

**Structure and Activity Investigations  
of the Cell Fate Determinant, SpoIIE,  
from *Bacillus subtilis*.**

**James Alan Tunaley**

**Thesis Submitted for the Degree of  
Doctor of Philosophy**

University of York

Department of Chemistry

September 2013

## Abstract

For many years the Gram positive bacterium *Bacillus subtilis* has been a model organism for prokaryotic cell and molecular biology. The asymmetric cell division which *B. subtilis* undergoes during sporulation is a simple system by which to study the process of cell differentiation. Sporulation is governed by a series of genetic temporal and spatial controls. Gene regulation brought about by a series of  $\sigma$  factors and transcriptional regulators is coupled to key morphological stages or checkpoints.  $\sigma^F$  initiates the first step in a cascade of complex genetic control which eventually produces a resilient endospore.

The activation of  $\sigma^F$ , the first compartment-specific sigma factor, in the forespore and its regulation through interaction between three proteins; SpoIIAA, SpoIIAB and SpoIIIE, is of particular interest. SpoIIIE, a protein phosphatase which binds to the asymmetric division septum, is a crucial factor in the selective activation of  $\sigma^F$  in the forespore. Of three putative domains in SpoIIIE only the C-terminal PP2C phosphatase domain has been structurally characterised. The central domain, domain II, of SpoIIIE has been assigned a role in interaction with the cell division machinery; however mutational studies have shown that, in addition, this domain is also responsible for the regulation of phosphatase activity.

This work describes the isolation and characterisation of three new fragments of SpoIIIE containing elements of the central cytoplasmic domain of SpoIIIE. These include a fragment found to accurately represent the N-terminal solubility limit of domain II which shows a high degree of oligomeric character. The fragments isolated show specific phosphatase activity against SpoIIAA~P, albeit at reduced rates compared to the free phosphatase domain, which indicates an inhibitory role for SpoIIIE domain II against the PP2C domain. Three ultimately unsuccessful approaches were attempted to isolate a co-complex of SpoIIIE and SpoIIAA~P for structural characterisation. A tendency for domain II-containing SpoIIIE fragments to precipitate in the presence of  $Mn^{2+}$  is also identified. An *in vivo* investigation into the sporulation efficiencies of amino acid substitutions in a potential regulatory interface between domains II and III of SpoIIIE indicated no strong sporulation defects.

# List of Contents

Abstract.....	1
List of Contents.....	2
List of Tables .....	7
List of Figures.....	8
Acknowledgements .....	12
Author's Declaration .....	13
Chapter 1: Introduction .....	14
1.1 Spore formation in Prokaryotes.....	14
1.2 <i>Bacillus</i> in industry and medicine .....	15
1.2.1 Risks posed by endospore-forming bacteria.....	17
1.3 <i>Bacillus subtilis</i> .....	18
1.4 Sporulation .....	21
1.4.1 Entry into the spore formation pathway .....	21
1.4.1.1 AbrB.....	22
1.4.1.2 SigH.....	24
1.4.1.3 CodY.....	24
1.4.2 Spo0A.....	26
1.4.3 The sporulation phosphorelay.....	29
1.5 Sigma factors.....	32
1.6 Morphological stages of sporulation.....	35
1.6.1 Coupling cell cycle to sporulation.....	36
1.6.2 Sporulation stage I .....	37
1.6.3 Sporulation stage II.....	38
1.6.4 Sporulation stage III .....	40
1.6.5 Sporulation stage IV.....	41
1.6.6 Sporulation stage V .....	42
1.6.7 Sporulation stage VI.....	43
1.6.8 Sporulation stage VII.....	43
1.7 The Compartment Specific $\sigma$ Factor Cascade .....	44
1.8 Activation of the First Compartment Specific $\sigma$ Factor, $\sigma^F$ .....	48
1.8.1 Effectors of $\sigma^F$ regulation.....	48
1.8.1.1 SpoIIAB.....	49
1.8.1.2 SpoIIAA.....	51

1.8.1.3 SpoIIE.....	53
1.9 The Z-ring and Asymmetric Cell Division .....	54
1.9.1 FtsZ.....	54
1.9.2 The <i>B. subtilis</i> divisome.....	57
1.9.3 Regulating Z-ring assembly during growth .....	58
1.9.4 Relocation of the division site during sporulation.....	59
1.10 SpoIIE, a Key Cell Fate Determinant.....	61
1.10.1 Developmental behaviour of SpoIIE .....	61
1.10.2 Three Domain structure.....	61
1.10.3 How compartment specific activation of $\sigma^F$ is achieved .....	63
1.11 Aims .....	66
Chapter 2: Materials and Methods .....	68
2.1 Agarose gel electrophoresis.....	68
2.2 Small Scale DNA Preparation .....	69
2.3 Construct preparation.....	70
2.3.1 Ligation Independent Cloning.....	70
2.3.2 Polymerase chain reaction.....	71
2.3.3 Oligonucleotide primers .....	73
2.3.4 LIC Vector preparation.....	75
2.3.4.1 pET-YSBLIC3C Preparation by BseRI digest.....	76
2.3.4.2 pET-YSBLIC3C Preparation by vector PCR.....	76
2.3.5 T4 DNA Polymerase Treatment.....	79
2.3.6 Annealing and transformation .....	80
2.3.6.1 Generation of TSS competent <i>E. coli</i> cells.....	81
2.3.7 Assessing Cloning Success .....	81
2.4 Site directed mutagenesis.....	83
2.4.1 Oligonucleotide Mutagenesis Primers .....	83
2.4.2 The Mutagenesis Reaction .....	83
2.4.3 DpnI Treatment.....	85
2.5 Protein Expression .....	87
2.5.1 Culture Growth .....	87
2.5.2 Induction of Protein Expression by IPTG .....	87
2.5.3 Harvesting Cells and Lysis by Sonication .....	89
2.5.4 Optimisation of Expression and Lysis Conditions.....	89
2.6 Protein Purification .....	91

2.6.1 Immobilised Metal Affinity Chromatography .....	91
2.6.2 HRV3C Protease Cleavage and Removal of 6xHis Affinity Tag .....	93
2.6.3 Size Exclusion Chromatography .....	94
2.6.4 Ion Exchange Chromatography .....	96
2.7 Protein Analysis .....	97
2.7.1 Manipulation of Protein Concentration .....	97
2.7.2 Determination of Protein Concentration .....	97
2.7.3 Polyacrylamide Gel Electrophoresis .....	98
2.7.3.1 Sodium Dodecyl Sulphate (SDS) PAGE .....	98
2.7.3.2 Native PAGE .....	99
2.7.4 Circular Dichroism .....	101
2.7.5 SEC-MALLS .....	102
2.8 Protein Crystallisation .....	104
2.8.1 Crystallisation Screening .....	105
2.8.2 Crystallisation Optimisation .....	106
Chapter 3: Investigations into SpoIIE fragments containing elements of the FtsZ binding domain. ....	107
3.1 Introduction .....	107
3.1.1 Previous approaches to soluble SpoIIE .....	107
3.2 Results .....	110
3.2.1 Isolation and Analysis of the B2-B1 fragment: SpoIIE <sub>375-590</sub> .....	110
3.3.3 Identification of the TM:FtsZ binding domain boundary .....	118
3.3.3 Analogue of SpoIIE from <i>Geobacillus stearothermophilus</i> .....	127
3.3.4 SpoIIE <sub>334-827</sub> fragment oligomerisation. ....	131
3.3.5 Possible homology of SpoIIE to the sporulation stage III protein, SpoIIIAH136 .....	140
3.2.6 The effect of Mn <sup>2+</sup> on SpoIIE Domain II fragments. ....	140
3.3 Summary and Discussion .....	143
Chapter 4: Activity-based and Structural Investigations into the Role of SpoIIE as a PP2C Phosphatase .....	145
4.1 Introduction .....	145
4.1.1 Serine/Threonine Phosphatases .....	145
4.1.2 The SpoIIE PP2C domain structure .....	146
4.1.3 The SpoIIE:SpoIIAA~P interaction .....	149
4.2 Methods .....	150
4.2.2 Native PAGE Phosphatase State Assay .....	150
4.2.3 Generation of Transition State Analogues .....	150

4.2.4 Thermafluor Assays.....	150
4.2.5 Silver Staining of PAGE gels.....	151
4.3 Results .....	152
4.3.1 Cloning, Expression and Purification of SpoIIAA and SpoIIAA~P with SpoIIAB .....	152
4.3.2.1 Purification of SpoIIAA .....	154
4.3.2.2 Purification of SpoIIAA~P and SpoIIAB.....	157
4.3.2.3 Analysis of Purified SpoIIAA, SpoIIAA~P and SpoIIAB Proteins.....	158
4.3.2 The SpoIIAA~P Dephosphorylation Reaction .....	160
4.3.2.1 Cognate vs. Orthologous Substrate .....	160
4.3.2.2 Metal Dependence of the SpoIIE PP2C domain. ....	162
4.3.2.3 Phosphatase activity of SpoIIE fragments. ....	164
4.3.3 Towards the Structural Elucidation of the SpoIIE:SpoIIAA Interaction .....	167
4.3.3.1 Phosphate Transition State Analogues .....	167
4.3.3.2 Phosphoserine mimics: Mutation of SpoIIAA.....	170
4.3.3.3 Use of an Inactive Active Site Metal to Trap a Pre-turnover Complex...	174
4.4 Summary and Discussion.....	179
Chapter 5: <i>In vivo</i> Investigation of SpoIIE Mutations; Regulation of Asymmetric Cell Division and Phosphatase Activity by SpoIIE.....	181
5.1 Introduction .....	181
5.1.1 Phenotypes of mutant <i>spoIIE</i> alleles.....	181
5.1.2 Competence in <i>B. subtilis</i> .....	186
5.1.2.1 DNA Uptake Machinery .....	187
5.1.2.2 Post-transformational Treatment of Exogenous DNA.....	188
5.1.2.3 Modes of Homologous Recombination .....	189
5.2 Methods.....	192
5.2.1 Generation of <i>B. subtilis</i> Competent Cells .....	192
5.2.2 The pUK-19 Integration Vector .....	193
5.2.3 Transformation of competent <i>B. subtilis</i> .....	194
5.2.4 Strains.....	194
5.2.5 Sporulation Efficiency Assay.....	195
5.3 Results .....	197
5.3.1 The Structural Context of Intragenic Suppressor Mutations in the SpoIIE Phosphatase Domain.....	197
5.3.2 Cloning of SpoIIE into an Integration Plasmid .....	198
5.3.3 Preparation of <i>B. subtilis</i> SpoIIE Mutant strains .....	204

5.3.4 Determination <i>spoIIE</i> Mutant Sporulation Efficiency.....	206
5.4 Discussion .....	209
Chapter 6: Conclusions and Future Work .....	210
6.1 The Behaviour of SpoIIE Domain II .....	210
6.2 SpoIIE Phosphatase Activity.....	211
6.3 An Interface for Intragenic Regulation of the Phosphatase Domain .....	212
6.4 How PP2C regulation might be achieved .....	213
6.5 Future work.....	214
Appendix I .....	217
The pET-28a Vector Map.....	217
Fragment Parameters (calculated using ProtParam (ExPASy)) .....	218
SpoIIE Secondary and Tertiary Structure Predictions.....	219
PsiPred (Secondary Structure) .....	219
GlobPlot 2.3 (Globular Domains and Transmembrane Helices).....	222
TMHMM predictor (Transmembrane Helices).....	222
Oligonucleotide Primer Sequences .....	223
Cloning .....	223
Mutagenesis .....	224
Solutions for generation of <i>B. subtilis</i> Competent cells.....	225
T base.....	225
SpC.....	225
SpII.....	225
Liquid Difco Sporulation Medium (DSM) a.k.a. Schaeffer's Sporulation Medium .....	226
List of abbreviations.....	227
References .....	230

## List of Tables

Table 1-1: Sporulation related $\sigma$ factors of <i>Bacillus subtilis</i> .....	33
Table 2-1: Required PCR components.....	72
Table 2-2: Typical PCR thermocycling parameters.....	73
Table 2-3: pET-YSBLIC3C linearisation by PCR.....	78
Table 2-4: T4 DNA Polymerase Treatment Reactions for LIC .....	79
Table 2-5: Reaction Parameters for Site Directed Mutagenesis .....	84
Table 2-6: Components for Polyacrylamide Gel Electrophoresis.....	100
Table 3-1: Buffers used in solubility screening of SpoIIE <sub>375-590</sub> .....	110
Table 5-1: <i>B. subtilis</i> strains used in this work.....	194

## List of Figures

Figure 1-1: Spores as a vehicle for targeted drug delivery. ....	16
Figure 1-2: Distribution of Gene Functions in <i>B. subtilis</i> .....	18
Figure 1-3: Example of a <i>B. subtilis</i> growth profile.....	21
Figure 1-4: Schematic representation of key parts of the AbrB regulon.....	23
Figure 1-5: Overlay of the CodY GAF domain in unliganded and ligand bound states...	25
Figure 1-6: DNA binding mode of Spo0A~P.....	27
Figure 1-7: The sporulation phosphorelay.....	29
Figure 1-8: Structures of the <i>Thermus aquaticus</i> ( <i>Taq</i> ) RNAP holoenzyme with (B) and without (A) a visible $\Delta 1.1$ truncation of <i>Taq</i> housekeeping $\sigma$ factor, $\sigma^A$ . ....	32
Figure 1-9: The morphological stages of sporulation. ....	35
Figure 1-10: Chromosomal organisation in pre-divisional vegetative and sporulating cells.....	37
Figure 1-11: Sporulation stage II. ....	39
Figure 1-12: Sporulation stage III.....	41
Figure 1-13: Late sporulation morphological developments. ....	42
Figure 1-14: The cascade of sporulation specific $\sigma$ factors.....	44
Figure 1-15: Transcriptional control of the $\sigma^F$ regulatory system.....	48
Figure 1-16: The SpoIIAB <sub>2</sub> : $\sigma^F$ crystal structure from <i>B. stearrowthermophilus</i> , PDB: 1L00. ....	50
Figure 1-17: C $\alpha$ trace superposition of SpoIIAA from <i>B. sphaericus</i> in both phosphorylated and unphosphorylated forms.....	50
Figure 1-18: The SpoIIAA <sub>2</sub> :SpoIIAB <sub>2</sub> complex from <i>B. stearrowthermophilus</i> . ....	51
Figure 1-19: Schematic of the induced release of $\sigma^F$ from SpoIIAB <sub>2</sub> : $\sigma^F$ by SpoIIAA. ....	52
Figure 1-20: Structure of FtsZ from <i>Methanococcus jannaschii</i> . ....	55
Figure 1-21: Functional behaviour of FtsZ proto-filaments.....	56
Figure 1-22: The <i>B. subtilis</i> divisome. ....	57
Figure 1-23: Direction of Z-ring formation to mid-cell in vegetative cell division by MinCD and Noc. ....	58
Figure 1-24: The three domain structure of the sporulation cell fate determinant, SpoIIE. ....	62
Figure 1-25: The state of $\sigma^F$ regulation equilibria in the sporangium before and after asymmetric septation.....	63
Figure 1-26: Possible mechanisms for the establishment of compartment specific gene expression directed by $\sigma^F$ .....	64
Figure 2-1: Ligation Independent Cloning.....	70

Figure 2-2: The pET- YSBLIC3C LIC cloning site .....	77
Figure 2-3: Site Directed Mutagenesis procedure. ....	85
Figure 2-4: The pET expression system.....	88
Figure 2-5: Purification by Nickel IMAC. ....	92
Figure 2-6: The Principle of Size Exclusion Chromatography. ....	95
Figure 2-7: Example CD spectra of protein secondary structure. ....	101
Figure 2-8: The crystallisation phase diagram.....	104
Figure 2-9: Sitting and hanging drop vapour diffusion experiments. ....	105
Figure 3-1: Soluble SpoIIE constructs identified by ESPRIT. ....	108
Figure 3-2: SDS-PAGE gel images showing total and soluble cell extracts from <i>E. coli</i> BL21 pLysS cultures harbouring the pYSBLIC3c_SpoIIE <sub>375-590</sub> plasmid. ....	111
Figure 3-3: SDS-PAGE gel image showing IMAC purification of the SpoIIE <sub>375-590</sub> .....	112
Figure 3-4: Size Exclusion purification of SpoIIE <sub>375-590</sub> peaks A (blue trace) and B (green trace). ....	113
Figure 3-5: SDS-PAGE gel image of purified aliquots of SpoIIE <sub>375-590</sub> . ....	114
Figure 3-6: SDS-PAGE gel image of HRV 3C cleavage of the 6xHis affinity tag of SpoIIE <sub>375-590</sub> .....	115
Figure 3-7: Size Exclusion purification of SpoIIE <sub>375-590</sub> after 6xHis affinity tag cleavage. .....	116
Figure 3-8: SEC-MALLS traces of SpoIIE fragments.....	117
Figure 3-9: Predictions of properties of SpoIIE in the 240 to 360 region with primary amino acid sequence. ....	119
Figure 3-10: PCR products of <i>spoIIE</i> amplification separated by electrophoresis on 1% agarose gels.....	120
Figure 3-11: SDS-PAGE images of SpoIIE fragment expression tests.....	121
Figure 3-12: The TM domain : Domain II boundary region.....	122
Figure 3-13: Size exclusion A <sub>280</sub> chromatograms of the SpoIIE <sub>334-827</sub> and SpoIIE <sub>339-827</sub> fragments with SDS-PAGE gel images.....	123
Figure 3-14: Primary sequence alignment of SpoIIE from <i>B. subtilis</i> and <i>G.</i> <i>stearothermophilus</i> . ....	127
Figure 3-15: Gel images of <i>GstSpoIIE</i> <sub>335-826</sub> cloning and expression.....	128
Figure 3-16: Size Exclusion purification of <i>GstSpoIIE</i> <sub>335-826</sub> . ....	129
Figure 3-17: Parallel purification of <i>BsuSpoIIE</i> <sub>334-827</sub> and <i>GstSpoIIE</i> <sub>335-826</sub> fragments on Superdex 200 gel filtration columns. ....	131
Figure 3-18: Analytical Ultra-Centrifugation analyses of the <i>BsuSpoIIE</i> <sub>334-827</sub> and <i>GstSpoIIE</i> <sub>335-826</sub> fragments at A <sub>280</sub> = 1.0, 0.3 and 0.1.....	133
Figure 3-19: SEC-MALLS analyses of <i>BsuSpoIIE</i> <sub>334-827</sub> and <i>GstSpoIIE</i> <sub>335-826</sub> . ....	134
Figure 3-20: Primary sequence alignment of SpoIIAH <sub>78-218</sub> and <i>BsuSpoIIE</i> <sub>457-590</sub> . ....	136

Figure 3-21: Images of 1% agarose gels of PCR amplifications of the <i>BsuSpoIIE</i> <sub>457-827</sub> coding nucleotide sequence visualised using SybrSafe DNA dye.....	137
Figure 3-22: Purification of the <i>BsuSpoIIE</i> <sub>457-827</sub> fragment. ....	138
Figure 3-23: Size exclusion chromatography of <i>GstIIE</i> <sub>457-826</sub> .....	140
Figure 3-24: SDS-PAGE analysis SpoIIE fragment pelleting on addition of MnCl <sub>2</sub> .....	141
Figure 3-25: Circular dichroism spectra of SpoIIE fragments on addition of Mn <sup>2+</sup> . ....	142
Figure 4-1: The phosphate containing active site of human PP2C $\alpha$ phosphatase. ....	146
Figure 4-2: Structural comparisons between SpoIIE <sub>590-827</sub> (PDB ID: 3T9Q) and <i>MtPstP</i> (PDB ID: 1TXO) PP2C $\alpha$ phosphatases.....	147
Figure 4-3: Structures of the SpoIIE <sub>590-827</sub> fragment. ....	148
Figure 4-4: Digests of putative <i>spoIIAA</i> and <i>spoIIAA-spoIIAB</i> containing plasmids. ....	153
Figure 4-5: SpoIIAA and SpoIIAA+SpoIIAB expression tests.....	154
Figure 4-6: Nickel IMAC purification of SpoIIAA.....	155
Figure 4-7: Size exclusion chromatograms showing purifications of SpoIIAA (red) and SpoIIAA~P (blue) as separated by a Superdex 75 16/60 gel filtration column.....	156
Figure 4-8: Size exclusion chromatogram of SpoIIAB.....	158
Figure 4-9: SDS- PAGE and Native PAGE analysis of SpoIIAA and SpoIIAB species from both <i>B. subtilis</i> and <i>B. sphaericus</i> .....	159
Figure 4-10: Native PAGE analysis of SpoIIAB kinase activity assay.....	159
Figure 4-11: Comparison of SpoIIE catalysed dephosphorylation of SpoIIAA~P from <i>B. subtilis</i> and <i>B. sphaericus</i> . ....	160
Figure 4-12: Sequence conservation of SpoIIAA from <i>B. subtilis</i> and <i>B. sphaericus</i> . ....	161
Figure 4-13: The metal dependence of SpoIIE <sub>590-827</sub> phosphatase activity <i>in vitro</i> .....	162
Figure 4-14: Manganese concentration dependence of SpoIIE <sub>590-827</sub> activity <i>in vitro</i> ... 163	
Figure 4-15: Comparison of SpoIIE fragment phosphatase activity. ....	164
Figure 4-16: Native PAGE analyses of SpoIIE fragment phosphatase activity at varying [Enzyme] to [Substrate] ratios.....	166
Figure 4-17: The transition state structural analogue of PSP.....	167
Figure 4-18: Effect of the phosphate transition state analogue, AlF <sub>4</sub> <sup>-</sup> , on SpoIIE:SpoIIAA complex formation.....	168
Figure 4-19: SEC analysis of SpoIIE:SpoIIAA interaction in the presence of AlF <sub>4</sub> <sup>-</sup> . ....	169
Figure 4-20: The chemical structures of Serine and Phosphoserine residues with the proposed structure mimics, Aspartate and Glutamate.....	171
Figure 4-21: Native PAGE analysis of SpoIIAA mutant interaction with SpoIIE <sub>590-827</sub> .. 172	
Figure 4-22: SEC-MALLS analysis of interaction between SpoIIE <sub>590-827</sub> and SpoIIAA (S58D). ....	173
Figure 4-23: Fluorescence based thermal shift assays of SpoIIE <sub>590-827</sub> stabilisation by Mn <sup>2+</sup> and Mg <sup>2+</sup> . ....	175

Figure 4-24: Bipyramidal crystal grown from SpoIIE <sub>590-827</sub> and SpoIIAA~P in the presence of 5 mM MgCl <sub>2</sub> from a well solution of 50 mM sodium citrate, 15% PEG 6K at pH 5. ....	176
Figure 4-25: Silver stained SDS PAGE analysis of crystals grown from SpoIIE and SpoIIAA~P in the presence of Mg <sup>2+</sup> . ....	177
Figure 4-26: SEC-MALLS analysis of SpoIIE <sub>590-827</sub> and SpoIIAA~P in the presence of Mg <sup>2+</sup> . ....	177
Figure 5-1: Location of sites of mutation and the associated phenotypes of SpoIIE Mutants.....	183
Figure 5-2: Regulation of ComK, the master transcriptional regulator of competence in <i>B. subtilis</i> .....	186
Figure 5-3: The DNA uptake apparatus produced by competent <i>B. subtilis</i> cells.....	187
Figure 5-4: RecA enables recognition of complementary DNA sequences to initiate homologous recombination. ....	189
Figure 5-5: Modes of Exogenous DNA Integration by Homologous Recombination in <i>B. subtilis</i> . ....	191
Figure 5-6: Plasmid map and Multiple Cloning Site of the <i>B. subtilis</i> Integration Vector: pUK-19.....	193
Figure 5-7: Distribution of SpoIIE S361F and Q483A Suppressor Mutations.....	198
Figure 5-8: Preparation of cloning reagents for the pUK-19- <i>spoIIE</i> plasmid.....	199
Figure 5-9: Schematic Representation of the XmaI/BamHI strategy for cloning of the pUK-19- <i>spoIIE</i> integration plasmid. ....	200
Figure 5-10: pUK-19- <i>spoIIE</i> cloning reagent activity tests. ....	201
Figure 5-11: Restriction Endonuclease flanking region conflict between BamHI and XmaI in the pUK-19 multiple cloning site. ....	202
Figure 5-12: Digestion to ascertain <i>spoIIE</i> fragment presence and orientation in recombinant pUK-19 plasmids.....	203
Figure 5-13: A schematic diagram of the genomic reorganisation caused by recombination of the <i>spoIIE</i> mutant harbouring pUK-19 plasmid. ....	204
Figure 5-14: The effect of culture container size on sporulation efficiency of the wild type IB333 strain of <i>B. subtilis</i> .....	206
Figure 5-15: Sporulation efficiencies of <i>spoIIE</i> mutant harbouring strains of <i>B. subtilis</i> relative to wild type. ....	207
Figure 6-1: A hypothesis for regulation of SpoIIE phosphatase activity.....	213
Figure 6-2: Preliminary data gathered on SpoIIE <sub>457-590</sub> , a target for NMR structural characterisation. ....	214

## Acknowledgements

I would first like to thank Prof. Tony Wilkinson for his indefatigable help, guidance and enthusiasm throughout my period of study at the YSBL. Thanks also to Marek Brzozowski for his guidance during my Thesis Advisory Panel meetings.

The work in this thesis could not have been carried out without the tireless efforts of Sally Lewis and Simon Grist who keep the YSBL labs in working order. Thank you to Vladimir Levdikov, my comrade in the study of SpoIIIE, and to Elena Blagova and Mark Fogg for getting me started in an unfamiliar lab environment. I also owe a great deal of thanks to Niels Bradshaw who has always brought a new point of view to the study of SpoIIIE and has been of great help.

I would very much like to express my gratitude to all the people, fellow students, post-docs and academics, who have made working at the YSBL such a pleasure. I am very glad to have been a part of the research community here and have found the environment extremely rewarding.

I would like to thank my Mum, Dad and brother Michael for their love and faith in me. Finally, and certainly not least, I would like to thank my Sarah for being such a wonderful person to be with.

## Author's Declaration

Chapters 1, 2, 5 and 6 are all my own work. Cloning and Expression testing of the SpoIIE<sub>334-827</sub> fragment was carried out alongside Emily Wheeler, a BSc project student. Analytical Ultracentrifugation experiments and data analysis were carried out by Dr Andrew Leech of the Molecular Interactions Lab, Dept. of Biology, University of York. The SpoIIE structures discussed within this work were solved by Vladimir Levnikov. During this work, a strong collaborative exchange of ideas was developed between the author and Niels Bradshaw, Harvard, however, except where stated the work presented here is my own. The co-expression of SpoIIAA and SpoIIAB was carried out alongside Katie Jameson, an MChem project student. The techniques used in the mutagenesis of *B. subtilis* were learned during a 3 week visit to the laboratory of Imrich Barak, Institute of Molecular Biology, Slovak Academy of Science. The experimental work in this study was all carried out by the author.

# Chapter 1: Introduction

## 1.1 Spore formation in Prokaryotes.

The vegetative growth cycle in prokaryotes, involving reproduction by binary fission, is well understood and also well conserved in genera of both Gram-positive and Gram-negative bacteria. Binary fission is a process in which a single mother cell undergoes symmetric cell division, producing two identical daughter cells. This cell cycle involves repeated iterations of cell elongation, chromosomal replication and cell division. While for most species of bacteria, cell division always occurs using this pathway, certain Gram-positive bacteria are able to undergo cell differentiation which results in the formation of spores. In contrast to the daughter cells produced by binary fission during vegetative growth, spores are desiccated and metabolically inactive on release. This endows them with increased resistance to adverse environmental conditions and extended periods of dormancy.

The *Streptomyces* genus of Gram-positive bacteria, belonging to the phylum Actinobacteria, commonly found in soil [1] undergoes sporulation as a means of reproduction. The use of spores in reproduction is commonly seen in species of fungi; however this behaviour is not the only peculiarity in the *Streptomyces* lifecycle. The vegetative growth of *Streptomyces* spp. occurs in repeatedly branching filaments, called vegetative hyphae, forming mycelia rather than in proliferation by iterative cell division [2]. Spores are developed by extension of so-called aerial hyphae, coated in a hydrophobic sheath, out of the liquid media (e.g. at the surface of a soil particle). After extension, multiple, simultaneous cell-divisions occur along the length of the hyphae resulting in separation into up to 50 asexual reproductive spores, or conidiospores [1, 2].

Conidiospores are released into the environment and each can germinate and form a new colony if in contact with a suitable substrate. Metabolically inactive, the spores can survive for long periods before germination. They are resistant to harmful conditions but far less so than the endospores produced by other Gram-positive bacteria [2]. Actinomycetes are biotechnologically and medically

important as they produce antibiotics such as erythromycin, streptomycin and chloramphenicol [1].

Genera of endospore-forming bacteria are large in number including *Bacilli* and *Clostridia*. In nutrient-depleted media, the formation of endospores allows preservation of the genetic information within an extremely resistant particle. Germination takes place when nutrient-rich conditions are restored. Dormant spores have been successfully revived after extraction from Dominican amber over 24 million years old [3]. The resistance of endospores to physical challenges such as heat, antibiotics and extremes of pH far surpasses that of non-extremophile vegetative cells.

## 1.2 *Bacillus* in industry and medicine.

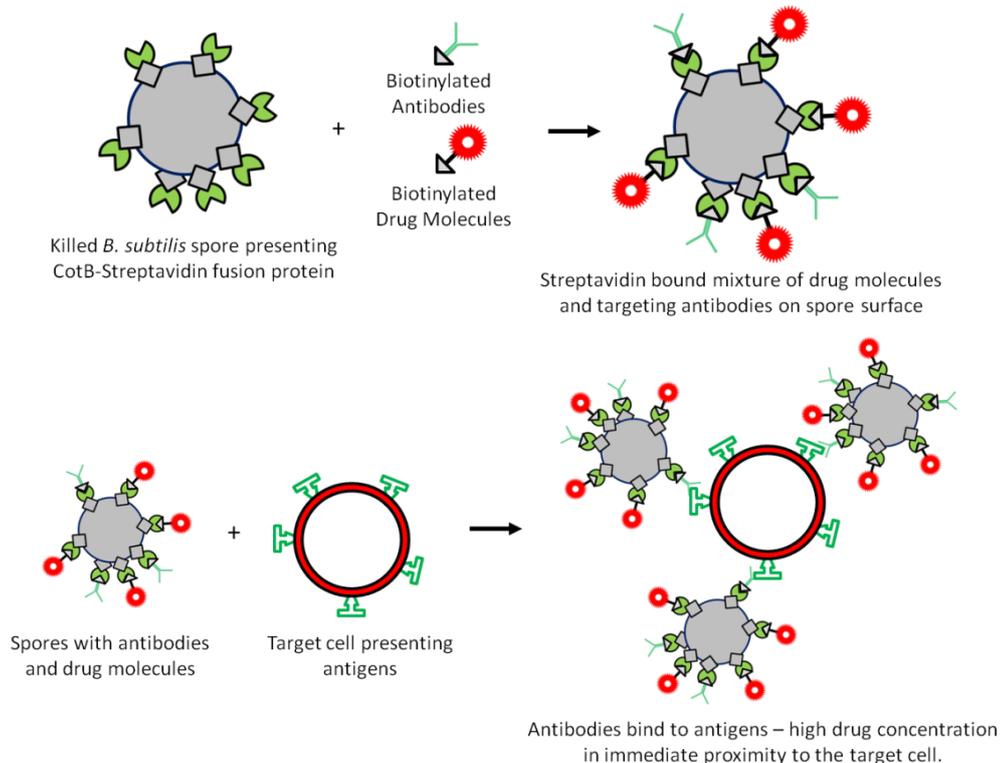
Strains of *Bacillus* are commonly used in biotechnological applications. *Bacilli* are naturally found in environments as diverse as soil, the guts of insects and hydrothermal vents. The variety of habitats is matched by diversity of metabolic capabilities [4]. This results in a diversity of applications of *Bacillus* spp.; *Bacilli* are used in the industrial production of key enzymes, vitamins, biosurfactants, pigments and biopolymers.

Spores formed by *Bacillus thuringensis* produce proteinaceous inclusions called  $\Delta$ -endotoxins. In certain *B. thuringensis* strains, these crystal (or Cry) proteins have specific insecticidal activities and are hence used as environmentally friendly insecticides. In addition the Cry genes have been produced in some plant species after genetic engineering to give inherent insecticidal properties [5, 6].

The concept of a probiotic has been in existence since the start of the 20<sup>th</sup> century, the modern definition; “A live microbial feed supplement which beneficially affects the host animal by improving its intestinal microbial balance” was suggested in 1989 [7, 8]. Species of the *Bacilli* and *Bifidobacteria* (a reclassification first applied to a *Lactobacillus* strain) genera represent a large portion of those microorganisms being used and studied as probiotics [9]. There is evidence that suggests that populations of probiotics in the gut

maintain the activity of the immune system allowing rapid response to new infections [10]. The therapeutic effects of *Lactobacillus* strains include prevention of antibiotic-associated diarrhoea, colon cancer and irritable bowel syndrome [11-13]. *Lactobacillus reuteri* CRL1098 strain and *Bifidobacterium adolescentis* and *pseudocatenulatum* species have been studied as probiotics for long term treatment and prevention of vitamin B12 and folic acid deficiencies, respectively [14, 15]. In each case, the bacterium in question is capable of synthesising and excreting these valuable molecules directly into the gut.

The ease of production and especially storage of spores of *Bacillus subtilis* and other endospore forming bacteria may have applications in the dissemination of vaccines to poorer areas. Dosing of mice with recombinant *B. subtilis* spores designed to present the tetanus toxin fragment C (TTFC) antigen on the outer coat enabled animals to survive a usually lethal challenge of tetanus toxin (50 x LD<sub>50</sub>) [16]. Similar methods have been used to induce immunity against *Clostridium difficile*, necrotic enteritis caused by *C. perfringens* and rotavirus [17-19].



**Figure 1-1: Spores as a vehicle for targeted drug delivery.**

Expression of CotB-streptavidin fusion proteins which are assembled on the surface of spores allows association of biotinylated moieties such as drugs and antibodies. Adapted from Nguyen et al., 2013 [20]

A similar technique using *B. subtilis* spores employs a killed form as a carrier of anti-cancer chemotherapy agents to cancer cells, specifically to target the epidermal growth factor receptor on HT 29 colon cancer cells [20]. Streptavidin is localised to the spore coat following its expression as a recombinant fusion protein with CotB. Any biotinylated moiety (i.e. drug molecules and antibodies) can then be attached to the spore surface allowing the system to potentially be used as a 'universal' targeted drug delivery system (Figure 1-1).

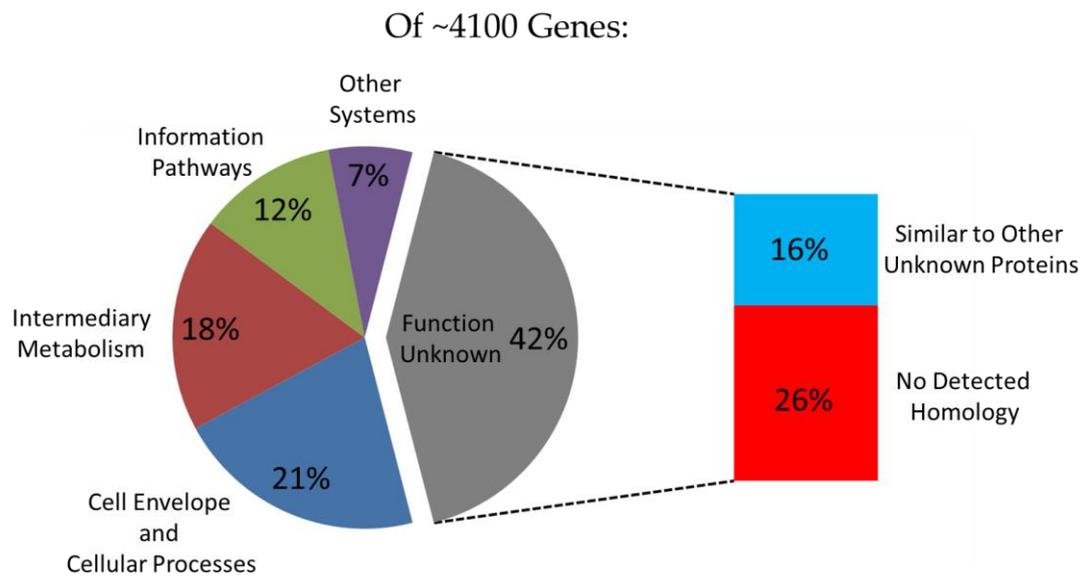
### 1.2.1 Risks posed by endospore-forming bacteria.

The increased resilience of endospores makes spore-forming species a potential hazard to human health. Common disinfectant techniques are often completely ineffective against them.

Some *Bacillus cereus* endospores are commonly present in foodstuffs and will survive temperatures of 100°C [21, 22]. Thorough cooking is required to disrupt the spores which will otherwise germinate in the gut. Enterotoxins produced during *B. cereus* proliferation may be responsible for up to 5% of foodborne illness [23, 24]. Endospore forming bacteria are also responsible for a number of serious and potentially fatal diseases. Toxins produced by the *Clostridium* species *C. botulinum* and *C. tetani* are harmful to humans and are the causes of botulism and tetanus respectively [25, 26]. *C. difficile* is a major concern in hospitals, affecting recovering patients. Resistant to most cleaning methods, *C. difficile* spores can quickly be ingested by multiple patients whose gut bacteria have been wiped out by the use of broad-spectrum antibiotics and cause severe diarrhoea potentially resulting in death [27]. *Bacillus anthracis*, the cause of anthrax, is the only obligate pathogen of the genus *Bacillus* [28]. Virulence is mediated by genes on two extra-genomic plasmids pX01 and pX02, allowing the bacteria to produce lethal toxins and evade the host immune system, respectively [29].

### 1.3 *Bacillus subtilis*

*Bacillus subtilis* is the model organism for Gram-positive bacteria. It is a rod-shaped, Gram-positive, bacterium of the genus *Bacillus*. Also known as grass *bacillus* and hay *bacillus*, *B. subtilis* has been considered as a soil-dwelling organism, however it has been argued persuasively that human gut populations are at levels suggesting that it can be a normal gut commensal [30].



**Figure 1-2: Distribution of Gene Functions in *B. subtilis*.**

The 'Other Systems' category consists of genes involved in stress response, antibiotic production, detoxification, phage-related functions and some other miscellaneous systems. From Kunst et al., 1997 [31].

The *B. subtilis* 168 genome consists of 4.214 megabase pairs of DNA comprising around 4,100 protein-coding genes with an average GC content of 43.5% [31]. Much of the genome (Figure 1-2) is devoted to the maintenance of the cell and the uptake and metabolism of carbon sources, especially those derived from plants. 42% of the genes are of unknown function and of these, more than half are not similar to any previously identified coding sequences. In 2003, 192 genes were shown to be essential, with another 79 predicted to be essential to the viability of *B. subtilis* on LB medium at 37°C [32]. As of February 2013, 261 genes encoding 259 proteins and two functional RNAs are regarded as essential after the elimination of some genes, due to their dispensability in other

conditions, and the addition of some genes previously unidentified [33]. These essential genes' functions are focused on protein synthesis, secretion and quality control, metabolism, the cell envelope and division, DNA maintenance and replication, cell protection and RNA synthesis and degradation. The functions of six essential genes remain unknown.

A factor influencing the use of *B. subtilis* as the model organism for Gram-positive bacteria is the ease with which genetic manipulation can be carried out. It is thought that natural genetic competence in *B. subtilis* confers advantages on the organism in nutrient limited conditions. The uptake and chromosomal inclusion of high molecular weight DNA fragments by recombination may provide an advantage to a struggling *B. subtilis* strain, perhaps exploiting previously unavailable metabolic pathways or conferring antibiotic resistance. The conditions which induce competence in *B. subtilis* are well characterised and in a laboratory environment are readily exploited to allow the development of mutant strains for further experimentation.

Industrially, *B. subtilis* is widely utilised and the pathway for secretion of so-called 'feeding' proteases has been used to effect the extracellular production of commercially valuable heterologous enzymes. While secreted proteins native to *B. subtilis* and some heterologous proteins can be collected from growth media in gram-per-litre quantities, the production of many proteins in this fashion is significantly less efficient [34]. The Sec-dependent secretion pathway in *B. subtilis* involves numerous checkpoints including signal peptide recognition, chaperones, translocation, propeptide- and enzyme-mediated folding and quality control proteases [35]. In order to achieve heterologous protein secretion in large quantities, the efficiency of each of these stages must be maintained. While the presence of biological regulatory elements currently complicates the application of *B. subtilis* secretion as a means of production of a wide range of proteins, strains with protease and chaperone-repression knockouts have been engineered. However, problems caused by incorrect translocation and extracellular protein folding are still present. In spite of the problems preventing the industrial production of many heterologous proteins via secretion in *B. subtilis*, the system is used widely and to great effect in the

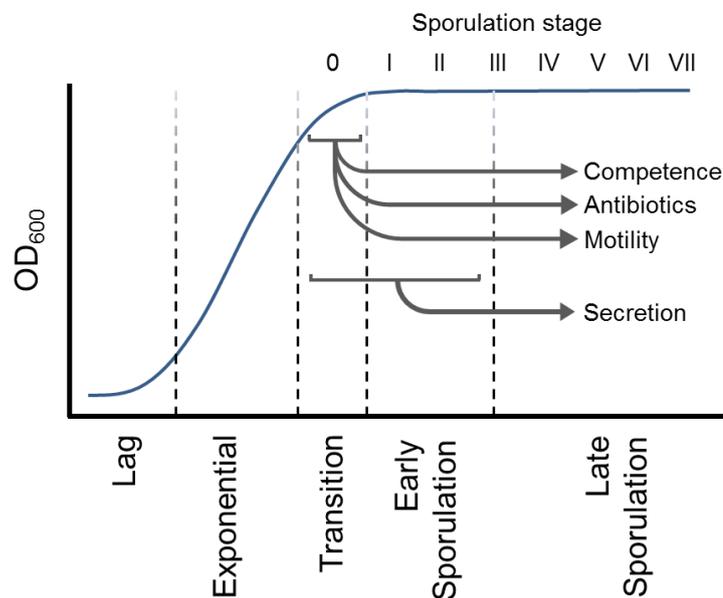
production of proteases and  $\alpha$ -amylases as well as numerous other commercially useful enzymes.

Upon encountering nutrient limitation, *B. subtilis* attempts to re-establish a food source through the development of motility (via flagella development and chemotaxis), genetic competence as well as production of extracellular hydrolases and antibiotics. The development of competence may allow *B. subtilis* to acquire and incorporate genetic material found in its environment, potentially conferring advantage in challenging conditions. Should these systems fail to redress the lack of nutrients, either through relocation to 'greener pastures' or the breakdown of material from other bacteria neutralised and digested by *Bacillus*' antibiotics and macromolecular scavenging systems, the bacterium embarks on a complex developmental process involving huge changes in morphology and genetic regulation in order to produce heat, chemical and irradiation resistant endospores. Over 500 genes are involved in the conversion of the vegetative cell into the dormant endospore via asymmetric cell division and the establishment of compartment specific gene expression [36]. Tight regulation of the processes involved in sporulation is key to its success.

## 1.4 Sporulation

### 1.4.1 Entry into the spore formation pathway

During growth, *B. subtilis* cells undergo binary fission which results in an exponential increase in cell density although this is dependent on sufficient nutrients being available. As nutrients are consumed, the cell cycle slows down and fewer cell division events occur (Figure 1-3). It is at this stage that *B. subtilis* utilises a number of adaptation strategies for re-establishing cell growth. At the same time, a process is initiated which will result in spore formation if the lack of nutrients persists.



**Figure 1-3: Example of a *B. subtilis* growth profile.**

The various phases of growth are shown on the x-axis. The adaptation strategies and their temporal utilisation are shown using grey bracketed arrows. The stages of sporulation are also shown indicating when sporulation becomes the sole survival strategy. Adapted from Phillips and Strauch, 2002 [37].

The formation of a resistant spore is energy and material intensive, involving the expression of scores of specific genes; furthermore it takes several hours to complete and, once Stage III is reached, is irreversible. To ensure that sporulation occurs only when all other survival strategies have been exhausted, sporulation-related *B. subtilis* gene expression is under tight transcriptional regulation, as described in section 1.4.2. Throughout nature, transcriptional

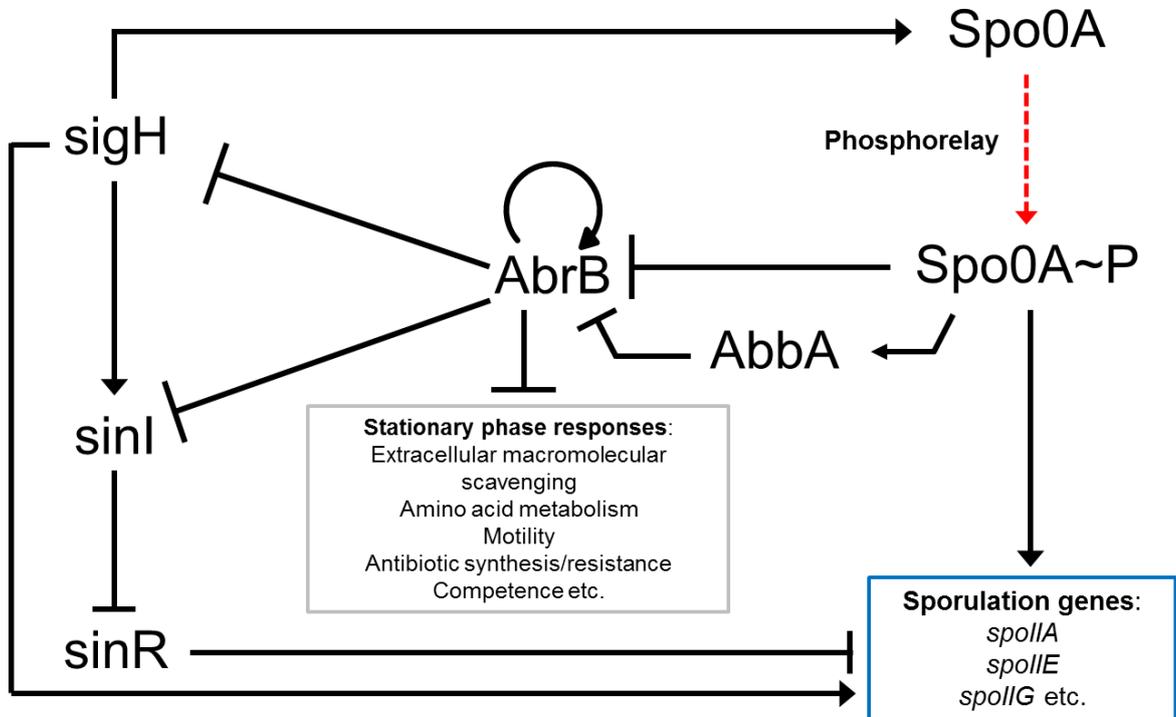
control can be enforced by global transcriptional regulators, a group of DNA binding proteins present throughout prokaryotic and eukaryotic biology. When these proteins are produced by the cell, they allow repression or activation of sub-sets of genes. It is by this mechanism that gene expression is directed throughout the stationary phase and towards the potential initiation of sporulation.

On transition into stationary phase, the transcriptional profile of *Bacillus subtilis* changes dramatically. During exponential phase of growth and division, genes encoding sugar utilisation and biosynthesis systems, are actively transcribed, whereas genes associated with motility, antibiotic production and nutrient scavenging are heavily repressed [37]. The opposite is true in stationary phase as the cell attempts to maintain resource levels; preserving rather than propagating its genetic material. The transition between these two states is mediated by global transcription regulators, specifically in the deactivation of certain repressors and the subsequent activation of genes related to stationary phase survival and sporulation.

#### **1.4.1.1 AbrB**

AbrB, named for its originally identified role in repressing antibiotic resistance, is a 10.4 kDa protein responsible for regulating transcription of over 40 different genes [37]. A number of these genes are themselves transcriptional regulators, while the remainder are involved in the majority of the adaptive responses that occur in stationary phase *B. subtilis* populations. This includes sporulation, with AbrB exerting repression of a number of sporulation related genes (Figure 1-4). *abrB* expression is autoregulated so as to maintain a stable intracellular concentration of the transcription factor throughout exponential phase. At the onset of stationary phase, the signal which interrupts expression of AbrB, and hence repression of stationary phase genes, is delivered by another global transcription regulator, Spo0A, in its phosphorylated form. Spo0A~P represses *abrB* transcription, causing intracellular concentrations of AbrB to be lowered and relieving AbrB's own repression of stationary phase system development. In addition Spo0A~P activates the transcription of *abbA*. The

protein, AbbA has been shown to bind directly to AbrB and prevent DNA-binding. This constitutes a second, parallel, mode of AbrB inhibition by Spo0A~P [38].



**Figure 1-4: Schematic representation of key parts of the AbrB regulon.**

A build-up of Spo0A~P causes a self-reinforcing cascade to deactivate repression by AbrB. Activation is depicted as black arrows; Repression is depicted as barred black lines. The dashed red arrow indicates the Spo0A phosphorelay which, via kinases and phosphoryl transferases, focuses nutrient deficiency signals towards the phosphorylation of Spo0A. AbrB is auto-regulated during the exponential phase. Inhibition of AbrB activity is carried out by Spo0A~P by two parallel pathways, by direct genetic repression and by production of AbbA, an AbrB inhibitor. Adapted from Philips and Strauch, 2002 [37].

It has been reported that AbrB is capable of differentially regulating multiple promoters. This results in different repressed systems being activated at distinct intracellular concentrations of the transcription regulator. Structural analyses of the protein have revealed a previously unidentified DNA binding motif which has been termed the looped-hinge helix fold [39]. It is thought that this fold affords the  $\alpha$ -helical and adjacent loop substructures involved in DNA binding a degree of flexibility in optimising DNA-protein interactions. The active DNA binding unit is a tetrameric species. In addition, a close homologue of AbrB, Abh, allows formation of AbrB:Abh heterodimers. As such, it is possible that differential transcriptional regulation of promoters affected by AbrB is

achieved by a complex system involving homo- and hetero- dimers and tetramers of AbrB and Abh, able to interact with a large number of promoters with a large range of affinities [40]. Over 40 chromosomal AbrB binding sites have been studied with no robust consensus sequence having been determined [41]. It is therefore thought that AbrB-DNA binding might also be dependent on specific three-dimensional DNA conformation.

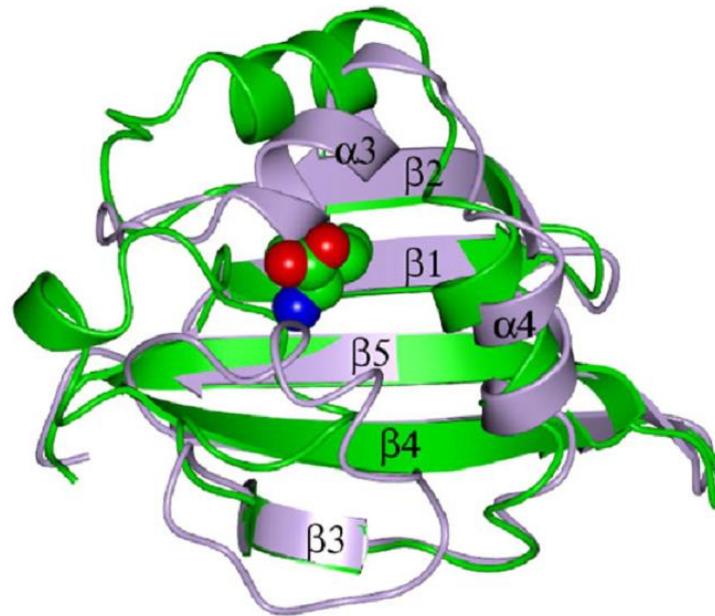
#### **1.4.1.2 SigH**

The late stationary phase sigma factor,  $\sigma^H$ , is key to the transition from vegetative growth to sporulation. The transcription of *spo0H*, which encodes  $\sigma^H$ , rises sharply at the onset of sporulation due to a positive feedback loop caused by generation of the activated master sporulation regulator, Spo0A~P. The regulon of  $\sigma^H$  consists of 87 or more genes under the control of up to 49 promoters [42]. This includes operons controlling genetic regulation (*spoIIA*) and the promotion of asymmetric cell division (*ftsAZ*) as well as genes encoding key components of the sporulation phosphorelay (*kinA* and *spo0F*), and *racA*, encoding the chromosomal reorganisation protein, RacA.  $\sigma^H$  mediated regulation is subject to catabolite repression and this regulation is abolished in the presence of a combination of 1% glucose and 0.1% glutamine.

#### **1.4.1.3 CodY**

Transition into the stationary growth phase is heralded by a drop in nutrient levels such that the bacteria can no longer maintain the rate of replication sustained throughout the exponential phase. Changes in intracellular energy levels and amino acid pool concentrations are stimuli which the cell responds to using a system of repressor:corepressor interactions. In prokaryotes, a corepressor is often a small molecule which binds to a repressor protein capable of inhibiting DNA transcription. Binding of the corepressor to the repressor increases the affinity of the latter for DNA leading to operator site binding and inhibition of transcription. Through this mechanism the depletion of a specific molecule, causing dissociation of the complex, can be redressed by

activation of genes capable of its synthesis or some orthologous method of cellular uptake.



**Figure 1-5: Overlay of the CodY GAF domain in unliganded and ligand bound states.**

The Isoleucine bound structure (green) shows a significant change in the conformation of helices  $\alpha 3$  and  $\alpha 4$  and the  $\beta 3$ - $\beta 4$  segment when compared the unliganded form (light blue). Taken from Levdikov et al., 2009 [43].

The CodY regulon comprises over 100 genes mainly related to nutrient deprivation stress responses and stationary phase. In addition, CodY represses production of an enzyme, KinB, and its counterpart, KapB, which promote sporulation. CodY is a major element of sporulation inhibition; CodY mutants sporulate at high efficiency even in nutrient rich media [44]. It is possible that other elements responsible for inhibition of sporulation are inactive in the  $\Delta codY$  strain (lacking any starvation stimuli), allowing unchecked Spo0A activity to bypass the usually stringent control of sporulation onset.

CodY is a 259 residue protein with an N-terminal GAF domain (a fold found in cGMP stimulated phosphodiesterases, adenylate cyclases and FhalA) and a C-terminal winged helix-turn-helix DNA-binding domain. It is responsible for the repression, during growth, of stationary phase and sporulation genes. CodY binds to GTP and the branched chain amino acids (BCAAs) isoleucine, leucine and valine, whose concentrations signal the energetic and metabolic status of

the cell. The BCAAs and GTP act as corepressors increasing the affinity of CodY for its binding sites situated at its target promoters [45]. As the intracellular concentrations of the corepressors fall (when the cells enter the transition phase) GTP and BCAAs dissociate from CodY leading to derepression of the CodY regulon. The corepressors of CodY are involved in protein synthesis: the branched chain amino acids being building blocks while GTP hydrolysis is responsible for providing the energy required for the initiation, elongation and translocation steps during RNA to protein translation [46]. They are hence a vital resource in the continuation of the cell cycle.

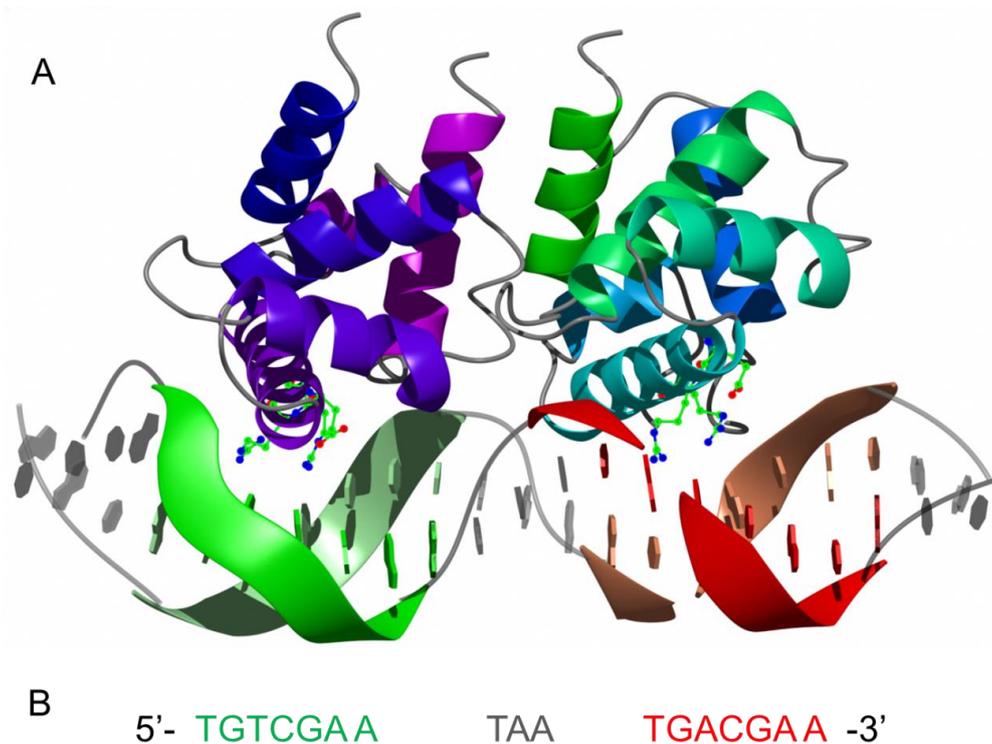
The C-terminal DNA-binding domain has a winged helix-turn-helix (HTH) motif, a recurring element in DNA binding proteins. The GAF-like N-terminal corepressor binding domain is assumed to bind BCAAs and GTP molecules in the same channel. Structures of the cofactor-binding domain of CodY in the unliganded form and in complex with isoleucine and valine have been determined [47]. The  $\alpha$ -amino and  $\alpha$ -carboxylate groups of the amino acid are stabilised by polar interactions with the Arg<sup>61</sup> side chain guanidinium group and the amide -NH and carbonyl groups of Val<sup>100</sup> and Thr<sup>96</sup>, respectively; A hydrophobic pocket is formed by the side chains of Met<sup>65</sup>, Phe<sup>71</sup>, Tyr<sup>75</sup> and Pro<sup>99</sup>. Comparison of these structures [43] reveals how ligand binding brings about drastic changes in the GAF domain (Figure 1-5). In the intact CodY molecule, these structural changes are expected to be propagated through the molecule to change the juxtaposition of the wHTH domains in the dimer. The details of the derepression mechanism remain to be elucidated.

### 1.4.2 Spo0A

If stationary phase survival systems fail to reverse the falling intracellular nutrient concentration levels, the continued confluence of nutrient deprivation signals will cause sporulation to be triggered. These signals are integrated by a complex phosphorelay culminating in the phosphorylation of the protein Spo0A. The phosphorylated form of Spo0A, Spo0A~P, a global transcription regulator, is capable of DNA binding and, as its concentration increases, effects

the activation of stationary phase nutrient deprivation responses. Ultimately, attainment of a threshold level of Spo0A~P, will cause the irrevocable initiation of sporulation [48].

Spo0A is a 29kDa protein which has been termed the master regulator of sporulation. A key route of early inquiry into the regulation of sporulation was into the *spo0* group of *Bacillus subtilis* mutants identified in the following loci: *spo0A*, *spo0B*, *spo0E*, *spo0F*, and *spo0H* [49]. These mutants were unable to initiate sporulation. Spo0A was identified as a key regulator; being either an activator or a repressor, depending on the regulatory target, of a large number of genes [50].



**Figure 1-6: DNA binding mode of Spo0A~P.**

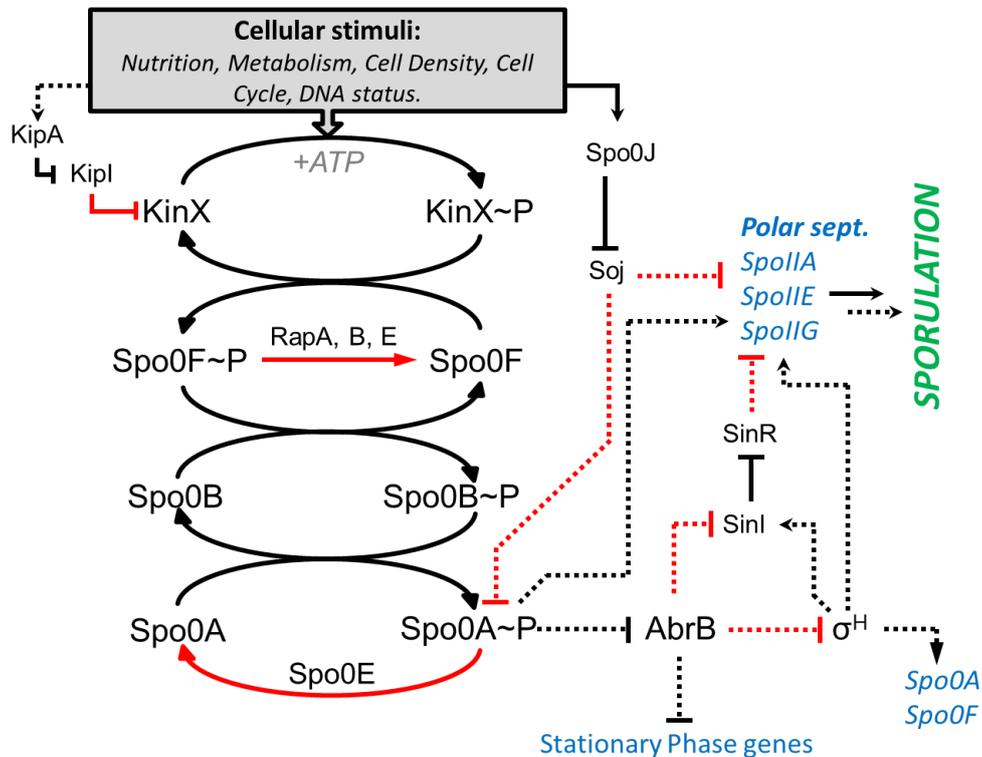
Dimeric C-terminal DNA-binding domains of Spo0A in complex with two aligned synthetic DNA duplexes (PDB: 1LQ1) [51]. Figure created using CCP4MG [52]. The upstream and downstream binding monomers of Spo0A are shown as ribbons with colours blended through the model from N- to C- termini; blue to magenta and green to blue, respectively. Atoms in the side chains of key nucleotide recognition residues Glutamate<sup>213</sup>, arginine<sup>214</sup> and arginine<sup>217</sup> are picked out as circles coloured by atom. Two duplexes form a long helix by associating in a head-to-tail manner; the orientation shows 5' to 3' on the sense strand as left to right. Spo0A boxes are shown as ribbons highlighted in green and light green at the upstream site and red and coral at the downstream site. The remaining base pairs are shown as partially transparent worms in grey. DNA base pairs are represented as nucleic acid blocks (same colours as phosphate backbone). (B) The two Spo0A boxes at the AbrB promoter site. The upstream site is shown in green and the downstream site in red. The interstitial 3 bases are shown in grey. This spacing is in evidence in the structure above.

Deletion of *spo0A* causes changes in the transcript levels of over 500 genes [36]. The 121 directly affected genes are arrayed across 54 genetic loci. Transcription of 80 genes expressed during vegetative growth is repressed by Spo0A. Conversely, the remaining 41 genes, involved in sporulation, are activated. 25 of the regulated genes, both positive and negative, are involved in further regulation of transcription. The direct and indirect transcriptional regulation effected by Spo0A is absolutely essential to both the initiation of sporulation and its successful progress and conclusion. Spo0A binds to a DNA binding site with a consensus sequence 5'-TGTCGAA-3' referred to as a '0A-box' [53]. These are present in two or more copies at most Spo0A-regulated promoters. The DNA-binding species of Spo0A is a dimer of Spo0A~P, as such Spo0A boxes appear in pairs with around 3 base pair separations [51] (Figure 1-6).

Expression of Spo0A is under transcriptional control of two promoters; one recognised by RNA polymerase containing  $\sigma^A$ , the other is recognised by RNA polymerase containing  $\sigma^H$ . Further regulation is applied by the formation of Spo0A~P, creating an autocatalytic feedback loop. A low level of *spo0A* is transcribed during vegetative phase and the transition phase, under regulation of  $\sigma^A$ . Phosphorylation of Spo0A then causes upregulation of production of *sigH*, previously repressed by AbrB.  $\sigma^H$  is capable of a higher level of *spo0A* transcription than  $\sigma^A$  allowing a rapid Spo0A build up once a small amount of Spo0A~P is created [48]. In addition to indirectly regulating its own transcription, via *abrB-sigH*, Spo0A~P also effects transcription of *kinA* and *spo0F*, whose protein products are integral parts of the Spo0A phosphorelay.

### 1.4.3 The sporulation phosphorelay

Spo0A is phosphorylated by an expanded two component system called a multi-component phosphorelay. This complexity enables the necessary tight regulation of Spo0A's activity (Figure 1-7).



**Figure 1-7: The sporulation phosphorelay.**

Lines with arrow heads indicate activation interactions, barred lines indicate inhibitory interactions. Solid lines indicate biochemical effects, dotted lines indicate effects on gene transcription. Red lines indicate interactions which inhibit the initiation of sporulation, black lines indicate interactions promoting the initiation of sporulation. Black text defines proteins. Blue text defines regulated genetic loci or wider groups of genes. KinX has been used to denote any of the five autokinases capable of phosphorylating Spo0F. KipI is an inhibitor only against KinA. As indicated, ATP is required for autophosphorylation of each kinase. Adapted from Philips and Strauch, 2002 [37].

The positive impetus for phosphorylation originates from five sensor kinases; KinA, KinB, KinC, KinD and KinE. The activities of these kinases are dependent on a variety of signals which serve to monitor the DNA replication status and metabolic and nutritional states, although the stimuli relevant to each kinase and their mode of transduction are unknown. Upon receipt of the appropriate stimulus, each kinase uses ATP to autophosphorylate a conserved histidine. The phosphoryl group then can then be transferred first to a conserved aspartate

residue of the response regulator, Spo0F, subsequently to a histidine residue on Spo0B and finally onto an aspartate on Spo0A.

Recent work indicates that KinA and KinB, which are under the transcriptional control of *sigH* and *sigA*, respectively, are capable of more extensive Spo0F phosphorylation compared to KinC, KinD and KinE and are responsible for driving the initiation of sporulation under laboratory conditions. Production of both kinases is affected by the stringent response, signalling nutrient deprivation, in which production of ppGpp is triggered. This molecule is capable of stalling the ribosome, preventing protein synthesis. There is also repression of *kinB* transcription by CodY [54]. Conversely, *kinC* and *kinD*, under the transcriptional control of the vegetative and early stationary phase transcription controller,  $\sigma^A$ , have slower kinetics and have been identified as maintaining low levels of Spo0A~P in the late exponential phase as well as being important for controlling biofilm formation [55, 56]. KinC and KinD have been reported to directly phosphorylate Spo0A with low activity *in vitro* [57]. The role of KinE is not known but its rate of phosphorylation is closer to those of KinC and KinD rather than those of KinA and KinB. KinA, among the five kinases, appears to be the only one with a co-expressed repression system in the form of the two proteins KipA and KipI, an anti-inhibitor and inhibitor, respectively. KinA and KinB are inhibited by the small protein, Sda, which is produced during DNA replication [58].

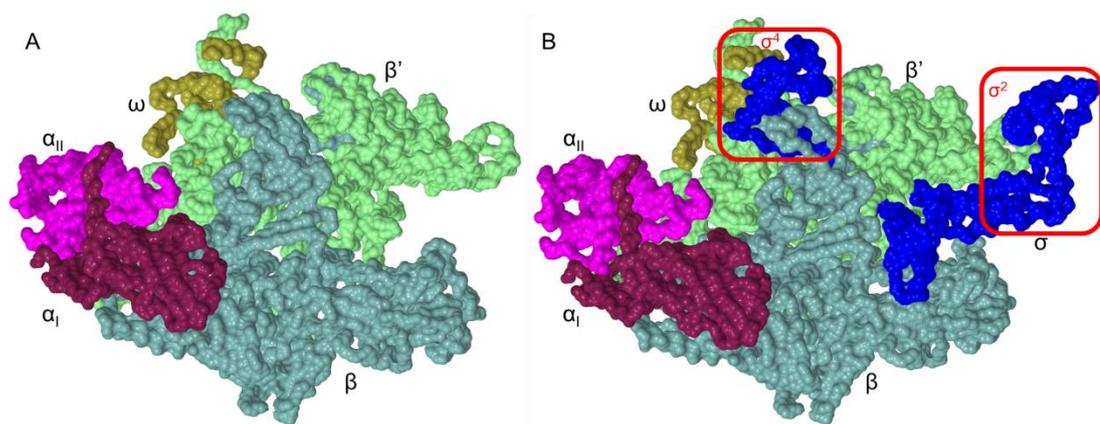
The response regulator, Spo0F, is phosphorylated by the above kinases at a conserved aspartate residue (Asp 54) [48]. Negative regulation of phosphorylation is exerted by the phosphatases; RapA, RapB and RapE. These enzymes are produced with a cognate Phr protein. The Phr proteins are pro-inhibitors of the Rap phosphatases. Pentapeptides derived from the Phr proteins form the active inhibition unit, however, development of these species requires sequential cleavage of the Phr protein on secretion from and, subsequently, reuptake into the cell. It is likely that this mechanism either allows stationary phase *B. subtilis* to measure local cell density (based on Phr derivative concentration) or that the process of secretion and uptake provides a timing mechanism for delaying inhibition of the Rap phosphatases [37].

Spo0F~P is acted on by the phosphotransferase, Spo0B which is under the transcriptional control of  $\sigma^A$ . No direct negative regulation of the phosphorelay has been identified acting on this factor, although it is considered likely that appropriate factors exist. Such a factor might be Obg, which is co-expressed with Spo0B, however, this has been found to bind to the ribosome in a guanine nucleotide-dependent manner [59]. Spo0B's activity is complete on transfer of the phosphate acquired from Spo0F to the conserved aspartate residue (Asp 56) of Spo0A. Phosphorylation induces dimerization of Spo0A.

Once produced, Spo0A~P is able to influence transcription, although further negative regulation is effected by the phosphatase Spo0E, and its two paralogues, YisI and YnzD, and the Spo0J/soj system. Spo0E, which is repressed by AbrB in exponential and early transition phase, is a 9 kDa phosphatase of which Spo0A~P is a substrate. Spo0E activity is able to catalyse the hydrolysis of the phosphate group from Spo0A~P to prevent Spo0A~P dependent regulation of transcription. YisI and YnzD are highly homologous to Spo0E and exhibit the same function in Spo0A~P dephosphorylation [60]. This is the last element of biochemical regulation of the phosphorelay. Spo0J and Soj mediate communication between transcriptional regulation and DNA replication and chromosomal segregation near the onset of sporulation. Soj binds to DNA as an ATP-bound dimer at a number of promoters regulated by Spo0A~P, causing inhibition and potentially capable of displacement of Spo0A directed transcriptional regulation [37, 61]. Soj also binds to a large number of sites around the origin of replication or to DnaA to promote and inhibit DNA replication, respectively [62]. Spo0J disrupts the DNA binding activity of Soj, allowing Spo0A-directed transcription to proceed.

## 1.5 Sigma factors

RNA polymerase (RNAP) is the enzyme responsible for the transcription of DNA into RNA, allowing the delivery of genetic information to the ribosome where proteins are produced. The template-directed phosphodiester bond formation reactions of RNAP are carried out by the 'core' enzyme consisting of five subunits:  $\alpha$ I,  $\alpha$ II,  $\beta$ ,  $\beta'$  and  $\omega$ . For localisation to DNA, upstream of the appropriate open reading frame, and initiation of transcription, a sixth, dissociable factor, called a  $\sigma$  factor, is required. The sigma ( $\sigma$ ) factor consists of determinants required for promoter recognition. When formed, the RNAP+ $\sigma$  complex, or holoenzyme (Figure 1-8), is capable of initiating transcription selectively on  $\sigma$ -specific promoters, by lowering the affinity of RNAP for non-specific DNA.  $\sigma$  factors belong to either the  $\sigma^{70}$  or  $\sigma^{54}$  families. These are mechanistically different and possess no sequence homology. Within a single organism, there will usually be no more than a single  $\sigma^{54}$  factor; conversely there can be dozens of  $\sigma^{70}$  factors. Through the conditional expression of alternate  $\sigma$  factors, the expression of genes in the cell can be efficiently controlled at the transcriptional level.



**Figure 1-8: Structures of the *Thermus aquaticus* (Taq) RNAP holoenzyme with (B) and without (A) a visible  $\Delta 1.1$  truncation of Taq housekeeping  $\sigma$  factor,  $\sigma^A$ .**

The  $\beta$  and  $\beta'$  subunits form a claw like cleft through which transcribing DNA is passed. The  $\sigma$  factor interacts extensively with the  $\beta$  and  $\beta'$  subunits of RNAP. Recognition of the DNA sequence at the -10 and -35 promoter elements is mediated by the  $\sigma$  domains,  $\sigma^2$  and  $\sigma^4$ , respectively, indicated in red. PDB ID: 1L9U.

$\sigma$  factors recognise two DNA sequences upstream of the start site for transcription. These sequences are referred to as the -10 and -35 consensus

sequences reflecting their *upstream* location. Association of the holoenzyme with the -10 and -35 sequences occurs through the  $\sigma^2$  and  $\sigma^4$  domains of the  $\sigma^{70}$  factor, respectively [63]. With a correctly positioned holoenzyme, the transcription start site of the target gene is placed in close proximity to the active site of RNAP in a so called closed promoter complex [64]. The DNA is then melted by RNAP to form the transcription bubble, allowing RNAP access to the nucleotide bases of the DNA. Numerous short abortive transcripts are produced before RNAP forms a stable functional interaction with both the DNA and nascent RNA chains; this forms the active elongation complex. Once synthesis of the RNA strand has commenced, the  $\sigma$  factor dissociates.

