB, a unique RNA-binding protein essential for the peptide-tagging activity of SsrA (tmRNA)." *EMBO Journal* **18**: 3796-3799.

Keiler, K. C., L. Sharpiro and K. P. Williams (2000). "tmRNAs that encode proteolysis-inducing tags are found in all known bacterial genomes: A two-piece tmRNA functions in *Caulobacter*." *Proceedings of the National Academy of Sciences of the United States of America* **97**: 7778-7783.

Keiler, K. C., P. R. Waller and R. T. Sauer (1996). "Role of a peptide tagging system in degradation of proteins synthesized from damage messenger RNA." *Science* **271**: 990-993.

Kim, J.-H. and M.-J. Hahn (2000). "Cloning and sequencing of the gene encoding the candidate HtrA of *Rickettsia typhi*." *Microbiology and Immunology* **44**: 275-278.

Kim, K. I., S.-C. Park, S. H. Kang, G.-W. Cheong and C. H. Chung (1999). "Selective degradation of unfolded proteins by the self-compartmentalizing HtrA protease, a periplasmic heat shock protein in *Escherichia coli*." *Journal of Molecular Biology* **294**: 1363-1374.

Kolmar, H., P. R. H. Waller and R. T. Sauer (1996). "The DegP and DegQ periplasmic endoproteases of *Escherichia coli*: Specificity for cleavage sites and substrate conformation." *Journal of Bacteriology* **178**: 5925-5929.

Komine, Y., M. Kitabatake, T. Yokogawa, K. Nishikawa and H. Inokuchi (1994). "A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*." *Proceedings of the National Academy of Science USA* **91**: 9223-9227.

Krojer, T., M. Garrido-Franco, R. Huber, M. Ehrmann and T. Clausen (2002). "Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine." *Nature* **416**: 455-459.

Kuroda, M., T. Ohta, I. Uchiyama, T. Baba, H. Yuzawa, I. Kobayashi, L. Cui, A. Oguchi, K.-i. Aoki, Y. Nagai, J. Lian, T. Ito, M. Kanamori, H. Matsumaru, A. Maruyama, H. Murakami, A. Hosoyama, Y. Mizutani-Ui, N. K. Takahashi, T. Sawano, R.-i. Inoue, C. Kaito, K. Sekimizu, H. Hirakawa, S. Kuhara, S. Goto, J. Yabuzaki, M. Kanehisa, A. Yamashita, K. Oshima, K. Furuya, C. Yoshino, T. Shiba, M. Hattori, N. Ogasawara, H. Hayashi and K. Hiramatsu (2001). "Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*." *The Lancet* **357**: 1225-1240.

Ladbury, J. E., M. A. Lemmon, M. Zhou, J. Green, M. C. Botfield and J. Schlessinger (1995). "Measurement of the binding of tyrosyl phosphopeptides to SH2 domains: A reappraisal." *Proceedings in the National Academy of Sciences of the United States of America* **92**: 3149-3155.

Laport, M. S., A. C. D. de Castro, A. Villardo, J. A. C. Lemos, M. d. C. F. Bastos and M. Giambiagi-deMarval (2001). "Expression of the major heat shock proteins DnaK and GroEL in *Streptococcus pyogenes*: A comparison to *Enterococcus faecalis* and *Staphylococcus aureus*." Current Microbiology **42**: 264-268.

Laskowska, E., D. Kuczuynksa-Wisnik, J. Skorko-Glonek and A. Taylor (1996). "Degradation by proteases Lon, Clp and HtrA, of *Escherichia coli* proteins aggregated *in vivo* by heat shock; HtrA protease action *in vivo* and *in vitro*." Molecular Microbiology **22**: 555-571.

Li, S.-R., N. Dorrell, P. H. Everest, G. Dougan and B. W. Wren (1996). "Construction and characterisation of a *Yersinia enterocolitica* O:8 high-temperature requirement (*htrA*) isogenic mutant." Infection and Immunity **64**: 2088-2094.

Lin-Chao, S., C.-L. Wei and Y.-T. Lin (1999). "RNase E is required for the maturation of *ssrA* RNA and normal *ssrA* RNA peptide-tagging activity." PNAS **96**: 12406-12411.

Lipinska, B., O. Fayet, L. Baird and C. Georgopoulos (1989). "Identification, characterisation, and mapping of the *Escherichia coli* *htrA* gene, whose product is essential for bacterial growth only at elevated temperatures." Journal of Bacteriology **171**: 1574-1584.

Lipinska, B., M. Zylicz and C. Georgopoulos (1990). "The HtrA (DegP) protein, essential for *Escherichia coli* survival at high temperatures, is an endopeptidase." Journal of Bacteriology **172**: 1791-1797.

Loosmore, S. M., Y.-P. Yang, R. Oomen, J. M. Shortreed, D. C. Coleman and M. H. Klein (1998). "The *Haemophilis influenzae* HtrA protein is a protective antigen." Infection and Immunity **66**: 899-906.

Malenka, R. C. and R. A. Nicoll (1999). "Long-term potentiation - a decade of progress." Science **285**: 1870-1874.

Mandal, P. K. (2002). "Complete NMR spectroscopic assignment of a neuronal transduction protein." *Monatshefte für Chemie* **133**: 205-217.

Manna, A. C., M. G. Bayer and A. L. Cheung (1998). "Transcriptional analysis of different promoters in the *sar* locus in *Staphylococcus aureus*." *Journal of Bacteriology* **180**: 3828-3836.

Mecasas, J., P. E. Rouviere, J. W. Erickson, T. J. Donohue and C. A. Gross (1993). "The activity of σ^E , an *Escherichia coli* heat-inducible σ -factor, is modulated by expression of outer membrane proteins." *Genes & Development* **7**: 2618-2628.

Missiakas, D., M. P. Mayer, M. Lemaire, C. Georgopoulos and S. Raina (1997). "Modulation of the *Escherichia coli* σ^E (RpoE) heat-shock transcription factor activity by RseA, RseB and RseC proteins." *Molecular Microbiology* **24**: 355-371.

Missiakas, D. and S. Raina (1997). "Signal transduction pathways in response to protein misfolding in the extracytoplasmic compartments of *E. coli*: role of two new phosphoprotein phosphatases PrpA and PrpB." *EMBO Journal* **16**: 1670-1685.

Muto, A., A. Fujihara, K.-i. Ito, J. Matsuno, C. Ushida and H. Himeno (2000). "Requirement of transfer-messenger RNA for the growth of *Bacillus subtilis* under stress." *Genes to Cells* **5**: 627-635.

