Structure and Composition of an Intercellular Channel in Sporulating *Bacillus subtilis*

Sonam Sidhu

PhD

Biology

December 2022

Abstract

In response to nutrient starvation, the Gram-positive soil bacterium *Bacillus subtilis* forms a dormant spore in a process called sporulation. Sporulation involves an asymmetric cell division that results in a larger mother cell and a smaller forespore. The mother cell then 'engulfs' the forespore in a phagocytotic like mechanism and nurtures it before releasing it into the environment. Engulfment occurs in the sporulation pathway of the disease-causing spore forming bacteria *Clostridium difficile* and *Bacillus anthracis* and results in the reorganisation of the sporangium from two cells that lie by side to a cell within a cell structure. In this thesis, *B. subtilis* is used as a model organism to study engulfment to determine how bacterial cells localise proteins, move macromolecules and catalyse membrane fusion.

Upon completion of engulfment, an intercellular channel is believed to form between the forespore protein SpoIIQ and the eight mother cell proteins encoded by the *spoIIIAA-spoIIIAH* operon. The channel is thought to allow the passage of a molecule important in forespore maturation, although the identity of this molecule is currently unknown. Whilst the structure of the complex formed between the extracellular C-terminal domains of SpoIIQ and SpoIIIAH has been determined by protein crystallography, no *in-vivo* experiments have been performed to address the stoichiometry of the channel complex in living cells. To determine the stoichiometry of the channel complex, high speed single molecule fluorescence microscopy was used to image live *B. subtilis* cells. Evidence was found to suggest that the channel complex assembles from SpoIIQ dimers and SpoIIIAH hexamers, suggesting that channel formation is more complicated than initially expected. SpoIIQ and SpoIIIAH were found to have low mobility and an average of 200-300 SpoIIQ and SpoIIIAH molecules were found in sporulating *B. subtilis* cells.

Table of Contents

Abstract 2

Table of contents 3

List of Tables 8

List of Figures 9

Acknowledgments 14

Authors Declaration 15

Chapter 1 – Biology Introduction 16

1.1 Cell Differentiation 16

1.2 Cell Differentiation in *B. subtilis* 16

1.3 Spore Forming Bacteria 17

1.4 Spore Forming Bacteria in Disease 18

1.5 *B. subtilis* Life Cycle 19

1.6 Signaling 20

1.7 Asymmetric Septation 21

1.8 Axial Filamentation 22

1.9 Differential Gene Expression 23

1.9.1 Activation of σ^F and σ^E 24

1.9.2 Activation of σ^G and σ^K 25

1.10 Cellular Dynamics of Engulfment 26

1.10.1 Cellular Engulfment: Role of SpoIIDPM 26

1.10.2 Cellular Engulfment: SpoIIQ and SpoIIIAH 28

1.10.3 Cellular Engulfment: Membrane Fission 33

1.11 Spore Cortex and Coat Assembly 33

1.12 Spore Germination 35

Chapter 2 – Physics Introduction 36

2.1 Introduction to Microscopy 36

2.2 Fluorescence 36

2.3 Fluorescent Proteins 37
2.4 Green Fluorescent proteins (GFPs) 38
2.5 Green Fluorescent protein (GFP) variants 39
2.6 Advantages and Disadvantages of Fluorescent Proteins 40
2.7 Why detect single molecules? 41
2.8 Super Resolution Microscopy 41
2.8.1 SIM Microscopy 43
2.8.2 Slimfield Microscopy 44
2.9 Aims 46
Chapter 3 - Methods 47
3.1 *mGFP-spoIIQ* Fusion Construction 47
3.2 Gibson Assembly 49
3.3 Colony PCR 49
3.4 Agarose Gel Electrophoresis 50
3.5 Preparation of *B. subtilis* competent cells 50
3.6 Transformation of *B. subtilis* 50
3.7 Confirming Transformation of *B. subtilis* 50
3.8 Fluorescence Imaging of *B. subtilis*: Sample Preparation 51
3.9 Fluorescence Imaging of *B. subtilis*: Structured Illumination Microscopy (SIM) 51
3.10 Fluorescence Imaging of *B. subtilis*: Slimfield Microscopy 52
3.11 Slimfield Microscopy Image Analysis 53
3.11.1 Image Analysis: Spot Correction and Subpixel localisation 53
3.11.2 Image Analysis: Isingle Determination 54
3.11.3 Image Analysis: Stoichiometry, Copy Number and Diffusion Coefficient Determination 55
3.12 Kernel Density Estimation (KDE) 55
3.12.1 Interpretation of Kernel Density Estimation (KDE) plots 56
3.13 Restriction Enzyme Digestion 57
3.14 PCR Amplification 58

3.15 Q5 Site Directed Mutagenesis 58

Chapter 4 – Stoichiometry, Diffusion Coefficient and Copy Number Analysis of SpoIIQ and SpoIIIAH in wildtype backgrounds 60

4.1 Introduction 60

4.2 mGFP- spoIIQ Fusion Construction 61

4.3 mGFP SpoIIQ Imaging – SIM Microscopy 65

4.4 *mGFP spoIIQ delspoIIQ* Strain Construction 66

4.5 SpoIIIAH Fusions 69

4.6 Slimfield Microscopy 70

4.7 SpoIIQ Imaging – Slimfield Microscopy 71

4.8 SpoIIIAH Imaging – Slimfield Microscopy 72

4.9 Stoichiometry analysis suggests that SpoIIQ is a dimer 73

4.10 Diffusion Coefficient analysis suggests that SpoIIQ has low mobility 75

4.11 Copy Number analysis suggests an average of 200 SpoIIQ molecules in sporulating *B.subtilis* cells 76

4.12 Stoichiometry analysis suggests that SpoIIIAH is a hexamer 76

4.13 Diffusion Coefficient analysis suggests that SpoIIIAH has low mobility. 78

4.14 Copy Number analysis suggests an average of 200 SpoIIIAH molecules in sporulating B.*subtilis* cells 80

4.15 Discussion 81

4.16 Dual Strain Construction 82

4.16.1 Slimfield Microscopy suggests that SpoIIQ and SpoIIIAH co-localise in live cells 83

4.16.2 Stoichiometry Analysis suggests that the Q-AH channel acts as a site of recruitment for SpoIIQ dimers and SpoIIIAH hexamers 84

Chapter 5 – Stoichiometry, Diffusion Coefficient and Copy Number Analysis of SpoIIQ and SpoIIIAH in mutant backgrounds 85

5.1 Introduction 85

5.2 Mutant Strain Construction 86

5.3 SIM Microscopy suggests that SpoIIQ is delocalised in the absence of SpoIIIAH 87

5.4 Slimfield Microscopy suggests that SpoIIQ is delocalised in the absence of SpoIIIAH 85

5.5 Slimfield Microscopy suggests that SpoIIIAH is delocalized in the absence of SpoIIQ 89
5.6 Stoichiometry Analysis suggests SpoIIQ oligomerises differently in the absence of SpoIIIAH 90
5.7 Diffusion Coefficient analysis suggests that SpoIIQ has low mobility in the absence of SpoIIIAH..91
5.8 Copy Number analysis suggests an average of 200 SpoIIQ molecules in the absence of SpoIIIAH 92
5.9 Stoichiometry analysis suggests that SpoIIIAH oligomerises differently in the absence of SpoIIQ 93
5.10 Diffusion Coefficient analysis suggests SpoIIIAH has low mobility in the absence of SpoIIQ 94
5.11 Copy Number Analysis suggests downregulation of SpoIIIAH in the absence of SpoIIQ..... 96
5.12 Discussion..... 96
Chapter 6 – Analysis of the SpoIIQ – SpoIIIAH Interface using Site Directed Mutagenesis 99
6.1 Introduction 99
6.2 The SpoIIQ – SpoIIIAH Interface..... 99
6.3 Core and Rim Residues 101
6.4 Rationale: Site Directed Mutagenesis..... 102
6.5 pSG1154-mGFP SpoIIQ Verification 102
6.6 pSG1154 Verification 104
6.7 PCR Amplification..... 105
6.8 SpeI and HindIII Restriction Enzyme Digest 106
6.9 Cloning into pCR- Blunt 107
6.10 Site Directed Mutagenesis 109
6.11 Sequencing 112
6.12 Restriction Enzyme Digestion of Mutated Plasmids 113
6.13 Ligation of Mutagenic Inserts to pSG1154..... 114
6.14 Discussion..... 115
Chapter 7 – Discussion and Future Work 117
7.1 Discussion of Results in Chapters 4, 5 and 6..... 117
7.2 Discussion of Results in the Context of Sporulation 119

7.3 Alternative methods to study the interactions between SpoIIQ and SpoIIIAH 121

7.4 Future Work 125

Chapter 8 – Appendix 127

8.1 *Bacillus subtilis* strains used in this thesis 127

8.2 Plasmids used in this thesis 127

8.3 Oligonucleotide Primers 128

8.3.1 Gibson Assembly Primers 128

8.3.2 Site Directed Mutagenesis Primers 128

8.3.3 SpoIIQ Sequencing Primers 128

8.3.4 M13 Forward and Reverse Sequencing Primers 129

8.4 *B. subtilis* liquid media and reagents 129

8.4.1 T Base 129

8.4.2 SpC 129

8.4.3 SpII 129

8.4.4 SpII+EDTA 129

8.4.5 Top Agar 130

8.4.6 Difco Sporulation Medium (DSM) 130

8.4.7 Spizizen Minimal Medium (SMN) 130

References 130

List of Tables

Table 3-1 Excitation and emission profiles of mGFP and mCherry fluorescent proteins 48

Table 3-2 *In-vivo* and *in-vitro* Isingle values for mCherry and mGFP 54

Table 3-3 PCR reaction set up 58

Table 3-4 Thermocycling conditions for PCR 58

Table 3-5 Q5 Site directed mutagenesis reaction composition 58

Table 3-6 Thermocycling conditions for Q5 site directed mutagenesis 59

Table 3-7 Kinase, Ligase and DpnI (KLD) treatment reaction composition 59

Table 4-1 List of strains imaged and analysed in thesis chapter 4 67

Table 4-2 List of strains used to construct the dual labelled strain 83

Table 5-1 List of strains imaged and analysed in thesis chapter 5 87

Table 6-1 List of mutagenic primers used in site directed mutagenesis 111

Table 6-2 Sequencing results from the first round of sequencing 112

Table 6-3 Sequencing results from the second round of sequencing 112

List of Figures

Figure 1-1 Gram-positive and Gram-negative bacterial membranes 16

Figure 1-2 Cell Differentiation in *B. subtilis* 17

Figure 1-3 Schematic diagram illustrating the *B. subtilis* vegetative and sporulation cycles 20

Figure 1-4 Diagram showing the role of the phosphorelay in sporulation initiation 21

Figure 1-5 Schematic representation of the asymmetric septation process 22

Figure 1-6 Chromosome translocation by the SpoIIIE DNA translocase 23

Figure 1-7 Activation of sporulation specific σ factors in the mother cell and forespore 24

Figure 1-8 Flow chart illustrating σ^F regulation 25

Figure 1-9 Engulfment during sporulation in *Bacillus subtilis* 26

Figure 1-10 Role of the SpollDPM complex in engulfment 27

Figure 1-11 X-ray crystal structure of the complex formed between the C-terminal extracellular domains of SpoIIQ and SpoIIIAH 28

Figure 1-12 Localisation of SpoIIQ 29

Figure 1-13 Localisation of SpoIIIAH 30

Figure 1-14 Zipper-like interaction between SpoIIQ and SpoIIIAH during engulfment membrane migration. 31

Figure 1-15 Secretion complex formed between the mother cell SpoIIIAA proteins and forespore SpoIIQ protein 32

Figure 1-16 Schematic representation of the proposed SpoIIQ-SpoIIIAH complex 33

Figure 1-17 Spore ultrastructure 34

Figure 2-1 Jablonski Diagram to illustrate the physics of fluorescence 37

Figure 2-2 Structure of the wtGFP chromophore 39

Figure 2-3 The excitation and emission spectra of BFP, CFP, GFP and mRFP1 40

Figure 2-4 Diagram showing the response of an optical imaging system to a single fluorophore 42

Figure 2-5 Fitting Single-Molecule Pixel Data to a Gaussian Function 43

Figure 2-6 Diagram illustrating the Moire effect and the principles of SIM Microscopy 44

Figure 2-7 Schematic diagram showing Slimfield illumination 45

Figure 3-1 *mGFP spoIIQ* fusion integration into the *B. subtilis* chromosome 47

Figure 3-2 Plasmid map of pSG1154 48

Figure 3-3 Schematic representation of the transformation selection method in *B. subtilis* 51

Figure 3-4 Diagram illustrating the Slimfield microscope in dual colour mode 52

Figure 3-5 Spot correction and subpixel localisation using custom built MATLAB code 53

Figure 3-6 Single characteristic Intensity (Isingle plot) generated by the getallspots code. 54

Figure 3-7 Objective rendering of single molecule parameters using Kernel Density Estimation 56

Figure 3-8 Stoichiometry distribution of SpoIIQ using Kernel Density Estimation (KDE) 57

Figure 4-1 Digestion of pSG1154 with SpeI and HindIII 61

Figure 4-2 PCR amplification of the DNA fragments required to make the *mGFP spoIIQ* fusion 62

Figure 4-3 Construction of the pSG1154 recombinant plasmid using Gibson Assembly 63

Figure 4-4 Agarose gel electrophoresis showing the PCR products from colony PCR of 15 colonies .. 63

Figure 4-5 Restriction enzyme digestion of pSG1154 recombinant plasmids with NotI and NdeI 64

Figure 4-6 *mGFP spoIIQ* fusion integration into the *B. subtilis* chromosome 65

Figure 4-7 Cartoon representation of the localisation of SpoIIQ during engulfment 66

Figure 4-8 SIM microscopy imaging of mGFP SpoIIQ..... 66

Figure 4-9 Agarose gel electrophoresis showing the PCR products from colony PCR of 13 colonies . 68

Figure 4-10 SIM microscopy imaging of mGFP SpoIIQ delSpoIIQ. 69

Figure 4-11 Schematic representation of the components required to make the SpoIIIAH fusions ... 70

Figure 4-12 Schematic representation of Widefield and Slimfield microscopy 71

Figure 4-13 Slimfield microscopy of mGFP SpoIIQ delSpoIIQ 71

Figure 4-14 Localisation of the mother cell protein SpoIIIAH during engulfment 72

Figure 4-15 Slimfield microscopy of fluorescently labelled SpoIIIAH strains during engulfment 73

Figure 4-16 Stoichiometry distributions of SpoIIQ using Kernel Density Estimation (KDE)..... 74

Figure 4-17 Determination of SpoIIQ stoichiometry by combining and averaging the KDE plots in Figure 4-16 74

Figure 4-18 SpoIIQ diffusion coefficient determination using Kernel Density Estimation (KDE). 75

Figure 4-19 Determination of SpoIIQ diffusion coefficient by combining and averaging the KDE plots in Figure 4-18 75

Figure 4-20 Determination of the SpoIIQ copy number using Kernel Density Estimation 76

Figure 4-21 Stoichiometry distributions of SpoIIIAH using Kernel Density Estimation (KDE)..... 77

Figure 4-22 Determination of SpoIIIAH stoichiometry by combining and averaging the KDE plots in Figure 4-21 78

Figure 4-23 SpoIIIAH diffusion coefficient determination using Kernel Density Estimation (KDE). 79

Figure 4-24 Determination of the SpoIIIAH diffusion coefficient by combining and averaging the KDE plots in Figure 4-23 79

Figure 4-25 Determination of the SpoIIIAH copy number using Kernel Density Estimation 80

Figure 4-26 Schematic representation of the assembly of the SpoIIQ-SpoIIIAH channel from SpoIIQ dimers and SpoIIIAH hexamers 81

Figure 4-27 Slimfield imaging of the mGFP SpoIIQ mCherry SpoIIIAH dual labelled strain 83

Figure 4-28 Linked and unlinked stoichiometry distributions of SpoIIIAH shown using Kernel Density Estimation (KDE) 84

Figure 5-1 *mGFP spoIIQ* fusion integration into the *B. subtilis* chromosome 86

Figure 5-2 Cartoon representation of SpoIIQ localisation in the absence of mother cell gene expression 87

Figure 5-3 SIM microscopy imaging of mGFP SpoIIQ delSpoIIIAH 88

Figure 5-4 Slimfield microscopy imaging of mGFP SpoIIQ delSpoIIIAH 89

Figure 5-5 Cartoon representation of SpoIIIAH localisation in cells without SpoIIQ 89

Figure 5-6 Slimfield microscopy imaging of mCherry SpoIIIAH delSpoIIQ 90

Figure 5-7 Stoichiometry distributions of SpoIIQ in the absence of SpoIIIAH shown using Kernel density estimation (KDE) 91

Figure 5-8 Determination of SpoIIQ stoichiometry in the absence of SpoIIIAH by combining and averaging the KDE plots in Figure 5-7 91

Figure 5-9 SpoIIQ diffusion coefficient determination using Kernel Density Estimation (KDE) in the absence of SpoIIIAH 92

Figure 5-10 Determination of the SpoIIQ diffusion coefficient in the absence of SpoIIIAH by combining and averaging the KDE plots in Figure 5-9 92

Figure 5-11 Determination of the SpoIIQ copy number using Kernel density estimation (KDE) in the absence of SpoIIIAH 93

Figure 5-12 Slimfield microscopy image of the mGFP SpoIIQ delSpoIIIAH strain showing cells with different SpoIIQ expression levels 93

Figure 5-13 Stoichiometry distributions of SpoIIIAH in the absence of SpoIIQ shown using Kernel Density Estimation (KDE) 94

Figure 5-14 Determination of SpoIIIAH stoichiometry in the absence of SpoIIQ by combining and averaging the KDE plots in Figure 5-13 94

Figure 5-15 SpoIIIAH diffusion coefficient determination using Kernel Density Estimation (KDE) in the absence of SpoIIQ 95

Figure 5-16 Determination of the SpoIIIAH diffusion coefficient in the absence of SpoIIQ by combining and averaging the KDE plots in Figure 5-15 95

Figure 5-17 Determination of the SpoIIIAH copy number using Kernel density estimation (KDE) in the absence of SpoIIQ 96

Figure 5-18 Stoichiometry distributions of SpoIIIAH in a wildtype SpoIIQ background shown using Kernel density estimation (KDE) 97

Figure 6-1 Diagrams illustrating the interface between the forespore protein SpoIIQ and mother cell protein SpoIIIAH 99

Figure 6-2 Interface summary table showing the number of atoms and residues of SpoIIQ and SpoIIIAH at the interface and surface 100

Figure 6-3 Interfacing residues table showing the core and rim residues in SpoIIQ and SpoIIIAH ... 101

Figure 6-4 Analysis of the pSG1154-mGFP spoIIQ plasmid using restriction enzyme digestion 103

Figure 6-5 Analysis of the pSG1154-mGFP spoIIQ plasmid using PCR 104

Figure 6-6 Analysis of the pSG1154 plasmid using restriction enzyme digestion 105

Figure 6-7 Agarose gel electrophoresis of the PCR amplification of the pspoIIQ-mGFP-spoIIQ insert from genomic DNA 106

Figure 6-8 Agarose gel electrophoresis of the restriction enzyme digestion of pSG1154 and the pspoIIQ-mGFP-spoIIQ PCR product with SpeI and HindIII 107

Figure 6-9 Plasmid map of pCR-Blunt 108

Figure 6-10 Ligation of pspoIIQ-mGFP-spoIIQ to pCR-Blunt using blunt ended cloning 109

Figure 6-11 Agarose gel electrophoresis of the restriction enzyme digestion of pCR-Blunt recombinant plasmids using SpeI and HindIII. 109

Figure 6-12 Primer design for the Q5 Site Directed Mutagenesis kit (Q5 SDM Kit) 110

Figure 6-13 Agarose gel electrophoresis of the restriction enzyme digestion of the mutated plasmids and the wildtype pCR18 parent plasmid with SpeI and HindIII 111

Figure 6-14 Agarose gel electrophoresis showing the gel extraction of the desired bands after restriction enzyme digestion of the mutated plasmids with SpeI and HindIII. 113

Figure 6-15 Transformation of *B. subtilis* strain IB714 with the mutated pSG1154-mGFP spoIIQ plasmids. 114

Figure 6-16 Agarose gel electrophoresis of the restriction enzyme digestion of the pSG1154-mGFP spoIIQ mutated plasmids and the wildtype pSG1154-mGFP spoIIQ plasmid with SpeI and HindIII . 115

Figure 7-1 Schematic representation of the assembly of the SpoIIQ-SpoIIIAH channel from SpoIIQ dimers and SpoIIIAH hexamers 118

Figure 7-2 Cryo-EM structure of SpoIIIAG 120

Figure 7-3 Co-immunoprecipitation (Co-IP) as a method to study the interactions between SpoIIQ and SpoIIIAH 121

Figure 7-4 Affinity chromatography to study the interactions between SpoIIQ and SpoIIIAH. 122

Figure 7-5 The Yeast 2 Hybrid system as a method to determine interactions between SpoIIQ and SpoIIIAH. 123

Figure 7-6 Fluorescence Resonance Energy Transfer (FRET) as a method to detect interactions between SpoIIQ and SpoIIIAH 124

Acknowledgments

Firstly, I would like to thank my supervisors Prof. Mark Leake and Prof. Tony Wilkinson for giving me the opportunity to pursue this PhD and for their constant support and encouragement over the last four years. Thanks, must also go to my Thesis Advisory Panel Prof. Daniella Barilla and Dr Ville-Petri Friman for their support and thought-provoking questions over the years. Further thanks go to my project collaborators Dr Imrich Barak, Dr Katarína Muchová and Dr Zuzana Chromíková from the Slovak Academy of Sciences for their assistance with the project and for hosting me for a research visit in their lab in Bratislava 😊

Starting a new project is always daunting and I am so grateful to Dr Aisha Syeda for introducing me to the lab and helping me hit the ground running. It has been a pleasure to share the lab bench with Aisha for the last four years and I so am grateful for everything she has taught me. I would also like to thank Dr Ji-Eun Lee for training me in Slimfield Microscopy and Slimfield Microscopy image analysis and for being a great mentor and friend throughout the training process. Further thanks go the post-docs in the Leake group especially Jack and Alex for helping me troubleshoot the MATLAB errors as and when they appear.

Thanks must go to Dr Jim Brannigan for his assistance with the molecular cloning in Chapter 6. I would like to thank Grant Calder from the University of York Technology Facility for his help with collecting the SIM Microscopy data discussed in Chapters 4 and 5. Thanks must also go to my postgraduate mentor Dr Julie Tucker and Monica Bandeira from Biology for their positivity and for believing in me when I didn't believe in myself!

Lastly, I would like to thank my friends and family for believing in me from the start and for their constant encouragement and support. Thank you to my best friend Elizabeth who assisted with creating some of the figures in Chapter 1. Thank you to my family especially my brother Dr. Amar Sidhu and my beautiful mother for their constant support, words of wisdom, encouragement and for motivating me to complete this degree. Lastly, thank you to Daniel for loving me and always having my back and keeping me going, particularly in the last few months. I can't thank you enough for everything you've done for me.

Authors Declaration

I declare that this thesis is a presentation of original work, and I am the sole author. Any contributions are acknowledged below or within the text. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

The SpoIIIAH strains imaged and analysed in Chapter 4 were made by Adam Hughes prior to the start of this work. The mutant strains imaged and analysed in Chapter 5 and the dual strain discussed in Chapter 7 were made by Dr Katarína Muchová at the Slovak Academy of Sciences in Bratislava, Slovakia. Dr Jim Brannigan assisted with the molecular cloning in Chapter 6 and performed the initial cloning into pCR-Blunt and the ligation of the mutated inserts into pSG1154. SIM Microscopy experiments were performed under the supervision of Grant Calder at the University of York Technology Facility.

1. Biology Introduction

1.1 Cell Differentiation

Cell differentiation is the process by which cells become specialised to perform specific functions (Cooper,2000). The development of multi-cellular eukaryotes relies on the formation of specialized cells from multi-potent stem cells. Cell differentiation is ubiquitous in single-celled organisms and allows microbial communities to optimize their resources, leading to more efficient proliferation and survival (Lopez *et al*.,2009). Bacteria can be classified as either Gram-positive or Gram-negative with both capable of differentiating into specialized cell types. Gram-positive bacteria have a single membrane and a relatively thick layer of peptidoglycan whereas Gram-negative bacteria have a thinner layer of peptidoglycan separating the two membranes (**Figure 1-1**) (Tocheva *et al*., 2016). *Caulobacter cresentus* is an example of a Gram-negative aquatic microorganism with the ability to differentiate. During replication, the microorganism divides asymmetrically producing two genetically identical yet morphologically distinct cell types; the flagellated swarmer cells and non-motile stalked cells (Adams,2000). Only stalked cells can undergo DNA replication meaning swarmer cells must first differentiate into stalked cells before replicating. In the photosynthetic cyanobacteria *Anabeana,* 5-10% of vegetative cells per filament differentiate into nitrogen-fixing cells called heterocysts when nitrogen is limited (Adams,2000). This differentiation is essential as heterocysts provide vegetative cells with fixed nitrogen but rely on them for carbon and reductants (Lopez *et al*., 2009)

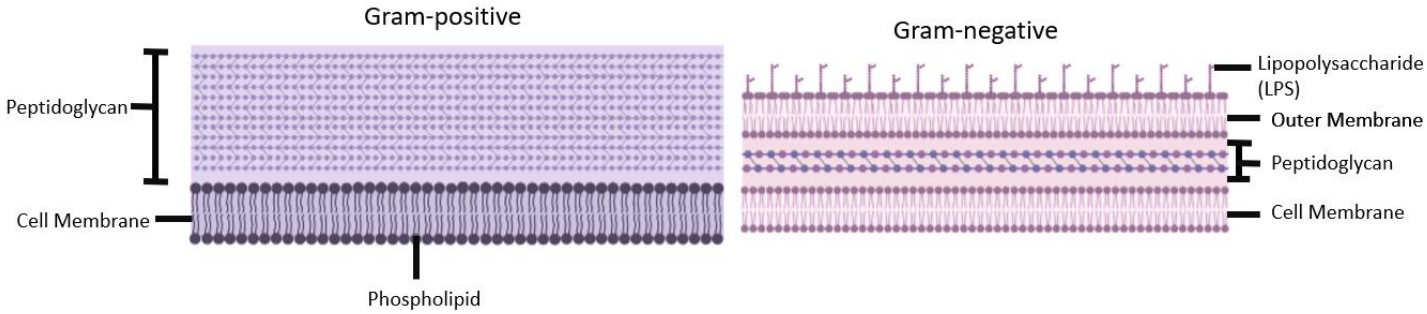


Figure 1-1: Gram-positive and Gram-negative bacterial membranes. Gram-positive bacteria have a single membrane and a thick layer of peptidoglycan whereas Gram-negative bacteria have a thinner layer of peptidoglycan separating two membranes. Gram-negative bacteria also have lipopolysaccharide on the outer membrane which plays a role in host-pathogen interactions with the immune system. Created using Bio Render.

1.2 Cell Differentiation in *B. subtilis*

Bacillus subtilis is a Gram-positive soil bacterium that displays many distinct cell types with differentiation occurring in *B. subtilis* cells growing exponentially as well as in those in the stationary phase (Narula *et al*.,2016). At the start of the stationary phase, *B. subtilis* differentiates into

competent cells, allowing DNA uptake from the environment (**Figure 1-2**)**.** When nutrients are depleted, vegetative *B. subtilis* cells differentiate into dormant spores in a process called sporulation, allowing them to withstand harsh environmental conditions (Lopez *et al*.,2009). A population of *B. subtilis* cells can secrete a toxin that acts as a 'killing factor' cannibalizing cells that have not yet initiated sporulation and allowing cannibalistic cells to delay their own entry into sporulation (Lopez *et al*.,2009) (**Figure 1-2**). During biofilm production, *B. subtilis* cells produce extracellular matrix proteins, allowing them to hold themselves together. Biofilm production is triggered by various environmental and cellular signals with serine starvation believed to induce biofilm production at the onset of stationary phase c The different cell types described above can be distinguished from one another by measuring cell-specific gene expression using reporter molecules (Lopez *et al*.,2009).

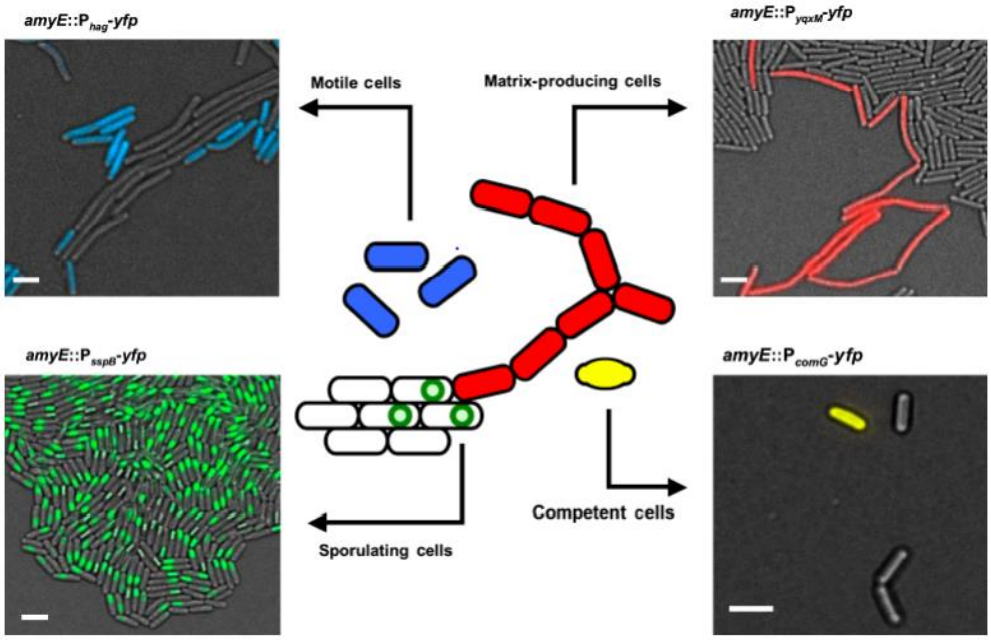


Figure 1-2: Cell differentiation in *B. subtilis* visualised in cells harbouring transcriptional reporter fusions to each cell type. Blue represents motile cells, red matrix-producing cells, green sporulating cells, and yellow competent cells. Each cell type has a characteristic phenotype that distinguishes it from the rest of the population (Taken from Lopez *et al*., 2009).

1.3 Spore Forming Bacteria

Bacterial endospores (spores) are extremely resistant cell types capable of surviving many environmental stresses including ionizing radiation, DNA-damaging agents, and heat (Hutchison *et al*., 2014). Christian Gottfried Ehrenberg was the first to observe spores in 1838, and nearly four decades later, spores were characterized by Robert Koch and Ferdinand Cohn (Riley *et al*., 2021). The scientists studied the spores of *B. subtilis* and *B. anthracis* and found that spores can survive

periods of boiling. They also discovered that spores and vegetative cells were different cellular forms of the same bacterial species and that vegetative cells could be converted into spores via sporulation and spores into vegetative cells via germination (Riley *et al*., 2021). During the next few decades, attempts were made to characterize the process of sporulation, but this was difficult due to limitations in optical microscopy and staining methods (Riley *et al*., 2021). The arrival of electron microscopy in the early 60s allowed a more detailed model of sporulation to be developed and by the mid-60s, Ryter defined the eight stages of sporulation that served as a framework for subsequent genetic dissection of spore formation (Ryter,1965).

Endospore formation is found in one of the earliest branching bacterial phyla, the *Firmicutes.* The *Firmicutes* phylum consists of five classes: *Bacilli*, *Clostridia*, *Erysipelotrichi*, *Negativicutes,* and *Thermolithobacteria* with endospore formers found in all classes except *Thermolithobacteria* (Hutchinson *et al*., 2014). Despite being most common in rod-shaped cells, endospores can be found in filamentous cells and cocci (Hutchinson *et al*., 2014). Many types of endospores exist in nature such as those from the *Sporospirillum spp* bacterium present in the gut of batrachian tadpoles. This species forms one or two endospore-like structures with its similarities to the *Firmicutes* not yet established. (Hutchinson *et al*., 2014).

1.4. Spore-Forming Bacteria in Disease

Many spore-forming microorganisms can cause infection in humans. Pathogenic spore formers include *Clostridium difficile*, *Clostridium tetani*, *Bacillus cereus, Clostridium botulinum,* and *Bacillus anthracis* which cause acute food poisoning, tetanus, botulism, and anthrax respectively. *Bacillus cereus* is a common contaminant of food that produces toxins that cause vomiting and diarrhoea in humans (Gauvry *et al*.,2019) (Nguyen and Talent, 2016). *C.tetani* is the bacterium that causes tetanus and often enters the body through a wound from a minor injury, before releasing a toxin that causes a tetanic spasm (Dong *et al*., 2019). *Clostridium botulinum* causes botulism, a rare and life-threatening condition caused by toxins produced by the bacteria (Dickson and Shevky,1923). The botulism toxin is the most potent toxin known to humans, lethal at 1.3–2.1 ng/kg (Dickson and Shevky,1923). *B. anthracis* is an anthrax disease-causing agent that has been used as a biological weapon due to its spores being infectious to humans and resistant to heat, drying, and disinfectants (Riley *et al*., 2021). *C.difficile* can be found in the human gastrointestinal (GI) tract causing *Clostridium difficile* infections (CDI) that range from mild diarrhoea to pseudomembranous colitis and death (Paredes-Sabja *et al*., 2014) (Burns and Minton,2011). *C.difficile* infections are problematic in hospitals as 20-30% of antibiotic-associated diarrhoeas in hospitals are caused by *C. difficile* spores (Paredes-Sabja *et al*., 2014) (Kelly and Salgado,2019). Unlike the organisms described above, *B.*

subtilis is a non-pathogenic, non-toxic bacterium highly amenable to genetic manipulation making it a model organism for laboratory studies on sporulation (Tan and Ramamurthi,2014).

1.5 *B. subtilis* life cycle

B. subtilis divides by binary fission every 30 minutes when nutrients are readily available (de Hoon *et al*., 2010). When it is starving, the microorganism undergoes a developmental process called sporulation that takes 8-10 hours to complete (Grossman and Losick,1998). Sporulation is a last resort adaptive response to starvation as the process uses a considerable amount of time and energy. *B. subtilis* commits to sporulation only when it has failed to deal with starvation in other ways, such as cannibalism or the formation of genetically competent cells (de Hoon *et al*.,2010). Spore formation is an effective way for *B. subtilis* to escape predation by higher organisms as spores are considerably more resistant than vegetative cells (de Hoon *et al*., 2010). Sporulation allows *B. subtilis* to escape predation by *Cupriavidus necator*, a non-obligate bacterial predator of Gram-positive and Gram-negative bacteria (Seccareccia *et al*.,2016).

Electron microscopy has been used to define the successive morphological stages of sporulation (**Figure 1-3**)**.** Sporulation begins with an asymmetric septation event that produces two daughter cells of unequal size: a smaller forespore and a larger mother cell (Errington, 2003). The mother cell then engulfs the forespore in a phagocytosis-like mechanism (Khanna *et al*., 2020). After the completion of engulfment, the forespore is a free protoplast inside the mother cell and is surrounded by an inner and outer membrane. Once inside the mother cell, the forespore develops into a mature spore and a series of protective layers are assembled around the spore core (**Figure 1-3**). The spore cortex is the first to be assembled and is a peptidoglycan layer between the two forespore membranes (de Hoon *et al*., 2010). Next, a proteinaceous spore coat made of 70 individual coat proteins is synthesised to encase the spore in a multi-layered structure. The forespore chromosome is saturated with small, acid-soluble proteins (SASPs) to protect the DNA from damage (de Hoon *et al*.,2010). The spore core becomes dehydrated as Ca^{2+} dipliconic acid (CaDPA) accumulates in the core and binds to free water molecules. The protective layers around the spore and the dehydrated spore core enable spores to withstand environmental stresses including heat, antibiotics, desiccation, radiation, and oxidation (Khanna *et al*., 2020). When nutrients become available again, they are sensed by membrane receptors, and *B. subtilis* exists the sporulation pathway and resumes vegetative growth in a process called germination (Khanna *et al*., 2020). The individual stages of sporulation will be discussed in more detail in the rest of the introduction.

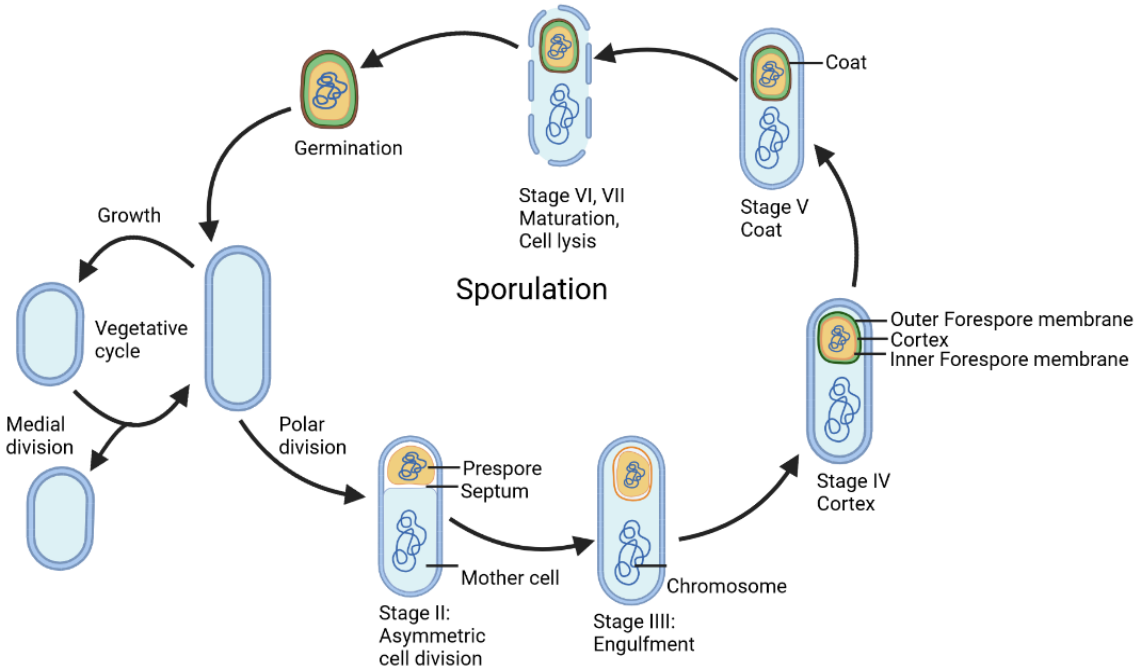


Figure 1-3: Schematic diagram illustrating the B. subtilis vegetative cycle (left) and sporulation cycle (right). Stage II marks the formation of the asymmetric septum which gives rise to the mother cell and forespore. Stage III marks the completion of the engulfment of the forespore by the mother cell. In stages IV and V, the spore cortex and coat are formed. Stage VI marks the release of the forespore into the environment upon mother cell lysis. Once nutrients are restored, spores germinate back into vegetative cells. (Adapted from McKenney et al., 2013).

1.6 Signalling

The Spo0A transcription factor controls sporulation initiation and is active in transcription when phosphorylated and inactive when dephosphorylated (Stephens, 1998). Spo0A is activated in a multi-step process called the phosphorelay (**Figure 1-4**)**.** In the first step of the phosphorelay, the cytoplasmic kinase KinA and membrane-bound kinase KinB respond to different environmental signals and phosphorylate Spo0F to Spo0F-P (Fujita and Losick, 2003) (Hoch, 1992). The *kinB* gene is encoded in an operon with the *kapB* gene and the product of this gene is responsible for the proper functioning of KinB (**Figure 1-4**) (Stephens,1998). Once phosphorylated, Spo0F-P acts as a secondary messenger and is a substrate for the Spo0B phosphotransferase. Spo0B mediates the transfer of phosphate from Spo0F-P to Spo0A via a phosphorylated enzyme intermediate (Trach *et al*.,1991). Spo0A then initiates sporulation by binding to the sequence GNCGAA (0A boxes) in the promoter regions of the genes it controls (Hoch, 1993) (**Figure 1-4**). High levels of Spo0A-P are required to initiate the transcription of sporulation genes *spoIIAA* and *spoIIE,* resulting in the production of sporulation-specific sigma factors and the transcription of genes involved in cellular

compartmentalization and sporulation (Hoch, 1993) (Sonenshien,2000). Spo0A directly governs the expression of about 121 genes (Molle *et al*., 2003).

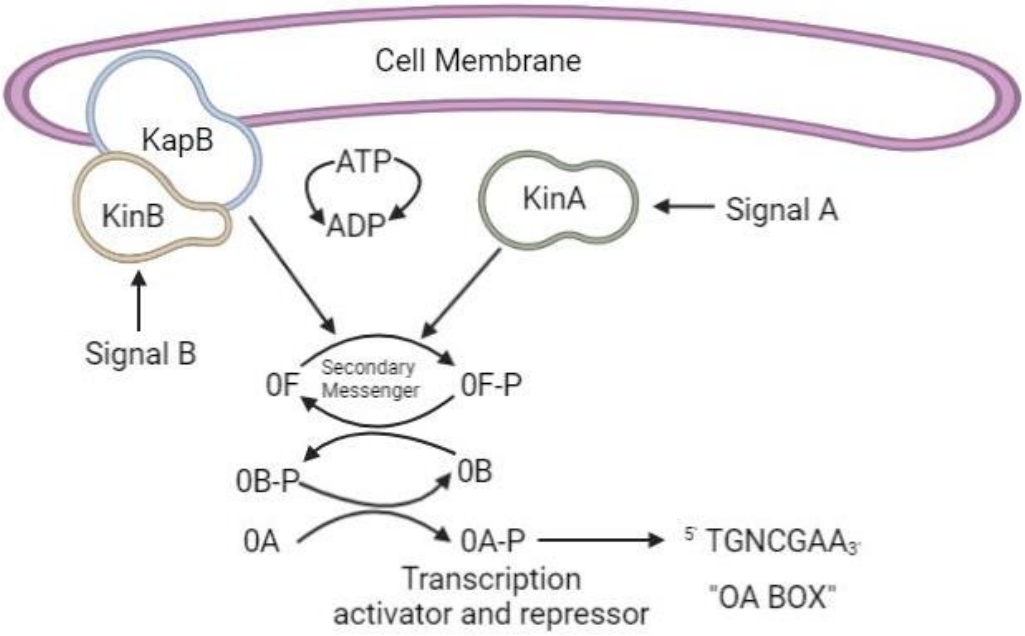


Figure 1-4: Diagram showing the role of the phosphorelay in sporulation initiation. Histidine kinases KinA and KinB provide phosphates to the secondary messenger Spo0F. Spo0F becomes phosphorylated to give Spo0F-P which acts as a substrate for Spo0B. Spo0B transfers phosphate from Spo0F to Spo0A to give Spo0A-P, a transcription factor that recognises 0A boxes on the promoters of the genes it controls. (Adapted from Hoch,1993)

1.7 Asymmetric Septation

In *B. subtilis* sporulation, a specialised septal structure is formed at one pole of the cell in a process called asymmetric septation. Asymmetric septation is irreversible and cells become committed to sporulation, even when transferred to a nutrient-rich medium (Dworkin and Losick,2005). The tubulin homologous protein FtsZ mediates septation by forming a ring-like structure called the Z ring (Barak and Muchova, 2018). The Z ring is made of treadmilling FtsZ filaments and serves as a scaffold for the recruitment of proteins involved in septal peptidoglycan synthesis and cell division. During vegetative growth, the Z ring assembles at mid-cell, generating two daughter cells of equal size (Khanna *et al*., 2020). During sporulation, however, the Z ring assembles closer to one of the cell poles as FtsZ interacts with SpoIIE and RefZ, two proteins synthesised under Spo0A control prior to asymmetric septation (Barak and Muchova, 2018) (**Figure 1-5**).

Asymmetric septation can be initiated with equal frequencies from both cell poles. In wild-type cells, the formation of a second septum is inhibited before the septation process is complete to ensure that only one septum forms during sporulation (Khanna *et al*., 2020) (**Figure 1-5**). Inhibition of the

second asymmetric septation event is mediated by MciZ and SpoIIDPM, two proteins produced in the mother cell after the first septation event. MciZ interacts with FtsZ and inhibits the formation of new Z rings and SpoIIDPM dissolves peptidoglycan in the partial septum (Khanna *et al*., 2020) (**Figure 1-5)**. The asymmetric septum is believed to be thinner than the vegetative septum [25 nm versus 50 nm] despite both septa being made from the same cell division machinery (Khanna *et al*., 2020) (Barak and Youngman, 1996). A thinner septum is likely to be physiologically important as it can bend and stretch easily during engulfment, facilitating the movement of the mother cell around the forespore (Khanna *et al*., 2020).

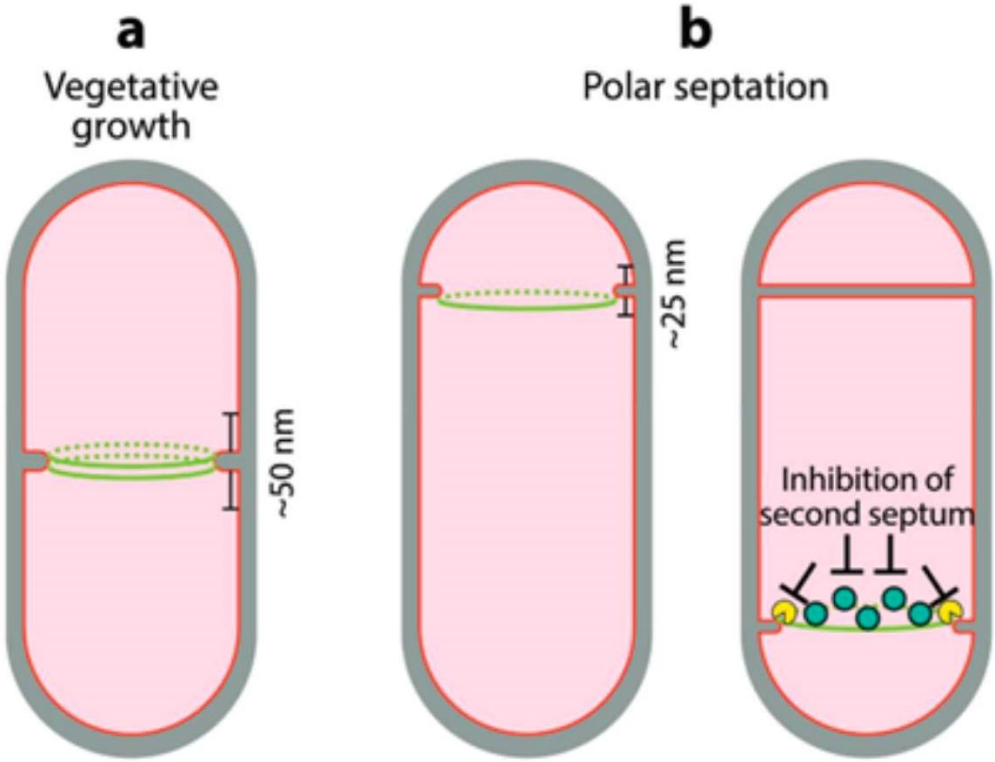


Figure 1-5: Schematic representation of the asymmetric septation process. A) During vegetative growth, FtsZ assembles Z rings at mid-cell whereas B) during polar septation, the Z rings assemble closer to one of the cell poles. The thickness of both A) and B) are indicated and the number of FtsZ filaments at the septum during vegetative growth is almost twice the number present during polar septation. B) After the formation of the first septum, the formation of a second septum is inhibited by SpollDPM (yellow) and MciZ (dark green). (Taken from Khanna *et al*., 2020).

1.8 Axial Filamentation

Before asymmetric septation, the chromosomes are remodelled to form a well-defined structure called an axial filament which runs from pole to pole, parallel to the long axis of the cell (Khanna *et al.*, 2020) (**Figure 1-6**)**.** The origins of replication of the chromosomes are at opposite poles of the cell and the proteins RacA, Spo0J and Soj are important in localizing these to the cell poles

(Grainge,2008). Asymmetric septation traps the forespore chromosome, as the septum bisects the chromosome leaving 30% of the chromosome in the forespore and 70% in the mother cell (**Figure 1-6**) (Khanna *et al*., 2020). SpoIIIE is the DNA translocase that pumps the remaining chromosome into the forespore. The SpoIIIE protein has an N-terminal transmembrane domain that anchors it to the septum and a cytoplasmic motor domain with ATPase activity (Khanna *et al*., 2020).In the presence of ATP, SpoIIIE moves along the DNA and translocates the 3Mb forespore chromosome into the forespore in approximately 15 minutes (Grainge, 2008). The SpoIIIE motor domains assemble into hexameric rings with the protein transferring both arms of the trapped chromosome into the forespore simultaneously, suggesting the presence of two SpoIIIE hexamers, one for each arm (**Figure 1-6**) (Grainge.,2008). Chromosome translocation by SpoIIIE is an energy-intensive process and requires 1.5 million ATP molecules to move 3Mb of DNA into the forespore (Khanna *et al*.,2020). Directionality of translocation is achieved as the SpoIIIE DNA translocase recognises the SpoIIIE recognition sequence GAGAAGGG using its C-terminal domain, which establishes active translocases in the mother cell and pumps DNA into the forespore (Grainge, 2008).

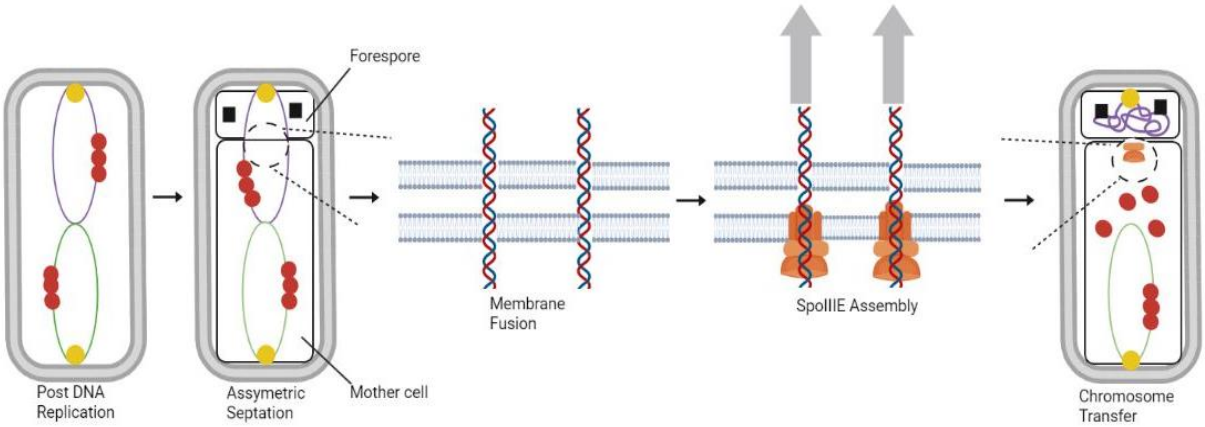


Figure 1-6: Chromosome translocation by the SpoIIIE DNA translocase. After replication, the cell has two chromosomes (purple and green). The origin (yellow circle) is localised to the cell poles. Asymmetric septation traps 1/3 of the chromosome in the forespore and forespore-specific proteins are produced (black squares). SpoIIIE assembles on the mother cell side as an active motor and pumps the DNA into the forespore (grey arrows show the direction of DNA movement). SpoIIIE removes any bound proteins (red circles) as it translocates the DNA into the forespore, ensuring mother cell-specific transcription factor complexes (red circles) do not enter the forespore. (Adapted from Grainge, 2008).

1.9 Differential Gene Expression

After asymmetric septation and chromosome translocation, cell-specific σ factors are activated, establishing different programs of gene expression in the mother cell and forespore (Hilbert and Piggot.,2004) (**Figure 1-7**)**.** After asymmetric septation, σ^F is activated in the forespore and σ^E is

activated in the mother cell. Later in sporulation, soon after the completion of engulfment, σ^G is activated in the forespore and σ^K is activated in the mother cell (Pogliano *et al*., 1998) (Hilbert and Piggot., 2004) (**Figure 1-7**)**.** The four σ factors are tightly regulated to ensure activation occurs in the order σ^F, σ^E, σ^G, σ^K. The mechanisms of activation of σ^F, $\sigma^{E,}$ and σ^K are well understood but it is less clear how σ^G is activated. Each σ factor will be outlined in more detail below.

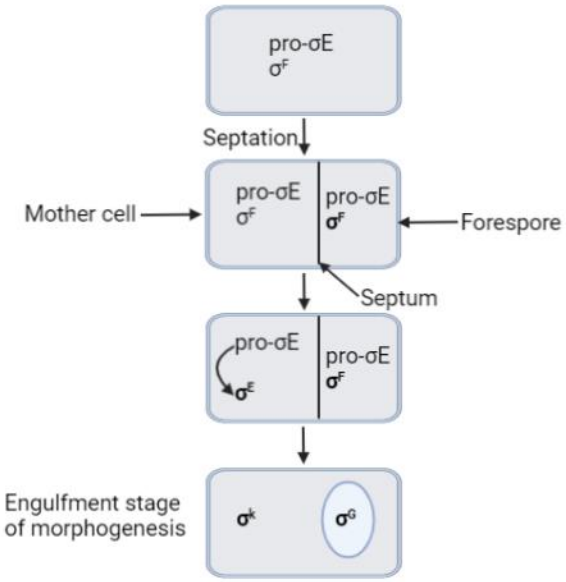


Figure 1-7: Activation of sporulation specific σ factors in the mother cell and forespore. σ^F and pro- σ^E are present in the pre-divisional cell in an inactive state. After asymmetric septation, σ^F is activated in the forespore only (shown by the bold lettering). σ^F then triggers the conversion of pro- σ^E to σ^E. σ^E is the active form of the transcription factor. During engulfment, σ^F in the forespore and σ^E in the mother cell trigger the activation of σ^G in the forespore and σ^K in the mother cell. (Adapted from Margolis *et al*.,1991).

1.9.1 Activation of σ^F and σ^E

σ^F is encoded by the third gene in the *spoIIAA* operon, *spoIIAC.* The *spoIIAA* operon is transcribed before asymmetric septation under the control of Spo0A, but σ^F is held inactive until asymmetric septation is complete (Hilbert and Piggot,2004) (**Figure 1-8**). SpoIIAA and SpoIIAB are the two other products of the *spoIIAA* operon and regulate the activity of σ^F. In pre-divisional cells, σ^F is held inactive by the anti-sigma factor SpoIIAB as SpoIIAB binds to σ^F, blocking its interaction with RNA polymerase (Arigoni *et al.,* 1999) (Kellner *et al*., 1996). The SpoIIAA anti-anti sigma factor is a SpoIIAB antagonist and displaces σ^F from the SpoIIAB: σ^F complex (**Figure 1-8**). SpoIIAB can neutralize SpoIIAA activity by phosphorylating one of its serine residues meaning SpoIIAA must be dephosphorylated to be active (Arigoni *et al*., 1999). SpoIIE is the serine phosphatase responsible for dephosphorylating and activating SpoIIAA (**Figure 1-8**). The SpoIIE protein is synthesised in the pre-divisional cell before accumulating at the sporulation septum and dephosphorylating the anti-anti sigma factor SpoIIAA in the forespore (Arigoni *et al*., 1999). This leads to σ^F activation and the

transcription of forespore-specific genes. σ^F is present in both the mother cell and forespore (**Figure 1-8**) but is held inactive in the mother cell by SpoIIAB (Barak *et al*., 1996).