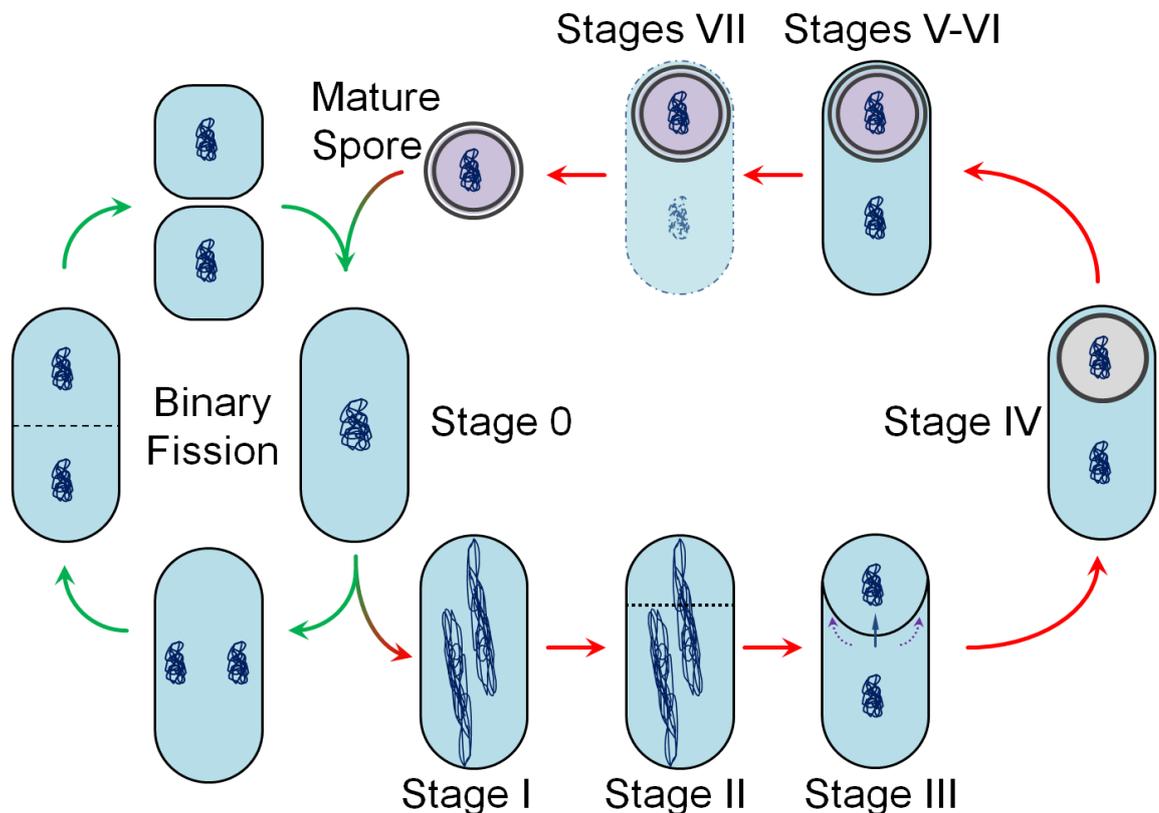
**Table 1-1: Sporulation related  $\sigma$  factors of *Bacillus subtilis*.**

$\sigma$ factor	Regulon function	Promoter sequence		
		-35 element	Spacer (bp)	-10 element
$\sigma^A$	Housekeeping/ early sporulation	TTGACA	17	TAT <u>A</u> AAT
$\sigma^E$	Early mother cell gene expression	ZHATAXX	14	CATA <u>C</u> AHT
$\sigma^F$	Early forespore gene expression	GCATR	15	GGHRAR <u>R</u> H TX
$\sigma^G$	Late forespore gene expression	GHATR	18	CAT <u>X</u> H TA
$\sigma^H$	Post exponential gene expression – competence and early sporulation	RWAGGAXXT	14	HGA <u>A</u> I
$\sigma^K$	Late mother cell gene expression	AC	17	CAT <u>A</u> NNNTA

Table adapted from Haldenwang, 1995. Nucleotide bases at the -10 site are underlined. H = A/C; N = A/G/C/T; R = A/G; W = A/G/C; X = A/T; Y = C/T; Z = T/G.

$\sigma$  factors play a key role in the sporulation pathway of *B. subtilis*. There are 10  $\sigma$  factors encoded in the genome of this organism [65]. Six of these  $\sigma$  factors play key roles in the initiation and completion of endospore formation (Table 1-1).  $\sigma^A$  is the *B. subtilis* housekeeping  $\sigma$  factor, responsible for directing transcription during growth and the early transition phase. The  $\sigma$  factors  $\sigma^E$ ,  $\sigma^F$ ,  $\sigma^G$ ,  $\sigma^H$  and  $\sigma^K$  direct transcription of stage and/or compartment specific sporulation genes and are activated in a sequential and temporal order referred to as a  $\sigma$  factor cascade during spore development.

## 1.6 Morphological stages of sporulation



**Figure 1-9: The morphological stages of sporulation.**

Chromosomes are shown in dark blue. The cell membrane and peptidoglycan are represented by a thin black line. Asymmetric septation is shown by a black dotted line in stage II. DNA translocation and cell membrane migration in stage III are represented by solid blue and dashed purple arrows, respectively. The spore cortex is shown in stage IV as a thick black line. The cortex and spore coat combination is shown as two thick black lines. Transparent elements with dashed lines undergo autolysis in stage VII.

Endospore formation, in *B. subtilis*, is an intricate process which progresses through seven morphologically defined stages, taking 7 to 8 hours to complete at 37°C (Figure 1-9). These stages were first categorised by electron microscopy studies of genetic mutants [66]. Stage 0 is represented by a cell not involved in sporulation or binary fission. In stage I DNA replication has been completed and the chromosomes have assumed an elongated form called the axial filament with their respective origins of replication at each of the cell poles. Stage II is defined by the development of the sporulation septum, at a polar site creating a small compartment; the forespore, and larger compartment; the mother cell.

The polar septum intersects with the extended forespore chromosome so that this chromosome is partitioned into both compartments with 70% of the chromosome initially occupying the mother cell compartment [67]. This distal portion of the chromosome is next translocated to the forespore through the action of the septum-located DNA translocase SpoIIIE. In stage III there is thinning of the peptidoglycan layer between the two daughter cells and migration of the mother cell membrane around the forespore in an engulfment process reminiscent of phagocytosis. Engulfment is completed when the migrating membranes meet on the distal side of the forespore where membrane fission takes place and the forespore is released into the mother cell cytoplasm as a cell-within-a-cell. Stages IV and V are characterised by the deposition of layers on the outer surface of the forespore; first a peptidoglycan cortex, followed by a proteinaceous spore coat. Spore maturation is stage VI, this stage exhibits no evident morphological alteration, however, spores which do not undergo maturation are less environmentally resistant than those which have been allowed to mature. In Stage VII, autolysis of the mother cell occurs; the now mature and resistant spore is released into the environment.

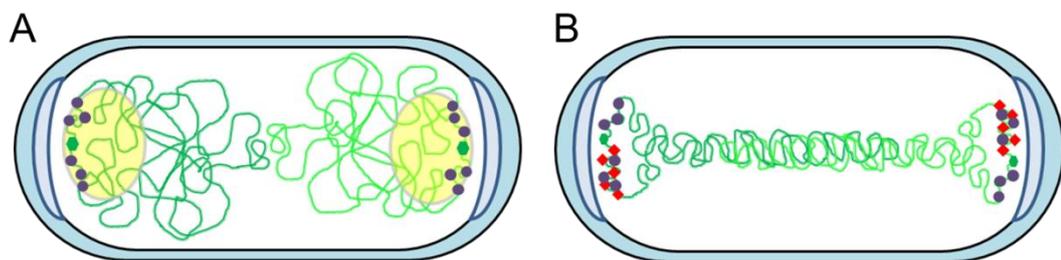
### 1.6.1 Coupling cell cycle to sporulation

Successful sporulation requires that following the asymmetric cell division, the mother cell and the forespore acquire a single complete chromosome. For this purpose, mechanisms are in place to ensure that cells undergoing chromosome replication do not initiate sporulation. DNA replication is initiated by DnaA, which assembles at the origin of replication, *oriC*, where it recruits the components of the replisome. DnaA is also a transcription factor that stimulates transcription of *sda* which encodes the 52 residue protein Sda which has been shown to interact with, and inhibit, the sensor kinases, KinA and KinB [68, 69]. This inhibition effectively blocks the development of threshold concentrations of phosphorylated Spo0A thus preventing initiation of sporulation while DNA replication is taking place. Sda inhibition is thought to be relieved by two mechanisms. Transcriptional activation of *sda* by DnaA is ATP dependent. DnaA-ATP is capable of initiating replication and activating *sda* transcription. As

replication continues the concentrations of DnaA-ATP are lowered by ATP hydrolysis to yield DnaA-ADP which is not capable of activating *sda*. Relief of KinA/KinB inhibition then occurs because Sda is an unstable protein which is susceptible to intracellular proteolysis by the ClpXP system [70, 71]. Once replication is complete, the diploid cell is free to pursue spore development.

### 1.6.2 Sporulation stage I

Stage I is the first stage in which a deviation from vegetative growth and symmetric cell division can be detected. Medial cell division occurring during growth is accompanied by markedly different chromosomal organisation in the pre-divisional cell to that in the pre-divisional sporangium (Figure 1-10).



**Figure 1-10: Chromosomal organisation in pre-divisional vegetative and sporulating cells.**

(A) Vegetative chromosomal segregation with barbell pattern of DNA. Chromosomes are represented by tangled lines in two shades of green. *oriC* regions are represented by green hexagons. Spo0J is shown as purple circles. DivIVA is shown as a light blue crescent outlined in dark blue. SMC is localised to Spo0J and represented by a yellow translucent circle outlined in grey. (B) Axial filamentation in a pre-divisional sporangium. RacA (red diamonds) associates DNA close to *oriC* and binds DivIVA at the cell poles. Adapted from Rudner, 2009; Pavlendova, 2007; and Errington, 2003.

Vegetative chromosomal segregation is hypothesised to occur throughout, and as a product of, replication in which the *oriC* regions of the nascent daughter chromosomes are driven or pulled away from the centre of the cell and reorganised by the structural maintenance of chromosomes complex (SMC). The exact mechanism by which this is achieved is unknown; however the deletion of either the *spo0J/soj* locus or the genes encoding the actin-homologues, MreB and Mbl, results in segregation defects. It is thought that Spo0J protects *oriC* from asynchronous replication initiation by forming an organised nucleoprotein structure in its vicinity [72], while also recruiting the SMC to compact the nascent chromosome [73]. DivIVA, a protein capable of

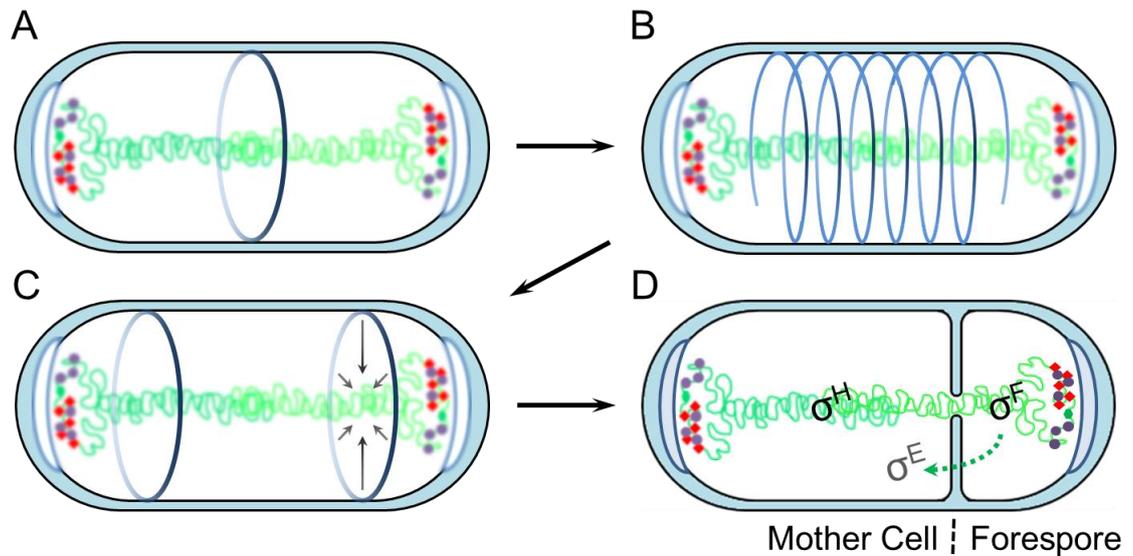
recognising cell membrane curvature [74], which localises in clusters at the cell poles, is implicated in the organisation of Spo0J and regulating its interactions with Soj, which directs the Spo0J/*oriC* complexes towards the cell poles [75]. The product of this chromosomal reorganisation is a barbell like structure with the greater proportion of each chromosome occupying its own half of the pre divisional cell.

Conversely, chromosomal reorganisation before asymmetric division in spore development results in a so-called axial filament. This reorganisation is dependent on two proteins, RacA and DivIVA. Production of RacA, named for its role in Reorganisation and chromosomal Anchoring, is directed by  $\sigma^H$ . The inhibitory effects on Spo0A phosphorylation by both the Spo0J/Soj and DnaA/Sda systems ensure that RacA is inactive until it can perform its role on the completely replicated chromosomes. The role of RacA is to collapse the amorphous nucleoid into a filament aligned with the long axis of the cell, a key feature of which is the anchoring of *oriC* to the cell pole. RacA binds to a number of sites on the chromosome close to *oriC* and also concomitantly to DivIVA [76]. This anchoring is also mediated by Soj and Spo0J, which are also required for *oriC* delocalisation from mid-cell.

### 1.6.3 Sporulation stage II

Asymmetric cell division (Figure 1-11) results in the formation of a smaller forespore and a larger mother cell which are divided by the sporulation septum, consisting of two membranes. This process requires a relocalisation of the cell division apparatus used in medial cell division. The cell division machinery, which is assembled following the formation of a structure called the Z-ring, is a complex conglomeration of proteins involved in maintaining localisation, gaining purchase on the cell membrane and driving contraction of the cell at the point of division. Up-regulation of FtsZ production and the presence of the sporulation-specific protein SpoIIE, both resulting from  $\sigma^H$  directed transcription, causes the cell division machinery to assemble around repositioned ring-like structures of FtsZ formed near to the two cell poles via a

helical intermediate state [77]. The cell division machinery consists of a large number of protein elements including FtsZ, FtsA, ZapA, EzrA, DivIB and peptidoglycan remodelling enzymes. Only one of the Z-rings will go on to form a site of cell division, dividing the mother cell and smaller forespore. The axial filament chromosomal arrangement means that the developing septum encloses ~30% of the forespore chromosome within that cell compartment.



**Figure 1-11: Sporulation stage II.**

(A) The medial Z-ring is formed at the vegetative division site. (B) A helical intermediate allows relocation of the Z-ring. (C) Two Z-rings form, one at each pole. Asymmetric cell division occurs as one of the two Z-rings contracts and forms a septum. (D) Only 30% of the chromosome is initially captured in the forespore compartment following septation.  $\sigma^F$  transcription, activated only in the forespore, signals the activation of  $\sigma^E$  in the mother cell. Adapted from Ben-Yehuda, 2002.

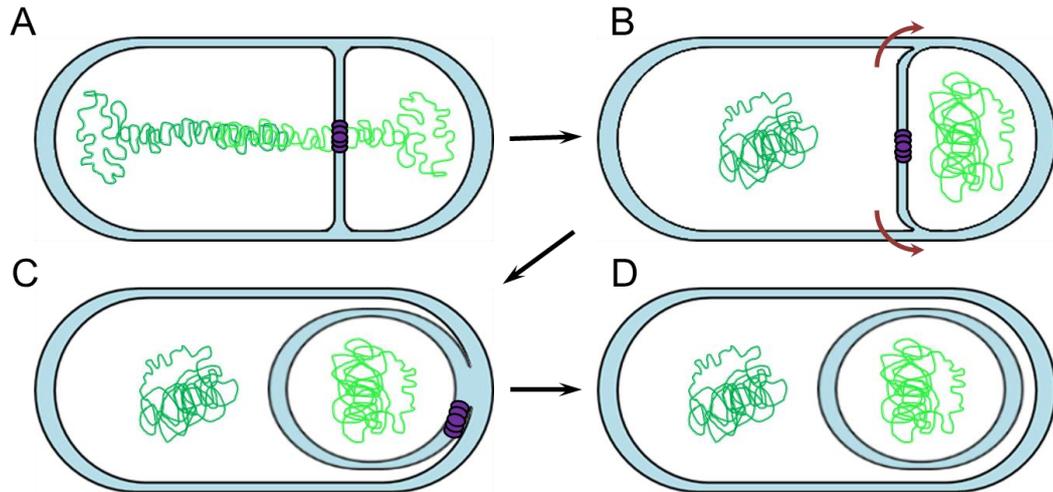
Stage II is also characterised by the activation of  $\sigma^F$ , the first compartment specific  $\sigma$  factor, in the forespore. This activation is mediated by a three component regulatory system consisting of SpoIIAA, SpoIIAB and SpoIIE. SpoIIAB is an anti-sigma factor which can sequester  $\sigma^F$  in an inactive complex [78]. SpoIIAA is an anti-anti-sigma factor which can break the inhibitory complex, but only in its dephosphorylated form [79-81]. SpoIIAA is phosphorylated and inactivated by SpoIIAB. This inactivation is overcome by the PP2C serine phosphatase, SpoIIE which dephosphorylates SpoIIAA~P [82-84]. Signalling proteins in the  $\sigma^F$  regulon communicate across the septum to activate  $\sigma^E$  in the mother cell. The mechanism by which  $\sigma^F$  activation is

restricted to the forespore is unknown but hypotheses include transient gene asymmetry between mother cell and forespore, the differential volume of the two compartments and compartment specific deactivation of SpoIIE [67, 85, 86].

#### 1.6.4 Sporulation stage III

Engulfment of the forespore by the mother cell occurs during the third stage of sporulation (Figure 1-12). Before this can occur, the trailing portion of the forespore chromosome must be translocated from the mother cell. This is accomplished through the activity of the ATP-dependent DNA translocase, SpoIIIE. This protein is produced before cell division occurs, under the control of  $\sigma^A$ , and the protein assembles as a complex at the annulus of the contracted sporulation septum. It has been shown recently that SpoIIIE multimers actively search for DNA sequences called SRSs (SpoIIIE Recognition Sequences) which dictate the polarity of translocation which is towards *oriC* in the forespore [87, 88]. In addition, SpoIIIE translocation strips RNAP and transcriptional regulators from the chromosome, effectively excluding them from the forespore [89]. This activity may have a role in the establishment of developmental gene expression as latent vegetative  $\sigma$  factors are excluded from the forespore.

The membrane of the mother cell next migrates towards the distal cell pole engulfing the forespore as it does so. Engulfment is complete when the mother cell membranes meet and undergo fission releasing the forespore as a cell within the mother cell. Membrane migration relies on proteins produced in both cell compartments, under the control of  $\sigma^F$  in the forespore, and  $\sigma^E$  in the mother cell. SpoIIB, SpoIID, SpoIIM and SpoIIP are produced in the mother cell and form a peptidoglycan remodelling complex. Null mutant studies show their role is in destabilising the sporulation septum to promote membrane migration to the cell pole [90].



**Figure 1-12: Sporulation stage III.**

(A) A SpoIIIE complex forms at the septum's annulus and translocates the lagging chromosome into the forespore. (B) Migration of the cell membrane towards the pole begins at the outer edges of the septum, the SpoIIIE locus remains and is carried towards the pole. (C) The separation of the membranes converges at the cell pole. (D) Membrane fission occurs completing engulfment of the forespore by the mother cell.

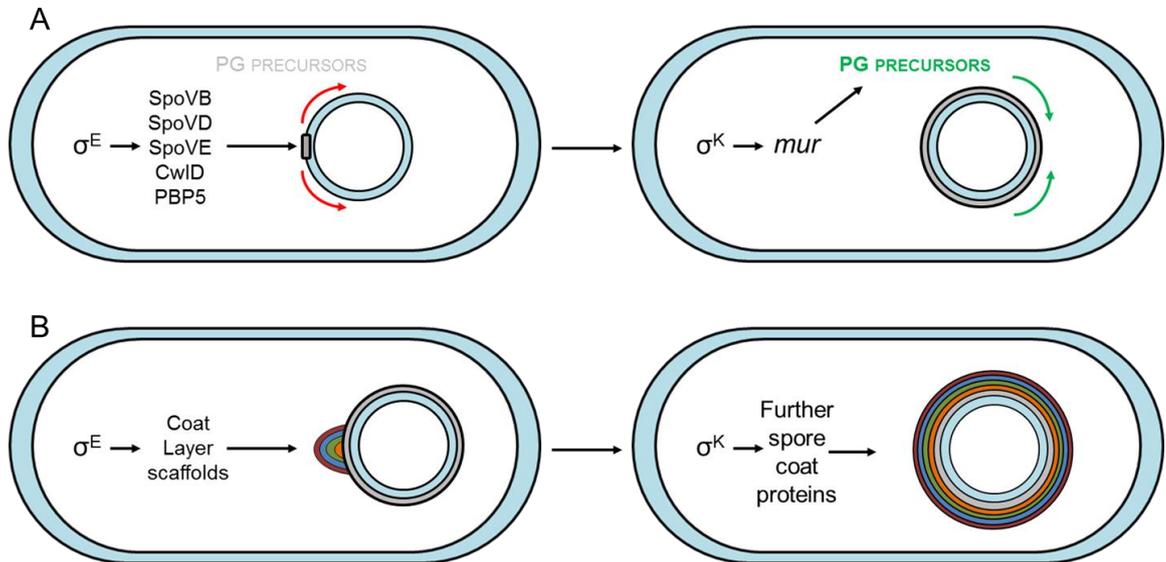
A molecular zipper complex formed by SpoIIQ and SpoIIIAH [91, 92], produced in the forespore and mother cell, respectively, promotes membrane migration. The synthesis of rigid peptidoglycan along the leading edge of the engulfing membrane is also thought to play an important role [93]. Membrane fission leads to separation of the two cells requiring SpoIIIE, the DNA translocase, SpoIIQ and FisB [94].

The completion of engulfment brings about activation of  $\sigma^G$  in the forespore, allowing expression of a new profile of genes. Similarly, signalling by  $\sigma^G$ -regulated genes causes activation of  $\sigma^K$  in the mother cell.

### 1.6.5 Sporulation stage IV

During stage IV (Figure 1-13A), the peptidoglycan (PG) spore cortex is produced [95]. The enzymes required for synthesis of this layer of the developing spore; SpoVB, SpoVD, SpoVE, CwlD and PBP5, are produced in the mother cell under the control of  $\sigma^E$  and localise to the engulfing membrane. Development of the spore cortex is, however, delayed until engulfment is complete. The enzymes regulated by  $\sigma^E$ , are unable to proceed due to limited

availability of the required substrates, peptidoglycan precursors; UDP-NAM and UDP-NAG. This shortage is redressed when the  $\sigma^K$  controlled genes in the *mur* locus are expressed. The produced enzymes synthesise the PG precursors required for construction of the spore cortex.



**Figure 1-13: Late sporulation morphological developments.**

(A) Stage IV, cortex formation.  $\sigma^E$  directs transcription of cortex forming proteins.  $\sigma^K$  causes transcription of the *mur* locus of peptidoglycan precursor synthesising enzymes, enabling cortex production. (B) Stage V, spore coat formation.  $\sigma^E$  directs transcription of localised spore coat scaffolds.  $\sigma^K$  causes transcription of spore coat proteins, enabling envelopment of the spore by the spore coat layers: basement layer; orange, inner coat; green, outer coat; blue, crust; red. Adapted from McKenney, 2013.

### 1.6.6 Sporulation stage V

Regulation of the construction of the multi-layered, proteinaceous spore coat, consisting of the basement layer, inner coat, outer coat and crust, occurs similarly to that of the cortex [96]. Clustered hubs of scaffold proteins for all four spore coat layers are developed at the mother cell proximal edge of the forespore under the transcriptional control of E- $\sigma^E$  (RNA polymerase containing  $\sigma^E$ ). Transcription control by E- $\sigma^K$  leads to the production of coat proteins which allow assembly to occur, thus enclosing the forespore (Figure 1-13B).

### 1.6.7 Sporulation stage VI

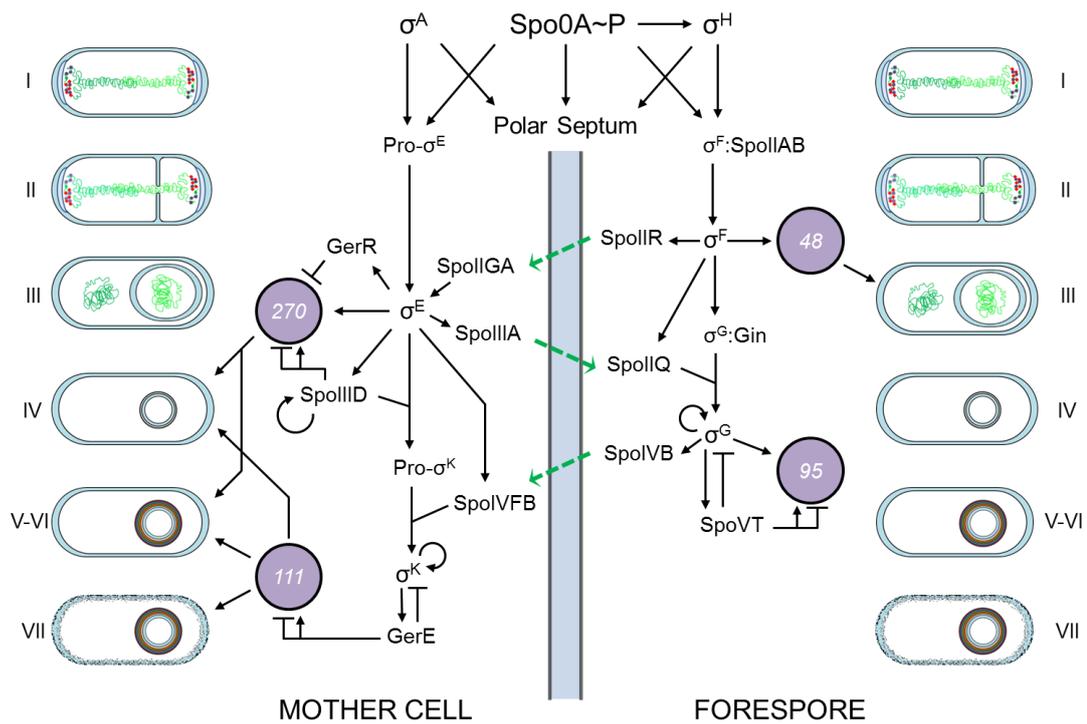
The processes contributing to spore maturation are unknown. However, investigations into the properties of mature and immature spores have suggested that during this stage, chemical modification of the spore cortex or coat takes place [97]. Some candidates for the enzymes involved include the spore coat proteins, Tgl and YabG. Spores which are not allowed to mature exhibit decreased wet heat and hypochlorite resistances.

### 1.6.8 Sporulation stage VII

Mother cell lysis is the last stage in endospore formation. The production of the peptidoglycan hydrolases, or autolysins, CwlB, CwlC and CwlH is upregulated at the point of mother cell lysis [98, 99]. These enzymes break down the mother cell wall as part of mother cell apoptosis. The mature spore is released and can remain dormant until germinant nutrients are encountered.

## 1.7 The Compartment Specific $\sigma$ Factor Cascade

Sporulation is directed temporally and morphologically by four compartment specific  $\sigma$  factors. After activation of  $\sigma^F$  in the forespore, each subsequent  $\sigma$  factor relies on transcription and activation by its predecessors. This results in a  $\sigma$  factor cascade (Figure 1-14). The activity of each  $\sigma$  factor is also under post-transcriptional control and each must be activated by intercompartmental signals.



**Figure 1-14: The cascade of sporulation specific  $\sigma$  factors.**

The activation and control of  $\sigma$  factors and key transcriptional regulators are shown. The polar septum is in the centre with the mother cell on the left and forespore on the right, as indicated. Green dashed arrows indicate an interaction across the intercellular membranes. Black arrows indicate activation, black barred lines indicate repression. Purple circles indicate the approximate gene numbers in a regulon. The stages of sporulation are shown at both edges; arrows from regulons to a stage indicate that genes in that regulon make a major contribution to the stage's development. Adapted from Kroos et al., 2007 [100]

$\sigma^F$  is the first sporulation specific sigma factor to be activated.  $\sigma^F$  is present in the pre-divisional cell and in both compartments following the asymmetric cell division but its activation is delayed pending the completion of the septum and restricted to the forespore. The protein product of *spoIIGB*,  $\text{pro-}\sigma^E$ , is similarly distributed before the sporulation cell division. The N-terminal 27 residues of

pro- $\sigma^E$  are removed to generate active  $\sigma^E$ , which serves as the first mother cell specific  $\sigma$  factor. Activation is mediated by SpoIIIGA, a protease, which cleaves the inhibitory propeptide of pro- $\sigma^E$ . For proteolysis to take place, SpoIIIGA requires a signal from the forespore in the form of SpoIIR [101, 102]. SpoIIR is produced in the forespore under direction of  $\sigma^F$  where it is initially repressed by RsfA. Its signal peptide directs it to the intermembrane space, where, by mechanisms which are not well understood, it is able to activate SpoIIIGA, and hence  $\sigma^E$ .

The regulons of  $\sigma^F$  and  $\sigma^E$  consist of around 48 and 270 genes, respectively, each controlling morphogenic proteins and further factors required for developmental regulation [100]. Under  $\sigma^F$  control are the genes *sigG* and *csfB* (*gin*), whose products,  $\sigma^G$  and Gin, form an inhibitory complex preventing formation of E- $\sigma^G$ . SpoIIQ, which is required for forespore engulfment, and SpoIIR are also under  $\sigma^F$  control and play key roles in intercellular signalling. The much larger regulon of  $\sigma^E$  encodes factors required for largely mother cell dependent processes, such as spore cortex and coat formation, occurring in late sporulation. The proteins produced under  $\sigma^E$  control include pro- $\sigma^K$  and two transcription regulators, GerR and SpoIID, as well as the *spoIIIA* operon, which codes for the eight proteins involved in engulfment and intercompartmental secretion and signalling.

Post-engulfment gene expression is orchestrated by  $\sigma^G$  in the forespore and  $\sigma^K$  in the mother cell. The trigger for  $\sigma^G$  activation in the forespore is the completion of spore engulfment. A multi-component secretion system is assembled between the two cells consisting of products of the E- $\sigma^E$  controlled *spoIIIA* locus in the mother cell and the product of the E- $\sigma^F$  regulated *spoIIQ* in the forespore [103]. It is thought that this apparatus plays a key role as a food channel in the nurturing of the engulfed forespore by the mother cell as well as contributing to signalling in  $\sigma^G$  activation [104]. Recent work has identified that transport might be activated by septal thinning caused by peptidoglycan hydrolysis which occurs during engulfment [105].

Aside from the assembly of the secretion system, the signal directing  $\sigma^G$  activation is not known. The complex formed between the  $\sigma^G$  specific inhibitor,

Gin [106, 107], and  $\sigma^G$  is disrupted allowing E- $\sigma^G$  to form. Auto regulation of *spoIIIG* causes a rapid rise in  $\sigma^G$  levels and hence in the expression of the  $\sigma^G$  regulon consisting of around 95 genes. Some of these genes encode proteins required to protect the DNA in the dormant spore from UV radiation, heat and desiccation, such as the small acid soluble proteins (SASPs) [108], or the proteases necessary to degrade them upon germination. The transcriptional regulator, SpoVT, is also under E- $\sigma^G$  control. It inhibits the positive feedback loop in  $\sigma^G$  expression, resulting in a pulse of  $\sigma^G$ -directed gene expression followed by a plateau as SpoVT represses *sigG* expression [109].

The regulatory mechanism preventing incorrect temporal activation of  $\sigma^K$  in the mother cell is more elaborate. Expression of *sigK*, the gene encoding  $\sigma^K$ , is directed by E- $\sigma^E$ . Remarkably, the *sigK* open reading frame is interrupted by a 42 kbp segment of DNA, termed the *skin* (SigK Intervening element which is now known to be a cryptic prophage). *spoIVCB* encodes the N-terminal portion of  $\sigma^K$ , and *SpoIVC* its C-terminal portion [110]. *spoIVCA* is located in the intervening DNA sequence and under E- $\sigma^E$  transcriptional direction, with further regulation exerted by SpoIIID. *spoIVCA* encodes a site-specific recombinase, SpoIVCA whose function is to excise the cryptic prophage DNA. The post excision chromosomal sequence, *sigK*, encodes the inactive  $\sigma$  factor precursor, pro- $\sigma^K$ .

The processing of the pro- $\sigma^K$  involves the removal of 20 residues from the N-terminus, catalysed by the  $\sigma^E$  directed gene product, SpoIVFB, an intramembrane metalloprotease. SpoIVFB is regulated by a three component inhibition system comprising SpoIVFB, BofA and SpoIVFA [111]. This complex assembles in the outer forespore membrane and is disrupted by SpoIVB, an intramembrane serine protease produced in the forespore under direction of  $\sigma^G$ . SpoIVB selectively degrades SpoIVFA and relieves inhibition of SpoIVFB [112]. Recent work has shown that pro- $\sigma^K$  is membrane associated and proteolysis is accompanied by release of active  $\sigma^K$  into the cytoplasm [113]. The 111 gene  $\sigma^K$  regulon encodes proteins required for cortex precursor synthesis (*mur* genes), spore coat formation (*cot* genes) and factors involved in apoptosis of the mother cell. The transcriptional regulator, GerE, regulates  $\sigma^K$  directed

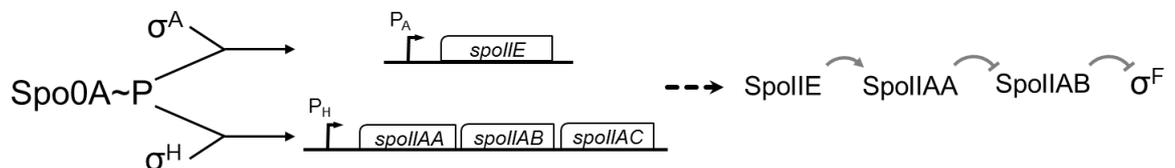
transcription in the mother cell playing a role analogous to that of SpoVT in  $\sigma^E$  regulation in the forespore.

It has been suggested that activation of  $\sigma^G$  or  $\sigma^K$  is insufficient to account for the high rates of  $\sigma^F$  and  $\sigma^E$  activity attenuation, respectively [114]. It is therefore posited that negative feedback loops are in place to deactivate  $\sigma^F$  and  $\sigma^E$  upon the advent of their successors' activity [115]. Fin (YabK) has been found to fulfil the role of  $\sigma^F$  deactivation. Expressed under the control of  $\sigma^G$ , Fin has been shown to inactivate  $\sigma^F$  selectively allowing  $\sigma^G$  to monopolise the available core RNAP [116]. The mechanism by which  $\sigma^E$  is inactivated upon  $\sigma^K$  activation is, as yet, not known.

## 1.8 Activation of the First Compartment Specific $\sigma$ Factor, $\sigma^F$

The gene expression programme followed by the two cells during spore formation relies on the correct temporal and spatial activation of  $\sigma^F$  in the forespore. The activation and control of sporulation-specific transcription factors downstream of  $\sigma^F$  takes place in a cascade. The *spoIIA* operon (*spoIIAC* encodes  $\sigma^F$ ) is transcribed by RNA polymerase containing  $\sigma^H$  and activated by Spo0A~P.  $\sigma^F$  is present therefore in the predivisional cell but inhibited in directing transcription. Upon asymmetric septation, this inhibition is relieved selectively in the forespore but maintained in the mother cell. As a result,  $\sigma^F$ -directed gene transcription is able to occur in a compartment-specific manner. Although it remains unknown how this inhibition is selectively overcome, the components of  $\sigma^F$  regulation are increasingly well understood.

### 1.8.1 Effectors of $\sigma^F$ regulation



**Figure 1-15: Transcriptional control of the  $\sigma^F$  regulatory system.**

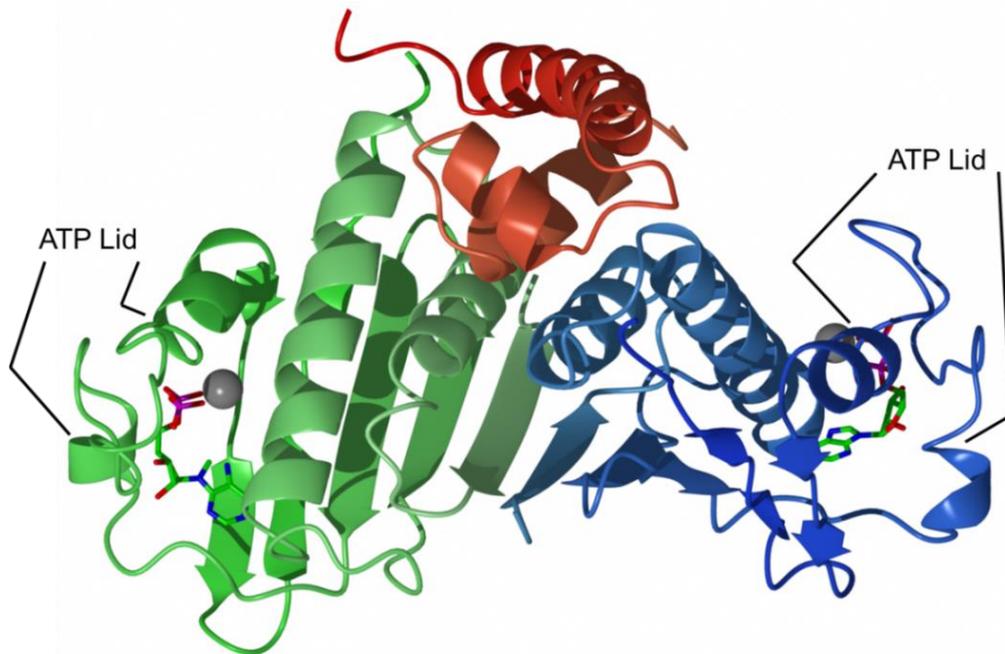
Solid black arrows show transcriptional activation. The black arrow indicates transcription and translation resulting in the four component protein system. The curved grey arrows indicate activation, whereas curved grey barred lines indicate inhibition. Adapted from Margolis, 1991.

Early genetic studies of this system identified a simple scheme for  $\sigma^F$  regulation [117] (Figure 1-15). The *spoIIA* operon consists of three genes; *spoIIAA*, *spoIIAB* and *spoIIAC* (*sigF*) [118]. SpoIIAA and SpoIIAB are an anti-anti-sigma factor and an anti-sigma factor, respectively. A third protein, encoded by the remote monocistronic *spoIIIE* gene, SpoIIIE, plays an important role in promoting SpoIIAA activity. Both the *spoIIA* and *spoIIIE* operons are under the transcriptional regulation of Spo0A~P, though *spoIIIE* is transcribed by E- $\sigma^A$ , rather than E- $\sigma^H$ .

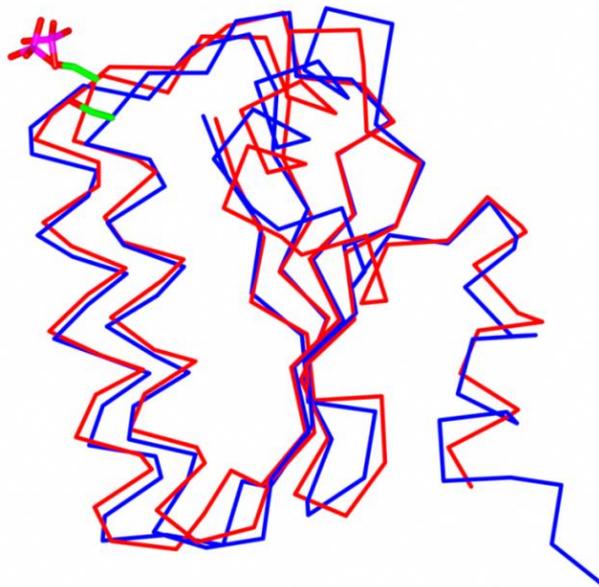
### 1.8.1.1 SpoIIAB

SpoIIAB, a homodimeric protein kinase with 16.2 kDa protomers, has been shown to inhibit  $\sigma^F$  directed transcription *in vitro* in an ATP-dependent manner [78]. This indicates a direct inhibition of  $\sigma^F$  by SpoIIAB-ATP. Further *in vitro* investigations of the SpoIIAB: $\sigma^F$  binding stoichiometry indicate a ratio of one SpoIIAB dimer to one molecule of  $\sigma^F$  in their complex [119]. This stoichiometry is observed in the co-crystal structure obtained of SpoIIAB: $\sigma^F$  from *B. stearothermophilus* [120] (Figure 1-16). The structure shows that SpoIIAB is a member of the GHKL superfamily (named after its members; Gyrase, HSP90, histidine kinase and MutL [121]) of phosphoryl transferases, all of which transfer the  $\gamma$ -phosphate of ATP to their substrate molecules. Structurally, the GHKL family is marked out by the Bergerat fold, an  $\alpha\beta$  sandwich consisting of four antiparallel  $\beta$ -strands and four  $\alpha$ -helices, which is important for ATP binding. Dimerisation of SpoIIAB occurs through the pairing of  $\beta$ 1-strands so that a continuous 10 stranded  $\beta$ -sheet spanning the dimer is formed. The small  $\alpha$ 3 and  $\alpha$ 4 helices form a pocket that constitutes the ATP binding site. The so called ATP lid is formed by the  $\alpha$ 3- $\alpha$ 4 loop which caps the pocket. This feature is thought to cause open and closed states which mediate nucleotide binding and cause a barrier to rapid nucleotide exchange [86].

$\sigma^F$  is largely disordered in the SpoIIAB<sub>2</sub>: $\sigma^F$  crystal structure such that <20% of the sigma factor, constituting the  $\sigma_3$  region, is defined in the electron density maps. This region, consisting of three  $\alpha$ -helices, binds across the interface of the SpoIIAB dimer. The  $\sigma_3$  region of  $\sigma^A$  was shown to be an integral component of the interface with the core enzyme in the structure of the RNAP holoenzyme from *T. aquaticus* (Figure 1-8). Therefore SpoIIAB appears to prevent formation of the holoenzyme, E- $\sigma^F$ . The active site of SpoIIAB consists of ATP and Mg<sup>2+</sup> binding residues and is remote from the  $\sigma^F$  binding site. Rather than acting on  $\sigma^F$ , the kinase activity of SpoIIAB has a different substrate, SpoIIAA, the anti-anti-sigma factor.



**Figure 1-16: The SpoIIAB<sub>2</sub>:σ<sup>F</sup> crystal structure from *B. stearothermophilus*, PDB: 1L00.** The two SpoIIAB protomers making up the dimer are coloured separately in green and blue. The ATP Lid loops are labelled for each monomer. The active site Mg<sup>2+</sup> ions are shown in grey and bound molecules of ADP are shown as cylinders coloured by atom. The visible portion of σ<sup>F</sup> is shown in red, the rest of σ<sup>F</sup> is disordered in the crystal.

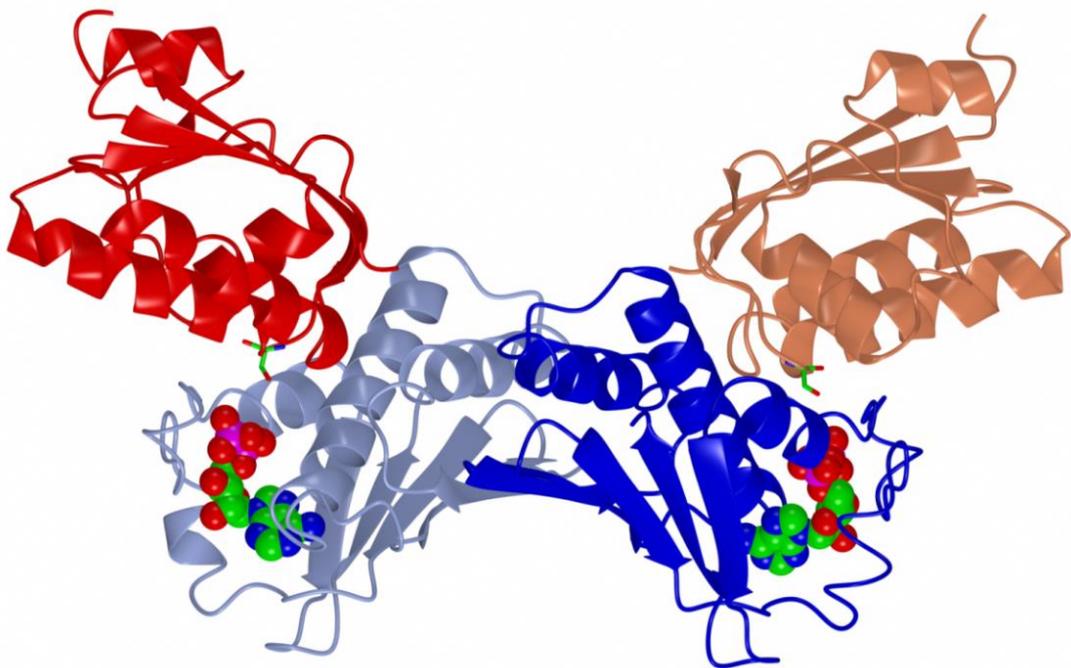


**Figure 1-17: α trace superposition of SpoIIAA from *B. sphaericus* in both phosphorylated and unphosphorylated forms.**

PDB IDs: 1H4X and 1H4Y. The α traces are coloured in blue and red for the phosphorylated and unphosphorylated molecules, respectively. The phosphorylatable residue, Ser57, is shown in cylinder format coloured by atom with and without the phosphate group. The phosphoserine has been modelled in two, half-occupancy, conformations.

### 1.8.1.2 SpoIIAA

SpoIIAA is a small, monomeric protein of 12.9 kDa. It binds to the protein kinase and anti-sigma factor, SpoIIAB, in the presence of ATP and ADP and is also its substrate [80, 81]. Inhibition of  $\sigma^F$  directed transcription *in vitro*, effected by SpoIIAB in an ATP dependent fashion, was shown to be relieved on addition of SpoIIAA [79]. In addition, it was observed that this relief is temporary, as  $\sigma^F$  inhibition was restored after a few minutes. SpoIIAA induces the release of  $\sigma^F$  from SpoIIAB. Accompanying this release is the phosphorylation of SpoIIAA by SpoIIAB on a conserved residue, Ser58 [122].



**Figure 1-18: The SpoIIAA<sub>2</sub>:SpoIIAB<sub>2</sub> complex from *B. stearothermophilus*.**

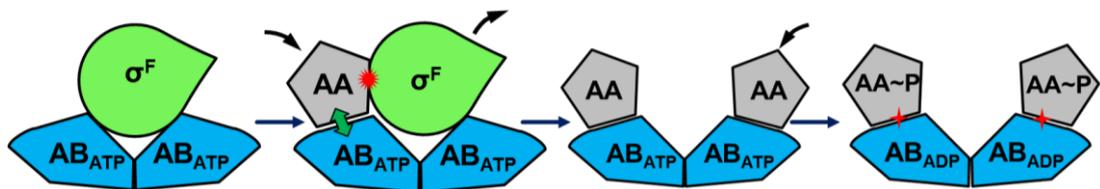
PDB Code: 1TH8. The two SpoIIAB molecules are shown as ribbons in shades of blue, the two SpoIIAA molecules are shown as ribbons in red and coral. The two ADP molecules bound by the SpoIIAB monomers are shown as spheres coloured by atom. The Ser57 residues are also shown as cylinders coloured by atom.

Comparison of crystal structures of *B. sphaericus* SpoIIAA in its phosphorylated and unphosphorylated forms shows very little conformational change which might bring about dissociation from SpoIIAB [123] (Figure 1-17). The SpoIIAA:SpoIIAB interaction, however, must be short-lived in order for  $\sigma^F$  inhibition to be reinstated after only a short time. The co-crystal structure of

SpoIIAA and SpoIIAB:ADP, obtained using proteins from *B. stearothermophilus*, offers some explanation as to how the SpoIIAA:SpoIIAB complex is destabilised upon phosphorylation of SpoIIAA [124] (Figure 1-18). The structure shows two molecules each of SpoIIAA in complex with SpoIIAB<sub>2</sub>. SpoIIAB has the same conformation as when in complex with  $\sigma^F$ .

SpoIIAA molecules interact with SpoIIAB using a surface which is adjacent to and overlapping with that used in  $\sigma^F$  binding with Ser57 (analogous to Ser58 in *B. subtilis* SpoIIAA) positioned near to the  $\beta$ -phosphate groups of ADP in the kinase active site. It is thought that the phosphorylation of SpoIIAA would result in the transferred phosphate group adopting a sterically and electrostatically unfavourable position with respect to the ADP  $\beta$ -phosphate and the side chain of glutamate-46 on SpoIIAB.

SpoIIAA is thought to induce release of  $\sigma^F$  from its inhibitory complex with SpoIIAB [124-126] (Figure 1-19). The asymmetry of the SpoIIAB: $\sigma^F$  interaction results in a larger SpoIIAA binding surface being available on one molecule of SpoIIAB than the other. Complete association of SpoIIAA with this SpoIIAB subunit cannot occur without dissociation of  $\sigma^F$ . As SpoIIAA docks with SpoIIAB in a zipper-like fashion, electrostatic and steric repulsion develops between SpoIIAA and  $\sigma^F$ . This induces the release of  $\sigma^F$  from the complex. A second molecule of SpoIIAA can then bind to the second SpoIIAB subunit. At this point phosphorylation of SpoIIAA can occur.



**Figure 1-19: Schematic of the induced release of  $\sigma^F$  from SpoIIAB<sub>2</sub>: $\sigma^F$  by SpoIIAA.** The stages involved are discussed in the text.

### **1.8.1.3 SpoIIE**

The activation of  $\sigma^F$  requires the sustained regeneration of SpoIIAA from SpoIIAA~P. SpoIIAA~P is the substrate of the protein serine phosphatase, SpoIIE [83]. This protein possesses a PP2C (protein phosphatase type 2C) phosphatase domain which catalyses the dephosphorylation of SpoIIAA~P. Mutational studies of SpoIIE have shown it plays a key role in both activation of  $\sigma^F$  and in the formation of the asymmetric division septum, both of which are hallmark features of sporulation stage two. This suggests SpoIIE has a role in co-ordinating the completion of asymmetric septum formation with activation of  $\sigma^F$ ; this is discussed further in section 1.10.

## 1.9 The Z-ring and Asymmetric Cell Division

Cell division has been intensively studied, as a key reproductive phenomenon in all living systems. In *B. subtilis*, for cell division during growth and sporulation, the cell division machinery is assembled at the site of formation of a ring-like structure by the key cell division protein, FtsZ (filamentation temperature sensitive Z) [127] (Figure 1-11A). While the machinery used is the same in both modes of cell division, the site of assembly is different and relocation must take place.

### 1.9.1 FtsZ

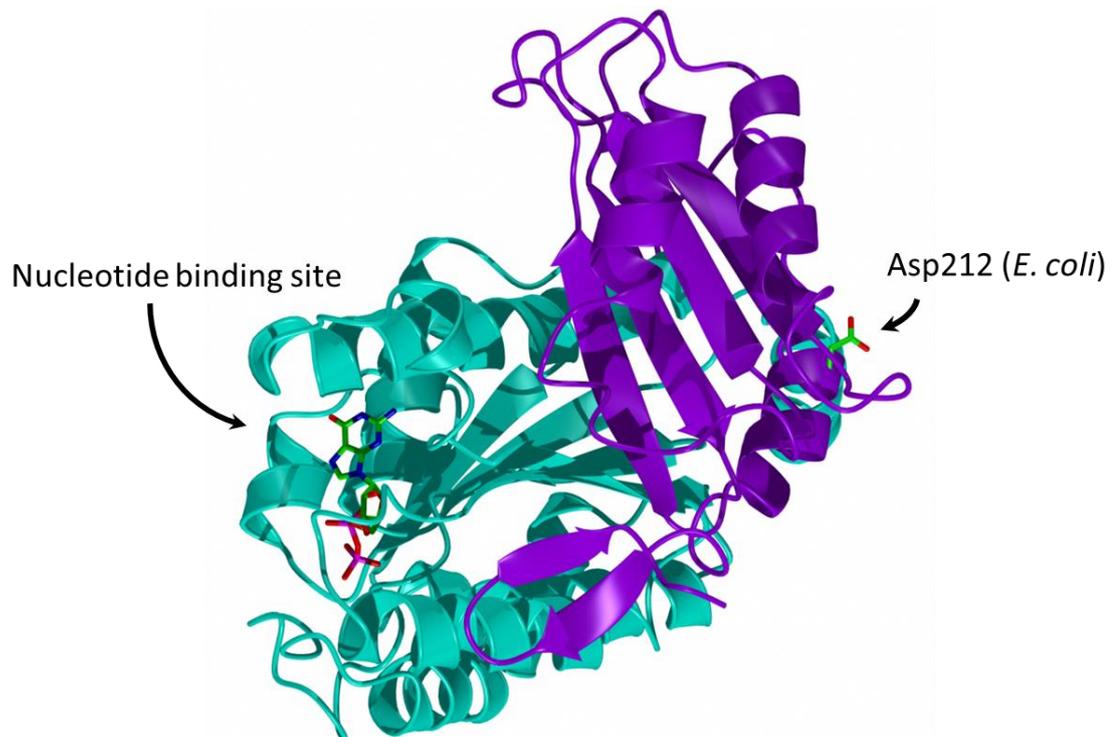
FtsZ is a protein of 40 kDa with 12% sequence identity to tubulin, a protein which forms microtubules during eukaryotic cell division. The level of sequence conservation is highest in regions of tubulin which mediate GTP binding. A glycine rich motif in FtsZ, GGGTGTG is a close match to the tubulin signature motif, (G/A)GGTGST which is part of the GTP binding sequence [128]. FtsZ is able to bind and hydrolyse GTP. FtsZ and tubulin-like proteins form a distinct family of GTPases with two domains connected by a central helix [129].

When the chromosomal copy of *ftsZ* is deleted and replaced with a copy under the control of an IPTG inducible promoter, it is possible to control FtsZ production through variation of the IPTG concentration [130]. When such cells were grown in the presence of IPTG, normal division was seen to occur, however, in the absence of IPTG, the cells were unable to divide and long filaments were observed. Sporulation of these cells was also impaired and asymmetric division septa were not formed.

Upon GTP binding, FtsZ forms long straight structures called proto-filaments, which have the diameter of a single FtsZ molecule [131, 132]. Under the same circumstances, tubulin forms hollow microtubules [133]. Hydrolysis of GTP induces curvature in the proto-filament. The formation of curved proto-filaments can also be induced on addition of GDP. No filament formation is induced in the presence of GMP, suggesting that formation of proto-filaments is

a product of nucleotide binding, rather than of nucleotide hydrolysis itself [134].

The FtsZ crystal structure (Figure 1-20) has been solved from multiple organisms, the first being *Methanococcus jannaschii*, a thermophilic methanogen [135]. The structure shows a highly conserved core consisting of two domains entitled Nt and Ct, and a bound molecule of GDP. Flexible peptide regions at both the N- and C- termini are not observed in the crystal structure. Truncations of FtsZ without the Ct domain are blocked in both polymerisation and GTP hydrolysis [136]. The GTP binding motif is present in the Nt domain while the key catalytic residue for GTP hydrolysis, Asp238 is in the Ct domain, these loci are conspicuously at opposite ends of the molecule.

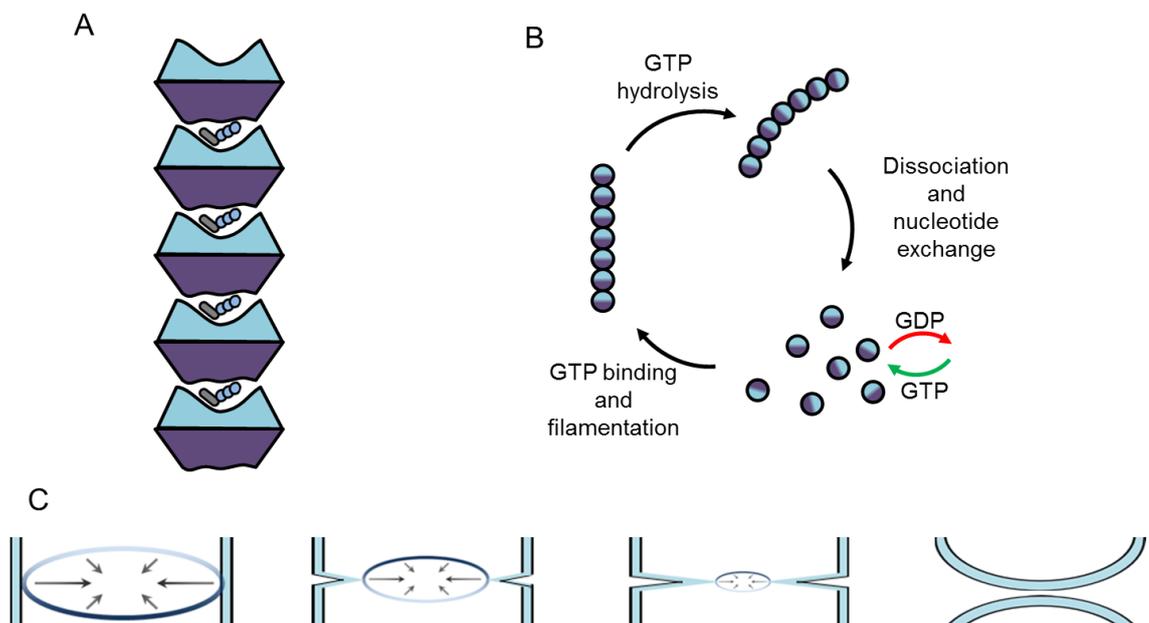


**Figure 1-20: Structure of FtsZ from *Methanococcus jannaschii*.**

PDB ID: 1FSZ. The Nt and Ct domains are coloured in cyan and purple, respectively. A bound GDP molecule is shown as cylinders coloured by atom in the nucleotide binding site. The side chain of Asp212 (*E. coli* numbering; Asp238 in *M. jannaschii*) is also shown as cylinders coloured by atom. Association of Asp212 with the nucleotide binding site completes the GTPase active site Lowe and Amos, 1998 [135].

The binding of GTP allows the Ct and Nt domains of separate monomers to associate about the GTPase active site resulting in the formation of a linear filament (Figure 1-21A). GTP hydrolysis leads to curvature of the filament as the

GDP occupancy increases. Depletion of GTP destabilises the filament causing its dissociation [137]. The nucleotide binding site of monomeric FtsZ is solvent accessible allowing efficient nucleotide exchange of GDP for GTP so that filaments can be reformed (Figure 1-21B). Recent work has indicated that a hinge opening mechanism operates in response to GTP hydrolysis [138]. The so-called T3 loop mediates binding of the GTP  $\gamma$ -phosphate by adopting a tense state which supports hydrophobic interactions with the T7 loop of the neighbouring FtsZ subunit. On GTP hydrolysis, the release of the  $\gamma$ -phosphate allows T3 to relax disrupting the interaction with T7. A nest of hydrophobic interactions on the opposite side of the GTP binding pocket, based around Leu296 on one protomer and Phe135, Leu167, Val174 and Leu176 on the other, forms a hinge which is pushed open by the disruption of the T3-T7 interface.



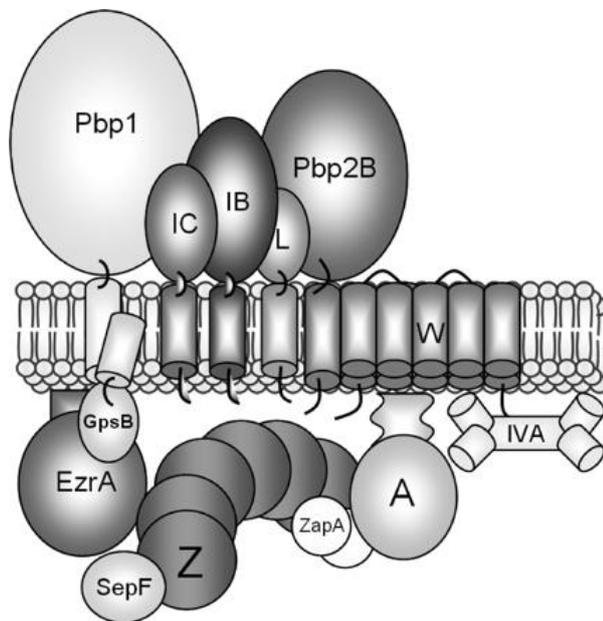
**Figure 1-21: Functional behaviour of FtsZ proto-filaments.**

The assembly of protofilaments by binding of GTP by the Nt subdomain, cyan, and active site completion by the Ct subdomain, purple, of another FtsZ monomer. (B) The GTP hydrolysis cycle of FtsZ features filamentation, hydrolysis, curvature, dissociation, nucleotide exchange and reassembly. (C) FtsZ drives constriction of the cell membrane towards cell division by repetition of the GTP hydrolysis cycle or ‘iterative pinching’.

It has been shown that the assembly, curvature, dissociation and reassembly of FtsZ proto-filaments provide all of the force necessary to drive membrane constriction at the cell division site through a process termed iterative pinching [139, 140] (Figure 1-21C). Although wild-type FtsZ alone is not capable of

interaction with the cell membrane, the addition of a C-terminal amphipathic helix and an N-terminal YFP (yellow fluorescent protein) domain caused the manifestation of indentations in liposomes in the presence of GTP. The indentations formed in locations with FtsZ-rings, as visualised by YFP fluorescence. The natural anchoring of FtsZ to the cell membrane is enabled by interaction with FtsA and numerous other proteins involved in a complex cell division machinery assembly.

### 1.9.2 The *B. subtilis* divisome



**Figure 1-22: The *B. subtilis* divisome.**

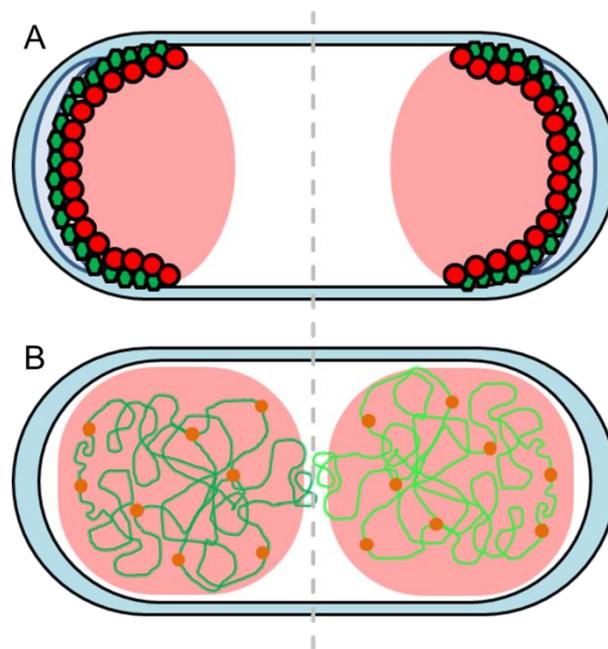
Protein label abbreviations are as follows: Z; FtsZ, A; FtsA, L; FtsL, W; FtsW, IB; DivIB, IC; DivIC, IVA, DivIVA. The first proteins to assemble are FtsZ, FtsA, ZapA and EzrA. The remaining proteins assemble some time later. Taken from Gamba, 2009 [141].

A complex cell division apparatus, called the divisome, assembles subsequent to Z-ring formation. Assembly takes place in two-steps [141]. Initially, FtsA, ZapA and EzrA localise with FtsZ. FtsA anchors the Z-ring to the cell membrane, ZapA is a promoter of Z-ring formation and EzrA is required for subsequent localization of cell-wall synthesis enzymes. A second wave of proteins is co-opted some time later and includes Pbp2B, FtsL, DivIB and DivIVA the roles of which are in cell-wall synthesis, regulation of cell division, coupling to cell cycle and division site regulation, respectively [142, 143]. The assembly of the second

set of divisome proteins seems highly concerted, suggestive of a complex web of interactions [141]. A full schematic of proteins comprising the *B. subtilis* divisome, originally presented by Gamba et al., 2009 [141], is shown in Figure 1-22.

### 1.9.3 Regulating Z-ring assembly during growth

Vegetative cell division requires the assembly of the cell division machinery at the middle of the cell. While FtsZ filaments can form only in the presence of GTP, which will be abundant in the cytoplasm, for much of the cell cycle Z-rings are not observed [144, 145]. This suggests the existence of one or more systems negatively regulating Z-ring formation. Two primary regulatory systems, MinCD and nucleoid occlusion (Figure 1-23) have been identified in *B. subtilis*.



**Figure 1-23: Direction of Z-ring formation to mid-cell in vegetative cell division by MinCD and Noc.**

The medial division site is indicated by the grey dashed line. Red transparent areas indicate regions where Z-ring formation is inhibited. (A) The MinCD system. MinC (red circles) and MinD (green hexagons) co-localise with DivIVA (light blue crescent) at the cell poles where MinC inhibits Z-ring formation. (B) The Nucleoid Occlusion (Noc) protein non-specifically binds to the chromosome and prevents cell division machinery assembly in its proximity. Adapted from Jamroskovic et al., 2012 [146] and Wu et al., 2004 [147].

The Min system was first characterised in *E. coli* and is so called because *min* mutants are associated with the formation of anucleate minicells [148, 149].

The three component system, encoded by the *minB* locus, prevents asymmetric cell division by localization of the Z-ring formation inhibitor, MinC, to the cell poles through interaction with a membrane bound ATPase, MinD, and a cell pole localising protein MinE [150-152].

In *B. subtilis*, there is an analogous system consisting of MinC and MinD. However the MinE is absent and its role appears to be played by DivIVA [146]. Interestingly for a protein which anchors proteins to the cell membrane, DivIVA does not possess an amphipathic membrane binding region; however, the crystal structure of *B. subtilis* DivIVA reveals a tetrameric assembly containing a coiled coil with distinct curvature [153]. The distribution of hydrophilic and hydrophobic residues around the helices suggests a directed interaction with the polar head groups and surface proximal lipids in the cell membrane, allowing recognition of concave cell membrane curvature. DivIVA has also been shown to exist as hexamers and octamers [154]. Through interaction with DivIVA, MinC and MinD are anchored to the cell poles where they prevent the formation of Z-rings (Figure 1-23A).

The Noc protein, named for nucleoid occlusion, has been observed to bind non-specifically throughout the chromosome and to prevent the assembly of the cell division apparatus (Figure 1-23B) [147]. This association directs cell division to the area with the lowest occupancy of chromosome. Post-DNA-replication, this uninhibited area is the medial division site. Additionally, Noc has a role in concentration of the divisome components by preventing their even distribution around the cell. In the absence of both MinCD and Noc, cells do not divide at all. In this case there is no preference for any cell division site and cell division machinery is too diffuse to allow assembly of a functional divisome.

#### 1.9.4 Relocation of the division site during sporulation

The relocation of the divisome to a polar site for the asymmetric cell division characteristic of sporulation requires the up-regulation of transcription of the *ftsAZ* operon encoding FtsZ and FtsA and the expression of *spoIIIE*. Three promoters govern transcription of *ftsAZ*, one of which, P<sup>H</sup>, is under the control

of  $\sigma^H$ , and hence activated during the initial stages of sporulation [155]. Asymmetric cell division is not severely affected by disruption of this promoter. However, when  $P^H$  disruption is coupled with deletion of *spoIIE*, the second effector of polar cell division, cells are severely deficient in asymmetric septum formation [77]. Deletion of *spoIIE* causes a delay in the formation of polar septa in comparison to wild-type. Upregulation of *ftsAZ* and activation of *spoIIE* is sufficient to force the formation of polar division septa during vegetative growth, suggesting that the up-regulation of both is a major driving force for divisome relocation.

In addition to the up-regulation of SpoIIE and FtsZ production, RacA has a role in promoting polar division [76]. RacA causes the formation of the axial filament in which chromosomal DNA is distributed evenly along the axis of the cell. RacA also anchors the chromosomes to opposite cell poles, through an interaction with sites around *oriC*, and DivIVA. This interaction could outcompete MinC and MinD resulting in non-specific localisation of the MinCD inhibitory complex throughout the cell and preventing Z-ring inhibition at the polar sites.

As occurs during binary fission, the Z-ring is initially formed at the mid-cell. The medial Z-ring disassembles and FtsZ relocates to the polar division sites on a helical trajectory. This trajectory is thought to be governed by lipid spirals in the cell membrane, to which FtsZ and FtsA are associated. One site develops into an asymmetric division septum while the other is abandoned and disassembles soon after compartment specific gene expression is established [77, 156].

It is not known how the cell selects a single Z-ring with which to complete division while abandoning the other. It seems that the formation of two Z-rings provides a safety net should the first division event be abortive. For example, RacA mutants fail to sequester DNA in the newly formed forespore. In this circumstance the second division site is activated, presumably in the hope of capturing a chromosome at the opposite cell pole. Additionally, a number of *spoII* mutant alleles show a phenotype in which two asymmetric septa form [157]. These cells are unable to complete sporulation, but the formation of two septa proves that both Z-rings are capable of supporting cell division.

## 1.10 SpoIIE, a Key Cell Fate Determinant

SpoIIE is a 92 kDa, 827 residue protein encoded by the gene, *spoIIE*. Its transcription is directed by E- $\sigma^A$ , activated by the master sporulation initiator, Spo0A~P [158]. SpoIIE is bifunctional, being required for both correct formation of the asymmetric division septum during sporulation, as described above, and in the indirect activation of the compartment specific  $\sigma$  factor,  $\sigma^F$ .

### 1.10.1 Developmental behaviour of SpoIIE

The subcellular localisation of SpoIIE and various mutants has been studied using GFP-fusion constructs and microscopic visualisation techniques [67, 159, 160]. SpoIIE, produced prior to the sporulation cell division, promotes the formation of a pair of polar Z-rings to which it co-localises [159, 161]. SpoIIE remains at the sporulation septum but dissociates from the abortive Z-ring. On completion of the asymmetric division septum, activation of  $\sigma^F$  is observed in the forespore. This indicates the release of some block on the phosphatase activity of SpoIIE. After activation, SpoIIE delocalises from the division septum becoming diffuse in the forespore membranes [162].

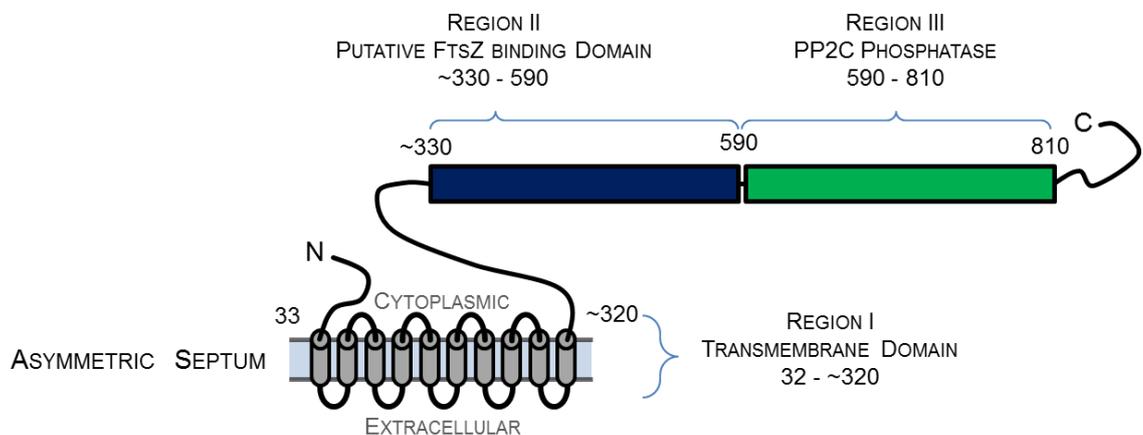
### 1.10.2 Three Domain structure

SpoIIE is thought to consist of three domains [163] (Figure 1-24). At the N-terminus, a 32 residue sequence precedes a transmembrane domain of around 300 residues. This region is highly hydrophobic and sequence analysis and topology studies of SpoIIE predicted the presence of 10 putative transmembrane helices [163]. The N-terminal peptide is predicted to lie on the cytoplasmic side of the membrane, as are the C-terminal domains II and III. This domain accounts for around two fifths of the protein.

The central portion of SpoIIE, from residues ~330 to ~590, is of unknown function, its primary sequence shows low homology to non-SpoIIE proteins and has the lowest level of conservation of the SpoIIE domains (17% between *B. subtilis* and *C. acetobutylicum*). There is evidence that SpoIIE interacts with FtsZ

from both yeast two-hybrid assays and size exclusion chromatography analysis. Both experiments indicate that domain II is required for this interaction which is not formed between FtsZ and the phosphatase domain (domain III) [164].

