Characterisation of a Putative *agr* System in *Clostridium botulinum* and *Clostridium sporogenes*

Clare Marie Cooksley, B.Sc.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy, September 2007

Declaration

Unless otherwise acknowledged, the work presented in this thesis is my own. No part has been submitted for another degree in the University of Nottingham or any other institute of learning.

C. Cooksley

Clare Cooksley

September 2007

Table of Contents

List of Figures.....vii

List of Tables.....x

Abbreviations.....xii

Abstract.....xv

Acknowledgements.....xvi

1.0 Introduction.....1

1.1 Introduction to clostridia.....1

1.2 *Clostridium botulinum*.....1

1.2.1 Discovery and occurrence.....1

1.2.2 Toxin types.....2

1.2.3 Botulinum neurotoxin structure.....3

1.2.4 Mechanism of toxin action.....5

1.2.5 Types of botulism and epidemiology.....6

1.2.5.1 Foodborne botulism.....6

1.2.5.2 Wound botulism.....8

1.2.5.3 Infant botulism.....8

1.2.6 Treatment.....9

1.2.7 Therapeutics.....10

1.2.8 Bioterrorism.....10

1.2.9 Regulation of botulinum neurotoxin expression in *C. botulinum*.....11

1.3 *Clostridium sporogenes*.....12

1.4 Sporulation in clostridia.....13

1.5 Quorum sensing.....14

1.5.1 The principle of quorum sensing.....14

1.5.2 Quorum sensing in Gram-positive bacteria.....16

1.5.3 Inactivation of the *agr* system.....22

1.5.4 Quorum sensing in clostridia.....22

1.5.5 Evidence for *agr* in clostridia.....23

1.6 Mutational studies in clostridia.....24

1.7 Project aims.....28

2.0 Materials and Methods.....29

2.1 Media.....29

2.1.1 Luria-Bertani medium.....29

2.1.2 TYG medium.....29

2.1.3 Brain Heart Infusion (BHI) agar.....29

2.1.4 Spizizen's Minimal Salts Medium (SMM).....29

2.1.5 *B. subtilis* Competence Medium.....30

2.1.6 *B. subtilis* Transformation Medium.....30

2.1.7 5 X Minimal salts.....30

2.1.8 *B. subtilis* Minimal growth medium.....30

2.1.9 *B. subtilis* Starvation medium.....31

2.2 Supplements.....31

2.3 Growth Conditions.....31

2.4 Bacterial strains.....33

2.5 Plasmids.....34

2.6 Oligonucleotide primers.....38

2.6.1 Primers used in Chapter 3.....38

2.6.2 Primers used in Chapter 4.....40

2.6.3 Primers used in Chapter 5.....43

2.7 Preparation and manipulation of DNA.....46

2.7.1 Preparation of plasmid DNA.....46

2.7.2 Preparation of chromosomal DNA (*E. coli* and *C. sporogenes*).....46

2.7.3 Preparation of chromosomal DNA (*C. botulinum*).....46

2.7.4 Restriction enzyme digests.....47

2.7.5 Blunt-ending of DNA.....47

2.7.6 De-phosphorylation of linearised DNA fragments.....47

2.7.7 Analysis of DNA by agarose gel electrophoresis.....48

2.7.8 Purification of DNA fragments.....48

2.7.9 Polymerase Chain Reaction.....48

2.7.10 DNA ligation.....49
2.7.11 TOPO TA cloning methods.....49
2.7.12 Preparation of electrocompetent *E. coli* cells.....49
2.7.13 Electroporation.....50
2.7.14 Transformation of DNA into *Bacillus subtilis* (method 1).....50
2.7.15 Transformation of naturally competent *B. subtilis* cells with DNA (method 2).....50
2.7.16 Blue/white colony selection.....51
2.7.17 Conjugation of *E. coli* CA434 (HB101) with *C. sporogenes* 10696 (filter mating method).....51
2.7.18 Conjugation of *E. coli* CA434 (HB101) with *C. sporogenes* 10696 or *C. botulinum* Hall A (spot method).....51
2.7.19 DNA sequencing.....52
2.7.20 Southern hybridisation.....52
2.8 RNA analysis.....53
2.8.1 Preparation of total RNA from *C. sporogenes* (method 1).....54
***2.8.1.1 Removal of contaminating genomic DNA from total RNA*.....54**
***2.8.1.2 DNase treated RNA sample clean up*.....55**
2.8.2 Preparation of total RNA from *C. sporogenes* (method 2).....55
***2.8.2.1 RNase inhibitor/ DNase treatment*.....56**
***2.8.2.2 Sample clean up*.....56**
2.8.3 Quantification of RNA and DNA.....56
2.8.4 One-step Reverse Transcriptase (RT)-PCR.....57
2.8.5 Real Time Quantitative PCR.....57
***2.8.5.1 Primers*.....57**
***2.8.5.2 RNA extraction*.....57**
***2.8.5.3 Reverse transcription*.....58**
***2.8.5.4 Real time PCR*.....58**
***2.8.5.5 Standard curve generation*.....59**
***2.8.5.6 Data analysis*.....59**
2.8.6 Northern hybridisation.....59

2.9 Estimation of segregational stability of plasmids in *Bacillus subtilis* and *Clostridium* spp.....61

2.10 Induction of gene replacement using integrative vector in *Clostridium sporogenes*.....62

2.11 Induction of target gene knockout using the ClosTron System.....63

2.11.1 Using plasmid pMTL007.....63

2.11.2 Using plasmid pMTL007CE-2.....63

2.12 Measurement of spore formation in *Clostridium* species.....64

3.0 Identification and Characterisation of Clostridial *agr* Systems.....65

3.1 Introduction.....65

3.2 Results.....68

3.2.1 Sequencing of the *agr* regions of *C. sporogenes*.....68

3.2.2 Comparison of *agr* region gene products amongst clostridia and other Gram-positive bacteria.....69

***3.2.2.1 AgrD alignments*.....69**

***3.2.2.2 AgrB alignments*.....75**

***3.2.2.3 Differences between the two intraspecies AgrD and AgrB proteins in* C. botulinum *and* C. sporogenes.....76**

***3.2.2.4 Analysis of regions flanking* agrB *and* agrD *in clostridia*.....76**

3.2.3 BLAST searches on *agr* and flanking regions.....78

3.2.4 Expression analysis of *agr* genes in *C. sporogenes*.....80

***3.2.4.1 Reverse Transcriptase PCR analysis to investigate linkage of* agr *gene expression in* C. sporogenes.....80**

3.2.4.2 Real Time RT-PCR to analyse expression

***of* agrD*1 and* agrD*2 in* C. sporogenes.....83**

3.3 Discussion.....86

3.3.1 Sequencing of the *agr* regions of *C. sporogenes*.....86

3.3.2 Comparison of AgrD and AgrB amongst clostridia

and other Gram-positive bacteria.....86

3.3.3 Analysis of regions flanking *agrB* and *agrD* in clostridia.....89

3.3.4 Expression analysis of *agr* genes in *C. sporogenes*.....93

4.0 Gene Modulation Techniques in *Clostridium sporogenes*.....98

4.1 Introduction.....98

4.2 Unstable plasmids as a tool for gene knockout in clostridia.....98

4.2.1 Results.....101

4.2.1.1 Plasmid segregational stability in

Bacillus subtilis.....105

4.2.1.2 Plasmid segregational stability in

Clostridium sporogenes.....105

***4.2.1.3 Construction and insertion of a* pyrF**

knockout cassette for use with integrative vector

***pMTL5112lacI*.....107**

4.2.2 Discussion.....114

4.3 Gene modulation in *C. sporogenes* using antisense RNA

technology.....117

4.3.1 Antisense target gene selection for this study.....119

4.3.2 Results.....121

***4.3.2.1 Construction of antisense 'mutants'*.....121**

***4.3.2.2 Effect of antisense production on sporulation*.....124**

4.3.3 Discussion.....127

5.0 Gene inactivation in *C. botulinum* and *C. sporogenes* using the

ClosTron System.....131

5.1 Introduction.....131

5.1.1 Adaptation of the group II intron for use in *Clostridium* spp.....135

5.1.2 Generation of multiple gene knockout mutants using the ClosTron System.....137

5.2 Results.....143

5.2.1 Generation of single knockout mutants.....143

5.2.2 Screening of putative single knockout mutants.....146

5.2.3 Generation of putative double knockout mutants.....153

***5.2.3.1 Step 1 – Generation of* agrD*1 mutant*.....153**

***5.2.3.2 Step 2 – Construction of plasmid expressing* Flp *recombinase*.....156**

***5.2.3.3 Step 3 – Excision of removable RAM*.....157**

5.2.4 Characterisation of clostridial mutants generated using the ClosTron system.....158

5.3 Discussion.....162

5.3.1 Future work.....170

6.0 General Discussion.....174

6.1 Introduction.....174

6.2 A functional *agr* system in *C. botulinum* and *C. sporogenes*.....174

6.3 Gene modulation techniques for determining *agr* gene function.....176

6.4 Construction of targeted knockout mutants for determining function of the *agr* genes.....178

6.5 Summary and future work.....179

Bibliography.....182

Appendix 1.....198

Appendix 2.....201

Appendix 3.....219

Appendix 4.....220

Appendix 5.....222

List of Figures

Chapter 1.0

Figure 1.1 Arrangement of neurotoxin genes in different *C. botulinum* types.....4

Figure 1.2 Diagram of the *agr* quorum sensing system from *S. aureus*.....18

Figure 1.3 Structures of the AIPs from the four *S. aureus agr* groups.....21

Figure 1.4 Schematic representation of single crossover recombination.....26

Figure 1.5 Schematic representation of double crossover recombination.....27

Chapter 3.0

Figure 3.1 *agr* gene regions in four species of clostridia.....67

Figure 3.2 Phylogenetic tree of AgrD proteins of various Gram-positive bacteria.....71

Figure 3.3 AgrD protein alignment of various Gram-positive bacteria.....72

Figure 3.4 AgrD protein alignments.....73

Figure 3.5 Phylogenetic tree of AgrB proteins of various Gram-positive bacteria.....75

Figure 3.6 Protein alignment of regions immediately upstream of AgrB in *C. botulinum*, *C. sporogenes*, *C. perfringens* and *C. acetobutylicum*.....77

Figure 3.7 Protein alignment of regions upstream of AgrB in *C. perfringens* and *C. botulinum*.....78

Figure 3.8 Reverse transcriptase PCR to show linkage of *agr* gene expression in *C. sporogenes*.....81

Figure 3.9 Real Time RT-PCR to assess *agrD* expression in *C. sporogenes*.....85

Chapter 4.0

Figure 4.1 Replication in rolling circle plasmids.....99

Figure 4.2 Strategy used to generate integrative vector pMTL5112lacI.....102

Figure 4.3 Representation of the chromosomal arrangements which may arise from two possible single crossover events.....109

Figure 4.4 Plasmid map and screening results on SC3 DNA.....111

Figure 4.5 Proposed mechanisms of action of antisense RNA.....118

Figure 4.6 Schematic representation of the *agr* regions in the *C. sporogenes* 10696 genome.....119

Figure 4.7 Schematic representation of the sporulation phosphorylation cascade in *B. subtilis*.....120

Figure 4.8 Plasmid map of inducible antisense delivery vector pMTL9361.....121

Figure 4.9 Confirmation of antisense RNA expression in *agrB*1 and *agrB*2 'mutants'.....123

Figure 4.10 Confirmation of antisense RNA expression in 336 and 340 'mutants'.....123

Figure 4.11 Sporulation in *C. sporogenes agrB* 'mutants' in the presence and absence of IPTG.....125

Figure 4.12 Sporulation in *C. sporogenes* 336 and 340 kinase 'mutants' in the presence and absence of IPTG.....126

Chapter 5.0

Figure 5.1 Secondary structure and recognition sites of the *L. lactis ltrB* group II intron.....133

Figure 5.2 Plasmid map of the pMTL007 ClosTron vector.....136

Figure 5.3 Plasmid map of pMTL007C-E2.....138

Figure 5.4A SOEing PCR carried out on *L. lactis ltrB* group II intron.....139

Figure 5.4B Ligation of retargeted vector into pMTL007 plasmid.....140

Figure 5.4C Expression of intron RNA and retrohoming of RNP.....141

Figure 5.4D Splicing of the ErmRAM in ClosTron mutants.....142
Figure 5.5 Schematic representation of the *agr* regions in the *C. botulinum* ATCC 3502 genome.....143
Figure 5.6 Products obtained from the re-targeting SOEing PCR carried out on the *L. lactis* group II intron.....145
Figure 5.7 Schematic representation of PCR screens carried out on putative integrants.....148
Figure 5.8 PCR screen 1 on four *C. botulinum* and two *C. sporogenes* mutants.....149
Figure 5.9A Southern blot to demonstrate the presence of a single intron insertion in four *C. botulinum* mutants.....151
Figure 5.9B Southern blot to demonstrate the presence of a single intron insertion in two *C. sporogenes* mutants.....152
Figure 5.10A PCR screen 1 on *C. botulinum agrD*1 mutant generated using pMTL007C-E2.....155
Figure 5.10B PCR screen 2 on *C. botulinum agrD*1 mutant generated using pMTL007C-E2.....155
Figure 5.10C PCR screen 3 on *C. botulinum agrD*1 mutant generated using pMTL007C-E2.....156
Figure 5.11 Plasmid map of vector pMTL5402F-FLP.....157
Figure 5.12 Growth of six clostridial mutants compared to wild type.....159
Figure 5.13 Sporulation in four *C. botulinum* mutants.....160
Figure 5.14 Sporulation in two *C. sporogenes* mutants.....161
Figure 5.15 Functional domains of the C. botulinum kinase proteins 0336 and 0340.....167
Figure 5.16 Schematic diagram depicting the insertion site of the group II intron in the *C. botulinum agrD* mutants, along with the possible transcriptional fusions which may occur.....168

List of Tables

Chapter 3.0

Table 3.1 Percentage protein identity between gene products in the *agr* regions of *C. sporogenes* and *C. botulinum*.....69

Table 3.2 Protein identity between the two AgrD and two AgrB proteins of *C. botulinum* and *C. sporogenes*.....76

Table 3.3 BLAST search results for Agr flanking regions in four clostridial species.....79

Chapter 4.0

Table 4.1 Segregational stability of integrative vector pMT5112lacI in *B. subtilis* 168, expressed as percentage plasmid loss.....105

Table 4.2 Segregational stability of integrative vector pMTL5112lacI in *C. sporogenes* cured strain, expressed as percentage plasmid loss.....106

Table 4.3 Segregational stability of integrative vector pMTL5112lacI containing *pyrF* knockout cassette in *C. sporogenes* cured strain, expressed as percentage plasmid loss.....108

Table 4.4 Segregational stability of integrative vector pMTL5112lacI containing *pyrF* knockout cassette in *C. sporogenes* SC3, expressed as percentage plasmid loss.....133

Table 4.5 Nomenclature of eight antisense 'mutants' created in *C. sporogenes*.....122

Chapter 5.0

Table 5.1 Target sites and primers utilised for gene knockout of four genes in *C. botulinum* ATCC 3502 and two genes in *C. sporogenes* 10696.....144

Table 5.2 Results from conjugation and integration experiments for knockout of genes in *C. botulinum* and *C. sporogenes*.....146

Table 5.3 Primers used in three different PCR screens of *C. botulinum* and *C. sporogenes* putative knockout mutants.....147

Table 5.4 Results from conjugation and integration experiments for the knockout of *agrD*1 in *C. botulinum* using plasmid pMTL007C-E2.....154

List of Abbreviations

ABC	ATP binding cassette
Agr	Accessory gene regulator
AHL	Acyl homoserine lactone
AI-1	Autoinducer 1
AI-2	Autoinducer 2
AIP	Autoinducing peptide
AQ	Absolute quantification
ATCC	American Type Culture Collection
BHI	Brain Heart Infusion
BLAST	Basic local alignment search tool
bp	Base pairs
cDNA	Complementary DNA
cfu	Colony-forming units
Cm	Chloramphenicol
C_T	Threshold cycle
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DNTP	Deoxyribonucleoside 5'- triphosphate
dso	Double stranded (+) origin
EBS	Exon binding site
EDTA	Ethylene diamine tetraacetic acid
Erm	Erythromycin
FRT	Flp recombinase target
g	Gram
g	Gravity
h	Hour
HA	Haemagglutinin
HPA	Health Protection Agency
IBS	Intron binding site
IEP	Intron encoded protein
IPTG	Isopropyl β-D-thiogalactopyranoside

Kb	Kilobase
k Da	Kilo Dalton
Kg	Kilogram
l	Litre
LB	Luria-Bertani
M	Molar
mg	Milligram
ml	Millilitre
mM	Millimolar
mm	Millimeter
mRNA	Messenger ribonucleic acid
NCBI	National Centre for Biotechnology Information
nM	Nanomolar
NTNH	Non toxic non haemagglutinin
OD	Optical density
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNPG	*p*-nitrophenyl α-D-glucopyranoside
QS	Quorum sensing
RAM	Retrotransposition Activated selectable Marker
RBS	Ribosome binding site
RC	Rolling circle
RCA	Reinforced Clostridial Agar
RIP	RNAIII inhibiting peptide
RNA	Ribonucleic acid
RNase	Ribonuclease
RNP	Ribonuclear protein
Rpm	Revolutions per minute
RQ	Relative quantification
rRNA	Ribosomal ribonucleic acid
RT-PCR	Reverse transcriptase polymerase chain reaction
SIDS	Sudden infant death syndrome

SNARE	Soluble NSF (N-ethyl maleimide sensitive fusion protein) Attachment protein Receptor
sso	Single stranded (-) origin
TAE	Tris-acetate-EDTA buffer
TeTx	Tetanus toxin
TSS	Toxic shock syndrome
TYG	Tryptone, yeast extract, sodium thioglycolate medium
µF	Microfarads
µg	Microgram
µl	Microlitre
µM	Micromolar
UV	Ultraviolet
V	Volts
VAMP	Vesicle associated membrane protein
v/v	Volume to volume
w/v	Weight to volume
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Abstract

Botulinum neurotoxin induces a potentially fatal paralytic condition in humans and various animal species collectively known as 'botulism'. It consequently poses a major problem to the food industry, due to the ability of its spores to survive in cooked foods. The incidence of wound botulism has also suffered a recent increase in the UK. The genome sequence of the *C. botulinum* Group I strain ATCC 3502 has recently been determined. *In silico* analysis has revealed the presence of two distinct loci capable of encoding proteins with homology to AgrB and AgrD of the *Staphylococcus aureus agr* quorum sensing system. The functional characterisation of these genes has been carried out in order to determine whether they play a role in quorum sensing. To simplify laboratory procedures, *C. sporogenes,* the non-toxic counterpart of *C. botulinum*, was initially focused on. The *agr* regions in *C. sporogenes* were sequenced and their proteins compared with those of *C. botulinum* and other Gram-positive bacteria. Regions of conservation were observed amongst the clostridia and, to a lesser extent, between clostridia and staphylococci.

Transcriptional linkage assays showed some of the genes of the *C. sporogenes agr* regions to be co-expressed, and Real-Time RT-PCR demonstrated the maximal expression of these genes during late exponential growth.

Modulation of the expression of the identified *agr* genes is a prerequisite to determining their function. Due to an initial lack of an effective gene knockout tool, antisense RNA expression was used for this purpose in *C. sporogenes*, and showed that down regulation of the *agrB* genes affects sporulation. The development of an integrative vector system for gene inactivation in *C. sporogenes* was also attempted.

Knockout mutants in *C. botulinum* and *C. sporogenes* were later constructed using the newly-developed ClosTron system. These mutants were used to demonstrate that AgrD1, AgrD2 and an orphan sensor kinase protein all play a role in the control of sporulation in *C. botulinum* and *C. sporogenes*.

Acknowledgements

I would like to thank my supervisors Professor Nigel Minton, Dr Weng Chan, Dr Alan Cockayne and Professor Mike Peck for their help and support throughout my PhD. Thank you also to the Clostridial Research Group (past and present), in particular Ian Davis, John Heap, Oliver Pennington, Steve Cartman and Glen Carter, for their help in the laboratory and their sense of humour! I would also like to thank Klaus Winzer for providing many of the AgrD and AgrB sequences for my alignments.

My thanks are also extended to Victoria Wright, who invested an awful lot of time, patience and enthusiasm in my Real Time RT-PCR work. I would also like to thank Saara Qazi for her assistance and advice in the lab, and Dave Mason and Andy Carter at IFR for their advice on how to handle *Clostridium botulinum*.

Last but not least, I would like to thank my parents and my fiancé Phil for their love, support, encouragement and patience through the good times and the bad. I certainly couldn't have done this without them.

Chapter 1.0

Introduction

Chapter 1 – Introduction

1.1 Introduction to clostridia

Clostridium is one of the largest prokaryotic genera, containing over 130 species of Gram-positive, anaerobic, spore forming bacilli (Andreesen *et al.*, 1989; MacFaddin, 2000). The pathogenic clostridia cause disease by the release of potent toxins. In fact, tetanus and botulinum neurotoxins, produced by *Clostridium tetani* and *Clostridium botulinum* respectively, are two of the most harmful natural compounds known (Durre, 2001). Tetanus toxin acts on inhibitory motor nerve endings in the central nervous system, causing spastic muscle paralysis. Botulinum toxin affects the peripheral nervous system at the endings of stimulatory motor neurones by preventing the calcium-dependent release of acetylcholine, and causes the flaccid paralysis of botulism (Durre, 2001).

Clostridium perfringens produces a number of toxins associated with gas gangrene and clostridial myonecrosis, while *Clostridium difficile* causes pseudomembranous colitis and antibiotic-associated diarrhoea, via the release of Toxins A and B (Durre, 2001).

Other, non-pathogenic, strains of clostridia also exist, including *Clostridium butyricum*, which has been approved as a probiotic against enterohaemorrhagic *Escherichia coli* in Japan. Others include *Clostridium acetobutylicum* and *Clostridium sporogenes*, which have been used in tumour therapy research. *C. acetobutylicum* has also been used in industrial solvent production (Durre, 2001).

1.2 *Clostridium botulinum*

1.2.1 Discovery and occurrence

C. botulinum was first isolated in 1895 by van Ermengem and was originally named *Bacillus botulinus* (Bell & Kyriakides, 2000). It was later placed into the genus

Clostridium by Bengtson in 1924 (Hatheway, 1993). The name botulism is derived from the Latin 'botulus' which means 'sausage'. Botulism was, for a long time, associated with the consumption of contaminated sausages and the commonly-used name for the disease was 'sausage poisoning' (Bell & Kyriakides, 2000).

C. botulinum occurs naturally in soil and lake sediments, and its spores are also found on plants growing in soil contaminated with faeces. It can be spread from plant to plant by bees that accumulate spores along with pollen, and honey is consequently frequently contaminated with *C. botulinum* spores (Salyers & Whitt, 2002). It is thought that *C. botulinum* spores are probably wide-spread in all parts of the world.

1.2.2 Toxin types

The *C. botulinum* neurotoxins are made up of seven structurally related proteins – serotypes A – G (Quinn & Minton, 2001). These toxin types differ in complex size, level of activation, substrate cleavage sites, receptor binding affinity, degree of muscle weakening and duration of action (Blocker *et al.*, 2003). The strains of *C. botulinum* can be placed into four physiological groups as follows:

Group I – Proteolytic strains producing toxin types A, B or F.

Group II – Non-proteolytic strains producing toxin types B, E or F.

Group III – Non-proteolytic strains producing toxin types C or D.

Group IV – Proteolytic strains producing toxin type G.

(Hatheway, 1993)

In addition to these four groups, strains of *Clostridium butyricum* (producing toxin type E) and *Clostridium barati* (producing toxin type F) have also been isolated from cases of infant botulism (Aureli *et al.*, 1986; Hall *et al.*, 1985).

The only groups documented as causing disease in humans are groups I and II, with the most common toxin types being A, B and E (also F but less frequently) (Quinn & Minton, 2001). All organisms isolated from cases of wound botulism have been members of group I (Hatheway, 1993). Group III strains tend to cause disease in birds and animals and group IV strains are not thought to cause disease at all (Quinn &

Minton, 2001). Occasionally, strains of *C. botulinum* may exist which produce a mixture of two different toxin types. These have been reported by Gimenez and Ciccarelli (1970), who isolated a strain producing types A and F; Poumeyrol *et al.* (1983), with a strain producing types A and B; and Hatheway and McCroskey (1989), with a strain producing types B and F. However, in strains containing two different toxin genes, the second gene can often be silent, due to the presence of nucleotide deletions or insertions which result in frameshift mutations. These strains therefore only produce one toxin type (Quinn & Minton, 2001).

1.2.3 Botulinum neurotoxin structure

Botulinum toxins form heterocomplexes with other non-toxic proteins (Quinn & Minton, 2001). This complex allows the toxin to remain stable, particularly under the acidic conditions encountered in the gastrointestinal tract (Hatheway, 1993). The neurotoxin is always found in association with a non-toxic, non-haemagglutinating protein (NTNH), which is of an equivalent size to the toxin. Other proteins with haemagglutinating activity may also be present, depending on the toxin serotype (Quinn & Minton, 2001). The botulinum neurotoxins are released as progenitor toxins, in complex with these non-toxic counterparts.

The mature 150 kDa botulinum toxin is originally synthesised and released from the bacteria as a single, inactive protein via cell lysis. Proteolytic clipping is required to form the active toxin, and may be carried out by a protease produced by the clostridia, or by a gastric protease (Salyers & Whitt, 2002). Dekleva and Dasgupta (1990) isolated a 62 kDa amidase from *C. botulinum* type A Hall strain, which had the capacity to process the single chain neurotoxin, resulting in the di-chain active form. The recent completion of the *C. botulinum* genome sequence has revealed a 62 kDa clostripain protein which may also be implicated in the proteolytic cleavage of the neurotoxin (Sebaihia *et al.*, 2007). This proteolysis results in a covalently linked hetero-dimeric polypeptide, held together by a disulphide bond (Quinn & Minton, 2001).

Hence the final modified, active toxin consists of a heavy chain and a light chain. The heavy chain is 100 kDa and is involved in the binding of the toxin to target cells (mediated by the carboxy terminal domain) and formation of channels in the neuronal

membrane to allow the light chain to enter the cell (mediated by the amino terminal domain). The 50 kDa light chain is the catalytic domain, exhibiting metalloprotease activity (Salyers & Whitt, 2002).

As far as the genetic organisation of the toxin genes is concerned, the gene encoding the NTNH protein immediately precedes the gene for the toxin. In clostridia that possess haemagglutinin activity, the encoding gene precedes the NTNH gene and evidence so far suggests that the haemagglutinin proteins are always encoded by the opposite DNA strand to the NTNH and neurotoxin genes (Quinn & Minton, 2001). The HA components are encoded by a tricistronic operon, while the NTNH and neurotoxin genes are transcribed alone, or form bicistronic operons (Marvaud *et al.*, 1998a). Figure 1.1 illustrates the arrangement of the neurotoxin and associated genes in the different *C. botulinum* types.

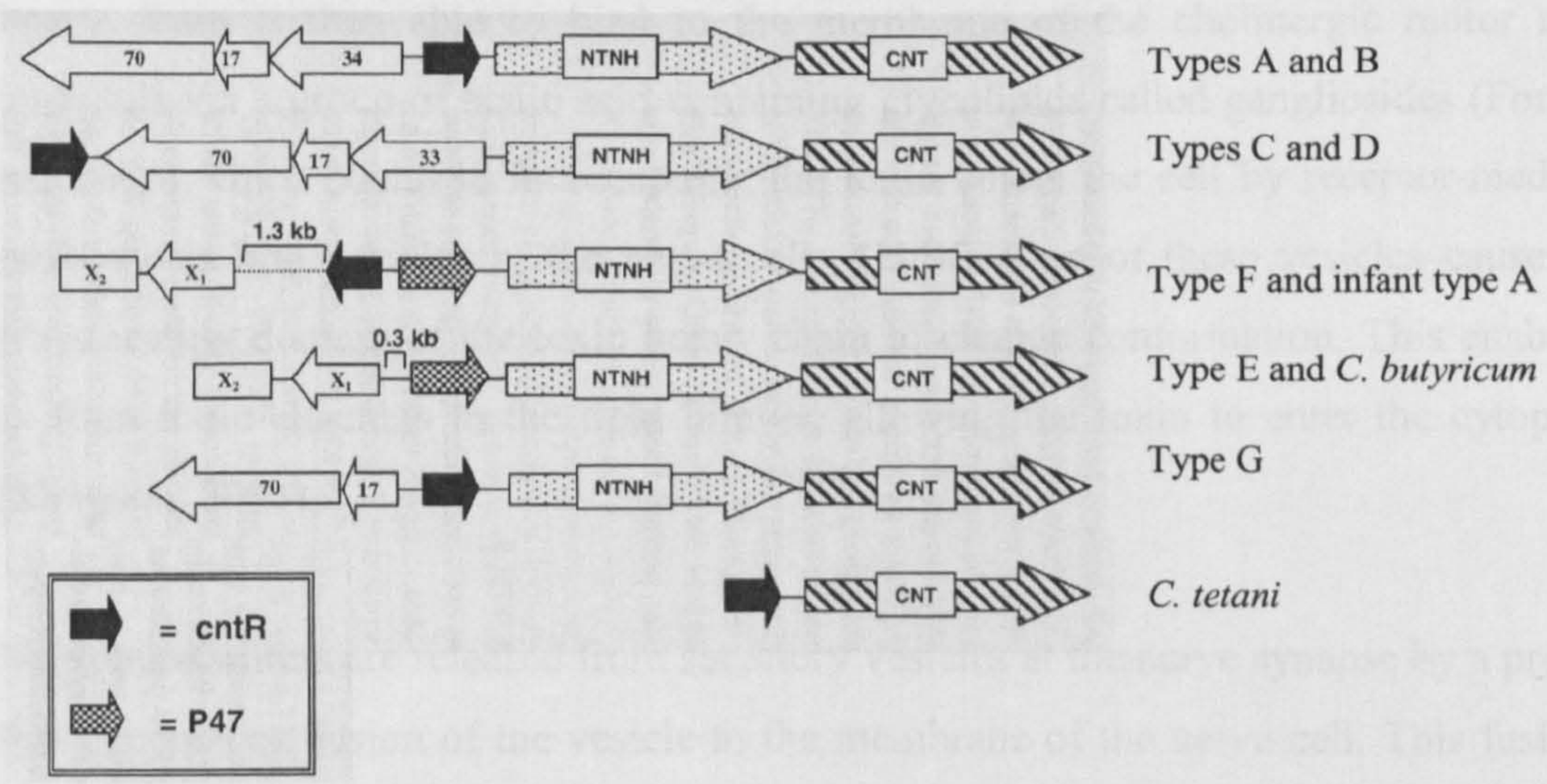


Figure 1.1 – Arrangement of neurotoxin genes in different *C. botulinum* types.
CNT, clostridial neurotoxin gene; NTNH, non-toxic, non-haemagglutinating protein gene; cntR, transcriptional regulator (botR); P47, an ORF encoding a 47 kDa protein with unknown function; X_1 and X_2, genes of unknown function; numbers indicate haemagglutinating protein genes. Adapted from Quinn & Minton (2001).

In some clostridia, the toxin genes are not carried on their chromosomes but may be carried on a lysogenic bacteriophage, which can integrate the gene into the bacterial chromosome. Only a bacterium infected with that phage is able to produce toxin, which is the case in *C. botulinum* types C and D (Salyers & Whitt, 2002). The loss of such phages renders the bacteria unable to produce toxin (Hatheway, 1993). The genes for types A, B, E and F toxins are not phage-encoded and are located on the bacterial chromosome. In contrast, Zhou *et al.* (1995) demonstrated that the genes for the toxin complex of *C. botulinum* type G are located on a large plasmid.

1.2.4 Mechanism of toxin action

Once the stable botulinum progenitor toxins have passed through the acidic environment of the stomach, they dissociate in the alkaline conditions of the intestines. The botulinum toxin is transcytosed from the mucosal side of the intestinal epithelium and passes into the bloodstream (Foran *et al.*, 2003). The binding domain of the toxin heavy chain is then able to bind to the membrane of the cholinergic motor nerve terminals via a group of sialic acid-containing glycolipids called gangliosides (Foran *et al.*, 2003). Once bound to its receptors, the toxin enters the cell by receptor-mediated endocytosis into vesicles in the nerve cell. Acidification of these vesicles causes the translocating domain of the toxin heavy chain to change conformation. This enables it to form ionic channels in the lipid bilayer, allowing the toxin to enter the cytoplasm (Simpson, 2004).

Neurotransmitters are released from secretory vesicles at the nerve synapse by a process that depends on fusion of the vesicle to the membrane of the nerve cell. This fusion is dependent on a conserved set of membrane proteins, which form a complex and are known as SNARE proteins (Soluble NSF [N-ethyl maleimide sensitive fusion protein] Attachment protein REceptor) proteins (Fasshauer *et al.*, 1998; Sutton *et al.*, 1998). These SNARE proteins are divided into groups and are essential components involved in vesicle fusion and secretory events in secretory cells. The SNARE proteins located in the nerve terminals are:

1. Vesicle Associated Membrane Protein (VAMP), also known as synaptobrevin
2. Syntaxin
3. SNAP-25 (a 25 kDa synaptosomal protein)

(Quinn & Minton, 2001)

The botulinum neurotoxins act on the cholinergic synapses of the peripheral nervous system by inhibiting the secretion of neurotransmitters. This inhibition occurs via the zinc-dependent hydrolysis of one or more of the SNARE proteins. Disabling any of these proteins prevents fusion of the vesicles to the nerve cell membrane and subsequent release of neurotransmitter. Mechanistically, the botulinum neurotoxins are able to bind to these SNARE proteins via a ten-residue consensus sequence of amino acids, of which all the SNARE proteins contain multiple copies (Herreros *et al.*, 1999). The blockage of the nerve synapses results in the characteristic flaccid paralysis observed in all forms of botulism.

1.2.5 Types of botulism and epidemiology

C. botulinum causes three naturally occurring types of disease: wound botulism, foodborne botulism and infant botulism. However, a fourth type, inhalational botulism, can also occur as a consequence of a man-made aerosolised version of the toxin (Coleman and Yergler, 2002). The only known case of this occurred in Germany in 1962 (Arnon *et al.*, 2001).

Botulism occurs after the toxin is absorbed into the circulation from a wound or a mucosal surface like the gut (Coleman & Yergler, 2002). Botulism is not an infection, it is in fact an intoxication; *C. botulinum* is unable to colonise the gut of healthy adults due to competition from the commensal flora. The three main types of botulism are outlined below.

1.2.5.1 Foodborne botulism

This form of the disease has historically been the most commonly occurring type in the UK and, consequently, *C. botulinum* has been the principal target of the food processing industry for almost a century. This form of the disease occurs following the

consumption of food that contains toxin produced by the bacterium. Contamination of food by *C. botulinum* is not uncommon, but simple ingestion does not usually cause the disease. Any bacteria germinating from spores cannot compete with the commensal flora of the gut and are therefore unable to grow to a high enough concentration to cause disease. Disease occurs when spores germinate in foods, producing toxin, which is subsequently ingested (Salyers & Whitt, 2002). The bacterium does not grow easily in foods but is most common in home-canned foods that have been heated, cooled and stored for long periods at room temperature. It is also an emerging issue with the increasing popularity of minimally heated, chilled foods, or 'ready-meals' (Peck, 2006). Growth and neurotoxin production can occur at 10-12°C in proteolytic *C. botulinum* (Peck & Stringer, 2005), but worryingly, growth has been reported for non-proteolytic *C. botulinum* at temperatures as low as 3°C (Graham *et al.*, 1997). This is well below the required 8°C, which applies to storage of chilled foods in England and Wales (Peck, 2006). The estimated lethal dose of a type A toxin for a human of 70 kg is approximately 70 μg orally and 0.7-0.9 μg if inhaled (Arnon *et al.*, 2001).

Initial neurological symptoms include nausea, diarrhoea, headache, dizziness, slurred speech and a dry mouth. If symptoms become more severe, the muscle paralysis can affect the muscles of the respiratory and cardiac systems, and death can subsequently occur if left untreated (Bell & Kyriakides, 2000). The symptoms of foodborne botulism appear approximately 12 to 36 hours after exposure to the toxin (McLauchlin *et al.*, 2006).

Although the occurrence of foodborne botulism is relatively rare (only 62 cases were reported between 1922 and 2006 (McLauchlin *et al.*, 2006)), the severity of its symptoms and prolonged treatment required by patients means that the cost to health services is extremely high. Setlow and Johnson (1997) predicted that in the USA, each case of botulism could cost approximately $30 million. It is therefore imperative that steps are taken to prevent the growth and toxin formation of *C. botulinum* in foods.

Recently, due to the treatment of foodborne botulism with equine antitoxin and supportive therapy, the fatality rate has fallen to 5-10% of cases (Peck, 2006). The introduction of effective control measures has also aided the reduction in the incidence

of foodborne botulism, with most cases now seemingly associated with home-prepared foods where the recommended precautions have not been adhered to.

1.2.5.2 Wound botulism

It is possible for *C. botulinum* to colonise a deep wound, as its environment may be sufficiently anaerobic to allow growth. Tissue destruction in the wound results in the blood supply to the area being substantially reduced. Any remaining oxygen is depleted by the body cells, leaving an anaerobic environment for the bacterium. The toxin produced then passes into the blood stream and causes disease, with similar symptoms to foodborne botulism. Previously, this type of botulism has been relatively rare in the UK, but there has been a recent increase in its incidence, which is largely due to its occurrence in intravenous drug users. *C. botulinum* wound infections had not been reported in the UK and Republic of Ireland until 2000 (Akbulut *et al.*, 2005). However, between the beginning of 2000 and the end of 2002 there were thirty-three clinically diagnosed cases of wound botulism in the UK and Ireland. The incidence is now rising rapidly with a further 15 cases in 2003, 40 in 2004 and at least 30 in 2005. All of these patients had injected heroin (Brett *et al.*, 2004; Akbulut *et al.*, 2005; McLauchlin *et al.*, 2006). Cases of wound botulism have also been previously reported when deep wounds have been heavily contaminated with soil containing *C. botulinum* (Salyers & Whitt, 2002).

1.2.5.3 Infant botulism

Infant botulism is relatively uncommon world-wide, but is the most common type of botulism in the USA (Quinn & Minton, 2001; Cox & Hinkle, 2002). In fact, since 1976, when the first two cases were diagnosed (Dodds, 1993), over 1000 cases have been reported world-wide, and over 90% of these were in the USA (Dodds, 1993; Cox & Hinkle, 2002). This form of the disease affects otherwise healthy children, who are usually less than one year old. At this age, the child's intestines are vulnerable to colonisation by *C. botulinum* because their commensal flora is not yet fully developed (Dodds, 1993). The most common cause of infant botulism seems to be spores in the general environment. These are ingested and can germinate in the anaerobic

environment of the gut, resulting in colonisation and toxin production (McMaster *et al.*, 2000). The only food so far suspected of causing infant botulism is honey, which can carry large numbers of bacterial spores (Dodds, 1993). In fact, honey has been implicated in a number of cases of infant botulism (Sonnabend *et al.*, 1985; Cox & Hinkle, 2002) and it is now recommended that parents do not feed honey to children below one year of age (Cox & Hinkle, 2002). A study by Cherington (1998) showed that up to 25% of honey products were found to contain spores.

Infant botulism has also been implicated in Sudden Infant Death Syndrome (SIDS), a suggestion which was first made by Midura and Arnon (1976). However, recent research has produced conflicting reports (Sonnabend *et al.*, 1985; Byard *et al.*, 1992).

1.2.6 Treatment

Once the toxin has bound to, and entered a neurone, any treatment is futile. Nerve regeneration may occur if a patient survives, but survivors often suffer irreversible neurological damage (Salyers & Whitt, 2002). The only treatment that can be given is supportive care during the paralysis, which may include mechanical ventilation in severe cases (Foran *et al.*, 2003). However, if administered promptly, antitoxin is able to neutralise any free toxin which has not yet bound to the nerve endings (McLauchlin *et al.*, 2006). This usually consists of polyvalent anti-toxin, as the specific toxin type may not be known (Bell & Kyriakides, 2000). However, as these antitoxins are only effective if administered promptly, and may cause allergic reactions (Breidenbach & Brunger, 2005), the development of curative treatments is crucial. Foran *et al.* (2003) reviewed some of the developing new treatments for botulism, which include the use of protease inhibitors and the genetic transfer of toxin-protease-resistant SNARES. Interestingly, the recent research of Satoh and colleagues suggests that the thearubigin fraction from black tea is highly efficient at blocking botulinum neurotoxin action *in vitro* (Satoh *et al.*, 2001; Satoh *et al.*, 2002; Satoh, 2005).

1.2.7 Therapeutics

Although, in its naturally occurring concentrations, botulinum neurotoxin is potentially lethal, the therapeutic use of neurotoxin type A was approved by the Food and Drug Administration in the USA in 1989 (Barnes, 2003). Highly diluted toxin is able to relieve muscle spasm within three days and has consequently been used to treat numerous spastic muscle conditions (Durre, 2001). These include spasms affecting vision (blepharospasm and strabismus); cerebral palsy (Durre, 2001); pharyngoesophageal spasm (Chao *et al.*, 2004); overactive bladder (Cruz, 2004); hemifacial spasm (Laskawi *et al.*, 1994); spasticity after stroke (Childers *et al.*, 2004); migraine (Blumenfeld *et al.*, 2004); excessive sweating (Naumann & Jost, 2004); whiplash injuries (Juan, 2004) and some of the symptoms of Parkinson's disease and Raynaud's syndrome (Cordivari *et al.*, 2004; Sycha *et al.*, 2004).

1.2.8 Bioterrorism

The extreme potency of the *C. botulinum* neurotoxins has rendered them a potential bioterrorist's weapon. They are among the most dangerous candidates for use in bioterrorism, along with anthrax and smallpox (Durre, 2001). The *C. botulinum* neurotoxin has been classed as a category A agent by the Centres for Disease Control in Atlanta, USA. These substances are the most serious threats and are 'easily disseminated, often transmissible from person to person, likely to cause high mortality, panic and disruption and would require some form of special intervention by the relevant public health authority (Lamb, 2001).

Botulinum neurotoxin could potentially be used in an aerosol, or to contaminate food supplies, and 1 g could be sufficient to kill one million people (Lamb, 2001). It was first developed as a biological weapon over sixty years ago and was released as an aerosol by the Aum Shinrikyo cult in Japan in the 1990s. However, no cases of botulism resulting from this release were reported (Whitby *et al.*, 2002). Iraq has also admitted to producing 19,000 litres of botulinum toxin and loading approximately half of this into missiles and bombs (Arnon *et al.*, 2001). It is also claimed that, incredibly, the amount of toxin which has still not been accounted for in Iraq is three times the amount required to kill every human being on earth (Arnon *et al.*, 2001).

Some comfort can, however, be taken from the fact that the distribution of the toxin via aerosol from a biological weapon would not be easy, and the production and distribution of large enough quantities would be difficult (Durre, 2001). The toxin can also be easily inactivated by standard water-cleaning procedures so it would be difficult to produce high enough concentrations to be effective. Botulism cannot be spread from person to person and is treatable, as described above, to a certain extent (Lamb, 2001). The fact that, generally, the mortality rate is less than 10% would also make it a less appealing agent to bioterrorists (Whitby *et al.*, 2002).

1.2.9 Regulation of botulinum neurotoxin expression in *C. botulinum*

Various factors have been implicated in the regulation of botulinum neurotoxin gene expression. These include two-component systems allowing the bacteria to sense changes in the population density of the external environment (known as Quorum Sensing, and discussed later), and the availability of exogenous carbon and nitrogen sources (Patterson-Curtis & Johnson, 1989).

C. botulinum types A, B, C, D, F and G possess a regulatory gene *botR* which is thought to regulate toxin expression. The BotR protein is approximately 21 kDa and shows characteristics of a DNA binding protein (Marvaud *et al.*, 1998b). In types C and D the *botR* gene is located upstream of both the HA and NTNH genes. In types A, B, F and G, the *botR* gene is found between the HA and NTNH genes (Marvaud *et al.*, 1998a) [see Figure 1.1]. Marvaud *et al.* (1998b) demonstrated that over-expression of the BotR/A protein in *C. botulinum* type A induced a higher level of production of botulinum neurotoxin A and associated non-toxic proteins, showing that BotR is indeed a positive regulator of toxin expression. They also used antisense mRNA to inhibit expression of the *botR/A* gene, which decreased the toxicity of the culture supernatant. A similar regulator, TetR, is found in *C. tetani.* Marvaud *et al.* (1998a) showed that BotR/A was also able to regulate the tetanus toxin (*TeNT*) gene in *C. tetani*. This was due to their high sequence homology, with TetR showing overall identity of 50-65% with BotR/C and BotR/F (Marvaud *et al.*, 1998a).

More recent research has shown that BotR and TetR regulate toxin production by functioning as alternative sigma factors. The BotR/A-core enzyme complex recognises two separate promoters; one driving transcription of the neurotoxin and NTNH genes and one driving transcription of the HA genes (Raffestin *et al.*, 2005). It has also been suggested that these regulators belong to the same family of alternative sigma factors in which *C. difficile* TcdR and *C. perfringens* UviA are found (Raffestin *et al.*, 2005; Garnier & Cole, 1988; Mani & Dupuy, 2001; Mani *et al.*, 2002). These sigma factors differ so much in their structure and function from major sigma factors identified in other prokaryotes that they have been assigned to their own group (group 5 of the σ^{70} family) (Dupuy & Matamouros, 2006).

1.3 *Clostridium sporogenes*

C. sporogenes is biochemically and morphologically similar to group I strains of *C. botulinum*, with up to 100% rRNA and 70-100% DNA homologies (McKinney *et al.*, 1993). As it does not produce a neurotoxin, it is an ideal candidate to study in place of *C. botulinum*. Results can be extrapolated to *C. botulinum*, but the hazards posed by working with the toxigenic species are avoided. Indeed it is routinely used by the food industry as a non-hazardous *C. botulinum* model.

McKinney *et al.* (1993) produced probes that were able to distinguish between *C. botulinum* and *C. sporogenes* using selective hybridisation. DNA sequences present in *C. botulinum* group I but absent in *C. sporogenes* could then be detected. The probes were not sequences of the toxin genes, indicating that there are additional differences, besides the absence of the toxin in *C. sporogenes*.

Although thought to be relatively non-pathogenic, *C. sporogenes* has occasionally been associated with various diseases including cerebrocortical necrosis, osteomylitis and antibiotic-associated haemorrhagic diarrhoea. The virulence factor associated with haemorrhagic diarrhoea is the haemorrhagic toxin, which shows collagenase activity (Hara-Kudo *et al.*, 1996).

1.4 Sporulation in clostridia

One of the most notable characteristics of the clostridia is their ability to form endospores. These dormant cells are generally formed under conditions unfavourable for growth and are highly resistant to harsh treatments like heat, freezing, chemicals, radiation and desiccation (Brown, 2000). In other spore-forming bacteria like *Bacillus* species, spores are generally produced in response to a lack of nutrients. However, in clostridia, carbon and energy sources are required in order for sporulation to take place, and other factors which may affect the cells' decision to sporulate include pH, temperature and exposure to oxygen (Durre, 2001).

The ability of *C. botulinum* spores to persist in cooked foods is a major factor responsible for cases of foodborne botulism. Indeed, spores of non-proteolytic strains of *C. botulinum* are able to survive the pasteurisation process and also to germinate even at refrigeration temperatures (Webb *et al.*, 2007; Graham *et al.*, 1997). As a result, a minimum standard heat treatment of 121.1°C for 3 minutes has been recommended as the treatment required in canned foods for a 'botulinum cook' (Peck, 2006; Stumbo *et al.*, 1975). Clearly this is not always adhered to during home-preparation of foods, which is when cases of foodborne botulism are usually observed.

Spore formation is a crucial survival mechanism utilised by the clostridia, which enables them to persist in an aerobic atmosphere until more favourable, anaerobic, conditions are found. It may also therefore be considered as a virulence factor, as, in the pathogenic clostridia, the formation of spores generally contributes to the survival of the organism in the environment and the subsequent development of disease. Indeed, once conditions become more favourable, spores are able to germinate back into vegetative cells, and subsequently cause disease by the production of toxins. In fact, it has been shown that some species of clostridia require sporulation to take place in order for toxins to be produced. For example, the *C. perfringens* enterotoxin is only synthesised in sporulating cells (Durre & Hollergschwandner, 2004). Any ingested bacteria sporulate in the intestines, with the concomitant production of enterotoxin. This toxin is responsible for causing histological damage to the intestinal epithelium, resulting in severe diarrhoea (McClane & Rood, 2001). *C. perfringens*-related food poisoning is currently the third most common foodborne disease in the USA (Li & McClane, 2006).

C. difficile is a major contributory organism in hospital acquired infections due to the resistance of its spores to many chemical treatments. Most commonly-used detergents, as well as ethanol, are inactive against its spores, which are consequently able to persist in the hospital environment. Spores are easily transmitted between patients, and also between healthcare workers and patients. Furthermore, is has been shown that *C. difficile* spores are able to survive in the environment for five months (Kim *et al.*, 1981). Because of this persistence, cases of *C. difficile*-related diarrhoea are increasing rapidly, with 51,690 UK cases being reported in 2005 alone (HPA report, 2006). The estimated cost of *C. difficile*-associated infections in the USA is $1.1 billion per year (McFarland *et al.*, 2007).

Clostridial spore formation is therefore an important virulence factor which, coupled with toxin production, allows these organisms to successfully persist in an aerobic environment and cause serious diseases. It is therefore vital to investigate the processes which control sporulation in order to better understand the control of clostridial virulence. Effective control measures may then be implemented in an effort to prevent or reduce spore formation.

1.5 Quorum sensing

1.5.1 The principle of quorum sensing

Quorum sensing is a mechanism of cell to cell communication among bacteria (Lerat & Moran, 2004). It is utilised by bacterial populations as a means of co-ordinately responding to external stimuli in a cell density-dependent manner. This occurs via signalling molecules known as autoinducers, and results in the regulation of specific genes. These signal molecules are released so that bacteria can detect the density of a population in a particular environment. The molecules are released during growth and accumulate in the environment. Once the concentration of the signal molecule reaches a threshold, or quorum, this triggers a response in the bacterial population (Winzer *et al.*, 2003). This usually involves a cascade of events which ultimately leads to changes in gene regulation (Winzer *et al.*, 2003). Genes regulated by quorum sensing include those

involved in virulence factor production (Winzer *et al.*, 2003), cell migration, formation of biofilms (Lerat & Moran, 2004), sporulation (Hilgers & Ludwig, 2001) and production of toxins (Kleerebezem & Quadri, 2001). It may be favourable for individual bacterial cells to suppress the expression of certain genes when the population density falls below a threshold. By waiting until the population size has increased, the concentration of virulence factor produced, and hence its effect, is much greater (Winzer *et al.*, 2003).

Quorum sensing was first described in the Gram-negative bacterium *Vibrio fischeri* (Nealson *et al.*, 1970). This bacterium lives on marine eukaryotes and produces light for its hosts. Light emission only occurs when the bacterial cell density reaches a threshold, and the luciferase operon is transcribed. The autoinducer is an acylated homoserine lactone (AHL) (Lerat & Moran, 2004). AHLs accumulate, and once a threshold is reached, activate a transcriptional regulator of the LuxR family (Winzer *et al.*, 2003). Quorum sensing systems have subsequently been discovered in many bacterial species.

Vibrio harveyi is another marine bacterium whose luminescence is regulated by quorum sensing. However, this organism has two quorum sensing systems (Bassler *et al.*, 1993). One is similar to that of *Vibrio fischeri* and uses the AHL signal molecule. The second system relies on the accumulation of a molecule called autoinducer 2 (AI-2), which is believed to be a furanosyl borate diester (Chen *et al.*, 2002). AI-2 is generated via the LuxS enzyme (Schauder *et al.*, 2001) and binds to a periplasmic protein LuxP, which then interacts with a membrane-bound histidine protein kinase, LuxQ. Following a series of dephosphorylation reactions, LuxO can no longer activate transcription of the repressor gene, and the *lux* operon is consequently transcribed (Winzer *et al.*, 2003).

A variety of Gram-positive and Gram-negative bacteria possess *luxS* homologues, similar to those found in *V. harveyi* (Winzer *et al.*, 2003). In fact, it has been claimed that BLAST searches have shown evidence of *luxS* homologues in the genomes of 'most bacterial species for which complete sequence data are available' (Hilgers & Ludwig, 2001). In addition, Xavier and Bassler (2003) found that the *luxS* gene exists in 35 of the 89 currently available complete bacterial genomes by database analysis. PCR, along with AI-2 activity assays, have allowed the identification of *luxS* in a number of species for which no genome sequences are available (Xavier & Bassler, 2003).

In some cases, signalling molecules can be recognised by different bacterial species. This is known as interspecies cross-talk and can either promote or inhibit gene expression (Sturme *et al.*, 2002). As *luxS* has been found in many different bacteria, it has recently been proposed that AI-2 could serve as a 'universal signal' for interspecies cross-talk (Xavier & Bassler, 2003). However, this is yet to be proved in species other than *Vibrio*.

Despite the proposed role of LuxS in quorum sensing, some researchers have recently questioned this theory and postulated that it is in fact a metabolic side product, produced during the activated methyl cycle (a recycling pathway involved in the metabolism of methionine) (Winzer *et al.*, 2002; Winzer *et al.*, 2003). Indeed, Doherty *et al.* (2006) demonstrated that the inactivation of *luxS* in *S. aureus* had no effect on virulence-associated traits. It is possible that the phenotypic changes observed in some bacteria following inactivation of *luxS*, may be due to the absence of the AI-2 signal molecule, but they could also be attributed to the disruption of the activated methyl cycle (Winzer *et al.*, 2003). The true role of this protein is yet to be elucidated.

Because quorum sensing controls the regulation of so many processes, including the production of virulence factors, it has been the focus of attention as a therapeutic target in the search for alternatives to conventional antibiotic therapy, which is becoming increasingly redundant due to the emergence of resistance. Areas attracting particular interest are the recent increase in Methicillin-Resistant *Staphylococcus aureus* (MRSA) (Abraham, 2006) and the problem of biofilm formation in *Pseudomonas aeruginosa* infections in cystic fibrosis patients (Lyczak *et al.*, 2000).

1.5.2 Quorum sensing in Gram-positive bacteria

Quorum sensing in Gram-positive bacteria occurs via a range of different signals but all systems involve a two-component regulatory signal transduction system, consisting of a membrane-bound receptor histidine protein kinase, and an intracellular response regulator (Sturme *et al.*, 2002). When this response regulator is phosphorylated by the receptor kinase, transcription of target genes is activated or repressed. Transcription of genes involved in the production of the autoinducing peptides is also activated (Sturme

et al., 2002). Most Gram-positive bacteria use signalling molecules that are created from large precursors and post-translationally modified (Fuqua & Greenberg, 1998). They are then exported using ATP binding cassette (ABC) export systems (Klein *et al.*, 1992; van der Meer *et al.*, 1993). Some bind to sensor kinases in the cell membrane, which transduce the signal across the membrane. Others bind to intracellular receptors after being transported into the cell by oligopeptide permeases.

The paradigm for Gram-positive quorum sensing is the staphylococcal *agr* (accessory gene regulator) system. This is the only known quorum sensing system in the staphylococci (Kies *et al.*, 2003) and was first identified in *S. aureus*. In late exponential to stationary phase, this system inhibits the expression of a number of surface proteins, for example fibronectin binding proteins A and B, and increases the expression of many secreted virulence factors (Yarwood & Schlievert, 2003). The system is able to switch on in the later parts of the growth phase by sensing the cell density of its surroundings (Otto *et al.*, 1998). The virulence factors controlled by *agr* include α, β and δ cytotoxins (haemolysins), serine protease, DNase, fibrinolysin, enterotoxins A-D, toxic shock syndrome (TSS) toxin 1 (Otto *et al.*, 1998), and leucocidins (Lina *et al.*, 1998). Other proteins whose production is regulated by *agr* in staphylococci include staphopain, lipase, phospholipase C, nuclease, hyaluronidase, coagulase, protein A (Novick, 2003), arginase, peptidase and chitinase B (Dunman *et al.*, 2001). In addition, *agr* is also involved in the invasion and apoptosis of epithelial cells (Yarwood & Schlievert, 2003).

The *agr* locus is approximately 3.5 Kb and is made up of two transcripts – RNAII and RNAIII, which are generated from the P2 and P3 promoters, respectively. The P2 operon codes for four proteins generating the *agr* sensing mechanism – Agr A, B, C and D (Novick *et al.*, 1995). The P3 operon codes for RNAIII, which is the effector molecule of the two-component system, and controls the expression of target genes. It also encodes staphylococcal δ toxin (Janzon *et al.*, 1989; Novick *et al.*, 1995). The roles of the components of the *agr* system are outlined below and illustrated in Figure 1.2.

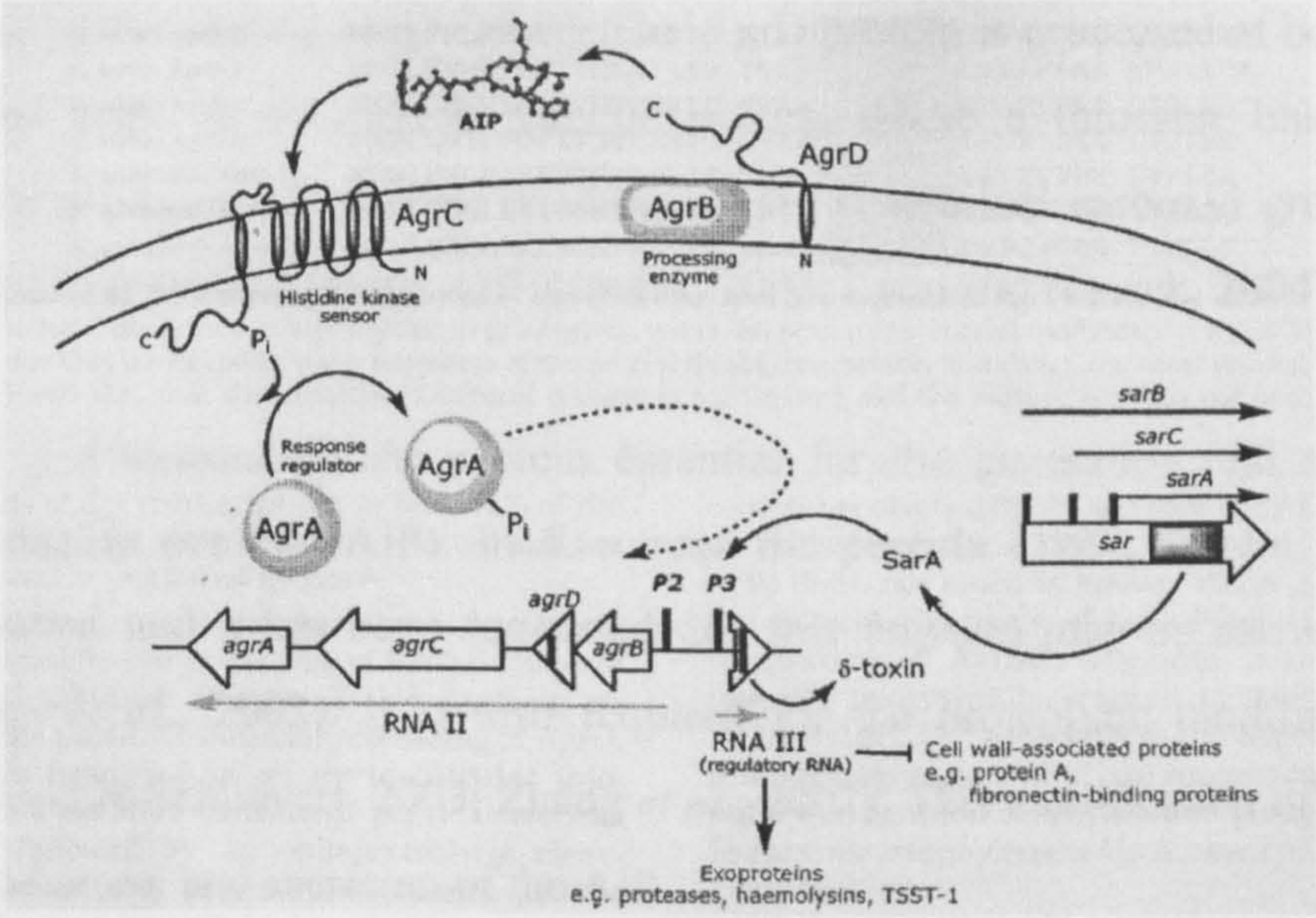


Figure 1.2 – Diagram of the *agr* quorum sensing system from *S. aureus*, taken from Chan *et al.* (2004). The AgrD propeptide is modified and secreted from the cell by AgrB and possibly another peptidase. This results in the formation of the mature AIP, which binds to the AgrC sensor kinase. Subsequent phosphorylation of AgrC activates the AgrA response regulator which promotes transcription of the *agrBDCA* operon. It also increases transcription of RNAIII, the effector molecule of the system.

AgrC is the transmembrane sensor kinase of a two-component system, whose N terminal interacts with a signal molecule. Activation of AgrC then occurs via a phosphorylation event. It was originally proposed that the C terminal was autophosphorylated at a histidine residue when stimulated by the signal molecule (Ji *et al.*, 1995; Novick *et al.*, 1995). However, subsequent investigations have found that AgrC is partially phosphorylated in the absence of any AIP. It has been suggested that the cytoplasmic domain of AgrC is dimeric, containing a four-helix bundle, and that autophosphorylation would occur as a *trans* process (Novick, 2003).

AgrA is the response regulator, which is required for the activation of the P2 and P3 promoters and is phosphorylated by AgrC. It binds with high affinity to the RNAIII-*agr* intergenic region, with this binding being localised to a pair of direct repeats in the P2 and P3 promoter regions (Koenig *et al.*, 2004).

AgrD is a propeptide, which is synthesised within the cell, processed and then secreted. AgrD requires modification by AgrB for production of the mature Autoinducing Peptide

(Ji *et al.*, 1995; Zhang *et al.*, 2002). The AgrD protein is processed at both the N and C terminal ends. A cyclisation reaction occurs, where a thioester linkage is formed between a conserved cysteine residue and the C terminal carboxyl group, in order to form the mature thiolactone AIP (Novick, 2003; Lyon and Novick, 2004).

AgrB is a transmembrane protein essential for the generation and secretion of the autoinducing peptide (AIP). In *S. aureus* the peptide export system is not an ABC transporter, and it has been suggested that this function may be carried out by AgrB (Sturme *et al.*, 2002). It is also required for the proteolytic modification of AgrD (Yarwood & Schlievert, 2003; Zhang *et al.*, 2002). This modification of AgrD results in the production and secretion of the AIP.

AIP, the autoinducing peptide of the two-component system, is a cyclic thiolactone of seven to nine amino acids, and is crucial for *agr* activity (Zhang *et al.*, 2002). The *agr* genes are constitutively expressed at low cell densities, but when the AIP level reaches a threshold, the *agr* response is activated (Ji *et al.*, 1995). Signalling by AgrA and AgrC results in an increase in the transcription of RNAII and RNAIII in a positive feedback manner (Yarwood & Schlievert, 2003).

RNAIII has two functions. It encodes the δ haemolysin toxin via the *hld* gene, but is also responsible for the increase in transcription of extracellular virulence factors, such as toxins, and the decrease in transcription of cell surface proteins (Yarwood & Schlievert, 2003). It has recently been demonstrated that RNAIII exerts these effects by its 3' domain acting as an antisense RNA, and repressing translation of mRNAs encoding virulence factors acting early in the infection process (Boisset *et al.*, 2007). In addition, the increase in transcription of secreted virulence factors observed in the later stages of infection was shown to occur by the antisense down-regulation of Rot (repressor of toxins) mRNA by RNAIII. Since Rot has been identified as an antagonist to the *agr* response, this de-repression permits the translation of many of the secreted virulence factors associated with the later stages of infection. It has also been demonstrated that the 3' domain of RNAIII by itself is able to promote the synthesis of several exoproteases and exotoxins (Boisset *et al.*, 2007).

Signal molecules produced by the *agr* system in *S. aureus* have been shown to stimulate *agr* activation in cells of their own population, but to inhibit expression of the *agr* system in cells of a different population (Kies *et al.*, 2003). This is largely due to the fact that variations in specific regions of *agrB*, *agrD* and *agrC* have resulted in the allocation of four *agr* specificity groups in *S. aureus* (Novick, 2003). The groups were determined by assessing the ability of their AIP molecules to inhibit the *agr* response in other groups (Ji *et al.*, 1997). These AIP groups show extremely high specificity, and just a single amino acid change can transform the AIP to a different group (Novick, 2003). Cross inhibition occurs across these four AIP groups, but the AIPs from groups I and IV are able to activate each other's *agr* system due to their strong similarity (Chan *et al.*, 2004). The structures of the AIPs from the four *S. aureus agr* groups are illustrated in Figure 1.3.

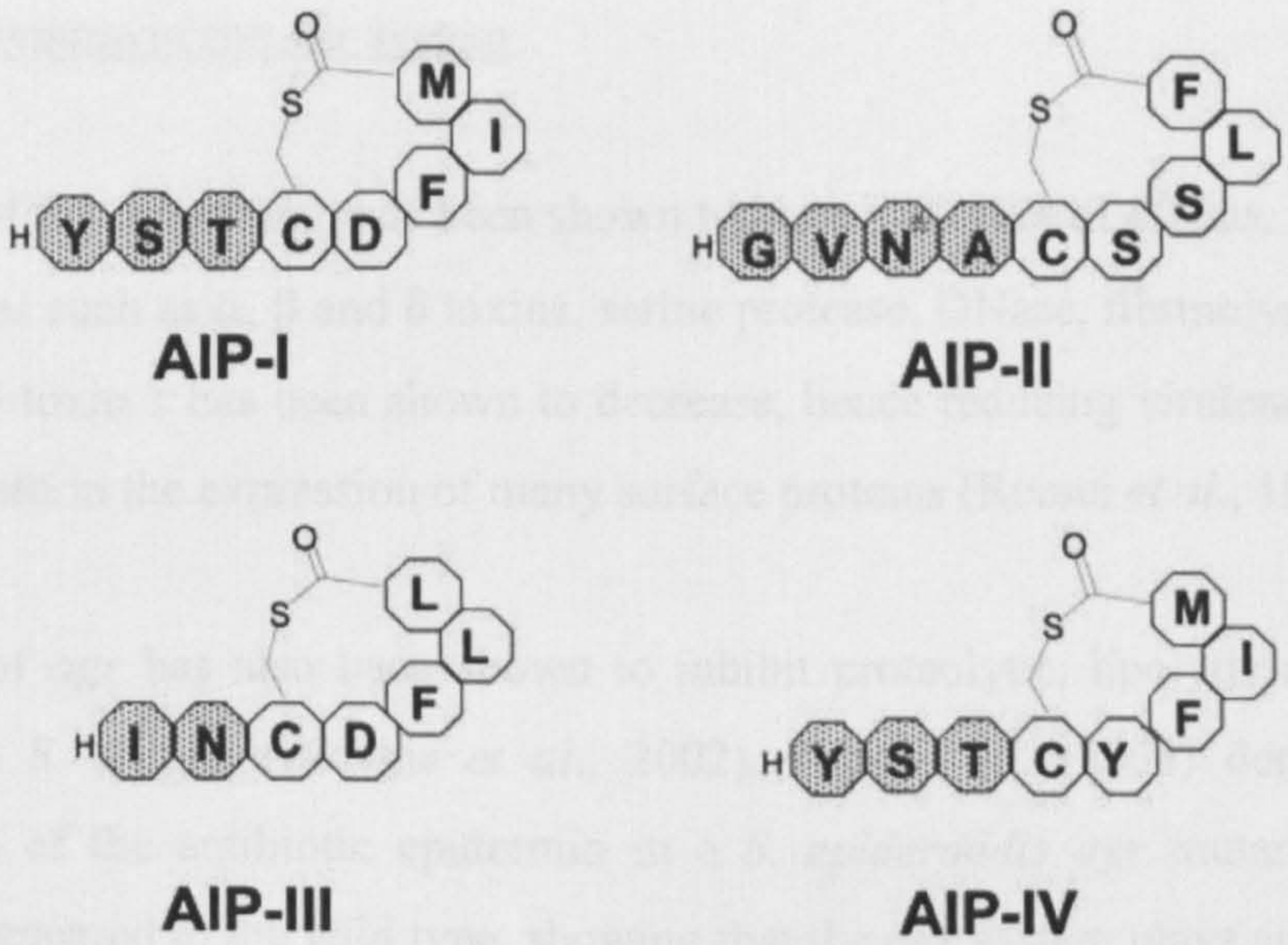


Figure 1.3 – Structures of the AIPs from the four *S. aureus agr* groups. Letters refer to amino acid residues. Exocyclic tail residues are shown in grey; endocyclic residues are shown in white. Taken from Lyon *et al.* (2002).

The production of RNAIII is affected by a second regulatory system – the staphylococcal accessory regulator *sar* (Blevins *et al.*, 2002). Indeed, Blevins *et al.* (2002), Bayer *et al.* (1996) and Dunman *et al.* (2001) demonstrated that mutating *sarA* resulted in a decrease in the production of RNAIII. Binding of SarA to the P2 and P3 promoters promotes transcription of RNAII and RNAIII, indirectly affecting virulence factor regulation. The *sar* locus also operates independently of *agr*, regulating other virulence factors directly (Dunman *et al.*, 2001).

1.5.3 Inactivation of the *agr* system

Mutation of the *agr* system has been shown to have a number of effects. The expression of exotoxins such as α, β and δ toxins, serine protease, DNase, fibrinolysin, enterotoxin B and TSS toxin 1 has been shown to decrease, hence reducing virulence. An increase was also seen in the expression of many surface proteins (Recsei *et al.*, 1986).

Mutation of *agr* has also been shown to inhibit proteolytic, lipolytic and haemolytic activity in *S. aureus* (Blevins *et al.*, 2002). Kies *et al.* (2003) demonstrated that production of the antibiotic epidermin in a *S. epidermidis agr* mutant was strongly reduced compared to the wild type, showing that the *agr* system plays an important role in epidermin biosynthesis.

The *agr* system therefore has potential for targeting in an attempt to attenuate virulence in staphylococci. However, there is an alternative argument that inactivating the *agr* system may, in fact, contribute to the conversion of an acute infection to a chronic one (Kong *et al.*, 2006). Some studies have shown that inhibiting the *agr* system enables staphylococci to form biofilms more readily. Vuong *et al.* (2004) demonstrated this in *S. epidermidis* and suggested that this may present problems when using medical implant devices.

1.5.4 Quorum sensing in clostridia

Several clostridia, including *C. botulinum*, have been found to possess LuxS homologues. The *C. botulinum luxS* gene shows 50.3% sequence identity with that of *V. harveyi*, but limited literature is available on its mechanism of control of cellular functions. To date, there is only one documented report on the function of *C. botulinum luxS* (Zhao *et al.*, 2006). This group demonstrated that AI-2 was present in *C. botulinum* culture supernatant and that this supernatant promoted the germination and outgrowth of spores. However, there is no evidence to suggest that this promotion was directly due to the effect of AI-2.

LuxS has also been investigated in other clostridial species and indeed, Ohtani *et al.* (2002b) showed that the *luxS* gene is involved in the regulation of toxin production in

C. perfringens via quorum sensing. This group also demonstrated that *luxS* mutants produced much less α, κ and θ toxins than the wild type, and that these mutants could be complemented with wild type culture supernatant to restore toxin production.

The study of quorum sensing in *C. difficile* is also gathering momentum, due to the emerging incidence of *C. difficile*-associated disease (CDAD) in hospitals. Carter *et al.* (2005) demonstrated that AI-2 is also produced by this clostridial species. In the region upstream of *luxS*, genes encoding a putative two component signal transduction system were identified, and the system was found to be involved in the regulation of AI-2 production. However, they suggested that AI-2 was not involved in the control of toxin production, as down-regulation of the *luxS* gene using antisense RNA did not cause a reduction in toxin production.

Lee & Song (2005) also carried out a similar study into *luxS* and quorum sensing in *C. difficile*, and demonstrated an up-regulation of toxin A mRNA production in response to exposure to AI-2-containing *C. difficile* supernatant. However, they could not demonstrate a concomitant increase in the level of the toxin. Their data do suggest, however, that *luxS* plays some role in the regulation of toxin production in *C. difficile*.

1.5.5 Evidence for *agr* in clostridia

Sequence alignments show that *agr* homologues occur in a number of bacteria besides staphylococci. In particular, they have been found in various clostridial species. They also occur in *Listeria monocytogenes* (Autret *et al.*, 2003), *Lactobacillus plantarum* (Diep *et al.*, 1994) and *Enterococcus faecalis* (Nakayama *et al.*, 2001).

Recent completion of the sequencing of the *C. botulinum* genome at the Sanger Institute, Cambridge, has revealed that it contains two *agrB* and two *agrD* homologues. However, there are currently no data available on the function of these *agr* genes, or whether they are involved in the control of virulence factor production. The investigation into the role of these genes could provide vital information on gene regulation in *C. botulinum*.

1.6 Mutational studies in clostridia

Historically, clostridia have been extremely difficult to manipulate genetically, which has hampered progress on the investigation of gene regulation in this genus. Currently, the only methods of transferring plasmid vectors into clostridia are by electroporation or conjugation from *E. coli* donors. However, electroporation is hampered in some strains of clostridia, for example *Clostridium acetobutylicum*, which possess a restriction system that cuts DNA originating from *E. coli*, or other GC rich DNA (Tummala *et al.*, 2001).

Typically, the study of gene function in prokaryotes is achieved by the mutation of the gene of interest, followed by phenotypic analysis. In many bacteria this is often achieved by a process known as allelic exchange mutagenesis, which relies on homologous recombination between similar sequences in the chromosome and in a plasmid vector. A non-functional copy of the target gene, often containing an antibiotic resistance marker, is introduced into a suicide plasmid, which is unable to replicate in the target host. Once this plasmid is introduced to the host, homologous recombination occurs between the host target gene and the inactive copy on the plasmid. This recombination may result in the inactivation of the chromosomal target gene, and will confer antibiotic resistance to the mutated cell.

Two different types of bacterial mutants can be created using allelic exchange mutagenesis; double crossover and single crossover mutants. Single crossover mutants arise from a single recombination event between the chromosomal target gene and a non-functional copy of the target gene, which is truncated at both the 5' and 3' termini. Figure 1.4 illustrates that, as well as a mutant version of the target gene, the plasmid used for its delivery is also integrated into the chromosome. This event results in the presence of two mutated copies of the gene on the chromosome; one truncated at the 5' end and one truncated at the 3' end (Carter, 2004). However, mutants obtained by this process exhibit instability due to the presence of large repeated flanking regions, and are able to revert to the wild type genotype (Carter, 2004).

The more favourable form of allelic exchange mutagenesis relies on the occurrence of two recombination events, resulting in a double crossover mutant. The target gene in this case is disrupted, usually by an antibiotic resistance cassette, and the vector sequence is excised from the chromosome during a second recombination event (Figure 1.5). This second event, which occurs between homologous regions in the chromosome, may result in the excision of both the vector and mutated target gene (resulting in reversion to the wild type genotype) or the removal of the vector sequence only (resulting in a target gene disrupted by the antibiotic resistance cassette) (Carter, 2004).

Although this form of mutagenesis has routinely been used in many species of eubacteria, it has been largely unsuccessful in the clostridia. In fact, until recently, there were no effective gene knockout techniques available for general use in clostridia. Over the past 15-20 years members of the clostridial scientific community have attempted various methods of gene inactivation; most of them unsuccessful or unreliable. Even though mutants have occasionally been created in clostridial species such as *C. perfringens* (Ohtani *et al.*, 2002a; Huang *et al.*, 2004) and *C. difficle* (Teichert *et al.*, 2006; O'Connor, *et al.*, 2006), knockout mutants in *C. botulinum* are yet to be obtained. In fact, even now, genetic manipulation in *C. botulinum* is in its infancy compared with that of *C. perfringens* and *C. acetobutylicum* (Bradshaw *et al.*, 1998).

For many years, the clostridial scientific community has been in dire need of a highly efficient clostridial universal knockout system, in order to rapidly advance the molecular genetic study of this genus.

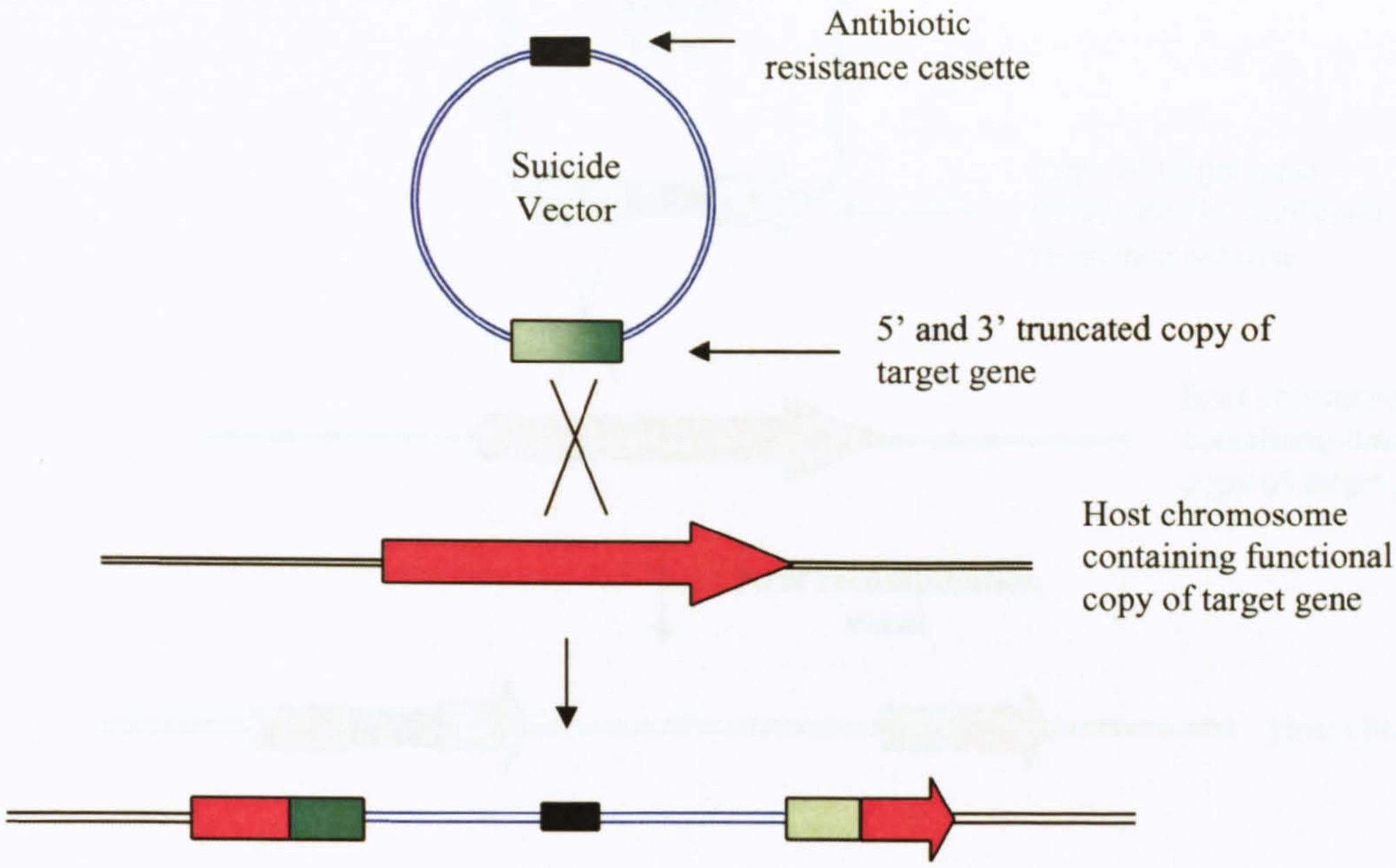


Figure 1.4 – Schematic representation of single crossover recombination (adapted from Carter, 2004 and Pennington, 2006). A non-functional copy of the target gene, truncated at the 5' and 3' termini, is introduced to the host on a plasmid. The whole plasmid is integrated into the chromosome by a single crossover recombination event, resulting in the presence of two truncated copies of the gene in the chromosome.