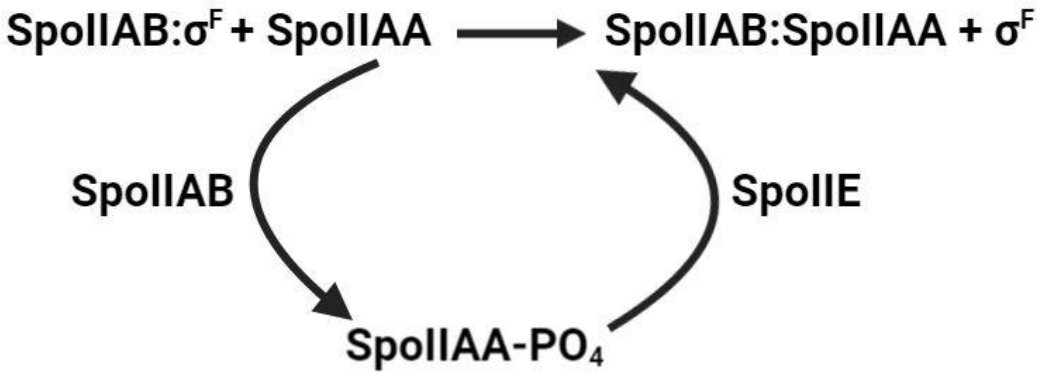


Figure 1-8: Flow chart illustrating σF regulation. Anti-sigma factor SpoIIAB binds σF and holds it inactive. SpoIIAA is an anti-anti-sigma factor that can reverse this inhibition. SpoIIAA is regulated by its phosphorylation state: it is inactive when phosphorylated by SpoIIAB and active when dephosphorylated by SpoIIE. Once activated, SpoIIAA can bind SpoIIAB, freeing σF, and leading to the transcription of forespore-specific genes (Adapted from Hilbert and Piggot, 2004).

Unlike σ^F, which is synthesised in an active form and held inactive by SpoIIAB, σ^E is synthesised in an inactive form and activated by proteolytic cleavage (Margolis *et al*.,1991). σ^E was the first sigma factor purified and pro- σ^E is the primary product of the gene encoding $\sigma^{E.}$ Pro- $\sigma^{E,}$ and σ^E are products of the *spoIIG* operon (Hilbert and Piggot,2004). The *spoIIG* operon encodes *spoIIGA,* the gene necessary for the processing of pro- σ^E to $\sigma^{E,}$ and *spoIIGB,* the gene encoding the pro- σ^E that is converted to σ^E in a *spoIIGA* dependent manner (Margolis *et al*.,.1991). The processing of pro- σ^E to σ^E involves the removal of 27 residues at the N-terminus and occurs only in the presence of the asymmetric septum, even though pro- σ^E and *spoIIGA* are present before septum formation (Margolis *et al*., 1991) (Higgins and Piggot.,1992).

1.9.2 Activation of σ^G and σ^K

Upon completion of engulfment, σ^G is activated in the forespore and σ^K is activated in the mother cell. σ^G is encoded by the *spoIIG* gene and once active, drives transcription of its own gene in a structural feedback loop that rapidly switches on σ^G activity (Margolis *et al*., 1991). The mechanisms of σ^G activation are poorly understood but are believed to involve the proteins SpoIIAB, CsfB, LonA, and Fin (Regan *et al*.,2012). SpoIIAB acts on σ^G as well as σ^F and inhibits σ^G activation in the mother cell. CsfB prevents the premature activation of σ^G and the LonA protease inhibits σ^G activation under non-sporulation conditions (Regan *et al*.,2012). Fin is believed to be necessary for the efficient transition from σ^F to σ^G in the forespore.

Similarly, to σ^E, σ^K is synthesised as the inactive precursor pro- σ^K and activation of σ^K requires the removal of 20 amino acid residues at the N terminus (Margolis *et al*.,1991). Mother cell proteins BofA, SpoIVFB, and SpoIVFA are involved in pro- σ^K processing and SpoIVFB is the enzyme that mediates the processing reaction (Hilbert and Piggot.,2004). SpoIVFB is held in an inactive complex with SpoIVFA and BofA and becomes activated through the actions of the σ^G activated SpoIVB protease. SpoIVB relieves the SpoIVFA and BofA mediated inhibition of SpoIVFB, leading to pro- σ^K processing and the activation of σ^K (Margolis *et al*.,1993) (Lu *et al*., 199) (Green and Cutting, 2000).

1.10 Cellular Dynamics of Engulfment

Engulfment is a process that resembles eukaryotic phagocytosis and takes 60 minutes to complete (Khanna *et al*., 2020). During engulfment, the mother cell membrane migrates around the forespore until the leading edges meet and fuse, releasing the forespore into the mother cell cytoplasm (**Figure 1-9**). Engulfment occurs in the sporulation pathway of *B. subtilis, C. difficile,* and *B. anthrax* and results in the reorganisation of the sporangium from two cells that lie side by side to a cell within a cell structure (Aung *et al*., 2007). Engulfment is the focus of this thesis as it provides an ideal system to study how bacterial cells move macromolecules, localize proteins, and catalyse membrane fusion (Aung *et al.,*2007). The key processes in cellular engulfment are outlined below.

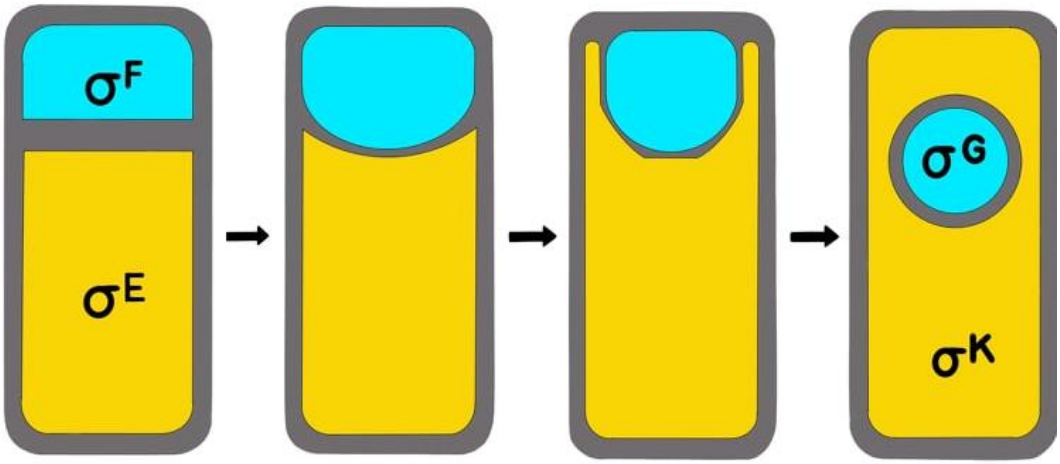


Figure 1-9: Engulfment during sporulation in *Bacillus subtilis*. From left to right: After asymmetric cell division the forespore (top) and mother cell (bottom) lie side by side and are separated by peptidoglycan (grey). Septal thinning and peptidoglycan degradation allows the migration of the mother cell membrane around the forespore. Eventually, the membranes fuse releasing the forespore into the mother cell cytoplasm. Compartment and stage-specific sigma factors are activated as indicated.

1.10.1 Cellular Engulfment: Role of SpoIIDPM

The first step of engulfment is septal thinning. Septal thinning is the process through which peptidoglycan is degraded from the middle towards the edges and the process is mediated by the

three mother cell proteins SpoIID, SpoIIP, and SpoIIM (**Figure 1-10**) (Aung *et al*.,2007). SpoIID, P and M localise to the centre of the septum in a hierarchical order and then relocalise to the leading edges of the mother cell membrane where they remain until engulfment is complete (**Figure 1-10**) (Khanna *et al*., 2020). SpoIIM recruits SpoIIP which in turn recruits SpoIID, and the three proteins form a complex commonly known as the DMP complex (Khanna *et al*., 2020) (Chastanet and Losick,2007). The DMP complex localises to the septal membrane via SpoIIB, a protein produced before asymmetric septation. SpoIIB is recruited to the invaginating septum in a FtsZ-dependent manner and extends the hierarchical order to SpoIIB →SpoIIM →SpoIIP →SpoIID (Chastanet and Losick,2007). The DMP complex has enzymatic activity as SpoIID and SpoIIP degrade septal peptidoglycan. SpoIIM has no enzymatic activity and serves as a scaffold for DMP complex assembly (Khanna *et al*.,2020).

The Burnt-bridge Brownian ratchet model of membrane migration is commonly used to explain the role of the DMP complex in engulfment (Gutierrez et *al.,*2010). In this model, the DMP complex functions as a motor, as peptidoglycan degradation by SpoIID and SpoIIP results in the absence of substrate in one direction and movement in the opposite direction (**Figure 1-10**) (Gutierrez et *al.,*2010). As SpoIID and SpoIIP move towards new targets, the mother cell membrane is dragged with them as the enzymes are anchored in the membrane (**Figure 1-10**) (Gutierrez et *al.,*2010).

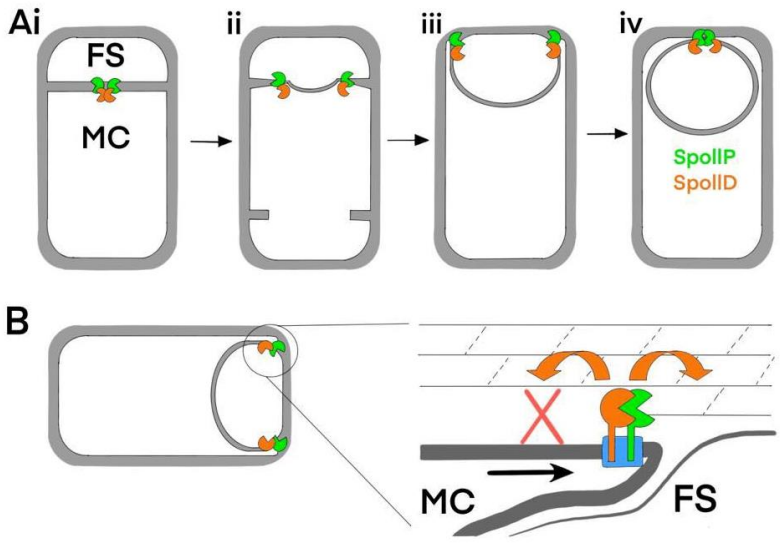


Figure 1-10: Role of the SpollDPM complex in engulfment. A (i) Engulfment splits the cell into two compartments: the mother cell (MC) and the forespore (FS). SpoIID (orange pacman) and SpoIIP (green pacman) are peptidoglycan degradation enzymes and localise to the centre of the septum. (ii) SpoIID and SpoIIP thin the septum by starting in the middle and moving towards the edges. (iii) SpoIID and SpoIIP localise to the leading edges during membrane migration (iv) Engulfment membrane fission occurs at the top of the forespore, releasing the forespore into the cytoplasm. B) Burnt-bridge Brownian ratchet model for membrane migration. As SpoIID (orange pacman) and SpoIIP (green pacman) are anchored in the membrane, they drag the engulfing membrane around the forespore as they degrade peptidoglycan (Adapted from Gutierrez *et al*., 2010).

1.10.2 Cellular Engulfment: SpoIIQ and SpoIIIAH

SpoIIQ and SpoIIIAH are two additional proteins required for engulfment and are the focus of this thesis. SpoIIQ is a forespore-specific protein transcribed under σ^F control and SpoIIIAH is a mother cell-specific protein transcribed under σ^E control (Londono-Vallejo *et al*.,1997). SpoIIQ has an N-terminal transmembrane domain and an extracellular Cterminal LytM domain (Meisner *et al*.,2012). LytM proteins are typically metalloendopeptidases that hydrolyze peptidoglycan in *E. coli,* but SpoIIQ does not exhibit metalloendopeptidase activity in its LytM domain as the two catalytic histidine residues required for Zn^{2+} binding are replaced by serines (Levdikov *et al.,*2012). SpoIIIAH has an N-terminal transmembrane domain and a Cterminal extracellular domain which has similarity to the YscJ/FliF-like proteins. YscJ/FliF-like proteins form multimeric rings in Type III secretion systems and flagella in Gram-negative bacteria (Meisner *et al*.,2012).

In 2012, two independent research groups solved the crystal structure of the complex of the C-terminal extracellular domains of SpoIIQ and SpoIIIAH (**Figure 1-11)** (Meisner *et al*.,2012) (Levdikov *et al.,*2012). The crystal structure showed that SpoIIQ is organised around a four-stranded anti-parallel β sheet (β6, β7, β8, and β11) which is flanked by two β strands (β1 and β4) on one side and a single β strand (β9) on the other side (Levdikov *et al.,*2012) (**Figure 1-11**). SpoIIIAH was found to consist of four α-helices and three β strands organised in the order α1- β1- β2- α3- α4- β3. SpoIIQ and SpoIIIAH interact via their Cterminal extracellular domains as the Nterminal domains of both proteins are anchored in the membrane (Sun *et al*., 2000) (Rubio and Pogliano, 2004) (Rodrigues *et al.,* 2013).

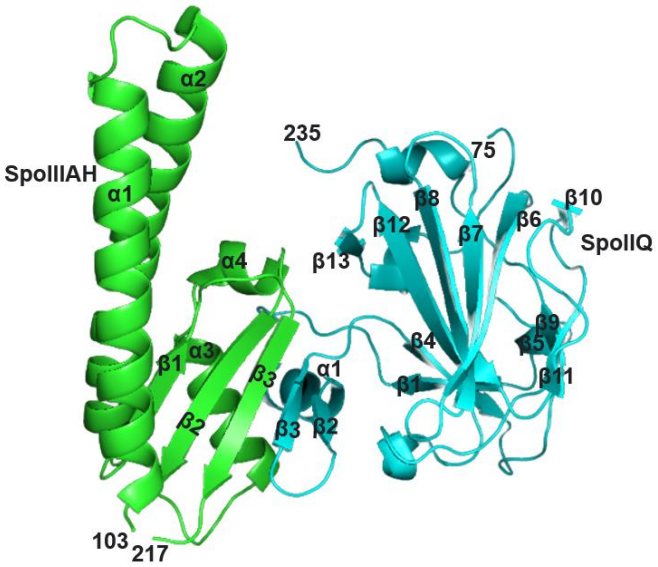


Figure 1-11: X-ray crystal structure of the complex formed between the C-terminal extracellular domains of SpoIIQ (residues 75-232) and SpoIIIAH (residues 103-217) (Created using Pymol). PDB = 3TUF

SpoIIQ and SpoIIIAH are important in σ^G activation and show a dynamic localisation pattern throughout engulfment (Rodrigues *et al*.,2013). Fluorescence microscopy imaging of GFP SpoIIQ fusions has shown that SpoIIQ initially localises at the sporulation septum before tracking the leading edges of the engulfing membrane by forming arcs and foci at the forespore/mother cell interface (**Figure 1-12**). After the completion of engulfment, GFP SpoIIQ is degraded by the serine protease SpoIVB on either side of its transmembrane domain, releasing GFP into the cytoplasm (**Figure 1-12**) (Rubio and Pogliano.,2004) (Jiang *et al*.,2005). SpoIIQ is randomly distributed throughout the forespore membrane in the absence of mother cell gene expression and engulfment is blocked (**Figure 1-12**) (Jiang *et al*.,2005) (Chiba *et al.,*2007). The protein responsible for SpoIIQ retention in the septum is currently unknown but the catalytic activities of the DMP complex are believed to contribute to SpoIIQ localisation as the DMP proteins degrade peptidoglycan, allowing a mother cell protein under σ^E control to anchor SpoIIQ into the septum (Rodrigues *et al*.,2013).

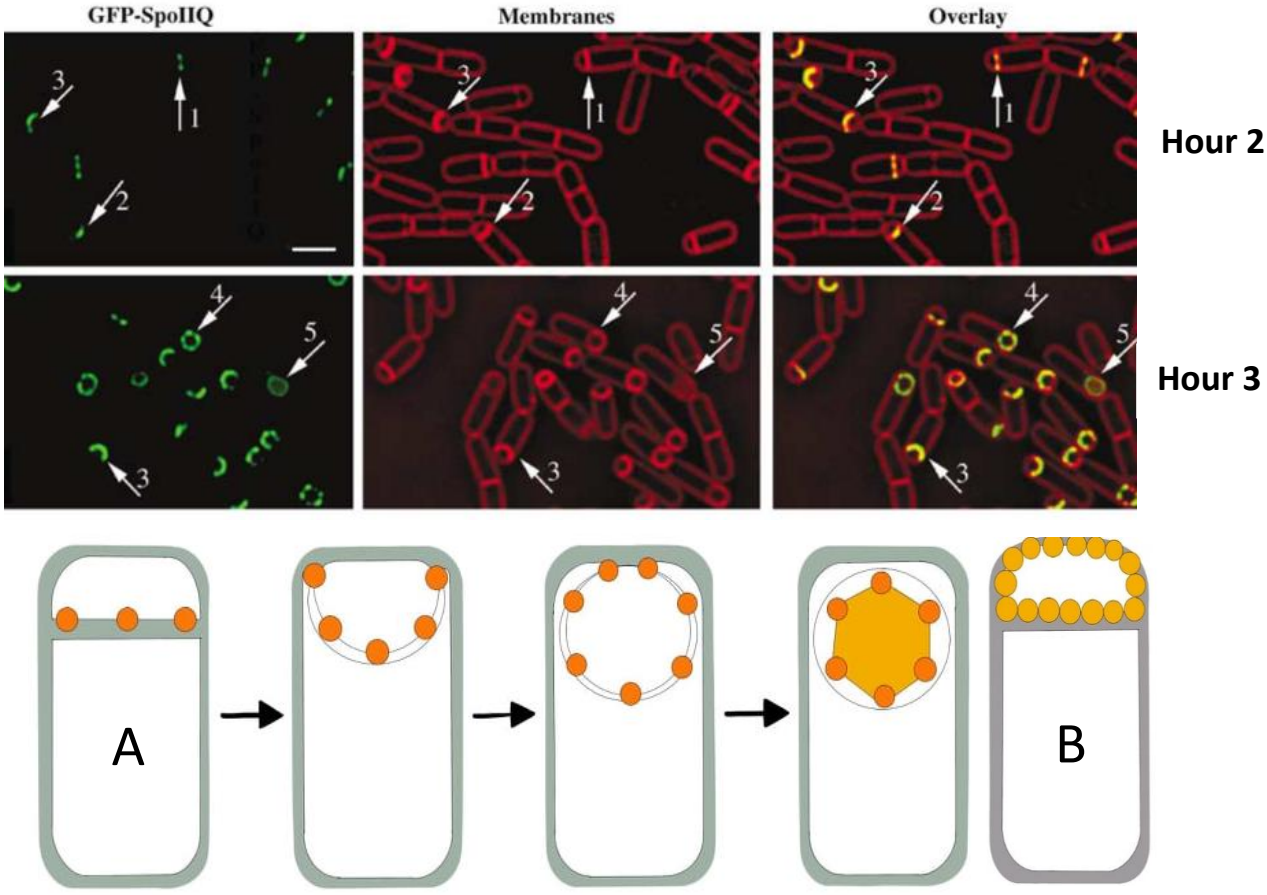


Figure 1-12: Localisation of SpoIIQ. Top) Localisation of GFP SpoIIQ (green) 2 and 3 hours after the onset of sporulation. Membranes stained red using FM-64 dye. Arrows 1,2 and 3 show early sporangia and arrows 4,5 represent sporangia later in engulfment. The release of GFP into the cytoplasm is not shown in this image (Taken from Rubio and Pogliano., 2004). Bottom) Cartoon representation of the localisation of SpoIIQ (orange circles) during engulfment. A) During engulfment, SpoIIQ initially localises at the sporulation septum and then forms arcs and foci at the mother cell/forespore interface. At the end of engulfment, SpoIIQ is degraded releasing GFP into the cytoplasm (shown by the orange shading). B) In the absence of mother cell gene

expression, SpoIIQ is randomly distributed throughout the forespore membrane and engulfment is blocked (Adapted from Rubio and Pogliano., 2004).

SpoIIIAH is the final gene in the *spoIIIAA* operon which encodes the eight genes *spoIIIAA-spoIIIAH*. SpoIIIAH is initially randomly distributed throughout the mother cell membrane. In cells that have initiated engulfment, SpoIIIAH localises at the septum, and as engulfment proceeds, the protein tracks the engulfing mother cell membrane as it moves around the forespore (Blaylock *et al*., 2004) (**Figure 1-13**)**.** Once engulfment is complete, SpoIIIAH localises around the forespore on the mother cell side (Blaylock *et al*.,2004). SpoIIQ is believed to be required for SpoIIIAH localisation as in the absence of SpoIIQ, SpoIIIAH is randomly distributed throughout the mother cell membrane (**Figure 1-13**) (Rodrigues *et al.,*2013) (Campo *et al*., 2008) (Doan *et al*.,2005).

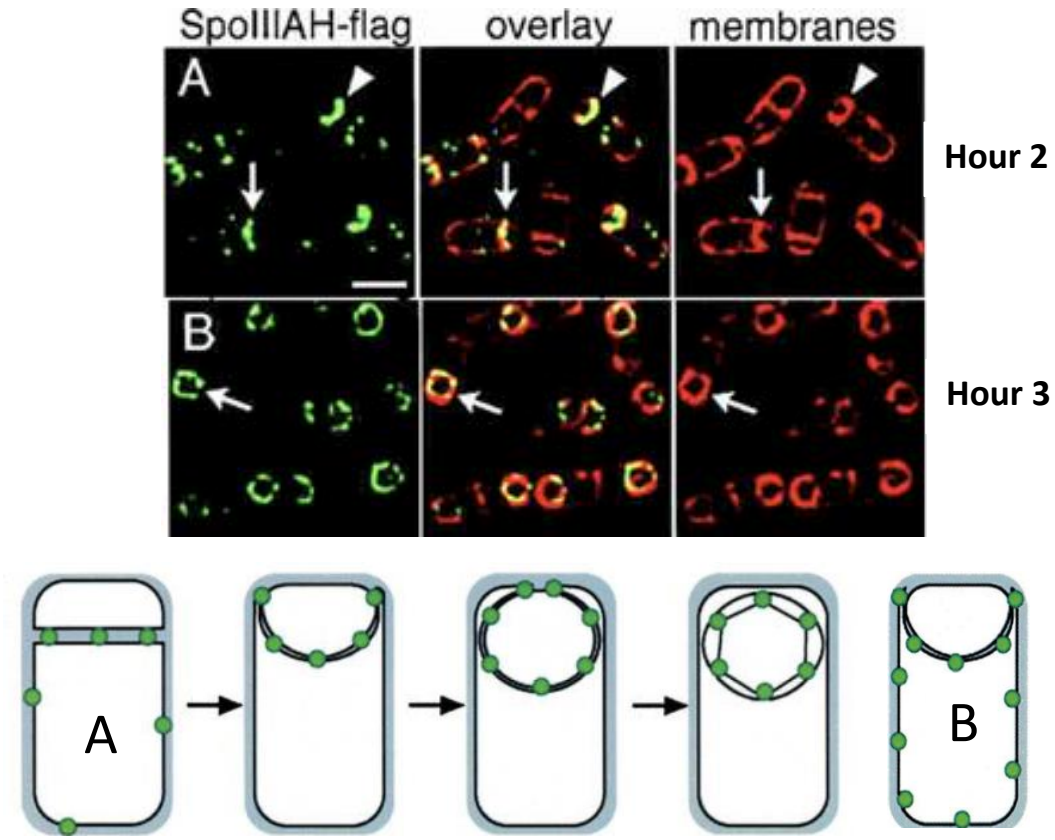


Figure 1-13: Localisation of SpoIIIAH. Top) Immunofluorescence microscopy imaging of SpoIIIAH-Flag (green) 2 and 3 hours after the onset of sporulation. Membranes stained red using FM-64 dye. Taken from (Blaylock *et al*.,2004). Bottom) A) Cartoon representation of the localisation of SpoIIIAH in wild-type cells B) Cartoon representation of SpoIIIAH localisation in cells without SpoIIQ showing that SpoIIIAH is randomly distributed in the mother cell membrane in the absence of SpoIIQ. (Adapted from Blaylock *et al*., 2004).

Co-immunoprecipitation assays have shown that SpoIIQ and SpoIIIAH directly interact and could form a complex (Blaylock *et al*.,2004). As SpoIIQ and SpoIIIAH are found in different cellular compartments and have a shared topology (Nterminus in, Cterminus out), and SpoIIIAH has similarity to the YscJ/FliF like proteins in type III secretion systems, the two proteins could form a complex that spans the intermembrane space (Blaylock *et al*.,2004). Several models have been proposed to explain the role of the SpoIIQ-SpoIIIAH complex during engulfment with one model suggesting that the complex acts as a zipper. In the zipper model, SpoIIQ initially localises at the sporulation septum and SpoIIIAH is randomly distributed throughout the mother cell membrane (**Figure 1-14**). Septal thinning allows SpoIIQ and SpoIIIAH to interact and SpoIIQ localises SpoIIIAH to the septum via diffusion and capture (Blaylock *et al*., 2004). The two proteins co-localize and form multiple foci until the final step of engulfment, membrane fusion, during which SpoIIQ is degraded and σ^G is activated (**Figure 1-14**)**.**

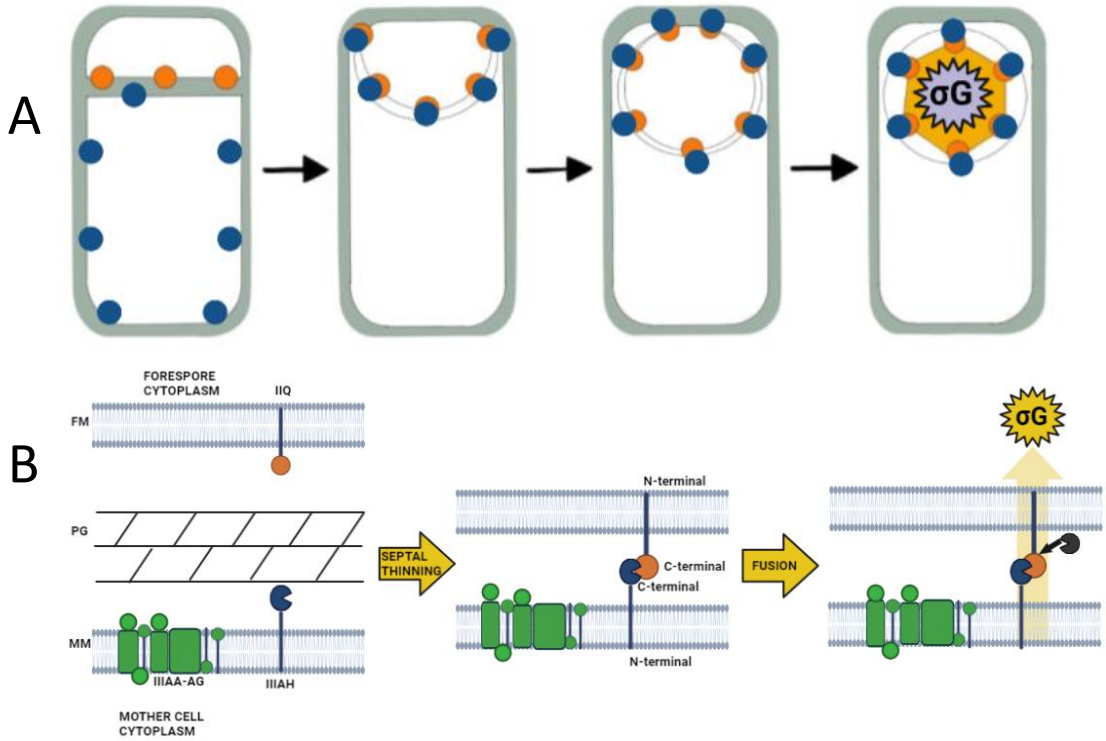


Figure 1-14: Zipper-like interaction between SpoIIQ (IIQ) and SpoIIIAH (IIIAH) during engulfment membrane migration. A) SpoIIQ (orange circles) initially localises at the sporulation septum and SpoIIIAH (blue circles) is randomly distributed throughout the mother cell membrane. Septal thinning allows SpoIIQ to interact with SpoIIIAH and the two co-localize in foci around the engulfing membrane. After membrane fusion, the final step in engulfment, SpoIIQ is degraded (orange shading). B) Enlarged view of the septum showing peptidoglycan (PG) between the forespore and mother cell membranes. PG inhibits the interaction between IIQ and IIIAH. After septal thinning, IIQ and IIIAH interact across the intermembrane space. After membrane fusion, SpoIIQ (orange) is degraded by an enzyme (black pacman) and SpoIIIAH remains in the membrane. σ^G is activated. N and C-termini labelled (Adapted from Blaylock *et al*., 2004).

The SpoIIQ – SpoIIIAH complex is believed to recruit the remaining SpoIIIAA proteins to form a multi-protein complex that spans the forespore-mother cell intermembrane space (Camp and Losick, 2008) (Camp and Losick, 2009). (**Figure 1-15**). The multi-protein complex has been suggested to form a secretion channel or 'feeding tube' between the mother cell and forespore allowing a molecule important for forespore maturation and σ^G activation to pass through (**Figure 1-15**). The nature of this molecule is currently unknown (Sogaard-Andersen,2013) (Morlot and Rodrigues,2018).

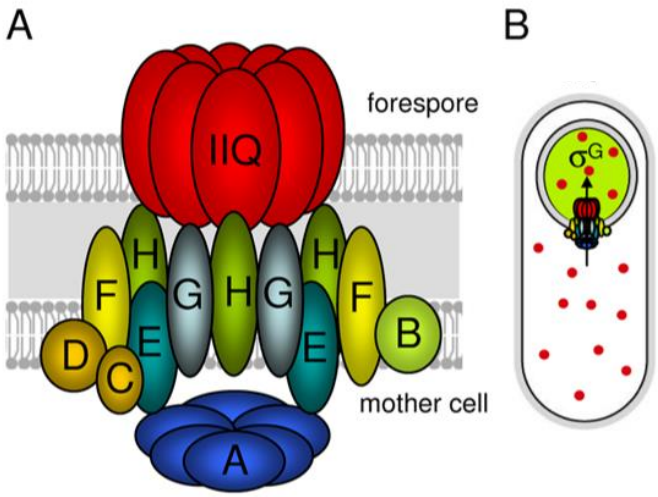


Figure 1-15: The mother cell SpoIIIAA proteins and forespore SpoIIQ protein form a secretion complex. (A) Schematic representation of the SpoIIIAA-SpoIIQ complex showing the predicted localisation of SpoIIIAA, SpoIIIAB, SpoIIIAC, SpoIIIAD, SpoIIIAE, SpoIIIAF, SpoIIIAG, and SpoIIIAH. SpoIIIAA is shown as a hexamer by analogy to other ATPases. SpoIIQ is shown as a multimeric pore. B) In the feeding tube model, the SpoIIIAA-SpoIIQ complex secretes a metabolite into the forespore important for maintaining forespore metabolic potential and activating σ^G (Taken from Doan *et al*., 2009).

In vivo biotinylation assays have been used to support the idea that SpoIIQ and SpoIIIAH form a channel complex (Meisner *et al*.,2008). In these assays, SpoIIIAH was tagged with the biotin acceptor peptide (BAP) at its C-terminus, and the *E. coli* biotin ligase BirA was expressed in the forespore under the control of the SpoIIQ promoter, *pspoIIQ.* BirA produced in the forespore was able to biotinylate BAP-tagged SpoIIIAH in the mother cell suggesting that the Cterminal domain of SpoIIIAH is accessible to the forespore cell cytoplasm (Meisner *et al*.,2008). This finding, when combined with the similarity of the SpoIIIAH Cterminal domain to YscJ/FliF-like proteins that form multimeric rings in type III secretion systems (T3SS) and flagella led to the hypothesis that SpoIIQ and SpoIIIAH form a channel complex that is gated (closed) on the mother cell side and open on the forespore side (**Figure 1-16**) (Meisner *et al*., 2008).

As SpoIIIAH is similar to the ring-forming proteins in T3SS and flagella, researchers have speculated that SpoIIIAH forms a circular complex with SpoIIQ. Using computational modelling, several ring-shaped models with higher order symmetries have been proposed, including a 12, 15, and 18

subunit model (Zeytuni *et al*.,2019) (Rodrigues *et al*., 2016). In the 12-subunit model, stacked rings of SpoIIQ and SpoIIIAH form a channel with an inner diameter of 58Å at its narrowest point, wide enough to allow the passage of small molecules and larger proteins (Levdikov *et al*.,2012). The inner diameter of the SpoIIIAH ring in the 15-subunit model is 82Å, which is similar to that of basal type III secretion systems (Meisner *et al*., 2012). The 18-subunit model has a larger inner diameter of 116Å (Meisner *et al*.,2012). In 2017, cryo-electron microscopy was used to determine the structure of the SpoIIIAG ring to a 3.5Å resolution, with researchers suggesting that the 30-mer SpoIIIAG ring structure best correlates with the 15-subunit model of the SpoIIQ – SpoIIIAH channel (Zeytuni *et al*., 2017). The SpoIIQ – SpoIIIAH ring models cause repulsion with the anionic bacterial membrane as they have outer surfaces that are largely electronegative, meaning additional SpoIIIAA proteins are likely to be required to form a continuous conduit between the mother cell and forespore membrane (Zeytuni *et al*.,2019).

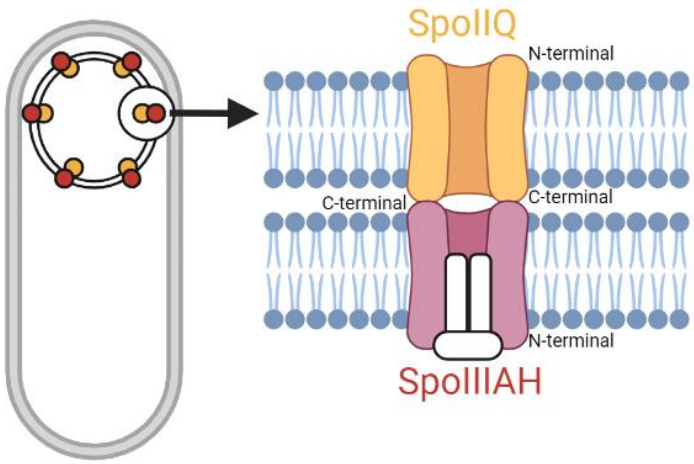


Figure 1-16: Schematic representation of the proposed SpoIIQ-SpoIIIAH complex showing that the complex is open on the forespore end but gated (closed) on the mother cell end (Adapted from Meisner *et al*.,2012)

1.10.3 Cellular Engulfment: Membrane Fission

Once the leading edge of the engulfing membrane reaches the forespore tip, lipids at the leading edge rearrange and the mother cell membrane undergoes membrane fission, releasing the forespore into the mother cell cytoplasm (Khanna *et al*.,2020). The protein FisB is synthesised in the mother cell under σ^E control and is important for membrane fission (Khanna *et al*.,2020). FisB recruits cardiolipin to the curved membrane at the site of membrane fission which destabilises the membrane, leading to lipid reorganisation and the separation of the mother cell and forespore membranes (Khanna *et al*.,2020).

1.11 Spore Cortex and Coat Assembly

After the completion of engulfment, the forespore develops into a mature spore inside the mother cell cytoplasm. The spore cortex and coat are then synthesised in the mother cell under the control of σ^E and σ^K (Setlow, 2011) (Driks,1999). The spore cortex is made from peptidoglycan precursors and resides between the inner and outer membrane (**Figure 1-17**) (McKenney *et al*.,2013). The spore coat is a multi-layered structure made from 70 individual proteins and is the outermost layer of the spore (**Figure 1-17**) (McKenney *et al*.,2013) (Takamastu and Watbe,2002). The *B. subtilis* spore coat consists of three layers: an inner coat, an outer coat, and crust and all layers are visible in thin-section microscopy. There is variation in the individual coat proteins present in spore-forming bacteria with variation believed to be important for defining the ecological niches of different bacteria (McKenney *et al*.,2013).

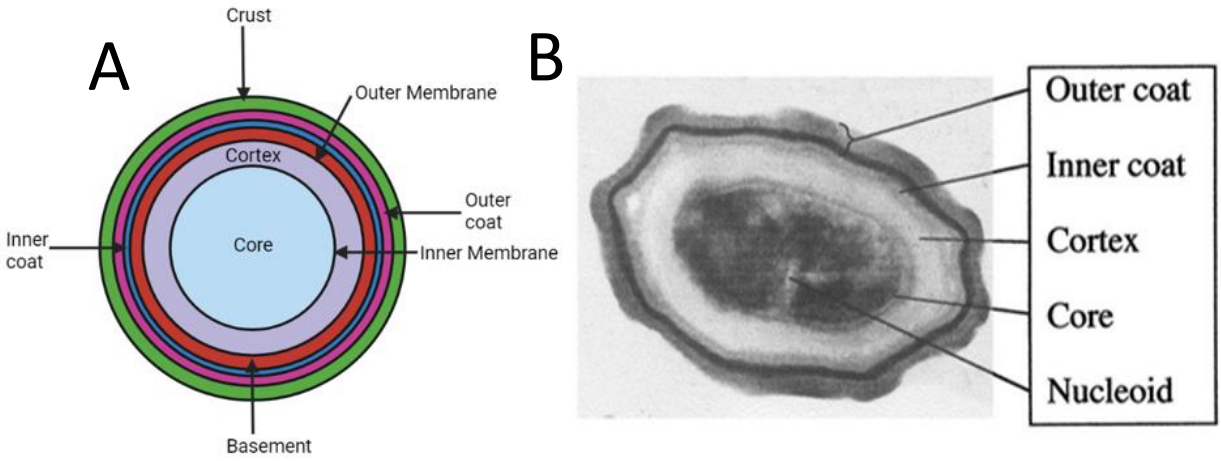


Figure 1-17: Spore ultrastructure. A) The multi-layered spore structure consists of the crust, outer coat, inner coat, cortex, and core. B) Electron micrograph of the *B. subtilis* spore. The DNA is contained within the nucleoid and the spore is surrounded by the peptidoglycan cortex, inner coat, and outer coat. (Taken from Nicholson, 2000).

Spore coat assembly occurs in five steps. The first two steps involve the SpoIVA protein which binds to the asymmetric division site on the mother cell side of the forespore membrane and covers the membrane (Takamatsu and Watabe,2002) (Henriques and Moran,2007). At the same time, mother cell proteins SpoVID, YrbA, and CotE localise at the SpoIVA layer, and these proteins comprise the basement layer on which the coat proteins are assembled. In step three, after the completion of engulfment and the activation of σ^K, CotA and other coat proteins are synthesised (Takamatsu and Watabe,2002) In step four, GerE-dependent coat proteins are synthesised and incorporated into the coat layers. In the final step, autolysins degrade the mother cell envelope, and mature spores are released from the mother cell into the environment (Takamatsu and Watabe,2002) (Henriques and Moran,2007).

Once released into the environment, spores are resistant to heat, desiccation, and UV radiation (Errington,2003). Spores are more resistant to UV radiation than vegetative cells due to the presence of CotA in the spore coat which confers resistance by generating a pigment similar to melanin (Driks,2002). The spore core is dehydrated as calcium dipicolinic acid (CaDPA) synthesised in the mother cell is taken up by the forespore. CaDPA binds to free water molecules and increases the heat resistance of spores. Small acid soluble proteins (SASPs) bind to DNA and protect it from damage (Driks,2002) (Henriques and Moran,2007). The *B. subtilis* spore can survive ingestion by bacterivores as the spore coat provides physical protection from predators (McKenney *et al*.,2013). *B. subtilis* spores can remain dormant for years and have been revived from dried plant samples from the 1640s (McKenney *et al*.,2013). They can also survive in a range of environments including hot springs, arctic sediments, and the mammalian GI tract.

1.12 Spore Germination

Germination occurs when spores quickly convert back to vegetative growth by responding to nutrients called germinants. Germinants are typically amino acids, purine nucleotides, and sugars, but spores can also germinate in response to high pressures, lysosomes, salts, dodecylamine, and CaDPA (Setlow,2003). Germinants must first penetrate the outer coat and cortex layers before they encounter germination receptors in the inner membrane (Xing *et al.,* 2020). Multiple germination receptors form clusters called germinosomes in the inner membrane which increases the concentration of germination receptors (Xing *et al*.,2020) (Setlow,2014). In *B. subtilis,* germination receptors are encoded by the tricistronic *gerA, gerB* and *gerK* operons expressed in the forespore, late in sporulation. Researchers believe that each germination receptor is made of three proteins encoded by each *ger* operon but the stoichiometry of proteins in the receptor and how the receptors interact with others in germinosomes is currently unknown (Setlow,2003) (Xing *et al*.,2020).

Spore germination takes place in multiple steps. Early in germination, the spore membrane increases in fluidity, and cations are released from the spore core into the supernatant. $Na^{+,}$ $K^{+,}$ and H^{+} ions are released first, and this is followed by the release of Ca^{2+}dipicolinic acid through channels made of seven SpoVA proteins in *B. subtilis* (Setlow, 2014) (Setlow,2003). The release of CaDPA completes the first stage of germination and triggers entry into stage two. In stage two, cortex lytic enzymes (CLEs) degrade the peptidoglycan cortex allowing the core to expand by taking up water (Moir,2006). The increased water content in the core allows metabolism to begin and this is followed by macromolecular synthesis that converts the germinated spore into a growing cell (Xing *et al*., 2020).

As described above, sporulation is a multi-step process that takes roughly 8 hours to complete in *Bacillus subtilis*. In this thesis, *B. subtilis* is used as a model organism to study the forespore protein SpoIIQ (Q) and mother cell protein SpoIIIAH (AH) which both show a dynamic localisation pattern during engulfment and are believed to form a channel complex through which the developing forespore is nurtured. Super-resolution microscopy and fluorescent protein labelling are used to determine the number of Q and AH molecules in single cells and the stoichiometry of the Q-AH channel complex. The following section of the introduction introduces microscopy and discusses fluorescence and fluorescent proteins. The advantages and disadvantages of fluorescent proteins are outlined and their use in single-molecule fluorescence microscopy is discussed. The last section of the introduction describes the need for super resolution microscopy and outlines the main principles of SIM and Slimfield microscopy.

2. Physics Introduction

2.1 Introduction to Microscopy

Microscopy has revolutionized the way researchers view nature and life, as it allows scientists to visualize objects that would be otherwise invisible. In the late 16th to 17th century, Hans Jansen and his son Zacharias developed the first functional compound microscope which allowed objects to be magnified up to 9 times (Vangindertael *et al*.,2016). In 1965, Robert Hooke published the book 'Micrographia' in which he described and illustrated fossils, insects, and plants that he had visualized using a hand-designed compound microscope, illustrating how light microscopy could be used in Biology (McNamara *et al*.,2017). Inspired by Hooke, Anton Van Leewenhoek designed a microscope with a single, small spherical lens which he used to observe specimens at 280x magnification (Gest,2004). He described his observations in letters to the Royal Society, giving him the unofficial title of 'father of microbiology' for the discovery of yeast and bacterial cells (Vangindertael *et al*.,2016) (Gest,2004).

Since these early discoveries, microscopy has evolved considerably as modern-day super-resolution microscopes are highly sophisticated and have many mechanical and optical components. Images are no longer visualized by the eye but are recorded by detectors and projected onto computer screens (Vangindertael *et al*.,2016). In Leeuwenhoek's microscope, samples were mounted onto needles, but this is no longer required as samples are treated with chemicals and resins before being placed onto electronically stabilized stages (Vangindertael *et al*.,2016). The amount of detail that can be visualized using microscopy has also dramatically increased from the micrometer to the nanometer range (Vangindertael *et al*.,2016).

2.2 Fluorescence

Microscopy aims to increase the signal between what is interesting (signal) and what is not (background) (Lichtman and Conchello,2005). Fluorescence microscopy allows the visualisation of objects of interest against a black background and its high sensitivity and specificity make it one of the most powerful techniques used in biological research (Lichtman and Conchello,2005). George Stokes first described the term 'fluorescence' in 1852 and used it to explain the process by which the mineral fluorite emits photons in the visible spectrum when irradiated with UV light (Stokes e*t al.*,1852). In 1933, Aleksander Jablonski developed the Jablonski diagram to illustrate the physics of fluorescence. In the Jablonski diagram, fluorescence occurs when an electron in a fluorophore undergoes an excitation transition to a higher energy state upon absorption of a single photon of light (Shashkova and Leake,2017) (**Figure 2-1**)**.** This state is transient as the excited electron quickly loses energy through vibrational losses (Shashkova and Leake,2017). The electron then undergoes an energy transition back to the ground state which is accompanied by the emission of a photon with a longer wavelength and smaller associated energy (**Figure 2-1**) (Shashkova and Leake,2017).

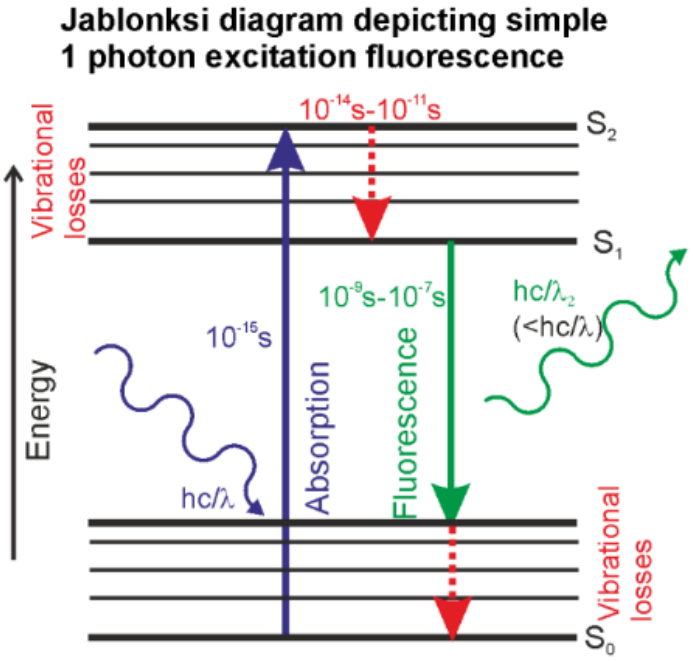


Figure 2-1: Jablonski Diagram to illustrate the physics of fluorescence. An electron of a fluorophore at the ground S_0 state moves to a higher energy state upon absorption of a single photon of light. When the excited electron relaxes back to the ground state due to vibrational losses, a photon of light with a longer wavelength and smaller associated energy is emitted (Taken from Shashkova and Leake,2017).

2.3 Fluorescent Proteins

Fluorophores are fluorescent molecules that respond to light. Protein fluorophores form a ϐ-barrel which forms before an auto-catalytic reaction that converts a tripeptide into a fluorophore

(Costantini and Snapp, 2013). The properties of a fluorophore can be modified by environmental changes with pH changes impacting the fluorescence of most fluorescent proteins. The pH range impacting a fluorophore's brightness is known as the fluorophore's pKa (Costantini and Snapp, 2013). The pKa is the pH at which the fluorophore's brightness decreases by one-half after which the fluorophore will continue to become darker at decreasing pH . A fluorophore's brightness is determined by its extinction coefficient and quantum yield and its relative brightness is calculated by multiplying these values (Toseland,2013). The extinction coefficient is a measure of how easily a fluorophore absorbs light and enters the excited state with fluorophores with a higher extinction coefficient absorbing more light (Toseland,2013). The quantum yield is the ratio of the number of photons emitted to the number of photons absorbed. Molecules with a lower rate of deactivation compared to fluorescence have a higher quantum yield meaning fluorophores with both a high quantum yield and extinction coefficient are the brightest (Toseland,2013). All fluorophores go through several cycles of excitation and emission before they are permanently deactivated or photobleached by reactive oxygen species (ROS) in the excited state, meaning fluorophores must be carefully chosen to ensure low bleaching rates (Toseland,2013).

2.4 Green fluorescent proteins (GFPs)

Osamu Shimomura isolated Green Fluorescent Protein (GFP) from the jellyfish *A. Victoria* in the 1960s. He found that the protein fluoresced when exposed to light in the blue-to-ultraviolet region of the electromagnetic spectrum in the absence of enzymes or co-factors (**Figure 2-2**) (Lippincoft-Schwartz and Patterson, 2003). Sometime later, Martin Chalfie cloned the GFP gene into bacteria and the nematode *C. elegans* and demonstrated that the GFP gene contained all the information needed for the synthesis of the fluorophore (Sample *et al*.,2009). Chalfie's findings suggested that GFP could be used as a fusion tag to track protein localization in live cells as the fluorophore retains its fluorescence when fused to proteins and does not affect the function of the fusion partner (Lippincoft-Schwartz and Patterson, 2003).

Site-directed mutagenesis studies have led to improvements in GFP characteristics giving rise to the new variant enhanced GFP (EGFP). EGFP has the mutation S65T ($Ser^{65} \rightarrow Thr^{65}$), which gives rise to a bright, stable protein by accelerating the speed of fluorophore formation (Lippincoft-Schwartz and Patterson, 2003). The mutation $Phe^{64} \rightarrow Leu^{64}$ allows the EGFP molecule to fold correctly at 37^{o}C and increases its temperature sensitivity (Lippincoft-Schwartz and Patterson, 2003) (Sample *et al.,*2009). In EGFP, the major and minor absorbance peaks of wtGFP are converted into a single absorbance peak at 489 nm, producing brighter molecules that retain wild-type fluorescence (**Figure 2-2**). EGFP

has been used in many experimental systems including whole-body imaging with intact animals, cytoskeletal imaging, and imaging of intracellular protein pathways (Hoffman,2005).

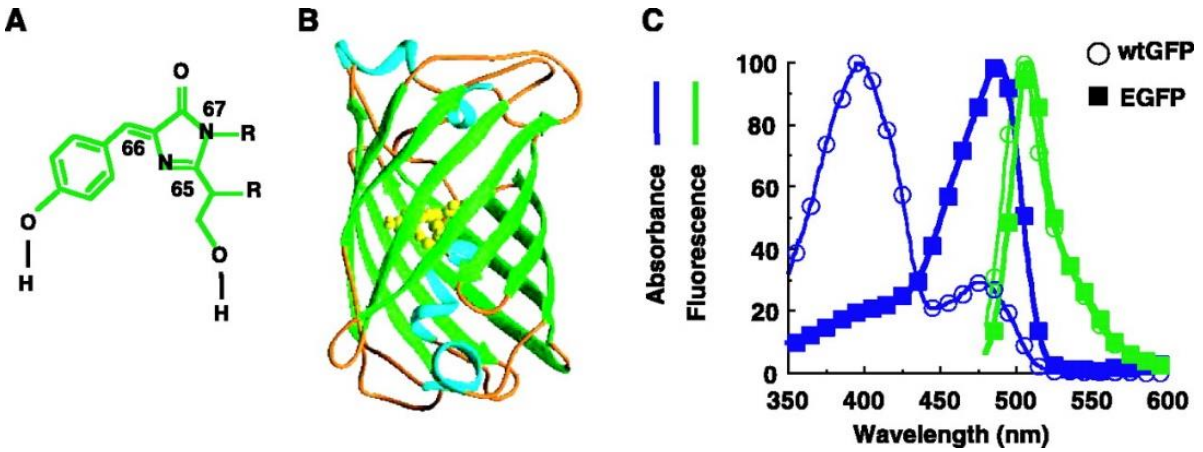


Figure 2-2: Structure of the wtGFP chromophore. A) The wtGFP chromophore consists of the three amino acids Ser65, Tyr66, and Gly67 which form a cyclic tripeptide. B) The chromophore is located inside the 11-strand β-barrel of GFP. C) wtGFP major and minor absorbance peaks (blue circles) and single fluorescence emission peak (green circles) compared with EGFPs single absorbance peak (blue squares) and single fluorescence emission peak (green squares) (Taken from Lippincoft-Schwartz and Patterson., 2003).

2.5 Green Fluorescent Protein (GFP) Variants

- GFP variants such as Cyan fluorescent protein (CFP) and Blue fluorescent protein (BFP) were discovered by screening *A. Victoria* GFP mutants (Lippincoft-Schwartz and Patterson, 2003)**.** BFP was one of the first EGFP variants used in multi-colour imaging and has an absorbance peak at 384 nm and an emission peak at 448 nm (**Figure 2-3**). BFP is dim and has a low quantum yield and photo bleaches easily. Its limited brightness is problematic in cellular imaging experiments as the UV light used to excite it also excites cellular biomolecules such as flavins contributing to high cellular autofluorescence (Sample *et al.*,2009). CFP has a spectrum between BFP and EGFP due to a substitution of Tyr^{66} to Trp^{66} and is brighter and more photostable than BFP (Sample *et al.,*2009). YFP has shifted emission and excitation spectra with respect to EGFP and other green fluorescent variants (**Figure 2-3**). YFP is brighter than EGFP but more sensitive to low pH and high halide concentrations. Monomeric (m)RFS have been developed and are useful experimentally in dual-labelling applications with GFP derived proteins. mCherry is brighter, more photostable and matures more rapidly than its predecessor mRFP1. mRFP1 is a monomer derived from the tetramer DsRed (Fink *et al.*, 2010).

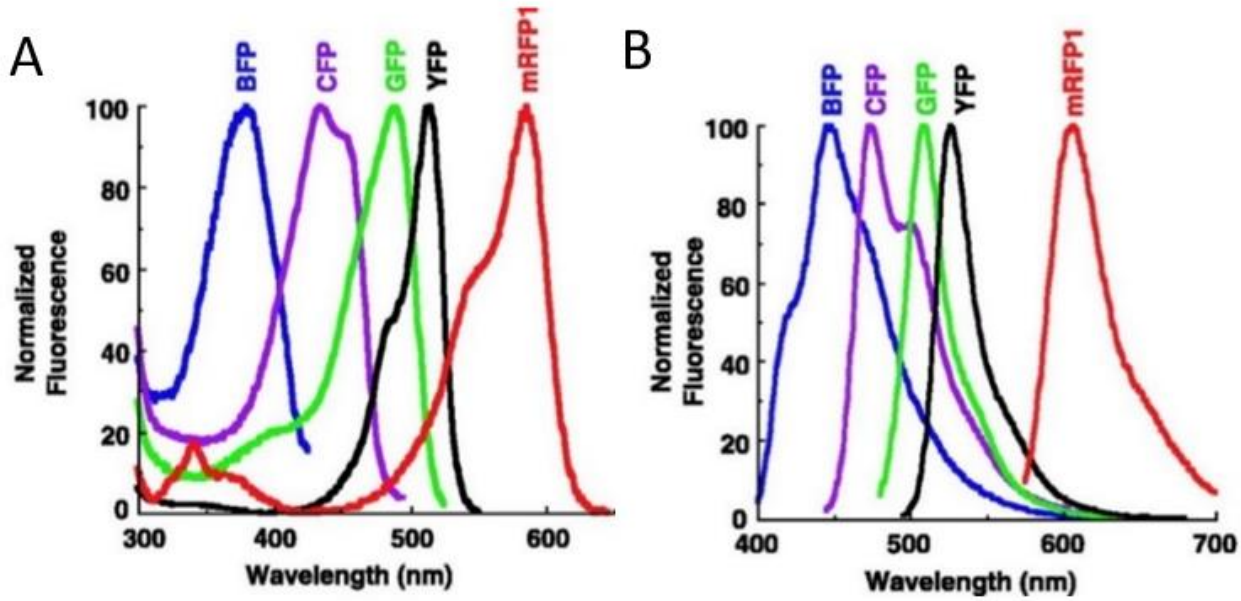


Figure 2-3: The excitation (A) and emission (B) spectra of blue (BFP), cyan (CFP), green (GFP), yellow (YFP), and red (mRFP1) fluorescent proteins. (Taken from Lippincoft-Schwartz and Patterson, 2003).