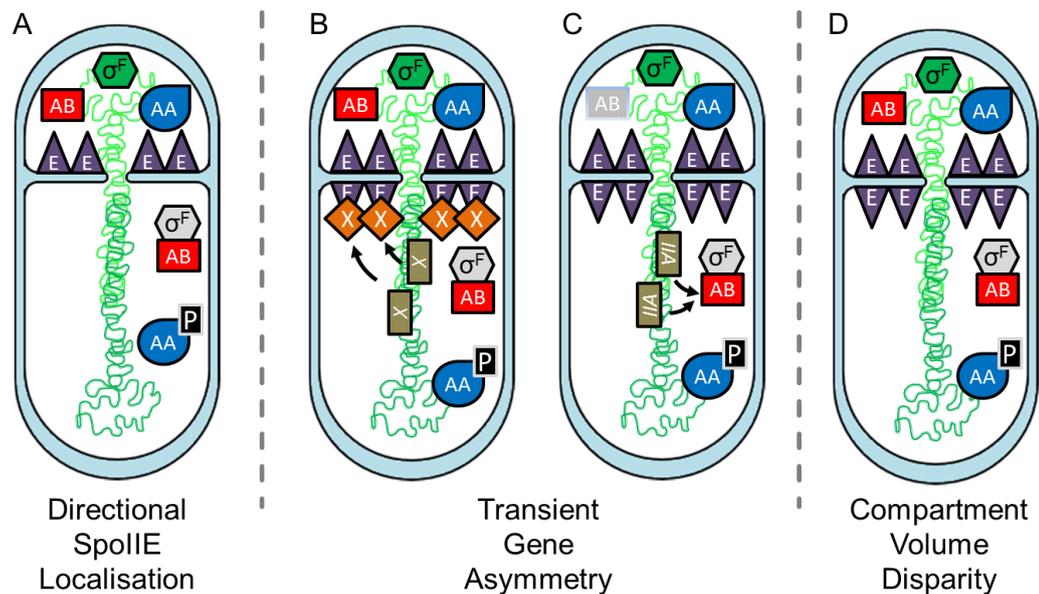
The C-terminal region of the protein is more highly conserved and sequence analysis indicates a PP2C phosphatase domain constituted of residues 590 to 810, with a number of highly conserved clusters of residues forming the signature of the PP2C family. The PP2C family of phosphatases are part of the broader PPM superfamily of serine/threonine phosphatases, found in both eukaryotes and prokaryotes. PPM phosphatases are metal dependent and in the case of PP2C domains two or three manganese ions are bound by aspartate or glutamate residues in the active site. Mechanistically, it is proposed that the metal cations activate water molecules for nucleophilic attack at the phosphorus of the Ser/Thr-O-PO<sub>3</sub> group during dephosphorylation [165, 166]. This domain of SpoIIIE catalyses the dephosphorylation of the phosphoserine (P-Ser58) of the anti-anti-sigma factor, SpoIIAA~P.



**Figure 1-24: The three domain structure of the sporulation cell fate determinant, SpoIIIE.** Shown are the three putative domains of SpoIIIE. The central region (II) and PP2C binding domain are shown as purple and green rectangles, respectively. The transmembrane domain is shown as a schematic of its membrane association pattern and orientation of the N- terminal peptide and C-terminal cytoplasmic domains is shown as in the cytoplasm rather than extracellular. The N- and C-termini and the domain boundaries, as well as they are understood, are also indicated.



Following asymmetric cell division (Figure 1-25B), dephosphorylation of SpoIIAA~P in the forespore by SpoIIIE allows the dissociation of SpoIIAB<sub>2</sub>:σ<sup>F</sup> by SpoIIAA. Although SpoIIAB can phosphorylate SpoIIAA, the rate of this reaction is limited by slow exchange of ADP for ATP. Thus in the forespore, SpoIIAA~P is converted to SpoIIAA by SpoIIIE, SpoIIAB is unable to sequester σ<sup>F</sup>, and the latter is free to combine with core RNAP and direct transcription (Figure 1-25C). The inhibitory equilibrium in the mother cell remains in place, although the process by which this is achieved is unknown.



**Figure 1-26: Possible mechanisms for the establishment of compartment specific gene expression directed by  $\sigma^F$ .**

Each model is described in the text. (A) Localisation of SpoIIIE to only the forespore side of the asymmetric division septum. No turnover of SpoIIAA~P occurs in the mother cell. The forespore inhibitory equilibrium is disrupted by production of SpoIIAA by SpoIIIE. (B) Enrichment of a putative SpoIIIE inhibitor, X, in the mother cell caused by transient gene asymmetry before forespore chromosome translocation. (C) Degradation of SpoIIAB by the ClpCP protease system. Transient gene asymmetry causes the presence of two spoIIA operons in the mother cell allows replenishment of degraded SpoIIAB. The SpoIIAB in the forespore is not replenished, preventing competition for  $\sigma^F$ . (D) The change in overall SpoIIAA concentration caused by SpoIIIE activity is much lower in the mother cell than the forespore due to the difference in compartment volume.

An early study in which mother cell and forespore compartments were converted into protoplasts by lysozyme treatment led to the proposal that SpoIIIE is confined to the forespore side of the polar septum [84] (Figure 1-26A). While this would lead to selective SpoIIAA~P dephosphorylation in the forespore, and therefore specific activation of  $\sigma^F$  in this compartment, later

observations indicate that SpoIIE is in fact present on both sides of the septum [159].

Another hypothesis for how SpoIIE might be activated selectively in the forespore is based on transient gene asymmetry between the newly formed forespore and the mother cell [85]. The asymmetric septum initially traps only 30% of the chromosome in the forespore [67]. The remaining 70% remains in the larger compartment and is pumped into the forespore by the septum localising DNA translocase SpoIIIE over a period of 20 minutes. During this period the forespore is deficient in genes on the distal portion of the chromosome, while the mother cell possesses two copies of these genes. It was proposed that an unidentified inhibitor of SpoIIE may be encoded on this region of the chromosome (Figure 1-26B). The production of the inhibitor would be confined initially to the mother cell allowing SpoIIE at the forespore side of the division septum to be activated selectively.

Another proposal relying on transient gene asymmetry is based on the sensitivity of SpoIIAB to proteolysis. The *spoIIA* operon encoding SpoIIAA, SpoIIAB and  $\sigma^F$ , is in the *oriC* distal region of the chromosome and will initially be present as two copies in the mother cell but absent in the forespore. SpoIIAB has been shown to be a target of the ClpCP protease system. It can be therefore be envisaged that SpoIIAB would be depleted in the forespore by ClpCP degradation leading to  $\sigma^F$  activation. This would be a selective action as SpoIIAB levels in the mother cell would be replenished by *spoIIA* gene expression [167] (Figure 1-26C).

The volume of the forespore is around one-eighth of that of the mother cell, a large disparity [86]. If the number of molecules of SpoIIE is the same on both sides of the septum, the concentration of SpoIIE in the mother cell would be eight times lower (Figure 1-26D). *In vitro* experiments have shown that a 10-fold increase in SpoIIE activity is sufficient to switch  $\sigma^F$  from being fully inhibited to being fully activated. The magnitude of this increase is similar to the volume disparity between the mother cell and forespore [168]. This hypothesis seems to be a very credible explanation of how compartment specific gene expression is initiated during sporulation of *B. subtilis*.

## 1.11 Aims

SpoIIE has been rigorously studied through genetic investigation [169]. However, to date only the C-terminal phosphatase domain has been structurally characterised [170]. Domain II of SpoIIE is required for localisation of the protein to the asymmetric division septum, but mutations within domain II have also been shown to have strong influences on phosphatase domain activity. This implies autoregulation of SpoIIE activity. The mutations identified require a structural context to be fully explained. Of particular interest is how domain II and the phosphatase domain (domain III) interact to inhibit phosphatase activity until asymmetric cell division has occurred, this being of paramount importance for the correct and compartmentalised activation of  $\sigma^F$  in the forespore.

The principal aim of my studies has been the cloning, overexpression and purification of soluble *B. subtilis* SpoIIE domain II constructs to allow structural characterisation and definition of the interaction between the two cytoplasmic domains. This involved the development of a fragment previously identified using a library-based fragment solubility screen. An investigation into the domain boundary between the transmembrane and cytoplasmic regions of the protein yielded a fragment containing the entirety of domain II. An additional fragment was developed, potentially isolating a sub-domain in SpoIIE domain II. Orthologues of SpoIIE fragments from *G. stearothermophilus* were also produced. This work is detailed in Chapter 3.

Another interaction in the regulation system of  $\sigma^F$  is that between SpoIIE and SpoIIAA~P. While structures of complexes have been solved to show the interactions between  $\sigma^F$  and SpoIIAB and SpoIIAB and SpoIIAA, the mode of interaction between SpoIIE and its specific substrate remains uncharacterised. The development of a co-expression system to produce large quantities of SpoIIAA~P and some attempts at isolating a SpoIIE:SpoIIAA~P complex for structural characterisation are detailed in Chapter 4.

Chapter 5 describes an *in vivo* mutagenesis study carried out, using some suppressor mutations identified in SpoIIE domain III, in order to provide insight into how domains II and II interact to regulate phosphatase activity.

## Chapter 2: Materials and Methods

This chapter covers core cloning, protein production and protein purification techniques underpinning the experimental programme of this thesis. The results chapters in this work contain additional methods sections specific to the investigations described.

### 2.1 Agarose gel electrophoresis

Experiments involving the manipulation and modification of DNA, such as molecular cloning and mutagenesis, require analysis of DNA fragment size and purity. The electrophoretic migration of DNA through a low percentage agarose gel matrix is very commonly used for this purpose. At neutral pH each phosphate group on DNA carries a negative charge so that DNA fragments of all sizes have a constant mass to charge ratio. As a result many of the most effective separation techniques for DNA use electrophoresis. As the negatively charged DNA migrates towards the positively charged cathode the larger fragments are hindered by the agarose matrix more regularly than the smaller fragments. This causes a separation of fragment sizes as the smallest fragments migrate furthest. Linear and circular DNA fragments of similar sizes also have different migration characteristics. Supercoiled, relaxed circular and linear DNA forms are readily resolved owing to their different levels of compactness. Resolved DNA species form distinct bands that can be visualised by the inclusion of SYBRsafe (Invitrogen), which intercalates the DNA bases leading to enhanced fluorescence emission in the visual region of the spectrum on absorbance of UV light. A DNA ladder, consisting of DNA fragments of known sizes, which encompass those expected of the analyte(s), is typically used to allow estimation of unknown DNA fragment sizes.

A 1% agarose gel is prepared by dissolving 1 g of agarose (a polysaccharide extracted from seaweed) in 100 ml of Tris-acetate-EDTA (TAE) buffer (40 mM Tris-acetate pH 7.6, 1 mM EDTA) by heating. SYBRsafe DNA stain is added to the cooling solution at a 10,000 fold dilution (according to manufacturer's instructions) before it is poured into a casting mould appropriate to the

electrophoresis apparatus. A comb is added to form wells for sample application. The solidified gel is placed in the electrophoresis apparatus containing TAE buffer. Sample preparation involves the addition of 1  $\mu\text{l}$  of a sample loading buffer, such as Blue Orange loading dye (Promega), to every 5  $\mu\text{l}$  of DNA sample. The sample loading buffer contains glycerol (to lend the samples viscosity to keep them in the wells) and visible dyes, indicating the progress of species through the gel through a visible dye front. A 100 V potential difference is applied across the gel for approximately 60 minutes, although this may vary depending on the size of the fragments being resolved and the extent of resolution required. DNA bands can be observed using a 302 nm UV trans-illuminator, as the DNA-bound SYBRsafe dye fluoresces.

## 2.2 Small Scale DNA Preparation

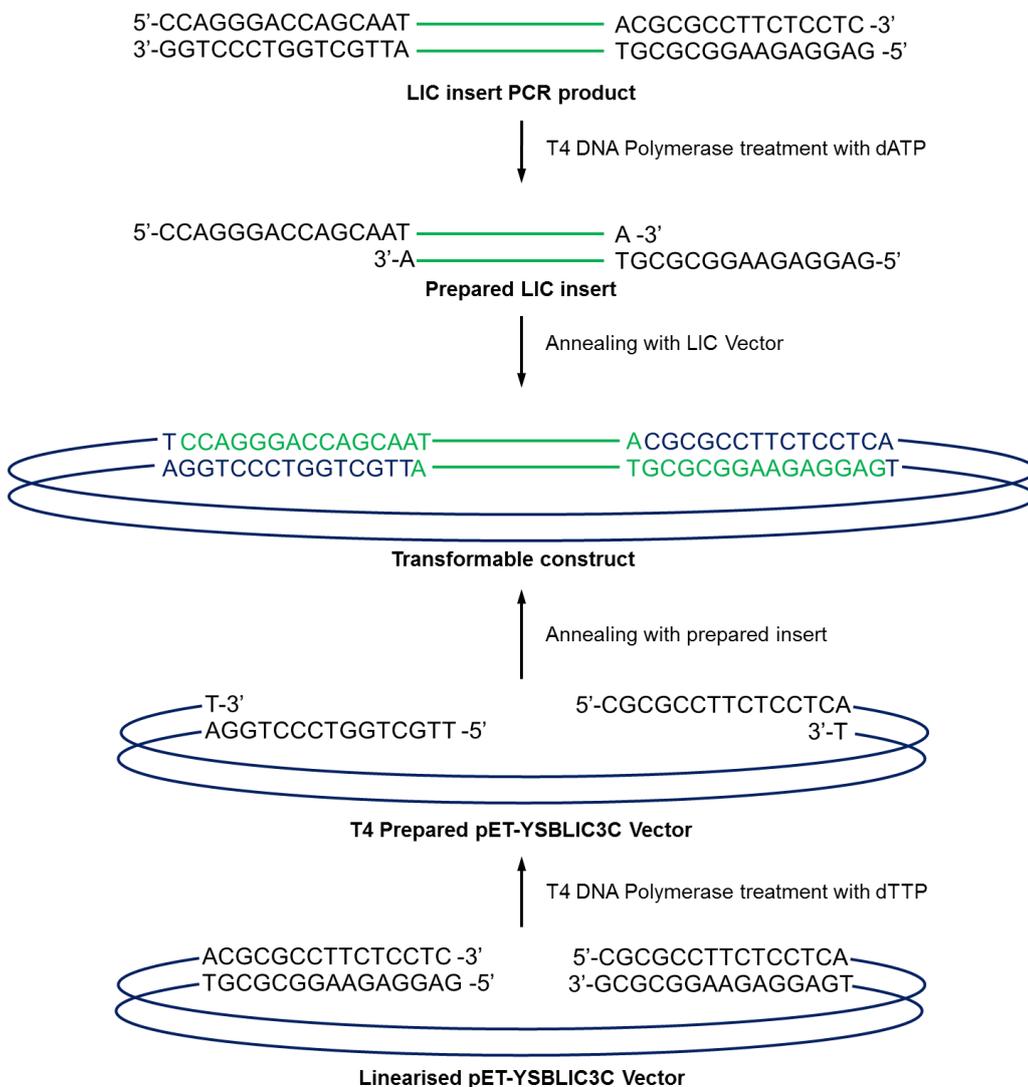
The isolation of plasmid DNA from transformed cell cultures is fundamental to the molecular cloning process. It is commonly used for screening of transformant colonies, in which DNA from 5 to 10 separate colonies is extracted for analysis, or for preparation of a plasmid from a previously transformed source when DNA from a single culture is isolated.

For each clone, 10 ml of LB media supplemented with 30  $\mu\text{g}/\text{ml}$  kanamycin is inoculated with cells from a colony using a sterile pipette tip. After overnight growth, the cultures are centrifuged for 10 minutes at 5,000 x g in a benchtop centrifuge and the supernatant discarded. The plasmid DNA from each cell pellet is then extracted and purified using a QIAprep miniprep kit (Qiagen). The concentration of DNA is determined by measuring the absorbance at a wavelength of 260 nm.

## 2.3 Construct preparation

### 2.3.1 Ligation Independent Cloning

Ligation Independent Cloning (LIC) is a method of incorporation of one DNA fragment (the insert) into a suitable linearised plasmid (the vector) resulting in a larger circular plasmid without the requirement of a DNA ligase (Figure 2-1). The modified pET-28a vector, pET-YSBLIC3C [171], is designed so the final construct encodes a Hisx6 affinity tag and an HRV3C protease (Human Rhinovirus 3C protease) cleavage site fused to the N-terminus of the target protein.



**Figure 2-1: Ligation Independent Cloning.**

A PCR amplified insert (top) and a linearised LIC vector, pET-YSBLIC3C (bottom), are separately prepared by exonuclease treatment with T4 DNA polymerase. When mixed, the two reagents anneal forming a stable species which can be taken up by cells.

The insert is produced by amplification of the target sequence from a DNA template by Polymerase Chain Reaction (PCR). By suitable design of the oligonucleotide primers used in the PCR, the amplified target DNA sequence is bracketed by regions homologous to those in the plasmid. The treatment of both the linearised plasmid and the PCR product with T4 DNA polymerase in the presence of a single nucleotide, resulting in exonuclease activity, yields complementary single stranded overhangs at the strand termini. These overhangs are designed to be sufficiently long that the insert and vector anneal when incubated together at room temperature. The base pair interactions in the annealed plasmid are stable enough to allow direct transformation of *E. coli* cells, after which *in vivo* ligation is carried out by the host cell.

### 2.3.2 Polymerase chain reaction

The amplification of a specific DNA sequence, like a gene of interest, from template DNA, such as a chromosome, plasmid or even a previous PCR product, can be performed using a Polymerase Chain Reaction (PCR) [172]. This technique is capable of biosynthetically producing relatively large quantities of the target DNA product, or amplicon, with high-fidelity. Two short single-stranded oligonucleotide primers are required to define the target region and these are extended by DNA polymerase during the reaction. The sequences of the primers are designed to match the nucleotide sequences at the extremes of the region of interest on opposing DNA strands. Numerous commercial DNA polymerases, such as KOD Hot Start Polymerase (Merck) are available, with various extension rates, fidelities and temperature optima. These DNA polymerases are from thermophilic sources and are active only at high temperatures, 72°C in the case of KOD polymerase from *Thermococcus kodakaraensis*.

The components of a typical PCR reaction are shown in Table 2-1. The forward and reverse primers are added in a  $10^6$  fold molar excess to the template DNA (calculated for the 4.12 Mbp *B. subtilis* chromosome). This excess ensures that the template-primer interaction is favoured over a template-template

interaction and that the amount of primer does not limit the yield of the experiment. The DNA polymerase substrates dGTP, dATP, dTTP and dCTP, are added in equimolar concentrations. DNA polymerase is a magnesium dependent enzyme and the concentration of MgSO<sub>4</sub> added to a PCR experiment is varied to optimise enzyme fidelity, facilitate the production of long amplicons or to minimise non-specific primer annealing, incomplete DNA melting and primer dimer formation [173-175].

**Table 2-1: Required PCR components**

Component	Stock Concentration	Volume (µl)
dNTP mix	2 mM	5
KOD DNA polymerase reaction buffer	10x concentrate	5
MgSO <sub>4</sub>	25 mM	2
DNA template	>50 ng/µl	0.5
Forward Primer	20 pmol/µl	1
Reverse Primer	20 pmol/µl	1
KOD DNA polymerase	1 U/µl*	1
milliQ Deionised Water	-	34.5

\*\* The enzyme unit (U) is the amount of an enzyme which increases the rate of reaction in a system by 1 µmol/minute [176].

The PCR reaction consists of 3 main steps (Table 2-2) that take place at different temperatures. Firstly, the dsDNA (double stranded DNA) template is melted at 94°C to produce ssDNA (single stranded DNA). Cooling allows annealing of primers to the target regions on the template ssDNA. The annealing temperature may be varied according to the primer T<sub>m</sub> (melting

temperature) values. After the annealing of primers to the template, the reaction is heated to the optimal temperature of KOD DNA polymerase, 72°C, and extension of the primers occurs in the 5' to 3' direction complementary to the template DNA and is continued for sufficient time to cover the whole region of interest at a rate of around 2 kb/min. Thermal cycling of these stages causes amplification of the target fragment and an exponential increase in the concentration of product.

**Table 2-2: Typical PCR thermocycling parameters**

Step	Temperature (°C)	Duration (s)	No. of Cycles
Initial Denaturation	94	120	1
Denaturation	94	30	35
Annealing	45	30	
Extension	72	40	
Final Extension	72	180	1
Hold	4	∞	1

These parameters are based on a PCR reaction to generate a 1 kb amplicon using KOD hotstart DNA polymerase.

### 2.3.3 Oligonucleotide primers

The addition of non-complementary nucleotide sequences to the 5' end of a primer leads to the incorporation of these sequences at the termini of the final PCR product. These flanking sequences are used in molecular cloning as a means of generating compatible ends for ligation of PCR products into appropriate vectors. In both LIC and ligation dependent cloning, the addition of a specific sequence at the PCR product termini is key to the cloning process.

LIC requires that the ends of both the vector and insert possess particular sequences. T4 DNA polymerase treatment in the presence of dATP produces complementary ends (Figure 2-1). All that is required is that two universal 5' overhang sequences are added to the forward and reverse oligonucleotide primers. The reverse primer extension includes a stop codon.

Forward extension → 5' – CCAGGGACCAGCA – 3'

Reverse extension → 5' – GAGGAGAAGGCGCGTTA – 3'

This method greatly simplifies primer design as cloning using restriction endonucleases requires consideration of open reading frames when positioning cleavage sites in primer overhangs. In contrast, LIC allows a much more formulaic approach to primer design making it an appealing strategy for high throughput cloning experiments [177]. One consideration is that the first codon of the gene of interest must be ATX (where X is any base). This form is required to ensure generation of the correct overhangs by T4 DNA polymerase. When the sequence in the gene of interest does not allow for this form, a common modification is the addition a methionine (Met; M) residue (codon = ATG) to the N-terminus of the fragment via further addition to the primer extension.

Primers for PCR require a length of nucleotide bases complementary in sequence to the DNA template. This complementary sequence directs the primer to bind specifically to the desired site on the template. The length and base composition of the complementary sequence determines the  $T_m$  of the primer – template DNA duplex.  $T_m$  can be calculated using the Wallace-Ikakura rule, where  $G + C$  is the total number of G and C bases and  $L$  is the total length of the primer:

$$T_m(^\circ\text{C}) = 64.9 + \frac{41(G + C - 16.4)}{L}$$

Both forward and reverse primers must be able to anneal to the template DNA simultaneously, so primer pairs should have  $T_m$  values within 5°C of one another. The annealing temperature used during thermocycling may also require modification to accommodate this requirement without allowing non-specific annealing (low temp.) or preventing annealing (high temp.). Ensuring that the terminal base at the 3' end of a primer is a G or C is also advisable. The GC base pair forms three hydrogen bonds rather than two and the additional H-bond makes a more stable interaction and gives more stable annealing at the site of DNA polymerase engagement with the primer-template.

Primer pairs can be designed to user-defined parameters using web based tools such as the NCBI hosted Primer-BLAST tool. These tools are capable of selecting priming sequence pairs from a given template sequence with matching  $T_m$  values and are also designed to identify primer pairs with the risk of forming primer dimers and DNA hairpins.

#### 2.3.4 LIC Vector preparation

The pYSBLIC3C vector, is a kanamycin resistance-conferring, pET-28a plasmid derivative (Appendix I) with a modified multiple cloning site (MCS)(Figure 2-2). pYSBLIC3C contains an engineered BseRI restriction site which allows plasmid linearisation. The DNA sequence encoding the His<sub>6</sub> affinity tag has been modified to encode a downstream HRV3C protease recognition motif (LEVL<sub>FQ</sub>/GP). This inclusion and further modifications beyond the BseRI restriction site provides the sequences necessary for T4 DNA polymerase truncation and vector - insert annealing. Growth and recovery of the pET-Y<sub>S</sub>BLIC3C vector is achieved by transformation of a cloning strain of *E. coli*, such as XL-10 Gold (Agilent), followed by plasmid isolation by DNA miniprep, section 2.2. Linearised vector is then prepared either by BseRI digestion of 50 µg of circular plasmid or by vector PCR using specially designed vector primers. The linearised vector is analysed by agarose gel electrophoresis and the appropriate band is purified by gel extraction.

#### **2.3.4.1 pET-YSBLIC3C Preparation by BseRI digest**

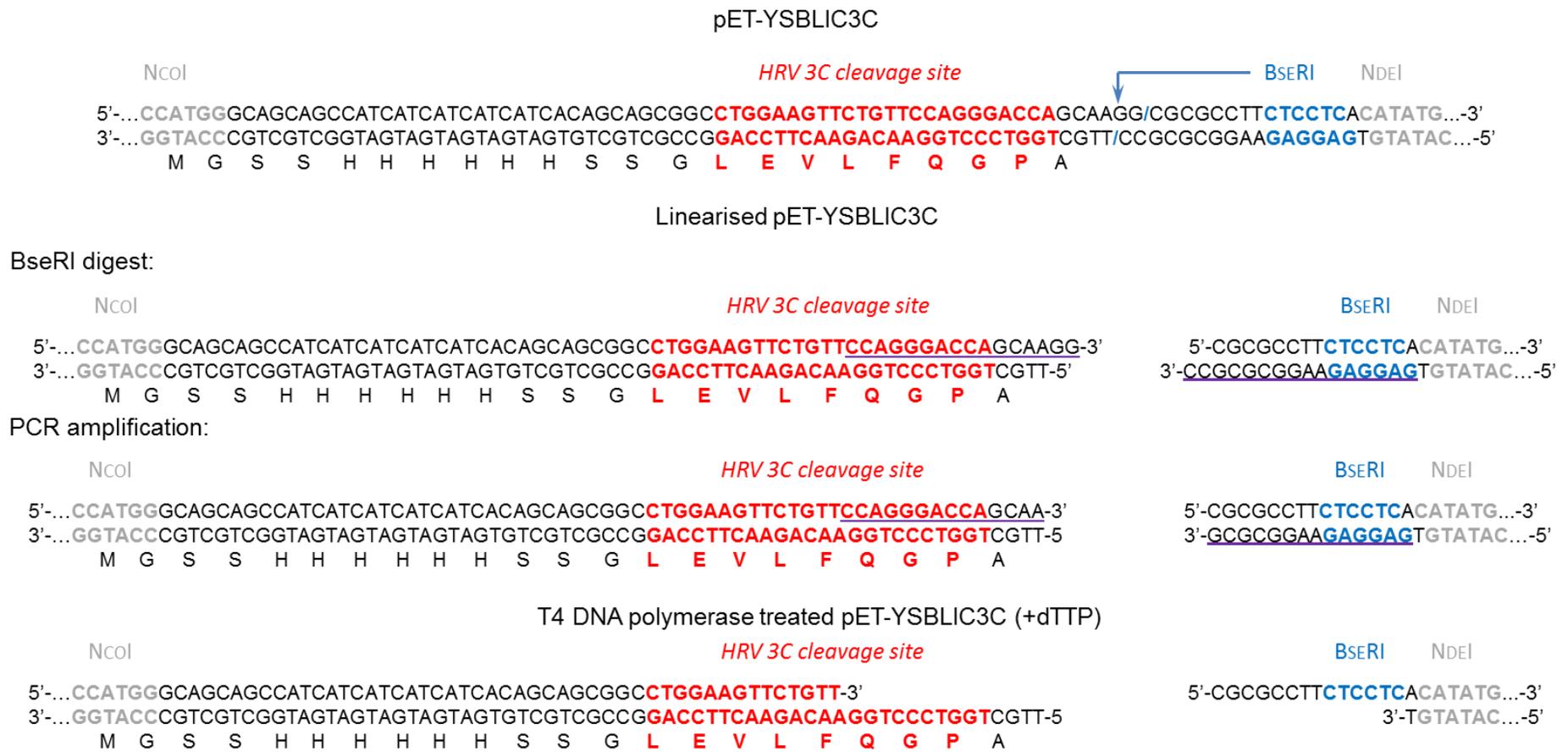
A large scale BseRI (New England Biosciences) digest of 50 µg of plasmid was carried out. The reaction mixture contained 50 µg pET-YSBLIC3C plasmid DNA, 50 U of BseRI restriction enzyme, 100 µl of 10x concentrated reaction buffer 2 (NEB) and milliQ H<sub>2</sub>O to a total volume of 1 ml. Reactions were incubated at 37°C for 110 minutes.

The reaction products were mixed with 200 µl 6x DNA loading dye and run on a 1% agarose gel for between 90 and 120 minutes at 110 V to achieve adequate separation of cut from uncut vector. The carefully excised band representing the linearised plasmid was then purified using Qiagen GenElute gel purification kit with elution of the final product in 30 µl of water.

#### **2.3.4.2 pET-YSBLIC3C Preparation by vector PCR**

The linearised vector can also be amplified by PCR. This method is favourable as a smaller quantity of template can be used and purification by agarose gel electrophoresis is not necessary. The reaction mixture and the thermocycling parameters are shown in Table 2-3.

After PCR the reaction products are analysed by agarose gel electrophoresis and treated with DpnI (discussed further in section 2.4.3) to digest the circular template DNA. This digest is carried out by direct addition of 2 µl of DpnI (NEB) and incubation at 37°C for 120 minutes. The use of DpnI reduces the risk of false positives but is not possible when plasmid linearisation is carried out by BseRI digest. After DpnI treatment the reaction products can be purified using a QIAquick PCR purification kit (QIAGEN).



**Figure 2-2: The pET- YSBLIC3C LIC cloning site.**

The cloning site of pET-YSBLIC3C is shown in three states; complete, linearised and after T4 DNA polymerase treatment. The BseRI restriction enzyme recognition sequence and cleavage site for plasmid linearisation is shown in blue. Beneath the DNA duplex sequences the translated open reading frame sequence is shown. The sequence encoding the Human Rhinovirus 3C protease cleavage site is shown in red. The regions to be truncated by T4 DNA polymerase are underlined in purple. NdeI and NcoI restriction sites are indicated in grey.

**Table 2-3: pET-YSBLIC3C linearisation by PCR**

<b>Reaction Mixture:</b>			
Component	Concentration	Volume ( $\mu$ l)	
dNTP mix	2 mM	5	
KOD DNA polymerase reaction buffer	10x concentrate	5	
MgSO <sub>4</sub>	25 mM	2	
pET-YSBLIC3C	>50 ng/ $\mu$ l	0.5	
Forward Primer	20 pmol/ $\mu$ l	1	
Reverse Primer	20 pmol/ $\mu$ l	1	
KOD DNA polymerase	1 U/ $\mu$ l*	1	
milliQ Deionised Water	-	34.5	

<b>Thermocycling parameters:</b>			
Step	Temperature ( $^{\circ}$ C)	Duration (s)	No. of Cycles
Initial Denature	94	120	1
Denature	94	30	30
Anneal	45	30	
Extension	72	120	
Final Extension	72	300	1
Hold	4	$\infty$	1

### 2.3.5 T4 DNA Polymerase Treatment

T4 DNA polymerase possesses proof-reading 3'-5' exonuclease activity. In the absence of dNTPs T4 DNA polymerase will remove 3'-terminal nucleotides processively from DNA fragments generating 5' overhanging ssDNA ends. In the presence of single dNTP the exonuclease cleavage reaction proceeds until the enzyme encounters a base on the overhanging strand where its polymerase can incorporate the added complementary nucleotide. This is exploited in LIC by the absence of the thymine bases on 3' termini of linearised pET-YSBLIC3C which can be digested in this fashion as shown in Figure 2-2. The insert is treated with T4 DNA polymerase in the presence of dATP while the linearised vector is treated with T4 DNA polymerase in the presence of dTTP. This leads to the generation of 5' overhangs on the insert which are complementary to those on the vector. This complementarity also ensures that the insert is incorporated into the plasmid in the correct orientation.

**Table 2-4: T4 DNA Polymerase Treatment Reactions for LIC**

Component	Insert Reaction	Vector Reaction
Linearised pET-YSBLIC3C	-	4 pmol
Purified LIC Insert	0.2 pmol	-
2.5U/ $\mu$ l T4 DNA Polymerase	0.4 $\mu$ l	8 $\mu$ l
25 mM dTTP	-	40 $\mu$ l
25 mM dATP	2 $\mu$ l	-
100 mM DTT	1 $\mu$ l	20 $\mu$ l
10x T4 DNA Polymerase Buffer	2 $\mu$ l	40 $\mu$ l
milliQ deionised water	To 20 $\mu$ l	To 400 $\mu$ l

A large scale T4 DNA polymerase (NEB) treatment was carried out on the vector before purification using a PCR purification kit. The prepared vector was then stored at -20°C and thawed for use in multiple cloning experiments.

Treatment of the LIC insert was carried out for each sample immediately before annealing. The reaction mixtures for both reactions are given in Table 2-4. Both reactions are incubated at 22°C for 30 minutes followed by thermal inactivation of T4 DNA polymerase at 75°C for 20 minutes.

### 2.3.6 Annealing and transformation

Annealing of the T4 DNA polymerase treated insert and vector components is carried out at room temperature for 20 minutes. During this time the ssDNA overhangs on the vector hybridise with those on the insert to form a nicked, yet stable, circular DNA complex. 2 µl of T4 DNA polymerase prepared vector at around 15 ng/µl is added to 1 µl of the above T4 DNA polymerase LIC insert reaction, now at 0.01 pmol/µl, and the mixture is incubated at room temperature. After 10 minutes, a further 1 µl of millQ deionised water is added before a further 10 minute incubation. The mixture is next used to transform competent cells of *E. coli*. A control reaction is set up with water in place of the insert DNA. This control demonstrates the background of false positives given by the vector DNA alone.

Transformation refers to the uptake of foreign genetic material by a cell. *E. coli* cells can be made competent through chemical treatment and commercially available competent cells can have very high transformation efficiencies. Competent XL-10 gold cells (Agilent) were prepared using the TSS competent cell generation protocol as described in section 2.3.6.1 [178].

1 µl of the annealing reaction (or control reaction) above is added to a 25 µl aliquot of XL-10 gold competent cells pre-incubated on ice. Incubation on ice is continued for a further 10 minutes. To enhance the transformation frequency the cells are heat-shocked for 30 seconds at 42°C followed by a further 5 minute incubation on ice. Competent cells are allowed to recover in 500 µl of a rich media, such as SOC, LB (Luria Bertani) or GS96 (QBiogene) at 37°C for 1 hour. This step allows the cells to recover metabolically and to express newly acquired antibiotic resistance genes, *kan<sup>R</sup>* in the case of pET-YSBLIC3C-based constructs.

After recovery, 100 µl of the culture is spread onto an LB-agar plate containing 30 µg/ml kanamycin. The plate is incubated overnight at 37°C. Colony-containing plates can be stored at 4°C for several days.

#### **2.3.6.1 Generation of TSS competent *E. coli* cells**

The TSS (Transformation and Storage Solution) competent cell protocol is a one-step method for generating competence and cryo-protection of *E. coli* cells. Commonly, competent cells are generated by the addition of CaCl<sub>2</sub> in a process requiring multiple resuspension steps in MgCl<sub>2</sub> and CaCl<sub>2</sub> containing buffers [172], however, for the TSS method transformation efficiencies of >10<sup>7</sup> colony forming units per µg of DNA can be achieved using a single step method [178]. This efficiency is only an order of magnitude lower than the efficiencies quoted for high transformation efficiency commercial competent cells.

0.2 ml of an overnight culture of the *E. coli* strain grown in 10 ml LB media at 37°C, is diluted into 20 ml fresh LB media and the culture is grown at 37°C for 2 hours (or 3 hours for XL strains i.e. XL-10 gold), followed by incubation on ice for 15 minutes. The culture is then centrifuged at 3500 x g for 10 minutes in a temperature controlled centrifuge at 4°C. The supernatant is discarded and the pellet resuspended in 950 µl ice cold TSS solution (10% PEG 3.35K and 50 mM MgCl<sub>2</sub> in LB media at pH 6.5) and 50 µl DMSO. The now competent cells are aliquoted in volumes of 25 µl and 100 µl and snap frozen in liquid N<sub>2</sub> for storage at -80°C.

#### **2.3.7 Assessing Cloning Success**

Colonies of bacteria which grow on the LB agar plates have been transformed to kanamycin resistance by uptake of the pET-YSBLIC3C vector DNA. Plasmid DNA is extracted from the colonies by miniprep as described in section 2.2. In many cases, transformant colonies may be false positives, having acquired a plasmid not containing the target insert. There are three main methods for analysing the success of a cloning experiment.

A plasmid PCR consists of a PCR analogous to that carried out earlier in the cloning process to amplify the insert. In this PCR, the template DNA is the plasmid DNA derived from the colonies. The primers used can be the insert specific primers from the original PCR or vector specific primers, which bracket the intended cloning site. The PCR reaction products are analysed by agarose gel electrophoresis as described in section 2.1. The observation of an appropriately sized fragment is evidence for cloning success.

Cloning success can also be analysed by restriction endonuclease digestion of product plasmids. Restriction endonucleases cleave dsDNA at specific recognition sequences. Computational analysis of the desired product DNA sequence using Serial Cloner (SerialBasics) generates a map of restriction enzyme cleavage sites and predicts the sizes of DNA fragments generated digestion by specific restriction endonucleases. The pET-YSBLIC3C plasmid has an NcoI restriction site upstream of the cloning site and an NdeI restriction site downstream of the cloning site (Figure 2-2) meaning that NcoI/NdeI co-digestion would generate a fragment of very similar length to the insert. An exception would be if the insert contains recognition sequences for these enzymes. 100 ng of plasmid DNA is mixed with 1 µl of each of NdeI and NcoI (NEB) and 2 µl of NEB buffer 4 (or new CutSmart buffer) in a total volume of 20 µl. The digestion reaction is incubated at 37°C for 30 minutes. The reaction products are analysed by agarose gel electrophoresis.

The most powerful technique for assessing clones is DNA sequencing, which when coupled with vector specific primers which prime outside the cloning site provides confirmation that the gene of interest lies within the plasmid. This technique is carried out by GATC biotech, a company allowing postal sample delivery and result delivery by e-mail. The sample requirements are 20 µl of plasmid DNA at 30-100 ng/µl (sufficient for up to 8 reactions). An array of common primers is available and the T7 and T7\_rev primers recognise the T7 promoter region and bracket the cloning sites of most pET expression system vectors, like pET-YSBLIC3C. The results give accurate DNA sequences as long as 1 kb if the template is abundant and pure.

## 2.4 Site directed mutagenesis

The QuikChange method of site-directed mutagenesis (SDM) is a powerful technique for the study of protein function and structure. The technique uses mutagenic oligonucleotide primers in which a small number of base mismatches, which direct the desired mutations, are flanked by regions matching the DNA template [179]. This technique can be used to change a single codon, multiple adjacent codons or it can be used to effect insertions and deletions. The QuikChange site directed mutagenesis kit (Agilent) was used to generate mutants in this work.

### 2.4.1 Oligonucleotide Mutagenesis Primers

Appropriate base exchanges are devised to implement the desired mutation. Mutagenic primers are designed in pairs consisting of 1-3 mismatched bases flanked by 10 to 15 bases complementary to the target template. According to the manufacturer's advice, the two primers should be 25-45 bases long, binding to complementary sequences on opposite strands of the template. It is also suggested that the  $T_m$  value for the primer pair should be greater than or equal to 78C°.

### 2.4.2 The Mutagenesis Reaction

The site directed mutagenesis reaction (Figure 2-3) is similar to a PCR, however, rather than amplification of a small portion of the template DNA the nascent DNA strand is extended around the entire plasmid stopping only when the 5' end of the initiating primer is reached (Table 2-5).

**Table 2-5: Reaction Parameters for Site Directed Mutagenesis**

<b>Reaction Mixture:</b>			
Component	Concentration	Volume ( $\mu$ l)	
dNTP mix	2 mM	5	
polymerase reaction buffer	10x concentrate	5	
Template Construct Plasmid	>50 ng/ $\mu$ l	1	
Forward Primer	125 ng/ $\mu$ l	1	
Reverse Primer	125 ng/ $\mu$ l	1	
<i>PfuTurbo</i> DNA polymerase	2.5 U/ $\mu$ l*	1	
milliQ Deionised Water	-	36	

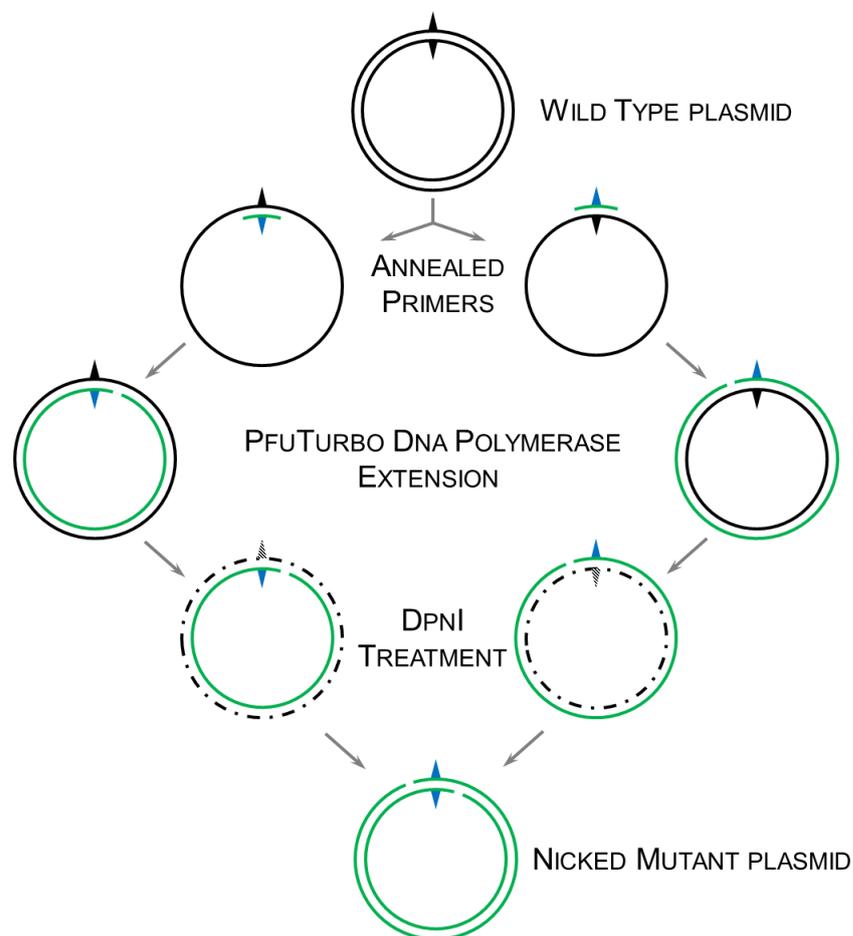
<b>Thermocycling parameters:</b>			
Step	Temperature ( $^{\circ}$ C)	Duration (s)	No. of Cycles
Initial Denature	95	30	1
Denature	95	30	
Anneal	55	60	12-18
Extension	68	60/kb*	
Hold	4	$\infty$	1

\* This parameter must be adjusted according to the size of the template plasmid.

The extension times are typically modified to allow extension round the entire template. *PfuTurbo* DNA polymerase used with the SDM kit extends at 1 kb/minute, allowing calculation of the appropriate extension time parameter. The number of reaction cycles is also modified according to the extent to which the mutant primers are mismatched to the template. Extended mismatching, as in primers used to introduce deletions or insertions of multiple codons, is more

effectively achieved using 18 cycles while a single base mismatch for a point mutation requires only 12 cycles. Thermocycling results in amplification of the synthesised mutant DNA strand, but the amount of product does not increase exponentially as the product strands cannot serve as a template in future polymerisation steps. Once the reaction is complete a 10 µl sample is analysed by agarose gel electrophoresis to confirm successful amplification of the mutant product.

### 2.4.3 DpnI Treatment



**Figure 2-3: Site Directed Mutagenesis procedure.**

The methylated wild type plasmid DNA isolated from a *dam*<sup>+</sup> *E. coli* strain is represented in black. The mutation site is indicated by triangles which are coloured black for unmutated and blue for mutated. DNA synthesised during the mutagenesis reaction is unmethylated and shown in green. Digestion of DNA by DpnI is indicated by the use of dashed lines.

A key step in the SDM procedure is the digestion of the original template DNA by DpnI. The transformation efficiency of nicked DNA is much lower than that of closed circular DNA and the presence of template DNA could give rise to a large number of wild type clones after transformation and growth. Isolation of the template plasmid from a DNA adenine methylation (*dam*<sup>+</sup>) *E. coli* strain, such as XL-10 gold, produces DNA in which the adenine base is methylated at the N6 position in each GATC base sequence. DpnI is a restriction enzyme which recognises and catalyses cleavage of a methylated GA(CH<sub>3</sub>)TC sequence but not an unmethylated GATC sequence. As the dATP molecules used in the reaction are unmethylated, the DNA synthesised during the *in vitro* reaction is also unmethylated. As a result DpnI treatment causes selective digestion of DNA strands synthesised *in vivo* by *E. coli*, leaving the newly synthesised DNA unaffected. 1 µl of DpnI at 10 U/µl is added to the SDM reaction products and the mixture is incubated for 1 hour at 37°C. 1 µl of the reaction products are then used to directly transform *E. coli* XL-10 gold competent cells, which are plated and grown on LB agar plates supplemented with kanamycin, as described in section 2.3.6. Assessment of the success of the mutagenesis reaction is achieved by chain terminating inhibitor-based DNA sequencing of plasmids isolated from single colonies.

## 2.5 Protein Expression

The pET derivative plasmid is introduced into an expression strain, commonly *E. coli* BL21 (DE3). This strain harbours the  $\lambda$ DE3 element, a phage lysogen which harbours the gene encoding T7 RNA polymerase under the control of a *lacUV5* promoter. T7 RNA polymerase drives transcription of coding sequences downstream of the pT7 promoter on pET-YSBLIC3C. Production of protein encoded by a pET construct is inducible by addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) to a growing bacterial culture.

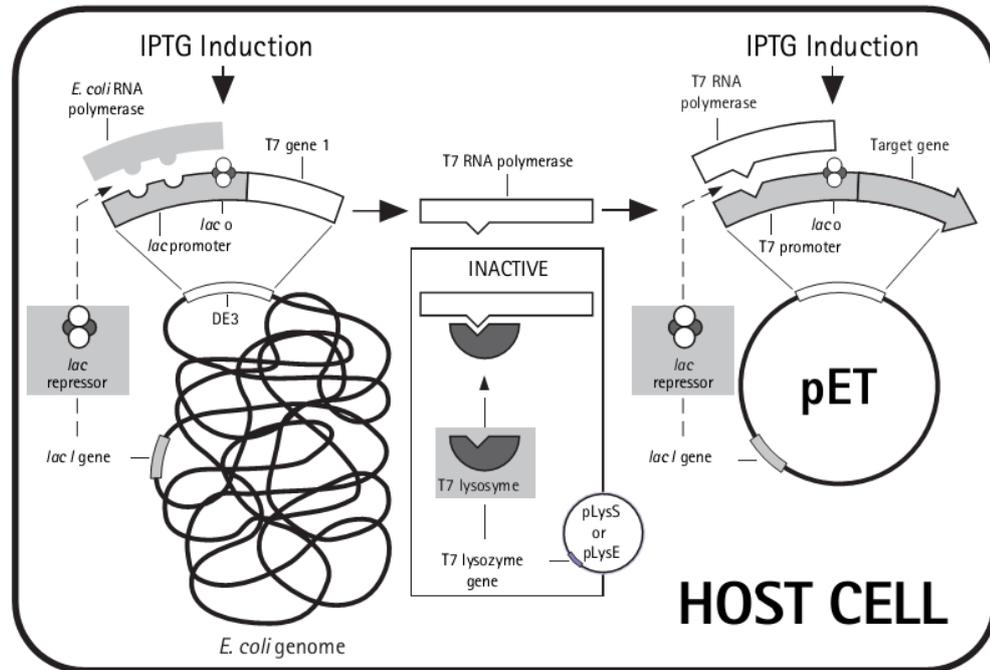
### 2.5.1 Culture Growth

The appropriate pET construct is used to transform competent *E. coli* BL21 (DE3) cells. The transformed cells are plated onto LB agar supplemented with kanamycin and grown overnight at 37°C. A single colony is used to inoculate 10 ml of LB liquid media supplemented with kanamycin at a final concentration of 30  $\mu$ g/ml; this culture is grown to high cell density overnight at 37°C with shaking at 180 rpm. This culture is diluted 100-fold into either 10 ml or 500 ml of fresh media for expression testing and large scale over expression, respectively. These cultures are similarly grown at 37°C with shaking at 180 rpm.

### 2.5.2 Induction of Protein Expression by IPTG

The pET expression system is based around the induced relief of transcriptional repression on a tightly controlled T7 RNA polymerase specific promoter (Figure 2-4). The T7 promoter is selectively recognised by T7 RNA polymerase encoded by bacteriophage T4 gene 1 carried by the DE3 element. The T7 RNA polymerase gene and the target gene on the pET vector are under lac operator control. Chromosomal and vector copies of the *lacI* gene ensure production of sufficient Lac repressor to repress transcription. The repression is overcome by addition of the non-hydrolysable allolactose mimic, IPTG. This leads to transcription of gene1 and production of T7 RNA polymerase which then

transcribes the target gene. As the pET T7 promoter is unique within the cell the T7 RNA polymerase produced is devoted to the expression of the plasmid-borne target gene. This generally leads to high level production of the recombinant protein which can constitute up to 50% of the total cell protein after a few hours of induction [180].



**Figure 2-4: The pET expression system.**

The host cell shown is harbouring a pET vector. During cell growth the pET vector T7 promoter and  $\lambda$ DE3 *lac* promoter are repressed by *lac* repressor at *lacO*. IPTG induction causes the relief of repression by *lac* repressor at both sites. T7 RNA polymerase is transcribed from the  $\lambda$ DE3 lysogen and its expression allows transcription of the target gene from the pET T7 promoter. The available pLysS and pLysE strains of BL21 (DE3) harbour an additional plasmid which encodes T7 lysosyme, an inhibitor of T7 RNA polymerase, which further prevents uninduced expression of the target gene. As presented in the Novagen pET System Manual, 2003.

Protein production was induced when the cell culture reached an  $OD_{600}$  of 0.6. 1 M IPTG was added directly to the growing culture to give a final concentration of 1 mM. On occasions, the final inducer concentration was varied in the range of 0.1 to 1.5 mM IPTG to control the rate of induction. After induction the culture was incubated at 16°C for 16 to 18 hours with shaking at 180 rpm. Protein production can also be carried out at up to 37°C for shorter periods down to 4 hours as the solubility and yield of overexpressed protein can vary greatly with the expression conditions.

### 2.5.3 Harvesting Cells and Lysis by Sonication

Cells from large cultures were harvested by centrifugation at 4,225 x g for 20 minutes in a Sorvall centrifuge. The pelleted cells were resuspended in 20 ml of lysis buffer by shaking and vortexing with periodic incubation on ice. A typical lysis buffer was 50 mM Tris-HCl, and 150 mM to 500 mM NaCl at a pH between 7.5 and 8.5. The resuspended cells were lysed by sonication at 15 Amplitude Microns for eight cycles of 30 seconds separated by 30 second intervals on ice. The contents of the disrupted cells, including proteins, membranes and DNA, become suspended in the lysis buffer. The insoluble elements are removed from the suspension by centrifugation at 35,000 x g for 25 minutes.

### 2.5.4 Optimisation of Expression and Lysis Conditions

The yield and solubility of recombinant proteins often varies with expression temperature, cell strain, expression media and lysis buffer. In order to determine the most appropriate expression conditions multiple 10 ml LB cultures with appropriate antibiotic were inoculated with 100 µl of a dense culture grown overnight from a single colony (section 2.5.1). These cultures are incubated at 37°C and their OD<sub>600</sub> monitored until it reaches 0.6. 1 ml of each sample was removed as an uninduced control. The remaining 9 ml of each culture was induced by addition of IPTG to a final concentration of 1 mM. Cultures were allowed to express protein with shaking at 180 rpm at different temperatures, most commonly only 16°C and 37°C for 16 and 4 hours, respectively. From each culture multiple 1 ml aliquots were taken and the cells harvested by centrifugation at 16,000 x g for 10 minutes in a Sigma™ microcentrifuge. Both uninduced and induced aliquots were resuspended in 150 µl of varying buffers and lysed by sonication for a total of 1 minute in 10 second bursts of 15 Amplitude Microns. A 10 µl aliquot was withdrawn to represent the total protein fraction (including both soluble and insoluble species). The remaining lysate was clarified by centrifugation in a microcentrifuge at 16,000 x g for 7 minutes after which a further 10 µl aliquot was taken to represent the soluble fraction. All aliquots were analysed by SDS-

PAGE (section 2.7.3.1) alongside similarly treated total and soluble fractions from the uninduced cell samples. Under ideal conditions there will be a roughly equal ratio of target protein in both total and soluble aliquots, indicating a negligible loss of protein due to insolubility. The insoluble fraction can also be analysed by resuspension of the lysate pellet in fresh buffer followed by a similar SDS-PAGE experiment.

## 2.6 Protein Purification

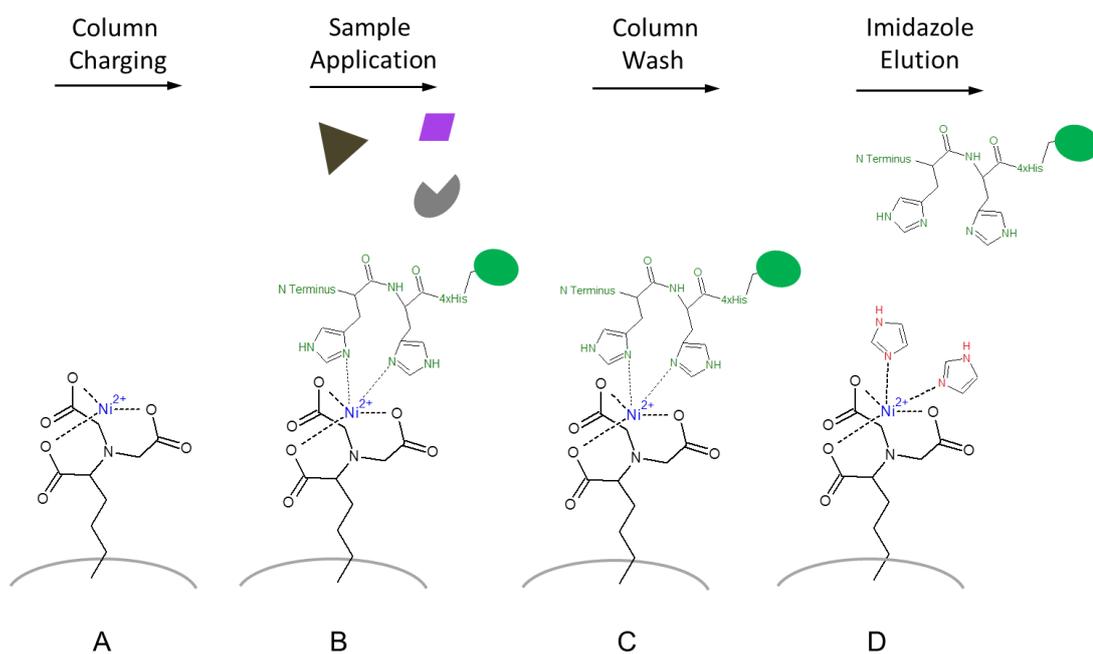
Structural and biophysical studies often require large amounts of pure protein. It is therefore necessary to resolve the target protein from the contaminating *E. coli* cell background. Separation techniques exist that separate proteins based on size, charge, hydrophobicity and the presence of engineered affinity tags. A two-step purification protocol was routinely used for the protein fragments presented in this work.

### 2.6.1 Immobilised Metal Affinity Chromatography

The pET-YSBLIC3C vector encodes an N-terminal 6xHis affinity tag which is fused to the target protein. This small peptide tag has a high affinity for divalent cations such as Ni<sup>2+</sup>. This affinity is conferred by the multiple imidazole moieties of the tandem histidine side chains [181]. Immobilised metal affinity chromatography (IMAC) utilises a resin which binds divalent cations via chelating groups, commonly nitrilotriacetic acid (NTA), packed into a sealed column to separate tagged target protein from untagged impurities. These HiTrap columns (GE Healthcare) are produced in volumes of 1 and 5 ml. The column resin is loaded with Ni<sup>2+</sup> by incubation with 100 mM NiSO<sub>4</sub>. When clarified cell lysate (Figure 2-5) is passed over the column, 6xHis tagged proteins are sequestered by the resin while impurities pass through without binding. After washing with multiple column volumes of buffer the bound protein(s) is eluted with an increasing concentration gradient of imidazole, which competes for the nickel with the protein histidine residues. The lysis buffer is supplemented with 20 mM imidazole which serves to reduce non-specific binding to the column [182] and the imidazole elution gradient increases to a final concentration of 500 mM.

The clarified cell lysate, generated in section 2.5.3, is loaded onto the column (pre-equilibrated in lysis buffer) at a flow rate of 2 ml/minute using a peristaltic pump. The material that does not bind to the column during loading is collected and is referred to as the flow-through. An 8 column volume washing step and a 16 column volume gradient elution are carried out at flow rate of 2 ml/minute

on an AKTA purifier system (GE Healthcare Life Sciences). The  $A_{280}$  and the  $A_{260}$  are continuously monitored to detect the presence of eluting protein and other absorbing species such as nucleic acids. The low imidazole concentration in the lysis buffer is used to wash the column until a stable baseline is observed, indicating that unbound proteins have been eluted from the column. A gradient elution of increasing imidazole concentration is achieved by pumping a mixture of low imidazole (20 mM) buffer with increasing proportions of high imidazole buffer (500 mM) through the column. As the imidazole concentration on the column increases, the bound histidine residues are displaced causing elution of the 6xHis tagged protein. Elution of the protein can be observed as a peak in the  $A_{280}$  trace. Fractionation is also carried out by the AKTA purifier and peak fractions are analysed by SDS-PAGE (section 2.7.3.1). Those fractions containing the target protein are pooled.



**Figure 2-5: Purification by Nickel IMAC.**

(A) A resin (grey arc) with attached chelating groups such as NTA (black) is charged with  $Ni^{2+}$ , a divalent cation (blue). (B) The mixture of proteins in the cell lysate is pumped through the column. 6xHis-tagged proteins (green) bind to the Ni-NTA and are immobilised, untagged proteins (brown, purple and grey shapes) remain in the mobile phase and are eluted in the flow through. (C) Unbound proteins are washed away thoroughly to ensure only sample bound to the stationary phase remains. (D) A linear imidazole concentration gradient is applied. As the imidazole concentration increases, Ni bound histidine residues are outcompeted and the tagged protein re-enters the mobile phase to be eluted.

## 2.6.2 HRV3C Protease Cleavage and Removal of 6xHis Affinity Tag

After IMAC, the 6xHis affinity tag is cleaved from the target protein by Human Rhinovirus 3C protease (HRV3C), a cysteine protease which recognises and cleaves the peptide sequence:



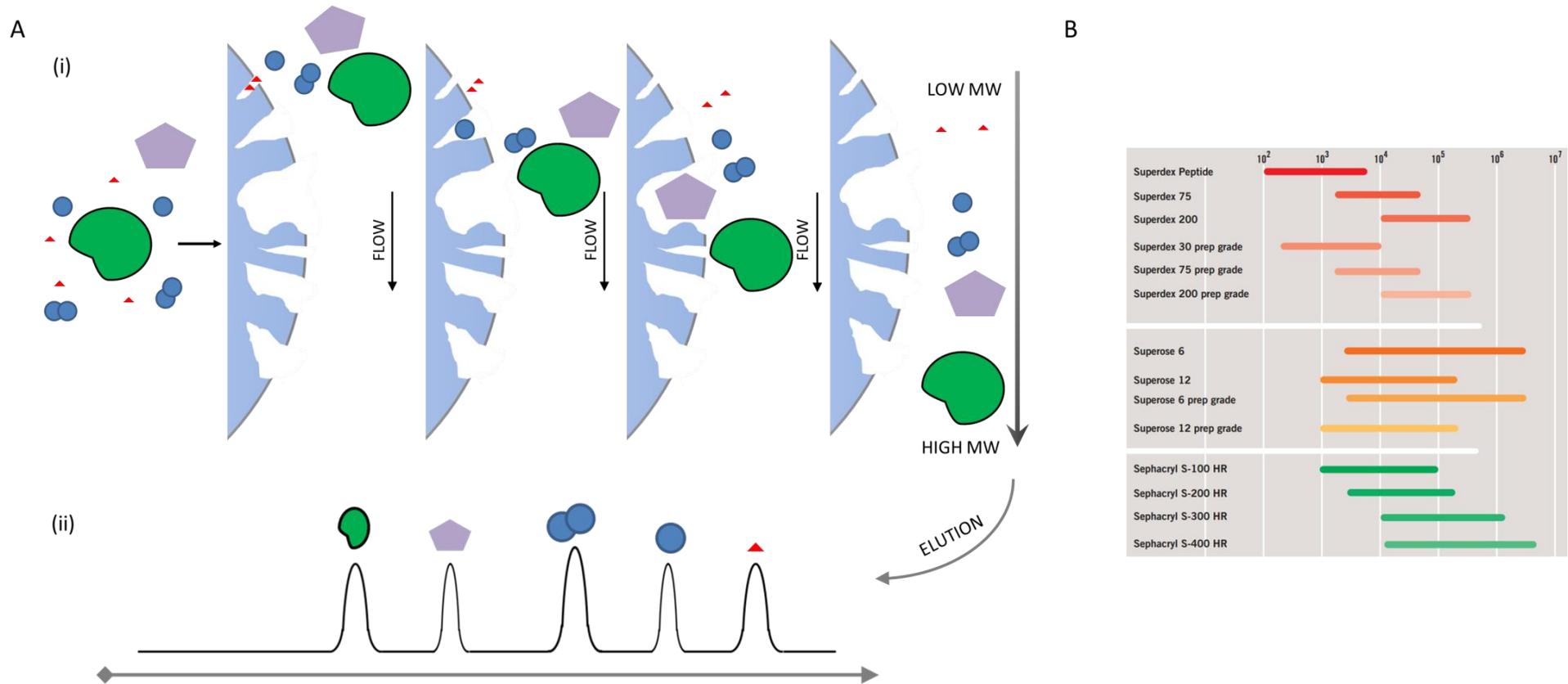
The coding sequence for this peptide in pET-YSBLIC3C (Figure 2-2) is flanked by 5' sequence encoding the 6xHis tag and 3' sequence encoding the recombinant protein. Thus HRV3C treatment of the recombinant fusion protein results in selective removal of the 6xHis tag, leaving the target protein with a 3 residue N-terminal addition of the sequence GPA. The HRV3C used in this work is produced with its own non-cleavable N-terminal 6xHis affinity tag. A second IMAC purification step removes any remaining 6xHis tagged protein including HRV3C, the 6xHis tags produced by HRV3C cleavage and any non-specifically binding contaminants.

Pooled peak fractions from the initial IMAC purification are mixed with a purified stock of 6xHis-HRV3C protease at 10 mg/ml. The HRV3C added is at a 1:50 mass ratio of enzyme to substrate. The cleavage reaction takes place during dialysis overnight against lysis buffer containing no imidazole. 10  $\mu$ l samples of the pooled fractions before and after HRV3C treatment are analysed by SDS-PAGE (section 2.7.3.1) to assess the progress of the cleavage reaction which is evidenced by a decrease in molecular weight of the post cleavage sample. The dialysate is then loaded onto a 5 ml HiTrap column equilibrated with lysis buffer on an AKTA purifier. Wash and imidazole gradient steps are then carried out as before. The cleaved target protein should elute in the flow through or early in the imidazole gradient, while contaminants and the 3C protease elute at higher concentrations of eluent. Fractions containing target protein are pooled.

### 2.6.3 Size Exclusion Chromatography

Size exclusion chromatography, also known as gel filtration, is a technique for separation of protein species based on size and shape (Figure 2-6). A size exclusion column is packed with cross-linked beads made from materials such as sepharose or superdex. The beads are porous each possessing multiple cavities of various sizes. As a mixture of proteins flows through the bead matrix or column bed, smaller proteins have lower mobility as they are retarded by frequent diffusion out of the mobile phase and into the pores. Meanwhile larger proteins are too bulky to enter many of the pores and thus achieve higher mobility. Therefore, elution occurs in the order from large species to small. The assorted column bed materials possess distinct pore size distributions and thus have capacities to resolve proteins in different size ranges [183]. Each column has a void volume, which is the elution volume of large materials that have not interacted with the pores. The remainder of the eluted bed volume of the column consists of species which have undergone resolution to some extent through pore interaction. An S200 HiLoad 16/60 column (GE healthcare), routinely used in this work, contains a superdex matrix optimised for separation of proteins between  $1 \times 10^4$  and around  $6 \times 10^5$  Da in a column bed 16 mm in diameter and 60 cm long. The total volume,  $V_t$ , is 120 ml, and the void volume,  $V_0$ , is 40 ml. To avoid solvent front effects on separation, gel filtration columns must be completely equilibrated in running buffer before the sample is applied. Diffusion effects reduce peak resolution, therefore, the sample should be injected onto the column in as small a volume as is practicable. In the case of 16/60 columns, the injection volume should be less than 3 ml.

Protein samples are concentrated to 2 ml and injected onto an S200 HiLoad 16/60 column pre-equilibrated in 0.22  $\mu\text{m}$  filtered and degassed running buffer at a flow rate of 1 ml/minute using an AKTA purifier. The  $A_{280}$  of the eluate is continuously monitored. 1.5 ml fractions are collected after 35 mls, and fractions with significant  $A_{280}$  readings are analysed using SDS-PAGE, section 2.7.3.1. Target protein containing fractions are pooled and concentrated for snap freezing and storage at  $-80^\circ\text{C}$ .



**Figure 2-6: The Principle of Size Exclusion Chromatography.**

(A) Separation of a mixture of differently sized protein species and multimeric states. (i) A mixture of proteins is resolved by retention of smaller species in pores of various sizes. The matrix beads (light blue) contain pores of various sizes. Larger species are less frequently retained. (ii) An example  $A_{280}$  trace for the demonstrated proteins, the largest species are eluted first with sequential elution of species of decreasing size. A monomer and dimer of the same protein (blue circles) can also be resolved. (B) Molecular Weight ranges for Gel Filtration Media. Resolution molecular weight ranges as presented by GE Healthcare [183].

## 2.6.4 Ion Exchange Chromatography

Ion exchange chromatography is a technique which can be used to separate molecules such as proteins based on their surface charge. The stationary phase is embedded with ionic functional groups which are positively or negatively charged in the case of anion and cation exchange, respectively. Molecules of an opposite charge will therefore bind to the column. The affinity of an analyte for the column matrix is determined by the density and distribution of its charged residues. Increasing ionic strength of the mobile phase, effected by a gradient increase in NaCl concentration, disrupts the electrostatic interactions between the protein and the resin and causes elution of the bound analytes in ascending order of affinity for the column.

The net charge on a protein is pH dependent. The isoelectric point, pI, of a protein represents the pH at which the protein has no net charge. Therefore, at pHs above the pI the protein is negatively charged and below the pI it is positively charged.

A 5 ml MonoQ (anion exchange) column (GE Healthcare) was equilibrated in 50 mM Tris-HCl buffer at pH 8.0. The protein sample was also dialysed into this salt-free buffer and loaded onto the column. An 8 column volume wash with salt-free buffer was used to elute unbound material before the column was developed with a 21 column volume 0-1 M NaCl linear gradient with 4 ml fractions being collected. The absorbance of eluate was continuously monitored at 280 nm and 254 nm. Peak fractions were analysed by SDS-PAGE, section 2.7.3.1.

## 2.7 Protein Analysis

### 2.7.1 Manipulation of Protein Concentration

Controlling protein concentration is important. Reduction of protein concentration simply requires dilution in a suitable buffer. Increases in protein concentration were achieved using an Amicon Ultra concentrator (Millipore). In these concentrators, a 15 ml sample reservoir is suspended in a 50 ml centrifuge tube. The base of the reservoir is constructed from semi permeable membrane panels. Concentrators are produced with various membrane permeabilities and are graded by Molecular Weight Cut-Off (MWCO), which describes the molecular weight above which a protein does not traverse the membrane. Centrifugation of the concentrator at 5,000 x g in a Harrier centrifuge forces the sample against the membrane which allows passage of water and small molecules but not species above the MWCO. The sample is retained in the sample reservoir, thus reducing the volume and increasing the concentration of the protein. Centrifugation is repeated at 10 minute intervals until the sample reaches the desired volume. The required time for concentration depends on the sample protein, buffer and concentrator MWCO. Between centrifugation steps it is important to check for protein precipitation and to resuspend any protein deposits that may build up at the membrane. The concentration of a 15 ml sample of protein can be increased 75-fold by concentration to 200  $\mu$ l in a short time without changing the buffer composition.

### 2.7.2 Determination of Protein Concentration

The UV absorbance of a protein solution at 280 nm can be used to estimate protein concentration. Each protein has a molar extinction coefficient ( $\epsilon$ ), which can be calculated based on the protein's amino acid sequence. Tryptophan, tyrosine and cystine residues contribute to  $\epsilon$  to different extents, their individual molar extinction coefficients are 5500, 1490 and 125  $M^{-1} cm^{-1}$ , respectively. The  $\epsilon$  for a protein fragment can be calculated using an online tool, Protparam (ExPASy) [184], which determines the relative concentrations of

each contributing residue in each mole of protein. The Beer-Lambert law (Equation 2-1) can be used to calculate protein concentration;  $c$ , based on absorbance at 280 nm;  $A$ , molar extinction coefficient;  $\epsilon$  and photocell pathlength;  $l$ .  $A_{280}$  readings are taken using an Eppendorf biophotometer, a buffer blank is also measured to provide a baseline reading.

$$A = \epsilon c l \quad \text{Equation 2-1}$$

### 2.7.3 Polyacrylamide Gel Electrophoresis

Polyacrylamide Gel Electrophoresis (PAGE) is a heavily used tool for analysis of proteins throughout expression optimisation, purification and subsequent studies. Polyacrylamide gels are produced by polymerisation of a mixture of acrylamide and bis-acrylamide initiated by ammonium persulphate and tetramethylethylenediamine (TEMED) which forms a matrix of pores of various sizes. When a potential difference is applied across the gel protein analytes undergo a size filtering effect as with agarose gel electrophoresis of DNA. The percentage concentration of acrylamide and bis-acrylamide in the polymerisation mixture can be adjusted to give higher or lower crosslinking density, thus changing the resolution properties of a polyacrylamide gel. Lower percentage gels allow resolution of larger macromolecules, while high percentage gels are able to separate smaller species more efficiently. Negatively charged proteins migrate towards the positively charged cathode at a migration velocity based on both the applied potential difference and on each protein's electrophoretic mobility. Gels and sample loading buffers are prepared using the components detailed in Table 2-6.

#### 2.7.3.1 Sodium Dodecyl Sulphate (SDS) PAGE

Polyacrylamide gel electrophoresis is most commonly used to analyse proteins in a denatured state where protein mobility is molecular weight dependent. The detergent, sodium dodecyl sulphate (SDS), possesses a 12 carbon hydrophobic chain that binds strongly to polypeptide chains. When protein is heat denatured

in the presence of SDS and a reducing agent,  $\beta$ -mercaptoethanol, a fully unfolded protein is uniformly coated with SDS. SDS has a negatively charged sulphate head group which masks the intrinsic charge of each protein and provides a constant mass to charge ratio amongst all analyte proteins. In SDS-PAGE, the gels, running buffer and sample loading dye all contain SDS.

Mini gel kits (Hoeffer) were used for this work to cast 1 mm thickness vertical polyacrylamide gels. A discontinuous gel system [185] is constructed by initial polymerisation of a resolving gel at pH 8.8, typically at a concentration of 10-17.5% acrylamide. Subsequently, a stacking gel is cast with 3% acrylamide at pH 6.8 and a comb to form sample loading wells. The low percentage stacking gel serves to concentrate samples into narrow bands before entry onto the resolving gel where the components are resolved, increasing the band resolution properties of the gel. Samples are prepared by mixing with 2 x SDS-PAGE Loading Dye and are boiled at 95°C for 5 minutes alongside a mixture of low molecular weight markers (Biorad), consisting of proteins of 14.4 kDa, 21.5 kDa, 31 kDa, 45 kDa, 66.2 kDa and 97.4 kDa. The marker and samples are loaded into the sample wells and a 200 V potential difference is applied for 52 minutes. On completion gels are stained with Coomassie Brilliant Blue, a colloidal dye that binds to protein but not to the gel. The gel is destained using a solution of propanol and acetic acid, leaving blue protein bands on the colourless and transparent gel.

### **2.7.3.2 Native PAGE**

Native PAGE is used to analyse proteins in their native state and the electrophoretic mobility of a sample depends on both charge and shape. Each protein has a pI (the pH at which it has no net charge), which is known as the isoelectric point. In the context of native PAGE the pI of a protein determines its direction of migration. At a pH above the pI of a protein there is a net negative charge and migration occurs in the direction of the cathode. Conversely, migration of a protein towards the anode occurs below its pI. This means successful analysis may require variation of experimental pH or inversion of the anode and cathode. The larger the net charge, the faster migration will occur,

however, the size of each protein serves as a retarding factor on migration as it is inhibited by the polyacrylamide matrix. As native proteins are much bulkier than their denatured counterparts, native gels are cast at lower percentage concentrations of acrylamide, namely 5-10%.