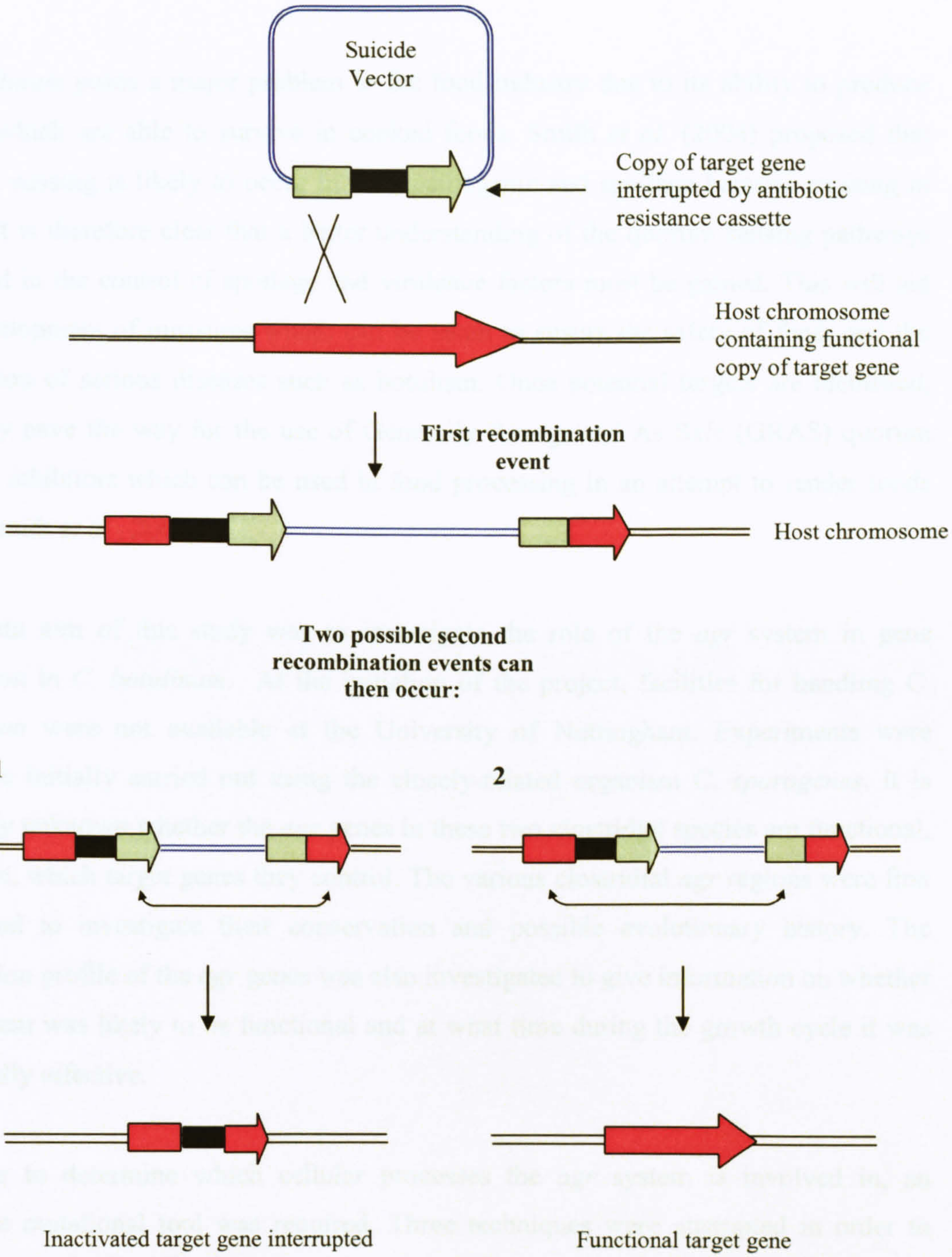


Figure 1.5 – Schematic representation of double crossover recombination (adapted from Carter, 2004). A copy of the target gene interrupted by an antibiotic resistance cassette is introduced into the host on a plasmid and a single crossover event occurs as shown in Figure 1.4. However, a second recombination event occurs, resulting in either: **1** – the removal of the vector sequence only, leaving the target gene disrupted by the antibiotic resistance cassette; or **2** - the excision of both the vector and mutated target gene and reversion to the wild type genome.

1.7 Project aims

C. botulinum poses a major problem to the food industry due to its ability to produce spores which are able to survive in cooked foods. Smith *et al.* (2004) proposed that quorum sensing is likely to occur in both pathogenic and spoilage bacteria existing in foods. It is therefore clear that a better understanding of the quorum sensing pathways involved in the control of spoilage and virulence factors must be gained. This will aid the development of measures which can be taken to ensure the safety of food, and the prevention of serious diseases such as botulism. Once potential targets are identified, this may pave the way for the use of Generally Recognised As Safe (GRAS) quorum sensing inhibitors which can be used in food processing in an attempt to render foods safer (Smith *et al.*, 2004).

The main aim of this study was to investigate the role of the *agr* system in gene regulation in *C. botulinum*. At the initiation of the project, facilities for handling C. *botulinum* were not available at the University of Nottingham. Experiments were therefore initially carried out using the closely-related organism *C. sporogenes*. It is currently unknown whether the *agr* genes in these two clostridial species are functional, and if so, which target genes they control. The various clostridial *agr* regions were first compared to investigate their conservation and possible evolutionary history. The expression profile of the *agr* genes was also investigated to give information on whether the system was likely to be functional and at what time during the growth cycle it was maximally effective.

In order to determine which cellular processes the *agr* system is involved in, an effective mutational tool was required. Three techniques were attempted in order to produce mutants in *C. sporogenes* and *C. botulinum*, which could be used to assess the functional role of the *agr* systems. The primary phenotype under investigation was sporulation.

Chapter 2.0

Materials and Methods

Chapter 2 - Materials and Methods

2.1 Media

All media were sterilised by autoclaving at 121°C and 15 psi for 20 minutes unless otherwise stated.

2.1.1 Luria-Bertani medium

Luria-Bertani (LB) broth contained 10 g of Tryptone (Difco Laboratories), 5 g of Yeast Extract (Difco Laboratories) and 5 g sodium chloride (Sigma Ltd) in 1 l of distilled water. LB agar was prepared by the addition of 1% (w/v) no. 1 Bacteriological agar (Oxoid).

2.1.2 TYG medium

TYG medium contained 30 g of Tryptone (Difco Laboratories), 20 g of Yeast Extract (Difco Laboratories) and 1 g of sodium thioglycollate (Sigma) in 1 l of distilled water. Sterile glucose was added to a final concentration of 0.5% (w/v) after autoclaving if necessary. TYG agar was prepared by the addition of 1% (w/v) no. 1 Bacteriological agar (Oxoid).

2.1.3 Brain Heart Infusion (BHI) agar

BHI agar contained 47 g BHI powder (Oxoid) in 1 l distilled water.

2.1.4 Spizizen's Minimal Salts Medium (SMM)

SMM was prepared according to Anagnostopoulos & Spizizen (1961) and contained 14 g anhydrous dibasic potassium phosphate, 6 g anhydrous monobasic potassium phosphate, 1 g sodium citrate and 0.2 g magnesium sulphate (heptahydrate) in 1 l distilled water. The pH was adjusted to 7.0 and the medium filter sterilised.

2.1.5 *B. subtilis* Competence Medium

Competence medium was prepared according to Anagnostopoulos & Spizizen (1961) and contained 10 ml SMM (see 2.1.4), 0.5 ml 10% glucose solution, 5 µl 0.1 M calcium chloride solution, 0.1 ml 250 mM magnesium sulphate solution, 0.1 ml 50 nM manganese chloride solution, 0.2 ml 10% casamino acids and 0.25 ml 10 mM tryptophan. The individual components were filter sterilised prior to use.

2.1.6 *B. subtilis* Transformation Medium

Transformation medium was prepared according to Anagnostopoulos & Spizizen (1961) and contained 1 ml SMM (see 2.1.4), 50 µl 10% (w/v) glucose solution, 20 µl 250 mM magnesium sulphate solution, 1 µl 10% (w/v) casamino acids and 25 µl 10 mM tryptophan. The individual components were filter sterilised prior to use.

2.1.7 5 X Minimal salts

5 X minimal salts was prepared as described by Bron (1990) and consisted of 10 g ammonium sulphate, 74 g anhydrous dibasic potassium phosphate, 27 g anhydrous monobasic potassium phosphate, 9.5 g trisodium citrate and 1 g magnesium sulphate heptahydrate. The solution was made up to 1 l with distilled water and the pH adjusted to 7.0.

2.1.8 *B. subtilis* Minimal growth medium

Minimal growth medium was prepared as described by Bron (1990) and consisted of 10 ml 5 X minimal salts, 500 µl 50% (w/v) glucose, 500 µl 2% (w/v) casamino acids, 100 µl 10 mg/ml L-tryptophan and 50 µl 2.2 mg/ml ferric ammonium citrate. The solution was made up to 50 ml with distilled water and filter sterilised prior to use.

2.1.9 *B. subtilis* Starvation medium

Starvation medium was prepared as described by Bron (1990) and consisted of 10 ml 5 X minimal salts and 500 µl 50% (w/v) glucose. The solution was made up to 50 ml with distilled water and filter sterilised prior to use.

2.2 Supplements

Media supplements were prepared and stored according to the manufacturer's instructions. Antibiotics were used at the following working concentrations:

Ampicillin	100 µg/ml
Erythromycin	500 µg/ml (*E. coli*), 10 µg/ml (*C. sporogenes* and *Bacillus*), 20 µg/ml (*C. botulinum*)
Cycloserine	250 µg/ml
Chloramphenicol	25 µg/ml
Thiamphenicol	15 µg/ml (*C. sporogenes*), 20 µg/ml (*C. botulinum*)

5-bromo-4-chloro-3 indolyl-β-D-galactoside (X-Gal) was used at 80 µg/ml.

2.3 Growth Conditions

All *E. coli* cultures were grown aerobically on LB medium and *Bacillus* on BHI or LB. Both were incubated at 37°C. Liquid cultures of *E. coli* were grown in LB broth with agitation at 200 rpm. *E. coli* and *Bacillus* strains were frozen at –80°C as bead stocks (PRO-LAB Diagnostics).

C. sporogenes was grown in an anaerobic cabinet (MK3 Anaerobic Work Station, Don Whitley Scientific Ltd) containing an atmosphere of 80% nitrogen, 10% hydrogen and 10% carbon dioxide. TYG medium containing cycloserine was used unless otherwise stated.

C. botulinum was grown in a MG1000 Anaerobic Work Station (Don Whitley Scientific Ltd) on TYG medium containing cycloserine.

Clostridial strains were maintained in cooked meat broths (Oxoid) and stored at room temperature. Spore deficient strains were kept in glycerol stocks at –80°C.

Growth of liquid cultures was monitored by measuring optical density at 600 nm (A_{600}) using a Pharmacia LKB Novaspec II.

2.4 Bacterial strains

Strain	Genotype/Description	Reference/Source
Clostridium sporogenes NCIMB 10696	Wild type strain	N. P. Minton, University of Nottingham, UK
Clostridium botulinum ATCC 3502	Wild type strain	M. W. Peck, Institute of Food Research, Norwich, UK
Clostridium sporogenes cured strain 10696-1	pIM13-based plasmid introduced into *C. sporogenes* 10696 wild type, which was subsequently cured of plasmid	Pennington, 2006
Clostridium sporogenes 0336 mutant	Group II intron inserted into CSP0336 gene	This study
Clostridium sporogenes 0340 mutant	Group II intron inserted into CSP0340 gene	This study
*Clostridium botulinum agrD*1 mutant	Group II intron inserted into agrD1 gene	This study
*Clostridium botulinum agrD*2 mutant	Group II intron inserted into agrD2 gene	This study
Clostridium botulinum 0336 mutant	Group II intron inserted into CBO0336 gene	This study
Clostridium botulinum 0340 mutant	Group II intron inserted into CBO0340 gene	This study
Bacillus subtilis Cu2189		A. Roberts, University College, London, UK
Escherichia coli TOP10	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZM15* Δ*lacX74 deoR recA1 araD139* Δ(*ara-leu*)*7697 galU galK rpsL* (Str^R) *endA1 nupG*	Invitrogen Ltd
Escherichia coli CA434 (HB101)	*thi-1 hsdS20* ($r-_B$, $m-_B$) *supE44 recAB ara-14 leuB5proA2 lacY1 galK rpsL20* (str^R) *xyl-5 mtl-1*	M. Young, UCW, Aberystwyth, UK

2.5 Plasmids

Plasmid	Genotype/Description	Reference/Source
pCR2.1-TOPO	*E. coli* PCR cloning vector, ColE1, Amp^R, Kan^R	Invitrogen
pMTL5112	Clostridial shuttle expression vector, ColE1, pIM13, Erm^R	Pennington, 2006
pMTL21Eβ10*lacIfacoriT*	Inducible clostridial vector, promoterless pAMβ1 replicon with upstream *fac*, *ptb*::*lacI* fusion, Erm^R	Carter, 2004
pMTL23	*E. coli* cloning vector, ColE1, Amp^R	Chambers *et al.*, 1988
pMTL5112*lacI*	Clostridial shuttle expression vector with *repL* under inducible control, ColE1, pIM13, Erm^R	This study
pCR2.1::*catP*Δ*Bsr*GI	PCR amplified *C. perfringens catP* (*AgeI*/*Pvu*II fragment) from pJIR418 with internal *Bsr*GI site deleted, ColE1, Amp^R, Kan^R	Pennington, 2006
pCR2.1::Cspo*pyrF*	PCR amplified *C. sporogenes pyrF* fragment, ColE1, Amp^R, Kan^R	Pennington, 2006
pMTL5112*lacI*::*pyrF*/*catP*	Clostridial shuttle expression vector with *repL* under inducible control and *pyrF* knockout cassette inserted ,ColE1, pIM13, Erm^R	This study
pMTL9361	Inducible clostridial expression vector, pCD6 replicon, P_{ptb}::*lacI*, P_{fac}, *lacZ*, Erm^R	Carter, 2004
pMTL9361::Cs*agrB*1AS1	Inducible clostridial expression vector containing *C. sporogenes agrB*1 small antisense *Not*I/*Kpn*I fragment	This study
pMTL9361::Cs*agrB*1AS2	Inducible clostridial expression vector containing *C. sporogenes agrB*1 large antisense *Not*I/*Kpn*I fragment	This study

Plasmid	Genotype/Description	Reference/Source
pMTL9361::Cs*agrB*2AS1	Inducible clostridial expression vector containing *C. sporogenes* *agrB*2 small antisense *Not*I/*Kpn*I fragment	This study
pMTL9361::Cs*agrB*2AS2	Inducible clostridial expression vector containing *C. sporogenes* *agrB*2 large antisense *Not*I/*Kpn*I fragment	This study
pMTL9361::Cs336AS1	Inducible clostridial expression vector containing *C. sporogenes* 0336 small antisense *Not*I/*Kpn*I fragment	This study
pMTL9361::Cs336AS2	Inducible clostridial expression vector containing *C. sporogenes* 0336 large antisense *Not*I/*Kpn*I fragment	This study
pMTL9361::Cs340AS1	Inducible clostridial expression vector containing *C. sporogenes* 0340 small antisense *Not*I/*Kpn*I fragment	This study
pMTL9361::Cs340AS2	Inducible clostridial expression vector containing *C. sporogenes* 0340 large antisense *Not*I/*Kpn*I fragment	This study
pMTL007	Inducible clostridial expression vector for expression of ClosTron, containing Erm RAM, ColE1, pCB102, Cm^R	Heap *et al.*, 2007
pMTL007C-E2	Constitutive clostridial expression vector for expression of ClosTron containing Erm RAM, flanked by FRT sites, ColE1, pCB102, Cm^R	J. T. Heap, University of Nottingham
pCP20	Vector expressing Flp recombinase from *Saccharomyces cerevisiae*, Cm^R, Amp^R	S. Pommier, University of Nottingham

Plasmid	Genotype/Description	Reference/Source
pMTL5402F	Inducible clostridial expression vector, ColE1, pCB102, Cm^R	Heap *et al.*, 2007
pMTL5402F-FLP	Inducible clostridial expression vector for expression of Flp, ColE1, pCB102, Cm^R	This study
pMTL007::Cbo*agrD*1-23a	Inducible clostridial expression vector for expression of ClosTron containing intron retargeted to *C. botulinum agrD*1	This study
pMTL007::Cbo*agrD*2-47a	Inducible clostridial expression vector for expression of ClosTron containing intron retargeted to *C. botulinum agrD*2	This study
pMTL007::Cbo0336-1203s	Inducible clostridial expression vector for expression of ClosTron containing intron retargeted to *C. botulinum* 0336	This study
pMTL007::Cbo0340-690s	Inducible clostridial expression vector for expression of ClosTron containing intron retargeted to *C. botulinum* 0340	This study
pMTL007C-E2::*Cboagr D*1-23a	Constitutive clostridial expression vector for expression of ClosTron containing intron retargeted to *C. botulinum agrD*1	This study
pMTL007C-E2::Csp0336-1207s	Constitutive clostridial expression vector for expression of ClosTron containing intron retargeted to *C. sporogenes* 0336	This study

Plasmid	Genotype/Description	Reference/Source
pMTL007C-E2::Csp0340-711s	Constitutive clostridial expression vector for expression of ClosTron containing intron retargeted to *C. sporogenes* 0340	This study

2.6 Oligonucleotide primers

Restriction sites are shown in bold

2.6.1 Primers used in Chapter 3

Primer Name	Sequence (5' – 3')	Use
CsTM331/1F	CAGCTTATTTGTGAGGGC	RT-PCR linkage of *C. sporogenes* 0331/*agrB*1 (forward)
CsTMAgrB1/1R	GGATATCAGCAAAGCCTC	RT-PCR linkage of *C. sporogenes* 0331/*agrB*1 (reverse)
CsTMAgrB1/1F	GTGTATATAGTGGAATAGTATGGC	RT-PCR linkage of *C. sporogenes agrB*1/*agrD*1 (forward)
CsTMAgrD1/1R	CTTCAGGTTGATACACGC	RT-PCR linkage of *C. sporogenes agrB*1/*agrD*1 (reverse)
CsTMAgrD1/1F	GGTATTAATGCTAGTTGCAAC	RT-PCR linkage of *C. sporogenes agrD*1/0333 (forward)
CsTM333/1R	GAACCACAAATGTTACTGATATC	RT-PCR linkage of *C. sporogenes agrD*1/0333 (reverse)
CsTM337/1F	GCTAAGTGAGGCTTTAAATTTG	RT-PCR linkage of *C. sporogenes* 0337/*agrB*2 (forward)
CsTMAgrB2/1R	GTATTCCTCCTGAATATTTCC	RT-PCR linkage of *C. sporogenes* 0337/*agrB*2 (reverse)
CsTM336/1F	GGGGAAGTTACTATAAATAGTGAAC	RT-PCR linkage of *C. sporogenes* 0336/0337 (forward)
CsTM337/1R	CAAATTTAAAGCCTCACTTAGC	RT-PCR linkage of *C. sporogenes* 0336/0337 (reverse)
CsTMAgrB2/1F	GAGGATGTAGATGAAAAGTTGAG	RT-PCR linkage of *C. sporogenes agrB*2/*agrD*2 (forward)
CsTMAgrD2/1R	GTTCTTTAGGTTGATATGCAC	RT-PCR linkage of *C. sporogenes agrB*2/*agrD*2 (reverse)
CsTMAgrD2/2F	CAATTAAAAGAAAAATGTG	RT-PCR linkage of *C. sporogenes agrD*2/0340 (forward)
CsTM340/2R	GCAATTATAATTAGCCATAG	RT-PCR linkage of *C. sporogenes agrD*2/0340 (reverse)
CsTM330/1F	GGTTATAGTGTTATGAGTATTGGAG	RT-PCR linkage of *C. sporogenes* 0330/0331 (forward)
CsTM331/1R	GAATACCATAACTGATTGGC	RT-PCR linkage of *C. sporogenes* 0330/0331 (reverse)
CsTM335/1F	GGTCATTACTGAGTCTTACACC	RT-PCR linkage of *C. sporogenes* 0335/0336 (forward)
CsTM336/1R	CTTCCTGACATTATTATTATGACTG	RT-PCR linkage of *C. sporogenes* 0335/0336 (reverse)
CsRTagrD1/1F	CTTGCATCCGTAGTAGCATCATC	Real time PCR on *C. sporogenes agrD*1 (forward)
CsRTagrD1/1R	TTTTGGCTCTTCAGGTTGATACA	Real time PCR on *C. sporogenes agrD*1 (reverse)
CsRTagrD2/1F	AAAAATGTGTAAAGGTTACTGCAAAA	Real time PCR on *C. sporogenes agrD*2 (forward)
CsRTagrD2/1R	ATGCACCAAATACACAAGCTGAA	Real time PCR on *C. sporogenes agrD*2 (reverse)

Primer Name	Sequence (5' – 3')	Use
CsRT16S/1F	CTGCAACTCGCCTACATGAAG	Real time PCR on *C. sporogenes* 16S (forward)
CsRT16S/1R	CTGCAACTCGCCTACATGAAG	Real time PCR on *C. sporogenes* 16S (reverse)

2.6.2 Primers used in Chapter 4

Primer Name	Sequence (5' – 3')	Use
RepL-P1	TCAG**TCTAGA**CCTCCCAATAACTACGTGGTG	Amplification of *repL* gene for integrative vector construction (forward)
RepL-P2	GACT**GATATC**ATAGTTTAAAATAAGGAGTGAG	Amplification of *repL* gene for integrative vector construction (reverse)
Ori-P1	ACTA**GCTAGC**AAGCTTCTAACTATATCATAACTGTTC	Amplification of + origin for integrative vector construction (forward)
Ori-P2	TCAT**GTCGAC**TATATAAATATGAGCGAAGCG	Amplification of + origin for integrative vector construction (reverse)
RNA/2F	GAGTTGGTAGCTCTTGATCC	Screening of integrative vector
PyrC/1F(R)	CTCCTATGGATGGTATGTG	Screening of integrative vector
PyrC/2F	GGTATTAATAGAGGCAATAAAGG	Screening of integrative vector
PyrK/1R(F)	GGTCTCCATAAAATAGGCTC	Screening of integrative vector
Ori/2R	CCATGCTCCAACAGC	Screening of integrative vector
AgrB1NotI/1F	TATT**GCGGCCGC**GAGAAGAGCTGCAATATGATTAATGCAG	Amplification of antisense fragments 1 and 2 of *C. sporogenes agrB*1 (forward)
AgrB1KpnI/1R	AGGA**GGTACC**GCTTATTGTAAAGGATATCAGCAAAGCCTC	Amplification of antisense fragment 1 of *C. sporogenes agrB*1 (reverse)
AgrB1KpnI/2R	AGGA**GGTACC**CTTTATAGGTTTAGCCTTAGAGTCTACAGG	Amplification of antisense fragment 2 of *C. sporogenes agrB*1 (reverse)
AgrB2NotI/2F	TATTGCGGCCGCAGGAGAGTAAAGGATGTTTTTTATTG	Amplification of antisense fragments 1 and 2 of *C. sporogenes agrB*2 (forward)
AgrB2KpnI/3R	AGGA**GGTACC**CCCTCCTGAATATTTCCTATATATCGC	Amplification of antisense fragment 1 of *C. sporogenes agrB*2 (reverse)
AgrB2KpnI/4R	AGGA**GGTACC**CTCAACTTTTCATCTACATCCTC	Amplification of antisense fragment 2 of *C. sporogenes agrB*2 (reverse)
Cs336NotI/1F	TATTGCGGCCGCAGGAGTTAAATTAATGAATTTAG	Amplification of antisense fragments 1 and 2 of *C. sporogenes* 0336 (forward)

Primer Name	Sequence (5' – 3')	Use
Cs336KpnI/1R	AGGA**GGTACC**CCTTCTGGTCTAATTCTGTAG	Amplification of antisense fragment 1 of *C. sporogenes* 0336 (reverse)
Cs336KpnI/2R	AGGA**GGTACC**CTTCCTTTTCATAGTCAACG	Amplification of antisense fragment 2 of *C. sporogenes* 0336 (reverse)
Cs340NotI/1F	TATT**GCGGCCGC**ACTATATGTTTATATATTAATATGACAG	Amplification of antisense fragments 1 and 2 of *C. sporogenes* 0340 (forward)
Cs340KpnI/1R	AGGA**GGTACC**CTTGTTTAAGTTTTCTAATTGG	Amplification of antisense fragment 1 of *C. sporogenes* 0340 (reverse)
Cs340KpnI/2R	AGGA**GGTACC**CTTTTTATCATAGGAGTTGC	Amplification of antisense fragment 2 of *C. sporogenes* 0340 (reverse)
CsAgrB1AS/1F	CAATATCCAATTCTGTTGC	Confirmation of antisense RNA production in *C. sporogenes* *agrB*1 mutants (forward)
CsAgrB2AS/1F	GGAAGTGAAATCTCTAGTAATTTAAGC	Confirmation of antisense RNA production in *C. sporogenes* *agrB*2 mutants (forward)
Cs336AS/1F	TAAAAATTGGAGTCATGAC	Confirmation of antisense RNA production in *C. sporogenes* 0336 mutants (forward)
Cs340AS/1F	GTGTGAGGTATATGAGCG	Confirmation of antisense RNA production in *C. sporogenes* 0340 mutants (forward)
GC4FSeq	GATACAATAAGTTATGGTTG	Screening of pMTL9361 for presence of antisense fragments
CodAS1F	CTTATGATTAAAATTTTAAGGAGGTG	Confirmation of antisense RNA production in *C. sporogenes* antisense mutants (reverse)
Rep/1F	CCTCCCAATAACTACGTG	Sequencing of *repL* gene in integrative vector pMTL5112lacI (forward)
Rep/1R	CGGATAACAATTCCAAC	Sequencing of *repL* gene in integrative vector pMTL5112lacI (reverse)
LacI/1F	GAAGCTTGCTAGAATTCC	Sequencing of *lacI* gene in integrative vector pMTL5112lacI (forward)

Primer Name	Sequence (5' – 3')	Use
LacI/1R	GAGGTGCAACATATGAAAC	Sequencing of *lacI* gene in integrative vector pMTL5112lacI (reverse)
Cs340KpnI/1R	AGGA**GGTACC**CTTGTTTAAGTTTTCTAATTGG	Amplification of antisense fragment 1 of *C. sporogenes* 0340 (reverse)
Cs340KpnI/2R	AGGA**GGTACC**CTTTTTATCATAGGAGTTGC	Amplification of antisense fragment 2 of *C. sporogenes* 0340 (reverse)
CsAgrB1AS/1F	CAATATCCAATTCTGTTGC	Confirmation of antisense RNA production in *C. sporogenes agrB*1 mutants (forward)
CsAgrB2AS/1F	GGAAGTGAAATCTCTAGTAATTTAAGC	Confirmation of antisense RNA production in *C. sporogenes agrB*2 mutants (forward)
Cs336AS/1F	TAAAAATTGGAGTCATGAC	Confirmation of antisense RNA production in *C. sporogenes* 0336 mutants (forward)
Cs340AS/1F	GTGTGAGGTATATGAGCG	Confirmation of antisense RNA production in *C. sporogenes* 0340 mutants (forward)
GC4FSeq	GATACAATAAGTTATGGTTG	Screening of pMTL9361 for presence of antisense fragments
CodAS1F	CTTATGATTAAAATTTTAAGGAGGTG	Confirmation of antisense RNA production in *C. sporogenes* antisense mutants (reverse)

2.6.3 Primers used in Chapter 5

Primer Name	Sequence (5' – 3')	Use
Cb-agrD1-2831a-IBS	AAAAAAGCTTATAATTATCCTTAACTAACATTAATGTGCGCCCAGATAGGGTG	Re-targeting PCR for generation of *C. botulinum agrD*1 mutant
Cb-agrD1-2831a-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCATTAATACTAACTTACCTTTCTTTGT	Re-targeting PCR for generation of *C. botulinum agrD*1 mutant
Cb-agrD1-2831a-EBS2	TGAACGCAAGTTTCTAATTTCGATTTTAGTTCGATAGAGGAAAGTGTCT	Re-targeting PCR for generation of *C. botulinum agrD*1 mutant
Cb-agrD2-2990a-IBS	AAAAAAGCTTATAATTATCCTTAACAGACTTAAGTGTGCGCCCAGATAGGGTG	Re-targeting PCR for generation of *C. botulinum agrD*2 mutant
Cb-agrD2-2990a-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCTTAAGTAATAACTTACCTTTCTTTGT	Re-targeting PCR for generation of *C. botulinum agrD*2 mutant
Cb-agrD2-2990a-EBS2	TGAACGCAAGTTTCTAATTTCGATTTCTGTTCGATAGAGGAAAGTGTCT	Re-targeting PCR for generation of *C. botulinum agrD*2 mutant
Cb-336-617s-IBS	AAAAAAGCTTATAATTATCCTTAAAACACGATAAAGTGCGCCCAGATAGGGTG	Re-targeting PCR for generation of *C. botulinum* 0336 mutant
Cb-336-617s-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGATAAATATAACTTACCTTTCTTTGT	Re-targeting PCR for generation of *C. botulinum* 0336 mutant
Cb-336-617s-EBS2	TGAACGCAAGTTTCTAATTTCGGTTTGTTTCCGATAGAGGAAAGTGTCT	Re-targeting PCR for generation of *C. botulinum* 0336 mutant
Cb-340-776s-IBS	AAAAAAGCTTATAATTATCCTTAAGAGACGCTATTGTGCGCCCAGATAGGGTG	Re-targeting PCR for generation of *C. botulinum* 0340 mutant
Cb-340-776s-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGCTATTATTAACTTACCTTTCTTTGT	Re-targeting PCR for generation of *C. botulinum* 0340 mutant
Cb-340-776s-EBS2	TGAACGCAAGTTTCTAATTTCGATTTCTCTTCGATAGAGGAAAGTGTCT	Re-targeting PCR for generation of *C. botulinum* 0340 mutant
Cs-336-617s-IBS	AAAAAAGCTTATAATTATCCTTAAAGGACGATAAAGTGCGCCCAGATAGGGTG	Re-targeting PCR for generation of *C. sporogenes* 0336 mutant

Primer Name	Sequence (5' – 3')	Use
Cs-336-617s-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGATAAATATAACTTACCTTTCTTTGT	Re-targeting PCR for generation of *C. sporogenes* 0336 mutant
Cs-336-617s-EBS2	TGAACGCAAGTTTCTAATTTCGATTTCCTTTCGATAGAGGAAAGTGTCT	Re-targeting PCR for generation of *C. sporogenes* 0336 mutant
Cs-340-776s-IBS	AAAAAAGCTTATAATTATCCTTAAGGGACGCTATTGTGCGCCCAGATAGGGTG	Re-targeting PCR for generation of *C. sporogenes* 0340 mutant
Cs-340-776s-EBS1d	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGCTATTATTAACTTACCTTTCTTTGT	Re-targeting PCR for generation of *C. sporogenes* 0340 mutant
Cs-340-776s-EBS2	TGAACGCAAGTTTCTAATTTCGATTTCCCTTCGATAGAGGAAAGTGTCT	Re-targeting PCR for generation of *C. sporogenes* 0340 mutant
EBS universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	Universal primer for re-targeting of group II intron
Cb-agrD1/R	GCACCAATAACACGCTGATG	Screening of *C. botulinum agrD*1 mutant (reverse)
Cb-agrD2/R	GATTTTGGTTCTTTGGGTTG	Screening of *C. botulinum agrD*2 mutant (reverse)
Cb-agrD1/F	GGTAGATATTAAAAAGGGGGATAAG	Screening of *C. botulinum agrD*1 mutant (forward)
Cb-agrD2/F	GTACAAAGGTTACTGCAAAATTAC	Screening of *C. botulinum agrD*2 mutant (forward)
Cb0336R	CTTAAATATCCTGAATCTAGTTTAGTTG	Screening of *C. botulinum* 0336 mutant (reverse)
Cb0340R	CTTGCTCAACACATACTATATCTCTC	Screening of *C. botulinum* 0340 mutant (reverse)
Cb0336F	CAAAATACATCTACTTATGTTATATATGATGGAATG	Screening of *C. botulinum* 0336 mutant (forward)
Cb0340F	CTTATTTTGAAAATGATTTAATTCTCTCAGG	Screening of *C. botulinum* 0340 mutant (forward)
Cs0336F	CAAAATACATCTACTTATGTTATATATAATGGAATG	Screening of *C. sporogenes* 0336 mutant (forward)
Cs0336R	TTTAAATATCCTGAATCTAATTTAGTTG	Screening of *C. sporogenes* 0336 mutant (reverse)
Cs0340F	CTTATTTTGAAAATGATTTAATTCTTTCGGG	Screening of *C. sporogenes* 0340 mutant (forward)
Cs0340R	GTATAATTTGCTCAACACATATCATATCAG	Screening of *C. sporogenes* 0340 mutant (reverse)
5402F-F1	TTAAGGAGGTGTATTTCATATGACCATGATTACG	Sequencing of intron region of pMTL007
pMTL007-R1	AGGGTATCCCCAGTTAGTGTTAAGTCTTGG	Sequencing of intron region of pMTL007 & pMTL007CE-2
Spofdx-seq-F1	GATGTAGATAGGATAATAGAATCCATAGAAAATATAGG	Sequencing of intron region of pMTL007CE-2
ErmRAM-F	ACGCGTTATATTGATAAAAATAATAATAGTGGG	Screening of Erm RAM in *C. botulinum* mutants (forward)
ErmRAM-R	ACGCGTGCGACTCATAGAATTATTTCCTCCCG	Screening of Erm RAM in *C. botulinum* mutants (reverse)

Primer Name	Sequence (5' – 3')	Use
FLP-F1	AATCTAGAAGGAGGTACTAGATGCCACAATTTGGTATATTA TGTAAAACACCACC	Amplifying *S. cerevisiae flp* gene from plasmid pCP20 (forward)
FLP-R1	TTCTCGAGTTATATGCGTCTATTTATGTAGGATGAAAGGTA GTCTAGTACC	Amplifying *S. cerevisiae flp* gene from plasmid pCP20 (reverse)

2.7 Preparation and manipulation of DNA

2.7.1 Preparation of plasmid DNA

Plasmid DNA was prepared using the Qiagen QIAprep Miniprep kit (Qiagen Ltd. UK) following the manufacturer's instructions. An additional step was included for Gram-positive bacteria; 20 mg/ml lysozyme (Sigma) was added to buffer P1 before resuspending the bacterial pellet. The lysate was incubated at 37°C for 5 minutes and then the Qiagen protocol continued as normal.

2.7.2 Preparation of chromosomal DNA (*E. coli* and *C. sporogenes*)

Chromosomal DNA was prepared using the Qiagen DNeasy Tissue kit (Qiagen Ltd. UK) following the manufacturer's instructions. An additional step was added after the 30-minute incubation at 37°C in lysis buffer. 4 µl 100 mg/ml RNase was added and the samples incubated at room temperature for 2 minutes.

2.7.3 Preparation of chromosomal DNA (*C. botulinum*)

Cells from 1-9 ml overnight *C. botulinum* culture were harvested by centrifugation at 5000 X g for 10 minutes, and resuspended in 180 µl enzymatic lysis buffer (20 mM Tris-Cl, pH 8, 2 mM sodium EDTA, 1.2% Triton X-100, 20 mg/ml lysozyme). The suspension was incubated at 37°C for 30 minutes. A 25 µl aliquot of proteinase K (600 mAU/ml) from the Qiagen DNeasy Tissue Kit was added, followed by 200 µl Buffer AL from the same kit. The cell suspensions were incubated at 73°C for 30 minutes, an equal volume of phenol:chloroform:isoamylalcohol (25:24:1 - Sigma) added and the lysate mixed by flicking the tube. The lysate was incubated for 15 minutes at room temperature, mixing the contents every 5 minutes to inactivate botulinum toxin. Tubes were centrifuged for 1 minute at 16,000 X *g* and the top aqueous phase transferred to a new tube. The efficacy of phenol destruction of spores was tested by plating 100 µl of the aqueous phase onto a TYG plate and incubating overnight. The aqueous phase was stored at –20°C overnight.

After overnight incubation, if no growth was visible on the TYG plate, the protocol was completed. 4 µl RNaseA (100 mg/ml) was added to the aqueous phase, vortexed and incubated at room temperature for 2 minutes. An equal volume of chloroform was added and centrifuged for 3 minutes at 16,000 X *g*. The aqueous phase was removed to a new tube. A 1/10 volume of 3 M sodium acetate, pH 5.2 was added to the solution of DNA and mixed by vortexing briefly. A total of 2.5 volumes of ice cold 100% ethanol was added, mixed by vortexing and placed at -80°C for 1 hour. Tubes were then centrifuged for 5 minutes at 16,000 X *g* and the supernatant discarded. A 1 ml aliquot of room temperature 70% ethanol was added, the tube mixed by inversion and centrifuged as before. The supernatant was discarded and the pellet air-dried before resuspending in TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8).

2.7.4 Restriction enzyme digests

Restriction enzymes were supplied by New England Biolabs Inc. and were used according to the manufacturer's instructions. Digest reactions contained 0.5-1 µg DNA, 1 µl restriction enzyme and 1 X reaction buffer made up to a final volume of 20 µl with sterile distilled water. Reactions were incubated at 37°C for 3 hours until digestion was complete (except for *Sma*I, which was incubated at 25°C).

2.7.5 Blunt-ending of DNA

Blunt-ending of DNA was carried out using T4 DNA polymerase or Klenow polymerase (New England Biolabs Inc). Both enzymes were used at 1/50 dilution in the supplied buffer, along with 1 µl 100 mM dNTPs. The reactions were incubated at 12°C (T4 DNA polymerase) or 25°C (Klenow) for 15 minutes.

2.7.6 De-phosphorylation of linearised DNA fragments

Where necessary, linearised DNA fragments were de-phosphorylated using alkaline phosphatase (New England Biolabs Inc.). The enzyme was used at 1/50 dilution in the supplied buffer and reactions were incubated at 37°C for 1 hour.

2.7.7 Analysis of DNA by agarose gel electrophoresis

Agarose gels were made up using 1% (w/v) agarose (VWR International Ltd) in 1 x TAE buffer (40 mM Tris, 1 mM EDTA and 0.1% (v/v) glacial acetic acid). Ethidium bromide was added to a concentration of 10 μg/ml. Gels were run at 80-100 V in 1 X TAE buffer. Before loading, 5 μl sample was added to 1 μl 6 X loading buffer (40% (w/v) sucrose, 0.25% (w/v) bromophenol blue). DNA was visualised on a UV transillumintor and a 1 Kb DNA ladder (New England Biolabs Inc) used at 0.5 μg/μl.

2.7.8 Purification of DNA fragments

PCR fragments were purified from agarose gel by extracting with a scalpel and using a Qiagen QIAquick Gel Extraction kit according to the manufacturer's instructions. PCR products were purified using the Qiagen QIAquick PCR Purification kit according to the manufacturer's instructions.

2.7.9 Polymerase Chain Reaction

Oligonucleotide primers were synthesised by Sigma-Genosys UK or Operon UK. Epicentre (Wisconsin, USA) PCR reagents were used and each reaction contained: 0.5 μl (1.25 units) FailSafe PCR enzyme mix (containing a proof-reading *Taq* DNA polymerase); 12.5 μl Failsafe PCR 2 X buffer E; 1 μl of each oligonucleotide primer (10 μM stock), 1 μl DNA template (100 ng) and distilled water to give a final reaction volume of 25 μl. Reactions were carried out in a Biometra T3000 thermocycler.

The following program was used: initial heating to 94°C for 4 minutes, then 30 cycles of the following: denaturation at 94°C for 1 minute, annealing at an appropriate temperature for 1 minute, and elongation at 72°C for an appropriate time. The final step in the reaction was a hold at 72°C for 7 minutes. The annealing temperature used was dependent on the melting temperatures of the primers used, and was set at 2-3°C below the melting temperature. The elongation time was dependent on the length of the template (1 minute per kilobase).

2.7.10 DNA ligation

Digested, purified vectors and inserts were ligated together at insert:vector ratios of 3:1 and 5:1. A control ligation without insert was also set up to assess any re-ligation of the vector. Ligation reactions were set up using 1 µl T4 DNA ligase and 1 X ligation buffer (New England Biolabs Inc.) and made up to 10 µl with sterile distilled water. Reactions were incubated at room temperature for 1 hour or overnight at 4°C.

2.7.11 TOPO TA cloning methods

The TOPO TA cloning kit (Invitrogen Life Technologies) was used, according to the manufacturer's instructions, to clone PCR fragments into the pCR2.1-TOPO vector. TOPO vectors are provided linearised with topoisomerase covalently bound to the 3' phosphate on each strand. This creates ddT overhangs, which are directly compatible with the deoxyadenosine overhangs of the 3' end of the PCR products (added by *Taq* polymerase). Ligation reactions were carried out as follows: 4 µl PCR product, 1 µl 1:4 salt solution and 1 µl pCR2.1-TOPO vector containing topoisomerase were mixed gently and incubated at room temperature for 30 minutes.

2.7.12 Preparation of electrocompetent *E. coli* cells

A 100 ml volume of LB broth was inoculated with 1 ml overnight broth culture of *E. coli* Top10 or CA434 and incubated shaking at 37 °C until an A_{600} of 0.5 – 0.8 was reached. Cultures were placed on ice for 15 minutes. Cells were harvested by centrifugation at 5000 X *g* for 15 minutes at 4°C, and then washed in 100 ml ice cold sterile distilled water. This wash step was repeated but the cells were resuspended in 2 ml 10% glycerol. The cells were harvested again and the final pellet resuspended in 300 µl 10% glycerol and stored at –80°C in 40 µl aliquots.

2.7.13 Electroporation

To remove salt prior to electroporation, ligation reactions were dialysed against distilled water for 30 minutes. Dialysis discs with 0.025 μm pores (Millipore, UK) were used. A 5 μl aliquot of dialysed ligation mix was added to 40 μl electrocompetent cells and the mixture transferred to an electroporation cuvette with a 2 mm gap (Biorad). Electroporation was carried out using a Gene Pulser (Biorad) set to 200 Ω, 2.5 kV and 25 μFD. After electroporation, 500 μl LB medium was added to the cells, which were then recovered by incubating shaking at 37°C for 1 hour, before being plated onto LB agar plates containing appropriate selection.

2.7.14 Transformation of DNA into *Bacillus subtilis* (method 1)

DNA was transformed into *Bacillus subtilis* as described by Anagnostopoulos & Spizizen (1961). A heavy inoculum of fresh overnight growth was added to 20 ml competence medium and incubated at 37°C with agitation until an A_{600} of 3 was achieved. The cells were diluted ten-fold in transformation medium and 1 ml aliquots added to 1 μg plasmid DNA. Cells were incubated at 37°C for a further 90 minutes. Aliquots were then spread onto LB agar plates containing the appropriate selection.

2.7.15 Transformation of naturally competent *B. subtilis* cells with DNA (method 2)

The transformation protocol was adapted from Bron (1990). 10 ml minimal growth medium was inoculated with a single colony of *B. subtilis* and incubated overnight at 37°C with agitation. A 1.4 ml aliquot of an overnight culture was used to inoculate 10 ml pre-warmed fresh minimal growth medium and the culture grown at 37°C with agitation for 3 hours. An 11 ml aliquot of pre-warmed starvation medium was added to the culture and growth continued at 37°C for 2 hours until competence was achieved. A 300 μl volume of competent cells was transferred to a sterile tube and 1 μg plasmid DNA in 10 μl TE buffer added. The cells were incubated at 37°C with agitation for 45 minutes. Pre-warmed LB broth (700 μl) was added and the cells incubated for a further 45-60 minutes. Cells were then plated out onto LB agar with appropriate selection in 10, 100 and 200 μl aliquots and incubated overnight at 37°C.

2.7.16 Blue/white colony selection

Blue/white selection was used when cloning fragments into the pCR2.1-TOPO vector. The successful insertion of a PCR product into this vector results in the disruption of the *lacZ* gene. X-Gal was added to the LB Amp plates, at a final concentration of 80μg/ml, to screen for β-galactosidase production. Any colonies producing this enzyme would cleave the X-Gal, resulting in the production of a blue coloration. This indicated that there was no insert present in these colonies. White colonies lacked β-galactosidase activity, indicating that the insert was usually present.

2.7.17 Conjugation of *E. coli* CA434 (HB101) with *C. sporogenes* 10696 (filter mating method)

Recipient and donor cells were harvested from 10 ml overnight broth cultures. The clostridial recipient cells were resuspended in 1 ml TYG broth and this same 1 ml suspension used to resuspend the *E. coli* donor cells. Nitrocellulose filters with 0.45 μm pores were placed onto TYG + 0.5% glucose plates and 200 μl mixed cell suspension spread onto the filters. Plates were incubated anaerobically overnight. Filters were placed in 1 ml TYG in 20 ml universals and vortexed for 1 minute to dislodge the cells. A 100 μl aliquot of resuspended cells was spread onto TYG plates without glucose, containing appropriate selection for the vector, and cycloserine, and incubated anaerobically overnight.

2.7.18 Conjugation of *E. coli* CA434 (HB101) with *C. sporogenes* 10696 or *C. botulinum* Hall A (spot method)

The 'spotting' method was carried out as described by Purdy *et al.* (2002). Donor and recipient cells were grown overnight in 5 ml broths (*E. coli* donor CA434 in LB, aerobic with agitation; clostridial recipient in TYG, anaerobic without agitation). A 1 ml aliquot of donor strain was taken and the cells harvested by centrifuging at 5000 X *g* for 1 minute. The pellet was resuspended in 1 ml sterile PBS and the cells harvested as before. The pellet was transferred to the anaerobic work station and resuspended in 200 μl clostridial recipient strain. The whole 200 μl cell suspension was spotted, in 10 μl

aliquots, onto non-selective TYG plates containing 0.5% glucose and incubated for 4-6 hours at 37°C in anaerobic conditions. The cells were harvested by flooding the plate with 1 ml sterile PBS and dislodging the cells with a plastic spreader. The resulting slurry was removed and spread onto selective plates, followed by incubation at 37°C overnight in anaerobic conditions.

2.7.19 DNA sequencing

Sequencing of DNA was carried out at the DNA Sequencing Laboratory, School of Biomedical Sciences, Queen's Medical Centre, Nottingham, using a 3100 Genetic Analyser (Applied Biosystems), or by GeneService (UK). Sequence analysis was carried out using the Lasergene DNAstar computer package or the Invitrogen Vector NTI package.

The National Centre for Biotechnology Information (NCBI) website was used to search sequence databases (http://www.ncbi.nlm.nih.gov/blast/).

2.7.20 Southern hybridisation

DNA was quantified according to 2.8.3 and diluted so that equal concentrations were loaded in each gel lane. Each DNA sample (1-3 μg) was digested overnight with appropriate restriction enzymes according to 2.7.4. Samples were run on a 0.8% agarose gel as described in 2.7.7. The blotting apparatus was then set up as follows. The gel tray used to cast the gel in was turned upside down in a tray containing 0.4 M sodium hydroxide. A layer of filter paper (Whatman) was placed on top to form a wick, and the gel placed on top of the paper. A piece of N + nylon membrane (Roche) was placed on top of the gel, followed by five pieces of blotting paper (Sigma), all of which had been soaked in 0.4 M sodium hydroxide. A weight was placed on the top to maximise blotting efficiency and the apparatus incubated at room temperature overnight.

Following transfer, the membrane was removed, wrapped in cling film and exposed to UV light for 2 minutes on each side to crosslink the DNA. The membrane was then rinsed in 2 X SSC (300 mM sodium chloride, 30 mM sodium citrate, pH 7). Blots were incubated in 10 ml DIG Easy Hyb (Roche) in hybridisation cylinders for 1 hour at 42°C

using a Techne Hybridiser HB-1D. The Easy Hyb was discarded and replaced with 10 ml DIG Easy Hyb containing 1 μg DIG-labelled DNA probe, which had been boiled for 10 minutes and transferred immediately to ice prior to use. Probes were generated by PCR, followed by purification and labelling by incorporating DIG-labelled dUTPs using the Roche DIG High Prime DNA Labelling and Detection Starter Kit II according to the manufacturer's instructions. The probe solution was incubated overnight at 42°C.

The membrane was washed twice in 2 X washing solution (2 X SSC, 0.1% (w/v) SDS (Sigma)) for 5 minutes at room temperature and then twice in 0.1 X washing solution for 20 minutes at 68°C. It was then equilibrated in washing buffer (0.1 M maleic acid, 0.15 M sodium chloride, pH 7, 0.3% (v/v) tween 20) for 1 minute at room temperature and blocked in blocking buffer (included in the Roche High Prime Kit) for 1 hour at room temperature. The membrane was incubated in blocking buffer containing anti-digoxigenin antibody (1:10,000 dilution [Roche]) for 30 minutes at room temperature, followed by two washes in washing buffer for 15 minutes at room temperature. It was then equilibrated in detection buffer (100 mM Tris HCl, pH 9.5, 100 mM sodium chloride) for 2 minutes at room temperature. The membrane was then placed on an acetate sheet and 1 ml of chemiluminescent substrate CSDP Star (Roche) was added, followed by a second acetate sheet. This was incubated for 5 minutes at room temperature followed by 10 minutes at 37°C, and exposed to X-ray film (Amersham Pharmacia Biotech).

2.8 RNA analysis

To minimise RNase contamination all RNA work was carried out in a designated clean area which was cleaned with RNaseZap (Sigma). All equipment was also wiped over with RNaseZap prior to use. RNase clean tips and microfuge tubes (Eppendorf) were used at all times. Water used to make up solutions was treated with DEPC by adding 1 ml DEPC to 1 litre of water, mixing and leaving at room temperature for approximately 3 hours before autoclaving.

2.8.1 Preparation of total RNA from *C. sporogenes* (method 1)

A 1 ml volume of *C. sporogenes* broth culture was added to 2 ml RNA Protect (Qiagen) in a 50 ml falcon tube and mixed by vortexing for 5 seconds. Cells were incubated for 5 minutes at room temperature and then harvested by centrifugation at 5000 X *g* for 10 minutes at room temperature. All of the supernatant was removed by dabbing the lip of the tube on a tissue. A 100 µl aliquot of 3 mg/ml lysozyme (Sigma), dissolved in DEPC treated water, was then used to resuspend the pellet. The cell suspension was transferred to a microfuge tube and mixed by vortexing for 10 seconds. Tubes were incubated at room temperature for 10 minutes, vortexing every 2 minutes. Total RNA was extracted from the cell lysate using Qiagen's RNeasy kit: 350 µl buffer RLT, to which β-mercaptoethanol had previously been added, was added to the lysate and mixed by vortexing. A 250 µl volume of cold ethanol was added and mixed by pipetting. The solution was then applied to an RNeasy column and centrifuged in a bench top centrifuge for 15 seconds at 8000 X *g*. The flow through was discarded and an on-column DNase treatment carried out as follows:

70 µl TURBO DNase buffer (Ambion) containing 10 µl TURBO DNase was added to the column and incubated at room temperature for 30 minutes. The Qiagen RNeasy protocol was then continued. A 350 µl volume of RW1 buffer was applied to the column and centrifuged for 15 seconds at 8000 X *g*. The flow through was discarded and the column washed by the addition of 500 µl RPE buffer. Centrifugation was carried out as before. The column was washed with 500 µl RPE buffer for a second time and centrifuged for 2 minutes at 8000 X *g*. The flow through was discarded and the column centrifuged for 1 minute at full speed. RNA was eluted by the addition of 50 µl RNase free water and centrifugation for 1 minute at 8000 X *g*.

2.8.1.1 Removal of contaminating genomic DNA from total RNA

Contaminating DNA was removed from RNA samples using TURBO DNase (Ambion). A 40 µl aliquot of RNA was mixed with 5 µl 10 X buffer and 5 µl TURBO DNase, and incubated at 37°C for 30 minutes.

2.8.1.2 DNase treated RNA sample clean up

RNA which had been treated with TURBO DNase was cleaned up prior to use in RT-PCR as described in the Qiagen RNeasy kit instructions. The RNA sample was adjusted to 100 µl with RNase free water. A 350 µl volume of RLT buffer was added to the sample and mixed, followed by addition of 250 µl cold ethanol. The solution was mixed by pipetting, applied to an RNeasy column and centrifuged for 15 seconds at 8000 X *g*. The column was washed by the addition of 500 µl RPE buffer and centrifuged as before. The column was washed a second time using another 500 µl RPE buffer and centrifuged for 2 minutes at 8000 X *g*. The RNA was eluted by the addition of 30 µl RNase free water and centrifugation for 1 minute at 8000 X *g*.

2.8.2 Preparation of total RNA from *C. sporogenes* (method 2)

Samples of *C. sporogenes* broth culture were collected in 10 ml volumes and 2 ml of phenol-ethanol solution [phenol pH 4.3-ethanol (1:9)] added immediately. The tubes were shaken well to halt any changes in gene expression or degradation of RNA. Tubes were placed on ice for 30 minutes before centrifuging at 3000 rpm for 20 minutes at 4°C. The supernatant was discarded and the cell pellets resuspended in 500 µl ice cold TES buffer (50 mM Tris, 5 mM EDTA, 50 mM NaCl). A 1 ml aliquot of silica beads was added to 2 ml tubes along with 600 µl acid phenol and 100 µl chloroform (all Sigma). The tubes were vortexed and the cell suspensions added. Tubes were heated to 60°C and vortexed for 3 times 30 seconds, before chilling on ice. Tubes were centrifuged at 5000 rpm for 10 minutes at 4°C. A 600 µl volume of acid phenol and 100 µl chloroform were added to 500 µl aqueous phase, before vortexing and centrifuging at 14000 rpm for 5 minutes at 4°C. A 600 µl aliquot of chloroform was added to 350 µl aqueous phase and the contents mixed and centrifuged as before. Finally, 20 µl 3M sodium acetete, pH 4.8 (Sigma) was added to 200 µl aqueous phase along with 550 µl ice cold 95% ethanol. The tubes were inverted to mix and the RNA precipitated at -80°C for 20 minutes, before centrifugation at 14 000 rpm for 20 minutes at 4°C. The supernatant was discarded and the pellet washed with 600 µl 70% ethanol. The tubes were once again centrifuged at 14 000 rpm for 10 minutes at 4°C before discarding the

supernatant and air-drying the pellet. This was then resuspended in 200 µl DEPC treated water.

2.8.2.1 RNase inhibitor/ DNase treatment

To 200 µl RNA solution was added: 4 µl RNase inhibitor mix (Promega), 20 µl TURBO DNase 10x reaction buffer (Ambion) and 4 µl TURBO DNase (Ambion). The mixture was incubated at 37°C for 30 minutes.

2.8.2.2 Sample clean up

A 0.5 ml volume of saturated phenol (pH 4.3):Chloroform:Isoamylalcohol (IAA) (25:24:1) (Sigma) was pipetted into 'Phase Lok Gel' tubes (Eppendorf) and the RNA samples added. Samples were mixed well by repeated inversion and centrifuged at 15 000 rpm for 5 minutes at 4°C. A 0.5 ml aliquot of Chloroform:IAA (24:1) (Sigma) was pipetted into fresh phase lock tubes and the top layer of the centrifuged sample added before mixing well. Samples were centrifuged as before. The supernatants were transferred to fresh eppendorf tubes and 50 µl 3M sodium acetate buffer solution (Sigma catalogue no. S-7899) added, along with 2.5 volumes of 96% ethanol. Samples were mixed and placed at -80°C for 1 hour. Tubes were centrifuged at 15 000 rpm at 2°C for 20 minutes before discarding the supernatant. A 1 ml volume of 70% ethanol was added and the tube inverted gently. Tubes were then centrifuged at 15000 rpm for 5 minutes at 4°C, the supernatant removed and tubes centrifuged again at the same speed for 30 seconds. Any remaining supernatant was removed and the tubes placed at 37°C to dry the RNA pellet. This was then dissolved in 100 µl DEPC water.

2.8.3 Quantification of RNA and DNA

RNA/DNA was quantified using a NanoDrop ND1000 spectrophotometer (NanoDrop Technologies, USA).

2.8.4 One-step Reverse Transcriptase (RT)-PCR

All RNA samples were prepared according to 2.8.1. RT-PCR was carried out using Qiagen's One-Step RT-PCR kit as described in the manufacturer's instructions. The reaction mix was composed of 10 µl 5 X buffer, 2 µl dNTP mix, 3 µl each primer (10 µM) and 2 µl enzyme mix and 20 ng (antisense expression) or 50 ng (linkage of expression) of RNA. The volume was adjusted to 50 µl with RNase free water. The reaction was started by heating to 50°C for 30 minutes to allow reverse transcriptase activity and then to 95°C for 15 minutes for DNA polymerase activation and reverse transcriptase inactivation. A total of 30 cycles of the following was then carried out: denaturation at 94°C for 1 minute, annealing at an appropriate temperature for 1 minute, and elongation at 72°C for an appropriate time. The final step in the reaction was a hold at 72°C for 10 minutes. The annealing temperature used was dependent on the melting temperatures of the primers used, and the elongation time was dependent on the length of the template (1 minute per kilobase).

RT-PCR products were visualised on an agarose gel as described in section 2.7.7.

2.8.5 Real Time Quantitative PCR

2.8.5.1 Primers

Three sets of PCR primers were designed using the Primer Express programme (Applied Biosystems, USA). The two sets of test gene primers were designed against the *agrD*1 and *agrD*2 genes of *C. sporogenes* (CsRTagrD1/1F & 1R and CsRTagrD2/1F & 1R). The set of control primers was designed against the 16S gene, which was to be used as the endogenous control (CsRT16S/1F & 1R). The Absolute Quantification (AQ) method was used.

2.8.5.2 RNA extraction

RNA was extracted at five time points during *C. sporogenes* growth. Triplicate 100 ml broth cultures were grown in TYG broth and growth monitored using OD. At each of the five time points, 10 ml culture was removed and RNA extracted according to 2.8.2.

RNA was quantified according to 2.8.3 and the quality checked on an Agilent Technologies 2100 Bioanalyser before proceeding to the reverse transcription step.

2.8.5.3 Reverse transcription

Reverse transcription was carried out using SuperScript II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. A final concentration of 5 μg RNA was used in the reaction, and after completion, the cDNA was purified according to 2.7.8.

2.8.5.4 Real time PCR

The real time PCR method used the double stranded DNA-specific dye SYBR Green. Real time PCR reaction mixes contained 12.5 μl Power SYBR Green PCR Master Mix (Applied Biosystems) and 2 μl cDNA template (0.05 ng/μl). Primers were used at a concentration of 10 μM stock. For the two sets of *agrD* primers, the reaction contained 0.75 μl forward primer and 2.25 μl reverse primer (300/900 ratio). For the 16S primers the reaction contained 2.25 μl forward primer and 0.75 μl reverse primer (900/300 ratio). The reaction mix was made up to 25 μl with water. The PCR programme was as follows:

60°C - 2 min
95°C – 10 min
95°C – 15 s
58°C – 1 min
} 40 cycles

After the 40 cycles were complete, a dissociation step was added to check for any non-specific product formation. This consisted of 95°C for 15 seconds, 60°C for 1 min and 95° C for 15 seconds. The reaction was carried out in an AB7500 Real Time PCR System (Applied Biosystems). Each reaction was set up in triplicate wells in a MicroAmp Optical 96-well reaction plate (Applied Biosystems).

2.8.5.5 Standard curve generation

When using the AQ method, a standard curve must be run for both the test gene and the control gene for every run carried out. Curves were constructed using 10-fold dilution series of *C. sporogenes* cDNA. Concentrations used ranged from 100 – 0.001 ng per reaction. The curve plotted gave the log of the concentration of cDNA in the sample against the threshold cycle (the cycle at which fluorescence rose above the background).

2.8.5.6 Data analysis

Analysis was carried out using the 7500 System software provided with the apparatus. The software provided amplification plots for each reaction, and for each gene tested, it determined the unknown concentrations of cDNA in each reaction. This was done by using the C_T value of the unknown sample, and reading the value for the cDNA concentration from the standard curves. The average value from each set of triplicate wells was calculated. Expression levels of the *agrD* genes were then normalised against the endogenous control, 16S, giving final values as relative amounts of *agrD* mRNA.

2.8.6 Northern hybridisation

RNA was quantified according to 2.8.3 and diluted so that equal concentrations were loaded in each gel lane. Two volumes of RNA sample buffer (50 µl 10X MOPS [200 mM MOPS, 50 mM sodium acetate, 10 mM EDTA, pH 7], 87.5 µl formaldehyde, 250 µl formamide (Sigma) and 112.5 µl DEPC treated water (Ambion)) were added to one volume of each RNA sample and the mixture heated to 65°C for 5 minutes. Loading dye (250 µl formamide, 83 µl formaldehyde, 50 µl 10 X MOPS, 0.1% bromophenol blue (Sigma), 50 µl glycerol (Courtin and Warner Ltd, UK) and 67 µl DEPC treated water) was added to each sample at a dilution of 1:10 before loading. Formaldehyde gels were prepared by melting 1.8 g agarose in 130.5 ml DEPC treated water, allowing to cool to 50°C and adding 6 µl ethidium bromide (10 mg/ml), 15 ml 10 X MOPS and 4.5 ml formaldehyde. Once mixed, the gels were poured and allowed to set. Samples were loaded and the gel run in 1 X MOPS at 100 V until the dye front had moved half way down the gel.

The gel was rinsed briefly in DEPC treated water and the blotting apparatus set up as follows. The gel tray used to cast the gel in was turned upside down in a tray containing 20 X SSC (3 M sodium chloride (Sigma), 300 mM sodium citrate (Sigma), pH 7). A layer of filter paper (Whatman) was placed on top to form a wick, and the gel placed on top of the paper. A piece of N + nylon membrane (Roche) was placed on top of the gel, followed by five pieces of blotting paper (Sigma), all of which had been soaked in 2 X SSC. A weight was placed on the top to maximise blotting efficiency and the apparatus incubated at room temperature overnight.

Following transfer, the membrane was removed, wrapped in cling film and exposed to UV light for 2 minutes on each side to crosslink the RNA. Blots were incubated in 8 ml DIG Easy Hyb (Roche) in hybridisation cylinders for 1 hour at 50°C using a Techne Hybridiser HB-1D. The Easy Hyb was discarded and replaced with 10 ml DIG Easy Hyb containing 1 μg DIG-labelled DNA probe, which had been boiled for 10 minutes and transferred immediately to ice prior to use. Probes were generated by PCR, followed by purification and labelling by incorporating DIG-labelled dUTPs using the Roche DIG High Prime DNA Labelling and Detection Starter Kit II according to the manufacturer's instructions. The probe solution was incubated overnight at 50°C.

The membrane was washed twice in 2 X washing solution (2 X SSC, 0.1% (w/v) SDS (Sigma)) for 5 minutes at room temperature and then twice in 0.1 X washing solution for 20 minutes at 68°C. It was then equilibrated in washing buffer (0.1 M maleic acid, 0.15 M sodium chloride, pH 7, 0.3% (v/v) tween 20) for 1 minute at room temperature and blocked in blocking buffer (Roche High Prime Kit) for 1 hour at room temperature. The membrane was incubated in blocking buffer containing anti-digoxigenin antibody (1:10,000 dilution [Roche]) for 30 minutes at room temperature, followed by two washes in washing buffer for 15 minutes at room temperature. It was then equilibrated in detection buffer (100 mM Tris HCl, pH 9.5, 100 mM sodium chloride) for 2 minutes at room temperature. The membrane was then placed on an acetate sheet and 1 ml of chemiluminescent substrate CSDP Star (Roche) was added, followed by a second acetate sheet. This was incubated for 5 minutes at room temperature followed by 10 minutes at 37°C, and exposed to X-ray film (Amersham Pharmacia Biotech).

2.9 Estimation of segregational stability of plasmids in *Bacillus subtilis* and *Clostridium* spp

Segregational stability was assessed using a method modified from Bron & Luxen (1985). *Bacillus* cultures were incubated aerobically, clostridial cultures anaerobically.

Day 1 - Two overnight broth cultures were grown in LB (*Bacillus*) or TYG (clostridia) with appropriate antibiotic; one containing 0.1 mM IPTG and one without IPTG.

Day 2 - 50 µl each overnight culture was added to 5 ml fresh culture medium without antibiotic, one with IPTG and one without. These were incubated at 37°C with agitation (*Bacillus*) or statically (clostridia) for 9 hours, by which time the culture had reached approximately 10 generations of growth. A 50 µl aliquot was subcultured into 5 ml fresh broth as before. These cells were again incubated overnight.

Day 3 - The following morning serial dilutions were carried out on each broth from 10^{-1} to 10^{-7} in fresh liquid medium. A 100 µl aliquot of each dilution was plated out as follows: broth containing IPTG; cells plated onto LB or TYG with selection and 0.1 mM IPTG, and onto LB or TYG alone. Broth without IPTG; cells plated onto LB or TYG with selection and onto LB or TYG alone. A 50 µl aliquot was subcultured into 5 ml fresh broth. After a further 9 hours incubation, 50 µl was again subcultured into 5 ml fresh broth.

Day 4 – The procedures carried out on day 3 were repeated

Day 5 – the procedures carried out on day 3 were repeated with the exception of the two 50 µl subcultures into 5 ml fresh broth.

All colonies were counted and the percentage plasmid loss was calculated using the difference between the total counts (on LB or TYG) and counts of resistant colonies (on LB or TYG with selection).

2.10 Induction of gene replacement using integrative vector in *Clostridium sporogenes*

Induction of integration of a knockout cassette into the host target gene, and loss of the unstable plasmid was carried out according to a method supplied by Philipe Soucaille, National Institute of Applied Science, Toulouse, France (personal communication).

Day 1 - Two overnight broth cultures were grown in TYG with appropriate antibiotic; one containing 0.1 mM IPTG and one without IPTG.

Day 2 - 50 µl each 24 h culture was added to 5 ml fresh culture medium without antibiotic, one with IPTG and one without. These were incubated at 37°C for 24 hours.

Day 3 - Serial dilutions of the 24 h culture were carried out from 10^{-1} to 10^{-6} in fresh liquid medium. A 100 µl aliquot of each dilution was plated out as follows: broth containing IPTG; cells plated onto TYG containing 0.1 mM IPTG and either thiamphenicol, erythromycin or no selection. Broth without IPTG; cells plated onto TYG containing either thiamphenicol, erythromycin or no selection. A 50 µl aliquot of 24 h culture was subcultured into 5 ml fresh broth and incubated for 24 h.

Day 4 and day 5 – the procedures carried out on day 3 were repeated.

Day 6 – the procedures carried out on days 4 and 5 were repeated with the exception of the 50 µl subculture into 5 ml fresh broth.

Colonies were enumerated after 48 hours' growth. Percentage plasmid loss was calculated and a selection of colonies was picked from plates obtained on days 4, 5 and 6, both with and without IPTG. These were streaked out in duplicate onto TYG agar containing either thiamphenicol or erythromycin. Any colonies exhibiting thiamphenicol resistance and erythromycin sensitivity could be treated as suspected integrative mutants, which had been cured of the plasmid.

2.11 Induction of target gene knockout using the ClosTron System

2.11.1 Using plasmid pMTL007

Transconjugants were obtained as described in the Chapter 5 results section. TYG broth (0.5 ml) containing 10 µg/ml (*C. botulinum*) or 7.5 µg/ml (*C. sporogenes*) thiamphenicol was inoculated with a clostridial transconjugant, containing the retargeted pMTL007 vector, and incubated overnight. A 100 µl aliquot of an overnight culture was used to inoculate 1 ml fresh TYG broth + thiamphenicol and the culture incubated until growth was visible (approximately 1 hour). IPTG was added to the culture to a final concentration of 1 mM to induce integration, and the culture incubated for 3 hours. Cells were harvested by centrifugation at 8000 rpm for 1 minute and the supernatant discarded. The cells were washed by resuspending in 0.5 ml sterile PBS. The cells were harvested and the supernatant discarded as before. The pellet was resuspended in 1 ml TYG broth without selection and incubated for a further 3 hours. A 100 µl aliquot of the integration mixture was plated out onto TYG agar containing 20 µg/ml (*C. botulinum*) or 10 µg/ml (*C. sporogenes*) erythromycin in neat, 10^{-2} and 5x concentrations. Plates were incubated for 48 hours.

2.11.2 Using plasmid pMTL007CE-2

A large colony of a clostridial transconjugant, containing the retargeted pMTL007CE-2 vector was thoroughly resuspended in 100 µl TYG broth. Serial dilutions from neat to 10^{-7} were carried out on this cell suspension and 100 µl each plated out onto agar containing 20 µg/ml (*C. botulinum*) or 10 µg/ml (*C. sporogenes*) erythromycin. Plates were incubated for 48 hours.

For both 2.11.1 and 2.11.2, frequency of integration was determined by carrying out serial dilutions of the integration mixture up to 10^{-7}, and plating out 100 µl of dilutions 10^{-4} to 10^{-7} onto TYG plates without selection. Plates were incubated overnight and colonies enumerated. The percentage frequency of integration was calculated by comparing these counts to those obtained from the selective integration plates.

Any potential integrants were re-streaked onto fresh selective TYG agar and their DNA extracted for PCR screening. Any correct integrants were subcultured several times onto non-selective TYG agar in order to cure them of the pMTL007 or pMTL007CE-2 plasmid.

2.12 Measurement of spore formation in *Clostridium* species

A 5 ml overnight culture was grown in TYG broth under anaerobic conditions. The following day the culture was diluted 1:100 in 5 ml fresh medium, and incubated for 72 hours. After incubation a 0.3 ml sample of culture was heated to 80°C for 20 minutes. Serial dilutions were then carried out and 100 µl aliquots of the heat-treated cell suspension plated onto TYG agar. Plates were incubated for 48 hours (*C. sporogenes*) or 72 hours (*C. botulinum*) and colonies enumerated.

Chapter 3.0

Identification and Characterisation of Clostridial *agr* Systems

Chapter 3 – Identification and Characterisation of Clostridial *agr* Systems

3.1 Introduction

The recent completion of the sequencing of the *C. botulinum* genome at the Sanger Institute, Cambridge UK, has revealed the presence of *agr* gene homologues which correspond to those existing in *S. aureus*. Figure 3.1 shows the *agr* regions identified to date in four different species of clostridia. Interestingly, *C. botulinum* possesses two *agrD* and two *agrB* homologues. To date, *C. acetobutylicum* is the only clostridial species found to possess all four *agr* genes corresponding to those identified in *S. aureus*. However, early studies suggest that some hypervirulent strains of *C. difficile* also possess homologues of all four *agr* genes (Ian Davis, University of Nottingham, UK, personal communication).

The *C. botulinum* genome contains 28 two-component systems, but none showing similarities to *agrA/agrC* are situated in the vicinity of the *agr* regions (Sebaihia *et al.*, 2007). The only two-component system identified in the *C. botulinum* genome which shows similarity to the staphylococcal *agrA* and *agrC* genes (genes CBO1186 and CBO1187) is present elsewhere in the genome (Sebaihia *et al.*, 2007). Furthermore, these genes exhibit relatively low percentage identities with the staphylococcal *agrC/A* genes. CBO1187 shares 13.4% protein identity with *S. aureus* AgrA, and CBO1186 shares 12.9% protein identity with *S. aureus* AgrC.

In *S. aureus*, the *agr* system is known to be a global regulator of a battery of virulence factors, but the mechanism of its function and the identification of its target genes in clostridia remain to be elucidated.

At the commencement of the project, facilities were not available for handling *C. botulinum*. The project was therefore initiated using *C. sporogenes*. The high sequence similarity and absence of toxin production render this species an ideal model organism

for *C. botulinum*. As the genome sequence of *C. sporogenes* has yet to be determined, the first step was to obtain the nucleic acid sequence of the *agr* regions in this *Clostridium*. This would allow the subsequent investigation of their expression profiles.

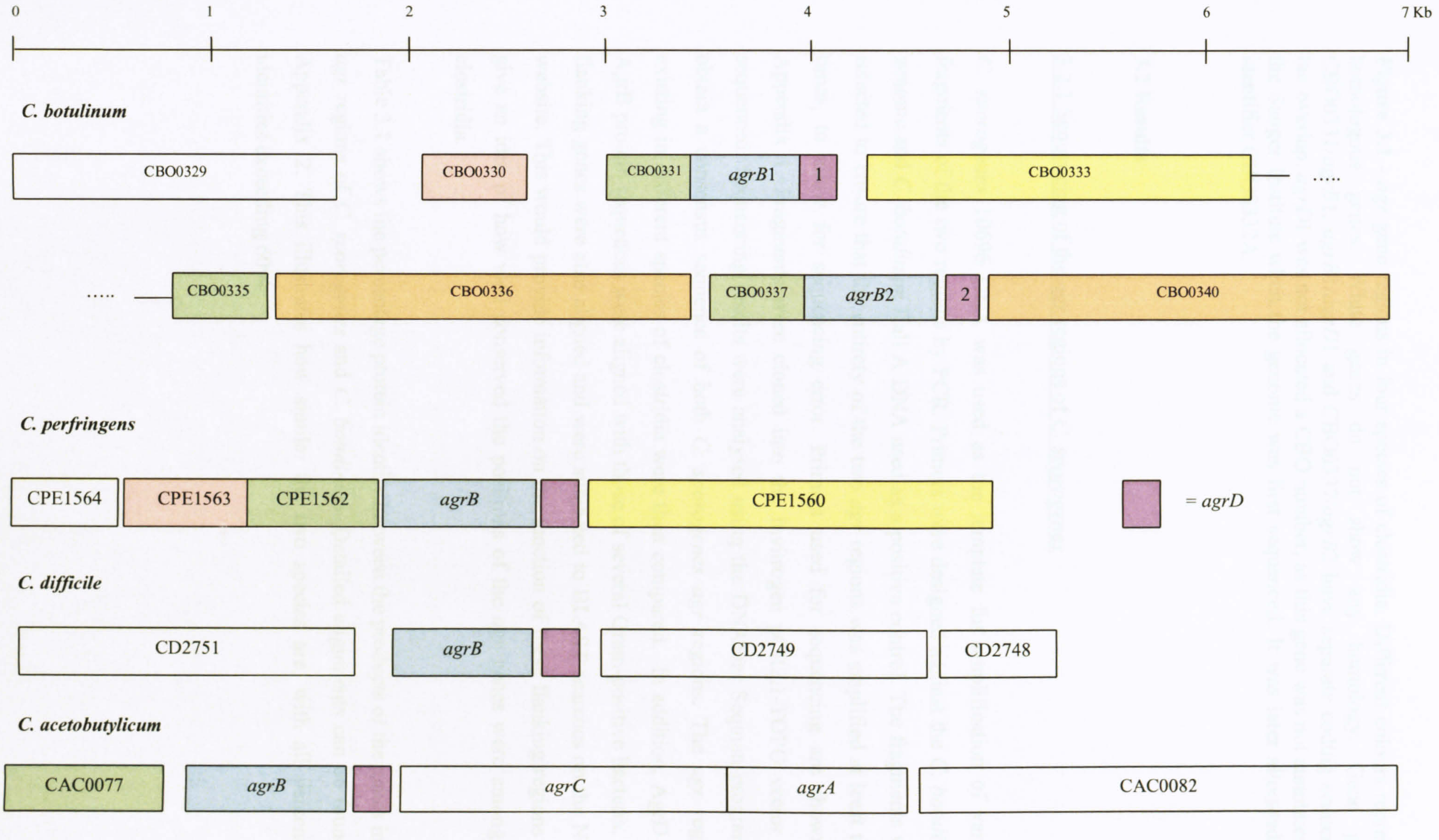


Figure 3.1 – *agr* gene regions in four species of clostridia. See next page for figure legend.

Figure 3.1 - *agr* gene regions in four species of clostridia. Different colours represent homologous genes. White genes do not show any homology. Gene pairs CBO0331/*agrB*1, *agrB*1/*agrD*1 and CBO0337/*agrB*2 have separate coding sequences but overlap. *agrD*1 was not allocated a CBO number, as this gene was not annotated by the Sanger Institute when the genome was first sequenced. It was later allocated the identifier CBO0332A.

3.2 Results

3.2.1 Sequencing of the *agr* regions of *C. sporogenes*

C. sporogenes 10696 DNA was used as the template for amplification of various fragments of the two regions by PCR. Primers were designed against the *C. botulinum* genome and *C. botulinum* Hall A DNA used as a positive control. The fragments were selected to ensure that the entirety of the two *agr* regions was amplified at least three times, to allow for sequencing error. Primers used for sequencing are shown in Appendix 1. Fragments were cloned into the Invitrogen pCR2.1-TOPO vector and sequenced. Sequencing results were analysed using the DNAStar Seqman program to obtain a consensus sequence of both *C. sporogenes agr* regions. The *agr* regions existing in different species of clostridia were then compared. In addition, AgrD and AgrB protein sequences were aligned with those of several Gram-positive bacteria. The flanking genes were also aligned and were subjected to BLAST searches on the NCBI website. This would provide information on the function of these flanking regions and give an idea of how well conserved the positions of the *agr* genes were among the clostridia.

Table 3.1 shows the percentage protein identity between the products of the genes in the *agr* regions of *C. sporogenes* and *C. botulinum*. Detailed alignments can be found in Appendix 2. This illustrates how similar the two species are, with all percentage identities exceeding 60%.

Protein	% Protein Identity
0330	89.2
0331	89.7
AgrB1	86.9
AgrD1	88.9
0333	84.0
0335	61.5
0336	84.8
0337	74.4
AgrB2	67.1
AgrD2	88.9
0340	90.6

Table 3.1 - Percentage protein identity between gene products in the *agr* regions of *C. sporogenes* and *C. botulinum*. Identifier numbers are those allocated to the *C. botulinum* genes by the Sanger Institute, but equivalent genes have been identified in *C. sporogenes*.

3.2.2 Comparison of *agr* region gene products amongst clostridia and other Gram-positive bacteria

Sequence alignments were carried out on AgrD, AgrB and flanking regions of various clostridia and other Gram-positive bacteria.

3.2.2.1 AgrD alignments

Figure 3.2 shows a phylogenetic tree of the AgrD proteins. It is evident that the *C. acetobutylicum* AgrD protein, and one of the AgrD proteins from *Clostridium beijerinckii* are less closely related than those from other clostridial species. *C. sporogenes* and *C. botulinum* AgrD1 proteins are more closely related to *C. perfringens* AgrD than to *C. sporogenes* and *C. botulinum* AgrD2. In fact, the clusters in the phylogenetic tree indicate that the AgrD1 and AgrD2 of *C. botulinum* and *C. sporogenes* are quite different. This is reinforced by the sequence identity data in Table 3.2.

