INVESTIGATING THE LIPOPROTEIN BIOSYNTHETIC PATHWAY IN STREPTOMYCES SPECIES.

Benjamin John Thompson

A thesis submitted in fulfilment of the requirements for a degree of Doctor of Philosophy at the University of East Anglia.

School of Biological Sciences, University of East Anglia.

June 2010

PUBLICATIONS ARISING FROM WORK IN THIS THESIS

Thompson, B.J., Widdick, D. A., Hicks, M. G., Chandra, G., Sutcliffe, I. C., Palmer, T. and Hutchings, M. I. (2010). Investigating lipoprotein biogenesis and function in the Gram-positive bacterium *Streptomyces coelicolor*. *Mol Microbiol* Accepted Article doi: 10.1111/j.1365-2958.2010.07261.x

The *Streptomyces coelicolor* work discussed in this thesis is also discussed in this paper. A manuscript covering the *S. scabies* work in the thesis is currently being prepared for submission to *Molecular Microbiology*.

Abstract

Lipoproteins are lipid modified proteins attached to the cytoplasmic membrane of both Gram-negative and Gram-positive bacteria. In Gram-negative bacteria, lipoproteins are produced in a four stage process. Firstly the protein is translocated across the cytoplasmic membrane, primarily by the Sec general secretory pathway, then lipidated at a conserved cysteine by the enzyme Lipoprotein diacylglyceryl transferase (Lgt). The signal peptide is cleaved by Lipoprotein signal peptidase II (Lsp), to leave the lipid modified cysteine at the N-terminus. Finally the lipoprotein is N-acylated by Lipoprotein N-acyl transferase (Lnt). This pathway is essential and occurs in the order listed. However the pathway in all Gram-positive bacteria tested to date is non-essential, and in low GC Gram-positive bacteria the Lnt enzyme is absent. This work concentrates on the lipoprotein biosynthetic pathway of *Streptomyces coelicolor*, the model organism for the high GC branch of Gram-positive bacteria and *Streptomyces scabies,* a pathogen of potatoes. Each of the genes encoding the enzymes listed above were disrupted to assess their importance to the bacteria. Strong evidence is shown in both species that the Lsp (and probably the Lgt) enzyme is essential, not seen in other Gram-positive bacteria. Evidence from the *S. scabies* work suggests the pathway occurs in a strict, regimented order. Both results indicate the lipoprotein biosynthetic pathway of *Streptomyces* is closer to that of the Gram-negative bacteria, than that of the other Gram-positives tested to date. This work also provides the first insights into the role of the lipoprotein biosynthetic pathway in plant pathogenesis. Studies in animal pathogens have shown that disrupting the pathway can have a variety of effects, ranging from avirulence to hypervirulence. This work shows that disrupting the lipoprotein processing enzymes has a moderate effect of the virulence of *S. scabies* on both potato tubers and radish seedlings.

Table of Contents

List of Figures and Tables

Appendix

Abbreviations

APS	ammonium persulphate
bp	base pairs
$CaCl_2$	calcium chloride
cm	centimeter
DNase	deoxyribonuclease
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotides
DMSO	dimethyl sulphoxide
dH_2O	distilled water
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
H_2O	water
HRP	horseradish peroxidase
HCl	hydrochloric acid
Kb	kilobase
kHz	kilohertz
kV	kilovolt
µF	microfarad
µg	microgram
µl	microlitre
µM	micromolar
mg	milligram
ml	microlitre
mm	millimetre
mM	millimolar
M	molar

NO	nitric oxide
nt	nucleotide
Ω	ohm
ORF	open reading frame
OD	optical density
PAI	pathogenicity island
pmole	picomole
PCR	polymerase chain reaction
PDVF	polyvinylidene difluoride
rpm	revolutions pre minute
RNase	ribonuclease
RNA	ribonucleic acid
SEM	scanning electron microscopy
NaCl	sodium chloride
SDS	sodium dodecyl sulphate
NaOH	sodium hydroxide
TEMED	tetramethylethylenediamine
TBE	Tris Borate EDTA
TBS	Tris buffered saline
TE	Tris EDTA
TGS	Tris Glycine SDS
TCB	Tris ultracentrifuge buffer
V	Volt
v/v	volume per volume
w/v	weight per volume

Acknowledgements

I extend my eternal thanks and gratitude to Matt Hutchings, for his amazing patience and guidance, without whom this work would never have been completed. Special thanks goes to Iain Sutcliffe, for his invaluable input from distance, and David Widdick for always being available to answer my stupid questions and for fixing the unfixable. Thanks to Matthew Hicks for his encyclopaedic knowledge of techniques and reagents. Many thanks go to the rest of the team: Ryan, Joerg, Felicity and Elaine, without whom the lab would have been a much less enjoyable place to be.

For their assistance with this project I would like to thank Govind Chandra for his help with all things computer based, Paul Thomas and Kim Findlay for teaching me how to use very expensive microscopes, and Dawn Bignell and Rose Loria for their assistance with the *S. scabies* work. I would also like to thank the BBSRC for their funding and the British Library for helping me find somewhere quiet to write up and preventing me from procrastinating too much.

Many thanks go to my family for their continuing support and encouragement, my dear friends at UEA and beyond, and to Sarah Hesketh for nagging/ encouraging (1-1 in book terms now though!).

Chapter 1 - Introduction

1.1 Introduction

Bacteria live in a variety of different environments. As such, they are subjected to a multitude of stresses to which they must sense and respond in order to survive. Both branches of bacteria, Gram-positive and Gram-negative, contain proteins localised within their membranes which can interact with the surrounding environment. This thesis will summarise work carried out on a specialised class of membrane proteins known as lipoproteins in the Gram-positive bacteria *Streptomyces coelicolor* and *Streptomyces scabies*.

1.2 Protein Translocation

The translocation of proteins across the cell membrane is essential for all cellular life. Although all proteins are synthesised cytoplasmically, many have extra-cellular functions, be they metabolic or structural, and their transport from the inside to the outside of the cell must be tightly regulated. There are two main translocation routes for proteins across the cytoplasmic membrane, i) The Sec (General Secretory) pathway and ii) The Tat (Twin Arginine Translocation) pathway.

1.2.1 The Sec Pathway

The Sec pathway has been studied for several decades, and is relatively well understood. This translocation pathway is involved in the secretion of unfolded proteins, as well as the insertion of proteins into the cell membrane [1]. The Sec pathway is found in all organisms, including the Endoplasmic Reticulum of eukaryotic cells and the thylakoid membranes of plant chloroplasts. Proteins

are targeted to the Sec machinery by a Sec signal sequence which is also conserved in all domains of life. The sequence is usually about twenty amino acids in length and does not contain any conserved sequence motifs. Instead there are three distinct regions: i) the *n*-region which consists of positively charged amino acids ii) the *h*-region consisting of hydrophobic amino acids, iii) the *c*-region or cleavage region (fig. 1.2), which is necessary for the action of signal peptidases.

The Sec machinery consists of several proteins: SecY*,* SecE and SecG form the hydrophilic protein conducting channel (PCC) and are associated with SecA, the motor protein which contains an ATPase domain. SecD and SecF are required for efficient translocation. Of these proteins, SecYEA are essential.

There are two methods of Sec mediated translocation, co-translational and post-translational. Co-translational protein translocation in bacteria is used by only a few secretory proteins [1], and is mainly used to insert integral membrane proteins into the membrane. In this method the Sec signal sequence of the protein to be translocated is recognised by the Signal Recognition Particle (SRP) whilst it is emerging from the ribosome. This entire complex is then transferred to the Sec machinery with the aid of the SRP receptor (fig. 1.1). The post-translational translocation in Gram-negative bacteria requires an additional chaperone protein, which in *E. coli* is named SecB. SecB is a secretion-specific chaperone which binds to aromatic and basic regions of the preprotein as it is emerging from the ribosome, stabilising it, and preventing folding. SecB interacts with SecA in the SecYEGA complex and passes the unfolded protein to SecA to allow for translocation. SecB is not essential for Sec translocation. It

appears that in its absence, the Sec signal peptide is enough to target the unfolded protein to the Sec machinery [1]. As SecB is only found in proteobacteria, it seems likely that bacteria lacking in this protein either rely on the signal peptide, or utilise other chaperones.

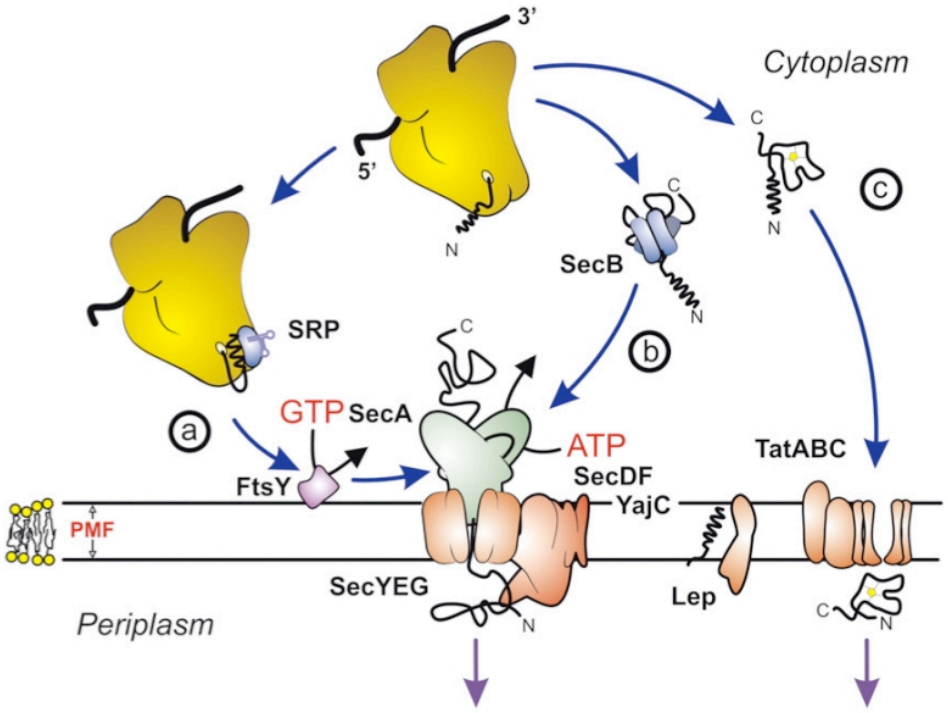


Figure 1.1 An overview of Sec and Tat dependent translocation. Sec dependent translocation can be co-translational (a), or post-translational (b) and involves unfolded proteins, whilst Tat dependent translocation (c) involves full folded proteins. From [1].

1.2.2 The Tat Pathway

The Tat pathway is used to transfer fully folded proteins across the cell membrane. Whilst the Sec system is ubiquitous, it appears that the Tat system is only encoded in about half of the currently sequenced bacterial genomes [2], although it must be stressed that there is a bias in the number of bacterial species sequenced towards the Gram-negatives, so this figure may not represent the true number of species containing the Tat system. Proteins which

are translocated by Tat can be identified by their signal sequence, which is larger than the Sec sequence, but still contains the tripartite structure described above. Unlike the Sec signal peptide however, the Tat signal peptide contains a very specific sequence which gives the system its name. A comparison of the two signal peptides can be seen in figure 1.2. The amino acid sequence spanning the division between *n* and *h* region contains the motif Ser/Thr-Arg-Arg-X-Φ-Φ (where X can be any amino acid and Φ is a hydrophobic residue) [3]. This sequence is essential for translocation, and the twin arginine residues are almost always conserved. To date only two Tat substrates have been found that lack the Arg-Arg motif [4]. The Tat system is able to translocate proteins that contain cofactors, which need to be folded before translocation. A molybdenum cofactor containing subunit of DMSO Reductase in *Shewanella oneidensis*, DmsA, has been shown to contain a twin arginine motif in its leader sequence [5], as has the HysB subunit of [NiFeSe] hydrogenase from *Desulfovibrio vulgaris* [6].

In *E. coli*, the Tat system consists of three proteins: TatA, TatB and TatC. A homologue of TatA, TatE, also exists but is encoded elsewhere in the genome. The transmembrane pore in the Tat complex is formed by the TatA protein. TatA is 89 amino acids in length and its functional domain is a single N-terminal transmembrane helix [4]. TatA is only approximately 9kDa in size, but individual TatA monomers are able to interact via their transmembrane domains to form complexes estimated to be up to 750kDa in size [7].

The TatC protein is 258 amino acids in length and contains six transmembrane helices. Both the N and C termini are located in the cytoplasm. This protein is

the most conserved Tat subunit throughout the bacterial kingdom [8]. TatC is essential for the formation of the Tat complex and is able to bind both the signal peptide of the protein to be transported and TatB. The structure of TatB is similar to that of TatA but its function is somewhat of a mystery. Many Gram-Positive bacteria, such as *B. subtilis,* lack TatB and are still able to export proteins via the Tat pathway. In *E. coli* however, the loss of TatB leads to a disruption of the translocation of some proteins, but not others [4, 7].

It is currently thought that in *E. coli*, when at rest, TatB and C from a complex which is ready to accept the Tat signal sequence. Once the TatBC-substrate complex is formed, TatA monomers are recruited, and the protein is translocated. Once this occurs the TatABC complex rapidly dissociates back to its resting state [7]. What is not known currently is whether a ready sized TatA pore is selected depending on the size of the substrate, or whether the TatA monomers form a pore when required. The exact role of TatB is yet to be elucidated. It is possible that this protein acts as a bridge between the TatC bound substrate and the TatA pore [3, 7].

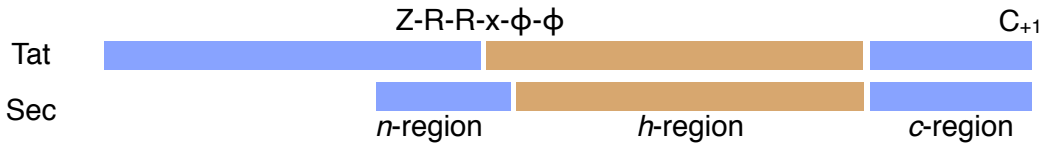


Figure 1.2 Comparison of the Tat and Sec secretion N-terminal signal sequences. The characteristic Twin arginine motif of a Tat substrate is shown. The cysteine at position +1 represents the first amino acid of a mature lipoprotein. Based upon [1].

1.3 Lipoproteins

In bacteria the addition of a lipid molecule to a protein to form a lipoprotein is tightly controlled as part of a multi-step reaction which occurs after translocation across the cytoplasmic membrane. This pathway appears to be unique to the Bacteria even though lipid modified proteins are also found in Archaea and Eukaryotes [9]**.** Most of the work on bacterial lipoproteins has been carried out in the Gram-negative bacterium *Escherichia coli*. The first lipoprotein discovered was the murein lipoprotein from *E. coli* also known as Braun's lipoprotein in 1969 [10].

The leader sequences of lipoproteins closely resemble those of other Sec transported proteins with the exception of an extra motif known as the lipobox. As before, the signal sequence contain three distinct regions: i) The *n*-region ii) The *h*-region iii) The *c*-region. The *c*, or cleavage, region contains the lipobox motif. This motif is essential for correct lipoprotein processing and contains a four amino acid sequence as follows: L_{-3}-[A/S/T]$_{-2}$-[G/A]$_{-1}$-C_{+1} [11].

The only invariant here is the cysteine residue at +1. The other three amino acids are occasionally different, but the frequency at which others occur is much lower [12]. The cysteine is invariant as it contains the sulphydryl group to which the diacylglyceride (lipid) moiety is covalently attached. It is labelled +1 as it will become the first amino acid (the N-terminus) of the mature lipoprotein. It has been noted that the lipoprotein signal sequence appears to be shorter than that of exported proteins that are not lipidated [13].

The production of a mature lipoprotein in Gram-negative bacteria is a four-stage process. The first stage involves export of the protein to be lipidated via the Sec (or Tat) pathway. The next three steps are processing steps which occur on the extracytoplasmic face of the membrane, each requiring a separate enzyme: i) Lipoprotein diacylglyceryl transferase (Lgt) ii) Lipoprotein signal peptidase II (Lsp) iii) Lipoprotein N-acyl transferase (Lnt). The pathway is tightly controlled and occurs strictly in the order listed [14] (figs. 1.3, 1.4). In Gram-positive bacteria, the pathway is different, due to the lack of an outer membrane.

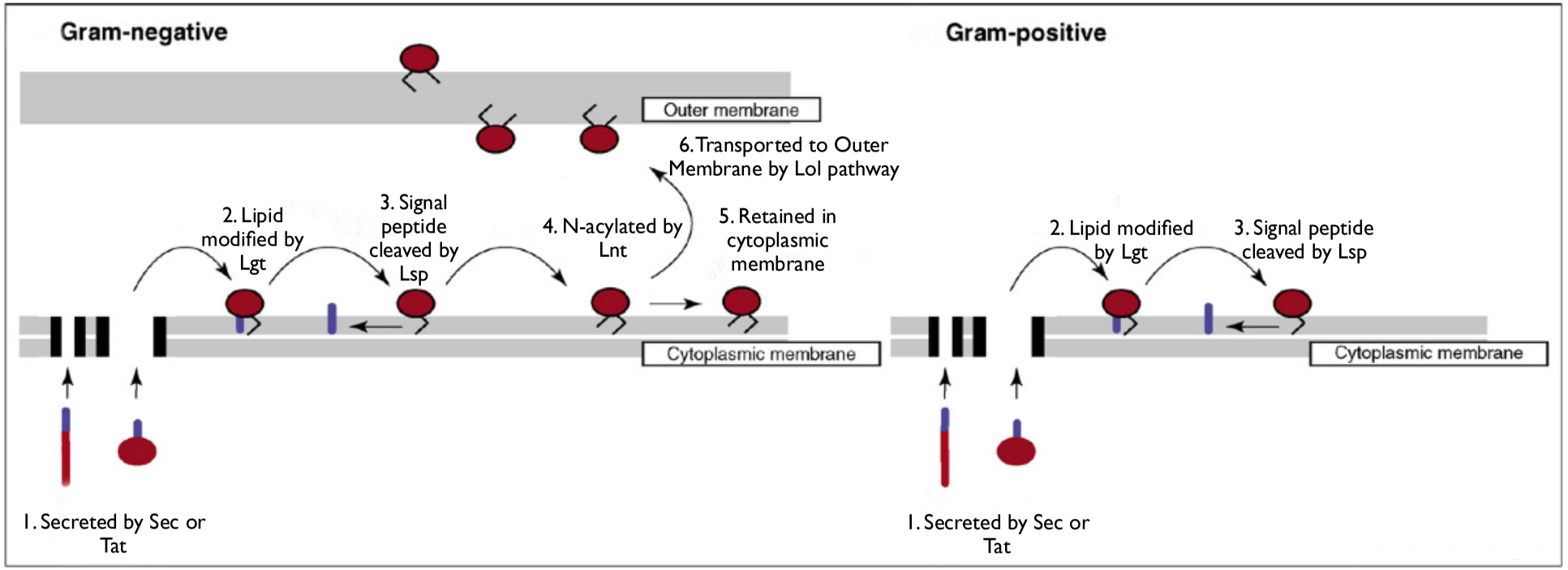


Figure 1.3 A comparison of Gram-negative and Gram-positive lipoprotein biosynthesis. From [11].

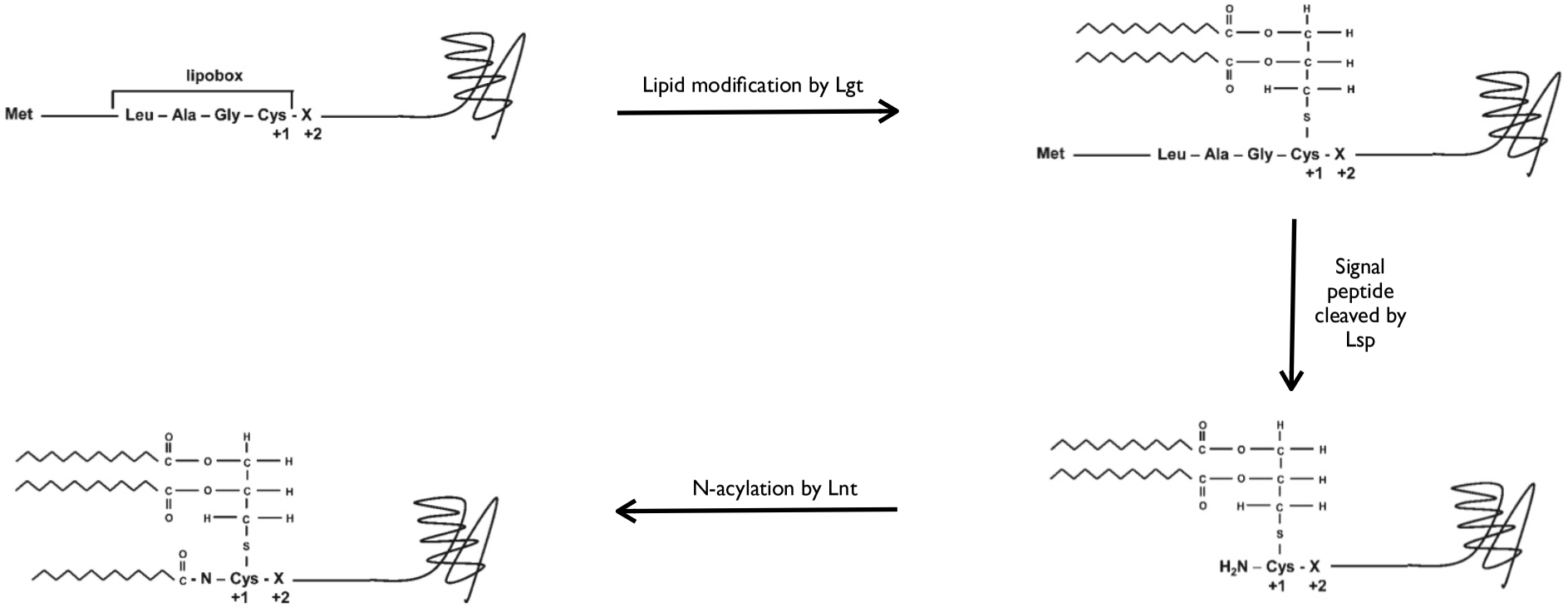


Figure 1.4 The lipoprotein biosynthetic pathway. From [18].

1.3.1 Lipoprotein diacylglyceryl transferase

The first enzyme in the pathway, Lgt, is responsible for adding the diacylglyceryl molecule to the cysteine in the lipobox via a thioether linkage. In *E. coli* it has been shown that this lipid moiety is transferred from phosphatidyl glycerol (PG) to the sulphhydrl group of the unmodified immature liprotein [12]. Lgt is a membrane bound enzyme which is highly conserved throughout the bacterial kingdom, with the structure containing large sections of hydrophobic amino acids separated by short hydrophilic sections. The large amounts of arginine and lysine lead to a net basic charge for the enzyme [15], estimated to have a pK_a value ~10 [16]. It is thought that the basic charge is important for interaction with acidic membrane phospholipids [14]. In *E. coli* a histidine residue at position 103 has been shown to be essential for enzyme activity and two other amino acids, His-196 and Tyr-235, important for function [17]. The generally accepted theory is that Lgt lipidates an immature lipoprotein that is anchored in the cytoplasmic membrane by its signal peptide. This would explain the presence of the hydrophobic region in the signal peptide. Confusingly however, a recent paper has suggested that the *h*-regions of some signal peptides may be more hydrophilic than previously thought and that the Lgt enzyme may be located on the inner face of the cytoplasmic membrane [16].

Lgt catalysed lipidation of the immature lipoproteins has been shown in various bacteria, including *L. monocytogenes*, *S. aureus*, *and S. equi*. *Δlgt* mutants grown in the presence of [^{14}C]-palmitic acid had no detectable lipoproteins in their membrane fractions, whereas they were detectable in the wild type controls [15, 18, 19].

1.3.2 Lipoprotein signal peptidase II

Lsp, also known as signal peptidase II, is the second enzyme in the pathway, and is a transmembrane protein responsible for removing the signal sequence from exported immature lipoproteins [19]. In eubacteria such as *B. subtilis* there is only a single copy of the *lsp* gene, whereas there are multiple signal peptidase I genes [20], which cleave the signal sequences from all non-lipoproteins translocated via Tat or Sec. This appears to be the case in the majority of bacterial species, although there do appear to be some examples of an organism having more than one *lsp* paralogue, such as in *Myxococcus xanthus* [21]. In *E. coli* the *lsp* gene is encoded in an operon with four other genes and is cotranscribed with *ileS*, which encodes the isoleucyl-tRNA synthetase [22], although there seems to be no obvious physiological connection. This cotranscription is conserved in a number of Gram-negative bacteria as are the other downstream open reading frames (ORF) [19]. In Gram-positive bacteria *lsp* and *ileS* are not found in the same operon [21]. To date an Lsp homologue has not been identified in eukaryotes or archaea but it has been shown that signal cleavage of lipoproteins in archaea is inhibited by globomycin, a potent inhibitor of Lsp [20, 23]. It has been shown that in the majority of bacteria, both Gram-positive and Gram-negative, a lipoprotein must be lipidated by Lgt before Lsp can recognise its substrate [20]. However, there are exceptions to this rule and *L. monocytogenes* Lsp can act on non-lipidated lipoproteins leading to their mass release into the growth medium [24]. This enzyme cleaves the signal sequence just prior to the +1 cysteine, which then becomes the N-terminal residue of the mature protein.

The sizes of the *lsp* genes sequenced so far encode Lsp enzymes ranging from 154 amino acids in *B. subtilis* to 181 amino acids *M. genitalium* [25]. Computer analysis of the protein suggests it contains four membrane spanning domains [19], consistent with the idea that this is a membrane-bound protein. The enzyme also contains several conserved amino acids that are necessary for function, including two aspartic acid residues which may form a catalytic dyad in the active site of the enzyme [19]. These residues are all predicted to be located on the external face of the cytoplasmic membrane in *B. subtilis* [26]. Indeed Lsp lacks the amino acids or associated metal ions required by several classes of proteases and is inhibited by Pepstatin, an inhibitor of aspartate peptidases, indicating that the enzyme belongs to this group, although it does not share the active sites found in plant or viral aspartate petidases, suggesting that Lsp may represent a novel class of these enzymes [26].

Many of the studies carried out on Lsp have been made possible by the action of the cyclic peptide antibiotic, globomycin. This antibiotic acts as a non-competitive analogue inhibitor of Lsp [27], and is lethal to Gram-negative bacteria as it leads to the accumulation of incorrectly processed lipoproteins in the inner membrane. Globomycin is lethal to *E. coli* at a concentration of 20-40 $\mu g\ ml^{-1}$ [28]. Overexpression of Lsp leads to increased resistance to globomycin, and this has been used to determine whether *lsp* genes from other bacteria expressed in *E. coli* can raise the level of globomycin resistance and therefore be shown to be active Lsp homologues as has been demonstrated in a variety of different species [19, 21, 25, 28].

1.3.3 Lipoprotein N-acyl transferase

The final enzyme involved in the maturation of a lipoprotein is Lnt. The function of this protein is to N-acylate the amino group found in the C_{+1} cysteine of the lipoprotein. Lnt was initially thought to be unique to Gram-negative bacteria as it had not been found in most Gram-positive bacteria, including *B. subtilis*, the model low GC Gram-positive bacterium [29]. However, recent genome sequence analysis has revealed that some members of the actinomycetes, the high-GC branch of Gram-positive bacteria, contain homologues of Lnt [30]. It is believed to be a member of the nitrilase superfamily of enzymes, which hydrolyse carbon-hydrogen bonds [30, 31].

Before Lnt can act on a lipoprotein it must first have its signal sequence removed by Lsp giving the strict sequence Lgt->Lsp->Lnt [14]. Once the signal peptide has been removed, Lnt catalyses the addition of an acyl group to the +1 cysteine residue in the lipoprotein [12]. In *E. coli* this acyl group is typically donated by phosphatidylglycerol, although phosphatidylethanolamine and cardiolipin can also be used [32].

Using sequence alignments and site-directed mutagenesis, in *E. coli* seven conserved amino acid residues have been discovered, three of which form a potential catalytic triad (E267, K335 and C387), and four which are found within the periplasm (W237, E343, Y388 and E389) [30]. Lnt is thought to be one of the few examples of a protein that exists in its reaction intermediary form *in vivo* [33]. The reason for this seems to be that due to the large number of lipoproteins found in a typical *E. coli* cell (~10^6 of the murein lipoprotein at any one time) there needs to be a rapid turn over of fatty acids, which has been

estimated at >2 x 10^4 min^{-1} [33].

As mentioned above, Lnt homologues have been identified in a number of Gram-positive bacteria. Intriguingly, indirect evidence of N-acylated lipoproteins has been discovered in *B. subtilis* [34], and in *S. aureus* the lipoprotein SitC has been shown to be triacylated, whereas others have been shown to only be diacylated [35, 36]. Both of these low GC species lack any homologues of *E. coli* Lnt, indicating that there may be other types of N-acetyltransferases currently unknown.

To date, only two confirmed Gram-positive Lnt proteins exist, found in *M. smegmatis* and *M. tuberculosis.* Homologues of this protein are found in several members of the Mycobacteriaceae. The Lnt from *M. tuberculosis* is interesting, as it has fused with a separate protein to become a two domain enzyme, known as Ppm1, a polyprenol monophosphomannose synthase. The Ppm1 domain is found at the c-terminus, and in *M. smegmatis* and *M. leprae* the two enzymes Lnt and Ppm are encoded by separate genes [37]. As well as acting as N-acyl transferase, Ppm1 is involved with the biosynthesis of lipomannan (LM) and lipoarabinomannan (LAM), cell envelope glycolipids involved in the infection process [37]. The confirmation of triacylation of lipoproteins in both species was carried out using the confirmed *M. tuberculosis* lipoprotein LppX expressed in *M. smegmatis*. Mass spectrometry of a trypsin digested LppX from a wild-type strain showed a much larger mass at the N-terminal cysteine than the predicted mass of the peptide, confirming there was an N-terminal modification. A *Δlnt* strain showed a mass smaller than the wild-type by a size corresponding to a C16 fatty acid. LppX run on an SDS-Page gel

was 0.3 kDa larger in a wild-type strain than in a *Δlnt* stain [38]. The *Δlnt* strain of *M. smegmatis* could be rescued by *M. tuberculosis* Ppm1, suggesting that part of it has the same function. Neither *M. smegmatis lnt* nor *ppm1* were able to rescue a *Δlnt* strain of *E. coli* [38]. The same can be said for the putative *lnt* genes from *Streptomyces coelicolor* and *Corynebacterium glutamicum* [30]. All of these homologues are missing the essential W237 and Y388 found in *E. coli* although the rest of the active site residues are conserved, and it is speculated that these differences are due to an altered substrate for each of the enzymes, given their difference in membrane fatty acids [30, 33].

1.3.4 The Lol System

In Gram-negative bacteria, following the action of Lnt, mature lipoproteins either remain anchored to the cytoplasmic membrane or, for the majority, are transported through the hydrophilic periplasm to the outer membrane via the lipoprotein localisation (Lol) transport system. The Lol system is comprised of five genes, ordered in the operon *lolABCDE*. The *lolCDE* genes encode an ATP-Binding Cassette (ABC) transporter consisting of two LolD, one LolC and one LolE proteins, whilst the two remaining genes encode the proteins LolA and LolB, responsible for transporting a lipoprotein across the periplasm and anchoring it into the outer membrane, respectively. The system is essential in *E. coli*, and is highly conserved in Gram-negative bacteria, although LolB does appear to be absent from several species [39]. The Lol system is absent in all Gram-positive bacteria checked to date [40], although a potential homologue of LolD has been found in *S. coelicolor.*

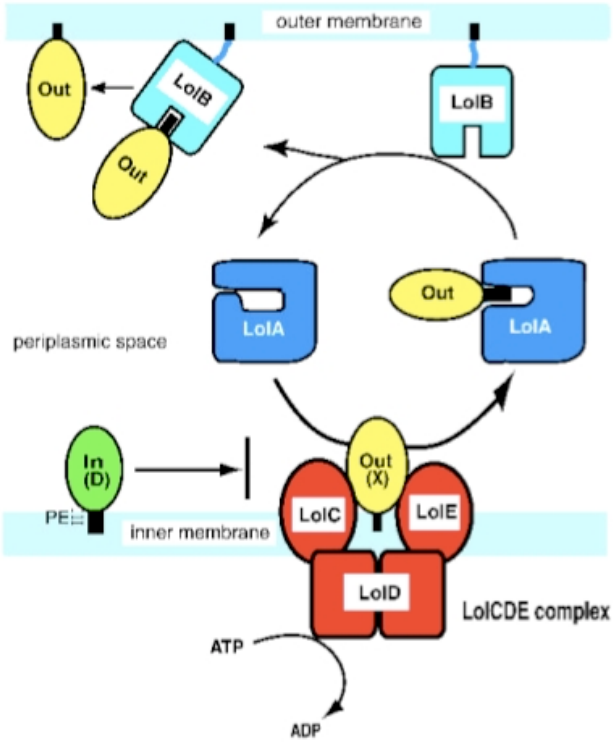


Figure 1.5 The Lol system from *E. coli* [13], the inner membrane retention +2 Aspartate (D) can be seen.

The LolCDE complex differs from other ABC transporters as it is not involved in the transmembrane transfer of a substrate and contains fewer membrane spanning domains (8) than ABC transporters usually have (≥10). It is one of only two essential ABC transporters found in *E. coli* [39]. To transport a lipoprotein across the periplasm, it is first accepted by the LolCDE complex, with LolD acting as the ATPase, releasing it from the cytoplasmic membrane. Since lipoproteins are hydrophobic, a chaperone is needed to move it across the periplasm. The LolA protein acts as the chaperone. This protein is 20 kDa and works as a monomer, consisting of an incomplete β-barrel with an α-helical lid, forming an internal hydrophobic cavity, which opens and closes to allow entrance and exit of the lipoprotein [41]. The LolA-lipoprotein complex is water soluble, presumably as the N-terminal lipid is shielded from the aqueous

environment by LolA [39]. The complex crosses the periplasm and passes the lipoprotein to LolB, a 23 kDa novel outer membrane lipoprotein. LolB transfers the lipoprotein to the inner leaflet of the outer membrane where it becomes firmly attached [39]. LolB has a similar structure to LolA and it contains the same hydrophobic pocket. It has a higher affinity for lipoproteins than LolA and it is believed that an arginine residue at +43 in LolA is important for the transfer of lipoproteins from LolA to LolB [41]. Direct mouth-to-mouth transfer of lipoproteins from LolA to LolB has been speculated, given that the opening of the hydrophobic pocket of LolA contains many negatively charged residues, whilst the opening of LolB has many positively charged residues, suggesting a direct interaction [42]. Currently it is not known how LolA accepts the lipoproteins from the LolCDE complex. The roles of LolC and LolE are unclear. Despite sharing ~26% identity they are both essential [39], and have ~19% identity to LolB, so it is possible that the same mouth-to-mouth transfer occurs between LolC/LolE-LolA as with LolA-LolB [42]. What is clear however, is that the transport across the periplasm is rapid and efficient. There is an estimated 10^6 molecules of the murein lipoprotein in the outer membrane in an *E. coli* bacterium, yet there are only a few hundred LolA and LolB enzymes [42]. Thus far no LolA-lipoprotein intermediates have been detected in periplasmic fractions of *E. coli* [41]. Triacylation by Lnt is a prerequisite for a lipoprotein being accepted by the Lol system, but the inner membrane retention signal has been shown to have no effect on the activity of Lnt [43].

E. coli has over 90 predicted lipoproteins with the vast majority expected to be found on the periplasmic face of the outer membrane [13]. It is vitally important that these outer membrane proteins are correctly localised as an accumulation

in the inner membrane leads to instability and, ultimately, cell death [44, 45]. In *E. coli* the location that lipoproteins ultimately end up in is determined by the amino acid residue found at position +2. Generally, a lipoprotein with an aspartate (Asp) residue at +2 will be retained in the cytoplasmic membrane, whilst the presence of another amino acid will lead to the lipoprotein becoming a substrate for the Lol machinery and transported to the outer membrane [13, 45]. However this +2 rule is not absolute, and it appears that the amino acid at position +3 has a bearing on the destination of the lipoprotein. An acidic or amide amino acid seems to enhance the effectiveness of Asp_{+2} whilst a histidine or lysine at +3 leads to the lipoprotein being localised to the outer membrane, regardless of the presence of Asp_{+2} [39]. Also, a number of synthetic inner membrane retention signals have been identified at +2, including phenylalanine, tryptophan, tyrosine, glycine and proline, although these do not occur naturally in *E. coli* lipoproteins [39]. The current hypothesis as to how the +2 rule works is that the Lol complex does not directly recognise the presence or absence of Asp at +2, but that the negative charge of the R group (carboxylic acid) is the correct distance away from its backbone carbon (Cα) to lead to a conformational change of the lipoprotein and Lol avoidance [13]. Also, Lol avoidance appears to involve an interaction between the Asp_{+2} and phosphatidylethanolamine found in the cytoplasmic membrane [46].

Other species of Gram-negative bacteria have been shown to possess a different set of retention signals, *Psuedomonas aeruginosa* uses the amino acids at +3 and +4, as well as the +2 Asp to determine whether a lipoprotein is retained or not [44]. *Borrelia burgdorferi,* the causative agent of Lyme disease, has no Lol avoidance system with all lipoproteins being targeted to the outer

membrane by default [47]. Interestingly, despite the lack of an outer membrane and the fact that all lipoproteins remain in the 'inner' membrane, there has yet to be a Gram-positive lipoprotein found with an Asp at residue +2 [45].

1.3.5 The essentiality of the lipoprotein pathway

The lipoprotein pathway is essential for viability in all the Gram-negative bacteria tested to date [48]. Mutants lacking the lipoprotein biosynthetic enzymes are likely to aggregate improperly processed lipoproteins in the cytoplasmic membrane or periplasm which ultimately leads to cell death [14]. A number of essential lipoproteins have been identified in *E. coli* such as LolB and Murein, as mentioned previously. The same is true of the Lol system [44]. Conversely it appears that Gram-positive bacteria are able to survive without this pathway, and in contrast to the rigid Gram-negative pathway, there is some flexibility as to the order of lipoprotein processing. As described above, the first enzyme in the pathway, Lgt, is a prerequisite for the action of the second enzyme Lsp in Gram-negative bacteria. A number of viable *Δlgt* Gram-positive mutants have been described, including *B. subtilis* [49], *L. monocytogenes* [24] and *S. agalactiae* [50], with the latter two examples having detectable, unlipidated, immature lipoproteins in their supernatant, showing that the action of Lsp is independent of Lgt. Examples of *Δlsp* Gram-positive bacteria can also be found, such as in *B. subtilis* [20], *L. monocytogenes* [51] and *M. tuberculosis* [18]. Confusingly, despite the non-essentiality of this pathway in Gram-positives, a number of essential lipoproteins do exist. In *Lactococcus lactis* two lipoproteins, PrtM and OppA, are essential for growth in milk but a *Δlsp* mutant is still viable and able to grow [52]. The proteins are lipidated, and remain active in their pro form, in the absence of Lsp. The same can be said of the *B.*

subtilis lipoprotein PrsA, involved in extracytoplasmic folding of secreted proteins [49]. In an *Δlgt* mutant the protein is unlipidated and remains attached to the cell membrane by its signal peptide unprocessed by Lsp, whereas in a *Δlsp* mutant the protein is found in both in its immature form, with an intact signal peptide and in its 'mature' form, suggesting that some alternative processing may be occurring in the absence of Lsp [49]. Examples of alternative processing of lipoproteins have been discovered in other Gram-positive bacteria, such as the Eep protein from *Streptococcus uberis* [53]. The reason for these alternative processing pathways is unclear, potentially they prevent the buildup of unprocessed lipoproteins in the cell membrane.

The final enzyme in the Gram-negative pathway Lnt is, as discussed, absent from the majority of Gram-positive bacteria. The reasons for its presence in the actinomycetes is unclear, especially given that it is not essential in *M. smegmatis* [38]. Recently however, it has been postulated that the outer most cell wall lipid in various mycobacteria forms a symmetrical bilayer that could be thought of as an outer membrane [54]. This, coupled with the fact that Cryo-Electron Microscopy has revealed a periplasm between the cell membrane and the cell wall in *S. aureus* [55] and in *B. subtilis* [56], leads to the tempting thought that the Lnt proteins found in Gram-positives thus far may be involved in the transport of lipoproteins to the outer cell wall. If this were true however, then the transport system is unknown, given the absence of the Lol system.

1.3.6 Functions of lipoproteins

Bacteria use lipoproteins to interact with their extracellular environment and they perform a variety of functions. The first identified lipoprotein, the murein

lipoprotein, was found to be anchored to the outer membrane of *E. coli* and is vital for its integrity. The C-terminal end is bound to the periplasmic peptidoglycan layer and stabilises the outer membrane [10]. A substantial number of Gram-positive lipoproteins are solute-binding proteins, a well studied example of which is the high affinity maltodextrin binding protein MalX of *Streptococcus pneumoniae* [12]. Solute-binding proteins in Gram-negative bacteria are located in the periplasm, and their function is to bind reversibly with their specific substrate, and pass it to their partner ABC transporters for uptake into the cytoplasm [57]. There have been many other identified lipoprotein functions including antibiotic resistance, sporulation and bacterial conjugation [58], but the most well studied is the pathogen-host interaction. Given that lipoproteins are non-essential in Gram-positive bacteria, the *lgt* or *lsp* genes have been disrupted in a number of pathogens with differing results. The loss of Lgt during *in vitro* experiments in *L. monocytogenes*, *S. aureus* and *S. agalactiae* lead to impaired growth in cell cultures, whilst *in vivo* tests in mice showed the two latter species to be hypervirulent [24, 50, 59]. In mice, a *S. equi Δlgt* mutant was attenuated, whilst in its natural host, ponies, the bacteria showed no change in virulence [60]. The loss of Lsp leads to both *in vitro* and *in vivo* attenuation of *L. monocytogenes* [51]. An *M. tuberculosis Δlsp* mutant, shows no growth impairment *in vitro* but is attenuated *in vivo* [61]. Other *Δlsp* mutants also show different phenotypes.

Aside from the species to species variation, the most striking thing about the results above is the hypervirulence of some of the *Δlgt* mutants. At first glance this seems counterintuitive. If, as described above, the loss of Lgt can lead to the loss of lipoproteins into a culture supernatant, one might assume that their

loss would lead to the reduced fitness of a bacterium and therefore a reduction in its ability to cause infection. However what seems to be happening is that the hypervirulent bacteria are evading the host's innate immune response. Bacteria can be recognised in their hosts by their lipoproteins. In vertebrates, lipoproteins are recognised by Toll-like receptor 2 (TLR2), which can recognise whether a lipoprotein is di- or triacylated and form heterodimers with TLR6 or TLR1 respectively, as can be seen in figure 1.6 [62]. For a TLR2-TLR1 heterodimer to be formed, an N-acyl fatty acid of at least eight carbons in length is needed to interact with a hydrophobic channel found in TLR1. This channel is not present in TLR6 and so it can only form the TLR2-TLR6 heterodimer in the presence of a diacylated lipoprotein. The immune response induced by both heterodimers appear to be the same and it is unknown if there are subtle differences in the response to each [62]. Hypervirulence in bacteria is presumably due to this pathway not being activated due to the lack of lipid molecules anchoring the protein into the membrane.

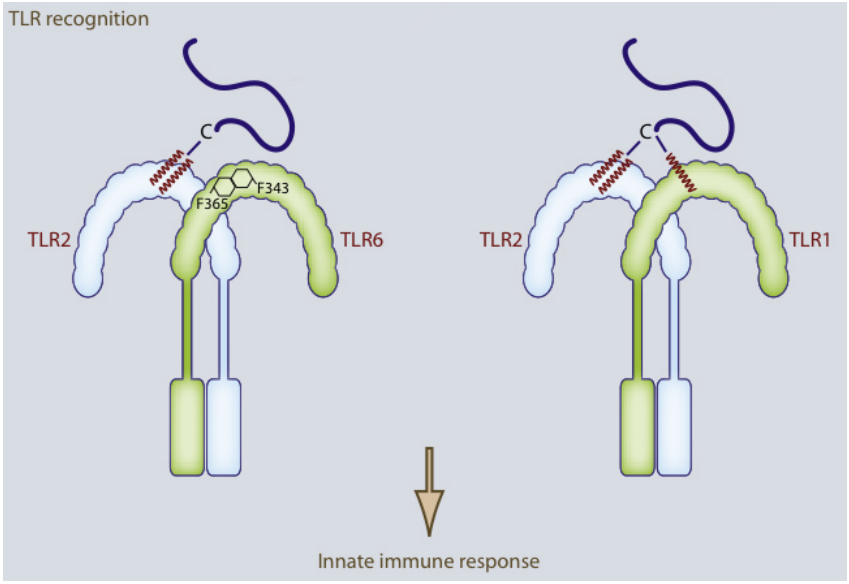


Figure 1.6. A diacylated or triacylated lipoproteins is recognised by either the TLR2-TLR6 or TLR2-TLR1 heterodimers respectively. TLR6 lacks the hydrophobic channel necessary to accept the N-acylated lipoprotein. Modified from [62].

1.3.7 Computational analysis of Lipoproteins

The rapid increase in the number of fully sequenced bacterial genomes has led to a huge amount of data available for mining. Currently more than 2000 lipoproteins have been identified *in silico* in a variety of species [63]. Lipoproteins are thought to represent between 0.5 and 8% of a bacterial genome [64], although the average is around 2.5%. Computer analysis often leads to the identification of false positives as the signal sequence can be highly variable. Pattern searches have been used to identify potential lipoproteins and most are based around the signal sequence, which contains the characteristic lipobox motif, L_{-3}-[A/S/T]$_{-2}$-[G/A]$_{-1}$-C_{+1} [11].

Studies on the Gram-positive signal sequence have led to the following observations. For a protein to be identified as a lipoprotein it must contain: i) no charged amino acids in the *h*-box ii) a cysteine residue at +1 but only between

position 15 and 50 in the leader peptide iii) only one Arg or Lys in the first seven amino acids [58]. Gram-negatives should have i) one or more charged amino acids in the first five to seven residues ii) the *h*-region should be 7 to 22 residues in length iii) the lipobox motif (shown above) should occur in the first 50 amino acids from the N-terminus [63].

1.4 *Streptomyces coelicolor*

Streptomyces coelicolor is the model organism for the high GC branch of Gram-positive bacteria known as the actinomycetes. It is a soil dwelling bacterium and has developed a complex life-cycle to enable it to survive in this nutrient scarce ecological niche.

One of the most striking things about *S. coelicolor* is its complex multistage life-cycle, which is highly unusual amongst prokaryotes. This bacterium begins life as a uninucleate spore that contains a hydrophobic coating and is resistant to desiccation. These spores germinate and establish a complex network of hyphae known as the substrate mycelium. These are often multinucleate with infrequent septa [65] with growth occurring at the hyphal tips. This method of growth initially lead to *Streptomyces* being mis-categorised as a fungus [66].

Specific environmental signals, including nutrient deprivation, activate the next stage of the developmental cycle. The substrate mycelial growth halts and the erection of aerial hyphae is initiated. The growth of these reproductive structures is fuelled by controlled lysis of the substrate mycelium [67], which provides nutrients to the aerial mycelium. The substrate hyphae are emptied of their cellular contents, but retain their cell wall integrity, presumably to prevent

collapse of the colony [68]. Emergence of the aerial hyphae is dependent on the production, and action, of the secreted lantibiotic (lanthionine-containing peptide antibiotics) SapB [69]. This protein reduces the surface tension at the water-air interface allowing the aerial hyphae to escape the aqueous environment of the colony. Other proteins known as chaplins act as hydrophobic sheaths, enhancing the growth of the hyphal tip [70].

When fully developed, the aerial mycelia undergo cell division by compartmentalising their chromosomes and laying down septa. These septated hyphae eventually form spore chains with each compartment a separate spore. The spore chains grow in a characteristic spiral shape, and their hydrophobic coating aids propagation as they can be carried long distances by water [67].

S. coelicolor is also unusual as it contains one of the largest genomes of any currently sequenced bacterium, with 8,667,507 base pairs encoding 7,825 genes [66], twice as many as the related actinomycete *M. tuberculosis*. The genome is linear, which is also very unusual amongst bacteria, and consists of a central core region flanked by two 'arms'. The core region is approximately half the genome in size (4.9 Mbp) and contains all the essential genes whilst the arms contain genes that are non-essential, only expressed under specific conditions and which were probably acquired through horizontal gene transfer (HGT). The left arm and right arm are of different lengths, being 1.5 Mbp and 2.3 Mbp respectively. Indeed, the core region has more in common with other actinomycetes such as *M. tuberculosis* or *Corynebacterium diptheriae* than the arms [66].

This large genome is certainly useful given the variability of the soil environment and *S. coelicolor* secretes a large number of proteins in order to break down a variety of potential nutrient sources. A predicted 10.5% of the encoded proteins are hydrolases (proteases, cellulases etc.) with an equally significant proportion of the proteome (7.8%) taken up by various transport systems, probably for uptake of the breakdown products [71].

It is generally accepted that the amount of regulatory genes in an organism increases proportionally to the increase in genome size and *S. coelicolor* contains a high proportion of regulatory proteins (12.3%) [66]. *S. coelicolor* encodes multiple two-component regulatory systems and contains 84 sensor kinases (SK) and 80 response regulators (RR), (67 paired SK/RRs and 17 unpaired RRs) comprising 0.86% of its total Open Reading Frames (ORFs). This value is 25% higher than the average non-pathogenic bacteria [72]. The high number of two-component systems is again almost certainly due to the variety of stresses the bacterium faces in the soil.

Another important set of genes in *S. coelicolor* are those which encode secondary metabolites. There are predicted to be 220 genes, in 22 clusters involved in the biosynthesis of secondary metabolites [71]. Commercially these metabolites are very important as approximately 70% of commercial antibiotics are derived from *Streptomyces* species as well as other pharmaceutically important compounds such as anti-tumour drugs [73]. These metabolites are most likely produced as a defence response to competitors, but may also serve other functions such as combating physical, chemical or biological stresses [66].

1.5 *Streptomyces scabies*

Like *S. coelicolor*, *S. scabies* is a soil dwelling saprophyte, and shares many elements of the multistage lifecycle seen in *S. coelicolor*. However, *S. scabies* is unusual in that it is a pathogenic *Streptomyces* species. *S. scabies* is predominantly a pathogen of potatoes (*Solanum tuberosum*), and has been identified worldwide [74]. The bacterium causes characteristic lesions on the surface of potato tubers, which reduces their market value. This disease is of great economic importance worldwide. As well as attacking potatoes and other tap root crops, *S. scabies* is also able to infect other monocotyledonous or dicotyledonous plants [67] and appears neither host, nor tissue, specific.

The ability to cause disease is due to the presence of a discrete cluster of genes, known as a pathogenicity island (PAI). The genes found in these islands can be moved between bacteria either individually or as a group, and as such the G+C content of the PAI can differ from the genome it is found in [75]. *S. scabies* contains a PAI which at 325 kb is currently the largest known bacterial PAI, with a G+C content substantially lower than the remaining genome [76]. The genes on this island encode the two main virulence factors for *S. scabies*, thaxtomin A and Nec1. Thaxtomin A is a nitrated dipeptide phytotoxin, capable of necrosing excised potato tissue and causing scabs on immature potatoes [67], whilst Nec1 is a novel virulence factor, and is a necrotic protein of unknown target [75]. Both factors are required for full plant virulence.

The mechanisms that *S. scabies* uses to invade plants is unknown, and being soil based offers some unique challenges that other plant pathogens do not face. Root structures lack the natural openings seen in other plant organs,

such as stomata found in leaves [77], and it appears that *S. scabies* employs specialised mechanisms to penetrate the potato tuber. Specialised hyphae have been visualised growing directly into a tuber, although it is not known whether enzymatic degradation of the plant cell wall occurs to allow the bacterial hyphae to enter the potato [77, 78].

1.6 Aims of this thesis

This work will investigate the lipoprotein biosynthetic pathways in both *S. coelicolor* and *S. scabies*. These pathways are unusual amongst bacteria, both contain one copy of *lsp*, whilst *S. coelicolor* contains two copies of *lgt*, and both contain two copies of *lnt.* There is also good evidence that *Streptomyces* species send large numbers of lipoproteins out via the Tat pathway which is highly unusual amongst bacteria.

Of the two *lgt* genes found in *S. coelicolor*, *lgt1* (*SCO2034*) is found in the central region of the chromosome, whilst *lgt2* (*SCO7822*) is found on one of the arms and has possibly been acquired by HGT. It is possible, therefore, that the gene may have an alternative function or be functionally redundant. It will be interesting to see which Lgt is able to add the diacylglyceride to an immature lipoprotein and, if both are able, whether there are specific proteins lipidated by each one. To test the functions of these enzymes it should be possible to examine whether Δ*lgt1 lgt2* mutants release unlipidated proteins into their growth media, as is the case with *L. monocytogenes* [24].

Why both *S. coelicolor* and *S. scabies* would have two homologues of *lnt* is a mystery. As has been previously stated, the Lnt protein was only thought to be

active in Gram-negative bacteria, until the recent discovery of an active Lnt in *M. smegmatis* [38]. As with the *M. smegmatis* Lnt, both of the *Streptomyces* proteins share five of the seven residues needed for activity in *E. coli.* It is currently unknown whether these proteins can act as true Lnts as the *M. smegmatis* Lnt does.

As discussed, *S. coelicolor*, *S. scabies* and other members of the actinomycetes are unusual in their lipoprotein biosynthesis and they may represent a new paradigm in this field of research. The *S. coelicolor* pathway will be analysed by a combination of experimental and *in silico* analysis. The *lgt* and *lsp* genes will be deleted from the chromosome of *S. coelicolor* and the mutant strains will be subjected to phenotypic analysis. Bioinformatic identification of lipoproteins and their roles in *S. coelicolor* will give clues about processes affected in strains which cannot synthesise lipoproteins. The macroscopic and microscopic phenotypes of the mutants will also be analysed to determine whether this pathway has any effect on development or antibiotic production. The same will be done for *S. scabies*, with the addition that the action of the lipoprotein pathway on the infection of plants will be assessed by the disruption of each of the lipoprotein processing genes. The action of these mutants on both living plants, and potato tubers will be tested. Whilst there have been multiple studies on disrupting lipoprotein biosynthesis in animal pathogens and the effect therein on virulence, to date there has been no such study on plant pathogens.

It seems likely that lipoproteins and the lipoprotein biosynthetic pathway will be of particular importance to *Streptomyces* species given their saprophytic

lifestyle. Nutrient scavenging in the soil requires a variety of exported and externally facing proteins, including plant cell wall degrading enzymes and substrate transport systems. It is likely that a number of these are lipoproteins, especially given the ability of this subset of proteins to remain anchored in the cell membrane in large numbers without mechanical disruption to the cell.

Chapter 2. Materials and Methods.

