University of Glasgow

McQuitty, Claire (2015) Characterisation of two key virulence factors in Escherichia coli O157:H7. MSc(R) thesis.

http://theses.gla.ac.uk/6943/

Characterisation of two key virulence factors in *Escherichia coli* O157:H7

A thesis submitted to the University of Glasgow for the degree of MSc by Research

Claire McQuitty BSc (Hons)

Submitted Date
25/11/2015

Institute of Infection, Immunity and Inflammation

College of Medical, Veterinary and Life Sciences

University of Glasgow

Abstract

Enterohaemorrhagic Escherichia coli (EHEC) are a subset of pathogenic *E. coli* which can cause diarrhoeal disease, with the majority of infections due to serovar O157:H7. Successful infection by EHEC is determined by the expression of two key virulence factors, flagella and the type three secretion system (T3SS), a bacterially encoded needle-like filament. Flagella are responsible for directional and targeted swimming of the bacteria as well as for initiating bacteria: host cell attachment. A subsequent switch to T3SS expression promotes enhanced binding via the transmission of bacterial proteins (termed effectors) into the host cell. Once in the host cell these effectors act to manipulate host cell pathways and further facilitate infection.

This thesis sets out to explore these two important virulence factors; firstly by exploring the regulation of flagella rotation, and thus motility, in a series of Acetaldehyde coA dehydrogenase (AdhE) deletion mutants; and secondly by assessing the potential of a novel fluorescent reporter molecule, LOV, in imaging the expression, translocation and host cell localisation of T3SS injected effector proteins.

Findings from this study show that the LOV domain can be used to fluorescently tag the bacterial effector, Tir, and monitor it's expression, translocation through the T3SS and localisation within the host cell, in real-time. This opens up exciting new possibilities in using LOV to fluorescently monitor the localisation of bacterial effector proteins within the host cell and thus give information on potential cellular partners and mechanism of action. Additionally, exploration of flagella rotation regulation, through the isolation and genomic sequencing of a set of AdhE deletion mutants, suggested a role for AdhE in bacterial motility via indirect acetylation of the chemotaxis protein CheY.

A better understanding of both of these virulence factors will render us more capable in our ability to carry out treatment and prevention of EHEC mediated disease.

Table of Contents   Page

List of Tables   **Page**

List of Figures   **Page**

Acknowledgements

Caitlin: Thank you for being endlessly enthusiastic and supportive while I was in Glasgow and for understanding that the real solution to most problems is tea and biscuits.

Author's declaration

I declare that, except where explicit reference is made to the contribution of others, this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or at any other institution.

Claire McQuitty

Abbreviations

°C	Degrees Celcius
AdhE	Bi-functional acetaldehyde/alcohol dehydrogenase
A/E Lesions	Attaching and Effacing Lesions
bp	Base Pair
BSA	Bovine Serum Albumin
Chl	Chloramphenicol
CO_2	Carbon dioxide
dH_2O	de-ionised H_2O
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's Phosphate Buffered Saline
ECL	Enhanced Chemi-luminescence
EDTA	Ethylenediaminetetraacetic acid
EHEC	*Enterohaemmorhagic E. coli*
EPEC	*Enteropathogenic E. coli*
FCS	Foetal Calf Serum
FMN	Flavin Mono Nucleotide
FRET	Fluorescence Resonance Energy Transfer
g	Grams
gDNA	Genomic Deoxyribonucleic acid
Gent	Gentamicin
GFP	Green Fluorescent Protein
h	Hours
H_2O	Water
ISP	Ion Sphere Particle
Kan	Kanamycin
Kb	Kilobase pair
kDa	Kilodalton
kV	Kilovolts
L	Litre
LB	Luria Broth
LEE	Locus of enterocyte effacement
Ler	LEE-encoded regulator
LOV	Light, Oxygen or voltage domain

LPS	Lipopolysaccharide
M	Molar
Map	Mitochondrial associated protein
MEM	Minimal Essential Media
mg	Milligram
min	Minute
ml	Millilitre
Na^+	Sodium
NHE3	Sodium Hydrogen antiporter 3 Cation proton pump
NLE	Non LEE encoded
nm	Nanometer
nM	Nanomolar
OD_{600}	Optical density at 600_{nm}
OT2	One Touch 2
PFA	Paraformaldehyde
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
ROI	Region of Interest
rpm	Revolutions per minute
RT	Room Temperature
SA	Salicylidene acylhydrazide
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SGLT 1	Sodium Glucose co-transporter 1
Snp	Single nucleotide polymorphism
Stx	Shiga toxin
T3SS	Type Three Secretion System
TAE	Tris-acetate-EDTA
TCA	Trichloracetic Acid
TER	Transepithelial electrical resistance
Tir	Translocated intimin receptor
UK	United Kingdom
USA	United States of America
V	Volts
w/v	Weight for Volume
Δ	Deletion

1. Introduction

1.1 *Escherichia coli* O157:H7

Escherichia coli (E. coli) are a common bacteria of the human gut and a component of the microbiota which colonise the mucosal layer of the colon (Kaper, Nataro, & Mobley, 2004). The vast majority of *E .coli* are commensals, existing symbiotically in the gut and rarely causing disease. Pathogenic *E. coli* have evolved to acquire a range of virulence factors that allow for invasion and colonisation of new niches within the human host. One example of a pathogenic *E. coli* is enterohaemmorhagic *E. coli* (EHEC), an intestinal pathogen that has been highlighted as a major cause of gastro-enteritis in developed countries (Naylor et al., 2007). The majority of cases in the UK and USA are due to infection with serovar O157:H7 (Nataro & Kaper, 1998). *E. coli* types can be grouped by their major surface antigens (Orskov et al,. 1977), "O157" refers to the O-antigen, a repetitive glycan polymer contained within the bacterial surface protein, lipopolysaccharide (LPS), and "H7" to the H–antigen, the flagellin protein. *E. coli* O157:H7, here after referred to as EHEC, colonises the mucosal surfaces of the terminal rectum of cattle, its natural reservoir (Croxen & Finlay, 2010; Naylor et al., 2003). The majority of disease outbreaks occur after consumption of contaminated food (Su & Brandt, 1995), such as meat, especially minced beef, dairy products and also fruit or vegetables contaminated by manure (Mahajan et al., 2009). EHEC can also be caught through direct contact with animals, and disease outbreaks associated with animal fairs and petting zoos are well reported (Heiman, K., et al, 2015). Additionally, there is risk of waterborne and secondary transmission of disease, particularly in establishments such as day care centres (Heiman, K., et al, 2015; Olsen et al., 2002). In the human host EHEC colonises the distal ileum and the large bowel and can cause abdominal cramps and bloody diarrhoea which, in more severe cases, can progress to haemolytic uremic syndrome (HUS) (Croxen & Finlay, 2010; Su & Brandt, 1995). HUS is a clinical syndrome that can result in acute kidney failure and is one of the most common causes of renal damage in children (Canpolat, 2015).

1.1.1 EHEC infection process

The EHEC infection process is shown in figure 1-1 and can be summarised as directional swimming of EHEC toward the host epithelial cell followed by bacterial attachment and injection of bacterial proteins. Initial infection is mediated by the expression of flagella, a bacterial cell surface organelle associated with motility, followed by switching to expression of the Type Three Secretion System (T3SS) to facilitate bacterial attachment to

the host cell (Tree et al., 2009). Flagella act as an adhesin in the terminal rectum of cattle (Mahajan et al., 2009) and are involved in directional "swimming" (chemotaxis) in the human host (Eisenbach and Caplan, 1998) (figure 1-1 A). During progression of infection EHEC undergoes a transition from expression of flagella to expression of the T3SS, a need-like filament encoded on a pathogenicity island, the Locus of Enterocyte Effacement (LEE) (Garmendia and Frankel, 2005). The T3SS is required for tight bacterial attachment to epithelial cells, via the injection of bacterial proteins, termed effectors, into the host cell (Jarvis et al., 1995) (figure 1-1 C). Injected effectors act on numerous host cell pathways and are essential to EHEC infection. Of particular importance is the effector protein Tir which, once injected, inserts into the host cell membrane to act as a receptor for the bacterial surface protein, intimin (Kenny et al., 1997). Tir also mediates rearrangement of host cell actin to form an actin pedestal, to which the bacteria is tightly bound (Divinney et al., 2001) (figure 1-1 D). It is important that switching from flagella to T3SS expression occurs at the right time and in the appropriate place in order to allow the progression of EHEC infection (Tree et al., 2009).

In addition to flagella and the T3SS, EHEC encodes a Shiga-like toxin (Stx) which is involved in the infection process. During colonisation a subset of bacteria are lysed, releasing Stx. The toxin acts on epithelial cells to promote relocation of cellular factors, such as nucleolin, to the host cell surface which facilitates intimate binding to the bacterial surface receptor, intimin (Tree et al., 2009) (figure 1-1 B).
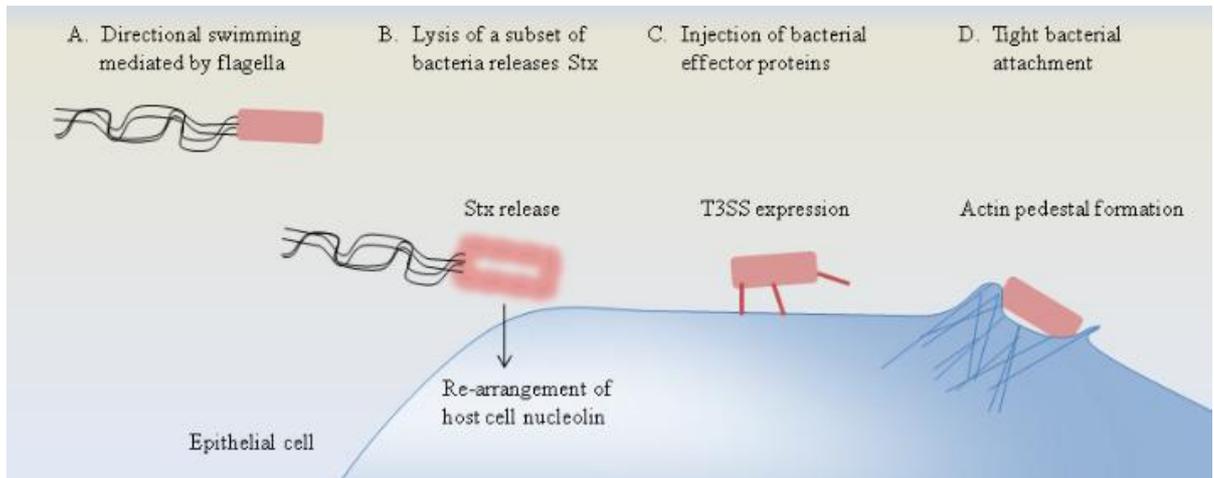
Figure 1-1: Schematic diagram of EHEC infection process. Initial infection is mediated by flagella, required for bacterial motility (A). During infection a subset of bacteria are lysed to release Stx, which acts on the host cell to rearrange nucleolin localisation to the host cell surface, where it acts as a receptor for the bacterial surface protein, intimin (B). A transition from flagella to T3SS expression occurs as infection progresses. The T3SS allows for injection of bacterial effector proteins into the host cell (C), resulting in the rearrangement of host cell actin and the formation of actin pedestals, to which EHEC is tightly bound (D).

1.1.2 The shiga-like toxin

EHEC encodes Stx on lambdoid-like phages that are integrated into the bacterial genome (Xu et al., 2012). The toxin is released upon bacteriophage- mediated lysis of the bacterial cell (Tree et al., 2009). Stx is composed of an enzymatically active "A" subunit and a pentamer of "B" subunits, which are involved in receptor binding (Croxen & Finlay, 2010). The toxin can be separated into two families; Stx 1 and Stx 2, with the majority of EHEC carrying the genes for both. In the context of disease, Stx producing bacteria are a major cause of HUS, with Stx 2 being more commonly associated with its onset (Croxen & Finlay, 2010 and Kimmitt, Harwood, & Barer, 2000). Toxin production is mediated by the SOS response, a global response to DNA damage, which stimulates bacteriophages (Kimmitt et al., 2000). DNA damage triggers a cascade of reactions resulting in cleavage of the SOS gene repressor, LexA (Janion, 2001). This stimulates initiation of the SOS response, triggering increased Stx production and release. Released toxin can act on two areas of the human gut: either directly on epithelial cells, or through binding to Gb3 receptors on Paneth cells located in the base of small intestinal crypts (Schüller et al., 2007). The Gb3 receptor role is currently unknown however cell lines and mouse models lacking Gb3 receptors are reported to have reduced sensitivity to Stx (Jacewicz, M. S, 1994, Melton-Celsa, A. R., 2014). The physiological role of Stx action on intestinal Paneth cells

has not been fully explored: however it also binds Gb3 receptors present on kidney epithelial cells and mediates necrosis leading to cell death (Croxen & Finlay, 2010).

1.1.3 Current treatment of EHEC

Due to the production of Stx, current treatment for EHEC is largely limited to alleviating the symptoms of the disease. Antibiotic treatment is associated with increased risk of developing HUS (Tarr, Gordon, & Chandler, 2005), possibly through triggering bacterial lysis resulting in toxin release, or by induction of the SOS response which results in increased toxin production via stimulation of bacteriophages which harbour Stx genes (Kimmitt et al., 2000).

The association of antibiotic treatment with increased production and release of Stx toxin presents a clear incentive for the development of alternative treatments for EHEC infection.

1.1.4 Anti-virulence compounds

An alternative to antibiotic treatment is to specifically target bacterial virulence factors important for pathogenicity, such as flagella, Stx and the T3SS. Not only would targeting virulence factors be beneficial for the treatment of EHEC, but non-lethal targeting of virulence factors as opposed to inducing bacterial death could reduce the selective pressure among bacteria that leads to development of resistance (Rasko & Sperandio, 2010).

While treatment with anti-virulence drugs may not kill bacteria directly, interference or inhibition of bacterial virulence factors could reduce host colonisation, for example by blocking recognition of host cell signalling factors that aid infection. Additionally, anti-virulence drugs selectively target the infecting pathogen, thus preventing damage to the host microbiota. In the context of disease, the host microbiota acts as a protective barrier against invading pathogens (Rasko & Sperandio, 2010).

There are several virulence targeting drugs under development which have potential for targeting EHEC. Synsorb-PK was developed in the 1990s as a Shiga toxin inhibitor (Armstrong et al., 1991). It acts as a Gb3 mimic by competitively binding free Shiga toxin in the gut. Although Snysorb-PK showed some success in removal of toxin from the gut,

phase one clinical trials noted no reduction in the on-set of HUS (Tarr et al., 2005, Rasko & Sperandio, 2010 and Scheiring et al., 2008) and subsequent use was discontinued.

The Salicylidene acylhydrazides (SA), a group of compounds identified by Kauppi et al., (2003), and Aurodox, a bacterially produced compound (Zambelloni et al., 2015), have been found to inhibit the Type Three Secretion System (T3SS). The T3SS is a needle like filament expressed by several gram negative bacteria, including EHEC, which facilitates the injection of bacterial proteins into the host cell. These injected proteins act to manipulate host cell processes and are required for successful infection (McGhie et al., 2009, Ogawa et al., 2008, Matsumoto, H. & Young, M., 2009) thus making the T3SS an attractive anti-virulence target.

The search for appropriate bacterial virulence factor targets and the development of anti-virulence drugs is an ongoing process, and one of particular importance in the era of antibiotic resistance.

1.2 Motility

The ability to be motile is a major factor in EHECs infection process. Motility is largely attributed to expression of flagella, long filamentous structures protruding from the bacterial cell surface. These act to propel the bacteria by either clockwise (CW) or counter-clockwise (CCW) rotation in response to a variety of external stimuli. Correct and appropriately timed expression of flagella by EHEC is important for successful infection. Expression of flagella are not only required for bacterial motility, but are also suggested to play a role in EHEC adherence to host cells (Mahajan et al., 2009). Mahajan et al., (2009) show that flagella are responsible for initial adherence to bovine terminal rectum epithelial cells, resulting in EHEC colonisation of the host reservoir. Similarly, flagella have been implicated in the adherence of Enteropathogenic *E.coli (*EPEC), a genetically similar pathogenic *E. coli* strain, to epithelial cells (Giron et al., 2002).

1.2.1 Flagella

The flagella machinery is composed of three main elements, the flagella filament, the flagella hook and the basal apparatus (or motor complex) (Figure 1-2). The filament and hook are constructed of multiple copies of FliC, flagella filament protein and FlgE, flagella hook protein. The basal apparatus is a complex structural component, spanning the

bacterial cytoplasmic membrane, which serves as the flagella motor. It is composed of four ring-like structures, the L-ring, P–ring and MS-ring, which are embedded in the bacterial cell wall, and the C-ring, comprising the "switch complex", situated inside the bacterial cell (Berg, 2003). The "switch complex" is composed of FliG, FliM and FliN proteins and acts to modulate switching between CW and CCW rotation of flagella. CW rotation results in bacteria "tumbling" and directional change where as CCW rotation mediates bacterial "running", propelled forward by flagella.

Switching is controlled partly by the chemotactic sensing protein CheY, which upon phosphorylation binds to FliM, increasing the probability of CW rotation (Bren & Eisenbach, 2000).
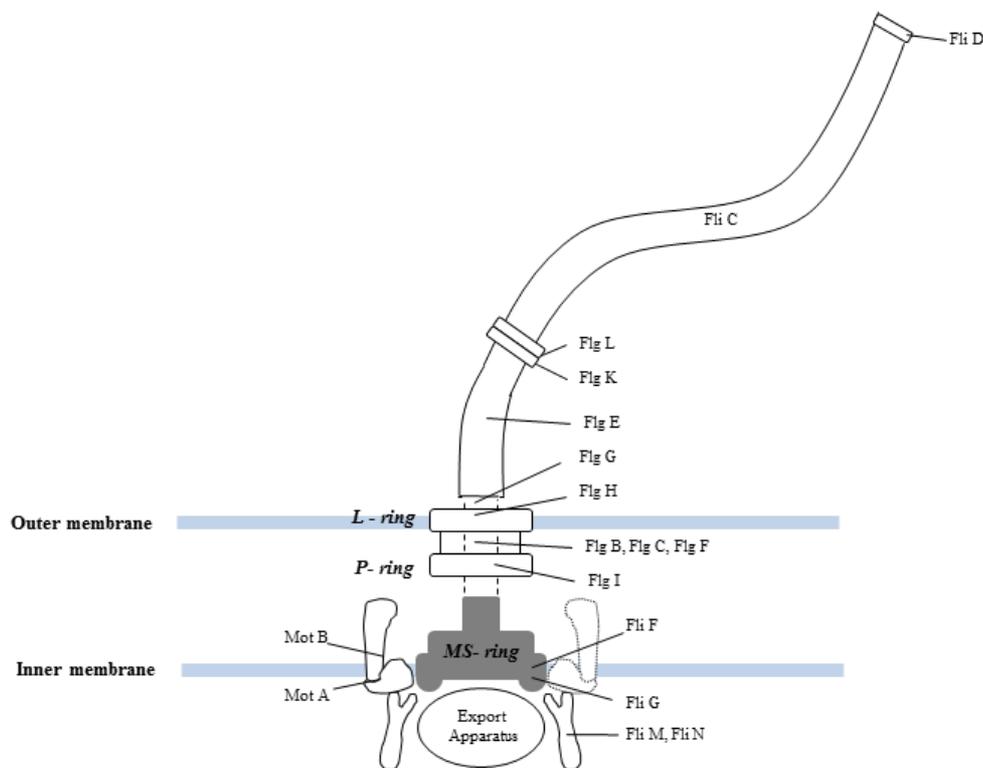
Figure 1-2: Schematic diagram of Flagella. Flagella drive bacteria motility and are composed of three main elements, the FliC flagella filament, FlgE flagella hook and a membrane embedded motor complex, the basal apparatus. Phosphorylated CheY, chemotactic protein, binds to FliM, a component of the motor complex, greatly increasing the probability of CW rotation.

1.2.2 The chemotaxis regulator CheY

Chemotaxis is the directional "swimming" of bacteria toward an attractant, or away from a repellant (Adler, 1965). CheY acts as a regulator of chemotaxis via its interaction with either CheA or FliM, dependent on its phosphorylation state (Barak & Eisenbach, 2004). In the non-phosphorylated state it binds to and is phosphorylated by the histidine kinase CheA. Upon phosphorylation, CheY dissociates from CheA and binds FliM, at a well conserved N-terminal segment (Bren & Eisenbach, 1998), increasing probability of CW flagella rotation. It is argued, however, that post translational modification of CheY by phosphorylation alone, is not sufficient for inducing CW rotation of flagella (Silversmith & Bourret, 1999 and Eisenbach & Caplan, 1998). Wolfe et al., 1988, showed that, in a genetically engineered *E. coli*, lacking the components of the chemotaxis pathway, expression of CheY alone was insufficient for generating CW rotation. However, the flagella motor could undergo a switch from CCW to CW biased rotation upon addition of acetate. CheY comprises several C- terminal acetylation sites; Lys 91, Lys 92, Lys 109, Lys 119, Lys 122 and Lys 126 (Barak et al., 2004 and Ramakrishnan et al., 1998), of which Lys 92, Lys 119 and Lys 122 are reported to be directly involved in FliM binding (Dyer et al., 2004). Furthermore, acetylation of CheY is reported to greatly increase the level of flagella CW rotation in comparison to non-acetylated CheY (Barak et al., 1992). A recent study by Fraiberg et al., (2015) determined Lys 91 and Lys 109 to be the major sites of CheY acetylation. More specifically they report Lys 91 to play an inhibitive role in CW rotation, which is relieved upon acetylation.

