Summary

A persistent goal of industrial biotechnology is to produce recombinant protein at reduced costs, while maintaining high product quality and yield. A high proportion of production costs are incurred due to downstream processing. One route to reduce these processing costs is to achieve secretion of proteins by the expression organism, an objective that has also been shown to increase product yield and quality. One route to achieve this is to re-engineer the bacterial flagella type III secretion system (FT3SS). This organelle provides motility to the cell through motor and filament structures -however in the context of biotechnology, the FT3SS provides a high capacity one step route for protein export from the cytoplasm to the extracellular space –exporting approximately 1000 protein subunits per minute- a rate which rivals other extracellular secretion systems in a range of biotechnologically relevant organisms.

This project aimed to optimise an *E. coli* strain for directed secretion of proteins from the cytoplasm directly to the extracellular environment, through a modified FT3SS. The flagella secretion system served as a platform technology for the secretion of a range of proteins. Initially native protein secretion through the modified secretion apparatus was investigated, using a flagellin monomer variant. Once established, directed recombinant protein secretion of an antibody fragment (CH2) through the modified FT3SS secretion apparatus was achieved with the aid of a secretion construct, which harboured recombinant cargo protein and was also comprised of elements of the native flagellin to aid secretion (the untranslated regions and secretion signal) and tags for purification. The secretion construct is modular and therefore amenable to alteration of the cargo protein –and also modification of the secretion construct for improved secretion. Strain optimisation for improved secretion was also implemented. With the generation of a wealth of strain improvements and secretion construct variants, a high-throughput assay was required to quantify secretion output of combinations of these secretion tools. An enzyme based secretion assay was developed to screen the multitude of strains and secretion plasmids in an efficient and accurate manner. This resulted in identification of a secretion strain and construct which resulted in high capacity secretion. When utilised together a 25 fold increase in the secretion capacity was observed. Finally expression of a 45kDa human collagen I α chain was investigated as extrusion of collagen from the FT3SS was an attractive possibility, while secretion was not achieved, strain improvements resulted in an increase in intracellular concentration of this large therapeutically relevant protein.

The yield of secreted protein was measured throughout and it is possible that with further improvements to strains and fermentation procedures that the FT3SS platform could be competitive in the field of industrial biotechnology

Acknowledgements

Firstly I would like to extend my thanks to The University of Sheffield for providing me with the opportunity and funding to carry out this PhD.

I would like to thank my supervisor Dr. Graham Stafford for his invaluable guidance and advice throughout my project. He has also provided me with the support, motivation and confidence to develop into the scientist I am today -your love of science is truly inspiring! I would also like to thank my secondary supervisor Prof. Phillip Wright, for his support and 'helicopter vision' throughout the project.

To Jason Heath and Brenka McCabe, the labs would be shambolic without you and I can't thank you enough for the continuous effort you both go to, to make things run smoothly. On top of this, the wealth of knowledge you both have had been a great help. You both also made the labs a genuinely great place to work -I've really valued your company over the past few years.

To my colleagues over the years -in the Stafford lab group, the Dental School and Chemical and Biological Engineering and the Synthetic Biology Network- on an academic level, thank you for your input and support, but I also greatly value the friendships I have made. A special mention goes to Andy Frey and Genevieve Melling –we've shared some excellent times and I'm certain I've gained some lifelong friends.

Kate and Liz. You are the most incredible beings and also excellent scientists. We've shared some amazing times over the past few years. You both bring out the best in me and it's been a privilege to go through this PhD experience together. I couldn't have done this without you, whether it's been rallying me along or celebrating my achievements, you've both always been there.

To Jack and Liz my eternal gratitude goes out to you for putting a roof over my head (and an airbed under my head!) for the past 9 months. I am incredibly grateful for you generosity and excellent company - it was also an honour to be so welcome in your house.

To my fellow Tamworth Tossers, your unique mix of love and support, with a good dose of mockery has made my PhD life a whole lot easier. Thank you!

Finally to Mum and Dad and Jackson -the love and support you give me has given me the confidence to excel throughout my academic career. You're all great inspirations to me and I'm glad I can do the Green family proud! All my love.

Table of contents

List of Figures

List of Tables

1. Introduction

Recombinant therapeutic protein production is a common application of industrial biotechnology. Recombinant therapeutics are of high economic value, however a persistent aim of the biotechnology industry is to increase product yields and reduce production costs. A high proportion of production costs are incurred due to downstream processing of recombinant proteins, i.e. separation from undesirable cellular contaminants (Lowe et al., 2001). One route to reduce these processing costs is to achieve secretion of proteins by the expression organism, an objective that has also been shown to increase product yield and quality (Georgiou & Segatori, 2005; Ni & Chen, 2009).

One possible solution is the high throughput one-step directed secretion of proteins from bacteria from the cell cytoplasm into the surrounding media from where it can then be harvested in a simplified manner. This also opens up the possibility of stable cell cultures that secrete protein continuously without the need for sacrifice of the seed culture during the process. Secretion into the media by Gram-negative bacteria is a major challenge in biotechnology and while it would be beneficial it is low yielding.

One route to achieving such an aim would be the re-engineering of an existing biological protein secretion system to reach this end. This thesis will outline work aimed at this end focussing on re-engineering of the common biotechnology employed bacterium, *Escherichia coli*, to directly secrete proteins in to growth medium using the flagella type III secretion system (FT3SS). This aim was addressed using a combination of traditional biological modification methods and emerging methodology associated with synthetic biology and certainly considers the ethos of directed engineering with a defined goal in mind. As a grand challenge for this secretion system and as an aim in itself, I set the objective of collagen production and secretion as an exemplar for this system, because it is an important therapeutic protein, there are challenges in its current production and the FT3SS has potential capacity to secrete it. Recombinant collagen production in *E. coli* requires additional metabolic pathway modifications. Routes for these improvements are identified and implemented through a combination of molecular and metabolic modelling techniques.

1.1. Synthetic Biology

Synthetic biology aims to design and construct novel biological parts, devices and systems and also redesign existing biological systems for useful purposes. In conjunction with these aims, synthetic biology aims to change the process and strategies previously employed in existing biological fields -such as genetic engineering- by drawing on elements of the design process synonymous to engineering. Biological work can be a time consuming process of trial and error. Whereas previously it was accepted that engineering biology could be cumbersome, synthetic biology aims to eliminate the uncertainty and lack of transferability of such work. The challenges associated with engineering biology are four-fold: (1) managing biology is difficult as it is unavoidably complex, (2) constructing and characterising biological systems is tedious and unreliable, (3) there is spontaneous variation in a biological system and (4) evolution is continuous (Endy, 2005). With these challenges in mind three principles, routed in engineering are applied to synthetic biology work: standardisation, decoupling and abstraction. The three ideas listed above are conceptual; however they provide a framework for implementation.

1.1.1. Standardisation

Standardisation is central to engineering; examples include screw threads and internet web addresses. If design and construction frameworks (standards) can be developed for biological systems then it would be possible to routinely assemble multi component genetic systems or 'circuits'. To allow this a universal standard must be abided to for the characterisation of all biological components. The universal nature of standard parts means that components are both consistent and compatible. Compatibility allows parts to be assembled and integrated into functional systems with other parts. Consistency means that biological parts function as expected in different circuits, organisms and laboratories. To achieve standardisation of biology, specifications must be developed which characterise biological parts quantitatively. Once a consensus on characterisation methods is agreed on it must be adhered to for all biological parts -although consideration can always be made to improvements or additions to characterisation requirements. The output of this work is datasheets, which can be arranged into a catalogue. Datasheets are

commonplace in engineering and concisely state the behaviour, requirements and compatibility of parts. In synthetic biology these components are biological parts, which are defined as genetically encoded components which confer a biological function (Canton et al., 2008). The behaviours of biological parts can be summarised by inputs and outputs. The parts are designed so that they possess certain properties which allow compatibility with other parts. One example of a biological standard is the BioBrick, which are catalogued in The Registry of Standard Biological Parts (http://parts.igem.org/Main_Page). Parts in the BioBricks repository are designed so that they are easily integrated into circuits due by 'prefix' and 'suffix' restriction enzyme sites on each part. The catalogue is easily navigable; parts are organised into components, such as promoters or ribosome binding sites -allowing circuits to be designed easily. Standard measurements such as reliability or 'polymerases per second (PoPS)' mean that the performance and suitability of each part can be evaluated. There are however difficulties associated with standardisation, for example there is a tendency for BioBricks to behave differently in different circuits, strains, or even laboratories (Kelly et al., 2009; Zucca et al., 2013).

1.1.2. Decoupling

Decoupling describes the separation of the design process and the actual fabrication of a system (Heinemann & Panke, 2006). Successful decoupling would allow design engineers with little biological knowledge to create functional biological circuits and likewise biologists to implement a blueprint. Furthermore it allows the separation of a large problem into many simple problems. Each problem can be solved independently, solutions can then be combined (Endy, 2005). This is exemplified by the building process, where architects, project managers, engineers, inspectors and builders all work on specialist areas to produce a final product. Decoupling relies on both abstraction and standardisation to deconstruct a problem and combine solutions respectively. While the combination of the two aids the design process through decoupling, this reliance results in the limitations of the two being relayed to the decoupling process, therefore this strategy is often ineffective.

1.1.3. Abstraction

Abstraction is a method that aims to reduce the complexity of biology, easing the process of design and analysis. It also describes the fact that a cell or organism can be deconstructed to increasingly simpler constituents. This allows the formation of abstraction hierarchies, where constituent parts are embedded in a more complex system, which confers context. At the top of the hierarchy is the chassis, (much like car) components can be loaded onto the chassis to change overall performance. Different chassis (bacterial, fungal, mammalian cell line) have different merits (growth, cost, yield, ease of scaling up) and can be selected according to purpose. At the bottom of the hierarchy are parts, which are assembled into devices, which combine to form systems which are loaded onto a chassis (Arkin, 2008). A part can be relevant in a number of devices and therefore systems. This hierarchy is applicable to any machine, whether it be a car or a bacteria (Figure 1.1). The use of hierarchical structure means that biology can be viewed at the lowest level of complexity possible for a task to be carried out. For example, while the design process can be viewed at the top of the hierarchy it is implemented from the bottom, at a DNA level (Andrianantoandro et al., 2006).

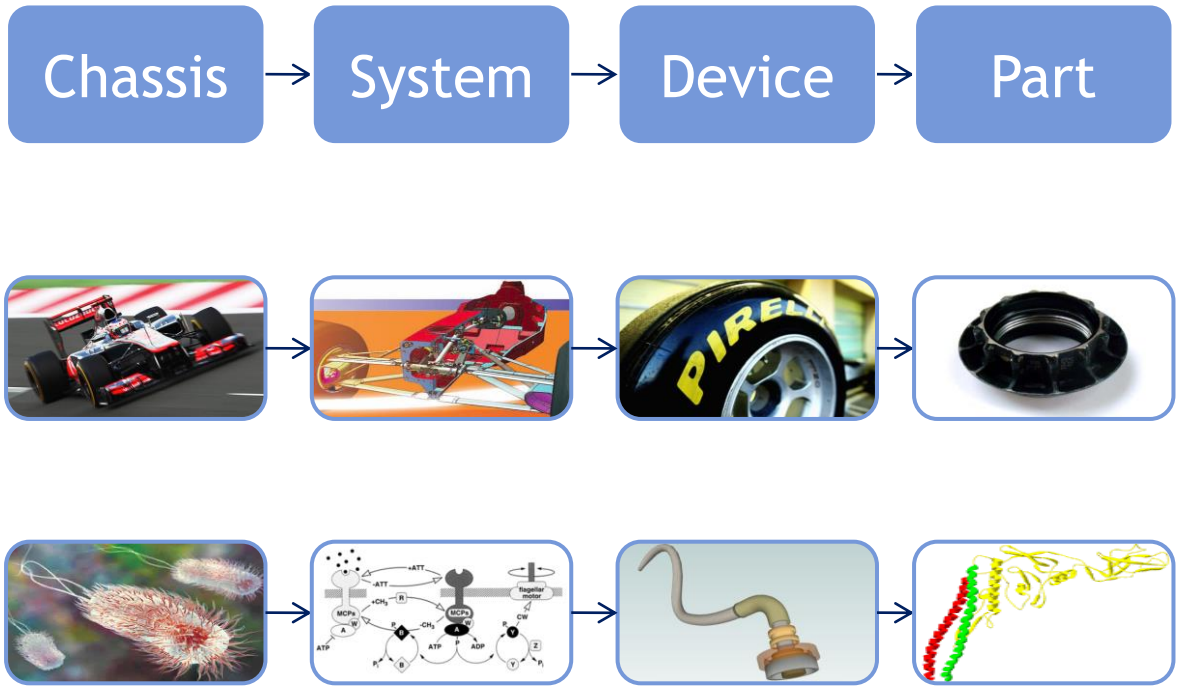


Figure 1.1: The abstraction hierarchy exemplified by a formula one car and *E. coli* as chassis.

Respective hierarchal components are the steering and motility system as systems, the wheel and bacterial flagella as devices and wheel nut and a flagellin subunit as parts.

Abstraction takes a reductionist perspective on biology. While this is convenient, it is not necessarily an accurate representation of biological systems. Reductionism versus a system wide view (holism) is a recurring debate in molecular biology (Gatherer, 2010). Holism is the basis of the rationale for systems biology, which involves observing a biological system as a whole (Mazzocchi, 2012). Application of systems biology to synthetic biology may aid integration of the two approaches.

1.1.4. The reality of synthetic biology

There are many problems associated with the synthetic biology vision of the future of biological engineering. Incomplete knowledge of biological systems contributes to much of this – something that the biological community are constantly striving to improve. Biological systems are not stable. Fluctuations in gene expression are common and strains accumulate mutations over generations, resulting in a change in phenotype (Kwok, 2010). Chassis organisms often have short generation times; therefore the rate at which mutations accumulate is amplified. In addition the rate at which mutations occur is increased in response to evolutionary pressures caused by the environment (Lenski & Travisano, 1994) and has been shown to increase with increased protein expression (Sleight et al., 2010). Genetic variation results in reduced predictability of phenotype and may amount to complete loss of function. Many issues stem from the lack of modularity within biological systems. While a circuit can be designed, it is not insulated from other cell processes. This impacts on the reliability and predictability of systems. Various methods can be employed to aid circuit insulation, these include deleting competing reactions, aiding protein interaction by protein fusion or rational engineering or localising components on a scaffold (Agapakis et al., 2010). An alternative solution to the variability seen is to complete tasks through a population of cells rather that single cells, thus decreasing overall variation (Andrianantoandro et al., 2006). Other steps, which are being taken to improve predictability, concern computational tools such as modelling. This involves genomic scale metabolic reconstructions, which are based on biological data. All known cell processes are annotated as a series of reactions which occur between metabolites (Andrianantoandro et al., 2006). Metabolic reconstructions include reaction inputs, outputs, rate constants and stoichiometry for each reaction, again based on experimental evidence. Reactions can be modified to account for change in culture media, gene knockouts or upregulation. Modelling methodologies can be applied to the reconstruction to simulate metabolic processes within the cell.

Modelling can be applied to many research questions, providing a dry-lab route of calculating the feasibility of systems, identifying routes for system improvement and also troubleshooting designs which do not perform as predicted. A common form of metabolic network simulation analysis is flux balance analysis (FBA), where the simulated flow of metabolites through a network is measured in response to an objective function (Reed et al., 2003; Orth et al., 2010). The system is subject to constraints therefore the flux of metabolites is limited to a solution space. Linear optimisation will produce a solution with a mini- or maximised objective function. The objective function may relate to biomass or be metabolite specific. Comparisons of solutions to different objective functions or alternatively the same objective function on different metabolic networks (Figure 1.2) are used to validate predictions, search for solutions or identify bottlenecks, therefore revealing targets for modification. Models do not always represent the biology well. In a bid to improve this, extensions of FBA have been developed, for example models such as 'FlexFlux' which involves regulatory metabolic network reconstruction (Marmiesse et al., 2015). This accounts for variable gene expression in a system in along with metabolic flux. Regulatory network construction based on knowledge of gene expression makes it possible to constrain a reaction to a certain percentage of its maximal flux value. This percentage corresponds to the estimated probability of activation of the gene associated to the reaction. Implementation of this allows more accurate models and for more complex phenotypes to be studied.

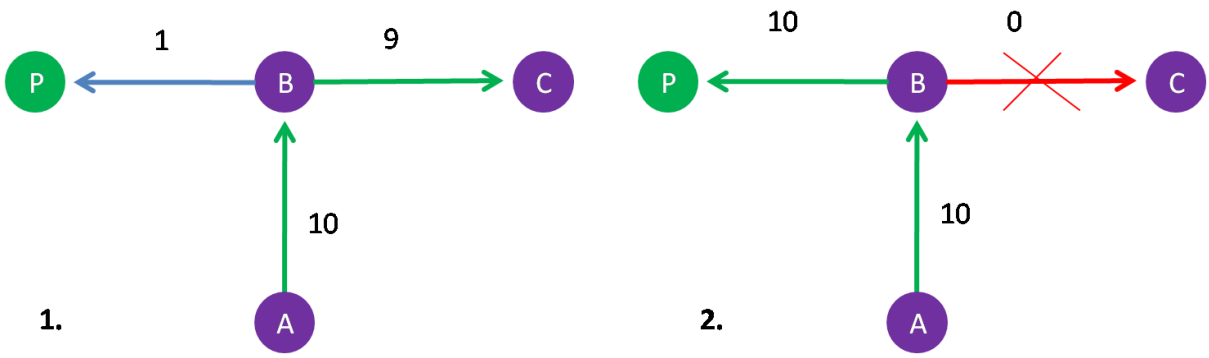


Figure 1.2: Schematic of a section of a metabolic reconstruction.

Flux balance analysis can be applied to measure the flux of metabolites through the metabolic reconstruction to maximise an objective function. In this schematic the objective function is to produce more 'P'. 1) In the model 9/10 metabolite flux goes through reaction B to C. 2) If the B to C reaction is knocked out of the model flux balance analysis shows that all metabolic flux will travel through the B to P reaction and more P is made.

1.1.5. Synthetic biology approaches: top down and bottom up

There are two approaches which can be taken within synthetic biology, top down and bottom up. Bottom up is currently far rarer and involves building an organism from scratch. Synthia is a rare example of this approach. Synthia was the product of the transplantation of a chemically derived minimal genome of *Mycoplasm mycodes* into a *M. capricolum* recipient cell, resulting in a cell which was self-replicating (Gibson et al. 2010). The *Mycoplasm* genus harbour the smallest complement of genes in any genus, furthermore the *M. mycodes* genome had previously been modified to remove all dispensable genomic material. Since this breakthrough there have been two iterations of removal of non-essential 'pseudogenes' from this chassis, resulting in the 531kbp, 473 gene Syn3.0. However it is noted that establishing a minimal genome relies on the top down approach; in fact the production of Syn3.0, was solely top down, relying on transposon mutagenesis of the existing Syn2.0 and strain characterisation (Hutchison et al., 2016). Work on minimal genomes is important to synthetic biology and much work is focusing on this in a variety of organisms (Juhas et al., 2011). The top down approach draws on the aforementioned points much more, repurposing existing biological function in an organism. In this approach, biological parts, devices or systems can be re-assembled, removed or added to those which exist in the chassis, to perform a specific task (Roberts et al., 2013). There is much confusion between the two approaches and the terms are often used interchangeably (Bedau et al. 2009; Purnick & Weiss, 2009).

1.1.6. Progressive synthetic biology

Synthetic biology is still very much emerging and evolving, but has matured over the past 15 years - improving on some of these 'hard truths'. For example stability of genetic circuits by implementing bidirectional promoters, where an essential gene (i.e. antibiotic resistance), protects a non-essential gene. Progression in the field has produced complex circuitry including, riboswitches, oscillating circuits, multiple input logic gates and impressive multicellular cross talk patterning systems (Cameron et al., 2008; Barcena Menendez et al., 2015). An area of research concerning posttranslational circuits is also emerging. The complexity of synthetic biology is increasing, for example a system based on inducible recombinases was able to implement a 16 gate Boolean logic device (Siuti et al., 2013). The integration of computational and systems biology has allowed much

progress in the field, for both designing and predicting biology in a 'top down' approach and also in analysing complex experimental outputs (Church et al. 2014; Takors & de Lorenzo, 2016). With a wealth of information being produced computationally, more efficient gene editing methods such as CRISPR/Cas allows the implementation of the ideas these generate (Jinek et al., 2012) . The field is still very collaborative for example the Sc 2.0' project, set out to design a minimal *Saccharomyces cerevisiae* genome from the 'bottom up', each of the 16 chromosomes is being worked on my a different institution (Callaway, 2014). Finally real world examples of synthetic biology are emerging, for example the chemical industry has been aided by synthetic biology in the incorporation of biosynthetic building blocks to produce small chemicals –this involves the combination or redesign of naturally occurring metabolic pathway (often from multiple species) to produce diverse metabolites or compounds which are made in a more cost effective manner to alternatives or otherwise inaccessible using naturally occurring enzymes of chemical processes, examples of this include food flavourings, biofuels and therapeutics (Szczebara et al. 2003; Winter & Tang, 2012; Breitling & Takano, 2015; Jullesson et al. 2015).

1.1.7. The salvation narrative of synthetic biology

It is envisioned that the field of synthetic biology will not only change the approach taken when working in biology, chemical engineering and manufacturing, but also provide solutions to some of the world's biggest problems. The synthetic biology community aim to positively contribute to a huge range of areas such as food production, renewable fuels, healthcare, computing and electronics – often with an emphasis on sustainability and a not-for-profit onus. While the philosophy of synthetic biology is to make a great difference to global society, it often falls short of this. Products of synthetic biology are often showpieces, which although demonstrate progress in science and are excellent tools for gaining interest in the field, do little to add to the revolutionary aims of synthetic biology. In addition many successful not-for-profit projects have spin out companies which do make profit. However there are some positive and successful examples of synthetic biology. Arguably the most impressive example was the large scale production of the anti-malarial plant derived drug artemisinin in yeast. This was scaled up by Sanofi-Aventis in a business model designed to stabilise world anti-malarial drug production, improve treatment and reduce mortality worldwide (Paddon et al., 2013). While this will improve healthcare worldwide, artemisinin

is a classic example of unequal distribution of technology. Synthetic biology derived artemisinin was created and produced in more developed countries and will compete with naturally derived artemisinin which is extracted from plants cultivated by farmers in less developed countries. While the producers of synthetic artemisinin set out on a not for profit business model (Peplow, 2013), the reality was that artemisinin crops were very successful during the years of synthetic biology and reduced artemisinin production (2014-2016), with the price of a kg of artemisinin reaching $250, much lower than the $350 'no profit-no less' margin of production and the factory has now ceased operations (Peplow, 2016). This is evidence of the sad reality of the combination of business and philanthropy. Perhaps less marketable in terms of salvation narrative, but still a solution to a global problem is the production of biofuels. While there is much debate about the true carbon footprint of 'green' fuel, the field of synthetic biology has excelled in producing renewable fuel. For example Solazyme recently produced 250,000 gallons of microalgae derived biodiesel for the US Navy (Mayfield, 2013).

1.2. Industrial Biotechnology

Industrial biotechnology represents the worldwide sector of utilising microorganisms or cell expression systems to produce a range of biological products ranging from therapeutics and biofuels to industrial enzymes and small chemicals (McCormick & Kautto, 2013). It also encompasses the production of chemicals and enzymes to treat textiles and paper. A large sector within industrial biotechnology involves the production of therapeutic biologics, in the USA alone this contributes $91 billion to the economy, which continues to rise (Carlson, 2016). Biologics accounted for five of the top ten prescribed drugs in the world in 2014 and included the largest selling drug, the monoclonal antibody Humira, which achieved $11.8 billion in sales for use in treatment of rheumatoid arthritis – amongst other conditions (Lindsley, 2015). While Humira is produced in Chinese Hamster Ovary (CHO) cells, the second biggest seller in this market is the synthetic long-acting insulin substitute Lantus, which is produced in *E. coli.* A staggering number of biologics are approved for therapeutic use, for *E. coli* alone this includes many hormones (largely insulin and human growth hormone), antibody fragments, interferons for cancer treatment, colony stimulating factors (which stimulate white blood cell production) and vaccines and this is set to grow over coming years (Walsh, 2014).

1.2.1. Industrial biotechnology expression chassis (organisms)

A range of chassis options are available for recombinant protein production in industry, all with differing advantages and disadvantages. The type of chassis depends on the nature of the recombinant protein and the specifications of the product. Commonly used expression systems are laid out below; as the focus of this project is to secrete protein, various secreted yields of protein product are also given. Many eukaryotic cells can be utilised for recombinant therapeutic protein production, these may include plants, microalgae, insect cells, fungus (often yeast) and animal cells, which are usually in the form of a cell line rather than a whole organism (spider silk (and other recombinant proteins) from goat milk, is one example of an exception to this (Clark, 1998; Service, 2002)). In general they are better at carrying out PTMs than prokaryotic cells. Mammalian cells are favoured because they are able to carry out PTMs and produce biologics with the correct glycoforms to be compatible in humans. They have been utilised to produce many antibodies and hormones at yields of up to 10g per Litre (Kim et al., 2012) with CHO cells being able to secrete recombinant protein, however yields for CHO are much lower -0.8g per Litre (Le Fourn et al., 2014). *Pichia pastoris* is another commonly used eukaryotic organism; it is also able to carry out PTMs and glycoengineering allows more human like glycosylation of protein, it also has cheaper growth requirements than CHO cells and is able to secrete protein into media efficiently –for example 1.5g secreted insulin per litre of cell culture following 80 hours growth (Weinacker et al., 2013; Fidan et al., 2015; Gurramkonda et al., 2010), although the maximum secreted yield in *Pichia* is around 15g L^{-1} 21kDa chains of rodent collagen (Werten et al., 1999). *Saccharomyces* is also used but has lower yields, for example 2.5mg L^{-1} secreted amylase after 96 hours (Rodríguez-Limas et al., 2015). Microalgae are also a desirable expression organism because they capture carbon dioxide and also utilise sunlight to acquire energy. Yields of 11.8g milk amyloid were achieved per kg dry weight total of *Chlamydomonas reinhardtii* –this was not secreted (Gimpel et al., 2015).

Prokaryotes are often favoured in biotechnology for a variety of reasons. Prokaryotes can be grown quickly and to high cell density on cheap carbon sources and secretion is possible (Ferrer-Miralles & Villaverde, 2013). Furthermore scale up of growth conditions is simple. Examples include *Bacillus,* which is widely reported to be proficient in recombinant protein secretion of certain enzymes achieving yields of up to 25g L^{-1}, following cell culture for 72 hours of more (Liu et al., 2013) – however it yields reported explicitly in the literature are much less impressive than this for example 1g L^{-1} proinsulin and up to 1g L^{-1} amylase (Heng et al. 2005; Olmos-Soto & Contreras-Flores, 2003).

Another candidate is *Pseudomonas* which can secrete up to 1g L^{-1}, of various proteins (Retallack et al., 2012). However when it comes to recombinant protein production *E. coli* is often seen as a 'workhorse' organism and is often favoured in biotechnology, and certainly by the often conservative contract manufacturing community, who develop expression (and secretion) platforms for a range of products. Reasons for this will be documented below.

1.3. *E. coli* as a chassis for synthetic biology and biotechnology

The use of *E. coli* is common in biotechnology. It has fast generation times and low growth requirements, therefore a large number of cells can be grown in a matter of hours. Alternatively, it is relatively cheap to make rich media, which allows cells to be grown to OD_{600} in excess of 100, for high protein expression in a small culture volume. Along with favourable growth requirements they are also well studied and there are a large number of tools available to both understand and alter *E. coli* gene and protein expression; these will be discussed in the next section. The first example of clinically available recombinant protein was human insulin from an *E. coli* host. This was followed by bovine growth hormone, again in *E. coli* (Swartz, 2001). Following years of optimisation, intracellular yields of up to 4.34g L^{-1} have been reported for human insulin, however the highest extracellular secreted is just 7mg L^{-1}, this was achieved through fusion to *Staphylococcal* protein A, which has been shown to result in the presence of extracellular protein in *E. coli*, although lysis is never excluded (Baeshen et al. 2014; Mergulhão et al. 2004). Secretion is documented in *E.coli;* however it is not as effective at secreting protein as other expression chassis, such as *Pichia* or *Bacillus* (Rosano & Ceccarelli, 2014). Furthermore the majority of secretion reported in biotechnology is for periplasmic secretion, not secretion in the media. *E. coli* do not naturally carry out PTMs and harbour endotoxins, however as stated above, a huge range of *E. coli* derived recombinant proteins are now licensed for use in humans and are produced on an industrial scale (Walsh, 2014). These examples demonstrate the amenability of *E. coli* to recombinant protein expression; it's suitability to industrial scale up and cost effectiveness as a chassis organism. Furthermore the fact that *E. coli* derived therapeutic products are licensed, highlights that products are effective without adverse characteristics and well viewed by regulatory authorities.

1.3.1. Enabling technologies

E. coli are particularly suitable for use in synthetic biology and biomanufacturing for a number of reasons. This includes intrinsic properties, such as good recombinant protein expression rate with low growth requirements and a high rate of cell growth up to optical densities in excess of OD_{600} 100 (Soini et al., 2008). Historical use of *E. coli* in industrial biotechnology is also beneficial as it is acceptable to stakeholders and companies and contract manufacturing organisations are equipped with expertise and machinery to grow *E. coli.* In terms of making improvements to the chassis, the genome, transcriptional regulation and metabolic pathways of *E. coli* are well characterised and available in easily accessible formats, for example the online database EcoCyc (http://www.ecocyc.org/). Furthermore there are online tools for visualisation of pathways such as Reactome (http://www.reactome.org/).

Molecular enabling technologies have contributed much to the amenability of *E. coli* to biotechnological applications. There are a large number of tools available for molecular manipulation of genomic and plasmid DNA (Sørensen & Mortensen, 2005; Rosano & Ceccarelli, 2014). The toolset available is also well tailored for achieving protein overexpression and technologies may be implemented for a range of tasks which include cloning, (recombinant) protein expression and protein purification and isolation. Many of these technologies are aided by the post genomic revolution, which has resulted in fast, cheap gene sequencing and synthesis. For example it was possible to sequence the genome of the parent strain which formed the basis of this study for £50, by Microbes NG (https://microbesng.uk/). In terms of recombinant protein expression, commercial cloning systems are widely available, common examples include blunt end vector cloning and gateway cloning (both Invitrogen). Genes for insertion can be assembled by PCR or synthesised de *novo* by services such as GeneArt® (Invitrogen) and GenScript. The costs of these commercial systems are continually decreasing, gene synthesis can cost from £0.23/bp (GenScript, 2016). Gene synthesis companies also offer codon optimisation for *E. coli*, this technology takes advantage of the fact that multiple codons encode the same amino acids and often realises higher expression levels- though this is not always a guarantee of functionality or solubility. The prevalence of tRNAs for each codon varies between organisms, therefore optimising codons to reflect this results in more efficient translation (Ikemura, 1981; Rosano & Ceccarelli, 2014).

Development in molecular methods has also greatly increased the ease of modifying chromosomal DNA of strains, allowing generation of knock out mutations or insertion of recombinant DNA. Traditionally a combination of restriction enzymes and ligase were used to generate mutations, however these methods can be considered cumbersome and inefficient (Warming et al., 2005). Recombineering techniques are ideal for removal or addition of genetic material into the *E. coli* chromosome. *E. coli* are not readily transformable with the presence of linear DNA alone. This is largely due to the presence of exonucleases (namely the RecBCD encoded exonuclease V) which digest linear DNA (Datsenko & Wanner, 2000). RecBCD deficient strains allow recombination, but with low efficiency. A RecD mutant was shown to be exonuclease V deficient, but still displayed high recombination efficiency (Biek & Cohen, 1986). Alternatively phage based recombineering methods provide simple, efficient means of introducing mutations. Lambda-Red plasmid based technology is widely used; the Red recombinase containing plasmid encodes proteins which inhibit RecCBD activity, separate linear DNA into single strands and anneal complementary single stranded DNA to chromosomal DNA. The plasmid is inducible, allowing control of recombination events (Murphy, 1998). One of the most commonly used versions of this protocol is the one generated by Datsenko & Wanner (2000) and will be discussed later in this thesis (Chapter 3) –although other methods are available which may result in more efficient or scarless mutagenesis, for example making use of the I-SceI meganuclease (Kim et al. 2014; Lee et al. 2009; Stringer et al. 2012). An alternative historical method for mutagenesis is phage transduction where bacteriophages transfer gene deletions from a donor bacteria to a host- usually P1 for *E.coli* (Donath et al., 2011), although this relies on the occurrence of a strain which already harbours the desired gene mutation, it is very efficient.

These tools can be utilised to overcome the challenges associated with protein production in *E. coli*. Protein overexpression often results in aggregates which form insoluble inclusion bodies rendering protein inactive. There are benefits to inclusion body formation, as it does hinder proteolysis, therefore product is not degraded (Sørensen & Mortensen, 2005; Rudolph & Lilie, 1996), however protein must be resolubilised to be functional and this complicates downstream processing. In addition, prokaryotic cells lack eukaryotic post translational machinery for glycosylation, phosphorylation, acylation and methylation, therefore recombinantly expressed proteins are often not correctly modified (Wacker et al., 2002; Yue et al., 2000; Sahdev et al., 2008). However steps have been made to implement and improve PTM capacity of *E. coli* which will be discussed later. A reduction in the rate of cell growth can arise when producing heterologous proteins; this is referred to as the metabolic burden and describes the amount of resources (both raw material and energy)

which are withheld from the host metabolism in order to express recombinant genes (Sørensen & Mortensen, 2005; Glick, 1995). Metabolic burden can be decreased by deleting unnecessary genes (including those for antibiotic resistance) or incorporating exogenous plasmid harboured DNA into the chromosome (Lee et al. 2009; Jones et al. 2000; Glick, 1995). Often, plasmid expression system characteristics allow controlled protein expression to overcome these challenges. Plasmids have different copy numbers and induction characteristics, allowing control of the timing and strength of protein expression; these can be tailored to promote the formation of high quality protein products without the formation of inclusion bodies or high metabolic burden (Rosano & Ceccarelli, 2014). Alternatively biological circuits can be designed to promote optimum protein expression by modifying promoters and ribosome binding sites (Young & Alper, 2010). The availability of catalogued and characterised promoters, ribosome binding sites and terminators in the BioBricks parts registry and on ribosome databases (for example http://rdp.cme.msu.edu/) means that designing circuits has, in theory been simplified over recent years. Co-expression of enzymes which aid correct folding (chaperones, hydrophilic proteins) and PTMs (N-linked glycosylation system of *Campylobacter jejuni* or prolyl-4-hydroxylase for proline hydroxylation) can also be incorporated into expression systems (Baneyx, 1999; Khokhlatchev et al. 1997; Wacker et al. 2002; Ikeno et al. 2013; Kranen et al. 2014; Pinkas et al. 2011). The combination of a large range of protein purification tags, all with different binding properties and choice of cleavage sites means that protein purification can be optimised for each expression system (Hunt, 2005). Protein tags or fusions can also aid protein folding and direct protein for secretion (Makrides, 1996; Rosano & Ceccarelli, 2014).

In addition to molecular tools, commercial strains have been optimised to aid the yield and quality of recombinant protein expression. These strains improvements are commonly concerned with posttranslational modifications. Mutant commercial strains such as Origami™ (Novagen) express thioredoxin reductase and glutathione reductase, which enhance disulphide bond formation in the cytoplasm, achieving correctly folded proteins (Prinz et al., 1997). Likewise the SHuffle® (NEB) expression system has enhanced cytoplasmic disulphide bond formation by expressing a disulphide bond isomerase, which aids production (Lobstein et al., 2012). BL21-CodonPlus (Agilent Technologies) and Rosetta™ (Novagen) enhance expression of proteins with rare *E. coli* codon sequences (Hunt, 2005; Sahdev et al. 2008). Combined, these enabling technologies allow controlled expression of recombinant proteins and effective protein purification in *E. coli* – reinforcing that this organism is an excellent chassis.

1.3.2. Cellular location and its effect of the protein product

Expressed protein can be directed to the periplasm, cytoplasm or secreted to the media (Cornelis, 2000). The preferred site of export is specific to the recombinant protein product. The cytoplasm is often favoured as it tends to result in a higher production yield, however there are also disadvantages to cytoplasmic location (Sahdev et al., 2008). Disulphide bond formation usually occurs in the periplasm. Therefore the formation of inclusion bodies is common if protein remains in the cytoplasm, as periplasmic molecular chaperones and folding enzymes are absent and incorrect folding and aggregation occur (Sørensen & Mortensen, 2005; Baeshen et al. 2015). Although it should be noted that some folding chaperones are present in the cytoplasm (Berlec & Štrukelj, 2013). The periplasm is also a reducing environment, which can aid protein folding (Eser et al., 2009). As discussed previously inclusion body formation can be beneficial; initial purification is simpler, N-terminal authenticity is greater and the protein is protected from proteolysis (Hannig & Makrides, 1998). However it also inhibits protein functionality, therefore inclusion bodies must be unfolded and refolded correctly, this is often complicated, time consuming, hard to predict and therefore costly (Singh & Panda, 2005). There are strategies to overcome protein aggregation, which include optimising protein expression with low copy number and promoter strength, growing culture at a low temperature and using strains that are designed to (over)express certain molecular chaperones, which reduce protein aggregation. These can be split into folding chaperones (such as ATP dependent Dnak and GroEL), holding chaperones (which hold unfolded proteins until chaperones are available) and disaggregating chaperones (such as ClpB, which solubilises aggregated protein (Mergulhão et al. 2005; Gupta et al. 2009; Berlec & Štrukelj, 2013). With the correct secretion signal protein can be secreted into the periplasm. This can be desirable as around 4% total cell protein is located here, thus downstream processing is simplified. Proteolysis is also reduced and the non-reducing environment allows formation of disulphide bonds (Berlec & Štrukelj, 2013; Yoon et al. 2010). Extracellular secretion would reduce the incidence of inclusion bodies, but also confer all of the positive attributes of inclusion body formation (Choi & Lee, 2004). Purification would be simpler; proteolysis less common and N-terminal authenticity would be maintained. Much is known about secretion signal peptides which direct proteins for export by secretion (Nielsen et al., 1997; Dyrløv Bendtsen et al., 2004) and these can be exploited to allow directed secretion of tagged proteins, through a secretion system of choice into the periplasm of *E.coli* with great efficiency. Additionally purification of extracellular protein, eliminates the steps usually implemented to first lyse cells and second remove cellular contaminants (van der Wal et al., 1998). Profit would be

increased due to the reduction in downstream processing costs and increase in product yield, as proteins are less likely to undergo proteolysis and more likely to be correctly folded (Choi & Lee, 2004). In summary secretion of protein into the extracellular media would be of great benefit in industrial biotechnology.

1.3.3. Post translational modifications

Problems at a post-translational level are rife when expressing recombinant protein in *E. coli*, as much of the post translational machinery required is absent. To circumvent this, machinery can be introduced on plasmid vectors. This technique is common, for example although *E. coli* is occasionally able to phosphorylate recombinant proteins using native kinases (Mijakovic et al., 2006), the co-expression of recombinant protein kinase (often fused to the target protein) confers reliable phosphorylation of proteins (Sahdev et al., 2008; Murata et al., 2008). This strategy has also been used for protein methylases and acetylases and their substrates (Yue et al., 2000; Acharya et al., 2005). Glycosylation is a complex PTM, which while common in eukaryotes is rarer in prokaryotes. There are two forms, N- and O-linked. The process of N-linked glycosylation in eukaryotes involves the oligosaccharyl intermediate (Glc3Man9GlcNAc2) being attached to an asparagine residue through the action of the enzyme oligosaccharyltransferase. The resulting structure is then trimmed by α-glucosidase I and II for quality control and can then be trimmed further or decorated with monosaccharides. O-linked glycosylation is much simpler and involves the attachment of a single GalNAc monosaccharide to a serine or threonine of the polypeptide backbone -next there is either the addition of a SA residue, which terminates the chain, or the addition of further monosaccharides to form a longer linear or branched chain (Brooks, 2004). While new examples of bacterial protein glycosylation are discovered all the time, their characteristics are significantly different from the common core and terminal glycans present in eukaryotic proteins- both therapeutic and non-therapeutic. This is problematic when producing heterologous proteins where glycosylation status is key to efficacy or function -effecting protein half-life, stability, solubility and immunogenicity (Baker et al., 2013); for example, N- linked glycosylation has also been shown to improve pharmacokinetic and biophysical properties of proteins. In addition recognition of N-glycans by carbohydrate binding proteins and lectins allows tissue and cell specific targeting of proteins (Lizak et al., 2011). O-linked glycosylation also aids protein stability and recognition, including the immune response (Steen et al., 2008). For these reasons correct glycosylation of

therapeutic products is advantageous as protein will be more characteristic of the native protein in structure and function and is more likely to interact with other proteins and less likely to elicit an immune response. Some non-enzymatic N- glycosylation of recombinant human interferon γ has been observed in *E. coli* (Mironova et al., 2005), however it is not a common feature. The discovery of an N-linked glycosylation system in *Campylobacter jejuni* which is transferable to *E. coli* has added to the possibility of bacteria derived glycoproteins (Wacker et al., 2002). While this research was central to enabling PTM's in *E. coli,* the efficiency of N-linked glycosylation was poor and occurs only in the periplasmic space. However additional work has demonstrated that this can be improved by codon optimisation of the oligosaccharyltransferase PglB, which transfers the *C. jejuni* heptasaccaride to asparagine (N) residues and by increasing expression of the glycosyltransferase WecA (Pandhal et al., 2012). *E. coli* can also be engineered to confer O-linked glycosylation with the expression of recombinant enzymes (Lubas, 2000; Henderson et al. 2011).

1.3.4. Immunogenicity

It is imperative that proteins derived for therapeutic uses have low or at least predictable immunogenicity. Differences in the 3D structure of proteins to those in humans are detected by antibodies, eliciting an immune response. The use of non-human expression systems can increase the immunogenic properties of therapeutic proteins. As explained previously, a recombinant human protein may not harbour all of the human PTMs; therefore an immunogenic response is likely. Additionally if the whole protein is not expressed or correctly folded it may result in exposure of new antigenic sites (Schellekens, 2002). While there are developments to reduce immunogenicity by implementing correct PTMs, as stated in the previous section, the production of small therapeutics which do not require PTMs such as glycosylation are advancing to clinical trials (Frenzel et al., 2013). Furthermore the storage of recombinant proteins may result in oxidation or aggregation which may elicit an immune response. The covalent attachment of polyethylene glycol (PEGylation) to recombinant therapeutic proteins can aid evasion of the immune system, as it shields sites of antigenicity (DeFrees et al., 2006). Furthermore PEGylation of protein can increase circulatory half-life, and protect the protein from enzyme degradation. It is also important to eliminate any *E. coli* derived contaminants which may induce an immune response (Schädlich et al., 2009). A common contaminant is endotoxin (often in the form of lipopolysaccharide chains in the case of *E.coli*) which is present on the *E. coli* outer membrane. Even minute amounts of endotoxins can have a large

effect in humans. Steps in the protein purification methods must ensure that such contaminants are removed (Petsch & Anspach, 2000). A widely used example of this is the use of polymyxin B which has a high binding affinity for the lipid A moiety of most endotoxins.

1.3.5. *E. coli* protein secretion systems

As stated, secretion of recombinant protein is desirable in industrial biotechnology, as product yield may be higher in concentration and quality and downstream processing simplified. Protein secretion in widespread in *E. coli* and can be active or inactive. For example periplasmic translocation without the release of cytoplasmic contaminants is observed when cells undergo osmotic shock (Jonasson et al., 2002). This section will focus on active routes of protein secretion to various cellular locations through a range of secretion systems (Figure 1.3). Together, these secretion systems allow protein secretion from the cytoplasm into the periplasm or extracellular space, in a one or two step secretion pathway.

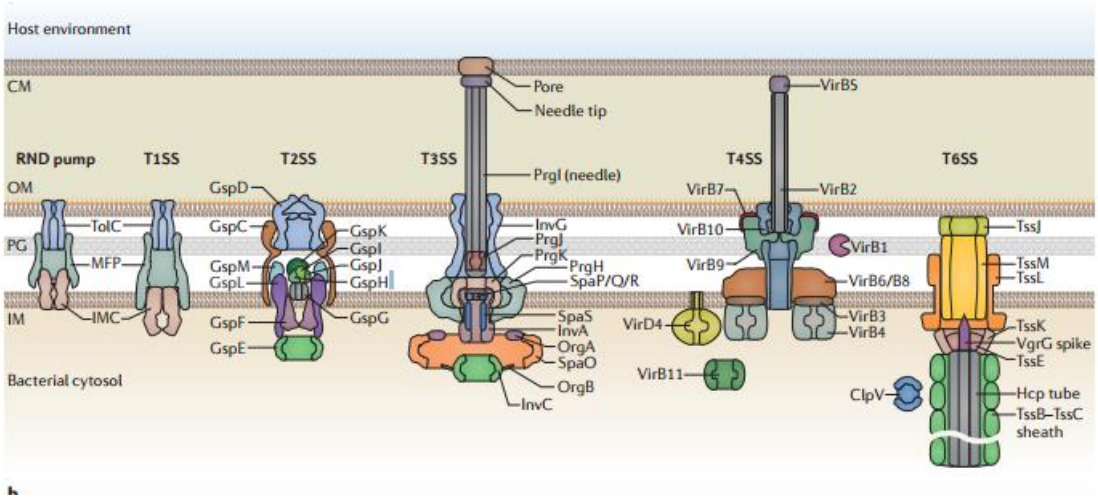


Figure 1.3: Secretion systems of *E. coli* and the proteins involved.

Showing the RND pump, type I, type II, injectisome type III, type IV and type VI secretion systems and their localisation in the inner membrane (IM), outer membrane (OM) and host cell membrane (CM). NOTE: Only the type III injectisome system is shown, not the flagella type III secretion system. Taken from (Costa et al., 2015)

1.3.5.1. Type I secretion system

Type I secretion in *E.coli* exports protein directly from the cytoplasm to the environment in one step. A number of substrates are secreted which are often associated with nutrient acquisition or virulence. In terms of their architecture they are also closely related to the resistance–nodulation–division (RND) family of multidrug efflux pumps (Costa et al., 2015). Both systems form a three component double-membrane-spanning channel with an ATP-binding cassette transporter (ABC) or RND efflux transporter which spans the inner membrane, a TolC membrane protein which is embedded in the outer membrane and an inner membrane anchored membrane fusion protein (MFP) which forms a tunnel like structure with TolC in its open form which spans the periplasm - linking the other two components (Tseng et al., 2009). TolC is a trimer that forms a β-barrel channel in the OM and a periplasmic α-helical barrel that associates with the MFP from multiple secretion pathways including the hemolysin (HylA), colicin V and RND (Delepelaire, 2004). Opening of the channel occurs at the periplasmic end of the α-barrel and is triggered by a twisting motion of the α-helix (Koronakis et al., 2004). Secretion is ATP facilitated by the inner membrane ABC. Both ABC and MFP carry out substrate recognition prior to secretion, whereas TolC can associate with multiple MFP complexes. Protein is directed for export by a non-cleavable C-terminal signal sequence, which binds to the ABC (Zhang et al., 1995). Secretion into the periplasmic cavity of the MFP is energized by the hydrolysis of ATP, which is mediated by the C terminus of the ATP-binding cassette complex (Thomas et al., 2014; Kanonenberg et al., 2013) –RND pumps also use a proton gradient to aid substrate secretion. This enables the ABC-MFP complex to associate with TolC. This triggers the opening of TolC and the subsequent release of the substrate into the extracellular space (Costa et al., 2015).

1.3.5.2. Type II secretion system

Secretion through the type II secretion system is a two-step process and machinery spans from the inner to outer membranes. T2SS are comprised of a channel like outer membrane complex (often referred to as the secretin), a periplasmic pseudopilus, which is anchored to an inner membrane platform -and is tightly associated with a cytoplasmic ATPase (Korotkov et al., 2012). First protein is exported through the Tat or Sec pathway from the cytoplasm to the periplasm, then protein is

exported to the environment (Mergulhão et al., 2005). T2SS substrates are transported to the periplasm as unfolded polypeptides by the Sec translocon or as folded proteins by the Tat transporter; protein is exported through the secretin in a folded state (Nivaskumar & Francetic, 2014). N-terminal secretion peptides direct protein for export to the periplasm through both the Sec and Tat pathways, which are cleaved following secretion into the periplasm (Choi et al., 2006; Patel et al., 2014). The secretion peptides of these two pathways are similar in structure; both harbour Ala-X-Ala motifs and hydrophobic and hydrophilic regions; however the Tat signal is less hydrophobic and also contains the twin arginine SRRxFLK sequence. Furthermore basic residues in the Tat signal, which are not present in the Sec signal, are thought to hinder interaction with Sec secretion apparatus. Substrates are recognised by TatC and assemble at the TatBC complex receptor where TatA is recruited and facilitates transport of folded protein across the inner membrane, this complex binding and possible translocation are driven by proton motive force (Alcock et al., 2013). The signal peptide is cleaved in the periplasm and following folding and PTM's may be directed for export through the secretin to the extracellular membrane by the chaperone AB_5 which is thought to recognise the tertiary or quaternary structure of folded proteins (Sandkvist, 2001). There are two targeting pathways for delivering protein to the Sec translocon, the SecA pathway is mainly used for periplasmic or outer membrane proteins and recognises proteins, which have a cleavable hydrophobic N-terminal Ala-X-Ala motive signal sequences. These proteins are transferred to SecA, which drives translocation through the SecYEG translocon channel by ATP hydrolysis (Denks et al., 2014). The SRP pathway mainly secretes inner membrane proteins; here ribosome bound SRP-signal nascent protein is directed to the membrane where it is recognised by the SRP membrane receptor: FtsY. FtsY binds to SecY and upon GTP hydrolysis promotes the dissociation of the SRP–FtsY complex delivering the protein cargo to the SecY channel (Chatzi et al., 2013). Once in the periplasm, N-terminal signal peptides are cleaved -at this stage proteins reside in the periplasm. Current understanding of how protein reaches the extracellular space is limited. The prevalent idea is that protein binds to the periplasmic domains of secretin and the tip of the pseudopilus, stimulating ATPase activity, subunits are added to the pseudopilus, and the growing pseudopilus functions as a piston, pushing substrates through secretin channel, however knowledge of how proteins are targeted to this complex is lacking (Nivaskumar et al., 2007; Korotkov et al., 2012).

1.3.5.3. Type III secretion system

The type III secretion system (T3SS) describes two distinct systems, the flagella and injectisome T3SS. The flagella T3SS shares many components with the injectisome T3SS and many of the proteins are homologous in structure and function (Erhardt et al. 2010). There is much discussion around the evolutionary divergence of the two T3SSs. With evidence that each diversified from the other or that they share a common ancestor (Abby & Rocha, 2012; Macnab, 2004; Gophna et al. 2003). Both systems span all of the cellular membranes and have the capacity to secrete a large number of protein subunits from the cytoplasm, direct to the supernatant or into host cells in one step.

1.3.5.3.1. Injectisome

The T3SS injectisome is present in several pathogenic bacteria (including enteropathogenic *E. coli* (Jarvis et al., 1995), much of the literature concerns *Salmonella* or *Yersinia*. The secretion system allows the bacteria to inject toxic proteins through the host eukaryotic cell membrane and is essential for infection, survival and pathogenicity (Cornelis & Van Gijsegem, 2000). The injectisome is comprised of around 25 proteins, which assemble in a highly regulated, ordered fashion. The central component is the needle complex, which spans both the inner and outer membranes and allows the translocation of secreted effector proteins from the cytoplasm to the extracellular environment (host cell) in one step (Diepold & Wagner, 2014). Other than the basal needle structure the injectisome is comprised of rings, export apparatus and cytostolic proteins (Galán & Wolf-Watz, 2006).

The basal structure is formed of an OM and IM ring which encompass the inner rod -the T3SS export apparatus is situated within the IM ring and is composed of two proteins (PrgK and PrgH) which form concentric rings. The secretion apparatus resides in the IM and forms a sorting platform for substrates and an ATPase complex mediates interaction with the export apparatus, which exhibits substrate recognition and unfolds substrates prior to export (Lara-Tejero et al., 2011). The export apparatus is thought to recognise a common N-terminal signal sequence, which is also the location of binding of T3SS chaperones for certain effector proteins (SopB and SigE for example (Roblin et al., 2015)), although in the case of some proteins a 5' UTR is thought to play a role, although this is

under debate (Munera et al., 2010; Lloyd et al., 2001; Anderson et al., 2002). The rod (PrgJ) assembles on the IM ring and finally the OM ring is made of InvG proteins from the secretin family, which are moved to the outer membrane with the aid of pilotins (InvH) (Koster et al. 1997; Crago & Koronakis, 1998). A neck region spans the periplasm and links the OM ring to the IM ring. The needle structure is then assembled on to the rod structure. This is comprised of around 100 <10kDa polymerised hair-pin like proteins called PrgI. Chaperones generally bind within the first 160 residues of the N-terminal of substrates. The ATPase complex (InvC ATPase) couples ATP hydrolysis to the unfolding and release of needle and effector proteins from their chaperones prior to secretion through the T3SS (Akeda & Galán, 2005); proton motive force may also aid secretion (Kosarewicz et al., 2012). The needle complex is traversed by a 28 Å diameter channel, through which substrates travel before assembly at the tip. At the tip of the needle assembles the translocon platform, upon contact with the host, two transmembrane proteins (IpaB and SipB) locate here to form the transposon and toxic proteins or 'effectors' are delivered from the bacterial cytoplasm to eukaryotic cells through this channel (Costa et al., 2015). These are analogous to the late (filament type) subunits of the flagella system these proteins are often chaperoned by specific chaperone proteins that bind in the N-terminal secretion signal region and are involved in docking at the export apparatus.

1.3.5.3.2. Flagella

The flagella T3SS is ultimately present to build a functional flagellum and provide motility to the cell, through motor and filament structures. The FT3SS shares morphology and mechanisms of assembly with the bacterial injectisome, with both assembling in a similar ordered fashion from the base upwards in steps which are controlled by checkpoints and a defined regulatory pathway, alongside several conserved proteins in the secretion apparatus- as outlined in Section 1.4.2 (Chilcott & Hughes, 2000). The structure consists of a basal body and cylindrical proteins which extend from the cytoplasm and protrude from the bacterial cell. These structures assemble at the distal end following secretion through the T3SS (Macnab, 2004). The FT3SS forms the basis of the next section of this introduction and will discussed in detail in Section 1.4

1.3.5.4. Type IV secretion system

The type IV secretion system is the most ubiquitous secretion system in nature and is able to transport both toxins, effector proteins and nucleic acid from bacteria to host cells. The T4SS spans both membranes of gram negative bacteria and transport is ATP dependent (Zechner et al., 2012). Subunits of six proteins form the scaffold and translocation apparatus is formed. Two types of subunit form the pilus that extends into the extracellular space and VirB1 is a periplasmic lytic transglycosylase that degrades the peptidoglycan layer and is required for pilus biogenesis. The system is powered by three ATPases located in the IM, which serve to provide energy for pilus biogenesis and substrate translocation (Costa et al. 2015; Fronzes et al. 2009; Cascales & Christie, 2004). The pili extend into the extracellular space and it is proposed that when they contact a host cell, the T4SS secretion switches to substrate translocation. C-terminal signal sequences which are comprised of unstructured clusters of positively charged or hydrophobic residues or internal signals, direct protein for export while nucleic acid export is thought to be contact-dependent initiated upon binding to a DNA dependent ATPase coupling protein (VirD4) –which favours GGGG nucleotide structures. This is directed for export through the T4SS (Darbari & Waksman, 2015). Much like other systems, proteins often require chaperones and adapter proteins for efficient secretion (Zechner et al., 2012; Christie et al., 2014). This general requirement for chaperones in protein secretion is both intriguing and seems to be a solution to prevent mis-folding and aggregation in these pathways since many of the proteins are destined for polymerisation.

1.3.5.5. Type V secretion system

There are now 5 classes of type V secretion systems types (a-e), however the best characterised are a and b (van Ulsen et al., 2014). The autotransporter (Va) and the two partner secretion pathway (Vb). Proteins secreted through the type V secretion pathway are comprised of passenger domains, signal sequences and β-domain regions (Henderson et al., 2004). Proteins are recognised by N-terminal signal sequences and directed into the periplasm by the Sec machinery and occasionally the SRP pathway. In the autotransporter system, once in the periplasm the signal sequences are cleaved and the β-domains assemble into β-barrel structures in the cell membrane. The passenger domain is then translocated across the membrane where it is released or remains anchored in the cell

envelope for surface display. In the Vb system the β-domain and passenger domain are translated as two peptides, as opposed to one in the autotransporter.

1.3.5.6. Type VI secretion system

The type VI secretion system is also an injectisome and analogous both structurally and mechanistically to an intracellular membrane-attached contractile T4 bacteriophage tail. It is comprised a tail complex which consists of structural elements that are equivalent to the contractile phage tail: a tail sheath, an inner tube which assemble on a baseplate (Ho et al., 2014). This is anchored to the cell by the membrane complex which extends from the inner to outer membrane. Extracellular signals cause a conformational change in the baseplate, which drives contraction of the sheath. The inner tube is transported out of the cell and effectors are recruited to the N- or C-terminals of the two proteins which form the spike domain at the tip (Costa et al., 2015). The T6SS therefore spans both of the bacterial membranes, transferring toxic effectors from the cytoplasm to eukaryotic cells in one step by propelling the structure along with effectors into mammalian or bacterial host cells; retraction of this requires ATP (Wandersman & Coulthurst, 2013). A single T6SS contraction event might translocate multiple effectors into a target cell (Shneider et al., 2013). The system appears to be a hybrid system, with components derived from the bacteriophage and type IV secretion systems. Its discovery and mechanism of action was a real surprise to the protein secretion field and is a wonderful example of biological evolution at work- though whether this comes from hijacking of a viral system by bacteria or vice versa is an open question.

1.3.6. Secretion of recombinant proteins

On average downstream processing can be said to account for 80% costs of biomanufacturing (Hellwig et al. 2004; Walsh, 2014). If recombinant proteins are secreted then downstream processing is simpler and therefore cheaper, furthermore cells can remain in culture without being disrupted and therefore continue to function. As discussed in section 1.3.2, following (or in the absence of) protein secretion, protein can be localised to the cytoplasm, periplasm or extracellular space. The properties of these environments affect protein quality and concentration. Secretion may aid correct protein folding, inhibit the formation of inclusion bodies and reduce product degradation (Mergulhão et al., 2005). Enhanced correct protein folding and lack of proteases also means that the

product is structurally robust and therefore biologically active. Secreted protein has better N-terminal authenticity, as cleavage of the secretion signal often occurs at the N-terminal following secretion, thus reducing the presence of N-terminal methionine residues at the N-terminus which can effect biological properties of the protein, including immunogenicity (Baneyx, 1999). Secretion of recombinant protein into the periplasm is commonly carried out in industry. This is because of the benefits which arise from non-cytoplasmic protein expression. Protein is harboured in a small and contained area, which aids initial isolation of protein, however periplasmic extraction must then be utilised to retrieve protein, followed by purification of the product from the remaining protein. As previously stated, secretion into the extracellular environment would be beneficial for both cost effectiveness of downstream processing and improved product quality. Much work on extracellular secreted protein relies more on the presence of secreted protein, without knowledge of the route by which it was localised there (Ni & Chen, 2009). This section of the review will focus on secretion of protein through known secretion pathways.

Protein secretion in *E. coli* is a complex process, therefore there are many issues surrounding the secretion of recombinant proteins. These include incompatibility for secretion, incomplete secretion, low capacity of secretion and proteolysis (Huang et al. 2001; Baneyx, 1999). For example the secretion of some large proteins may be impossible, as this is limited by secretion system capacity. When protein secretion demand exceeds the capabilities of the cell, intracellular accumulation occurs, leading to the formation of inclusion bodies, therefore the two must be evenly matched (Schlegel et al., 2013). To date recombinant products have been secreted through the type I, II, III and V secretion systems in *E. coli,* all with limited success.

Recombinant proteins have been secreted via the 3.5nm diameter type I secretion system by fusing them to the HlyA secretion signal (Blight & Holland, 1994). As protein transits in an unfolded state, the secretion system is capable of secreting large proteins. The T1SS is commonly used due to its simplicity; a range of proteins (including ScFvs) have been secreted through the system with secretion efficiencies into the extracellular space varying from 1% to 90%, with yields of around 2mg L^{-1} reported (Fernández, 2004; Fernandez & de Lorenzo, 2001). As lipases are a natural substrate of the type I secretion system, particular success has arisen when expressing recombinant lipases (Chung et al., 2009). As there are so many proteins which rely on the T1SS for secretion and its use as a means to efflux toxic compounds, limited recombinant protein secretion can occur due to competitive exclusion –this is also true of TolC in the T2SS. Other negatives of this expression system

include the common occurrence of cytoplasmic accumulation of recombinant protein and the dependence of the T1SS on the growth phase and oxygen availability (Hahn & Specht, 2003).

The type II secretion system is also capable of secreting recombinant proteins. Secretion is directed by N-terminal signal peptides, which can be fused to recombinant proteins, as a range of native substrates exist a number of secretion signals can be utilised (Berlec & Štrukelj, 2013). Following secretion they are cleaved which is a benefit of the system when expressing recombinant protein. As explained previously the T2SS is a two stage secretion system, with different secretion steps for each membrane. Folded protein is secreted through the TAT and secreton secretion systems, but not the Sec or SRP systems. Many recombinant proteins will fold in the cytoplasm, therefore the only route for the export of these proteins by the T2SS is the TAT system which is slow and easily saturated (Georgiou & Segatori, 2005). However many recombinant proteins have been successfully secreted through the TAT system by fusion to the twin-arginine signal peptide, including ScFv (up to 4.4mg L^{-1} following periplasmic extraction), GFP (up to 1g L^{-1} following periplasmic extraction), human growth hormone and interleukin (Choi & Lee, 2004; Matos et al. 2014; Matos et al. 2012). The Sec and SRP systems are ineffective if protein folding occurs in the cytoplasm. The SRP pathway has been utilised to secrete human growth hormone and thioredoxin into the periplasm, by fusion with a DsbA signal peptide, with yields of 4mg L^{-1} observed following periplasmic extraction (Soares et al., 2003). A yield of 60 mg L^{-1} was achieved for GFP in small batch culture or 1g L^{-1} following fed batch fermentation to optical densities upwards of OD_{600}150 (Matos et al., 2012). Transport from the periplasm through the outer membrane via TolC is easily saturated; low recombinant protein translation rates must be used to overcome this, therefore yield is low (Mergulhão et al., 2005). However the majority of uses of the T2SS for recombinant protein secretion concern secretion to the periplasm, as explained previously this enables advantageous qualities such as correct protein folding and lack of proteolysis, however the cells must be lysed to obtain protein. If yield and protein quality are high this method if favourable in industry, however there is a clear case for extracellular protein secretion.

However the research in this thesis focuses on the FT3SS. The FT3SS has previously been modified for protein secretion with limited success -but great promise. A 35 kDa recombinant lipase PalB was secreted in small amounts in by a from a flagella CAPless (Δ*fliCD*) *E. coli* MKS12 mutant after 32 hours, however the paper does not eliminate cell lysis as the reason for the occurrence of PalB in the media or show that the protein was secreted by the modified FT3SS specifically (Narayanan et al.,

2010). A selection of recombinant amino acid motifs and genes were fused to the a secretion signal and terminator sequence and secreted from the same mutant strain (Majander et al., 2005). The largest secreted protein was a 43 kDa enolase and the highest yield was 12mg L^{-1} of 28 kDa Peb1. Another study showed that sea snail venom (25 amino acids long) could be secreted by the *Salmonella enterica* FT3SS when fused to a protein which is natively secreted by the FT3SS (Singer et al., 2012). Spider silk has also been secreted from the *Salmonella* injectisome T3SS. Three different spider silk proteins were fused to an N-terminal secretion tag and a chaperone (Widmaier et al., 2009). The largest protein secreted was 75kDa at a yield of 6.9 $nmol^{-1}$ L^{-1} Hour. The highest yield of recombinant protein secretion achieved was 1.8 mg L^{-1} $Hour^{-1}$. A range of proteins (GFP, human complement protein) have also been successfully secreted in *Salmonella* with the partial protein secretion tag (Végh et al., 2006; Dobó et al., 2010). The intricacies of these modes of recombinant FT3SS in terms of modifications to the FT3SS and use of secretion signals will be discussed in a later section, once the FT3SS has been introduced more thoroughly.

The autotransporter type V secretion system has potential for recombinant protein secretion applications; although this is more commonly related to cell surface expression of protein, rather than full secretion into the media. For example replacing the passenger domain of the autotransporter protein can result in surface display on the outside of the bacteria –this has value in vaccine development (Jong et al., 2012). Recombinant ScFv, lipase and toxins have all been displayed on the surface of *E. coli* following secretion through the TVSS (van Ulsen et al., 2014).

In this thesis the secretion of recombinant proteins via a modified FT3SS will be investigated. As described above, secretion of recombinant proteins has been achieved previously in a number of other secretion systems in addition to the FT3SS. However this secretion system was selected due to various regulatory and structural characteristics of the FT3SS. In the next section the secretion system will be detailed and reference will be given to why this secretion system was selected to form the basis of the secretion platform. A strategy for modification of this protein secretion system to generate a secretion apparatus for protein expression will also be given.

1.4. The bacterial flagellum

The bacterial flagella is an organelle which allows cell motility in many Gram negative bacteria (Wang et al., 2012). Each flagella is a helical filament which is driven by an ion flux powered rotary motor (Berg, 2003). On average *E. coli* have 4-10 flagellum –which can be around 20µm in length (Lino, 1974). Movement is directional and driven by a sensory system which dictates whether the motor spins clockwise or anticlockwise. With clockwise flagella rotation, *E. coli* conduct tumble or a random search moving 30 diameters per second in one direction – then the other (Adler, 1975). Receptors on the cell surface detect chemical stimulants, such as sugars and amino acids –and chemical deterrents. In this event the motor will switch to anticlockwise rotation and steady one-directional movement known as a run is achieved, this causes the individual flagella to assemble into a bundle and allows the bacteria to move to a more beneficial environment. The bacterial flagella is the product of billions of years of evolution – it is comprised of over 50 proteins, which assemble in an orderly fashion, controlled by checkpoints (Chilcott & Hughes, 2000). Proteins assemble initially at the inner membrane and proceed to assemble to the cell exterior in a distal to proximal fashion. The majority of flagella proteins are secreted as unfolded subunits through the existing central channel and proceed to the distal end of the existing flagella structure for assembly beneath various FT3SS cap foldases.

1.4.1. Structure and assembly

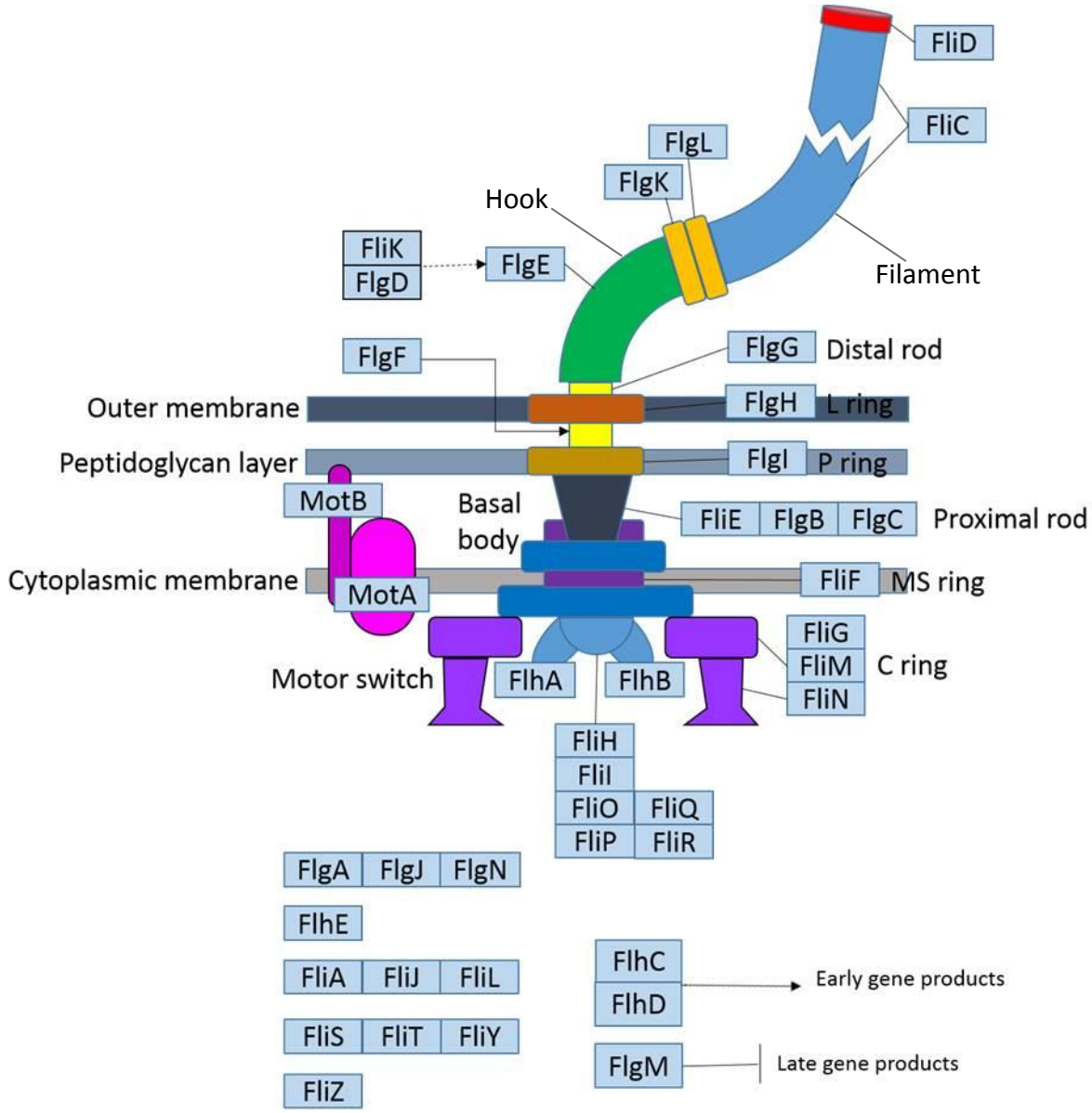


Figure 1.4: The distribution of proteins in the flagella type II secretion system.

Adapted from (Liu & Ochman, 2007)

This section will give an overview of the mechanism of assembly and structure of the FT3SS (Figure 1.4), for specific protein name and function see Appendix 1. The basic structure of the flagella is of a basal body, hook and filament. The basal body acts as the motor and the secretion apparatus, the hook as a universal joint and the filament as a propeller. Flagella assembly initiates at the MS ring,

which is comprised of FliF subunits. The MS ring is a component of the basal body which houses the T3SS element of the bacterial flagella. FlhB, FliO, FliP, FlhA, FliQ, and FliR are integral membrane proteins which assemble in the vicinity of the MS ring and either directly or indirectly all six interact with each other, forming the export gate. The latter three associate with the MS ring directly, anchoring the export gate in the central pore of the MS ring (Minamino, 2014). Once assembled these form a pore of about 2.0nm diameter, through which the majority of the remaining flagella proteins are exported (Yonekura et al., 2003). Once *in situ* the C ring proteins FliG, FliM and FliN then attach to the MS ring and cytoplasmic domain of FlhA protrudes into this cavity formed within the C ring (Kawamoto et al. 2013). The cytoplasmic domain of FlhA forms the binding site for ATPase, chaperones and export substrates. FlhB is also located in the export apparatus and has a membrane and cytoplasmic domain connected by a flexible linker. The C ring proteins form the rotor/switch complex, which contribute to control of rotation direction. Next the motor proteins MotA and MotB, which form the $MotA_4MotB_2$ complex attach to the C ring via MotA. This complex forms the stator and also the proton channel. Prior to assembly unbound MotB blocks proton flow through the proton channel, however when MotA collides with the basal body motor (FliG), a conformational change removes the block and proton flow is initiated (Hosking et al. 2006). This allows proton transport down their electrochemical potential gradient through the cell membrane, providing energy (known as proton motive force) which drives both secretion of flagella substrates and rotation of the flagella motor (Hosking et al. 2006) Recently sodium motive force was also shown to drive secretion (Minamino et al. 2016). T3SS of the rod, hook and filament proteins occurs through the six protein export gate described above, with the aid of the soluble FliH, FliI and FliJ proteins, which form the ATPase complex and at least partially unfold substrates prior to secretion. The proximal and distal rod proteins (FliE, FlgB, FlgC, FlgF, FlgG) are secreted through the T3SS in succession and assemble to span the peptidoglycan and outer membrane. Rod proteins assemble under the rod cap foldase FlgJ which and has dual function as it is also a β-N-acetylglucosaminidase, which degrades the peptidoglycan layer, to allow export of the rod proteins into the periplasmic space (Nambu et al., 1999; Herlihey et al., 2014). Formation of the P and L ring proteins which assemble in the peptidoglycan and outer membrane respectively, are dependent on the rod proteins being in place beforehand. The P and L rings surround the other T3SS transmembrane proteins, providing a channel for both assembly and rotation (Chaban et al., 2015).

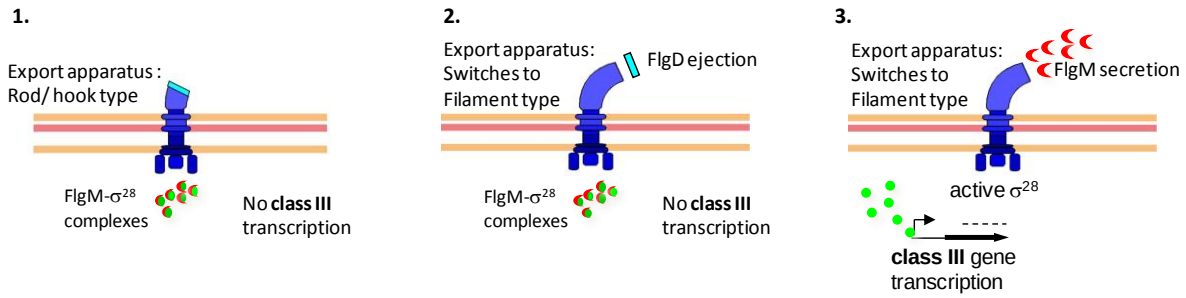


Figure 1.5: The switch from class II to class III flagella gene expression.

The hook grows under the FlhD cap. FlgM binds to σ^{28}, inhibiting it from initiating transcription of class III genes 2) Following completion of the hook, the hook cap is ejected 3) FlgM is secreted, σ^{28} is uninhibited and class III gene transcription is initiated

---

The hook is then assembled under the FlgD cap, from FlgE subunits to a length of 55nm, under the control of the hook length regulator protein FliK to form the universal joint (Erhardt et al. 2010). Following the completion of the hook, a signal is sent from the hook length control protein FliK to the membrane export component FlhB, which results in a conformational change -switching export substrate specificity from early to late substrate, allowing secretion of the hook-junction and filament proteins (Williams et al., 1996; Evans et al., 2014). This change is brought about because following FlhB auto cleavage a region is opened which is comprised of a basic patch formed by the α-helix and β-sheets which is adjacent to the NPTH loop. It is thought that an acidic loop of FliK is able to 'wedge' into the cleft formed by the basic patch -this drives the substrate specificity switch (Mizuno et al., 2011). This is critical for the secretion of late substrates, as prevention of auto cleavage maintains the secretion apparatus early locked (Fraser et al., 2003), suggesting that FlhB is essential for recognition of export substrates. Current theory suggests that FliK is intermittently secreted through the hook structure, immediately triggering the switch from early substrate (FlgE included) secretion to late when the hook reaches its physiological length of 55nm. While the hook extends at a constant rate, secretion of FliK is inversely proportional to this and it is thought that this low rate of secretion towards the end of hook extension, results in the substrate specificity switch as FliK is more likely to interact with FlhB (Erhardt et al., 2011). In synchrony following completion of the hook, secretion of the hook filament junction protein, FlgK displaces the FlgD cap from the hook tip (Ohnishi et al., 1994). This permits secretion of the anti σ^{28} factor FlgM, resulting in uninhibited intracellular FliA (Hughes et al., 1993). Prior to this FlgM binds σ^{28} factor –FliA (Figure 1.5), inhibiting

class III gene expression (and also protecting in FliA from proteolysis) (Barembruch & Hengge, 2007). FliA initiates class III based transcription of proteins including the hook associated proteins (HAPs) FlgK, FlgL and FliD and the 52kDa flagellin protein FliC (Daniell et al., 2003). Secretion of the late subunits is dependent on chaperones FlgN (for FlgK and FlgL), FliT (for FliD) and FliS (for FliC). The hook-filament junction proteins (FlgK and FlgL) assemble on the hook and finally the filament grows as flagellin monomers are transported from the cytoplasm through the central channel and then to the distal end (Yonekura et al., 2000). Flagellin monomers assemble under the FliD cap which rotates as flagellin units assemble and aids polymerisation of the flagellin units. 20,000-30,000 flagellin monomers self-assemble in one flagella at a rate up to 1000 per minute to form a filament 10-20µm long, with a diameter of 120-240nm (Erhardt et al. 2010).

Efficient secretion through the FT3SS is accomplished by the presence of chaperones, secretion signals, ATPases, proton and sodium motive force and subunit interaction. Together the export gate including the protruding FlhA cytoplasmic domain -which binds secretion substrates, chaperones and ATPase- form the ‘export cage'. (Evans et al., 2014). FlhB is also located in the export apparatus and has a membrane and cytoplasmic domain connected by a flexible linker. FlhB auto cleavage occurs at the NPTH loop the cytoplasmic domain. The cytoplasmic domain remains associated with the membrane domain and this flexibility between the two substrates allows secretion to occur (Minamino & Macnab, 2000). Situated below the export cage is the ATPase complex (comprised of FliI, FliH and FliJ). The ATPase regulator FliH interacts with the ATPase FliI; anchoring it to the C-ring and preventing ATPase activity and interaction with the export cage until the secretion structure is constituted (Minamino & Macnab, 2000). To initiate subunit export, FliJ is required to span the gap between FliH-FliI and FlhA (Abrusci et al., 2013).

As described, the flagella assembles in a sequential manner; during secretion in the early phase, unchaperoned early substrates interact with the ATPase FliI and its regulator FliH (Auvray et al. 2002; Thomas et al. 2004). Following docking at FliI, substrates are sorted for early or late export by the N-terminal secretion signal residues of substrates (Stafford et al. 2007). FliI is also thought to mediate partial unfolding of the substrate (Akeda & Galán, 2005). Along with FlhA, early substrates also interact with FlhB prior to export. Autocleaved FlhB has a hydrophobic pocket in the cytoplasmic region; the presence of this is essential for the secretion of rod and hook proteins. It is though that this is a binding site for N-terminal gate recognition motifs (GRM) of early secretion substrates (Evans et al., 2013).

Late substrate soluble components are pre-bound to their respective chaperones at the C-terminus (Bennett & Hughes, 2000). Chaperone binding prevents polymerisation of monomeric substrates prior to export and degradation by proteolytic enzymes (Fraser et al., 1999). The chaperones target chaperone-substrate complexes towards the export machinery, where the chaperone N-terminals associate with the ATPase (FliI), which is present as a FliI-FliH complex, this complex directs the N-terminal of the substrate to FlhA, in unison the FliJ drives the N-terminal of chaperones to interact with the cytoplasmic domain of FlhA (Thomas et al. 2004; Kinoshita et al. 2013; Minamino & Namba, 2008). FliJ interacts with FlhA and has affinity for the FlgKL and FliD chaperones FlgN and FliT, but not the FliC chaperone FliS suggesting that FliJ enables efficient export of the hook associated proteins prior to the secretion of FliC (Evans et al. 2006). It is thought that rotation of FliJ may drive conformational change in FlhA, which results in release of the subunit by its cognate chaperone and further unfolding (Abrusci et al., 2013).

The chaperones direct late substrates for export, prior to release, but they do not serve as signals for secretion. Disordered N-terminus domains serve as secretion signals for protein export (Kuwajima et al., 1989). No defined cleavable signal peptide sequence is conserved in T3SS export proteins, therefore it is thought that a recognition is achieved at a higher level in the disordered N-terminus domains (Homma et al., 1990; Vonderviszt et al., 1992). Despite this, studies have previously identified the peptide regions responsible for directed secretion by investigating the secretion of protein with N-terminal residue deletion mutants. For example it was found that only residues 26-47 of FliC were required for secretion, furthermore fusing these residues to a homologous protein directed it for FT3SS (Végh et al., 2006). There is also much debate over whether the 5’ untranslated region of secretion substrate mRNA is required as a secretion signal, while it is certainly capable of driving secretion it is not essential (Cheng & Schneewind, 2000; Majander et al. 2005; Dobó et al. 2010). As previously explained early subunits harbour GRMs which have affinity for the export gate protein FlhB (Evans et al., 2013).

Secretion through the FT3SS and torque generation of the flagella is powered by ATP and proton and sodium motive force. This is under tight regulation as it is costly. As described the FliH inhibits premature ATP expenditure by the ATPase complex and MotB blocks proton flow through the proton channel (Minamino & Macnab, 2000; Hosking et al. 2006). The ATPase complex facilitates substrate association with the export gate, however the ATPase is not essential for protein export,

although it is thought that following interaction with the substrate-chaperone complex, ATP energy is required to release the substrate (Minamino & Namba, 2008). It is now known that the PMF also provides energy for secretion through the FT3SS; this is comprised of two components: the electrochemical potential difference (ΔΨ) and the proton gradient (ΔpH) (Galperin et al. 1982). It was shown that in the absence of the ATPase complex, low efficiency ΔpH transport occurs, however when the FliJ component of the ATPase complex interacts with FlhA high capacity ΔΨ and ΔpH driven secretion is enabled (Erhardt et al., 2014). There is also evidence that FlhA acts as a channel for sodium, thus facilitating a sodium motive force. While this is usually inactive, in the absence of the ATPase complex it is activated –this suggests a dual powered system, which is effective in the absence or presence of ATPase (Minamino et al. 2016). Once subunits are secreted they must move through the existing structure before they are crystallised at the tip, as the filament can reach lengths of 20μm –as the structure elongates the rate of protein secretion remains constant (Turner et al., 2012). This constant rate of secretion is driven by head to tail linkage of both FliC and FlgE protein subunits, which results in a chain (Evans et al., 2013). This chain is formed by linkage of the N- and C-terminals of adjacent subunits; this enables subunit transit through the existing flagella structure by a pulling force which is generated as the N-terminus of the foremost subunit shortens during crystallisation at the tip -this results in subunits being pulled from the export apparatus into the existing flagella structure.

As stated previously, with modifications there is scope to secrete recombinant proteins through a modified structure of the FT3SS. Not only would this enable one step secretion of protein from the cytoplasm to the extracellular space, but if the secretion capacity could be fully harnessed, then the yield of secreted recombinant proteins could be of industrial biotechnological relevance. If this were achieved this work would serve as a pertinent example of a top down synthetic biology approach to achieving recombinant protein secretion in an *E. coli* chassis. To enable this, knowledge was drawn on from the literature of the genetic regulation of FT3SS structure and previous flagella mutants which enabled monomeric secretion of subunits through the FT3SS and into the media.

1.4.2. Comparison to injectisome T3SS

There are many similarities between the Injectisome and flagella T3SS. There are many homologous proteins and both assemble from the base to the tip (Erhardt et al. 2010). The majority of major structures are in place in both, however there are some differences. In the IT3SS the MS ring is

comprised of secretins -it is also void of a motor. The export apparatus is situated within the MS ring of both structures however the injectisome lacks a robust C-ring –although a homologue SpaO is present and this forms the sorting platform along with OrgA and OrgB (Lara-Tejero et al., 2011). Injectisome gene expression is less stepwise in regulation, as all components are transcribed from class II operons. However there are still checkpoints in terms of assembly, due to the affinity of substrates and chaperones for the secretion apparatus. Chaperones protect substrates from premature polymerisation and degradation and direct substrates for export at the export apparatus in both systems and this is facilitated by the ATPase and proton motive force. The IT3SS ATPase InvC is analogous to FliI and also unfolds protein before transit which is aided by OrgA and OrgB, the former is homologous to FliH, however the latter has no known homologue in the FT3SS (Akeda & Galán, 2005). However while FT3SS chaperones bind to the C-terminal of substrates, IT3SS chaperones bind to the N-terminal, however N-terminal signal peptides are present in both systems (Kosarewicz et al., 2012). If injectisome chaperones are removed this can lead to IT3SS substrates being exported through the FT3SS, suggesting homology of the N-terminal secretion signal and T3SS and also demonstrating the importance of chaperones (Lee & Galán, 2004). The main differences arise in terms of the extracellular structure and function, however subunits are still analogous. The needle (PrgI) and hook proteins are analogous and the length of both are controlled by an accessory protein (InvJ in the injectisome), resulting in either an 80nm needle or 55nm hook. InvJ also drives a substrate specific switch following hook completion and late substrates are exported (Kawamoto et al. 2013; Wee & Hughes, 2015). Rather than the hook filament junction, a translocon platform is then assembled on the tip of the needle and upon contact with a host cell, this results in the secretion of two translocon proteins, followed by the secretion of effector proteins through the lumen of the existing structure into a host cell (Galán & Wolf-Watz, 2006). Whereas the hook-filament junction proteins assemble on the hook tip and filament proteins are secreted through the existing structure –rather than be released into the environment, they are polymerised under a cap protein. Both are high capacity secretion systems.

1.4.3. Flagella gene hierarchy

As production of flagella is an energetically expensive process it is under tight regulatory control. Flagella genes are assembled into a transcriptional hierarchy (Chilcott & Hughes, 2000). Three classes exist within the hierarchy; the class I operon *flhDC* is at the top of the hierarchy (Figure 1.6). The class I operon product is the $FlhD_4C_2$ complex which activates the class II genes which encode

the flagella hook basal body (HBB) proteins and some regulatory proteins, including sigma factor σ28 (FliA) and anti-sigma factor σ28 FlgM. FlgM suppresses FliA during hook basal body formation, upon completion FlgM is secreted into the media, allowing FliA to promote the transcription of the class III genes which encode the filament, motor and chemotaxis proteins (Kutsukake, 1994). *flhDC* expression is regulated by various proteins which are expressed in response to environmental cues including pH, temperature, nutrient availability and osmotic pressure (Soutourina & Bertin, 2003). These proteins may either directly or indirectly result in up- or down- transcriptional regulation of *flhDC* or act on the $FlhD_4C_2$ complex to either block it from activating gene expression of class II promoters or actively degrade the complex. Transcriptional regulation is mediated through multiple sites in the large upstream region of the *flhDC* promoter (Fahrner & Berg, 2015). Additionally feedback expression loops of flagella genes occur; where protein produced as a result of *flhDC* expression go on to repress or activate master regulator or class II or III protein expression (see Figure 1.6). These proteins are the protein chaperones FlgN, FliS and FliT. FlgN regulates FlgM, therefore affecting class III gene expression (Karlinsey et al. 2000). FliS negatively regulates the export of FlgM (Yokoseki et al., 1996). FliT negatively regulates class II gene expression by binding to $FlhD_4C_2$ (Yamamoto and Kutsukake, 2006). Alternatively proteins can directly digest the $FlhD_4C_2$ complex. The effects of these proteins on flagella gene expression are outlined in detail in Table 1.1 with reference to the phenotype of strains following mutagenesis. Along with flagella biogenesis $FlhD_4C_2$ also regulates promoters in a wider regulon. While it binds to a multitude of promoter regions it does not cause a large transcriptional effect (Stafford et al., 2005), nether-the-less it has been shown to repress cell division, upregulate galactose transport, downregulate aerobic respiration (Prüß et al. 2001; Prüss & Matsumura, 1996).

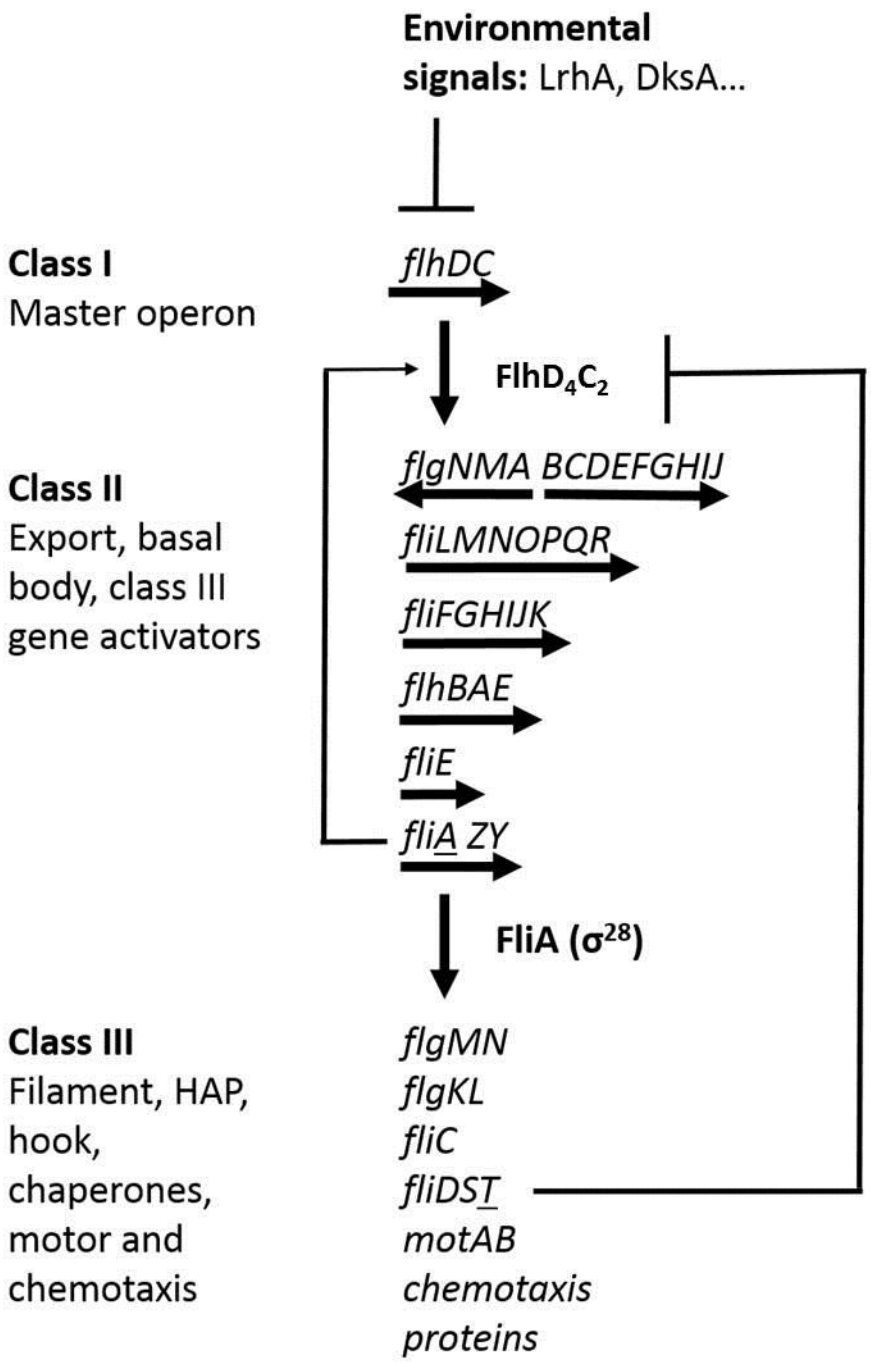


Figure 1.6: Flagella genes organised into operons within flagella regulatory classes.

Feedback inhibition is denoted with blunt ended arrows. Adapted from Claret & Hughes, 2002; Karlinsey et al. 2000.

The master regulator forms a point of fulcrum control over the FT3SS; therefore it is a prime point investigation into improved FT3SS performance with regards to protein secretion, whether this concerns transcriptional regulators and proteases which respond to environmental cues or negative feedback from within the flagella regulatory hierarchy. Previous experimental work will now be summarised which has 1) modified the FT3SS to initiate the secretion of recombinant proteins and 2) reported gene modifications which have resulted in increased master regulator activity. This will form the basis of experimental work in the next chapter, as combined it will aid the production of a strain which is capable of high capacity secretion of recombinant proteins through the FT3SS.

1.4.4 Previous FT3SS gene deletions that result in secretion of monomeric flagella subunits

Various gene deletions are outlined in the literature, which result in the secretion of monomeric flagellin into the extracellular media. This will be exploited for the development of a modified FT3SS which is able to secrete recombinant proteins into the media. As stated previously, bacterial FT3SS genes and function are highly conserved among gram negative bacteria, therefore deletions reported in the literature regarding various bacterial species are also applicable to *E. coli*. Deletions of the flagella, cap and hook-filament junction proteins produce immotile mutants as the flagella fail to polymerise, these are outlined below –with additional reference made to beneficial knockout mutations in terms of increasing FT3SS secretion.

As discussed the tip of the flagella is formed of the cap protein FliD. Flagellin subunits assemble under the cap. In FliD mutants the hook assembles, however polymerised FliC filaments do not form, instead FliC is secreted into the culture media as monomers (Yokoseki et al., 1995). FliD mutants are mainly immotile, however it has been shown that when supplemented with exogenous flagellin, filaments will polymerise on the hook and some motility is restored. *E. coli* which lack FliC, do not have filaments and are essentially immotile (Li et al., 1993). Class II promoter activity is seen to increase in Δ*fliD* (Brown et al., 2008). This is thought to be because *fliD* is in the same operon as *fliT*, which represses master regulator gene expression. Furthermore an increased number of flagella basal bodies were observed following the deletion of *fliCD*, this is thought to be due to the removal of secretion competitors to FlgM (Singer et al. 2012).

The hook-filament junction is absent in Δ*flgKL* bacteria. These mutants produce FliC however it does not form a filament, instead the unpolymerised FliC monomers are secreted into the culture media (Homma, 1984). Increased extracellular homologous protein (FlgM, FliD) has been observed in cells with *flgKL* deletions in comparison to wild type, suggesting that secretion of some protein subunits is higher in Δ*flgKL* (Brown et al., 2008).