2.1 Strains and culture conditions.

The bacterial strains, and plasmids used or constructed in this study are listed in table 2.2. Growth media used are listed in table 2.3. Liquid cultures of *E. coli* were routinely grown shaking at 250 rpm, in LB broth at 37°C unless stated. Liquid cultures of *S. coelicolor* or *S. scabies* were grown at 30°C, shaking at 250 rpm. Typically 10 ml of liquid culture was grown. Cultures grown on solid media were grown at the same temperatures listed above, unless stated. Where necessary, cultures were supplemented with antibiotics at concentrations listed in table 2.1.

Antibiotic	Stock Concentration (mg/ml)	Working Concentration (µg/ml) for media	Working Concentration (mg/ml) for overlays
Ampicillin	100	100	
Apramycin	50	50	1.25
Chloramphenicol	25	25	
Hygromycin	25	12.5	0.625
Kanamycin	50	50	
Nalidixic acid	25	25	0.5
Vancomycin	25	25	

Table 2.1 Concentrations of antibiotics used during this thesis.

Strain	Description	Cosmid/Plasmid used to manipulate parent strain	Resistance	Reference
E. coli				
DH5α	General Cloning Strain			Promega
One Shot Top 10	Cloning host for TOPO cloning kit			Invitrogen
BW25113	BW25113 containing λ RED recombination plasmid pIJ790		$Chlor^R$	[84]
ET12567	Et12567 containing helper plasmid pUZ8002		$Chlor^R$ /Tet^R	[84]
BT340	DH5α containing temperature sensitive FLP recombination plasmid BT340		$Carb^R$/$Chlor^R$	[84]
S.coelicolor **A3 (2)**				
M145	$SCP1^-$, $SCP2^-$ *S. coelicolor* wild-type strain			[79]
J2172	M600 *ΔcseA* $SCP1^-$, $SCP2^-$			[93]
BJT1000	M145 *Δlsp::apr*	4A10 *lsp::apr*	$Apra^R$	This work
BJT1001	M145 *Δlsp*Flp			This Work
BJT1002	M145 *Δlgt1*::apr	4G6 *lgt1::apr*	$Apra^R$	This Work
BJT1004	M145 *vanJ*-Sco *lsp*	pBT106	Hyg^R	This Work
BJT1005	M145 *Δlsp*FLP + *vanJ* promoter Sco *lsp* fusion	pTDW188	Hyg^R	This Work
BJT1006	M145 *Δlsp*::*apr* + Sco *lsp* cis	4A10 *lsp*::*apr* + 4A10 *bla->hyg*		This Work
BJT1007	M145 *Δlsp*FLP + Sco *lsp in trans*	pBT100	$Apra^R$	This Work
BJT1008	M145 *Δlsp*FLP + N10 Sco *lsp*	pBT102	$Apra^R$	This Work
BJT1009	M145 *Δlsp*FLP + N20 Sco *lsp*	pBT103	$Apra^R$	This Work
BJT1010	M145 *Δlsp*FLP + N30 Sco *lsp*	pBT104	$Apra^R$	This Work
BJT1011	M145 *Δlsp*FLP + N40 Sco *lsp*	pBT105	$Apra^R$	This Work
BJT1012	M145 *Δlsp*FLP + D148A/D177A Sco *lsp*	pBT101	$Apra^R$	This Work
BJT1013	M145 + *eGFP* Sco *lgt1* fusion	pBT111	$Apra^R$	This Work
BJT1014	M145 + mCherry Sco *lgt1* fusion	pBT115	$Apra^R$	This Work
BJT1015	M145 + *eGFP* Sco *lgt2* fusion	pBT112	$Apra^R$	This Work
BJT1016	M145 + *mCherry* Sco *lgt2* fusion	pBT116	$Apra^R$	This Work
BJT1017	M145 + *eGFP* Sco *lsp* fusion	pBT113	$Apra^R$	This Work

BJT1018	M145 + *mCherry* Sco *lsp* fusion	pBT117	Apra^R	This Work
BJT1019	M145 + *mCherry* N40 Sco *lsp* fusion	pBT118	Apra^R	This Work
BJT1021	M145 *Δlgt1* + Sco *lgt1 in trans*	4G6 *lgt1::apr* + pBT119	Hyg^R	This Work
BJT1022	M145 *Δlgt1* + Sco *lgt1* c-trunc	4G6 *lgt1::apr* + pBT120	Hyg^R	This Work
BJT1023	M145 *Δlgt1* + *eGFP* Sco *lgt1* fusion	4G6 *lgt1::apr* + pBT111	Apra^R	This Work
BJT1024	M145 *Δlgt1* + *mCherry* Sco *lgt1* fusion	4G6 *lgt1::apr* + pBT115	Apra^R	This Work
BJT1025	M145 *Δlgt1* + *mCherry* Sco *lsp* fusion	4G6 *lgt1::apr* + pBT117	Apra^R	This Work
BJT1026	M145 *Δlgt1* + Sco3484-His fusion	4G6 *lgt1::apr* + pTDW188	Hyg^R	This Work
BJT1027	M145 *Δlgt2::apr*	8E7 *lgt2::apr*	Apra^R	This Work
BJT1028	M145 *Δlgt2* + Sco *lgt2 in trans*	8E7 *lgt2::apr* + pBT121	Hyg^R	This Work
BJT1029	M145 *Δlgt2* + Sco *lgt2* c-trunc	8E7 *lgt2::apr* + pBT122	Hyg^R	This Work
BJT1030	M145 *Δlgt2* + *eGFP* Sco *lgt2* fusion	8E7 *lgt2::apr* + pBT112	Apra^R	This Work
BJT1031	M145 *Δlgt2* + *mCherry* Sco *lgt2* fusion	8E7 *lgt2::apr* + pBT116	Apra^R	This Work
BJT1032	M145 *Δlgt2* + *mCherry* Sco *lsp* fusion	8E7 *lgt2::apr* + pBT117	Apra^R	This Work
BJT1033	M145 *Δlgt2* + Sco3484-His fusion	8E7 *lgt2::apr* + pTDW188	Hyg^R	This Work
BJT1034	M145 *Δlsp*Flp + *eGFP* Sco *lsp* fusion	pBT113	Apra^R	This Work
BJT1035	M145 *Δlsp*Flp + *mCherry* Sco *lsp* fusion	pBT117	Apra^R	This Work
BJT1036	M145 *Δlsp*Flp + *eGFP* N40 Sco *lsp* fusion	pBT114	Apra^R	This Work
BJT1037	M145 *Δlsp*Flp + *mCherry* N40 Sco *lsp* fusion	pBT118	Apra^R	This Work
BJT1039	M145 *Δlsp*Flp+ Sco *lsp in trans* + Sco3484-His fu	pBT100 + pTDW188	Hyg^R	This Work
DW1006	M145 + Sco3484-His fusion	pTDW188	Hyg^R	[82]
DW1007	M145 *Δlsp*Flp + Sco3484-His fusion	pTDW188	Hyg^R	[82]
	M145 *ΔtatB*Flp + *eGFP* fused *tatB*		Apra^R	This Work
S. scabies				
87-22	Wild-type			[127]
BJT1040	Scabies *Δlgt::apr*	Scab 139 *lgt::apr*	Apra^R	This Work
BJT1041	Scabies *Δlgt::apr* + Scab *lgt in trans*	Scab 139 *lgt::apr* + pBT124	Apra^R Hyg^R	This Work
BJT1042	Scabies *Δlgt::apr* + Sco *lgt1 in trans*	Scab 139 *lgt::apr* + pBT119	Apra^R Hyg^R	This Work
BJT1043	Scabies *Δlgt::apr* + Sco *lgt2 in trans*	Scab 139 *lgt::apr* + pBT121	Apra^R Hyg^R	This Work
BJT1044	Scabies *Δlsp::apr*	Scab 45 *lsp::apr*	Apra^R	This Work

BJT1045	Scabies *Δlsp::apr* + Scab *lsp in trans*	Scab 45 *lsp::apr* + pBT123	Apra^R Hyg^R	This Work
BJT1046	Scabies *Δlsp::apr* + Sco *lsp* in trans	Scab 45 *lsp::apr* + pBT110	Apra^R Hyg^R	This Work
BJT1047	Scabies *Δlnt1::apr*	Scab 351 *lnt1::apr*	Apra^R	This Work
BJT1048	Scabies *Δlnt2::apr*	Scab 2255 *lnt2::apr*	Apra^R	This Work
BJT1049	Scabies *Δlnt2::hyg*	Scab 2255 *lnt2::hyg*	Hyg^R	This Work
BJT1050	Scabies *Δlnt1::apr Δlnt2::hyg*	Scab 351 *lnt1::apr* + Scab 2255 *lnt2::hyg*	Apra^R Hyg^R	This Work
Plasmids				
pSET152	Integrative Streptomyces vector		Apra^R	[79]
pMS82	Integrative Streptomyces vector		Hyg^R	[129]
pSETSORhyg	pSET152 with *apra* replaced by *hyg*		Hyg^R	[80]
pIJ8660	eGFP containing plasmid		Apra^R	[81]
pFK210	pBluescript containing mCherry red fluorescent gene		Apra^R	K. Flärdh Unpublished
pMH034	pMS82 containing vanJ promoter		Hyg^R	M. Hutchings Unpublished
pBAD24-NdeI	pBAD24 with the NcoI site changed to NdeI		Amp^R	M. Hutchings Unpublished
pTDW134	based on pIJ8600, contains eGFP fused *tatB*		Apra^R	D. Widdick Unpublished
pBT100	pSET152 + full length Sco *lsp*		Apra^R	This work
pBT101	pSET152 + D148A, D177A Sco *lsp* point mutant		Apra^R	This work
pBT102	pSET152 + N10 Sco *lsp* truncation		Apra^R	This work
pBT103	pSET152 + N20 Sco *lsp* truncation		Apra^R	This work
pBT104	pSET152 + N30 Sco *lsp* truncation		Apra^R	This work
pBT105	pSET152 + N40 Sco *lsp* truncation		Apra^R	This work
pBT106	pMS82 vanJp-Sco *lsp* fusion		Hyg^R	This work
pBT107	pBAD24-NdeI + E. coli *lsp*		Amp^R	This work
pBT108	pBAD24-NdeI + Sco *lsp*		Amp^R	This work
pBT109	pBAD24-NdeI +Sco N40 *lsp*		Amp^R	This work
pBT110	pSETSORHYG + Sco *lsp*		Hyg^R	This work
pBT111	pIJ8660 eGFP-*lgt1* fusion		Apra^R	This work

pBT112	pIJ8660 eGFP-*lgt2* fusion	Apra^R	This work
pBT113	pIJ8660 eGFP-*lsp* fusion	Apra^R	This work
pBT114	pIJ8660 eGFP-N40 Sco *lsp* truncation fusion	Apra^R	This work
pBT115	pFK210 mCherry-Sco *lgt1* fusion	Apra^R	This work
pBT116	pFK210 mCherry-Sco *lgt2* fusion	Apra^R	This work
pBT117	pFK210 mCherry-Sco *lsp* fusion	Apra^R	This work
pBT118	pFK210 mCherry-N40 Sco *lsp* truncation fusion	Apra^R	This work
pBT119	pMS82 + full length Sco *lgt1*	Hyg^R	This work
pBT120	pMS82 + c-truncated Sco *lgt1*	Hyg^R	This work
pBT121	pMS82 + full length Sco *lgt2*	Hyg^R	This work
pBT122	pMS82 + c-truncated Sco *lgt2*	Hyg^R	This work
pTDW188	pSETSORhyg + penta Histidine fused Sco3483	Hyg^R	[82]
pBT123	pMS82 + full length *S. scabies lsp*	Hyg^R	This work
pBT124	pSETSORHYG + full length *S. scabies lgt*	Hyg^R	This work
pBT125	pMS82 + full length *S. scabies lnt1*	Hyg^R	This work
Cosmids			
4G6	cosmid + full length *lgt1*	Kan^R, Amp^R	[83]
4G6 lgt1::apr	4G6 cosmid + *apra* marked *lgt1* deletion	Apra^R, Kan^R, Amp	This work
8E7	cosmid + full length *lgt2*	Kan^R, Amp^R	[83]
8E7 *lgt2*::apr	8E7 cosmid + *apra* marked *lgt2* deletion	Apra^R, Kan^R, Amp	This work
4A10	cosmid + full length *lsp*	Kan^R, Amp^R	[83]
4A10 *lsp*::apr	4A10 cosmid + *apra* marked *lsp* deletion	Apra^R, Kan^R, Amp	This work
4A10 *lsp*Flp	4A10 cosmid + in frame *lsp* deletion	Kan^R, Amp^R	This work
4A10 bla->hyg	4A10 cosmid with *bla->hyg* replacement	Kan^R, Hyg^R	This work
Scab 139	Cosmid + full length Scabies *lgt*	Kan^R, Amp^R	Cornell University
Scab 139 lgt::ap	139 cosmid + *apra* marked Scabies *lgt*	Apra^R, Kan^R, Amp	This work
Scab 45	Cosmid + full length Scabies *lsp*	Kan^R, Amp^R	Cornell University
Scab 45 lsp::apr	45 cosmid + *apra* marked Scabies *lsp*	Apra^R, Kan^R, Amp	This work
Scab 351	Cosmid + full length Scabies *lnt1*	Kan^R, Amp^R	Cornell University
Scab 351 lnt1::a	351 cosmid + *apra* marked Scabies *lnt1*	Apra^R, Kan^R, Amp	This work

Scab 2255	Cosmid + full length Scabies *Int2*	Kan^R, Amp^R	Cornell University
Scab 2255 Int2::*apra*	2255 cosmid + *apra* marked Scabies *Int2*	$Apra^R$, Kan^R, Amp	This work

Table 2.2 Strains, plasmids and cosmids used during this thesis.

Media	Composition	Weight, %v/v, %w/v or mM per 1 litre
Lennox Broth (LB)	Tryptone	10 g
	Yeast Extract	5 g
	NaCl	5 g
	dH_2O	1000 ml
LB Agar	Tryptone	10 g
	Yeast Extract	5 g
	NaCl	5 g
	Agar	15 g
	dH_2O	1000 ml
Soy Flour Media (SFM)	Mannitol	20 g
	Soy Flour	20 g
	Agar	20 g
	Tap Water	1000 ml
Tryptone Soya Broth (TSB)	TSB powder (Oxoid)	30 g
	dH_2O	1000 ml
Yeast extract Malt extract medium (YEME)	Yeast Extract	3 g
	Peptone	5 g
	Malt Extract	3 g
	Glucose	10 g
	dH_2O	1000 ml
Difco Nutrient Broth (DNB) Agar	DNB powder (BD)	4 g
	Agar	20 g
Minimal Media (pH 7.0)	*L*-asparagine	0.5 g
	di-potassium hydrogen orthophosphate	0.5 g
	Magnesium Sulpate heptahydrate	0.2 g
	Iron (II) sulphate heptahydrate	0.01 g
	Agar	10 g
	dH_2O	1000 ml
	Before dispensing, 4 ml 50% glucose added per 200ml	
Instant Mash Agar (IMA)	SMASH potato powder	20 g
	Agar	20 g
	Tap Water	1000 ml
Murashige & Skoog (MS) Medium	MS powder (Melford)	2.2 g
	Sucrose	2%
	MES Sodium Salt pH 5.9 (Sigma Aldrich)	0.5 g
	dH_2O	1000 ml

Table 2.3 Growth media used during this thesis.

2.1.1 Preparation of *Streptomyces* spores.

Spores from *S. coelicolor* or *S. scabies* were plated onto Soya Flour Media (SFM), or Instant Mash Agar (IMA) respectively, with the spores streaked out to grow a confluent lawn on the plate. Plates were incubated at 30°C for 5 nights. Spores were harvested by placing 1 ml of sterile 20% glycerol (2G) on the plate, and sloughing off the spores with a sterile cotton bud. A further 1 ml of 2G was then added and the spore suspension was removed by pipetting and stored in a 2 ml centrifuge tube at -20° C.

2.1.2 Glycerol stocks.

Glycerol stocks of *E. coli* strains were made by centrifuging 1ml of culture in a desktop centrifuge for 30 seconds. Once centrifuged the supernatant was poured off, 0.5ml of fresh LB broth and 0.5 ml 40% glycerol were added. The pellet was resuspended and frozen.

2.2 Genetic Manipulations.

2.2.1 Plasmid preparation.

Qiaprep Spin Miniprep kits (Qiagen) were used to prepare plasmid DNA from 10 ml overnight cultures, as per the manufacturers instructions. Plasmids were eluted in autoclaved distilled water and stored at -20°C. A number of plasmids were synthesised by Genscript USA Inc. (section 2.4.1).

2.2.2 Cosmid preparation.

In order to recover cosmid DNA from *E. coli*, 1.5 ml of overnight culture containing the cosmid was transferred to a 1.6 ml microfuge tube and cells were recovered by centrifugation at 13,000 rpm for 1 minute. The cell pellet was

immediately resuspended in 100 µl of Solution 1 (table 2.5). Once resuspended, 200 µl Solution 2 (table 2.5) was added and the tube was mixed 10 times by inversion. After mixing, 150 µl ice cold Solution 3 (table 2.5) was added and mixing was achieved by inverting the tubes five times. The solution was centrifuged for five minutes at 13,000 rpm and the supernatant was transferred to a fresh tube. Immediately after centrifugation 400 µl 1:1 phenol/ chloroform was added to the supernatant and the solution was mixed for 2 minutes on a vortex mixer. The tube was then centrifuged as above. The upper phase was transferred to a fresh tube and 600 µl ice-cold 2-propanol was added. The tube was left on ice for 10 minutes. Once cooled the tube was centrifuged as above to spin down the precipitated cosmid DNA. All liquid was removed from the tube and the pellet, containing cosmid DNA and total RNA, was washed with 200 µl 70% ethanol. The tube was centrifuged as above, all ethanol was removed and the tube was left open on its side for five minutes to air dry the pellet and remove any residual ethanol. The pellet was resuspended in 50 µl TE buffer plus 2 µl DNase free RNase (Sigma-Aldrich) and the tube was incubated at 37°C for 15 minutes. The cosmid DNA was then stored at -20°C.

2.2.3 Polymerase Chain Reaction.

GoTaq polymerase (Promega) was used for colony PCR, with TaqExpand Long Template (with buffer 3) (Roche) used for high fidelity PCR and cloning. A typical reaction contained buffer supplied by the manufacturer, 50 µM dNTP mix (12.5 µM each dNTP), 5% DMSO, 20 pmoles of each primer (Invitrogen), 2.5 units polymerase in a 50 µl reaction. If the template was a plasmid, 1 µl of a stock at approximately 50 ng/µl was used. When genomic DNA was used, a

bacterial colony was picked with a sterile wooden toothpick and scraped into a PCR tube. To amplify DNA a BioRad DNA Engine thermocycler was used. A typical PCR program included an initial denaturing step of 94°C/5 minutes followed by 25 cycles with a denaturing step of 94°C/30 seconds, a primer annealing step of 63°C/90 seconds and an extension step of 72°C/2 minutes, and a final single extension of 72°C/5 minutes. The annealing temperatures and extension times were tailored to suit each primer set.

2.2.4 DNA sequencing.

All cloned PCR products were sequenced using the ABI Big Dye 3.1 die-terminator reaction mix, according to the manufacturers instructions. When sequencing from a plasmid, a 10 µl reaction containing; 1 µl of DNA, 1 µl of primer (3.2 pm/µl), 1 µl reaction mix, 1.5 µl reaction buffer, 5.5 µl dH_2O was used. The thermo-cycler conditions included an initial denaturing step of 96°C/1 minute followed by 25 cycles with a denaturing step of 96°C/10 seconds, a primer annealing step of 50°C/5 seconds and an extension step of 60°C/4 minutes, and a final single extension of 60°C/10 seconds. The sequence analysis was carried out at The Genome Analysis Centre (www.jicgenomelab.co.uk) using 3730XL sequencers (Life Technologies). Sequence trace files were analysed using 4Peaks software (mekentosj.com/science/4peaks/).

Name	Sequence
Sco Lgt1 KO For	gcgccccccgtccgacacggtagcgtcgaccctgccatgattccggggatccgtcgacc
Sco Lgt1 KO Rev	taccgggcgccctcggcggtgtgtcttgcgggcgggtcatgtaggctggagctgcttc
Sco Lgt1 Test For	gcgccccccgtccgacacgg
Sco Lgt1 Test Rev	taccgggcgccctcggcggt
Sco Lgt1 complementation For	gctgaccgacgccgggcagctcaag
Sco Lgt1 complementation Rev	agacgatcgtcaaggcggccacggt
Sco Lgt1 c-trunc Rev	tcacggctccaccacggcctc
Sco Lgt1 gfp For	ccccggatcccatatggcccttcttggcggagtcggcgccg
Sco Lgt1 gfp Rev	ggccggatccgctgaccgacgccgggcagctcaag
Sco Lgt1 c-trunc gfp Rev	ccccggatcccatatgcggctccaccacggcctc
Sco Lgt2 KO For	gaaacccctccacgacctcgaccaaggctctcgatcatgattccggggatccgtcgacc
Sco Lgt2 KO Rev	ttcgcaccagcaccgccggtcgcctgtacgagcgcatcatgtaggctggagctgcttc
Sco Lgt2 Test For	gaaacccctccacgacctcg
Sco Lgt2 Test Rev	ttcgcaccagcaccgccggt
Sco Lgt2 complementation For	caggagtggccgtatgacaccccct
Sco Lgt2 ctrunc Rev	tcaattgaaaatgttgctcgt
Sco LgtLgt2 gfp For	ggccggatcccaggagtggccgtatgacaccccct
Sco Lgt2 gfp Rev	ccccggatcccatatgcgatctgttacggacgttccaggac
Sco Lgt2 n-trunc gfp Rev	ccccggatcccatatgattgaaaatgttgctcgt
Sco Lsp KO For	tcgtgctcagtcaaggacctaggctgagggactcacgtgattccggggatccgtcgacc
Sco Lso KO Rev	gacaaccagtccctgtggacagccggaccggaggggtcatgtaggctggagctgcttc
Sco Lsp Test For	tcgtgctcagtcaaggacct
Sco Lsp Test Rev	gacaaccagtccctgtggac

Sco Lsp complementation For	gctgttcccaaggcccgcggcaccg
Sco N40 Truncation For	ggcgcccatatgcgcacgcggggcaagcggcggg
Sco Lsp gfp For	ggccggatccgctgttcccaaggcccgcggcaccg
Sco Lsp gfp Rev	ccccggatcccatatggtccttgtggacggtcccgtccggg
Sco Lsp Comp KO For	gatgctcgaccagaccgagcgcgccctggaccggctggcattccggggatccgtcgacc
Sco Lsp Comp KO Rev	gggaatcgtgctcacccgtcgagtatgccggacgggtgatgtaggctggagctgcttc
Sco Lsp Comp KO test For	gatgctcgaccagaccgagc
Sco Lsp Comp KO test Rev	gggaatcgtgctcacccgtc
Scab lgt KO For	gtcggccgggcccgacagggtagcgtcgaccctgccatgattccggggatccgtcgacc
Scab lgt KO Rev	tgggcaggcatcacatcgagcgaaccggctgtcgcgtcatgtaggctggagctgcttc
Scab lgt test For	ggcggcagcggctcccgcaaggccg
Scab lgt test Rev	tgggcaggcatcacatcgagcgaac
Scab lgt comp For	acatctacgcggaccagtccatgac
Scab lgt comp Rev	tgggcaggcatcacatcgagcgaac
Scab lsp KO For	ttgtcctcagtcaggaacctaggttgagggactcacgtgattccggggatccgtcgacc
Scab lsp KO Rev	cctgtggacaacctccccggcgcggcccggacgtcctcatgtaggctggagctgcttc
Scab lsp test For	acgacgggagccacgacagtggttg
Scab lsp test Rev	cctgtggacaacctccccggcgcgg
Scab lsp comp For	ctcgggcacgtacggcctctgcgag
Scab lsp comp Rev	gaggcgcccgaccgtgcctgtggac
Scab Lnt Ko For	tgtcggcgccggggtgacgcgccgtacactccccgagtgattccggggatccgtcgacc
Scab Lnt Ko Rev	ccccgccgttccggcacgcgcacgcgaggcgtacggctatgtaggctggagctgcttc
Scab lnt comp For	gcctctcgcctgtgccgggcctcga

Scab lnt comp Rev	tggagccgaccctaacgggccggtt
Scab lnt test For	gatcagcgccgagctccccacggtg
Scab lnt test Rev	ccccgccgttccggcacgcgcacgc
Scab Lnt2 KO For	actggagggacgcaccgggtggcagggaggcgctgtatgattccggggatccgtcgacc
Scab Lnt2 KO Rev	gggcagggtgctcagcgcccgggacgccctggccactcatgtaggctggagctgcttc
Scab lnt2 TEST For	actggagggacgcaccgggt
Scab lnt2 TEST Rev	gggcagggtgctcagcgccc
Scab lnt2 comp FOR	ggtaccgccaaagagcacagccccaccgaca
Scab lnt2 comp REV	aagcttccgtcctcgccttcctcgacgccgt
EcolspA KO For	caacgtcgccggtgacggtgaaaaacgtaagtttgcctgattccggggatccgtcgacc
EcolspA KO Rev	ccggataagatgcgtcagcatcgcatccggcagggtttatgtaggctggagctgcttc
EcolspA For	ggcgcccatatgagtcaatcgatctgttca
EcolspA Rev	ggcgccggatccttattgtttttcgctctaga
ScoFL LspA For	ggcgcccatatggcagaggcggagcgcatcatcg
Sco LspA Rev	ggcgccggatcctcagtccttgtggacggtcccgtcc
ScoN40 LspA For	ggcgcccatatgcgcacgcggggcaagcggcggg
EcolspA test For	agggctgaaagtcgcgttgagtaaa
EcolspA test Rev	acggcgctattgctctgtacagatt
P1	attccggggatccgtcgacc
P2	tgtaggctggagctgcttc
bla replacement For	ccctgataaatgcttcaataatattgaaaaaggaag
bla replacement Rev	aatcaatctaaagtatatatgagtaaacttggtctgacag
bla Test For	cctagatccttttaaattaaaaatg

bla Test Rev	tcaaatatgtatccgctcatgagac
Scab Nec1 test for	atgagcgcgaacggaagcccc
Scab Nec1 test rev	ctactttctcgttatccatat
agarase test for	gtggtcaaccgacgtgatctc
agarase test rev	ctacacggcctgatacgtcct

Table 2.4 Oligonucleotides used during this thesis.

Solution	Composition
Tris/Borate/EDTA buffer (TBE)	90 mM Tris Base 90 mM Boric Acid 2 mM EDTA
TE buffer	10 mM Tris-HCl (pH 7.0) 1 mM EDTA
DNA loading buffer	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene-cyanol blue 40% (w/v) sucrose in water
DNA extraction 'Solution 1'	50 mM Tris-HCl pH 8.0 10 mM EDTA
DNA extraction 'Solution 2'	200 mM NaOH 1% (v/v) SDS
DNA extraction 'Solution 3'	3 M Potassium Acetate pH 5.5
Tris/Glycine/SDS (TGS) Buffer	25 mM Tris 192 mM glycine 0.1% (w/v) SDS
Transfer Buffer	25 mM Tris 192 mM glycine 20% (v/v) ethanol
Tris buffered saline (TBS) buffer	20 mM Tris-HCl (pH 7.6) 137 mM NaCl
TBS Tween buffer	TBS + 0.1% (v/v) Tween
Blocking solution	TBS + 0.1% (v/v) Tween 5% Skimmed Milk Powder
Enhanced Chemiluminescence (ECL) 'Solution A'	100 mM Tris-HCl pH 8.5 0.4 mM coumaric acid in DMSO 2.5 mM Luminol
ECL 'Solution B'	100 mM Tris-HCl pH 8.5 0.02% (v/v) Hydrogen Peroxide
TCB ultracentrifuge buffer	100 mM Tris pH 8.0 50 mM NaCl

Table 2.5 Solutions and reagents used during this thesis.

2.2.5 General restriction digest.

When digesting DNA with a single restriction enzyme typically a 20 µl reaction was set up using; 1 µl of the required enzyme, 2µl of the corresponding buffer (Roche), 10µl of DNA and 7 µl sH_2O. A digest was incubated at 37°C for 1 hour. When a double digest was required, 1 µl of the second enzyme was added at the expense of 1 µl dH_2O.

2.2.6 DNA ligation.

DNA ligation reactions took place in a total volume of 10 µl containing 1x ligation buffer and 1 unit of T4 DNA ligase (Roche), tubes were placed in a float in a beaker of room temperature water and chilled in a refrigerator to 4°C overnight. Concentrations of both vector and insert DNA were obtained using a nanodrop ND2000c (Thermo Scientific), with volumes of vector and insert being adjusted so that the molar ratio was approximately 3:1, with total DNA concentration being approximately 10 µg/ml. The following day the ligations were PCR purified (Qiagen) as per the manufacturers instructions, and electroporated into *E. coli* strain DH5α.

2.2.7 Preparing and transforming $CaCl_2$ competent cells.

An *E. coli* overnight culture was diluted 1/50 in 10 ml fresh LB and grown at 37°C until the OD_{600} reached between 0.3-0.6. The cells were then recovered by centrifugation, the media decanted and the cell pellet resuspended in 5ml ice cold $CaCl_2$ (glycerol). The cells were left on ice for a minimum of 30 minutes. The cells were then recovered by centrifugation and resuspended in 1 ml ice cold $CaCl_2$ (glycerol). The cells were left on ice for a minimum of 10 minutes then aliquoted out into 500 ml aliquots for storage at -80°C. 10 µl of the plasmid to be ligated was mixed with 200 µl of $CaCl_2$ and plated onto pre-warmed plates containing the appropriate antibiotics.

2.2.8 Preparation of electrocompetent cells.

E. coli cells to be electroporated were grown overnight from glycerol stocks with the relevant antibiotics in 10 ml of growth medium. The overnight culture was diluted 1/100 into 10 ml of fresh LB and grown shaking for 3-4 hours until they

reached an OD_{600} of approximately 0.4. The cells were recovered by centrifugation at 4000 rpm in a desktop centrifuge at 4°C. Once recovered the supernatant was poured off and the cell pellet was gently resuspended in 10 ml of ice-cold 10% glycerol. The cells were recovered by centrifugation as above, with the pellet resuspended in 5 ml of ice-cold glycerol. The cells were again centrifuged, the supernatant decanted and the cells resuspended in the remaining volume of glycerol (~100 µl). Cells were then divided into 50 µl aliquots and stored at -20°C.

2.2.9 Electroporating cells.

Cosmid DNA (2 µg) was used to transform 50 µl of pre-prepared electrocompetent cells in an ice cold electroporation cuvette using the BioRad® Electroporator set to: 200 Ω, 25 µF and 2.5 kV. Immediately following electroporation, 1 ml of ice cold LB broth was added to the cuvette. The entire contents of the cuvette was transferred to a 1.5 ml microfuge tube and the cells were incubated shaking at 30ºC for 1 hour. A 0.5 ml sample of these cells were plated onto an agar plate containing the relevant antibiotics and incubated overnight.

2.2.10 Agarose Gel electrophoresis.

Assessing the sizes of DNA fragments required running them in agarose gels containing 1% (w/v) agarose in 1x TBE buffer (table 2.5) and 1 µl ethidium bromide. Samples were mixed with 0.25 volumes of DNA loading buffer (table 2.5) and the gels were run in 1x TBE buffer at 100v for approximately 1 h. A 1Kb DNA marker ladder (Roche) was co-electrophoresed alongside the samples, and the DNA was visualised by exposure to UV light.

2.2.11 Extraction of DNA from agarose gels.

DNA fragments of interest were excised from an agarose gel using a scalpel and extracted using a Qiaquick Gel Extraction Kit (Qiagen), according to the manufacturer's instructions. The DNA was eluted in 50 µl dH_2O.

2.2.12 PCR purification.

The Qiaquick PCR purification kit (Qiagen) was used to remove any incorporated nucleotides, primers, enzymes etc. from a PCR reaction. The DNA was eluted in 50 µl dH_2O.

2.2.13 Precipitating DNA with Ethanol.

To precipitate DNA from a solution, 1/10 volume of 3M Sodium Acetate was added to the sample which was then mixed by inversion. Two volumes of ice cold 100% ethanol was then added followed by mixing by inversion. The mixture was centrifuged at 13000 rpm in a bench top centrifuge for 10 minutes. Post centrifugation, all ethanol was removed by pipetting and 200 µl of 70% ethanol was added to the pellet which was mixed by inversion. The sample was centrifuged at 13000 rpm for 5 minutes, and all ethanol was removed. The microfuge tube was left open for 2 minutes to air dry the DNA pellet which was then resuspended in either 30 µl TE buffer (table 2.5) containing 2 µl RNase (Sigma-Aldrich), or 30 µl autoclaved distilled water.

2.2.14 Chromosomal DNA preparation from *Streptomyces.*

An overnight culture of *Streptomyces* grown in 50% TSB/YEME media at 30 °C was centrifuged at 4000 rpm in a desktop centrifuge for 5 minutes. The supernatant was discarded and the cell was resuspended in 500 µl of Solution

1 (table 2.5) and transferred to a 1.5 ml microfuge tube. 10 µl of filter sterilised lysozyme (30 mg/ml) and 5 µl DNase-free RNase (10 mg/ml) (Roche) were added and the mycelial pellet was incubated for 1 hour at 30°C. After incubation, 5 µl of 20% SDS was added and the solution was mixed by inversion. One volume of 1:1 phenol-chloroform (approximately 500 µl) was added and the solution was mixed thoroughly by vortexing for 1 minute and centrifuged on a bench-top centrifuge at full speed for 5 minutes. Post-centrifugation, the solution had separated and the upper-aqueous level containing the DNA was carefully removed and transferred to a fresh microfuge tube. The phenol-chloroform steps were repeated until the top aqueous level was clear, indicating the all protein had been removed. The clear aqueous layer was transferred to a fresh microfuge tube, and 1 ml of 100% ethanol was added. The tube was mixed by inverting the tube several times followed by centrifuging at maximum speed for 5 minutes. All ethanol was removed and the resultant cell pellet was washed in 200 µl 70% ethanol and centrifuged for 2 minutes at maximum speed. All ethanol was removed and the pellet was dried for one minute at room temperature. The DNA was resuspended in 50 µl of sterile water. Where necessary the microfuge tube was warmed to 37 °C for 15 minutes to dissolve the DNA pellet.

2.3 Constructing gene knockouts.

2.3.1 Generating a knockout PCR product.

The antibiotic resistance cassette of choice, typically apramycin (*apr)* or hygromycin (*hyg*), containing the origin of transfer *oriT* was amplified using knockout (KO) primers specific for the gene to be disrupted (see table 2.4). The details of the plasmids containing the antibiotic cassettes, as well as instructions

regarding the design of the KO cosmids was found at http://streptomyces.org.uk/. The forward primer contains 39 nucleotides (nt) of the upstream coding region ending ATG, the translational start codon of each gene, plus a 20 nt P1 sequence corresponding to the 5′ end of the antibiotic resistance cassette. The reverse primer had 39 nt of antisense sequence ending TGA, the translational stop codon of each gene plus a 19 nt sequence corresponding to the end of the antibiotic resistance cassette. The PCR amplification included an initial denaturing step of 94°C/2 minutes followed by 10 cycles with a denaturing step of 94°C/45 seconds, a primer annealing step of 50°C/45 seconds and an extension step of 72°C/90 seconds, followed by 15 cycles with a denaturing step of 94°C/45 seconds, a primer annealing step of 55°C/45 seconds and an extension step of 72°C/90 seconds with and a final single extension of 72°C/5 minutes. PCR products were checked by agarose gel electrophoresis, excised and gel extracted then stored at -20ºC until use.

2.3.2 Introducing cosmids into *E. coli*.

Cosmids from *S. coelicolor* or *S. scabies* were obtained from the *Streptomyces* group at the John Innes Centre (JIC) or Prof Rose Loria's group at Cornell University, respectively. These cosmids contained the wild-type copy of the gene to be targeted for replacement. To check the *S. coelicolor* cosmids were correct they were digested with *Sac*I (1µl *Sac*I, 2 µl Buffer A, 17 µl cosmid DNA) at 37ºC for 1 hour. The *S. scabies* cosmids were digested with *Bam*HI (1µl *Bam*HI, 2 µl Buffer B, 17 µl cosmid DNA) at 37ºC for 1 hour. All digests were separated on a 1% agarose gel with the results checked either *in silico* at http://streptomyces.org.uk/redirect/res_cosmid2.html or, for *S. scabies* cosmids, via information received from Prof Rose Loria's group at Cornell university.

Cosmid DNA (2 µg) was used to transform 50 µl electrocompetent *E. coli* BW25113/pIJ790 as described in section 2.2.9. The entire contents of the cuvette were transferred to a 1.5ml microfuge tube and the cells were incubated shaking at 30ºC for 1 hour. A 0.5 ml sample of these cells were plated onto LB-agar containing; ampicillin, kanamycin (to select for the incoming cosmid) and chloramphenicol (to select for pIJ790), at concentrations described in table. The plates were incubated at 30°C overnight. A single colony was picked from each plate and inoculated into a 10ml LB broth containing the same antibiotics. These cultures were grown shaking overnight at 30°C.

2.3.3 PCR-targeting the *S. coelicolor* cosmid.

An overnight culture (0.1 ml) of *E. coli* BW25113/pIJ790 containing the cosmid of choice was inoculated into 10 ml LB broth plus antibiotics (as in the previous section) and 100 µl 1M L-arabinose, and grown shaking at 30°C for 4 hours. The arabinose induces the λ red genes carried on pIJ790 allowing transformation with linear DNA [84].

The cells were made electrocompetent and 50µl cells were electroporated with 2 µl of the antibiotic knockout PCR corresponding to the gene of choice. The cells were incubated shaking at 37°C for 1 hour and then plated onto LB agar containing ampicillin and kanamycin, to select for the cosmid, and apramycin, to select for the gene disruption. The plates were incubated at 37°C overnight to select for the loss of pIJ790. Single colonies were selected and grown in 10 ml LB broth plus antibiotics for 15 hours at 37°C.

2.3.4 Checking the mutagenised cosmid.

The cosmids were extracted from overnight cultures as in section 2.2.2 and were checked for gene disruption using PCR. Test primers for the disrupted genes of choice were used in combination with primers P1 and P2 which anneal to the antibiotic resistance cassette (see table 2.4) to ensure the knock-out was successful. The cosmids were also digested with *Sac*I or *Bam*HI to check for disruption.

2.3.5 Conjugating the cosmids into *Streptomyces*.

Both *S. coelicolor* and *S. scabies* contain a methyl-sensing restriction system. The disrupted cosmid must therefore be passaged through the non-methylating *E. coli* strain ET12567 before introduction into the target *Streptomyces* species.

Cosmid DNA (2 µl) was electroporated into 50 µl of electrocompetent ET12567 cells containing the driver plasmid pUZ8002 (table 2.2). Following electroporation the cells were plated onto LB agar containing; ampicillin, apramycin and chloramphenicol to select for the incoming cosmid, and to retain the *dam* mutation. The plates were incubated at 37ºC overnight. Single colonies were selected and grown in 10 ml LB broth plus antibiotics at 37ºC overnight.

Overnight culture (300 µl) was diluted into 10 ml fresh LB broth plus antibiotics and grown shaking for 4 hours at 37ºC. Cultures were centrifuged for 5 minutes at 13,000 rpm to recover the cells. The pellet was then washed twice with fresh LB broth, to remove any antibiotics potentially harmful to *S. coelicolor* or *S. scabies*. The cell pellet was then resuspended in 1 ml fresh LB broth.

Spores (10 µl) of the required strain were added to 500 µl LB broth, and heat shocked at 50°C in a water bath for 10 minutes. They were left to cool for 15 minutes at room temperature.

Washed *E. coli* cells (500 µl) were mixed with 500 µl of heat shocked spores. The mixture was centrifuged briefly, the supernatant was poured off and the remaining volume (about 50 µl) was used to resuspend the pellet. The entire mixture was then plated out onto SFM agar for *S. coelicolor*, or IMA agar for *S. scabies* and incubated at 30°C for 16-20 hours.

The following day the SFM or IMA plates were overlaid with 1 ml of sterile H_2O containing 0.5 mg naladixic acid (an antibiotic used to selectively kill *E. coli*) and either 1.25 mg apramycin or 0.625 mg hygromycin (to select for recombination of the incoming cosmid with the streptomycete chromosome). A sterile spreader was used to gently distribute the antibiotic solution evenly. Once overlaid the plate was incubated at 30°C for four days or until colonies appeared. These colonies were replicated using velvets onto DNA agar plates containing nalidixic acid and apramycin (or hygromycin) and DNA plates containing nalidixic acid and kanamycin. These were incubated at 30°C for two days.

Any double cross-over exconjugants would be apramycin or hygromycin resistant and kanamycin sensitive as the *apr* or *hyg* cassette has replaced the chromosomal copy of gene in the genome. These colonies were picked from the original conjugation plate and plated onto an SFM or IMA plate containing

apramycin or hygromycin and nalidixic acid (as above). This plate was grown at 30°C for five days. This cycle of apramycin/hygromycin selection followed by velvet replica plating was done a minimum of three times to ensure loss of all kanamycin resistant single crossovers. Spores of these double cross-over exconjugants were prepared and stored at -20°C.

2.3.6 Excising the gene disruption cassette using FLP recombinase.

The antibiotic disruption cassette found in the gene disruption contains FLP recognition targets (FRT) sites at the 5′ and 3′ ends. The enzyme FLP recombinase is able to remove the cassette between these FRTs leaving an 81bp 'scar'.

E. coli DH5α/BT340 (table 2.2) was grown shaking at 30°C in 10 ml LB broth containing chloramphenicol (to retain the BT340 plasmid) overnight. A fresh 10 ml LB broth was inoculated using 100 µl of this overnight culture and grown with chloramphenicol for 4 hours. These cells were made electrocompetent and electroporated with the apramycin cassette disrupted cosmid. The electroporation was plated out onto LB agar containing apramycin (to select for the incoming cosmid) and chloramphenicol (to retain BT340). The plate was incubated at 30°C for 2 days. A single colony was selected and restreaked onto LB agar containing no antibiotics, which was in turn plated at 42°C for 15 hours to induce expression of FLP recombinase followed by loss of the BT340 plasmid. The cosmids were checked using PCR primers against the gene of choice. If the *apra* cassette had been removed by the FLP recombinase an 81bp PCR fragment was amplified and could be detected by electrophoresis.

2.3.7 Introduction of *oriT* to a FLP cosmid.

The removal of the apramycin cassette by the action of FLP recombinase also removes the origin of transfer (*oriT*) found within it. In order for the knockout FLP cosmid to be introduced into the *S. coelicolor* chromosome, a new *oriT* must be introduced to the cosmid. This was done by amplifying the apramycin cassette from plasmid pIJ773 (table 2.2) which contains an *oriT* using the apramycin specific primers (known as P1 and P2) with flanking sequence against the *bla* gene in the cosmid backbone (table 2.4). Once amplified, this cassette was transformed into electrocompetent *E. coli* BW25113/pIJ790 (as described above) containing the FLP knockout cosmid of choice. Following electroporation the cells were plated onto LB agar containing ampicillin and apramycin to select for the altered cosmid. The plates were incubated at 37ºC overnight.

Once checked by PCR, the FLP cosmids were transformed into electrocompetent Et12567/pUZ8002 as before, and stored at -20°C. This strain was conjugated into the required *Streptomyces* species to create an unmarked gene deletion. Single exconjugants were selected by growing potential unmarked deletion mutants on SFM media containing kanamycin to select for the incoming FLP cosmid. Kanamycin resistant colonies were selected and streaked for single colonies onto SFM containing no antibiotics. These plates were grown for 4-5 nights then the colonies were replica plated onto DNA agar plates containing either apramycin or kanamycin. Colonies that were sensitive to both represented potential double exconjugants and were selected and tested by PCR for the FLP scar.

2.4 Gene Complementations

2.4.1 Complementation with full length and truncated *lsp*.

The unmarked deletion strain of *S. coelicolor lsp, lsp*Flp was complemented with both full length, and N-terminally truncated versions of the gene. The truncations removed either the first 10, 20, 30 or 40 amino acids of the gene, and were designed so that the first codon was in the same position as the annotated *lsp* start codon with an additional 300 bp upstream DNA, containing the *lsp* promoter. The constructs were synthesised by Genscript USA Inc. and cloned into the vector pUC57. Each allele was excised from the pUC57 by digestion with *BamHI* and *EcoRI* restriction enzymes and sub-cloned into pSET152, to make the plasmids listed in table 2.2. Each plasmid was transformed into ET12567 / pUZ8002 by electroporation, and transformants were selected on LB agar containing apramycin. The transformed strains were conjugated into *S. coelicolor* M145 *ΔlspFLP* and plated onto SFM medium. Exconjugants were selected using apramycin resistance.

2.4.2 Complementation of *lsp::*apr with 4A10.

In order to reverse the apramycin disruption of the *lsp* gene, the wild type gene was reintroduced into the strain M145 *lsp*::*apr.* The cosmid 4A10, containing the *lsp* gene, was electroporated into *E. coli* strain BW25113/pIJ790 and plated onto LB-agar containing ampicillin, kanamycin and chloramphenicol. The plate was incubated at 30°C overnight. A single transformant was picked and grown in 10 ml LB broth containing the same antibiotics and grown shaking overnight at 30°C. Overnight culture (0.1 ml) was diluted into 10 ml fresh LB broth plus antibiotics plus and 100 µl 1M L-arabinose and grown for approximately 4 hours at 30°C. These cells were made electrocompetent, and the ampicillin

resistance gene (*bla*) found on the backbone of the cosmid was replaced by electroporating the cells with a hygromycin resistance cassette (containing *oriT*) that was amplified with primers bla replacement For and bla replacement Rev (table 2.4). Hygromycin resistant, ampicillin sensitive colonies were selected and confirmed by PCR using the bla Test For and bla Test Rev primers (table 2.4). The cosmid was extracted and transformed into *E. coli* ET12567 / pUZ8002 by electroporation and conjugated with the strain M145 *lsp*Flp Selection for single exconjugants involved picking colonies that were hygromycin resistant, kanamycin resistant and apramycin sensitive. After growth on SFM agar in the absence of antibiotics, double exconjugants were selected by identifying colonies that were hygromycin sensitive, kanamycin sensitive and apramycin sensitive, with genomic DNA tested by PCR, using the Lsp Test For and Lsp Test Rev primers to ensure the full length gene was present.

2.4.3 *E. coli* complementation.

E.coli lsp was amplified from genomic DNA using primers EcolspA For and EcolspA Rev. *S. coelicolor lsp* and *S. coelicolor* N40 *lsp* were amplified using pBT100 and pBT105 as templates, respectively. PCR primers ScoFL LspA For and Sco LspA Rev were used for *S. coelicolor lspA* with ScoN40 LspA For and Sco LspA Rev used for *S. coelicolor* N40 *lspA.* Each of the forward primers incorporated an *Nde*I restriction site, and each of the reverse primers incorporated a *Bam*HI restriction site. Once amplified, each PCR product run on a 1% agarose gel, extracted and gel purified (sections 2.11, 2.12) was cloned into PCR2.1-Topo (Invitrogen). Following cloning, each plasmid was transformed into competent Top10 *E. coli* cells (Invitrogen) and plated onto LB-

agar, containing ampicillin, X-Gal and IPTG and incubated at 37°C overnight. White colonies were selected from the overnight plates, grown for 15 hours in LB-media and ampicillin at 37°C. Plasmids were extracted from each overnight, digested with *Nde*I and *Bam*HI and ligated into the vector pBAD24-*Nde*I to make vectors pBT106-109 (table 2.2), such that each inserted gene is under the control of the arabinose-inducible pBAD promoter. Each clone was confirmed by restriction digest and sequencing and then used to transform *E. coli* strain BW25113 / pIJ790 by electroporation. The incoming plasmid was selected for by growth on LB-agar containing ampicillin, with kanamycin and chloramphenicol, and grown at 30 °C in order to maintain pIJ790. The chromosomal copy of *lsp* was disrupted by PCR targeting it using an apramycin cassette containing *lspA* flanking DNA amplified using the EcolspA KO For and EcolspA KO Rev PCR primers. Transformants were plated onto LB agar containing ampicillin, apramycin, and arabinose, to express the *lsp* alleles *in trans*.

2.5 Protein Methods

2.5.1 Protein preparation for electrophoresis.

Mycelium from an overnight culture grown shaking at 30°C in 10 ml of 50% TSB/YEME media was harvested by centrifuging in a bench-top centrifuge at 4000 rpm for 5 minutes. All medium was discarded and the pellets were washed in 1 ml of TCB (table 2.5) and transferred to a microfuge tube. The pellet was centrifuged briefly, the TCB removed and was resuspended in a volume of fresh TCB ranging from 100µl to 500µl depending on the size of the pellet. The required volume of 50x EDTA free protease inhibitor (Roche) was added to the mix.

The suspensions were sonicated, on ice, at 50 kHz for 5 seconds, followed by one minute chilling on ice. Each sample was sonicated an average of 5 times. Following sonication the samples were centrifuged in a bench-top centrifuge at 15,000 rpm for 1 minute and the supernatant stored at -20°C as crude protein extract.

2.5.2 SDS-PAGE.

15% SDS polyacrylamide gels were cast using a mini Protean 3 system as per the manufacturer's instructions (BioRad). In addition to the 15% polyacrylamide, the resolving gels contained 0.325M Tris-HCl pH 8.8, 0.1% (w/v) SDS, 0.01% (v/v) TEMED and 0.1% (v/v) ammonium persulphate (APS). The mixture was poured between two glass plates to a distance of approximately 2.5 cm from the top of the plates and was layered with 100% ethanol. Once polymerisation had occurred, the ethanol was removed and the gel was washed with dH_2O. All water was removed using filter paper, and the stacking gel was layered on top. The stacking gel consists of 4% polyacrylamide, 0.125M Tris-HCl pH 6.8, 0.1% (w/v) SDS, 0.01% (v/v) TEMED and 0.1% (w/v) APS. A comb was inserted into the top layer, and once polymerisation had occurred, the comb was removed and the gel was fastened into its electrode, and placed into the electrophoresis tank, which was filled with TGS buffer (table 2.5). The concentration of protein in samples was ascertained using a BioRad protein concentration assay kit. Equal concentrations of protein were loaded into each well, having first been mixed with one volume of Laemmli buffer (BioRad) + 5% 2-mercaptoethanol and boiled at 95°C for 2 minutes. The samples were run at 200 volts for 1.5 hours.

2.5.3 Semi-dry immunoblotting.

For immunoblotting, proteins immobilised within the gel were transferred to a polyvinylidene difluoride (PDVF) membrane using a Bio-Rad Semi Dry Transfer Cell. Following electroporation, the gel was washed in transfer buffer (table 2.5), along with 12 pieces of 1 mm filter paper (Whattman) cut slightly larger then the gel. Six pieces of filter paper were placed in the Transfer Cell with a sheet of PVDF membrane (BioRad), cut to a slightly smaller size and presoaked in methanol and then transfer buffer was lain on top. The gel was then placed on top of the membrane, with the final 6 sheets of filter paper rested on top. Pressure was placed on this sandwich to remove any air bubbles, and any excess transfer buffer was removed. The Transfer Cell was assembled and run at 15 volts for 1 hour. Following the run, the membrane was blocked overnight, shaking, in 50 ml blocking solution (table 2.5). The following day the blocking solution was removed, and 10 ml of fresh blocking solution containing either 8 µl of CseA primary antibody or 2 µl Anti-His-HRP antibody was added. The membrane was incubated in this solution on a shaker for 1 hour at room temperature. After incubation the membrane was washed twice in 20 ml 1 x TBS + 1% Tween. Each wash was 10 minutes long. Following these washes, the TBS/Tween was removed and 10 ml of fresh blocking solution containing 2.5 µl of secondary antibody (HRP-linked anti-rabbit IgG) was added to the CseA blot. The Anti-His antibody is an HRP conjugate and requires no secondary antibody. The membrane was incubated for 1 hour. Once this incubation was complete the membrane was washed twice as above. 2 ml of developing solution was mixed immediately prior to development (table 2.5), and then pipetted over the membrane, ensuring the entire surface was covered, and then incubated at room temperature for 1 minute. Once incubated the

membrane was placed in an Hypercassette (GE Healthcare), covered with clingfilm and overlaid with X-ray film (FujiFilm). The X-ray film was exposed for 1 minute, with fresh exposures taken at increasing/decreasing time points as required. Films were developed using a Konica-Minolta SRX-101A development machine.

2.5.4 Sub-cellular fractionations.

In order to separate the cell membrane from the cytoplasm, crude cell extracts were ultracentrifuged. Crude extracts were placed in a 1 mm thick walled centrifuge tube (Beckman) and were balanced by volume. The tubes were spun for 1 hour at 80,000 rpm at 4°C in a Beckman centrifuge. Following centrifugation the supernatant (cytoplasmic fraction) was removed and stored at -20°C. 200 µl of fresh TCB buffer was added and the tubes were respun as above. The supernatant (wash fraction) was removed and discarded, leaving the cell membrane pellet. This was resuspended by gentle pipetting in 50 - 200 µl of TCB + 1% Sarcosyl depending on the size of the pellet. This membrane fraction was then stored at -20°C.

2.6 Microscopy.

2.6.1 Light Microscopy.

Brightfield images were acquired using a Zeiss M2 Bio Quad SV11 stereomicroscope. The samples were illuminated with a halogen lamp and reflected-light images captured with an AxioCam HRc CCD camera and AxioVision software (Carl Zeiss, Welwyn Garden City, UK).

2.6.2 Scanning Electron Microscopy.

Samples were mounted on an aluminium stub using Tissue Tek^{R} (BDH Laboratory Supplies, Poole, UK). The stub was then immediately plunged into liquid nitrogen slush at approximately -210°C to cryo-preserve the material. The sample was transferred, onto the cryostage of an ALTO 2500 cryo-transfer system (Gatan, Oxford, UK) attached to a Zeiss Supra 55 VP FEG scanning electron microscope (Zeiss SMT, Germany). Sublimation of surface frost was performed at -95°C for three minutes before sputter coating the sample with platinum for 3 mins at 10mA, at colder than -110°C. After sputter-coating, the sample was moved onto the cryo-stage in the main chamber of the microscope, held at approximately -130°C. The sample was imaged at 3 kV and digital TIFF files were stored.

2.7 Phenotype Assays.

2.7.1 Antibiotic sensitivity.

All phenotype assays were conducted in a standard 12 well cell culture plate. Each well contained 3 ml minimal media (+1% glucose). When strains were being tested for sensitivity to antibiotics, the required concentration of the relevant antibiotic was dissolved into the minimal medium before it was poured into the 12 well plate. The highest concentration was found in well A1, whilst the lowest was found in well C3. A control well containing no antibiotics was found in C4. 100 µl of sterile water containing 1 µl spores was placed onto each well and thoroughly spread to cover the entire surface of the growth media. The plates were sealed and incubated at 30°C for 5 days. Growth was assessed as either ‘confluent’, ‘weak’ where there were sufficient colonies to count individually, or ‘none’ where there was an absence of any colonies.