1.2.3 Acetylation of CheY

Available acetate is in part regulated by the enzyme Acetyl-coA (AdhE). AdhE is involved in the conversion of acetyl-coA to acetaldehyde, which is subsequently converted to ethanol. A previous study by Beckham et al., (2015) showed that deletion of *adh*E in an EHEC model (ΔAdhE) resulted in elevated levels of extracellular acetate and strong suppression of the T3SS. In addition it was noted that ΔAdhE bacteria showed over-expression of flagella but were completely immotile. The authors speculated that this immotile phenotype may be due to inappropriate acetylation of one or more flagellar protein(s).

Motility plays an important role in EHECs infection process. A better understanding of the components underlying regulation of motility could present potential new targets for preventing and treating EHEC infection.

1.3 The Type Three Secretion System (T3SS)

Another key player in EHEC infection is the T3SS. The T3SS is involved in both the establishment and progression of infection. Initial attachment of EHEC is mediated by surface proteins; such as flagella (Mahajan et al., 2009) followed by a more intimate attachment with the host cell established by the T3SS (Jarvis et al., 1995). The T3SS machinery is encoded on the Locus of Enterocyte Effacement (LEE) bacterial pathogenicity island (Figure 1-4) (Garmendia & Frankel, 2005). It is composed of a needle–like filament (Figure 1-3) that facilitates the injection of bacterial proteins, termed effectors, across the cell membrane and into the host cell (Tree et al., 2009). Injected effectors act to manipulate host cell signalling pathways and mediate tight attachment of bacteria, via the recruitment of host cell actin to form pedestals, termed attachment and effacement (A/E) lesions.
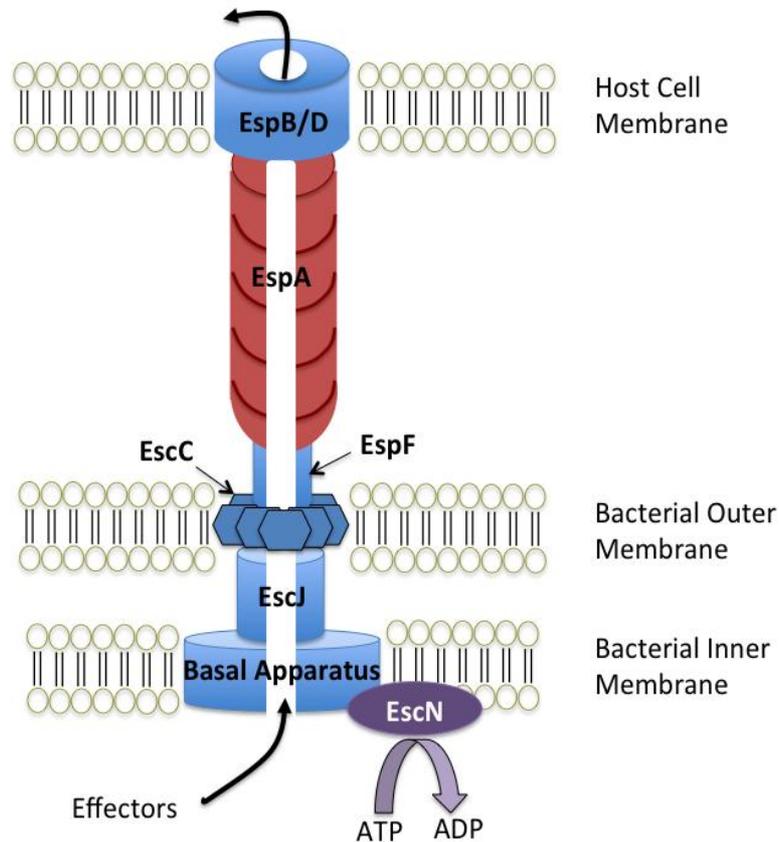
Figure 1-3: The Type Three Secretion System. The T3SS apparatus is composed of an inner membrane basal apparatus, an outer membrane Esc C ring and a filamentous needle structure. The basal apparatus is composed of seven Esc proteins, held in a ring like structure within the bacterial inner membrane, associated with which is the ATPase Esc N. On top of the basal apparatus, and spanning the space between the inner and outer membrane, lies a super-structure of about 24 Esc J subunits. Esc C forms an outer membrane ring from which protrudes the Esp F needle filament (Wang et al., 2009).
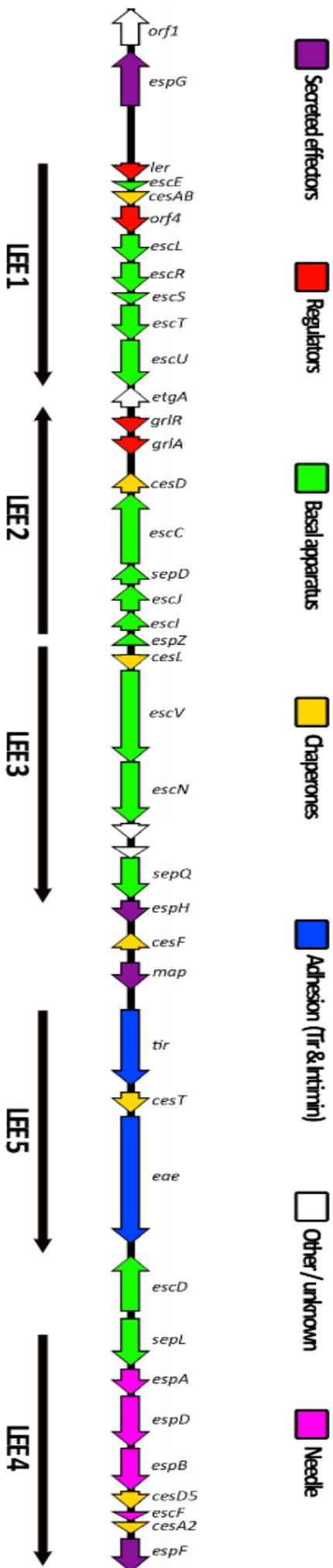
Figure 1-4: Locus of Enterocyte Effacement (LEE) (Figure from James Connolly, PhD, University of Glasgow)
The LEE encodes 41 genes important in EHEC infection. These include, the T3SS (green and pink) and secreted bacterial effector proteins (purple, and Tir in blue).

1.3.1 The T3SS machinery

The structural component of the T3SS is composed of EspA, EspB and EspD, encoded on the LEE (Knutton et al., 1998). EspA, encoded within the region between *espB* and *eae* (encoding intimin), forms the filamentous needle-like structure that translocates effectors from the bacterium into host cells (Kenny, 2002 and Sekiya et al., 2001). EspB and EspD are translocated via the EspA filament and have been detected in host cell membranes. EspD has been reported to be involved in regulating secretion of EspA and EspB, as shown by reduced levels of EspA and EspB secretion in an *espD* mutant (Lai et al., 1997).

1.3.2 EHEC effector proteins

Tobe et al., (2006), report a total of 39 EHEC effector proteins, which can be sorted in to more than 20 different protein families. These can be categorized as LEE-encoded, those encoded within the 35 kb LEE pathogenicity island (Figure 1-4), and non-LEE (Nle) encoded effector proteins, encoded on prophages throughout the bacterial genome.

The LEE-encoded proteins; Tir, Map, EspB, EspF, EspH, EspZ and EspG, are exported via the T3SS during infection (Dean and Kenny 2009, Leong et al, 2003), along with an additional Nle effector protein, EspFu, found on prophage 933U (Campellone et al, 2004). These exported effector proteins are multifunctional and can target and manipulate host cell machinery and pathways to aid infection.

The roles of several effectors and their importance during infection are highlighted below.

Tir and EspFu facilitate bacteria attachment

Tir is a ~78 kDa bacterial protein which, after injection, is inserted into the host cell membrane where it acts as a receptor for the bacterial surface protein, intimin (Kenny et al., 1997). Intimin-Tir interaction results in clustering of Tir within the host cell and facilitates tight attachment of the bacteria via the accumulation of host cell actin to form an actin pedestal (DeVinney et al., 2001). The formation of actin pedestals (also termed A/E lesions) in EHEC requires an additional injected effector protein, EspFu (Campellone et al., 2004). The C-terminal NYP$_{458}$ tri-peptide in Tir interacts with two host cell proteins; IRSp53 and IRTKs, which in turn recruit EspFu. EspFu binds a GTPase binding domain within N-WASP, resulting in a conformational change that ultimately stimulates activation

of the Arp2/3 complex leading to actin recruitment and assembly of actin pedestals (Campellone et al., 2004 and DeVinney et al., 2001) .

EspB and EspF inhibit phagocytosis

In addition to facilitating bacterial attachment, effector proteins can act to perturb host cell defence mechanisms by blocking phagocytosis. Phagocytosis is a key process in the host defence against invading pathogens. Professional phagocytes, dendritic cells and macrophages, internalise and degrade bacteria upon recognition of bacterial surface proteins such as lipopolysaccharide (LPS) (Peiser 2000). This is achieved by rearrangement of the actin cytoskeleton to form a pseudopod followed by expansion of cellular membrane. The bacteria are internalised into a phagosome, the contents of which are degraded (Freeman & Grinstein, 2014). EspB and EspF inhibit phagocytosis. EspB competitively binds to myosin through a specific myosin binding domain. It inhibits pseudopod formation and prevents phagosome closure (Lizumi et al., 2007, Santos & Finlay, 2015). EspF perturbs phagocytosis through inhibition of cytoskeletal rearrangement in phagocytic cells (Quitard et al., 2006). Inhibition of phagocytosis is a valuable defence mechanism as it provides EHEC with protection from destruction by the host cell.

Map and EspF interfere with multiple cellular processes

In 2000 Kenny and Jepson reported EspB-dependent delivery of effector protein Orf19, encoded directly upstream of *tir*. Orf19, which re-locates to host cell mitochondria and causes loss of mitochondrial membrane stability, was subsequently re-named Mitochondria-Associated Protein, Map. The role of Map has been predominantly studied in EPEC, however a similar role may be present in EHEC. Map has been reported to target and disrupt the function of host cell mitochondria (Kenny & Jepson, 2000 and Ma et al., 2006), by disrupting mitochondrial membrane potential, leading to cytochrome c release and resulting in apoptosis (Santos & Finlay, 2015.) Map has also been noted to induce Cdc42-dependent filopodia formation at the site of infection (Kenny et al., 2002), disrupt tight junctions (Dean & Kenny, 2004) and is linked to up-take of EPEC by non-phagocytic cells (Jepson et al., 2003). During EPEC infection cytoskeletal rearrangement to form filopodia-like structures occurs at the site of infection. Filopodia formation was found to be dependent on Cdc42, a small cellular GTPase involved in filopodia formation under normal cell conditions, (Kenny et al., 2002) although the mechanism by which Map and Cdc42 interact is currently unknown. Tir has also been reported to play a role in down-

regulating filopodia formation (Kenny et al., 2002), suggesting an overlap in effector protein action on host cell pathways.

Other injected effector proteins play a crucial role in EHEC infection. Notably EspF is reported to interact with numerous host cell proteins, via 4 proline-rich motifs, resembling eukaryotic proline-rich sequences, recognised by cellular proteins bearing SH3 or tryptophan-tryptophan (WW) domains (Nougayrede et al., 2004.) EspF is reported to mediate nucleolar and mitochondrial disruption (Dean et al., 2010 and Nougayrede et al., 2004), inactivate NHE3 (Hodges et al., 2008) and SGLT-1 (Dean et al., 2005) intestinal pumps and disrupt epithelial tight junctions (Viswanathan et al., 2004) contributing to diarrhoeal symptoms.

EHEC effector proteins act on multiple pathways in the host cell. They are required for host cell attachment and phagocytic cell evasion. They cause apoptosis and contribute to the unpleasant symptoms associated with the disease. There is still much to be learnt about the specific roles and host cell targets of the effectors and improved knowledge in this area could allow us to develop new treatments and preventative methods against EHEC. To this end, visualisation of translocation and host cell localisation of the effectors proteins will provide us with valuable information on EHEC disease.

1.3.3 Monitoring translocation of effector proteins

Fluorescent imaging of bacterial effector proteins has allowed us to advance in our knowledge of certain diseases. *Salmonella enterica*, uses a T3SS to inject the effector protein SipA into the host, where it binds to and modulates host cell actin leading to bacterial invasion (Zhou, 1999). Similarly *Shigella flexneri*, relies on T3SS mediated translocation of the effector IpaC into the host cell to induce membrane ruffling, a pre-cursor to bacterial invasion (Blocker et al., 1999). Both *Salmonella enteric* and *Shigella flexneri* effector proteins have been visualised by fluorescent microscopy (discussed below) - aiding us in our understanding of the infection process.

Green fluorescent protein (GFP), isolated from the jellyfish *Aequorea victoria*, was first used as a fluorescent marker to study gene expression in *Caenorhabditis elegans* (Tsien, 1998), and has since been widely reported as a fluorescent marker to visualise protein localisation and gene expression (Chalfie et al., 2009). For example, Schlumberger et al.,

(2005), developed a host cell GFP-chaperone model, in which InvB, the host cell-derived chaperone for Salmonella effector SipA, was fused with GFP, allowing imaging of SipA localisation within the cell. This method allows for indirect imaging of bacterial effector proteins post-translocation, by fluorescently tagging the effector chaperone, an option that is available for a only a limited number of well-characterised effectors. A more attractive method to explore effector dynamics in the context of infection is to image bacterial effector expression, translocation and host cell localisation; and requires the production of a fluorescently tagged effector. GFP, although well established as a fluorescent reporter, has three major limitations that hinder its use as an effector protein tag. Firstly, GFP chromophore formation is oxygen-dependent, restricting its use to aerobic systems (Tsien, 1998); secondly, GFP fluorescence is unstable at low pH; thirdly, and most crucial in the context of imaging effector translocation, GFP forms a very stable barrel shaped protein and cannot be unfolded, excluding it from successful translocation through the T3SS pore (Akeda & Galán, 2005).

To overcome this limitation Van Engelenburg and Palmer (2008) adapted a split-GFP system to monitor translocation of three *Salmonella enterica* effectors exported by the T3SS. This was achieved by creation of effector proteins fused with a single strand of the GFP β- barrel along with loading of the complementary GFP β- barrel fragment into the host cell. Upon translocation, the complementary GFP fragments combine and fluorescence is detected (Van Engelenburg & Palmer, 2010). The split-GFP system supplies information on cellular localisation of effector proteins post-translocation, but does not allow for imaging of translocation in real-time.

Fluorescent labelling techniques such as Fluorescence resonance energy transfer (FRET) and Cys:FlAsH have also been used to image bacterial effector translocation (Ehsani et al., 2009); however, both have their limitations. FRET is a technique that is based on the direct transfer of light energy from a donor fluorophore to an acceptor fluorophore, and its subsequent fluorescence (Wallrabe & Periasme, 2005). FRET can be used to monitor bacterial effector translocation via a system in which effector proteins are tagged with β-lactamase and the host cell is loaded with a cleavable FRET substrate. Upon injection of effector, β-lactamase cleaves the FRET substrate and a change in fluorescence is detected (Enninga et al., 2007). The FRET system is effective for detecting the presence of effector proteins post-translocation, but doesn't allow for real-time imaging of effector translocation, as there is a slight delay between translocation and FRET cleavage.

Real-time imaging of translocation can be achieved by tagging effector proteins with a tetracysteine motif tag (4Cys). 4Cys is able to form a tight association with the small bi-arsenic dye FlAsH, and produce fluorescence. The 4Cys:FlAsH labelling system has been used successfully to monitor translocation of effector proteins in *Shigella flexneri*, by production of 4Cys-effector fusions (Sansonetti et al., 2005) and in *Salmonella*, via genomic incorporation of the 4Cys tag (Van Engelenburg & Palmer, 2010). Although the 4Cys:FlAsH system is an improvement on the FRET and GFP systems, in that it can offer real-time imaging, it has limited sensitivity and toxicity to the bacteria from the FlAsH dye has to be tightly controlled (Ehsani et al., 2009).

The above techniques allow us to image effector proteins throughout infection and have helped us to explore the mechanism of infection for several pathogens. However, we lack a fluorescent reporter that allows us to visualise the entire T3SS mediated infection process-from effector expression, to translocation, to host cell localisation. One such candidate is the small, plant derived, domain; LOV.

1.3.4 A novel reporter for visualising effector proteins throughout infection

The Light, Oxygen or Voltage (LOV) domain is a flavin-mononucleotide (FMN) binding domain derived from the N-terminus region of plant blue-light receptors. Upon illumination with blue-light it undergoes a reversible photocycle resulting in the formation of a covalent bond between the FMN chromophore and a cysteine residue within the LOV domain (Figure 1-5) (Christie, 2007). FMN is a blue-light absorbing chromophore, which fluoresces green upon excitation, affording it potential as a fluorescent reporter molecule. LOV offers several advantages over other fluorescent reporter molecules. Firstly, LOV is relatively small in size, ~11 kDa compared with ~25 kDa for GFP. This small size offers potential for fluorescent labelling of proteins which pass through the T3SS, something which is currently achieved in the split GFP system and 4cys-FlAsH system, described above. Secondly, the fluorescence of the LOV domain is oxygen independent (Drepper et al., 2007) offering it considerable potential as a fluorescent reporter in *in vivo* and anaerobic bacterial systems. In the context of visualising bacterial or cellular protein dynamics, these two factors provide LOV based reporters an advantage over currently used fluorescent imaging techniques.

A fluorescent reporter that has the potential to monitor effector proteins in real-time both in the bacterial and host cell is an exciting prospect. The use of LOV as this reporter will be evaluated in this thesis.
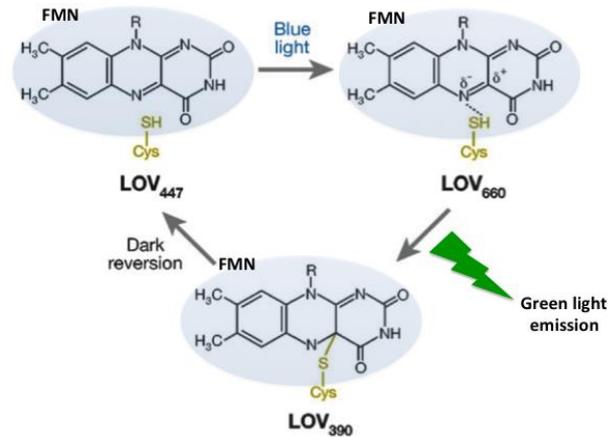
Figure 1-5: LOV reversible photocycle. (Figure adapted from Christie, 2007.) In dark conditions FMN is non-covalently bound to a conserved cysteine residue within the LOV domain. Upon illumination with blue/UV light a highly reactive LOV_{660}:FMN compound is formed, which rapidly decays to LOV_{390} with formation of FMN covalent bond and emission of green light.

1.4 Aims and Questions

The aim of this thesis is to explore two key areas of EHEC infection: the regulation of flagella mediated motility, using a number of AdhE deletion mutants and T3SS mediated effector protein translocation, using the novel fluorescent tool, LOV. To do this we posed two questions:

Could acetylation of flagella proteins affect motility in EHEC?

This thesis sets out to further explore the Δ*adh*E immotile phenotype, with the hypothesis that, in the *adh*E deletion strain, an elevated level of acetate may cause post-translational modifications to the chemotaxis protein CheY, resulting in non-motile bacteria. If this is the case, could AdhE be a potential anti-virulence drug target?

Is LOV an effective fluorescent reporter for monitoring bacterial effector proteins in real time, throughout infection?

This study evaluates the use of phiLOV as a novel fluorescent reporter for tagging bacterial effector proteins and monitoring their exportation, translocation and host cell re-localisation. PhiLOV is a derivative of LOV, which possesses several amino acid changes in the FMN chromophore binding region resulting in increased physical constraint of the FMN molecule and improved photostability (Christie et al., 2012).

2. Materials and Methods

2. Materials and Methods

2.1 Reagents

Table 2-1: Reagents and supplier details

Reagent	Source
20 x MES Running Buffer	Life Technologies
20 x Transfer Buffer	Life Technologies
Alexafluor (488, 555, 647)	Life Technologies
αO157 antibody	MAST assure
BSA	Sigma Aldrich
Bug Buster	Life Technologies
DAKO mounting media	Dako
DNA Ligase Kit	Roche
E-gel size select agarose	Life Technologies
EDTA (TAE)	
Fast DNA fragmentation and Library Prep set	New England Biolabs
Gotaq Mastermix	Promega
Heat inactivated foetal calf serum	SigmaAldrich
HeLa cells	European Collection of cell cultures
Ion Xpress Barcode Adaptors, 1 – 16 kit	Life Technologies
KOD buffer	Toyoba
KOD taq polymerase	Toyoba
Library amplification primer mix	
MES Buffer	Invitrogen
Minimal Essential Media, Hepes modification, (MEM-HEPES; Sigma M7278)	Sigma Aldrich
DNA, RNA and Protein Purification kit	Qiagen
NEBNext Fast DNA Fragmentation and Library Prep set kit	New England Biolabs
Nitrocellulose Membrane	Life Technologies
NuPAGE 4 – 12 % Bis- Tris pre-cast gel	Amersham
NuPAGE 4X sample buffer	Invitrogen
Phalloidin	Life Technologies
Platinum PCR supermix high fidelity	Life Technologies
QIAEX II Gel extraction kit	Qiagen
Qiagen spin mini prep	Qiagen
SeeBlue plus2	Invitrogen
Skimmed Milk Powder	Marvel
Super Optimal Media (S.O.C)	Invitrogen
Super Signal West Pico Chemiluminescent substrate	Thermo Scientific
T4 DNA ligase	New England Biolabs
T4 DNA ligation buffer	New England Biolabs
T4 polynucleotide ligase	New England Biolabs
Trichlorocetic Acid (TCA)	Life Technologies
Triton-X 100	Sigma Aldrich
XL GOLD ultracompetent cells	Life Technologies
1 Kb plus Ladder	Invitrogen
10 X DNA loading buffer	Invitrogen
10mM dNTPs	Life Technologies
10 X PFU Buffer	Thermo Scientific

Table 2-2: Solutions

Solution	Composition
PBS-Tween (PBST)	137 mM NaCl, 2.7 mM KCl, 10 mM NaHPO$_4$, 1.8 mM KH$_2$PO$_4$
Coomassie Blue	500ml dH2O, 400ml methanol, 100 ml acetic acid, 0.5g Coomassie blue R250

2.2 Media, Strains, Growth of bacteria and cell culture

2.2.1 Growth Media

Minimal Essential Media, Hepes modification, (MEM-HEPES; Sigma M7278) was purchased from Sigma-Aldrich. Luria- Bertani (LB) and Terrific broth (TB) growth media were prepared using the components in table 2-3, in de-ionised water, and sterilised by autoclaving. Solid media was prepared by adding 15 g/L agar with de-ionised water and autoclaving. "Motility agar" (0.25 % agar) was prepared by adding 2.5 g agar to 1 L de-ionised water, prior to autoclaving, followed by addition of 4 ml 50 % v/v filter sterilised glucose once media had cooled to between 45 – 55°C.