All of the AgrD proteins of the staphylococci are closely grouped together, but AgrD of *Enterococcus faecalis* is not closely related to any of the other proteins, and is found on a branch of its own. This is probably due to the fact that the mature signalling peptide formed in this organism is actually longer than those found in other Gram-positive bacteria; it is, in fact, made up of eleven amino acids (Nakayama *et al.*, 2006). Generally the staphylococcal AgrD proteins show a greater number of highly conserved regions than those of the clostridia (Figure 3.3).

Figure 3.4 shows an alignment of the putative clostridial AIPs, along with those from the four *S. aureus* groups. All proteins in this alignment show a conserved cysteine residue (shown in bold red), and the clostridial proteins appear to show their own conserved sequences, which do not correspond to those observed in the staphylococcal AIP groups. For example, almost all of the clostridial sequences exhibit a serine and an alanine residue immediately N terminal to the conserved cysteine. In addition, the clostridial proteins mostly share a common glutamine and proline residue six residues C terminal to the conserved cysteine residue. Neither of these conserved sequences is observed in the *S. aureus* AIP groups.

Interestingly, the putative AgrD protein from *C. acetobutylicum* is quite different to those of the other clostridial species. Figure 3.4B demonstrates that there is an additional residue involved in the formation of the thiolactone ring structure. In the other clostridial species there appear to be four residues C terminal to the conserved cysteine, which would potentially form the thiolactone ring (shown in purple in Figure 3.4B). In *C. acetobutylicum*, there is an additional leucine residue present, which would produce a larger ring structure. This clostridial species is also the only one known to possess homologues of all four *S. aureus agr* genes (B, D, A and C). Its *agr* system clearly differs significantly to those identified in other clostridial species.

The possible tail sequence of the mature clostridial AIPs is also indicated in Figure 3.4B in green. There appears to be very high conservation among these residues. However, it has yet to be confirmed how many of these residues are indeed involved in the formation of the tail. This number may vary, as is seen in *S. aureus*, with the different AIP groups possessing tails from two to four amino acids in length.

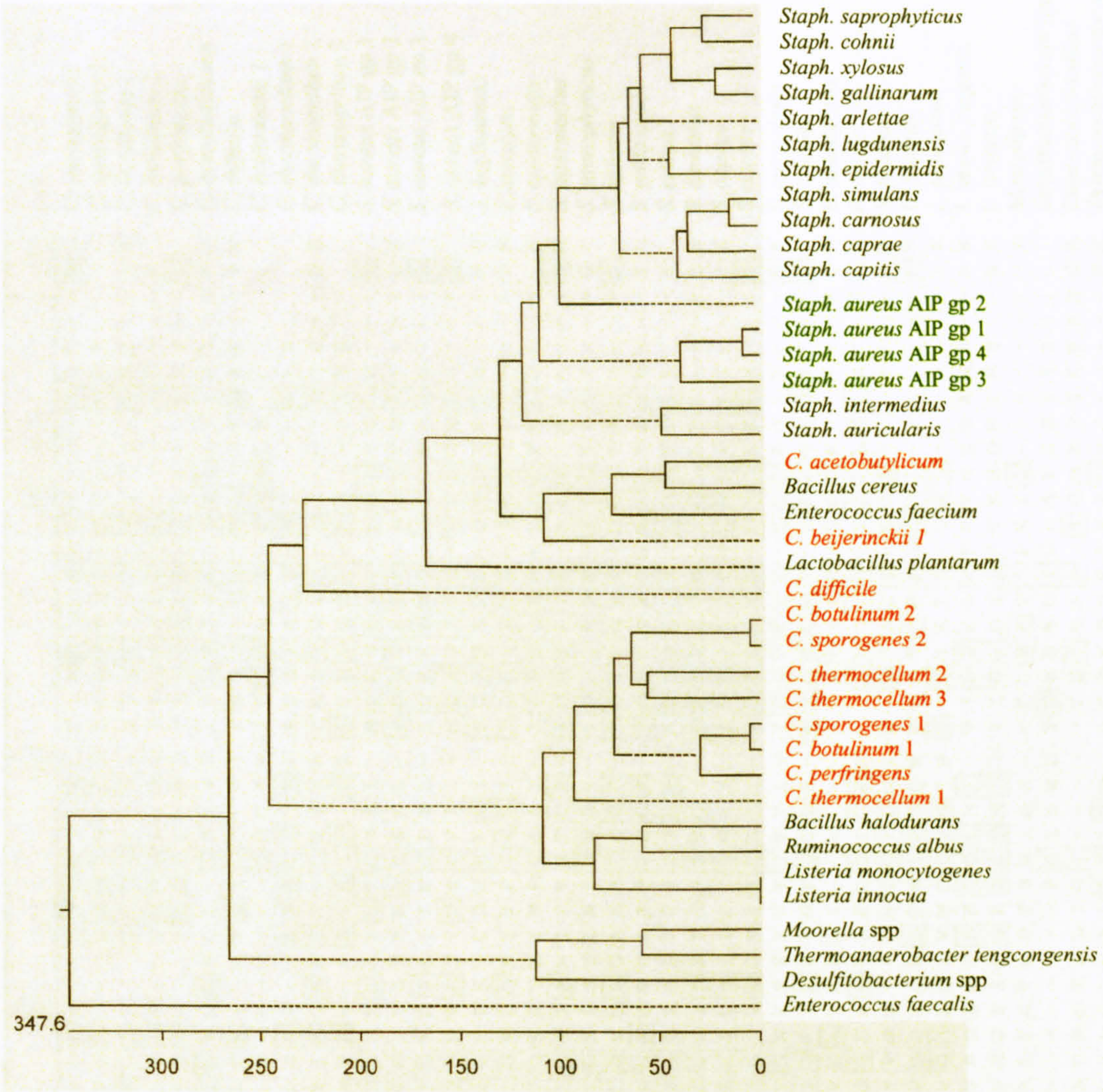


Figure 3.2 - Phylogenetic tree of AgrD proteins of various Gram-positive bacteria. Clostridial species are highlighted in red; *S. aureus* in green.

Figure 3.3 – AgrD protein alignment of various Gram-positive bacteria. The cysteine residue at position 36 is conserved in almost all proteins in the alignment.

A

```
C.spor AgrD1   MKK-LSKKVLMLVATFTTLLASVVASSACVWCVYQPEEPKCLREE
C.spor AgrD2   MKKQLKEKCVKVTAKLLKSVAYSTADSACVFGAYQPKEPKSLRK
C.bot AgrD1    MKK-LNKKVLMLVATFTTLLASIVASSACYWCVYQPKEPKCLREE
C.bot AgrD2    MKKQLKEKCTKVTAKLLKSVAYSTADSACHLGIYQPKEPKSLRK
C.perf AgrD    MKK-LNKNLLTLFAALTTVVATTVATSACLWFTHQPEEPKSLRDE
C.diff AgrD    MKK-FIVRFMKFASSLALSTAILSANSTCPWIIHQPKVPKEISNLKKTN
S.a AIP gpI    MNT-LFNLFFDFITGILKNIGNIAAYSTCDFIMDEVEVPKELTQLHE
S.a AIP gpII   MNT-LVNMFFDFIIKLAKAIGIVGGVNACSSLFDEPKVPAELTNLYDK
S.a AIP gpIII MKK-LLNKVIELLVDFFNSIGYRAAYINCDFLLDEAEVPKELTQLHE
S.a AIP gpIV   MNT-LLNIFFDFITGVLKNIGNVASYSTCYFIMDEVEVPKELTQLHE
```

B

```
C.spor AgrD1    -MKK-LSKKVLMLVATFTT-L--LASVVASSACVWCV-YQPEEPK---CLREE
C.spor AgrD2    -MKKQLKEKCVKVTAKLLK-S--VAYSTADSACVFGA-YQPKEPK---SLRK
C.bot AgrD1     -MKK-LNKKVLMLVATFTT-L--LASIVASSACYWCV-YQPKEPK---CLREE
C.bot AgrD2     -MKKQLKEKCTKVTAKLLK-S--VAYSTADSACHLGI-YQPKEPK---SLRK
C.perf AgrD     -MKK-LNKNLLTLFAALTT-V--VATTVATSACLWFT-HQPEEPK---SLRDE
C.diff AgrD     -MKK-FIVRFMKFASSLAL-S--TAILSANSTCPWII-HQPKVPKEISNLKKTN
C.aceto AgrD    MNLKEQLNKVNDKFIKGLGKASMKIGEQANGKCVLVTLYEPKMPE---ELLKENIDK
```

Figure 3.4 – AgrD protein alignments. **A** – Alignment of putative clostridial AgrD proteins with those of the four *S. aureus* AIP groups. Bold residues indicate the sequence of the mature AIP. **B** - Alignment of putative AgrD proteins from five clostridial species. C. spor, *C. sporogenes*; C. bot, *C. botulinum*; C. perf, *C. perfringens*; C. diff, *C. difficile*; S. a, *S. aureus*; C. aceto, *C. acetobutylicum*. Residues coloured red are completely conserved. Residues coloured blue are partially conserved. Bold residues indicate the possible AIP sequence, with purple residues forming the thiolactone ring and green residues forming the tail. The length of the tail may, however, be shorter than indicated.

3.2.2.2 AgrB alignments

Figure 3.5 shows a phylogenetic tree of the AgrB proteins. As with AgrD, the *C. acetobutylicum* and one of the *C. beijerinckii* AgrB proteins are less closely related than the proteins of the other clostridial species, and the AgrB1 of *C. botulinum* and *C. sporogenes* are more closely related to *C. perfringens* AgrB than to the AgrB2 in the same organisms. Detailed alignments of the AgrB proteins can be found in Appendix 2. Again, there are several regions that are highly conserved amongst Gram-positive bacteria.

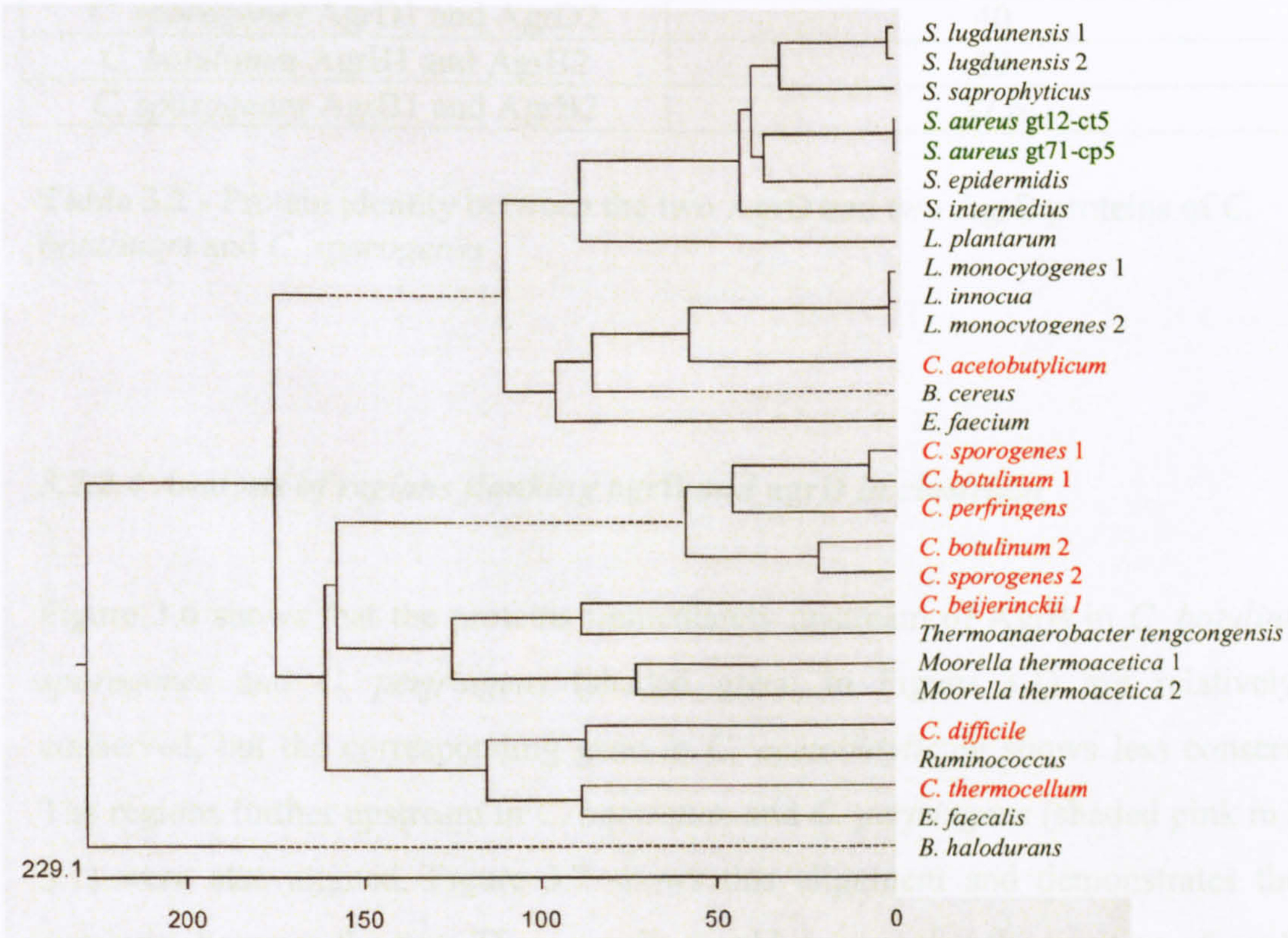


Figure 3.5 - Phylogenetic tree of AgrB proteins of various Gram-positive bacteria. Clostridial species are highlighted in red; *S. aureus* in green.

3.2.2.3 Differences between the two intraspecies AgrD and AgrB proteins in C. botulinum *and* C. sporogenes

Alignments revealed that the two AgrB proteins in *C. botulinum* and *C. sporogenes* are quite different (39% and 37.6% protein identity – see Table 3.2). The two AgrD proteins show slightly higher identity but still show significant differences. Alignments can be found in Appendix 2.

Proteins compared	Protein identity (%)
C. botulinum AgrD1 and AgrD2	40
C. sporogenes AgrD1 and AgrD2	40
C. botulinum AgrB1 and AgrB2	39
C. sporogenes AgrB1 and AgrB2	37.6

Table 3.2 - Protein identity between the two AgrD and two AgrB proteins of *C. botulinum* and *C. sporogenes*

3.2.2.4 Analysis of regions flanking agrB *and* agrD *in clostridia*

Figure 3.6 shows that the proteins immediately upstream of AgrB in *C. botulinum*, *C. sporogenes* and *C. perfringens* (shaded green in Figure 3.1) are relatively well conserved, but the corresponding gene in *C. acetobutylicum* shows less conservation. The regions further upstream in *C. botulinum* and *C. perfringens* (shaded pink in Figure 3.1) were also aligned. Figure 3.7 shows this alignment and demonstrates the high similarity between the two. These results would suggest that the position of *agrB* in *C. sporogenes*, *C. botulinum* and *C. perfringens* is well conserved, as they are all flanked by similar genes.

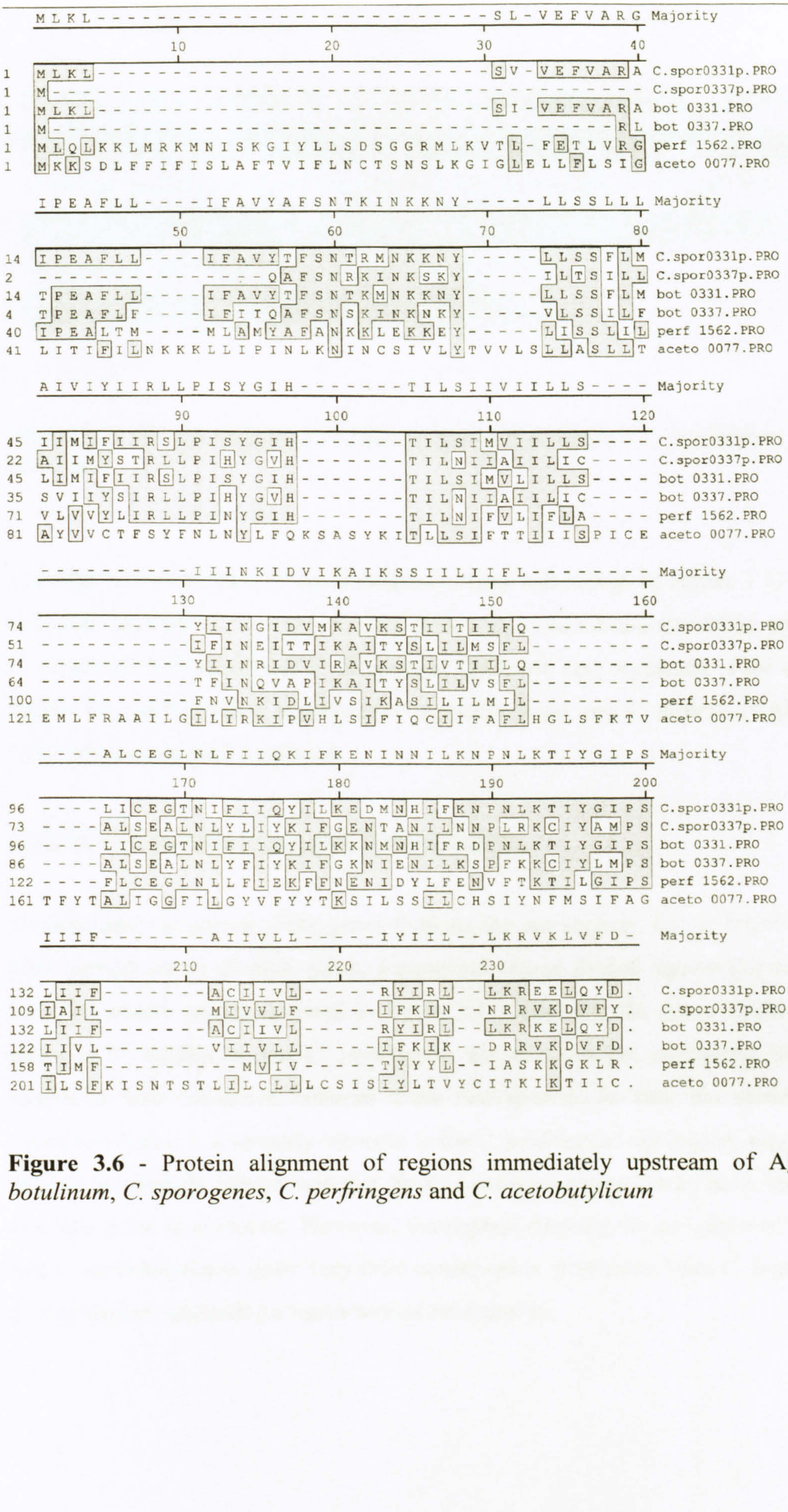


Figure 3.6 - Protein alignment of regions immediately upstream of AgrB in *C. botulinum*, *C. sporogenes*, *C. perfringens* and *C. acetobutylicum*

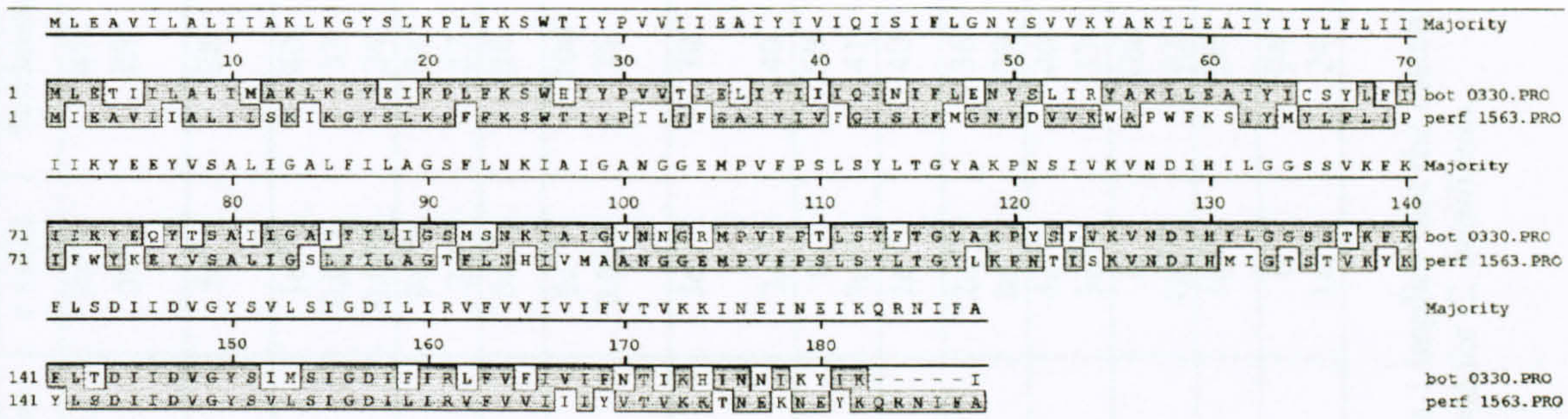


Figure 3.7 - Protein alignment of regions upstream of AgrB in *C. perfringens* and *C. botulinum*

The regions downstream of *agrD* (shaded yellow and orange in Figure 3.1) show high similarity between *C. perfringens* CPE1560 and *C. botulinum* CBO0333 and between *C. botulinum* CBO0340 and *C. botulinum* CBO0336 (see Appendix 2 for alignment). Again, the corresponding regions in *C. acetobutylicum* and *C. difficile* (CAC0082 and CD2749) are less well conserved.

3.2.3 BLAST searches on *agr* and flanking regions

To determine the nature of the genes flanking the *agr* regions, Entrez BLAST searches were carried out on all pink, green, yellow and orange shaded regions shown in Figure 3.1. The results are summarised in Table 3.3. Generally, the genes flanking the *agr* genes in *C. botulinum* and *C. perfringens* are very similar, suggesting that the *agr* system is well conserved between these two species. In fact, the assembly of *C. botulinum* region 1 is virtually identical to the *C. perfringens agr* region, suggesting that the two are directly related, and that the *C. botulinum* region 2 may have been inserted at a later point in evolution. However, the regions flanking the *agr* genes of *C. difficile* and *C. acetobutylicum* show very little conservation with those from *C. botulinum* and *C. perfringens*, indicating a much weaker relationship.

Region	Closest match from BLAST search	Bit score	e value	% identity
C. botulinum CBO0330	• Hypothetical protein CPE1563 from *C. perfringens* strain 13 • ABC-type branched chain amino acid transport system, protease component *from C. thermocellum* ATCC 27405	162 45.8	$3e^{-39}$ $5e^{-4}$	47 28
C. perfringens CPE1563	• ABC-type branched chain amino acid transport system, protease component *from C. thermocellum* ATCC 27405	55.8	$5e^{-7}$	28
C. botulinum CBO0331	• Hypothetical protein CPE1562 from *C. perfringens* strain 13 • Predicted membrane metal-binding protein from *C. thermocellum* ATCC 27405 • AgrC from *S. epidermidis*	140 39.7 38.9	$1e^{-32}$ 0.021 0.036	40 33 24
C. botulinum CBO0337	• Hypothetical protein CPE1562 from *C. perfringens* strain 13 • Two component sensor histidine kinase from *C. perfringens* strain 13	206 32.7	$9e^{-16}$ 2.4	31 27
C. botulinum CBO0335	• Hypothetical protein CPE1562 from *C. perfringens* strain 13	79.9	$5e^{-14}$	28
C. perfringens CPE1562	• Hypothetical protein CPE1562 from *C. perfringens* strain 13 • Signal transduction histidine kinase regulating citrate/malate metabolism from *Syntrophomonas wolfei*	286 42	$5e^{-76}$ 0.015	100 25
C. acetobutylicum CAC0077	• Conserved membrane protein, possible homology of CAAX-like membrane endopeptidase from *C. acetobutylicum* ATCC 824 • Predicted metal-dependent membrane protease from *Nostoc punctiforme* PCC 73102	309 50.8	$3e^{-83}$ $3e^{-5}$	70 40
C. botulinum CBO0333	• Hypothetical protein CPE1560 from *C. perfringens* strain 13 • HD-GYP domain from *C. thermocellum* ATCC 27405	662 306	0 $7e^{-82}$	59 41
C. perfringens CPE1560	• HD-GYP domain from *C. thermocellum* ATCC 27405	302	$1e^{-80}$	40
C. botulinum CBO0336	• Sensory transduction histidine kinase from *C. acetobutylicum* ATCC 824 • Two-component sensor histidine kinase from *C. perfringens* strain 13	292 286	$2e^{-77}$ $8e^{-76}$	36 39
C. botulinum CBO0340	• Sensory transduction histidine kinase from *C. acetobutylicum* ATCC 824 • Two-component sensor histidine kinase from *C. perfringens* strain 13	311 275	$4e^{-83}$ $3e^{-72}$	40 37
C. difficile CD2749	• Putative radical SAM family protein from *C. difficile* 630 • Arylsulfatase regulator (Fe-oxidoreductase) from *Clostridium* spp	915 599	0 $1e^{-169}$	100 62
C. difficile CD2751	• Hypothetical protein CPE0655 from *C. perfringens* strain 13	107	$1e^{-21}$	26
C. acetobutylicum CAC0082	• Predicted membrane protein from *C. acetobutylicum* ATCC 824 • Lantibiotic mersacidin modifying enzyme from *Bacillus halodurans*	1280 134	0 $1e^{-29}$	96 24

Table 3.3 – BLAST search results for proteins encoded by Agr flanking regions in four clostridial species. An e value greater than 1 implies that the match is poor and the identification should be interpreted with caution. Search results for *C. sporogenes* gave the same identification as those for *C. botulinum*.

3.2.4 Expression analysis of *agr* genes in *C. sporogenes*

Little is currently known about how, when and, indeed, if the *agr* genes are expressed in clostridia. The control, timing and mechanism of expression of the system in *S. aureus* have been well characterised (Arvidson, 1983; Janzon & Arvidson, 1990; Novick *et al.*, 1995), but no published data on clostridial *agr* expression is currently available. Various methods can be utilised to assess gene expression, including reverse transcriptase PCR (Carter *et al.*, 2005), Northern blotting (Ji *et al.*, 1995) and real time RT-PCR (Couesnon *et al.*, 2006).

3.2.4.1 Reverse Transcriptase PCR analysis to investigate linkage of agr *gene expression in* C. sporogenes

RT-PCR was used to ascertain whether there was transcriptional linkage between the *agr* genes within each of the two regions. RNA was extracted, according to 2.8.1, during early and late exponential growth, when OD readings were 0.53 and 2.21. The same amount of total RNA (50 ng) was used in each reaction. Nine sets of primers were designed, each of which spanned two genes within the *agr* regions (Figure 3.8(a)). Primer sequences can be found in section 2.6.1. A positive result from the RT-PCR would indicate that a pair of genes was co-expressed. Control reactions were carried out without the reverse transcriptase enzyme, to confirm that the presence of a product was not due to DNA contamination. A positive control was also set up using DNA as the template to confirm that the PCR conditions used were effective.

Figure 3.8(b) suggests that all of the genes in *agr* region 1 are transcriptionally linked, and that transcription occurs during both early and late exponential growth. The RT-PCR results also suggest that, in region 2, gene pairs Cs0336/Cs0337 and *agrD*2/Cs0340 are not co-expressed, but the other genes in region 2 appear to be co-expressed. Again, all pairs which are co-expressed show expression throughout early and late exponential growth (pairs Cs0335/Cs0336, Cs0337/*agrB*2 and *agrB*2/*agrD*2).

(a)

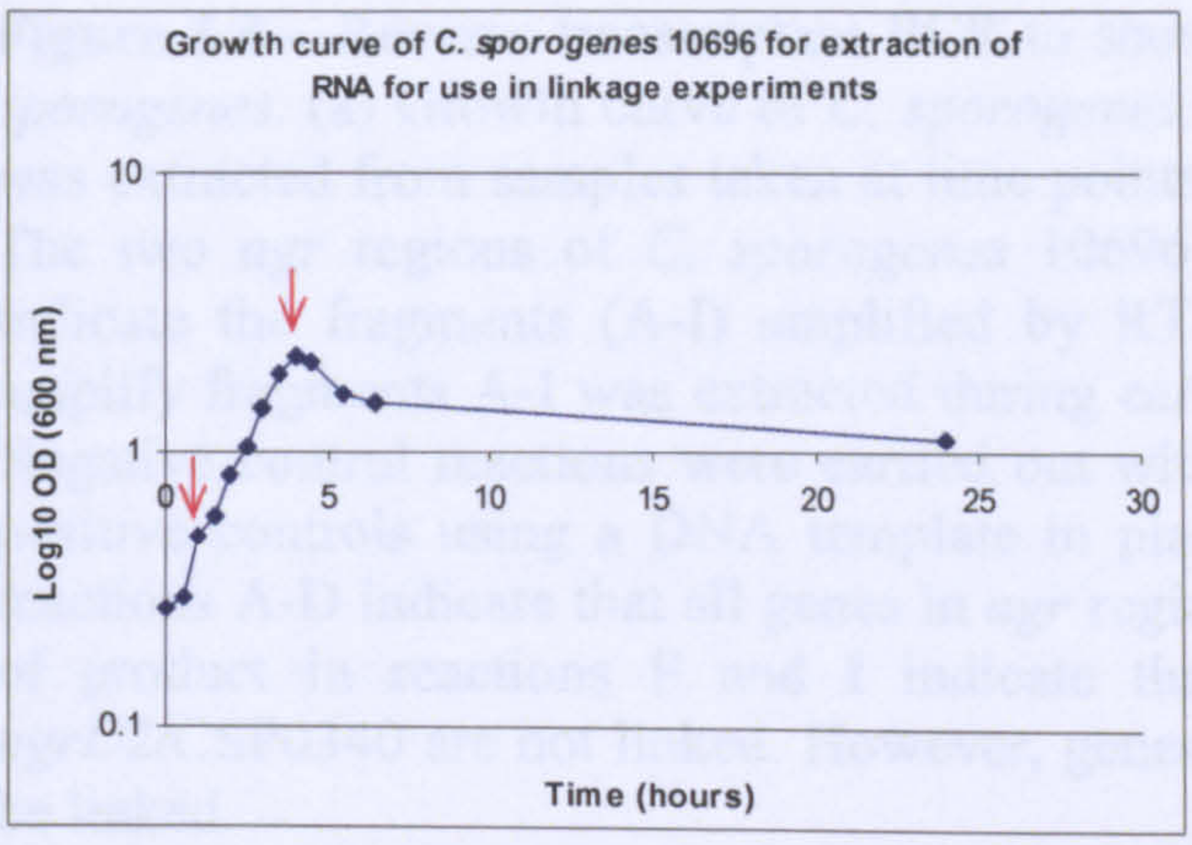
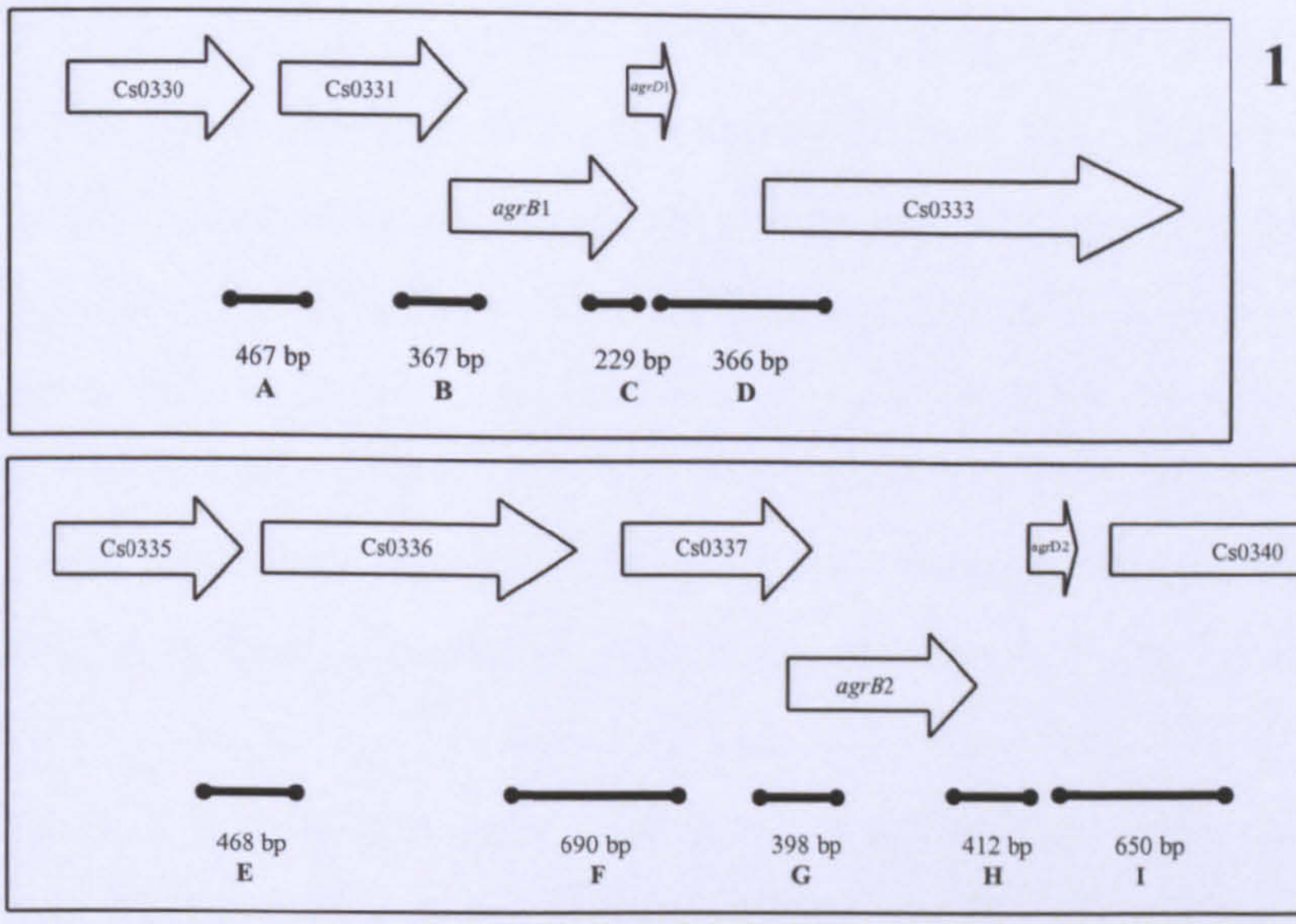


Figure 3.8 – Reverse transcriptase PCR to show linkage of *agr* gene expression in *C. sporogenes*. For full legend see next page.

(b)



(c)

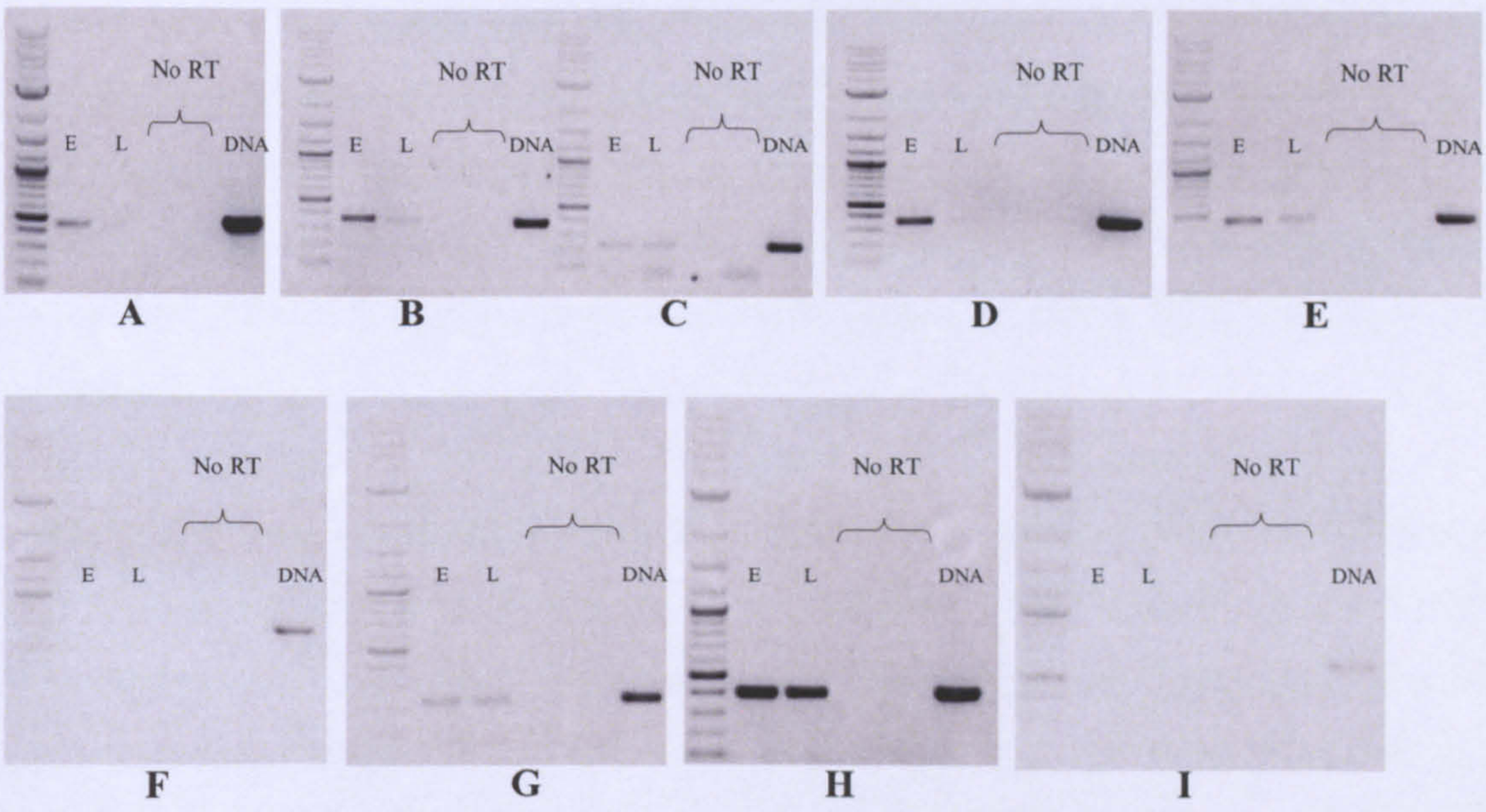


Figure 3.8 – Reverse transcriptase PCR to show linkage of *agr* gene expression in *C. sporogenes*. (a) Growth curve of *C. sporogenes* culture used for RNA extraction. RNA was extracted from samples taken at time points 3 and 9 – indicated by red arrows. (b) The two *agr* regions of *C. sporogenes* 10696 are indicated by 1 and 2. Bold bars indicate the fragments (A-I) amplified by RT-PCR. (c) RNA used in RT-PCRs to amplify fragments A-I was extracted during early (E) and late (L) exponential growth. Negative control reactions were carried out with no reverse transcriptase (no RT) and positive controls using a DNA template in place of RNA (DNA). Positive results in reactions A-D indicate that all genes in *agr* region 1 are transcriptionally linked. A lack of product in reactions F and I indicate that gene pairs CSP0336/CSP0337 and *agrD*2/CSP0340 are not linked. However, genes CSP0337, *agrB*2 and *agrD*2 appear to be linked.

3.2.4.2 Real Time RT-PCR to analyse expression of agrD*1 and* agrD*2 in* C. sporogenes

Since RT-PCR is only a crudely quantitative technique, further investigations were carried out in order to quantitate expression of the *agr* genes throughout the growth curve. Northern blots were attempted to confirm the linkage of expression of the *agr* genes, and to give an indication of the transcript sizes. However, this method proved unsuccessful after several attempts.

An effective and sensitive way of assessing levels of gene expression is by Real Time Reverse Transcriptase PCR. This method avoids the visualisation of PCR products on an agarose gel using ethidium bromide staining, which can be insensitive, and is also a much more sensitive and quantitative method than Northern blotting. It allows the amplification and detection of even tiny amounts of mRNA and enables the quantification of differences in mRNA expression. PCR also has the added advantage of being able to discriminate between very similar mRNAs (Hunt, 2006), and so is useful in this instance where the two *agr* systems will be studied. Real time RT-PCR has been used to assess the levels of expression of various genes, including the neurotoxin genes of *C. botulinum* (Couesnon *et al.*, 2006; Shin *et al.*, 2006; Lovenklev *et al.*, 2004).

The real time PCR method used in this study utilised the double stranded DNA-specific dye SYBR Green. This dye binds to the amplified double stranded DNA and emits a signal 1000 times greater than when unbound (Stratagene, 2007). As the amount of amplified product increases, the amount of fluorescence increases, until it reaches a threshold where it can be detected. The cycle of amplification at which this occurs is referred to as the C_T or threshold cycle, and is determined by the amount of starting template present. A reaction profile obtained from a real time PCR can be found in Appendix 3.

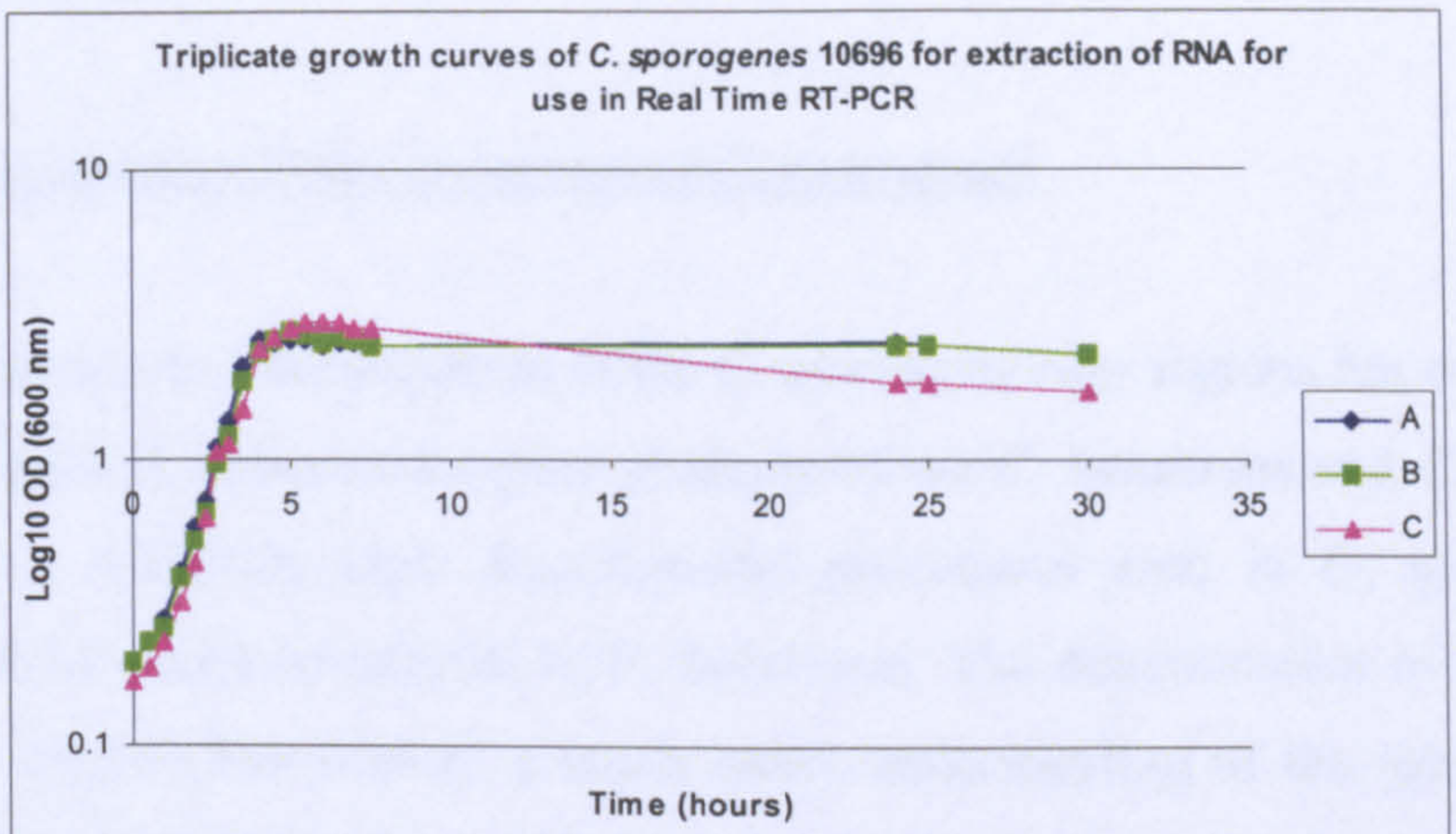
This method was employed to assess the levels of expression of the two *agrD* genes in *C. sporogenes*. Initial problems were encountered whilst attempting the Relative Quantification (RQ) method, favoured by many research groups. The possible low level of expression and the AT-rich sequences in the clostridial *agr* regions resulted in the seeking of an alternative method. Eventually, the Absolute Quantification (AQ) method

was used to successfully assess the level of expression at five time points during growth. Primers were designed against the two *agrD* genes and the 16S rRNA gene, which was to be used as the endogenous control. 16S is a popular choice of endogenous control as its expression does not change throughout growth (Kouguchi *et al.*, 2006; Couesnon *et al.*, 2006; Lovenklev *et al.*, 2004).

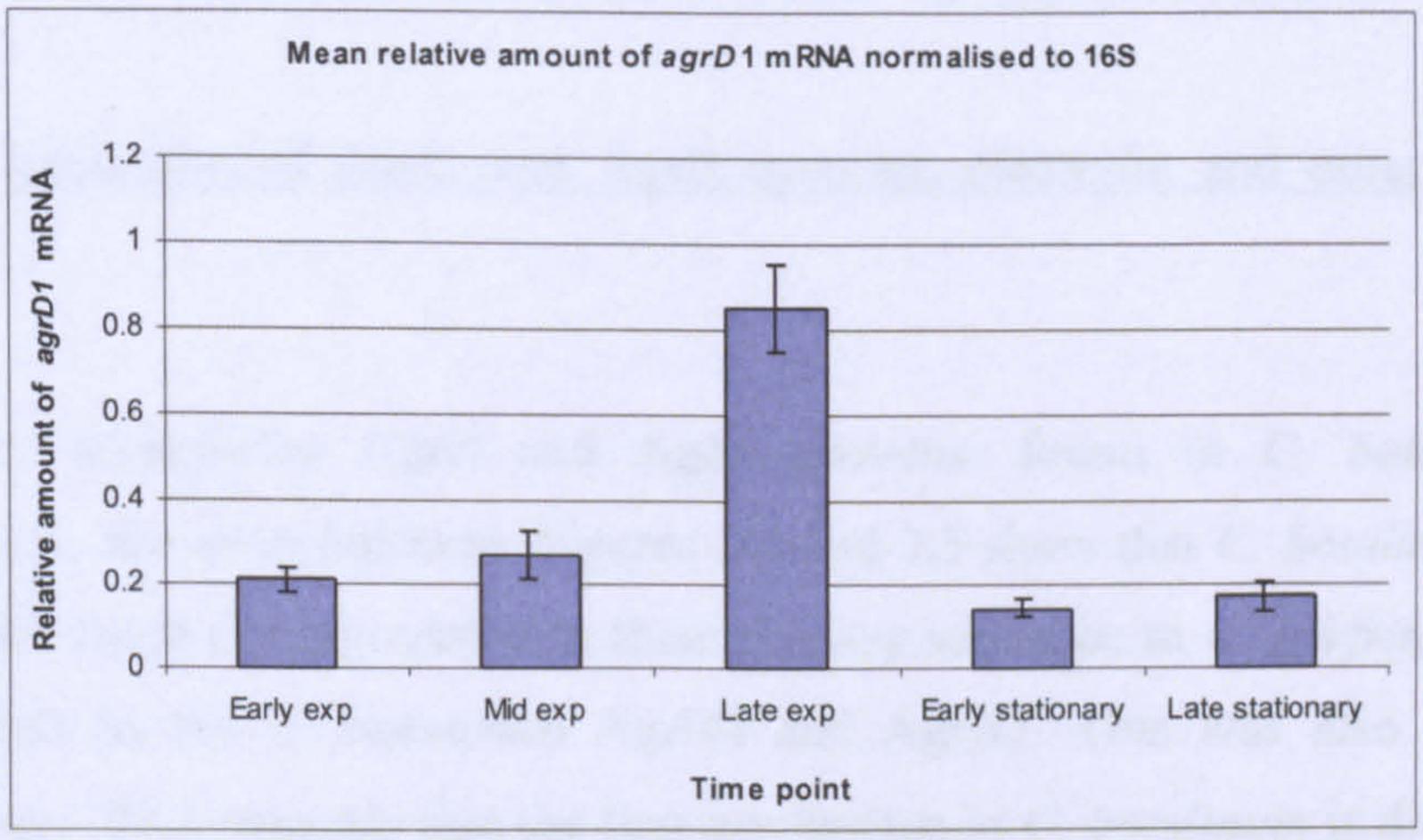
Triplicate 100 ml broth cultures of *C. sporogenes* were grown in TYG broth and growth was monitored using OD (Figure 3.9A). At each of five time points, 10 ml culture was removed and RNA extracted according to 2.8.2. Samples were taken during early, mid and late exponential growth, along with early and late stationary growth. OD readings at these time points were approximately 0.4, 1.1, 2.0, 2.6 and 2.5 respectively.

The AQ method necessitates the construction of a standard curve for both the test gene and the control gene for every run carried out. Curves were constructed using 10-fold dilution series of *C. sporogenes* cDNA. Concentrations used ranged from 100 – 0.001 ng per reaction. The curve was constructed by plotting the log of the concentration of cDNA in the sample against the threshold cycle. A typical standard curve graph obtained is shown in Appendix 3. The level of expression of each *agrD* gene could then be calculated using the standard curves, and this was normalised to the level of expression of the 16S gene. This was carried out for all three samples from each of the five time points, giving average values as relative amounts of *agrD* mRNA.

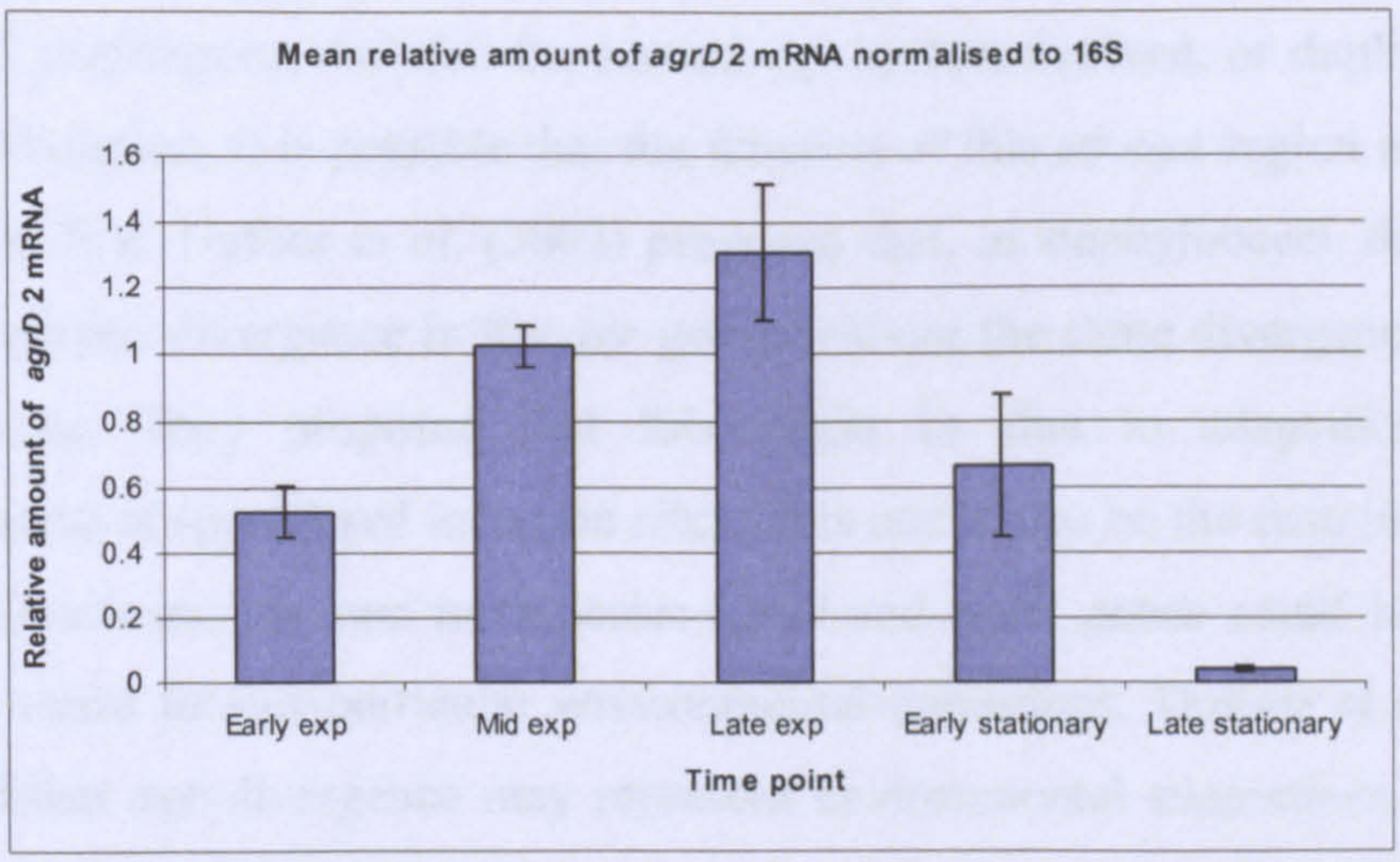
Figure 3.9 (B and C) suggests that *agrD* gene expression increases with time during exponential growth, and then falls dramatically once stationary phase is reached. This corresponds with what has previously been demonstrated in *S. aureus*, where the genes of the *agr* system are maximally expressed towards late exponential to early stationary growth (Arvidson, 1983; Janzon & Arvidson, 1990; Novick *et al.*, 1995).



A



B



C

Figure 3.9 – Real Time RT-PCR to assess *agrD* expression in *C. sporogenes*. A, growth curve of triplicate *C. sporogenes* broth cultures used for RNA extraction. 10 ml samples were taken at early, mid and late exponential growth, along with early and late stationary growth. B, relative *agrD*1 mRNA expression at the five time points. C, relative *agrD*2 mRNA expression at the five time points. Both *agrD* genes show an increase in expression throughout exponential growth, reaching maximal expression at late exponential growth.

3.3 Discussion

3.3.1 Sequencing of the *agr* regions of *C. sporogenes*

The completion of the sequence of the *C. sporogenes agr* regions has confirmed that the protein identity between the gene products of the *C. botulinum* and *C. sporogenes agr* regions is relatively high. Experimental procedures used in *C. sporogenes* should therefore be easily transferred to *C. botulinum*. The determination of the sequences of the *agr* regions has allowed a much better understanding of the genes that comprise them, and their relationship with other clostridial *agr* gene homologues.

3.3.2 Comparison of AgrD and AgrB amongst clostridia and other Gram-positive bacteria

The two intraspecies AgrD and AgrB proteins, found in *C. botulinum* and *C. sporogenes*, are quite different. Figures 3.2 and 3.5 show that *C. botulinum* AgrD1 and AgrB1 are more closely related in their primary sequence to *C. perfringens* AgrD and AgrB than to the *C. botulinum* AgrD2 and AgrB2. This was also the case in *C. sporogenes*. This suggests that the first *agr* system in *C. botulinum* is directly related to that of *C. perfringens*, and that the second *agr* system evolved, or duplicated at a later point in evolution. It is possible that the function of this second region may differ from that of the first. Dufour *et al.* (2002) proposed that, in staphylococci, different species have developed divergence in the *agr* genes without the same divergence in the rest of the genome. They proposed that this might be due to adaptation to different environments at specialised infection sites. This could also be the case in *C. sporogenes* and *C. botulinum*; the two intraspecies *agrD* and *agrB* genes could have evolved at different times to suit particular environmental conditions. Dufour *et al.* (2002) also proposed that *agr* divergence may represent environmental adaptations, whose effects on pathogenesis could be merely incidental, and that the locus is also important in the control of catabolic pathways, nutrient uptake and energy metabolism.

Interestingly, the *E. faecalis* AgrD was not closely related to that of any of the other Gram-positive bacteria, as shown in the alignment in Figure 3.3. It is known that *E. faecalis* AgrD is very different to other AgrD proteins (Sturme *et al.*, 2002; Nakayama

et al., 2006). This organism uses a GBAP (gelatinase biosynthesis-activating pheromone) cyclic lactone AIP, whose precursor is coded for by the *fsrD* gene, and which does not have the well-conserved cysteine that AIPs of most other Gram-positive bacteria share. This is commonly the key residue in the formation of the thiolactone ring of the AIP, but in *E. faecalis* formation involves a serine residue. Figure 3.3 shows that the proteins in the alignment that do not possess the highly conserved cysteine residue belong to: one strain of *C. thermocellum*, *Staphylococcus intermedius* and *Moorella* spp (which have a serine in its place) and *E. faecalis*. AIPs of strains of *S. intermedius* lack the conserved cysteine residue and contain a cyclic lactone, rather than a thiolactone (Kalkum *et al.*, 2003).

Also highly conserved among the staphylocccal AgrD proteins is the region immediately downstream of the AIP (Figure 3.4). In *S. aureus* groups I and IV this sequence is DEVEVP, and varies only slightly in groups I and III. This region is likely to be involved in C terminal processing in the formation of the AIP (Lyon & Novick, 2004). Interestingly, this sequence is not found in any of the clostridial species analysed; the only conserved residue is the proline. However, the clostridia appear to share their own conserved putative processing site based on the sequence YQPKEP. The clostridial AIP sequences show little similarity to those found in *S. aureus*. This is not surprising, as the two genera are found in completely different environments, where they will have developed their own systems to suit very different environmental niches.

The N terminal processing site is not as conserved in either the staphylococci or the clostridia. It is understandable that more conservation is found in the C terminal regions as, in staphylococci, the C terminal of the protein is located in the cytoplasm and the group specific processing of AgrD is determined by this terminal. The conservation of the N terminal is therefore not so critical (Zhang *et al.*, 2004).

The processing of the AgrD propeptide to form the mature AIP differs from other quorum sensing systems. AgrD is subject to two proteolytic cleavage events, with the AIP sequence being located in the middle of the protein. In other systems, the precursor is subject to only one cleavage, with the C terminal portion serving as the signal molecule (Qiu *et al.*, 2005).

Novick (2003) demonstrated that the last two residues of the AIPs in *S. aureus* were usually limited to the hydrophobic residues F, L, V, Y and sometimes M. This is the case with the putative AIPs in *C. botulinum* and *C. sporogenes*, but not in *C. perfringens* or *C. difficile*. In fact, none of the clostridial putative AIPs show much similarity to the *S. aureus* AIPs, except for the conserved cysteine; again suggesting that the two species' *agr* systems are highly divergent. The different putative clostridial AIPs are also very different from each other, which could be explained by the different species of clostridia occupying different environmental niches and therefore developing different systems to suit that niche. The similar AIPs found in *S. aureus* all belong to one species, whereas the different putative AIPs analysed in the clostridia all belong to different species. It is therefore no surprise that they show limited similarity.

In general, the staphylococci seem to show more regions of conservation in AgrD than the clostridia. The staphylococci also share a conserved glycine eight or nine residues N terminal to the conserved cysteine (Lyon *et al.*, 2004). Again, this tight conservation is not seen in the clostridia.

There are extensive regions of conservation in the AgrB proteins of the staphylococci. However, many of these are not shared by the clostridia, which generally show fewer regions of similarity. There are, however, some regions that are highly conserved among all organisms in the alignment (see Appendix 2). Generally, the AgrB proteins show more regions of conservation throughout both the staphylococci and the clostridia, than were observed in the AgrD proteins. This is to be expected, as the sequence of the AIP must be highly specific in order to bind to the receptor region of AgrC. AgrB is involved in the processing of AgrD so it is not so critical that its sequence must differ greatly from other AgrB proteins. Indeed, the interaction between AIP and AgrC is specific between staphylococcal AIP groups, but the specificity of the interaction between AgrB and AgrD is not so strict (Zhang & Ji, 2004).

The most notable conserved residue amongst the AgrB proteins is the cysteine at position 116, which is absolutely conserved among all organisms in the alignment (see Appendix 2). Lyon and Novick (2004) proposed that this conserved cysteine in the AgrB proteins of *S. aureus* may play a role in the formation of a thioester acyl-enzyme intermediate, which is important in the processing of AgrD.

Also absolutely conserved among the AgrB proteins of all organisms in the alignment is a histidine at position 109. Qiu *et al.* (2005) found that replacing the conserved histidine or cysteine with an alanine residue in *S. aureus* group I AgrB totally eliminated the ability to process AgrD, showing that the two residues are required for post-translational processing. All *S. aureus* AIP groups also share a serine residue at position 113, but the only clostridia that have this residue are *C. perfringens* and *C. acetobutylicum*. Qiu *et al.* (2005) proposed that this residue is also required for processing of AgrD, but found that its replacement with alanine had no inhibitory effect on processing. This indicated that this serine is not required for post-translational processing, and explains why it is not conserved among the clostridia. Qiu *et al.* (2005) concluded that the histidine and cysteine residues are putative catalytic residues that form the catalytic centre, and are located on the inner surface of the cytoplasmic membrane. However, it is so far unclear how the interaction between AgrB and AgrD occurs, and how this results in the generation of the mature AIP (Zhang & Ji, 2004).

3.3.3 Analysis of regions flanking *agrB* and *agrD* in clostridia

The results obtained from alignments of the proteins encoded by the genes flanking the *agr* genes indicate that the position of *agrB* and *agrD* in *C. sporogenes*, *C. botulinum* and *C. perfringens* is well-conserved. The corresponding regions in *C. acetobutylicum* and *C. difficile* are less well conserved, but this analysis confirms that all five clostridial species contain regions showing some homology to the *S. aureus agr* regions.

The BLAST searches carried out on proteins encoded by the flanking regions confirmed the conservation of the position of the *agr* genes. In the searches on the regions upstream of *agrB* the closest match for CB00331, CBO0335 and CB00337 was CPE1562 from *C. perfringens*. This protein also shows some similarity to a histidine kinase (Table 3.1), which may play some role in the *agr* system (although it does not bear any resemblance to any *agrC* gene products). The proteins encoded further upstream (CBO0330 and CPE1563) also showed high similarities. This confirms that the two regions upstream of *agrB* in *C. perfringens*, and *agrB*1 in *C. botulinum* and *C. sporogenes* are all homologous. CPE1563 was identified as the putative protease component of an ABC-type branched-chain amino acid transport system. ABC-type

transport systems are known to be involved in quorum sensing in Gram-positive bacteria, but as yet, these genes' involvement in the *agr* system of clostridia is unknown.

As yet, it has not been determined precisely how the AgrD protein is processed and exported. Until recently, the only published hypothesis proposed that the AgrB protein is responsible for both of these processes (Sturme *et al.*, 2002). However, Saenz *et al.*, (2000) demonstrated that over-expression of AgrB and AgrD in *S. epidermidis* did not cause an increase in the production of AIP. This led them to the conclusion that the processing or secretion of the AIP might require an additional factor besides AgrB. Indeed, a recent publication by Kavanaugh *et al.* (2007) has proposed that a type I signal peptidase, SpsB, may in fact be responsible for the N-terminal cleavage of AgrD. The fact that its activity was unaltered in an *agr* mutant indicated that the peptidase is chromosomally encoded outside the *agr* locus (Kavanaugh *et al.*, 2007). However, this group did not determine the mechanism by which the AIP is exported.

Most quorum sensing systems identified to date in Gram-positive bacteria involve the processing of the AIP, followed by its secretion via an ABC transporter. Usually the proteins involved in these processes are encoded by genes in the same operons as those that encode the pro-peptide and the two component signal transduction systems (for example the *L. lactis* nisin operon) (Zhang *et al.*, 2002; Kleerebezem & Quadri, 2001). This fact, taken with the proposal that there may be some other factors involved in the processing and excretion of the AIP, may provide some substance to the theory that an ABC transporter is involved in *C. botulinum* and *C. sporogenes*. Interestingly, genes 0330 and 0331 (identified as an ABC-transporter and sensor kinase respectively) were shown to be transcriptionally linked to the *agr* genes in 3.2.4.1. Indeed, *agrB* genes investigated by Zhang *et al.* (2002) did not exhibit a recognisable ATP binding motif. It is therefore possible that this function is carried out by a separate ABC transporter gene (0330 in the case of *C. botulinum* and *C. sporogenes*).

The gene products encoded upstream of AgrB in *C. difficile* and *C. acetobutylicum* seem less closely related to those of *C. botulinum*, *C. sporogenes* and *C. perfringens*, but were both matched with proteases.

The regions downstream of *agrD* show high similarity in *C. botulinum* (CBO0333) and *C. perfringens* (CPE1560). The nearest match for the proteins of both genes was CPE1560. The top hit in the BLAST search of this protein is a HD-GYP domain from *Clostridium thermocellum*. HD proteins are a super family of metal-dependent phosphohydrolases. A sub-class of these proteins possess an additional GYP motif, and evidence suggests that this class of proteins is involved in bacterial signal transduction. In the plant pathogen *Xanthomonas campestris* pv. *campestris*, virulence factor production and biofilm dispersal are positively controlled by a two component signal transduction system whose response regulator, RpfG contains such a domain (Ryan *et al.*, 2006). It was demonstrated that the HD-GYP domain of the RpfG protein was involved in the degradation of cyclic di-GMP, and that mutation of the H and D residues resulted in the loss of enzymatic activity against cyclic di-GMP and the regulation of virulence factor production. Increased levels of the cyclic di-GMP second messenger promoted biofilm formation, whereas reduced levels promoted virulence factor production (Ryan *et al.*, 2006). HD-GYP domains have been shown to be conserved among a wide range of bacteria and are therefore of significant importance in gene regulation. These proteins of the *agr* system exhibiting homology to a HD-GYP region could therefore be involved in signal transduction or possibly quorum sensing.

The products of the two genes either side of *agr* region 2 in *C. botulinum* (CBO0336 and CBO0340) are also similar, and were matched against a sensory transduction histidine kinase of *C. acetobutylicum*. Again, histidine kinases are also known to be involved in quorum sensing.

Generally, there appear to be several genes in the *agr* regions of *C. botulinum and C. perfringens*, which show similarities to genes involved in signal transduction. Their role in the *agr* response in clostridia is not yet known, but it is possible that they are somehow involved in the signalling of this system, particularly in view of the fact that the identification of *agrA* and *agrC* homologues has not been confirmed in these species. The data obtained also suggest that the *agr* region 1 in *C. botulinum* is directly related to that of *C. perfringens*. The arrangement of the genes in these two species is identical, as are their identifications from the BLAST search. It appears that somewhere during evolution, the *C. botulinum agr* system has been duplicated and inserted further downstream.

The downstream regions of *C. acetobutylicum* and *C. difficile* did not show close matches with any of the other clostridial *agr* region genes investigated. This would indicate that the *agr* region position is not well conserved between these species and *C. perfringens*, *C. botulinum* and *C. sporogenes*. *C. acetobutylicum* has been shown to possess its own *agrA* and *agrC* genes, and so there may not be the need for other genes involved in signal transduction in close proximity. Furthermore, the putative AgrD protein in *C. acetobutylicum* shows very little similarity to other clostridial AgrD proteins, indicating that this species' *agr* system evolved quite separately from the others.

It appears that the *C. difficile agrB* and *agrD* genes are isolated, and are not immediately flanked by any genes that could possibly be involved in signal transduction. Their presence could therefore be due to an insertion or a duplication somewhere in evolution, leaving any related genes elsewhere in the chromosome.

In summary, further work is clearly necessary to investigate the precise mechanism of action of the *agr* systems in the clostridia. It is also, as yet, unclear which genes or processes the systems control. The *agr* systems which exist in the clostridia bear a strong resemblance to those in the staphylococci, although their mechanism of action is clearly different. It is not clear how the flanking genes are involved in the *agr* response, but it is quite apparent that, certainly in *C. botulinum* and *C. perfringens*, the *agr* systems are closely related and that there are various genes flanking them which may be involved in signal transduction. The system identified in *C. difficile* is much less closely related, and *C. acetobutylicum* appears to have evolved independently of the other clostridial species; a hypothesis which is reinforced by the presence of all four *agr* genes.

3.3.4 Expression analysis of *agr* genes in *C. sporogenes*

RT-PCR analysis indicated transcriptional linkage of all genes in *agr* region 1 of *C. sporogenes* 10696. This is consistent with expression in *S. aureus*, where the *agr ACDB* genes are transcribed together as part of the P2 operon (Novick *et al.*, 1995). All reactions demonstrated the presence of this transcript during both early and late exponential growth. Fragments A, B and D (Figure 3.8 (b)) appear to show more intense bands in the reactions using RNA from early exponential growth. However, because of the limitations of this technology, these data cannot be used to draw conclusions on the quantity of expression of this transcript. This theory is reinforced by the fact that the genes spanning fragment C (*agrB*1/*agrD*1) appear to be co-expressed (albeit weakly) at equal levels during early and late exponential growth (Figure 3.8 (b) C). However, their flanking genes (which, according to the results, appear to be co-expressed) show a higher intensity band during early exponential phase. A more reliable technique should therefore be used for quantification of the expression of these genes.

All of the genes in *agr* region 2 do not appear to be co-expressed as a single operon, as pairs Cs0336/Cs0337 and *agrD*2/Cs0340 did not show co-expression. It is no coincidence that genes Cs0336 and Cs0340 are both orphan kinases, which have an, as yet, unknown function in the *agr* regions of *C. sporogenes* and *C. botulinum*. To date, *agrC* and *agrA* homologues have not been identified in these *agr* systems. It could be proposed that one, or both, of the Cs0336 and Cs0340 genes, which are very closely homologous, may function as the *agrC* of the system. However, these results suggest that, if this is indeed the case, the kinase(s) are not transcribed with the other genes of the *agr* system, in contrast to what is observed in *S. aureus*.

It is certain that these results must be confirmed in further investigations. Although RT-PCR is an extremely sensitive technique, and has been used previously to demonstrate transcriptional linkage in *C. difficile* (Carter *et al.*, 2005), it cannot be relied on as a quantitative technique, and the reverse transcriptase reaction may be inhibited by the presence of contaminating RNase enzymes. The technique may also be hindered by difficulties in amplifying templates with large amounts of secondary structure (Roche, 2006). Northern blot analysis was therefore attempted to confirm the RT-PCR results, and ascertain the definitive transcript sizes. After a number of attempts, no bands were

observed in Northern blot experiments, despite loading up to 15 μg RNA onto the gel. This is probably because the *agr* genes, as in *S. aureus*, are naturally expressed at very low levels, which are too low to be detected using Northern blotting. Groups who have successfully used Northern blotting to demonstrate *agr* expression have all probed against either the RNAIII gene (Traber & Novick, 2006; Ji *et al.*, 1995) or the α or δ haemolysin genes (Janzon *et al.*, 1989; Janzon & Arvidson, 1990; Traber & Novick, 2006). These genes are not available for targeting in the clostridia. In addition, most groups to carry out this technique in staphylococci have used laboratory strains of *S. aureus*, which express the *agr* genes at higher than natural levels (Geisinger *et al.*, 2006; Ji, *et al.*, 1995; Janzon *et al.*, 1989). This is because most laboratory strains are derivatives of strain NCTC 8325, which has recently been shown to behave differently to clinical isolates of *S. aureus* (Cassat *et al.*, 2006). Derivatives of this strain, such as RN6390, contain a mutation in the *rsbU* gene, which reduces *sigB* expression and in turn results in higher expression of RNAIII than that found in clinical isolates (Cassat *et al.*, 2006). This is a result of *sigB* being a negative regulator of the *agr* system. Indeed, Cassat *et al.* (2006) demonstrated that RN6390 showed higher RNAIII and protease expression than the clinical isolate UAMS-1. It was confirmed, using quantitative PCR, that RN6390 produced approximately 200-fold more RNAIII than UAMS-1 during exponential growth. Expression was also compared to nine other clinical isolates and found to be higher than all of these. The absence of RNAIII and haemolysin genes in clostridial species meant that there was no other option but to probe for the *agr* genes themselves. Unfortunately, it appeared that their low level of expression was not sufficient to enable detection using this technique.

To determine reliably the precise time of expression of the *agrD* genes, quantitative real time PCR was the method of choice. This is an excellent method for quantifying changes in gene expression (Ambion, 2006) and was used by Couesnon *et al.* (2006) to demonstrate that expression of the *C. botulinum botR/A*, *bont/A* and *antp* (associated non-toxic protein) genes occurred concomitantly at the transition between exponential and stationary growth. They were also able to demonstrate that the *botR/A* expression level was approximately 100-fold less than those of the *bont/A* and *antp* genes. This group also used the absolute quantification technique, indicating that problems may have been encountered with the more favourable relative quantification technique.

Quantitative RT-PCR data can be expressed relative to an internal standard (relative quantification) or an external standard curve (absolute quantification) (Bustin & Nolan, 2004). The RQ technique relies upon consistent and highly efficient amplification of the cDNA template. Initial standard curve reactions are carried out to demonstrate that the efficiency of the PCR of the control gene is similar to that of the test gene. Software provided with the real time PCR machine plots a standard curve of cDNA template concentration against threshold cycle. The efficiency of the reaction can be determined from the slope of the line. A slope of –3.32 indicates a PCR efficiency of 100%. A more negative slope indicates a less efficient reaction. This method can only be used accurately if the standard curves of both the test and control gene are efficient, and the two are similar. This means that the results from the RQ run for each gene can be directly compared.

Several attempts at standard curve generation were made, using both the *agrB* and *agrD* genes as target genes. Unfortunately, the standard curves obtained consistently showed very poor PCR efficiency. There was also no similarity between the test and control curves and the curves were not reproducible. This meant that the RQ method could not be employed using the *C. sporogenes* RNA. These findings were unusual, but not altogether surprising. Clostridial DNA is notoriously AT rich, which causes problems when designing efficient PCR primers (the overall GC content of the *C. botulinum* genome is only 28.2% [Sebaihia *et al.*, 2007]). Although the primers were designed using the Primer Express software (Applied Biosystems), initial searches did not yield any primers. The parameters were therefore altered in order for suitable primers to be obtained. This meant that even the best possible primers from the sequence provided were not ideal and there was no guarantee that they would produce an efficient PCR. The priming of the primers greatly affects the efficiency of the PCR, which would explain why the reaction efficiency was poor; there was a poor choice of primers in the AT rich sequence. Primers were re-designed, but the efficiency did not improve.

Altering the concentrations of cDNA used to produce the standard curves also failed to improve them. The only option was therefore to use the absolute quantification method. This method is not ideal, as it involves the use of many more wells than the RQ method and if the efficiencies are different each time, one run cannot be easily compared to another. Standard curves for the test and control genes must be amplified in the same

plate, and the test data are then normalised to the control gene by dividing the target amount by the control gene amount. The fact that the efficiencies were so different for each run carried out meant that they could not easily be compared between runs. However, this problem is overcome to a certain extent by normalising each set of data to the control gene. Other groups carrying out real time PCR on clostridial RNA have also resorted to alternative methods, indicating that they may also have experienced problems with the RQ method. Along with Couesnon *et al.* (2006), who used the absolute quantification method, Kouguchi *et al.* (2006) used a similar method where the standard curve was constructed using plasmids containing the gene of interest, so that the exact template copy number in the reaction could be determined.

The problems encountered with the real time PCR experiments must be borne in mind when interpreting the results. Standard curves in this study were generally of poor quality, indicating the poor efficiency of the reactions. This occasionally necessitated the removal of some of the data points from the curve, in order to improve it. In doing so, there was the danger that the test samples may not have fallen into the range of the standard curve. This problem was accounted for by reducing the quantity of template added to each test reaction to 0.1 ng. In theory, if inhibitors were present in the samples, and were responsible for the poor efficiency of reaction, reducing the amount of template would therefore reduce the amount of inhibitor and improve the efficiency of the reaction. As this had no effect on efficiency it can be concluded that the poor standard curves were probably not due to the presence of inhibitors in the test samples. With AQ, since small differences in nucleic acid input can lead to large differences in PCR product yield, the amount of starting material must be quantified extremely accurately in order to normalise sample data and correct for tube to tube differences (Bustin & Nolan, 2004). However, the cDNA was quantified using the extremely accurate NanoDrop spectrophotometer before being added to the PCR so this should not have been a problem.

One of the most obvious explanations for the poor efficiency is the AT rich target sequence, which could have resulted in poor primer annealing. The temperature used for annealing was reduced from 60°C to 58°C but this failed to improve the efficiency.

In summary, the results obtained using the AQ method suggest that *agrD* expression is maximal during late exponential growth and then drops off significantly once stationary phase is reached. The higher expression of the *agr* genes later on in growth seems logical, as this suggests similarities to the expression levels observed in the staphylococci. In addition, as the *agr* system in *C. sporogenes* appears to be involved in sporulation (see Chapter 4), it seems logical that its genes are expressed maximally towards late exponential growth when sporulation is initiated. The real time PCR reactions were repeated and, although the magnitude of the results varied, the expression pattern did not, leading to the conclusion that the *agr* system in *C. sporogenes* is maximally expressed during late exponential growth.

Chapter 4.0

Gene Modulation Techniques in *Clostridium sporogenes*

Chapter 4 – Gene Modulation Techniques in *Clostridium sporogenes*

4.1 Introduction

Until recently, methods for knockout mutagenesis in clostridia through allelic exchange had long eluded the clostridial scientific community. Most attempts have been based on the inactivation of target genes through homologous recombination but efforts over the last 20 years have been largely unsuccessful. Recently, two areas have been focussed on as alternative methods of gene inactivation or modulation in clostridia. One area utilises the replication machinery of unstable plasmid vectors; the other relies on the production of antisense RNA fragments which interfere with translation of target proteins.

4.2 Unstable plasmids as a tool for gene knockout in clostridia

Focus of attention has recently been turned to integrative vector systems in an attempt to achieve gene knockout through double crossover allelic replacement. This strategy exploits the replication mechanism of the unstable single stranded or rolling circle (RC) plasmids. Early investigations into the problems of cloning in *B. subtilis* indicated that a contributory factor could be the fact that the vector plasmids used in the cloning experiments accumulated large amounts of single stranded DNA which was unstable and recombinogenic (Niaudet *et al.*, 1982; Projan *et al.*, 1987). It was later demonstrated that replication of unstable plasmids stimulates DNA recombination (Noirot *et al.*, 1987). In fact, Gruss & Ehrlich (1989) demonstrated that recombination was stimulated 150-1500-fold in rolling circle plasmids, compared with the chromosome or plasmids that do not generate single stranded DNA. Many Gram-positive bacterial plasmids replicate via the RC mechanism, which is illustrated in Figure 4.1 and proceeds as follows.

A plasmid-encoded initiator protein, known as a Rep protein, and possessing specific nicking activity, binds to a specific binding site of approximately 30 base pairs, contained within the double stranded origin (dso) or + origin of the plasmid (Khan, 1996; Khan, 2000). The Rep protein then nicks the leading strand of the double stranded DNA, causing it to bend at the origin of replication. Cleavage by the Rep protein results

in a free 3' OH group, whilst the Rep protein is attached to the 5' phosphate of the DNA. A host helicase unwinds the DNA, allowing the 3' OH group to serve as a primer for synthesis by DNA polymerase III (Khan, 1996; Khan, 2005).