1.4.5. FT3SS recombinant protein secretion

As described, the FT3SS provides a high capacity route of protein export from the cytoplasm to the extracellular space – the filament is comprised of up to 30,000 flagellin monomers (Erhardt et al. 2010). If generation time is assumed to be 30 minutes, this amounts to an export rate of approximately 1000 subunits per minute, which equates to approximately 0.5g L^{-1} $Hour^{-1}$ -a rate which rivals that achieved by secretion efficient yeast strains such as *Pichia Pastoris* or secretion competent *Bacilus* genus (Love et al., 2012; Chen et al., 2012; Liu et al., 2013) However it has since been demonstrated experimentally that the rate of secretion is one monomer every two seconds, equating to 15 mg^{-1} $Hour^{-1}$ (Turner et al. 2012)–while this is less impressive than the previous estimation, it is still an remarkable rate. This capacity can be exploited for high throughput secretion of recombinant proteins. In addition, well understood natural regulatory systems are in place for flagella gene expression control and lend themselves to manipulation and control. Specifically, they are under direct or indirect control of the master regulator complex $FlhD_4C_2$, which itself is transcriptionally regulated by proteins in response to environmental cues. Various attempts have been made to convert the FT3SS into a high capacity secretion system. These efforts mostly rely on the fact that removal of filament and cap proteins results in a modified FT3SS that is capable of secreting both native and recombinant proteins (Majander et al. 2005; Narayanan et al. 2010; Singer et al. 2012).

This project aims to optimise an *E. coli* strain for directed secretion of proteins from the cytoplasm directly to the extracellular environment, through a modified FT3SS. The flagella secretion system will serve as a platform technology for the secretion of a range of proteins. Strain optimisation by knockout mutagenesis aims to improve secretion output by altering regulatory gene expression and also removing superfluous genes to reduce metabolic burden. Assays will be developed to quantify secretion output of strains; this will be coupled with an assay to measure flagella gene expression.

Initially native protein secretion through the modified secretion apparatus will be investigated and an SDS-PAGE based protein secretion assay will be developed. Once this is established directed recombinant protein secretion will be implemented with the aid of a secretion construct, which will harbour recombinant cargo protein and is also comprised of a secretion signal and tags for purification. The secretion construct will be modular and therefore allow alteration of the cargo protein –and later modification of the secretion construct for improved secretion. This will be employed to secrete a recombinant enzyme which will enable a high throughput enzyme based secretion assay to be developed. This assay will allow efficient and accurate screening of a multitude of strains and secretion plasmids to identify the most productive combination in terms of secretion.

This work is in the context of previous attempts to modify the FT3SS to secrete recombinant protein. In section 1.3.6 yields of secreted recombinant protein through the FT3SS were reported; here I will outline how they were implemented, along with additional examples.

ΔfliCD mutants were successfully utilised to secrete a number or recombinant proteins (including enolase, GFP, the Peb1 adhesin, and PalB lipase) through the FT3SS into the media. This was enabled by fusing recombinant protein to the 5' and 3' untranslated regions of the FliC gene (Majander et al., 2005; Narayanan et al., 2010). This background was also used to secrete FlgM secretion signal tagged recombinant toxin proteins; while secretion was achieved without this mutation, it was greatly improved with it (Singer et al., 2012).

ΔfliC mutants were utilised to investigate secretion of FliC 47 amino acid secretion signal peptide fused recombinant proteins in the absence of the UTRs. Here it was concluded that the 5' UTR was not necessary to secrete protein through the FT3SS. In addition to this it was established that only of residues 26-47 of the secretion signal were required to incur secretion (Dobó et al., 2010; Végh et al., 2006).

The literature here has focused on secretion of late secretion substrates in mutants with intact hooks, however it is also possible to bypass this and secrete protein linked to early secretion substrates in a hook mutant, this forgoes the control of FlgM on protein secretion. This has been achieved by fusing recombinant protein to either FlgD or FlgE secretion signals or the FlgD 5' UTR in a Δ*flgD* or Δ*flgE* background (Singer et al., 2014)

1.4.6. Master regulator effects of FT3SS expression

Once the FT3SS has been successfully modified to allow recombinant protein secretion, increasing the number of basal flagella systems in the organism will allow more secretion into the culture media. Some gene deletions may increase the number of flagella present. The main route to enabling this is to reduce the prevalence of negative regulators of the master regulator and increase activators. An alternative is to remove proteins which actively degrade the $FlhD_4C_2$ complex or inhibit the complex from inducing class II gene expression. If this is achieved the increased abundance of $FlhD_4C_2$ will directly or indirectly promote expression of all remaining flagella genes in the hierarchy. Motility of cells is commonly used to ascertain flagella gene expression, as the output of increased flagella gene expression in cells which produce full length filaments is generally an increase in motility. Previous gene deletions which resulted in increased flagella expression (or motility) as a result of either increased *flhDC* expression or decreased $FlhD_4C_2$ complex proteolysis or binding are compiled in Table 1.1. It is hoped that these can be later implemented to improve secretion capacity of the modified FT3SS. Likewise gene deletions which resulted in reduced flagella gene expression are also included, to give a more complete picture of regulation and because these could also be utilised later -in promoter replacement strategies for example. This table only concerns flagella synthesis and does not account for the wider regulatory effects which $FlhD_4C_2$ initiates transcription throughout the genome –which included protein folding machinery, polysaccharide synthesis, membrane transport (Stafford et al., 2005; Fitzgerald et al., 2014). It should be noted that whereas previously literature describing assembly and regulation of the FT3SS was drawn on for both *Salmonella* and *E. coli* due to similarities, environmental cue determinants of *flhDC* expression vary in the two species. This is because in low nutrient availability *E. coli* flagella gene expression is upregulated, whereas in *Salmonella* it is down regulated, therefore care must be taken when interpreting data, depending on which species was investigated and whether nutrient availability impacted on results (Wada et al., 2012). For example in *Salmonella* YdiV inhibits $FlhD_4C_2$ class II gene expression in low nutrient concentrations. This can be reversed by upregulating the flagella protein FliZ, as it represses *ydiV.* However the *E. coli ydiV* is non-functional, therefore neither of these effects are observed in *E. coli* (Wada et al., 2012) –that aside FliZ does negatively regulate *flhDC* expression directly in *E. coli* (Pesavento & Hengge, 2012). Literature referring to environmental regulators of *Salmonella* was excluded for this summary (for example RflM and EcnR (Wozniak et al., 2009; Singer et al., 2013)), with the exception of information relating to FliT and Dnak as they do not concern environmental regulation of *flhDC.*

Table 1.1: Gene knockouts and reported flagella gene and phenotypic effects

Gene	Organism	Function in WT	Reported gene and phenotypic effects in mutant	Reference
A. Knockout mutations result in increased flagella gene expression or motility				
fliT	*Salmonella enterica*	Prevents $FlhD_4C_2$ from binding to class II promoters, also releases DNA-bound $FlhD_4C_2$ which results in increased ClpXP proteolysis	Higher class II gene expression and increased motility, despite less $FlhD_4C_2$	Yamamoto & Kutsukake, 2006
fliZ	*E. coli*	Binds to the *flhDC* promotor region	Increased *flhDC* expression in the *fliZ knockout mutant*	Pesavento & Hengge, 2012
lrhA	*E. coli*	Binds to *flhD*	3.5 reduction in FlhDC expression and hypermotile cells	Lehnen et al. 2002
dksA	*E. coli*	Enhances activity of ribosome synthesis molecules pppGpp and ppGpp, which reduce flagella gene expression	Increased *flhDC* expression (Lemke et al., 2009), hyperflagelated and hypermotile cells (Aberg et al., 2009) OR less motile cells (Magnusson et al., 2007)	(Lemke et al., 2009; Aberg et al., 2009; Magnusson et al., 2007)
clpXP	*E. coli*	Degrades $FlhD_4C_2$ protein (is a protease)	Hyperflagellated	Kitagawa et al. 2011; Tomoyasu et al. 2002
matA	*E. coli*	Represses *flhDC* transcription	Mutants were slightly more motile	Lehti et al. 2012

rscB	*E. coli* and *S.enterica*	Binds to the *flhDC* regulon and negatively regulates *flhDC*, positively regulates *fliPQR*	Decreased *flhDC* gene expression when over expressed in *E. coli,* increased motility in *Salmonella* knockout mutant	Wang et al. 2007; Francez-Charlot et al. 2004
hdfR	*E. coli*	Binds $FlhD_4C_2$	In H-NS deficient mutants. HdfR is negatively controlled by H-NS. Therefore H-NS mutants have more active HdfR and mutants are nonflagellated	Ko & Park, 2000
B. **Knockout mutations result in decreased flagella gene expression or motility**				
sirA, csrB & csrC	*E. coli*	csrA stabilises $FlhD_4C_2$ and increases its expression, csrB and csrC are antagonists of csrA. sirA activates csrBC translation	Reduced *flhDC* gene expression and motility and less *flhDC* mRNA in late exponential phase	Wei et al. 2001
DnaK	*Salmonella enterica*	This chaperone aids correct assembly of $FlhD_4C_2$ allowing it to interact with promoters	Lack flagella and class II and III gene expression	Takaya et al. 2006
CRP	*E. coli*	Increases *flhD* operon transcription in the absence of glucose	*crp* mutant was non-motile	Soutourina et al. 1999; Zhao et al. 2007
QseB	*E. coli*	Increases *flhD* operon transcription in the presence of FliA	Less flagellin and less transcription of *flhD, fliA, motA and fliC*	Sperandio et al. 2002
H-NS	*E. coli*	Increases *flhD* operon transcription when all H-NS binding sites are bound by H-NS	*hns* mutant was non-motile	Soutourina et al. 1999

With this information in mind it seems that there is great potential to modify the FT3SS into a secretion apparatus, for the secretion of recombinant protein. Secretion occurs from the cytoplasm directly to the extracellular environment, the system is high throughput and there is good control of secretion with the use of protein secretion signals. There is a good understanding of structure and genetic regulation of the FT3SS and the regulatory network is well suited for genetic manipulation. In addition, there are previous examples of recombinant protein secretion by this route, demonstrating that manipulations can result in positive results. Once an optimised protein secretion system has been produced, strain improvements and plasmid based improvements for increased secretion will be investigated and recombinant proteins will be expressed for directed secretion through the modified FT3SS.

1.5. Collagen

This project had the main aim of improving the flagella type III secretion system (FT3SS) for recombinant protein secretion, however the crowning aim was to secrete human collagen as an exemplar of this secretion platform. Collagen is the most abundant and widely distributed class of mammalian proteins, comprising of around one third of all protein, it constitutes three-quarters of dry skin weight and is the most abundant component of the extracellular matrix (Shoulders & Raines, 2009). Collagen describes a class of proteins which form supercoiled triple α helical supramolecular structures. Collagen proteins are comprised of a combination of 46 distinct polypeptide chains which trimerise homo- or hetero-trimerically to form the 28 distinct collagen types, all of which vary in size, function and distribution. While the different collagen types have different structures and properties all are structural fibrous proteins which are usually located in the extracellular matrix. Collagen most commonly serves as a biomechanical scaffold, giving tensile strength, allowing cell attachment and macromolecule anchorage, therefore defining the form of associated tissues (Kadler et al., 1996). Collagen is also involved in cell migration, cancer, angiogenesis and tissue morphogenesis and repair (Kadler et al., 2007). As different collagen types have different properties, tissue types have different collagen type compositions. Type I collagen is the most abundant collagen and is found in all tissues except cartilage. For this reason it will be the focus of the work carried out in this thesis. As described below Type I collagen is extremely useful, however there are issues with deriving it from animal sources, therefore recombinant sources would be favourable.

1.5.1. Basic structure

The three polypeptide α chains which form a collagen triple helix molecule are about 1000 amino acids long and all conform to a strict repeating amino acid motif of $(Gly\text{-}X\text{-}Y)_n$ (Brodsky & Persikov, 2005). Most commonly the motif repeated is Gly-Pro-Hyp (Ramshaw et al., 1998). Each α chain has a pitch of 18 amino acids per turn, these assemble trimerically and are staggered relative to each other by one amino acid (Schulz & Bader, 2007). Collagens are not consistently triple helical molecules – all collagens have non-helical domains at the N- and C- termini. Furthermore only type I collagen is in absence of any imperfections in the triple helix. Other collagens have interruptions in the triple helix, this is more common in non-fibril forming collagens (Kadler et al., 2007).

1.5.2. Collagen biosynthesis

I will focus on fibril forming collagen formation – primarily of type I collagen, as they are the most abundant and the majority of literature also focuses on them. However it is likely that the basic mechanism of triple helix formation applies to all collagen types (Gelse et al., 2003).

Collagens are mainly produced in cells of connective tissues and secreted into the extracellular space, where they assemble into collagen fibrils. The steps of collagen biosynthesis are summarised in Figure 1.7. Collagen mRNA is transcribed in the nucleus, capping at the 5' end and polyadenylation at the 3' end occurs as may alternative splicing. mRNA is then transported to the cytoplasm and translated at the rough endoplasmic reticulum (RER) (Gelse et al., 2003). Prior to trimerisation extensive post-translational modifications occur in the endoplasmic reticulum. Single chain preprocollagen projects from the RER, following the removal of signal peptides for RER protrusion the molecule is known as procollagen. Some proline and lysine residues are hydroxylated in the presence of cofactors by prolyl 3-hydroxylase (P3H) and prolyl 4-hydroxylase (P4H) or lysyl 3-hydroxylase respectively (LH3). The hydroxylation of proline residues is essential to collagen thermostability. Around 100 proline residues are converted to hydroxyproline in each collagen α chain (Olsen, 2003). Some hydroxylysines are also O-linked glycosylated (Perdivara et al., 2013). The terminal pro-peptides have a single N-linked oligosaccharide, which is added by oligosaccharide transfer complex (Clark, 1979). The aforementioned PTMs are essential for the formation of intermolecular hydrogen bonds between α chains and later for intermolecular crosslinking and

carbohydrate attachment. The propeptide chains are flanked by N- and C- terminus telopeptide regions. Three procollagen α chains align at the C-terminal and aided by enzymes (peptidyl-prolyl *cis-trans*-isomerase, protein disulphide isomerase) and molecular chaperones (heat shock protein 47), trimerise and fold, progressing to the N terminal to form an α triple helix. These triple helical procollagen molecules are now packaged into secretory vesicles in the Golgi apparatus and secreted into the extracellular space. N- and C-metalloproteinases then cleave the procollagen molecule forming tropocollagen. Lysyl oxidase catalyses the formation of aldehydes from lysine and hydroxylysine residues of telopeptides.

Collagen fibril formation is largely a self-assembly process, driven by entropy (much like flagella). As described previously type I collagen is comprised of uninterrupted triple helical collagens, with non-helical flanking N- and C- terminus regions -which are on average 300nm length and 1.5nm in diameter. The fibrils orientate differently depending on the tissue, but type I tropocollagens align in parallel and display D periodicity of 67nm. As the length of tropocollagen monomers is not a multiple of D this suggests a regular pattern of gaps and overlaps in the collagen fibril (Shoulders & Raines, 2009). Once assembled fibrils may crosslink with other fibrils of the same or different collagen types or be decorated with fibril associated collagens.

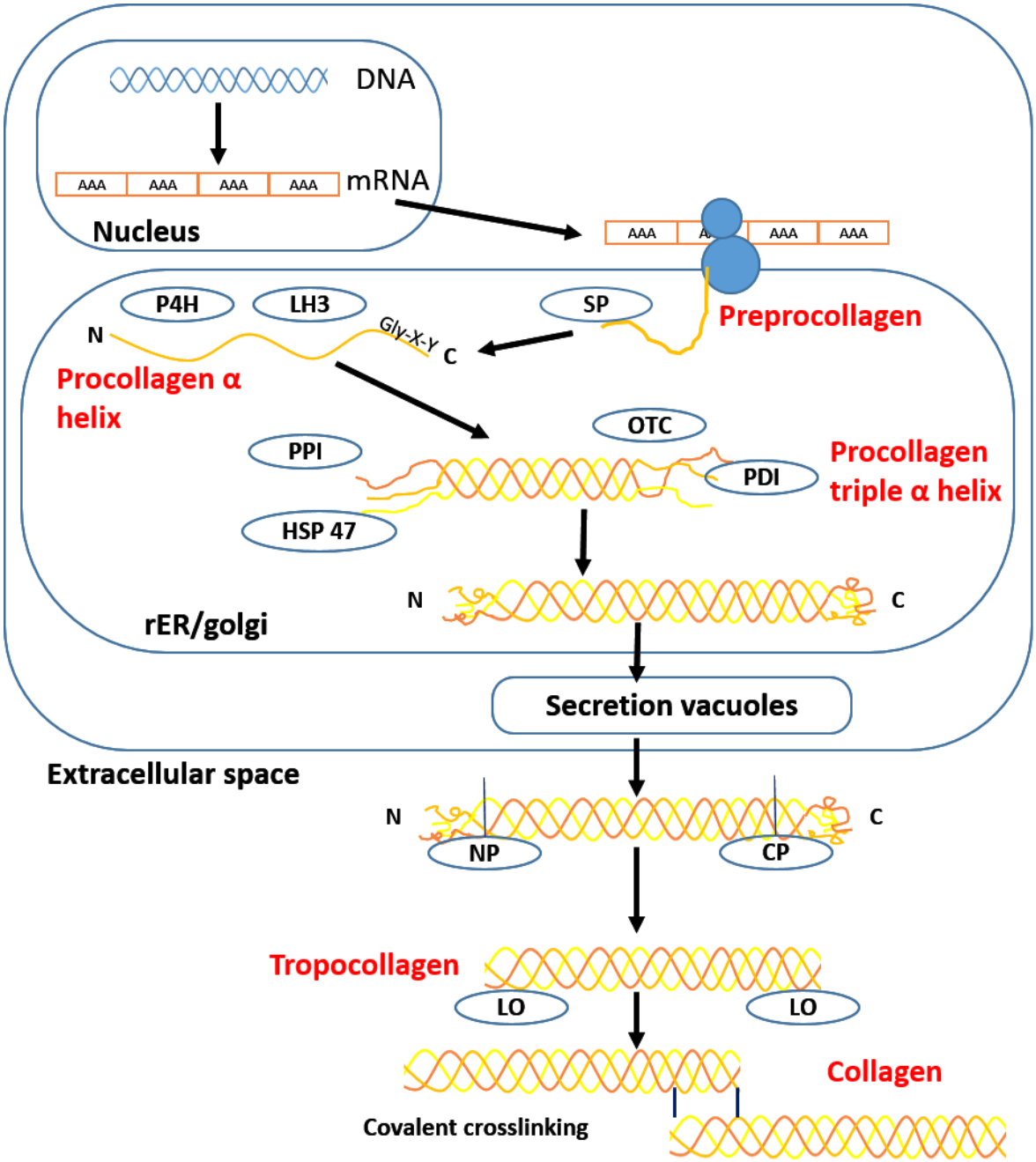


Figure 1.7: Collagen (fibril) biosynthesis in eukaryotic cells.

Including transcription, translation, post translational modifications, secretion and fibril formation. SP: signal peptidase, LH3: lysyl hydroxylase, P4H: prolyl 4-hydroxylase, OTC: oligosaccharyl transferase complex, PDI: protein disulphide isomerase, PPI: peptidyl-prolyl *cis-trans*-isomerase, NP : procollagen *N*-proteinase, CP: procollagen C-proteinase, LO: lysyl oxidase, HSP47: heat shock protein 47. Adapted from Gelse et al., 2003.

1.5.3. Therapeutic uses of collagen

Collagen is widely used as a medical material, due to good biocompatibility, biodegradability and low antigenicity. In addition the nature of collagen to form strongly associated fibres by self-aggregation and crosslinking, means that collagen has excellent strength. It also controls fluid loss and provides thermoregulation (Patino et al., 2002). Therapeutic collagen promotes cell integration; on administration exogenous collagen activates a cellular response, this is product specific but can involve the penetration of a range of cells including fibroblasts, macrophages and neutrophils. The degradation time of collagen products can be altered by crosslinking and pre-treatment procedures (Friess, 1998). These characteristics lend collagen for use in a large number of medical applications. Most commonly collagen is used as a drug delivery system, however collagen is increasingly being used as a biomaterial for tissue therapy (Lee et al. 2001). Collagen can be manipulated into many physical states including solutions gels, fibres, membranes, sponges and tubing for a range of medical procedures including dressing for burns and wounds following injury or surgery, as fillers and haemostatic agents. Collagen has also been used as a dressing for vascular prostheses, as a guide for peripheral nerve generation and in various ophthalmology applications (Patino et al., 2002; Bushnell et al., 2008). Additionally collagen can be used for the delivery of a range of drugs (Friess, 1998). For example collagen microspheres and gels have been utilised for the delivery of chemotherapy drugs (Panduranga Rao, 1996; Kojima et al. 2013). Type I collagen is useful as a scaffold for tissue engineering as it is biocompatible and also abundant and ubiquitous –therefore obtainable on a large scale and suitable for widespread use. In tissue engineering collagen is used as a scaffold in the form of a hydrogel or lattice. Although it can be used intact, removing the N- and C-telopeptides can reduce antigenicity (Glowacki & Mizuno, 2008). Additional crosslinking and copolymerisation of molecules such as chondroitin 6-sulphate may be favoured prior to use of collagen in tissue engineering, as this has been shown to increase cell proliferation within the scaffold and aid retention of proteoglycans –which are essential for cell signalling, migration and uptake of proteins (van Susante et al. 2001; Ferdous & Grande-Allen, 2007).

1.5.4. Sources of therapeutic collagen

1.5.4.1. Animal Sources

The global demand for collagen is huge, annually an estimated 50,000 tonnes of collagen (including gelatine) are derived from animal sources for use in medical applications (Olsen, 2003). Collagen is sourced from bovine skin, intestine and tendons (Patino et al., 2002). While animals may be a cheap source of collagenous material, there are many concerns surrounding various unfavourable characteristics of animal derived collagen. Disadvantageous qualities include poor biocompatibility, homogeneity, characterisation, immunogenicity and the transfer of disease from animal to patient. The main point of concern regarding disease and animal derived collagen is bovine spongiform encephalopathy (BSE). Although ante-mortem BSE tests are available they are not very reliable, therefore the risk of BSE in collagen therapies can only be managed and reduced, it cannot be guaranteed (Lasch et al., 2003). This is achieved by: (1) sourcing bovine collagen from certified 'BSE free' herds (2) using material (usually tendons) which carry a lower risk of infectivity (3) carrying out risk reducing procedures in the manufacturing process, such as alkali treatment (4) extracting collagen from tissue which has a high collagen content (again tendon) therefore reducing the quantity of tissue used (5) limiting the number of collagen treatments to a patient (6) only using certain routes of delivery. With these procedures in place the risk of contracting BSE from bovine sourced collagen products is reduced to less than one in a million (Friess, 1998). Immunological responses to collagen are variable and are dependent on the route of administration and nature of the collagen based therapeutic, for example 2.8% patients treated with an injectable collagen filler developed inflammation and 6.8% people treated with a collagen bone filler developed allergies (Charriere et al., 1989; Lynn et al., 2004). Preparing animal collagen for clinical applications is a complicated process. The presence of covalent crosslinks between collagen molecules is the major factor which impedes the breakdown of collagen when preparing it for clinical use (Friess, 1998). Additionally collagen is insoluble in organic solvents making downstream processing difficult. Once non-crosslinked, collagen molecules may be modified to be acylated, methylated or have polymers attached to give different properties. Crosslinks are then reintroduced to confer the structure required. This is achieved by aldehyde treatment or by physical treatment with UV or heat. Collagen is then sterilised by irradiation or ethylene oxide.

1.5.4.2. Recombinant sources of human collagen

Recombinant collagen has been produced in a number of organisms (Table 1.3). The major barrier to recombinant collagen is the lack of adequate post translational modifications in various recombinant expression systems. Without PTMs such as glycosylation and hydroxylation, collagen structure and function varies. In some recombinant collagen expression systems collagen genes are co-expressed with enzymes for proline hydroxylation. The key enzyme for this process is prolyl-4-hydroxylase (P4H). Co-expression of collagen with human P4H has shown to improve proline hydroxylation, although not necessarily to the degree observed in native collagens. Alternatively inadequate proline hydroxylation can be avoided by expressing collagen in a mammalian cell line which natively expresses the enzyme. Plant cells have been also been shown to express low levels of endogenous PH4 but it is not high enough to sufficiently hydroxylate proline. Lysyl-hydroxylase-3 (LH3) is important for lysine hydroxylation, however it is less commonly coexpressed in cells. Collagen is able to trimerise without the mammalian folding chaperone HSP47 (Vuorela et al., 1997). Animal model experiments have been carried out using recombinant collagen type VII from CHO cells and show promising results, although this collagen is glycosylated in a mammalian way (Hou et al., 2015). Another study investigated the effect of prolyl-4-hydroxylase expressing *Pichia* derived recombinant collagen III procollagen, which was matured to collagen III *in vitro.* When applied to wounds in a pig model, healing was enhanced (Nuutila et al., 2015). This demonstrates demand and increased acceptability of these treatments.

Table 1.2: Recombinant human collagen production.

Collagen type, product and titre are listed. Co-expression of enzymes and posttranslational modification are noted, along with the occurrence of extracellular secretion. rP4H denotes co expression with recombinant prolyl-4-hydroxylase, rLH3 with recombinant lysyl-hydorxylase-3. Note: Only recombinant human protein listed, recombinant C and N collagen telopeptides are not included in this table.

Cell	Collagen type	Collagen product	Product titre	rP4H?	rLH3?	Post translational modifications?	Extracellular secretion?	Reference
Insect cell/ baculovirus	III	Full length triple helical collagen	40mg/L media	Y	N	Proline hydroxylation	N	Helaakoski, 1996
Human embryonic kidney cell	VI	Full length α1(VI) and α2(VI) chains (would not form collagen trimers)	50-150µL/ml media	N	N	50% successful proline and lysine hydroxylation	Y	Tillet et al. 1994
Pichia pastoris	III	Full length triple helixes	~15mg/L	Y	N	42% successful proline hydroxylation	N	Vuorela et al. 1997
Pichia pastoris	I	Full length triple helixes with pepsin treatment (no N-propeptide)	200-500mg/L	Y	N	Suggests 100% successful proline hydroxylation	N	Nokelainen et al. 2001

Human embryonic kidney cell	V	Full length triple helix	15µg/ml	N	N	Some lysine hydroxylation, glycosylation	Y	Fichard et al. 1997
Insect cell/ baculovirus	II	Full length triple helix	50mg/L	Y	Y	Lysine hydroxylation, glycosylation	15% secreted	Nokelainen et al. 1998
Insect cell/ baculovirus	II	Information not available, but product formed collagen fibrils	Not reported	Y	N	Lysine hydroxylation, glycosylation	N	Notbohm, 1999
HEK	XVI	Full length triple helix	80-100µL/L media	N	N	N-glycosylation, proline and lysine hydroxylation	Y	Kassner et al. 2004
HEK	VII	Full length triple helix	2-5mg/L	N	N		Y	Chen et al. 2002
Saccharomyces cerevisiae	I	Full length triple helix	3-4µg/mg total protein	Y	N	Proline hydroxylation	N	Toman et al. 2000
CHO	IV	Full length α1(IV) and α2(IV) chains	Not reported	N	N		Y	Fukuda et al. 1997
Corn	I	Full length α1(I) chain	120mg/kg corn germ	N	N	Low level of lysine hydroxylation	N	Zhang et al. 2009

Barley	I	Full length and 45kDa α1(I) chain	140mg/kg seed	N	N	Very low level of proline hydroxylation	N	Eskelin et al. 2009
Maize	I	Full length α1(I) chain	4mg/kg	Y	N	High level of proline hydroxylation	N	Xu et al. 2011
Tobacco plant	I	Full length triple helix	20mg procollagen/ kg leaves	Y	Y	Some proline and lysine hydroxylation	N	Stein et al. 2009
Pichia pastoris	III	Full length α1(III) chain	Not reported	Y	N	Proline hydroxylation	Not reported	Xu et al. 2015
Pichia pastoris	I	Full length α1(I) chain	89mg from 30g cells	N	N	33% proline residues hydroxylated	Not reported	Setina et al. 2016
E. coli	I	Fragment of α1(I) chain (193 residues = around 23kDa)	10% total cell protein	Y	N	'High'	Not reported	Buechter et al. 2003
E. coli	III	38kDa fragment of αI (III) chain	90mg/Litre cell culture	Y	Y	26% proline and lysine residues hydroxylated	Not reported	Rutschmann et al. 2014

Heterologous collagen expression has been demonstrated in a range of expression systems, however there are common issues. Secreted collagen is rarely observed, therefore cells must be lysed or in the case of plant material ground to extract collagen. This is costly and time consuming.

1.5.5. Recombinant collagen production in *E. coli*

Many bacteria express collagen –like proteins which have the same Gly-Xaa-Yaa repeating motif of residues. These have been shown to form trimers when expressed in native or recombinant (i.e. *E. coli*) hosts. These could have technological applications themselves (An et al., 2016), but also suggests the bacteria have a natural affinity for producing these kinds of proteins (Yu et al., 2014). There are no reports of successful expression of recombinant full length human collagen in *E. coli.* Attempts to produce recombinant collagen in *E. coli* have only amounted to the production of human like collagen. This may be synthetically derived repeats of the triple motif commonly seen in collagen (Goldberg et al., 1989). Alternatively this may be expressed from the cDNA of fragments of human collagen mRNA. This type of collagen forms triple helixes similar to human collagen and has been shown to have low immunogenicity and is used in some therapeutic technologies (Xu et al. 2012). Expression of this kind of human like collagen requires extremely high cell density culture (~OD_{600} 90). Fragments of human type 1 and III collagen α chains have also been expressed intracellularly in in *E. coli* (Buechter et al., 2003; Rutschmann et al., 2014).

1.5.5.1. Posttranslational modifications

As discussed, post translational modifications, particularly hydroxylation are important in collagen biosynthesis. This requires enzymes which are not natively expressed by E. coli. The expression of recombinant P4H in *E. coli* is well reported. *E coli* derived recombinant P4H has been shown to hydroxylate proline residues of collagenous protein in vitro (Pinkas et al., 2011; Neubauer et al., 2005). Recombinant LH3 expression was reported to result in hydroxylation of

lysine residues of a collagen fragment in *E. coli* (Rutschmann et al., 2014). O- and N- linked glycosylation is also important in collagen biosynthesis. As explained previously pathways for O- and N-linked glycosylation can be recombinantly engineered into *E. col*i (Wacker et al. 2002; Lubas, 2000; Henderson et al. 2011).

1.5.5.2. Proline metabolism

Collagen is a proline rich protein. *E. coli* are able to produce proline autotrophically , however production is limited by a negative feedback loop of proline on glutamate-5-kinase (Adams & Frank, 1980). This lack of intracellular proline is likely to limit the yield of recombinant collagen. A number of methods have successfully increased the amount of internal proline content of *E. coli*. One method is to remove the feedback inhibition; this has been achieved on two occasions by two different point mutations in the *proB* gene, which encodes glutamate-5-kinase (Csonka et al., 1988; Rushlow et al., 1985). Alternatively completely removing the organism's ability to produce proline, supplementing the media with proline and giving an osmotic shock has been shown to cause cells to uptake proline from the environment, leading to high internal proline levels -this has been demonstrated with hydroxyproline (Buechter et al., 2003). This demonstrates that the limiting step in proline rich protein production in *E. coli* is the availability of proline and not a lack of tRNA (Jonasson et al., 2002). While effective this method is not suited to scale up, which would be essential for large scale recombinant production. Another strategy for improving the amount of intracellular proline is to increase the expression of proteins which produce endogenous proline or transport proline from the extracellular to intracellular environment and to reduce the expression of proteins which cause a reduction in intracellular proline. This may involve knocking out genes which are not conducive to proline production or accumulation or changing the promoters of genes which aid intracellular proline production or accumulation to confer increased intracellular proline. Routes for this strategy can be identified by searching the literature or by modelling the proline biosynthesis pathway and are discussed in Chapter 6.

Chapter 2. Materials and methods

2.1. Strains

Table 2.1: *E. coli* strains used in the study.

Strains were either purchased commercially, donated by University of Sheffield (UoS) staff or generated in the study

Strain	Source
MC1000: F- *Δ(araA-leu)7697 [araD139]$_{B/r}$, Δ(codB-lacI)3 galK16 galE15*(GalS) *λ$^-$ e14 relA1 rpsL150*(strR) *spoT1 mcrB1* (CGSC)	Dr Graham Stafford, UoS
MG1655: F- LAM- rph-1 (CGSC)	Dr Graham Stafford, UoS
MG1655 Δ*clpX*	Professor Jeff Green, UoS
(DH5α) NEB 5-alpha Competent *E. coli*: *fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi^{-1} hsdR17*	New England Biotechnologies
BL21 (DE3): *F$^-$ompT gal dcm lon hsdS$_B$(r$_B^-$ m$_B^-$) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])*	New England Biotechnologies
MC1000 *ΔflhDC::*FRT-*Km*-FRT	Dr Graham Stafford, UoS
MC1000 *ΔfliC::*FRT-*Km*-FRT	Dr Matthew Hicks, UoS
MC1000 *ΔfliCD::*FRT-*Km*-FRT	Dr Matthew Hicks, UoS
MC1000 *ΔfliC ΔflgKL ΔfliCD::*FRT-*Km*-FRT	This study
*MC1000 ΔfliC ΔflgKL::*FRT-*Km*-FRT	Dr Matthew Hicks, UoS
MC1000 ΔfliC ΔflgKL	Dr Matthew Hicks, UoS
MC1000 *ΔfliC ΔflgKL ΔlrhA::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔdksA::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔclpX::Km*	This study

MC1000 *ΔfliC ΔflgKL ΔfliDST::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔflgMN::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔmotAB::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔfliDST::*FRT-*Cm*-FRT *ΔflgMN::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔmotAB::*FRT-*Cm*-FRT *ΔfliDST::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔmotAB::*FRT-*Cm*-FRT *ΔflgMN::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔclpX:: Km ΔfliDST::*FRT-*Cm*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔclpX:: Km ΔflgMN ::*FRT-*Cm*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔclpX:: Km ΔmotAB::*FRT-*Cm*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔclpX:: Km ΔmotAB ΔfliDST*	This study
MC1000 *ΔfliC ΔflgKL ΔclpX::Km ΔmotAB ΔflgMN*	This study
MC1000 *ΔfliC ΔflgKL ΔflgDE::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔproB::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔproB* E153A*::*FRT-*Cm*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔputA::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔgdhA::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔgdhA ΔputA::*FRT-*Km*-FRT	This study
MC1000 *ΔfliC ΔflgKL ΔproB ΔputA::*FRT-*Km*-FRT	This study

2.2. Plasmids

Table 2. 2: Plasmids used in this study.

Plasmids were either purchased commercially, kindly gifted, donated by University of Sheffield (UoS) staff or generated in the study

Plasmid	Description/function	Resistance	Source
pTrc99a-FF	pTrc99a with *Nde*I site in multiple cloning region. Empty vector for protein expression. Referred to as pTrc99a in text.	Ampicillin	Gillian Fraser. Oshima et al. 2006
pTrc99a-E2	*fliC* Δ191-280 inserted between *Nde*I and *Bam*HI of pTrc99a-FF	Ampicillin	Dr. Matthew Hicks, UoS
pJexpress-404-1-47 *fliC*	*fliC* 5' UTR- 1-47 secretion signal residues- *XhoI*-TEV-*Eco*R1- *Pst*I- TEV-FLAG tag-Step tag- *Xba*I- *Hind*III- *fliC* 3' UTR inserted between *Nde*I and *Bam*HI of pJexpress-404-1-47 *fliC* colA1	Ampicillin	Dr. Matthew Hicks, UoS
pJexpress-404-1-47 *fliC* COL1A1	pJexpress 404 with *Nde*I- *fliC* 5' UTR- 1-47 secretion signal residues- *XhoI*- TEV- *Eco*R1- 45 kDa COL1A1- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *fliC* 3' UTR- *Bam*HI	Ampicillin	DNA 2.0, USA
pJexpress-404-1-47 *fliC*-Scl2	Scl2 inserted between *Eco*RI and *Pst*I of pJexpress-404-1-47	Ampicillin	Dr. Matthew Hicks, UoS

		fliC		
pJexpress-404-1-47 CH2	*fliC*-CH2	CH2 inserted between *Eco*RI and *Pst*I of pJexpress-404-1-47 *fliC*	Ampicillin	Dr. Matthew Hicks, UoS, CH2: Dr Jagroop Pandhal
pJexpress-404-1-47 CH2 (no 5'UTR)	*fliC*-CH2 (no 5'UTR)	PCR amplified *Nde*I- 1-47 secretion signal residues-*XhoI*- TEV- *Eco*R1- CH2- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *fliC* 3' UTR- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-CH2 (no 5'UTR)		PCR amplified *Nde*I-*XhoI*- TEV-*Eco*R1- CH2- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *fliC* 3' UTR- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-1-47 cutinase	*fliC*-cutinase	cutinase inserted between *Eco*RI and *Pst*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-1-47 cutinase (no 3' UTR)	*fliC*-cutinase (no 3' UTR)	PCR amplified *Nde*I- *fliC* 5' UTR 1-47 secretion signal residues-*XhoI*- TEV- *Eco*R1- cutinase-*Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study

pJexpress-404-1-47 *fliC*-cutinase (no 5’ UTR)	PCR amplified *Nde*I- 1-47 secretion signal residues- *Xho*I- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag- *Xba*I- *Hind*III- *fliC* 3’ UTR- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-1-47 *fliC*-cutinase (no 5’ or 3’ UTR)	PCR amplified *Nde*I- 1-47 secretion signal residues- *Xho*I- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag- *Xba*I- *Hind*III- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-26-47 *fliC*-cutinase (no 5’ UTR)	PCR amplified *Nde*I- 26-47 secretion signal residues- *Xho*I- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag- *Xba*I- *Hind*III- *fliC* 3’ UTR- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-26-47 *fliC*-cutinase (no 5’ or 3’ UTR)	PCR amplified *Nde*I- 26-47 secretion signal residues- *Xho*I- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag- *Xba*I- *Hind*III- - *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study

pJexpress-404-sal 26-47 *fliC*-cutinase (no 5' UTR)	PCR amplified *Nde*I- 26-47 *Salmonella* secretion signal residues- *XhoI*- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *fliC* 3' UTR- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-sal 26-47 *fliC*-cutinase (no 5' or 3' UTR)	PCR amplified *Nde*I- 26-47 *Salmonella* secretion signal residues- *XhoI*- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- - *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404-*fliC*-cutinase (no 5'UTR)	PCR amplified *Nde*I- *XhoI*- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *fliC* 3' UTR- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pJexpress-404- *fliC*-cutinase (no 5' or 3' UTR)	PCR amplified *Nde*I- *XhoI*- TEV- *Eco*R1- cutinase- *Pst*I- TEV- FLAG tag- Step tag-*Xba*I- *Hind*III- *Bam*HI inserted between *NdeI* and *BamH*I of pJexpress-404-1-47 *fliC*	Ampicillin	This study
pKD46	Expression of λ-RED recombinase system for	Ampicillin	Barry Wanner,

	recombineering		USA
pCP20	Expression of FLP recombinase for excision of FRT-flanked markers	Ampicillin	Barry Wanner, USA
pKD3	Template for overlap PCR and cassette construction	Chloramphenicol	Barry Wanner, USA
pKD4	Template for overlap PCR and cassette construction	Kanamycin	Barry Wanner, USA
pKD13	Template for overlap PCR and cassette construction	Kanamycin	Barry Wanner, USA
pKD32	Template for overlap PCR and cassette construction	Chloramphenicol	Barry Wanner, USA
pGPS123-lacZ	pGPS123 constructed by *Xho*I and *HindIII* excision of Km and insertion of Gm in pRS551. *lacZ* inserted between *Eco*RI and *Pst*I in pGPS123.	Ampicillin, Gentamycin	Dr Graham Stafford, UoS
pGPS123-*fliA*-lacZ	*fliA-lacZ* inserted between *Eco*RI and *Pst*I in pGPS123	Ampicillin, Gentamycin	Dr Graham Stafford, UoS
pGPS123-*fliL*-lacZ	*fliL-lacZ* inserted between *Eco*RI and *Pst*I in pGPS123	Ampicillin, Gentamycin	Dr Graham Stafford, UoS
pGPS123-*flhB*-lacZ	*flhB-lacZ* inserted between *Eco*RI and *Pst*I in pGPS123	Ampicillin, Gentamycin	Dr Graham Stafford, UoS

2.3. Chemicals, reagents and buffers

Unless stated all chemicals and reagents were purchased from Sigma-Aldrich or Fisher Scientific. All reagents used were of analytical grade and were purchased from the following suppliers (Table 2.3). Buffers were prepared in distilled water and pH adjusted with HCl or NaOH. Solutions were sterilised by autoclave or sterile filtration (0.22µm). The manufacturer's instructions were followed during the use of all molecular biology kits.

Table 2.3: Manufacturers and suppliers of chemicals, reagents, kits and equipment.

Supplier	Location
BDH Laboratory Supplies	Poole, UK
Beckman Coulter	High Wycombe, UK
Bioline	London, UK
BioRad Laboratories	Hertfordshire, UK
BMG Labtech	Ortenberg, Germany
Calbiochem (Merck Millipore)	Watford, UK
Cell Signalling Technology	Danvers, Massachusetts, USA
Eppendorf	Hamburg, Germany
Expedeon	Swavesey,Cambridgeshire, UK
Fisher Scientific	Loughborough, UK
FlowGen Biosciences (SLS Life Sciences)	Hassle, East Riding of Yorkshire, UK
GATC Biotech	Konstanz, Germany
GE Healthcare Life Science	Buckinghamshire, UK
IBA Lifesciences	Goettingen, Germany

Life Technologies	Carlsbad, California, USA
Merck Millipore	Watford, UK
New England Biolabs (NEB)	Hitchin, Hertfordshire, UK
Norgen	Thorold, Canada
Promega	Southampton, UK
Sigma-Aldrich	Poole, UK
Statens Serum Institut	Copenhagen, Denmark
Syngene	Cambridge, UK
Tecan Group Ltd.	Männedorf, Switzerland
Thermo Scientific	Leicestershire, UK
Waring Laboratory Science	Torrington, Connecticut, USA
Xograph	Gloucestershire, UK

2.4. Microbiological culture methods

2.4.1. Bacterial culture

Stocks of bacteria were stored in 12.5% glycerol (v/v) in cryo-vials at -80°C. Luria Broth (LB) Agar supplemented with necessary antibiotics (ampicillin: 100µg/mL LB, kanamycin: 50µg/mL LB, chloramphenicol: 25µg/mL LB. All filter sterilised) were inoculated with *E. coli* either from cryo-vials or cell pellets following liquid culture. Cells were incubated statically at 37°C overnight and inoculated LB agar plates stored at 4°C for up to one month. Single *E. coli* colonies (or around 5 colonies for a secretion assay experiment) were selected from LB agar plates and grown in 5mL LB broth in a 20mL universal tube overnight. LB was supplemented with relevant antibiotics (aforementioned concentrations). Overnight starter cultures were then used to inoculate fresh LB, supplemented with relevant antibiotics. Starter cultures were

diluted 1 in 100 in 10mL LB broth in a 50mL Falcon tube. Exceptions to this involved the inoculation of 500mL LB broth in a 2.5L conical flask for protein overexpression, 100mL LB broth in a 500mL conical flask for secretion or gene expression assay growth curve experiments or 200µL LB broth per well for 96 well plate growth curve experiments. If necessary, additional supplements were added to initiate promoter induction of plasmids – these included (isopropyl-β-d-galactopyranoside) IPTG (Calbiochem) and L-arabinose (both filter sterilised) Cells were cultured at 37°C unless otherwise stated, with 180rpm agitation (140rpm for conical flask cultures). Optical density of cells was measured in a 1mL cuvette at OD_{600} using a spectrophotometer.

2.4.2. Transformation by heat shock

For transformation of plasmids into *E. coli* heat shock was used. 10mL cultures were grown as described in 2.4.1 to OD_{600} 0.4-0.6 and incubated on ice for 20 minutes. Cells were centrifuged in a pre-chilled (4°C) centrifuge for 10 minutes, 3500rpm to pellet cells. The pellet was resuspended in 10mL chilled 0.1M $CaCl_2$ and incubated on ice for 10 minutes. The centrifugation step was repeated and the pellet was resuspended in 1mL 0.1M $CaCl_2$. 3µL plasmid DNA was transferred into 200µL cell aliquots in Eppendorfs and incubated on ice for 30 minutes. Eppendorfs were transferred to a 42°C water bath for 90 seconds and then incubated on ice for 2 minutes. 1mL LB was added and cells were left to recover at 37°C for 1 hour and cells plated onto LB agar plates with the relevant antibiotics. Alternatively at the final resuspension step, the cell pellet was resuspended in 1mL 0.1 $CaCl_2$ 10% (v/v) glycerol. 200 µL cell aliquots were transferred to Eppendorfs and stored at -80°C. Following defrost of these aliquots on ice; the protocol could then be resumed.

2.4.3. Transformation by electroporation

Plasmids were also transformed into *E. coli* by electroporation. 10mL cells were grown as described in 2.4.1 to OD_{600} 0.4-0.6 and incubated on ice for 20 minutes. Cells were centrifuged in a pre-chilled centrifuge at 7500rpm for 5 minutes, supernatant was resuspended in 10mL ice

cold dH_2O. This step was repeated three times. Following the final spin cells were resuspended in up to 500µL dH_2O. 1µL purified PCR product or plasmid DNA was added to 100µL cell aliquots, contents were transferred to a pre-chilled electroporation cuvette (Flowgen Biosciences). Cells were electroporated in a BIO-RAD MicroPulser™ series 411BR (BioRad Laboratories) (EC1 setting, PL5, potential difference: 2.5kV, resistance: 200Ω and capacitance: 25µF). Cells were recovered in 1mL LB for one hour at 37°C, unless otherwise stated and cells plated onto LB agar plates with the relevant antibiotics. For storage at -80°C, at the final resuspension step, the cell pellet was instead resuspended in 10% (v/v) glycerol. 100 µL cell aliquots were transferred to Eppendorfs and stored at -80°C. Following the defrost of these aliquots on ice; the protocol could then be resumed.

2.4.4. Chromosomal mutagenesis by Lambda Red recombination.

Chromosomal mutagenesis was achieved by the method outlined by Datsenko & Wanner (2000). Chromosomal disruptions were implemented by Lambda Red recombinase induced homologous recombination. Chromosomal disruptions include knockout mutagenesis and genomic insertions. For knockout mutagenesis PCR products were generated which harboured FRT (FLP recombinase recognition site) flanked antibiotic resistance cassettes (derived from template DNA of pKD3, pKD4 or pKD13) and ~35bp homologous nucleotides to regions adjacent to the gene which was inactivated. For genomic insertion, additional DNA (a gene or a promoter region) was also incorporated into the PCR product, adjacent to the antibiotic resistance cassette. Mutagenesis events occur following electroporation of the PCR product into arabinose induced Lambda Red recombinase plasmid pKD46 (a derivative of pINT-ts) containing cells. The plasmid harbours γ, ß and exo (λ) (Murphy, 1998). γ inhibits RecBCD activity, which digests linear DNA. Exo encodes an exonuclease which digests the 5'-3' double stranded linear DNA. ß is a binding protein capable of annealing complementary single stranded DNA.

Bacterial cultures expressing the Lambda Red recombinase plasmid pKD46 were supplemented with 1mM L-arabinose and cultured at 30°C with agitation. Cells underwent transformation by electroporation with PCR derived linear FRT flanked DNA. Following incubation for one hour at 30°C and then overnight at room temperature, cells were spread on to LB agar plates

supplemented with relevant antibiotics. Following successful mutagenesis the antibiotic resistance site could be removed by a FLP recombinase carrying helper plasmid (pCP20) which acts on the FRT sites. The Red and helper plasmids are cured at 43°C. Plasmids were cured so that no further recombination events could occur. The antibiotic resistance gene flanked by FRT sites was removed from the chromosome of cells expressing the pCP20 plasmid when incubated at 43°C.

2.4.5. Chromosomal mutagenesis by P1 phage transduction

P1 phage transduction utilises P1*vir* phage to move regions of chromosomal DNA from one *E. coli* genome to another. As P1*vir* is capable of packaging up to 90,000kb DNA, this tool can be harnessed to transduce genes linked to selectable genetic markers (e.g. antibiotic resistance cassettes).

2.4.5.1. Lysate preparation

An overnight culture of the donor strain was diluted 1 in 100 in LB broth with 5mM $CaCl_2$ and 0.2% (w/v) glucose and incubated at 37°C, 150rpm for 1 hour. 100µL P1 phage lysate was added to the culture and incubated for 1-3 hours (until the media was clear). Several drops of chloroform were added and the culture was vortexed and centrifuged for 2 minutes, 13000rpm. The supernatant was transferred to a fresh tube and a few drops of chloroform were added. The lysate was stored at 4°C.

2.4.5.2. Transduction

An overnight culture of the recipient strain was centrifuged to pellet the cells. The cell pellet was resuspended in half the volume of LB broth with 100mM $MgSO_4$ and 5mM $CaCl_2$. 100µL aliquots of culture were transferred into Eppendorf tubes. The Eppendorfs contained (1) Donor strain P1 lysate (2) 50µL donor strain P1 lysate. Negative controls contained (3) 100µL recipient cells or (4) 100µL donor strain P1 lysate only. Aliquots were incubated at 30°C for 30 minutes without agitation. 1mL LB broth was added and cells were incubated at 37°C for 1 hour. Cells

were spread on LB agar plates containing relevant antibiotics and 10mM sodium citrate (to chelate the calcium and inhibit reinfection). Successful transformants were passaged on LB agar with the appropriate antibiotic and sodium citrate three times to ensure the eradication of P1.

2.4.6. Strain storage

Following the conformation of a novel strain or plasmid, cells were stored at -80°C. Single colonies were picked from an LB agar plate and grown overnight in LB broth with relevant antibiotics. 1mL media was transferred to a cryo-tube with 0.5mL 50% (v/v) glycerol LB media and stored at -80°C.

2.4.7. Plate reader growth curve

For 96 well plate growth curve experiments cells were incubated in the Tecan Infinite 200 Pro plate reader (Tecan Group Ltd.) at 37°C, with 6mm orbital agitation –OD_{600} readings were taken every 30 minutes.

2.4.8. Gram Staining

Cell morphology was visualised by microscopy following Gram staining. Cells were grown in LB media and following centrifugation a scoop of the cell pellet was placed on the slide. Cells were heat fixed over a Bunsen burner flame. Cells were flooded with crystal violet and incubated for one minute. Residual stain was washed off under a running tap; cells were then covered in Gram's iodine and left for one minute. The wash step was repeated and cells were decolourised with a few drops of acetone and left for 30 seconds. The slide was rinsed again and counter stained with carbol fuchsin for 30 seconds. The slide was rinsed and dried with care. Slides were visualised under a Nikon Eclipse TS100 microscope with a digital camera attachment.

Crystal Violet

The stain is a 0.1 % Crystal Violet stain made up of 20 ml Solution A (5 g crystal violet (BDH Laboratory supplies) in 95 mL ethanol) and 80 ml Solution B (1% Aqueous Ammonium oxalate in H_2O).

Gram's iodine

1g iodine and 3g potassium iodide were added to 300mL distilled water.

Carbol Fuchsin

1g carbol fuchsin was added to 100mL 95% ethanol

2.5. Molecular biology methods

2.5.1. DNA methods

2.5.1.1. Buffers and reagents

10 x TAE buffer

DNA agarose gels were prepared with and run in TAE buffer. 48.4g Tris base, 20mL 0.5M EDTA, pH 8 and 11.44mL glacial acetic acid were dissolved in up to 1 litre of dH_2O. Prior to use a 1 x TAE buffer was prepared in dH_2O.

6x DNA loading buffer

DNA was prepared for gel electrophoresis in 6x DNA loading buffer. 100mL in dH_2O with 60mL glycerol, 6mL Tris, pH 8, 1.2mL 0.5M EDTA, pH 8 and 60mg bromophenol blue.

2.5.1.2. DNA agarose gel electrophoresis

Unless specified, DNA analysis was carried out on 1% (w/v) TAE agarose gels. Agarose powder was dissolved in 1 x TAE buffer and supplemented with 0.5 µl of ethidium bromide. DNA gels were set in casting trays. 6x DNA loading buffer was added to DNA samples which were run alongside either GeneRuler 1kb DNA ladder (Thermo Scientific) or PCR Ranger 100bp DNA

ladder (Norgen). DNA gels were run at 100V until adequate separation of DNA fragments was achieved. DNA was observed under UV light using a G:BOX (Syngene). If DNA bands required isolation, bands were cut from the gel with a scalpel. DNA was purified with the ISOLATE II PCR and Gel Kit (Bioline). DNA was stored at -20°C.

2.5.1.3. Bacterial plasmid DNA extraction

Plasmid DNA was isolated from liquid cell culture with the Isolate II Plasmid Mini Kit (Bioline).

2.5.1.4. Bacterial chromosomal DNA extraction

Chromosomal DNA was isolated from liquid cell culture with Wizard® Genomic DNA Purification Kit (Promega).

2.5.1.5. Quantification of DNA concentration

Following calibration with nuclease free water, the concentration (ng μL^{-1}) 1.5µL isolated DNA was measured by a Nanodrop 1000 spectrophotometer (Thermo Scientific).

2.5.1.6. Polymerase chain reaction

Polymerase chain reaction (PCR) was carried out to amplify existing DNA templates. Primers were designed and then synthesised by Sigma-Aldrich. Upon arrival nuclease free water was added to the dry oligonucleotides to result in a 10µM solution. Primers were stored at -20°C. Prior to use primers were diluted 1 in 10 to form a 100µM solution. Primers with extensions to the homologous region at the 5' end were used to add nucleotides (e.g. restriction enzyme sites) to the ends of template DNA. Primers are specified in Appendix 2.

2.5.1.6.1. High-fidelity PCR

Phusion® High-Fidelity DNA Polymerase (New England Biosciences) was used to ensure accurate DNA amplification for vector cloning, chromosomal mutagenesis or DNA sequencing. PCR reactions for Phusion® reactions were set up as specified in Table 2.4 and

Table 2.5.

Table 2.4: Reagents and quantities for PCR with Phusion® High-Fidelity DNA Polymerase

50 µL PCR reaction mixes were prepared. For convenience master mixes were also prepared when multiple PCR reactions were being carried out with common components.

Reagent	**µL**
NF dH_20	35.5
5 x Phusion® High-Fidelity Buffer	10
DMSO	1
DNA template (~10µM)	0.5
Primer 1 (100µM)	0.5
Primer 2 (100µM)	0.5
dNTPs (100µM)	0.5

Phusion® High-Fidelity DNA Polymerase	0.5

Table 2.5: Settings for PCR with Phusion® High-Fidelity DNA Polymerase.

Annealing temperature and duration varied based on the primers. If no product was derived following an annealing setting of 55°C for 1 minute (a) the reaction was repeated with an annealing setting of 36°C and 2 minutes (b). The extension time varied based on the size of the PCR product. There were 30 cycles of the middle three steps.

Step	Temperature (°C)	Time (m:s)
Initial Denaturing	98	3:00
Denaturing	98	0:30
Annealing	a. 55 or b. 36	a. 1:00 or b. 2:00
Extension	72	0:30/kb product
Final Extension	72	7:00

2.5.1.6.2. Overlap PCR

Overlap PCR was used to adjoin DNA templates or implement single point substitutions. As the product was then incorporated into genomic DNA for transcription, it was essential that replication fidelity was high; therefore Phusion® High-Fidelity DNA Polymerase was used. Reactions were set up as described in Table 2.4; however DNA from multiple templates was added. The settings for thermocycling were identical to those in those in

Table 2.5, with consideration being made for the product length of the adjoined templates when calculating extension time.

2.5.1.6.3. Colony PCR

Colony PCR involves the PCR based confirmation of chromosomal mutagenesis or vector DNA ligation in individual *E. coli* colonies. As fidelity was not essential DreamTaq DNA Polymerase (Thermo Scientific) was utilised for PCR reactions which were set up as specified in Table 2.6 and Table 2.7.

Table 2.6: Reagents and quantities for PCR with DreamTaq DNA Polymerase.

20 µL PCR reaction mixes were prepared. The DreamTaq PCR Master Mix comprises of buffer, polymerase and dNTPs. For convenience master mixes were also prepared when multiple PCR

reactions were being carried out with common components. A small scoop of DNA from a single colony was added to the PCR tube instead of isolated template DNA.

Reagent	µL
NT dH_20	9.7
2 x DreamTaq PCR Master Mix	10
DMSO	0.1
DNA template	N/A
Primer 1 (100µM)	0.1
Primer 2 (100µM)	0.1

Table 2.7: Settings for PCR with DreamTaq DNA Polymerase.

Annealing temperature and duration varied based on the primers. If no product was derived following an annealing setting of 55°C for 30 seconds (a) the reaction was repeated with an annealing setting of 36°C and 1 minute (b). The extension time varied based on the size of the PCR product. There were 30 cycles of the middle three steps. In colony PCR the initial denaturing time was extended to 5 minutes, to ensure cell lysis occurred.

Step	Temperature (°C)	Time (m:s)
Initial Denaturing	95	5:00
Denaturing	95	0:30
Annealing	a. 55 or b. 36	a. 0:30 or b. 1:00
Extension	72	1:00
Final Extension	72	7:00

2.5.1.7. Blunt end cloning

Linear DNA (for example, PCR products derived following overlap PCR) was blunt end cloned into a storage vector using the CloneJET PCR Cloning Kit (Thermo Scientific). Efficient ligation of linear DNA into the vector was ensured through the induction of a suicide gene in re-circularised vector. Blunt end cloning and use of proofreading enzymes insured that the entire DNA fragment including ends was correctly incorporated into the vector. Following transformation into cells, this allowed storage of DNA and a simple means of DNA amplification, by cell division. Insert DNA was isolated from the vector by either PCR of restriction digestion.

2.5.1.8. DNA ligation

Restriction digestion and ligation were used to clone DNA into vectors. Cloning vectors (acceptors) and insert DNA underwent double restriction digest with appropriate restriction enzymes and buffers. Insert DNA was derived either from plasmid DNA or PCR reaction. Antarctic phosphatase was added to restriction digest reactions of acceptor vector DNA to prevent self-ligation of the vector. All restriction enzymes, buffers and Antarctic phosphatase were purchased from New England Biosciences. Restriction digest reactions were set up as stated in

Table 2.8. Reactions were incubated for 2 hours at 37°C. Following incubation, 6µL 10x Antarctic Phosphatase Reaction Buffer and 2µL Antarctic Phosphatase enzyme were added to acceptor vector DNA digest reactions and the mix was incubated at 37°C for a further hour. 5µL acceptor vector DNA reaction mix was run on a TAE agarose DNA gel alongside undigested donor DNA to confirm successful digestion. Restriction enzymes in acceptor vector DNA reaction mixes were heat inactivated by incubating at 65°C for 20 minutes, with the exception of *Bam*HI which required inactivation by clean up with the ISOLATE II PCR and Gel Kit (Bioline). All of the insert DNA reaction mixture was loaded on to a TAE agarose DNA gel, DNA product of the expected size was excised and purified with the ISOLATE II PCR and Gel Kit (Bioline).

Table 2.8: Reagents and quantities to restriction enzyme digest.

The appropriate buffer for the two restriction enzymes was selected using the New England Biosciences Double Digest Finder. Reaction mixtures were set up in Eppendorf tubes.

Reagent	**µL**
NF dH_20	Up to a total of 60
10 x Buffer	6
Restriction enzyme 1	2
Restriction enzyme 2	2
DNA	1µg total -for example 20µL of 50ng/µL

Following successful digestion, vector and insert DNA were ligated using T4 DNA ligase (New England Biosciences). Reaction mixes were set up as listed in Table 2.9 and incubated at room temperature overnight.

Table 2.9: Reagents and quantities for DNA ligation.

Reactions were set up in 0.5mL Eppendorf tubes. A 1:1 ratio of vector to insert sticky ends is desirable; as vector DNA is bigger (kbp) than insert DNA, DNA was added to result in a 1:3 ratio.

Reagent	**µL**
NF dH_20	Up to 20
10 x T4 DNA ligase buffer	2
Vector DNA	As required. 1 part

Insert DNA	As required. 3 parts
T4 DNA ligase buffer	1

1µL ligation mix was transformed into NEB 5-alpha Competent *E. coli* (New England Biosciences) and screened by plating on LB agar plates supplemented with the relevant antibiotic based on the acceptor vector. Sussessful transformants were screened by colony PCR. Colonies which yielded positive PCR results were grown for plasmid extraction to undergo confirmatory restriction digest and DNA sequening to confirm the correct sequence of DNA.

2.5.1.9. DNA sequencing

Both chromosomal and plasmid DNA were sequenced using the LIGHTRUN™ Sequencing Service (GATC Biotech). The region of chromosomal DNA which required sequencing was first amplified using primers which annealed either side of the region and high-fidelity polymerase. In an Eppendorf, 5µL of either 80-100ng μL^{-1} purified plasmid DNA or 20-80ng μL^{-1} purified PCR product DNA was added to 5µL of 5µM primer. Reaction mixes were sent to GATC Biotech. The resulting nucleotide sequences were visualised using Finch TV. Multiple reaction mixes were prepared for each DNA template, with different primers to yield full coverage of DNA based on at least two different primers (which were designed to anneal at either end of the DNA template). Sequence coverage was achieved for around 1Kb of DNA; for DNA which exceeded this length, primers were also designed to anneal within the DNA template.

2.5.1.10. Gene synthesis

Synthetic genes were designed and sent for synthesis by GeneArt® Strings DNA Fragments (Life Technologies). Upon delivery, dried DNA was resuspended in 50µL nuclease free water and blunt end cloned into a storage vector using the CloneJET PCR Cloning Kit (Thermo Scientific). The ligation mix was transformed into NEB 5-alpha Competent *E. coli* (New England Biosciences) and screened by plating on an ampicillin supplemented LB agar plate. Transformants were confirmed by DNA sequencing.

2.5.2. Protein methods

2.5.2.1. Buffers and reagents

SDS-PAGE lower Tris buffer

For use in SDS-PAGE resolving gel, 6.06g Tris Base and 0.4g Sodium dodecyl sulphate (SDS) were dissolved in 100mL dH_2O before adjusting the pH to 6.8.

SDS-PAGE upper Tris buffer

For use in the SDS-PAGE stacking gel, 18.17g Tris base and 0.4g Sodium dodecyl sulphate (SDS) were dissolved in 100mL dH_2O before adjusting the pH to 8.8.

10 x SDS-PAGE running buffer

SDS-PAGE gels were run in tanks with SDS-PAGE running buffer. 30g Tris base, 144g glycine and 10g SDS was made up to a volume of 1000mL with dH_2O. 1 x SDS-PAGE running buffer was prepared in dH_2O.

2x SDS loading buffer

Protein samples were prepared for SDS-PAGE in 2x SDS loading buffer. This was prepared with 10mL glycerol, 1g SDS and 0.1g bromophenol blue to a total of 50mL 100mL Tris, pH 6.8. Prior to use DTT from a 1M stock was added to a concentration of 200mM.

Semi dry transfer buffer

Semi dry transfer buffer was used during protein transfer for Western blot and was made with 2.9g Tris base, 1.45g glycine, 1.85mL 10% (w/v) SDS and 100mL methanol. The reagents were mixed and the total made up to 500mL with H_2O.

10 x TBS buffer

TBS was used for incubation of nitrocellulose membranes. 24g Tris base and 88g NaCl was dissolved in dH_2O and pH adjusted to 7.6 before adding dH_2O to a final volume of 1 litre. Prior to use a 1 x solution of TBS was prepared in dH_2O.

2.5.2.2. Bacterial protein precipitation

Protein analysis work was carried out with SDS-PAGE gels and Western blot. Following cell culture samples were prepared as follows. Following centrifugation bacterial cell culture cell pellets were resuspended in 200µL 2 x SDS loading buffer per OD Unit (1mL cell culture at OD_{600} 1.0 is equivalent to 1 OD Unit) of cells and heated at 95°C for 10 minutes. For total supernatant protein precipitation as part of the secretion assays, 10% v/v trichloroacetic acid (TCA) was added to supernatant fractions and incubated on ice for 30 minutes. The mixture was pelleted by centrifugation (4°C) and the supernatant was discarded. The pellet was resuspended in the same volume of acetone, vortexed and the mixture was centrifuged before supernatant being discarded and the pellet air dried. The pellet was suspended in 50µL 2x SDS loading buffer/OD unit and heated at 95°C for 10 minutes. For samples derived from protein purification 2x SDS loading buffer was added at a 1:1 ratio and heated for 5 minutes at 95°C.

2.5.2.3. SDS-PAGE

Protein analysis work was carried out on SDS-PAGE gels. SDS-PAGE gels were prepared using equipment from the Mini-PROTEAN® (Bio-Rad) casting kit. Long plates with 1.5mm spacers and short plates were cleaned with 70% industrial methylated spirits and assembled on casting module. The resolving gel was prepared and poured between the plates. Recipes for SDS-PAGE resolving gels are listed in

Table 2.10; an appropriate % crosslinking was selected based on the size of the protein of interest. Isopropanol was layered on top to level the gel and protect it from the air interface. Once set, isopropanol was removed, the stacking gel (0.975mL acrylamide, 2.1mL 0.5M Tris pH 6.8 (upper Tris), 4.725mL H_2O, 17µL TEMED, 0.1mL 10% (w/v) ammonium persulphate) poured and the comb inserted. Samples were loaded on to SDS-PAGE gels with EZ-Run™ Prestained Rec Protein Ladder (Fisher Scientific) and run in a Mini-PROTEAN ® Tetra Vertical Electrophoresis Cell in 1 x SDS running buffer. Gels were washed in dH_20 three times and

strained with either InstantBlue™ (Expedeon) or Silver Stain Plus™ (Bio-Rad) reagents and washed three times in dH_2O before imaging with a scanner.

Table 2.10: Reagents and quantities for polyacrylamide resolving gels.

Size of the protein of interest determined the % acrylamide in the resolving gel. 5µL TEMED and 0.35mL 10% (w/v) ammonium persulphate were added to the reagents listed below to initiate crosslinking.

% crosslinking	Acrylamide (mL)	Lower Tris (1.5M, pH 8.8)(mL)	H_20 (mL)
10	2.475	2.5	4.825
12	3	2.5	4.3
15	3.75	2.5	3.55

2.5.2.4. Western blot

An SDS-PAGE gel was stacked with a nitrocellulose membrane (GE Healthcare Life Science) sandwiched in 3 layers of chromatography paper (GE Healthcare Life Science). All were saturated with 1 x semi-dry transfer buffer. The layers were assembled on a BioRad Trans-Blot Semi Dry Transfer Cell (BioRad Laboratories), which was run at 10v for one hour. The nitrocellulose membrane was placed in 50mL 3% w/v BSA in 1 x TBS/0.1% TWEEN and left on rollers for two hours. The membrane was washed in 1 x TBS/0.1% TWEEN 3 x 10 minutes. 20mL 1 x TBS/0.1% TWEEN containing the primary antibody was added to the membrane which was incubated on rollers for one hour. Wash steps were repeated. If necessary the antibody incubation step and wash steps were repeated with a secondary antibody diluted in 1 x TBS/0.1% TWEEN. Antibodies are listed in Table 2.11. The membrane was soaked in Pierce® ECL Western Blotting Substrates (Thermo Scientific) and placed in cling film and moved into an X-Ray cassette folder (Kodak). Sheets of CL-Xposure™ X-ray films (Thermo Scientific) were placed in the cassette on the membrane for varying lengths of time. Films were developed using a Compact X4 X-ray Film Processor (Xograph). If necessary, nitrocellulose membranes were stripped of antibodies to allow probing with alternative antibodies. Membranes were

washed in 1 x TBS/0.1% TWEEN 3x 10 minutes and incubated in 10mL Restore™ Western Blot Stripping Buffer (Thermo Scientific) for 15 minutes. Wash steps were repeated and the process could then be initiated at the blocking stage.

Table 2.11: Antibodies for probing nitrocellulose membranes for Western blotting.

The antibody, manufacturer and concentration of antibody in in 1 x TBS/0.1% TWEEN is listed

Antibody	Manufacturer	Concentration
H48 Monospecific H Rabbit Antiserum, *E. coli*	Statens Serum Institut	1 in 1000
Anti-rabbit HRP (horse radish peroxidase).	Cell Signalling Technology	1 in 3000
Monoclonal ANTI-FLAG®M2 antibody produced in mouse	Sigma-Aldrich	1 in 1000
Monoclonal Anti-COL1A1 antibody produced in mouse	Sigma-Aldrich	1 in 500
Anti-mouse HRP (horse radish peroxidase)	Cell Signalling Technology	1 in 3000
E. coli GroEL Rabbit IgG	Sigma-Aldrich	1 in 1000

2.5.2.5. Cell lysis

Cell lysis was required prior to the purification of intracellular proteins. Cell cultures underwent centrifugation, supernatant was discarded and the cell pellet was resuspended in the appropriate buffer for subsequent protein purification. Cells which overexpressed streptavidin II tagged proteins were lysed by a French pressure cell. Cells underwent three repetitions of cell disruption by French press at 1000psi. Cells were incubated on ice for 1 minute between applications.

2.5.2.6. Concentration of protein

Amicon® Ultra 0.5mL Filters (Merck Millipore) allowed concentration of protein by size exclusion either prior to or following protein purification. Up to 500µL protein was applied to the column and centrifuged at 13,000rpm, 4°C until the desired volume of sample was achieved. Cycles of concentration of protein followed by the addition of an alternative buffer and further concentration allowed buffer exchange using this system. For larger volumes of protein, Vivaspin 20 Protein Concentrator Spin Columns, MWCO 10000 (GE Healthcare Life Science) were utilised, allowing the concentration of volumes up to 20mL.

2.5.2.7. Protein purification

Protein was purified either for use as a protein standard or to demonstrate TEV cleavage. When purifying Streptavidin II tagged protein from the secreted fraction of cell culture protein was suspended in 100mM Tris-HCl, 150mM NaCl, 1mM EDTA pH 8 and purified using *Strep*-Tactin® Spin Column (IBA Lifesciences) and eluted with 2mM D-biotin in the appropriate buffer for subsequent experimental work. Streptavidin II tagged protein from the intracellular fraction only was purified on a 1mL StrepTrap™ HP column (GE Healthcare Life Science) according to the manufacturer's instructions. Protein was suspended in 100mM Tris-HCl, 150nM NaCl, 1mM EDTA, pH8 and applied to the column coupled to a P-1 peristaltic pump (GE Healthcare Life Science) according to the manufacturer's instructions. Protein was eluted with 2.5mM desthiobiotin and underwent dialysis overnight to yield purified protein in 100mM Tris-HCl, pH 8. For SDS-PAGE protein standards, protein was immediately prepared in 2x SDS loading buffer, aliquoted and stored at -20°C.

2.5.2.8. Quantification of protein concentration

Purified protein concentration was quantified with the Pierce™ BCA Protein Assay Kit (Thermo Scientific) in a clear 96 well plate. Absorbance of solutions was measured by a FLUOstar plate reader (BMG Labtech). Results were plotted on a scatter graph and the concentration of unknown samples was calculated from the linear equation derived from absorbance of the protein standards.

2.5.2.9. TEV cleavage

Purified protein which harboured TEV (Tobacco Etch Virus) sites was cleaved by AcTEV™ Protease (Thermo Scientific). The manufacturer's protocol was followed however on occasions protein was initially suspended in 50mM Tris-HCl, 0.5mM EDTA, pH 8 rather than in 100mM Tris-HCl, 150mM NaCl, 1mM EDTA, pH 8 as recommended. Around 20µg protein was incubated with 1µL AcTEV™ Protease, 7.5µL 10x TEV buffer, 1,5µL 0.1M DTT and nuclease free dH_2O to a volume of 150µL for 2 hours at room temperature.

2.6. Assays to quantify type III secretion system capacity for protein secretion.

2.6.1. Gene expression assay

Gene expression was measured by gene promoter activity using the Miller assay (Miller, 1972). The assay measures β galactosidase activity of whole cells which harbour lacZ gene fusion plasmids. The assay measures the catalysis of ONPG to ONP. This is dependent on β galactosidase activity, which is relative to lacZ activity, which is dependent on fusion gene activity.

LB media was inoculated with overnight cell culture to a starting OD_{600} of 0.05 along with antibiotics. Culture was grown 37°C, 140rpm. A sample of cell culture was transferred to an Eppendorf and incubated on ice for 20 minutes. The OD_{600} was measured. Cells were pelleted by centrifugation and resuspended in the same volume of Z buffer. 100µL chloroform, 50µL 0.1% (w/v) SDS was added and vortexed for 30 seconds. Eppendorfs were equilibrated at 28°C for 5 minutes. 200µL ONPG substrate was added, samples were vortexed and the stop clock was started. When the solution developed a yellow colour 500µL 0.1M $NaCO_3$ was added, the mixture was vortexed and the clock stopped. Samples were centrifuged and OD_{420} and OD_{550} were measured against a Z buffer blank.

Miller units were calculated as follows: 1000 x [(OD_{420} -1.75 x OD_{550})/(T x V x OD_{600})]

Where T = time of ONPG assay in minutes and V = volume of culture in mL

Z buffer

0.42g Na_2HPO_4, 0.28g $NaH_2PO_4.H_2O$ and 0.5mL 1M KCl was added to 40mL dH_2O. 135µL 2-mercaptocethanol was added fresh. The pH was adjusted to 7 and water added to 50mL.

ONPG

ONPG (2-nitrophenyl β-D-galactopyranoside) was dissolved in phosphate buffer to a concentration of 4mg mL^{-1} fresh every day. Phosphate buffer was prepared by adding 0.85g Na_2HPO_4 and 1.03g $NaH_2PO_4.2H_2O$ to 90mL dH_2O. Adjust pH to 7 and add dH_2O to a final volume of 100mL.

2.6.2. Protein secretion assay

As described in section 2.4.1. Bacterial culture, LB media was inoculated with overnight cell culture along with antibiotics and IPTG if necessary. During incubation OD_{600} was measured and 1 OD unit sample taken either every hour or when OD_{600} reached 1.0 or 1.5 (for E2 protein). Samples were prepared for SDS-PAGE and Western blotting as described in 2.5.2.2. Bacterial protein precipitation. A protein standard was added to quantify relative concentrations of protein loaded onto the SDS-PAGE gel. This was be calculated by densitometry of western blot bands using ImageJ (http://rsbweb.nih.gov/ij/).

2.6.3. Flagellin standard to quantify secretion

To quantitatively measure flagellin secretion by strains a flagellin standard (E2) was used. pTrc99a harbouring E2 was freshly transformed into MC1000 Δ*fliC*. 100µL overnight culture was plated onto semisoft agar plates (1% w/v tryptone, 0.5% w/v NaCl, 0.8% w/v bacteriological agar) and incubated (along with a flask containing water to ensure humidity) at 37°C for 24 hours. Cells were scraped off motility agar plates with a spreader and transferred into 12mL 50mM Tris-HCl, pH 7.8. The solution was agitated in a lab blender (Waring

Laboratory Science) for 3 minutes on full power. Contents were transferred into Eppendorfs and centrifuged on full speed for 5 minutes at 4°C to remove bacterial cells. The supernatant was transferred into Beckman high speed centrifuge tubes (Beckman Coulter) and centrifuged at 67500g for 15 minutes at 4°C to collect flagella filaments. The supernatant was discarded and the pellet resuspended in the residual supernatant. The centrifugation step was repeated, the supernatant removed and the pellet resuspended in 50mM Tris-HCl, pH 7.8.

To monomerise E2 protein, isolated E2 was sonicated twice for 10 seconds in a sonication water bath and then heated at 50°C for 10 minutes. Following centrifugation at 67,500g for 1 hour supernatant was transferred to a new Eppendorf. E2 concentration is quantified with the Pierce™ BCA Protein Assay Kit (Thermo Scientific).

2.6.4. Motility assay

Swimming behaviour of cells was ascertained by measuring the zone of motility which arose on a motility agar plate following inoculation with a single colony. A single colony was taken from an LB agar plate using a sterile pipette tip and inoculated into a motility assay agar plate (1g LB agar in 100mL dH_2O for a final concentration of 0.04% agar). Plates were incubated at 30°C in a stationary incubator alongside a flask of water to ensure humidity. The radius of the zone of motility was measured using a ruler at time points and plates returned to the incubator.

2.6.5. Flagella expression assay

The levels of assembled flagella, was established by shearing flagella from cells by vortexing before running on Coomassie gel. The concentration of flagellin was ascertained by densitometry of the Coomassie stain or following Western blot and probing with anti-H48 (FliC) antibody and a suitable secondary antibody. In brief, Cells were incubated in 5mL LB media overnight at 30°C, 50rpm. OD_{600} was measured; cells were centrifuged for 10 minutes, 3500rpm, 4°C and resuspended in 0.5mL 100mM Tris-HCl, pH 7.8 per OD unit to normalise for variation in OD_{600} in cell cultures. Cells underwent 6 x 10 second rounds of application to the

vortex with 30 seconds incubation on ice in between, to shear flagella from cells. The suspension was transferred into Beckman high speed centrifuge tubes (Beckman Coulter) and centrifuged at 67500g for 1 hour at 4°C. The supernatant was removed and the pellet resuspended in the residual supernatant. The centrifugation step was repeated, the supernatant removed and the pellet resuspended in 50µL 100mM Tris-HCl, pH 7.8.

2.6.6. MUB assay for cutinase activity

The 4-methylumbelliferyl butyrate (MUB) assay was used to measure the activity of secreted cutinase. 10mL cell cultures of cells expressing cutinase were induced with 0.05mM IPTG and grown to OD_{600} 1. Supernatant was prepared by centrifugation (15 minutes, 13,000rpm, 4°C) of 1mL cell culture sample followed by the removal of 900µL supernatant. The use of sterile filtration as an alternative was also tested, however it was no more effective at removing cells than centrifugation, therefore was abandoned. LB broth was prepared in tandem to facilitate the removal of background interference by subtraction. Supernatant samples were normalised by OD_{600} with the addition of LB broth.

Solutions of MUB were prepared fresh daily. 250mM MUB was dissolved in dimethylformamide (DMF) with 1% Triton X-100. This mixture was then diluted in 0.05M phosphate citrate (pH5) buffer to a concentration of 500µM MUB. A 500µM solution of 4-methylumbelliferone (4-MU) was also prepared through the same method, to serve as a positive control. 40µL prepared supernatant was added to 96 well plate wells and the reaction was initiated with the addition of 160µL 500µM MUB. In addition to wells which contained supernatant and MUB substrate –wells were also filled with (1) LB and substrate, (2) substrate only, (3) LB and phosphate citrate buffer, (4) phosphate citrate buffer only and (5) 4-MU to aid the removal of background interference through subtraction and to serve as a positive control. 96 well plates were gently rocked for one minute to facilitate mixing of solutions and then incubated at 30°C for 30 minutes.

Following incubation, fluorescence was measured, either by imaging 96 well plates under UV light (G:BOX, Syngene) or by quantification using a Tecan Infinite 200 Pro plate reader (Tecan Group Ltd.). Based on the manufacturer instructions and corroboration following excitation and emission scans, measurements of fluorescence were recorded at an emission setting of 446nm, following excitation at either 302nm or 365nm. If 365nm excitation was used 50µL

reaction mixture was quenched with 100µL 0.1M sodium carbonate buffer, pH 10.5 prior to measurement of florescence. Buffering was required as fluorescence measured at this excitation wavelength is pH dependent.

Chapter 3. Strain improvements for increased secretion capacity of native substrates through the T3SS

This study aimed to maximise protein secretion through a modified *E. coli* flagella type III secretion system (FT3SS) through a series of gene modifications. High capacity directed secretion of protein through the FT3SS could result in a supernatant fraction comprised of secreted protein which is both at a concentration which is competitive with that achieved in industry and also free of contaminant protein. For example the highest concentration of extracellularly secreted recombinant protein in *E. coli* to date was 12mg L^{-1} (Majander et al., 2005), however this was only 50% pure. The highest yield of protein secreted to the periplasm is around 1g L^{-1}, however this requires processing to purify protein from the cell and then from periplasmic protein (Matos et al., 2012). While the incubation method was optimised to ensure protein was expressed to a satisfactory level, the main focus of the study focused was on strain based improvements to protein production and secretion. To assess the performance of mutant strains, assays were used to measure flagella gene expression and protein secretion. The initial focus was to investigate secretion of the native FT3SS substrate, the flagellin (FliC) monomer. It is rational to begin investigating the amenability of a modified FT3SS for protein secretion with the native substrate, firstly because it is sensible to modify one variable at a time to simplify any troubleshooting that may be required and secondly because of the high affinity of the FT3SS for FliC protein secretion. As described FliC is the most prevalent FT3SS substrate; filaments are comprised of up to 30,000 FliC monomers, equating to a secretion capacity of around 1000 subunits per minute (Erhardt et al. 2010). In wild type *E. coli* FliC monomers assemble at the tip under the cap protein FliD, which aids polymerisation. However with modifications it is also possible to secrete FliC monomers through the FT3SS into the media without polymerisation; these include removing the FliD cap protein or either of the hook filament junction proteins (FlgK or FlgL) (Ikeda et al. 1987; Homma et al. 1984). Flagella filament length is not regulated, furthermore if flagella are mechanically broken, then additional FliC monomers assemble at the tip, resulting in regrowth (Homma & Lino, 1984). These characteristics can be exploited to result in the continual secretion of monomers through the FT3SS into the media.

3.1. Secretion platform for secretion of native protein

The study first required the construction of a modified FT3SS to allow efficient transport of protein monomers through the structure into the media. This also required the production of an inducible plasmid based expression system harbouring the flagellin protein tagged for FT3SS secretion.

3.1.1. Construct for flagellin secretion

To effectively measure secretion capacity of flagellin in mutant strains, the expression of flagellin was decoupled from the native flagella gene expression pathway. Additionally a deletion of 89 amino acids was implemented from the central region of the flagellin monomer, resulting in the protein known as E2 (Figure 3. 1). E2 is a synthetic test secretion construct which describes a flagellin like protein, which is encoded by the *E. coli* flagellin gene *fliC* Δ191-280 that lacks residues 191-280 in the central antigenic region of the construct but assembles a fully functional flagella monomer swimming at approximately 80% of wild-type levels (personal communication, G Stafford). E2 is similar to full-length native *E. coli* flagellin; it is readily secreted and assembles into filaments and therefore is an excellent model protein for testing secretion. While it is referred to as a 'native' protein, to distinguish it from entirely recombinant protein, it is a native variant protein and therefore also is a good test protein to assess the potential to utilise this system to secrete a range of heterologous proteins. Furthermore as the E2 protein is 13kDa smaller than native flagellin, the two are distinguishable and so any possibility of the strains to revert to wild-type or contamination of cultures can be distinguished by a marked size difference in SDS-PAGE and Western blots using a FliC specific antibody. The modified flagellin protein was cloned into the pTrc99a plasmid, which harbours a gene for ampicillin resistance and allows IPTG induced control of expression of the protein, via the *trc* promoter. The 47 amino acid N-terminal signal peptide of flagellin is also present, as this is thought to act as a non-cleavable signal peptide and therefore confers targeted FT3SS export protein, but the 5' UTR is absent (Dobó et al., 2010).

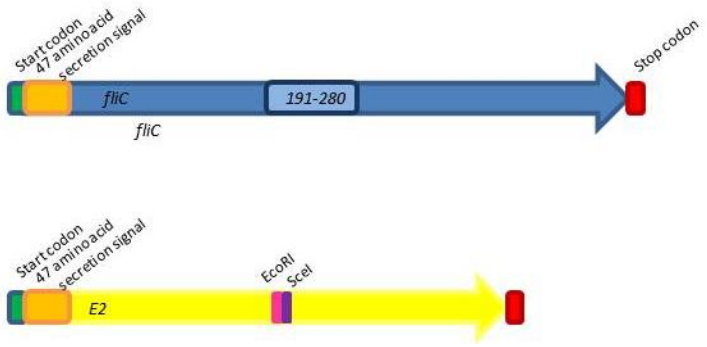


Figure 3. 1: Schematic of the native *E. coli fliC* gene and the E2 variant

Common features of the two open reading frames include the start and stop codons, the 47 amino acid FliC secretion signal and the majority of the FliC gene. Transcription and translation of the *fliC* genetic construct (top) will result in the formation of full FliC monomer. Residues 191-280 are highlighted in the *fliC* gene. In the E2 secretion construct (bottom) residues 191-280 are removed and replaced with an *Eco*RI and *Sce*I site –this will allow further modification of the construct to harbour 'cargo'. Expression of this construct will result in a truncated FliC monomer.

3.1.2 Platform strain for flagellin secretion: HAP-less and CAP-less flagella

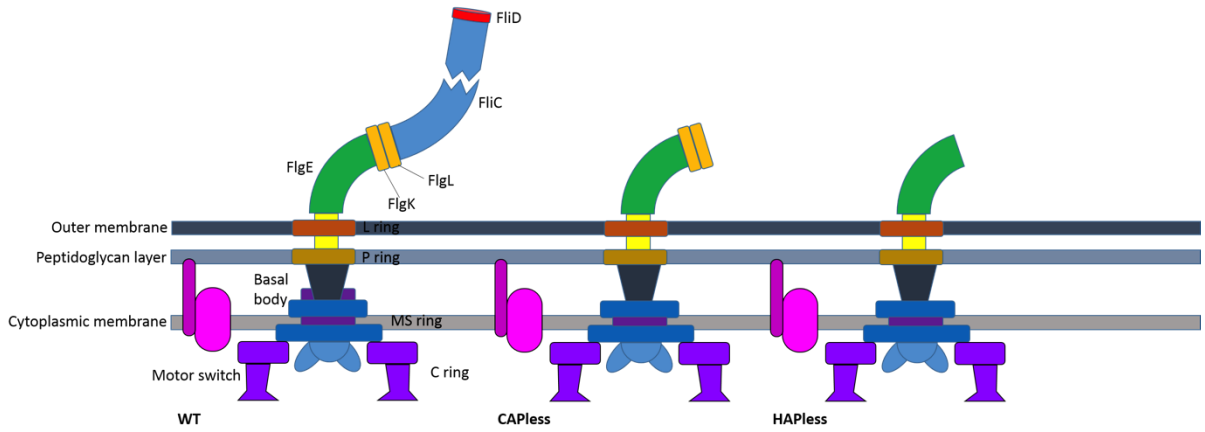


Figure 3.2. The structural differences in wild type flagella type III secretion system and the CAPless and HAPless (secretor strain) mutants.

The wildtype (left), Δ*fliCD* (middle) and Δ*fliC* Δ*flgKL* (right) FT3SS structures are shown. Protein components of the wildtype structure are labelled to indicate the mutant strategies.

A number of mutants were created for improved secretion by the bacterial FT3SS. The rationale behind mutants was inferred from the literature; while much of this concerns *Salmonella,* the data is relevant to *E. coli* too, as the two genera are phylogenetically similar and the FT3SS highly conserved between the two, although some subtle differences exist. It is reported that the removal of FliD, FlgK and FlgL proteins in *Salmonella* results in secretion of FliC into the media in (Homma, 1984; Stafford et al. 2007; Yokoseki et al. 1995) Furthermore with the removal of FliC, recombinant proteins can be secreted through the FT3SS without competition from the native substrate (Majander et al. 2005; Narayanan et al. 2010; Singer et al. 2012). With this in mind two flagella mutant MC1000 strains (Δ*fliC* Δ*flgKL* and Δ*fliCD*) were generated (by Dr. Matt Hicks and in this study respectively) to compare secretion. These are referred to as the CAPless and HAPless strains respectively (Figure 3.). MC1000 was selected as the parent strain for a number of reasons. MC1000 is a K-12 strain, which is similar to MG1655. Like MG1655 it is well studied and requires low biosafety measures, however additional mutations to the *lac* and *ara* operons aid expression of *lac* and *ara* inducible plasmids in cells (Casadaban & Cohen, 1980). Finally MC1000 is a leucine auxotroph and therefore compliant with the requirements of the University of Sheffield Genetic Modification Committee (i.e. reducing risk of release given its auxotrophy), furthermore it also adhered to the recommendations for 'regulating synthetic biology' which were outlined in the Synthetic Biology Roadmap for the UK (http://www.rcuk.ac.uk/documents/publications/syntheticbiologyroadmap-pdf/). The CAPless and HAPless strains lack some of the proteins which would be present in flagella. Removal of FliC was essential to both strains, as FliC expression is competitive to the production and secretion of other proteins. The CAPless strain is also deficient of the FliD cap protein, which polymerises flagellin, therefore is not required in monomeric secretion. The HAPless stain has no hook junction proteins, upon which polymerisation of FliC monomers would occur, even in the absence of FliD (Homma, 1984). Rather than flagella, the idea was that the FT3SSs would now function as secretion apparatus, which secrete protein monomers from the cytoplasm directly to the extracellular environment.

3.1.2.1 PCR product generation for mutagenesis: MC1000 Δ*fliCD*

Knockout mutagenesis was implemented by Lambda Red homologous recombination into the *E.* coli chromosome using the method of Datsenko and Wanner (2000) (Figure 3.). The *fliCD* gene knockout was executed in WT MC1000. A PCR reaction was carried out to produce a DNA construct for knockout mutagenesis by homologous recombination into the *E. coli* chromosome. The primer set was designed with the following in mind: 50bp gene specific flanking sequences to the intended deletion and 20bp sequences that allow amplification of the antibiotic resistance gene and FRT cassettes which allow removal of the antibiotic cassette from the template DNA plasmid (pKD3). The PCR reaction was considered successful if a 1.1 kbp band was present when PCR reaction mixtures were run on a DNA gel. A positive result was achieved and bands of the correct size were cut out and DNA isolated.

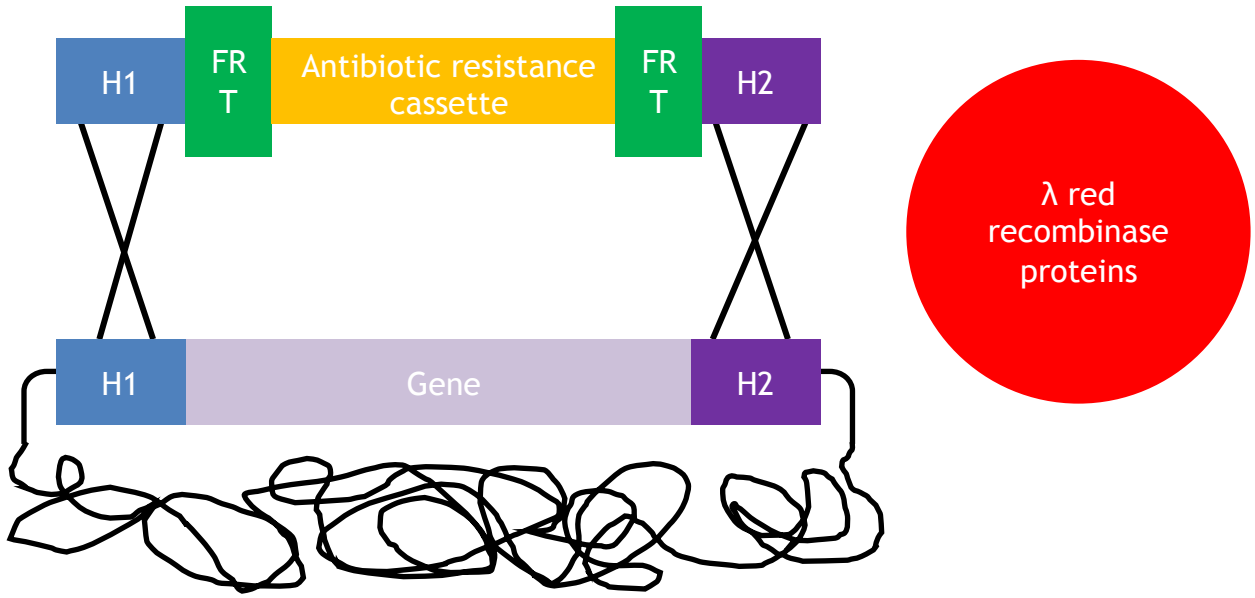


Figure 3.3: Lambda Red recombinase knock out mutagenesis.

Schematic of a cell harbouring plasmid derived Lambda Red recombinase proteins, PCR product template (top) and chromosomal DNA (bottom). The isolated PCR product construct contains the antibiotic resistance cassette, flanked with FLP recombinase target (FRT) sites, and homologous regions (H1 and H2) to chromosomal DNA adjoining the gene which will be knocked out.

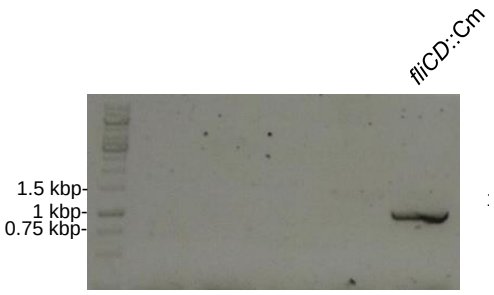


Figure 3.4: DNA gel of fliCD gene knockout construct PCR product.

PCR reactions were carried out using the templates and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and the image inverted. Blank lanes were loaded with samples relating to other experiments. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of a gene knockout construct for *fliCD* harbouring the chloramphenicol antibiotic resistance cassette (*fliCD*::Cm). The PCR product was isolated for transformation cells expressing Lambda Red recombinase.