Table 2-3: Growth media recipes

Media	Component added per litre
LB	10 g tryptone
	10 g NaCl
	5 g yeast extract
TB	12 g tryptone
	24 g yeast extract

2.2.2 Strains

Strains used in this study are detailed in table 2-4. Throughout experiments TUV 93-0 and Zap193 were used as "wildtype" strains.

Table 2-4: Strains

Strain	Details
TUV 93-0	Derivative of *E. coli* O157:H7 EDL933, shiga toxin negative
Zap193	NCTC 12900, O157:H7 shiga toxin negative
ΔTir	TUV 93-0, Tir deletion (Roe laboratory)
ΔAdhE	TUV 93-0, AdhE deletion (Emmerson et al., 2006)
TUV RFP	TUV 93-0, chromosomally expressing RFP (Roe laboratory)
ΔTir RFP	TUV 93-0 Tir deletion, chromosomally expressing RFP (Roe laboratory)

2.2.3 Storage of bacteria

Strains were stored as glycerol stocks at -80 °C and re-streaked from stock onto LB plates with appropriate antibiotics. Glycerol stocks were produced by growing bacteria in 5 ml LB overnight with appropriate antibiotics, and mixing with a final volume of 25 % v/v glycerol before transfer to -80 °C. LB plates were kept for a maximum of 2 weeks at 4 °C before being re-streaked from stock.

2.2.4 Growth of bacteria

Strains were inoculated into LB media or streaked onto LB plates with antibiotic when required. Antibiotics were used at a final concentration of; Kanamycin 50 μg / ml, Chloramphenicol 25 μg / ml and Ampicillin 100 μg / ml. For most experiments 5 ml broth was inoculated with a single colony and grown overnight at 37 °C, with shaking (220 rpm).

2.2.5 Cell culture

HeLa cells, purchased from European Collection of Cell cultures (ECACC), were maintained in Minimal Essential Media, Hepes modification, (MEM-HEPES; Sigma M7278) supplemented with 1 % L-Glutamine, 1% Penicillin / Streptomycin and 10 % heat inactivated Foetal Calf Serum (FCS). Cells were incubated at 37 °C, 5 % CO_2 and split at between 75-90 % confluency.

2.3 Molecular techniques

2.3.1 Plasmids

Plasmids used in this study are listed in table 2-5.

Table 2-5: Plasmids used in this study

Plasmid	Antibiotic resistance	Source
Tir-phiLOV	Kanamycin	Created by Jayde Gawthorne
Map-phiLOV	Kanamycin	Created during this study
EspF-phiLOV	Kanamycin	Gawthorne et al
TirGFP	Chloramphenicol	Roe et al., 2004
MapGFP	Chloramphenicol	Roe et al., 2004
TirSTOP	Kanamycin	Created during this study
dsRed	Ampicillin	Fredlund, Pasteur
RFP	Kanamycin	Beckham, Glasgow University

2.3.2 Extraction of plasmid DNA

Plasmid DNA was extracted from overnight cultures from a single bacteria colony, grown with shaking (200 rpm) in 5ml LB at 37 °C, using QIAprep Spin Miniprep kit.

2.3.3 Oligonucleotide primers

Oligonucleotide primers used in this study are listed in table 2-6. All primers were synthesized by Invitrogen and used to amplify DNA by PCR.

Table 2-6: Primers used in this study

Name	Sequence 5' to 3'	Purpose
Map 5'	CGA GAT CTG CAC ACT CCA GTA TCC ATT CA	To amplify *map* region for cloning
Map 3'	CGG GTA CCC AAT CGG GTA TCC TGT ACA TG	of *map*LOV
MapLOV1	CCA GAC GAA GAT GAT AGA	Screen for
MapLOV2	CGA AGA TGA TCG GAT AAT	MapLOV
AdhE F	CAA TAC GCC TTT TGA CAG CA	Colony PCR to
AdhE R	TAA TGG CGT TAC GGG TCT TC	confirm ΔAdhE
ΔAdhE R	ACC TTA CCG ATT TTC GAG GTG	escape mutants
TirSTOP F	CTG ATC GAG AAG AGC TTT GTG ATT ACC GAC CCG CGT CAG ATC	To produce TirSTOP insert
TirSTOP R	TGA CGC GGG TCG GTA ATC ACA AAG CTC TTC TCG ATC AGG TAC	

2.3.4 Agarose gel electrophoresis

To produce a 1 % w/v agarose gel, 1 g of agarose was added to 100 ml 1x Tris-acetate-EDTA (TAE) and dissolved by heating. Upon cooling, and prior to pouring the gel, 10μl SYBR®- Safe DNA gel stain (Thermo Fischer) was added. Samples were loaded in the appropriate volume of 10 X loading buffer (Invitrogen) and run at 100 V for 45-50 minutes. 1 kb plus ladder (Invitrogen) was run next to samples and gel was viewed using an Alphalmager transilluminator (Alpha Innotech, UK.)

2.3.5 DNA gel extraction and purification

DNA was separated by electrophoresis and bands were viewed using a transilluminator. Target bands were cut using a sterile scalpel and purified using QIAEX II Gel extraction kit, as per manufacturers' instructions. The final volume was eluted in 50 μl nuclease free water.

2.3.6 Restriction enzyme digest

Restriction digests were carried out in a total volume of 30 μl at 37 °C for a maximum of 5 hours, depending on buffer. The reagents used are listed in table 2-7. Typically a reaction consisted of 1 μl enzyme, 3 μl appropriate buffer, 20-25 μl DNA and made up to 30 μl with nuclease free water.

Table 2-7: Restriction enzyme digest reaction mix

Reagent	Source
BamH1- HF	NEB
Kpn1-FD	Fermentas
Bgl2	NEB
Cutsmart buffer	NEB
FD 10X Buffer	Fermentas

2.3.7 Ligation reaction

Ligation reactions were carried out at room temperature for 30 minutes using a rapid DNA ligase kit (Roche). The ligation mix was composed of 10 μl of 2X T4 DNA ligation buffer, 1 μl T4 DNA Ligase, a ratio of insert : vector, 3 : 1 and made up to 20 μl with de-ionised water.

2.3.8 Heat shock transformation

XL-10 GOLD ultracompetent cells, 25 μl, were thawed on ice and mixed with 5 μl of ligation product. Reaction mix was incubated on ice for 10 minutes, at 42 °C for 1 minute, then on ice for 2 minutes before being mixed with 300 μl pre-warmed Super Optimal Media (S.O.C) (Invitrogen). The reaction mix was incubated with shaking (200 rpm) for 30 minutes to 1 hour at 37 °C, before being plated onto appropriate antibiotic plates.

2.3.9 Production of electrocompetent E.coli

Strains from overnight cultures were sub-cultured at OD_{600} 0.05 in 25 ml LB plus appropriate antibiotic, at 37 °C, 200 rpm and grown to OD_{600} 0.6. All following steps were carried out on ice, with ice cold glycerol and centrifugation at 4 °C. Bacteria was incubated at 4 °C for 15 minutes, pelleted by centrifugation, 10 minutes, 5000 x g and re-suspended in 10 ml 10 % v/v glycerol. This step was repeated 3 times after which bacteria were re-suspended in 1 ml 10 % v/v glycerol and pelleted at 12,000 x g for 1 minute; repeated 5

times. Bacteria were re-suspended in a final volume of 150 μl 10 % v/v glycerol and used immediately in 40 μl aliquots.

2.3.10 Electroporation transformation

2 μl of plasmid or ligation mix (as a negative control) was mixed with 40 μl fresh electrompetent cells in an electroporation cuvette on ice. Transformation was carried out by electroporation using an eppendorf electroporator 2510 (Invitrogen, UK) at 2.5 kV for 4 seconds. 1 ml of LB was immediately added to the transformation mix and cells were incubated at 37 °C, with shaking, 200 rpm, for 90 minutes, after which 50 μl was plated onto LB plates with the appropriate antibiotic.

2.3.11 Isolation of ΔadhE escape mutants

Greater than 20 individual ΔadhE colonies were cultured overnight and inoculated onto LB motility agar plates as per motility assay protocol. Plates were incubated for >15 hours at 30 °C. Bacteria that were seen to regain motility, determined by spreading across the plate, were re-streaked onto fresh LB plates. To confirm the motile phenotype motility assay was repeated 3 times for each newly isolated escape mutant.

2.3.12 Confirmation of ΔadhE escape mutants

ΔadhE escape mutants were confirmed by colony PCR, using primers spanning the adhE region and reaction mix and conditions detailed in Table 5-1 and Table 5-2, Appendix I. For each escape mutant one colony was picked into 100 μl nuclease free water and incubated at 95 °C for 10 minutes and 2 μl was used in the PCR reaction.

2.3.13 Creation of Map-phiLOV

To investigate the potential of LOV as a fluorescent reporter to monitor expression and translocation of EHEC effector proteins, a Map-phiLOV plasmid was created.

The *map* region of pACYC MapGFP plasmid was amplified by PCR (reaction mix and conditions detailed in Tables 5-3 and 5-4, Appendix I) to produce a fragment of 1002 bp flanked by restriction sites 5' Bgl II and 3' Kpn I. The *map* fragment was isolated by gel electrophoresis and purified using QIAEX II Gel extraction kit, then digested with Bgl II and Kpn I.

pACYC *Tir*-phiLOV was digested sequentially with BamH I and Kpn I and run on an agarose gel to obtain the LOV backbone, product size 2880 bp. The plasmid backbone was purified using QIAEX II Gel extraction kit, as per manufacturers' instructions.

In between each step in digestion, the reaction mix was treated with MN: DNA, RNA and Protein purification kit (Macherey- Nagel GimbH & Co.) to clean-up and purify product.

Ligation of the map fragment and LOV backbone was carried out to create Map-phiLOV, which was subsequently transformed into XL GOLD ultra-competent cells by heat shock transformation. Map-phiLOV positive colonies were confirmed by colony PCR (reaction mix and conditions in Tables 5-5 and 5-6, Appendix I) and double digest with BamH I and Kpn I. The plasmid was purified and transformed into competent E. *coli* Zap193 by electroporation.

2.3.14 Creation of Tir-STOP

A TirSTOP plasmid was engineered to explore the efficiency of plasmid derived Tir expression and translocation. The plasmid pACYC Tir-phiLOV was digested with Bgl II and Kpn I and run on agarose gel to isolate the pACYC-Tir plasmid backbone (lacking phiLOV). Self-annealing primers containing a STOP codon were allowed to anneal and then phosphorylated to produce a STOP codon insert. Tir-STOP was produced by ligation of the Tir backbone and Tir-STOP insert and transformed into XL GOLD ultra-competent cells by heat shock transformation

2.4 Characterisation of strains

2.4.1 Ion Torrent Sequencing

To explore the reason behind the difference in bacteria motility between TUV 93-0, Δ*adh*E and escape mutants, whole bacterial genomes of Δ*adh*E and escape mutants M1-M7 were sequenced using an Ion Torrent PGM, to identify any differences between the bacterial genomes. This was done in collaboration with Leena Nieminin at the University of Strathclyde, Glasgow.

2.4.1.1 DNA library preparation and barcoding

Strains M1 – M7 were selected for sequencing as they presented a diverse range in regain of motility in comparison with Δ*adh*E and TUV 93-0 (as determined by motility assay),

with M2 and M4 showing greater motility than TUV 93-0 (55.0 mm and 63.75 mm diameter compared to TUV 93-0, 46.08 mm diameter). Prior to sequencing, a 200 bp gDNA library was prepared for each strain. Genomic DNA was enzymatically fragmented to blunt-ended DNA fragments, using NEBNext Fast DNA Fragmentation and Library Prep set. The fragmented gDNA was ligated to ion-compatible barcode adapters, supplied in Ion Xpress Barcode Adaptors, 1-16 kit. gDNA from a total of 8 strains was fragmented and ligated to individual barcodes, 1 to 8. Barcoded library was amplified by PCR (conditions in Table 5-7, Appendix I) and quantified using Quibit 2.0 fluorometer. Amplified barcoded library was size-selected for 350 bp, by running on E-gel SizeSelect Agarose Gel and iBase unit against 50 bp DNA ladder. Library integrity was tested using a Bioanalyzer. Figure 2-1 summarises gDNA library preparation workflow.

2.4.1.2 Library quality assessment

Barcoded libraries were pooled into groups of 4 samples and diluted to a concentration of 26 pM, then run on Ion OneTouch 2 (OT2) system. OT2 prepares Ion sphere particles (ISPs) containing clonally amplified DNA library. Following OT2 a quality control step, using Quibit 2.0 fluorometer, was performed to determine the amount of library bound to ISPs (template positive ISPs). In order to proceed to the next step a value of between 10 – 30 %, template positive ISPs was required. Lower than 10 % is indicative of low loading quality and greater than 30 % suggests an excess of polyclonals in the library. Template-positive ISPs were enriched using Ion OneTouch ES to leave only poly and monoclonal ISPs. The above steps were carried out as per the manufacturer instructions.

2.4.1.3 Ion Torrent sequencing

Pooled libraries were loaded onto an Ion 316 Chip (4 strains per chip) and run using Ion Torrent PGM, as per instructions from Life Technologies. An overview of the Ion Torrent sequencing process is displayed in Figure 2-2.
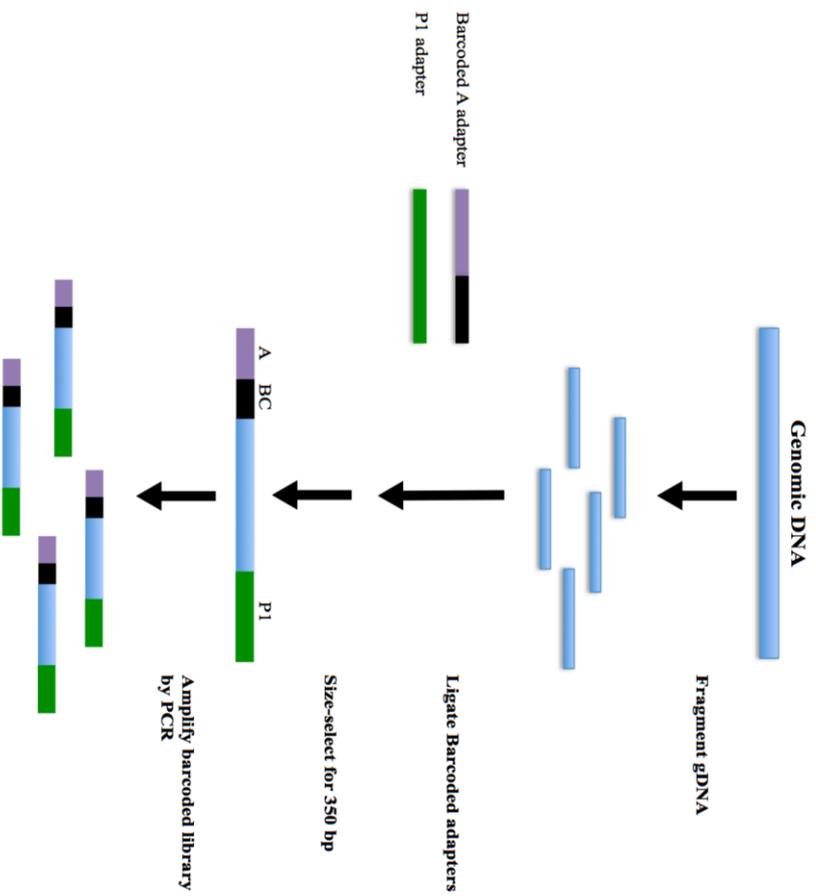
Figure 2-1: gDNA library preparation work flow.
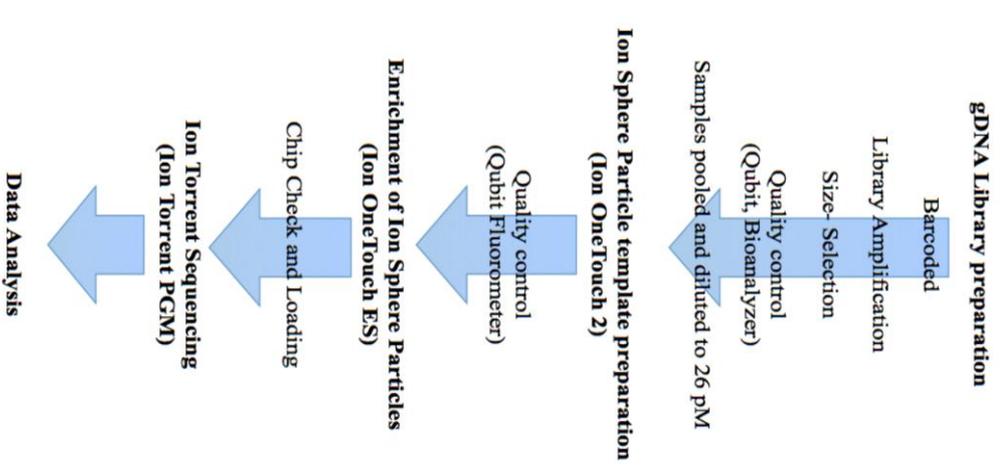
Figure 2-2: Ion torrent sequencing overview.

2.4.2 Protein secretion assay for determining expression of phiLOV constructs

To determine whether the phiLOV recombinant proteins were secreted by Zap193 bacteria a secretion assay was performed with Zap193 bacteria bearing either the *Map*-phiLOV or *Tir*-phiLOV plasmid. Zap193 was used in these experiments as this strain is shown to have higher expression of Tir than TUV 93-0 (Roe et al., 2004).

Strains from overnight cultures of Zap193, Zap193 *Tir*-phiLOV and Zap193 *Map*-phiLOV, were sub cultured, at an OD_{600} of 0.05, into 50 ml pre-warmed MEM supplemented with Hepes, to promote T3SS expression. Cultures were incubated at 37 °C, with shaking (200 rpm) until growth reached an OD_{600} of between 0.6 and 0.8. Bacteria was pelleted by centrifugation at 4000 x g for 10 minutes and supernatant was collected, carefully as to not disturb pellet, and incubated at 4 °C overnight with 10 % v/v Trichlorocetic Acid (TCA) to precipitate secreted proteins. Precipitated proteins from the secreted protein supernatant were collected by centrifugation at 4000 xg for 1 hour, 4 °C, dried and re-suspended in 1.5 M Tris (pH 8.8). Secreted proteins were then analysed by SDS PAGE and Western blot analysis for anti phi LOV.

2.4.3 Phi-LOV fusion expression assay, by measure of fluorescence

To assess expression of the phiLOV constructs, Zap193 transformed with either *Map*-phiLOV, *Tir*-phiLOV or *Tir*-GFP (as a positive control for fluorescence) were sub-cultured from overnight cultures into 50 ml MEM supplemented with Hepes, as per secretion assay. At 0, 2, 3, 4, 5 and 6 hours after sub culture optical density of the culture was recorded and fluorescence was measured using FLUOstar microplate reader (OPTIMA). Cultures were grown to a maximum of OD_{600} 1.0.

2.4.4 SDS PAGE

Samples were mixed with NuPAGE 4X sample buffer (Invitrogen) at a ratio of 3 : 1 and heated at 95 °C for 10 minutes. 5 μl of sample was loaded onto a NuPAGE 4-12 % Bis-Tris pre-cast gel (Invitrogen) and run at 150 V for 50 minutes in MES Buffer (Invitrogen). SeeBlueplus 2 ladder was run to determine molecular weight of sample bands. SDS-PAGE gels were stained with Coomassie blue stain (500 ml dH2O, 400 ml methanol, 100 ml acetic acid, 0.5 g Coomassie blue R250) for 30 minutes and de-stained overnight with de-stain solution (500 ml dH2O, 400 ml methanol, 100 ml acetic acid).

2.4.5 Western blot analysis

Proteins, separated on SDS-PAGE gel, were transferred onto a nitrocellulose membrane (Amersham, UK) in MES transfer buffer at 30 V for 90 minutes. The membrane was blocked overnight in PBS-Tween (PBS-T) (0.01%), 8 % skimmed milk (Marvel, UK). The membrane was incubated with primary antibody, anti phiLOV (gift from Jayde Gawthorne), (1:1000 in PBS-T, 8% skimmed milk) for 1 hour at room temperature, gentle rocking, washed 3 times for 10 minutes each in PBS-T, and incubated with secondary antibody, anti Rabbit (1:10,000 in PBS-T, 8% skimmed milk), for 1 hour room temperature, gentle rocking. Western blots were developed with 4 ml v/v Super Signal West Pico Chemiluminescent Substrate (Thermo Scientific, UK).