**Table 2-6: Components for Polyacrylamide Gel Electrophoresis**

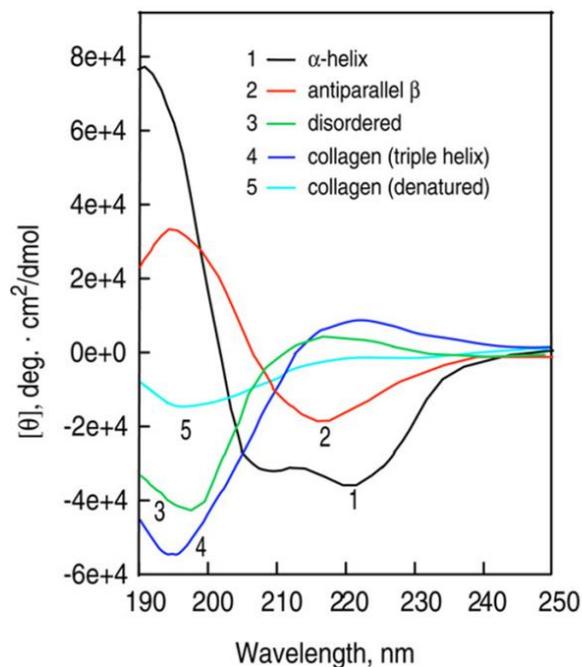
<b>Solution</b>	<b>Component</b>	<b>Concentration</b>
Resolving Gel	Tris-HCl pH 8.8	0.38 M
	Acrylamide Stock*	10 – 17.5 %
	Water	To 8.2 ml
	<b>SDS</b>	<b>0.1 %</b>
	Ammonium Persulphate	0.025 %
	TEMED	8 µl
<b>Stacking Gel</b>	Tris-HCl pH 6.8	0.13 M
	Acrylamide Stock*	3 %
	Water	To 5 ml
	SDS	0.1 %
	Ammonium Persulphate	0.025 %
	TEMED	8 µl
Loading Dye [x2]	Tris-HCl pH 6.8	30 mM
	Glycerol	5 %
	Bromophenol blue	0.01%
	<b>β-mercaptoethanol</b>	<b>5 %</b>
	<b>SDS</b>	<b>1 %</b>
Running Buffer	Tris-HCl pH 8.8	25 mM
	Glycine	200 mM
	<b>SDS</b>	<b>0.1 %</b>

Components in **bold red** text are not used when carrying out Native PAGE.

\*The acrylamide stock used contains 30% (w/v) Acrylamide and 0.8% (w/v) bis-Acrylamide.

A Native PAGE gel is cast with no stacking gel and its components contain no SDS. No protein markers are used but when analysing a mixture of species single pure proteins are used as referential controls. Native sample loading dye also contains no SDS and samples are not boiled, so that the proteins' native states are retained. The gel is run at 100 V at 4°C for 2 hours; the reduced temperature and voltage serves to prevent the temperature rising in the gel that would cause sample denaturation. Native gels are also visualised using Coomassie Brilliant Blue dye.

### 2.7.4 Circular Dichroism



**Figure 2-7: Example CD spectra of protein secondary structure.**

Spectra shown are recorded for poly-L-lysine in  $\alpha$ -helical (1), antiparallel  $\beta$ -sheet (2) and extended (disordered) (3) conformations and for collagen in native (triple helical)(4) and denatured (5) forms. Taken from by Greenfield, 2006 [186].

Circular Dichroism (CD) measures the difference in the absorption of right- and left-handed circularly polarised light. Asymmetry in protein molecules results in a differential absorbance of each polarisation of light by the peptide bond chromophore that absorbs in the far UV spectral region (190-250 nm). This difference is plotted as molar ellipticity,  $\theta$ , against the wavelength of the

incident light. Various secondary structure elements are defined by different alignments of contiguous peptide bonds relative to one another, which causes variation in both absorbed wavelength and molar ellipticity (Figure 2-7). Even a denatured protein exhibits a CD spectrum due to conformational preferences between adjacent residues.

CD spectra were measured with a Jasco J810 CD Spectrophotometer (JASCO Inc.). A xenon lamp and prism monochromator were used to select different wavelengths. To avoid interference from the short wavelength UV absorbance of O<sub>2</sub> the chamber was continually swept with N<sub>2</sub> during the experiment. Temperature control to avoid variations in protein folding was achieved using a Peltier temperature control unit. CD spectra were measured at a protein concentration of 0.2 mg/ml in 1 mm pathlength quartz cuvettes with a sample volume of 400 µl. Data were collected over the wavelengths range 180 - 260 nm. A buffer scan was subtracted from the protein spectra as a baseline correction.

### 2.7.5 SEC-MALLS

Size Exclusion Chromatography - Multi-Angle Laser Light Scattering (SEC-MALLS) is a technique which allows molecular weight (MW) estimation of the component species of a solution. Particles, such as proteins, which have a different polarisability from the solvent they are suspended in will scatter light from a beam passed through the solution. The amount of light scattered depends on concentration but is also proportional to molecular weight. Scattered light is measured by multiple detectors set at various angles to the incident light. Concentration is measured by differential refractive index. Linking the detection of both concentration and light scattering to gel filtration chromatography can be used to estimate molecular weight.

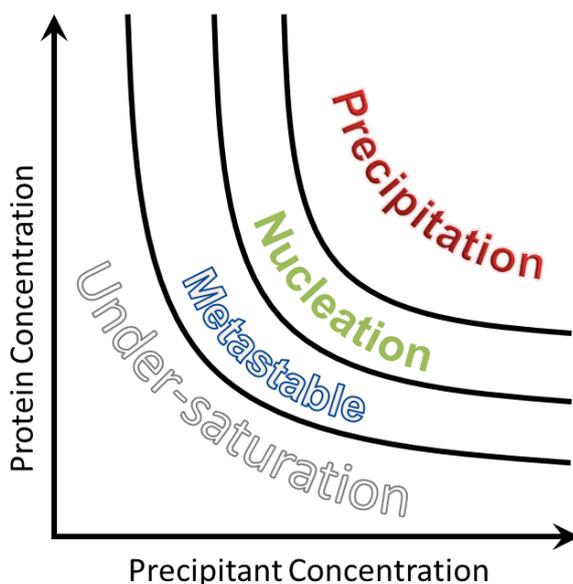
Experiments were conducted on a system comprising a Wyatt HELEOS-II multi-angle light scattering detector and a Wyatt rEX refractive index detector linked to a Shimadzu HPLC system (SPD-20A UV detector, LC20-AD isocratic pump system, DGU-20A3 degasser and SIL-20A autosampler). Work was conducted at

room temperature ( $20 \pm 2^\circ\text{C}$ ). The column used was a Superdex S200 10/300 GL column (G.E. Healthcare) unless otherwise stated. The column was equilibrated with solvent before use and flow was continued at the working flow rate of 0.5 ml/min until baselines for UV, light scattering and refractive index detectors were all stable.

The sample injection volume was 100  $\mu\text{l}$ . Shimadzu LC Solutions software was used to control the HPLC and Astra V software for the HELEOS-II and rEX detectors. The Astra data collection was 1 minute shorter than the LC solutions run to maintain synchronisation. Blank buffer injections were used as appropriate to check for carry-over between sample runs. Data were analysed using the Astra V software. MWs were estimated using the Zimm fit method with degree 1. A value of 0.19 was used for protein refractive index increment ( $\text{dn}/\text{dc}$ ).

## 2.8 Protein Crystallisation

The formation of well-ordered protein crystals can occur when an aqueous solution transitions from under-saturation to super-saturation due to a change in protein or precipitant concentration, pH or temperature. Several states of saturation define the process of crystal formation and are dependent on relative protein and precipitant concentration (Figure 2-8). In the precipitation zone protein is not stable in solution and precipitates in an amorphous mass. The nucleation zone, existing at slightly lower super-saturation levels, results in a slower precipitation that can give rise to small ordered crystal nuclei. The metastable zone exists at low levels of super-saturation. In this zone protein is unable to form nuclei but can extend crystal nuclei, causing crystal growth.

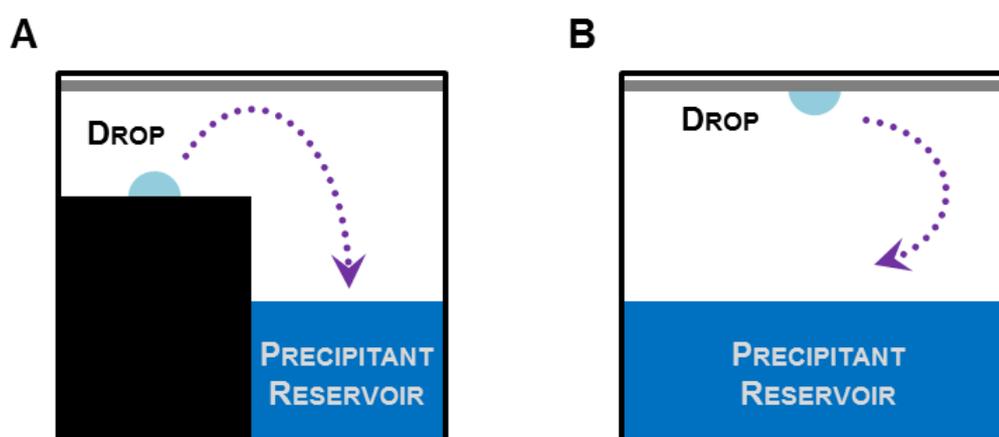


**Figure 2-8: The crystallisation phase diagram.**

The diagram indicates the various levels of super-saturation. Crystal nuclei are formed in the nucleation zone and crystal growth occurs in the metastable zone. Amorphous precipitate is formed in the precipitation zone.

In a successful crystallisation experiment the precipitant causes the protein solution to enter the nucleation zone where a small number of nuclei will form, leading to a reduction in the protein concentration. This concentration reduction in favourable cases will cause the solution to enter the metastable zone at which point the nuclei will slowly grow into large and well-ordered crystals suitable for X-ray crystallographic structure determination.

The crystallisation experiments in this work are based on vapour diffusion in either sitting or hanging drops (Figure 2-9). Both experiments consist of a sealed well containing a reservoir of precipitant solution and a separate drop of much smaller volume of protein mixed in a 1:1 ratio with the precipitant solution. An osmotic gradient is formed between the low precipitant concentration drop and the high precipitant concentration reservoir. As a result water evaporates from the drop increasing both precipitant and protein concentration.



**Figure 2-9: Sitting and hanging drop vapour diffusion experiments.**

Schematic diagrams of sitting drop (A) and hanging drop (B) vapour diffusion experiments. Purple arrows indicate diffusion of water from the low precipitant concentration drop towards the high precipitant concentration reservoir.

### 2.8.1 Crystallisation Screening

In order to find conditions in which protein crystals are formed it is necessary to test a variety of types and concentrations of precipitants, pHs and additives usually requiring 10s of milligrams of protein. Broad ranges of conditions can be sampled using commercially available screens. 96-condition screens commonly aim to either cover a wide range of precipitants, such as the Index and Hampton I & II screens (Hampton Research) or focus on a particular chemical subset like PEGs in the PACT premier screen (Molecular Dimensions).

The solutions from the chosen screen are transferred to an MRC-Wilden 96-well crystallisation plate in 54  $\mu$ l aliquots using a Hydra-96 pipetting robot (Robbins scientific). A 300 nl sitting drop experiment is aliquoted using a Mosquito

nanolitre pipetting robot (TTP LabTech). 150 nl each of precipitant solution and 150 nl protein solution are pipetted into raised wells. The plate is sealed using an optical grade self-adhesive plastic lid. Plates are stored at 20°C and checked for crystal formation at 2 or 3 day intervals.

### 2.8.2 Crystallisation Optimisation

Conditions which successfully produce protein crystals, or 'hits', are optimised to increase the size, quality and reproducibility of the crystals. The optimisations in this work were carried out in a 24-well plate (Falcon) with hanging drops of 2 µl over a 1 ml precipitant solution reservoir. The precipitant solutions in the optimisation plate generally mimic the initial hit with modification of one or two variables, such as precipitant concentration and pH. Each well can accommodate up to three drops which might be used to optimise protein concentration. Drops consisting of a 1:1 ratio of protein and precipitant solutions are placed onto the surface of siliconised glass cover slips, prepared with Aquasil (Pierce). The cover-slips are sealed over the wells using vacuum grease. Plates are incubated at 20°C.

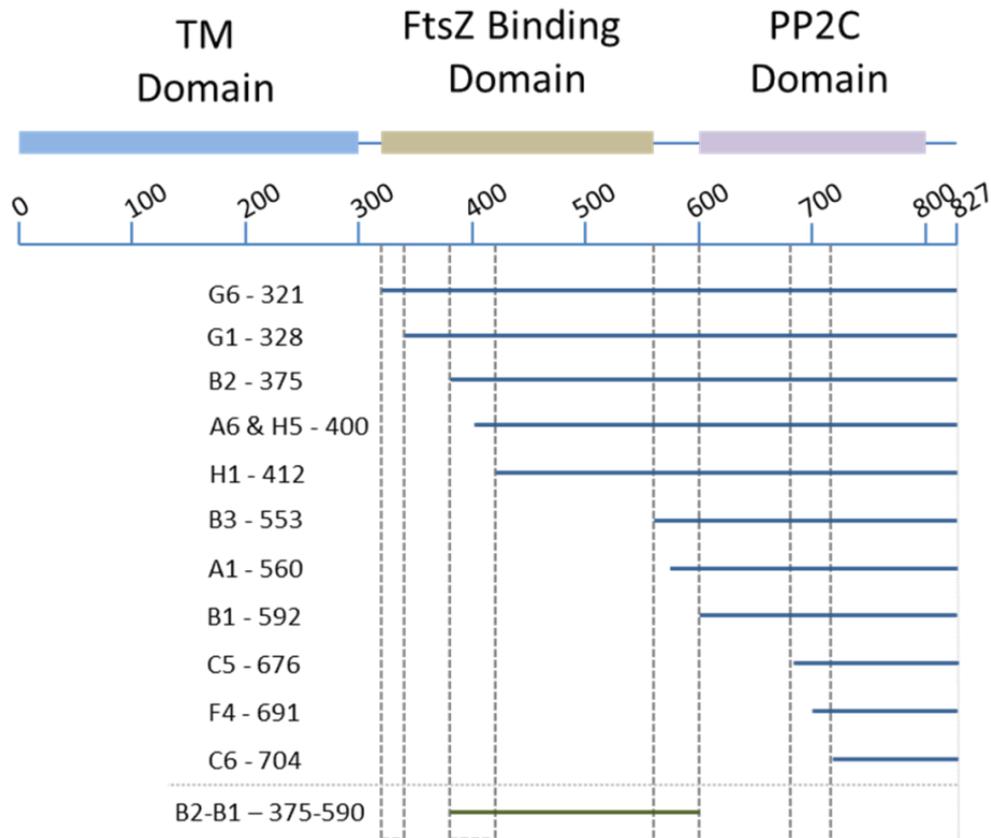
## Chapter 3: Investigations into SpoIIE fragments containing elements of the FtsZ binding domain.

### 3.1 Introduction

The multi-domain protein, SpoIIE, possesses two putative cytoplasmic domains in addition to a multiple helix membrane spanning domain. The C-terminal cytoplasmic domain is easily identifiable as a PP2C $\alpha$  protein phosphatase domain. However, the remaining 290 residues show little sequence homology to proteins of known structure or identified function beyond SpoIIE orthologues. This central domain, termed the FtsZ binding domain, is a target of great interest for how SpoIIE's phosphatase activity is regulated. Structures have been obtained by X-ray crystallography of every element of the sigF regulation system excepting the transmembrane and FtsZ binding domains of SpoIIE. This chapter describes the cloning, isolation and characterisation of a number of fragments of SpoIIE encompassing the FtsZ binding domain. The goal of these studies is the elucidation of the structure of the FtsZ binding domain so that its interactions with the PP2C $\alpha$  phosphatase domain may be understood.

#### 3.1.1 Previous approaches to soluble SpoIIE

An enduring obstacle to structural studies of SpoIIE has been obtaining enough soluble protein to support a programme of protein crystallisation experiments. Elucidation of the domain boundaries within a protein chain is often an important step on the pathway to the expression of soluble fragments of a larger recalcitrant protein. Domain boundaries can sometimes be inferred from a bioinformatics analysis by searching for segments of disorder, which may constitute inter-domain loops. Previously, analysis of SpoIIE using the RONN disorder predictor, carried out in this lab, yielded 16 possible targets. From these the SpoIIE<sub>590-810</sub> fragment, encompassing the PP2C $\alpha$  phosphatase domain, proved to be soluble. This construct was modified to include the 17 truncated C-terminal residues, was successfully crystallised and the structure solved as a domain-swapped dimer, discussed further in Chapter 4 [170].



**Figure 3-1: Soluble SpoIIE constructs identified by ESPRIT.**

A truncation library based solubility screen, ESPRIT (Expression of Soluble Protein by Random Incremental Truncation) [187], was carried out using a cloned *spoIIE* fragment encompassing base pairs 901 to 2481 in order to determine solubility windows across the cytosolic SpoIIE domains [188]. After screening and following up hits by recloning and DNA sequencing, the 12 most soluble fragments spanning the entire region were studied further (Figure 3-1). While these fragments had shown solubility in the library screen, further study involving larger scale expression demonstrated that only two of the 6 fragments which spanned the FtsZ domain could be overexpressed in *E. coli* in a soluble form; SpoIIE<sub>375-827</sub> and SpoIIE<sub>412-827</sub>, named B2 and H1 respectively according to their ESPRIT screen designations. These fragments showed a tendency to oligomerise and when examined by Size Exclusion Chromatography formed oligomers of molecular mass >600 kDa as well as numerous smaller species. Neither the oligomeric fractions of B2, nor the isolatable dimer fractions of H1

yielded any protein crystals in crystallisation screens (Andrea Rawlings, unpublished).

Six constructs, derived by taking the N-termini of ESPRIT hits and coupling them with the N-terminal boundary of the PP2C $\alpha$  domain at residue 590, were tested for expression and solubility. Of these only one showed adequate solubility for study. This construct comprised SpoIIE residues 375-590 and was termed B2-B1 according to the ESPRIT hits it best represented.

## 3.2 Results

### 3.2.1 Isolation and Analysis of the B2-B1 fragment: SpoIIE<sub>375-590</sub>

The B2-B1 (SpoIIE<sub>375-590</sub>) fragment had previously been identified in high throughput screening as soluble following over-expression in the *E. coli* BL21 pLysS expression strain. The construct consisted of the kanamycin resistant pET-YSBLIC3C plasmid (pYSBLIC3c\_SpoIIE<sub>375-590</sub>) bearing nucleotides 1126-1770 of *spoIIE* from the *B. subtilis* coding sequence downstream of sequence encoding an N-terminal 6xHis affinity tag and a Human Rhinovirus 3C protease cleavage site. The encoded protein has a molecular mass of 28.6 kDa with a calculated isoelectric point (pI) of 5.52. The solubility screen was carried out in high throughput lysis buffer of 20 mM Tris-HCl, 150 mM NaCl at pH 7.4. Expression tests were carried out to discover optimum solubility conditions for the SpoIIE<sub>375-590</sub> fragment.

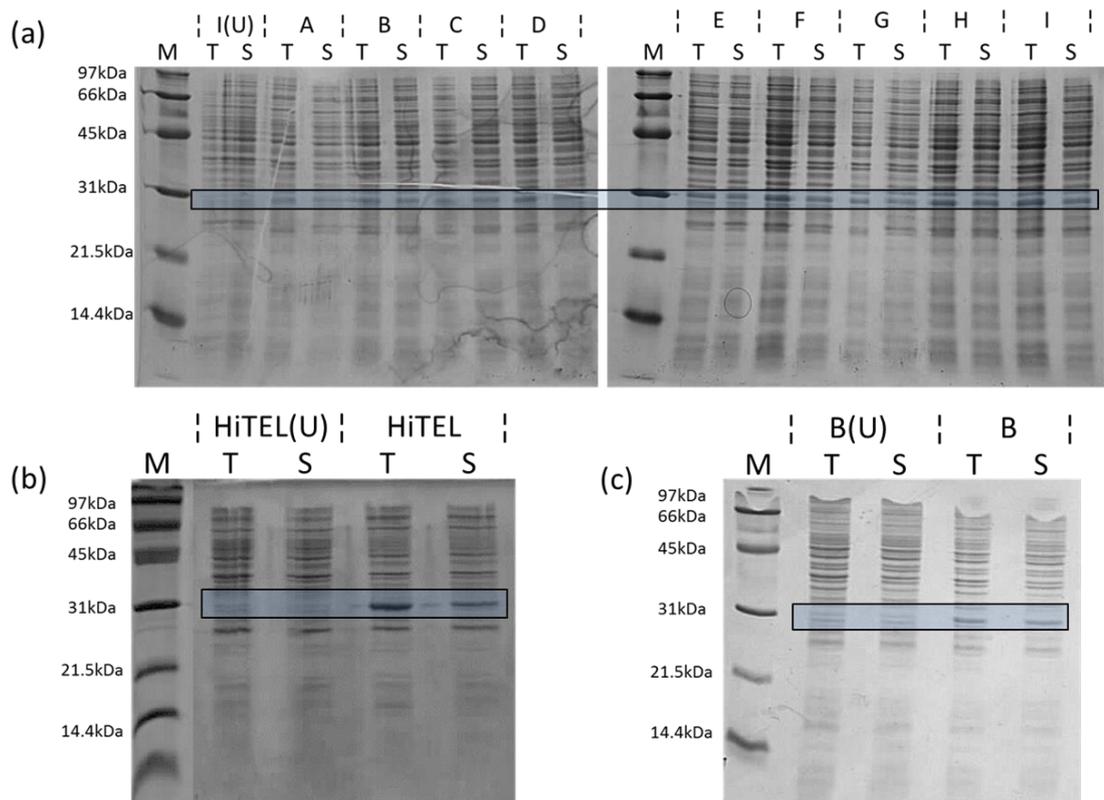
**Table 3-1: Buffers used in solubility screening of SpoIIE<sub>375-590</sub>.**

Buffer	Insert Reaction	pH
A	50 mM Tris-HCl, 500 mM NaCl, 30 mM Imidazole	8.0
B	20 mM NaH <sub>2</sub> PO <sub>4</sub> , 150 mM KCl, 30 mM Imidazole	7.6
C	60 mM HEPES, 150 mM KCl	7.5
D	100 mM BTP, 150 mM KCl	7.5
E	20 mM K <sub>2</sub> HPO <sub>4</sub> , 150 mM KCl,	7.5
F	20 mM Na <sub>2</sub> HPO <sub>4</sub> , 150 mM KCl, 20 mM Imidazole	7.6
G	20 mM Na <sub>2</sub> HPO <sub>4</sub> mM, 500 mM KCl, 20 mM Imidazole	7.6
H	50 mM Tris-HCl, 150 mM KCl	8.0
I	20 mM Tris-HCl, 150 mM NaCl	7.4

*E. coli* BL21 pLysS cells harbouring the pYSBLIC3c\_SpoIIE<sub>375-590</sub> plasmid were cultured in 50 ml falcon tubes containing 10 mls of LB with 30 µg/ml

kanamycin. Expression was carried out at 16°C overnight after induction at an optical density of 600 nm ( $OD_{600}$ ) of 0.6 with IPTG to a final concentration of 1 mM as described in section 2.5.2. 1 ml aliquots of cells were spun down for 1 min at 16,000 x g in a bench top microcentrifuge. Cell aliquots were then resuspended in an array of buffers (Table 3-1) before lysis by sonication. A 'total' protein fraction and a 'soluble' protein fraction, isolated by centrifugation of the lysate, were analysed by SDS polyacrylamide gel electrophoresis (SDS-PAGE) on a 15% polyacrylamide gel.

The expression tests indicated that lysis in all the buffers tested led to production of soluble SpoII $E_{375-590}$  (Figure 3-2). It was decided that buffer B (20 mM  $NaH_2PO_4$ , 150 mM KCl at pH 7.6) showed the most equal ratio of soluble to total SpoII $E_{375-590}$  and shows better solubility than the HiTEL (**H**igh **T**hroughput **E**xpression **L**aboratory) buffer used initially.

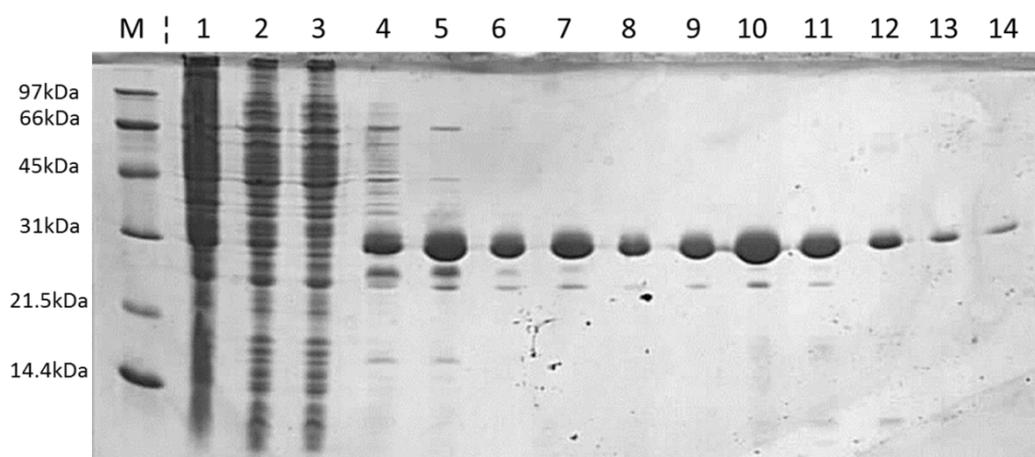


**Figure 3-2: SDS-PAGE gel images showing total and soluble cell extracts from *E. coli* BL21 pLysS cultures harbouring the pYSBLIC3c\_SpoII $E_{375-590}$  plasmid.**

Blue shaped boxes indicate the band associated with SpoII $E_{375-590}$  (a) Cells lysed in the buffers shown in Table 3-1. (U) indicates an uninduced sample. T and S indicate total and soluble cell samples, respectively. The letters A to I indicate the buffers used which can be referenced in Table 3-1. M indicates Biorad Low Molecular weight markers. (b) Cells lysed in the original screening buffer. (c) Cells lysed in the optimised buffer – B.

The SpoIIE<sub>375-590</sub> fragment expression tests were scaled up in order to produce the quantities of protein required for crystallisation experiments. Four 2 l flasks containing 500 ml of LB media were inoculated with *E. coli* BL21(DE3) pLysS cells harbouring the pYSBLIC3c\_SpoIIE<sub>375-590</sub> plasmid and cultured at 37°C as described in section 2.5.1. Production of SpoIIE<sub>375-590</sub> was induced with 1 mM IPTG when the OD<sub>600</sub> reached 0.6 with cultures subsequently being incubated at 16°C overnight.

Cultured cells were harvested by centrifugation for 15 minutes at 4,255 x g before resuspension in lysis buffer B with the addition of 30 mM imidazole, henceforth Buffer A. This modification is to prevent non-specific binding of proteins to a Nickel Immobilised Metal Affinity Chromatography (IMAC) column, described in section 2.6.1. Resuspended cells were then lysed by sonication as described in section 2.5.3. Insoluble factors in the lysate were cleared by centrifugation at 30,600 x g for 30 minutes.

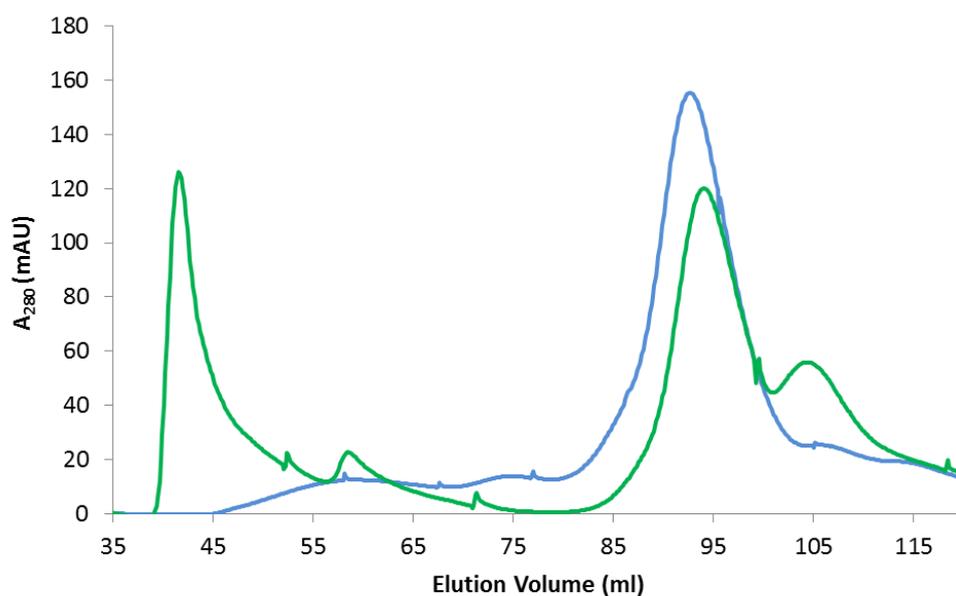


**Figure 3-3: SDS-PAGE gel image showing IMAC purification of the SpoIIE<sub>375-590</sub>.**

M indicates Bio-rad broad range Molecular weight markers. Lane 1; Total Cell extract, lane 2; Soluble cell extract [Load], lane 3; unbound flow through, lanes 4-8; fractions from Peak A eluting at around 60 mM imidazole, lanes 9-14; fractions from Peak B eluting at around 120 mM imidazole.

The clarified lysate was applied to a 5 ml HisTrap nickel-charged column (GE Biosciences) and washed with Buffer A until to remove unbound proteins. A linear gradient 30-500 mM imidazole in buffer A was then used to develop the column. The A<sub>280</sub> chromatogram revealed protein eluting in two peaks, a small peak (Peak A) at around 60 mM imidazole with a second peak (Peak B) at an

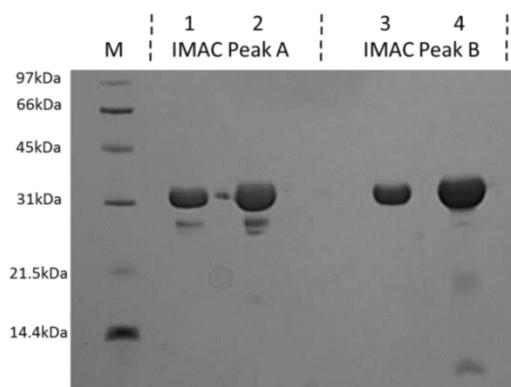
imidazole concentration of around 120 mM. When analysed by SDS-PAGE, the peak fractions were each shown to contain a strong band at around the expected mass of 27.6 kDa (Figure 3-3). The peak fractions were separately pooled and concentrated to a volume of < 4 mls for further purification by size exclusion chromatography.



**Figure 3-4: Size Exclusion purification of SpoIIE<sub>375-590</sub> peaks A (blue trace) and B (green trace).**

Concentrated fractions were loaded onto a Sephacryl S-400 16/60 gel filtration column (GE Biosciences) in two separate size exclusion chromatography experiments (Figure 3-4), as described in section 2.6.3. The running buffer used in this step was 20 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM KCl at pH 7.6. Peak A resolved into five species; two small broad peaks with retention volumes of 59 and 75 mls, one major peak at 90 mls and finally two shoulder peaks at 108 and 115 mls. SDS-page analysis of these peaks indicated that only the major peak at 90 mls contained an appreciable concentration of SpoIIE<sub>375-590</sub>. In contrast, B resolved into four species; one large peak in the column's void volume at 40 mls, a small peak at 59 mls, a major peak at 90 mls and a moderately sized shoulder peak at 105 mls. SDS-PAGE analysis showed that all of these peaks contained the SpoIIE<sub>375-590</sub> fragment in a largely pure form, with the exception of the shoulder peak (105 mls) which had a number of contaminants including a major band at

around 25 kDa. The major peaks, eluting at 90 mls, from each gel filtration experiment, were isolated and concentrated to 13.6 and 16.9 mg/ml, respectively. The concentrated protein stocks were analysed by SDS-PAGE, both showing minor contaminants at lower molecular masses (Figure 3-5).



**Figure 3-5: SDS-PAGE gel image of purified aliquots of SpoIIE<sub>375-590</sub>.**

M represents Biorad LMW protein markers. Lanes 1 and 2; post gel filtration aliquots of protein from Peak A at one tenth and one fifth concentration respectively. Lanes 3 and 4; post gel filtration aliquots of protein from Peak B at one tenth and one fifth concentration, respectively.

Both of these samples of SpoIIE<sub>375-590</sub> were put into crystallisation screens, as described in section 2.8.1. The proteins were screened in tandem using an MRC Wilden 96-well plate with two sample wells per condition against the Hampton I/II (Hampton research), PACT (molecular dimensions), Index (Hampton research) and CSS I/II (molecular dimensions) screens. The CSS I/II screen uses 48 conditions modified by two base buffers giving 96 conditions in total, the buffers used were; 1 M Tris-HCl at pH 8 and 1 M Tris-HCl at pH 5.6. Analysis of the screens showed a number of seemingly promising conditions. However, when tested for X-Ray diffraction, the crystals proved to be salt. Optimisation of promising conditions varying precipitant and buffer concentrations and pH were unable to reproduce the results shown in the original screens.

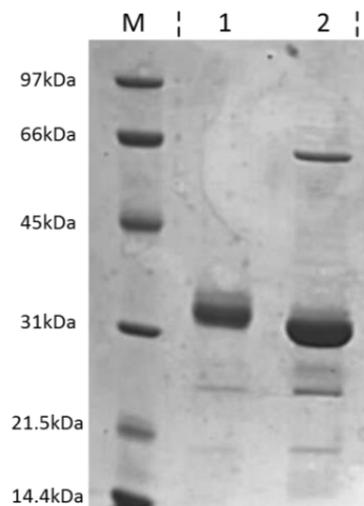
The formation of salt crystals in protein crystallisation screens of SpoIIE<sub>375-590</sub> prompted the removal of a NaH<sub>2</sub>PO<sub>4</sub> based buffer from the later stages of purification. Further purification of the fragment following buffer exchange into 50 mM Tris-HCl, 150 mM NaCl at pH 8.0, a buffer similar to that previously used successfully in the purification of the SpoIIE<sub>590-827</sub> fragment representing the PP2C phosphatase domain.

Nickel IMAC purification of cell lysate from 2 l of *E. coli* BL21 pLysS cells induced to produce SpoIIE<sub>375-590</sub> was carried out as above. SDS-PAGE analysis showed a large portion of SpoIIE<sub>375-590</sub> eluting in a peak at around 120 mM imidazole. The peak fractions were isolated and buffer exchanged, using a 12,000-14,000 MW cut off membrane, into the above buffer with the addition of 20 mM imidazole. A 1:50 ratio of Human Rhinovirus 3C (HRV 3C) protease with a 6xHis affinity tag was added to the dialysis in order to effect cleavage of the His<sub>6</sub> tag from SpoIIE<sub>375-590</sub>, as described in section 2.6.2.

The IMAC purified SpoIIE<sub>375-590</sub> bears an eight residue protease recognition sequence sandwiched between the N-terminal 6xHis tag and the beginning of the SpoIIE sequence:

→ MGSSHHHHHSSGLEVLFQ**GPA**-SpoIIE

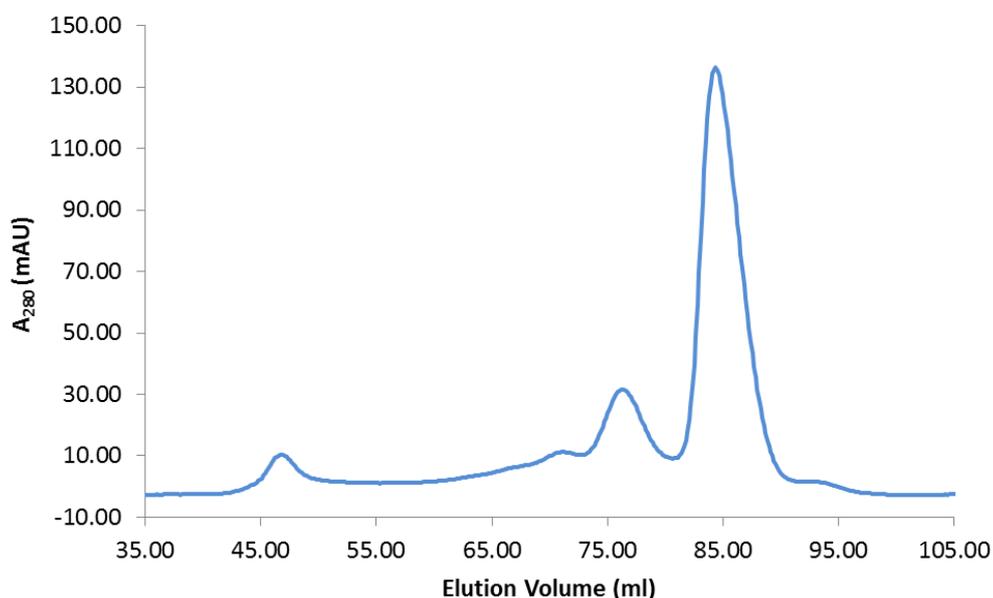
The underlined sequence is the recognition site for HRV 3C protease. It uses a catalytic cysteine residue to cleave to substrate polypeptide between the glutamine and glycine residues of the recognition site. The remaining SpoIIE fragment is represented in bold with a glycine, proline, alanine sequence of residues remaining at the N-terminus.



**Figure 3-6: SDS-PAGE gel image of HRV 3C cleavage of the 6xHis affinity tag of SpoIIE<sub>375-590</sub>.**

M represents Biorad LMW protein markers. Lane 1 shows the SpoIIE<sub>375-590</sub> fragment before addition of HRV 3C protease, lane 2 shows SpoIIE<sub>375-590</sub> after buffer exchange and HRV 3C protease cleavage. The cleaved SpoIIE<sub>375-590</sub> fragment shows a lower molecular mass, the HRV 3C protease is also visible as a band at around 60 kDa.

The success of the cleavage of 6xHis-SpoIIE was analysed by SDS-PAGE, showing cleavage had occurred efficiently (Figure 3-6). The cleaved SpoIIE<sub>375-590</sub> was applied to a 5 ml Nickel charged HisTrap column and the column washed before a linear gradient to a concentration 500 mM imidazole was applied. The A<sub>280</sub> chromatogram showed a large peak in the low imidazole wash with three overlapping peaks eluting from the column at imidazole concentrations between 60 and 150 mM. Analysis by SDS-PAGE shows that the peaks in the imidazole gradient contained very low concentrations of SpoIIE<sub>375-590</sub> with the large wash peak showing a high concentration of the SpoIIE fragment. The high concentration peak was isolated, concentrated and applied to a Superdex 200 (S200) 16/60 gel filtration column (GE Biosciences) (Figure 3-7). The A<sub>280</sub> chromatogram and SDS-PAGE analysis revealed a major SpoIIE<sub>375-590</sub> peak eluting at 90 mls with a much less intense SpoIIE<sub>375-590</sub> containing peak at 78 mls. The major peak, having very minor contaminants, was concentrated to 13.6 mg/ml in 400 µl and stored at -80°C, a total yield of around 5.5 mg of purified SpoIIE<sub>375-590</sub>.

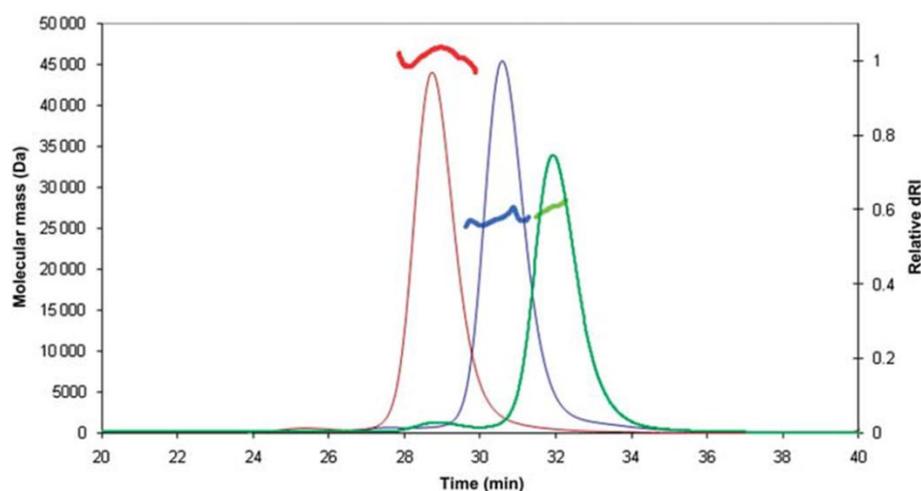


**Figure 3-7: Size Exclusion purification of SpoIIE<sub>375-590</sub> after 6xHis affinity tag cleavage.** The A<sub>280</sub> chromatogram is shown as a blue trace.

The purified SpoIIE<sub>375-590</sub> fragment was placed in crystallisation trials against the Wizard (emerald biosystems), Natrix (hampton research) and Hampton I/II

screens at protein concentrations of 14 and 10 mg/ml (this second concentration with added 5 mM TCEP, a reducing agent). Analysis of the screens showed no promising hits, with around 50% of conditions yielding amorphous precipitate.

Along with the previously isolated SpoIIE fragment H1 (412-827) and the PP2C $\alpha$  domain (590-827), SpoIIE<sub>375-590</sub> was analysed by Size Exclusion Chromatography with Multi-Angle Laser Light Scattering, described in section 2.7.5. This method allows molecular weight calculation of species as they elute from a size exclusion column. This calculation is based on the relationship between the refractive index of a protein solution and the extent to which larger molecules in the solution scatter light. The fragments were injected at a concentration of 1 mg/ml onto an S200 10/300 gel filtration column (GE Biosciences) equilibrated at 0.5 ml/min with a mobile phase consisting of 50 mM Tris-HCl and 150 mM NaCl at pH 8.0. The theoretically calculated molecular weights of H1, PP2C $\alpha$  and SpoIIE<sub>375-590</sub> are 47 kDa, 26.5 kDa and 25.4 kDa, respectively. The SEC-MALLS results show that each fragment elutes as a single major peak with an experimentally calculated molecular weight close to that calculated from their primary sequences; all fragments appear therefore to be monomeric (Figure 3-8).



**Figure 3-8: SEC-MALLS traces of SpoIIE fragments.**

The molecular mass (left vertical axis) and differential refractive index (dRI) (right vertical axis) are shown. Experiments carried out using a Superdex S200 column. The bold lines give molecular mass of the eluting species calculated from measurements of the refractive index and the multi-angle laser light scattering. Three traces for the (i) H1 (red), (ii) PP2C (green) and (iii) B2- B1 (blue) fragments are overlaid.

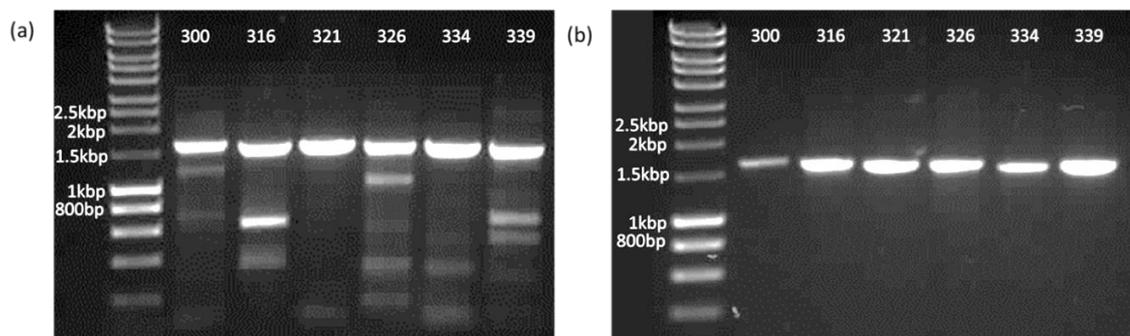
### 3.3.3 Identification of the TM:FtsZ binding domain boundary

The domain boundaries within SpoIIE are not well defined. Bioinformatics and library screen based methods have failed to identify a soluble SpoIIE construct containing the whole cytoplasmic domain. The N-terminal membrane-spanning domain extends to somewhere between the amino acid residues 300 and 340. This section describes multiple bioinformatics analyses which were used to guide the design and testing fragments of SpoIIE with the goal of isolating the largest possible soluble domain II and III construct.

The sequence of SpoIIE was analysed using three web based bioinformatics tools; GlobPlot 2.3 [189], TMHMM predictor [190] and PSIPRED [191]. GlobPlot 2.3 predicts the likely domain distribution in a given amino acid sequence, accompanying literature states that GlobPlot excels in identifying linear motifs of residues which often constitute domain boundaries. The GlobPlot analysis did not predict any disordered regions in the N-terminal 575 residues of the SpoIIE sequence. GlobPlot identified nine transmembrane helices at the N-terminus and the PP2C $\alpha$  domain (aas 592-804) at the C-terminus. The programme further described residues 322-802 as a PP2C $\alpha$  domain. A search for known PP2C $\alpha$  domains yields no sequences of around 480 residues in length, the average being around 240 residues. Numerous BLAST (Basic Local Alignment Search Tool) searches have also failed to find any homology for the sequence between 322 and 590 of SpoIIE. It is possible that the algorithm has recognized a linear motif and assigned a long domain containing homology to a PP2C $\alpha$  domain, another isoform of PP2C. The C-terminal end of the final predicted transmembrane helix is at residue 302.

TMHMM predictor is designed to recognise sequences of around 20 residues likely to fold into hydrophobic  $\alpha$ -helices which are stabilised in and traverse cell membranes. This tool was chosen in order to identify the C-terminal limit of the three-hundred plus residue region of SpoIIE consisting of these helices, the transmembrane domain. TMHMM predicted nine transmembrane helices, with a probability of around 1, ending very close to residue 310. A further 20 residue section of the sequence was assigned a low probability of forming a transmembrane helix.



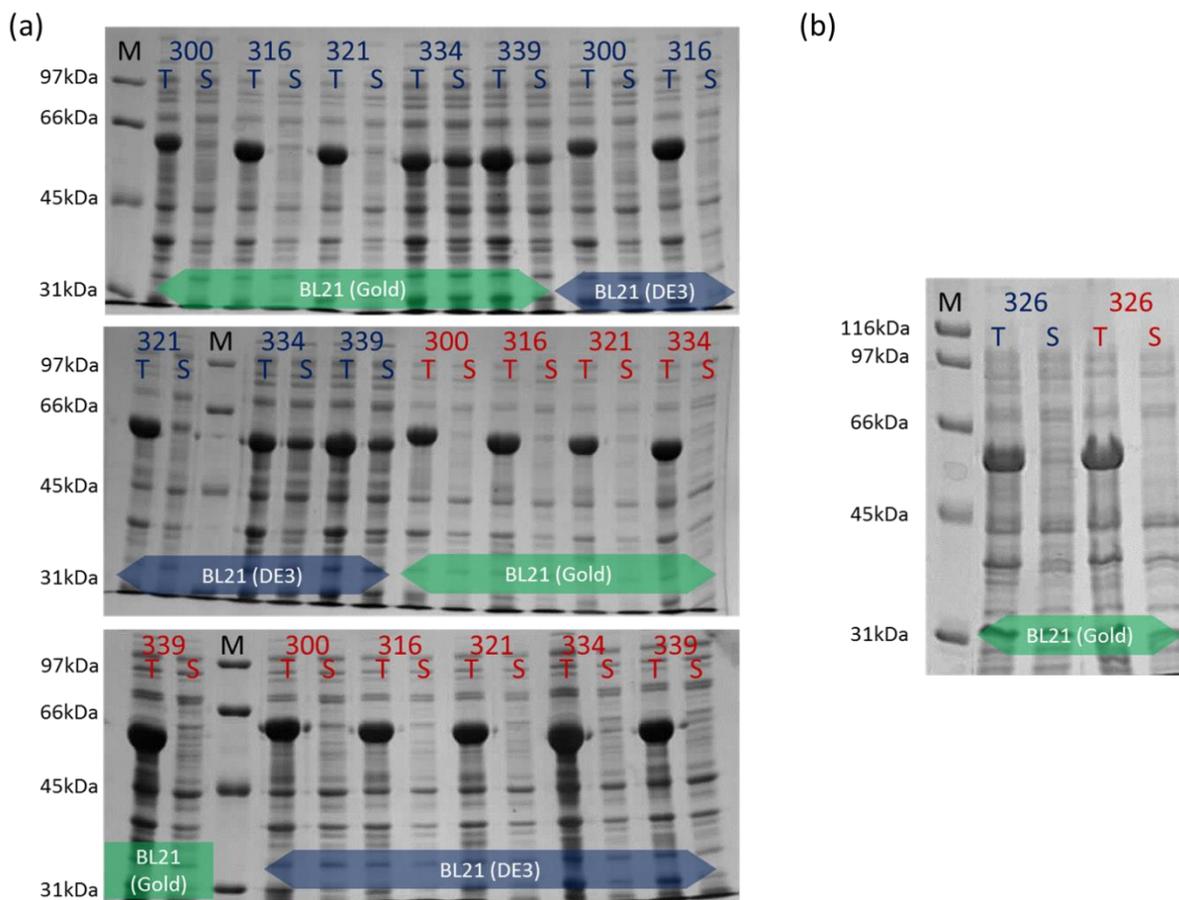


**Figure 3-10: PCR products of *spoIIE* amplification separated by electrophoresis on 1% agarose gels.**

SybrSafe, a DNA binding UV fluorescent dye was used for visualisation. The DNA ladder used is the Bioline Hyperladder I. (a) Raw PCR reaction products. (b) Products after purification by Gel Extraction.

PCR amplification of the *SpoIIE* gene fragments was initially carried out with the following program: (5 min; 94°C, 35x[40 s; 94°C, 30 s; 40°C, 25 s; 72°C], 3 min; 72°C, HOLD; 4°C). However, these PCRs failed and fragments of the expected length, (~ 1.6 kbp), were absent in some cases. As changing the Mg<sup>2+</sup> concentration can affect template binding by PCR primers and increase product specificity, the PCRs were repeated with the MgSO<sub>4</sub> concentration doubled [192]. The high Mg<sup>2+</sup> concentration PCR produced fragments of the expected size as the major product in each case (Figure 3-10). The target bands were excised and purified using a QiaQuik Gel Extraction Kit.

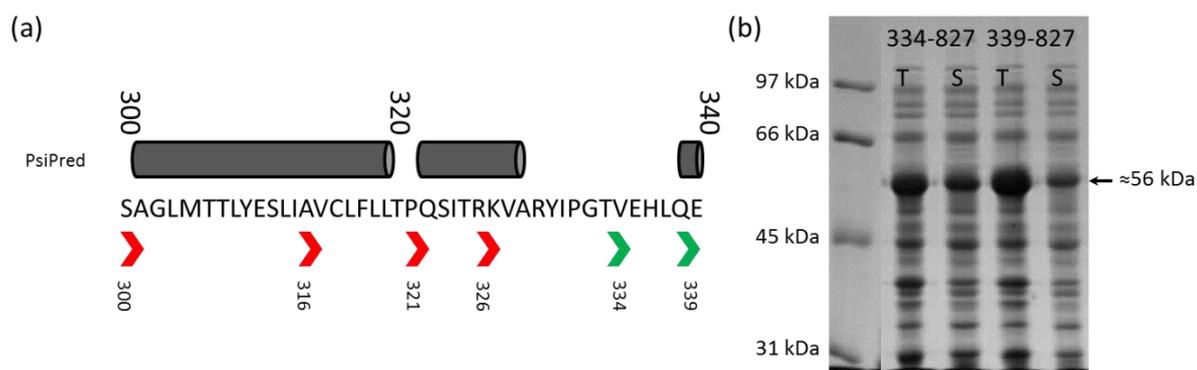
After treatment with T4 DNA polymerase, as described in section 2.3.5, the amplified gene fragments were annealed to the pET-YSBLIC3C vector, and the products were used to transform competent *E. coli* XL-10 Gold cells as described in section 2.3.6. The sequences of the recombinant plasmids obtained were determined to confirm that the expected fragments had been cloned in the correct reading frame. Next, the plasmids were introduced into *E. coli* strains BL21 (DE3) and BL21 (Gold) for expression of the target protein fragments. Small scale cultures, grown at 16°C overnight and at 37°C for four hours, were used to test for expression of soluble protein by SDS-PAGE analysis (Figure 3-11).



**Figure 3-11: SDS-PAGE images of SpoIIE fragment expression tests.**

M indicates Biorad protein markers. Lane headings in blue represent expression carried out at 16°C overnight. Lane headings in red represent expression at 37°C for four hours. Total and soluble cell fractions are shown for each condition. Lysis buffer used is 50 mM Tris-HCl, 150 mM NaCl at pH 8.0. Cell strains are shown by blue or green banners in the bottom of each image. (a) Expression tests for fragments beginning at 300, 316, 321, 334 and 339. (b) Expression test for fragment beginning at 326.

All test expression cultures were found to contain overexpressed protein of the expected molecular weight; however, all but two of the twelve constructs showed the target protein to be insoluble. The only soluble SpoIIE fragments were expressed at 16°C, these being the two shortest fragments SpoIIE<sub>334-827</sub> and SpoIIE<sub>339-827</sub>. Both of these soluble fragments have their N-termini in an eleven amino acid residue stretch (329-340) that lacks predicted secondary structure. The pattern of soluble fragments would suggest that the small  $\alpha$ -helix predicted between 321 and 327 is the last element of secondary structure preventing solubility (Figure 3-12) and suggests the 9 residue putative loop preceding 340 as the interface between domains 1 and 2 of SpoIIE.



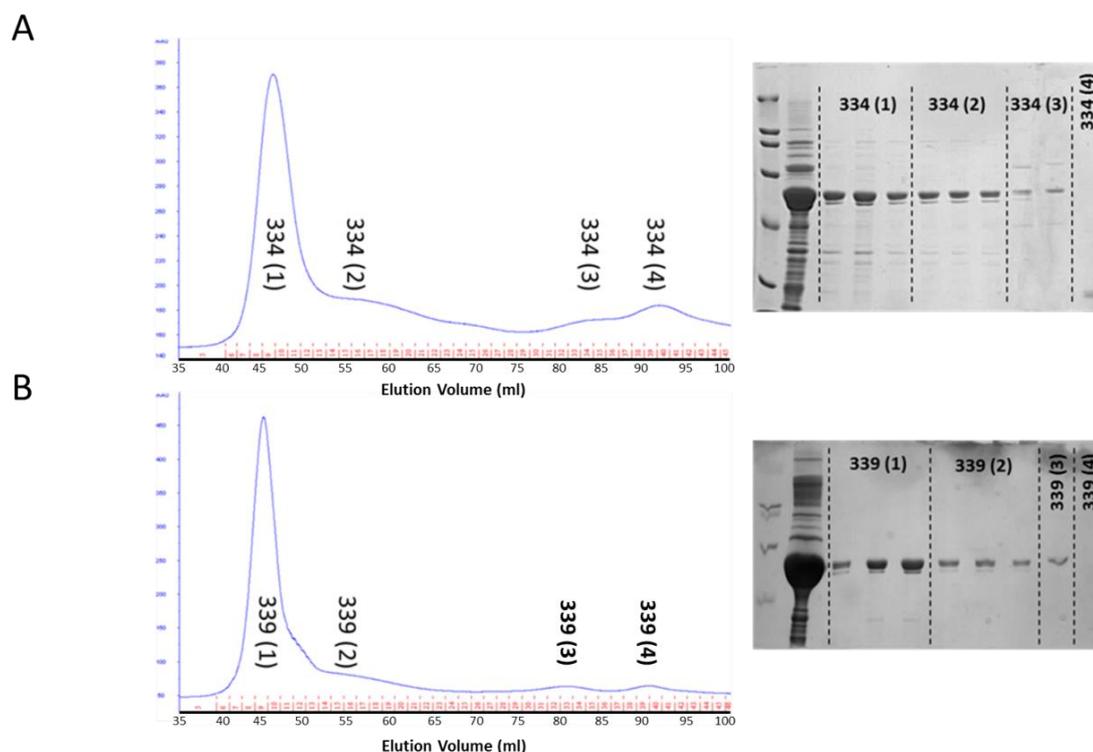
**Figure 3-12: The TM domain : Domain II boundary region.**

(a) Schematic representation of the 300 – 340 region of SpoIIE. The PSIPRED secondary structure prediction is shown as grey cylinders for  $\alpha$ -helices. The constructs studied are represented by arrowheads at the relevant N-terminal residue. These arrows are colour coded; red for insoluble and green for soluble. The amino acid sequence in this region is also shown. (b) SDS-PAGE image of soluble expression of the two constructs; 334-827 and 339-827.

The SpoIIE<sub>334-827</sub> and SpoIIE<sub>339-827</sub> fragments represent the longest SpoIIE soluble fragments identified to date. While the level of solubility seen in the above expression tests was acceptable, it was thought that improved solubility would simplify purification and give stable protein. Therefore the SpoIIE<sub>334-827</sub> fragment was screened using the sparse matrix solubility screen, consisting of 30 lysis buffers [193]. This led to a buffer based on 100 mM triethanolamine, 50 mM LiCl, 5 mM EDTA at pH 8.5 being chosen for future work.

*E. coli* BL21 (Gold) cells harbouring the SpoIIE<sub>334-827</sub> and SpoIIE<sub>339-827</sub> fragments were each cultured in 4 flasks containing 750 ml LB media supplemented with 30  $\mu$ g/ml kanamycin. Production of SpoIIE<sub>334-827</sub> was induced, as described in section 2.5.2, with expression overnight at 16°C before harvesting of cells by centrifugation and their lysis by sonication in 100 mM triethanolamine, 200 mM NaCl, 30 mM imidazole at pH 8.5. The omission of EDTA from the buffer was in order to avoid abstraction of Ni<sup>2+</sup> during IMAC column chromatography; no adverse effect on solubility was detected. Increasing the salt concentration to 200 mM NaCl improved the extent of purification achieved by IMAC. Lysate, clarified by centrifugation, was applied to a 5 ml HisTrap column for nickel IMAC purification. After washing unbound protein from the column with low imidazole buffer, bound protein was eluted using a linear (30-500 mM) imidazole gradient. Both fragments eluted in the range 125 - 200 mM imidazole

as a single broad peak. Cleavage of the N-terminal 6xHis affinity tag by HRV 3C protease led to degradation of the fragments so this step was omitted. One observation made throughout the purification was that filters in the AKTA FPLC became rapidly blocked, due to either SpoIIIE<sub>334-827</sub> aggregation, to the high viscosity of the triethanolamine in the buffers or to a combination of both.



**Figure 3-13: Size exclusion A<sub>280</sub> chromatograms of the SpoIIIE<sub>334-827</sub> and SpoIIIE<sub>339-827</sub> fragments with SDS-PAGE gel images.**

(a) SpoIIIE<sub>334-827</sub>. (b) SpoIIIE<sub>339-827</sub>.

The peak fractions from nickel IMAC were concentrated and applied to a 16/60 S200 gel filtration column (GE Biosciences) with a mobile phase of 100 mM Triethanolamine, 200 mM NaCl, 5 mM DTT at pH 8.5. The separation range of this column is given as 10 – 600 kDa with the highest selectivity between 30 and 250 kDa. The A<sub>280</sub> chromatograms of each fragment (Figure 3-13) showed a large portion of protein in the void volume of the column, at 45 mls, which could consist of species ranging between near the separation limit of 600 kDa to multi-megadalton aggregates. A broad shoulder peak to the void volume was in the cases of both fragments centred at 56 mls; this elution might suggest 6-8mer species of around 400 kDa in size. A further two peaks were seen in each

chromatogram at 80 and 90 mls, estimated molecular masses for these elution volumes would be around 50 and 25 kDa, respectively. The peak eluting at 90 mls would represent a molecule smaller than a monomer and, indeed, SDS-PAGE analysis showed no target protein in the peak. The peak at 80 mls is at the expected elution volume for a monomeric species, according to manufacturer's calibration data. All other peaks showed the presence of target protein with a small impurity at a very slightly lower molecular weight.

Both the void volume peak and the second peak for each fragment were isolated and concentrated. In the case of both fragments the void volume peak precipitated to a large extent during concentration and a large amount of protein was lost. The final yields were around 5 mg of protein for both fragments; the void volume peak could be concentrated to around 6 mg/ml and the high MW peak to around 4 mg/ml.

The protein concentrations achieved are roughly half that used as a starting point in protein crystallisation experiments of around 10 mg/ml; however some proteins will crystallise at much lower concentrations. Another complication is that the void volume peak could contain any number of species. Crystallisation, and more importantly growth of crystals that will diffract, requires the slow and ordered assembly of homologous subunits. The subunits in large oligomers or aggregates will generally cause problems in this process. Regardless, protein from both fractions of each fragment was placed in crystallisation trials at the concentrations achieved. The screens used were PACT, Hampton I/II and JCSG (molecular dimensions). Analysis of the screens showed no promising hits with a large number of conditions demonstrating amorphous or sandy precipitate.

The oligomerisation of SpoIIE has been observed in every fragment with an N-terminus between 334 and 412. The PP2C $\alpha$  domain fragment, SpoIIE<sub>590-827</sub>, did not oligomerise. The B2 (SpoIIE<sub>375-827</sub>) and H1 (SpoIIE<sub>412-827</sub>) fragments both demonstrated oligomeric behaviour, with much protein eluting in the void volume of an S200 gel filtration column. These fragments also have a dimer fraction (estimated by elution volume). The B2-B1 (SpoIIE<sub>375-590</sub>) fragment eluted in the S200 void volume with what is likely to be a monomeric species

eluting at 90 mls. It should be noted that the SEC-MALLS experiments are carried out at lower protein concentrations. The 1 mg/ml protein sample loaded will have been diluted by a factor of fifteen if it elutes as a 1.5 ml peak; a concentration of 0.07 mg/ml. Analysis of the B2 fragment on an S200 gel filtration column shows that at an injection concentration of 7.4 mg/ml protein elutes in three species including a putative dimer, whereas at a higher injection concentration of 20 mg/ml the protein elutes almost entirely in the void volume. Oligomerisation is of course concentration dependent, however the SpoIIE<sub>334-827</sub> and SpoIIE<sub>339-827</sub> fragments of SpoIIE were never observed with a major species of lower molecular weight than the void volume.

The initial aim of this section was to identify long and soluble fragments of SpoIIE in order to extend the very small number of SpoIIE reagents available for study. This aim was achieved with two constructs being produced which we believe to represent the greater proportion of the cytoplasmic domain of SpoIIE. The constructs behave as high order oligomers in solution and this is likely to cause problems in crystallisation and potentially in biophysical assays where precipitation might be a problem. As both constructs seem to behave similarly the decision was taken to continue study of the SpoIIE<sub>334-827</sub> fragment, the longest SpoIIE fragment yet investigated, while delaying work on the SpoIIE<sub>339-827</sub> fragment. During purification it was identified that cleavage of the 6xHis affinity tag caused instability and degradation, in uncleaved samples. Whether the tag causes misfolding; resulting in instability on cleavage, or the N-terminal segment, which is exposed on cleavage, is susceptible to degradation is unknown.

A close collaborator, Niels Bradshaw, from the laboratory of Richard Losick, Harvard, communicated his observation that long fragments of SpoIIE form oligomers and aggregate in a Mn<sup>2+</sup> dependent fashion. A SpoIIE fragment he purified using a 6xHis-SUMO domain solubility tag precipitated when exposed to millimolar concentrations of MnCl<sub>2</sub>. Manganese (II) can be chelated by EDTA and hence the use of this additive could decrease the fraction of SpoIIE which forms oligomers. This assumes, however, that the oligomers seen in size exclusion chromatography are due to Mn<sup>2+</sup> sequestered by SpoIIE during

expression. The addition of 10 mM EDTA to the solution before size exclusion chromatography did not alter the observed oligomeric state of *BsuSpoIIE*<sub>334-827</sub> in subsequent isolations.

### 3.3.3 Analogue of SpoIIE from *Geobacillus stearothermophilus*

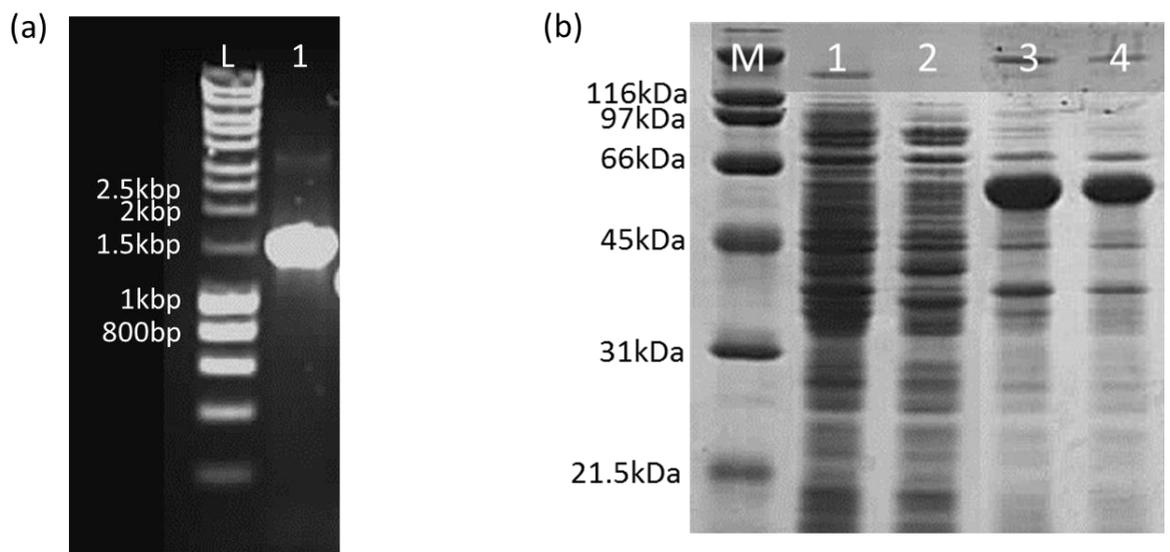
In an attempt to increase the stability of the SpoIIE<sub>334-827</sub> fragment, we sought the thermophilic homologue from *Geobacillus stearothermophilus* (*Gst*), a spore-forming thermophile with an optimum growth temperature of between 50 and 65°C [194]. This strategy has been used successfully in other systems [195, 196].



**Figure 3-14: Primary sequence alignment of SpoIIE from *B. subtilis* and *G. stearothermophilus*.**

Alignment carried out using ClustalW [197]. Visualisation and structural annotation of PP2Ca domain carried out using ESPript (<http://esprict.ibcp.fr>) [198].

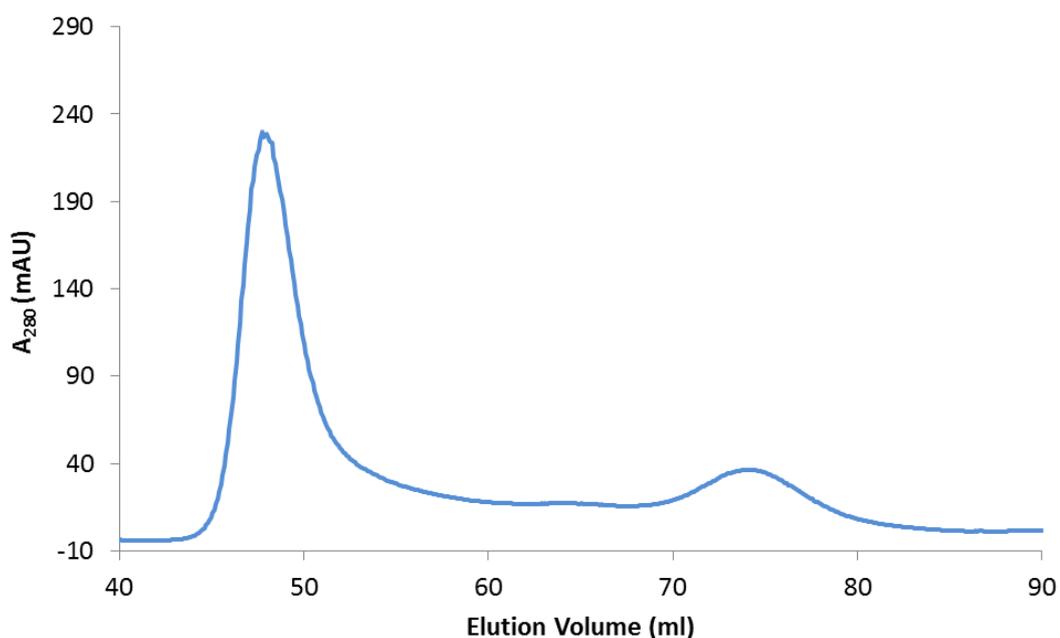
Sequencing and annotation of the *G. stearothermophilus* genome is incomplete. The greater part of the *G. stearothermophilus* strain 10 genome has been sequenced in the lab of Bruce Roe, University of Oklahoma, and the results made available online at <http://www.genome.ou.edu/bstearo.html> (*Bacillus (Geobacillus) stearothermophilus* Genome Sequencing Project funded by NSF EPSCoR Program). In order to find the nucleotide sequence of *spoIIE* in the unannotated genome sequence, it was necessary to carry out a BLAST search. The search algorithm used was tblastx. The input, the *BsuSpoIIE* amino acid sequence, is translated into a nucleotide sequence in every possible permutation of codon usage and reading frame; sequences showing good alignment to the query are reported to the user. An Open Reading Frame (ORF) encoding an 826 residue protein was highlighted in contig.543 encompassing bases 109825 to 112275. The predicted protein has 56% sequence identity to *BsuSpoIIE* (Figure 3-14).



**Figure 3-15: Gel images of *GstSpoIIE*<sub>335-826</sub> cloning and expression.**

(a) PCR amplification products of *spoIIE*<sub>1003-2481</sub> on 1% agarose gel. L represents Bioline Hyperladder DNA markers. Lane 1 shows PCR products; the expected fragment size is 1.47 kbp. (b) Expression test of *GstSpoIIE*<sub>335-826</sub>. M represents Biorad broad range markers; lane 1, Uninduced total cell fraction; lane 2, Uninduced soluble cell fraction; lane 3, Induced total cell fraction; lane 4, Induced soluble cell fraction.

The *Gst spoIIIE* gene fragment encoding *GstSpoIIIE*<sub>335-826</sub> was amplified from *G. stearothermophilus* strain 10 genomic DNA by PCR using primers designed to facilitate cloning into the pYSBLIC3C expression vector. The nucleotide fragment was readily amplified and after cloning and transformation of *E. coli* BL21 (Gold) cells, production of the desired 55 kDa protein was achieved. The protein was soluble in a lysis buffer of 50 mM Tris-HCl, 200 mM NaCl at pH 8.5 (Figure 3-15).



**Figure 3-16: Size Exclusion purification of *GstSpoIIIE*<sub>335-826</sub>.**  
The A<sub>280</sub> chromatogram is shown as a blue trace.

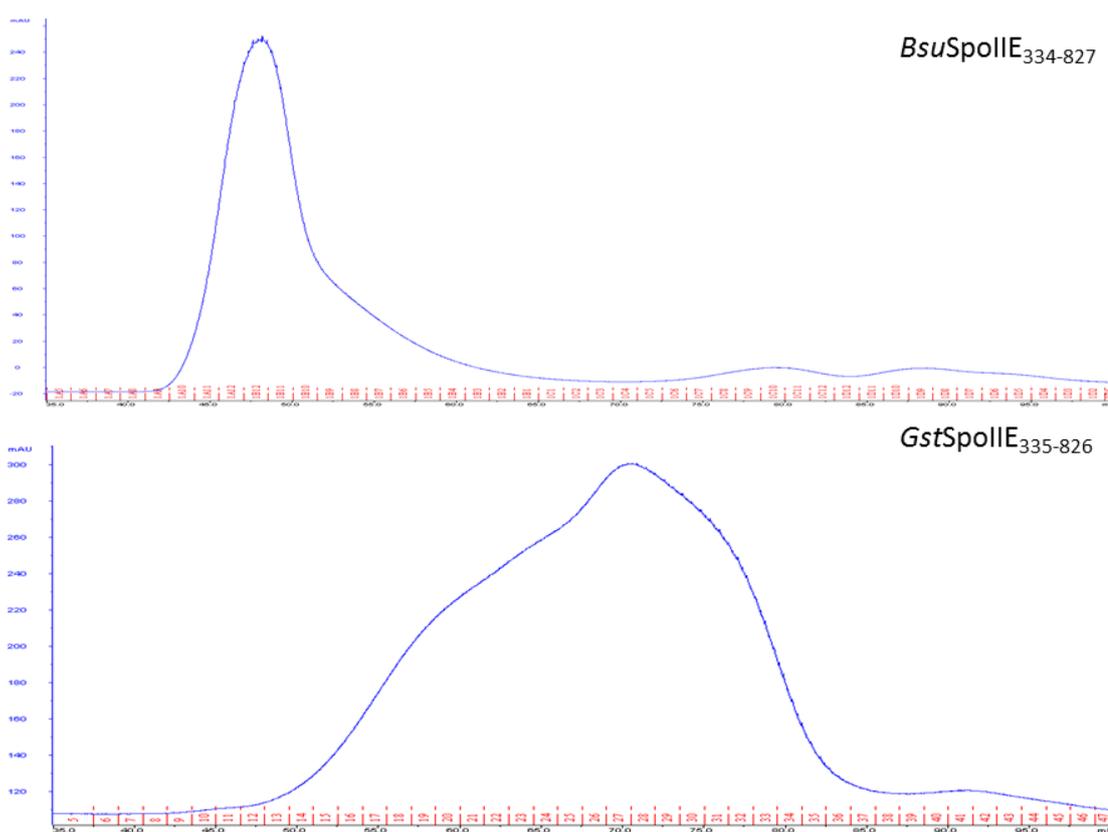
A large scale preparation of the fragment *GstSpoIIIE*<sub>335-826</sub> was carried out from 4 flasks each containing 750 ml of LB media. Harvested cells were lysed in 50 mM Tris-HCl, 200 mM NaCl, 20 mM imidazole at pH 8.5. Lysate, clarified by centrifugation, was applied to a 5 ml HisTrap nickel IMAC column. The IMAC column was washed to remove unbound protein before a linear gradient to a final concentration of 500 mM imidazole was applied. A peak eluting at 140 mM imidazole was found by SDS-PAGE analysis to contain the target protein fragment. Some protein was lost in the concentration of the pooled peak fractions before purification by size exclusion chromatography. After S200 gel filtration column chromatography (Figure 3-16) in 50 mM Tris-HCl, 200 mM NaCl, 20 mM EDTA at pH 8.5, SDS-PAGE analysis showed a large proportion of

protein in the column void volume with a smaller peak eluting at around 75 mls, the elution volume expected for a monomer. The addition of EDTA did not reduce the formation of oligomers, as judged by the size of the void volume peak in a second purification by size exclusion chromatography.