After the new leading strand has been completely synthesised, the Rep protein carries out a second nicking event, releasing a single stranded DNA molecule, corresponding to the original leading strand of the plasmid. This molecule contains a single stranded origin (sso) or – origin, which is required for the conversion of this single stranded molecule back to a double stranded one (Khan, 2005). The single stranded DNA intermediate is highly unstable and highly recombinogenic. These properties can be exploited in an attempt to induce homologous recombination with the host chromosome, and subsequent inactivation of a target gene. The unstable plasmid should then be easily lost from the population.

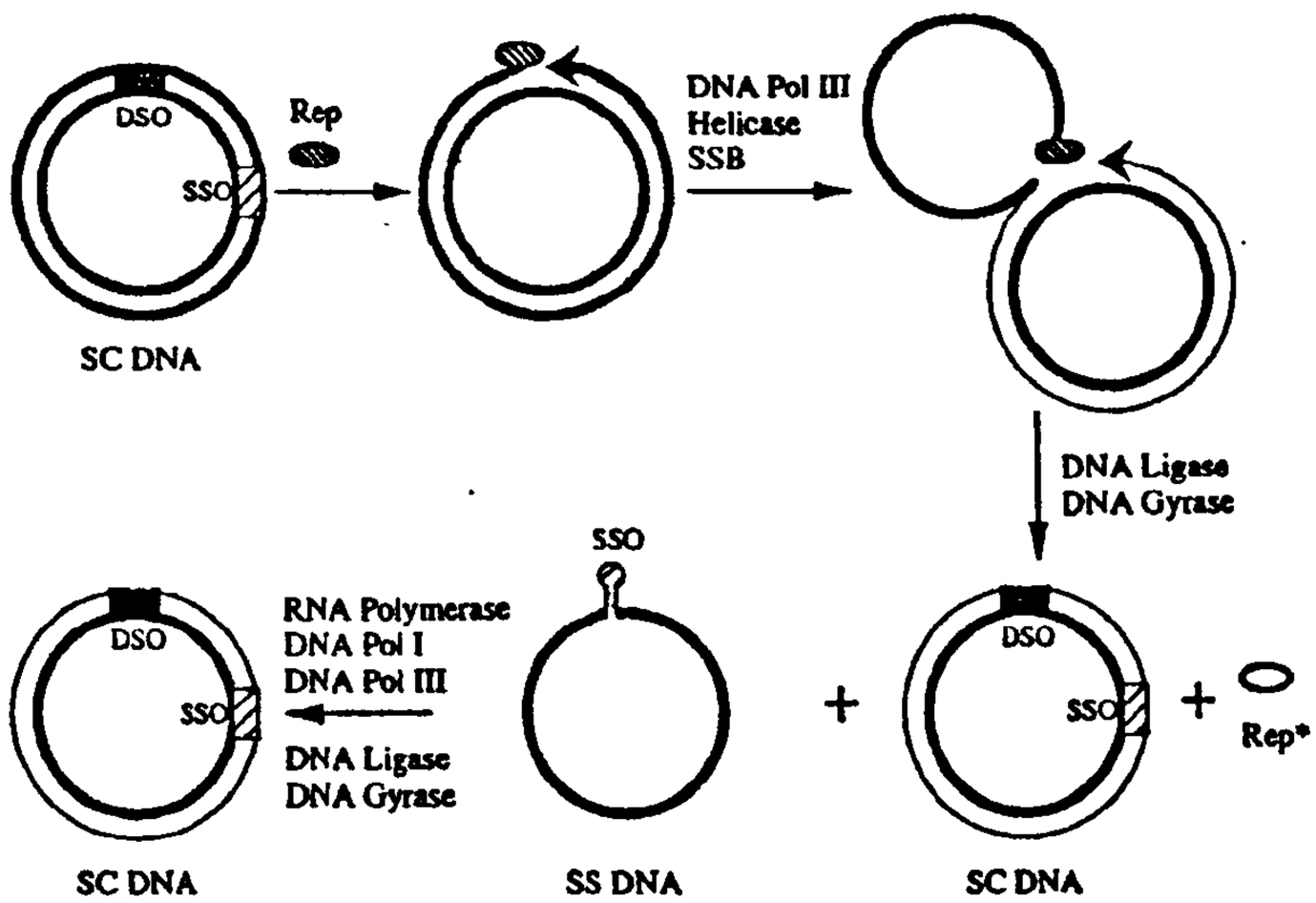


Figure 4.1 - Replication in rolling circle plasmids (Khan, 1996). A single stranded intermediate is produced which is highly unstable and therefore highly recombinogenic.

An attempt to utilise the properties of RC plasmid replication to induce double crossover events was previously carried out by Pennington (2006). The plasmid constructed (pMTL5112) possessed the *B. subtilis* pIM13 replication region and an erythromycin resistance cassette. The plasmid was deficient in the negative origin so

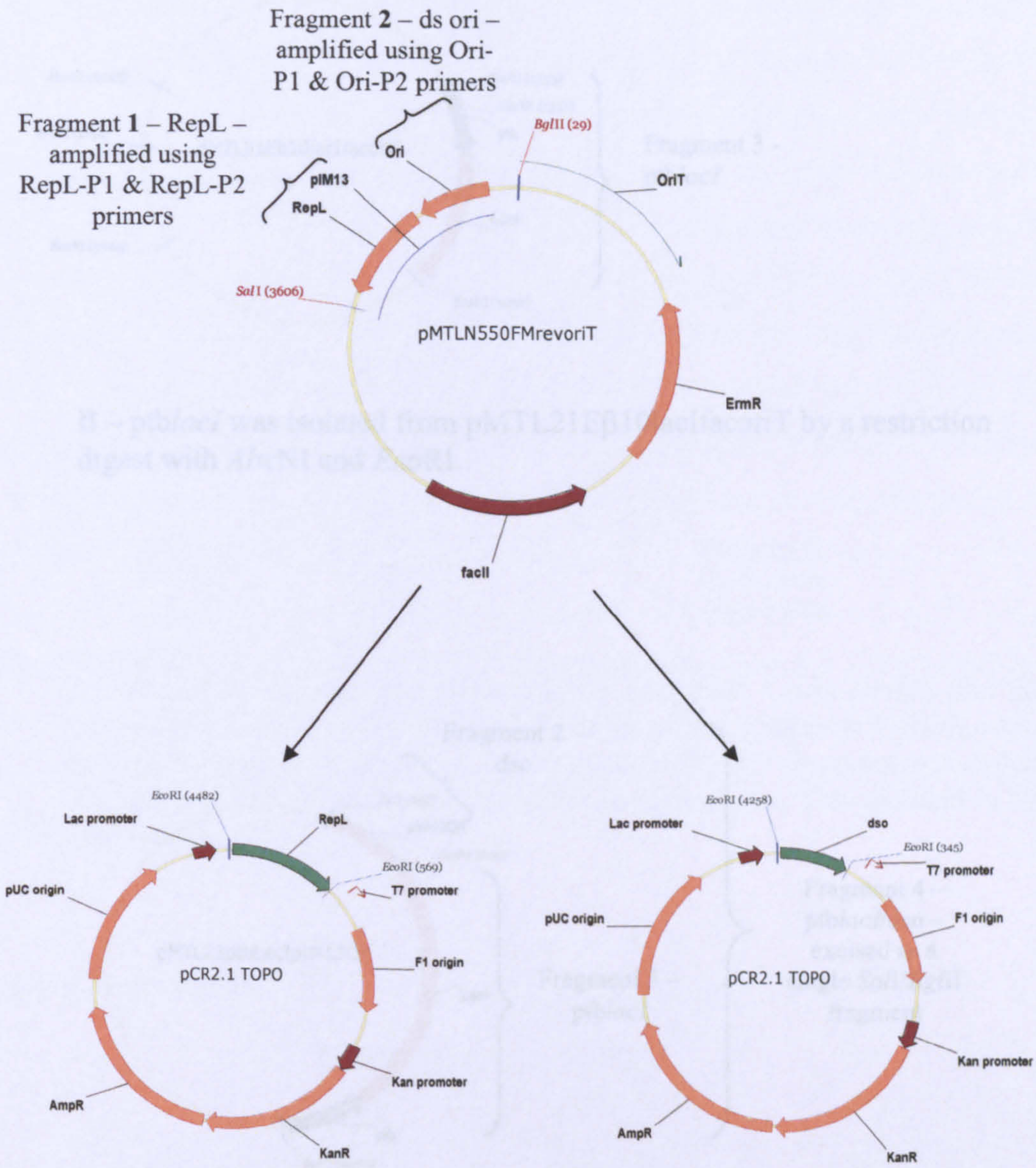
that any single stranded DNA produced could not be converted back into double stranded DNA and may integrate into the host chromosome, inactivating any given target gene. This attempt was unsuccessful, so the plasmid was developed further in this investigation to render *repL* under inducible control.

The overall strategy involved the replacement of the *B. subtilis* pIM13 replicon region of plasmid pMTL5112 with the pIM13 double stranded origin and the *lac* operon repressor *lacI*, which is under control of the *C. acetobutylicum ptb* (phosphotransbutyrylase) promoter. *repL* from pIM13 was then inserted under the control of the *Clostridium pasteurianum* ferredoxin promoter *facII*. This promoter is fused to the *lac* operator, rendering it inducible by IPTG. The plasmid should then function in the following way to achieve double crossover mutagenesis:

In the absence of IPTG, LacI should bind to the *lac* operator, preventing transcription of the *lac* operon. On addition of IPTG, an artificial inducer of the *lac* operon, LacI will be displaced, allowing transcription to proceed from the *facII* promoter, driving the transcription of *repL*. The RepL protein will act *in trans* on the double stranded origin, allowing replication to proceed. The single stranded intermediate generated via this replication process is lacking in the – origin, and therefore fails to be converted back into double stranded DNA. Consequently, the cell should contain large amounts of highly unstable, highly recombinogenic single stranded DNA. A mutant version of a target gene of interest, containing an antibiotic cassette, may be inserted into the *Bam*HI site of the integrative vector. Once this is introduced into a clostridial strain, and induced, single stranded DNA containing the mutant gene should be produced and double crossover recombination should take place between the single stranded mutant gene and the chromosomal target gene. The unstable single stranded DNA should be lost after allelic exchange, leaving only the clostridial chromosome containing the mutant gene.

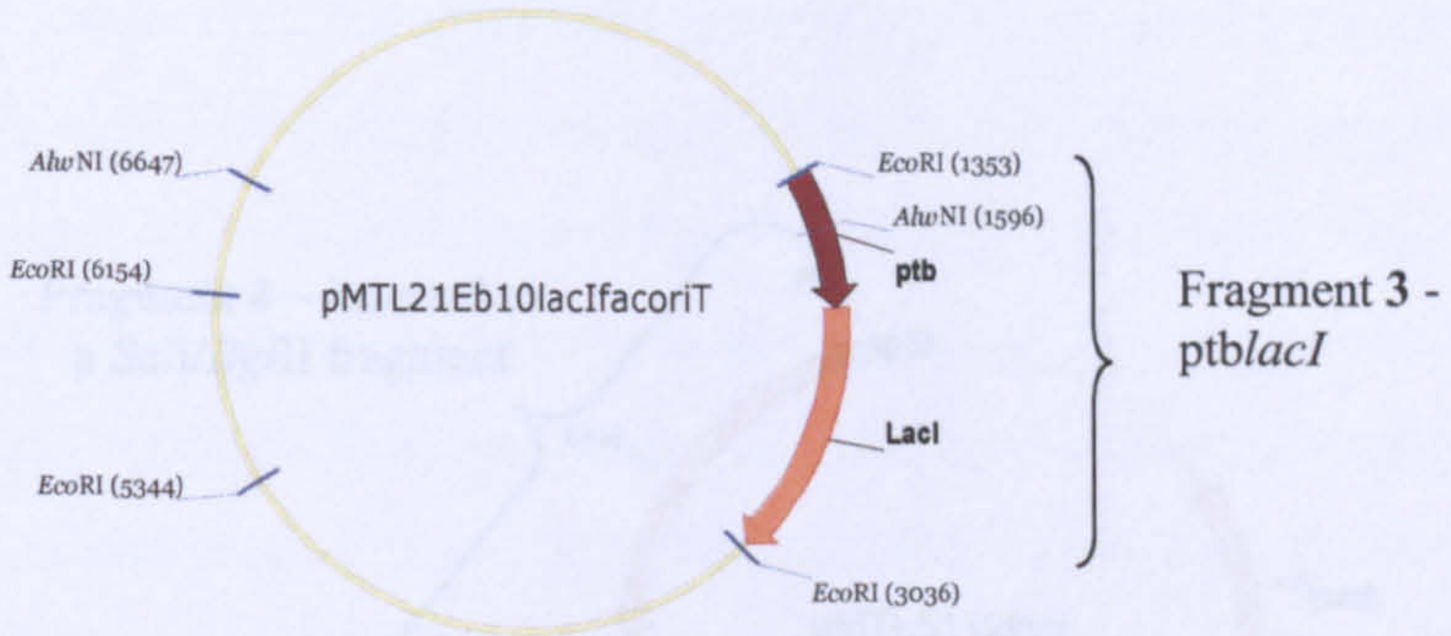
4.2.1 Results

The strategy used to generate the integrative vector is shown in Figure 4.2. First, the *repL* gene and the dso from plasmid pMTL5112 were amplified by PCR and cloned into the pCR2.1-TOPO vector (Figure 4.2A). The primers for the *repL* gene (RepL-P1 and RepL-P2) were designed to incorporate *Xba*I and *Eco*RV restriction sites at the 5' and 3' termini of the PCR product. Those for the dso (Ori-P1 and Ori-P2) were designed to incorporate *Nhe*I and *Sal*I sites at the 5' and 3' termini. *ptblacI* was then isolated from the pAMβ1-based vector pMTL21Eβ10lacIfacoriT using *Alw*NI (before blunt ending) and *Eco*RI (Figure 4.2B). *ptblacI* was cloned into the pMTL23 'holding' vector as a *Sma*I/*Eco*RI fragment. pIM13 dso was then cloned as a *Sal*I/*Nhe*I fragment into the same vector, downstream of *ptblacI* (Figure 4.2C). The whole *ptb*/*lacI*/dso fragment was then excised as a *Sal*I/*Bgl*II fragment and cloned back into the original pMTL5112 vector. Finally, the pIM13 *repL* fragment was isolated as an *Xba*I/*Eco*RV fragment from pCR2.1-TOPO and cloned into the final vector downstream of the *facII* promoter. This resulted in the integrative plasmid pMTL5112lacI (Figure 4.2D). Three separate double restriction digests, using *Nde*I/*Sal*I, *Hind*III/*Bam*HI and *Nde*I/*Xba*I, were carried out to confirm the correct assembly of the final integrative vector.

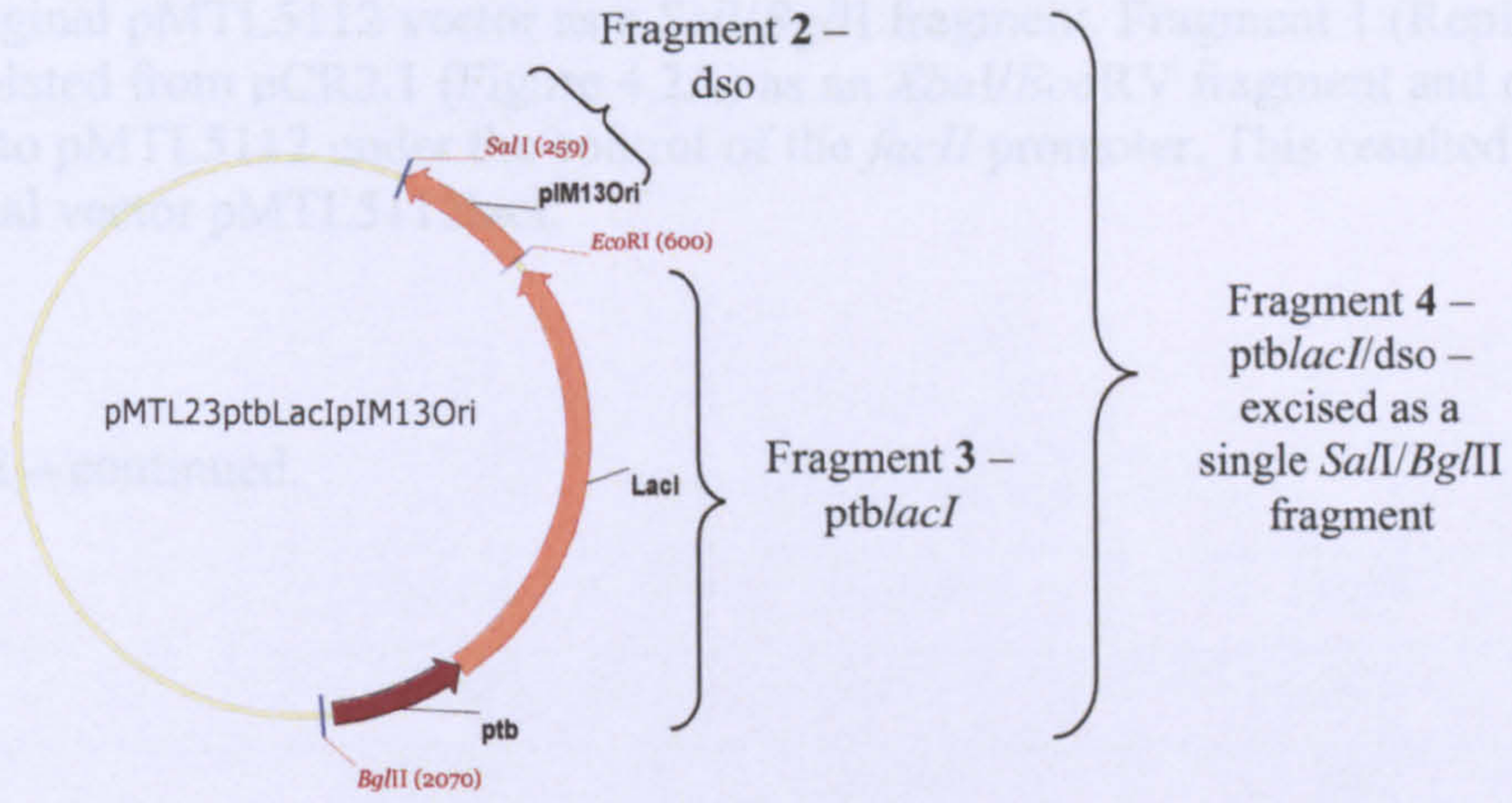


A – *repL* and pIM13 ds ori were amplified from pMTL5112 by PCR. The two fragments were cloned separately into the pCR2.1 TOPO vector.

Figure 4.2 – Strategy used to generate integrative vector pMTL5112lacI.

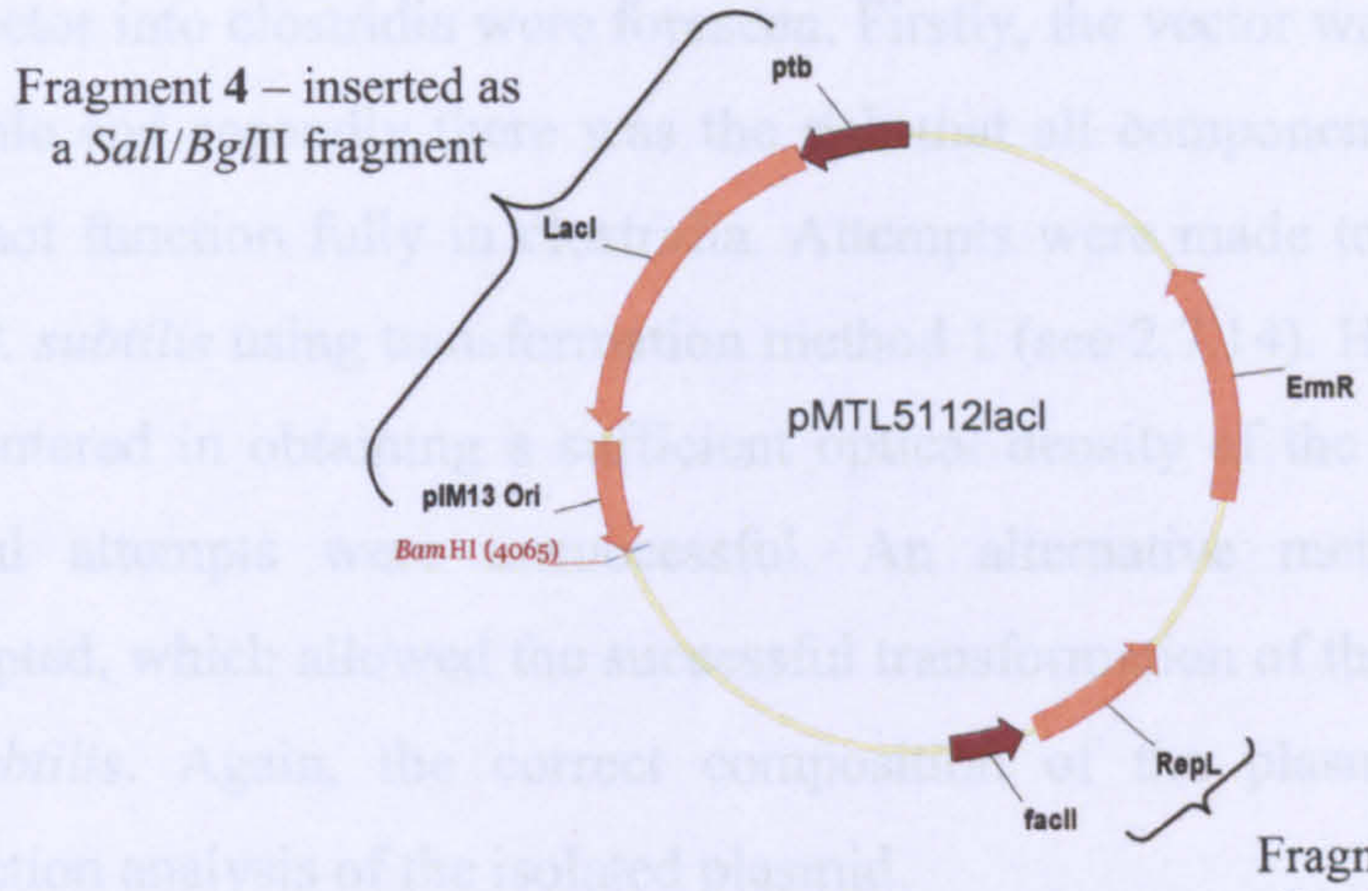


B – ptb*lacI* was isolated from pMTL21Eβ10lacIfacoriT by a restriction digest with *Alw*NI and *Eco*RI.



C – dso fragment 2 and ptb*lacI* fragment 3 were inserted into the pMTL23 holding vector. Both were then excised together as *Sal*I/*Bgl*II fragment 4.

Figure 4.2 – continued.



D – Assembly of final integrative vector. Fragment 4 was cloned into the original pMTL5112 vector as a *Sal*I/*Bgl*II fragment. Fragment 1 (RepL) was isolated from pCR2.1 (Figure 4.2A) as an *Xba*I/*Eco*RV fragment and cloned in to pMTL5112 under the control of the *facII* promoter. This resulted in the final vector pMTL5112lacI.

Figure 4.2 – continued.

4.2.1.1 Plasmid segregational stability in Bacillus subtilis

Confirmation of the plasmid's instability in clostridia needed to be carried out. *Bacillus subtilis* Cu2189 was used as an initial recipient, as problems in introducing the vector into clostridia were foreseen. Firstly, the vector was expected to be highly unstable and secondly there was the risk that all components of the vector system may not function fully in clostridia. Attempts were made to transform the plasmid into *B. subtilis* using transformation method 1 (see 2.7.14). However, problems were encountered in obtaining a sufficient optical density of the *B. subtilis* culture, and several attempts were unsuccessful. An alternative method (see 2.7.15) was attempted, which allowed the successful transformation of the integrative vector into *B. subtilis*. Again, the correct composition of the plasmid was confirmed by restriction analysis of the isolated plasmid.

Assessment of segregational stability of the plasmid was carried out according to method 2.9. Stability is expressed as percentage plasmid loss in Table 4.1, which illustrates that the plasmid proved highly unstable, even in the absence of IPTG induction.

IPTG	Percentage Plasmid Loss		
	Day 3	Day 4	Day5
+	99.997	100	100
-	99.990	99.999	100

Table 4.1 - Segregational stability of integrative vector pMT5112lacI in *B. subtilis* 168, expressed as percentage plasmid loss.

4.2.1.2 Plasmid segregational stability in Clostridium sporogenes

The promising instability of the plasmid in *B. subtilis* allowed the progression to the introduction of the plasmid into *C. sporogenes*, the organism originally intended as the target. The plasmid was first transformed into the conjugation donor *E. coli* CA434 and the plasmid isolated from resulting clones in order to confirm its identity using restriction digests. Conjugation was then attempted with *C. sporogenes* 10696

using two different conjugation methods (methods 2.7.17 and 2.7.18), and the addition of IPTG concentrations of 0, 0.01 and 0.1 mM to the agar culture media. Several attempts using both methods proved unsuccessful. As a last resort, the conjugation was repeated using an uncharacterised mutant strain of *C. sporogenes*. This strain had previously contained a plasmid over-expressing a recombinant nitroreductase enzyme (Pennington, 2006), and had then been cured of the plasmid. The strain was found to be much more receptive to conjugation, which was possibly due to an unidentified mutation. Conjugation into this strain (*C. sporogenes* 10696-1) using the 'spot' method was successful in producing up to 310 transconjugants/ml. Plasmids of those subcultured were screened using restriction analysis and sequencing of the *lacI* and *repL* genes.

Segregational stability was determined using the same method as for *B. subtilis* and is shown in Table 4.2. The plasmid appeared to be slightly more stable in *C. sporogenes* 10696-1 than had previously been observed in *B. subtilis*, showing 0.00005% (+ IPTG) and 0.00006% (- IPTG) of the plasmid remaining after day 6, whereas complete loss was observed by day 5 in *B. subtilis*. However, the plasmid was still highly unstable.

IPTG	Percentage Plasmid Loss			
	Day 3	Day 4	Day 5	Day 6
+	99.62069	99.98808	99.99956	99.99995
-	99.85100	99.99083	99.99952	99.99994

Table 4.2 - Segregational stability of integrative vector pMTL5112lacI in *C. sporogenes* 10696-1, expressed as percentage plasmid loss.

4.2.1.3 Construction and insertion of a pyrF *knockout cassette for use with integrative vector pMTL5112lacI*

The final stage of this investigation was to attempt to produce a double crossover knockout mutant using pMTL5112lacI, which had proved highly unstable; a property required for integration into the host chromosome and plasmid loss to occur. The *pyrF* gene was selected as a target for knockout, as a relatively simple assay can be utilised to test for mutants. *pyrF* codes for the orotidine 5'-phosphate decarboxylase protein, which is an essential enzyme in pyrimidine biosynthesis. Inactivation of this gene would result in uracil auxotrophic mutants, which could be screened by a simple assay, streaking putative mutants onto minimal agar medium in the presence and absence of uracil. The knockout cassette was constructed by first isolating the *catP* (chloramphenicol acetyltransferase) gene which had previously been cloned from *C. perfringens* into the pCR2.1 TOPO vector to make pCR2.1CatPΔBsrGI (Pennington, 2006). This was to be used as the antibiotic resistance cassette. The *catP* fragment was excised by restriction digestion with *Pvu*II, resulting in a blunt-ended fragment. The *C. sporogenes pyrF* gene, along with partial sequences of its flanking genes *pyrK* and *pyrC*, had also been previously cloned into pCR2.1, resulting in pCR2.1::*pyrF* (Pennington, 2006). The *pyrF* gene was cut centrally, using *Acc*I, and blunt ended. The *catP* gene was then ligated into the middle of the *pyrF* gene, resulting in an inactive *pyrF* gene contained within the pCR2.1 vector. Clones were verified using restriction analysis and sequencing, and then the whole knockout cassette excised as an *Eco*RI fragment, before blunt ending. The integrative vector pMTL5112lacI was digested at the *Bam*HI site and also blunt ended. The *pyrF* knockout cassette was then ligated into the integrative vector, and verified using restriction analysis (using *Nde*I/*Sal*I, *Bsr*GI/*Hpa*I and *Not*I/*Bgl*II), PCR screening and sequencing. After transforming the completed vector into *E. coli* CA434, conjugation was carried out using the 'spot' method with *C. sporogenes* 10696 wild type and *C. sporogenes* 10696-1. As was encountered before the knockout cassette was inserted, the conjugation was only successful when the cured strain was used; no transconjugants were obtained from *C. sporogenes* 10696 wild type. Plasmids were isolated, transformed back into *E. coli* TOP10 and confirmed using restriction analysis. Induction of plasmid loss and integration was carried out according to method 2.10 which was supplied by Philippe Soucaille (personal communication).

Plasmid loss was not nearly as high as before the knockout cassette was inserted, suggesting that the plasmid had become more stable (Table 4.3). Numbers of colonies isolated on TYG + erythromycin and TYG + thiamphenicol were very similar (data not shown), which would indicate that there were no thiamphenicol resistant, erythromycin sensitive colonies present. Ten colonies were picked from various plates during the induction experiment and plated in duplicate onto both TYG containing erythromycin and TYG containing thiamphenicol. All isolates grew on both plates, confirming that the plasmid had not been lost from these isolates. The induction of plasmid loss and integration was repeated but similar, erratic results were obtained.

IPTG	Percentage Plasmid Loss			
	Day 3	Day 4	Day 5	Day 6
+	45.143	8.995	57.647	18.750
-	9.615	0.829	-143.836	1.429

Table 4.3 - Segregational stability of integrative vector pMTL5112lacI containing *pyrF* knockout cassette in *C. sporogenes* cured strain, expressed as percentage plasmid loss.

Another possibility was that a single crossover event had occurred instead of a double crossover. This would mean that the plasmid itself had inserted into the chromosome, and would explain the existence of thiamphenicol resistant, erthromycin resistant colonies. The two possibilities which could occur from a single crossover event are shown in Figure 4.3. In order to investigate this possibility, a colony from day 6 of the induction experiment (named SC3) was subcultured and used in further investigations. The plasmid DNA from this strain was isolated and transformed into *E. coli* TOP10, after which two clones were analysed. PCR screening was carried out to verify whether the plasmid was still present. Primers RNA/2F and PyrC/1F(R) (which both bind in the plasmid) were used to screen both the *E. coli* clones containing the re-transformed plasmid (ECSC3a and ECSC3b), along with the plasmid prepared straight from the clostridial SC3 isolate. Results of this PCR (screen 1) are shown in Figure 4.4B, along with the binding sites of the primers (Figure 4.4A).

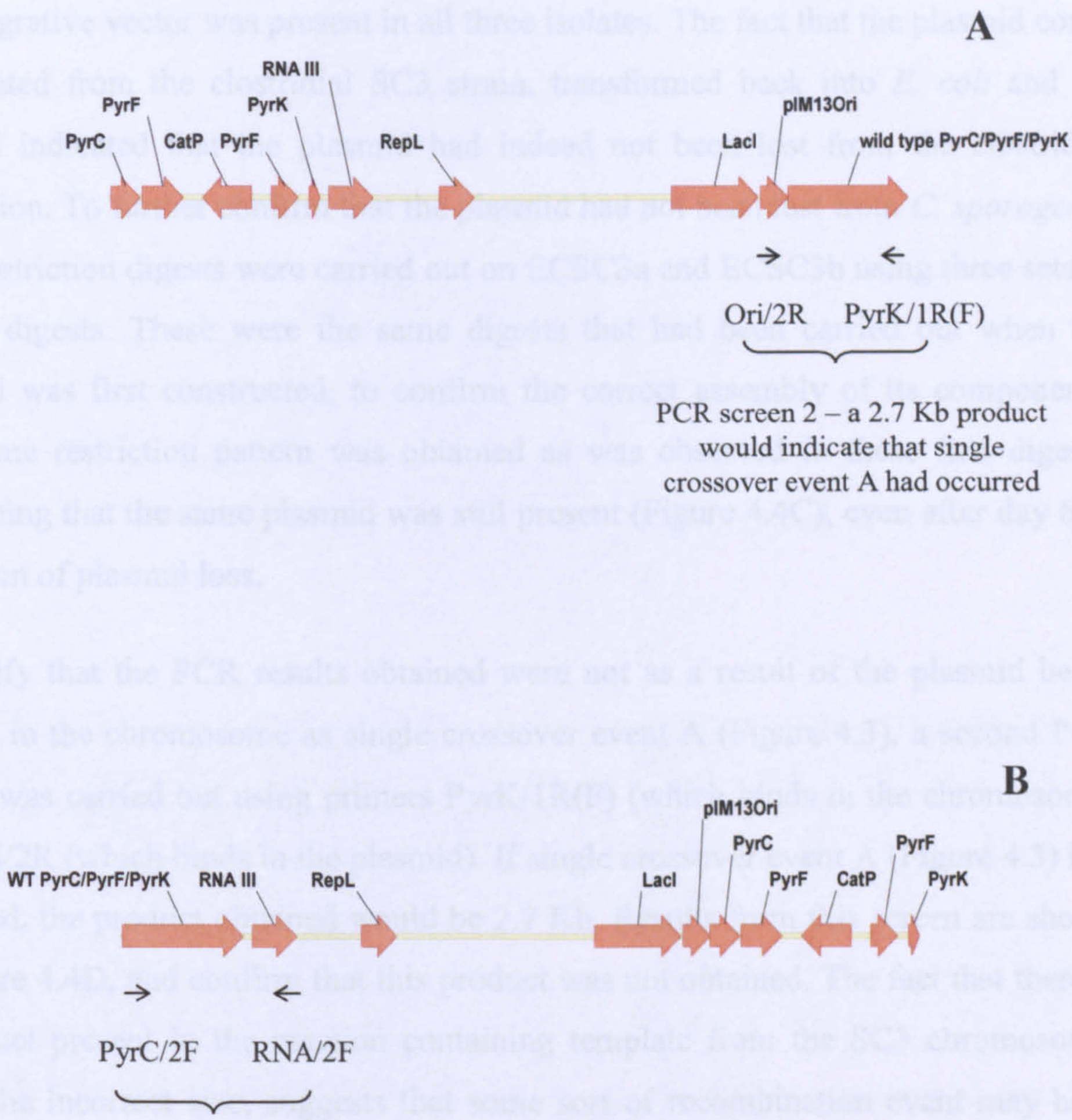


Figure 4.3 - Representation of the chromosomal arrangements which may arise from the two possible single crossover events (A and B). PCR screen 2 using primers Ori/2R (binds in the plasmid) and PyrK/1R(F) (binds in chromosome) would give a 2.7 Kb product if single crossover event A had occurred. PCR screen 3 using primers PyrC/2F (binds in the chromosome) and RNA/2F (binds in plasmid) would give a 2.8 Kb product if single crossover event B had occurred.

The expected 2.5 Kb products were observed in all three reactions confirming that the integrative vector was present in all three isolates. The fact that the plasmid could be isolated from the clostridial SC3 strain, transformed back into *E. coli* and re-isolated indicated that the plasmid had indeed not been lost from the clostridial population. To further confirm that the plasmid had not been lost from *C. sporogenes* SC3, restriction digests were carried out on ECSC3a and ECSC3b using three sets of double digests. These were the same digests that had been carried out when the plasmid was first constructed, to confirm the correct assembly of its components. The same restriction pattern was obtained as was observed in these first digests, confirming that the same plasmid was still present (Figure 4.4C), even after day 6 of induction of plasmid loss.

To verify that the PCR results obtained were not as a result of the plasmid being present in the chromosome as single crossover event A (Figure 4.3), a second PCR screen was carried out using primers PyrK/1R(F) (which binds in the chromosome) and Ori/2R (which binds in the plasmid). If single crossover event A (Figure 4.3) had occurred, the product obtained would be 2.7 Kb. Results from this screen are shown in Figure 4.4D, and confirm that this product was not obtained. The fact that there is a product present in the reaction containing template from the SC3 chromosome, albeit the incorrect size, suggests that some sort of recombination event may have occurred, which does not correspond to either of those outlined in Figure 4.3.

A third PCR screen was then carried out to confirm that the plasmid was not present as single crossover event B (Figure 4.3). Primers PyrC/2F (which binds in the chromosome) and RNA/2F (which binds in the plasmid) were used. If event B (Figure 4.3) had occurred, a 2.8 Kb product would be obtained. Results from this screen are shown in Figure 4.4E, and again, confirm that this product was not obtained.

It is impossible to ascertain precisely the nature of the rearrangements which have occurred. The only certainties are that the integration had not occurred as intended, the plasmid had not been lost and that *pyrF* knockout mutants were not obtained.

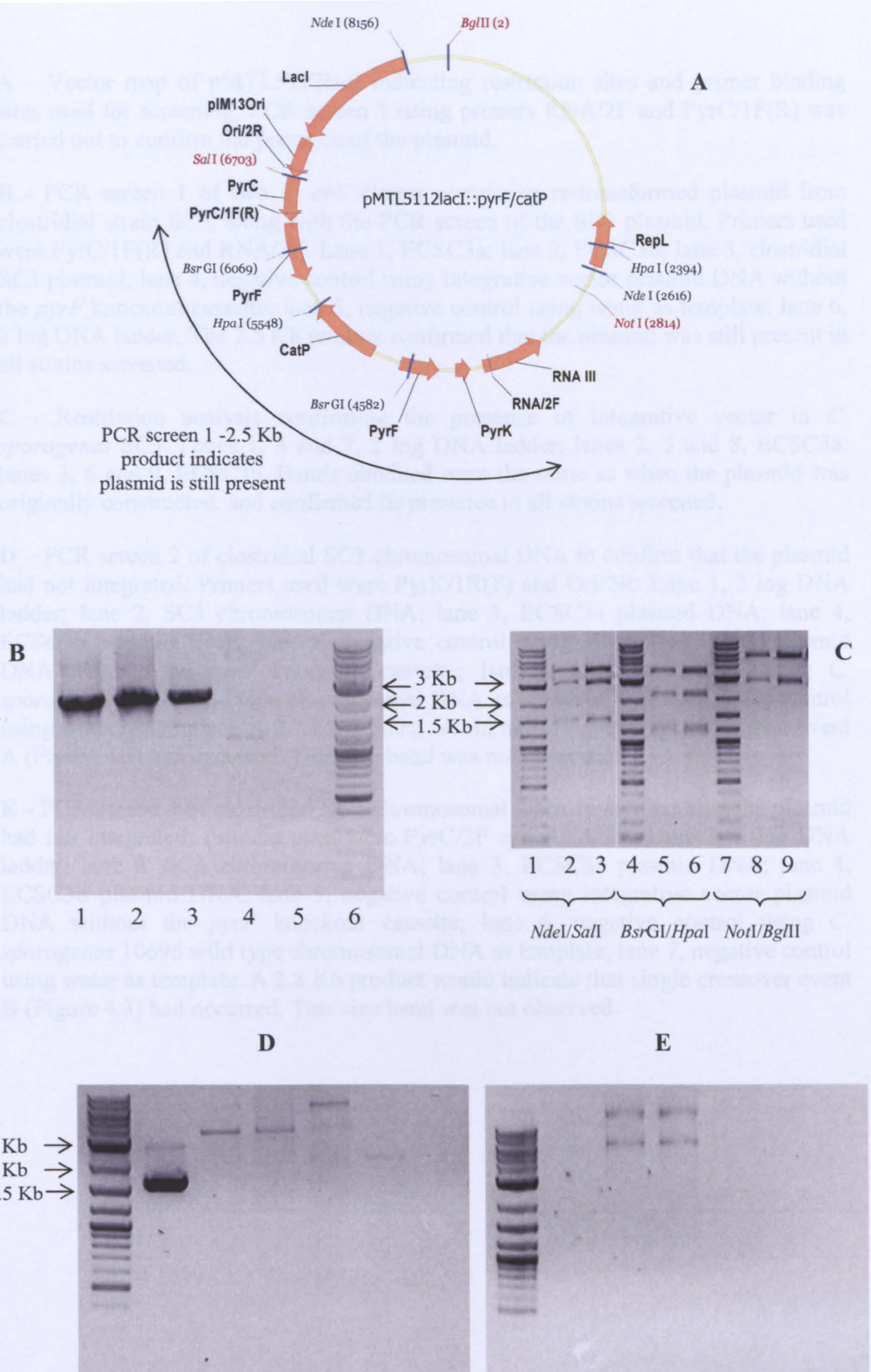


Figure 4.4 - Plasmid map and screening results on SC3 DNA

A – Vector map of pMTL5112lacI indicating restriction sites and primer binding sites used for screening. PCR screen 1 using primers RNA/2F and PyrC/1F(R) was carried out to confirm the presence of the plasmid.

B - PCR screen 1 of two *E. coli* clones containing re-transformed plasmid from clostridial strain SC3, along with the PCR screen of the SC3 plasmid. Primers used were PyrC/1F(R) and RNA/2F. Lane 1, ECSC3a; lane 2, ECSC3b; lane 3, clostridial SC3 plasmid; lane 4, negative control using integrative vector plasmid DNA without the *pyrF* knockout cassette; lane 5, negative control using water as template; lane 6, 2 log DNA ladder. The 2.5 Kb product confirmed that the plasmid was still present in all strains screened.

C - Restriction analysis confirming the presence of integrative vector in *C. sporogenes* SC3. Lanes 1, 4 and 7, 2 log DNA ladder; lanes 2, 5 and 8, ECSC3a; lanes 3, 6 and 9, ECSC3b. Bands obtained were the same as when the plasmid was originally constructed, and confirmed its presence in all strains screened.

D – PCR screen 2 of clostridial SC3 chromosomal DNA to confirm that the plasmid had not integrated. Primers used were PyrK/1R(F) and Ori/2R. Lane 1, 2 log DNA ladder; lane 2, SC3 chromosomal DNA; lane 3, ECSC3a plasmid DNA; lane 4, ECSC3b plasmid DNA; lane 5, negative control using integrative vector plasmid DNA without the *pyrF* knockout cassette; lane 6, negative control using *C. sporogenes* 10696 wild type chromosomal DNA as template; lane 7, negative control using water as template. A 2.7 Kb product would indicate that single crossover event A (Figure 4.3) had occurred. This size band was not observed.

E - PCR screen 3 of clostridial SC3 chromosomal DNA to confirm that the plasmid had not integrated. Primers used were PyrC/2F and RNA/2F. Lane 1, 2 log DNA ladder; lane 2, SC3 chromosomal DNA; lane 3, ECSC3a plasmid DNA; lane 4, ECSC3b plasmid DNA; lane 5, negative control using integrative vector plasmid DNA without the *pyrF* knockout cassette; lane 6, negative control using *C. sporogenes* 10696 wild type chromosomal DNA as template; lane 7, negative control using water as template. A 2.8 Kb product would indicate that single crossover event B (Figure 4.3) had occurred. This size band was not observed.

Final confirmation that the plasmid was still unstable and un-integrated was carried out using segregational stability tests, as described in method 2.9. Results indicated that the maximum plasmid loss was only 65.31% (Table 4.4). If the plasmid had integrated, it would be expected that loss would be much greater than this. These results therefore confirm that the plasmid was not lost from the population, suggesting that it had not integrated into the host chromosome and that this method for attempted gene knockout had been unsuccessful.

IPTG	Percentage Plasmid Loss			
	Day 3	Day 4	Day 5	Day 6
+	2.00	40.91	65.31	60.94
-	33.33	8.82	21.05	44.44

Table 4.4 - Segregational stability of integrative vector pMTL5112lacI containing *pyrF* knockout cassette in *C. sporogenes* SC3, expressed as percentage plasmid loss.

4.2.2 Discussion

Along with Philippe Soucaille, who has successfully used a rolling circle plasmid to generate knockout mutants in *C. acetobutylicum* (Saint-Prix *et al.*, 2006) two other groups have recently published attempts to exploit the properties of unstable plasmids in order to generate knockout mutants in clostridia. Harris *et al.* (2002) used an unstable vector to introduce a mutation into the *spo0A* gene of *C. acetobutylicum* by double crossover recombination. The resulting mutant was sporulation deficient and showed reduced solvent production. Although the mutant was obtained, sequencing demonstrated that the second crossover event did not take place as predicted, due to a crossover event taking place between two 10 base pair homologous sequences in the 3' end of the *repL* structural gene. This confirms the unpredictable nature of this method.

The second group to publish an attempt at this method ironically obtained their mutant accidentally (O'Connor *et al.* 2006). This group introduced an unstable plasmid containing target gene sequences into *C. difficile*, and found subsequently that the plasmid could not be recovered. Further investigations confirmed that the plasmid had integrated into the host chromosome, producing two unstable mutants of the *rgaR* and *rgbR* genes. These mutants were, however, produced by a single crossover event, rather than the more desirable double crossover, and PCR screens showed that there was a small amount of wild type gene product present.

Both of these groups demonstrated that, although unstable plasmids can be used to generate mutants in clostridia, the unstable, and therefore unpredictable, replication mechanism of these plasmids means that the process does not always proceed as predicted. This may be the reason for the failure to obtain mutants in this study. Indeed, Gruss & Ehrlich (1989) emphasised the infidelity of the rolling circle mechanism of replication and noted that almost every step in the mechanism can digress from its usual function, therefore affecting the rearrangements that occur. It is impossible to control this mechanism of replication, due to the fact that the abundance of single stranded DNA stimulates intermolecular homologous recombination, which does not necessarily occur in the desired manner. Another contributory factor could be that the single stranded DNA produced might have

formed secondary structures with itself, thus preventing it from effectively integrating into the host chromosome. The *catP* sequence possesses significant secondary structure which may form a hairpin loop. This could have been acting as a negative origin, hence rendering the system inefficient. Negative origins contain numerous imperfect inverted repeats, forming hairpin structures (Ehrlich *et al.*, 1991). The mechanism of plasmid pMTL5112lacI is based on the absence of the negative origin preventing the conversion of single stranded to double stranded DNA. The *catP* sequence acting as a negative origin would mean that the single stranded DNA could be converted back, preventing the accumulation of single stranded DNA and reducing the occurrence of recombination. This would explain why the plasmid was lost at a much less efficient rate once the knockout cassette was introduced; the restoration of a single stranded origin-like structure allowed replication to proceed as normal, and stabilised the plasmid.

Another possibility could be the unusually high expression of *repL* in plasmid pMTL5112lacI. Gruss & Ehrlich (1989) proposed that a factor which prevented runaway replication was that the Rep protein in *E. coli* plasmid RK6 was less active at high concentrations. The fact that *repL* is under control of the strong *facII* promoter in pMTL5112lacI means that *repL* is over-expressed and therefore may be less active, causing a decrease in replication and therefore a decrease in single stranded DNA production.

A problem also obviously lies with the fact that difficulties were encountered when transferring the plasmid into its intended host. Several attempts were made before it was successfully transformed into *B. subtilis* and it was never successfully conjugated into *C. sporogenes* 10696. This is not a practical situation to have if such a plasmid is to be used as a routine knockout tool. The only way the plasmid could be conjugated into *C. sporogenes* was to use the cured strain as a host. This is an uncharacterised mutant, which obviously has some differences to the wild type. It is therefore impossible to determine why the conjugation was successful in this strain, and also whether an unknown mutation could be responsible for the integration being unsuccessful. The fact that this strain is more susceptible to conjugation than the wild type would also indicate that plasmids are more stable in this strain. This is not ideal when using an integrative plasmid system that relies on instability.

Moreover, the high instability of the plasmid may mean that it was too unstable to replicate successfully in the wild type and explain why conjugation was not successful.

Clearly there are many unanswered questions surrounding the use of integrative vector pMTL5112lacI. The mechanisms underlying the functioning of the plasmid remain to be explained. Future investigations could be carried out to aid the characterisation of the mechanism of action of this plasmid. Southern blots would confirm that single stranded DNA is indeed being produced. The efficiency of the induction of *repL* could be determined using reverse transcriptase PCR on RNA from induced and uninduced cultures. The problem of the *catP* gene possibly acting as the negative origin could also be addressed. Gruss & Ehrlich (1989) demonstrated that rifampin addition inhibited negative origin function. All of the negative origins tested in their study were dependent on RNA polymerase action, which is inhibited by rifampin. Therefore its addition to the culture rendered the negative origin inactive, allowing the accumulation of single stranded DNA. This could be attempted to prevent the possibility of *catP* acting as a negative origin.

These further investigations may contribute to the understanding of why this method was unsuccessful, and suggest future directions for the pursuit of an integrative vector system which can be used for the routine inactivation of genes in clostridia.

4.3 Gene modulation in *C. sporogenes* using antisense RNA technology

An alternative gene modulation strategy utilises 'antisense' oligonucleotides, which can block the expression of a single gene, or multiple genes of an operon (Kernodle *et al.*, 1997; Schlingensiepen & Schlingensiepen, 1997). Antisense oligonucleotides (oligos) are short stretches of nucleotides which are complementary to a region of mRNA in the gene of interest (Jen & Gewirtz, 2000). These oligos bind specifically to the target mRNA, interfering with translation into protein. It is thought that this leads to the destruction of the mRNA (Jen & Gewirtz, 2000). Antisense RNA can be expressed from a plasmid vector which has been introduced into the bacterial cell of interest (Jen & Gewirtz, 2000).

Bacteria naturally produce antisense RNA to regulate their own gene expression (Schlingensiepen & Schlingensiepen, 1997) and one of the first groups to exploit this fact was Kernodle *et al.* (1997). They targeted the α toxin *hla* gene of *S. aureus* with antisense RNA molecules and observed a 16-fold decrease in the amount of α toxin detected in culture supernatants. There was also a corresponding decrease in the bacterium's lethal activity in mice. Antisense RNA has also been produced against the *hla* gene downstream of a tetracycline promoter, resulting in an inducible antisense system (Ji *et al.*, 1999).

Several mechanisms of action of antisense RNA have been proposed (Schlingensiepen & Schlingensiepen, 1997):

1. The antisense RNA molecules prevent elongation of the polypeptide chain by steric block of the ribosome on the RNA (Figure 4.5 A).
2. Initiation of translation is prevented by the antisense RNA molecule binding to the promoter region or near the initiation codon (Figure 4.5 B).
3. The cellular enzyme RNaseH recognises the duplex of the oligonucleotide and the mRNA, and cleaves the mRNA, leaving the oligonucleotide intact (Figure 4.5 C).

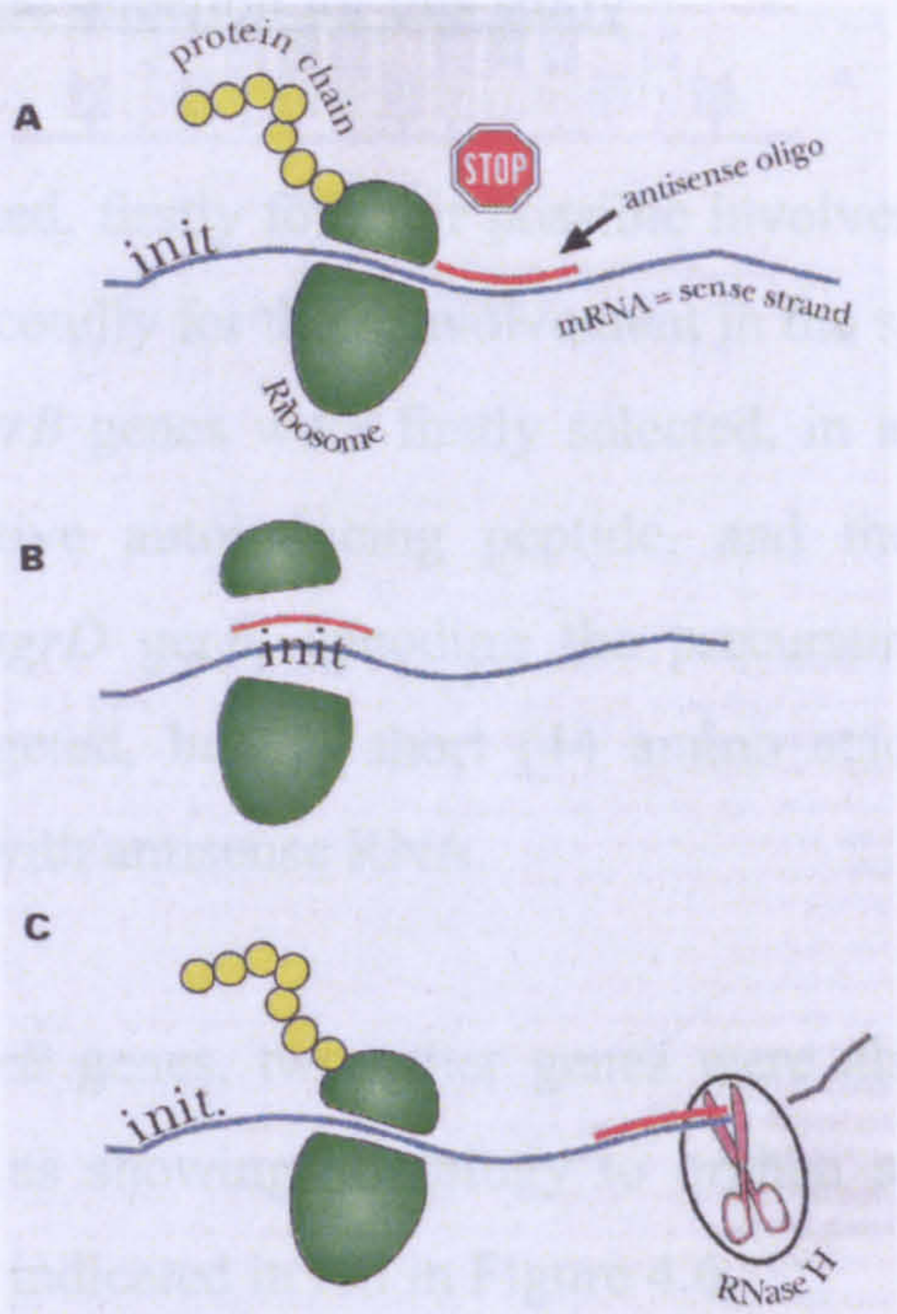


Figure 4.5 – Proposed mechanisms of action of antisense RNA (Schlingensiepen & Schlingensiepen, 1997). The antisense RNA may inhibit translation by several mechanisms. A, antisense RNA prevents elongation of the polypeptide chain by blocking movement of the ribosome along the mRNA; B, the antisense RNA blocks binding of the ribosome to the mRNA near to the initiation codon; C, a cellular RNaseH recognises the duplex of antisense RNA and sense mRNA and cleaves the mRNA.

Since the introduction of antisense technology to modulate staphylococcal virulence gene expression in 1997, various groups have applied this technology in clostridia. Tummala *et al.* (2003b) produced antisense RNA against both acetoacetate decarboxylase and co-enzyme A-transferase in *C. acetobutylicum* and demonstrated that co-enzyme A-transferase is the rate-limiting enzyme in acetone formation. The same group went on to use antisense RNA, combined with gene over-expression, to increase the butanol/acetone ratio in *C. acetobutylicum* (Tummala *et al.*, 2003a). Marvaud *et al.* (1998b) demonstrated partial inhibition of the *botR/A* regulatory gene in *C. botulinum*, which resulted in a reduction in neurotoxin production.

4.3.1 Antisense target gene selection for this study

Target genes were selected, firstly for their possible involvement in the *agr* quorum sensing response, and secondly for their involvement in the sporulation cascade in *C. sporogenes*. The two *agrB* genes were firstly selected, in an attempt to inhibit the processing of the putative autoinducing peptide, and therefore inhibit the *agr* response. Ideally, the *agrD* gene, encoding the precursor of the AIP signalling molecule should be targeted, but its short (44 amino acid) sequence rendered it unsuitable for targeting with antisense RNA.

Along with the two *agrB* genes, two other genes were also selected. These were identified in Chapter 3 as showing homology to orphan sensor kinase genes. All target genes selected are indicated in red in Figure 4.6.

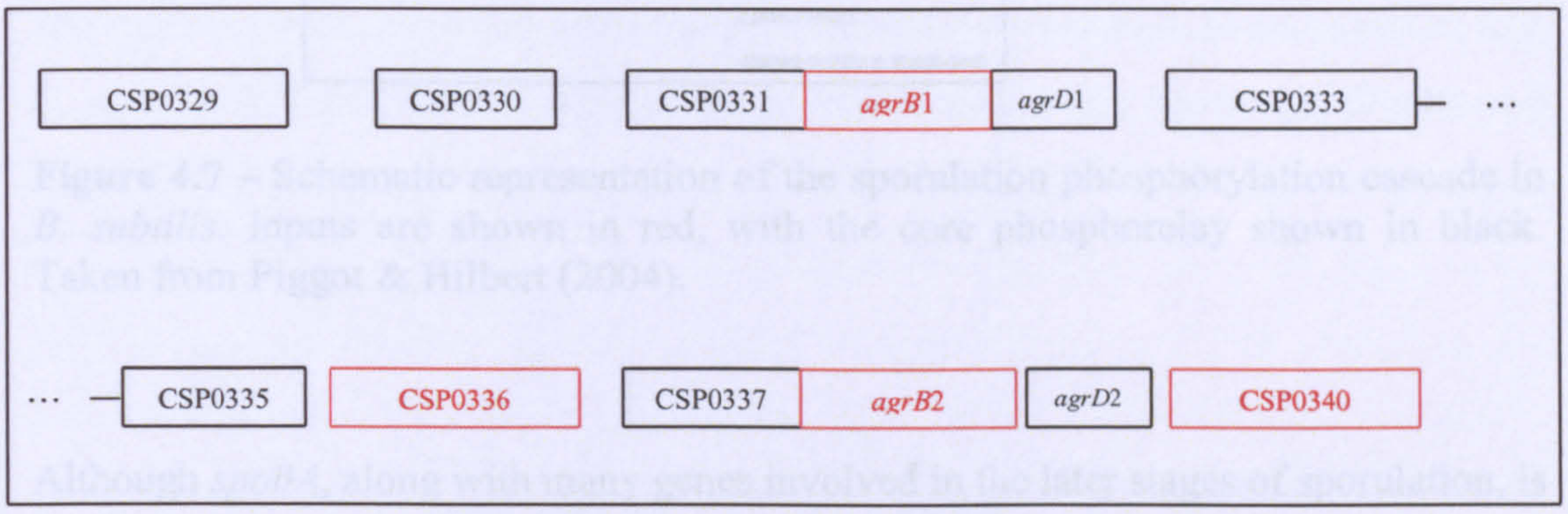


Figure 4.6 – Schematic representation of the *agr* regions in the *C. sporogenes* 10696 genome (not to scale). Genes targeted for antisense modulation are shown in red.

Extensive research has been carried out on the initiation of sporulation in *B. subtilis* and to a lesser extent in the clostridia. Research shows that, in *B. subtilis*, external sporulation signals are received by five orphan histidine kinase proteins (KinA – KinE), usually under conditions of nutrient deprivation (Hoch, 1993). These proteins in turn trigger a phosphorylation cascade involving the transfer of a phosphate group from Spo0F to SpoB, and finally to Spo0A, the master regulator of sporulation (Piggot & Hilbert, 2004). Many other sigma factors are involved in the process, but ultimately Spo0A in its phosphorylated form allows sporulation to proceed. This process is depicted in Figure 4.7.

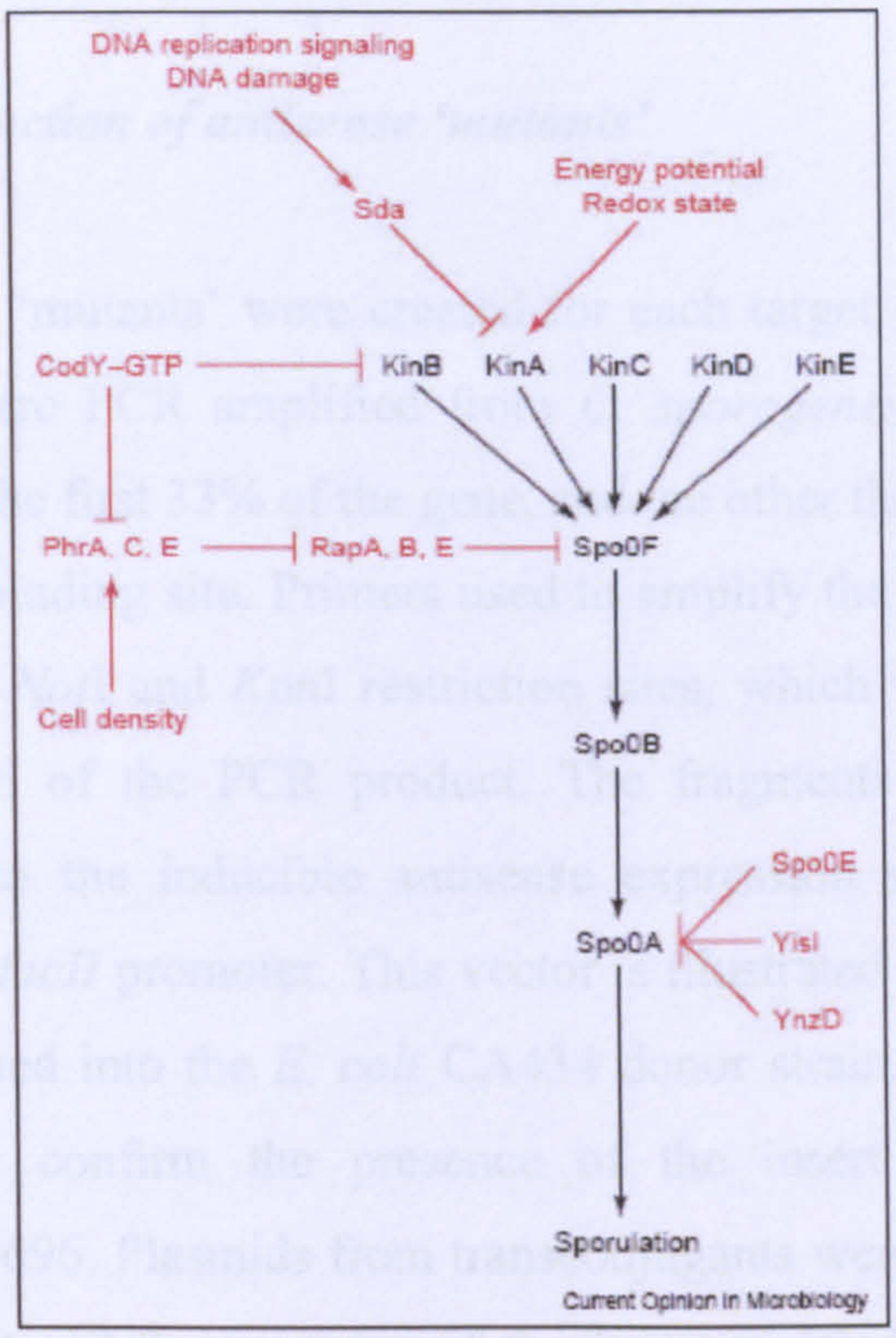


Figure 4.7 – Schematic representation of the sporulation phosphorylation cascade in *B. subtilis*. Inputs are shown in red, with the core phosphorelay shown in black. Taken from Piggot & Hilbert (2004).

Although *spo0A*, along with many genes involved in the later stages of sporulation, is also found in the sequenced species of clostridia, most of the genes involved in the earlier stages of sporulation have not been identified (Parades *et al.*, 2005). This includes the five *kin* genes, along with *spo0F* and *spo0B*. A *spo0B* homologue has been identified in *C. tetani* but it is not known whether it is functional (Stephenson & Lewis, 2005). The fact that Spo0A has been shown to be involved in the initiation of sporulation in various species of clostridia (Huang & Sarker, 2006; Ravagnani *et al.*, 2000) would indicate that the initiation of sporulation in clostridia occurs through a different mechanism. This may occur via a novel phosphorelay system, but the most likely hypothesis is that Spo0A is phosphorylated directly by a sensor kinase (Parades *et al.*, 2005). Indeed, in the *C. botulinum* genome, five orphan kinases have been identified, which may be involved in the phosphorylation of Spo0A (Worner *et al.*, 2006). Two of these orphan kinase genes (0336 and 0340) are found within the *agr* regions of *C. botulinum* and *C. sporogenes* and were therefore selected as targets for antisense modulation.

4.3.2 Results

4.3.2.1 Construction of antisense 'mutants'

Two antisense 'mutants' were created for each target gene. Two fragments of each target gene were PCR amplified from *C. sporogenes* 10696 genomic DNA. One encompassed the first 33% of the gene, and the other the first 66%, and both included the ribosome binding site. Primers used to amplify the fragments are shown in 2.6.2 and contained *Not*I and *Kpn*I restriction sites, which were incorporated into the 5' and 3' termini of the PCR product. The fragments were cloned in the reverse orientation into the inducible antisense expression vector pMTL9361 under the control of the *facII* promoter. This vector is illustrated in Figure 4.8. The vector was then transformed into the *E. coli* CA434 donor strain. Restriction digests and PCR were used to confirm the presence of the insert before conjugating into *C. sporogenes* 10696. Plasmids from transconjugants were isolated, re-transformed into *E. coli* TOP10 and the presence of the insert verified using restriction digests and PCR. The resulting 'mutants' were named as illustrated in Table 4.5.

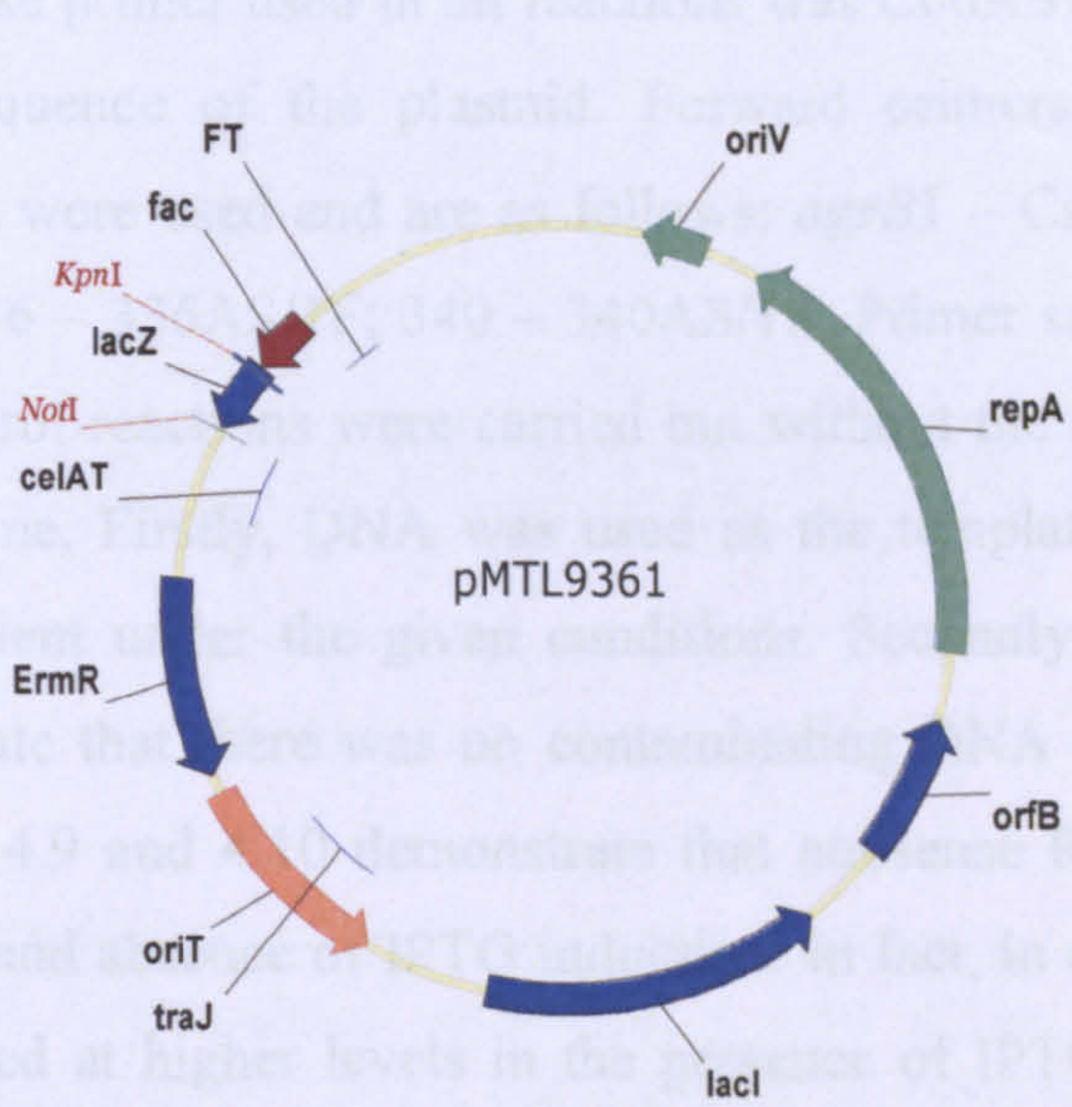


Figure 4.8 – Plasmid map of inducible antisense delivery vector pMTL9361. The antisense fragments are inserted between the *Not*I and *Kpn*I restriction sites. Expression of the antisense RNA fragment is driven from the *fac* promoter on addition of IPTG and de-repression of LacI from the *lac* operator. The *celAT* terminator prevents transcriptional read-through. An erythromycin resistance cassette is located on the plasmid backbone for positive selection of clones carrying the antisense vector.

Target Gene	Name of 'Mutant'	
	Small antisense fragment	Large antisense fragment
*agrB*1	*agrB*1 AS1	*agrB*1 AS2
*agrB*2	*agrB*2 AS1	*agrB*2 AS2
CSP0336	336 AS1	336 AS2
CSP0340	340 AS1	340 AS2

Table 4.5 – Nomenclature of eight antisense 'mutants' created in *C. sporogenes*

The inducible production of antisense RNA in all four 'mutants' was then confirmed using RT-PCR. 5 ml broth cultures of all 'mutants' were grown in the presence and absence of 1 mM IPTG for 6 hours. RNA was extracted from each strain according to 2.8.1 and 20 ng RNA from each 'mutant' was subjected to RT-PCR as described in 2.8.4. The reverse primer used in all reactions was CodAS1/F, which binds in the *facII* promoter sequence of the plasmid. Forward primers which bound in the antisense fragment were used and are as follows: *agrB*1 – CsagrB1AS/1F; *agrB*2 – CsagrB2AS/1F; 336 – 336AS/1F; 340 – 340AS/1F. Primer sequences can be found in 2.6.2. Two control reactions were carried out without the addition of the reverse transcriptase enzyme. Firstly, DNA was used as the template to illustrate that the reaction was efficient under the given conditions. Secondly, RNA was used as a template to illustrate that there was no contaminating DNA in the sample. Results shown in Figures 4.9 and 4.10 demonstrate that antisense RNA was expressed in both the presence and absence of IPTG induction. In fact, in all 'mutants', antisense RNA was expressed at higher levels in the presence of IPTG than in the absence. This demonstrates the inducible nature of the pMTL9361 expression vector. However, these results also demonstrate that in all cases, except 340 AS2, expression of antisense RNA was still able to occur even in the absence of IPTG, indicating that 'leaky' transcription was occurring from the *fac* promoter.

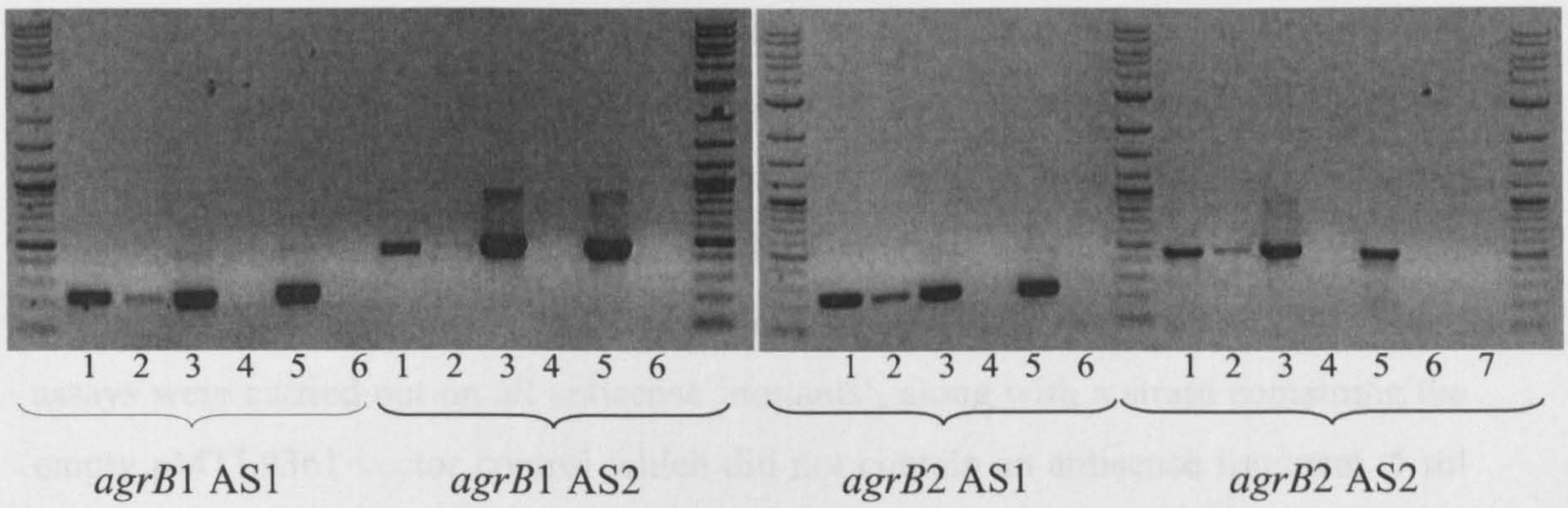


Figure 4.9 – Confirmation of antisense RNA expression in *agrB*1 (**A**) and *agrB*2 (**B**) 'mutants'. 1, RNA from culture containing 1 mM IPTG; 2, RNA from culture without IPTG; 3, DNA from + IPTG culture, with no reverse transcriptase; 4, RNA from + IPTG culture with no reverse transcriptase; 5, DNA from - IPTG culture, with no reverse transcriptase; 6, RNA from - IPTG culture with no reverse transcriptase; 7, negative control using water as template. All 'mutants' analysed demonstrated higher levels of antisense RNA expression under IPTG induction.

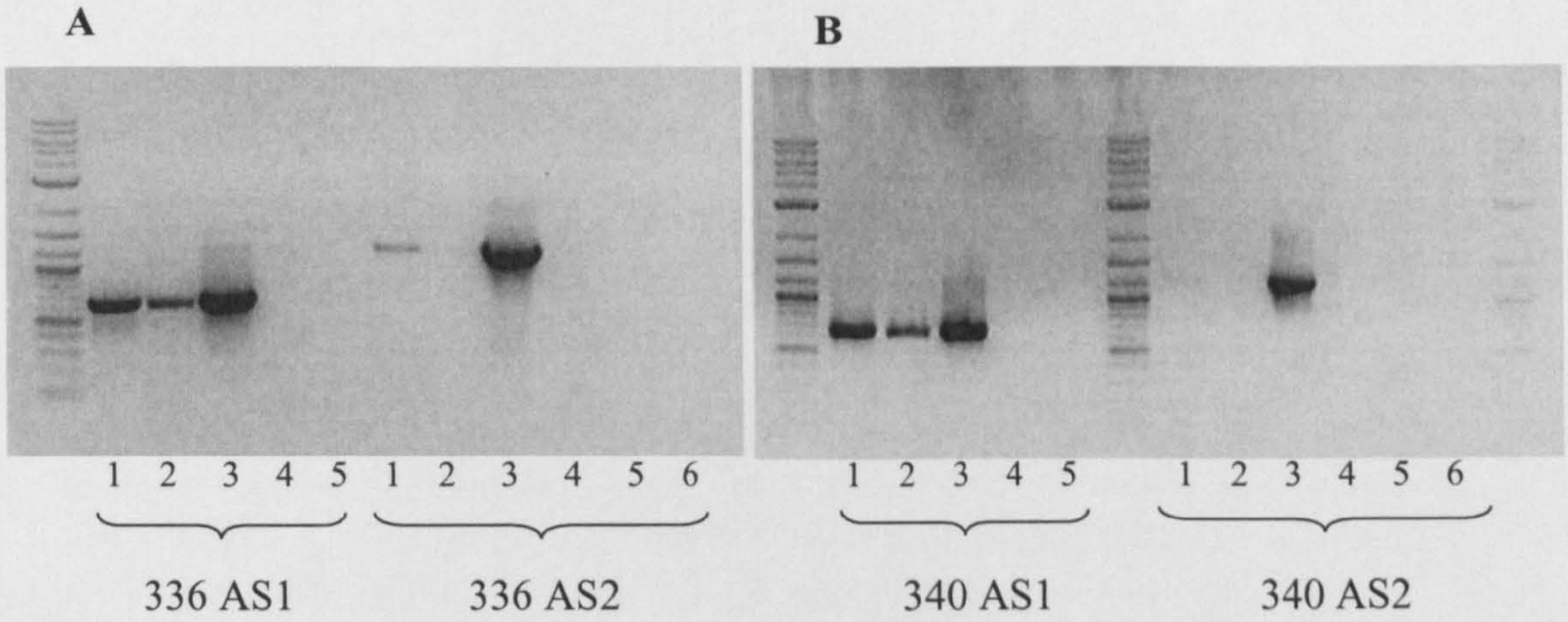
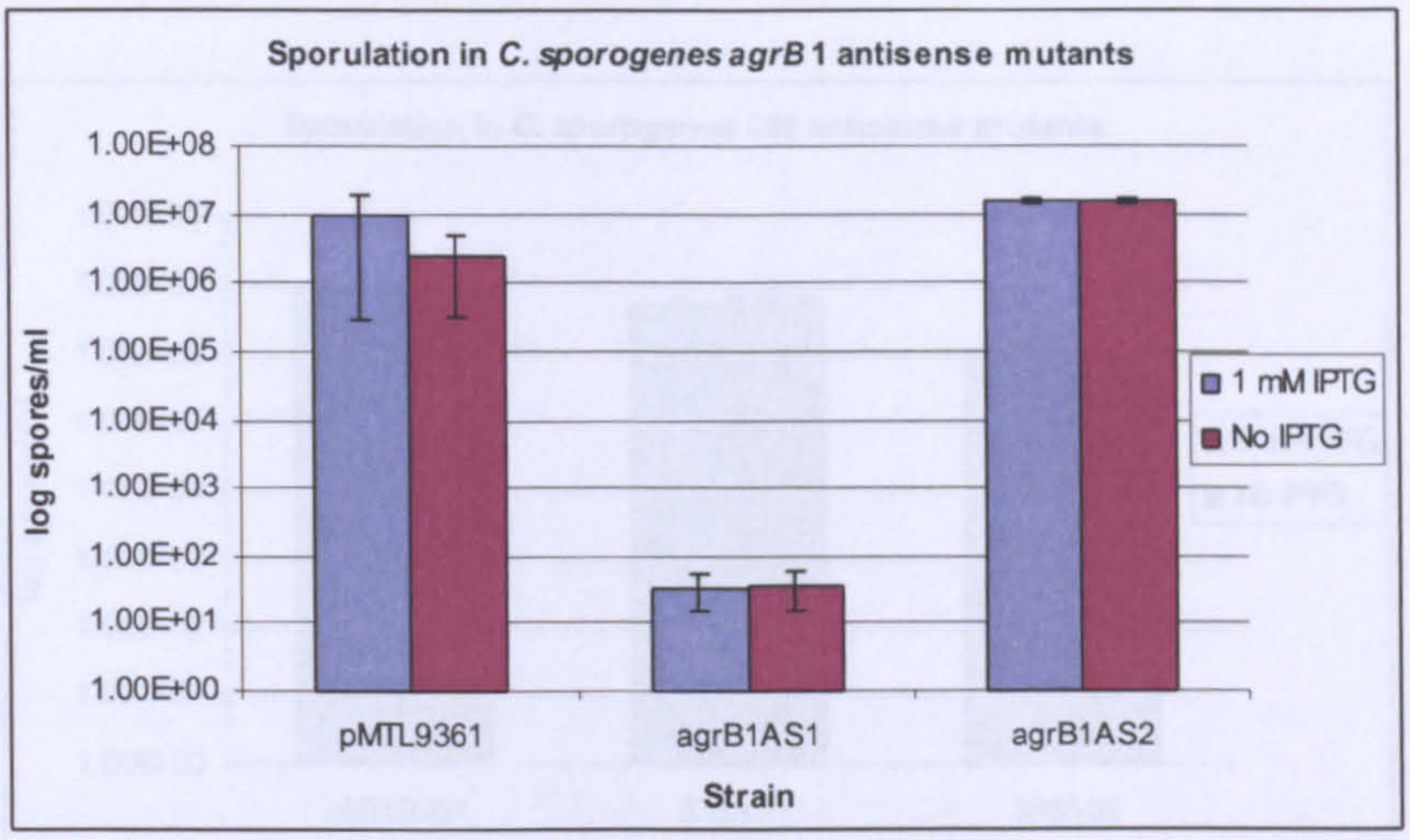


Figure 4.10 - Confirmation of antisense RNA expression in 336 (**A**) and 340 (**B**) 'mutants'. 1, RNA from culture containing 1 mM IPTG; 2, RNA from culture without IPTG; 3, DNA from + IPTG culture, with no reverse transcriptase; 4, RNA from + IPTG culture with no reverse transcriptase; 5, RNA from - IPTG culture, with no reverse transcriptase; 6, negative control using water as template. All 'mutants' analysed demonstrated higher levels of antisense RNA expression under IPTG induction.