3.1.2.2. MC1000 Δ*fliCD* mutant generation via Lambda Red recombinase method

Lambda Red recombinase induces chromosomal disruptions through the insertion of linear DNA by homologous recombination. Lambda Red recombinase enzymes are harboured in the arabinose inducible plasmid pKD46. When expressed they catalyse the conversion of linear DNA to single stranded DNA and the annealing of complementary single stranded DNA. This results in the linear DNA template being inserted into chromosomal DNA at homologous regions. The pKD46 plasmid is cured at 43°C without selection for ampicillin resistance, so that no further recombination events occur and ampicillin plasmids can be used in future studies using these strains. Chromosomal disruption is confirmed by the acquisition of antibiotic resistance encoded by the linear DNA. Once the chromosomal deletion is confirmed by PCR the antibiotic resistance cassette is removed by the activity of FLP recombinase

Following the isolation of PCR product DNA, cells were prepared for mutagenesis. The Lambda Red recombinase and ampicillin resistance harbouring plasmid pKD46 was freshly transformed into the parent strain by electroporation. Successful transformants were induced with 1mM

arabinose and grown in liquid culture at 30°C for 4 hours and then prepared for electroporation. 1µL PCR product (approximately 15ng/µL) was electroporated with 100µL fresh electrocompetant cells. Cells were incubated in 1mL LB for 1 hour at 30°C and then left at room temperature overnight. Cells were spread on agar plates the following day and incubated at 30°C.

The presence of positive colonies on chloramphenicol supplemented LB plates inoculated with cells following electroporation with *fliCD*::Cm allowed selection of successful transformants. This was coupled to the absence of colonies on the negative control chloramphenicol supplemented plate following plating of cells which had not been transformed with the knockout construct (to detect spontaneous antibiotic resistant colonies or contamination by Cm resistant bacteria). This indicates that antibiotic resistance has been acquired via homologous recombination of the PCR construct and indicates that the gene deletion has been successful. To ensure no further recombination events occurred in the Lambda Red recombinase derived mutants, positive colonies were grown in liquid culture without selection for ampicillin resistance at 43°C to cure the Red plasmid (pKD46). This was verified by the loss of ampicillin resistance after one further passage at 37°C.

3.1.2.3. MC1000 Δ*fliCD* mutant conformation by PCR

To show that successful mutants had the correct chromosomal structure, PCR reactions were carried out with chromosomal DNA from the mutant as a template and primers which anneal to chromosomal DNA upstream and downstream of the gene deletion (Figure 3.). The size of the PCR product determines whether the target gene has been deleted. A positive result indicates the presence of the antibiotic resistance cassette and removal of the gene.

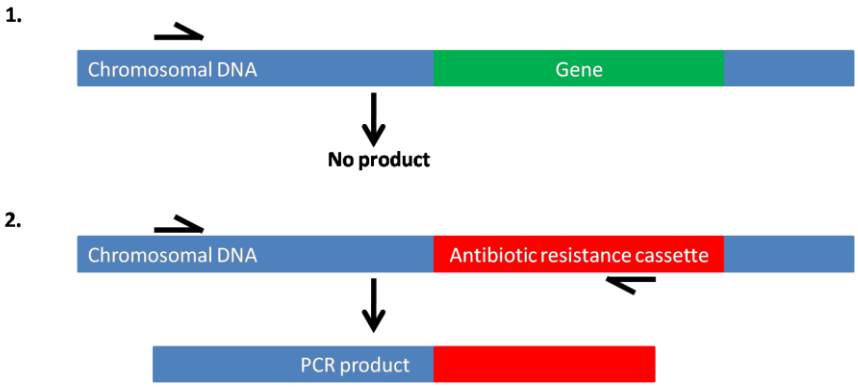


Figure 3.5: PCR verification of gene knockouts.

PCR reactions with chromosomal DNA templates from cells following Lambda Red recombineering which resulted in 1) an unsuccessful homologous recombination event 2) a successful homologous recombination event, which has resulted in gene mutagenesis and acquisition of antibiotic resistance.

A PCR product of 1.1kbp was obtained following PCR indicating that MC1000 *ΔfliCD* was successfully generated (

Figure 3.). If knockout mutagenesis was unsuccessful then a PCR product of 3.1kbp would have been observed.

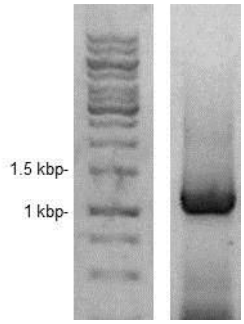


Figure 3.6: DNA gel to confirm successful gene knockout mutagenesis of the *fliCD* genes in MC1000 through homologous recombination.

PCR reactions were carried out using the templates and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and the image inverted. Samples were run with GeneRuler™ 1kb DNA ladder. Image cropped to exclude other

experiments to aid ease of viewing. Successful recombination resulted in the formation of a 1.1kbp PCR product.

Successful production and conformation of MC1000 Δ*fliC* Δ*flgKL* was achieved by M. Hicks in previous work in the laboratory but was used in this study- this was reconfirmed in a similar manner. Once established, the capacity for protein secretion and growth phenotypes were ascertained for both the HAP-less and CAP-less MC1000 based flagella variants, to observe which is more suitable for use as the platform strain for protein secretion. This required the development of an assay to accurately measure protein secretion.

3.2. Development of a standardised protein secretion assay to quantify flagellin secretion

To quantify the secretion profile of strains a secretion assay was designed to measure intra- and extracellular concentrations of specific proteins. Initial investigation into the secretion capacity via the two secretion platforms was carried out using the ptrc99a E2 plasmid which expresses the FliC *Δ191-280* construct, described previously. E2 protein was detected either by Coomassie stain or a flagellin (H48) antibody and a HRP linked secondary. In conjunction with E2 detection, an E2 flagellin monomer protein standard was prepared so that the yield of both intracellular and secreted protein could be calculated.

3.2.1. Isolating monomeric flagellin to quantify secretion capacity

To determine quantitative measurements (w/v) of protein expression, a flagellin like (E2) protein standard of known concentration was derived. Once a w/v unit measurement was quantified this allowed intra- and inter-experimental comparisons of secretion profiles. In addition the E2 protein standard serves as a positive control in both Coomassie stain and Western blot based experiments. Secreted E2 protein was purified from a flagellin deficient (Δ*fliC*) strain expressing the E2 harbouring plasmid. Following growth of cells on semisoft agar

plates, cells were scraped off and subject to agitation in a lab blender to shear flagella from cells. Following isolation and monomerisation by sonication and heating to 50°C, E2 protein was present in the soluble fraction (Figure 3.). Native *E. coli* flagellin is a 52 kDa protein (Daniell et al., 2003) however as the E2 protein is encoded by a shortened version of the flagellin gene, its expected size is 42kDa. The band of E2 protein runs at around 42 kDa on the SDS-PAGE gel.

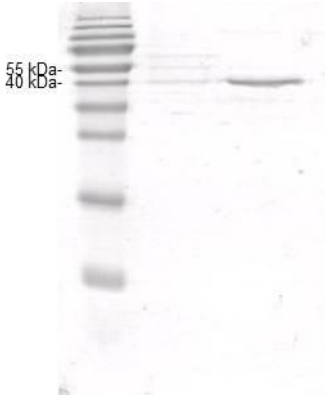


Figure 3.7: SDS-PAGE protein gel of isolated momomeric E2

E.coli MC1000 *ΔfliC* containing the plasmid ptrc99a E2 was grown, cells were harvested and monomeric E2 was isolated by shearing, sonication and high speed centrifugation, before SDS-PAGE. Lanes: (1) EZ-Run™ Prestained Rec Protein Ladder, (2) insoluble protein, (3) soluble protein.

A BCA assay was performed to quantify the concentration of E2 in the soluble fraction. The concentration of the E2 protein standard was typically in the region of 0.2-0.5µg/mL.

3.2.2. Development of a secretion assay to quantify platform strain suitability

An assay to measure both the intracellular and extracellular concentration of E2 protein directed for secretion through the modified flagella appendages is detailed here. The pTrc99a E2 plasmid was transformed into *E. coli* strains by heat shock transformation. Following transformation and conformation by antibiotic resistance acquisition, E2 expression was

induced with 0.05mM IPTG as this concentration had been shown to be effective in the production of ptrc99a harboured proteins in the laboratory. Cells were incubated at 37°C, 180rpm. Once cells reached OD_{600} 1.5 (stationary phase) 1.0 OD unit of cells (i.e. 1ml at OD 1) were collected and then prepared for SDS-PAGE of the supernatant and cells to isolate secreted proteins. Cells harbouring the empty pTrc99a plasmid were also grown as a negative control. The E2 protein standard was also loaded (diluted in 2x SDS loading buffer 2:8 for SDS-PAGE and 1:9 for Western Blot SDS-PAGE). Western blots were also carried out. Nitrocellular membranes were incubated in 1x TBS/0.1% TWEEN with 1 in 1000 anti-FliC antibody (H48) as a primary antibody and 1 in 3000 anti-rabbit HRP as a secondary. The antibody was generated against full length flagellin, however adequate antigenic affinity is conferred by the E2 protein. Furthermore while the antibody is intended for use as an agglutination agent for screening for the presence of clinical *E. coli* strains, this work established that it can also be used effectively in Western blotting experiments. Densitometry of each band on the Western blots or Coomassie stained electrophoresis gels was measured using ImageJ. Densitometry values were then calibrated to the E2 protein standard to calculate the concentration of secreted and intracellular E2 in the two prototype platform strains for flagella secretion.

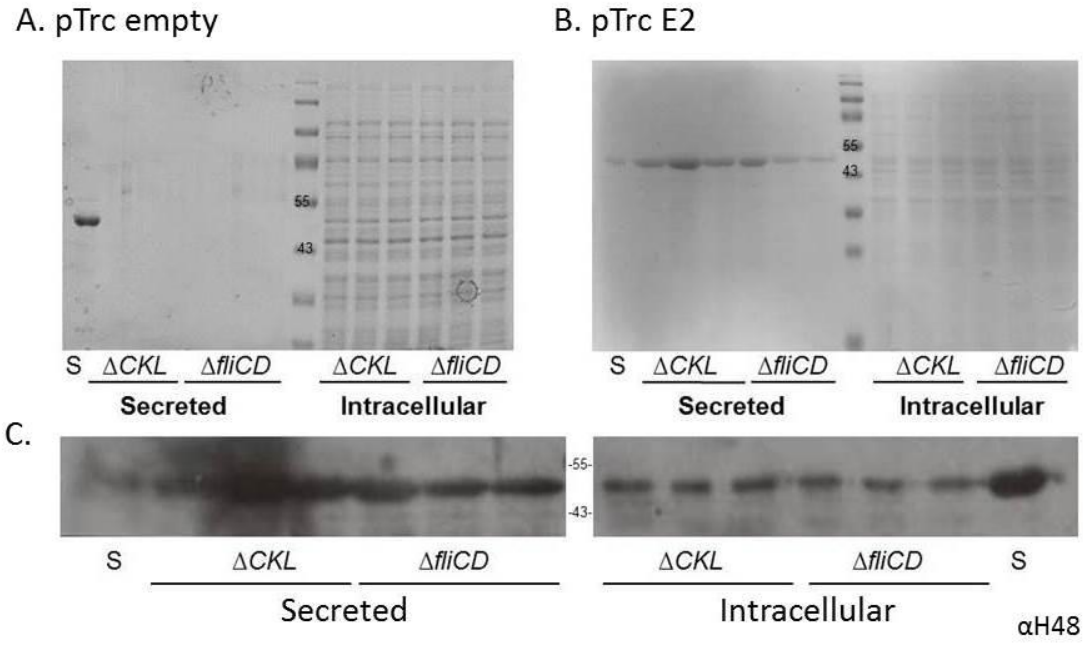


Figure 3.8: Intracellular and secreted fractions of MC1000 Δ*fliC* Δ*flgKL* and Δ*fliCD*.

E.coli ΔflgKL ΔfliC (Δ*CKL*) or Δ*fliCD* containing either the plasmid pTrc empty or pTrc E2 was grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.5. 1 OD unit of cells were prepared for SDS-PAGE and the supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western blot analysis. An E2 protein standard (S) was loaded to allow quantification of intracellular and secreted protein concentration. Coomassie stained SDS-PAGE of cells harbouring (A) empty pTrc99a plasmid and (B) pTrc99a E2. Intracellular: 5µL; secreted: 15µL; standard 5 (left), 10 (right). (C) Western blot with αH48 and appropriate secondary antibody following SDS-PAGE of pTrc99a E2 positive cells. Intracellular: 2µL; secreted: 15µL; standard 1 (left), 2 (right). Three biological repeats for each strain and plasmid combination.

Both strains demonstrated a similar intracellular protein expression pattern (Figure 3.A and Figure 3.B), suggesting that both are viable strains. Protein expression of E2 was similar in both strains (Figure 3.C). E2 protein is visible in the supernatant of cells expressing the protein (Figure 3.C), this suggests that the E2 protein is being successfully targeted for secretion through the flagella variants of both platform strains. In support of this cell lysis was negligible in both strains and for both plasmids. This is apparent when observing the lack of protein (aside from E2) in the secreted fraction of media (Figure 3.A and Figure 3.B). The results of the

Western blot shown here highlight the issues that can arise when using this method, as a high background signal is observed in Figure 3.C, which masks the true immunogenic derived signal. Densitometry based analysis was carried out on images (where protein bands were not obscured) using ImageJ. A worked example is given (Figure 3.).

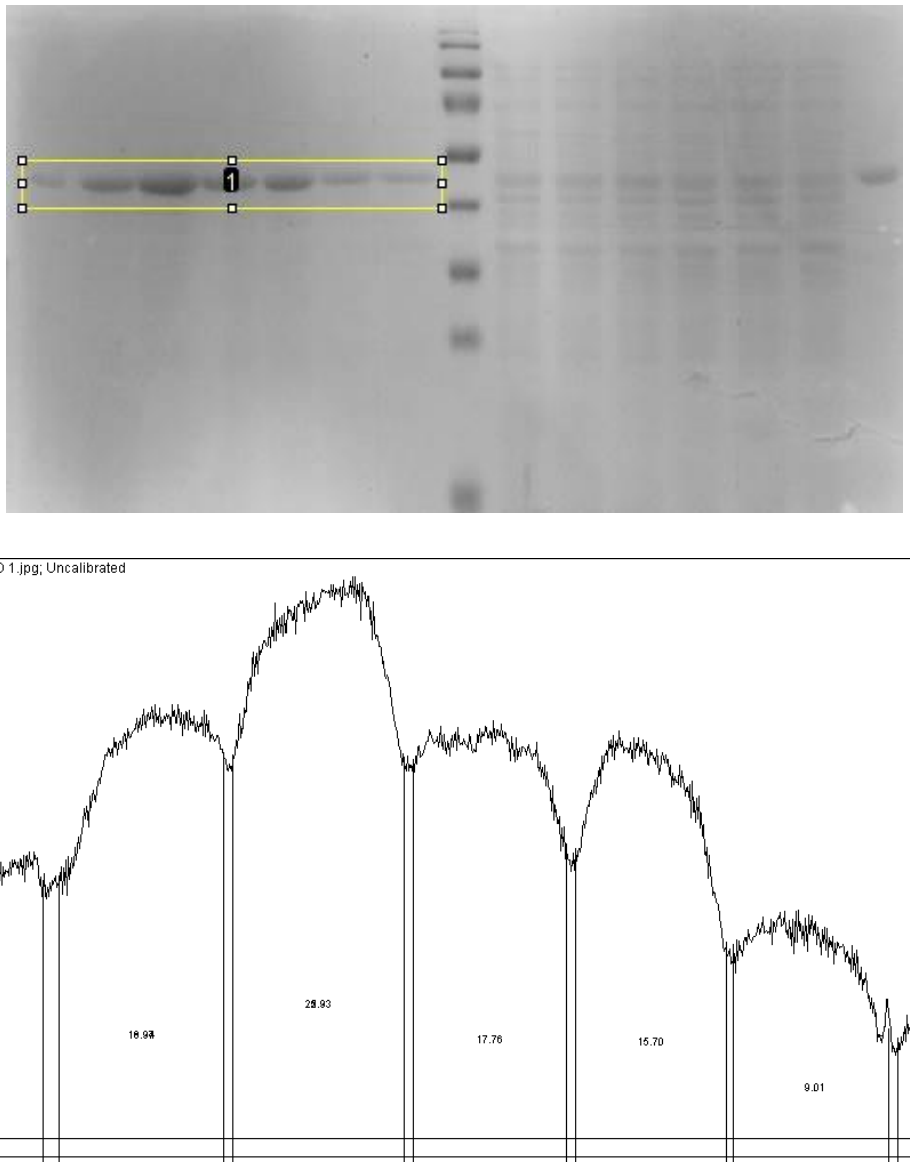


Figure 3.9: Worked example of densitometry analysis of secreted E2 from a Coomassie stained SDS-PAGE (Figure 3.B).

A horizontal box is drawn around the protein of interest, ImageJ then plots the densitometry outline for this selection. The lanes are divided up and using the image for reference. Gaps are left to show the zones of separation between lanes. Individual lanes are then plotted and densitometry values derived for each.

Using the values derived by the route given in Figure 3., and with knowledge of how much E2 standard was loaded onto the SDS-PAGE, the exact amount of E2 present in each lane can be derived. This can then be extrapolated to calculate the amount of protein present in both the supernatant (from both Coomassie stained SDS-PAGE and Western blot analysis) and intracellular fractions (from Western blot analysis only). When possible, Coomassie stained SDS-PAGE gels are preferentially used as there are fewer experimental steps to obtain this, therefore the possibility of experimental error is reduced, further to this a linear range of detection is insured, whereas this is not always true during Western blotting. Figure 3. shows the results derived from the Coomassie stained supernatant samples. Following 3 biological replicates per mutant strain, a higher concentration of E2 was observed in the supernatant of all Δ*CKL* cultures, in comparison to Δ*fliCD*. The highest concentration of E2 in Δ*CKL* and Δ*fliCD* cultures was 9.19 and 5.57 mg L^{-1} respectively. When combined, (Figure 3.) the concentration of secreted E2 was significantly higher in Δ*CKL* cultures.

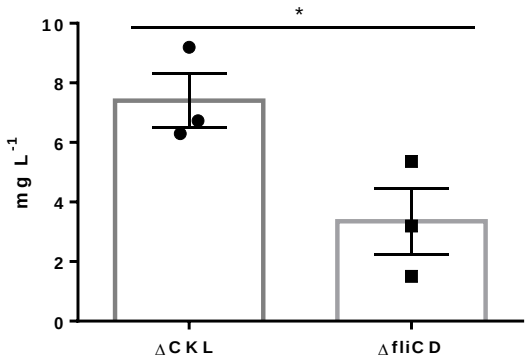


Figure 3.10: Results following densitometry analysis of the supernatant samples of the Coomassie stained SDS-PAGE seen in Figure 3..

E.coli ΔflgKL ΔfliC (Δ*CKL*) or *ΔfliCD* containing the plasmid pTrc E2 was grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.5. The supernatant from 1 OD unit of cells were prepared for SDS-PAGE and Coomassie stain. An E2 protein standard (S) was loaded to allow quantification of secreted protein concentration following densitometry analysis using Image J. Average from combined concentration of secreted E2 protein per Litre of cell culture. 3 biological replicates per mutant strain, grown on the same day. Standard error of the mean and result of two way t-test shown (* = $p<0.05$)

Densitometry analysis of the Western blot (Figure 3.C), found that there was a significantly higher concentration of E2 in Δ*CKL* cells. On average almost twice as much E2 was located in Δ*CKL* cells (91.36 mg L^{-1}) in comparison to Δ*fliCD*. Secreted and intracellular concentrations of E2 were paired for each biological replicate to calculate the percentage of total E2 which was secreted. On average Δ*CKL* was more efficient at secreting E2 (7.92%) than Δ*fliCD* (5.88%).

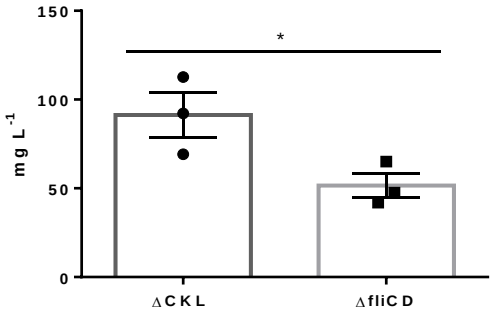


Figure 3.11: Results following densitometry analysis of the intracellular samples of Western blot seen in Figure 3..

E.coli ΔflgKL ΔfliC (Δ*CKL*) or Δ*fliCD* containing the plasmid pTrc E2 was grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.5. The cells from 1 OD unit of cells were prepared for SDS-PAGE and Western blot analysis using anti-flagellin (H48) antibody and an HRP secondary. An E2 protein standard (S) was loaded to allow quantification of intracellular E2 protein concentration per Litre of cell culture, following densitometry analysis using Image J. 3 biological replicates, grown on the same day. Standard error of the mean and result of two way t-test shown (* = $p<0.05$)

3.2.3. Growth phenotype of secretion platform strains

The deletion of genes often correlates to a change in growth phenotype. This may result in an increased rate of growth if genes are removed which previously resulted in the production of a protein which was either metabolically costly to produce or resulted in a metabolically costly function. Conversely a decreased rate of growth may occur if the mutation is detrimental or if metabolic burden is in fact increased by the antibiotic resistance cassette acquired during knockout mutagenesis. To ascertain whether this was true of either candidate platform

secretion strain, growth was measured over time to visualise this. Cells were grown in 200µL LB cultures in a 96 well plate in a TECAN plate reader at 37°C with constant shaking (6mm orbital).

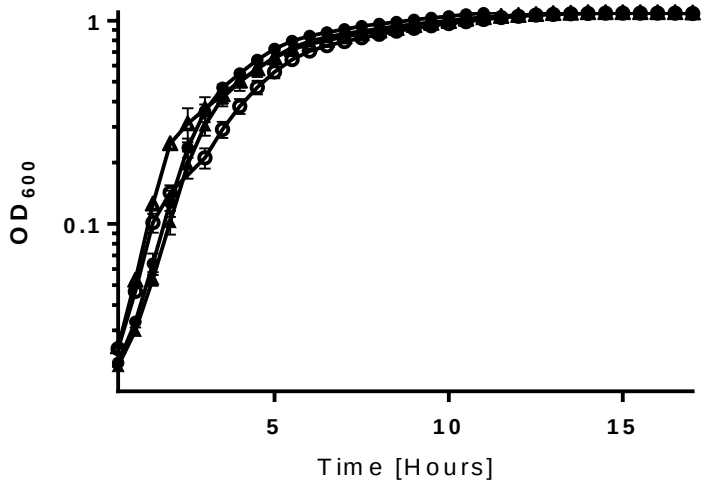


Figure 3.12: Growth of candidate platform strains for flagella mediated protein secretion over time in a 96 well plate. Calibrated to LB media.

Starter cultures of cells expressing either pTrc empty or pTrc E2 were supplemented with ampicillin and 0.05mM IPTG and aliquoted into 96 well plate wells. The 96 well plate was incubated in the TECAN plate reader at 37°C with 6mm orbital shaking. OD_{600} measurements were recorded every 30 minutes. From three technical repeats of two biological repeats. *ΔCKL* + pTrc E2: ▲, *ΔCKL* + pTrc empty: △, Δ*fliCD* + pTrc E2: ●, Δ*fliCD* + pTrc empty: ○.

The optical density of a solution is calculated by spectroscopy and based on the Beer-Lambert law. The principle of optical density relies on passing wavelengths of light through a solution and measuring the intensity of emitted light. The composition of the solution determines how much light is emitted –or scattered in the case of a bacterial suspension (Hall et al., 2014). It is therefore dependent on both the composition of the solution (density of culture) and the path length of the solution which the light passes through. In a standard cuvette, this is 1cm – however when measuring the absorbance of a solution in a 96 well plate this will differ, thus giving an incorrect measurement of optical density. Using a method for path length correction the path length of 200µL bacterial culture was found to be 5.95mm

(https://tools.thermofisher.com/content/sfs/brochures/AN-SkanIT-Microplate-Based-Pathlength-Correction-Technical-Note-EN.pdf). While using the Beer-Lambert equation to calculate the precise adjustment multiple would be more accurate this required knowledge of the molar absorptivity for a bacterial suspension in LB media. As a reliable value for this could not be located in the literature an alternative method was chosen where the OD_{600} output of the same bacterial cultures in both a 1mL cuvette and a 96 well plate well was measured. Using a multiplication factor of 2.19 the OD_{600} acquired from a 96 well plate could be adjusted to give an accurate measurement of optical density obtained from a 1mL cuvette. While this calculation was not made routinely, when applied to results this adjustment gives OD_{600} readings which are more in line with those we are accustomed to, in terms of growth phase - for example final OD_{600} is around 2.0.

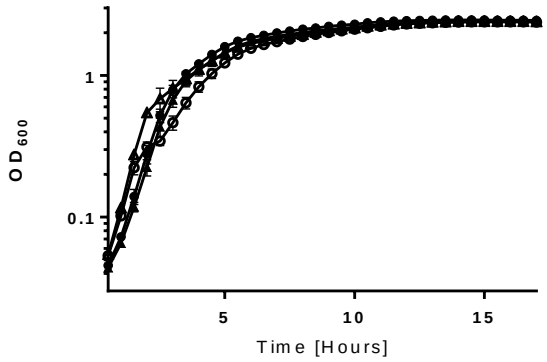


Figure 3.13: Growth of candidate platform strains for flagella mediated protein secretion over time in a 96 well plate. Calibrated to LB media and OD_{600} adjusted for path length.

Starter cultures of cells expressing either pTrc empty or pTrc E2 were supplemented with ampicillin and 0.05mM IPTG and aliquoted into 96 well plate wells. The 96 well plate was incubated in the TECAN plate reader at 37°C with 6mm orbital shaking. OD_{600} measurements were recorded every 30 minutes and multiplied by 2.19. *ΔCKL* + pTrc E2: ▲, *ΔCKL* + pTrc empty: △, Δ*fliCD* + pTrc E2: ●, Δ*fliCD* + pTrc empty: ○

The results of growth curve experiments (Figure 3., Figure 3.) show that the initial rate of growth is higher in both strains when expressing empty vector (initial (0.5 to 1.5 hours) mean generation time from 42 to 70 minutes for *ΔCKL* and from 49 to 61 in Δ*fliCD*). However in the late log phase Δ*fliCD* reaches a higher OD at time point 2.5 when expressing E2 protein in

comparison to empty vector (statistically significant: p = <0.0001 –paired t-test of average OD_{600} for each time point from 2.5 to 5.5 hours). Cells expressing E2 reach early stationary phase first. When expressing empty vector, Δ*CKL* consistently displays a higher optical density than Δ*fliCD* in the log stage (p = <0.001 for 0.5 to 5.5 hours). When expressing E2 protein the two strains grew at a similar rate, however Δ*fliCD* has a higher OD_{600} in late log-early stationary phase (also p = <0.0001 for 3.5 to 7.5 hours). Overall neither mutation resulted in cells which were not viable or markedly altered in growth characteristics.

3.2.4. Selection of secretion platform strain

Observations show that while the growth of Δ*CKL* expressing E2 protein is slightly poorer than Δ*fliCD*, Δ*CKL* expresses and secretes (2.16 and 1.77 times more E2 respectively) significantly more E2 than Δ*fliCD*. In addition it is on average 1.35 times more efficient at secreting E2. Based on this the HAP-less *CKL* strain was selected as the platform strain for this work. This strain will form the basis of all further secretion work and improvements. This MC1000 Δ*flgKL* Δ*fliC* strain may be referred to as Δ*CKL*, the (original) secretor strain or (original) secretion apparatus strain.

3.3. Optimisation of a standardised protein secretion assay to quantify flagellin secretion

While the main focus of this work is to improve protein secretion through the flagella type III secretion system through a number of strain based improvements, it was necessary to first optimise the protein secretion assay to ensure initial protein secretion is satisfactory, this will both ensure protein secretion is measurable and have the added benefit of improving protein secretion capacity somewhat prior to the strain based strategy. Further to this is was necessary to ensure that cultures were grown for an adequate amount of time prior to harvesting to ensure that the amount of E2 detected was representative of the maximum output of cells (i.e. secretion has occurred and proteolysis is low).

3.3.1. Refining induction of expression vector

As described, ptrc99a harbouring E2 plasmid is IPTG inducible. E2 protein expression was tuned with IPTG concentration to maximise protein secretion in the Δ*CKL* strain. E2 protein expression and secretion were measured as before, however cultures were supplemented with 0, 0.05, 0.1, 0.5 or 1mM IPTG.

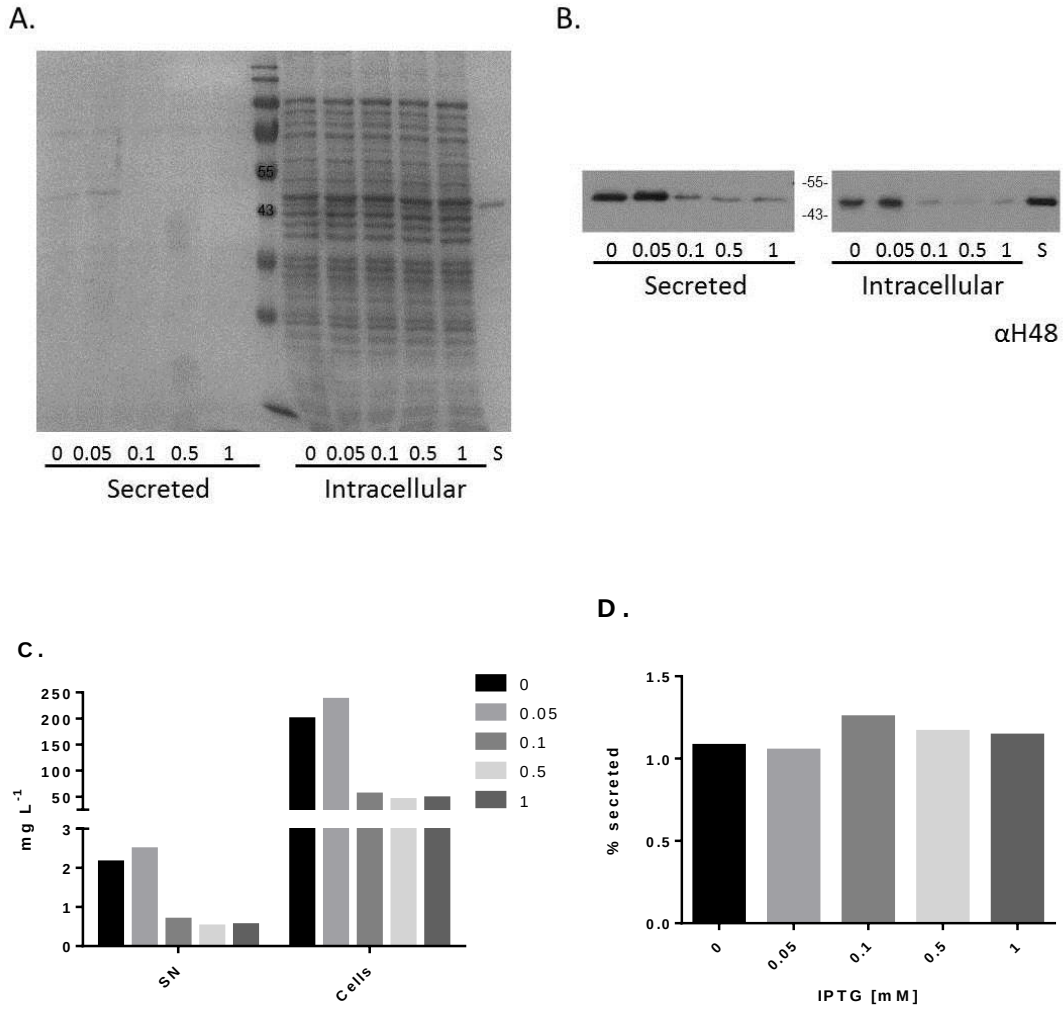


Figure 3.14: Intracellular and secreted fractions of MC1000 Δ*fliC* Δ*flgKL* supplemented with different concentrations of IPTG.

E.coli ΔCKL containing the plasmid pTrc E2 was grown in LB supplemented with 0, 0.05, 0.1, 0.5 or 1mM IPTG. Cells were harvested at OD_{600} 1.5. The cells and supernatant from 1 OD unit of cells were prepared for (A) SDS-PAGE and (B) Western blot analysis of supernatant using anti-flagellin (H48) antibody and an HRP secondary. An E2 protein standard (S) was loaded to allow

quantification of intracellular and secreted protein concentration following densitometry analysis using Image J. (C) Concentration of intracellular and secreted E2 protein per mL cell culture. (D) Secreted protein as a percentage of total protein (both secreted and intracellular).

The absence of proteins in the secreted fraction of protein on the Coomassie stained SDS-PAGE (Figure 3.A) shows that cell lysis is not evident at any concentration of IPTG. E2 protein is visible in the secreted fraction of cells where either 0 or 0.05mM IPTG was added to media, however none is visible for increased concentrations of IPTG, suggesting that the concentration of E2 secreted was lower than the detection threshold of the Coomassie stain. This is confirmed in the corresponding Western blot (Figure 3.B), where it is observed that more E2 protein is both expressed and secreted with the addition of 0 or 0.05mM IPTG. Densitometry (Figure 3.C) revealed that the addition of 0.05mM IPTG resulted in 236mg L^{-1} E2 accumulating intracellularly and 2.48mg L^{-1} being secreted into the culture media. The next most favourable IPTG concentration for protein expression and secretion was to supplement cells with 0.00mM IPTG. The addition of 0.1mM IPTG or more resulted in a reduction in E2 protein expression and secretion; high IPTG can result in cell lysis however this was not evident here. It has previously been reported that induction of pTrc99a with IPTG concentrations in excess of 0.1mM resulted in reduced protein production and the formation of inclusion bodies, which could be subject to proteolysis (Choi et al., 2006; Jürgen et al., 2010). The results from densitometry were also used to assess the secretion capacity of cells (Figure 3.D). While variation was seen in the amount of protein expressed and secreted, the amount of protein secreted as a percentage of total protein expressed and secreted is consistent with all concentrations of IPTG, with around 1% total E2 protein produced by the cell being secreted. This suggests that the flagella type III secretion system may have a maximum secretion capacity and that while increased expression will result in increased secretion, secretion capacity will limit the amount of protein which is secreted. This theory will be investigated throughout this project.

Plasmid protein expression was achieved without IPTG induction- this is because the pTrc99a vector is known to exhibit leaky expression of protein (Greer-Phillips et al., 2003). With IPTG concentrations above 0.1mM both protein expression and secretion were reduced. The amount of E2 both expressed and secreted by cells supplemented with 0.1-1mM IPTG was very similar, suggesting that E2 production can become saturated with higher concentrations of

IPTG induction. As 0.05mM IPTG resulted in both the highest level of expression and secretion of E2 protein, cells were supplemented with this concentration of IPTG throughout protein secretion assays concerning the ptrc99a E2 vector.

It is evident that both the expression and secretion capacity of Δ*CKL* is variable. In the results derived for the data shown in Figure 3. there was almost three times the concentration of E2 in the supernatant and just over twice the level of the intracellular E2 compared to Figure 3. (0.05mM IPTG only). Furthermore it was calculated that cells were almost 8 times more efficient at secreting E2, than those in Figure 3.. Possible reasons for this will be discussed later.

With the protein secretion assay optimised to ensure that E2 expression and secretion are maximised, the focus of the study can now shift to the production of mutant strains with increased secretion capacity.

3.4. Improved secretion: master regulator repressor mutants

3.4.1. Rationale for mutant generation: *lrhA, dksA, clpX*

As explained in (Figure 1.6) flagella gene expression is organised into a pyramid like hierarchal structure, with all genes either directly or indirectly under the control of the master regulator complex, $FlhD_4C_2$. As a fulcrum point of control, the master regulator complex serves as an excellent point to alter flagella gene expression. It is hypothesised that with increasing $FlhD_4C_2$, this should in turn increase the number of secretion apparatus expressed and therefore the amount of protein secreted through the FT3SS. Many environmental signals cause transcriptional repression of *flhDC* gene expression; furthermore the resulting $FlhD_4C_2$ protein complex is actively degraded by ClpXP protease. It is hoped that removing transcriptional repressors and inhibiting protease targeting of $FlhD_4C_2$, may result in increased flagella gene expression. The strategies to implement this are explained below.

3.4.1.1. *lrhA*

The literature reports that *lrhA* knock out mutants exhibited higher activity of flagella class I and II gene promoters (Lehnen et al., 2002) where *flhDC* promoter activity was 3.5 x higher in the *lrhA* knockout mutant. If *lrhA* is removed from the secretor strain this may result in more secretion apparatus being present, this may result in increased protein secretion into the media.

3.4.1.2. *dksA*

Previous studies conclude that *dksA* knock out mutants exhibit higher activity of flagella gene promoters for all three classes of flagella proteins (Lemke et al., 2009). *lacZ* reporter gene fusions show that *flhDC* promoter activity was 2.3 x higher in log phase and 9 x higher in stationary phase. Higher gene expression can be extrapolated to more secretion apparatus and therefore more secretion, particularly throughout the stationary phase, where both motility and flagella gene expression has previously been found to decline (Amsler et al., 1993).

3.4.1.3. *clpX*

Along with ClpP, ClpX forms the ClpXP protease complex. While ClpP is the proteolytic component of the complex, ClpX is the ATPase which confers substrate specificity –one of which is $FlhD_4C_2$. ClpX has been shown to actively bind to and denature $FlhD_4C_2$, before ClpP degrades the protein, resulting in reduced flagella gene expression. Δ*clpX* mutants were shown to accumulate $FlhD_4C_2$, the complex had increased half-life and cells were more motile (Kitagawa et al., 2011). Removal of *clpX* has also been shown to result in decreased degradation of $FlhD_4C_2$, but also increased secretion of the anti-sigma factor FlgM, which represses flagella class III transcription (Guo et al. 2014). The removal of *clpX* resulted in a 'greatly increased' concentration of extracellular FliC and increased flagella gene expression (Tomoyasu et al., 2002). Therefore removing Δ*clpX* from the secretor strain may result in increased flagella gene expression and more secretion. The protease ClpP also interacts with ClpA, which targets ClpP to different proteins (Gottesman, 1996), therefore ClpP will remain in the genome so that these cellular processes are not altered.

3.4.2. Mutant generation: *lrhA, dksA, clpX*

3.4.2.1. Knockout mutagenesis via the Lambda Red recombinase method: lrhA and dksA

lrhA and *dksA* knockout mutants were derived in using the FLP-recombinase method outlined previously. The parent strain was MC1000 Δ*fliC* Δ*flgKL*. Primers were designed to amplify kanamycin cassettes with homologous ends to those up and downstream of *lrhA* and *dksA.* The template for the *dksA* mutant was pKD4 as more genes are located downstream in the operon, therefore it is important that translation signals are not disrupted. The scar that arises following knock out mutagenesis using the pKD4 template harbours an RBS site and start codon, therefore downstream gene expression of other genes in the operon will occur, as in the wildtype. The PCR product for *lrhA* was derived from the pKD13 plasmid, as it is situated at the end of the operon, therefore deletion of *lrhA* should have no downstream deleterious effects.

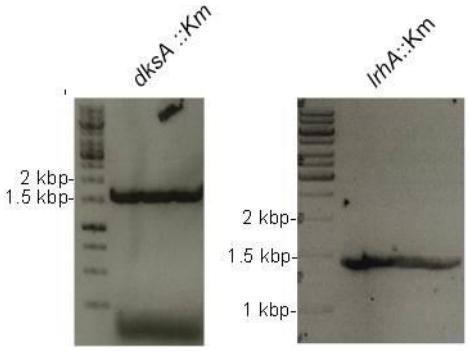


Figure 3.15: Agarose DNA gels showing PCR products to initiate gene knock out mutagenesis for (A) *dksA* and (B) *lrhA*

PCR reactions were carried out using chromosomal DNA as template and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder.

1.6 and 1.4 kbp PCR products were obtained (Figure 3.) and excised from the gel. Using electroporation these linear fragments of DNA were then transformed into arabinose induced,

pKD46 harbouring parent cells. Following antibiotic resistance screening on kanamycin supplemented agar plates, positive colonies were selected for further characterisation by PCR screening to ensure that *dksA* and *lrhA* had been successfully deleted from the chromosome. Once positive colonies were isolated the pKD46 vector was immediately cured from cells as previously described to ensure no further recombination events could occur.

3.4.2.2. PCR conformation of knockout mutagenesis: lrhA and dksA

The acquisition of antibiotic resistance is a good indicator of a successful knockout mutagenesis event; however PCR reactions were also carried out to ensure that recombination had occurred in the correct location of the chromosome. PCR reactions using chromosomal DNA and a combination of primers designed to anneal either within the kanamycin cassette or chromosomal region (up or downstream of the original location of the mutant gene) were carried out to further characterise knockout mutagenesis.

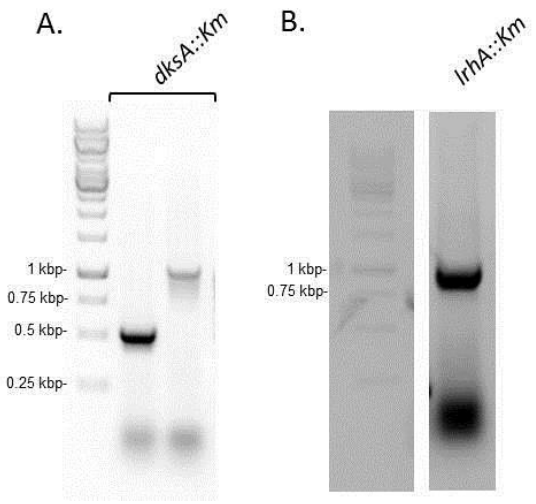


Figure 3.16: Agarose DNA gels showing PCR products to confirm gene knock out mutagenesis for (A) *dksA* and (B) *lrhA*

PCR reactions were carried out using chromosomal DNA as template and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. (1) Δ*CKL* Δ*dksA* chromosomal template. Expected product- lane 1: 548kbp, lane 2: 1109kbp. (2) Δ*CKL* Δ*lrhA* chromosomal template. Expected product- lane 1: 944kbp

Figure 3. indicates that knockout mutagenesis was successful. PCR products were obtained using primer sets in which one primer annealed to a region of chromosomal DNA and another to the kanamycin cassette. The size of the DNA products indicate that homologous recombination occurred in the correct region of chromosomal DNA.

3.4.2.3. Knockout mutagenesis via phage transduction: *clpX*

The *clpX* mutant was derived through an alternative mutagenesis method. As an MG1655 Δ*clpX* mutant was readily accessible (gifted by J. Green, University of Sheffield), it was not necessary to generate the *clpX* knockout mutant *de novo*, instead P1 phage transduction could be harnessed to transfer the kanamycin linked *ΔclpX* mutation from MG1655 into the MC1000 Δ*fliC* Δ*flgKL* parent strain to yield a *ΔclpX* mutant. In addition to this a MC1000 parent strain was also used, resulting in a MC1000 *ΔclpX* strain with viable flagella to allow motility based phenotypic characterisation of a single *ΔclpX* mutation in an isogenic background.

Following P1 phage transduction of MG1655 Δ*clpX*, liquid lysates of the donor strain were derived. The recipient strains were inoculated with the lysate, initiating phage transduction of donor genetic material to chromosomes of recipient strains. The presence of colonies following inoculation of LB plates supplemented with citrate and kanamycin by recipient cells signified successful Δ*clpX* transduction events. This was coupled to the absence of colonies on kanamycin supplemented negative control plates which were inoculated with phage or cells only. Transformants were passaged on citrate containing agar plates to eliminate the phage. The phage transduction derived mutant cells were passaged on citrate and antibiotic containing plates to ensure the phage was eliminated and no further transduction events occurred. As the *clpX* gene is not within 100,000bp of the *fliC* or *flgKL* genes, it was not necessary to confirm that these genes had been reinstated by the P1 phage during transduction, as the phage is not able to package sections of genomic DNA larger than this.

The phenotype of the Δ*clpX* mutant was verified to ensure that the mutation is functional in the wild-type MC1000 strain. As the secretor strain cells lack filaments they have no motility. Therefore the effect of Δ*clpX* on motility in wild type MC1000 compared to MC1000 Δ*clpX* on motility agar plates was assessed. The latter should display increased motility due to the increased concentration of $FlhD_4C_2$ which has been reported in *E. coli* Δ*clpX* mutants (Kitagawa et al., 2011). If this is shown in MC1000 the phenotype may extend to the secretor strain.

3.4.3. Mutant characterisation: *lrhA*

Following verification of mutants, strains were characterised. Secretion was profiled by the E2 protein secretion assay and flagella gene expression assay. Although the protein standard allows the comparison of protein secretion between different experiments, the Δ*CKL* strain was run in tandem to provide a simple visual comparison, so that the standard was not the sole point of context. In the flagella gene expression assay the Δ*CKL* strain experiments must be run in parallel to allow comparison between mutant strains against the Δ*CKL* strain. The literature reports that *lrhA* knock out mutants exhibited higher activity of flagella class I and II gene promoters (Lehnen et al., 2002). *flhDC* promoter activity was 3.5 x higher in the *lrhA* knockout mutant. Therefore it is expected that both gene expression and secretion assays will display higher performance in the Δ*lrhA* mutant strain in comparison to the Δ*CKL* strain.

3.4.3.1. Effects of deletion of *lrhA* on FT3SS secretion

Following verification of the *ΔCKL* Δ*lrhA* strain by antibiotic resistance and PCR, the secretion assay was performed. A plasmid harbouring the E2 secretion construct was transformed by heat shock into Δ*CKL* and Δ*CKL* Δ*lrhA* cells. The supernatant was analysed by Western blot (Figure 3.). Bands were present in all samples with the exception of Δ*CKL* at hour 4.

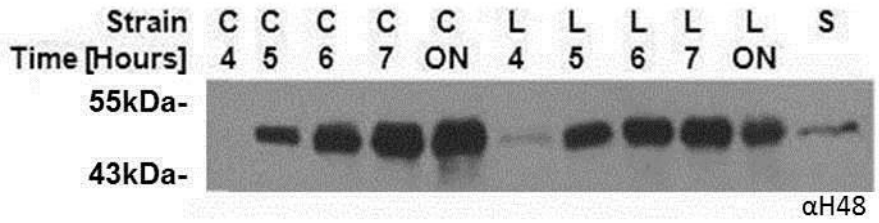


Figure 3.17: Secretion of E2 from the secretor strain and Δ*lrhA* mutant during the growth curve.

E.coli Δ*CKL* (C) or Δ*CKL* Δ*lrhA* (L) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at hourly intervals and the OD_{600} measured. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western analysis of supernatant using anti-flagellin (H48) antibody and an HRP secondary. An E2 protein standard (S) was loaded to allow quantification of secreted protein concentration.

Densitometry of each band on the Western blot (Figure 3.) was measured using ImageJ and values were calibrated to the E2 protein standard to calculate the concentration of E2 protein in samples from both strains at hour time points. Results reflect the amount of E2 secreted per OD unit, however if this value is adjusted to account for the volume removed to amount to 1 OD unit, a mg L^{-1} culture can be calculated (Figure 3.). Cell densities (OD_{600}) of the two cultures were very similar throughout the growth curve, demonstrating that there was no detrimental effect of the mutation on strains. In the Δ*CKL* strain there was no E2 secreted at 4 hours. From 5 to 24 hours the concentration of E2 in secretor strain culture media increased to a maximum of 8.742mg L^{-1}. E2 was present in the media of *ΔCKL* Δ*lrhA* after 4 hours. This continued to increase up to 7 hours where there was 6.633 mg L^{-1} E2 in the media. A decrease was observed following 24 hours growth. From 4-6 hours protein secretion was higher in *ΔCKL* Δ*lrhA* however at 7 and 24 hours more extracellular E2 was present in Δ*CKL* strain media. Extracellular E2 concentration continued to increase during the stationary phase for both strains, only decreasing after 24 hours in *ΔCKL* Δ*lrhA*.

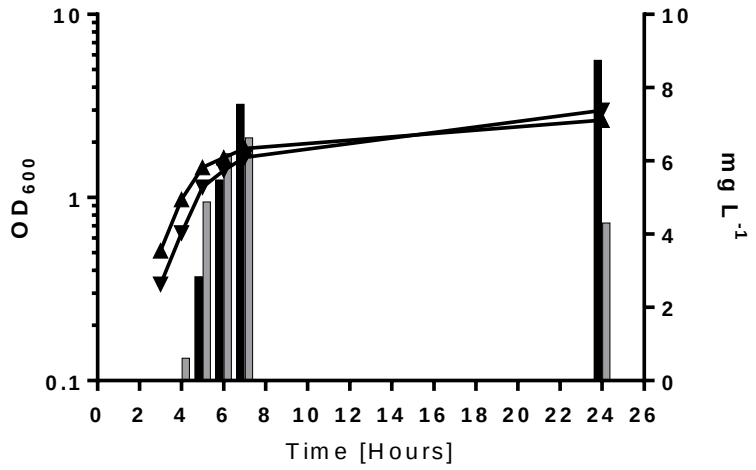


Figure 3.18: Secretion of E2 from the secretor strain and Δ*lrhA* mutant during the growth curve

E.coli ΔCKL and *ΔCKL* Δ*lrhA* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (*ΔCKL*: ▲, *ΔCKL* Δ*lrhA*: ▼, left hand axis). In parallel cells were harvested and the supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) before SDS-PAGE and Western analysis using anti-flagellin (H48) antibody and an HRP secondary. The amount of protein secreted is shown in mg L^{-1} cell culture as calculated from a standardised amount of E2 protein using ImageJ based densitometry (Δ*CKL*: black bar, Δ*CKL* Δ*lrhA*: grey bar, right hand axis)

The results shown in Figure 3. show the cumulative amount of E2 present in the media, however this is not a true representation of secretion over time. By subtracting the total protein at a given hour by that calculated for the previous a rate of protein secretion per litre of cells per hour can be determined (Figure 3.). For both strains the rate of E2 protein secretion into the supernatant increased up until 5 hours, where the rate of protein secretion was 2.834 mg L^{-1} $Hour^{-1}$ in the Δ*CKL* strain and 4.260 mg L^{-1} $Hour^{-1}$ in the Δ*CKL* Δ*lrhA* mutant. After 5 hours the rate of protein secretion began to decrease. At 24 hours the rate of protein secretion was 0.075mg L^{-1} $Hour^{-1}$ in the Δ*CKL* strain and -0.145 mg L^{-1} $Hour^{-1}$ in the *lrhA* mutant. A negative value for the rate of protein secretion is indicative of protein degradation exceeding secretion. Overall the rate of protein secretion in the LrhA mutant initially exceeded that of the Δ*CKL* strain from 4 to 5 hours. Δ*CKL* Δ*lrhA* also demonstrated the highest rate of protein secretion observed at 4.260 mg L^{-1} $Hour^{-1}$, however after 5 hours, while the rate of

secretion decreased in both strain, it decreased at a faster rate in Δ*CKL* Δ*lrhA*. At 24 hours Δ*CKL* Δ*lrhA* performed less efficiently to the Δ*CKL* strain as a negative rate of secretion was observed, indicating that secretion was low and degradation high. This information can also be used to allude to the impact of protein degradation following secretion –manifesting in a negative value for the rate of protein yielded. Protein degradation can be masked by protein secretion if secretion exceeds degradation, as the rate of protein secreted will appear as a positive number. While it is not possible to conclude how much these impact on the observed yield it would seem appropriate to suggest that degradation occurs more in later stages of protein secretion, as at 24 hours there a negligible or negative (Δ*CKL* and Δ*CKL* Δ*lrhA* respectively) rate of protein secretion. Further to this at 6 and 7 hours the rate of protein secretion was lower than that seen at the previous hour –however it is not possible to conclude whether this is due to a decreased rate of protein secretion or an increased rate of protein degradation.

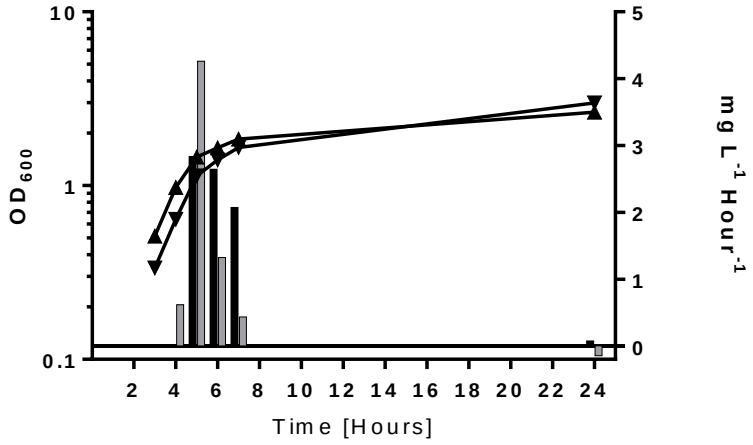


Figure 3.19: Secretion of E2 from the secretor strain and Δ*lrhA* mutant during the growth curve per hour

E.coli Δ*CKL* and Δ*CKL* Δ*lrhA* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (Δ*CKL*: ▲, Δ*CKL* Δ*lrhA*: ▼, left hand axis). In parallel cells were harvested and the supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) before SDS-PAGE and Western analysis using anti-flagellin (H48) antibody and an HRP secondary. The amount of protein secreted per hour is shown in mg L^{-1} $Hour^{-1}$ cell culture as calculated from a standardised amount of E2 protein using ImageJ based densitometry (Δ*CKL*: black bar, Δ*CKL* Δ*lrhA*: grey bar, right hand axis)

These results were inconsistent with the literature – protein secretion was predicted to be higher in the Δ*lrhA* mutant, as it was reported that *lrhA* knockout mutants exhibited 3.5x increased flagella gene expression and were hypermotile (Lehnen et al., 2002). Flagella gene expression assays were carried out to assess whether this was a reflection of gene expression or whether it was a result of poor secretion capacity.

3.4.3.2. Effect of deletion of *lrhA* on flagella gene expression

Increase in flagella gene expression is indicative of an increased number of flagella in the phenotype. To measure flagella gene expression, flagella *lacZ* reporter gene fusions which had previously been used by Stafford et al. (2005) were transformed into strains. *lacZ* encodes the β-galactosidase alpha chain which restores enzymatic activity to a defective β-galactosidase enzyme. This was implemented by plasmids harbouring the σ70 *fliA* promoter fragment fused to a truncated *lacZ* gene. Initiation of transcription at the promoter results in the translation of the β-galactosidase alpha chain, following complementation and formation of the functional enzyme, this cleaves substrates including lactose and o-nitrophenyl-β-d-galactopyranoside (ONPG) (Figure 3.). *fliA* expression is under the control of two promoters: σ70 and σ28. σ70 is directed by the presence of the $FlhD_4C_2$ complex, to initiate *fliA* transcription (Aldridge et al. 2006.). FliA is also known as σ28, therefore it is also able to self-promote gene transcription. However σ28 (FliA) is bound by the anti- σ28 FlgM, therefore only promotes the expression of class III genes (*fliA* included) when FlgM has been exported from the cell, in response to hook completion (Figure 1.5). As flagella gene expression is all under a common regulation system any flagella gene is a good indicator of overall flagella gene activity. Increase in gene expression of *fliA* by σ70 is indicative of more $FlhD_4C_2$ being present and therefore upregulated gene expression throughout the flagella gene hierarchy. This should result in the more modified flagella being translated, which will allow more product to be secreted.

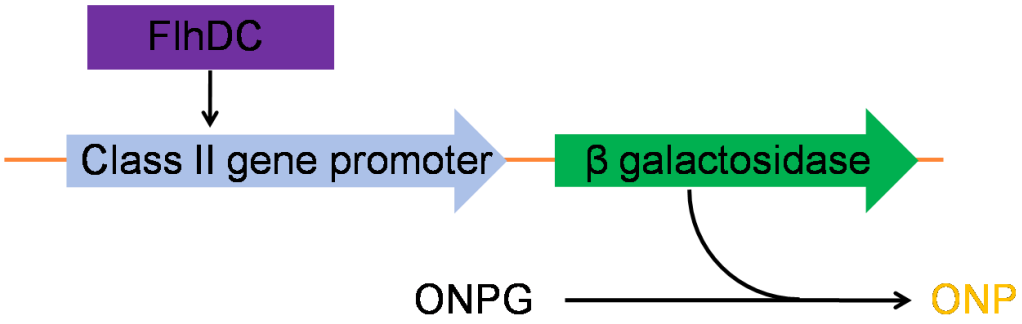


Figure 3.20: **Schematic of the gene expression assay.**

Class II promoters are fused to *lacZ* in the plasmid. Therefore when the master regulator $FlhD_4C_2$ activates the flagella class II gene promoter, β galactosidase is expressed which catalyses the hydrolysis of ONPG into ONP which produces a yellow colour.

A key measurement in the assay is the time taken for a reagent to turn yellow as a result of ONPG cleavage and ONP production. This is catalysed by the β galactosidase produced in response to promoter activation. As the assay is measuring gene expression over time and in multiple strains a range of β galactosidase concentrations will be present, therefore a range times and OD results will be observed. It is essential that the assay is accurate through the range of concentrations of β galactosidase that may be observed. The assay was optimised so that the time taken for the yellow colour to arise was within a reasonable time and colour change was measurable with a spectrophotometer. This was particularly relevant for high OD assays as the colour change occurs at a high rate. Preliminary results indicated that performing the assay with the same volume of culture to Z buffer was a suitable ratio.

Flagella gene expression assays were carried out on Δ*CKL* and Δ*CKL* Δ*lrhA* strains to quantify flagella gene expression through the growth curve (Figure 3.). Samples of cell culture were removed every hour to access gene promoter activity throughout the growth curve.

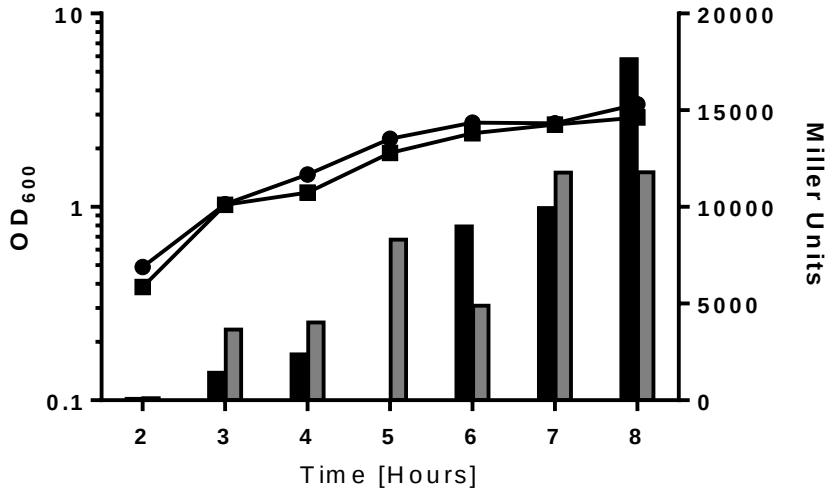


Figure 3.21: Flagella gene expression throughout the growth curve.

E. coli Δ*CKL* and Δ*CKL* Δ*lrhA* containing the plasmid pGPS*fliA* was grown in LB and the OD_{600} assayed (Δ*CKL*: ●, Δ*CKL* Δ*lrhA*: ■, left hand axis). In parallel cells were harvested and lysed. A β-galactosidase assay was carried out with ONPG as the substrate. Miller Units were derived from assays (Δ*CKL*: black bar Δ*CKL* Δ*lrhA*: grey bar -right hand axis). One biological replicate of each. NOTE: no reading obtained for Δ*CKL* at 5 hours.

Growth curves of the two strains were very similar for both strains. The overall trend for both strains was an increase in *fliA* gene expression as OD_{600} increased. However *fliA* gene expression fluctuated throughout the growth curve in Δ*CKL* Δ*lrhA*. Δ*CKL* promoter activity exhibited a constant increase from a negligible amount in the early exponential phase to a high amount in the stationary phase. Δ*CKL* Δ*lrha fliA* promoter activity was consistently higher and showed an overall increase during the exponential phase. At 6 hours activity decreased and was lower than Δ*CKL* activity. This was followed by an increase at 7 hours. Δ*CKL* Δ*lrha fliA* promoter activity levelled off from 7 to 8 hours and was lower than Δ*CKL* fliA activity.

lrhA knock out mutants are reported to exhibit higher activity of flagella class I and II gene promoters and master regulator promoter activity was 3.5 x higher in the *lrhA* knockout mutant (Lehnen et al., 2002). Therefore it was expected that both gene expression and secretion assays would show higher performance of the Δ*lrhA* mutant strain. This effect was not seen in secretion assay results, neither was it observed in gene expression results. Despite

this Δ*CKL* Δ*lrhA* gene expression results do not reflect the secretion assay results. Based on the secretion assay, gene regulation would be expected to increase up to 7 hours then tail off. However gene regulation fluctuated through the growth curve. Consistency between secretion and gene expression results was observed in the earlier growth phases (up to 5 hours), as Δ*CKL* Δ*lrhA* outperformed the Δ*CKL* strain in both assays. However this was short lived and potentially inaccurate, given the poor reliability of results obtained for the remainder of the growth curve.

The regulation phenotype of *ΔlrhA* was different to that reported in the literature. These inconsistencies may be due to the genotype of the strain, which may incur genetic variation over time. This can result in the accumulation of suppressor mutations which reverse the phenotype. Another possible cause of the unexpected results in that an insertion sequence is present in strains. Insertion sequences (IS) are mobile genetic elements which cause genomic plasticity (Naas et al., 1994). As some strains of MC1000 have been shown to harbour an IS5 element in one of two 'hot spots' ~300 or ~250 bps upstream of the *flhD* promoter (Barker et al. 2004; Wang & Wood, 2011), Experimental work will turn to investigating whether this is true of the strain utilised in this study.

3.4.4. Presence of insertion sequence 5 in the *flhDC* promoter region

While IS5 has been shown to 'hop' upstream of *flhD* in motile strains, it was found that it would not insert in structurally non motile mutants (i.e *flgK,* motA) therefore if present , the IS5 element will have be an inherited from the parent MC1000 strain (Wang & Wood, 2011). IS5 elements caused increased motility in these strains, by reducing transcriptional repression. IS5 insertion does not alter *flhD* or the promoter region *per se*, instead it is thought to alter the AT-rich region upstream of *flhDC.* Through this modification it is assumed that IS5 uncouples binding of the LrhA from transcriptional regulation of *flhDC* (Barker et al., 2004). If the IS5 element is upstream of *flhD* in the MC1000 strain used in this thesis, this may account for the unexpected results obtained for both secretion and flagella gene expression in the *lrhA* mutant.

3.4.4.1. Testing for the presence of insertion sequence 5 in the *flhDC* promoter region

The presence of the IS5 element upstream of *flhD* was confirmed when the entire MC1000 Δ*CKL* strain was sequenced by the MicrobesNG service, University of Birmingham. The genomic region is constructed below (Figure 3.).

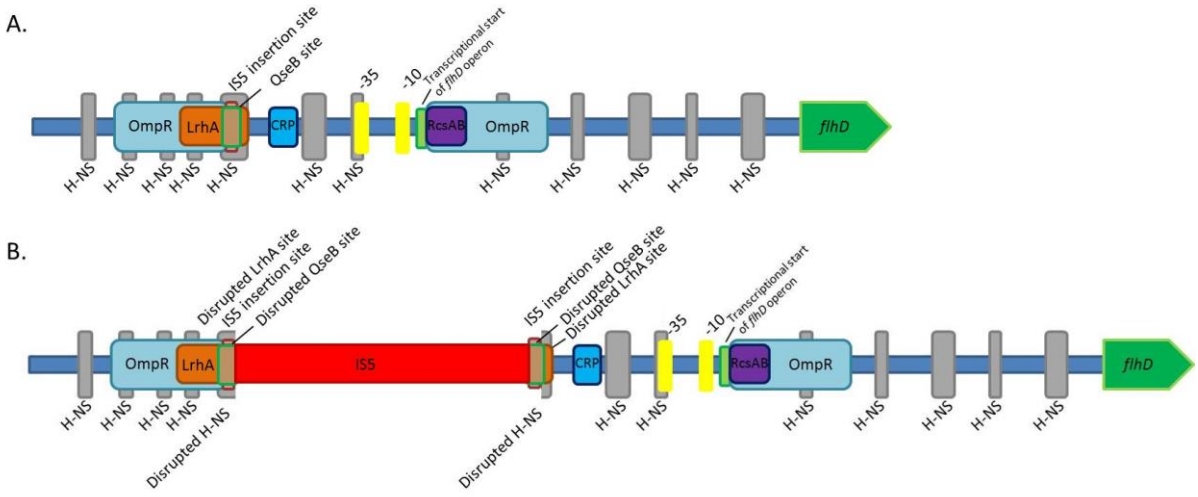


Figure 3.22: Schematic of the genetic region upstream of the *flhD* operon with and without the IS5 element.

Construction of the upstream region of *flhD* in (A) MC1000 without any IS elements, based on literature and genomic databases and (B) MC1000 Δ*CKL* based on sequencing of the whole genome. Binding sites for regulatory proteins shown for: H-NS (histone-like nucleoid-structuring protein), OmpR, LrhA, CRP (cAMP-receptor protein), RscAB and QseB. The *flhD* gene, transcriptional start site of the *flhD* operon and the -10 and -35 promoter sequences. Also shown are the IS5 insertion site and in (2) the IS5 insertion and duplication of the IS5 insertion site to flank the IS5 element (Yanagihara et al. 1999; Barker et al. 2004).

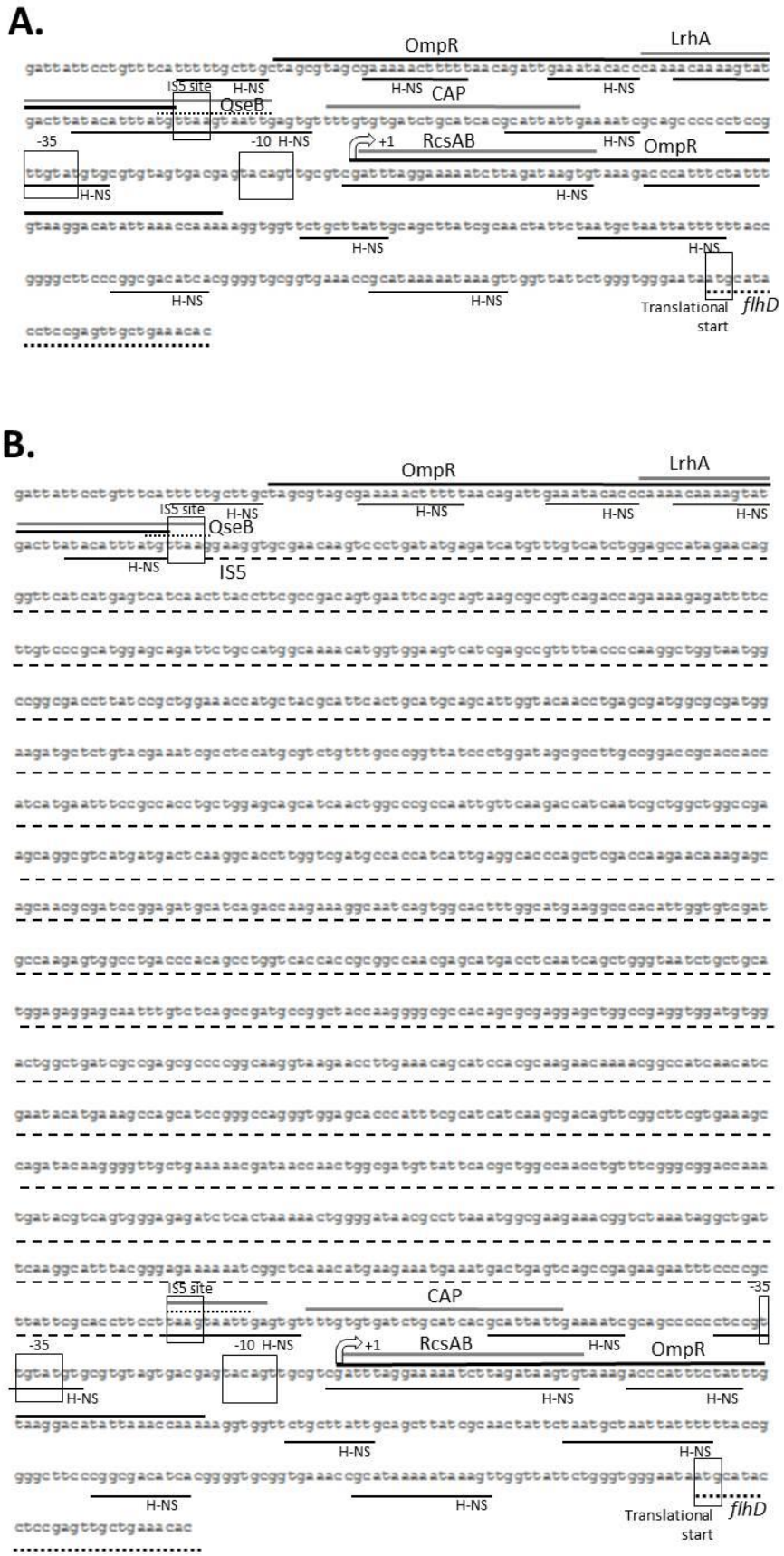

A.
OmpR
LrhA
IS5 site
H-NS
QseB
CAP
-35
-10
+1
RcsAB
Translational start
flhD
B.
IS5
-35
-10
Translational start
flhD

Figure 3.23: Annotated nucleotide sequence of the genetic region upstream of the *flhD* operon with and without the IS5 element.

Construction of the upstream region of *flhD* in (A) MC1000 without any IS elements, based on literature and genomic databases and (B) MC1000 Δ*CKL* based on sequencing of the whole genome. Binding sites for regulatory proteins shown for: H-NS (histone-like nucleoid-structuring protein), OmpR, LrhA, CRP (cAMP-receptor protein), RscAB and QseB. The *flhD* gene, transcriptional start site (+1) of the *flhD* operon and the -10 and -35 promoter sequences are also labelled. Also shown are the IS5 insertion site and in (2) the IS5 insertion and duplication of the IS5 insertion site to flank the IS5 element (Yanagihara et al. 1999; Barker et al. 2004; Clarke & Sperandio, 2005)

As shown in Figure 3. and Figure 3., the IS5 element is present upstream of *flhD* in the MC1000 Δ*CKL* strain. This leads to the disruption of the LrhA binding site, one of the H-NS sites and one of the QseB sites. It also results in the decoupling of further H-NS sites, one of the OmpR binding sites and one of the QseB sites (not shown) from the *flhD* operon. With regards to LrhA and OmpR the IS5 insertion has been shown to result in increased transcriptional activation of *flhD,* presumably due to the inability of transcriptional repressors to bind to sites (Barker et al. 2004). When osmolality is high, phosphorylated OmpR binds to two sites in the *flhD* promoter region, causing a reduction in *flhDC* gene expression (Shin & Park, 1995). It is unclear if OmpR binding is required at both sites to instigate this effect in *E. coli*. It was reported that OmpR was required at both binding sites to initiate an effect in *Yersinia,* so this may well be the case in *E. coli* -although it is important to note that in *Yersinia* OmpR binding causes upregulation of *flhDC* expression therefore conservation between the two species may be low (Hu et al., 2009). The H-NS is a positive regulator of the *flhD* operon, and requires binding at all H-NS sites (Soutourina et al., 1999). Throughout the *E. coli* genome H-NS binds at multiple sites in promoter regions, lateral interaction of H-NS proteins causes altered transcriptional control (Oshima et al., 2006). Usually H-NS binding exerts negative regulatory control, as it either traps or inhibits binding of RNA polymerase; it is suggested that in flagella gene expression H-NS binding allows additional transcriptional activators to bind to the promoter region. Interestingly HN-S also positively regulates motility, by binding to the FliG motor protein promoter region (Donato & Kawula, 1998). The quorum sensing regulator QseB is also a positive regulator which binds at two sites in the *flhD* promoter region -the proximal weak affinity site was disrupted and the distal high affinity site (not shown as it is 650+ bp upstream

of the *flhD* transcriptional start site) was decoupled (Clarke & Sperandio, 2005). However as the overall effect of IS5 insertion is for a 2.7 fold increase in *flhD* transcription the negative effect of decoupling H-NS and QseB transcriptional control seems to be negated. Some transcriptional binding sites are upstream of the IS5 site therefore unaffected by the insertion in the MC1000 Δ*CKL* strain. The binding site for the negative *flhDC* transcription factor RscAB complex remains *in situ* and can therefore potentially negatively regulate *flhD.* However the CRP binding site is still intact, which is beneficial in terms of flagella gene expression upregulation, as in the absence of glucose CRP activity increases *flhDC* expression (Zhao et al., 2007). In has been shown *in vitro* that the presence of CRP can relieve the repression observed in an H-NS knockout mutant. This was not shown *in vivo* but may suggest that a combination of transcription factors act to negate the negative effects of HN-S decoupling from *flhD* by positively increasing gene *flhD* expression beyond their usual capacity (Soutourina et al., 1999).

Therefore it is possible to conclude that MC1000 is a hypermotile strain and flagella gene expression is already upregulated. Therefore the removal of master regulator repressors such as LrhA is ineffective in increasing flagella gene expression. While this renders the strategy to increase flagella gene expression through the removal of *lrhA* void, this is not an issue as flagella gene expression is already high. Investigation will continue, to observe whether this can be improved upon further with the removal of genes which encode transcriptional repressors which are not affected by the IS5 insertion or act directly on *flhDC/* $FlhD_4C_2$ (DksA and ClpX respectively).

3.4.5. Mutant characterisation: *dksA*

Previous studies have shown that flagella gene expression is higher in *dksA* knock out mutants (Lemke et al., 2009). *lacZ* reporter gene fusions show that *flhDC* promoter activity was 2.3 x higher in log phase and 9 x higher in stationary phase. Here investigation focused on whether the Δ*CKL* Δ*lrhA* strain exhibited higher flagella gene expression and whether this translates to more secretion apparatus and therefore more secretion. This mutant holds promise as it was found that protein secretion in the secretor strain decreases in the stationary phase (Figure 3.), therefore this mutant has potential to oppose this characteristic and improve protein secretion throughout the stationary phase. DksA acts by binding near the active site of RNA polymerase

and therefore decreasing the 'lifespan' of open RNAP (Paul et al. 2004). This in turn results in reduced RNAP-*flhDC* promoter DNA complexes and therefore less *flhDC* gene expression. As DksA is indirectly linked to *flhDC* promoter DNA (through decreased open RNAP), its route of action is not altered by the presence of IS5, which only concerns transcriptional control upstream of the *flhDC* promoter.

3.4.5.1. Effects of deletion of *dksA* on FT3SS secretion

Once the *dksA* knockout was confirmed by PCR (Figure 3.) a secretion assay was carried out to establish the secretion profile of the strain. As it has been reported that flagella gene expression increases in the stationary phase of Δ*dksA* mutants; the secretion was measured through the growth curve, with samples being taken every hour. Δ*CKL* expressing E2 protein were also grown in tandem so that the secretion capacities of the two could be compared. A protein standard was also loaded onto the SDS-PAGE gel to allow yields to be quantified. Following Western blotting E2 protein could be observed after 5 hours in both strains (Figure 3.). More protein was seen in the supernatant of Δ*CKL* cells than Δ*dksA* at all time points.

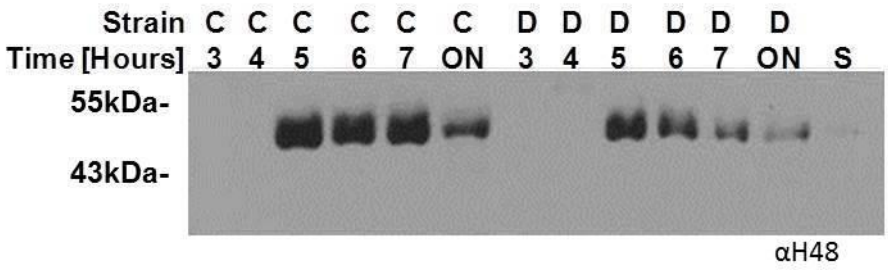


Figure 3.24: Secretion of E2 from the secretor strain and Δ*dksA* mutant during the growth curve.

E.coli Δ*CKL* (C) or Δ*CKL* Δ*dksA* (D) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at hourly intervals and the OD_{600} measured. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western analysis of supernatant using anti-flagellin (H48) antibody and an HRP secondary. An E2 protein standard (S) was loaded to allow quantification of secreted protein concentration.

Densitometry analysis was carried out on the Western blot. While the standard appears only faintly on the Western blot, it was detectable in ImageJ software analysis; therefore secreted protein yield could be calculated and plotted against OD_{600} (Figure 3.). The two mutant strains grew similarly, suggesting that removing *dksA* does not confer a change in growth phenotype. The concentration of E2 found in culture supernatant was consistently higher in Δ*CKL* in comparison to Δ*CKL* Δ*dksA*. The highest concentration of protein observed was 93.117 mg L^{-1} in Δ*CKL* following 6 hours of growth. At every time point following 6 hours the concentration of E2 protein decreased -with the exception of Δ*CKL* after 8 hours, which exhibited a slight increase. After 24 hours E2 protein was still detected in the supernatant of both strains at concentrations of 48.158 mg L^{-1} and 15.646 mg L^{-1} for Δ*CKL* and Δ*CKL* Δ*dksA* respectively.

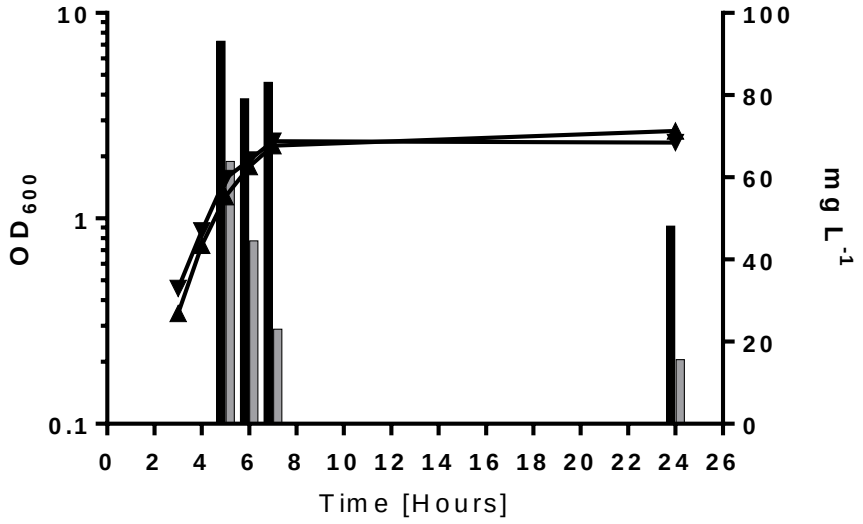


Figure 3.25: Secretion of E2 from the secretor strain and Δ*dksA* mutant during the growth curve

E.coli Δ*CKL* and Δ*CKL* Δ*dksA* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (Δ*CKL*: ▲, Δ*CKL* Δ*dksA*: ▼, left hand axis). In parallel cells were harvested and the supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) before SDS-PAGE and Western analysis using anti-flagellin (H48) antibody and an HRP secondary. The amount of protein secreted is shown in mg L^{-1} cell culture as calculated from a standardised amount of E2 protein using ImageJ based densitometry (Δ*CKL*: black bar, Δ*CKL* Δ*dksA*: grey bar, right hand axis)

When visualised as secretion per hour (Figure 3.) it is evident that the majority of E2 protein is secreted into the media between hour 4 and 5, where the rate of secretion is 93.117 mg L^{-1} $Hour^{-1}$ and 63.876 mg L^{-1} $Hour^{-1}$ for Δ*CKL* and Δ*CKL* Δ*dksA* respectively. Following this (with the exception of Δ*CKL* at hour 7, where a slight positive value for the rate of secretion is observed) the values obtained for the rate of protein secretion is always negative, indicating protein degradation. The concentration of E2 protein observed here is markedly higher than that seen in the previous experiment (Figure 3.). Potential reasons for this will be discussed in a later section.

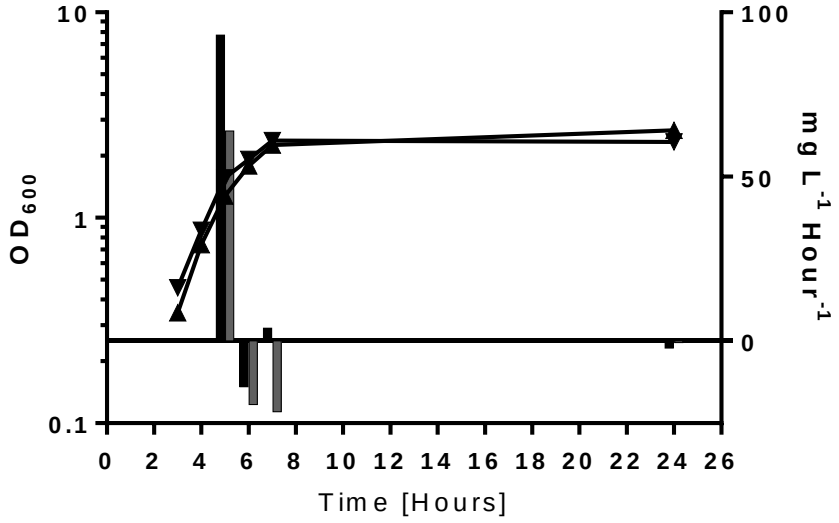


Figure 3.26: Secretion of E2 from the secretor strain and Δ*dksA* mutant during the growth curve per hour

E.coli Δ*CKL* and Δ*CKL* Δ*dksA* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (Δ*CKL*: ▲, Δ*CKL* Δ*dksA*: ▼, left hand axis). In parallel cells were harvested and the supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) before SDS-PAGE and Western analysis using anti-flagellin (H48) antibody and an HRP secondary. The amount of protein secreted per hour is shown in mg L^{-1} $Hour^{-1}$ cell culture as calculated from a standardised amount of E2 protein using ImageJ based densitometry (Δ*CKL*: black bar, Δ*CKL* Δ*dksA*: grey bar, right hand axis)

Following the initial positive rate of protein secretion at 5 hours, protein was secreted at an increasingly reduced rate; this was opposite to what was expected if Δ*CKL* Δ*dksA* had performed as anticipated. If the *dksA* knockout resulted in increased flagella gene expression

in the stationary phase as reported in the literature, increased secretion apparatus formation and therefore E2 secretion would have been observed. To investigate if flagella gene expression was higher in the *dksA* mutant a gene expression assay was carried out.

3.4.5.2. Effects of deletion of *dksA* on flagella gene expression

Flagella gene expression was measured throughout the growth curve of both Δ*CKL* and Δ*CKL* Δ*dksA.*

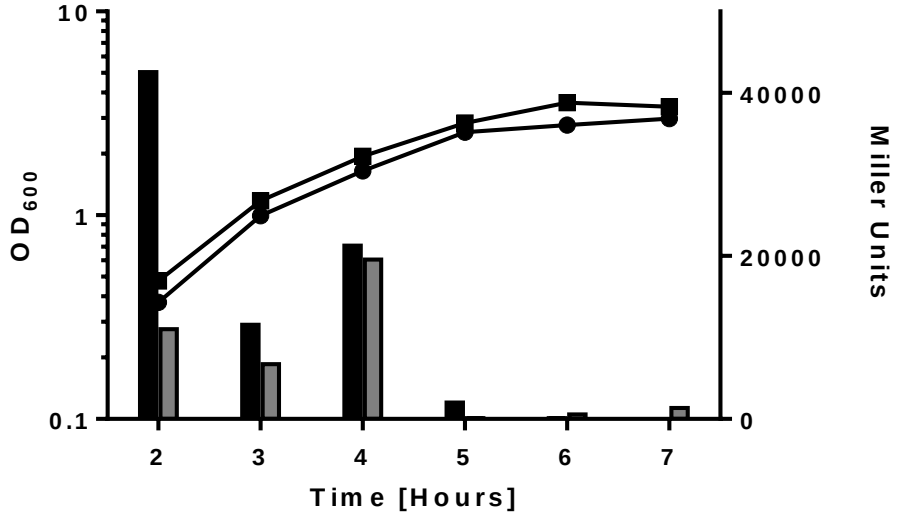


Figure 3.27: Flagella gene expression throughout the growth curve.

E. coli Δ*CKL* and Δ*CKL* Δ*dksA* containing the plasmid pGPS*fliA* was grown in LB and the OD_{600} assayed (Δ*CKL*: ●, Δ*CKL* Δ*dksA*: ■, left hand axis). In parallel cells were harvested and lysed. A β-galactosidase assay was carried out with ONPG as the substrate. Miller Units were derived from assays (Δ*CKL*: black bar, Δ*CKL* Δ*dksA*: grey bar, right hand axis)

Growth curves of the two strains were very similar for both strains. The overall trend for both strains was a decrease in *fliA* gene expression as OD_{600} increased. However between hour 3 and 4 (late log stage) *fliA* gene expression increased in both *ΔCKL dksA* and *ΔCKL*. After 2 hours the promoter activity of *ΔCKL* was higher than *ΔCKL dksA*, following 3 hours this was still true,

however the difference had reduced. From 4 hours the promoter activity of both strains was very similar –both exhibiting a steady decrease until the stationary phase where gene expression was negligible in both strains.

It was expected that both gene expression and secretion assays would be higher in Δ*dksA* mutant strain, particularly in the stationary phase. This effect was not seen in either secretion assay results or gene expression results, however *ΔCKL* Δ*dksA* gene expression results do reflect the secretion assay results, both showing a decrease through the growth curve in *fliA* promoter activity and E2 protein secretion respectively. In the early stationary phase both strains displayed an increase in gene expression from 3 to 4 hours (Figure 3.); this coincides with the high initial rate of E2 secretion seen in both *ΔCKL* and *ΔCKL* Δ*dksA* during the early stationary phase in Figure 3..

3.4.6. Mutant characterisation: *clpX*

As ClpX has been shown to actively degrade the $FlhD_4C_2$ complex (Tomoyasu et al., 2002), the ClpX knockout mutant is expected to exhibit increased flagella gene expression and therefore secretion. This has been shown experimentally as *ΔclpX* mutants were shown to accumulate $FlhD_4C_2$, the complex had increased half-life and cells were more motile (Kitagawa et al., 2011). Therefore removing *ΔclpX* from the Δ*CKL* strain may result in increased flagella gene expression and more secretion. Furthermore as this concerns the reduction in proteolysis of the $FlhD_4C_2$ protein complex, the presence of IS5 is not relevant to this investigation.

Prior to the investigation of the *clpX* knockout mutation in the immotile secretor strain, the effect of the same mutation was investigated in the wild-type MC1000 strain. As the MC1000 Δ*clpX* strain is able to form functional flagella, it enables comparable measurements of motility and quantification of the abundance of filament protein in wild type MC1000 and the MC1000 Δ*clpX* mutant.

3.4.6.1. Effects of deletion of *clpX* on the abundance of filament protein

Increased flagella gene expression will result in an increase in the amount of filament (FliC) protein which is presented on the surface of cells. Relative concentration of filament protein associated with the cell surface of MC1000 with and without the Δ*clpX* mutation was

measured to ascertain how the Δ*clpX* mutation alters flagella assembly and by proxy, type III secretion. To achieve this strains were grown with low agitation at 30°C to ensure that flagella filaments remained intact. Cells were then carefully collected from the supernatant to ensure filaments remained intact and normalised for OD_{600} to allow direct comparison of the concentration of filament proteins obtained from each inoculate. Following vortexing to shear the filaments from the cells the filament proteins were isolated via several centrifugation steps, ensuring no intracellular filaments contaminated the sample. Isolated filaments were then analysed by SDS-PAGE (Figure 3.).

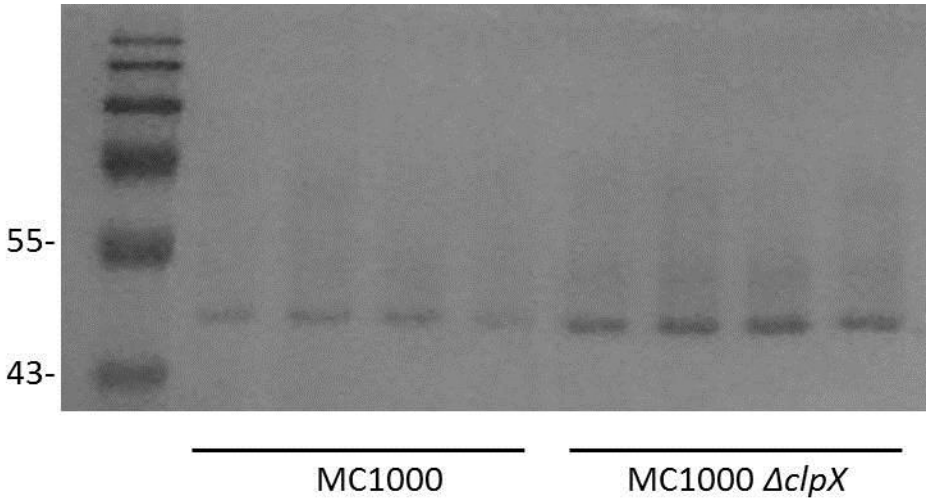


Figure 3.28: Concentration of sheared flagella filaments from WT MC1000 and MC1000 Δ*clpX*.

E.coli MC1000 and MC1000 Δ*clpX* were grown in LB and the OD_{600} assayed. Cells were harvested and normalised according OD_{600} of cultures from which they were derived, filaments were isolated and loaded onto an SDS-PAGE and stained with Instant Blue. Results from four biological replicates.

Densitometry analysis was then carried out on the image of the SDS-PAGE gel, to allow quantification of the concentration of filament proteins obtained from the two strains. As a protein standard was not loaded onto the SDS-PAGE it is not possible to calculate a numerical value for the yield of filament proteins; however a relative value is also informative as to the effect of the Δ*clpX* mutation on the concentration of FliC protein secreted through the flagella

type III secretion system. The *clpX* knockout mutation resulted in an average of 1.95 times more filament protein being secreted through the flagella type III secretion system (Figure 3.).

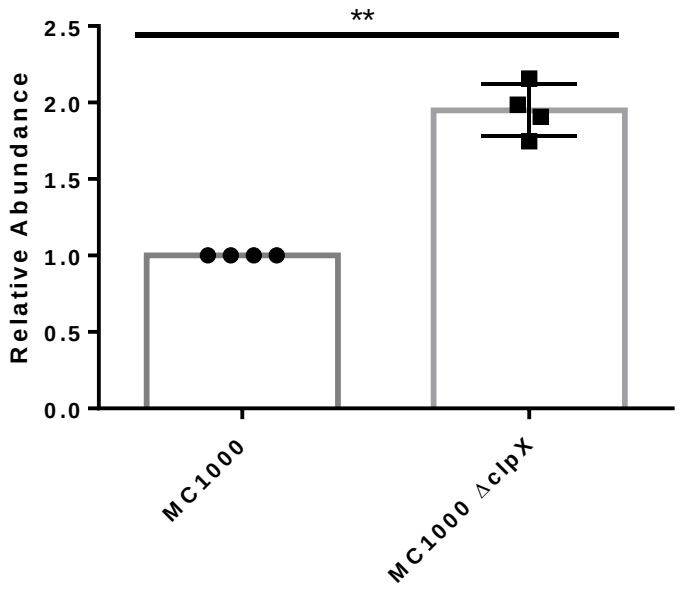


Figure 3.29: Relative concentration of sheared flagella filaments from WT MC1000 and MC1000 Δ*clpX*.

E.coli MC1000 and MC1000 Δ*clpX* were grown in LB and the OD_{600} assayed. Cells were harvested and normalised according OD_{600} of cultures from which they were derived, filaments were isolated and loaded onto an SDS-PAGE. Using ImageJ based densitometry the abundance of filament protein was quantified; results were paired to allow the calculation of relative abundance of filament protein to MC1000. Results from four biological replicates, grown on the same day. A paired t-test was carried out for the pre-normalised values, to give more power to the calculation. Individual data points, mean and standard error of the mean and result of two way t-test shown (** = $p<0.01$)

Significantly more (2 fold) FliC monomers were isolated from the Δ*clpX* mutant, suggesting that more filament proteins assemble in the mutant as opposed to the wild type. As they are secreted through the fT3SS, this would suggest higher secretion capacity in Δ*clpX* mutants.

3.4.6.2. Effects of deletion of *clpX* on motility

Motility agar plates were inoculated with either WT MC1000 or MC1000 Δ*clpX*, to allow measurements of motility. As flagella are intact in these strains, increased flagella gene expression will manifest in increased motility. The swimming motility of Δ*CKL* mutants was also assessed. As these mutant strains do not form functional filaments, it is not expected that they will exhibit any motility behaviour. Inoculates from all four strains were stabbed into each plate to reduce any variability that may arise from the environment. Following inoculation by stabbing individual colonies into the motility agar, plates were incubated at 30°C in humid conditions, allowing swimming motility to occur (Figure 3.). The radius of the swimming zone from the inoculation point was measured at time intervals and plates were returned to the incubator. Over time the swimming zone of motile strains extended and was in some cases obscured by the perimeter of the agar plate, therefore a measurement of radius was chosen over circumference. Results were used to calculate the rate of swimming.

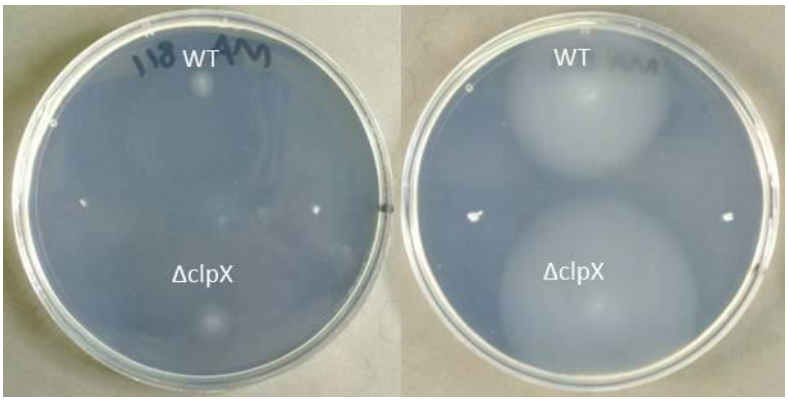


Figure 3.30: The effect of *clpX* on the swimming motility of MC1000 and MC1000 Δ*flgKL* Δ*fliC* strains.

0.25% LB agar plates were inoculated with (clockwise from the top), MC1000, MC1000 Δ*flgKL* Δ*fliC,* MC1000 Δ*clpX* and MC1000 Δ*flgKL* Δ*fliC* Δ*clpX.* Swimming motility was measured at time intervals during incubation at 30°C. Images were recorded following 1 hour (left) and 24 hours (right).