Muyrers, J. P. P., Y. Zhang, G. Testa and A. F. Stewart (1999). "Rapid modification of bacterial artificial chromosomes by ET recombination." *Nucleic Acids Research* **27**: 1555-1557.

Naberhaus, F., W. Weiglhofer, H.-M. Fischer and H. Hennecke (1998). "Identification of the *Bradyrhizobium japonicum* *degP* gene as part of an operon containing small heat-shock protein genes." *Archives of Microbiology* **169**: 89-97.

Noone, D., A. Howell, R. Collery and K. M. Devine (2001). "YkdA and YvtA, HtrA-like serine proteases in *Bacillus subtilis* engage in negative autoregulation and reciprocal cross-regulation of *ykdA* and *yvtA* gene expression." *Journal of Bacteriology* **183**: 654-663.

Noone, D., A. Howell and K. M. Devine (2000). "Expression of *ykdA*, encoding a *Bacillus subtilis* homologue of HtrA is heat shock inducible and negatively autoregulated." *Journal of Bacteriology* **182**: 1592-1599.

Novick, R. P. (1991). "Genetic systems in Staphylococci." *Methods in Enzymology* **204**: 587-636.

Ohta, T., K. Saito, M. Kuronda, K. Honda, H. Hirata and H. Hayashi (1994). "Molecular cloning of two heat shock genes related to the *hsp70* genes in *Staphylococcus aureus*." *Journal of Bacteriology* **176**: 4779-4783.

Oschkinat, H. (1999). "A new type of PDZ domain recognition." *Nature Structural Biology* **6**: 408-410.

Pallen, M. J. and C. P. Ponting (1997). "MicroCorrespondence: PDZ domains in bacterial proteins." *Molecular Microbiology* **26**: 411-415.

Pallen, M. J. and B. W. Wren (1997). "The HtrA family of serine proteases." *Molecular Microbiology* **26**: 209-221.

Pedersen, L. L., M. Radulic, M. Doric and Y. A. Kwaik (2001). "HtrA homologue of *Legionella pneumophila*: an indispensable element for intracellular infection of mammalian but not protozoan cells." *Infection and Immunity* **69**: 2569-2579.

Phillips, R. W. and R. M. Roop, II (2001). "*Brucella abortus* HtrA functions as an authentic stress response protease but is not required for wild-type virulence in BALB/c mice." *Infection and Immunity* **69**: 5911-5913.

Piserchio, A., M. Pellegrin, S. Mehta, S. M. Blackman, E. P. Garcia, J. Marshall and D. F. Mierke (2002). "The PDZ1 domain of SAP90 - characterisation of structure and binding." *Journal of Biological Chemistry* **277**: 6967-973.

Pogliano, J., S. A. Lynch, D. Belin, E. C. C. Lin and J. Beckwith (1997). "Regulation of *Escherichia coli* cell envelope proteins involved in protein folding and degradation by the Cpx two-component system." *Genes & Development* **11**: 1169-1182.

Poquet, I., A. Bolotin and A. Gruss (2001). "Optimising the production of heterologous exported proteins in *Lactococcus lactis* by inactivation of HtrA, the unique housekeeping surface protease." *LAIT* **81**: 37-47.

Poquet, I., V. Saint, E. Seznec, N. Simoes, A. Bolotin and A. Gruss (2000). "HtrA is the unique surface housekeeping protease in *Lactococcus lactis* and is required for natural protein processing." *Molecular Microbiology* **35**: 1042-1051.

Qoronfleh, M. W., C. A. Bortner, P. Schwarzborg and B. J. Wilkinson (1998). "Enhanced levels of *Staphylococcus aureus* stress protein GroEL and DnaK homologs early in infection of human epithelial cells." *Infection and Immunity* **66**: 3024-3027.

Raina, S., D. Missiakas and C. Georgopoulos (1995). "The rpoE gene encoding the σ^E (σ^{24}) heat shock sigma factor of *Escherichia coli*." *EMBO Journal* **14**: 1043-1055.

Resto-Ruiz, S. I., D. Sweger, R. H. Widen, N. Valkov and B. E. Anderson (2000). "Transcriptional activation of the *htrA* (High-Temperature Requirement A) gene from *Bartonella henselae*." *Infection and Immunity* **68**: 5970-5978.

Roche, E. D. and R. T. Sauer (1999). "SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity." *EMBO Journal* **18**: 4579-4589.

Roche, E. D. and R. T. Sauer (2001). "Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons." *Journal of Biological Chemistry* **276**: 28509-28515.

Rouvière, P. E., A. D. L. Peñas, J. Mecsas, C. Z. Lu, K. E. Rudd and C. A. Gross (1995). "*rpoE*, the gene encoding the second heat-shock sigma factor, σ^E , in *Escherichia coli*." *EMBO Journal* **14**: 1032-1042.

Sassoon, N., J.-P. Arié and J.-M. Betton (1999). "PDZ domains determine the native oligomeric structure of the DegP (HtrA) protease." *Molecular Microbiology* **33**: 583-589.

Schurr, M. J., H. Yu, J. M. Martinez-Salazar, J. C. Boucher and V. Deretic (1996). "Control of AlgU, a member of the σ^E -like family of stress sigma factors, by the negative regulators MucA and MucB and *Pseudomonas aeruginosa* conversion to mucoidy in cystic fibrosis." *Journal of Bacteriology* **178**.

Sebert, M. E., L. M. Palmer, M. Rosenberg and J. N. Weiser (2002). "Microarray-based identification of *htrA*, a *Streptococcus pneumoniae* gene that is regulated by the CiaRH two-component system and contributes to nasopharyngeal colonization." *Infection and Immunity* **70**: 4059-4067.

Seol, J. H., S. K. Woo, E. M. Jung, S. J. Yoo, C. S. Lee, K. Kim, K. Tanaka, A. Ichihara, D. B. Ha and C. H. Chung (1991). "Protease Do is essential for survival of *Escherichia coli* at high temperatures: its identity with the *htrA* gene product." *Biochemical and Biophysical Research Communications* **176**: 730-736.

Skórko-Glonek, J., K. Krzewski, B. Lipinska, E. Bertoli and F. Tanfani (1995). "Comparison of the structure of wild-type HtrA heat shock protease and mutant HtrA proteins. A Fourier transform infrared spectroscopic study." *Journal of Biological Chemistry* **270**: 11140-11146.