2.4.6 Motility assay

10 µl of overnight culture of TUV 93-0, Δ*adh*E and escape mutants 1-10 grown in LB, was normalised to an OD_{600} 0.7 by addition of LB and inoculated onto the centre of a fresh LB motility agar plate (recipe in Table 2-3) and incubated overnight (12 hours) at 30 °C. The diameter of the bacteria swim zone was measured. This was repeated at least 3 times for each strain.

2.5 Immunofluorescense microscopy

2.5.1 Flagella staining

Bacteria, TUV 93-0, Δ*adh*E and escape mutants M1-M10, were grown in 5 ml LB overnight then sub-cultured at an OD_{600} 0.05 and grown to OD_{600} 0.6, at 37 °C with shaking, 200 rpm. Culture was diluted 1/100 in PBS/2% PFA and air dried onto coverslips. Coverslips were washed 3 times to remove PFA and flagella were stained with anti H7 (1:1000 in PBS), 45 mins, then alexafluor 488 (1: 2500 in PBS), 45 mins. Coverslips were mounted using Dako fluorescent mounting media and imaged using a Zeiss M1 Axioskopp microscope at objective x 100. De-convolution and post processing of images was carried out using Zen image analysis software.

2.5.2 Bacteria adhesion to host cells and cell lesion assay

HeLa cells were seeded into 12 well plates at 5 x 10^4 cells/well and grown for a minimum of 20 hours in 37 °C, 5 % CO_2 prior to infection. Bacterial strains, TUV 93-0, Δ*tir* or Δ*tir*-phiLOV, were cultured in 5 ml LB, plus appropriate antibiotics, (37 °C, 200 rpm) to OD_{600} 0.6, then sub cultured, 10 µl into 5 ml MEM-Hepes, and incubated for 12 hours, 37 °C,

5 % CO_2 non-shaking. Cells were washed once with 1 X PBS and infected at a multiplicity of infection (MOI) of 10:1 bacteria: cell (5 x 10^5 bacteria / ml in MEM-Hepes plus 0.01% glucose). Plates were centrifuged at 800 x g and incubated at 37 °C, 5 % CO_2 for 2 hours. Infection was carried out for a total of 5 hours, during which cells were washed once with PBS after 2 hours infection to remove unbound bacteria.

2.5.3 Fixation and staining for fluorescent imaging of bacterial adhesion to host cells

Cultures on coverslips in 12 well plates were fixed with 200 μl 2% PFA for 30 min (15 min on bench, 12 min gentle rocking). Coverslips were washed 4 times with 1 X PBS then cells permeabilised with 200 μl, 0.05% Triton X-100, for 5 min with gentle rocking. Wells were washed 4 times with 1 X PBS then blocked for 45 min with 0.1 % BSA / PBS, rocking. Bacteria were stained with primary antibody, α O157 (dilution 1: 100) in 0.01% BSA / PBS for 45 mins, rocking, then washed 3 times with 1 X PBS and incubated with secondary antibody, rabbit alexafluor 555 (1:1000). Host cell actin was stained with phalloidin 488 or 647 (1:1000) in 0.01% BSA / PBS, at room temperature, for 45 min, rocking. Coverslips were washed 3 times with 1 X PBS for 5 min each then mounted onto 5 μl Dako mounting media on glass slides.

2.6 Data collecting and processing

2.6.1 phiLOV construct expression

Reporter fusion expression was measured by monitoring fluorescence (as described in section 2.4.3) and plotted as Fluorescence over OD_{600}. Experiments were repeated 3 times and mean fluorescent values were plotted, with error bars showing standard error of mean (SEM).

2.6.2 A/E Lesion quantification

A/E lesions were quantified using Fiji / image J for 100 cells per condition. Bacteria were highlighted and set as a ROI, and the total area of bacterial pixels was recorded. Within this ROI a threshold was set to select actin pedestals pixels. A ratio of the total pixel area of the pedestals / total pixel area of bacteria was calculated per cell. Uninfected cells were also taken into account and included in the 100 cells counted. The ratio was plotted per condition along with standard error of the mean (SEM) for the samples.

2.6.3 Sequencing Alignment and single-nucleotide polymorphism (snp) identification

Whole bacterial genomes were sequenced using Ion Torrent PGM, as described in section 2.4.1. Reads were trimmed using CLC then "groomed" by FastQ groomer to convert them to a format compatible with SPADES. SPADES was used for de novo sequencing of reads and Mauve used to re-order contigs against a reference genome. The reference genome used in this case was *E. coli* O157: H7 EDL 933, downloaded from NCBI. Re-ordered contigs were annotated using Prokka.

3. Results

3.1 A role for AdhE in EHEC motility.

An *adh*E knock out mutant, Δ*adh*E, created by Baba et al., 2006, was shown to display down-regulated T3SS expression and up-regulated flagella expression (Beckham et al., 2014). However, despite up-regulation of flagella, the bacteria are immobile. Work carried out in this study isolated several Δ*adh*E escape mutants (M1 – M10) which had re-gained motility. These were phenotypically characterised and sequenced to determine the mechanism by which motility was re-gained and thus explore the role of AdhE, and acetylation, in flagella mediated motility.

3.1.1 Isolation of ΔadhE escape mutants.

When grown on LB motility agar for a period of >17 hours, *E. coli* Δ*adh*E mutants were seen to spontaneously regain motility. Ten Δ*adh*E escape mutants, were isolated (Materials and Methods 2.3.12) and the *adh*E gene deletion confirmed by colony PCR (Figure 3-2). Two PCR reactions were performed for each escape mutant with the primers shown in Figure 3-1. The extent by which each mutant had regained motility was examined by motility assay. M1 – M10 were grown on motility agar for 12 hours at 30 °C and the diameter of growth was measured and compared to that of TUV 93-0 and Δ*adh*E (Materials and Methods 2.4.6). The diameter of the bacterial growth was taken as an indication of the extent of bacterial motility, a larger diameter being indicative of greater mobility. Motility assay was repeated 3 times for each strain. Each escape mutant regained motility to a different extent but each was significantly more motile than Δ*adh*E (un-paired t-test: $p < 0.01$). M2, M3 and M4 were found to be most motile with M2 and M4 (55.0 mm and 63.75 mm) spreading further than TUV 93-0 (50.97mm) (Figure 3-3).
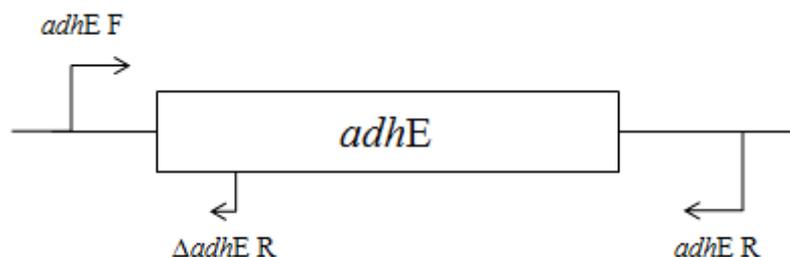
Figure 3-1: Primers for confirmation of AdhE deletion by PCR reaction.

For each strain PCR reaction was set up with forward primer *adh*E F and either Δ*adh*E R or *adh*E R. If *adh*E is present then amplification with *adh*E F and Δ*adh*E R yields a product of 900 bp and amplification with *adh*E F and *adh*E R yields a product of 2680 bp. If *adh*E is absent amplification with *adh*E F and *adh*E R yields a product of 1000 bp.
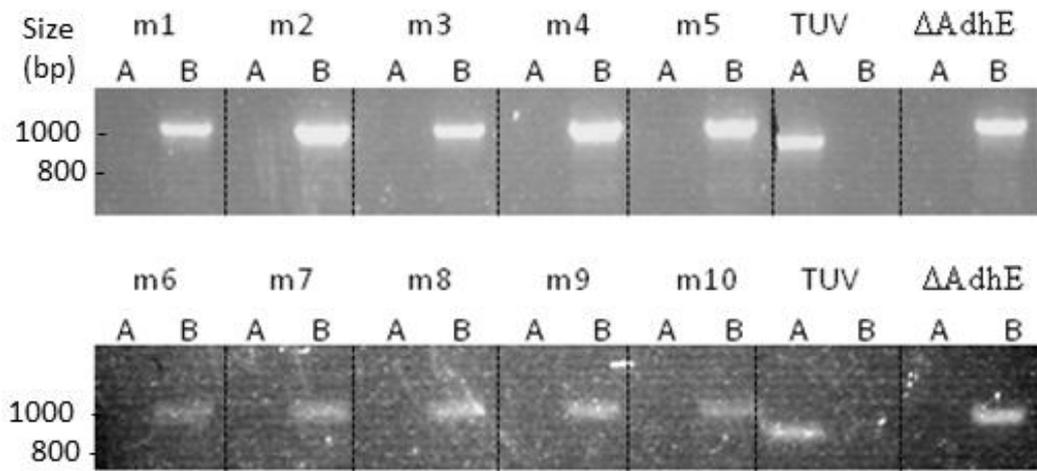
Figure 3-2: Colony PCR confirmation of Δ*adh*E escape mutants. Colony PCR confirmed the deletion of AdhE in the ΔAdhE escape mutants. PCR reaction with primers *adh*E F and Δ*adh*E R, lane "A", show AdhE is present in TUV 93-0 but not in Δ *adh*E or M1 – M10, indicated by a band at 900 bp. PCR amplification with primers *adh*E F and *adh*E R, lane "B", show a band at 1000 bp in M1 –M10 and Δ*adh*E, indicating that the AdhE gene is absent.
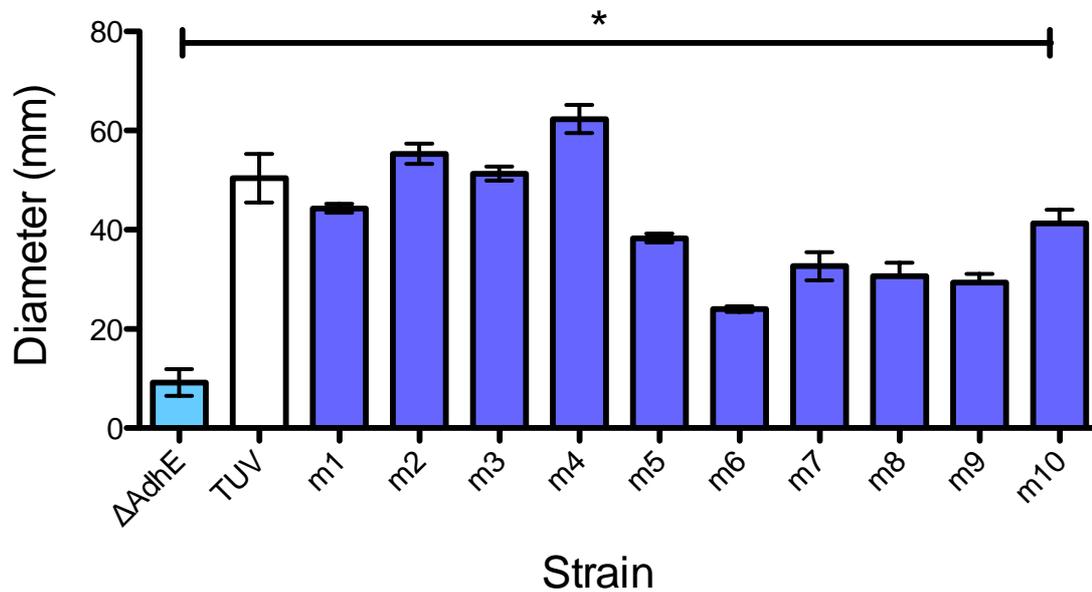
Unpaired T-Test p value < 0.01

Figure 3-3: motility of Δ*adh*E escape mutants

The motility of M1-M10, Δ*adh*E and TUV 93-0 was compared by motility assay, Materials and Methods 2.4.6. Motility assay was performed in triplicate and the mean diameter of swimming zone is displayed in mm. Error bars show SEM from 3 individual experiments. An unpaired t-test was performed and shows that Δ*adh*E is significantly less motile than TUV 93-0 and M1 – M10.

3.1.2 Flagella expression is up-regulated in the ΔAdhE escape mutants

Flagella expression is up-regulated in Δ*adh*E, (Beckham et al., 2014), thus it was important to establish whether M1 − M10 displayed a similar expression profile. Protein secretion profiles of M1- M10 showed a large band at 62 KDa, confirmed to be FliC, the flagellin component, by western blot analysis. This band was also seen for Δ*adh*E, but not TUV 93-0 (Figure 3-4 A and B). Additionally, flagella were visualised by fluorescent microscopy, by staining with anti-FliC primary antibody and alexfluor 488 secondary antibody (Materials and Methods 2.5.1). It was seen that M1 − M10 express numerous and long flagella, similar to that of the Δ*adh*E mutant. Western blot analysis revealed up-regulation of flagella in Δ*adh*E and all of the Δ*adh*E escape mutants, compared with TUV 93-0, with the exception of M1 (Figure 3-4). However, flagella up-regulation was observed in M1, both by immuno-fluorescent microscopy and protein secretion profile, thus it would be beneficial to repeat western blot analysis for FliC for this sample.

A.
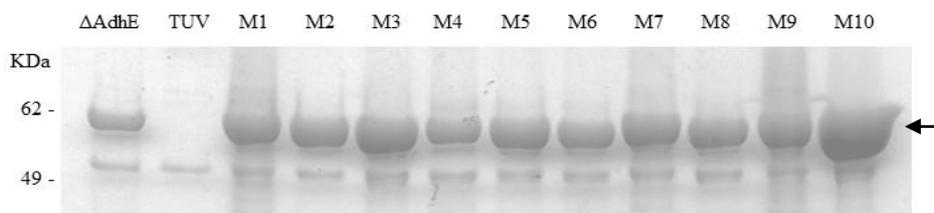
B.
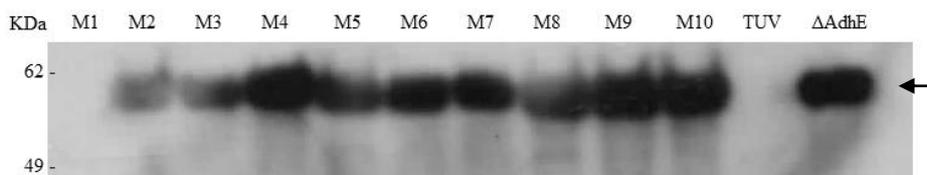
Figure 3-4: Flagella expression. Protein secretion profile of Δ*adh*E and M1- M10 showed the presence of a large band at 62 KDa which was lacking in TUV (A). Western blot analysis for anti-fliC confirmed this band to be FliC, a structural component of the flagella, suggesting up-regulation of flagella in Δ*adh*E and the majority of the escape mutants (B).

3.1.3 Whole genome sequencing of ΔadhE escape mutants identified snps in motility genes

It was established that the escape mutants display a similar profile to ΔadhE in terms of flagella expression; however, spontaneous mutation has given rise to regain of motility. To identify gene regions in which SNPs mutations resulted in regain the motility phenotype, seven ΔadhE escape mutants (M1 – M7) along with ΔadhE mutant strain, were selected to be sequenced using Ion Torrent PGM sequencing. A bar coded genomic DNA library was prepared for each strain and libraries were pooled, M1, M2, M3 and M4 or M5, M6, M7 and ΔadhE. Two Ion 316 chips were loaded, chip A with pooled libraries M1 – M4 and chip B with pooled libraries M5 – M7 and ΔadhE (Materials and Methods 2.4.1). The quality of the sequencing data was reported in a PGM sequencing Run Report. Detailed information on chip loading, library quality and alignment to the reference strain for libraries on both 316 chips can be found in Appendix II. Overall, the resulting data from sequencing was of sufficiently high quality for post processing and further analysis.

Prior to comparative analysis, the quality of the sequencing data was assessed and only sequences from those strains with high sequencing coverage were included in the final data assessment. The strain coverage indicates the number of times that each nucleotide is sequenced, and a higher coverage reduces the risk of interpreting sequencing error as genuine variance between sequenced strain and reference strain. The coverage for the majority of strains on both chips was seen to be close to 30 X. However, strains M1 and M3 showed particularly low coverage, 11.53 and 2.29 respectively, and were thus excluded from further analysis (Table 3-1).

Table 3-1: Coverage and Mean read length of sequenced strains.

Strain	Coverage	Mean Read Length (bp)
ΔAdhE	31.00	200
M1	11.53	205
M2	38.88	195
M3	2.29	173
M4	66.00	181
M5	30.97	215
M6	29.49	217
M7	29.01	213

3.1.4 Genome Assembly and Alignment

Sequencing data output were in FastQ file format and were trimmed using CLC v 7.0, parameters set to exclude lengths of below 75 bp and above 250 bp. Trimmed files were assembled using SPADES (using default parameters) and information on the largest, smallest and mean contig length, along with the N50 and the N90 was calculated (Table 3-2).

Table 3-2: Sequencing statistics. Information in the Table below gives an indication of the quality of the sequencing data, as determined by the number and size of contigs for each strain, along with the N50 and N90. The N50 is the length of the smallest contig, in a set of contigs whose combined lengths total at least 50 %, or 90 % for the N90, of the assembly.

Strain	Total contigs	Largest (bp)	Smallest (bp)	Mean (bp)	N50	N90
ΔAdhE	848	374865	58	6109.6	107951	18196
M2	678	374839	56	7621.4	118573	20649
M4	493	374862	56	10446.9	123468	25166
M5	1151	342108	57	4534.8	109409	13324
M6	1149	374799	57	4539.3	109261	16643
M7	1063	286214	57	4900.2	107153	18194

3.1.5 Sequence alignment and annotation

To align the escape mutants, the Δ*adh*E genomic data was firstly aligned to the reference strain EDL 933, after which escape mutants were aligned to Δ*adh*E. Alignment of contigs was carried out using Mauve and contigs were concatenated with Artemis. Finally, PROKKA was used to annotate each strain allowing for genes containing SNPs to be identified.

3.1.6 Nucleotide variation in motility genes

As the phenotype most readily observed for the Δ*adh*E escape mutants was motility, it made most sense for initial searches for mutations to focus on the motility genes. A map of the motility genes in which initial searching was carried out is shown in Figure 3-5. M2, M4 and M6 were found to encode a 65 amino acid deletion at the C-terminal of CheY which resulted in a truncated version of the protein that lacked several residues involved in FliM binding. Analysis of M5 showed an insertion of 23 amino acids in FliM, and M7 showed a point mutation in FliC resulting in amino acid substitution of Alanine to Valine

at position 504. Full amino acid sequence alignments can be found in Appendix III and a summary of mutations in escape mutants can be found in Table 3-3.

In addition to this there are multiple 'silent' SNPs mutations in many of the genes associated with motility however, no others result in amino acid sequence alteration.
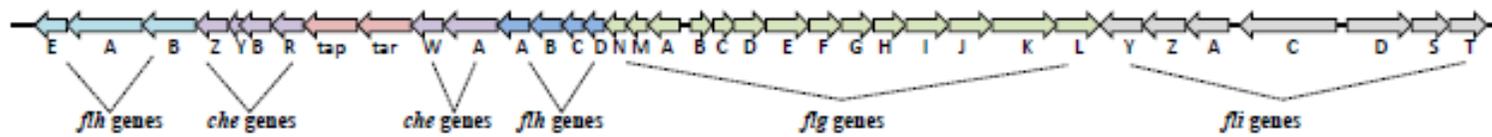
Figure 3-5: Map of *E. coli* O157:H7 motility genes. Sequencing analysis showed snps in multiple motility genes.

Table 3-3: Summary of mutations in motility genes of Δa*dh*E escape mutants. The underlined sections highlight the FliM and CheY binding site. Sections in red denote amino acid substitutions. *K* signifies lysine residues in CheY.

Gene	Strain	Sequence
CheY	EDL933, Δa*dh*E, M5, M7	MAD*K*EL*K*FLVVDDFSTMRRIVRNLL*K*ELGFNNVEEAEDGVDALN*K*LQAGGYGFVISDWNMPNMDGLELL*K*TIRADGA MSALPVLMVTAEA*KK*ENIIAAAQAGASGYVV*K*PFTAATLEE*K*LN*K*IFE*K*LGM
	M2, M4, M6	MAD*K*EL*K*FLVVDDFSTMRRIVRNLL*K*ELGFNNVEEAEDGVDALN*K*LQAGGYGFVISDWNMPNMDGLELL*K*TIRADGA MSALPVLMVTAEA*KK*ENI*MLRRKRGPVAMW*
FliM	EDL933, Δa*dh*E, M2, M4, M6, M7	MGDSILSQAEIDALLNGDSEVKDEPTASVSGESDIRPYDPNTQRRVVRERLQALEIINERFARHFRMGLFNLLRRSPDITV GXIRIQPYHEFARNLPVPTNLNLIHLKPLRGTGLVVFSPSLVFIAVDNLFGGDGRFPTKVEGREFTHTEQRVINRMLKLAL EGYSDAWKAINPLEVEYVRSEMQVKFTNITTSPNDIVVNTPFHVEIGNLTGEFNICLPFSMIEPLRELLVNPPLENSRNEDQ NWRDNLVRQVQHSQLELVANFADISLRLSQILKLKPGDVLPIEKPDRIIAHVDGVPVLTSQYGTLNGQYALRIEHLINPIL NSLNEEQPK
	M5	MGDSILSQAEIDALLNGDSEVKDEPTASVSGESDIRPYDPNTQRRVVRERLQALEIINERFARHFRMGLFNLLRRSPDITV GAIRIQPYHEFARNLPVPTNLNLIHLKPLRGTGLVVFSPSLVFIAVDNLFGGDG*ASRPKWKVASLPIPNSASSTACM*EGREFT HTEQRVINRMLKLALEGYSDAWKAINPLEVEYVRSEMQVKFTNITTSPNDIVVNTPFHVEIGNLTGEFNICLPFSMIEPLR ELLVNPPLENSRNEDQNWRDNLVRQVQHSQLELVANFADISLRLSQILKLKPGDVLPIEKPDRIIAHVDGVPVLTSQYGT LNGQYALRIEHLINPILNSLNEEQPK
FliC	M7	Alanine 504 → Valine

3.2 LOV: a novel fluorescent reporter to monitor effector proteins throughout infection.

The development of a fluorescent reporter that allows us to monitor the expression, translocation and host cell localisation of bacterial effector proteins throughout infection will greatly aid us in our understanding of the mode of action of EHEC, as well as other T3SS bearing bacteria. This thesis evaluates the use of LOV as a fluorescent reporter to tag bacterial effectors.