Results were used to calculate the rate of swimming. Secretion apparatus (Δ*flgKL* Δ*fliC)* mutants always exhibited a lower rate of swimming than strains with intact flagella (Figure 3.). After 1 hour a 0.625 and 0.875 mm $Hour^{-1}$ rate of swimming was observed for the Δ*CKL* mutants, however from 2 hours onwards this rate reduced to around 0 for both strains. This

indicated that the Δ*CKL* mutation resulted in a strain which was entirely immotile, presumably due to the absence of flagella. Both the wildtype flagella strains displayed swimming motility throughout the experiment –this demonstrates that growth conditions were suitable for motility to occur, confirming the immotile phenotype of the secretion apparatus mutants. After 1 and 3 hours the rate of swimming in MC1000 strains was slightly higher that MC1000 Δ*clpX,* however at 2 and 20 hours the rate of motility was significantly higher in MC1000 cells with the Δ*clpX* mutation (p = <0.01. Paired t-test). After 20 hours the average rate of motility for MC1000 Δ*clpX* cells was 1.13 mm $Hour^{-1}$ -1.33 times that of MC1000. These results demonstrate that the removal of *clpX* from the MC1000 chromosome results in increased motility in cells which produce functional flagella.

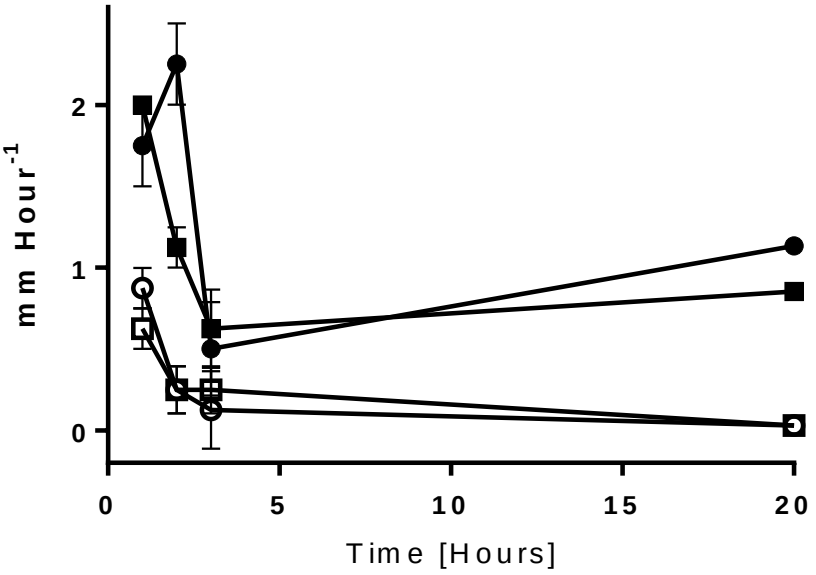


Figure 3.31: The effect of *clpX* on swimming rate of MC1000 and MC1000 Δ*flgKL* Δ*fliC* strains.

0.25% LB agar plates were inoculated with individual colonies. Swimming motility was measured at time intervals during incubation at 30°C and recorded, allowing the rate of swimming to be calculated. MC1000: ■, MC1000 Δ*fliC* Δ*flgKL:* □*,* MC1000 Δ*clpX*: ● and MC1000 Δ*fliC* Δ*flgKL* Δ*clpX:* ○. Four biological replicates for each strain.

The secretion capacity of the Δ*CKL* Δ*clpX* mutant was then investigated to see if the phenotypic effects observed in Figure 3. extend to the modified FT3SS secretion platform.

3.4.6.3. Effects of deletion of *clpX* on FT3SS secretion

As discussed, the literature suggests that the removal of *clpX* from the chromosome will result in an increase in $FlhD_4C_2$ complex, therefore increased flagella gene expression. The secretion capacity of Δ*CKL* Δ*clpX* in comparison to the Δ*CKL* strain was investigated throughout the growth curve. A plasmid harbouring the E2 secretion construct was electroporated into *ΔCKL* and *ΔCKL ΔclpX* cells. Following inoculation and induction with 0.05mM IPTG measurements of OD_{600} were taken every hour and the supernatant and cells were prepared for SDS-PAGE and Western blot (Figure 3.), to ascertain the concentration of E2 protein. Previously only the supernatant of cells has been analysed by Western blot in secretion experiments through the growth curve, however more information can be obtained by observing the concentration of E2 located intracellularly. Bands were present in all samples at each time point. A higher concentration of E2 protein was seen both intracellularly and in the secreted fraction of Δ*CKL* Δ*clpX* mutants in comparison to Δ*CKL* at all times.

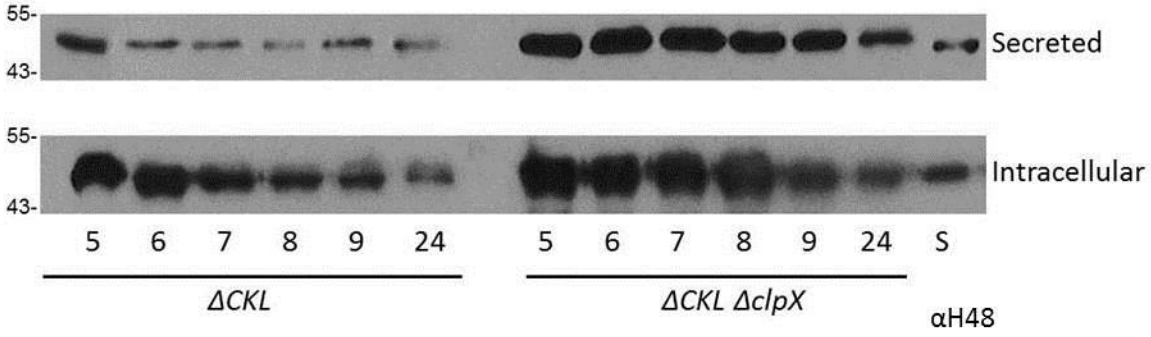


Figure 3.32: Intracellular and secreted E2 from the secretor strain and Δ*clpX* mutant during the growth curve.

E.coli Δ*CKL* (C) or Δ*CKL* Δ*clpX* (C) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at hourly intervals and the OD_{600} measured. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using anti-flagellin (H48) antibody and an HRP secondary. An E2 protein standard (S) was loaded to allow quantification of intracellular and secreted protein concentration.

Densitometry of each band on the Western blot (Figure 3.) was measured using ImageJ and values were calibrated to the E2 protein standard to calculate the concentration of E2 protein

in samples from both strains at hour time points. In addition to the results obtained by the Western blot shown in Figure 3., a further biological repeat was analysed and is included in the following results. Results reflect the amount of E2 secreted per OD unit, however if this value is adjusted to account for the volume removed to amount to 1 OD unit, a mg L^{-1} culture can be calculated (Figure 3.) for both intracellular and secreted E2 protein. Cell densities (OD_{600}) of the two cultures were very similar throughout the growth curve, demonstrating that there was no detrimental effect of the mutation on strains. The concentration of E2 in the supernatant of Δ*CKL* Δ*clpX* was always greater than that in Δ*CKL* (Figure 3.: left). The highest concentration of secreted E2 in the Δ*CKL* strain was observed in the mid log phase at 5 hours (1.237 mg L^{-1}); after which slight fluctuations in E2 concentration were observed, however the concentration was always lower in the hours following this initial measurement. From 5 to 6 hours the concentration of E2 in the secreted fraction of Δ*CKL* Δ*clpX* cells increased from 3.029 to 4.220 mg L^{-1} and then began to decrease to 2.324 mg L^{-1} secreted E2 following 24 hours growth. As observed in the secreted fraction (Figure 3.: right), the highest intracellular concentration of E2 (1.748 g L^{-1}) was observed at 5 hours in Δ*CKL*. After 24 hours there was 0.440 g L^{-1} E2 protein present in cells. There was always more E2 present in Δ*CKL* Δ*clpX* cells than Δ*CKL*. The highest concentration of intracellular protein in Δ*CKL* Δ*clpX* was observed after 5 hours (2.633 g L^{-1}) and decreased every hour following this, with the exception of 7 hours where there was a slight increase. After 24 hours 1.439 g L^{-1} E2 was visible in cells.

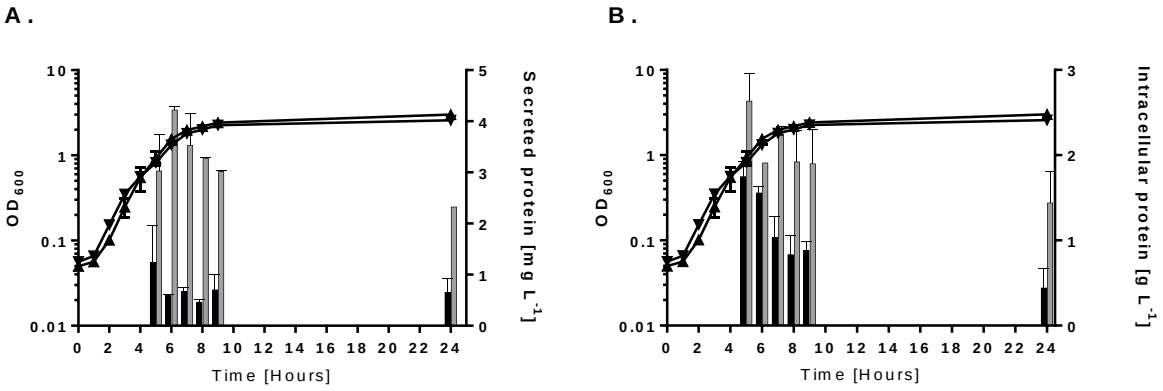


Figure 3.33: Intracellular and secreted E2 from the secretor strain and Δ*clpX* mutant during the growth curve

E.coli Δ*CKL* and Δ*CKL* Δ*clpX* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (Δ*CKL*: ▲, Δ*CKL* Δ*dksA*: ▼, left hand axis). In parallel 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) and prepared with 2x SDS-PAGE loading buffer. Samples underwent SDS-PAGE and Western blot analysis using anti-flagellin (H48) antibody and an HRP secondary. The amount of protein secreted (A) or located intracellularly (B) is shown in mg L^{-1} cell culture as calculated from a standardised amount of E2 protein using ImageJ based densitometry (Δ*CKL*: black bar, Δ*CKL* Δ*dksA*: grey bar, right hand axis). Results from two biological replicates.

By subtracting the concentration of E2 present in the supernatant at the previous hour from each value it is possible to observe the rate of protein secretion at each time point in mg L^{-1} $Hour^{-1}$, rather than a cumulative view. A positive value shows that protein secretion exceeds degradation, while a negative value suggests the opposite. The highest rate of E2 secreted by Δ*CKL* was 1.971 mg L^{-1} $Hour^{-1}$ and 3.745 mg L^{-1} $Hour^{-1}$ for Δ*CKL* Δ*clpX*. For both strains this high rate of E2 secretion was observed from 4 to 5 hours (the mid log phase). As the concentration of E2 was not measured at 4 hours, it is not possible to ascertain whether the rate observed at 5 hours is accurate or whether protein was also secreted prior to 4 hours. After 5 hours the rate of E2 secreted by Δ*CKL* was either negative or negligible, suggesting that secretion was low and degradation high. After 6 hours cells reached early stationary phase the rate of secretion observed in the *clpX* knockout mutant was reduced, but positive –suggesting that either (or both) secretion was slightly lower or degradation slightly higher than at 5 hours. After 6 hours Δ*CKL* Δ*clpX* demonstrated a negative rate of protein secretion, suggesting protein degradation was higher than secretion of E2. The rate of secretion observed between 9 and 24 hours for both strain was around 0 –suggesting that both E2 protein secretion and degradation were matched.

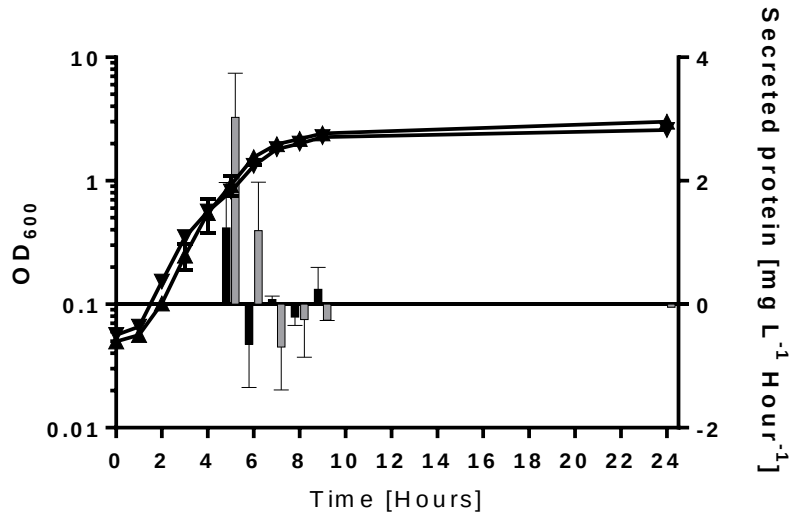


Figure 3.34: Secretion of E2 from the secretor strain and Δ*clpX* mutant during the growth curve per hour

E.coli Δ*CKL* and Δ*CKL* Δ*clpX* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (Δ*CKL*: ▲, Δ*CKL* Δ*clpX*: ▼, left hand axis). In parallel cells were harvested and the supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) before SDS-PAGE and Western blot analysis using anti-flagellin (H48) antibody and an HRP secondary. The amount of protein secreted per hour is shown in mg L^{-1} $Hour^{-1}$ cell culture as calculated from a standardised amount of E2 protein using ImageJ based densitometry (Δ*CKL*: black bar, Δ*CKL* Δ*clpX*: grey bar, right hand axis) Results from two biological replicates.

One of the advantages of measuring both secreted and intracellular protein is that a broader picture of strain effects can be inferred, to include translational and secretion capacity of strains. By calculating the total concentration of protein in cell cultures (both secreted and intracellular), it is then possible to calculate the proportion of total protein which is secreted (expressed as a percentage of total protein). This provided information as to which strain is the most efficient secretor. This is informative as it is probable that the tuning of protein production to secretion capacity may be beneficial to improving the rate of protein secretion. For example if the cell invests metabolic energy in producing recombinant protein, less energy is available to produce secretion machinery.

Figure 3. shows the secretion efficiencies of both Δ*CKL* and Δ*CKL* Δ*clpX* over the growth curve. The Δ*clpX* mutant is more efficient at protein secretion than the *ΔCKL* strain throughout the growth curve. At 6 hours (late log phase) it is over 5 times more efficient at secreting E2 protein than *ΔCKL.* Although it is reasonably consistent, the overall trend is for the efficiency of E2 protein secretion to increase through the growth curve. The highest proportion of secreted protein observed was 0.217%. This indicates very low efficiency of protein secretion. With additional strain improvements and alterations to the fermentation method, it may be possible to improve this.

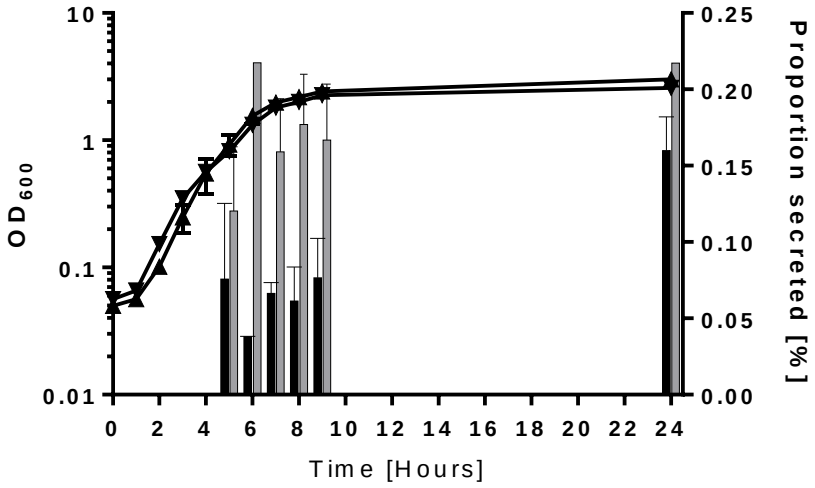


Figure 3.35: The proportion of E2 secreted from the secretor strain and Δ*clpX* mutant during the growth curve

E.coli Δ*CKL* and Δ*CKL* Δ*clpX* containing the plasmid pTrc E2 was grown in LB and the OD_{600} assayed (Δ*CKL*: ▲, Δ*CKL* Δ*clpX*: ▼, left hand axis). In parallel 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells precipitated with TCA (10% v/v) and prepared with 2x SDS-PAGE loading buffer. Samples underwent SDS-PAGE and Western blot analysis using anti-flagellin (H48) antibody and an HRP secondary. Following the calculation of the amount of protein secreted or located intracellularly (mg L^{-1} cell culture), the percentage of secreted protein as a proportion of total protein was ascertained (Δ*CKL*: black bar, Δ*CKL* Δ*dksA*: grey bar, right hand axis). Results from two biological replicates.

As the measure of secretion efficiency of the FT3SS is devised by comparing secreted and intracellular protein, it cannot distinguish the route by which protein appears in the supernatant. A high value for secretion efficiency may suggest that cell lysis is occurring. This

was investigated by stripping the nitrocellulose membrane used for Western blotting of antibodies and reprobing this with an anti-GroEL and suitable secondary-HRP conjugated antibody (Figure 3.). As GroEL is a cytoplasmic contaminant, the presence of it in the secreted fraction suggests some level cell of lysis has occurred- although this cannot be related to cell numbers or concentrations of GroEL. As a degree of cell lysis is expected in cell culture, a low level of GroEL is always expected to be present in samples (visible band at all time points, in both strains in Figure 3.) and is never observed at a high enough level to suggest it accounts for the differences in secretion levels observed. The exception to this is seen following 24 hours growth, as a higher concentration of GroEL is detected in the secreted fraction (more so with the *clpX* knockout mutation), suggesting an increased rate of cell lysis. This means that E2 found in the extracellular fraction following 24 hours growth, may in part be due to cells lysis, rather than directed FT3SS secretion.

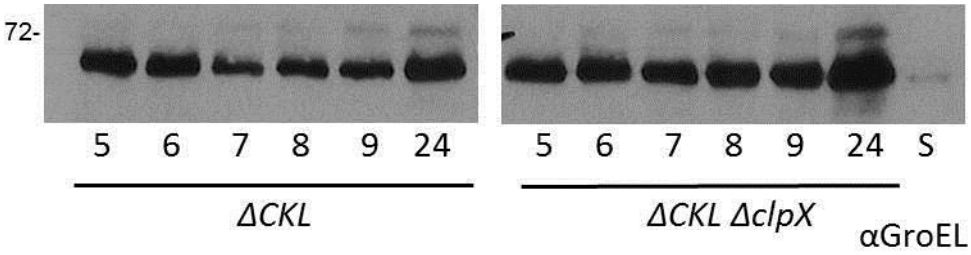


Figure 3.36: Abundance of GroEL in the supernatant of *ΔCKL* and the Δ*clpX* mutant through the growth curve

E.coli ΔCKL (C) or *ΔCKL ΔclpX* (C) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at hourly intervals and the OD_{600} measured. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. An E2 protein standard (S) serves as a negative control. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using anti-GroEL antibody and a HRP secondary.

The effect of the Δ*clpX* mutation appears to have a positive effect of both the amount of E2 protein produced intracellularly, the concentration secreted protein, the rate at which this

protein is secreted and finally the proportion of total protein that is secreted. This is very promising in terms of the aim of improving protein secretion through the fT3SS.

It was hypothesised that an increase in FT3SS secretion would be observed in *ΔclpX* mutants due to decreased $FlhD_4C_2$ complex degradation and therefore increased $FlhD_4C_2$ half-life, which will amount to increased flagella gene expression. In order to investigate whether this increase in protein secretion is due to increased flagella gene expression, gene expression assays were carried out. As described previously, the use plasmids which harbour a flagella gene promoter fragment, fused to a truncated *lacZ* gene (which encodes β galactosidase) allowed gene expression to be measured based on the rate of ONPG cleavage. Despite several rounds of optimisation and the use of plasmids harbouring either the *lacZ-fliA* promoter fusion used previously, or a *lacZ-flhB* promoter fusion it was not possible to obtain results for gene expression in the *ΔclpX* mutant. As assays of β galactosidase activity through *lacZ* fusions in Δ*clpX* mutants is reported in the literature, it is unlikely due to incompatibility. In fact one example investigates β galactosidase activity through a *fliA-lacZ* fusion in Δ*clpXP* mutant strains, and concludes that *fliA* gene expression is increased almost 4 fold with the Δ*clpXP* mutation (Kitagawa et al., 2011). While this investigation was carried out in enterohaemorrhagic *E. coli,* the experimental procedure was otherwise very similar, although this was in a Δ*clpXP* not Δ*clpX* background, so potentially free ClpP could be affecting β galactosidase, although this has never been reported. Another possibility is that flagella gene expression has not been increased, instead the increase in E2 expression may be due to the absence of ClpX directed degradation of FliC (E2), RNA polymerase or ribosome degradation for example (although there is no experimental evidence for this in the literature). As it is evident from the literature that the removal of *clpX* from the chromosome of *E. coli* resulted in increased flagella gene expression (Tomoyasu et al., 2003; Kitagawa et al., 2011), it is postulated that this is the route by which E2 protein secretion was increased in the Δ*CKL ΔclpX* strain.

While *ΔclpX* had a positive effect of the concentration of E2 protein secreted through the FT3SS into the media, the literature also highlights that ClpX protease acts on many proteins in addition to $FlhD_4C_2$ and alters many gene regulatory pathways in addition to flagella. ClpX protease confers protein quality control, control of stress response and DNA damage, stationary phase gene expression and cell division (Flynn et al. 2003; Baker & Sauer, 2012) It has been previously shown that as ClpX modulates the balance of cell division proteins, the

absence of ClpX protein leads to elongated filamentous *E. coli* (MC784 strain) and that cell division is delayed (Camberg et al., 2011). The phenotype of Δ*CKL* Δ*clpX* was investigated to observe whether it is phenotypically altered in any way –both by observing the growth and visual phenotype of the strain in comparison to Δ*CKL*.

3.4.6.4. Effects of deletion of *clpX* on growth phenotype

A number of factors can alter the growth phenotype of cells (chromosomal mutations, the production of plasmid derived protein, growth conditions). The effect of this can vary from marginally increasing or decreasing the growth rate, altering the time at which stages of the growth curve are reached or entirely compromising cell growth, should the strain be non-viable. The *clpX* knockout mutant proved beneficial in increasing the concentration of E2 protein both produced and secreted -this modification may also alter the growth phenotype of the mutant in comparison to the Δ*CKL* strain. In addition to this the expression of plasmid DNA and production of protein can also change the growth phenotype of cells. This effect is usually detrimental to cell growth, as metabolic energy is diverted to non-chromosomal protein production and away from growth processed. Both the effect of Δ*clpX* and the production of plasmid derived E2 was investigated with regard to cell growth.

Figure 3. shows that the induction of gene expression in cells harbouring pTrc E2 resulted in a decrease in growth rate. Cells exhibited a prolonged lag phase (until 1.5 or 2.5 hours, as opposed to 1 hour) and the rate of growth (mean generation time increased 1.36 and 1.49 fold for Δ*CKL* and Δ*clpX* respectively) in the log phase was reduced. However in late stationary phase similar final OD_{600} values for the cultures were achieved. This effect was more pronounced in Δ*clpX* cells, which were shown to exhibit significantly lower optical densities throughout the growth curve when expressing pTrc E2 in comparison to empty vector ($p = <0.0001$. Paired t-test of average OD_{600} values at each time point). The introduction of Δ*clpX* to cells resulted in a lower rate of growth in cells expressing either empty pTrc or pTrc E2. In cells expressing empty vector, this effect was not significant –while the addition of Δ*clpX* did compromise growth rate it was not sufficient enough to suggest the mutation is very detrimental to cell growth. Optical densities at all times were significantly lower for Δ*clpX* + pTrc E2 than all other strain and plasmid combinations vector ($p = <0.0001$. Paired t-test of average OD_{600} values at each time point). The lag phase was 1.5 hours longer, growth rate in

log phase slower and the time taken to reach stationary phase longer. This is likely to be an additive effect of the reduced growth rate of both cells expressing pTrc E2 and the Δ*clpX* mutation, although it would appear that the production of E2 protein contributes more to this effect. Growth of Δ*clpX* +pTrc E2 was still satisfactory, however it reveals that while the Δ*clpX* strain may be preferable for secretion of E2 protein, it may require a longer incubation time, as the growth lags.

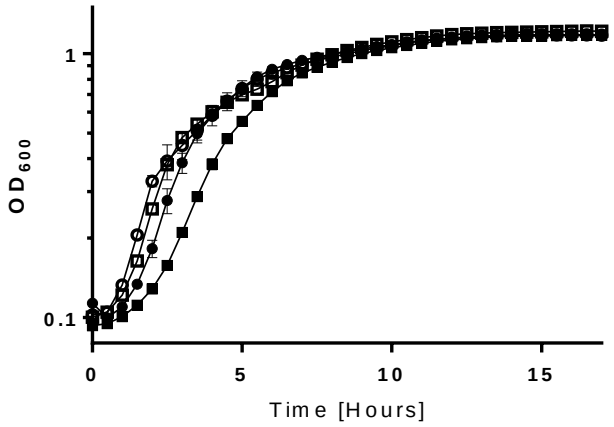


Figure 3.37: Growth of the secretor strain and the Δ*clpX* mutant with pTrc or pTrc E2 plasmid in LB media over time in a 96 well plate. Calibrated to LB media.

E.coli Δ*CKL* and Δ*CKL* Δ*clpX* containing either empty pTrc vector or the pTrc E2 were grown in 96 well plate wells in LB and the OD_{600} assayed. Cells were supplemented with 0.05mM IPTG to induce plasmid expression. Results were calibrated to values for LB only values. Averages were calculated to account for technical repeats and then biological repeats. The standard error of the mean is shown for biological repeats (N=3). Δ*CKL* + pTrc E2: ●, Δ*CKL* + pTrc empty: ○, Δ*CKL* Δ*clpX* +pTrc E2: ■ , Δ*CKL* Δ*clpX* +pTrc empty: □

3.4.6.5. Effects of deletion of *clpX* on cell morphology

Previous studies have shown that the absence of ClpX protein leads to elongated filamentous *E. coli* (Camberg et al., 2011). This observation was made in a MG1655 background strain with a Δ*minC* deletion, this knockout mutation resulted in elongated cell phenotype, presumably because MinC inhibits the cell division protein FtsZ. The removal of ClpX in the Δ*minC* strain resulted in increased cell elongation. Therefore the overall conclusion was that removal of ClpX

resulted in a higher abundance of proteins which delay cell division and lead to cell elongation. This was investigated by measuring the length of Δ*CKL* cells -with and without the Δ*clpX* mutation- following growth in liquid culture, prior to fixing cells onto microscope slides and Gram staining cells. Images were captured at 100x magnification and a scale was added to the images (Figure 3.). Cell length was measured using ImageJ from these images. A range of lengths are visible by eye and care was taken not to preferentially select cells of a certain length, instead a small area was selected and 33 cells lengths were measured for each biological sample. Cell morphology looks similar by eye, although it is evident that there are some very elongated cells present in the Δ*CKL* Δ*clpX* images (i.e. Figure 3.: Δ*CKL* Δ*clpX* 3).

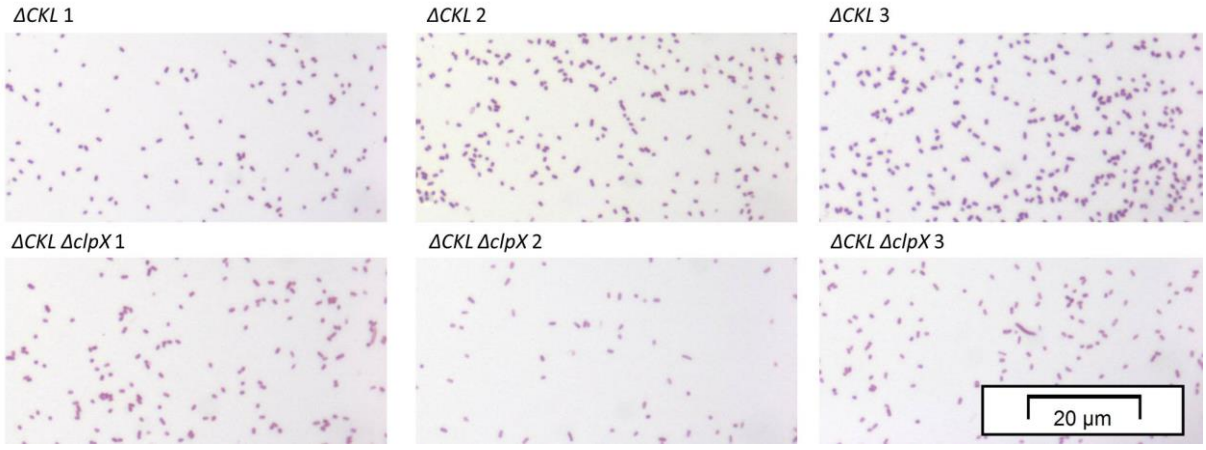


Figure 3.38: Microscope images of Gram stained Δ*CKL* and Δ*CKL* Δ*clpX* cells.

Following liquid culture cells were fixed onto microscope slides and Gram stained before imaging under 100x magnification. Three biological replicates of each strain: top: Δ*CKL,* bottom: Δ*CKL* Δ*clpX.* A 20µm scale was added to images. Cell length was measured using ImageJ and adjusted by calibration to the scale to calculate the length of cells.

The results following measurement of cell length show that Δ*clpX* results in elongation of cells from an average of 0.611µm for Δ*CKL* to 0.782µm for Δ*CKL* Δ*clpX*, this represents a 1.28 fold increase which was found to be statistically significant (Figure 3.: left). Sorting this data into bins with 0.1µM increments for the centre of each bin, revealed that the distribution of sizes differs for the two strains (Figure 3.: right). Δ*CKL* Δ*clpX* were more commonly larger in length

that Δ*CKL.* The modal length of Δ*CKL* cells was centred at 0.5µM, whereas for Δ*CKL* Δ*clpX* it was 0.9µM. The distribution of Δ*CKL* Δ*clpX* cell length extended beyond the highest length seen for Δ*CKL* (1.0µM), to 1.5 µM.

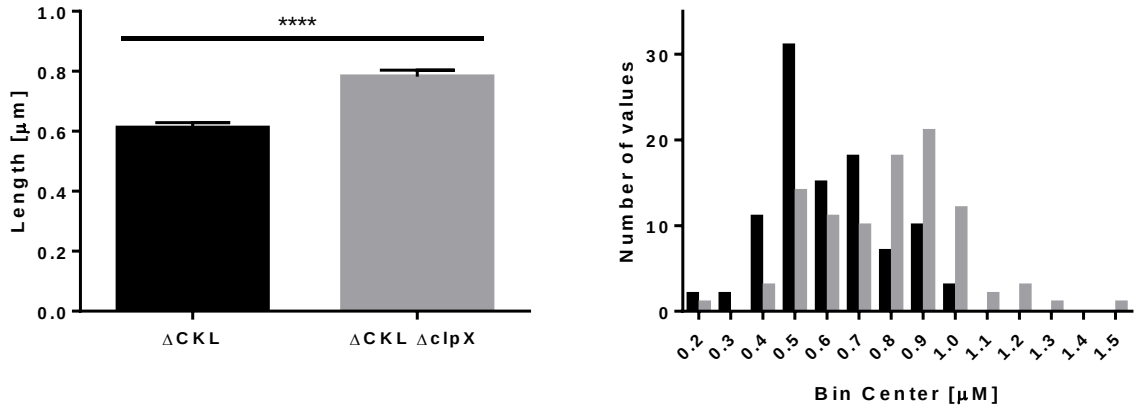


Figure 3.39: Length of Gram stained Δ*CKL* and Δ*CKL* Δ*clpX* cells and histogram following bin sorting for length.

Following liquid culture cells were fixed onto microscope slides and Gram stained before imaging under 100x magnification and length under analysis with ImageJ. Cell length was adjusted by calibration to the scale to calculate the length of cells. N = 99 (33 measurements from each biological replicate; three biological replicates of each strain). Left: Average length of cells. Standard error of the mean shown. **** = p <0.0001. Unpaired t-test. Right: Results sorted into bins for length (0.1µM centres). Black: Δ*CKL*. Grey: Δ*CKL* Δ*clpX*

3.5. Negative control of FT3SS secretion

Altering the secretion capacity of the FT3SS by altered *flhDC* expression was not feasible in the case of *lrhA* or *dksA,* however some progress was made with Δ*clpX.* Prior to implementing additional strategies to improve FT3SS secretion, negative control of FT3SS secretion was investigated, to ensure the presence of E2 in the supernatant is indicative of FT3SS directed secretion and no other route.

3.5.1. Rationale for mutagenesis: negative control of protein secretion through the FT3SS

As the E2 protein harbours the secretion signal for FT3SS, it is expected that this is the route for protein secretion into the supernatant. Currently the low and consistent concentration of GroEL detected in the supernatant of cells, suggests that cell lysis is not prevalent enough to be responsible for the presence of the majority of E2 visible in the supernatant of cells. To confirm this with more confidence, strains were developed which do not form fully functional secretion apparatus to observe whether E2 protein would be found in the supernatant of these cultures.

3.5.1.1. *flhDC*

As described the class I operon *flhDC* sits at the top of the flagella gene transcriptional hierarchy (Figure 1.6) and initiates all flagella gene expression (whether directly or indirectly). The removal of *flhDC* in the Δ*CKL* strain will result in the total absence of flagella gene expression and thus lack flagella structures (i.e. secretion apparatus). The basis of the secretion platform is the directed secretion of proteins which harbour the 47 amino acid FliC secretion signal though the FT3SS, therefore secretion of FliC secretion signal tagged proteins will not be achieved in strains void of secretion apparatus. Any protein observed in the supernatant will be a result of undirected extracellular transit i.e. cell lysis, cytoplasmic leakage.

3.5.1.2. *flgDE*

The rationale behind the Δ*flgDE* mutant concerns the early and late substrates of flagella gene expression. As flagella synthesis is a metabolically expensive process, it is under tight regulatory control, with checkpoints throughout. A key point of control of flagella gene expression is the inhibition of class III gene transcription until the hook structure is *in situ* (Figure 1.5). Prior to the formation of the hook structure FlgM is retained by the cell and inhibits class III gene expression by binding to the transcriptional initiator of class III gene expression: FliA (σ^{28} factor). Upon completion FlgM is secreted through the hook structure and

class III gene transcription is initiated by free σ^{28} (Claret & Hughes, 2002; Karlinsey et al. 2000). The Δ*flgDE* strain should not secrete FliC, although this is not just due to inhibition of class III gene expression, the strain will also be 'early locked' in terms of substrates and late substrates (FliC included) will not be secreted. Previous work has demonstrated that FliC secretion signal tagged proteins are not secreted through hook deficient strains. As stated, FliC would not usually be present in a Δ*flgDE* strain (Stafford et al., 2007). However as the design of the secretion construct allows induction of plasmid expression via the IPTG inducible *trc* promoter of the pTrc vector, it will be possible to express the FliC variant E2 in this strain, regardless of the incomplete hook-basal body structure.

3.5.2. Mutant generation: *flhDC* and *flgDE*

3.5.2.1. Knockout mutagenesis via phage transduction: *flhDC*

Attempts to use P1 phage transduction to transfer a kanamycin linked Δ*flhDC* mutation from an MC1000 Δ*flhDC* parent stain into MC1000 Δ*CKL* consistently resulted in strains with poor growth in liquid culture –this could possibly be attributed to the retention of phage, however the sodium citrate procedure was carried out multiple times to eliminate phage. Therefore the MC1000 Δ*flhDC* parent strain was utilised as a negative control. While the absence of *flgKL* and *fliC* from the chromosome would have represented a true negative control, the fact that the absence of *flhDC* will in turn ensure that these genes (along with all flagella genes) are not expressed –will effectively result in the same phenotype.

3.5.2.2. Knockout mutagenesis via the Lambda Red recombinase method: *flgDE*

As a strain with the *flgDE* knockout mutation was not readily available, it was necessary to derive Δ*CKL* Δ*flgDE de novo* using the Lambda Red recombineering method. Primers were designed to amplify the FRT flanked resistance gene from the appropriate template plasmid (pKD4). Successful PCR reactions were confirmed by the presence of product DNA which corresponded to the anticipated PCR product size of 1.6 kbp. A positive result was achieved (Figure 3.), the band was excised and DNA isolated. Product DNA comprises of FRT flanked

kanamycin resistance cassette, with homologous ends to the up- and downstream regions of *flgDE.*

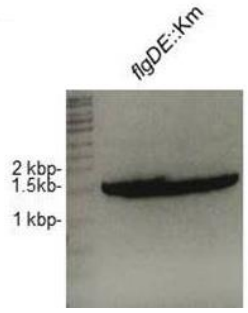


Figure 3.40: DNA gel of gene knockout construct PCR products.

PCR reactions were carried out using the pKD4 template and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of the *flgDE* gene knockout construct harbouring the kanamycin resistance cassette

---

Once isolated, the linear PCR derived DNA was electroporated into parent strains expressing Lambda Red recombinase enzymes. This resulted in the linear DNA template being inserted into chromosomal DNA at homologous regions. Successful recombination events were screened by plating on agar plates supplemented with kanamycin. Positive colonies were observed for all and following the curing of the Red recombinase plasmid so that no further recombination events occur, the deletion of *flgDE* was confirmed with PCR.

3.5.2.3. PCR conformation of knockout mutagenesis: *flgDE*

To confirm that the acquisition of antibiotic resistance was indicative of the removal of the gene targeted PCR reactions using combinations of primers which anneal either in the antibiotic resistance cassette or the region of chromosomal DNA upstream of *flgDE* were carried out. The presence of a 1333kbp PCR product confirmed *flgDE::Km* knockout mutagenesis (Figure 3.).

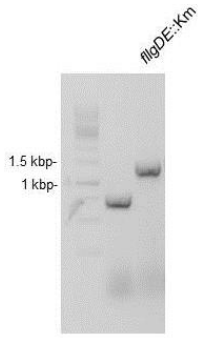


Figure 3.41: Agarose DNA gel of PCR product to confirm successful knockout mutagenesis of *flgDE*.

PCR reactions were carried out using the templates and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of a 1333 kbp PCR product.

3.5.3. Mutant characterisation: Further optimisation of the secretion assay

Prior to the assessment of the secretion capacity of the negative control strains, the secretion assay was optimised to improve throughput. While informative, this secretion assay was time consuming, as samples and measurements must be taken every hour during growth. In addition, as sample must be removed from the cultures every hour, cultures must be a large volume –this limits the number of experiments which can be carried out in tandem, due to limited incubator space. These two factors limit the throughput of the secretion assay, therefore the assay was modified so that secretion was assessed based on a single time point, allowing smaller culture volumes to be grown.

As the secretor strain was grown alongside all of the growth curve secretion assays for the strains which were designed to increase flagella master regulator expression (Δ*lrhA,* Δ*dksA*, Δ*clpX*), a wealth of information about the secretion capacity of the original secretor strain

through the growth curve has been gathered in this thesis (Figure 3., Figure 3. and Figure 3.). Overall, so far in this thesis, the highest concentration of E2 in the supernatant of Δ*CKL* and fastest secretion rate was observed between OD_{600} 1 and 2. Therefore OD_{600} 1.5 was selected as the optimal time point to measure secretion capacity of strains expressing E2 protein herein. The volume of cells grown was scaled down from 100mL to 10mL cultures, allowing many experimental variables and repeats to be tested in one experiment.

3.5.4. Negative control mutant characterisation: Effects of deletion of *flhDC* and *flgDE* on FT3SS secretion

While the secretion capacity for E2 protein by the original secretor strain and mutants has been investigated it cannot be confirmed that cell lysis, a leaky cell membrane or undirected protein secretion through other secretion systems are not the cause of this. Routine Coomassie stain SDS-PAGE gels of the supernatant of cell cultures provides evidence cell lysis is not prevalent, as supernatant fractions are clear of contaminating proteins. Further to this Western blotting of supernatant fractions and probing with an antibody for the cytoplasmic chaperone protein GroEL does not reveal major cytoplasmic contamination of the supernatant prior to the late stationary phase (Figure 3.). Despite this information, the utilisation of a negative control for FT3SS secretion would add more power to this assumption. Protein expression and secretion of E2 was measured in the original secretion strain MC1000 Δ*CKL*, the most promising mutant strain MC1000 Δ*CKL* Δ*clpX* and the two negative controls MC1000 Δ*flhDC* and MC1000 Δ*CKL* Δ*flgDE* (Figure 3.).

Aside from the 42kDa E2 protein in the supernatant fractions no other protein is visible in the Coomassie stained SDS-PAGE, showing that the supernatant is free of protein contaminants. It is apparent that E2 is overexpressed in the cell fraction of all strains; however it is only visible in the supernatant of Δ*CKL* and Δ*CKL* Δ*clpX*. In this experiment the concentration of secreted E2 is higher in the Δ*clpX* mutant. The same pattern of expression and secretion is observed in the H48 antibody probed Western Blot of the same protein samples. This shows that the two methods give an accurate representation of the concentration of E2 protein present both in the cell and secreted fractions of all strains. Despite the E2 seen in the Coomassie strained SDS-PAGE being more linear, the Western blot underwent densitometry analysis, as the cell

fractions of the Coomassie stain are not suitable *for* analysis by this method, due to the abundance of other protein. However it is evident from the Coomassie stained SDS-PAGE (Figure 3.: top), that E2 is visible in the intracellular fractions of all strains at a similar concentration, with the exception of Δ*CKL* Δ*flgDE* where the concentration is lower. While no E2 is seen in the secreted fractions of Δ*flhDC* and Δ*flgDE* mutants, around four times more is seen in the supernatant of Δ*CKL* Δ*clpX* cells in comparison to Δ*CKL*. Densitometry analysis of Coomassie stained SDS-PAGE shown confirmed this to be 3.8 times more E2.

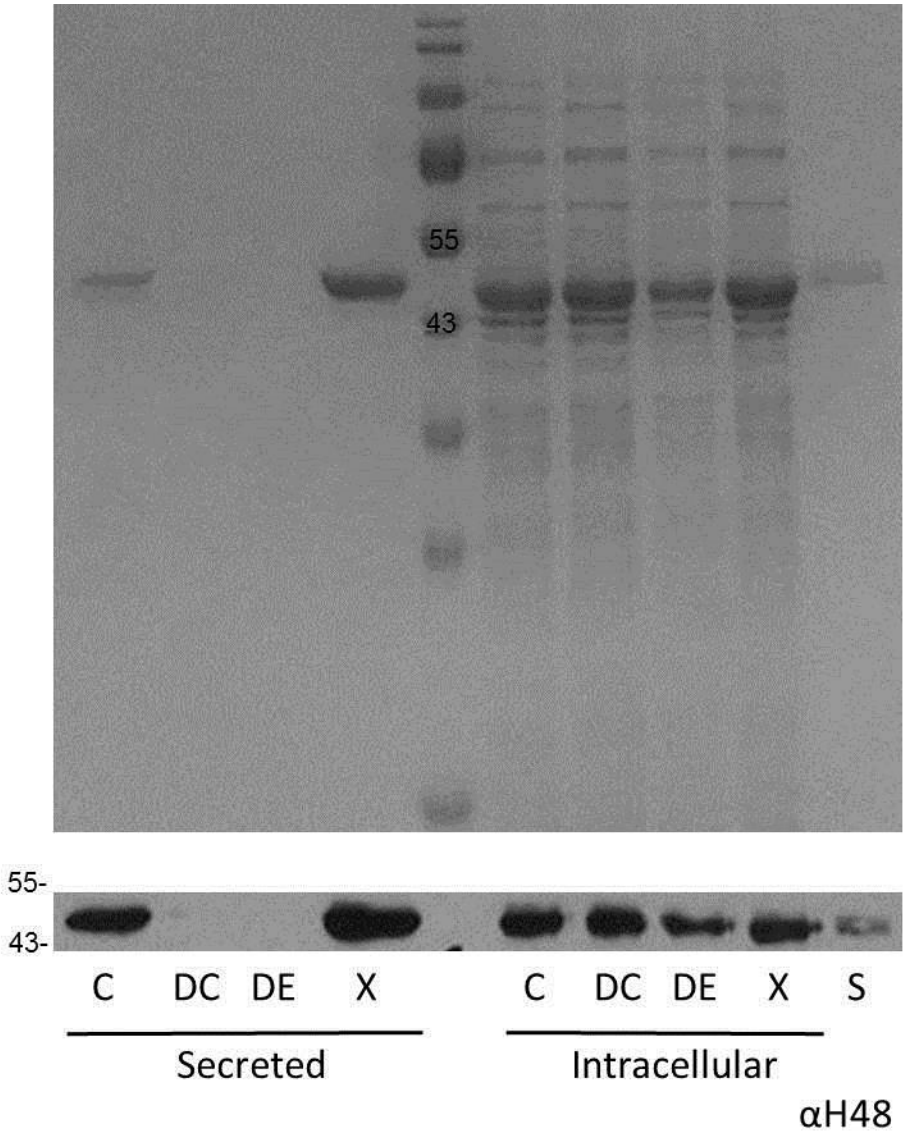


Figure 3.42: Intracellular and secreted E2 from the secretor strain, Δ*flhDC,* Δ*flgDE* and Δ*clpX* mutants.

E.coli Δ*CKL* (C), Δ*flhDC* (DC)*,* Δ*CKL* Δ*flgDE* (DE) or Δ*CKL* Δ*clpX* (X) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at OD_{600} 1.5. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and either Coomassie staining with Instant Blue or Western blot analysis of cells and supernatant using anti-flagellin (H48) antibody and an HRP secondary. An E2 protein standard (S) was loaded to allow quantification of intracellular and secreted protein concentration. Intracellular: 5µL; secreted: 15µL; standard: 5µL.

Densitometry analysis by ImageJ was carried out on two biological repeats of the experiment. The use of the E2 protein standard allowed the protein concentration to be quantified for both intracellular and secreted protein, this enabled values from the two replicates to be combined (Figure 3.). A higher concentration of E2 was found in the intracellular fractions of mutant strains in comparison to the Δ*CKL.* Δ*clpX* displayed both the highest intracellular (1350 mg L^{-1} cell culture) and extracellular (23.457 mg L^{-1}) E2 concentration. No E2 protein was found in the secreted fraction of either MC1000 Δ*flhDC* or MC1000 Δ*CKL* Δ*flgDE.*

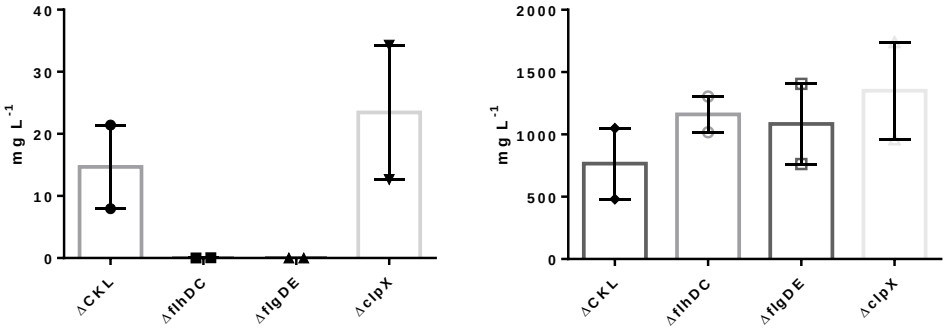


Figure 3.43: Intracellular and secreted E2 from the secretor strain, Δ*flhDC,* Δ*flgDE* and Δ*clpX* mutants.

E.coli Δ*CKL*, Δ*flhDC,* Δ*CKL* Δ*flgDE* or Δ*CKL* Δ*clpX* containing the plasmid pTrc E2 were grown in LB to OD_{600} 1.5. An E2 protein standard was loaded to allow quantification of intracellular and secreted protein concentration. Following Western blot analysis of cells and supernatant using anti-flagellin (H48) antibody and an HRP secondary, densitometry analysis was carried out using ImageJ. The amount of protein located intracellularly (cells) or secreted (SN) is shown in mg L^{-1} cell culture as calculated from a standardised amount of E2 protein. Data from two biological replicates. Individual data points, mean and standard error of the mean are given.

In addition to this the same protein samples were also examined for the presence of GroEL by stripping the Western blot probed with αH48 antibody and reprobing with αGroEL antibody. As a cytoplasmic chaperone protein GroEL is expected to be present in similar concentrations in intracellular protein fractions across strain. The presence of GroEL in the secreted fraction

indicated some leakage of protein from the cytoplasm into the supernatant. It is apparent that the abundance of GroEL in the intracellular fractions of the strains is very similar –while this is expected it is also a useful indicator that protein loading onto the SDS-PAGE is accurate (Figure 3.). Less GroEL is present in the supernatant in comparison to cells; there is more GroEL present in the supernatant of MC1000 Δ*flhDC* than the other three strains.

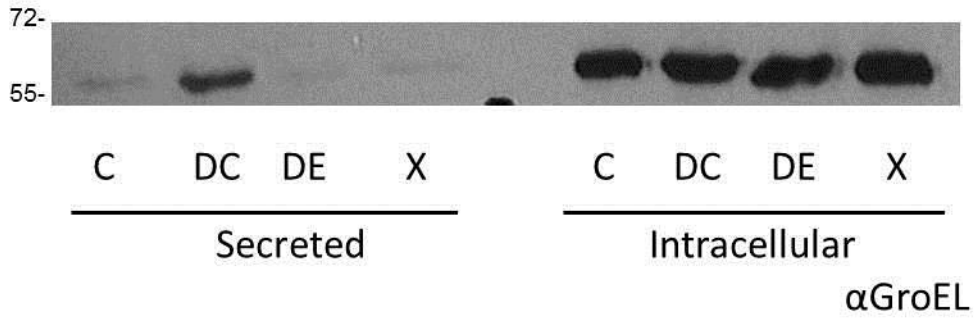


Figure 3.44: Abundance of GroEL in the intracellular and supernatant fractions of *ΔCKL* and the Δ*flhDC, ΔflgDE* and Δ*clpX* mutants, following expression and secretion of E2 protein through the FT3SS.

E.coli ΔCKL (C), *ΔflhDC* (DC)*, ΔCKL ΔflgDE* (DE) or *ΔCKL ΔclpX* (X) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at OD_{600} 1.5. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using anti-groEL antibody and an HRP secondary. Intracellular: 5µL; secreted: 15µL.

Densitometry allowed comparison of the abundance of GroEL in both the cell and secreted fractions of all strains in both biological replicates. As a GroEL protein standard was not used it was not possible to absolutely quantify the concentration of GroEL, therefore to enable the comparison of results across biological replicates densitometry results were normalised to the value calculated for either secreted or intracellular Δ*CKL* (Figure 3.). As seen in Figure 3., the abundance of intracellular GroEL of all strains was very similar and demonstrated low variability –suggesting accurate loading of cell protein onto the SDS-PAGE. This is interesting as it has been reported that GroEL levels are elevated in *clpX* knockout mutants (Weichart et al., 2003). In secreted fractions of protein very low levels of GroEL were detected in both Δ*flgDE* and Δ*clpX* mutants, suggesting that cell lysis and cytoplasmic leakage were low. The concentration of GroEL in Δ*flhDC* was slightly higher overall than in the secretion apparatus

mutant. Despite the presence of cytoplasmic protein in the supernatant of Δ*flhDC*, E2 protein was not detected (Figure 3.). This suggests E2 detected in the supernatant of other strains (Δ*CKL* and Δ*CKL* Δ*clpX*) has been secreted through the modified FT3SS.

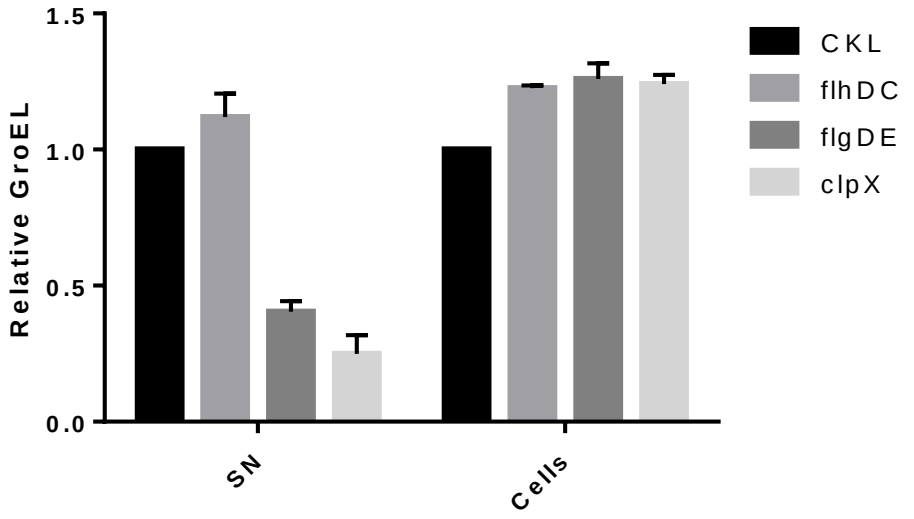


Figure 3.45: Densitometry analysis of the abundance of GroEL in the intracellular and supernatant fractions of *ΔCKL* and the Δ*flhDC, ΔflgDE* and Δ*clpX* mutants, following expression and secretion of E2 protein through the FT3SS.

E.coli ΔCKL, *ΔflhDC, ΔCKL ΔflgDE,* or *ΔCKL ΔclpX* containing the plasmid pTrc E2 was grown in LB. Cells were harvested at OD_{600} 1.5. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant (SN) using anti-GroEL antibody and an HRP secondary. Densitometry analysis was then carried out, and results normalised to the value derived for Δ*CKL* to allow comparison across biological replicates (N=2).

The complete absence of E2 protein observed in the supernatants of both Δ*flhDC* and Δ*CKL* Δ*flgDE* confirm that they are both suitable strains to demonstrate negative control of FliC secretion signal directed E2 protein secretion through the FT3SS. Complementary to this finding is the evidence that cell lysis or cytoplasmic leakage are not prevalent in the positive secretion strains; contaminating protein is not observed in the Coomassie strains of the supernatant fractions and the concentration of GroEL is not higher in positive strains in comparison to negative strains.

Now that negative control of E2 protein secretion has been established, it can be concluded with confidence that E2 found in the supernatant fraction is present due to directed FT3SS secretion.

3.6. Improved FT3SS secretion: reduced metabolic burden and negative feedback

3.6.1. Rationale for mutants: reduced metabolic burden and feedback mutants

An alternative strategy is to improve flagella gene expression by reducing internal negative feedback which occurs within the FT3SS gene expression hierarchy. An additional strategy would be to remove FT3SS genes which are obsolete in the modified FT3SS mutant. This may include genes linked to motility or precursors or chaperones to deleted genes. It is presumed that the removal of any gene will result in decreased metabolic cell burden and therefore increased energy for other processes (such cell growth, the production of secretion apparatus or plasmid protein). The rationale for the mutants (Δ*CKL*) Δ*motAB,* Δ*fliDST* and Δ*flgMN* comprise of a combination of these strategies.

3.6.1.1. *motAB*

The MotAB complex attaches to the cell wall via the MotB protein forming the flagella motor stator. In addition the complex forms the channel which allows proton and sodium transport down their electrochemical potential gradient through the cell membrane. This process provides energy (known as proton or sodium motive force) which drives the flagella motor (Hosking et al. 2006; Minamino et al. 2016). MotB blocks this channel prior to the other FT3SS elements being in place, only unlocking it when it comes into contact with the basal body motor. Overexpression of MotAB increases the proton permeability of cells and causes a reduction in proton motive force, which results in growth impairment (Blair & Berg, 1990;

Zhou et al. 1998). Whereas, increased proton motive force has been shown to increase secretion of FT3SS substrates through the existing structure (Minamino et al. 2016). MotAB has been shown to be essential in flagella rotation but not assembly, and therefore by proxy one can infer secretion (Blair, 1995). This suggests that deleting MotAB in the Δ*CKL* strain should have no effect on the secretion properties of the secretion apparatus. The Δ*CKL* Δ*motAB* strain may have a reduced metabolic burden, therefore direct more resources to cell growth and potentially flagella and secretion protein production. In addition to this the knockout mutation may result in reduced proton permeability of cells, therefore an increase in proton motive force and an increase in growth rate and improved secretion due to more secretion apparatus being *in situ*

3.6.1.2. *flgMN*

FlgN is a chaperone to FlgK and FlgL therefore unnecessary in the secretor strain. It is also a translational regulator of class II anti-sigma 28 factor FlgM (Figure 3.) (Aldridge et al., 2003). Loss of FlgN leads to reduced FlgM (FlgM is still present as it is still translated from the class II promoter), resulting in higher levels of sigma 28 gene expression (Karlinsey et al. 2000). Conversely it is reported that reduced levels of sigma 28 are observed in FlgM mutants, this is because FlgM tightly binds sigma 28 and while this inhibits function, it also protects it from proteolysis (Barembruch & Hengge, 2007). Sigma 28 is transcribed from *fliA*, which also has two promoters – one class II and one class III. When uninhibited by FlgM, sigma 28 activates class III gene expression. While the majority of class III genes are involved in rotation or hook and filament proteins and are therefore not conducive in the secretor strain; sigma 28 has been shown to activate transcription of the master regulator *flhDC*, increasing flagella numbers (Clarke & Sperandio, 2005). Deletion of FlgM and FlgN in the secretor strain should have no effect on the existing secretion apparatus structure. It should result in the metabolic burden being reduced and if *fliA* activity is increased, an increase in class III and *flhDC* gene expression, the latter of which may result in an increased number of secretion apparatus.

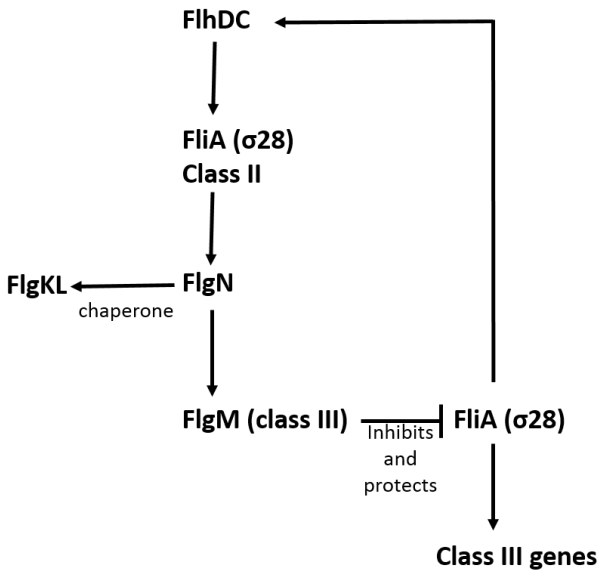


Figure 3.46: Schematic of the gene regulatory network of the FlgMN proteins

3.6.1.3. *fliDST*

Another operon which if removed has potential to both reduce metabolic burden and have an effect on regulation of the master regulator, is *fliDST* (Figure 3.). As discussed previously FliD is the cap protein which aids the assembly of FliC proteins to form the filament. Therefore in a FliC deficient mutant, FliD and the corresponding chaperone FliT, are not necessary (Fraser et al., 1999). Furthermore FliT has been shown to be an anti $FlhD_4C_2$ factor, preventing the protein binding to class II promoters (Yamamoto and Kutsukake, 2006). FliS is the chaperone of FliC, therefore unnecessary in the Δ*fliC* secretor strain (Yokoseki et al., 1995). FliS has also been shown to bind and stabilise flgM in the cytoplasm prior to *fliA* expression (FliA displaces FliS) and supress secretion of FlgM upon hook-basal body completion (Galeva et al., 2014). The removal of all genes in the *fliD* operon have been shown to result in an increase in secreted FlgM (Yokoseki et al., 1996), which as discussed leads to an increase in class III flagella gene transcription. Δ*fliDST* mutants are non-motile, however as motility is not required this is not problematic. The main advantage of this strategy is that FliT will not inhibit $FlhD_4C_2$ and FliS will not bind FlgM, therefore class II and III gene expression will be increased respectively, this may result in the presence of more secretion apparatus (Aldridge et al., 2010).

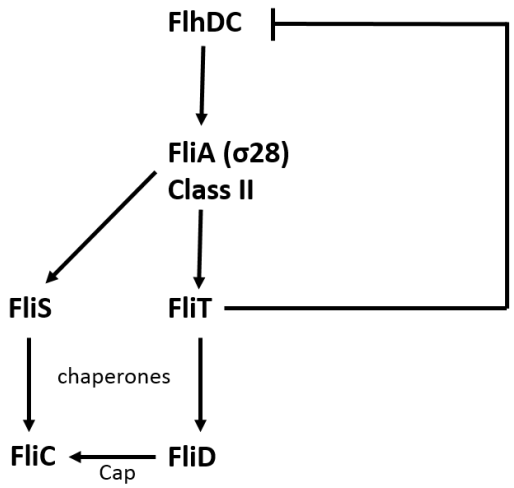


Figure 3.47: **Schematic of the gene regulatory network of *fliDST***

3.6.2. Mutant generation: *fliDST, motAB, flgMN*

3.6.2.1. Knockout mutagenesis via the Lambda Red recombinase method: *fliDST, motAB, flgMN*

Primers were designed to amplify the FRT flanked resistance genes from the appropriate template plasmid. Successful PCR reactions were confirmed by the presence of product DNA which corresponded to the anticipated PCR product size (Table 6.).

Table 6.1: PCR products for gene knockouts

Gene knockout	Antibiotic resistance	Template plasmid	Size of product (kbp)
motAB	Cm	pKD3	1.1
motAB	Km	pKD4	1.6
fliDST	Km	pKD13	1.4
flgMN	Cm	pKD3	1.1
flgMN	Km	pKD4	1.6

Positive results for all PCR experiments were achieved (Figure 3.) -bands of the correct size were cut out and DNA isolated. Product DNA comprises of FRT flanked antibiotic resistance cassettes, with homologous ends to the up- and downstream regions of chromosomal DNA of the gene selected for knockout mutagenesis.

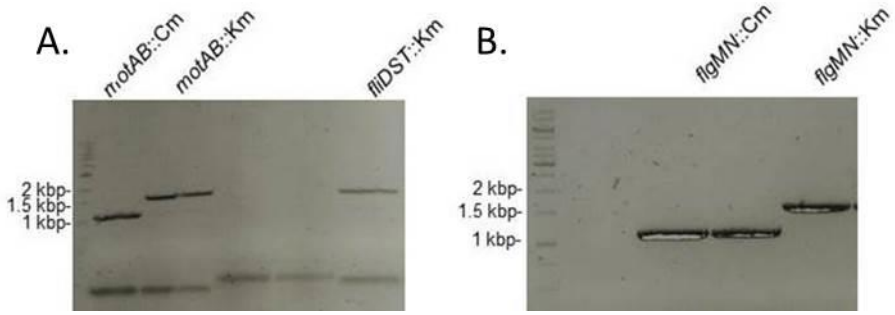


Figure 3.48: DNA gels of gene knockout construct PCR products.

PCR reactions were carried out using the templates and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of the gene knockout constructs harbouring antibiotic resistance cassettes (as listed in Table 6.). A. *motAB*::Cm, *motAB*::Km, *fliDST*::Km. B. *flgMN*::Cm , *flgMN*::Km.

Once isolated, the linear PCR derived DNA was electroporated into parent strains expressing Lambda Red recombinase enzymes. This resulted in the linear DNA template being inserted into chromosomal DNA at homologous regions. Successful recombination events were screened by plating on agar plates supplemented with the relevant antibiotic (chloramphenicol or kanamycin). Positive colonies were observed for all recombination strains and following the curing of the Red recombinase plasmid so that no further recombination events occur, the deletion of chromosomal genes was confirmed with PCR.

3.6.2.2. PCR conformation of knockout mutagenesis: *fliDST, motAB, flgMN*

To confirm that the acquisition of antibiotic resistance was indicative of the removal of the gene targeted PCR reactions using combinations of primers which anneal either in the antibiotic resistance cassette or the region of chromosomal DNA either side of the target knockout gene. Either the size or the presence of the resulting PCR products allowed further confirmation of knockout mutagenesis.

Table 6.2: PCR products for confirmation of knockout mutagenesis in the secretor strain.

Positive results were obtained for the following knockout mutagenesis conformations. The size of PCR products which denote positive results is given along with the size of the PCR product (if relevant) expected if knockout mutagenesis was not successful.

Gene knockout	Antibiotic Resistance	Size of product if positive (kbp)	Size of product if negative (kbp)
motAB	Cm	492	N/A
motAB	Km	777	N/A
fliDST	Km	993	N/A
flgMN	Km	1698	813

Ideally the primer pair used to confirm mutagenesis is comprised of one primer which anneals in the antibiotic resistance cassette and one in the chromosomal region flanking the target knockout gene. When knockout mutagenesis by homologous recombination of the antibiotic resistance cassette into the chromosome is successful a positive result will arise. The size of the PCR product can be predicted. If knockout mutagenesis is not achieved this will result in a null PCR result as the primer designed to anneal to the ABC will not. This is denoted by 'N/A' in Table 6.. In the case of Δ*flgMN::Km* PCR conformation by this method was not successful , instead a primer pair which annealed in chromosomal DNA either side of the target gene was utilised. In this instance a positive PCR result will arise whether mutagenesis is achieved or not.

Here knockout mutagenesis can be confirmed by the size of the resulting product, as the expected size differs depending on whether the primers anneal to chromosomal DNA flanking either the flagella gene or the antibiotic resistance gene.

Positive PCR results were achieved for all flagella gene knockout mutants with the exception of *flgMN*::*Cm*.

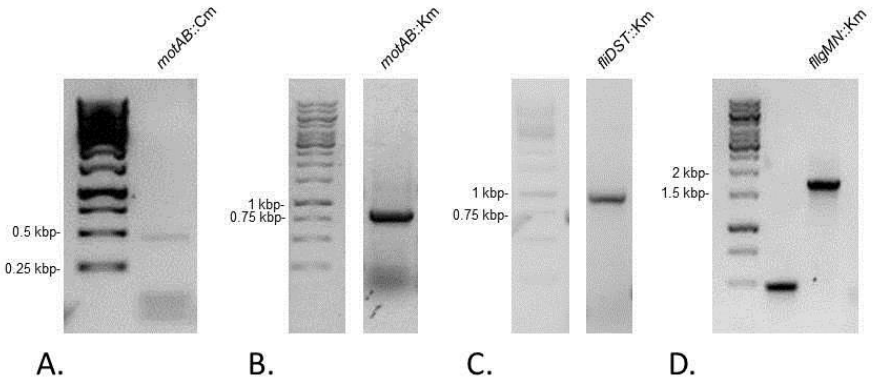


Figure 3.49: Agarose DNA gels of PCR products to confirm successful knockout mutagenesis.

PCR reactions were carried out using the templates and primers in **Appendix 2** with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of products of the sizes listed in Table 6.. A. *motAB*::Cm, B. *motAB*::Km, C. *fliDST*::Km, D. *flgMN*::Km

---

The secretion capacity of these strains was then assessed, to observe whether they resulted in improved rates of protein secretion via the FT3SS.

3.6.3. Reduced metabolic burden and feedback mutants: Effects of deletion of *motAB, flgMN* and *fliDST* on FT3SS secretion

Once derived, the expression and secretion capacity of E2 protein was measured in these strains to ascertain the suitability of these mutants. The expression of pTrc E2 in Δ*motAB* and Δ*flgMN* resulted in secretion of E2 into the media of cells (Figure 3.). E2 was also secreted in

the original secretor strain and the Δ*CKL* Δ*clpX* mutant, allowing comparison of secretion capacity to be made in relation to strains already investigated (Figure 3.). Densitometry analysis was carried out on the secreted fractions of protein shown in Figure 3.. As an E2 protein standard was not included on this SDS-PAGE, the concentration of E2 protein secreted cannot be quantified *per se*, however relative improvement in comparison to Δ*CKL* can be ascertained. This confirmed that Δ*flgMN* secreted a very similar amount (0.94) of E2 protein to Δ*CKL*. Secretion of E2 was 1.52 times higher in the Δ*motAB* mutant in comparison to Δ*CKL*, however the Δ*clpX* mutant secreted the highest concentration of protein (1.82 times Δ*CKL*), which is similar to the fold improvement seen in Figure 3.. The lack of protein (aside from E2) in the supernatant of all strains, suggests that lysis is not prevalent in any of the mutant stains. Due to experimental failure a reliable result could not be obtained for a αH48 (flagellin) antibody probed Western blot, it is not possible to measure the concentration of intracellular E2 protein in strains. The concentration of E2 in the cells of strains appears to be similar, as there is no clear overexpression in any of the strains.

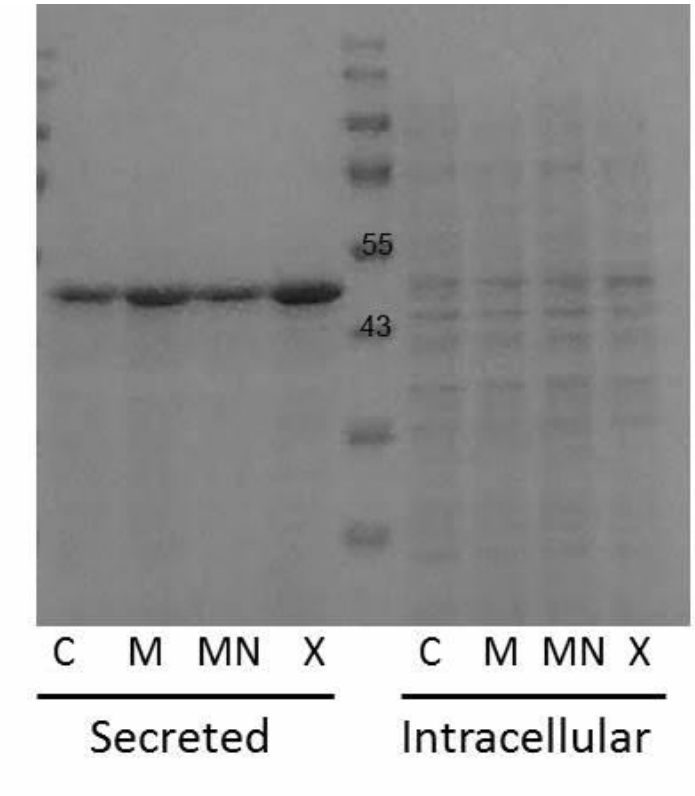


Figure 3.50: Intracellular and secreted E2 from the Δ*CKL* strain, Δ*motAB,* Δ*flgMN* and Δ*clpX* mutants.

E.coli Δ*CKL* (C), Δ*CKL* Δ*motAB* (M)*,* Δ*CKL* Δ*flgMN*(MN) or Δ*CKL* Δ*clpX* (X) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at OD_{600} 1.5. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and Coomassie staining with Instant Blue.

E2 secretion was then quantified in the most promising novel strain: Δ*CKL* Δ*motAB* (Figure 3.) and the as yet untested Δ*CKL* Δ*fliDST* strain. While the concentration of E2 located in the media of Δ*motAB* cells in comparison to the secretor strain looked lower in the Coomassie stained SDS-PAGE, it appeared much higher in the Western blot (Figure 3.). This is an example of the lack of linearity of Western blots. Intracellularly, less E2 was observed in the Δ*motAB* mutant, suggesting that it is also more efficient at secreting E2. In contrast the Δ*fliDST* mutant expressed less E2 than the secretor strain and secreted a small amount (some observed in the Coomassie stain, none detected in the Western blot.) This is presumably due to the similarity of E2 to the FliS chaperoned FliC, this will be discussed later.

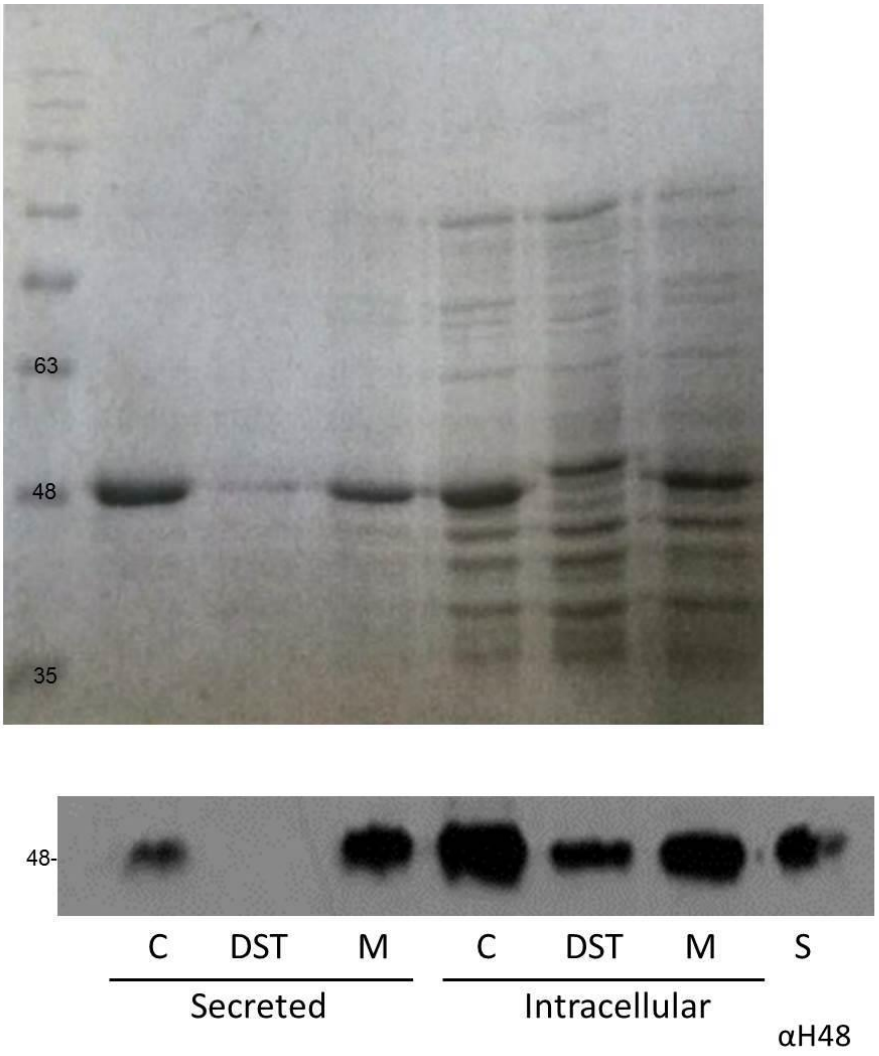


Figure 3.51: Intracellular and secreted E2 from the secretor strain, Δ*fliDST* and Δ*motAB* mutants.

E.coli Δ*CKL* (C), Δ*CKL* Δ*fli*DST (DST) or Δ*CKL* Δ*motAB* (M) containing the plasmid pTrc E2 was grown in LB. Cells were harvested at OD_{600} 1.5. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and either Coomassie staining with Instant Blue or Western blot analysis of cells and supernatant using anti-flagellin (H48) antibody and an HRP secondary. The BLUeye Prestained Protein Ladder was loaded, as was an E2 protein standard (S) to allow quantification of intracellular and secreted protein concentration.

These observations were confirmed with densitometry analysis of the Western blot (Figure 3.). Very little E2 protein was detected in the supernatant of Δ*fliDST* cells, whereas E2 secretion was improved over two fold, from 0.612 µg mL^{-1} in the secretor strain to 1.534 µg mL^{-1} in Δ*motAB.* When secreted E2 is calculated as a proportion of total E2 protein it is evident that Δ*motAB* is not only the strain with the highest secretion capacity, but also the most efficient protein secretor –secreting over 17% of the total E2 protein expressed (Figure 3.).

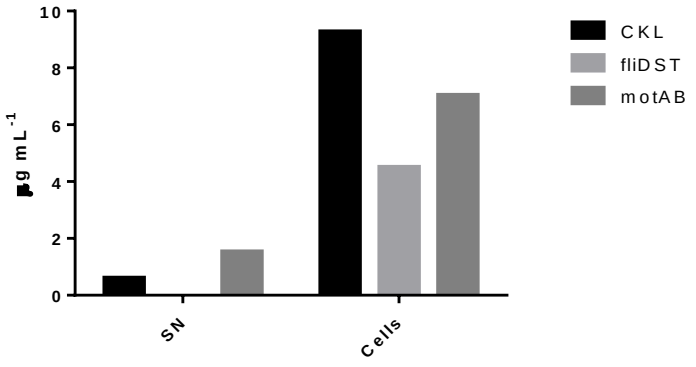


Figure 3.52: Intracellular and secreted E2 from the secretor strain, Δ*fliDST* and Δ*motAB* mutants.

E.coli Δ*CKL*, Δ*CKL* Δ*fliDST* or Δ*CKL* Δ*motAB* containing the plasmid pTrc E2 were grown in LB to OD_{600} 1.5. An E2 protein standard (S) was loaded to allow quantification of intracellular and secreted protein concentration. Following Western blot analysis of cells and supernatant using anti-flagellin (H48) antibody and an HRP secondary, densitometry analysis was carried out using ImageJ. The amount of protein located intracellularly (cells) or secreted (SN) is shown in µg L^{-1} cell culture as calculated from a standardised amount of E2 protein. Δ*flgMN* and Δ*clpX* mutants, following expression and secretion of E2 protein through the FT3SS. .

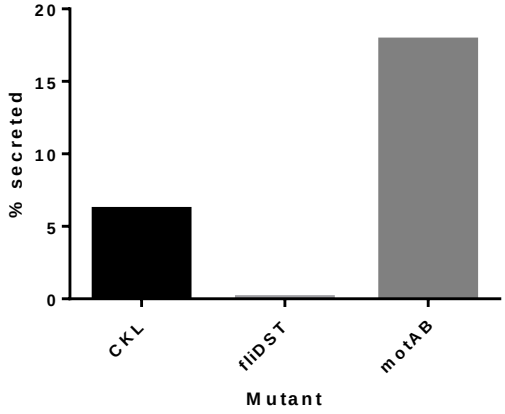


Figure 3.53: The proportion of secreted E2 as a percentage of total E2 expressed and secreted by the secretor strain, Δ*fliDST* and Δ*motAB* mutants, following expression and secretion of E2 protein through the FT3SS.

E.coli Δ*CKL*, Δ*CKL* Δ*fliDST* or Δ*CKL* Δ*motAB,* containing the plasmid pTrc E2 was grown in LB. Cells were harvested at OD_{600} 1.5. Samples underwent SDS-PAGE. Densitometry analysis was then carried out using ImageJ and the proportion of secreted protein calculated as the concentration of secreted protein in comparison to the concentration of total E2 protein (both secreted and intracellular).

3.7. Discussion: a prototype secretion platform for secretion of native protein

The focus of this chapter has been to produce a prototype platform *E. coli* strain for secretion from a base chassis strain. MC1000 was chosen on the basis of its broad similarity to a 'standard' K12 strain but also since it is well characterised and contains a leucine and thiamine auxotrophy for safety. Secretion of protein from the cytoplasm to the extracellular media was achieved in the chassis strain by modifying the flagella type III secretion system (FT3SS) to form a pipe like secretion apparatus. Strain improvements were then employed to this strain with the aim of increasing the amount of secretion apparatus expressed or by removing unnecessary flagella genes to reduce the metabolic burden. It was hypothesised that this would result in more protein being secreted into the media.

Development of assays that quantified secretion and flagella gene expression in strains was central to all of the work in this study. Without an accurate assay method, suitable for application to all strains it would not be possible to assess if strain modification resulted in improved secretion. Protein secretion via the modified FT3SS has been measured in a number of mutant strains -including strains which were designed to exhibit total absence of FT3SS secretion, which serve as negative controls. This could be paired with information on flagella gene expression to give an extensive view of flagella protein secretion in strains. These results are used to assess whether strain improvements have been effective.

3.7.1. Protein secretion assay development

Through knockout mutagenesis it was possible to develop a novel platform strain capable of secreting a variant (E2) of the native flagella filament protein FliC through truncated FT3SS and into the media of cells –this forms the basis of the protein secretion assay, which allowed accurate quantification of protein expression and secretion in mutant strains. The secretion assay provides a method of quantifying protein secretion in strains either as a concentration per volume of media or as a secretion rate or efficiency. The secretion assay was optimised to

report protein secretion in strains throughout the growth curve. Utilisation of a protein standard allowed the calculation of a weight per OD unit or weight per volume of cell culture concentration of protein to be derived from Coomassie stained SDS-PAGE or Western blots. Quantifications of secretion rates are powerful as they provide information on the change in secretion capacity of cells over time or with the implementation of mutations. The w/v measurement is advantageous because it allows direct comparison of secretion efficiency between experiments and is a common industry standard used by contract manufacturing organisations to compare protein yields from fermentations -which is ultimately the aim of this work. A variation of the protein secretion assay involved growing cell cultures in smaller volumes and measuring secretion at one time point; this improved the throughput of the assay, allowing protein secretion to be measured in more mutant strains in parallel experiments. These points combined mean that a wealth of information can be gained from the secretion assay, to aid secretor strain improvement.

3.7.2. A modified FT3SS platform secretion strain

Strain improvements were made to *E. coli* to implement protein secretion through the FT3SS. It was established that E2 protein secretion occurred in both Δ*fliC* Δ*flgKL* (HAPless) and the Δ*fliCD* (CAPless) mutant *E. coli* (Figure 3.). As both the volume of protein produced and secreted was around double that observed in the HAPless strain this formed the basis of further strain optimisation in this study (Figure 3.). This effect may be due to the shorter distance that secreted protein has to pass through to reach the extracellular environment, although it has been demonstrated that length of the existing flagella structure does not alter the rate of secretion (Evans et al., 2013). Additionally it may be due to the effect of FliT on the master regulator. FliT is the chaperone to FliD, therefore in the absence of FliD, more FliT is available to negatively regulate *flhDC*, which in turn will result in downregulation of all flagella genes (Yamamoto and Kutsukake, 2006). Less flagella gene expression will lead to less secretion apparatus and therefore a reduction in secretion. In the HAPless mutant FliD is expressed, therefore FliT is less readily available for interaction with *flhDC*. FlgK and FlgL are secretion competitors to FliC (and any other protein directed for secretion through the FT3SS, furthermore it has been shown that they are preferentially secreted, as their chaperone FlgN

has a higher affinity for the secretion apparatus (Evans et al. 2006); therefore this may account for the reduced concentration of E2 secreted through the CAPless strain, as *flgKL* is still present. Finally as FlgN is the chaperone to FlgK and FlgL, there is more free FlgN when these proteins are absent. As FlgN is a translational regulator of FlgM, this will lead to increased FliA inhibition (but with increased protection), which has been shown to result in increased flagella gene expression (Aldridge et al., 2003).

3.7.3. Negative control of FT3SS secretion

Negative control of E2 secretion through the modified FT3SS was established, confirming that the presence of E2 in the supernatant of strains was due to directed secretion through secretion apparatus and not through cell lysis or cytoplasmic leakage. This was characterised by an absence of E2 protein in the supernatant of cultures of Δ*flhDC* and Δ*CKL* Δ*flgDE,* paired with the presence of E2 in the intracellular fraction (Figure 3.). To ensure the absence of secreted E2 was not due to experimental failure, samples from strains with functional secretion apparatus were run in tandem. Following Western blotting the nitrocellulose membranes were then stripped and reprobed with α-GroEL, to demonstrate that cytoplasmic contaminants (i.e. GroEL) were not present at a high concentration in the supernatant of strains with or without functional secretion apparatus (Figure 3.), confirming that directed FT3SS secretion and not cell lysis or membrane leakage is the route of E2 transport into the supernatant. Although some lysis was GroEL was detected, this antibody is known to be very sensitive; therefore it is unlikely to represent a lot of cytoplasmic derived protein –this is supported by the lack of cellular contaminants seen in the secreted fraction of Coomassie stained SDS-PAGE gels. It would be beneficial to carry out a complete cell lysis to see how this compared to the concentration of GroEL observed in these blots, this is true for GroEL detection throughout this thesis.

3.7.4. Strain improvements: reduced negative regulation of the master regulator

Based on the literature, additional strain improvements (Δ*lrhA,* Δ*dksA,* Δ*clpX*) were made to Δ*CKL* to decrease negative transcriptional control of *flhDC* or degradation of the $FlhD_4C_2$ complex. This should result in an increased number of secretion apparatus and therefore an increase in the amount of protein secreted. With the exception of Δ*clpX* which was derived by phage transduction, all gene knock outs in this chapter were implemented with lambda Red recombinase (Datsenko & Wanner, 2000). P1 phage was successfully eradicated from Δ*CKL* Δ*clpX.* The 'gene-doctoring' variation on this method (where linear DNA is incorporated into a plasmid until cleavage by SceI is induced *in vivo* prior to homologous recombination) was considered and there are now versions of this method which do not result in FRT scars in the chromosome (Kim et al. 2014; Lee et al. 2009), however as the method used was selected as it was well established in the laboratory. All mutant strains have been verified and assay based secretion characterisation was carried out the strains.

3.7.4.1. Effects of deletion of *lrhA* on FT3SS secretion

ΔCKL Δ*lrhA* secretion (Figure 3.) and flagella gene expression (Figure 3.) results showed disparity with one another and were inconsistent with those reported in the literature (Lehnen et al., 2002). This is due the presence of an IS5 element upstream of *flhD* (Figure 3.), therefore reduced transcriptional repression of the master regulator is already active as the IS5 site decouples negative transcriptional regulation sites from the promoter. Consequentially removing master regulator transcription repressors had limited effect. The presence of the IS5 element is not a drawback as is beneficial to improved secretion. Evolution has already resulted in an improved strain by IS5 insertion. With this evidence in mind it seemed more beneficial to focus on other routes to improving flagella gene expression, because while there is still scope for some increased transcription, this will be incremental. This strain can be improved further by other strategies such as reducing metabolic burden which is still valid as the effect of IS5 is not concerned here.

3.7.4.2. Effects of deletion of *dksA* on FT3SS secretion

Although gene expression (Figure 3.) and secretion (Figure 3.) results complement each other in the Δ*CKL* Δ*dksA* mutant, the observations are not concordant with expectations based on the literature. E2 secretion was much lower in Δ*CKL* Δ*dksA* in comparison to the Δ*CKL* strain and showed a negative correlation in the amount secreted over the growth curve (i.e. in the stationary phase). However Lemke et al., 2009 reported that flagella gene expression for all three classes of flagella genes were increased with the removal of Δ*dksA* from strains. Furthermore the literature states that flagella gene activity increased during the stationary phase (from 2 times to 9 times wild type gene expression), therefore an increase in secretion should be expected in the stationary phase, not the decrease observed in this thesis. As DksA inserts into RNA polymerase (RNAP) to cause a structural change, which in turn acts on the *flhDC* promoter. The presence of the IS5 element is not responsible for these differing observations as it concerns genetic elements upstream of the promoter, whereas DksA induced disruptions in RNAP-promoter binding occurs in the -6 to +6 positions (relative to the +1 transcriptional start site) (Rutherford et al. 2009).

This unexpected effect may be due to the nature of the assay. DksA inhibits ribosomal RNA promoter activity and also binds to RNAP, reducing the lifespan of open RNAP. DksA also amplifies the activity of the alarmone ppGpp, which also binds to RNAP. Lemke et al. (2009) found that this effect was much more pronounced when genes which encode the alarmone ppGpp was also deleted from the genome, however it is assumed that ppGpp was intact in this work. Another reason may be linked to the fact that experiments in this thesis were carried out in rich media. While DksA is produced at a relatively constant rate ppGpp is linked to the stringent response and produced in response to low nutrient concentration (Lemke et al. 2009; Magnusson et al. 2007; Paul et al. 2004). ppGpp causes the redirection of a cells resources away from costly biosynthetic processes. Therefore rather than producing more flagella to move to an area where nutrients are more abundant, ppGpp initiates an alternative strategy where the cell represses non-essential functions (such as motility) in a bid to conserve nutrients. Lemke et al. (2009) performed experiments in minimal media where nutrients were low. Therefore the effect of Δ*dksA* in this study may be not as expected because experiments were carried out in rich media, therefore the stringent response would not be in effect and therefore ppGpp would be present at a low abundance. Further searching of the literature has uncovered results which suggest that DksA activity may differ to that reported in the Lemke et al. (2009) paper. One study concluded that Δ*dksA* mutants were immotile and that DksA

positively regulated flagella expression (Magnusson et al., 2007). Reports on the effect of DksA on flagella expression are very conflicting, suggesting that DksA may not be a very suitable target. It should be considered that in this work or any of the other studies that the mutant may have reverted to be *+dksA* and this may be the reason for conflicting results. Finally, the unexpected secretion and gene expression results may have come about due to the fact that, as DksA alters gene expression throughout the genome via RNAP it is likely that there are other phenotypic alterations to the strain beyond flagella gene expression. Many promoters are directly affected by DksA conformational change of RNAP, these include promoters involved in ribosome synthesis, membrane stress response and transport and amino acid biosynthesis and transcription may either be activated or inhibited (Lemke et al., 2009; Rutherford et al., 2009). While the growth of the Δ*CKL* Δ*dksA* strain was not hindered in comparison to Δ*CKL*, the remit for additional strain effects to cause unexpected observations linked to any of the aforementioned DksA regulated genes is likely. Due to the poor secretion performance of the strain and the inconclusiveness of the literature, focus will concern other strategies to increase flagella expression.

3.7.4.3. Effects of deletion of *clpX* on FT3SS secretion

The effect of the removal of *clpX* on the FT3SS was tested in both strains which produced wildtype flagella and in the modified FT3SS secretion apparatus strain (Δ*CKL*). Both the increased concentration of sheared filament protein (Figure 3.) and swimming motility (Figure 3.) in *clpX* knockout mutants in the MC1000 strain background and the increased E2 secretion capacity of Δ*CKL* strains with the *clpX* knockout mutation, suggest that the removal of *clpX* from the chromosome results in increased flagella gene expression. This is likely due to reduction in negative regulation of flagella gene expression, which is usually downregulated due to ClpX directed proteolysis of the $FlhD_4C_2$ complex by ClpP (Tomoyasu et al., 2002; Takaya et al., 2012). This is independent of the presence of the IS5 element, as it concerns the $FlhD_4C_2$ complex. More $FlhD_4C_2$ is present in Δ*clpX* mutants, which demonstrate upregulated flagella gene expression, and increased flagella type III secretion (Kitagawa et al. 2011; Tomoyasu et al. 2003). However as reliable results were not obtained for flagella gene expression assays in strains with the Δ*clpX* mutation, this cannot be confirmed experimentally in this study. In further work it would be desirable to assess flagella gene expression using RT-qPCR. This technique involves first converting cell mRNA to cDNA using reverse transcription, then using

primers to amplify cDNA. The incorporation of fluorophores results in fluorescently labelled DNA and allows the concentration of product to me measured at each cycle. With each cycle DNA and therefore fluorescence increase –as this is proportional to the initial concentration of mRNA in the sample, this can be used gene expression. This would allow quantification of the concentration of flagella gene mRNA in cells -allowing direct comparison of strains for flagella gene expression. Furthermore as primers can be designed to amplify any cDNA, the gene expression of a number of flagella genes could be assessed, providing a more detailed picture of gene expression throughout the hierarchy. As ClpX proteolysis is evident throughout the cell, there is a possibility that flagella gene expression is not increased due reduced $FlhD_4C_2$ degradation, but instead due to an indirect effect –for example ClpX has been previously been shown to degrade a subunit of the RpoS RNA polymerase (Pesavento & Hengge, 2012), therefore it would be interesting to carry out proteomic analysis on *clpX* knockout mutants to investigate changes at a proteomic level also, so see if flagella protein is effected in a different manner to that expected with information on genomics.

While an increase FT3SS secretion was observed in Δ*clpX* mutants, the fact that ClpX acts as a cell wide protease means that the mutation may have other effects on cell phenotype. ClpX targets SsrA- and non-SsrA tagged proteins for degradation, these include RpoS, a sigma factor that induces cell stress resistance (Schweder et al., 1996). RpoS has also been shown to negatively regulate flagella biosynthesis, however if this was the case in this mutant, this effect if presumably compensated elsewhere (Patten et al., 2004). ClpX recognises $FlhD_4C_2$ directly, however degradation is accelerated with the YdiV adapter protein, as it directs ClpXP to proteins to undergo degradation (Takaya et al., 2012) and no SsrA site is present in the protein sequence of FlhD or FlhC. Reduced ClpX mediated proteolysis could also be beneficial i.e. less degradation of other flagella proteins (although there is no experimental evidence for this). SsrA is added post translationally to incompletely translated protein at the ribosome. This occurs in 0.5% proteins, of which >90% are degraded by ClpXP (Lies & Maurizi, 2008). Therefore ClpX is capable of degrading a wide range of protein, although whether these incompletely translated proteins would become released from ribosomes and complete enough to be functional is not certain. A notable reported phenotype of *clpX* mutant is a reduced optical density in the stationary phase and decreased survival of cells if stationary phase is prolonged; this was linked to an increase in growth-phase regulated proteins (Weichart et al., 2003). This was alluded to in the experimental work carried out in this thesis as it was observed that the Δ*clpX* mutant had a detrimental effect of growth in comparison to

strains which did not carry this mutation, however it was not a large difference. An additional phenotypic observation is the elongation of cells with *clpX* knockout mutations, due to the role of ClpX in the regulation of protein associated with cell division (Camberg et al., 2011). While in the experiments for this thesis, cells looked similar in microscope images, analysis of length found that cells with the Δ*clpX* mutation were 1.28 times longer than those without (Figure 3. and Figure 3.), thus confirming this report and highlighting that the removal of *clpX* does result in wide phenotypic effects. As the Camberg et al. (2011) study was carried out on cells with a Δ*minC* background (which is therefore already susceptible to altered phenotype related to cell division), this is the first time this effect has been reported in WT cells. Interestingly ClpXP has been shown to degrade DksA (Flynn et al., 2003), however as results for the effect of the removal of *dksA* on flagella gene expression were inconclusive, it is not possible to comment on this based on this study.