Skórko-Glonek, J., A. Wawrzynów, K. Krzewski, K. Kurpierz and B. Lipinska (1995). "Site-directed mutagenesis of the HtrA (DegP) serine protease, whose proteolytic activity is indispensable for *Escherichia coli* survival at elevated temperatures." *Gene* **163**: 47-52.

Skórko-Glonek, J., D. Zurawa, E. Kuczwara, M. Wozniak, Z. Wypych and B. Lipinska (1999). "The *Escherichia coli* heat shock protease HtrA participates in defence against oxidative stress." *Molecular Genes and Genetics* **1999**: 342-350.

Smeds, A., P. Varmenen and A. Palva (1998). "Molecular characterization of a stress-inducible gene from *Lactobacillus helveticus*." *Journal of Bacteriology* **180**: 6148-6153.

Smith, C. K., T. A. Baker and R. T. Sauer (1999). "Lon and Clp family proteases and chaperones share homologous substrate-recognition domains." *Proceedings of the National Academy of Sciences of the United States of America* **96**: 6678-6682.

Songyang, Z., A. S. Fanning, C. Fu, J. Xu, S. M. Marfatia, A. H. Chishi, A. Crompton, A. C. Chan, J. M. Anderson and L. C. Cantley (1997). "Recognition of unique carboxyl-terminal motifs by distinct PDZ domains." *Science* **275**: 73-77.

Speiss, C., A. Beil and M. Ehrmann (1999). "A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein." *Cell* **97**: 339-347.

Spiers, A., H. K. Lamb, S. Cocklin, K. A. Wheeler, J. Budworth, A. L. Dodds, M. J. Pallen, D. J. Maskell, I. G. Charles and A. R. Hawkins (2002). "PDZ domains facilitate binding of High-temperature requirement protease A (HtrA) and tail-specific protease (Tsp) to heterologous substrates through recognition of the Small Stable RNA A (*ssrA*)-encoded peptide." *Journal of Biological Chemistry* **277**: 39443-39449.

St Geme, J. W., II and S. Grass (1998). "Secretion of the *Haemophilus influenzae* HMW1 and HMW2 adhesins involves a periplasmic intermediate and requires the HMWB and HMWC proteins." *Molecular Microbiology* **27**: 617-630.

Steiner, K. and H. Malke (2001). "*relA*-Independent amino acid starvation response network of *Streptococcus pyogenes*." *Journal of Bacteriology* **183**: 7354-7364.

Strauch, K. L. and J. Beckwith (1988). "An *Escherichia coli* mutation preventing degradation of abnormal periplasmic proteins." *Proceedings of the National Academy of Science USA* **85**: 1576-1580.

Strauch, K. L., K. Johnson and J. Beckwith (1989). "Characterization of *degP*, a gene required for proteolysis in the cell envelope and essential for growth of *Escherichia coli* at high temperature." *Journal of Bacteriology* **171**: 2689-2696.

Swarmy, K. H. S., C. J. Chung and A. L. Goldberg (1983). "Isolation and characterization of Protease Do from *Escherichia coli*, a large serine protease containing multiple subunits." *Archives of Biochemistry and Biophysics* **224**: 543-554.

Swarmy, K. H. S. and A. L. Goldberg (1981). "*E. coli* contains eight soluble proteolytic activities, one being ATP dependent." *Nature* **292**: 652-654.

Thao, M. L. and P. Baumann (1998). "Sequence analysis of a DNA fragment from *Buchnera aphidicola* (aphid endosymbiont) containing the genes *dapD-htrA-ilvH-ftsL-ftsI-murE*." *Current Microbiology* **37**: 214-216.

Tochio, H., F. Hung, M. Li, D. S. Bredt and M. Zhang (2000). "Solution structure and backbone dynamics of second PDZ domain of post-synaptic density-95." *Journal of Molecular Biology* **295**: 225-237.

Tochio, H., Q. Zhang, P. Mandal, M. Li and M. Zhang (1999). "Solution structure of the extended neuronal nitric oxide synthase PDZ domain complexed with an associated peptide." *Nature Structural Biology* **6**: 417-421.

Tu, G.-F., G. E. Reid, J.-G. Zhang, R. L. Moritz and R. J. Simpson (1995). "C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide." *Journal of Biological Chemistry* **270**: 9322-9326.

Waller, P. R. H. and R. T. Sauer (1996). "Characterisation of *degQ* and *degS*, *Escherichia coli* genes encoding homologs of the DegP protease." *Journal of Bacteriology* **178**: 1146-1153.

Walma, T., C. A. E. Spronk, M. Tessari, J. Aelen, J. Schepens, W. Hendricks and G. Vuister (2002). "Structure, dynamics and binding characteristics of the second PDZ domain of PTP-BL." *Journal of Molecular Biology* **316**: 1101-1110.

Wang, R. F. and J. M. Kaguni (1989). "A novel sigma factor is involved in expression of the *rpoH* gene of *Escherichia coli*." *Journal of Bacteriology* **171**: 4248-4253.

Wiegert, T. and W. Schumann (2001). "SsrA-mediated tagging in *Bacillus subtilis*." *Journal of Bacteriology* **183**: 3885-3889.

Williams, K., P. C. F. Oyston, N. Dorrell, S.-R. Li, R. W. Titball and B. W. Wren (2000). "Investigation into the role of the serine protease HtrA in *Yersinia pestis* pathogenesis." *FEMS Microbiology Letters* **186**: 281-286.

Yamamoto, T., T. Hanawa, S. Ogata and S. Kamiya (1996). "Identification and characterization of the *Yersinia enterocolitica* *grsA* gene, which protectively responds to intracellular stress induced by macrophage phagocytosis and to extracellular environmental stress." *Infection and Immunity* **64**: 2980-2987.

Yamamoto, Y., N. Inagaki and K. Satoh (2001). "Overexpression and characterization of the C-terminal processing protease for precursor D1 protein - Regulation of enzyme-substrate interaction by molecular environments." *Journal of Biological Chemistry* **276**: 7518-7525.