4.3.2.2 Effect of antisense production on sporulation

Potential phenotypes which may be affected by *agrB* modulation were considered and sporulation selected as a possible candidate. Sporulation was also the obvious phenotype to assay in the 336 and 340 'mutants' due to their possible role in the phosphorylation cascade in the initiation of sporulation. This is also a relatively easy phenotype to assay, using the method of heat inactivation of vegetative cells. Spore assays were carried out on all antisense 'mutants', along with a strain containing the empty pMTL9361 vector control which did not contain an antisense fragment. 5 ml broth cultures were grown for 72 hours and spore assays carried out in triplicate according to 2.12. Growth was monitored throughout the incubation period before heat treatment to ensure that all strains grew at a similar rate. Spore assay results are shown in Figures 4.11 and 4.12.



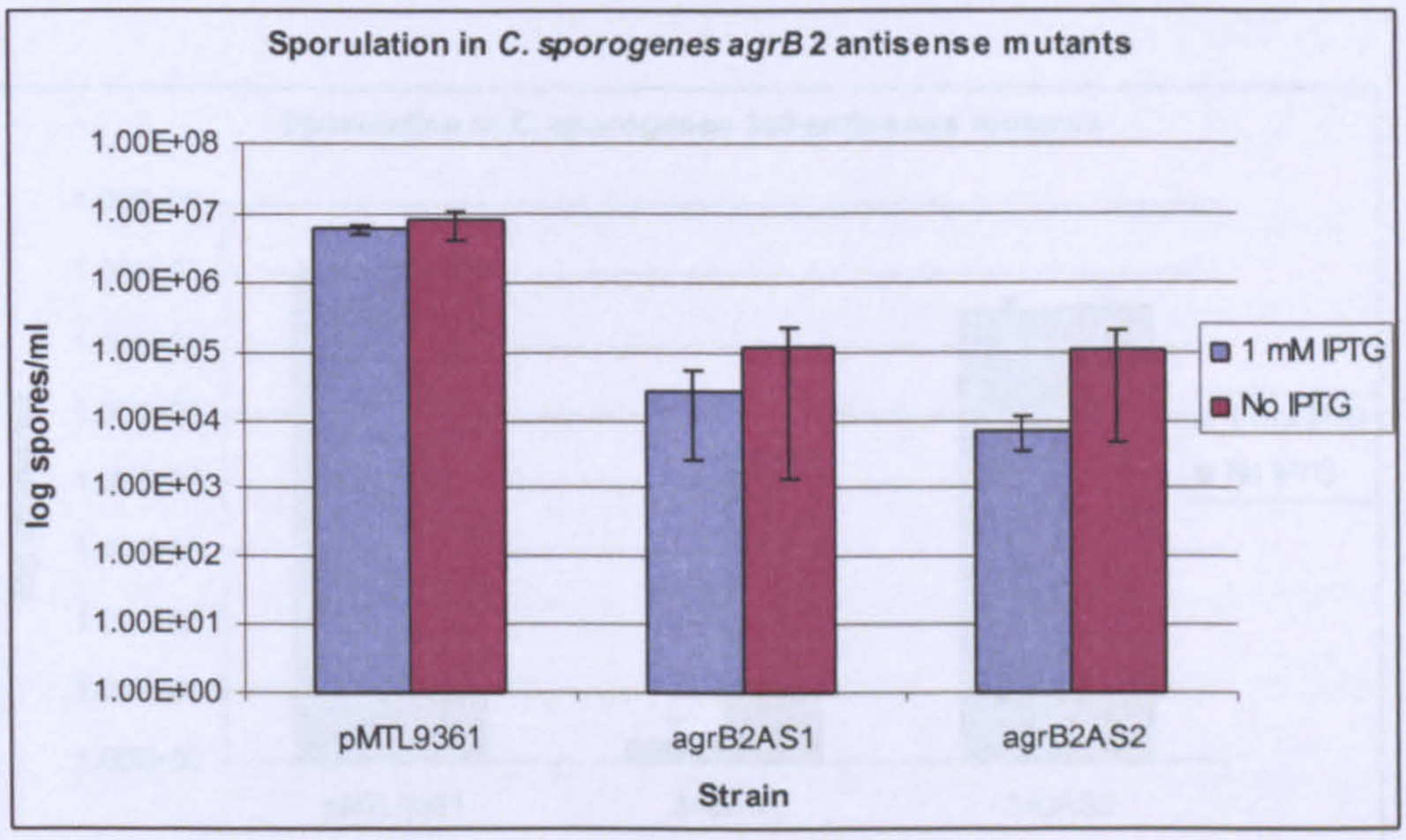
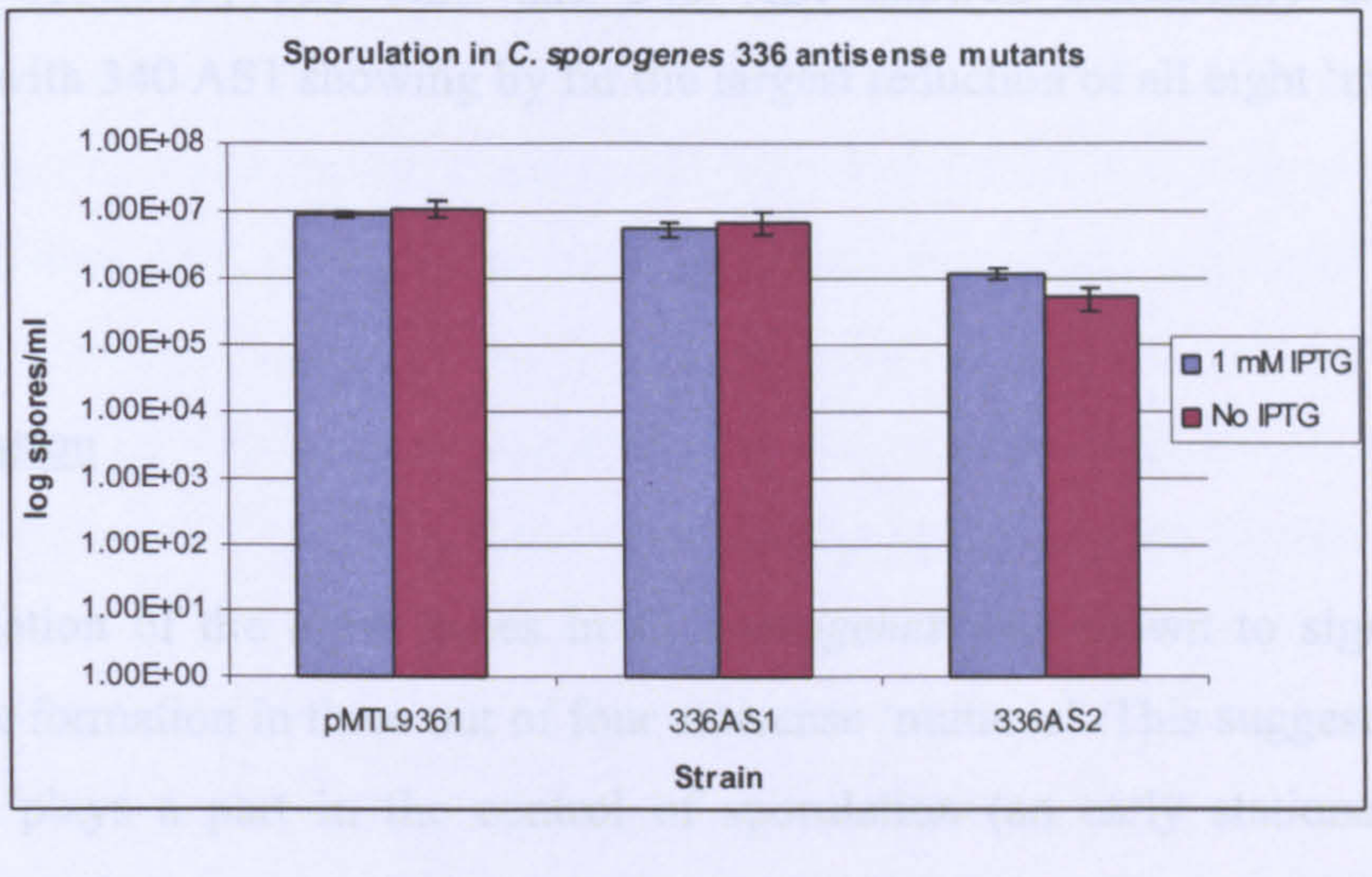


Figure 4.11 – Sporulation in *C. sporogenes agrB* 'mutants' in the presence and absence of IPTG. In the *agrB*1 AS1 'mutant', the gene was down-regulated even in the absence of IPTG induction, resulting in an equivalent reduction in sporulation with or without IPTG. The *agrB*1 AS2 fragment appears to be ineffective in down-regulating the *agrB*1 gene. In both *agrB*2 'mutants' down-regulation of sporulation was more apparent in the presence of IPTG induction of antisense RNA.



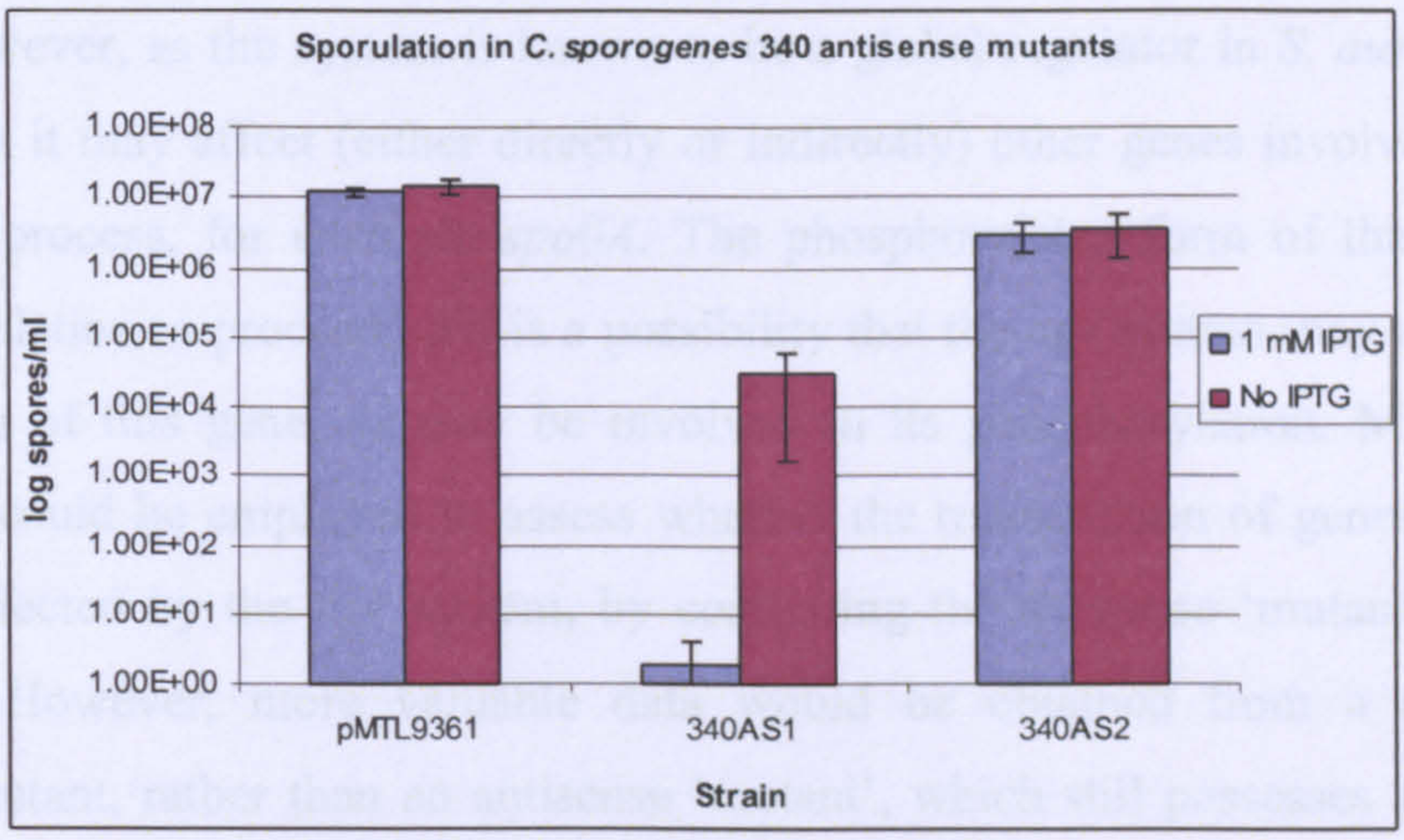


Figure 4.12 – Sporulation in *C. sporogenes* 336 and 340 kinase 'mutants' in the presence and absence of IPTG. The 336 AS1 fragment appears to be inefficient, as does the 340 AS2 fragment. A small reduction in sporulation in the 336 AS2 'mutant' may indicate an antisense fragment with poor efficiency, or that this gene plays little part in sporulation. A huge reduction in sporulation was observed under IPTG induction of the 340 AS1 fragment.

Figure 4.11 illustrates that there was a significant reduction in sporulation in the *agrB*1 AS1 'mutant' in the presence of IPTG, compared to the pMTL9361 control. Sporulation in both *agrB*2 'mutants' was also reduced by at least two orders of magnitude. All three of these results were shown to be statistically significant at a confidence interval of 95% using the pair wise Student's T test. However, the *agrB*1 AS2 'mutant' showed no reduction in sporulation.

Figure 4.12 demonstrates that 336 AS1 and 340 AS2 showed no reduction in sporulation. However, 336 AS2 and 340 AS1 showed statistically significant reductions, with 340 AS1 showing by far the largest reduction of all eight 'mutants'.

4.3.3 Discussion

Down-regulation of the *agrB* genes in *C. sporogenes* was shown to significantly inhibit spore formation in three out of four antisense 'mutants'. This suggests that the *agr* system plays a part in the control of sporulation (an early stationary phase process). From these results it cannot be determined how the *agr* system affects this process; however, as the system is known to be a global regulator in *S. aureus*, it is possible that it may affect (either directly or indirectly) other genes involved in the sporulation process, for example *spo0A*. The phosphorylated form of this protein allows sporulation to proceed so it is a possibility that the *agr* system may affect the transcription of this gene, or may be involved in its phosphorylation. Microarray technology could be employed to assess whether the transcription of genes such as *spo0A* is affected by the *agr* system, by comparing the antisense 'mutants' to the wild type. However, more valuable data would be obtained from a complete knockout mutant, rather than an antisense 'mutant', which still possesses some *agr* activity.

The only significant reduction in sporulation in the kinase 'mutants' was observed in 340 AS1. This showed a reduction of over six orders of magnitude in the presence of IPTG induction, which suggests that, firstly, the 340 kinase is involved in the control of sporulation, and that, secondly, the larger antisense fragment used against this target gene was not effective. It is therefore a possibility that gene 340 may be involved (possibly with other orphan kinase genes) in the initiation of sporulation via the phosphorylation cascade. It is not possible to say from these results whether the 336 gene is involved in sporulation or not. Its down-regulation did not appear to significantly affect sporulation, but this could be due to the selection of ineffective antisense RNA fragments. The only way to address this would be to completely knock out this gene.

The recent sequencing of the *C. botulinum* genome has revealed the presence of three other orphan kinase genes - CBO0780, CBO2762 and CBO1120 (Worner *et al.*, 2006), which can be assumed to also be present in *C. sporogenes*. It is therefore possible that one or more of these could also be involved in the initiation of sporulation, along with 0336 and/or 0340. An alternative hypothesis could be that a single kinase is involved in the phosphorylation of Spo0A, rather than the involvement of a cascade, as identified in *B. subtilis*. Indeed, it has been demonstrated that in *B. subtilis*, KinA can directly phosphorylate Spo0A, albeit at low levels (Trach & Hoch, 1993). Therefore the possibility of a single kinase directly phosphorylating Spo0A in order to trigger sporulation cannot be discounted. Clearly more work must be carried out on this process, investigating all five orphan kinases, in order to determine exactly how sporulation is controlled in *C. sporogenes* and *C. botulinum*.

The different *agrB* 'mutants' showed inhibition of sporulation to varying degrees, indicating that some antisense RNA oligonucleotides selected were more effective than others. The antisense oligonucleotides selected covered 33% and 66% of the *agrB* gene, based on antisense RNA generation by Tummala *et al.* (2003b). However, without computational analysis of the antisense oligonucleotides selected, it is difficult to determine how effective the oligonucleotide will prove. Indeed, the complex secondary structure of mRNA molecules means that much of the molecule may be inaccessible to antisense RNA. Scherr & Rossi (1998) introduced a system to determine the most effective target sites by using RNaseH cleavage. This showed which sites were accessible to the enzyme and therefore which sites would be accessible to antisense RNA. Tummala *et al.* (2003b) also sought ways of increasing the efficiency of antisense RNA. They used computer programs and theoretical analyses to obtain important information on the structure of antisense RNA, which may affect its hybridisation with mRNA. They concluded that, as the (no. components)/(no. total nucleotides) ratio decreased, the number of free nucleotides decreased, meaning more nucleotides were intramolecularly bound, making it more difficult for RNases to degrade the antisense RNA. (Free nucleotides are nucleotides in an antisense RNA molecule that are not involved in intramolecular bonding, providing potential interaction sites with target mRNA. Components are structural features that contain regions of high complementarity with an antisense RNA molecule (Tummala *et al.*, 2003b)). In the future, the use of *in silico* analysis will

provide information on optimal antisense olignucleotides to use. In fact, Bo *et al.* (2006) reported a database (AOBase) which stores published oligonucleotide sequences and aids selection of efficient antisense oligonucleotides. However, computer programs were not available for use in this instance, so the fragments were selected purely on the basis of their covering the first 33% and first 66% of the target gene. It is therefore possible that the fragments selected were not ideal, and this would explain the differing effects that the two antisense fragments had on each target gene. An ideal solution would be to produce knockout mutants of the four target genes, which would give a much 'cleaner' picture of their true function. It must be remembered that the use of antisense technology does not result in the complete abolition of target protein synthesis, but merely down-regulates it. This can be an advantage when modulating genes which may be essential, but also hinders the interpretation of phenotypic assays.

Another interesting factor to note is the need for IPTG induction to produce the antisense RNA. The RT-PCR results indicated that in all 'mutants', antisense RNA was produced, even in the absence of IPTG induction. This is a result of 'leaky' expression from the *fac* promoter, which is not tightly controlled. This was confirmed in the spore assays by the fact that, when a reduction in sporulation was observed, this occurred in both the presence and absence of IPTG, albeit to a lesser extent in its absence in most cases. This would indicate that even a small amount of antisense RNA is sufficient to have some down-regulatory effect. However, it would be more useful for the expression of the antisense RNA to be more tightly regulated, in order to control the exact timing of target gene down-regulation.

Although these findings indicate that *agrB* is involved in sporulation, they do not prove conclusively that an AIP is involved. Cross-feeding assays, using wild type supernatant, could not demonstrate the restoration of sporulation in *agrB*1AS1 (data not shown). To ascertain whether there is in fact a signalling molecule involved, it would need to be detected and ideally purified. Multistage mass spectrometry has previously been used for precisely this purpose (Kalkum *et al.*, 2003) and enabled the characterisation of secreted quorum sensing peptides in *S. aureus*. This would also provide more evidence as to whether sporulation is under the influence of a system utilising a secreted AIP.

In summary, the data obtained in this study provide important information as to how sporulation is controlled in *C. sporogenes*. This may also prove relevant in *C. botulinum*, which would be of great value to the food industry in the search for methods of inhibiting sporulation, and therefore improving food safety. However, antisense technology is not an ideal method of gene modulation, and further work should preferably be carried out using knockout mutants.

Chapter 5.0

Gene Inactivation in *Clostridium botulinum* and *Clostridium sporogenes* using the ClosTron System

Chapter 5 - Gene Inactivation in *C. botulinum* and *C. sporogenes* Using the ClosTron System

5.1 Introduction

Until recently, the lack of an effective method of clostridial gene knockout has meant that other methods have had to suffice. However, neither of the alternatives described in Chapter 4 are ideal. The use of unstable plasmids is unreliable in that the homologous recombination events occurring are not tightly controlled and therefore undesired recombination events may occur. Antisense RNA technology also has drawbacks, as absolute gene inactivation is not achieved. In addition, the intended target gene may not be the only gene affected by the antisense fragment. Indeed, the *agrB*1 and *agrB*2 antisense fragments used in the construction of the 'mutants' in Chapter 4 show relatively high similarity, and so the possibility of either fragment acting on both *agrB* genes cannot be discounted. Similarly, the smaller antisense fragments used in the two orphan kinase mutants exhibit 59.2% sequence identity, again suggesting that it may be possible for a given fragment to interfere with the translation of both proteins.

As discussed in Chapter 1, the use of suicide vectors to induce double crossover allelic exchange mutagenesis is practised in many bacteria, but has been largely unsuccessful in the clostridia. The only clostridial species on which successful reports of double crossover mutations can be found is *C. perfringens*. Sarker *et al.* (1999) and Huang *et al.* (2004) used suicide vectors to construct mutants against the enterotoxin gene and the *spo0A* gene respectively. However, the latter group found it necessary to screen 3000 colonies before a double crossover mutant was isolated. To date, mutants obtained in all other clostridial species by this method have all been through single crossover events. Clearly, a much more efficient alternative to this method must be utilised in order to obtain stable clostridial knockout mutants.

An alternative method of gene inactivation in Gram-positive bacteria has been the focus of attention over the past few years and recently has been developed into a universal

clostridial gene knockout system (Heap *et al.* 2007). This system utilises a mobile group II intron, originally isolated from *Lactococcus lactis*, to induce site-specific mutations in the bacterial chromosome. Group II introns are perhaps more commonly recognised in eukaryotic cells, and are nucleotide sequences which interrupt protein-coding genes. However, they are also found in bacterial genomes where they are able to self-splice out of the RNA transcripts of the host gene. Some group II introns encode an Intron Encoded Protein (IEP) which contains various enzymatic activities (reverse transcriptase, maturase [promotion of RNA splicing] and site-specific DNA endonuclease). These render the intron a mobile element, allowing it to insert into intron-less copies of the host gene within the same cell, by a highly efficient and specific process known as retrohoming. The host DNA target site is recognised by a Ribo Nuclear Protein (RNP) complex, consisting of the IEP and the excised intron RNA. The intron RNA reverse splices into the double stranded target DNA and is then reverse transcribed by the intron-encoded reverse transcriptase (Mohr *et al.*, 2000).

This insertion ability of mobile group II introns was originally capitalised on by Alan Lambowitz and co-workers at the University of Texas, who proposed that group II introns might be adapted for use in targeted gene disruption in bacterial genomes (Mohr *et al.*, 2000). They utilised the mobile group II intron from the *ltrB* gene of *Lactococcus lactis*. This gene is found on conjugative plasmids, allowing its insertion into intron-less copies of the gene after introduction into a new host cell by conjugation. The IEP of this system is the LtrA protein, whose ORF is located within the group II intron. This research group proposed that altering the intron recognition sites would allow the re-targeting of the group II intron to almost any desired DNA target (Mohr *et al.*, 2000; Karberg *et al.*, 2001).

Recognition of the DNA target site occurs primarily through sequences in the intron RNA. Three sites, known as exon binding sites (EBS) 1, 2 and 1δ, which are approximately 5 bases long, base pair with complementary sequences in the target DNA known as intron binding sites (IBS) 1, 2 and 1δ' (Mohr *et al.*, 2000). It was discovered that, as the IEP recognises only a few additional bases in the flanking regions of the DNA target site, altering the sequence of the EBS1 and EBS2 regions of the intron altered its specificity, allowing it to re-target to a given gene selected for knockout (Karberg *et al.*, 2001). These re-targeted introns became known as 'TargeTrons'. The

structure and recognition sites of the *L. lactis ltrB* group II intron are shown in Figure 5.1.

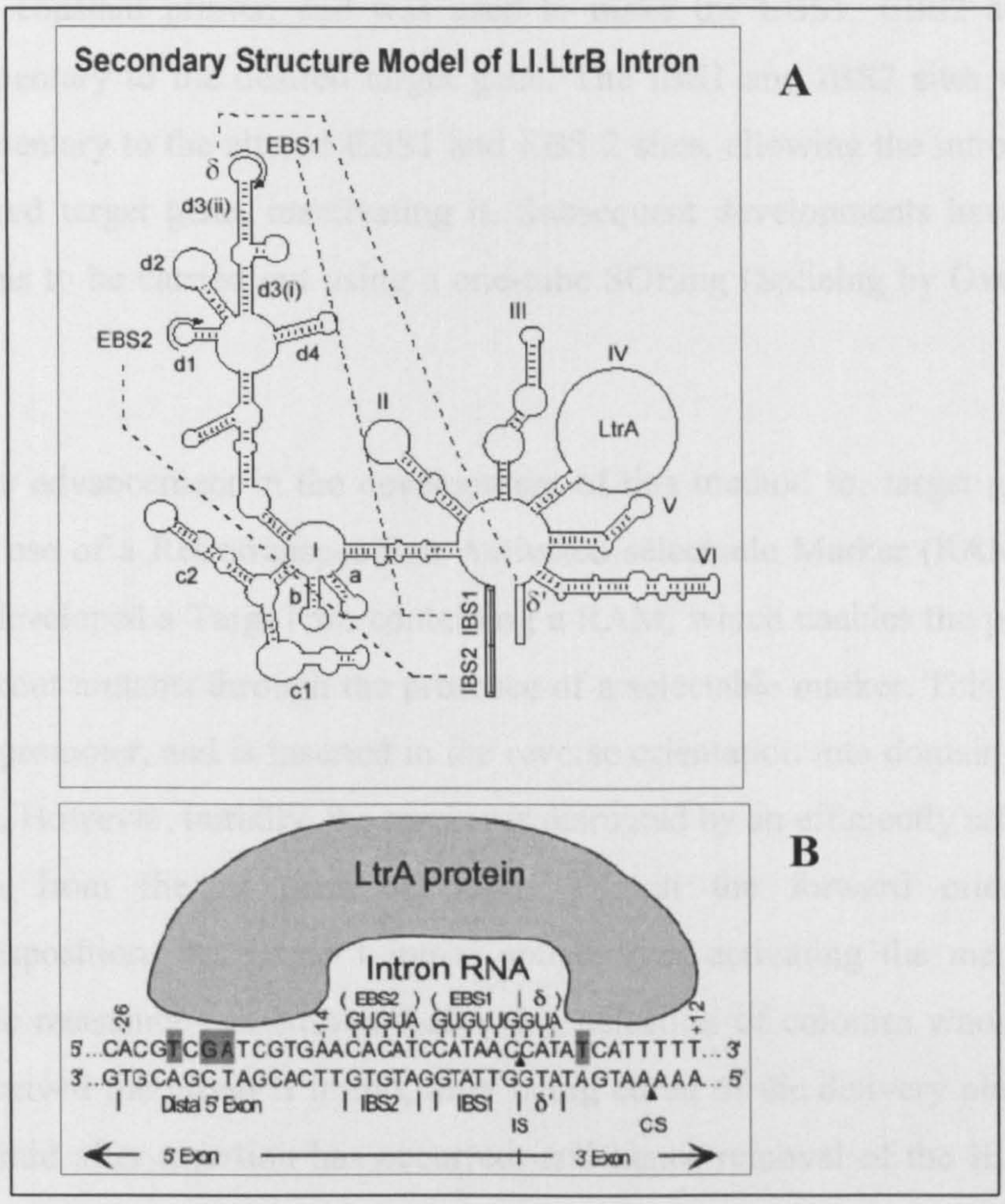


Figure 5.1 – Secondary structure and recognition sites of the *L. lactis ltrB* group II intron.

A – Secondary structure of the intron. The EBS1, EBS2 and δ recognition sites are indicated in domain 1. The LtrA IEP is indicated in the non-structural loop of domain IV.

B – Interaction of the EBS1/IBS1, EBS2/IBS2 and δ/δ' sites in the intron and target gene respectively. Interactions between the IEP and target nucleotides are shown in grey.

Adapted from Mohr *et al.*, 2000 and Zhong *et al.*, 2003.

The fact that the mechanism has only minimal dependence on host factors makes it an attractive method for use in a wide range of bacteria. For example, the *L. lactis ltrB* intron has been utilised for gene knockout in both *E. coli* and Gram positive bacteria (Karberg *et al.*, 2001). This group went on to develop a PCR-based method for generating the re-targeted introns. This was a two-step PCR, using three unique primers and one constant primer, and was used to make the EBS1, EBS2 and δ sequences complementary to the desired target gene. The IBS1 and IBS2 sites were then made complementary to the altered EBS1 and EBS 2 sites, allowing the intron to re-target to the desired target gene, inactivating it. Subsequent developments have allowed these alterations to be carried out using a one-tube SOEing (Splicing by Overlap Extension) PCR.

A further advancement in the development of this method for target gene inactivation was the use of a Retrotransposition Activated selectable Marker (RAM). Zhong *et al.* (2003) developed a TargeTron containing a RAM, which enables the positive selection of knockout mutants through the presence of a selectable marker. This marker contains its own promoter, and is inserted in the reverse orientation into domain IV of the group II intron. However, initially, the marker is disrupted by an efficiently self-splicing group I intron from the *td* gene of phage T4, in the forward orientation. During retrotransposition, the group I intron splices out, activating the marker (usually an antibiotic resistance cassette) and allowing selection of colonies whose chromosomes have received the group II intron, after being cured of the delivery plasmid. Curing of the plasmid after insertion has occurred, and hence removal of the IEP, also prevents any further mobilisation of the intron, and any further mutations from occurring (Heap, 2006).

The use of these re-targeted group II introns was simplified by the development of a computer algorithm which predicts optimal *L. lactis ltrB* intron insertion sites, and designs primers for modifying the intron to re-target to these sites (Perutka *et al.*, 2004).

5.1.1 Adaptation of the group II intron for use in *Clostridium* spp.

The construction of the group II intron delivery vector pMTL007 has recently been carried out by Heap *et al.* (2007). This group modified the commercialised 'TargeTron' vector pACD4K-C, provided by Sigma, for use in clostridial species, as some of its features were non-functional in clostridia. In the pMTL007 vector, the sequence encoding the IEP, LtrA has been removed from within the intron, and relocated further downstream in the plasmid. This relocation allows the protein to act *in trans*, without affecting mobility of the intron.

Expression of the intron in pACD4K-C is driven by the T7 promoter, which does not function in clostridia. Therefore, Heap *et al.* (2007) replaced this promoter with the ferredoxin promoter from *C. pasteurianum*, which functions strongly in a wide range of clostridial species. The *lacI* gene was introduced to allow for inducible expression of the intron, as the ferredoxin promoter is fused to the *lac* operator. They also introduced the pCB102 replicon region to allow replication in clostridia. The *catP* cassette (coding for chloramphenicol acetyltransferase from *C. perfringens*) was also introduced into the plasmid backbone to allow for selection of thiamphenicol resistance in clostridia. The RAM used in vector pACD4K-C contains a kanamycin resistance cassette, which is ineffective in clostridia due to their natural kanamycin resistance. A new RAM containing an erythromycin resistance cassette from the *E. faecalis* pAMβ1 plasmid, and driven by the thiolase promoter was therefore constructed, which would function efficiently in clostridial species. This 'ErmRAM' was inserted into the non-structural loop in which the *ltrA* sequence had previously been located. The final vector was named pMTL007 (Figure 5.2), and the modified group II intron which could be successfully used in clostridial species was termed the 'ClosTron'.

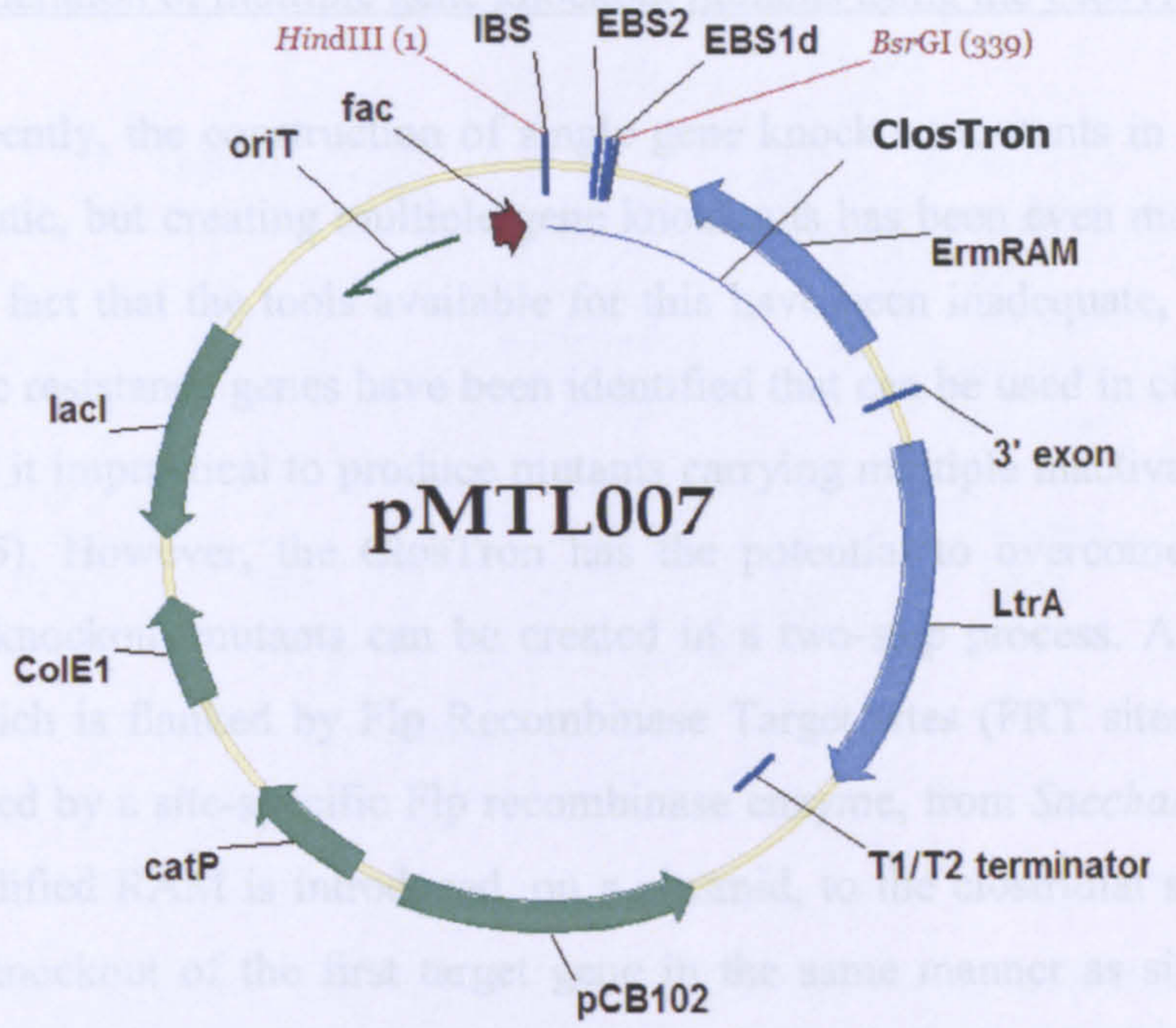


Figure 5.2 – Plasmid map of the pMTL007 ClosTron vector. The vector contains both Gram-positive and Gram-negative replication machinery for use in both *E. coli* and clostridia. The 350 bp region of the intron, retargeted to the gene of choice, is inserted between the *Hin*dIII and *Bsr*GI sites. Inducible expression of the retargeted intron is carried out from the *fac* promoter once LacI is de-repressed by IPTG induction. The *catP* gene is used to positively select for the presence of the vector by chloramphenicol/thiamphenicol resistance. Once the ErmRAM is spliced and inserted chromosomally, erythromycin resistance is conferred. The LtrA protein is located outside of the intron region but is able to act *in trans* with the spliced intron and facilitate the retrohoming process.

5.1.2 Generation of multiple gene knockout mutants using the ClosTron System

Until recently, the construction of single gene knockout mutants in clostridia has been problematic, but creating multiple gene knockouts has been even more difficult. Along with the fact that the tools available for this have been inadequate, only a few natural antibiotic resistance genes have been identified that can be used in clostridia, which has rendered it impractical to produce mutants carrying multiple inactivated genes (Chen *et al.*, 2005). However, the ClosTron has the potential to overcome these limitations. Double knockout mutants can be created in a two-step process. A modified RAM is used which is flanked by Flp Recombinase Target sites (FRT sites). These sites are recognised by a site-specific Flp recombinase enzyme, from *Saccharomyces cerevisiae*. The modified RAM is introduced, on a plasmid, to the clostridial strain of interest, to induce knockout of the first target gene in the same manner as single knockouts are created. However, the mutants obtained are subsequently cured of this first plasmid and a second plasmid expressing Flp recombinase is introduced. This enzyme induces the excision of the first RAM by catalysing homologous recombination between the two FRT sites. After curing of the recombinase-carrying plasmid, the process is repeated to induce disruption of a second target gene. This process has been used in *E. coli* (Zhong *et al.*, 2003), but until recently has not been attempted in clostridia. In this study it was used in an attempt to create an *agrD*1/*agrD*2 double mutant in *C. botulinum*, using plasmid pMTL007C-E2. This plasmid is illustrated in Figure 5.3.

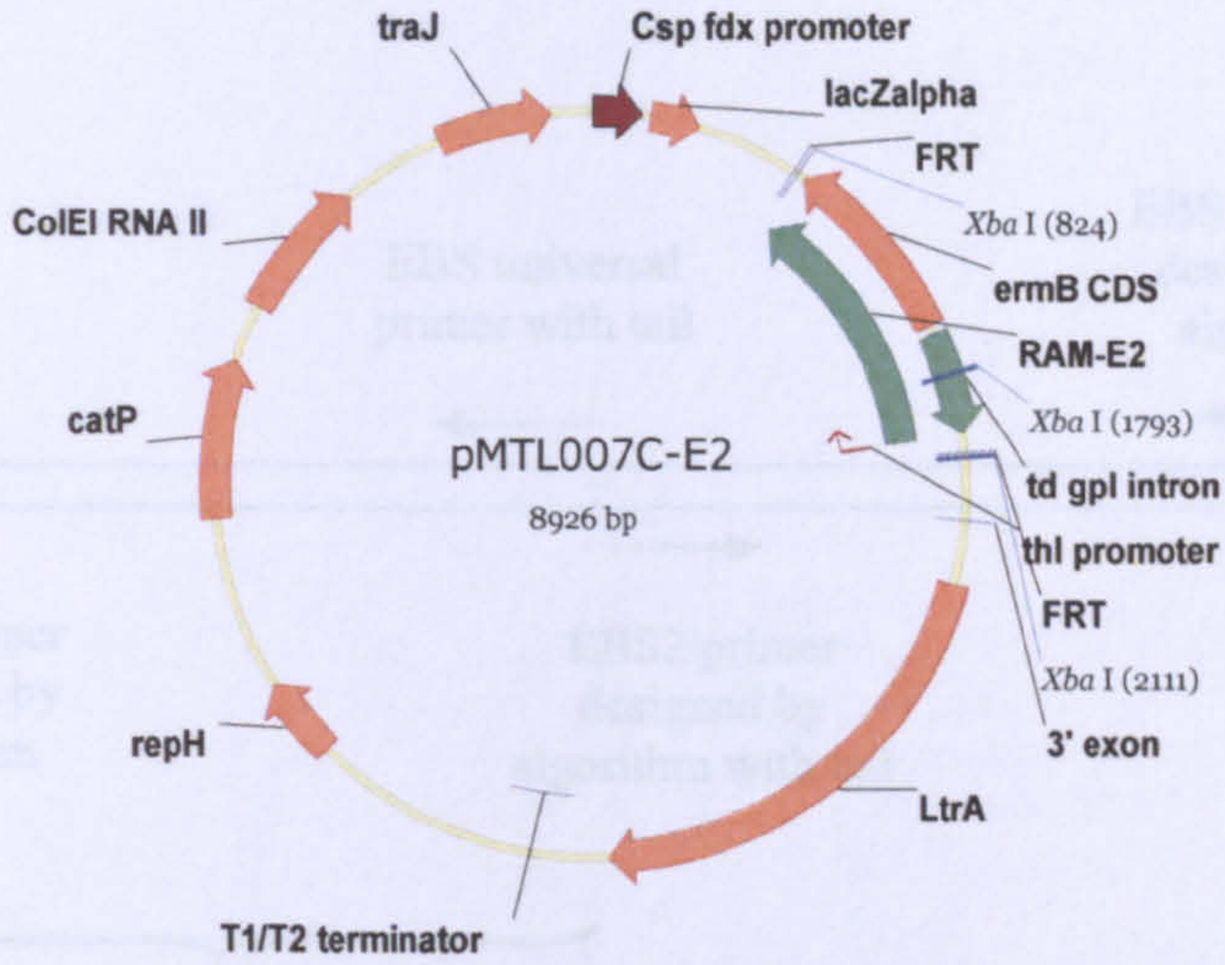


Figure 5.3 – Plasmid map of pMTL007C-E2. The vector is used to deliver retargeted group II introns in the generation of multiple gene knockout mutants. The vector differs from the original pMTL007 vector in that the expression of the group II intron is under constitutive control from the *spofdx* promoter. The ErmRAM is also flanked by FRT sites which should allow the RAM to be removed by the Flp recombinase enzyme. This would allow subsequent genes to be inactivated using the same procedure.

A detailed description of the generation of mutants using the ClosTron system is illustrated in Figure 5.4.

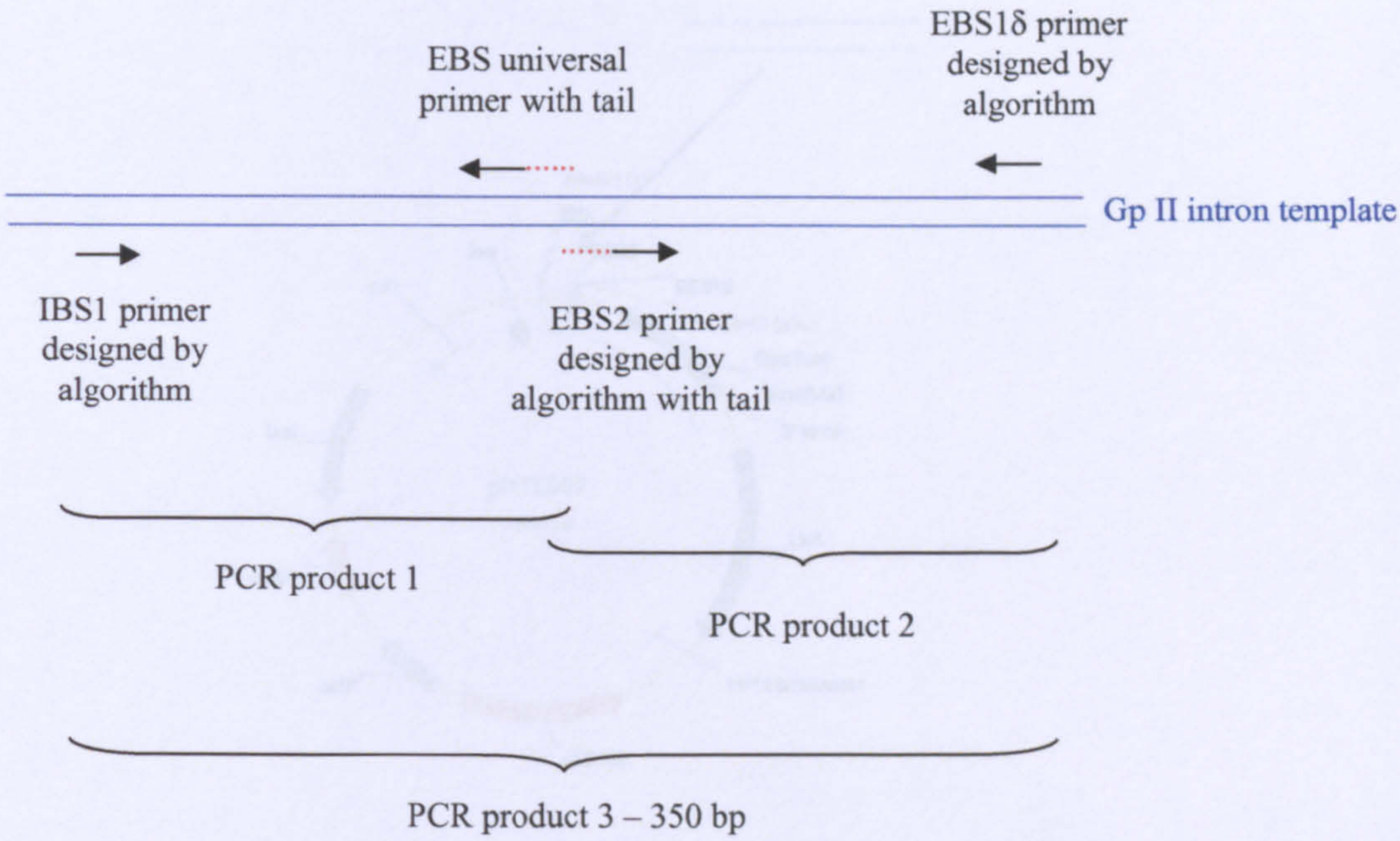


Figure 5.4A - SOEing PCR carried out on *L. lactis ltrB* group II intron.
The resulting 'retargeted' intron is rendered specific for the target gene of interest. Initially products 1 and 2 are amplified from the IBS1/EBS and EBS2/EBS1δ primer pairs respectively. The tails on the EBS universal and EBS2 primers cause a mutation in the two products. The 3' end of product 1 is complementary to the 5' end of product 2, resulting in the hybridisation of the two products. The full 350 bp product is then amplified by the IBS1 and EBS1δ primers. This all occurs in a one-tube reaction.

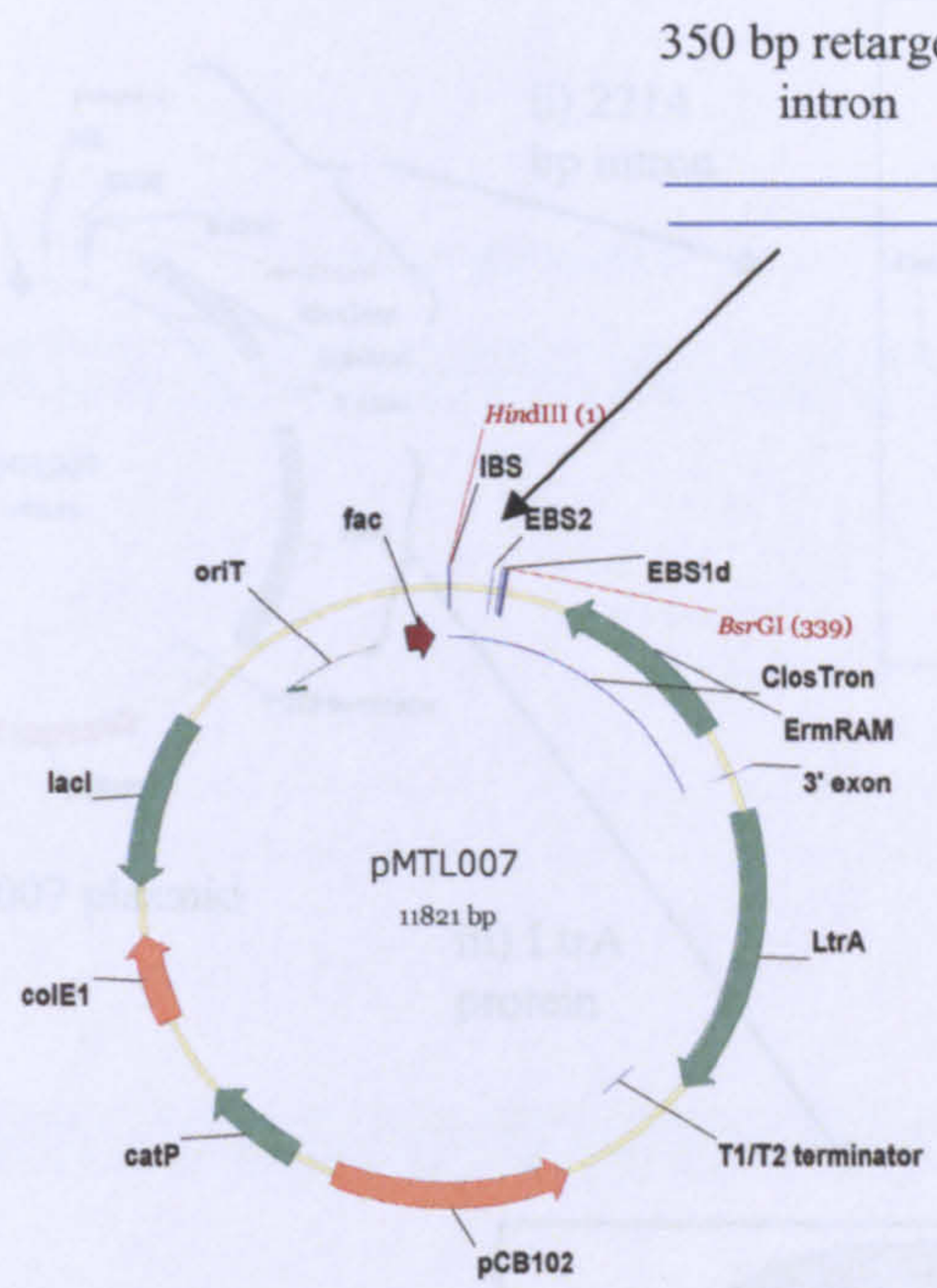


Figure 5.4B – Ligation of retargeted vector into pMTL007 plasmid.
The 350 bp retargeted region of the intron is purified, digested and ligated into the pMTL007/pMTL007C-E2 plasmid between the *Hin*dIII and *Bsr*GI restriction sites. The retargeted vector is then introduced into the target organism by conjugation. Cells carrying the plasmid are selected for using thiamphenicol.

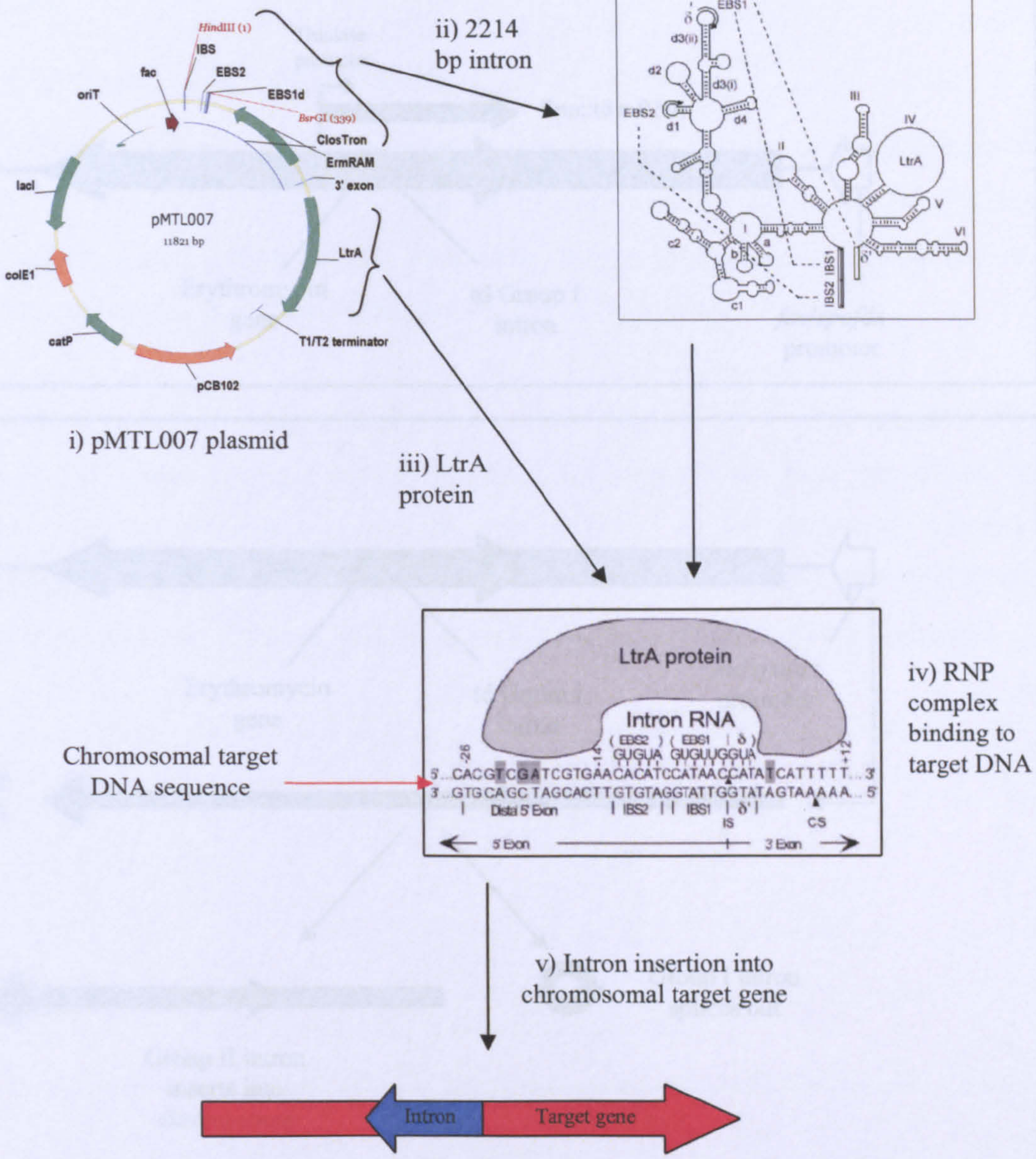


Figure 5.4C – Expression of intron RNA and retrohoming of RNP.
i) Expression of the intron from the pMTL007 plasmid is induced using IPTG. Expression from the pMTL007C-E2 plasmid is constitutive. Once the intron is expressed, retrohoming can occur. **At this point, the group I intron is spliced out of the erythromycin resistance gene of the RAM – see Figure 5.4D**.
ii) The intron RNA folds into its functional configuration and binds to the LtrA protein.
iii) This results in the ribonuclear protein RNP complex consisting of the LtrA protein with the intron RNA wound around it.
iv) The RNP complex scans the chromosomal DNA in search of a target site. The LtrA protein first recognises the four bases shown in grey, which allows unwinding of the target DNA. The intron RNA can then bind to the 13 bases of the IBS2, IBS1 and IBS1δ sites.
v) The intron RNA reverse splices into the double stranded target DNA and is reverse transcribed by the intron-encoded reverse transcriptase. Host enzymes repair the DNA to create a permanent insertion.

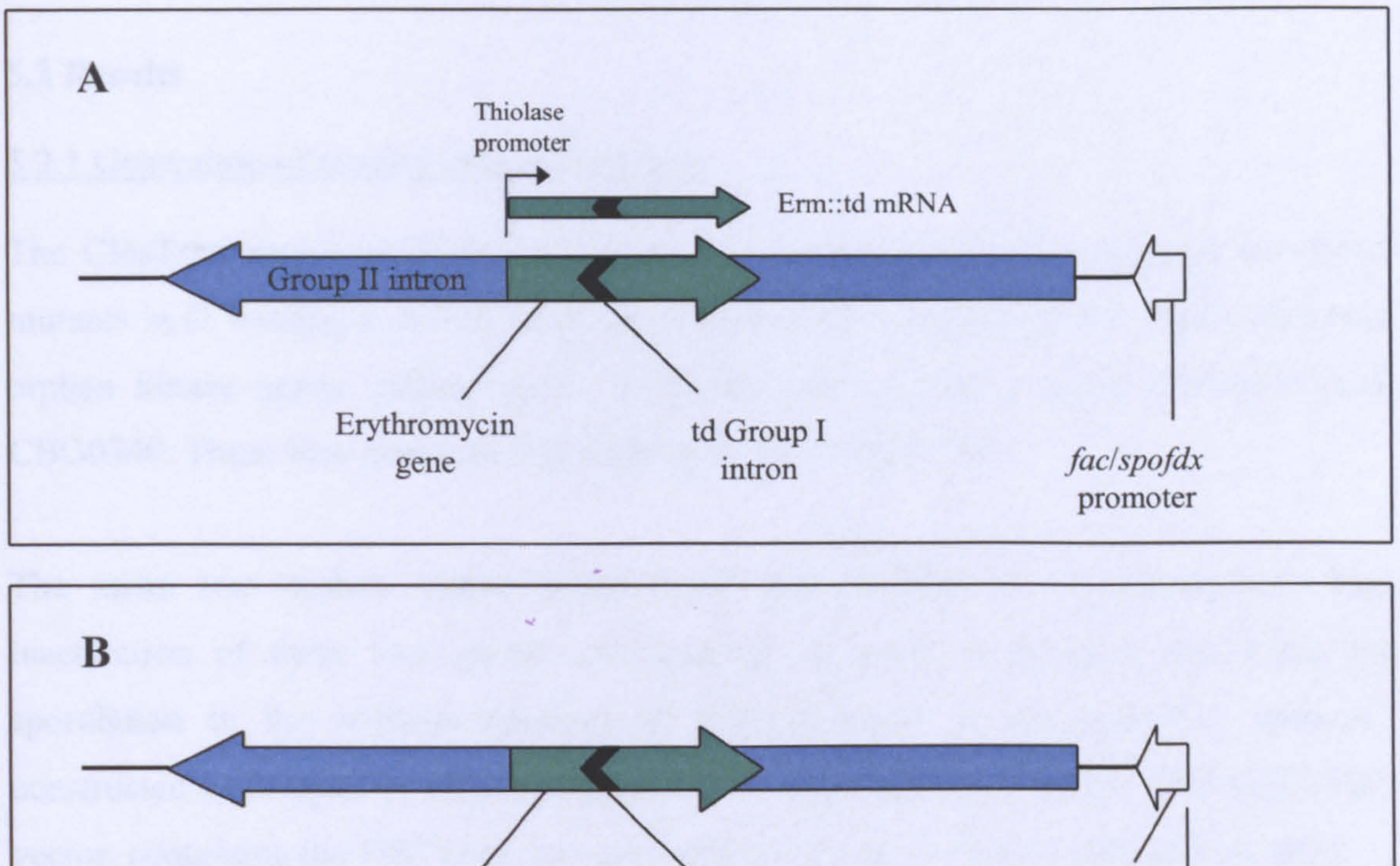


Figure 5.4D – Splicing of the ErmRAM in ClosTron mutants. Adapted from Heap (2006).

The only scenario in which erythromycin resistance is conferred on the cell is when the group I intron is spliced out of the RAM and the group II intron is inserted into the chromosome.

A. Expression of the ErmRAM from the pMTL007/pMTL007C-E2 plasmid. No erythromycin resistance is conferred because the mRNA produced retains the *td* group I intron insertion. This element cannot splice out because it is in the wrong orientation.

B. Once group II intron production is induced, transcription proceeds from the *fac/spofdx* promoter. The *td* group I intron is now in the correct orientation and splices out. The group II intron inserts into the chromosome, conferring erythromycin resistance on to the cell, due to the transcription of the erythromycin resistance gene from the thiolase promoter.

5.2 Results

5.2.1 Generation of single knockout mutants

The ClosTron vector pMTL007 was used to generate the first documented knockout mutants in *C. botulinum* ATCC 3502. Four genes were targeted; *agrD*1, *agrD*2 and two orphan kinase genes within the *C. botulinum agr* regions – genes CBO0336 and CBO0340. These four genes are highlighted in red in Figure 5.5.

The same two orphan kinase genes were also targeted in *C. sporogenes*. The inactivation of these two genes was required in order to compare the effect on sporulation in the mutants obtained to that observed in the antisense 'mutants' constructed in Chapter 4. In the construction of these two mutants the pMTL007C-E2 vector, containing the FRT sites, was used rather than the original pMTL007 plasmid.

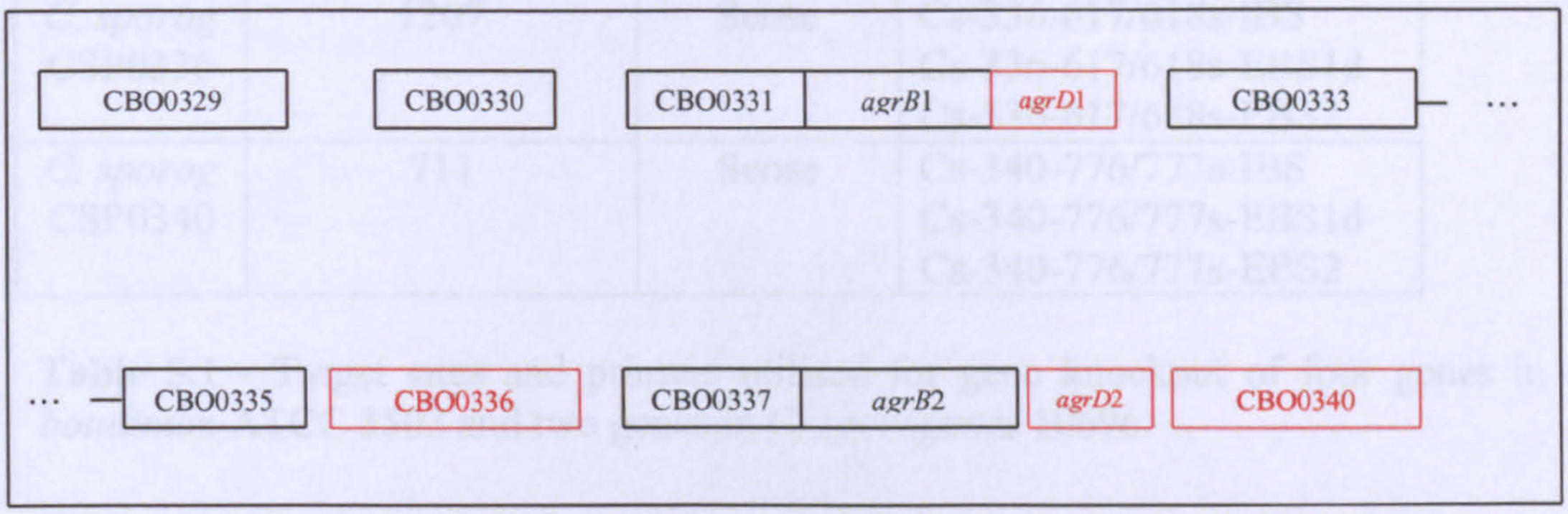


Figure 5.5 – Schematic representation of the *agr* regions in the *C. botulinum* ATCC 3502 genome (not to scale). Genes targeted for knockout using the ClosTron system are shown in red. Genes 0336 and 0340 were also targeted in *C. sporogenes*.

The target sites selected and the primers designed by the computer algorithm (available at http://www.sigma-genosys.com/targetron/) for use in the re-targeting PCR are shown in Table 5.1. The SOEing PCRs used the three unique primers indicated, along with the EBS universal primer.

Target Gene	Intron insertion site (bp into gene)	Orientation of intron	Primers used to generate re-targeted intron
*C. bot agrD*1	23	Antisense	Cb-agrD1-2831a-IBS Cb-agrD1-2831a-EBS1d Cb-agrD1-2831a-EBS2
*C. bot agrD*2	47	Antisense	Cb-agrD2-2990a-IBS Cb-agrD2-2990a-EBS1d Cb-agrD2-2990a-EBS2
C. bot CBO0336	1203	Sense	Cb-336-617/618s-IBS Cb-336-617/618s-EBS1d Cb-336-617/618s-EBS2
C. bot CBO0340	690	Sense	Cb-340-776/777s-IBS Cb-340-776/777s-EBS1d Cb-340-776/777s-EBS2
C. sporog CSP0336	1207	Sense	Cs-336-617/618s-IBS Cs-336-617/618s-EBS1d Cs-336-617/618s-EBS2
C. sporog CSP0340	711	Sense	Cs-340-776/777s-IBS Cs-340-776/777s-EBS1d Cs-340-776/777s-EBS2

Table 5.1 - Target sites and primers utilised for gene knockout of four genes in *C. botulinum* ATCC 3502 and two genes in *C. sporogenes* 10696.

The SOEing PCR was then carried out to alter the region of the group II intron responsible for its specificity. This was done using the primers indicated in Table 5.1, along with the EBS universal primer, and was carried out in a single one-tube reaction. A four-primer mix was first set up as follows:

2 µl 100 µM IBS primer

2 µl 100 µM EBS1d primer

2 µl 20 µM EBS2 primer

2 µl 20 µM EBS universal primer

12 µl water

This primer mix was then used to set up a 50 µl PCR reaction mix (see method 2.7.9) using 5 µl four-primer mix and 1 µl 1:10 Intron PCR Template, provided in the Sigma TargeTron Gene Knockout System Kit. The PCR was then carried out in triplicate using the following conditions:

94°C for 30 seconds

94°C for 15 seconds
55°C for 30 seconds } 30 cycles
72°C for 30 seconds

72°C for 2 minutes

An example of the PCR products generated in these reactions is shown in Figure 5.6. Triplicate reactions were pooled before running on an agarose gel and purifying the required 350 bp product. The products were ligated into the pCR2.1 TOPO vector before sequencing to confirm the correct product.

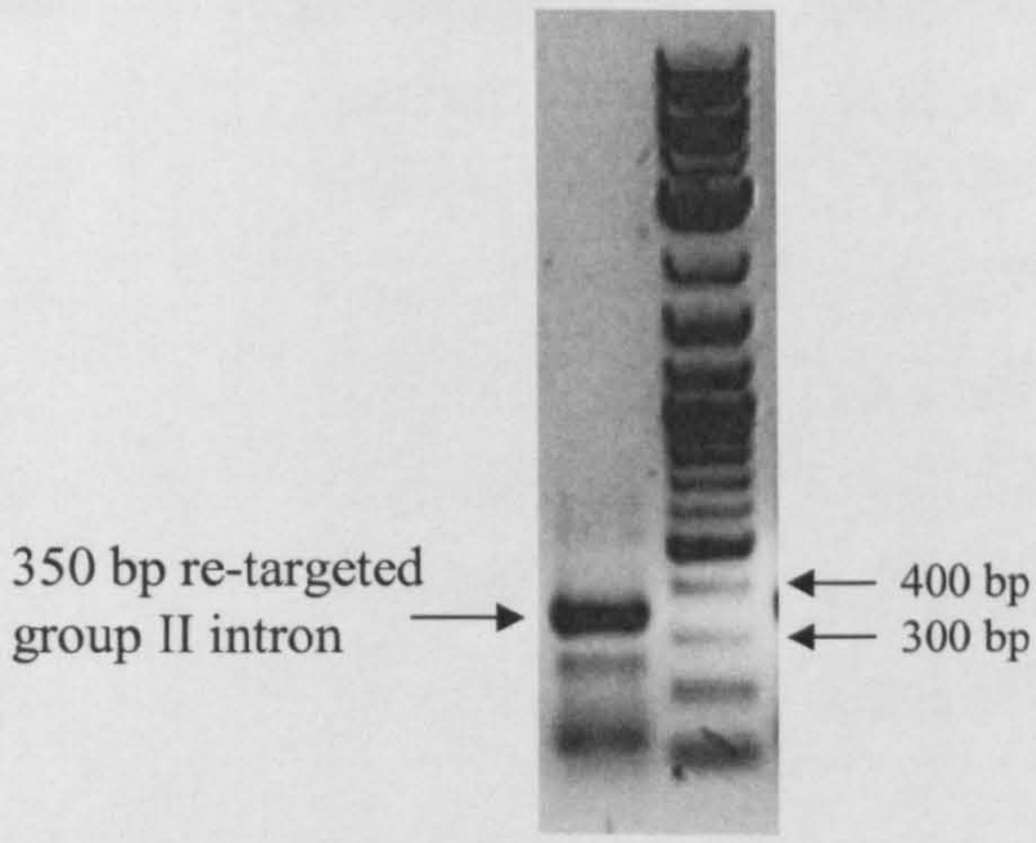


Figure 5.6 – Products obtained from the re-targeting SOEing PCR carried out on the *L. lactis* group II intron. The example shown is from the *C. botulinum agrD*1 PCR and the 350 bp product was purified.

The required products were re-isolated from the pCR2.1 vector by digesting with *Hin*dIII and *Bsr*GI. The pMTL007 or pMTL007C-E2 vectors were cut with the same two enzymes, allowing each of the inserts to be ligated into the vectors. This resulted in the intron delivery vectors containing a re-targeted intron with specificity for one of the six target genes. The vectors were then transformed into *E. coli* TOP10 by electroporation, and selected for using chloramphenicol. Plasmids were isolated from the resulting colonies and sequencing carried out again to further confirm the correct re-targeted intron sequence. Primers used for sequencing were 5402F-F1 & pMTL007-R1 (for pMTL007 plasmid) and Spofdx-seq-F1 & pMTL007-R1 (for pMTL007C-E2 plasmid). The plasmids were finally transformed into the conjugation donor strain *E. coli* CA434, before being conjugated into *C. botulinum* or *C. sporogenes* using the 'spot' method. Transconjugants were selected for using thiamphenicol. Induction of target gene knockout by integration of the group II intron was then carried out as described in methods 2.11.1 and 2.11.2. Results from the conjugation and integration experiments are shown in Table 5.2.

Target Gene	Transconjugants (cfu/ml)	Integrants (cfu/ml)	Frequency of integration
*C. bot agrD*1	2.73×10^{3}	1.2×10^{4}	6.32×10^{-4}
*C. bot agrD*2	1.84×10^{3}	1.38×10^{2}	1.45×10^{-5}
CBO0336	5.6×10^{3}	1.6×10^{2}	8.0×10^{-6}
CBO0340	1.8×10^{3}	1.52×10^{5}	7.0×10^{-3}
CSP0336	8.4×10^{2}	3.93×10^{3}	7.87×10^{-4}
CSP0340	9.2×10^{2}	2.26×10^{7}	7.53

Table 5.2 – Results from conjugation and integration experiments for knockout of genes in *C. botulinum* and *C. sporogenes.*

5.2.2 Screening of putative single knockout mutants

Four to six putative knockout mutants were picked for each gene. Genomic DNA was prepared from each isolate and used to perform three PCR screenings, which are

outlined in Table 5.3 and Figure 5.7. A pool was also screened, where the colonies from an entire induction agar plate were harvested using 1 ml PBS, and the DNA extracted. This was to determine whether a correct integrant was present in the population if none of the individual colonies screened were correct.

PCR Screen	Target Gene	Primers	Expected product (bp)
1 – Amplifies product across intron-exon junction and confirms that intron has inserted into desired gene.	*C. bot agrD*1	• Target R(Cb-agrD1/R) • EBS universal	289
	*C. bot agrD*2	• Target R(Cb-agrD2/R) • EBS universal	294
	CBO0336	• Target F(Cb0336F) • EBS universal	492
	CBO0340	• Target F(Cb0340F) • EBS universal	449
	CSP0336	• Target F(Cs0336F) • EBS universal	492
	CSP0340	• Target F(Cs0340F) • EBS universal	448
2 – Confirms that intron is present in the desired target gene and that RAM is spliced.	*C. bot agrD*1	• Target F (Cb-agrD1/F) • Target R (Cb-agrD1/R)	1700
	*C. bot agrD*2	• Target F (Cb-agrD2/2F) • Target R (Cb-agrD2/R)	1700
	CBO0336	• Target F (Cb0336F) • Target R(Cb0336R)	2000
	CBO0340	• Target F (Cb0340F) • Target R (Cb0340R)	2000
	CSP0336	• Target F (Cs0336F) • Target R(Cs0336R)	2158
	CSP0340	• Target F (Cs0340F) • Target R (Cs0340R)	2240
3 – Indicates that RAM is either spliced and therefore integrated, or presence of full length, unspliced RAM.	*C. bot agrD*1	• ErmRAM F • ErmRAM R	900
	*C bot agrD*2		900
	CBO0336		900
	CBO0340		900
	CSP0336		900
	CSP0340		900

Table 5.3 – Primers used in three different PCR screens of *C. botulinum* and *C. sporogenes* putative knockout mutants. Primer sequences can be found in 2.6.

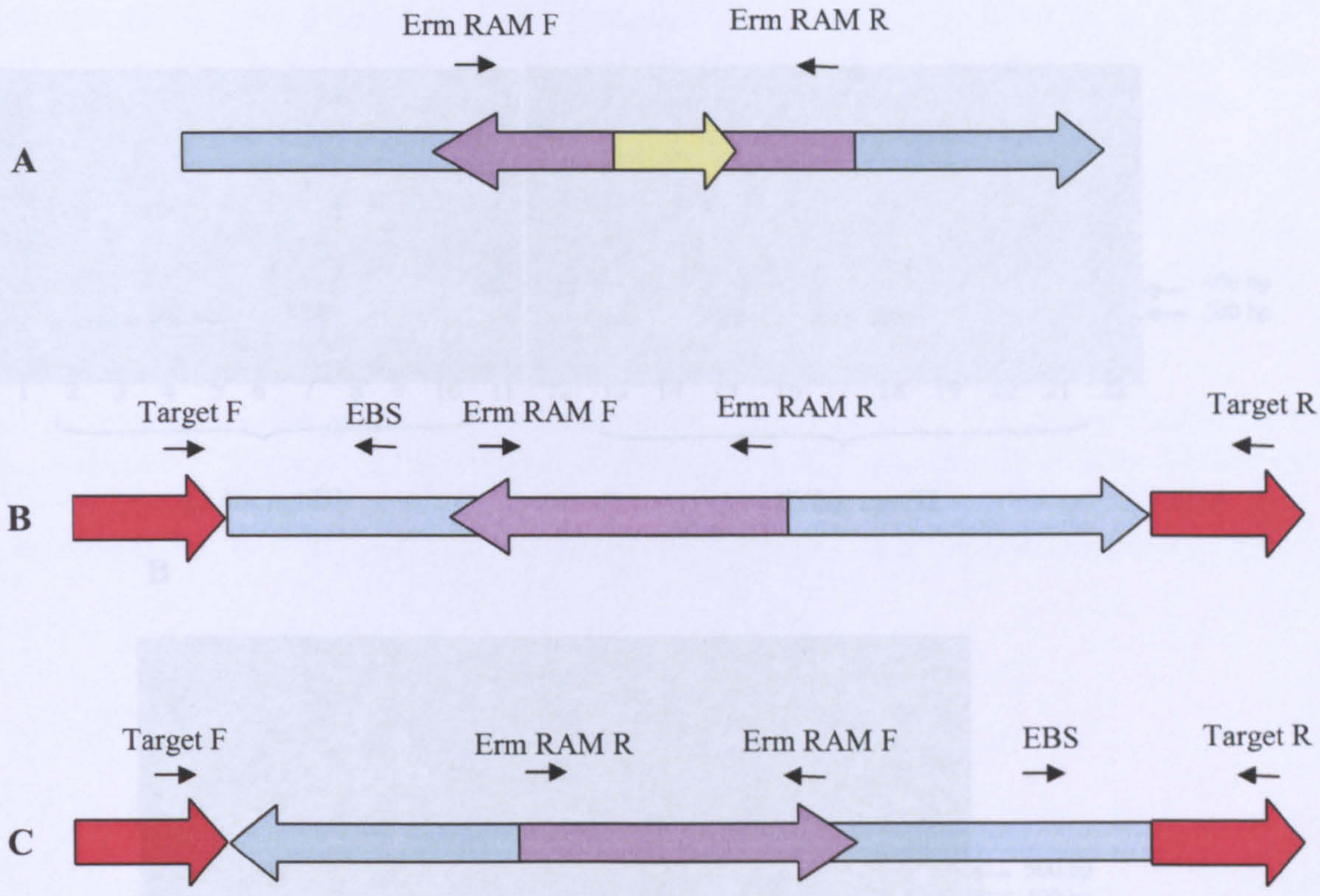
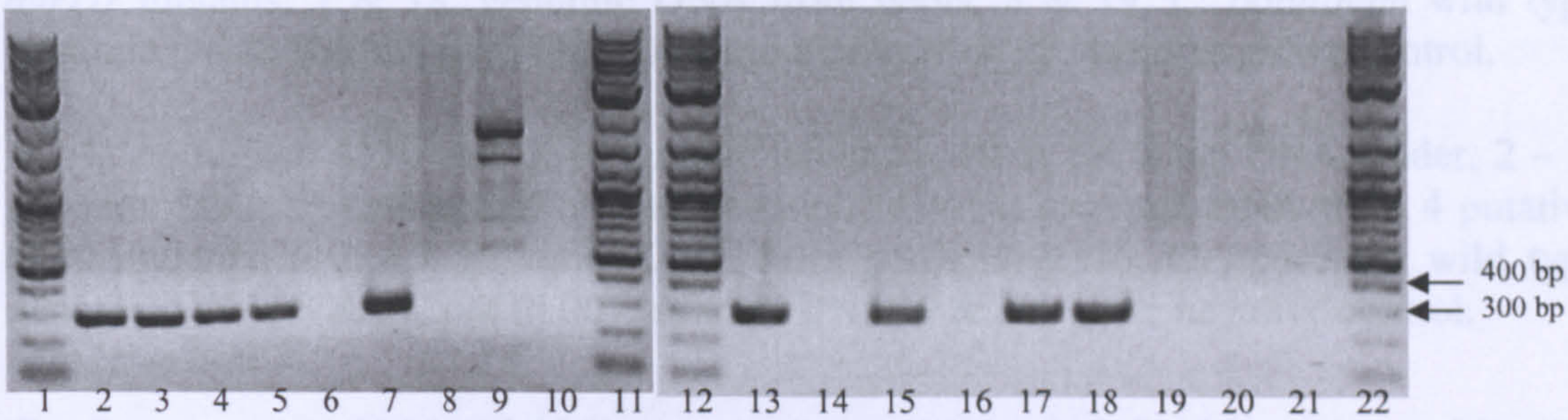


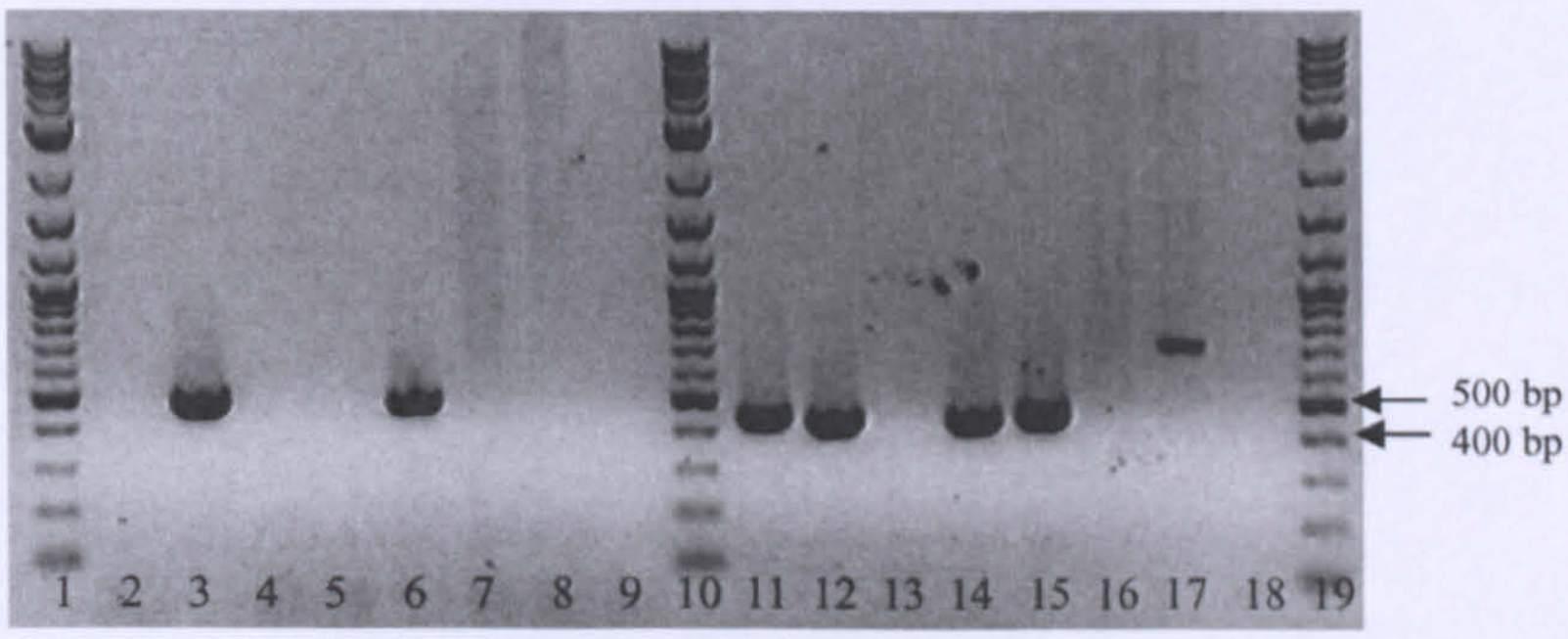
Figure 5.7 – Schematic representation of PCR screens carried out on putative integrants. A, pMTL007/pMTL007C-E2 plasmid DNA; B, chromosomal DNA with the ClosTron inserted into a target gene in the sense orientation (e.g. CBO/CSP0336 and CBO/CSP0340); C, chromosomal DNA with the ClosTron inserted into a target gene in the antisense orientation (e.g. *C. bot agrD*1 and *agrD*2). Red, target gene; blue, group II intron; purple, Erm RAM; yellow, group I intron. Primers are represented by black arrows. Expected product sizes and regions amplified are shown in Table 5.3.