2.7.2 Lysozyme/SDS sensitivity.

When strains were assessed for sensitivity to lysozyme or SDS, they were they were tested in three different ways. Firstly 100 µl of sterile water containing 1 µl spores was placed onto each well of a plate containing 3 ml minimal media. The plates were dried in a sterile flow hood, and once fully dried, they were overlaid with sterile water containing differing concentrations of lysozyme, or percentages of SDS. The plates were dried as before, sealed and incubated at 30°C for 5 days. A variance of this method involved the spores being placed onto the plates and dried as above, but grown for 15 hours at 30°C prior to overlay with the lysozyme or SDS. Once overlayed the plates were reincubated for a further 5 days. The final assay involved preincubation of the spores in sterile water containing differing concentrations of lysozyme or percentages of SDS for one hour, prior to plating onto 12 well plates.

2.8 Virulence assays.

2.8.1 Potato disc tuber assay.

A potato tuber of the cultivar Maris Piper, was washed and peeled. The tuber was surface sterilised by immersion in 2% bleach for 5 minutes, with stirring. The potato was transferred using sterile forceps into a beaker containing 500 ml of sterile water in a laminar flow hood. The potato was rinsed for 2 minutes, and transferred to a second beaker of water where it was washed as before. After washing was completed, the potato was lain on some sterile paper towels and cores were taken using a sturdy potato peeler (sterilised with ethanol). These cores were sliced into discs 0.5 cm thick using a sterile scalpel. The slices were placed onto sterile filter paper, prewetted with 2 ml of sterile water, in a petri dish. Four potato slices were placed in each dish. Plates of the *S.*

scabies strains were prepared on IM agar. Spores were plated in order to obtain confluent lawns of bacteria. The plates were incubated for 5 nights, and agar plugs (1 cm^2) were cut using a sterile scalpel. A plug was placed on each potato piece, spore side down. Uninoculated IM agar was used as a negative control. The plates were sealed and incubated at 30 °C in the dark for either 2 nights, or 7 nights. The agar plugs were then removed and the potato slices were viewed under a microscope to investigate signs of necrosis.

2.8.2 Radish seedling root virulence assay.

Radish seeds of the cultivar Scarlet Globe were soaked in 70% ethanol for 10 minutes, the ethanol removed, then soaked in 13% bleach for a further 10 minutes. The bleach was removed, and the seeds were washed copiously in sterile H_2O. The seeds were placed onto sterile filter paper in a petri dish, prewetted with 2 ml sterile H_2O. The seeds were incubated in the dark at room temperature for 24 hours. Cultures of various *S. scabies* strains were set up by growing 1 x 10^6 spores in 10 ml of 50% TSB/YEME media, with the relevant antibiotics at 30° C for 20 hours. After this time had elapsed the mycelium was collected by centrifugation and washed twice in TSB. After the final wash the mycelium was resuspended in 1 ml of TSB broth, with 500 µl being spread over half a square petri dish containing plant MS medium (table 2.3) (1% agar). The remaining 500 µl of mycelium were placed in a 12 well cell culture plate, into which seedlings at the same stage of germination were immersed prior to being placed in the petri dish. These dishes were sealed and stored vertically in a incubation chamber at 21° C with a day length of 12 hours.

Chapter 3 - The lipoprotein biosynthetic pathway of *S. coelicolor.*

3.1 Introduction

This chapter focuses on the biosynthetic pathway by which *S. coelicolor* processes its lipoproteins, from the cytoplasm to the fully functional cell surface protein. The lipoprotein biosynthetic pathway in *S. coelicolor* consists of several different steps, starting with translocation of the protein across the cell membrane. This usually occurs via the Sec pathway but *S. coelicolor* can also export fully folded, but unlipidated lipoproteins through the twin arginine translocase (Tat) pathway and this appears to be unusual amongst bacteria [82]. Following translocation the protein is anchored in the bacterial membrane by its signal sequence. The protein is tethered to the membrane by the addition of a diacylglyceryl lipid moiety to the conserved cysteine residue found in all lipoprotein signal sequences. Addition of this lipid is catalysed by the membrane bound enzyme Lgt (Lipoprotein diaclyglyceryl transferase). Once lipidated, the Lsp enzyme (Lipoprotein signal peptidase) cleaves the signal peptide from the anchored, lipidated protein to produce the mature lipoprotein. As mentioned in Chapter 1 this pathway is strictly ordered in Gram negative bacteria [14] and appears, with few exceptions, to be similarly controlled in Gram positive bacteria. There are some unusual aspects to the lipoprotein biosynthetic pathway in *S. coelicolor* which will be examined in detail below. Whilst there is only one copy of the *lsp* gene as in other bacteria, there are two putative *lgt* genes, named in this study as *lgt1* (*Sco2034*) and *lgt2* (*Sco7822*). There are also two copies of *lnt*, *lnt1* (*Sco1014*) and *lnt2* (*Sco1336*). What role these genes have is currently unknown, however, as discussed in Chapter 1, the recent discovery of a fully functional Lnt protein in *M. smegmatis*, does suggest that one or both may serve their predicted function.

3.2 Aims.

The aims for this chapter are two-fold. Firstly, it will focus on the bioinformatic analysis of the lipoprotein biosynthetic enzymes from a variety of bacterial species and analysis of the functions of the predicted 223 lipoproteins found in *S. coelicolor.* These functions will be related to any potential developmental phenotypes gained from the second aim, which is to obtain and analyse mutants in each of the lipoprotein synthesis genes found in *S. coelicolor*, using the REDIRECT system. These mutants will be examined both macroscopically and microscopically, using light and electron microscopy. Two model lipoproteins, CseA (SCO3357) and SCO3484 will be immunoblotted to examine their processing in each of the mutants, either natively, or fused to a hexa-histidine peptide. Each of the mutants will be subjected to a variety of stresses, to gauge the effect the mutations are having on the bacteria. Any mutants which exhibit a drastic phenotype will be complemented to see whether it is possible to restore the wild-type phenotype.

3.3 Results.

3.3.1 Identification and comparison of the lipoprotein biosynthetic enzymes.

In order to learn more about the lipoprotein biosynthetic enzymes in *S. coelicolor* the sequences were first identified by BLASTP analysis and then aligned with the primary sequences of homologous enzymes from other bacteria. In total, eight species were chosen for sequence comparisons. Alongside *S. coelicolor* (Sco), three other streptomycetes were chosen: *S. scabies* (Scab), *S. avermitilis* (Sav) and *S. griseus* (Sgr). Two related actinomycetes *M. tuberculosis* (Mtb) and *M. smegmatis* (Msm) were also

chosen, primarily due to the confirmed Lnt enzyme in *M. smegmatis*. *B. subtilis* (Bsu) was used to represent the low GC Gram-positive bacteria. Finally, the model Gram-negative bacterium *E. coli* (Eco) was chosen, as much of the work involved with identifying essential residues in Lgt and Lnt has been carried out in *E. coli*.

3.3.1.1 Lipoprotein diacyglycerol transferase (Lgt).

The two Lgt enzymes from *S. coelicolor* share 60% identity. Both are predicted to contain seven transmembrane helices (http://www.enzim.hu/hmmtop/, http://www.cbs.dtu.dk/services/TMHMM-2.0/, http://www.ch.embnet.org/software/TMPRED_form.html). Both *S. coelicolor* enzymes show a high degree of identity with the other *Streptomyces* species (>70%), and both have ≥45% similarity with the MtbLgt enzymes. Both ScoLgt1 nor ScoLgt2 share a low identity with the Lgt enzymes from *E. coli* and *B. subtilis* (<30%), however the enzymes from these two species are equally dissimilar.

The Histidine residue at position 103 in the *E. coli* Lgt enzyme has been shown to be essential for function [17], but interestingly it is absent from all of the actinomycete proteins compared, where it is a tryptophan (fig. 3.1). Conversely, the two other amino acids important for function in *E. coli* , His-196 and Tyr-235 [17], are both present in all of the bacteria, with the exception of H196Q in *M. tuberculosis*.

```
ScoLgt1 ---MELAFIPSPSRGVLHLGPVPLRGYAFCIIIGVFVAVWLGNKRWVARGG--RPGTVAD 55
ScoLgt2 ---MDLAYLPSPSTGVLHLGPIPLRAYAFCIILGVFAAVWLGNRRWVARGG--KQGVIAD 55
ScabLgt ---MELAYIPSPSHGVYYLGPVPLRGYAFCIIIGVFVAVWLGNKRWVARGG--QAGTIAD 55
SavLgt  ---MELAYIPSPARGVLYLGPIPLRGYAFCIIIGVFVAVWLGNKRWVARGG--RPGTVAD 55
SgrLgt  ---MNLAFIPSPSTGVIELGPIPLRGYAFCIIIGVFVAVWFGNKRWVARGG--KAGTVAD 55
MtbLgt  -MRMLPSYIPSPPRGVWYLGPLPVRAYAVCVITGIIVALLIGDRRLTARGG--ERGMTYD 57
MsmLgt  MTTTVLAYLPSPSQGVWHLGPVPIRAYALCIIVGIVAALVIGDRRWQARGG--EPGVIYD 58
EcoLgt  -MTSSYLHFPEFDPVIFSIGPVALHWYGLMYLVGFIFAMWLATRRANRPGSGWTKNEVEN 59
BsuLgt  ----MNEAIEPLNPIAFQLGPLAVHWYGIIIGLGALLGLWIAMRESEKRGL--QKDTFID 54

ScoLgt1 IAVWAVPFGLIGGRL-YHVITDYQLYFSE-GRDWVDAFKIWEGGLGIWGAIAFGAVGAWI 113
ScoLgt2 VTLWAVPFGLVGGRL-YHVFTSPDAYFGE-RGEPVRALYVWEGGLGIWGAIALGAVGAWI 113
ScabLgt IAVWAVPFGLVGGRL-YHVITDYELYFSE-GRDWVDAFKVWEGGLGIWGAIALGAVGAWI 113
SavLgt  IAVWAVPFGLVGGRL-YHVITDYELYFSE-GRDWVDAFKIWEGGLGIWGAIALGAVGAWI 113
SgrLgt  VAVWAVPFGLVGGRL-YHVITDYQLYFSD-GEDWVDAFKIWEGGLGIWGAIAFGAVGAWI 113
MtbLgt  IALWAVPFGLIGGRL-YHLATDWRTYFGDGGAGLAAALRIWDGGLGIWGAVTLGVMGAWI 116
MsmLgt  IALWAVPFGLAGGRI-YHVITDWKTYFGPTGKGFGAALQIWEGGLGIWGAVAFGAVGAWI 117
EcoLgt  LLYAGFLGVFLGGRIGYVLFYNFPQFMAD----PLYLFRVWDGGMSFHGGLIGVIVVMII 115
BsuLgt  LVLFAIPIAIICARI-YYVAFEWDYYAAH----PGEIIKIWKGGIAIHGGLIGAILTGYV 109

ScoLgt1 GARRRGVPMPAYADAVAPGIALAQAIGRWGNWFNQELYGKAT-DLPWAVEIT-------- 164
ScoLgt2 GCRRHRIPLPAFADAVAPGIVLAQAIGRWGNWFNQELYGRPT-TLPWGLEIDR------- 165
ScabLgt GCRRRGIPLPAYADAVAPGIALAQAIGRWGNWFNQELYGKPT-DLPWAVEIT-------- 164
SavLgt  GCRRRGIPLPAWADAVAPGIAFAQAFGRWGNWFNQELYGRET-HVPWALHIT-------- 164
SgrLgt  ACRRRGIPLPAWADALAPGIAIAQAIGRWGNWFNQELYGKPT-DLPWALEIS-------- 164
MtbLgt  GCRRCGIPLPVLLDAVAPGVVLAQAIGRLGNYFNQELYGRET-TMPWGLEIFYRRDPSGF 175
MsmLgt  ACRLRGIPLPAFGDAIAPGIILAQGIGRLGNYFNQELYGRPT-DVPWGLEIYERLN--KF 174
EcoLgt  FARRTKRSFFQVSDFIAPLIPFGLGAGRLGNFINGELWGRVDPNFPFAMLFPGSRTEDIL 175
BsuLgt  FSRVKNLSFWKLADIAAPSILLGQAIGRWGNFMNQEAHGEAV-SRAFLENLHLP------ 162

ScoLgt1 ---------STADGRVPGTYHPTFLYESLWCIGVALLVIWADRRFKLGHGRAFALYVAAY 215
ScoLgt2 -------AHRPAGTLDIATYHPTFLYESLWNIGVAALILWAAKRFPLGHGRTFALYVAAY 218
ScabLgt ---------SSTDGRLPGTYHPTFLYESLWCIGVALLVIWADRRFTLGHGRAFALYVASY 215
SavLgt  ---------SSTDGRVPGYYHPTFLYESLWCVGVGFLVIWADRRFKLGHGRAFALYVAAY 215
SgrLgt  ---------EGPN-RVAGTYHPTFLYESLWCIGVALLVIWADRRFKLGHGRAFALYVAGY 214
MtbLgt  DVPN--SLDGVSTGQVAFVVQPTFLYELIWNVLVFVALIYIDRRFIIGHGRLFGFYVAFY 233
MsmLgt  GQSD--QLNGVSTGQVTAVVHPTFLYELIWNIAVFGFLIWVDRKFRIGHGRLFALYVASY 232
EcoLgt  LLQTNPQWQSIFDTYGVLPRHPSQLYELLLEGVVLFIILNLYIRKPRPMGAVSGLFLIGY 235
BsuLgt  ------EFIINQMYINGQYYHPTFLYESLWSF-VGVIVLLLLRRANLRRGEMFLIYIIWY 215

ScoLgt1 CAGRFWIEYMRVDDAHHILG----LRLNNWTALFVFLLAVLYIVLSARKR---------- 261
ScoLgt2 TVGRFGTEYLRIDEAHTFLG----LRLNNWTSVLVFLGAVACLVVSAHRH---------- 264
ScabLgt CTGRFWIEYMRVDEAHHILG----LRLNNWTALFVFVLAVIYMVLSARKR---------- 261
SavLgt  CVGRAWIEYMRVDDAHHILG----VRLNDWTAIAVFLLAVLYIVLSSRKR---------- 261
SgrLgt  CAGRGWIEYMRVDEAHHILG----LRLNVWTAIVVFILAVVYIVISAKIR---------- 260
MtbLgt  CAGRFCVELLRDDPATLIAG----IRINSFTSTFVFIGAVVYIILAPKGREAPGALRG-- 287
MsmLgt  CVGRFWVELMRSDTATEFAG----IRVNTFTSTFVFIGAVVYIMLAPKGREEPESLRGKA 288
EcoLgt  GAFRIIVEFFRQPDAQFTGAWVQYISMGQILSIPMIVAGVIMMVWAYRRS---------- 285
BsuLgt  SIGRYFIEGMRTDSLMLTDS----LRIAQVISIVLIVLAVAAIIFRRVKG---------- 261

ScoLgt1 ------------------------------------------------------------PGR- 264
ScoLgt2 ------------------------------------------------------------PGI- 267
ScabLgt ------------------------------------------------------------PGR- 264
SavLgt  ------------------------------------------------------------PGR- 264
SgrLgt  ------------------------------------------------------------PGR- 263
MtbLgt  --------SEYVVDEALEREPAELAAAAVASAASAVG----------------PVGPGEP 323
MsmLgt  ADETEEGDEESLVDEAGKELVAAAAGTGVVAAATAAAREDTDDTDGTTDPAAASDEPGDE 348
EcoLgt  ------------------------------------------------------------
BsuLgt  ------------------------------------------------------------

ScoLgt1 ---------------------------------------------EAVVEPGAETAAGDS 279
ScoLgt2 ---------------------------------------------ENVAR--------LQ 274
ScabLgt ---------------------------------------------EEVVEPGVSDGDGDE 279
SavLgt  ---------------------------------------------EEIVEPGAS----DT 275
SgrLgt  ---------------------------------------------EEIVEPDRD------ 272
MtbLgt  NQPD----------------------DVAEAVKAEVAEVTDEVAAESVVQVADRDGESTP 361
MsmLgt  AEPETEAATAVLTADGAEAMSVDEQPDDDEAVDEEAGQEAEEPADELVELADEPEGEPEP 408
EcoLgt  ------------------------------------------------------------
BsuLgt  ------------------------------------------------------------
```

```
ScoLgt1 GSAADKDVKGTKDAE-------DAEGAEDGAEKTDASGATEAPEDTSGADEADAAKDAEG 332
ScoLgt2 GAGADGRTD---------------DPRPADASVGLASGPPGNSTPRRATESWNVRNRS-- 317
ScabLgt GAGAAAPVE-------------DSETKSESKAKPGES-EAKAEGDAGASGGGGEKKEPE- 324
SavLgt  GTGADDPVDLGKD---------EDKATTDKATATDTSTTTDKSTDRGKNEDENEGEDAEP 326
SgrLgt  ATPAEKDGSG------------EDGSGEKGVAKADAAAKDPLTKDEPGKDATAENAGAAG 320
MtbLgt  AVEETSEADIERE---------QPGDLAGQAPAAHQVDAEAASAAPEEPAALASEAHDET 412
MsmLgt  AAAEEAEPAETIEPDEDEIFDAELAEALAEAAEDFAVVPESAGSDEAEVAAAESAADAED 468
EcoLgt  ---PQQHVS---------------------------------------------------- 291
BsuLgt  -YSKERYAE---------------------------------------------------- 269

ScoLgt1 ---VTNGADSAKKG---------------------------------------------- 343
ScoLgt2 ------------------------------------------------------------
ScabLgt ---------SAKKS---------------------------------------------- 329
SavLgt  SEKTESAAESAKKV---------------------------------------------- 340
SgrLgt  ---------AAEKA---------------------------------------------- 325
MtbLgt  EP------EVPEKAAP----IPDPAK---------------------------------- 428
MsmLgt  GPSDAESAEVPAEAAADDAVVSEPAGTVDETEVAEEAAAEAADLEFESYDAELAEALAEA 528
EcoLgt  ------------------------------------------------------------
BsuLgt  ------------------------------------------------------------

ScoLgt1 ------------------------------------------------------------
ScoLgt2 ------------------------------------------------------------
ScabLgt ------------------------------------------------------------
SavLgt  ------------------------------------------------------------
SgrLgt  ------------------------------------------------------------
MtbLgt  ---------PDELAVAGPGDDPAEPDGIRRQDD-------------------------- 452
MsmLgt  AEDMAVVVAPAETAESDEAREPVETDATESEEPSEPEEPDAPEAVEAPEALDENADETRA 588
EcoLgt  ------------------------------------------------------------
BsuLgt  ------------------------------------------------------------

ScoLgt1 ----------------------------
ScoLgt2 ----------------------------
ScabLgt ----------------------------
SavLgt  ----------------------------
SgrLgt  ----------------------------
MtbLgt  ------------FSSRRRRWWRLRRRRQ 468
MsmLgt  EPAAPAAVATAPVEPEKGRLRRWLRRNR 616
EcoLgt  ----------------------------
BsuLgt  ----------------------------
```

Figure 3.1 Amino acid sequence alignment of the Lgt enzymes. Potential transmembrane domains are underlined. The essential *E. coli* H103 is highlighted in red whilst the H196 and Y235, important for function in *E. coli* are highlighted in blue. Sequences compared using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

What is immediately obvious when comparing the actinomycetes to *E. coli* and *B. subtilis* is that Lgt proteins from the actinomycetes contain elongated, cytoplasmic, C-termini which are not present in *E. coli* and *B. subtilis*. Interestingly, these elongated sequences contain very little similarity to each other, with the *Mycobacterium* proteins being substantially longer than those from *Streptomyces*. When compared from amino acid 292 onwards (EcoLgt

terminates at amino acid 291), the two *S. coelicolor* enzymes only share 15% similarity, whilst the 30% similarity shared between ScoLgt2 and MtbLgt is the highest similarity found between any two actinomycete Lgt proteins. The amino acid comparison can be seen in figure 3.2.

```
ScoLgt1 ------------------------------------------------------------
ScoLgt2 ------------------------------------------------------------
ScabLgt ------------------------------------------------------------
SavLgt  ------------------------------------------------------------
SgrLgt  ------------------------------------------------------------
MtbLgt  ---------VDEALEREPAELAAAAVASAASAVGPVGPGEPN------------------ 33
MsmLgt  TEEGDEESLVDEAGKELVAAAAGTGVVAAATAAAREDTDDTDGTTDPAAASDEPGDEAEP 60

ScoLgt1 ------------------------------------------------------------
ScoLgt2 ------------------------------------------------------------
ScabLgt ------------------------------------------------------------
SavLgt  ------------------------------------------------------------
SgrLgt  ------------------------------------------------------------
MtbLgt  --------------------QPDDVAEAVKAEVAEVTDEVAAESVVQVADRDGESTPAVE 73
MsmLgt  ETEAATAVLTADGAEAMSVDEQPDDDEAVDEEAGQEAEEPADELVELADEPEGEPEPAAA 120

ScoLgt1 ----------------------------------------------------DAEDAEGAED 10
ScoLgt2 ------------------------------------------------------------
ScabLgt ------------------------------------------------------------
SavLgt  ----------------------------------------------------ATTDKATATD 10
SgrLgt  ------------------------------------------------------------
MtbLgt  ETSE---------ADIEREQPGDLAGQAPAAHQVDAEAASAAPEEPAALASEAHDETEPE 124
MsmLgt  EEAEPAETIEPDEDEIFDAELAEALAEAAEDFAVVPESAGSDEAEVAAAESAADAEDGPS 180

ScoLgt1 GAEKTDASGATEAPEDTSGADEADAAKDAEGVTNGADS---------AKKG--------- 52
ScoLgt2 --------------VGLASGPPGN--STP---RRATES-------WNVRNRS-------- 26
ScabLgt --TKSESKAKPGESEAKAEGDAGA--SGGGGEKKEPES---------AKKS--------- 38
SavLgt  TSTTTDKSTDRGKNEDENEGEDAE---PSEKTESAAES---------AKKV--------- 49
SgrLgt  ------AKADAAAKDPLTKDEPGK---DATAENAGAAG--------AAEKA--------- 34
MtbLgt  VPEKAAPIPDPAKPDELAVAGPGDDPAEPDGIRRQDDFSSRRRRWWRLRRRRQ------- 177
MsmLgt  DAESAEVPAEAAADDAVVSEPAGTVDETEVAEEAAAEAADLEFESYDAELAEALAEAAED 240

ScoLgt1 ------------------------------------------------------------
ScoLgt2 ------------------------------------------------------------
ScabLgt ------------------------------------------------------------
SavLgt  ------------------------------------------------------------
SgrLgt  ------------------------------------------------------------
MtbLgt  ------------------------------------------------------------
MsmLgt  MAVVVAPAETAESDEAREPVETDATESEEPSEPEEPDAPEAVEAPEALDENADETRAEPA 300

ScoLgt1 -------------------------
ScoLgt2 -------------------------
ScabLgt -------------------------
SavLgt  -------------------------
SgrLgt  -------------------------
MtbLgt  -------------------------
MsmLgt  APAAVATAPVEPEKGRLRRWLRRNR 325
```

Figure 3.2 Amino acid sequence alignment of the C-termini of the Actinomycete Lgt enzymes. Sequences compared using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

3.3.1.2 Lipoprotein signal peptidase (Lsp).

Of the species tested, the Lsp proteins from the *Streptomyces* species share the highest identity (≥75%). ScoLsp showed relatively weak identity with the Lsp from *Mycobacterium* species (~41%) and even lower to both EcoLsp (36%) and BsuLsp (34%). Currently it is believed that there are five regions of the Lsp protein necessary for protein function (fig. 3.3).

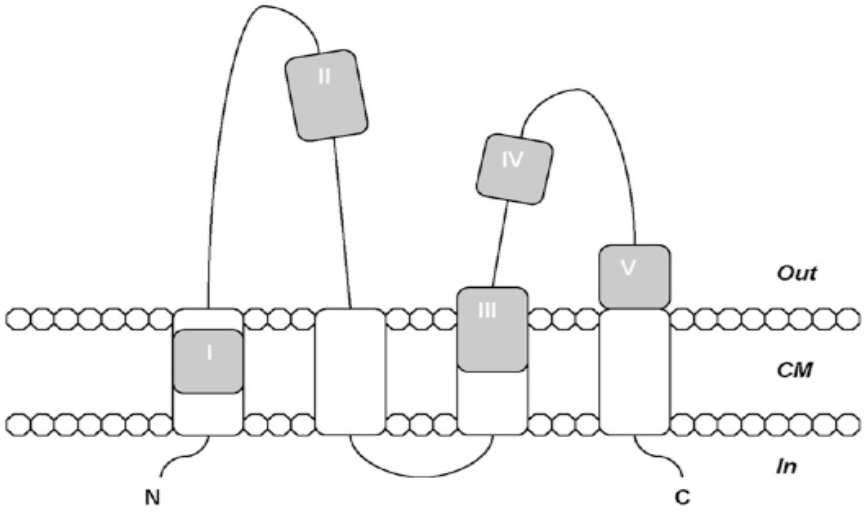


Figure 3.3. The predicted membrane topology of *L. pneumophila* Lsp showing the five essential regions I->V. From [19].

Within these conserved regions, the amino acid residues that are critical for function can be found in domains I, where an aspartic acid (D) is necessary for enzyme stability [26], as are the residues NXXD in region III and FNXAD in region V. It is thought that the aspartic acid residues in III and V may form a catalytic dyad which is the active site of the protease [19, 28]. Whilst the amino acid sequences of regions I->V are inconsistent between species, all the residues needed for stability or catalysis in *B. subtilis* [26] are conserved in all species (fig. 3.4).

```
ScoLsp    MAEAERIIGTPDIPDAAGEGQERPDADPEREQQEQEQAPERTRGKRRVAVLFAVALFAYL 60
ScabLsp   MAEAERIIGTPDIPDATGAGPEQSDDGAGAATAEPADAAARPRGRRRIAVLFGVAALAYA 60
SavLsp    MAEAERIIGTPDIPEAAGAEPEQADGESGGAGAG----TERPKGRRRIAVLFAVAALAYA 56
SgrLsp    MAEAERIIGMPENPDVDGTDEGGSTAADAAVNAG--------RGKRKILALLSVAVVAYL 52
MtbLsp    MPDEP--TG-SADPLTSTEEAGGAGEPNA------------PAPPRRLRMLLSVAVVVLT 45
MsmLsp    MTDET--SG-PAEPVT--DAPGDAESP--------------AQPKRRLRLLLTVAAVVLF 41
EcoLsp    ---------------------------------MSQSICSTG-------LRWLWLVVVVLI 21
BsuLsp    ------------------------------------------------MLYYMIALLIIA 12

ScoLsp    LDLGSKMLVVAKLEHHEPIEIIGDWLRFAAIRNAGAAFGFGEA----FTIIFTVIAAAVI 116
ScabLsp   LDLVSKMIVVARLEHHEPIEIIGEWLKFEAIRNAGAAFGFGEA----FTIIFTVIATIVI 116
SavLsp    FDLVSKLIVVAKLEHHAPIEIIGDWLRFEAIRNAGAAFGFGEA----FTVIFTVIAAAVI 112
SgrLsp    LDLGSKMLVVAKLEHQPPIDIIGDWLQFRAIRNPGAAFGIGEA----FTVIFTIIATGVI 108
MtbLsp    LDIVTKVVAVQLLPPGQPVSIIGDTVTWTLVRNSGAAFSMATG----YTWVLTLIATGVV 101
MsmLsp    LDVVTKVLAVRLLTPGQPVSIIGDTVTWTLVRNSGAAFSMATG----YTWVLTLVATGVV 97
EcoLsp    IDLGSKYLILQNFALGDTVPLFP-SLNLHYARNYGAAFSFLADSGGWQRWFFAGIAIGIS 80
BsuLsp    ADQLTKWLVVKNMELGQSIPIIDQVFYITSHRNTGAAWGILAG----QMWFFYLITTAVI 68

ScoLsp    VVIARLARKLHS--LPWAIALGLLLGGALGNLTDRIFRAPGVFEGAVVDFIAPKHFAVFN 174
ScabLsp   VVIARLARKLYS--LPWAIALGLLLGGALGNLTDRIFRAPGVFEGAVVDFIAPKGFAVFN 174
SavLsp    VVIARLARKLYS--LPWAIALGLLLGGALGNLTDRIFRSPGVFEGAVVDFIAPKHFAVFN 170
SgrLsp    VVIFRIARKLYS--LPWAIALGLLLGGALGNLTDRIFRAPGVFEGAVVDFIAPKNSAVFN 166
MtbLsp    VGIFWMGRRLVS--PWWALGLGMILGGAMGNLVDRFFRAPGPLRGHVVDFLSVGWWPVFN 159
MsmLsp    IGIIWMGRRLVS--PWWALGLGLILGGATGNLVDRFFRSPGPLRGHVVDFFSVGWWPVFN 155
EcoLsp    VILAVMMYRSKATQKLNNIAYALIIGGALGNLFDRLWH--GFVVDMIDFYVGDWHFATFN 138
BsuLsp    IGIVYYIQRYTKGQRLLGVALGLMLGGAIGNFIDRAVR--QEVVDFIHVIIVNYNYPIFN 126

ScoLsp    LADSAIVCGGILIVILSFRGLD----------------------------PDG------- 199
ScabLsp   LADSAIVCGGILIVLLSFRGLD----------------------------PDG------- 199
SavLsp    LADSAIVCGGILIVLLSFRGLD----------------------------PDG------- 195
SgrLsp    LADSAIVCGGILIVILSFKGLD----------------------------PDG------- 191
MtbLsp    VADPSVVGGAILLVILSIFGFDFDTVGRR--------------------HADG------- 192
MsmLsp    VADPSVVGGAILLVALSLFGFDFDTVGRRRPGEDAEPSAGASDSTPEAPAADGPDKPAGP 215
EcoLsp    LADTAICVGAALIVLEGFLPSR---------------------------------------- 160
BsuLsp    IADSSLCVGVMLLFIQMLLDSG---------------------------------------- 148

ScoLsp    -----------TVHKD---- 204
ScabLsp   -----------TVHKG---- 204
SavLsp    -----------TVHKD---- 200
SgrLsp    -----------TVHKD---- 196
MtbLsp    ----------DTVGRRKADG 202
MsmLsp    VGPEDAAEESKTVGHQAEPS 235
EcoLsp    ------------AKKQ---- 164
BsuLsp    -----------KKKKEQ--- 154
```

Figure 3.4. Amino acid sequence alignment of the Lsp enzymes. Potential transmembrane domains are underlined, highlighted in yellow are regions believed to be important in *B. subtilis* Lsp. Amino Acids in blue are to be necessary for enzyme stability in *B. subtilis* whilst those in red are necessary for catalysis. Sequences compared using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

As with the Lgt enzymes, there is a marked protein elongation between the Lsp enzymes from the actinomycetes and those from *E. coli* and *B. subtilis*. This time however the elongated sections are found in the N-termini. Like the C-terminal tails of Lgt these extensions are predicted to be cytoplasmic, however

unlike the Lgt tails they show a higher level of similarity. Between the *Streptomyces* species the similarity ranges from 45-68%, whilst between the *Mycobacterium* species the similarity is 35%. However, there is low similarity between the two groups. There are two predicted amino acids that are conserved between the species. (fig. 3.5).

```
ScoLsp  MAEAERIIGTPDIPDAAGEGQERPDADPEREQQEQEQAP 39
ScabLsp MAEAERIIGTPDIPDATGAGPEQSDDGAGAATAEPADAA 39
SavLsp  MAEAERIIGTPDIPEAAGAEPEQADGESGGAGAGT---- 35
SgrLsp  MAEAERIIGMPENPDVDGTD----EGGSTAADAAV---- 31
MtbLsp  MPDEP--TG-SADPLTS------TEEAGGAGEP------ 24
MsmLsp  MTDET--SG-PAEPVT---------DAPGDAE------- 20
```

Figure 3.5. Amino acid sequence alignment of the N-terminal extensions of the Lsp enzymes from the actinomycetes tested. Predicted conserved residues are highlighted in red.

3.3.1.3 Lipoprotein N-acyl transferase (Lnt).

As has been discussed, *S. coelicolor* is unusual as it contains two potential copies of the *lnt* gene. The ScoLnt1 protein shares a high degree of identity with ScabLnt1 (86%) and a lower level with the *Mycobacterium* Lnt enzymes (~42%). It only shares 29% identity with ScoLnt2, which is itself more similar to the remaining *Streptomyces* Lnt enzymes. There is no BsuLnt homologue, as discussed in Chapter 1.

Currently, the only experimentally confirmed Gram-positive Lnt enzymes are found in *M. smegmatis*, and *M. tuberculosis* [38]. MsmLnt, ScoLnt1 and ScoLnt2 all conserve the three residues (E267, K355, C387) which form the enzyme's catalytic triad in *E. coli* [30], as do all the other species tested (fig. 3.6). However, of the four remaining essential amino acids found in *E. coli* Lnt, only E343 and E389 are conserved between all species. In both MsmLnt and

ScoLnt1, a valine is present in place of *E.* coli W237 whilst in ScoLnt2 a serine is present. Y388 is conserved between EcoLnt and ScoLnt1, with a phenylalanine located here in ScoLnt2 or a tryptophan in MsmLnt. As can be seen in figure 3.6, MtbLnt has an extended C-terminus. As described in chapter 1, this is as a result of a fusion between *lnt* and the *ppm1* genes [37], which has lead to a multi domain protein.

```
ScoLnt1  ------------------------------------------------------------MTA 3
ScoLnt2  ------------------------------------------------------------
ScabLnt1 ------------------------------------------------------------MTV 3
ScabLnt2 ------------------------------------------------------------
SavLnt   ------------------------------------------------------------
SgrLnt   -----------------------------MRIRGARLPARRHAAGTAAAGGSGGRAGDD 30
MtbLnt   ------------------------------------------------------------
MsmLnt   MADDRARRFDRFRVRPEEITEVIPAVTDDDPLEDPLDDDVAPGLDDAEPEPEPRDEHDEP 60
EcoLnt   ------------------------------------------------------------

ScoLnt1  TATTVGEPDRTQPQTTPASRAVSRLARLLPAAAAALSGVLLYISFPPRTLWWLALPAFAV 63
ScoLnt2  ----------------MKTLDRRLASPWRRSGLAVLAGALPVLAFPGPALWWWAW--FAL 42
ScabLnt1 TATPVDEPEQLEPQAAPVSRVSRWAARLLPAAAAALSGVLLYVSFPPRTLWWLALPAFAV 63
ScabLnt2 ----------------MRTP-DWIASPWRRRSAVVVAGALPVLAFPAPGLWWSAY--GAL 41
SavLnt   ----------------MRMPEDWLGSRWWRGAAAVLAGALPMLAFPAPSWWWFAY--VAL 42
SgrLnt   ASGDGPPDDGPGPRGRGSAREGALSSAWARGAAALLAGALPALAFPAPGLWWFAY--VAL 88
MtbLnt   ------MKLGAWVAAQLPTTRTAVRTRLTRLVVSIVAGLLLYASFPPRNCWWAAVVALAL 54
MsmLnt   SRPATGSRIGGWVARRGSRFGKGVLDRCAPLSAAIGGGLALWLSFPPIGWWFTAFPGLAL 120
EcoLnt   -----------------MAFASLIERQRIRLLLALLFGACGTLAFSPYDVWPAAIISLMG 43

ScoLnt1  FGWVLRGR----GWK--AGLGLG-YLFGLGFLLPLLVWTGVEVGPGPWLALAAIEALFVA 116
ScoLnt2  VPWILLAR----TAPGGKRAAYDGWCGGFGFVLAMHHWLLPNLHVFTFVIAALLGALWVP 98
ScabLnt1 LGWVLRGR----GWK--AGLGLG-YLFGLGFLLPLLVWTGVEVGPGPWLALAVIEAVFVA 116
ScabLnt2 VPWILLLR----AAPTGRRAAYDGWLGGFAFMLAMHHWLLPNLHVFTFLIAGLLGALWAP 97
SavLnt   VPWILLAR----TAPTGGRAAYDGWLGGLGFMLAVHHWLLPSLHVFTVLIAALLGALWAP 98
SgrLnt   VPWLLLIR----GARSPRRAALDGWIGGIGFVVAVHHWLMPSLHVFIVLLAALLGLLWAP 144
MtbLnt   LAWVLTHR----ATTPVGGLGYG-LLFGLVFYVSLLPWIGELVGPGPWLALATTCALFPG 109
MsmLnt   LGWVLTRT----ATTKAGGFGYG-VLFGLAFYVPLLPWISGLVGAVPWLALAFAESLFCG 175
EcoLnt   LQALTFNRRPLQSAAIGFCWGFGLFGSGINWVYVSIATFGGMPGPVNIFLVVLLAAYLSL 103

ScoLnt1  AVGAGVAAVSKLPGSP------VWAAAVWIAGEAARARAPFEGFPWGKIAFG-QADGVFL 169
ScoLnt2  WGWLVHRTLGGTPSSRRVAAALVVLPSGWLLAELVRSWQGLGG-PWGMLGASQWQVAPAL 157
ScabLnt1 LVGAGVAVVSKLPGWP------VWAAALWVAGEAARARAPFHGFPWGKIAFG-QADGVFL 169
ScabLnt2 WAWLAHRFLAGSPSSGRVAAALLVVPSGWLMIELVRSWQGLGG-PWGLLGSSQWEVEAAL 156
SavLnt   WGWLVRRFLAGVPSPGRVAAAMLVLPSGWLMVELVRSWQGLGG-PWGLLGSSQWQVEPAL 157
SgrLnt   WGLLVARLLGGSPSAGRAVAAVVVVPSGWLMIELVRSWEGLGG-PWGLLGASQWDVAPAL 203
MtbLnt   IFGLFAVVVRLLPGWP------IWFAVGWAAQEWLKSILPFGGFPWGSVAFG-QAEGPLL 162
MsmLnt   LFGLGAVVVVRLPGWP------LWFATLWVAAEWAKSTFPFGGFPWGASSYG-QTNGPLL 228
EcoLnt   YTGLFAGVLSRLWPKTTWLRVAIAAPALWQVTEFLRGWV-LTGFPWLQFGYS-QIDGPLK 161

ScoLnt1  PLAAVGGTPVLGFAVVLCGFALHEAVRLAVRARR-GDEVRRGAAAVALLGVAVPVVGAVA 228
ScoLnt2  RLASVGGVWLLSFLVVAVNAALAVLVAVRRARVP-------ALAGLVATAAATSAAWVWS 210
ScabLnt1 PLAALGGTPVLGFAVVLCGFGLCEAVRLLVANRRTGEAVRRGTAAVAALSVAVPLAGAFA 229
ScabLnt2 RLASVGGVWLLSFLVVAVNVAVTVLVSVRASRAP-------ALAGLVVVAVATSAVWLWS 209
SavLnt   RLASVGGVWLLSFLVVTVNIAVAVLVSVRASRMP-------ALAGLVATAAATSAAWAWS 210
SgrLnt   RVASVGGVWLVSLLVLAVNTGVALLVADRAARRT-------GGVLLAVCALSVAAVWAWA 256
MtbLnt   PLVQLGGVALLSTGVALVGCGLTAIALEIEKWWRTGGQ-GDAPPAVVLPAACICLVLFAA 221
MsmLnt   ALARIGGAPLVSFAVALIGFSLTLLTAQIVWWWRHGHKPGVPAPAVMLPGVAIAASLLVT 288
EcoLnt   GLAPIMGVEAINFLLMMVS-GLLALALVKRNWRP-------LVVAVVLFALPFPLRYIQW 213
```

```
ScoLnt1  A------RPLVSDTAEDGTATVAVIQGNVPRAGLGFNAQRRAVLDYHARETQRLADEVKA 282
ScoLnt2  P-----------RPDTDERAAIAVVQPGV--------VAGADSADRRFDREEQLTRRLAD 251
ScabLnt1 S------RALVSDTAEDGTATVAVIQGNVPRLGLDFNAQRRAVLDYHAKETERLAAQVKA 283
ScabLnt2 P-----------RPDVRGEVRIAVVQPGV--------TDGP---DARFAREEALTRRLAG 247
SavLnt   P-----------RPDHDGDMRIAIVQPGV--------IDGTGSADRRFDREEQLTRELAG 251
SgrLnt   P-----------RPTETGAVRIAVVQPGV--------VEGPGSVARRFDRGEELTRALRG 297
MtbLnt   IVVWPQVRHAGSGSGGEPTVTVAVVQGNVPRLGLDFNAQRRAVLDNHVEETLRLAADVHA 281
MsmLnt   ALVWPQVRQSGTGAGDDTAVTVAAVQGNVPRLGLEFNAQRRAVLDNHVKETLRLADDVKA 348
EcoLnt   F-----------TPQPEKTIQVSMVQGDIP--------QSLKWDEGQLLNTLKIYYNATA 254

ScoLnt1  GKVARPDFVLWPENSSDIDPFANADARLVIDRAAKAVGAPISVGGVVERDGK------LL 336
ScoLnt2  ---RDLDLIVWGESSVGFDLDDRPDLARRLAALSRETGADILVN-VDARRSD---KPGIY 304
ScabLnt1 GKVARPDFVLWPENSSDIDPFANADARAVIDRAATAIGAPISVGGVVERDGK------LY 337
ScabLnt2 ---QDVDLIVWGESSVGHDLADRPDLSDRIAALARAADSDILVN-VDARRSD---RPGIY 300
SavLnt   ---QNVDLVVWGESSVGFDLADRPDLARRIAGLSERVGADILVN-VDARRSD---RPGIY 304
SgrLnt   ---RGVDLVVWGESSIGAGAWERPETARRLAGLSRLVGADLLVN-VDARQTDGSGRSGIF 353
MtbLnt   GLAQQPQFVIWPENSSDIDPFVNPDAGQRISAAAEAIGAPILIGTLMDVPGRPRENPEWT 341
MsmLnt   GRAAQPMFVIWPENSSDIDPLLNADASAQITTAAEAIDAPILVGGVVRADGYTPDNPVAN 408
EcoLnt   PLMGKSSLIIWPESAITDLEINQQPFLKALDGELRDKGSSLVTGIVDARLNK-QNRYDTY 313

ScoLnt1  NEQILWDPDKG----PVDTYDKRQIQPFGEYLPLRSLIGAINDEWTSMVSRDFSRGTEPG 392
ScoLnt2  KSSVLVGPQGP----TGDRYDKMRLVPFGEYVPFRSLLGWATSVGKAAG-EDRRQGTEQV 359
ScabLnt1 NEQILWDPAKG----PVDTYDKRQIQPFGEYLPLRSLIGAINGEWTSMVRKDFSRGTEPG 393
ScabLnt2 KSSVLVGPDGP----TGDRYDKMRLVPFGEYIPMRSLLGWATSVGEAAG-EDRRRGTEQV 355
SavLnt   KSSVLVGPHGP----TGDRYDKMRLVPFGEYIPARSLLGWATSVGKAAG-EDRRRGSEQV 359
SgrLnt   KSAVLVGPDGP----TGDRYDKMRLVPFGEYVPARSLLGWATSVGKAAG-EDRLRGDRQV 408
MtbLnt   NTAIVWNPGTG----PADRHDKAIVQPFGEYLPMPWLFRHLSGYADRAG--HFVPGNGTG 395
MsmLnt   NTVIVWEPTDG----PGERHDKQIVQPFGEYLPWRGFFKHLSSYADRAG--YFVPGTGTG 462
EcoLnt   NTIITLGKGAPYSYESADRYNKNHLVPFGEFVPLESILRPLAPFFDLPMSSFSRGPYIQP 373

ScoLnt1  VFT----MAGTKVGLVTCYEAAFDWAVRSEVTDGAQMISVPSNNATFDRSEMTYQQLAMS 448
ScoLnt2  VM---DVGDGLRIGPMVCFESAFPDMSRSLVADGAQVLVAQSSTSTFQHTWAPEQHASLA 416
ScabLnt1 VFT----MHGAKIGLVTCYEAAFDWAVRSEVTDGAQLISVPSNNATFDRSEMTYQQLAMS 449
ScabLnt2 VF---DAGKGLRIGPMVCFESAFPDMSRQLALDGAELLLAQSATSSFQQSWAPEQHATLA 412
SavLnt   VM---NAGHGLRIGPMVCFETAFPDMSRHLAEDGAEVLLAQSSTSTFQQSWAPEQHASLA 416
SgrLnt   VMTLPDGARGLRIGPLVCFETAFPDMSRRLVRDGAQVIVAQSATSTFQHSWAPAQHASLG 468
MtbLnt   VVR----IAGVPVGVATCWEVIFDRAPRKSILGGAQLLTVPSNNATFNKT-MSEQQLAFA 450
MsmLnt   VVH----AAGVPIGITTCWEVIFDRAARESVLNGAQVLAVPSNNATFDEA-MSAQQLAFG 517
EcoLnt   PLS----ANGIELTAAICYEIILGEQVRDNFRPDTDYLLTISNDAWFGKSIGPWQHFQMA 429

ScoLnt1  RIRAVEHSRTVTVPVTSGVSAIIMPDG-RITQKTGMFVADSLVQEVPLRSSETPATRLGI 507
ScoLnt2  ALRAAETGRPMVHATLTGVSAVYDANGARIGSWLGTDASASRVYEVPVTHGTTPYVRYGD 476
ScabLnt1 RVRAVEHSRTVTVPVTSGVSAIIMPDG-RIAQKTGMFVPDSLVQKVPLRSSQTPATRYGI 508
ScabLnt2 ALRAAETGRPMVHATLTGVSAVYGPSGERVGPWLGTDASEAAVYRVPMAGGTTPYVRFGE 472
SavLnt   ALRAAETGRPMVHATLTGVSAVYGPSGERVGSWLGTGKSTSAVYDVPLARGVTPYVRFGD 476
SgrLnt   ALRAAENGRPMVHATLTGISAAYGPRGERVGRPLGTDASAAEVFDLPLAGGSTLYNRLGD 528
MtbLnt   KVRAVEHDRYVVVAGTTGISAVIAPDG-GELIRTDFFQPAYLDSQVRLKTRLTPATRWGP 509
MsmLnt   KLRAVEHDRYVVVAGTTGISAVIAPDG-HEISRTEWFQPAYLDNQIRLKTDLTPATKWGP 576
EcoLnt   RMRALELARPLLRSTNNGITAVIGPQG-EIQAMIPQFTREVLTTNVTPTTGLTPYARTGN 488

ScoLnt1  APEIALVLVAAGGLGWAVGAG---------------------------------------- 528
ScoLnt2  WTVYAALGILAAWGAAEGVRTVRLRRNRPARPGPP---------------------------- 511
ScabLnt1 APEMLLVLVAAGGLGWAIGAG---------------------------------------- 529
ScabLnt2 WPVQAALLVLVAWGAVEGVRALRLRRQAG-PRPPA------------------------------ 506
SavLnt   WPVHAALLILAALCTAEGVREFRLRRTVPSPRVPP---------------------------- 511
SgrLnt   WPVYGALAALAVLCAVEGLR--ALRRPAPGPPGPP---------------------------- 561
MtbLnt   ILQWILVGAAAAVVLVAMRQNGWFPRPRRSEPKGENDDSDAPPGRSEASGPPALSESDDE 569
MsmLnt   IVQAVLVIAGVAVLLIAILHNGRFAP---------------------------------- 602
EcoLnt   WPLWVLTALFGFAAVLMSLRQRRK------------------------------------- 512

ScoLnt1  ------------------VRGRRARDV---------------------------------- 537
ScoLnt2  ------------------ARTAH-GSPARPGH------------------------------ 524
ScabLnt1 ------------------VRGRRAGGV---------------------------------- 538
ScabLnt2 ------------------AQSADPSSTARTTRSHSS------------------------- 524
SavLnt   ------------------ARTVR-ESPARPGR------------------------------ 524
SgrLnt   ------------------ARTAH-GSPGSPGH------------------------------ 574
MtbLnt   LIQPEQGGRHSSGFGRHRATSRSYMTTGQPAPPAPGNRPSQRVLVIIPTFNERENLPVIH 629
MsmLnt   -----------------RMLRRRSATTVKR------------------------------ 615
EcoLnt   ------------------------------------------------------------
```

```
ScoLnt1  ------------------------------------------------------------
ScoLnt2  ------------------------------------------------------------
ScabLnt1 ------------------------------------------------------------
ScabLnt2 ------------------------------------------------------------
SavLnt   ------------------------------------------------------------
SgrLnt   ------------------------------------------------------------
MtbLnt   RRLTQACPAVHVLVVDDSSPDGTGQLADELAQADPGRTHVMHRTAKNGLGAAYLAGFAWG 689
MsmLnt   ------------------------------------------------------------
EcoLnt   ------------------------------------------------------------

ScoLnt1  ------------------------------------------------------------
ScoLnt2  ------------------------------------------------------------
ScabLnt1 ------------------------------------------------------------
ScabLnt2 ------------------------------------------------------------
SavLnt   ------------------------------------------------------------
SgrLnt   ------------------------------------------------------------
MtbLnt   LSREYSVLVEMDADGSHAPEQLQRLLDAVDAGADLAIGSRYVAGGTVRNWPWRRLVLSKT 749
MsmLnt   ------------------------------------------------------------
EcoLnt   ------------------------------------------------------------

ScoLnt1  ------------------------------------------------------------
ScoLnt2  ------------------------------------------------------------
ScabLnt1 ------------------------------------------------------------
ScabLnt2 ------------------------------------------------------------
SavLnt   ------------------------------------------------------------
SgrLnt   ------------------------------------------------------------
MtbLnt   ANTYSRLALGIGIHDITAGYRAYRREALEAIDLDGVDSKGYCFQIDLTWRTVSNGFVVTE 809
MsmLnt   ------------------------------------------------------------
EcoLnt   ------------------------------------------------------------

ScoLnt1  ------------------------------------------------------------
ScoLnt2  ------------------------------------------------------------
ScabLnt1 ------------------------------------------------------------
ScabLnt2 ------------------------------------------------------------
SavLnt   ------------------------------------------------------------
SgrLnt   ------------------------------------------------------------
MtbLnt   VPITFTERELGVSKMSGSNIREALVKVARWGIEGRLSRSDHARARPDIARPGAGGSRVSR 869
MsmLnt   ------------------------------------------------------------
EcoLnt   ------------------------------------------------------------

ScoLnt1  -----
ScoLnt2  -----
ScabLnt1 -----
ScabLnt2 -----
SavLnt   -----
SgrLnt   -----
MtbLnt   ADVTE 874
MsmLnt   -----
EcoLnt   -----
```

Figure 3.6. Amino acid sequence alignment of the Lnt enzymes. Potential transmembrane domains are underlined. Amino Acids in blue are to be necessary for enzyme stability in *E. coli* whilst those in red are necessary for catalysis. Sequences compared using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

3.3.2 Bioinformatic analysis of lipoproteins in *S. coelicolor.*

3.3.2.1 Identification of Lipoproteins.

Analysis of the *S. coelicolor* genome sequence by Professor Iain Sutcliffe (University of Northumbria, Newcastle), Dr Matt Hutchings (UEA) and Dr. Govind Chandra (John Innes Centre, Norwich) led to the identification of up to 223 putative lipoproteins (appendix 1). All *S. coelicolor* proteins (http://strepdb.streptomyces.org.uk/) were assessed, and only those with a cysteine within the first 50 amino acids (representing a potential lipobox C_{+1}, characteristic of a lipoprotein) were accepted for further analysis. These proteins were matched against the G+LPP pattern, used for the identification of Gram positive lipoprotein sequences [64], and the revision G+LPPv2 [11, 85]. These patterns are used to filter out any proteins that do not contain a potential lipobox (see chapter 1). In tandem with this, potential lipoproteins were also tested against the prosite pattern PS51257 (http://www.expasy.ch/tools/scanprosite/). Those which passed both tests were submitted to the following tests: SignalP v3.0 [86], Phobius [87], LipoP [88] and Predlipo [89]. The combined use of these tests is useful for accurate identification of lipoproteins. In addition lipoproteins secreted by the Tat pathway (Chapter 1), were identified using TatFind [90] and TatP [91].

3.3.2.2 Functional analysis of lipoproteins.

The 223 potential lipoproteins identified in the *S. coelicolor* genome using the methods described above represent ~2.7% of the *S. coelicolor* proteome. Surprisingly, fifty of the putative lipoproteins passed one or other of the Tat tests mentioned above, with 38 passing both, which would represent 17% of the total *S. coelicolor* lipoproteins. Even when taking into account the genome size of *S.*

coelicolor, this is a large amount. *M. tuberculosis* is estimated to have 10-15% Tat secreted lipoproteins [11]. In comparison, *E.coli* is predicted to have 28 Tat secreted proteins in total, and *B. subtilis* is predicted to only secrete ≤7 proteins via Tat [2].

Of the putative *S. coelicolor* lipoproteins, it is unsurprising that 35% are of unknown function and several have no homologues outside of the genus *Streptomyces*. ABC transporters make up a further 41% of the lipoproteome. Of these, more than half are needed for carbohydrate transport, and given the complex nature of the soil, and the saprophytic nature of *S. coelicolor* cannot be surprising.

Function	**Numbers**	**%**
Solute binding proteins	92	41%
Putative enzymes	34	15%
Redox processes	6	3%
Signal transduction ('three component' systems)	6	3%
Cell envelope processes	7	3%
Function unknown	78	35%

Table 3.1. Overview of lipoprotein functions in *S. coelicolor.* Number of proteins and % of total lipoproteins are shown.

3.3.3 Disrupting the *S. coelicolor* biosynthetic genes.

Cosmids containing the *S. coelicolor* genes *lgt1* (*Sco2034*), *lgt2* (*Sco7822*) and *lsp* (*Sco2074*), were obtained from the *Streptomyces* group at the John Innes Centre. Genes were identified, by homology searches against the proteomes of *E. coli*, *B. subtilis* and *M. tuberculosis* (see section 3.3.1 above). The cosmids were checked by restriction digest, to ensure their fidelity. Digestion patterns were obtained from http://streptomyces.org.uk/.

Apramycin marked cassettes for each of the genes were generated using 'KO' primers specific to each gene (table 2.4) and introduced into their respective cosmids as explained in chapter 2. The disrupted cosmids were checked by PCR, and the correct position of the antibiotic disruption cassette was ensured by using several combinations of primers. Firstly, short test primers annealing to flanking regions of the gene were used to confirm that there was a size difference between the wild type and the disrupted gene, with a band at ~1400 bp representing the *apr* cassette. Secondly primers specific to the *apr* cassette were used to show that it did indeed exist somewhere in the cosmid, and finally combinations of the two were used to show that it was in the right place. PCR primers are listed in table 2.4 and results for each of the genes can be seen in figures 3.7 and 3.8.