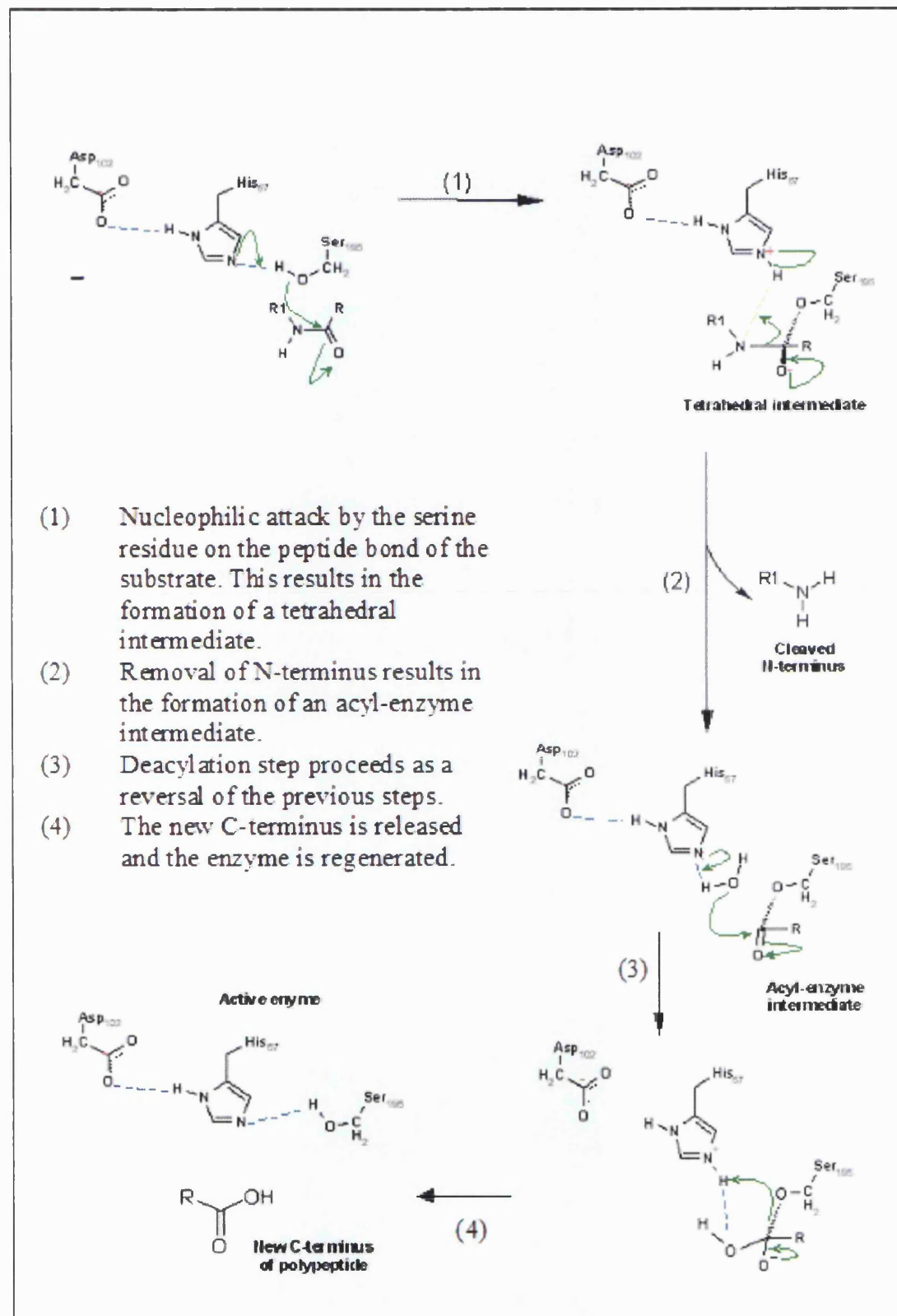
Yu, D., H. M. Ellis, N. A. Jenkins, N. G. Copeland and D. L. Court (2000). "An efficient recombination system for chromosome engineering in *Escherichia coli*." *Proceedings of the National Academy of Sciences USA* **97**: 5978-5983.

Zhang, Y., J. P. P. Muirers, G. Testa and A. F. Stewart (2000). "DNA cloning by homologous recombination in *Escherichia coli*." *Nature Biotechnology* **18**: 1314-1317.

APPENDIX

A.1	Mechanism of serine proteases: the catalytic triad.....	272
A.2	Primers used for the production of recombinant proteins	273
A.3	Bacterial strains used.....	274
A.4	Primers used for primer extension experiments.....	274
A.5	Primers used for sequencing reactions	275
A.6	Primers used for making DNA probes for Northern blot analysis.....	276
A.7	Primers used for constructing linear DNA fragment for allelic replacement	276
A.8	Primers for colony PCR to verify allelic replacement	277
A.9	Sequenced ORF of <i>sahtrA1</i>	278
A.10	Sequenced ORF of <i>sahtrA2</i>	279
A.11	Sequenced ORF of <i>sthhoA</i>	280
A.12	Sequenced ORF of <i>sthhoB</i>	281

A.1 Mechanism of serine proteases: the catalytic triad



A.2 Primers used for the production of recombinant proteins

Primers	Primer sequence	Encoded restriction sites
EMR1	CCCGGGGATACATGTCAGATTTTAATC	<i>Afl</i> III
EMR2	GTACTTTTCTGTCAAGCTTTTATCTAAAGAA	<i>Hind</i> III
EMR3	CATATAATGGAGGGAATCCATGTCAGATTTTAATC	<i>Bam</i> HI
EMR4	CATATAATGGAGGATCCTATGTCAGATTTTAATC	<i>Bam</i> HI
EMR5	TTCAATTGTTACCTCGAGTTAATTACTTGAATAG	<i>Xho</i> I
EMR6	TTTGCTATTCCAGGATCCGAAGTTAAAGTAAC	<i>Bam</i> HI
EMR7	GTACTTTTCTGTCTCGAGTTATCTAAAGAA	<i>Xho</i> I
EMR8	TTTGCTATTCCGGATCCTGAAGTTAAAGTAAC	<i>Bam</i> HI
EMR9	GATAAGTGATTTTATTGCCATGGAGTGGAC	<i>Nco</i> I
EMR10	ATTGACACGTTGTCTAGATCAACATTATTTAG	<i>Xba</i> I
EMR11	CGTCAACAATATGATTTTCTCTCATGACG	N/A
EMR12	GGTGATAGCGAACAAAATGACAAGTC	N/A
EMR13	GATAAGTGATTTGGATCCATGGAGTGGAC	<i>Bam</i> HI
EMR14	ATTGACACGTTGCCATGGTCAACATTATTTAG	<i>Nco</i> I
EMR15	GATAAGTGATTTGGATCCGATGGAGTGGACATTAG	<i>Bam</i> HI
EMR16	CATTTGCAATACCTGGATCCGAAGTACAAAG	<i>Bam</i> HI
EMR17	ATTGACACGTTGCCCCGGTCAACATTATTTTAG	<i>Sma</i> I
EMR18	CATTTGCAATACCGGATCCTGAAGTACAAAG	<i>Bam</i> HI
EMR19	GTACGAGAGCAGGATCCATGAAAAAACAC	<i>Bam</i> HI
EMR20	CGGTTTAAGAAGCTTTTAGCGCAGCAATAG	<i>Hind</i> III
EMR21	AGGTACGAGAGCGGATCCCATGAAAAAACAC	<i>Bam</i> HI
EMR22	GCAGCATCCATGGAAAAACACACCC	<i>Nco</i> I
EMR23	GCCCCGTTTAAGGGATCCTTAGCGCAGCAATAG	<i>Bam</i> HI
EMR24	TTTAACGACGCCGGATCCATGTTTGTGAAGC	<i>Bam</i> HI
EMR25	TTTTAACGACGCGGATCCCATGTTTGTGAAGC	<i>Bam</i> HI
EMR26	GGCAAAAGCCGCCTGCAGTTAGTTCGACGC	<i>Pst</i> I
EMR27	ACGACGCCTCCACCATGGTTGTGAAGCTC	<i>Nco</i> I