3.2.1 Production of Map-phiLOV

map-phiLOV was created by subcloning the gene encoding Map from an existing *map*-GFP plasmid into a vector containing phiLOV. *map* was amplified by PCR (Figure 3-6 A) and isolated by gel extraction. Tir was excised from *tir*-phiLOV and replaced with the *map* insert (Materials and Methods 2.3.14) to create *map*-phiLOV. The *map*-phiLOV plasmid was confirmed by restriction digest (Figure 3-6 B) and sequencing. Plasmid maps of both *map*-phiLOV and *tir*-phiLOV are shown in Figure 3-7.

3.2.2 Tir-phiLOV is expressed in Zap193 and TUV 93-0 background

To examine expression of effector-phiLOV fusions, Zap193 transformed with either *tir*-phiLOV or *map*-phiLOV was grown in T3SS-inducing conditions (Materials and Methods 2.4.3) and fluorescence was measured throughout growth. Both *tir*-phiLOV and *map*-phiLOV strains showed a gradual increase in fluorescence over time, indicating that both effector-phiLOV fusions are expressed in the Zap193 background (Figure 3-8 A). As a positive fluorescence control, TirGFP expression was also monitored (Figure 3-8 B) and seen to be much higher than both Tir-phiLOV and Map-phiLOV. Western blot analysis further confirmed the expression of Tir-phiLOV in Zap193 (Figure 3-9). Tir-phiLOV expression was found to be higher than that of Map-phiLOV, as indicated by a more pronounced increase in fluorescence, and was therefore selected for use in subsequent experiments. Consistent with this, western blot for Map-phiLOV produced very little signal compared to Tir-phiLOV, suggesting the fusion was poorly expressed (Figure 3-9).

To determine whether Tir-PhiLOV was expressed during infection, Tir-phiLOV was transformed into *tir* deletion TUV 93-0 strain, Δ*tir*, to produce Δ*tir*/tir-phiLOV, and used to infect HeLa cells. *Δtir*/tir-phiLOV was grown in LB to an OD_{600} of 0.6 and sub-cultured

over night in T3SS inducing conditions (Materials and Methods 2.4.2). HeLa cells, grown to ~70% confluency on cover slips in 12-well culture dishes, were infected at an MOI of 10 and contact between bacteria and host cell was promoted by centrifugation at 800 x g for 5 minutes. After 5 hours of infection, cells were fixed and stained for immunofluorescent microscopy (Materials and Methods 2.5.2 and 2.5.3).

From this experiment Tir-PhiLOV was seen to accumulate as a specific foci co-localised with the bacteria (Figures 3-10 and 3-11), confirming that Tir-phiLOV is expressed in the Δ*tir* background during infection. Previous work by Gawthorne et al,. (unpublished), confirmed that foci are due to phiLOV bound to Tir, and that non effector-bound phiLOV does not accumulate, as seen by an absence of foci in bacteria transformed with free phiLOV.
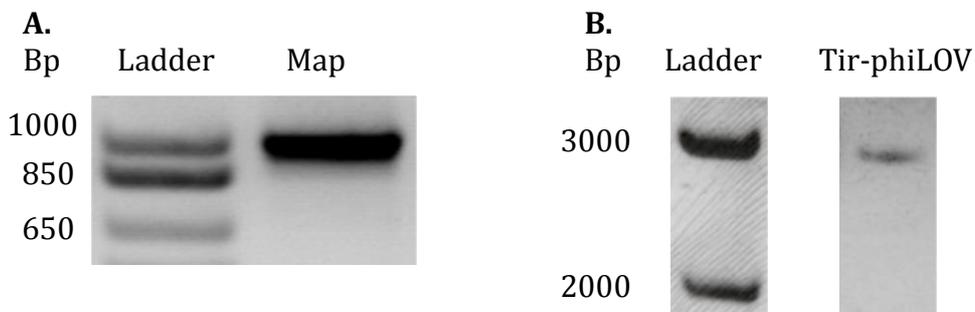
Figure 3-6: Production of *map*-phiLOV plasmid. (A) Visualisation of Map product at 1002 bp after amplification from *Map*GFP plasmid following separation by agarose gel electrophoresis. (B) Restriction enzyme digest of *Tir*-phiLOV with Kpn1 and BamH1 to isolate LOV backbone, 2880 bp.
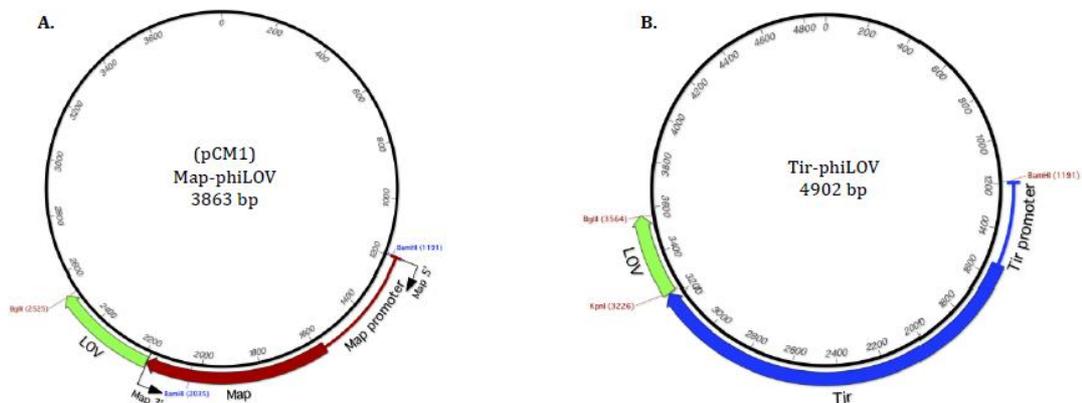
Figure 3-7: Plasmid maps. (A) *map*-phiLOV and (B) *tir*-phiLOV. The plasmid backbone for both *map*-phiLOV and *tir*-phiLOV is pACYC

A.

Fluorescence corrected for background

▲ Tir- phiLOV
■ Map- phiLOV

B.

Fluorescence corrected for background

● TirGFP
▲ Tir- phiLOV

* p < 0.05
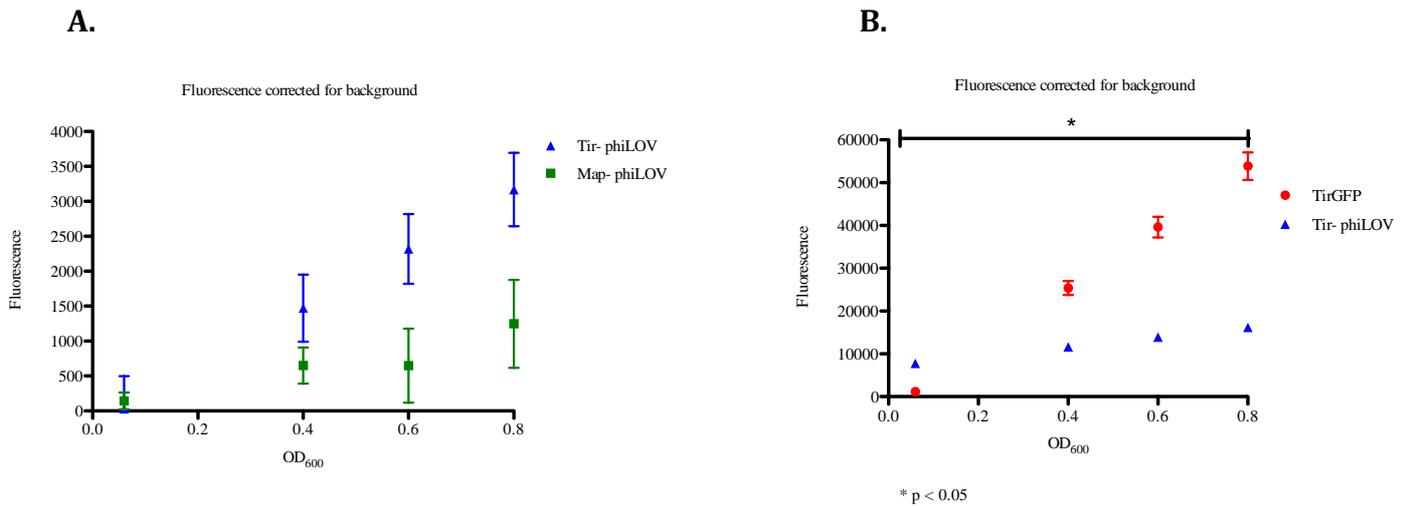
Figure 3-8 Expression of effector-LOV fusions in Zap193. Zap193 transformed with (A) the *tir*-phiLOV or *map*-phiLOV plasmid and (B) *tir*-phiLOV or *ti*rGFP plasmid were grown in T3SS-inducing conditions. Fluorescence was measured periodically throughout growth and corrected for background fluorescence, by subtracting fluorescence of Zap193 with no plasmid. TirGFP fluorescence was monitored as a positive control. TirGFP fluorescence was significantly higher than that of Tir-phiLOV ($p \leq 0.05$) at all recorded time points. Error bars show SEM of fluorescence values from 3 separate experiments.
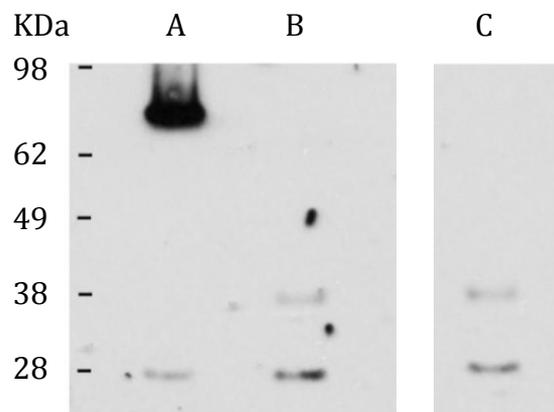
Figure 3-9: Tir-phiLOV is expressed in Zap193. Zap193 transformed with (A) *tir*-phiLOV, (B) *map*-phiLOV or (C) carrying no plasmid, were grown in T3SS-inducing conditions (Materials and Methods 2.4.2). Cell lysates were analysed by Western blot and effector-LOV fusions were detected with an αLOV antibody (Materials and Methods 2.4.5). The band in lane A indicates strong expression of Tir-phiLOV construct, highlighted by the arrow.
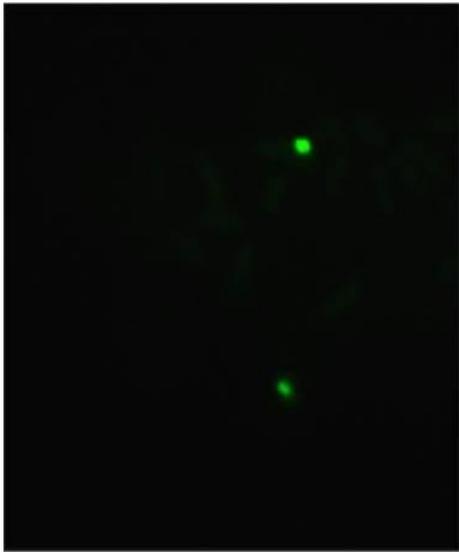
A.Tir-phiLOV **B. *E.coli* O157:H7** **C. Actin** **D. Merge**
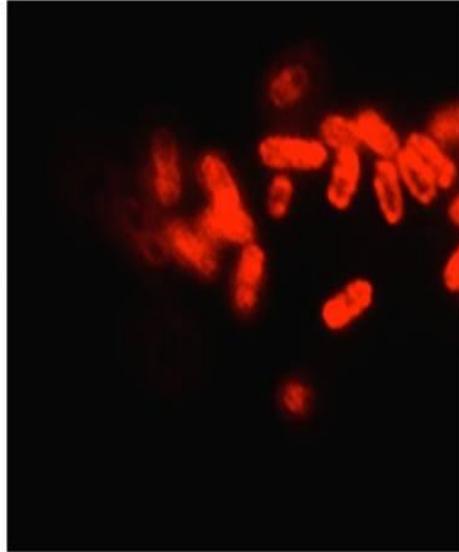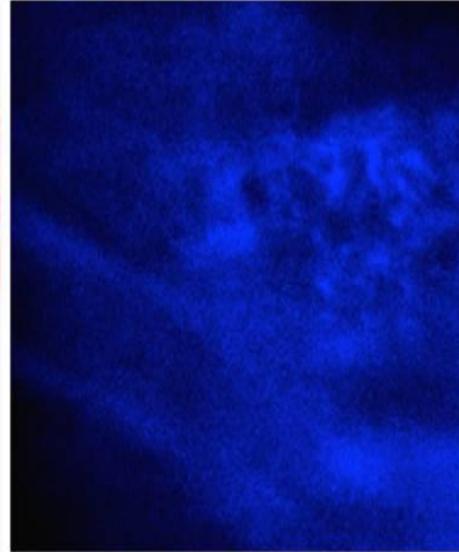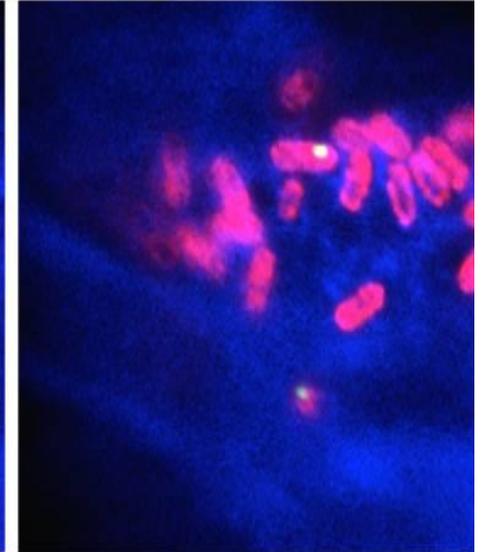
Figure 3-10 Fluorescent image of Tir-phiLOV expression. HeLa cells infected with Δ*tir/*tir-phiLOV for 5 hours were fixed and stained for fluorescent microscopy. Bacteria were treated with αO157 and stained with alexafluor 555, and actin stained with phalloidin 647. Images show Tir-phiLOV foci (A), *Δtir/*Tir-phiLOV outer membrane (B), host cell actin (C) and merge (D).
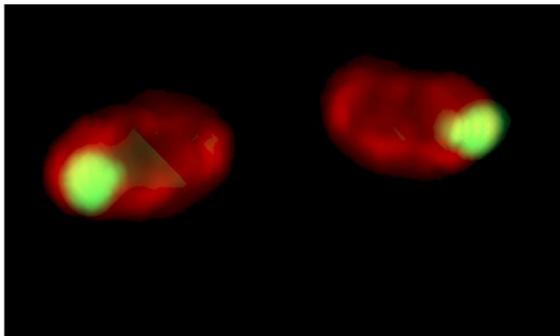
Figure 3-11 Tir-phiLOV is expressed in *Δtir* background. HeLa cells were infected with Δ*tir*/tir-phiLOV and fixed and stained with alexafluor555 (red) for bacteria. Tir-phiLOV can be seen as distinct green foci within the bacteria. The black and white image was false coloured using Zen image software, and rotated to show bacteria in 3D.

3.2.3 Tir- phiLOV can be translocated by the T3SS and remains functional

Having established that Tir-phiLOV is expressed by *E. coli* Zap193, fluorescent microscopy was used to determine whether Tir-phiLOV is exported by T3SS. Host cell lesion assays were set up using Δ*tir/*tir-phiLOV (Materials and Methods 2.5.2) and samples were fixed and imaged after an infection period of 5 hours. Post processing of the images showed Tir-phiLOV to be distinct and separate from the bacteria in the area between the bacterium and the host cell pedestal (at the bottom of the z-stack) (Figure 3-12), indicating that Tir-phiLOV had been exported.

Functionality of Tir-phiLOV was confirmed by examination of A/E lesion formation. A/E lesion formation in Δ*tir/*tir-phiLOV was compared with TUV 93-0. Cell adhesion assays used Δ*tir/*tir-phiLOV, TUV 93-0 or Δ*tir* (as a negative control) and infected at an MOI of 10, as described in section 3.2.2. After an infection of 5 hours, host cell actin was stained with phalloidin-647 and bacteria with anti-O157: alexafluor-555 (Materials and Methods 2.5.3). A/E lesions were indicated by an accumulation of actin beneath attached bacteria.

In cells infected with Δ*tir*, no pedestals were observed and few bacteria were found attached to cells, confirming that Tir is required for actin pedestal formation and tight bacterial attachment to cells. When Δ*tir* was complemented with *tir*-phiLOV pedestal formation was observed (Figure 3-13), demonstrating that Tir-phiLOV is functional.

To determine whether phiLOV interferes with Tir functionality A/E lesions were quantified. Greater than 100 cells per condition were imaged and A/E lesions quantified for each condition by determining the ratio of pedestal to bacteria pixel area (Materials and Methods 2.6.2). Figure 3-14 shows that Δ*tir* cells expressing Tir-phiLOV partially recovered the ability to form A/E lesions compared to TUV 93-0, indicated by a decrease in the ratio of pedestal to bacteria pixel area. This suggests that phiLOV partly interferes with Tir functionality.
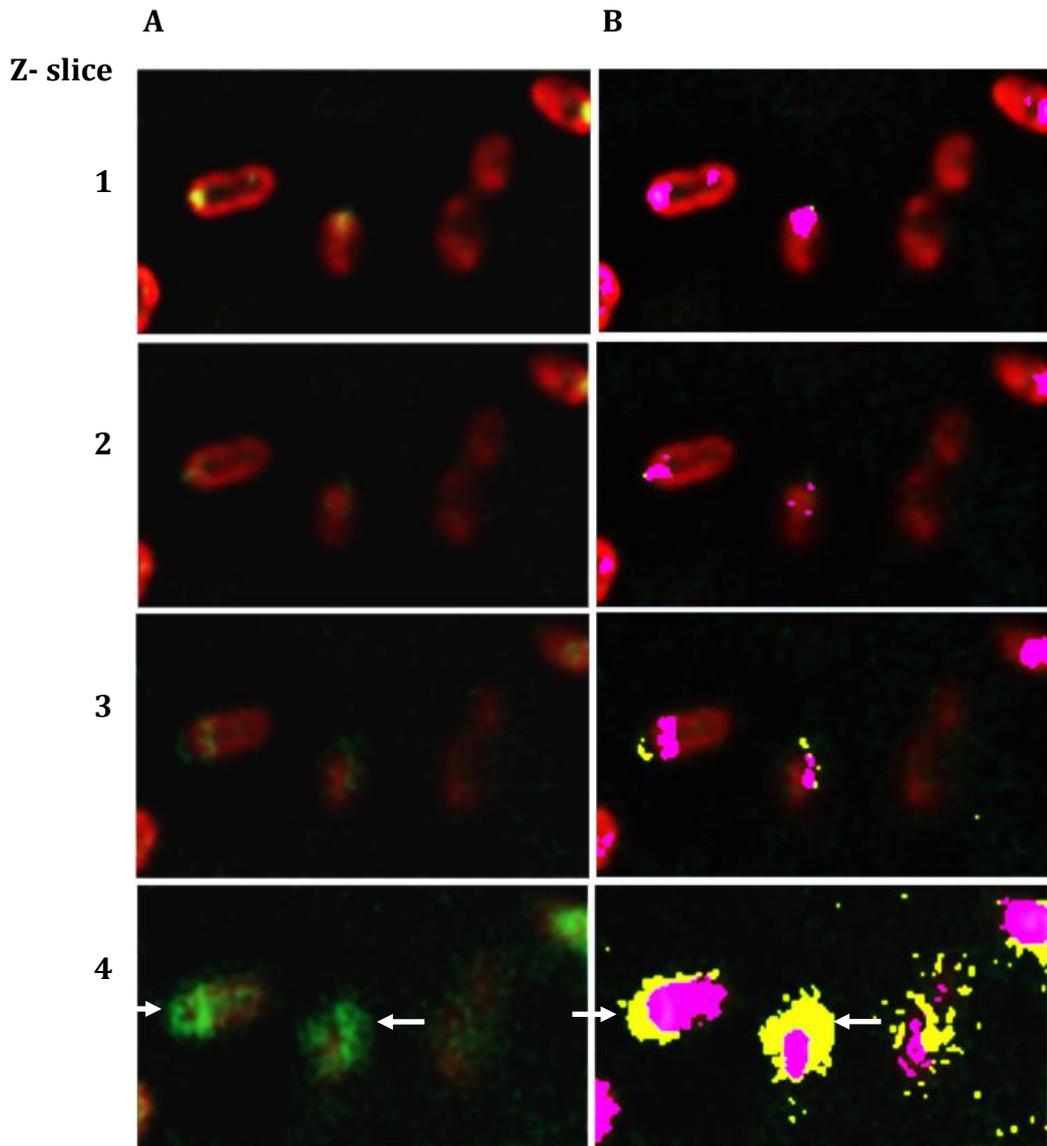
Figure 3-12 Tir-phiLOV is translocated. Expression and localisation of Tir-phiLOV in attached bacteria. The image shows Z-slices from the centre, slice 1, to the bottom, slice 4, of the bacteria (closest to the host cell). (A) Bacterial surface protein antigen O157 stained with alexafluor 555 (red) and Tir-phiLOV (green). Tir-phiLOV is seen as a specific foci in the centre of the bacteria then as a diffuse spread beneath the bacterium. (B) Image was false coloured to show co-localisation between Tir-phiLOV and O157. Pink shows areas where Tir-phiLOV is co-localised with the bacterial stain and yellow shows areas where there is no co-localisation. Note that beneath the bacterium, and closest to the host cell, the signal from Tir-phiLOV is distinct and separate (indicated by the white arrows), indicating that it has been translocated.
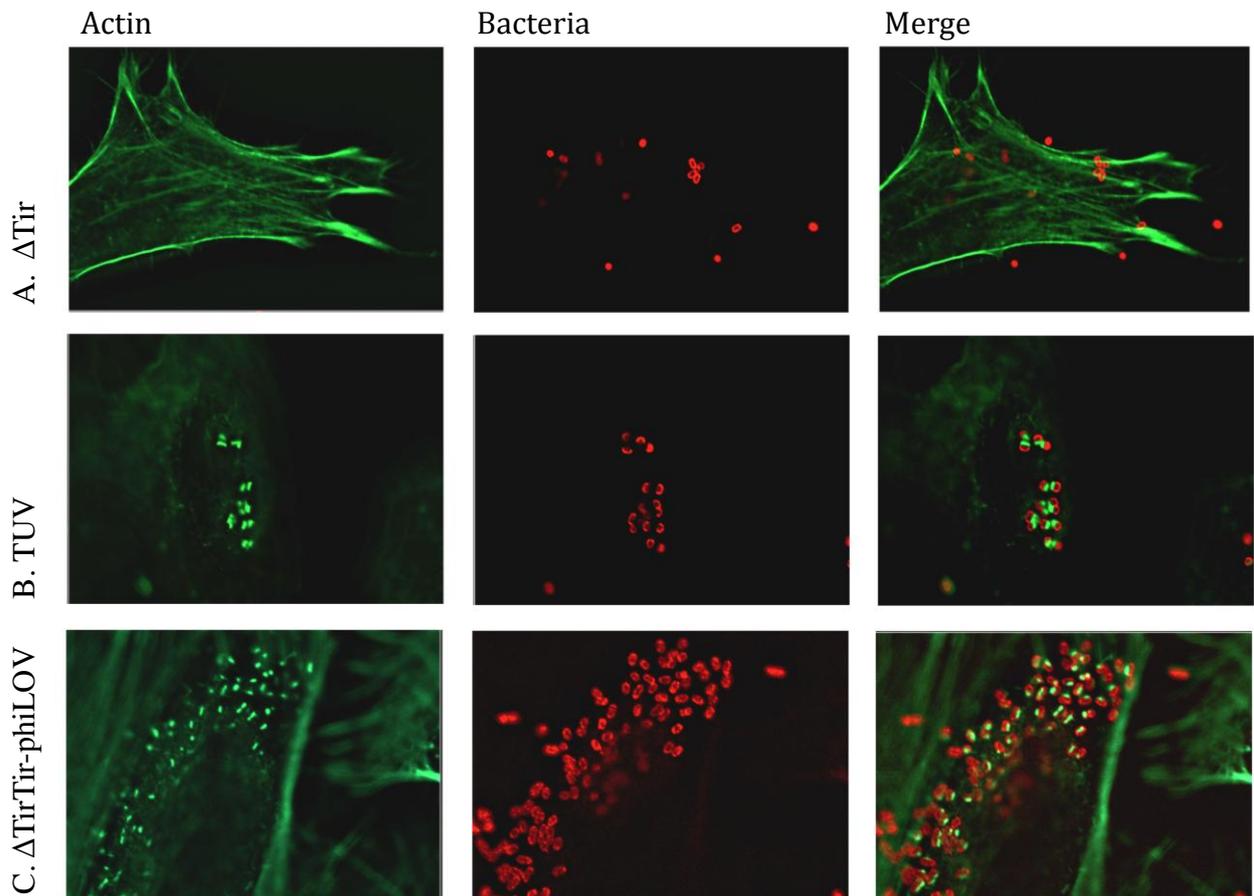
Figure 3-13: *Δtir/*tir-phiLOVis translocated and is able to generate A/E lesions.