PCR screen 1 results are shown in Figure 5.8. Results of PCR screens 2 and 3 can be found in Appendix 4.

A



B



C

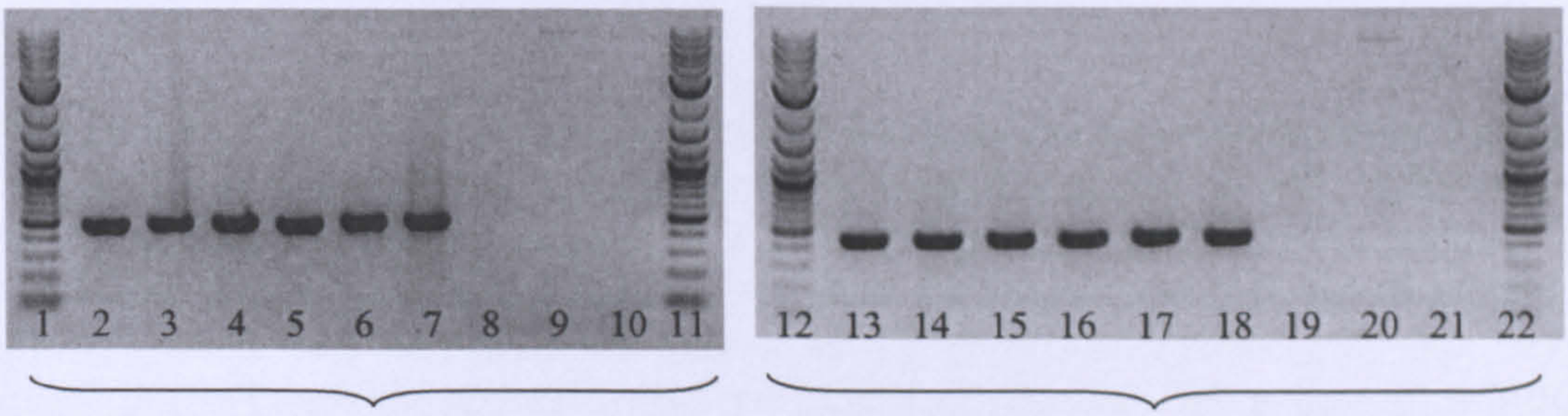


Figure 5.8 – PCR screen 1 on four *C. botulinum* and two *C. sporogenes* mutants. Primers used and expected product sizes are shown in Table 5.3 and Figure 5.7. Non-specific binding of primers to the plasmid control is occasionally observed. This was previously been seen during development of the ClosTron knockout method (Heap *et al.*, 2007), but does not affect interpretation of results.

A - *C. botulinum agrD* mutant screen: 1, 11, 12 & 22, 2 log DNA ladder; 2 – 6, genomic DNA from 5 putative *agrD*1 mutants; 13 – 17, genomic DNA from 5 putative *agrD*2 mutants; 7 & 18, genomic DNA from pools; 8 & 19, *C. botulinum* wild type genomic DNA; 9 & 20, pMTL007 plasmid DNA, 10 & 21, water negative control.

B - *C. botulinum* 0336 and 0340 mutant screen: 1, 10 & 19, 2 log DNA ladder; 2 – 5, genomic DNA from 4 putative 0336 mutants; 11 – 14, genomic DNA from 4 putative 0340 mutants; 6 & 15, genomic DNA from pools; 7 & 16, *C. botulinum* wild type genomic DNA; 8 & 17, pMTL007 plasmid DNA; 9 & 18, water negative control.

C - *C. sporogenes* 0336 and 0340 mutant screen: 1, 11, 12 & 22, 2 log DNA ladder; 2 – 6, genomic DNA from 5 putative 0336 mutants; 13 – 17, genomic DNA from 5 putative 0340 mutants; 7 & 18, genomic DNA from pools; 8 & 19, *C. sporogenes* wild type genomic DNA; 9 & 20, pMTL007C-E2 plasmid DNA; 10 & 21, water negative control.

In addition, Southern hybridisation was used to confirm that only a single intron insertion had occurred. Wild type and mutant chromosomal DNA was digested using *Hin*dIII, which fragments the DNA but does not cut inside the inserted group II intron. The probe used was designed to hybridise to the spliced ErmRAM sequence. Figure 5.9 confirms that only one copy of the ErmRAM, and therefore one intron is present in each of the mutants.

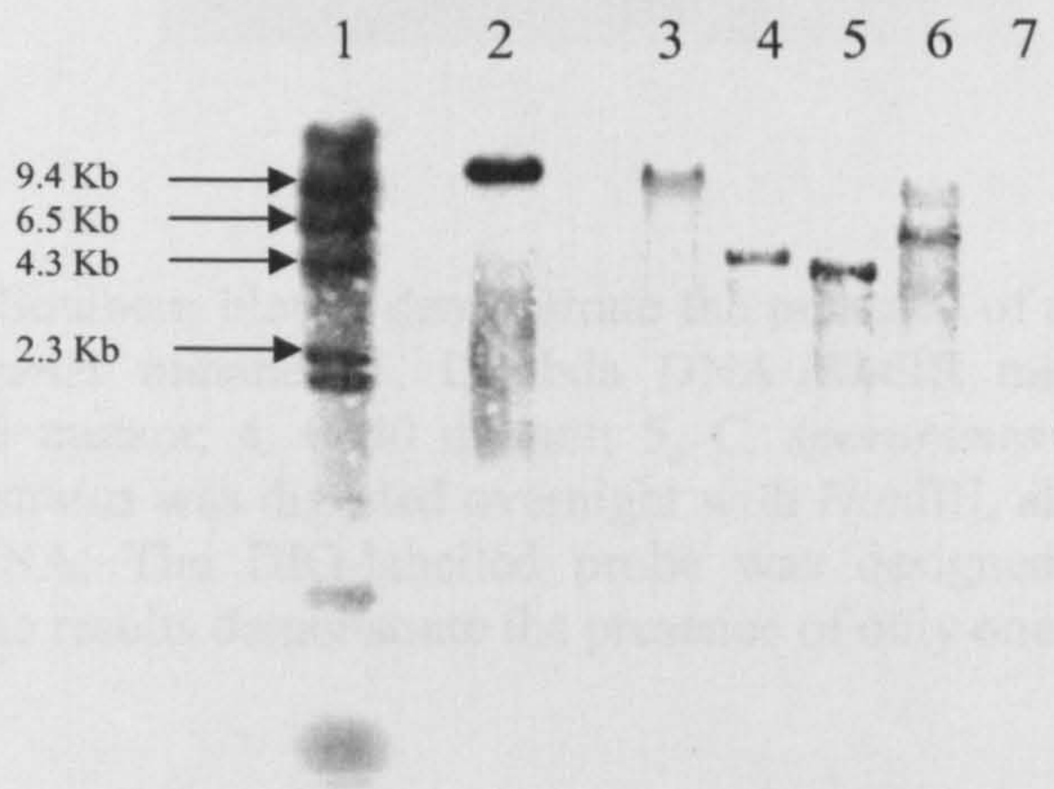


Figure 5.9A – Southern blot to demonstrate the presence of a single intron insertion in four *C. botulinum* mutants. 1, Lambda DNA/*Hin*dIII markers; 2, pMTL007 vector; 3, *agrD*1 mutant; 4, *agrD*2 mutant; 5, CBO0336 mutant; 6, CBO0340 mutant; 7, *C. botulinum* wild type. Chromosomal DNA from all strains was digested overnight with *Hin*dIII, along with the pMTL007 plasmid DNA. The DIG-labelled probe was designed to target the ErmRAM sequence and the results demonstrate the presence of only one RAM in each mutant.

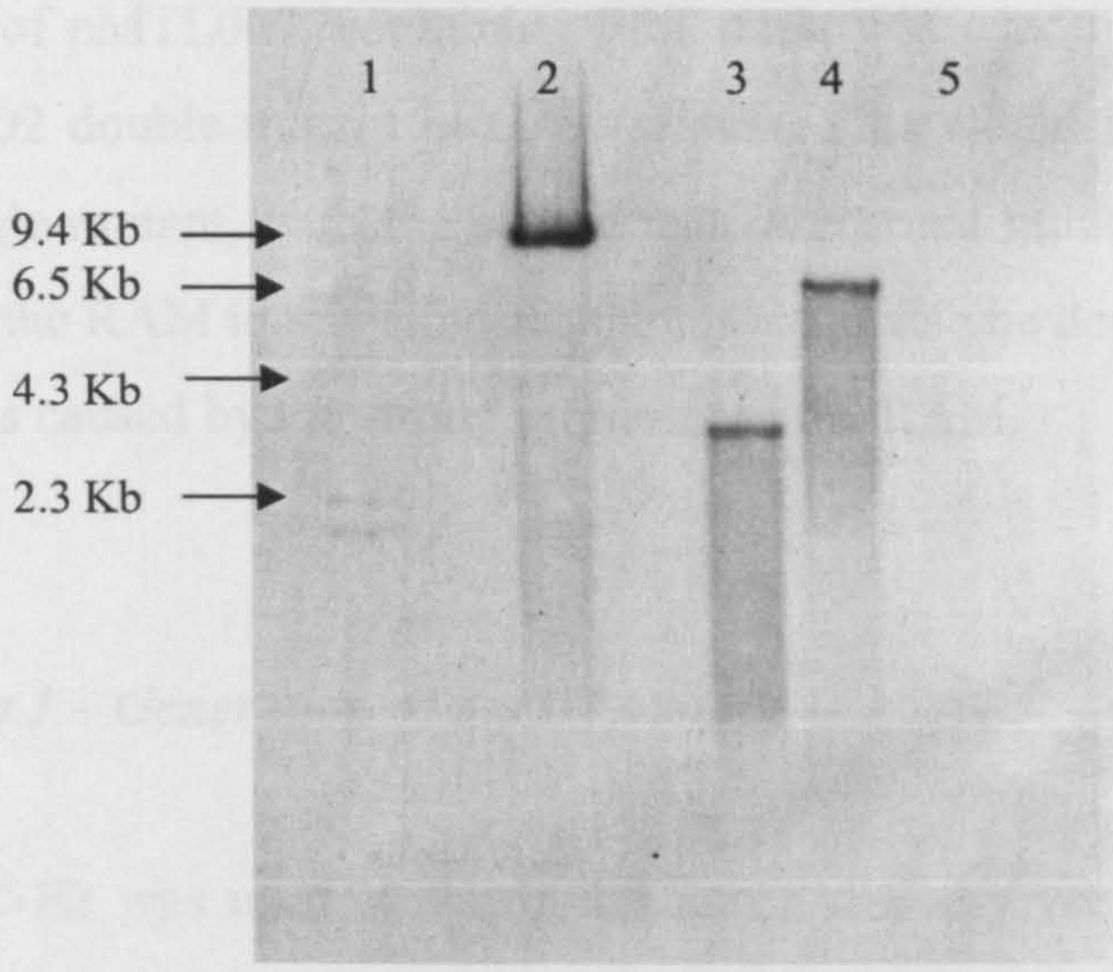


Figure 5.9B – Southern blot to demonstrate the presence of a single intron insertion in two *C. sporogenes* mutants. 1, Lambda DNA/*Hin*dIII markers; 2, pMTL007C-E2 vector; 3, 0336 mutant; 4, 0340 mutant; 5, *C. sporogenes* wild type. Chromosomal DNA from all strains was digested overnight with *Hin*dIII, along with the pMTL007C-E2 plasmid DNA. The DIG-labelled probe was designed to target the ErmRAM sequence and the results demonstrate the presence of only one RAM in each mutant.

Confirmation of the stability of the mutants was carried out by growing a 48-hour broth culture of the mutant and plating this out onto TYG agar in duplicate, in the presence and absence of erythromycin selection. Both plates gave similar colony counts, indicating that the mutants were stable and that erythromycin selection was not required to maintain the mutations.

Genomic DNA from all *C. botulinum* mutants was used to amplify both intron-exon junctions. The products were then sequenced and were found to be correct (data not shown).

5.2.3 Generation of putative double knockout mutants

A derivate of pMTL007, containing FRT sites, was used in an attempt to generate an *agrD*1/*agrD*2 double mutant in *C. botulinum*. This would involve removing the RAM from a single mutant, with the subsequent insertional mutation of a second gene. The removal of the RAM in any single mutants would also be desirable in order to avoid any polar effects caused by the strong promoter of the RAM.

5.2.3.1 Step 1 – Generation of agrD*1 mutant*

pMTL007C-E2 was used as the initial intron delivery vector to inactivate the *agrD*1 gene in the same manner as for the generation of the single mutants. The vector map is shown in Figure 5.3. The retargeted *agrD*1 sequence which had been cloned into the pCR2.1 TOPO vector during the original single knockout process was re-isolated and cloned into the pMTL007C-E2 vector. This was subsequently transformed into the *E. coli* conjugation donor CA434 and conjugated into *C. botulinum* using the single mutant methodology. Since the intron in pMTL007C-E2 is constitutively expressed, induction with IPTG was not required. Instead, one thiamphenicol resistant transconjugant was resuspended in 100 µl TYG broth. This was then serially diluted and 100 µl of each dilution spread onto TYG plates containing erythromycin. Dilutions were also plated onto TYG without selection to allow calculation of integration frequencies. Any putative knockouts were repeatedly subcultured onto fresh TYG erythromycin or thiamphenicol plates, until thiamphenicol sensitivity was detected, indicating that plasmid loss had occurred. Results from the conjugation and integration experiments are shown in Table 5.4.

Target Gene	Transconjugants (cfu/ml)	Integrants (cfu/ml)	Frequency of integration
*agrD*1	1.6×10^{3}	4.3×10^{5}	4.7×10^{-3}

Table 5.4 – Results from conjugation and integration experiments for the knockout of *agrD*1 in *C. botulinum* using plasmid pMTL007C-E2. This was the first stage of the double knockout process.

PCR screens were then carried out on genomic DNA isolated from five putative knockout mutants along with a pool. This was carried out using the same screens used for the single knockout mutant, described in 5.2.2. Results of the screen are shown in Figure 5.10A-C, and indicate that four out of the five putative knockout mutants were correct.

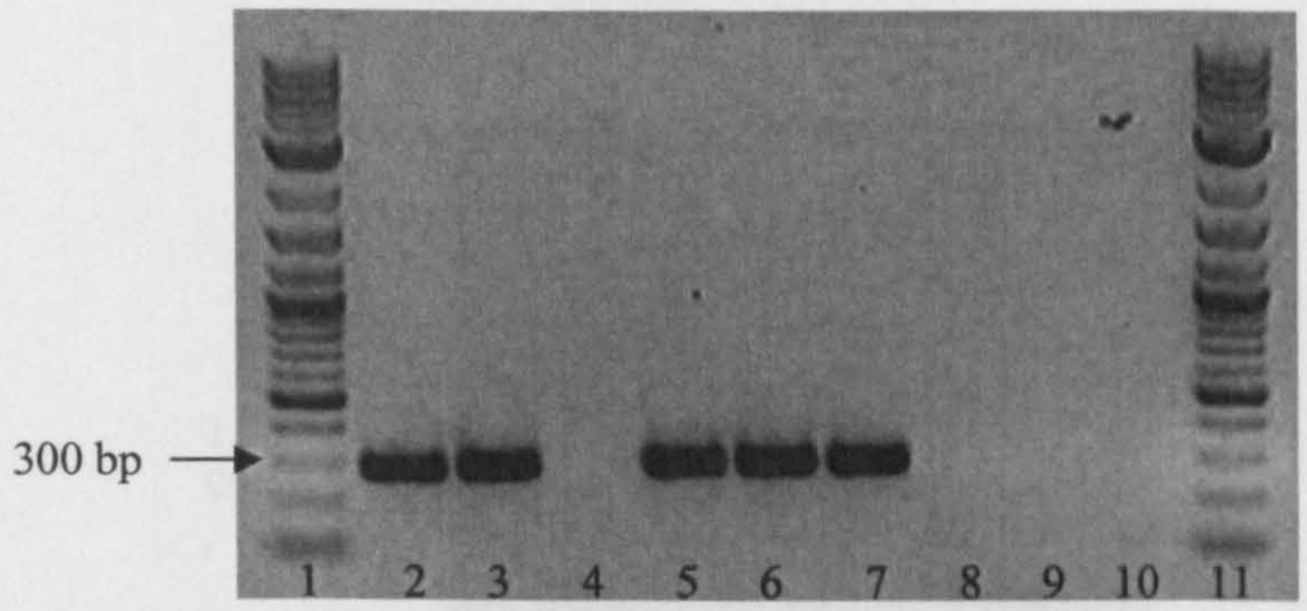


Figure 5.10A – PCR screen 1 on *C. botulinum agrD*1 mutant generated using pMTL007C-E2. Primers used were the same as those given in Table 5.3. 1 & 11, 2 log DNA ladder; 2 – 6, genomic DNA from 5 putative *agrD*1 mutants; 7, genomic DNA from pool; 8, *C. botulinum* wild type genomic DNA; 9, pMTL007C-E2 plasmid DNA, 10, water negative control. Results show that four out of the five putative mutants gave correct products.

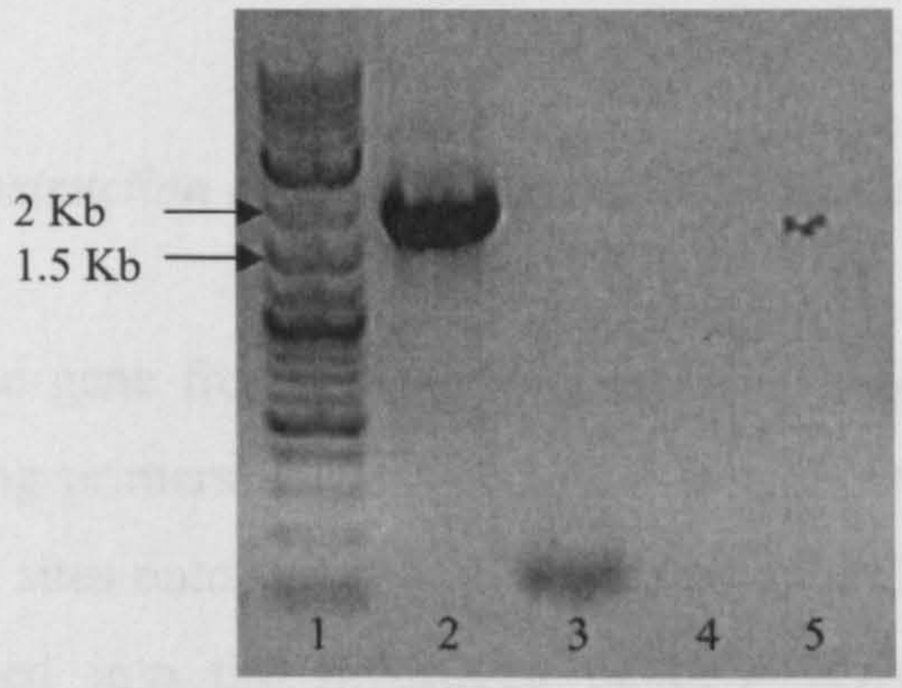


Figure 5.10B – PCR screen 2 on *C. botulinum agrD*1 mutant generated using pMTL007C-E2. Primers used were the same as those given in Table 5.3. 1, 2 log DNA ladder; 2, genomic DNA from putative *agrD*1 mutant 1; 3, *C. botulinum* wild type genomic DNA; 4, pMTL007C-E2 plasmid DNA, 5, water negative control. Genomic DNA from putative mutant 1 gave the correct size product.

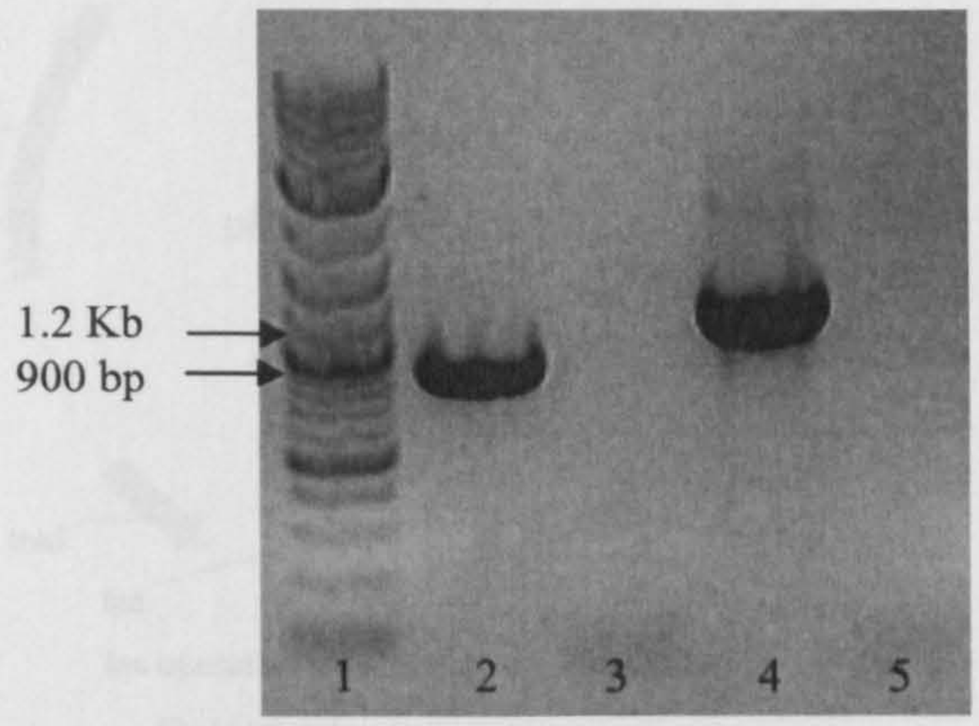


Figure 5.10C – PCR screen 3 on *C. botulinum agrD*1 mutant generated using pMTL007C-E2. Primers used were the same as those given in Table 5.3. 1, 2 log DNA ladder; 2, genomic DNA from putative *agrD*1 mutant 1; 3, *C. botulinum* wild type genomic DNA; 4, pMTL007C-E2 plasmid DNA, 5, water negative control. Genomic DNA from putative mutant 1 gave the correct size product.

5.2.3.2 Step 2 – Construction of plasmid expressing Flp *recombinase*

The *flp* recombinase gene from *Saccharomyces cerevisiae* was PCR amplified from plasmid pCP20 using primers FLP-F1 and FLP-R1. These primers incorporated *Xba*I and *Xho*I restriction sites onto the 5' and 3' ends of the PCR product respectively. The fragment was cloned into the Invitrogen pCR2.1 TOPO vector and checked by sequencing. It was then excised from the pCR2.1 vector and sub-cloned as an *Xba*I/*Xho*I fragment into plasmid pMTL5402F, under the inducible control of the *fac* promoter. This gave plasmid pMTL5402F-FLP, which is shown in Figure 5.11.

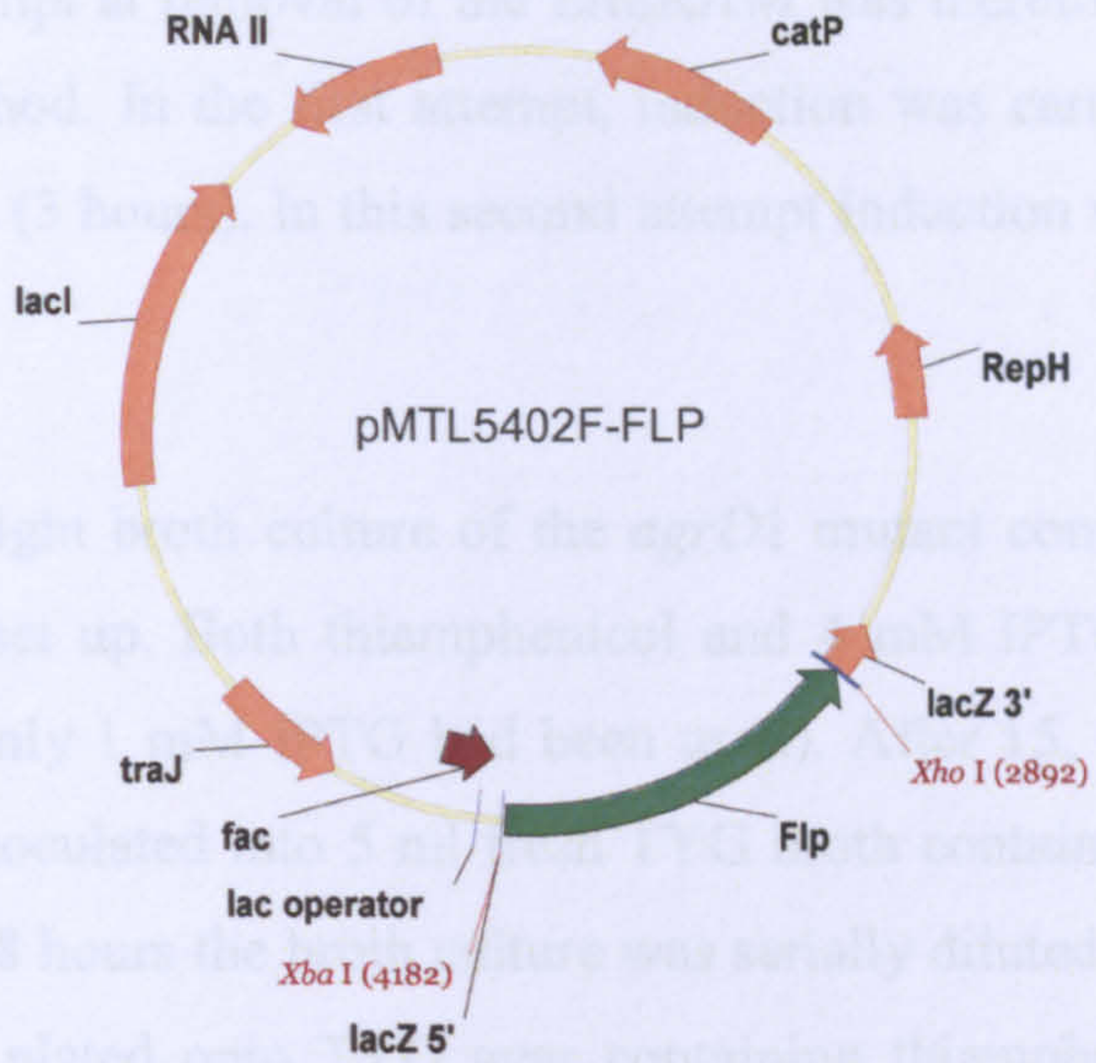


Figure 5.11 – Plasmid map of vector pMTL5402F-FLP. The vector was used to express the Flp protein and therefore remove the ErmRAM in an attempt to construct a *C. botulinum* double *agrD* mutant. The *Saccharomyces cerevisiae flp* recombinase gene is under inducible control of the *fac* promoter. Addition of IPTG removes repression by LacI, allowing expression of *flp*, which can then, in theory, 'flip' out the ErmRAM from the chromosome of the first mutant by recombination of its flanking FRT sites.

5.2.3.3 Step 3 – Excision of removable RAM

Plasmid pMTL5402F-FLP was conjugated into the *agrD*1 mutant using the standard conjugation method. Transconjugants were selected for using thiamphenicol. Induction of the expression of the *flp* recombinase gene was then carried out using IPTG as described in 2.11.1. This was permitted by the fact that the gene is under the inducible control of the fac promoter. Colonies were selected for using thiamphenicol and 80 were picked for further investigation. These colonies were streaked in duplicate onto TYG agar containing erythromycin or thiamphenicol. Colonies exhibiting erythromycin sensitivity, combined with thiamphenicol resistance had potentially lost the ErmRAM. However, after overnight incubation, all colonies showed erythromycin resistance, indicating that the ErmRAM had not been lost.

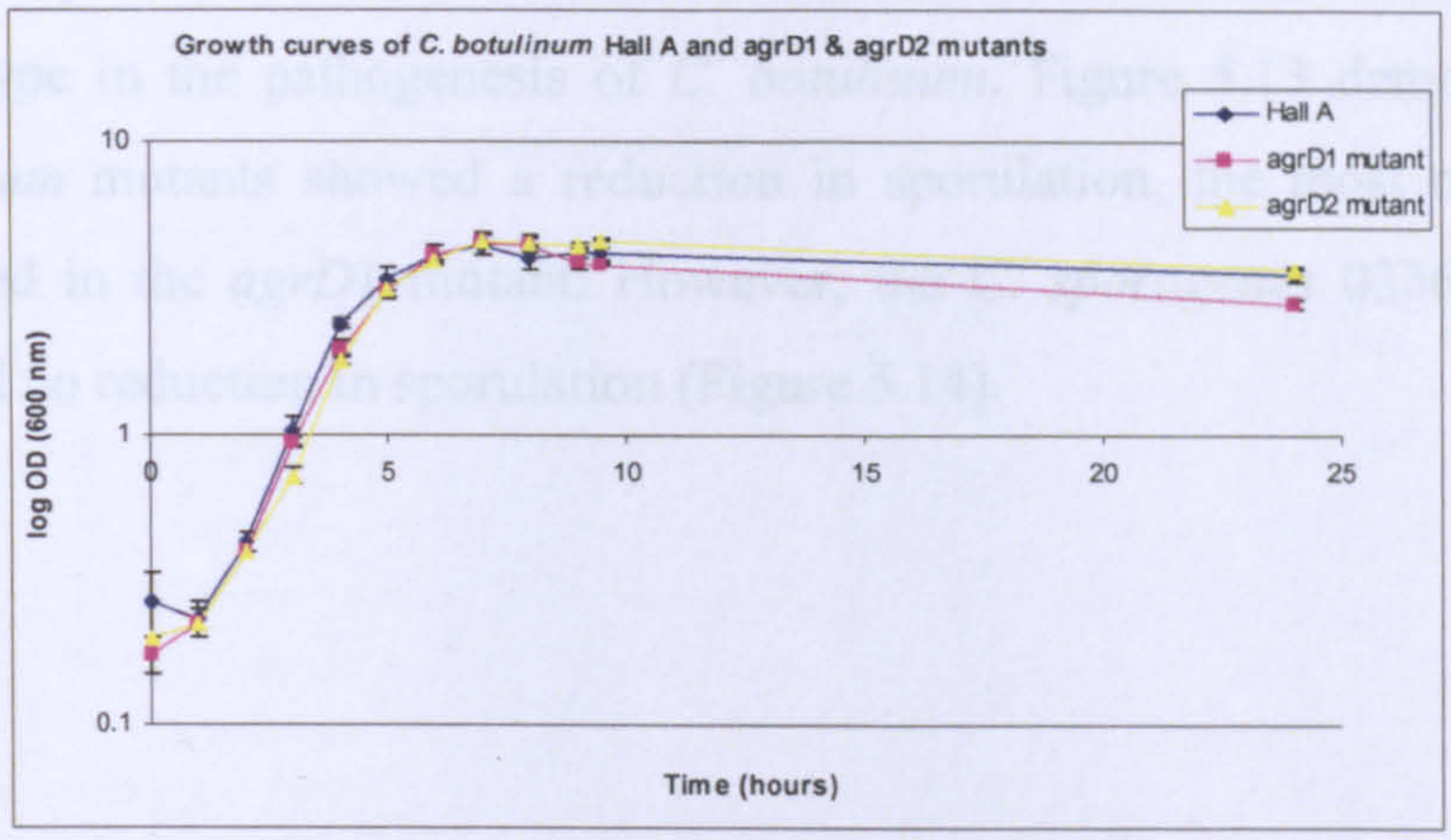
A second attempt at removal of the ErmRAM was therefore made, using an alternative induction method. In the first attempt, induction was carried out for a relatively short period of time (3 hours). In this second attempt induction was carried out over 48 hours as follows:

A 5 ml overnight broth culture of the *agrD*1 mutant containing the pMTL5402F-FLP plasmid was set up. Both thiamphenicol and 4 mM IPTG were added to this culture (previously, only 1 mM IPTG had been used). After 15, 23 and 39 hours, 50 µl broth culture was inoculated into 5 ml fresh TYG broth containing thiamphenicol and 4 mM IPTG. After 48 hours the broth culture was serially diluted to 10^{-7} and 100 µl aliquots of each dilution plated onto TYG agar containing thiamphenicol. After incubation, 105 colonies were streaked in duplicate onto TYG agar containing erythromycin or thiamphenicol as in the previous attempt. Again, after overnight incubation, all isolates exhibited erythromycin resistance indicating that the ErmRAM had not been removed.

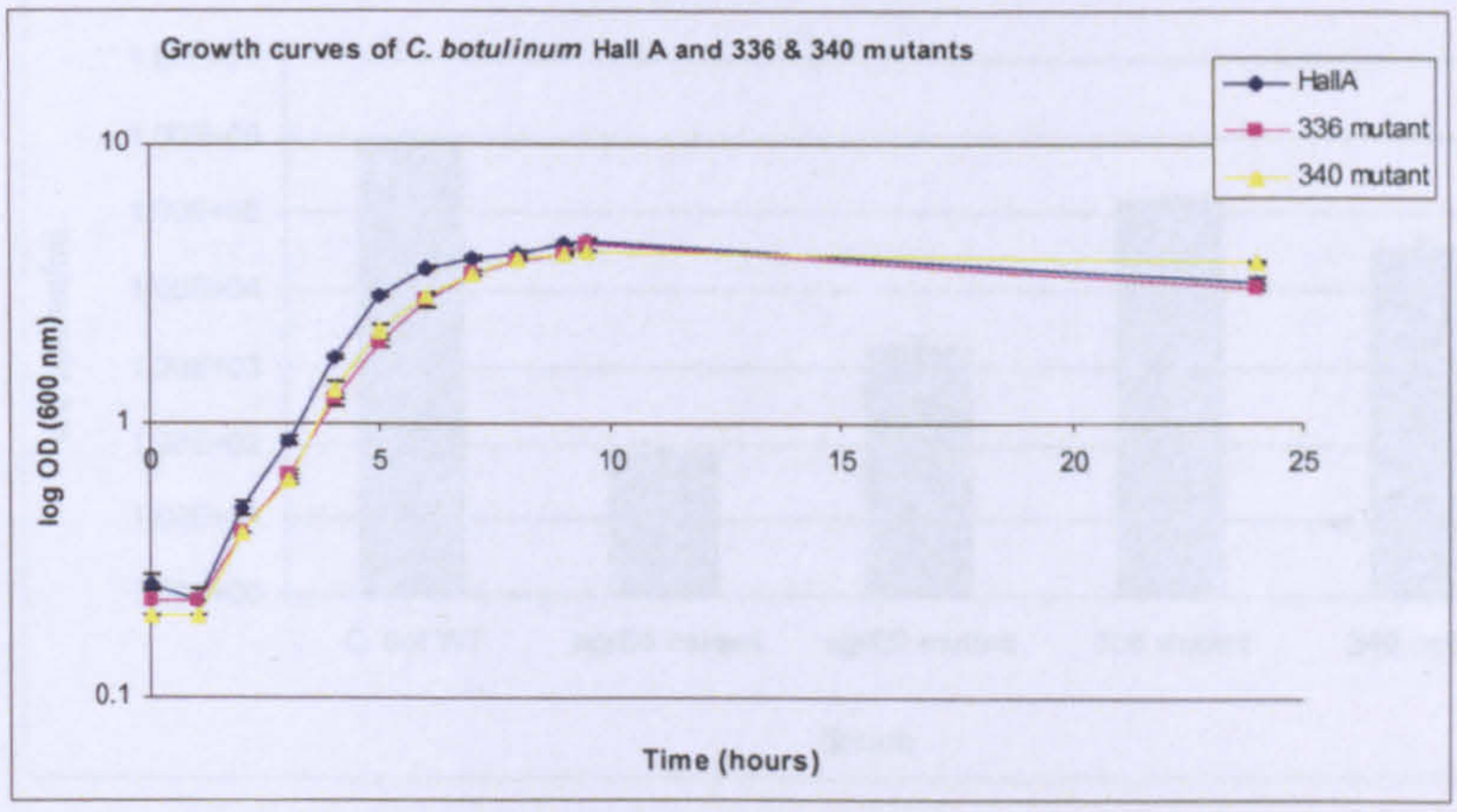
This attempt at marker removal and the subsequent generation of a double knockout mutant was therefore unsuccessful.

5.2.4 Characterisation of clostridial mutants generated using the ClosTron system

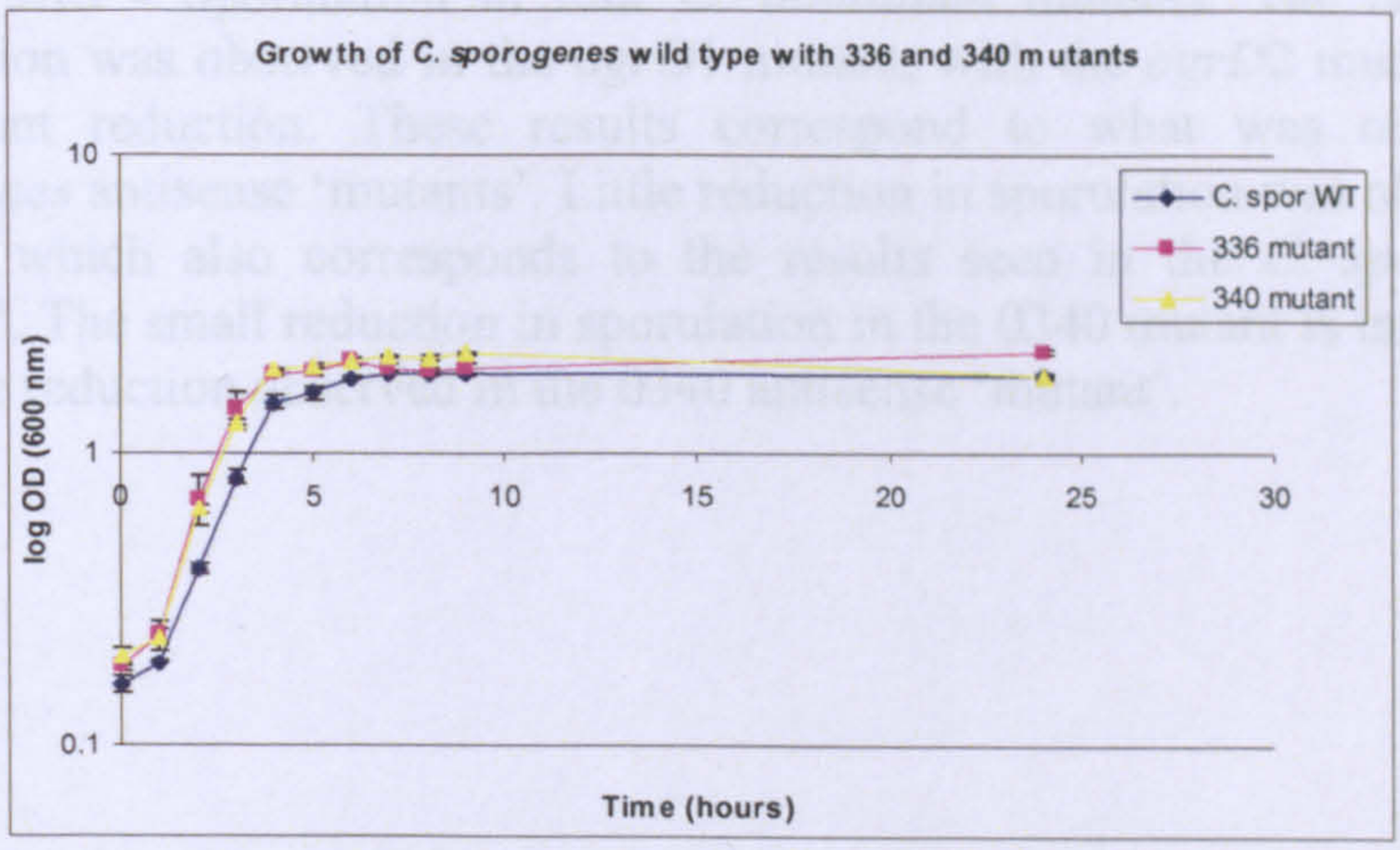
Mutants generated were phenotypically characterised. Firstly, growth, in TYG broth, of all mutants was assessed by following optical density hourly. Growth was not affected in the *C. botulinum agrD*1 and *agrD*2 mutants (Figure 5.12A). The *C. botulinum* 0336 and 0340 mutants showed a reduction in growth rate compared to the wild type, but the final OD obtained was no different (Figure 5.12B). Curiously, the *C. sporogenes* 0336 and 0340 mutants showed a slight increase in growth rate compared to the wild type (Figure 5.12C).



A



B



C

Figure 5.12 – Growth of six clostridial mutants compared to wild type. A – *C. botulinum agrD*1 and *agrD*2 mutants. B – *C. botulinum* 0336 and 0340 mutants. C – *C. sporogenes* 0336 and 0340 mutants.

The ability to sporulate was then assessed in all mutants, due to the importance of this phenotype in the pathogenesis of *C. botulinum*. Figure 5.13 demonstrates that all *C. botulinum* mutants showed a reduction in sporulation, the most marked effect being observed in the *agrD*1 mutant. However, the *C. sporogenes* 0336 and 0340 mutants showed no reduction in sporulation (Figure 5.14).

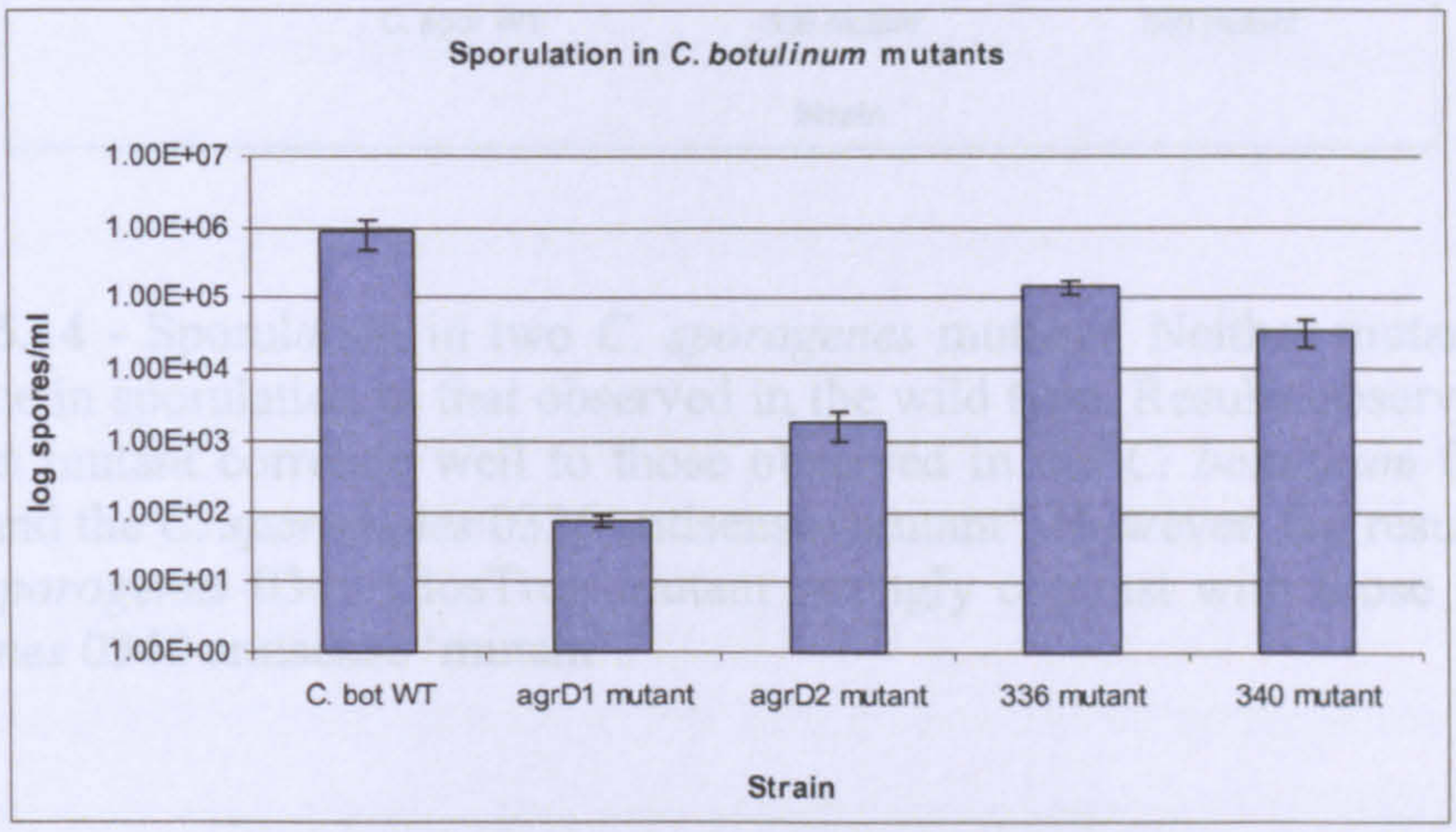


Figure 5.13 - Sporulation in four *C. botulinum* mutants. The largest reduction in sporulation was observed in the *agrD*1 mutant, with the *agrD*2 mutant also showing a significant reduction. These results correspond to what was observed in the *C. sporogenes* antisense 'mutants'. Little reduction in sporulation was observed in the 0336 mutant, which also corresponds to the results seen in the *C. sporogenes* antisense 'mutant'. The small reduction in sporulation in the 0340 mutant is in marked contrast to the huge reduction observed in the 0340 antisense 'mutant'.

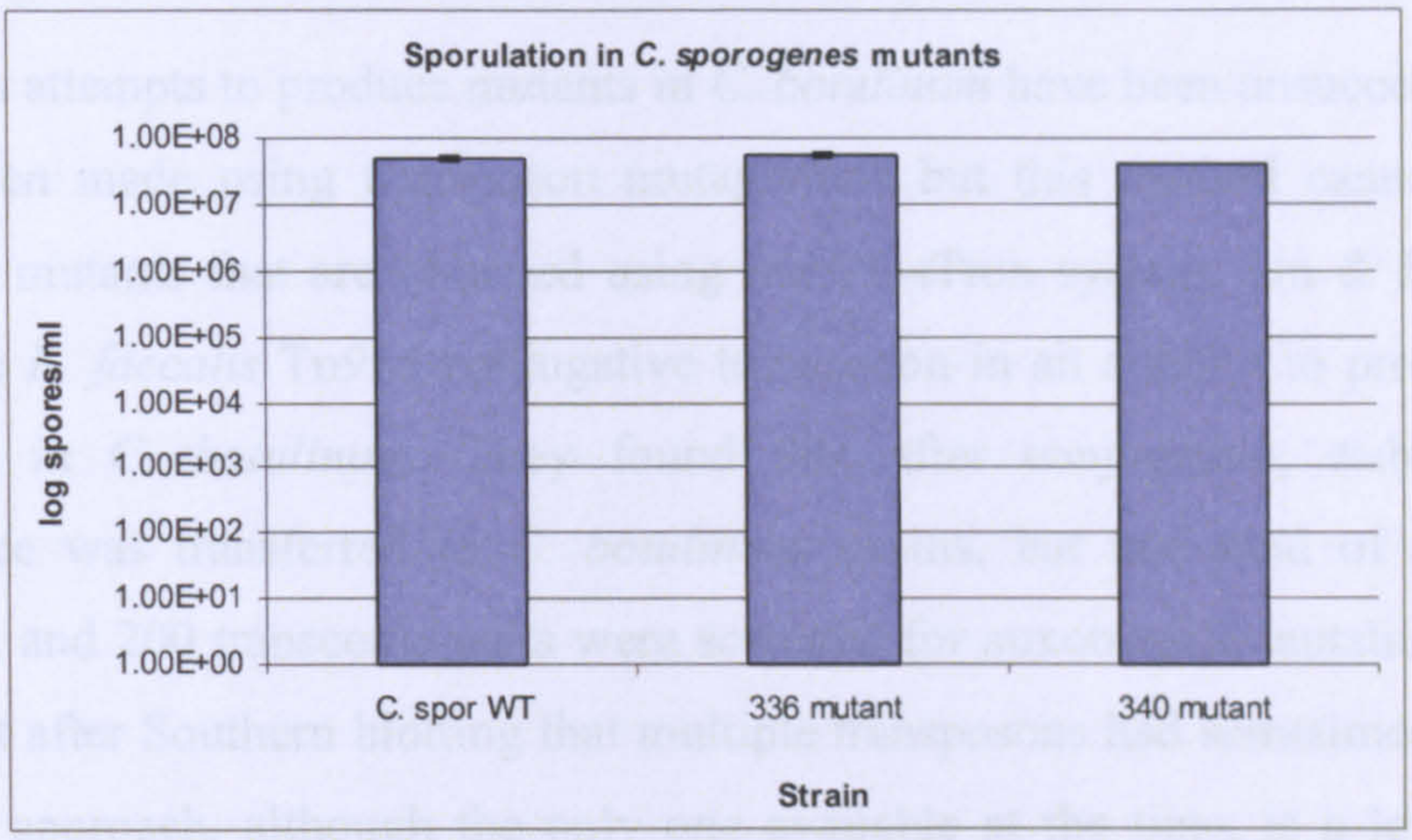


Figure 5.14 - Sporulation in two *C. sporogenes* mutants. Neither mutant showed any difference in sporulation to that observed in the wild type. Results observed in the 0336 ClosTron mutant correlate well to those observed in the *C. botulinum* 0336 ClosTron mutant and the *C. sporogenes* 0336 antisense 'mutant'. However, the results observed in the *C. sporogenes* 0340 ClosTron mutant strongly contrast with those seen in the *C. sporogenes* 0340 antisense 'mutant'.

5.3 Discussion

Previous attempts to produce mutants in *C. botulinum* have been unsuccessful. Attempts have been made using transposon mutagenesis but this method cannot produce the directed mutants that are obtained using the ClosTron system. Lin & Johnson (1991) used the *E. faecalis* Tn*916* conjugative transposon in an attempt to produce knockout mutants in *C. botulinum*. They found that after conjugation, stable tetracycline resistance was transferred to *C. botulinum* strains, but this kind of mutagenesis is random, and 200 transconjugants were screened for auxotrophic mutations. It was also apparent after Southern blotting that multiple transposons had sometimes inserted. This kind of approach, although the only one available at the time, is a long-winded and random way of producing mutants, and has many limitations.

Other groups have been forced to rely on the use of antisense RNA technology for the modulation of gene expression in *C. botulinum*. Marvaud *et al.* (1998b) used this technique in their investigations into the function of *botR/A* in *C. botulinum* toxin regulation. They down-regulated expression of *botR/A* using antisense RNA, but this inhibition was only partial, and one can never be sure that antisense RNA is not affecting the expression of other genes or processes.

There has been one published example of gene knockout in clostridia using the group II intron technology, but this method stopped short of utilising a RAM for selection of mutants. Chen *et al.* (2005) used the *L. lactis* group II intron to disrupt the *plc* gene of *C. perfringens*, but the absence of a RAM necessitated the screening of many colonies in order to detect a mutant. In this instance, screening was made easier by the fact that alpha toxin activity could be screened for by observation of a halo around the colonies on egg yolk agar. However, if other genes were targeted which were not so easily screened, this would mean a very time-consuming PCR screen of many colonies.

This study is the first documented case of directed gene knockout in *C. botulinum* and this method promises to open up many more areas of research into the understanding of gene expression in this organism.

Sporulation in the *C. botulinum agrD* knockout mutants corresponds to that observed in the *C. sporogenes agrB* antisense 'mutants'. If the *agr* systems in *C. botulinum* and *C. sporogenes* function in a similar way to that of *S. aureus*, then essentially, inactivating the *agrB* or the *agrD* gene should have a similar effect; inhibiting the production of the AIP signalling molecule. The *C. sporogenes agrB*1 and *agrB*2 antisense 'mutants' showed a reduction in sporulation of 5.5 logs and 2.5 logs respectively. The *C. botulinum agrD*1 and *agrD*2 mutants showed a reduction in sporulation of 4 logs and 2.5 logs respectively. The differences may be attributed to the experiments being conducted in two different species of clostridia, and two different gene modulation techniques being employed. These data suggest that the *agrB/D* genes are involved in the control of sporulation in *C. botulinum* and *C. sporogenes*, with the *agrB*1/*D*1 genes having the greatest effect. To our knowledge, this is the first time that a link between the *agr* quorum sensing system and the control of sporulation in these organisms has been made.

The spore assay results obtained from the kinase mutants were not so easy to interpret. The *C. sporogenes* 0336 antisense 'mutants' showed no reduction in sporulation, and the same result was observed in the knockout mutant. There was a small reduction in sporulation in the *C. botulinum* knockout mutant, but this difference could be attributed to the mutant being produced in a different species of clostridia.

The spore assay results observed in the 0340 mutants of *C. sporogenes* and *C. botulinum* did not correlate well, suggesting that there may be some other factor affecting the behaviour of these mutants. The *C. sporogenes* 0340 antisense 'mutant' showed more than a 6 log reduction in sporulation compared to the wild type (see Chapter 4). The observed reduction in the *C. botulinum* knockout mutant was only 1.5 logs. This was a huge difference in results, so, in an attempt to confirm which was actually the case, the 0340 gene was knocked out in *C. sporogenes* using the ClosTron. However, rather than confirming one of the previous results, these spore assays produced a third result – no effect on sporulation at all. This then posed the question: are these mutants giving a clear picture of how the 0340 kinase is involved in sporulation in *C. botulinum* and *C. sporogenes*? Clearly the methods used must be called into question, and the most reliable one determined. It could be proposed that antisense modulation is not an ideal method, as discussed in Chapter 4. The fact that a

huge reduction in sporulation was observed in the 0340 antisense 'mutant' might be explained by the antisense fragment possibly affecting other genes which were involved in sporulation, as well as the intended target gene. This may be a contributory factor, but it seems unlikely that such a huge difference in results (i.e. a 6 log reduction, compared to zero reduction) could be caused solely by the non-specific binding of an antisense fragment. The two ClosTron mutants were also produced using two different plasmids containing slightly different ErmRAMs (pMTL007C-E2 contains the ErmRAM flanked by the FRT sites). This may be a contributory factor, along with the fact that the mutants were constructed in two different organisms. However, again, it is unlikely that these theories fully explain the difference.

Another possibility is that the genes targeted by the ClosTron system are not completely inactivated. In theory, knockout mutants give a much 'cleaner' picture of any phenotypic changes than antisense 'mutants', as they completely inactivate the target gene. However mutants obtained using the ClosTron are actually insertional mutants, and when targeting proteins composed of different functional domains, the nature of the insertion of the group II intron may mean that the whole gene is not inactivated.

Histidine kinase proteins are composed of several domains, which carry out different functions. The 0336 and 0340 genes investigated in *C. sporogenes* and *C. botulinum* were found to contain a PAS domain. These domains are sensory modules which typically sense signals in the form of oxygen tension, redox potential or light intensity (Vreede *et al.*, 2003). They also possessed a histidine kinase domain, which is the key element in two-component signal transduction and is responsible for the autophosphorylation of the histidine with ATP, and the phosphotransfer from the kinase to an aspartate residue of a response regulator (Schultz *et al.*, 1998). The domains of the 0336 and 0340 kinase proteins are outlined in Figure 5.15. These domains were analysed using the SMART website (Schultz *et al.*, 1998). In the 0336 mutants, the group II intron was inserted into the histidine kinase domain, which would inactivate this domain, preventing phosphotransfer and therefore subsequent activation of a response regulator. It can be assumed, therefore, that the 0336 genes were effectively inactivated in both *C. botulinum* and *C. sporogenes*.

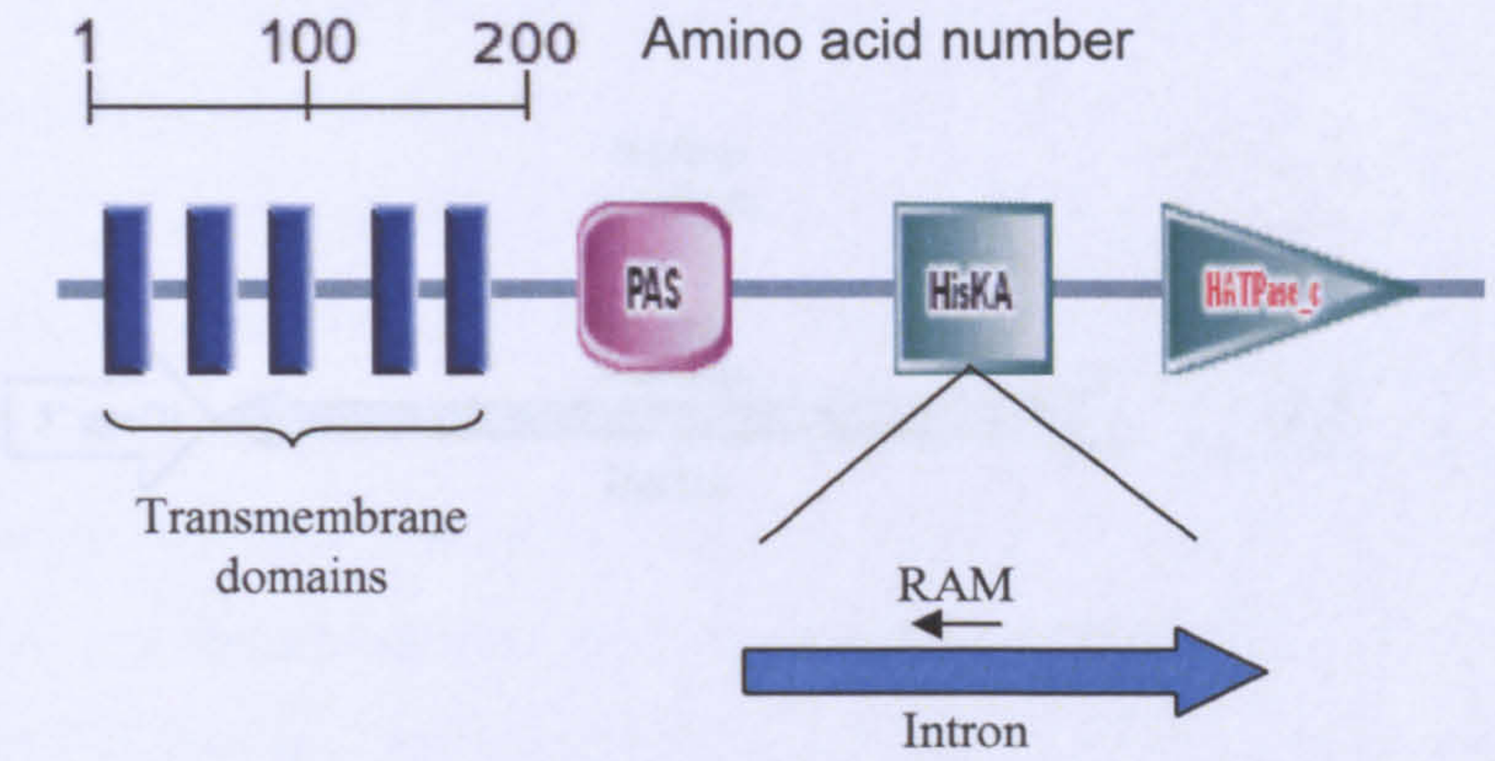
However, in the 0340 mutants, the intron was inserted further upstream in the PAS domain. It is therefore possible that a transcriptional fusion could be produced, which included part of the intron, and the downstream portion of the kinase gene, encompassing the histidine kinase domain. In-frame start codons exist in this fusion, but it is unlikely that any fusion would be transcribed due to the lack of an efficient ribosome binding site. However, if a weak RBS was utilised, a truncated protein could in theory be produced, with a functional histidine kinase domain. It is unlikely that autophosphorylation could occur, due to the insertion of the intron in the signal sensing domain, but in the absence of any clear explanation of these results, this possibility cannot be discounted. The spore assay results on the 0340 mutants may therefore give an inaccurate picture of the effects of inactivating the 0340 gene. Complete inactivation of this gene may have caused a much larger inhibition of sporulation, if this gene is indeed involved in the phosphorylation of Spo0A.

The possibility of transcriptional fusions complicating knockout procedures may also be applicable to the *agrD* ClosTron mutants. The nature of the location of the AIP sequence in the middle of the *agrD* gene means that the AIP may not be completely inactivated. Due to the small size of the AgrD protein, there was little choice of intron insertion site. Indeed, the algorithm only detected one retargeting site in *agrD*2 and two in *agrD*1. The *agrD*1 site selected was chosen because of its lower e value, indicating that there would be a greater chance of successful retargeting. Therefore the sites selected in the *agrD* genes were in fact upstream of the AIP sequence. This is depicted in Figure 5.16. In the *agrD*1 mutant, there is an in-frame methionine start codon 15 residues upstream of the AIP sequence. Again, there is no clear ribosome binding site (the sequence 9 bp upstream of the start codon is ATCTGGGC), but if there were to be weak binding of the ribosome to this sequence, then in theory, an N-terminally truncated *agrD* gene, containing the AIP sequence may be translated, allowing some *agr* signalling to occur. The strong thiolase promoter of the ErmRAM may also be a contributory factor, allowing transcription of the fused intron and *agrD* genes. This may explain why sporulation still occurred in the mutant (which showed approximately 10^2 spores/ml). It is therefore unknown whether the *agrD*1 gene was completely inactivated, and whether a greater decrease in sporulation may have been observed in an absolute knockout.

However, it is highly unlikely that this would occur, as the N-terminal of the AgrD protein has been shown to be vital for processing into the mature AIP to occur (Zhang *et al.*, 2004). The AgrD protein was found to integrate into the cytoplasmic membrane via an α-helical amphipathic motif in its N-terminal region, which was required for stabilisation of AgrD, in order for processing by the membrane bound AgrB to occur. This region would be missing in the truncated protein, and processing of the AgrD propeptide would therefore almost certainly not occur.

The transcriptional fusion issue was not encountered with the *agrD*2 mutant, as no start codon was present between the final stop codon and the AIP sequence, preventing any such transcriptional fusion from occurring.

***C. botulinum* 0336**



***C. botulinum* 0340**

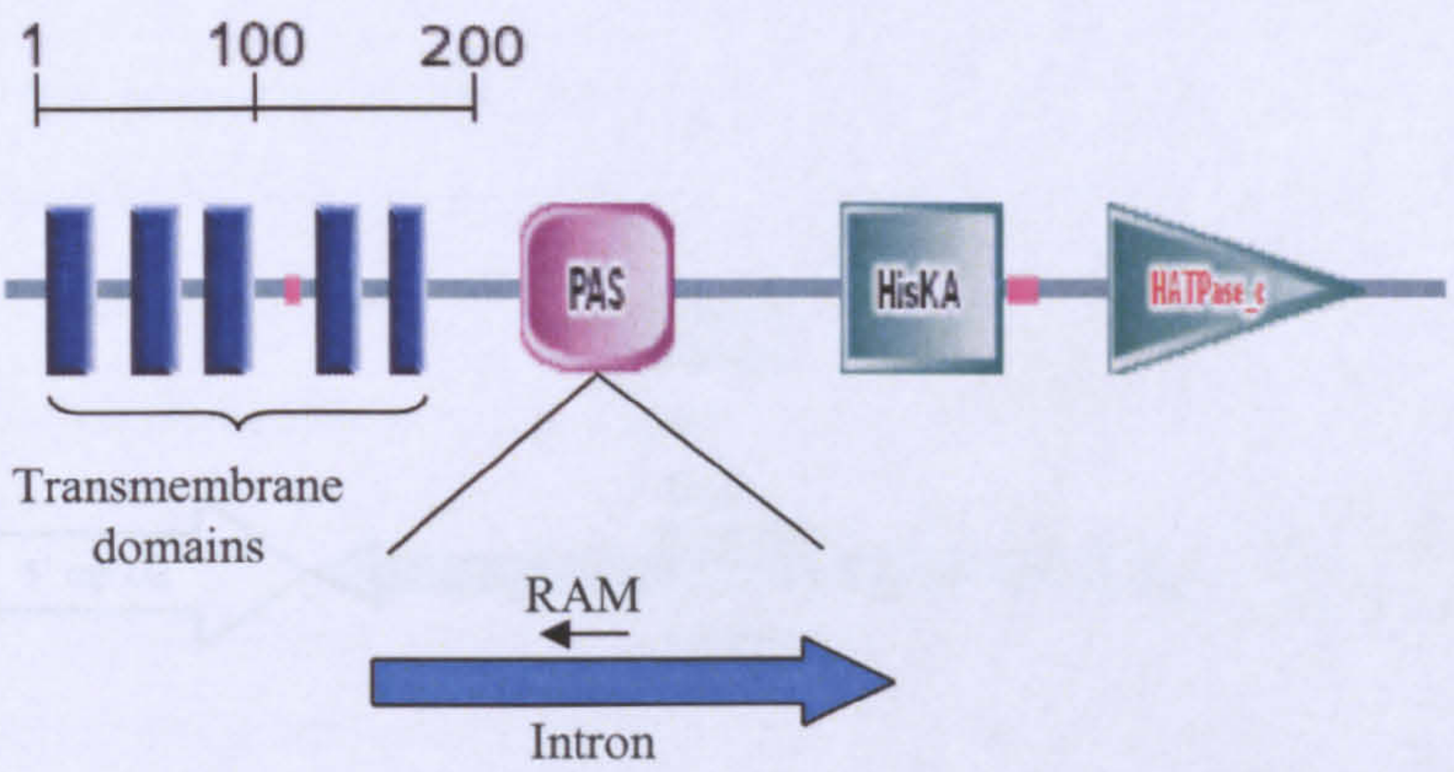
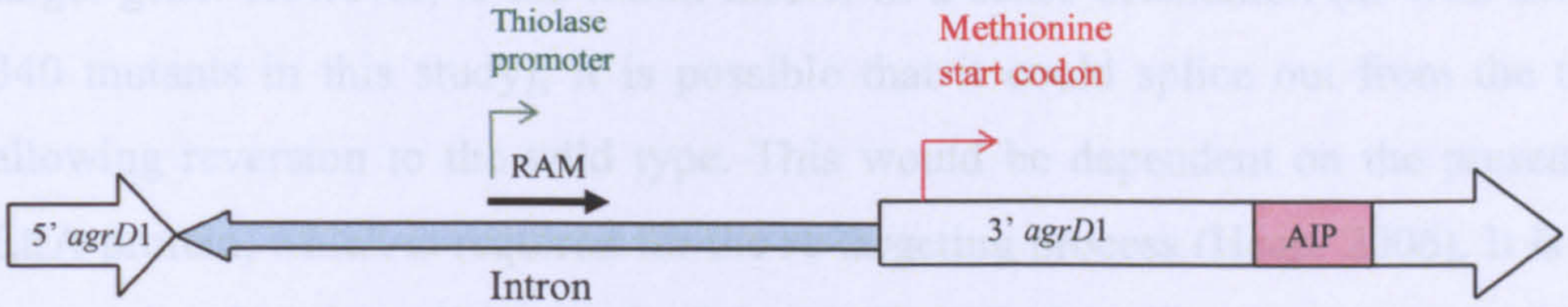


Figure 5.15 – Functional domains of the *C. botulinum* kinase proteins 0336 and 0340, as analysed on the SMART website (http://smart.embl-heidelberg.de/). The position of the intron insertion is indicated by the blue arrow. In CBO0336 the intron inserts into the HisKA (histidine kinase) domain. In CBO0340 the intron inserts into the PAS domain. Green triangles indicate histidine kinase-like ATPase domains. Small pink regions indicate segments of low compositional complexity.

C. botulinum agrD1



C. botulinum agrD2

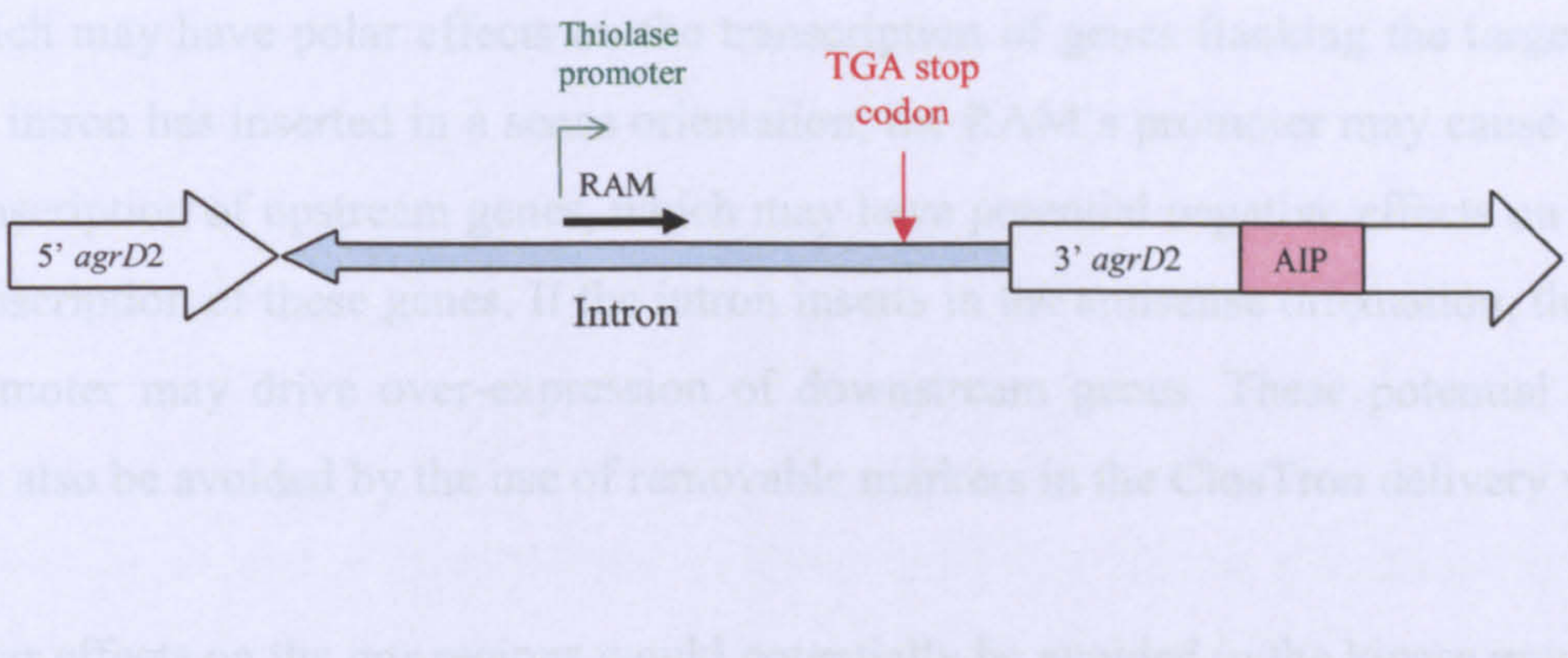


Figure 5.16 - Schematic diagram depicting the insertion site of the group II intron in the *C. botulinum agrD* mutants, along with the possible transcriptional fusions which may occur. In the *agrD*1 mutant, a transcriptional fusion may occur between part of the intron and the 3' segment of the *agrD*1 gene. Transcription may be driven from the strong thiolase promoter of the RAM. A methionine start codon found within the fusion may result in translation of the fusion, incorporating the AIP region. This should not occur in the *agrD*2 mutant as there is a stop codon downstream of the thiolase promoter which would prevent translation of the truncated *agrD*2 sequence.

There are also other issues which should be borne in mind when using the ClosTron system. When an intron is inserted into a target gene in the antisense orientation (as with the *agrD* mutants in this study), it is unable to splice out from the sense transcript of the target gene. However, if the intron inserts in a sense orientation (as with the 336 and 340 mutants in this study), it is possible that it could splice out from the transcript, allowing reversion to the wild type. This would be dependent on the presence of the LtrA protein, which is required for the re-targeting process (Heap, 2006). It is therefore imperative that the cell is cured of the pMTL007 plasmid as soon as possible, as this ensures the absence of the LtrA protein, preventing any further, unwanted, mobilisation of the intron. This would be a potential problem when introducing a second plasmid containing the LtrA protein during the production of a double mutant. However, this problem is prevented by rapid curing of the first plasmid, as once the intron has inserted into the chromosomal DNA, it cannot be spliced out. It would also be avoided by the use of a removable ErmRAM as in plasmid pMTL007C-E2.

Another potential limitation of the system is the possible polar effects caused by the insertion of a RAM (Heap 2006). The ErmRAM possesses its own, strong, promoter which may have polar effects on the transcription of genes flanking the target gene. If the intron has inserted in a sense orientation, the RAM's promoter may cause antisense transcription of upstream genes, which may have potential negative effects on the sense transcription of these genes. If the intron inserts in the antisense orientation, the RAM's promoter may drive over-expression of downstream genes. These potential problems can also be avoided by the use of removable markers in the ClosTron delivery vectors.

Polar effects on the *agr* regions would potentially be avoided in the kinase mutants, as it has been shown in Chapter 3 that these genes are not transcriptionally linked to the *agr* region genes. There may therefore be transcriptional terminator sequences between the kinase genes and the *agr* genes. Again, in the *agrD*2 mutant, downstream polar effects may also be avoided by the fact that the gene immediately downstream of *agrD*2 is kinase 0340, which is not transcriptionally linked and hopefully separated by a terminator sequence. However, polar effects may possibly occur in the *agrD*1 mutant, as the gene immediately downstream (0333) is transcriptionally linked, and may suffer over-expression due to the strong thiolase promoter of the ErmRAM.

In summary, clearly this method of gene inactivation is not perfect, and several factors must be considered when choosing a target site, and interpreting results of phenotypic analyses. The data obtained in this study would indicate that the *agrD* genes are involved in the control of sporulation in *C. botulinum*, but the magnitude of the effect of *agrD*1 cannot be fully assessed due to the possible transcriptional fusion which may result in the production of low levels of AIP in this mutant. It would also appear that the 0336 kinase protein is not involved in the initiation of sporulation in *C. botulium* and *C. sporogenes*. However, it is difficult to draw any definite conclusions on the role of the 0340 gene. It would appear that it is involved to a certain degree in sporulation, according to the *C. botulinum* results, but this gene may not have been completely inactivated in either species, leading to difficulties when interpreting the two sets of spore assay results.

Despite these drawbacks, ClosTron technology is in its infancy, and has proved to be of ground-breaking importance as a clostridial knockout tool. It has enabled the relatively straight-forward generation of four directed insertional knockout mutants in *C. botulinum*, which has never before been achieved. The technique will also enable a wealth of knowledge to be obtained on the molecular mechanisms of gene expression in clostridia, which has for so long eluded the clostridial scientific community.

5.3.1 Future Work

Future developments in ClosTron technology will enable the inactivation of multiple target genes, and will avoid the problems of polar effects discussed earlier. This methodology was first documented in *E. coli* by Zhong *et al.* (2003), who created an *E. coli lacZ/argF* double mutant by removing the first RAM using the Flp recombinase enzyme. The pMTL007C-E2 plasmid has been constructed for this purpose (J. T. Heap, unpublished work) but, to date, attempts at multiple gene knockout using this system have been unsuccessful. The only clostridial species in which the RAM has been successfully removed is *C. acetobutylicum*. Jamie Scott (personal communication) at the University of Nottingham has removed the RAM from mutants created using the ClosTron. Saint-Prix *et al.* (2006) also removed a marker from a mutant constructed using an unstable plasmid. In both cases, the plasmid used to express the Flp protein was one constructed by the group of Philippe Soucaille (Saint-Prix *et al.*, 2006). This

would indicate that it is indeed possible to remove the RAM from a ClosTron mutant, and the problem encountered in this study lies with the Flp expression vector.