Comparing the sequences of the two fragments reveals high sequence identity. Conspicuously however, the number of cysteine residues differs. *BsuSpoIIE*<sub>334-827</sub> contains five cysteine residues while *GstSpoIIE*<sub>335-826</sub> contains eight. Four of these cysteine residues, in the region between residues 399 and 446 of SpoIIE are conserved between species. *BsuSpoIIE*<sub>334-827</sub> has a single cysteine residue in the phosphatase domain. It is possible that additional cysteine residues may contribute to the thermal stability of the *Gst* SpoIIE fragment via the formation of disulphide bridges.

### 3.3.4 SpoIIE<sub>334-827</sub> fragment oligomerisation.

A parallel purification of *BsuSpoIIE*<sub>334-827</sub> and *GstSpoIIE*<sub>334-827</sub> was carried out; the same buffer batches were used over the same time period with the only variation being in the FPLC machine used. Cells from 4 x 750 ml cultures, were lysed by sonication in 50 mM Tris-HCl, 200 mM NaCl, 2 mM β-mercaptoethanol (β-ME), 20 mM imidazole and the clarified lysates were applied to 5 ml HisTrap nickel IMAC columns. After washing, bound protein was eluted from the column using a linear gradient with a final concentration of 500 mM imidazole. Both fragments eluted at around 140 mM imidazole. The nickel IMAC peak fractions were pooled, 5 mM (final concentration) EDTA was added the samples were concentrated to ~ 3 ml.

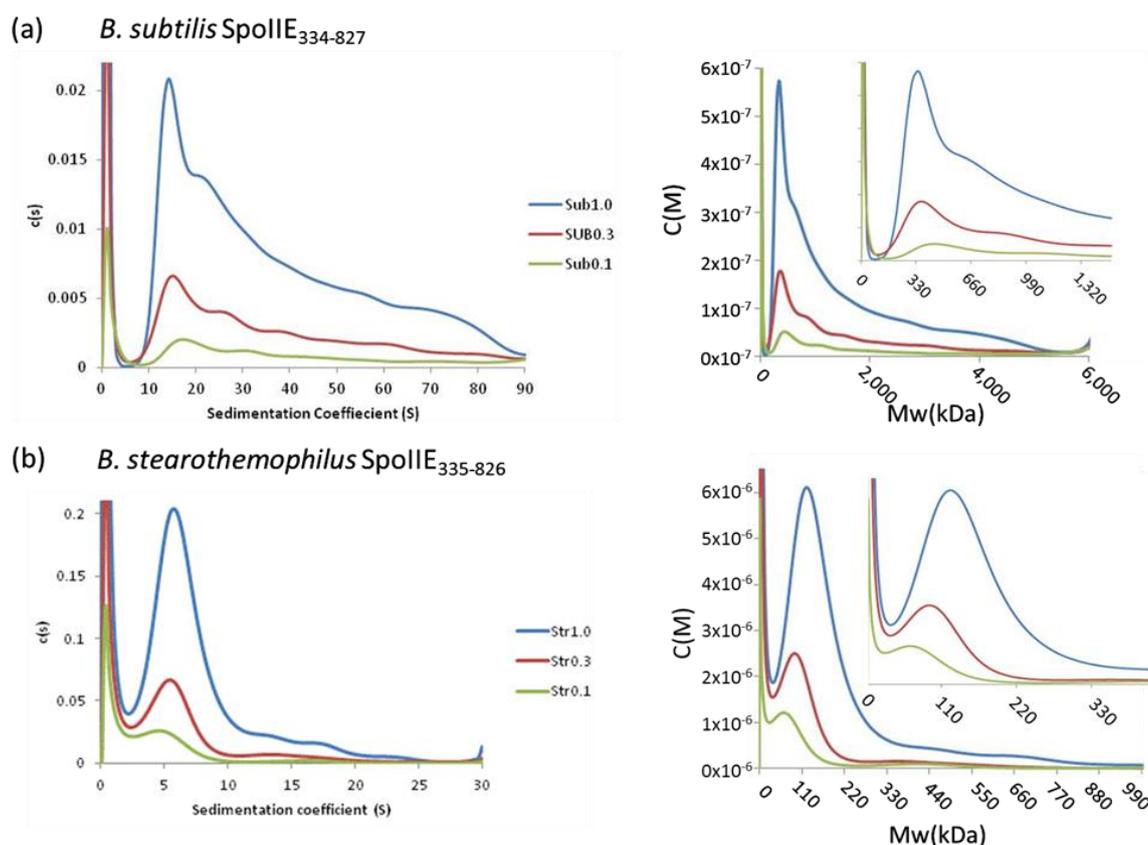


**Figure 3-17: Parallel purification of *BsuSpoIIE*<sub>334-827</sub> and *GstSpoIIE*<sub>335-826</sub> fragments on Superdex 200 gel filtration columns.**

A running buffer of 50 mM Tris-HCl, 200 NaCl, 5 mM EDTA, 2 mM β-ME was used. A<sub>280</sub> chromatograms are shown as blue traces. The *BsuSpoIIE*<sub>334-827</sub> elutes almost completely in the column void volume. The *GstSpoIIE*<sub>335-826</sub> fragment elutes as two broad species centred at 60 ml and 75 ml, likely representing dimers and monomers, respectively.

The concentrated protein samples were each applied to an S200 gel filtration column equilibrated in 50 mM Tris-HCl, 200 mM NaCl, 5 mM EDTA, 2 mM  $\beta$ -ME at pH 8.5. When compared to previous purifications, the oligomeric behaviour of the *BsuSpoIIE*<sub>334-827</sub> fragment was unchanged whereas the behaviour of *GstSpoIIE*<sub>335-826</sub> changed markedly; no oligomers were detected in the void volume instead two overlapping peaks centred at 60 and 75 ml were observed, likely representing dimers and monomers, respectively (Figure 3-17). The observed change in the *Gst* fragment's oligomeric behaviour may be attributed to the inclusion in the buffer of  $\beta$ -ME, a reducing agent.

The void volume fraction from the *Bsu* fragment and the combined monomer and dimer fractions from the *Gst* fragment were concentrated to 2.9 mg/ml and 3.5 mg/ml, respectively. The fragments were next analysed by Analytical Ultracentrifugation (AUC) a technique which measures the rates of sedimentation of macromolecular species and gives an accurate estimation of their size. In order to ensure no variation in components between the baseline and sample buffers, dialysis was carried out into freshly prepared 20 mM Tris-HCl, 100 mM NaCl, 2 mM DTT at pH 8.5. The same batch of buffer was then used for diluting samples and as a baseline cell buffer. Three concentrations of each fragment were run at  $A_{280}$  values of 1, 0.3 and 0.1. The extinction coefficients of *BsuSpoIIE*<sub>334-827</sub> and *GstSpoIIE*<sub>335-826</sub> differ and the calculated protein concentrations of the samples are 1.16 mg/ml, 0.48 mg/ml and 0.16 mg/ml, respectively, for *BsuIIE*<sub>334-827</sub> and 1.11 mg/ml, 0.33 mg/ml and 0.11 mg/ml, respectively, for *GstIIE*<sub>335-827</sub>. The samples were centrifuged at 20,000 rpm in a Beckman Optima XL/I analytical ultracentrifuge, using Beckman cells with 12 mm path length, double sector charcoal-filled Epon centrepieces and sapphire windows, in an AN-50Ti rotor. The collected AUC data were analysed by Andrew Leech using sedfit AUC analysis software [199]. Fitting of the data was carried out first to calculate the distribution of sedimentation coefficients, representing the rates at which different species travel from the inner to the outer radii of the cells.



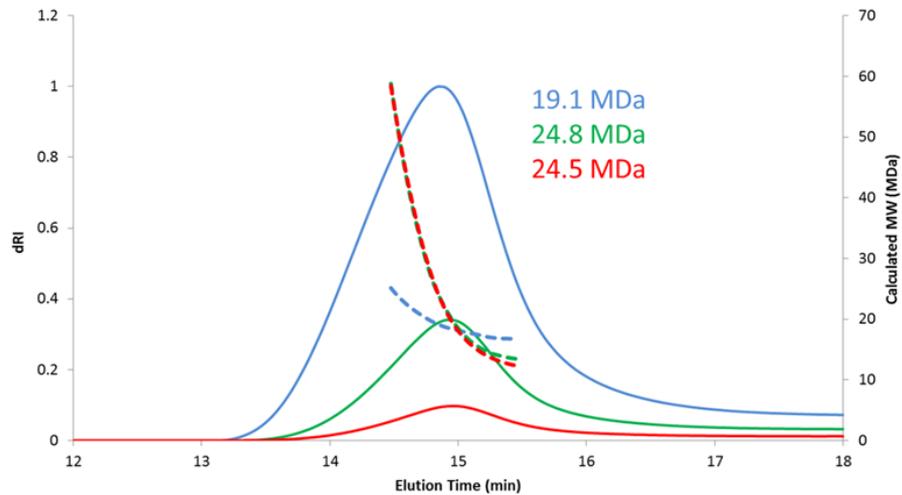
**Figure 3-18: Analytical Ultra-Centrifugation analyses of the *BsuSpoIIE*<sub>334-827</sub> and *GstSpoIIE*<sub>335-826</sub> fragments at  $A_{280} = 1.0, 0.3$  and  $0.1$ .**

Fitting and calculation carried out to determine sedimentation coefficient (S), left, and molecular weight (Mw), right. (a) Analyses of *BsuSpoIIE*<sub>334-827</sub> show multiple species with a 330 kDa periodical increase in Mw. (b) Analyses of *GstSpoIIE*<sub>335-826</sub> suggest equilibrium between monomeric and dimeric species at the studied concentrations.

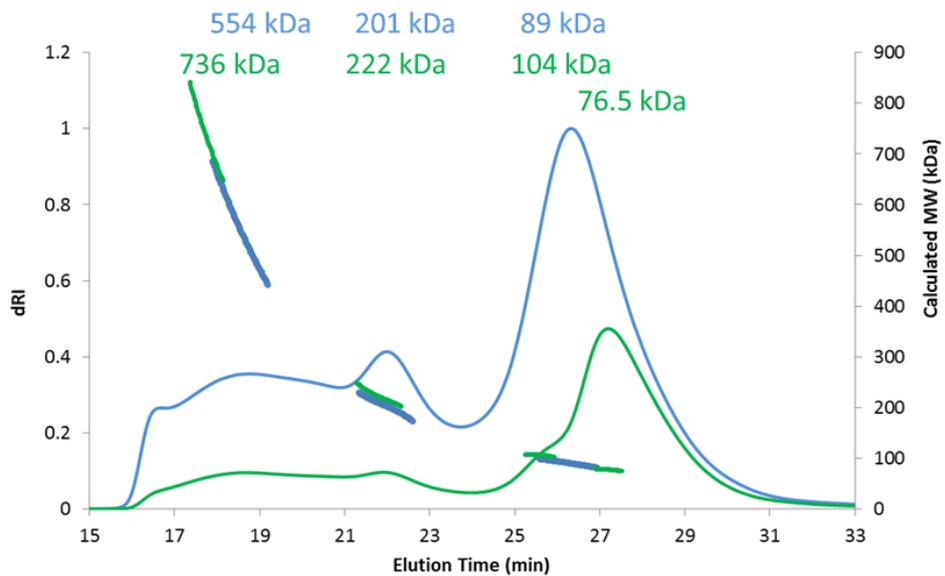
The AUC data analysis results (Figure 3-18), in agreement with the size exclusion chromatography results, show that *BsuSpoIIE*<sub>334-827</sub> has characteristics of a large number of species, whereas the *GstSpoIIE*<sub>335-826</sub> has a narrower species distribution. Further analysis of calculated sedimentation coefficients was then used to predict the probable molecular weights of the species. *BsuSpoIIE*<sub>334-827</sub> shows species at multiples of around 330 kDa. The calculated molecular mass is 55 kDa, hence 330 kDa represents a hexameric unit. The 660 kDa species might represent two hexameric units stacking together. As mentioned, very little is known of the mode of action of SpoIIE *in vivo* other than the requirement of localisation to the forming septum. The observed behaviour here could represent functional units of SpoIIE.

*GstSpoII*E<sub>335-826</sub> shows a single different species in each concentration. Molecular weight analysis predicts species at around 55, 80 and 110 kDa; monomer, 1.5mer and dimer, respectively. The 1.5mer prediction can be explained by the monomer and dimer species being in dynamic exchange at this concentration.

**A**



**B**



**Figure 3-19: SEC-MALLS analyses of *BsuSpoII*E<sub>334-827</sub> and *GstSpoII*E<sub>335-826</sub>.**

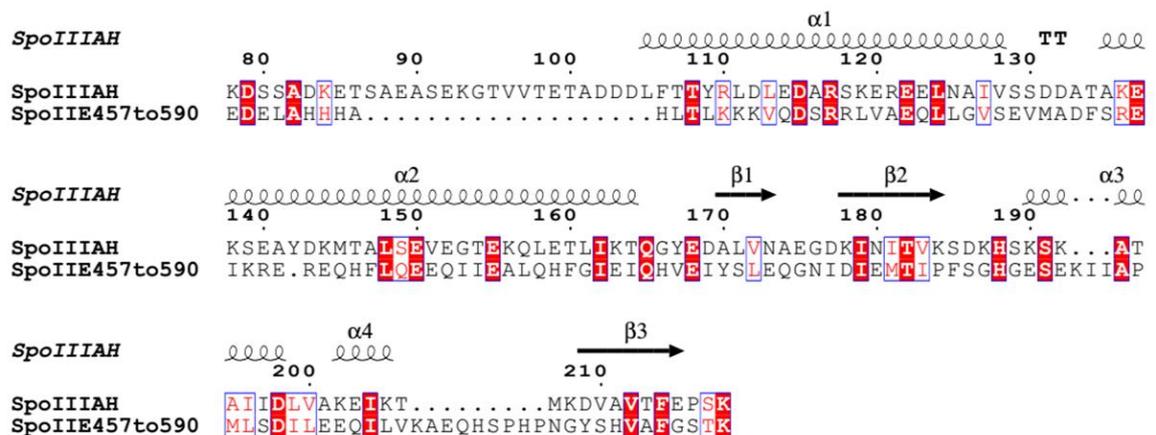
(A) Differential Refractive Index chromatograms of *BsuSpoII*E<sub>334-827</sub> at 3 mg/ml (blue line), 1 mg/ml (green line) and 0.1 mg/ml (red line) as separated by a Superose 6 10/300 gel filtration column. Calculated MW data are show as thick dashed lines with annotated average peak MWs in the same colours. (B) Differential Refractive Index chromatograms of *GstSpoII*E<sub>335-826</sub> at 3 mg/ml (blue line) and 1 mg/ml (green line) as separated by a Superdex 200 10/300 gel filtration column. Calculated MW data are show as thick lines with annotated average peak MWs in the same colours.

*BsuSpoIIE*<sub>334-827</sub> was analysed by SEC-MALLS in 20 mM Tris-HCl, 100 mM NaCl, 2 mM DTT at pH 8.0 using a Superose 6 10/300 gel filtration column (GE biosciences), capable of resolving proteins below 40,000 kDa (Figure 3-19A). Injections were made at concentrations of 3, 1 and 0.3 mg/ml. Each  $A_{280}$  trace showed a single peak eluting after 15 minutes; around the column void volume, 7.5 mls. Molecular mass predictions for these peaks ranged between 19 and 25 MDa. This supports observations made regarding the tendency of this fragment to form large oligomers. Analysis of *GstSpoIIE*<sub>335-826</sub> by SEC-MALLS was carried out with a running buffer of 20 mM Tris-HCl, 100 mM NaCl, 2 mM DTT at pH 8.0 and separation on an S200 10/300 gel filtration column (Figure 3-19B). Injections were made at 1 and 3 mg/ml. Both  $A_{280}$  traces show a range of peaks eluting across the separation volume of the column. Predicted molecular masses for these species suggest monomer/dimer equilibrium as well as small populations of tetramers and higher order species. These observations are at odds with those seen by AUC and are unlikely to be explained by the small change in pH.

### 3.3.5 Possible homology of SpoIIE to the sporulation stage III protein, SpoIIIAH

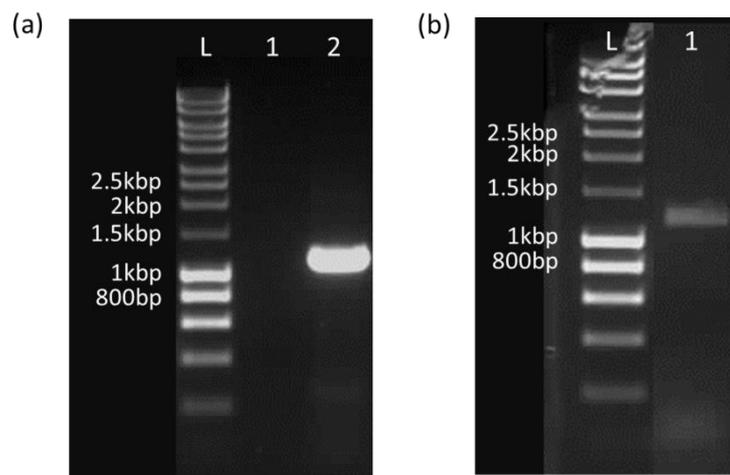
A possible homologous relationship between SpoIIIAH, a protein involved in the engulfment stage (III) of sporulation, and residues 457-590 of SpoIIE was communicated to us by Niels Bradshaw, Harvard (personal communication). This was a particularly interesting observation because SpoIIIAH and SpoIIE are known and proposed binding partners of SpoIIQ, respectively [91, 162]. SpoIIIAH and SpoIIQ associate to form a ‘ratchet-like’ mechanism across the intercellular space to assist membrane migration during engulfment and are also proposed to be integral components of a feeding tube structure allowing nutrient exchange between the mother cell and engulfed forespore. Fluorescence microscopy studies have shown that SpoIIE is released from the division septum after  $\sigma^F$  activation, but subsequently relocates to the engulfing septum in a SpoIIQ-dependent manner.

The sequence identity between the SpoIIIAH<sub>78-218</sub> and SpoIIE<sub>457-590</sub> sequences is 17% with regions of similarity generally clustered in areas of secondary structure (Figure 3-20). The 457-827 fragment of SpoIIE was produced with good solubility in *E. coli* with a 6xHis-SUMO tag (Niels Bradshaw, per. comm.). It therefore represents a good candidate for structural studies.



**Figure 3-20: Primary sequence alignment of SpoIIIAH<sub>78-218</sub> and BsuSpoIIE<sub>457-590</sub>.** Alignment carried out using T-Coffee [200]. Visualisation and SpoIIIAH structural annotation (PDB ID: 3TUF) carried out using ESPrnt (<http://esprnt.ibcp.fr>) [198].

Fragments encoding SpoIIE<sub>457-827</sub> from *B. subtilis* and SpoIIE<sub>457-826</sub> from *G. stearothermophilus* were therefore cloned into the pYSBLIC3C expression vector. PCR amplification of the target gene fragments from *Bsu* and *Gst* genomic DNA produced single products of the expected size; 1.1 kbp. After ligation independent cloning, the same fragment could be amplified by PCR from the recovered plasmids indicating the presence of the SpoIIE fragment insert (Figure 3-21).



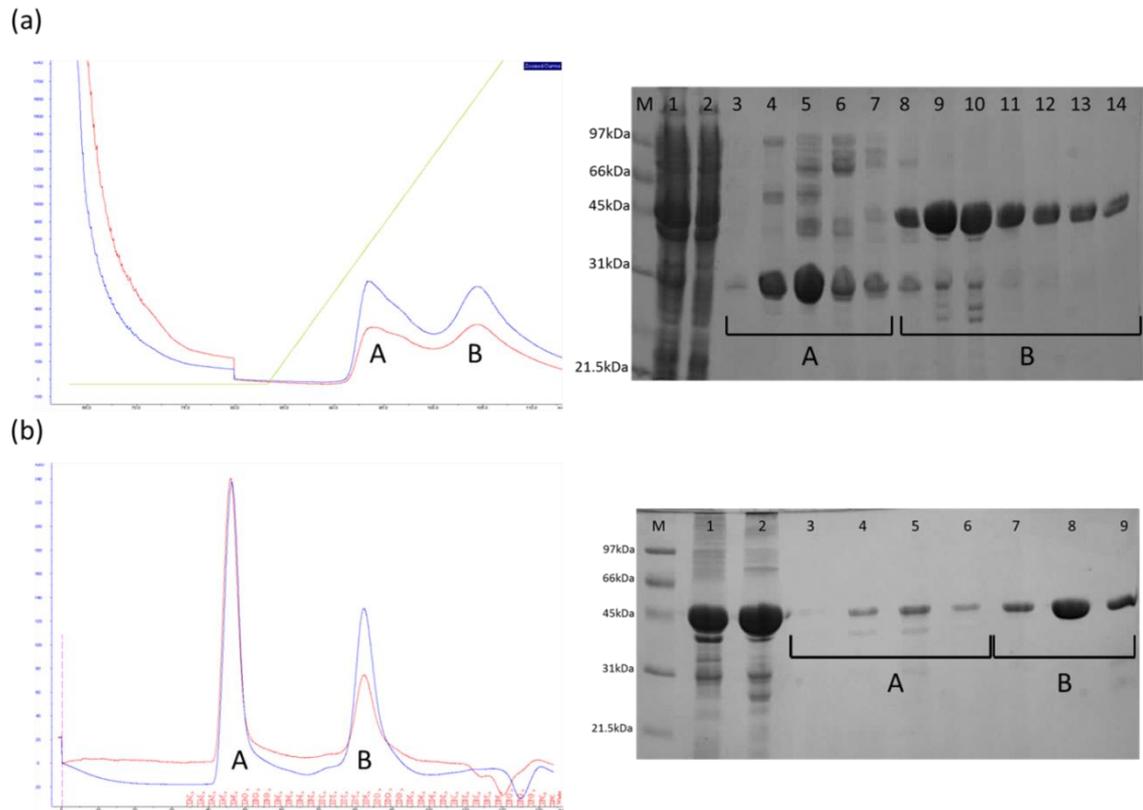
**Figure 3-21: Images of 1% agarose gels of PCR amplifications of the *BsuSpoIIE*<sub>457-827</sub> coding nucleotide sequence visualised using SybrSafe DNA dye.**

L represents Bioline Hyperladder I DNA markers (a) PCR amplification from *Bsu* genomic DNA. Lane 1; no template (negative control), lane 2; PCR amplification. (b) PCR amplification from plasmid isolated by miniprep after ligation independent cloning. Lane 1; PCR amplification.

Plasmids harbouring the target nucleotide sequences were used in transformation of *E. coli* BL21 (Gold) competent cells and expression of the target protein fragments was induced. The *BsuSpoIIE*<sub>457-827</sub> and *GstSpoIIE*<sub>457-826</sub> fragments have calculated molecular masses of 41.3 and 41.0 kDa, respectively. Both fragments were produced at moderate levels but with good solubility following cell growth overnight at 16°C and lysis in 50 mM Tris-HCl, 100 mM NaCl at pH 8.0.

The *BsuSpoIIE*<sub>457-827</sub> fragment was overproduced on a larger scale by IPTG induction overnight at 16°C in 2 flasks each containing 500 ml LB media. Cells were harvested by centrifugation and resuspended in a lysis buffer consisting of 50 mM Tris-HCl, 100 mM NaCl, 20 mM imidazole at pH 8.0 before lysis by sonication. Lysate clarified by centrifugation was applied to a 5 ml HisTrap

nickel IMAC column equilibrated in lysis buffer. Unbound protein was washed from the column before bound protein was eluted using a linear gradient with a final concentration of 500 mM imidazole.



**Figure 3-22: Purification of the *BsuSpoIIE*<sub>457-827</sub> fragment.**

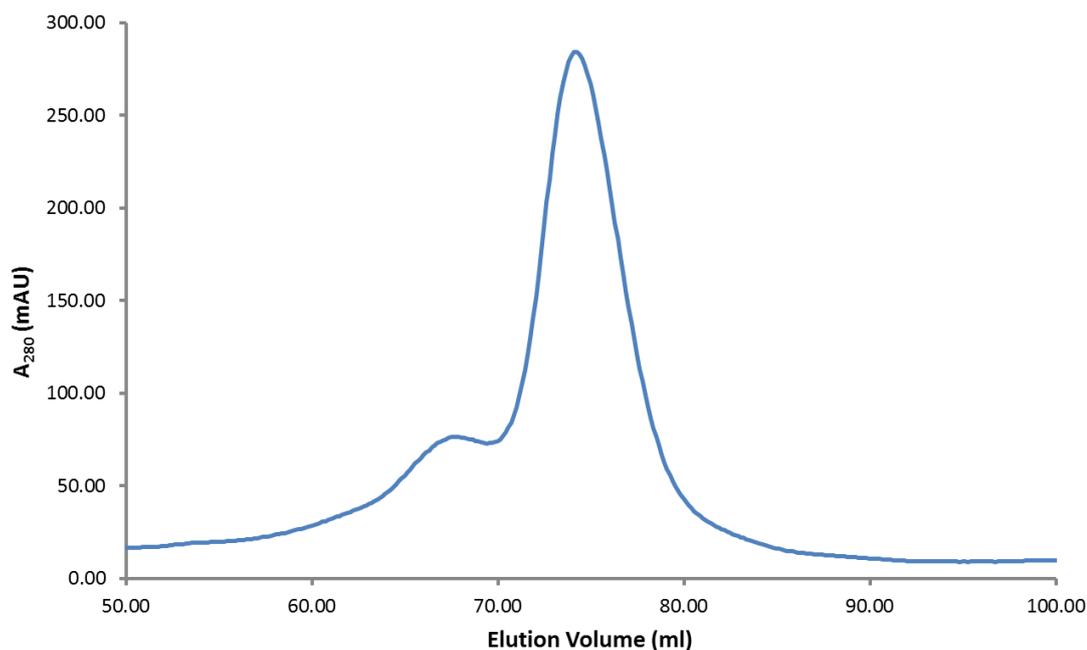
(a) A<sub>280</sub> and A<sub>260</sub> chromatograms of nickel IMAC purification with SDS page analysis. Peak A contains a 31 kDa contaminant. Peak B contains mainly target protein. (b) A<sub>280</sub> and A<sub>260</sub> chromatograms of size exclusion purification with SDS page analysis. Only a small concentration of *BsuSpoIIE*<sub>457-827</sub> is present in peak A with a major concentration of the fragment in peak B.

The IMAC purification A<sub>280</sub> chromatogram (Figure 3-22A), showed two peaks eluting at around 116, and 164 mM imidazole. The first peak exhibited a strong yellow colour. SDS-PAGE analysis of fractions from these peaks shows that the peak eluting at 116 mM imidazole contained a high concentration of an unknown 31 kDa protein. Fractions from the peak eluting at 164 mM imidazole showed a high concentration of protein with a mobility similar to the 45 kDa protein marker, consistent with the 41.3 kDa molecular mass of the *BsuSpoIIE*<sub>457-827</sub> fragment. Fractions in the second peak contained only small amounts of the 31 kDa contaminant and so the two species were easily separated. Fractions containing the target protein fragment were concentrated

from 16.5 to 3 ml by centrifugation in 10K MWCO Amicon concentration cells. The protein was then applied to an S200 16/60 gel filtration column in 50 mM Tris-HCl, 100 mM NaCl, 2 mM DTT at pH 8.0 running buffer (Figure 3-22B). The  $A_{280}$  chromatogram showed a large void volume peak and a smaller peak eluting at 82 ml. SDS-PAGE analysis showed a small concentration of target protein in the void volume with a much larger concentration present in the peak eluting at 82 ml.

*BsuSpoIIE*<sub>457-827</sub> containing fractions eluting at 82 ml, with an estimated molecular weight around 39 kDa, were isolated and concentrated to around 20 mg/ml. Protein crystallisation screening was carried out on this proteins and on a stock of *BsuSpoIIE*<sub>457-827</sub>, provided by Niels Bradshaw. The screens applied were Hampton I/II, JCSG and PEG/ion (Hampton) with protein concentrations at 8 and 15 mg/ml. Condition F7 in the JCSG screen (0.8 M succinic acid at pH 7.0) showed promising microcrystals for both protein isolations at both concentrations after one night. Optimisation was carried out on this condition, varying protein and succinic acid concentration, pH and presence of additives, yielded larger crystals at protein concentrations around 6 mg/ml. X-ray analysis indicated that these were not salt crystals but nor was clear protein crystalline diffraction seen. Further optimization yielded no improvement.

An analogous purification of *GstIIE*<sub>457-826</sub> was carried out. The size exclusion chromatogram showed the *GstIIE*<sub>457-826</sub> fragment to elute at 75 ml (Figure 3-23). This elution volume corresponds to a molecular weight estimate around 47 kDa. A void volume peak was also observed but did not contain any *GstIIE*<sub>457-826</sub>, as analysed by SDS-PAGE. Peak fractions were combined and concentrated to around 20 mg/ml before storage at -80°C. This protein was used to carry out crystallisation screening which returned no optimisable hits.



**Figure 3-23: Size exclusion chromatography of *GstIIE*<sub>457-826</sub>.**

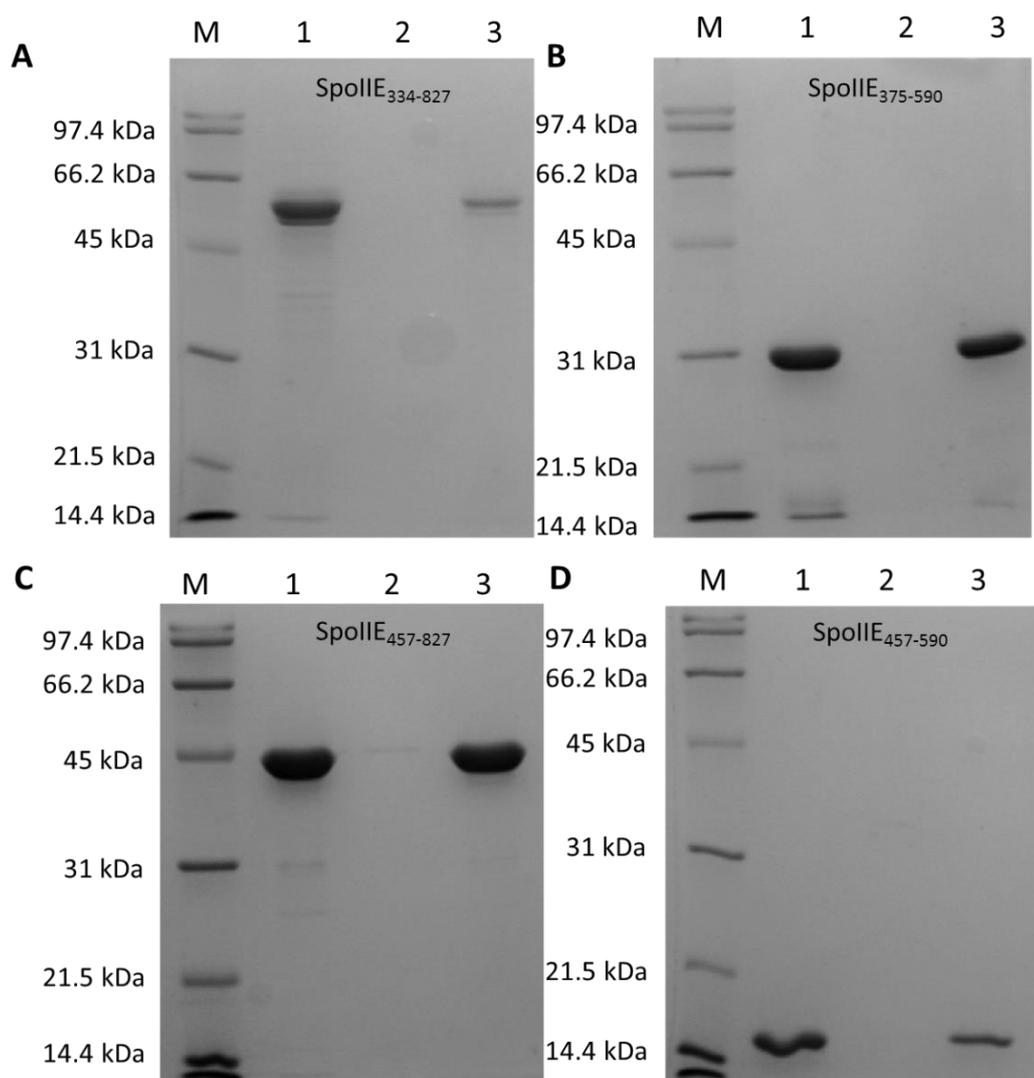
The A280 chromatogram of the elution peaks is shown. Purification was carried out using an S200 16/60 gel filtration column.

### 3.2.6 The effect of $Mn^{2+}$ on SpoIIE Domain II fragments.

A fragment of *Bsu*SpoIIE consisting of residues 320-827 was observed to precipitate on addition of  $MnCl_2$  (Niels Bradshaw, Harvard, Personal Communication). This was measured by centrifugation and resuspension before SDS-PAGE analysis.

A similar experiment was carried out to analyse how the available domain II containing fragments of *Bsu*SpoIIE behave on addition of  $Mn^{2+}$ . A final concentration of 5 mM  $MnCl_2$  was added 50  $\mu$ l samples of the SpoIIE<sub>334-827</sub>, SpoIIE<sub>375-590</sub> and SpoIIE<sub>457-827</sub> fragments at 1 mg/ml. A newly produced fragment, *Bsu*SpoIIE<sub>457-590</sub>, not previously discussed in this work, was also examined in this way. A visible flocculent precipitate was observed in the cases of SpoIIE<sub>457-590</sub>, SpoIIE<sub>375-590</sub> and SpoIIE<sub>457-827</sub>, but not for SpoIIE<sub>334-827</sub>. These samples were then centrifuged using a benchtop microfuge at 16,000 x g for 5 minutes. The supernatant was removed and any pellet was resuspended by vortexing for 30 seconds in fresh  $MnCl_2$ -free buffer. 10  $\mu$ l samples of the resuspended pellets were analysed by SDS-PAGE (Figure 3-24).

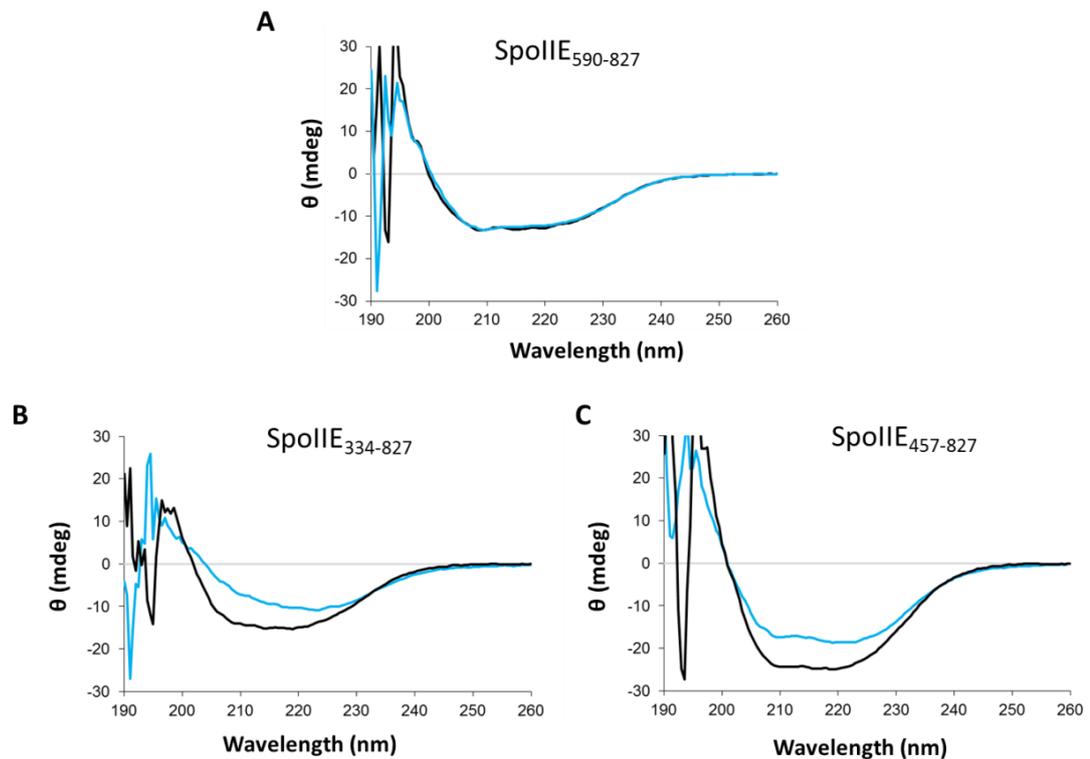
This experiment showed that SpoIIE<sub>334-827</sub> exhibits the lowest proportion of precipitation on addition of Mn<sup>2+</sup>. In contrast, the other three fragments, which are all monomeric in solution, show a higher proportion of precipitated protein. This proportion could be up to 100% in the cases of SpoIIE<sub>375-590</sub> and SpoIIE<sub>457-827</sub>.



**Figure 3-24: SDS-PAGE analysis SpoIIE fragment pelleting on addition of MnCl<sub>2</sub>.** M; LMW range markers (Bio-rad), lane 1; 1 mg/ml SpoIIE - no Mn<sup>2+</sup> or centrifugation, lane 2; 1 mg/ml SpoIIE - no Mn<sup>2+</sup> with centrifugation, lane 3; 1 mg/ml SpoIIE - with 5 mM Mn<sup>2+</sup> and centrifugation. (A) SpoIIE<sub>334-827</sub>. (B) SpoIIE<sub>375-590</sub>. (C) SpoIIE<sub>457-827</sub>. (D) SpoIIE<sub>457-590</sub>.

The circular dichroism (CD) spectra of SpoIIE fragments which contain elements of domain II were also shown to change on addition of MnCl<sub>2</sub>. Samples of SpoIIE<sub>590-827</sub>, SpoIIE<sub>334-827</sub>, SpoIIE<sub>375-590</sub> and SpoIIE<sub>457-827</sub> were analysed by CD, as described in section 2.7.4. A second set of data was then

recorded for these fragments with 2 mM added  $\text{MnCl}_2$ . Comparison of the spectra shows (Figure 3-25) a reduction in secondary structure signal is seen on addition on  $\text{Mn}^{2+}$  for fragments containing elements of SpoIIIE domain II but not for SpoIIIE<sub>590-827</sub>, which consists of only the phosphatase domain. This suggests a change in conformation of the protein induced by the addition of  $\text{Mn}^{2+}$ , but this, strikingly, is caused by domain II rather than the  $\text{Mn}^{2+}$  binding phosphatase domain. It is unclear whether the detected change in conformation is due to protein unfolding or simply a reorganisation of secondary structure.



**Figure 3-25: Circular dichroism spectra of SpoIIIE fragments on addition of  $\text{Mn}^{2+}$ .** Examined fragments: (a) SpoIIIE<sub>590-827</sub> (B) SpoIIIE<sub>334-827</sub> (C) SpoIIIE<sub>457-827</sub>. Black traces are recorded for protein solutions at 0.2 mg/ml with no added  $\text{MnCl}_2$ . Blue traces are spectra recorded in the presence of 2 mM  $\text{MnCl}_2$ .

### 3.3 Summary and Discussion

The isolation and characterisation of three novel fragments of SpoIIE from *B. subtilis* was carried out. These fragments were produced with the aim of carrying out structural studies of SpoIIE domain II. SpoIIE<sub>375-590</sub> is a monomer in solution but size exclusion chromatography shows some tendency towards aggregate formation. The difference in the oligomeric behaviour observed for these new fragments; SpoIIE<sub>375-590</sub>, SpoIIE<sub>334-827</sub> and SpoIIE<sub>457-827</sub>, suggests that the region between residues 334 and 457 of SpoIIE induces oligomerisation. This is supported by the observed behaviour of previously studied fragments, H1 (SpoIIE<sub>412-827</sub>) and B2 (SpoIIE<sub>375-827</sub>). The secondary structure in the region between residues 334 and 375 has been predicted (Appendix I) to consist of a long  $\alpha$ -helix and the primary sequence shows a roughly even distribution between hydrophobic and hydrophilic residues. That SpoIIE<sub>457-827</sub> is monomeric could suggest that 457-590 is a stable and soluble subdomain of SpoIIE. Recent attempts to isolate this fragment on its own, carried out jointly by the author and Vladimir Levnikov (unpublished), have been successful and it is thought that it might prove to be a good reagent for structural elucidation by NMR, being only 14 kDa and monomeric.

Analogous fragments to SpoIIE<sub>334-827</sub> and SpoIIE<sub>457-827</sub> were isolated from *Geobacillus stearothermophilus*. The *Gst*SpoIIE<sub>457-826</sub> fragment showed similar oligomeric character to *Bsu*SpoIIE<sub>457-827</sub>. Conversely, the oligomeric character of *Bsu*SpoIIE<sub>334-827</sub> and *Gst*SpoIIE<sub>335-826</sub> showed some differences. *Gst*SpoIIE<sub>335-826</sub> did not show the tendency to favour very high MW species exhibited by *Bsu*SpoIIE<sub>334-827</sub>. Instead, this fragment seems to exist in equilibrium between monomeric and dimeric states.

The presence of  $Mn^{2+}$  seems to have a pronounced effect on domain II. Fragments which contain elements of domain II showed a change in CD spectra and a tendency to precipitate on addition of millimolar concentrations of  $MnCl_2$ . When the sequence of SpoIIE is analysed using Metal Detector, an online tool for detecting cysteine and histidine residues which form transition metal binding sites, the 4 conserved cysteine residues (399, 402, 408 and 446) are identified as forming metal binding sites. It is unlikely that the observed phenomena

caused by addition of  $Mn^{2+}$  to SpoIIE domain II fragments is due to this putative metal binding site, as it occurs whether or not these residues are present. Attempts at protein crystallisation with all fragments discussed in this chapter were unsuccessful.

# Chapter 4: Activity-based and Structural Investigations into the Role of SpoIIE as a PP2C Phosphatase

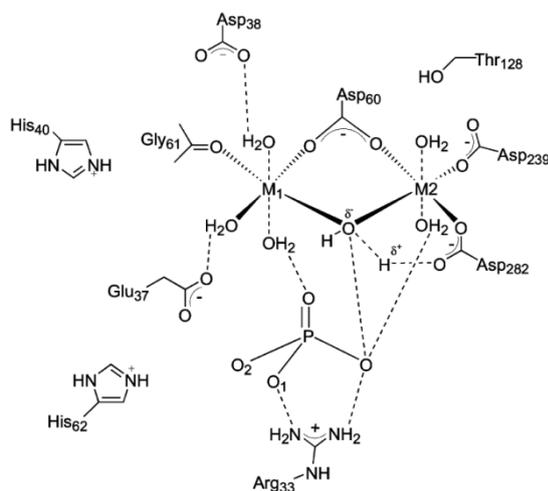
## 4.1 Introduction

### 4.1.1 Serine/Threonine Phosphatases

Signal transduction by reversible phosphorylation is critical in the regulation of cell function throughout nature [165]. Phosphorylation is carried out by protein kinases on serine, threonine or tyrosine residues. The induced changes in conformation and function of the substrate allows transduction of a signal. This method of signalling is made more powerful by its reversibility, which allows fine control of downstream effects by competitive dephosphorylation by protein phosphatases. The phosphatases which act to dephosphorylate phosphotyrosine in proteins, belonging to the Protein Tyrosine Phosphatase superfamily (PTP), differ from those which act upon phosphorylated serine and threonine residues, which belong to the PPP and PPM families [201]. The most prevalent ser/thr phosphatases in eukaryotes belong to the PPP family and are separated into PP1, PP2A and PP2B categories while the PPM family mainly consists of PP2C phosphatases. Each family possesses strong homology suggesting conserved tertiary structures and catalytic mechanisms. Diversity between phosphatases of the same family with respect to substrate and activity is conferred by regulatory proteins and/or subdomains.

Although the PP2C phosphatases of the PPM family possess no sequence homology with the PPP family, the two families show remarkable structural similarity [166]. The PP2C phosphatase domain has an  $\alpha\beta\alpha$ -structure where two antiparallel  $\beta$ -sheets are apposed to one another and flanked by  $\alpha$ -helices. The core PP2C phosphatase domain consists of 6  $\alpha$ -helices and 11  $\beta$ -strands. The active site is positioned at the top of the central sandwich and contains two  $Mn^{2+}$  ions and 4 invariant aspartate residues (Figure 4-1). The structure of the human PP2C $\alpha$  phosphatase reveals 3 water molecules interacting with the two metal ions. The crystal was grown in 50 mM potassium phosphate and a

molecule of inorganic phosphate ( $P_i$ ) was observed bound in the active site. As phosphate is a reaction product its presence provides insight into the active site configuration after hydrolysis. The interactions suggest that the reaction mechanism could be initiated by nucleophilic attack by a water molecule activated by interaction with both  $Mn^{2+}$  ions.



**Figure 4-1: The phosphate containing active site of human PP2C $\alpha$  phosphatase.**

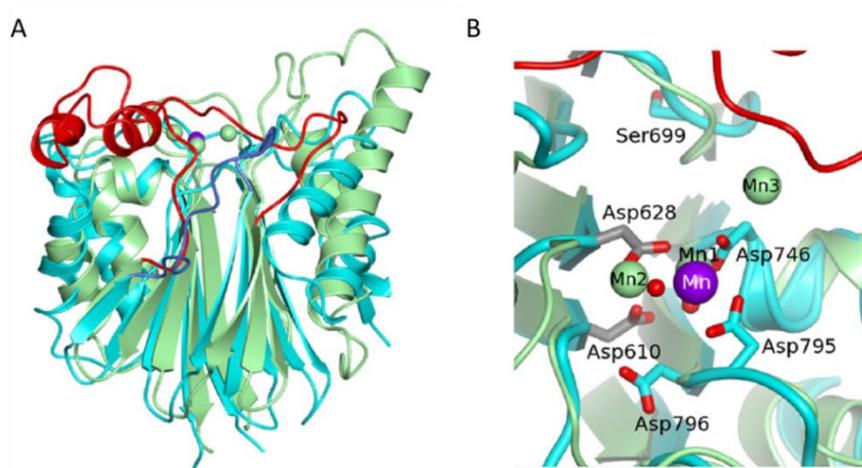
As presented by Jackson et al., 2003 [202]. Structure originally presented by Das et al., 1996 [166]. PDB ID: 1A6Q. The active site contains two  $Mn^{2+}$  ions (M1 and M2) held by conserved Asp, Glu and Gly residues. Electrostatic interactions with a bound phosphate group occur via 3 of the 6 water molecules present. One of these water molecules interacts with both M1 and M2 and also undergoes by a polarising activation by Asp<sub>282</sub>.

#### 4.1.2 The SpoIIE PP2C domain structure

The X-ray structure of SpoIIE<sub>590-827</sub> had been solved by Vladimir Levдикov after crystallisation from Tris-buffered sodium citrate solutions at pH 8.5 (PDB ID: 3T9Q, 3T91) [170]. The structure reveals a domain swapped dimer in which the  $\beta_1$  -  $\beta_3$  strands of each subunit form part of the central  $\beta$  sandwich of the other subunit (Figure 4-3A). The active sites are situated at opposite ends of the dimer. In this structure, each active site contains a single  $Mn^{2+}$  ion. Additional electron density observed between the two subunits was observed into which a molecule of aldohexose was modelled.

In [170], a structural search for similar domains yielded human PP2C $\alpha$  as well as the bacterial phosphatases: PstP from *Mycobacterium tuberculosis* [203],

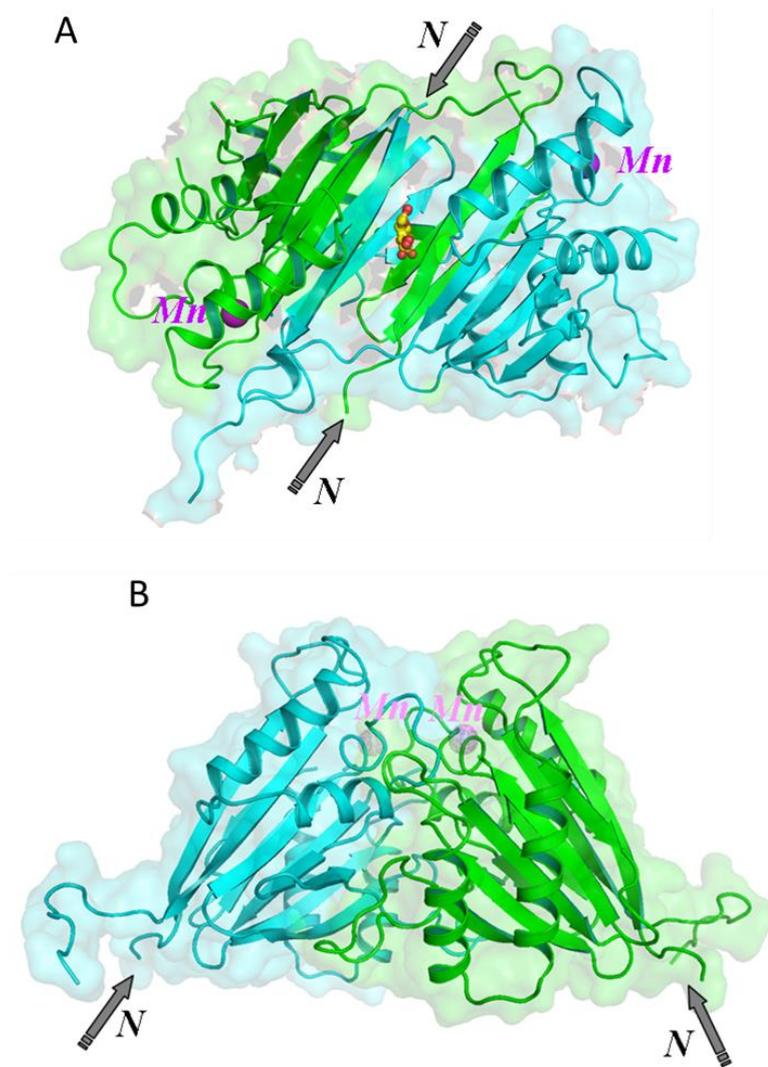
MspP from *Mycobacterium smegmatis* [204], STP from *Streptococcus agalactiae* [205] and tPphA from *Thermosynechococcus elongates* [206]. Comparison of the MtPstP (PDB ID: 1TXO) and SpoIIE structures shows a close overlay of the  $\beta$ -sandwich cores, however significant differences can be seen in the surrounding regions (Figure 4-2). The bacterial PP2C $\alpha$  phosphatases of known structure each possess 3 active site manganese ions. While the binding site for the third manganese ion is largely formed by residues in the  $\beta$ 7- $\beta$ 8 region, which is truncated in SpoIIE, the binding site for a second manganese ion consisting of D628, G629 and D610 is conserved. The lack of a second bound manganese atom in the structure could be for a combination of reasons. Firstly, the domain swapped dimer configuration of the structure, if not biologically relevant, could diminish the Mn<sup>2+</sup> affinity of the second metal binding site. One key observation in support of this is that D628, G629 and D610 all lie on the domain swapped portion of each monomer. Secondly, the crystallisation conditions from which the crystal was grown contained 200 mM sodium citrate which can chelate manganese and compete with the PP2C phosphatase.



**Figure 4-2: Structural comparisons between SpoIIE<sub>590-827</sub> (PDB ID: 3T9Q) and MtPstP (PDB ID: 1TXO) PP2C $\alpha$  phosphatases.**

(A) Overlay of a reconstituted (non-domain swapped) SpoIIE PP2C monomer (cyan) with MtPstP (light green). Manganese ions are indicated as balls coloured purple for SpoIIE and light green for MtPstP. The  $\beta$ 7- $\beta$ 8 segment is truncated in SpoIIE (blue) in comparison to MtPstP (red), this long loop forms part of the third metal binding site which is not present in SpoIIE. (B) Active site following overlay as in (A). The manganese atom present in SpoIIE (Mn) overlays with Mn1 from MtPstP. Residues Asp610 and Asp628, required for binding of a second manganese atom are indicated in grey. The Mn3 site of MtPstP is formed by the  $\beta$ 7- $\beta$ 8 loop and Asp118 which aligns with SpoIIE Ser699. The latter residue points away from the active site in SpoIIE indicating that it does not play a role in metal binding.

A second structure of SpoIIE<sub>590-827</sub> has recently been solved, by Vladimir Levnikov, in which there is no domain swapping and a different dimer is formed (unpublished). In this structure no manganese ions are seen in the active sites which point towards, and occlude, each other (Figure 4-3B).



**Figure 4-3: Structures of the SpoIIE<sub>590-827</sub> fragment.**

PDB IDs: (A) 3T9Q and (B) not deposited. Separate subunits are shown in light blue and green. The N-termini of each subunit are labelled. (A) The domain swapped dimer. The active sites show by the presence of a single Mn atom (purple). The modelled aldohexose sugar is shown in yellow (carbon) and red (oxygen). (B) The occluded active site dimer. No Mn atoms are bound in the active sites, which are indicated by transparent labels.

### 4.1.3 The SpoIIE:SpoIIAA~P interaction

A three-dimensional structure of a PP2C protein phosphatase and its phosphoprotein substrate has, to the best of this author's knowledge, not been reported. The structures of the PP2C domain of SpoIIE [170] and SpoIIAA~P have been determined [123]. Attempts to generate a convincing model of the SpoIIE:SpoIIAA~P interaction have not so far been successful. As described in Levdikov et al., 2012 [170], the SpoIIE PP2C and human PP2C $\alpha$  structures were overlaid by least-squares superposition applied to common secondary structure elements. Meanwhile, the phosphate species in the human PP2C $\alpha$  active site was superposed onto the phosphate of the phosphoserine in the *Bsph*SpoIIAA~P structure. Of the three possible conformations in which the SpoIIAA~P phosphate group O atoms are precisely overlaid with those in the human PP2C structure, two lead to large steric clashes between the SpoIIE and SpoIIAA domains. The remaining conformation shows no contacts between the two proteins. Rigid body energy minimization was attempted by Vladimir Levdikov but without success. Possible explanations are:

- (i) The use of a structure from an orthologous rather than native substrate - *Bsph*SpoIIAA~P is a poor substrate for *Bsu*SpoIIE.
- (ii) Many of the loops surrounding the active site of SpoIIE are disordered in the crystal structure and these loops may become ordered upon complex formation.
- (iii) Formation of the SpoIIE:SpoIIAA complex could require conformational change in either or both proteins - related to (ii).

## 4.2 Methods

### 4.2.2 Native PAGE Phosphatase State Assay

The phosphatase assays used in this work take advantage of the difference in mobility observed between SpoIIAA in its phosphorylated and unphosphorylated forms [188]. Except where otherwise stated, a 20  $\mu$ l phosphatase reaction is set up in 20 mM Tris-HCl 100 mM NaCl at pH 8.0 which contains:

- 1  $\mu$ l of 10.2 mg/ml SpoIIAA~P; giving a final concentration of 39  $\mu$ M.
- A quantity of the assayed SpoIIE fragment dependent on the required enzyme-to-substrate ratio.
- 1  $\mu$ l of 200 mM MnCl<sub>2</sub>.

The duration of the reaction can be varied. When finished, 10  $\mu$ l of each reaction is analysed by native PAGE on a 7.5% acrylamide gel, as described in section 2.7.3.2.

### 4.2.3 Generation of Transition State Analogues

The AlF<sub>4</sub><sup>-</sup> and MgF<sub>3</sub> transition state analogues are spontaneously generated at a concentration of 10 mM *in situ* by mixing of an equal volume of 20 mM AlCl<sub>3</sub> or MgCl<sub>2</sub> and a three- and four-fold molar excess of NaF, respectively.

### 4.2.4 Thermafluor Assays

Thermafluor thermal melting assays use hydrophobic dyes to measure protein unfolding as a function of temperature [207]. The dyes exhibit increased fluorescence when bound to exposed hydrophobic patches on proteins. In analysis of proteins, as the temperature rises stepwise from 25°C to 95°C an increase in observed fluorescence can be seen at the melting temperature, T<sub>m</sub>, as the hydrophobic dye molecules bind to the exposed core of the denaturing

sample. After melting, the fluorescence should decrease as the temperature increases due to thermal displacement of bound dye molecules.

In this work 0.5 mg/ml SpoIIIE<sub>590-827</sub> was analysed in a buffer of 20 mM Tris-HCl, 100 mM at pH 8.25. The dye used was SYPROOrange (Sigma-Aldrich) added at a 2000-fold dilution from the commercial stock, as per manufacturer's instructions. The fluorescence of the dye was probed using excitation and emission wavelengths of 517 nm and 585 nm, respectively. Temperature and fluorescence were controlled and recorded, respectively, at intervals of 1°C between 25°C and 95°C using a Mx3005P qPCR system and MxPro QPCR software.

#### 4.2.5 Silver Staining of PAGE gels

Silver staining of PAGE gels provides much more sensitive protein visualisation than Coomassie staining [208]. In the process, silver ions are bound by carboxylate (Asp/Glu) and sulfhydryl (Cys) groups on amino acid side chains. The silver is then reduced to free metallic silver which is visible.

The gel to be stained is first washed with a solution of 50% (v/v) methanol and 10% (v/v) acetic acid for 30 minutes. This is followed by 15 minute incubation in a solution of 5% (v/v) methanol. The gel is washed by 3 sequential incubations in milliQ-H<sub>2</sub>O for 5 minutes. A 2 minute incubation in 0.2 g/l sodium thiosulphate is followed by another wash with H<sub>2</sub>O for 3 x 30 seconds. The gel is then incubated with 2 g/l silver nitrate for 25 minutes and then washed with water for 3 x 1 minute. A solution of 30 g/l sodium carbonate, 0.004 g/l sodium thiosulphate and 0.2% (v/v) formaldehyde is then applied for 8 minutes. A wash in 14 g/l EDTA over 10 minutes is followed by transfer to H<sub>2</sub>O. Protein containing bands will be stained orange and the gel will be otherwise clear.

## 4.3 Results

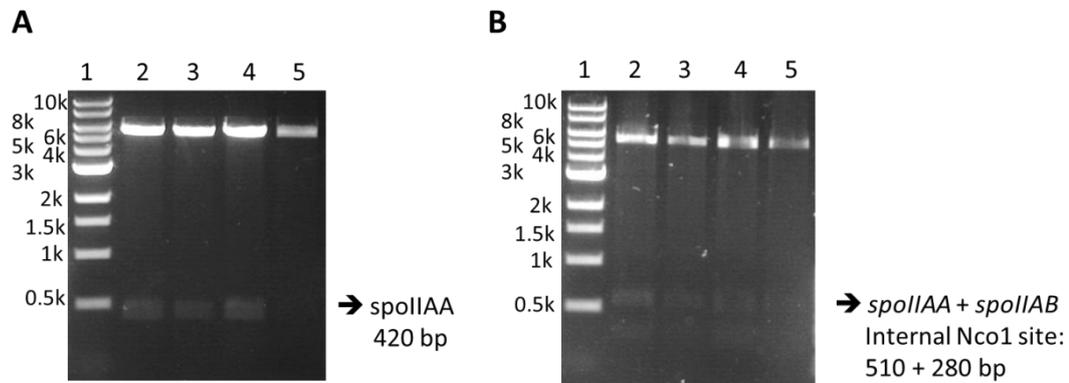
### 4.3.1 Cloning, Expression and Purification of SpoIIAA and SpoIIAA~P with SpoIIAB

The work described in this chapter, up to section 4.3.2.2, was carried out in association with Katie Jameson, a final year MChem project student. In order to isolate the native substrate of SpoIIE from *B. subtilis*, the genes *spoIIAA* and *spoIIAB* were co-overexpressed in *E. coli*. *spoIIAA* and *spoIIAB* are the first and second of three cistrons in the *spoIIA* locus. SpoIIAA, SpoIIAB and  $\sigma^F$  are co-expressed by the cell, under the transcriptional control of  $\sigma^H$  and Spo0A~P. *In vivo*, SpoIIAB phosphorylates SpoIIAA to form SpoIIAA~P and this activity can be transplanted into *E. coli*. Primers were designed to amplify *spoIIAA* and the two cistron segment *spoIIAA-spoIIAB* for cloning into the pET-YSB LIC3C vector. This cloning strategy aimed to produce SpoIIAA~P with an HRV 3C protease cleavable 6xHis affinity tag and native SpoIIAB to allow simple separation of the co-expressed species. The primer sequences are shown in Appendix I.

The coding sequences were amplified by PCR, as described in section 2.3.2, with a modified annealing temperature of 56°C. Expected product sizes were 340 base pairs for the *spoIIAA* only fragment and 750 base pairs for the *spoIIAA-spoIIAB* amplicon. The products of both PCRs were analysed by 1% agarose gel electrophoresis and showed amplicons of the expected size and good product specificity. Purification was carried out by PCR clean-up. T4 DNA polymerase treatment of both fragments and annealing with prepared pET-YSB LIC3C was carried out as described in sections 2.3.5 and 2.3.6. Annealed samples were used to transform competent *E. coli* XL-10 Gold cells. For each construct, plasmid DNA was extracted by miniprep as described in section 2.2. Extracted plasmids were identified as AA1 to AA4 for *spoIIAA* harbouring constructs and AAAB1 to AAAB4 for constructs harbouring the dual cistron sequence. Double digests using NcoI and NdeI were carried out on around 200 ng of each isolated plasmid and analysed by 1% agarose gel electrophoresis (Figure 4-4).

The expected insert band sizes were 420 bp for the AA construct and 510 + 280 bp for the AAAB constructs, due to an Nco1 cleavage site within *spoIIAB*.

Analysis of the test digests indicated that constructs AA1, AA2 and AA3 and AAAB1, AAAB2 and AAAB3 carried the desired inserts; neither AA4 nor AAAB4 exhibited the expected DNA fragment sizes, suggesting no insert was present.

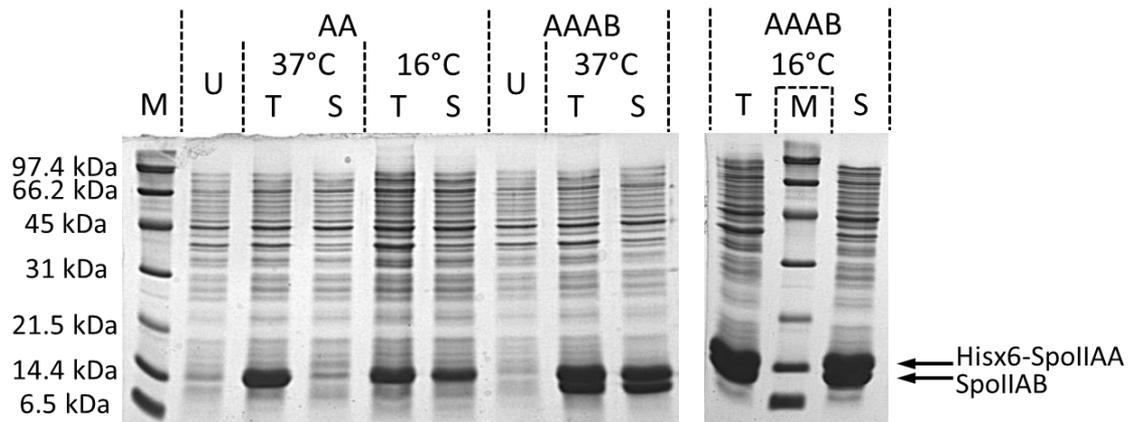


**Figure 4-4: Digests of putative *spoIIAA* and *spoIIAA-spoIIAB* containing plasmids.** 4 clones of each construct were analysed after double digest with *NcoI* and *NdeI*. (A) *SpoIIAA* production construct analysis: Lane 1; 1 kb DNA ladder, lanes 2-5; constructs AA1-AA4. Expected insert length: 420 bps. (B) *SpoIIAA+AB* production construct analysis: Lane 1; 1 kb DNA ladder, lanes 2-5; constructs AAAB1-AAAB4. *spoIIAB* has an internal *NcoI* site causing two insert bands. Expected insert band sizes: 510 and 280 bps.

The AA1-3 and AAAB1-3 constructs were submitted for DNA sequencing using T7 promoter specific primers (primer sequences in Appendix I). The results showed that the expected sequences were present in AA1, AA2 and AAAB1. The remaining constructs had acquired the following mutations: AA3; point mutation encoding G4E mutant (GGA to GAA), AAAB2; silent stop codon mutation (TAA to TGA), AAAB3; single base deletion resulting in frame shift mutation upstream of 6<sup>th</sup> *spoIIAA* codon.

The AA1 and AAAB1 plasmids were used to transform competent *E. coli* B834 (DE3) cells and the transformants plated onto LB agar supplemented with 30 µg/ml kanamycin. Single kanamycin resistant colonies were used to inoculate overnight LB cultures from which expression testing was carried out as described in section 2.4.4, varying the expression temperature. Cultures were grown to OD<sub>600</sub> = 0.6 at which point uninduced samples were collected before induction with IPTG (section 2.4.2). Induced cultures for each construct were grown at 16°C and 37°C for 22 and 4 hours respectively before harvesting and

lysis. The lysis buffer used was 50 mM Tris-HCl, 500 mM NaCl at pH 8.0. Total and soluble samples were taken after sonication and after centrifugation respectively and analysed by 15% SDS-PAGE (Figure 4-5) as described in section 2.6.3. The expected masses are SpoIIAA; 15.3 kDa and SpoIIAB; 16.3 kDa.



**Figure 4-5: SpoIIAA and SpoIIAA+SpoIIAB expression tests.**

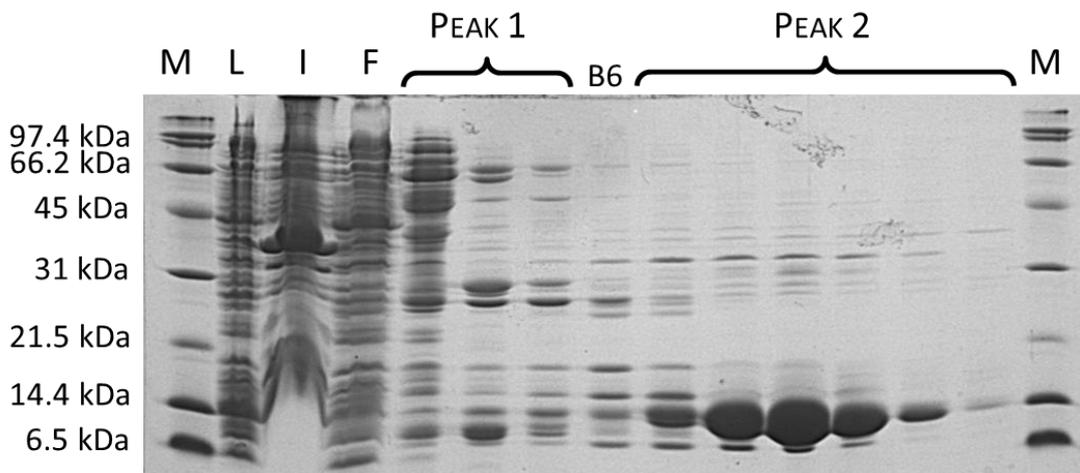
M indicates Bio-Rad broad range markers, marker masses are listed. U, T and S indicate uninduced, total and soluble samples, respectively.

The SDS gels show that SpoIIAA is expressed with around 60% solubility, estimated by band intensity, at 16°C but is insoluble when expressed at 37°C. Soluble expression of both SpoIIAA and SpoIIAB from the AAAB construct is exhibited at both 16°C and 37°C. The intensity of the SpoIIAA (SpoIIAA~P) bands in the total and soluble samples indicate a much higher solubility than is exhibited by the construct expressing SpoIIAA alone.

#### 4.3.2.1 Purification of SpoIIAA

4 flasks containing 500 ml LB media supplemented with 30 µg/ml kanamycin were inoculated with 10 ml of LB overnight culture of B834 (DE3) harbouring the pET-YSBLIC3C-SpoIIAA construct. The cultures were grown at 37°C to an OD<sub>600</sub> of 0.6 and protein production induced by addition of IPTG to a final concentration of 1 mM. Protein expression was carried out at 16°C for 22.5 hours. Cells were harvested by centrifugation and combined by resuspension in

a lysis and purification buffer of 50 mM Tris-HCl, 500 mM NaCl, 50 mM imidazole at pH 8.0. Resuspended cells were lysed by sonication and the insoluble materials removed by centrifugation at 30,600 x g for 30 minutes in an SS-34 rotor. The supernatant was loaded onto a 5 ml HiTrap (GE Healthcare) IMAC column charged with Ni<sup>2+</sup> using a peristaltic pump. After a 15 column volume wash with loading buffer the column was developed with a linear 50 – 500 mM gradient of imidazole in loading buffer using an AKTA purifier. 8 ml fractions were collected during the wash and 4ml fractions were collected after the start of the gradient. The absorbance of the eluate was monitored at wavelengths of 280 nm and 254 nm to analyse protein and DNA-containing species simultaneously. Two A<sub>280</sub> peaks were observed. The first was a large peak eluting at 130 mM imidazole which also showed a high A<sub>260</sub> indicating that the peak likely contained non-protein contaminants. The second peak which eluted at 220 mM imidazole was much smaller.



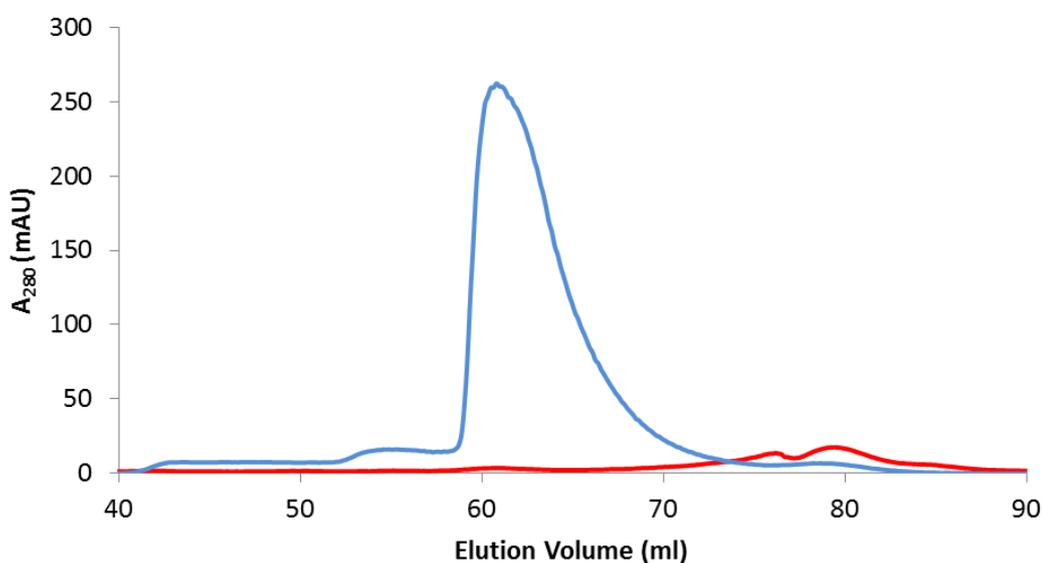
**Figure 4-6: Nickel IMAC purification of SpoIIAA.**

M indicates Bio-Rad broad range markers. L, I and F indicate samples of load solution, resuspended insoluble lysate and flowthrough, respectively. Fractions from both peaks described in the text are indicated. Fraction B6 lay in between peaks 1 and 2.

Samples from the load solution, resuspended insoluble lysate, flowthrough and peak fractions were analysed by 15% SDS-PAGE (Figure 4-6). The SDS-PAGE gels showed that little SpoIIAA was present in the flowthrough, suggesting tight binding to the column. The first eluted peak contained multiple protein species with no strong band matching the expected MW of the target protein. The

second peak had significant enrichment of SpoIIAA with some medium and low molecular weight contaminants. The low  $A_{280}$  measured for the SpoIIAA containing peak can be attributed to its very low extinction coefficient;  $\epsilon = 1615 \text{ M}^{-1} \text{ cm}^{-1}$ .

The fractions comprising peak 2 were combined and treated with HRV3C protease to cleave the 6xHis affinity tag, as described in section 2.5.2. The 6xHis cleavage was almost complete as analysed by SDS-PAGE. SpoIIAA eluted from a second IMAC purification step in a peak at 70 mM imidazole, HRV3C protease and some contaminants eluted at 140 mM imidazole (data not shown). The peak fractions containing SpoIIAA were pooled and concentrated to  $\sim 2$  mls and further purified by size exclusion chromatography using a HiLoad Superdex 75 PG gel filtration column (GE Healthcare) pre-equilibrated with a running buffer of 20 mM Tris-HCl, 150 mM NaCl at pH 8.0 (Figure 4-7). 4 ml fractions were collected. SpoIIAA eluted in one major peak after 78 ml. Peak fractions were pooled and concentrated to 5 mg/ml, analysed by native and SDS-PAGE, as shown later (Figure 4-9), snap frozen in liquid  $\text{N}_2$  and stored at  $-80^\circ\text{C}$ .