2.6 Advantages and Disadvantages of Fluorescent Proteins

Fluorescent proteins are advantageous when compared to other methods of protein labelling. Firstly, fluorescent proteins generate their chromophores autocatalytically which leaves no free dye in the solution and makes them simpler to use than organic fluorophores (Sample *et al*.,2009). When compared to the luciferase reporter, GFP is more suitable for whole-cell imaging as it has a stronger signal due to its high quantum yield (Toseland,2013). The protein sequence of GFP can be highly expressed in mammalian cells and the proteins' structure is not significantly affected by the external environment as the chromophore is protected within the 3D structure of the protein (Hoffman, 2005) (**Figure 2-2**). GFPs' long excitation wavelength (490 nm) allows long-term measurements to be made as the protein does not quench easily. Fluorescent protein labelling is commonly used for *in-vivo* imaging as fluorescent proteins can be coupled to proteins of interest with 100% specificity and are bright enough to be distinguished from background autofluorescence (Toseland,2013) (Xue *et al*.,2011).

Despite having many advantages, fluorescent proteins are not without their disadvantages. Many researchers are conflicted as to whether GFP is toxic to cells as some believe that GFP aggregation causes cellular toxicity and others believe that exciting GFP for long periods of time may generate free radicals, but the consensus belief is that most GFP applications do not result in toxicity (Jensen,2012). Attachment of fluorescent proteins to proteins of interest can affect protein and cellular function, particularly if fluorescent proteins oligomerize, motivating researchers to express the smallest amount of fusion protein required for imaging (Jensen,2012). Fusion proteins may also

break down in the cellular environment generating background signals and poor/diffuse localization (Jensen,2012).

2.7 Why detect single molecules?

The single molecule is the fundamental minimal unit in most biological processes (Lenn and Leake,2012). Single-molecule methods allow the investigation of molecular heterogeneity whereas traditional biochemical methods rely on the average properties of large proportions of molecules. An example of this is a sample of 1 µl of water which contains approximately 10^9 molecules. Ensemble methods would assume that the average value from the 10^9 molecules represents the entire population of molecules. Biomolecules, however, exist in many conformational states, and so the mean conformation from ensemble methods provides information on the most stable of the many conformations but not on the other short-lived but essential states (Lenn and Leake,2012) (Leake,2013). Bulk ensemble methods cannot probe different states in an experimental system which is problematic as molecules within a large population are often doing different things at different times (Lenn and Leake,2012) (Leake,2013).

Single-molecule approaches are used in experimental systems to study the biochemical and physical properties of individual molecules one by one (Lenn and Leake,2012) (Leake,2013). By studying single molecules, biological processes can be studied on a smaller spatial, temporal, and numerical scale than ensemble experiments, but many observations are needed to construct the underlying distribution of molecular properties (Lenn and Leake,2012) (Leake,2013). Single-molecule approaches are used in biology as biological processes often occur at the level of one or few molecules with very few molecules in each cell (Leake,2010). Single-molecule studies also allow the identification of unpredictable behaviour masked by the ensemble averaging process which is useful in pathology as the number of disease-causing molecules is often low (Harriman and Leake,2011).

Single-molecule approaches have many advantages when compared to ensemble methods. Single-molecule studies allow the investigation of multiple folding states and rare states that would be averaged in ensemble experiments, allowing the determination of the connectivity and kinetic rate constants between different states (Deniz *et al*.,2007). The high level of sensitivity in single-molecule experiments is advantageous as some biomolecules aggregate at ensemble concentrations (Deniz *et al*.,2007).

2.8 Super Resolution Microscopy

Light microscopes are often used by scientists to visualise whole cells and electron microscopes are used to visualise subcellular structures. Scientists are unable to see subcellular organelles with light microscopy as light microscopes have a poorer resolution than electron microscopes (Schermelleh *et*

al.,2019). Resolution is defined as the smallest distance between two objects that allows the objects to be observed as separate entities. The resolution of conventional light microscopy techniques such as widefield and confocal microscopy is 200-250 nm, which is half the smallest wavelength of visible light (400 nm) (Schermelleh *et al*.,2019). This means that any points separated by a distance less than 200 nm cannot be resolved using a light microscope as the two points will be indistinguishable.

When an image is taken of a single point source of light (e.g., a fluorophore) using an imaging system such a light microscope, the single point source of light appears blurry due to the diffraction limit of light. The blur in the image is called the point spread function (PSF) (**Figure 2-4**). In the XY cross section of the PSF, the PSF forms an Airy pattern consisting of concentric rings (Galbraith and Galbraith, 2011). In the XZ profile, the PSF is much taller than it is wider in the XY profile, so the Z resolution of most microscopes is worse than the XY resolution. Every dye molecule in the sample is substituted in the image with a PSF and the size of the PSF influences the resolving power. The Abbes diffraction formula for lateral (i.e., XY) resolution is d= λ/2NA. λ is the wavelength of light and typically the shortest wavelength of light used for imaging in fluorescence microscopy is between 400-500 nm (Galbraith and Galbraith, 2011). NA is the numerical aperture and is typically around 1 meaning the equation can be simplified to d (lateral [XY] resolution) = λ/2, giving a value of 200-250 nm. To get beyond this resolution limit, super resolution techniques need to be used. In 2014, the Nobel Prize in Chemistry was awarded to Eric Betzig, Stefan W. Hell, and William E Moerner for the development of super resolution fluorescence microscopy (Galbraith and Galbraith, 2011).

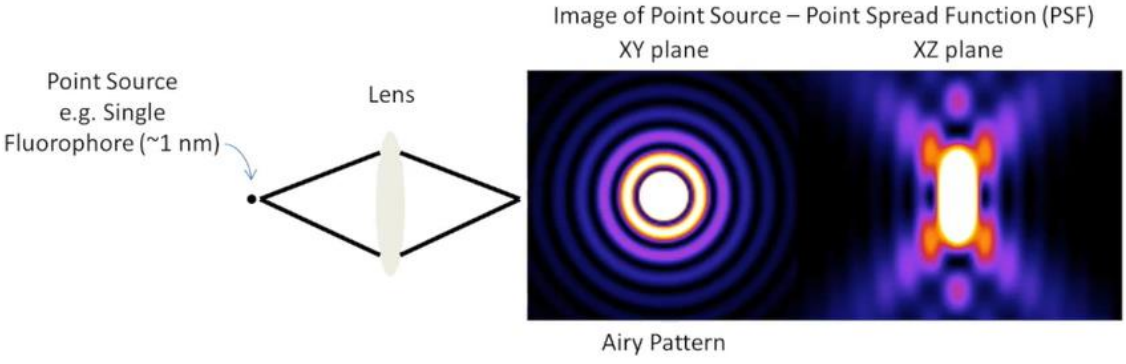


Figure 2-4: Diagram showing the response of an optical imaging system to a point source (i.e., a single fluorophore). The point spread function (PSF) in the XY and XZ planes is shown. (Adapted from Teledyne Photometrics).

Diffraction limited spots take the shape of an Airy Disk that consists of a bright central spot with concentric rings (Figure 2-4). Using statistics, a fluorophore displaying an Airy pattern can be localised precisely by fitting the central spot to a Gaussian function. The Gaussian is fitted to the centre of the disk as the central spot is the brightest and most localised to the true location of the fluorophore (Teledyne Photometrics, 2022). Gaussian fitting results in a 3D representation of the Airy disk and the higher and thinner the central peak, the more accurately the fluorophore can be localised (Figure 2-5). If two Airy disks are very close together, they merge into one, making it impossible to tell close objects apart. These merged disks are counted as one data point and information is lost, even with statistical fitting (Teledyne Photometrics, 2022). Super resolution localisation microscopy overcomes the problem of overlapping fluorophores by using fluorescent molecules that switch on and off at random. By taking thousands of images, enough data is captured where all the molecules are "on" at least once. Molecules close together in space, are unlikely to both be "on" at the same time in the same image frame and can therefore be separated by time. All super-resolution localization microscopy techniques create a super-resolution image by individually localizing single fluorophores while separating them temporally. The way they differ is in the methods they use to switch the fluorophores on and off (Teledyne Photometrics, 2022).

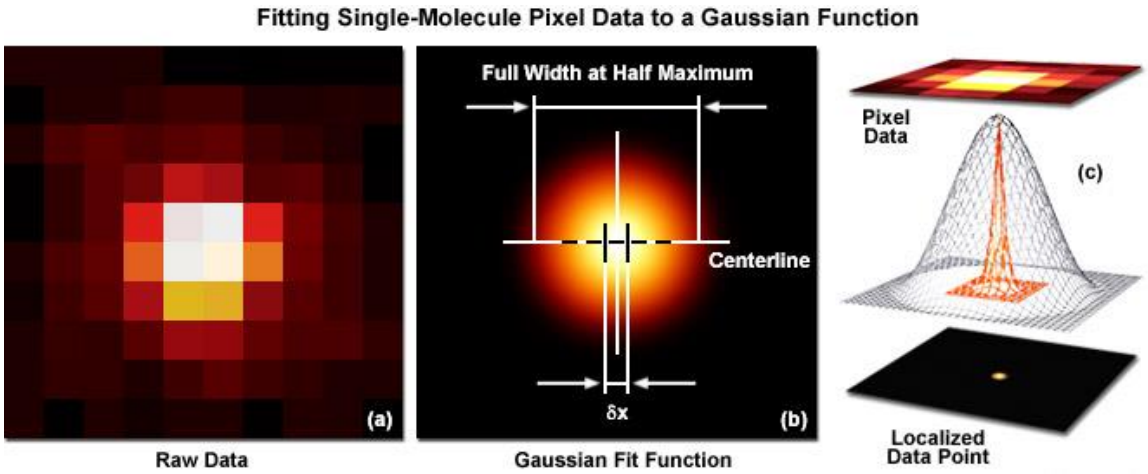


Figure 2-5: Fitting Single-Molecule Pixel Data to a Gaussian Function. A) The raw data appears as a pixelated Airy disk. B) By applying a Gaussian Function, the pixels can be fit to the intensity of light, indicating the most intense spot of light. C) The data can be used to localise the central data point where the fluorophore is located (Taken from Teledyne Photometrics, 2022).

2.8.1 SIM Microscopy

SIM microscopy is a super resolution technique that allows optical sectioning and improvements in both lateral and axial resolution. In SIM microscopy, the sample is illuminated with patterned light and not uniform light (Galbraith and Galbraith, 2011). When the patterned illumination light (**Figure 2-6**) interacts with the high spatial frequency structures in the sample that are too fine to be

ordinarily resolved, the two produce an interference pattern called a Moire fringe (**Figure 2-6**). The moire fringe is coarser than either pattern alone and is detected by the microscope (Galbraith and Galbraith, 2011). The patterned illumination light is translated and rotated to give a series of images with different moire fringes which are then reconstructed to give the final super resolution image. As the patterned illumination is known, it can be mathematically removed from the Moire to gain access to the normally irresolvable higher resolution information in the sample. SIM increases resolution to 100 nm in the x-y (lateral) direction and 250 nm axially (Galbraith and Galbraith, 2011).

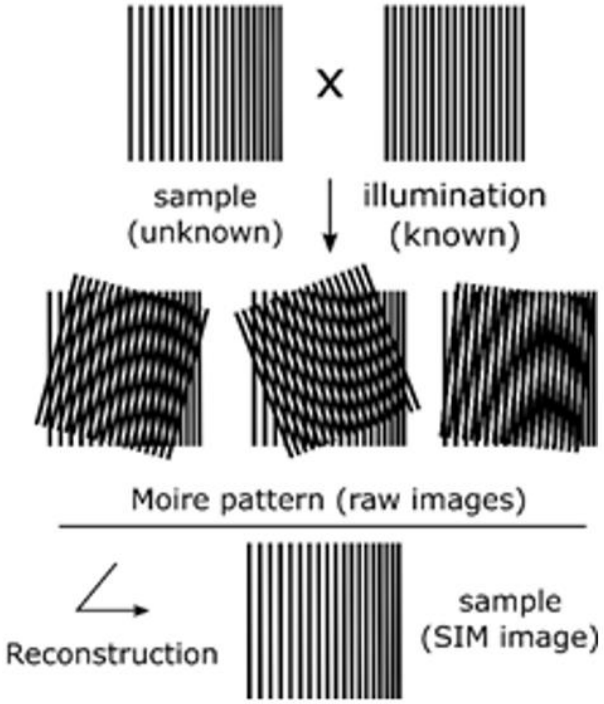


Figure 2-6: Diagram illustrating the Moire effect and the principles of SIM Microscopy (Adapted from Cherry Biotech)

2.8.2 Slimfield Microscopy

Slimfield microscopy is another type of super resolution microscopy used to investigate biological processes. With Slimfield microscopy, the normal fluorescence excitation field is reduced to encompass single bacterial cells which produces an excitation field of 100-1000x the excitation intensity of standard epifluorescence microscopy (Wollman *et al*.,2015). This allows imaging at much greater signal intensity relative to normal camera imaging noise and facilitates the millisecond time scale imaging of single fluorescently labelled proteins (Wollman *et al*.,2015). Slimfield microscopy has been used to study components of the *E. coli* replisome when fused to the fluorescent protein YPet. In this study, 2 distinct spots were observed in 75% of cells. A single spot was observed in cells that had not moved far enough from the origin of replication (**Figure 2-7**) (Wollman *et al*.,2015).

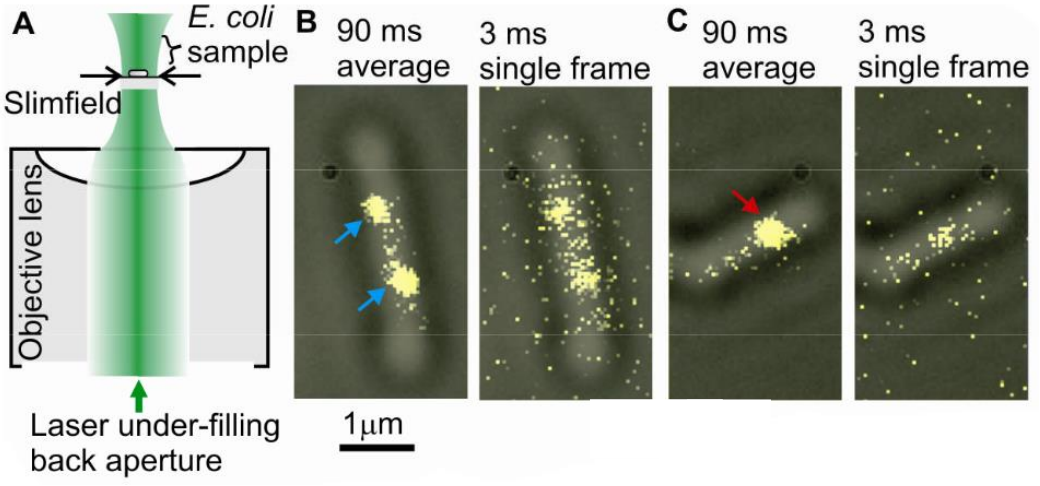


Figure 2-7: A) Schematic diagram showing Slimfield illumination. Images of bacterial cells in which DnaQ, a component of DNA polymerase is fused to YPet, having either (B) 2 spots or (C) 1 spot per cell. Taken from (Wollman *et al*.,2015).

Slimfield microscopy allows dual colour imaging and was used to study the bacterial condensin complex MukBEF in *E. coli.* Pairs of GFP and mCherry labelled MukB, MukE, and MukF proteins were imaged and yielded a stoichiometry of 4:2:2 (Wollman *et al*.,2015). Slimfield microscopy has also been used to investigate the role of SpoIIIE in cellular development in live *B. subtilis* cells (Wollman *et al*.,2020). SpoIIIE was found to form tetramers capable of cell and stage-dependent clustering, with its copy number increasing to 700 as sporulation progresses (Wollman *et al*.,2020). The Slimfield microscope used in this thesis allows the detection of fluorescently labelled proteins to a 40 nm localisation precision as Slimfeld produces a relatively high local laser excitation field that allows for smaller sampling times for single fluorescent molecules. The 40 nm spatial precision is obtained in the same way as any other localization microscopy-based approach - a Gaussian function is fitted to the intensity profile of each fluorescent spot which allows you to pinpoint the centre of its intensity peak to roughly an order of magnitude better than the "width" of the spot itself.

2.9 Aims

The aim of this thesis is to investigate the role of SpoIIQ and SpoIIIAH throughout engulfment, with a particular focus on the SpoIIQ-SpoIIIAH channel. As mentioned above, Levdikov *et al* and Meisner *et al* used protein crystallography and computational modelling methods to generate ring shaped models of the SpoIIQ-SpoIIIAH channel. The two research groups generated slightly different results with Levdikov and colleagues proposing a 12-mer model and Meisner *et al* proposing either a 15 or 18 subunit model. Despite these inconsistencies in findings, many researchers believe the channel is composed of 12 to 24 subunits. The stoichiometry of subunits in the SpoIIQ-SpoIIIAH channel is currently unknown and no *in-vivo* experiments have been performed to address this problem in living cells. Slimfield microscopy has been used to determine the stoichiometry of other protein complexes (Wollman *et al.,* 2015) and is used in this thesis to study the SpoIIQ-SpoIIIAH channel.

This thesis aims to determine the stoichiometry of SpoIIQ and SpoIIIAH in live cells in both wildtype and mutant backgrounds. The thesis also aims to determine the mobility of SpoIIQ and SpoIIIAH in live cells in wildtype and mutant backgrounds. The number of molecules of SpoIIQ and SpoIIIAH in wildtype and mutant backgrounds is calculated in this thesis. Lastly, the thesis aims to examine the SpoIIQ – SpoIIIAH interface and the residues contributing to the interface.

3. Methods

3.1 *mGFP-spoIIQ* Fusion Construction

To investigate the stoichiometry of the SpoIIQ – SpoIIIAH channel, a plasmid containing an *mGFP-spoIIQ* fusion was constructed and integrated into the *B. subtilis* chromosome at the *amyE* locus. The *amyE* locus is a genetic locus commonly used for integration into *B. subtilis.* The *B. subtilis* strain IB714 contains an erythromycin cassette disrupting the *amyE* locus and this strain was used to integrate the mGFP-*spoIIQ* fusion into the *B. subtilis* chromosome via a double cross over event. As illustrated in **Figure 3.1**, integration of the *mGFP-spoIIQ* fusion results in a switch in antibiotic resistance from erythromycin to spectinomycin as the erythromycin resistance from the chromosome is lost and spectinomycin resistance from the plasmid containing the *mGFP-spoIIQ* is retained.

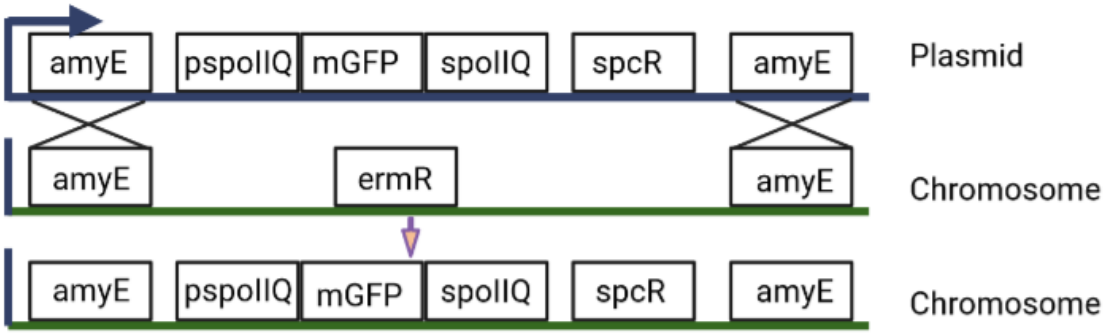


Figure 3-1: Diagram showing how the *mGFP-spoIIQ* fusion integrates into the *B. subtilis* chromosome at the *amyE* locus via a double cross over event. There is a switch in antibiotic resistance from erythromycin to spectinomycin.

mGFP was fused to the N terminus of SpoIIQ, as the C-terminal domain of SpoIIQ interacts with partner protein, SpoIIIAH (Lewis and Marston, 1998). The sequence of the native *spoIIQ* promoter, *pspoIIQ* was cloned immediately upstream of the mGFP sequence, allowing the promoter to drive transcription of the fusion protein. The plasmid containing the *mGFP-spoIIQ* gene fusion was designed to contain a spectinomycin resistance cassette (spcR). Prior to the start of this work, mGFP and mCherry SpoIIIAH fusion proteins were prepared by Adam Hughes. SpoIIIAH fusions were designed for integration at the native *spoIIIAH* locus as *spoIIIAH* is the final gene in the *spoIIIAA-AH* operon and it was impractical to clone the entire *spoIIIAA-AH* operon at the *amyE* locus. Since *spoIIIAH* is the final gene in the *spoIIIAA-AH* operon, integration of the *spoIIIAH* fusion at the native *spoIIIAH* locus required a section of *spoIIIAG* immediately upstream of the spoIIIAH fusion. SpoIIIAH fusions were designed to contain a chloramphenicol (camR) resistance cassette. Fluorescent proteins mGFP and mCherry were used to construct the plamids encoding for the SpoIIQ and SpoIIIAH fusion proteins. mGFP is a monomeric green fluorescent protein and mCherry is a monomeric red

fluorescent protein. The excitation and emission wavelengths of the two fluorescent proteins are shown in **Table 3-1**.

Fluorescent Protein	Maximum excitation wavelength (nm)	Maximum emission wavelength (nm)
mGFP	488	507
mCherry	587	610

Table 3-1: Excitation and emission profiles of mGFP and mCherry fluorescent proteins

pSG1154 (**Figure 3-2**) is the vector into which the *mGFP-spoIIQ* gene fusion construct was integrated. pSG1154 was obtainted from the *Bacillus subtilis* Stock Centre and contains an *amyE* open reading frame split in two by a spectinomycin resistance cassette and a sequence encoding a GFP mutant. The strain harbouring the pSG1154 plasmid was first inoculated in 50 ml LB with 50 µl ampicillin and then incubated overnight at 37°C. The next day, the pSG1154 plasmid was isolated using the Qiagen Mini Prep Kit as per the manufacturer's specifications. 1.5 µg of pSG1154 was digested with HindIII and SpeI, allowing excision of the sequence encoding the GFP mutant.

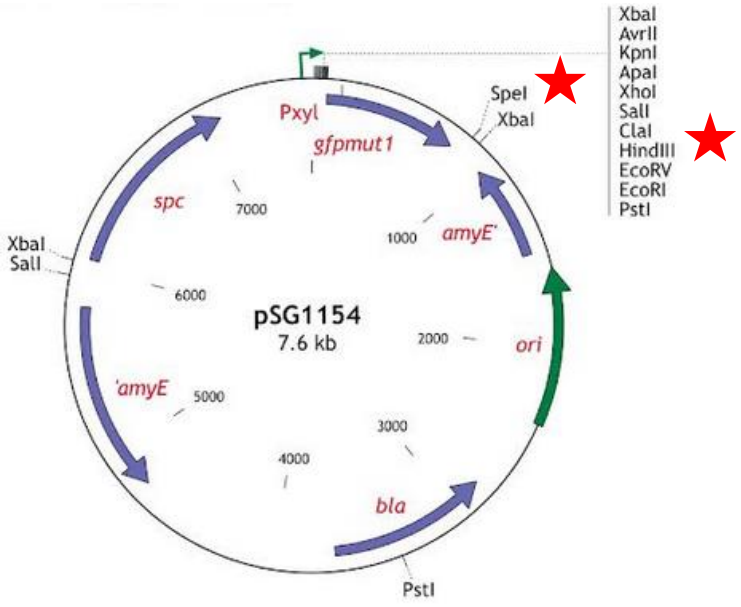


Figure 3-2: Plasmid map of pSG1154. In the pSG1154 plasmid, the *amyE* open reading frame is split in two by a spectinomycin resistance cassette and a sequence encoding a GFP mutant. Spc is the spectinomycin resistance cassette, ori the origin of replication and bla the ampicillin resistance cassette. HindIII and SpeI restriction sites are highlighted with red stars.

The pspoIIQ, mGFP and spoIIQ sequences required for construction of the plasmid containing the *mGFP-spoIIQ* fusion were amplified in 50 µl PCR reactions (primers 20 pmoles/µl, 2.5 mM dNTPS and 50 ng/µl template DNA) using 2 units of Pfu polymerase. The thermocycler profile consisted of 29 cycles of 94°C for 2 minutes, 53°C for 30 seconds and 72°C for 30 seconds followed by a 5-minute

final extension at 72°C. PCR products were analysed using agarose gel electrophoresis and purified using the Qiagen PCR purification kit as per the manufacturer's specifications. The mGFP sequence was amplified from plasmid pSG1729. The *pspoIIQ* sequence and the spoIIQ coding sequence were amplified from *B. subtilis* chromosomal DNA. Primers required for the amplification of each component were designed using serial cloner and each primer was designed to have a 3' sequence complementary to the adjacent fragment, allowing the three fragments to be assembled into a linear array using Gibson Assembly. The DNA fragment encoding *pspoIIQ-mGFP-spoIIQ* genes was then further amplified by PCR and integrated into the pSG1154 plasmid using Gibson Assembly (Gibson *et al*;2009).

3.2 Gibson Assembly

Gibson Assembly joins multiple overlapping DNA fragments in a single isothermal reaction. The Gibson assembly master mix includes three different enzymatic activities in a single compatible buffer. The exonuclease creates single stranded 3' overhangs that facilitate the annealing of fragments that are complementary at one end (the overlap region). The DNA polymerase fills in the gaps between the annealed fragments and the DNA ligase seals the resulting nicks in the backbone. A 20 µl Gibson assembly reaction was set up with the following four fragments: pspoIIQ (358 ng/µl), SpoIIQ (319 ng/µl), mGFP (316 ng/µl) and pSG1154 (116 ng/µl) each 1 µl, 10 µl Gibson assembly master mix and 6 µl deionized H_20. This sample, along with the NEB positive control was incubated in a thermocycler at 50°C for 60 minutes. Following incubation, NEB 5 alpha competent cells were mixed with 2 µl of the assembly reaction. 100 µl of cells were spread onto LB plates with ampicillin (100 µg/ml) and incubated overnight at 37°C.

3.3 Colony PCR

Colony PCR was used to determine whether the resulting antibiotic-resistant colonies contained the desired. 15 colonies were selected from the plates and each colony was resuspended in 20 µl of water. The tubes were warmed for 5 minutes at 95°C and 2.5 mM dNTPs, 20 pmoles/µl primers (AH59 and AH72, **Appendix 1**) and Pfu polymerase were added to the reaction. PCR was performed as described above and the PCR products were analysed using agarose gel electrophoresis. Having identified which colonies contained the desired 2kb insert, recombinant plasmids were isolated using the Qiagen Mini-Prep kit as per the manufacturer's specifications. To confirm the presence of the insert, isolated plasmids were digested with restriction enzymes NotI and NdeI. Restriction enzyme digestion reactions were incubated at 37°C for 3 hours and fragments were separated by agarose gel electrophoresis.

3.4 Agarose Gel Electrophoresis

0.5 g of agarose was weighed and placed in a 250 ml conical flask. 50 ml 0.5X TBE was added to the agarose and the solution was swirled to mix. The solution was heated in a microwave for 1 minute to dissolve the agarose. The solution was then left on a bench for 5 minutes to cool before 2 µl of SYBR safe dye was added. The gel was then slowly poured into a gel tank with any bubbles pushed to the side using a disposable tip. The comb was then inserted at the correct position and the gel was left to set for 30 minutes. After 30 minutes, 0.5X TBE running buffer was poured into the gel tank to submerge the gel to 2-5 mm depth. Samples were then loaded into the wells and the gel was ran at 110V for 1hr before being visualised under UV light.

3.5 Preparation of *B. subtilis* competent cells

A single colony of the required *B. subtilis* strain was resuspended in 100 µl LB and then spread onto a LB plate with the appropriate antibiotic before being incubated at 30°C overnight. The next day, cells were scraped from the plate and used to inoculate 40 ml of freshly prepared SpC medium and grown at 37°C to an optical density (OD_{600}) of 0.4. Growth was maintained at 37°C for 2 hours until cells became naturally competent. After this, 1 ml of culture was diluted into freshly prepared and pre-warmed SpII medium and allowed to grow at 37°C for 90 minutes. Cells were pelleted by centrifugation in a bench-top centrifuge at 3000rpm for 15 minutes at 24°C. After centrifugation, cells were resuspended in 8 ml of the supernatant and 2 ml 80% glycerol was added. The cells were then aliquoted and stored at -80°C.

3.6 Transformation of *B. subtilis*

An aliquot of competent cells was thawed at room temperature. 250 µl of cells were mixed with 250 µl SpII+EGTA solution and 20-200 ng of plasmid or chromosomal DNA in a microcentrifuge tube. Samples were incubated at 37°C for 30 minutes with shaking at 180RPM. The transformation mix was then suspended in 4.5 ml pre-warmed top agar supplemented with appropriate antibiotics and rapidly plated onto LB plates for overnight incubation at 37°C.

3.7 Confirming Transformation of *B. subtilis*

B. subtilis strain IB714 contains an erythromycin cassette at the *amyE* locus. Transformation into this strain can be checked by looking for antibiotic resistance switching following integration of the desired plasmid. In the case of a double cross over event, the antibiotic resistance cassette on the chromosomal locus is replaced by that on the desired plasmid (**Figure 3-1**). Colonies were picked and re-plated onto spectinomycin and erythromycin and incubated overnight at 37°C to check for the desired phenotype (**Figure 3-3**)

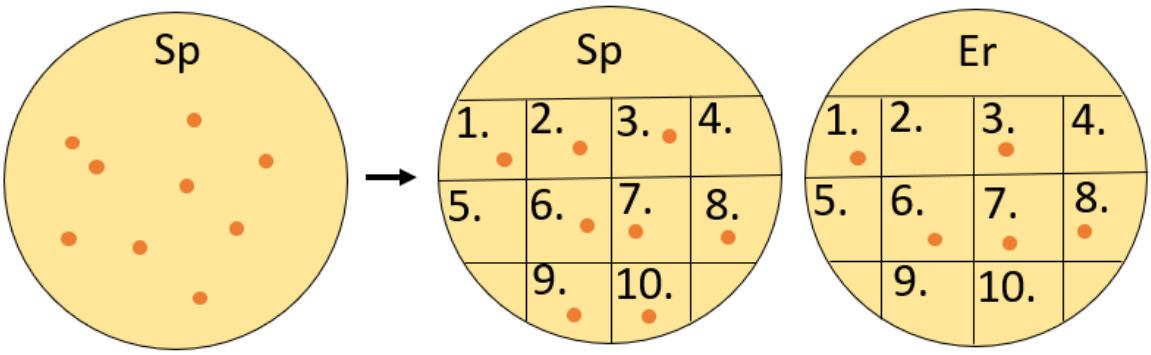


Figure 3-3: Schematic representation of the transformation selection method in *B. subtilis* in which the erythromycin resistance cassette at the chromosomal locus is replaced by a spectinomycin resistance cassette derived from the integrated plasmid. Colonies from a transformed *B. subtilis* strain containing a plasmid with a spectinomycin resistance cassette designed to replace an erythromycin resistance cassette on the chromosomal locus are plated onto LB agar supplemented with spectinomycin and erythromycin. Colonies that are spectinomycin resistant and erythromycin sensitive display the desired double cross over phenotype. Colonies resistant to both spectinomycin and erythromycin display a single cross over phenotype.

3.8 Fluorescence Imaging of *B. subtilis*: Sample Preparation

A single colony of the desired *Bacillus subtilis* strain was resuspended in 100 µl LB and then spread onto an LB plate with the appropriate antibiotic before being incubated at 30°C overnight. The next day, cells were scraped from the plate and used to inoculate 10 ml Difco Sporulation Medium (DSM) to an optical density (OD_{600}) of approximately 0.1. The culture was grown at 37°C for 8 hours until cells reached the engulfment stage of sporulation. The cells were then pelleted in a bench-top centrifuge set to 5000rpm for 3 minutes. The supernatant was removed, and cells were resuspended in 200 µl Spizizen minimal medium (SMN). The samples were then re-pelleted and resuspended in SMN again. Samples were then pelleted a third time and resuspended in 20 µl SMN. After this, samples were spotted onto a 1% agarose/SMN bed on a glass slide before being sealed with a coverslip.

3.9 Fluorescence Imaging of *B. subtilis*: Structured Illumination Microscopy (SIM)

The Zeiss Elyra 7 with Lattice SIM^2 is the SIM microscope that was used to image B. *subtilis* strains in this thesis. In Lattice SIM^2, the sample area is illuminated with a lattice spot pattern instead of the rotating grid lines used in conventional SIM (Chapter 2, **Figure 2-5)** which increases the imaging speed and provides a higher contrast to allow for more robust image reconstruction (Zeiss). The sampling efficiency of lattice pattern illumination is 2X higher than classical SIM meaning less laser dosage is needed for sample illumination. The strongly improved photon efficiency of lattice illumination allows for increased imaging speed, higher contrast, and lower photon dosage. The SIM^2 algorithm was used for image reconstruction as it reconstructs all structured illumination acquisition

data from the Elyra 7 with minimal artefacts for living samples. SIM^2 is a two-step image reconstruction algorithm and is superior to conventional one-step image reconstruction methods in terms of resolution, sectioning, and robustness (Zeiss).

3.10 Fluorescence Imaging of *B. subtilis*: Slimfield Microscopy

Single molecule Slimfield imaging allows the detection of fluorescently labelled proteins with millisecond sampling to within 40 nm precision, enabling the real time quantification of the stoichiometry and mobility of tracked molecular complexes (Plank *et al*.,2009) (Syeda *et al*., 2019). A custom built; dual colour bespoke single molecule microscope that uses a narrow 10µm at full width half maximum excitation field at the sample plane was used (**Figure 3-4**). A bespoke GFP/mCherry channel splitter was used as described previously (Wollman *et al*.,2015) (Syeda *et al*.,2019). The set up included a high magnification objective (NA 1.49 Apo TIRF 100× oil immersion, Nikon) and the detector was an Andor iXon 128 emCCD camera operating at a 12-bit gain at 180 Hz and 5ms exposure/frame (Payne-Dwyer *et al*.,2022). The samples were illuminated for single colour imaging in either brightfield or for Slimfield illumination in camera triggered frames by a 20mW 480 nm wavelength continuous wave OPSL laser to a power density of 4kW/cm^2 or a 40mW 561 nm wavelength continuous wave OPSL laser to a power density of 23kW/cm^2. For dual colour imaging excitation was from 20mW 488 nm and 40mW 561 nm OPSL lasers digitally modulated to produce an alternating laser excitation within a 5ms period. The number of frames per acquisition was 1000 for the mGFP and mCherry strains.

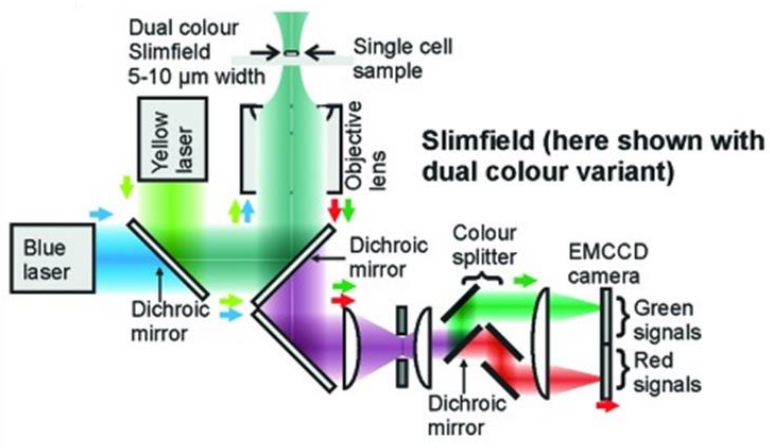


Figure 3-4: Diagram illustrating the Slimfield microscope in dual colour mode. (Taken from Shashkova and Leake,2017)

3.11 Slimfield Microscopy Image analysis

Slimfield microscopy images were processed using the ADEMs code software in MATLAB (MathWorks) (GitHub). The main code for detecting single molecule spots is 'tracker'. All the codes, including 'tracker' are available in the Leake group (https://github.com/awollman/single-molecule-tools).

3.11.1 Image Analysis – Spot Correction and Subpixel localisation

Using the 'tracker' code, bright foci in the images were first identified using a series of image transformations and thresholding steps, allowing identification of the candidate foci. Centroids of the candidate foci were determined using Gaussian masking and accepted if their intensity was greater than a signal to noise ratio (SNR) of 0.4 (Syeda *et al*.,2019) (Payne-Dwyer *et al*.,2022). Intensity was calculated by subtracting the sum of the pixel intensities inside a 5-pixel circle region of interest (ROI) from the pixel intensities in an outer square ROI of 17x17 pixels (background) (**Figure 3-5**). The SNR was defined as the mean background corrected pixel intensity in the circular ROI divided by the standard deviation of the pixel intensities in the square ROI (Syeda *et al*.,2019) (Payne-Dwyer *et al*.,2022). Foci were linked together into trajectories between frames if they were within 5 pixels of each other. Linked foci were accepted as "tracks" if they persisted for 5 frames or more at 5ms exposure to ensure less noise was picked up (Syeda *et al*.,2019) (Payne-Dwyer *et al*., 2022).

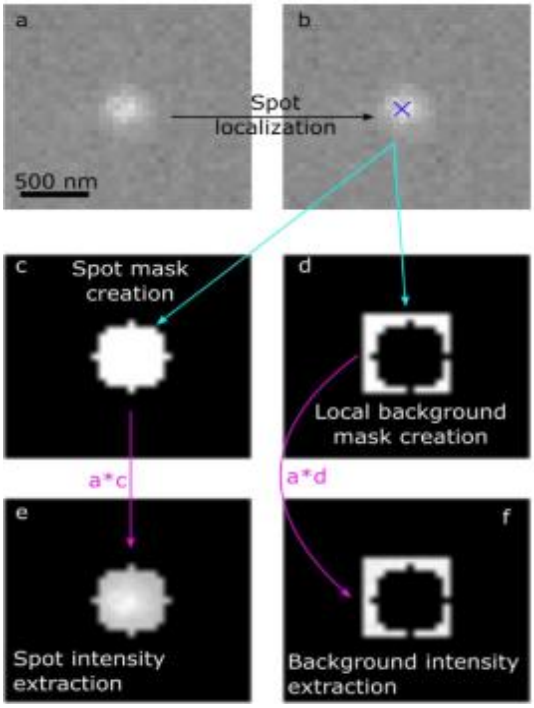


Figure 3-5: Spot correction and subpixel localisation using custom built MATLAB code. The process by which the MATLAB 'tracker' code identifies candidate foci and subtracts the background is illustrated above (Taken from Shepherd *et al*., 2021).

3.11.2 Image Analysis – Isingle Determination

Once the single molecules had been detected, the brightness (intensity) of a single mGFP and mCherry molecule was determined from *in vivo* Slimfield data and then compared to values obtained from *in-vitro* immobilisation assays which used isolated, individual fluorophores (**Table 3-2**). A comparison was made as fluorophores can dimerise in *vivo* making it difficult to determine the brightness of a single fluorophore molecule. To determine the brightness (intensity) of single fluorophores the 'getallspots' code was used. This code takes all the spots of mGFP or mCherry throughout the frames and identifies the fluorescent proteins in all the frames. After 1/3rd of the time of the experiment (approximately 300 frames in out of 1000, so all spots after 1/3rd bleach), the code reaches the photo blinking region in which single molecules are turning on and off. Spots in this region are therefore believed to have characteristics of a single fluorophore (Leake *et al*.,2006). For all spots after 300 spots, a plot was generated showing the intensity vs the probability with the peak of this plot corresponding to the brightness of a single fluorophore. The brightness (intensity) of a single fluorophore is called the Isingle value. The getallspots code generates the plot shown in **Figure 3-6** by default.

Fluorophore	*In vivo* Isingle (counts, A.U)	*In vitro* Isingle (Counts, A.U)
mGFP	142	155
mCherry	118	128

Table 3-2: *In vivo* and *in vitro* Isingle values for mCherry and mGFP. The Isingle value is the brightness of a single fluorophore molecule. Counts are fluorescence arbitrary units (A.U).

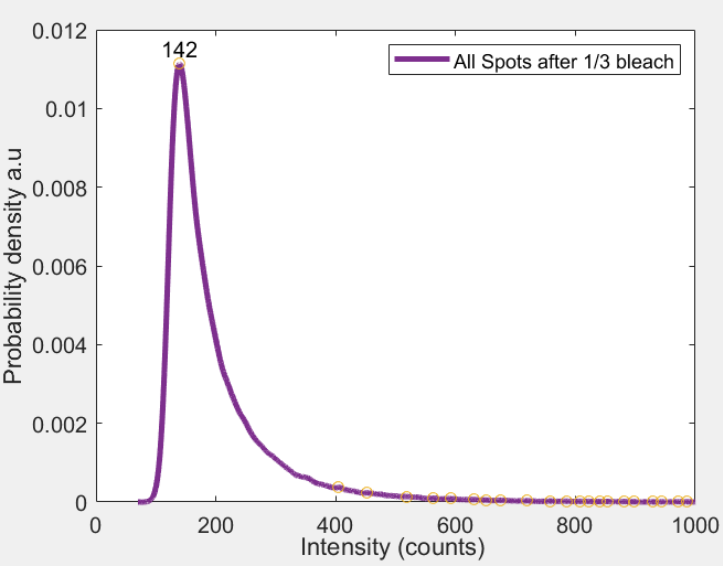


Figure 3-6: Single characteristic Intensity (Isingle plot) generated by the getallspots code. A graph generated by the getallspots code showing the brightness (intensity) of an mGFP molecule to be 142. Probability density is defined such that the area under the plot is always equal to 1.

3.11.3 Image Analysis - Stoichiometry, Copy Number & Diffusion Coefficient Determination

Before analysing the fluorescent signals inside the cell, the cell boundary was determined. Manual segmentation was performed using the fluorescent images and was used as a cell mask to identify where in the images to look for bright spots that correspond to individual molecular complexes. After cell segmentation and the determination of the brightness of single mGFP and mCherry fluorophores, the MATLAB code was used to calculate the stoichiometry, microscopic diffusion coefficient and copy number of SpoIIQ and SpoIIIAH in wildtype and mutant backgrounds. The copy number refers to the total number of SpoIIQ or SpoIIIAH molecules in cells and was calculated using the CoPro code. After background subtraction, the CoPro code adds up all the pixel values within the cell mask in question in the first bright frame of acquisition and divides this by the intensity of a single fluorophore molecule (Isingle) (Shashkova *et al*.,2021). Stoichiometry and diffusion coefficient values were calculated using the sampletrackanalyser code. The stoichiometry is the number of SpoIIQ or SpoIIIAH molecules per fluorescent spot and is calculated by dividing the summed pixel intensity values associated with the unbleached brightness of each foci by the Isingle value (Shashkova *et al*.,2021). The diffusion coefficient is a measure of a protein's mobility inside live cells. To calculate the microscopic diffusion coefficient, the centroids of the foci within each track, as generated from the ADEMScode tracking analysis above were used to calculate the displacements over the length of each track in a chronological sequence. From these the mean squared displacement of each track was calculated and the code takes a gradient of the mean squared displacement (M) as a function of time (t) to calculate the diffusion coefficient (Shashkova *et al*.,2021). Negative values in the diffusion coefficient plots arise as the code measures the gradient of M vs t but sometimes some tracks look like they have a negative gradient simply because diffusion as a process is stochastic (random with time).

3.12 Kernel Density Estimation (KDE)

Diffusion coefficients and molecular stoichiometries can be rendered as either a histogram or as a Kernel Density Estimation (KDE) plot. With histograms, data is binned across the observed distribution, but the size of the histogram and its position can lead to subjective bias. Subjective bias can lead to errors in the interpretation of molecular heterogeneity as distributions may be more complex than that assumed by histograms. As shown in **Figure 3-7**, if a histogram has too many bins it suggests the presence of more heterogeneity than there is, whereas too few bins hides molecular heterogeneity (Leake,2014). A KDE plot generates the most objective distribution from any data set for any single molecule property as the width of the KDE plot is set to the measurement error for the

property in that particular experiment (Leake,2014). In KDE plots, the position of the identified peaks can also be robustly quantified to identify single molecule states (Leake,2014). All single molecule data in this thesis is presented in KDE plots. The kernel width of the stoichiometry plots is 0.7 as the noise that comes from the intensity measurement of a single fluorophore molecule is typically 70% of the Isingle value. The kernel width of the diffusion coefficient plots is 0.05 and this comes from the error on the mobility measurements which come from the localisation precision on the tracking.

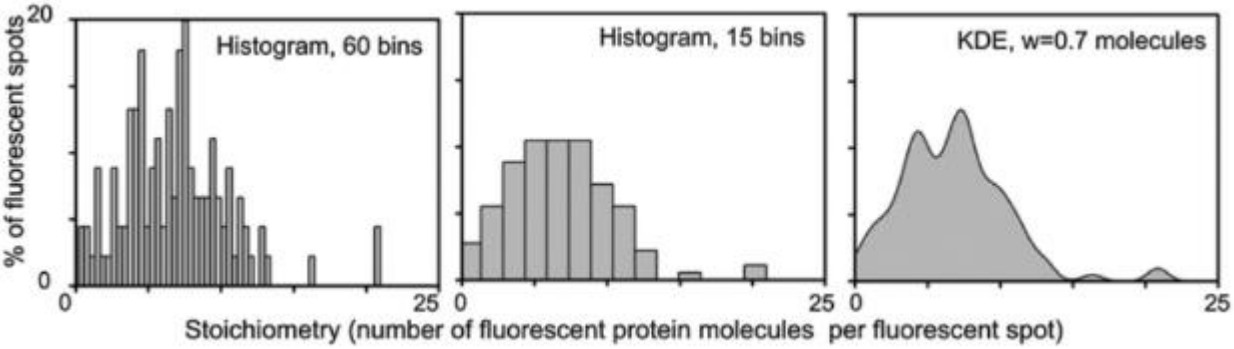


Figure 3-7: Objective rendering of single molecule parameters using Kernel Density Estimation (KDE). Histogram with too many bins (left) suggesting the presence of heterogeneity that is not there. Histogram with too few bins (middle) hiding heterogeneity. KDE plot (right) is the most objective way to render a single molecule property. The KDE plot above has a width of 0.7 molecules and represents the stoichiometry of a bacterial flagellar motor (Taken from Leake, 2014).

3.12.1 Interpretation of Kernel Density Estimation (KDE) plots

Stoichiometry KDE plots often show clear gaps. These gaps occur as each detected fluorescent focus may contain more than one subunit of a given protein (**Figure 3-8**). The smallest reproducible peak on a KDE plot is representative of the subunit stoichiometry of a protein (red arrow, **Figure 3-8**) and the periodic peaks on the stoichiometry distribution are integer multiples of the subunit stoichiometry (Syeda *et al*.,2019) (**Figure 3-8**). Peaks in the KDE plots are detected by MATLABs find peaks function which examines the overall noise trace and identifies possible peaks above the noise. Using the automatic find peaks function, the software pulls out what appears to be a likely subunit at a certain probability. The 'periodicity spacing' refers to the spacing between the peaks (blue arrows, **Figure 3-8**) and is often consistent with the subunit stoichiometry (the smallest peak on the KDE plot) (Syeda *et al*.,2019). In some instances, there may be differences between the 'periodicity spacing' and the size of the fundamental subunit stoichiometry. This may occur as spots with higher stoichiometry are brighter and easier for the software to detect above the noise. Dimmer spots are more difficult for the software to detect above the noise, especially if they are moving quickly. In this scenario, the sensitivity of the detection function may not be high enough to detect the fundamental peak meaning a clear

peak at low stoichiometry is not observed. A periodicity spacing is observed as this is the difference in stoichiometries from the very bright peaks. A clear agreement between the fundamental subunit stoichiometry peak and the periodicity spacing gives a high level of confidence that a real subunit has been detected. Inconsistencies between the fundamental subunit stoichiometry and periodicity spacing mean a lower level of confidence when interpreting the subunit size. The KDE plot in Figure 3-8 concludes that SpoIIQ is a dimer.

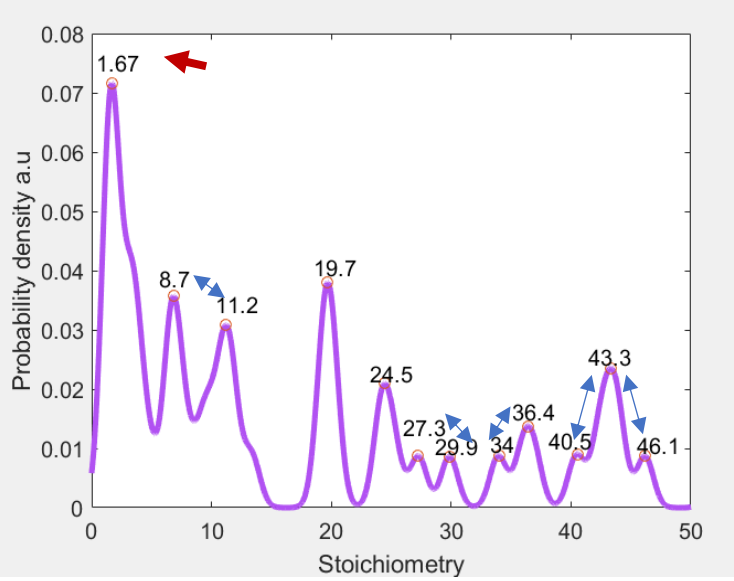


Figure 3-8: Stoichiometry distribution of SpoIIQ using Kernel Density Estimation (KDE). KDE plot showing the smallest peak at 1.67 (highlighted by the red arrow) which suggests that the fundamental subunit stoichiometry of SpoIIQ is two, which is most consistent with SpoIIQ forming a dimer. The peaks at higher stoichiometries are integer multiples of the subunit stoichiometry. The periodicity spacing is the spacing between the peaks and is approximately two molecules within the margin of error (blue arrows). The use of 'probability density' reflects the fact that each distribution is continuous and the area under the plot is always equal to 1.

3.13 Restriction Enzyme Digestion

0.5-1 µg of plasmid DNA was used for the diagnostic restriction enzyme digestions to verify plasmids. 2-5 units of restriction enzyme was used in a total reaction volume of 20 µl containing 2 µl of the corresponding buffer supplied at 10x concentration by the manufacturer. Restriction enzyme digestion reactions were incubated for 1 hour and fragments were analysed using agarose gel electrophoresis. Commercially available DNA size markers were run alongside the digestion samples to estimate the sizes of the restriction fragments. For cloning, 2-5 µg of plasmid DNA was used in larger total reaction volumes (50 µl) to maximise the concentration of the desired product after gel extraction.

3.14 PCR Amplification

Genomic DNA was isolated from the *mGFP spoIIQ delspoIIQ* strain (Strain SS102, **Appendix 1**) using the Qiagen Genomic DNA (gDNA) Extraction kit as per the manufacturer's specifications. PCR was used to amplify the pspoIIQ-mGFP-spoIIQ insert from genomic DNA using the pspoIIQ primer AH59 (**Appendix 1**) and reverse spoIIQ primer AH72 (**Appendix 1**). 4 50 µl PCR reactions were set up as outlined in Table 3-3 and the samples were subjected to 35 cycles of amplification using the protocol conditions in Table 3-4.

Component	50µl Reaction	Final Concentration
5X Q5 Reaction Buffer	10µl	1X
10 mM dNTPs	1µl	200 µM
10 µM Forward Primer	2.5 µl	0.5 µM
10 µM Reverse Primer	2.5 µl	0.5 µM
Template DNA	2 µl	0.5-1 µg
Q5 High-Fidelity DNA Polymerase	1 µl	2 units
5X Q5 High GC Enhancer	10 µl	1X
Nuclease-Free Water	Up to 50 µl	

Table 3-3: PCR reaction set up.

Step	Temp	Time
Initial Denaturation	98°C	1 min
35 Cycles	98°C 72°C 72°C	30 seconds 30 seconds 4 mins
Final Extension	72°C	10 mins
Hold	4°C	Hold

Table 3-4**:** Thermocycling Conditions for PCR

3.15 Q5 Site Directed Mutagenesis

Q5 Site Directed Mutagenesis was used to mutate the two SpoIIQ leucine residues: Leu109 and Leu118 to alanine, phenylalanine, and glutamic acid. Non overlapping primers (**Appendix 1**) were designed using the NEBaseChanger tool and annealing temperatures were calculated using the NEB Tm Calculator. The reaction compositions and thermocycling conditions for Q5 site-directed mutagenesis are shown in **Table 3-5** and **Table 3-6** respectively.

	25 µl RXN	Final Concentration
Q5 Hot Start High-Fidelity 2X Master Mix	12.5 µl	1X
10 µM Forward Primer	1.25 µl	0.5 µM
10 µM Reverse Primer	1.25 µl	0.5 µM
Template DNA (1-25 ng/µl)	1 µl	1-25 ng
Nuclease-free water	9.0 µl	

Table 3-5: Q5 Site directed mutagenesis reaction composition.

	Temperature	Time
Initial Denaturation	98°C	30 seconds
25 cycles	98°C 58-72°C 72°C	10 seconds 30 seconds 4 minutes
Final Extension	72°C	2 minutes
Hold	4°C	

Table 3-6: Thermocycling conditions for Q5 site directed mutagenesis.

The PCR product was then phosphorylated, ligated, and digested with DpnI using the Kinase, Ligase & DpnI (KLD) treatment mix. Reactions were set up as illustrated in **Table 3-7** and reagents were mixed well by pipetting gently up and down before a 5-minute incubation at room temperature.

	10 µl RXN	Final Concentration
PCR Product	1 µl	
2X KLD Reaction Buffer	5 µl	1X
10X KLD Enzyme Buffer	1 µl	1X
Nuclease-free water	3 µl	

Table 3-7: Kinase, Ligase & DpnI (KLD) treatment reaction composition

After KLD treatment, 5 µl of the KLD reaction mixture was added to 50 µl of NEB chemically competent cells and samples were incubated on ice for 30 minutes. After 30 minutes, samples were heat shocked in a 42°C water bath for 30 seconds and then incubated on ice again for 5 minutes. Following this, 950 µl of SOC was added to the cells and the cells were shaken gently in an incubator for 1 hour at 37°C. Cells were then spun in a centrifuge set to 6000rpm for 5 minutes. After centrifugation, the supernatant was discarded, and the pellet resuspended in 200 µl SOC. 100 µl of the mixture was spread onto LB Kanamycin (30 µg/ml) plates and incubated at 37°C overnight. The remaining 100 µl was stored in the fridge at 4°C.