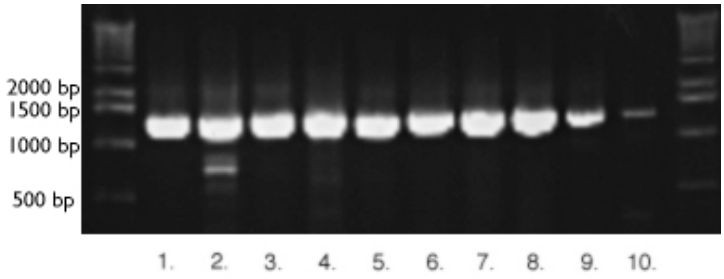


Figure 3.7. Confirmation of the disrupted *lgt1::apr* (Lanes 1. -> 4.) and *lgt2::apr* (Lanes 4->8) in cosmids 4G6 and 8E7 respectively. Primer combinations in each lane (expected band sizes shown in brackets): 1. = P1/P2 (1423 bp), 2. = P1/Sco Lgt1 Test Rev (1424 bp), 3. = P2/Sco Lgt1 Test For (1423 bp), 4. = Sco Lgt1 Test For/Sco Lgt1 Test Rev (1424 bp), 5. = P1/P2 (1423 bp), 6 = P1/Sco Lgt2 Test Rev (1424 bp), 7. = P2/Sco Lgt1 Test Rev (1423 bp), 8. = Sco Lgt2 Test For/Sco Lgt2 Test Rev (1423 bp). Lanes 9 and 10 are positive controls of the *lgt1::apr* and *lgt2::apr* PCR products using the P1 + P2 primers (1423 bp).

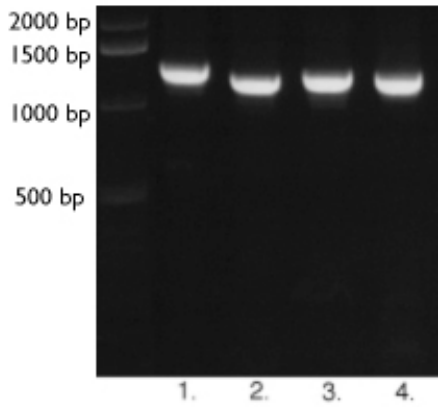


Figure 3.8. Confirmation of the disrupted *lsp::apr* in cosmid 4A10. Primer combinations in each lane (expected band sizes shown in brackets): 1. = Sco LspTest For/Sco Lsp Test Rev (1424 bp), 2. = P1/P2 (1423 bp), 3. = P2/Sco Lsp Test For (1423 bp), 4. = P1/Sco Lsp Test Rev (1424 bp).

When the mutagenised cosmids were transferred into *S. coelicolor* by conjugation to give in-frame gene disruptions (section 2.3.5), genomic DNA was isolated from the exconjugants selected on apramycin and checked by PCR as before (Figs. 3.9 and 3.18).

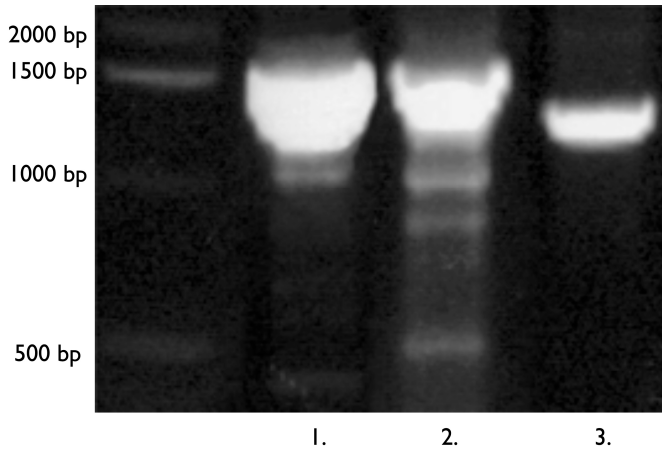


Figure 3.9. PCR of S. coelicolor genomic DNA, to confirm apramycin disruption of *lgt1*. Primer combinations in each lane (expected band sizes shown in brackets): 1. = Sco Lgt1 Comp For/Rev (1736 bp), 2. = P1/Sco Lgt1 Comp Rev (1611 bp), 3. = P1/P2 (1423 bp).

A marker-less deletion of *lsp* was also made, using the FLP recombinase method described in Chapter 2. Removing the *apr* cassette, left a signature 81bp 'scar', which was detectable by PCR, as can be seen in figure 3.10. This strain was used throughout the experiments described below.

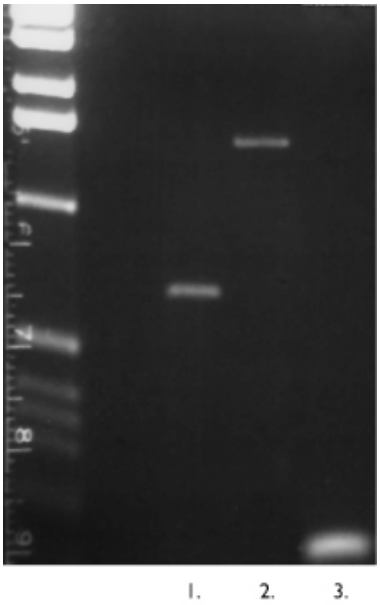


Figure 3.10. PCR of *S. coelicolor* genomic DNA using the Sco LspTest For/Sco Lsp Test Rev primers (expected band sizes shown in brackets). Lane 1. = wild-type M145 (654 bp). Lane 2. = *Δlsp::apr* (1424 bp). Lane 3. = *Δlsp*FLP (81 bp).

Initial attempts to isolate a deletion of *lgt2* proved fruitless. The mutation was successfully made in the cosmid 8E7 but could not be conjugated into *S. coelicolor* M145, probably due to the presence of three transposons in the cosmid carrying the *lgt2* gene. Instead, the *lgt2* gene deletion was carried out by Dr David Widdick (UEA). A suicide vector containing the *apr* disrupted *lgt2* gene was used to make the mutant *S. coelicolor* strain, which was thoroughly tested to ensure it was correct [82]. A hygromycin cassette was also used to make an *lgt2::hyg* disruption strain. Creation of the *S. coelicolor lgt* double mutant using *lgt1::apr* and *lgt2::hyg* proved impossible despite repeated attempts. The deletion of *lsp* also proved difficult, and resulted in numerous

growth phenotypes that are described below.

3.3.3.1 Development and colony morphology.

Of all the null mutants obtained, only the *Δlsp* strain (BJT1001) showed a marked phenotype. This strain is slower growing, and sporulates poorly. This strain frequently overproduces the blue-pigmented antibiotic actinorhodin (act) when grown on a variety of growth media, which is often a sign of stress [92]. The macroscopic phenotype of the *lgt1::apr* and *lgt2::apr* strains appear similar to the wild-type.

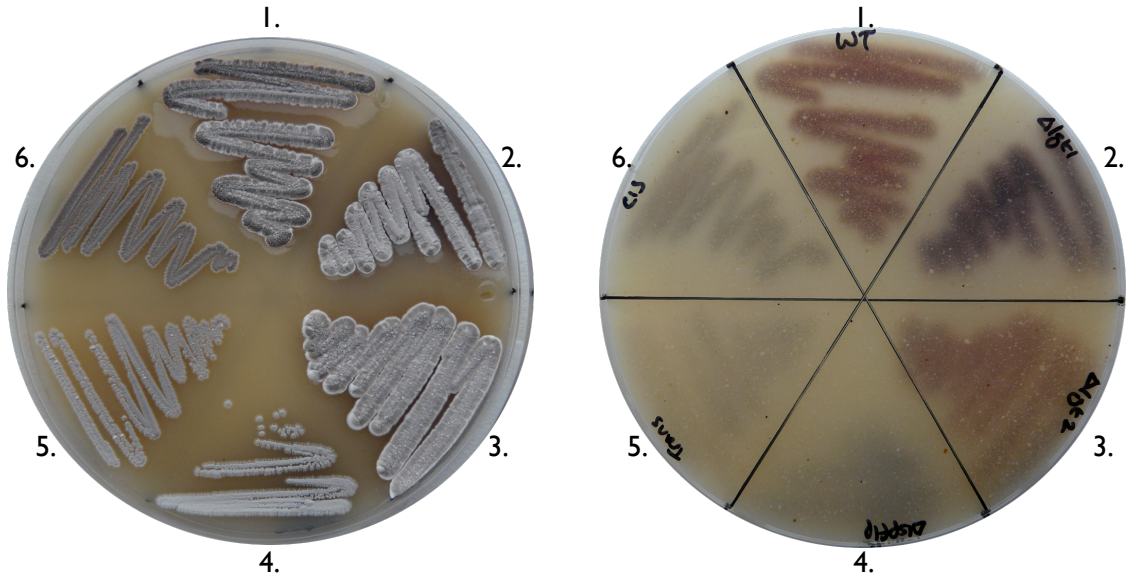
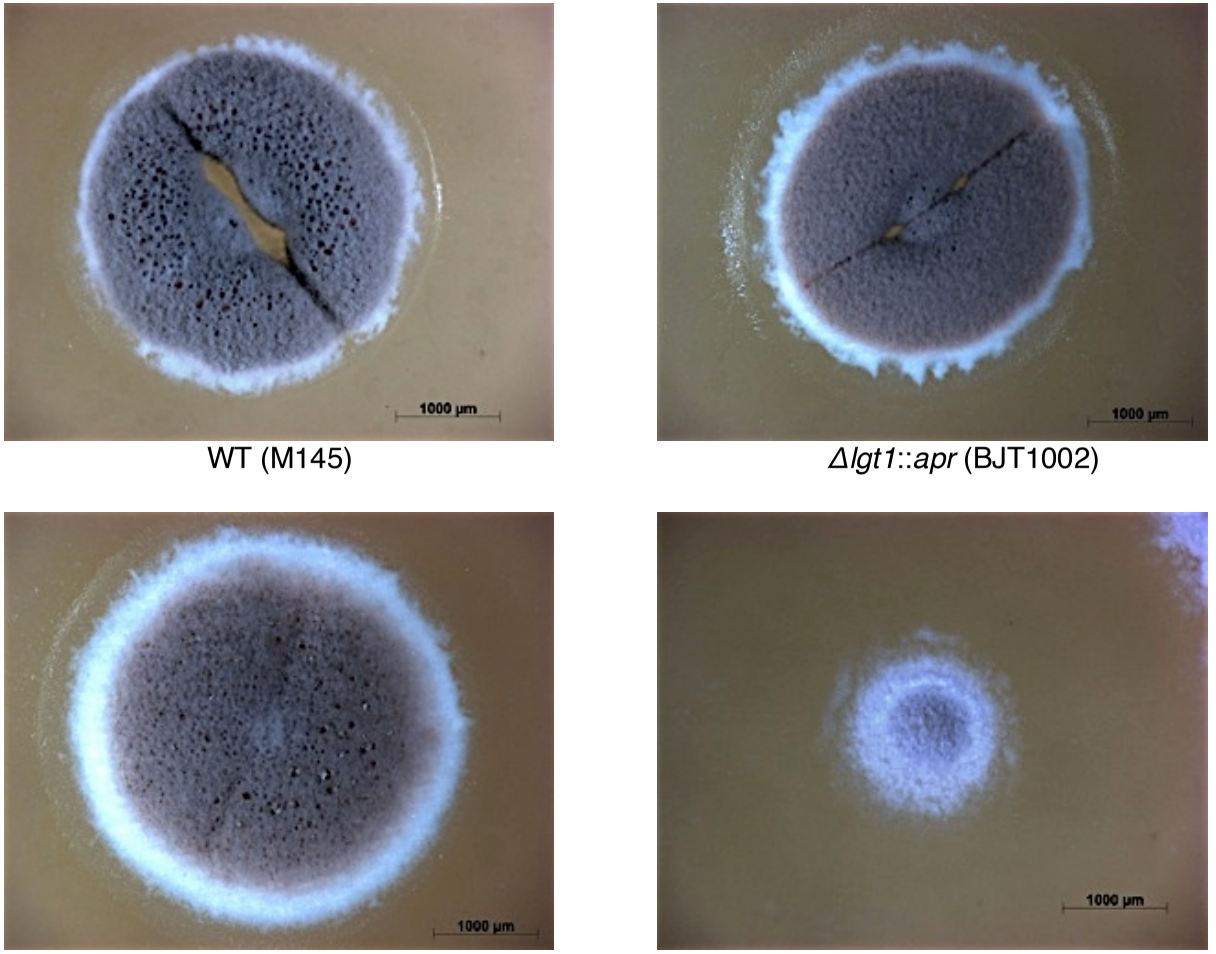


Figure 3.11. Unmagnified phenotypes of each of the *S. coelicolor* strains grown on SFM media. The sections are as follows: 1. = WT (M145), 2. = M145 *Δlgt1::apr* (BJT1002), 3. = M145 *Δlgt2::apr* (BJT1027), 4. = M145 *Δlsp*Flp (BJT1001), 5. = M145 *Δlsp* + *lsp in trans* complementation (BJT1007), 6. = M145 *Δlsp* + *lsp* cis complementation (BJT1006). The growth retardation of the *Δlsp* strain is clearly visible, whilst the complemented strains look more closer to wild-type. Neither *Δlgt1*, nor *Δlgt2* display any obvious phenotype.

All strains were visualised by both light and scanning electron microscopy. Both of these techniques confirmed that the *Δlsp* mutant (BJT1001) had a drastic phenotype in comparison to the other strains. Individual colonies viewed under the light microscope can be seen in figure 3.12. After 5 nights growth on SFM growth media (table 2.3), the wild-type, *Δlgt1* (BJT1002) and *Δlgt2* (BJT1027) mutants were typically between 3 and 4 mm in diameter, whilst the diameter of the *Δlsp* mutant was between 1 and 2 mm. The *Δlsp* mutant also lacks depth when compared to the other strains, appearing very flat against the growth media. As this mutant sporulates very poorly it is unsurprising to see a larger proportion of white areas around the colony, representing undifferentiated substrate hyphae, than grey areas, representing mature spore chains. The bisections shown in the WT and *Δlgt1* colonies below are unexplained, and appears to occur at random in these strains as well as in the *Δlgt2* mutant. A colony divided into four was also seen regularly. These crenellations were not seen in the *Δlsp* mutant and could potentially be caused by the colony collapsing in upon itself, due to its increased height and mass.



WT (M145)

Δlgt1::*apr* (BJT1002)

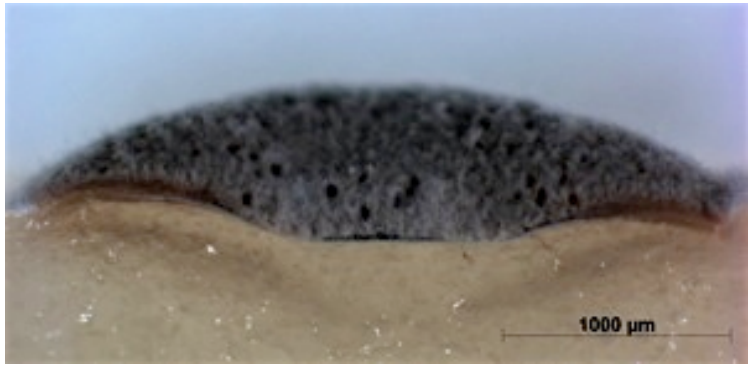
Δlgt2::*apr* (BJT1027)

Δ*lsp*Flp (BJT1000)

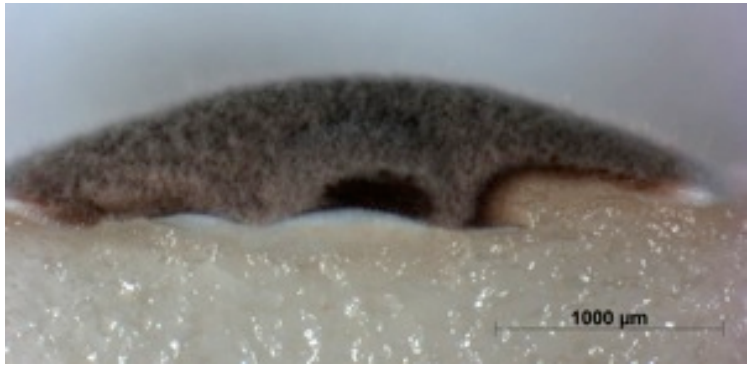
Figure 3.12. Light microscopy of each of the colonies tested, the *Δlsp* mutant is markedly smaller than the other mutants, which show no obvious difference to the wild-type. The cause of the crenellations is unknown. Strain names are shown in brackets.

Colonies dissected bilaterally can be seen in figure 3.13. As with the top-down images, the difference between the *Δlsp* mutant and the other strains is conspicuous. These results again highlight the lack of depth in this strain, and its overall reduced size. Also apparent, is the reduction in the layer of undecylprodigiosin, the red pigmented antibiotic, in the *Δlsp* mutant. Whether or not this is down to the delayed development in the strain is unknown.

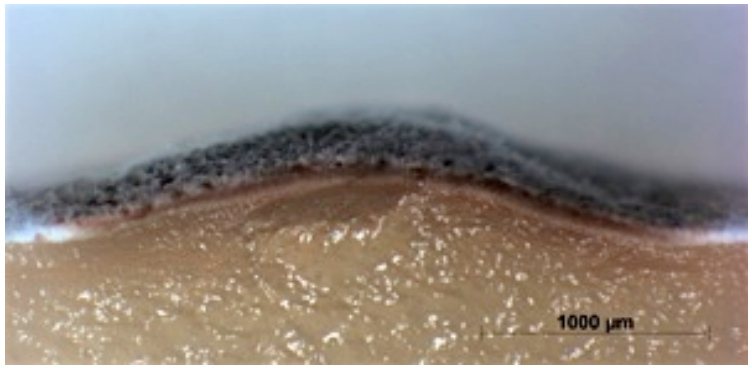
Viewing colonies grown for 5 nights under SEM (fig. 3.14) shows some differences between the samples. As with the light microscopy results above, the wild-type sample appears very similar to both the *Δlgt1* and *Δlgt2* mutants. At first glance, this appears to be the case for the *Δlsp* mutant as well. However, on closer inspection there are very few spiral spore chains in comparison to the other strains. The overall lack of these structures would explain the reduction of sporulation in the *Δlsp* mutant, in comparison to the wild-type. The spore chains in *S. coelicolor* begin as undifferentiated aerial hyphae (see chapter 1), which grow vertically in comparison to the substrate mycelium. The lack of this vertical growth may also be a contributing factor to the overall lack of depth in the *Δlsp* mutant.



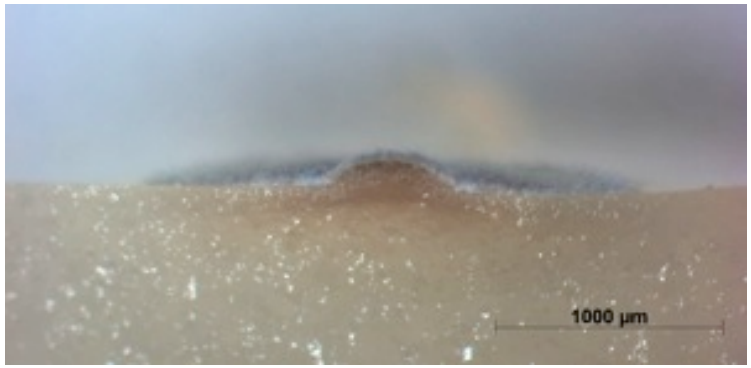
WT (M145)



Δlgt1::*apr* (BJT1002)



Δlgt2::*apr* (BJT1027)



Δ*lsp*Flp (BJT1000)

Figure 3.13. A cross section of the colonies from fig. 3.12, viewed under a light microscope. The thin layer of red undecylprodigiosin can be seen between the grey spores and the growth media in the wild-type, *Δlgt1* and *Δlgt2* mutants. It is reduced in the *Δlsp*FLP strain. Strain names are shown in brackets.

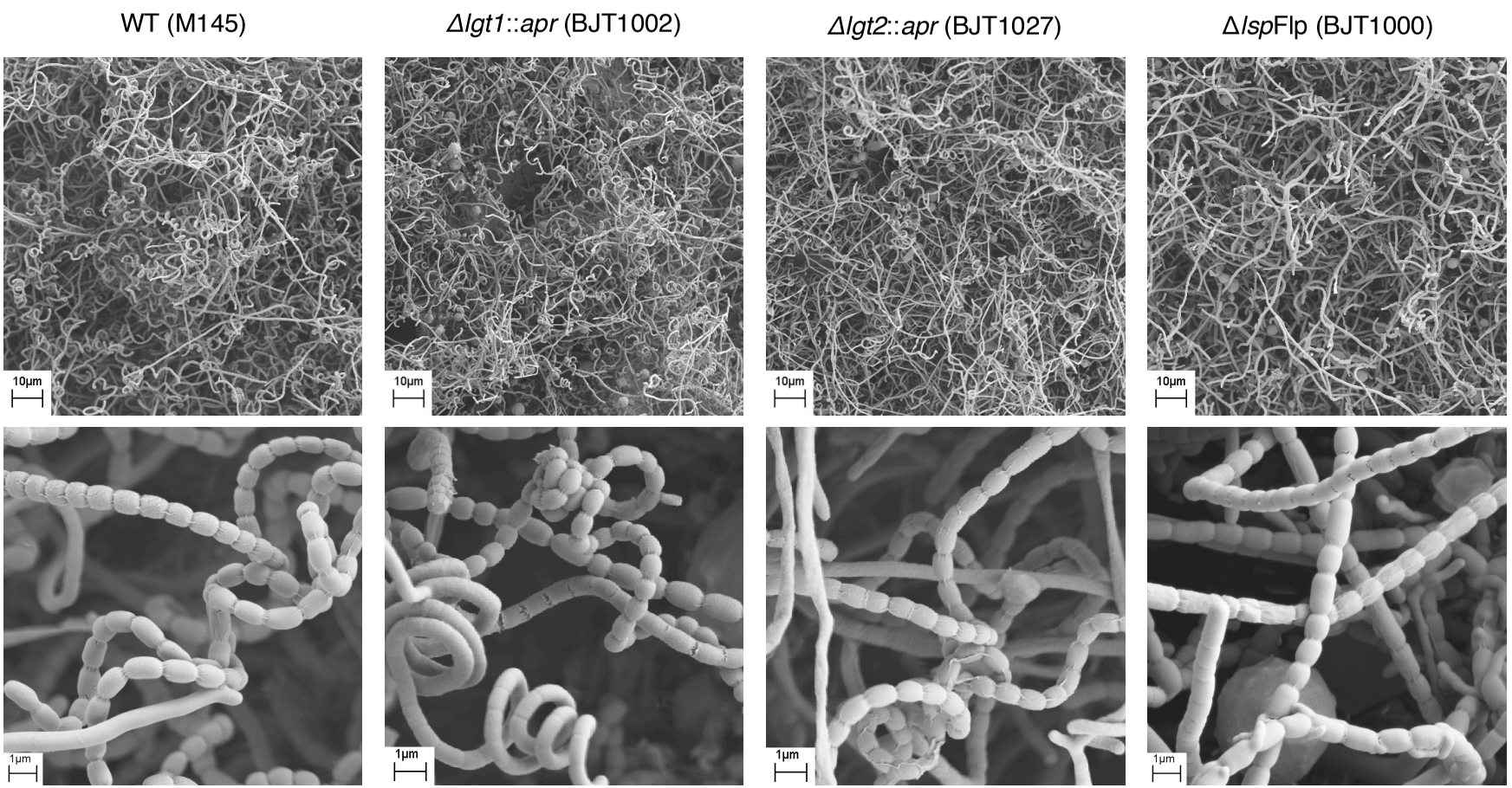


Figure 3.14. Scanning electron microscopy of each of the strains grown for 5 nights on SFM media. The *Δlsp* mutant contains less spore chains than the other mutants, which look similar to the wild-type. Strain names are shown in brackets.

3.3.4 Effect of gene disruption on lipoprotein processing.

3.3.4.1 Western blots.

Spores of each strain were grown for 15 hours in 50% TSB/YEME growth media (table 2.3) and the resultant cell extract was fractionated by ultracentrifugation to obtain separate cytoplasmic and cell membrane fractions. Each group of proteins was run on a separate 15% SDS PAGE gel, with the amount protein loaded in each lane normalised. The results for each of the fractions can be seen in figures 3.15 and 3.16.

3.3.4.2 CseA.

The first lipoprotein assayed was CseA (SCO3357), a Sec translocated, experimentally confirmed lipoprotein from *S. coelicolor*, with no known homologues outside of the streptomycetes [93]. CseA is part of the CseABC-σ^E regulatory system which is involved in sensing of cell envelope stress in *S. coelicolor*. Good polyclonal antibodies exist against this lipoprotein [93]. The CseA protein runs as a band at 21 kDa on a protein gel. This band was absent from all the cytoplasmic fractions (fig. 3.15). The band was also undetectable in the precipitated growth medium (not shown). Being a membrane bound protein, CseA was present as expected in the membrane fractions of wild-type *S. coelicolor*. It is absent from the *ΔcseA* null mutant (J2172 - see table 2.2). In the membrane fractions of both the *Δlgt1* (BJT1002) and *Δlgt2* (BJT1027) strains CseA is present, with no detectable size shift in comparison to the wild-type. However, in the *Δlsp*FLP mutant (BJT1001), CseA was undetectable.

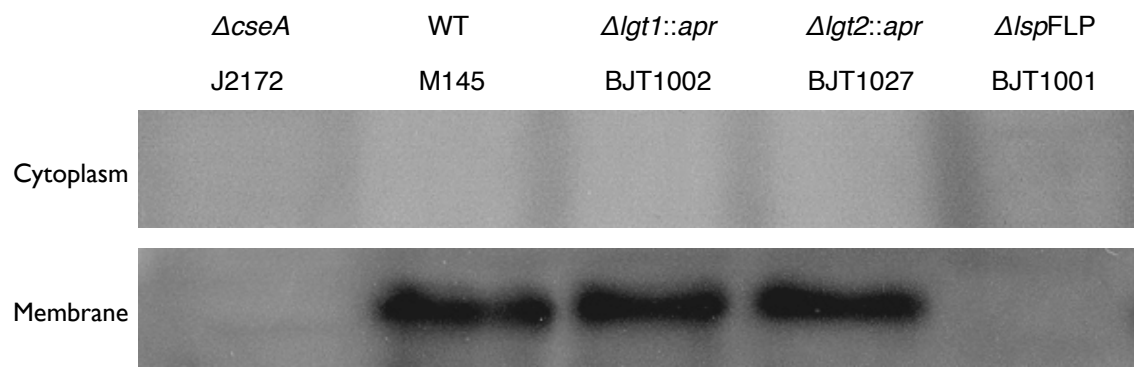


Figure 3.15. Western blot analysis of the cytoplasm and membrane fractions of the various *S. coelicolor* strains. CseA can be seen in the WT, *Δlgt1* and *Δlgt2* membrane samples. The fractions were incubated against an anti-CseA antibody. Strain names are as labelled.

3.3.4.3 SCO3484

The second lipoprotein assayed was SCO3484, a Tat translocated, confirmed lipoprotein [2, 82], and a substrate binding protein of unknown function. An effective antibody against this lipoprotein is unavailable at the current time, so a fusion was created whereby the gene was cloned with six histidine codons at its 3′ end (pTDW188 - see table 2.2). Constructs carrying this altered gene were conjugated into each *S. coelicolor* strain in single copy, integrating using the ϕC31 *attP-int* locus, and membrane fractions of each of the strains were prepared, as discussed previously. These fractions were immunoblotted with a commercially available, monoclonal anti-His antibody (Qiagen). This antibody is highly sensitive and is able to detect proteins at a very low level. The results (fig 3.16), show that as with CseA, the processing of SCO3484 in both the *Δlgt1* and *Δlgt2* mutants appears to be the same as the wild-type. However, the processing in the *Δlsp* strain is different. In this case the protein, whilst present at a low level, is detectable although it is larger than the protein from the wild-type. This suggests that it retains its signal peptide due to the lack of Lsp and

that the majority of the protein is proteolysed to remove it from the membrane. The size of the protein in an *in trans* complemented *Δlsp* mutant is returned to normal. This will be discussed in section 3.3.5.2 below.

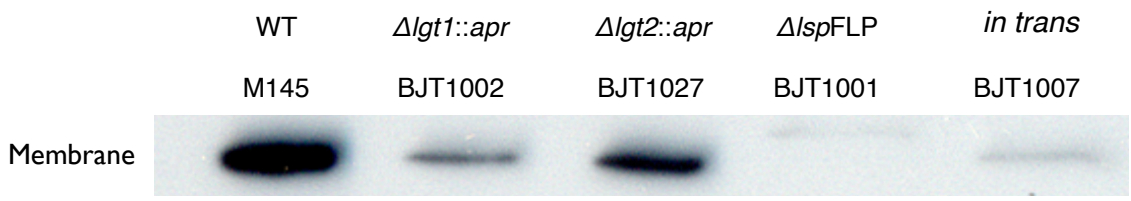


Figure 3.16. Western blot analysis of His-tagged SCO3484 using an anti-His antibody in the membranes of each of the strains listed. The protein in the *Δlsp* mutant has shifted in comparison to the other strains, suggesting it is improperly processed. It has reverted to its fully processed size in the complemented strain. Strain names are as labelled.

3.3.5 Complementation of the *Δlsp* mutant.

Of the three gene disruptions made, only the *Δlsp* mutant displays a severe phenotype, as described above. In an attempt to reverse these mutations, a full length copy of the *lsp* gene was replaced into the *Δlsp* strain in order to see if could restore the mutant back to a wild-type phenotype. This was done in two ways, with a full length copy of the gene and its promoter being reintroduced into the mutant on an integrative plasmid, or by reversing the REDIRECT mediated apramycin gene disruption (section 2.4), to reintroduce a full length *lsp* gene back into the bacterial chromosome, thus reconstructing the wild-type strain. A copy of the full length *lsp* was synthesised by Genscript USA inc. with an additional 300 bp of upstream DNA incorporating the *lsp* promoter (see Chapter 2).

The *in trans* complementation of the *Δlsp* mutant involved subcloning the full length *lsp* gene, under the control of its own promoter, into the vector pSET152 to form vector pBT100 (see table 2.2) and transforming it into the unmarked *Δlsp* deletion strain to create the strain BJT1007.

Replacing the *lsp* gene into the chromosome of the mutant to make a *cis* complemented strain (section 2.4.2) involved replacing the ampicillin resistance gene (*bla*) found on the backbone of the cosmid 4A10 (which contains full length *lsp*) with the hygromycin resistance cassette (*hyg*) containing the origin of transfer (fig. 3.17). This altered cosmid, pBT110, was conjugated into M145 *lsp::apr* (BJT1000) and double exconjugants were selected by choosing colonies that were hygromycin sensitive and apramycin sensitive (BJT1006). Genomic DNA was tested by PCR (fig. 3.18), as described earlier, to ensure that the full length *lsp* was present.

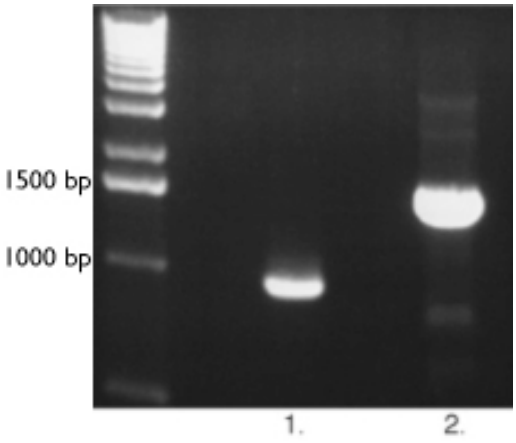


Figure 3.17 PCR confirmation of the replacement of the *bla* gene in the backbone of cosmid 4A10 with the *hyg* cassette. Lane 1. = cosmid 4A10, lane 2. = cosmid 4A10 *bla->hyg*. The *bla* Test forward and reverse primers were used. The band in lane 1 (~1 Kb) is the *bla* gene and flanking DNA, absent in the altered cosmid and replaced by the *hyg* cassette (~1.5 Kb).

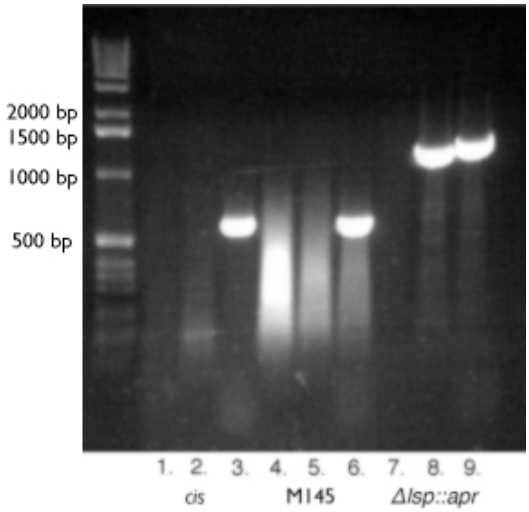
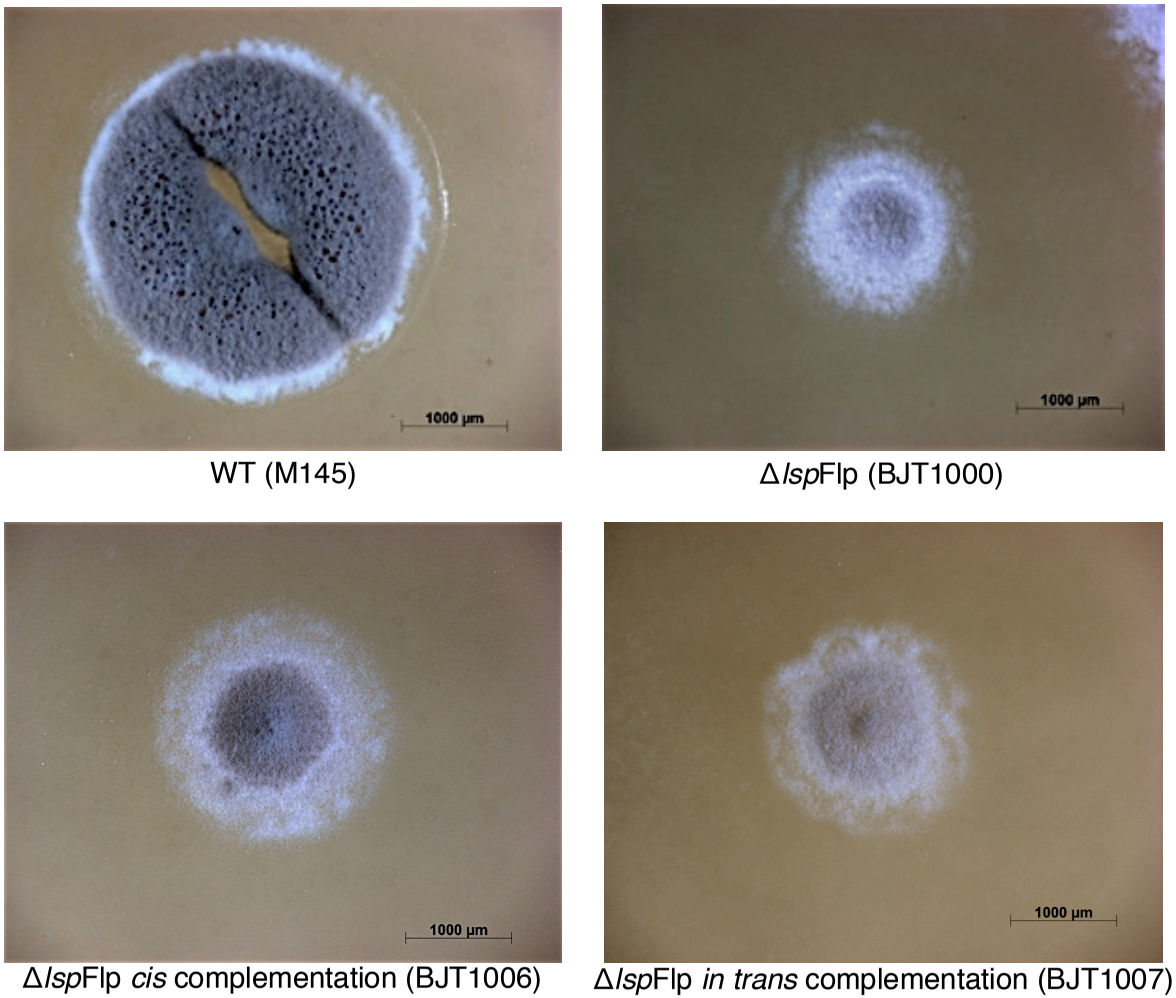


Figure 3.18. PCR of genomic DNA to confirm the replacement of full length *lsp* into the *Δlsp::apr* mutant. Lanes 1-3 = *cis* complementation, 4-6 = M145 wild-type, 7-9 = M145 *lsp::apr.* Primer combinations in lane 1,4,7 = *hyg* test forward/reverse. Lanes 2,5,8 = P1 and P2. Lanes 3,6,9 = *lsp* Test forward/reverse. A band representing full length *lsp* (614 bp) is present in the wild-type and *cis* complementation. This band is absent in the *Δlsp::apr* mutant, where the *apr* gene is present (1424 bp), replacing the *lsp* gene.

3.3.5.1 Development and colony morphology.

When plated as a confluent lawn, as in figure 3.11 the complementations both look closer to the wild-type than the *Δlsp*FLP mutant (BJT1001). They appear to have grown more than the mutant, and are sporulating to a greater extent. When individual colonies were viewed under the light microscope however, the similarities between the wild-type and complemented strains end. The results for colonies grown for 5 nights can be seen in figure 3.19. As before the *Δlsp* mutant colony is dwarfed by the wild-type. Neither of the complemented strains have returned to the size of the wild-type, and appear to approximately the same size as the deletion mutant. The *cis* complemented strain (BJT1006) is markedly more grey than the *in trans* complemented strain (BJT1007) (fig. 3.19)

and is producing a higher number of spores, which is confirmed in the SEM images (fig. 3.22). Both complemented colonies retain the large proportion of white areas seen in the *Δlsp* mutant but absent in the wild-type. The cross-sections of each of the colonies (fig. 3.20) confirms the drastic difference with both complemented strains show the same lack of depth observed in the *Δlsp* mutant.



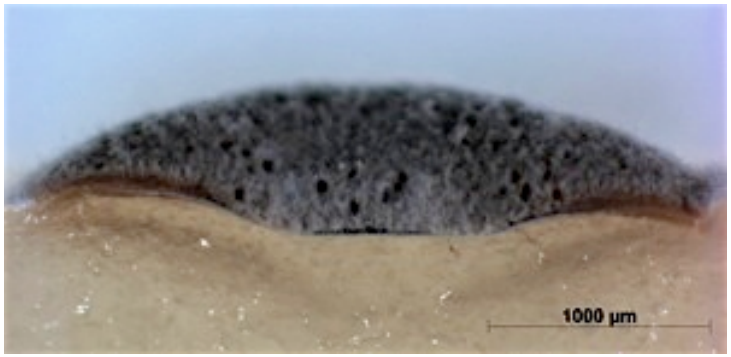
WT (M145)

Δ*lsp*Flp (BJT1000)

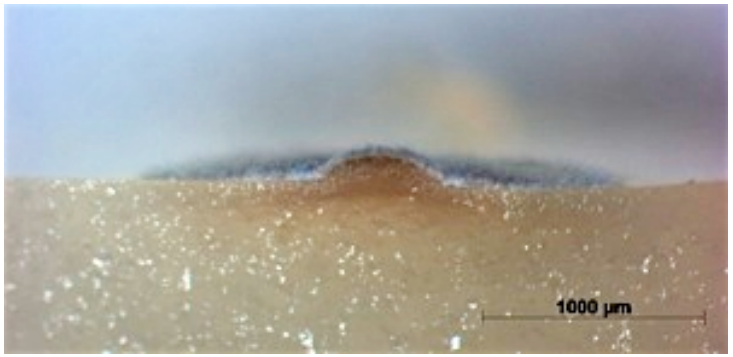
Δ*lsp*Flp *cis* complementation (BJT1006)

Δ*lsp*Flp *in trans* complementation (BJT1007)

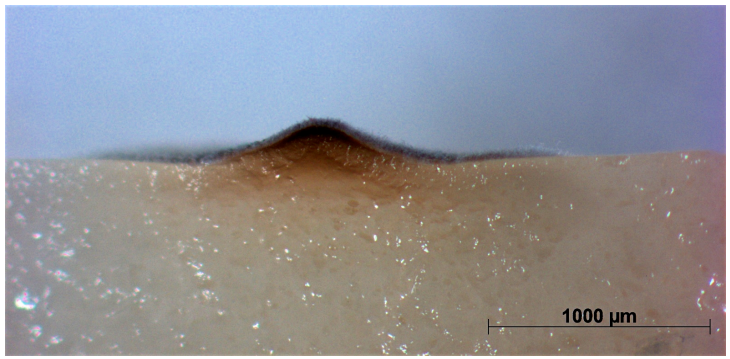
Figure 3.19. Light microscopy of the wild-type, *Δlsp* and complemented strains. Strain names are shown in brackets.



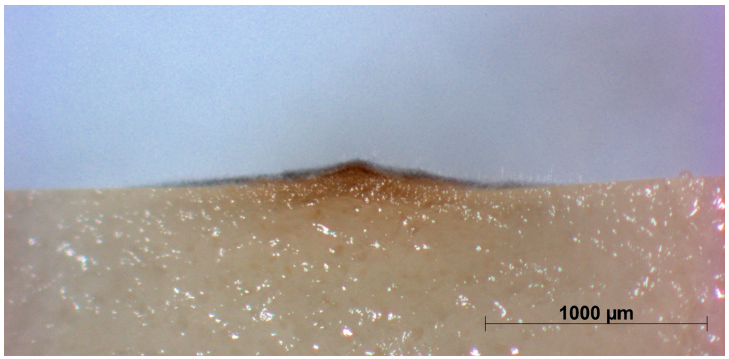
WT (M145)



Δ*lsp*Flp (BJT1000)



Δ*lsp*Flp *cis* complementation (BJT1006)



Δ*lsp*Flp *in trans* complementation (BJT1007)

Figure 3.20 Cross sections of the wild-type, *Δlsp* and complemented strains. Strain names are shown in brackets.

Whilst light microscopy shows very little difference between the *Δlsp* and the complemented strains, the use of SEM revealed some variations unseen at a low magnification. Colonies viewed after both 2 and 5 nights can be seen in figures 3.21 and 3.22.

After 2 nights, the *cis* complementation (BJT1006) resembles the wild-type, with a large amount of aerial hyphae present. As with the wild-type, there are no spore chains visible. The *in trans* complemented strain (BJT1007) however closely resembles *Δlsp* strain and shares the distinctive lack of aerial hyphae. High magnification of these strains show there is no vertical growth of aerial hyphae and any depth of the colony appears to be a result of substrate hyphae aggregating on top of each other. After 5 nights, all the strains look broadly similar, but closer inspection reveals that whilst both the wild-type and *cis* complemented strains share an abundant number of curly aerial hyphae, they are not present in either the *Δlsp* mutant or the *in trans* complemented strain. The highest magnification does reveal though that mature spores are present in all four of the strains. If you were to base the analysis of the complemented strains solely on these SEM data it would appear that the *cis* complementation closely resembles the wild-type, whilst the *in trans* complementation is closer to the *Δlsp* mutant.

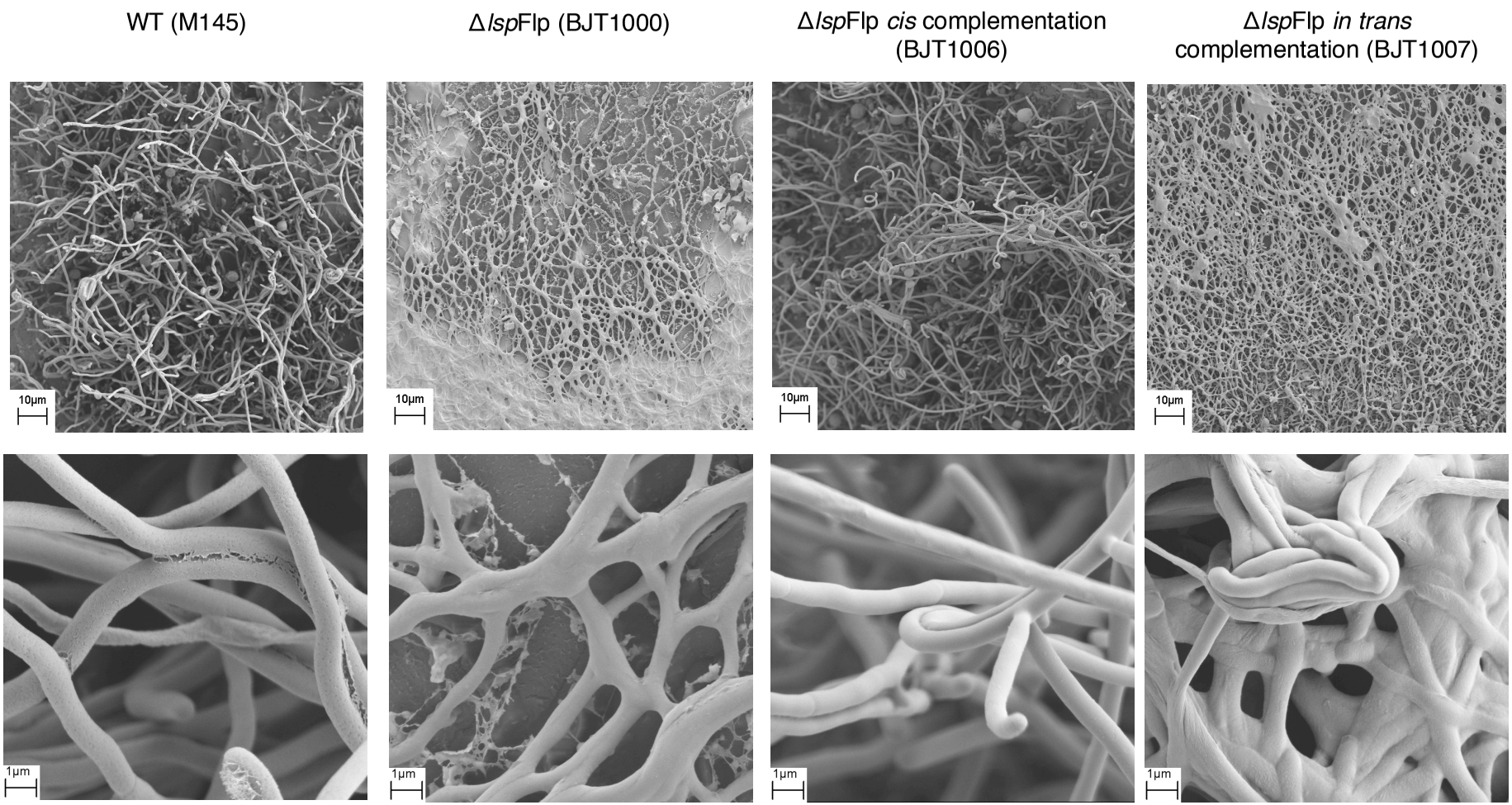


Figure 3.21. Scanning electron microscopy of each of the wild-type, *Δlsp* and complemented strains grown for 2 nights on SFM media. Strain names are shown in brackets.

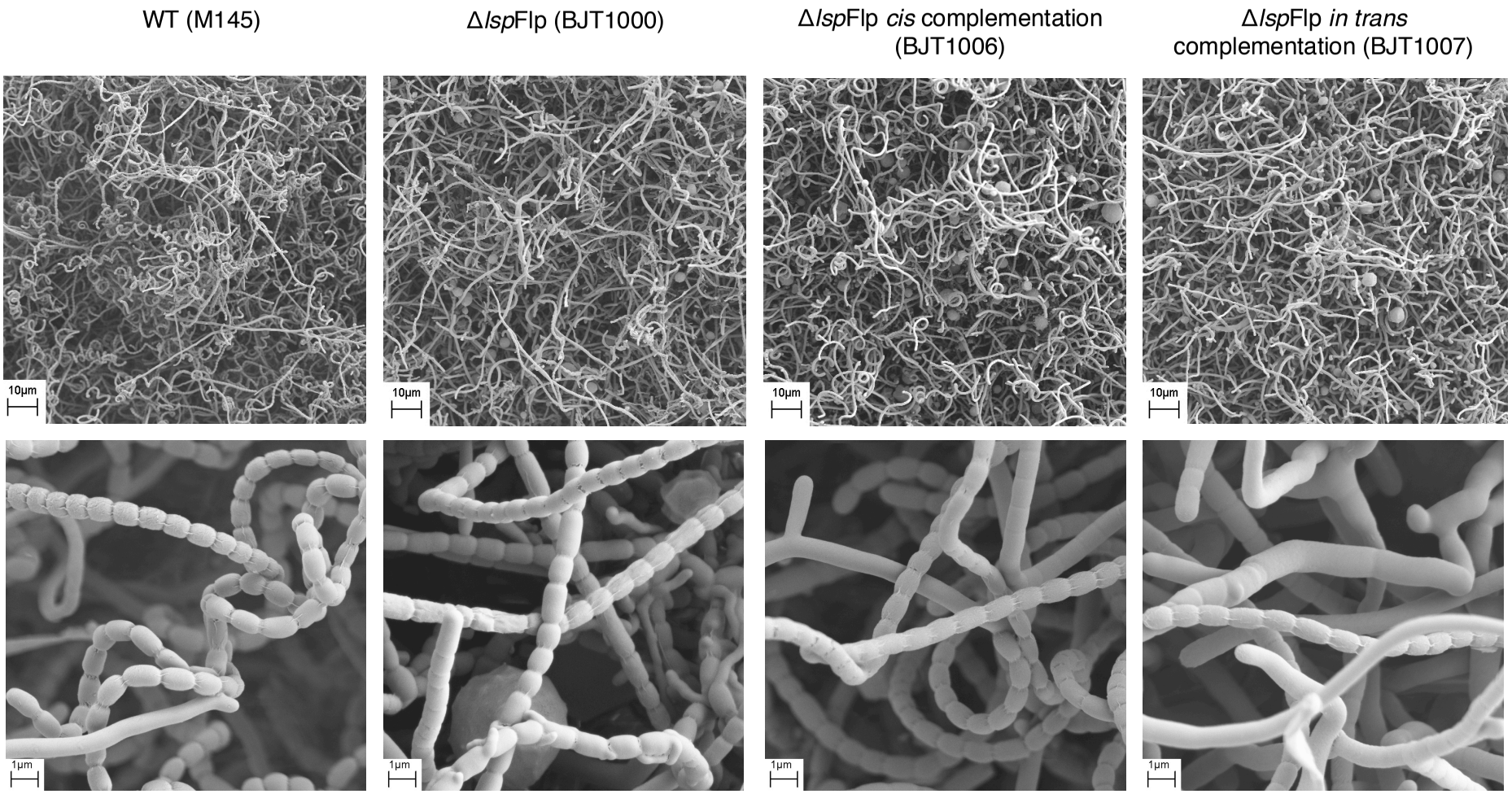


Figure 3.22. Scanning electron microscopy of each of the wild-type, *Δlsp* and complemented strains grown for 5 nights on SFM media. Strain names are shown in brackets.

3.3.5.2 Effect of gene disruption on lipoprotein processing.

As can be seen in figure 3.16, *in trans* complementation of the *Δlsp* mutant (BJT1007) is able to restore the processing of SCO3484 (the *cis* complementation could not be tested at the time due to a clash of antibiotic resistance cassettes in the integrative vectors). In order to confirm this, membrane fractions from the wild-type, *Δlsp* mutant and complemented strains were collected as before and tested against an unaltered native lipoprotein, in this case CseA. The results can be seen in figure 3.24 below. As before, the band representing CseA in the wild-type membrane is absent from the *Δlsp* membrane. It is however, detectable in both of the complemented strains at the same size as the wild-type band.

3.3.6 Phenotype tests.

Many lipoproteins are involved in maintaining cell envelope integrity. In *E. coli*, the murein lipoprotein (Lpp) is needed to covalently attach the peptidoglycan cell wall to the inner membrane. It seems feasible that disruption of the lipoprotein biosynthetic pathway may lead to a weakened cell envelope due to the incomplete processing, or absence of lipoproteins. To test this hypothesis the wild type and mutant strains were tested for sensitivity to a variety of stresses. Identical numbers of spores were plated into each well of a 12 well cell-culture plate containing 3ml of minimal growth media (+1% glucose). Each well contained a decreasing concentration of the compound to be tested, with the highest concentration being well A1 and the control well being C4 (fig. 3.23). Growth was scored as 2 if the growth was confluent, 1 if individual colonies were sparse enough to be counted by eye, or 0 for the absence of growth (for full results see appendix 2).

Name/Strain	WT	Δ*lgt1*::*apr*	Δ*lgt2*::*apr*	Δ*lsp*FLP	*cis* comp	*in trans* comp
	M145	BJT1002	BJT1027	BJT1001	BJT1006	BJT1007
Bacitracin						
Confluent Growth (2) up to	50 µg/ml	30 µg/ml	100 µg/ml	20 µg/ml	35 µg/ml	35 µg/ml
Weak Growth (1) up to	N/A	40 µg/ml	500 µg/ml	30 µg/ml	40 µg/ml	50 µg/ml
+15 Hour SDS						
Confluent Growth (2) up to	0.2%	0.2%	0.2%	0.05%	0.06%	0.04%
Weak Growth (1) up to	N/A	N/A	N/A	0.1%	0.1%	0.1%
0 hour lysozyme						
Confluent Growth (2) up to	0.5 µg/ml	0.5 µg/ml	0.5 µg/ml	0.3 µg/ml	0.75 µg/ml	0.75 µg/ml
Weak Growth (1) up to	1.25 µg/ml	2.5 µg/ml	2.5 µg/ml	0.75 µg/ml	1 µg/ml	1.25 µg/ml
+15 Hour Lysozyme						
Confluent Growth (2) up to	1 µg/ml	0.5 µg/ml	0.3 µg/ml	0.4 µg/ml	0.3 µg/ml	0.4 µg/ml
Weak Growth (1) up to	1.25 µg/ml	1.25 µg/ml	1.25 µg/ml	1.25 µg/ml	2.5 µg/ml	1.25 µg/ml

Table 3.2. Abbreviated results of the stress test carried out on each of the *S. coelicolor* strains. As described, confluent growth was scores as a 2, whilst weak growth was scored as a 1. N/A represents confluent growth to the highest level tested.