3.7.4.4. Promoter replacement or promoter region modification for increased transcription of the master regulator

As discussed improvements to FT3SS secretion through increased *flhDC* expression are hampered slightly by the presence of IS5, however improvement can still be made. The promoter region could be modified to remove the negative regulators (RscAB) still present and restore IS5 disrupted or decoupled positive regulatory binding sites (i.e. H-NS, QseB, OmpR). In isolation -or in combination this may result in improved *flhDC* expression. A promoter replacement strategy may provide a route of greater change. Replacement of native *flhD* promoter with a strong promoter may result in increased flagella gene expression in excess of that reached through the presence of the IS5 element. In another study an arabinose inducible extra copy of the *flhD* promoter was inserted into the chromosome, this resulted in a 60 fold mRNA increase, however only 5 fold change was seen in class II flagella gene expression (Erhardt & Hughes, 2010), suggesting that there may be limits to how much increased *flhDC* expression can increase gene expression throughout the hierarchy –due to feedback regulation for example. Promoters of varying strengths are characterised and available as BioBricks on The Registry of Standard Parts and would be ideal to carry out this work. A clear choice for this would be to place a strong sigma 70 promoter upstream of *flhDC* to overexpress the $FlhD_4C_2$ complex and in turn upregulate all flagella genes. This may result in more flagella and therefore more secretion. Additionally there are benefits in using inducible promoters so that

there is an element of control of gene and therefore protein expression with regards to strength and timing of induction (Dubendorf & Studier, 1991; Dragosits et al. 2012). It must be noted that increasing promoter strength may have detrimental effects on the cell with regard to homologous protein production and therefore cell growth and viability (Dong et al., 1995). If this occurs, the promoter could be exchanged for a weaker promoter, which would still upregulate gene activity in comparison to wild type levels.

3.7.5. Strain improvements: reduced metabolic burden and negative feedback

Alternative strategies to increase flagella gene expression independent of the master regulator included reducing metabolic burden by removing unnecessary genes. It was hoped this would result in either the redistribution of energy to processes concerning growth, secretion apparatus and plasmid protein production or a reduction in negative feedback on flagella gene expression.

3.7.5.1. Effects of deletion of *motAB* on FT3SS secretion

It was predicted that the removal of *motAB* from cells would result in an increase in protein secretion due to reduced metabolic burden and increased proton motive force. The removal of *motAB* from Δ*CKL* resulted in just over double the concentration of secreted E2 protein, despite the concentration of intracellular E2 being 0.75 fold that seen in Δ*CKL* (Figure 3. and Figure 3.), this demonstrates that the removal of *motAB* from cells increased the efficiency of E2 secretion through the secretion apparatus. This is possibly due to the reduction in metabolic burden due to the absence of MotAB allowing more metabolic expenditure on other cellular activities, including cell growth, and expression of E2 and the modified FT3SS apparatus. It is also anticipated that as the MotAB complex acts as a proton conductive pathway, the absence of MotAB will result in reduced drain on proton motive force. Cells have a more robust growth phenotype, which may aid protein production as cells will be more productive, furthermore as PMF is required for secretion of substrates, this may facilitate increased protein secretion (Minamino et al. 2016; Zhou et al. 1998; Blair & Berg, 1990; Hosking et al. 2006; Minamino & Namba, 2008). As supernatant fractions were free of cellular

contaminants (Figure 3. and Figure 3.), it was not evident that the absence of MotAB caused membrane breakage or permeability, although it would be beneficial to confirm this with a more sensitive method, such as Western blot analysis of supernatant fractions with an α-GroEL antibody. It would also be preferential to have growth curve data for this strain, to observe whether growth rate is increased when *motAB* is removed from the genome. These factors will be investigated in Chapter Four.

3.7.5.2. Effects of deletion of *flgMN* on FT3SS secretion

The main focus of this strategy was the removal of the anti-sigma 28 factor (FliA) FlgM. Prior to hook completion FlgM binds FliA and inhibits both flagella class III gene and *flhDC* transcription. While FlgM inhibits FliA, it also protects it from proteolysis and the literature reports either increased or decreased *flhDC* expression following the removal of FlgM (Karlinsey et al. 2000; Barembruch & Hengge, 2007). The removal of FlgN was also implemented as it is a chaperone to FlgK and FlgL and a transcriptional regulator of FlgM (Aldridge et al., 2003), therefore unnecessary in the Δ*CKL* Δ*flgM* mutant. While only one biological repeat was carried out and it was not possible to measure the concentration of intracellular E2 in the mutant, it would appear that it did not result in an increase in secretion, as less E2 (0.94 fold) was seen in the Δ*CKL* Δ*flgMN* strain compared to Δ*CKL*(Figure 3.). This is likely due to the role of FlgM in protecting FliA from degradation –in the absence of FlgM, the concentration of FliA should be reduced and therefore *flhDC* gene expression was reduced. Investigation into the gene activity of this strain with the *lacZ* fusion assay used in this thesis, would give insight into whether this was true here, however since the aim of this project is to investigate strains which result in increased secretion, this was not a priority

3.7.5.3. Effects of deletion of *fliDST* on FT3SS secretion

Many strategies interplayed to support the removal of *fliDST* from Δ*CKL*. At the most basic this involved the removal of chaperone and cap proteins to deleted proteins to reduce metabolic burden. Other intricacies included the removal of *flhDC* inhibitor FliT and FlgM binding FliS with the inclination that this would cause class II and III gene expression will be increased respectively, which may result in the presence of more secretion apparatus (Aldridge et al. 2010; Yamamoto & Kutsukake 2006; Yokoseki et al. 1996).

The failure of the *fliDST* knockout mutation to result in improved secretion of E2 protein centres on the inability of the cell to accumulate intracellular E2 (Figure 3.). It is beyond the scope of the experiments conducted to ascertain whether E2 gene expression was reduced in the Δ*fliDST* strain, therefore the focus of my discussion on this will be at a protein level. Knowledge from the mechanism of flagella protein expression in the literature, would suggest that the absence of the flagellin chaperone protein FliS is the route of this phenomenon. As E2 is a variant of the native FliC subunit, and the chaperone binding region is conserved, it is likely that it benefits from the chaperone effects of FliS. While the removal of *fliD* or *fliT* does not inhibit the production of viable flagella in certain environments, the removal of *fliS* was shown to greatly reduce the length to flagella (Yokoseki et al., 1995). FliS prevents premature polymerisation of FliC in the cytoplasm (Auvray et al., 2001), this should also be true of E2. If E2 was present in the intracellular fraction and not the secreted (as found with FliC in the Yokoseki et al. (1995) study), this could be because of secretion inhibition of E2 due to polymerisation in the absence of FliS. However as it is also displays reduced concentration intracellularly it is likely that E2 is degraded more readily in the FliS mutant. However while the disordered C-terminal region of FliC (present in E2 also) is degraded in the absence of FliS, this is not true of the whole FliC protein and resulted in a 5kDa smaller product (Ozin et al., 2003). This discrepancy in size was not observed with E2 protein. In addition to this the negative secretion strains Δ*flhDC* and Δ*flgDE* should also be FliS deficient, as class III gene transcription will not occur in either. Slightly less E2 was seen in the cells of Δ*flgDE,* however the concentration in Δ*flhDC* was similar (Figure 3.). A low rate of secretion can also be attributed to the absence of FliS, as FliC secretion is initiated by FliS. In a FliS-FliC complex, FliS binds to docking platform FlhA, anchoring FliC to the export gate, where FliC is unfolded for efficient export (Furukawa et al., 2016). It was recently found that a major rate limiting factor in reduced FliC secretion was lack of FliS, resulting in inadequate unfolding of FliC (Furukawa et al., 2016). In summary, as with FliC, both E2 intracellular accumulation and export may be hindered in the FliS mutant. In later chapters the chaperone binding site is absent in secretion constructs, therefore chaperone-decoupled expression and secretion in this strain can be investigated.

3.7.6. Limitations of the E2 protein secretion assay

3.7.6.1. Variation in results

Inter-experimental variation of yields of both intracellular and secreted E2 was observed; while variation is common in biological systems (Endy, 2005), discussion will focus on experimental sources of variation. In the growth curve E2 protein secretion assays variation was common, for ease only the highest recorded secreted E2 yield is discussed. Figure 3. (below) shows that there was large variation between experiments –as the results for Δ*CKL* yield are biological replicates. While great care was taken to ensure that experimental procedure was consistent, the yields vary dramatically. Steps taken to ensure consistency across experiments included, thoroughly rinsing glassware of detergent or residual salts before use and using the same brand of LB media and agar throughout; these steps were of particular importance because it is reported that increased concentrations of NaCl results in elongated hooks and reduced flagellin secretion –this was in a *flgE* pseudo-revertant background, however still applicable (Saito et al., 1998). Frozen aliquots of IPTG and antibiotics were used to ensure concentration remained consistent. And as temperature alters flagella assembly (Gerber et al., 1973), cells were incubated in the same incubator every time, as temperature (and agitation) was found to vary between incubators, despite the settings reporting the same setup. pH has also been shown to result in differential flagella gene expression (Maurer et al., 2005), therefore in the future less variable results may be obtained following buffering of LB broth.

A source of error relates to the lack of linearity exhibited by Western blot analysis of protein concentration. This refers to the disparity in concentration of protein and chemiluminescent output. It is known that results derived by Western blot show limited linearity, this leads to inaccuracies in the estimation of concentration (Taylor et al., 2013). This could be investigated further by measuring chemiluminescent output related to protein concentration, by producing a standard curve of dilutions of E2 and relating to densitometry results. Alternatively Western blot technology which relies on infrared signal, rather than chemiluminescence could be utilised instead, however it was not possible to access one for this work. This would allow confidence that the protein levels detected are within the linear dynamic range of the Western blot. Even if this was implemented the nature of Western blots lends itself to inaccuracy due to the number of steps associated with the experiments, there are multiple opportunities for variation to occur. Use of Coomassie stain eliminates inaccuracy which may occur due to lack of linearity of Western blots and also has less experimental steps, therefore less sources of

potential error; however in low concentrations of secreted E2 it is unsuitable as it is not detectable. Additionally, Western blots provide information on the concentration of intracellular E2, this is not possible with Coomassie stained SDS-PAGE gels. It should also be noted that use of densitometry based analysis serves as a source of error, while care is taken to remove background saturation and define the boundaries of protein bands, this is subjective. These facts do not make the method void, but it is important to be aware that there are sources of error associated.

Another source of inconstistency is likely to have arisen from the E2 protein standard. Protein was not initially stored in aliquots therefore was subject to successive freeze thawing, which can degrade protein (Cao et al., 2003). This would affect the quantitative yield values derived, meaning that absolute protein secretion values cannot accurately be derived; however quantitative comparisons between experiments are reliable. While the concentration of the protein standard may be variable between experiments, intra-experimental results are proportional; therefore this will not affect the relationships observed. Following growth curve secretion experiments, protein standards were aliquoted to eliminate any freeze-thaw effects. While preventing E2 standard degradation would reduce variation in quantification of yield, the use of a protein standard itself is another source of inaccuracy as it contributes another variable in measurement of E2 concentration. This is evident in Figure 3.; where the pattern of results for strains in comparison to Δ*CKL* vary depending on whether secreted E2 yield was measured in mg L^{-1} or as a relative concentration (Figure 3.), For this reason Δ*CKL* is always grown, prepared and analysed in tandem so that relative concentrations can also be derived, so that the concentration of E2 can be assessed based on two points of reference and very variable results (i.e. very low quantitative yields of Δ*motAB* and Δ*flgMN* in Figure 3.A) can be excluded.

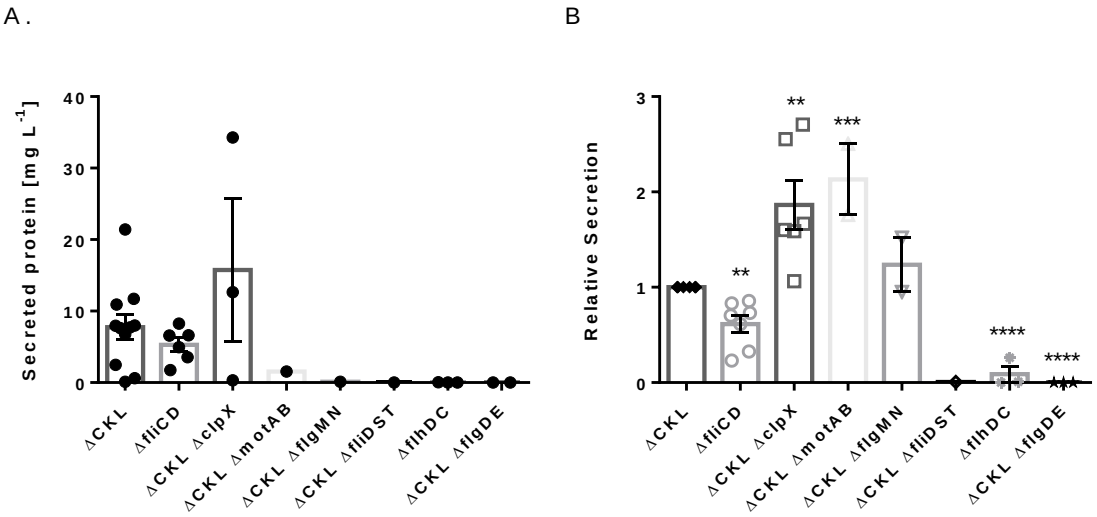


Figure 3.54: Combined results from E2 secretion assays in this chapter.

Both absolute quantifications (A and B) and relative yields in comparison to the one achieved for Δ*CKL* in that experiment (C and D) are given for each strain. Results are comprised of secreted (A and C) and intracellular (B and D). Note: only results for cultures grown in 10mL LB broth, supplemented with the 0.05mM IPTG are included. Standard error of the mean is shown. Unpaired t-test of Δ*CKL* to each strain: **, *** and **** denote p = <0.01, 0.001 and 0.0001 respectively.

3.7.6.2. E2 as a substrate

The use of E2 as a substrate allowed quantification of intracellular and secreted protein to be measured in a number of strains, however as a variant of the native FliC protein it is subject to biological factors associated with the cell environment (as seen in with the removal of FliS in Figure 3.). Production of a non-native protein would be desirable as it may reduce the effect of native biological control on protein expression, therefore reducing variables from strain to strain. Furthermore, the ability of the modified FT3SS to secrete recombinant proteins, would be beneficial both in terms of showcasing the capability of the modified FT3SS structure and also in opening up the strain for further applications concerning the production and secretion of industrially relevant proteins.

3.7.7. Outcomes of strain improvements to the FT3SS platform secretion strain

Despite the limitations described the addition of knockout mutations to the Δ*CKL* strain did lead to improvements in E2 concentration. Overall the implementation of Δ*clpX* resulted in great improvements to the concentration of E2 both expressed and secreted whether measured by quantification or relative E2 (Figure 3.). The addition of Δ*motAB* resulted in the highest average yield of secreted E2 relative to that achieved for Δ*CKL* grown in the same technical replicate. This shows that it is possible to implement additional gene knockouts in the secretor strain to further improve the directed secretion of E2 through the secretion apparatus. Based on the difference in E2 concentrations in intracellular and secreted protein fractions, it is clear that it is not the production, but the secretion capacity of E2 which limits the yield of secreted E2 protein. Through the addition of knockout mutations (Δ*motAB*), it is possible to improve the secretion capacity of the secretion apparatus (more secreted in comparison to intracellular). This may be indicative of improved capacity of the secretion apparatus itself or that show that tuning of E2 expression to secretion capacity is important so as not to overload cells with high concentrations of intracellular E2. This will be investigated in the next two chapters.

In 10mL cultures (with a non-freeze thawed E2 protein standard) the highest yield of secreted E2 protein achieved was just under 35 mg L^{-1} following 6 hours of culture of Δ*CKL* Δ*clpX* cells. This is an improvement to the concentrations reported in studies which have previously modified the FT3SS for protein secretion (12 mg L^{-1} following at least 16 hours growth in Majander et al. 2005 and 1.8 mg L^{-1} $Hour^{-1}$ in Widmaier et al. 2009, which would equate to 10.8 mg L^{-1} in this system, assuming cultures are grown for six hours). This shows great promise of this protein secretion platform in relation to secretion capacities previously reported.

In this chapter a modified FT3SS secretion platform for the secretion of protein into the extracellular media was established. A protein secretion assay was developed and it was found that the addition of knockout mutations to the Δ*CKL* strain resulted in variation in the secretion capacity of the secretion apparatus. While some of these mutations appear to increase the secretion capacity of the modified FT3SS, further improvement is required for this secretion system to be competitive with industry and competitors in different research

groups–for example the 60mg L^{-1} yield of periplasmic secreted GFP in a small batch fermentation (Matos et al., 2012). The work in this chapter laid a good foundation for both measuring and improving secretion in strains. However as it was found that the use of E2 to measure secretion is not compatible with all strains, the next chapter will investigate the development of a secretion assay for a non-native industrially relevant protein -which is of course main aim of the thesis. This should allow quantification of expression and secretion in all strains and also show scope for secretion of a range of proteins through the secretion apparatus.

Chapter 4. Further development and testing of the engineered FT3SS system: non-native protein secretion and assay development

Following the successful development of a platform strain for the secretion of native protein through a modified flagella type III secretion system (FT3SS), it was desirable to make further improvements to extend the secretion capacity of the platform technology to non-native recombinant proteins. This was for a multitude of reasons; firstly, it was found that production and secretion of the native E2 protein was not compatible with all strains and therefore constituted a limitation of the suitability of the protein secretion assay which was developed to assess secretion capacity of strains. Second, it will showcase that the secretion platform design permits secretion of a range of proteins. Finally, secretion of recombinant proteins through the FT3SS is central to the suitability of this platform technology for bio-manufacturing.

4.1. Design of a synthetic modular secretion construct for FT3SS secretion

Previously, secretion of a near-native flagellin termed E2 was achieved by expressing E2 protein fused to a 47 amino acid FliC secretion signal peptide. The signal peptide could be fused to a recombinant protein and the same principle could be applied, however while this strategy was successful, there are some limitations. Protein detection by Western blot with H48 (flagellin) antibody was protein specific; therefore it would be necessary to use a novel antibody for each recombinant protein. This would be costly, time consuming (as each antibody would require optimisation) and limit the range of recombinant proteins to those with corresponding antibodies. Another limitation concerns downstream processing of protein following secretion. Currently secreted E2 is not readily isolated from the media; furthermore if E2 protein was isolated it is not able to be cleaved from the FT3SS signal peptide to yield a purified product and is prone to polymerisation due to it being similar to native FliC, flagellin. Therefore to facilitate directed secretion of recombinant proteins, followed by purification

then isolation, a synthetic modular secretion construct was generated. This was achieved previously in the laboratory by a process of *in silico* design and synthesis (by DNA 2.0) and was the basis for studies here.

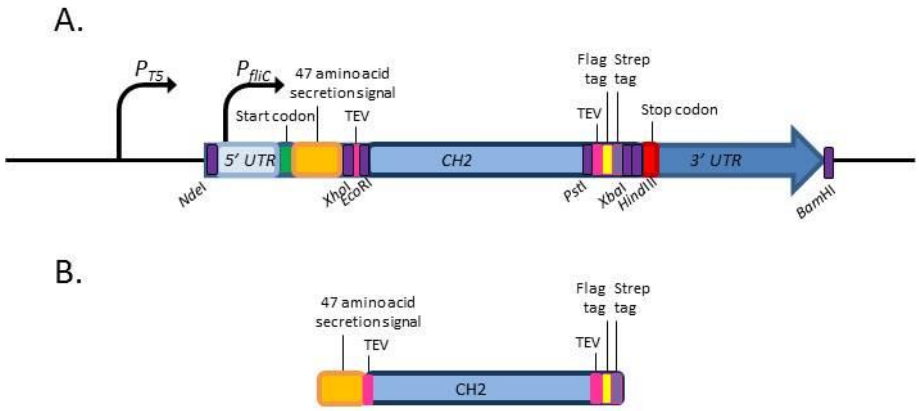


Figure 4.1: Schematic of (A) the prototype synthetic modular secretion construct inserted into an IPTG inducible plasmid and (B) the resulting peptide.

(A) The genomic prototype modular secretion construct is comprised of the transcribed elements of the secretion construct along with the *fliC* 5' and 3' untranslated regions (UTRs) and the start and stop codons. Restriction enzyme sites are incorporated throughout to allow modification of the secretion construct in a modular fashion. The prototype secretion construct is under transcriptional control of the IPTG inducible T5 promoter and the sigma factor-28 (FliA) inducible *fliC* promoter which is harboured in the 5' UTR. The secretion construct shown harbours CH2 cargo (B) The resulting peptide comprises of the FliC 47 amino acid secretion signal, tobacco etch virus protease (TEV) sites, a FLAG tag, a streptavidin II tag and the protein for secretion –here shown with CH2 cargo. Note: it is also possible to express an 'empty' secretion construct, with no protein cargo between the *Eco*RI and *Pst*I sites.

The genetic prototype secretion construct consists of the 5' UTR region of FliC, the 47 amino acid N-terminal secretion signal peptide of FliC, restriction sites to incorporate interchangeable protein for secretion, TEV protease recognition sites, a FLAG antigen epitope tag (FLAG tag) and Streptavidin-II purification tag (Strep tag) and the 3' UTR of the *fliC* gene (Figure 4.). A TEV protease site was selected as the incorporation of just seven amino acid residues, allows highly site specific enzyme cleavage, furthermore its use is very well characterised and although

other protease are available which are much more efficient at cleaving protein, TEV protease activity is adequate (Frey & Görlich, 2014; Waugh, 2011). Both the Strep II and FLAG tags are also small and readily expressed in *E. coli* and are compatible a large range of laboratory based systems for purification and antibody detection, additionally although in principle they may be a bit costly, they are acceptable in industry (Saraswat et al., 2013). Strep II does not interfere with folding or activity of fusion proteins or induce protein aggregation and facilitates one-step purification of high yields of protein. The hydrophilic nature of FLAG means that it is presented on the outside of protein, so easily assessable to antibodies (Zhao et al., 2013). The construct has restriction enzyme sites to insert the whole construct into a vector (the IPTG inducible pJexpress in this study) and to interchange the protein for secretion via central *PstI* and *EcoR1* sites. The FliC UTR regions were incorporated into the secretion construct as they have been shown to be critical in the implementation of FT3SS secretion of native or non-native proteins, possibly due to the presence of a secretion signal (Majander et al., 2005) –this will be examined more thoroughly in the next chapter. Furthermore the 5' UTR harbours the *fliC* promoter; therefore it was incorporated to allow natural induction of the secretion construct by the cell –as it seemed sensible to make use of native flagella gene regulation as well as being able to induce higher expression using the IPTG-responsive *T5* promoter on this plasmid backbone (pJexpress404). Translation results in a multifunctional protein suitable for secretion by the FT3SS. The N-terminal 47 amino acids of FliC are thought to act as a non-cleavable signal peptide which confer targeted FT3SS export of the construct (Dobó et al., 2010). The FLAG tag provides a means of detection of secreted protein by FLAG specific antibody. The Strep tag allows purification of the secretion construct following secretion into the media by affinity chromatography. Finally the tobacco etch virus (TEV) protease recognition site allows the cleavage of the purification tags and signal peptide following purification, so that the final product is comprised of the protein of interest only.

4.2. Antibody fragment within the synthetic modular secretion construct

Once the synthetic modular secretion construct (secretion construct) was developed a number of proteins (including enzymes and antibody fragments) were inserted by DNA ligation between the *Pst*I and *Eco*RI sites. This included the antibody fragment 'C_H2', which was generously donated by Dr Jagroop Pandhal, University of Sheffield. Following preliminary protein secretion tests, the antibody fragment C_H2 (CH2), which had been codon optimised for *E. coli,* was found to express and secrete reliably and for this reason forms the basis for the investigation of the secretion of non-native protein in through the modified FT3SS. CH2 refers to the second (of three) constant heavy (CH) domain of an immunoglobulin antibody. CH2 is of particular interest because unlike other antibody domains it is able to form a stable monomer. CH2 is amenable to modification for target binding and also for use as a scaffold, examples of which include HIV binding therapeutic molecules therefore has many uses in biotechnology (Gehlsen et al., 2012; Ying et al., 2014).

4.3. Investigation of the performance of the synthetic modular secretion construct

Prior to the investigation of the secretion capacity of CH2 through the modified FT3SS of mutants, initial research focused on the robustness of the secretion construct. This involved testing the suitability of the FLAG tag for antibody based detection of protein, the performance of the Step tag during protein purification and the effectiveness of the TEV sites in isolating protein product from the secretion construct.

4.3.1. Protein purification of the CH2 harbouring secretion construct

To confirm functionality of the Strep tag located in the secretion construct, CH2 was overexpressed in BL21 (DE3). Cells were harvested and lysed by French press, before undergoing Strep tag purification on a Strep-trap column. The insoluble and insoluble fractions

of cell lysate, along with samples collected from the flow through, wash and elution steps of the Strep-trap column protocol underwent SDS-PAGE and Coomassie staining.

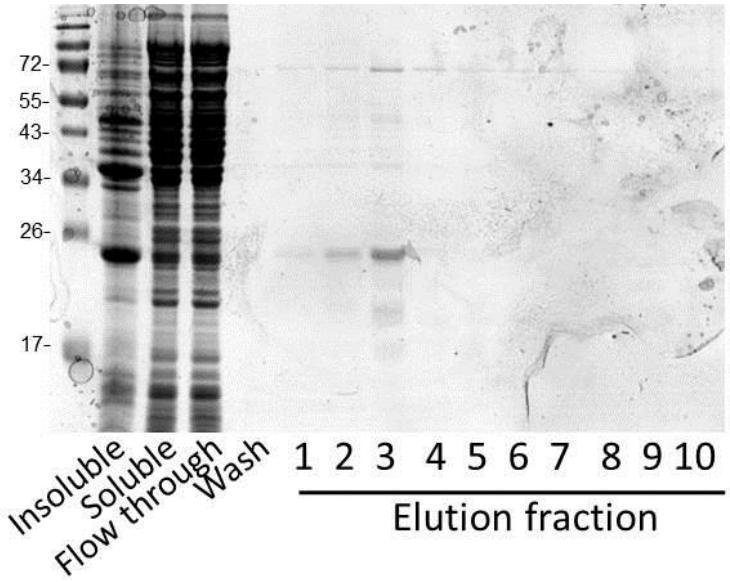


Figure 4.2: Protein fractions following overexpression and purification of Strep tagged CH2 protein.

E.coli BL21 (DE3) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep was grown in LB to OD_{600} 0.6-0.8 before induction with 1mM IPTG. Cells were harvested 3 hours after induction and resuspended in 100mM Tris-HCl, 150mM NaCl, 1mM EDTA pH 8, prior to cell lysis by French pressure cell. The resulting lysate was centrifuged for 10 minutes, 13,000g, 4°C to remove cell debris. The resulting supernatant was then centrifuged for a further 30 minutes and to yield insoluble (pellet) and soluble (supernatant) cell lysate protein fractions. Soluble Strep tagged CH2 protein was purified using a Strep-trap column. Protein was applied to the column; residual protein was washed off with 10 column volumes of 100mM Tris-HCl, 150mM NaCl, 1mM EDTA, pH 8 buffer and eluted with 2mM D-biotin. The flow through, wash and elution fractions were collected. 2 x SDS loading buffer was added to protein samples at a 1:1 ratio. Samples were run on 15% SDS-PAGE gels. Protein samples were loaded as follows. Lane: (1) EZ-Run™ Prestained Rec Protein Ladder, (2) insoluble cell lysate, (2) soluble cell lysate, (3) Strep column flow through, (4) elution following wash step, (5-15) elution fractions 1-10. 5µL sample was loaded in lanes 2-4 and 15µL in lanes 5-15. SDS-PAGE gels were then stained with InstantBlue™ Coomassie stain.

The expected size of the CH2 protein is 12.6kDa however within the secretion construct it is predicted to be 22.4 kDa. An overexpressed protein can be seen clearly in the insoluble protein and in elution fractions 1-3. As the protein ran just below the 26 kDa marker, it can be assumed that this is the CH2 protein. Despite around 50% of CH2 protein being insoluble in this strain when expressed at these high levels, adequate soluble protein allowed protein purification on the Strep-trap column. The wash fraction is free of CH2 protein, suggesting that CH2 remains bound once associated with the Strep-trap column. Elution of CH2 was achieved within the application of the first 1mL elution buffer. The highest concentration of CH2 was eluted in the third elution fraction; following this no CH2 was detected in elution fractions. There are some additional protein contaminants in the elution fractions which contain CH2, which are seen at 72kDa, around 55 and 36kDa and in the third elution fraction only, around 20kDa. The concentration of these contaminants seems to increase, with increased concentration of CH2. Despite efforts to remove these contaminants it was not possible. This included altering the Strep column protocol, by using different buffers for binding, different concentrations of D-biotin in the elution buffer and also by applying additional downstream steps. These additional steps included dialysis, with the best results achieved following overnight dialysis (40000 MWCO) in 100mM Tris-HCl, pH 8 (Figure 4.); the 72 and 36kDa protein contaminants are no longer visible and the 20kDa protein contaminant can be seen faintly in elution fractions 2 and 3. The major remaining protein contaminant is seen around 55kDa and is present at a similar concentration to the CH2 protein, which runs at 26kDa here. It is possible that this could be a dimeric version of the 26kDa protein, although dimers should be absent following preparation with DTT and SDS it is around double the size. There was less total protein in elution fraction 1 prior to dialysis; following dialysis very little protein is visible. Elution fractions 2 and 3 both contain adequate protein for detection by Coomassie stain. There is a higher concentration of both CH2 and the 55kDa protein contaminant in elution fraction 3.

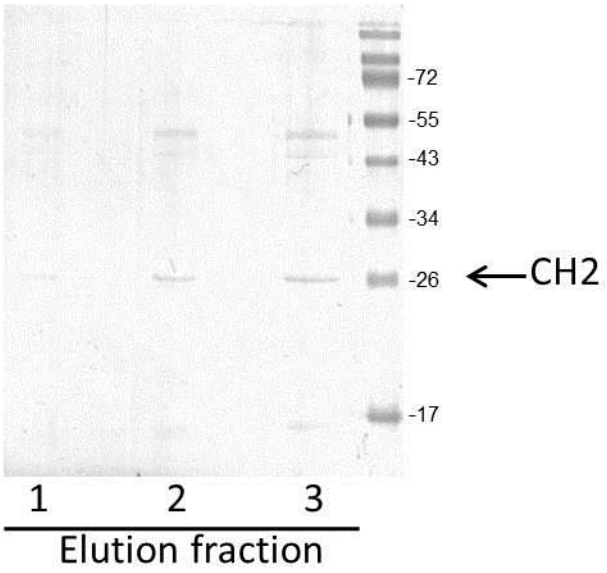


Figure 4.3: CH2 containing protein fractions following purification and dialysis.

Soluble Strep tagged CH2 protein was purified using a Strep-trap column. Fractions which contained CH2 protein were dialysed overnight in 100mM Tris, pH 8. 2 x SDS loading buffer was added to protein samples at a 1:1 ratio. Samples were run on 15% SDS-PAGE gels, with empty lanes between each. Protein samples were loaded as follows. Lane: (1) elution fraction 1, (2) empty lane, (3) elution fraction 2, (4) empty lane, (5) elution fraction 3, (6) EZ-Run™ Prestained Rec Protein Ladder. 30µL sample was loaded. SDS-PAGE gels were then stained with InstantBlue™ Coomassie stain. CH2 protein indicated by arrow.

4.3.2. Antibody detection of CH2 protein in the secretion construct

While dialysis aided the removal of some of the protein contaminants in purified CH2, it was not possible to remove them all (Figure 4.). Protein fractions were analysed by Western blot to investigate whether the FLAG-tag was functional in CH2 and to see if protein contaminants harboured a FLAG-tag (Figure 4.).

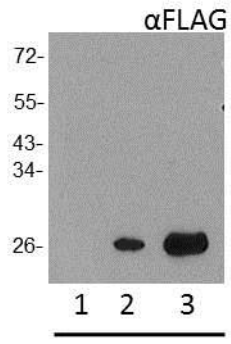


Figure 4.4: Detection of the FLAG-tag in CH2 containing protein fractions following purification and dialysis.

Soluble Strep tagged CH2 protein was purified using a Strep-trap column. Fractions which contained CH2 protein were dialysed overnight in 100mM Tris, pH 8. 2 x SDS loading buffer was added to protein samples at a 1:1 ratio. Samples were run on 15% SDS-PAGE gels. Protein samples were loaded as follows. Lane: (1) elution fraction 1, (2) elution fraction 2, (3) elution fraction 3. 5µL sample was loaded. SDS-PAGE gels then underwent Western blotting with an anti-FLAG antibody.

The FLAG-tag was detected at around the 26kDa marker as before, supporting the observation from the Coomassie stain. This was the only FLAG-tag signal detected, demonstrating that the contaminant proteins, including the 55kDa protein did not harbour FLAG-tags. In keeping with observations from the Coomassie stained SDS-PAGE gel, there was an increasing concentration of CH2 from elution fraction 1 to 3. In elution fraction 1 is can be assumed that the concentration of FLAG-tagged protein was below that of the detection threshold of the anti-FLAG antibody. As CH2 was visible in elution fraction 1 on the Coomassie stained SDS-PAGE, this would suggest poor transfer efficiency of protein on to the nitrocellulose membrane.

4.3.3. CH2 protein isolation from the secretion construct

TEV sites were incorporated into the secretion construct to facilitate cleavage of accessory peptides in final produced proteins (i.e. the FT3SS signal peptide, FLAG and Strep tags). The effectiveness of this strategy was assessed by incubating purified CH2 protein from elution

fractions with TEV enzyme. Following a repeat of the purification protocol (protein eluted in fractions 2 and 3 again, as seen in Figure 4.), protein was then incubated in buffer with AcTEV™ Protease and DTT for 2 hours at room temperature. As a negative control, tandem reactions were set up without TEV protease. The resulting mixtures were prepared for SDS-PAGE followed by Coomassie stain or Western blot to ascertain whether cleavage at the TEV sites had occurred in the presence of TEV protease (Figure 4.). In the elution fractions following dialysis there is more total protein in elution fraction 3, both in terms of CH2 and also additional protein contaminants. The TEV protease is visible on the Coomassie stained SDS-PAGE gel and runs at 27kDa. When fused to the secretion construct CH2 runs just below the 26kDa marker but following successful TEV cleavage it should run at 13kDa. TEV protease is visible in lanes 9, 11 and 13. CH2 protein fused to the secretion construct is visible following incubation without TEV protease (lanes 10 and 12), but not in fractions where TEV protease was present. A protein with a small molecular weight (around 13kDa) is seen in elution fraction 2, following TEV cleavage. This same protein is seen in all fractions which contain elution fraction 3 (whether TEV treated or not), suggesting some non TEV directed cleavage of CH2 protein. This could also be a contaminant protein, however as a higher concentration of this band is present following incubation with TEV as opposed to without (Figure 4.: last four lanes in Coomassie), this is more likely to be the TEV cleavable CH2 protein. Furthermore if protein dilution during incubation with (or without) TEV protease is considered, a higher concentration of the small band is seen following TEV cleavage in comparison to that seen in the elution fraction 3 only. In the Western blot when a high concentration of protein was loaded (pre-TEV treated samples) a HRP fluorescent signal was detected for multiple bands, suggesting that the anti-FLAG antibody bound non-specifically to protein. In samples with CH2 protein and no TEV protease, an antibody signal is detected. This is not as strong as the signal seen in elution fraction prior to incubation with TEV protease, due to dilution of protein and perhaps protein degradation during incubation. No FLAG-tag is detected at any other molecular weights following incubation with or without TEV protease. When CH2 protein is incubated with TEV protease, no anti-FLAG-HRP signal is detected, this shows that FLAG-tag fused CH2 is no longer present in the protein fraction, demonstrating successful TEV protease cleavage of CH2 from the FLAG tag. This combined with the appearance of a higher concentration of protein which is

around 13kDa, suggests TEV cleavage is successful at both TEV sites in the secretion construct.

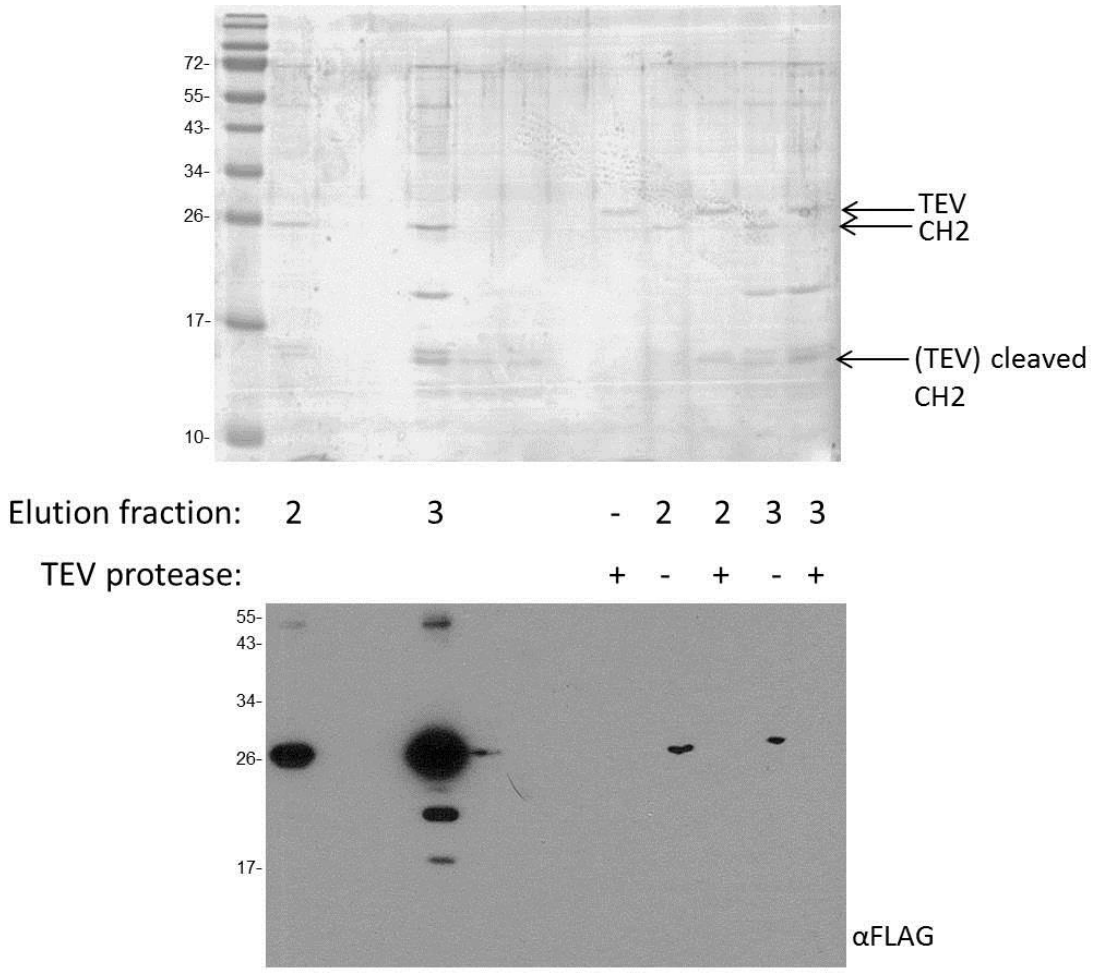


Figure 4.5: Protein fractions following overexpression and purification and TEV cleavage of CH2.

CH2 protein was overexpressed in BL21 (DE3) and soluble protein was purified using a Strep-trap column. Fractions which contained CH2 protein were dialysed overnight in 100mM Tris-HCl, 150mM NaCl, 1mM EDTA, pH 8 and then incubated for two hours either with or without TEV protease. 2x SDS loading buffer was added to protein samples at a 1:1 ratio. Samples were run on 15% SDS-PAGE gels. Protein samples were loaded as follows. Lane: (1) EZ-Run™ Prestained Rec Protein Ladder, (2) elution fraction 2 following dialysis, (3 and 4) empty, (5) elution fraction 3 following dialysis, (6-8) empty, (9) TEV protease only, (10) elution fraction 2 only, (11) elution fraction 2 and TEV protease, (11) elution fraction 3 only, (12) elution fraction 3 and TEV protease. 20µL sample was loaded. SDS-PAGE gels were then stained with InstantBlue™ Coomassie stain (top) or underwent Western blotting with anti-FLAG antibody (bottom).

4.3.4. CH2 protein structural properties

CH2 can also dimerise, due to the weak carbohydrate-mediated interchain protein–protein interactions that occur between two CH2 domains at residue 297 in addition one intra- and one inter- chain disulphide bonds form (Chintalacharuvu et al., 2002) -the interchain bond forms between the hinge regions associated with CH2 (Prabakaran et al., 2008). As *E. coli* are not capable of glycosylation or disulphide bond formation (without the aid of strains such as CyDisCo, which produces catalysts for disulphide bond formation and isomerisation or expression of the N-linked glycosylation system in *Campylobacter jejuni* (Wacker et al., 2002; Matos et al., 2014)), it was not known if intracellular expression of CH2 would result in disulphide bond formation. Formation of disulphide bridge requires a non-reducing environment; however as CH2 protein is routinely prepared for SDS-PAGE in 2x SDS loading buffer with the reducing agent DTT, only the monomer has been detected. To investigate whether recombinant CH2 harboured in the secretion construct was capable of forming dimeric structures, purified CH2 was prepared for SDS-PAGE in 2x SDS loading buffer with and without the DTT component and run on SDS-PAGE gels. On the Coomassie stained SDS-PAGE gel it is evident that this round of protein purification resulted in an elution fraction with less contaminant protein. -with DTT a band is seen at 26kDa, however the band at 55kDa is not present. Without DTT the 26kDa band is not present, instead a band is seen at around 15 and 20kDa (both too small to be the predicted 22.6kDa CH2-secretion construct protein). When probed with anti-FLAG antibody on a Western blot (Figure 4.), the 26kDa band is detected and a very faint signal just below 55kDa is detected with DTT. In the absence of DTT the band which is present at around 26kDa is not present, whereas a band is visible at around 50kDa, which could suggest a dimeric form of the 26kDa CH2 protein.

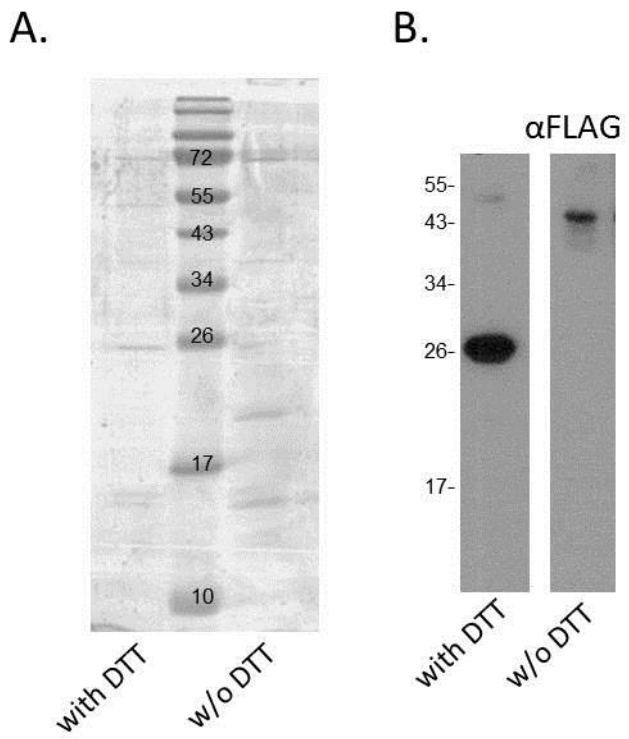


Figure 4.6: CH2 protein following overexpression and purification prepared for SDS-PAGE followed by Coomassie stain or Western blotting in both the presence and absence of DTT.

Soluble Strep tagged CH2 protein was overexpressed in BL21 (DE3) and purified using a Strep column. Fractions which contained CH2 protein were dialysed overnight in 100mM Tris-HCl, pH 8. 2x SDS loading buffer either with or without DTT was added to protein samples at a 1:1 ratio. Samples were run on 15% SDS-PAGE gels and underwent (A) Coomassie staining or (B) Western blotting with anti-FLAG antibody. The resulting Western blot was cropped to remove irrelevant lanes. Lanes shown are (left) CH2 with DTT, (right) CH2 without DTT. 20µL sample was loaded.

Experimental work has shown that CH2 protein expressed within the secretion construct is amenable to overexpression, Strep-tag mediated purification, antibody detection by FLAG-tag and TEV protease cleavage of accessory peptides. In addition to these secretion construct facilitated characteristics, the CH2 protein was also shown to be functional in forming dimeric structures, presumably by disulphide formation. While successful, these findings all concern intracellular derived protein, focus will now move to secreting this protein through the FT3SS

into the extracellular media but they at least show the feasibility of the use of this construct and some element of native function in terms of dimerisation.

4.4. Investigation of secretion of the synthetic modular secretion construct through the FT3SS.

4.4.1. Secretion of the CH2 harbouring synthetic modular secretion construct through the modified FT3SS

In the previous section, CH2 was overexpressed in BL21 (DE3). It was confirmed that the secretion construct plasmid expressed protein upon induction and also established that purification (Strep II) and antigen tags (FLAG) were faithfully presented. TEV cleavage capability was also verified. However, investigation of the potential of the modified FT3SS to secrete a heterologous protein within this prototype secretion construct required expression of the pJexpress-FliC-CH2-FLAG-Strep plasmid in FT3SS secretion strains. Initial investigation of CH2 expression and secretion was carried out in the original *ΔCKL* secretor strain as a starting point for studies. Following cell culture of the secretor strain, with induction of the plasmid harbouring the FliC secretion signal peptide-linked CH2 in the secretion construct; FLAG-tagged protein was visible by Western blot in both the intracellular and supernatant (secreted) fractions (Figure 4.). Less CH2 protein was seen in the secreted fraction and no FLAG signal was detected in the cells or supernatant of cells expressing the empty vector.

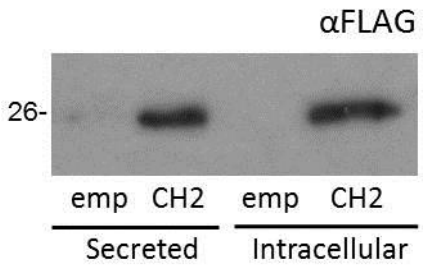


Figure 4.7: CH2 protein in the secreted and intracellular fractions of *ΔCKL* secretor strain cell culture.

E.coli ΔCKL containing either the plasmid pJexpress-FliC-CH2-FLAG-Strep or pJexpress-empty vector were induced with 0.05mM IPTG and grown in LB at 37°C with 180rpm agitation. Cells were harvested when cell culture reached OD_{600} 1.5. 1 OD unit of cells were prepared for SDS-PAGE and the supernatant (SN) from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western blot analysis with anti-FLAG antibody. Protein samples were loaded as follows. Lane: (1) CH2 –SN, (2) empty vector –SN, (3) CH2 –cells, (4) empty vector –cells.

4.4.2. Optimisation of a standardised protein secretion assay to quantify secretion of CH2 through the modified FT3SS: induction

The data presented above showed that the secretor strain was effective in both expressing and secreting CH2 protein, however the expression protocol carried out was based on that derived from the optimisation of secretion of the E2 monomer following expression of pTrc E2. As CH2 secretion involves the expression of an entirely different protein and more importantly from a different plasmid backbone (pJexpress), a range of expression protocols was tested to ensure that CH2 expression and secretion are robust before the effect of strain improvements is assessed. As described, there are two means of expression of genes within pJexpress-FliC-CH2-FLAG-Strep plasmid: with IPTG via the *T5* promoter and through native expression of the FliC promoter which is still incorporated in the FliC 5'UTR. While induction through the FliC promoter is regulated by the flagella genetic regulon via *flhDC* and *fliA*, CH2 protein expression was tuned with IPTG to maximise protein secretion in the *ΔCKL* strain. CH2 protein expression and secretion were measured as before, however cultures were supplemented with 0, 0.01, 0.05, 0.1 or 1mM IPTG.

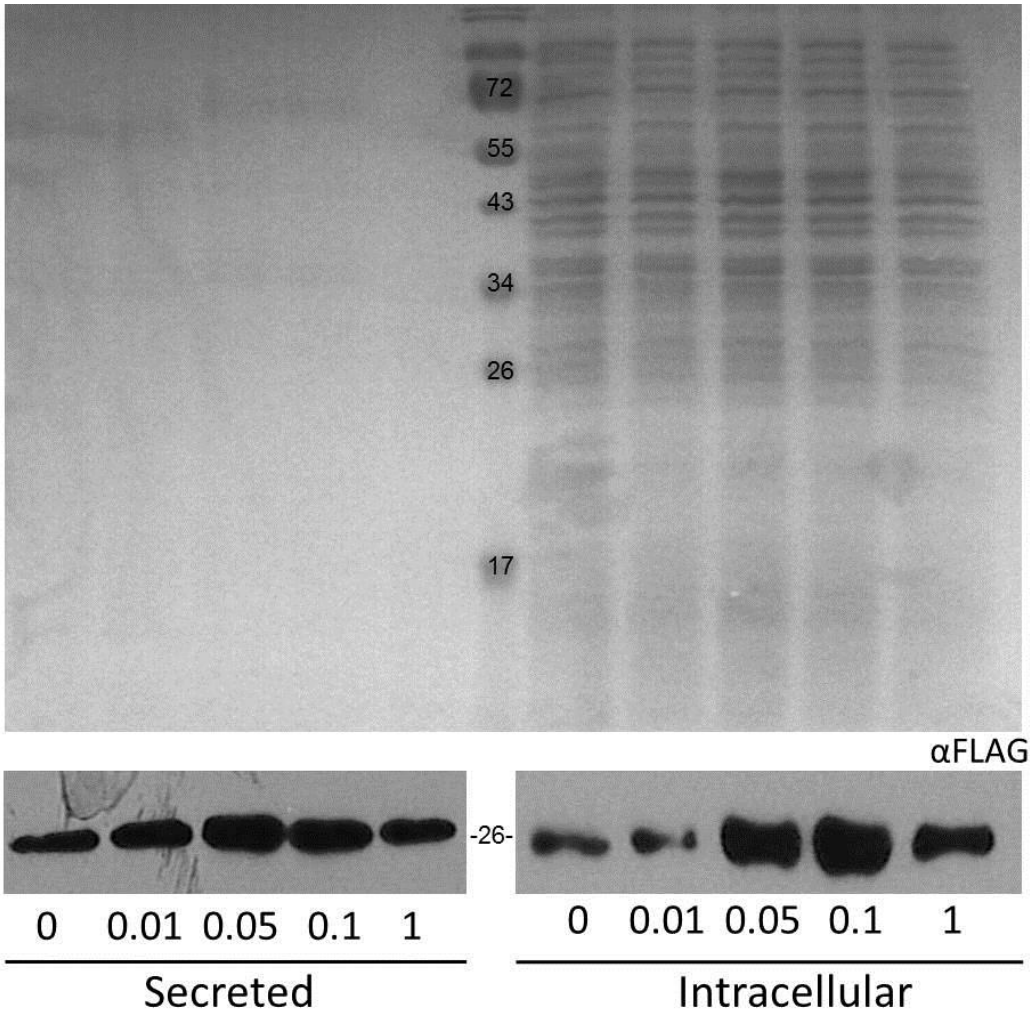


Figure 4.8: Intracellular and secreted fractions of MC1000 *ΔfliC ΔflgKL* supplemented with different concentrations of IPTG.

E.coli Δ*CKL* containing the plasmid pJexpress-FliC-CH2-FLAG-Strep was grown in LB supplemented with 0, 0.01, 0.05, 0.1 or 1mM IPTG. Cells were harvested at OD_{600} 1.5. The cells and supernatant from 1 OD unit of cells were prepared for (top) SDS-PAGE and (bottom) Western blot analysis of supernatant using anti-FLAG-HRP antibody. 15µL supernatant derived protein was loaded onto SDS-PAGE gels and 5µL and 2µL cell derived protein was loaded on to SDS-PAGE gels for Coomassie staining and Western-blotting respectively.

The Coomassie stained SDS-PAGE gel shows that the addition of different concentrations of IPTG did not alter levels of total intracellular protein expression and that overexpressed CH2 protein was not visible, suggesting expression may not be as strong as in BL21(DE3) (Figure 4.: top). Secreted fractions are clear of protein, demonstrating that cell lysis did not occur at a

detectable level with the induction of recombinant protein expression with any concentration of IPTG. The absence of any visible protein in secreted fractions also suggests that the concentration of secreted CH2 was below the detection threshold of Coomassie stain. The results of the Western blot show that CH2 was both produced intracellularly and secreted without and with IPTG induction (Figure 4.: bottom). CH2 expression was lowest with the addition of 0 or 0.01mM IPTG and was optimal at 0.1mM but surprisingly dropped off at 1mM IPTG. This was mirrored with secreted CH2, where CH2 was detected in the supernatant increased from 0 to 0.05mM IPTG where after the concentration of secreted CH2 decreased with the addition of more IPTG. This was confirmed following analysis with densitometry (Figure 4.). Secretion efficiency was calculated by adding the values derived for secreted and intracellular CH2 (total CH2) in a strain and then dividing this by the value for secreted CH2. This showed that the most efficient secretion occurred with the addition of 0.01mM IPTG (Figure 4.C). There was little difference in the secretion efficiency of cultures supplemented with 0.05 to 1mM IPTG, however 0.05mM resulted in the next most efficient secretion capacity of CH2.

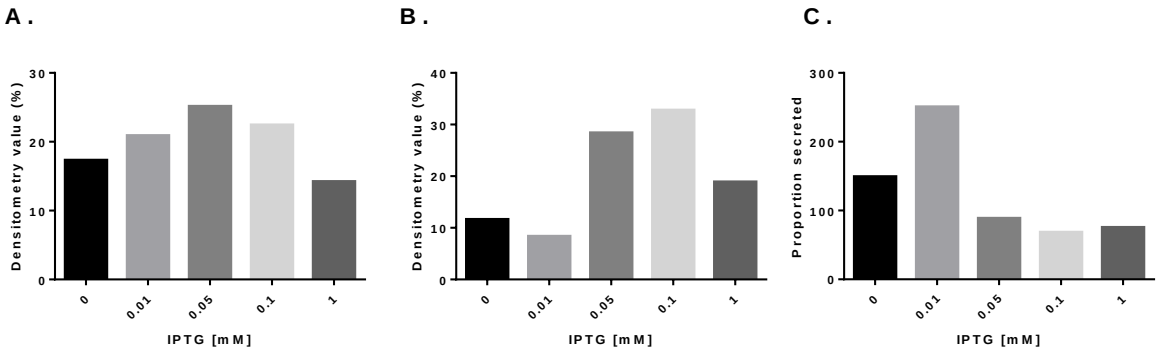


Figure 4.9: Expression and secretion capacities of MC1000 *ΔfliC ΔflgKL* supplemented with different concentrations of IPTG.

E.coli Δ*CKL* containing the plasmid pJexpress-FliC-CH2-FLAG-Strep was grown in LB supplemented with 0, 0.01, 0.05, 0.1 or 1mM IPTG. Cells and supernatant were harvested and prepared for SDS-PAGE and Western blotting with an anti-FLAG-HRP antibody, Densitometry analysis was carried out on the resulting Western blot using Image J. Results are shown for (A) densitometry value of secreted protein as a percentage of total secreted CH2 protein for all IPTG induction concentrations, (B) as stated, but for intracellular CH2, (C) as a proportion of CH2 detected in the supernatant in comparison to the intracellular fraction.

In subsequent CH2 secretion assay experiments, media was supplemented with 0.05mM IPTG. Although secretion wasn't most efficient at this level of induction, it did result in the second highest intracellular and highest secreted concentration of CH2. During the investigation of E2 secretion samples were taken once cultures reached OD_{600} 1.5, here samples will be taken at OD_{600} 1.0 due to the time constraints associated.

4.4.3. Isolating monomeric CH2 to quantify secretion capacity

As with the E2 secretion assay, it was beneficial to derive a CH2 protein standard to allow quantitative measurements of CH2 expression and secretion. This will give an indication of w/v yields of CH2 protein and also allow inter-experimental comparison of expression and secretion capacity. CH2 protein was overexpressed in BL21 (DE3) and purified and dialysed as described in previous sections. The concentration of CH2 was measured by BCA assay, however as multiple protein bands were seen in purified protein the total concentration of protein calculated by BCA assay cannot all be attributed to CH2. To account for this densitometry analysis was carried out on the Coomassie stained SDS-PAGE gel of CH2 protein standards (Figure 4.), this then allowed the proportion of CH2 in total protein to be determined and therefore the concentration of CH2 to be calculated.

The total concentration of protein found in elution franction 3 of the protein loaded onto Figure 4. was 41.4µg mL^{-1}. Densitometry calcualted that the band present at around 26kDa (CH2) comprised 33.708% total protein seen on the Coomassie stained SDS-PAGE gel. Therefore the concentration of CH2 protein in the protein standard is 14µg mL^{-1}. Therefore following Western blotting with the anti-FLAG-HRP antibody it can be concluded the intensity of the band seen at around 26kDa for the protein standard is due to a 14µg mL^{-1} concentration of CH2. This allows calculation of the concentration of CH2 in other samples due to proportional densitometry.

4.5. Improved secretion of recombinant protein through the FT3SS

With the protein secretion assay optimised to ensure that CH2 expression and secretion through the modified FT3SS are maximised and measurable in a quantitative manner, the focus of the study could now shift to testing the secretion capacity of CH2 in the secretor strain mutants, with the aim of improving CH2 protein secretion.

4.5.1. Establishing that secretion in engineered strains is truly FT3SS dependent

Prior to the investigation of improved secretion of CH2 through the modified FT3SS of strains with additional knock out mutations, it was important to ensure that CH2 secretion was facilitated by the FT3SS and not through cell lysis or other routes of non-directed protein secretion. This was investigated by measuring protein secretion in the negative secretion strains MC1000 *ΔflhDC* and MC1000 *ΔflgKL ΔfliC ΔflgDE.*

As seen in Figure 4., CH2 was detectable by anti-FLAG-HRP in the intracellular fraction of all strains, however in strains designed to act as negative controls for secretion of CH2 through the FT3SS (*ΔflhDC* and *ΔflgDE*) less CH2 was produced. In terms of strains which should demonstrate CH2 secretion, the *ΔflgMN* mutation resulted in slightly reduced CH2 expression in comparison to *ΔCKL*. These observations were confirmed with densitometry analysis of Western blots (two biological replicates) and quantification of CH2 protein concentration by calibration with the protein standard. In terms of intracellular protein, 54.8 mg L^{-1} and 33.9 mg L^{-1} CH2 was produced. While the *ΔflgMN* mutation did not result in more intracellular CH2 it did result in more secreted CH2: on average *ΔCKL* secreted 69.31 µg L^{-1} CH2 and *ΔCKL ΔflgMN* 128.21 µg L^{-1}, suggesting it is also more efficient at secreting protein than *ΔCKL.* Although a maximum of 0.51% total protein was secreted into the supernatant.

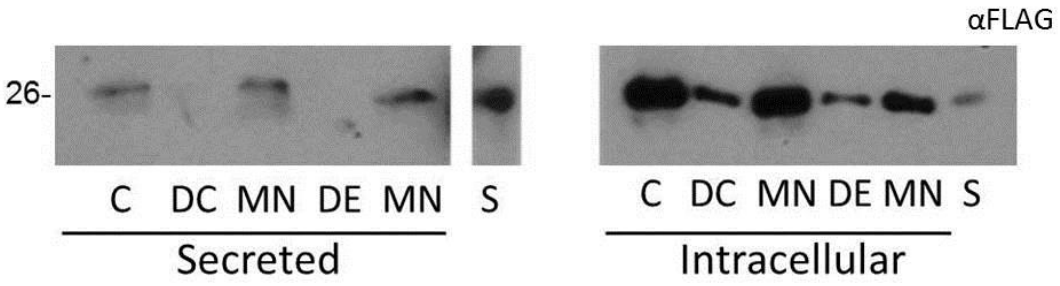


Figure 4.10: Intracellular and secreted CH2 from the *ΔCKL*, *ΔflhDC* and *ΔCKL ΔflgMN* and *ΔCKL ΔflgDE* mutants.

E.coli Δ*CKL* (C), Δ*flhDC* (DC), Δ*CKL* Δ*flgMN* (MN) or Δ*CKL* Δ*flgDE* (DE) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep was grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.0. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using an anti-FLAG-HRP antibody. A CH2 protein standard (S) was loaded to allow quantification of intracellular and secreted protein concentration. Samples were loaded as follows: Supernatant: 20µL, cells: 2µL, standard: 5µL. Note supernatant and cell fractions were run on separate SDS-PAGE gels.

To measure cell lysis, Western blots were stripped of anti-FLAG-HRP and reprobed with anti-GroEL to observe the presence of the cytoplasmic GroEL protein. Roughly similar concentrations were seen in the intracellular fractions of the different strains, however in the secreted fractions more GroEL was present with the addition of the *ΔflgMN* and Δ*flgDE* (where no secretion of CH2 was seen) mutations. As in Chapter 3, in the absence of a total cell lysis sample, it is not possible to calculate how much lysis this represents, however on account of the lack of protein in the supernatant fractions, this is unlikely to be very high at all.

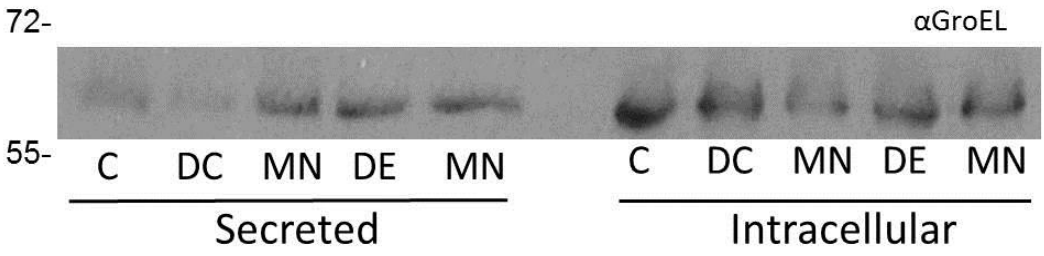


Figure 4.11: Abundance of GroEL in the intracellular and supernatant fractions of the secretor strain, Δ*flhDC,* Δ*flgMN* and Δ*flgDE* mutants, following expression and secretion of CH2 protein through the FT3SS.

E.coli ΔCKL (C), *ΔflhDC* (DC), Δ*CKL ΔflgMN* (MN) or *ΔCKL ΔflgDE* (DE) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB. Cells were harvested at OD_{600} 1.0. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using anti-GroEL antibody and an HRP secondary. Samples were loaded as follows: Supernatant: 20µL, cells: 2µL.

While CH2 is seen in the supernatant of positive secretion strains and is negligible in negative secretion strains, it is not possible to conclude that this is due to the FT3SS secretion alone because negative secretion strains express much less CH2 protein (Figure 4.). Therefore at this stage it is not possible to confirm that CH2 protein seen in the media is due to FT3SS directed secretion. The reduced level of CH2 expression in negative secretion strains is due to the presence of the *fliC* promoter in the secretion construct. The *fliC* promoter is subject to native transcriptional control of the cell, therefore as the *ΔflhDC and ΔflgDE* mutations result in reduced class III gene transcription due to trapping of FlgM inside the cell, less plasmid derived protein is expressed. To counteract this, the media of negative secretion strains was supplemented with additional IPTG, with the aim of compensating for reduced *fliC* promoter induction via increased T5 promoter induction. An alternative to this would be to produce a Δ*flgDE* Δ*flgM* mutant, resulting in free σ^{28} and therefore decoupling hook completion and the initiation of class III gene transcription. Despite several attempts to produce this mutant through Lambda-Red recombineering and phage transduction it was not achieved. It is not thought that this mutation combination would be lethal; therefore no explanation for this can be given, other than the common synthetic biology concept of the unpredictability of biology.

To investigate whether increased induction of the T5 promoter in the CH2 secretion construct harbouring plasmid, would result in increased CH2 expression in the negative control strains (*ΔflhDC* and *ΔflgDE*) the media was supplemented with additional IPTG. Following SDS-PAGE and Coomassie staining no visible differences in total intracellular protein were observed. CH2 was not visibly overexpressed with the addition of increasing concentrations of IPTG (Figure 4.B), neither was it observable in the supernatant fractions of cell culture (Figure 4.A). There is evidence of some intracellular protein contamination of the supernatant of *ΔflhDC* in all induction conditions. This is also true of *ΔflgDE* but is less pronounced. The result of the Western blot (Figure 4.C) showed that increasing concentrations of IPTG correlated with an increase in intracellular CH2 concentration. While *ΔflhDC* and *ΔflgDE* expressed less CH2 than the secretor strain with the addition of 0.05mM IPTG, this could be compensated with the addition of 1mM IPTG. This was confirmed by densitometry analysis (Figure 4.). CH2 was found in the supernatant of the secretor strain (837.36 µg L^{-1}) and in low quantities in *ΔflhDC* (108.36 µg L^{-1}) when supplemented with 1mM IPTG -in what was almost certainly a phenomenon caused by cell lysis in these cells (as illustrated by anti-GroEL Western below -Figure 4.). This amounted to 4.59% CH2 protein being secreted in the secretor strain, as opposed to 0.64% in *ΔflhDC* or 0% in *ΔflgDE* with the addition of 1mM IPTG.

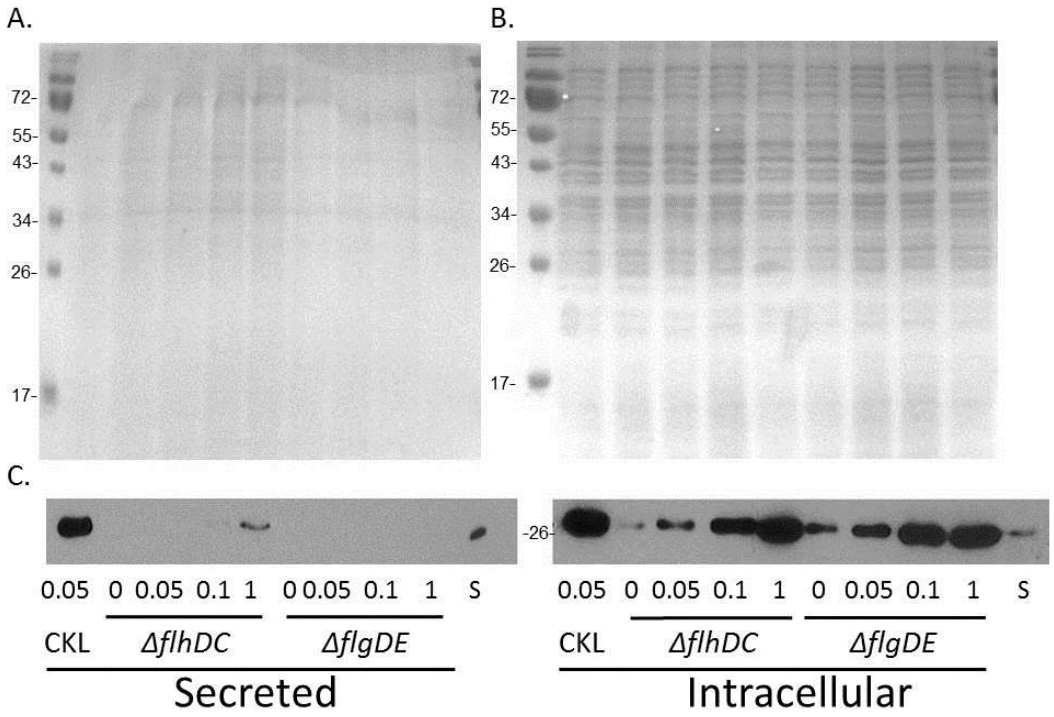


Figure 4.12: Intracellular and secreted CH2 from the secretor strain and the ΔflhDC and Δ*flgDE* mutants when supplemented with increasing concentrations of IPTG

E.coli Δ*CKL* (C), Δ*flhDC* (DC) or Δ*CKL* Δ*flgDE* (DE) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep was grown in LB supplemented various concentrations of IPTG (see annotation). Cells were harvested at OD_{600} 1.0. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain (A and B) or Western blot (C) analysis of cells and supernatant using an anti-FLAG-HRP antibody. A CH2 protein standard (S) was loaded to allow quantification of intracellular and secreted protein concentration during Western blot analysis. Note supernatant and cell fractions were run on separate SDS-PAGE gels. Samples were loaded to as follows: Supernatant: 20µL, standard 10µL. Cells: 2µL (Western blot), 5µL (Coomassie stain), standard: 5µL.

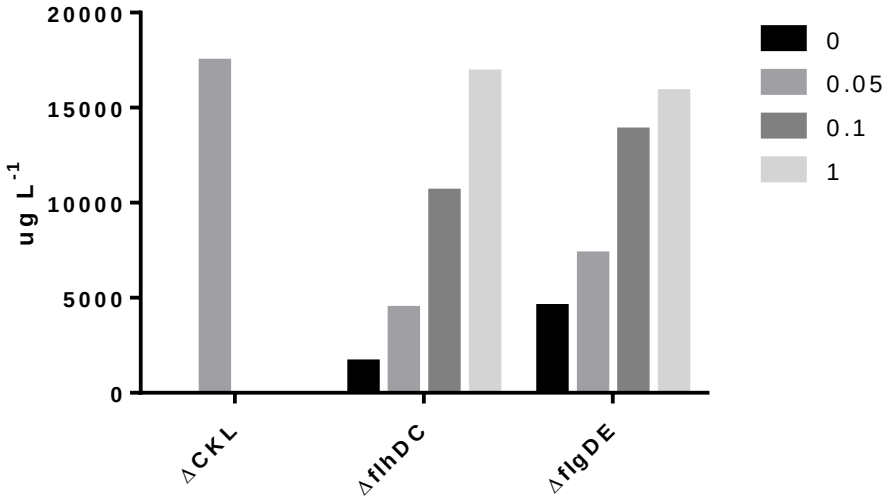


Figure 4.13: Quantitative measurement of intracellular CH2 from the secretor strain, Δ*flhDC* and Δ*flgDE* mutants following the addition of various concentrations of IPTG.

E.coli Δ*CKL* (C), Δ*flhDC* (DC)*,* Δ*CKL*) or Δ*CKL* Δ*flgDE* (DE) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB supplemented with various concentrations of IPTG to OD_{600} 1.0. A CH2 protein standard (S) was loaded to allow quantification of intracellular protein concentration. Following Western blot analysis of cells using anti-FLAG-HRP, densitometry analysis was carried out using ImageJ to calculate the concentration of CH2 in cells per litre of cell culture.

Cell lysis was measured by measuring supernatant fractions for the presence of the cytoplasmic chaperone protein GroEL (Figure 4.). The secretor strain and Δ*flgDE* mutant had similar concentrations of GroEL in the supernatant of cell culture, whereas Δ*flhDC* mutants had higher concentrations suggesting that cell lysis was higher in this strain. This complemented results seen in the Coomassie stain (Figure 4.). The addition of increasing concentrations of IPTG did not result in increased GroEL. Cell lysis suggests that cells are not healthy, however it is not documented that the *flhDC* mutant should have poor fitness, in fact it is a more competitive in animal and plant models (Gauger et al., 2007; Tans-Kersten et al., 2004). It is possible that some deleterious mutations have occurred in the strain; therefore it may be favourable to generate a fresh version of this mutant knockout. However for the purpose of this investigation lysis was still relatively low and presence of E2 protein in the supernatant was only visible following 1mM IPTG induction, suggesting that cytoplasmic leakage was very

low. For these reasons it was concluded that the strain is still suitable for use as a negative control. Most importantly little GroEL was detected in the supernatant of Δ*CKL* cells, suggesting that any protein seen in the supernatant of this sample is due to directed secretion and not lysis.

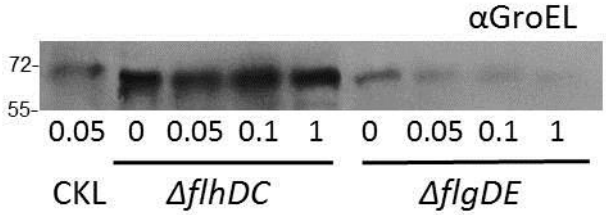


Figure 4.14: GroEL in the supernatant fraction of the CH2 expressing secretor strain and the ΔflhDC and Δ*flgDE* mutants with the addition of IPTG.

E.coli Δ*CKL* (C), Δ*flhDC* (DC) or Δ*CKL* Δ*flgDE* (DE) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep was grown in LB supplemented various concentrations of IPTG (see annotation). Cells were harvested at OD_{600} 1.0. Samples underwent SDS-PAGE and Western blot analysis using anti-groEL antibody and an HRP secondary.

Addition of higher concentrations of IPTG is an effective means of counteracting reduced *fliC* promoter activity in negative secretion strains. This allowed conformation that FT3SS secretion signal peptide tagged protein is not present in the supernatant of strains which lack fully functional secretion apparatus. With negative control established it can be confidently stated that the presence of these proteins in the supernatant of strains with functional secretion apparatus is due to directed FT3SS secretion alone.

4.5.2. Secretion construct modification for negative control of FT3SS secretion

While negative control of FT3SS secretion has been demonstrated by the absence of FliC secretion signal peptide tagged protein in the supernatant of strains with either non-functional

or incomplete secretion apparatus, the specificity of the secretion signal peptide has not been investigated. The creation of a secretion construct without the secretion signal peptide will serve as a negative control for protein secretion through the FT3SS. In addition to this the removal of the 5' UTR may alleviate the issue of cell mediated negative regulation on plasmid expression through the *fliC* promoter, as the promoter is situated in this 5' UTR region.

4.5.2.1. Production of the modified secretion construct

The region of the prototype secretion construct upstream of the protein cargo is currently comprised of the 5' UTR region of *fliC* followed by the 47 amino acid FliC secretion peptide. Primers were designed to amplify DNA from the plasmid pJexpress-FliC-CH2-FLAG-Strep without the 5' UTR or without the 5' UTR and the 47 amino acid FliC secretion signal peptide. As the secretion construct ribosome binding site and start codon were incorporated into the 5' UTR and signal peptide respectively, it was required to integrate them into the forward primers so that PCR products would harbour extensions of the original template with these motifs. The forward primer is capped with an *Nde*I restriction enzyme site and the reverse a *Bam*HI restriction enzyme site, this will allow the PCR product to be cloned into pJexpress with *Nde*I and *Bam*HI sticky ends.

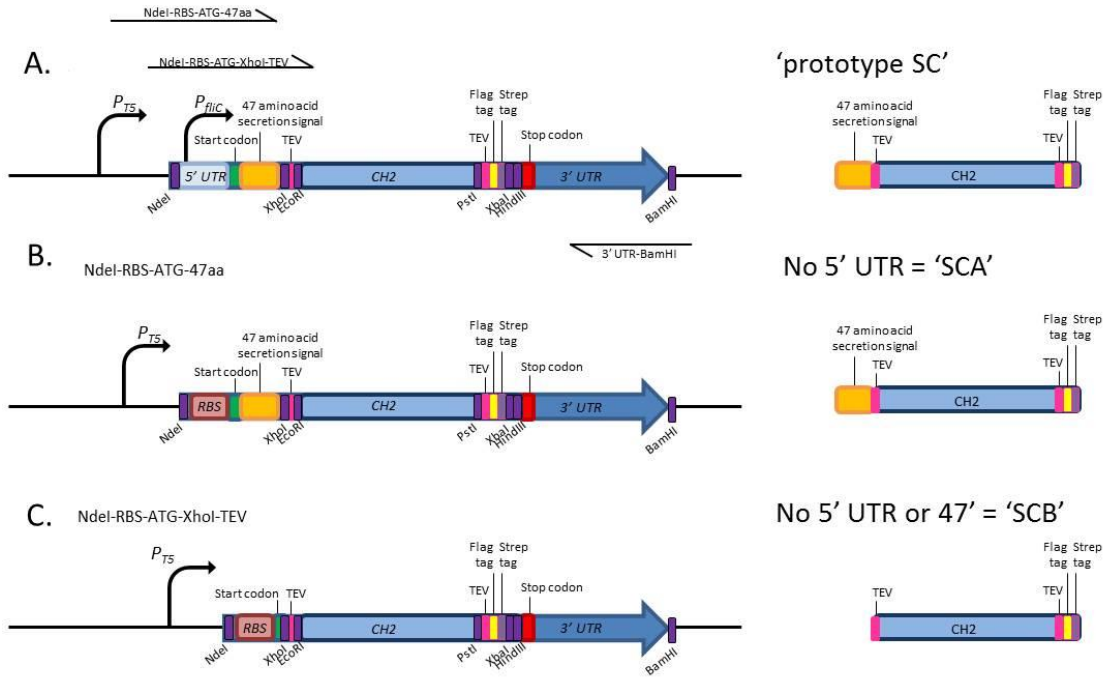


Figure 4.15: Schematic of the existing prototype secretion construct harbouring CH2 and the alternative secretion constructs (SCA and SCB) following the removal of the 5' UTR or 47 amino acid secretion peptide.

The genetic construct (left) and resulting protein based secretion construct (right) are given for three variations of CH2 harbouring constructs. The original prototype CH2 secretion construct which contains the 5' UTR and 47 amino acid secretion signal (A) was used as a template for PCR using one of two forward and one reverse primers with elongated ends. Restriction enzyme sites featured in the elongated ends of all primers and ribosome binding site (RBS) and start codon featured in the two forward primers. This resulted in the generation of donor DNA, which following restriction digest with *Nde*I and *Bam*HI was ligated into a *Nde*I/*Bam*HI cut plasmid harbouring the empty secretion construct. This resulted in the generation of two novel plasmids which novel genetic constructs (B and C), known as SCA and SCB. Expression of these resulted in the peptide products denoted on the right for each.

PCR products for both primer sets were derived by high-fidelity PCR. The size of PCR products when run on a DNA agarose gel confirmed the successful amplification of template DNA with the aforementioned primer mediated DNA extensions. PCR products were excised from the gel and isolated.

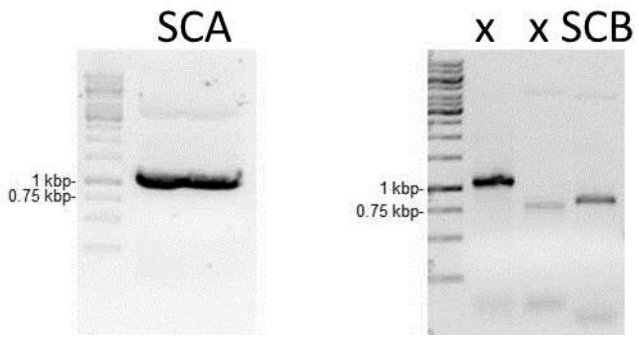


Figure 4.16: Agarose DNA gels showing SCA and SCB PCR products for ligation into *Nde*I and *Bam*HI cut pJexpress-FliC-empty-FLAG-Strep

PCR reactions were carried out using pJexpress-FliC-CH2-FLAG-Strep as template and primers in **Appendix 2** with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. Left: 50µL *Nde*I-RBS-ATG-47aa-CH2-Strep-FLAG-3' UTR-*Bam*HI (964bp). Right: (1) Ladder, (2 and 3) not relevant to this experiment, (4) 10µL *Nde*I-RBS-ATG-CH2-Strep-FLAG-3' UTR-*Bam*HI (823bp)

Isolated PCR products and pJexpress-FliC-empty-FLAG-Strep underwent restriction digest with *Nde*I and *Bam*HI in CutSmart buffer. Following this incubation step the acceptor vector was treated with Antarctic phosphatase to discourage re-circularisation and then run on a DNA agarose gel to confirm successful restriction digest. As a control, acceptor vector which had been incubated without restriction digest enzymes was also loaded on to the gel. As opposed to the smaller supercoiled circular DNA seen in lane 1, two DNA products are visible in the second lane, showing the cut acceptor vector (3967kbp) and the 772kbp section of DNA comprising of the empty secretion construct.

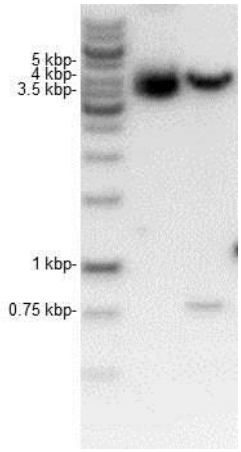


Figure 4.17: Agarose DNA gel showing uncut and *Nde*I and *Bam*HI cut pJexpress-FliC-empty-FLAG-Strep

Restriction digest reactions were carried out both with and without the addition of restriction enzymes. 5µL restriction digest mixture was prepared for DNA electrophoresis and loaded on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide and visualised under UV light and inverted. Lanes: (1) GeneRuler™ 1kb DNA ladder, (2) uncut vector, (3) *Nde*I and *Bam*HI cut vector

---

As the product resulting from restriction digest of PCR products was only 18bp shorter than the original product, these restriction digest mixtures were not visualised by DNA electrophoresis. All restriction digest reactions underwent PCR clean up to inactivate restriction enzymes, as *Bam*HI is not amenable to heat inactivation. DNA concentration of both acceptor vector and PCR products following restriction digest and clean up was measured and DNA ligation reactions were set up to contain the cut acceptor vector and either SCA or SCB DNA. Following incubation overnight, ligation reactions were transformed into NEB 5-alpha Competent *E. coli* and screened by plating on LB agar plates supplemented with ampicillin. Positive colonies were screened by colony PCR to confirm successful ligation of DNA.

Plamid DNA was isolated from colonies which yeilded positive colony PCR results and were sent for sequencing with primers which annealed to DNA regions flanking the insertion site of donor DNA following ligation. Sequences from both plasmids alligned to the expected sequence exactly. Protein expression and secretion can now be assessed for these newly derived plasmids.

4.5.2.2. Testing the modified secretion construct

Following the successful cloning of two versions of the CH2 harbouring secretion construct; one without the FliC 5' UTR (SCA) and one lacking both the FliC 5' UTR and the 47 amino acid secretion signal peptide (SCB) the expression and secretion of these plasmids was investigated in both the secretor strain and the negative secretion strain *ΔflgDE*.

As seen in the Coomassie stained SDS-PAGE gel, CH2 was not distinguishable in either the cell or secreted protein fractions (Figure 4.). The secreted protein fraction was clear of other protein contaminants aside from a large protein seen at around 72kDa. Despite this, as the remainder of the fraction as clear, it can be assumed that cell lysis did not occur in cell culture for any combination of plasmid or strain. Western blot analysis showed that CH2 harboured in the original secretion construct was produced intracellularly and secreted in *ΔCKL* cells (Figure 4.: bottom), whereas (as seen previously in Figure 4.) a very low concentration was produced in the *ΔflgDE* mutant. In the absence of the FliC 5' UTR the same concentration of CH2 was produced in *ΔCKL* cells as was when it was present. In the *ΔflgDE* mutant the absence of the 5' UTR resulted in an increase in CH2 expression as opposed to when it was present –in fact the concentration of CH2 was in line with that seen in *ΔCKL* cells. Without the 5' UTR CH2 was absent in the supernatant of both strains. In the absence of the FliC 5' UTR and the 47 amino acid secretion signal peptide, it was predicted that the size of the CH2 product would reduce to 13.5kDa; this was seen in the Western blot. Δ*flgDE* expressed slightly more than *ΔCKL;* however both strains had much more intracellular CH2 in the absence of the secretion signal, this was not expected. As expected in the absence of the secretion signal the supernatant of cells cultures of both *ΔCKL* and *ΔflgDE* was void of CH2 protein. This pattern was seen in two biological replicates.

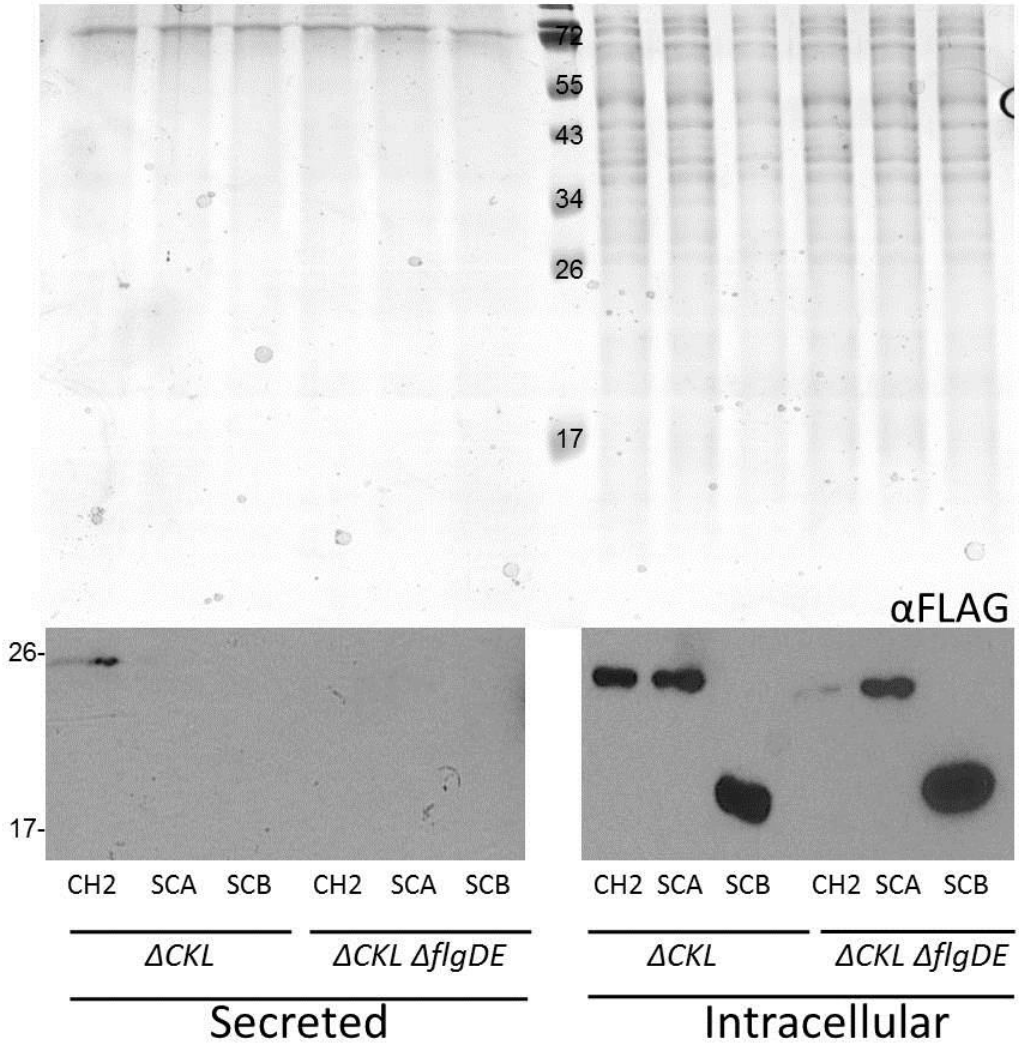


Figure 4.18: Intracellular and secreted CH2 from the secretor strain and *ΔCKL ΔflgDE* mutant with variable FliC 5' UTR and FliC secretion signal peptide sequences

E.coli Δ*CKL* (C) or Δ*CKL* Δ*flgDE* (DE) containing the either the plasmid pJexpress-FliC-CH2-FLAG-Strep or the SCA or SCB variant plasmid was grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.0. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain (top) or Western blot (bottom) analysis of cells and supernatant using an anti-FLAG-HRP antibody. Samples were loaded to as follows: Supernatant: 20µL. Cells: 2µL (Western blot), 5µL (Coomassie stain).

4.5.3. Reduced metabolic burden and negative regulation mutants: Effects of deletion of *motAB, flgMN, fliDST* and *clpX* on FT3SS secretion of CH2

With the development of a protein secretion assay for the non-native protein CH2 and the confirmation of control of directed FT3SS secretion with the negative secretion strains *ΔflhDC* and *ΔCKL ΔflgDE,* the investigation into the secretion capacity of CH2 in the suite of promising secretion strains which have been generated could begin. CH2 protein secretion assays were carried out to investigate the expression and secretion capacity of the *ΔCKL* strain with the addition of *ΔmotAB,* Δ*fliDST,* Δ*flgMN* and *ΔclpX* mutations. *ΔCKL* was assessed in tandem for reference and the CH2 protein standard often utilised to calculate both expression and secretion yields.

Figure 4. serves as an example of one of a number of biological replicates of CH2 protein secretion assay experiments, which were carried out. As seen in the Coomassie stained SDS-PAGE gel, while CH2 was not visible, the supernatant protein fraction was routinely clear of additional protein contaminants, suggesting that cell lysis was not commonplace. Western blot analysis demonstrated that intracellular CH2 concentration is variable between strains, however reasonable amounts are expressed in all, suggesting that CH2 expression is compatible in all mutant strains tested. This is also true of secreted protein; all mutant strains are capable of secreting CH2 through the modified FT3SS and into the media. As a CH2 protein standard was not loaded onto this Western blot it is not possible to quantify the concentration of CH2 protein loaded onto the gel to give a yield, however it is possible to use densitometry analysis to calculate intracellular or secreted CH2 in strains relative to that in the secretor strain. Samples for each strain on an SDS-PAGE gel were derived in tandem -from cell culture to sample preparation and finally Western blot analysis, therefore it is valid to normalise results for each strain to *ΔCKL*. As *ΔCKL* is common to every experiment, normalised results can be compared across experiments.

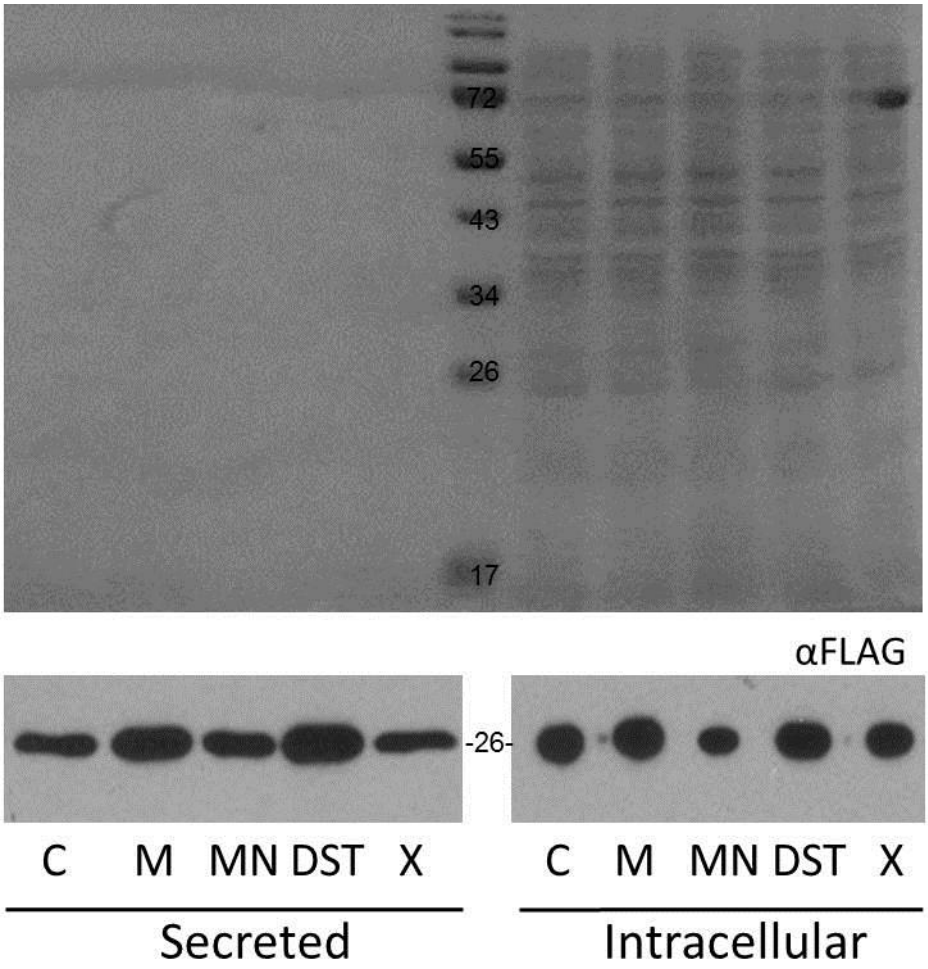


Figure 4.19: Intracellular and secreted CH2 from the secretor strain and Δ*motAB, ΔflgMN, ΔfliDST* and Δ*clpX* and mutants.

E.coli Δ*CKL* (C), Δ*CKL* Δ*motAB* (M), Δ*CKL* Δ*flgMN* (MN), Δ*CKL* Δ*fli*DST(DST) or Δ*CKL* Δ*clpX* (X) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.0. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. The supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) and prepared with 2x SDS-loading buffer. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain (top) or Western blot (bottom) analysis of cells and supernatant using an anti-FLAG-HRP antibody. Samples were loaded to as follows: Supernatant: 15µL. Cells: 2µL (Western blot), 5µL (Coomassie stain).

Following densitometry analysis of the Western blot shown in Figure 4. along with additional biological replicates, relative levels of CH2 expression and secretion in comparison to the secretor strain were derived (Figure 4.). Aside from secreted CH2 in the *ΔmotAB* mutant, there is quite a lot of variation observed in the relative CH2 concentrations in comparison to *ΔCKL*. The trend observed is that all additional knockout mutations result in on average, less intracellular CH2 and more secreted CH2 in comparison to the *ΔCKL*. *ΔCKL ΔflgMN* produced 0.71 CH2 protein in comparison to the *ΔCKL*, this was significantly less. In terms of secreted CH2, the *ΔfliDST* mutation resulted in an average of 1.64 CH2 relative to *ΔCKL*, however this was not found to be significant. The three other knockout mutations all resulted in significantly more secreted CH2 relative to *ΔCKL* (*ΔCKL ΔmotAB* = 1.53*, ΔCKL ΔflgMN* =1.46, *ΔCKL ΔclpX* = 1.55). These findings suggest that all four strategies to improve CH2 are effective, as more CH2 is secreted and furthermore as less CH2 is found intracellularly, these strains are more efficient at secreting CH2 than the original secretor strain.