A.3 Bacterial strains used

Bacterial strain	Genotype
<i>E. coli strains</i>	
BL21	F <i>ompT hsdS_B(r_B⁻m_B⁻) gal dcm</i>
BL21 Star	F <i>ompT hsdS_B(r_B⁻m_B⁻) gal dcm rne131(DE3)</i>
BL21(DE3)	F <i>ompT hsdS_B(r_B⁻m_B⁻) gal dcm (DE3)</i>
BL21(DE3)pLysS	F <i>ompT hsdS_B(r_B⁻m_B⁻) gal dcm (DE3) pLysS</i>
BL21(DE3)pLysS Star	F <i>ompT hsdS_B(r_B⁻m_B⁻) gal dcm rne131(DE3) pLysS (Cam^R)</i>
DH5 _α	F ϕ 80dlacZΔM15 Δ(<i>lacZYA-argF</i>)U169 <i>deoR recA1 endA1 hsdR17(r_K⁻m_K⁻) phoA supE44 λ⁻ thi-1 gyrA96 relA1</i>
JM109	<i>endA1 recA1 gyrA96 hsdR17 (r_K⁻m_K⁻) relA1 supE44 Δ(lac-proAB) [F' traD36 proAB lacI^qΔM15)</i>
SG1009	<i>Nal^S str^S rif^S thi⁻ lac⁻ ara⁺ gal⁺ mtl⁻ recA⁺ uvr⁺ lon⁺</i>
M15	
<i>S. typhimurium strains</i>	
SL1344	<i>hisG46</i> virulent mouse strain
SL3261	<i>hisG46 aroA554</i> attenuated mouse strain
<i>S. aureus strain</i>	
8325 PS47 (36) 8409 *47	Possesses no resistance to antibiotics

A.4 Primers used for primer extension experiments

Primer	Sequence	Used for:
PEXT-13	CCATGGAAATTTAGGTCTTCTG	<i>sahtrA1</i>
PEXT-15	CAATGCAACGATGACTGTTTAAACC	<i>sahtrA1</i>
PEXT-21A	CCAATATCCACTAATGTCCACTCC	<i>sahtrA2</i>

A.5 Primers used for sequencing reactions

Gene	Primer	Sequence
<i>sahtrA1</i>	ALLN1A-SEQF0	TCTGTCTCCCCTACAATTGTTGGAG
	ALLN1A-SEQF1	CGATAGCGTATTCGCAATGGGTAAC
	ALLN1-SEQF1	CTGGTGGCAATACAAAAGTTAGC
	ALLN1A-SEQF2	CGCCAAAGCTGATAGTGATATTGATC
	ALLN1-SEQF2	GAACAACCAAAACGTCAAAGCCG
	ALLN1A-SEQF3	CACGTGCTTGTCGTTTAGAAGGAG
	ALLN1A-SEQF4	CAGCATCTAACCAACAGCACCTG
	ALLN1-SEQR1	GAAGCACCGCATTTACGTGAAGCTC
	ALLN1-SEQR2	CTATCGCCTTGGCCAGGTGATTAC
	ALLN1-SEQR3	GAGTTACCTGGGTTAATAGCAGC
<i>sahtrA2</i>	ALLN2-SEQF1	GAATCAACTGAAATGAGAAAGCTAAG
	ALLN2-SEQF2	CAGATGACGATCGAAAACCTTGATG
	ALLN2-SEQF3	GATAAAGCATCTCAAGACGAAGTGGG
	ALLN2-SEQR1	CTCTATTTGACGACAGCACCACCCG
	ALLN2-SEQR2	CGACCGATTAACTGTGTTATTGGC
<i>hhoA</i>	DEGQ-SEQF1	ACGCAAATTGCCATCGCCGATTCC

Plasmid	Primer	Sequence
pGEX-4T-1	pGEX-FOR	GGGCTGGCAAGCCACGTTTGGTG
	pGEX-REV	CCGGGAGCTGCATGTGTCAGAGG
pTrc99A	pTrc-FOR	TTGACAATTAATCATCCGGCTCG
	pTrc-REV	GTGCGTTCTGATTAACTGTATC
pGEM-T	pGEM-SEQR1	GCATCCAACGCGTTGGGAGCTCTC
pQE-30	pQE-SEQFOR	GTGAGCGGATAACAATTTACAG
	pQE-SEQREV	GTTCTGAGGTCATTACTGGATCTATC
pRSET-B	pRSET-SEQR	CTAGTTATTGCTCAGCGGTGGCAG

A.6 Primers used for making DNA probes for Northern blot analysis

Probe		Primers		
PDZ1		EMR-6		EMR-7
PDZ2		EMR-16		EMR-17
DnaK	FOR	GGAGAAACTCAAGT	REV	CACCGGCTGTTGATAGTAC
		TGGTGAAG		TTCG
23SrRNA	FOR	GCTTGCTTGTTGGGG	REV	GTCTACGACCAAATACTA
		TTGTAGGACAC		AACGCCC