3.3.6.1 Cell envelope specific antibiotics.

Of the antibiotics tested, bacitracin, carbenicillin and vancomycin, the most obvious effect was observed with bacitracin. Bacitracin is a cyclic dodecapeptide produced by *Bacillus* species, and an unusual antibiotic because it requires a divalent metal ion (for example Zn^{2+}) in order to function [94]. Bacitracin inhibits cell wall biosynthesis in Gram-positive bacteria by interfering with the transport of peptidoglycan precursors, leading to the cessation of cell wall synthesis and eventually cell lysis (fig. 3.27) [95]. Wild-type *S. coelicolor* exhibited growth at ≤50 µg ml^{-1} bacitracin but higher concentrations were lethal. The *Δlgt1* mutant (strain BJT1002 - see table 2.2) showed slightly lower resistance, with strong growth at ≤30 µg ml^{-1} and weak growth up to 40 µg ml^{-1}. The *Δlgt2* mutant (BJT1027) showed a higher level of resistance than the wild-type, growing strongly at ≤100 µg ml^{-1} and growing

weakly at 500 μg ml^{-1}. In contrast, *Δlsp* (BJT1001) showed decreased resistance compared to the wild-type, only showing strong growth at ≤20 μg ml^{-1} and weak growth up to 30 μg ml^{-1}.

3.3.6.2 SDS Sensitivity.

The detergent Sodium Dodecyl Sulphate (SDS) is an anionic surfactant. It is used frequently in experiments to denature proteins by disrupting their non-covalent bonds. As a detergent, SDS also affects lipid-lipid interaction in the cell membrane phospholipid-bilayer [96], potentially leading to a weakened cell membrane. Of the three SDS experiments carried out, neither incubating the spores for 1 hour in SDS nor overlaying the spores with SDS after 0 hours lead to any discernible effect on the strains. Overlaying the plates after 15 hours growth (newly germinated spores) had no effect on the wild-type, *Δlgt1* (BJT1002) or *Δlgt2* (BJT1027) strains, each of which was resistant to ≤0.2% SDS. The *Δlsp* strain (BJT1001) showed strong growth at ≤0.05% SDS and weak growth up to 0.1% SDS. There was no growth at 0.2% SDS. These results suggest that *S. coelicolor* spores are very resistant to SDS as are newly germinated spores under normal conditions but deletion of *lsp* increases SDS sensitivity.

3.3.6.3 Lysozyme sensitivity.

Lysozyme belongs to a family of enzymes known as the 1,4-β-*N*-acetylmuramidases which damage Gram-positive bacterial cell walls [97]. This enzyme attacks peptidoglycan by hydrolysing a glycosidic bond between *N*-acetylmuramic acid and *N*-acetylglucosamine [98]. The weakening of the cell wall means that an affected bacterium is more susceptible to lysis caused by turgor pressure. Lysozyme had an effect on the strains both when overlain after

0 hours and 15 hours. After 0 hours, the wild-type showed confluent growth on ≤0.5 μg ml^{-1} and weak growth up to 1.25 μg ml^{-1}. The *Δlgt1* (BJT1002) and *Δlgt2* (BJT1027) mutants were broadly similar. The *Δlsp* mutant (BJT1001) was the most sensitive, showing confluent growth up to ≤0.3 μg ml^{-1} and weak growth up to 0.75 μg ml^{-1}. This pattern is changed in the +15 hour overlay experiment. In this case, the wild-type and *Δlsp* mutant both show higher levels of resistance than at 0 hours, with confluent growth at ≤1.0 μg ml^{-1} and ≤0.4 μg ml^{-1} respectively. The *Δlgt1* mutant shows a similar level of resistance to the 0 hour overlay, but the *Δlgt2* mutant shows considerably lower resistance, with confluent growth at ≤0.3 μg ml^{-1} and weak growth up to 0.75 μg ml^{-1}.

3.3.6.4 Complemented *Δlsp* strain phenotypes.

As the *Δlsp* mutant typically displayed enhanced sensitivity to stresses compared to the wild-type, the same phenotype assays were carried out using the *cis* and *in trans* complemented *lsp* strains (BJT1006 and BJT1007 respectively). The resistance of both complemented strains to bacitracin increased. Both showed a similar level of resistance, with confluent growth seen at ≤35 μg ml^{-1} compared to ≤20 μg ml^{-1} for the *Δlsp* mutant (BJT1001) and ≤100 μg ml^{-1} for the wild-type (M145).

The *Δlsp* mutant showed confluent growth on 0.05% SDS when overlain after 15 hours. The complemented strains both showed approximately the same level of resistance, with the *cis* complementation slightly higher at 0.06% and the *in trans* complementation slightly lower at 0.04%. These results are still far lower than the wild-type, which showed confluent growth at the highest level of SDS tested, 0.2%.

The pattern of resistance for the complemented *Δlsp* strains when stressed by lysozyme differed between the 0 hour and 15 hour experiments. When overlain after 0 hours these strains showed levels of resistance closer to the wild-type whilst after 15 hours the levels were closer to the *Δlsp* mutant. After 0 hours, the wild-type showed confluent growth at ≤0.5 μg ml^{-1}. This was increased to ≤0.75 μg ml^{-1} in both the *cis* and *in trans* complementations, the *Δlsp* mutant grew confluently ≤0.3 μg ml^{-1}. When overlain after 15 hours, the wild-type grew confluently to ≤1.25 μg ml^{-1}, whilst both the *Δlsp* mutant and *in trans* complementation grew to ≤0.4 μg ml^{-1}. The *cis* complementation was even lower, only growing confluently to ≤0.3 μg ml^{-1}.

3.3.7 Truncating the N-terminus of Lsp.

As has been shown in figure 3.5, the Lsp protein from *S. coelicolor* contains an elongated N-terminus, when compared to both *E. coli* and *B. subtilis*. Similar elongations are also present in the Lsp enzymes of the other actinomycetes, as described above. In order to find out whether this extension is needed for the enzyme to function, deletions were made in the coding sequence of *lsp* in order to produce truncated enzymes. These truncated genes were synthesised by Genscript, sub-cloned into pSET152 to form vectors pBT101-105 (table 2.2), and introduced into the unmarked *Δlsp* deletion mutant (BJT1001) in single copy, integrating using the ϕC31 *attP-int* locus, to make strains which were identical to the *in trans* complemented strain (BJT1007), except for the loss of some, or all of the N-terminal extension. In total, four truncations were made, with enzymes starting at amino acid 10 (N10), 20 (N20), 30 (N30) and 40 (N40) (strains BJT1008-1011). The first amino acid of the N40 truncation aligns to the first amino acid of the *E. coli* Lsp (fig. 3.23). In addition to these truncations, a

double point mutant was constructed, where two active site aspartate residues, D148 and D177, were change to alanine (D148A/D177A), in order to create an inactive enzyme (BJT1012). These residues are conserved in all genomes checked and thought to make up a catalytic dyad forming the active site of the protein in *B. subtilis* (section 3.3.1.2).

```
BsLsp  ----------------------------------------------MLYYMIALLIIA 12
EcLsp  -----------------------------------MSQSICSTGLRWLWLVVVVLI 21
ScLsp  VAEAERIIGTPDIPDAAGEGQERPDADPERPEQQEQEQAPERTRGKRRVAVLFAVALFAYL 60
           N10       N20       N30       N40
```

Figure 3.23. The N-terminal extensions of Lsp from *B. subtilis*, *E. coli* and *S. coelicolor.* The four truncations are shown, with the amino acid which becomes the start of the protein highlighted in red. The codons for each of these amino acids were changed to encode methionine in the truncated forms.

To test if the truncations or the point mutant were able to complement the *Δlsp* mutant, membrane fractions were prepared and immunoblotted with the CseA antibody as described in section 3.3.4.2. The results can be seen in figure 3.24. As before, mature CseA is present in the wild-type and complemented strains and absent in the *Δlsp* strain. The fully processed protein is also present in the N10 truncation. There is a band in the N30 lane, but it is slightly larger than those seen in the other lanes. This would be consistent with an unprocessed form of CseA which retains its signal peptide. The protein is undetectable in the N20, N40 and D148A/D177A complementations. As before, CseA was undetectable in cytoplasmic fractions or precipitated supernatant (not shown).

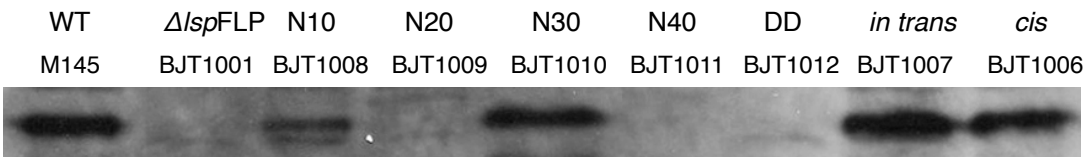


Figure 3.24. Membrane fractions from Wild-type, *Δlsp*, N-truncated Lsps, point mutant (DD), *cis* and *in trans* complementations. The band visible represents CseA. Strain names are as indicated.

3.3.8 Lsp specific antibody.

Whilst the results shown in section 3.3.7 demonstrated whether the CseA protein was processed correctly or not by the N-truncated Lsp enzymes, the absence of detectable lipoprotein neither confirms nor denies the presence of the Lsp enzyme itself. In order to see whether the absence of the model lipoprotein CseA was down to the absence of Lsp, an antibody was raised against the synthesised peptide KLEHHEPIEIIGDWLRFA, which corresponds to residues 72 to 89 in an extracytoplasmic loop of Lsp. Disappointingly this antibody was unable to detect the Lsp protein in *S. coelicolor* cell extracts of membrane fractions. This also thwarted a plan to visualise potential direct interactions between Lgt1/Lgt2 and Lsp using membrane crosslinking.

3.3.9 Complementing an *E. coli lsp* disruption.

The Lsp protein is encoded in only single copy in the majority of bacterial species studied, such as *B. subtilis* [20]. The protein is essential in *E. coli*, as has been shown when the protein is inhibited by the antibiotic globomycin. As has been discussed above, the *S. coelicolor lsp* gene has been correctly identified, and disrupting this gene leads to altered lipoprotein processing. It has also been shown that the N40 truncated Lsp protein is unable to rescue a *S. coelicolor Δlsp* mutant (section 3.3.7). To further investigate the role the N-

terminal extension plays in *S. coelicolor* Lsp, both the full length (FL) protein and N40 truncation were tested to see whether they could rescue an *E. coli ΔlspA* strain. Typically, to test whether a potential protein is a true Lsp homologue, it is expressed in *E. coli* to see whether the excess of the enzyme leads to an increase in resistance to globomycin. This approach was not chosen here as globomycin is in very short supply worldwide and is not commercially available. Instead, the full length *E. coli lspA*, FL *S. coelicolor lspA* and the N40 truncation gene were subcloned into the cosmid pBAD24-*Nde*I containing an arabinose inducible promotor, to create vectors pBT107-109 (table 2.2) and introduced into *E. coli* strain BW25113 / pIJ790 (see section 2.4.3). The *E. coli* chromosomal copy was disrupted with an apramycin cassette whilst the *in trans* alleles were expressed by the addition of arabinose.

Despite repeated attempts, a knockout of the *E. coli lspA* gene was not obtained. However, apramycin resistant colonies of the strains complemented with the N40 or FL *S. coelicolor lspA* gene were obtained, and an initial PCR using primers specific for the apramycin cassette confirmed it was present, as can be seen in figure 3.25 below.

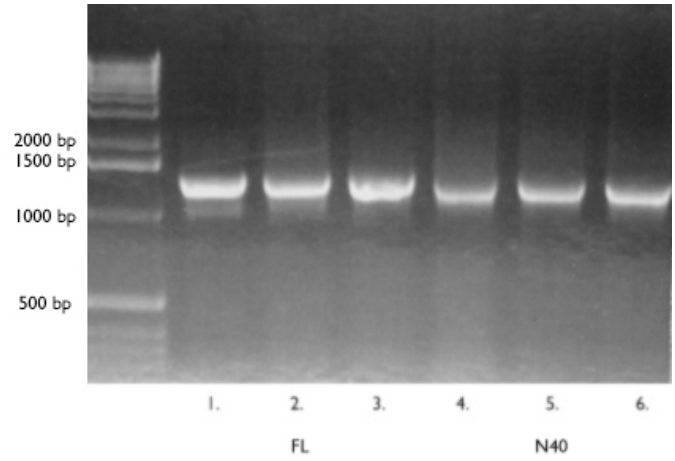


Figure 3.25. PCR showing the presence of the *apr* cassette in six potential *E. coli ΔIsp* colonies. Lanes 1->3 were complemented with full length *S. coelicolor lsp* and lanes 4->6 were complemented with N40 *lsp.* Bands at ~1400 bp represent the *apr* cassette.

Although these results showed that the apramycin cassette was present in both the FL and N40 complemented strains, it did not confirm that the *E. coli* chromosomal *lsp* gene had been disrupted. Using primers specific to the *E. coli lsp* gene as well as those specific for the apramycin cassette, showed that whilst the apramycin cassette is present in each (as before), a band representing the *E. coli lsp* gene is also present, suggesting that the antibiotic cassette has gone in to the chromosome in an incorrect location. Repeated attempts produced the same result suggesting that the *E. coli lsp* gene is very difficult to disrupt (fig. 3.26).

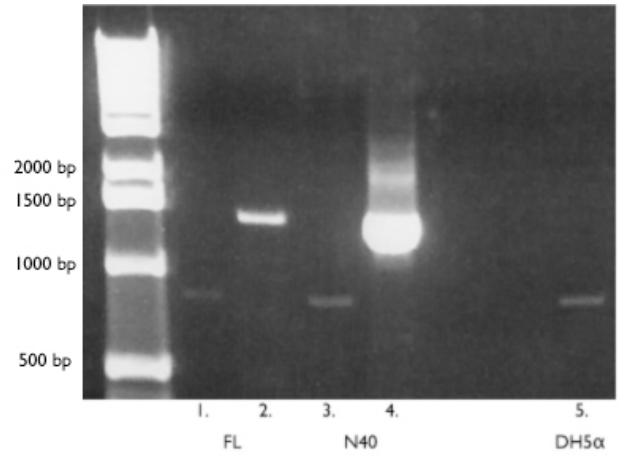


Figure 3.26. PCR of *E. coli* genomic DNA showing our inability to disrupt the *E. coli lsp* gene, despite the presence of the *apr* cassette. FL = full length *S. coelicolor lsp* complementation (Lanes 1 + 2). N40 = truncated *S. coelicolor lsp* complementation (lanes 3 + 4). Lanes 1 + 3 use the *E. coli lsp* Test forward/reverse primers. Lanes 2 + 4 use *apr* cassette P1 + P2 primers. Lane 5 shows the size of the *lsp* gene in *E. coli* DH5α using the *E. coli lsp* Test forward/reverse primers.

3.4 Discussion

The data presented in this chapter shows that disrupting the lipoprotein biosynthetic pathway in *S. coelicolor* has a deleterious effect on the overall fitness, and phenotype, of the bacterium. From a macroscopic perspective it is obvious that the *Δlsp* mutant (BJT1001 - see table 2.2) has a number of severe growth phenotypes, whilst the *Δlgt1* (BJT1002) and *Δlgt2* (BJT1027) mutants appear quite normal. This observation is backed up by both the light and scanning electron microscopy, both of which show that *Δlsp* is retarded in every aspect of its growth and development. This strain frequently overproduced the antibiotic actinorhodin, which is often seen as a response to cellular stress [92]. At this point it is timely to note that *S. coelicolor* has a tendency to show random growth phenotypes, however time was taken to ensure that the microscopic phenotypes described within this chapter were characteristic for the majority of

colonies/cells and for multiple, independently isolated, *Δlsp* colonies.

Bioinformatic analysis of each of the proteins making up the *S. coelicolor* lipoprotein biosynthetic pathway reveal that there are a number of differences compared to the published and studied enzymes from other bacterial species. Typically, once an immature lipoprotein has been exported out of the cell by Sec or Tat and is anchored into the cell membrane by its signal sequence, the first enzyme to act upon it is Lgt, which adds the diacylglyceryl moiety to the proprotein. *S. coelicolor* is unusual in that it contains two homologues of the Lgt protein, whereas *E. coli*, *B. subtilis* and the selection of actinomycetes chosen (fig. 3.2) only have one. Both of the *S. coelicolor* Lgt enzymes share a good level of similarity (~58%) and both share the His-196 and Tyr-235 residues that have been shown to be important for function in *E. coli* Lgt [17] which suggests that they are true Lgt enzymes. However His-103, which is essential for function in *E. coli* Lgt, is present in *B. subtilis*, yet absent from *S. coelicolor* Lgt1 and Lgt2 (as well as the other actinomycete Lgt enzymes) where it is a tryptophan. The reasons for this are unknown. Lgt in *E. coli* is used to transfer the lipid moiety from phosphatidyl glycerol (PG) to the sulphydryl group of the immature lipoprotein [12]. Perhaps, given the difference in membrane composition between the actinomycetes and *E. coli*, a different phosholipid is used as a substrate by their respective Lgt enzymes, and therefore the amino acid substitution is needed.

In contrast to Lgt, the *S. coelicolor* Lsp enzyme, which removes the signal peptide from the immature lipoprotein contains all the experimentally verified essential residues from another organism, in this case *B. subtilis*, as do all the

other species tested (fig. 3.4). The *S. coelicolor* Lsp shows a low similarity to both *E. coli* and *B. subtilis* Lsp (~30%), and the conserved regions I->V probably account for a large proportion of this similarity.

The essential amino acids of the final enzyme in the lipoprotein pathway, Lnt, have been experimentally verified in *E. coli* [30]. This enzyme was thought not to exist in Gram-positive bacteria, but has recently been identified in *M. smegmatis* and *M. tuberculosis* [38]. The *M. smegmatis* enzyme contains the three essential amino acids thought to form the catalytic region in *E. coli* [30], but only contains two of the other four essential *E. coli* residues (fig. 3.6). Given the lack of an outer membrane, the reason *S. coelicolor* contains two potential Lnt homologues is unclear, especially as at they appear, at first glance, to be functionally redundant. However, whilst *S. coelicolor* Lnt1 and Lnt2 share a low level of similarity (26%) to each other, the catalytic triad found in *E. coli* and *M. smegmatis* Lnt enzymes is present in both *S. coelicolor* homologues. Lnt1 shares three out of the other four essential *E. coli* residues, whist Lnt2 has two. Sadly, it was not possible to see whether either, or both of these proteins are able to N-acylate a lipoprotein, or whether their deletion alters the way lipoproteins are processed in *S. coelicolor* as the genes could not be disrupted in *S. coelicolor* (D. Widdick, unpublished).

One of the more striking aspects of the *S. coelicolor* proteins, is that both the Lgt enzymes, and Lsp contain extended C, or N-termini respectively, which are absent from *E. coli*. The functions of these extensions is unclear and will be discussed below.

As mentioned above, disrupting the *lgt* genes in *S. coelicolor* had very little effect on the bacterium, whilst disrupting *lsp* led to a highly pleiotropic phenotype. In order to try to explain why this might be the case, it is necessary to again consider the pathway by which lipoproteins are created. The loss of Lgt1 or Lgt2 had no effect on the processing of either of the model lipoproteins tested (CseA, SCO3484), both of which were detected in membrane fractions. Neither mutant had a detectable growth phenotype, with each strain looking comparable to the wild-type under both light and electron microscopy. Given that both of the lipoproteins tested are the same size as the wild-type protein it seems likely that they are correctly lipidated (fig. 3.15 and 3.16). In *B. subtilis* it has been shown that the signal sequence of an unlipidated lipoprotein is enough to anchor it to the cytoplasmic membrane [99]. If this were the case in either of the *Δlgt* mutants however, you would expect a size shift in the band representing the immature lipoprotein as it would have an increased mass due to the presence of the signal sequence. This is clearly not the case in either of the mutants tested.

A likely explanation for this lack of phenotype and the correct lipoprotein processing observed in both single mutants is that the the two Lgt homologues in *S. coelicolor* are complementing each other. In the absence of Lgt1 the lipoproteins are lipidated by Lgt2, and *vice versa*. Both Lgt enzymes are able to partially complement a *S. scabies Δlgt* mutant (BJT1040) (Chapter 4), which is good evidence that both are true Lgt homologues. However, these results raise several questions. Firstly, is the lipidation of of proteins so important to the viability of a *S. coelicolor* cell that it needs two copies of the *lgt* gene, in case one is lost? The *lgt2* (*SCO7822*) gene is found in one of the arms of the

chromosome, which typically contain non-essential genes most likely acquired through HGT [66]. Conversely, *lgt1* is found in the core-region of the chromosome, which contains the essential genes. Regardless of the position of the genes, it appears that both are retained by *S. coelicolor* suggesting that the lipidation step of the lipoprotein biosynthetic pathway is of great importance to the bacterium. The second question that is unanswered, is whether or not each Lgt enzyme has a specific 'sub-set' of lipoproteins to lipidate. *S. coelicolor* has an estimated 223 lipoproteins, and to test the processing of each one in both the *Δlgt1* and *Δlgt2* mutants (BJT1002 and BJT1027 repectively) would be an arduous task. Again though, this scenario seems unlikely, given that of the two lipoproteins tested neither showed a difference in either the *Δlgt1* or the *Δlgt2* mutant. If both were from the Lgt1 subset, you would expect a difference in the Lgt2 membrane and *vice versa*. Also, 2 dimensional gel electrophoresis carried out by D. Widdick shows that there is little difference between the proteomes of the membranes of wild-type and *Δlgt1*/*Δlgt2* mutants implying that the single mutants are complementing one another [82]. Our hypothesis that one enzyme caters to Sec translocated lipoproteins and the other caters to Tat translocated lipoproteins seems unlikely. CseA is Sec dependent, whilst SCO3484 is Tat dependent and both were were processed correctly in the *Δlgt1* and *Δlgt2* mutant. Given these results, it appears that, for reasons unknown, *S. coelicolor* has two functioning copies of the Lgt protein.

What is also unknown based upon these results, is whether lipidation by Lgt is a prerequisite for the action of Lsp, as it is in the majority of bacteria tested to date [14], with the exception of *L. monocytogenes*[24]. In order to find out if this is the case in *S. coelicolor* a double Lgt mutant is needed. Despite numerous

attempts, the construction of this mutant proved impossible. If the order of enzyme activity is tightly controlled, then the loss of lipidation will mean that Lsp is unable to function, leaving the immature lipoprotein bound to the membrane by its signal sequence with a detectable shift in size in the lipoproteins tested, as mentioned above. If the processing pathway is not tightly controlled then it might be expected that the lipoproteins tested would be undetectable in the membrane fraction of a *Δlgt1 lgt2* mutant as Lsp would act on the proprotein to cleave its signal sequence leading to its release into the extracellular medium as has been shown in *L. monocytogenes* [24]. Alternatively, the loss of lipidation may be fatal, which would explain the inability to construct the double mutant. An inability to construct a double *lgt* mutant in *S. coelicolor* suggests that Lgt function is essential in this bacterium.

The loss of Lsp had a large effect on the growth of *S. coelicolor* presumably due to the loss of lipoprotein processing. This is immediately obvious simply looking at the colonies by eye, without the need for magnification. Compared to the wild-type (M145) the *Δlsp*FLP (BJT1001) colonies are small, and overproduce actinorhodin. Under the light microscope the severity of the growth retardation is even clearer. The colonies are a fraction of the size of the wild-type, and are lacking any depth, appearing to grow almost flush to the surface of the growth media (figs. 3.12 and 3.13). They also sporulate poorly, and the shortage of aerial hyphae and spore chains can be observed when a colony is viewed under SEM (fig. 3.14).

The two lipoproteins tested were processed in different ways in the *Δlsp* mutant. SCO3484 was present in the cell membrane fraction (fig. 3.16) but was seen at

a larger size than in the wild-type, suggesting that it still contained an uncleaved signal peptide. CseA on the other hand, was completely undetectable in both the cytoplasmic and membrane fractions as well as the precipitated supernatant. The reasons for this difference are unknown, but it has been noted that the same lipoprotein can be processed in different ways, even in the same mutant [11]. Why there is this difference in processing is difficult to explain. A large amount of unprocessed lipoproteins anchored in the cell membrane by both their signal sequence and lipid moiety is likely to a detrimental effect on the integrity of the membrane. These accumulated proteins are likely to be broken down by extra-cellular proteases. If this were happening in the *S. coelicolor Δlsp* mutant (BJT1001) it could explain why no CseA was detectable. However as SCO3484 is present, then not all incorrectly anchored lipoproteins are proteolysed, or at least not all at the same rate.

Even though the lipoprotein biosynthetic pathway is non-essential in the Gram-positive bacteria tested to date, a number of essential lipoproteins do exist, such as PrsA in *B. subtilis,* required for the correct folding of translocated proteins [49], or PrtM and OppA from *Lactococcus lactis* [52]. The retention of one lipoprotein (SCO3484) and the loss of another (CseA) may indicate that there are a subset of lipoproteins that are essential or more important to the viability of the bacterium. For example, the *S. coelicolor* lipoprotein SCO1639 is involved in the correct folding of translocated proteins and its homologue found in *Streptomyces anulatus* (formerly *S. chrysomallus*) could not be deleted [100], suggesting that essential lipoproteins probably do exist in *Streptomyces*, although SCO1639 can be disrupted in *S. coelicolor* [82].

If the absence of CseA is representative of a number of lipoproteins in the *Δlsp* mutant (BJT1001) it might explain the dramatic phenotypes seen. Complementing the mutant both *cis* and *in trans* (BJT1006 and BJT1007 respectively) restored the correct processing of CseA and it was detectable in the membrane fractions of each at the same size as in the wild-type (fig. 3.24). The *in trans* complementation also lead to the correct processing of SCO3484, which was detectable at the same size as in the wild-type, not the increased size seen in the *Δlsp* mutant, suggesting that the signal sequence is again being cleaved (fig. 3.16). Two dimensional electrophoresis of lipoproteins from the wild-type, *Δlsp* and complemented strains performed by D. Widdick confirms that whilst there is a large subset of lipoproteins lost from the *Δlsp* membrane when compared to the wild-type, they are restored in both the complemented strains [82].

Whilst the processing of lipoproteins was restored by complementing the loss of Lsp, the retarded growth phenotype remained [82]. The *cis* complementation (BJT1006) seems to produce a higher number of spores than the *in trans* complementation (BJT1007) although both show no increase in colony size compared to the *Δlsp* mutant, and colonies of both complemented strains remain much smaller than the wild-type (fig. 3.19). Both complemented strains also retain the characteristic flat colony phenotype shown by the *Δlsp* mutant (fig. 3.20). Surprising differences between the two types of complementations were seen when viewed under SEM (figs. 3.21 and 3.22). After 5 nights of growth, the wild-type, *Δlsp* (BJT1001) and complemented strains (BJT1006 and BJT1007) looked quite similar, but after only 2 nights, the *cis* strain closely resembled the wild-type, whilst the *in trans* strain was much closer to the *Δlsp*

mutant, showing that replacing Lsp in different ways has lead to different results. This, and the inability to restore a wild-type phenotype in a complemented strain, suggests that the growth defects seen in the *Δlsp* mutant are not solely down to the loss of Lsp. Deletion of *lsp* appears to result in spontaneous secondary mutations that may suppress an otherwise lethal phenotype. If lipoprotein processing is essential in *S. coelicolor* this might also explain why it encodes two copies of Lgt and why a double *lgt* mutant could not be isolated in this study.

This idea is further backed up by the stress tests carried out. Figure 3.27 below shows the actions of several antibiotics that effect the cell envelope in various ways. The majority of the antibiotics tested had no effect on any of the *S. coelicolor* strains (data not shown). For example, *S. coelicolor* has a natural resistance to the β-lactam group of antibiotics due to its constitutive expression of β-lactamases [101]. *S. coelicolor* also contains a gene cluster which confers resistance to the glycopeptide antibiotic vancomycin, likely transferred by HGT from other members of the actinomycetes [93].

The most drastic effect was seen with bacitracin. Whilst the wild-type and *Δlgt2* mutant showed good growth at high concentrations of bacitracin (100 and 200 μg/ml respectively), the *Δlsp* mutant is much more sensitive, only showing strong growth at 20 μg/ml. The *Δlgt1* strain was in between showing strong growth up to 30 μg/ml. As has been mentioned, bacitracin, a cyclic dodecapeptide produced by *Bacillus* species, inhibits cell wall biosynthesis in Gram-Positive bacteria by binding to the lipid carrier undecaprenyl pyrophospate (UP), preventing it from being dephosphoylated by membrane

associated pyrophosphatases [94, 95]. Preventing this dephosphorylation means that the UP cannot be recycled, reducing the amount available for transport of the newly synthesised peptidoglycan precursors, leading to the cessation of cell wall synthesis and eventually cell lysis (fig. 3.27) [103]. Currently, there are four known resistance mechanisms to bacitracin: removal of the antibiotic by an ABC transporter, overexpression of UP, expression of alternative UP phosphatase or exopolysaccaride production [95].

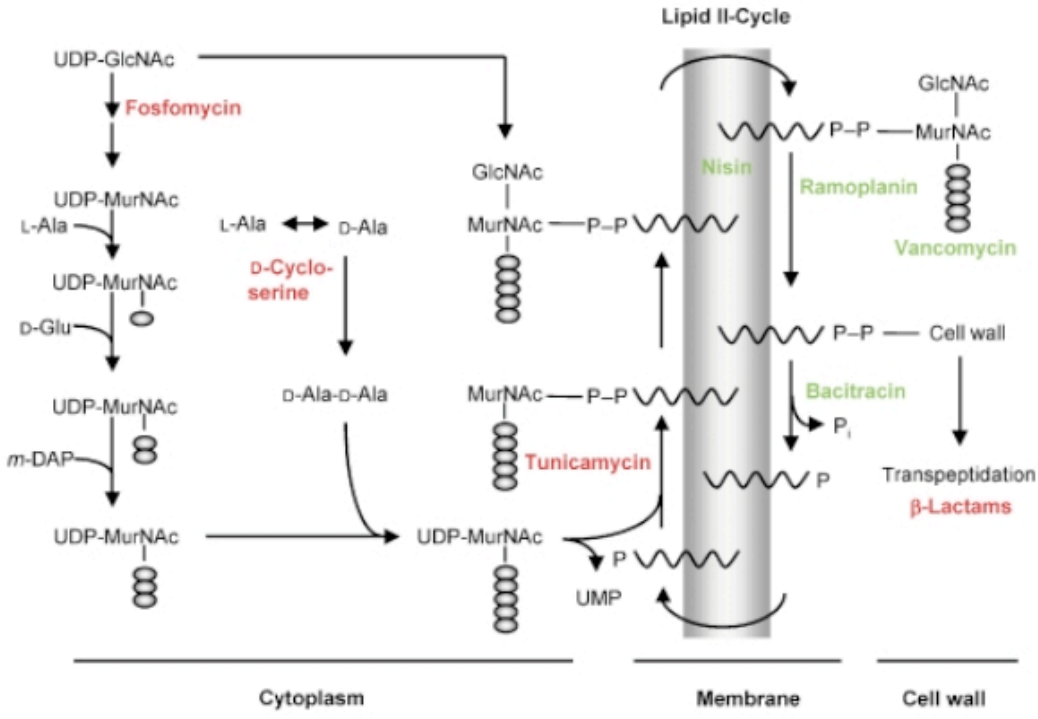


Figure 3.27. Diagram showing the effects of a variety of antibiotics on cell wall biosynthesis. From [102].

Both the detergent SDS, and the muramidase lysozyme, had a severe effect on the *Δlsp* mutant. When overlain after 15 hours with SDS the wild-type and *Δlgt* mutants all survived the highest % tested (0.2%), whilst the *Δlsp* mutant grew confluently at a quarter of this. The spores of all of the strains were impervious

to SDS and the results show that the *Δlsp* mutant has a compromised cell membrane.

The lysozyme results are curious. When overlain after 0 hours, the *Δlsp* mutant (BJT1001) is more susceptible than the other strains which are all resistant to the same concentration. When overlain over 15 hours however, the results are different, with all the mutants more susceptible than the wild-type. The lysozyme in the 0 hour overlay affects the germinating spores, and as we have seen the growth of the *Δlsp* strain is much more feeble than the others, suggesting that the newly emerging germling cell walls may be more susceptible to cell wall damage than the other strains. When the strains are overlain after 15 hours the resistance in the *Δlsp* strain increases relative to the 0 hour overlay, either because; there is a higher density of cells, and therefore there is a much larger amount of cell wall present raising the resistance to the lysozyme or the mycelial cell walls are simply more resistant than the germling cell walls. Quite why the resistance of the *Δlgt2* mutant drops is unknown, although reproducible.

As has been discussed, complementing the *Δlsp* strain leads to the resumption of correct lipoprotein processing, but shows mixed results in the phenotype assay tests. Both the *cis* and *in trans* complementations have a higher resistance to bacitracin than the *Δlsp* mutant, although they do not match the wild-type levels. This is also true of the 0 hour Lysozyme overlay. Conversely neither complementation shows an increased resistance to SDS, nor 15 hour lysozyme overlay.

Recent work has uncovered a new type of enzyme that may shed some light on why the *Δlsp* mutant is so damaged, and why complementing the mutant restores lipoprotein processing, but only some of its ability to resist cell envelope stress.

Typically, in the final stage of cell wall synthesis, glycan chains are polymerised by transglycosylation and transpeptidation, in which peptide side chains are cross-linked to form the rigid cell wall peptidoglycan (fig. 3.28) [104]. The enzymes that facilitate this reaction are transglycosylases/transpeptidases, also know as Penicillin Binding Proteins (PBPs). The peptidoglycan is crosslinked by transpeptidases between Ala_4 and Lys_3 leading to a D,D (also known as 3,4) transpeptidation. PBPs are targets for the β-lactam antibiotics which inhibit cell wall synthesis by acting as analogues of the D-Ala-D-Ala portion of peptidoglycan, covalently altering the transpeptidase active site [95] and preventing peptidoglycan cross-linking, leading to a mechanically weakened cell wall.

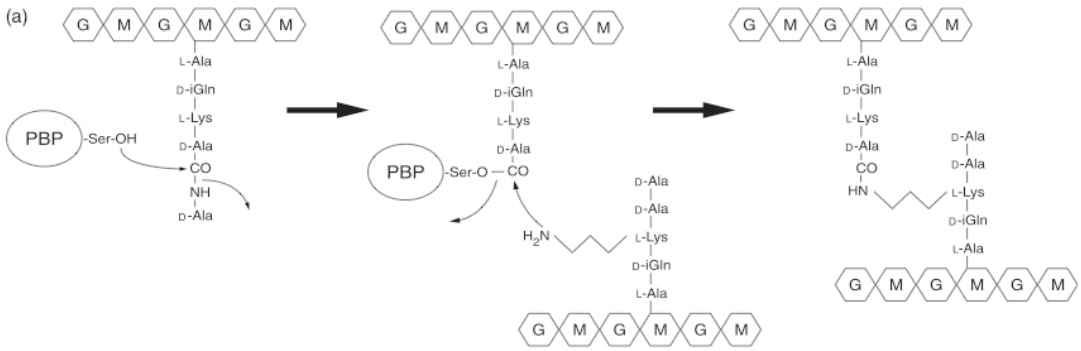


Figure 3.28. The transpeptidation reaction. From [105].

An enzyme domain known as the YkuD domain (formerly known as the ErfK/YbiS/YcfS/YhnG domain) is ubiquitous amongst prokaryotes and appears to be involved in alternative cross-linking of peptidoglycan and in attachment of

proteins to the cell wall. The YkuD family can be recognised by the conserved amino acid sequence ΦGΦHGTX$_{10}$(S/T)XGCΦR(M/L) (where Φ is a hydrophobic amino acid and X is any amino acid) [106]. The first protein containing this domain whose function was elucidated is L,D transpeptidase (Ldt$_{fm}$) from *Enterococcus faecium* [107]. This enzyme catalyses the modification of the bacterial peptidoglycan, acting as an alternate transpeptidase cross-linking peptidoglycan between Lys$_3$ and Lys$_3$. This leads to L,D (also known as 3,3) cross-links, rather than the D,D links created by PBPs (fig. 3.29).

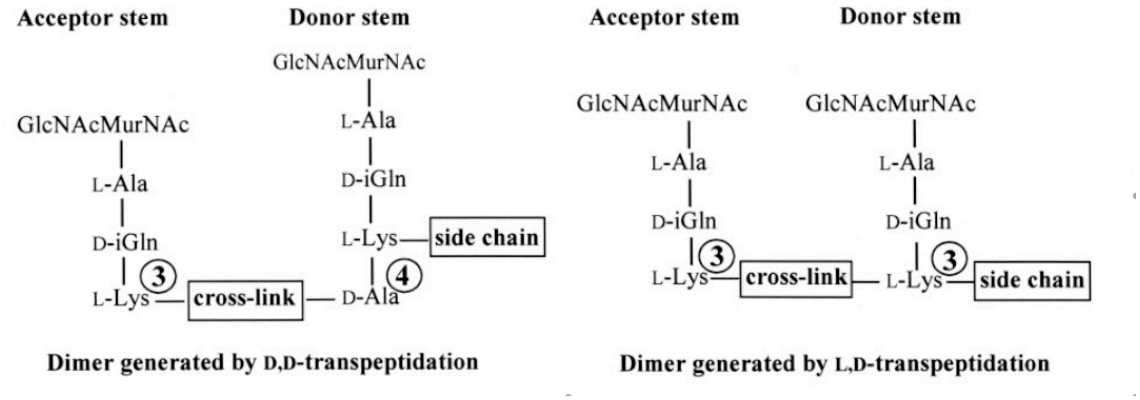


Figure 3.29. Comparison of the two different forms of transpeptidation: D,D-transpeptidation, facilitated by PBPs, and L,D-transpeptidation, as carried out by Ldt enzymes containing a YkuD domain. From [108].

In *E. faecium* Ldt was shown to confer resistance to β-lactam antibiotics as the protein has no affinity for these compounds. Ldt-catalysed remodelling of the peptidoglycan is an alternative pathway that negates the use of PBPs which are often sensitive to β-lactam antibiotics. Further studies have revealed that Ldt proteins exist in other bacterial species, including *E. coli* [109] and *B. subtilis* [108]. During stationary phase, *M. tuberculosis* contains up to 80% L,D cross-

links in its peptidoglycan [110] and recent work has identified the protein Ldt_{Mt2}. Deleting the gene encoding this protein lead to mutants which were small and smooth and had stunted aerial growth, despite the lack of difference in the cell wall constituents [111]. *Streptomyces albus* G has been shown to contain L,D cross-links, but only as a minor percentage of the overall cross-linking in the cell wall [112].

These findings are important to this work as homology searching has shown that *S. coelicolor* possesses six lipoproteins containing the YkuD domain (appendix 1). All are of unknown function, although one of the proteins (SCO4868) also contains a peptidoglycan binding domain. Assuming they are Ldt enzymes, the potential disruption of the processing of these lipoproteins in the *Δlsp* mutant, as seen with CseA, is likely to have a detrimental effect on the cell wall and could explain the extreme sensitivity of the *Δlsp* mutant to bacitracin. As discussed above, bacitracin stops the recycling of UP and therefore prevents any new peptidoglycan precursors being transported out of the cell, ceasing the formation of any new D,D cross-linked peptidoglycan. The L,D transpeptidases are the only group of enzymes capable of modifying the cell wall in the absence of new peptidoglycan precursors [110], and the absence of these lipoproteins in the *Δlsp* mutant would suggest that, in the presence of bacitracin, no peptidoglycan cross-links of any kind are being made, leading to substantial weakening of the cell wall. The small flat growth phenotype seen in the *Δlsp* mutant is also analogous to the growth phenotype seen in the *M. tuberculosis* strain lacking Ldt_{Mt2}. As has been shown, complementing the *Δlsp* mutant restores the correct processing of CseA and SCO3484. It also raises the resistance of the complemented strains to bacitracin, suggesting that the

YkuD domain containing lipoproteins have also been restored and that L,D peptidoglycan cross-linking has resumed, although these lipoproteins were not identified in the membrane proteome of wild-type or complemented *Δlsp* mutants, and the colonies retain their flat phenotype.

Whilst the Ldt homologues in Gram-positive bacteria are involved in remodelling the cell wall, a homologue in the Gram-negative *E. coli* has been shown to have a separate function, catalysing the cross-linking of Braun's Lipoprotein (Lpp) to the periplasmic peptidoglycan. The Ldt enzyme anchors the C-terminus of the Lpp to the peptidoglycan [109], while the N-terminus is secured in the outer membrane by the lipid moiety. If any of the *S. coelicolor* YkuD containing lipoproteins are involved with linking lipoproteins anchored in the cytoplasmic membrane to the cell wall peptidoglycan, then the loss of anchoring may also contribute to the weakened cell wall of the *Δlsp* strain.

Restoring the Lsp enzyme restores resistance to bacitracin, as well as resistance to lysozyme overlay after 0 hours, but not resistance to 15 hour SDS or lysozyme overlay. This suggests that the mycelia immediately post-germination are as healthy as the wild-type, but later develop flaws in their cell envelope.

The inability of the complemented strains to restore both the growth phenotype of the wild-type and the integrity of the cell envelope, but to restore the correct processing of lipoproteins, suggests that at least one secondary mutation has arisen in the *S. coelicolor* genome. If this has occurred spontaneously to compensate for the loss of correct lipoprotein processing it suggests that *lsp* is

an essential gene in *S. coelicolor* which would prove unique amongst the Gram-positive bacteria tested to date. The difficulty in disrupting the *lsp* gene suggests that it is very important to *S. coelicolor*. The genome of this mutant will have to be sequenced, along with that of the wild-type, to ascertain where the potential mutations may have arisen.

As has been mentioned previously, the Lgt enzymes in *S. coelicolor* both contain an extended C-terminal whilst the Lsp protein contains an extended N-terminus. These extensions are not present in the Gram-negatives nor the low GC Gram-positives. In order to assess the roles that these extensions perform, truncations were made of both the Lsp and the Lgt proteins. Given that both extensions are cytoplasmic, and absent from *E. coli*, it was thought unlikely that they performed an enzymatic function, but more likely a structural one. Removing the C-terminal extension from either of the Lgt enzymes had no effect on the processing of CseA (data not shown). This result is not unexpected, given that deletion of either protein in its entirety has no effect on lipoprotein processing, as discussed. To ascertain whether the N-terminal extension of Lsp was necessary for enzyme activity a series of truncation mutants were created, and the results are shown in figure 3.24. The first 10 amino acids of Lsp appear to be unnecessary for function as properly processed CseA is detectable. Confusingly, deleting the first 20 amino acids renders Lsp inactive, with CseA undetectable, yet removing a further 10 residues allows CseA to be detected again, albeit at a larger size, presumably with its signal sequence still attached. Truncating the N-terminus to the same size as *E. coli* also leads to the loss of CseA.

These results suggest that the N-terminal extension of Lsp is required for enzyme function. What cannot be proved however, is whether or not the extension is required for enzyme stability. The inactivity of the anti-Lsp antibody means we were unable to show whether the enzyme is present *in situ* but functionally inactive in the N20 and N40 truncation mutants or whether it is not expressed. The lack of an anti-Lsp antibody also meant that the planned experiment to directly visualise Lgt1/Lgt2-Lsp interaction was not possible. The plan was to chemically link proteins in the cell membrane fraction, then probe the fraction with the anti-Lsp antibody. If there was a direct Lgt-Lsp interaction, the band present would be larger than the size of the Lsp protein. This could have been carried out in both the *Δlgt1* and *Δlgt2* strain to see whether either, or both proteins interacted with Lsp. This interaction seems plausible, as passive diffusion of an immature lipoprotein between processing enzymes seems energetically wasteful. Attempts to fuse Lgt1, Lgt2 and Lsp to the fluorescent reporter protein eGFP, to visualise their position in the bacterial hyphae also proved fruitless, presumably due to the low level of their production (data not shown). It was postulated that these proteins might be localised to the hyphal tip, given that this is where the Tat translocation machinery is located (D. Widdick and T. Palmer unpublished).

The D148A/D177A (DD) point mutant was created to see whether the loss of Lsp acts as a 'master switch', halting the transcription/translation of non-essential lipoproteins, to prevent the cell membrane being damaged by having too many embedded signal sequences. The thought was that the presence of the inactive DD enzyme in the cell membrane might mean that the switch remained 'on' and a higher molecular weight band comprising of CseA with its

intact signal peptide would be detectable. This does not appear to be the case as CseA is clearly not detectable in the membrane although the lack the anti-Lsp antibody means we were unable to show whether the enzyme had been expressed or not.

In summary, the work in this chapter describes the lipoprotein biosynthetic pathway in the Gram-positive bacterium *S. coelicolor*. This pathway is unusual amongst bacteria, as it encodes two Lgt homologues, and two potential Lnt homologues. It appears that the Lgt proteins are functionally redundant because both seem capable of lipidating a lipoprotein. This work suggests, but does not definitively prove, that the *lsp* enzyme is essential in *S. coelicolor*. Deleting *lsp* does prove seriously detrimental to the bacterium, altering both the processing of lipoproteins and the growth of the cell. These data suggest that Lsp is an essential enzyme in *S. coelicolor* but the bacterium has gained one or more spontaneous suppressor mutations to rescue the strain. Also the inability to isolate a double *lgt* mutant suggests that the lipidation step of the lipoprotein biosynthetic pathway is essential for survival. Future work will be aimed at mapping these suppressor mutations and this work is ongoing. Deletion strains of each of the six YkuD lipoproteins will also have to be made to see whether they contribute to the remodelling of *S. coelicolor* peptidoglycan and lead to any of the growth phenotype seen in the *lsp* deletion strain.

Chapter 4 - The lipoprotein biosynthetic pathway of *S. scabies.*

4.1 Introduction

The previous chapter investigated the lipoprotein biosynthetic pathway in the model streptomycete *S. coelicolor.* The work showed strong evidence as to the essentiality of the enzyme Lsp, but left some unanswered questions, notably as to the phenotype of a mutant lacking Lgt, and the role played by the Lnt homologues encoded by streptomycete genomes. This chapter attempts to rectify this by extending the study of the same pathway to the plant pathogenic bacteria *S. scabies*.

Whilst the majority of well studied pathogenic bacteria are Gram-negative, the Gram-positive actinomycetes do include a number of animal pathogens (e.g. *Rhodococcus equi, M. tuberculosis*) and plant pathogens (e.g. *Leifsonia xyli, Clavibacter michiganensis*) [113]. *S. scabies* is a plant pathogen, predominantly of potatoes (*Solanum tuberosum*). *S. scabies* is one of only a few characterised pathogenic *Streptomyces* species and has been identified in Europe, Asia, Africa as well as North and South America [74]. The main phenotype of an infected potato plant is the presence of lesions or scabs on the surface of the potato tuber (fig. 4.1). These scabs are usually based on the surface of the plant and have a round appearance but can overlap to cover a significant proportion of the surface of the tuber [67]. These lesions are often only identified at the time of harvesting [74] reducing the marketing value of the potato. As a direct result this disease is of great economic importance to farmers worldwide.

Figure 4.1. Typical symptoms of a scabies infected potato [67].

Like *S. coelicolor*, *S. scabies* is a soil dwelling saprophytic bacterium, and has a very similar developmental cycle. However *S. scabies* has gained a set of genes which has enabled it to colonise a variety of plants. Whilst the bacteria usually infects tap root crops such as potato or radish, the bacterium appears to be neither host, nor tissue, specific and will infect seedlings of monocotyledonous or dicotyledonous plants [67]. Other plant pathogens, such as members of the *Xanthomonas* or the *Pseudomonas* genera, attack the leaves or stems of plants. There is an advantage in doing this, as these structures contain an abundance of natural openings (such as the stomata) which the invading pathogen can utilise. In contrast *S. scabies*, as a soil dwelling bacterium, infects the roots or tubers of plants and causes scabs on the latter. These plant structures have a lower amount of openings [77], and the bacterium has developed a specialised way of entering the plant. The main weapons for the pathogenicity of *S. scabies* are two compounds, thaxtomin A, and Nec1.

4.1.1 The *S. scabies* pathogenicity island.

Bacterial pathogens have a very specialised niche, and the ability to cause disease is a highly evolved phenotype. The genes responsible for this phenotype can be moved between bacteria, either individually or as a group, with the ability to utilise another source of nutrients potentially conferring extra fitness upon the donor strain [114]. Virulence genes are often arranged in a bacterial genome in a discrete cluster, known as a pathogenicity island (PAI). Pathogenicity islands have been discovered in a wide variety of Gram-positive pathogens including *Staphylococcus aureus* and *Bacillus anthracis*. Given their ability to move from one species to another, the genetic background of a PAI often differs from the species of bacteria it is found in and the G+C content of a PAI is frequently different from the genome surrounding it [75]. The movement of PAIs from one species to another is often described as 'Horizontal Gene Transfer', but in the strictest sense, it should be described as 'Lateral Gene Transfer' as the transfer is intergenetic, rather than interdomain [115].

S. scabies has acquired a set of genes which has allowed it to become a very successful plant pathogen. These genes are arranged into a PAI, which at 325 kb is currently the largest known bacterial PAI. At 54%, the G+C content of the PAI is considerably lower than the rest of the *S. scabies* genome, which is 71.45% [76] (www.sanger.ac.uk/projects/S_scabies). The PAI contains multiple genes responsible for pathogenicity, including those needed for the biosynthesis of thaxtomin A and the *nec1* gene, which has the lowest G+C content of any currently sequenced streptomycete gene [113]. The DNA in the areas around these genes contains multiple transposons, and insertion elements (IS). This is a typical characteristic of a PAI, as are the presence of truncated ORFs [77],

presumably a relic of the transfer of genes from one species to another.

S. scabies is related to the other scab forming streptomycetes, *S. turgidiscabies* and *S. acidscabies* but the PAI responsible for the disease phenotype was initially passed from *S. scabies* to the other varieties [76]. These genes offered the recipients a selective advantage and were maintained in the newer pathogens. Simply moving the PAI from one species of streptomycete to another does not confer the ability to infect plants however. *S. coelicolor* cannot be made a pathogen simply by artificially introducing the PAI from *S. turgidiscabies* [76]. The reasons for this are unclear, but what is known is that the PAI integrates into an 11bp region of the *S. coelicolor* gene *bacA* (SCO1326), which shows similarity to an *E. coli* undecaprenyl pyrophosphate phosphatase, a gene involved in resistance to the antibiotic bacitracin. The insertion of the PAI does not affect the resistance of *S. coelicolor* to bacitracin. The DNA flanking *bacA*, in *S. coelicolor* has a low G+C content, suggesting that this site may have been used for lateral gene transfer previously [67]. Whilst it may be impossible to artificially 'weaponise' *S. coelicolor*, this is not true of all *Streptomyces* species. Transfer of the *S. turgidiscabies* PAI into *S. diastatochromogenes* allowed the recipient to colonise potato tubers [67]. The reasons for the selectivity of recipient strains is unclear. As mentioned earlier, there are few pathogenic streptomycetes and it may be that there is some underlying genetic background that prevents most species from accepting, or retaining, a PAI. Perhaps another element is needed in order to activate the pathogenicity genes, it has been suggested that perhaps some PAI need concomitant 'metabolic islands' (MAIs) for optimal pathogen fitness when invading a host [115]. These may be absent from one recipient strain, but

present in another.

4.1.2. Thaxtomin A.

The most well described virulence factor from *S. scabies* is thaxtomin A. This molecule is a nitrated dipeptide phytotoxin, capable of necrosing excised potato tissue and causing scabs on immature potatoes [67]. Thaxtomin A is produced by *S. scabies*, *S. turgidiscabies* and *S. acidscabies,* with other variations produced by other *Streptomycetes* (e.g. thaxtomin C, produced by *S. ipomoeae*) and is the primary pathogenicity determinant in *Streptomyces* [116]. Thaxtomin A appears to have a novel mechanism of action, whereby it inhibits cellulose biosynthesis and induces hypertrophy in plant cells [114] and it can be purified from both infected potatoes and growth media, with nanomolar concentrations able to induce irregular, binucleate, cells in onion root tips [67]. As a virulence factor, thaxtomin A exhibits no antimicrobial activity and is purely plant specific [113]. This result suggests that it may be involved in cytokinesis, and it is tempting to think that this inhibition of cell wall synthesis would make cells weaker, and therefore more susceptible to the invading *S. scabies* [77].

The pathway for the production of thaxtomin A requires multiple genes, and is summarised in figure 4.2, it requires conserved non-ribosomal peptidase synthases, a P450 monooxygenase and Nitric Oxide synthase (NOS) genes [113]. Expression of these genes is regulated by TxtR a member of the AraC/XylS family of regulators. TxtR binds a disaccharide, cellubiose, which is a component of cellulose. Disruption of the *txtR* gene leads to a cessation of thaxtomin A production. However, as well as cellubiose, the plant polymer suberin is also required to induce thaxtomin production [117].

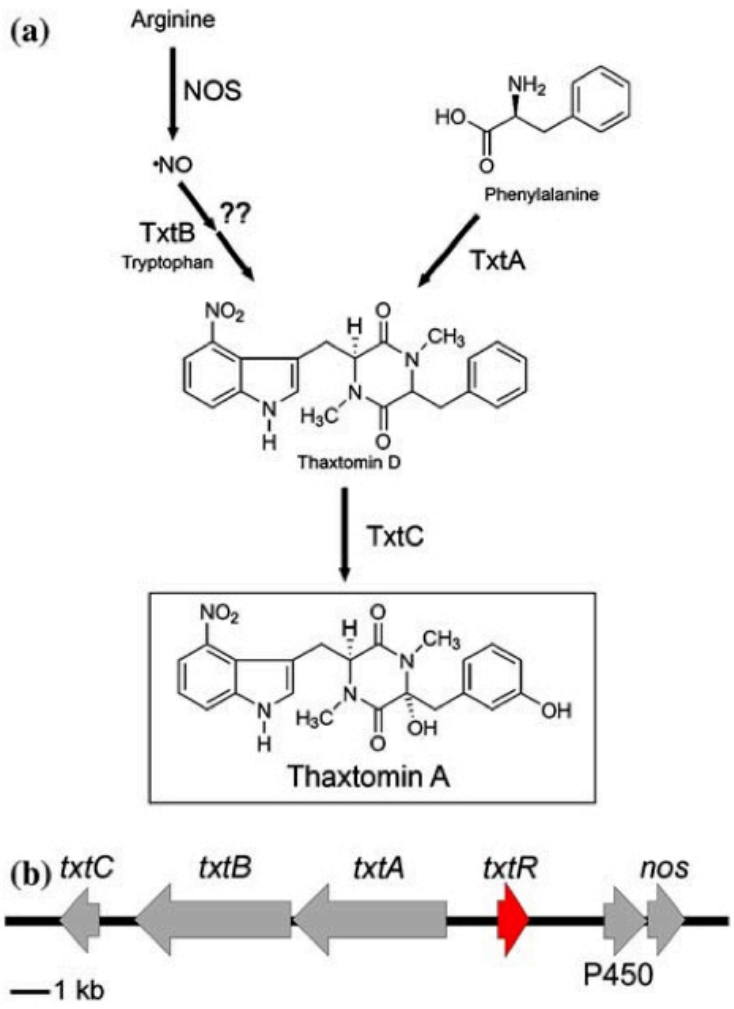


Figure 4.2. a) The biosynthetic pathway for the production of thaxtomin A. b) The *txt* gene cluster. From [113].