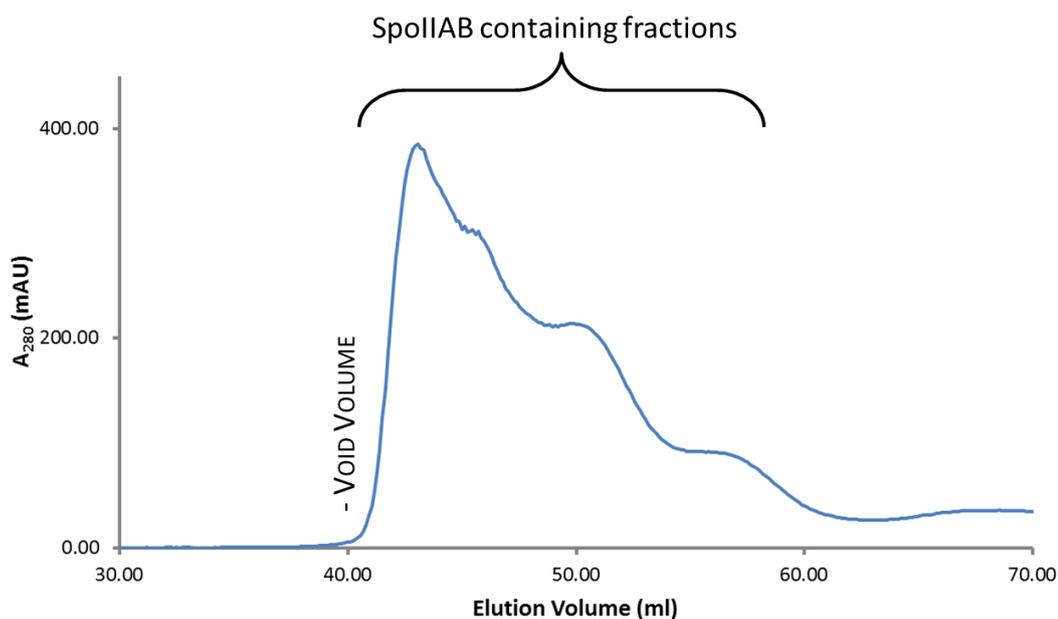
**Figure 4-7: Size exclusion chromatograms showing purifications of SpoIIAA (red) and SpoIIAA~P (blue) as separated by a Superdex 75 16/60 gel filtration column.**

The variation in elution volume between SpoIIAA (78 ml) and SpoIIAA~P (60 ml) suggests a difference in multimeric state between the two species. SpoIIAA is likely a monomer, while SpoIIAA~P is likely to be dimeric. The difference in maximal absorbance at 280 nm highlights the difference in expression yield between the two species, whose extinction coefficients are equal.

#### **4.3.2.2 Purification of SpoIIAA~P and SpoIIAB**

Purification of SpoIIAA~P was carried out in an analogous fashion to that of SpoIIAA. SpoIIAB did not bind to the HiTrap IMAC purification column and so was easily separated from SpoIIAA~P, which eluted from the column at an imidazole concentration of 225 mM. The 6xHis tag was cleaved using HRV3C protease. Subsequent purification, using a HiLoad Superdex 75 PG gel filtration column, showed SpoIIAA~P to elute after 60 ml; 18 ml earlier than SpoIIAA (Figure 4-7). This indicates a significant mobility difference between SpoIIAA and SpoIIAA~P, suggesting that SpoIIAA~P forms a dimer while SpoIIAA is monomeric, as reported earlier [209]. SpoIIAA~P was analysed by SDS and native PAGE (as shown later in Figure 4-9), concentrated to 102 mg/ml and aliquoted and stored at -80°C.

The SpoIIAB containing flowthrough from IMAC purification of SpoIIAA~P was purified by Ion Exchange Chromatography using a 5 ml Resource Q anion exchange column (GE Healthcare), as described in section 4.2.1. The running buffer was 50 mM Tris-HCl at pH 8.0 and an elution gradient between 0 M and 1 M NaCl was applied. SpoIIAB eluted at around 320 mM NaCl in the second half of a broad A<sub>280</sub> peak. Purification by size exclusion chromatography, using a HiLoad Superdex 75 PG gel filtration column with the running buffer used for purification of SpoIIAA and SpoIIAA~P, showed a number of high MW states of SpoIIAB eluting from immediately after the 40 ml void volume of the column to 60 ml (Figure 4-8). The AB containing fractions were combined and concentrated to 91 mg/ml before analysis by SDS and native PAGE (Figure 4-9) and storage at -80°C.

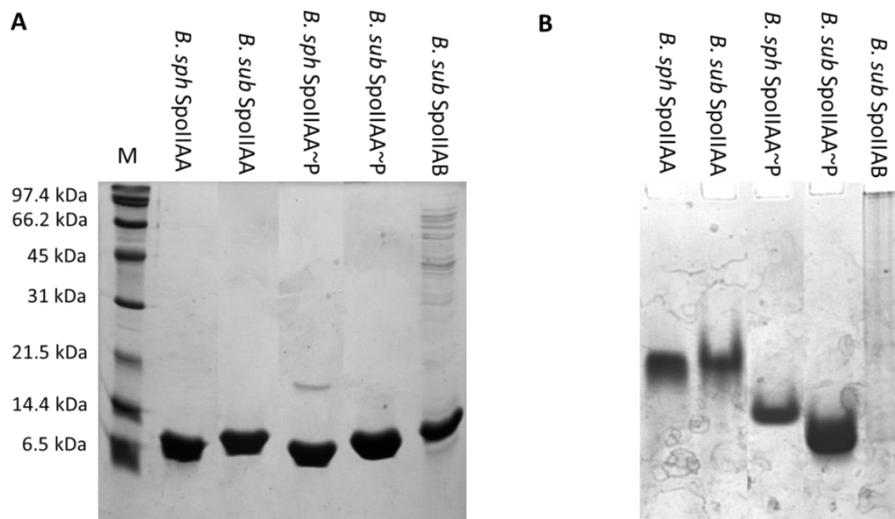


**Figure 4-8: Size exclusion chromatogram of SpoIIAB.**

The  $A_{280}$  trace is shown as a blue line. The void volume of the column (40 ml) is indicated. SpoIIAB seems to elute as a small number of high molecular weight species.

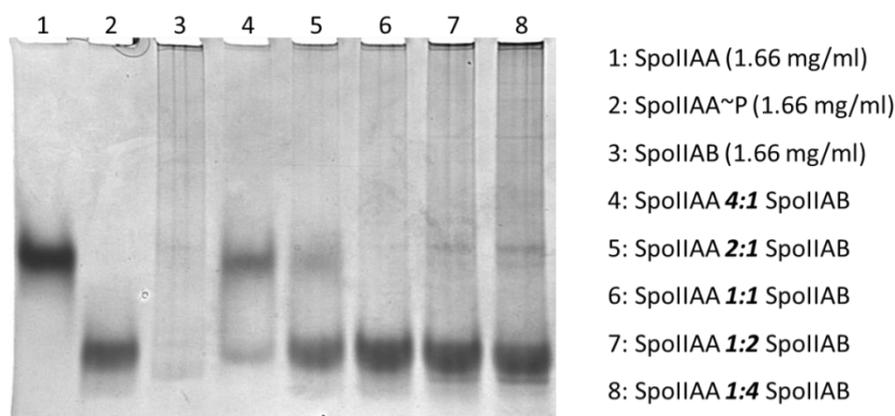
#### 4.3.2.3 Analysis of Purified SpoIIAA, SpoIIAA~P and SpoIIAB Proteins

Samples of purified SpoIIAA and SpoIIAB proteins, including the SpoIIAA analogue from *B. sphaericus*, were analysed by SDS and Native PAGE (Figure 4-9). Both SpoIIAA and SpoIIAA~P from *B. subtilis* were shown to be pure by both SDS and Native PAGE. A significant increase in native gel mobility is seen for SpoIIAA~P in comparison to SpoIIAA, which could be attributed to the additional negatively charged phosphate moiety. The final yield of isolated SpoIIAA~P (~100 mg) far exceeds that of SpoIIAA (~5 mg), agreeing with the observations made during expression testing. SpoIIAB was not as pure as the SpoIIAA species. Moreover, SpoIIAB does not enter the native polyacrylamide gels, this may be due to the high MW species observed during size exclusion.



**Figure 4-9: SDS- PAGE and Native PAGE analysis of SpoIIAA and SpoIIAB species from both *B. subtilis* and *B. sphaericus*.**

(A) 15% SDS-PAGE gel with indicated samples each containing approximately 20  $\mu\text{g}$  of protein. M indicates Bio-Rad Broad range markers. (B) 7.5% Native PAGE gel with indicated samples.



**Figure 4-10: Native PAGE analysis of SpoIIAB kinase activity assay.**

The contents of numbered lanes are given on the right. Molar ratios are shown.

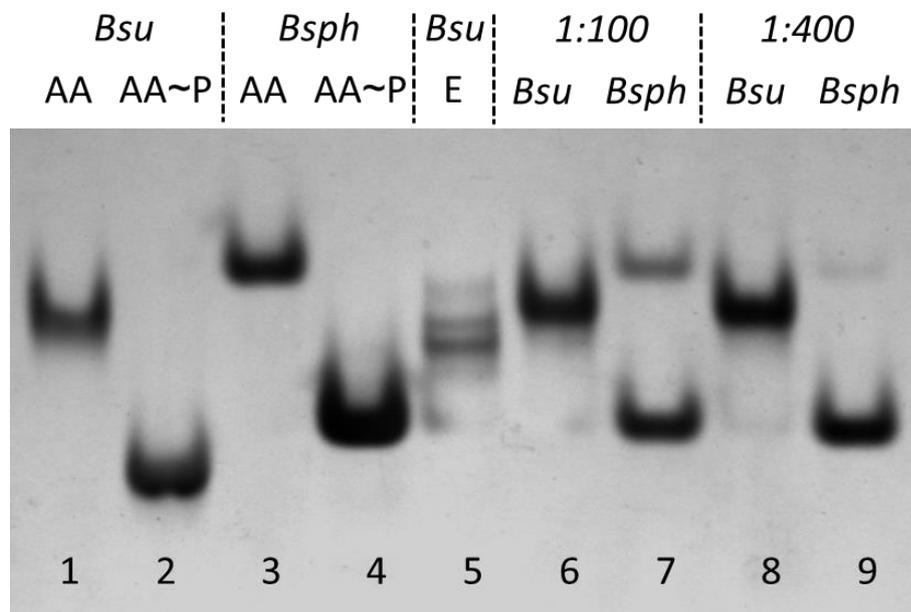
The oligomeric nature of the purified SpoIIAB might suggest some unfolding and aggregation which could affect the protein's enzymatic activity. The kinase activity of SpoIIAB against its cognate substrate, SpoIIAA, was analysed by a previously unused Native PAGE assay developed specifically for visualising the phosphorylation state of SpoIIAA. Varying ratios of SpoIIAA and SpoIIAB were incubated at room temperature for 45 minutes in 20 mM Tris-HCl, 100 mM NaCl at pH 8.0 with 10 mM ATP. The products were analysed by 7.5% native PAGE (Figure 4-10). The gel shows that the purified SpoIIAB is capable of phosphorylating SpoIIAA *in vitro*. Conversion of SpoIIAA to SpoIIAA~P appears

complete in reactions with an equimolar ratio or excess of SpoIIAB, both reactions with excess SpoIIAA show remaining substrate. The relative band intensities allow rough estimation of SpoIIAB activity, and that approximately  $5.5 \times 10^{-4}$  molecules of SpoIIAA are turned over per SpoIIAB molecule per second, this value is comparable to a previously published rate of  $8 \times 10^{-4} \text{ s}^{-1}$  [86].

### 4.3.2 The SpoIIAA~P Dephosphorylation Reaction

#### 4.3.2.1 Cognate vs. Orthologous Substrate

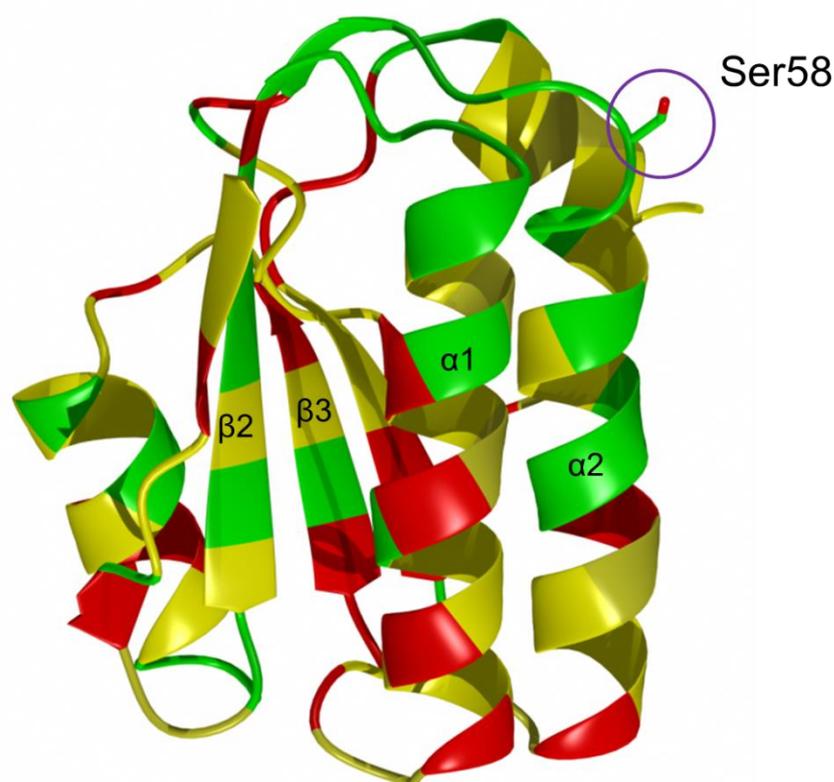
A Native PAGE phosphatase activity assay comparing the rate at which SpoIIE<sub>590-827</sub> dephosphorylates *Bsu*SpoIIAA~P and *Bsph*SpoIIAA~P was carried out as described in section 4.2.2. Reactions were carried out using 100 and 400 times molar excess of SpoIIAA~P over SpoIIE<sub>(590-827)</sub>. This experiment showed that the rate of *Bsu*SpoIIAA~P dephosphorylation far exceeds that of *Bsph*SpoIIAA~P. Full conversion of *Bsu*SpoIIAA~P to product is achieved at both enzyme to substrate ratios whereas partial dephosphorylation of *Bsph*SpoIIAA~P occurs in both cases (Figure 4-11).



**Figure 4-11: Comparison of SpoIIE catalysed dephosphorylation of SpoIIAA~P from *B. subtilis* and *B. sphaericus*.**

Lane 1; *Bsu*SpoIIAA (16 µg), lane 2; *Bsu*SpoIIAA~P (16 µg), lane 3; *Bsph*SpoIIAA (16 µg), lane 4; *Bsph*SpoIIAA~P (16 µg), lane 5; SpoIIE<sub>590-827</sub> (20 µg), lane 6; *Bsu*SpoIIAA~P (20 µg) + SpoIIE<sub>590-827</sub> (0.4 µg), lane 7; *Bsph*SpoIIAA~P (20 µg) + SpoIIE<sub>590-827</sub> (0.4 µg), lane 8; *Bsu*SpoIIAA~P (20 µg) + SpoIIE<sub>590-827</sub> (0.1 µg), lane 9; *Bsph*SpoIIAA~P (20 µg) + SpoIIE<sub>590-827</sub> (0.1 µg).

The efficiency of the dephosphorylation reaction catalysed by *Bsu*SpoIIE<sub>590-827</sub> is much higher for *Bsu*SpoIIAA~P than it is for *Bsph*SpoIIAA~P. The two SpoIIAA proteins possess 35% sequence identity and a further 30% sequence similarity. When the *Bsph*SpoIIAA~P is coloured by sequence conservation it can be seen that the conserved residues are strongly clustered around the loops between  $\beta 2$  and  $\alpha 1$  and  $\beta 3$  and  $\alpha 2$  which are in the region of the phosphorylatable serine (Figure 4-12). Investigations using NMR into the likely dimer interface of SpoIIAA have highlighted these features along with helix  $\alpha 3$  as being directly involved in dimerisation. It is thought that both the addition of a highly charged group and small conformational changes which occur in SpoIIAA on phosphorylation determine the higher affinity of SpoIIE for SpoIIAA~P relative to SpoIIAA [209]. That the activity of *Bsu*SpoIIE on *Bsph*SpoIIAA~P is so much lower compared to *Bsu*SpoIIAA~P suggests that the conformational element of the interaction is not adequately mimicked by the orthologue.



**Figure 4-12: Sequence conservation of SpoIIAA from *B. subtilis* and *B. sphaericus*.** PDB ID: 1H4Z. Sequence alignment and formatting carried out using CCP4MG. The *B. sphaericus* SpoIIAA structure is shown as ribbons coloured by sequence conservation. Conserved, chemically similar and unconserved residues are coloured in green, yellow and red, respectively. The side chain of the phosphorylatable serine (Ser58) is shown as cylinders and indicated with a purple circle.

In order to investigate the enzymatic characteristics of SpoIIE further, *Bsph*SpoIIAA was set aside in favour of the native substrate, *Bsu*SpoIIAA. Henceforth, unless specified all references to SpoIIAA species are to *Bsu*SpoIIAA.

#### 4.3.2.2 Metal Dependence of the SpoIIE PP2C domain.

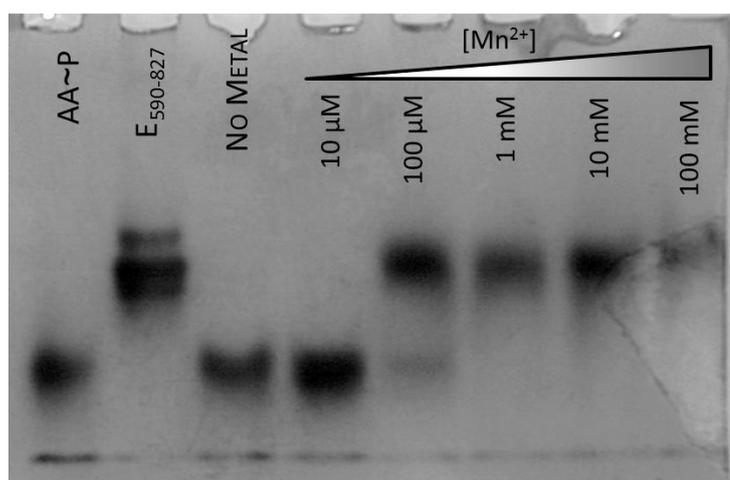
The PP2C family of protein Serine/Threonine phosphatases are widely understood to be manganese or magnesium dependent enzymes. The phosphatase activity of SpoIIE was initially followed in the presence of 2 mM  $\text{MnCl}_2$  [82]. In order to confirm this dependence, the phosphatase activity of SpoIIE<sub>590-827</sub> was analysed in the presence of six divalent metal cation chloride salts. A native PAGE phosphatase assay was carried out, as described in section 4.2.2.1, with a 1:200 molar excess of SpoIIAA~P to SpoIIE<sub>590-827</sub>. The various metal chlorides were added at a concentration of 10 mM and reactions were allowed to proceed for 60 minutes. Of the divalent cations added, only  $\text{Mn}^{2+}$  supported observable activity in dephosphorylation of SpoIIAA~P as judged by the decrease in substrate mobility (Figure 4-13).



**Figure 4-13: The metal dependence of SpoIIE<sub>590-827</sub> phosphatase activity *in vitro*.** SpoIIAA~P, SpoIIAA and SpoIIE<sub>590-827</sub> are shown, as indicated, on the left hand side. Reactions containing no added metal, 10 mM  $\text{MnCl}_2$ , 10 mM  $\text{MgCl}_2$ , 10 mM  $\text{CaCl}_2$ , 10 mM  $\text{CoCl}_2$ , 10 mM  $\text{CuCl}_2$  and 10 mM  $\text{ZnCl}_2$ .

It was also noted that in the absence of metal ions, the dephosphorylation reaction does not occur. This indicates that  $Mn^{2+}$  ions do not carry through the purification steps with the SpoIIE<sub>590-827</sub>, suggesting their binding is weak. This was confirmed by atomic absorption spectrometry analysis of the Mn content of a sample of SpoIIE<sub>590-827</sub> incubated with 10 mM  $MnCl_2$  for 20 minutes followed by passage through a Superdex S200 10/300 GL gel filtration column in an  $MnCl_2$  free buffer. No retained  $Mn^{2+}$  was observed when measured absorbance was compared to a  $Mn^{2+}$  standard curve (data not shown).

Failure to retain  $Mn^{2+}$  may indicate a regulatory mechanism in which the affinity of SpoIIE for  $Mn^{2+}$  is controlled. Chelation by citrate notwithstanding, low affinity binding would also explain the lack of observed metal ions in the crystal structure of SpoIIE<sub>590-827</sub>. To determine how variation in  $Mn^{2+}$  concentration affects the activity of SpoIIE<sub>590-827</sub>, dephosphorylation of SpoIIAA~P was measured over 45 minutes at an enzyme to substrate ratio of 1:100 with  $MnCl_2$  concentrations ranging between 10  $\mu$ M and 100 mM (Figure 4-14).



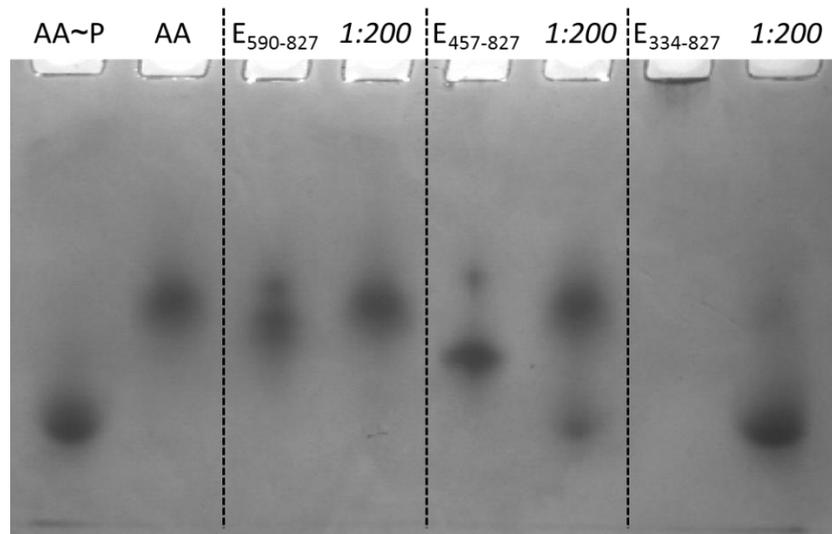
**Figure 4-14: Manganese concentration dependence of SpoIIE<sub>590-827</sub> activity *in vitro*.** SpoIIAA~P and SpoIIE<sub>590-827</sub> are shown, as indicated, on the left hand side. Reactions containing no added metal, 10  $\mu$ M, 100  $\mu$ M, 1 mM, 10 mM and 100 mM  $MnCl_2$  are shown to the right.

The lowest concentration of  $Mn^{2+}$  to allow full conversion of SpoIIAA~P to SpoIIAA was 1 mM. A large degree of conversion was observed at a concentration of 100  $\mu$ M but none at 10  $\mu$ M. This suggests that the SpoIIE<sub>590-827</sub>

active site is incapable of binding sufficient  $Mn^{2+}$  at 10  $\mu M$  to support dephosphorylation.

#### 4.3.2.3 Phosphatase activity of SpoII E fragments.

The *in vitro* phosphatase activity observed for the SpoII E<sub>590-827</sub> fragment shows clearly that the PP2C domain fragment is capable of dephosphorylating SpoII AA~P in the presence of  $Mn^{2+}$ . In the assays described above, the  $Mn^{2+}$  concentration greatly exceeds that present in a *B. subtilis* cell. However, the level or mechanism of regulation of SpoII E phosphatase activity cannot be determined by observing only the PP2C domain's behaviour. Genetic experiments indicate that the SpoII E phosphatase activity *in vivo* is regulated by interactions involving residues in other domains of the protein; this is discussed in more detail in Chapter 5. Until the completion of the asymmetric cell division septum, SpoII E must remain inactive. Therefore, either SpoII E must be capable of carrying out dephosphorylation in free solution and is inhibited by interaction with the cell division machinery until septum completion or it must be inactive until activated by a signal communicating septum completion.



**Figure 4-15: Comparison of SpoII E fragment phosphatase activity.**

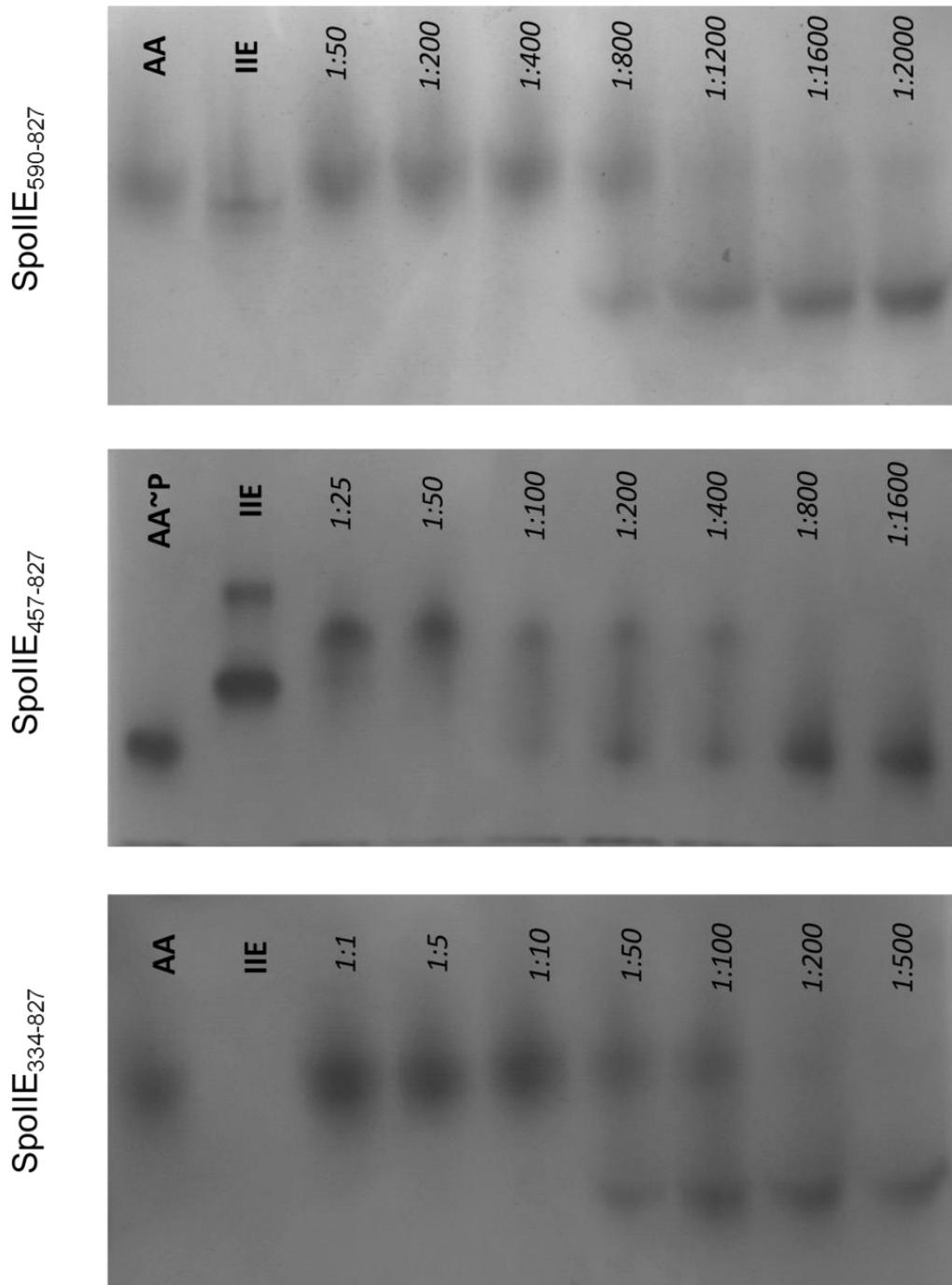
Control lanes containing SpoII AA~P (5  $\mu g$ ) and SpoII AA (5  $\mu g$ ) are on the left. Each subsequent pair of lanes hold a control lane of the SpoII E fragment (5  $\mu g$ ) followed by a sample of the 1:200 ratio phosphatase reaction.

The addition of elements of SpoIIE domain II, which is thought to interact with the divisome, to the PP2C domain could have an effect on its phosphatase activity. In order to determine whether domain II has an inhibitory or activating effect, the phosphatase activity of 3 phosphatase domain containing SpoIIE fragments; SpoIIE<sub>334-827</sub>, SpoIIE<sub>457-827</sub> and SpoIIE<sub>590-827</sub> was analysed using the native PAGE phosphatase assay described in section 4.2.2. A comparative experiment was carried out measuring dephosphorylation of SpoIIAA~P at an enzyme to substrate ratio of 1:200 over 60 minutes (Figure 4-15).

The levels of dephosphorylation observed indicate the following activity hierarchy: SpoIIE<sub>590-827</sub> > SpoIIE<sub>457-827</sub> > SpoIIE<sub>334-827</sub>. These fragments were further analysed at a wider range of ratios to gain a better appreciation of reaction rates (Figure 4-16).

The reaction rates were roughly compared by determining the lowest concentration of SpoIIE which drove the reaction to completion. This occurred at ratios of 1:400, 1:50 and 1:10 for IIE<sub>590-827</sub>, IIE<sub>457-827</sub> and IIE<sub>334-827</sub>, respectively.

Although these data are only semi-quantitative, it is apparent that SpoIIE<sub>590-827</sub> and SpoIIE<sub>457-827</sub> are ~50 and ~5 times more active than SpoIIE<sub>334-827</sub>. This shows that SpoIIE domain II has an inhibitory effect on the PP2C phosphatase activity. The data for the SpoIIE<sub>334-827</sub> fragment are more difficult to interpret due to its oligomeric nature, as detailed in section 3.3.4. SpoIIE<sub>457-827</sub> in contrast, has been shown to be monomeric, as analysed by SEC-MALLS, supporting the observed reduction in phosphatase activity.



**Figure 4-16: Native PAGE analyses of SpoIIE fragment phosphatase activity at varying [Enzyme] to [Substrate] ratios.**

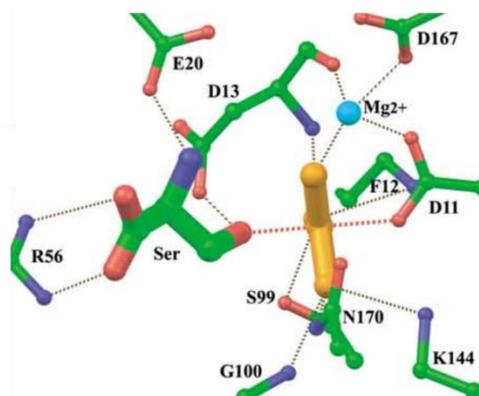
The SpoIIE fragment analysed is indicated to the left of each gel. Either SpoIIAA or SpoIIAA~P is used as a control for band mobility in the first lane of each gel. The second lane of each gel contains 10 µg of the analysed SpoIIE fragment. Ratios above each lane indicate the tested Enzyme:Substrate ratio. All reactions were incubated for 80 minutes before 7.5% Native PAGE analysis.

### 4.3.3 Towards the Structural Elucidation of the SpoIIE:SpoIIAA Interaction

An aim of this project was to gain structural insight into the mode of interaction between SpoIIE and SpoIIAA. The following section describes attempts to isolate the complex using transition state analogues, mutation of Ser58 in SpoIIAA and replacement of  $Mn^{2+}$  in SpoIIE with  $Mg^{2+}$ , a metal which inhibits enzymatic turnover, as described in section 4.3.2.2.

#### 4.3.3.1 Phosphate Transition State Analogues

The structural study of transition states of nucleotide metabolising enzymes commonly exploits charged inorganic moieties which mimic the  $\gamma$ - and  $\beta$ -phosphates of the nucleotide [210, 211]. These mimics bind in place of phosphate groups to mediate transition state interactions with surrounding residues. The use of free serine and  $AlF_3$  in a study of phosphoserine phosphatase (PSP) showed an  $AlF_3$  mediated interaction between the serine and PSP (Figure 4-17) [212].

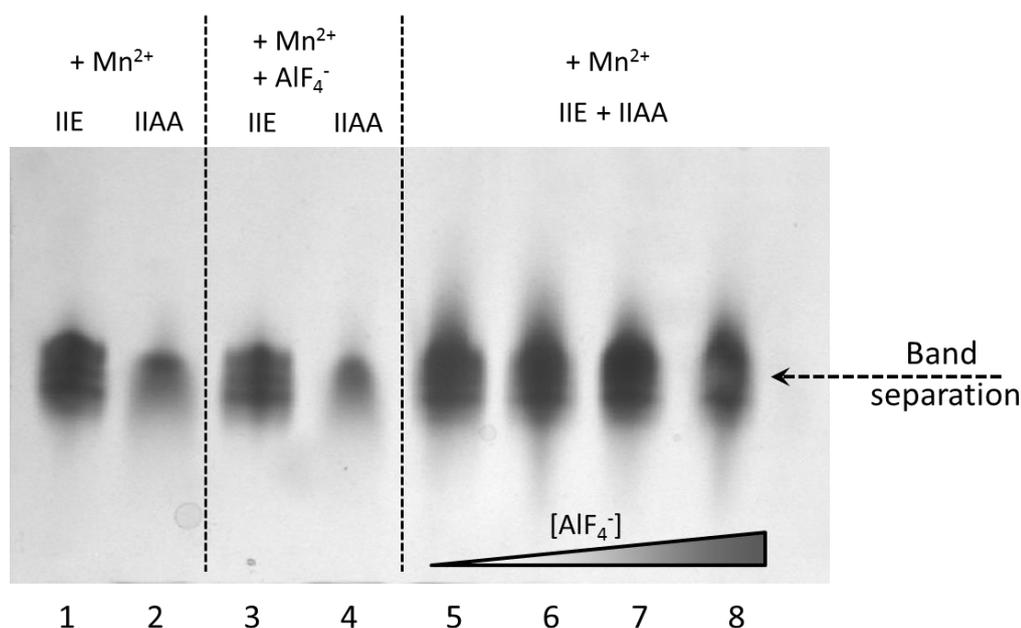


**Figure 4-17: The transition state structural analogue of PSP.**

PDB ID: 1L7N.  $AlF_3$  is shown mimicking a trigonal planar transition state of the substrate phosphate group. As presented by Wang et al., 2002 [212].

While SpoIIAA~P is a much larger substrate molecule than phosphoserine, stabilisation of the active site in a transition state conformation might similarly promote the interaction between SpoIIE and SpoIIAA. SpoIIE<sub>590-827</sub> and SpoIIAA

at an equimolar concentration of 80  $\mu\text{M}$  were mixed in the presence of 10 mM  $\text{MnCl}_2$  and transition state mimics  $\text{AlF}_4^-$  and  $\text{MgF}_3^-$  at varying concentrations. The mixtures were incubated at room temperature for 15 minutes before analysis by native PAGE as described in section 2.7.3.2. No significant change was seen between the bands observed for SpoIIE<sub>590-827</sub> and SpoIIAA with or without transition state analogue. It is hard to interpret this observation.



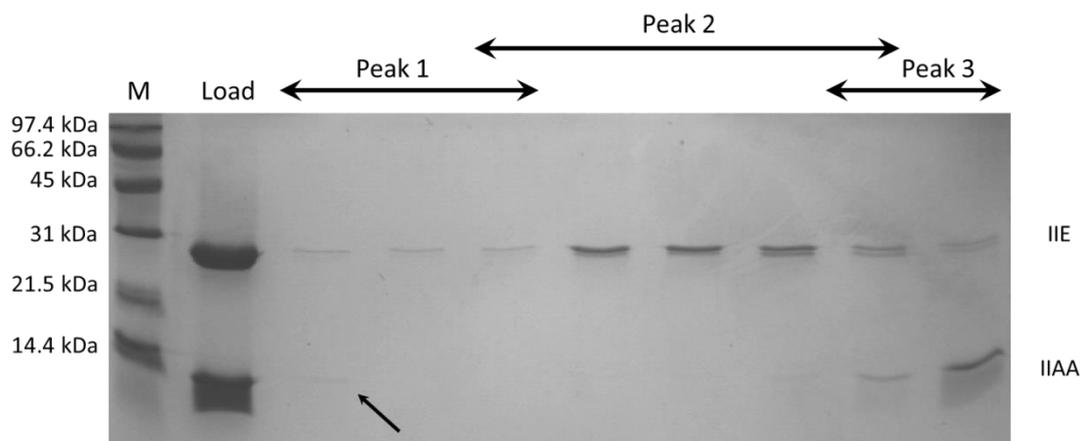
**Figure 4-18: Effect of the phosphate transition state analogue,  $\text{AlF}_4^-$ , on SpoIIE:SpoIIAA complex formation.**

Samples analysed by 7.5% Native PAGE. Lane 1; 20  $\mu\text{g}$  SpoIIE<sub>590-827</sub> + 10 mM  $\text{MnCl}_2$ , lane 2; 10  $\mu\text{g}$  SpoIIAA + 10 mM  $\text{MnCl}_2$ , lane 3; 20  $\mu\text{g}$  SpoIIE<sub>590-827</sub> + 10 mM  $\text{MnCl}_2$  + 2 mM  $\text{AlF}_4^-$ , lane 4; 10  $\mu\text{g}$  SpoIIAA + 10 mM  $\text{MnCl}_2$  + 2 mM  $\text{AlF}_4^-$ , lane 5; 20  $\mu\text{g}$  SpoIIE<sub>590-827</sub> + 10  $\mu\text{g}$  SpoIIAA + 10 mM  $\text{MnCl}_2$  + 0.5 mM  $\text{AlF}_4^-$ , lane 6; 20  $\mu\text{g}$  SpoIIE<sub>590-827</sub> + 10  $\mu\text{g}$  SpoIIAA + 10 mM  $\text{MnCl}_2$  + 1 mM  $\text{AlF}_4^-$ , lane 7; 20  $\mu\text{g}$  SpoIIE<sub>590-827</sub> + 10  $\mu\text{g}$  SpoIIAA + 10 mM  $\text{MnCl}_2$  + 2 mM  $\text{AlF}_4^-$  and lane 8; 20  $\mu\text{g}$  SpoIIE<sub>590-827</sub> + 10  $\mu\text{g}$  SpoIIAA + 10 mM  $\text{MnCl}_2$  + 3.33 mM  $\text{AlF}_4^-$ .

Primarily, the native gel mobility of SpoIIE<sub>590-827</sub> and SpoIIAA are very similar and the observation of changes in band strength is difficult. Additionally, any newly formed complex might possess a very similar mass to charge ratio and, hence, have a similar mobility. The transition state analogues are small inorganic anions and when exposed to a potential difference will have a very high mobility causing rapid titration away from the proteinaceous species. As native PAGE gels are run over two hours, only stable long-lived complexes can

be observed by this technique. While the experimental conditions are unfavourable to complex formation, the bands observed for SpoIIE and SpoIIAA mixed with  $\text{AlF}_4^-$  and  $\text{MnCl}_2$  showed a decrease in band resolution when compared to that of the individual proteins (Figure 4-18). This is not strong evidence; however, the smearing effect may be due to an equilibrium between two species with different mobilities. This effect seemed stronger at 60 molar equivalents of  $\text{AlF}_4^-$ , the highest concentration used, where some weak band separation could be discerned. If this interpretation is valid, it suggests that some stable complex may have been formed.

Native PAGE was unable to provide sufficient evidence for SpoIIE:SpoIIAA complex formation in the presence of  $\text{AlF}_4^-$ . Size exclusion chromatography was therefore used to further investigate whether complex formation might be taking place with  $\text{AlF}_4^-$ . A 100  $\mu\text{l}$  sample of SpoIIE<sub>590-827</sub> and SpoIIAA, each at a concentration of 1 mg/ml, with 10 mM  $\text{MnCl}_2$  and 5 mM  $\text{AlF}_4^-$  was injected onto a Superdex S200 10/300 GL gel filtration column equilibrated in 20 mM Tris-HCl, 100 mM NaCl, 10 mM  $\text{MnCl}_2$  and 5 mM  $\text{AlF}_4^-$  at pH 8.0. The  $A_{280}$  chromatogram indicated protein elution in three partially overlapping peaks. The peak fractions were analysed by SDS-PAGE (Figure 4-19).



**Figure 4-19: SEC analysis of SpoIIE:SpoIIAA interaction in the presence of  $\text{AlF}_4^-$ .** Peak fraction analysis carried out by 12% SDS PAGE. M represents Bio-Rad Low Range markers. Load represents the sample loaded onto the size exclusion column.

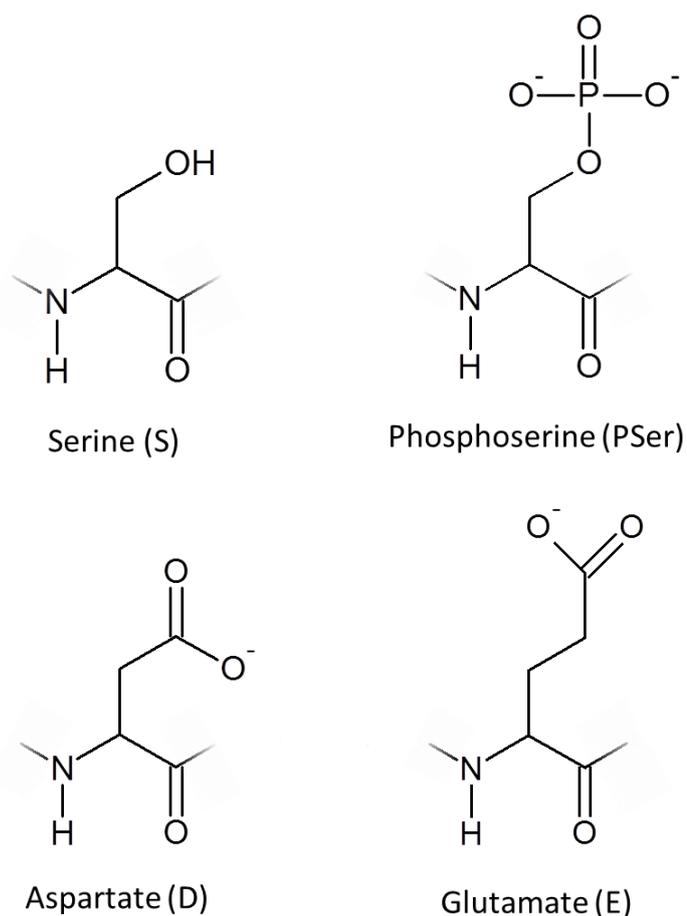
Peak 2 exhibits a higher  $A_{280}$  signal and the SDS PAGE analysis shows that it is dominated by SpoIIE<sub>590-827</sub> which has a higher extinction coefficient than SpoIIAA. Peak 3 contains the highest concentration of SpoIIAA. The SDS PAGE gel suggests that peak 1 is formed by significant constructive interference, in which the perceived peak is a sum of the absorbance caused by the large peak 2 and the much smaller peak eluting immediately before. The first fraction analysed from peak 1 shows the presence of SpoIIE<sub>590-827</sub>, which could represent the leading edge of peak 2, however, a band of around half the intensity is also seen at the same mobility as SpoIIAA. The presence of SpoIIAA in the highest MW peak of the chromatogram suggests that some interaction between SpoIIAA and SpoIIE<sub>590-827</sub> is occurring which might be attributable to  $AlF_4^-$ . However the quantity of material which exhibits this interaction is certainly below 2% of the total protein, indicating a very weak interaction.

Overall the evidence for formation of a stable complex between SpoIIE<sub>590-827</sub> and SpoIIAA in the presence of inorganic phosphate transition state analogues is rather weak although some evidence for a weak interaction does exist. The logical next step in this line of investigation is the co-crystallisation screening of SpoIIAA and SpoIIE<sub>590-827</sub> in the presence of  $AlF_4^-$  generated *in situ*. The process of crystal formation, should it occur, may select for the inclusion protein complex under the correct crystallisation conditions. It should be noted that this strategy would require large, but not prohibitive, quantities of SpoIIAA and SpoIIE<sub>590-827</sub> as well as an extended crystal optimisation process and it cannot be relied upon to ever produce the desired results.

#### **4.3.3.2 Phosphoserine mimics: Mutation of SpoIIAA**

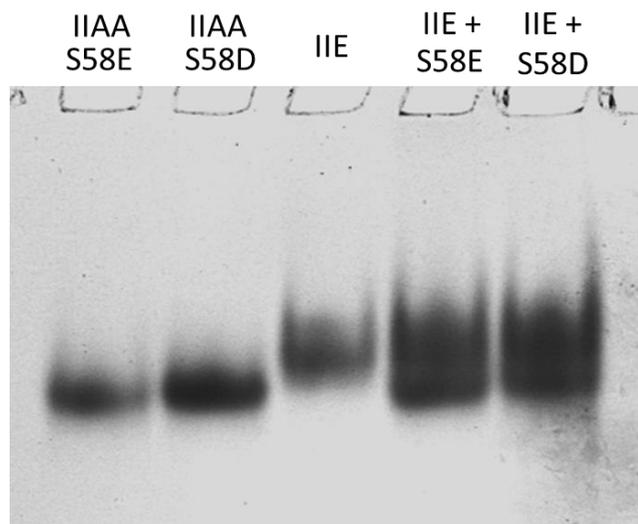
A phosphorylated serine residue on the surface of a protein presents a doubly negatively charged species on the side chain's  $\delta$ -position. For this group to be removed from SpoIIAA~P, it is anticipated the phosphate moiety must reach into the active site of SpoIIE. It is at this stage that the enzyme-substrate interaction should be strongest. When a hydrolysable group is replaced by a non-reactive mimic, the reaction cannot proceed and a long-lived non-productive complex may be formed.

Glutamate has been seen to mimic phosphoserine in other systems [213]. In an attempt to mimic the SpoIIAA phosphoserine, serine 58 was mutated to both aspartate and glutamate. Aspartate and glutamate present an anionic carboxylate group in the  $\gamma$  and  $\delta$  positions, respectively, which could interact with elements of the active site in an analogous fashion to a phosphate group (Figure 4-20).



**Figure 4-20: The chemical structures of Serine and Phosphoserine residues with the proposed structure mimics, Aspartate and Glutamate.**

Oligonucleotide primers were designed to mutate the wild type AGC codon to GAC and GAA encoding aspartate and glutamate, respectively. The mutagenesis reaction as described in section 2.3 was carried out on the pET-YSBLIC3C\_SpoIIAA vector using site directed mutagenesis. DNA sequencing was carried out on the isolated mutant clones and showed the presence of the intended mutation in both cases.



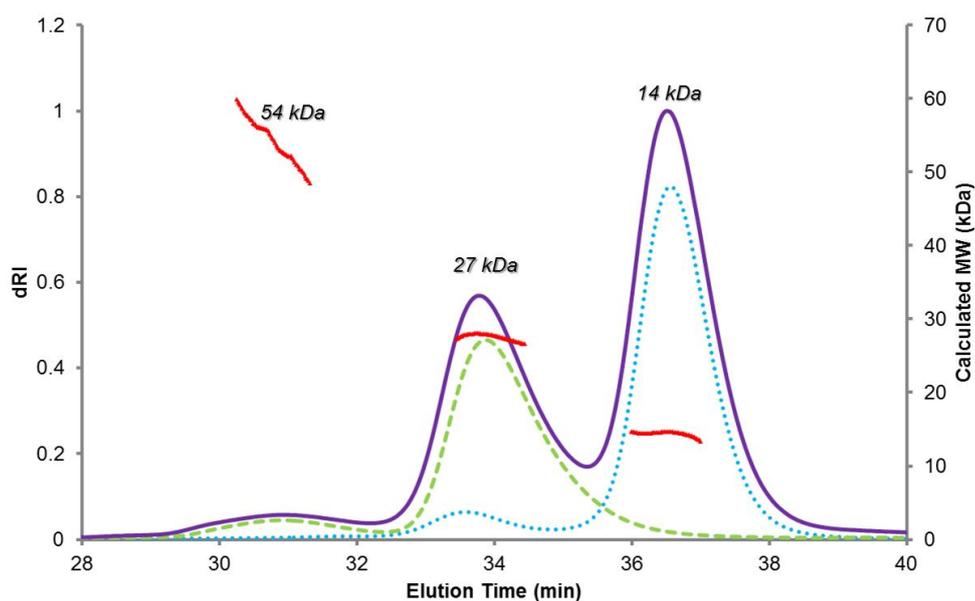
**Figure 4-21: Native PAGE analysis of SpoIIAA mutant interaction with SpoIIE<sub>590-827</sub>.**  
 In each lane IIE indicates 8 µg of SpoIIE<sub>590-827</sub>. SpoIIAA mutant containing lanes each represent 10 µg of protein. All samples contained 10 mM MnCl<sub>2</sub>.

Competent *E. coli* B834 (DE3) cells were transformed with constructs harbouring SpoIIAA(S58D) and SpoIIAA(S58E). Expression and solubility testing showed that on induction with IPTG, S58D is isolated with the highest solubility at 16°C after 20 hours with lysis in 50 mM Tris-HCl, 50 mM NaCl at pH 8.0, while the S58E induction proceeded at 37°C for 4 hours with lysis in 50 mM Tris-HCl, 500 mM NaCl at pH 8.5. The solubility observed for both mutant proteins was notably lower than that of wild type SpoIIAA. A purification process analogous to that of SpoIIAA was carried out for both S58D and S58E. Both mutants eluted from a HiLoad Superdex 75 PG gel filtration column at around 75 ml, consistent with a monomeric species, as seen with SpoIIAA, rather than a dimeric, SpoIIAA~P-like, species.

Native PAGE showed that a mixture of either mutant with SpoIIE<sub>590-827</sub> and MnCl<sub>2</sub> does not result in formation of a new band which might indicate complex formation (Figure 4-21).

As previously mentioned, observation of a species by Native PAGE indicates a long lifetime. In order to further investigate potential complex formation, the SpoIIAA (S58D) mutant was analysed by SEC-MALLS alongside SpoIIE<sub>590-827</sub> as described in section 2.7.5. The running buffer used was 50 mM Tris-HCl, 100 mM NaCl, 10 mM MnCl<sub>2</sub> at pH 8.0. Unfortunately, major precipitation of SpoIIAA

(S58E) on thawing prevented a parallel investigation with this mutant. The differential refractive index chromatogram indicated the elution of three protein-containing peaks (Figure 4-22). This chromatogram overlays well with a combination of the chromatograms recorded for SpoIIE<sub>590-827</sub> and for SpoIIAA (S58D) separately. The calculated molecular weights agree well with a SpoIIE<sub>590-827</sub> dimer (theoretical MW: 53 kDa), a SpoIIE<sub>590-827</sub> monomer (theoretical MW: 26.5 kDa) and a SpoIIAA (S58D) monomer (theoretical MW: 13.2 kDa) in order of elution. The calculated mass of a small peak eluting at 33.5 min in the SpoIIAA (S58D) chromatogram was 27.6 kDa which suggests a small proportion of the SpoIIAA (S58D) exists as a dimer (theoretical MW: 26.4 kDa).



**Figure 4-22: SEC-MALLS analysis of interaction between SpoIIE<sub>590-827</sub> and SpoIIAA (S58D).** Differential refractive index traces for IIE<sub>590-827</sub>, IIAA (S58D) and IIE<sub>590-827</sub> + IIAA(S58D) are shown as green dashed, blue dotted and purple solid lines, respectively. Peak MW calculations from the IIE<sub>590-827</sub> + IIAA(S58D) experiment are plotted as red lines, average calculated MW for each peak is annotated. All samples contained 10 mM MnCl<sub>2</sub>.

The evidence suggests that the S58D and S58E mutants of SpoIIAA are unable to form a long-lived complex with SpoIIE<sub>590-827</sub>. Of the two mutants, S58E theoretically seemed the more promising as the anionic carboxylate group was presented at the same distance from the SpoIIAA backbone as the phosphate group. Additionally, the phosphate group presented to SpoIIE on phosphoserine is tetrahedral rather than trigonal planar in nature and possesses two formal

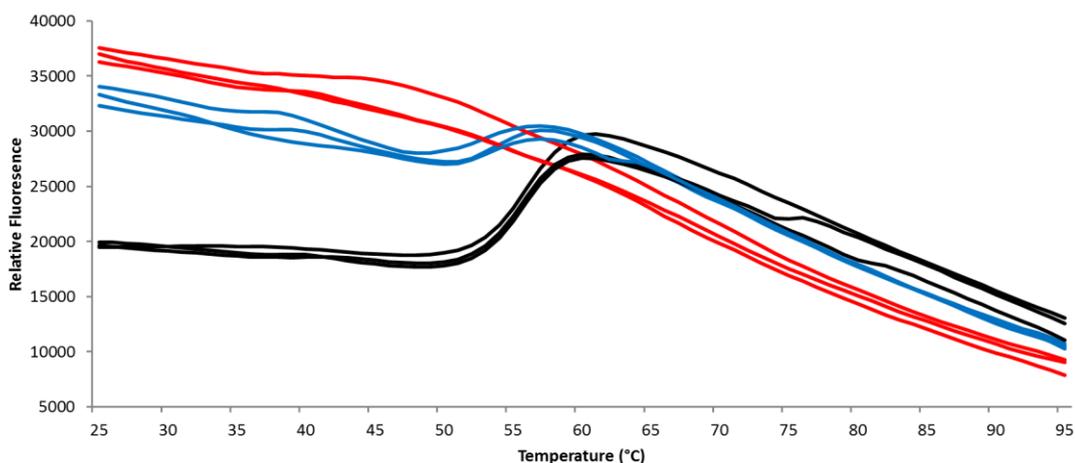
negative charges rather than one. Stabilisation of the phosphate group by residues or water molecules in the active site could very well be inadequately mimicked by a carboxylate group's disparate charge and geometry.

#### **4.3.3.3 Use of an Inactive Active Site Metal to Trap a Pre-turnover Complex**

Earlier, in section 4.3.2.2, I showed that  $Mn^{2+}$  supports much higher reaction rates than other divalent cations tested. This observation supports two possible explanations:

1. SpoIIE only binds  $Mn^{2+}$ .
2. SpoIIE may bind other divalent cations but these cannot support dephosphorylation of SpoIIAA~P efficiently.

In order to test which hypothesis pertains, the melting temperature of SpoIIE<sub>590-827</sub> was analysed in the absence of metal ions and in the presence of  $Mn^{2+}$  or  $Mg^{2+}$  using ThermoFluor as described in section 4.2.4 (Figure 4-23). The data recorded for SpoIIE<sub>590-827</sub> with no added metal suggest that either the protein is natively disordered or becomes disordered by the addition of the hydrophobic dye. This is suggested by the high fluorescence at low temperature, indicating that a large quantity of dye is bound to exposed hydrophobic residues in the protein. In contrast, a much lower starting fluorescence is observed when 10 mM  $MnCl_2$  is present. These curves also exhibit an increase in fluorescence at around 55°C indicating that SpoIIE<sub>590-827</sub> becomes thermally denatured at this temperature. The data recorded from SpoIIE<sub>590-827</sub> in the presence of 10 mM  $MgCl_2$  shows a similar pattern, however to a lesser extent. The starting fluorescence is higher than for  $Mn^{2+}$  containing samples, suggesting that either partial denaturation of all of the protein, or full denaturation of some of the protein has occurred even at lower temperatures. A similar increase in observed fluorescence at around 55°C suggests that some protein has been stabilised to the same extent observed with  $Mn^{2+}$ .

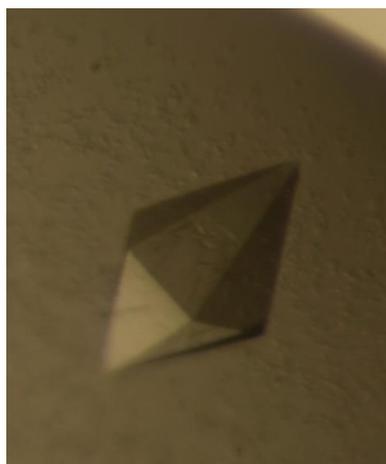


**Figure 4-23: Fluorescence based thermal shift assays of SpoIIE<sub>590-827</sub> stabilisation by Mn<sup>2+</sup> and Mg<sup>2+</sup>.**

Three measured spectra are shown for SpoIIE<sub>590-827</sub> in the presence of no added metal (red), 10 mM MnCl<sub>2</sub> (black) and 10 mM MgCl<sub>2</sub> (blue).

The thermal melt data suggest that SpoIIE<sub>590-827</sub> binds, and is stabilised by, Mn<sup>2+</sup> and to a lesser extent Mg<sup>2+</sup>. This, coupled with the relative enzymatic activity of SpoIIE<sub>590-827</sub> when incubated with these divalent cations, suggests that Mg<sup>2+</sup> can occupy the active site of SpoIIE but is much less able to turnover SpoIIAA~P than Mn<sup>2+</sup>.

Co-crystallisation screening experiments were set-up using SpoIIE<sub>590-827</sub> and SpoIIAA~P including 5 mM and 10 mM concentrations of either MnCl<sub>2</sub> or MgCl<sub>2</sub>. Formation of small crystals was observed in both MnCl<sub>2</sub> and MgCl<sub>2</sub> containing conditions. However, the number of hits observed in the presence of both 5 mM and 10 mM MgCl<sub>2</sub> was over twice that observed for MnCl<sub>2</sub> containing conditions. The crystals grown were bipyramidal in morphology which is similar to the shape observed in past crystallisations of the SpoIIE<sub>590-827</sub> fragment alone. Optimisation of hits in crystal screening produced large bipyramidal crystals grown in a hanging drop experiment, as described in section 2.8.2, which were grown from a well solution of 50 mM NaCitrate, 15% PEG 6K at pH 5 with a protein solution of 5 mg/ml of both SpoIIE<sub>590-827</sub> and SpoIIAA~P in 20 mM Tris-HCl 100 mM NaCl, 5 mM MgCl<sub>2</sub> at pH 8 (Figure 4-24). The crystals were tested for X-ray diffraction in house and showed diffraction to between 4 Å and 5 Å.

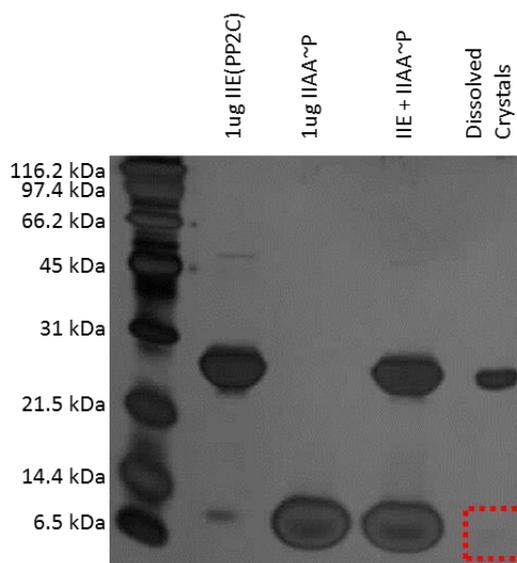


**Figure 4-24: Bipyramidal crystal grown from SpoIIE<sub>590-827</sub> and SpoIIAA~P in the presence of 5 mM MgCl<sub>2</sub> from a well solution of 50 mM sodium citrate, 15% PEG 6K at pH 5.**

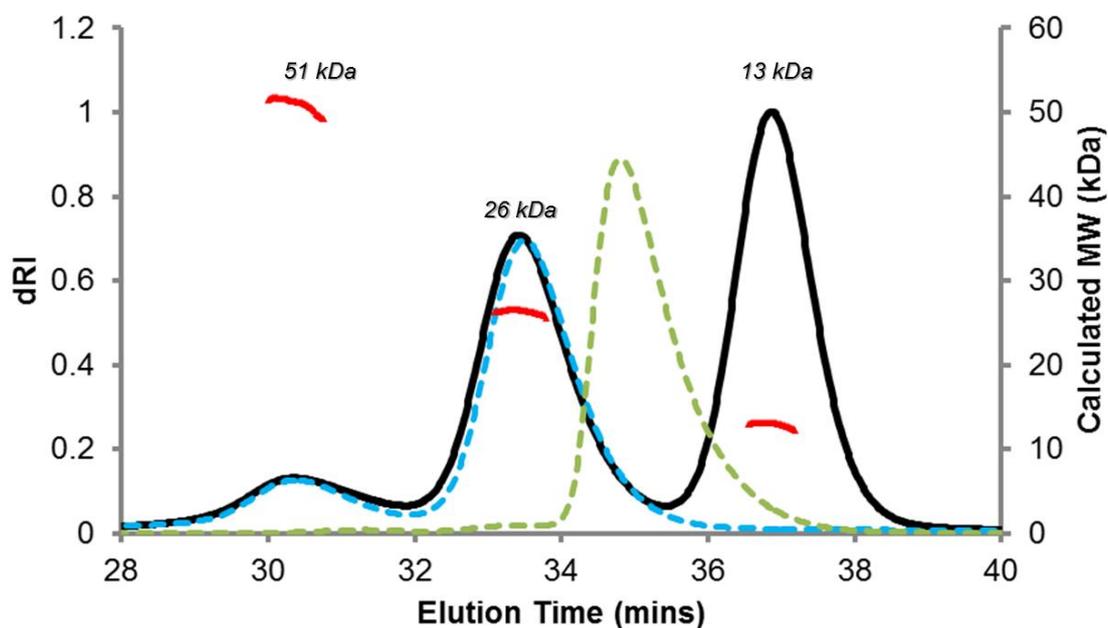
Two crystals were sent to the Diamond Light Source, UK, for data collection on the i04 beamline. The X-ray wavelength used was 0.92 Å. Datasets were collected to resolutions of 3.4 Å and 3.8 Å. The data were processed with XDS using the xia2 program [214, 215], which calculated a unit cell with parameters;  $a = 87.14 \text{ \AA}$ ,  $b = 87.14 \text{ \AA}$ ,  $c = 159.10 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$  and  $\gamma = 120^\circ$  in a primitive hexagonal space group ( $P6_{2/4}22$ ). These unit cell dimensions are almost identical to those of the PP2C domain swapped dimer structure, however in that case  $c = 320 \text{ \AA}$ ; twice the value determined for the co-crystal. A solution by molecular replacement was attempted using MOLREP and both the SpoIIE<sub>590-827</sub> domain swapped dimer and active site-occluded dimer co-ordinates, however no solution was found.

In order to determine the protein content of these crystals, two crystals fished from the same crystallisation drop as those tested at the synchrotron were dissolved in 25  $\mu\text{l}$  of 50 mM Tris-HCl, 100 mM NaCl buffer at pH 8.0. This solution was then examined by SDS PAGE alongside SpoIIE and SpoIIAA~P (Figure 4-25).

If the crystals contained a 1:1 complex, a SpoIIAA band with intensity equal to roughly half of that exhibited for SpoIIE<sub>590-827</sub> would be visible. The lack of a SpoIIAA band (lane 5) indicates that the crystals contain only SpoIIE<sub>590-827</sub>.



**Figure 4-25: Silver stained SDS PAGE analysis of crystals grown from SpoIIE and SpoIIAA~P in the presence of Mg<sup>2+</sup>.**  
The expected SpoIIAA~P band position in lane 5 is highlighted in red.



**Figure 4-26: SEC-MALLS analysis of SpoIIE<sub>590-827</sub> and SpoIIAA~P in the presence of Mg<sup>2+</sup>.**  
Chromatograms plotting dRI for IIE<sub>590-827</sub>, IIAA~P and IIE<sub>590-827</sub> + IIAA~P are shown as blue dashed, green dashed and black solid lines, respectively. Peak MW calculations from the IIE<sub>590-827</sub> + IIAA~P experiment are plotted as red lines, average calculated MW for each peak is annotated. All samples contained 10 mM MgCl<sub>2</sub>.

To see if an interaction could be observed in the solution state, SpoIIE<sub>590-827</sub> and SpoIIAA~P at a 1:2 molar ratio in the presence of 10 mM Mg<sup>2+</sup> was analysed by SEC-MALLS. This indicated that a species is generated with a calculated molecular weight of 13 kDa (Figure 4-26). The peak representing a SpoIIAA~P dimer (calculated MW = 21 kDa) was not present suggesting that dephosphorylation has taken place to yield an unphosphorylated SpoIIAA monomer (as observed in comparable experiments where Mn<sup>2+</sup> replaces Mg<sup>2+</sup>). This evidence suggests that while the rate of catalysis carried out by SpoIIE is much lower in the presence of Mg<sup>2+</sup> than Mn<sup>2+</sup> (Figure 4-13) exchange does not sufficiently retard the reaction to prevent dephosphorylation of SpoIIAA~P at comparable molar ratios over a duration of around 3 hours; the period between initialisation of the SEC-MALLS sample batch and injection of this sample.

## 4.4 Summary and Discussion

The cloning and isolation of SpoIIAA, SpoIIAB and SpoIIAA~P from *B. subtilis* was carried out. The efficiency of SpoIIE<sub>590-827</sub> catalysed dephosphorylation of *Bsu*SpoIIAA~P was markedly higher than that for *Bsph*SpoIIAA~P. Isolated SpoIIAB, despite oligomerisation, was shown to have kinase activity; producing SpoIIAA~P from SpoIIAA in the presence of ATP.

Investigation into the reaction rates of various fragments of SpoIIE indicated a drop in activity on inclusion of domain 2 elements. A 50-fold reduction in dephosphorylation of SpoIIAA~P was observed between *Bsu*SpoIIE<sub>590-827</sub> and *Bsu*SpoIIE<sub>334-827</sub>. An intermediate 10-fold reduction was observed for *Bsu*SpoIIE<sub>457-827</sub>. While the reduction observed for the *Bsu*SpoIIE<sub>334-827</sub> fragment could, in part, be due to oligomerisation, the activity observed for the *Bsu*SpoIIE<sub>457-827</sub> fragment clearly suggests direct inhibition of SpoIIE's phosphatase activity by SpoIIE domain 2. Interestingly, this observation is in contrast to a previous investigation which showed that full length SpoIIE possessed a higher dephosphorylation rate than that observed for a fragment representing the PP2C domain alone in the presence of 2 mM Mn<sup>2+</sup> [216].

The phosphatase activity of the SpoIIE<sub>590-827</sub> fragment was not detectable at a Mn<sup>2+</sup> concentration below 100 µM. It could be that full length SpoIIE is required to create an active site with high affinity for manganese. The optimum Mn<sup>2+</sup> concentration for sporulation of *B. stearothermophilus* has been reported as between 10 and 30 µM, which supports the hypothesis that a higher Mn<sup>2+</sup> affinity state for the phosphatase domain must exist [217].

Attempts to form a stable SpoIIE<sub>590-827</sub>:SpoIIAA complex were, unfortunately, largely fruitless. Some evidence for a weak interaction in the presence of the phosphate transition state analogue, AlF<sub>4</sub><sup>-</sup>, was observed. It is possible that the use of this additive in co-crystallisation screening might cause *in situ* complex formation and provide crystals. SpoIIAA mutants designed to be mimics of SpoIIAA~P did not show any evidence of interaction with SpoIIE and it is thought that while in other systems an interaction has been observed based on the use glutamate mimics of phosphoserine, the specificity of interaction

between SpoIIE and SpoIIAA~P does not accommodate these variations in substrate. The use of a non-native divalent cation to slow the dephosphorylation of SpoIIAA~P was unsuccessful.

The methods which are suitable for attempted isolation of a phosphatase-phosphoprotein substrate complex are few, explaining to some extent the lack of published structures. A promising but impractical system uses a non-hydrolysable phosphoserine mimic, pCF<sub>2</sub>-Ser, to construct substrate mimic peptides. These mimics have been shown to effect melatonin production and p53 tumor suppressor protein activity on interaction with target proteins *in vivo* [218]. This mimic would potentially serve as a tool for isolating a SpoIIE:SpoIIAA complex structure as it is reported to support native interactions relying on the presence of a phosphorylated serine, however introduction of this moiety in place of SpoIIAA Ser58 by recombinant protein production techniques or by chemical modification is not possible and it would be necessary to carry out a chemical synthesis of the whole protein.

## Chapter 5: *In vivo* Investigation of SpoIIE Mutations; Regulation of Asymmetric Cell Division and Phosphatase Activity by SpoIIE.

### 5.1 Introduction

The importance of SpoIIE in sporulation is not simply underlined by its requirement to activate  $\sigma^F$ , but also in the mediation between the genetic and cytokinetic states of the cell through regulation of  $\sigma^F$  and control of asymmetric septation.

This section describes a number of *spoIIE* mutant phenotypes which show evidence for intragenic regulation of phosphatase activity and inter-communication between domains. The competence system of *B. subtilis* is also described, as it is of relevance to the *in vivo* mutagenesis work documented in this chapter.

#### 5.1.1 Phenotypes of mutant *spoIIE* alleles

SpoIIE's role in asymmetric septum formation, activation of  $\sigma^F$  and, importantly, coupling of the two has been extensively studied through the identification and analysis of mutants. Figure 5-1 shows a representation of some of these mutations and their phenotypes. The data shown has been collated from various literature sources and from Niels Bradshaw (Harvard, personal communication) [159, 163, 219-221].

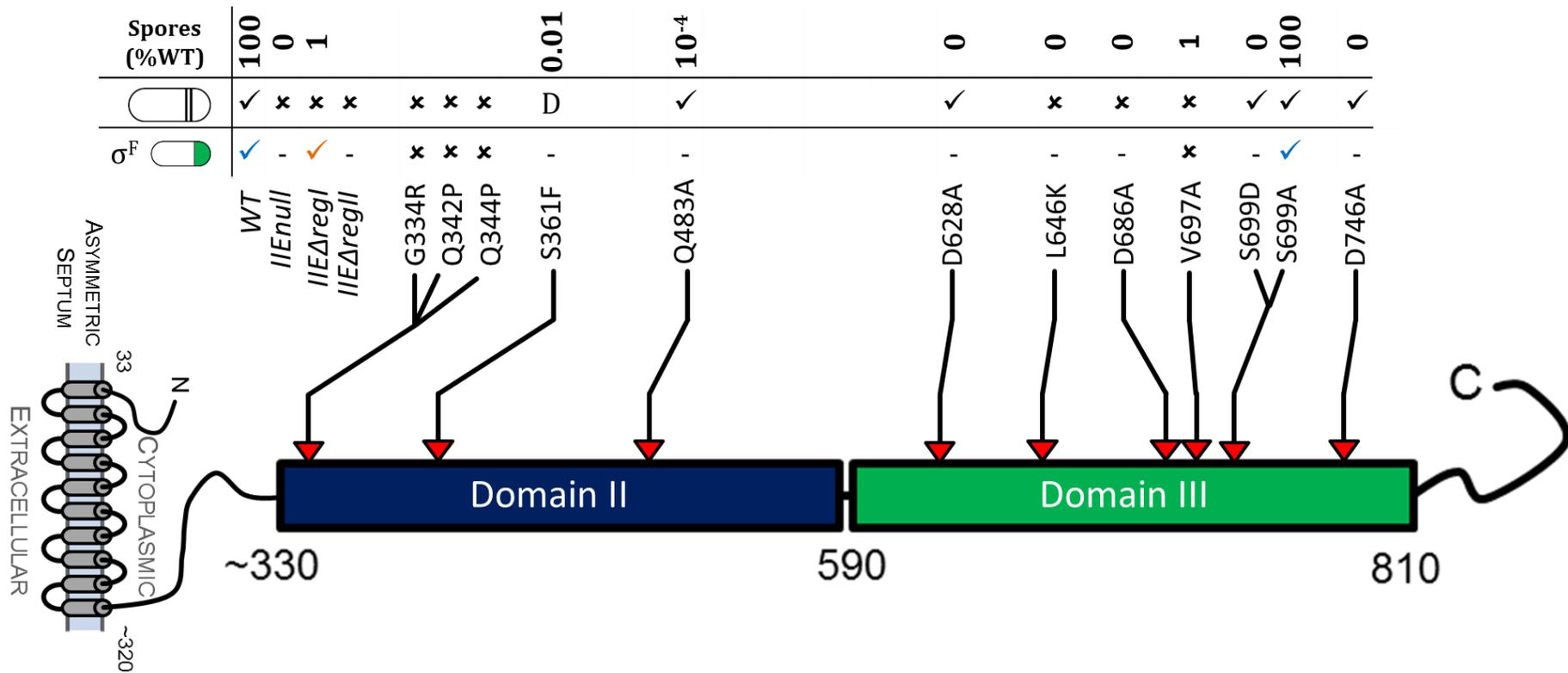
*spoIIE* null mutant cells display aberrant septation, in which roughly 50% of the cells fail to form asymmetric septa at T<sub>2</sub> (two hours after the initiation of sporulation). When cells do form an asymmetric septum, the absence of SpoIIE means that  $\sigma^F$  cannot be activated. Eventually some cells will form a second division septum at the opposite cell pole [220, 222, 223]. The asymmetric septa of *spoIIE* null mutants are much thicker than the normal sporulation septa of wild type cells. This implicates SpoIIE in structural reorganisation of the forming septum. Thick septa are seen in vegetative cell division as a

consequence of the higher peptidoglycan content in preparation for cell fission. During sporulation the thinner septa are thought to afford elasticity to the membrane in order to facilitate engulfment.