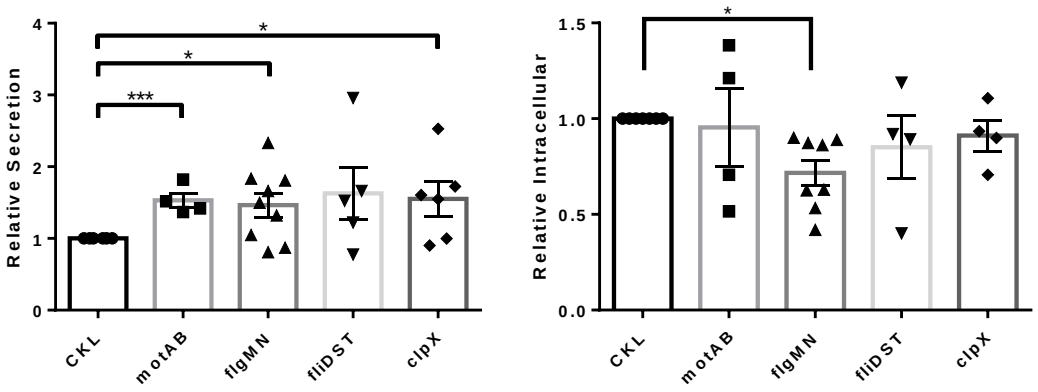


Figure 4.20**: Western blot densitometry analysis of intracellular and secreted protein fractions of CH2 expressing secretor strain and Δ*motAB, Δflg*MN*, Δfli*DST and Δ*clpX* and mutants**.

E.coli Δ*CKL*, Δ*CKL* Δ*motAB*, Δ*CKL* Δ*flgMN,* Δ*CKL* Δ*fli*DST, or *CKL* Δ*clpX* containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.0. Following Western blot analysis of cells and supernatant using anti-FLAG-HRP, densitometry analysis was carried out using ImageJ. Results were then normalised to the value obtained for either secreted or intracellular CH2 in the Δ*CKL* strain (1). Individual data points, average and standard error of the mean are given for secreted (left) and intracellular (right) relative concentration of CH2. Unpaired t-test was carried out (* = $p<0.05$, *** = $p<0.0005$). Biological replicates (secreted, intracellular): Δ*CKL* (6, 7), Δ*CKL* Δ*motAB* (4, 4), Δ*CKL* Δ*flgMN* (9, 8), Δ*CKL* Δ*fli*DST (5, 4), *CKL* Δ*clpX* (6, 4)

While relative concentrations of intracellular and secreted CH2 are informative, it is desirable to obtain quantitative yields so that productivity can be compared to industry standards. CH2 protein standard was loaded alongside many biological samples; however the high incidence of high background and weak fluorescent signal resulted in many being unsuitable for analysis by densitometry. From the results obtained, the overall trend was the same as that observed in Figure 4.: the addition of all mutations resulted in the average yield of secreted protein being higher in the secreted fraction and lower in the intracellular fraction (Figure 4.: A and B). In addition, all mutations resulted in more efficient secretion of CH2 into the media of cell cultures (Figure 4.C), the highest being *ΔCKL ΔclpX* which secreted 1.39% total CH2 protein into the media. *ΔCKL ΔflgMN* achieved both the highest yield and average yield of secreted protein

(219.66 and 128.64 µg L^{-1}). *ΔCKL* demonstrated the highest level of CH2 protein expression (maximum: 75.66 mg L^{-1}, average: 41.83 mg L^{-1}).

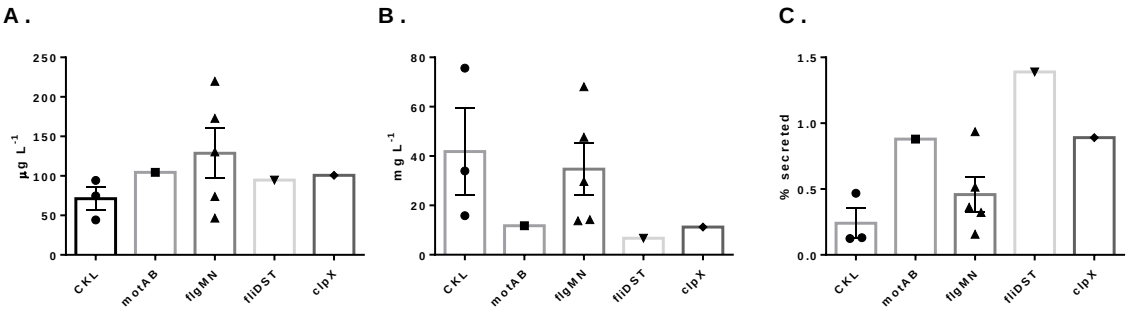


Figure 4.21**: CH2 yield in the intracellular and secreted protein fractions of the secretor strain and Δ*motAB,* Δ*flgMN,* Δ*fliDST* and Δ*clpX* mutants**.

E.coli Δ*CKL*, Δ*CKL* Δ*motAB*, Δ*CKL* Δ*flgMN,* Δ*CKL* Δ*fli*DST, or *CKL* Δ*clpX* containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.0. A CH2 protein standard was loaded onto SDS-PAGE gels along with biological samples. Following Western blot analysis of cells and supernatant using anti-FLAG-HRP, densitometry analysis was carried out using ImageJ and CH2 protein yields were calculated per Litre of cell culture for secreted (left) and intracellular (middle) protein fractions. The proportion of secreted protein was also calculated as a % of total CH2 protein both secreted and retained intracellularly. Biological replicates (secreted, intracellular): Δ*CKL* (3), Δ*CKL* Δ*motAB* (1), Δ*CKL* Δ*flgMN* (5), Δ*CKL* Δ*fli*DST (1), *CKL* Δ*clpX* (1). No significant results were found.

4.5.4. Reduced metabolic burden: Effects of deletion of *motAB, flgMN, fliDST* and *clpX* on growth phenotype

Gene deletions may lead to reduced metabolic burden as protein production is a metabolically costly process. It was hypothesised that the mutant strains would result in reduced metabolite expenditure and the reallocation of metabolites to essential processes such as cell growth but also to producing the remaining flagella proteins. Conversely the production of recombinant protein (i.e. CH2) will add to metabolic burden and may lead to a reduction in growth rate. Finally gene deletion may lead to strains which are not viable or less healthy if essential genes or processes are affected. To evaluate the effect of gene deletion and CH2 expression on growth phenotype, growth curves were obtained for *ΔCKL* and Δ*CKL* Δ*motAB*, Δ*CKL* Δ*flgMN,* Δ*CKL* Δ*fli*DST, or *CKL* Δ*clpX* expressing either CH2 or empty vector.

When expressing empty secretion construct *ΔCKL* had highest OD_{600} at all times (Figure 4.). The next was Δ*fliDST*, followed by Δ*flgMN* and Δ*motAB.* A lower OD_{600} for Δ*clpX* was consistently recorded, both in the log phase and stationary (statistically significant for both parts of the growth phase- Paired t-test of average OD_{600} values for each strain. p = <0.005). While all other strains reached a similar final stationary phase OD_{600}, Δ*clpX* was significantly lower. When expressing CH2 the growth phenotype of all strains was very similar. In the late log phase Δ*fliDST* (Paired t-test of average OD_{600} values for each strain from 5 to 10 hours. p = <0.0001) had a slightly higher OD_{600} and Δ*motAB* slightly lower OD_{600} than the other strains (Paired t-test of average OD_{600} values for each strain from 6 to 11 hours. All at least p = <0.01). The mean generation time for Δ*CKL* is 1.15 fold higher when expressing CH2, suggesting that overall expressing recombinant CH2 does lower growth slightly -this was investigated further in Figure 4..

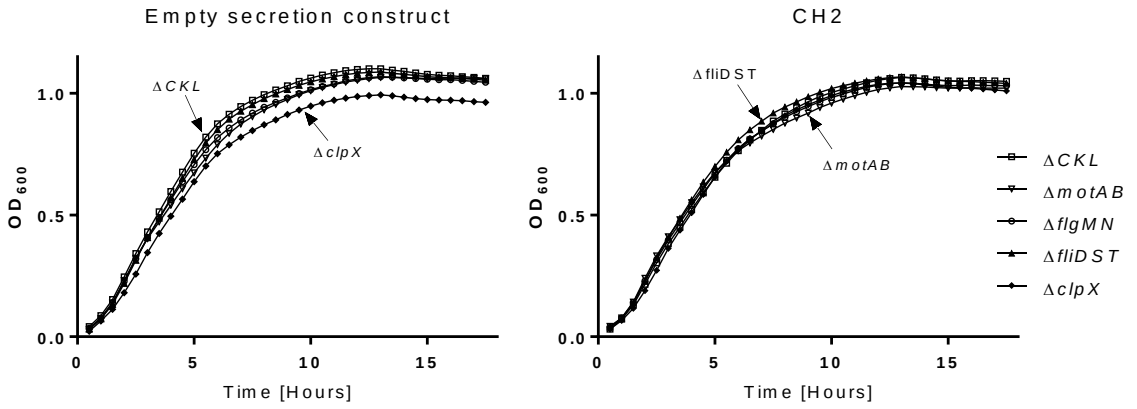


Figure 4.22: Growth of the Δ*CKL* strain and (Δ*CKL)* Δ*motAB,* Δ*flgMN,* Δ*fliDST* and Δ*clpX* mutant strains expressing empty or CH2 harbouring secretion construct over time in a 96 well plate. Calibrated to LB media.

Starter cultures of cells expressing either pJexpress-FliC-empty-FLAG-Strep (left) or pJexpress-FliC-CH2-FLAG-Strep (right) were grown in LB supplemented with ampicillin and 0.05mM IPTG and aliquoted into 96 well plate wells. The 96 well plate was incubated in the TECAN plate reader at 37°C with 6mm orbital shaking. OD_{600} measurements were recorded every 30 minutes. Δ*CKL*: □, Δ*motAB:* ▽ , Δ*flgMN:* ○, Δ*fliDST:* ▲, Δ*clpX:* ♦. Results represent the combination of three technical repeats, of six biological repeats.

When visualised by strain it is possible to evaluate the effect of CH2 expression on cell growth phenotype (Figure 4.). Both Δ*CKL* and Δ*fliDST* have similar early log and stationary phase OD_{600}, however in the late log phase optical density was significantly reduced in cells expressing CH2 (Paired t-test of average OD_{600} values for strain expressing empty or CH2 vector from 4 to 12 and 5.5 to 10.5 hours respectively. Both p = <0.0001). Δ*motAB* and Δ*flgMN* had similar early log optical densities whether expressing CH2 or not, however in the late log and stationary phase (7 hours onwards) cells expressing CH2 exhibited significantly lower cell density (Paired t-test of average OD_{600} values for strain expressing empty or CH2 vector. Both p = <0.0001). The growth rate of Δ*clpX* cells in the log phase was comparable whether expressing CH2 or empty vector (mean generation time 1.03 fold higher for empty secretion construct), however in the very late log and stationary phase cells expressing CH2 had a higher OD_{600} (Paired t-test

of average OD_{600} values for strain expressing empty or CH2 vector from 8 hours onwards. Both p = <0.0001). The main finding of this data is that CH2 is not toxic to cells and that cells from all strains are healthy for the duration of the growth curve. When expressing empty secretion construct Δ*clpX* does not grow very well, however this observation is off set when expressing CH2 protein as Δ*clpX* has a higher rate of growth when expressing CH2. This shows that not only were knockout mutations non-deleterious to strains, but that cell death (and therefore lysis) is not prevalent, therefore FT3SS secretion is true.

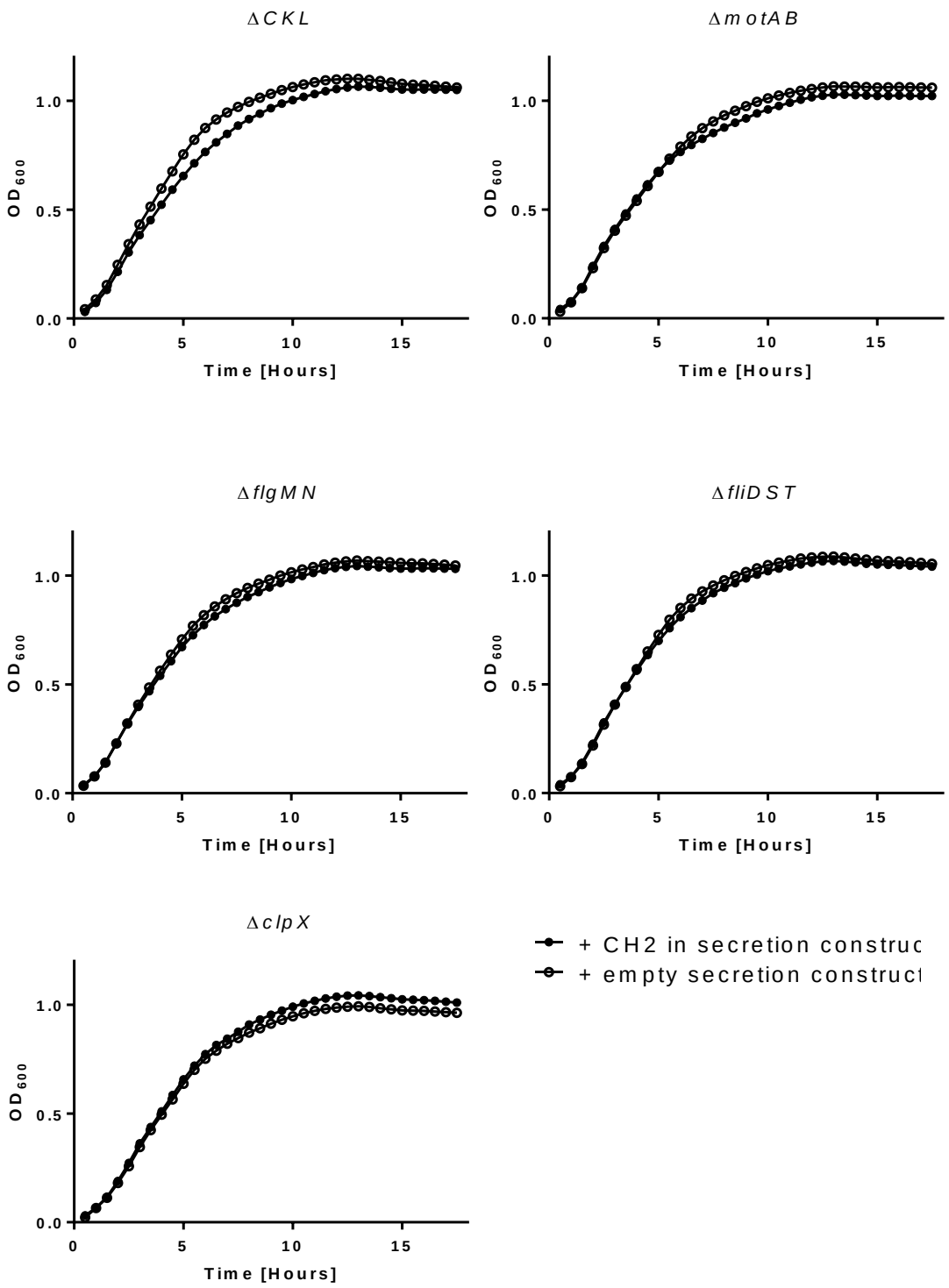

ΔCKL
ΔmotAB
ΔflgMN
ΔfliDST
ΔclpX
OD600
Time [Hours]
0.0
0.5
1.0
0
5
10
15
+ CH2 in secretion construc
+ empty secretion construct

Figure 4.23: Growth of Δ*CKL* and Δ*motAB, ΔflgMN, ΔfliDST* and Δ*clpX* mutant strains expressing empty or CH2 harbouring secretion construct over time in a 96 well plate. Calibrated to LB media.

Starter cultures of cells expressing either pJexpress-FliC-empty-FLAG-Strep (white fill) or pJexpress-FliC-CH2-FLAG-Strep (black fill) were grown in LB supplemented with ampicillin and 0.05mM IPTG and aliquoted into 96 well plate wells. The 96 well plate was incubated in the TECAN plate reader at 37°C with 6mm orbital shaking. OD_{600} measurements were recorded every 30 minutes. Results represent the combination of three technical repeats, of six biological repeats.

4.6. Discussion: Secretion of a non-native antibody fragment through the FT3SS

4.6.1. Performance of the synthetic modular secretion construct

The prototype secretion construct was designed to mediate secretion of recombinant protein through the modified FT3SS. Design features were included to allow protein purification, antibody detection and finally isolation of the protein product from the remainder of the secretion construct. Initial experimental work set out to assess the effectiveness of these features. While a number of recombinant proteins could have been selected, preliminary tests found that expression of the antibody fragment, CH2 was both efficient and consistent in Δ*CKL* cells, therefore the focus of this work was the expression and secretion of the prototype secretion construct with CH2 cargo (shortened to 'CH2' for the purpose of this discussion).

Prior to investigation of recombinant CH2 secretion through the modified FT3SS, the features of the secretion construct were tested to ensure they were effective, as they would enable all characterisation and downstream processing of CH2 following secretion. Design features included a Strep-tag for protein purification, FLAG-tag for antibody detection and TEV protease sites to yield pure CH2 product, free of the secretion construct. Protein purification of intracellularly expressed CH2 was achieved using a StrepTrap™ HP column. This resulted in purification of CH2 protein (Figure 4.). The CH2 protein was predicted to be 22.4kDa by ExPASy, however the prominent band seen following purification ran slightly larger than this. Despite this discrepancy in size it can be assumed that this was the full CH2 protein as it is common for proteins to run at slightly different sizes to those predicted on SDS-PAGE gel, this may be due to a number of factors, for example the amount of SDS which binds to the protein may vary to that predicted, giving the protein a different negative charge, resulting in a different movement through the SDS-PAGE gel matrix. Other proteins consistently co-eluted with CH2, they had no affinity for anti-FLAG-HRP therefore they are not aggregated or degraded CH2 protein. While steps were explored to reduce contaminants in the elution fractions or to remove them with dialysis or chromatography, they were not effective. This showed that it is possible to express CH2 in the secretion construct at high levels and at least

partially purify protein- although more work is needed to understand how to improve this. Since the ultimate goal is secretion, purification from the cytoplasm is relevant, but not the ultimate goal so it was not pursued further. It does highlight issues with cytoplasmic protein purification and demonstrates one of the reasons why industrial biotechnology companies are so interested in true secretion. In the future more methods could be explored, for example purifying CH2 from a different expression strain to BL21 (DE3). However it still allowed the derivation of a CH2 protein standard to quantify yield of secreted CH2, therefore these observations were not limitations to the study.

CH2 detection was facilitated by the FLAG-tag. Anti-FLAG-HRP proved effective in Western blot analysis of CH2 protein (Figure 4.). Again the corresponding band was larger than predicted, however as it corroborates with the size of the protein seen in Coomassie stained SDS-PAGE gels following protein purification (Figure 4. and Figure 4.), this further confirmed that this protein was CH2. When high concentrations of CH2 protein were loaded onto the SDS-PAGE gel some additional HRP signal was detected for different sized proteins (Figure 4.), one of which ran at around 55kDa and may be the dimeric form or CH2. The other bands were small and therefore likely to be degradation products of CH2. As CH2 is the most abundant protein which anti-FLAG-HRP detects, this is not an issue experimentally.

Isolation of CH2 from the rest of the secretion construct with TEV protease was effective. The results of Western blot analysis clearly show the loss of the full size CH2 protein following incubation with TEV protease (Figure 4.). The Coomassie stain was less conclusive, but suggested that some non TEV directed cleavage of CH2 protein occurred, however a higher concentration was always present following incubation with TEV as opposed to without. Analysis of TEV cleavage of CH2 by Coomassie stain was hindered for a number of reasons; firstly the initial concentration of CH2 was low, therefore not clearly visible. Additionally, the sizes of the CH2 and TEV protease were similar, meaning it was difficult to distinguish them. In the future loading a higher concentration of CH2 onto SDS-PAGE gels would be preferential in terms of visibility. In would also be desirable to remove the TEV protease from the reaction mixture, so that it does not obscure the visualisation of CH2 –this would also allow pure CH2 to be isolated from the resulting solution by centrifugation. This could easily be achieved as a His-tag is incorporated into AcTEV™ Protease to facilitate its removal from solution.

Overall the secretion construct was effective in facilitating CH2 overexpression, protein purification, antibody detection and finally isolation of CH2 from the secretion construct. Furthermore the CH2 protein is able to form dimeric structures as seen in Figure 4..

4.6.2. Secretion of CH2 through the modified FT3SS

Expression and secretion of CH2 was achieved in all strains with functional secretion apparatus (i.e. not negative controls). The anti-FLAG-HRP antibody was effective at detecting CH2 both in the intracellular and secreted cell culture protein fractions (Figure 4.). It was found that the addition of 0.05mM IPTG to cell cultures resulted in the highest concentration of secreted CH2 (Figure 4.) therefore all further work was carried out with this induction protocol as increased secreted protein is the ultimate goal of this project. For ease of experimental work it was decided to grow cells to OD_{600} 1.0, so that time constraints were limited and therefore experimental output could be high. A CH2 protein standard was derived and although the issues of co-elution were also a factor here, the combination of densitometry analysis and protein concentration assay allowed the calculation of the concentration of CH2 protein. This could be correlated to the HRP signal which resulted from detection of CH2 to accurately measure the concentration of CH2 in biological samples. These considerations resulted in a protein secretion assay which is suitable to measure variable CH2 expression and secretion in modified FT3SS secretion strains, furthermore as the cargo protein can be substituted, this secretion assay is suitable for detection of a multitude of recombinant proteins, should they be inserted into the secretion construct.

4.6.3. Secretion of CH2 through the modified FT3SS: strain mediated negative control

It was imperative to ensure that CH2 detected in the media of cell culture resulted from directed FT3SS secretion and not through other routes. While other research groups have demonstrated secretion of recombinant protein through the FT3SS, they rarely take adequate

steps to ensure negative control of protein secretion, which effectively exclude cell lysis as a route for extracellular protein.

Negative control strains were derived (*ΔflhDC* and *ΔCKL ΔflgDE*) and following supplementation with additional IPTG -to compensate for flagella gene expression being negatively regulated via the *fliC* promoter in the negative secretion strains- absence of CH2 protein in the supernatant of negative secretion strains was confirmed (Figure 4.). The comparable intracellular CH2 expression in intracellular protein fractions of all strains and the absence of cell lysis in *ΔCKL* cells (Figure 4.), gave added weight to this conclusion, as it confirms that protein seen in the supernatant of these cells was directed by secretion and not cell lysis. Some GroEL was detected in the supernatant of *ΔflhDC* cells cultures suggesting that cell lysis had occurred, however as GroEL was detected at a low concentration for *ΔflhDC* and high for *ΔflgDE* cells in Figure 4., this suggests that strains do not inherently lyse and therefore it can be concluded with confidence that any increase in CH2 in the supernatant is due to increased capacity of the FT3SS.

The second control strategy involved the removal of the FliC 5' UTR from the secretion construct (SCA). In addition to this a secretion construct which also lacked the 47 amino acid FliC secretion signal peptide (SCB) was also constructed to confirm that secretion was under the control of the signal peptide. The modified secretion constructs were successfully produced and expressed in both the *ΔCKL* and the negative secretion strain *ΔCKL ΔflgDE* (Figure 4.). Removal of the 5' UTR did alleviate the reduction in CH2 expression previously observed in negative secretion strains (Figure 4.), as a similar concentration of CH2 was expressed in *ΔflgDE* cells as in *ΔCKL* when this SCA was expressed. Interestingly the removal of the *fliC* promoter did not cause reduced CH2 expression in *ΔCKL* cells, whereas it might be hypothesised that the removal of one of the two promoter sites for plasmid gene expression would result in less gene expression.

As CH2 was visible in the secreted fraction of secreted *ΔCKL* cells when expressed in the prototype secretion construct, but not SCA or SCB, this suggests that the 5' UTR is essential for secretion through the FT3SS, this is in line with the findings of (Majander et al., 2005; Narayanan et al., 2010). As both the 5' UTR and 47 amino acid secretion signal were absent in SCB, it is not possible to conclude whether the 47 amino acid secretion signal is essential for secretion. In the future it would be beneficial to produce a secretion construct which harbours the 5' UTR without the 47 amino acid secretion signal, to enable this investigation. This effect

of the presence and absence of the UTRs and the 47 amino acid secretion signal will be investigated further in the next chapter.

4.6.4. Effects of reduced metabolic burden and negative regulation mutants on FT3SS secretion of CH2

Once the non-native FT3SS protein secretion assay had been optimised and control of secretion established with negative secretion strains, it was possible to begin to test the secretion capacity of modified FT3SS secretion strains with confidence that CH2 present in the supernatant was due to directed secretion through the modified FT3SS. Unlike the E2 protein (Figure 3. and Figure 3.) CH2 was effectively expressed and secreted in all secretor strain mutants (Figure 4.). This was expected due to the decoupling of CH2 from any native chaperones. This also demonstrates that a substantial concentration of CH2 was maintained intracellularly without the requirement of a chaperone. On average all mutations resulted in strains that secreted a higher concentration of CH2 through secretion apparatus into the media (Figure 4.). Therefore the strain improvement mutagenesis strategies which were theorised to result in increased secretion capacity, were effective. Mutant strains were reasonably similar in their secretion output, ranging from 1.46 to 1.64 times the output of *ΔCKL*. While *ΔCKL ΔfliDST* produced the highest single and average relative CH2 secretion value, it was not significant, whereas the other mutant strains were significantly more effective at secreting CH2 than *ΔCKL*. Conversely all mutant strains produced less intracellular CH2 in comparison to *ΔCKL*, however aside from *ΔflgMN* (which produced 0.71 fold intracellular CH2 in comparison to *ΔCKL*) these were not statistically significant observations. As the percentage of secreted protein through the FT3SS is consistently very low, it is unlikely that this reduction in intracellular CH2 is wholly an artefact of high secretion through the secretion apparatus. That aside –combined, decreased intracellular and increased secreted concentrations of CH2 does mean that mutant strains were more efficient at secreting CH2 than *ΔCKL*, as a higher percentage of total CH2 was located in the supernatant (Figure 4.C).

In terms of yields, while it would be desirable to obtain some more measurements it can be concluded that the concentration of intracellular CH2 ranges from an average of 41.83 mg L^{-1} in *ΔCKL* to 6.72 mg L^{-1} in *ΔfliDST* (Figure 4.). The highest single concentration of secreted CH2

observed was 219 µg L^{-1} in *ΔflgMN*, this strain also demonstrated the highest average at 128.64 µg L^{-1}. This compared to 71.08 µg L^{-1} in *ΔCKL* (meaning either a three or two fold increase in yield of secreted protein with the *ΔflgMN* mutation depending on whether you use highest or average yield figures). This demonstrates that secretion capacity of strains may be substrate dependent as in the previous chapter both *ΔmotAB* and *ΔclpX* secreted higher concentrations of E2 into the media (Figure 3.).

In comparison to the yields achieved for E2, CH2 was much lower by all measurements. For example *ΔCKL* cells produced an average of 443.92 mg L^{-1} E2, in comparison to 41.83 mg L^{-1} – around 10 times less. This was expected as CH2 is a non-native protein; therefore the cell is unlikely to be as effective at producing it. More striking is the inefficiency of the modified FT3SS to secrete CH2 in comparison to E2. While an average of 3.06% total E2 protein was secreted into the media of *ΔCKL* cells, only 0.24% CH2 was. Reduced expression and secretion manifested in a reduced yield of secreted CH2 (71.08 µg L^{-1}) by around 100 times, in comparison to E2 (7.75 mg L^{-1}). Considering that E2 is essentially the native substrate of the FT3SS, whereas the only commonality of CH2 is the UTR regions and secretion signal harboured in the secretion construct, this was logical. This demonstrates the limitations of using recombinant expression systems and also highlights that while the FT3SS has affinity for non-native proteins, it is not as efficient as secreting them as it is native substrates. In comparison to other reports of FT3SS secretion of recombinant protein (12mg L^{-1} and 10.8mg L^{-1} for Majander et al. (2005) and Widmaier et al. (2009) respectively), the yield achieved here was around 150 times lower. In comparison to the 2mg L^{-1} yields reported for T1SS secretion of recombinant protein into the extracellular space this was around 30 times less (Fernández, 2004; Fernandez & de Lorenzo, 2001). However aside from the T1SS data, the yields reported in the literature were calculated based on a period of growth around three times longer than observed here, therefore yields could be improved upon if cell culture was prolonged. The concentration of CH2 reported for T2SS secretion is higher again; however as it is not extracellular secretion this was not strictly comparable.

The yields achieved for the production and secretion of CH2 were also low in comparison to industrial standards for other extracellular protein secretion expression strains –for example 1.5g insulin or 2.5g amylase secreted per Litre in *Pichia* or *Saccharomyces* respectively, although again this was over a period of 80 or 96 hours and at high cell density as opposed to around 6 hours and an optical density of 1.0 in these experiments (Gurramkonda et al., 2010;

Rodríguez-Limas et al., 2015). In terms of intracellular production of CH2 protein, this was in line with yields reported for antibody fragment production in the literature, for example a yield of 79mg L^{-1} for a Fab antibody fragment was achieved in *E. coli* (49mg L^{-1} was secreted into the periplasm by an unknown route) which was comparable to the 75.66 mg L^{-1} yield attained here (Jalalirad, 2013).

As stated the highest single intracellular and secreted CH2 yields were 75.66 mg L^{-1} *(ΔCKL*) and 219 µg L^{-1} (*ΔflgMN*) respectively. This yield relates to the weight of CH2 per Litre of cell culture at OD_{600} 1.0. However to make the production of low yield, low value products economically feasible, industry routinely grow their cell cultures to optical densities in excess of OD_{600} 100 to reduce production costs (Soini et al. 2008; Shiloach & Fass, 2005), therefore a litre of cell culture, will contain exponentially more cells that in the cultures grown in this study. Therefore the weight per litre yield is misleading. For example a crude calculation estimates that 1mL of culture at OD_{600} 1.0 will contain 8.0 x 10^8 *E. coli*, compared to 8.0 x 10^{10} at OD_{600} 100. If the same yield was reported per litre of cell culture for cells at both OD_{600} 1.0 and 100, the cells in the OD_{600} 100 culture would be 100 times less efficient. Culture of *E. coli* to a high optical density, requires adequate aeration and nutrient supply (often supplemented with glucose for example) and requires a longer period of cell culture (e.g. 24 hours minimum to reach OD_{600} 100), compared to cell cultures in this thesis (Soini et al., 2008). This is unlikely to be a linear effect as other factors such as limited oxygen and nutrient supply, along with build-up of growth limiting chemicals such as acetate, will interplay at different optical densities; however this is an important conceptual point. The main finding here is that a modification of the FT3SS followed by strain improvements and the construction of a modular secretion construct resulted in an intracellular concentration of the antibody fragment CH2 which could rival other yields reported in the literature, however the next chapter will focus on the implementation and combination of more strategies to increase FT3SS secretion, with the aim of improving this yield.

It is evident that these experiments produce variable results. Nevertheless, statistically significant differences were obtained for some strains with regards to secreted and intracellular protein. Efforts to reduce variability have been put in place both experimentally and through data analysis. Experimentally secretion assays for different strains are run in tandem. A stock solution of LB media with antibiotics and IPTG is aliquoted into individual cell culture vessels prior to inoculation to ensure concentrations of supplements are uniform.

Cultures are incubated in the same incubator, so that heat and agitation are the same. Protein samples are also prepared in tandem, in an effort to ensure treatment is the same. In terms of analysis while quantification of yield is more informative, relative concentrations in comparison to *ΔCKL* secretor strain are beneficial in terms of reducing variability between experiments, as day to day overall fluxes in productivity are reduced. These steps aim to reduce intra-experimental variation, however despite meticulous repetition of this protocol inter-experimental variation does occur. This resonates with the synthetic biology concept of unpredictability in biological systems (Endy, 2005). In addition to this the use of Western blots is likely to be a big source of variation, as discussed in the previous chapter, lack of linearity and the sheer number of experimental steps account for error. With this in mind it would be preferential to develop a non-Western blot based secretion assay. This will be in the form of an enzyme based secretion assay and will have the added benefit of allowing larger scale testing of constructs because throughput should be reduced; this will be investigated in the next chapter.

While it was hypothesised that the deletion of flagella linked genes may lead to reduced metabolic burden in cells and therefore more metabolic energy available for growth processes, the growth phenotypes of mutant strains were for optical density to be slightly lower than *ΔCKL* when expressing empty secretion construct (Figure 4.), suggesting that the knockout mutations slightly compromised the fitness of cells through pleiotropic effects. Despite this, this effect was very slight it was concluded that none of the strains were compromised enough to consider them unsuitable. In Figure 4. it is evident that expressing CH2 results in a slightly lower optical density of cells at some point in the growth curve of strains, this is expected as the expression of CH2 requires metabolic energy, therefore energy is diverted from cell growth. Interestingly the opposite is true of *ΔclpX.* The optical densities of *ΔclpX* cultures expressing empty vector were consistently lower than other strains, particularly in the stationary phase. However this was not the case when *ΔclpX* was expressing CH2. Optical densities of *ΔclpX* were higher when expressing CH2 (especially in the late-log and stationary phase) as opposed to empty vector (Figure 4.), suggesting that the production of CH2 may stabilise the growth phenotype of *ΔclpX.* This effect may be because in the absence of ClpX, there is an increase in free ClpP. In a wild type strain ClpP would be directed to specific protein degradation targets by ClpX (Kitagawa et al. 2011; Gottesman, 1996). It is possible that in the absence of ClpX, non-specific (or ClpA directed) ClpP proteolysis is more common, leading to a reduction in growth rate in cells (as seen in empty secretion construct expressing cells).

However, when CH2 is expressed, it is possible that this recombinant protein is targeted by ClpP instead of native protein, therefore growth is restored. The only caveat to this is that intracellular CH2 concentration is similar in Δ*clpX* strains in comparison to +*clpX* strains.

The mutations implemented in the strains tested in this chapter (aside from the negative secretor strains) were designed to improve the secretion of CH2 by increasing the number of secretion apparatus. The rationale behind mutants either involved the reduction of metabolic burden by the removal of genes which encode proteins which are now redundant in the *ΔCKL* mutant so that more metabolic energy was available to expend of secretion apparatus production (and to recombinant protein production) or the removal of genes which should result in the upregulation of flagella gene expression and therefore result in more secretion apparatus. These strategies were effective, as CH2 secretion was increased in all mutant strains, however it is beyond the scope of this study to determine which of the two strategies is more effective at increasing FT3SS secretion, however further investigation of improved secretion following the implementation of these strategies and also combinations of these strategies in the next chapter should reveal more information on this.

The implementation of negative control means that it can be confidently stated that the presence of CH2 in the media is a result of FT3SS, however is not possible to confirm whether increased secreted CH2 is due to the incidence of more secretion apparatus or whether individual apparatus are capable of secreting a greater number of CH2 protein subunits. To ascertain this it would be beneficial to image the FT3SS hook basal bodies of strains, this would be possible with a number of microscopy techniques, for example electron microscopy or fluorescence microscopy using fluorescent antibodies which have affinity for the hook protein or primary anti-hook anitbodies. This could also be investigated by cell membrane preparation and purification of hook subunits, which could be visualised by Coomassie or Western blot analysis with an anti-FlgE antibody (many commercially available). Finally at a genetic level qPRC could be utilised to investigate whether *flgE* expression was increased in mutant strains in comparison to Δ*CKL.*

The implementation of the CH2 protein secretion assay utilising the modular synthetic secretion construct has allowed accurate and quantifiable measurements of intracellular and secreted CH2 in all mutant strains, along with the calculation of secretion efficiencies. This reassuringly mirrors the overall effect of secretion observed with the E2 substrate. Each strategy for improved protein secretion through the modified FT3SS resulted in increased

protein secretion capacity and increased protein secretion efficiency; however the yields obtained are still low in comparison to those reported by industry. While the combination of these knockout mutations will be investigated in the next chapter, to see if the increases in secretion capacity will be additive. This will involve the construction and characterisation of many mutant secretor strains. Furthermore alteration will be made to the secretion construct to see if secretion can be improved through this route. While the assay was effective, it does show variability and is throughput limited as the assay is lengthy. Therefore work in the next chapter aimed to harness the modularity of the secretion construct to substitute the cargo protein from CH2 to an enzyme, with the view of developing an enzyme based protein secretion assay. It is hypothesised that this will be more accurate, less variable and more high throughput for the testing of a wider range of strains and constructs.

Chapter 5: Development of a high-throughput assay for FT3SS secretion to screen a multitude of secretion strains and plasmids

5.1. An enzyme based FT3SS secretion assay

The successful development of protein secretion assays in Chapters Three and Four and subsequent preliminary development of secretion competent strains allowed measurement of secretion capacity of a modified FT3SS secretion apparatus. It was shown that the implementation of strain improvements increased secretion capacity of native flagellin variant (E2) protein through the modified FT3SS (Δ*CKL*), with additional (Δ*CKL*) Δ*clpX* (Δ*CKL*) Δ*motAB* and (Δ*CKL*) Δ*flgMN* mutants showing promise as high level secretion strains. Following this it was shown that recombinant protein could also be secreted and that strain improvements resulted in around double the secreted yield of CH2 antibody fragment protein. These data were largely based on use of Western blotting and Coomassie staining to assess protein secretion, both of which have limitations in terms of throughput. They also do not assess the folding-state of the proteins since no attached function or enzymatic activity can be measured.

This chapter therefore focuses on the development of an enzyme based secretion assay to achieve this higher throughput nature of an assay, in order to assess both a large array of strains and constructs but also to build a robust system for future studies in this area. Hence the aim was to express and secrete a recombinant enzyme through the FT3SS and then analyse enzyme activity of the culture supernatant as a proxy for protein secretion and visualisation by other means. A subsequent outcome of such a goal would also be the broadening of substrate range for the FT3SS system and also indicate the ability of proteins to refold in a functional state post-secretion. As functionality of protein is important in biotechnology (i.e. medicines), this folding is required should this secretion system be utilised in the biotechnology industry. It should also enable quick and accurate screening of secretion, allowing a number of strains and plasmids to be tested, with the aim of identifying combinations which allow high levels of FT3SS secretion. The design of the synthetic modular

secretion construct allows simple substitution of the protein cargo –this will be exploited to incorporate an enzyme into the secretion construct -as shown below.

Measurement of enzyme activity in biological systems is ubiquitous in biological research, investigation may concern an organism's natural capacity to produce an enzyme (for example the production of digestive enzymes in filamentous fungi (Beneyton et al., 2016)) or involve production of non-native enzymes to aid experimentation. Examples of this have already been utilised in this work, for example the production of enzymes to confer antibiotic resistance markers (i.e. β-lactamase conferring ampicillin resistance) or as a reporter (i.e. expression of β-galactosidase linked to promoter fragments.). Use of enzyme activity to measure secretion is established in a range of organisms including *Bacillus, Pichia* and *E. coli* (Eom et al. 2005; Degering et al. 2010; Lee et al. 2001; Weis et al. 2004). In some studies the enzyme secreted is of industrial importance, in others it is used solely as a reporter for protein secretion; in this study both are relevant for the enzyme. Other screening assays for secretion capacity have focused on secretion of intrinsically fluorescent protein (i.e. GFP) or protein which is suitable for fluorescent labelling (such as tetracysteine) (Haitjema et al., 2014; DeLisa et al., 2002).

5.2. Choice of enzyme for production within the synthetic modular secretion construct

For this part of the project, an enzyme with robust properties in terms of environmental stability, biochemical characterisation and useful facile enzyme assay was required. One such enzyme was the cutinase from *Fusarium solani,* which was chosen for recombinant production and secretion in the synthetic modular secretion construct because it is well studied, useful in industry, suitable for ester based assays (Liu et al. 2009) and previously expressed in *E. coli.* Cutinase describes a group of hydrolases which occur in plants, fungi and yeast, while their primary role is to degrade plant cutin (a polyester), they are in the hydrolase family and are multifunctional polyesterases, which have also been shown to behave like lipases and therefore catalyse hydrolysis reactions in a range of substrates including triglycerides, water soluble esters and plastics (Ahmed Al-Tammar et al., 2016; Carvalho et al., 1998; Fojan et al., 2000). Cutinase is used in a wide range of industries, including food, agriculture, chemical,

textiles and detergent. Cutinase activity can be measured through biocatalysis reactions of a range of substrates, such as *p*-nitrophenyl butyrate and 4-methylumbelliferyl butyrate or acetate which result in reaction products that can be assayed by production of a fluorescent coloured end-product (Griswold et al., 2003; Yang et al., 2013). The cutinase used in this work is derived from the filamentous fungi *Fusarium solani pisi* which has been well characterised since its discovery in the 1960s and become a model system for the study of cutinases. *F. solani* cutinase is a stable, 197 residue, 22kDa –it is active as a monomer and the active site is at the site of a triad of residues (Ser 120, Asp 175, His 188). The formation of two intramolecular disulphide bridges is important for maintaining catalytic activity (Carvalho et al., 1998). This cutinase was previously utilised in the cotton and synthetic textile industry (Chen et al., 2013). Cutinase has been expressed recombinantly in a range of fungi, yeast and bacteria (Carvalho et al., 1998; Griswold et al., 2003).

5.2.1. Cloning cutinase into the synthetic modular secretion construct

5.2.1.1. Design of synthetic cutinase gene

The *F. solani* cutinase in this study has previously been expressed in *E. coli,* however it was reported that codon optimisation for *E. coli* was essential to initiate secretion of signal peptide tagged-recombinant cutinase into the periplasm, so a similar approach was utilised in this work (Griswold et al., 2003). Furthermore this is a commonly implemented strategy to improve protein expression of recombinant protein (Gustafsson et al., 2004). –therefore codon optimisation was implemented in this work.

The sequence for *F. solani* cutinase was obtained from the NCBI database (accession no. K02640), codon optimised for *E.coli* using the GeneOptimizer® sliding window algorithm and sent for synthesis by GeneArt® Strings. In the design, *Eco*RI and *Pst*I restriction enzyme sites flanked the cutinase gene and consideration was taken to exclude any other restriction enzyme sites that are present in the secretion vector or constructs (*Nde*I, *Xho*I, *EcoR*I, *Pst*I, *Xba*I, *Hind*III, *Bam*HI). Upon arrival DNA was immediately cloned into a blunt end cloning storage vector and stored in NEB 5-alpha. Following colony PCR with primers which annealed to either side of the cloning site, it was confirmed that the cutinase gene was harboured in the storage vector (Figure 5.) and cells were stored at -80°C.

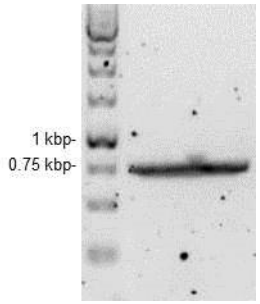


Figure 5.1: Agarose DNA gel showing a PCR product following colony PCR of cells to confirm blunt end cloning of the synthesised cutinase gene into the storage vector.

PCR reactions were carried out using plasmid harbouring *E. coli* colonies as template and primers in Appendix 2 with DreamTaq DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. Lane 1: successful PCR resulted in a 736bp DNA product.

5.2.1.2. Production of the cutinase harbouring secretion construct

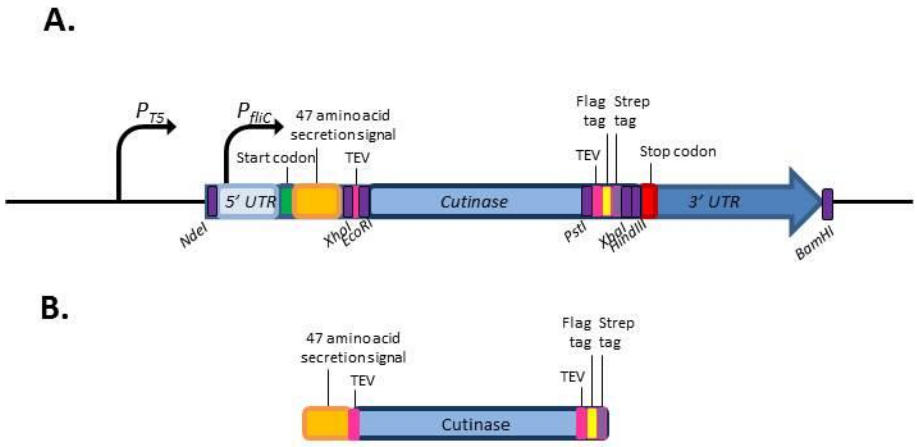


Figure 5.2: Genetic and protein schematic of the secretion construct harbouring cutinase

Following ligation into pJexpress-FliC-empty-FLAG-Strep the cutinase harbouring secretion construct was derived (A), which resulted in the expression of a FT3SS signal peptide cutinase flanked by FLAG and Strep tags and TEV protease sites (B)

Ligation of donor DNA (from pJET 1.2 cutinase) into the pJexpress-FliC-empty-FLAG-Strep acceptor should result in the product seen in Figure 5.. Both donor and acceptor DNA underwent restriction digest with *EcoRI* and *PstI*, and ligation overnight. Following transformation and antibiotic resistance screening, positive colonies were screened by restriction digest. Successful ligation of cutinase into the acceptor vector, will result in the presence of a 617 bp DNA product following restriction digest with *EcoRI* and *PstI*. Positive results were obtained for colony 2 and 6 (Figure 5.), which were then sequenced using primers which anneal to secretion construct DNA flanking the cutinase cargo. The resulting sequences both aligned exactly with the predicted DNA sequence, therefore ligation was successful.

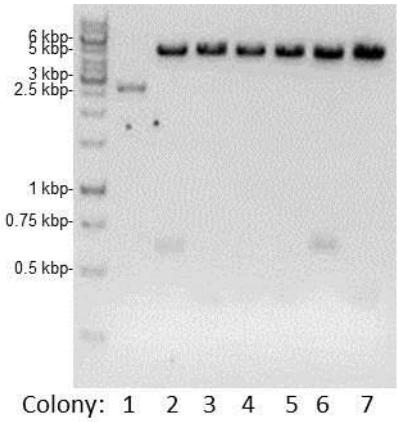


Figure 5.3: Agarose DNA gel showing plasmid from colonies DNA digested with *Eco*RI and *Pst*I.

Plasmid DNA was isolated from positive colonies following ligation and antibiotic screening. Restriction digest reactions were carried out with *Eco*RI and *Pst*I. 5µL restriction digest mixture was prepared for DNA electrophoresis and loaded along with GeneRuler™ 1kb DNA ladder on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide and visualised under UV light and inverted.

Once cloned into the secretion construct the expression and secretion of cutinase through the modified FT3SS was investigated. While secreted cutinase is intended to be utilised for an enzyme based protein secretion assay, as cutinase is harboured in the secretion construct, investigation of cutinase expression and secretion is also enabled by the FLAG-tag. The concentration of both intracellular and secreted cutinase will be measured using Western

blotting and antibody detection with αFLAG-HRP; in addition an assay will be developed to measure the amount of secreted cutinase by enzyme activity.

5.2.2. Secretion of the cutinase harbouring synthetic modular secretion construct through the modified FT3SS

Initial investigation of cutinase expression and secretion was carried out in the original Δ*CKL* secretor strain. Following cell culture and IPTG induction of Δ*CKL* freshly transformed with either empty or cutinase harbouring secretion construct; FLAG-tagged protein was visible in both the intracellular and supernatant (secreted) fractions of cells expressing cutinase (Figure 5.). The size of the cutinase-secretion construct protein was predicted to be 30.6 kDa, this corresponds to the size of the protein identified by the αFLAG-HRP antibody. A lower concentration of cutinase protein was observed in the secreted fraction and there was no non-specific binding of the FLAG antibody to protein in cells expressing cutinase or empty secretion construct. In the accompanying Coomassie stained SDS-PAGE gel (not shown), the secreted fraction was clear of all protein, demonstrating that lysis was not prevalent in cell culture, but also that secreted cutinase is present at a concentration below that detectable by Coomassie stain.

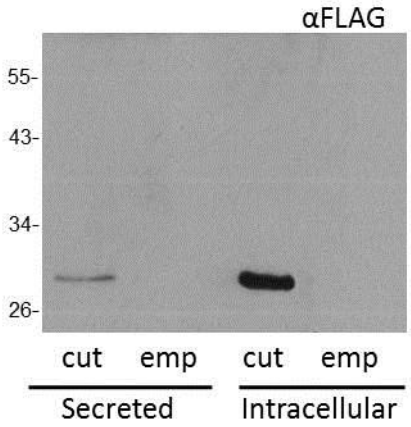


Figure 5.4: Cutinase protein in the secreted and intracellular fractions of Δ*CKL* secretor strain cell culture.

E.coli Δ*CKL* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pJexpress-empty vector were induced with 0.05mM IPTG and grown in LB at 37°C with 180rpm agitation. Cells were harvested when cell culture reached OD_{600} 1.0. 1 OD unit of cells (intracellular) were prepared for SDS-PAGE and the supernatant (secreted) from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western blot analysis with αFLAG-HRP antibody. Protein samples were loaded as follows. Lane: Secreted -(1) cutinase, (2) empty vector. Intracellular -(3) cutinase, (4) empty vector.

5.3. Development of a standardised protein secretion assay to quantify cutinase secretion

Following the confirmation that cutinase was adequately expressed and secreted in the Δ*CKL* strain, efforts were made to develop an enzyme based secretion assay.

5.3.1. Trail development of a tributyrin and pNPB protein secretion assay

It was reported that cutinase activity could be measured by the breakdown of tributyrin substrates. Methods describe that organisms secreting cutinase can be plated onto agar plates supplemented with tributyrin and that zones of clearing will develop around colonies which secrete cutinase (Kwon et al., 2009). Despite efforts to optimise this assay no zones of clearing

were evident. However, *F. solani* cutinase (either from *F. solani* or recombinantly expressed in bacteria) hydrolytic activity has been reported for *p*-nitrophenyl (*p*NP)-esters, where cutinase cleaves the substrate to yield a 4-nitrophenol molecule, which presents as a yellow colour detectable by absorbance at 405nm (Liu et al. 2009; Chen et al. 2008). Attempts to develop a *p*NP-butyrate assay resulted in positive results in terms of change in absorbance (405nm) in the presence of the supernatant of cell culture secreting cutinase, however results were inconsistent and optimisation proved difficult, therefore this assay was abandoned and an alternative substrate chosen based on an analogous ability to cleave ester type bonds between a reporter substrate and a 4-carbon butyrate chain- in this case, 4-methylumbelliferyl butyrate (MUB).

5.3.2. Development of a MUB based cutinase protein secretion assay

Cutinase catalyses the cleavage of MUB to yield a fluorescent 4-methylumbelliferone (4-MU) molecule (Yang et al., 2013). Fluorescence is visible under UV light excitation visually and qualitatively and detectable in a quantitative manner using a plate reader. The principle of this assay relies on the fact that the amount of MUB cleaved and therefore 4-methylumbelliferone released can be measured using fluorescence and detected. An assay was developed to measure the abundance of cutinase in the secreted fraction of cell culture, as a factor of fluorescence. This involved the preparation of cell culture supernatant, addition of substrate, incubation and finally measurement of fluorescent output. It was hoped that this assay would provide a means of increasing the throughput and accuracy of the assessment of relative protein secretion through the modified FT3SS. Assay development was important, as it is imperative that the assay was both reliable and accurate and also able to detect a wide range of concentrations of cutinase protein. Consideration was also made to the ease of experimental set up and robustness, as this will aid the high-throughput nature of the assay and therefore enable screening the secretion capacity of many strains for future workers. To ensure that results from fluorescent cutinase protein secretion assays were representative of the concentration of cutinase located in the supernatant of cell cultures, Western blot analysis was also carried out in tandem. The assay was developed with reference to the Sigma-Aldrich product guidelines for the substrate and by following literature and laboratory expertise with methylumbelliferyl-sialic acids.

5.3.3. Preliminary MUB assay: active secreted cutinase

Prior to assay optimisation it was first established that the recombinant cutinase could effectively cleave MUB and that this was detectable by fluorescence. First LB and supernatant from cells expressing either empty or cutinase harbouring secretion construct was added to MUB dissolved in Tris Buffer, pH 8. Reaction mixes were loaded into wells of a black walled, clear bottom 96 well plate to allow UV excitation through the clear bottom, without cross excitation or emission from adjacent wells. Samples were collected from the same cell culture analysed by Western blot in Figure 5., therefore the presence (and absence) of cutinase in culture supernatant was confirmed. Following incubation at room temperature and visualisation by image capture under UV light (approximately 302nm), it was evident that the addition of supernatant which contained cutinase resulted in the cleavage of MUB and therefore production of a fluorescent emission when excited with UV light (Figure 5.). Conversely this was not seen following the addition of LB media or the supernatant of cells which expressed empty secretion construct. This demonstrates that the assay is effective in reporting the presence of FT3SS secreted cutinase in the crude supernatant of cells. It also shows that it was possible to secrete a recombinant protein through the FT3SS and that it was able to fold correctly in the supernatant to result in an active enzyme.

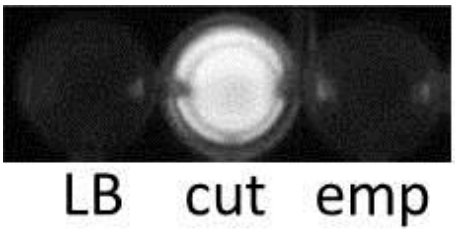


Figure 5.5: Fluorescence emission of reaction mixtures containing MUB and either LB media or cell culture supernatant.

E.coli Δ*CKL* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pJexpress-empty vector were induced with 0.05mM IPTG and grown in LB at 37°C with 180rpm agitation. Cells were harvested when cell culture reached OD_{600} 1.0. The supernatant was obtained by centrifugation. Either LB media or cell supernatant was added to 500mM MUB in Tris, pH 8 and added to a 96 well plate. Following incubation the black walled clear bottom, 96 well plate was imaged with a G:BOX during excitation with UV light. Well: (1) LB + MUB, (2), supernatant (cutinase) + MUB, (3) supernatant (empty) + MUB

5.3.4. Development of the MUB assay: early considerations

Once the assay had been shown to be effective in this preliminary context, some early considerations to experimental procedure were made. Initial thought was given to the preparation of supernatant, the ratio of substrate to enzyme and the adequate buffering of solutions. Substrate comprised of buffered MUB and 'enzyme' (secreted protein in LB media and buffer). In all optimisation experiments enzyme containing supernatant was freshly derived from cultures of Δ*CKL* cells expressing either cutinase harbouring or empty secretion construct; supernatant was obtained by centrifugation. This was as a proxy for pure enzyme, as it was important experiments were optimised in the context of cell supernatants. Dissolving MUB in buffer solution proved difficult, but was facilitated by the addition of Triton X-100 and (dimethylformamide) DMF (Vaneechoutte et al., 1988). It was desirable for substrate to be in excess so that the reaction was not rate limited by its availability. Additionally as it was observed that LB emitted background fluorescence following excitation at 365 nm, preliminary investigation found that results were more satisfactory when the concentration of LB (supernatant) in reaction mixtures was low. With these considerations and preliminary evidence, a ratio of 20:80 supernatant (equivalent to enzyme): buffered substrate solution was used throughout –it was also confirmed experimentally that this was sufficient to effectively buffer reaction mixes, to the given pH. Finally following some investigation it was established that adequate separation of cells from supernatant was possible with centrifugation at 13,000rpm for 15 minutes and that steps such as sterile filtration of supernatant was not necessary. With a basic assay framework in place steps were taken to optimise the assay for accurate measurement of fluorescent output as a factor of cutinase cleavage of MUB. All work was carried out in black 96 well plates to avoid cross- excitation and emission from adjoining plates.

5.3.5. Development of the MUB assay: buffer pH

The recommended pH-dependent excitation wavelength reported by Sigma-Aldrich for both MUB and 4-MU is 365nm. Therefore this was used during assay development. The pH of the enzyme-substrate reaction mix required some optimisation, as the optimum pH for cutinase and MUB, do not necessarily concord. According to the literature, *F. solani* cutinase is active

from pH 2-12, however maximal enzymatic activity occurs at pH 8.5 for a range of substrates (Petersen et al., 2001; Ni & Chen, 2009). Sigma report excitation and emission data for MUB for Tris pH 8, however the literature reports fluorescence of negative controls is high when MUB is buffered at pH 8, presumably due to non-enzymatic cleavage of MUB, whereas at pH 5 less background fluorescence is observed (Vaneechoutte et al., 1988). High fluorescence of negative controls (i.e. LB + MUB or supernatant of cells secreting empty vector) when incubated in pH 8 buffer was also observed in preliminary investigation during this work, therefore the remainder of experiments were carried out in pH 5 buffered solution. The optimal temperature for native cutinase directed substrate cleavage is dependent on the substrate but ranges from 30 to 60°C; however the literature does not report an optimum temperature for *F. solani* activity with MUB substrate (Chen et al., 2008). Experiments were carried out at 30°C, as at this temperature the reaction proceeded fast enough to measure but not too fast to overflow the plate reader detector during the experiments. Quenching the reaction with the addition of a high pH buffer is a standard procedure for MU-based substrates as the fluorescence output of the product 4-MU under certain excitation emissions (365nm included), is pH dependent and maximal in the range of pH 10 upwards (Fink & Koehler, 1970; Zhi et al. 2013).

5.3.6. Development of the MUB assay: substrate concentration

Another factor of importance is that the concentration of substrate remains high enough throughout the course of the experiment to not limit the rate of reaction either in terms of affinity of binding to the enzyme but also that it does not become exhausted in solution resulting in false reduction in activity if measured over longer time periods. Equally as the plate reader is only sensitive within a range of fluorescent intensities it is important that the output of experiments does not exceed these limits –this is particularly relevant to the upper limits of detection, as it is hoped that this assay will be utilised to measure increased cutinase secretion and therefore increased AU readings. A range of substrate concentrations were investigated (500, 100, 50, 10, 5 µM) and LB or supernatant from cell culture of cutinase or empty secretion construct harbouring Δ*CKL* cells was added. A MUB only control was also included. Mixtures were left overnight to allow more time for cutinase to catalyse 4-MU production and to

investigate whether MUB self cleaves over a prolonged period of time, note that this exceeds the time of later experiments. Excitation was carried out at 365nm, and emission was scanned so that the peak emission could be established for future work. This was essential as it is reported that the emission wavelength of 4-MU varies in different solutions (e.g. LB v water), therefore the manufacturer's guidelines could not be relied on (Fior et al., 2009). 500µM MUB resulted in emission intensities higher than the detection limit of the plate reader. At 5 and 10µM low substrate concentration appeared to be rate limiting on the reaction as low fluorescence output was measured. At 50µM and 100µM fluorescence readings for samples incubated with cutinase were much higher than without and while LB and empty secretion construct resulted in higher fluorescence than MUB alone, it was very low (Figure 5.). 446nm resulted in the highest output and was selected as the emission wavelength for future measurements of fluorescence. Readings at 446nm are situated at the peak of the emission spectra curve and therefore result in larger differences between results for different reaction mixes; this will enable investigation into the differences between 4-MU concentrations (and therefore cutinase concentration) in experiments. The difference between MUB incubated with or without cutinase is larger for 100µM MUB than 50 µM (60833 compared to 41481 at 446nm), therefore it was concluded that this concentration was the most advisable to use as differences between reaction mixtures were larger, but all within the range of detection by the plate reader.

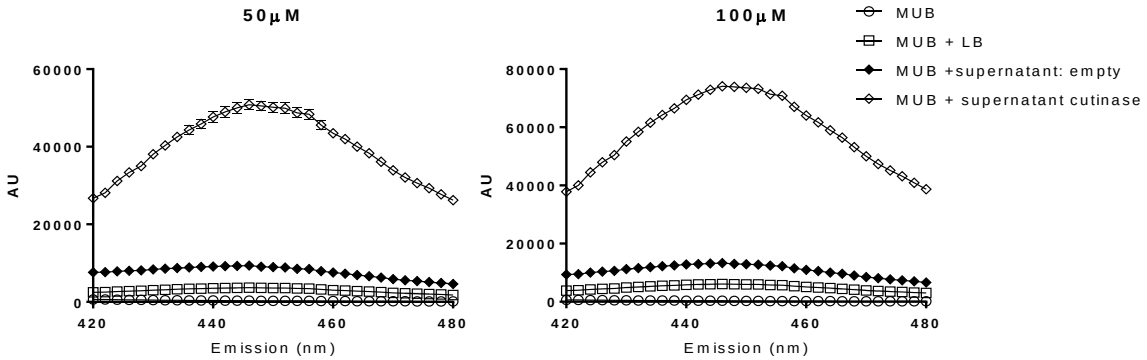


Figure 5.6: Emission spectra for pH 5 buffered 50µM and 100µM MUB with LB media or cell culture supernatant from cells expressing cutinase or empty FT3SS signal peptide tagged secretion constructs.

E.coli Δ*CKL* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pJexpress-empty vector were induced with 0.05mM IPTG and harvested as in Figure 5.. Either LB media or cell supernatant (40µL) was added to 160µL pH 5 buffered, 50µM or 100 µM MUB solution and added to a 96 well plate. A MUB only and MUB + LB reaction were also run to observe fluorescence of all components. Following incubation at 30°C overnight, the reaction mix was added to quenching buffer (1:2 ratio) fluorescence was measured from 420 to 480nm following excitation at 365nm. Results from one biological replicate, with three technical repeats. Standard error of the mean displayed.

5.3.7. Development of the MUB assay: time and sensitivity to a range of concentrations of cutinase

Length of an enzyme catalysed reaction is an important consideration. Reactions undergo a period of linear rate of reaction followed by a plateau where the rate of reaction slows as substrate runs out. This plateau can be linked a reduction in either substrate (due to product formation) or enzyme (due to degradation). Fluorescence output is more accurately linked to 4-MU and therefore cutinase concentration, when the rate of reaction is linear- i.e. the initial phase; therefore it is important to calculate when this occurs in a given enzyme-substrate reaction. This was investigated by measuring the fluorescence output of reactions over time. In addition to investigation of the appropriate reaction length, the effect of different

concentrations of enzyme containing supernatant was also investigated to establish if this linearly reduced the amount of product produced. This is of particular relevance of the aim of this assay is to measure varying concentrations of cutinase supernatant. To mimic this, the cutinase containing supernatant was diluted 2 fold with empty secretion construct supernatant.

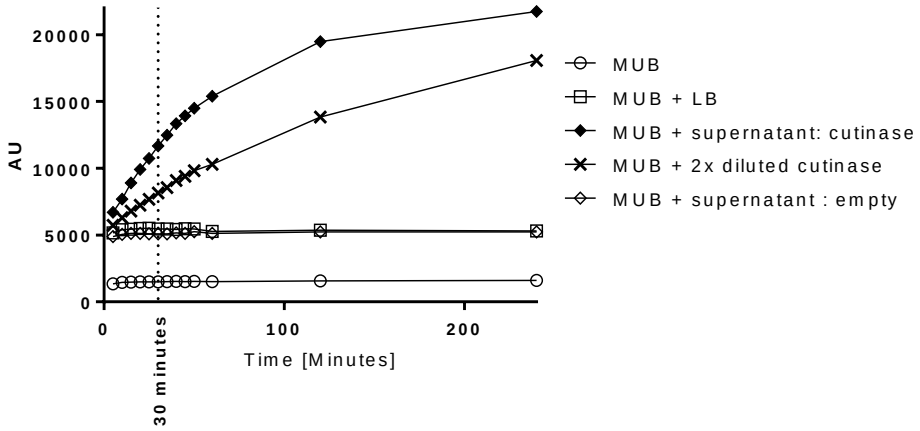


Figure 5.7: Fluorescence output (AU) of reaction mixtures over time.

E.coli Δ*CKL* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pJexpress-empty vector were induced with 0.05mM IPTG and grown and prepared as in Figure 5.. Either LB media or cell supernatant (40µL) was added to 160µL 100µM MUB solution (0.05M phosphate citrate buffer, pH 5) and added to a 96 well plate. A MUB only reaction was also run to observe fluorescence of all components. Measurements were taken every 5 minutes for the first 60 minutes and then at 120 and 240 minutes during incubation at 30°C. Fluorescence was measured at 446nm following excitation at 365nm.

Fluorescence output of reaction mixtures was measured every 5 minutes from 5 to 50 minutes and then at 60, 120 and 240 minutes (Figure 5.). LB and supernatant do contribute to fluorescence, however it is evident that in the absence of cutinase the emission from reaction mixtures remains consistent throughout the experiment (Figure 5.: MUB, MUB + LB, MUB + empty). This is indicative that MUB is not subject to breakdown to form 4-MU in the absence of cutinase. Furthermore LB and the supernatant of cells void of cutinase does not facilitate the production of 4-MU during two hours of incubation, instead the emission observed is intrinsic to the solution. This observation gives confidence to the finding that cutinase is

actively cleaving MUB to yield fluorescent 4-MU, as increased fluorescence output of reaction mixtures over time is only evident in the presence of cutinase. Therefore cutinase catalysed MUB cleavage is the route for 4-MU derivation and therefore fluorescence. Furthermore the assay is sensitive to different concentrations of cutinase in reaction mixtures, as incubation of MUB with diluted cutinase supernatant resulted in linearly roughly 2-fold lower fluorescence readings, this shows that the assay is sensitive enough to detect different concentrations of secreted cutinase. The rate of fluorescence output for mixtures containing cutinase was linear from 5 to 40 minutes with the higher concentration of cutinase and 5 to 50 minutes when twice diluted. Following this the rate of reaction reduced, although fluorescence continued to increase up to 240 minutes. 30 minutes would be the optimum time to take fluorescence readings, as it is evident that the rate of reaction is linear for a range of enzyme concentrations, with substrate concentration not rate limiting at this stage. However in reality this was rarely effective -an example is given below.

A small scale investigation was carried out to confirm negative regulation of cutinase secretion and also measure cutinase secretion in two of the strains which exhibited increased concentration of secreted E2 and CH2 proteins, to see if the MUB protein secretion assay could detect these differences. Δ*CKL,* Δ*CKL* Δ*flgMN* and Δ*CKL* Δ*clpX* and the negative secretion strains Δ*flhDC and* Δ*CKL* Δ*flgDE* were transformed with either empty or cutinase harbouring secretion construct, supplemented with the appropriate concentration of IPTG to induce plasmid expression and incubated at 37°C, with agitation. When cultures reached OD_{600} 1.0 samples were then taken and prepared the for the MUB protein secretion assay. Following centrifugation and the addition of LB to account normalise for the variation in OD_{600} measured in different cell cultures, supernatant was mixed with 100µM MUB and incubated for 30 minutes at 30°C, the reaction was then quenched with sodium carbonate pH 10.5 and fluorescence was measured at 365nm excitation, 446nm (Figure 5.). Following 30 minutes of incubation fluorescence readings for reaction mixtures containing potential cutinase containing supernatants were often only marginally higher that the results obtained for those without (for example Δ*CKL* + empty secretion construct or either value for the negative secretion strain Δ*flhDC*). In addition the differences in fluorescence reading for reaction mixes which presumably had variable cutinase concentrations were both small, as the aim is to screen a large number of strains and secretion constructs by this method it would be preferential for more variation to be seen between strains. If reaction mixtures were incubated for a longer period of time the differences would be more pronounced (as seen in

Figure 5.), however the rate of reaction would not be linear, and therefore results would not be as accurate. Instead the MUB assay was modified slightly, so that the excitation wavelength was in the UV range –this was based on the observed fluorescence following UV excitation in preliminary MUB-cutinase investigation (Figure 5.). Fluorescence was high in the well containing supernatant with cutinase and absent in the wells without.

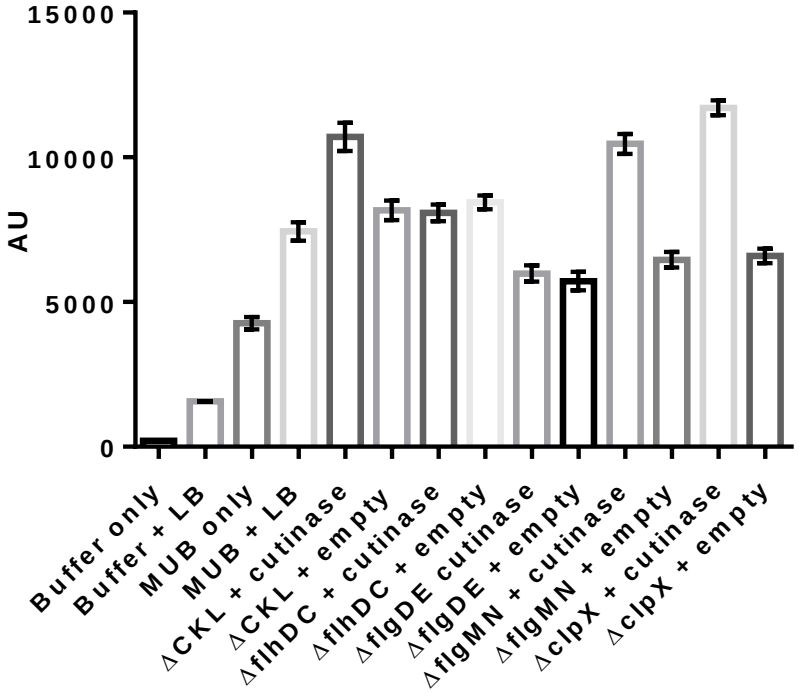


Figure 5.8: Fluorescence output (AU) of reaction mixtures following incubation of MUB with the secreted fraction of *ΔCKL, ΔflhDC, ΔCKL ΔflgDE, ΔCKL ΔflgMN* or *ΔCKL ΔclpX* cells secreting empty or cutinase harbouring secretion construct. Excitation 365nm, Emission 446nm.

E.coli Δ*CKL, ΔflhDC, ΔCKL ΔflgDE, ΔCKL ΔflgMN* or *ΔCKL ΔclpX* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep (cutinase) or pJexpress-FliC-empty-FLAG-Strep (empty) were induced with 0.05mM IPTG (or 1mM for *ΔflhDC* and *ΔflgDE*) and grown in LB at 37°C with 180rpm agitation. Cells were harvested when cell culture reached OD_{600} 1.0. The supernatant was obtained by centrifugation. Either LB media or cell supernatant (40µL) was added to 160µL 100µM MUB solution (0.05M phosphate citrate buffer, pH 5) and added to a 96 well plate. Citrate buffer, buffer + LB, 100µM MUB and MUB + LB reaction mixes were also included to provide information on the AU of all components of the reaction. Following incubation at 30°C

for 30 minutes, the reaction mix was added to quenching buffer (1:2 ratio) and fluorescence was measured at 446nm following excitation at 365nm. Results from one biological replicate, with three technical repeats. Standard error of the mean displayed. Two-way ANOVA (variables: strain and plasmid) p= <0.0001.

5.3.8. Development of the MUB assay: excitation in the UV region of the spectrum

One final consideration concerned the fact that the image in Figure 5. was taken by a camera attached to a standard UV-transilluminator-digital camera (UV-filter attached) (G:Box). Therefore the possibility of using both a transilluminator based assay and potentially a plate reader assay using UV excitation wavelengths was explored as this may have given better results than the traditional 365nm excitation. The UV-transilluminator (Syngene) used has an excitation wavelength of 302nm. Therefore fluorescence was measured following excitation at 302nm – UV excitation occurs from the bottom up in the plate reader, therefore this required the use of black walled, clear bottom 96 well plates, to allow measurement of fluorescence for each well following UV excitation, without interference from adjoining wells. Excitation at 302nm is also pH dependent, however at excitations below 320nm 4-MU fluorophore emission is highest at pH 5-6, quenching is not required, as reaction mixtures are already buffered to pH 5. The only consideration that was checked was the appropriate emission wavelength, again because there is evidence that MU has differing emission properties in different solutions (Fior et al., 2009). Supernatant of Δ*CKL* cells from the cultures expressing cutinase or empty secretion construct were incubated with 100µM MUB for 30 minutes at 30°C prior to an emission scan following excitation at 302nm (**Error! Reference source not found.** 5.9). At all emission wavelengths fluorescence was highest for MUB incubated with cutinase containing supernatant. This resulted in maximal emission at 446nm, as before -this was also a wavelength at which the contribution of LB to the fluorescence output was reduced (i.e. MUB + LB has a negative linear distribution, however at 446nm it is relatively low, in comparison to 420nm for example). Finally, the MUB only sample emitted very little fluorescence at any emission wavelength. This suggests that LB not MUB contributes to background fluorescence. At 302nm excitation and 446nm emission the difference between MUB + cutinase and MUB + empty (the next highest) was 3648 units. This is larger than the difference seen between these two reaction mixes when excited at 365nm in Figure 5., which amounted to 3236.

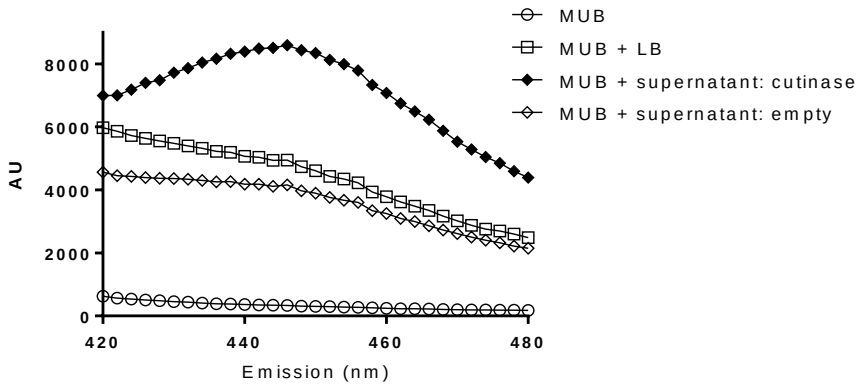


Figure 5.8: Emission spectra for pH 5 buffered MUB with LB media or cell culture supernatant from cells expressing cutinase or empty FT3SS signal peptide tagged secretion constructs- excitation at 302nm.

E.coli Δ*CKL* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pJexpress-empty vector were induced with 0.05mM IPTG and grown and prepared for the MUB protein secretion assay as in Figure 5.. A MUB only and MUB + LB reaction was also run to observe fluorescence of all components. Following incubation at 30°C for 30 minutes, fluorescence was measured from 420 to 480nm following excitation at 302nm. Results from one biological replicate, with two technical repeats. Standard error of the mean displayed.

Efforts were then made to assess whether an image based MUB protein secretion assay could be utilised, where MUB and cell culture was incubated and then a photo taken as it Figure 5.. It was hoped that densitometry could be used to quantify fluorescence, however shadowing in wells was an issue and the sensitivity was not high enough to distinguish between different fluorescent outputs of different mutant strains -though the difference between control and experimental samples for secretion was always clear- as in Figure 5.. This method would be a useful for screening mutants (from a mutant library for example), to ascertain whether mutants exhibited secretion of cutinase occurred or not. The results from one of these experiments is shown in **Appendix 5**.

Carrying out the MUB assay in a plate reader with fluorescence excitation at 302nm, should result in more pronounced differences between readings for reaction mixtures when cutinase is present or absent. Additionally as quenching is not required, the assay is simplified, thus

improving the high throughput nature further. Investigation into the performance of this plate reader MUB protein secretion assay was then tested using the same biological samples used in Figure 5.. In addition intracellular fractions and secreted fraction of cell cultures were prepared for Western blot analysis, to provide information on the concentration of intracellular protein and also to give confidence that the relative changes in fluorescence reported by MUB assay did relate to secreted cutinase concentration.

The results of Western blot analysis show that cutinase was expressed in all strains, although it was quite low in the Δ*flgDE* mutant, despite the addition of extra IPTG (Figure 5.). Secreted cutinase was visible in the supernatant of Δ*CKL,* around twice as much was visible in the secreted fractions of Δ*flgMN* and Δ*clpX,* whereas cutinase was absent from the supernatant of Δ*flhDC and* Δ*flgDE* cells. This confirms that excitation at 365nm produced unreliable results as Figure 5. suggest that Δ*flgMN* cells secreted less cutinase than Δ*CKL* and that Δ*clpX* secreted marginally more.

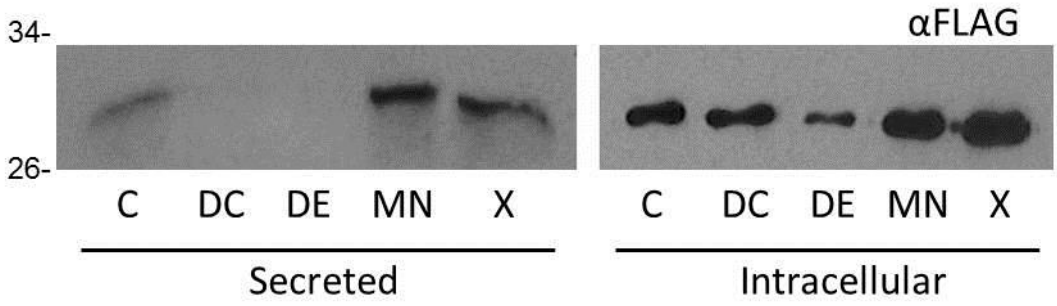


Figure 5.10: Intracellular and secreted cutinase from the secretor strain and *ΔflhDC, ΔflgDE, ΔflgMN* and *ΔclpX* and mutants.

E.coli ΔCKL (C), *ΔflhDC* (DC), *ΔCKL ΔflgDE* (DE), *ΔCKL ΔflgMN* (MN) or *ΔCKL ΔclpX* (X) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB supplemented with 0.05mM IPTG (or 1mM for DC and DE). Cells were prepared for Western blot analysis with anti-FLAG-HRP antibody as in Figure 5.. Samples were loaded to as follows: Supernatant: 20µL. Cells: 2µL

Following centrifugation and the addition of LB to normalise for the variation in OD_{600} measured in different cell cultures, supernatant was mixed with 100µM MUB and incubated.

When reaction mixes are excited at 302nm again the highest fluorescence readings are derived for reaction mixes containing the supernatant of cutinase expressing Δ*CKL*, *ΔflgMN* and *ΔclpX* cells (Figure 5.). It is evident that the fluorescence readings for all other reaction mixes containing cell culture supernatant or LB are uniformly lower (and not statistically different from each other (Two-way ANOVA: Tukey's multiple comparison test). It is also evident that background fluorescence is due to LB alone (MUB only low, buffer + LB and MUB + LB equally high). This is not a limitation as the concentration of LB (or supernatant) is constant in reaction mixes. In terms of results for MUB incubated with supernatant containing cutinase the highest was for *ΔclpX* (6788), followed by *ΔflgMN* (5748) and 4712 for *ΔCKL*. This assay was sensitive enough to differentiate differences between different concentrations of secreted cutinase; all cutinase containing supernatant reactions resulted in fluorescence values which were statistically different to each other ($p = <0.0001$. Two way ANOVA: Tukey's multiple comparison test). Finally fluorescence output of negative control strains expressing cutinase, was statistically lower than those measured for *ΔCKL* –demonstrating that negative control is established and that cutinase in the supernatant arises by directed secretion through the FT3SS alone.

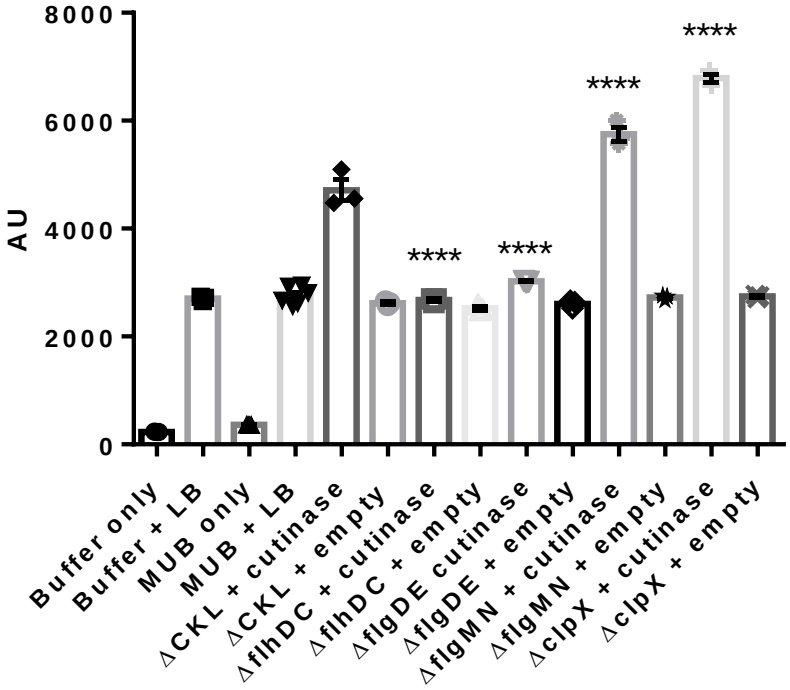


Figure 5.11: Fluorescence output (AU) of reaction mixtures following incubation of MUB with the secreted fraction of *ΔCKL, ΔflhDC, ΔCKL ΔflgDE, ΔCKL ΔflgMN* or *ΔCKL ΔclpX* cells secreting empty or cutinase harbouring secretion construct. Excitation 302nm, Emission 446nm.

E.coli Δ*CKL, ΔflhDC, ΔCKL ΔflgDE, ΔCKL ΔflgMN* or *ΔCKL ΔclpX* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep (cutinase) or pJexpress-FliC-empty-FLAG-Strep (empty) were induced with 0.05mM IPTG (or 1mM for *ΔflhDC* and *ΔflgDE*) and grown and prepared for the MUB protein secretion assay as in Figure 5.. Citrate buffer, buffer + LB, 100µM MUB and MUB + LB reaction mixes were also included to provide information on the AU of all components of the reaction. Following incubation at 30°C for 30 minutes, fluorescence was measured at 446nm following excitation at 302nm. Results from one biological replicate, with three technical repeats. Standard error of the mean displayed. Two-way ANOVA (variables: strain and plasmid) and Tukey's multiple comparison test: **** = $p < 0.0001$.

These results reflected the results of the Western blot, *ΔCKL* was found to have the lowest concentration of secreted cutinase for both experiments. *ΔflgMN* and *ΔclpX* should display twice the cutinase activity of *ΔCKL* as around twice as much cutinase was visible in the Western blot. For this purpose the average MUB + LB background was removed to give a clearer representation of the effect of cutinase on fluorescence output; this showed a 2.06 fold increase in AU for *ΔclpX* and a 1.53 fold increase for *ΔflgMN.* For *ΔclpX* this is concordant with the Western blot, it is a little lower for *ΔflgMN,* however still representative of the fact that removing *flgMN* from *ΔCKL* resulted in improved cutinase secretion.

The establishment of a reliable, high-throughput MUB assay to measure cutinase secretion through the modified FT3SS allowed investigation of the secretion output of a number of strains in one quick and accurate experiment. This can now be utilised to screen a multitude of potential strain and plasmid based improvements. Throughout all of the results chapters, either Western blot analysis or the MUB protein secretion assay have demonstrated that gene knockout mutation based strain improvements can result in increased secretion of FT3SS secretion signal peptide tagged protein through the FT3SS. It was hoped that the combination of these knock out mutations which had previously been found to improve secretion, would result in an additive effect on secretion capacity.

5.4. Combination of strain improvement strategies for increased FT3SS capacity

The existing gene knockouts which had been found to result in increased secretion output through Western blot analysis of secreted E2, CH2 or cutinase or by MUB based cutinase secretion assay, were combined in the hope that these effects would be incremental. These were *fliDST, motAB, flgMN* and *clpX*. As strains with antibiotic linked gene knockouts or DNA templates of FRT flanked antibiotic resistance cassettes (ABCs) for gene knockout implementation are already established, it was possible to generate combination mutants with reasonable ease. The aim was to generate strains which incorporated multiple gene knockouts, to assess the effect of the consolidation of two or three strategies at once.

5.4.1. Mutant generation: multiple gene knockout strain improvements

5.4.1.1. Knockout mutagenesis via the Lambda Red recombinase method

The strains (Δ*CKL*) Δ*clpX* Δ*motAB,* Δ*clpX* Δ*flgMN,* Δ*clpX* Δ*fliDST,* Δ*motAB* Δ*flgMN* and Δ*motAB* Δ*fliDST* were all generated using this method. In Chapter 3 linear DNA was derived by PCR which comprised of FRT flanked antibiotic resistance genes -with homologous ends to the chromosomal region either side of the target gene for knockout mutagenesis. These linear DNA fragments were electroporated into the appropriate parent strain expressing Lambda Red recombinase enzymes (Table 5.). Successful recombination events resulted in linear DNA template being inserted into chromosomal DNA at homologous regions. Successful recombination was screened for by plating on agar plates supplemented with the relevant antibiotic (chloramphenicol or kanamycin). Positive colonies were observed for all recombination strains and following the curing of the Red recombinase plasmid so that no further recombination events occur; the deletion of chromosomal genes was confirmed with PCR (Figure 5.).

Table 5.1: Parent strain and the additional gene knockout targets, with PCR products for confirmation of knockout mutagenesis in the various parent strains.

Parent strains already harboured antibiotic resistance, due to prior gene knockout mutagenesis by FRT flanked ABC cassette homologous recombination. Additional gene knockouts were successfully implemented with different ABCs in the following strains. Following mutagenesis colony PCR was carried out on successful transformants. The size of PCR products which denote a positive result for each new mutant is given, along with the size of the PCR product (if relevant) expected if knockout mutagenesis was not successful.

Parent Strain: MC1000 Δ*CKL* Δ*flgKL*	Resistance	Additional knockout mutation	Resistance and template plasmid	PCR conformation primers (forward, reverse)	Size of PCR product if knockout:	
					Successful	Unsuccessful
Δ*clpX*	Km	Δ*motAB*	Cm, pKD3	motAB F, motAB R	1198	1913
Δ*clpX*	Km	Δ*flgMN*	Cm, pKD3	flgMN F and flgMN R	1198	813
Δ*clpX*	Km	Δ*fliDST*	Cm, pKD32	fliDST F, C1	264	N/A
Δ*motAB*	Cm	Δ*flgMN*	Km, pKD4	flgMN confo, kt	1338	N/A
Δ*motAB*	Cm	Δ*fliDST*	Km, pKD13	fliDST confo and k2	993	N/A

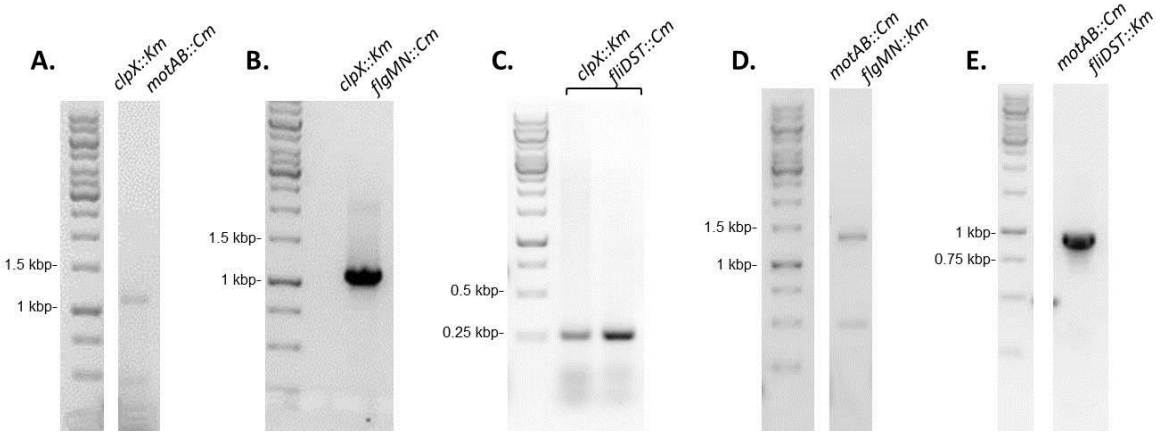


Figure 5.12: Agarose DNA gels of PCR products to confirm successful knockout mutagenesis.

PCR reactions were carried out using the templates and primers in Appendix 2 with DreamTaq DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of products of the sizes listed in Table 5.. This confirmed the successful knockout mutagenesis and therefore the generation of the following strains: A. *clpX::Km motAB::Cm,* B. *clpX::Km flgMN::Cm,* C. *clpX::Km fliDST::Cm,* D. *motAB::Cm flgMN::Km,* E. *motAB::Cm fliDST::Km*

5.4.1.2. Knockout mutagenesis via the P1 phage transduction method:

While Lambda-Red recombineering is the preferred method for knockout mutagenesis (due to the absence of phage), the efficiency is poor and therefore it can take multiple attempts of the protocol to generate mutants- when necessary P1 phage transduction was carried out to achieve knockout mutants in a short space of time. This resulted in the successful generation of (Δ*CKL*) Δ*flgMN* Δ*fliDST,* Δ*clpX* Δ*motAB* Δ*flgMN* and Δ*clpX* Δ*motAB* Δ*fliDST* from the parent strains and P1 phage lysates listed in Table 5.. As the number of ABCs which could be linked to gene knockout DNA constructs was limited to kanamycin or chloramphenicol, FRT linked antibiotic resistance was sometimes removed from the chromosome using the Flp recombinase gene harbouring pCP20 plasmid. This resulted in the excision of the ABC leaving an 82 or 85 bp FRT scar site for chloramphenicol or kanamycin cassettes respectively. This method was implemented in all three of the parent strains listed in 5.Table 5..

Table 5.2: Parent strain and associated knockout mutations implemented by P1 phage transduction with chromosomal DNA packaged from donor strains.

Parent strain. Δ*fliC* Δ*flgKL:*	P1 transduced mutation	Donor strain. Δ*fliC* Δ*flgKL:*
Δ*flgMN*	Δ*fliDST:: Km*	Δ*fliDST:: Km*
Δ*motAB* Δ*flgMN*	Δ*clpX::Km*	Δ*clpX::Km*
Δ*motAB* Δ*fliDST*	Δ*clpX::Km*	Δ*clpX::Km*

Once successful transformatants were obtained cells were passaged on sodium citrate plates, to eradicate phage. Growth was always normal in these strains, suggesting that phage was eliminated from these strains. Care was taken to ensure that there was no possibility of existing genes knockouts in the parent strain being reinstated by chromosomal DNA packaged by the P1 phage during homologous recombination. The chromosomal location of genes targeted for knockout mutagenesis are listed in Table 5.. P1 phage is capable of packaging and initiating the homologous recombination of up to 100,000bp of donor chromosomal DNA into a parent strain. Therefore if the gene selected for knockout mutagenesis by P1 phage was within 100,000bp of the site of any existing gene knockouts in the donor strain, PCR conformation would be required to ensure this gene was not restored by P1 phage transduction in the newly derived strain. The site of Δ*clpX* (and therefore Δ*clpX::Km*) is distant from all other target genes. *fliDST* (and Δ*fliDST::Km*) is within 100,000bp of *fliC* therefore there would be potential for restoration of the Δ*fliC* mutation, however as donor DNA was derived from a Δ*fliC* strain, this is not possible. Therefore PCR conformation of retention of knockout mutagenesis genotype was not required.

Table 5.3: Chromosomal location (bp) of genes which underwent knockout mutagenesis and existing gene knockouts.

Gene	Chromosomal location	
	Start	**End**
Δ*fliC*	2053401	2054069
Δ*flgKL*	1140598	1140943
Δ*motAB*	2024187	2025997
Δ*flgMN*	1131700	1132404
Δ*fliDST*	2054069	2056275
Δ*clpX*	412775	414049

It was hoped that either Lambda-Red recombination or P1 phage transduction could be utilised to also produce Δ*fliC* Δ*flgKL* Δ*motAB* Δ*flgMN* Δ*fliDST* and Δ*fliC* Δ*flgKL* Δ*motAB* Δ*flgMN* Δ*fliDST* Δ*clpX* however attempts were unsuccessful. Following either mutagenesis protocol positive colonies were not detected following antibiotic screening. It is not through that these combinations of mutations should result in lethality; therefore it is believed that with more time these strains could be generated. Despite absence of the full complement of combinations of knockout mutant strategies, a plethora of combinations were derived. These were characterised by growth phenotype and FT3SS secretion capacity using the MUB plate reader cutinase protein secretion assay. For ease abbreviated versions of these strains will be used in figures and in text throughout this section (Table 5.).

Table 5.4: Strains tested for expression and secretion of cutinase through the modified FT3SS.

All strains are in the MC1000 background. Abbreviations are often shorted in figure labels to exclude the '*Δ'* from the notation. e.g. *ΔfliC ΔflgKL ΔfliDST ΔflgMN* or *ΔDST ΔMN* or DST MN

Strain (MC1000)	**Abbreviation**
ΔfliC ΔflgKL	*ΔCKL*
ΔfliC ΔflgKL ΔclpX	*ΔX*
ΔfliC ΔflgKL ΔfliDST	*ΔDST*
ΔfliC ΔflgKL ΔflgMN	*ΔMN*
ΔfliC ΔflgKL ΔmotAB	*ΔM*
ΔfliC ΔflgKL ΔflgMN ΔfliDST	*ΔMN ΔDST*
ΔfliC ΔflgKL ΔmotAB ΔfliDST	*ΔM ΔDST*
ΔfliC ΔflgKL ΔmotAB ΔflgMN	*ΔM ΔMN*
ΔfliC ΔflgKL ΔclpX ΔfliDST	*ΔX ΔDST*
ΔfliC ΔflgKL ΔclpX ΔflgMN	*ΔX ΔMN*
ΔfliC ΔflgKL ΔclpX ΔmotAB	*ΔX ΔM*
ΔfliC ΔflgKL ΔclpX ΔmotAB ΔfliDST	*ΔX ΔM ΔDST*
ΔfliC ΔflgKL ΔclpX ΔmotAB ΔflgMN	*ΔX ΔM ΔDST*
ΔflhDC	*ΔDC*

5.4.2. Combined strategies for improved FT3SS secretion: Effects of multiple gene deletions on growth phenotype

As previously described, gene deletions may lead to reduced metabolic burden and therefore increased cell growth. Equally gene deletion may lead to strains which are not viable or less healthy if essential genes or processes are affected. In addition the production of recombinant protein (i.e. cutinase) will add to metabolic burden and may lead to a reduction in growth rate, for example it was found that the growth phenotypes of mutant strains expressing empty or CH2 harbouring secretion construct grew slightly slower or to a lower final optical density than Δ*CKL* in the majority of strains (Figure 4.). Additionally when expressing recombinant protein, optical densities of all strains (with the exception of Δ*clpX*) were slightly lower at some point in the growth curve in comparison to when the same strains were expressing empty secretion construct (Figure 4.). As the gene knockouts investigated here were combined, it was important to observe whether they have an additive effect on growth phenotype or whether extensive mutagenesis had resulted in strains which were non-viable. To evaluate the effect of gene deletion and cutinase expression on growth phenotype, growth curves were obtained for *ΔCKL* and the newly derived mutant strains expressing either empty or cutinase harbouring secretion construct.