The conserved non-ribosomal peptidase synthases (TxtA and TxtB) are required for production of the *N*-methylated cyclic peptide backbone of the molecule, whilst post-cyclisation hydroxylation steps are carried out by the P450 monooxygenase (TxtC) [113]. Perhaps the most interesting part of the pathway is related to the *nos* gene. Nitric oxide (NO) is an intracellular signalling molecule in mammals, involved in a variety of situations, including regulation of blood pressure and the immune system [118]. Recently NOS-like proteins have been discovered in several Gram-positive bacteria, including *B. subtilis,* that show homology to the mammalian NOS proteins [119]. The function of these

bacterial NOS (bNOS) proteins is unknown, as unlike *S. scabies* they do not produce thaxtomin [113]. In mammals NO is produced by the oxidation of L-arginine to L-citrulline and NO, via an enzyme bound intermediate, catalysed by three isoenzymes [118, 119]. Like the NOS proteins from other bacteria, *S. scabies* share the same key residues as those found in the mammalian enzymes and, typical of bacterial NOS, they lack the mammalian N-terminal Zn^{2+} and carboxy-terminal flavoprotein reductase domains. However, unlike other bacteria the *S. scabies* bNOS has an elongated N-terminus [118]. The *nos* gene in *S. scabies* is necessary for thaxtomin A production, as it is in the other scab causing *Streptomyces* species, as they are highly conserved and their deletion results in a drastic drop in thaxtomin A production and a loss of virulence [118].

4.1.3 Nec1

Alongside thaxtomin, Nec1 is the other significant virulence factor found in *S. scabies*. As has been previously mentioned, the *nec1* gene has a particularly low G+C content, and unusually, although it is conserved amongst scab causing bacteria, it shows no homology to any other genes or gene products currently sequenced [75]. Thus, it can be thought of as a novel virulence factor [67]. Usefully, as *nec1* is unique amongst the scab causing streptomycetes, it can be used as a marker for the rapid detection of *S. scabies* and related species by conventional or real-time PCR, regardless of whether scab symptoms are present on a potato or not [114]**.** The Nec1 protein is secreted, typically during the early log phase, after approximately 20 hours [67, 75]. Thaxtomin is produced after approximately 48 hours, suggesting that Nec1 is involved with early interaction with the plant to be colonised. Thaxtomin production is

unaffected by Nec1 production [76], but it does appear that a combination of the two is required for a pathogen to be truly effective. *Arabidopsis thaliana* plants infected with a *Δnec1* strain of *S. turgidiscabies*, show mild symptoms of root damage, but recover, whilst wild-type infected strains invariably die [75]. The reasons for the need for the combination of the two virulence factors is unclear. Certainly the gap between their secretion is likely to be important, and it may be that Nec1 has a role in suppressing plant cell defences. Thaxtomin produces a rapid (within one minute of inoculation) Ca^{2+} influx , followed by a net efflux of H^{+} in root cells, part of the plant cell defence [67], potentially Nec1 suppresses the cell defences induced by the acidification of the cell wall.

4.1.4. Mechanisms of *S. scabies* infection

Whilst the tools used by *S. scabies* to infect potatoes are known and (at least partially) understood, the exact method by which the pathogen interacts with its host are not. As mentioned earlier, root systems in plants offer unique challenges for invading pathogens not seen in other plant structures, mainly due to their lack of natural openings, such as the multitudinous stomata found in leaves [77]. However as a place to live, the soil does offer some distinct advantages when compared to a leaf, as it has a relatively stable temperature day and night, and young roots are an excellent source of nutrients [77].

When *S. scabies* hyphae come into contact with a potato, specialised mechanisms must be employed by the bacteria in order to penetrate the tuber. *S. ipomoea*, a pathogen of sweet potato (*Ipomoea batatas*), appears to grow specialised hyphae that branch off directly into the sweet potato. These lateral hyphae are able to both pierce, and grow within, the plant cell wall [78]. More

recent work has showed the same structures in *S. scabies* growing on potato tubers (fig. 4.3) [77]. There is some suggestion that there is degradation of the cell wall where these hyphae penetrate, which would presumably be enzymatic [78]. It appears that this method of infection differs from that of fungal plant pathogens, as there are no appressoria or infection cushions to be seen.

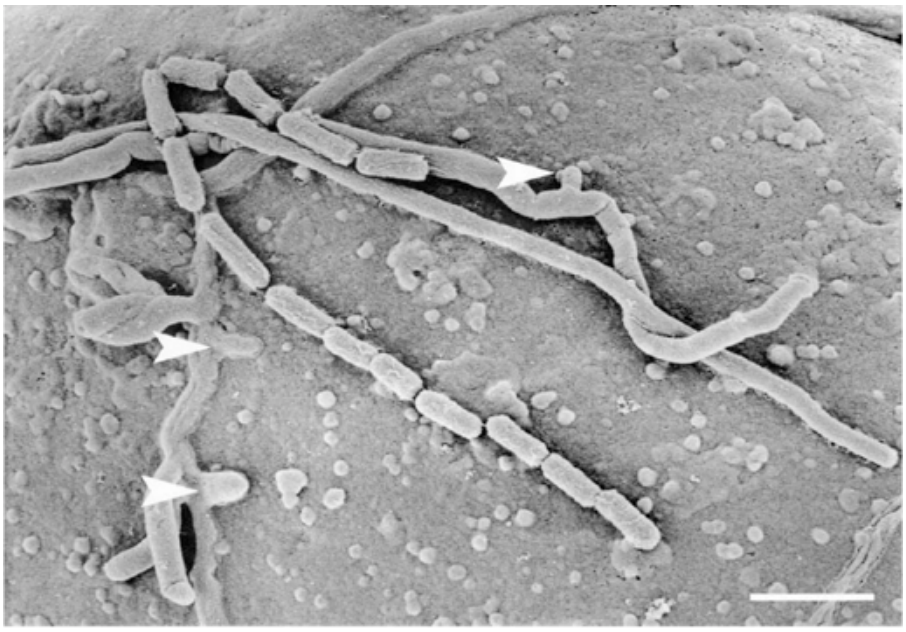


Figure 4.3. SEM image of *S. scabies* hyphae growing on a potato tuber. Infection hyphae are shown penetrating the surface of the tubers (white arrows). Scale bar = 2µm. From [77].

Whatever the mechanism is for breaching the host cell, it is clear that *S. scabies* does significant damage to the root system of the potato plant. A recent paper has used Computed Tomograph (CT) scanning to investigate the damage caused by scabies to potato plant roots in a non-invasive way, *in situ*. The results of this study showed that a potato plant growing in *S. scabies* infected sand had a significantly less complex root system than a control plant after four weeks of growth, and the growth of the infected plant was reduced after this time [74].

Once the bacterium is growing within the host potato, the next stage in the infection process is to secrete the relevant pathogenicity factors. In Gram-negative plant pathogens, such as *Pseudomonas syringae*, this process has been studied in detail, with these bacteria using well characterised methods of neutralising plant host defences based around the Type III secretion system (T3SS). This complex of approximately 30 proteins forms a needle like structure allows pathogens to directly inject proteins which disrupt the hosts defence signalling into a host cell. These are known as Type III secreted effectors [120]. Gram-positive bacteria are completely lacking in T3SS, and the exact methods by which pathogenicity factors are exported is unknown, other than the Tat or Sec systems, and the specialised ESAT-6 system, also known as type VII secretion systems, discovered in *M. tuberculosis* [121]. Large numbers of *S. scabies* lipoproteins (~20%) are translocated by the Tat system [122], and a recent study has shown that the loss of Tat leads to an avirulent phenotype in *S. scabies* [80]. Seven Tat substrate virulence factors were identified, and individual mutants in each coding gene showed reduced virulence. One of these proteins, putative spermidine/putrescine transporter peptide binding protein (SCAB81041) has been identified as a lipoprotein [122].

4.2 Aims

The aims of this chapter are to analyse the lipoprotein biosynthetic pathway of *S. scabies*. Unlike *S. coelicolor*, *S. scabies* only contains one potential *lgt* gene (*SCAB68531*) and, like most bacteria, one potential *lsp* gene (*SCAB68121*). Like *S. coelicolor,* S*. scabies* contains two putative *lnt* genes, named in this study as *lnt1* (*SCAB83111*) and *lnt2* (*SCAB76621*), and the *S. scabies* cosmids containing these genes were provided by Professor Rose Loria (Cornell).

Mutants in each of the lipoprotein processing genes will be examined, both macro and microscopically, as will the effects these mutations have on the processing of lipoproteins. This work with *S. scabies* will allow us to confirm and extend our findings on lipoprotein biogenesis in *Streptomyces* gained using *S. coelicolor* as a model (chapter 3). Additionally, given the lack of research on the role of lipoproteins in plant infection, each mutant will be assessed for virulence in both potato tubers and whole plants, and compared to a wild-type *S. scabies* infection. As discussed in chapter 1, disruption of the lipoprotein biosynthetic genes can attenuate bacterial animal pathogens, or make them hypervirulent. Given the obvious differences between plant and animal defence responses, it will be interesting to see the role lipoproteins play in bacteria-host interaction.

4.3 Results

4.3.1 Identification and comparison of the lipoprotein biosynthetic enzymes.

The lipoprotein biosynthetic enzymes from *S. scabies* were initially identified by BLASTP analysis, and the primary sequence of each enzyme was aligned with the equivalent primary sequences from: *S. coelicolor* (Sco), *S. avermitilis* (Sav), *S. griseus* (Sgr), *M. tuberculosis* (Mtb), *M. smegmatis* (Msm) *B. subtilis* (Bsu) and *E. coli* (Eco).

4.3.1.1 Lipoprotein diacyglycerol transferase (Lgt).

The Lgt enzyme from *S. scabies* shares a high amount of identity (≥74%) to all the *Streptomyces* Lgt enzymes, with the exception of ScoLgt2 (60%). The identity to the *M. tuberculosis* and *M. smegmatis* Lgt enzymes is lower at 53%

and 57% respectively, whilst the lowest identity is to BsuLgt and EcoLgt (31% and 26% respectively) (fig. 3.1).

The essential *E. coli* His-103 is absent from *S. scabies* where it is a tryptophan, although both His-196, and Tyr-235, needed for *E. coli* function are conserved [17], as in the other actinomycetes tested. Also present is the elongated C-terminus discussed in chapter 3.

4.3.1.2 Lipoprotein signal peptidase (Lsp).

S. scabies Lsp shares a high level of identity to the enzymes from the other *Streptomyces* species (≥74%), and a low identity to the other species tested (≤44%). As with *S. coelicolor* it conserves the NXXD and FNXAD residues necessary for catalysis in *B. subtilis* [19, 28], as well as the Asp-15 needed for enzyme stability [26]. Also present is the extended N-terminus, characteristic of the actinomycete enzymes (fig. 3.4).

4.3.1.3 Lipoprotein N-acyl transferase (Lnt).

The *S. scabies* Lnt1 and Lnt2 enzymes share a low identity to each other (31%), but both also show low levels of identity to *E. coli* Lnt (22% and 28%). Both enzymes conserve the three residues (E267, K355, C387) thought to form the enzyme's catalytic triad in *E. coli* [30]. The *S. coelicolor* Lnt enzymes also show a low level of identity to each other (29%) and *E. coli* Lnt (~25%), but when directly compared, ScabLnt1 and ScoLnt1 share a high level of identity (86%), as do ScabLnt2 and ScoLnt2 (72%).

4.3.2 Disrupting the *S. scabies* biosynthetic genes.

Cosmids containing the *S. scabies lgt* (*SCAB68531*), *lsp* (*SCAB68121*), *lnt1* (*SCAB83111*) and *lnt2* (*SCAB76621*) were obtained from Cornell University. Each gene was replaced with an antibiotic resistance cassette. Apramycin was used to replace each of the genes individually, to form cosmids Scab 139 *lgt::apr*, Scab 45 *lsp::apr*, Scab 351 *lnt1::apr* and Scab 2255 *lnt2::apr*. A hygromycin marked *lnt2* deletion strain was also made, cosmid Scab 2255 *lnt2::hyg*. in order to construct an *lnt1*::apr, *lnt2::hyg* double mutant (D. Widdick, unpublished).

The mutagenised cosmids were transferred into *S. scabies* by conjugation, to form strains BJT1040, BJT1044, BJT1047, BJT1048, BJT1049 and BJT1050 (see table 2.2) as described in chapter 2, and potential mutants were selected on apramycin or hygromycin. The genomic DNA from potential mutants was checked by PCR. An important observation was that the isolation of an *Δlsp* mutant was much easier than it was in *S. coelicolor* (see previous chapter).

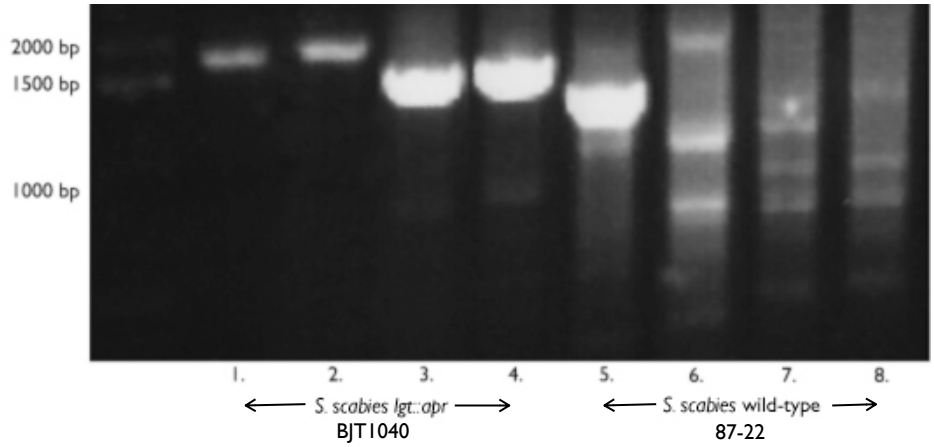


Figure 4.4. PCR of *S. scabies* genomic DNA showing disruption of the *lgt* gene with the *apr* cassette. Lanes 1->4 show the mutant, lanes 5->8 are the wild type control. Primer combinations in each lane (expected band sizes shown in brackets): 1+5 = ScabLgt comp forward/ScabLgt test reverse (1642 bp/1493 bp). Lanes 2 + 6 = ScabLgt comp forward/P2 (1602 bp/no band). Lanes 3 + 7 = ScabLgt Test reverse/P1 (1439 bp/no band). Lanes 4 + 8 = P1/P2 (1423 bp/no band).

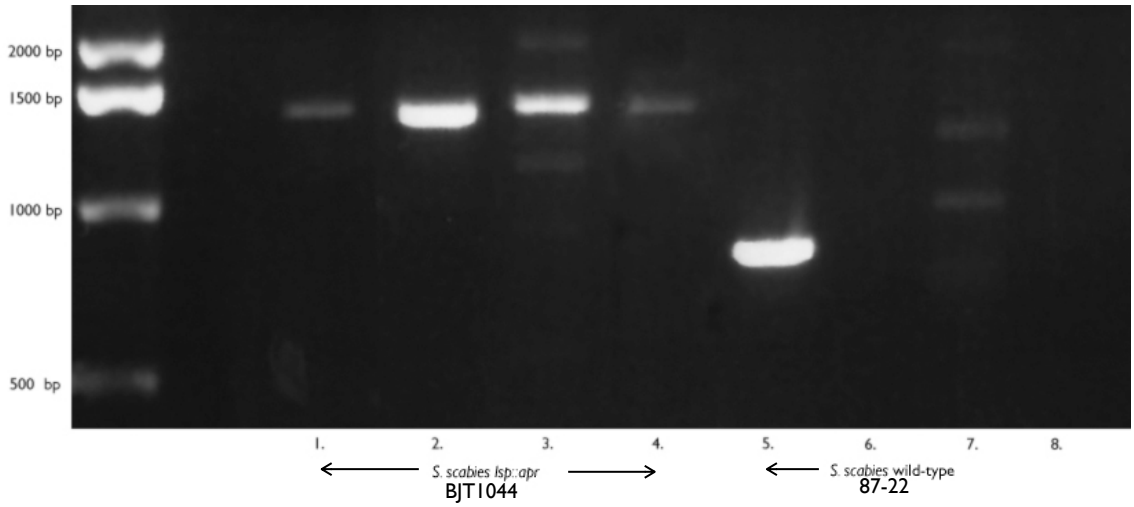


Figure 4.5. PCR of *S. scabies* genomic DNA showing disruption of the *lsp* gene with the *apr* cassette. Lanes 1->4 show the mutant, lanes 5->8 are the wild type control. Primer combinations in each lane (expected band sizes shown in brackets): 1+5 = ScabLsp Test forward/reverse (1437 bp/651 bp). Lanes 2 + 6 = ScabLsp test Forward/ P2 (1418 bp / no band). Lane 3 + 7 = ScabLsp Test reverse/P1 (1419 bp /no band). Lanes 4 + 8 = P1/P2 (1423 bp/no band).

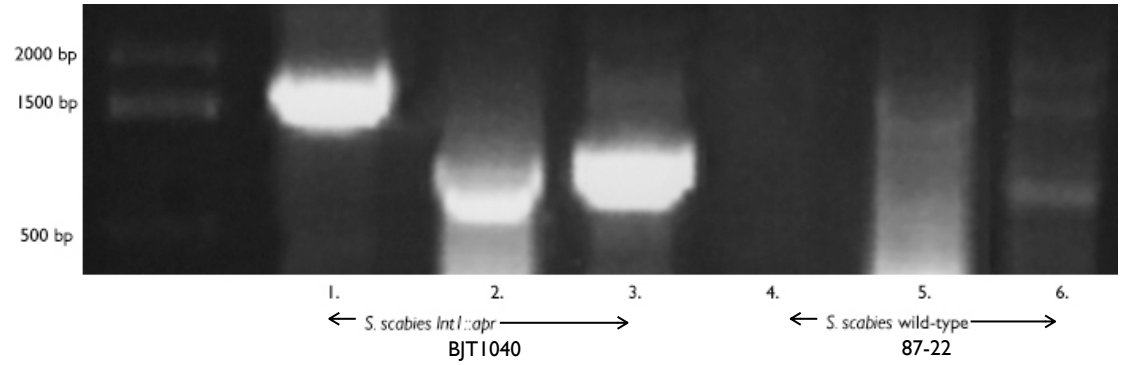


Figure 4.6. PCR of *S. scabies* genomic DNA showing disruption of the *lnt1* gene with the *apr* cassette. Lanes 1->3 show the mutant, whilst lanes 4->6 are the wild-type control. Primer combinations in each lane (expected band sizes shown in brackets): 1 + 4 = ScabLnt1 Complementation Forward/P2 (1602 bp/no band). 2 + 5 = ScabLnt1 Test reverse/P1 (1440 bp/no band). 3 + 6 = P1/P2 (1423 bp/no band).

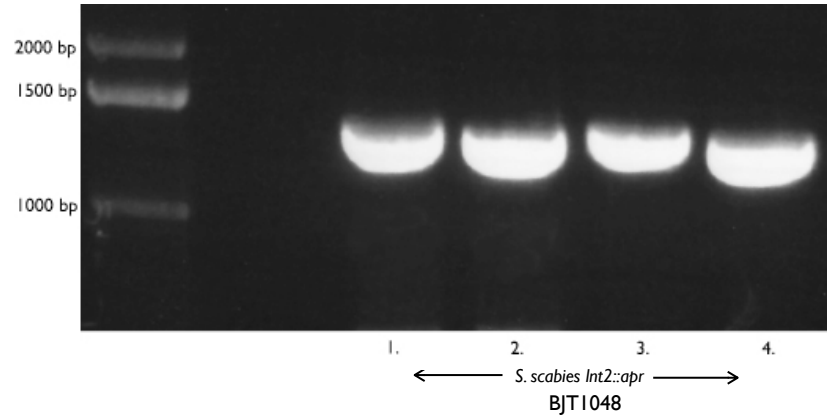


Figure 4.7. PCR of *S. scabies* genomic DNA showing disruption of the *lnt2* gene with the *apr* cassette. Primer combinations in each lane (expected band sizes shown in brackets): Lane 1 = ScabLnt2 Test Forward/Reverse (1574 bp). Lane 2. = can Lnt2 Test Forward/P1 (1420 bp). Lane 3. = ScabLnt2 Test reverse/P1 (1425 bp). Lane 4. = P1/P2 (1423 bp).

4.3.2.1 Confirmation of *nec1.*

Before any work was carried out on the strains created, a final check was carried out to ensure that all the strains created were *S. scabies* and not *S.*

coelicolor contamination had occurred. A simple test using primers specific to the *nec1* gene [114] (table 2.4) was used. This gene is present in *S. scabies* and absent in *S. coelicolor*. The results shown in figure 4.8 show that a band representing both *nec1* (666bp) is present in the genomic DNA of each of the *S. scabies* strains, and absent from the negative control, *S. coelicolor* genomic DNA.

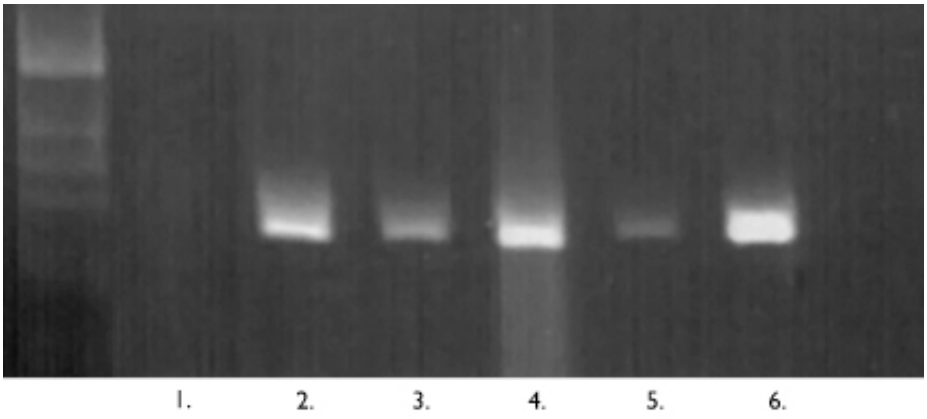


Figure 4.8. PCR of *S. scabies* genomic DNA to confirm presence of *nec1*. Lane 1. = Wild-type *S. coelicolor* (M145). Lane 2. = Wild-type *S. scabies* (87-22). Lane 3. = *S. scabies ΔIgt::apr* (BJT1040). Lane 4 = *S. scabies ΔIsp::apr* (BJT1044). Lane 5. = *S. scabies ΔInt1::apr* (BJT1047). Lane 6. = *S. scabies ΔInt2::apr* (BJT1048).

4.3.2.2 Development and colony morphology.

Once each of the null mutants, and the *ΔInt* double mutant were created, they were viewed under the light microscope (fig. 4.9). The typical morphology of a wild-type colony, is a flat circular disc, about 3-4 mm in diameter, after 5 nights growth, with a raised section in the centre. The shape is akin to that of a fried egg. The colonies were grey, indicating the presence of mature spores, but typically contained a white outline, presumably representing areas of undifferentiated aerial hyphae. The *ΔIgt* mutant (BJT1040) had a very different growth phenotype, growing slower than the wild-type, sporulating poorly, as

suggested by a much lighter shade of grey. Often the size of an *Δlgt* colony was smaller than that of the wild-type, but not consistently. This strain also lacked the raised area in the centre of the colony and the surface was frequently pock marked. The *Δlsp* strain (BJT1044) had a similar, although more pronounced phenotype than the *Δlgt* mutant. The strain sporulated very poorly, which lead to an almost entirely white colony, and was smaller than the wild-type, usually between 2-3 mm. In contrast to the 'fried egg' growth pattern seen in the wild-type, the *Δlsp* mutant grew in a dome structure, frequently pock-marked, like the *Δlgt* mutant. These holes in the surface of the colony were often filled with a clear liquid, which appeared to be covered in a 'skin' of come kind (see section 4.3.2.3 below). The various *Δlnt* mutants all display a more subtle growth phenotype than either the *Δlgt* or *Δlsp* mutants, but nevertheless they do look different to the wild-type (fig. 4.9). All sporulate to a good level, giving the strains an overall grey appearance, and each retains the white border seen in the wild-type. Although the *Δlnt1* (BJT1047) and *Δlnt2* (BJT1048) mutant colonies are of a similar size to the wild-type, the edges are much more ragged and the colonies are irregularly shaped. Frequently the raised section in the centre of the colony is missing from these mutants. The *Δlnt* double mutant (BJT1050) contains both the raised area from the wild-type, and the irregular colony shape of the *Δlnt* single mutants. The colony size is slightly larger than the wild-type, although the growth seems more dispersed, and there are often patches where no aerial hyphae/spore chains are present.

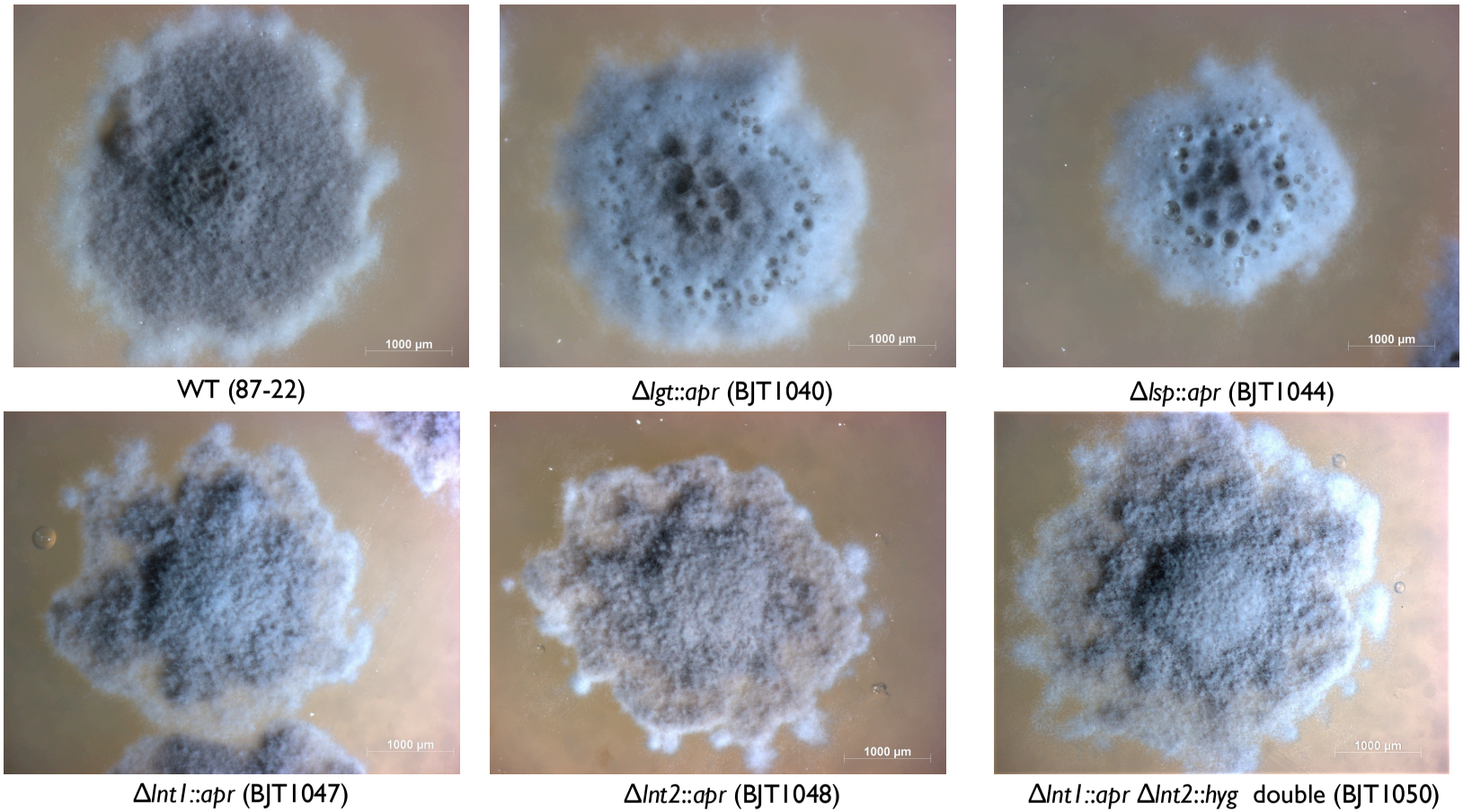


Figure 4.9. Light microscopy of each of the *S. scabies* strains. Strain names shown in brackets.

Viewing the strains grown for 5 nights under SEM showed that the *Δlsp* mutant (BJT1044) contains fewer hyphae and thus fewer spores than the other strains, consistent with the overall whiteness of the colony seen under light microscopy. The wide-angle view (fig. 4.10) shows that all of the other strains contain a dense mass of hyphae, with the wild-type showing the largest amount of spiral spore chains. Although the *Δlgt* (BJT1040) and *Δlsp* mutants share a similar growth phenotype, the *Δlgt* mutant is showing a higher amount of spore chains, although not nearly as many as the wild-type. Time constraints prevented there being any images taken of the *Δlnt* double mutant. Given its lack of an obvious growth phenotype, this was not considered an issue. Neither of the *Δlnt* mutants (BJT1047, BJT1048) display an obvious phenotype when viewed under SEM.

The SEM closeups of the hyphae revealed that *S. scabies* spores appear to be a different shape to those seen in *S. coelicolor*, with the former being much more cylindrical. This is not overly surprising given the variety of spore shapes found in the streptomycetes [123]. This highly magnified view confirmed the observation that the *Δlsp* mutant is sporulating poorly. Finding spore chains from each of the other strains to image was relatively easy, whilst it proved very difficult for the *Δlsp* mutant. The immature spore chain shown in figure 4.11, was the closest identified, and this shows a further phenotype for the *Δlsp* mutant. As shown, the immature spores are irregularly sized compared to the wild-type. This was seen frequently in the *Δlsp* mutant, but was also seen in the *Δlgt* mutant, although with less regularity.

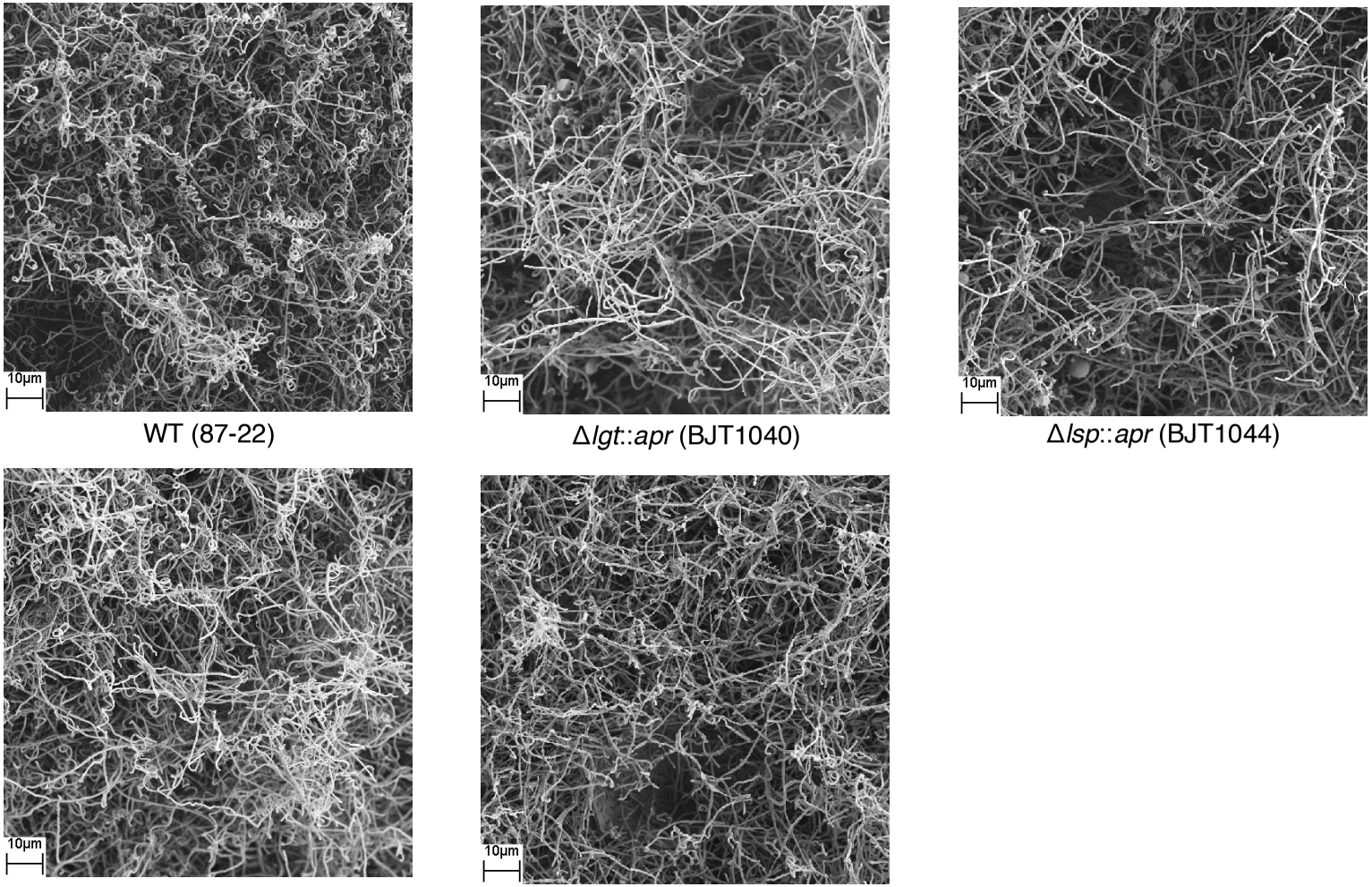
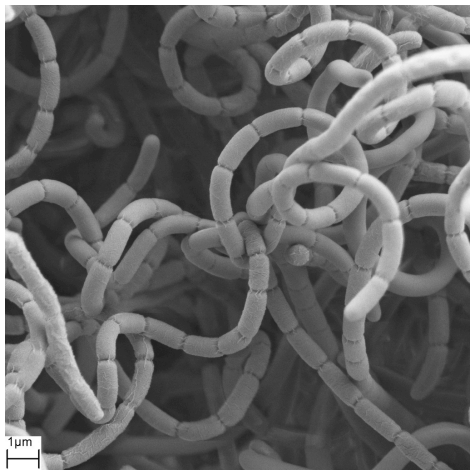
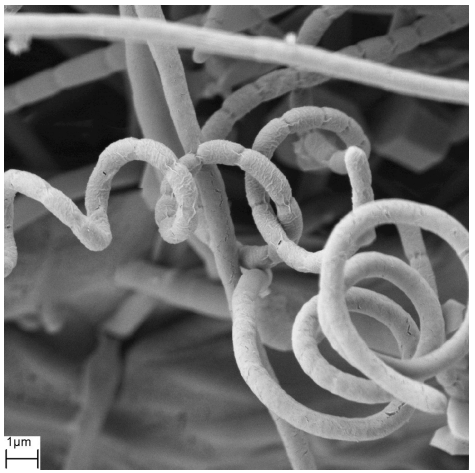


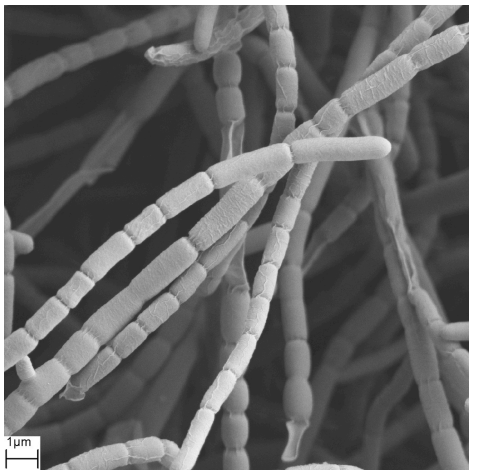
Figure 4.10. Wide angle scanning electron microscopy of each of the strains grown for 5 nights on IMA media. Strain names are shown in brackets.



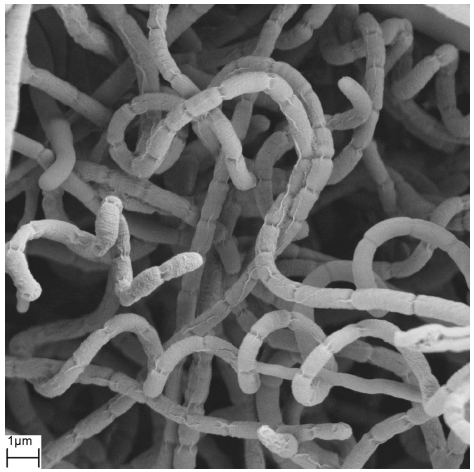
WT (87-22)



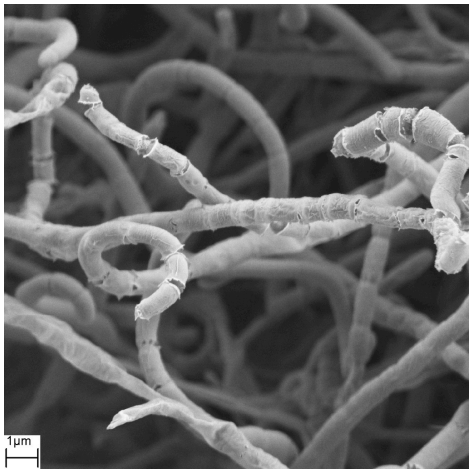
Δ*lgt*::*apr* (BJT1040)



Δ*lsp*::*apr* (BJT1044)



Δ*lnt1*::*apr* (BJT1047)

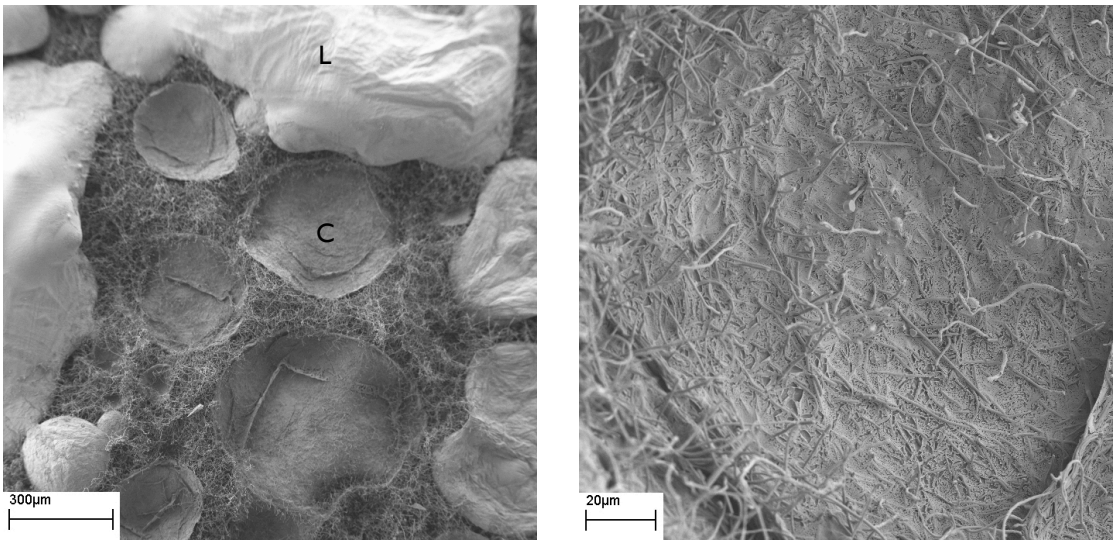


Δ*lnt2*::*apr* (BJT1048)

Figure 4.11. Zoomed scanning electron microscopy images of the spore chains of each of the strains grown for 5 nights on IMA media. Strain names are shown in brackets.

4.3.2.3 Liquid on the surface of the *Δlsp* mutant

As discussed in section 4.3.2.2 above, light microscopy of the *Δlsp* strain revealed the presence of liquid droplets unseen in the wild-type. These appeared to have a 'skin' and were assessed under SEM (fig. 4.12). The wide angle image shows that these droplets do indeed possess a covering of some sort, as a number of crenellations can be seen on their surface. Where this liquid has been displaced, a crater is left in the surface of the colony. A zoomed in view of these craters reveals that they are formed by the remains of the matrix surrounding the liquid and are frequently punctured by mycelia. The identity of this liquid is unknown, although confluent lawns of both wild-type and the *Δlsp* mutant have been washed with sterile water in order to remove any surface associated liquids and consequently discover their identity. This work is currently ongoing.



Wide Zoomed

Figure 4.12. Scanning electron microscopy images of the liquid droplets found on the surface of the *Δlsp* mutant (BJT1044), marked L, and the craters they leave in the surface of the colony, marked C.

4.3.3 Effect of gene disruption on lipoprotein processing.

4.3.3.1 Western blots.

As with *S. coelicolor,* each *S. scabies* mutant was was grown for 15 hours in liquid culture, before being fractionated by ultracentrifugation as described in chapter 2. Both cytoplasmic and membrane fractions were normalised and run on a 15% SDS-PAGE gel. Two model lipoproteins were used and were produced in each *S. scabies* strain with C-terminal His tags to facilitate Western blotting with monoclonal anti-His antibodies (Qiagen).

4.3.3.2 SCO3484.

Given the absence of any confirmed *S. scabies* lipoproteins, and the inability of the anti-CseA antibody to detect the *S. scabies* homologue of CseA, the confirmed Tat secreted *S. coelicolor* lipoprotein, SCO3484 was used as a model lipoprotein. As in chapter 3, this protein was fused to a penta-Histidine peptide at its C-terminus and expressed in each of the *S. scabies* strains. This allowed the monoclonal anti-His antibody (Qiagen) to be used. Given the sensitivity of this antibody, typically a very short exposure time was used. This can be seen in figure 4.13. If the band in the membrane of the wild-type fraction represents the mature, processed, form of the protein then this mature protein is also detectable in the cytoplasm of the wild-type. The cytoplasm of the *Δlgt* strain (BJT1040 - see table 2.2) contains two forms of the lipoprotein, presumably with and without signal sequence. Only the unprocessed form is detectable in the membrane of the *Δlgt* mutant in a 2 second exposure. The cytoplasm of the *Δlsp* strain (BJT1044) contains the fully processed protein, whilst the unprocessed form is seen in the membrane. The *Δlnt1* (BJT1047) cytoplasm contains both forms of SCO3484, whilst only the mature form is seen in the

membrane. The same result is seen in the membranes of the *Δlnt2* (BJT1048) and *Δlnt* double (BJT1050) mutants (not shown).

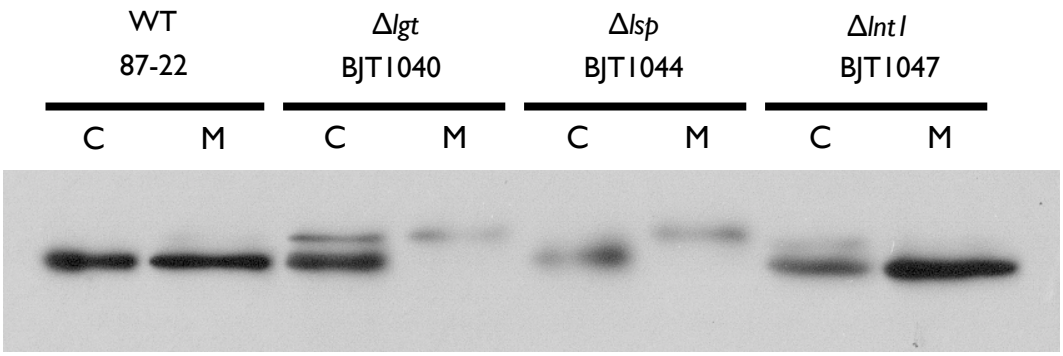


Figure 4.13. Western blot analysis, using anti-his antibody of the cytoplasm (c) and membrane (m) fractions of the different *S. scabies* strains, as listed.

4.3.3.3 PstS-His.

A second *S. coelicolor* lipoprotein, PstS (phosphate specific transport), was used to study the processing of *S. scabies* lipoproteins. This lipoprotein is part of the high affinity, low velocity system for the transport of P_i ions across the cytoplasmic membrane. This system has been well characterised in *E. coli* [124], *B. subtilis* [125] and *S. coelicolor* [126]. The PstS lipoprotein (SCO4142) was produced with a penta-histidine tag at its C-terminus, and expressed in each of the *S. scabies* strains. The results, shown in figure 4.14 show that the fully processed protein can be seen in the membranes of the wild-type, *Δlnt1* (BJT1047), *Δlnt2* (BJT1048), and *Δlnt* double (BJT1050) mutants, although the protein is detectable at a lower level in these mutants. PstS is undetectable in the membrane of the *Δlgt* mutant (BJT1040), but does appear in the supernatant of the same strain (M. Hicks, unpublished). The protein is larger in the *Δlsp* mutant (BJT1044), and presumably retains its signal sequence.

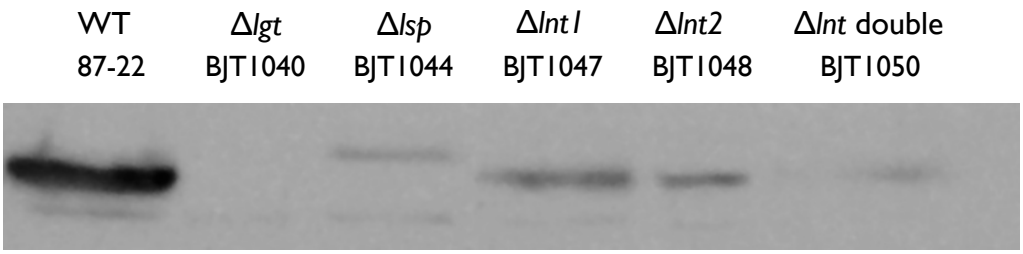


Figure 4.14. Western blot analysis, using an anti-his antibody, of the membrane fractions of the different *S. scabies* strains, as listed.

4.3.4 Plant virulence assays.

As has been mentioned, to date there has been no investigation into the role that lipoproteins play in plant pathogenicity although numerous studies have been done on animal pathogens (chapter 1). The aim of this work is to assess the role that *S. scabies* lipoproteins play in the plant infection process. Two approaches were chosen. The first involved overlaying sterilised discs of potato with agar plugs containing a confluent lawn of growth of the wild type, *Δlgt, Δlsp, Δlnt1, Δlnt2* or *Δlnt* double mutant *strains of S. scabies*. The second involved assessing the effect that each of the mutants had on a whole plant, in this case newly germinated radish seedlings (*Raphanus sativus*) (Chapter 2). These methods were adapted from previous work on the infection pathway of *S. scabies* [127]. As has been noted, *S. scabies* is capable of infecting a wide number of plants, both monocotyledonous and dicotyledonous [67] and radish plants were chosen due to their rapid germination and growth. The potato cultivar Maris Piper was used in the potato disc experiments due to its low immunity to *S. scabies*, as documented by the British Potato Council (http:// varieties.potato.org.uk).

4.3.4.1 Potato tuber assay.

The effect that each of the *S. scabies* strains had on potato slices was assessed after both 2 and 5 nights. The results can be seen in figures 4.15 and 4.16 below. Circular discs were cut from a sterile potato tuber and had an agar plug containing a confluent lawn of each *S. scabies* strain placed upon them (section 2.8.1). After two nights the control, overlain with sterile agar only, showed no evidence of necrosis. It is difficult to see where the agar was placed. In contrast, a brown square is clearly visible in the wild-type infected potato disc, and represents an area of necrosis. This is also visible in each of the mutant strains (fig. 4.15), although it does appear to be lighter in colour in the *Δlnt2* mutant (BJT1048 - see table 2.2) than in the other strains. Whilst these areas of necrosis are present in each strain, there is an absence of any obvious breakdown of the tuber surface, which remains a uniform flatness.

After 5 nights growth however, there are large areas of potato tissue degradation caused by each of the *S. scabies* strains (fig. 4.16). As before, the negative control remains totally clear of any areas of necrosis, and it is still difficult to see where the agar plug has been placed. For each of the infected strains, removing the agar square has also removed the top layer of potato, either completely as seen in the wild-type or *Δlgt* (BJT1040) infected potatoes, or partially, as seen in the *Δlnt1* (BJT1047) or *Δlnt* double (BJT1050) mutants. The loss of this top layer is presumably due to the bacteria becoming tightly associated with the surface of the potato and potentially growing into the disc, in a manner similar to that seen in figure 4.3. A large covering of bacteria remains on each potato, unseen after 2 nights. There is no obvious difference in the rate of growth suggesting that the loss of each respective lipoprotein processing

enzyme has had no effect on the ability of the strain to grow on, and necrose, the potato tissue. In the wild-type infected potato, large areas of black are seen, corresponding to severe necrosis, likely caused by thaxtomin A. These areas are beyond the edges of where the agar plug was placed, suggesting that either the thaxtomin A is readily diffusible, or that there is a large amount of unseen substrate mycelial growth, and it is from here that the toxin is secreted. These black areas are also seen in the *Δlgt* mutant, and the *Δlnt1* mutant, although to a lesser extent. They are absent from the remaining mutants, which retain the brown areas seen after two nights infection.

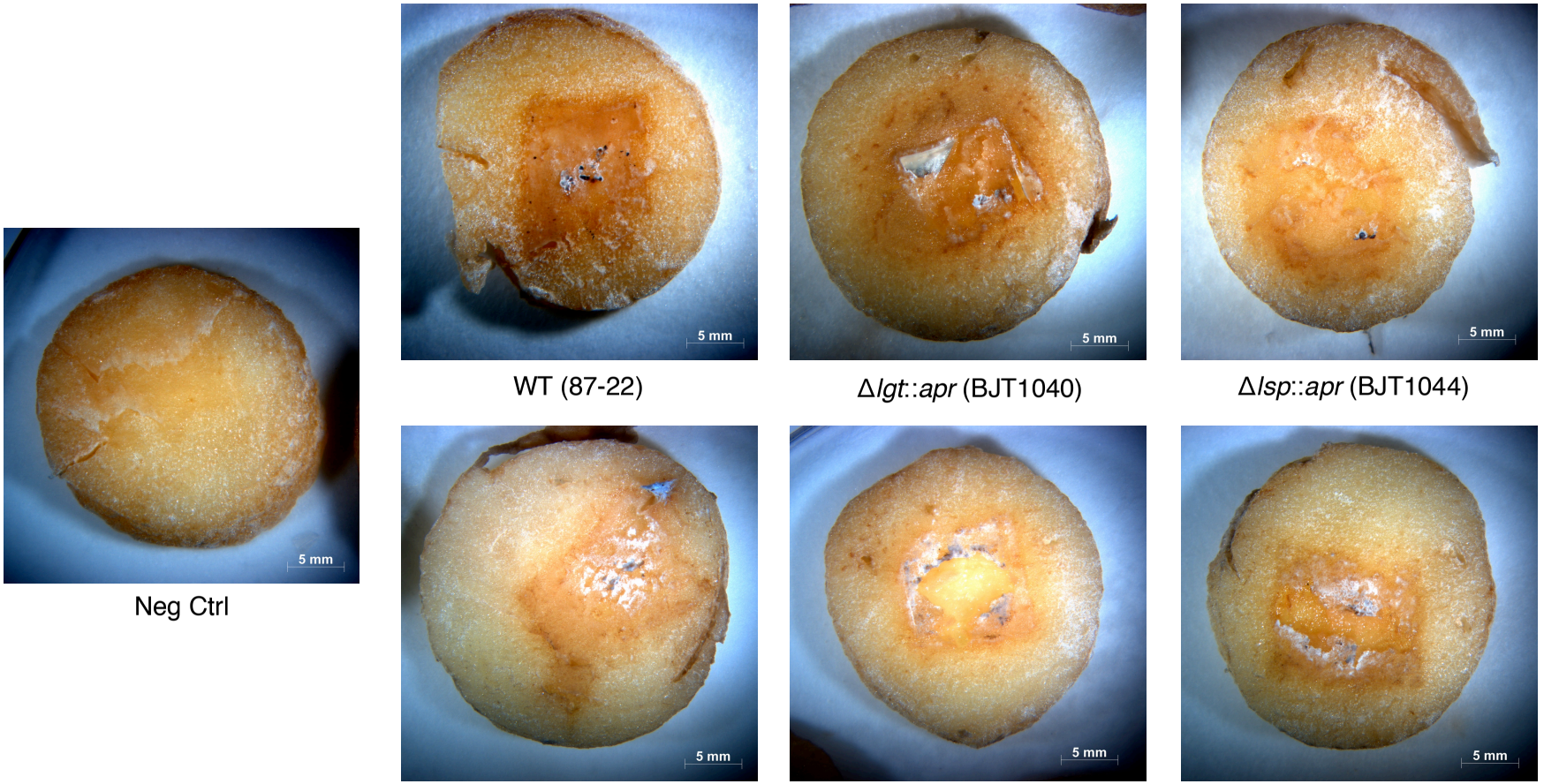


Figure 4.15. 2 night infection of sterile potato discs. The brown area, visible in all but the negative control represents tissue necrosis and shows where the agar plug has been placed. Strain names are shown in brackets.

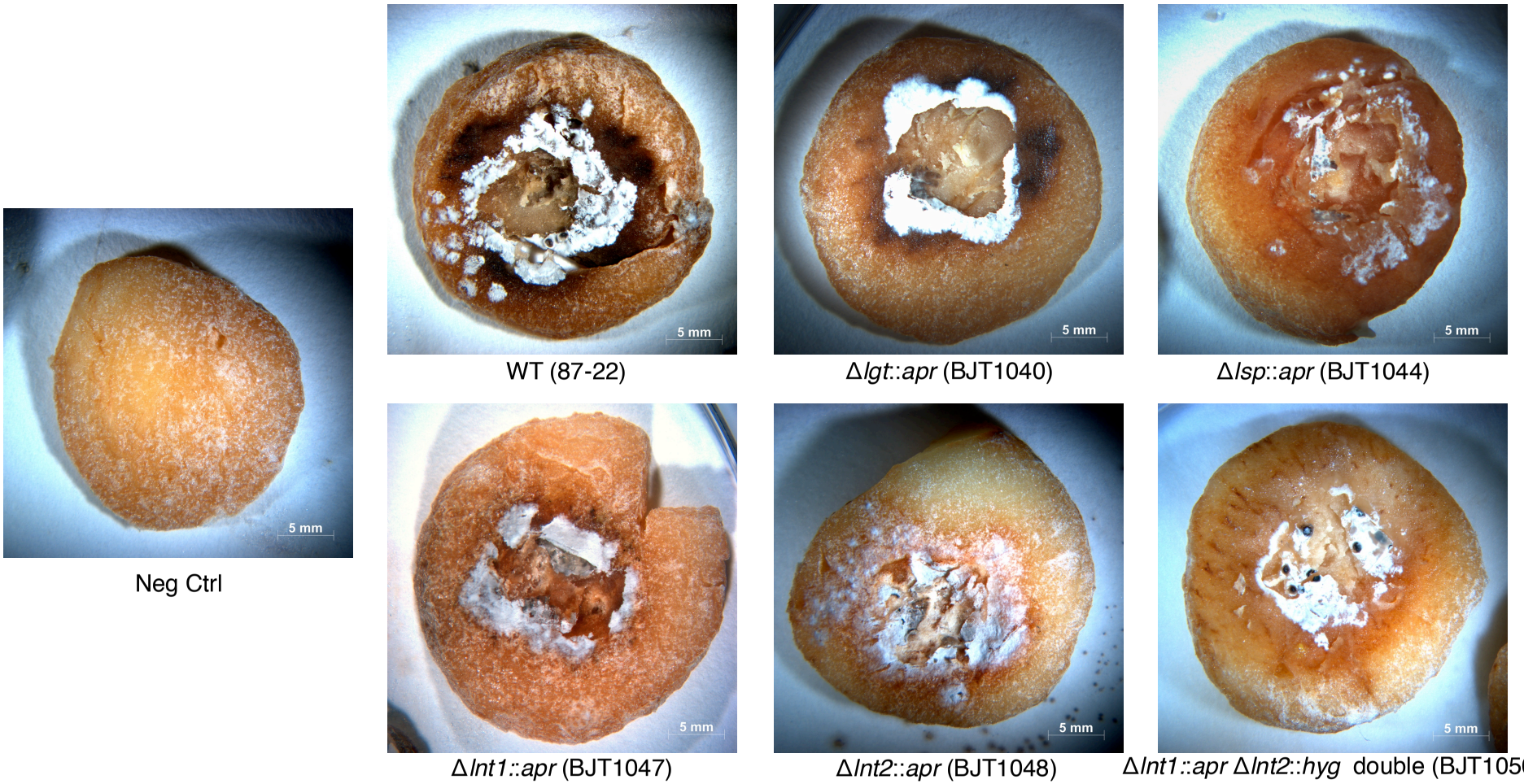


Figure 4.16. 5 night infection of sterile potato discs. Black areas representing tissue necrosis can be seen in the wild-type, *Δlgt* and *Δlnt1* infected discs. Strain names are shown in brackets.