Chapter 4- Stoichiometry, Diffusion Coefficient and Copy Number Analysis of SpoIIQ and SpoIIIAH in wildtype backgrounds

4.1 Introduction

In *B. subtilis*, the forespore protein SpoIIQ and mother cell protein SpoIIIAH are involved in engulfment and membrane migration. During engulfment, both proteins initially localise at the sporulation septum before forming arcs and foci around the engulfing membrane (Rubio and Pogliano, 2004) (Blaylock *et al*.,2004). In 2008, it was shown that SpoIIIAH was accessible to enzymatic modification by biotin ligase produced in the forespore cytoplasm. This, combined with the similarity of the SpoIIIAH C-terminal domain to YscJ/FliF like proteins that form multimeric rings in type III secretion systems led to the hypothesis that SpoIIQ and SpoIIIAH form a channel complex spanning the forespore-mother cell intermembrane space (Meisner *et al*., 2008). SpoIIQ is a membrane protein expressed under σ^F control and has an N-terminal transmembrane segment and C-terminal extracellular domain (Meisner *et al*.,2012). SpoIIIAH is encoded by the *spoIIIAH* gene, the last gene in the *spoIIIAA* operon. The SpoIIIAH protein is expressed under σ^E control and has an N-terminal transmembrane segment and C-terminal extracellular domain. (Meisner *et al*.,2012). The C-terminal domain shows similarity to the YscJ/FliF family of type III secretion systems with proteins of this family forming multimeric rings in secretion complexes in Gram negative bacteria (Meisner *et al*.,2012).

Slimfield microscopy allows millisecond time scale imaging of single fluorescently labelled proteins (Wollman *et al*.,2015). In this chapter, SpoIIQ and SpoIIIAH were fused to fluorescent proteins and single fluorescently labelled SpoIIQ and SpoIIIAH foci were tracked to super resolution precision in live *B. subtilis* cells. The stoichiometry of SpoIIQ and SpoIIIAH was determined, giving an indication of the single-subunit stoichiometry of both proteins. The precise number of molecules of SpoIIQ in the forespore and SpoIIIAH in the mother cell was calculated with copy number estimates suggesting equal numbers of SpoIIQ and SpoIIIAH in live *B. subtilis* cells. The mobility of SpoIIQ and SpoIIIAH in live *B. subtilis* cells was measured through calculation of the microscopic diffusion coefficient and suggested both proteins have a relatively low mobility, a finding expected for membrane proteins believed to form a channel complex.

4.2 *mGFP-spoIIQ* Fusion Construction

A plasmid containing an *mGFP-spoIIQ* fusion was constructed for the integration of the *mGFP-spoIIQ* fusion into the *B. subtilis* chromosome at the *amyE* locus. This method is advantageous when compared to integration at the native *spoIIQ* locus for a double cross over event, as *amyE* integration does not require insert – chromosome homology (Shimotsu and Henner, 1986). To begin cloning, the pSG1154 plasmid shown in **Figure 4-1** was digested with the restriction enzymes SpeI and HindIII and the fragments were separated by agarose gel electrophoresis. As shown in **Figure 4-1**, a double digest of pSG1154 with SpeI and HindIII is expected to yield two fragments: a 7kb fragment corresponding to the linearised vector backbone and a 600bp fragment corresponding to the gfpmut1 sequence. The agarose gel in **Figure 4-1** therefore suggests that restriction enzyme digestion of pSG1154 was successful as both bands matched the expected sizes. Following gel electrophoresis, the 7kb linearised pSG1154 vector backbone was extracted from the gel using the Qiagen Gel Extraction kit as per the manufacturer's specifications for use in subsequent cloning steps. The 600bp fragment encompassing the gfpmut1 sequence was discarded.

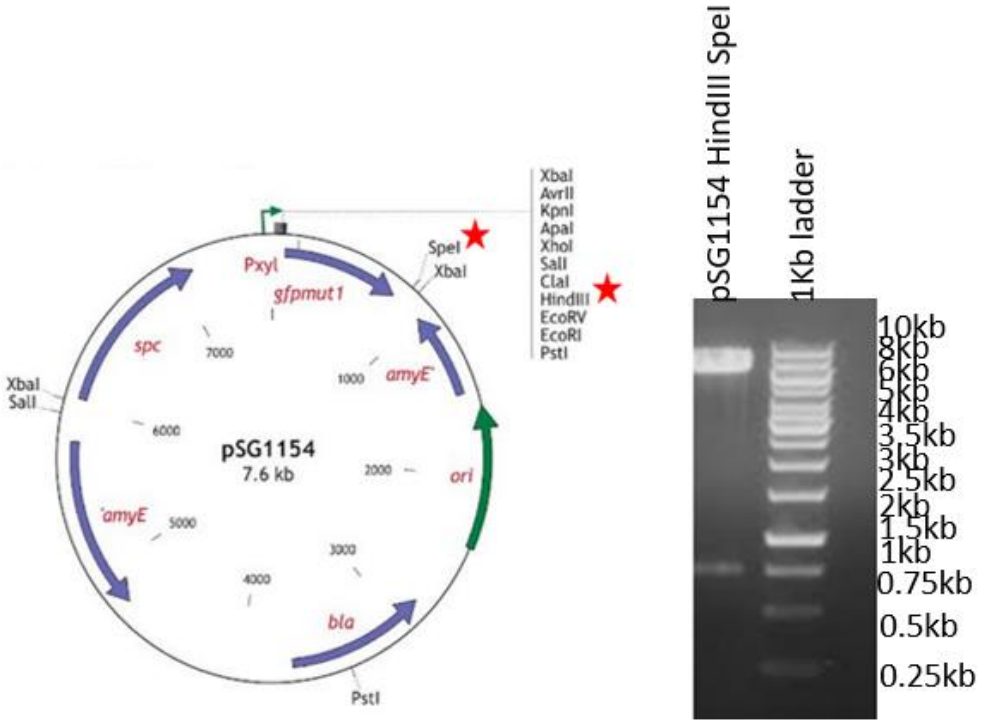


Figure 4-1: Digestion of pSG1154 with SpeI and HindIII. Left) Plasmid map of pSG1154 showing that a double digest of pSG1154 with SpeI and HindIII is expected to yield two bands: a 7kb fragment corresponding to the linearised vector backbone and a 600bp fragment corresponding to the gfpmut1 sequence. SpeI and HindIII restriction sites are highlighted with red stars. Right) Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of pSG1154 with SpeI and HindIII showing successful digestion as the bands matched the expected sizes.

After the pSG1154 linearised vector backbone was extracted from the gel, the three fragments required to construct the *mGFP-spoIIQ* fusion were prepared by PCR (**Figure 4-2**). Two PCR reactions were set up to amplify each fragment and agarose gel electrophoresis was used to analyse the PCR products. As shown in **Figure 4-2**, amplification of *pspoIIQ* from *B. subtilis* chromosomal DNA yields two 300bp fragments as *pspoIIQ* is 343bp long. Amplification of spoIIQ yields two DNA fragments at 1kb as the spoIIQ coding sequence is 1kb long. Amplification of mGFP from pSG1729 yields two 700bp fragments as the mGFP sequence is 760bp long. After gel electrophoresis, the PCR products were cleaned up using the Qiagen PCR Purification kit as per the manufacturer's specifications before being used in the Gibson Assembly reaction.

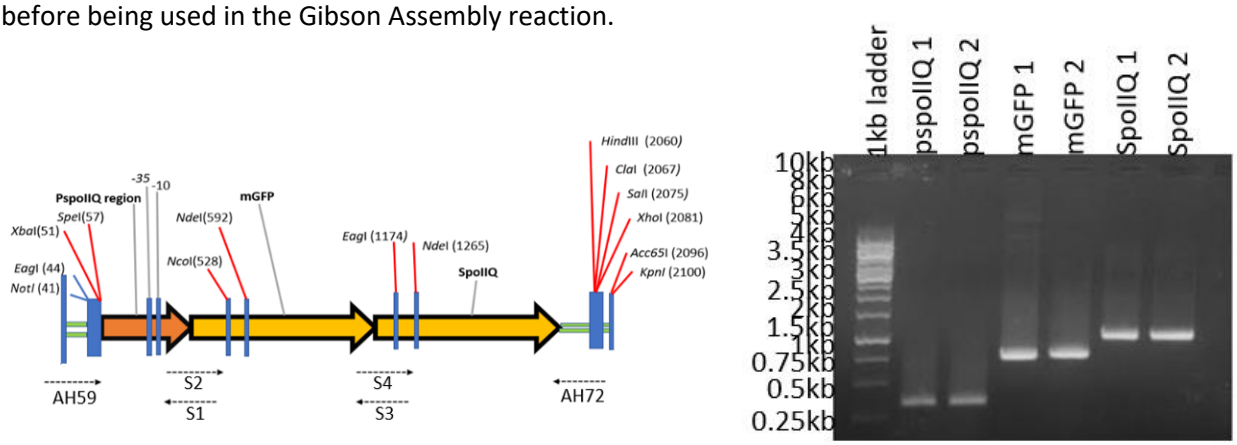
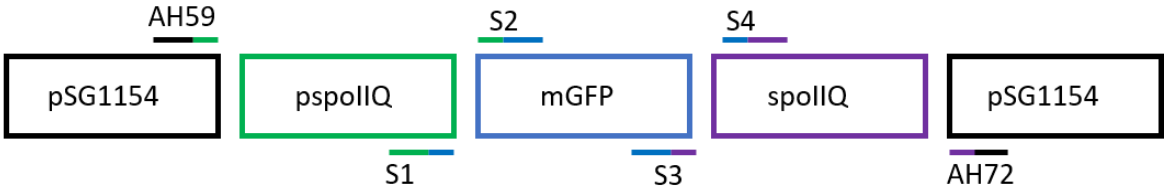


Figure 4-2: PCR amplification of the three DNA fragments required to make the *mGFP-spoIIQ* fusion. Left) Components for the *mGFP-spoIIQ* fusion showing the three DNA fragments required to construct the integration plasmid and location of primers. Right) Agarose gel electrophoresis using a 1% agarose gel showing the products of PCR amplification of pspoIIQ, mGFP and spoIIQ. Two reactions were performed for each amplification and fragments closely matched expected sizes indicating successful amplification.

Gibson assembly was next used to assemble the DNA fragments required to construct the pSG1154-mGFP-spoIIQ plasmid (**Figure 4-3**). Gibson Assembly was performed as described in section 3.2 of the methods chapter and the reaction used pSG1154, pspoIIQ, SpoIIQ and mGFP fragments (**Figure 4-3**). The assembly reaction was mixed with NEB 5 alpha competent cells and the cells were plated on LB ampicillin plates and incubated overnight at 37°C. The next day colonies were used in colony PCR to establish the presence of the pspoIIQ-mGFP-spoIIQ insert using the forward pspoIIQ primer AH59 (**Appendix 1**) and reverse spoIIQ primer AH72 (**Appendix 1**). 15 colonies were used in the colony PCR reaction and the PCR products were analysed using agarose gel electrophoresis. As shown in **Figure 4-4**, the PCR products from colonies 7 and 13 showed bands with a mobility relative to the markers of 2kb, suggesting the presence of the desired recombinant plasmid in these colonies. The PCR products from all other colonies showed a single

band with a mobility relative to the markers of 1.2kb and were discarded as this band was considered a non-specific PCR artefact.



Primer Name	Sequence	Forward/Reverse	Overlaps	Anneals
AH59	**AATTCGCGGCCGCTCTAGAACTAGT**CATAGAGCGGACGTGACTC	Forward	pSG1154	pspoIIQ
S1	GTTCTTCTCCTTTACTCATTGTTTCATCACCTCAGCAAC	Reverse	mGFP	pspoIIQ
S2	TGTTGCTGAGGTGATGAAACAATGAGTAAAGGAGAAGAAC	Forward	pspoIIQ	mGFP
S3	GTTTTCTTTTCTTCCTCTCTCATTTTTTTGTATGGTTCATC	Reverse	spoIIQ	mGFP
S4	GATGAACCATACAAAAAAATGAGAGAGGAAGAAAAGAAAAC	Forward	mGFP	spoIIQ
AH72	**GAGGTCGACGGTATCGATAAAGCTT**GAGCAATAGGCGGTTCATG	Reverse	pSG1154	spoIIQ

Figure 4-3: Construction of the pSG1154 recombinant plasmid using Gibson Assembly. Top) Schematic representation of the four fragments required to construct the pSG1154-mGFP spoIIQ plasmid. Coloured lines above each fragment represent the primers required for PCR amplification with complementary represented by the corresponding colours. Bottom) Table showing the sequences of the primers required for Gibson Assembly. The direction of the primers as well as the sequences they overlap and anneal too are highlighted.

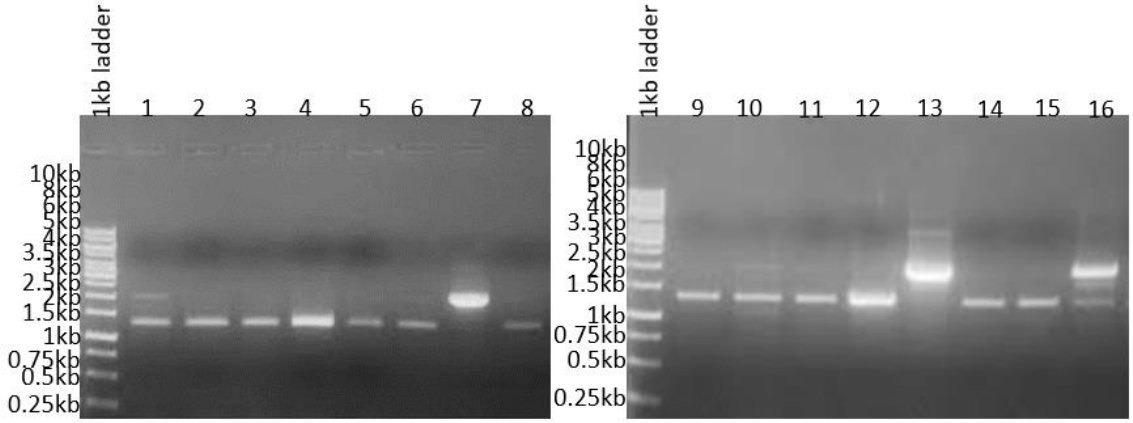


Figure 4-4: Agarose gel electrophoresis using a 1% agarose gel of the PCR products from the colony PCR of 15 colonies numbered 1-15. PCR products from colonies 7 and 13 show bands with a mobility relative to the markers of 2kb suggesting that the desired recombinant plasmid containing the 2kb pspoIIQ-mGFP-spoIIQ insert is present in these colonies. All other colonies show a band with a mobility relative to the markers of 1.2kb suggesting these colonies do not contain the desired recombinant plasmid and this band results from the non-specific binding of primers.

Colonies 7 and 13 were grown in LB supplemented with 100 µg/ml ampicillin overnight at 37°C and plasmids were isolated the following day using the Qiagen Mini Prep kit as per the manufacturer's specifications. To further confirm the presence of the pspoIIQ-mGFP-spoIIQ insert, plasmids 7 and 13 were digested with the restriction enzymes NotI and NdeI and fragments were separated by agarose gel electrophoresis (**Figure 4-5**). As shown in the plasmid map of the desired psG1154-mGFP spoIIQ recombinant plasmid (**Figure 4-5**), a single digest with NdeI is expected to yield three fragments: fragments at 3kb and 4kb, corresponding to pSG1154 vector backbone and a 700bp fragment corresponding to the mGFP sequence. A double digest with NotI and NdeI is expected to yield four fragments: fragments at 3kb and 4kb corresponding to the pSG1154 vector backbone, a 500bp fragment corresponding to the pspoIIQ sequence and a 700bp fragment corresponding to the mGFP sequence (**Figure 4-5**). The agarose gel in **Figure 4-5** suggests that the restriction enzyme cleavage of plasmids 7 and 13 was successful as fragments matched the expected sizes**.**

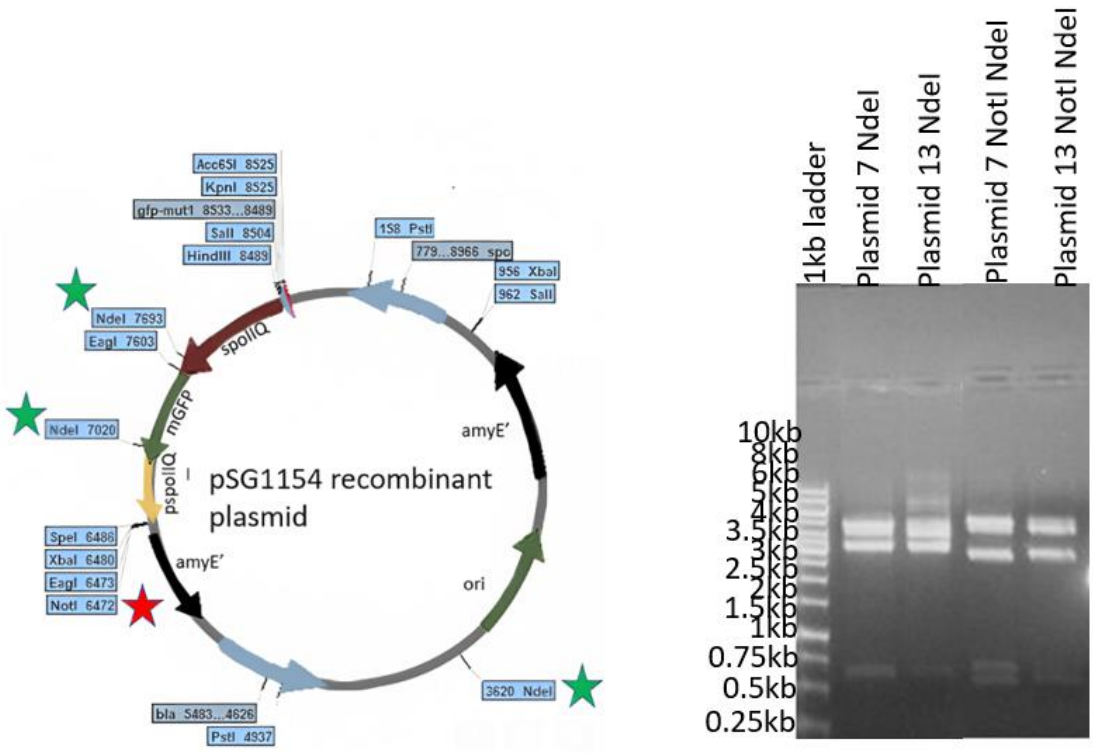


Figure 4-5: Restriction enzyme digestion of pSG1154 recombinant plasmids 7 and 13 with NotI and NdeI. Left) Plasmid map of the desired pSG1154-mGFP spoIIQ plasmid showing the location of NdeI restriction sites (green stars) and NotI restriction sites (red stars). Digestion of the pSG1154-mGFP spoIIQ plasmid with NdeI is expected to yield three fragments: fragments at 3kb and 4kb corresponding to the pSG1154 vector backbone and a 700bp fragment corresponding to the mGFP sequence. Digestion of the pSG1154-mGFP spoIIQ plasmid with NotI and NdeI is expected to yield four fragments: fragments at 3kb and 4kb corresponding to the pSG1154 vector backbone, a 700bp fragment corresponding to the mGFP sequence and a 500bp fragment corresponding to the pspoIIQ sequence. Right) Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of plasmids 7 and 13 with NotI and NdeI. The agarose gel suggests restriction enzyme digestion was successful as bands matched the expected sizes.

Plasmids 7 and 13 were next used to transform *B. subtilis* strain IB714 (**Table 4-1**) as outlined in section 3.6 of the methods chapter. *B. subtilis* strain IB1714 was chosen for transformation as it contains an erythromycin cassette at the *amyE* locus allowing integration of the *mGFP-spoIIQ* fusion into the *B. subtilis* chromosome at *amyE* via a double cross over event. As illustrated in **Figure 4-6**, integration of the fusion results in a switch in antibiotic resistance from erythromycin to spectinomycin as the erythromycin resistance from the chromosome is lost and spectinomycin resistance from the plasmid is gained.

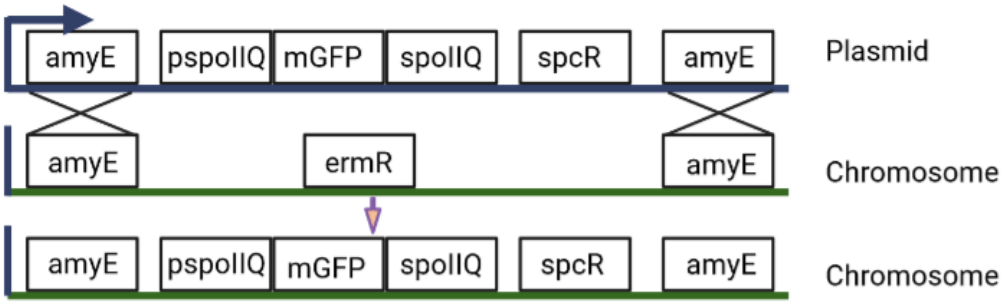


Figure 4-6: Diagram showing how the *mGFP spoIIQ* fusion integrates into the *B. subtilis* chromosome at the *amyE* locus via a double cross over event. There is a switch in antibiotic resistance from erythromycin to spectinomycin.

4.3 mGFP SpoIIQ imaging – SIM microscopy

To check if labelling SpoIIQ with mGFP at the *amyE* locus affected sporulation and the localisation of SpoIIQ, the resulting strain (SS101, **Table 4-1**) was imaged using Structured Illumination Microscopy (SIM). SIM was used to image the *mGFP spoIIQ* strain as the technique allows imaging of fluorescently labelled biological samples with a spatial resolution improved by a factor of two compared to traditional optical microscopy methods (Shaw *et al*., 2015). Research studies using GFP SpoIIQ fusions have shown that SpoIIQ shows a dynamic localisation pattern during engulfment (Blaylock *et al*., 2004). As shown in **Figure 4-7,** SpoIIQ initially localises at the sporulation septum before forming arcs and foci at the forespore/mother cell interface (Blaylock *et al*., 2004). SIM microscopy imaging of mGFP SpoIIQ shows that SpoIIQ localisation is as expected as the arcs and foci observed in **Figure 4-8** are consistent with those expected to form during engulfment and membrane migration (**Figure 4-7**). SIM microscopy imaging therefore suggests that the *mGFP spoIIQ* strain (SS101, **Table 4-1**) is ready to use in subsequent steps as it is sporulating and the localisation of SpoIIQ is as expected. This experiment was a single experiment and was not compared quantitatively to previous published results with other strains.

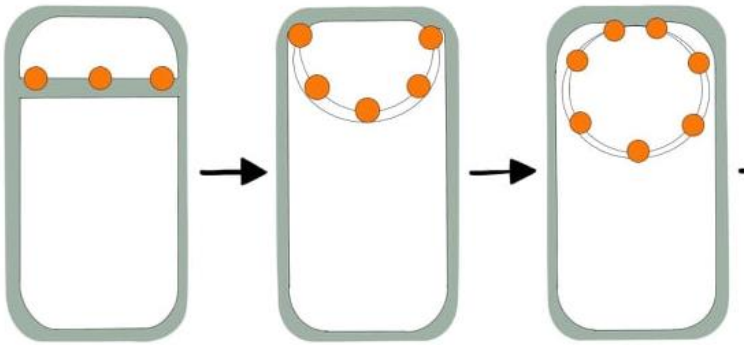


Figure 4-7: Cartoon representation of the localisation of SpoIIQ during engulfment. SpoIIQ (orange circles) initially localises at the sporulation septum before forming arcs and foci at the forespore/mother cell interface (Adapted from Blaylock *et al*.,2004).

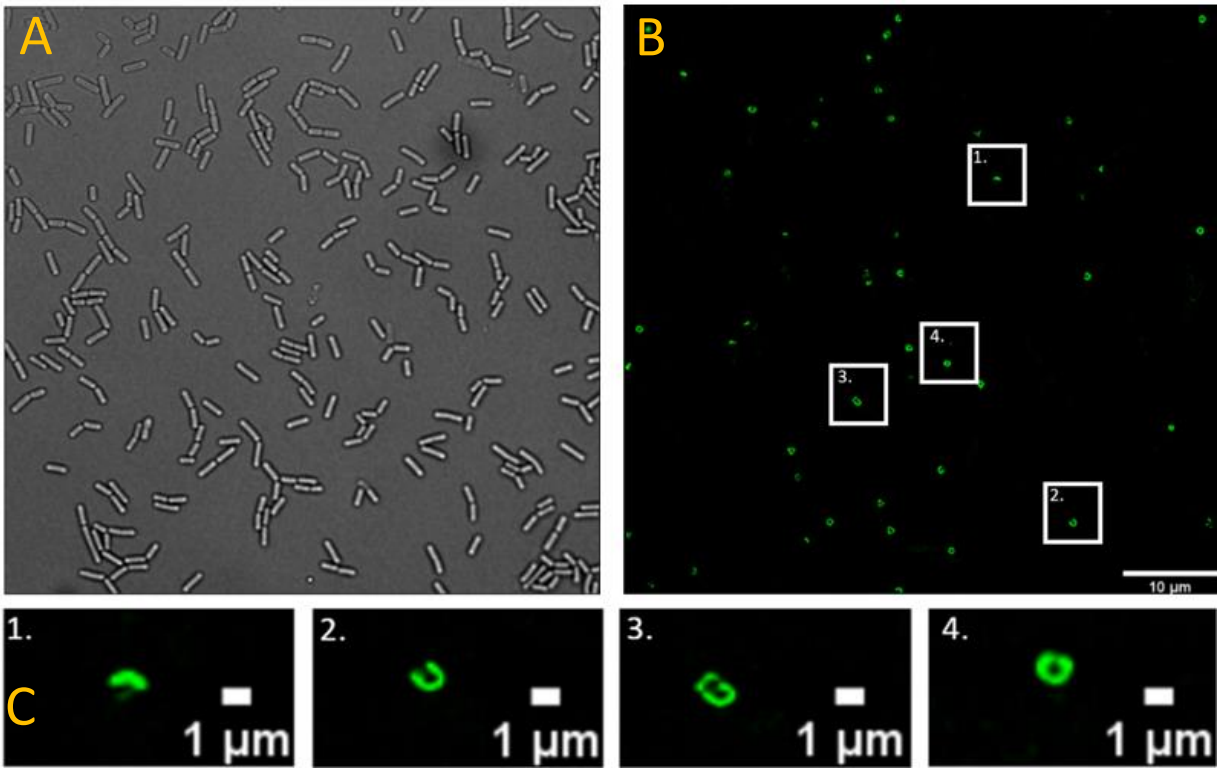


Figure 4-8: SIM microscopy imaging of mGFP SpoIIQ. A) Brightfield image showing live *B. subtilis* cells. B) SIM microscopy fluorescent image showing strong fluorescent signals and the presence of SpoIIQ around the engulfing and engulfed forespore membranes. C) Close up fluorescent images showing the dynamic localisation pattern of SpoIIQ during engulfment (left to right, 1-4). SpoIIQ first localises at the sporulation septum and then forms arcs and foci as engulfment proceeds until completion.

4.4 mGFP SpoIIQ delSpoIIQ Strain Construction

The strain described above has *mGFP spoIIQ* at the *amyE* locus and wildtype *spoIIQ* at the native *spoIIQ* locus (Strain SS101, **Table 4-1**). The native *spoIIQ* gene was next disrupted using chromosomal DNA from the spoIIQ null mutant strain BKE36550 (*Bacillus subtilis* Stock Centre) in which the *spoIIQ* gene is disrupted by an erythromycin cassette. The native *spoIIQ* gene was disrupted to ensure that

only fluorescently labelled SpoIIQ was present which would allow clear interpretation of the biological function of SpoIIQ in live *B. subtilis* cells.

Competent cells were first prepared from the existing mGFP spoIIQ strain (SS101, **Table 4-1**), and chromosomal DNA from the spoIIQ null mutant BKE36550 was used for transformation. Transformation was carried out using different volumes of the delspoIIQ chromosomal DNA as described in section 3.6 of the methods chapter. After an overnight incubation at 37°C, 32 colonies were observed across the test plates and no colonies were observed on the negative control plate in which the delspoIIQ chromosomal DNA was excluded, as expected. The colonies were next streaked out on LB spectinomycin and LB erythromycin plates and all colonies were expected to grow on both antibiotics as erythromycin resistance is gained from the delspoIIQ chromosomal DNA and spectinomycin resistance is present in strain SS101 before transformation and should be retained. 31 colonies grew on both LB spectinomycin and LB erythromycin meaning one colony was discarded. Before SIM was used for imaging, colony PCR was performed using primers AH59 and AH72 to ensure that transformation had not caused a loss of GFP fluorescence at the *amyE* locus (**Figure 4-3**). For colony PCR, 13 colonies were selected from the 31 colonies growing on both spectinomycin and erythromycin and colony DNA from the mGFP spoIIQ strain (SS101) was used as a positive control in the PCR reaction. The PCR products were analysed using agarose gel electrophoresis and as shown in **Figure 4-9**, the PCR products from all colonies besides colony 7 showed a band with a mobility relative to the markers of 2kb, corresponding to the pspoIIQ-mGFP-spoIIQ insert. PCR products from the positive control (colony DNA from strain SS101) were loaded in lane 1 adjacent to the 1kb ladder and showed a bright 2kb band as expected. The genotype of the resulting strain SS102 is shown in **Table 4-1**.

Strain Identifier	Genotype
IB714	*amyE::erm*
BKE36550	*spoIIQ:: erm*
SS101	*IB714 amyE::mGFP-spoIIQ-spcR spoIIQ:: spoIIQ*
SS102	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ::erm*
AH103	*spoIIIAH:: mGFP-spoIIIAH-camR*
AH104	*spoIIIAH:: mCherry-spoIIIAH-camR*

Table 4-1: List of strains imaged and analysed in thesis chapter 4.

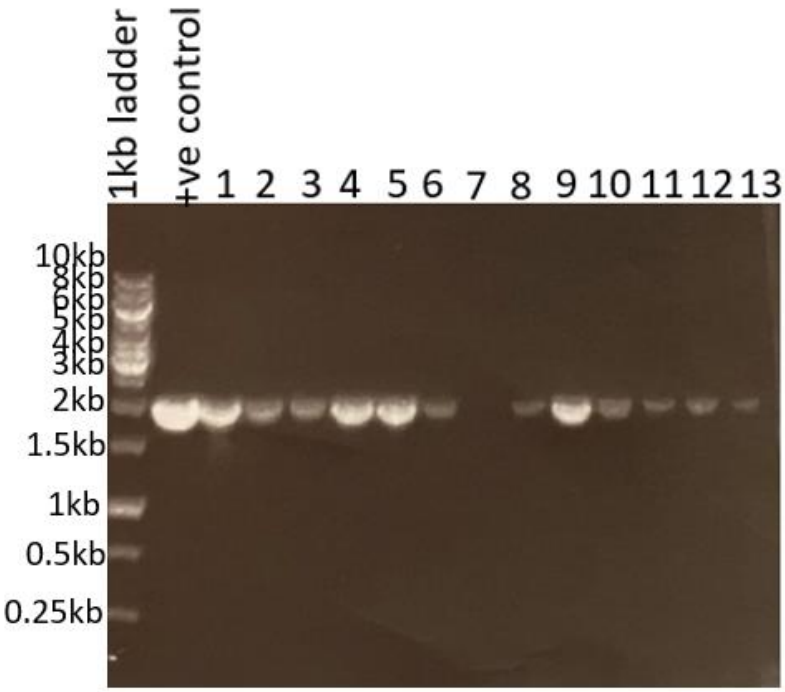


Figure 4-9: Agarose gel electrophoresis using a 1% agarose gel of the PCR products from the colony PCR of 13 colonies numbered 1 -13. PCR products from all colonies besides colony 7 show a band with a mobility relative to the markers of 2kb, corresponding to the desired pspoIIQ-mGFP-spoIIQ insert. Colony DNA was taken from the mGFP spoIIQ strain SS101 (Table 4-1) and used as positive control with the PCR products from this reaction showing a bright band at 2kb as expected.

Colonies 1,4,5 and 9 were selected and grown up for SIM microscopy based on the intensity of the desired 2kb bands on the agarose gel in **Figure 4-9**. Overnight cultures of colonies 1, 4, 5 and 9 were diluted in Difco Sporulation Medium (DSM) and grown for 8 hours to reach the engulfment stage of sporulation before being harvested and spotted onto 1% agarose pads. As shown in **Figure 4-10**, disrupting the native *spoIIQ* locus did not affect sporulation or the localisation of SpoIIQ. The localisation of SpoIIQ in the mGFP SpoIIQ delSpoIIQ strain (SS102, **Table 4-1**) is the same as the localisation of SpoIIQ in the mGFP SpoIIQ strain (SS101, **Table 4-1**) shown in **Figure 4-8.** As expected in both **Figures 4-8** and **4-10**, SpoIIQ initially localises at the sporulation septum and then forms arcs and foci around the engulfing and engulfed forespore membranes. The strong mGFP signal and correct localisation of SpoIIQ in **Figure 4-10** confirms that the mGFP SpoIIQ delSpoIIQ strain (SS102) is sporulating and ready to use in Slimfield experiments.

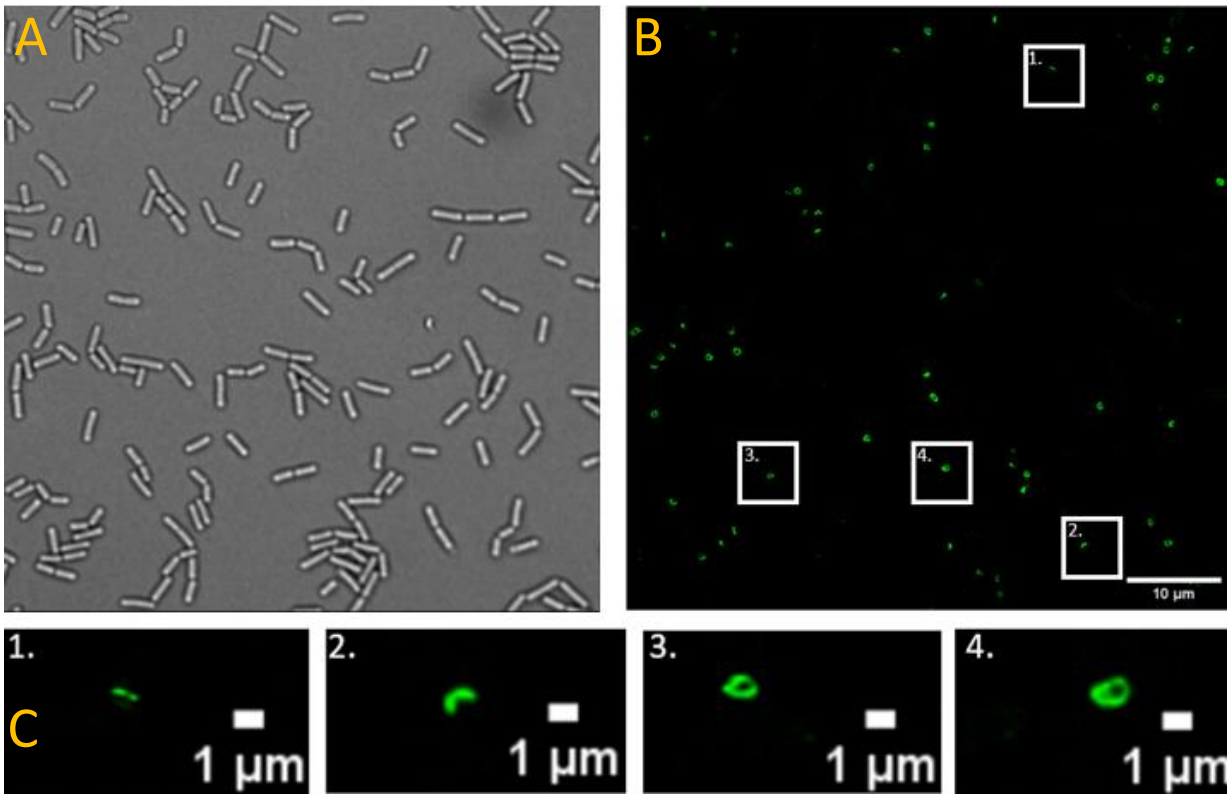


Figure 4-10: SIM microscopy imaging of mGFP SpoIIQ delSpoIIQ. A) Brightfield image showing live *B. subtilis* cells. B) SIM microscopy fluorescent image showing strong fluorescent signals and the presence of SpoIIQ around the engulfing and engulfed forespore membranes. C) Close up fluorescent images showing the dynamic localisation pattern of SpoIIQ during engulfment (left to right, 1-4). SpoIIQ first localises at the sporulation septum and then forms arcs and foci as engulfment proceeds until completion.

4.5 SpoIIIAH fusions

The *spoIIIAH* fusions imaged and analysed in this thesis were prepared by Adam Hughes prior to the start of this work and are listed in **Table 4-1**. SpoIIIAH fusions were designed for integration at the native *spoIIIAH* locus as *spoIIIAH* is the final gene in the *spoIIIAA* operon and it is impractical to clone the entire *spoIIIAA* operon at the *amyE* locus. As shown in **Figure 4-11**, *spoIIIAH* fusions contained a 500bp homologous sequence upstream and downstream of *spoIIIAH* (dspoIIIAH), with the upstream sequence corresponding to the 3' end of *spoIIIAG*. These sequences allowed the spoIIIAH fusions to recombine at the native *spoIIIAH* locus.

Figure 4-11: Schematic representation of the components required to make the spoIIIAH fusions. Coloured lines above fragments represent primers used for PCR amplification and black lines represent restriction endonuclease sites. *camR* corresponds to a chloramphenicol resistance cassette, *dspoIIIAH* corresponds to 500 base pairs downstream of *spoIIIAH*. (Taken from Adam Hughes, University of York, PhD Thesis 2018).

4.6 Slimfield Microscopy

Slimfield microscopy is a relatively inexpensive and simple super resolution microscopy technique that permits single molecule detection at high speed, allowing the molecular make up and time resolved features of *B. subtilis* sporulation to be resolved (Plank *et al*., 2009). In Slimfield microscopy, the normal fluorescence excitation field is reduced to encompass a single cell (**Figure 4-12**). This produces an excitation field with 100-1000x the intensity of standard epifluorescence microscopy and facilitates millisecond time scale imaging of single fluorescently labelled proteins by allowing for much greater signal intensity relative to normal camera imaging noise (Wollman *et al*.,2015). Millisecond timescale imaging is advantageous when studying sporulation in live *B. subtilis* cells as the time scale is fast enough not only to keep up with the diffusional motion present in the cytoplasm of cells, but it can also sample fast molecular transitions that occur in the millisecond regime (Wollman *et al*.,2015). In this thesis, Slimfield microscopy was used to determine the stoichiometry of SpoIIQ and SpoIIIAH fluorescent foci to the precision of single molecules allowing the single subunit stoichiometry of both proteins to be determined. Slimfield microscopy analysis also yielded information about the mobility of SpoIIQ and SpoIIIAH through calculation of the microscopic diffusion coefficient as well as the total number of molecules of both proteins in live cells. The data obtained from Slimfield microscopy experiments is computationally intensive, requiring a significant amount of in silico analysis, which was performed using MATLAB software available in the laboratory.

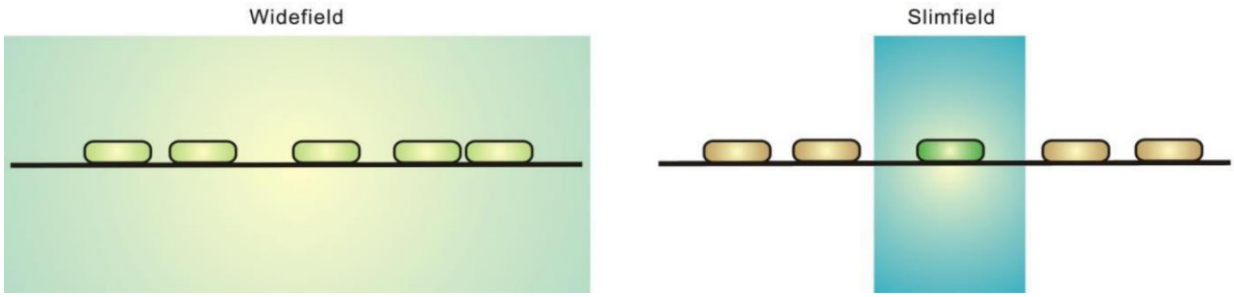


Figure 4-12: Schematic representation of Widefield and Slimfield microscopy. Conventional wide-field microscopy illuminates a wider field of the sample with lower excitation intensity whereas in Slimfield microscopy, the excitation light is concentrated in a smaller area providing a much greater excitation intensity in this area. (Taken from Chiu and Leake, 2011).

4.7 SpoIIQ Imaging – Slimfield Microscopy

The mGFP spoIIQ delspoIIQ strain (Strain SS102, **Table 4-1**) described above was imaged using Slimfield microscopy to confirm the robustness of the technique for imaging live *B. subtilis* cells. An overnight culture of the SS102 strain was diluted in Difco Sporulation Medium (DSM) and grown for 8 hours to reach the engulfment stage of sporulation before being harvested and spotted onto 1% agarose pads. As shown in **Figure 4-13**, the strain shows a strong fluorescent signal consistent with the presence of mGFP and the expected localisation of SpoIIQ. mGFP SpoIIQ is distributed around the engulfing and engulfed forespore membranes with some cells showing SpoIIQ localised at the asymmetric septum (Blaylock *et al*., 2004). This is expected as SpoIIQ shows a dynamic localisation pattern during engulfment and the process of engulfment is not always synchronous amongst cells in a population.

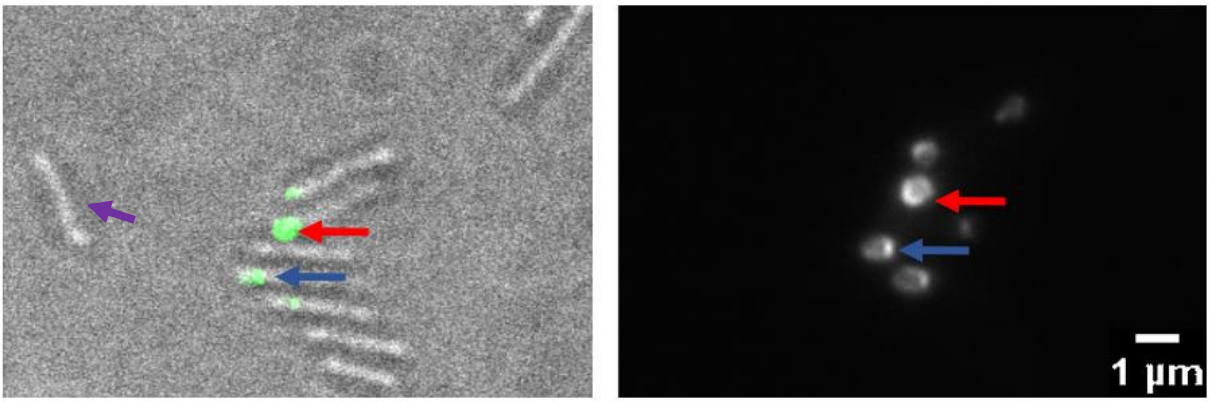


Figure 4-13: Slimfield microscopy of mGFP spoIIQ delspoIIQ. Left) Brightfield and fluorescent images superimposed showing SpoIIQ localised at the asymmetric septum (blue arrow) and at the engulfing and engulfed forespore membranes in arcs and foci (red arrow). Right) Fluorescent image showing SpoIIQ localised at the septum (blue) and forespore membranes (red). The cell highlighted with the purple arrow did not sporulate.

4.8 SpoIIIAH Imaging – Slimfield Microscopy

Two SpoIIIAH strains (AH103 and AH104, **Table 4-1**) were imaged using Slimfield microscopy to establish whether fluorescently labelling SpoIIIAH affected sporulation and the localisation of SpoIIIAH. Research studies using immunofluorescence microscopy have been used to visualise SpoIIIAH-Flag during engulfment (Blaylock *et al*., 2004). SpoIIIAH is initially randomly distributed throughout the mother cell membrane. In cells that have initiated engulfment, SpoIIIAH localises to the sporulation septum before tracking the engulfing mother cell membrane as it moves around the forespore (Blaylock *et al*., 2004) (**Figure 4-14**). At the end of engulfment, SpoIIIAH is localised around the forespore on the mother cell side. Slimfield microscopy images of mGFP and mCherry SpoIIIAH show that SpoIIIAH localisation is as expected as the bright fluorescent signals are localised to the engulfing and engulfed forespore membranes on the mother cell side. Some fluorescent signal is also observed in the mother cell membrane which is expected as SpoIIIAH is a mother cell protein under σ^E control (Blaylock *et al*., 2004) (**Figure 4-15**). SpoIIIAH strains AH103 and AH104 were imaged using SIM microscopy but image reconstruction was poor due to weak fluorescent signals and excessive background noise (data not shown).

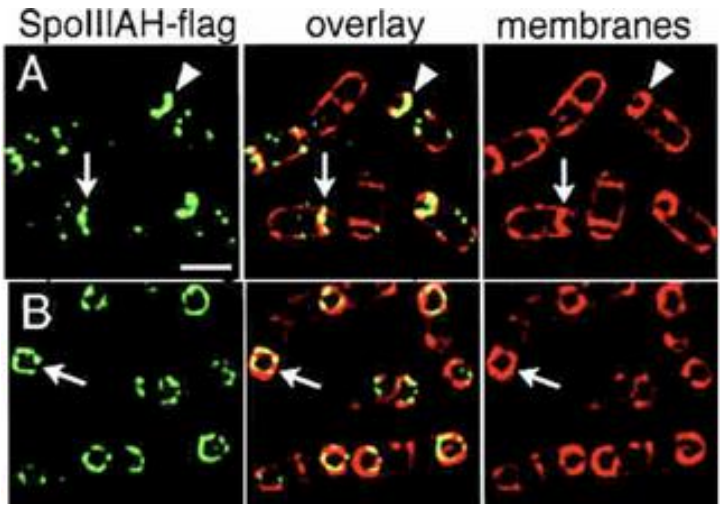


Figure 4-14: Localisation of the mother cell protein SpoIIIAH during engulfment. Immunofluorescence microscopy was used to determine the localisation of SpoIIIAH-Flag (green) and FM-64 which stains membranes red. Images were taken 2hrs (A) and 3hrs (B) after the onset of sporulation. (Taken from Blaylock *et al*.,2004)

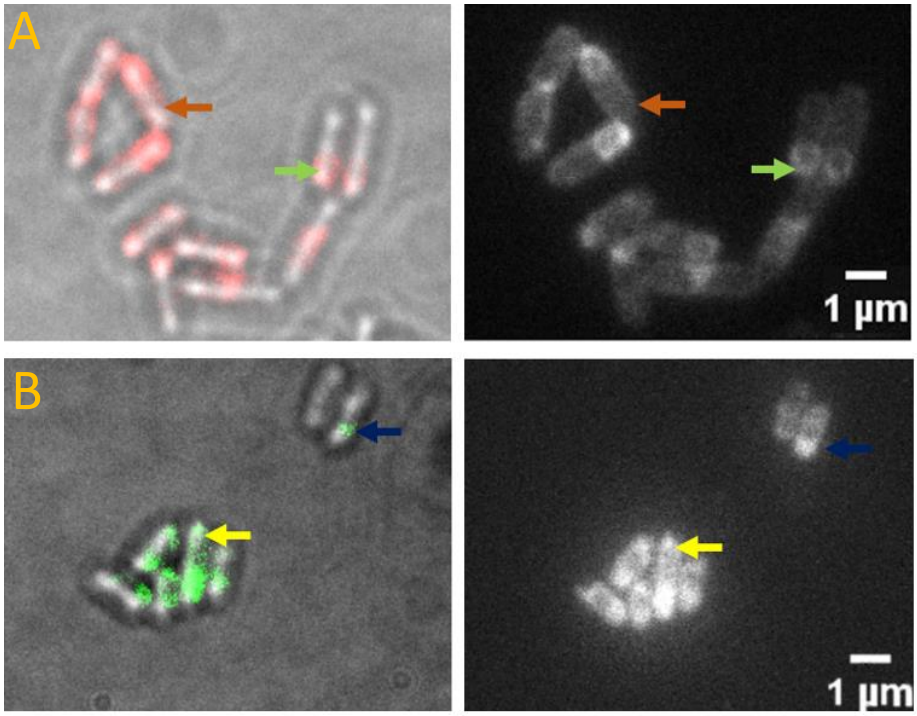


Figure 4-15: Slimfield microscopy imaging of fluorescently labelled SpoIIIAH strains during engulfment. Slimfield microscopy imaging of mCherry SpoIIIAH (A) and mGFP SpoIIIAH (B). Brightfield and fluorescent images superimposed showing the localisation of SpoIIIAH at the engulfing and engulfed forespore membranes on the mother cell side (green and blue arrows). Some SpoIIIAH localises around the mother cell membrane as expected (orange and yellow arrows) Right) Fluorescent images showing the localisation of SpoIIIAH around the forespore membrane on the mother cell side (green and blue arrows) and around the mother cell membranes (orange and yellow arrows).

4.9 Stoichiometry analysis suggests that SpoIIQ is a dimer

The mGFP SpoIIQ delSpoIIQ strain (Strain SS102, **Table 4-1**) was imaged using Slimfield microscopy and molecular stoichiometries were determined using the MATLAB code. When collecting the data, three replicates were taken to ensure the reliability and validity of stoichiometry estimates. Stoichiometry estimates are displayed in Kernel Density Estimation (KDE) plots, and the plots were analysed as described in **Section 3.12.1** of the methods chapter. As shown in **Figure 4-16**, KDE plots of molecular stoichiometries for the three datasets show an array of stoichiometries within the range of 0-50. The three stoichiometry plots show clear and reproducible peaks with the measured stoichiometry appearing as periodic peaks on the stoichiometry distribution plots. The smallest reproducible peak on the KDE plot represents the subunit stoichiometry of a protein and the periodic peaks on the stoichiometry distribution are integer multiples of the subunit stoichiometry. The periodicity spacing is the spacing between the peaks and is often consistent with the subunit stoichiometry **(Chapter 3, section 3.12.1**).

Figure 4-16 shows the SpoIIQ stoichiometry estimates from three independent datasets using Kernel Density Estimation (KDE). The plots show a single fundamental peak around 2 (the smallest peak, highlighted with red arrows) and a periodicity spacing of 2 across the three datasets which is most consistent with SpoIIQ forming dimers which then form part of a larger assembly of SpoIIQ molecules in the Q-AH channel. As the three plots shown in **Figure 4-16** are relatively similar in terms of the single subunit stoichiometry (the smallest peak) and the periodicity spacing, the plots were combined and averaged to produce the stoichiometry plot shown in **Figure 4-17**. The combined plot shows a sharp peak at 2.19± 0.4 and a periodicity spacing of 2 molecules within the margin of error (error being the standard deviation of the three 2-mer peak values shown in **Figure 4-16**) suggesting SpoIIQ forms a dimer.

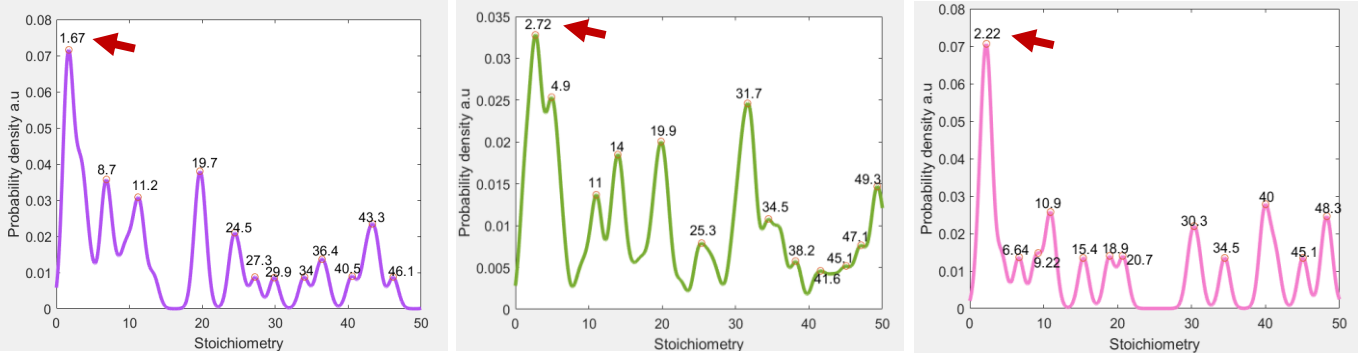


Figure 4-16: Stoichiometry distributions of SpoIIQ using Kernel Density Estimation (KDE). KDE plots of three independent data sets showing the presence of a single fundamental peak around 2 (red arrows) and a periodicity spacing of approximately 2 molecules. These plots suggest that SpoIIQ forms a dimer. . Kernel width 0.7.

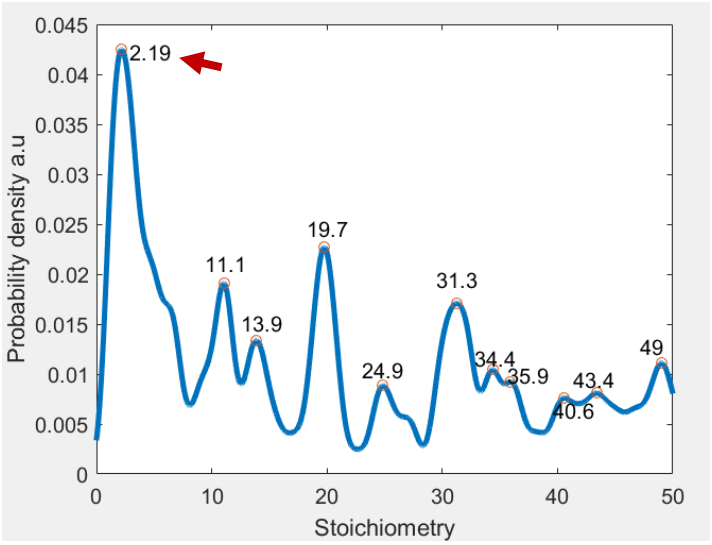


Figure 4-17: Determination of the SpoIIQ stoichiometry by combining and averaging the Kernel Density Estimation plots shown in Figure 4-16. The plot shows a single fundamental peak at 2.19 ± 0.4 (red arrow) and a periodicity spacing of approximately 2 molecules suggesting SpoIIQ forms a dimer. Kernel width 0.7.