Although the Flp recombinase enzyme has been successfully used for marker removal in *C. acetobutylicum*, an attempt to use this method to construct a *C. botulinum* *agrD*1/*D*2 double mutant in this study was unsuccessful. This was due to the inability to remove the RAM from the *agrD*1 mutant. The modified ErmRAM in the pMTL007C-E2 plasmid was flanked by FRT sites, which should be recognised by the Flp recombinase enzyme. However, on introduction of plasmid pMTL5402F-FLP, and subsequent induction of *flp* expression, the RAM remained *in situ*. A factor affecting this result may be the thermolability of the Flp protein. It has been shown to function maximally at 35°C, and not at all at temperatures exceeding 39°C (Buchholz *et al.*, 1996). Indeed the successful removal of the RAM by Jamie Scott (personal communication) was carried out at 30°C.

Buchholz *et al.*, (1996) also concluded that the Flp:substrate ratio played a major role in the efficiency of the enzyme, as did levels of its expression. Neither of these factors was tightly regulated in the attempts to remove the ErmRAM in this study. An alternative may be to use a different recombinase protein such as Cre from bacteriophage P1. This protein was shown to function most efficiently at and above 37°C and is therefore suitable for use in hosts which grow at these temperatures (Buchholz *et al.*, 1996). The same group proposed that Cre may function more efficiently than Flp in *E. coli* at low expression levels.

Clearly further developmental research must be carried out in order to create an efficient marker removal system to be used in conjunction with the ClosTron. The efficiency of marker removal may possibly be improved by using a plasmid which expresses the Flp protein constitutively, which would avoid any problems encountered with induction times and IPTG concentrations.

An alternative gene knockout technique has recently been developed, which may be of relevance to this study. Suzuki *et al.* (2005) developed a system for gene knockout in *Corynebacterium glutamicum* which combined the marker removal principle using a

recombinase enzyme, with a method of DNA removal from the chromosome using the meganuclease I-*Sce*I. Bacterial genomes do not naturally contain I-*Sce*I sites and this nuclease has also been successfully used in *E. coli* to produce knockout mutants induced by homologous recombination (Posfai *et al.*, 1999).

Suzuki *et al.* (2005) utilised this system to knock out a region of DNA totalling 67 Kb. A plasmid containing two Cre recombinase recognition sites and two I-*Sce*I recognition sites was first introduced into the chromosome by homologous recombination. On introduction of a plasmid expressing Cre and I-*Sce*I, a large segment of chromosomal DNA was first excised by recombination of the Cre sites. This left the remaining segment of target DNA flanked by I-*Sce*I sites, which was then cut by the I-*Sce*I enzyme, resulting in the excision of a large portion of DNA. The fact that all recognition sites were removed during the process meant that the method could be used repeatedly to remove subsequent portions of DNA. This principal could ideally be used to produce knockout mutants in clostridia, where the target gene was not merely interrupted, but entirely removed. This would circumvent problems encountered with polar effects and incomplete inactivation of the target gene, and would also allow the investigation into the ability of the 0340 kinase mutants to initiate sporulation. The whole kinase gene should ideally be removed using the above method, and the downstream portion containing only the histidine kinase domain inserted in its place. Subsequent spore assays would confirm whether that portion of the protein is in fact able to initiate sporulation, and would ascertain whether the 0340 mutants constructed in this study were 'true' mutants. This is clearly a technology for the future of clostridial gene knockout.

Finally, it would be of great interest to investigate the ability of the *C. botulinum* mutants to produce neurotoxin. Along with sporulation, this is a major virulence factor of this organism. Toxin is produced temporally throughout growth, but reaches maximal levels towards late exponential growth (Rao *et al.*, 2007). Considering the *agr* system in *S. aureus* is involved in the control of toxin production, it is a distinct possibility that this is also the case in *C. botulinum*. Indeed Sebaihia *et al.* (2007) postulated that, since maximal neurotoxin production takes place in late exponential and early stationary growth, 'it seems likely that neurotoxin gene expression may be subject to quorum sensing'.

Another area of investigation should centre round the relationship between toxin production and sporulation. In some clostridial species, for example *C. perfringens*, toxins are produced in association with sporulation (Mitchell, 2001). However, it has been proposed that, in *C. botulinum*, the highest levels of toxin production were generally found in cells undergoing rapid autolysis, which did not sporulate (Davis *et al.*, 2000; Schantz & Johnson, 1992). Furthermore, Roa *et al.* (2007) also demonstrated that, although type A toxin production was maximal towards late exponential growth, cells were yet to sporulate at that time point. This suggests that in *C. botulinum*, sporulation is not a prerequisite for toxin production, and there is a possibility that the cells either carry out one of these processes or the other. Indeed, in *B. anthracis* it has been demonstrated that when cells were cultured in toxin-inducing conditions, sporulation efficiency was significantly reduced in comparison to *B. subtilis* under the same conditions (Hadjifrangiskou *et al.*, 2007). In addition, Takumi *et al.* (1980) isolated non-toxic variants of type A *C. botulinum* and demonstrated that these exhibited enhanced sporulation. It would therefore be interesting to observe whether the *C. botulinum* mutants in this study, which displayed a reduced ability to sporulate, produced toxin at higher levels than the wild type. This remains to be investigated, as facilities for a neurotoxin assay were not available at the University of Nottingham during this project.

Currently, methods employed to assess toxin production are Real Time RT-PCR and the mouse bioassay, both of which are time-consuming and expensive. The Clostridial Research Group at the University of Nottingham is currently establishing an ELISA assay which will give a rapid and quantitative measure of neurotoxin production.

Chapter 6.0

General Discussion

Chapter 6 – General Discussion

6.1 Introduction

The role of quorum sensing in the control of virulence factor production in *C. botulinum* clearly needs to be elucidated. Due to the serious nature of the disease caused by this organism, it is imperative to understand how its virulence is controlled, in order to develop suitable preventative measures. The aims of this study were to establish that a functional *agr* system exists in *C. botulinum* and to investigate the nature of its mechanism of action. In order to determine which processes the system controlled, an effective gene inactivation tool was sought.

6.2 A functional *agr* system in *C. botulinum* and *C. sporogenes*

Before the commencement of this study, there was no published evidence to suggest that *C. botulinum* possessed a functional *agr* system. The genome sequence clearly confirmed the presence of two *agr* regions but it was unknown whether they were actually expressed, and what their mechanism of action was.

Sequencing of the *C. sporogenes agr* regions, along with subsequent analysis revealed that the genes of the *agr* systems were well conserved in the clostridia included in the analysis, particularly between *C. botulinum*, *C. sporogenes* and *C. perfringens*. Some similarities also existed with the staphylococcal *agr* genes, particularly *agrB* and *agrD*. Alignments of AgrD proteins from *S. aureus* and various clostridial species revealed that a putative AIP was present in the clostridia, which showed some degree of similarity to the AIPs produced by the four *S. aureus* groups. Most importantly, all of the clostridial proteins included in the alignment showed the presence of the conserved cysteine residue, known to be essential for formation of a functional AIP in *S. aureus* (Zhang *et al.*, 2002). There is also some evidence to suggest that putative processing sites are present on either side of the clostridial AIP sequences. This was the first line of evidence to suggest that *C. botulinum* and *C. sporogenes* do indeed possess *agrD* genes

which could potentially produce a functional AIP. However, since the AgrD and AgrB proteins in the two *agr* regions of *C. botulinum* and *C. sporogenes* were quite different (possessing approximately 40% protein identity) this suggested that the two *agr* systems could function independently, and may have evolved separately. The fact that there are two separate systems present renders the elucidation of their function more complex.

Although comprehensive sequence alignments and BLAST searches were carried out, a definitive identification of *agrA* and *agrC* genes in *C. botulinum* and *C. sporogenes* was not achieved. This could be because the *agr* systems do not directly utilise a two component system, but it is also possible that signal transduction takes place via a slightly different mechanism to that observed in *S. aureus*. Indeed, there are several genes, whose allocated function was related to signal transduction, found in close proximity to the *agr* genes. However, their involvement in the *C. botulinum agr* system was not elucidated in this study.

The second line of evidence to support the theory of a functional *agr* system in *C. botulinum* and *C. sporogenes* was obtained from transcriptional studies of the *C. sporogenes agr* systems. RT-PCR experiments indicated that the genes of the *agr* system were indeed expressed. This had not been previously confirmed and added evidence to the theory that this system is indeed functional. The genes of *agr* region 1 were shown to be transcribed together as an operon. This corresponds to what has been previously demonstrated in *S. aureus*, where the *agr* genes are transcribed together from the P2 promoter (Novick *et al.*, 1995). However, the *agr* genes in region 2 of *C. sporogenes* were not found to be transcribed together. The two orphan kinase genes flanking the region appeared to be transcribed separately from the *agr* genes. This would suggest that these kinase genes are not involved in the *agr* response, and indeed, later studies (see Chapters 4 and 5) suggested that these genes could in fact be involved in the control of sporulation, and might be implicated in the equivalent of the phosphorelay system of *Bacillus* species.

Finally, Real Time RT-PCR was used to demonstrate that the *agrD* genes were maximally expressed towards late exponential growth. This also corresponds with what has been previously demonstrated in *S. aureus*, where the *agr* system is maximally expressed towards late exponential to early stationary growth phase (Yarwood &

Schlievert, 2003). Problems were, however, encountered with the Real-Time PCR experiments, which were likely due to the high AT content of clostridial DNA. This necessitated the use of the Absolute Quantification method, rather than the Relative Quantification method, but this has also been used by other groups attempting Real Time PCR on clostridial DNA (Couesnon *et al.*, 2006; Kouguchi *et al.*, 2006).

The data obtained from this study gave clear evidence for functional *agr* systems in *C. botulinum* and *C. sporogenes*. This evidence then needed to be taken further in order to ascertain which functions these *agr* systems controlled.

6.3 Gene modulation techniques for determining *agr* function

In order to assess the function of the *agr* genes, it was necessary to inactivate some of them, allowing the appropriate phenotypic tests to be carried out. Historically, the inactivation of genes in the clostridia has been extremely problematic, with only the occasional report of such an achievement. Alternative strategies were therefore sought. The first of these was the construction of an integrative vector, which would introduce a mutant copy of a target gene in the form of an unstable single stranded DNA species. This would potentially integrate into the host chromosome by homologous recombination.

Several problems were encountered with this methodology. Firstly, it proved impossible to conjugate the plasmid into the target organism – *C. sporogenes*. This was possibly because of the highly unstable nature of the plasmid. Eventually, conjugation into an uncharacterised mutant of *C. sporogenes* proved successful, but was not ideal, as the mutant strain appeared to stabilise the plasmid, which conflicted with the overall strategy of using an unstable plasmid. After introducing a knockout cassette into the integrative vector and attempting to induce recombination with the chromosome, the nature of the recombination event which had occurred could not be determined. This confirms that the use of these types of vectors for gene inactivation is not ideal, as it is almost impossible to predict the recombination events which may occur. It has been previously shown that the occurrence of recombination events is highly unpredictable (Gruss & Ehrlich, 1989). The fact that the mutant strain of *C. sporogenes* used was

uncharacterised also contributed to the complexity of the methodology, as it is possible that some unknown mutation in the strain was responsible for the unsuccessful recombination. Few groups using unstable plasmids for the purpose of gene knockout have induced predictable recombination events. In summary, this method of gene inactivation is not the method of choice and in this instance was unsuccessful.

Another method of gene modulation employed was the use of antisense RNA. This method has its advantages in that it does not give absolute inactivation of the target gene; protein production is merely down-regulated. This is ideal for the study of essential genes. Another advantage to its use was exploited in this study, where an inducible expression vector was used for antisense delivery. This means that the time at which antisense production is induced can be controlled to a certain degree, although it was shown that some antisense expression was still allowed to occur in the absence of induction, due to the use of a 'leaky' promoter in the vector.

Disadvantages of this technique include the difficulty in interpreting phenotypic assay data. The fact that the target gene is only partially inactivated means that the full magnitude of its inactivation cannot be assessed. An absolutely quantitative result cannot be gained, as there is still some target gene expression occurring. In order to obtain a quantitative result from phenotypic tests, the gene would need to be knocked out. A further disadvantage to the technique is that it cannot be confirmed whether an antisense fragment is only acting on the intended target gene. Indeed, in this study, genes were targeted which show significant regions of homology with other genes in the *agr* region. For example the antisense fragment designed to target the *agrB*1 gene in *C. sporogenes* could also have been acting on *agrB*2. This could also be the case with the two orphan kinase genes (0336 and 0340). Again, this should be borne in mind when interpreting phenotypic assay results, and ideally, a targeted knockout mutant should be obtained.

Nevertheless, the construction of antisense 'mutants' gave extremely important preliminary data, which suggested that both the *agrB* genes and the 0340 orphan kinase gene were involved in the control of sporulation in *C. sporogenes*. Previously there had been no evidence to suggest a link between the *agr* system and sporulation in clostridia.

6.4 Construction of targeted knockout mutants for determining function of the *agr* genes

This study has produced the first ever documented targeted knockout mutants in *C. botulinum* using the newly-developed ClosTron system. This technology has revolutionised clostridial gene manipulation methodology and has the potential to be used in any clostridial species.

As the *agr* system in *S. aureus* has been shown to control the expression of a number of secreted virulence factors, such as haemolysins, toxic shock syndrome toxin (Otto *et al.*, 1998) and coagulase (Novick, 2003), it would seem appropriate to assay for secreted proteins in the *C. botulinum* and *C. sporogenes* mutants constructed. Collagenase assays carried out on the *C. sporogenes* antisense mutants failed to yield any definitive conclusions. Furthermore, the facilities to assay botulinum neurotoxin were not available during this project. It was therefore decided, as an alternative, to investigate sporulation. Although this was not the ideal phenotype to test for, considering it was not connected with a secreted virulence factor, the assays did in fact provide some hugely important data on the control of sporulation in *C. botulinum* and *C. sporogenes*. The knockout mutants produced have allowed a link between the *agr* system and sporulation to be established for the first time, and have indicated that both of the *agrD* genes in *C. botulinum* are involved in this process.

However, this study has also highlighted that the ClosTron system is not perfect, and that the nature of any target genes should be fully assessed before attempting to inactivate them using this methodology. The orphan kinase proteins targeted in this study exhibit a number of different functional domains. It became apparent, after mutant generation, that the position of these domains in relation to the intron insertion site may have an effect on the efficiency of inactivation of the gene. It is possible that some of the domains of the protein may still be synthsised and hence be functional. In the future generation of such mutants, this must be investigated before commencement of the procedure.

This issue is also relevant to proteins which are post-translationally processed, which is the case with many signalling molecules utilised by Gram-positive bacteria (Fuqua &

Greenberg, 1998). In the construction of the *agrD*1 mutant, the group II intron was inserted upstream of the AIP sequence, which allowed the possibility of some low levels of transcription of the AIP sequence, resulting in incomplete inactivation of the *agrD*1 gene.

Despite these drawbacks, the data obtained from the mutants generated in this study has indicated that the *agr* system is involved in the control of sporulation in *C. botulinum*, which has not previously been described.

Data obtained from the kinase mutants has indicated that the 0336 kinase is not involved in the control of sporulation in *C. botulinum* or *C. sporogenes*. It would appear that the 0340 kinase gene may be involved, but the degree of involvement could not be determined for certain, as it could not be confirmed that the complete inactivation of these genes was in fact achieved in the generation of the insertion mutants.

6.5 Summary and future work

This study has allowed the characterisation of an *agr* system in *C. botulinum* and *C. sporogenes*, of which very little was previously known. It has indicated that one of the cellular functions controlled by this system is sporulation – an essential survival mechanism in this organism. This may contribute to the future development of control measures against foodborne botulism.

Results obtained in this study also indicate high similarities between the *agr* systems in *C. botulinum*, *C. sporogenes* and *S. aureus*, including the timing of the expression of their genes. This may suggest that virulence factor production is controlled in a similar manner in all three organisms, although a two-component system has yet to be identified in *C. botulinum* and *C. sporogenes*.

There is huge potential for future work on this project. It has yet to be confirmed that a diffusible signal molecule is utilised by these *agr* systems. In order to confirm this, the peptide should be purified and further characterised to compare it to those produced in *S. aureus* and other Gram-positive bacteria. This has recently been undertaken in

Streptococcus thermophilus by Ibrahim *et al.* (2007). This group have begun the characterisation of a novel quorum sensing system and subjected *S. thermophilus* culture supernatant to HPLC analysis in order to search for signalling molecules. Sequencing and mass spectrometry was used to further identify these peptides. They focussed their study on a new modified peptide termed Pep1357C and found that the production of this signalling molecule required the activity of both a transcriptional regulator and an oligopeptide transport system. This is clearly the next stage in the investigation into the *agr* system in *C. botulinum*. Either natural or synthetic peptides would ideally then be used in complementation experiments to demonstrate the restoration of sporulation to wild type levels in the knockout mutants.

The addition of wild type supernatant to the *agrB* antisense 'mutants' failed to restore sporulation in these strains, but as discussed previously, antisense mutants are not simple mutants and a much cleaner picture would probably be obtained using knockout mutants.

Since a link between the *agr* system and sporulation has been demonstrated, it seems likely that other late exponential phase processes may also be under the control of this system. In order to confirm this, microarray analysis should be carried out. This may also indicate how toxin gene expression is affected by knocking out the *agr* system. Along with sporulation, toxin production is an important virulence factor in *C. botulinum*, and the inhibition of its production could also lead to the development of effective control measures for use in the food industry. Toxin production in many clostridia is maximal during late exponential to early stationary growth (Dupuy & Sonenshein, 1998; McClane & Rood, 2001). Since the *agr* system has been shown to be maximally expressed during this phase, it is possible that it is involved in the control of *C. botulinum* toxin production. Microarray data would also indicate whether expression of the *botR* gene is affected by knocking out the *agr* system. This sigma factor has been previously shown to up-regulate toxin expression in *C. botulinum* (Marvaud *et al.*, 1998b).

It is not known what effect inactivation of the *agr* system may have on toxin production in *C. botulinum*. It has been shown in some clostridia that toxin expression is sporulation-dependent (Durre & Hollergschwandner, 2004) indicating that a reduction

in sporulation in the *agrD* mutants may result in a reduction in toxin production. However, in *B. anthracis*, it has recently been demonstrated that, when this organism is cultured in toxin-inducing conditions, a reduction in sporulation is observed compared to *B. subtilis* under the same conditions (Hadjifrangiskou *et al.*, 2007). This would indicate that this organism makes the decision to either sporulate or produce toxin. The explanation offered was that the bacterium may maximise toxin production when needed (i.e. during infection in the case of *B. anthracis*), whilst maintaining the ability to sporulate if a change in environmental conditions necessitates it (Hadjifrangiskou *et al.*, 2007). It can therefore not be predicted which would be the case in *C. botulinum*.

To confirm for certain whether inactivating the *agr* system resulted in a change in neurotoxin expression in *C. botulinum*, toxin assays should be carried out on the *agrD* mutants. Unfortunately, facilities for carrying out such an assay were not available at the University of Nottingham until after the completion of this project.

Another obvious area of this project to be continued is the construction of an *agrD* double mutant in *C. botulinum*. This would indicate whether the *agr* system is essential to the cell, and would provide interesting information on whether the complete knockout of both systems resulted in the total abolition of sporulation and/or toxin production. Complementation vectors could also be introduced into all mutants constructed in an attempt to restore a wild type phenotype.

Finally, there is huge scope for the investigation of the mechanism of sporulation initiation in *C. botulinum*. Initial experiments in this study have indicated that kinase 0340 may be involved in this process in *C. botulinum* and *C. sporogenes*, but the nature of the mutants produced needs to be confirmed, in order to ascertain whether this target gene was truly inactivated. As discussed in Chapter 5, I-*Sce*I technology could be employed to completely remove this gene, giving a clearer picture of its function.

Bibliography

Bibliography

Abraham, W. R. (2006). Controlling biofilms of gram-positive pathogenic bacteria. *Curr Med Chem* **13**, 1509-1524.

Akbulut, D., Grant, K. A. & McLauchlin, J. (2005). Improvement in laboratory diagnosis of wound botulism and tetanus among injecting illicit-drug users by use of real-time PCR assays for neurotoxin gene fragments. *J Clin Microbiol* **43**, 4342-4348.

Ambion (2006).RT-PCR: the basics. http://www.ambion.com/techlib/basics/rtpcr/ Accessed 6/5/06

Anagnostopoulos, C. & Spizizen, J. (1961). Requirements for transformation in *Bacillus subtilis*. *Journal of Bacteriology* **81**, 741-746.

Andreesen, J. R., Bahl, H. & Gottschalk, G. (1989). Introduction to the physiology and biochemistry of the genus *Clostridium*. In *Clostridia*, pp. 27-62. Edited by N. P. Minton & D. J. Clarke: Plenum Press.

Arnon, S. S., Schechter, R., Inglesby, T. V. & other authors (2001). Botulinum toxin as a biological weapon: medical and public health management. *Jama* **285**, 1059-1070.

Arvidson, S. O. (1983). Extracellular enzymes from *Staphylococcus aureus*. In *Staphylococci and staphylococcal infections*, pp. 745-808. Edited by C. F. Easmon & C. Adlam. London: Academic Press.

Aureli, P., Fenicia, L., Pasolini, B., Gianfranceschi, M., McCroskey, L. M. & Hatheway, C. L. (1986). Two cases of type E infant botulism caused by neurotoxigenic *Clostridium butyricum* in Italy. *J Infect Dis* **154**, 207-211.

Autret, N., Raynaud, C., Dubail, I., Berche, P. & Charbit, A. (2003). Identification of the *agr* locus of *Listeria monocytogenes*: role in bacterial virulence. *Infect Immun* **71**, 4463-4471.

Barnes, M. (2003). Botulinum toxin--mechanisms of action and clinical use in spasticity. *J Rehabil Med*, 56-59.

Bassler, B. L., Wright, M., Showalter, R. E. & Silverman, M. R. (1993). Intercellular signalling in *Vibrio harveyi*: sequence and function of genes regulating expression of luminescence. *Mol Microbiol* **9**, 773-786.

Bayer, M. G., Heinrichs, J. H. & Cheung, A. L. (1996). The molecular architecture of the *sar* locus in *Staphylococcus aureus*. *J Bacteriol* **178**, 4563-4570.

Bell, C. & Kyriakides, A. (2000). *Clostridium botulinum. A practical approach to the organism and its control in foods*: Blackwell Science.

Blevins, J. S., Beenken, K. E., Elasri, M. O., Hurlburt, B. K. & Smeltzer, M. S. (2002). Strain-dependent differences in the regulatory roles of *sarA* and *agr* in *Staphylococcus aureus*. *Infect Immun* **70**, 470-480.

Blocker, D., Pohlmann, K., Haug, G., Bachmeyer, C., Benz, R., Aktories, K. & Barth, H. (2003). *Clostridium botulinum* C2 toxin: low pH-induced pore formation is required for translocation of the enzyme component C2I into the cytosol of host cells. *J Biol Chem* **278**, 37360-37367.

Blumenfeld, A. M., Dodick, D. W. & Silberstein, S. D. (2004). Botulinum neurotoxin for the treatment of migraine and other primary headache disorders. *Dermatol Clin* **22**, 167-175.

Bo, X., Lou, S., Sun, D., Yang, J. & Wang, S. (2006). AOBase: a database for antisense oligonucleotides selection and design. *Nucleic acids research* **34**, D664-667.

Boisset, S., Geissmann, T., Huntzinger, E. & other authors (2007). Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & development* **21**, 1353-1366.

Bradshaw, M., Goodnough, M. C. & Johnson, E. A. (1998). Conjugative transfer of the *Escherichia coli-Clostridium perfringens* shuttle vector pJIR1457 to *Clostridium botulinum* type A strains. *Plasmid* **40**, 233-237.

Breidenbach, M. A. & Brunger, A. T. (2005). New insights into clostridial neurotoxin-SNARE interactions. *Trends Mol Med* **11**, 377-381.

Brett, M. M., Hallas, G. & Mpamugo, O. (2004). Wound botulism in the UK and Ireland. *J Med Microbiol* **53**, 555-561.

Bron, S. & Luxen, E. (1985). Segregational instability of pUB110-derived recombinant plasmids in *Bacillus subtilis*. *Plasmid* **14**, 235-244.

Bron, S. (1990). Plasmids. In *Molecular biological methods for Bacillus*, pp. 146-147. Edited by C. R. Harwood & S. M. Cutting. Chichester: John Wiley & Sons.

Brown, K. L. (2000). Control of bacterial spores. *Br Med Bull* **56**, 158-171.

Buchholz, F., Ringrose, L., Angrand, P. O., Rossi, F. & Stewart, A. F. (1996). Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Research* **24**, 4256-4262.

Bustin, S. A. & Nolan, T. (2004). Analysis of mRNA expression by real-time PCR. In *Real time PCR An essential guide*, pp. 125-184. Edited by K. Edwards, J. Logan & N. Saunders. Norfolk: Horizon Bioscience.

Byard, R. W., Moore, L., Bourne, A. J., Lawrence, A. J. & Goldwater, P. N. (1992). *Clostridium botulinum* and sudden infant death syndrome: a 10 year prospective study. *J Paediatr Child Health* **28**, 156-157.

Carter, G. P. (2004). The development of molecular tools and the study of LuxS in *Clostridium difficile*. PhD thesis. *Institute of Infection, Immunity and Inflammation*, University of Nottingham.

Carter, G. P., Purdy, D., Williams, P. & Minton, N. P. (2005). Quorum sensing in *Clostridium difficile*: analysis of a luxS-type signalling system. *J Med Microbiol* **54**, 119-127.

Cassat, J., Dunman, P. M., Murphy, E. & other authors (2006). Transcriptional profiling of a *Staphylococcus aureus* clinical isolate and its isogenic *agr* and sarA mutants reveals global differences in comparison to the laboratory strain RN6390. *Microbiology* **152**, 3075-3090.

Chambers, S. P., Prior, S. E., Barstow, D. A. & Minton, N. P. (1988). The pMTL nic- cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* **68**, 139-149.

Chan, W. C., Coyle, B. J. & Williams, P. (2004). Virulence regulation and quorum sensing in staphylococcal infections: competitive AgrC antagonists as quorum sensing inhibitors. *J Med Chem* **47**, 4633-4641.

Chao, S. S., Graham, S. M. & Hoffman, H. T. (2004). Management of pharyngoesophageal spasm with Botox. *Otolaryngol Clin North Am* **37**, 559-566.

Chen, X., Schauder, S., Potier, N., Van Dorsselaer, A., Pelczer, I., Bassler, B. L. & Hughson, F. M. (2002). Structural identification of a bacterial quorum-sensing signal containing boron. *Nature* **415**, 545-549.

Chen, Y., McClane, B. A., Fisher, D. J., Rood, J. I. & Gupta, P. (2005). Construction of an alpha toxin gene knockout mutant of *Clostridium perfringens* type A by use of a mobile group II intron. *Appl Environ Microbiol* **71**, 7542-7547.

Cherington, M. (1998). Clinical spectrum of botulism. *Muscle Nerve* **21**, 701-710.

Childers, M. K., Brashear, A., Jozefczyk, P. & other authors (2004). Dose-dependent response to intramuscular botulinum toxin type A for upper-limb spasticity in patients after a stroke. *Arch Phys Med Rehabil* **85**, 1063-1069.

Coleman, E. A. & Yergler, M. E. (2002). Botulism. *Am J Nurs* **102**, 44-47.

Cordivari, C., Misra, V. P., Catania, S. & Lees, A. J. (2004). New therapeutic indications for botulinum toxins. *Mov Disord* **19 Suppl 8**, S157-161.

Couesnon, A., Raffestin, S. & Popoff, M. R. (2006). Expression of botulinum neurotoxins A and E, and associated non-toxin genes, during the transition phase and stability at high temperature: analysis by quantitative reverse transcription-PCR. *Microbiology* **152**, 759-770.

Cox, N. & Hinkle, R. (2002). Infant botulism. *Am Fam Physician* **65**, 1388-1392.

Cruz, F. (2004). Mechanisms involved in new therapies for overactive bladder. *Urology* **63**, 65-73.

Davis, T. O., Henderson, I., Brehm, J. K. & Minton, N. P. (2000). Development of a transformation and gene reporter system for group II, non-proteolytic *Clostridium botulinum* type B strains. *Journal of Molecular Microbiology and Biotechnology* **2**, 59-69.

Dekleva, M. L. & Dasgupta, B. R. (1990). Purification and characterization of a protease from *Clostridium botulinum* type A that nicks single-chain type A botulinum neurotoxin into the di-chain form. *J Bacteriol* **172**, 2498-2503.

Diep, D. B., Havarstein, L. S., Nissen-Meyer, J. & Nes, I. F. (1994). The gene encoding plantaricin A, a bacteriocin from *Lactobacillus plantarum* C11, is located on the same transcription unit as an agr-like regulatory system. *Appl Environ Microbiol* **60**, 160-166.

Dodds, K. L. (1993). Worldwide incidence and ecology of infant botulism. In *Clostridium botulinum Ecology and control in foods*, pp. 105-117. Edited by A. H. W. Hauschild & K. L. Dodds: Marcel Dekker Inc.

Doherty, N., Holden, M. T., Qazi, S. N., Williams, P. & Winzer, K. (2006). Functional analysis of *luxS* in *Staphylococcus aureus* reveals a role in metabolism but not quorum sensing. *J Bacteriol* **188**, 2885-2897.

Dufour, P., Jarraud, S., Vandenesch, F., Greenland, T., Novick, R. P., Bes, M., Etienne, J. & Lina, G. (2002). High genetic variability of the *agr* locus in *Staphylococcus* species. *J Bacteriol* **184**, 1180-1186.

Dunman, P. M., Murphy, E., Haney, S. & other authors (2001). Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. *J Bacteriol* **183**, 7341-7353.

Dupuy, B. & Sonenshein, A. L. (1998). Regulated transcription of *Clostridium difficile* toxin genes. *Mol Microbiol* **27**, 107-120.

Dupuy, B. & Matamouros, S. (2006). Regulation of toxin and bacteriocin synthesis in *Clostridium* species by a new subgroup of RNA polymerase sigma-factors. *Research in microbiology* **157**, 201-205.

Durre, P. (2001). From Pandora's box to cornucopia: clostridia - a historical perspective. In *Clostridia Biotechnology and medical applications*, pp. 8-17. Edited by H. Bahl & P. Durre: Wiley-VCH, Germany.

Durre, P. & Hollergschwandner, C. (2004). Initiation of endospore formation in *Clostridium acetobutylicum*. *Anaerobe* **10**, 69-74.

Ehrlich, S. D., Bruand, C., Sozhamannan, S., Dabert, P., Gros, M. F., Janniere, L. & Gruss, A. (1991). Plasmid replication and structural stability in *Bacillus subtilis*. *Research in Microbiology* **142**, 869-873.

Fasshauer, D., Sutton, R. B., Brunger, A. T. & Jahn, R. (1998). Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 15781-15786.

Foran, P. G., Davletov, B. & Meunier, F. A. (2003). Getting muscles moving again after botulinum toxin: novel therapeutic challenges. *Trends Mol Med* **9**, 291-299.

Fuqua, C. & Greenberg, E. P. (1998). Self perception in bacteria: quorum sensing with acylated homoserine lactones. *Curr Opin Microbiol* **1**, 183-189.

Garnier, T. & Cole, S. T. (1988). Studies of UV-inducible promoters from *Clostridium perfringens* in vivo and in vitro. *Mol Microbiol* **2**, 607-614.

Geisinger, E., Adhikari, R. P., Jin, R., Ross, H. F. & Novick, R. P. (2006). Inhibition of rot translation by RNAIII, a key feature of *agr* function. *Mol Microbiol* **61**, 1038-1048.

Gimenez, D. F. & Ciccarelli, A. S. (1970). Studies on strain 84 of *Clostridium botulinum*. *Zentralbl Bakteriol [Orig]* **215**, 212-220.

Graham, A. F., Mason, D. R., Maxwell, F. J. & Peck, M. W. (1997). Effect of pH and NaCl on growth from spores of non-proteolytic *Clostridium botulinum* at chill temperature. *Lett Appl Microbiol* **24**, 95-100.

Gruss, A. & Ehrlich, S. D. (1989). The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol Rev* **53**, 231-241.

Hadjifrangiskou, M., Chen, Y. & Koehler, T. M. (2007). The alternative sigma factor sigmaH is required for toxin gene expression by *Bacillus anthracis*. *J Bacteriol* **189**, 1874-1883.

Hall, J. D., McCroskey, L. M., Pincomb, B. J. & Hatheway, C. L. (1985). Isolation of an organism resembling *Clostridium barati* which produces type F botulinal toxin from an infant with botulism. *J Clin Microbiol* **21**, 654-655.

Hara-Kudo, Y., Yamakawa, Y. & Kumagai, S. (1996). Purification and some properties of *Clostridium sporogenes* hemorrhagic toxin. *Biochem Biophys Res Commun* **227**, 413-418.

Harris, L. M., Welker, N. E. & Papoutsakis, E. T. (2002). Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* **184**, 3586-3597.

Hatheway, C. L. & McCroskey, L. M. (1989). Unusual neurotoxigenic *Clostridium* recovered from human fecal specimens in the investigation of botulism. In *Recent advances in microbial ecology*, pp. 477-481. Edited by T. E. A. Hattori. Tokyo: Japan Scientific Society Press.

Hatheway, C. L. (1993). *Clostridium botulinum* and other clostridia that produce neurotoxins. In *Clostridium botulinum Ecology and control in foods*, pp. 3-20. Edited by A. H. W. Hauschild & K. L. Dodds: Marcel Dekker Inc.

Heap, J. T. (2006). Protocol for Clostridial Gene Knockout using pMTL007. Nottingham: University of Nottingham.

Heap, J. T., Pennington, O. J., Cartman, S. T., Carter, G. P. & Minton, N. P. (2007). The ClosTron: A universal gene knock-out system for the genus *Clostridium*. *Journal of Microbiological Methods* **70**, 452-464.

Herreros, J., Lalli, G., Montecucco, C. & Schiavo, G. (1999). Pathophysiological properties of clostridial neurotoxins. In *The comprehensive sourcebook of bacterial protein toxins*, pp. 202-228. Edited by J. E. Alouf & J. H. Freer. London: Academic Press.

Hilgers, M. T. & Ludwig, M. L. (2001). Crystal structure of the quorum-sensing protein LuxS reveals a catalytic metal site. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11169-11174.

Hoch, J. A. (1993). Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol* **47**, 441-465.

HPA (2006). Mandatory surveillance of healthcare accociated infections report 2006, pp. 1-63. London: Health Protection Agency.

Huang, I. H., Waters, M., Grau, R. R. & Sarker, M. R. (2004). Disruption of the gene (*spo0A*) encoding sporulation transcription factor blocks endospore formation and enterotoxin production in enterotoxigenic *Clostridium perfringens* type A. *FEMS Microbiol Lett* **233**, 233-240.

Huang, I. H. & Sarker, M. R. (2006). Complementation of a *Clostridium perfringens spo0A* mutant with wild-type *spo0A* from other *Clostridium* species. *Appl Environ Microbiol* **72**, 6388-6393.

Hunt, M. (2006). Real Time PCR. In *Microbiology and Immunology online*: University of South Carolina School of Medicine. http://pathmicro.med.sc.edu/pcr/realtime-home.htm. Accessed 10/10/06

Ibrahim, M., Guillot, A., Wessner, F., Algaron, F., Besset, C., Courtin, P., Gardan, R. & Monnet, V. (2007). Synthesis of a new modified peptide by *Streptococcus thermophilus* depends on a Rgg regulator and on the oligopeptide transport system. In *4th Conference on Functional Genomics of Gram-positive Microorganisms*. Tirennia, Italy.

Janzon, L., Lofdahl, S. & Arvidson, S. (1989). Identification and nucleotide sequence of the delta-lysin gene, *hld*, adjacent to the accessory gene regulator (*agr*) of *Staphylococcus aureus*. *Mol Gen Genet* **219**, 480-485.

Janzon, L. & Arvidson, S. (1990). The role of the delta-lysin gene (*hld*) in the regulation of virulence genes by the accessory gene regulator (*agr*) in *Staphylococcus aureus*. *Embo J* **9**, 1391-1399.

Jen, K. Y. & Gewirtz, A. M. (2000). Suppression of gene expression by targeted disruption of messenger RNA: available options and current strategies. *Stem Cells* **18**, 307-319.

Ji, G., Beavis, R. C. & Novick, R. P. (1995). Cell density control of staphylococcal virulence mediated by an octapeptide pheromone. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 12055-12059.

Ji, G., Beavis, R. & Novick, R. P. (1997). Bacterial interference caused by autoinducing peptide variants. *Science* **276**, 2027-2030.

Ji, Y., Marra, A., Rosenberg, M. & Woodnutt, G. (1999). Regulated antisense RNA eliminates alpha-toxin virulence in *Staphylococcus aureus* infection. *J Bacteriol* **181**, 6585-6590.

Juan, F. J. (2004). Use of botulinum toxin-A for musculoskeletal pain in patients with whiplash associated disorders [ISRCTN68653575]. *BMC Musculoskelet Disord* **5**, 5.

Kalkum, M., Lyon, G. J. & Chait, B. T. (2003). Detection of secreted peptides by using hypothesis-driven multistage mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2795-2800.

Karberg, M., Guo, H., Zhong, J., Coon, R., Perutka, J. & Lambowitz, A. M. (2001). Group II introns as controllable gene targeting vectors for genetic manipulation of bacteria. *Nat Biotechnol* **19**, 1162-1167.

Kavanaugh, J. S., Thoendel, M. & Horswill, A. R. (2007). A role for type I signal peptidase in *Staphylococcus aureus* quorum sensing. *Mol Microbiol* **65**, 780-798.

Kernodle, D. S., Voladri, R. K., Menzies, B. E., Hager, C. C. & Edwards, K. M. (1997). Expression of an antisense *hla* fragment in *Staphylococcus aureus* reduces alpha-toxin production *in vitro* and attenuates lethal activity in a murine model. *Infect Immun* **65**, 179-184.

Khan, S. A. (1996). Mechanism of replication and copy number control of plasmids in gram-positive bacteria. *Genet Eng (N Y)* **18**, 183-201.

Khan, S. A. (2000). Plasmid rolling-circle replication: recent developments. *Mol Microbiol* **37**, 477-484.

Khan, S. A. (2005). Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* **53**, 126-136.

Kies, S., Vuong, C., Hille, M., Peschel, A., Meyer, C., Gotz, F. & Otto, M. (2003). Control of antimicrobial peptide synthesis by the agr quorum sensing system in *Staphylococcus epidermidis*: activity of the lantibiotic epidermin is regulated at the level of precursor peptide processing. *Peptides* **24**, 329-338.

Kim, K. H., Fekety, R., Batts, D. H., Brown, D., Cudmore, M., Silva, J., Jr. & Waters, D. (1981). Isolation of *Clostridium difficile* from the environment and contacts of patients with antibiotic-associated colitis. *J Infect Dis* **143**, 42-50.

Kleerebezem, M. & Quadri, L. E. (2001). Peptide pheromone-dependent regulation of antimicrobial peptide production in Gram-positive bacteria: a case of multicellular behavior. *Peptides* **22**, 1579-1596.

Klein, C., Kaletta, C., Schnell, N. & Entian, K. D. (1992). Analysis of genes involved in biosynthesis of the lantibiotic subtilin. *Appl Environ Microbiol* **58**, 132-142.

Koenig, R. L., Ray, J. L., Maleki, S. J., Smeltzer, M. S. & Hurlburt, B. K. (2004). Staphylococcus aureus AgrA binding to the RNAIII-agr regulatory region. *Journal of bacteriology* **186**, 7549-7555.

Kong, K. F., Vuong, C. & Otto, M. (2006). *Staphylococcus* quorum sensing in biofilm formation and infection. *Int J Med Microbiol.* **296**, 133-139

Kouguchi, H., Suzuki, T., Hasegawa, K., Mutoh, S., Watanabe, T., Niwa, K., Yoneyama, T., Katoh, Y. & Ohyama, T. (2006). Quantitative detection of gene expression and toxin complex produced by *Clostridium botulinum* serotype D strain 4947. *J Microbiol Methods.* **67**, 416-423

Lamb, A. (2001). Biological weapons: the facts not the fiction. *Clin Med* **1**, 502-504.

Laskawi, R., Ellies, M., Drobik, C. & Batz, A. (1994). Botulinum toxin treatment in patients with hemifacial spasm. *Eur Arch Otorhinolaryngol* **251**, 271-274.

Lee, A. S. & Song, K. P. (2005). LuxS/autoinducer-2 quorum sensing molecule regulates transcriptional virulence gene expression in *Clostridium difficile*. *Biochem Biophys Res Commun* **335**, 659-666.

Lerat, E. & Moran, N. A. (2004). The evolutionary history of quorum-sensing systems in bacteria. *Mol Biol Evol* **21**, 903-913.

Li, J. & McClane, B. A. (2006). Comparative effects of osmotic, sodium nitrite-induced, and pH-induced stress on growth and survival of *Clostridium perfringens* type A isolates carrying chromosomal or plasmid-borne enterotoxin genes. *Applied and Environmental Microbiology* **72**, 7620-7625.

Lin, W. J. & Johnson, E. A. (1991). Transposon Tn*916* mutagenesis in *Clostridium botulinum*. *Appl Environ Microbiol* **57**, 2946-2950.

Lina, G., Jarraud, S., Ji, G., Greenland, T., Pedraza, A., Etienne, J., Novick, R. P. & Vandenesch, F. (1998). Transmembrane topology and histidine protein kinase activity of AgrC, the agr signal receptor in *Staphylococcus aureus*. *Mol Microbiol* **28**, 655-662.

Lovenklev, M., Holst, E., Borch, E. & Radstrom, P. (2004). Relative neurotoxin gene expression in *Clostridium botulinum* type B, determined using quantitative reverse transcription-PCR. *Appl Environ Microbiol* **70**, 2919-2927.

Lyczak, J. B., Cannon, C. L. & Pier, G. B. (2000). Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. *Microbes Infect* **2**, 1051-1060.

Lyon, G. J., Wright, J. S., Muir, T. W. & Novick, R. P. (2002). Key determinants of receptor activation in the agr autoinducing peptides of *Staphylococcus aureus*. *Biochemistry* **41**, 10095-10104.

Lyon, G. J. & Novick, R. P. (2004). Peptide signaling in *Staphylococcus aureus* and other Gram-positive bacteria. *Peptides* **25**, 1389-1403.

MacFaddin, J. F. (2000). *Biochemical tests for identification of medical bacteria*, Third edn. Philadelphia: Lippincott Williams and Wilkins.

Mani, N. & Dupuy, B. (2001). Regulation of toxin synthesis in *Clostridium difficile* by an alternative RNA polymerase sigma factor. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 5844-5849.

Mani, N., Lyras, D., Barroso, L., Howarth, P., Wilkins, T., Rood, J. I., Sonenshein, A. L. & Dupuy, B. (2002). Environmental response and autoregulation of *Clostridium difficile* TxeR, a sigma factor for toxin gene expression. *J Bacteriol* **184**, 5971-5978.

Marvaud, J. C., Eisel, U., Binz, T., Niemann, H. & Popoff, M. R. (1998a). TetR is a positive regulator of the tetanus toxin gene in *Clostridium tetani* and is homologous to BotR. *Infect Immun* **66**, 5698-5702.

Marvaud, J. C., Gibert, M., Inoue, K., Fujinaga, Y., Oguma, K. & Popoff, M. R. (1998b). *botR/A* is a positive regulator of botulinum neurotoxin and associated non-toxin protein genes in *Clostridium botulinum* A. *Mol Microbiol* **29**, 1009-1018.

McClane, B. A. & Rood, J. I. (2001). Clostridial toxins involved in human enteric and histotoxic infections. In *Clostridia Biotechnology and medical applications*, pp. 169-209. Edited by H. Bahl & P. Durre. Weinheim: Wiley VCH.

McFarland, L. V., Beneda, H. W., Clarridge, J. E. & Raugi, G. J. (2007). Implications of the changing face of *Clostridium difficile* disease for health care practitioners. *Am J Infect Control* **35**, 237-253.

McKinney, M. W., Levett, P. N. & Haylock, R. W. (1993). Cloning of a DNA sequence unique to *Clostridium botulinum* group I by selective hybridization. *J Clin Microbiol* **31**, 1845-1849.

McLauchlin, J., Grant, K. A. & Little, C. L. (2006). Food-borne botulism in the United Kingdom. *J Public Health (Oxf)* **28**, 337-342.

McMaster, P., Piper, S., Schell, D., Gillis, J. & Chong, A. (2000). A taste of honey. *J Paediatr Child Health* **36**, 596-597.

Midura, T. F. & Arnon, S. S. (1976). Infant botulism. Identification of *Clostridium botulinum* and its toxins in faeces. *Lancet* **2**, 934-936.

Mitchell, W. J. (2001). Biology and physiology. In *Clostridia Biotechnology and medical applications*, pp. 50-104. Edited by H. Bahl & P. Durre: Wiley-VCH, Germany.

Mohr, G., Smith, D., Belfort, M. & Lambowitz, A. M. (2000). Rules for DNA target-site recognition by a lactococcal group II intron enable retargeting of the intron to specific DNA sequences. *Genes Dev* **14**, 559-573.

Nakayama, J., Cao, Y., Horii, T., Sakuda, S., Akkermans, A. D. & de Vos, W. M. (2001). Gelatinase biosynthesis-activating pheromone: a peptide lactone that mediates quorum sensing in *Enterococcus faecalis*. *Molecular Microbiology* **41**, 145-154.

Nakayama, J., Chen, S., Oyama, N., Nishiguchi, K., Azab, E. A., Tanaka, E., Kariyama, R. & Sonomoto, K. (2006). Revised model for *Enterococcus faecalis* fsr quorum-sensing system: the small open reading frame *fsrD* encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal AgrD. *J Bacteriol* **188**, 8321-8326.

Naumann, M. & Jost, W. (2004). Botulinum toxin treatment of secretory disorders. *Mov Disord* **19 Suppl 8**, S137-141.

Nealson, K. H., Platt, T. & Hastings, J. W. (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. *J Bacteriol* **104**, 313-322.

Niaudet, B., Goze, A. & Ehrlich, S. D. (1982). Insertional mutagenesis in *Bacillus subtilis*: mechanism and use in gene cloning. *Gene* **19**, 277-284.

Noirot, P., Petit, M. A. & Ehrlich, S. D. (1987). Plasmid replication stimulates DNA recombination in *Bacillus subtilis*. *J Mol Biol* **196**, 39-48.

Novick, R. P., Projan, S. J., Kornblum, J., Ross, H. F., Ji, G., Kreiswirth, B., Vandenesch, F. & Moghazeh, S. (1995). The *agr* P2 operon: an autocatalytic sensory transduction system in *Staphylococcus aureus*. *Mol Gen Genet* **248**, 446-458.

Novick, R. P. (2003). Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* **48**, 1429-1449.

O'Connor, J. R., Lyras, D., Farrow, K. A., Adams, V., Powell, D. R., Hinds, J., Cheung, J. K. & Rood, J. I. (2006). Construction and analysis of chromosomal *Clostridium difficile* mutants. *Molecular Microbiology* **61**, 1335-1351.

Ohtani, K., Bhowmik, S. K., Hayashi, H. & Shimizu, T. (2002a). Identification of a novel locus that regulates expression of toxin genes in *Clostridium perfringens*. *FEMS Microbiol Lett* **209**, 113-118.

Ohtani, K., Hayashi, H. & Shimizu, T. (2002b). The *luxS* gene is involved in cell-cell signalling for toxin production in *Clostridium perfringens*. *Mol Microbiol* **44**, 171-179.

Otto, M., Sussmuth, R., Jung, G. & Gotz, F. (1998). Structure of the pheromone peptide of the *Staphylococcus epidermidis agr* system. *FEBS Lett* **424**, 89-94.

Paredes, C. J., Alsaker, K. V. & Papoutsakis, E. T. (2005). A comparative genomic view of clostridial sporulation and physiology. *Nat Rev Microbiol* **3**, 969-978.

Patterson-Curtis, S. I. & Johnson, E. A. (1989). Regulation of neurotoxin and protease formation in *Clostridium botulinum* Okra B and Hall A by arginine. *Appl Environ Microbiol* **55**, 1544-1548.

Peck, M. W. & Stringer, S. C. (2005). The safety of pasteurised in-pack chilled meat products with respect to the foodborne botulism hazard. *Meat Science* **70**, 461-475.

Peck, M. W. (2006). *Clostridium botulinum* and the safety of minimally heated, chilled foods: an emerging issue? *J Appl Microbiol* **101**, 556-570.

Pennington, O. J. (2006). The development of molecular tools for the expression of prodrug converting enzymes in *Clostridium sporogenes*. PhD thesis. *Institute of Infections, Immunity and Inflammation*. University of Nottingham.

Perutka, J., Wang, W., Goerlitz, D. & Lambowitz, A. M. (2004). Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J Mol Biol* **336**, 421-439.

Piggot, P. J. & Hilbert, D. W. (2004). Sporulation of *Bacillus subtilis*. *Curr Opin Microbiol* **7**, 579-586.

Posfai, G., Kolisnychenko, V., Bereczki, Z. & Blattner, F. R. (1999). Markerless gene replacement in *Escherichia coli* stimulated by a double-strand break in the chromosome. *Nucleic Acids Research* **27**, 4409-4415.

Poumeyrol, M. *et al.* (1983). Intoxication botulique mortelle due a une souche de *Clostridium botulinum* de type AB. *Med Malad Infec* **13**, 750.

Projan, S. J., Monod, M., Narayanan, C. S. & Dubnau, D. (1987). Replication properties of pIM13, a naturally occurring plasmid found in *Bacillus subtilis*, and of its close relative pE5, a plasmid native to *Staphylococcus aureus*. *J Bacteriol* **169**, 5131-5139.

Purdy, D., O'Keeffe, T. A., Elmore, M., Herbert, M., McLeod, A., Bokori-Brown, M., Ostrowski, A. & Minton, N. P. (2002). Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. *Mol Microbiol* **46**, 439-452.

Qiu, R., Pei, W., Zhang, L., Lin, J. & Ji, G. (2005). Identification of the putative staphylococcal AgrB catalytic residues involving the proteolytic cleavage of AgrD to generate autoinducing peptide. *J Biol Chem* **280**, 16695-16704.

Quinn, C. P. & Minton, N. P. (2001). Clostridial neurotoxins. In *Clostridia Biotechnology and medical applications*, pp. 211-250. Edited by H. Bahl & P. Durre. Germany: Wiley-VCH.

Raffestin, S., Dupuy, B., Marvaud, J. C. & Popoff, M. R. (2005). BotR/A and TetR are alternative RNA polymerase sigma factors controlling the expression of the neurotoxin and associated protein genes in *Clostridium botulinum* type A and *Clostridium tetani*. *Mol Microbiol* **55**, 235-249.

Rao, S., Starr, R. L., Morris, M. G. & Lin, W. J. (2007). Variations in expression and release of botulinum neurotoxin in *Clostridium botulinum* type A strains. *Foodborne Pathogens and Disease* **4**, 201-207.

Ravagnani, A., Jennert, K. C., Steiner, E. & other authors (2000). Spo0A directly controls the switch from acid to solvent production in solvent-forming clostridia. *Mol Microbiol* **37**, 1172-1185.

Recsei, P., Kreiswirth, B., O'Reilly, M., Schlievert, P., Gruss, A. & Novick, R. P. (1986). Regulation of exoprotein gene expression in *Staphylococcus aureus* by agar. *Mol Gen Genet* **202**, 58-61.

Roche (2006). Transcriptor Reverse Transcriptase online manual: Roche Applied Science.
http://www.rocheappliedscience.com/packinsert/3531317a.pdf#search='Reverse%20transcriptase%20PCR' Accessed 6/5/06

Ryan, R. P., Fouhy, Y., Lucey, J. F. & other authors (2006). Cell-cell signaling in Xanthomonas campestris involves an HD-GYP domain protein that functions in cyclic di-GMP turnover. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6712-6717.

Saenz, H. L., Augsburger, V., Vuong, C., Jack, R. W., Gotz, F. & Otto, M. (2000). Inducible expression and cellular location of AgrB, a protein involved in the maturation of the staphylococcal quorum-sensing pheromone. *Arch Microbiol* **174**, 452-455.

Saint-Prix, F., Gonzalez-Pajuelo, M., Croux, C., Girbal, L. & Soucaille, P. (2006). Development of a system for chromosomal integration and gene deletion in *Clostridium acetobutylicum* ATCC 824. In *Clostridium IX Ninth International Workshop on the Regulation of Metabolism, Genetics and Development of the Solvent- and Acid-forming Clostridia*. Rice University, Houston, Texas, USA.

Salyers, A. A. & Whitt, D. D. (2002). *Bacterial pathogenesis. A molecular approach*, Second edn: ASM Press.

Sarker, M. R., Carman, R. J. & McClane, B. A. (1999). Inactivation of the gene (*cpe*) encoding *Clostridium perfringens* enterotoxin eliminates the ability of two *cpe*-positive *C. perfringens* type A human gastrointestinal disease isolates to affect rabbit ileal loops. *Mol Microbiol* **33**, 946-958.

Satoh, E., Ishii, T., Shimizu, Y., Sawamura, S. & Nishimura, M. (2001). Black tea extract, thearubigin fraction, counteract the effects of botulinum neurotoxins in mice. *Br J Pharmacol* **132**, 797-798.

Satoh, E., Ishii, T., Shimizu, Y., Sawamura, S. & Nishimura, M. (2002). The mechanism underlying the protective effect of the thearubigin fraction of black tea (*Camellia sinensis*) extract against the neuromuscular blocking action of botulinum neurotoxins. *Pharmacol Toxicol* **90**, 199-202.

Satoh, E. (2005). Ethyl acetate extract from black tea prevents neuromuscular blockade by botulinum neurotoxin type A in vitro. *Int J Food Sci Nutr* **56**, 543-550.

Schantz, E. J. & Johnson, E. A. (1992). Properties and use of botulinum toxin and other microbial neurotoxins in medicine. *Microbiol Rev* **56**, 80-99.

Schauder, S., Shokat, K., Surette, M. G. & Bassler, B. L. (2001). The LuxS family of bacterial autoinducers: biosynthesis of a novel quorum-sensing signal molecule. *Mol Microbiol* **41**, 463-476.

Scherr, M. & Rossi, J. J. (1998). Rapid determination and quantitation of the accessibility to native RNAs by antisense oligodeoxynucleotides in murine cell extracts. *Nucleic acids research* **26**, 5079-5085.

Schlingensiepen, R. & Schlingensiepen, K. (1997). Technical aspects of antisense oligonucleotides. In *Antisense - from technology to therapy*, pp. 3-28. Edited by R. Schlingensiepen, W. Brysch & K. Schlingensiepen. Berlin: Blackwell Wissenschaft.

Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 5857-5864.

Sebaihia, M., Peck, M. W., Minton, N. P. & other authors (2007). Genome sequence of a proteolytic (Group I) *Clostridium botulinum* strain Hall A and comparative analysis of the clostridial genomes. *Genome Res.* **17,** 1082-1092

Setlow, P. & Johnson, E. A. (1997). Spores and their significance. In *Food Microbiology: Fundamentals and Frontiers*, pp. 33-70. Edited by M. P. Doyle, L. R. Beuchat & T. J. Montville. Washington DC: ASM Press.

Shin, N. R., Shin, J. H., Chun, J. H., Yoon, S. Y., Kim, B. S., Oh, H. B. & Rhie, G. E. (2006). Determination of neurotoxin gene expression in *Clostridium botulinum* Type A by Quantitative RT-PCR. *Mol Cells* **22**, 336-342.

Simpson, L. L. (2004). Identification of the major steps in botulinum toxin action. *Annu Rev Pharmacol Toxicol* **44**, 167-193.

Slater, H., Alvarez-Morales, A., Barber, C. E., Daniels, M. J. & Dow, J. M. (2000). A two-component system involving an HD-GYP domain protein links cell-cell signalling to pathogenicity gene expression in *Xanthomonas campestris*. *Mol Microbiol* **38**, 986-1003.

Smith, J. L., Fratamico, P. M. & Novak, J. S. (2004). Quorum sensing: a primer for food microbiologists. *J Food Prot* **67**, 1053-1070.

Sonnabend, O. A., Sonnabend, W. F., Krech, U., Molz, G. & Sigrist, T. (1985). Continuous microbiological and pathological study of 70 sudden and unexpected infant deaths: toxigenic intestinal *Clostridium botulinum* infection in 9 cases of sudden infant death syndrome. *Lancet* **1**, 237-241.

Stephenson, K. & Lewis, R. J. (2005). Molecular insights into the initiation of sporulation in Gram-positive bacteria: new technologies for an old phenomenon. *FEMS Microbiol Rev* **29**, 281-301.

Stratagene (2007). Real-Time PCR Systems Application Note. http://www.umass.edu/research/genomics/files/Real-Time_PCR_with_SYBR_Green_Dye.pdf. Accessed 8/5/07

Stumbo, C. R., Purohit, K. S. & Ramakrishnan, T. V. (1975). Thermal process lethality guide for low-acid foods in metal containers. *Journal of Food Science* **40**, 1316-1323.

Sturme, M. H., Kleerebezem, M., Nakayama, J., Akkermans, A. D., Vaugha, E. E. & de Vos, W. M. (2002). Cell to cell communication by autoinducing peptides in gram-positive bacteria. *Antonie Van Leeuwenhoek* **81**, 233-243.

Sutton, R. B., Fasshauer, D., Jahn, R. & Brunger, A. T. (1998). Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 A resolution. *Nature* **395**, 347-353.

Suzuki, N., Nonaka, H., Tsuge, Y., Okayama, S., Inui, M. & Yukawa, H. (2005). Multiple large segment deletion method for *Corynebacterium glutamicum*. *Applied Microbiology and Biotechnology* **69**, 151-161.

Sycha, T., Graninger, M., Auff, E. & Schnider, P. (2004). Botulinum toxin in the treatment of Raynaud's phenomenon: a pilot study. *Eur J Clin Invest* **34**, 312-313.

Takumi, K., Kinouchi, T. & Kawata, T. (1980). Isolation of nontoxigenic variants associated with enhanced sporulation and alteration in the cell wall from *Clostridium botulinum* type a 190L by treatment with detergents. *Microbiol Immunol* **24**, 469-477.

Teichert, M., Tatge, H., Schoentaube, J., Just, I. & Gerhard, R. (2006). Application of mutated *Clostridium difficile* toxin A for determination of glucosyltransferase-dependent effects. *Infect Immun* **74**, 6006-6010.

Traber, K. & Novick, R. (2006). A slipped-mispairing mutation in AgrA of laboratory strains and clinical isolates results in delayed activation of *agr* and failure to translate delta- and alpha-haemolysins. *Mol Microbiol* **59**, 1519-1530.

Trach, K. A. & Hoch, J. A. (1993). Multisensory activation of the phosphorelay initiating sporulation in *Bacillus subtilis*: identification and sequence of the protein kinase of the alternate pathway. *Molecular Microbiology* **8**, 69-79.

Tummala, S. B., Tomas, C., Harris, L. M., Welker, N. E., Rudolph, F. B., Bennett, G. N. & Papoutsakis, E. T. (2001). Genetic tools for solventogenic clostridia. In *Clostridia Biotechnology and medical applications*, pp. 105-123. Edited by H. Bahl & P. Durre. Germany: Wiley-VCH.

Tummala, S. B., Junne, S. G. & Papoutsakis, E. T. (2003a). Antisense RNA downregulation of coenzyme A transferase combined with alcohol-aldehyde dehydrogenase overexpression leads to predominantly alcohologenic *Clostridium acetobutylicum* fermentations. *J Bacteriol* **185**, 3644-3653.

Tummala, S. B., Welker, N. E. & Papoutsakis, E. T. (2003b). Design of antisense RNA constructs for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*. *J Bacteriol* **185**, 1923-1934.

van der Meer, J. R., Polman, J., Beerthuyzen, M. M., Siezen, R. J., Kuipers, O. P. & De Vos, W. M. (1993). Characterization of the *Lactococcus lactis* nisin A operon genes *nisP*, encoding a subtilisin-like serine protease involved in precursor processing, and *nisR*, encoding a regulatory protein involved in nisin biosynthesis. *J Bacteriol* **175**, 2578-2588.

Vreede, J., van der Horst, M. A., Hellingwerf, K. J., Crielaard, W. & van Aalten, D. M. (2003). PAS domains. Common structure and common flexibility. *The Journal of biological chemistry* **278**, 18434-18439.

Vuong, C., Kocianova, S., Yao, Y., Carmody, A. B. & Otto, M. (2004). Increased colonization of indwelling medical devices by quorum-sensing mutants of *Staphylococcus epidermidis* in vivo. *J Infect Dis* **190**, 1498-1505.

Webb, M. D., Pin, C., Peck, M. W. & Stringer, S. C. (2007). Historical and contemporary NaCl concentrations affect the duration and distribution of lag times from individual spores of non-proteolytic *Clostridium botulinum*. *Appl Environ Microbiol.* **73**, 2118-2127

Whitby, M., Street, A. C., Ruff, T. A. & Fenner, F. (2002). Biological agents as weapons 1: smallpox and botulism. *Med J Aust* **176**, 431-433.

Winzer, K., Hardie, K. R. & Williams, P. (2002). Bacterial cell-to-cell communication: sorry, can't talk now - gone to lunch! *Curr Opin Microbiol* **5**, 216-222.

Winzer, K., Hardie, K. R. & Williams, P. (2003). LuxS and autoinducer-2: their contribution to quorum sensing and metabolism in bacteria. *Adv Appl Microbiol* **53**, 291-396.

Worner, K., Szurmant, H., Chiang, C. & Hoch, J. A. (2006). Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from *Clostridium botulinum*. *Mol Microbiol* **59**, 1000-1012.

Xavier, K. B. & Bassler, B. L. (2003). LuxS quorum sensing: more than just a numbers game. *Curr Opin Microbiol* **6**, 191-197.

Yarwood, J. M. & Schlievert, P. M. (2003). Quorum sensing in *Staphylococcus* infections. *J Clin Invest* **112**, 1620-1625.

Zhang, L., Gray, L., Novick, R. P. & Ji, G. (2002). Transmembrane topology of AgrB, the protein involved in the post-translational modification of AgrD in *Staphylococcus aureus*. *J Biol Chem* **277**, 34736-34742.

Zhang, L. & Ji, G. (2004). Identification of a staphylococcal AgrB segment(s) responsible for group-specific processing of AgrD by gene swapping. *J Bacteriol* **186**, 6706-6713.

Zhang, L., Lin, J. & Ji, G. (2004). Membrane anchoring of the AgrD N-terminal amphipathic region is required for its processing to produce a quorum-sensing pheromone in *Staphylococcus aureus*. *J Biol Chem* **279**, 19448-19456.

Zhao, L., Montville, T. J. & Schaffner, D. W. (2006). Evidence for quorum sensing in *Clostridium botulinum* 56A. *Lett Appl Microbiol* **42**, 54-58.

Zhong, J., Karberg, M. & Lambowitz, A. M. (2003). Targeted and random bacterial gene disruption using a group II intron (targetron) vector containing a retrotransposition-activated selectable marker. *Nucleic Acids Research* **31**, 1656-1664.

Zhou, Y., Sugiyama, H., Nakano, H. & Johnson, E. A. (1995). The genes for the *Clostridium botulinum* type G toxin complex are on a plasmid. *Infect Immun* **63**, 2087-2091.

Appendices

Appendix 1 – Primers used for sequencing of *C. sporogenes agr* genes

A – Primers used for first round of sequencing

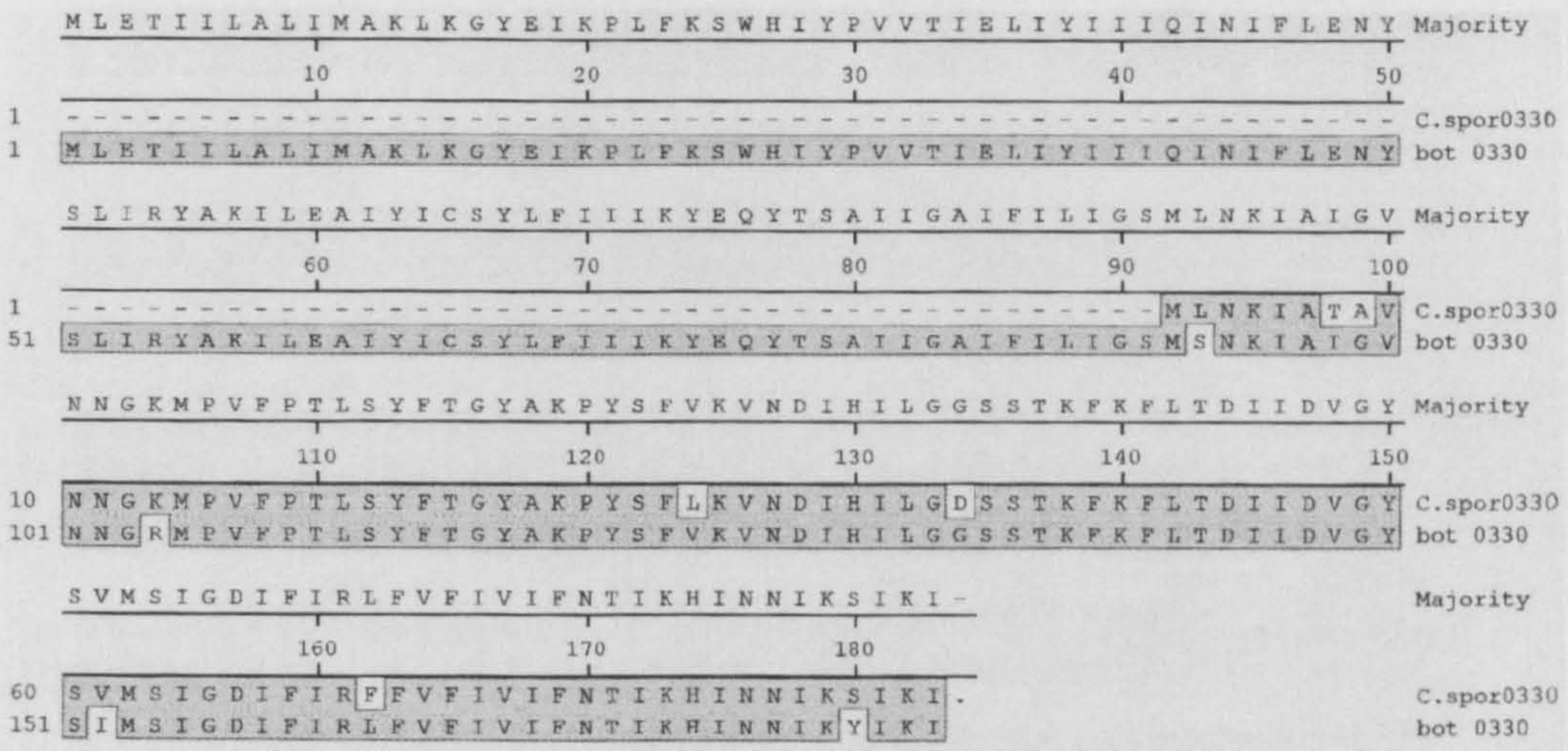
Primer name	Sequence (5' – 3')	Gene targeted
Cb330/1F	ATAAAGCCATTATTTAAATCATGG	CBO0330
Cb331/1F	CCAATCAGTTATGGTATTCATAC	CBO0331
Cb331/1R	GTATGAATACCATAACTGATTGG	CBO0331
Cb332/1R	GGTTTAGCCTTAGAGTCTACAGG	CBO0332
Cb333/1R	TTACATAAAAGCCTAAAGGCCA	CBO0333
Cs333/7F	GTTGAAAGGGTAGTTCTTTATAGTAG	CBO0333
Cb334/7R	CAATTATTGCAGCATCTG	CBO0334
Cb334/14F	GAATTAGCAAAAGAAATGC	CBO0334
Cb335/1F	GATGCCATTTCTTACAATATGGT	CBO0335
Cb336/1F	GAATTGTTAAATGAAAGTAGAGA	CBO0336
Cb336/1R	CTCTACTTTCATTTAACAATTCTAT	CBO0336
Cs336/7F	CCTTACATTAATTCCTAAAAGTG	CSP0336
Cs336/11F	GAAAATATAATTAAATTTCAGCATGG	CSP0336
Cs336/14R	GCACCTACACAAAAAAGC	CSP0336
Cb337/1R	GTATGTACACCATAGTGTATTGG	CBO0337
Cb337/1R	GTATGTACACCATAGTGTATTGG	CBO0337
Cb338/1F	CCAGTAGATACGAAAGCTAAGCC	CBO0338
Cs338/4R	CTATTTTATTTGAAATCTTTTC	CSP0338
Cb340/1R	CTAATATCTGCACTGCTG	CBO0340
Cb340/3F	GAGATATTACTTCACAAAAACAGG	CBO0340
Cb341/1R	GGTACTAAAACCTCACAACC	CBO0341

B – Walking primers used for subsequent rounds of sequencing

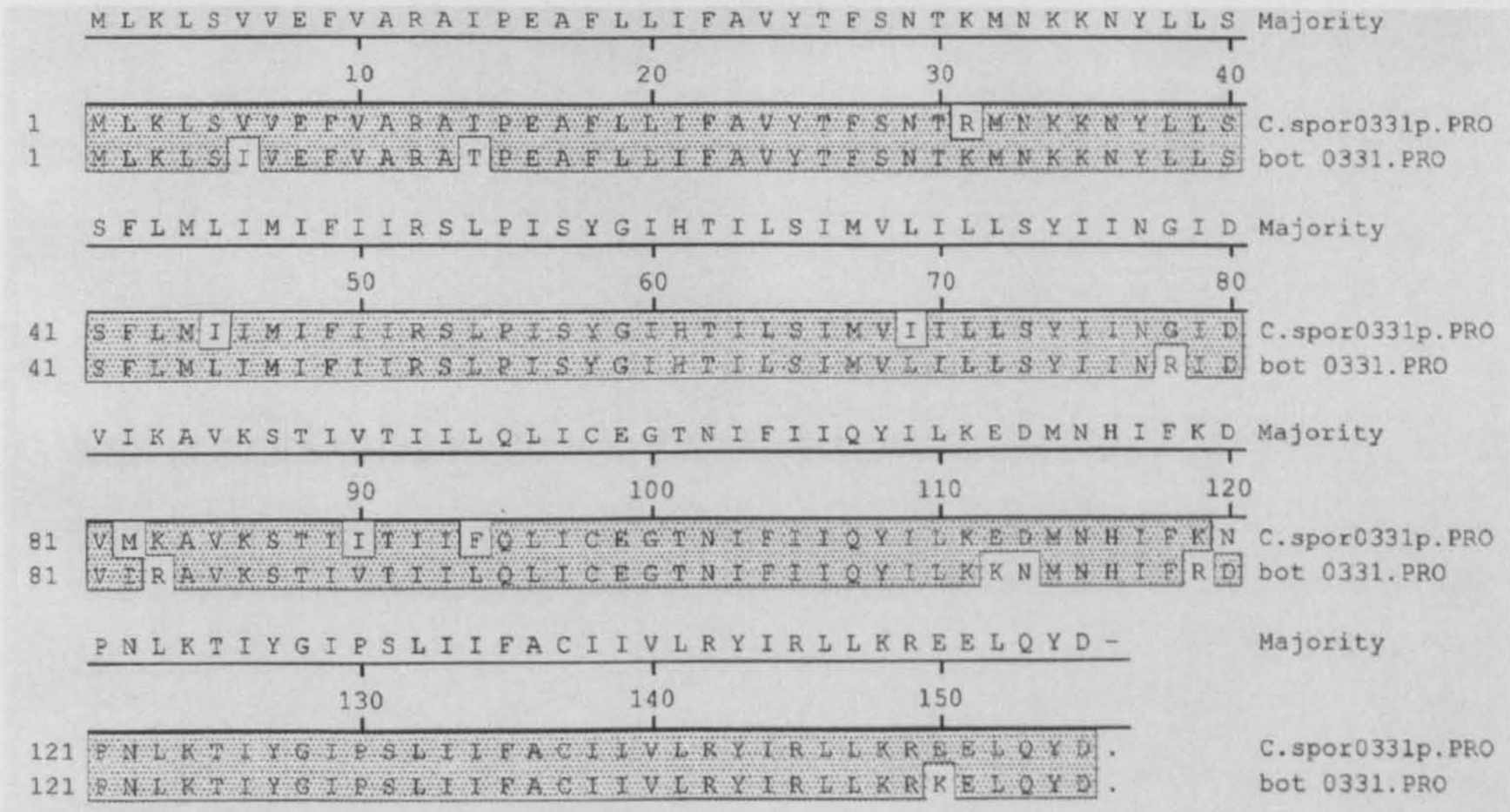
Primer name	Sequence (5' – 3')	Gene targeted
Cs330/2F	ATTATTGAAAAATATATATAAATTGC	CSP0330
Cs333/2R	GCTACCTCATGTTCTCAC	CSP0333
Cs332/2F	GAGGCTTTGCTGATATCC	CSP0332
Cs333/3R	CTTAGGGAATGCCCATATTCC	CSP0333
CsAgrD/1F	AAGATCAGTTTATAACTGATCTTTC	CSP*agrD*1
Cs333/4R	CCTATAGTTTTAAGTACCTTATCTCC	CSP0333
Cs330/3F	AAATTAGGATTCTTAAATATATGATTC	CSP0330
Cs331/2R	TTTAAAATAATTTCTAATTATTAGGATG	CSP0331
Cs336/2F	CCCTGGAGGTAATATATTTG	CSP0336
Cs337/2R	CAAACTATAGCAGGCAATTTC	CSP0337
Cs336/3F	CTTATTTGGATGTTTATATCAATAAA	CSP0336
Cs336/2R	TCTTATTACAAACTATAGCAGGC	CSP0336
Cs336/4F	GATGCATTAAATTTAGCTATTATTATAAT	CSP0336
Cs336/5F	GATGCATTAAATTTAGCTATTATTAT	CSP0336
Cs336/3R	CCTCTTTTTGTCTTAAATCCTC	CSP0336
Cs333/1F	GCAGGAGTGTTTATACTTACAGC	CSP0333
Cs333/5R	GAACCACAAATGTTACTGATATC	CSP0333
Cs333/2F	GCTACTATGTCTTCTTCTCTTGTAT	CSP0333
Cs333/6R	CTTCTTGCATCCGTAGTAGC	CSP0333
Cs337/2F	CAATAGGAGAGTAAAGGATGTT	CSP0337
Cs338/2F	CTTATGGAATAATAATATAATTTGC	CSP0338
Cs340/3R	CATATTTTGAAGTTCAGCATTAT	CSP0340
Cs336/8F	GCTTTTTTGTGTAGGTGCTAT	CSP0336
Cs336/4R	CCTTCTGGTCTAATTCTGTAG	CSP0336
Cs340/4R	GATGCACTTTCATTGGC	CSP0340
Cs338/3F	GCATTAATTTATTATGGAAGTTG	CSP0338

Primer name	Sequence (5' – 3')	Gene targeted
Cs339/1F	GTTACTGCAAAATTACTTAAATC	CSP0339
Cs340/1F	CAGTTTTATTTTTGTAAACCAAT	CSP0340
Cs338/5R	CACTTATGGAGCATTTGTG	CSP0338
Cs336/9F	GTTGATAGTAAAATTTATTATATTGC	CSP0336
Cs337/3R	GAATTAATGTAAGGTTATTTTTCTATT	CSP0337
Cs338/4F	CTTATGGAGCATTTGTGG	CSP0338
Cs340/5R	CTTTTTCTGTTCTTCTTGTAC	CSP0340
Cs340/2F	CAGATTTTAGCCCTCTTTG	CSP0340
Cs338/6R	CTTTACACTTACTCCTGCAG	CSP0338
Cs336/12F	CTAAGTGTAATCCCATACCAG	CSP0336
Cs336/5R	GAATATTCAAAAAGAGGATTTAAG	CSP0336
Cs338/5F	GGCAAAGCTTTACACTTACTC	CSP0338
Cs340/6R	CTGATGCACTTTCATTGG	CSP0340
Cs336/6R	CTATGTACTAATATAGCATCCTTTG	CSP0336
Cs336/13F	GGATACAACTAAATTAGATTCAG	CSP0336
Cs336/7R	CTGAAAATTCCATATTTATACG	CSP0336
Cs337/3F	CTCTAAAATTGTATAATACAATTTTAG	CSP0337
Cs336/8R	GATACAACTAAATTAGATTCAGG	CSP0336

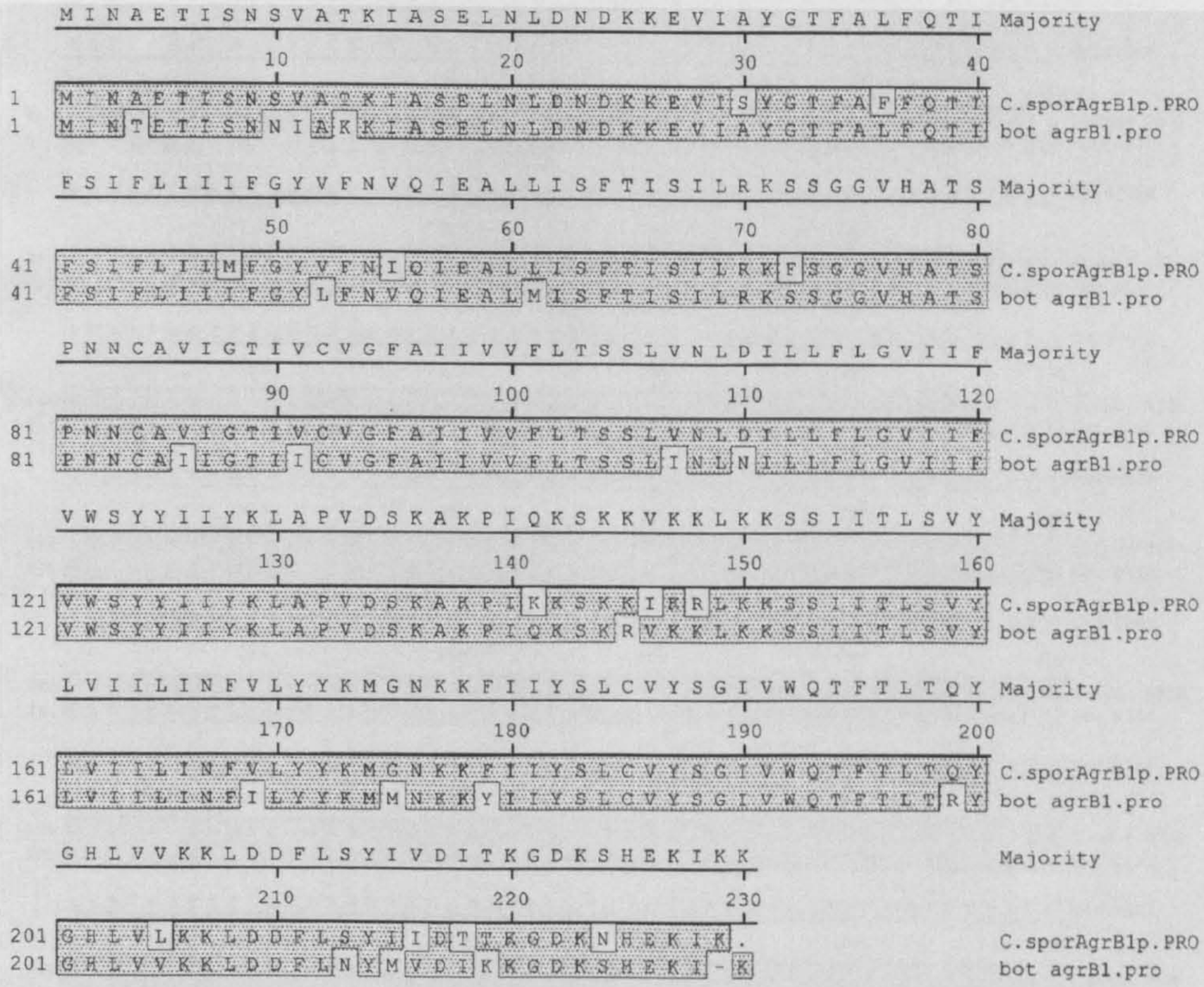
Appendix 2 – Sequence alignments for Chapter 3