Figure 5. shows that the growth phenotype of all strains was fairly similar and all strains were viable. In the log phase lower optical density readings were found for all strains in comparison to Δ*CKL* whether cells were expressing empty or cutinase containing secretion construct. With the exception of Δ*X* Δ*M* when expressing cutinase harbouring secretion construct (Figure 5.B), Δ*M* Δ*DST* had lower optical densities than all other strains when expressing either plasmid (Figure 5.A and B).

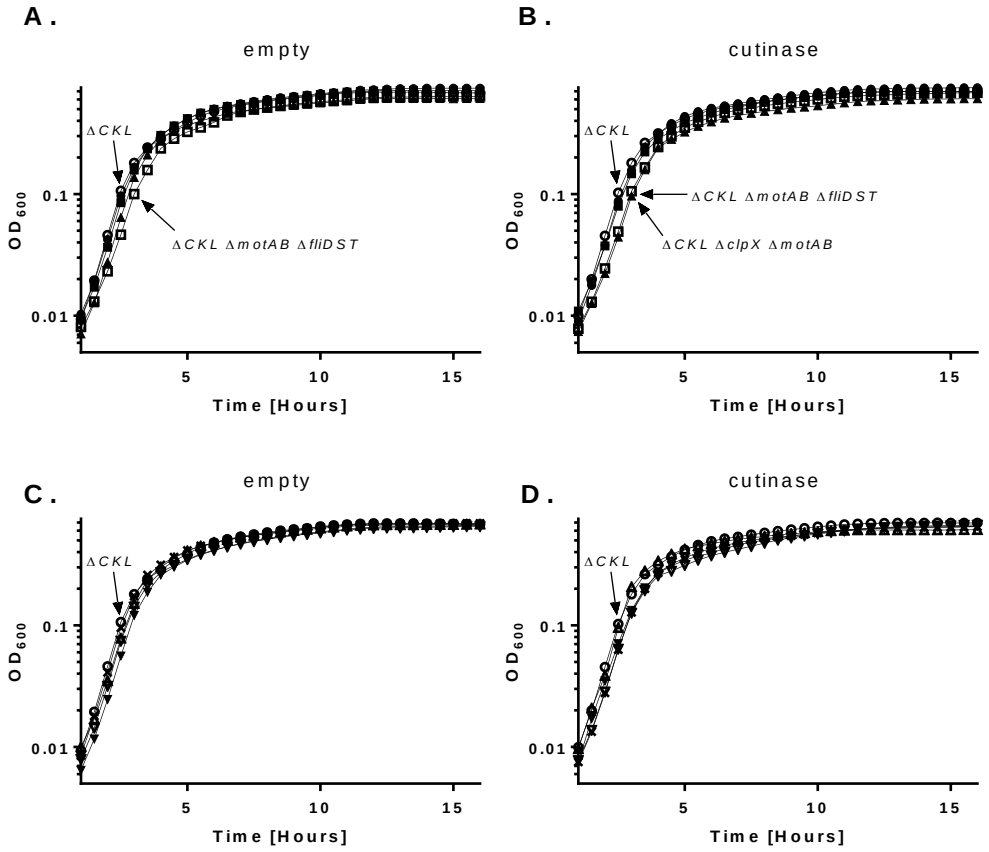


Figure 5.13: Growth of Δ*CKL* and (Δ*CKL*) Δ*MN* Δ*DST,* Δ*M* Δ*MN,* Δ*M* Δ*DST,* Δ*X* Δ*M,* Δ*X* Δ*MN,* Δ*X* Δ*DST,* Δ*X* Δ*M* Δ*MN* and Δ*X* Δ*M* Δ*DST* mutant strains expressing either empty or cutinase harbouring secretion construct over time in a 96 well plate. Calibrated to LB media.

Starter cultures of cells expressing either pJexpress-FliC-empty-FLAG-Strep or pJexpress-FliC-cutinase-FLAG-Strep were grown in LB supplemented with ampicillin and 0.05mM IPTG and aliquoted into 96 well plate wells. The 96 well plate was incubated in the TECAN plate reader at 37°C with 6mm orbital shaking. OD_{600} measurements were recorded every 30 minutes. Results split into four graphs for ease of viewing - Δ*CKL* with relevant plasmid on all graphs to allow comparison. A and C: Δ*CKL*: ○, Δ*flgMN* Δ*fliDST:* ●, Δ*motAB* Δ*flgMN:* ■, Δ*motAB* Δ*fliDST:* □, Δ*clpX* Δ*motAB:* ▲. B and D: Δ*CKL*: ○, Δ*clpX* Δ*flgMN:* △, Δ*clpX* Δ*fliDST:* ▼, Δ*clpX* Δ*motAB* Δ*flgMN:* ▽ and Δ*clpX* Δ*motAB* Δ*fliDST:* X. Results represent the combination of two biological repeats, with three technical replicates of each biological.

Viewing that data for the growth curve of each strain individually allows observations to be made on the effect of recombinant protein expression on cell growth (Figure 5.). The differences in OD_{600} for Δ*MN* Δ*DST* expressing either plasmid were alike. Δ*X* Δ*M* grew to slightly higher and Δ*X* Δ*MN* slightly lower optical densities in the late log phase when expressing empty secretion construct. Δ*X* Δ*DST,* Δ*X* Δ*M* Δ*MN* and Δ*X* Δ*M* Δ*DST* all displayed higher OD_{600} readings throughout the log and early stationary phase when not expressing cutinase, whereas Δ*M* Δ*MN* and Δ*M* Δ*DST* had slightly higher optical densities in the log phase when expressing cutinase. In summary the differences between strain growth phenotype when expressing empty or cutinase containing secretion construct was small and all strains were shown to be viable.

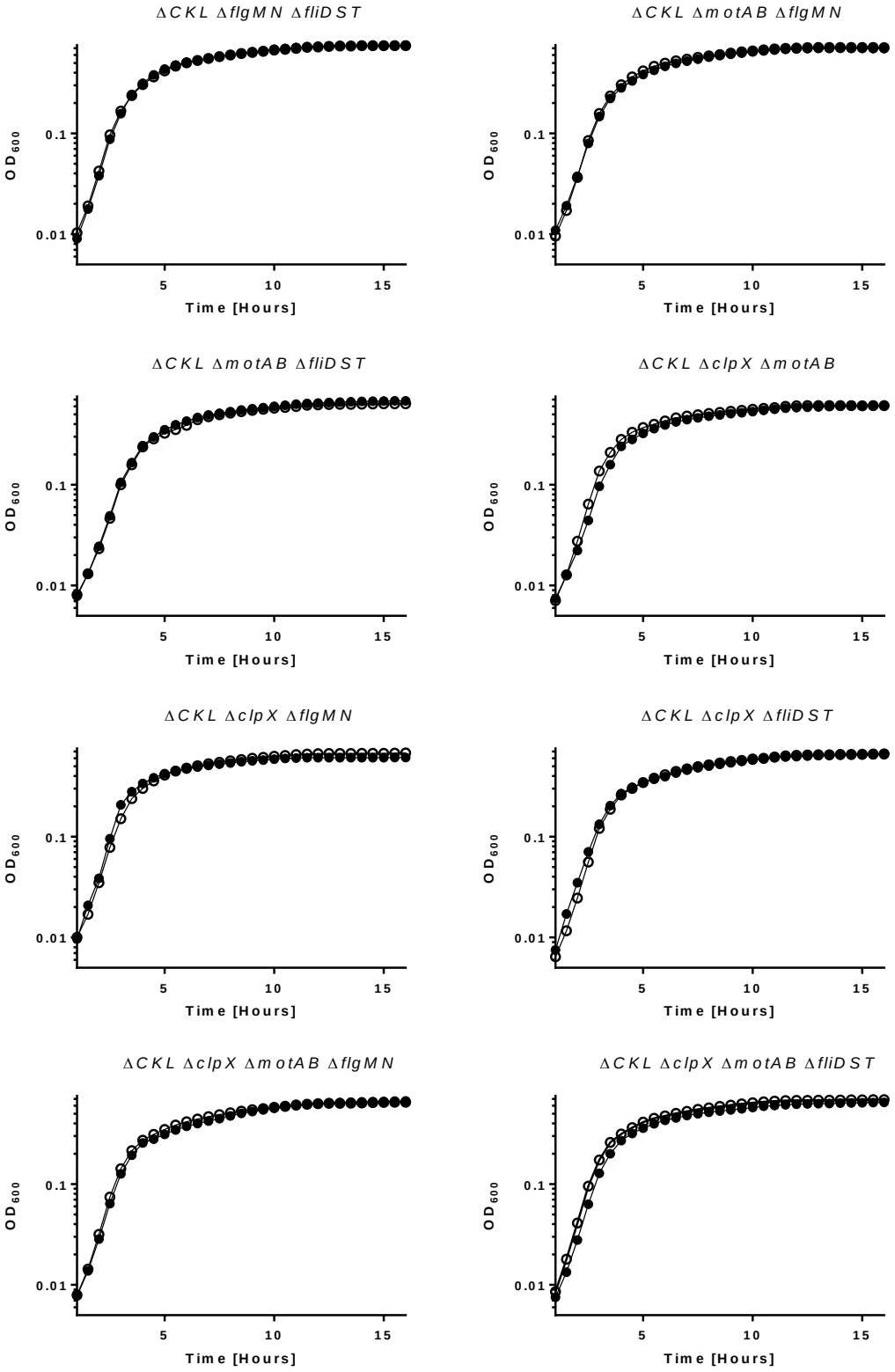

ΔCKL ΔflgMN ΔfliDST
ΔCKL ΔmotAB ΔflgMN
ΔCKL ΔmotAB ΔfliDST
ΔCKL ΔclpX ΔmotAB
ΔCKL ΔclpX ΔflgMN
ΔCKL ΔclpX ΔfliDST
ΔCKL ΔclpX ΔmotAB ΔflgMN
ΔCKL ΔclpX ΔmotAB ΔfliDST
OD600
0.1
0.01
5
10
15
Time [Hours]

Figure 5.14: Growth of Δ*CKL* and (Δ*CKL*) Δ*MN* Δ*DST,* Δ*M* Δ*MN,* Δ*M* Δ*DST,* Δ*X* Δ*M,* Δ*X* Δ*MN,* Δ*X* Δ*DST,* Δ*X* Δ*M* Δ*MN* and Δ*X* Δ*M* Δ*DST* mutant strains expressing either empty (no fill) or cutinase (black) harbouring secretion construct over time in a 96 well plate. Calibrated to LB media.

Starter cultures of cells expressing either pJexpress-FliC-empty-FLAG-Strep (white fill) or pJexpress-FliC-cutinase-FLAG-Strep (black fill) were grown in LB supplemented with ampicillin and 0.05mM IPTG and aliquoted into 96 well plate wells. The 96 well plate was incubated in the TECAN plate reader at 37°C with 6mm orbital shaking. OD_{600} measurements were recorded every 30 minutes. Results represent the combination of two biological repeats, with three technical replicates of each biological.

No combination of mutant knockouts resulted in strains which grew poorly, therefore all strains had potential to be viable secretion platform strains. The FT3SS secretion phenotype was investigated in all strains to investigate whether secretion capacity was improved.

5.4.3. Combining strategies for improved FT3SS secretion: Effects of multiple gene deletions on FT3SS secretion of cutinase

With a high-throughput assay established for the measurement of directed secretion through the modified FT3SS, the suite of mutant strains which had been generated could be quickly characterised in terms of secretion capacity. While the cutinase secretion assay is effective at measuring secreted cutinase, it is not able to determine the concentration of intracellular cutinase. Previously information on intracellular and secreted protein have been combined to calculate the secretion efficiency of strains, therefore to give an idea of this, the cell pellets from cultures were prepared for Western blot analysis. While this did limit the throughput somewhat, it was worthwhile, furthermore it was the preparation of the supernatant fraction for SDS-PAGE, which was most time consuming and therefore throughput limiting. All 13 Δ*CKL* background mutant strains which had been generated were tested, along with Δ*flhDC* to ensure negative control of secretion was established. Cells cultures were grown in tandem and

following harvesting, samples were prepared for either SDS-PAGE or Coomassie staining or Western blotting (cell pellets) or MUB protein secretion assay (supernatant).

The intracellular protein fractions of cell cultures were run on SDS-PAGE gels. Following Coomassie staining (Figure 5. top) it was evident that all fractions look similar in terms of the pattern and relative concentration of bands. The exception to this was a band between 17 and 26kDa which was overexpressed in *ΔX ΔM, ΔX ΔDST* and *ΔM ΔDST.* This is too small to be full length cutinase in the secretion construct as this is expected (and has been shown by Western blot) to be 30.6kDa, furthermore Western blot analysis does not detect this protein (not shown as image cropped for ease of viewing). Western blot analysis shows that the concentration of intracellular cutinase is fairly uniform in all strains (Figure 5. bottom), including those which displayed overexpression of the smaller protein.

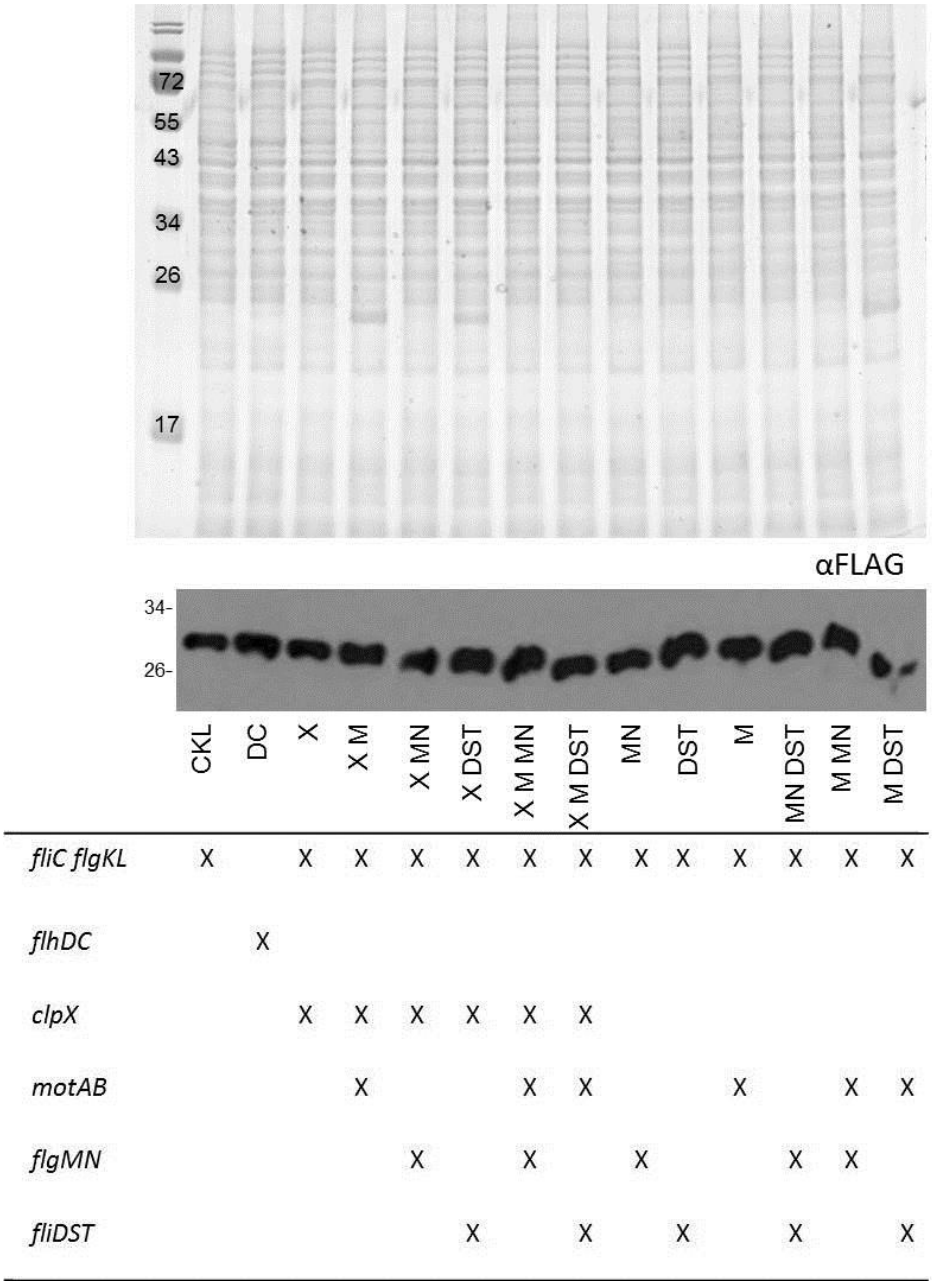


Figure 5.15: Intracellular cutinase from Δ*CKL* and all additional mutant strains.

The *E.coli* strains listed in Table 5. containing the plasmid pJexpress-FliC-cutinase-FLAG-Strep were grown in LB supplemented 0.05mM IPTG (or 1mM for *ΔflhDC*). Cells were harvested at OD_{600} 1.0. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain (top) or Western blot (bottom) analysis of cells and supernatant using an anti-FLAG-HRP antibody. Samples were loaded as follows: 2µL (Western blot), 5µL (Coomassie stain). One biological replicate of each strain shown.

Densitometry analysis was carried out following three biological replicates of each strain to quantify the concentration of intracellular cutinase in each strain relative to *ΔCKL* from Western blot (Figure 5.). Although no statistically significant differences in intracellular levels were noted, the results suggest that the trend is for the removal of genes (in addition to *ΔCKL*) to result in expression of a higher concentration of intracellular cutinase in comparison to *ΔCKL* alone –the only exception to this was *ΔX ΔM ΔDST.* The highest individual intracellular cutinase concentration was observed in a *ΔM ΔDST* strain and was 5.88 fold higher than that measured for *ΔCKL* in that experiment. The second highest was a 5.68 fold increase on *ΔCKL* for a *ΔmotAB* strain. *ΔmotAB* also exhibited the highest average fold increase (3.42 higher than that seen in *ΔCKL*).

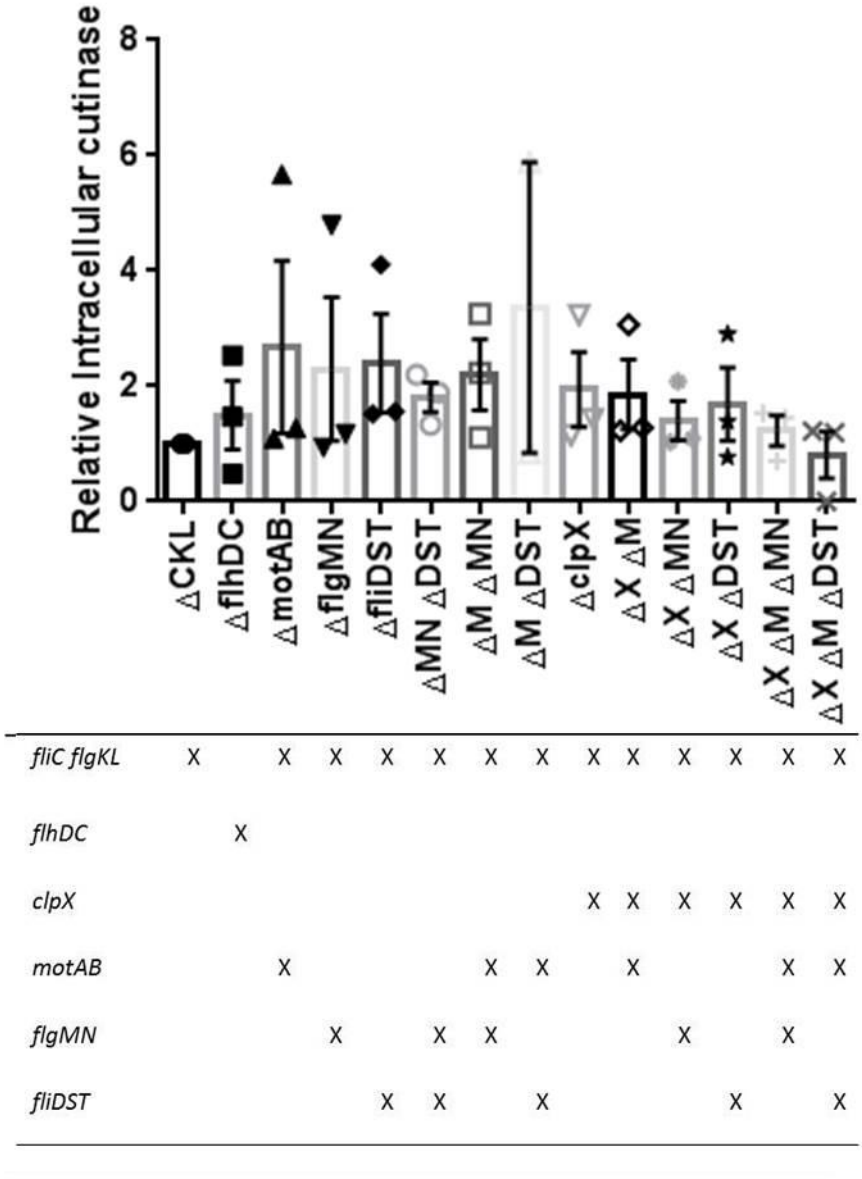


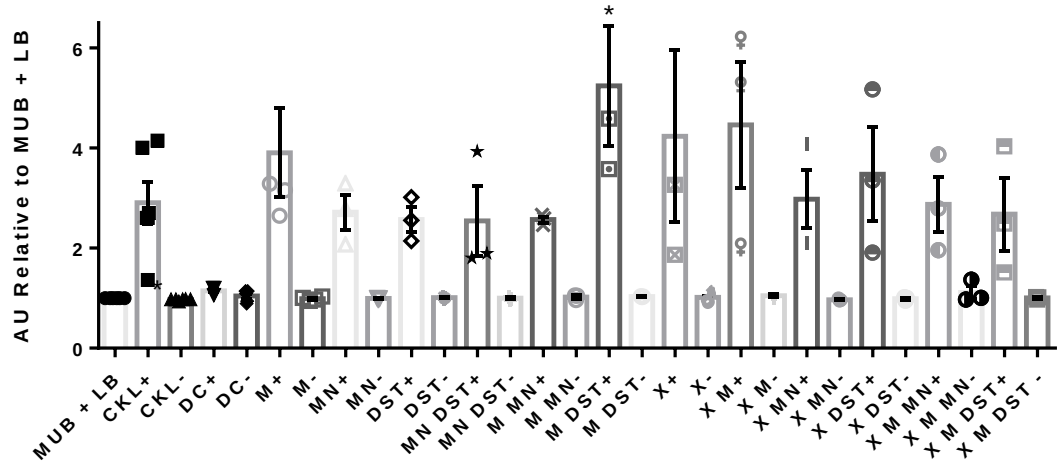
Figure 5.16: Western blot densitometry analysis of intracellular protein fractions of cutinase expressing *ΔCKL, ΔflhDC* and *ΔCKL* strains with additional mutations

The *E.coli* strains listed in Table 5. containing the plasmid pJexpress-FliC-cutinase-FLAG-Strep were grown in LB supplemented 0.05mM IPTG (or 1mM for *ΔflhDC*). Cells were harvested at OD_{600} 1.0. Following Western blot analysis of cells and supernatant using anti-FLAG-HRP, densitometry analysis was carried out using ImageJ. Results were then normalised to the value obtained for intracellular cutinase in the *ΔCKL* strain. Individual data points, average and standard error of the mean are given. Three biological replicates were carried out for each strain.

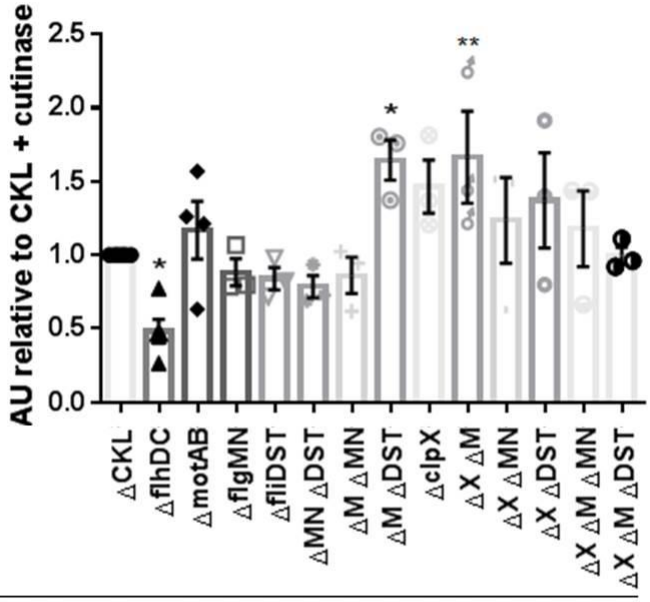
To measure the concentration of secreted cutinase, the supernatant from these cell cultures (Figure 5.) and cell cultures of all strains secreting empty-secretion construct was prepared for the cutinase protein secretion MUB assay. Following incubation of culture supernatant with MUB, the fluorescent output measured following excitation at 302nm can be related to the presence of 4-MU, therefore the amount of cutinase cleaved MUB and therefore the amount of cutinase present. Data was handled in a number of ways to allow normalisation and therefore combination of results from different experiments and also to remove background fluorescence. This involved calculating AU relative to that measured for MUB + LB in each replicate (Figure 5.9A) or for cutinase expressing cell cultures, following adjustment for MUB + LB, the relative AU value of strains compared to that seen for Δ*CKL* expressing cutinase (Figure 5.9B).

Figure 5.9A shows that once normalised to MUB + LB the AU values derived for MUB incubated with either the supernatant of strains expressing empty-secretion construct or Δ*flhDC* expressing cutinase were all similar to both each other and MUB + LB itself. This demonstrates that once background fluorescence had been accounted for, fluorescence emission was low in the absence of cutinase. On average Δ*MN,* Δ*DST,* Δ*MN* Δ*DST,* Δ*M* Δ*MN,* Δ*X* Δ*M* Δ*MN* and Δ*X* Δ*M* Δ*DST* all had lower emission readings than Δ*CKL* (2.90 fold higher than MUB + LB)*,* suggesting less cutinase was present in the supernatant of these cells. The AU of Δ*X* Δ*MN* was very similar to Δ*CKL,* Δ*X* Δ*DST* was the next highest, followed by Δ*M.* The three highest emission readings were measured in Δ*X,* Δ*X* Δ*M* and Δ*M* Δ*DST* –with the latter being found to be significantly higher (5.24 fold higher) than that measured for Δ*CKL* expressing cutinase (Tukey's multiple comparison test). A slightly different pattern of results was seen when these results were then normalised relative to the value derived for Δ*CKL* + cutinase in each set of experiments (Figure 5.9B). Variation was reduced and the results on a one-way ANOVA were found to be significant. Δ*flhDC* resulted in a significantly lower fluorescent output than Δ*CKL.* Δ*MN,* Δ*DST,* Δ*MN* Δ*DST,* Δ*M* Δ*MN* and Δ*X* Δ*M* Δ*DST* again had lower average fluorescent values that Δ*CKL.* However as opposed to in Figure 5.9A, with this calculation, Δ*X* Δ*M* Δ*MN* fluorescence was higher than Δ*CKL.* This demonstrates the effect data handling can have on the interpretation of results. All other strains had higher emission outputs than Δ*CKL,* the highest two averages were derived from Δ*M* Δ*DST* (1.64 fold higher) and Δ*X* Δ*M* (1.67 fold higher), which were found to be significant.

A.



B.



fliC flgKL	X		X	X	X	X	X	X	X	X	X	X	X	X
flhDC		X												
clpX									X	X	X	X	X	X
motAB			X				X	X		X			X	X
flgMN				X		X	X				X		X	
fliDST					X	X		X				X		X

Figure 5.9: Fluorescence output (AU) of reaction mixtures following incubation of MUB with the secreted fraction of Δ*CKL* and all Δ*CKL* background mutant strains secreting empty (-) or cutinase (+) harbouring secretion construct.

The *E. coli* strains listed in Table 5. containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep (+) or pJexpress-FliC-empty- FLAG -Strep (-) were induced with 0.05mM IPTG (or 1mM for *ΔflhDC*) and grown and prepared for the MUB protein secretion assay as in Figure 5.. MUB + LB reaction mixes were also included to provide information on the AU of components which provide background fluorescence. Following incubation at 30°C for 30 minutes fluorescence was measured at 446nm following excitation at 302nm. Results from three biological replicates, with three technical repeats for each. Individual data points, average and standard error of the mean are given. To allow comparison of results between experiments, results were either (A) normalised to the average value for MUB + LB in each experiment, (B) values calculated in 'A' normalised to the average value for Δ*CKL* cells secreting cutinase. Standard error of the mean displayed. Two-way ANOVA of A: not significant. One-way ANOVA of B: p = <0.0001. Tukey's multiple comparison test of cutinase expressing strain compared to cutinase expressing Δ*CKL* (*: p = <0.05, **: p = < 0.01)

As Δ*motAB* Δ*fliDST* and Δ*clpX* Δ*motAB* were found to result in consistently high statistically significant fluorescent outputs whichever way the data was handled these two strains became candidate high secretion platform strains. Fresh transformed cells from these strains were grown in tandem with Δ*CKL* and prepared for Western blot analysis of supernatant, to confirm that the high fluorescence output observed was related to increased secreted cutinase concentration. From the Coomassie stain it is again evident that the protein at around 23kDa is overexpressed in the Δ*clpX* Δ*motAB* strain, however it is not evident in Δ*motAB* Δ*fliDST* as it was in Figure 5.. The supernatant is free of cellular contaminants suggesting that cell lysis is not prevalent in any of the strains. Western blot analysis of cell culture fractions with an antibody for the cytoplasmic chaperone protein GroEL did reveal that cell lysis or cytoplasmic membrane leakage was increased in Δ*motAB* Δ*fliDST* and Δ*clpX* Δ*motAB* (Figure 5.11), however as this antibody is very sensitive and the Coomassie stained SDS PAGE gel does not suggest cell lysis occurred at a high rate, it can be assumed that the majority of cutinase found in the supernatant of cell culture has been secreted through the FT3SS. In terms of secreted cutinase, achieving a satisfactory result was difficult; in one biological replicate (not shown) cutinase was

only seen in the secreted fraction of Δ*clpX* Δ*motAB* cells. In the other biological replicate it was necessary to overexpose the X-ray film for a prolonged period of time to achieve a result where cutinase could be seen in the secreted fraction of all strains (Figure 5.10 top). By eye it appears that Δ*CKL* secretes the least, Δ*motAB* Δ*fliDST* secreted slightly more and Δ*clpX* Δ*motAB* secreted around four times as much cutinase. The Western blot shows that the concentration of cutinase in the intracellular fractions is very similar for all strains. Although the background exposure is high, densitometry analysis was carried out and found there to be 1.02 and 0.84 fold more cutinase in the cells of Δ*clpX* Δ*motAB* and Δ*motAB* Δ*fliDST* respectively. Whereas Δ*clpX* Δ*motAB* was shown to have secreted 3.58 times the amount of cutinase that Δ*CKL* did, which looks accurate, however densitometry analysis calculated that Δ*motAB* Δ*fliDST* secreted 2.2 times the concentration Δ*CKL* did, which does seem higher than what the image suggests.

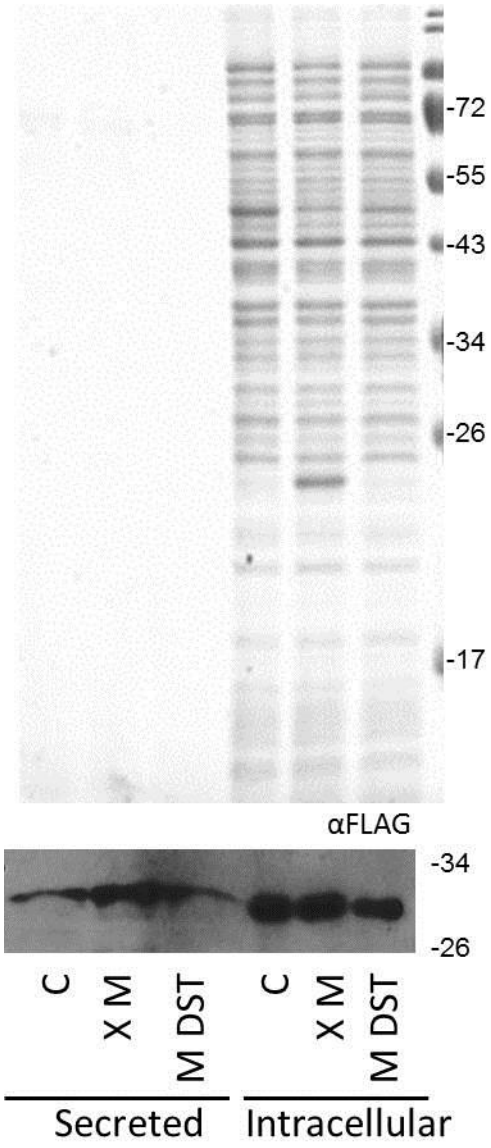


Figure 5.10: Intracellular and secreted cutinase from Δ*CKL* and the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* mutants

E.coli ΔCKL (C), *ΔCKL ΔclpX ΔmotAB* (X M) or *ΔCKL ΔmotAB ΔfliDST* (M DST) containing the plasmid pJexpress-FliC-cutinase-FLAG-Strep was grown in LB supplemented with 0.05 IPTG. Cells were prepared for SDS-PAGE as in Figure 5. and were either stained with Instant Blue Coomassie stain or Western blot analysis of cells and supernatant using an anti-FLAG-HRP antibody. Samples were loaded to as follows: Supernatant: 20µL. Cells: 2µL (Western blot), 5µL (Coomassie stain).

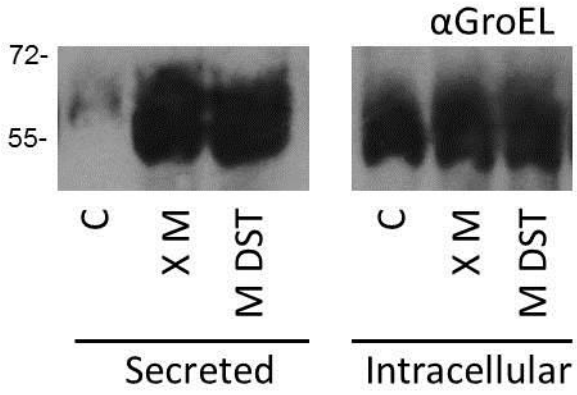


Figure 5.11: Abundance of GroEL in the intracellular and supernatant fractions of Δ*CKL* and the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* mutants, following expression and secretion of cutinase protein through the FT3SS.

E.coli ΔCKL (C), *ΔCKL ΔclpX ΔmotAB* (X M) or *ΔCKL ΔmotAB ΔfliDST* (M DST) containing the plasmid pJexpress-FliC-cutinase-FLAG-Strep was grown in LB. Cells were harvested at OD_{600} 1.0. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using anti-GroEL antibody and an HRP secondary. Samples were loaded as follows: Supernatant: 20µL, cells: 2µL.

In addition to analysis by Western blot (Figure 5.10) supernatant was also used in a MUB assay, along with the supernatant from the strains expressing empty secretion construct. If results are concordant this will validate the results of the Western and MUB assay. Normalisation to MUB + LB eliminated any background fluorescence in the assay. Overall the fluorescent output of cells expressing empty secretion construct was lower than that observed in MUB + LB for all strains (Figure 5.). *ΔCKL* exhibited cutinase activity, however the fluorescent output measured for the reaction mixtures incubated with *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* was significantly higher than *ΔCKL.* This amounted to a 5.35 or 4.74 fold increase in fluorescent output when normalised to *ΔCKL.*

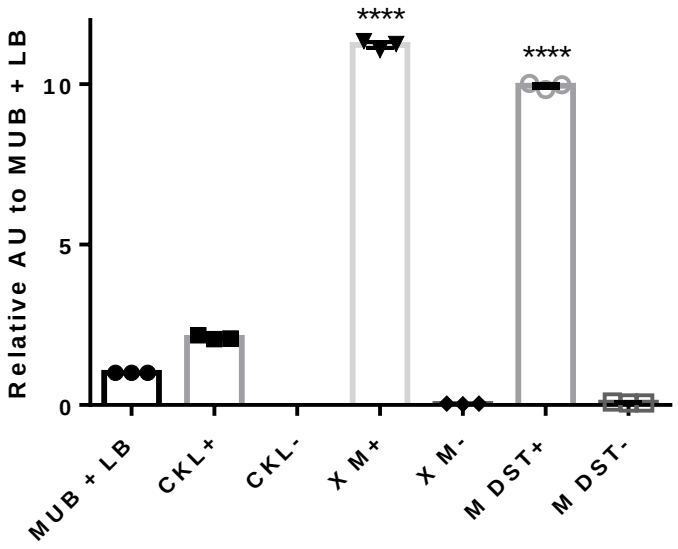


Figure 5.20: Relative fluorescence output (AU) of reaction mixtures following incubation of MUB with the secreted fraction of Δ*CKL, ΔflhDC, ΔclpX ΔmotAB* or *ΔCKL ΔmotAB ΔfliDST* secreting empty (-) or cutinase (+) harbouring secretion construct.

E. coli Δ*CKL, ΔflhDC, ΔclpX ΔmotAB* or *ΔCKL ΔmotAB ΔfliDST* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep (+) or pJexpress-FliC-empty-FLAG-Strep (-) were induced with 0.05mM IPTG (or 1mM for Δ*flhDC*) and grown and prepared for the MUB protein secretion assay as in Figure 5.9. Fluorescence was measured at 446nm following excitation at 302nm. Results from one biological replicate, with three technical repeats for each. Standard error of the mean is displayed. To allow comparison of results between experiments, results were either normalised to the average value for MUB + LB in each experiment. A two-way ANOVA found the variance of data to be statistically significant both between strains and secretion construct cargo. ($p = <0.0001$). Tukey's multiple comparison test of cutinase expressing strains compared to cutinase expressing Δ*CKL* (****: $p = <0.0001$)

The MUB protein secretion assay enabled secretion capacity in 13 strains to be assessed in a time effective manner. Both the MUB and antibody detection protein secretion assay showed that more cutinase was secreted in the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* strains, with *ΔclpX ΔmotAB* consistently secreting the highest concentration of cutinase. Currently *ΔclpX ΔmotAB* is the best secretion strain, based on the large improvements to secretion capacity. However

this may be cutinase substrate specific, therefore secretion capacity of both *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* was investigated for E2 and CH2 protein.

5.4.4. Candidate high capacity secretion strains: E2 and CH2 secretion

To investigate whether *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* had a high capacity to secrete either E2 or CH2 through the modified FT3SS, the relevant plasmids were transformed into these strains and intracellular and secreted concentrations of protein were measured by Western blot analysis.

5.4.4.1. Secretion of E2 through the modified FT3SS of the best candidate secretion strains

E2 protein (pTrc E2) and empty vector was expressed in *ΔCKL* and the two most promising secretion strains. E2 was visible in the supernatant of all strains expressing E2 (Figure 5. top); however it was not as distinguishable as it has been in past experiments shown in Chapter 3. This was in part due to the low concentration of E2 in general (low concentration observed in *ΔCKL* too) and in part due to the incidence of some cell lysis, especially in *ΔclpX ΔmotAB.* This level of cell lysis was seen in all biological replicates. The highest concentration of E2 was visible in the supernatant of *ΔclpX ΔmotAB*, followed by *ΔCKL,* then *ΔmotAB ΔfliDST* No E2 protein or evidence of cell lysis was observed when strains expressed empty vector. Intracellularly, as seen in Figure 5. and Figure 5.10 a protein of around 22kDa was visible in the intracellular fractions of *ΔclpX ΔmotAB* when expressing E2 protein or empty vector – suggesting it is not a result of recombinant protein expression (whether cutinase or E2) the potential origin of this will be discussed later. The Western blot (Figure 5. bottom) confirmed that there was more E2 in the supernatant of *ΔclpX ΔmotAB* cells, similar concentrations were seen in *ΔCKL* and *ΔmotAB ΔfliDST.* Intracellularly very similar concentrations of E2 were seen in all strains.

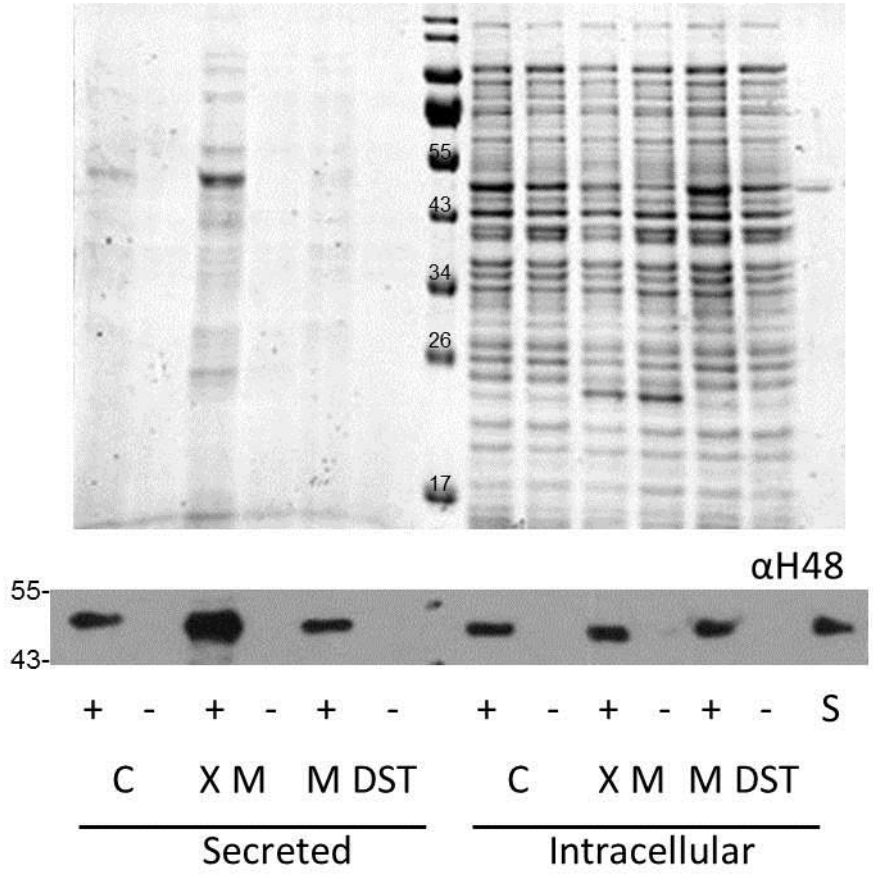


Figure 5.21: Intracellular and secreted E2 from *ΔCKL, ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST*.

E.coli ΔCKL (C), *ΔclpX ΔmotAB* (X M) and *ΔmotAB ΔfliDST* (M DST)containing the plasmid pTrc E2 or pTrc empty was grown in LB. Cells were harvested at OD_{600} 1.5 and prepared for SDS-PAGE as in Figure 5.. Samples underwent SDS-PAGE and either Coomassie staining with Instant Blue (constrast adjusted) or Western blot analysis of cells and supernatant using anti-flagellin (H48) antibody and an HRP secondary. EZ-Run™ Prestained Rec Protein Ladder was loaded, as was an E2 protein standard (S) to allow quantification of intracellular and secreted protein concentration. Sample loaded (supernatant, cells, standard): for Coomassie: 15, 5, 5. Western: 10, 1, 1.

Densitometry analysis was carried out the Western blot shown here and three other biological replicates. The results show that *ΔmotAB ΔfliDST* secretes the least E2 protein, this is expected as discussed previously (Figure 3.) the absence of FliS reduced E2 expression and secretion. *ΔCKL* secreted an average of 1.06µg E2 per Litre of culture supernatant as opposed to 3.98µg for *ΔclpX ΔmotAB.* This represented a significant 3.75 fold increase, however it is evident that cell lysis did occur, therefore how much of this can be attributed to FT3SS secretion is unclear. Intracellular E2 was lowest in *ΔCKL* 54.36mg L^{-1}. The other strains expressed around twice this with averages of 126.63 and 109.00 mg L^{-1} being observed for *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* respectively. In terms of the proportion of secreted E2 *ΔclpX ΔmotAB* was the most efficient secretor, secreting 5.73% total E2 protein.

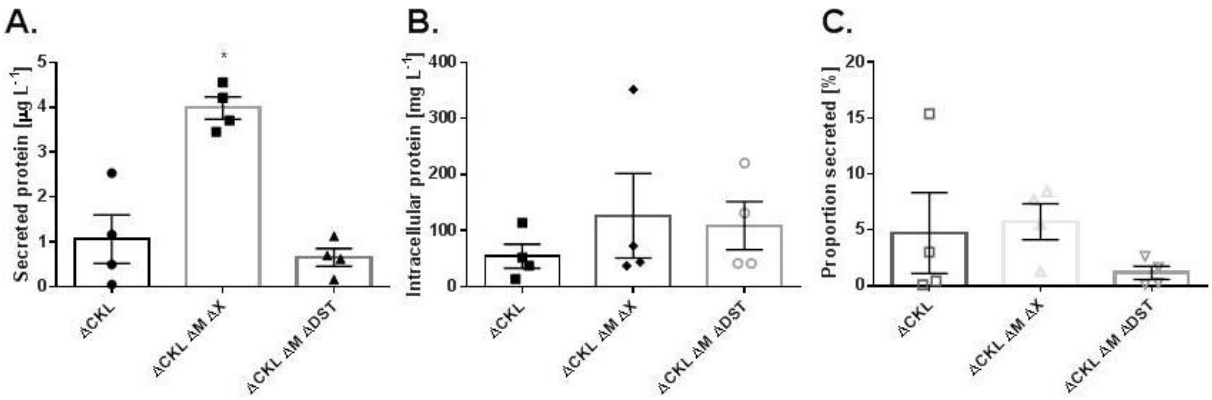


Figure 5.22: Western blot densitometry analysis of intracellular and secreted protein fractions of E2 expressing *ΔCKL, ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST*.

E.coli ΔCKL, ΔclpX ΔmotAB or *ΔmotAB ΔfliDST* containing the plasmid pTrc E2 were grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.5. Following Western blot analysis of cells and supernatant using an anti-flagellin (H48) antibody and an HRP secondary, densitometry analysis was carried out using ImageJ. An E2 protein standard allowed exact quantification of E2 in cell culture. Individual data points, average and standard error of the mean are given for (A) secreted and (B) intracellular E2. (C) Proportion of secreted E2 as a proportion of total E2 both secreted and retained in cells. Four biological replicates for each strain. One-way ANOVA and Tukey's multiple comparison test carried out (*: p = <0.05)

5.4.4.2. Secretion of CH2 through the modified FT3SS of the best candidate secretion strains

CH2 protein was also secreted in the two best candidate secretion strains. Cells expressing empty secretion construct were also grown, to assess the effect of CH2 expression on total protein expression and cell lysis in strains. Figure 5. (top) shows that the supernatant was free of intracellular contaminants, suggesting cell lysis was not prevalent. In the intracellular fractions of *ΔclpX ΔmotAB* the 22kDa protein was again overexpressed, whether cells were expressing CH2 or empty secretion construct. From the Western blot (Figure 5.-bottom) it is apparent that less CH2 was expressed in *ΔmotAB ΔfliDST* compared to the other strains which are quite similar. *ΔclpX ΔmotAB* secretes much more CH2 than the other two strains.

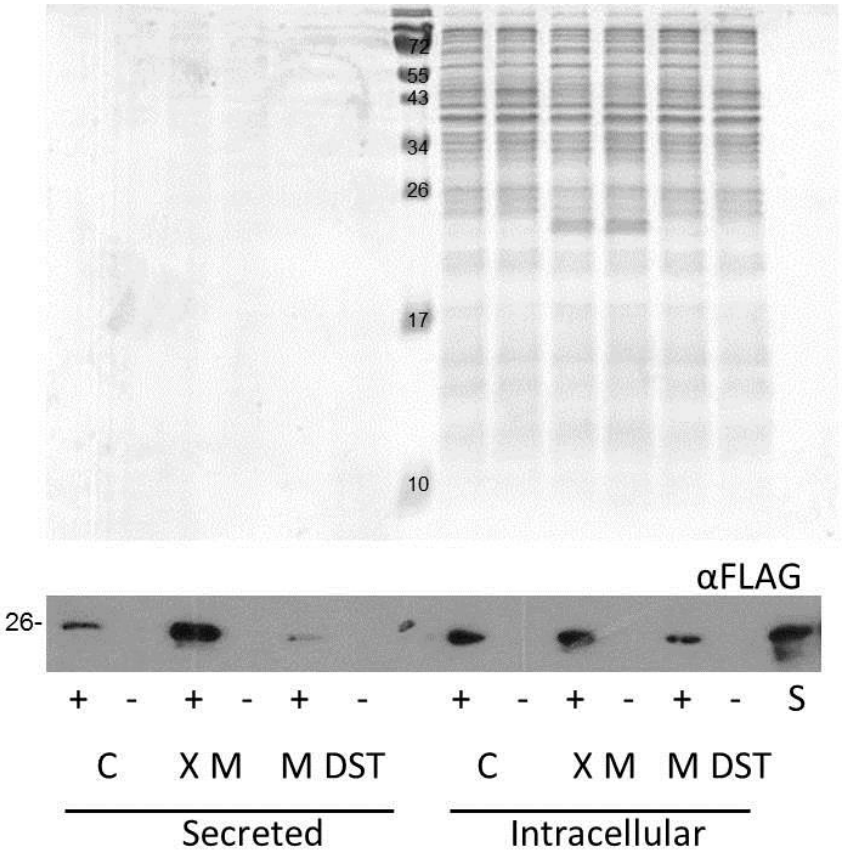


Figure 5.23: Intracellular and secreted CH2 from *ΔCKL, ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST*

E.coli ΔCKL (C), *ΔclpX ΔmotAB* (X M) and *ΔmotAB ΔfliDST* (M DST) containing the plasmid pJexpress-FliC-CH2-FLAG-Strep or pJexpress-FliC-empty-FLAG-Strep were grown in LB supplemented with 0.05mM IPTG. Cells were grown and prepared as in Figure 5.. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain (top) or Western blot (bottom) analysis of cells and supernatant using an anti-FLAG-HRP antibody. Samples were loaded to as follows: Supernatant: 20µL. Cells: 2µL (Western blot), 5µL (Coomassie stain). Standard: 5µL

Two biological replicates were carried out however secreted CH2 was only detected in the blot shown in Figure 5.. Densitometry analysis was carried out when CH2 protein was visible. It was established that 86.76µg CH2 was secreted into a Litre of media of *ΔclpX ΔmotAB* cells, this was 2.56 fold higher than *ΔCKL* which secreted 33.94µg L^{-1} and much higher than the 4.85 µg L^{-1} measured in *ΔmotAB ΔfliDST.* Intracellularly *ΔclpX ΔmotAB* expressed the most CH2 (4.64mg L^{-1} on average), despite this it still demonstrated the highest proportion for secreted protein as

a proportion of total CH2 in cell culture, secreting 0.024% CH2 -this was very low in comparison to values derived in the previous chapter.

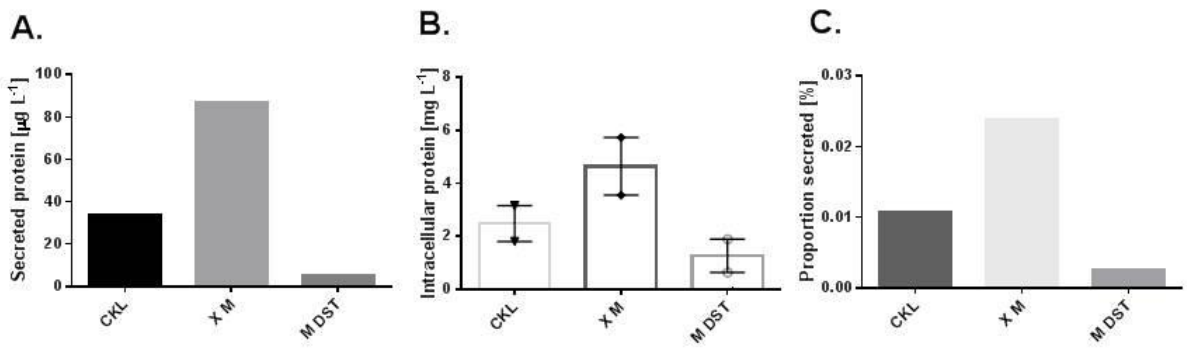


Figure 5.24: Western blot densitometry analysis of intracellular and secreted protein fractions of CH2 expressing *ΔCKL, ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST*.

E.coli ΔCKL, ΔclpX ΔmotAB and *ΔmotAB ΔfliDST* containing the plasmid pJexpress-FliC-CH2-FLAG-Strep were grown in LB supplemented with 0.05mM IPTG. Cells were harvested at OD_{600} 1.0. Following Western blot analysis of cells and supernatant using anti-FLAG-HRP, densitometry analysis was carried out using ImageJ. A CH2 protein standard allowed exact quantification of CH2 in cell culture. Individual data points, average and standard error of the mean are given when possible for (A) secreted and (B) intracellular CH2. (C) Proportion of secreted CH2 as a proportion of total CH2 both secreted and retained in cells. One biological replicate for A and C, two biological replicates for B.

Despite the promise of *ΔmotAB ΔfliDST* with respect to increased secretion capacity for cutinase protein, it is evident that this secretion capacity ability did not extend to all proteins. However the discovery of the high capacity cutinase secretor *ΔclpX ΔmotAB* also proved to be a good platform strain for the secretion of E2 and CH2 protein. Therefore from the 13 modified FT3SS tested, *ΔCKL ΔclpX ΔmotAB* was the most promising, due to its high capacity to secrete a range of protein substrates.

To date investigation has focused on strain improvement to increase secretion capacity of the FT3SS. This has been successful and the combination of strategies has been shown to culminate in an additive effect of secretion improvement, leading to the identification of two high secretion strains. Focus will now shift to investigate whether the secretion capacity can be

improved by modifying the secretion construct. In Chapter 4 is was shown that the secretion construct could be modified to inhibit FT3SS secretion, in this Chapter modifications will be implemented with the aim of improving FT3SS secretion

5.5. Modification of the synthetic modular secretion construct for increased secretion through the FT3SS

The term 'secretion construct' has been commonly used to refer to the transcribed and translated protein secretion construct. For the purpose of this section this will also refer to the untranslated (UTR) regions of the genetic secretion construct too. The current secretion construct harbours the 5' and 3' UTR region and a 47 amino acid secretion signal and was designed as a prototype. In the previous chapter it was found that altering these regions resulted in differential expression and secretion. The 5' UTR region of FliC contains the *fliC* promoter and ribosome binding site (RBS). It is reported that is also harbours a signal for secretion (in *Salmonella*) and that secretion can be directed with the 5' UTR alone (Majander et al., 2005; Anton et al., 2010). However, necessity of the 5' UTR was unsupported for both *Salmonella* and *E. coli* in the Végh et al. 2006 study. In the previous Chapter it was found that the removal of the 5' UTR resulted in increased intracellular expression of protein, as it alleviated cell mediated negative regulation on plasmid expression through the *fliC* promoter, as the promoter is situated in this 5' UTR region (Figure 4.). However absence of the 5' UTR did seem to inhibit secretion. While the 5' UTR may permit some secretion, the 47amino acid secretion signal is attributed as the primary facilitator of FT3SS secretion (Dobó et al., 2010). Removal of the signal peptide has been shown (here (Figure 4.) and in the literature) to result in inhibition of secretion, however the literature suggests that the entire signal peptide is not required to mediate FT3SS secretion. Instead, secretion can be achieved with just the 26-47 amino acid region of the signal peptide in both *Salmonella* and *E. coli* (Végh et al., 2006). The 3' UTR does not appear to be essential for secretion through the FT3SS (Anton et al., 2010; Végh et al., 2006).

Despite the sometimes conflicting evidence, it is apparent that the presence or absence of UTR region and secretion signal peptide can incur variation in secretion, therefore a number of variant secretion constructs were generated. The aim was to assess whether secretion could

be improved with (a combination of) the removal of either UTR or the 1-25 N terminal amino acid sequence of the secretion signal peptide. Since much of the investigation into FT3SS secretion concerned *Salmonella,* a secretion construct was also designed to harbour the 26-47 *Salmonella* FliC secretion signal peptide. It was hoped that one (or more) of these variant secretion constructs would surpass the secretion capacity facilitated by the prototype secretion construct and therefore serve as an improved secretion construct.

5.5.1. Production of variants of the modified secretion construct

Nine variations of the secretion construct were produced (Table 5.). Cutinase cargo was used so that the high-throughput MUB protein secretion assay could be implemented, to efficiently screen secretion construct mediated variation in FT3SS secretion. Primers were designed to amplify the relevant section of the secretion construct with restriction enzyme sites flanking the product, to allow cloning into a restriction enzyme digested pJexpress-FliC-empty vector. If necessary an RBS site and a start codon or stop codon was incorporated into the primers –this was essential if the 5' or 3' UTR were removed respectively, as they harbour these important genetic features. The 26-47 amino acid signal of *E. coli* and *Salmonella typhimurium* is conserved from 29-47, however while 26-28 encode SSS in *E. coli* it is GTA in *S. typhimurium.* A primer was designed which was homologous to the secretion signal from the 28th residue, however the 5' UTR region included the nucleotides to transcribe GTA. Nucleotides were codon optimised for *E. coli* to aid translation. A graphical interpretation of primers and homologous sites is shown in Figure 5..

Table 5.5: Guide to the production of variations of the pJexpress-FliC-cutinase-FLAG-Strep plasmid.

Nine variations of the original pJexpress harbouring secretion construct were produced by PCR product amplification, restriction digest and ligation. PCR reactions (primers and size of product) are listed for each product and were carried out with the pJexpress-FliC-cutinase-FLAG-Strep plasmid as a template.

Reaction/secretion construct number	Product (in pJexpress)/secretion construct (SC)	F	R	Size (bp)
1	*Nde*I- 5' UTR-47aa-cutinase-Strep-FLAG-*Bam*HI	5' UTR 47aa F	no 3' R	1158
2	*Nde*I-RBS-ATG-47aa-cutinase-Strep-FLAG-3' UTR-*Bam*HI	47aa F	3' UTR R	1226
3	*Nde*I-RBS-ATG-47aa-cutinase-Strep-FLAG-*Bam*HI	47aa F	no 3' R	897
4	*Nde*I-RBS-ATG-*E. coli* 26-47-cutinase-Strep-FLAG-3' UTR-*Bam*HI	coli 26-47 F	3' UTR R	1148
5	*Nde*I-RBS-ATG-*E. coli* 26-47-cutinase-Strep-FLAG-*Bam*HI	coli 26-47 F	no 3' R	825
6	*Nde*I-RBS-ATG-*Salmonella* 26-47-cutinase-Strep-FLAG-3' UTR-*Bam*HI	sal 26-47 F	3' UTR R	1148
7	*Nde*I-RBS-ATG-*Salmonella* 26-47-cutinase-Strep-FLAG-*Bam*HI	sal 26-47 F	no 3' R	825
8	*Nde*I-RBS-ATG-cutinase-Strep-FLAG-3' UTR-*Bam*HI	No 5' or 47aa	3' UTR R	1084
9	*Nde*I-RBS-ATG-cutinase-Strep-FLAG-*Bam*HI	No 5' or 47aa	no 3' R	761

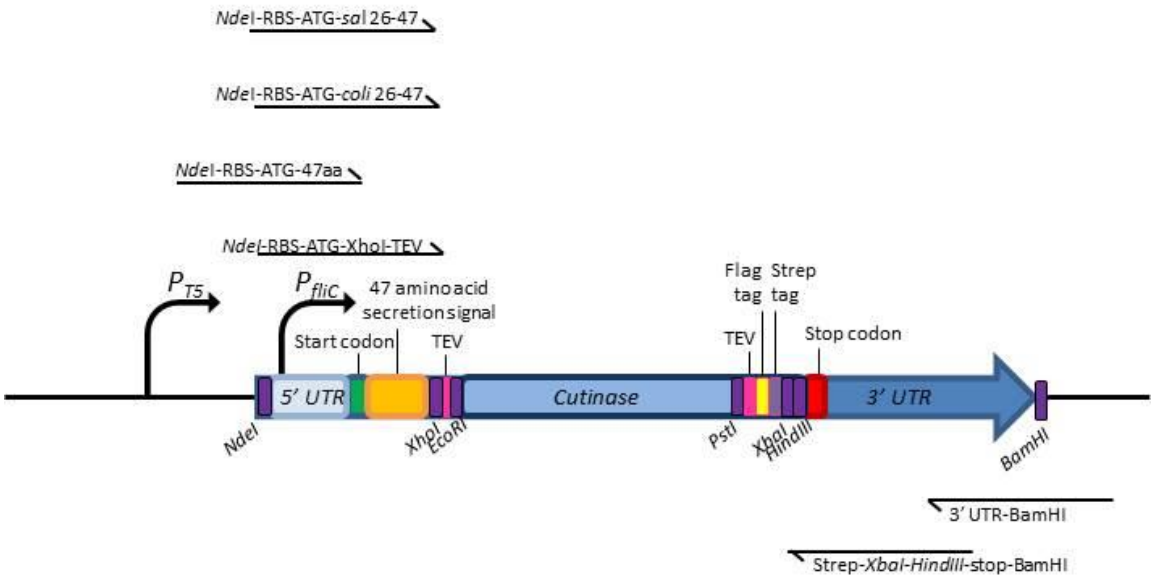


Figure 5.25: Schematic of the existing secretion construct harbouring cutinase and the primers used to generate alternative secretion constructs following the removal or modification of the 5' or 3' UTR regions or the 47 amino acid secretion peptide.

The genetic version of the cutinase harbouring secretion construct is shown. This contains the 5' and 3' UTR and 47 amino acid secretion signal was used as a template for PCR using a combination of the forward and reverse primers with elongated ends shown here. Restriction enzyme sites featured in the elongated ends of all primers. RBS site, start or stop codons or codon optimised regions featured in primers if necessary. This resulted in the generation of donor DNA, which following restriction digest with *Nde*I and *Bam*HI was ligated into an *Nde*I/*Bam*HI digested plasmid harbouring the empty secretion construct. This resulted in the generation of novel plasmids which harbour novel genetic secretion constructs (Table 5.).

Following high-fidelity PCR, DNA products of the anticipated size (Table 5.) were derived for all primer combinations. This was confirmed by running PCR products on a DNA agarose gel (Figure 5.). PCR products were excised from the gel and isolated, for restriction digestion.

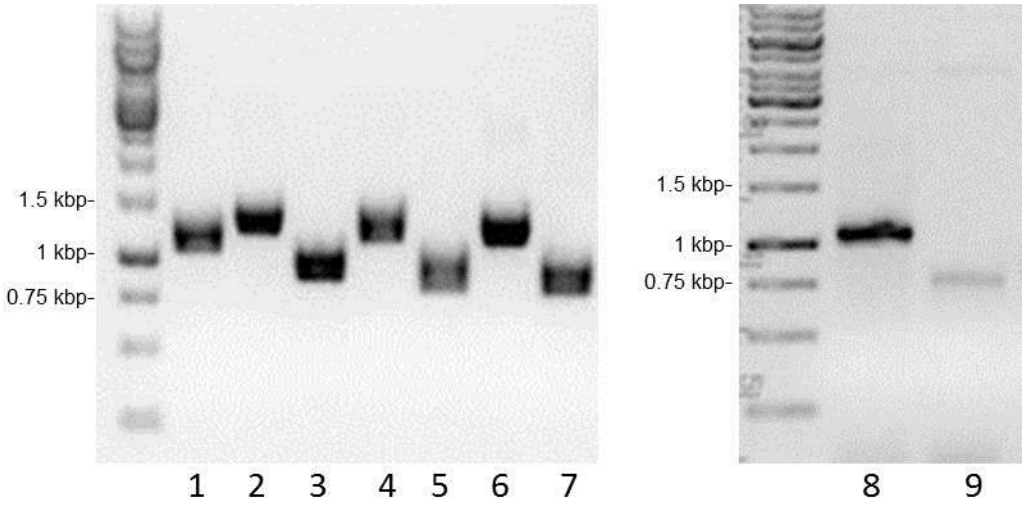


Figure 5.26: Agarose DNA gels showing PCR products for ligation into *Nde*I and *Bam*HI digested pJexpress-FliC-empty-FLAG-Strep

PCR reactions were carried out using pJexpress-FliC-cutinase-FLAG-Strep as template and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. The lane annotations refer to the product number listed in Table 5.; the sizes of products are also given in this table.

Isolated PCR products and pJexpress-FliC-empty-FLAG-Strep underwent restriction digest with *Nde*I and *Bam*HI in CutSmart buffer. Following this incubation step the acceptor vector was treated with Antarctic phosphatase to discourage re-circularisation and then run on a DNA agarose gel to confirm successful restriction digest. As a control, acceptor vector which had been incubated without restriction digest enzymes was also loaded on to the gel. As opposed to the smaller supercoiled circular DNA seen in lane 1, two DNA products are visible in the

second lane, showing the cut acceptor vector and the faint 772kbp section of DNA comprising of the empty secretion construct (Figure 5.). Although the digested acceptor vector appears to run bigger than the 3967kbp expected, the increase in size compared to the undigested and presence of the excised secretion construct give confidence that it is the correct product. As the product resulting from restriction digest of the PCR products was only 18bp shorter than the original product, these restriction digest mixtures were not visualised by DNA electrophoresis. As *Bam*HI cannot be heat inactivated all digested DNA underwent PCR clean up prior to ligation.

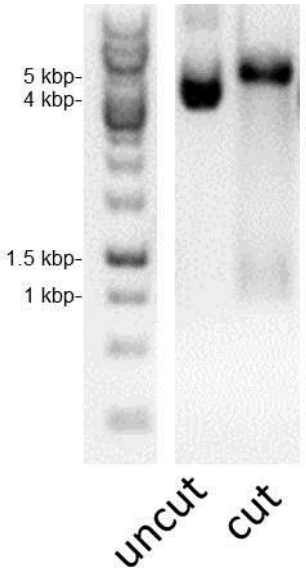


Figure 5.27: Agarose DNA gel showing uncut and *Nde*I and *Bam*HI cut pJexpress-FliC-empty-FLAG-Strep

Restriction digest reactions were carried out both with and without the addition of restriction enzymes. 5µL restriction digest mixture was prepared for DNA electrophoresis and loaded on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide and visualised under UV light and inverted. Lanes: (1) GeneRuler™ 1kb DNA ladder, (2) uncut vector, (3) *Nde*I and *Bam*HI cut vector

Following overnight ligation of all PCR products into pJexpress-FliC-empty-FLAG-Strep, ligation reactions were transformed into NEB 5-alpha Competent *E. coli* and screened by plating on LB agar plates supplemented with ampicillin. Positive colonies were screened by colony PCR with DreamTaq DNA polymerase and relevant primers to confirm successful ligation of DNA. Positive colonies were then sequenced and full sequence alligment was confirmed for at least one plasmid for each product.

5.5.2. Production of variants of the modified secretion construct

Once the full set of secretion construct variants had been established expression and secretion of cutinase was investigated to see if any of the newly derived constructs resulted in improved FT3SS secretion. Plasmids harbouring all nine of the variant secretion constructs –along with the original prototype secretion construct were freshly transformed in to Δ*CKL.* The Δ*CKL* strain was chosen so that secretion of variant secretion constructs could be related to the best characterised secretion strain. Δ*flhDC* secreting cutinase and Δ*CKL* secreting empty secretion construct were included as negative controls; Δ*clpX* secreting cutinase was also included as a positive reference, so that Δ*CKL* secreting the original secretion construct was not the sole point of context to previous experiments. The MUB protein secretion assay was utilised to investigate secreted cutinase and Western blot analysis was used to investigate expression.

The Coomassie stain (Figure 5. top) reveals that intracellular protein expression was uniform in all strains when expressing all plasmids. Expression of cutinase SC8 did result in a visible protein band at around 26kDa. Following Western blot analysis it was evident that when expressed from the prototype secretion construct cutinise was present in all strains. Modification of the secretion construct resulted in variable intracellular expression of cutinase, the highest being when SC8 was expressed. In some strains very little cutinase was observable in cells. In comparison to the 30.6kDa full length prototype secretion construct, candidate secretion constructs 4-7 were all slightly smaller (predicted size 27.8kDa), owing to the fact residues 1-25 were not present and SC8 and SC9 smaller again because of the absence of all 47 residues of the secretion signal peptide (predicted size 25.8kDa). As expected removal of UTRs did not alter protein size.

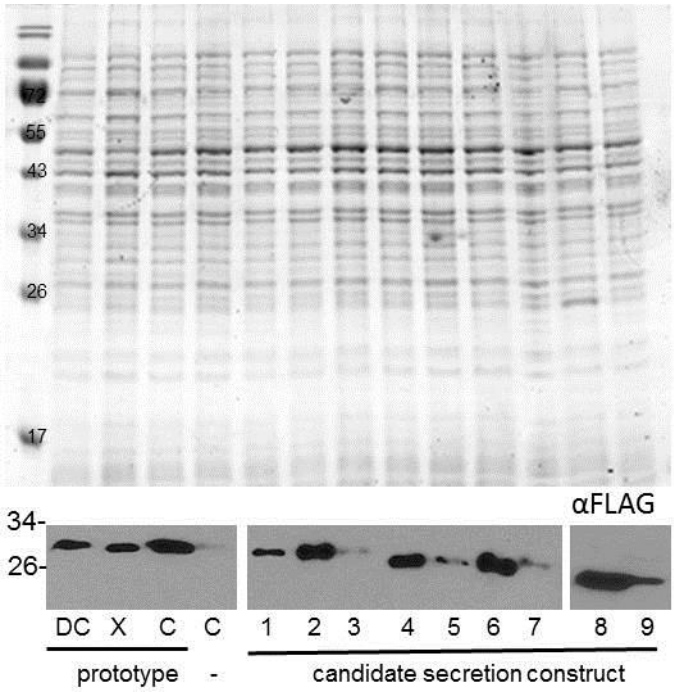


Figure 5.28: Intracellular cutinase in original prototype cutinase-secretion construct expressing *ΔflhDC* and *ΔCKL ΔclpX* and *ΔCKL* expressing the prototype secretion construct and nine candidate secretion constructs

Both *E.coli* strains *ΔflhDC*, *ΔCKL ΔclpX* and *ΔCKL* expressing pJexpress-FliC-cutinase-FLAG-Strep and *ΔCKL* expressing pSC1-9 were grown in LB supplemented 0.05mM IPTG (or 1mM for *ΔflhDC*). See Table 5. for numerical labelling of secretion constructs. Cells were harvested at OD_{600} 1.0. 1 OD unit of cells were prepared with 2x SDS-PAGE loading buffer. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain (top) or Western blot (bottom) analysis of cells and supernatant using an anti-FLAG-HRP antibody. Samples were loaded to as follows: 2µL (Western blot), 5µL (Coomassie stain). One biological replicate of each strain shown.

Results in Figure 5. were analysed by densitometry and combined with another biological replicate by normalising all values to that measured for *ΔCKL* expressing the original cutinase harbouring secretion construct (Figure 5.). This confirmed that significantly less cutinase was present in the intracellular fractions of cells expressing secretion construct variants which do not have the 3' UTR (SC1, 3, 5, 7). In contrast expressing SC8 resulted in significantly higher intracellular cutinase expression (2.19 fold).

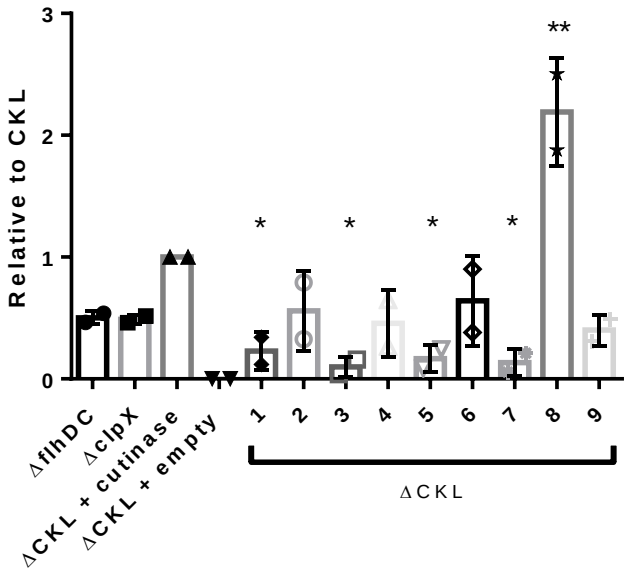


Figure 5.29: Western blot densitometry analysis of intracellular protein fractions of prototype cutinase-secretion construct expressing *ΔflhDC* and *ΔCKL* *ΔclpX* and *ΔCKL* expressing the prototype secretion construct and nine candidate variant secretion constructs.

The Western blot derived image from Figure 5. and one other biological replicate following the same protocol underwent densitometry analysis with ImageJ. Results were then normalised to the value obtained for intracellular cutinase in the *ΔCKL* strain expressing the original prototype cutinase harbouring secretion construct. Individual data points, average and standard error of the mean are given. Two biological replicates were carried out for each strain. One-way ANOVA carried out on data derived from *ΔCKL* cells expressing cutinase (i.e. variable is secretion construct) p = < 0.0001. Tukey's multiple comparison test in comparison to *ΔCKL* secreting original prototype secretion construct. *: p = <0.05, **: p = <0.01.

Next a MUB protein secretion assay was performed on the supernatant fraction of cells expressing the candidate secretion constructs. Following incubation of the supernatant fractions of the cell cultures described in Figure 5. and Figure 5. with MUB, reaction mixes were excited at 302nm and emission at 446nm was read. Results were normalised to the average AU for MUB + LB in each experiment and then normalised to the AU for *ΔCKL* cells expressing cutinase in the original secretion construct to allow the combination of biological replicates and ease of data interpretation. Florescent output values were low for all negative controls. The AU for *ΔclpX* expressing cutinase was on average 1.61 fold higher than that seen in *ΔCKL,* this is in line with previous observations, suggesting that the fluorescent output recorded for *ΔCKL* is dependable on for comparison of cutinase secretion of secretion construct variants. Secretion constructs 2-9 all had emission values which were higher than the values measured for *ΔCKL* secreting empty secretion construct, however they were significantly lower than the emission value for *ΔCKL* with the prototype secretion construct, suggesting that secretion was hindered in these secretion constructs in comparison to the prototype secretion construct. The only variant to result in higher AU readings than the prototype was SC1, which was on average 1.78 times higher –this was found to be significant.

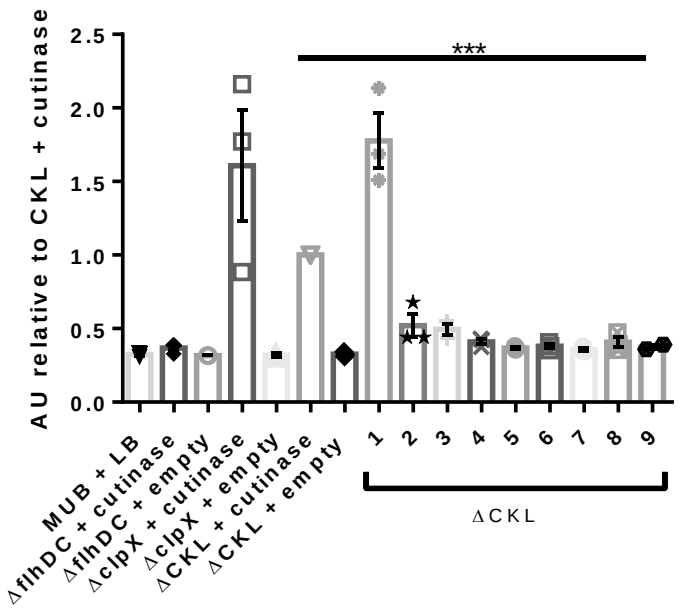


Figure 5.30: Relative fluorescence output (AU) of reaction mixtures following incubation of MUB with the secreted fraction of prototype secretion construct expressing Δ*flhDC* and *ΔCKL* *ΔclpX* and Δ*CKL* expressing the prototype secretion construct and nine candidate variant secretion constructs.

Both *E.coli* strains *ΔflhDC*, *ΔCKL ΔclpX* and *ΔCKL* expressing pJexpress-FliC-cutinase-FLAG-Strep or pJexpress-FliC-empty-FLAG-Strep and *ΔCKL* expressing pSC1-9 were grown in LB supplemented 0.05mM IPTG (or 1mM for *ΔflhDC*). See Table 5. for numerical labelling of secretion constructs. Cells were grown and prepared for the MUB protein secretion assay as in Figure 5.9. Fluorescence was measured at 446nm following excitation at 302nm. Results from three biological replicates, with three technical repeats for each. Individual data points, the mean and standard error of the mean is displayed. To allow comparison of results between experiments, results were normalised to the average value for *ΔCKL* secreting the original secretion construct once adjusted for MUB + LB background AU in each experiment. One-way ANOVA carried out on data derived from *ΔCKL* cells expressing cutinase (i.e. variable is secretion construct) $p = < 0.0001$. Tukey's multiple comparison test in comparison to *ΔCKL* secreting original prototype secretion construct. All at least $p = < 0.001$ (***).

Investigations found that SC8 expressed significantly more cutinase than any other secretion construct, however it secreted little. In contrast SC1 secreted a high concentration of cutinase, however little was found intracellularly. Interestingly the second highest intracellular and secreted concentration of cutinase were both recorded in the original secretion construct in SC8 refers to the secretion construct without the 5' or 3' UTR regions or 47 amino acid signal and SC1 the secretion construct without the 3' UTR region. Reasons for these observations will be discussed later.

5.6. Investigation of secretion through the FT3SS following both strain and secretion construct improvements

While *ΔclpX ΔmotAB* was found to have the highest secretion capacity, *ΔmotAB ΔfliDST* was also effective at secreting cutinase (if not the other two substrates), therefore was also included in this section of investigation. Expression and secretion were investigated with combinations of the two strains with the highest secretion capacity (*ΔmotAB ΔfliDST* and *ΔclpX ΔmotAB*) and the secretion construct which resulted in the highest secretion capacity (SC1). In the absence of a protein standard to quantify absolute yield, both the prototype secretion strain (*ΔCKL*) and prototype secretion construct were used in conjunction to allow comparison of secretion improvements. Cells were cultured and prepared for MUB secretion assay and Western blot analysis as before. The secreted fraction of cell cultures was also prepared for SDS-PAGE to visualise secreted cutinase by another means, but mainly to assess whether cell lysis occurred with any combination of strain or secretion construct.

The absence of intracellular protein in the Coomassie stained SDS-PAGE gel of secreted protein fractions show that cell lysis was not prevalent in any of the strain or plasmid combinations (Figure 5.A). A band is present at around 72kDa, however as this was also present in lanes with no sample loaded, this is likely to be an artefact of buffers used to make and run the SDS-PAGE gels. The intracellular fractions look similar, aside from the commonly overexpressed protein in *ΔclpX ΔmotAB* cells expressing any plasmid (Figure 5.C). Western blot analysis shows that cutinase was secreted in all strain combinations aside from *ΔmotAB ΔfliDST* with SC1 (Figure

5.B). A very small amount of cutinase was present in the supernatant of *ΔCKL* cells, *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* secreted a reasonable amount of cutinase, however the highest concentrations were seen when *ΔCKL* or *ΔclpX ΔmotAB* expressed SC1 - *ΔclpX ΔmotAB* resulting in the highest concentration. Intracellularly cutinase was not visible in the cells of *ΔmotAB ΔfliDST* expressing SC1, which explains why none was secreted in this strain. *ΔclpX ΔmotAB* expressed more cutinase that *ΔCKL* with both secretion constructs and strains expressed more cutinase with the prototype secretion construct (Figure 5.D).

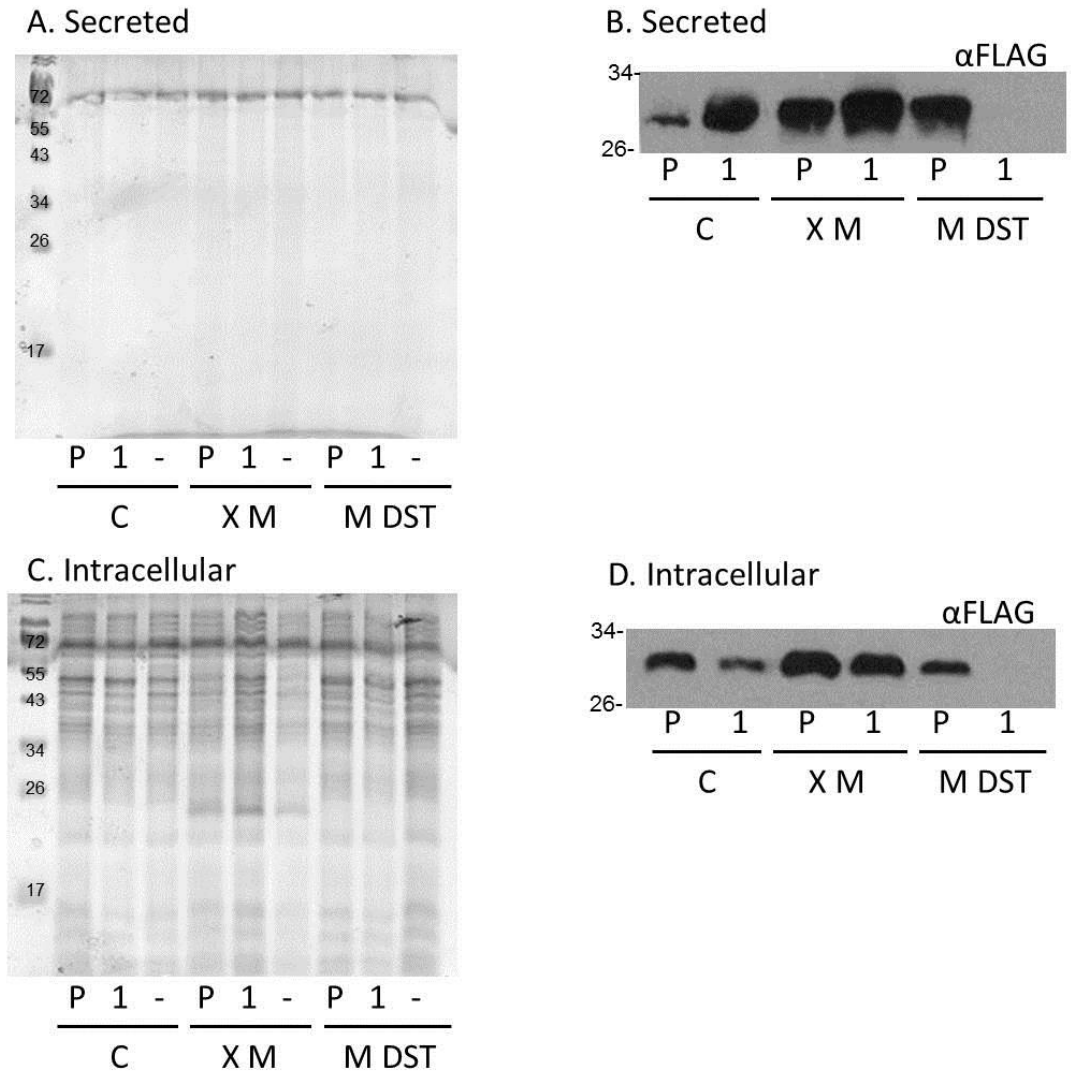


Figure 5.31: Intracellular and secreted cutinase from Δ*CKL* and the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* mutants expressing either cutinase harboured in either the prototype (P) secretion construct or SC1 (1).

E.coli ΔCKL (C), *ΔCKL ΔclpX ΔmotAB* (X M) or *ΔCKL ΔmotAB ΔfliDST* (M DST) expressing either pJexpress-FliC-cutinase-FLAG-Strep (P), pSC1 (1) or pJexpress-FliC-empty-FLAG-Strep (-) was grown in LB supplemented with 0.05 IPTG. Cells were grown and prepared as in Figure 5.. Samples underwent SDS-PAGE and either staining with Instant Blue Coomassie stain or Western blot analysis of cells and supernatant using an anti-FLAG-HRP antibody. (A) Samples from all cultures were loaded onto the Coomassie stained SDS-PAGE ((A).Supernatant: 20µL. (C) cells: 5µL), samples from cultures expressing cutinase (P or 1) were loaded on to the SDS-PAGE gel for Western blotting: (B) Supernatant: 20µL. (D) Cells: 2µL.

Densitometry allowed the results from two biological replicates of the data shown in Figure 5. to be combined. This was achieved by normalising results to *ΔCKL* expressing the prototype secretion construct. The relative abundance of secreted protein was variable (Figure 5.). The highest observed concentration of secreted cutinase was derived from *ΔclpX ΔmotAB* when expressing pSC1 – 4.36 times more on average in comparison to the 2.81 fold increase seen when *ΔCKL* expressing pSC1 compared to the prototype. Cutinase was not observed in the intracellular or secreted fractions of *ΔmotAB ΔfliDST* cells when expressing pSC1 for either biological replicate. Intracellularly the concentration of cutinase in the other two strains was similar whether expressing cutinase in the prototype or pSC1 secretion construct –the concentration was always higher in *ΔclpX ΔmotAB.*

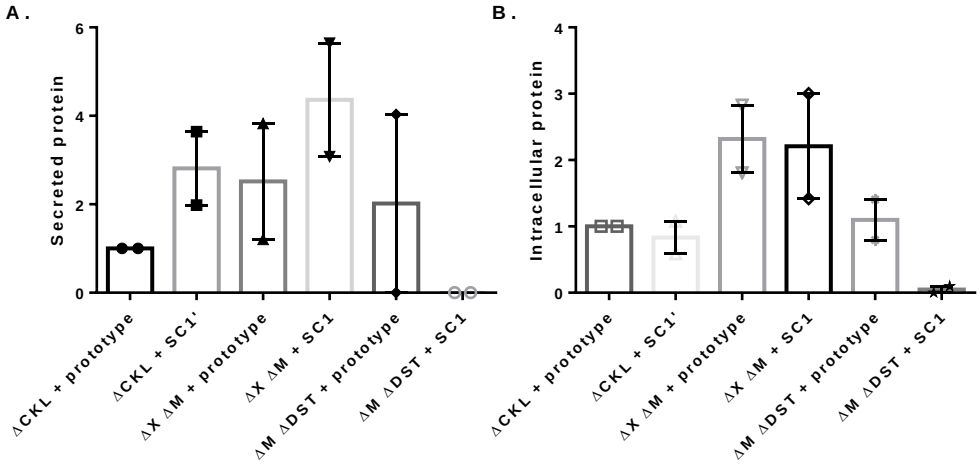


Figure 5.32: Western blot densitometry analysis of intracellular and secreted protein fractions of Δ*CKL* and the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* mutants expressing cutinase harboured in either the prototype secretion construct or SC1.

E.coli ΔCKL, *ΔCKL ΔclpX ΔmotAB* or *ΔCKL ΔmotAB ΔfliDST* containing the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pSC1 was grown in LB supplemented with 0.05 IPTG. Cells were harvested at OD_{600} 1.0. Following Western blot analysis of cells and supernatant using anti-FLAG-HRP, densitometry analysis was carried out using ImageJ and results were normalised to *ΔCKL* expressing cutinase in the prototype secretion construct in each experiment. Individual data points, average and standard error of the mean are given when possible for (A) secreted and (B) intracellular cutinase. Two biological replicates of for each. Two-way ANOVA found that variance was significant between strains for both secreted and intracellular cutinase ($p = <0.05$)

The results of Western blot following probing with the cytoplasmic protein antibody anti-GroEL demonstrate that there was little cell lysis or membrane leakage in the *ΔclpX ΔmotAB* cells when expressing either secretion construct (slightly higher with the prototype, but still negligible) (Figure 5.). Both *ΔCKL* and *ΔmotAB ΔfliDST* have slight cytoplasmic contamination when expressing all both secretion constructs, a slightly higher level of cytoplasmic contamination was seen in the supernatant of *ΔmotAB ΔfliDST* when expressing SC1. However given the sensitivity of this antibody, it can be concluded that cell lysis or membrane leakage is not high in any of the cell cultures shown here.

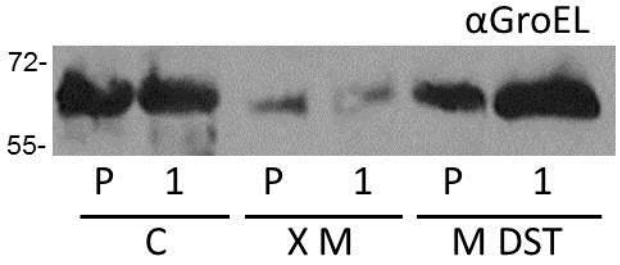


Figure 5.33: Abundance of GroEL in the intracellular and supernatant fractions of Δ*CKL* and the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* mutants, following expression and secretion of cutinase protein either in the prototype secretion construct or SC1 through the FT3SS.

E.coli ΔCKL (C), *ΔCKL ΔclpX ΔmotAB* (X M) or *ΔCKL ΔmotAB ΔfliDST* (M DST) containing the plasmid pJexpress-FliC-cutinase-FLAG-Strep or pSC1 was grown in LB. Cells were harvested at OD_{600} 1.0. Samples underwent SDS-PAGE and Western blot analysis of cells and supernatant using anti-GroEL antibody and an HRP secondary. Samples were loaded as follows: Supernatant: 20µL, cells: 2µL.

Following incubation of the supernatant of the cell cultures analysed in Figure 5.-Figure 5. with MUB, emission values were derived. To combine biological replicates, results were subtracted by the average fluorescence reading for MUB + LB to reduce the effect of background fluorescence, finally results were normalised to that calculated for *ΔCKL* expressing the prototype secretion construct in each experiment, to aid interpretation. Both strain and secretion construct resulted in significant variation in results (Two-way ANOVA. $p = <0.0001$) (Figure 5.). Empty secretion construct resulted in low fluorescent outputs as did *ΔmotAB ΔfliDST* when expressing either of the cutinase containing secretion constructs. However

compelling results were observed for *ΔclpX ΔmotAB* and SC1. In terms of the prototype secretion construct *ΔclpX ΔmotAB* had 7.67 times more fluorescent output than *ΔCKL.* When expressing SC1 this resulted in a 15.70 and 23.78 fold increase in fluorescent output for the *ΔCKL* and *ΔclpX ΔmotAB* strains respectively. All of these increments were shown to be significantly higher than *ΔCKL* with the prototype secretor strain. Additionally the difference between *ΔclpX ΔmotAB* secreting the two secretion constructs was significant, as was the difference or *ΔCKL* and *ΔclpX ΔmotAB* secreting SC1. These amounted to a 3.10 and 1.51 fold change for each of the aforementioned significant increases in emission output. Therefore not only were great improvements made in terms of secretion strain and construct, they are also additive.

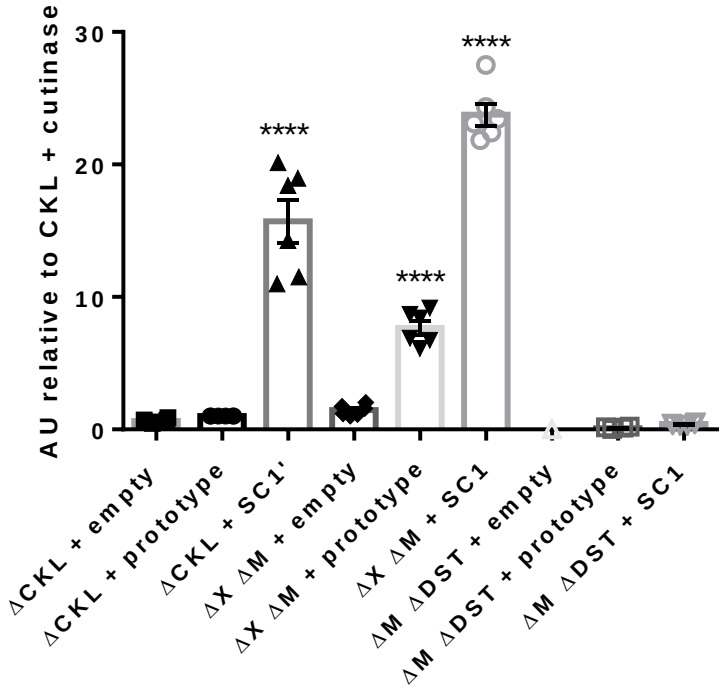


Figure 5.34: Relative fluorescence output (AU) of reaction mixtures following incubation of MUB with the secreted fraction of Δ*CKL* and the *ΔclpX ΔmotAB* and *ΔmotAB ΔfliDST* mutants, following expression and secretion of either the empty secretion construct or cutinase protein either in the prototype secretion construct or SC1.

E. coli Δ*CKL, ΔflhDC*, *ΔclpX ΔmotAB* or *ΔCKL ΔmotAB ΔfliDST* containing either the plasmid pJexpress-FliC-cutinase-FLAG-Strep, pSC1 or pJexpress-FliC-empty-FLAG-Strep were induced with 0.05mM IPTG and grown and prepared for the MUB protein secretion assay as in Figure 5.9. Fluorescence was measured at 446nm following excitation at 302nm. Results from two biological replicates, with three technical repeats for each. Individual data points, mean and standard error of the mean is displayed. To allow comparison of results between experiments, results were normalised to the average value for MUB + LB in each experiment and then normalised to the results for Δ*CKL* expressing the prototype cutinase construct in each experiment. A two-way ANOVA found the variance of data to be statistically significant both between strains and secretion construct cargo. ($p = <0.0001$). Tukey's multiple comparison test of cutinase expressing strains compared to the prototype cutinase secretion construct expressing Δ*CKL* (****: $p = <0.0001$)

5.7. Discussion: Development of a high-throughput assay for FT3SS secretion to screen a multitude of secretion strains and plasmids

5.7.1. Secretion of active cutinase through the modified FT3SS

This Chapter aimed to incorporate the esterase cutinase into the secretion construct to continue to characterise secretion of recombinant protein through the modified FT3SS. While cutinase is an industrially relevant enzyme, the major motivation to investigate the secretion of this enzyme was to develop a high-throughput assay. The assay should efficiently and accurately measure secretion capacity with the aim of screening a large number of potential strains and variants of the secretion construct to establish the most promising combination for achieving a high concentration of secreted protein. It also allowed demonstration that a recombinant secreted protein could fold correctly once secreted into the media, as enzyme activity is dependent on structure.