4.10 Diffusion Coefficient analysis suggests that SpoIIQ has low mobility

The MATLAB code was next used to calculate the SpoIIQ diffusion coefficient to estimate the mobility of SpoIIQ in living cells in $\mu m^2/s$. Similarly, to molecular stoichiometry analysis, diffusion coefficient determination was carried out on three independent data sets to ensure that patterns and trends in analysis were repeatable and valid. As shown in **Figure 4-18**, KDE plots from the microscopic diffusion coefficient analysis show a range of values from 0 to $2\mu m^2/s$ suggesting the presence of a mixed population with some static and immobile SpoIIQ molecules and other molecules having greater mobility. In **Figure 4-18**, the three plots have a single peak around zero suggesting that most SpoIIQ molecules have low mobility. Similarly, to stoichiometry analysis, the three plots shown in **Figure 4-18** were combined and averaged to produce the single diffusion coefficient KDE plot shown in **Figure 4-19**. This plot shows a range of values from 0 to $2\mu m^2/s$ with the majority of SpoIIQ molecules having a mobility of 0.015 $\mu m^2/s$.

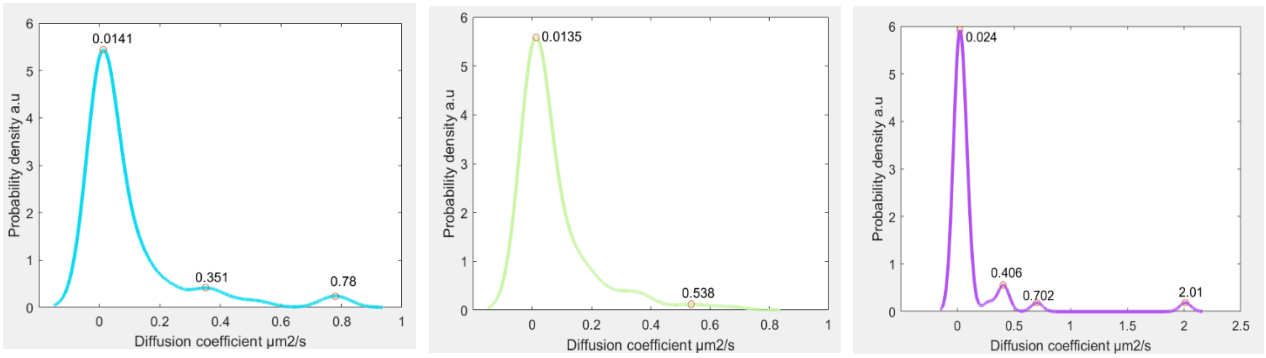


Figure 4-18: SpoIIQ diffusion coefficient determination using Kernel Density Estimation (KDE). KDE plots from three independent data sets showing the presence of a single peak around zero and a range of mobilities from 0 -2 $\mu m^2/s$. These plots suggest that SpoIIQ has low mobility with most molecules in the population having a mobility between 0.01-0.02 $\mu m^2/s$. Kernel width 0.05.

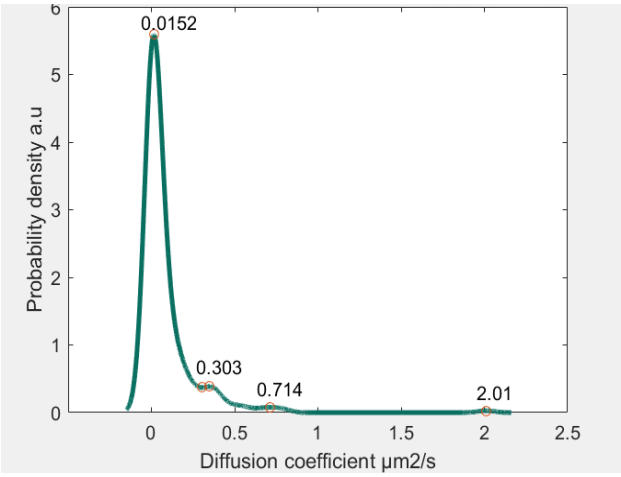


Figure 4-19: Determination of the SpoIIQ diffusion coefficient by combining and averaging the Kernel Density Estimation plots shown in Figure 4-18. The plot shows a single peak at 0.0152 $\mu m^2/s$ and a range of mobilities from 0-2 $\mu m^2/s$. Kernel width 0.05.

4.11 Copy Number analysis suggests an average of 200 SpoIIQ molecules in sporulating *B.subtilis* cells

The MATLAB code was used to calculate the total number of SpoIIQ molecules present in single, sporulating *B. subtilis* cells. Similarly, to stoichiometry and diffusion coefficient analysis, three replicates were analysed and as shown in **Figure 4-20,** a single sporulating cell contains an average of 232± 24 SpoIIQ molecules. As well as the sharp peak values around 200, **Figure 4-20** shows that some cells have a much higher copy number of several thousand, suggesting that there is a large distribution in SpoIIQ expression levels. This finding is supported by the Slimfield image of the mGFP SpoIIQ delSpoIIQ strain (Strain SS102, **Table 4-1**) (**Figure 4-13**) in which some cells are brighter than others suggesting that brighter cells have higher SpoIIQ expression levels. Cells with a copy number of zero have not yet reached the engulfment stage of sporulation in which SpoIIQ is expressed.

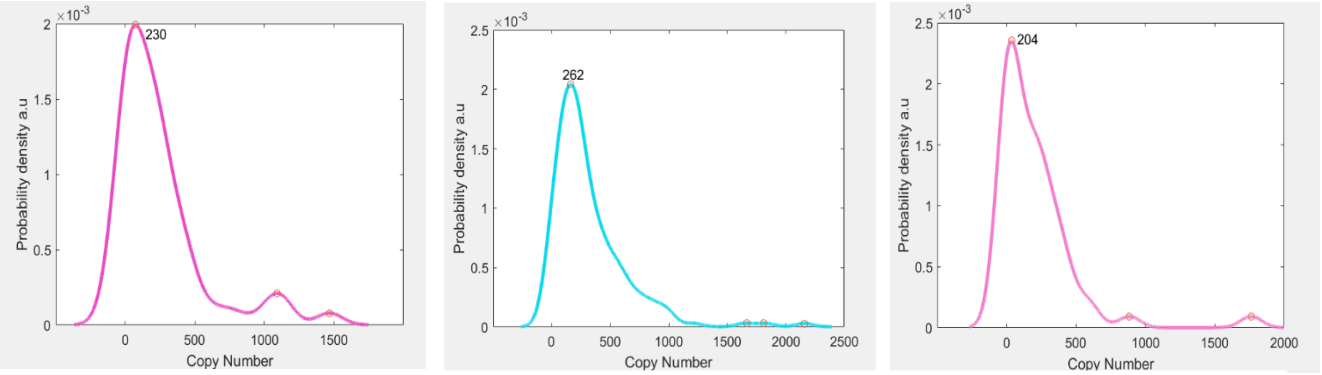


Figure 4-20: Determination of the SpoIIQ copy number using Kernel Density Estimation (KDE). KDE plots from three independent data sets showing the presence of a single peak around 200 and a range of copy numbers from 0-2000. These plots suggest that most cells have an average of 232± 24 SpoIIQ molecules per cell.

4.12 Stoichiometry analysis suggests that SpoIIIAH is a hexamer

The MATLAB code was used to determine the stoichiometry of SpoIIIAH when the protein was labelled with either mCherry or mGFP. Similarly, to SpoIIQ stoichiometry analysis, three replicates were taken to validate the consistency of molecular stoichiometry estimates. **Figure 4-21** shows the SpoIIIAH stoichiometry estimates from independent datasets using Kernel Density Estimation (KDE). Most plots have a single fundamental peak around 6 (highlighted with red arrows) and a periodicity spacing of 3-6 molecules, suggesting that SpoIIIAH predominately forms a hexamer. The plots also show evidence of trimers in some traces as several plots have smaller, shoulder peaks around 3 (highlighted with green arrows).

The replicates were then combined and averaged to produce a single stoichiometry plot for mCherry SpoIIIAH (**Figure 4-22A**) and mGFP SpoIIIAH (**Figure 4-22B**). As shown in **Figure 4-22**, both SpoIIIAH stoichiometry plots show a range of stoichiometries between 0-50 and have a similar fundamental

peaks and periodicity spacing within the margin of error. The single fundamental peaks of the plots are 5.79± 0.3 and 5.21± 0.7 suggesting that within the margin of error, SpoIIIAH predominately forms hexamers which then further assemble into a multi-subunit structure during Q-AH channel formation. As the two stoichiometry plots shown in **Figure 4-22** are similar to one another it suggests that fluorescent protein labelling does not affect the stoichiometry of SpoIIIAH as the two SpoIIIAH fusions behave similarly.

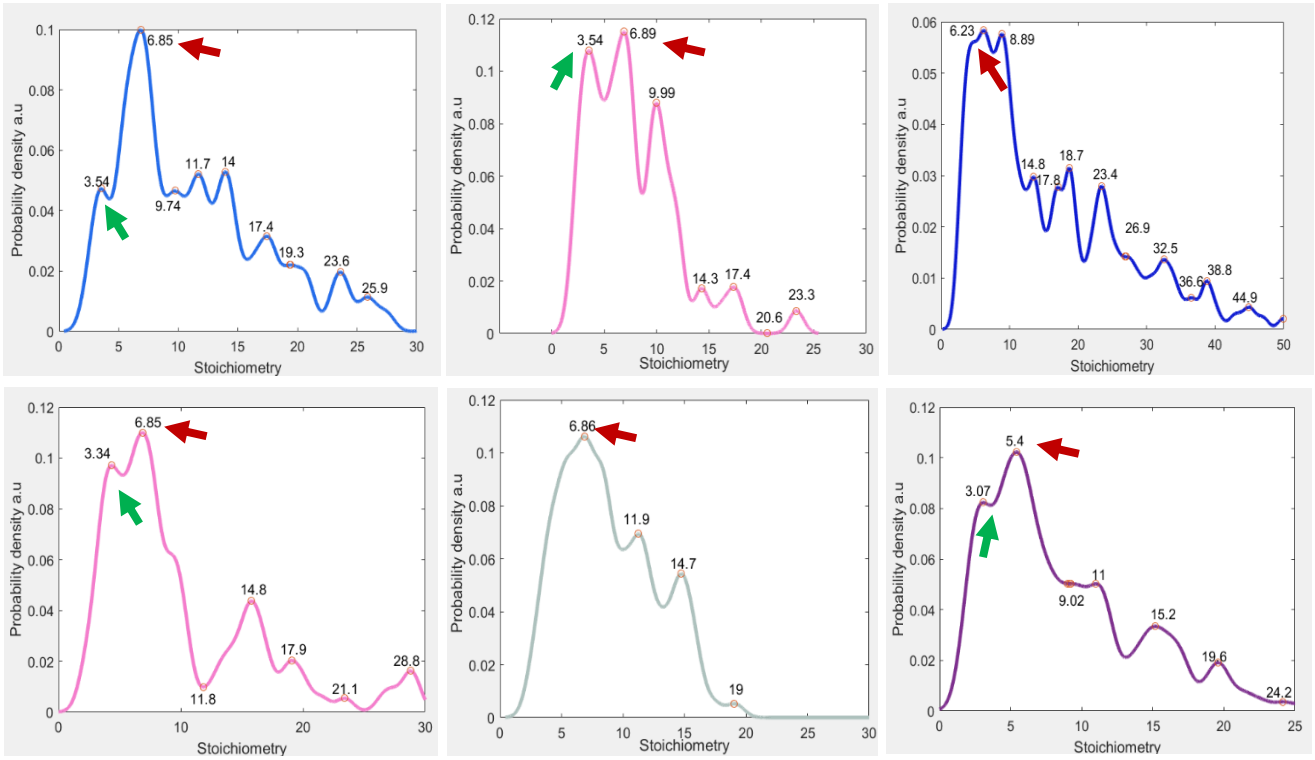
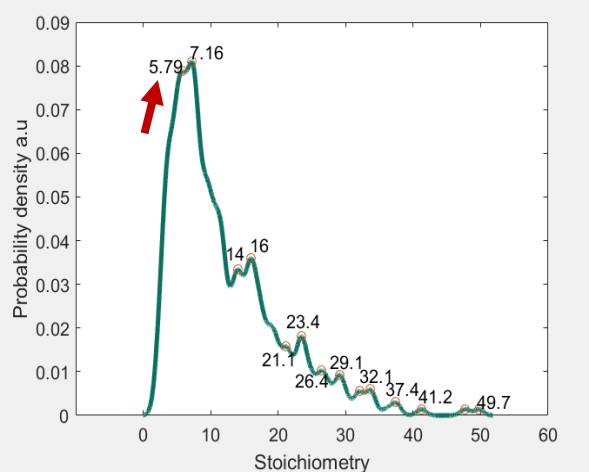


Figure 4-21: Stoichiometry distributions of SpoIIIAH using Kernel Density Estimation (KDE). KDE plots of three independent data sets for mCherry SpoIIIAH (Top) and mGFP SpoIIIAH (Bottom) showing the presence of a single fundamental peak around 6 (red arrows) and smaller shoulder peaks around 3 in some traces (green arrows). This suggests SpoIIIAH predominately forms a hexamer which may be assembled from smaller trimers Kernel width 0.7.



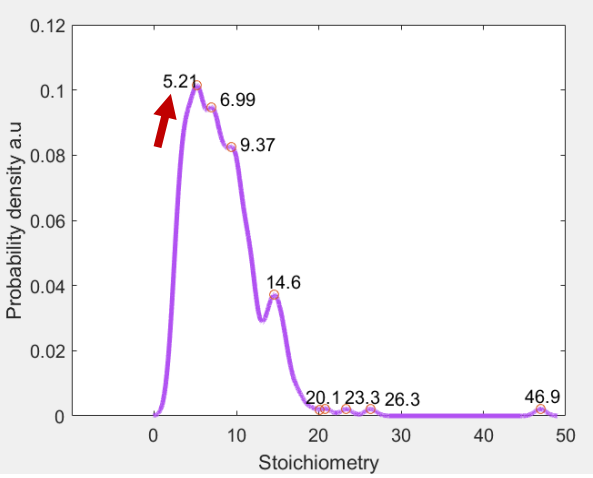


Figure 4-22: Determination of the SpoIIIAH stoichiometry using Kernel Density Estimation (KDE) by combining and averaging the KDE plots shown in Figure 4-21. Left) KDE stoichiometry plot of SpoIIIAH labelled with mCherry. The plot shows a single fundamental peak at 5.79± 0.3 (red arrow) and a periodicity spacing of approximately 3-6 molecules suggesting that SpoIIIAH forms a hexamer that may assemble from smaller trimers. Right) KDE stoichiometry plot of SpoIIIAH labelled with mGFP. The plot shows a single fundamental peak at 5.21± 0.7 (red arrow) and a periodicity spacing of approximately 3-6 molecules suggesting that SpoIIIAH forms a hexamer which may assemble from trimers. Kernel width 0.7.

4.13 Diffusion Coefficient analysis suggests that SpoIIIAH has low mobility

The MATLAB code was next used to calculate the SpoIIIAH diffusion coefficient to estimate the mobility of SpoIIIAH in living cells in $\mu m^2/s$. Similarly, to molecular stoichiometry analysis, diffusion coefficient analysis was carried out on three independent data sets for mCherry SpoIIIAH and mGFP SpoIIIAH (**Figure 4-23**) and individual plots were combined and averaged to produce single diffusion coefficient KDE plots for both SpoIIIAH strains (**Figure 4-24**). As shown in **Figure 4-24**, SpoIIIAH has low mobility as both plots have sharp peaks around zero. The plots also suggest the presence of a mixed population with some SpoIIIAH molecules moving faster than others as the diffusion coefficients range from 0-3 $\mu m^2/s$.

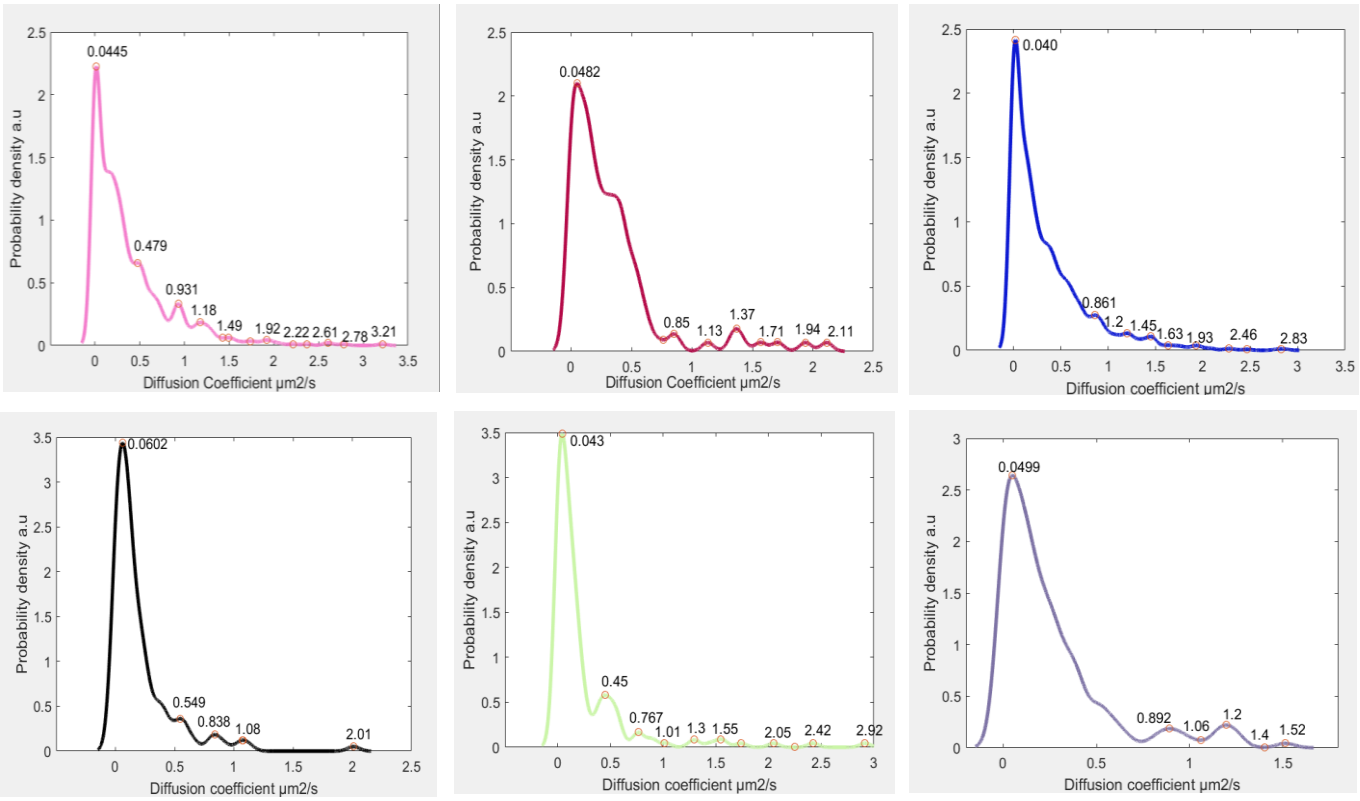


Figure 4-23: SpoIIIAH diffusion coefficient determination using Kernel Density Estimation (KDE). KDE plots of three independent data sets of mCherry SpoIIIAH (Top) and mGFP SpoIIIAH (Bottom) showing the presence of a single peak around 0 and a range of values from 0-3 $\mu m^2/s$. These plots suggest that SpoIIIAH has low mobility in sporulating *B. subtilis* cells. Kernel width 0.05.

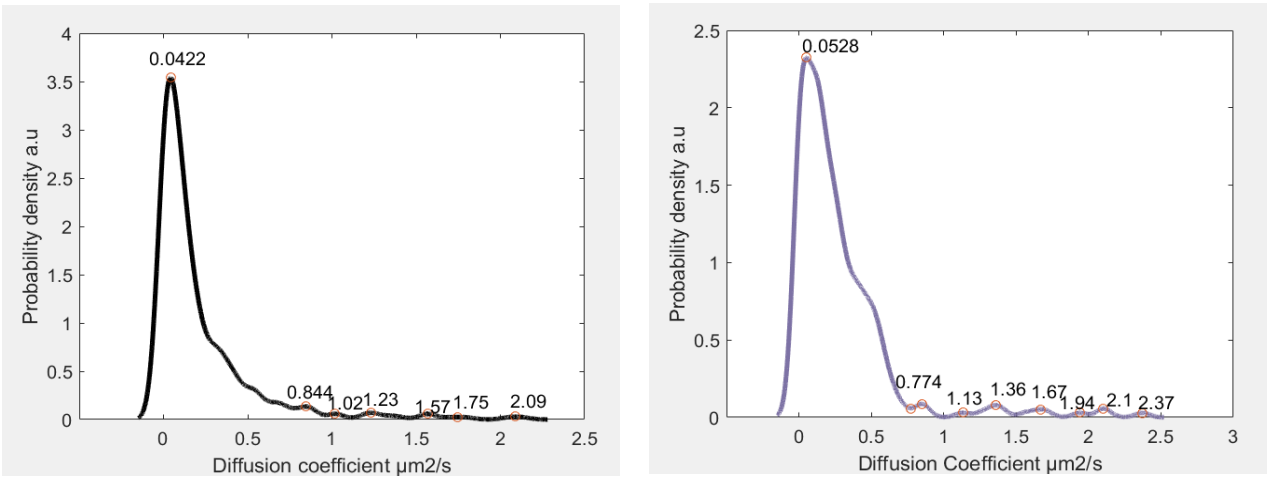


Figure 4-24: Determination of the SpoIIIAH diffusion coefficient by combining and averaging the KDE plots shown in Figure 4-23 Left) Determination of the mCherry SpoIIIAH diffusion coefficient using Kernel Density Estimation (KDE). The plot shows a single peak at 0.0422 $\mu m^2/s$ and a range of values from 0-2 $\mu m^2/s$. Right) Determination of the mGFP SpoIIIAH diffusion coefficient using Kernel Density Estimation (KDE). The plot shows a single peak at 0.0528 $\mu m^2/s$ and a range of values from 0-2.5 $\mu m^2/s$. Kernel width 0.05.

4.14 Copy Number analysis suggests an average of 200 SpoIIIAH molecules in sporulating B.*subtilis* cells

The MATLAB code was used to calculate the total number of SpoIIIAH molecules in single, sporulating *B. subtilis* cells. Similarly, to stoichiometry and diffusion coefficient analysis, three independent datasets were analysed for both the mGFP and mCherry SpoIIIAH strains. As shown in **Figure 4-25**, a single sporulating cell contains an average of 200-300 SpoIIIAH molecules as the mean number of SpoIIIAH molecules per cell was 275±22 in the mGFP SpoIIIAH strain (AH103, **Table 4-1**) and 263±49 in the mCherry SpoIIIAH strain (AH104, **Table 4-1**). **Figure 4-25** also shows some cells with an anomalously high copy number suggesting there is a large distribution in SpoIIIAH expression levels. This finding is supported by the Slimfield images in **Figure 4-15** in which some cells are brighter than others suggesting that brighter cells have higher SpoIIIAH expression levels. Cells with a copy number of zero have not yet reached the engulfment stage of sporulation in which SpoIIIAH is expressed. This analysis is possible despite the low signal observed using SIM microscopy as although SIM is a super-resolution technique its sensitivity is poorer than that of Slimfield microscopy at the low light (low stoichiometry) end, meaning Slimfield microscopy can be used for quantification whereas SIM microscopy cannot.

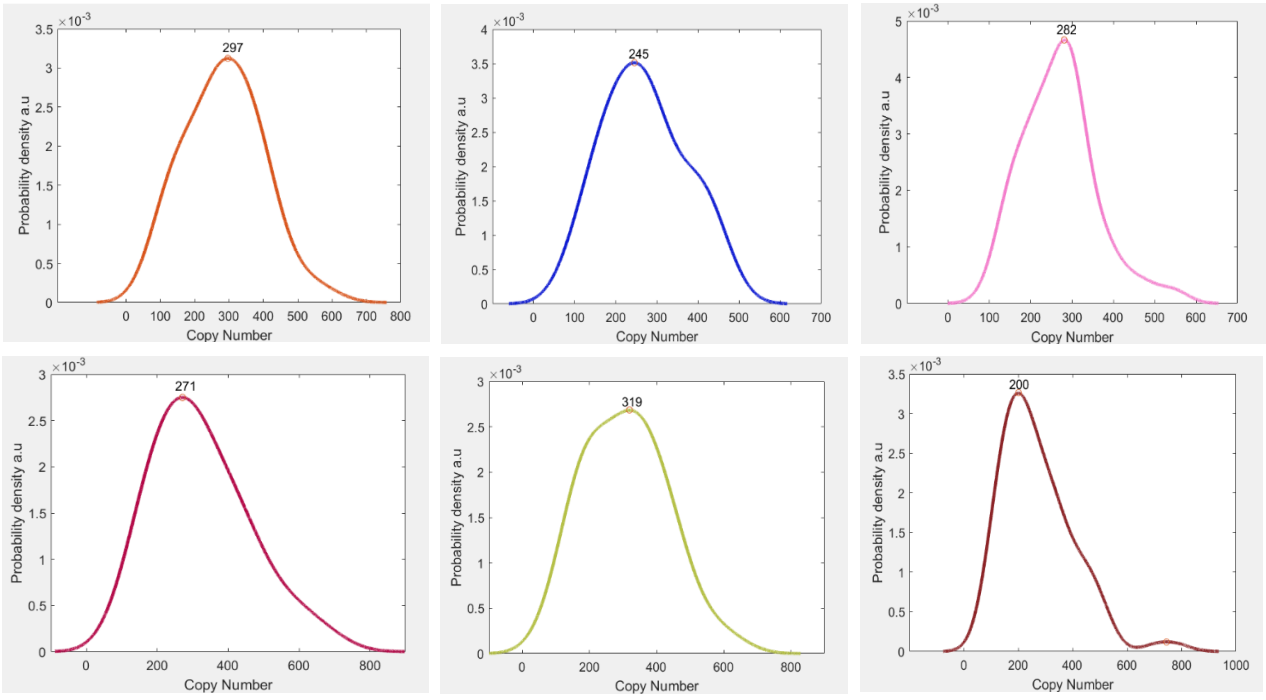


Figure 4-25: Determination of the SpoIIIAH copy number using Kernel Density Estimation (KDE) Top) KDE plots of three independent mGFP SpoIIIAH data sets showing the presence of a single peak around 200 and a range of copy numbers from 0-600 suggesting that most cells have an average of 275±22 SpoIIIAH molecules per cell. Bottom) KDE plots of three independent mCherry SpoIIIAH data sets showing the presence of a single peak around 200-300 and range of copy numbers from 0-800, suggesting most cells have an average of 263±49 SpoIIIAH molecules per cell.

4.15 Discussion

mGFP SpoIIQ strains SS101 and SS102 (**Table 4-1**) were sporulating and showed the expected SpoIIQ localisation when imaged with SIM and Slimfield microscopy. When fluorescently labelled with mGFP, SpoIIQ showed a dynamic localisation pattern during engulfment as the protein initially localised at the asymmetric septum before forming arcs and foci at the engulfing and engulfed forespore membranes. SpoIIIAH strains AH103 and AH104 (**Table 4-1**) were also sporulating and showed the expected SpoIIIAH localisation when imaged with Slimfield microscopy. When imaged with SIM microscopy, image reconstruction was poor due to weak mGFP and mCherry fluorescent signals and excessive background noise. When labelled with mGFP and mCherry, SpoIIIAH localised to the engulfing membrane on the mother cell side with some fluorescence observed in the mother cell. This is to be expected as SpoIIIAH is a mother cell protein that moves to the asymmetric septum via a diffusion and capture mechanism (Blaylock *et al*., 2004) (Doan *et al*., 2005).

The MATLAB code was used to determine the molecular stoichiometries of SpoIIQ and SpoIIIAH to calculate the stoichiometry of both proteins in living cells. MATLAB analysis showed that SpoIIQ forms a dimer and SpoIIIAH predominately forms a hexamer, suggesting that the Q-AH channel is formed from dimers of SpoIIQ and hexamers of SpoIIIAH. This is consistent with both the 12-mer subunit model proposed by *Levidikov et al* and 18-mer subunit model proposed by *Meisner et al* as these models can be assembled from dimers and hexamers (**Figure 4-26**). The 15-mer subunit model proposed by *Meisner et al* cannot be assembled from dimers and hexamers and so the findings are inconsistent with this model (Meisner *et al*.,2012) (Levidikov *et al*.,2012). **Figure 4-26** suggests that assembly of the Q-AH channel is more complex than initially expected due to SpoIIQ and SpoIIIAH having multiple subunits rather than the expected singular monomeric subunit (Meisner *et al*.,2012) (Levidikov *et al*.,2012). SpoIIIAH hexamers may assemble from SpoIIIAH trimers as evidence of trimer formation was found in some traces.

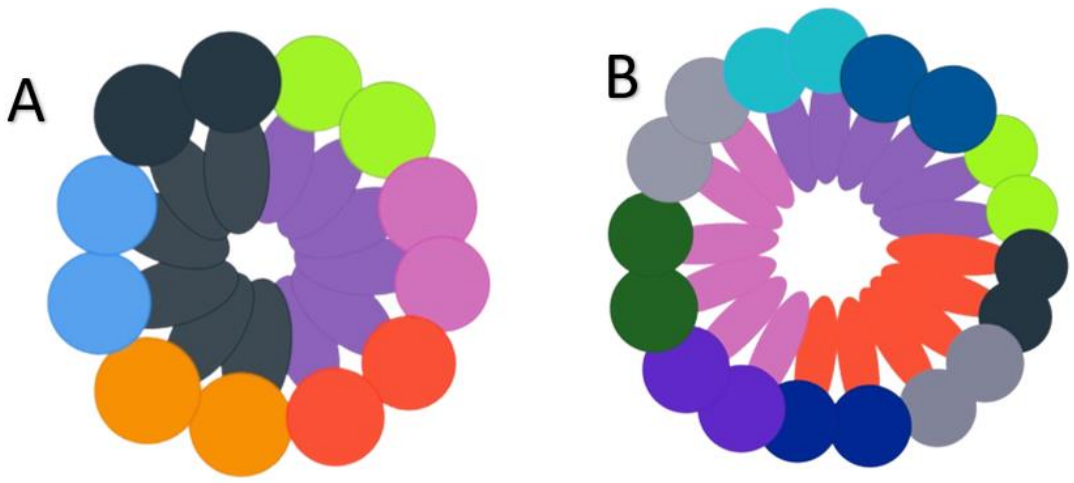


Figure 4-26: Schematic representation of the assembly of the SpoIIQ-SpoIIIAH channel from SpoIIQ dimers and SpoIIIAH hexamers. A) 6 SpoIIQ dimers and 2 SpoIIIAH hexamers are required to assemble the 12-mer subunit model proposed by Levidikov *et al.* B) 9 SpoIIQ dimers and 3 SpoIIIAH hexamers are required to assemble the 18-mer subunit model proposed by *Meisner et al. (*(Meisner *et al*.,2012) (Levidikov *et al*.,2012).

Diffusion coefficient analysis was performed to investigate the mobility of SpoIIQ and SpoIIIAH in live *B. subtilis* cells. SpoIIQ and SpoIIIAH were both found to have low mobility, a finding that might be expected for membrane proteins believed to form a channel. Despite the majority of SpoIIQ and SpoIIIAH molecules being stationary, the Kernel Density Estimation (KDE) plots showed the presence of a mixed population with some molecules moving faster than others. This suggests the presence of turnover with molecules not yet incorporated into the channel complex freely diffusing in the unbound pool. SpoIIIAH was also found to have a slightly higher mobility than SpoIIQ which is to be expected as SpoIIIAH resides in the mother cell membrane before diffusing to the asymmetric septum via diffusion and capture (Blaylock *et al*., 2004) (Doan *et al*., 2005).

Copy number analysis was used to determine the total number of SpoIIQ and SpoIIIAH molecules in single, sporulating *B. subtilis* cells. MATLAB analysis found that on average there are 200-300 molecules of SpoIIQ and SpoIIIAH in sporulating *B. subtilis* cells. Copy number values ranged from 0-800 for SpoIIIAH and 0-2000 for SpoIIQ suggesting that there is a very large distribution in SpoIIQ and SpoIIIAH expression levels in live cells, a finding supported by the Slimfield images which show some cells being exceptionally brighter than others.

4.16 Dual Strain Construction

To determine whether SpoIIQ and SpoIIIAH co-localise in live cells, a dual labelled strain was constructed. In this strain (SS105, **Table 4-2**), SpoIIQ and SpoIIIAH were fused to fluorescent proteins with different spectral characteristics, allowing both proteins to be visualised independently in the same cell. To construct the dual labelled strain, genomic DNA taken from the mCherry SpoIIIAH strain (AH104, **Table 4-2**) constructed by Adam Hughes prior to the start of this work was transformed into the mGFP SpoIIQ strain SS102 (**Table 4-2**). The resulting strain has *mGFP spoIIQ* at the *amyE* locus and *mCherry spoIIIAH* at the native *spoIIIAH* locus. The native *spoIIQ* locus is disrupted by an erythromycin cassette in the dual strain (SS105). The dual strain is resistant to spectinomycin (from the *mGFP spoIIQ* fusion at the *amyE* locus) and chloramphenicol (from the *mCherry spoIIIAH* fusion at the native *spoIIIAH* locus).

Strain Identifier	Genotype
SS102	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ::erm*
AH104	*spoIIIAH:: mCherry-spoIIIAH-camR*
SS105	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ::erm spoIIIAH:: mCherry-spoIIIAH-camR*

Table 4-2 : Strains used to construct the dual labelled strain.

4.16.1 Slimfield Microscopy suggests that SpoIIQ and SpoIIIAH co-localise in live cells.

The dual strain was imaged using Slimfield Microscopy to confirm the robustness of the technique for imaging live *B. subtilis* cells. Imaging of the dual strain confirmed that the strain was sporulating and the localisation of SpoIIQ and SpoIIIAH was not affected by the fluorescent protein labelling of the partner protein. As shown in **Figure 4-27** , mGFP SpoIIQ and mCherry SpoIIIAH co-localise at the sporulation septum and around the forespore membrane. The images suggest the presence of a mixed population as SpoIIQ and SpoIIIAH are not always co-localised and bound to each other in live *B. subtilis* cells. This is shown in **Figure 4-27**, as SpoIIIAH is randomly distributed throughout the mother cell membrane in some cells. Slimfield microscopy imaging suggests that the concentration of SpoIIQ and SpoIIIAH is the greatest at the sporulation septum and forespore membrane, as expected.

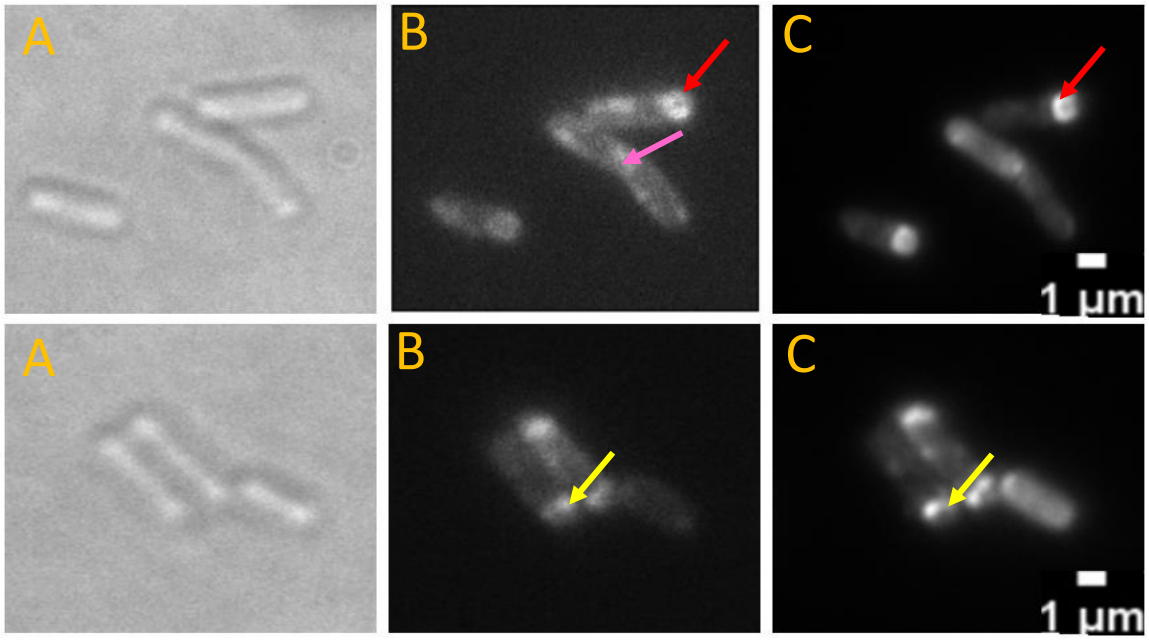


Figure 4-27: Slimfield imaging of the mGFP SpoIIQ mCherry SpoIIIAH dual labelled strain (SS105). A) Brightfield images and Slimfield fluorescent images of the B) red channel and C) green channel showing that SpoIIIAH (B) and SpoIIQ (C) co-localise at the sporulation septum (yellow arrows) and forespore membrane (red arrows). The images suggest that there is a mixed population, and the two proteins are not always co-localised as SpoIIIAH is randomly distributed throughout the mother cell membrane in some cells (pink arrow).

4.16.2 Stoichiometry Analysis suggests that the Q-AH channel acts as a site of recruitment for SpoIIQ dimers and SpoIIIAH hexamers.

A preliminary attempt was made to analyse the Slimfield images using custom built MATLAB code to determine the stoichiometry of SpoIIIAH in the dual labelled strain. The Kernel Density Estimation (KDE) plot shown in **Figure 4-28**, shows the linked (red) and unlinked (blue) stoichiometry of SpoIIIAH. The linked stoichiometry refers to the SpoIIIAH bound to SpoIIQ, and the unlinked stoichiometry refers to the SpoIIIAH not in complex with SpoIIQ. Analysis of the linked and unlinked stoichiometries give an indication of how the Q-AH channel complex is assembled. If the co-localised (linked) and not co-localised (unlinked) stoichiometry of SpoIIIAH are similar to one another, it suggests that SpoIIQ and SpoIIIAH may form pre-assembled, functional units which are then used to form the Q-AH channel. If the co-localised (linked) and not co-localised (unlinked) SpoIIIAH stoichiometry are different to one another, the forming channel may act as a site of recruitment and one protein may act as a scaffold and recruit the other protein. **Figure 4-28** shows that the peak positions between the linked and unlinked stoichiometries for SpoIIIAH remain constant but the positions are shifted towards higher apparent stoichiometries for linked stoichiometry. Higher linked stoichiometry suggests that the channel acts as an assembly point and that SpoIIQ recruits SpoIIIAH to the channel. This would mean that the SpoIIIAH stoichiometry builds up as the channel is assembled.

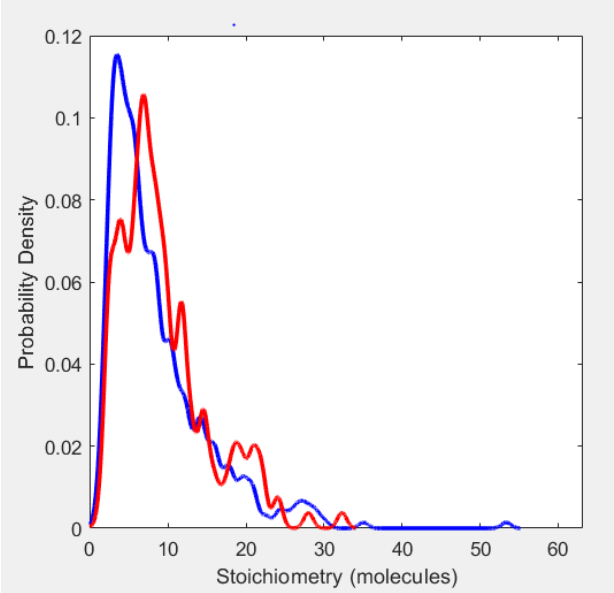


Figure 4-28: Linked and unlinked stoichiometry distributions of SpoIIIAH shown using Kernel Density Estimation (KDE). The plot shows the linked stoichiometry (red) and unlinked stoichiometry (blue). Linked stoichiometry refers to SpoIIIAH co-localised with SpoIIQ and unlinked stoichiometry refers to the SpoIIIAH that is not co-localised with SpoIIQ. The plot shows similar peak positions for linked and unlinked stoichiometries and a shift towards higher apparent stoichiometries for the linked SpoIIIAH stoichiometry. Kernel width 0.7.

Chapter 5- Stoichiometry, Diffusion Coefficient and Copy Number Analysis of SpoIIQ and SpoIIIAH in mutant backgrounds

5.1 Introduction

SpoIIQ and SpoIIIAH are two proteins required for engulfment and membrane migration in *B. subtilis.* (Londono-Vallejo *et al*.,1997). SpoIIQ is a forespore specific protein under σ^F control whereas SpoIIIAH is a mother cell protein under σ^E control (Londono-Vallejo *et al*.,1997). Both proteins use their N-terminal transmembrane domains to attach to the cell membrane and interact with one another via their C-terminal extracellular domains (Meisner *et al*.,2012). As described in Chapter 4, SpoIIQ and SpoIIIAH show a dynamic localisation pattern throughout engulfment and are important in σ^G activation (Rodrigues *et al*.,2013). Slimfield microscopy with mGFP spoIIQ fusions has shown that in the presence of SpoIIIAH, SpoIIQ initially localises at the sporulation septum and then forms arcs and foci at the forespore/mother cell interface.

SpoIIIAH is the final gene in the *spoIIIAA* operon which consists of the eight genes *spoIIIAA-spoIIIAH* (Blaylock *et al*., 2004). Slimfield microscopy with mGFP and mCherry SpoIIIAH fusions has shown that in the presence of SpoIIQ, SpoIIIAH is initially randomly distributed throughout the mother cell membrane. In cells that have initiated engulfment, SpoIIIAH localises in discrete foci and as engulfment proceeds, the protein tracks the engulfing mother cell membrane as it moves around the forespore.

As SpoIIQ and SpoIIIAH are believed to interact with one another via their C-terminal extracellular domains, this chapter aims to investigate how the localisation of one protein is affected by the absence of the partner protein. In this chapter, SpoIIQ and SpoIIIAH were fused to fluorescent proteins and single fluorescently labelled foci of both proteins were tracked to super-resolution precision in live *B. subtilis* cells. The chapter illustrates how stoichiometry, diffusion coefficient, and copy number values change in a mutant background when the partner protein of either SpoIIQ or SpoIIIAH is missing. Single subunit stoichiometry analysis of SpoIIQ and SpoIIIAH suggests that both proteins form different fundamental subunits in a mutant background compared to a wild-type background. Copy number analysis suggests that the number of molecules of SpoIIIAH decreases in a mutant SpoIIQ background, illustrating the possibility of downregulation in SpoIIIAH expression in the absence of SpoIIQ. The mobility of SpoIIQ and SpoIIIAH in mutant backgrounds was measured through calculation of the microscopic diffusion coefficient and suggested that both proteins have low mobility in the absence of their partner protein.

5.2 Mutant Strain Construction

mGFP SpoIIQ delSpoIIIAH (SS103, **Table 5-1**) and mCherry SpoIIIAH delSpoIIQ (SS104, **Table 5-1**) are the two mutant strains imaged and analysed in this thesis chapter. Both strains were constructed by the Barak laboratory in Bratislava, Slovakia, and were used to analyse the effects of perturbing the interactions between SpoIIQ and SpoIIIAH. To construct strain SS103, the pSG1154-mGFP spoIIQ plasmid described in Chapter 4 was transformed into *B. subtilis* strain IB714 (**Table 5-1**) which contains an erythromycin resistance cassette at the *amyE* locus. This allowed the *mGFP spoIIQ* fusion to be integrated into the *B. subtilis* chromosome at the a*myE* locus, resulting in a switch in antibiotic resistance from erythromycin to spectinomycin (**Figure 5-1**). The resulting strain has *mGFP spoIIQ* at the *amyE* locus and wildtype 'dark' *spoIIQ* at the native *spoIIQ* locus. delspoIIIAH genomic DNA isolated from the SpoIIIAH null mutant strain BKE24360 was next transformed into this strain resulting in the mGFP SpoIIQ delSpoIIIAH mutant strain. This strain has resistance to spectinomycin (from the mGFP spoIIQ fusion at *amyE*) and resistance to erythromycin (from the delspoIIIAH genomic DNA).

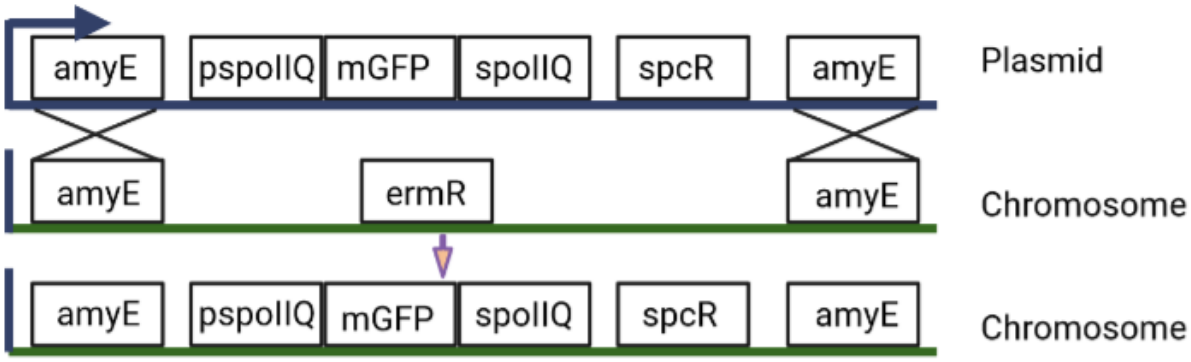


Figure 5-1: Diagram showing how the *mGFP spoIIQ* fusion integrates into the *B. subtilis* chromosome at the *amyE* locus via a double cross-over event. There is a switch in antibiotic resistance from erythromycin to spectinomycin.

To construct the mCherry SpoIIIAH delSpoIIQ strain, the mCherry-spoIIIAH pAHIIIAH04 plasmid prepared by Adam Hughes prior to the start of this work was transformed into the spoIIQ null mutant strain BKE36550 (*Bacillus subtilis* Stock Centre) in which the *spoIIQ* gene is disrupted by an erythromycin cassette. The resulting strain, SS104, has resistance to chloramphenicol (from the mCherry spoIIIAH fusion at the native *spoIIIAH* locus) and resistance to erythromycin (from the delspoIIQ genomic DNA).

Strain Identifier	Genotype
IB714	*amyE::erm*
BKE36550	*spoIIQ:: erm trpC2*
BKE24360	*spoIIIAH:: erm trpC2*
SS103	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ:: spoIIQ spoIIIAH::erm*
SS104	*IB714 spoIIIAH:: mCherry-spoIIIAH-camR spoIIQ:: erm*

Table 5-1: List of strains imaged and analysed in thesis chapter 5.

5.3 SIM Microscopy suggests that SpoIIQ is delocalised in the absence of SpoIIIAH

Fluorescence microscopy imaging of GFP SpoIIQ fusions has shown that in the absence of mother cell gene expression, SpoIIQ is randomly distributed throughout the forespore membrane. The protein fails to localise at the sporulation septum and does not form arcs/foci at the engulfing membrane on the forespore side (**Figure 5-2**) (Rodrigues *et al*.,2013). The protein required for the retention of SpoIIQ at the sporulation septum is currently unknown, but it is believed that the catalytic activities of the SpoIID, P, M (DMP complex) allow a mother cell protein under σ^E control to anchor SpoIIQ into the septal membrane (Rodrigues *et al*.,2013).To establish whether SpoIIIAH deletion affected sporulation and the localisation of mGFP SpoIIQ, the mGFP SpoIIQ delSpoIIIAH strain (SS103) was imaged using Structured Illumination Microscopy (SIM). SIM microscopy imaging showed that SpoIIQ is delocalised in the absence of SpoIIIAH as without SpoIIIAH, SpoIIQ is distributed throughout the forespore and fails to localise at the sporulation septum. **Figure 5-3** therefore suggests that SpoIIQ forms arcs and foci around the engulfing forespore membrane in a SpoIIIAH dependent manner. The strong mGFP signal visualised in **Figure 5-3** suggests that the mGFP spoIIQ delspoIIIAH strain is functional and ready to use in subsequent steps.

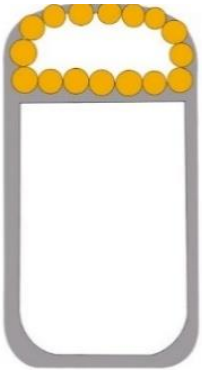


Figure 5-2: Cartoon representation showing SpoIIQ localisation (orange circles) in the absence of mother cell gene expression. SpoIIQ is distributed throughout the forespore membrane and fails to localise at the sporulation septum or in arcs/foci at the engulfing forespore membrane. In the absence of mother cell gene engulfment is blocked.

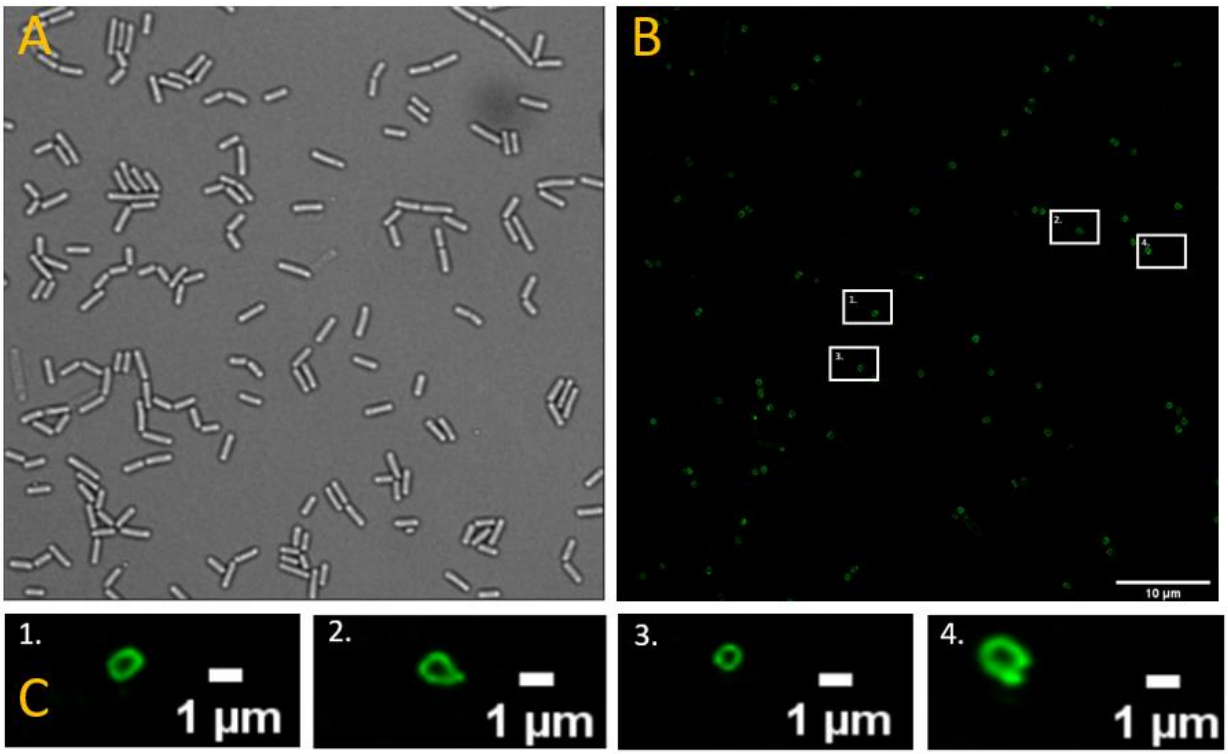


Figure 5-3: SIM microscopy imaging of mGFP SpoIIQ delSpoIIIAH. A) Brightfield image showing live *B. subtilis* cells. B) SIM microscopy fluorescent image showing strong fluorescent signals and the delocalisation of SpoIIQ in the absence of SpoIIIAH. C) Close-up fluorescent images of mGFP spoIIQ delspoIIIAH showing that in the absence of SpoIIIAH, SpoIIQ is localised throughout the forespore membrane and fails to localise at the sporulation septum or in arcs/foci at the engulfing forespore membrane.

5.4 Slimfield Microscopy suggests that SpoIIQ is delocalised in the absence of SpoIIIAH

The mGFP SpoIIQ delSpoIIIAH strain described above was imaged using Slimfield microscopy to confirm the robustness of the technique for imaging live *B. subtilis* cells. For Slimfield microscopy, an overnight culture of the mGFP spoIIQ delspoIIIAH strain was diluted in Difco Sporulation Medium (DSM) and grown for 8 hours to reach the engulfment stage of sporulation before being harvested and spotted onto 1% agarose pads. As shown in **Figure 5-4**, Slimfield microscopy yields similar results to SIM microscopy and shows that in the absence of SpoIIIAH, SpoIIQ is distributed throughout the forespore and not present at the sporulation septum in any cells. SpoIIQ is unable to form arcs/foci at the engulfing forespore membrane in the absence of SpoIIIAH, suggesting that SpoIIIAH is important for the correct localisation of SpoIIQ. In **Figure 5-4**, the fluorescent signal is consistent with the presence of mGFP but the mGFP signal is delocalised.

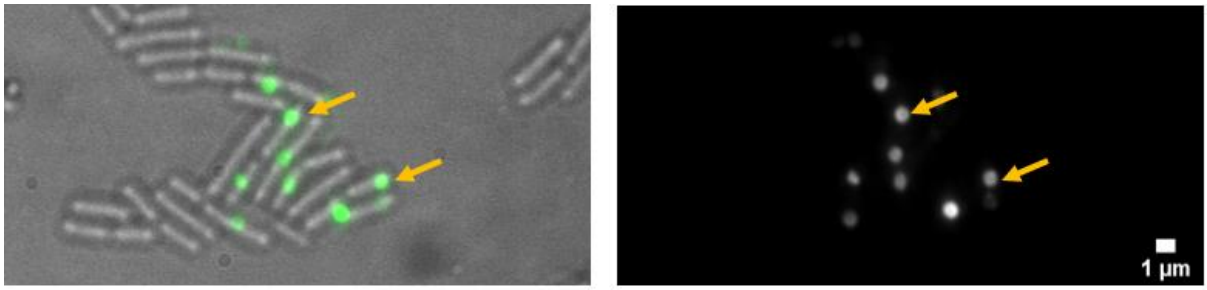


Figure 5-4: Slimfield microscopy imaging of mGFP spoIIQ delspoIIIAH. Left) Brightfield and fluorescent images superimposed showing that SpoIIQ localises around the forespore in the absence of SpoIIIAH (orange arrows). Right) Fluorescent image showing mGFP spoIIQ around the forespore (orange arrows) and not at the sporulation septum or engulfing forespore membrane in arcs/foci, which is expected in the presence of SpoIIIAH.