A.7 Primers used for constructing linear DNA fragment for allelic replacement

Primer	Sequence	Used for
AR-1	ATAAAATGAATCTGACGTACACAGCAATTTTGCGTTACCTGTTA ATCGAGATTGAAACACgcccgcgtcccgtcaagtcagcg	<i>htrA::KAN</i>
AR-2	TGGGGGGTTTCACAGAAAAGTGTTGCCCCCTTCCATGGCGGAAG GGGGACAAAGGTGAAttccccggatccgtcgacctgcag	<i>htrA::KAN</i>
AR-3	ATAAAATGAATCTGACGTACACAGCAATTTTGCGTTACCTGTTA ATCGAGATTGAAACACctcgaattgacataagcctg	<i>htrA::GM</i>
AR-4	TGGGGGGTTTCACAGAAAAGTGTTGCCCCCTTCCATGGCGGAAG GGGGACAAAGGTGAgaagccgatctcgcttgaac	<i>htrA::GM</i>
AR-5	ATAAAATGAATCTGACGTACACAGCAATTTTGCGTTACCTGTTA ATCGAGATTGAAACACgcggtatttaacgacctg	<i>htrA::CAT</i>
AR-6	TGGGGGGTTTCACAGAAAAGTGTTGCCCCCTTCCATGGCGGAAG GGGGACAAAGGTGAaccatcatacactaatcag	<i>htrA::CAT</i>
AR-7	TTAAGTGTGCGGTAACTCTTTTCGGCGCCGTTTCCAGCCCTTGCA TCGATACCAGGCCAGgcccgcgtcccgtcaagtcagcg	<i>hhoA::KAN</i> <i>hhoAB::KAN</i>
AR-8	CATACCATGAGTTAACCGGACATCACATCGCAGGCTGATGCCCCG GTTTAAGACAGGATTAccccggatccgtcgacctgcagg	<i>hhoA::KAN</i>
AR-9	TGATGTCCGGTTAACTCATGGTATGCTGCTGCCGTTCCCTTTTTT AACGACGCCTCCATCgcccgcgtcccgtcaagtcagcg	<i>hhoB::KAN</i>
AR-10	GCAGGGCCAGTCTTTATTGACTAGTCGTCGAACGACGCGGGCAA AAGCCGCGTCTGTTTTAccccggatccgtcgacctgcagg	<i>hhoB::KAN</i> <i>hhoAB::KAN</i>

A.8 Primers for colony PCR to verify allelic replacement

Primer	Sequence
ARKT-1	ATTCCTGTTTGTAATTGTCC
ARKT-2	ATCCATGTTGGAATTTAATCGCGG
ARH-1	GACTGACCTTTACGCATGGG
ARH-2	CACAAAACGGTGCAAATAAGTAT
ARA-1	TTAAGTGTCGGGTAACTCT
ARA-2	CATACCATGAGTTAACCGGA
ARB-1	TGATGTCCGGTTAACTCATG
ARB-2	GCAGGGCCAGTCTTTATTGA

Mutation		Primers	
<i>htrA::KAN</i>	5'	ARKT-1	ARH-1
	3'	ARKT-2	ARH-2
<i>hhoA::KAN</i>	5'	ARKT-1	ARA-1
	3'	ARKT-2	ARA-2
<i>hhoB::KAN</i>	5'	ARKT-1	ARB-1
	3'	ARKT-2	ARB-2
<i>hhoAB::KAN</i>	5'	ARKT-1	ARA-1
	3'	ARKT-2	ARB-2
<i>htrA::GM</i>		ARH-1	ARH-2
<i>htrA::CAT</i>		ARH-1	ARH-2

A.9 Sequenced ORF of *sahtrA1*

```

sahtrA1 1   ATGTCAGATT TTAATCATAC AGATCATTCT ACAACAAACC ATAGCCAAAC ACCTAGATAC AGAAGACCTA 70
sahtrA1 71   AATTTCCATG GTTTAAAACA GTCATCGTTG CATTGATTGC TGGAATTATT GGTGCACTTC TAGTACTTGG 140
sahtrA1 141  TATAGGCAAA GTATTAAATA GTACAATTTT AAATAAAGAT GGTTCAACTG TTCAGACAAAC AAATAATAAA 210
sahtrA1 211  GGTGGCAATC AATTAGACGG TCAAAGCAAG AAATTCGGTA CCGTTCATGA AATGATAAAA TCTGTCTCCC 280
sahtrA1 281  CTACAATTGT TGGAGTTATT AACATGCAAA AAGCATCAAG TGTAGACGAC TTATTTAAAG GCAAATCATC 350
sahtrA1 351  TAAACCATCT GAAGCTGGAG TAGGTTCAGG TGTATCTAT CAAATAAACA ACAATTCAGC TTATATCGTT 420
sahtrA1 421  ACAAAACAAT ATGTTATTGA TGGCGCAAAT GAAATTAGAG TCCAATTACA TAATAAAAAA CAAGTTAAAG 490
sahtrA1 491  CGAAATTAGT TGGTAAAGAT GCAGTAACTG ATATTGCTGT ACTTAAATTT GAAAATACAA AAGGTATTAA 560
sahtrA1 561  AGCGATTCAA TTTGCCAACT CTTCAAAAGT ACAAACTGGC GATAGCGTAT TCGCAATGGG TAACCCATTA 630
sahtrA1 631  GGATTACAAT TTGCTAACTC TGAACATCT GGTATCATT CAGCAAGCGA ACGTACGATT GACGCTGAGA 700
sahtrA1 701  CAACTGGTGG CAATACAAAA GTTAGCGTTC TTCAAACAGA TGCTGCTATT AACCCAGGTA ACTCAGGTGG 770
sahtrA1 771  CGCATTAGTA GATATTAATG GTAATTTAGT TGGTATTAAC TCAATGAAAA TTGCTGCGAC ACAAGTTGAA 840
sahtrA1 841  GGTATCGGGT TTGCTATTCC AAGTAATGAA GTTAAAGTAA CAATTGAACA ACTTGTAATA CATGGTAAAA 910
sahtrA1 911  TTGACCGCCC TTCGATTGGT ATTGGTTTAA TTAATTTGAA AGATATTCCT GAAGAAGAGC GCGAGCAACT 980
sahtrA1 981  TCATACTGAT AGAGAAGACG GTATTTATGT CGCCAAAGCT GATAGTGATA TTGATCTTAA AAAAGGTGAT 1050
sahtrA1 1051 ATTATTACAG AAATTGATGG CAAGAAAAAT AAAGATGATG TTGATTTAAG AAGCTATTTA TATGAAAATA 1120
sahtrA1 1121 AAAAACTGG TGAATCAGTC ACTGTTACCG TTATCCGTGA TGGTAAAACA AAAGAAGTTA AAGTGAAATT 1190
sahtrA1 1191 AAAACAACAA AAAGAACAAC CAAAACGTCA AAGCCGATCA GAACGTCAAT CACCTGGCCA AGGCGATAGA 1260
sahtrA1 1261 GATTTCTTTA GATAA 1275