4.3.4.2 Radish seedling assay.

Plant infection assays involved germinating sterile seeds of the Radish cultivar Scarlet Globe, and immersing the seedlings in a broth of *S. scabies* (wild-type or mutants), before plating them in sterile containers on Murashige and Skoog medium containing 2% sucrose (table 2.3) which had previously been overlaid with the same *S. scabies* strain. The seedlings were grown for 7 nights in a controlled environment chamber, before being assessed (section 2.8.2). An overview of the effect of each of the mutations on virulence in these seedlings can be seen in figure 4.17 below. The negative control plant, overlaid with sterile TSB growth medium only, shows healthy growth, with a long primary root, and multiple secondary roots. After 7 days the primary roots were typically between 11-12 cm in length. The leaves are healthy and new leaf growth can be seen at the apical meristem**.** The stem of the plant was typically between 1.3 and 1.6 cm. This growth pattern is also seen in plants treated with the non-pathogenic *S. coelicolor*. However, plants infected with wild-type *S. scabies* are vastly different. The growth is stunted, with a primary root of between 4-6 cm. There is very little secondary root growth evident, and those that are present are much smaller than in the negative control. Very small leaves were present and there was no evidence of further leaf growth. The stem was shorter than the negative control, rarely exceeding 1.5 cm in length. This phenotype was consistent in each of the *S. scabies* mutants tested, as can be seen in figure 4.17.

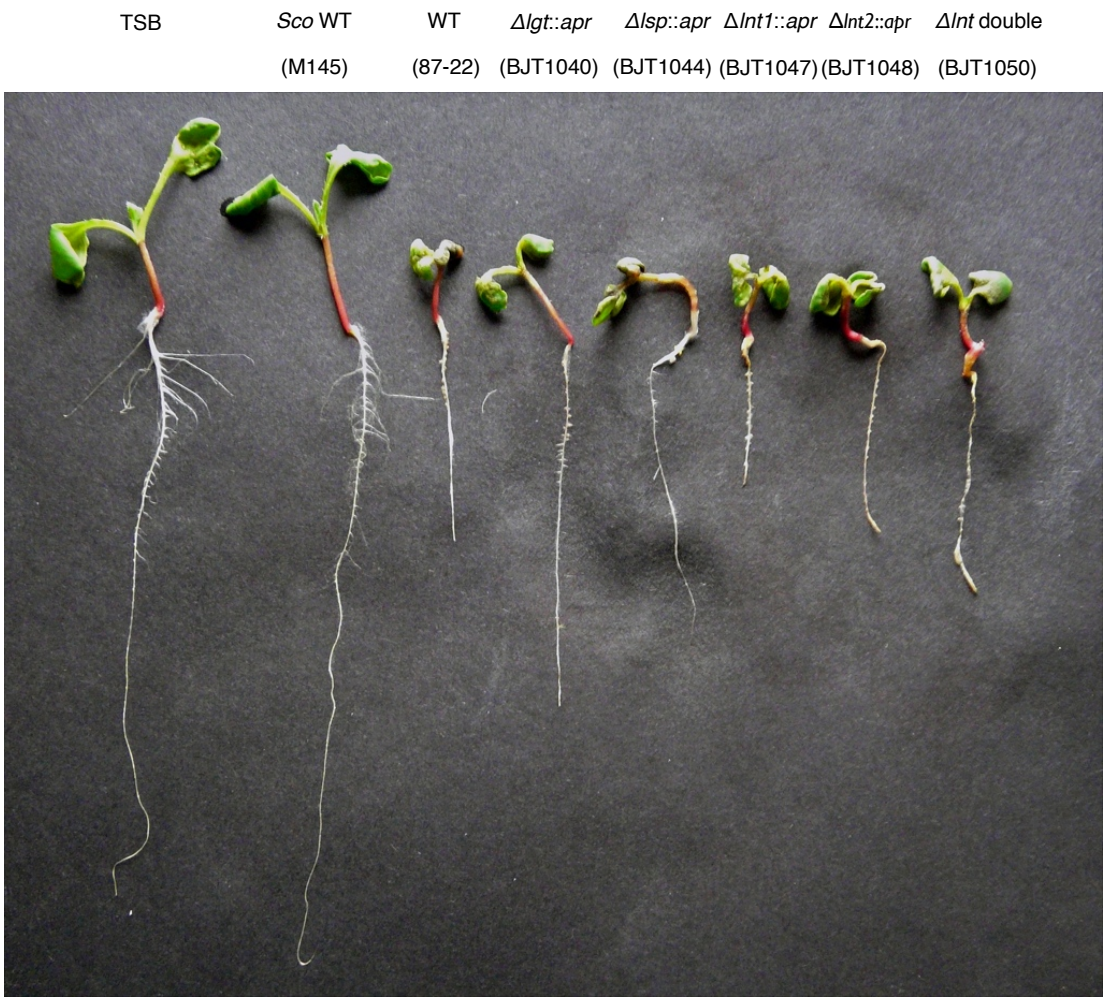


Figure 4.17. Overview of Radish seedlings grown for 7 nights and either: mock infected (TSB), infected with *S. coelicolor* wild-type (Sco WT), or the *S. scabies* strains. Strain names are shown in brackets.

4.3.4.3 Analyses of root structures.

As shown above, each of the *S. scabies* strains has a deleterious effect on the root systems of Radish seedlings not seen in mock-inoculated plants or those infected with the non-pathogenic *S. coelicolor.* The root systems of each plant were studied under a light microscope to view any subtle changes not seen in the overview above. The results can be seen in figure 4.18 below. A complex network of root hairs, invisible in the overview photo above can be see in each *S. scabies* strain. As above however, each of the mutant strains shows a severe growth phenotype compared to the negative controls. Both the TSB treated, and *S. coelicolor* treated negative controls have a complex root system consisting of many secondary roots emerging at regular intervals from the primary root. The entire root system grew into the agar making the plants difficult to remove without also removing a large amount of agar. In contrast, the root system in each of the *S. scabies* infected strains was much less complex. In the wild-type infected plant the root system comprises only the single, stunted, primary root, with multiple brown nodules which represent failed secondary roots. The root grows on the surface of the growth medium, and lacks the penetration seen in the wild-type. Only the fragile attachments of the root hairs allowed the plant to remained bound. This weak attachment is seen in all of the *S. scabies* strains. Whilst the *Δlgt* strain (BJT1040) does share the stunted primary root seen in the wild-type, the overall phenotype does not seem as severe. Multiple secondary roots are present, although they are significantly more stunted than in the wild-type. This phenotype is shared in the *Δlsp* mutant (BJT1044). The roots of both the *Δlnt1* (BJT1047) and *Δlnt2* (BJT1048) infected plants look like those seen in the wild-type infection, with the presence of nodules representing failed secondary root growth. The phenotype of the

Δlnt double mutant (BJT1050) is similar to that seen in the *Δlgt* and *Δlsp* mutants. Multiple secondary roots are seen at regular intervals. Again however, these are consistently shorter than those found in the negative controls.

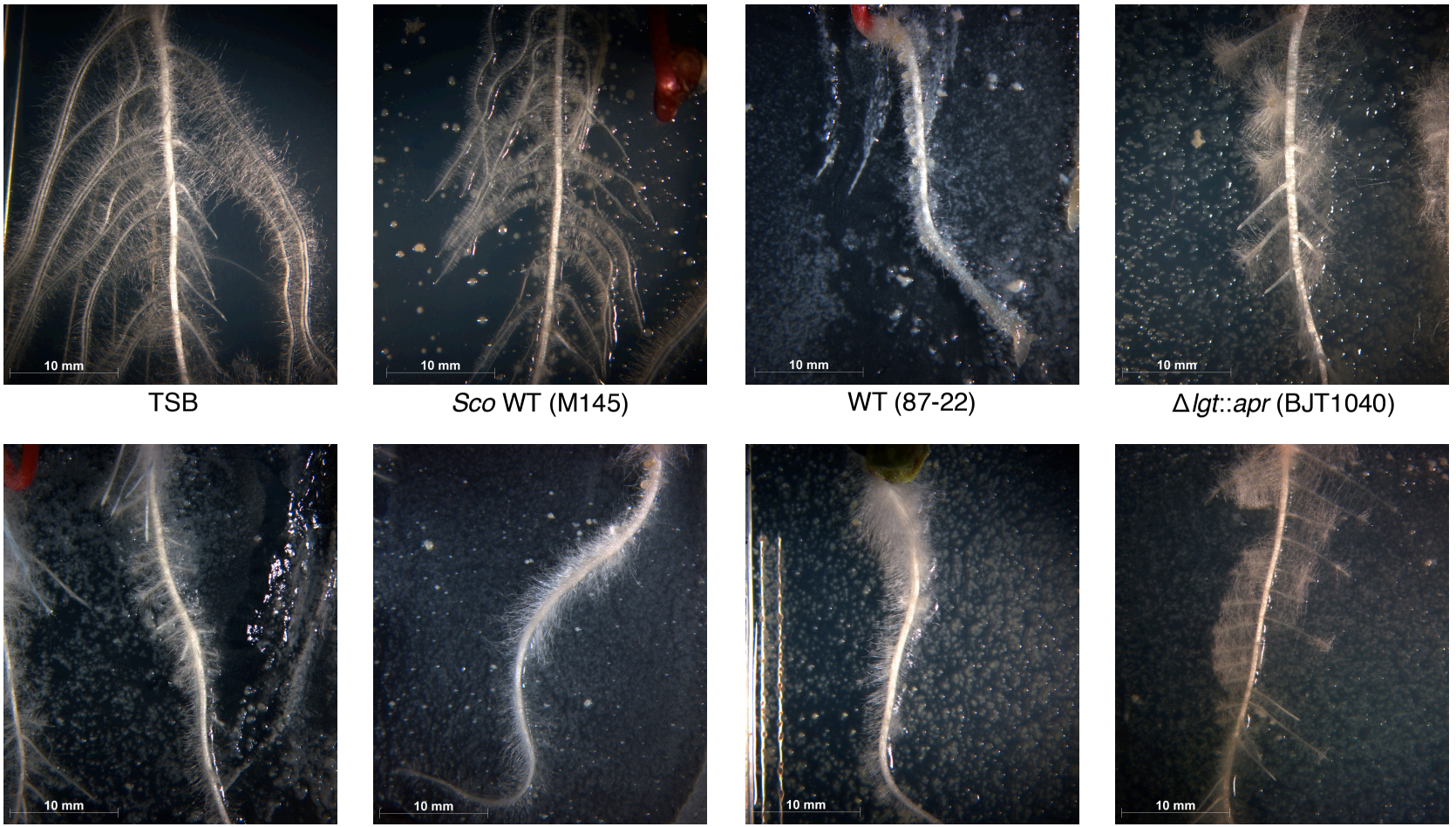


Figure 4.18. Magnified view of Radish seedling roots grown for 7 nights and either: mock infected (TSB), infected with *S. coelicolor* wild-type (Sco WT), or the Scabies strains. Strain names are shown in brackets.

4.3.5 Complementing the *S. scabies ΔIgt* and *ΔIsp* mutant.

In the previous chapter, we showed that disrupting the *lsp* gene in *S. coelicolor* caused a severe growth phenotype. This was not the case with the disruption of either of the *lgt* genes which both appeared as the wild-type, potentially due to the self complementation of the disruption. *S. scabies* is unlike *S. coelicolor* as it only contains one copy of the *lgt* gene (section 4.2). Disrupting this gene causes a growth phenotype similar to that of the *ΔIsp* mutant. Both of the *ΔIgt* and *ΔIsp* strains were only partially complemented by replacing a full length copy of the disrupted gene back into the mutant *in trans*. The *S. coelicolor lgt1*, *lgt2* and *lsp* genes were also used to see if they were able to complement the *S. scabies* mutants. Using the *lgt* genes from *S. coelicolor* allowed us to further investigate whether either, or both genes encoded a true Lgt homologue. The results for colonies grown for 5 nights on IM agar can be seen in figures 4.19 and 4.20 below.

4.3.5.1 *S. scabies ΔIgt* complementation.

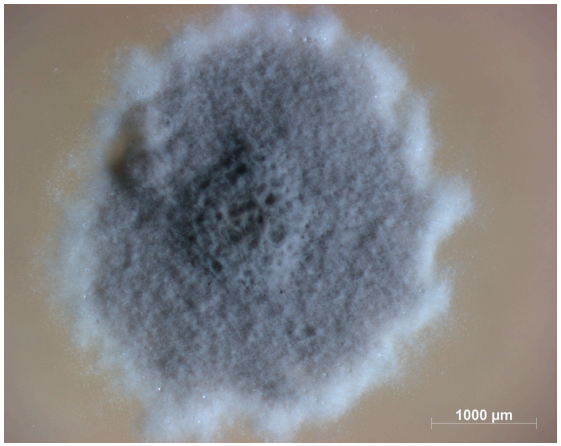
When compared to the wild-type under light microscopy, the *ΔIgt* mutant (BJT1040) showed a lower level of sporulation leading to a lighter colony appearance, spore chains were observed under SEM however. Unlike wild-type colonies which are typically flat with a raised centre, the *ΔIgt* mutant appears to have more depth and often contains depressions upon its surface similar to those described in the *ΔIsp* strain (section 4.3.2.2). Complementing the mutant with the *lgt* genes from either *S. scabies* (BJT1041 - see table 2.2) or *S. coelicolor* (BJT1042 and BJT1043 for Sco *lgt1* and Sco *lgt2* respectively) does go someway to restoring the growth phenotype to wild-type, as can be seen below. When complemented with the *S. scabies* gene, the colony regains

both the flatness, and halo of white substrate mycelium, but fails to regain the raised area seen in the middle of a wild-type colony. This complementation also appears to have a more 'ragged' appearance than the uniform shape of the wild-type. This ragged appearance is even more pronounced in the mutant complemented with the *S. coelicolor lgt1* gene. This strain lacks both a defined edge, and uniform topography, containing many raised areas. In contrast, the *S. coelicolor lgt2* complemented strain appears much closer to the wild-type. The strain is flat, and a uniform grey with a white halo surrounding it. However, the raised area in the centre of the wild-type colony is again absent from this complementation, and it was frequently smaller in size than the wild-type.

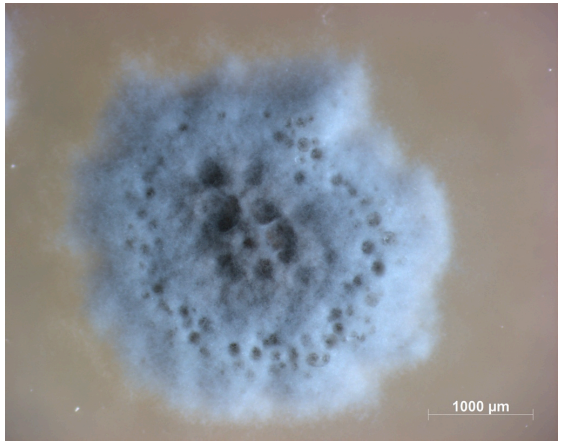
4.3.5.2 *S. scabies ΔIsp* complementation.

As discussed in section 4.3.2.2 the *S. scabies ΔIsp* mutant (BJT1044) has a severe growth phenotype, with colony sizes much smaller than the wild-type, and the presence of an unknown liquid on the colony surface (section 4.3.2.3). Substantially less sporulation occurs in the mutant, leading to an overall white appearance. Complementing the mutant with either the *S. scabies*, or *S. coelicolor lsp* gene (BJT1045 and BJT1046 - respectively) had an effect on the mutant, but neither was able to completely rescue the strain (fig. 4.20). In both cases the complemented strains showed an increase in colony size, with both being flat like the wild type, as opposed to the domed shape seen in the *ΔIsp* mutant. However, the shape of the colonies in both complementations was different to the wild-type. Unlike the broadly circular wild-type colony, the *S. scabies lsp* complementation had a misshapen central grey area, surrounded by diffuse substrate mycelium. This area of substrate mycelia was thin enough that the growth media could be seen below. There is no raised area in the

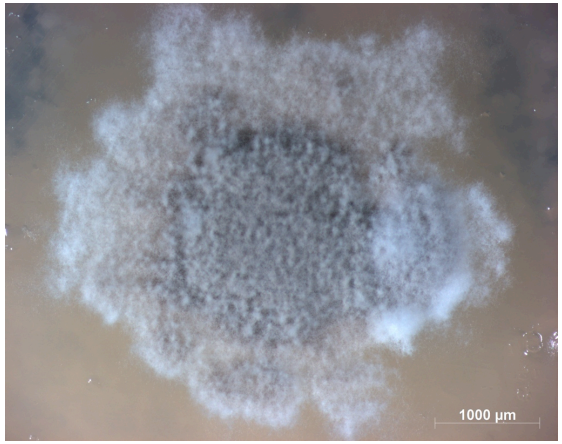
middle of the colony, instead holes are seen similar (albeit smaller) to those seen in the *Δlsp* and *Δlgt* mutants. The *S. coelicolor lsp* complemented mutant shares the large area of diffuse substrate mycelia, but in this case the centre of the colony is more spherical than the *S. scabies* complementation. However, this is divided into an outer ring and an inner circle. The outer ring is white, presumably consisting of undifferentiated mycelium, whilst the centre is grey, indicating mature spore chains are present. The lower amount of aerial hyphae corresponds with the observation that this strain sporulates poorly in comparison to the *S. scabies lsp* complementation.



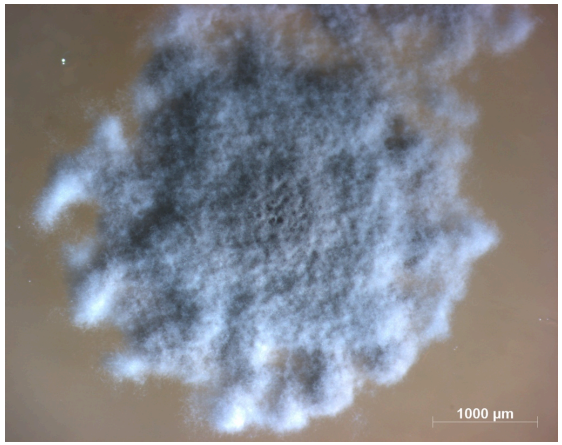
WT (87-22)



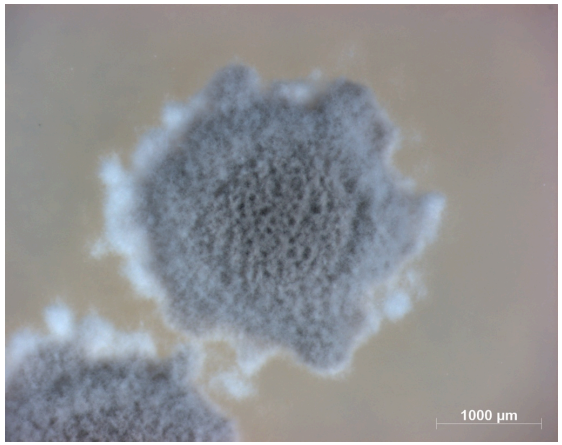
Δ*lgt*::*apr* (BJT1040)



Δ*lgt*::*apr* + Scab*lgt* (BJT1041)



Δ*lgt*::*apr* + Sco*lgt1* (BJT1042)



Δ*lgt*::*apr* + Sco*lgt2* (BJT1043)

4.19. Light microscopy of the *S. scabies Δlgt* complementation strains. The wild-type, and *Δlgt* colonies are also shown for reference. Strain names are shown in brackets.

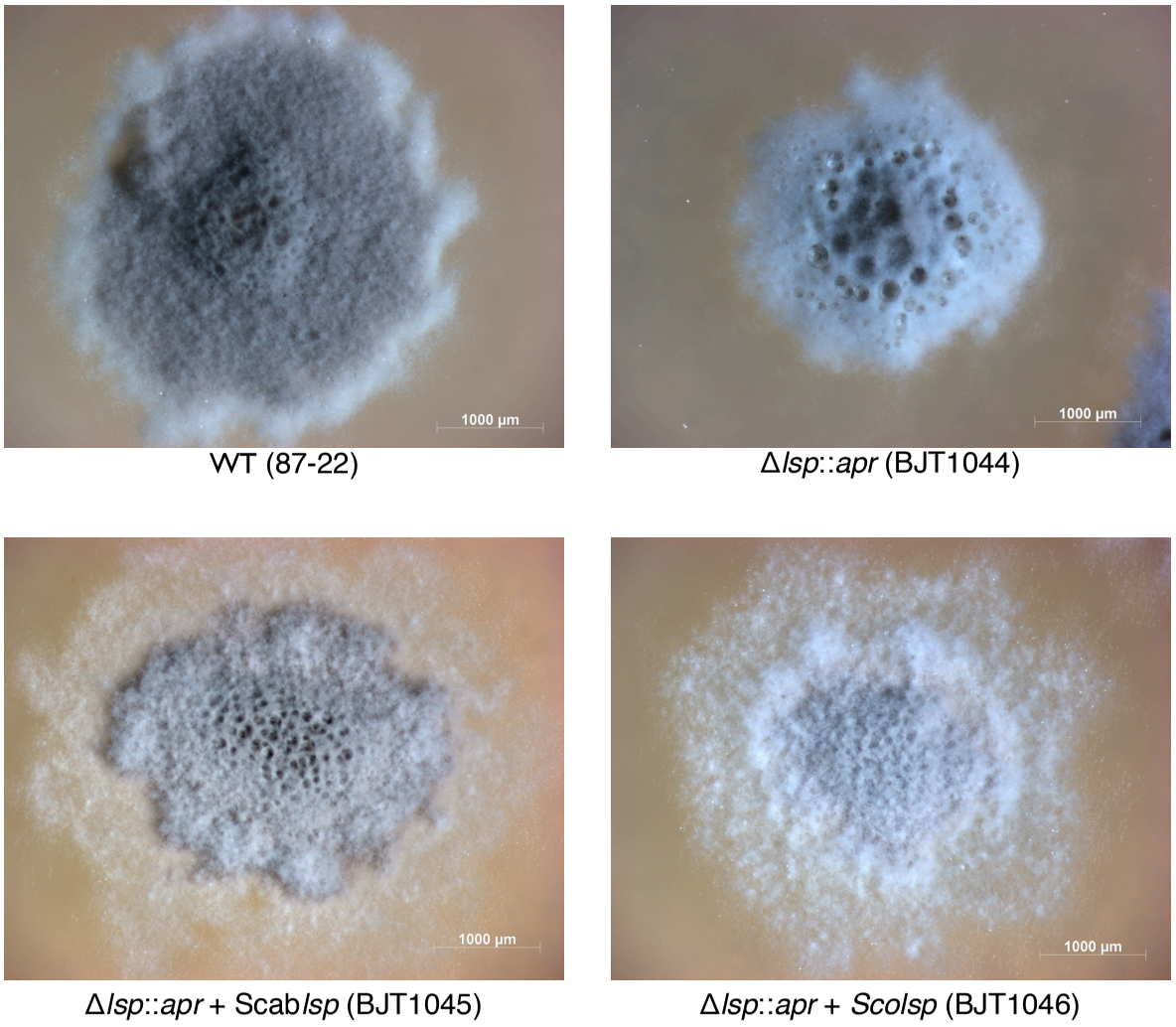


Figure 4.20. Light microscopy of the *S. scabies Δlsp* complementation strains. The wild-type, and *Δlsp* colonies are shown for reference.

4.4 Discussion

The experiments presented in this chapter were focussed on the plant pathogen, *S. scabies*. The aim was to investigate the effect that the loss of lipoprotein processing had on the bacterium, compare this to results from *S. coelicolor* (chapter 3) and determine whether this had a deleterious effect on its ability to infect plants. The role of lipoproteins in plant pathogenesis is unknown.

The bioinformatic analysis shows that, unlike in *S. coelicolor*, *S. scabies* only encodes one Lgt homologue which shares a high level of identity with both *S. coelicolor* Lgt1 and Lgt2. The *S. scabies* Lsp enzyme also shows a high level of similarity to the *S. coelicolor* enzyme. The *S. scabies* Lnt enzymes are curious, insomuch that although they share the potential catalytic triad from *E. coli*, they share a low level of identity with each other (32%). The genes encoding both Lnt1 and Lnt2 are found in the right arm of the *S. scabies* genome, suggesting that they may have been acquired by HGT and, given their low similarity, could have been acquired from different sources. Both the *lgt* and *lsp* genes are found more centrally in the chromosome.

Disrupting the lipoprotein processing enzymes in *S. scabies* was relatively simple when compared to *S. coelicolor*. However, it should be noted that recombination efficiency is much higher in *S. scabies* compared to *S. coelicolor,* such that single crossover events are rarely seen in *S. scabies* (R. Seipke, personal communication). However, observing the mutant colony phenotypes (fig. 4.9) shows there are several parallels when compared to the *S. coelicolor* mutants. The *Δlsp S. scabies* mutant is much smaller than the wild-type, and

sporulates poorly, in a manner similar to the *S. coelicolor* mutant. The *Δlgt* mutant (BJT1040) shares many of the growth deformities, although is slightly larger than the *Δlsp* strain (BJT1044). Both are discernibly different to the wild-type. The same cannot be said for the disruptions in the *lnt* genes. Neither of the single mutants (BJT1047 and BJT1048), nor the double (BJT1050) show any size reduction compared to the wild-type, although both of the single *lnt* mutants lack the raised area in the centre of the colony, characteristic of wild-type *S. scabies*. The edges of each of the *lnt* mutant colonies does look more dispersed than the wild-type, leading to an irregular shape. A lack of phenotype was observed under SEM, where the *Δlnt1* mutant (BJT1047) showed no obvious differences to the wild-type. However, as with the light microscope, the *Δlsp* mutant (BJT1044) sporulated very poorly, and the spores that were observed under SEM were often misshapen. Surprisingly, given their similarity when viewed under light microscope, the *Δlgt* strain (BJT1040) did not look nearly as damaged when viewed under SEM. One curious discovery made by SEM is the discovery that the liquid on the surface of the *Δlsp* mutant (BJT1044) is sequestered into covered vesicles (fig. 4.12). The 'skin' forming this coating is of unknown consistency, but it is punctured multiple times by the mycelia. *S. scabies* is known not to produce any antibiotics, so initially this liquid was assumed to be water. The identity of this liquid remains unknown, but is due to be investigated. Potentially it could represent a cryptic secondary metabolic pathway switched on in the absence of *lsp* or *lgt*.

The effect that the loss of the lipoprotein biosynthetic enzymes had on the processing of the two lipoproteins tested is intriguing. It is important to note that both SCO3484 and PstS (SCO4142) are native *S. coelicolor* lipoproteins, and

therefore may be processed differently in the heterologous host *S. scabies* than they are in their native *S. coelicolor*. Looking at the *S. scabies* membrane fractions (fig. 4.13), the mature SCO3484 lipoprotein is present in the wild-type, with a larger unprocessed form seen in the membranes of the *Δlgt* (BJT1040) and *Δlsp* mutant (BJT1044). The protein is properly processed in the *Δlnt1* (BJT1047), *Δlnt2* (BJT1048) and *Δlnt* double (BJT1050) mutants**.** This pattern would be expected if the lipoprotein biosynthetic pathway is ordered tightly as it is in Gram-negative bacteria. The unprocessed band seen in the membrane of the *Δlgt* strain shows that lipidation by Lgt is required before Lsp can cleave the signal peptide. The unprocessed band in the *Δlsp* strain is also seen in *S. coelicolor.* However, the presence of processed lipoproteins in the cytoplasms of each strain is difficult to explain. In the wild-type cytoplasm processed SCO3484 is detectable, as it is in the *Δlsp* cytoplasm. In the cytoplasms of the *Δlgt* and *Δlnt*s both processed and unprocessed forms can be seen. The reasons for these extra bands are unclear, although the result is reproducible. Experimental error can be discounted, given that the pattern is different between strains, and fractions, and any problem with the fractionating process would give a uniform pattern for each lane. There should be no bands detectable in the cytoplasm, as seen in *S. coelicolor* (fig. 3.16).

The processing of PstS (fig. 4.14), is different from that seen with SCO3484. In this case, PstS is undetectable in the membrane of the *Δlgt* mutant (BJT1040). However, a band corresponding to the processed form is detectable in the supernatant of the *Δlgt* mutant, showing that it has been cleaved. This could suggest that Lsp is able to act on non-lipidated proteins. However, mass spec analysis of another lipoprotein found in the supernatant of the *Δlgt* mutant

revealed that the signal peptide has been cleaved just downstream of the lipobox, removing the C_{+1} (M. Hicks, unpublished) suggesting that processing is not due to Lsp but due to an alternative signal peptidase, potentially a signal peptidase I (SpaseI). In Gram-negative bacteria, one copy of SpaseI exists. However in *B. subtilis* and several *Streptomyces* species, including *S. coelicolor*, multiple homologues (SCO5596, 5597, 5598, 5599) exist [128]. Another possible explanation would be that an alternative, as yet undiscovered Lsp, such as the Eep protein from *Streptococcus uberis* [53] is cleaving the lipoprotein. The use of alternative processing indicates again that the lipoprotein biosynthetic pathway is potentially ordered, given that Lsp does not seem to act on a non-lipidated protein. The band representing PstS in the *Δlsp* mutant (BJT1044) is larger than seen in the wild-type, suggesting that, as with SCO3484, it retains its signal sequence. The deletion of the *lnt* genes, either individually or together, appears to have no effect on the processing of PstS, as was the case with SCO3484. As both of these proteins are from *S. coelicolor*, future work will involve using native *S. scabies* lipoproteins to see whether the results are the same as those described here.

Complementing both the *Δlgt* and *Δlsp* mutants only partially rescues each strain. The similarity of the phenotypes in both the *Δlgt* and *Δlsp* mutants suggest that they share a defective lipoprotein biosynthesis pathway, which in turn suggests that the pathway in *S. scabies* has a regimented order as found in Gram-negative bacteria. This result is also strong evidence that the two Lgt enzymes found in *S. coelicolor* are complementing each other (chapter 4).

Recently it has been shown that the *S. coelicolor* lipoprotein SCO3484 is triacylated in wild-type *S. scabies*, whilst in the *Δlgt* mutant (BJT1040) it is unlipidated (M. Hicks unpublished). This shows that *S. scabies* contains at least one functional Lnt enzyme although what role this N-acylation plays is unknown. Given that the double *Δlnt* disruption mutant displays no severe growth phenotype, the importance of triacylation of lipoproteins in *S. scabies* must be questioned. Perhaps only a subset of lipoproteins are triacylated in *Streptomyces.*

What is clear from this work is that disrupting lipoprotein processing seems to have an effect on plant pathogenicity. Each of the mutant strains was able to infect both potato tubers and radish seedlings to a similar extent to the wild-type although the roots of the radish seedlings were less stunted in the *Δlgt* (BJT1040), *Δlsp* (BJT1044) and *Δlnt* double (BJT1050) mutants compared to wild type *S. scabies* suggesting a potential reduction in virulence. Deleting the Tat secreted lipoprotein SCAB81041 has been shown to have a moderate effect on the virulence of *S. scabies* [80] in agreement with this work. That disrupting the lipoprotein processing pathway in *S. scabies* does not have a larger impact on virulence is surprising, given that three lipoprotein genes are found on the *S. scabies* PAI island: *SCAB77471*, encoding a substrate binding protein (SBP) also found in *S. avermitilis*, *SCAB77361* and *SCAB77271*, both encoding SBP, although unique to *S. scabies*. However, there is no evidence that these proteins are virulence factors. Given that the improper processing of lipoproteins does not have a large impact on virulence, it appears that lipoproteins are not a determining factor in *S. scabies* pathogenesis. Neither of the two main pathogenicity factors, thaxtomin A and Nec1, are lipoproteins and

it seems that the loss of the lipoprotein processing enzymes does not effect their action. However, Nec1 production seems to be stimulated by glucose [75] and thaxtomin A production appears to be induced by xylan and glucan [130], so the lack of effect seen through disruption of lipoprotein biosynthetic genes is surprising given the large number of lipoproteins are involved in carbohydrate transport in the streptomycetes (chapter 3). There is no evidence of hypervirulence in any of the strains, suggesting that lipoproteins are not the determining factor for the initiation of the plant defence response, as their loss does not allow the bacteria to infect the host plant undetected. The radish assays above had a seven night duration, after which the virulence was assessed. Future work should include a longer term experiment to see whether any of the plants are able to recover from infection.

Chapter 5. Discussion

The aim of this thesis was to investigate the lipoprotein processing pathway in *Streptomyces*. Two *Streptomyces* species were used as models, *S. coelicolor*, a model organism for the high GC Gram-positive actinomycetes and *S. scabies*, one of the causative agents of scab formation on potato tubers, a disease of significant economic value worldwide. Both species are soil dwelling saprophytes, with complex multistage lifecycles. Their genomes are linear and considerable larger than most other bacteria, and presumably this linearity allows the acquisition of new genes, making the species more successful soil saprophytes. As little is known about the lipoprotein processing pathway in the high GC Gram-positive bacteria this work, which involved disrupting the genes encoding each of the biosynthetic enzymes in both S. coelicolor and S. scabies, allowed us to observe whether the pathway is essential and ordered, as in the Gram-negatives, or dispensable, as it is in the low GC Gram-positives.

S. coelicolor contains two Lgt homologues, and both *S. coelicolor* and *S. scabies* contain two Lnt homologues. The reasons for this are unclear. The presence of two Lgt homologues seems unusual amongst *Streptomyces* species, with only *S. coelicolor* and *S. clavuligerus* containing the duplicates, the latter on a megaplasmid, whilst the presence of two Lnt homologues seems ubiquitous throughout the species. The presence of multiple enzymes made it difficult to determine whether Lgt has an essential function in *S. coelicolor*. Disruption of one or other of the *S. coelicolor lgt* genes had very little effect on the growth of the bacterium, nor did it seem to effect the processing of the two model lipoproteins tested or on the lipoproteome, as detected by 2D gel analysis [82]. The two *lgt* genes are therefore self complementing in *S.*

coelicolor. The double *lgt* mutant was impossible to isolate in this study, which in itself suggests that the lipidation step of the lipoprotein biosynthetic pathway may be essential in *S. coelicolor* and that therefore the entire pathway is also essential. The second *lgt* gene (SCO7822) is found very close to the end of one of the *S. coelicolor* chromosome arms, and these are the regions of the chromosome which typically contain non-essential genes acquired through HGT [66]. Why *S. coelicolor* has sought to maintain this gene when a second homologue is encoded elsewhere on the chromosome does suggest that the Lgt function is of great importance to the bacterium. One hypothesis was that each enzyme caters to a different subset of lipoproteins, for example Tat or Sec secreted lipoproteins, but this now seems unlikely, given that the two lipoproteins tested (one Sec- and one Tat-dependent) were processed correctly in both the *Δlgt1* and *Δlgt2* mutant.

S. scabies only contains a single *lgt* gene, and it was possible to delete this gene, which indicates that Lgt function is not essential in *S. scabies*. However, the growth phenotype exhibited in this mutant is very similar to that observed upon disrupting the *lsp* gene, suggesting that they share incorrect lipoprotein processing which in turn has an effect on growth and development, as also observed in *S. coelicolor*. In the *Δlgt* deletion strain, a band consistent with unprocessed SCO3484 still containing its signal peptide was seen in the cell membrane, suggesting that the action of Lgt is a required in order for Lsp to recognise its substrate and cleave the signal sequence, indicating an ordered processing of lipoproteins. PstS is processed differently in the *Δlgt* mutant, and was undetectable in the cell membrane. The unlipidated protein was probably cleaved by an alternative signal peptidase, showing that, for this lipoprotein at

least, Lsp is unable to function in the absence of Lgt, again suggesting a strict order in the lipoprotein processing pathway. Why there is a difference in the processing of the two lipoproteins is unclear at this time, although there was also a difference seen in the processing of CseA and SCO3484 in *S. coelicolor*.

Disrupting the *lsp* gene from both *S. coelicolor* and *S. scabies* resulted in a growth phenotype very different from the wild-type, both growing and sporulating poorly. In *S. coelicolor* the mutation could not be completely rescued by complementation, either *cis* or *in trans* with the loss of the gene causing the improper processing of the two model lipoproteins tested, with one (CseA) becoming undetectable in the cell membrane, and one (SCO3484) retaining its signal sequence. Which of these two outcomes is more prevalent is unknown, and more lipoproteins will have to be tested to see which is the norm. Either way, both forms of misprocessing are likely to have a significant effect on the bacterium, either leading to the loss of swathes of lipoproteins, or having numerous lipoproteins anchored into the cell membrane by both their lipid moiety and signal sequence.

As discussed in chapter 4, recent results have proved the presence of at least one true Lnt enzyme in *S. scabies*. In the *Δlgt* deletion mutant the model lipoprotein tested is unlipidated and has been non-specifically processed just downstream of the lipobox, whilst the same lipoprotein is triacylated in the wild-type (M. Hicks, unpublished). Currently, both the *Δlnt1* and *Δlnt2* deletion strains are being tested for N-acylation of lipoproteins to see which, if not both, acts as a true Lnt enzyme. Neither *S. scabies lnt* mutant exhibited a severe growth phenotype probably due to self complementation, as seen in the *S.*

coelicolor Δlgt mutants, although the phenotype of the *lnt* double mutant is not as severe as that of the *lgt* and *lsp* mutants. Now that functional Lnt enzymes have been identified in *M. smegmatis* and *S. scabies,* with very similar proteins present in a variety of actinomycetes including *S. coelicolor* it must be asked 'Why does this branch of the Gram-positives need an Lnt enzyme?'. In Gram-negative bacteria, this enzyme is a prerequisite for the activity of the Lol pathway, needed to transport a lipoprotein to the outer membrane. In Gram-positive bacteria the Lol pathway is missing, given the lack of outer membrane, and it could therefore be assumed that the Lnt enzyme function is not required. However, given the complex nature of the cell envelope in various Gram-positives, particularly *M. tuberculosis*, and the evidence of a potential pseudo-periplasmic space discovered in others, such as *S. aureus* [55] and *B. subtilis* [56], it is tempting to think that there may be a targeting system, as yet unknown, in the Gram-positive branch of the bacteria in order to shuttle lipoproteins to the 'outer' membrane.

Another aspect of lipoprotein processing that this work uncovered is the presence of elongated sections of the *Streptomyces* enzymes. In both species the Lgt and Lsp enzymes contain elongated sections, at the C- and N-terminals respectively. These extensions are absent from both the Gram-negative and low GC Gram-positive bacteria assessed. These areas are both found cytoplasmically suggesting, initially at least, that they serve no enzymatic function, given that the active sites of both proteins are on the external face of the cell membrane (chapter 1). The initial theory was that these extensions represent sites of interaction between the Lgt and Lsp enzyme. This is attractive given that the Gram-negative pathway is well ordered (as may be the

pathway in *S. scabies* as described above), and that passive diffusion of unprocessed lipoproteins between the processing enzymes seems energetically wasteful. In opposition to this theory is the lack of extensions in the Lnt enzymes and the translocation machinery. The majority of this work was done on the *S. coelicolor* Lsp protein, as truncating either of the Lgt enzymes singularly has no effect, presumably due to the self complementation mentioned previously. It was thought that if the Lgt and Lsp proteins were interacting directly, truncating the N-terminus of the Lsp enzyme would impede or stop this interaction, and the processing of the lipoprotein tested, CseA, would cease in a manner similar to the deletion strain. Removing the first 10 amino acids of the protein had no effect on CseA processing, yet removing 20 lead to it being undetectable. Removing 30 amino acids lead to a size shift in CseA, suggesting the enzyme was inactive, whilst removing 40 amino acids again leads to CseA being undetectable. With the exception of the N10 truncation, removing the N-terminus of Lsp results in the loss of enzyme activity. It is not known however, whether these truncated enzymes are expressed or not.

Attempts to fuse both of the *S. coelicolor* Lgt and Lsp proteins to both eGFP and mCherry fluorescent markers to visualise them were unsuccessful, given their low expression and the natural background fluorescence of *S. coelicolor*. It was thought that the lipoprotein processing machinery may have been localised at the growing hyphal tip, given that this is the location of new bacterial growth. The Tat machinery has been found at the hyphal tip (D. Widdick, unpublished) and to find the lipoprotein processing enzymes there too would lend weight to the theory that the whole pathway is linked.

Although they share the same environment, *S. scabies* differs from *S. coelicolor* as it has gained the ability to infect plants. Two main compounds, thaxtomin A and Nec1 are the main elicitors of necrosis in potato tubers. Previous to this work, no research had been undertaken on the role lipoproteins have on plant pathogenesis. Much has been done on their action in animal pathogenesis, and the results of disrupting their biosynthesis range from attenuation to hypervirulence. What is clear though is that the loss of Lgt or Lsp is not lethal in the Gram-positive pathogens tested to date. This is also the case in *S. scabies.* Both the *Δlgt* and *Δlsp* deletion strains show a similar growth phenotype as described above. Neither mutation can be fully complemented, similar to the *S. coelicolor Δlsp* mutant, suggesting the strains have also gained secondary mutations in order to survive. The disruption of these enzymes had a clear effect on the processing of the two model lipoproteins tested, with examples of improper localisation shown in each deletion strain. It is important to note that both SCO3484 and PstS (SCO4142) are native *S. coelicolor* lipoproteins, and therefore may be processed differently in the heterologous host *S. scabies* than they are in their native *S. coelicolor*. This is certainly the case for SCO3484 as can be seen in figure 3.16 (*S. coelicolor*) and figure 4.13 (*S. scabies*). Future work will have to involve native *S. scabies* lipoproteins. Despite this incorrect processing**,** none of the mutants showed any avirulence in the two plant assays carried out, with each deletion strain able to causes disease in both potatoes and radishes.

The Radish seedlings were also susceptible to each of the *S. scabies* strains. The general pattern of growth seen in the wild-type infected plant: short stem, withered leaves and a short primary root is replicated in each infection, although

the overall severity of virulence caused by the *Δlgt*, *Δlsp* and *Δlnt* double mutants seemed less than the wild-type, with more secondary roots seen, in contrast to the complete failure of secondary root growth in the other strains. This growth stunting was clearly a result of the *S. scabies* as the non-pathogenic *S. coelicolor* produced no evidence of damaged growth. Given that the reduction of secondary root formation was frequently seen in each of the deletion strains, and the 'nodules' representing this failed growth are thought to be the area where *S. scabies* gains access to the plant [77], it appears as though disrupting the *S. scabies* lipoprotein machinery has had no effect on the bacteria's ability to colonise root structures. That disrupting the lipoprotein processing pathway does not have a significant effect on virulence is surprising, given that three lipoproteins encoding SBP (SCAB77471, SCAB77361, SCAB77271) are present in the *S. scabies* pathogenicity island, although as noted above, there is no evidence that these are virulence factors. The loss of the Tat dependent lipoprotein SCAB81041 has been shown to cause a moderate decrease in virulence in *A. thaliana* [80]. A moderate decrease in virulence is also seen in the *Δlgt*, *Δlsp* and *Δlnt* double mutants, showing that the lipoprotein processing pathway does play a role in plant pathogenesis. A further gene involved in the pathogenicity process *SCAB78931* has been discovered. This gene encodes a cutinase, an enzyme used by pathogenic fungi to break down the protective layer of cutin found on the aerial organs of plants. The improper processing of this lipoprotein does not seem to impede the infection process, and its role is still unclear, given that the target organs for the bacterium are found underground. This gene does offer a potential insight into the origins of the pathogenicity genes found in *S. scabies* given that its closest homologue is found in the fungus *Phytophthora infestans*, the causative

agent of potato blight. The large numbers of potentially improperly processed substrate binding proteins also does not effect the production of pathogenicity elicitors. As has been noted, both Nec1 and thaxtomin A appear to be induced by carbohydrates [75, 130], and as we have seen, large numbers of lipoproteins, in *S. coelicolor* at least, are involved with carbohydrate transport. Again though, as there is no detriment to virulence it seems likely that either non-lipoprotein transporters are used for carbohydrate uptake, or any lipoproteins necessary remain active.

To conclude, this work has analysed the lipoprotein processing pathways of both *S. coelicolor* and *S. scabies* and shown strong evidence that the lipoprotein biosynthesis pathway is essential in *S. coelicolor*, and preliminary evidence in the latter that the pathway by which lipoproteins are processed occurs in a strict, regimented order.

Bibliography

1. Natale, P., Brüser, T. and Driessen, A.J.M. (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane-Distinct translocases and mechanisms. *Biochim Biophys Acta* **1778**: 1735-1756.
2. Widdick, D.A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschröder, M. and Palmer, T. (2006) The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. *Proc Natl Acad Sci USA* **103**: 17927-17932.
3. De Keersmaeker, S., Vrancken, K., Van Mellaert, L., Anné, J. and Geukens, N. (2007) The Tat pathway in Streptomyces lividans: interaction of Tat subunits and their role in translocation. *Microbiology (Reading, Engl)* **153**: 1087-1094.
4. Lee, P.A., Tullman-Ercek, D. and Georgiou, G. (2006) The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol* **60**: 373-395.
5. Gralnick, J.A., Vali, H., Lies, D.P. and Newman, D.K. (2006) Extracellular respiration of dimethyl sulfoxide by Shewanella oneidensis strain MR-1. *Proc Natl Acad Sci USA* **103**: 4669-4674.
6. Valente, F.M.A., Pereira, P.M., Venceslau, S.S., Regalla, M., Coelho, A.V. and Pereira, I.A.C. (2007) The [NiFeSe] hydrogenase from Desulfovibrio vulgaris Hildenborough is a bacterial lipoprotein lacking a typical lipoprotein signal peptide. *FEBS Letters* **581**: 3341-3344.
7. Sargent, F., Berks, B.C. and Palmer, T. (2006) Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. *FEMS Microbiology Letters* **254**: 198-207.

8. Lee, N.-K., Yun, C.-W., Kim, S.W., Chang, H.-I., Kang, C.-W. and Paik, H.-D. (2008) Screening of Lactobacilli derived from chicken feces and partial characterization of Lactobacillus acidophilus A12 as an animal probiotics. *J Microbiol Biotechnol* **18**: 338-342.

9. Berg, J., Tymocozko, J.L. and Stryer, L., (2007) *Biochemistry*. W. H. Freeman.

10. Braun, V. and Wu, H.C. (1994) Lipoproteins, structure, function, biosynthesis and model for protein export. *New Compr Biochem* **27**: 319-341.

11. Hutchings, M.I., Palmer, T., Harrington, D.J. and Sutcliffe, I.C. (2009) Lipoprotein biogenesis in Gram-positive bacteria: knowing when to hold 'em, knowing when to fold 'em. *Trends Microbiol* **17**: 13-21.

12. Hayashi, S. and Wu, H.C. (1990) Lipoproteins in bacteria. *J Bioenerg Biomembr* **22**: 451-471.

13. Narita, S.-I. and Tokuda, H. (2006) An ABC transporter mediating the membrane detachment of bacterial lipoproteins depending on their sorting signals. *FEBS Letters* **580**: 1164-1170.

14. García-del Portillo, F. and Cossart, P. (2007) An important step in listeria lipoprotein research. *Journal of Bacteriology* **189**: 294-297.

15. Qi, H.Y., Sankaran, K., Gan, K. and Wu, H.C. (1995) Structure-function relationship of bacterial prolipoprotein diacylglyceryl transferase: functionally significant conserved regions. *Journal of Bacteriology* **177**: 6820-6824.

16. Selvan, A. and Sankaran, K. (2008) Localization and characterization of prolipoprotein diacylglyceryl transferase (Lgt) critical in bacterial lipoprotein biosynthesis. *Biochimie* **90**: 1647-1655.

17. Sankaran, K., Gan, K., Rash, B., Qi, H.Y., Wu, H.C. and Rick, P.D. (1997) Roles of histidine-103 and tyrosine-235 in the function of the prolipoprotein diacylglyceryl transferase of Escherichia coli. *Journal of Bacteriology* **179**: 2944-2948.

18. Rezwan, M., Grau, T., Tschumi, A. and Sander, P. (2007) Lipoprotein synthesis in mycobacteria. *Microbiology (Reading, Engl)* **153**: 652-658.

19. Geukens, N., De Buck, E., Meyen, E., Maes, L., Vranckx, L., Van Mellaert, L., Anné, J. and Lammertyn, E. (2006) The type II signal peptidase of Legionella pneumophila. *Research in Microbiology* **157**: 836-841.

20. Tjalsma, H., Kontinen, V.P., Prágai, Z., Wu, H., Meima, R., Venema, G., Bron, S., Sarvas, M. and van Dijl, J.M. (1999) The role of lipoprotein processing by signal peptidase II in the Gram-positive eubacterium bacillus subtilis. Signal peptidase II is required for the efficient secretion of alpha-amylase, a non-lipoprotein. *J Biol Chem* **274**: 1698-1707.

21. Paitan, Y., Orr, E., Ron, E.Z. and Rosenberg, E. (1999) A nonessential signal peptidase II (Lsp) of Myxococcus xanthus might be involved in biosynthesis of the polyketide antibiotic TA. *Journal of Bacteriology* **181**: 5644-5651.

22. Miller, K.W. and Wu, H.C. (1987) Cotranscription of the Escherichia coli isoleucyl-tRNA synthetase (ileS) and prolipoprotein signal peptidase (lsp) genes. Fine-structure mapping of the lsp internal promoter. *J Biol Chem* **262**: 389-393.

23. Giménez, M.I., Dilks, K. and Pohlschröder, M. (2007) Haloferax volcanii twin-arginine translocation substates include secreted soluble, C-terminally anchored and lipoproteins. *Mol Microbiol* **66**: 1597-1606.

24. Baumgärtner, M., Kärst, U., Gerstel, B., Loessner, M., Wehland, J. and Jänsch, L. (2007) Inactivation of Lgt allows systematic characterization of lipoproteins from Listeria monocytogenes. *Journal of Bacteriology* **189**: 313-324.

25. Prágai, Z., Tjalsma, H., Bolhuis, A., van Dijl, J.M., Venema, G. and Bron, S. (1997) The signal peptidase II (lsp) gene of Bacillus subtilis. *Microbiology (Reading, Engl)* **143 (Pt 4)**: 1327-1333.

26. Tjalsma, H., Zanen, G., Venema, G., Bron, S. and van Dijl, J.M. (1999) The potential active site of the lipoprotein-specific (type II) signal peptidase of Bacillus subtilis. *J Biol Chem* **274**: 28191-28197.

27. Rahman, M.S., Ceraul, S.M., Dreher-Lesnick, S.M., Beier, M.S. and Azad, A.F. (2007) The lspA gene, encoding the type II signal peptidase of Rickettsia typhi: transcriptional and functional analysis. *Journal of Bacteriology* **189**: 336-341.

28. De Greeff, A., Hamilton, A., Sutcliffe, I.C., Buys, H., Van Alphen, L. and Smith, H.E. (2003) Lipoprotein signal peptidase of Streptococcus suis serotype 2. *Microbiology (Reading, Engl)* **149**: 1399-1407.

29. Robichon, C., Vidal-Ingigliardi, D. and Pugsley, A.P. (2005) Depletion of apolipoprotein N-acyltransferase causes mislocalization of outer membrane lipoproteins in Escherichia coli. *J Biol Chem* **280**: 974-983.

30. Vidal-Ingigliardi, D., Lewenza, S. and Buddelmeijer, N. (2007) Identification of essential residues in apolipoprotein N-acyl transferase, a member of the CN hydrolase family. *Journal of Bacteriology* **189**: 4456-4464.

31. Pace, H.C. and Brenner, C. (2001) The nitrilase superfamily: classification, structure and function. *Genome Biol* **2**: REVIEWS0001.

32. Gupta, S.D., Dowhan, W. and Wu, H.C. (1991) Phosphatidylethanolamine is not essential for the N-acylation of apolipoprotein in Escherichia coli. *J Biol Chem* **266**: 9983-9986.

33. Buddelmeijer, N. and Young, R. (2010) The essential Escherichia coli apolipoprotein N-acyltransferase (Lnt) exists as an extracytoplasmic thioester acyl-enzyme intermediate. *Biochemistry* **49**: 341-346.

34. Hayashi, S., Chang, S.Y., Chang, S., Giam, C.Z. and Wu, H.C. (1985) Modification and processing of internalized signal sequences of prolipoprotein in Escherichia coli and in Bacillus subtilis. *J Biol Chem* **260**: 5753-5759.

35. Navarre, W.W., Daefler, S. and Schneewind, O. (1996) Cell wall sorting of lipoproteins in Staphylococcus aureus. *J Bacteriol* **178**: 441-446.

36. Kurokawa, K., Lee, H., Roh, K.-B., Asanuma, M., Kim, Y.S., Nakayama, H., Shiratsuchi, A., Choi, Y., Takeuchi, O., Kang, H.J., Dohmae, N., Nakanishi, Y., Akira, S., Sekimizu, K. and Lee, B.L. (2009) The Triacylated ATP Binding Cluster Transporter Substrate-binding Lipoprotein of Staphylococcus aureus Functions as a Native Ligand for Toll-like Receptor 2. *J Biol Chem* **284**: 8406-8411.

37. Gurcha, S.S., Baulard, A.R., Kremer, L., Locht, C., Moody, D.B., Muhlecker, W., Costello, C.E., Crick, D.C., Brennan, P.J. and Besra, G.S. (2002) Ppm1, a novel polyprenol monophosphomannose synthase from Mycobacterium tuberculosis. *Biochem J* **365**: 441-450.