5.5 Slimfield Microscopy suggests that SpoIIIAH is delocalised in the absence of SpoIIQ

Immunofluorescence microscopy has been used to show that SpoIIIAH is randomly distributed throughout the mother cell membrane in the absence of SpoIIQ (**Figure 5-5**) suggesting that SpoIIQ is important for the appropriate localisation of SpoIIIAH (Blaylock *et al*.,2004). The mCherry spoIIIAH delspoIIQ (SS104) strain was imaged using Slimfield microscopy to confirm the robustness of the technique for imaging live *B.subtilis.* For Slimfield microscopy, an overnight culture of the mCherry SpoIIIAH delSpoIIQ strain was diluted in Difco Sporulation Medium (DSM) and grown for 8 hours to reach the engulfment stage of sporulation before being harvested and spotted onto 1% agarose pads. As shown in **Figure 5-6**, Slimfield microscopy imaging shows that SpoIIIAH is delocalised in the absence of SpoIIQ as the mCherry signal is weak and dispersed throughout the mother cell. This suggests that SpoIIIAH localises to the engulfing membrane on the mother cell side in a SpoIIQ dependent manner. The mCherry SpoIIIAH delSpoIIQ strain was imaged using SIM microscopy but image reconstruction was poor due to weak fluorescent signals and excessive background noise (data not shown).

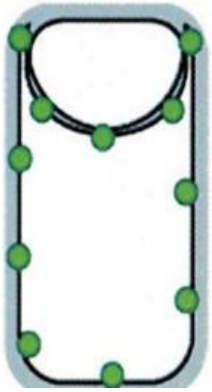


Figure 5-5: Cartoon representation of SpoIIIAH localisation in cells without SpoIIQ visualised by immunofluorescence microscopy. In the absence of SpoIIQ, SpoIIIAH is dispersed throughout the mother cell membrane (Adapted from Blaylock *et al*.,2004)

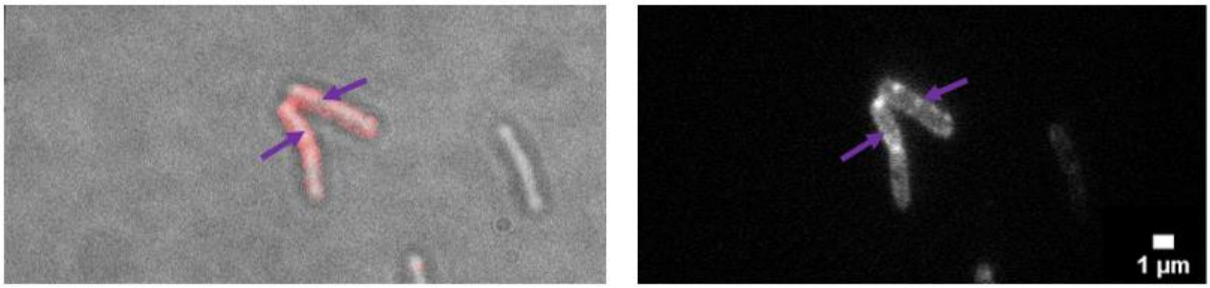


Figure 5-6: Slimfield microscopy imaging of mCherry spoIIIAH delspoIIQ. Left) Brightfield and fluorescent images superimposed showing SpoIIIAH localised around the mother cell membrane in the absence of SpoIIQ (purple arrows). Right) Fluorescent image showing mCherry spoIIIAH randomly dispersed throughout the mother cell membrane (purple arrows) and not at the sporulation septum or around the engulfing mother cell membrane as expected in the presence of SpoIIQ.

5.6 Stoichiometry Analysis suggests SpoIIQ oligomerises differently in the absence of SpoIIIAH

The mGFP SpoIIQ delSpoIIIAH strain (SS103) was imaged using Slimfield Microscopy and molecular stoichiometries were determined using the MATLAB code. When collecting the data three independent replicates were taken to ensure the reliability of any stoichiometry estimates and Kernel Density Estimation (KDE) plots were analysed as described in **Section 3.12.1** of the methods chapter. As shown in **Figure 5-7,** Kernel Density Estimation (KDE) plots for molecular stoichiometries show a range of stoichiometries within the range of 0-25. Similarly, to the stoichiometry analysis performed in Chapter 4, the three mGFP SpoIIQ delSpoIIIAH plots show clear and reproducible peaks with the measured stoichiometry appearing as a periodic ripple on the stoichiometry plots. The spacing between the peaks is known as the periodicity spacing and the periodic peaks on the stoichiometry distribution are integer multiples of the subunit stoichiometry within measurement error. In **Figure 5-7,** mGFP SpoIIQ has a periodicity of 5-6 across the three datasets when SpoIIIAH is absent suggesting that SpoIIQ oligomerises differently in the absence of its partner protein. This is because in the presence of SpoIIIAH, SpoIIQ forms a dimer (Chapter 4) whereas in the absence of SpoIIIAH, SpoIIQ forms either a pentamer or a hexamer. As the three plots in **Figure 5-7** are similar in terms of their single subunit stoichiometry (the smallest peak) and periodicity spacing, the plots were combined and averaged to produce the stoichiometry plot shown in **Figure 5-8**. The combined plot shows a sharp peak at 5.23±0.5 and a periodicity spacing of 5-6 molecules within the margin of error (error being the standard deviation of the 3 peak values shown in **Figure 5-7**) suggesting that SpoIIQ forms either a pentamer or a hexamer in the absence of SpoIIIAH.

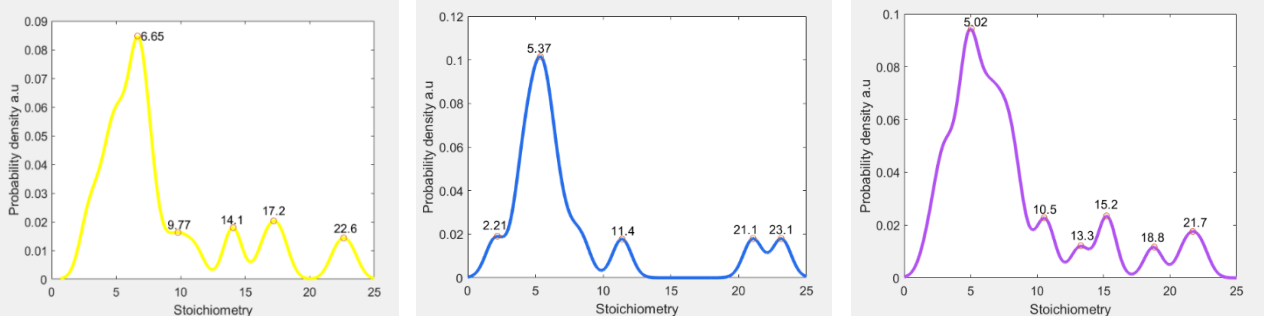


Figure 5-7: Stoichiometry distributions of SpoIIQ in the absence of SpoIIIAH shown using Kernel density estimation (KDE). KDE plots of three independent mGFP SpoIIQ delSpoIIIAH data sets showing the presence of a single fundamental peak between 5-6 and a periodicity spacing of approximately 5-6 molecules. These plots suggest SpoIIQ forms either a pentamer or hexamer in the absence of SpoIIIAH. Kernel width 0.7

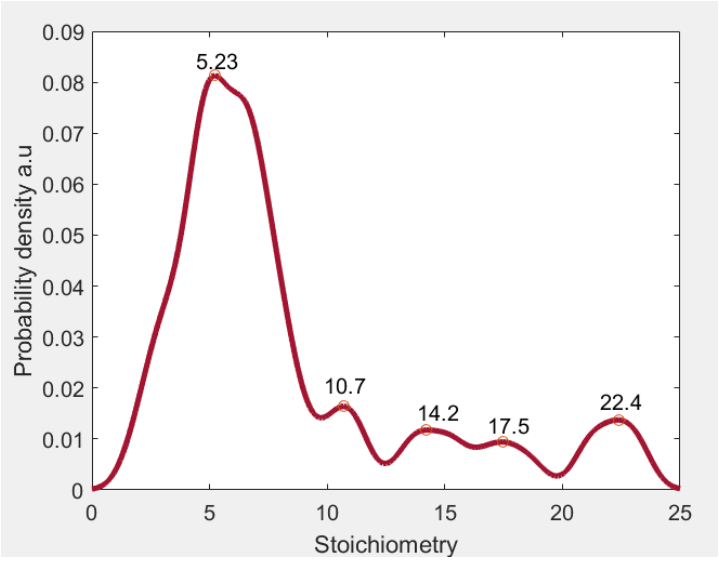


Figure 5-8: Determination of SpoIIQ stoichiometry in the absence of SpoIIIAH by combining and averaging the Kernel Density Estimation plots (KDE) shown in Figure 5-7. The plot shows a single fundamental peak at 5.23±0.5 and a periodicity spacing of approximately 5-6 molecules suggesting that SpoIIQ forms either a pentamer or a hexamer in the absence of partner protein, SpoIIIAH. Kernel width 0.7.

5.7 Diffusion Coefficient analysis suggests that SpoIIQ has low mobility in the absence of SpoIIIAH

The MATLAB code was used to determine the diffusion coefficient of SpoIIQ in the absence of SpoIIIAH, to determine whether the mobility of SpoIIQ changes in a mutant background. As for the molecular stoichiometry analysis, diffusion coefficient determination was carried out on three independent data sets to ensure the reliability and validity of diffusion coefficient estimates. As shown in **Figure 5-9,** KDE plots from microscopic diffusion coefficient analysis show a range of values from 0-1.1 $\mu m^2/s$ suggesting that some SpoIIQ molecules are more mobile than others. All 3 plots in **Figure 5-9** have a peak around 0.01 $\mu m^2/s$ suggesting that most SpoIIQ molecules in the population are essentially immobile. Similarly, to the stoichiometry analysis above, the three plots in **Figure 5-9** were combined and averaged to produce the single diffusion coefficient KDE plot shown in **Figure 5-**

10. **Figure 5-10** shows a range of values from 0- 1.4µm^2/s with the majority of SpoIIQ molecules having a mobility of 0.016 µm^2/s. These results suggest that SpoIIQ has a low mobility in the presence and absence of SpoIIIAH as most SpoIIQ molecules had a mobility of 0.015 µm^2/s. in the presence of SpoIIIAH (Chapter 4).

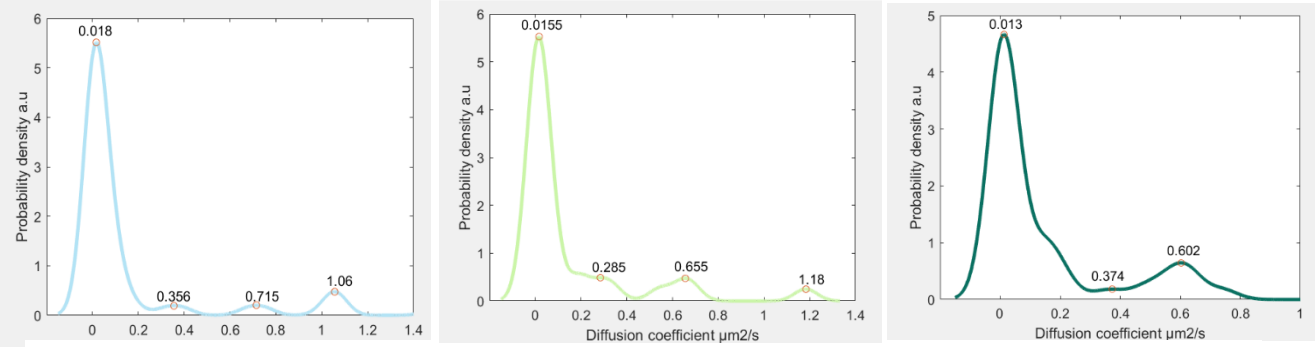


Figure 5-9: SpoIIQ diffusion coefficient determination using Kernel Density Estimation (KDE) in the absence of SpoIIIAH. KDE plots of three independent mGFP spoIIQ delspoIIIAH data sets showing the presence of a single peak around 0.01 and a range of mobilities from 0- 1.1 µm^2/s. These plots suggest that SpoIIQ is largely immobile with most molecules in the population having a low mobility of 0.01µm^2/s. Kernel width 0.05.

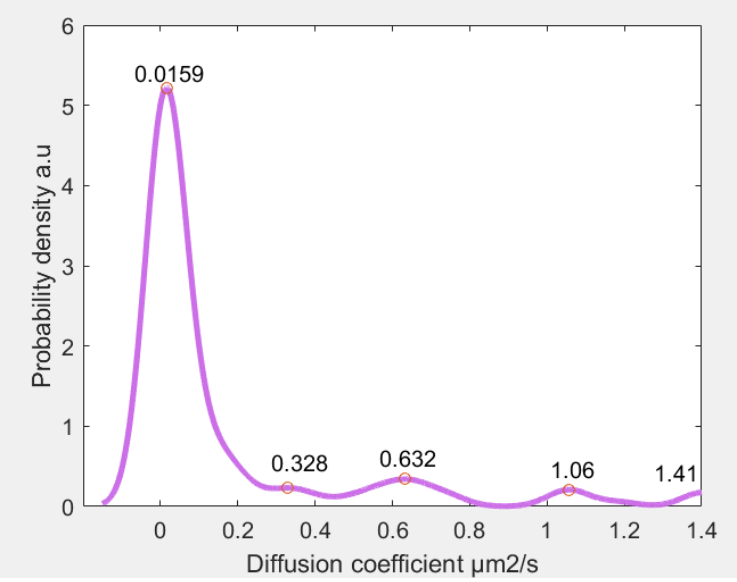


Figure 5-10: Determination of the SpoIIQ diffusion coefficient in the absence of SpoIIIAH by combining and averaging the KDE plots shown in Figure 5-9. The plot shows a single peak at 0.0159 µm^2/s and a range of values from 0-1.4 µm^2/s. Kernel width 0.05.

5.8 Copy Number analysis suggests an average of 200 SpoIIQ molecules in the absence of SpoIIIAH

The MATLAB code was used to calculate the total number of SpoIIQ molecules present in single *B. subtilis* cells in the absence of SpoIIIAH. Similarly, to stoichiometry and diffusion coefficient analysis, three replicates were analysed and as shown in **Figure 5-11,** a single cell contains an average of 228±67 SpoIIQ molecules when SpoIIIAH is missing. As well as the sharp peak values around 200, **Figure 5-11** shows that some cells have an anomalously high copy number of several thousand

suggesting there is a large distribution in SpoIIQ expression levels. This finding is supported by the Slimfield image of the mGFP SpoIIQ delSpoIIIAH strain in **Figure 5-12** as some cells in this raw image are exceptionally brighter than others suggesting that brighter cells have higher SpoIIQ expression levels. These results suggest that the SpoIIQ copy number is not influenced by the presence or absence of SpoIIIAH, as in the presence of SpoIIIAH a single cell contains 232± 24 SpoIIQ molecules (Chapter 4) whereas in the absence of SpoIIIAH a single cell contains an average of 228±67 SpoIIQ molecules.

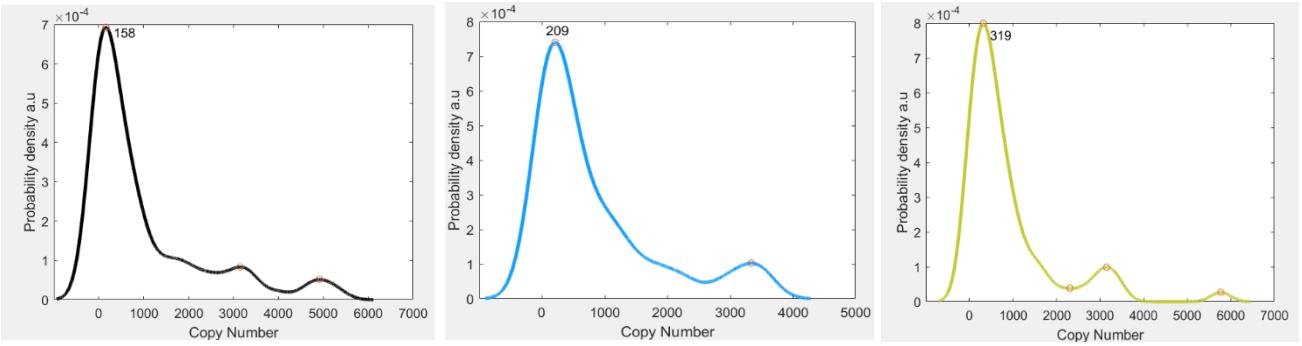


Figure 5-11: Determination of the SpoIIQ copy number using Kernel Density Estimation (KDE) in the absence of SpoIIIAH. KDE plots of three independent mGFP spoIIQ delspoIIIAH data sets showing the presence of a single peak around 200-300 and a range of copy numbers from 0 to several thousand, suggesting most cells have an average of 228±67 SpoIIQ molecules per cell.

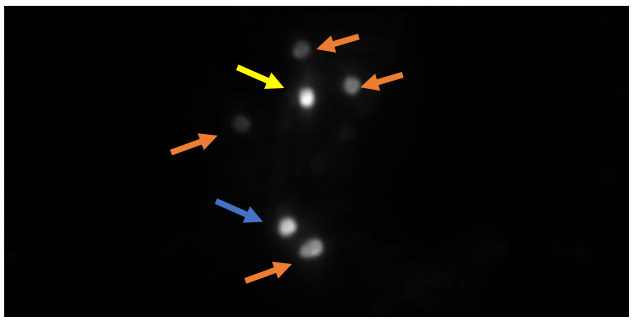


Figure 5-12: Slimfield microscopy image of the mGFP SpoIIQ delSpoIIIAH strain showing cells with different SpoIIQ expression levels. Exceptionally bright cells (yellow arrow) are likely to have 4000-6000 SpoIIQ molecules. Moderately bright cells (blue arrow) are likely to have around 1000-2000 SpoIIQ molecules. Most cells are not very bright (orange arrow) and have around 200-300 SpoIIQ molecules.

5.9 Stoichiometry analysis suggests that SpoIIIAH oligomerises differently in the absence of SpoIIQ

The mCherry SpoIIIAH delSpoIIQ strain (SS104), was imaged using Slimfield Microscopy and molecular stoichiometries were determined using the MATLAB code. Similarly, to the analysis of the mGFP SpoIIQ delSpoIIIAH strain, three replicates were analysed to validate the consistency of molecular stoichiometry estimates (**Figure 5-13**). In **Figure 5-13**, all three plots show a single fundamental peak around 3 suggesting that SpoIIIAH forms a trimer in the absence of SpoIIQ. The

three replicates were then combined and averaged to produce the single stoichiometry plot in **Figure 5-14**. In **Figure 5-14,** the KDE plot shows a range of stoichiometries from 0-30. There is a fundamental peak at 3.73±0.2 and a periodicity spacing of approximately 3 molecules, suggesting that SpoIIIAH forms a different fundamental subunit in the absence of its partner protein. This is concluded as SpoIIIAH forms a hexamer in the presence of SpoIIQ (Chapter 4) and a trimer in the absence of SpoIIQ (**Figure 5-14**).

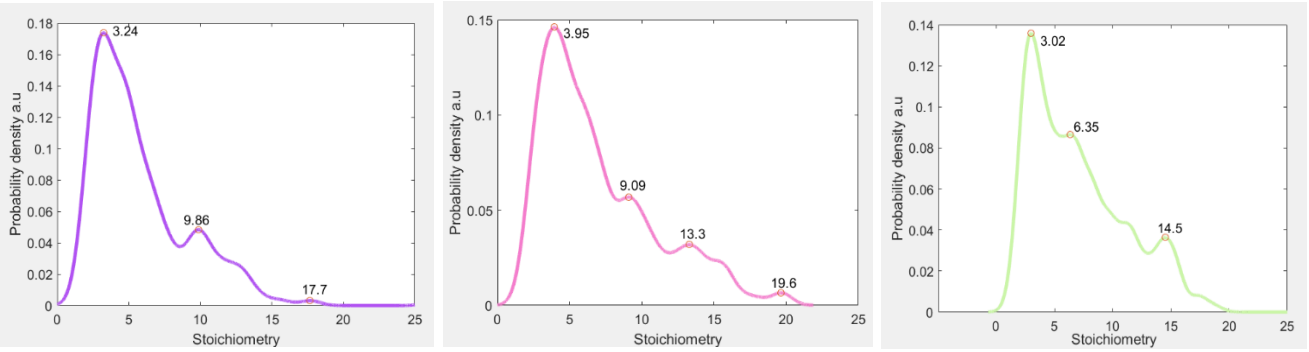


Figure 5-13: Stoichiometry distributions of SpoIIIAH in the absence of SpoIIQ shown using Kernel Density Estimation (KDE). KDE plots of three independent mCherry SpoIIIAH delSpoIIQ data sets showing the presence of a single fundamental peak around 3 suggesting that SpoIIIAH forms a trimer in the absence of SpoIIQ. Kernel width 0.7.

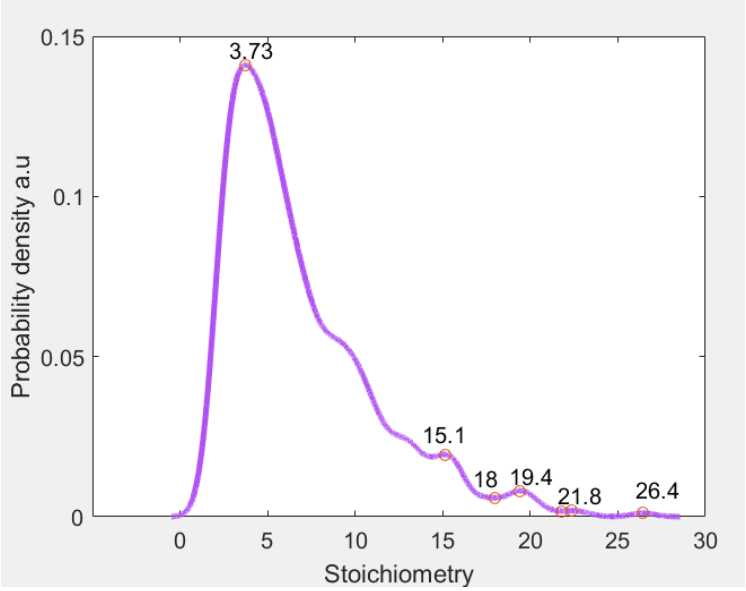


Figure 5-14: Determination of the SpoIIIAH stoichiometry in the absence of SpoIIQ by combining and averaging the Kernel Density Estimation plots (KDE) shown in Figure 5-13. The plot shows a single fundamental peak at 3.73±0.2 and a periodicity spacing of approximately 3 molecules, suggesting SpoIIIAH forms a trimer in the absence of its partner protein, SpoIIQ. Kernel Width 0.7.

5.10 Diffusion Coefficient analysis suggests SpoIIIAH has low mobility in the absence of SpoIIQ

The MATLAB code was used to determine the microscopic diffusion coefficient of SpoIIIAH in the absence of SpoIIQ to determine the protein's mobility in the absence of its partner protein. Diffusion

coefficients were calculated for three individual replicates with replicates looking broadly similar to one another (**Figure 5-15**). As shown in **Figure 5-15**, KDE plots from the microscopic diffusion coefficient analysis show a range of values from 0-3 $\mu m^2/s$ suggesting that some SpoIIIAH molecules are more mobile than others. All 3 plots in **Figure 5-15** have a peak close to zero suggesting that most SpoIIIAH molecules in the population have low mobility. The diffusion coefficient plots were then combined and averaged to produce the single plot shown in **Figure 5-16.** In **Figure 5-16**, there is a range of diffusion coefficient values from 0-3$\mu m^2/s$ suggesting the presence of a mixed population with some SpoIIIAH molecules moving significantly faster than others. The plot also shows a sharp peak at 0.05 $\mu m^2/s$ suggesting most SpoIIIAH molecules are relatively immobile. These results suggest that SpoIIIAH has low mobility in the presence and absence of SpoIIQ as in the presence of SpoIIQ most SpoIIIAH molecules had a mobility of 0.0422 $\mu m^2/s$.

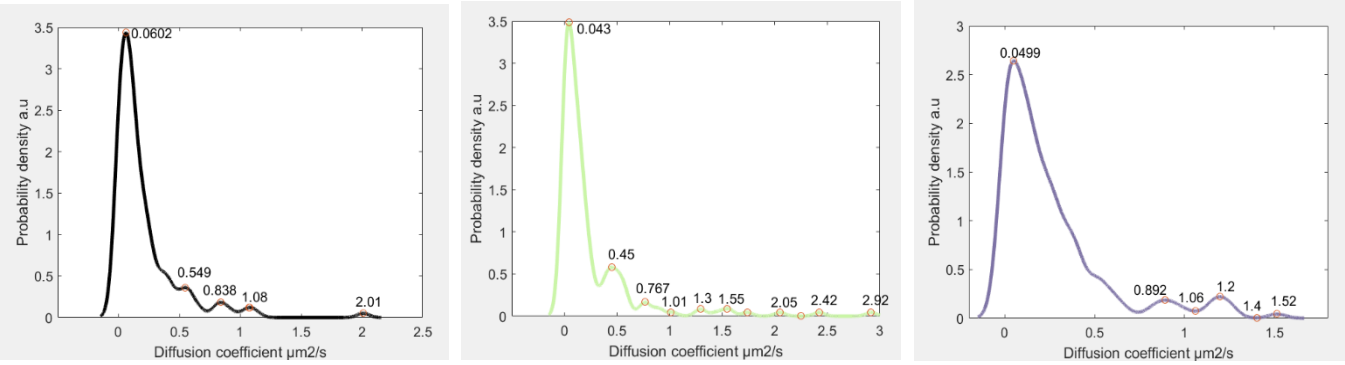


Figure 5-15: SpoIIIAH diffusion coefficient determination using Kernel Density Estimation (KDE) in the absence of SpoIIQ. KDE plots of three independent mCherry SpoIIIAH delSpoIIQ data sets showing the presence of a single peak close to zero and a range of mobilities from 0 -3 $\mu m^2/s$. These plots suggest that SpoIIIAH is largely immobile as most molecules in the population have a low mobility of 0.04-0.06 $\mu m^2/s$. Kernel width 0.05.

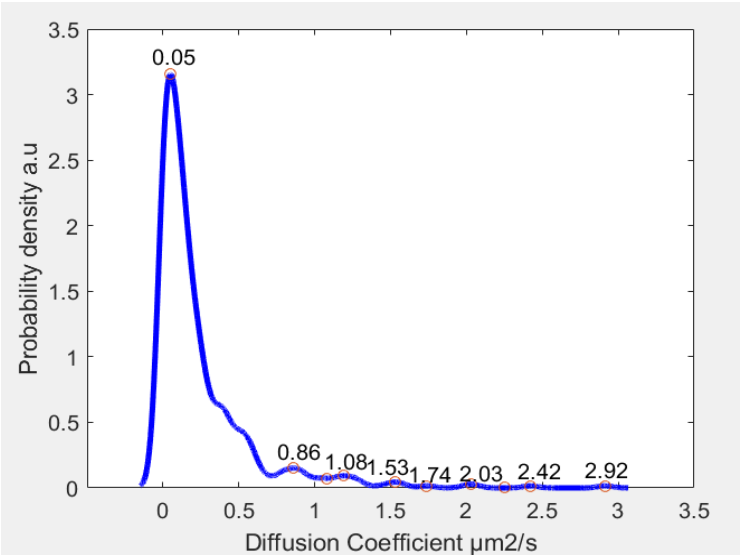


Figure 5-16: Determination of the SpoIIIAH diffusion coefficient in the absence of SpoIIQ by combining and averaging the Kernel Density Estimation plots (KDE) shown in Figure 5-15. The plot shows a single peak at 0.05 $\mu m^2/s$ and a range of values from 0-3 $\mu m^2/s$. Kernel width 0.05.

5.11 Copy Number Analysis suggests downregulation of SpoIIIAH in the absence of SpoIIQ

The MATLAB code was used to calculate the total number of SpoIIIAH molecules in the absence of SpoIIQ, in live *B. subtilis* cells. Similarly, to the stoichiometry and diffusion coefficient analysis, three replicates were analysed, and as shown in **Figure 5-17**, a single cell contains an average of 110 ± 20 SpoIIIAH molecules when SpoIIQ is missing. As well as the sharp peak values around 100, **Figure 5-17** shows some cells with higher copy numbers which is to be expected as not all cells are identical and some cells may be brighter than others, leading to greater SpoIIIAH expression in these cells. These results suggest that SpoIIIAH is downregulated in the absence of SpoIIQ as the number of SpoIIIAH molecules was 263±49 in the presence of SpoIIQ and 110 ± 20 in the absence of SpoIIQ.

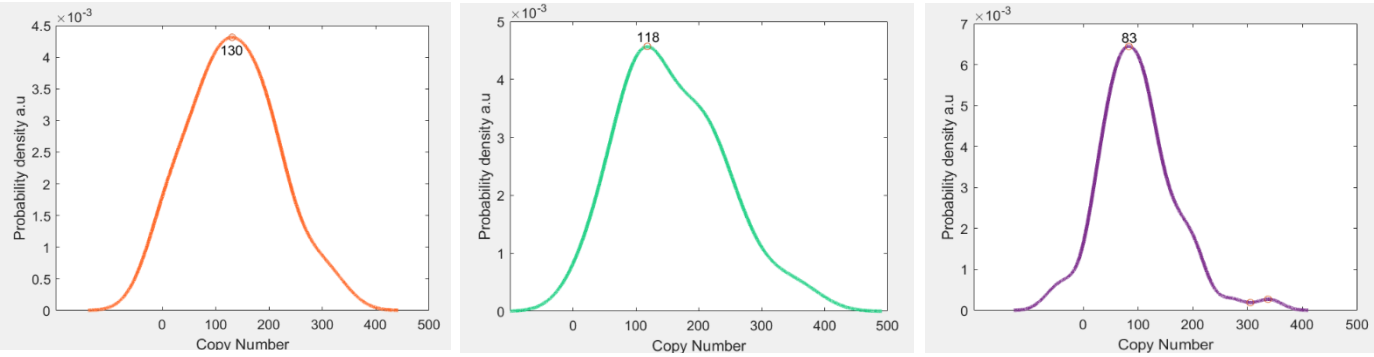


Figure 5-17: Determination of the SpoIIIAH copy number using Kernel Density Estimation (KDE) in the absence of SpoIIQ. KDE plots of three independent mCherry spoIIIAH delspoIIQ data sets showing the presence of a single peak around 100 and a range of copy numbers from 0 to 400 suggesting most cells have an average of 110±20 SpoIIIAH molecules per cell.

5.12 Discussion

In this chapter, Slimfield microscopy was used to characterise the effects of perturbing the interaction between SpoIIQ and SpoIIIAH. Imaging of the mGFP SpoIIQ delSpoIIIAH strain using SIM and Slimfield microscopy revealed that in the absence of SpoIIIAH, SpoIIQ was randomly distributed in the forespore and not present at the sporulation septum in any cells. SpoIIQ was unable to form arcs/foci at the engulfing forespore membrane in the absence of SpoIIIAH, suggesting that SpoIIIAH is important for the appropriate localisation of SpoIIQ during engulfment. A similar result was found for partner protein SpoIIIAH, as Slimfield microscopy imaging of the mCherry SpoIIIAH delSpoIIQ strain showed that SpoIIIAH was randomly dispersed in the mother cell membrane in the absence of SpoIIQ. Without SpoIIQ, SpoIIIAH was unable to localise to the sporulation septum and to the engulfing membrane on the mother cell side, suggesting SpoIIIAH localisation occurs in a SpoIIQ-dependent manner.

Custom-built MATLAB code was used to determine the molecular stoichiometries of SpoIIQ and SpoIIIAH in mutant backgrounds when the partner protein of either SpoIIQ or SpoIIIAH is missing.

The mGFP SpoIIQ delSpoIIIAH strain imaged and analysed in this thesis chapter has 'light' fluorescently labelled *spoIIQ* at the *amyE* locus and 'dark' wildtype *spoIIQ* at the native *spoIIQ* locus, with both 'light' and 'dark' spoIIQ expressed from the wildtype spoIIQ promoter. In this thesis, imaging of the mGFP SpoIIQ delSpoIIIAH strain has been treated as a preliminary attempt as the native unlabelled *spoIIQ* gene is present, making data interpretation difficult. The mGFP SpoIIQ delSpoIIIAH KDE stoichiometry plot (**Figure 5-8**), shows a broad fundamental peak between 5-6 and this broad peak could be hiding lots of smaller peaks at smaller stoichiometries. This is because when the kernel width is set to 0.7 the fundamental peak between 5-6 is very broad and if a single population of SpoIIQ molecules forming either a pentamer or a hexamer was present the peak would be expected to be narrower in terms of its width. The presence of a broad fundamental peak may suggest the presence of multiple populations and not a single population forming a pentamer/hexamer. It is difficult to determine the distributions of 'light' and 'dark' SpoIIQ which makes it difficult to draw quantitative conclusions and motivates the need to have a strain in which the native *spoIIQ* gene is disrupted.

SpoIIIAH oligomerises differently in the absence of SpoIIQ. In the absence of SpoIIQ, the fundamental subunit size of SpoIIIAH changes from 5.79± 0.3 in a wildtype SpoIIQ background to 3.73±0.2 when SpoIIQ is removed from the system, suggesting lower SpoIIIAH stoichiometry in the absence of SpoIIQ. This can be explained by assuming that the SpoIIIAH hexamers assemble from SpoIIIAH trimers. This assumption is supported by the stoichiometry plots shown in **Figure 5-18**, which show that in a wildtype SpoIIQ background, SpoIIIAH is predominately a hexamer but there is evidence of trimers in some traces. Trimers are shown by the shoulder peaks around 3 in some of the individual replicates (**Figure 5-18).** This suggests that in a wildtype SpoIIQ background, SpoIIQ mediates the interaction between SpoIIIAH trimers which come together to form the SpoIIIAH hexamers needed for Q-AH channel formation. In the absence of SpoIIQ, however, SpoIIIAH trimers are randomly dispersed throughout the mother cell membrane as the Q-AH channel complex cannot form.

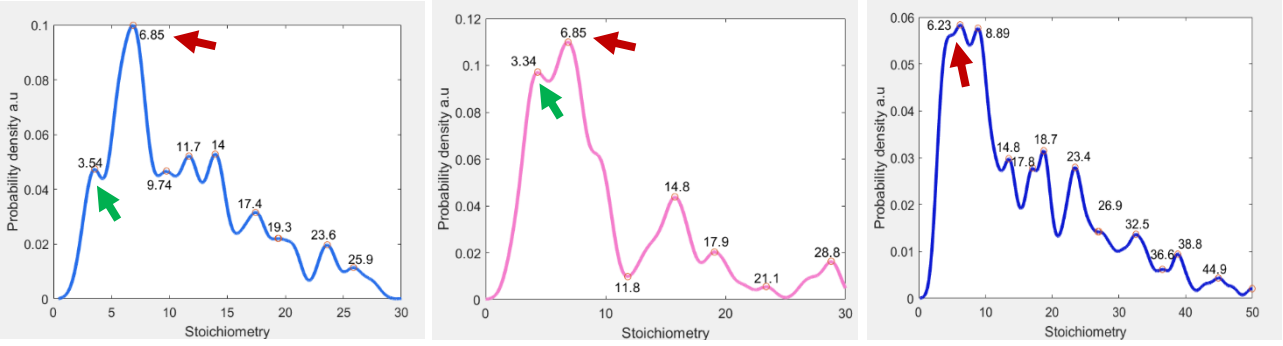


Figure 5-18: Stoichiometry distributions of SpoIIIAH in a wildtype SpoIIQ background shown using Kernel Density Estimation (KDE). KDE plots of three independent mCherry SpoIIIAH data sets showing the presence of a single fundamental peak at 6 (red arrows) and a shoulder peak around 3 (green arrows) in 2 out of 3 replicates. These plots suggest that SpoIIIAH is predominately forms hexamers, and the hexamers may assemble from trimers. Kernel Width 0.7.

Diffusion coefficient analysis was performed to investigate how the mobility of SpoIIQ and SpoIIIAH changes in a mutant background compared to a wild-type background. SpoIIQ has a low mobility in the presence and absence of its partner protein. This is shown by the peak at 0.0152 $\mu m^2/s$ in the presence of SpoIIIAH (**Figure 4-19**, Chapter 4) and 0.0159 $\mu m^2/s$ in the SpoIIIAH deficient strain (**Figure 5-10**). SpoIIQ appears to have higher mobility in the absence of SpoIIIAH. This is because there are more populations of SpoIIQ molecules with higher diffusion coefficients in the SpoIIIAH deficient background (**Figure 5-10**) compared to the SpoIIIAH wildtype background (**Figure 4-16**, Chapter 4). This change in mobility can be explained by the fact that SpoIIQ forms a very large interface with SpoIIIAH, and so if SpoIIIAH is absent, SpoIIQ cannot lock into the Q-AH channel complex present at the engulfing membrane. Similarly, to SpoIIQ, SpoIIIAH has low mobility in the presence and absence of its partner protein. This is shown by the sharp peak at 0.0422 $\mu m^2/s$ in the presence of SpoIIQ (**Figure 4-24**, Chapter 4) and 0.05 $\mu m^2/s$ in the SpoIIQ deficient strain (**Figure 5-16)**. SpoIIIAH appears to be more mobile than SpoIIQ in the presence and absence of its partner protein. This suggests that in the presence of SpoIIQ, SpoIIIAH trimers diffuse from the mother cell membrane to the sporulation septum via 'diffusion and capture' and in the absence of SpoIIQ, SpoIIIAH trimers are randomly dispersed throughout the mother cell.

Copy number analysis was used to determine whether the total number of SpoIIQ and SpoIIIAH molecules changes in a mutant background compared to a wild-type background. The average number of SpoIIIAH molecules decreases from 263±49 in a wildtype SpoIIQ background (**Figure 4-25**, Chapter 4) to 110±20 in a mutant background (**Figure 5-17**) suggesting there is downregulation of SpoIIIAH expression in the absence of SpoIIQ. The average number of SpoIIQ molecules per cell stays roughly the same from 232± 24 in a wild-type SpoIIIAH background (**Figure 4-20**, Chapter 4) to 228±67 in a mutant SpoIIIAH background (**Figure 5-11**).

Chapter 6: Analysis of the SpoIIQ – SpoIIIAH interface using Site Directed Mutagenesis

6.1 Introduction

SpoIIQ and SpoIIIAH are two proteins required for engulfment and membrane migration in *B. subtilis.* SpoIIQ is a forespore specific protein transcribed under σ^F control whereas SpoIIIAH is a mother cell protein transcribed under σ^E control. Both proteins use their N-terminal transmembrane domains to anchor themselves into the membrane and interact with one another via their C-terminal extracellular domains. The C-terminal extracellular domain of SpoIIQ comprises 282 residues with residues 75 to 232 arranged to form 13 β strands in 3 β sheets and a single α helix (**Figure 6-1**) (Levdikov *et al*., 2012) (Meisner *et al.,*2012). The C-terminal extracellular domain of SpoIIIAH consists of four α-helices and three β strands organised in the order α1- β1- β2- α3- α4- β3 (**Figure 6-1**) (Levdikov *et al*., 2012). SpoIIQ and SpoIIIAH form an interface with one another as illustrated in the ribbon and space filling diagrams shown in **Figure 6-1**. The ribbon diagram shows that at the SpoIIQ–SpoIIIAH interface, β3 from SpoIIIAH aligns with β3 from SpoIIQ and helix α3 from SpoIIIAH packs across helix α1 from SpoIIQ (Levdikov *et al*., 2012).

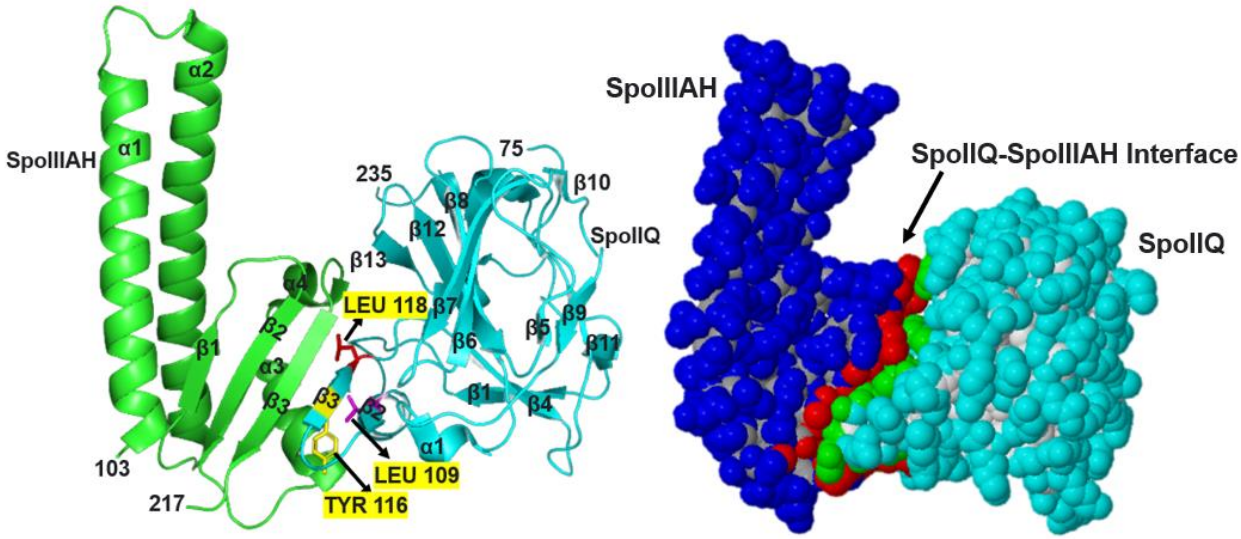


Figure 6-1: Diagrams illustrating the interface between the forespore protein SpoIIQ and mother cell protein SpoIIIAH. A) Ribbon diagram of the X-ray crystal structure of the complex formed between the C-terminal extracellular domains of SpoIIQ (residues 75-232) (blue) and SpoIIIAH (residues 103-217) (green) with the 3 core SpoIIQ residues Leu 118 (red), Leu 109 (pink) and Tyr (yellow) highlighted (Created using PyMoL). B) Space filling diagram of the SpoIIQ-SpoIIIAH interface taken from PISA (Krissinel and Henrick, 2007). PDB = 3TUF

6.2 The SpoIIQ – SpoIIIAH Interface

The SpoIIQ – SpoIIIAH interface was further investigated using the online tool PISA (Proteins, Interfaces, Structures, Assemblies) (Krissinel and Henrick, 2007), which allows exploration of

macromolecular interactions using data from protein crystallography. The interface summary table in **Figure 6-2** shows that there are differences in the number of atoms of SpoIIQ and SpoIIIAH at the interface and surface. Atoms not found at the interface or surface are buried within the core of the proteins. SpoIIQ comprises 1202 atoms with 674 atoms on the surface, of which 82 contribute to the interface with SpoIIIAH. SpoIIIAH has 894 atoms with 615 atoms totally or partially buried at the surface of which 89 contribute to the SpoIIQ – SpoIIIAH interface. As well as the differences in the number of atoms, there are differences in the number of residues of SpoIIQ and SpoIIIAH at the interface and at the surface. **Figure 6-2** shows that 23 SpoIIQ residues interact with 20 SpoIIIAH residues at the SpoIIQ – SpoIIIAH interface. The table shows that SpoIIQ comprises 158 residues with 141 residues at the surface, of which 23 residues contribute to the interface meaning 17 residues are buried within the core of the protein. SpoIIIAH comprises 115 residues with 114 residues at the surface, of which 20 residues contribute to the interface meaning 1 residue is buried within the core of the protein. PISA was next used to determine how much of each protein's surface area is buried within the interface. As illustrated in **Figure 6-2**, SpoIIQ has a total surface area of 8248Å with 807Å buried within the interface. SpoIIIAH, on the other hand, has a total surface area of 7690Å with 744 Å buried within the interface. These calculations suggest that at the SpoIIQ – SpoIIIAH interface 1500 Å of total surface area is buried. This is sufficient for a stable and specific association between SpoIIQ and SpoIIIAH and is in the range of buried surface areas observed for non-obligate protein - protein complexes (Levdikov *et al*., 2012).

	SpoIIIAH		SpoIIQ	
Number of atoms				
interface	89	10.0%	82	6.8%
surface	615	68.8%	674	56.1%
total	894	100.0%	1202	100.0%
Number of residues				
interface	20	17.4%	23	14.6%
surface	114	99.1%	141	89.2%
total	115	100.0%	158	100.0%
Solvent-accessible area, Å				
interface	744.0	9.7%	806.7	9.8%
total	7690.1	100.0%	8248.0	100.0%

Figure 6-2: Interface summary table showing the number of atoms and residues of SpoIIQ and SpoIIIAH at the interface and at the surface. The table shows how much of each protein's surface area is buried within the interface.

6.3 Core and Rim Residues

The interface residues in SpoIIQ and SpoIIIAH can be further subdivided into core and rim residues (Janin *et al.*,2008). Core residues are those which become completely buried upon formation of a protein – protein complex whereas rim residues are residues on the surface that are not completely buried. The interfacing residues table in **Figure 6-3** allows easy identification of the core and rim residues in SpoIIQ and SpoIIIAH as column 2 is the total accessible surface area of the residue and column 3 is the accessible surface area buried within the interface. As shown in **Figure 6-3**, of the C-terminal SpoIIIAH residues involved in binding SpoIIQ, Ala193, Thr194, Ile197, Val210, Ala211 and Val212 are the core residues as these residues become completely buried within the interface and make a significant contribution to the interface. The SpoIIIAH rim residues are His188, Lys190, Lys205, Asp209, Thr213 and Phe214. **Figure 6-3** shows that the core residues in SpoIIQ involved in binding SpoIIIAH are Leu109, Tyr116 and Leu118, as these residues have essentially all their surface buried within the interface. The rim residues in SpoIIQ are Tyr96, Thr98, Glu106, Asn114, Thr115, Ser117 and Lys120.

A) SpoIIIAH Core and Rim Residues

		2	3
A:HIS 188	H	128.78	62.32 |||||
A:SER 189		50.65	10.38 |||
A:LYS 190		185.42	100.84 ||||||
A:SER 191	H	78.07	32.06 |||||
A:LYS 192		76.99	0.00
A:ALA 193		20.92	20.92 |||||||||
A:THR 194	H	70.67	67.99 |||||||||
A:ALA 195		59.53	0.00
A:ILE 196		2.34	0.00
A:ILE 197		53.74	53.74 |||||||||
A:ASP 198	HS	97.88	31.87 ||||
A:LEU 199		54.34	0.00
A:VAL 200		6.61	0.00
A:ALA 201		46.16	0.00
A:LYS 202		158.51	23.87 ||
A:GLU 203		109.96	0.00
A:ILE 204		39.06	0.00
A:LYS 205	H	72.78	58.20 |||||||
A:THR 206		92.82	26.21 |||
A:MET 207		146.33	43.10 |||
A:LYS 208		67.81	11.40 ||
A:ASP 209	H	94.75	65.30 ||||||
A:VAL 210	H	19.97	19.97 |||||||||
A:ALA 211		28.09	26.43 |||||||||
A:VAL 212	H	29.39	29.39 |||||||||
A:THR 213	H	68.69	31.31 |||||
A:PHE 214	H	33.67	27.98 ||||||||
A:GLU 215		51.66	0.00
A:PRO 216		102.65	0.67 |
A:SER 217		79.98	0.00

B) SpoIIQ Core and Rim Residues

		2	3
B:TYR 96	H	40.32	31.91 |||||||
B:GLU 97		66.40	0.00
B:THR 98		112.27	75.98 ||||||
B:ASP 99		130.73	2.68 |
B:ALA 100		31.67	0.00
B:ALA 101		66.92	0.00
B:LYS 102	H	143.78	26.33 ||
B:GLU 103		119.22	0.00
B:GLU 104		84.76	0.00
B:LYS 105	HS	56.94	30.45 |||||
B:GLU 106	H	92.33	68.03 |||||||
B:ALA 107		69.64	6.01 |
B:ALA 108		3.19	0.00
B:LEU 109		47.53	47.15 ||||||||
B:VAL 110		8.61	0.00
B:THR 111		49.59	16.04 ||||
B:TYR 112		83.02	0.00
B:ASN 113	H	104.09	6.57 |
B:ASN 114	H	140.14	79.45 |||||
B:THR 115	H	74.06	40.78 |||||
B:TYR 116	H	139.56	131.00 ||||||||
B:SER 117	H	40.17	32.66 ||||||||
B:LEU 118	H	83.79	79.22 ||||||||
B:SER 119		0.00	0.00
B:LYS 120	H	78.48	58.36 |||||||

Figure 6-3: Interfacing residues table showing the core and rim residues in SpoIIQ and SpoIIIAH. Core residues are those which are significantly buried within the interface (blue stars) whereas rim residues are solvent accessible and not buried within the interface (green stars). The table allows easy identification of the core and rim residues as column 2 is the total accessible surface area of the residue and column 3 is the accessible surface buried within the interface. A) SpoIIIAH core residues are Ala193, Thr194, Ile197, Val210, Ala211 and Val212 and the rim residues are His188, Lys190, Lys205, Asp209, Thr213 and Phe214. B) SpoIIQ core residues are Leu109, Tyr116 and Leu118 and the rim residues are Tyr96, Thr98, Glu106, Asn114, Thr115, Ser117 and Lys120. Vertical lines represent how much of the accessible surface of a residue is buried. The greater the number of vertical lines, more of the accessible surface area of a residue is buried.

6.4 Rationale: Site Directed Mutagenesis

In Chapter 5, Slimfield microscopy was used to show that SpoIIIAH localises to the engulfing mother cell membrane in a SpoIIQ dependent manner as the protein was randomly distributed throughout the mother cell membrane when SpoIIQ was absent. This suggests that SpoIIQ is important in localising SpoIIIAH to the sporulation septum, consistent with the diffusion and capture mechanism proposed previously (Blaylock *et al*.,2004). The effect of SpoIIIAH absence on SpoIIQ localisation is more difficult to determine as in the absence of SpoIIIAH, SpoIIQ localises throughout the forespore membrane as expected for a forespore protein. Core residues Leu109 and Leu118 in SpoIIQ were mutated to alanine, phenylalanine, and glutamic acid to investigate the effect of mutagenesis on sporulation and to determine whether the interactions between SpoIIQ and SpoIIIAH seen in the crystal structure (**Figure 6-1**) are important *in vivo*. Site directed mutagenesis is predicted to mediate the loss of interaction between SpoIIQ and SpoIIIAH as the Q-AH channel complex cannot form, preventing cells from sporulating efficiently. Leu109 and Leu118 were mutated to phenylalanine as this would introduce a bulky phenyl group which would be expected to prevent SpoIIQ from locking onto SpoIIIAH and forming a tight interaction at the interface. The two leucine residues were mutated to glutamic acid as this would introduce a negative charge which would not be accommodated in the hydrophobic interface. Leu109 and Leu118 were also mutated to alanine as alanine is expected to truncate the side chain allowing side chain contribution to be effectively measured.

6.5 pSG1154-mGFP SpoIIQ Verification

The pSG1154-mGFP spoIIQ plasmid described in Chapter 4 (**Figure 6-4**) was decided as the template for mutagenesis of the two core leucine residues in SpoIIQ. This plasmid was first digested with PstI to establish its integrity as a template for mutagenesis. PstI digest was expected to yield DNA fragments of 4.7kb and 4.3kb (**Figure 6-4**). After cleavage, the fragments were separated by agarose gel electrophoresis and as shown in **Figure 6-4** two bands with mobilities of 5kb and 2.5kb were observed. This suggested that the pSG1154-mGFP spoIIQ plasmid was smaller than expected as not

enough kb of DNA was accounted for. PCR was next used to establish the presence of the pspoIIQ-mGFP-spoIIQ insert using the forward pspoIIQ primer AH59 (**Appendix 1**) and reverse SpoIIQ primer AH72 (**Appendix 1**). As the pspoIIQ-mGFP-spoIIQ insert is approximately 2kb in size, the agarose gel in **Figure 6-5** suggests that amplification of the insert was successful. A 2kb fragment was not observed in the control lane lacking the DNA template, as expected. These findings suggested that the loss of DNA seen with the PstI digest was not due to the absence of the insert but instead due to the presence of a smaller than expected pSG1154 vector backbone. The pSG1154-mGFP spoIIQ plasmid was next discarded and not used in subsequent cloning steps.

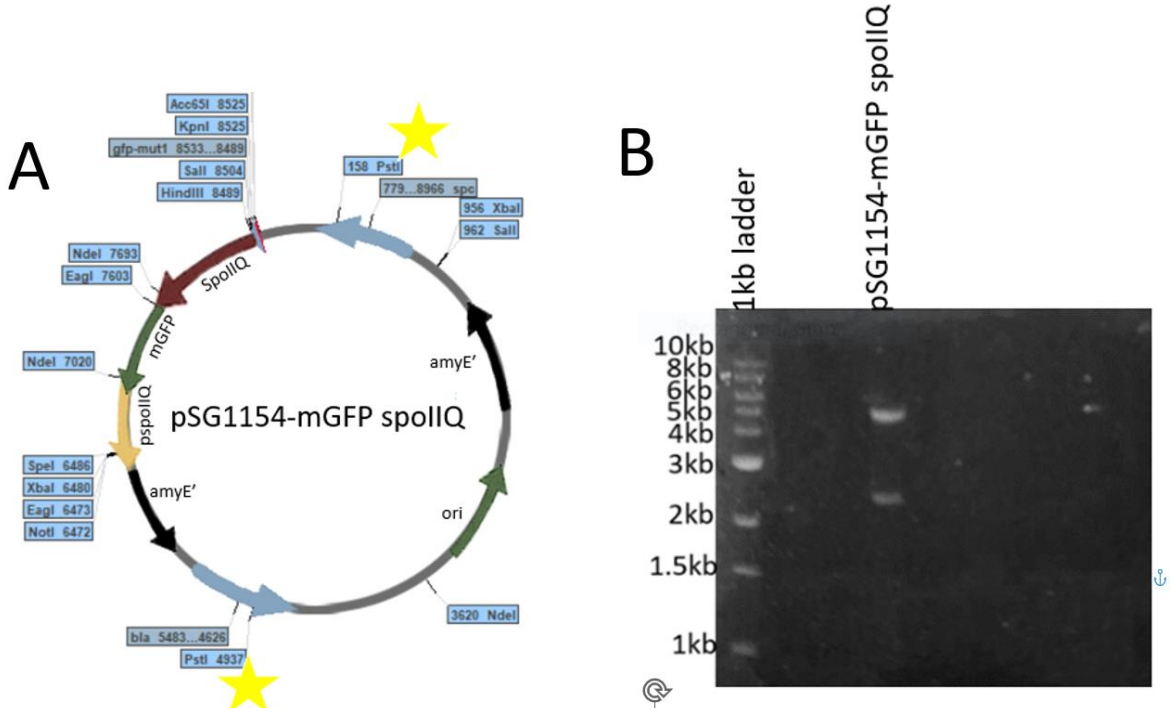


Figure 6-4: Analysis of the pSG1154-mGFP spoIIQ plasmid using restriction enzyme digestion. A) Plasmid map of the pSG1154-mGFP spoIIQ plasmid showing the pspoIIQ-mGFP-spoIIQ insert flanked by sequences from the amyE' locus. PstI restriction sites are highlighted with yellow stars and PstI digest is expected to yield two fragments of 4.7kb and 4.3kb. B) Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion with PstI. Digestion of the pSG1154-mGFP spoIIQ plasmid with PstI yields two bands with mobilities relative to the markers of 5kb and 2.5kb suggesting that the plasmid is smaller than expected.