HeLa cells infected with (A) Δ*tir*, (B) TUV 93-0 or (C) *Δtir/*tir-phiLOV, were fixed and stained for actin (green) with phalloidin 488 and bacteria (red) with anti O157: alexafluor555. Image shows that (A) deletion of *Tir* results in an inability to form A/E lesions, and (C) complementation with *tir*-phiLOV in the Δ*tir* strain results in regain of A/E lesion formation. TUV 93-0 was used as a control in this experiment as both Δ*tir* and Δ*tir* expressing Tir-phiLOV were generated from a TUV 93-0 background.
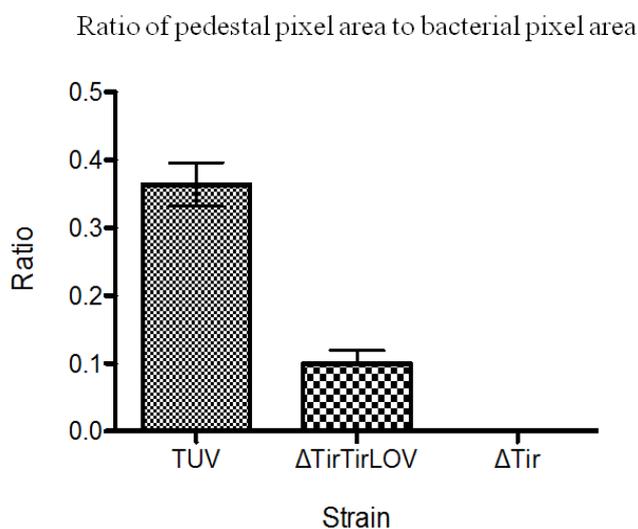
Figure 3-14: Quantification of A/E lesions. The ratio of pedestal to bacterial pixel area was quantified for greater than 100 cells per condition using ImageJ, and plotted to give an indication of each strains ability to form A/E lesions. In the Δ*tir* strain A/E lesion formation is completely inhibited. Partial regain of A/E lesion formation is seen in the *Δtir/*tir-phiLOV complement when compared to TUV 93-0. Error bars show the SEM for the mean ratio for each strain over 3 separate experiments.

4 Discussion

4.1 Discussion

A deeper understanding of the regulation and expression of EHEC virulence factors will lead to improved treatment and prevention of this disease. This thesis focuses on two major EHEC virulence factors and serves to provide insight into the regulation of flagella motility, as well as introduce a novel tool to fluorescently monitor the T3SS.

Firstly, this thesis explores the role of AdhE in EHEC flagella rotation regulation through the comparison of three phenotypically different EHEC strains: TUV 93-0 (wildtype), non-motile Δ*adh*E and motile Δ*adh*E escape mutants. The findings from this study suggest that AdhE may have an indirect effect on EHEC motility via its role in the breakdown of acetyl coA. In wild type bacteria, AdhE catalyses the breakdown of acetyl coA to acetaldehyde and ethanol. In this situation bacteria are motile and flagella rotation switching can occur between CW and CCW. Deletion of *adh*E results in non-motile bacteria. It was hypothesised (Beckham et al, 2014) that deletion of *adh*E results in accumulation of acetyl coA which can be inter-converted to acetate via two pathways yielding the highly reactive intermediate substrates, acetyl-phosphate (Acetyl-P) and acetyl adenylate (acetyl-amp). Acetyl-P is a substrate for phosphorylation of CheY Asp57, phosphorylation of which is recognised to activate CheY and induce CW rotation of flagella (Ramakrishnan et al., 1998 and Welch et al., 1993). Similarly acetyl-amp acetylates CheY and significantly increases its CW activity (Barak et al., 1992). Thus it is proposed that in the Δ*adh*E mutant, an increase in acetyl coA results in increased production of Acetyl-P and acetyl-amp leading to greater probability of CW rotation. The inability of these bacteria to be motile would therefore be due to excessive CW rotation. This is summarised in Figure 4-1.

The Δ*adh*E escape mutants regained motility spontaneously. Whole genomic sequencing of these escape mutants showed that, in several mutants, CheY was truncated and lacked the FliM binding site. We hypothesise that the absence of this binding site rescues the Δ*adh*E phenotype in motility. Two non-mutually-exclusive mechanisms could explain this: Firstly, the FliM binding site is known to contain several lysine residues that are key for acetylation of the CheY protein. As mentioned above, immotility of Δ*adh*E mutants could be due to over production of acetyl-amp that would increase the levels of acetylated CheY, without these lysine residues CheY cannot be acetylated. Secondly, the binding of acetylated/phosphorylated CheY to FliM increases the probability of CW rotation. It has been proposed that in the Δ*adh*E mutant this interaction is increased, whilst in the Δ*adh*E

59

escape mutants, the absence of the FliM binding site in the CheY protein does not allow CheY and FliM to bind, and thus the inhibitory action on on CW rotation would be suppressed. These two hypotheses need to be further tested.
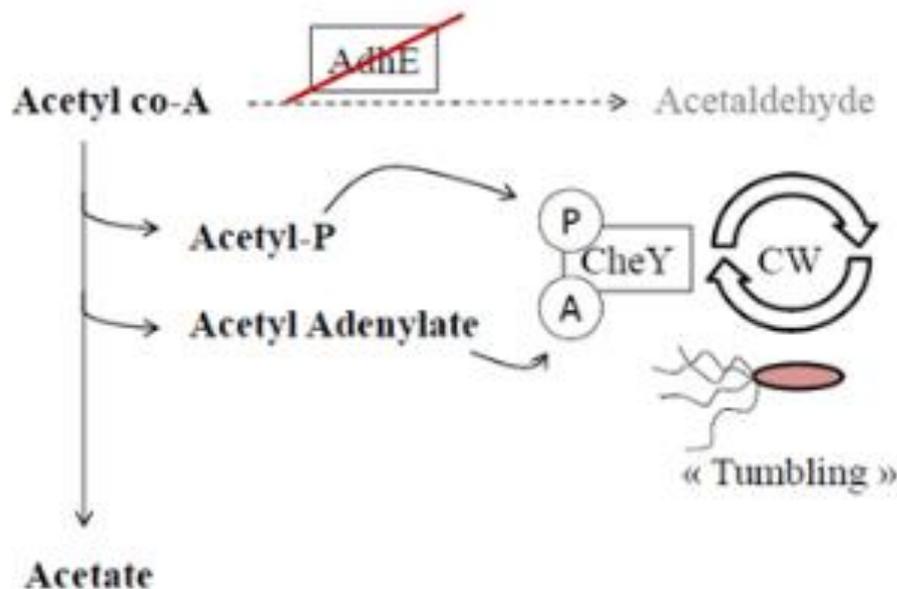
Figure 4-1. A proposed role for *adh*E in regulating EHEC motility. Motility is mediated by flagella rotation allowing bacteria to switch between "running" (CCW rotation of flagella) and "tumbling" (CW rotation of flagella) states. CheY negatively regulates motility through binding to flagella motor, FliM, and increasing probability of CW rotation of flagella (bacteria tumbling) (Bren and Eisenbach, 2000). AdhE is an enzyme that converts acetyl co-A to acetaldehyde, the deletion of *adh*E results in an accumulation of acetyl coA which can be converted to acetate yielding two highly reactive intermediate substrates, acetyl-P and acetyl adenylate. Acetyl-P and acetyl adenylate can phosphorylate or acetylate CheY to increase CheY activity and thus increase probability of CW rotation.

In addition to this, the M5 escape mutant shows deletion of a 7 residue section followed by the insertion of a 23 residue section at position 135 in FliM. It is reasonable to assume that an insertion of this length could perturb the function of the protein. Initial search using Interpro (www.ebi.ac.uk/interpro/) showed this mutation to be located in a Che-C like domain. FliM is composed of three domains: the N-terminal CheY-P binding domain, a central Che-C like domain (residues 43- 240) and a C-terminal FliN binding domain (Park et al, 2006 and Toker et al., 1996). Interestingly, Toker et al., (1996) noted that deletion

mutation in the central domain of FliM, deletion of 10 residue length sections, between residues 101 and 260, resulted in mutants with greatly reduced motility and/or rotation switch bias. M5 is mutated within this region.

The mutations in the escape mutants mentioned above are either in CheY or FliM, two proteins known to be involved in flagellar mediated motility. Mutations in CheY interfere with the FliM binding site, therefore presumably perturbing CheY binding to FliM and thus interfering with motility. The mutation in FliM does not interfere with the CheY binding site but also has an effect on motility, suggesting that this region of FliM could also be involved in motility via a non-CheY binding mediated mechanism. Finally, escape mutant M7 has a single point mutation in FliC. FliC forms the structural component of the flagella and it is possible that a mutation in this protein may perturb flagella function. The sequencing results of this mutant highlight a single amino acid substitution from alanine to valine in FliC, Several publications report that substitution of alanine to valine can alter protein function (Li et al., 2004 and Chen et al., 2005), further examination of this specific point mutation could prove to shed light on the mechanism of flagellar motility regulation.

The use of LOV as a reporter was also evaluated within the context of this thesis and data presented here shows its potential use in monitoring effector protein expression and translocation. The ability to fluorescently label and monitor expression and translocation of effector proteins in real-time is arguably beneficial to researchers. This work has demonstrated that LOV can be used to fluorescently tag the EHEC effector protein, Tir, allowing us to visualise its expression, translocation and host cell localisation in real-time. Central to this conclusion is the observation that the Tir-phiLOV construct was capable of restoring the ability of bacteria to bind and interact with the host cell actin cytoskeleton. Using LOV as a fluorescent reporter opens the opportunity for monitoring of effectors throughout the entire infection process, from expression to localisation in the host cell. Obvious advantages of using such a reporter is the range of information that it gives on the infection process and the limited interference it has with "natural" infection. Furthermore, LOV is oxygen independent and pH stable allowing it considerable potential for use in anaerobic systems, such as anaerobic bacteria as well as in animal models. *In vivo* models could utilise this system by fluorescently tagging proteins with LOV, which could give information on protein expression and localisation within tissues in animal models of infection.

4.2 Considerations and Future work

This thesis sets the foundations for future work on the mechanism of CheY mediated flagella rotation. Yan et al., (2008) highlight several lysine residues, Lys 92, Lys 119 and Lys 122, which are subject to acetylation and involved in FliM binding, and thus induction of CW rotation. Several of the ΔadhE escape mutants, M2, M4 and M6, display a truncated CheY, lacking the FliM binding site and residues Lys 119 and Lys 122. We hypothesise that the removal of acetylation of these residues could decrease CheY binding to FliM and thus decrease probability of CW flagella rotation, thus conferring to more motile bacteria. To examine this further it would serve as useful to introduce truncated CheY in the ΔadhE background and see if the motility phenotype can be rescued. Additionally, co-immunoprecipation of FliM and truncated or non-truncated CheY would shed light into the importance of these regions for the binding of the two proteins. Furthermore it has been reported by Fraiberg et al., 2015 that acetylation of lysine residue 109 stimulates generation of CW rotation and that substitution of this residue results in an inability to support CW rotation. ΔadhE escape mutants with truncated CheY are lacking this lysine residue. This raises the possibility that it is the absence of this lysine that is rescuing motility. To test this, flagellar motility should be assessed in ΔadhE lacking Lysine 109.

Additionally, of the analysed escape mutants, three showed previously published or suggested mutations (Fraiberg et al., 2015 and Yan et al., 2008). This suggests that whole genome sequencing by Ion Torrent Sequencing of motility defective escape mutants is a useful tool to explore the mechanisms that regulate the process of flagellar motility, and that it could be used to find new regulators of this process.

Another aim of this thesis was to evaluate LOV as a novel fluorescent effector tag. Although LOV has proved to show success as a fluorescent reporter it is important to acknowledge that the fluorophore also holds certain limitations. Firstly, fluorescence of LOV is arguably weaker than that of GFP, as seen when comparing fluorescence during TirGFP expression versus Tir-phiLOV expression. Despite this, fluorescence of Tir-phiLOV was readily detected during infection, however, a word of caution; Tir is highly expressed and localises in a well-defined area within the host cell, lending to easy visualisation. It is therefore wise to establish whether LOV fluorescence is sufficient to monitor host cell localisation of more diffusely spread effectors proteins. To this end, the

map-phiLOV plasmid was created and will prove a useful tool in further establishing the potential of LOV as a fluorescent marker. Map has multiple functions in the host cell among which is the targeting and disruption of host cell mitochondria (Kenny and Jepson, 2000 and Ma et al., 2006). Future work aiming to visualise Map-phiLOV host-cell localisation will stand as a good test to further assess LOVs' ability as a reporter molecule.

A further challenge, in regard to LOV fluorescence, is in out-competing background fluorescence in tissue and cell culture samples. Numerous cellular and tissue components auto-fluoresce (reviewed by Monici, 2005) providing an additional challenge when imaging LOV tagged proteins. PhiLOV is an enhanced version of the LOV molecule with increased photo-stability and fluorescence (Christie et al., 2012). The use of LOV as a fluorescent marker is relatively new, therefore perhaps there is opportunity for additional enhancement and further increase in fluorescence.

Finally, results from this study showed that Tir functionality was slightly impaired after conjugation with phiLOV, indicated by partial regain of A/E lesion formation in cells infected with Δ*tir* bacteria transformed with the Tir-phiLOV. PhiLOV is C-terminally conjugated to Tir. Tir inserts into the host cell membrane by forming a hairpin loop structure with intracellular N and C terminus (Luo et al., 2000). Thus, the Tir-phiLOV fusion is likely to interfere partly with Tir structural arrangement in the plasma membrane and slightly diminish A/E lesion formation ability. Insertion of phiLOV into another region of the Tir protein could solve this issue. A strategy that could be undertaken is to look at regions of the protein that are not evolutionary conserved, and to insert the *lov* gene in this location.. However, although non-conserved regions could be seen as regions that are not as important, they may play a role in the final folding/structure and activity of the protein. Additionally, Tir-phiLOV was expressed from the low copy number plasmid, pACYC (copy number 10-12). Therefore expression of Tir may not be comparable to native level, explaining the reduction of A/E lesion formation in the cells infected with the Δ*tir* strain expressing Tir-phiLOV. To test this, a *tir*-STOP (Tir plus stop codon) plasmid was created. Comparison of A/E lesion formation between these strains will serve to determine whether reduction in A/E lesion formation is partly due to reduced Tir expression.

Overall, the work presented in this thesis serves to open a platform for further investigating two essential EHEC virulence factors. Of particular importance is the validation of the novel LOV tools which could be used to explore the regulation and translocation of

effector proteins in EHEC. It would be fascinating to apply the LOV technology to examine the mechanism by which the ΔAdhE mutant results in reduced secretion of effector proteins. The LOV system would help distinguish between transcriptional, translational and post-translational mechanisms of action.

References

Akeda, Y., & Galán, J. E. (2005). Chaperone release and unfolding of substrates in type III secretion. Nature, 437(7060), 911–5.

Barak, R., & Eisenbach, M. (2004). Co-regulation of acetylation and phosphorylation of CheY, a response regulator in chemotaxis of *Escherichia coli*. Journal of Molecular Biology, 342(2), 375–81.

Barak, R., Prasad, K., Shainskaya, A., Wolfe, A. J., & Eisenbach, M. (2004). Acetylation of the chemotaxis response regulator CheY by acetyl-CoA synthetase purified from *Escherichia coli*. Journal of Molecular Biology, 342(2), 383–401.

Barak, R., Welch, M., Yanovsky, A., Oosawa, K., & Eisenbach, M. (1992). Acetyladenylate or Its Derivative Acetylates the Chemotaxis Protein CheY in Vitro and Increases Its Activity at the Flagellar Switch ?, Biochemistry, 10099–10107.

Beckham, K. S. H.,Connolly, J. P. R. Ritchie, J. M. Wang, D., Gawthorne, J., Tahoun, A., Gally, D. L., Burgess, K., Burchmore, R. J., Smith, B. O.,Beatson, S., Byron, O., Wolfe, A. J., Douce, G. R. & Roe,A. J. (2014) The metabolic enzyme AdhE controls the virulence of Escherichia coli O157:H7. Molecular Microbiology 93 (1), 199-211

Berg, H. (2003) The rotary motor of bacterial flagella. Annual Review of Biochemistry 72, 19-54

Blocker, A., Gounon, P., Larquet, E., Niebuhr, K., Cabiaux, V., Parsot, C., & Sansonetti, P. (1999). The tripartite Type II Secretion System of *Shigella flexneri* inserts IpaB and IpaC into Host Membranes, The Journal of Cell Biology 147(3), 683–693.

Bren, a, & Eisenbach, M. (1998). The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. Journal of Molecular Biology, 278(3), 507–14.

Campellone, K. G., Robbins, D., & Leong, J. M. (2004). EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. Developmental Cell, 7(2), 217–28.

Canpolat, N. (2015). Hemolytic uremic syndrome. Turk Pediatri Ars, 50(2), 73-82

Chalfie, M., Verne, J., & Thousand, T. (2009). GFP : Lighting up life. PNAS 106(25), 10073–10080.

Christie, J. M. (2007). Phototropin blue-light receptors. Annual Review of Plant Biology, 58, 21–45.

Christie, J. M., Hitomi, K., Arvai, A. S., Hartfield, K. a, Mettlen, M., Pratt, A. J.& Getzoff, E. D. (2012). Structural tuning of the fluorescent protein iLOV for improved photostability. The Journal of Biological Chemistry, 287(26), 22295–304.

Croxen, M. a, & Finlay, B. B. (2010). Molecular mechanisms *of Escherichia coli* pathogenicity. Nature Reviews. Microbiology, 8(1), 26–38.

DeVinney, R., Puente, J. L., Gauthier, a, Goosney, D., & Finlay, B. B. (2001). Enterohaemorrhagic and enteropathogenic *Escherichia coli* use a different Tir-based mechanism for pedestal formation. Molecular Microbiology, 41(6), 1445–58.

Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauss, U., Guterl, J. K., Wendorff, M., Losi, A., & Gartner (2007). Reporter proteins for in vivo fluorescence without oxygen. Nature Biotechnology, 25, 443-445

Dyer, C. M., Quillin, M. L., Campos, A., Lu, J., McEvoy, M. M., Hausrath, A. C., & Dahlquist, F. W. (2004). Structure of the constitutively active double mutant CheYD13K Y106W alone and in complex with a FliM peptide. Journal of Molecular Biology, 342(4), 1325–35.

Ehsani, S., Rodrigues, C. D., & Enninga, J. (2009). Turning on the spotlight--using light to monitor and characterize bacterial effector secretion and translocation. Current Opinion in Microbiology, 12(1), 24–30.

Eisenbach, M., & Caplan, S. R. (1998). Bacterial chemotaxis: Unsolved mystery of the flagellar switch. Current Biology, 8(13), R444–R446.

Emmerson, J.R., Gally, D.L., and Roe, A.J. (2006) Genera- tion of gene deletions and gene replacements in *Escherichia coli* O157:H7 using a temperature sensitive allelic exchange system. Biological Procedures Online 8: 153–162.

Enninga, J., Sansonetti, P., & Tournebize, R. (2007). Roundtrip explorations of bacterial infection: from single cells to the entire host and back. Trends in Microbiology, 15(11), 483–90.

Fraiberg, M., Afanzar, O., Cassidy, C. K., Gabashvili, A., Schulten, K., Levin, Y., & Eisenbach, M. (2015). CheY's acetylation sites responsible for generating clockwise flagellar rotation in *Escherichia coli*. Molecular Microbiology, 95(2), 231–44.

Freeman, S. A. & Grinstein, S. (2014). Phagocytosis: receptors, signal integration, and the cytoskeleton. Immunology Reviews, 262, 193-215

Garmendia, J., & Frankel, G. (2005). Escherichia coli Infections : Translocation , MINIREVIEW Enteropathogenic and Enterohemorrhagic *Escherichia coli* Infections. Infection and Immunity 73(5), 2573-2585.

Heiman, K., Mody, R. K., Shacara, D. J., Griffin, P. M. & Gould, L. H. (2015). Escherichia coli O157 Outbreaks in the United States 2003-2012. Emerging Infectious Diseases, 21(8),1293-1301

Jacewicz M.S., Mobassaleh M., Gross S. K., Balasubramanian K. A., Daniel P. F., Raghavan S., McCluer R. H., Keusch G. T., (1994). Pathogenesis of Shigella diarrhea: XVII. A mammalian cell membrane glycolipid, Gb3, is required but not sufficient to confer sensitivity to Shiga toxin. Journal of Infectious Diseases, 169(3),538-46.

Janion, C. (2001). Some aspects of the SOS response system--a critical survey. Acta Biochimica Polonica, 48(3), 599–610.

Jarvis, K. G., Girón, J., Jerse, E., McDaniel, T. K., Donnenberg, M. S., & Kaper, J. B. (1995). Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. PNAS, 92(17), 7996–8000.

Kaper, J. B., Nataro, J. P., & Mobley, H. L. (2004). Pathogenic *Escherichia coli*. Nature Reviews. Microbiology, 2(2), 123–40.

Kauppi, A. M., Nordfelth, R., Uvell H., Wolf-Waltz, H. & Elofsson, M. (2003) Targeting bacterial virulence: inhibitors of type III secretion in *Yersinia*. Chem Biol 10, 241-249

Kenny, B., DeVinney, R., Stein, M., Reinscheid, D. J., Frey, E. & Finlay, B. B. (1997). Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell, 91(4), 511–20.

Kimmitt, P. T., Harwood, C. R., & Barer, M. R. (2000). Toxin gene expression by shiga toxin-producing *Escherichia coli*: the role of antibiotics and the bacterial SOS response. Emerging Infectious Diseases, 6(5), 458–65.

Knutton, S., Rosenshine, I., Pallen, M. J., Nisan, I., Neves, B. C., Bain, C., & Frankel, G. (1998). A novel EspA-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. The EMBO Journal, 17(8), 2166–76.

Lizumi, Y., Sagara, H., Kabe, Y., Azuma, M., Kume, K., Ogawa, M., Nagai, T., Gillespie, P. G., Sasakawa, C. & Handa, H. (2007). The Enteropathogenic *E. coli* Effectore EspB facilitates Microvillus Effacing and Antiphagocytosis by Inhibiting Myosin Function. Cell Host & Microbe 2(6), 383-392

Luo, Y., Frey, E. , Pfuetzner, R. , Creagh, L., Knoechel, D. G., Haynes, C. , & Strynadka, N. C. (2000). Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. Nature, 405 (6790), 1073–7.

Mahajan, A., Currie, C. G., Mackie, S., Tree, J., McAteer, S., McKendrick, I. & Smith, D. G. E. (2009). An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157: H7 with bovine intestinal epithelium. Cellular Microbiology, 11(1), 121–137.

Matsumoto, H. & Young, G. M. (2009). Translocated effectors of *Yersinia*. Current opinion in Microbiology, 12(1), 94-100.

McGhie, E. J., Brawn, L. C., Hume, P. J., Humphreys, D. & Koronakis, V. (2009) *Salmonella* takes control: effector-driven manipulation of the host. Current opinion in Microbiology 12 (1), 117-124.

Mele, C., Remuzzi, G. & Noris M (2014). Hemolytic uremic syndrome. Seminars Immunopathology, 36, 399-420.

Melton-Celsa, A. R., (2014). Shiga Toxin (Stx) Classification, Structure and Function. Microbiology Spectrum, 2(2).

Monici, M. (2005). Cell and tissue autofluorescence, research and diagnosis applications. Biotechnology Annual Reviews 11, 227-256.

Nataro, J. P., & Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. Clinical Microbiology Reviews, 11(1), 142–201.

Naylor, S. W., Flockhart, A., Nart, P., Smith, D. G. E., Huntley, J., Gally, D. L., & Low, J. C. (2007). Shedding of *Escherichia coli* O157:H7 in calves is reduced by prior colonization with the homologous strain. Applied and Environmental Microbiology, 73(11), 3765–7.

Naylor, S. W., Low, J. C., Besser, T. E., Mahajan, A., Gunn, G. J., Pearce, M. C., & Mckendrick, I. J. (2003). Lymphoid Follicle-Dense Mucosa at the Terminal Rectum Is the Principal Site of Colonization of Enterohemorrhagic *Escherichia coli* O157 : H7 In the Bovine Host. Infection and Immunity,71(3), 1505-1502.

Ogawa, M., Handa, Y., Ashida, H., Suzuki, M. & Sasakawa, C. (2008). The versatility of *shigella* effectors. Nature Reviews Microbiology 6, 11-16.

Olsen, S. J., Miller, G., Breuer, T., Kennedy, M., Higgins, C., Walford, J., & Mead, P. (2002). A waterborne outbreak of *Escherichia coli* O157:H7 infections and hemolytic uremic syndrome: implications for rural water systems. Emerging Infectious Diseases, 8(4), 370–5.

Orskov, I ., Frits, O., Jann, B. & Klaus, J. (1977). Serology, Chemistry, and Genetics of O and K Antigens of *Escherichia coli*. Bacteriology Reviews, 41(3), 667-710

Peiser, L., Gough, P. J., Kodama, T. & Gordon, S. (2000). Macrophage class A scavenger receptor-mediated phagocytosis of *Escherichia coli*: role of heterogeneity, microbial strain, and culture conditions in vitro. Infection and Immunity, 68, 1953-1963.

Ramakrishnan, R., Schuster, M., & Bourret, R. B. (1998). Acetylation at Lys-92 enhances signaling by the chemotaxis response regulator protein CheY, PNAS 95, 4918–4923.

Rasko, D., & Sperandio, V. (2010). Anti-virulence strategies to combat bacteria-mediated disease. Nature Reviews. Drug Discovery, 9(2), 117–28.

Roe, A. J., Naylor, S. W., Spears, K. J., Yull, H. M., Dransfield, T., Oxford, M. & Gally, D. L. (2004). Co-ordinate single-cell expression of LEE4- and LEE5-encoded proteins of *Escherichia coli* O157:H7. Molecular Microbiology, 54(2), 337–52.

Roe, A., Naylor, S., Spears, K., Spears, H., Yull, H., Dransfield, M., Oxford, M., McKendrick, I., Porter, M., Woodward, M., Smith, D. G. E. & Gally, D.(2002) Coordinate single-cell expression of LEE4- and LEE5-encoded proteins of *Escherichia coli* O157:H7. Molecular Microbiology 54(2), 337–352.

Sansonetti, P., Nhieu, G. T. Van, & Enninga, J. (2005). Secretion of type III effectors into host cells in real time Nature Methods 2(12), 959–965.

Santos, A. S & Finlay, B. B (2015) Bringing down the host: enteropathogenic and enterohaemorrhagic *Escherichia coli* effector- mediated subversion of host immune pathways. Cellular Microbiology, 17 (3), 318-332.

Scheiring, J., Andreoli, S. P. & Zimmerhackl, L. B. (2008). Treatment and outcome of Shiga-toxin-associated haemolytic uremic syndrome (HUS). Pediatric Nephrology 23, 1749-1760

Schüller, S., Heuschkel, R., Torrente, F., Kaper, J. B., & Phillips, A. D. (2007). Shiga toxin binding in normal and inflamed human intestinal mucosa. Microbes and Infection / Institut Pasteur, 9(1), 35–9.

Silversmith, R. E., & Bourret, R. B. (1999). Throwing the switch in bacterial chemotaxis. Trends in Microbiology (17), 16–22.

Su, C., & Brandt, L. J. (1995). Escherichia coli 0157 : H7 Infection in Humans. Ann Intern Med, 123(36), 698–714.

Tarr, P. I., Gordon, C. a, & Chandler, W. L. (2005). Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. Lancet, 365(9464), 1073–86.

Tree, J. J., Wolfson, E. B., Wang, D., Roe, A. J., & Gally, D. L. (2009). Controlling injection: regulation of type III secretion in enterohaemorrhagic Escherichia coli. Trends in Microbiology, 17(8), 361–370.

Tsien, R. Y. (1998). The Green Fluorescent Protein. Annual Review of Biochemistry 67, 509-544.

Van Engelenburg, S. B., & Palmer, A. E. (2008). Quantification of real-time *Salmonella* effector type III secretion kinetics reveals differential secretion rates for SopE2 and SptP. Chemistry & Biology, 15(6), 619–28.

W., & Jaegar, K. (2007) Reporter proteins for in vivo fluorescence without oxygen. Nature Biotechnology 25 (4), 443-445.

Webster K. & Schnitzler, E. (2014). Hemolytic uremic syndrome. Handbook of Clinical Neurology, 120, 1113-1123.

Welch, M., Oosawat, K., & Aizawat, S. (1993). Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria, 90, 8787–8791.

Xu, X., McAteer, S. P., Tree, J. J., Shaw, D. J., Wolfson, E. B. K., Beatson, S. a, & Gally, D. L. (2012). Lysogeny with Shiga toxin 2-encoding bacteriophages represses type III secretion in enterohemorrhagic Escherichia coli. PLoS Pathogens, 8(5), e1002672

Yan, J., Barak, R., Liarzi, O., Shainskaya, A., & Eisenbach, M. (2008). In vivo acetylation of CheY, a response regulator in chemotaxis of *Escherichia coli*. Journal of Molecular Biology, 376(5), 1260–71.

Zhou, D. (1999). Role of the S. typhimurium Actin-Binding Protein SipA in Bacterial Internalization. Science, 283(5410), 2092–2095.

5. Appendix I: Polymerase chain reaction conditions

Table 5-1: Colony PCR reaction: confirmation of Δ*adh*E escape mutants

Reagent	Volume per reaction
Gotaq Mastermix	12.5 μl
Nuclease free water	8.5 μl
25 μM *adh*E F and *adh*E R or Δ*adh*E R primers	1 μl of each
Template	2 μl
Total reaction volume	25 μl

Table 5-2: Colony PCR conditions: confirmation of *adh*E escape mutants

Temperature	Time	
98 °C	2 minutes	
98 °C	45 seconds	
55 °C	45 seconds	X 35
72 °C	1 min 30 secs	
72 °C	10 minutes	

Table 5-3: PCR reaction: map insert amplification

Reagent	Volume per reaction
10mM dNTPs	1 μl
10X PFU Buffer	5 μl
PFU Turbo Taq	0.5 μl
Nuclease free water	40.5 μl
25 μM Map 5' and Map 3' primers	1 μl of each
Template	1 μl
Total reaction volume	50 μl

Table 5-4: PCR conditions: map insert amplification

Temperature	Time	
95°C	2 minutes	
95 °C	30 seconds	
50 °C	30 seconds	X 30
72 °C	1 minute	
72 °C	7 minutes	

Table 5-5: KOD colony PCR reaction: confirm map-phiLOV plasmid

Reagent	Volume per reaction
10mM dNTPs	5 μl
KOD Buffer	2 μl
KOD taq	0.5 μl
Nuclease free water	8.5 μl
MgCl2	1 μl
25 μM Map 5' and MapLOV 3' primers	0.5 μl of each
Template	2 μl
Total reaction volume	20 μl

Table 5-6: KOD colony PCR conditions: confirm map-phiLOV plasmid

Temperature	Time	
98°C	1 minute	
98 °C	15 seconds	
52 °C	4 seconds	X 25
72 °C	20 seconds	
72 °C	7 minutes	

Table 5-7: gDNA library amplification PCR reaction mix

Reagent	Volume per reaction
Platinum PCR SuperMix High fidelity	200 μl
Library Amplification Primer Mix	10 μl
Unamplified library	50 μl
Total reaction volume	**350 μl**

6. Appendix II Supplementary information on Ion Torrent Sequencing results

6.1 Chip loading: ISP Loading percentage and ISP Density

Chip loading quality is indicated by ISP loading percentage, the percentage of wells in the chip that contain either a template or non-templated ISP, and ISP density. High ISP loading is desirable and both chips show loading of greater than 80% (83 % Chip A, 86 % Chip B) (Figure 6-1). The ISP density is displayed as an image showing the percentage loading across the surface of the chip. Red and orange indicate areas of good loading and blue and green indicate areas were loading is very low. Both chips were seen to have good loading density (Figure 6-1). Additional filtering separates chip wells into different classes (Table 6-1.)

A.   **B.**
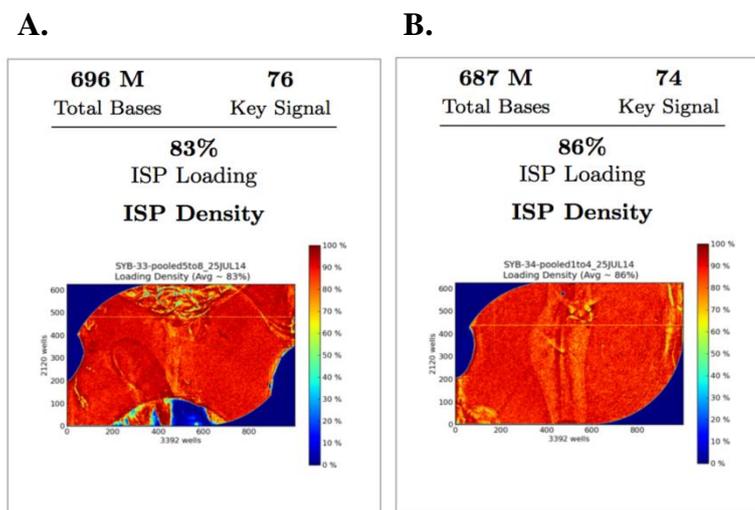
Figure 6-1 Ion Torrent 316 chip Loading. The percentage of wells loaded with sample for each chip is indicated by the ISP loading, and can be seen to be high for both chips. Chip A bears pooled library from strains M1 –M4, and has a loading percentage of 83 % (A) and chip B, bears libraries from strains M5 – M6 and ΔAdhE, and has a loading percentage of 86 % (B). The ISP loading density is shown for each chip, red/orange indicates good loading, yellow indicates passable and green and blue shows areas on the chip where loading is very low (normally due to presence of air bubbles) .

6.2 Read filtering and trimming

After establishing chip loading quality the identified ISPs bearing library are filtered and trimmed. After trimming any reads shorter than 4 bases are filtered into "Low Quality" category. Additional filters remove polyclonals and primer:dimers, ensuring a high quality library. The percentage of polyclonals, primer:dimers and low quality reads for each chip are displayed in Table 6-2.

Table 6-1: Characterisation of Chip wells

Parameter	Chip A (%)	Chip B (%)	Description
Percentage of addressable wells:			
With ISPs	85.8	83.5	Percentage of wells that are ISP positive ie. An ISP is present in the well. A higher percentage of ISP positive wells indicates better chip loading quality.
Live	99.9	99.9	Number and percentage of wells containing an ISP with sufficient signal to be associated with the Library or Test Fragment key.
Test Fragment	00.7	00.4	Number and percentage of Live ISPs with a key signal that is identical to the Test fragment key signal.
Library	99.3	99.6	Number and percentage of Live ISPs with an identical key signal to the library key signal. This gives an indication of the quality of the reads.

Table 6-2 ISP library quality. ISP library filters excludes ISPs bearing Polyclonal, low quality or primer:dimer library, as specified by the criteria below.

Parameter	Chip A (%)	Chip B (%)	Description
Percentage of library ISPs that are:			
Filtered: Polyclonal	12.2	21.1	ISPs carrying clones from two or more templates. Filters out reads with > 1 library population.
Filtered: Low Quality	20.1	16.3	Low or un recognisable signal, or reads of < 4 bp
Filtered: Primer dimer	00.0	00.0	Filters out reads with an insert length of < 8 bp
Final Library ISPs	**67.7**	**62.6**	Number and percentage of reads that pass above filters.

6.3 Sequence Alignment

During sequencing, genomes were aligned to reference strain *Escherichia coli* O157:H7 strain EDL933 (NCBI). An assessment of alignment was performed automatically during sequencing and is shown in Figure 6-2. Furthermore, alignment to reference genome was assessed by determining the Aligned Quality at a phred score of 20 (AQ20) (Table 6-3), which gives information on the accuracy of the base calls within a read. Libraries from both chips were aligned successfully to the reference strain genome.

A.   **B.**
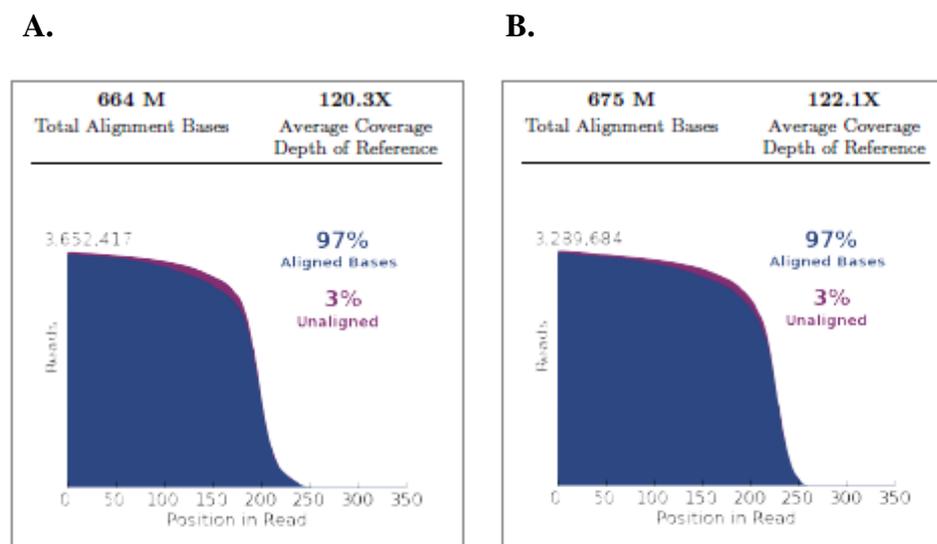
Figure 6-2. Sequencing Read alignment to Reference strain. The alignment of reads to the reference strain gives an indication of the quality of the sequencing data and can be seen to be high for both chips. Chip A, pooled libraries M1 – M4, shows an alignment value of 97 % and a coverage of 120.3 X. Chip B, pooled libraries M5, M6 and ΔAdhE shows an alignment of 97 % with coverage of 122.1 X.

Table 6-3: Summary of DNA sequence quality. During sequencing, reads were aligned to the reference genome, *Escherichia coli* O157:H7 strain EDL933 and the quality of alignment is recorded as the AQ20. The AQ20 is The Aligned Quality (AQ) at a Phred Score of 20. In other words, the number of bases with a base call accuracy of 99%.