38. Tschumi, A., Nai, C., Auchli, Y., Hunziker, P., Gehrig, P., Keller, P., Grau, T. and Sander, P. (2009) Identification of apolipoprotein n-acyltransferase (LNT) in mycobacteria. *J Biol Chem*.

39. Tokuda, H. and Matsuyama, S.-I. (2004) Sorting of lipoproteins to the outer membrane in E. coli. *Biochim Biophys Acta* **1693**: 5-13.

40. Pathania, R., Zlitni, S., Barker, C., Das, R., Gerritsma, D.A., Lebert, J., Awuah, E., Melacini, G., Capretta, F.A. and Brown, E.D. (2009) Chemical genomics in Escherichia coli identifies an inhibitor of bacterial lipoprotein targeting. *Nat Chem Biol* **5**: 849-856.

41. Watanabe, S., Oguchi, Y., Takeda, K., Miki, K. and Tokuda, H. (2008) Introduction of a Lethal Redox Switch That Controls the Opening and Closing of the Hydrophobic Cavity in LolA. *J Biol Chem* **283**: 25421-25427.

42. Okuda, S. and Tokuda, H. (2009) Model of mouth-to-mouth transfer of bacterial lipoproteins through inner membrane LolC, periplasmic LolA, and outer membrane LolB. *Proc Natl Acad Sci USA* **106**: 5877-5882.

43. Fukuda, A., Matsuyama, S.-I., Hara, T., Nakayama, J., Nagasawa, H. and Tokuda, H. (2002) Aminoacylation of the N-terminal cysteine is essential for Lol-dependent release of lipoproteins from membranes but does not depend on lipoprotein sorting signals. *J Biol Chem* **277**: 43512-43518.

44. Tanaka, K., Matsuyama, S.I. and Tokuda, H. (2001) Deletion of lolB, encoding an outer membrane lipoprotein, is lethal for Escherichia coli and causes accumulation of lipoprotein localization intermediates in the periplasm. *Journal of Bacteriology* **183**: 6538-6542.

45. Seydel, A., Gounon, P. and Pugsley, A.P. (1999) Testing the '+2 rule' for lipoprotein sorting in the Escherichia coli cell envelope with a new genetic selection. *Mol Microbiol* **34**: 810-821.

46. Hara, T., Matsuyama, S. and Tokuda, H. (2003) Mechanism underlying the inner membrane retention of Escherichia coli lipoproteins caused by Lol avoidance signals. *J Biol Chem* **278**: 40408-40414.

47. Schulze, R.J. and Zuckert, W.R. (2006) Borrelia burgdorferi lipoproteins are secreted to the outer surface by default. *Mol Microbiol* **59**: 1473-1484.

48. Sankaran, K. and Wu, H.C. (1994) Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *J Biol Chem* **269**: 19701-19706.

49. Leskelä, S., Wahlström, E., Kontinen, V.P. and Sarvas, M. (1999) Lipid modification of prelipoproteins is dispensable for growth but essential for efficient protein secretion in Bacillus subtilis: characterization of the Lgt gene. *Mol Microbiol* **31**: 1075-1085.

50. Henneke, P., Dramsi, S., Mancuso, G., Chraibi, K., Pellegrini, E., Theilacker, C., Hübner, J., Santos-Sierra, S., Teti, G., Golenbock, D.T., Poyart, C. and Trieu-Cuot, P. (2008) Lipoproteins are critical TLR2 activating toxins in group B streptococcal sepsis. *J Immunol* **180**: 6149-6158.

51. Réglier-Poupet, H., Frehel, C., Dubail, I., Beretti, J.-L., Berche, P., Charbit, A. and Raynaud, C. (2003) Maturation of lipoproteins by type II signal peptidase is required for phagosomal escape of Listeria monocytogenes. *J Biol Chem* **278**: 49469-49477.

52. Venema, R., Tjalsma, H., van Dijl, J.M., de Jong, A., Leenhouts, K., Buist, G. and Venema, G. (2003) Active lipoprotein precursors in the Gram-positive eubacterium Lactococcus lactis. *J Biol Chem* **278**: 14739-14746.

53. Denham, E.L., Ward, P.N. and Leigh, J.A. (2008) Lipoprotein signal peptides are processed by Lsp and Eep of Streptococcus uberis. *Journal of Bacteriology* **190**: 4641-4647.

54. Hoffmann, C., Leis, A., Niederweis, M., Plitzko, J.M. and Engelhardt, H. (2008) Disclosure of the mycobacterial outer membrane: cryo-electron tomography and vitreous sections reveal the lipid bilayer structure. *Proc Natl Acad Sci USA* **105**: 3963-3967.

55. Matias, V.R.F. and Beveridge, T.J. (2006) Native cell wall organization shown by cryo-electron microscopy confirms the existence of a periplasmic space in Staphylococcus aureus. *Journal of Bacteriology* **188**: 1011-1021.

56. Matias, V.R.F. and Beveridge, T.J. (2005) Cryo-electron microscopy reveals native polymeric cell wall structure in Bacillus subtilis 168 and the existence of a periplasmic space. *Mol Microbiol* **56**: 240-251.

57. Perry, J., Staley, J. and Lory, S., (2002) *Microbial Life*. Sinauer Associates.

58. Sutcliffe, I.C. and Russell, R.R. (1995) Lipoproteins of gram-positive bacteria. *J Bacteriol* **177**: 1123-1128.

59. Bubeck Wardenburg, J., Williams, W.A. and Missiakas, D. (2006) Host defenses against Staphylococcus aureus infection require recognition of bacterial lipoproteins. *Proc Natl Acad Sci USA* **103**: 13831-13836.

60. Hamilton, A., Robinson, C., Sutcliffe, I.C., Slater, J., Maskell, D.J., Davis-Poynter, N., Smith, K., Waller, A. and Harrington, D.J. (2006) Mutation of the maturase lipoprotein attenuates the virulence of Streptococcus equi to a greater extent than does loss of general lipoprotein lipidation. *Infection and Immunity* **74**: 6907-6919.

61. Sander, P., Rezwan, M., Walker, B., Rampini, S.K., Kroppenstedt, R.M., Ehlers, S., Keller, C., Keeble, J.R., Hagemeier, M., Colston, M.J., Springer, B. and Böttger, E.C. (2004) Lipoprotein processing is required for virulence of Mycobacterium tuberculosis. *Mol Microbiol* **52**: 1543-1552.

62. Schenk, M., Belisle, J.T. and Modlin, R.L. (2009) TLR2 looks at lipoproteins. *Immunity* **31**: 847-849.

63. Babu, M.M., Priya, M.L., Selvan, A.T., Madera, M., Gough, J., Aravind, L. and Sankaran, K. (2006) A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *Journal of Bacteriology* **188**: 2761-2773.

64. Sutcliffe, I.C. and Harrington, D.J. (2002) Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. *Microbiology* **148**: 2065-2077.

65. Elliot, M.A., Buttner, M.J. and Nodwell, J.R., (2007) Multicellular development in Streptomyces Myxobacteria: Multicellularity and Differentiation. In. W.D. E. (ed). ASM Press, pp. 419-438.

66. Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A.-M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.-H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.-A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J. and

Hopwood, D.A. (2002) Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* **417**: 141-147.

67. Loria, R., Kers, J. and Joshi, M. (2006) Evolution of plant pathogenicity in Streptomyces. *Annu. Rev. Phytopathol.* **44**: 469-487.

68. Miguelez, E.M., Hardisson, C. and Manzanal, M.B. (1999) Hyphal death during colony development in Streptomyces antibioticus: morphological evidence for the existence of a process of cell deletion in a multicellular prokaryote. *J Cell Biol* **145**: 515-525.

69. Straight, P.D., Willey, J.M. and Kolter, R. (2006) Interactions between Streptomyces coelicolor and Bacillus subtilis: Role of surfactants in raising aerial structures. *Journal of Bacteriology* **188**: 4918-4925.

70. Elliot, M.A., Karoonuthaisiri, N., Huang, J., Bibb, M.J., Cohen, S.N., Kao, C.M. and Buttner, M.J. (2003) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in Streptomyces coelicolor. *Genes & Development* **17**: 1727-1740.

71. Bentley, S.D., Brosch, R., Gordon, S.V., Hopwood, D.A., Cole, S.T. (2004) Genomics of Actinobacteria, The High G+C Gram-Positive Bacteria. *Microbial Genomes*: 333-360.

72. Hutchings, M.I., Hoskisson, P.A., Chandra, G. and Buttner, M.J. (2004) Sensing and responding to diverse extracellular signals? Analysis of the sensor kinases and response regulators of Streptomyces coelicolor A3 (2). *Microbiology* **150**: 2795-2806.

73. Komatsu, M., Takano, H., Hiratsuka, T., Ishigaki, Y., Shimada, K., Beppu, T. and Ueda, K. (2006) Proteins encoded by the conservon of Streptomyces coelicolor A3(2) comprise a membrane-associated

heterocomplex that resembles eukaryotic G protein-coupled regulatory system. *Mol Microbiol* **62**: 1534-1546.

74. Han, L., Dutilleul, P., Prasher, S., Beaulieu, C. and Smith, D. (2008) Assessment of Common Scab-Inducing Pathogen Effects on Potato Underground Organs Via Computed Tomography Scanning. *Phytopathology* **98**: 1118-1125.

75. Joshi, M., Rong, X., Moll, S., Kers, J., Franco, C. and Loria, R. (2007) Streptomyces turgidiscabies secretes a novel virulence protein, Nec1, which facilitates infection. *Mol Plant Microbe Interact* **20**: 599-608.

76. Kers, J.A., Cameron, K.D., Joshi, M.V., Bukhalid, R.A., Morello, J.E., Wach, M.J., Gibson, D.M. and Loria, R. (2005) A large, mobile pathogenicity island confers plant pathogenicity on Streptomyces species. *Mol Microbiol* **55**: 1025-1033.

77. Loria, R., Coombs, J., Yoshida, M., Kers, J. and Bukhalid, R. (2003) A paucity of bacterial root diseases: Streptomyces succeeds where others fail. *Physiological and Molecular Plant Pathology* **62**: 65-72.

78. Clark, C. and Matthews, S. (1987) Histopathology of sweet potato root infection by Streptomyces ipomoea. *Phytopathology* **77**: 1418-1423.

79. Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F. and Hopwood, D.A., (2000) *Practical streptomyces genetics*. John Innes Foundation.

80. Joshi, M.V., Mann, S.G., Antelmann, H., Widdick, D.A., Fyans, J.K., Chandra, G., Hutchings, M.I., Toth, I., Hecker, M., Loria, R. and Palmer, T. (2010) The Tat pathway exports multiple virulence proteins in the plant pathogen Streptomyces scabies. *Mol Microbiol* **Accepted article**.

81. Sun, J., Kelemen, G.H., Fernández-Abalos, J.M. and Bibb, M.J. (1999) Green fluorescent protein as a reporter for spatial and temporal gene

expression in Streptomyces coelicolor A3(2). *Microbiology (Reading, Engl)* **145 (Pt 9)**: 2221-2227.

82. Thompson, B.J., Widdick, D.A., Hicks, M.G., Chandra, G., Sutcliffe, I.C., Palmer, T. and Hutchings, M.I. (2010) Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor Mol Microbiol* **Accepted Article**.

83. Redenbach, M., Kieser, H.M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H. and Hopwood, D.A. (1996) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome. *Mol Microbiol* **21**: 77-96.

84. Gust, B., Chandra, G., Jakimowicz, D., Yuqing, T., Bruton, C.J. and Chater, K.F. (2004) Lambda red-mediated genetic manipulation of antibiotic-producing Streptomyces. *Adv Appl Microbiol* **54**: 107-128.

85. Rahman, O., Cummings, S.P., Harrington, D.J. and Sutcliffe, I.C. (2008) Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. *World Journal of Microbiology & Biotechnology* **24**: 2377-2382.

86. Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**: 783-795.

87. Kall, L., Krogh, A. and Sonnhammer, E.L. (2004) A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**: 1027-1036.

88. Juncker, A.S., Willenbrock, H., Von Heijne, G., Brunak, S., Nielsen, H. and Krogh, A. (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**: 1652-1662.

89. Bagos, P.G., Tsirigos, K.D., Liakopoulos, T.D. and Hamodrakas, S.J. (2008) Prediction of lipoprotein signal peptides in Gram-positive bacteria with a Hidden Markov Model. *J Proteome Res* **7**: 5082-5093.

90. Rose, R.W., Bruser, T., Kissinger, J.C. and Pohlschroder, M. (2002) Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol* **45**: 943-950.

91. Bendtsen, J.D., Nielsen, H., Widdick, D., Palmer, T. and Brunak, S. (2005) Prediction of twin-arginine signal peptides. *BMC Bioinformatics* **6**: 167.

92. Paget, M.S., Chamberlin, L., Atrih, A., Foster, S.J. and Buttner, M.J. (1999) Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in Streptomyces coelicolor A3(2). *Journal of Bacteriology* **181**: 204-211.

93. Hutchings, M.I., Hong, H.-J., Leibovitz, E., Sutcliffe, I.C. and Buttner, M.J. (2006) The sigma(E) cell envelope stress response of Streptomyces coelicolor is influenced by a novel lipoprotein, CseA. *Journal of Bacteriology* **188**: 7222-7229.

94. Ming, L.-J. (2003) Structure and function of "metalloantibiotics". *Med. Res. Rev.* **23**: 697-762.

95. Jordan, S., Hutchings, M.I. and Mascher, T. (2008) Cell envelope stress response in Gram-positive bacteria. *FEMS Microbiology Reviews* **32**: 107-146.

96. Schuck, S., Honsho, M., Ekroos, K., Shevchenko, A. and Simons, K. (2003) Resistance of cell membranes to different detergents. *Proc Natl Acad Sci USA* **100**: 5795-5800.

97. Nakimbugwe, D., Masschalck, B., Deckers, D., Callewaert, L., Aertsen, A. and Michiels, C.W. (2006) Cell wall substrate specificity of six different lysozymes and lysozyme inhibitory activity of bacterial extracts. *FEMS Microbiol Lett* **259**: 41-46.

98. Callewaert, L., Aertsen, A., Deckers, D., Vanoirbeek, K.G., Vanderkelen, L., Van Herreweghe, J.M., Masschalck, B., Nakimbugwe, D., Robben, J. and Michiels, C.W. (2008) A new family of lysozyme inhibitors contributing to lysozyme tolerance in gram-negative bacteria. *PLoS Pathog* **4**: e1000019.

99. Denham, E.L., Ward, P.N. and Leigh, J.A. (2009) In the absence of Lgt, lipoproteins are shed from Streptococcus uberis independently of Lsp. *Microbiology (Reading, Engl)* **155**: 134-141.

100. Pahl, A. and Keller, U. (1994) Streptomyces chrysomallus FKBP-33 is a novel immunophilin consisting of two FK506 binding domains; its gene is transcriptionally coupled to the FKBP-12 gene. *EMBO J* **13**: 3472-3480.

101. Ogawara, H. (1975) Production and property of beta-lactamases in Streptomyces. *Antimicrob Agents Chemother* **8**: 402-408.

102. Silver, L.L. (2003) Novel inhibitors of bacterial cell wall synthesis. *Curr Opin Microbiol* **6**: 431-438.

103. Bernard, R., Guiseppi, A., Chippaux, M., Foglino, M. and Denizot, F. (2007) Resistance to bacitracin in Bacillus subtilis: unexpected requirement of the BceAB ABC transporter in the control of expression of its own structural genes. *Journal of Bacteriology* **189**: 8636-8642.

104. Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A. and Charlier, P. (2008) The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews* **32**: 234-258.

105. Zapun, A., Contreras-Martel, C. and Vernet, T. (2008) Penicillin-binding proteins and beta-lactam resistance. *FEMS Microbiology Reviews* **32**: 361-385.

106. Bielnicki, J., Devedjiev, Y., Derewenda, U., Dauter, Z., Joachimiak, A. and Derewenda, Z.S. (2006) B. subtilis ykuD protein at 2.0 A resolution: insights into the structure and function of a novel, ubiquitous family of bacterial enzymes. *Proteins* **62**: 144-151.

107. Biarrotte-Sorin, S., Hugonnet, J.-E., Delfosse, V., Mainardi, J.-L., Gutmann, L., Arthur, M. and Mayer, C. (2006) Crystal structure of a novel beta-lactam-insensitive peptidoglycan transpeptidase. *Journal of Molecular Biology* **359**: 533-538.

108. Magnet, S., Arbeloa, A., Mainardi, J.-L., Hugonnet, J.-E., Fourgeaud, M., Dubost, L., Marie, A., Delfosse, V., Mayer, C., Rice, L.B. and Arthur, M. (2007) Specificity of L,D-transpeptidases from gram-positive bacteria producing different peptidoglycan chemotypes. *J Biol Chem* **282**: 13151-13159.

109. Magnet, S., Bellais, S., Dubost, L., Fourgeaud, M., Mainardi, J.-L., Petit-Frère, S., Marie, A., Mengin-Lecreulx, D., Arthur, M. and Gutmann, L. (2007) Identification of the L,D-transpeptidases responsible for attachment of the Braun lipoprotein to Escherichia coli peptidoglycan. *Journal of Bacteriology* **189**: 3927-3931.

110. Lavollay, M., Arthur, M., Fourgeaud, M., Dubost, L., Marie, A., Veziris, N., Blanot, D., Gutmann, L. and Mainardi, J.-L. (2008) The peptidoglycan of stationary-phase Mycobacterium tuberculosis predominantly contains cross-links generated by L,D-transpeptidation. *Journal of Bacteriology* **190**: 4360-4366.

111. Gupta, R., Lavollay, M., Mainardi, J.L., Arthur, M., Bishai, W.R. and Lamichhane, G. (2010) The Mycobacterium tuberculosis protein LdtMt2 is a nonclassical transpeptidase required for virulence and resistance to amoxicillin. *Nat Med* **16**: 466-469.

112. Leyh-Bouille, M., Bonaly, R., Ghuysen, J.M., Tinelli, R. and Tipper, D. (1970) LL-diaminopimelic acid containing peptidoglycans in walls of Streptomyces sp. and of Clostridium perfringens (type A). *Biochemistry* **9**: 2944-2952.

113. Loria, R., Bignell, D., Moll, S., Huguet-Tapia, J., Joshi, M., Johnson, E., Seipke, R. and Gibson, D. (2008) Thaxtomin biosynthesis: the path to plant pathogenicity in the genus Streptomyces. *Antonie van Leeuwenhoek* **94**: 3-10.

114. Cullen, D.W. and Lees, A.K. (2007) Detection of the nec1 virulence gene and its correlation with pathogenicity in Streptomyces species on potato tubers and in soil using conventional and real-time PCR. *J Appl Microbiol* **102**: 1082-1094.

115. Kado, C. (2009) Horizontal gene transfer: sustaining pathogenicity and optimizing host-pathogen interactions. *Mol Plant Pathol* **10**: 143-150.

116. Healy, F., Wach, M., Krasnoff, S., Gibson, D. and Loria, R. (2000) The txtAB genes of the plant pathogen Streptomyces acidiscabies encode a peptide synthetase required for phytotoxin thaxtomin A production and pathogenicity. *Mol Microbiol* **38**: 794-804.

117. Lerat, S., Simao-Beaunoir, A.M. and Beaulieu, C. (2009) Genetic and physiological determinants of Streptomyces scabies pathogenicity. *Mol Plant Pathol* **10**: 579-585.

118. Kers, J.A., Wach, M.J., Krasnoff, S.B., Widom, J., Cameron, K.D., Bukhalid, R.A., Gibson, D.M., Crane, B.R. and Loria, R. (2004) Nitration of a peptide phytotoxin by bacterial nitric oxide synthase. *Nature* **429**: 79-82.

119. Adak, S., Aulak, K. and Stuehr, D. (2002) Direct evidence for nitric oxide production by a nitric-oxide synthase-like protein from Bacillus subtilis. *Journal of Biological Chemistry* **277**: 16167-16171.

120. McCann, H. and Guttman, D. (2008) Evolution of the type III secretion system and its effectors in plant-microbe interactions. *New Phytologist* **177**: 33-47.

121. Simeone, R., Bottai, D. and Brosch, R. (2009) ESX/type VII secretion systems and their role in host-pathogen interaction. *Curr Opin Microbiol* **12**: 4-10.

122. Widdick, D., Thompson, B.J., Hicks, M.G., Chandra, G., Bignell, D., Loria, R., Sutcliffe, I.C. and Hutchings, M.I. (2010) Analysing the biogenesis of lipoproteins and their role in virulence in the plant pathogen Streptomyces scabies. *Manuscript in Preparation*.

123. Tresner, H.D., Davies, M.C. and Backus, E.J. (1961) Electron microscopy of Streptomyces spore morphology and its role in species differentiation. *J Bacteriol* **81**: 70-80.

124. Lefèvre, P., Braibant, M., de Wit, L., Kalai, M., Röeper, D., Grötzinger, J., Delville, J.P., Peirs, P., Ooms, J., Huygen, K. and Content, J. (1997) Three different putative phosphate transport receptors are encoded by the Mycobacterium tuberculosis genome and are present at the surface of Mycobacterium bovis BCG. *Journal of Bacteriology* **179**: 2900-2906.

125. Qi, Y., Kobayashi, Y. and Hulett, F.M. (1997) The pst operon of Bacillus subtilis has a phosphate-regulated promoter and is involved in phosphate transport but not in regulation of the pho regulon. *Journal of Bacteriology* **179**: 2534-2539.

126. Díaz, M., Esteban, A., Fernández-Abalos, J.M. and Santamaría, R.I. (2005) The high-affinity phosphate-binding protein PstS is accumulated under high fructose concentrations and mutation of the corresponding gene affects differentiation in Streptomyces lividans. *Microbiology (Reading, Engl)* **151**: 2583-2592.

127. Loria, R., Bukhalid, R.A., Creath, R.A., Leiner, R.H., Olivier, M. and Steffens, J.C. (1995) Differential Production of Thaxtomins by Pathogenic Streptomyces Species In Vitro. *Phytopathology*: 537-541.

128. Parro, V., Schacht, S., Anne, J. and Mellado, R.P. (1999) Four genes encoding different type I signal peptidases are organized in a cluster in Streptomyces lividans TK21. *Microbiology* **145 (Pt 9)**: 2255-2263.

129. Gregory, M.A., Till, R. and Smith, M.C.M. (2003) Integration site for Streptomyces phage phiBT1 and development of site-specific integrating vectors. *Journal of Bacteriology* **185**: 5320-5323.

130. Wach, M.J., Krasnoff, S.B., Loria, R. and Gibson, D.M. (2007) Effect of carbohydrates on the production of thaxtomin A by Streptomyces acidiscabies. *Arch Microbiol* **188**: 81-88.

131. Barona-Gomez, F., Lautru, S., Francou, F.X., Leblond, P., Pernodet, J.L. and Challis, G.L. (2006) Multiple biosynthetic and uptake systems mediate siderophore-dependent iron acquisition in Streptomyces coelicolor A3(2) and Streptomyces ambofaciens ATCC 23877. *Microbiology* **152**: 3355-3366.

132. Li, H., Jacques, P.E., Ghinet, M.G., Brzezinski, R. and Morosoli, R. (2005) Determining the functionality of putative Tat-dependent signal peptides in Streptomyces coelicolor A3(2) by using two different reporter proteins. *Microbiology* **151**: 2189-2198.

133. Yeats, C., Bentley, S. and Bateman, A. (2003) New knowledge from old: in silico discovery of novel protein domains in Streptomyces coelicolor. *BMC Microbiol* **3**: 3.

134. Gordon, N.D., Ottaviano, G.L., Connell, S.E., Tobkin, G.V., Son, C.H., Shterental, S. and Gehring, A.M. (2008) Secreted-protein response to sigmaU activity in Streptomyces coelicolor. *Journal of Bacteriology* **190**: 894-904.

135. van Wezel, G.P., White, J., Bibb, M.J. and Postma, P.W. (1997) The malEFG gene cluster of Streptomyces coelicolor A3(2): characterization, disruption and transcriptional analysis. *Mol Gen Genet* **254**: 604-608.

136. Lun, S. and Bishai, W.R. (2007) Characterization of a novel cell wall-anchored protein with carboxylesterase activity required for virulence in Mycobacterium tuberculosis. *J Biol Chem* **282**: 18348-18356.

137. Park, H.S., Shin, S.K., Yang, Y.Y., Kwon, H.J. and Suh, J.W. (2005) Accumulation of S-adenosylmethionine induced oligopeptide transporters including BldK to regulate differentiation events in Streptomyces coelicolor M145. *FEMS Microbiol Lett* **249**: 199-206.

138. Bertram, R., Schlicht, M., Mahr, K., Nothaft, H., Saier, M.H. and Titgemeyer, F. (2004) In silico and transcriptional analysis of carbohydrate uptake systems of Streptomyces coelicolor A3(2). *Journal of Bacteriology* **186**: 1362-1373.

139. Sugawara, K., Dohmae, N., Kasai, K., Saido-Sakanaka, H., Okamoto, S., Takio, K. and Ochi, K. (2002) Isolation and identification of novel ADP-ribosylated proteins from Streptomyces coelicolor A3(2). *Biosci Biotechnol Biochem* **66**: 2292-2296.

Table A1. List of putative *S. coelicolor* lipoproteins.

Swiss-Prot	Verdict	Functional analysis	Genomic context
SCO0026	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO0065	Unclear	**SBP PF01547▲**	**In operon with SCO0066-SCO0066 but locus incomplete; adjacent to sugar isomerase SCO0063 and glycosyl hydrolase SCO0068**
SCO0273	Lpp	**SBP PF01547▲**	**Linked with SCO0271-SCO0272 but locus incomplete**
SCO0290	Lpp	**SBP PF01547, BxlE2▲**	**In operon with SCO0291 but locus incomplete**
SCO0301	Lpp	enzyme - esterase/lyase?	Monomer
SCO0352	Lpp	**SBP PF00528 sugars?▲**	**In operon with SCO0353-SCO0354 but locus incomplete**
SCO0389	Lpp	**CASH domain - carbohydrate active enzyme?**	**Part of large operon SCO0387-SCO0401; polysacch synthesis ?**
SCO0449	Lpp	**SBP rhamnose/sugars?**	**Monomer, locus incomplete**
SCO0453	Lpp	**SBP PF01547▲**	**Linked with SCO0454-SCO0456 but locus incomplete**
SCO0472	Lpp	Matches InterPro 11044 Quinohemoprotein amine dehydrogenase ?	SCO0472-0474 = 3 tandem & overlapping Lpp ORFs, in operon with SCO0475 ABC perm/ATP fusion
SCO0473	Lpp	**SBP PF01297, divalent cation substrate**	**SCO0472-0474 = 3 tandem & overlapping Lpp ORFs, in operon with SCO0475 ABC perm/ATP fusion**
SCO0474	Lpp	Matches InterPro 11044 Quinohemoprotein amine dehydrogenase ?	SCO0472-0474 = 3 tandem & overlapping Lpp ORFs, in operon with SCO0475 ABC perm/ATP fusion
SCO0494	Lpp	**SBP PF01497 Fe-siderophore**	**Locus complete, in operon with SCO0493-0498; CchF coelichelin uptake system [131].**
SCO0516	Lpp	Matches InterPro 11044 Quinohemoprotein amine dehydrogenase ? Beta rpt protein	Monomer
SCO0528	Lpp	Conserved hypothetical; highly homologous to SCO1146	Monomer
SCO0531	Lpp	**SBP PF01547▲**	**Linked with SCO0532-SCO0533 but locus incomplete**
SCO0538	Lpp	**SBP PF01547▲; TAT secreted [132].**	**In operon with SCO0537-SCO0540 but locus incomplete**
SCO0607	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO0638	Lpp	PF0249 fascilin domain protein, MPT83 homologue	Monomer
SCO0660	Lpp	**SBP PF01547▲; also has ankyrin domains?**	**In operon with SCO0661-SCO0664 but locus incomplete**
SCO0706	Lpp	**SBP, matches COG0683 LivK**	**Locus complete, in operon with SCO0707-0710 for ABC branched chain aa uptake**
SCO0768	Lpp	Conserved hypothetical, limited actinomycete distribution	Monomer
SCO0798	Unclear	Conserved hypothetical, matches COG2706 carboxymuconate cyclase	In operon with SCO0797 conserved hypothetical
SCO0808	Lpp	**SBP rhamnose/pentoses; RbsE3▲**	**Locus complete, in operon with SCO0809-0811; adjacent to putative sugar isomerase**
SCO0914	Lpp	**SBP PF01547▲**	**Monomer, locus incomplete**
SCO0930	Lpp	**PF03640 Y-X4-D motif protein, as is SCO7673 [133]**	**Monomer; part of SigU regulon [134]**
SCO0933	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO0941	Lpp	Conserved hypothetical, distant relationship to poly-gamma-glutamate biosynthesis prote	Monomer
SCO0952	Lpp	**SBP PF01547▲**	**In operon with SCO0949-SCO0951 but locus incomplete, includes conserved hypothetical**
SCO0996	Lpp	**SBP PF01497 Fe-siderophore**	**Locus complete, in operon with SCO0997-0998**
SCO1056	Lpp	**SBP PF01547▲**	**In operon with SCO1057-SCO1059 but locus incomplete (includes beta-glucosidase)**
SCO1065	Lpp	**SBP PF01547▲**	**In operon with SCO1063-SCO1064 but locus incomplete**
SCO1106	Lpp	PF03713 DUF305 protein with conserved HH motif [133]	Monomer
SCO1146	Lpp	Conserved hypothetical; highly homologous to SCO0528	Monomer
SCO1341	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO1344	Lpp	**SBP PF00496 family 5**	**Monomer, locus incomplete**
SCO1356	Lpp	**PF00355 Fe-S reiske domain protein**	**In operon with SCO1357 cytoplasmic pyridoxamine 5'-phosphate oxidase; part of SigU regulon [134]**
SCO1539	Lpp	**SBP PF01547▲**	**In operon with SCO1538-SCO1539 but locus incomplete**
SCO1557	Lpp	**SBP PF03180 NlpA**	**Locus completed by SCO1558-1559**
SCO1639	Lpp	PF00254 FKBP-type PPIase (homologue of FKBP-33 [100])	Monomer
SCO1655	Lpp	**SBP PF00496 family 5**	**Monomer, locus incomplete**
SCO1714	Lpp	Conserved hypothetical, Streptomyces restricted?	In operon with SCO1713 small orf?
SCO1763	Lpp	PF00355 Fe-S reiske domain protein	Monomer
SCO1882	Lpp	**SBP PF01547▲**	**Monomer, linmked to locus containing SCO1885-SCO1886 but locus incomplete**
SCO1898	Unclear	**SBP PF01547, SmoE▲**	**In operon with SCO1899-SCO1902 but locus incomplete**
SCO1940	Lpp	thioredoxin-like	Monomer
SCO1955	Lpp	PF00355 Fe-S reiske domain protein	Monomer; TAT secreted [132]
SCO2008	Lpp	SBP, PF01094 branched chain amino acids■	Locus complete, in operon with SCO2009-2012 for ABC branched chain aa uptake
SCO2096	Lpp	Protease, PF01841 transglutaminase family	In operon with SCO2094 regulator and SCO2095 DUF58 protein
SCO2134ext	Lpp	Conserved hypothetical, limited actinomycete distribution	
SCO2153	Lpp	PF03734 YkuD/COG1376 ErfK family, putative peptidoglycan active enzyme	Monomer
SCO2156	Unclear	**CtaC**	**Component of cytochrome bc1-aa3 supercomplex**

SCO2171	Lpp	Putative peptidase with PDZ putative target binding domain (PF00595)	Monomer
SCO2178	Lpp	Conserved hypothetical COG3868, putative carbohydrate active enzyme	Monomer
SCO2218	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO2231	Lpp	**SBP PF01547 , MalE [135]▲■**	**Linked to SCO2229-SCO2230 but locus incomplete**
SCO2272	Lpp	**SBP PF01497 Fe-siderophore**	**Locus complete, in operon with SCO2273-SCO2274**
SCO2275	Lpp	**PF09375 Peptidase M75 (imelysin) family**	**In operon with SCO2276 Dyp-peroxidase and SCO2277 Ftr1 family protein**
SCO2278ext	Unclear	Conserved hypothetical, Streptomyces restricted?	Homologue of SCO6871; overlap with SCO2275 Lpp-SCO2277 operon
SCO2355	Lpp	Conserved hypothetical	Monomer
SCO2392	Lpp	PF00657 Lipase	In operon with SCO2393 putative aldose-1-epimerase
SCO2404	Lpp	**SBP PF00532 family 1, pentoses▲**	**Locus complete, in operon with SCO2405-2406; adjacent to putative aldose epimerase**
SCO2457	Lpp	Conserved hypothetical, limited actinomycete distribution	In operon with SCO2458
SCO2505	Lpp	**SBP PF01297, divalent cation substrate**	**Locus completed by SCO2506-SCO2507**
SCO2515	Lpp	Hypothetical protein, unique to SCO	In operon with SCO2513 transcriptional regulator and SCO2514
SCO2550	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO2658	Lpp	**SBP PF01547▲**	**In operon with SCO2659-SCO2661 but locus incomplete**
SCO2673	Lpp	**PF02687 DUF214 permease, 9 other MSD**	**Monomer**
SCO2679	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO2721	Lpp	PF0092 Von Willebrand factor domain (ligand binding?)	Monomer
SCO2725	Lpp	Conserved hypothetical, limited actinomycete distribution	Monomer; Family with SCO4650, SCO4651
SCO2760	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO2764	Unclear	**Conserved hypothetical COG4222**	**Monomer; adjacent to fused permease/ABC protein SCO2763**
SCO2780	Lpp	**SBP PF01497 Fe-siderophore**	**Linked to adjacent SCO2781 siderophore interacting protein & SCO2782-2785 siderophore biosynthesis locus; DesE [131].**
SCO2795	Lpp	**SBP PF01547, Ceb1E▲**	**In operon with SCO2796-SCO2797 but locus incomplete**
SCO2828	Unclear	**SBP PF00497 family 3**	**Locus complete, in operon with SCO2829-SCO2832**
SCO2838	Lpp	Glycosyl hydrolase family 6 PF01341	Monomer
SCO2839	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO2903	Unclear	Conserved hypothetical, Streptomyces restricted?	Monomer; SAV5171 homologue likely Lpp
SCO2933	Lpp	**SBP PF04069 ProXL**	**Locus complete, in operon SCO2930-SCO2933**
SCO2946	Lpp	**SBP PF01547▲**	**In operon with SCO2943-SCO2945 but locus incomplete**
SCO2978	Lpp	**SBP PF01547, AglE2▲**	**Linked to SCO2979-SCO2980 but locus incomplete**
SCO3011	Lpp	**LpqB Conserved hypothetical, limited actinomycete distribution**	**In operon with SCO3012-SCO3014 = MtrA-MtrB3CS**
SCO3053	Unclear	Serine protease with lectin-like domain and FG-GAP repeats (ligand binding?)	Monomer
SCO3086	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO3101	Lpp	Conserved hypothetical; very distantly related to N-terminal domain of putative PpiC Acidothermus Ace1_1911	Monomer
SCO3106	Lpp	Conserved hypothetical; homologous to SCO7393 Lpp	Monomer
SCO3107	Lpp	Conserved hypothetical; PF09352 DUF1994 domain	Monomer
SCO3194	Lpp	PF03734 ErfK/YbiS/YnhG family, putative peptidoglycan active enzyme	Monomer
SCO3300rev	Lpp	Conserved hypothetical, Streptomyces restricted? Serine-rich protein	In locus with adjacent SCO3301 and SC3302 integral membrane protein
SCO3301	Lpp	Conserved hypothetical, Streptomyces restricted?	In locus with adjacent SCO3300 and SC3302 integral membrane protein
SCO3310ext	Lpp	homologue of SAV4743 Lpp - unique to Streptomyces?	Nterminal extension requires reannotation
SCO3342	Lpp	Serine Glycine rich protein, Streptomyces restricted?	Monomer
SCO3357	Lpp	CseA proven Lpp	In SigE operon; TCS accessory protein
SCO3431rev	Lpp	Conserved hypothetical	Misannotated in genome with N-terminal extension?
SCO3456	Lpp	**SBP PF01547**	**In operon with SCO3453-SCO3455; TOBE domain in SCO3453 suggest Mo/sulphate or related substrate**
SCO3484	Lpp	**SBP PF01547▲**	**In operon with SCO3481-SCO3483 but locus incomplete; TAT substrate ?**
SCO3505	Lpp	**SBP PF01547▲**	**In operon with SCO3504 but locus incomplete; TAT substrate ?**
SCO3540	Lpp	Putative protease, related to SCO4241, SCO5179 & SCO5180 Lpp	Monomer: MTB Rv2224c homologous putative Lpp is a carboxyesterase [136]
SCO3546	Lpp	Conserved hypothetical, limited actinomycete distribution	Monomer
SCO3561	Lpp	Conserved hypothetical, related to SCO4471 Lpp	Monomer
SCO3667	Lpp	**SBP COG4213 XylF/sugars family**	**Orphan but in operon with SCO3666 DNA-binding regulatory protein**
SCO3674	Lpp	Conserved hypothetical with 5 FG-GAP repeats, ligand binding?	Monomer
SCO3704	Lpp	**SBP PF01547, molybdate?**	**Locus complete, in operon with SCO3705-3706; likely association with SCO3703 TOBE-domain protein**
SCO3707	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO3737	Lpp	**Sortase, homologous to SAV92 putative Lpp sortase**	**In operon with SCO3738, possible DPAGT sortase substrate unique to SCO (1 other MSD)**

SCO3772	Lpp	Putative hydrolase	In operon with SCO3771 penicillin binding protein
SCO3781	Lpp	Conserved hypothetical, Streptomyces restricted? Homologue of SCO6963 Lpp	Monomer
SCO3789	Unclear	PF03372 Endonuclease/Exonuclease/phosphatase family; 2 or 3 MSD topology unclear	In operon with SCO3787 & SCO3788
SCO3965	Lpp	**DUF461 protein, conserved hypothetical**	**TAT Lpp? In operon with SCO3966 Trx-like Lpp; putative Cu resistance locus**
SCO3966	Lpp	**Trx-like fold, SCO domain, cytochrome c biogenesis**	**In putative Cu resistance operon with SCO3965 DUF461 putative Lpp & SCO3964 false +ve**
SCO3968	Unclear	Conserved hypothetical, Streptomyces restricted?	Monomer; part of SCO3965/SCO3966 Cu resistance locus?
SCO3998	Lpp	Conserved hypothetical, related to quinoprotein dehydrogenase?	Monomer
SCO3999	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO4085	Lpp	PF03724 META (DUF306) family protein	In operon with SCO4086-4067 purine biosynthesis proteins PurF & PurM
SCO4095	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO4134	Lpp	FAD-linked oxidase with berberine enzyme domain [alkaloid biosyntheis?]	Monomer
SCO4135	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO4142	Lpp	**SBP PstS phosphate**	**Locus completed by SCO4139-SC4141**
SCO4231	Lpp	Conserved hypothetical, limited actinomycete distribution	Adjacent to but genetically distinct from SCO4229-SCO4230 TCS
SCO4241	Lpp	Putative protease, related to SCO3540, SCO5179 & SCO5180 Lpp	Monomer
SCO4273	Lpp	Zinc metalloprotease, homolgous to poly-lysine degrading enzyme of Str albus	Monomer
SCO4286	Lpp	**SBP PF01547▲**	**Locus incomplete, orphan in operon with SCO4287 conserved hypothetical**
SCO4289	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO4291	Lpp	Hypothetical protein, unique to SCO; TP repeat sequences	Monomer
SCO4328	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO4401	Unclear	**SBP? CD06341 related to SBP for amino acids/peptides**	**Monomer**
SCO4458	Lpp	PF03713 DUF305 protein with conserved HH motif [133]	Monomer
SCO4471	Lpp	Conserved hypothetical, 49% identical to SCO3561 Lpp	Monomer
SCO4472	Lpp	**ResA homologue, cytochrome C biogenesis**	**In operon with SCO4473-SCO4474. Cytochrome C maturation locus**
SCO4546	Lpp	**Member of protein family unique to streptomyces c.f SCO5020-5022**	**In operon with SCO4547 secreted protein related to SCO5023**
SCO4552	Lpp	Conserved hypothetical	Monomer
SCO4650	Lpp	Conserved hypothetical, limited actinomycete distribution	Monomer; Family with SCO2725, SCO4651
SCO4651	Lpp	Conserved hypothetical, limited actinomycete distribution	Monomer; Family with SCO2725, SCO4650
SCO4739	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO4759	Lpp	Putative polysaccharide deacetylase (PF01522); SCO4760 paralogue	In operon with SCO4760 Lpp paralogue
SCO4760	Lpp	Putative polysaccharide deacetylase (PF01522); SCO4759 paralogue	In operon with SCO4759 Lpp paralogue
SCO4832	Lpp	**SBP PF04069 OpuAC glycine betaine-related**	**Locus complete, in operon with SCO4830-SCO4831 + SCO4829 putative oxidoreductase**
SCO4868	Lpp	PF03734 ErfK/YbiS/YnhG family, putative peptidoglycan active enzyme	Monomer
SCO4884	Lpp	**SBP PF02608 Bmp/ribonucleosides**	**Locus completed by SCO4886-4890; ribonucleoside uptake & metabolism locus**
SCO4885	Lpp	**SBP PF02608 Bmp/ribonucleosides**	**Locus completed by SCO4886-4890; ribonucleoside uptake & metabolism locus**
SCO4905	Lpp	AfsQ3 putative TCS accessory protein	In operon with SCO4906-SCO4907 TCS
SCO4934	Lpp	PF03734 ErfK/YbiS/YnhG family, putative peptidoglycan active enzyme	Monomer
SCO5014	Lpp	Conserved hypothetical, Streptomyces restricted?	In operon with membrane protein, secreted protein and OMP-like protein (SCO5012-SCO5015)
SCO5020ext	Lpp	Member of protein family unique to streptomyces ? Cf SCO5021-5022; SCO4546	Nterminal extension requires reannotation
SCO5021	Lpp	**Member of protein family unique to streptomyces c.f. SCO4546, SCO5022**	**Linked to adjacent SCO5020 and SCO5022 Lpp**
SCO5022	Lpp	**Member of protein family unique to streptomyces c.f. SCO4546, SCO5021**	**In operon with SCO5023, relative of SCO4547**
SCO5110	Lpp	SBP PF00496 family 5	Monomer, linked to adjacent loci ?
SCO5113	Lpp	**SBP PF00496 family 5 BldKB■**	**Locus completed by SCO5112-SCO5115 BldKA-BldKD**
SCO5117	Lpp	**SBP PF00496 family 5**	**Locus completed by SCO5118-SCO5121**
SCO5155	Lpp	Conserved hypothetical, limited actinomycete distribution	Monomer; distantly related to SCO7393
SCO5179	Lpp	Putative protease, related to SCO3540, SCO4241 & adjacent SCO5180 Lpp	Monomer
SCO5180	Lpp	Putative protease, related to SCO3540, SCO4241 & adjacent SCO5179 Lpp	Monomer
SCO5232	Lpp	**SBP PF01547▲**	**Linked to SCO5233-SCO5235 but locus incomplete**
SCO5260	Lpp	**SBP PF00497 family 3; SAM induced [137]**	**Locus complete, in operon with SCO5258-SCO5259; adjacent SCO5257 is a putative SAM dependent methyltransferase**
SCO5305	Lpp	**Conserved hypothetical, Streptomyces restricted?**	**In operon with SCO5305 unpaired sensor kinase**
SCO5430	Lpp	**SBP PF01547▲**	**In operon with SCO5428-SCO5429 but locus incomplete**
SCO5457	Lpp	**PF03734 ErfK/YbiS/YnhG family, putative peptidoglycan active enzyme**	**Monomer; SCO5458 Lpp paralogue adjacent**
SCO5458	Lpp	PF03734 ErfK/YbiS/YnhG family, putative peptidoglycan active enzyme	Monomer; SCO5457 anom prob paralogue adjacent
SCO5477	Lpp	**SBP PF00496 family 5; SAM induced [137]**	**Locus completed by SCO5476-SCO5480**
SCO5588	Lpp	Hypothetical protein, unique to SCO	Monomer

SCO5646	Lpp	**SBP PF01547**	**Locus complete; in operon with SCO5646-SCO5649, including actinomyecete specific DUF103 protein; ferric transporter ?**
SCO5658	Lpp	**SBP COG0687 polyamine binding**	**In operon with SCO5659 but locus incomplete**
SCO5667	Lpp	**SBP polyamine (PotD) related**	**Locus complete, in operon with SCO5668-SCO5671, including putative oxidoreductase**
SCO5686	Lpp	**SBP COG1653 Ugp**	**Locus incomplete; in large operon SCO5696-SCO5691 of carbohydrate active proteins; missed in▲**
SCO5702	Lpp	Conserved hypothetical, Streptomyces restricted? Alanine-rich protein	In operon with SCO5701 Alanine rich and SCO5700 aminoglycoside resistance kinase
SCO5776	Lpp	**SBP PF00497 family 3; GluB, glutamine**	**Locus completed by SCO5774-SCO5777 GluA-D**
SCO5797	Lpp	Conserved hypothetical, matches IPR009003 serine/cysteine protease	SCO5798 anom prob Lpp paralogue adjacent
SCO5798	Lpp	Conserved hypothetical, matches IPR009003 serine/cysteine protease	SCO5797 Lpp paralogue adjacent
SCO5995	Lpp	Hypothetical protein, unique to SCO	In operon with SCO5996 WD40 repeat protein; sensing system?
SCO6005	Lpp	**SBP PF01547, NgcE▲**	**In operon with SCO6006-SCO6007 but locus incomplete**
SCO6009	Lpp	**SBP pentoses (matches COG4213, XylF/COG1879, RbsB); XylF▲■**	**Locus completed by SCO6010-SCO6011**
SCO6040	Lpp	Conserved hypothetical	In operon with SCO6039 putative flavoprotein oxidoreductase
SCO6050	Lpp	Putative polysaccharide deacetylase (PF01522)	In operon with SCO6049 small hypothetical protein
SCO6065	Lpp	**SBP PF04069 OpuAC glycine betaine-related**	**Locus completed by SCO6062-SCO6063**
SCO6070	Lpp	SBP pentoses (matches COG1879, RbsB)	Orphan, missed in▲
SCO6088	Lpp	**SBP PF01547**	**In operon with SCO6086-SCO6087 but locus incomplete**
SCO6096	Lpp	**SBP sulphonates or related substrates**	**Locus complete, in operon with SCO6093-SCO6095; adjacent to locus for sulphate assimilation SCO**
SCO6114	Lpp	**SBP PF00496 family 5**	**Locus completed by SCO6111-SCO6113**
SCO6125	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO6178	Lpp	Putative polysaccharide deacetylase (PF01522)	In operon with SCO6177 conserved hypothetical protein
SCO6221	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO6231	Lpp	**SBP PF01547▲**	**Linked to SCO6029-SCO6030 but locus incomplete; adjacent to mannosidase locus**
SCO6257	Lpp	**SBP PF00532 family 1; sugars, RbsE2▲**	**Locus complete, in operon with SCO6258-SCO6259; adjacent to SCO6260 putative sugar kinase**
SCO6289	Lpp	**Putative FAD dependent oxidoreductase (PF01266)**	**In operon SCO6289-SCO6293 including dihydropicolinate synthase and proline racemase; metabolite biosynthesis & export ?**
SCO6368	Lpp	**Putative SBP, restricted distribution**	**In operon with SCO6366-SCO6367; ABC locus completed by SCO6365**
SCO6377	Lpp	Putative Ricin B-like lectin	Monomer
SCO6381	Lpp	Conserved hypothetical	Monomer
SCO6451	Lpp	**SBP PF00496 family 5**	**Locus complete, in operon with SCO6452-SCO6455**
SCO6569	Lpp	SBP PF00532 family 1, pentoses▲	Locus complete, in operon with SCO6568-SCO6567; adjacent to SCO6570 putative carbohydrate active oxidoreductase
SCO6592	Lpp	Hypothetical protein, unique to SCO	In operon with SCO6593 ALF repeat (DUF312) protein
SCO6601	Lpp	**SBP PF01547▲**	**Linked to SCO6602-SCO6603 but locus incomplete; adjacent to beta-glucosidase SCO6604**
SCO6644	Lpp	**SBP PF00496 family 5**	**In operon with SCO6645-SCO6646 but locus incomplete; SCO6646 is a major faciliator superfamily permease**
SCO6705	Unclear	Conserved hypothetical	Monomer
SCO6816	Lpp	**SBP PF01547**	**Locus complete: in operon with SCO6814-SCO6815; Part of large genomic area of lateral gene transfer (SCO6806-6953)**
SCO6871	Lpp	Conserved hypothetical, Streptomyces restricted?	In operon with SCO6870-SCO6873; Part of large genomic area of lateral gene transfer (SCO6806-6953)
SCO6916	Lpp	Conserved hypothetical, Streptomyces restricted?	Monomer
SCO6963	Lpp	Conserved hypothetical, Streptomyces restricted? Homologue of SCO3781 Lpp	Monomer
SCO6979	Lpp	**SBP pentoses (matches COG1879, RbsB)▲**	**Locus complete in operon with SCO6980-SCO6982; SCO6982 related to xylose isomerase PF01261**
SCO7013	Lpp	**SBP PF01547, AglE1▲**	**Linked to SCO7011-SCO7012 but locus incomplete**
SCO7028	Lpp	**SBP PF01547, BxlE1▲**	**In operon with SCO7029-SCO7030 but locus incomplete; orthologue of S. griseus BxlE (Tsujibo et al. 2004)**
SCO7069	Lpp	Fibronectin type III domain - ligand binding?	Monomer
SCO7167	Lpp	**SBP PF01547▲**	**In operon with SCO7166-SCO7165 but locus incomplete**
SCO7185	Lpp	**SBP, PF01094 branched chain amino acids**	**Locus complete, in operon with SCO7183-SCO7184 IMPS & SCO7181-SCO7182 ATPases; ABC branched chain aa uptake**
SCO7218	Lpp	**SBP PF01497 Fe-siderophore**	**Locus completed by SCO7216-SCO7217**
SCO7232	Lpp	Conserved hypothetical, limited actinomycete distribution	In operon with SCO7230-SCO7231 TCS = accessory protein?
SCO7393	Lpp	distantly related to SCO5155	Monomer
SCO7399	Lpp	**CdtB SBP (PF01497 Fe-siderophore)**	**Locus complete, in operon with SCO7398-SCO7400; Bunet et al. 2006**
SCO7408	Lpp	**SBP PF01547, LacE▲**	**Linked to SCO7409-SCO7410 but locus incomplete**
SCO7434	Lpp	SspA, sporulation specific Lpp (Tzanis et al. 2006 poster)	Monomer
SCO7460	Lpp	Hypothetical protein, unique to SCO	Monomer
SCO7489	Lpp	**SBP PF01547▲**	**In operon SCO7484-SCO7489 but locus incomplete; SCO7485 putative carbohydrate active oxidoreductase**
SCO7503	Unclear	**SBP PF01547▲**	**In operon with SCO7504-SCO7505 but locus incomplete**

SCO7532	Lpp	Conserved hypothetical, limited actinomycete distribution; hydrolase domain?	Monomer
SCO7535	Lpp	Conserved hypothetical, limited actinomycete distribution	In operon with SCO7533-SCO7534 TCS = accessory protein?
SCO7546	Lpp	**SBP PF04069 OpuAC glycine betaine-related**	**Locus complete, in operon with SCO7544-SCO7547, including putative sulphatase, + SCO7548 DUF323 protein**
SCO7555	Lpp	**SBP PF01547, Ceb2E▲**	**In operon with SCO7556-SCO7557 but locus incomplete**
SCO7563	Lpp	**SBP PF00496 family 5**	**Locus complete, in operon with SCO7564-SCO7567**
SCO7583	Lpp	Glycosyl hydrolase family 9 PF00759	Monomer
SCO7673	Lpp	PF03640 Y-X4-D motif protein, as is SCO0930 [133]	In operon with SCO7674-7675
SCO7674	Lpp	PF00127 Copper binding protein	In operon with SCO7673-7675 (4 MSD): copper acquisition locus ?
SCO7677	Lpp	**SBP PF00496 family 5**	**Locus complete, in operon with SCO7678-SCO7680**
SCP1.165	Lpp	Conserved hypothetical, limited actinomycete distribution	In operon with SCP1.166; locus distantly related to SCO4546 protein family

All lipoproteins highlighted in bold represent potential substrate binding proteins.

▲ Identified in genomic survey of carbohydrate binding SBP [138]

■ Identified as an ASDP-ribosylated protein [139]

Table A2. Stress Phenotype Results.

Name	Conc/Amount		WT			*Δlgt1::apr*			*Δlgt2::apr*			*Δlsp*FLP			*cis**			*in trans**		
			A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
Bacitracin	µg/ml	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		10	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		15	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		20	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		25	2	2	2	2	2	2	2	2	2	1	1	1	2	2	2	2	2	2
		30	2	2	2	2	2	2	2	2	2	1	1	1	2	2	2	2	2	2
		35	2	1	2	1	1	1	2	2	2	0	0	0	1	2	2	2	2	2
		40	2	1	2	1	1	1	2	2	2	0	0	0	1	2	1	2	1	1
		50	2	0	2	1	0	0	2	2	2	0	0	0	1	1	0	1	1	0
		100	0	0	0	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0
		200	0	0	0	0	0	0	1	1	2	0	0	0	0	0	0	0	0	0
		300	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
		400	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
		500	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
		1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
+15 Hour SDS	%	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.01	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.02	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.03	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.04	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.05	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	1	1	1
		0.06	2	2	2	2	2	2	2	2	2	1	1	1	2	2	1	1	1	1
		0.07	2	2	2	2	2	2	2	2	2	1	1	1	1	1	0	1	1	1
		0.08	2	2	2	2	2	2	2	2	2	1	1	1	1	1	0	1	1	1
		0.09	2	2	2	2	2	2	2	2	2	1	1	1	1	1	0	1	1	1
		0.1	2	2	2	2	2	2	2	2	2	1	0	1	1	1	0	1	1	1
		0.2	2	1	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0
0 hour lysozyme	µg/ml	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.3	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2
		0.4	2	2	2	2	2	2	2	1	2	1	1	1	2	2	2	2	2	2
		0.5	2	2	2	2	2	2	2	1	2	1	1	1	2	2	2	2	2	2
		0.75	1	1	1	1	1	1	1	1	2	1	1	1	1	2	2	2	2	2
		1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1
		1.25	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1
		2.5	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0
		5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
+15 Hour Lysozyme	µg/ml	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
		0.4	2	2	2	2	2	2	1	1	1	2	2	1	1	2	1	1	2	2
		0.5	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1
		0.75	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		1	2	2	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1
		1.25	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1
		2.5	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0
		5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

**cis = Δlsp* cis complementation
in trans = Δlsp in trans complementation