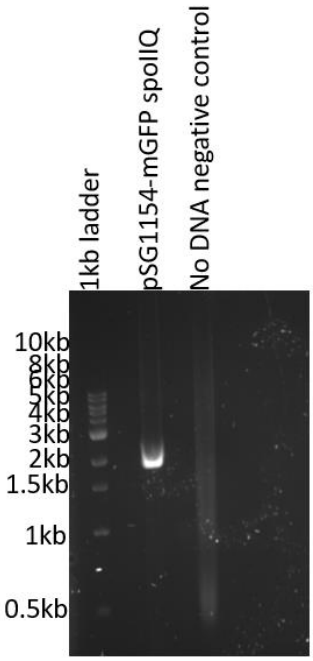


Figure 6-5: Analysis of the pSG1154-mGFP spoIIQ plasmid using PCR. Agarose gel electrophoresis using a 1% agarose gel showing the presence of the 2kb insert in the pSG1154-mGFP spoIIQ plasmid. No 2kb fragment is observed in the no DNA negative control lane as expected.

6.6 pSG1154 Verification

A fresh stock of pSG1154 was taken and digested with restriction enzymes to confirm the size of the plasmid before its use in subsequent cloning steps and the resulting fragments were separated by agarose gel electrophoresis (**Figure 6-6**). As shown in **Figure 6-6**, a single digest of pSG1154 with SpeI or HindIII is expected to yield a single fragment of 7kb. A double digest with both restriction enzymes is expected to yield two fragments: a 7kb fragment corresponding to the vector backbone and a 600bp fragment encompassing the gfpmut1 sequence. The agarose gel in **Figure 6-6** shows that the restriction enzyme digestion of pSG1154 produces fragments matching the expected sizes.

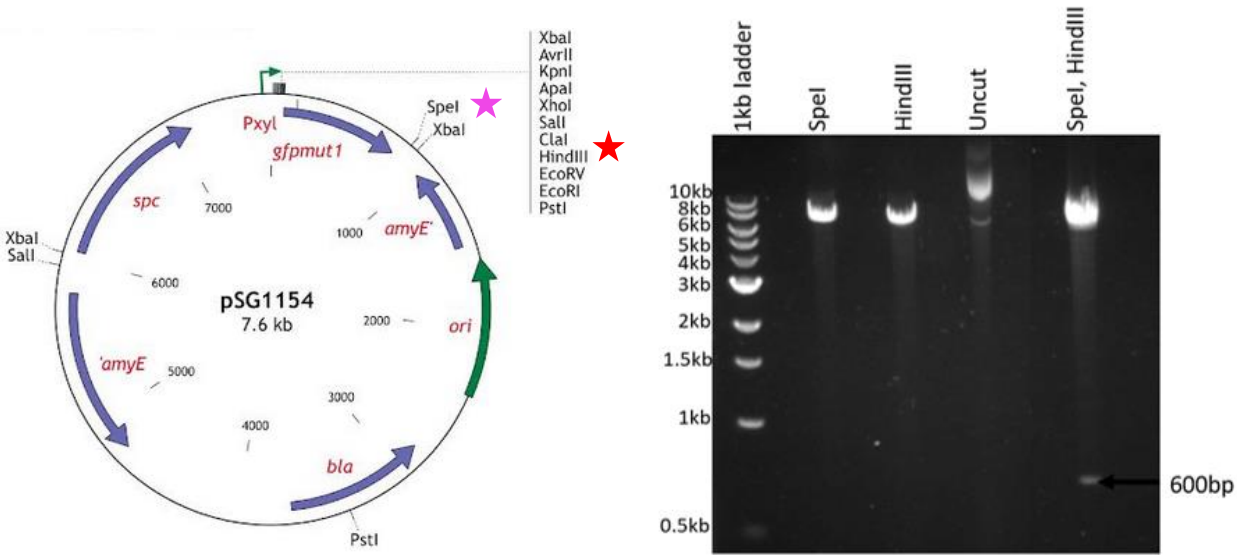


Figure 6-6: Analysis of the pSG1154 plasmid using restriction enzyme digestion. Left) Plasmid map of the pSG1154 plasmid showing the location of key restriction sites including SpeI (pink stars) and HindIII (red stars) Right) Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of pSG1154 with SpeI, HindIII and a double digest with SpeI and HindIII. A single digest with SpeI or HindIII is expected to yield a single fragment of 7kb. A double digest with both restriction enzymes is expected to yield two fragments: a 7kb fragment corresponding to the linearised vector backbone and a 600bp fragment encompassing the gfpmut1 sequence. Uncut pSG1154 was ran on the agarose gel alongside cut pSG1154 and has a different mobility due to circular DNA taking a different conformation to linearised DNA.

6.7 PCR Amplification

PCR was next used to amplify the 2kb pspoIIQ-mGFP-spoIIQ insert from genomic DNA taken from the mGFP spoIIQ delspoIIQ strain (Strain SS102, **Appendix 1**). The insert was taken from genomic DNA and not from the pSG1154-mGFP spoIIQ plasmid as this plasmid had previously been shown to be the wrong size (**Figure 6-4**). Using the pspoIIQ primer AH59 (**Appendix 1**) and reverse SpoIIQ primer AH72 (**Appendix 1**), 4 identical PCR reactions were set up to amplify the insert from genomic DNA. PCR products were then analysed using agarose gel electrophoresis and as shown in **Figure 6-7**, PCR amplification yielded product bands with mobilities relative to the markers of 2kb. PCR products were next extracted from the agarose gel using the Qiagen Gel Extraction Kit as per the manufacturer's specifications.

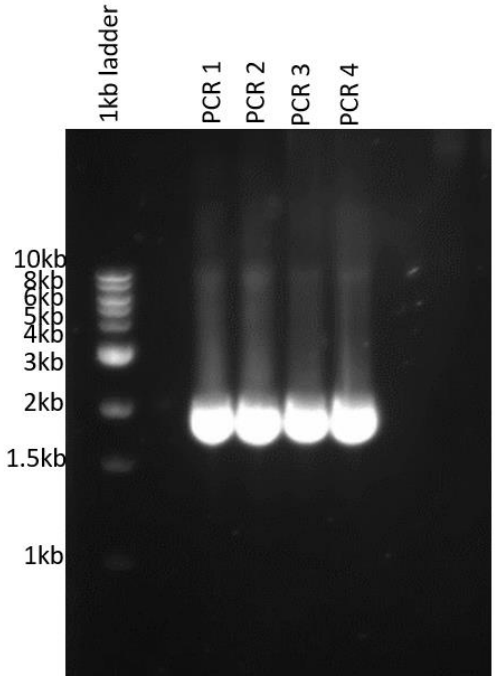


Figure 6-7: Agarose gel electrophoresis using a 1% agarose gel of the PCR amplification of the 2kb pspoIIQ-mGFP-spoIIQ insert from genomic DNA taken from the mGFP spoIIQ delspoIIQ strain (SS102)

6.8 SpeI and HindIII Restriction Enzyme Digest

Samples of the pSG1154 plasmid DNA and the pspoIIQ-mGFP-spoIIQ PCR product were each digested with SpeI and HindIII. A double digest was set up to digest pSG1154 and the pspoIIQ-mGFP-spoIIQ insert and the products were separated by agarose gel electrophoresis (**Figure 6-8**). A double digest of pSG1154 with SpeI and HindIII yields two fragments: a 7kb fragment corresponding to the linearised vector backbone and a 600bp fragment encompassing the gfpmut1 sequence. A double digest of the PCR product with SpeI and HindIII is expected to yield a single 2kb fragment indistinguishable in mobility from the starting materials. After restriction enzyme digestion, the 7kb linearised vector backbone and 2kb insert were extracted from the agarose gel using the Qiagen Gel Extraction Kit as per the manufacturer's specifications. Several attempts were made to ligate the fragments together which proved unsuccessful despite the fragments having complementary sticky ends, meaning an alternative cloning strategy using pCR-Blunt was implemented.

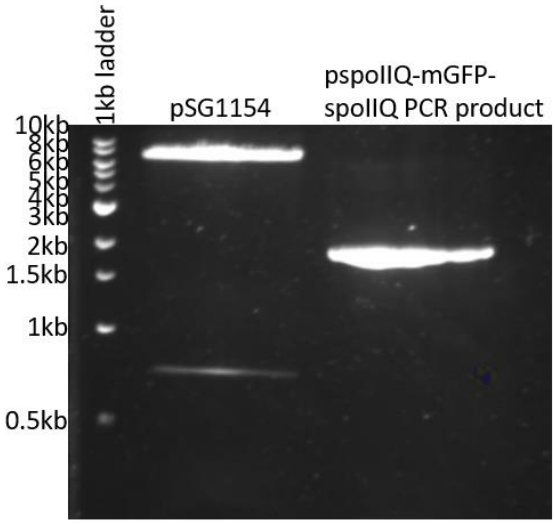


Figure 6-8: Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of pSG1154 and the pspoIIQ-mGFP-spoIIQ PCR product with SpeI and HindIII. Digestion of pSG1154 with SpeI and HindIII yields two fragments: a 7kb fragment corresponding to the linearised vector backbone and a 600bp fragment encompassing the gfpmut1 sequence. Digestion of the pspoIIQ-mGFP-spoIIQ PCR product yields a single 2kb fragment.

6.9 Cloning into pCR-Blunt

As ligation of the pSG1154 vector and pspoIIQ-mGFP-spoIIQ PCR product was proving difficult via restriction enzyme cloning, Dr Jim Brannigan cloned the PCR product into pCR-Blunt using the Zero Blunt™ PCR Cloning kit. As shown in **Figure 6-8**, the pCR-Blunt vector contains the lethal *E. coli ccdB* gene fused to the sequence coding for the C terminus of LacZa (Bernard *et al*., 1994). Ligation of a blunt PCR product to the pCR-Blunt vector disrupts expression of the *lacZa-ccdB* gene fusion, allowing recombinants to grow upon transformation. Cells containing non-recombinant vector are killed when the transformation mixture is plated, as transformation of the non-recombinant vector leads to expression of the *ccdB* gene. The *ccdB* gene encodes the lethal CcdB protein which inhibits DNA gyrase, resulting in DNA breakage and cell death. For this cloning step, PCR was used to amplify the pspoIIQ-mGFP-spoIIQ insert from genomic DNA taken from the mGFP spoIIQ delspoIIQ strain (Strain SS102, **Appendix 1**). The PCR reaction was carried out using Q5 DNA polymerase and primers AH59 and AH72 (**Appendix 1**) to generate a blunt ended insert. The blunt ended insert was then ligated to linearised pCR-Blunt using the ExpressLink™ T4 DNA Ligase enzyme. The ligation mixture was used to transform *E. coli* DH10B competent cells and after an overnight incubation at 37^{o}C, colonies were selected, and plasmids were isolated using the Qiagen Mini Prep kit as per the manufacturer's specifications.

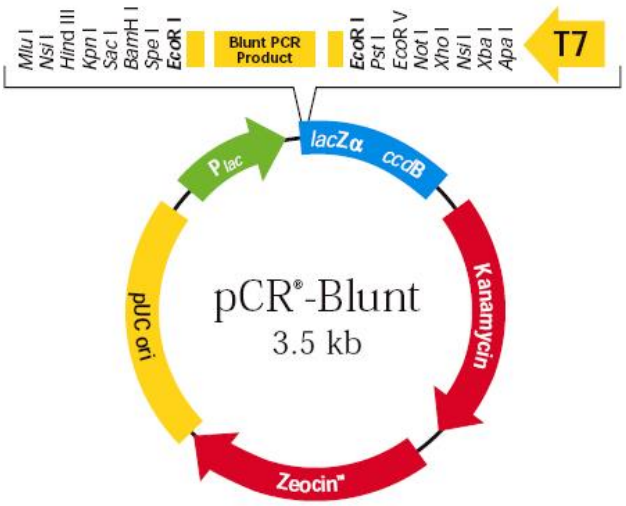


Figure 6-9: Plasmid map of pCR-Blunt showing the location of the *lacZa-ccdB* gene fusion (blue), pUC origin of replication (yellow) and antibiotic resistance genes (red). SpeI and HindIII restriction sites are shown and appear on the same side of the Blunt PCR Product

After the plasmids were isolated from cells using the Qiagen Mini Prep kit, individual recombinant plasmids were digested with SpeI and HindIII to determine the orientation of the insert. As shown in **Figure 6-10**, the blunt ended pspoIIQ-mGFP-spoIIQ insert can ligate into the vector in one of two directions, meaning restriction enzyme digestion is needed to confirm insert orientation as well as insert size. **Figure 6-11** shows that blunt ended cloning generated four clones with two clones in each orientation. pCR24 and pCR28 are in orientation 1 and pCR17 and pCR18 are in orientation 2. The agarose gel in **Figure 6-11** suggests that the insert size is as expected as lanes 1,2,7 and 8 show bands with mobilities relative to the markers of 2kb. Digestion of pCR24 and pCR28 with SpeI yields two fragments: a 3.5kb fragment corresponding to the pCR-Blunt vector backbone and a 2kb fragment corresponding to the insert. Digestion of these plasmids with HindIII yields a single fragment of 5.5kb corresponding to the total size of the pCR-Blunt recombinant plasmid (**Figure 6-11**). This occurs as the HindIII sites are in close proximity so the DNA fragment between the sites is not visualised on the agarose gel. Digestion of pCR17 and pCR18 with HindIII yields two fragments: a 3.5kb fragment corresponding to the pCR-Blunt vector backbone and a 2kb fragment corresponding to the insert. Digestion of these plasmids with SpeI yields a single fragment of 5.5kb corresponding to the total size of the pCR-Blunt recombinant plasmid. This occurs as the SpeI sites are in close proximity meaning the DNA fragment between the sites is not visualised on the agarose gel. The four recombinant plasmids were next sent for sequencing using the M13 forward and M13 reverse sequencing primers (**Appendix 1**). Sequencing confirmed both the insert sequence and orientation as suggested by restriction enzyme digestion.

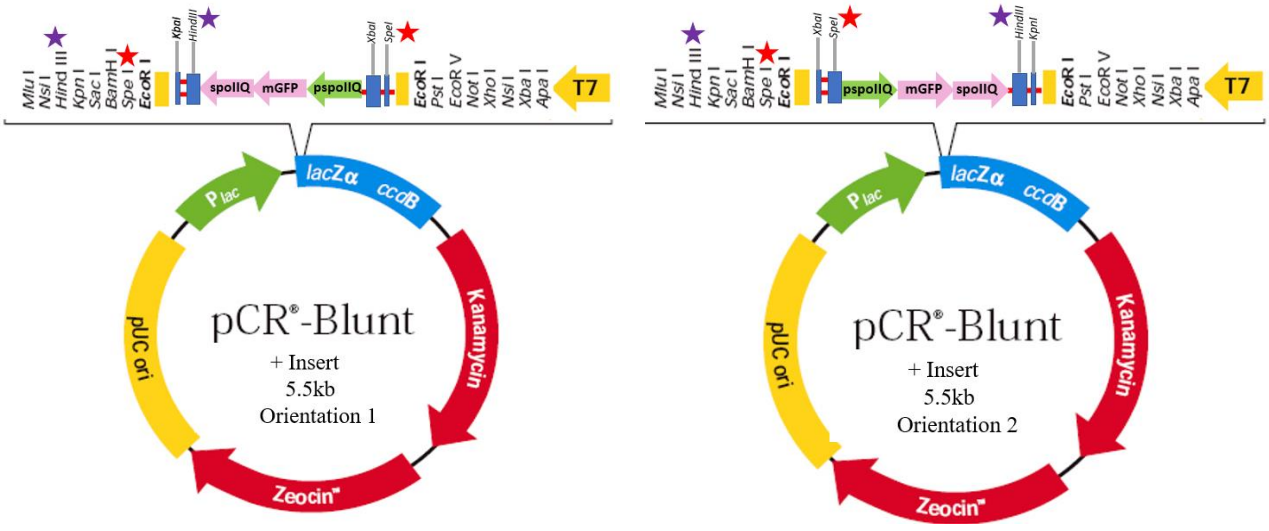


Figure 6-10: Ligation of pspoIIQ-mGFP-spoIIQ to pCR-Blunt using blunt ended cloning. Plasmid maps of pCR-Blunt showing the presence of the pspoIIQ-mGFP-spoIIQ insert and the two orientations the insert can take when ligated to pCR-Blunt using Blunt ended cloning. The total size of the recombinant plasmid is 5.5kb. HindIII sites are highlighted with purple stars and SpeI sites with red stars

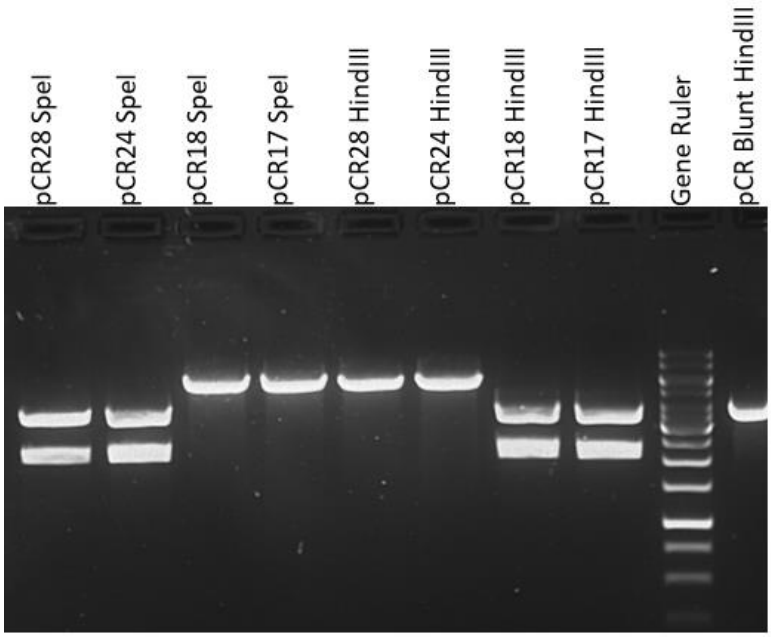


Figure 6-11: Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of pCR-Blunt recombinant plasmids using SpeI and HindIII. Digestion of pCR24 and pCR28 with SpeI yields two fragments: one at 3.5kb corresponding to the pCR-Blunt vector backbone and the other at 2kb corresponding to the insert. Digestion of pCR18 and pCR17 with HindIII yields two fragments: one at 3.5kb (vector backbone) and the other at 2kb (insert). Digestion of pCR28 and pCR24 with HindIII and pCR18 and pCR17 with SpeI generates a single fragment of 5.5kb corresponding to the total size of the pCR-Blunt recombinant plasmid. pCR blunt empty vector cut with HindIII yields a single band at 3.5kb and the GeneRuler shows the insert bands between 2.0 and 2.5kb.

6.10 Site Directed Mutagenesis

The recombinant plasmid pCR18 was used as the template for mutagenesis of the SpoIIQ core residues Leu109 and Leu118 to alanine, phenylalanine, and glutamic acid. The Q5 Site Directed

Mutagenesis kit (Q5 SDM Kit) was used for mutagenesis and allows rapid site-specific mutagenesis of double stranded plasmid DNA in less than 2 hours. The Q5 SDM Kit utilises the robust Q5 Hot Start High-Fidelity DNA polymerase and custom mutagenic primers created using the NEBBaseChanger online primer design tool (**Table 6-1**). Unlike kits that rely on linear amplification, primers designed for the Q5 SDM Kit do not overlap. As shown in **Figure 6-12**, substitutions in SpoIIQ are created by incorporating the desired nucleotide changes in the centre of the forward primer with at least 10 complementary nucleotides included on the 3' side of the mutation. The reverse primer is designed so that the 5' ends of the two primers anneal back-to back. The Q5 Master Mix, primers and template were then combined in a single tube, allowing exponential amplification to occur. Next the PCR product was combined with a KLD enzyme mix comprised of a kinase, ligase and Dpn1 and the mixture was incubated for 5 minutes. During this time, rapid phosphorylation and ligation occurred which circularised the PCR product and removed the template. The reaction was then transformed into NEB 5-alpha competent *E. coli* cells and cells were plated on LB Kanamycin (30 µg/ml) plates before overnight incubation at 37°C. The next day, mutated plasmids were isolated from cells using the Qiagen Mini Prep kit and analysed using restriction enzyme digestion before being sent for sequencing to confirm the presence of the mutations.

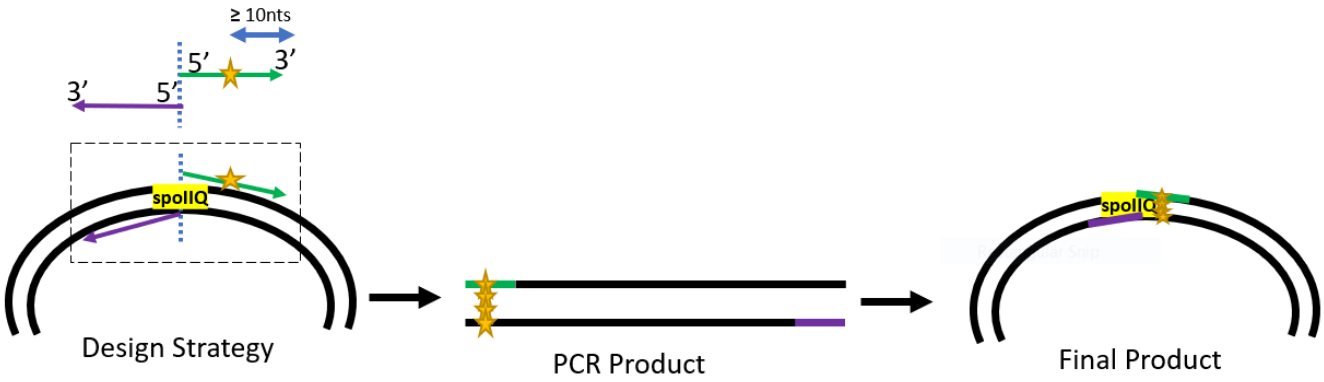


Figure 6-12: Primer design for the Q5 Site Directed Mutagenesis kit (Q5 SDM Kit). The Q5 SDM Kit was used to incorporate substitutions into the *spoIIQ* gene using specifically designed forward (green) and reverse (purple) primers. Primers designed for Q5 SDM are non-overlapping allowing exponential amplification of the DNA template. Substitutions are created by incorporating the desired nucleotide changes (yellow stars) in the centre of the forward primer with at least 10 complementary nucleotides on the 3' side of the mutation.

Mutation	Mutagenic Primers
L109F-F L109F-R	AGAAGCAGCAttcGTTACCTATAATAAC TTCTCTTCTTTTGCGGCATC
L109A-F L109A-R	AGAAGCAGCAgccGTTACCTATAATAAC TTCTCTTCTTTTGCGGCATC
L109E-F L109E-R	AGAAGCAGCAgagGTTACCTATAATAACACGTAC TTCTCTTCTTTTGCGGCATC
L118F-F L118F-R	CACGTACAGCtttAGCAAAGGAATTG TTATTATAGGTAACGAGTGC
L118A-F L118A-R	CACGTACAGCgcaAGCAAAGGAATTG TTATTATAGGTAACGAGTGC
L118E-F L118E-R	CACGTACAGCgaaAGCAAAGGAATTG TTATTATAGGTAACGAGTGC

Table 6-1: List of mutagenic primers used in site directed mutagenesis. Primers were designed using NEBBaseChanger. Mutated codons are highlighted in yellow lowercase text.

The mutated plasmids were digested with SpeI and HindIII and fragments were separated by agarose gel electrophoresis. As shown in **Figure 6-13**, digestion of the mutated plasmids with a double digest of SpeI and HindIII yields two fragments: a 3.5kb fragment corresponding to the pCR-Blunt vector backbone and a 2kb fragment corresponding to the insert, suggesting that the insert is of expected size. The wildtype parent plasmid pCR18 was also digested with SpeI and HindIII with digestion yielding two fragments of 3.5kb and 2kb.

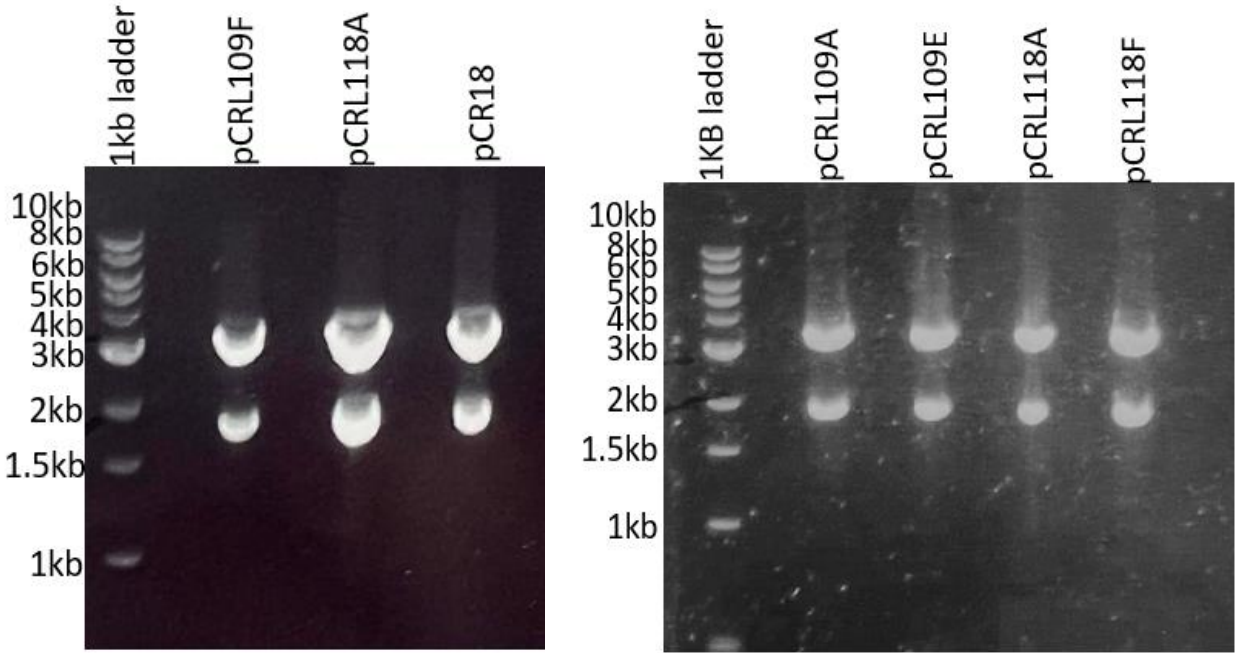


Figure 6-13: Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of the mutated plasmids and the wildtype pCR18 parent plasmid with SpeI and HindIII. Every lane in the agarose gel shows two fragments: a 3.5kb fragment corresponding to the pCR-Blunt vector backbone and a 2kb fragment corresponding to the insert suggesting the insert is of expected size.

6.11 Sequencing

The mutated plasmids were sent for sequencing to confirm the presence of the mutation using the forward and reverse SpoIIQ sequencing primers (**Appendix 1**). In the first round of sequencing, only one of each plasmid from a single colony was sent for sequencing (**Table 6-2**), and only two plasmids out of six had both the desired mutation and the correct SpoIIQ sequence. In the second round of sequencing, three of each plasmid were sent for sequencing and the sequencing results of this attempt are summarised in **Table 6-3**.

Mutated Plasmid	Presence of the mutation √ or X	Verification of the insert sequence √ or X	Remark
pCRL109A	X	√	Wildtype codon present
pCRL109F	√	√	Proceed
pCRL109E	X	X	Leucine CTC mutated to Valine GTG
pCRL118F	X	X	Sequence shorter than expected
pCRL118A	√	√	Proceed
pCRL118E	√	X	Base deletion 9bp upstream

Table 6-2: Sequencing results from the first round of sequencing of the six mutated plasmids. Mutated plasmids with the desired mutation and correct SpoIIQ sequence are highlighted in yellow.

Mutated Plasmid	Presence of the mutation √ or X	Verification of the insert sequence √ or X	Remark
1. pCRL109A	X	X	Sequence shorter than expected
2. pCRL109A	X	√	Wildtype codon present
3. pCRL109A	√	√	Proceed
4. pCRL109E	√	√	Proceed
5. pCRL109E	√	√	Proceed
6. pCRL109E	√	X	Base deletion 9bp upstream
7. pCRL118F	√	X	Base deletion 9bp upstream
8. pCRL118F	√	√	Proceed
9. pCRL118F	X	√	Wildtype codon present
10. pCRL118E	X	√	Wildtype codon present
11. pCRL118E	√	√	Proceed
12. pCRL118E	√	X	Extra sequence taken up from pCR Blunt

Table 6-3: Sequencing results from the second round of sequencing of the four mutated plasmids which failed in the first round of sequencing. Three of each plasmid were sent for sequencing to maximise the chances of selecting plasmids with the desired mutation and the correct SpoIIQ sequence. Mutated plasmids with the desired mutation and correct SpoIIQ sequence are highlighted in yellow.

6.12 Restriction Enzyme Digestion of Mutated Plasmids

After the presence of the desired mutations was confirmed by sequencing, the insert was isolated from the pCR18 mutated recombinant plasmids using a double digest of SpeI and HindIII. The six mutated plasmids were first re-transformed into NEB competent cells and after an overnight incubation at 37^{o}C, plasmids were isolated using the Qiagen Mini Prep Kit as per the manufacturer's specifications. pSG1154, the six mutated plasmids and the wildtype pCR18 parent plasmid were digested with SpeI and HindIII with the resultant fragments separated by agarose gel electrophoresis. Following gel electrophoresis, the desired bands were isolated from the gel using the Qiagen Gel Extraction kit as per the manufacturer's specifications. The concentrations of the gel extracted fragments were checked using Nanodrop and were found to be in the range of 40-160 ng/µl. As shown in **Figure 6-14**, digestion of pSG1154 with SpeI and HindIII yields two fragments: a 7kb fragment corresponding to the linearised vector backbone and a 600bp fragment encompassing the gfpmut1 sequence. The 7kb fragment was extracted from the gel and kept for subsequent ligation steps. Digestion of the six mutated plasmids and the wildtype pCR18 parent plasmid with SpeI and HindIII yields two fragments: a 3.5kb fragment corresponding to the pCR-Blunt vector backbone and a 2kb fragment corresponding to the insert. The 2kb fragments in each gel were extracted and kept for subsequent ligation steps and the 3.5kb band was discarded.

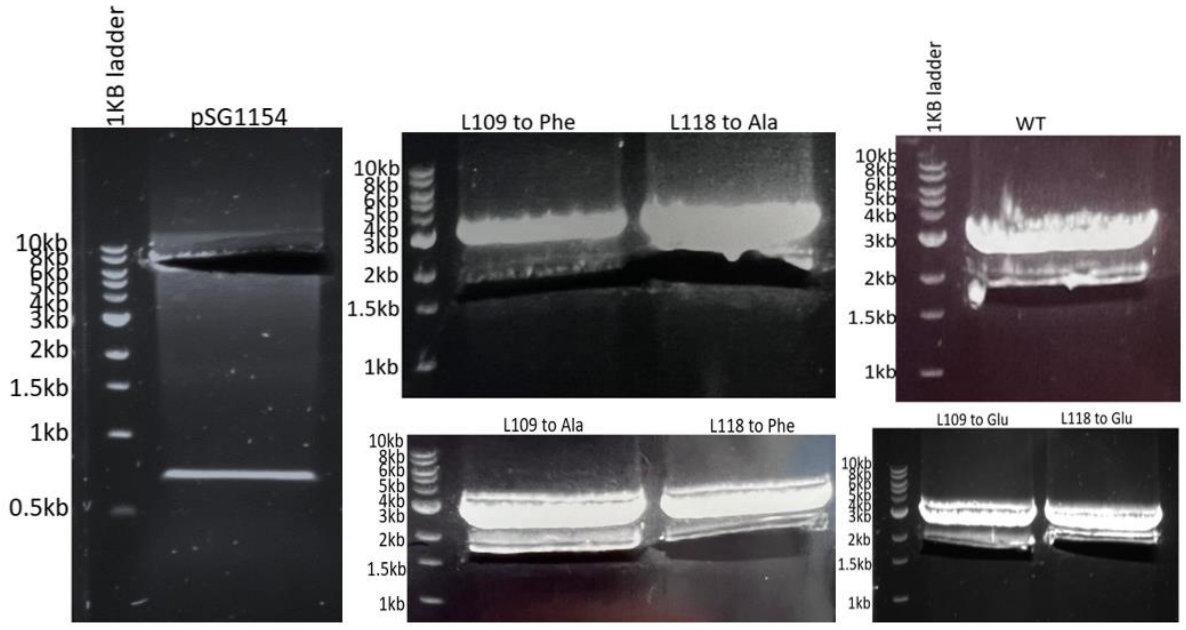


Figure 6-14: Agarose gel electrophoresis using a 1% agarose gel showing the gel extraction of the desired bands after restriction enzyme digestion of the mutated plasmids with SpeI and HindIII. Digestion of pSG1154 with SpeI and HindIII yields two fragments: a 7kb fragment corresponding to the linearised vector backbone which was kept for subsequent ligation steps and a 600bp fragment encompassing the gfpmut1 sequence which was discarded. Digestion of the mutated plasmids and the wildtype pCR18 parent plasmid with SpeI and HindIII yields two fragments: a 3.5kb fragment corresponding to the pCR-Blunt vector backbone and a 2kb fragment corresponding to the insert. 2kb fragments were kept for subsequent ligation steps.

6.13 Ligation of Mutagenic Inserts to pSG1154

After the mutated plasmids and the pSG1154 vector were digested with SpeI and HindIII, the mutagenic 2kb inserts were ligated to the 7kb linearised pSG1154 vector backbone using the NEB Quick Ligase enzyme by Dr Jim Brannigan. This was carried out as ligation into pSG1154 would allow integration of the mutagenic mGFP spoIIQ fusions into the *B. subtilis* chromosome at the *amyE* locus via double cross over, by transforming *B. subtilis* strain IB714 with the mutated pSG1154-mGFP spoIIQ plasmids (**Figure 6-15**).

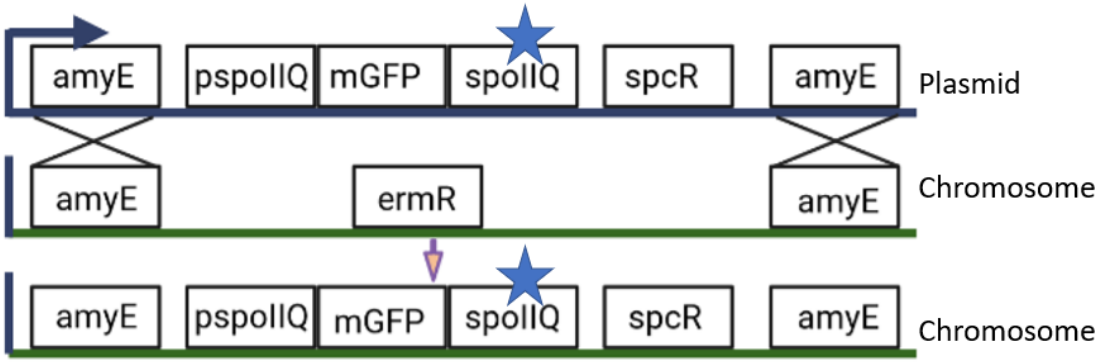


Figure 6-15: Transformation of *B. subtilis* strain IB714 with the mutated pSG1154-mGFP spoIIQ plasmids. The mGFP spoIIQ fusion integrates into the *B. subtilis* chromosome at the *amyE* locus via a double cross over event. There is a switch in antibiotic resistance from erythromycin to spectinomycin. The blue star illustrates the desired mutation in the SpoIIQ core residues Leu109 or Leu118.

The NEB Quick Ligase enzyme was used to ligate the mutagenic inserts (2kb) to the pSG1154 vector backbone (7kb) in 10 µl ligation reactions which were incubated at room temperature for 30 minutes. 50ng of cut vector and 40ng insert were used to give a 3:1 molar ratio of insert to vector. Alongside the test reactions, two control ligations; C1 (cut vector, no ligase) and C2 (cut vector, plus ligase) were set up to verify that the vector had been completely digested. 2.5 µl (1/4 of the ligation reactions) were transformed into NEB C2987 competent cells. For transformation, cells were placed on ice for 30 minutes. Following this, the cells were transferred to a 42°C water bath for 30 seconds before being returned to ice for 5 minutes. 450 µl SOC media (provided in the Q5 SDM Kit) was added to the cells and the cells were incubated at 37°C for 1 hour. After this, 100 µl ($1/5^{th}$ of the samples) was plated onto LB spectinomycin (100 µg/ml) plates and incubated overnight at 37°C. After an overnight incubation at 37°C, colonies were counted. Control plate C1 had 5 colonies and control plate C2 had 8 colonies. Ligations with mutagenic inserts ranged from 50-200 colonies per plate, suggesting that the ligation reactions had worked well.

Following ligation, the pSG1154-mGFP spoIIQ mutated plasmids were isolated from cells using the Qiagen Mini Prep kit as per the manufacturer's specifications. Ligation was verified using restriction enzyme digestion with SpeI and HindIII and fragments were separated by agarose gel

electrophoresis. Digestion of both the pSG1154-mGFP spoIIQ mutated plasmids and the wildtype pSG1154-mGFP spoIIQ plasmid with HindIII and SpeI is expected to yield two fragments: a 7kb fragment corresponding to the pSG1154 vector backbone and a 2kb fragment corresponding to the insert. The agarose gel in **Figure 6-16** shows that ligation was successful as the bands matched the expected sizes. The band corresponding to the insert in lane 2 of the agarose gel (pSG1154-L118E) is slightly bigger than expected and sequencing confirmed that this was due to the presence of an extra sequence that had been taken up from the pCR Blunt vector backbone. Another clone of the pCRL118E plasmid was then digested with HindIII and SpeI to isolate the mutagenic insert before ligation into pSG1154 was repeated for this mutagenic insert. Once all the pSG1154-mGFP spoIIQ mutated plasmids were confirmed by restriction enzyme digestion, plasmids were sent for sequencing using the forward and reverse SpoIIQ sequencing primers (**Appendix 1**). After sequencing, the plasmids were sent to the Barak laboratory in Bratislava, Slovakia for transformation into *B. subtilis* strain IB714, after which sporulation assays would be used determine whether mutating SpoIIQ core residues affected sporulation efficiency in *B. subtilis.*

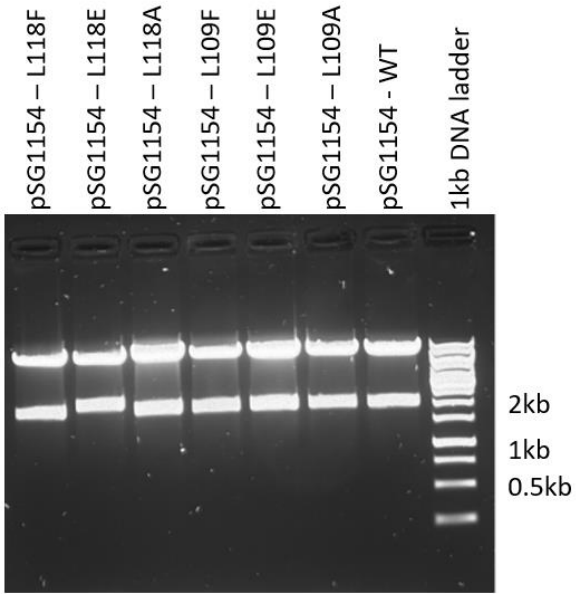


Figure 6-16: Agarose gel electrophoresis using a 1% agarose gel of the restriction enzyme digestion of the six pSG1154-mGFP spoIIQ mutated plasmids and the wildtype pSG1154-mGFP spoIIQ plasmid with a double digest of SpeI and HindIII. Every lane in the agarose gel shows two fragments: a 7kb fragment corresponding to the pSG1154 vector backbone and a 2kb fragment corresponding to the insert suggesting that ligation was successful.

6.14 Discussion

In this chapter, the online tool PISA was used to examine the SpoIIQ-SpoIIIAH interface and identify the core and rim residues in SpoIIQ and SpoIIIAH. SpoIIQ core residues Leu109 and Leu118 were mutated to alanine, phenylalanine and glutamic acid using site directed mutagenesis to investigate

the effect of mutagenesis on sporulation and to determine whether the interactions between SpoIIQ and SpoIIIAH seen in the crystal structure are important *in vivo* (**Figure 6-1**). Site directed mutagenesis is predicted to mediate the loss of interaction between SpoIIQ and SpoIIIAH as the Q-AH channel complex cannot form, preventing cells from sporulating efficiently.

After the pSG1154-mGFP spoIIQ plasmid described in Chapter 4 was found to be an unfit template for site directed mutagenesis, an attempt was made to re-create the pSG1154-mGFP spoIIQ recombinant plasmid. A fresh stock of pSG1154 was digested with HindIII and SpeI, and attempts were made to ligate the 7kb linearised vector backbone (**Figure 6-8**) to the 2kb HindIII and SpeI digested pspoIIQ-mGFP spoIIQ PCR product (**Figure 6-8**) amplified from genomic DNA taken from strain SS102 (**Appendix 1**). Ligation proved unsuccessful on several occasions, and the reactions may have failed as the SpeI and HindIII sites are very close to the ends of the insert making it difficult to determine whether the insert had been fully digested by the restriction enzymes. An alternative cloning strategy using pCR-Blunt was implemented, as pCR-Blunt is a high copy number plasmid and smaller in size than pSG1154, making it an ideal template for site directed mutagenesis. Cloning using pCR-Blunt ensured that the insert was fully digested with HindIII and SpeI and that the Quick ligase enzyme was working as expected.

pCR-Blunt was digested using HindIII and SpeI and the insert was ligated to the vector backbone using blunt ended cloning. The pCR-Blunt recombinant plasmid pCR18 was used as the template for site directed mutagenesis of the SpoIIQ core residues Leu109 and Leu118 to alanine, phenylalanine, and glutamic acid. Mutagenesis was performed using the Q5SDM kit and was confirmed using restriction enzyme digestion and sequencing.

Once mutagenesis was confirmed by sequencing, the mutated plasmids and the pSG1154 plasmid were digested with HindIII and SpeI and the mutagenic inserts were ligated to the 7kb linearised pSG1154 vector backbone using the NEB Quick Ligase enzyme. This was carried out as the ultimate goal of the chapter was to transform the pSG1154-mGFP spoIIQ mutated plasmids into *B. subtilis* strain IB714 in collaboration with the Barak laboratory in Bratislava, Slovakia. Transformation into IB714 would allow the mGFP spoIIQ mutagenic inserts to integrate into the *B. subtilis* chromosome at the *amyE* locus, after which sporulation assays would be used determine whether mutating SpoIIQ core residues Leu109 and Leu118 affected sporulation efficiency in *B. subtilis.*

Chapter 7 – Discussion and Future Work

7.1 Discussion of Results in Chapters 4, 5 and 6

In this thesis, fluorescently labelled SpoIIQ and SpoIIIAH strains were imaged using Slimfield microscopy to determine the localisation of SpoIIQ and SpoIIIAH in wildtype and mutant backgrounds. In the presence of SpoIIIAH, SpoIIQ localises at the sporulation septum before forming arcs and foci at the engulfing forespore membrane. In the absence of SpoIIIAH, however, SpoIIQ is randomly distributed in the forespore and not present at the sporulation septum in any cells. In the presence of SpoIIQ, SpoIIIAH localises to the engulfing membrane on the mother cell side, whereas in the absence of SpoIIQ, SpoIIIAH is randomly dispersed in the mother cell membrane. This suggests that SpoIIIAH localisation to the sporulation septum and engulfing membrane occurs in a SpoIIQ dependent manner.

Custom built MATLAB code was used to determine the single subunit stoichiometry of SpoIIQ and SpoIIIAH in wildtype and mutant backgrounds. In a wildtype background, SpoIIQ was found to form a dimer and SpoIIIAH was found to form a hexamer that may assemble from two SpoIIIAH trimers. These findings suggest that the Q-AH channel complex is assembled from dimers of SpoIIQ and hexamers of SpoIIIAH. This is consistent with both the 12-mer subunit model proposed by Levidikov *et al* and 18-mer subunit proposed by Meisner *et al* but inconsistent with the 15-mer subunit model proposed by *Meisner et al* (Meisner *et al*.,2012) (Levidikov *et al.*,2012) (**Figure 7-1**). As SpoIIQ and SpoIIIAH have multiple subunits rather than being monomeric, assembly of the Q-AH channel complex may be more complicated than initially expected (Meisner et *al*.,2012) (Levidikov *et al*.,2012). In the absence of SpoIIQ, SpoIIIAH was found to form a trimer which suggests that SpoIIQ mediates the interaction between SpoIIIAH trimers to form the SpoIIIAH hexamers needed for Q-AH channel formation. In the absence of SpoIIQ, SpoIIIAH trimers are randomly distributed throughout the mother cell membrane as the Q-AH channel complex cannot form. The stoichiometry of SpoIIQ in the absence of SpoIIIAH was not accurately determined as the mGFP spoIIQ delspoIIIAH strain (SS103, **Appendix 1**) imaged and analysed in this thesis has 'light' fluorescently labelled *spoIIQ* at the *amyE* locus and 'dark' wildtype *spoIIQ* at the native *spoIIQ* locus. It was difficult to draw quantitative conclusions from the analysis of this strain as the distributions of 'light' and 'dark' spoIIQ were hard to determine.

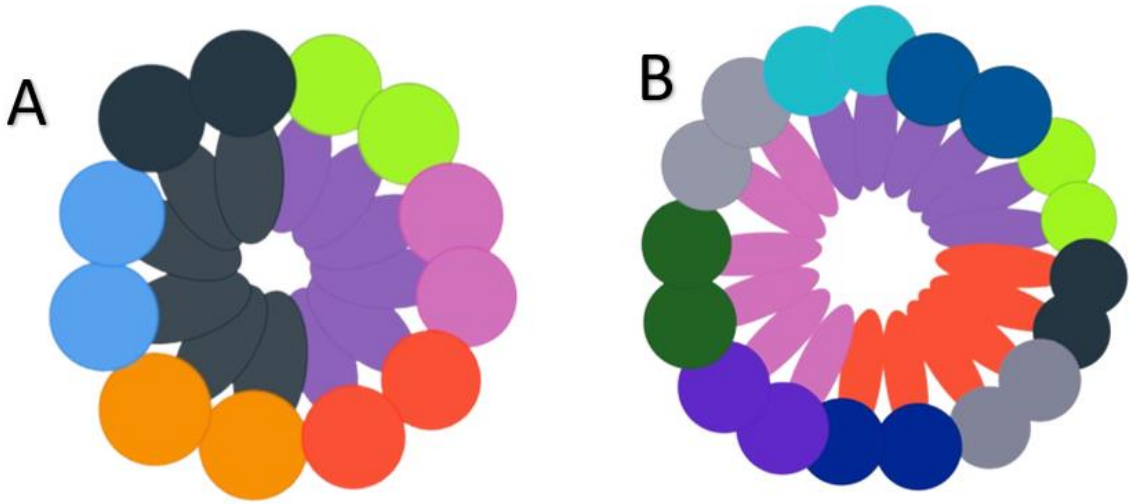


Figure 7-1: Schematic representation of the assembly of the SpoIIQ-SpoIIIAH channel from SpoIIQ dimers and SpoIIIAH hexamers. SpoIIIAH hexamers are believed to be assembled from SpoIIIAH trimers. A) 6 SpoIIQ dimers and 2 SpoIIIAH hexamers made from 4 SpoIIIAH trimers are required for a 12-mer submit model. B) 9 SpoIIQ dimers and 3 SpoIIIAH hexamers made from 6 SpoIIIAH trimers are required for a 15-mer subunit model.

The MATLAB code was used to calculate the diffusion coefficient of SpoIIQ and SpoIIIAH in wildtype and mutant backgrounds to estimate the mobility of the proteins in living cells. In wildtype and mutant backgrounds, SpoIIQ and SpoIIIAH were found to have low mobility, a finding expected for proteins believed to form a channel complex. Interestingly, SpoIIIAH was found to have slightly higher mobility than SpoIIQ, in both wildtype and mutant backgrounds. This suggests that in the presence of SpoIIQ, SpoIIIAH trimers move to the sporulation septum whereas in the absence of SpoIIQ, SpoIIIAH trimers are randomly dispersed in the mother cell.

Copy number analysis was used to estimate the total number of SpoIIQ and SpoIIIAH molecules in sporulating *B. subtilis* cells. In a wildtype background, there are approximately 200-300 molecules of SpoIIQ and SpoIIIAH. This is lower than the copy number of SpoIIIE as the number of SpoIIIE molecules rises to 700 as sporulation progresses (Wollman *et al*., 2020). In the absence of SpoIIQ, the number of SpoIIIAH molecules decreased, suggesting there is downregulation in SpoIIIAH expression when SpoIIQ is absent. In both wildtype and mutant backgrounds there is variation in the number of SpoIIQ and SpoIIIAH molecules, suggesting that there is a very large distribution in SpoIIQ and SpoIIIAH expression levels in live cells.

The SpoIIQ-SpoIIIAH interface was investigated using the online tool PISA and the core and rim residues in both SpoIIQ and SpoIIIAH were identified. SpoIIQ core residues Leu109 and Leu118 were mutated to alanine, phenylalanine and glutamic acid, and a set of pSG1154-mGFP spoIIQ mutated plasmids were created using restriction enzyme cloning and site directed mutagenesis. The mutated

plasmids were sent to the Barak laboratory in Bratislava, Slovakia for transformation into the *B. subtilis* strain 1B714. Transformation into IB714 would allow the mGFP spoIIQ mutagenic inserts to integrate into the *B. subtilis* chromosome at the *amyE* locus, after which sporulation assays would be used determine whether mutating the SpoIIQ core residues affected sporulation efficiency in *B. subtilis.*

7.2 Discussion of Results in the Context of Sporulation

The results of this thesis support the idea that SpoIIQ and SpoIIIAH form a channel complex that connects the forespore and mother cell, with the complex being gated on the mother cell side. Research in the field has shown that the SpoIIIAA proteins SpoIIIAA, SpoIIIAB, SpoIIIAC, SpoIIIAD, SpoIIIAE and SpoIIIAG share similarity with components of secretion systems, suggesting that these proteins are involved in regulating the Q-AH channel from the mother cell side (Meisner *et al*.,2008). SpoIIIAA resembles ATPases found in type II and type IV secretion systems. SpoIIIAB shares weak similarity with GspF and TadB/C proteins found in type II and type IV secretion systems. SpoIIIAG shares weak similarity with the YscJ/FliF protein family (Meisner *et al*.,2008). SpoIIIAF shares weak similarity with FlhB (YscU) found in type III secretion systems. SpoIIIAE is similar to ABC-type permeases involved in Type I secretion systems (Meisner *et al*., 2008).

In 2017, Zeytuni *et al* determined the structure of SpoIIIAG. SpoIIIAG forms a large, ring like structure in the intermembrane space between the forespore and mother cell. The SpoIIIAG ring was found to interact with SpoIIQ and the other SpoIIIAA proteins to form a continuous channel between the mother cell and forespore (Zeytuni *et al*., 2017). Using single particle cryo-electron microscopy, the high-resolution 3D structure of the SpoIIIAG channel was determined, revealing that SpoIIIAG is composed of 30 repeating units, representing the highest level of symmetry that has been reported within any protein assembly (**Figure 7-2**) (Kora, 2018) (Zeytuni *et al*., 2017). Each component in the SpoIIIAG channel contains a 'b-triangle motif', which is proposed to be the driving force for assembly of the ring complex (Kora, 2018) (Zeytuni *et al*., 2017). The researchers believe that the 30-mer SpoIIIAG ring best correlates with the 15-mer SpoIIQ-SpoIIIAH ring model as the two have similar dimensions, orientations, and charge compatibility, suggesting the possibility of an interaction interface between the two complexes when forming the sporulation complex (Zeytuni *et al*., 2017). The data in this thesis disagrees with this as it found that SpoIIQ forms dimers and SpoIIIAH forms hexamers with the channel being assembled from dimers of SpoIIQ and hexamers of SpoIIIAH. As the 15-mer subunit model cannot be assembled from dimers and hexamers, it suggests that the interaction between the SpoIIQ-SpoIIIAH ring model and the SpoIIIAG ring model is more complex than that assumed by Zeytuni *et al* (Zeytuni *et al*., 2017).

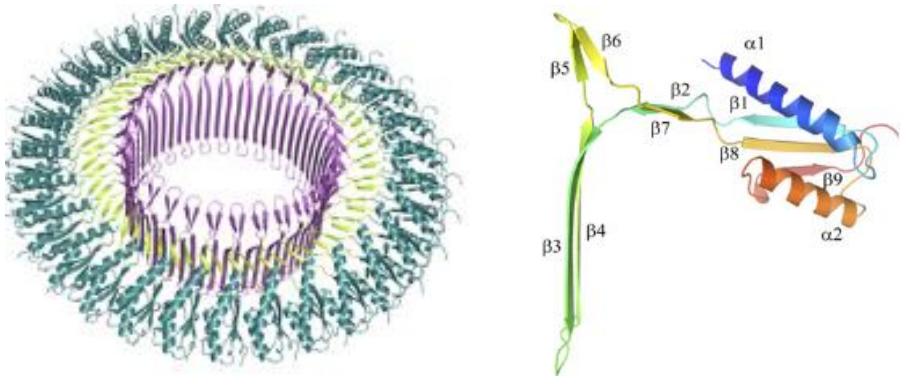


Figure 7-2: Cryo-EM structure of SpoIIIAG. Left) Overview of the 3D structure of SpoIIIAG – structure colour coded by distinctive ring elements. Right) The monomeric structure comprises two motifs, the RBM α1β1β8α2β9 fold and the β2–8 -triangle motif (Taken from Zeytuni *et al*., 2017).

Research has suggested that whilst SpoIIIAH lies at the heart of the complex and links the SpoIIIAA mother cell proteins to SpoIIQ in the forespore, cells lacking SpoIIIAH have a sporulation efficiency of 5% whilst all other mutants in the complex are 1,000-fold worse (Doan *et al*., 2009). This is interesting as Slimfield microscopy imaging of the mGFP SpoIIQ delSpoIIIAH strain in Chapter 5 showed that in the absence of SpoIIIAH, engulfment was blocked as SpoIIQ failed to localise at the sporulation septum and at the engulfing/engulfed forespore membrane in any cells. The 1,000-fold worse sporulation efficiency in cells lacking any SpoIIIAA protein besides SpoIIIAH can be explained by assuming that SpoIIIAH tethers the SpoIIIAA proteins to SpoIIQ (Doan *et al*., 2009) (Meisner *et al*., 2008). In this instance, in the absence of SpoIIIAH the other SpoIIIAA proteins could interact weakly with SpoIIQ or the SpoIIIAG protein, which, like SpoIIIAH, shares weak homology with the YscJ/FliF protein family and might function in place of SpoIIIAH (Doan *et al*., 2009) (Meisner *et al*., 2008).

The intercellular channel complex formed between the forespore protein SpoIIQ and the eight mother cell proteins SpoIIIAA-SpoIIIAH is believed to function as a 'secretion apparatus' that secretes osmolytes and/or nutrients from the mother cell into the forespore (Doan *et al*. 2009). This allows the complex to 'feed' the forespore in the final stages of preparation for dormancy. Research has supported the idea of the 'secretion complex' by showing that in the absence of the SpoIIIAA proteins and SpoIIQ, the forespore develops large invaginations and appears to shrink/collapse when engulfment is complete (Doan *et al*. 2009). The collapse of the forespore membrane may lead to instability of the forespore and be triggered by a lack of osmolytes and/or metabolic potential. This idea is further supported as it has been shown mutations that impair glycerol metabolism in *B. subtilis* result in membrane collapse (Doan *et al*. 2009).