Deletion of the transmembrane region of SpoIIE, *spoIIE $\Delta$ regI*, leads to failure of SpoIIE to localise to the septum and a reduced frequency of asymmetric division [163]. When compared to cells with wild-type *spoIIE*, a 56% reduction in the proportion of cells with  $\sigma^F$  activated only in the forespore is also observed. Other cells showed activation of  $\sigma^F$  throughout the whole cell. This indicates a defect in the regulation of SpoIIE phosphatase activity so that SpoIIAA~P is dephosphorylated throughout the cell.

The replacement of region I with two transmembrane helices from the *E. coli* protein, MalF, shows similar properties to wild type *spoIIE* in promotion of polar septum formation in vegetative cells which overproduce FtsZ [169]. This evidence suggests that the role of SpoIIE region I is limited to anchoring the protein to the membrane and that the cytoplasmic domains are responsible for regulation. In sporulating cells, however, this mutant causes much reduced formation of asymmetric septa and  $\sigma^F$  is prematurely activated throughout the whole cell. It was observed that the levels of MalF-SpoIIE present in the cell were much higher than those seen for wild type SpoIIE. Therefore, the replacement of the transmembrane domain with the MalF helices stabilises SpoIIE. An increase in SpoIIE levels is sufficient to explain the uncompartimentalised  $\sigma^F$  activation which would cause asymmetric cell division to be abandoned.

Deletion of the central domain of SpoIIE in the *IIE $\Delta$ regII* mutant prevents subcellular localisation to the asymmetric septum [169], and the observed reduction in the formation of asymmetric septa is ~95% when compared to wild type. Localisation experiments in the same work indicated a diffuse distribution of SpoIIE $\Delta$ regII-GFP throughout the cell membranes rather than at the sites of asymmetric division. This evidence implicates domain II of SpoIIE in interaction with the cell division machinery which seemingly directs localisation.



**Figure 5-1: Location of sites of mutation and the associated phenotypes of SpoIIE Mutants.**

SpoIIE's three domain structure is shown. Phenotypes associated with mutations in SpoIIE are compared (top) by Spores (%WT); % sporulation efficiency compared to Wild Type (where data are available), ; frequency of asymmetric septation and  $\sigma^F$  ; compartment specific  $\sigma^F$  activation. The phenotypes of Wild Type (WT), SpoIIE deletion (*IIEnull*) and domain deletion mutants (*IIEΔregI* and *IIEΔregII*) are shown to the left. The phenotypes of SpoIIE point mutations in the cytoplasmic domains are given along with an indication of their location in the primary sequence. For frequency of asymmetric septation, ✓ and ✗ indicate frequencies comparable to WT and *IIEnull* mutants, respectively, D indicates a delay in septum formation which is recovered from over time. For compartment specific  $\sigma^F$  activation, - indicates that little activation of  $\sigma^F$  is seen in any cells, ✓ indicates that the specificity is comparable to wild type, ✓ indicates that specificity is maintained but at a lower level than with wild type and ✗ indicates that specificity is abolished and that  $\sigma^F$  is activated throughout the cell.

Many *spoIIE* point mutations have been identified and analysed giving insight into SpoIIE's function during sporulation. A number of mutations in the PP2C domain (domain III) exhibit defects in  $\sigma^F$  activation. For example D686A and D746A are unable to activate  $\sigma^F$  [169]. The crystal structure of the SpoIIE phosphatase PP2C domain shows that D746 co-ordinates a  $Mn^{2+}$  ion in the active site. Mutation of this residue would disrupt metal coordination in the active site, explaining the phenotype of the D746A mutation. D686A prevents SpoIIAA~P dephosphorylation and causes a severe defect in sporulation septum formation; even more pronounced than that shown in a *spoIIE null* mutant. This indicates that elements of region III contribute to and potentially co-ordinate cytokinesis and the enzymatic activity of SpoIIE.

A V697A mutant results in hyperactivation of SpoIIE phosphatase activity. This causes  $\sigma^F$  activity throughout the cell before asymmetric division takes place because SpoIIAA~P is dephosphorylated throughout the cell [219]. This suggests that this mutant bypasses some form of phosphatase activity regulation, which is key for correct asymmetric septation and compartmentalised  $\sigma^F$  activity. This mutation was studied in combination with a  $\sigma^F$  mutant (V233M) which causes large shifts in E- $\sigma^F$  affinity for various promoters. The phenotype of the double mutant is identical that of V697A. In contrast, when *spoIIE(V697A)* was combined with a  $\sigma^F$  knockout mutant asymmetric septation was restored. This indicates that the signal to abandon asymmetric septation is due to a gene which is hyperactivated by the  $\sigma^F(V233M)$  mutant.

The interdependent roles played by domains II and III of SpoIIE, implied by the D686A mutant, are further supported by two key mutations found in region II, S361F and Q483A. Each of these mutations prevents activation of  $\sigma^F$ , and hence exerts an effect through the phosphatase domain. However the mode by which  $\sigma^F$  activation is blocked differs for the two mutants. While SpoIIE(Q483A) exhibits no *in vitro* phosphatase activity, explaining  $\sigma^F$ 's inactive state, SpoIIE(S361F) is able to carry out dephosphorylation of SpoIIAA~P indicating that perhaps *in vivo* this mutant is unable to release dephosphorylated SpoIIAA after turnover [159, 221]. The Q483A mutant shows no defect in asymmetric

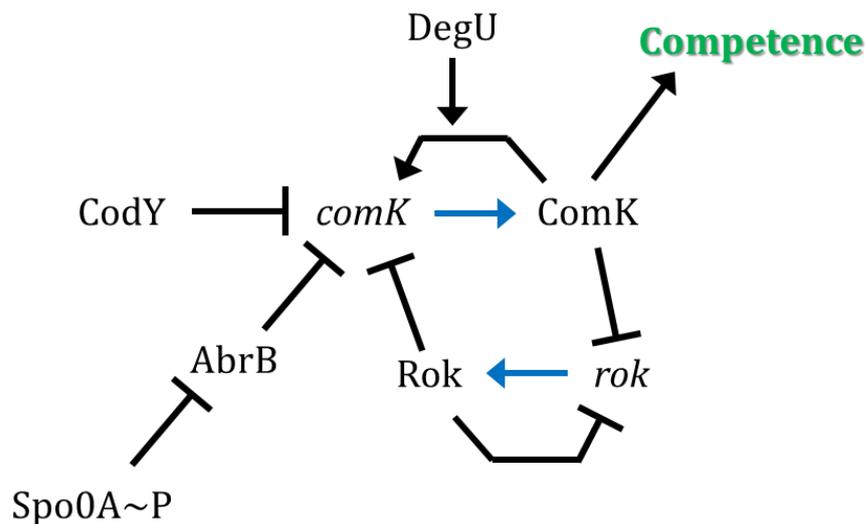
septation whereas, the S361F mutant was shown to exhibit delayed asymmetric septation. At 90 minutes after sporulation initiation, the fraction of cells that had undergone asymmetric cell division was comparable to that seen in a SpoIIE null mutant, however, at 150 minutes the number of cells with asymmetric septa was similar to wild type strains.

The S361F and Q483A mutations are suppressed by point mutations elsewhere in the *spoIIE* gene [221]. The sites of the suppressor mutations map to domains II and III, again suggesting interaction between these two domains. Of the 10 identified suppressors of S361F, six are located in the phosphatase domain. There were only 3 suppressors identified of Q483A; domain II harbours Q342P and R502L, while the previously mentioned V697A mutation suppresses Q483A from domain III. Further analysis of the Q342P suppressor showed that the in presence of Q342P as a single mutation in *spoIIE*,  $\sigma^F$  is activated earlier and to higher levels than is normal, as observed with V697A.

Early observations assigned discrete functions to the putative domains of SpoIIE: domain I, membrane association; domain II, interaction with the cell division machinery; domain III,  $\sigma^F$  activation. The phenotypes exhibited by *spoIIE* mutants challenge this assignment. For example, mutations in one cytoplasmic domain can have an effect on the assigned function of the other. This is indicative of inter-dependency between the roles of both cytoplasmic domains of SpoIIE. Further, this suggests a direct communication of the state of septation through domain II to the PP2C domain, a signal which clearly has a strong influence on  $\sigma^F$  activation. There is the possibility that regulation of phosphatase turnover is caused by block of SpoIIAA release by SpoIIE after turnover, as indicated by the observed S361F phenotype. There is a notable absence of identified mutations in the transmembrane domain, short of its deletion, which cause aberrant sporulation phenotypes.

### 5.1.2 Competence in *B. subtilis*

*B. subtilis* cells can develop competence spontaneously as a survival strategy. On entry into the stationary phase, the transcription of over 20 competence specific or related genes is upregulated under the control of ComK, a master transcriptional regulator [224]. Interestingly, only 10-20% of cells in stationary phase cultures are generally observed to possess competence. This is due to a noise-induced bistability which is explicable through the way *comK* transcription is regulated [225] (Figure 5-2).



**Figure 5-2: Regulation of ComK, the master transcriptional regulator of competence in *B. subtilis*.**

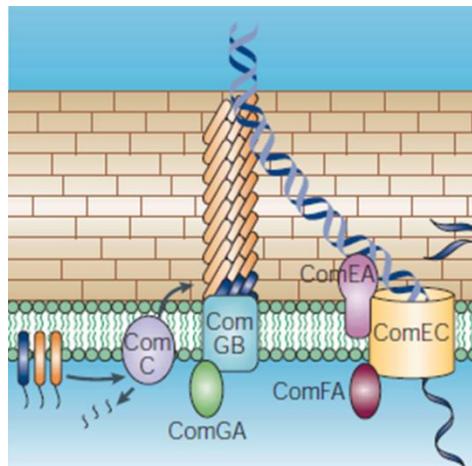
Black barred lines represent repression. Black arrows indicate activation. Blue arrows represent gene expression. Adapted from Maamar et al., 2005 [225].

*comK* is repressed by CodY and AbrB (the activities of which are depleted in stationary phase conditions) and by Rok [226]. ComK itself is a repressor of transcription of *rok*. After the relief of AbrB and CodY repression, ComK levels might reach a threshold sufficient for the relief of its own repression by Rok through a positive feedback loop, much like the relief of AbrB by Spo0A~P via  $\sigma^H$ . Promotion of *comK* expression is provided by DegU which increases the affinity of ComK for its own promoter but cannot cause transcription of ComK in its absence. The basal level of the ComK regulator is therefore the determining

factor in its own regulation. In a stationary phase culture, the basal levels of ComK are in a range where 10-20% of cells can overcome repression and fully adopt competence.

### 5.1.2.1 DNA Uptake Machinery

In competent cells of *B. subtilis*, DNA uptake machinery assembles at the cell poles and consists of two macromolecular structures [96]. The first, the pseudopilus, is responsible for the binding of extracellular DNA termini at the cell wall and their direction towards a membrane bound transport apparatus. The pseudopilus is a tower-like apparatus constructed from major and minor pseudopilins (Major: ComGC – Minor: ComGD, ComGE and ComGG) by the assembly and anchoring proteins; ComC, ComGA and ComGB (Figure 5-3). It has been argued that the pseudopilus may be able to extend and retract through the cell wall to facilitate the relocation of captured DNA to the cell membrane [227].



**Figure 5-3: The DNA uptake apparatus produced by competent *B. subtilis* cells.**

The pseudopilus, left, which recognises dsDNA termini is anchored by ComGB and ComGA. Pseudopilins (major – orange, minor – blue) are added to the forming pseudopilus by ComC. DNA transport is carried out by ComEC, a membrane channel forming protein, assisted by the DNA binding protein, ComEA, and putative ATP-dependent translocase ComFA. NucA (not shown), a DNase, is thought to cleave dsDNA to increase the population of DNA termini for recognition by the pseudopilus. As presented by Chen et al., 2004 [228].

Captured DNA is guided to the second structure, a transport portal, through which it is taken up by the cell (Figure 5-3). The portal consists of ComEA, a

DNA binding receptor; ComEC, a membrane channel and ComFA, an ATP-binding protein homologous to an *E. coli* translocase, PriA. Double-stranded DNA is bound by ComEA and passed through the membrane channel, ComEC, presumably driven by the putative translocase, ComFA. During this process, one strand of the dsDNA is degraded and released into the extracellular milieu; the remaining ssDNA is transported into the cytoplasm.

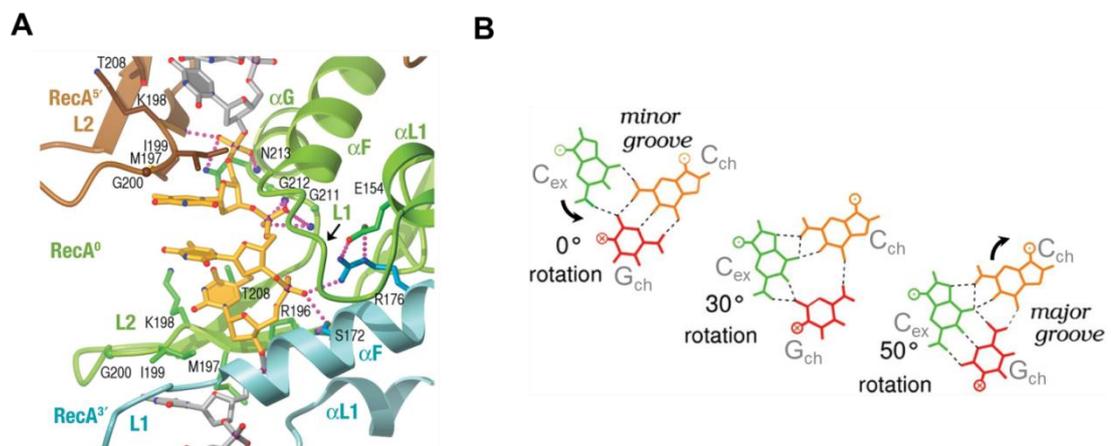
Another protein having a large impact on the efficiency of DNA uptake is, NucA, a manganese-dependent DNase, which is believed to play a role in cleavage of DNA to provide a larger population of DNA termini for recognition by the pseudopili. Deletion of *nucA* results in a large reduction in DNA uptake. This reduction is much more severe during the transformation of large DNA fragments and less severe with small DNA fragments.

#### **5.1.2.2 Post-transformational Treatment of Exogenous DNA.**

The roles of the proteins which chaperone acquired ssDNA which has been taken up are not well defined, with the exception of RecA, which is directly involved in homologous recombination. The internal portion of the DNA uptake membrane channel associates with RecA, as well as DprA, SsbB and YjbF. It is thought SsbB and DprA, ssDNA binding proteins, play a role in protecting ssDNA from nuclease degradation and in mediating the correct formation of a DNA-RecA filament. The role of YjbF is not well understood; however, a homologue from *Streptococcus pneumoniae*, CoiA, has been shown to have no role in formation of the DNA-RecA filament and it is argued that its role is in mediating the recombination process.

When exogenous DNA is taken up by the cell a nucleoprotein complex, the DNA-RecA filament, is produced. This is the active structure which both searches for homology between exogenous ssDNA and the cell's own chromosome and initiates recombination. The RecA-ssDNA filament adopts an extended kinked conformation in which the average number of nucleotides per turn is dramatically increased but local nucleotide triplet configuration (Figure 5-4A) is remarkably similar to B-DNA. This conformation is induced by the interaction of three RecA protomers with each triplet. When this filament interacts with

dsDNA which remains in the native B-configuration at a second RecA binding site, the B-DNA like triplets can parse the genomic DNA through Hoogsteen base pair interactions to search for homology [229]. False pairings are disrupted by the R169 residue of RecA. When identity is found, RecA also mediates a 50 degree rotation of the complementary strand's glycosidic bond to cause pairing exchange resulting in formation of a double helix with the exogenous DNA [230] (Figure 5-4B).



**Figure 5-4: RecA enables recognition of complementary DNA sequences to initiate homologous recombination.**

(A) The conformation of a B-DNA-like triplet (gold) in the nucleoprotein filament. Three RecA protomers induce this conformation: RecA<sup>5'</sup> (brown), RecA<sup>0</sup> (green) and RecA<sup>3'</sup> (light blue). Each RecA protomer interacts with three triplets, once in each of the roles. The ssDNA continues in the 5' and 3' directions (grey). Taken from Chen et al., 2008 [229]. (B) Example of a Watson-Crick C-G base pair exchange which might be induced by RecA. The RecA-bound exogenous cytosine base (C<sub>ex</sub>) recognises a chromosomal cytosine-guanine base pair (C<sub>ch</sub> and G<sub>ch</sub>). A 50° rotation of the glycosidic bond of the guanine-containing strand promotes formation of a Watson-Crick base pair between complementary chromosomal and exogenous DNA strands. C<sub>ch</sub> is displaced from the newly-formed major groove. Taken from Saladin et al., 2010 [230].

### 5.1.2.3 Modes of Homologous Recombination

In *Bacillus*, homologous recombination of exogenous DNA can occur in two modes: double crossover and Campbell-like, also known as single-crossover, integration [231]. The recombination event which takes place in *Bacillus* is dependent on the homology detected in the exogenous DNA.

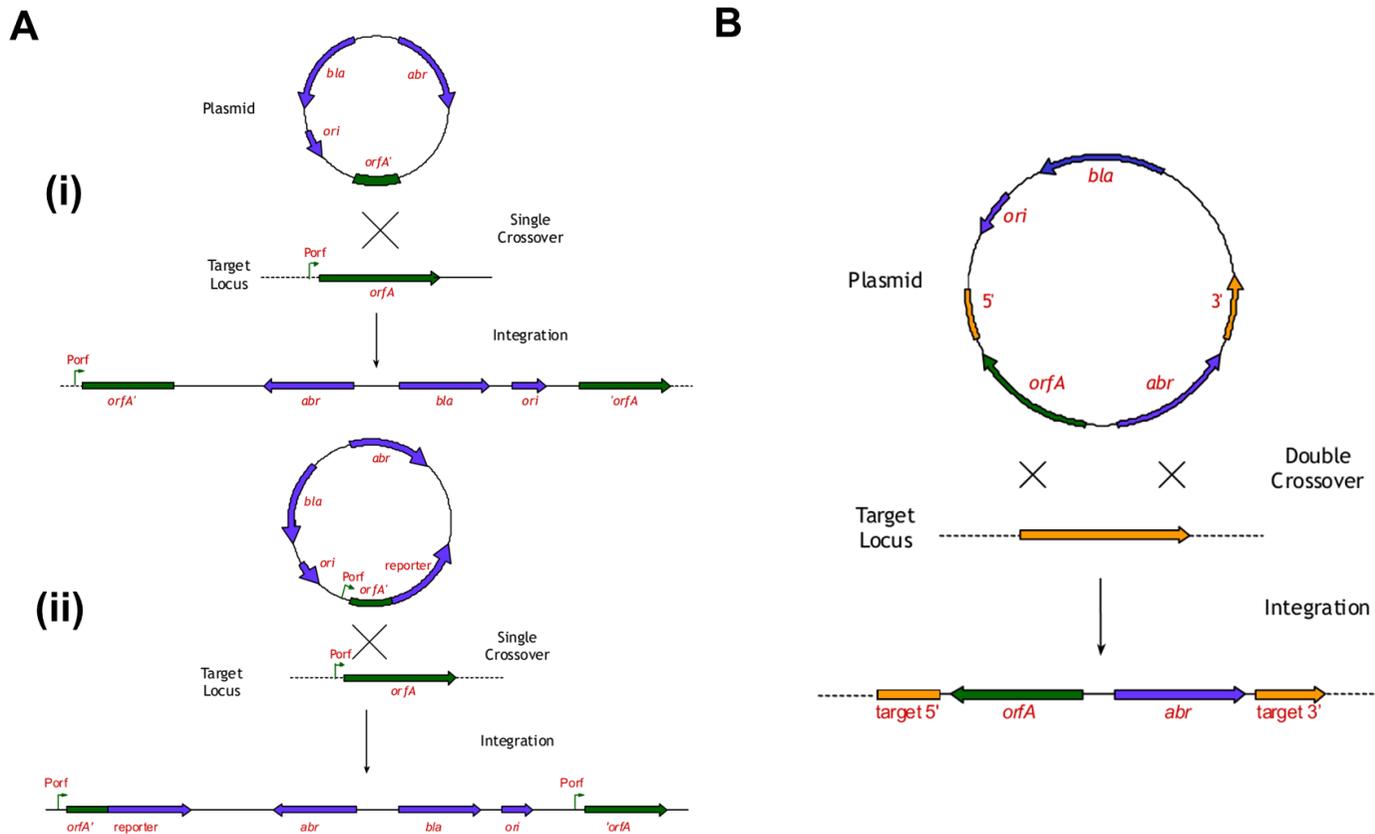
Where there is one region of homology, single crossover integration takes place. The entire DNA fragment is incorporated into the chromosome starting with the homologous segment. It should be noted that because any circular exogenous

DNA is linearised before uptake, the single crossover can occur at any point in the region of homology.

Where there are two regions of sequence homology, a replacement of the sequence between the homologous loci, or double crossover, occurs.

These modes of integration are commonly utilised in the genetic study of *B. subtilis* through introduction of integration vectors which have been engineered to harbour sequences homologous to target genes in *B. subtilis*. Each integration vector contains an *E. coli oriC* region which allows replication of the plasmid in *E. coli* but not in *B. subtilis*. This means that in order for a strain of *Bacillus* to retain the genetic information on an artificially introduced plasmid it must be integrated into the chromosome, hence integration vector.

In addition to the *oriC* region, integration vectors have the *bla* ampicillin resistance cassette for selection in *E. coli* and a second antibiotic resistance cassette, *abr*, for selection in the Gram-positive *B. subtilis*. The desired homologous DNA and any additional features are cloned into the vector at a Multiple Cloning Site (MCS). Three hypothetical integration plasmid constructs and their uses [232] are described in Figure 5-5.



**Figure 5-5: Modes of Exogenous DNA Integration by Homologous Recombination in *B. subtilis*.**

*orfA* is a hypothetical open reading frame and  $P_{orf}$  is its promoter. (A) Single crossover (Campbell-like) integration. Each plasmid has only one region of homology (i) Knockout mutation of *orfA*, a portion of *orfA*, *orfA'*, is present on the plasmid, integration results in a displaced full *orfA* gene being separated from its promoter. (ii) Inclusion of  $P_{orf}$  on the plasmid and addition of a reporter gene (e.g. *lacZ*) after the region of homology leads to the one copy of *orfA* and one reporter gene fusion under the same transcriptional control. (B) Knockout by locus interruption. A double crossover occurs which interrupts the chromosomal target gene with *orfA* and *abr*. The integrant has acquired antibiotic resistance and the insertion of *orfA* but no longer has a functioning target gene. Taken from Zeigler, 2002 [232].

## 5.2 Methods

### 5.2.1 Generation of *B. subtilis* Competent Cells

A single colony of the *B. subtilis* IB333 strain was resuspended in 100  $\mu$ l LB and plated on LB agar. The plate was incubated at 30 °C overnight. The grown cells were resuspended in 1 ml LB and used to inoculate 20 ml of freshly prepared SpC (Appendix I) medium to an OD<sub>600</sub> of 0.4. This culture was then grown for approximately 3 hours until the cells reached the stationary phase (i.e. when the OD<sub>600</sub> remains unchanged for 20-30 min), where they become naturally competent. Thereafter, 2 ml of the culture was diluted into freshly prepared and prewarmed SpII (Appendix I) medium and allowed to grow for 90 min at 37 °C. The cells were then pelleted by centrifugation (2500 x g, 20 min at standard temperature) and resuspended in a mixture of 8 ml of supernatant and 2 ml of 80% glycerol and aliquoted for storage at -80°C. SpC and SpII media are always prepared fresh from the sterile stock solutions.



### 5.2.3 Transformation of competent *B. subtilis*

250 µl of competent cells were thawed at standard temperature and mixed with 250 µl of CaCl<sub>2</sub>-free SpII with 2 mM EGTA and approximately 1 µg of desired plasmid or chromosomal DNA in a 1.5 ml Eppendorf tube. Samples were incubated at 37°C for 30 min in an orbital shaker. TOP agar (0.7% (w/v) LB agar) is melted in a microwave and allowed to cool to 42°C. The transformation mix is added to 4.5 ml TOP agar supplemented with 10 µg/ml kanamycin and the mix is plated onto LB agar supplemented with 10 µg/ml kanamycin for overnight incubation at 37°C. Colonies represent strains which have acquired and integrated the plasmid.

### 5.2.4 Strains

The strains of *B. subtilis* used here are detailed in Table 5-1.

**Table 5-1: *B. subtilis* strains used in this work.**

Strain	Genotype	Description
<b>Obtained:</b>		
IB133	PY79 and <i>Bacillus subtilis</i> 168 prototroph	Wild Type strain
IB614	<i>thrA5, spo0A::kan</i>	Spo0A interrupted by kan resistance, Spo-strain
<b>Novel:</b>		
JTE001	As IB133, <i>spoIIE kan</i>	Kanamycin resistant <i>spoIIE</i> integrant
JTE002	As IB133, <i>spoIIE- K649T kan</i>	<i>spoIIE</i> K649T mutant
JTE003	As IB133, <i>spoIIE- I650A kan</i>	<i>spoIIE</i> K650I mutant
JTE004	As IB133, <i>spoIIE- L695W kan</i>	<i>spoIIE</i> K695W mutant
JTE005	As IB133, <i>spoIIE- V697A kan</i>	<i>spoIIE</i> K697A mutant
JTE006	As IB133, <i>spoIIE- V728A kan</i>	<i>spoIIE</i> K728A mutant

### 5.2.5 Sporulation Efficiency Assay

The assay used to determine sporulation efficiency takes advantage of the heat resistance of mature spores. Cultures of *B. subtilis* were induced to sporulate in nutrient-limiting DSM (Defined Sporulation Media) as follows. Stocks of *B. subtilis* strains to be tested were streaked onto LB agar plates supplemented with appropriate antibiotics (10 µg/ml kanamycin for strains with kanamycin resistance) and grown overnight at 37°C. For each strain, one colony from this plate was resuspended in 50 µl LB liquid media and spread onto a fresh LB agar plate (with appropriate antibiotics), before incubation overnight at 30°C. 600µl of DSM was used to harvest the cells from each LB plate by resuspension with a glass spreader. The dense cell suspension was pipetted into an Eppendorf tube and used to inoculate a 5 ml DSM culture to OD<sub>600</sub> = 0.1 for each sample. This culture was grown for 24 hours at 37°C with shaking at 180 rpm. During this time sporulation is induced and the final culture consists of a distribution of both vegetative cells and spores.

Sporulation efficiency is calculated in spores/ml. This value was determined using the following method. 1 ml of sporulation induced DSM culture was heated for 10 minutes at 80°C in a heating block. 10x serial dilutions in phosphate-buffered saline (PBS) (0.01 M phosphate buffer, 3 mM KCl and 140 mM NaCl at pH 7.4) of the heat-treated culture were carried out. 50 µl of each dilution was then spread onto a separate LB agar plate supplemented with appropriate antibiotics and grown at 37°C overnight. The number of colonies on each plate was counted. Where the number of colonies was between 20 and 200 it was used to determine a value for colony forming units per ml (CFU/ml) with the following formula:

$$CFU.ml^{-1} = \frac{[No. Colonies] \times 20}{[Dilution Factor]}$$

As this value is determined from a culture in which only spores remain viable, this value also represents spores/ml. The sporulation efficiency within a culture, when required, was determined by analysis of a non-heat killed sample of sporulating culture. The obtained CFU/ml represents both spores and

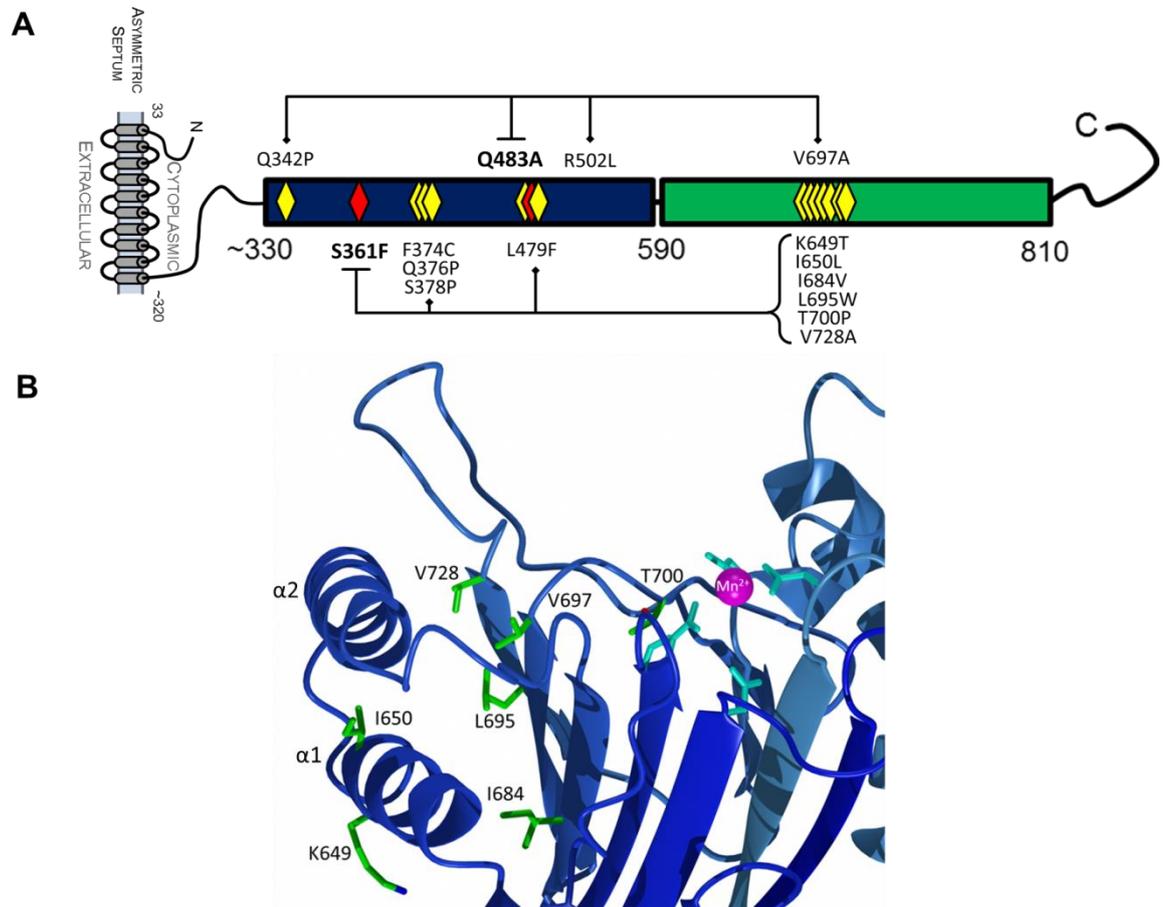
vegetative cells, hence, the proportion of spore vs. vegetative cells can be determined.

## 5.3 Results

### 5.3.1 The Structural Context of Intragenic Suppressor Mutations in the SpoIIE Phosphatase Domain

A previous experiment seeking suppressor mutations of two key SpoIIE mutations in domain II, S361F and Q483A, which cause severe sporulation defects, revealed 13 suppressor mutants within the *spoIIE* ORF [221] (Figure 5-7A). These mutations, when present alongside either S361F or Q483A, restore sporulation. Of the 13 mutants, 7 cause suppression from domain III. That all of these inter-domain suppressors are quite localised in the PP2C domain primary sequence, between residues 649 and 728, suggests a region through which signals might be passed to and from the phosphatase domain. Indeed, when the seven suppressor mutations in the PP2C domain are mapped onto the crystal structure, they all lie within a relatively small region of the structure around helices  $\alpha 1$  and  $\alpha 2$ . Helices  $\alpha 1$  and  $\alpha 2$  lie across one face of the central  $\beta$ -sandwich of the PP2C domain. It is on the outer face of these helices and in the hydrophobic interaction between the helices and the  $\beta$ -sheet that the mutations are located (Figure 5-7B).

The phenotypes presented by S361F and Q483A, discussed in section 5.1.1, strongly suggest signalling between the two cytoplasmic domains of SpoIIE. The structural clustering of their suppressor mutations, further suggests that the  $\alpha 1$ - $\alpha 2$  region of the SpoIIE phosphatase domain is responsible for mediating this signal transduction. The V697A mutant, a suppressor of Q483A, was shown to be hyperactive in  $\sigma^F$  activation as a single mutant and possess around 1% of the sporulation efficiency of wild type SpoIIE [219]. In order to determine if the amino acid substitutions in the region around helices  $\alpha 1$  and  $\alpha 2$  of the SpoIIE PP2C are associated with similar sporulation defects when present alone, two mutant alleles in the  $\alpha 1/\alpha 2$  -  $\beta$ -sheet interface; corresponding to substitutions L695W and V728A, and the two mutant alleles corresponding to substitutions on the outer surface of the  $\alpha 1$ - $\alpha 2$  segment; K649T and I650A, were analysed. Accordingly, single mutants were prepared and the sporulation efficiency was determined.



**Figure 5-7: Distribution of SpoIIIE S361F and Q483A Suppressor Mutations.**

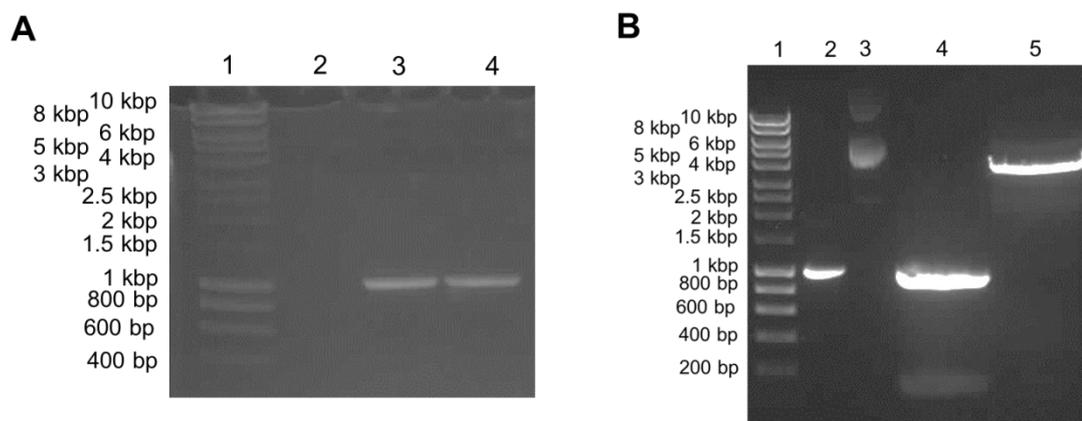
(A) Key mutations (red diamonds) and their suppressors (yellow diamonds). (B) Mapping of suppressor mutations onto the SpoIIIE<sub>590-827</sub> domain-swapped dimer structure (PDB ID: 3T9Q). The PP2C domain is shown as a blue ribbon. The  $\alpha 1$  and  $\alpha 2$  helices are labelled. The active site residues (cyan cylinders) and Mn atom (magenta) are indicated. The wild type residues for suppressor mutations are shown as cylinders coloured by atom.

### 5.3.2 Cloning of SpoIIIE into an Integration Plasmid

A strategy was devised for cloning a *spoIIIE* gene fragment consisting of base pairs 1581-2481 into the pUK-19 integration vector. Mutation of the recombinant plasmid by SDM and subsequent transformation of *B. subtilis* cells would result in strains harbouring *spoIIIE* alleles with the selected point mutations.

Amplification of a *spoIIIE* fragment was originally planned with primers introducing an EcoRI cleavage site downstream of the SpoIIIE stop codon. However, the *spoIIIE* sequence has an internal EcoRI site and so this strategy was modified in favour of an XmaI cleavage site. The *spoIIIE*(1504-2481) gene

fragment was amplified by PCR, as described in section 2.3.2, using a LIC forward primer, IIE\_LIC\_502F and a reverse primer designed to introduce the CCCGGG XmaI recognition sequence. The 5' end of the insert was generated by a BglII site within the amplified *spoIIE* fragment (at base pair 1580). The PCR products were analysed by agarose gel electrophoresis, as described in section 2.1, and a band was observed matching the expected amplicon size, 977 bp (Figure 5-8A). The products were then purified using a PCR purification kit (Qiagen).



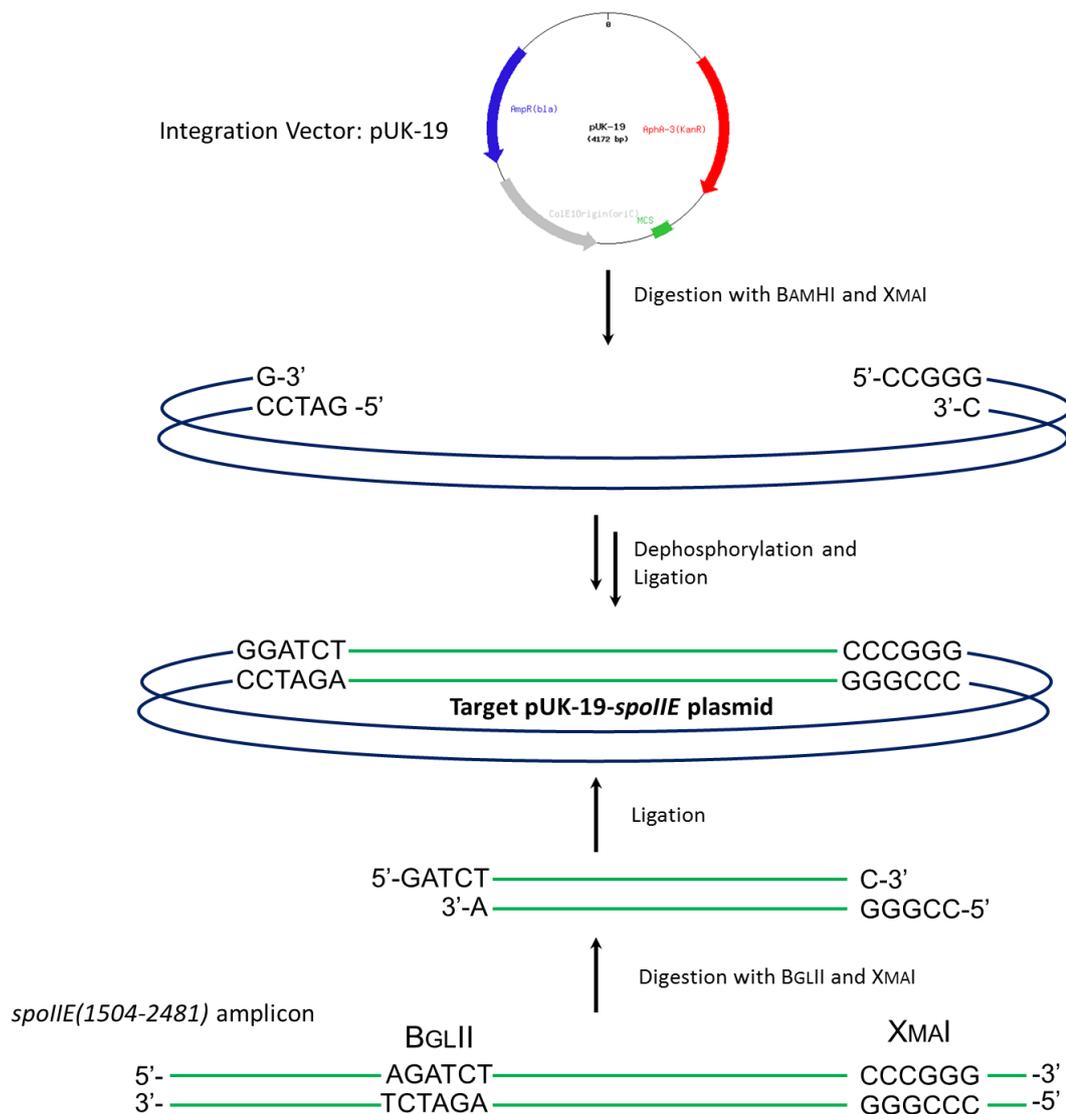
**Figure 5-8: Preparation of cloning reagents for the pUK-19-*spoIIE* plasmid.**

(A) Amplification of the *spoIIE*(1504-2841) fragment by PCR. Lane 1; Hyperladder I (ladder sizes to the left), lane 2; Negative control PCR – no template DNA, lanes 3 and 4; PCR products at around 1 kbp in length. (B) Restriction enzyme digests of both pUK-19 vector and the *spoIIE* fragment insert. Lane 1; Hyperladder I (Bioline) (ladder sizes to the left), lane 2; undigested insert, lane 3; undigested vector, lane 4; digested insert, lane 5; digested vector.

The pUK-19 vector was isolated by miniprep, section 2.2, from transformed *E. coli* XL-10 Gold cells. A double restriction enzyme digest of 450 µg pUK-19 using 1 unit each of BamHI and XmaI was set up in a volume of 20 µl. A second double digest was set up of 500 ng of the PCR amplified insert using 1 unit each of XmaI and BglII in 20 µl. The products of both digestions were separated by agarose gel electrophoresis (Figure 5-8B). The products of the insert digest showed a band at the expected fragment size, 900 bp, and one below 200 bp, representing the cleaved upstream 77 bps, which has been cleaved by BglII. The products of the vector digest show a single band at 4 kbp representing the linearised vector.

The digest product bands were excised from the agarose gel and purified using a QIAquick Gel Extraction Kit (Qiagen).

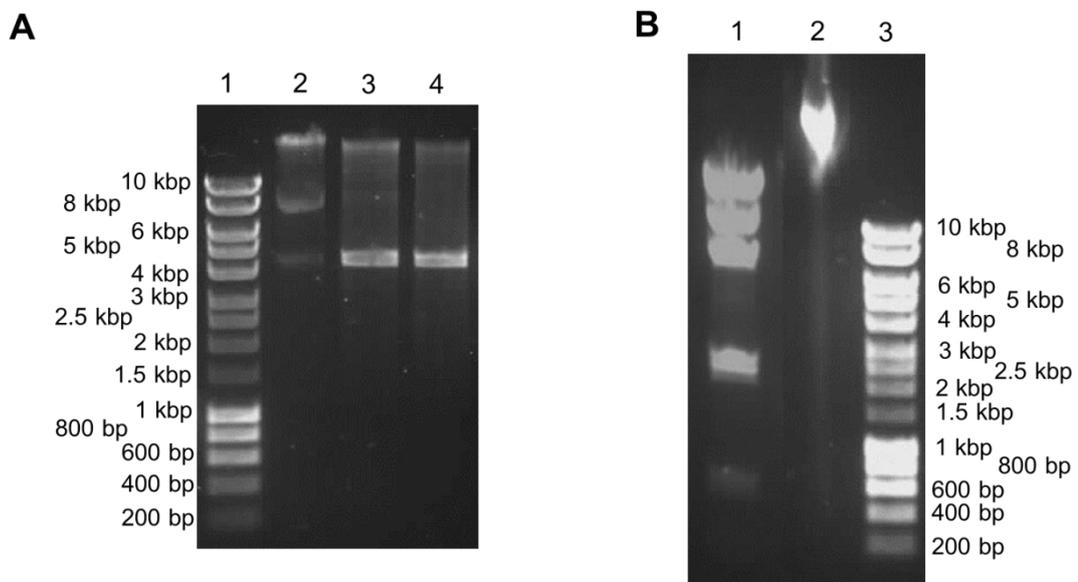
XmaI cleaves the sequence CCGGG to produce 5'-CCGG overhangs. BglII, the recognition sequence of which is AGATCT, is an isocaudomer of BamHI, GGATCC, and produces the same 5'-GATC overhang on cleavage. Because of this, cleavage of the vector with BamHI and the insert with BglII will still yield complementary ends allowing ligation of the vector and insert (Figure 5-9).



**Figure 5-9: Schematic Representation of the XmaI/BamHI strategy for cloning of the pUK-19-spoII E integration plasmid.**

The gel-extracted, purified vector was treated with shrimp alkaline phosphatase at 37°C for 45 minutes after which the enzyme was heat inactivated at 65°C for 5 minutes. This step removes the 5'-phosphate group from the DNA termini of the linearised vector in order to prevent self-religation. A ligation reaction containing 1 µl T4 DNA Ligase, 65 ng of dephosphorylated pUK-19 vector and 50 ng digested *spoIIE* insert in 20 µl was carried out overnight at 16°C. 2 µl of the reaction was added to competent *E. coli* XL-10 Gold cells, following the transformation protocol described in section 2.3.6, the cells were plated out on LB agar supplemented with 100 µg/ml ampicillin and incubated overnight at 37°C. No colonies were seen on the plate.

In order to ascertain whether the restriction enzyme stocks used to digest the vector and insert had not lost their activity, separate digests of the vector were carried out using XmaI and BamHI (Figure 5-10A). This digest showed that both XmaI and BamHI were capable of significant linearisation of pUK-19, and therefore should be generating the complementary overhangs for ligation to the vector. A test to verify the T4 DNA ligase activity was carried out on cut λ-phage DNA. This showed that the T4 DNA ligase was functional as the treated λ-DNA showed ligation of smaller fragments into one large species (Figure 5-10B).

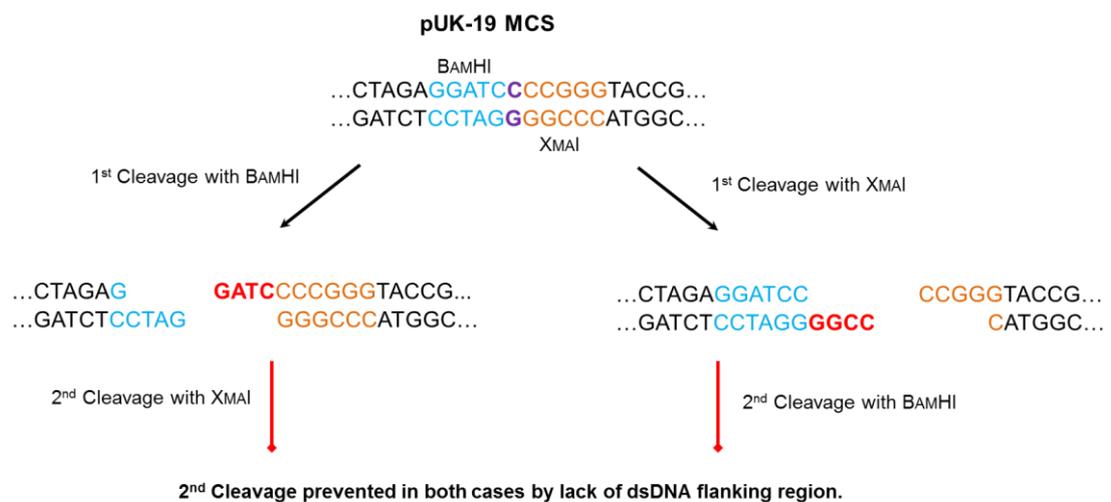


**Figure 5-10: pUK-19-*spoIIE* cloning reagent activity tests.**

(A) Test digests of pUK-19 using XmaI and BamHI. Lane 1; Hyperladder I (ladder sizes to the left), lane 2; uncut pUK-19, lane 3; pUK-19 after digestion with BamHI, lane 4; pUK-19 after digestion with XmaI. (B) Ligase activity test on cut DNA from the 48 kb λ-phage. Lane 1; Cut λ-phage DNA, lane 2; Ligated λ-phage DNA, lane 3; Hyperladder I (ladder sizes to the right).

As all reagents were shown to be functional, fresh insert was produced by PCR and the cloning process repeated with variations in vector to insert ratio during ligation of 1:1, 1:3, 1:5 and 1:10. No colonies were seen on the LB agar plates after transformation.

While a small proportion of restriction enzymes will cleave substrate DNA when their recognition site is at the DNA terminus, many require a 1-6 nucleotide flanking region of double stranded DNA on either side of their recognition site [233]. Closer analysis of the pUK-19 multiple cloning site suggested that cleavage with either restriction enzyme would cause a disruption in the recognition site for the other due to a 1 base recognition sequence overlap. Cleavage by either enzyme removes all flanking dsDNA from the second cleavage site (Figure 5-11). BamHI has been found to show more than 50% reduction in cleavage efficiency of sequences with fewer than 3 flanking bases.

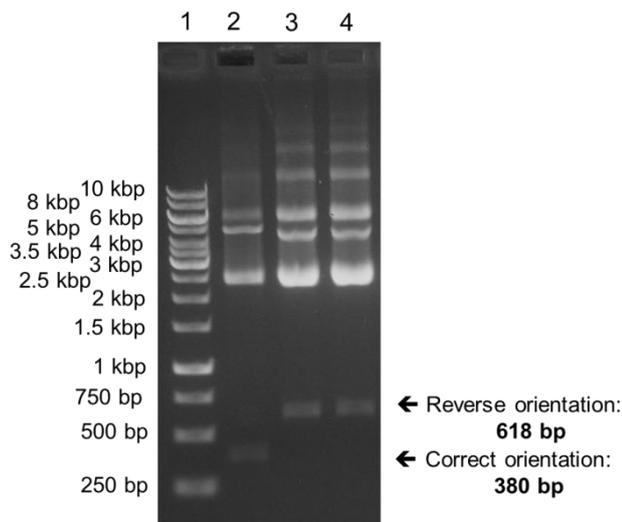


**Figure 5-11: Restriction Endonuclease flanking region conflict between BamHI and XmaI in the pUK-19 multiple cloning site.**

BamHI and XmaI recognition sequences are coloured blue and orange, respectively. The nucleotide belonging to both sequences is coloured purple. Single stranded DNA overhangs are coloured red.

A new cloning strategy was therefore devised in which the desired *spoIIE* fragment is generated from a full-length *spoIIE* harbouring pET-11a expression construct. 300 ng of pET-11a-*spoIIE* was digested using BglII and BamHI to yield a 952 bp fragment. This fragment was purified by agarose gel

electrophoresis followed by gel extraction purification, and ligated into dephosphorylated BamHI linearised pUK-19 vector at an insert to vector ratio of 3:1 and digestion products were used to transform competent *E. coli* XL-10 Gold cells as described in section 2.3.6. Plasmid DNA from obtained colonies was isolated by miniprep, described in section 2.2. The ligation reaction could have produced plasmid with this insert in either orientation as all produced DNA termini possess the same 5'-GATC overhang. The recombinant plasmid was expected to possess two EcoRI recognition sequences; one from the vector and one from the insert. The internal EcoRI site is closer to the 3' end of the *spoIIE* coding sequence and hence the direction of insert ligation into the plasmid could be determined from the fragment sizes of the products of EcoRI digestion. Test digestion with EcoRI (Figure 5-12) showed that one recombinant pUK-19 vector had the desired orientation, exhibiting an excised band of approximately 380 bp; the remaining two plasmids exhibited excised bands of around 600 bp, indicating that the insert had been ligated in the opposite orientation. DNA sequencing of the region between the M13 primer binding sites confirmed the correct orientation of the *spoIIE* sequence into the pUK-19 plasmid which showed the EcoRI excision band at 380 bp.

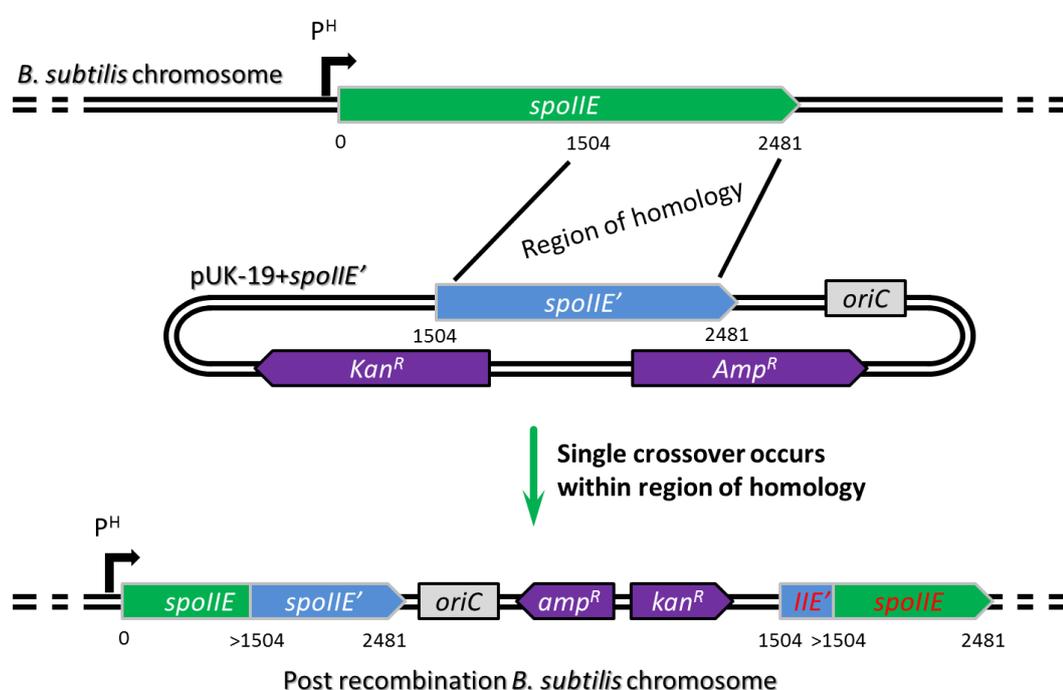


**Figure 5-12: Digestion to ascertain *spoIIE* fragment presence and orientation in recombinant pUK-19 plasmids.**

Lane 1; GeneRuler™ 1kb DNA Ladder, lanes 2-4; EcoRI digests of isolated recombinant pUK-19 plasmids.

### 5.3.3 Preparation of *B. subtilis* SpoIIIE Mutant strains

The plasmid harbouring the desired *spoIIIE* insert was mutated by site directed mutagenesis (SDM) to produce the target single residue substitutions as described in section 2.4. The nucleotide base exchanges were as follows: K649T; AAA to ACG at base 1945, I650A; ATC to GCC at base 1948, L695W; TTG to TGG at base 2083, V697A; GTT to GCC at base 2089 and V728A; GTG to GCC at base 2182. Mutagenic primer pairs were designed using the online PrimerX tool [234]. Mutations were checked by DNA sequencing using M13\_forward and M13\_reverse primers. All of the desired mutations had been successfully introduced.



**Figure 5-13: A schematic diagram of the genomic reorganisation caused by recombination of the *spoIIIE* mutant harbouring pUK-19 plasmid.**

The *B. subtilis* chromosome harbours a single copy of *spoIIIE* (green), the transcription of which is dictated by a  $\sigma^H$  specific promoter ( $P^H$ ). The pUK-19+*spoIIIE'* plasmid harbours *spoIIIE'* (blue), a mutated gene fragment of *spoIIIE* encompassing base pairs 1504-2481. Due to linearisation of the plasmid DNA on uptake, recombination can initiate at any point in the indicated region of homology. There is therefore the chance that mutants can be missed during recombination. The reorganised genome after recombination possesses a  $P^H$ -controlled, modified copy of the complete *spoIIIE* gene, the upstream and downstream portions of which are from genomic *spoIIIE* and exogenous *spoIIIE'*, respectively. The entire sequence of the recombined plasmid is incorporated into the genome, resulting in inclusion of the ampicillin ( $amp^R$ ) and kanamycin ( $kan^R$ ) resistance cassettes as well as the exogenous *oriC*. A deactivated gene fragment of *spoIIIE* without a promoter is also present.

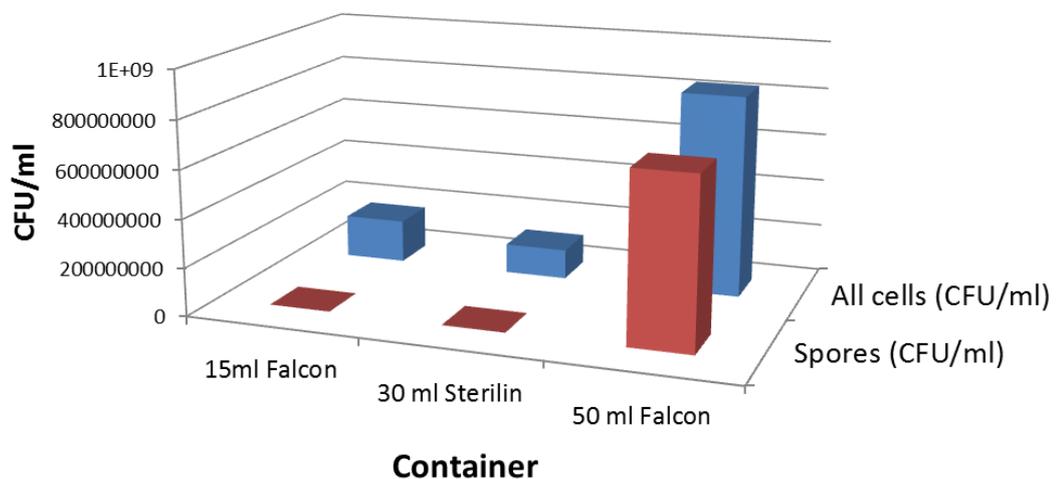
The efficient transformation of *B. subtilis* is aided by a low population of multimeric plasmid species. Plasmids isolated from *RecA*<sup>+</sup> strains of *E. coli* have

a much reduced proportion of multimeric forms compared to those isolated from *RecA1* (*RecA*<sup>-</sup>) strains. The *E. coli* XL-10 Gold strain has the *RecA1* phenotype and hence plasmids isolated from these cells will produce multimeric plasmids and have low transformation efficiency into *B. subtilis*. In order to achieve optimum transformation efficiency, the *spoIIE* mutant containing plasmids were used to transform *E. coli* MM294 competent cells, which have the *RecA*<sup>+</sup> phenotype, and re-isolated by miniprep as described in section 2.2.

Plasmids harbouring mutant *spoIIE* fragments, isolated from *E. coli* MM294, were used to transform *B. subtilis* IB333 competent cells as described in section 5.2.3. Transformants were plated using TOP agar onto LB agar plates supplemented with 10 µg/ml kanamycin and incubated overnight at 37°C. Transformations carried out using all plasmids were successful with upwards of 50 colonies on each plate indicating that the *B. subtilis* competent cells had successfully taken up and integrated the applied mutant *spoIIE* plasmid (as depicted in Figure 5-13). Stocks of the new mutant strains (listed in Table 5-1) were created by overnight growth in 10 ml LB cultures supplemented with 10 µg/ml kanamycin at 37°C. 800 µl of this culture was mixed with 200 µl 80% glycerol and frozen at -80°C. The genomes of the mutant strains were not sequenced to ensure mutations were present in the active *spoIIE* ORF. However, the 441 base pair region of wild-type *spoIIE* sequence upstream of the first amino acid substitution, K649T, is sufficient to make mutant allele inclusion very likely. Future experimentation would be required to absolutely confirm this.

### 5.3.4 Determination *spoIIIE* Mutant Sporulation Efficiency

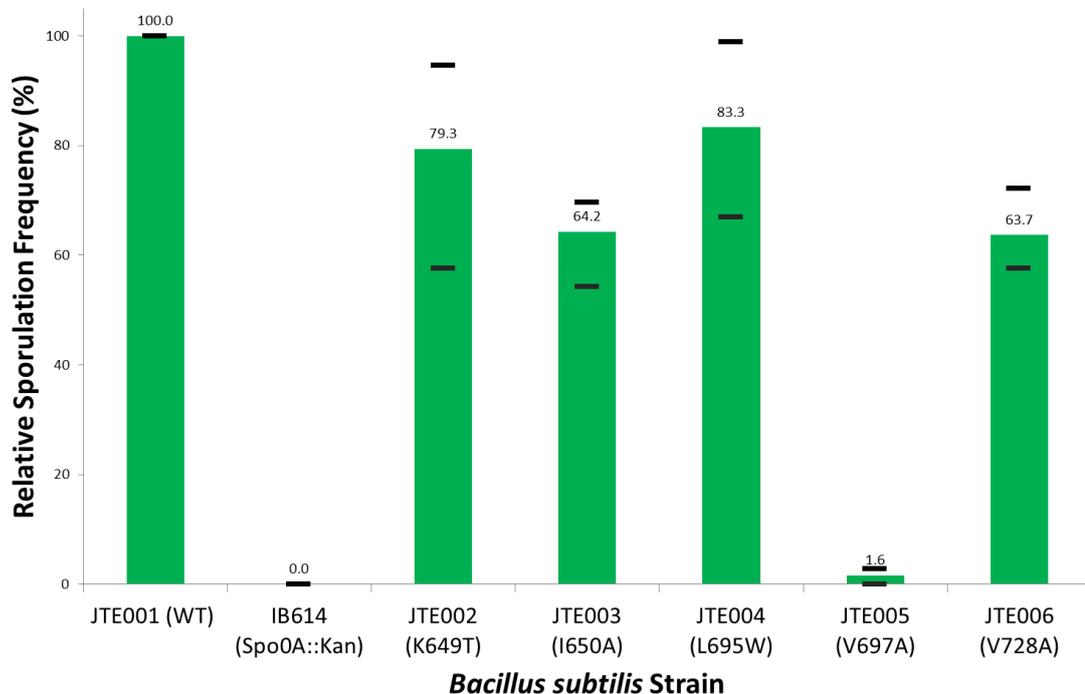
5 ml cultures of DSM in 30 ml Sterilin tubes inoculated with IB333, IB614 and JTE001-6 strains of *B. subtilis* were induced to sporulate and their sporulation efficiency was assayed as described in section 5.2.5. The data obtained for the wild type strain, IB333, showed that of  $2.5 \times 10^8$  CFU/ml in the DSM culture, only  $3.7 \times 10^6$  CFU/ml were due to heat-resistant spores. This suggests a sporulation efficiency of around 1.46%, much lower than the expected value of greater than 70% provided by Imrich Barak (personal communication). It was thought that the severely reduced sporulation efficiency might be due to incorrectly prepared DSM media or errors introduced during serial dilution of the culture, however rigorous examination of these factors showed that the sporulation efficiency of the IB333 wild type strain remained at between 0.1 and 2%. However, an investigation into the effect of container size, and hence, aeration of the culture showed that the low sporulation efficiency was exhibited in DSM sporulation cultures grown in 15 ml Falcon tubes and 30 ml Sterilin tubes, but not in 50 ml Falcon tubes in which sporulation efficiency was found to be 83% (Figure 5-14). The investigation showed that cultures grown in 15 ml and 30 ml containers showed an almost 10-fold reduction in measured vegetative colony forming units and a 3,000-fold lower spore count than those grown in 50 ml tubes.



**Figure 5-14: The effect of culture container size on sporulation efficiency of the wild type IB333 strain of *B. subtilis*.**

The JTE001-6 *spoIIE* mutant strains and the IB614, *spo0A* knockout, strain of *B. subtilis* were induced to sporulate in 50 ml Falcon tubes and assayed for sporulation efficiency, as described in section 5.2.5.

The JTE001 strain was the product of integration of the unmutated pUK-19-*spoIIE* plasmid into the chromosome; this strain has acquired kanamycin resistance without mutations in the *spoIIE* gene. This strain was used as a positive control and all sporulation efficiencies reported are normalised to the values obtained for this strain. The *spo0A* knockout strain, IB614, was used as a negative control. This strain has the *spo0A* gene interrupted by a kanamycin antibiotic resistance gene resulting in the entry to sporulation being inhibited at the earliest stage. Sporulation efficiencies measured in triplicate for the *spo0A* knockout strain were between  $5 \times 10^{-4} \%$  and  $1 \times 10^{-2} \%$  relative to the JTE001 wild type strain which consistently formed a fraction of spores/ml to total CFU/ml above 70 % between measurements. The relative sporulation efficiencies recorded for the 5 mutant *spoIIE* harbouring strains are shown in Figure 5-15.



**Figure 5-15: Sporulation efficiencies of *spoIIE* mutant harbouring strains of *B. subtilis* relative to wild type.**

The green histogram shows the mean value obtained for sporulation efficiency across three experiments. Black bars indicate the minimum and maximum recorded relative sporulation efficiency recorded for each strain.

The strain harbouring the V697A mutant, JTE005, showed sporulation efficiencies between 0% and 2.6% of wild type. These values are consistent with those expected for this mutant (Figure 5-1). The sporulation defects exhibited by the 4 other mutants; K649T, I650A, L695W and V728A, are comparatively mild showing reductions in efficiency to between 63.7% and 83.3% of wild type. These measured reductions in sporulation efficiency are not significant as the sporulating cultures are still producing above  $1 \times 10^8$  spores/ml. A reduction of an order of magnitude or more would suggest a more than a slight loss of efficiency in SpoIIE function due to the mutants. The observed reductions in sporulation efficiency compared to wild type are sufficient to indicate that recombination events including the desired mutations in the chromosomal copy of *spoIIE* have occurred and that recombination events were not initiated downstream of the mutant codons which would result in chromosomal wild-type *spoIIE* (as described by Figure 5-13).

## 5.4 Discussion

The data recorded here show that the S361F suppressor mutations in *spoIIE* domain III; K649T, I650A, L695W and V728A, are not sufficient to show a strong sporulation defect when present alone. In contrast, the Q483A suppressor mutation in *spoIIE* domain III, V697A, exhibits a strong sporulation defect on the order of a 100 fold reduction in sporulation efficiency, as previously reported [219].

The hypothesis that the  $\alpha 1$ - $\alpha 2$  region of the PP2C domain acts as an intragenic regulatory interface is not challenged by the data collected in this investigation. Instead, it suggests that the defects in SpoIIE function caused by the S361F mutation require only mild changes to correct the defect in efficient sporulation. On the other hand the Q483A mutation is only suppressed by a much more severe response in the form, V697A a hyperactive phosphatase activity mutant.

The SpoIIE V697A mutation has been shown to cause hyperactivation of the PP2C phosphatase domain which results in uncompartimentalised  $\sigma^F$  activation. Q483A on the other hand causes a complete loss of phosphatase activity and, hence, no  $\sigma^F$  activation. These phenotypes are diametrically opposed which may explain why V697A suppresses Q483A.

In contrast to Q483A, S361F has been shown to be capable of SpoIIAA~P dephosphorylation but it has been proposed that it is unable to release the product, SpoIIAA. This observation prompted the notion that wild type SpoIIE is capable of dephosphorylating SpoIIAA~P at the septum, but is prevented from releasing the product until the asymmetric septum is completed. This would require a signal to be passed through SpoIIE domain II to effect release of the product, this signal is presumably inhibited in S361F and recovered in the mutants tested in this investigation.

## Chapter 6: Conclusions and Future Work

### 6.1 The Behaviour of SpoIIE Domain II

Three new soluble fragments of SpoIIE have been isolated. A key objective in the creation of these fragments was the elucidation of the structure of SpoIIE domain II. Unfortunately, structural studies have been stymied by ubiquitously unsuccessful protein crystallisation experiments. It was, however, observed that the oligomerisation seen in some fragments containing domain II elements can be attributed to the sequence between residues 334 and 375. SpoIIE<sub>457-827</sub> has been found to be monomeric in solution; however, SpoIIE<sub>334-827</sub> has been identified using AUC as forming oligomers which seem to have periodic stacking of hexamers.

It is possible that this oligomerisation has a functional role. A SpoIIE mutant isolated by Niels Bradshaw, Harvard (personal communication), K356D, has been observed to reduce the frequency of asymmetric septation and severely inhibit the activation of  $\sigma^F$  during sporulation *in vivo*. A more interesting feature of this mutant is that it shows monomeric, rather than oligomeric, character in solution when expressed as part of a SpoIIE<sub>320-827</sub> fragment. It is likely that multimeric SpoIIE plays a part in function at the cell division septum in addition to the expected interaction with parts of the cell division machinery.

A difference between the oligomeric states of *Bsu*SpoIIE<sub>334-827</sub> and *Gst*SpoIIE<sub>335-826</sub> was observed. *Gst*SpoIIE<sub>335-826</sub> shows a much reduced tendency to form large oligomers, instead favouring monomer/dimer equilibrium. As the two orthologues possess a high degree of sequence identity, the cause of this change in behaviour is unknown. However, conspicuously, K356 is not conserved and is instead arginine in the *Gst*SpoIIE sequence.

SpoIIE domain II containing fragments showed a tendency to precipitate on addition of Mn<sup>2+</sup>. This does not seem to be an exacerbation of the oligomeric character shown by SpoIIE<sub>334-827</sub> as the removal of trace manganese using EDTA had no discernible effect on this trait. A putative metal binding site was identified on analysis of the SpoIIE sequence using a web based tool, Metal

Detector. The observed behaviour cannot be ascribed to this however, as even fragments without the cysteine residues implicated, e.g. SpoIIE<sub>457-827</sub>, exhibit it.

Considering the disparate behaviour exhibited by domain II containing SpoIIE fragments it seems likely that domain II might be made of two smaller sub domains: an oligomerisation inducing domain between residues 334 and 457 and a subdomain which might hold similarities to SpoIIIAH from residue 457 to the N-terminal limit of the phosphatase domain at residue 590. Amino acid substitutions in each of these regions (S361F and Q483A) result in defects in the correct activation of  $\sigma^F$  and these regions can therefore be expected to play a role in regulation of the phosphatase domain.

## 6.2 SpoIIE Phosphatase Activity

SpoIIE is a key player in the instatement of the complex sporulation sigma factor relay, a co-ordination of asymmetric gene expression between two cells; the mother cell and forespore. It is crucial to sporulation that  $\sigma^F$ , the first compartment specific sigma factor, is activated only when asymmetric septation is complete at which point separate patterns of gene expression can be embarked upon. SpoIIE is responsible for activating  $\sigma^F$  by dephosphorylation of SpoIIAA. However, this phosphatase activity must be under regulation until cell division occurs as uncompartimentalised  $\sigma^F$  activity results in abortive sporulation. How this regulation is imposed is unknown.

The work carried out here has briefly looked at the relative rates of some SpoIIE fragments in dephosphorylation of SpoIIAA. The observations made have indicated that fragments of SpoIIE with elements of domain II are less active than a fragment consisting of only the phosphatase domain.

Elucidation of the structure of a SpoIIE:SpoIIAA~P complex describing the dephosphorylation reaction has not been successful. The structure of such a phosphatase:phosphoprotein substrate complex has, to this author's knowledge, never been solved.

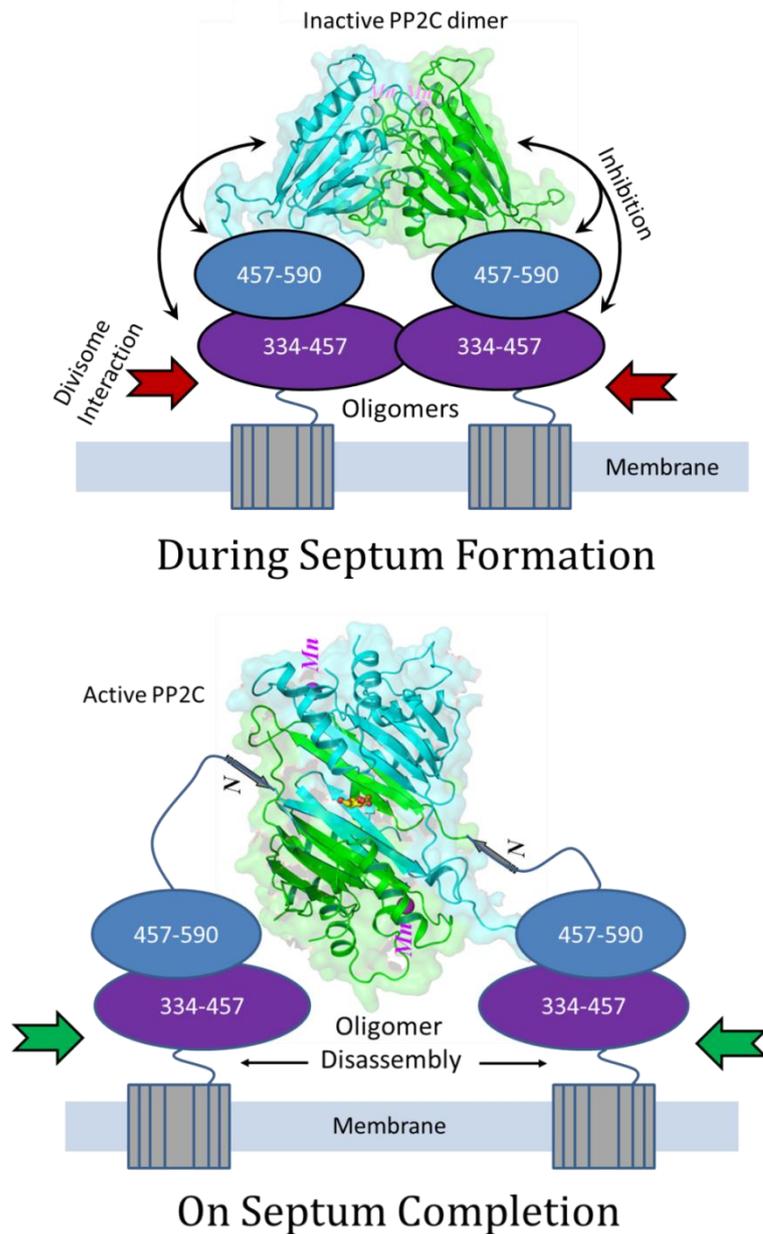
It has been noted throughout studies of SpoIIE fragments that the affinity of the active site for  $Mn^{2+}$  is low. Over 100  $\mu M$   $MnCl_2$  is required for maximum phosphatase activity of SpoIIE<sub>590-827</sub>. In addition, the X-ray crystal structure of this fragment shows only a single manganese atom in the active site. This is in contrast to the 2 or 3 manganese atoms observed in structures of other PP2C phosphatases. It could be that low manganese affinity provides an inhibitory mechanism to phosphatase activity and is reinstated on completion of the asymmetric septum or that full length SpoIIE is the species which promotes full  $Mn^{2+}$  affinity. The second explanation is supported by previous work [216].