```

A.10 Sequenced ORF of *sahtrA2*

```

sahtrA2 1   ATGGAGTGGA CATTAGTGGA TATTGGTAAA AAACATGTAA TTCCTAAAAG TCAGTACCGA CGTAAGCGTC 70
sahtrA2 71   GTGAATTCTT CCACAACGAA GACAGAGAAG AAAATTTAAA TCAACATCAA GATAAACAAA ATATAGATAA 140
sahtrA2 141  TACAACATCA AAAAAAGCAG ATAAGCAAAT ACATAAAGAT TCAATTGATA AGCACGAACG TTTTAAAAAT 210
sahtrA2 211  AGTTTATCAT CGCATTTAGA ACAGAGAAAC CGTGATGTTA ATGAGAATAA AGCTGAAGAA AGTAAAAGTA 280
sahtrA2 281  ATCAGGATAG TAAGTCAGCA TATAACAGAG ATCATTATTT AACAGACGAT GTATCTAAAA AACAAAATTC 350
sahtrA2 351  ATTAGATTCA GTGGACCAAG ATACAGAGAA ATCAAAATAT TATGAGCAAA ATTCTGAAGC GACTTTATCA 420
sahtrA2 421  ACTAAATCAA CCGATAAAGT AGAATCAACT GAAATGAGAA AGCTAAGTTC AGATAAAAAAC AAAGTTGGTC 490
sahtrA2 491  ATGAAGAGCA ACATGTACTT TCTAAACCTT CAGAACATGA TAAAGAGACT AGAATTGATT CTGAGTCTTC 560
sahtrA2 561  AAGAAGTGAT TCAGACAGCT CGATGCAGAC AGAGAAAATA AAAAAAGACA GTTCAGATGG AAATAAAAGT 630
sahtrA2 631  AGTAATCTGA AATCTGAAGT AATATCAGAC AAATCAAATA CAGTACCAAA ATTGTCGGAA TCTGATGATG 700
sahtrA2 701  AAGTAAATAA TCAGAAGCCA TTAACCTTAC CGGAAGAACA GAAATTGAAA AGACAGCAAA GTCAAAATGA 770
sahtrA2 771  GCAAACAAAA ACCTATACAT ATGGTGATAG CGAACAAAAT GACAAGTCTA ATCATGAAAA TGATTTAAGT 840
sahtrA2 841  CATCATACAC CATCGATAAG TGATGATAAA GATAACGTTA TGAGAGAAAA TCATATTGTT GACGATAATC 910
sahtrA2 911  CTGATAATGA TATCAATACA CTATCATTAT CAAAAATAGA TGACGATCGA AAACCTTGATG AAAAAATTCA 980
sahtrA2 981  TGTGAAGAT AAACATAAAC AAAATGCAGA CTCGTCTGAA ACGGTGGGAT ATCAAAGTCA GTCAACTGCA 1050
sahtrA2 1051 TCTCATCGTA GCACTGAAAA AAGAAATATT TCTATTAATG ACCATGATAA ATTAACGGT CAAAAACAA 1120
sahtrA2 1121 ATACAAAGAC ATCGGCAAAT AATAATCAAA AAAAGGCTAC ATCAAAATTG AACAAAGGCG GCGCTACGAA 1190
sahtrA2 1191 TAATAATTAT AGTGACATTT TGA AAAAGTT TTGGATGATG TATTGGCTA AATTAGTTAT TCTAATGGGT 1260
sahtrA2 1261 ATTATTATTC TAATTGTTAT TTTGAATGCC ATTTTAAATA ATGTGAACAA AAATGATCGC ATGAATGATA 1330
sahtrA2 1331 ATAATGATGC AGATGCTCAA AAATATACGA CAACGATGAA AAATGCCAAT AACACAGTTA AATCGGTGCT 1400
sahtrA2 1401 TACAGTTGAA AATGAACAT CAAAAGATTG ATCATTACCT AAAGATAAAG CATCTCAAGA CGAAGTGGGA 1470
sahtrA2 1471 TCAGGTGTTG TATATAAAAA ATCTGGAGAT ACGTTATATA TTGTTACGAA TGCACACGTT GTCGGTGATA 1540
sahtrA2 1541 AAGAAAATCA AAAAAAATCT TTCTCGAATA ATAAAAGTGT TGTGGGAAA GTGCTTGGA AAGATAAATG 1610
sahtrA2 1611 GTCAGATTTA GCTGTTGTTA AAGCAACTTC TTCAGACAGT TCAGTGAAAG AGATAGCTAT TGGAGATTCA 1680
sahtrA2 1681 AATAATTTAG TGTTAGGAGA GCCAATATTA GTCGTAGGTA ATCCACTTGG TGTAGACTTT AAAGGCACTG 1750
sahtrA2 1751 TGACAGAAGG TATTATTTCA GGTCTGAACA GAAATGTTCC TATTGATTTT GATAAAGATA ATAAATATGA 1820
sahtrA2 1821 TATGTTGATG AAAGCTTTCC AAATTGATGC ATCAGTAAAT CCAGGTAAC TCGGTGGTGC TGTCGTCAAT 1890
sahtrA2 1891 AGAGAAGGAA AATTAATAGG TGTAGTTGCA GCTAAAATTA GTATGCCAAA CGTTGAAAT ATGTCAATTG 1960
sahtrA2 1961 CAATACCTGT TAATGAAGTA CAAAAGATTG TAAAAGATTT AGAAACAAAA GGTA AAATTG ACTATCCCGA 2030
sahtrA2 2031 TGTAGGTGTT AAAATGAAGA ATATTGCCAG TCTAAATAGT TTTGAAAGAC AAGCAGTTAA ATTGCCAGGA 2100
sahtrA2 2101 AAAGTTAAGA ACGGTGTTGT TGTAGATCAA GTTGACAACA ATGGTTTAGC AGATCAATCT GGTCTGAAAA 2170
sahtrA2 2171 AAGGTGATGT AATTACTGAA TTAGATGGCA AACTTTTAGA AGATGATTTA CGCTTTAGGC AGATTATATT 2240
sahtrA2 2241 TAGTCATAAA GATGACTTGA AATCAATTAC AGCGAAGATT TATAGAGATG GTAAAGAGAA AGAAATTAAT 2310
sahtrA2 2311 ATTAACTAA AATAA 2325