After engulfment, the forespore becomes fully engulfed inside the mother cell and isolated from the external environment and the forespore developmental program might involve downregulation (or

shut-off) of genes required for certain metabolic functions (Doan *et al*. 2009). Under these circumstances, the forespore becomes dependent on the mother cell for nutrients and/or osmolytes as the forespore has lost the ability to provide for itself and has become dependent on the mother cell for its final stages of maturation (Doan *et al*. 2009).

7.3 Alternative methods to study the interactions between SpoIIQ and SpoIIIAH

Both existing research in the field and the results of this thesis support the idea that SpoIIQ and SpoIIIAH interact and form a channel complex spanning the forespore - mother cell intermembrane space. To further confirm this theory, *in vitro* biochemical methods could be used to verify the interaction between SpoIIQ and SpoIIIAH. Co-immunoprecipitation (Co-IP) is a technique commonly used to study protein-protein interactions and relies on the use of antibodies to visualise proteins of interest. In Co-IP, sporulating *B. subtilis* cells are first lysed to produce the desired cell lysate (Figure 7-3). A primary antibody recognising SpoIIQ is then added to the cell lysate as the identity of SpoIIQ is known and a specific antibody recognising SpoIIQ can be engineered. Next, a secondary antibody that only recognises and binds to the primary antibody is added to the cell lysate. Agarose beads that recognise the secondary antibody are then added and make the bound complex heavier in solution. The solution is centrifuged allowing the heavy agarose beads with the bound antibodies and proteins to form a pellet and sink to the bottom of the tube. To separate the antibodies and proteins from each other typically heat and washing is applied. The identity of any proteins binding to SpoIIQ (i.e., SpoIIIAH) can be determined using Mass Spectrometry or Western Blotting.

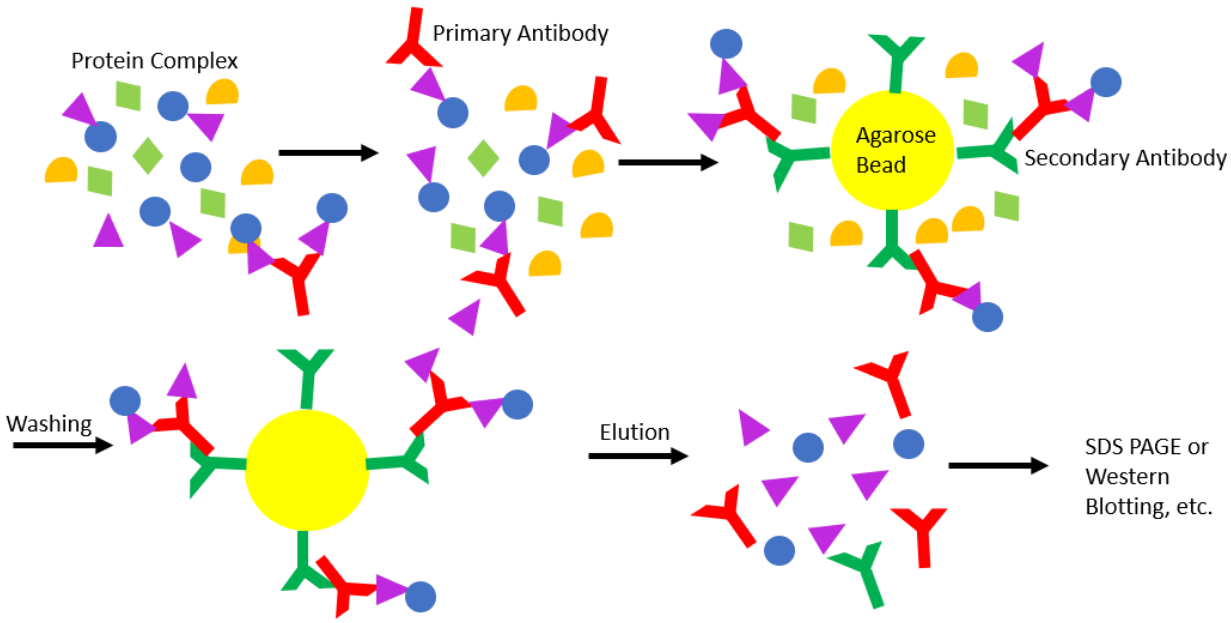


Figure 7-3: Co-immunoprecipitation (Co-IP) as a method to study the interactions between SpoIIQ and SpoIIIAH. SpoIIQ (purple) is bound by a specific primary antibody (red). SpoIIIAH (blue) is bound to SpoIIQ (purple) in the cell lysate. A secondary antibody (green) that recognises the primary antibody (red) is then added to the cell lysate. Agarose beads (yellow) that bind to the secondary antibody (green) are added to make the complex heavier in solution. Washing allows the removal of unbound proteins and elution separates the bound proteins (i.e., SpoIIIAH, blue) from the antibodies. The identity of the bound proteins are confirmed using SDS PAGE or Western Blotting.

Affinity chromatography, also known as the pull-down assay, is another *in vitro* technique that could be used to verify the interaction between the forespore protein SpoIIQ and mother cell protein SpoIIIAH. A pull-down assay uses a column with tiny beads to which the 'bait' protein SpoIIQ is attached via a tag, allowing SpoIIQ to remain immobilized to the column (**Figure 7-4**). Sporulating *B. subtilis* cells are then lysed which releases all the proteins from inside the cells. The cell lysate is then run through the column and when SpoIIIAH encounters SpoIIQ, the two proteins should interact (**Figure 7-4**). As SpoIIQ is the bait protein stuck to the column, if SpoIIIAH binds to SpoIIQ it will also be stuck to the column. A wash buffer is then added which removes any unbound proteins which are then discarded. SpoIIQ and any bound proteins (i.e., SpoIIIAH) are then eluted from the column and Western Blotting is used to confirm the identity of SpoIIIAH as a SpoIIQ binding partner (**Figure 7-4**). This involves denaturing the proteins, running them on a gel and probing the gel with antibodies allowing us to tag and visualise our proteins of interest (SpoIIQ and SpoIIIAH).

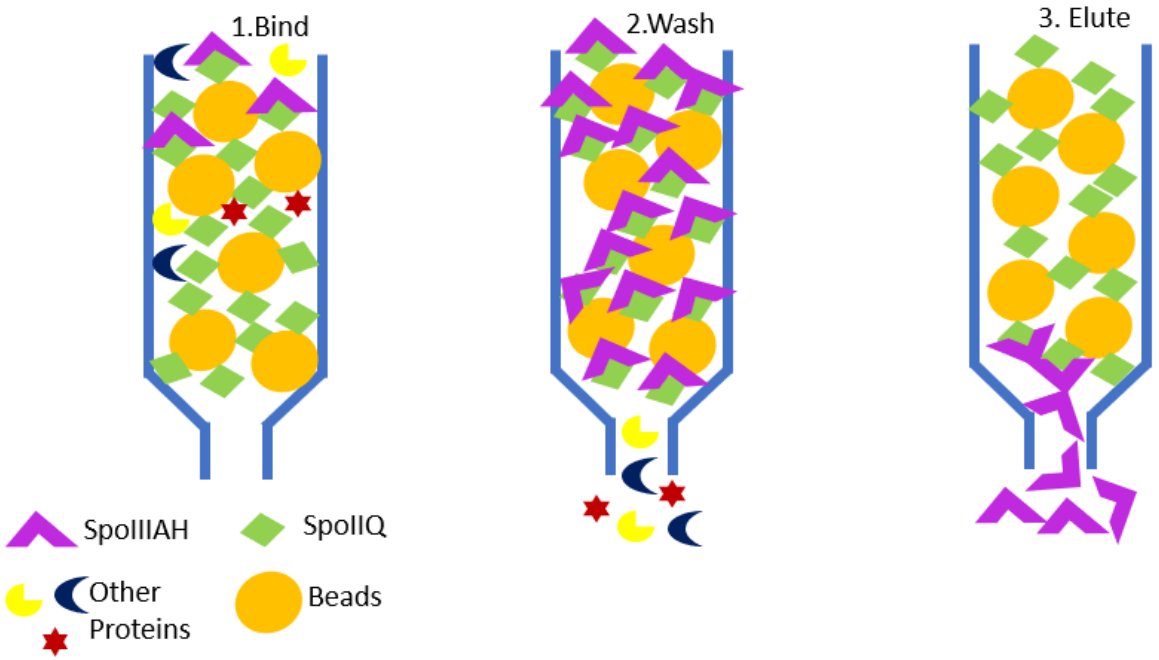


Figure 7-4: Affinity chromatography to study the interactions between SpoIIQ and SpoIIIAH. SpoIIQ (green) is bound to the tiny beads on the column. The cell lysate is passed through the column and any proteins interacting with SpoIIQ (i.e., SpoIIIAH, purple) bind to SpoIIQ. A wash buffer is added to remove any unbound proteins. SpoIIQ binding partners (i.e., SpoIIIAH) are then eluted from the column and analysed to confirm their identity.

The Yeast 2 hybrid system is an *in-vivo* method that could be used to identify a possible interaction between SpoIIQ and SpoIIIAH. The Yeast 2 hybrid system uses reporter genes to show the interactions between proteins. If SpoIIQ and SpoIIIAH interact, then a desired reporter gene will be expressed. If the two proteins do not interact, the reporter gene is not expressed. The Yeast 2 Hybrid system relies on the process of transcriptional activation and manipulates transcription factors. Transcription factors can be split into two domains; a DNA binding domain that binds to DNA and an activator domain that activates transcription by recruiting RNA Polymerase. To look for a potential interaction between SpoIIQ and SpoIIIAH, SpoIIQ is fused to the DNA binding domain and SpoIIIAH is fused to the activator domain (**Figure 7-5**). The DNA binding domain binds to the DNA in the promoter region and as SpoIIQ and SpoIIIAH interact, the activator domain comes into close proximity with the gene region and RNA polymerase is recruited, leading to transcription of the gene (**Figure 7-5**). A common reporter gene used in the Yeast 2 Hybrid system is the HIS3 gene which is required for histidine biosynthesis and without this gene, cells are killed when plated on histidine depleted media.

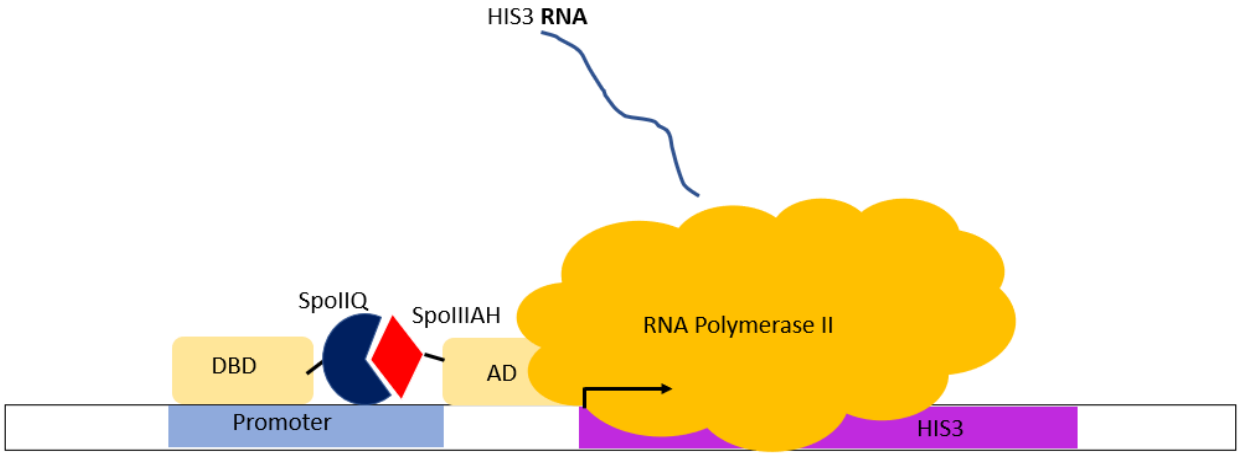


Figure 7-5: The Yeast 2 Hybrid system as a method to determine interactions between SpoIIQ and SpoIIIAH. SpoIIQ is fused to the transcriptional factor DNA binding domain (DBD) and SpoIIIAH to the activator domain (AD). If SpoIIQ and SpoIIIAH interact, the DBD domain binds to the promoter region and the AD encounters the gene. RNA polymerase II recruited and the HIS3 gene is transcribed leading to the production of HIS3 RNA. The HIS3 gene is required for histidine biosynthesis.

Another *in-vivo* technique that can be used to examine interactions between SpoIIQ and SpoIIIAH is Fluorescence Resonance Energy Transfer (FRET). FRET allows the study of molecular interactions in living cells between interacting molecules that are less than 10 nm apart. In FRET, SpoIIQ would be tagged with CFP and SpoIIIAH with YFP (**Figure 7-6**). CFP absorbs light at 440 nm and emits light at 490 nm which is cyan fluorescence. Similarly, YFP absorbs light at 490 nm and emits at 527 nm which is yellow fluorescence. The emission wavelength of CFP acts as excitation wavelength for YFP as the two show spectral overlap (**Figure 7-6**). If SpoIIQ and SpoIIIAH interact with one another, they come

close enough for FRET to occur and yellow fluorescence is observed (**Figure 7-6**). If SpoIIQ and SpoIIIAH do not interact, FRET will not occur, and only cyan fluorescence is produced.

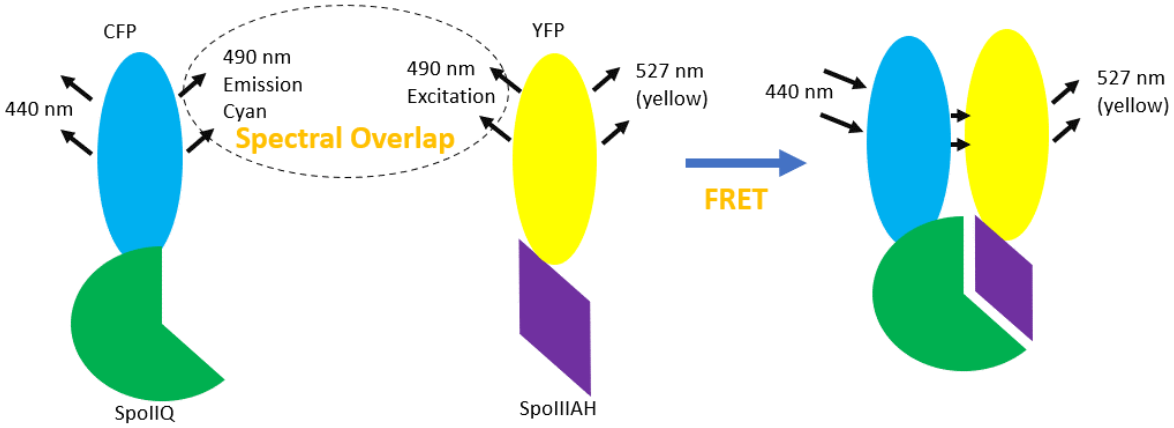


Figure 7-6: Fluorescence Resonance Energy Transfer (FRET) as a method to detect interactions between SpoIIQ and SpoIIIAH. If SpoIIQ and SpoIIIAH interact and are less than 10 nm apart FRET occurs and energy is transferred from the donor molecule CFP to the acceptor molecule YFP. Yellow fluorescence from the YFP is produced.

Once the interaction between SpoIIQ and SpoIIIAH has been further confirmed using both *in-vivo* and *in-vitro* methods, the oligomeric state of both proteins can be further investigated. Interestingly, data in this thesis showed that SpoIIQ forms dimers and SpoIIIAH forms hexamers assembled from trimers. This finding differs from the findings of Levdikov *et al* who used size exclusion chromatography with multiangle laser scattering (SEC-MALS) to determine the quaternary structure of SpoIIQ and SpoIIIAH and their respective complex. SEC MALS is a technique used to determine the spread of different molecular weights in a solution allowing the absolute characterisation of macromolecules in terms of their size (Levdikov *et al*.,2012). Levdikov *et al* analysed the extracellular domains of both SpoIIQ (43-283) and SpoIIIAH (32-218) and mixtures of the two proteins. The researchers found traces that showed single peaks associated with molecular masses of 25.8 kDa (SpoIIQ) and 20.8 kDa (SpoIIIAH) (Levdikov *et al*., 2012). The calculated masses of the recombinant proteins are 26.4 kDa and 21.2 kDa suggesting that both proteins are monomers. When mixed together, the trace showed a new peak with an average molecular mass of 47.6 kDa and as the calculated molecular mass of a 1:1 complex is 47.6 kDa, the researchers concluded that SpoIIQ and SpoIIIAH form a heterodimer (Levdikov *et al*., 2012). The results found in this thesis may differ from these results as the SEC-MALS experiments performed by Levdikov *et al* were performed in solution meaning the effect of the cellular environment was missed. Levdikov *et al* also only examined the extracellular C-terminal domains of SpoIIQ and SpoIIIAH which suggests that the presence of the transmembrane N-terminal domains in live cells may affect the oligomeric stoichiometry of both SpoIIQ and SpoIIIAH

7.4 Future Work

In Chapter 4 of this thesis, a dual labelled mGFP SpoIIQ mCherry SpoIIIAH strain was imaged and analysed. Slimfield microscopy images (**Figure 4-27**) showed that on a qualitative level SpoIIQ and SpoIIIAH co-localise at the sporulation septum and around the engulfing/engulfed forespore membranes. SpoIIQ and SpoIIIAH are not always co-localised in live cells as SpoIIIAH is randomly distributed around the mother cell in some cells (**Figure 4-27**). To expand on this finding, quantitative experiments examining the co-localisation of mGFP SpoIIQ and mCherry SpoIIIAH could be conducted. This would allow observation of the spatial overlap between mGFP and mCherry, as each fluorescent protein has a different emission wavelength allowing us to determine quantitatively whether SpoIIQ and SpoIIIAH are in the same area of the cell or very near to one another.

The preliminary image analysis of the dual labelled strain in Chapter 4 determined the stoichiometry of SpoIIIAH and found differences between the linked and unlinked SpoIIIAH stoichiometry. These findings suggested that the growing channel acts as a site of recruitment and SpoIIIAH stoichiometry builds up as the channel is assembled. In the future, it would be interesting to determine the SpoIIQ stoichiometry in the same way to see whether there are differences in the linked and unlinked SpoIIQ stoichiometry. Differences are to be expected as this thesis has suggested that SpoIIQ and SpoIIAH do not form pre-assembled units with one another but instead are recruited to the growing channel as dimers and trimers. The SpoIIIAH trimers then form hexamers which are used to assemble the SpoIIQ-SpoIIIAH channel complex. As well as stoichiometry analysis, it would be interesting to determine the diffusion coefficient and copy number of SpoIIQ and SpoIIIAH in the dual labelled strain to see whether these are consistent with the findings from the single labelled strains. The diffusion coefficient and copy number values of SpoIIQ and SpoIIIAH in the dual labelled strain are expected to be consistent with those from the single labelled strains as fluorescently labelling the partner protein of either SpoIIQ or SpoIIIAH is not expected to affect the mobility or number of molecules of either protein.

In chapter 5, two mutant strains were imaged and analysed to determine whether the localisation of SpoIIQ and SpoIIIAH was affected by the absence of their partner proteins. It was found that both proteins become delocalised in the absence of their partner proteins and that the stoichiometry, diffusion coefficient and copy number values of SpoIIQ and SpoIIIAH change in mutant backgrounds. The mGFP spoIIQ delspoIIIAH strain imaged and analysed in this thesis chapter has 'light' fluorescently labelled *spoIIQ* at the *amyE* locus and 'dark' wildtype *spoIIQ* at the native *spoIIQ* locus, with both 'light' and 'dark' spoIIQ expressed from the wildtype spoIIQ promoter. As it was difficult to

quantitatively determine the distributions of 'light' and 'dark' *spoIIQ,* a strain in which the native *spoIIQ* gene is disrupted could be constructed in the future. Analysis of this strain would allow more accurate determination of how the stoichiometry, diffusion coefficient and copy number values of SpoIIQ change in a SpoIIIAH deficient background.

In chapter 6, site directed mutagenesis was used to produce a set of pSG1154-mGFP-spoIIQ mutated plasmids in which the SpoIIQ core residues Leu109 and Leu118 were mutated to alanine, phenylalanine, and glutamic acid. The mutated plasmids were produced with the goal of transforming the plasmids into *B. subtilis* to determine whether mutating SpoIIQ core residues affected the sporulation efficiency of *B. subtilis.* The plasmids were sent to the Barak lab in Bratislava, Slovakia who transformed the plasmids into the SpoIIQ null mutant strain BKE36550. Fluorescence microscopy was performed on the strains and intermediate delocalisation of the mGFP-SpoIIQ signal was observed. This meant that mutating SpoIIQ core residues Leu109 and Leu118 caused some SpoIIQ molecules to become delocalised into the forespore whilst other SpoIIQ molecules localised correctly at the sporulation septum and around the engulfing/engulfed forespore membranes. This is interesting and suggests an intermediate phenotype between the correctly localised mGFP SpoIIQ signal observed in Chapter 4 and the completely delocalised mGFP SpoIIQ signal observed in the SpoIIIAH deficient background in Chapter 5. To build on this finding, sporulation assays could be used to determine the sporulation efficiency of *B. subtilis* strains with wildtype SpoIIQ, SpoIIQ in the absence of SpoIIIAH and SpoIIQ with mutated core residues to determine how the sporulation efficiency is affected in each case. It is expected that the *B. subtilis* strain with wildtype SpoIIQ would have the highest sporulation efficiency and the spoIIQ delspoIIIAH strain would have the lowest sporulation efficiency. The strain with mutated SpoIIQ core residues is expected to have an intermediate sporulation efficiency.

Chapter 8 – Appendix

8.1 Plasmids used in this thesis.

Plasmid	Plasmid Features
pSG1154 Chapter 4 Figure 4-1	The pSG1154 plasmid contains an *amyE* open reading frame split in two by a spectinomycin resistance cassette and a sequence encoding a GFP mutant. pSG1154 also has an origin of replication (ori) and an ampicillin resistance cassette (bla).
pSG1154-mGFP spoIIQ Chapter 4 Figure 4-5	In the pSG1154-mGFP spoIIQ recombinant plasmid, a sequence encoding an mGFP spoIIQ fusion replaces the gfpmut1 sequence in the *amyE* open reading frame. The pSG1154-mGFP spoIIQ plasmid is resistant to spectinomycin and ampicillin and has an origin of replication (ori).
pCR Blunt Chapter 6 Figure 6-10	The pCR-Blunt vector contains the lethal *E. coli ccdB* gene fused to the sequence coding for the C terminus of LacZa. The *ccdB* gene encodes the lethal CcdB protein which inhibits DNA gyrase, resulting in DNA breakage and cell death. Ligation of a blunt PCR product to the pCR-Blunt vector disrupts expression of the *lacZa-ccdB* gene fusion, allowing recombinants to grow upon transformation.

8.2 *Bacillus subtilis* strains used in this thesis.

Strain Identifier	Genotype	Source/Reference
IB714	*amyE::erm*	*Bacillus subtilis* Stock Centre
BKE36550	*spoIIQ:: erm trpC2*	*Bacillus subtilis* Stock Centre
BKE24360	*spoIIIAH:: erm trpC2*	Koo et al., 2017
AH103	*spoIIIAH:: mGFP-spoIIIAH-camR*	Adam Hughes, PhD Thesis, University of York 2018
AH104	*spoIIIAH:: mCherry-spoIIIAH-camR*	Adam Hughes, PhD Thesis, University of York 2018
SS101	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ:: spoIIQ*	pSG1154-mGFP spoIIQ plasmid transformed into IB714
SS102	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ:: erm*	SS101 transformed with genomic DNA from BKE36550
SS103	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ:: spoIIQ spoIIIAH::erm*	SS101 transformed with genomic DNA from BKE24360
SS104	*IB714 spoIIIAH:: mCherry-spoIIIAH-camR spoIIQ:: erm*	AH104 transformed with genomic DNA from BKE36550
SS105	*IB714 amyE:: mGFP-spoIIQ-spcR spoIIQ::erm spoIIIAH:: mCherry-spoIIIAH-camR*	SS102 transformed with genomic DNA from AH104

8.3 Oligonucleotide Primers

8.3.1 Gibson Assembly Primers

Primer Name	Sequence	Forward /Reverse	Overlaps (lowercase)	Anneals (uppercase)
AH59	aattcgcggccgctctagaactagtCATAGAGCGGACG TGACTC	Forward	pSG1154	pspoIIQ
S1	gttcttctcctttactcatTGTTTCATCACCTCAGCAAC	Reverse	mGFP	pspoIIQ
S2	tgttggctgaggtgatgaaacaATGAGTAAAGGAGAA GAAC	Forward	pspoIIQ	mGFP
S3	gttttcttttcttcctctctcaTTTTTTTGTATGGTTCATC	Reverse	spoIIQ	mGFP
S4	gatgaaccatacaaaaaaATGAGAGAGGAAGAAAA GAAAAC	Forward	mGFP	spoIIQ
AH72	gaggtcgacggtatcgataaagcttGAGCAATAGGCGG TTCATG	Reverse	pSG1154	spoIIQ

8.3.2 Site Directed Mutagenesis Primers

Mutation	Mutagenic Primers (mutation highlighted in yellow lowercase text)
L109F-F L109F-R	AGAAGCAGCAttcGTTACCTATAATAAC TTCTCTTCTTTTGCGGCATC
L109A-F L109A-R	AGAAGCAGCAgccGTTACCTATAATAAC TTCTCTTCTTTTGCGGCATC
L109E-F L109E-R	AGAAGCAGCAgagGTTACCTATAATAACACGTAC TTCTCTTCTTTTGCGGCATC
L118F-F L118F-R	CACGTACAGCtttAGCAAAGGAATTG TTATTATAGGTAACGAGTGC
L118A-F L118A-R	CACGTACAGCgcaAGCAAAGGAATTG TTATTATAGGTAACGAGTGC
L118E-F L118E-R	CACGTACAGCgaaAGCAAAGGAATTG TTATTATAGGTAACGAGTGC

8.3.3 SpoIIQ Sequencing Primers

Primer	Primer Sequence
FWD SpoIIQ Sequencing Primer	CGGCCGTCATTTTAACAG
REV SpoIIQ Sequencing Primer	TGTCACCTTGCTCTAC

8.3.4 M13 Forward and Reverse Sequencing Primers

Primer	Primer Sequence
M13 Forward	GTAAAACGACGGCCAG
M13 Reverse	CAGGAAACAGCTATGAC

8.4 *Bacillus subtilis* liquid media and reagents

8.4.1 T base

Component	Volume added
$(NH_4)_2SO_4$	1 g
$K_2HPO_4.3H_2O$	9.15 g
KH_2PO_4	3 g
trisodium citrate.$2H_2O$	0.5 g
Ultrapure water	Up to 500 ml (and autoclave at 121 °C for 20 minutes)

8.4.2 SpC

Component	Volume added
T base	20 ml
50% (w/v) glucose	0.2 ml
1.2% (w/v) $MgSO_4.7H_2O$	0.3 ml
10% (w/v) bacto yeast extract	0.4 ml
1% (w/v) casamino acids	0.5 ml
Ultrapure water	18.6 ml

8.4.3 SpII

Component	Volume added
T base	100 ml
50% (w/v) glucose	1 ml
1.2% (w/v) $MgSO_4.7H_2O$	7 ml
10% (w/v) bacto yeast extract	1 ml
1% (w/v) casamino acids	1 ml
0.1 M $CaCl_2$	0.5 ml
Ultrapure water	89.5 ml

8.4.4 SpII+EGTA

Component	Volume added
SpII (without $CaCl_2$)	100 ml
0.1 M ethylene glycol-bis (β-aminoethyl ether)-N, N, N',N'-tetra acetic acid (EGTA) pH 8.0	4 ml

8.4.5 Top Agar

Component	Volume added
LB medium	LB medium plus 0.7% (w/v) agar.

8.4.6 Difco Sporulation Medium (DSM)

Component	Volume added
Bacto-nutrient broth	8 g
10% (w/v) KCl	10 ml
1.2% (w/v) $MgSO_4.7H_2O$	10 ml
1 M NaOH	0.5 ml
Ultrapure water	Make up to IL (and autoclave at 121 °C for 20 minutes)

The following sterile solutions were added after autoclaving:

Component	Volume added
1M Ca $(NO_3)_4$	1 ml
0.1 M $MnCl_2$	1 ml
1 mM $FeSO_4$	1 ml

8.4.7 Spizizen Minimal Medium (SMM)

Component	Volume added
$(NH_4)_2SO_4$	2 g
K_2HPO_4	14 g
KH_2PO_4	6 g
trisodium citrate.$2H_2O$	1 g
$MgSO_4.7H_2O$	0.2 g
Ultrapure water	Make up to IL (and autoclave at 121 °C for 20 minutes)

The following sterile solutions were added after autoclaving:

Component	Volume added
50% (w/v) D-glucose	10 ml
5 mg/ml L-tryptophan	10 ml

References

- Adams, DG. (2000) Heterocyst formation in cyanobacteria. *Curr Opin Microbiol* 3: 618-624.
- Arigoni, F., Guerout-Fleury, A., Barak, I. and Stragier, P. (1999) The SpoIIE phosphatase, the sporulation septum, and the establishment of forespore-specific transcription in Bacillus subtilis: a reassessment. *Molecular Microbiology*, 31(5), pp.1407-1415.
- Aung, S., Shum, J., Abanes-De Mello, A., Broder, D., Fredlund-Gutierrez, J., Chiba, S. and Pogliano, K. (2007) Dual localization pathways for the engulfment proteins during Bacillus subtilis sporulation. *Molecular Microbiology*, 65(6), pp.1534-1546.
- AWOLLMAN/single-molecule-tools: A series of MATLAB functions and scripts for tracking and quantification of spots in images, deconvolution, segmentation and simulation., GitHub. Available at: https://github.com/awollman/single-molecule-tools (Accessed: December 20, 2022).
- Barák, I. and Muchová, K. (2018) The positioning of the asymmetric septum during sporulation in *Bacillus subtilis*. PLoS ONE 13(8): e0201979.
- Barák, I. and Youngman, P. (1996) SpoIIE mutants of Bacillus subtilis comprise two distinct phenotypic classes consistent with a dual functional role for the SpoIIE protein. *Journal of Bacteriology*, 178(16), pp.4984-4989.
- Barák, I., Behari, J., Olmedo, G., Guzmán, P., Brown, D., Castro, E., Walker, D., Westpheling, J. and Youngman, P. (1996) Structure and function of the Bacillus SpoIIE protein and its localization to sites of sporulation septum assembly. *Molecular Microbiology*, 19(5), pp.1047-1060.
- Bernard, P., Gabant, P., Bahassi, EM. and Couturier, M. (1994) Positive-selection vectors using F plasmid ccdB killer gene. *Gene* 148:71-74.
- Blaylock, B., Jiang, X., Rubio, A., Moran, C. and Pogliano, K. (2004) Zipper-like interaction between proteins in adjacent daughter cells mediates protein localization. *Genes & Development*, 18(23), pp.2916-2928.
- Burns, D. and Minton, N. (2011) Sporulation studies in Clostridium difficile. *Journal of Microbiological Methods*, 87(2), pp.133-138.
- Camp, A. and Losick, R. (2008) A novel pathway of intercellular signalling in Bacillus subtilis involves a protein with similarity to a component of type III secretion channels. *Molecular Microbiology*, 69(2), pp.402-417.
- Camp, A. and Losick, R. (2009) A feeding tube model for activation of a cell-specific transcription factor during sporulation in Bacillus subtilis. *Genes & Development*, 23(8), pp.1014-1024.

- Campo, N., Marquis, K. and Rudner, D. (2008) SpoIIQ Anchors Membrane Proteins on Both Sides of the Sporulation Septum in Bacillus subtilis. *Journal of Biological Chemistry*, 283(8), pp.4975-4982.
- Chastanet, A., and Losick, R. (2007) Engulfment during sporulation in *Bacillus subtilis* is governed by a multi-protein complex containing tandemly acting autolysins. *Mol Microbiol* 64: 139– 152.
- Chiba, S., Coleman, K. and Pogliano, K. (2007). Impact of Membrane Fusion and Proteolysis on SpoIIQ Dynamics and Interaction with SpoIIIAH. *Journal of Biological Chemistry*, 282(4), pp.2576-2586.
- Chiu, SW and Leake, M.C. (2011) Functioning nanomachines seen in real-time in living bacteria using single-molecule and Super-Resolution fluorescence imaging. *International Journal of Molecular Sciences*, 12(4), pp. 2518–2542.
- Cooper, GM. The Cell: A Molecular Approach. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. Cell Proliferation in Development and Differentiation. Available from: https://www.ncbi.nlm.nih.gov/books/NBK9906/.
- Costantini, LM. and Snapp, EL. (2013) Fluorescent proteins in cellular organelles: Serious pitfalls and some solutions. *DNA and Cell Biology*, 32(11), pp. 622–627.
- de Hoon, M., Eichenberger, P. and Vitkup, D. (2010) Hierarchical Evolution of the Bacterial Sporulation Network. *Current Biology*, 20(17), pp. R735-R745.
- Deniz, A., Mukhopadhyay, S. and Lemke, E. (2007) Single-molecule biophysics: at the interface of biology, physics, and chemistry. *Journal of The Royal Society Interface*, 5(18), pp.15-45.
- Dickson, E. and Shevky, R. (1923) BOTULISM. STUDIES ON THE MANNER IN WHICH THE TOXIN OF CLOSTRIDIUM BOTULINUM ACTS UPON THE BODY. *Journal of Experimental Medicine*, 37(5), pp.711-731.
- Doan, T., Marquis, K. and Rudner, D. (2005) Subcellular localization of a sporulation membrane protein is achieved through a network of interactions along and across the septum. *Molecular Microbiology*, 55(6), pp.1767-1781.
- Doan, T., Morlot, C., Meisner, J., Serrano, M., Henriques, A., Moran, C. and Rudner, D (2009) Novel Secretion Apparatus Maintains Spore Integrity and Developmental Gene Expression in Bacillus subtilis. *PLoS Genetics*, 5(7), p.e1000566.
- Dong, M., Masuyer, G. and Stenmark, P. (2019) Botulinum and Tetanus Neurotoxins. *Annual Review of Biochemistry*, 88(1), pp.811-837.

- Driks, A. (2002) Maximum shields: the assembly and function of the bacterial spore coat. *Trends in Microbiology*, 10(6), pp.251-254.
- Driks, A. (1999) Bacillus subtilis Spore Coat. *Microbiology and Molecular Biology Reviews*, 63(1), pp.1-20.
- Dworkin, J. and Losick, R. (2005) Developmental Commitment in a Bacterium. *Cell*, 121(3), pp.401-409.
- Errington, J. (2003) Regulation of endospore formation in Bacillus subtilis. *Nature Reviews Microbiology*, 1(2), pp.117-126.
- Fink, D., Wohrer, S., Pfeffer, M., Tombe, T., Ong, C.J. and Sorensen, P.H.B. (2010). Ubiquitous expression of the monomeric red fluorescent protein mcherry in transgenic mice. *genesis*, 48(12), pp.723–729. doi:https://doi.org/10.1002/dvg.20677.
- Fujita, M. and Losick, R. (2003) The master regulator for entry into sporulation in Bacillus subtilis becomes a cell-specific transcription factor after asymmetric division. *Genes & Development*, 17(9), pp.1166-1174.
- Galbraith, CG. and Galbraith, JA. (2011) Super-resolution microscopy at a glance, *Journal of Cell Science*, 124(10), pp. 1607–1611.
- Gauvry, E., Mathot, A., Couvert, O., Leguérinel, I., Jules, M. and Coroller, L. (2019) Differentiation of Vegetative Cells into Spores: a Kinetic Model Applied to Bacillus subtilis. *Applied and Environmental Microbiology*, 85(10).
- Gest H. (2004) The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes Rec R Soc Lond.* (2):187-201
- Gibson, DG., Young, L., Chuang, RY., Venter, JC., Hutchison, CA 3rd. and Smith HO. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* (5):343-5.
- Grainge, I. (2008) Sporulation: SpoIIIE Is the Key to Cell Differentiation. *Current Biology*, 18(18), pp. R871-R872.
- Green, DH. and Cutting, SM. (2000) Membrane topology of the Bacillus subtilis pro-sigma(K) processing complex. *Journal of Bacteriology*, 182(2), pp. 278–285
- Greenwich, J., Reverdy, A., Gozzi, K., Di Cecco, G., Tashjian, T., Godoy-Carter, V. and Chai, V. (2019). A Decrease in Serine Levels during Growth Transition Triggers Biofilm Formation in Bacillus subtilis. 201(15). doi:https://doi.org/10.1128/jb.00155-19.
- Grossman, A. and Losick, R. (1988) Extracellular control of spore formation in Bacillus subtilis. *Proceedings of the National Academy of Sciences*, 85(12), pp.4369-4373.

- Gutierrez, J., Smith, R. and Pogliano, K. (2010) SpoIID-Mediated Peptidoglycan Degradation Is Required throughout Engulfment during Bacillus subtilis Sporulation. *Journal of Bacteriology*, 192(12), pp.3174-3186.
- Harriman, O. and Leake, M. (2011) Single molecule experimentation in biological physics: exploring the living component of soft condensed matter one molecule at a time. *Journal of Physics: Condensed Matter*, 23(50), p.503101.
- Henriques, A. and Moran, C. (2007) Structure, Assembly, and Function of the Spore Surface Layers. *Annual Review of Microbiology*, 61(1), pp.555-588.
- Hidenori, S. and Henner, DJ. (1986) Construction of a single-copy integration vector and its use in analysis of regulation of the TRP operon of bacillus subtilis. *Gene*, 43(1-2), pp. 85–94.
- Higgins, M. and Piggot, P. (1992) Septal membrane fusion? a pivotal event in bacterial spore formation?. *Molecular Microbiology*, 6(18), pp.2565-2571.
- Hilbert, D. and Piggot, P. (2004) Compartmentalization of Gene Expression during Bacillus subtilis Spore Formation. *Microbiology and Molecular Biology Reviews*, 68(2), pp.234-262.
- Hoch, J. (1992) The phosphorelay signal transduction pathway in the initiation of Bacillus subtilis sporulation. *Journal of Cellular Biochemistry*, 51(1), pp.55-61.
- Hoch, J. (1993) REGULATION OF THE PHOSPHORELAY AND THE INITIATION OF SPORULATION IN BACILLUS SUBTILIS. *Annual Review of Microbiology*, 47(1), pp.441-465.
- Hoffman, R. (2005) Advantages of multi-color fluorescent proteins for whole-body and in vivo cellular imaging. *Journal of Biomedical Optics*, 10(4), p.041202.
- Hutchison, E., Miller, D. and Angert, E. (2014) Sporulation in Bacteria: Beyond the Standard Model. *Microbiology Spectrum*, 2(5).
- Janin, J., Bahadur, RP., Chakrabarti, P. (2008) Protein-protein interaction and quaternary structure. *Q Rev Biophys.* 41(2):133-80.
- Jensen, E. (2012) Use of Fluorescent Probes: Their Effect on Cell Biology and Limitations. *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology*, 295(12), pp.2031-2036.
- Jiang, X., Rubio, A., Chiba, S. and Pogliano, K. (2005) Engulfment-regulated proteolysis of SpoIIQ: evidence that dual checkpoints control σK activity. *Molecular Microbiology*, 58(1), pp.102-115.4
- Kellner, E., Decatur, A. and Moran, C. (1996) Two-stage regulation of an anti-sigma factor determines developmental fate during bacterial endospore formation. *Molecular Microbiology*, 21(5), pp.913-924.

- Kelly, A. and Salgado, P. (2019) The engulfasome in *C. difficile*: Variations on protein machineries. *Anaerobe*, 60, p.102091.
- Khanna, K., Lopez-Garrido, J. and Pogliano, K. (2020) Shaping an Endospore: Architectural Transformations During Bacillus subtilis Sporulation. *Annual Review of Microbiology*, 74(1), pp.361-386.
- Khanna, K., Lopez-Garrido, J., Zhao, Z., Watanabe, R., Yuan, Y., Sugie, J., Pogliano, K. and Villa, E. (2019) The molecular architecture of engulfment during *Bacillus subtilis* sporulation. *eLife*, 8.
- Kora, JK. (2018) Feeding the Spore: A molecular structure of the SpoIIIAG channel from Bacillus subtilis. [online] Centre for Blood Research. Available at https://cbr.ubc.ca/feeding-the-spore-a-molecular-structure-of-the-spoiiiag-channel-from-bacillus-subtilis/. [Accessed 21 May 2023].
- Leake, MC. (2010) Shining the spotlight on functional molecular complexes. *Communicative & Integrative Biology*, 3(5), pp.415-418.
- Leake, MC. (2013) The physics of life: one molecule at a time. *Philosophical Transactions of the Royal Society Biological Sciences*, 368(1611), p.20120248.
- Leake, MC., Chandler, J., Wadhams, G., Bai, F., Berry, R. and Armitage, J. (2006) Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature*, 443(7109), pp.355-358.
- Leake, MC. (2014). Analytical tools for single-molecule fluorescence imaging in cellulo, *Phys. Chem. Chem. Phys.*, 16(25), pp. 12635–12647.
- Lenn, T. and Leake, M. (2012). Experimental approaches for addressing fundamental biological questions in living, functioning cells with single molecule precision. *Open Biology*, 2(6), p.120090.
- Levdikov, V., Blagova, E., McFeat, A., Fogg, M., Wilson, K. and Wilkinson, A. (2012). Structure of components of an intercellular channel complex in sporulating Bacillus subtilis. *Proceedings of the National Academy of Sciences*, 109(14), pp.5441-5445.
- Lewis, PJ. and Marston, AL. (1998). GFP vectors for controlled expression and dual labelling of protein fusion in *Bacillus subtilis*. *GENE*. 227(1999), pp.101-109.
- Lichtman, JW. and Conchello, JA. (2005) Fluorescence microscopy. *Nat Methods.* (12):910-9.
- Lippincott-Schwartz, J. and Patterson, GH. (2003) Development and use of fluorescent protein markers in living cells. *Science.* 300(5616):87-91.
- Londono-Vallejo, J., Fréhel, C. and Stragier, P. (1997) spoIIQ, a forespore-expressed gene required for engulfment in Bacillus subtilis. *Molecular Microbiology*, 24(1), pp.29-39.

- Lopez, D., Vlamakis, H. and Kolter, R. (2009) Generation of multiple cell types in Bacillus subtilis. *FEMS Microbiology Reviews*, 33(1), pp.152-163.
- Lu, S., Cutting, S. and Kroos, L. (1995) Sporulation protein SpoIVFB from *Bacillus subtilis* enhances processing of the sigma factor precursor Pro-sigma K in the absence of other sporulation gene products. *Journal of Bacteriology* 177, 1082-1085
- Margolis, P., Driks, A. and Losick, R., (1991) Differentiation and the establishment of cell type during sporulation in Bacillus subtilis. *Current Opinion in Genetics & Development*, 1(3), pp.330-335.
- McKenney, P., Driks, A. and Eichenberger, P. (2013) The Bacillus subtilis endospore: assembly and functions of the multilayered coat. *Nature Reviews Microbiology*, 11(1), pp.33-44.
- McNamara, G., Difilippantonio, M., Ried, T. and Bieber, FR. (2017) Microscopy and image analysis. *Current Protocols in Human Genetics*, 94, 4.4.1–4.4.89.
- Meisner, J., Maehigashi, T., Andre, I., Dunham, C. and Moran, C. (2012). Structure of the basal components of a bacterial transporter. *Proceedings of the National Academy of Sciences*, 109(14), pp.5446-5451.
- Meisner, J., Wang, X., Serrano, M., Henriques, A. and Moran, C. (2008) A channel connecting the mother cell and forespore during bacterial endospore formation. *Proceedings of the National Academy of Sciences*, 105(39), pp.15100-15105.
- Moir, A. (2006) How do spores germinate?. *Journal of Applied Microbiology*, 101(3), pp.526-530.
- Molle, V., Fujita, M., Jensen, ST., Eichenberger, P., Gonzalez-Pastor, JE., Liu, JS. and Losick, R.(2003) 'The SPO0A regulon of bacillus subtilis', *Molecular Microbiology*, 50(5), pp. 1683–1701. doi:10.1046/j.1365-2958.2003.03818.x.
- Morlot, C. and Rodrigues, C. (2018) The New Kid on the Block: A Specialized Secretion System during Bacterial Sporulation. *Trends in Microbiology*, 26(8), pp.663-676.
- Narula, J., Fujita, M. and Igoshin, O. (2016) Functional requirements of cellular differentiation: lessons from Bacillus subtilis. *Current Opinion in Microbiology*, 34, pp.38-46.
- Nguyen, A. and Tallent, S. (2019) Screening food for Bacillus cereus toxins using whole genome sequencing. *Food Microbiology*, 78, pp.164-170.
- Nicholson, W., Munakata, N., Horneck, G., Melosh, H. and Setlow, P. (2000) Resistance of Bacillus Endospores to Extreme Terrestrial and Extraterrestrial Environments. *Microbiology and Molecular Biology Reviews*, 64(3), pp.548-572.

- Paredes-Sabja, D., Shen, A. and Sorg, J. (2014) Clostridium difficile spore biology: sporulation, germination, and spore structural proteins. *Trends in Microbiology*, 22(7), pp.406-416.
- Payne-Dwyer, AL., Syeda, AH., Shepherd, JW., Frame, L. and Leake, MC. (2022) RecA and RecB: probing complexes of DNA repair proteins with mitomycin C in live *Escherichia coli* with single molecule sensitivity. *J. R. Soc. Interface.* 19: 20220437.
- Plank, M., Wadhams, GH. and Leake, MC. (2009) Millisecond timescale Slimfield imaging and automated quantification of single fluorescent protein molecules for use in probing complex biological processes. *Integr Biol (Camb).* 1(10):602-12.
- 'Protein interfaces, surfaces and assemblies' service PISA at the European Bioinformatics Institute. (http://www.ebi.ac.uk/pdbe/prot_int/pistart.html), Krissinel, E and Henrick, K. (2007) *Inference of macromolecular assemblies from crystalline state.* J. Mol. Biol. 372, 774--797.
- Regan, G., Itaya, M. and Piggot, P. (2012) Coupling of σ G Activation to Completion of Engulfment during Sporulation of Bacillus subtilis Survives Large Perturbations to DNA Translocation and Replication. *Journal of Bacteriology*, 194(22), pp.6264-6271.
- Riley, E., Schwarz, C., Derman, A. and Lopez-Garrido, J. (2021) Milestones in Bacillus subtilis sporulation research. *Microbial Cell*, 8(1), pp.1-16.
- Rodrigues, C., Henry, X., Neumann, E., Kurauskas, V., Bellard, L., Fichou, Y., Schanda, P., Schoehn, G., Rudner, D. and Morlot, C. (2016) A ring-shaped conduit connects the mother cell and forespore during sporulation in Bacillus subtilis. *Proceedings of the National Academy of Sciences*, 113(41), pp.11585-11590.
- Rodrigues, C., Marquis, K., Meisner, J. and Rudner, D. (2013) Peptidoglycan hydrolysis is required for assembly and activity of the transenvelope secretion complex during sporulation in Bacillus subtilis. *Molecular Microbiology*, 89(6), pp.1039-1052.
- Rodrigues, C., Ramírez-Guadiana, F., Meeske, A., Wang, X. and Rudner, D. (2016) GerM is required to assemble the basal platform of the SpoIIIA-SpoIIQ transenvelope complex during sporulation in Bacillus subtilis. *Molecular Microbiology*, 102(2), pp.260-273.
- Rubio, A. and Pogliano, K. (2004) Septal localization of forespore membrane proteins during engulfment in Bacillus subtilis. *The EMBO Journal*, 23(7), pp.1636-1646.
- Ryter A. (1965) Étude morphologique de la sporulation de *Bacillus subtilis*. *Ann Inst Pasteur*. 108:40–60.
- Sample, V., Newman, R. and Zhang, J. (2009) The structure and function of fluorescent proteins. *Chemical Society Reviews*, 38(10), p.2852.

- Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmaier, O. and Drummen GPC. (2019) Super-resolution microscopy demystified. *Nat Cell Biol.* 21(1):72-84.
- Seccareccia, I., Kovacs, AT., Gallegos-Monterrosa, R. and Nett, M. (2016) 'Unraveling the predator-prey relationship of cupriavidus necator and bacillus subtilis', *Microbiological Research*, 192, pp. 231–238. doi:10.1016/j.micres.2016.07.007.

- Setlow, P. (2003) Spore Germination. *Current Opinion in Microbiology*, 6, pp.550-556.
- Setlow, P. (2011) Dynamics of the assembly of a complex macromolecular structure - the coat of spores of the bacterium Bacillus subtilis. *Molecular Microbiology*, 83(2), pp.241-244.
- Setlow, P. (2014) Germination of Spores of Bacillus Species: What We Know and Do Not Know. *Journal of Bacteriology*, 196(7), pp.1297-1305.
- Shashkova, S., Andersson, M., Hohmann, S. and Leake, MC. (2021) Correlating single-molecule characteristics of the yeast aquaglyceroporin Fps1 with environmental perturbations directly in living cells. *Methods*, 193, pp. 46-53.
- Shashkova., S. and Leake, MC. (2007) Single-molecule fluorescence microscopy review: shedding new light on old problems. *Biosci Rep.* 37 (4): BSR20170031.
- Shaw M., Zajiczek, L. and O'Holleran K. (2015) High speed structured illumination microscopy in optically thick samples *Methods* 88 11–9.
- Shepherd, JW., Payne-Dwyer, A., Lee, JE., Syeda, A. and Leake, MC. (2021) Combining single-molecule super-resolved localization microscopy with fluorescence polarization imaging to study cellular processes. J *Phys. Photonics* 3 034010.
- Søgaard-Andersen, L. (2013) Stably bridging a great divide: localization of the SpoIIQ landmark protein in Bacillus subtilis. *Molecular Microbiology*, 89(6), pp.1019-1024.
- Sonenshein, A. (2000) Control of sporulation initiation in Bacillus subtilis. *Current Opinion in Microbiology*, 3(6), pp.561-566.
- Stephens, C. (1998) Bacterial sporulation: A question of commitment?. *Current Biology*, 8(2), pp. R45-R48.
- Stokes, GG., Trans, P. and Lond RS. (1852) On the change of refrangibility of light. *Philos. Trans. R. Soc. Lond*. 142, 463-562.
- Sun, Y., Sharp, M. and Pogliano, K. (2000) A Dispensable Role for Forespore-Specific Gene Expression in Engulfment of the Forespore during Sporulation of Bacillus subtilis. *Journal of Bacteriology*, 182(10), pp.2919-2927.

- *Super resolution localization microscopy* (2022) *Teledyne Photometrics*. Available at: https://www.photometrics.com/learn/super-resolution-microscopy/super-resolution-localization-microscopy (Accessed: May 6, 2023).
- Super-resolution structured illumination microscopy (SR-SIM) Cherry Biotech. Available at: https://www.cherrybiotech.com/scientific-note/microscopy/super-resolution-structured-illumination-microscopy-sr-sim (Accessed: December 20, 2022).
- Syeda, AH., Wollman, AJM., Hargreaves, AL., Howard, JAL., Brüning, JG., McGlynn, P. and Leake MC. (2019). Single-molecule live cell imaging of Rep reveals the dynamic interplay between an accessory replicative helicase and the replisome. *Nucleic Acids Res.* 47(12):6287-6298.
- Takamatsu, H. and Watabe, K. (2002) Assembly and genetics of spore protective structures. *Cellular and Molecular Life Sciences (CMLS)*, 59(3), pp.434-444.
- Technical note - image restoration through deconvolution (2022) *Teledyne Photometrics*. Available at: https://www.photometrics.com/learn/spinning-disk-confocal-microscopy/image-restoration-through-deconvolution (Accessed: December 22, 2022).
- Tocheva, E., Lopez-Garrido, J., Hughes, H., Fredlund, J., Kuru, E., VanNieuwenhze, M., Brun, Y., Pogliano, K. and Jensen, G. (2013). Peptidoglycan transformations during Bacillus subtilis sporulation. *Molecular Microbiology*, 88(4), pp.673-686.
- Tocheva, E., Ortega, D. and Jensen, G. (2016) Sporulation, bacterial cell envelopes and the origin of life. *Nature Reviews Microbiology*, 14(8), pp.535-542.
- Toseland, CP. (2013) Fluorescent labelling and modification of proteins. *J Chem Biol.* 6(3):85-95.
- Trach, K., Burbulys, D., Strauch, M., Wu, J., Dhillon, N., Jonas, R., Hanstein, C., Kallio, P., Perego, M., Bird, T., Spiegelman, G., Fogher, C. and Hoch, J. (1991) Control of the initiation of sporulation in Bacillus subtilis by a phosphorelay. *Research in Microbiology*, 142(7-8), pp.815-823.
- Vangindertael, J., Camacho, R., Sempels, W., Mizuno, H., Dedecker, P. and Janssen, KPF. (2018) An introduction to optical super-resolution microscopy for the adventurous biologist. *Methods Appl Fluoresc.* 6(2):022003.
- Wollman, A., Miller, H., Zhou, Z. and Leake, M. (2015) Probing DNA interactions with proteins using a single-molecule toolbox: inside the cell, in a test tube and in a computer. *Biochemical Society Transactions*, 43(2), pp.139-145.
- Wollman, A., Muchová, K., Chromiková, Z., Wilkinson, A., Barák, I. and Leake, M. (2020) Single-molecule optical microscopy of protein dynamics and computational analysis of

images to determine cell structure development in differentiating Bacillus subtilis. *Computational and Structural Biotechnology Journal*, 18, pp.1474-1486.

- Xing, Y., and Harper, W. (2020) Bacillus spore awakening: recent discoveries and technological developments. *Current Opinion in Biotechnology*, 64, pp.110-115.
- Xue, Q., Harriman, O. and Leake, M. (2011) Using bespoke fluorescence microscopy to study the soft condensed matter of living cells at the single molecule level. *Journal of Physics: Conference Series*, 286, p.012001.
- Zeiss Elyra 7 with lattice sim² super-resolution microscope. Available at: https://www.zeiss.com/microscopy/en/products/light-microscopes/widefield-microscopes/elyra-7.html (Accessed: December 20, 2022).
- Zeytuni, N. and Strynadka, N. (2019) A Hybrid Secretion System Facilitates Bacterial Sporulation: A Structural Perspective. *Microbiology Spectrum*, 7(1).
- Zeytuni, N., Hong, C., Flanagan, K., Worrall, L., Theiltges, K., Vuckovic, M., Huang, R., Massoni, S., Camp, A., Yu, Z. and Strynadka, N. (2017) Near-atomic resolution cryoelectron microscopy structure of the 30-fold homooligomeric SpoIIIAG channel essential to spore formation in *Bacillus subtilis. Proceedings of the National Academy of Sciences*, 114(34), pp. E7073-E7081.