A second crystal structure (Vladimir Levdikov, unpublished) has been solved showing a dimer in which the active sites are proximate and occluded. No manganese ions are seen in this structure. It is possible that this dimer is representative of an inhibited phosphatase domain.

### 6.3 An Interface for Intragenic Regulation of the Phosphatase Domain

Strains of *B. subtilis* which harboured mutations in the phosphatase domain of SpoIIE known to suppress domain II mutation phenotypes showed no significant defects in sporulation. A likely explanation for this is that these mutations behave only as corrections and are not prohibitive to normal function in the absence of S361F. One hypothesis which has been presented based on observation of SpoIIE(S361F) is that wild type SpoIIE's phosphatase activity is not inhibited but release of dephosphorylated SpoIIAA is prevented. It is therefore suggested that the signal from completion of asymmetric cell division serves to allow release of SpoIIAA which has been turned over. The suppression of the S361F phenotype could simply be due to a reconstitution of this release mechanism, through modification of the phosphatase domain.

## 6.4 How PP2C regulation might be achieved



**Figure 6-1: A hypothesis for regulation of SpoIIE phosphatase activity.**

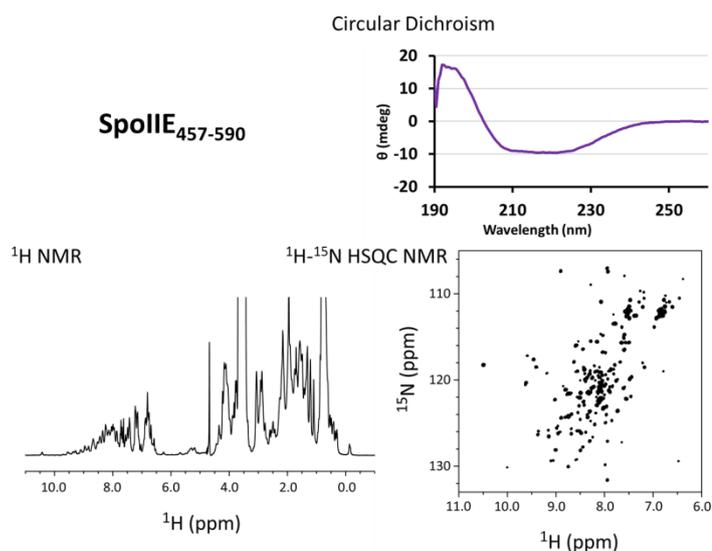
Before and during septum formation, SpoIIE phosphatase activity is held inactive by inhibition from domain II which in turn is controlled by an interaction with the divisome. On septum completion, the interaction controlling inhibition is removed or modified resulting in the relief of phosphatase activity inhibition. The oligomeric state of SpoIIE is likely to play a key role in regulation of phosphatase activity.

The evidence presented in previous studies of SpoIIE indicates that SpoIIE's phosphatase activity is held in check until cell division is complete. They also support a mechanism where the phosphatase domain is inhibited by domain II.

A signal for relief of this inhibition could be passed through the cell division machinery. Based on the two crystal structures obtained of the phosphatase domain, it is hypothesised by the author that an inactive and active form exists and that this state is controlled by interaction with domain II (Figure 6-1). The oligomerisation observed for SpoIIE fragments containing regions of domain II could be a functional element of this regulation, especially if the oligomerisation state of SpoIIE can be controlled through an interaction with the divisome. The significance of  $Mn^{2+}$  dependent precipitation of SpoIIE is a mystery. However, a feature of the oligomerisation of SpoIIE could be the storage of manganese to ensure a high local concentration when inhibition of the phosphatase domain is relieved.

## 6.5 Future work

Recent work has produced a new soluble fragment of SpoIIE: SpoIIE<sub>457-590</sub>. Crystallisation attempts of this fragment have not been successful, however the fragment is small enough (14 kDa) to consider structural characterisation by NMR. To date, the fragment has been shown to be structured by CD and both  $^1H$  one dimensional and  $^1H/^{15}N$  HSQC NMR experiments have been carried out. These preliminary results show promise (Figure 6-2) and a solution state structure of this domain could therefore be solved in the not too distant future.



**Figure 6-2: Preliminary data gathered on SpoIIE<sub>457-590</sub>, a target for NMR structural characterisation.**

Another fragment representing only SpoIIE domain II (SpoIIE<sub>334-590</sub>) has also been successfully expressed recently as an MBP (maltose binding protein) fusion. While the requirement for chimeric construct suggests the fragment might be heavily oligomerised, this is yet another fragment with which to study the behaviour of SpoIIE.

The behaviour observed for the SpoIIE(K356D) mutant by Niels Bradshaw provides an interesting line of inquiry for structural characterisation. If this substitution is applied to the SpoIIE<sub>334-827</sub> fragment, and also now potentially MBP-SpoIIE<sub>334-590</sub>, it should promote a monomeric state which will be a much more promising prospect for crystallisation.

The currently used native gel assay for determining phosphatase activity provides highly qualitative results which are based upon visual analysis of relative band intensities corresponding to SpoIIAA and SpoIIAA~P. The establishment of a more quantitative assay would allow a more rigorous approach to determination of the kinetics involved in dephosphorylation of SpoIIAA~P by SpoIIE. Two colorimetric assays are commercially available for the detection of free phosphate, another stoichiometric product of the desphosphorylation reaction: PiColorLock Gold (Innova Biosciences), a malachite green end point assay; and EnzChek (Molecular Probes), a real time assay utilising purine nucleoside phosphorylase. Either could be used effectively to gain a more comprehensive understanding of the enzyme kinetics specific to different fragments of SpoIIE.

The data presented in this thesis represent an important advance in terms of the production of soluble fragments of the cytosolic domains of SpoIIE for functional and structural studies. The identification of the TM:Domain 2 solubility boundary at residue 334 will allow future investigators to approach structural characterisation of domain 2 with fragments representing the whole domain. The characterisation of fragments in this work also means that oligomerisation observed amongst SpoIIE domain 2-containing fragments can now be attributed to the residues between 334 and 457. Studies of the dephosphorylation reaction between SpoIIE fragments and SpoIIAA~P have indicated an inhibitory effect by SpoIIE domain 2 on the activity of the PP2C

phosphatase domain. This observation supports the hypothesis that PP2C activity is held in check through interaction with domain 2 before asymmetric septation is complete.

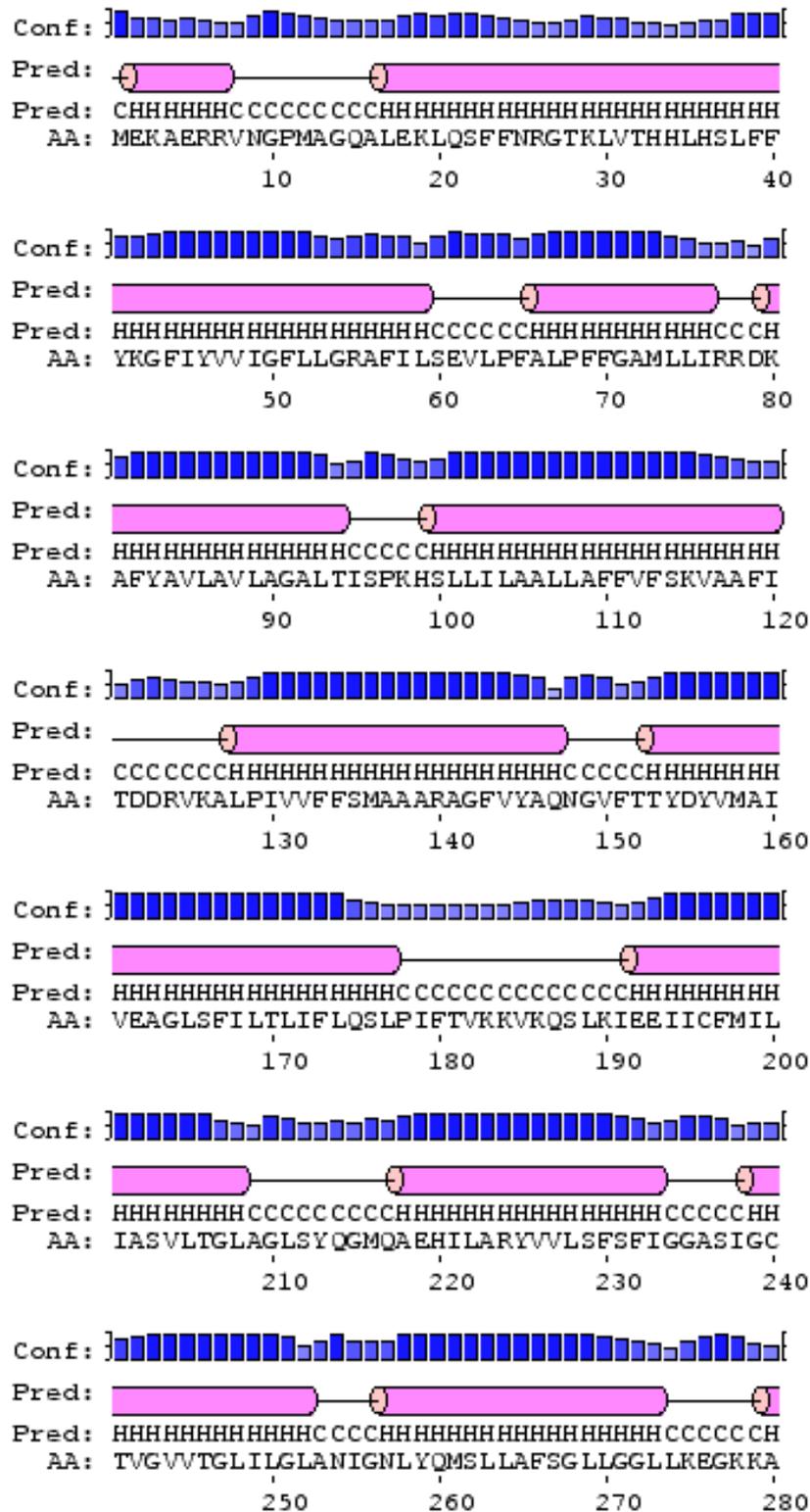


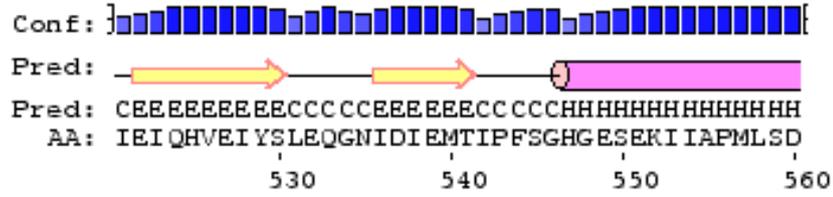
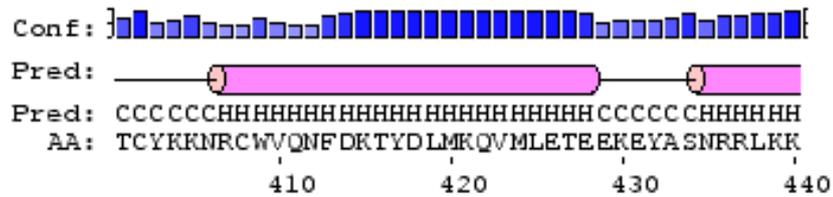
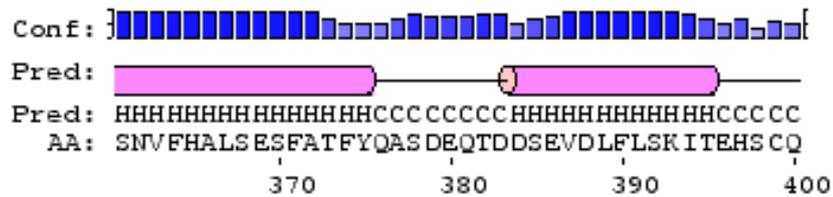
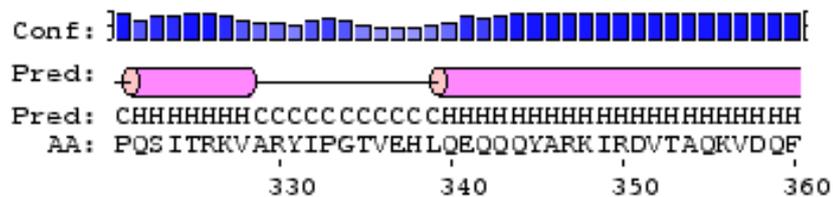
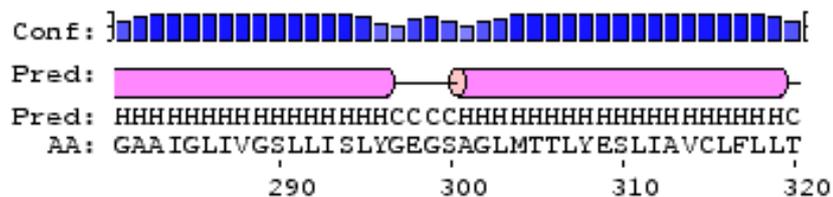
## Fragment Parameters (calculated using ProtParam (ExPASy))

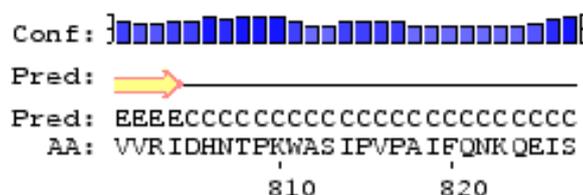
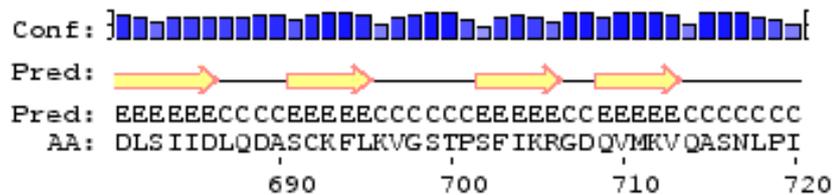
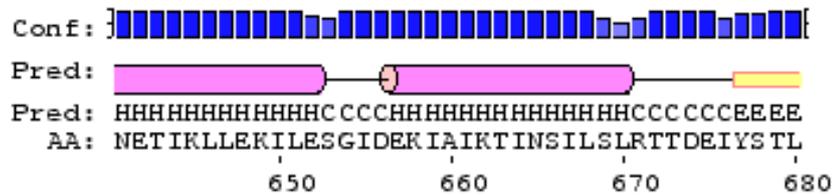
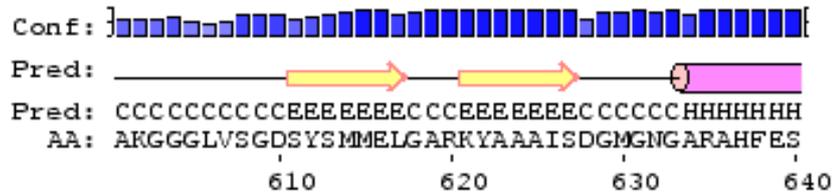
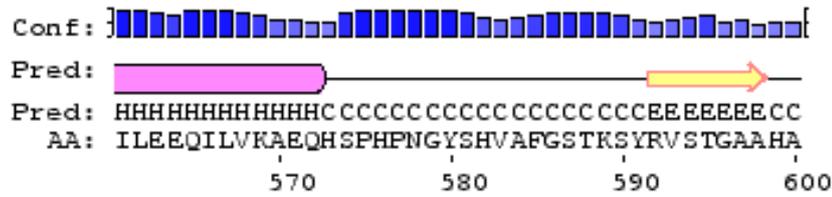
Protein	Fragment	Cleaved 6xHis Tag	Theoretical MW (Da)	Theoretical pI	$\epsilon(A_{280})$ ( $M^{-1} \text{ cm}^{-1}$ )	
<i>Organism: B. subtilis</i>						
SpoIIE	590-827	No	28654	5.76	16960	
		Yes	26508	5.32	16960	
	334-827	No	58540	5.49	34630	
		Yes	56393	5.32	34630	
	457-827	No	43888	5.52	19940	
		Yes	41742	5.29	19940	
	375-590	No	25522	5.23	16180	
		Yes	27669	5.52	16180	
	SpoIIAA	-	No	15361	6.28	1615
		-	Yes	13215	5.90	1615
	SpoIIAB	-	N/A	16335	4.58	4595
	<i>Organism: G. stearothermophilus</i>					
SpoIIE	335-826	No	58224	5.18	50310	
	457-826	No	47378	5.25	36245	
	457-826	Yes	45231	5.04	36245	

# SpolIE Secondary and Tertiary Structure Predictions

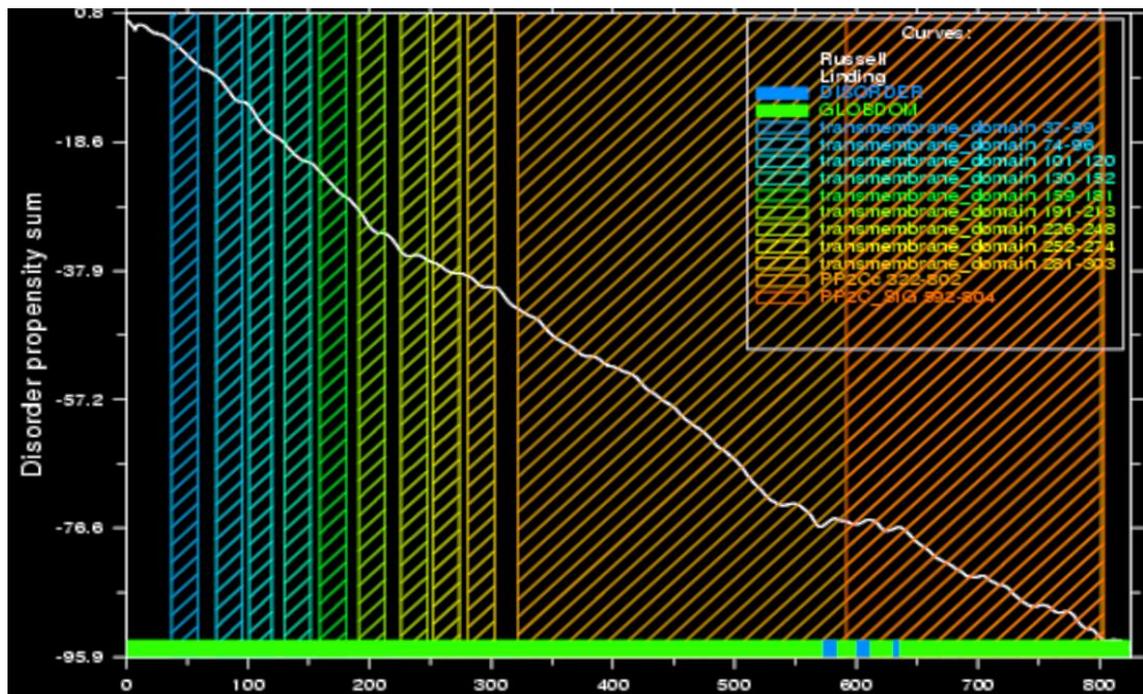
## PsiPred (Secondary Structure)



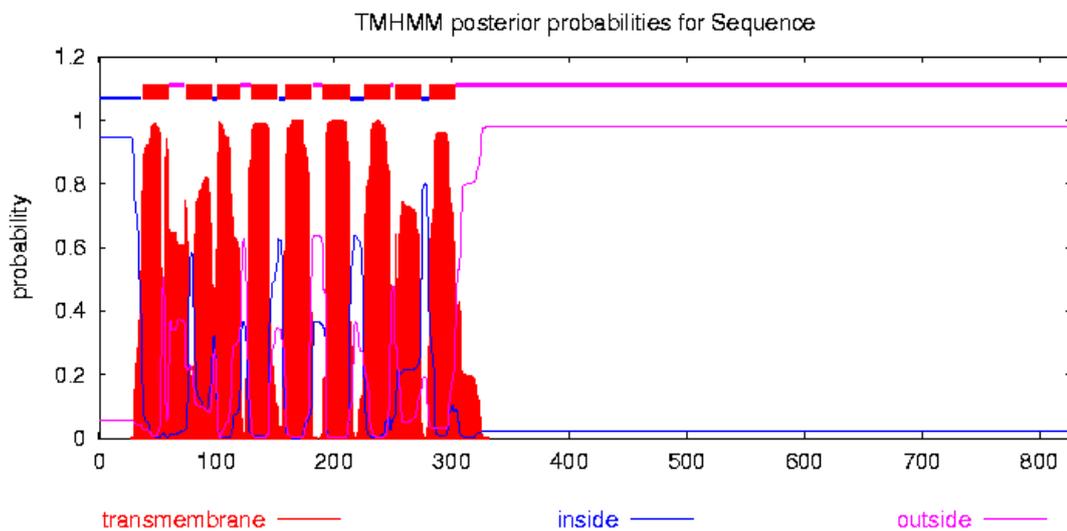




## GlobPlot 2.3 (Globular Domains and Transmembrane Helices)



## TMHMM predictor (Transmembrane Helices)



## Oligonucleotide Primer Sequences

### Cloning

#### ***Bsu*SpoII $\text{E}$**

*Tm* : Domain II boundary fragment

IIE_300_3C_FWD	CCAGGGACCAGCAATGTCTGCCGGTCTGATG
IIE_316_3C_FWD	CCAGGGACCAGCAATGCTGTTTTTGCTCACACC
IIE_321_3C_FWD	CCAGGGACCAGCAATGCCTCAATCTATTACGAGG
IIE_326_3C_FWD	CCAGGGACCAGCAATGAGGAAAGTGGCGAG
IIE_334_3C_FWD	CCAGGGACCAGCAATGGGAACTGTCGAGCATC
IIE_339_3C_FWD	CCAGGGACCAGCAATGCTTCAAGAGCAACAGC
IIE3C457_FWD	CCAGGGACCAGCAATGGAGGATGAGCTCGCACATCATCATGC

#### ***Gst*SpoII $\text{E}$**

GSII $\text{E}$ 3C335FWD	CCAGGGACCAGCAATGGGCACGGCAGAATATATC
GSII $\text{E}$ 3C457FWD	CCAGGGACCAGCAATGATTGAAAAAGAGATGGTTG
GSII $\text{E}$ 826LICRV	GAGGAGAAGGCGCGTTATTGCGCCTTTTTTCATATAC

#### ***Bsu*SpoII $\text{A}$ + *Bsu*SpoII $\text{A}$ B**

S2AX_0_LICp_FWD	CCAGGGACCAGCAATGAGCCTTGAATTGACATGAATG
S2AA_LIC_REV	GAGGAGAAGGCGCGTCATGATGCCACCCCCAGTGT
S2AAB_LIC_REV	GAGGAGAAGGCGCGTTAATTACAAAGCGCTTTGCTTTTTG

## Mutagenesis

### ***Bsu*SpoII E**

#### In vivo investigation mutagenic primers

S2E_V697A_FWD	GTAAATTTTTGAAGCCGGATCGACGCCAG
S2E_V697A_REV	CTGGGCGTCGATCCGGCCTTCAAAAATTTAC
S2E_L695W_FWD	CCAGCTGTAAATTTTGGAAAGTTGGATCGAC
S2E_L695W_REV	GTCGATCCAACCTTCCAAAATTTACAGCTGG
S2E_V728A_FWD	GTATTATTAATGAATTCGATGCCGAGGTTGTGAG TGAACAGCTG
S2E_V728A_REV	CAGCTGTTCACTCACAACTCGGCATCGAATTCAT TAATAATAC
S2E_I650A_FWD	CAAGCTTCTTGAAAAAGCCCTTGAATCGGGCATTG
S2E_I650A_REV	CAATGCCCGATTCAAGGGCTTTTTCAAGAAGCTTG
S2E_K649T_FWD	CGATCAAGCTTCTTGAAACGATCCTTGAATCGGGCATTG
S2E_K649T_REV	CAATGCCCGATTCAAGGATCGTTTTCAAGAAGCTTGATCG

### ***Bsu*SpoII A**

S2AA-S58D_FWD	CTTTCCTTTATGGACGACTCGGGGCTTGCG
S2AA-S58D_REV	CGCCAAGCCCCGAGTCGTCCATAAAGGAAAG
S2AA-S58E_FWD	GAGGACCTTTCCTTTATGGACGAATCGGGGCTTGG CGTTATTTTAG
S2AA-S58E_REV	CTAAAATAACGCCAAGCCCCGATTCGTCCATAAAG GAAAGGTCCT

## Solutions for generation of *B. subtilis* Competent cells

### **T base**

0.20% (w/v)  $(\text{NH}_4)_2\text{SO}_4$

1.83%(w/v)  $\text{K}_2\text{HPO}_4 \cdot 3\text{H}_2\text{O}$

0.60% (w/v)  $\text{KH}_2\text{PO}_4$

0.10% (w/v) trisodium citrate.2H<sub>2</sub>O

### **SpC**

20 ml T base

0.2 ml 50% (w/v) glucose

0.3 ml 1.2% (w/v)  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$

0.4 ml 10% (w/v) bacto yeast extract

0.5 ml 1% (w/v) casamino acids

### **SpII**

100 ml T base

1 ml 50% (w/v) glucose

7 ml 1.2% (w/v)  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$

1 ml 10% (w/v) bacto yeast extract

1 ml 1% (w/v) casamino acids

0.5 ml 0.1 mol/l  $\text{CaCl}_2$

## Liquid Difco Sporulation Medium (DSM) a.k.a. Schaeffer's Sporulation Medium

Recipe for 1L of DSM Base [231]:

8 g	Bacto-nutrient broth
10 ml	10% (w/v) KCl
10 ml	1.2% (w/v) MgSO <sub>4</sub> ·7H <sub>2</sub> O
0.5 ml	NaOH

Additional solutions are to be autoclaved separately and added to autoclaved DSM base immediately before use of the DSM:

1 ml	1 M Ca(NO <sub>3</sub> ) <sub>2</sub>
1 ml	0.01 M MnCl <sub>2</sub>
1 ml	1 mM FeSO <sub>4</sub>

When autoclaved, FeSO<sub>4</sub> forms a pinkish precipitate, this precipitate should be resuspended by thorough vortexing prior to addition to DSM base.

## List of abbreviations

aas	Amino Acids
ADP	Adenosine di-phosphate
ATP	Adenosine tri-phosphate
BCAA	Branched Chain Amino Acid
bp	Base pairs
<i>Bsph</i>	<i>Bacillus sphaericus</i>
<i>Bsu</i>	<i>Bacillus subtilis</i>
Da	Dalton
DNA	Deoxyribonucleic acid
dsDNA	Double Stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethyleneglycoltetraacetic acid
FPLC	Fast Protein Liquid Chromatography
GDP	Guanosine di-phosphate
GFP	Green Fluorescence Protein
<i>Gst</i>	<i>Geobacillus stearothermophilus</i>
GTP	Guanosine tri-phosphate
HRV3C	Human Rhinovirus 3C protease

IMAC	Immobilised Metal Affinity Chromatography
IIAA	SpoIIAA
IIAA~P	Phospho-SpoIIAA
IIAB	SpoIIAB
IIE	SpoIIE
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
kbp	Kilo-Base pairs
kDa	Kilo-Dalton
LD <sub>50</sub>	Median Lethal Dose
M	Molar
MDa	Mega-Dalton
mM	Millimolar
nM	Nanomolar
MW	Molecular Weight
MWCO	Molecular Weight Cut-Off
NMR	Nuclear magnetic resonance
PCR	Polymerase Chain Reaction
pI	Isoelectric point
PG	Peptidoglycan
RNAP	RNA Polymerase
SDS	Sodium Dodecyl Sulphate

spp.	Species
SRS	SpoIIIE Recognition Sequence
ssDNA	Single Stranded DNA
TM	Transmembrane
TTP	Thymidine tri-phosphate
WT	Wild-type
$\mu\text{M}$	Micromolar

## References

1. Tortora, G., B. Funke, and C. Case, *Microbiology: An Introduction, 9th Edition (Book & CD-ROM)*. 2006: Benjamin Cummings.
2. Flardh, K. and M.J. Buttner, *Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium*. *Nat Rev Micro*, 2009. **7**(1): p. 36-49.
3. Cano, R. and M. Borucki, *Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber*. *Science*, 1995. **268**(5213): p. 1060-1064.
4. Sansinenea, E., *The Most Important Bacillus Species in Biotechnology*, in *Bacillus thuringiensis Biotechnology*, E. Sansinenea, Editor. 2012, Springer Netherlands. p. 329-341.
5. HÖFte, H., et al., *Structural and functional analysis of a cloned delta endotoxin of Bacillus thuringiensis berliner 1715*. *European Journal of Biochemistry*, 1986. **161**(2): p. 273-280.
6. Vaeck, M., et al., *Transgenic plants protected from insect attack*. *Nature*, 1987. **328**(6125): p. 33-37.
7. Metchnikoff, E., *The prolongation of life*. Optimistic studies. Translated and edited by P. Chalmers Mitchell, 1907.
8. Fuller, R., *Probiotics in man and animals*. *J Appl Bacteriol*, 1989. **66**(5): p. 365-78.
9. Cutting, S.M., *Bacillus probiotics*. *Food Microbiology*, 2011. **28**(2): p. 214-220.
10. Medicine, U.o.P.S.o. 'Good' bacteria keep immune system primed to fight future infections. 2010, February 3; Available from: <http://www.sciencedaily.com/releases/2010/01/100127095945.htm>.
11. Arvola, T., et al., *Prophylactic Lactobacillus GG Reduces Antibiotic-Associated Diarrhea in Children With Respiratory Infections: A Randomized Study*. *Pediatrics*, 1999. **104**(5): p. e64.
12. Brady, L.J., D.D. Gallaher, and F.F. Busta, *The Role of Probiotic Cultures in the Prevention of Colon Cancer*. *The Journal of Nutrition*, 2000. **130**(2): p. 410.
13. Niedzielin, K., H. Kordecki, and B. Birkenfeld, *A controlled, double-blind, randomized study on the efficacy of Lactobacillus plantarum 299V in patients with irritable bowel syndrome*. *Eur J Gastroenterol Hepatol.*, 2001. **13**(10): p. 1143-7.
14. Strozzi, G. and L. Mogna, *Quantification of folic acid in human feces after administration of Bifidobacterium probiotic strains*. *J Clin Gastroenterol*, 2008. **42** (Suppl 3 Pt 2:S179-84).
15. Molina, V.C., et al., *Lactobacillus reuteri CRL 1098 prevents side effects produced by a nutritional vitamin B12 deficiency*. *Journal of Applied Microbiology*, 2009. **106**(2): p. 467-473.
16. Duc, L.H., et al., *Bacterial Spores as Vaccine Vehicles*. *Infection and Immunity*, 2003. **71**(5): p. 2810-2818.
17. Hoang, T.H., et al., *Recombinant Bacillus subtilis Expressing the Clostridium perfringens Alpha Toxoid Is a Candidate Orally Delivered Vaccine against Necrotic Enteritis*. *Infection and Immunity*, 2008. **76**(11): p. 5257-5265.

18. Lee, S., et al., *Development of a Bacillus subtilis-Based Rotavirus Vaccine*. Clinical and Vaccine Immunology, 2010. **17**(11): p. 1647-1655.
19. Permpoonpattana, P., et al., *Immunization with Bacillus Spores Expressing Toxin A Peptide Repeats Protects against Infection with Clostridium difficile Strains Producing Toxins A and B*. Infection and Immunity, 2011. **79**(6): p. 2295-2302.
20. Nguyen, V.A.T., et al., *Killed Bacillus subtilis spores expressing streptavidin: a novel carrier of drugs to target cancer cells*. Journal of Drug Targeting, 2013. **21**(6): p. 528-541.
21. Kotiranta, A., K. Lounatmaa, and M. Haapasalo, *Epidemiology and pathogenesis of Bacillus cereus infections*. Microbes and Infection, 2000. **2**(2): p. 189-198.
22. Roberts, T.A.B.-P., A. C.; Tompkin, R. B., *Characteristics of microbial pathogens*, in *Microorganisms in Foods 5*. 1996, London: Blackie Academic & Professional. p. 24.
23. Turnbull, P., *Bacillus*, Chapter 15, in *Medical Microbiology*, S. Baron, Editor. 1996, Galveston (TX): University of Texas Medical Branch at Galveston.
24. Ehling-Schulz, M., M. Fricker, and S. Scherer, *Bacillus cereus, the causative agent of an emetic type of food-borne illness*. Molecular Nutrition & Food Research, 2004. **48**(7): p. 479-487.
25. Ergmengem., E.v., *Über einen neuen anaeroben Bacillus und seine Beziehungen Zum Botulismus*. Zentralbl. Hyg. Infektionskr. , 1897. . **26**: p. 1-8.
26. Wells, C. and T. Wilkins, *Clostridia: Sporeforming Anaerobic Bacilli*, Chapter 18, in *Medical Microbiology*, S. Baron, Editor. 1996, Galveston (TX): University of Texas Medical Branch at Galveston.
27. *Pathogenic Bacteria*, in *Sherris Medical Microbiology: An Introduction to Infectious Diseases*, Ryan KJ and R. CG, Editors. 2004, McGraw Hill. p. 322-324.
28. Spencer, R., *Bacillus anthracis*. J Clin Pathol, 2003(56): p. 182-187.
29. Kolstø, A.-B., N.J. Tourasse, and O.A. Økstad, *What Sets Bacillus anthracis Apart from Other Bacillus Species?* Annual Review of Microbiology, 2009. **63**(1): p. 451-476.
30. Hong, H.A., et al., *Bacillus subtilis isolated from the human gastrointestinal tract*. Research in Microbiology, 2009. **160**(2): p. 134-143.
31. Kunst, F., et al., *The complete genome sequence of the Gram-positive bacterium Bacillus subtilis*. Nature, 1997. **390**(6657): p. 249-256.
32. Kobayashi, K., et al., *Essential Bacillus subtilis genes*. Proceedings of the National Academy of Sciences, 2003. **100**(8): p. 4678-4683.
33. Commichau, F.M., N. Pietack, and J. Stulke, *Essential genes in Bacillus subtilis: a re-evaluation after ten years*. Molecular BioSystems, 2013. **9**(6): p. 1068-1075.
34. Pohl, S. and C.R. Harwood, *Chapter 1 - Heterologous Protein Secretion by Bacillus Species: From the Cradle to the Grave*, in *Advances in Applied Microbiology*, S.S. Allen I. Laskin and M.G. Geoffrey, Editors. 2010, Academic Press. p. 1-25.
35. Harwood, C.R. and R. Cranenburgh, *Bacillus protein secretion: an unfolding story*. Trends in Microbiology, 2008. **16**(2): p. 73-79.
36. Molle, V., et al., *The Spo0A regulon of Bacillus subtilis*. Molecular Microbiology, 2003. **50**(5): p. 1683-1701.

37. Phillips, Z.E.V. and M.A. Strauch, *Bacillus subtilis sporulation and stationary phase gene expression*. Cellular and Molecular Life Sciences CMLS, 2002. **59**(3): p. 392-402.
38. Banse, A.V., et al., *Parallel pathways of repression and antirepression governing the transition to stationary phase in Bacillus subtilis*. Proceedings of the National Academy of Sciences, 2008. **105**(40): p. 15547-15552.
39. Vaughn, J.L., et al., *Novel DNA binding domain and genetic regulation model of Bacillus subtilis transition state regulator AbrB*. Nat Struct Mol Biol, 2000. **7**(12): p. 1139-1146.
40. Sullivan, D.M., et al., *Insights into the Nature of DNA Binding of AbrB-like Transcription Factors*. Structure, 2008. **16**(11): p. 1702-1713.
41. Chumsakul, O., et al., *Genome-wide binding profiles of the Bacillus subtilis transition state regulator AbrB and its homolog Abh reveals their interactive role in transcriptional regulation*. Nucleic Acids Research, 2011. **39**(2): p. 414-428.
42. Healy, J., et al., *Post-transcriptional control of a sporulation regulatory gene encoding transcription factor sigma H in Bacillus subtilis*. Mol Microbiol, 1991. **5**(2): p. 477-87.
43. Levnikov, V.M., et al., *Structural rearrangement accompanying ligand binding in the GAF domain of CodY from Bacillus subtilis*. J Mol Biol, 2009. **390**(5): p. 1007-18.
44. Molle, V., et al., *Additional Targets of the Bacillus subtilis Global Regulator CodY Identified by Chromatin Immunoprecipitation and Genome-Wide Transcript Analysis*. Journal of Bacteriology, 2003. **185**(6): p. 1911-1922.
45. Belitsky, B.R. and A.L. Sonenshein, *Genome-wide identification of Bacillus subtilis CodY-binding sites at single-nucleotide resolution*. Proceedings of the National Academy of Sciences, 2013. **110**(17): p. 7026-7031.
46. Berg, J.M., J.L. Tymoczko, and L. Stryer, *Biochemistry*, p. 830-834. 5th ed. 2007: W. H. Freeman & Co Ltd.
47. Levnikov, V.M., et al., *The Structure of CodY, a GTP- and Isoleucine-responsive Regulator of Stationary Phase and Virulence in Gram-positive Bacteria*. Journal of Biological Chemistry, 2006. **281**(16): p. 11366-11373.
48. Hoch, J.A., *Regulation of the Phosphorelay and the Initiation of Sporulation in Bacillus Subtilis*. Annual Review of Microbiology, 1993. **47**(1): p. 441-465.
49. Hoch, J., *Genetics of bacterial sporulation*. Advances in Genetics Vol. 18. 1976.
50. Strauch, M., et al., *The SpoOA protein of Bacillus subtilis is a repressor of the abrB gene*. Proceedings of the National Academy of Sciences, 1990. **87**(5): p. 1801-1805.
51. Zhao, H., et al., *DNA Complexed Structure of the Key Transcription Factor Initiating Development in Sporulating Bacteria*. Structure, 2002. **10**(8): p. 1041-1050.
52. McNicholas, S., et al., *Presenting your structures: the CCP4mg molecular-graphics software*. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 386-94.
53. Baldus, J.M., et al., *Phosphorylation of Bacillus subtilis transcription factor SpoOA stimulates transcription from the spoIIg promoter by enhancing binding to weak OA boxes*. Journal of Bacteriology, 1994. **176**(2): p. 296-306.

54. Tojo, S., K. Hirooka, and Y. Fujita, *Expression of kinA and kinB of Bacillus subtilis, Necessary for Sporulation Initiation, Is under Positive Stringent Transcription Control*. Journal of Bacteriology, 2013. **195**(8): p. 1656-1665.
55. Shemesh, M., R. Kolter, and R. Losick, *The Biocide Chlorine Dioxide Stimulates Biofilm Formation in Bacillus subtilis by Activation of the Histidine Kinase KinC*. Journal of Bacteriology, 2010. **192**(24): p. 6352-6356.
56. López, D., et al., *Structurally diverse natural products that cause potassium leakage trigger multicellularity in Bacillus subtilis*. Proceedings of the National Academy of Sciences, 2009. **106**(1): p. 280-285.
57. Jiang, M., et al., *Multiple histidine kinases regulate entry into stationary phase and sporulation in Bacillus subtilis*. Molecular Microbiology, 2000. **38**(3): p. 535-542.
58. Veening, J.W., H. Murray, and J. Errington, *A mechanism for cell cycle regulation of sporulation initiation in Bacillus subtilis*. Genes Dev, 2009. **23**(16): p. 1959-70.
59. Zhang, S. and W.G. Haldenwang, *Guanine nucleotides stabilize the binding of Bacillus subtilis Obg to ribosomes*. Biochemical and Biophysical Research Communications, 2004. **322**(2): p. 565-569.
60. Perego, M., *A new family of aspartyl phosphate phosphatases targeting the sporulation transcription factor Spo0A of Bacillus subtilis*. Molecular Microbiology, 2001. **42**(1): p. 133-143.
61. Quisel, J.D. and A.D. Grossman, *Control of Sporulation Gene Expression in Bacillus subtilis by the Chromosome Partitioning Proteins Soj (ParA) and Spo0J (ParB)*. Journal of Bacteriology, 2000. **182**(12): p. 3446-3451.
62. Scholefield, G., et al., *Spo0J regulates the oligomeric state of Soj to trigger its switch from an activator to an inhibitor of DNA replication initiation*. Molecular Microbiology, 2011. **79**(4): p. 1089-1100.
63. Gruber, T.M. and C.A. Gross, *Multiple Sigma Subunits and the Partitioning of Bacterial Transcription Space*. Annual Review of Microbiology, 2003. **57**(1): p. 441-466.
64. Murakami, K.S., S. Masuda, and S.A. Darst, *Structural Basis of Transcription Initiation: RNA Polymerase Holoenzyme at 4 Å Resolution*. Science, 2002. **296**(5571): p. 1280-1284.
65. Haldenwang, W.G., *The sigma factors of Bacillus subtilis*. Microbiological Reviews, 1995. **59**(1): p. 1-30.
66. Kay, D. and S.C. Warren, *Sporulation in Bacillus subtilis. Morphological changes*. Biochem J, 1968. **109**(5): p. 819-24.
67. Wu, L.J. and J. Errington, *Use of asymmetric cell division and spoIIIE mutants to probe chromosome orientation and organization in Bacillus subtilis*. Molecular Microbiology, 1998. **27**(4): p. 777-786.
68. Whitten, A.E., et al., *The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition*. J Mol Biol, 2007. **368**(2): p. 407-20.
69. Jacques, D.A., et al., *Structure of the sporulation histidine kinase inhibitor Sda from Bacillus subtilis and insights into its solution state*. Acta Crystallographica Section D, 2009. **65**(6): p. 574-581.

70. Michael, W.M., *Cell cycle: Connecting DNA replication to sporulation in Bacillus*. Current Biology, 2001. **11**(11): p. R443-R445.
71. Ruvolo, M.V., K.E. Mach, and W.F. Burkholder, *Proteolysis of the replication checkpoint protein Sda is necessary for the efficient initiation of sporulation after transient replication stress in Bacillus subtilis*. Mol Microbiol, 2006. **60**(6): p. 1490-508.
72. Glaser, P., et al., *Dynamic, mitotic-like behavior of a bacterial protein required for accurate chromosome partitioning*. Genes & Development, 1997. **11**(9): p. 1160-1168.
73. Sullivan, N.L., K.A. Marquis, and D.Z. Rudner, *Recruitment of SMC by ParB-parS Organizes the Origin Region and Promotes Efficient Chromosome Segregation*. Cell, 2009. **137**(4): p. 697-707.
74. Ramamurthi, K.S. and R. Losick, *Negative membrane curvature as a cue for subcellular localization of a bacterial protein*. Proceedings of the National Academy of Sciences, 2009. **106**(32): p. 13541-13545.
75. Pavlendová, N., K. Muchová, and I. Barák, *Chromosome segregation in Bacillus subtilis*. Folia Microbiologica, 2007. **52**(6): p. 563-572.
76. Ben-Yehuda, S., D.Z. Rudner, and R. Losick, *RacA, a Bacterial Protein That Anchors Chromosomes to the Cell Poles*. Science, 2003. **299**(5606): p. 532-536.
77. Ben-Yehuda, S. and R. Losick, *Asymmetric cell division in B. subtilis involves a spiral-like intermediate of the cytokinetic protein FtsZ*. Cell, 2002. **109**(2): p. 257-66.
78. Duncan, L. and R. Losick, *SpollAB is an anti-sigma factor that binds to and inhibits transcription by regulatory protein sigma F from Bacillus subtilis*. Proceedings of the National Academy of Sciences, 1993. **90**(6): p. 2325-2329.
79. Duncan, L., S. Alper, and R. Losick, *SpollAA Governs the Release of the Cell-type Specific Transcription Factor  $\sigma^F$  from its Anti-sigma Factor SpollAB*. Journal of Molecular Biology, 1996. **260**(2): p. 147-164.
80. Diederich, B., et al., *Role of interactions between SpollAA and SpollAB in regulating cell-specific transcription factor sigma F of Bacillus subtilis*. Genes & Development, 1994. **8**(21): p. 2653-2663.
81. Min, K.-T., et al.,  *$\sigma^F$ , the first compartment-specific transcription factor of B. subtilis, is regulated by an anti- $\sigma$  factor that is also a protein kinase*. Cell, 1993. **74**(4): p. 735-742.
82. Duncan, L., et al., *Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division*. Science, 1995. **270**(5236): p. 641-4.
83. Arigoni, F., et al., *SpollE governs the phosphorylation state of a protein regulating transcription factor sigma F during sporulation in Bacillus subtilis*. Proceedings of the National Academy of Sciences, 1996. **93**(8): p. 3238-3242.
84. Feucht, A., et al., *Bifunctional protein required for asymmetric cell division and cell-specific transcription in Bacillus subtilis*. Genes & Development, 1996. **10**(7): p. 794-803.
85. Barak, I. and A.J. Wilkinson, *Where asymmetry in gene expression originates*. Molecular Microbiology, 2005. **57**(3): p. 611-20.

86. Clarkson, J., et al., *Fluorescence and Kinetic Analysis of the SpoIIAB Phosphorylation Reaction, a Key Regulator of Sporulation in Bacillus subtilis*. *Biochemistry*, 2004. **43**(11): p. 3120-3128.
87. Cattoni, D.I., et al., *SpoIIIE mechanism of directional translocation involves target search coupled to sequence-dependent motor stimulation*. *EMBO Rep*, 2013. **14**(5): p. 473-479.
88. Becker, E.C. and K. Pogliano, *Cell-specific SpoIIIE assembly and DNA translocation polarity are dictated by chromosome orientation*. *Molecular Microbiology*, 2007. **66**(5): p. 1066-1079.
89. Marquis, K.A., et al., *SpoIIIE strips proteins off the DNA during chromosome translocation*. *Genes & Development*, 2008. **22**(13): p. 1786-1795.
90. Chastanet, A. and R. Losick, *Engulfment during sporulation in Bacillus subtilis is governed by a multi-protein complex containing tandemly acting autolysins*. *Molecular Microbiology*, 2007. **64**(1): p. 139-152.
91. Levnikov, V.M., et al., *Structure of components of an intercellular channel complex in sporulating Bacillus subtilis*. *Proceedings of the National Academy of Sciences*, 2012. **109**(14): p. 5441-5445.
92. Meisner, J., et al., *Structure of the basal components of a bacterial transporter*. *Proceedings of the National Academy of Sciences*, 2012. **109**(14): p. 5446-5451.
93. Meyer, P., et al., *Cell wall synthesis is necessary for membrane dynamics during sporulation of Bacillus subtilis*. *Molecular Microbiology*, 2010. **76**(4): p. 956-970.
94. Hilbert, D.W. and P.J. Piggot, *Compartmentalization of Gene Expression during Bacillus subtilis Spore Formation*. *Microbiology and Molecular Biology Reviews*, 2004. **68**(2): p. 234-262.
95. Vasudevan, P., et al., *Spore cortex formation in Bacillus subtilis is regulated by accumulation of peptidoglycan precursors under the control of sigma K*. *Molecular Microbiology*, 2007. **65**(6): p. 1582-1594.
96. McKenney, P.T., A. Driks, and P. Eichenberger, *The Bacillus subtilis endospore: assembly and functions of the multilayered coat*. *Nat Rev Micro*, 2013. **11**(1): p. 33-44.
97. Sanchez-Salas, J.L., et al., *Maturation of released spores is necessary for acquisition of full spore heat resistance during Bacillus subtilis sporulation*. *Appl Environ Microbiol*, 2011. **77**(19): p. 6746-54.
98. Yang, J., et al., *Transcriptional Regulation and Characteristics of a Novel N-Acetylmuramoyl-L-Alanine Amidase Gene Involved in Bacillus thuringiensis Mother Cell Lysis*. *Journal of Bacteriology*, 2013. **195**(12): p. 2887-2897.
99. Nugroho, F.A., et al., *Characterization of a New Sigma-K-Dependent Peptidoglycan Hydrolase Gene That Plays a Role in Bacillus subtilis Mother Cell Lysis*. *Journal of Bacteriology*, 1999. **181**(20): p. 6230-6237.
100. Kroos, L., *The Bacillus and Myxococcus Developmental Networks and Their Transcriptional Regulators*. *Annual Review of Genetics*, 2007. **41**(1): p. 13-39.
101. Jonas, R.M., et al., *The Bacillus subtilis spoIIIG operon encodes both sigma E and a gene necessary for sigma E activation*. *J Bacteriol*, 1988. **170**(2): p. 507-11.

102. Hofmeister, A.E.M., et al., *Extracellular signal protein triggering the proteolytic activation of a developmental transcription factor in B. subtilis*. Cell, 1995. **83**(2): p. 219-226.
103. Doan, T., et al., *Novel secretion apparatus maintains spore integrity and developmental gene expression in Bacillus subtilis*. PLoS Genet, 2009. **5**(7): p. e1000566.
104. Camp, A.H. and R. Losick, *A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis*. Genes Dev, 2009. **23**(8): p. 1014-24.
105. Rodrigues, C.D., et al., *Peptidoglycan hydrolysis is required for assembly and activity of the transenvelope secretion complex during sporulation in Bacillus subtilis*. Mol Microbiol, 2013.
106. Rhayat, L., et al., *Genetic dissection of an inhibitor of the sporulation sigma factor sigma(G)*. J Mol Biol, 2009. **390**(5): p. 835-44.
107. Karmazyn-Campelli, C., et al., *How the early sporulation sigma factor sigmaF delays the switch to late development in Bacillus subtilis*. Mol Microbiol, 2008. **67**(5): p. 1169-80.
108. Fairhead, H., B. Setlow, and P. Setlow, *Prevention of DNA damage in spores and in vitro by small, acid-soluble proteins from Bacillus species*. J Bacteriol, 1993. **175**(5): p. 1367-74.
109. Wang, S.T., et al., *The forespore line of gene expression in Bacillus subtilis*. J Mol Biol, 2006. **358**(1): p. 16-37.
110. Kunkel, B., R. Losick, and P. Stragier, *The Bacillus subtilis gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene*. Genes Dev, 1990. **4**(4): p. 525-35.
111. Rudner, D.Z. and R. Losick, *A sporulation membrane protein tethers the pro-sigmaK processing enzyme to its inhibitor and dictates its subcellular localization*. Genes Dev, 2002. **16**(8): p. 1007-18.
112. Dong, T.C. and S.M. Cutting, *SpoIVB-mediated cleavage of SpoIVFA could provide the intercellular signal to activate processing of Pro-sigmaK in Bacillus subtilis*. Mol Microbiol, 2003. **49**(5): p. 1425-34.
113. Zhou, R., et al., *Features of Pro-sigmaK important for cleavage by SpoIVFB, an intramembrane metalloprotease*. J Bacteriol, 2013. **195**(12): p. 2793-806.
114. Li, Z. and P.J. Piggot, *Development of a two-part transcription probe to determine the completeness of temporal and spatial compartmentalization of gene expression during bacterial development*. Proceedings of the National Academy of Sciences, 2001. **98**(22): p. 12538-12543.
115. Zhang, B. and L. Kroos, *A feedback loop regulates the switch from one sigma factor to the next in the cascade controlling Bacillus subtilis mother cell gene expression*. Journal of Bacteriology, 1997. **179**(19): p. 6138-44.
116. Camp, A.H., A.F. Wang, and R. Losick, *A small protein required for the switch from {sigma}F to {sigma}G during sporulation in Bacillus subtilis*. J Bacteriol, 2011. **193**(1): p. 116-24.

117. Margolis, P., A. Driks, and R. Losick, *Establishment of cell type by compartmentalized activation of a transcription factor*. *Science*, 1991. **254**(5031): p. 562-5.
118. Fort, P. and P.J. Piggot, *Nucleotide sequence of sporulation locus *spoIIA* in *Bacillus subtilis**. *J Gen Microbiol*, 1984. **130**(8): p. 2147-53.
119. Campbell, E.A. and S.A. Darst, *The anti-sigma factor *SpoIIAB* forms a 2:1 complex with *sigma(F)*, contacting multiple conserved regions of the sigma factor*. *J Mol Biol*, 2000. **300**(1): p. 17-28.
120. Campbell, E.A., et al., *Crystal structure of the *Bacillus stearothermophilus* anti-sigma factor *SpoIIAB* with the sporulation sigma factor *sigmaF**. *Cell*, 2002. **108**(6): p. 795-807.
121. Dutta, R. and M. Inouye, *GHL, an emergent ATPase/kinase superfamily*. *Trends in Biochemical Sciences*, 2000. **25**(1): p. 24-28.
122. Najafi, S.M., A.C. Willis, and M.D. Yudkin, *Site of phosphorylation of *SpoIIAA*, the anti-anti-sigma factor for sporulation-specific sigma F of *Bacillus subtilis**. *J Bacteriol*, 1995. **177**(10): p. 2912-3.
123. Seavers, P.R., et al., *Structure of the *Bacillus* Cell Fate Determinant *SpoIIAA* in Phosphorylated and Unphosphorylated Forms*. *Structure*, 2001. **9**(7): p. 605-614.
124. Masuda, S., et al., *Crystal structures of the ADP and ATP bound forms of the *Bacillus* anti-sigma factor *SpoIIAB* in complex with the anti-anti-sigma *SpoIIAA**. *J Mol Biol*, 2004. **340**(5): p. 941-56.
125. Yudkin, M.D. and J. Clarkson, *Differential gene expression in genetically identical sister cells: the initiation of sporulation in *Bacillus subtilis**. *Mol Microbiol*, 2005. **56**(3): p. 578-89.
126. Shu, J.C., J. Clarkson, and M.D. Yudkin, *Studies of *SpoIIAB* mutant proteins elucidate the mechanisms that regulate the developmental transcription factor *sigmaF* in *Bacillus subtilis**. *Biochem J*, 2004. **384**(Pt 1): p. 169-78.
127. Addinall, S.G., E. Bi, and J. Lutkenhaus, *FtsZ ring formation in *fts* mutants*. *J Bacteriol*, 1996. **178**(13): p. 3877-84.
128. de Boer, P., R. Crossley, and L. Rothfield, *The essential bacterial cell-division protein *FtsZ* is a GTPase*. *Nature*, 1992. **359**(6392): p. 254-6.
129. Nogales, E., et al., *Tubulin and *FtsZ* form a distinct family of GTPases*. *Nat Struct Biol*, 1998. **5**(6): p. 451-8.
130. Beall, B. and J. Lutkenhaus, *FtsZ in *Bacillus subtilis* is required for vegetative septation and for asymmetric septation during sporulation*. *Genes Dev*, 1991. **5**(3): p. 447-55.
131. Lu, C., M. Reedy, and H.P. Erickson, *Straight and curved conformations of *FtsZ* are regulated by GTP hydrolysis*. *J Bacteriol*, 2000. **182**(1): p. 164-70.
132. Erickson, H.P., et al., *Bacterial cell division protein *FtsZ* assembles into protofilament sheets and minirings, structural homologs of tubulin polymers*. *Proc Natl Acad Sci U S A*, 1996. **93**(1): p. 519-23.
133. Brinkley, W., *Microtubules: a brief historical perspective*. *J Struct Biol*, 1997. **118**(2): p. 84-6.

134. Mukherjee, A. and J. Lutkenhaus, *Guanine nucleotide-dependent assembly of FtsZ into filaments*. J Bacteriol, 1994. **176**(9): p. 2754-8.
135. Lowe, J. and L.A. Amos, *Crystal structure of the bacterial cell-division protein FtsZ*. Nature, 1998. **391**(6663): p. 203-6.
136. Oliva, M.A., S.C. Cordell, and J. Lowe, *Structural insights into FtsZ protofilament formation*. Nat Struct Mol Biol, 2004. **11**(12): p. 1243-50.
137. Mukherjee, A. and J. Lutkenhaus, *Dynamic assembly of FtsZ regulated by GTP hydrolysis*. EMBO J, 1998. **17**(2): p. 462-9.
138. Li, Y., et al., *FtsZ Protofilaments Use a Hinge-Opening Mechanism for Constrictive Force Generation*. Science, 2013. **341**(6144): p. 392-395.
139. Osawa, M., D.E. Anderson, and H.P. Erickson, *Reconstitution of contractile FtsZ rings in liposomes*. Science, 2008. **320**(5877): p. 792-4.
140. Li, Z., et al., *The structure of FtsZ filaments in vivo suggests a force-generating role in cell division*. EMBO J, 2007. **26**(22): p. 4694-708.
141. Gamba, P., et al., *Two-step assembly dynamics of the Bacillus subtilis divisome*. J Bacteriol, 2009. **191**(13): p. 4186-94.
142. Bramkamp, M., et al., *Regulated intramembrane proteolysis of FtsL protein and the control of cell division in Bacillus subtilis*. Mol Microbiol, 2006. **62**(2): p. 580-91.
143. Real, G., et al., *Cell division protein DivIB influences the Spo0J/Soj system of chromosome segregation in Bacillus subtilis*. Molecular Microbiology, 2005. **55**(2): p. 349-367.
144. Margolin, W., *FtsZ and the division of prokaryotic cells and organelles*. Nat Rev Mol Cell Biol, 2005. **6**(11): p. 862-71.
145. Romberg, L. and P.A. Levin, *Assembly dynamics of the bacterial cell division protein FTSZ: poised at the edge of stability*. Annu Rev Microbiol, 2003. **57**: p. 125-54.
146. Jamroskovic, J., et al., *An oscillating Min system in Bacillus subtilis influences asymmetrical septation during sporulation*. Microbiology, 2012. **158**(Pt 8): p. 1972-81.
147. Wu, L.J. and J. Errington, *Coordination of cell division and chromosome segregation by a nucleoid occlusion protein in Bacillus subtilis*. Cell, 2004. **117**(7): p. 915-25.
148. Adler, H.I., et al., *Miniature Escherichia coli Cells Deficient in DNA*. Proc Natl Acad Sci U S A, 1967. **57**(2): p. 321-6.
149. de Boer, P.A., R.E. Crossley, and L.I. Rothfield, *A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli*. Cell, 1989. **56**(4): p. 641-9.
150. Dajkovic, A., et al., *MinC spatially controls bacterial cytokinesis by antagonizing the scaffolding function of FtsZ*. Curr Biol, 2008. **18**(4): p. 235-44.
151. de Boer, P.A., et al., *The MinD protein is a membrane ATPase required for the correct placement of the Escherichia coli division site*. EMBO J, 1991. **10**(13): p. 4371-80.
152. Raskin, D.M. and P.A. de Boer, *Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli*. Proc Natl Acad Sci U S A, 1999. **96**(9): p. 4971-6.

153. Oliva, M.A., et al., *Features critical for membrane binding revealed by DivIVA crystal structure*. EMBO J, 2010. **29**(12): p. 1988-2001.
154. Stahlberg, H., et al., *Oligomeric structure of the Bacillus subtilis cell division protein DivIVA determined by transmission electron microscopy*. Mol Microbiol, 2004. **52**(5): p. 1281-90.
155. Gonzy-Treboul, G., C. Karmazyn-Campelli, and P. Stragier, *Developmental regulation of transcription of the Bacillus subtilis ftsAZ operon*. J Mol Biol, 1992. **224**(4): p. 967-79.
156. Levin, P.A. and R. Losick, *Transcription factor Spo0A switches the localization of the cell division protein FtsZ from a medial to a bipolar pattern in Bacillus subtilis*. Genes & Development, 1996. **10**(4): p. 478-88.
157. Ryter, A., P. Schaeffer, and H. Ionesco, *[Cytologic classification, by their blockage stage, of sporulation mutants of Bacillus subtilis Marburg]*. Ann Inst Pasteur (Paris), 1966. **110**(3): p. 305-15.
158. York, K., et al., *Spo0A controls the sigma A-dependent activation of Bacillus subtilis sporulation-specific transcription unit spoII E*. J Bacteriol, 1992. **174**(8): p. 2648-58.
159. King, N., et al., *Septation, dephosphorylation, and the activation of sigmaF during sporulation in Bacillus subtilis*. Genes Dev, 1999. **13**(9): p. 1156-67.
160. Barak, I., et al., *Structure and function of the Bacillus SpoII E protein and its localization to sites of sporulation septum assembly*. Molecular Microbiology, 1996. **19**(5): p. 1047-60.
161. Levin, P.A., et al., *Localization of the sporulation protein SpoII E in Bacillus subtilis is dependent upon the cell division protein FtsZ*. Mol Microbiol, 1997. **25**(5): p. 839-46.
162. Campo, N., K.A. Marquis, and D.Z. Rudner, *SpoII Q anchors membrane proteins on both sides of the sporulation septum in Bacillus subtilis*. J Biol Chem, 2008. **283**(8): p. 4975-82.
163. Arigoni, F., et al., *The SpoII E phosphatase, the sporulation septum and the establishment of forespore-specific transcription in Bacillus subtilis: a reassessment*. Molecular Microbiology, 1999. **31**(5): p. 1407-15.
164. Lucet, I., et al., *Direct interaction between the cell division protein FtsZ and the cell differentiation protein SpoII E*. The EMBO Journal, 2000. **19**(7): p. 1467-1475.
165. Barford, D., *Molecular mechanisms of the protein serine/threonine phosphatases*. Trends in Biochemical Sciences, 1996. **21**(11): p. 407-12.
166. Das, A.K., et al., *Crystal structure of the protein serine/threonine phosphatase 2C at 2.0 Å resolution*. EMBO Journal, 1996. **15**(24): p. 6798-809.
167. Pan, Q., D.A. Garsin, and R. Losick, *Self-reinforcing activation of a cell-specific transcription factor by proteolysis of an anti-sigma factor in B. subtilis*. Mol Cell, 2001. **8**(4): p. 873-83.
168. Clarkson, J., I.D. Campbell, and M.D. Yudkin, *Efficient regulation of sigmaF, the first sporulation-specific sigma factor in B. subtilis*. J Mol Biol, 2004. **342**(4): p. 1187-95.
169. Carniol, K., et al., *Genetic dissection of the sporulation protein SpoII E and its role in asymmetric division in Bacillus subtilis*. Journal of Bacteriology, 2005. **187**(10): p. 3511-20.

170. Levdikov, V.M., et al., *Structure of the phosphatase domain of the cell fate determinant SpoII E from Bacillus subtilis*. J Mol Biol, 2012. **415**(2): p. 343-58.
171. Fogg, M.J., et al., *Application of the use of high-throughput technologies to the determination of protein structures of bacterial and viral pathogens*. Acta Crystallogr D Biol Crystallogr, 2006. **62**(Pt 10): p. 1196-207.
172. Sambrook, J. and D. Russell, *Molecular Cloning: A Laboratory Manual*. 2001: Cold Spring Harbor Laboratory Press.
173. Promega. *PCR Amplification*. n.d. [cited 2013 24/08]; Available from: <http://www.promega.co.uk/resources/product-guides-and-selectors/protocols-and-applications-guide/pcr-amplification/>.
174. Roux, K.H., *Optimization and Troubleshooting in PCR*. Cold Spring Harbor Protocols, 2009. **2009**(4): p. pdb.ip66.
175. Markoulatos, P., N. Siafakas, and M. Moncany, *Multiplex polymerase chain reaction: a practical approach*. J Clin Lab Anal, 2002. **16**(1): p. 47-51.
176. NC-IUB, *Units of Enzyme Activity*. European Journal of Biochemistry, 1979. **97**(2): p. 319-320.
177. Aslanidis, C. and P.J. de Jong, *Ligation-independent cloning of PCR products (LIC-PCR)*. Nucleic Acids Research, 1990. **18**(20): p. 6069-74.
178. Chung, C.T., S.L. Niemela, and R.H. Miller, *One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution*. Proc Natl Acad Sci U S A, 1989. **86**(7): p. 2172-5.
179. Weiner, M.P., et al., *Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction*. Gene, 1994. **151**(1-2): p. 119-23.
180. Novagen, *pET System Manual, 2nd Printing, 10th Edition*. 2003.
181. Hengen, P., *Purification of His-Tag fusion proteins from Escherichia coli*. Trends Biochem Sci, 1995. **20**(7): p. 285-6.
182. *Addition of imidazole during binding improves purity of histidine-tagged proteins*. [cited 2013 31/08]; Available from: [https://www.gelifesciences.com/gehcls\\_images/GELS/Related%20Content/Files/1314742967685/litdoc28406741AA\\_20110831015342.pdf](https://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314742967685/litdoc28406741AA_20110831015342.pdf).
183. *Gel Filtration: Principles and Methods*. [cited 2013 31/08]; Available from: [https://www.gelifesciences.com/gehcls\\_images/GELS/Related%20Content/Files/1314807262343/litdoc18102218AK\\_20110831220049.pdf](https://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314807262343/litdoc18102218AK_20110831220049.pdf).
184. Gasteiger, E., et al., *ExpASY: The proteomics server for in-depth protein knowledge and analysis*. Nucleic Acids Res, 2003. **31**(13): p. 3784-8.
185. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 1970. **227**(5259): p. 680-5.
186. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure*. Nat Protoc, 2006. **1**(6): p. 2876-90.
187. Yumerefendi, H., et al., *ESPRIT: An automated, library-based method for mapping and soluble expression of protein domains from challenging targets*. Journal of Structural Biology, 2010. **172**(1): p. 66-74.

188. Rawlings, A.E., et al., *Expression of soluble, active fragments of the morphogenetic protein SpoIIE from Bacillus subtilis using a library-based construct screen*. Protein Engineering, Design & Selection, 2010. **23**(11): p. 817-25.
189. Linding, R., et al., *GlobPlot: Exploring protein sequences for globularity and disorder*. Nucleic Acids Res, 2003. **31**(13): p. 3701-8.
190. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes*. J Mol Biol, 2001. **305**(3): p. 567-80.
191. Buchan, D.W., et al., *Protein annotation and modelling servers at University College London*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W563-8.
192. Cobb, B.D. and J.M. Clarkson, *A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods*. Nucleic Acids Res, 1994. **22**(18): p. 3801-5.
193. Lindwall, G., et al., *A sparse matrix approach to the solubilization of overexpressed proteins*. Protein Eng, 2000. **13**(1): p. 67-71.
194. Gordon, R.E. and N.R. Smith, *Aerobic sporeforming bacteria capable of growth at high temperatures*. J Bacteriol, 1949. **58**(3): p. 327-41.
195. Liu, B., W.K. Eliason, and T.A. Steitz, *Structure of a helicase-helicase loader complex reveals insights into the mechanism of bacterial primosome assembly*. Nat Commun, 2013. **4**: p. 2495.
196. Wendrich, T.M., C.L. Beckering, and M.A. Marahiel, *Characterization of the relA/spoT gene from Bacillus stearothermophilus*. FEMS Microbiol Lett, 2000. **190**(2): p. 195-201.
197. Larkin, M.A., et al., *Clustal W and Clustal X version 2.0*. Bioinformatics, 2007. **23**(21): p. 2947-8.
198. Gouet, P., X. Robert, and E. Courcelle, *ESPrpt/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins*. Nucleic Acids Research, 2003. **31**(13): p. 3320-3323.
199. Schuck, P., *Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling*. Biophys J, 2000. **78**(3): p. 1606-19.
200. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment*. J Mol Biol, 2000. **302**(1): p. 205-17.
201. Fauman, E.B. and M.A. Saper, *Structure and function of the protein tyrosine phosphatases*. Trends Biochem Sci, 1996. **21**(11): p. 413-7.
202. Jackson, M.D., C.C. Fjeld, and J.M. Denu, *Probing the function of conserved residues in the serine/threonine phosphatase PP2C $\alpha$* . Biochemistry, 2003. **42**(28): p. 8513-21.
203. Pullen, K.E., et al., *An alternate conformation and a third metal in PstP/Ppp, the M. tuberculosis PP2C-Family Ser/Thr protein phosphatase*. Structure, 2004. **12**(11): p. 1947-54.
204. Wehenkel, A., et al., *Structural and binding studies of the three-metal center in two mycobacterial PPM Ser/Thr protein phosphatases*. J Mol Biol, 2007. **374**(4): p. 890-8.

205. Rantanen, M.K., et al., *Structure of Streptococcus agalactiae serine/threonine phosphatase. The subdomain conformation is coupled to the binding of a third metal ion.* FEBS J, 2007. **274**(12): p. 3128-37.
206. Schlicker, C., et al., *Structural analysis of the PP2C phosphatase tPphA from Thermosynechococcus elongatus: a flexible flap subdomain controls access to the catalytic site.* J Mol Biol, 2008. **376**(2): p. 570-81.
207. Ericsson, U.B., et al., *Thermofluor-based high-throughput stability optimization of proteins for structural studies.* Anal Biochem, 2006. **357**(2): p. 289-98.
208. Switzer, R.C., 3rd, C.R. Merrill, and S. Shifrin, *A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels.* Anal Biochem, 1979. **98**(1): p. 231-7.
209. Clarkson, J., I.D. Campbell, and M.D. Yudkin, *Phosphorylation induces subtle structural changes in SpoIIAA, a key regulator of sporulation.* Biochem J, 2003. **372**(Pt 1): p. 113-9.
210. Cliff, M.J., et al., *Transition state analogue structures of human phosphoglycerate kinase establish the importance of charge balance in catalysis.* J Am Chem Soc, 2010. **132**(18): p. 6507-16.
211. Chappie, J.S., et al., *G domain dimerization controls dynamin's assembly-stimulated GTPase activity.* Nature, 2010. **465**(7297): p. 435-40.
212. Wang, W., et al., *Structural characterization of the reaction pathway in phosphoserine phosphatase: crystallographic "snapshots" of intermediate states.* Journal of Molecular Biology, 2002. **319**(2): p. 421-31.
213. Suzuki, T., et al., *Phosphorylation of three regulatory serines of Tob by Erk1 and Erk2 is required for Ras-mediated cell proliferation and transformation.* Genes & Development, 2002. **16**(11): p. 1356-1370.
214. Kabsch, W., *XDS.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.
215. Winter, G., C.M. Lobley, and S.M. Prince, *Decision making in xia2.* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 7): p. 1260-73.
216. Lucet, I., R. Borriss, and M.D. Yudkin, *Purification, kinetic properties, and intracellular concentration of SpoIIIE, an integral membrane protein that regulates sporulation in Bacillus subtilis.* J Bacteriol, 1999. **181**(10): p. 3242-5.
217. Cheung, H.Y., L. Vitkovič, and M.R.W. Brown, *Toxic Effect of Manganese on Growth and Sporulation of Bacillus stearothermophilus.* Journal of General Microbiology, 1982. **128**(10): p. 2395-2402.
218. Panigrahi, K., et al., *The alpha, alpha-difluorinated phosphonate L-pSer-analogue: an accessible chemical tool for studying kinase-dependent signal transduction.* Chem Biol, 2009. **16**(9): p. 928-36.
219. Hilbert, D.W. and P.J. Piggot, *Novel spoIIIE Mutation That Causes Uncompartmentalized sigma F Activation in Bacillus subtilis.* Journal of Bacteriology, 2003. **185**(5): p. 1590-1598.
220. Barák, I. and P. Youngman, *SpoIIIE mutants of Bacillus subtilis comprise two distinct phenotypic classes consistent with a dual functional role for the SpoIIIE protein.* Journal of Bacteriology, 1996. **178**(16): p. 4984-9.

221. Carniol, K., P. Eichenberger, and R. Losick, *A threshold mechanism governing activation of the developmental regulatory protein sigma F in Bacillus subtilis*. *Journal of Biological Chemistry*, 2004. **279**(15): p. 14860-70.
222. Illing, N. and J. Errington, *Genetic regulation of morphogenesis in Bacillus subtilis: roles of sigma E and sigma F in prespore engulfment*. *J Bacteriol*, 1991. **173**(10): p. 3159-69.
223. Waites, W.M., et al., *Sporulation in Bacillus subtilis. Correlation of biochemical events with morphological changes in asporogenous mutants*. *Biochem J*, 1970. **118**(4): p. 667-76.
224. Kramer, N., J. Hahn, and D. Dubnau, *Multiple interactions among the competence proteins of Bacillus subtilis*. *Mol Microbiol*, 2007. **65**(2): p. 454-64.
225. Maamar, H. and D. Dubnau, *Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop*. *Mol Microbiol*, 2005. **56**(3): p. 615-24.
226. Hamoen, L.W., G. Venema, and O.P. Kuipers, *Controlling competence in Bacillus subtilis: shared use of regulators*. *Microbiology*, 2003. **149**(Pt 1): p. 9-17.
227. Kaufenstein, M., M. van der Laan, and P.L. Graumann, *The three-layered DNA uptake machinery at the cell pole in competent Bacillus subtilis cells is a stable complex*. *J Bacteriol*, 2011. **193**(7): p. 1633-42.
228. Chen, I. and D. Dubnau, *DNA uptake during bacterial transformation*. *Nat Rev Microbiol*, 2004. **2**(3): p. 241-9.
229. Chen, Z., H. Yang, and N.P. Pavletich, *Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures*. *Nature*, 2008. **453**(7194): p. 489-4.
230. Saladin, A., et al., *Modeling the early stage of DNA sequence recognition within RecA nucleoprotein filaments*. *Nucleic Acids Res*, 2010. **38**(19): p. 6313-23.
231. *Molecular Biological Methods for Bacillus (Modern Microbiological Methods)*, ed. C. Harwood and S. Cutting. 1990: Wiley.
232. Daniel R. Zeigler, P.D., *Integration Vectors for Gram-Positive Organisms*, in *Catalog of Strains*, B.G.S. Center, Editor 2002.
233. Zimmermann, K., D. Schogl, and J.W. Mannhalter, *Digestion of terminal restriction endonuclease recognition sites on PCR products*. *Biotechniques*, 1998. **24**(4): p. 582-4.
234. Labid, C. and Y. Gao. 14/08/2006 [cited 2013 24/09]; Available from: <http://www.bioinformatics.org/primerx/>.