	AQ20	
	Chip A	**Chip B**
Total Number of Bases (Mbp)	600	617
Mean Length (bp)	176	201
Longest Alignment (bp)	285	277
Mean Coverage Depth	108.7	111.7

7. Appendix III Sequence Alignments

```
        1       10        20        30        40        50        60        70        80        90       100       110       120      130
        |-------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_CheY ATGGCGGATAAAGAACTTAAATTTTTGGTTGTGGATGACTTTTCCACCATGCGACGCATAGTGCGTAACCTGCTGAAAGAGCTGGGATTCAATAATGTTGAGGAAGCGGAAGATGGCGTCGACGCTCTCA
M2_CheY   ATGGCGGATAAAGAACTTAAATTTTTGGTTGTGGATGACTTTTCCACCATGCGACGCATAGTGCGTAACCTGCTGAAAGAGCTGGGATTCAATAATGTTGAGGAAGCGGAAGATGGCGTCGACGCTCTCA
M4_CheY   ATGGCGGATAAAGAACTTAAATTTTTGGTTGTGGATGACTTTTCCACCATGCGACGCATAGTGCGTAACCTGCTGAAAGAGCTGGGATTCAATAATGTTGAGGAAGCGGAAGATGGCGTCGACGCTCTCA
M6_CheY   ATGGCGGATAAAGAACTTAAATTTTTGGTTGTGGATGACTTTTCCACCATGCGACGCATAGTGCGTAACCTGCTGAAAGAGCTGGGATTCAATAATGTTGAGGAAGCGGAAGATGGCGTCGACGCTCTCA
Consensus ATGGCGGATAAAGAACTTAAATTTTTGGTTGTGGATGACTTTTCCACCATGCGACGCATAGTGCGTAACCTGCTGAAAGAGCTGGGATTCAATAATGTTGAGGAAGCGGAAGATGGCGTCGACGCTCTCA

        131      140       150       160       170       180       190       200       210       220       230       240       250      260
        |-------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_CheY ATAAGTTGCAGGCAGGCGGTTATGGATTTGTTATCTCCGACTGGAACATGCCCAACATGGATGGCCTGGAATTGCTGAAAACGATTCGTGCGGATGGCGCGATGTCGGCATTGCCAGTGTTAATGGTGAC
M2_CheY   ATAAGTTGCAGGCAGGCGGTTATGGATTTGTTATCTCCGACTGGAACATGCCCAACATGGATGGCCTGGAATTGCTGAAAACGATTCGTGCGGATGGCGCGATGTCGGCATTGCCAGTGTTAATGGTGAC
M4_CheY   ATAAGTTGCAGGCAGGCGGTTATGGATTTGTTATCTCCGACTGGAACATGCCCAACATGGATGGCCTGGAATTGCTGAAAACGATTCGTGCGGATGGCGCGATGTCGGCATTGCCAGTGTTAATGGTGAC
M6_CheY   ATAAGTTGCAGGCAGGCGGTTATGGATTTGTTATCTCCGACTGGAACATGCCCAACATGGATGGCCTGGAATTGCTGAAAACGATTCGTGCGGATGGCGCGATGTCGGCATTGCCAGTGTTAATGGTGAC
Consensus ATAAGTTGCAGGCAGGCGGTTATGGATTTGTTATCTCCGACTGGAACATGCCCAACATGGATGGCCTGGAATTGCTGAAAACGATTCGTGCGGATGGCGCGATGTCGGCATTGCCAGTGTTAATGGTGAC

        261      270       280       290       300       310       320       330       340       350       360       370       380      390
        |-------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_CheY TGCAGAAGCGAAGAAAGAGAACATCATTGCTGCGGCGCAAGCGGGGGCCAGTGGCTATGTGGTGAAGCCATTTACCGCCGCGACGCTGGAGGAAAAACTCAACAAAATCTTTGAGAAACTGGGCATGTGA
M2_CheY   TGCAGAAGCGAAGAAAGAGAACATCAT-GCTGCGGCGCAAGCGGGGGCCAGTGGCTATGTGGTGA
M4_CheY   TGCAGAAGCGAAGAAAGAGAACATCAT-GCTGCGGCGCAAGCGGGGGCCAGTGGCTATGTGGTGA
M6_CheY   TGCAGAAGCGAAGAAAGAGAACATCAT-GCTGCGGCGCAAGCGGGGGCCAGTGGCTATGTGGTGA
Consensus TGCAGAAGCGAAGAAAGAGAACATCAT.GCTGCGGCGCAAGCGGGGGCCAGTGGCTATGTGGTGA.......................................................
```

Figure 7-1 Sequence alignment of ΔAdhE, m2, m4, m6 for CheY

The figure above shows the amino acid sequence alignment between strains ΔAdhE, m2, m4 and m6 for the gene CheY. Strains m2, m4 and m6 show deletion of 65 amino acids at the C-terminal end of CheY resulting in a truncated protein. This figure was generated using MultAlin (University Toulouse), with sequence input settings on "Auto" from FASTA file of the CheY sequence as determined by Ion torrent sequencing (Materials and Methods section 2.4).

```
                  1        10        20        30        40        50        60        70        80        90       100       110       120       130
                  |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  ATGGGCGATAGTATTCTTTCTCAAGCTGAAATTGATGCGCTGTTGAATGGTGACAGCGAAGTCAAAGACGAACCGACAGCCAGTGTTAGCGGCGAAAGTGACATTCGTCCGTACGATCCGAATACCCAAC
 M5_FliM   ATGGGCGATAGTATTCTTTCTCAAGCTGAAATTGATGCGCTGTTGAATGGTGACAGCGAAGTCAAAGACGAACCGACAGCCAGTGTTAGCGGCGAAAGTGACATTCGTCCGTACGATCCGAATACCCAAC
Consensus  ATGGGCGATAGTATTCTTTCTCAAGCTGAAATTGATGCGCTGTTGAATGGTGACAGCGAAGTCAAAGACGAACCGACAGCCAGTGTTAGCGGCGAAAGTGACATTCGTCCGTACGATCCGAATACCCAAC

                131       140       150       160       170       180       190       200       210       220       230       240       250       260
                 |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  GACGGGTTGTGCGCGAACGTTTGCAGGCGCTGGAAATCATTAATGAGCGCTTTGCCCGCCATTTTCGTATGGGGCTGTTCAACCTGCTGCGTCGTAGCCCGGATATAACCGTCGGGGCCATCCGCATTCA
 M5_FliM   GACGGGTTGTGCGCGAACGTTTGCAGGCGCTGGAAATCATTAATGAGCGCTTTGCCCGCCATTTTCGTATGGGGCTGTTCAACCTGCTGCGTCGTAGCCCGGATATAACCGTCGGGGCCATCCGCATTCA
Consensus  GACGGGTTGTGCGCGAACGTTTGCAGGCGCTGGAAATCATTAATGAGCGCTTTGCCCGCCATTTTCGTATGGGGCTGTTCAACCTGCTGCGTCGTAGCCCGGATATAACCGTCGGGGCCATCCGCATTCA

                261       270       280       290       300       310       320       330       340       350       360       370       380       390
                 |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  GCCGTACCATGAATTTGCCCGCAACCTGCCGGTGCCGACCAACCTGAACCTTATCCATCTAAAACCGCTGCGCGGCACTGGGCTGGTGGTGTTCTCACCGAGTCTGGTGTTTATCGCCGTGGATAACCTG
 M5_FliM   GCCGTACCATGAATTTGCCCGCAACCTGCCGGTGCCGACCAACCTGAACCTTATCCATCTAAAACCGCTGCGCGGCACTGGGCTGGTGGTGTTCTCACCGAGTCTGGTGTTTATCGCCGTGGATAACCTG
Consensus  GCCGTACCATGAATTTGCCCGCAACCTGCCGGTGCCGACCAACCTGAACCTTATCCATCTAAAACCGCTGCGCGGCACTGGGCTGGTGGTGTTCTCACCGAGTCTGGTGTTTATCGCCGTGGATAACCTG

                391       400       410       420       430       440       450       460       470       480       490       500       510       520
                 |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  TTTGGCGGCGATGGACGCTTCCCGACCAAAGTGGAAGGTCGCGAGTTTACCCATACCGAACAGCGCGTCATCAACCGCATGTTGA------------------------------------------------
 M5_FliM   TTTGGCGGCGATGGA-GCTTCCCGACCAAAGTGGAAGGTCGCGAGTTTACCCATACCGAACAGCGCGTCATCAACCGCATGTTGAGTGGAAGGTCGCGAGTTTACCCATACCGAACAGCGCGTCATCAAC
Consensus  TTTGGCGGCGATGGA.GCTTCCCGACCAAAGTGGAAGGTCGCGAGTTTACCCATACCGAACAGCGCGTCATCAACCGCATGTTGA...........................................

                521       530       540       550       560       570       580       590       600       610       620       630       640       650
                 |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  -----------AACTGGCGCTTGAAGGCTATAGCGACGCATGGAAGGCGATTAATCCGCTGGAAGTTGAGTACGTGCGTTCGGAAATGCAGGTGAAATTTACCAATATCACCACCTCGCCGAACGACATTG
 M5_FliM   CGCATGTTGAAACTGGCGCTTGAAGGCTATAGCGACGCATGGAAGGCGATTAATCCGCTGGAAGTTGAGTACGTGCGTTCGGAAATGCAGGTGAAATTTACCAATATCACCACCTCGCCGAACGACATTG
Consensus  ..........AACTGGCGCTTGAAGGCTATAGCGACGCATGGAAGGCGATTAATCCGCTGGAAGTTGAGTACGTGCGTTCGGAAATGCAGGTGAAATTTACCAATATCACCACCTCGCCGAACGACATTG
```

75

```
      651       660        670        680        690        700        710        720        730        740        750        760        770        780
      |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  TGGTTAACACGCCGTTCCATGTGGAGATTGGCAACCTGACCGGCGAATTTAATATCTGCCTGCCATTCAGCATGATCGAGCCGCTACGGGAATTGTTGGTTAACCCGCCGCTGGAAAACTCGCGTAATGA
  M5_FliM  TGGTTAACACGCCGTTCCATGTGGAGATTGGCAACCTGACCGGCGAATTTAATATCTGCCTGCCATTCAGCATGATCGAGCCGCTACGGGAATTGTTGGTTAACCCGCCGCTGGAAAACTCGCGTAATGA
Consensus  TGGTTAACACGCCGTTCCATGTGGAGATTGGCAACCTGACCGGCGAATTTAATATCTGCCTGCCATTCAGCATGATCGAGCCGCTACGGGAATTGTTGGTTAACCCGCCGCTGGAAAACTCGCGTAATGA

      781       790        800        810        820        830        840        850        860        870        880        890        900        910
      |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  AGATCAGAACTGGCGCGATAACCTGGTGCGCCAGGTGCAGCATTCACAGCTGGAGCTGGTCGCCAACTTTGCCGATATCTCGCTACGCCTGTCGCAGATTTTAAAACTGAAACCCGGTGACGTCCTGCCG
  M5_FliM  AGATCAGAACTGGCGCGATAACCTGGTGCGCCAGGTGCAGCATTCACAGCTGGAGCTGGTCGCCAACTTTGCCGATATCTCGCTACGCCTGTCGCAGATTTTAAAACTGAAACCCGGTGACGTCCTGCCG
Consensus  AGATCAGAACTGGCGCGATAACCTGGTGCGCCAGGTGCAGCATTCACAGCTGGAGCTGGTCGCCAACTTTGCCGATATCTCGCTACGCCTGTCGCAGATTTTAAAACTGAAACCCGGTGACGTCCTGCCG

      911       920        930        940        950        960        970        980        990       1000       1010       1020       1030      1040
      |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliM  ATAGAAAAACCCGACCGCATCATCGCCCATGTTGACGGCGTCCCGGTGCTGACCAGTCAGTACGGCACTCTCAACGGTCAGTATGCGTTACGGATAGAACATTTGATTAACCCGATTTTAAATTCTCTGA
  M5_FliM  ATAGAAAAACCCGACCGCATCATCGCCCATGTTGACGGCGTCCCGGTGCTGACCAGTCAGTACGGCACTCTCAACGGTCAGTATGCGTTACGGATAGAACATTTGATTAACCCGATTTTAAATTCTCTGA
Consensus  ATAGAAAAACCCGACCGCATCATCGCCCATGTTGACGGCGTCCCGGTGCTGACCAGTCAGTACGGCACTCTCAACGGTCAGTATGCGTTACGGATAGAACATTTGATTAACCCGATTTTAAATTCTCTGA

     1041   1050        1060
      |---------+---------|
AdhE_FliM  ACGAGGAACAGCCCAAATGA
  M5_FliM  ACGAGGAACAGCCCAAATGA
Consensus  ACGAGGAACAGCCCAAATGA
```

Figure 7-2 Sequence alignment of ΔAdhE and m5 for FliM

The above figure shows the amino acid sequence alignment of ΔAdhE and m5 for the gene FliM. M5 has an insertion of 23 amino acids. This figure was generated using MultAlin (University Toulouse), with sequence input settings on "Auto", from FASTA file of the FliM sequence as determined by Ion torrent sequencing (Materials and Methods section 2.4).

```
              1        10        20        30        40        50        60        70        80        90       100       110       120       130
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC     ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATG
M7_FliC       ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATG
Consensus     ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATG

             131       140       150       160       170       180       190       200       210       220       230       240       250       260
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC     ACGCCGTAGGTCAGGCGATTGCTAACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATTTCTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAA
M7_FliC       ACGCCGTAGGTCAGGCGATTGCTAACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATTTCTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAA
Consensus     ACGCCGTAGGTCAGGCGATTGCTAACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATTTCTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAA

             261       270       280       290       300       310       320       330       340       350       360       370       380       390
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC     CAACTTACAGCGTATTCGTGAACTGACGGTTCAGGCCACTACAGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCAAATCTCGTCTTGATGAAATTGACCGCGTATCCGGCCAGACC
M7_FliC       CAACTTACAGCGTATTCGTGAACTGACGGTTCAGGCCACTACAGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCAAATCTCGTCTTGATGAAATTGACCGCGTATCCGGCCAGACC
Consensus     CAACTTACAGCGTATTCGTGAACTGACGGTTCAGGCCACTACAGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCAAATCTCGTCTTGATGAAATTGACCGCGTATCCGGCCAGACC

             391       400       410       420       430       440       450       460       470       480       490       500       510       520
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC     CAGTTCAACGGCGTGAACGTGCTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGCGAATGACGGCGAAACCATCACGATCGACCTGAAAAAAATCGATTCTGATACTCTGGGTCTGAATGGCTTTA
M7_FliC       CAGTTCAACGGCGTGAACGTGCTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGCGAATGACGGCGAAACCACCACGATCGACCTGAAAAAAATCGATTCTGATACTCTGGGTCTGAATGGCTTTA
Consensus     CAGTTCAACGGCGTGAACGTGCTGGCGAAAGACGGTTCAATGAAAATTCAGGTTGGTGCGAATGACGGCGAAACCAcCACGATCGACCTGAAAAAAATCGATTCTGATACTCTGGGTCTGAATGGCTTTA

             521       530       540       550       560       570       580       590       600       610       620       630       640       650
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC     ACGTAAATGGTAAAGGTACTATTACCAACAAAGCTGCAACGGTAAGTGATTTAACTTCTGCTGGCGCGAAGTTAAACACCACGACAGGTCTTTATGATCTGAAAACCGAAAATACCTTGTTAACTACCGA
M7_FliC       ACGTAAATGGTAAAGGTACTATTACCAACAAAGCTGCAACGGTAAGTGATTTAACTTCTGCTGGCGCGAAGTTAAACACCACGACAGGTCTTTATGATCTGAAAACCGAAAATACCTTGTTAACTACCGA
Consensus     ACGTAAATGGTAAAGGTACTATTACCAACAAAGCTGCAACGGTAAGTGATTTAACTTCTGCTGGCGCGAAGTTAAACACCACGACAGGTCTTTATGATCTGAAAACCGAAAATACCTTGTTAACTACCGA

             651       660       670       680       690       700       710       720       730       740       750       760       770       780
              |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC     TGCTGCATTCGATAAATTAGGGAATGGCGATAAAGTCACAGTTGGCGGCGTAGATTATACTTACAACGCTAAATCTGGTGATTTTACTACCACTAAATCTACTGCTGGTACGGGTGTAGACGCCGCGGCG
M7_FliC       TGCTGCATTCGATAAATTAGGGAATGGCGATAAAGTCACAGTTGGCGGCGTAGATTATACTTACAACGCTAAATCTGGTGATTTTACTACCACTAAATCTACTGCTGGTACGGGTGTAGACGCCGCGGCG
Consensus     TGCTGCATTCGATAAATTAGGGAATGGCGATAAAGTCACAGTTGGCGGCGTAGATTATACTTACAACGCTAAATCTGGTGATTTTACTACCACTAAATCTACTGCTGGTACGGGTGTAGACGCCGCGGCG
```

```
        781      790        800        810        820        830        840        850        860        870        880        890        900        910
        |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC   CAGGCTGCTGATTCAGCTTCAAAACGTGATGCGTTAGCTGCCACCCTTCATGCTGATGTGGGTAAATCTGTTAATGGTTCTTACACCACAAAAGATGGTACTGTTTCTTTCGAAACGGATTCAGCAGGTA
M7_FliC     CAGGCTGCTGATTCAGCTTCAAAACGTGATGCGTTAGCTGCCACCCTTCATGCTGATGTGGGTAAATCTGTTAATGGTTCTTACACCACAAAAGATGGTACTGTTTCTTTCGAAACGGATTCAGCAGGTA
Consensus   CAGGCTGCTGATTCAGCTTCAAAACGTGATGCGTTAGCTGCCACCCTTCATGCTGATGTGGGTAAATCTGTTAATGGTTCTTACACCACAAAAGATGGTACTGTTTCTTTCGAAACGGATTCAGCAGGTA

        911      920        930        940        950        960        970        980        990       1000       1010       1020       1030       1040
        |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC   ATATCACCATCGGTGGAAGCCAGGCATACGTAGACGATGCAGGCAACTTGACGACTAACAACGCTGGTAGCGCAGCTAAAGCTGATATGAAAGCGCTGCTCAAAGCAGCGAGCGAAGGTAGTGACGGTGC
M7_FliC     ATATCACCATCGGTGGAAGCCAGGCATACGTAGACGATGCAGGCAACTTGACGACTAACAACGCTGGTAGCGCAGCTAAAGCTGATATGAAAGCGCTGCTCAAAGCAGCGAGCGAAGGTAGTGACGGTGC
Consensus   ATATCACCATCGGTGGAAGCCAGGCATACGTAGACGATGCAGGCAACTTGACGACTAACAACGCTGGTAGCGCAGCTAAAGCTGATATGAAAGCGCTGCTCAAAGCAGCGAGCGAAGGTAGTGACGGTGC

       1041     1050       1060       1070       1080       1090       1100       1110       1120       1130       1140       1150       1160       1170
        |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC   CTCTCTGACATTCAATGGCACAGAATATACCATCGCAAAAGCAACTCCTGCGACAACCACTCCAGTAGCTCCGTTAATCCCTGGTGGGATTACTTATCAGGCTACAGTGAGTAAAGATGTAGTATTGAGC
M7_FliC     CTCTCTGACATTCAATGGCACAGAATATACCATCGCAAAAGCAACTCCTGCGACAACCACTCCAGTAGCTCCGTTAATCCCTGGTGGGATTACTTATCAGGCTACAGTGAGTAAAGATGTAGTATTGAGC
Consensus   CTCTCTGACATTCAATGGCACAGAATATACCATCGCAAAAGCAACTCCTGCGACAACCACTCCAGTAGCTCCGTTAATCCCTGGTGGGATTACTTATCAGGCTACAGTGAGTAAAGATGTAGTATTGAGC

       1171     1180       1190       1200       1210       1220       1230       1240       1250       1260       1270       1280       1290       1300
        |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC   GAAACCAAAGCGGCTGCCGCGACATCTTCAATTACCTTTAATTCCGGTGTACTGAGCAAAACTATTGGGTTTACCGCGGGTGAATCCAGTGATGCTGCGAAGTCTTATGTGGATGATAAAGGTGGTATCA
M7_FliC     GAAACCAAAGCGGCTGCCGCGACATCTTCAATTACCTTTAATTCCGGTGTACTGAGCAAAACTATTGGGTTTACCGCGGGTGAATCCAGTGATGCTGCGAAGTCTTATGTGGATGATAAAGGTGGTATCA
Consensus   GAAACCAAAGCGGCTGCCGCGACATCTTCAATTACCTTTAATTCCGGTGTACTGAGCAAAACTATTGGGTTTACCGCGGGTGAATCCAGTGATGCTGCGAAGTCTTATGTGGATGATAAAGGTGGTATCA

       1301     1310       1320       1330       1340       1350       1360       1370       1380       1390       1400       1410       1420       1430
        |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC   CTAACGTTGCCGACTATACAGTCTCTTACAGCGTTAACAAGGATAACGGCTCTGTGACTGTTGCCGGGTATGCTTCAGCGACTGATACCAATAAAGATTATGCTCCAGCAATTGGTACTGCTGTAAATGT
M7_FliC     CTAACGTTGCCGACTATACAGTCTCTTACAGCGTTAACAAGGATAACGGCTCTGTGACTGTTGCCGGGTATGCTTCAGCGACTGATACCAATAAAGATTATGCTCCAGCAATTGGTACTGCTGTAAATGT
Consensus   CTAACGTTGCCGACTATACAGTCTCTTACAGCGTTAACAAGGATAACGGCTCTGTGACTGTTGCCGGGTATGCTTCAGCGACTGATACCAATAAAGATTATGCTCCAGCAATTGGTACTGCTGTAAATGT

       1431     1440       1450       1460       1470       1480       1490       1500       1510       1520       1530       1540       1550       1560
        |---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
AdhE_FliC   GAACTCCGCGGGTAAAATCACTACTGAGACTACCAGTGCTGGTTCTGCAACGACCAACCCGCTTGCTGCCCTGGACGACGCAATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAAC
M7_FliC     GAACTCCGCGGGTAAAATCACTACTGAGACTACCAGTGCTGGTTCTGCAACGACCAACCCGCTTGCTGCCCTGGACGACGCAATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAAC
Consensus   GAACTCCGCGGGTAAAATCACTACTGAGACTACCAGTGCTGGTTCTGCAACGACCAACCCGCTTGCTGCCCTGGACGACGCAATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAAC
```

Figure 7-3 Sequence alignment of ΔAdhE and m7 for FliC

Amino acid sequence alignment of ΔAdhE and m7 for FliC. M7shows an Single nucleotide polymorphism (SNP) at position 467, specifically base pair exchange from Thymine (T) to Cytosine (C). This figure was generated using MultAlin (University Toulouse), with sequence input settings on "Auto", from FASTA file of the FliC sequence as determined by Ion torrent sequencing (Materials and Methods section 2.4).