```

A.11 Sequenced ORF of *sthhoA*

```

sthhoA 1   ATGAAAAAAC ACACCCAGCT GTTAAGTGCA TTAGCGTTAA GTGTCGGGTT AACTCTTTTCG GCGCCGTTTC 70
sthhoA 71   CAGCCCTTGC ATCGATACCA GGCCAGGTGC CAGGCCAGGC GACGCTGCCA AGCCTTGCCC CTATGCTGGA 140
sthhoA 141  GAAAGTGCTG CCTGCTGTCG TCAGCGTAAA AGTCGAGGGA ACGGCCGCCC AGAGCCAAAA AGTGCCGGAG 210
sthhoA 211  GAGTTTAAAA AATTCTTTGG CGAGGATCTG CCAGACCAGC CGTCCAGCC GTTTGAAGGA CTCGGTTGG 280
sthhoA 281  GGGTGATTAT CGATGCCGCG AAAGGCTATG TATTAACCAA TAATCATGTG ATTAATCAGG CACAGAAGAT 350
sthhoA 351  CAGCATTCAA CTGAATGACG GACGCGAATT CGACGCGAAG CAGATCGGCG GCGACGACCA GAGCGATATC 420
sthhoA 421  GCTCTGTTAC AAATTCAGAA TCCAGCAAG TTAACGCAAA TTGCCATCGC CGATTCCGAC AAACCTCCGC 490
sthhoA 491  TCGGCGATTT CGCCGTGGCG GTCGGTAATC CGTTTGGTCT TGGACAAACC GCCACCTCCG GGATTATTTT 560
sthhoA 561  AGCGCTGGGA CGCAGCGGGC TTAATCTGGA AGGCTTGAG AACTTTATTC AAACCGATGC CTCTATTAAC 630
sthhoA 631  CGCGGCAACT CCGGCGGCGC GCTGCTTAAC CTGAACGCGC AGCTGATCGG GATTAATACC GCGATCCTCG 700
sthhoA 701  CGCCAGGGGG CGGGAGCATC GGCATTGGCT TTGCTATTCC TTCCAATATG GCGCAGACGC TGGCGCAGCA 770
sthhoA 771  GTTGATTCAG TTCGGCGAAA TCAAACGCGG ATTGCTGGGA ATTAAAGGCA CTGAAATGAC CGCTGATATC 840
sthhoA 841  GCCAAGGCAT TCAAACGAA CGTTCAGCGT GCGCTTTTG TCAGCGAGGT TTTACCCAAT TCAGGTTGCG 910
sthhoA 911  CGAAGGCCGG GGTGAAATCC GGAGACGTGA TTATCAGTCT TAACGGTAAG CAGCTGAATA GCTTTGCCGA 980
sthhoA 981  ACTGCGTTCA CGTATCGCCA CCACCGAACC GGGCACGAAA GTGAAGCTGG GCCTGCTGCG CGATGGTAAG 1050
sthhoA 1051 CCGCTGGAGG TGGAAATCAC GCTGGATTCC AATACCTCTT CTTCGCCCAG CGCCGAAATG ATCGCCCCGG 1120
sthhoA 1121 CGTTGCAAGG CGCAACGTTG AGCGACGGCC AACTGAAAGA CGGGACGAAA GCGGTTAAGG TTGATAGCGT 1190
sthhoA 1191 CGAAAAAAGC AGTCCTGCCG CGCAGGCCGG TTTGCAAAAA GATGATGTTA TCATCGGCGT TAATCGCGAT 1260
sthhoA 1261 CGCATCAGTT CTATCGCCGA AATGCGCAAA GTGATGGCGG CAAAACCGTC CATCATTGCT CTTAGGTTAG 1330
sthhoA 1331 TACGCGGCAA CGAGAACATT TATCTATTGC TCGCTAA 1368

```

A.12 Sequenced ORF of *sthoB*

```

sthoB 1   ATGTTTGTGA AGCTCTTACG TTCGGTCGCA ATAGSTTTAA TTGTGGGGGC TATTCTGTTG GCGGTGATGC 70
sthoB 71   CTTCTTTTGG CAAAATTAAT CCTATCGCGG TCCCGCAATT CGACAGTACC GATGAGACGC CAGCCAGTTA 140
sthoB 141  TAATTTTGGG GTTCGGCGCG CCGCGCCTGC CGTCGTCAAT GTCTATAACC GCAGTATGAA CAGTACCGCG 210
sthoB 211  CATAATCAAC TGGAGATCCG CACGCTGGGT TCCGGCGTGA TCATGGATCA ACGCGGTTAT ATTATTACCA 280
sthoB 281  ACAAACACGT GATTAACGAT GCGGATCAGA TTATCGTCGC GCTACAGGAT GGCGCGCTCT TTGAAGCGCT 350
sthoB 351  ACTGGTTGGC TCGGATTGCG TTACCGATCT GCGGCTGCTG AAGATCAACG CCACTGGCGG GCTGCCCTACC 420
sthoB 421  ATCCCGATTA ATACAAAGCG TACACCGCAT ATTGGCGACG TCGTACTGGC TATCGGCAAC CCATATAATC 490
sthoB 491  TGGGACAGAC CATCACCCAG GGGATCATCA GCGCAACGGG TCGTATCGGC CTGAACCCGA CGGGGCGACA 560
sthoB 561  GAATTTTCTC CAGACCGACG CCTCGATTAA CCACGGTAAT TCCGGCGGCG CGCTGGTCAA CTCGTTAGGC 630
sthoB 631  GAACTGATGG GGATCAACAC CCTCTCTTTT GATAAGAGTA ACGATGGCGA AACGCCGGAA GGCCTTGSTT 700
sthoB 701  TTGCGATTCC CTTCCAGCTA GCCACGAAAA TTATGGATAA GCTTATCCGC GACGGTCGCG TGATTGCGCG 770
sthoB 771  CTATATCGGT ATTGGCGGAC GAGAAATCGC GCCGCTGCAC GCGCAGCAGG GTAGCGGCAT GGACCCGATT 840
sthoB 841  CAGGGCATTG TCGTTAATGA AGTGACGCCA AACGGCCCCG CCGCGCTTGC CGGTATTGAG GTTAATGATT 910
sthoB 911  TGATTATTTT GGTCAATAAT AAACCCGCGG TGTCGCGCTT GGAGACGATG GATCAGGTGG CGGAAATCCG 980
sthoB 981  CCGGGGCTCC GTCATTCCGG TCGTGGTAAT GCGGGATGAT AAGCAGCTCA CGTTCCAGGT GACGGTGACG 1050
sthoB 1051 GAATACCCGG CGTCGAACTA A 1071

```