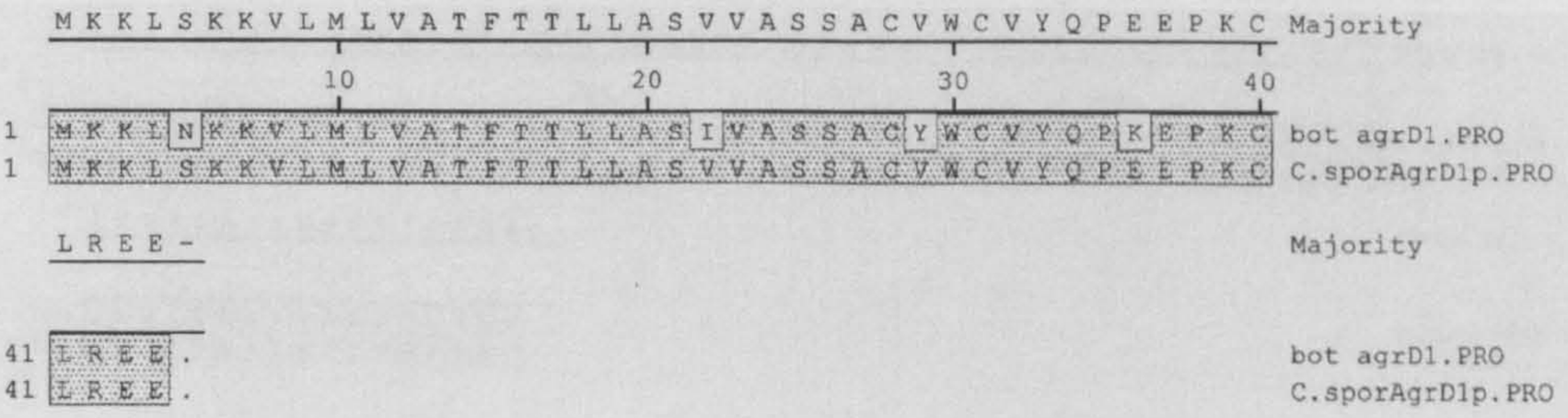
Protein alignment of *C. botulinum* and *C. sporogenes* 0330



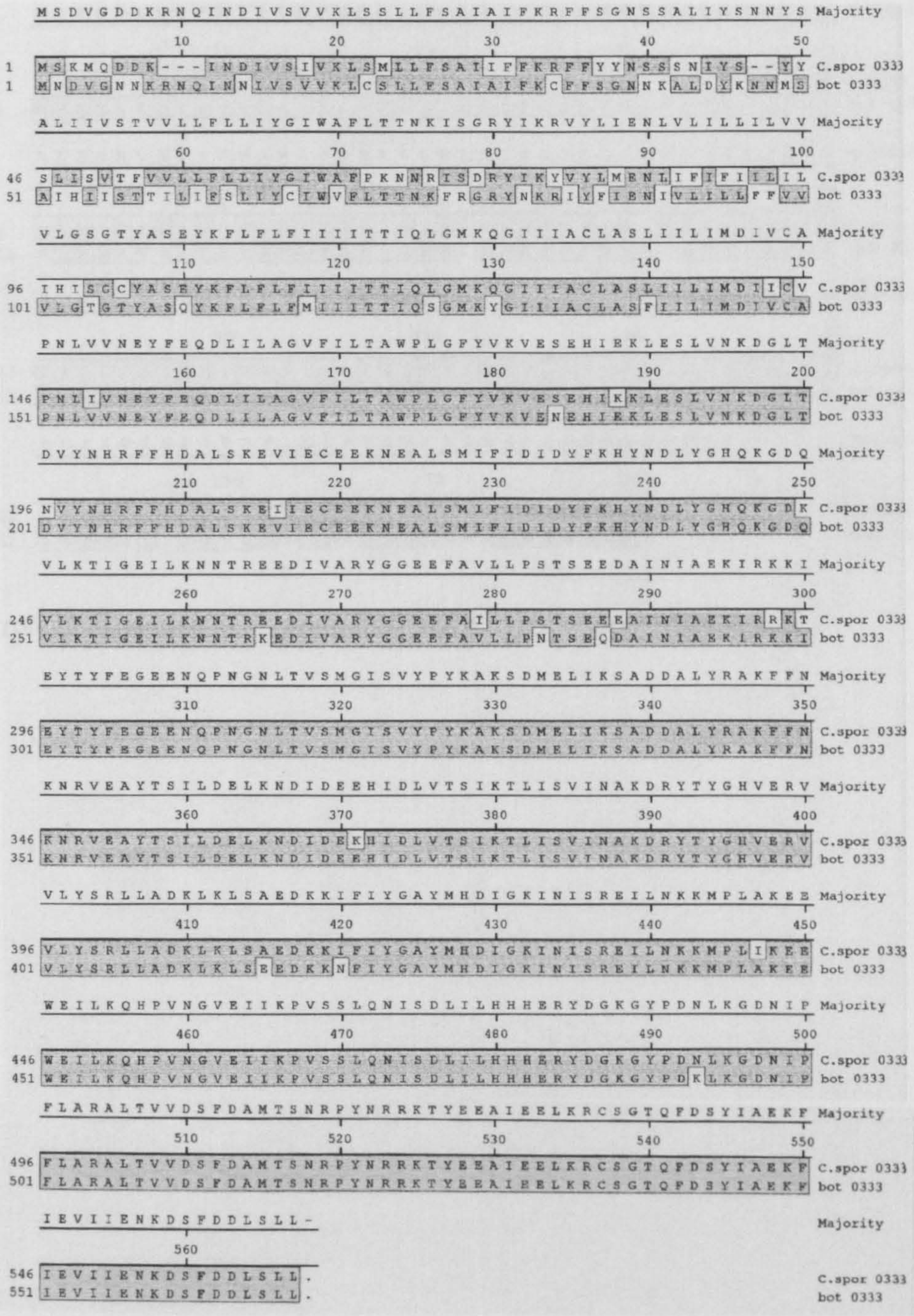
Protein alignment of *C. botulinum* and *C. sporogenes* 0331



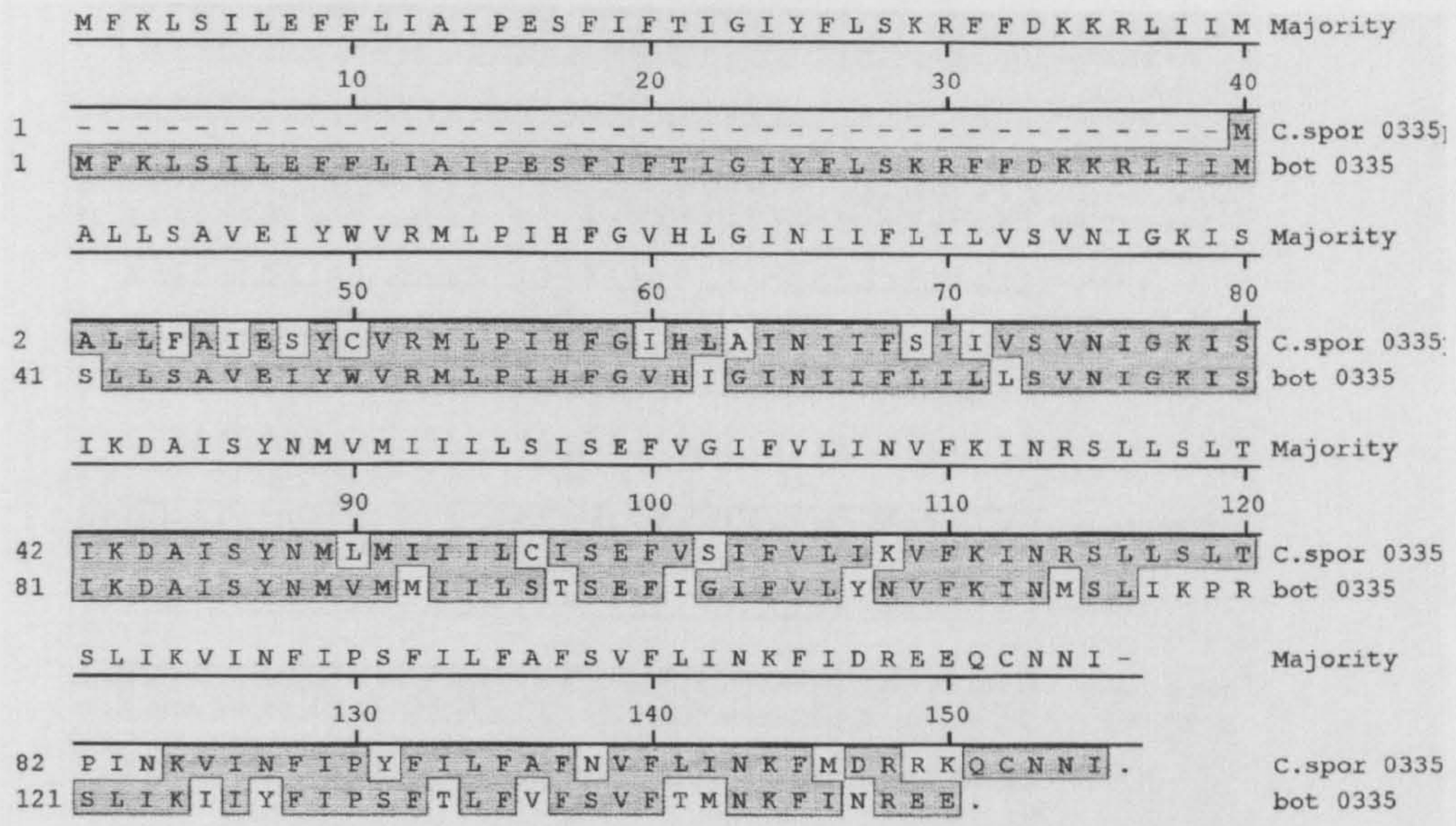
Protein alignment of *C. botulinum* and *C. sporogenes* AgrB1



Protein alignment of *C. botulinum* and *C. sporogenes* AgrD1

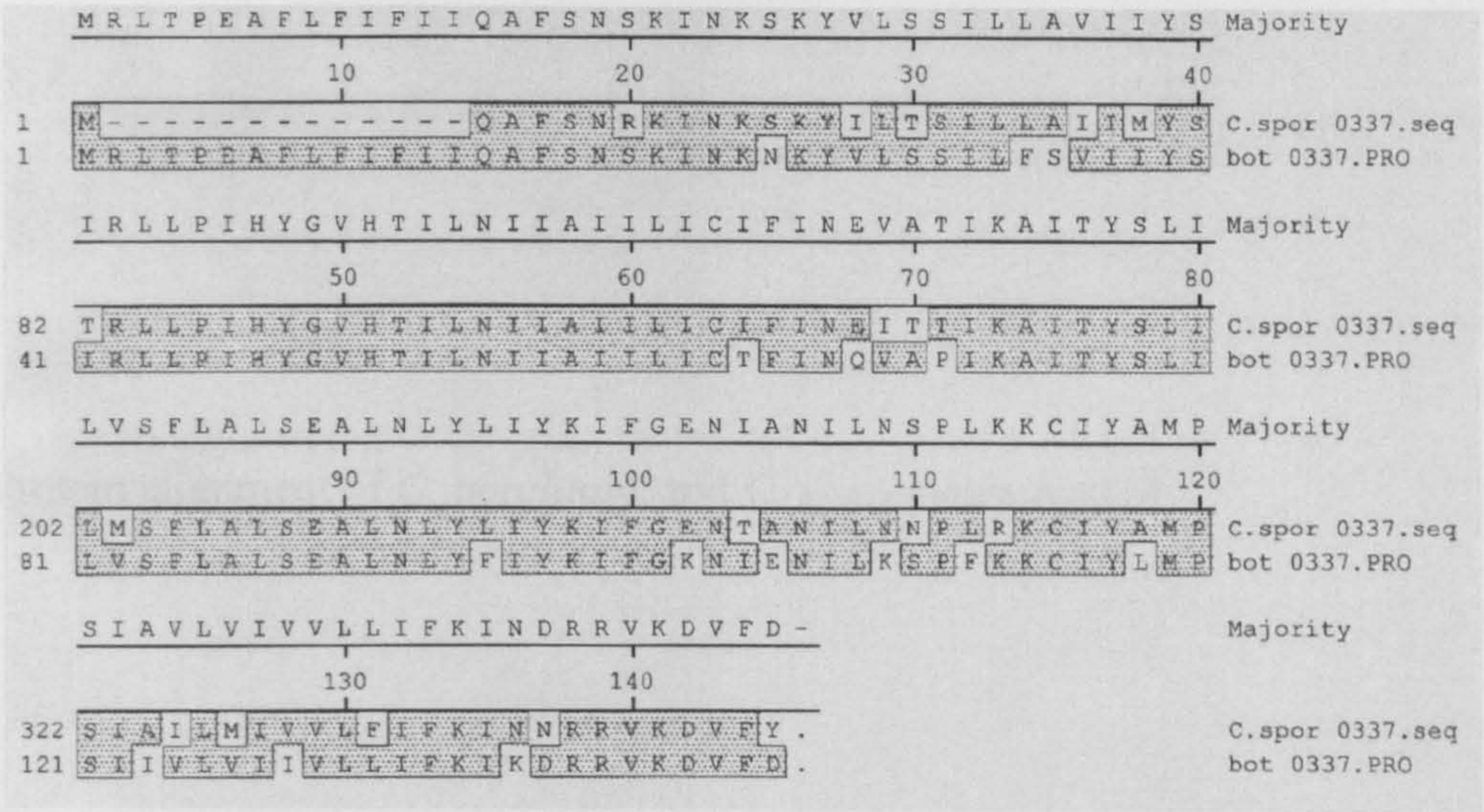


Protein alignment of *C. botulinum* and *C. sporogenes* 0333

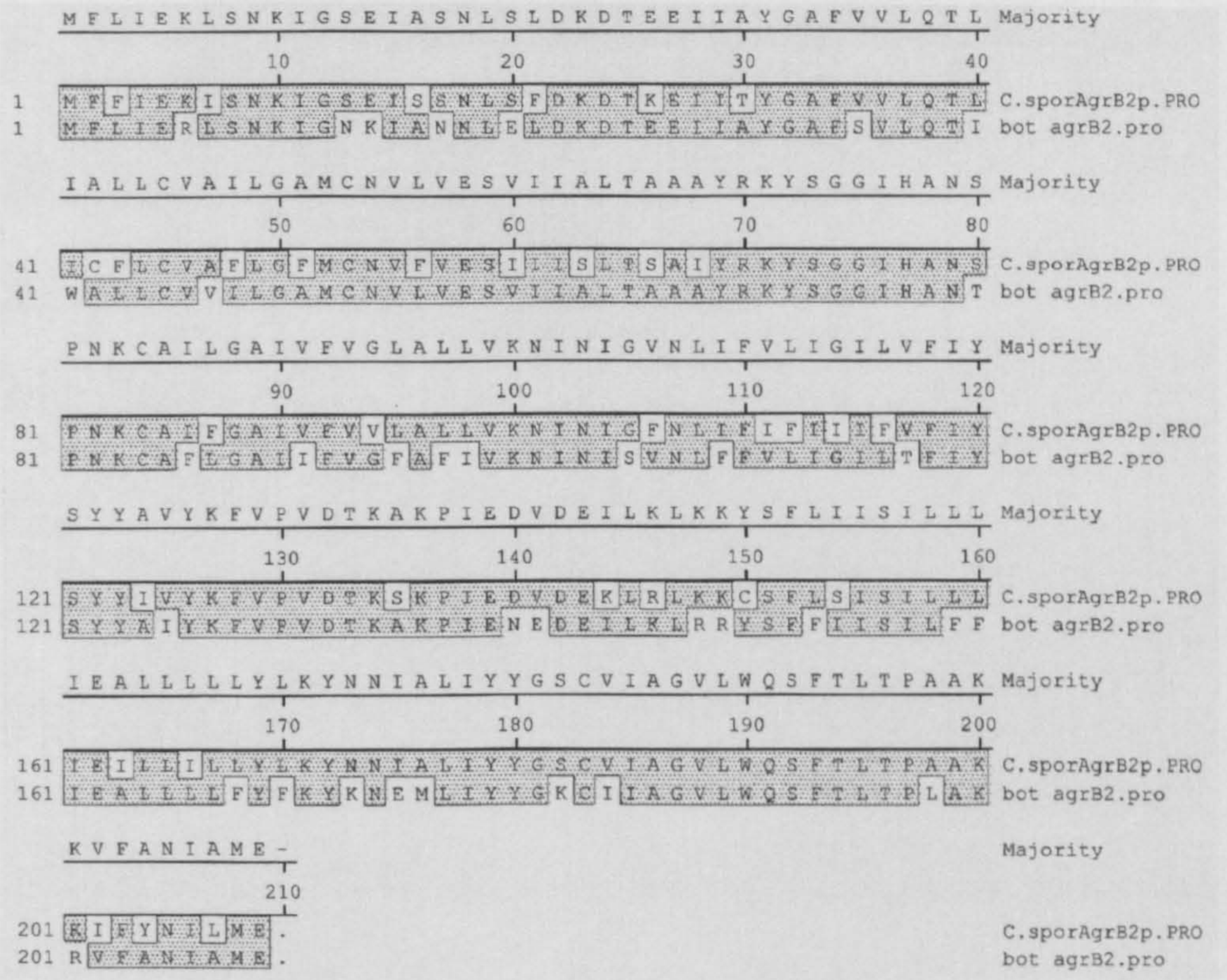


Protein alignment of *C. botulinum* and *C. sporogenes* 0335

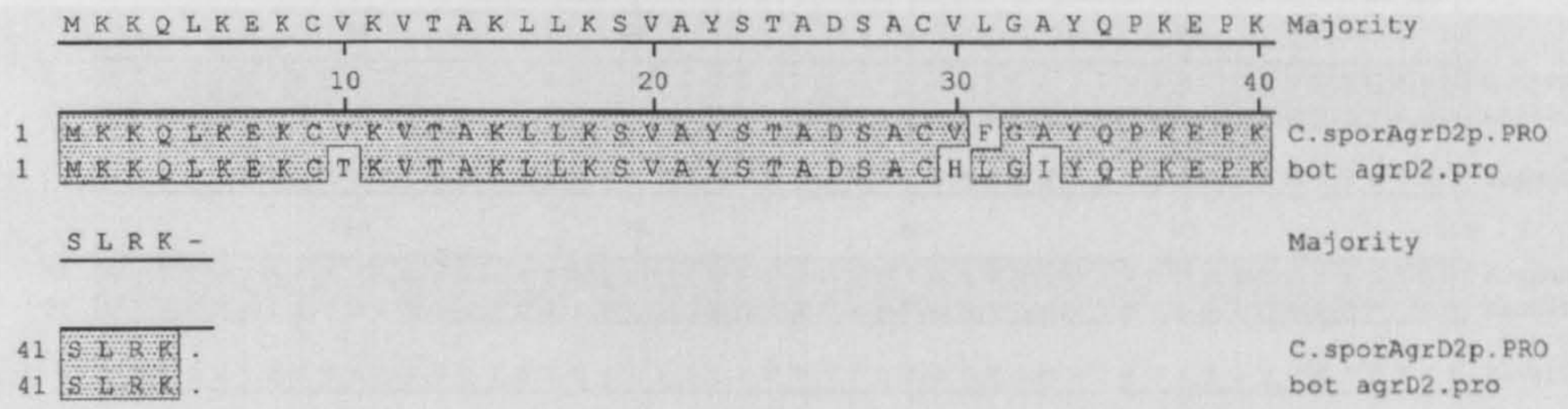
Protein alignment of *C. botulinum* and *C. sporogenes* 0336



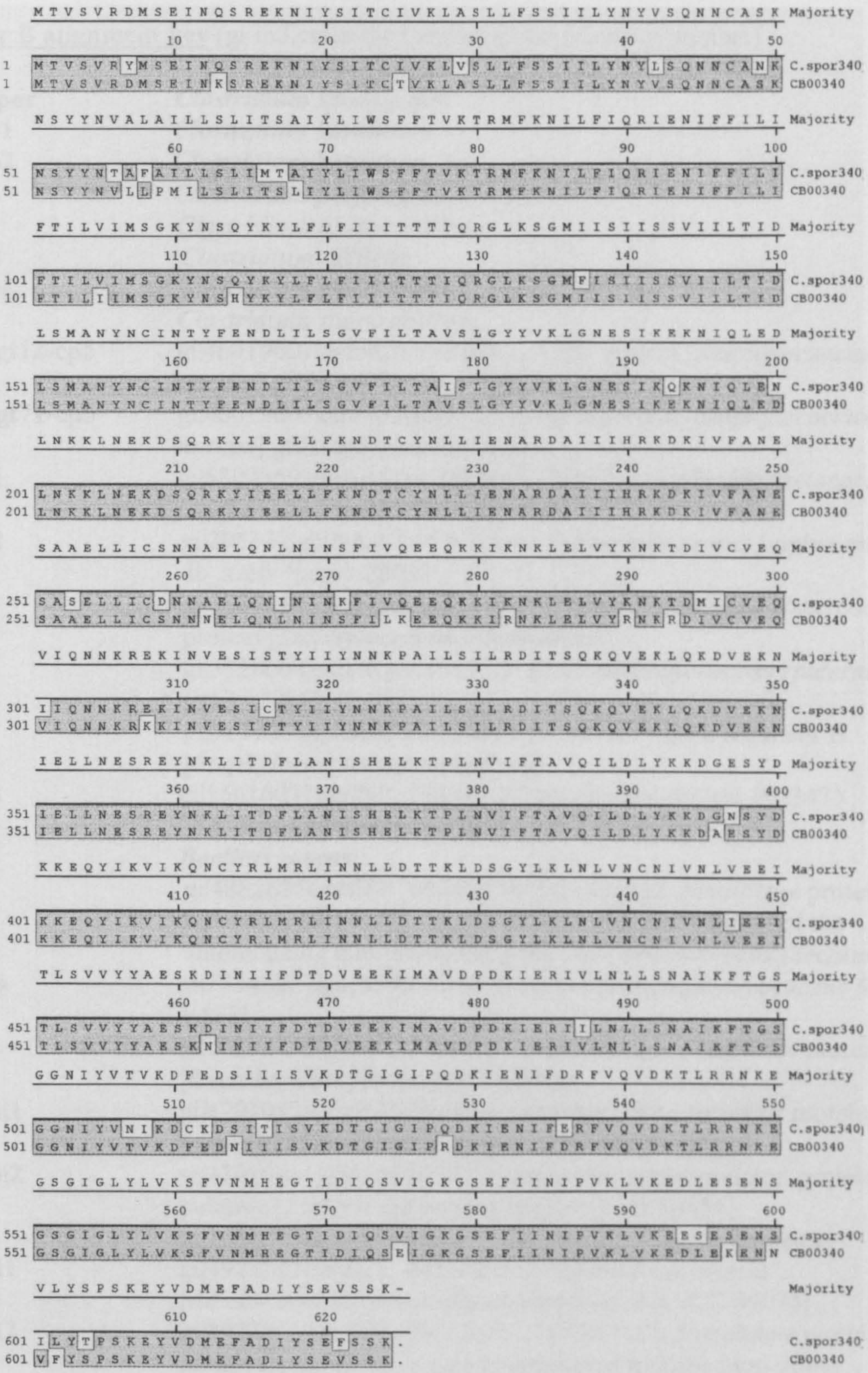
Protein alignment of *C. botulinum* and *C. sporogenes* 0337



Protein alignment of *C. botulinum* and *C. sporogenes* AgrB2



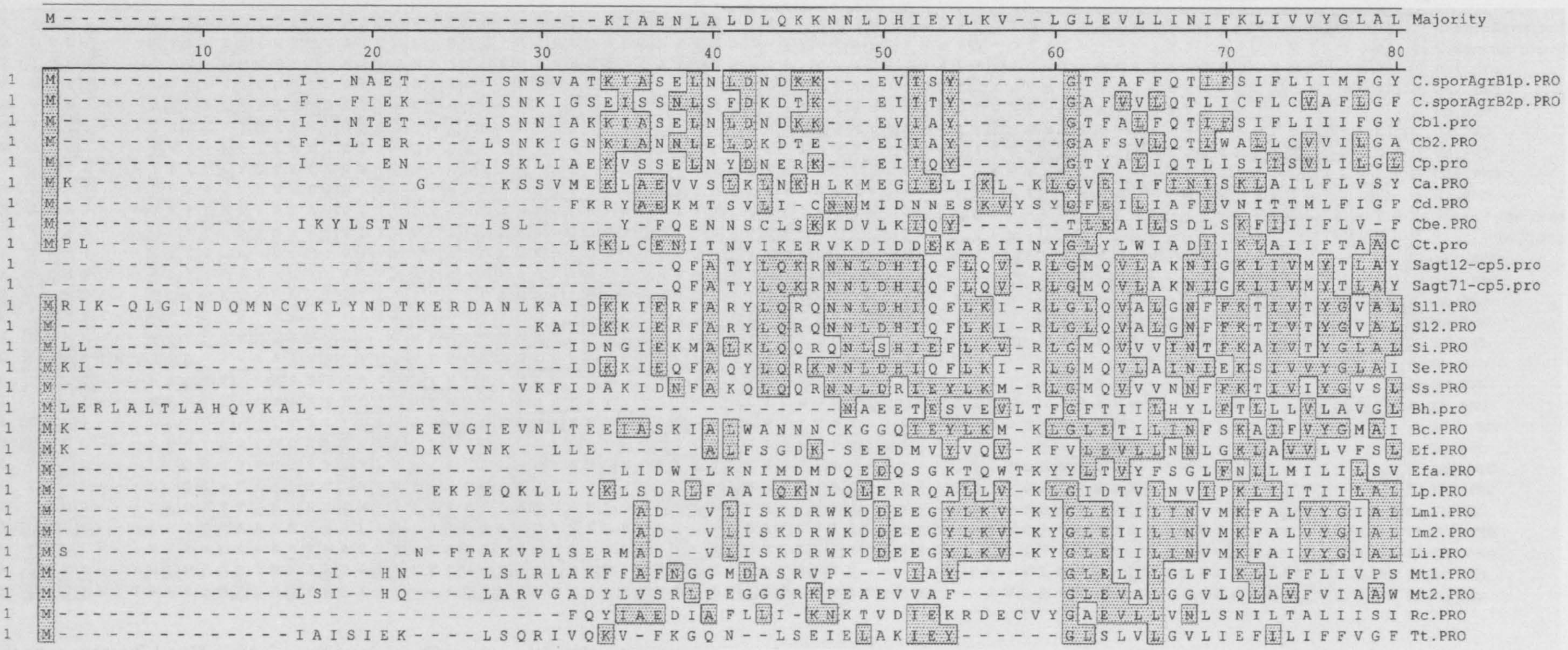
Protein alignment of *C. botulinum* and *C. sporogenes* AgrD2



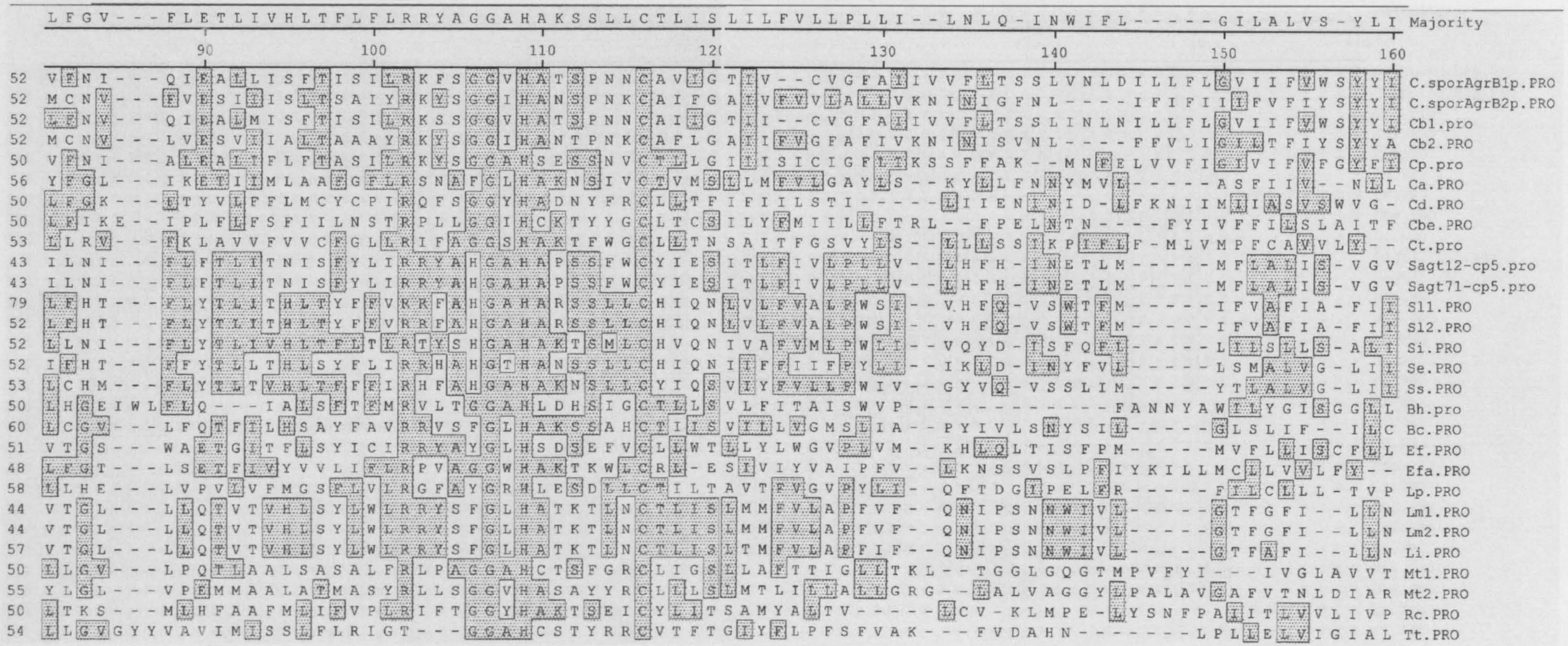
Protein alignment of *C. botulinum* and *C. sporogenes* 0340

Agr B alignment key (gi indicates the GenBank identification number)

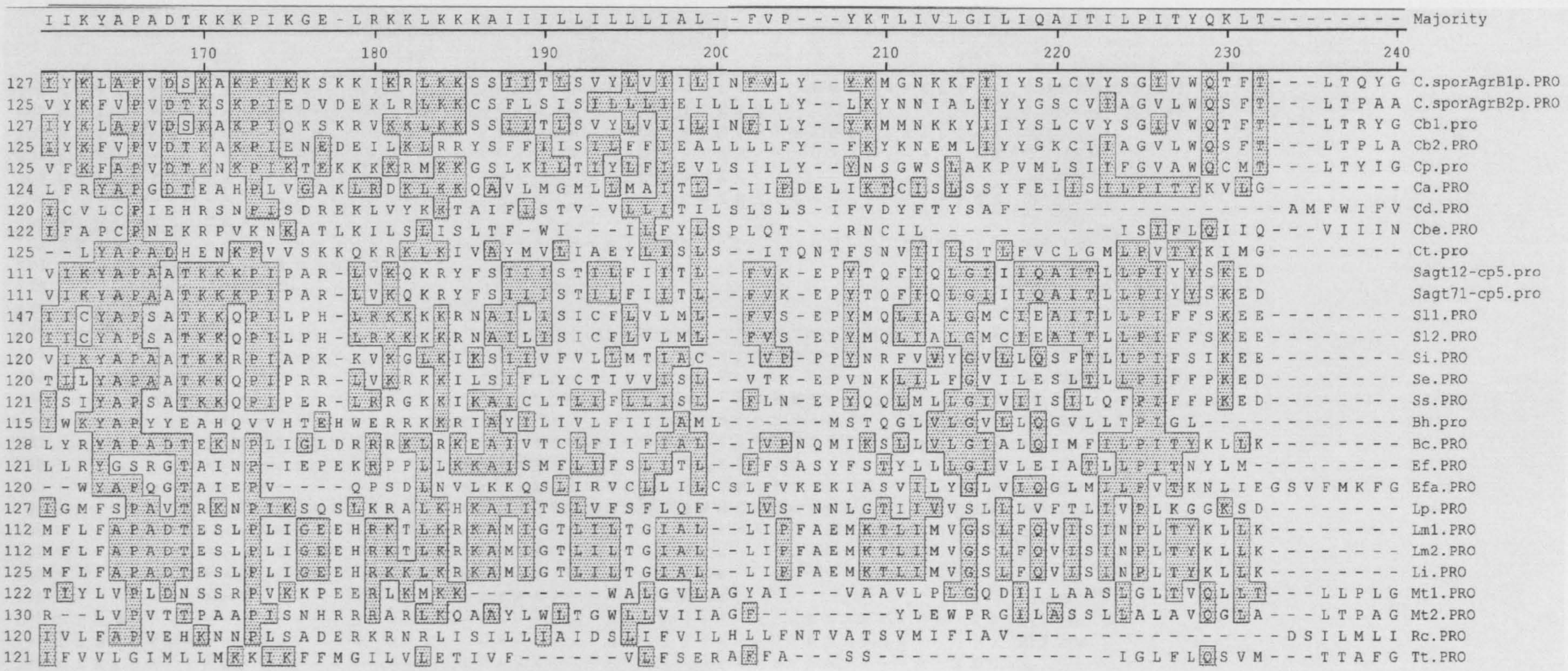
Cspor	***Clostridium sporogenes***
Cb1	***Clostridium botulinum*** 1
Cb2	***Clostridium botulinum*** 2
Cp	***Clostridium perfringens***
Ca	***Clostridium acetobutylicum***
Cd	***Clostridium difficile***
Cbe	***Clostridium beijerinckii***
Ct	***Clostridium thermocellum***
Sagt12-cp5	gi|46019607|emb|CAE92765.1| AgrB protein [***Staphylococcus aureus***] gt12-cp5"isolate="i81
Sagt71-cp5	gi|46019603|emb|CAE92762.1| AgrB protein [***Staphylococcus aureus***] gt71-cp5"isolate="m55
Sl1	gi|5802559|gb|AAD51708.1|AF173933_1 AgrB [***Staphylococcus lugdunensis***] strain="307; RN8160
Sl2	gi|295137|gb|AAA71976.1| ORFD ***Staphylococcus lugdunensis*** db_xref="taxon:28035
Si	gi|45479582|gb|AAS66745.1| putative AIP processing/secretion protein [***Staphylococcus intermedius***]
Se	gi|3320008|emb|CAA89193.1| agrB [***Staphylococcus epidermidis***] strain="0869.12.80"
Ss	gi|51490658|emb|CAG25888.1| Accessory gene regulator B [***Staphylococcus saprophyticus***]
Bh	gi|15616037|ref|NP_244342.1| hypothetical protein BH3475 [***Bacillus halodurans*** C-125]
Bc	***Bacillus cereus***
Ef	gi|48826256|ref|ZP_00287478.1| COG4512: Membrane protein putatively involved in post-translational modification of the autoinducing quorum-sensing peptide [***Enterococcus faecium***]
Efa	gi|29343827|gb|AAO81588.1| agrBfs protein [***Enterococcus faecalis*** V583]
Lp	gi|28272732|emb|CAD65661.1| accesory gene regulator protein B [***Lactobacillus plantarum*** WCFS1]
Lm1	gi|47019375|gb|EAL10116.1| accessory gene regulator protein B, putative [***Listeria monocytogenes*** str. 4b H7858]
Lm2	gi|47016461|gb|EAL07382.1| accessory gene regulator protein B, putative [***Listeria monocytogenes*** str. 1/2a F6854]
Li	gi|16412462|emb|CAC95274.1| lin0041 [***Listeria innocua***]
Mt1	gi|49237216|ref|ZP_00331271.1| hypothetical protein Mther02000212 [***Moorella thermoacetica*** ATCC 39073]
Mt2	gi|49236510|ref|ZP_00330569.1| COG4512: Membrane protein putatively involved in post-translational modification of the autoinducing quorum-sensing peptide [***Moorella thermoacetica*** ATCC 39073]
Rc	***Ruminococcus***
Tt	gi|20515553|gb|AAM23839.1| hypothetical protein TTE0563 [***Thermoanaerobacter tengcongensis*** MB4]



Alignments of AgrB proteins of various Gram positive bacteria



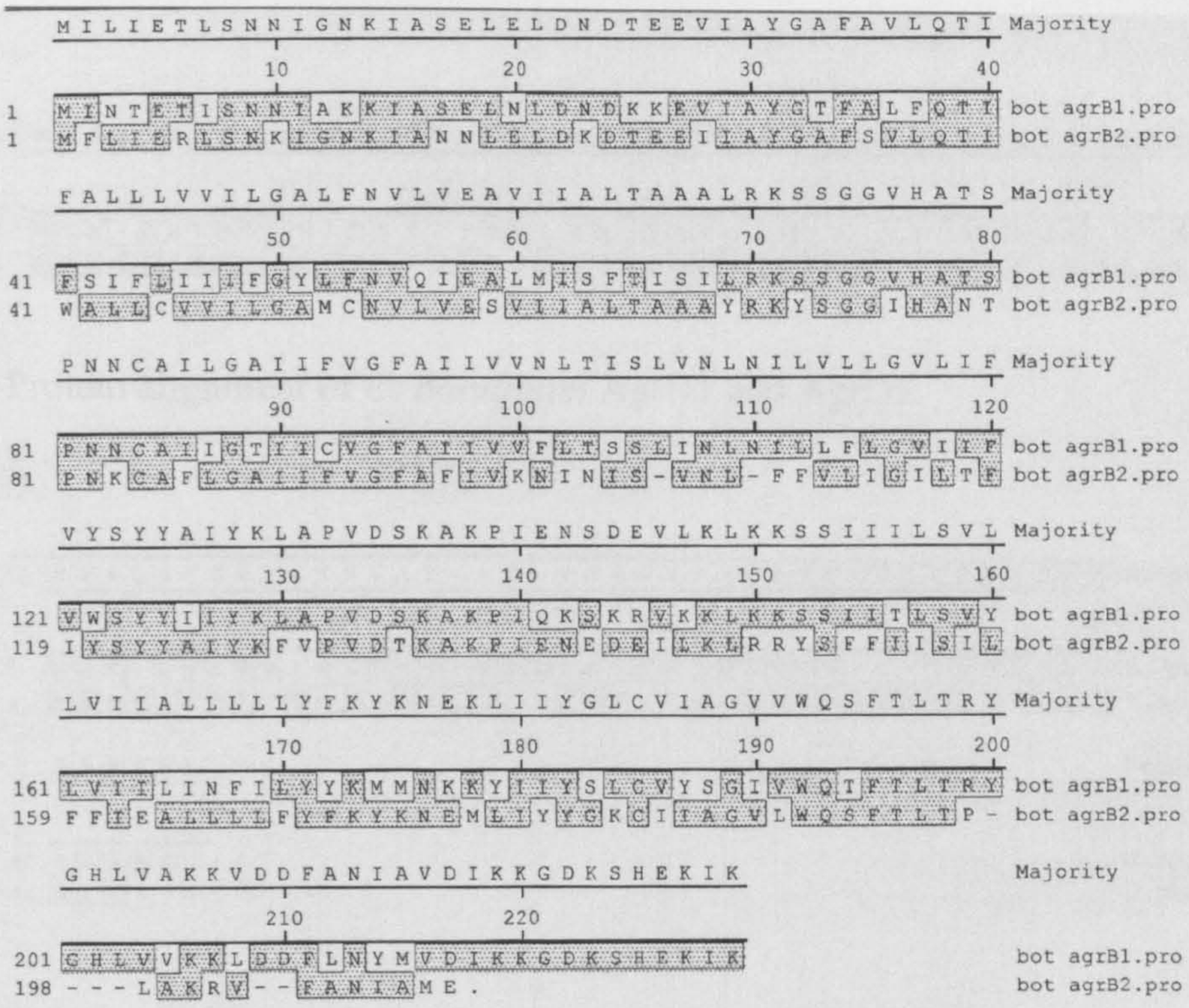
Alignments of AgrB proteins of various Gram positive bacteria continued



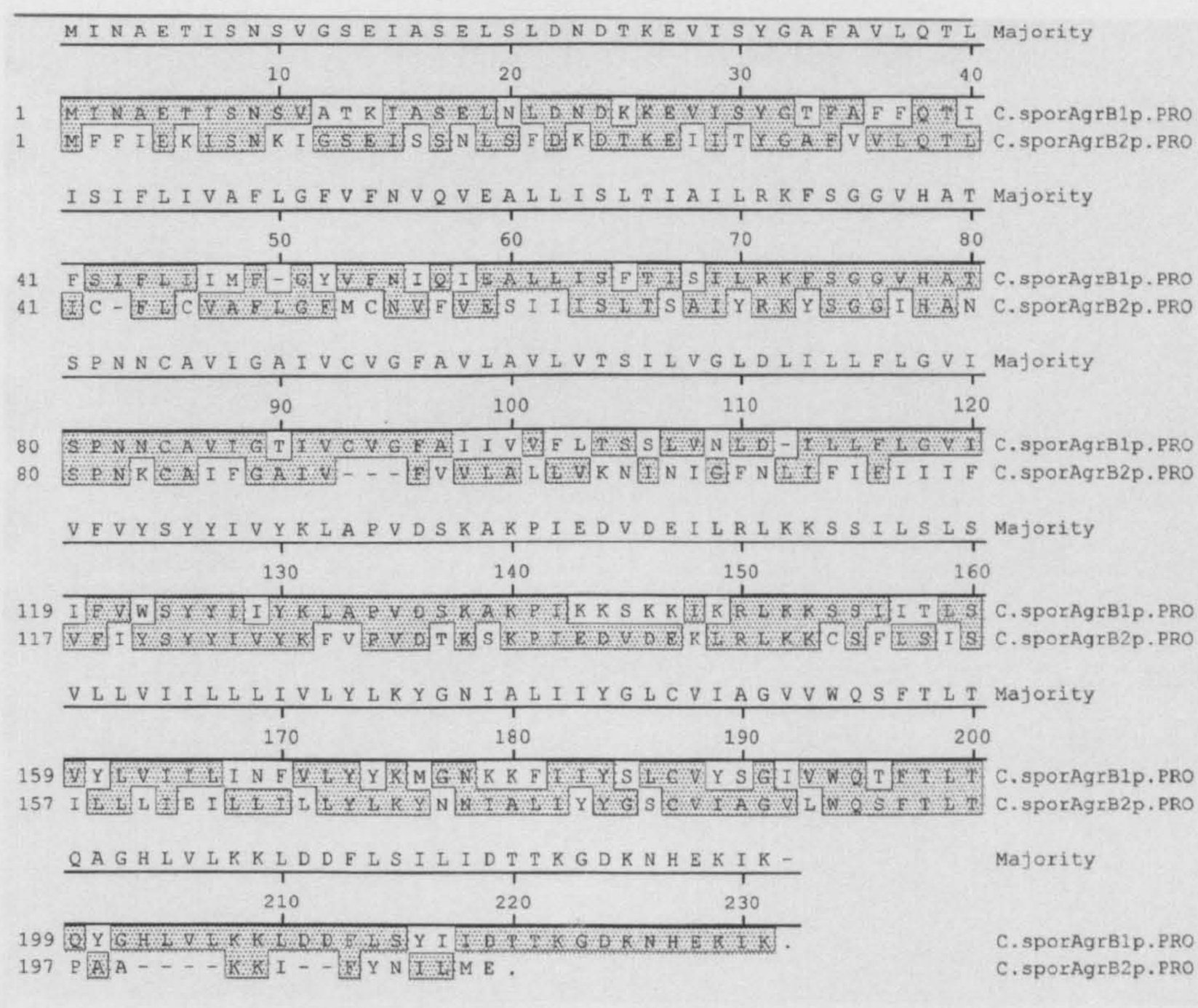
Alignments of AgrB proteins of various Gram positive bacteria continued

```
                                                                                                      Majority
    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - F - - - - - - - - - - - - - - - - -
                    250                 260                 270                 280                 290
202 H L V L K K L - - - - - - - - - - - - - - - - - - - - - - D D F L S Y I I D T T K G D K N H E K I K .   C.sporAgrB1p.PRO
200 K K I - - - - - - - - - - - - - - - - - - - - - - - - - - - - F Y N I L M E .                           C.sporAgrB2p.PRO
202 H L V V K K L - - - - - - - - - - - - - - - - - - - - - - D D F L N Y M V D I K K G D K S H E K I - K   Cb1.pro
200 K R V - - - - - - - - - - - - - - - - - - - - - - - - - - - - F A N I A M - - - - - - - - - - - - - E   Cb2.PRO
200 N I L L K T I - - - - - - - - - - - - - - - - - - - - - - D S F T N K L L                               Cp.pro
194 - - - - - - - - - - - - - - - - - - - - - R R Y K N - - - - - Y Y E F E R T I K - - - - - - - - Q S     Ca.PRO
182 M L V L G K L - K A K V                                                                                Cd.PRO
181 T K G V - - - - - - - - - - - - - - - - - - - - - - - - - - I F N A - - - - - - K N N K T F F N R T T   Cbe.PRO
193 - - - - - - - - - - - A R H G N G L                                                                    Ct.pro
178                                                                                                        Sagt12-cp5.pro
178                                                                                                        Sagt71-cp5.pro
215 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - T     Sl1.PRO
188 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - T     Sl2.PRO
188 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - V                                 Si.PRO
188 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - I N H G K - - - - - - - - - H F   Se.PRO
189 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - Y     Ss.PRO
177 - - - - - - - - - - K V T R Q - - - - - - - - - - - - - - - - L N R F I L K G G E T N E E - - - - N S   Bh.pro
198 - - - - - - - - - - - - - - - - - - - - - R G R N N - - - - - Y E K Y E E E I N - - - - - - - - - G     Bc.PRO
189 - - - - - - - - - - - - - - - - - - - - - - - - - N - - - - - F G G I F H E K N - - - - - - - - - N H   Ef.PRO
194 K K I I K N V I E K R V A K V S D G V G T K P R L N Q N S P N I F G Q W M G Q T E K P K K - - - N I E K   Efa.PRO
196 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - E A E - - - - - - - - - - N V   Lp.PRO
182 - - - - - - - - - - - - - - - - - - - - - R R Y R N - - - - - Y E K Y E                                 Lm1.PRO
182 - - - - - - - - - - - - - - - - - - - - - R R Y R N - - - - - Y E K Y E                                 Lm2.PRO
195 - - - - - - - - - - - - - - - - - - - - - R R Y R N - - - - - Y E K Y E                                 Li.PRO
188 H K A M Y Y L - - - - - - - - - - - - - - - - - - - - - - D Y L L I - - - - - - K L S P H I T R A G R   Mt1.PRO
197 F A A V G W A - - - - - - - - - - - - - - - - - - - - - - D S L L Q - - - - - - R I L P L G             Mt2.PRO
184 G L L I N R I E K S K I - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M E E   Rc.PRO
175 E R L V - - - - - - - - - - - - - - - - - - - - - - - - - - N L - - - - - - - - - A D N L M A K I G I     Tt.PRO
```

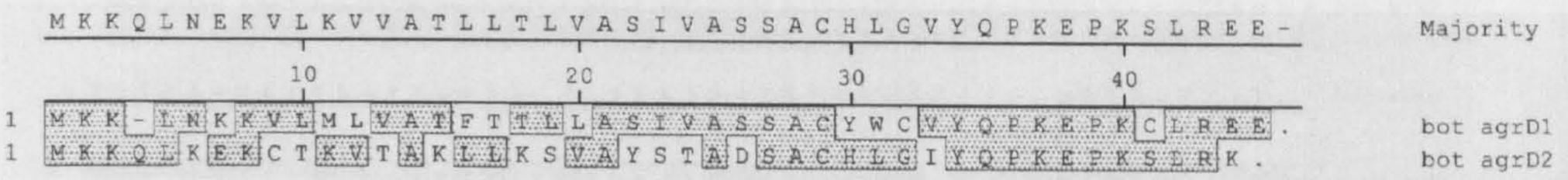
Alignments of AgrB proteins of various Gram positive bacteria continued



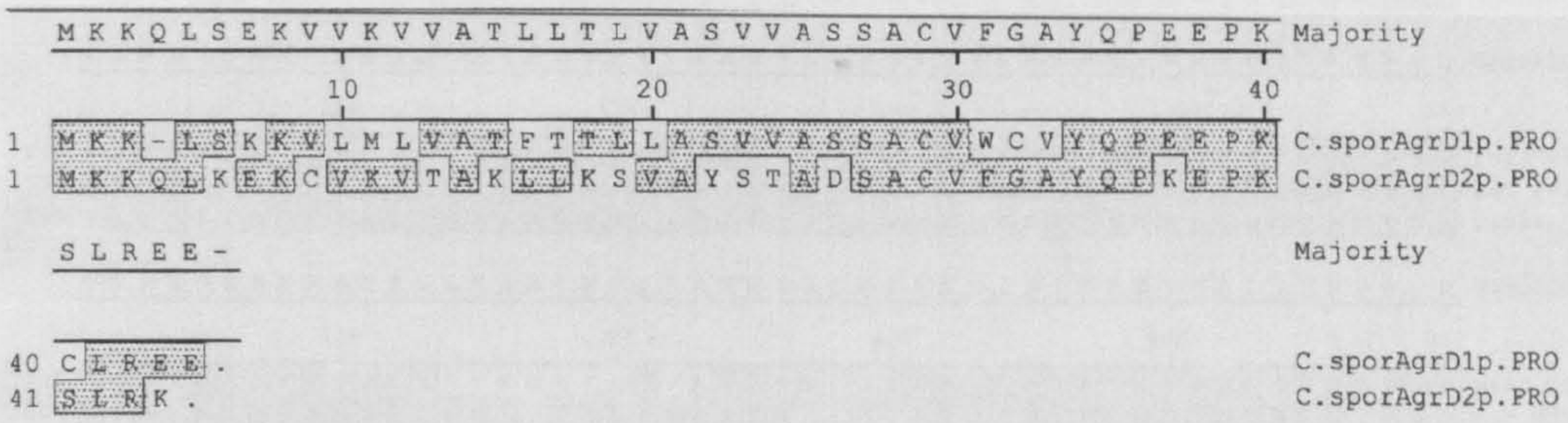
Protein alignment of *C. botulinum* AgrB1 and AgrB2



Protein alignment of *C. sporogenes* AgrB1 and AgrB2

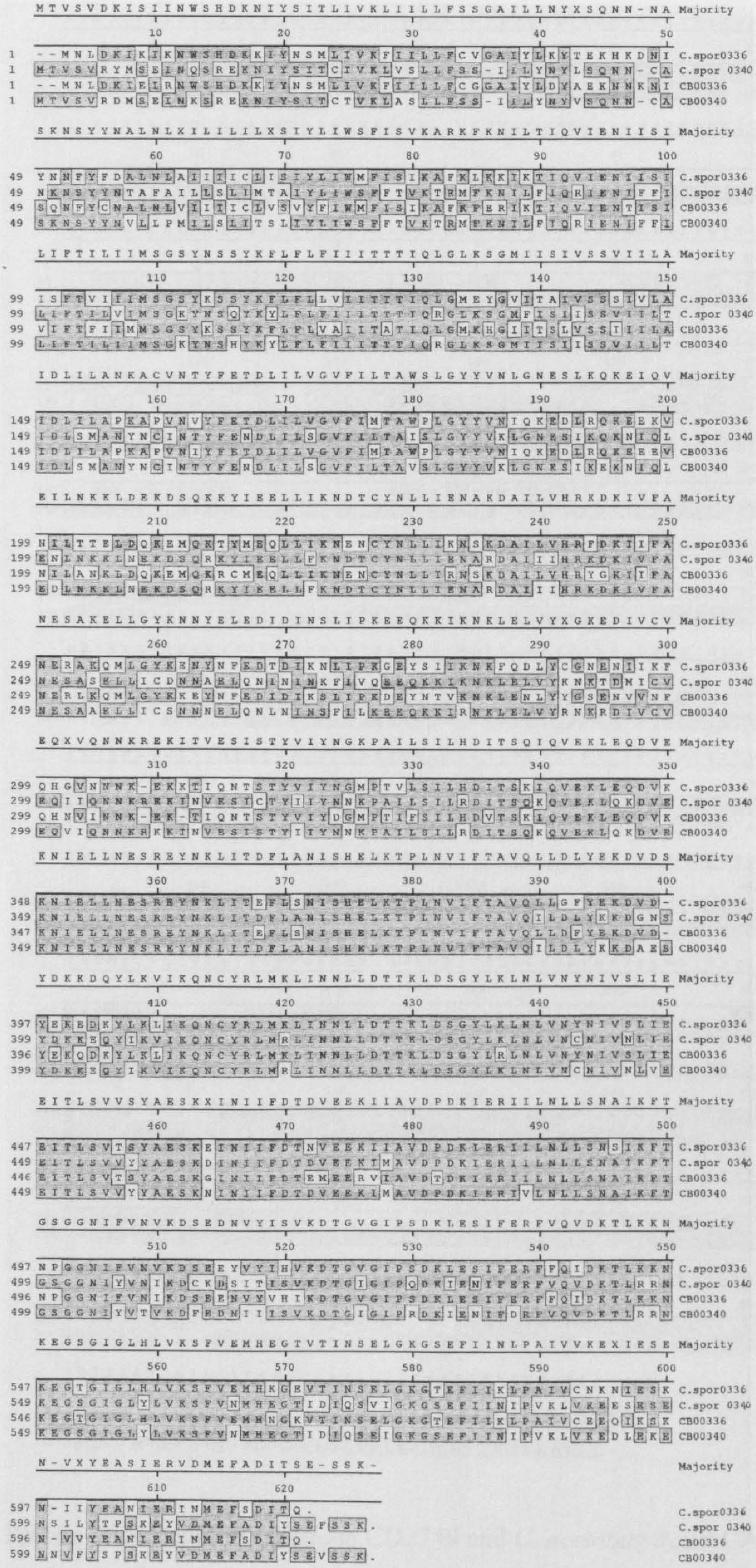


Protein alignment of *C. botulinum* AgrD1 and AgrD2

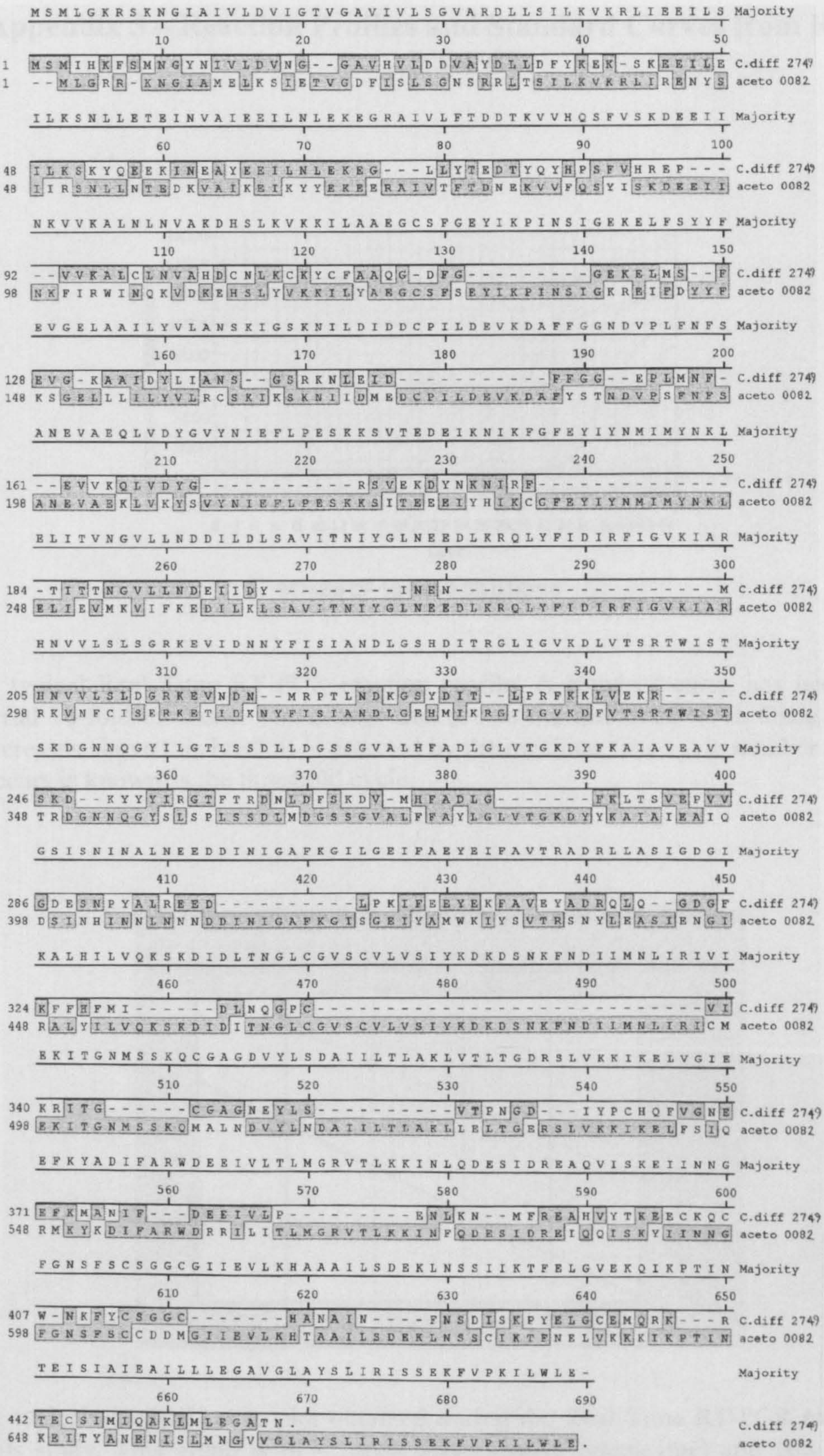


Protein alignment of *C. sporogenes* AgrD1 and AgrD2

Protein alignment of *C. perfringens* CPE1560, *C. botulinum* CBO0333 and *C. sporogenes* CS0333

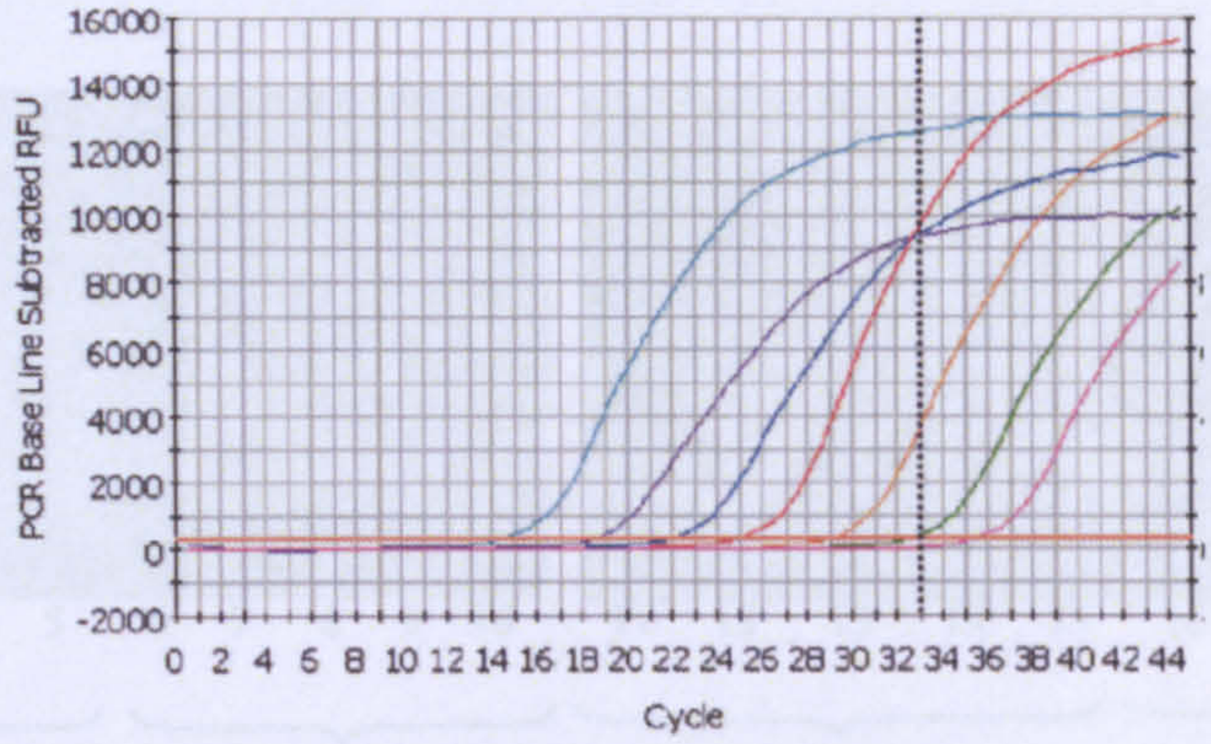


Protein alignment of *C. botulinum* and *C. sporogenes* 0336 and 0340



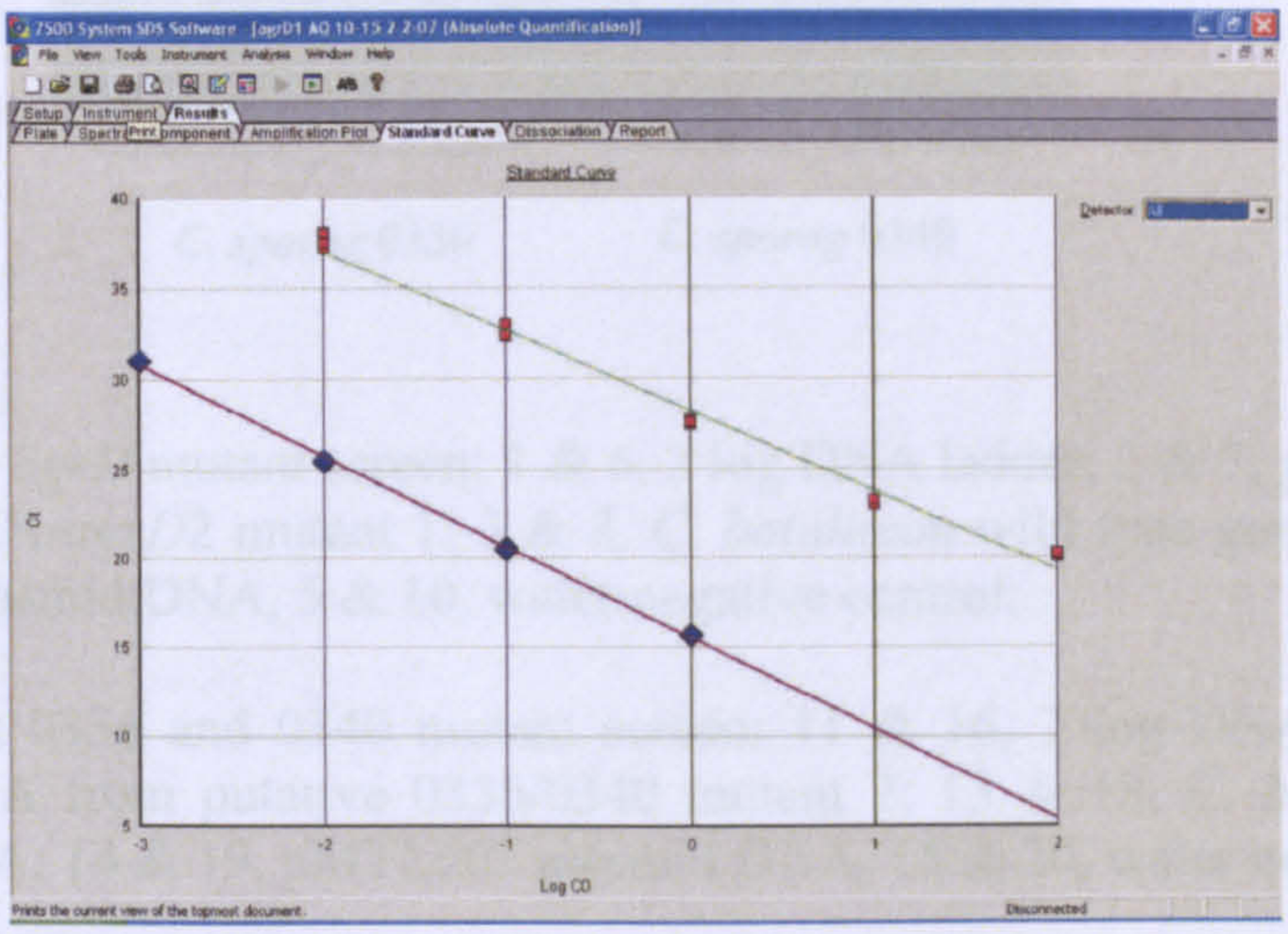
Protein alignment of *C. difficile* CD2749 and *C. acetobutylicum* CAC0082

Appendix 3 – Reaction Profiles and Standard Curves from Real Time RT-PCR



SERIES OF 10-FOLD DILUTIONS

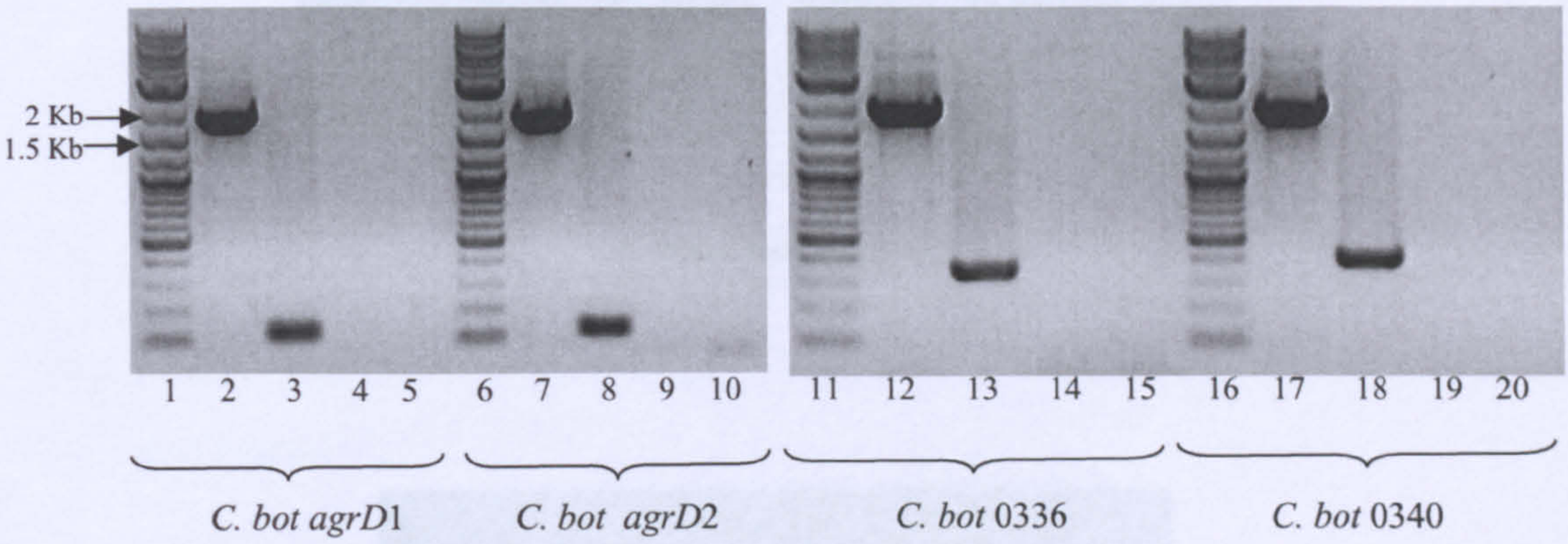
A typical Real Time RT-PCR reaction profile. A standard curve has been run using serial 10-fold dilutions. The orange line shows the threshold level. Once fluorescence increases above this level, it is detected by the machine. The cycle number at which this occurs is known as the threshold cycle.

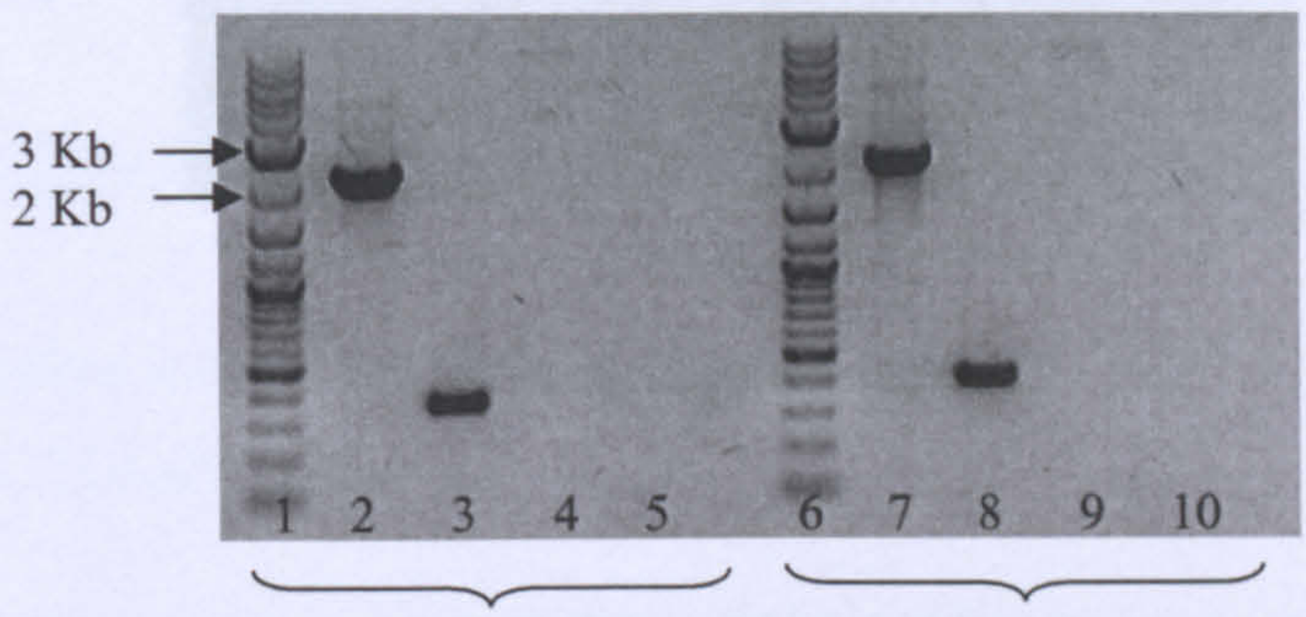


A typical standard curve plot obtained during the Real Time RT-PCR experiments in this study. This graph is of *C. sporogenes agrD*1 (green line) and 16S (purple line) quantitation. A 100% efficient reaction would give a slope of -3.32. In this reaction, *agrD*1 gave a slope of -4.428 and 16S gave a slope of -5.073, indicating an inefficient reaction.

Appendix 4 – PCR Screen Results from Chapter 5 (5.2.2)

PCR Screen 2 on four *C. botulinum* and two *C. sporogenes* mutants



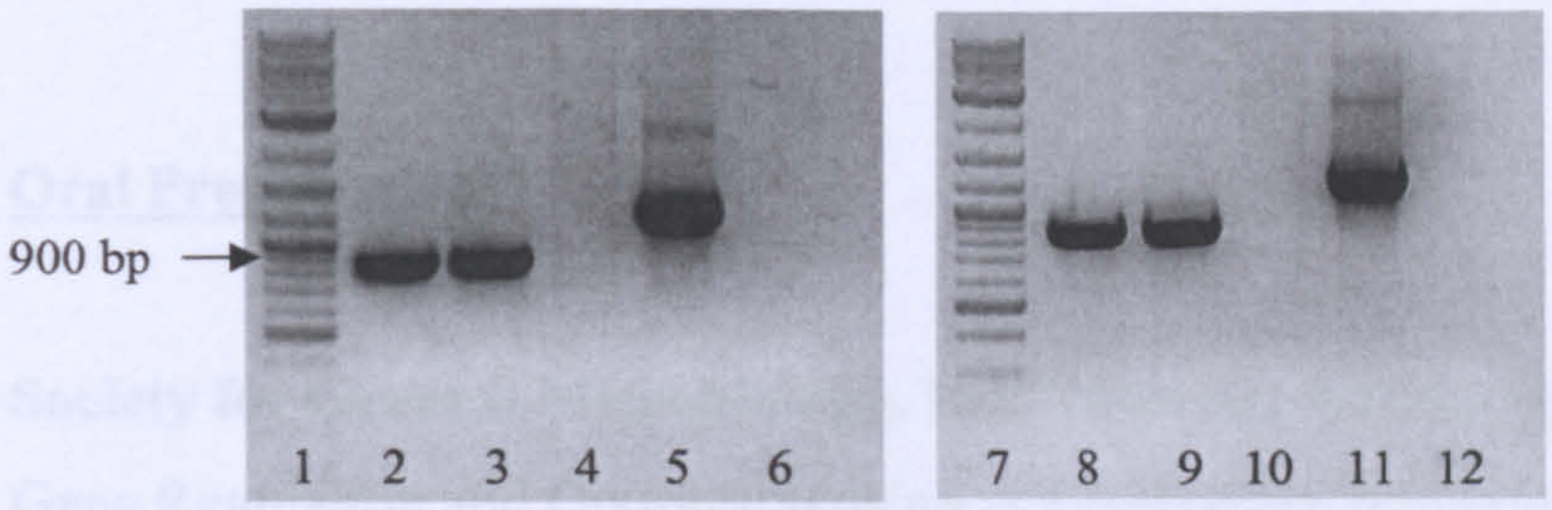


C. botulinum agrD mutant screen: 1 & 6, 2 log DNA ladder; 2 & 7, genomic DNA from putative *agrD*1/*agrD*2 mutant 1; 3 & 8, *C. botulinum* wild type genomic DNA; 4 & 9, pMTL007 plasmid DNA, 5 & 10, water negative control.

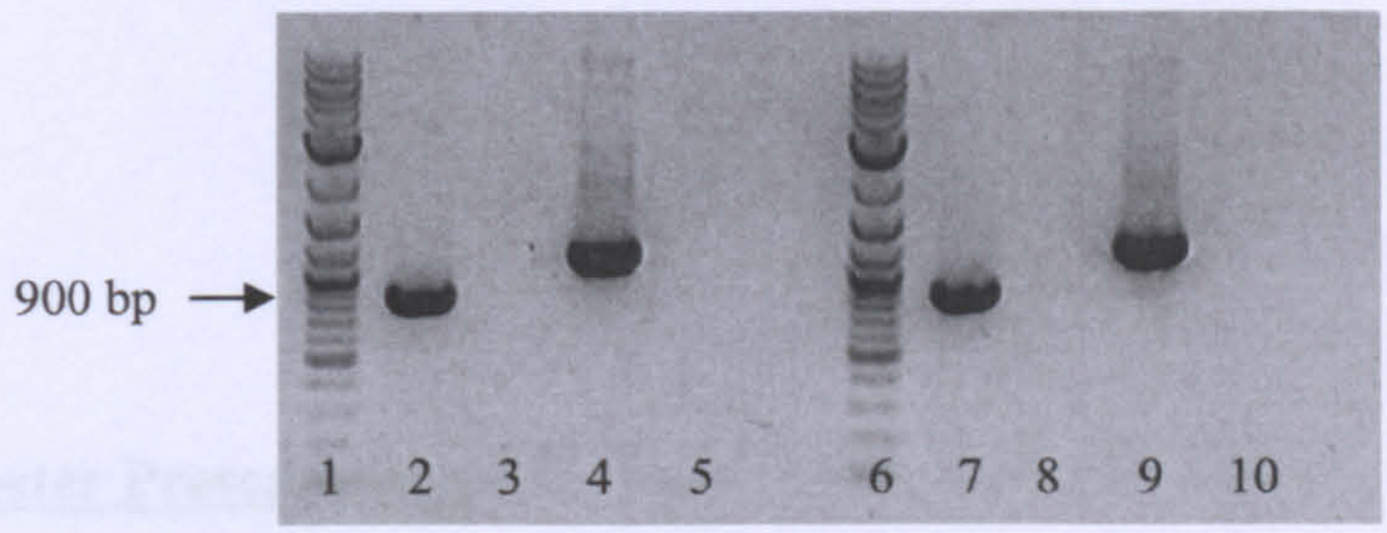
C. botulinum 0336 and 0340 mutant screen: 11 & 16, 2 log DNA ladder; 12 & 17, genomic DNA from putative 0336/0340 mutant 2; 13 & 18, *C. botulinum* wild type genomic DNA; 14 & 19, pMTL007 plasmid DNA, 15 & 20, water negative control.

C. sporogenes 0336 and 0340 mutant screen: 1 & 6, 2 log DNA ladder; 2 & 7, genomic DNA from putative 0336/0340 mutant 1; 3 & 8, *C. sporogenes* wild type genomic DNA; 4 & 9, pMTL007C-E2 plasmid DNA, 5 & 10, water negative control.

PCR Screen 3 on four *C. botulinum* and two *C. sporogenes* mutants



C. botulinum mutants



C. sporogenes mutants

C. botulinum mutants: 1 & 7, 2 log DNA ladder; 2, genomic DNA from putative *agrD*1 mutant 1; 3, genomic DNA from putative *agrD*2 mutant 1; 4 & 10, *C. botulinum* wild type genomic DNA; 5 & 11, pMTL007 plasmid DNA ; 6 & 12, water negative control; 8, genomic DNA from putative 0336 mutant 2; 9, genomic DNA from putative 0340 mutant 2.

C. sporogenes mutants: 1 & 6, 2 log DNA ladder; 2, genomic DNA from putative 0336 mutant 1; 3 & 8, *C. sporogenes* wild type genomic DNA; 4 & 9, pMTL007C-E2 plasmid DNA ; 5 & 10, water negative control; 7, genomic DNA from putative 0340 mutant 1.

Appendix 5 – Presentations

Oral Presentations

Society for General Microbiology, 161st Meeting, Edinburgh, UK

Gene Regulation and Quorum Sensing in *Clostridium botulinum*

Society for General Microbiology, 160th Meeting, Manchester, UK

Characterisation of a putative *agr* system in *Clostridium botulinum* and *Clostridium sporogenes*

Poster Presentations

4th Conference on Functional Genomics of Gram-Positive Micro-organisms, Green Park Resort, Tirrenia, Italy

Clostpath 2006 – 5th International Meeting on the Molecular Biology and Pathogenesis of the Clostridia, University of Nottingham, UK

Bacterial Conversations – Talking, Listening and Eavesdropping, Royal Society, London, UK