The recombinant *F. solani* cutinase protein cargo was designed and synthesised *de novo* and successfully cloned into the secretion construct. While Western blot analysis had been successfully utilised in Chapters 3 and 4 to ascertain the concentration of secreted protein and therefore secretion capacity of strain, there were some limitations –namely related to the time consumed and inaccuracies which arise due to inherent properties of the procedure, for example lack of linearity of chemiluminescent signal. Development of an enzyme based secretion assay was key to improving the throughput of secretion capacity characterisation; the specifications were that it should be time efficient and provide accurate results. A multitude of assays have been reported for the cutinase enzyme –although fewer have been reported suitable for the *F. solani* cutinase specifically (Kwon et al. 2009; Liu et al. 2009; Chen et al. 2008). The use of tributyrin, Rhodamin B and pNPB were all tested in preliminary experiments, however MUB proved to yield the best results in preliminary testing (Figure 5.), therefore was chosen as the substrate for the development of the cutinase protein secretion assay. This result was impressive as it indicated that the modified FT3SS had the capacity to secrete a protein, which was able to fold correctly in the extracellular space to exhibit

functionality. SDS-PAGE based analysis of cell cultures was still of importance in this work to ensure that cell lysis wasn't commonplace and to visualise intracellular cutinase concentrations, but also to show that the MUB secretion assay produced results which were concordant and therefore producing reliable results. In line with the MUB protein secretion assay result, cutinase was shown to be expressed and secreted in sufficient concentrations to be detectable through Western blot analysis with anti-FLAG-HRP antibody (Figure 5.). The MUB assay was not suitable for measuring intracellular cutinase; a trial run was carried out where cells were pelleted and resuspended in MUB, however results were similar for strains whether secreting cutinase or not. Although cells were not lysed here, the lack of negative control at this stage made it an undesirable line of investigation.

In previous chapters a protein standard has been derived to quantify yields of both expressed and secreted protein. However this was not possible for cutinase; following overexpression in BL21 (DE3), sufficient cutinase was found in the soluble fraction, however when the purification protocol was carried out, cutinase did not bind to the Strep-trap column efficiently and the low concentration which was recovered quickly degraded. However as cutinase secretion was primarily investigated as a tool to measure relative changes in FT3SS protein secretion, it is not so essential to quantify exact concentrations. That said, it would still be beneficial, therefore expressing cutinase in a different protein expression strain would be desirable to see if any more cutinase could be recovered. Another option would be to attempt to purify cutinase on a column with affinity for the FLAG-tag.

Steps were taken to ensure that the assay was fully optimised to account for substrate concentration (Figure 5.), time of reaction (Figure 5.) and the correct wavelength to measure fluorescent output (AU) following excitation (Figure 5. and Figure 5.8). pH was also investigated and pH 5 settled on due to reduced background fluorescence. While the optimal pH for the assay was not in line with the optimal pH for cutinase, it was evident that cutinase was still active in these conditions and as the assay measures relative fluorescence, the priority is for the assay to be accurate and reproducible, therefore it is not essential that cutinase is optimally functional. Initially the assay was carried out with the addition of quenching buffer and excitation at 365nm, however this produced results with high background fluorescence, which masked the true pattern of variation in the data (Figure 5.). As an alternative reaction mixtures were instead subject to excitation at the UV light range, at 302nm. At this excitation wavelength background fluorescence was both reduced and more uniform, it was observed

that the majority of this was due to the presence of LB in reaction mixes (Figure 5.). This effect may be reduced if cell cultures were grown in a minimal media (e.g. M9) to alleviate the contribution of LB to background fluorescence, however as this investigation concerns the secretion of protein in industrially relevant conditions, the data derived would not be applicable, as industry use rich media in cell culture. As LB was constant in all reaction mixtures, this was not a limitation and with excitation at 302nm the results from reaction mixes which contained cutinase were clearly pronounced from other reaction mixtures. Further to this the differences between 'positive' results were larger, this made data analysis and interpretation much simpler. Finally as pH quenching was not necessary, throughput was again improved.

The image based MUB secretion assay (Appendix 5) was useful as an indicator of whether cutinase was present or absent, however not sensitive enough to distinguish between different concentrations of cutinase by eye. Densitometry was not possible due to the presence of shadows, so this could not be developed into a high throughput test, however there certainly is scope to utilise the image based assay to screen for presence or absence of secretion, for example when screening a transposon library.

Once optimised the assay was utilised to demonstrate that cutinase secretion was directed through the modified FT3SS. This was confirmed by the the absence of cutinase in the supernatant of negative secretion strains (Figure 5.). It was essential that the assay did not report false positive results of cutinase secretion, so this was ideal. It was also demonstrated that the results of the MUB protein secretion assay were concordant with Western blot analysis. Higher fluorescent output (Figure 5.) and higher antibody signal (Figure 5.) are both indicative of a higher concentration of secreted cutinase. Both assays reported Δ*CKL* to secrete the least cutinase and that Δ*CKL* Δ*clpX* secreted twice as much. There was a slight mismatch in Δ*CKL* Δ*flgMN* as the Western blot suggested around twice as much cutinase was secreted, compared to the MUB secretion assay which calculated it to be 1.53 fold higher. However the trend was still correct. The fact that negative control, sensitivity and accuracy have been achieved in the MUB secretion assay is essential to producing reliable results. Throughput was also greatly increased with the utilisation of this assay in comparison to Western blot analysis, however it was hoped that this could be extended by growing cell cultures in a 96 well plate and then centrifuging the plate to isolate cell culture supernatant. This could then be directly applied to MUB to carry out the assay. This modification to the assay was trailed in tandem

with 10mL cultures, however as it was evidently important to normalise culture supernatants with the addition of LB media to account for the variation in OD_{600} of cell cultures, therefore this method was not suitable for high throughput use, as this step hindered the process. As OD_{600} were generally within 10% of each other, this could be overlooked, however it was decided that this should be accounted for to improve accuracy.

With the establishment that the MUB protein secretion assay enabled negative control and the distinction of differing concentrations of secreted cutinase, the assay was utilised for the investigation of secretion capacity of a number of strains and secretion construct variants.

5.7.2. Combination of strain improvement strategies for increased FT3SS capacity

Following on from confirmation in Chapters 3, 4 and 5 that four of the strategies (Δ*motAB,* Δ*flgMN,* Δ*fliDST and* Δ*clpX*) to improve the secretion capacity of the FT3SS were successful, it was desirable to combine these knockout mutations to investigate whether the improvements to secretion would be cumulative. Strains were generated through a combination of Lambda-Red recombineering and P1 phage transduction. This resulted in the generation of eight new strains which carried a combination of knockout mutagenesis strategies which had been shown to increase protein secretion through the FT3SS. The growth curve of strains suggested that knockout mutations generally resulted in lower optical densities of cells compared to Δ*CKL*, however it did not appear that any of the combinations of mutations were lethal or very detrimental to cell growth (Figure 5.). The effect of producing cutinase was negligible on cell growth for all strains (Figure 5.4); suggesting that overexpression of a recombinant enzyme had no negative effects on cell viability.

Once cell viability was confirmed the collection of knockout mutants could be screened to ascertain the most effective secretors of cutinase. This involved testing the eight newly generated strains, the four previously tested secretion strains, the original Δ*CKL* strain and a negative control strain must also be included. To test the secretion capacity of 14 strains using the Western blot based secretion assay would have been extremely time consuming, and additionally limiting as not all of the experiments could be analysed in tandem, whereas all can

be screened on one 96 well plate. The concentration of intracellular cutinase produced in all strains was fairly similar –all were capable of expressing a sufficient amount with induction with 0.05mM IPTG (Figure 5. and Figure 5.). The variability of intracellular cutinase between biological repeats was high (Figure 5.6), this was due to the dependence on the concentration of intracellular cutinase in Δ*CKL* on results. As all results are normalised to this value, any fluctuations in the concentration of cutinase in Δ*CKL* has a large effect on all other results. In future a protein standard would be useful to reduce the variation, as values could be normalised to a constant concentration of cutinase.

Despite the thorough optimisation of the MUB protein secretion assay, results for the concentration of secreted cutinase in all strains were variable (Figure 5.9). However the assay was able to uncover strains which secreted significantly higher cutinase than Δ*CKL.* Previously Δ*flgMN* had been shown to secrete more cutinase than Δ*CKL* (Figure 5.); however this was not the case here, it is likely that this is the result of biological variation –as the results here are derived from three biological replicates, whereas previous reading were derived from one, this data is more reliable. In Figure 5.97, Δ*flgMN* did not contribute to increased secretion capacity with regard to cutinase secretion (in Chapter 3 and 4 E2 secretion in Δ*flgMN* was similar to Δ*CKL* and for CH2 better Figure 3. and Figure 4.). Δ*flgMN* fluorescence was lower than Δ*CKL* alone and Δ*motAB* Δ*flgMN* and Δ*clpX* Δ*flgMN* was low, despite Δ*motAB* and Δ*clpX* being high. This suggests that Δ*flgMN* actually has a negative effect on cutinase secretion through the FT3SS. It has previously been shown that the removal of *clpX* results in increased secretion of FlgM (Guo et al. 2014). Therefore Δ*clpX* will already demonstrate some of the benefits of the strategy which is implemented by removing *flgMN.* This would explain why the combination of the two does not lead to increased secretion in comparison to the two individually, but not why it would be decreased.

The strains which secreted significantly higher concentrations of cutinase were Δ*motAB* Δ*fliDST* and Δ*clpX* Δ*motAB.* Δ*clpX* and Δ*motAB* were both high capacity cutinase secretors individually, so it was anticipated that Δ*clpX* Δ*motAB* would secrete a high concentration of cutinase, however it is an incremental increase, not a cumulative increase. Suggesting that one of these mutations is limiting in terms of secretion capacity. Alone, Δ*fliDST* did not secrete a high concentration of cutinase, however when combined with Δ*motAB* it improved the secretion capacity over Δ*motAB* alone. Interestingly the combination of these mutations did not result in improved secretion capacity of the FT3SS, as Δ*clpX* Δ*motAB* Δ*fliDST* secreted less than

Δ*motAB* Δ*fliDST* and Δ*clpX* Δ*motAB.* If the three mutation strategies concerned flagella proteins, it might be theorised that only so many proteins can be removed before the structure or regulation is comprised, however ClpX is not a direct component of this. An alternative perspective is that there is a point where *flhDC* expression cannot be improved or that additional $FlhD_4C_2$ becomes detrimental to the cell. It was not possible to generate Δ*fliC* Δ*flgKL* Δ*motAB* Δ*flgMN* Δ*fliDST* and Δ*fliC* Δ*flgKL* Δ*motAB* Δ*flgMN* Δ*fliDST* Δ*clpX* in the time of this project, however the poor secretion capacity of Δ*flgMN* Δ*fliDST* and Δ*clpX* Δ*motAB* Δ*flgMN* or Δ*clpX* Δ*motAB* Δ*fliDST* suggest that these strains would not have resulted in improved secretion capacity of the FT3SS.

Of the 14 strains in the Δ*CKL* background, which carry additional gene knockouts, the two best strains (Δ*motAB* Δ*fliDST* and Δ*clpX* Δ*motAB*) –as established by MUB secretion assay- were investigated more thoroughly to confirm the improvement in secretion capacity of cutinase by Western blot and to extend measurement of secretion capacity to additional protein substrates. Western blot analysis of the secreted fraction of these strains found that in comparison to Δ*CKL,* Δ*motAB* Δ*fliDST* secreted slightly more cutinase, whereas Δ*clpX* Δ*motAB* secreted around four times as much cutinase (Figure 5.10). These results were concordant in terms of Δ*clpX* Δ*motAB* with the MUB protein secretion assay (Figure 5.0), however Δ*motAB* Δ*fliDST* had a higher fluorescent output that would be expected based on the Western blot. The rationale for the development of the MUB protein secretion assay was in part due to variable and potentially incorrect results obtained with Western blot analysis, therefore it is more likely that it is the Western blot which is giving inaccurate results. This is particularly relevant when protein abundance is high as lack of linearity in the chemiluminescent signal is more pronounced, leading to incorrect reporting of protein concentration (Gilda et al., 2015). This highlights a major issue when using Western blot analysis to screen unknown and high ranging quantities of protein, Western blots are better optimised by loading similar concentrations of antibody detected protein in each lane of the SDS-PAGE gel. However not only is this unknown, it is also unlikely in this kind of screening experiment. The MUB secretion assay is much better suited to detecting a wide range of concentration of secreted proteins of unknown and high ranging concentration.

In general obtaining a satisfactory Western blot of secreted cutinase was often difficult and experimental failure due to lack of antibody chemiluminescent signal was commonplace; it is possible that cutinase is secreted into the supernatant in very low concentrations, which

hinders detection by antibody binding. An alternative is that the FLAG-tag is for some reason not always accessible or present for antibody binding, however this is unlikely as a the FLAG-tag is detectable in some capacity, therefore must be present. While the overall aim of the work is to achieve a high yield of secreted protein, it is not a limitation if low concentrations of cutinase are secreted because the purpose of this assay is to screen FT3SS secretion capacity to assess improvements, to generate a high capacity platform strain for secretion. Whether the FT3SS has a low affinity for cutinase or whether cells generally express less cutinase than other secretion substrates (E2 and CH2) is hard to assess as a protein standard was not generated for cutinase, therefore an absolute concentration cannot be established. However based on comparison of the antibody signal derived from intracellular and secretion fractions, following loading of the same volume of sample from cutinase or CH2 expressing cells onto an SDS-PAGE gel, it would seem that cells express similar amounts, however less cutinase is secreted (i.e. Figure 4.). Many of the characteristics of CH2 are common to cutinase (i.e. non-native protein, lack of chaperone), therefore many reasons for this can be ruled out. As the secretion apparatus is evidently capable of secreting cutinase (and larger proteins), low secretion capacity is not due to size exclusion, instead it could be due to the characteristics such as hydrophobicity of cutinase residues. The fact that the inner channel of the FT3SS consists of mainly polar residues (Samatey et al., 2001), suggests that substrate hydrophobicity may have resulted in selective pressure, resulting in this evolutionary adaptation. So while substrate hydrophobicity is clearly important, it should not hinder secretion as the FT3SS is already adapted to overcome this.

Secretion in the two most promising strains, Δ*motAB* Δ*fliDST* and Δ*clpX* Δ*motAB* was investigated with the E2 and CH2 secretion substrates, to investigate whether secretion capacity was substrate specific. It is known that E2 is both expressed in low quantities and absent in the supernatant of Δ*fliDST* mutants (Figure 3.), however some E2 was seen in the secreted fraction of Δ*motAB* Δ*fliDST* (Figure 5.). While it was still poor in comparison to Δ*CKL* this does suggest that the increased secretion capacity achieved with Δ*motAB* can somewhat compensate for the loss of FliS chaperone. Secretion was improved 3.75 fold in Δ*clpX* Δ*motAB* and while the concentration of E2 yielded was not very high (1.06µg L^{-1}) in this biological replicate (Figure 5.2), there is potential for a very high yield to be achieved –for example a 3.75 fold increase on the highest achieved yield of secreted E2 in Δ*CKL* grown in 10mL culture (14.674mg L^{-1}), would result in a yield of 55.028 mg L^{-1} secreted E2 protein.

Δ*motAB* Δ*fliDST* secreted CH2 very poorly in comparison to Δ*CKL*, as expression was also reduced this is likely to be the reason for this observation (Figure 5.). This suggests that the capacity of strains to express protein may be substrate specific, however as this is only based on two biological replicates and was not statistically significant; it may not be a true link. Δ*clpX* Δ*motAB* both made and secreted more CH2 than Δ*CKL.* 86.76µg CH2 was secreted into a Litre of media which was high, but not the highest recorded yield (220µg L^{-1} in Δ*flgMN*), however as this was 2.56 fold higher than Δ*CKL* it clearly has a high capacity for CH2 secretion. For example this would extend the highest yield of secreted CH2 in Δ*CKL* from 94.27 µg L^{-1} to 241.33 µg L^{-1}. Whether these yields could actually be achieved remains to be seen, but it does suggest promise for future yields of secreted protein in the Δ*clpX* Δ*motAB* strain.

Δ*clpX* Δ*motAB* is consistently efficient at secreting protein through secretion apparatus, whereas Δ*motAB* Δ*fliDST* was only effective at secreting cutinase. This does highlight a limitation of the cutinase secretion assay as results are not necessarily transferable to different substrates, this may result in investing time in strains when secretion capacity is not transferable to other substrates or conversely result in strains which are poor secretors of cutinase being overlooked, when they have potential to be efficient at secreting a different substrate. However as the MUB protein secretion assay did enable the fast screening of 15 strains conclude that one (Δ*clpX* Δ*motAB*) was very efficient at secreting a number of proteins, the assay is an excellent tool for investigating secretion apparatus capacity. With two strains identified which have a high capacity to secrete cutinase, focus shifted to modification of the secretion construct to improve protein secretion by that route.

5.7.3. Modification of the synthetic modular secretion construct for increased FT3SS secretion

Finally the MUB protein secretion assay was utilised to screen the effect of various features of the secretion construct on expression and secretion of cutinase through the FT3SS. The aim was to improve secretion by modifying the prototype secretion construct. Nine variations of the prototype secretion construct were constructed and expressed in Δ*CKL* -once again the MUB secretion assay was effective in quickly and accurately screening a large number of biological samples. In terms of expression it was found that the absence of the 3' UTR resulted

in poor expression (Figure 5.). The role of the 3' UTR in prokaryotes is not well understood. The stop codon is located here; however this was incorporated into the secretion construct by primer extension. While eukaryote translational control is mainly linked to the 3' UTR, this is not the case in prokaryotes. However there is some evidence that the 3' UTR protects the mRNA sequence from degradation by 3' exonucleases (Belasco, 2010), so it would be logical so assume that the removal of the 3' UTR results in less protein expression because there is reduced mRNA. Further investigation with RT-PCR would be very interesting so observe whether the concentration of secretion construct mRNA is lower in the absence of the 3' UTR.

The removal of the 5' UTR and the 47 amino acid signal peptide resulted in a large increase in cutinase expression (SC8), however removal of the 5' UTR alone (SC2) did not. The removal of the 47 amino acid signal peptide but with the 5' UTR intact was not investigated, it would be informative to assess this in the future, to gain more information on the role of the 5' UTR on secretion, especially as there is much in the literature to suggest that a secretion signal is also harboured in this region (Majander et al. 2005; Aldridge et al. 2006). However is seems possible that the presence of the 47 amino acid signal peptide either hinders expression or directs a high concentration of protein for secretion. If the latter is true then secreted protein degradation must be very high, as the relationship of intracellular to secreted protein is not in linear. Another explanation would be that the secretion signal peptide reduces the formation of inclusion bodies; therefore removal of the secretion signal results in a higher rate of formation of inclusion bodies, which are less susceptible to proteolysis, therefore more protein is present in the intracellular fraction. This was not seen in other studies which investigated the effect of the removal of the 5' UTR and secretion signal, however this work was carried out in *Salmonella* (Végh et al., 2006). Other reasons could be that it improved RNA stability or translation efficiency.

In terms of secreted cutinase, the majority of UTR variants secreted less cutinase than the original prototype secretion construct, aside from the secretion construct which was only defective of the 3' UTR (SR1) (Figure 5.). Following the MUB protein secretion assay, fluorescence output was so low in the other UTR variants that is was more similar to that recorded for Δ*CKL* secreting empty vector than cutinase, suggesting that secretion was almost absent in these strains. This confirmed that in this secretion strain the entire 47 amino acid secretion signal (and not just residues 26-47) is essential for secretion of protein thought the FT3SS in this secretion platform strain, this is differing to what has been reported in the

literature for *Salmonella* (Végh et al., 2006; Dobó et al., 2010). Again in the future it would be interesting to investigate secretion in a secretion construct with the 5' UTR only, as this has been reported to facilitate secretion despite the absence of the 47 amino acid secretion signal. The removal of the 3' UTR alone (SC1) resulted in a large improvement to FT3SS secretion, despite resulting in low expression. This is unexpected as the protein structure of the secretion construct should be the same regardless of the presence or absence of the 3' UTR, unless the absence of the 3' UTR caused it to be degraded by 3' exonucleases (Belasco, 2010). This may be evidence that secretion capacity is tuned to protein expression and is increased when the concentration of intracellular protein tagged for secretion is at a specific concentration (i.e. enough to secrete, not so much that secretion apparatus become 'blocked'), however as the concentration of native protein for FT3SS in a wild type strain is dynamic, this is unlikely. An explanation may be linked to the evidence that overexpression of *fliC* mRNA results in reduced FlgM secretion, if this is due to the 3' UTR region then this could explain improved secretion due to increased induction of late substrate secretion –including protein harbouring the FliC secretion signal (Guo et al. 2014). Interestingly expression of the prototype secretion construct resulted in the second best concentration of intracellular and secreted cutinase, after SC8 and SC1 respectively, suggesting that the prototype secretion construct which harbours both the 5' and 47 amino acid secretion signal and the 3' UTR perhaps results in a balance of the high levels of expression versus secretion. As the highest levels were observed in the absence of the 5' and 47 amino acid secretion signal and the 3' UTR respectively. This may suggest that the two UTR regions and the 47 amino acid secretion signal act on expression and secretion respectively to form an equilibrium of secreted versus intracellular FliC protein in the cell.

5.7.4. Combination of strain and secretion construct based strategies for improved FT3SS secretion

The ultimate aim of this chapter was use the MUB protein secretion assay to screen many different strains and secretion constructs to locate the best performing of the two. This was achieved and the most efficient secretion strain (Δ*CKL* Δ*clpX* Δ*motAB*) and secretion construct (SC1) have now been identified, leading to the designation of these as the mark 2 secretion platform strain and mark 2 secretion construct, with reference to the prototypes; these two

successful strategies to improve FT3SS were then combined to see it the observed effects would be cumulative and secretion capacity realty improved in comparison to the prototype secretion strain and secretion construct (*ΔCKL*). This was achieved as the combination of the Δ*clpX* Δ*motAB* strain secreting the mark 2 secretion construct resulted in a 23.78 fold increase in secreted cutinase (Figure 5.). If this is true of other protein substrates this could result in a large yield of secreted protein and would give this secretion system a good standing in terms of industry standard. Altering the cargo of the mark 2 secretion construct to CH2, E2 and any other protein followed by expression in the mark 2 secretion strain would be very desirable to see if it resulted in a similar output.

5.7.5. Future utilisation of the MUB secretion assay

With the development of the MUB protein secretion assay there is much scope to investigate a number of other strategies to improve secretion through the FT3SS. Strategies can be wide ranging, so long as the implementation of it can be measured by change in secretion. These may improve further strain improvements, for example in terms of flagella gene regulation the flagella protein FliZ is reported to inhibit $FlhD_4C_2$ from binding to flagella class II promoters, thus reducing flagella gene expression (Wada et al., 2012). Other points of manipulation could involve the removal of the other class III genes, which are involved in chemotaxis in response to unfavourable environments and therefore not necessary in a shake flask with LB media; these include CheA-Z (mutants remained motile), Aer –which is linked to anaerobic respiration and also the Entner-Doudoroff pathway, which involves the conversion of sugar acids into energy and therefore should not be an issue in cultures grown in LB with mixing and finally Tar, Tap, Trg and Tsr, which are chemoreceptors for specific monosaccharides, amino acids or peptides (Prüss et al. 2003; Weerasuriya et al. 1998; Grebe & Stock, 1998; Parkinson, 1978).

There is information in the literature that suggests that hook length alternation and straightening are possible. for example deletion of some of the C-terminal residues of *flgE* resulted in straight hooks in *Salmonella* (Kutsukake et al., 1979) -it would be interesting to see if this effected secretion capacity. It would also be interesting to see if increased expression of the ATPase complex (fliHIJ) would result in an increase in substrate sorting and secretion. Another option would be to implement a change in strategy and focus on secretion using an

alternative secretion signal -for example secretion of recombinant protein through the FT3SS with FlgM has previously been achieved (Singer et al. 2012). As FlgM secretion is required prior to FliC secretion, it may result in improved secretion. An alternative would be to engineer directed secretion in a hook-less secretion strain using early secretion substrates, this has been implemented using FlgD and FlgE secretion signal peptides and is also being investigated in this laboratory. Modification of the chromosome by insertion of DNA is another option, whether to implement *flhDC* promoter region modification or promoter replacement as discussed in Chapter 3, or to insert secretion construct DNA for more stable expression.

There would also be merit to constructing a transposon library of mutants with a range of mutations, these could then be screened my MUB protein secretion assay to identify novel sources or combination of gene knockout to aid increased protein secretion. To date motility based assays been used to identify novel sources of either improved or reduced motility, however this could serve as a more high-throughput alternative (Girgis et al., 2007; Cameron et al., 2008). In the spirit of synthetic biology and reductionist biology (e.g. Synthia and Venter (Gibson et al. 2010; Hutchison et al. 2016)), it would also be interesting to investigate the 'minimal flagella' required to mediate directed secretion. Not only would this be interesting but it could reveal unknown information about flagella protein function and essentiality. Another synthetic biology example would be to investigate if refactoring the flagella operons could aid secretion (Temme et al., 2012). This is based on the idea that the chromosome can be streamlined for more controlled and possibly improved expression. This involves taking a relevant operon and removing non-coding or regulatory genes, next each coding gene is recoded with different codons –in an effort to reduce all native regulation, these genes are then reordered and put under the control of synthetic promotors, ribosome binding sites and spacer sequences. The product is a refactored gene cluster organised into well characterised genetic parts with no native regulation. This results in greater control of gene expression and aid investigation into function of the individual components of the system –this may result in more efficient gene expression. A model predicted that the most efficient way of ordering the flagella operons was (predictably) from class I, to class II to class III (Kalir et al., 2001).

The assay could be modified to gauge more information on protein secretion in terms of yield over time. Following culture of cells secreting cutinase (or any substrate if the MUB secretion assay isn't being deployed), cells could be pelleted and supernatant removed for measurement of secreted protein. The cell pellet could then be resuspended in fresh media and incubated

for a further period of time, repetition of supernatant retrieval would allow secretion over time to be accurately assessed. The assay could also be deployed to investigate the effect of cell culture method protein secretion. For example the effect of temperature, agitation, culture size, different media on protein secretion.

In summary with the advent of the MUB protein secretion assay there is huge scope to improve or just investigate FT3SS secretion further, with hope of producing a high capacity secretion strain suitable for use in industry or facilitating the furthering of knowledge on the FT3SS.

Strain improvements could be based on further logical and informed strategies from the literature, this may relate to an extension of the 'reduced metabolic burden' strategy, for example removing the remaining chemotaxis linked class III genes (as in (Singer et al. 2012). Alternatively it could take a more engineering, synthetic biology based strategy, this has particular relevance when it is considered that a low percentage of total protein was secreted through the FT3SS. Steps could be taken in an attempt to improve this, for example increasing expression of the ATPase complex (fliHIJ) to see if substrate sorting and secretion could be increased or by investigating whether straight hooks could be implemented with deletion of some of the C-terminal residues of *flgE* as seen in *Salmonella* (Kutsukake et al., 1979). Another alternative would be to implement modelling to see if there are any bottlenecks occur, for example proton or sodium motive force could be investigated, to identify routes to improve secretion (Benedict et al., 2012). Proteomics could also be utilised to identify sources of improvement, for example an iTRAQ experiment of cells secreting the secretion construct and empty vector could infer routes for improvement. Finally a non-directed approach could be taken to identify mutant strains which are efficient at secreting protein. This could be carried out by generating a transposon library of mutants from the Δ*CKL* background and then screening secretion via a cutinase based MUB assay. This would generate strains with mutations throughout the chromosome, which is beneficial because favourable mutations may not actually be connected to flagella gene expression -for example they may result in the loss of a non-essential metabolically expensive function. In addition to improvements to the strain, further improvements may be valuable to the secretion construct. A secretion construct plasmid library could also be assembled and tested, this could include many combinations of secretion signals and UTRs from different flagella substrates, both in their entirety and in sections. It could also be expanded to investigate the effect on induction of the secretion

construct, as this was a theme throughout this thesis. A range of promoter and RBS codons could be utilised from the BioBricks Catalogue for example –again the MUB –protein secretion assay would aid the investigation here.

The MUB secretion assay could also be utilised to implement non-flagella gene related improvements, for example investigating whether more cutinase is secreted when co-expressed with enzymes which have been shown to aid recombinant protein production. The effect of cell culture environment could also be screened, whether this is to investigate whether secretion is improved in various medias (relevant in terms of chemotaxis based gene expression) or induction protocols, but also in an effort to alter flagella hook structure, as certain pH, temperatures and salt concentrations have been shown to result in straight hooks, although it should be noted that as this was carried out *in vitro* it may not have the same effect on live cell culture (Kato et al., 1984; Hirano et al., 1994).

Chapter 6: 45kDa human collagen expression for FT3SS secretion

There are two outlooks for developing profitable industrial biotechnology platforms; one is to produce a high yield of product and the other to produce valuable product. In the previous results Chapters it was demonstrated that the FT3SS could potentially be modified into a high capacity secretion machine. Through this system it was shown that both native and recombinant proteins could be secreted into the extracellular media and that with improvements to both the strain and secretion construct almost 25 times the original secretion capacity could be achieved. If this was transferable to the yields seen prior to improvement this could result up to 55.028 mg L^{-1} E2 protein or 241.33 µg L^{-1} recombinant CH2 protein being secreted. The focus of this chapter will be to attempt to secrete a pharmaceutically relevant 45kDa human pro α (I) collagen chain in the section platform strain. This was previously expressed in barley (Eskelin et al., 2009) and represents half of the full length chain, but retains many of the important collagen 1 cell adhesion properties. As explained in the introduction collagen has many therapeutic uses and is currently obtained from animal sources, however there is a drive to produce recombinant collagen as this would resolve issues of biocompatibility, safety and homogeneity (Olsen, 2003; Nuutila et al. 2015). Current recombinant sources include plants, animal cells and *Pichia.* The highest yield reported is 500 mg L^{-1} in *Pichia* (Nokelainen et al., 2001) and the highest secreted 150 mg L^{-1} in HEK cells, but at a high cost/ yield ratio (Tillet et al. 1994). Secretion has only been reported in mammalian cells, therefore is would be beneficial to produce and secrete collagen in strains which harbour the modified FT3SS, as the helical nature of collagen 1 lends itself to potential flagella secretion as it has a diameter that easily fits within the 2nm pore of the secretion apparatus. While this would incur all of the benefits relating to *E. coli* based expression systems (low cost, fast growth etc.), there are some negatives surrounding post translational modifications however these can be partly overcome with the co-expression of prolyl-4-hydroxylase and lysyl-hydroxylase-3, which have been shown to effective in producing collagen with correct hydroxylation modifications of prolyl and lysyl residues (Figure 1.7). To date the only reports of expression of chains of human collagen reported in *E. coli* are for a 38kDa fragment of human pro α (III) collagen chain. When co-expressed with prolyl-4-hydroxylase and lysyl-hydroxylase-3 a yield of 90mg L^{-1} intracellular COL3A1 collagen was achieved with a hydroxylation pattern

similar to the human form (Rutschmann et al., 2014). A 23kDa fragment of human pro α (I) collagen chain was also expressed to 10% of total cellular protein and was also highly hydroxylated with co-expression of prolyl-4-hydroxylase (Buechter et al., 2003). Collagen is a proline rich protein and while *E. coli* are able to produce proline autotrophically, production is limited by a negative feedback loop of proline on glutamate-5-kinase (Adams & Frank, 1980). This lack of intracellular proline is likely to limit the yield of recombinant collagen, therefore experimental design and strain improvements were designed to try and improve the concentration of intracellular proline, in an attempt to increase collagen expression.

6.1. Production of collagen within the synthetic modular secretion construct

The aim of this chapter was to express a 45kDa human pro α (I) collagen chain in the original secretion platform strain (Δ*CKL*) in the first instance. This was partly chosen as the 45 KDa fragment is a similar size to flagellin (the most abundant substrate of the FT3SS) While expression would be impressive, secretion would also be desirable- as the idea of extruding collagen fibres from flagella export apparatus would be a great achievement. In addition, as described in the last section, it would be the largest fragment of collagen to be produced in *E. coli.* The pro α (I) collagen chain was chosen as collagen I is the most abundant collagen protein, therefore the applications of recombinant collagen I would be wide- for example in tissue engineering -both *in vitro* (for cell culture) but also *in vivo* for aid in skin graft treatments. From herein the 45kDa pro α (I) collagen chain will be referred to as 'COL1A1'.

6.1.1. A synthetic collagen gene in the synthetic modular secretion construct

Prior to the experimental work carried out in this thesis the sequence for the 45kDa human pro α (I) collagen chain or COL1A1 (accession number: NP000088) was obtained from the NCBI database. The gene was codon optimised for *E.coli* using the Genes® algorithm and then

incorporated into the prototype secretion construct sequence between the *Eco*RI and *Pst*I restriction enzyme sites and consideration was taken to exclude any other restriction enzyme sites that are present in the secretion vector or constructs (*Nde*I, *Xho*I, *EcoR*I, *Pst*I, *Xba*I, *Hind*III, *Bam*HI). The resulting sequence was sent for synthesis by DNA 2.0. The gene was synthesised in an IPTG inducible pJexpress404 vector and the resulting genetic secretion construct and translated protein product is shown in Figure 5..

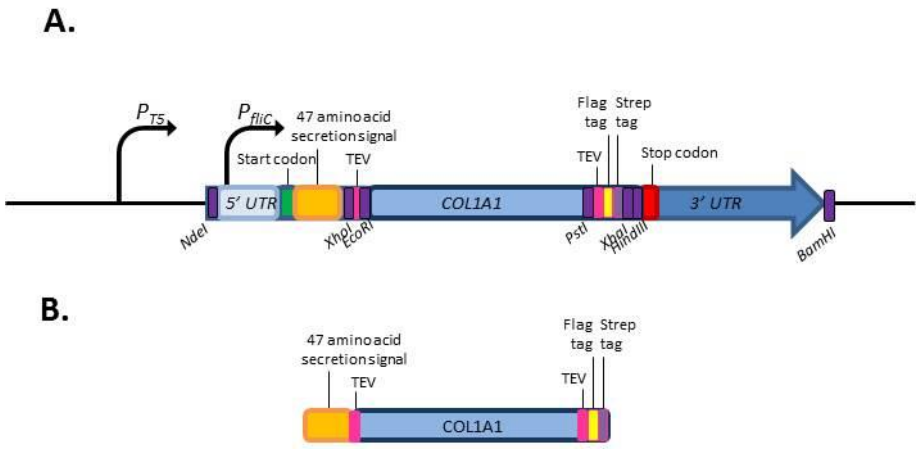


Figure 6.12: Genetic and protein schematic of the secretion construct harbouring COL1A1

(A) Genetic schematic of pJexpress-FliC-COL1A1-FLAG-Strep, which resulted in the expression (B) of a FT3SS signal peptide COL1A1 flanked by FLAG and Strep tags and TEV protease sites

6.1.2. Expression of the COL1A1 harbouring synthetic modular secretion construct: hyperosmotic shock

The protein secretion assay was utilised to observe whether COL1A1 was expressed in the secretion system expression platform. As human collagen has proved difficult to express in *E. coli* it was not anticipated that secretion would necessarily occur, as yields were likely to be low. Initial investigation of COL1A1 expression was carried out in the original Δ*CKL* secretor strain. In preliminary work by M. Hicks (in our laboratory team) COL1A1 had been successfully expressed but at low levels in the MC1000 Δ*CKL* strain when cells were incubated overnight in liquid culture before induction with 1mM IPTG, however no secretion was observed. A priority is to increase the concentration of collagen in cells, as it has been observed throughout

experimentation in this thesis that the percentage of secreted protein in relation to total protein is consistently low (around 5% the maximum observed) in any variant of the secretion strain, therefore if the intracellular yield is low it is unlikely that collagen will be detectable in the secreted fraction.

Proline is an osmoprotectant, it has previously been shown that supplementing the media with proline and applying osmotic shock has been shown to cause cells to uptake proline from the environment, leading to high internal proline levels (Buechter et al., 2003); this lead to the successful production of hydroxyproline. In *E. coli* hyperosmotic shock causes cells to upregulate ProU and ProP proline transporter expression, resulting in increased transport and therefore increased accumulation of intracellular proline (Wood, 1988). Therefore it was investigated whether this would aid intracellular collagen production. Δ*CKL* freshly transformed with either empty or COL1A1 harbouring secretion construct cells were incubated for 16 hours at 37°C with 180rpm agitation, cells underwent osmotic shock with 500mM NaCl (or dH_20 as a negative control) and returned to the shaking incubator. After 30 minutes cultures were supplemented with 1mM IPTG and returned to the shaking incubator for 3 hours. Cultures were then prepared for analysis by SDS-PAGE and Coomassie stain or Western blot analysis. Following this procedure the OD_{600} of cultures was generally around OD_{600} 3.5-4, with or without the application of osmotic shock.

The Coomassie stained SDS-PAGE gels show that a low concentration of intracellular protein was seen in the supernatant of cells expressing empty secretion construct following hyperosmotic shock, however aside from that no other protein was visible (Figure 6.A). In the Western blot probed with the anti-FLAG-HRP antibody a band was detected in samples of cells expressing COL1A1, which was absent in cells expressing empty secretion construct, between 72 and 55kDa at around 65kDa –as expected no FLAG was detected in the collagen standard (Figure 6.C). Likewise, in Figure 6.D anti-COL1A1 detected a band in these samples at around 65kDa (indicated by an arrow) which is also absent in the negative control-empty vector lanes. The antibody binding pattern is slightly different in the standard, which was expected as the collagen is derived from a complex biological source from a different species (rat), however the collagen motif is evidently present in this sample, therefore it serves as a good positive control. Full length COL1A1 is a 45kDa protein, however with the secretion construct accessory peptides it is predicted to be 54.5kDa. Although it appears slightly larger than expected, the use of two antibodies and negative controls mean that it can be concluded that this is COL1A1

(N.B. this was also confirmed by M Hicks using MS/MS on Strep-tag purified COL1A1). In Figure 6.C it is evident that hyperosmotic shock led to a reduced concentration of intracellular COL1A1 in comparison to without. Supernatant fractions were probed with anti-FLAG-HRP for the presence of COL1A1, however no FLAG tag was detected (not shown).

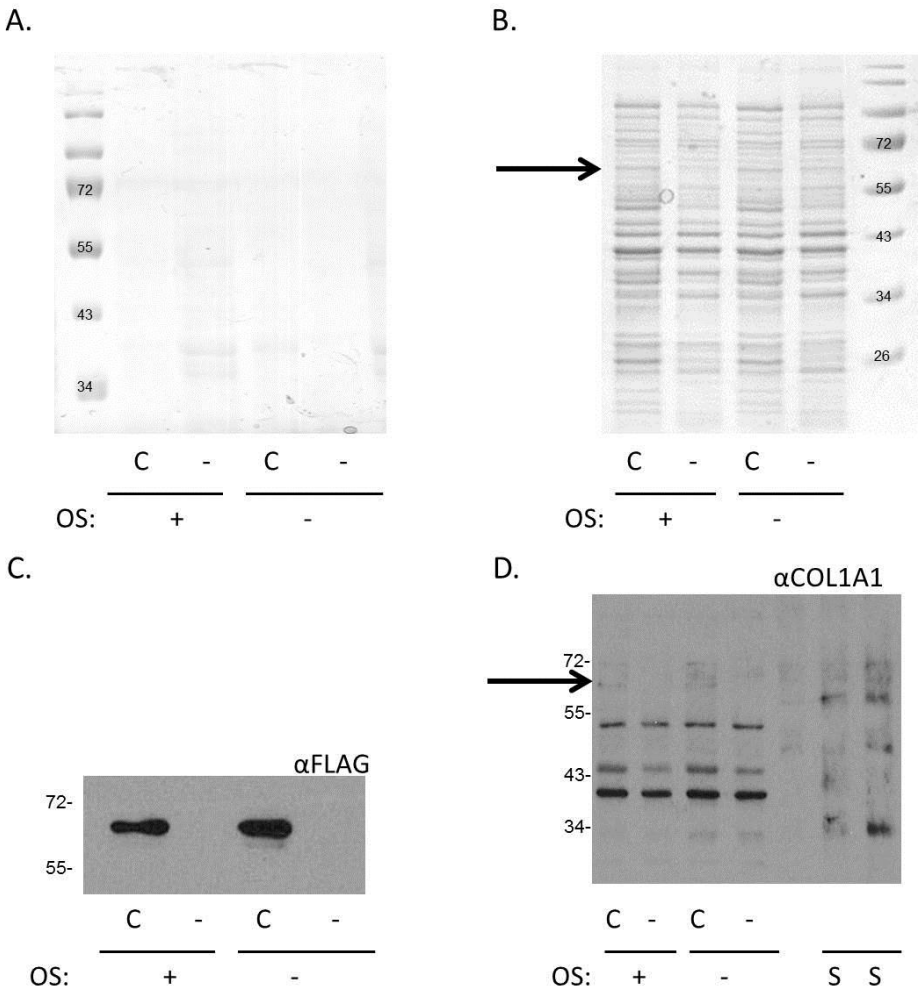


Figure 6.2: Intracellular and secreted fractions of Δ*CKL* expressing either empty (-) or COL1A1 (C) harbouring secretion construct following osmotic shock.

E.coli ΔfliC ΔflgKL (Δ*CKL*) harbouring either the plasmid pJexpress-FliC-COL1A1-FLAG-Strep (C) or pJexpress-FliC-empty-FLAG-Strep (-) were grown in LB overnight at 37°C, 180 rpm, following induction of hyperosmotic shock with 500mM NaCl or the same volume of dH_20, cells were incubated for 30 minutes and then supplemented with IPTG. Cells were incubated for 3 hours and then harvested. 1 OD unit of cells were prepared for SDS-PAGE and the supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western blot

analysis. A rat tail COL1 standard was loaded as a positive control. Coomassie stained SDS-PAGE of the (A) supernatant and (B) intracellular fractions of cell cultures. Western blot of the intracellular fractions probed with (C) anti-FLAG-HRP and (D) anti-COL1A1 and appropriate secondary antibody. Osmotic shock (OS): + or -. 10µL sample loaded for each sample. 10µL rat tail COL1A1 on (D). Arrows denote presence of COL1A1.

Following densitometry analysis of two biological replicates, the application of osmotic shock was shown to result in a reduction of COL1A1 expressed in cells (Figure 6.) –following osmotic shock cells expressed half the concentration of COL1A1 in comparison to controls. This was unexpected, although it is not known if this did relate to increased intracellular proline or not.

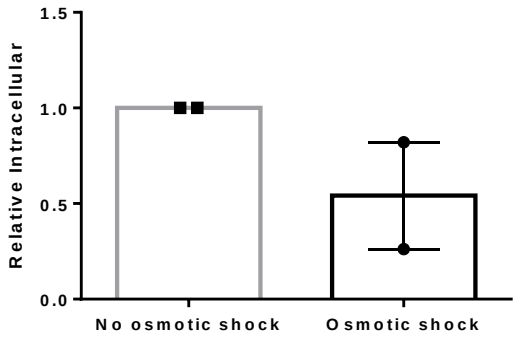


Figure 6.3: Results following densitometry analysis of the intracellular samples of anti-FLAG-HRP probed Western blots of the intracellular fractions of COL1A1 expressing cells following cell culture with or without osmotic shock.

Following cell culture and initiation of osmotic shock if appropriate as in Figure 6., densitometry analysis was carried out on Western blots following probing with anti-FLAG-HRP. Results from two biological replicates, normalised to the result for COL1A1 without osmotic shock. No statistical significant found.

6.2. Strain improvements for increased intracellular proline

An alternative strategy to improve intracellular proline and therefore possible COL1A1 production is to make strain improvements to the expression strain. Strategies to implement this were identified by searching the literature and examining the metabolic network. The metabolic network for proline metabolism is shown in Figure 6.. There are three membrane transporters of proline Put, ProU and ProP. All transport proline into the cell. ProP is an ABC transporter and expressed through a wide range of osmolalities, ProU is a dual proline/betaine transporter and is transcriptionally regulated in response to high osmolality (Mellies et al., 1994; Wood et al., 2001). PutP is very tightly regulated in response to PutA concentration, therefore not directly linked to osmolality (Dunlap & Csonka, 1985). Pyrroline-5-carboxylate synthase synthesises intracellular proline and several pathways result in the production of the proline precursor 1-pyrroline-5-carboxylate. Additionally PutA results in two enzymes products (proline dehydrogenase and 1-pyroline-5-carboxylate dehydrogenase) which result in proline and proline 1-pyrroline-5-carboxylate catabolism resulting in less proline. This demonstrates that the concentration of proline could be increased intracellularly by (1) increasing the amount of proline being transported into the cell (2) to increase the amount of proline produced by the cell's metabolic pathways (3) to reduce proline catabolism. Strategies will be implemented by gene knock out mutagenesis or by changing promoters to cause stronger gene activity. Strategies 2 and 3 were implemented in this thesis and will be discussed in more detail in the next section, with support from the literature.

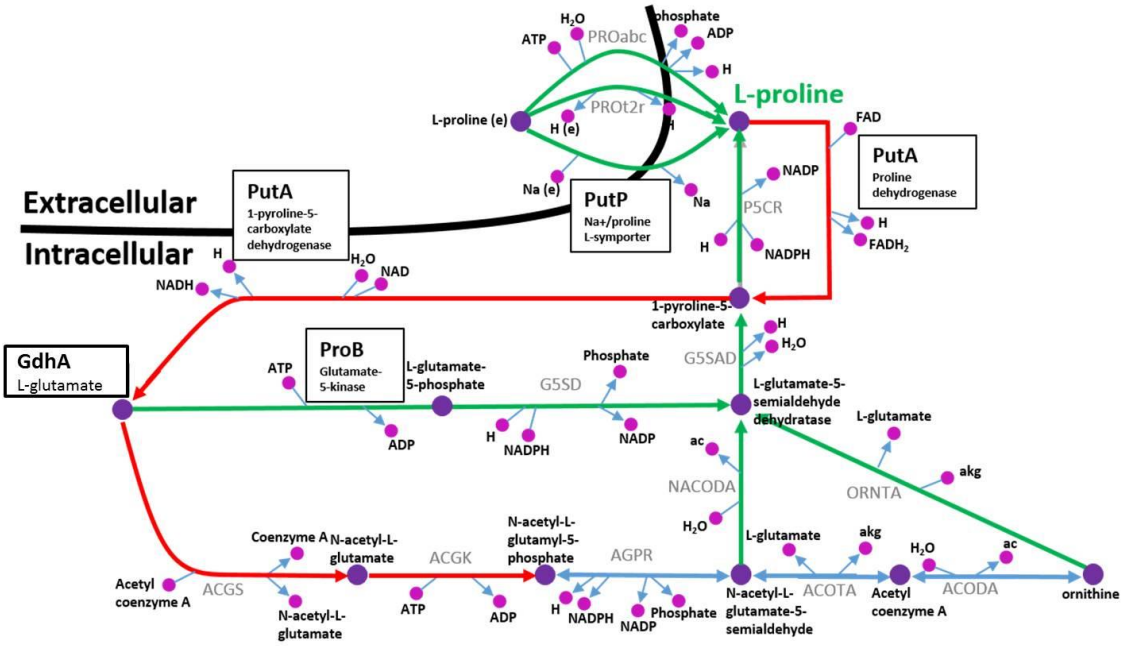


Figure 6.4: The proline metabolism network in *E. coli*

Green arrows denote reactions which result in the production or transport of proline and its precursors, red arrows denote the opposite. Intracellular proline is both derived and used intracellularly and transported into the cell from the extracellular environment. Proteins of interest include: the transporters PROabc (ProU), PROt2r (ProP) and PutP. Proteins: PutA, ProB and metabolite: L-glutamate also shown which is metabolically relevant in other metabolic networks. Adapted from Reed et al. 2003

6.2.1. Removal of feedback inhibition of proline accumulation: ProB

As described, *E. coli* are able to produce proline autotrophically , however production is limited by a negative feedback loop of proline on glutamate-5-kinase (Adams & Frank, 1980). Proline binds to glutamate kinase (ProB), changing the shape of the active site, causing inactivity of glutamate kinase. This is known as allosteric inhibition. Removal of feedback inhibition has been achieved on two occasions by two different point mutations in the *proB* gene, which encodes glutamate-5-kinase. These point mutations were implemented at codon 107 (gat -> aat, D -> N) or codon 143 (gag -> gcg, E -> A). (Csonka et al., 1988; Rushlow et al., 1985). Intracellular proline levels were not reported in the literature; however it is

hypothesised that this will occur, potentially resulting in increased COL1A1 expression. These two point substations will therefore be implemented in the Δ*CKL* strain using Lambda-Red recombinase based knockout mutagenesis followed by Lambda-Red recombinase based insertion of a PCR derived *proB* gene with the SNPs *in situ.* This strategy involves the production of Δ*proB* prior to the insertion of mutant *proB* genes. As a bonus this Δ*proB* mutant may also be useful in increasing internal *E. coli* proline. The ability of a proline auxotroph (Δ*proB*) to accumulate proline from the environment following hyperosmotic shock has also been demonstrated previously (Buechter et al., 2003). Although no value in osmotic shock was found here, complete removal of the *proB* gene may be of value; while production of proline via this metabolic route will be inhibited, it may have a positive consequence elsewhere if other pathways or transporters compensate for this.

6.2.2. Removal of proline catabolism: PutA

putA encodes a multifunctional enzyme which catalyses both of the proline catabolism reactions in the proline metabolic network (Ling et al., 1994). The two enzyme activities are proline dehydrogenase and P5C dehydrogenase. Introducing a *putA* knockout in a proline overproducing strain (*proB* D107N) was shown to increase intracellular proline levels further (Csonka, 1988). *putA* has also been shown to inhibit expression of the proline transporter *putP* which transports proline into the cell from the extracellular environment (Nakao et al. 1988; Falcioni et al. 2013) therefore its removal may increase proline transportation into the cell.

6.2.3. Removal of proline catabolism for other metabolic processes: GdhA

Following PutA initiated catabolism of proline and it's metabolic precursor 1-pyrroline-5-carboxylate, a source of metabolite flow is through the glutamate pathway via glutamate dehydrogenase (GdhA). There are two routes to the production of glutamate in *E. coli* and when nitrogen is abundant (as it would be in LB media) the glutamate dehydrogenase pathway

is implemented to yield glutamate. It would seem logical that loss of metabolites from this network into other cellular metabolic networks would not be conducive to the production of proline (Kumar et al., 2010). Furthermore as second route to glutamate production is present cells will not become glutamate autotrophs.

6.2.4. Mutagenesis for increased intracellular proline: *proB, putA, gdhA*

Once potential strain modifications had been identified which may result in an increase the amount of internal proline, these strategies were implemented by gene knock out or insertion mutagenesis. Δ*CKL* was used as the parent strain, so that COL1A1 secretion could be investigated throughout. These were Δ*proB, proB D107N, proB E153A,* Δ*putA* and Δ*gdhA.* Combination of these strategies was also implemented with P1 phage transduction which led to the generation of Δ*putA* Δ*proB* and Δ*putA* Δ*gdhA.*

6.2.4.1. Knockout mutagenesis via the Lambda Red recombinase method: *proB, putA, gdhA*

Primers were designed to amplify the FRT flanked resistance genes from the appropriate template plasmid. Successful PCR reactions were confirmed by the presence of product DNA which corresponded to the anticipated PCR product size (Table 6.).

Table 6.12: PCR products for gene knockouts

Gene knockout	Antibiotic resistance	Template plasmid	Size of product (kbp)
proB	Km	pKD4	1.6
putA	Km	pKD13	1.4
gdhA	Km	pKD13	1.4

Positive results for all PCR experiments were achieved (Figure 6.5) -bands of the correct size were cut out and DNA isolated. Product DNA comprises of FRT flanked antibiotic resistance cassettes, with homologous ends to the up- and downstream regions of chromosomal DNA of the gene selected for knockout mutagenesis.

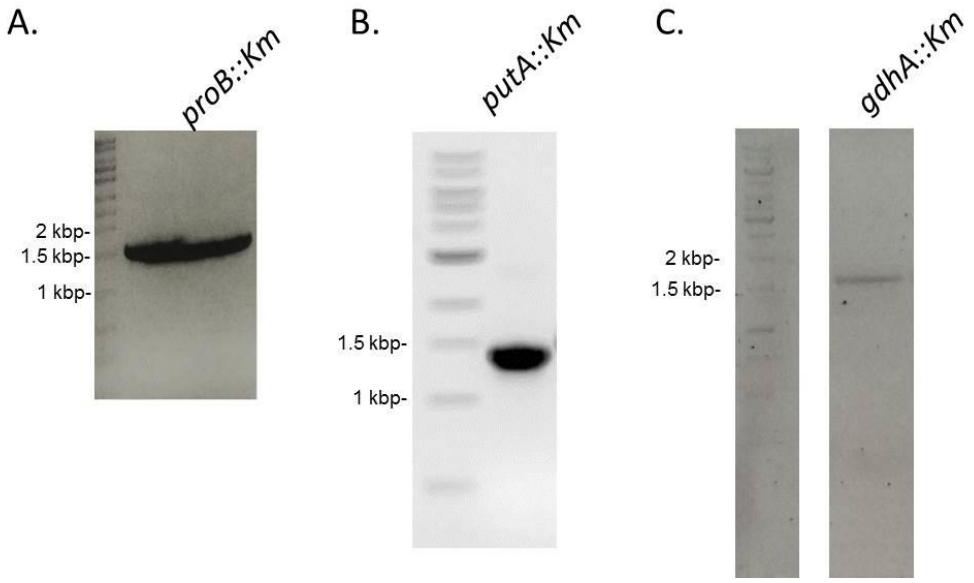


Figure 6.5: DNA gels of gene knockout construct PCR products.

PCR reactions were carried out using the templates and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples in A and C were run with GeneRuler™ 1kb DNA ladder; sample in B run with Quickload Purple 1kb ladder (NEB). PCR resulted in the formation of the gene knockout constructs harbouring antibiotic resistance cassettes (as listed in Table 6.). A. *proB*::Km, B. *putA*::Km, C. *gdhA*::Km.

Once isolated, the linear PCR derived DNA was electroporated into Δ*CKL* parent strains expressing Lambda Red recombinase enzymes. Successful recombination events were screened by plating on agar plates supplemented with kanamycin. Positive colonies were observed for all recombination strains and following the curing of the Red recombinase plasmid so that no further recombination events occur, the deletion of chromosomal genes was confirmed with PCR.

6.2.4.2. PCR conformation of knockout mutagenesis: *proB, putA, gdhA*

To confirm that the acquisition of antibiotic resistance was indicative of the removal of the gene targeted PCR reactions using combinations of primers which anneal either in the antibiotic resistance cassette or the region of chromosomal DNA either side of the target knockout gene. Either the size or the presence of the resulting PCR products allowed further confirmation of knockout mutagenesis.

Table 6.13: PCR products for confirmation of knockout mutagenesis in Δ*CKL*.

Positive results were obtained for the following knockout mutagenesis conformations. The size of PCR products which denote positive results is given along with the size of the PCR product (if relevant) expected if knockout mutagenesis was not successful.

Gene knockout	Antibiotic Resistance	Size of product if positive (kbp)	Size of product if negative (kbp)
proB	Km	1260	N/A
putA	Km	1465	4164
gdhA	Km	758	N/A

The size of the PCR product which would arise for primer pairs following successful knockout mutagenesis was predicted (Table 6.). Following colony PCR or successful transformants, positive PCR results were achieved for all proline metabolism gene knockout mutants.

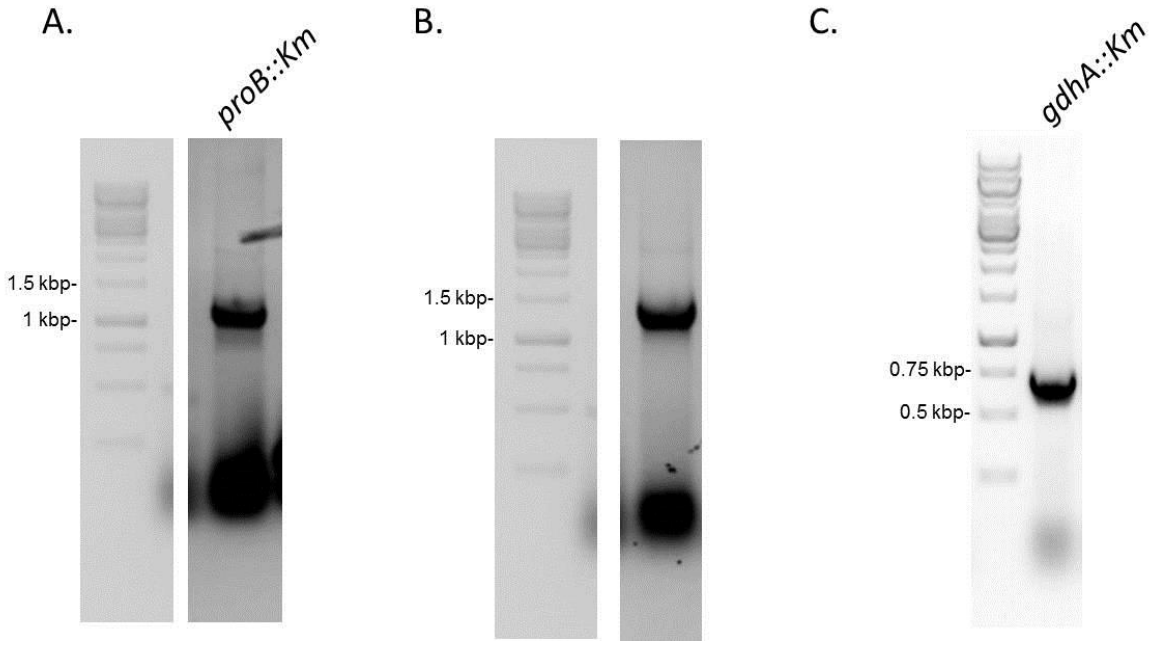


Figure 6.6: Agarose DNA gels of PCR products to confirm successful knockout mutagenesis.

PCR reactions were carried out using chromosomal DNA templates and primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of products of the sizes listed in Table 6.. A. *proB*::Km, B. *putA*::Km, C. *gdhA*::Km.

6.2.4.3. Knockout mutagenesis via the P1 phage transduction method

While Lambda-Red recombineering is the preferred method for knockout mutagenesis (due to the absence of phage), the efficiency is poor and therefore it can take multiple attempts of the protocol to generate mutants- when necessary P1 phage transduction was carried out to achieve knockout mutants in a short space of time. This resulted in the successful generation of the combination strategy mutants (Δ*CKL*) Δ*putA* Δ*proB* and Δ*putA* Δ*gdhA* from the Δ*CKL* Δ*putA* parent strain following the removal of FRT linked antibiotic resistance using the Flp recombinase gene harbouring pCP20 plasmid. P1 phage lysates were derived from Δ*CKL* Δ*proB* and Δ*CKL* Δ*gdhA.*

Once successful transformatants were obtained cells were passaged on sodium citrate plates, to eradicate phage. Growth was always normal in these strains, suggesting that phage was

eliminated from these strains. Care was taken to ensure that there was no possibility of existing genes knockouts in the parent strain being reinstated by chromosomal DNA packaged by the P1 phage during homologous recombination. The chromosomal location of genes targeted for knockout mutagenesis are listed in Table 5.. P1 phage is capable of packaging and initiating the homologous recombination of up to 100,000bp of donor chromosomal DNA into a parent strain. However *proB* and *gdhA* are not within this limit, therefore additional conformation of existing gene knockouts of the parent strain was not required.

Table 6.3: Chromosomal location (bp) of genes which underwent knockout mutagenesis and existing gene knockouts.

Gene	**Chromosomal location**	
	Start	**End**
Δ*fliC*	2053401	2054069
Δ*flgKL*	1140598	1140943
Δ*putA*	1072681	1076643
Δ*proB*	242522	243625
Δ*gdhA*	1877223	1878566

It was hoped that either Lambda-Red recombination or P1 phage transduction could be utilised to also produce (Δ*CKL*) Δ*proB* Δ*gdhA* and (Δ*CKL*) Δ*putA* Δ*gdhA* Δ*proB* however this was not successful in the time.

6.2.4.4. Site directed mutagenesis

To implement the D107N and E153A nucleotide substitutions, an extension of the Lambda Red recombinase method was implemented, to enable site directed mutagenesis. The *proB* gene was amplified from the genome of Δ*CKL* using primer pairs; one primer was homologous to the

region where the nucleotide substitution was to be made (i.e. residue 107 or 163). This primer contained an single nucleotide polymorphism to implement the following nucleotide alteration at: codon 107 (gat -> aat, D -> N) or codon 143 (gag -> gcg, E -> A). Overlap PCR was carried out with primers to yield a PCR product which harboured a chloramphenicol resistance cassette and the *proB* gene with the given SNP. As the *proB* gene knockout mutant in the Δ*CKL* background has been derived, this PCR construct was inserted into this parent background by a second round of Lambda-Red recombineering. This resulted in the excision of the kanamycin resistance gene and the insertion of the *proB D107N::*Cm or *proB E153A::*Cm cassette, which could be screened for by plating on chloramphenicol supplemented LB agar plates. The PCR reaction carried out are outlined in Figure 6..

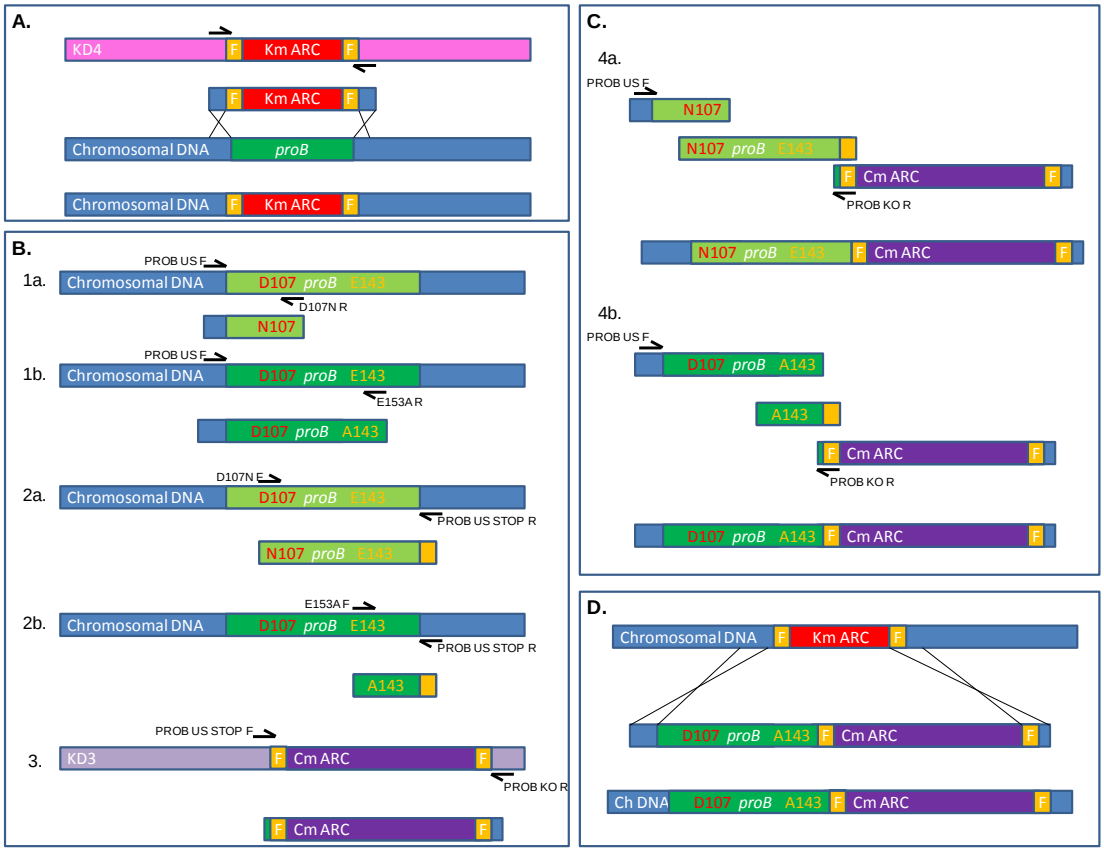


Figure 6.7: Strategy to implement *proB* substitutions by overlap PCR.

Schematic of site directed mutagenesis implemented by Lambda Red recombineering and overlap PCR. (A) *proB* knockout mutagenesis. A DNA construct which harbours the kanamycin resistance cassette (Km ARC) with ends homologous to chromosomal DNA up- and down-stream of *proB* is produced by PCR with custom primers and pKD4 as a template. This construct was inserted into MC1000 Δ*CKL* chromosomal DNA by the Lambda Red recombineering method. (B) PCR reactions to create DNA fragments necessary for overlap PCR for site directed mutagenesis. Template was MC1000 Δ*CKL* for 1-2 and pKD3 for 3. D107A and E143A primers contained single nucleotide polymorphisms so that nucleotide substitutions would be implemented during extension (C) PCR products from B are the templates for C. (C 4a) DNA templates from reactions B1a, B2a and B3 are assembled by overlap PCR. C 4b) DNA templates from reaction B1b, B2b and B3 are assembled by overlap PCR. D) The products of B are incorporated into the chromosomal DNA of MC1000 Δ*CKL* Δ*proB* by the Lambda Red recombineering method. NOTE: all primers and PCR reactions are listed in Appendix 2 and 4. Yellow boxes indicate non-antibiotic resistance coding sections of pKD3. 'F' indicates the flp recombinase target (FRT) site of this coding region.

To implement the D107N and E143A substitutions in Δ*CKL* Δ*proB* site directed mutagenesis was carried out. PCR reactions were carried out to produce the products shown in Figure 6.. PCR reactions were considered successful if bands of the sizes listed in (

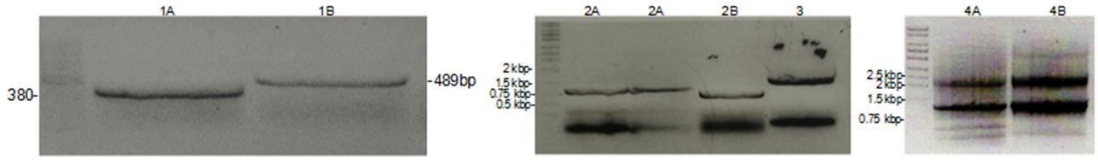


Figure 6.C). Although bands at additional sizes were present the 2.2 kbp products were excised. The bands should be the DNA constructs of the *proB* gene with an SNP at either codon 107 or 153, adjoined to a chloramphenicol cassette which is flanked by FRT sites. To ensure purity and to verify the identity of these bands, the 2.2 kbp bands, were excised, isolated, cloned into pJET2.1 and transformed into NEB 5-alpha Competent *E. coli* before sequencing – both constructs proved to align to the predicted sequences, including the nucleotide substitutions.

Table 6.4: Size of PCR products in overlap PCR to generate D107N and E143A *proB* substitutions

PCR reaction	PCR product (kbp)
1A	380
1B	489
2A	830
2B	725

3	1099
4A	2211
4B	2211

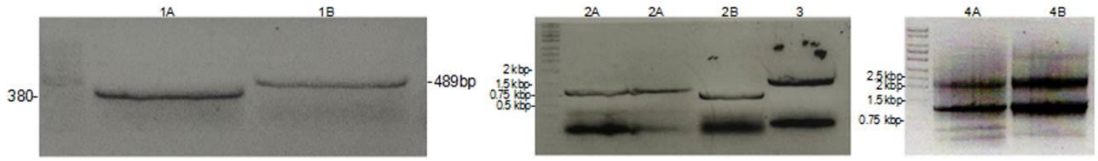


Figure 6.8: DNA gels of DNA template PCR products for overlap PCR to implement single nucleotide substitutions.

PCR reactions were carried out using the templates and primers in Appendix 4 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of all DNA templates for overlap PCR (1A-4B, see Table 6. for reference)

Following conformation by sequencing, FRT-ABC-*proB D107N/E153A-FRT* was amplified from pJET2.1 with the primers used to generate the overlap product initially. This linear DNA was

then into electroporated into Δ*CKL* parent strains expressing Lambda Red recombinase enzymes. Successful recombination events were screened by plating on agar plates supplemented with chloramphenicol. Positive colonies were observed for all recombination strains and following the curing of the Red recombinase plasmid so that no further recombination events occur, loss of kanamycin resistance was confirmed in both strains and the insertion of the *proB D107N* or *proB E153A* genes was confirmed with PCR (Figure 6.) and sent for sequencing. As the product was so long, multiple sequencing reactions were carried out with different primers which annealed throughout the PCR product. Sequencing revealed that in the colonies isolated the D107N substitution has not been implemented, however the E153A substitution had.

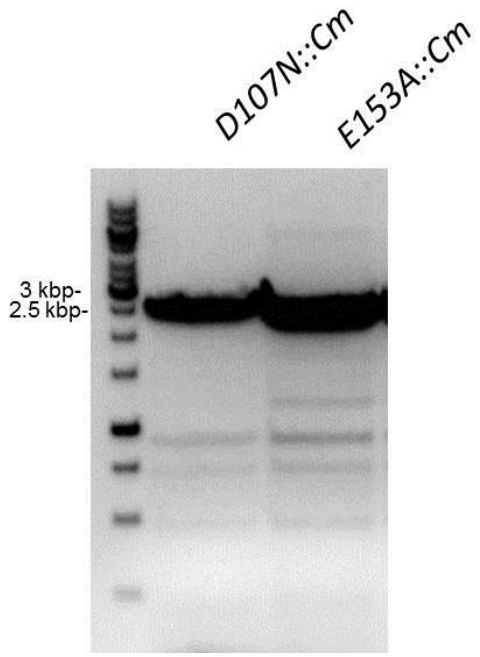


Figure 6.9: Chromosomal DNA amplified from the chromosome of *proB D107N* and *proB E153A*

PCR reactions were carried out chromosomal DNA from single colonies of potential *proB D107N* and *proB E153A* mutants using Wizard® Genomic DNA Purification Kit. Primers in Appendix 2 with Phusion® High-Fidelity DNA Polymerase. PCR mixtures were analysed on a 1% TAE agarose DNA gel supplemented with a trace of ethidium bromide, visualised under UV light and inverted. Samples were run with GeneRuler™ 1kb DNA ladder. PCR resulted in the formation of 2488bp DNA for all reactions.

6.3. Collagen expression in strains with proline metabolism based improvements for increased intracellular proline

With the strain improvements designed to result in increased intracellular proline confirmed COL1A1 expression was measured in all of the strains. As the hyperosmotic shock protocol was shown to have a negative effect of the expression of COL1A1, this strategy was not implemented in these strains. The intracellular fraction protein from cell cultures all looked uniform (Figure 6.A), no protein was detected in the supernatant fractions of cell cultures following Coomassie stain –not shown. The relative chemiluminescent signal detected for each sample following antibody probing with either anti-FLAG-HRP or anti-COL1A1 was very similar, suggesting both gave an accurate representation of the concentration of COL1A1 protein in each sample (Figure 6.B and C). No signal was detected in the cells of the Δ*CKL* strain expressing empty secretion construct. In was evident that COL1A1 was not expressed in Δ*proB E153A* or Δ*putA.* The strains which expressed the least COL1A1 were Δ*CKL,* Δ*putA* Δ*gdhA* and Δ*proB* Δ*putA.* Both Δ*proB* and Δ*gdhA* expressed the highest concentration of COL1A1.

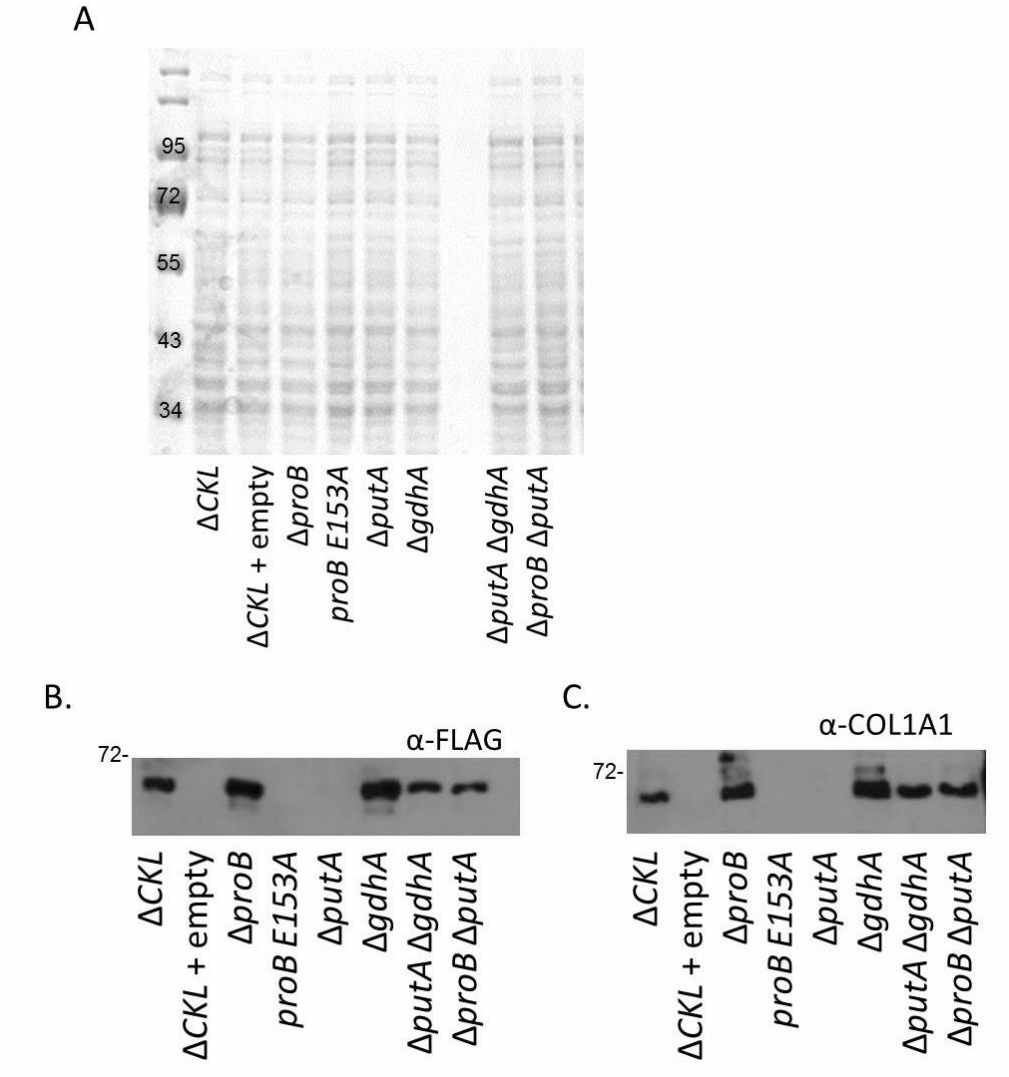


Figure 6.10: Intracellular and secreted fractions of Δ*CKL* expressing either empty or COL1A1 harbouring secretion construct and proline metabolism mutant strains: Δ*proB*, *proB E153A*, Δ*putA*, Δ*gdhA*, Δ*putA* Δ*gdhA* and Δ*proB* Δ*putA*

E.coli Δ*CKL* harbouring either the plasmid pJexpress-FliC-COL1A1-FLAG-Strep or pJexpress-FliC-empty-FLAG-Strep (empty) and (Δ*CKL*) Δ*proB*, *proB E153A*, Δ*putA,* Δ*gdhA,* Δ*putA* Δ*gdhA* and Δ*proB* Δ*putA* expressing pJexpress-FliC-COL1A1-FLAG-Strep were grown in LB overnight at 37°C, 180 rpm, following induction with 1mM IPTG cells were incubated for three hours and then harvested. 1 OD unit of cells were prepared for SDS-PAGE and the supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western blot analysis. (A) Coomassie stained SDS-PAGE of the intracellular fractions of cell cultures, including one lane left blank. Western blot of the intracellular fractions probed with (B) anti-FLAG-HRP and (C) anti-COL1A1 and appropriate secondary antibody.

Densitometry analysis was carried out based on two biological replicates, probed with either anti-FLAG-HRP or anti-COL1A1. The reported relative densitometry for strains in comparison to that recorded for Δ*CKL* resulted in a similar pattern of results. Δ*ghdA* expressed the most collagen (1.47 and 1.70 fold more than Δ*CKL* respectively) followed by Δ*proB.* Δ*CKL* and Δ*putA* Δ*gdhA* expressed similar concentration of COL1A1. When calculating using anti-FLAG-HRP it was show that Δ*putA* Δ*proB* did not express a high concentration of COL1A1, however with anti-COL1A1 it was calculated to result in 1.39 fold more intracellular collagen. As less variability was observed using the anti-COL1A1, these results may be more reliable.

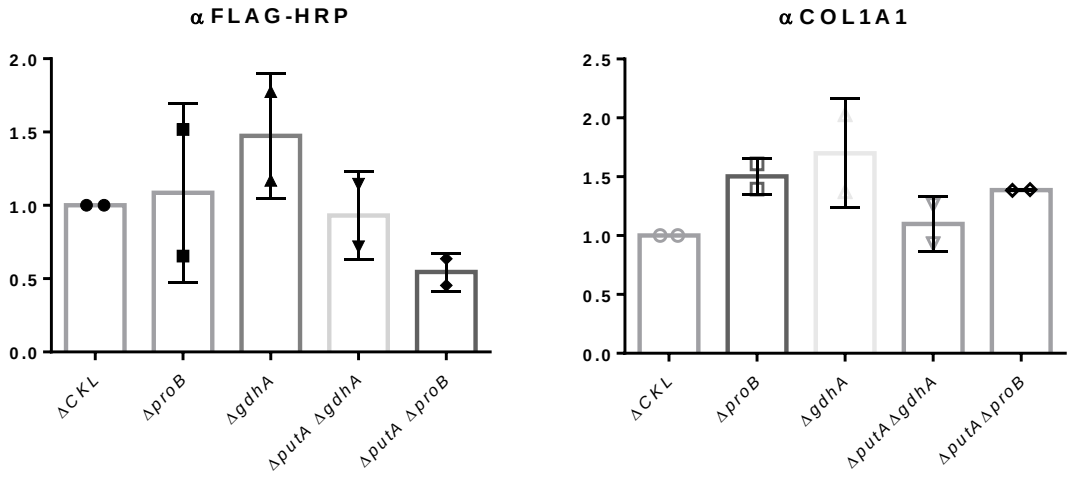


Figure 6.13: Results following densitometry analysis of the intracellular samples of anti-FLAG-HRP and anti-COL1A1 probed Western blots of Δ*CKL,* Δ*proB,* Δ*gdhA,* Δ*putA* Δ*gdhA* and Δ*proB* Δ*putA* COL1A1 expressing cells following cell culture

Following cell culture and Western blot analysis as described in the figure legend of Figure 6., densitometry analysis was carried out on Western blots following probing with anti-FLAG-HRP or anti-COL1A1 and appropriate secondary antibody. Results from two biological replicates, normalised to the result for Δ*CKL* in each repeat. Strains which expressed no COL1A1 are not shown. One-way ANOVA: no significance.

6.4. Secretion of a bacterial collagen like protein in the FT3SS secretion strain

Secretion of COL1A1 was not observed, however it was of interest to investigate the capacity of the prototype and mark 2 secretion strains to express and secrete a collagen like protein to establish whether it was amenable to secreting this kind of protein. As an alternative to the 45kDA COL1A1 expression and secretion of the *Streptococcus pyogenes* collagen-like surface protein 2 (Scl2) was investigated. The protein harbours the Gly-X-X triplet amino acid motifs, characteristic of human collagen and has previously been expressed in *E. coli* with yields of 10g L^{-1} at high cell density - 0.3g L^{-1} at low- but not secreted (Chen et al., 2010; An et al., 2014; Peng et al., 2012). There is suggestion that these collagen-like proteins could also be used for therapeutic applications, for example it has been incorporated into a hydrogel and found to aid wound healing in rats (Cereceres et al., 2015). Prior to the work in this thesis, the Scl2 protein (Accession: AY069936) was cloned into the prototype secretion construct with the *Eco*RI and *Pst*I restriction enzyme sites as with COL1A1 in Figure 5.. Expression was induced with 0.05mM IPTG, which was in-keeping with the expression protocol in Chapter 4 and 5. The predicted size is 22.6 kDa for Scl2 or 32kDa in the secretion construct. Following expression in Δ*CKL* and Δ*CKL* Δ*clpX* Δ*motAB,* no protein was seen in the Coomassie stained secreted fraction of cell cultures (Figure 6.14A). Western blot analysis with anti-FLAG-HRP found that Scl2 was both expressed and secreted in both strains; around twice the concentration of Scl2 was seen in the supernatant of Δ*CKL* Δ*clpX* Δ*motAB* cells (Figure 6.14B). The Scl2 protein ran at a slightly larger size in the supernatant as opposed to the intracellular fraction. The size of the band in the intracellular fraction was closer to the expected size. When stripped and probed with anti-GroEL, similar concentrations of GroEL were seen in the secreted fractions of both strains.

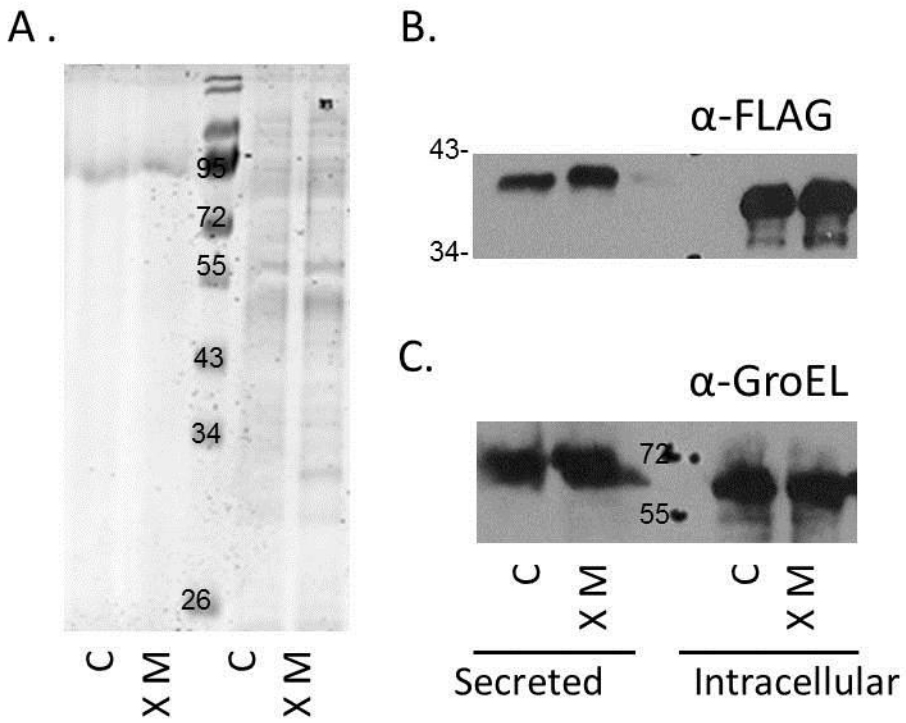


Figure 6.14: Intracellular and secreted fractions of Δ*CKL* or Δ*CKL* Δ*clpX* Δ*motAB* expressing Scl2.

E.coli Δ*CKL* (C) or Δ*CKL* Δ*clpX* Δ*motAB* (XM) harbouring pJexpress-FliC-Scl2-FLAG-Strep were grown in LB overnight at 37°C, 180 rpm, following induction with 0.05mM IPTG cells were until they reached OD 1.0. 1 OD unit of cells were prepared for SDS-PAGE and the supernatant from 1 OD unit of cells was precipitated with TCA (10% v/v) before SDS-PAGE and Western blot analysis. (A) Coomassie stained SDS-PAGE of the secreted and intracellular fractions of cell cultures. Western blot of the intracellular fractions probed with (B) anti-FLAG-HRP and (C) anti-GroEL and appropriate secondary antibody. Samples loaded: Supernatant: 20µL. Cells: 5µL or 2µL for Coomassie and Western respectively

Densitometry analysis based on two biological replicated showed that on average cells from both strains expressed similar concentrations of intracellular Scl2, however Δ*CKL* Δ*clpX* Δ*motAB* cell culture resulted in 1.49 times more Scl2 protein being secreted into the media (Figure 6.15).

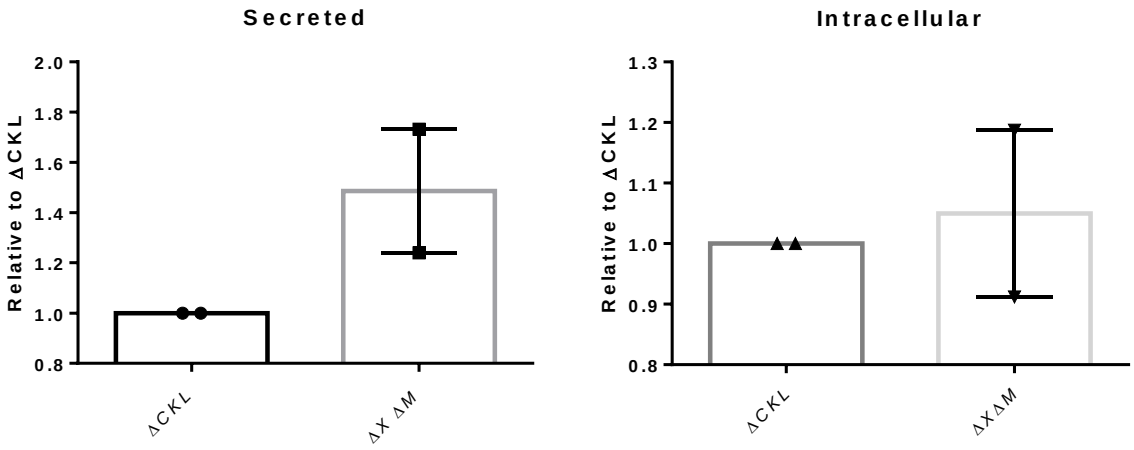


Figure 6.15: Results following densitometry analysis of the intracellular samples of anti-FLAG-HRP probed Western blots of Δ*CKL* (C) and Δ*CKL* Δ*clpX* Δ*motAB* (XM) Scl2 expressing cells following cell culture

Following cell culture and Western blot analysis as described in the figure legend of Figure 6.14, densitometry analysis was carried out on Western blots following probing with anti-FLAG-HRP. Results from two biological replicates, normalised to the result for Δ*CKL* in each repeat. One-way ANOVA: no significance.

Although not attempted here it would be beneficial to attempt to purify either of these collagen based products using the Strep-tag and TEV cleavage sites. This would allow investigation as to whether the purified product would have any beneficial properties in terms of cell adhesion, if this was found steps could be taken to utilise the collagen product to investigate wound healing for example.

6.5. Discussion: Collagen expression for FT3SS

Collagen has many important therapeutic applications and is currently obtained from animal sources, therefore there are issues with safety and biocompatibility (Olsen, 2003). This has led to the generation of expression strains to generate recombinant collagen for use in therapeutics (Nuutila et al., 2015). Most of these expression systems involve whole plants, animal cells or *Pichia* (which are not reported to secrete the collagen produced). An *E.coli* based system could improve many of the issues associated with production of collagen from animal sources or these cell lines. Previously this has been achieved with a 23 and 38kDa chain of human pro α (I) or (III) collagen chain fused to glutathione S-transferase (Buechter et al., 2003; Rutschmann et al., 2014), however this thesis aimed to express and/or secrete a 45kDa chain of human pro α (I) chain. As collagen is a proline rich protein and *E. coli* do not accumulate proline, much emphasis was put on improving the concentration of intracellular proline.

6.5.1. Expression of human COL1A1 in the FT3SS secretion strain

Initially efforts to express COL1A1 were linked to investigating whether expression levels could be increased with hyperosmotic shock. This has previously been shown to increase the concentration of intracellular proline and also aid the production of a fragment of collagen (Buechter et al., 2003). While COL1A1 was expressed in both strains, the use of hyperosmotic shock led to a reduction in the concentration of COL1A1 seen in cells by around half (Figure 6.C and D and Figure 6.). This may have been due to the fact that the report in the literature was linked to a proline feedback mechanism knockout mutant (Δ*proB),* therefore there may have been limited effect in a +*proB* strain as additional proline transported in the cell due to hyperosmotic shock may have been immediately led to feedback inhibition and therefore less intracellular proline. Alternatively the addition of NaCl may have caused the cells fitness to be reduced, however (slight) cell lysis was only observed in cells expressing the empty secretion construct following osmotic shock (Figure 6.A). Regardless of the effect of hyperosmotic shock, this result does represent the largest fragment of human collagen to be expressed in *E.*

*coli.*COL1A1 was not detected in the secreted fraction of cells, therefore steps were required to improve this. It was decided that first improving expression may aid this and that the best route to achieve this would be to increase intracellular proline concentration.

6.5.2. Proline metabolism strain improvements

The hyperosmotic shock method was not successful in improving intracellular COL1A1 concentrations; therefore strategies to improve intracellular proline which was intrinsic to the cell metabolism were implemented. Mining the literature and surveying the proline metabolic network resulted in the generation of three strategies. This was to remove negative feedback of intracellular proline accumulation and could be implemented by removing *proB* or generating strains with an SNP in the *proB* gene (Buechter et al., 2003; Csonka et al., 1988; Rushlow et al., 1985). Another strategy involved the removal of the dual function *putA* gene which catalyses two reactions which catabolise proline or the proline precursor 1-pyrroline-5-carboxylate. Removal of *putA* has been shown to increase intracellular proline levels and result in increased proline transporters, which move extracellular proline into the cell (Csonka 1988; Nakao et al. 1988). Finally it was hypothesised that the removal of *gdhA* would result in less proline precursor metabolites being directed to other metabolic processes.

Aside from the *proB D107N* mutant all of these strains were generated and Δ*putA* Δ*proB* and Δ*putA* Δ*gdhA* were successfully combined. It was shown that the Δ*proB E153A* and Δ*putA* mutants did not express COL1A1 (Figure 6.B and C). The intracellular protein fraction of these strains did not look altered compared to controls (Figure 6.A) and growth rates seemed to be unaffected- as observed by similar final OD_{600} – indicating these cells were not perturbed in their fitness in this media. However the other strains all produced COL1A1 (Figure 6. B and C). Results for densitometry was less variable when calculated based the Western bolts derived by the anti-COL1A1 antibody as opposed to anti-FLAG-HRP, furthermore as this was an antibody raised against a motif of residues which is present in the COL1A1, it is specific it also provides more evidence of quality control of COL1A1 production as the antibody was raised against an 88 amino acid section of COL1A1, which is present in the N-terminus of the protein and therefore present in the 45kDa COL1A1 protein. For these reasons I will describe results based on the findings following anti-COL1A1 probing of the nitrocellulose membrane. The highest

increase in intracellular COL1A1 was observed in the *gdhA* knockout mutant, resulting in an average of 1.7 fold more expression. The next was in Δ*proB,* with a 1.5 fold increase Δ*putA* Δ*proB* and ΔputA Δ*gdhA* were the next. All strain improvements to proline metabolism (which expressed COL1A1) resulted in increased intracellular COL1A1 –while it is not known if they actually resulted in an increase in intracellular proline, it does suggest potential. Without the full set of combinations or information on the intracellular proline concentrations of strains it is difficult to speculate on the effect of each knock out mutation strategy, however it would seem that the addition of Δ*putA* results in a reduced concentration of COL1A1 (i.e. Δ*putA* Δ*gdhA* and Δ*putA* Δ*proB,* lower than Δ*gdhA* and Δ*proB* alone. This is surprising since the literature would predict that a putA mutant would accumulate excess proline (Csonka et al. 1988; Nakao et al. 1988). This may allude to the fact that intracellular proline levels do not alter COL1A1 expression at all –however as discussed more information is required to confirm this.

As metabolic networks are complex it is difficult to pull apart the reasons for the success or failure of these strategies to result in increased COL1A1 production. A good starting point would be to see in these mutants did result in variation in intracellular proline levels. This could be achieved by GC-MS or more traditional methods, such as detection with ninhydrin. Ninhydrin reacts with free amines to yield a colour change, while it reacts with all amino acids, proline is the only amino acid to result in an absorption peak at 520nm (Paes et al., 2013). However if GC-MS was combined with an internal standard this could provide quantitative measurements of intracellular proline. GC-MS was attempted in this body of work, although results required further optimisation as results were 'noisy' and masked the presence of proline in biological samples, this could be improved with improved sample preparation i.e. more biological sample, use of minimal media, not LB.

Although improvements were made in the concentration of COL1A1 found in cells, it was not detected in the secreted fractions of cell culture. This has not previously been an issue for other substrates and it is unclear whether FT3SS secretion of COL1A1 is ultimately not possible or whether induction conditions or secretion capacity could be modified to facilitate this. To gain an idea of whether the FT3SS is capable of secreting COL1A1 expression and secretion of a collagen-like protein was investigated.

6.5.3. Secretion of a collagen-like protein through the FT3SS

Scl2 was successfully expressed and secreted in both the prototype and mark 2 secretion platform strains. Despite similar intracellular concentrations more was secreted in the mark 2 strain suggesting the improved secretion capacity characteristics of the Δ*clpX* Δ*motAB* applied for this substrate. The size of the band in the intracellular fraction was smaller than the one observed in the extracellular fraction of protein. This may be due to proteolysis occurring to the protein in the cell, as this is commonly reported when recombinant protein is expressed in the cytoplasm. The use of mass spectrometry could aid investigation into whether this was the case. As an alternative, intracellular fractions could be extracted following a shorter period of time to investigate whether less small protein was detected with Western blotting.

This experimental work demonstrates that the FT3SS is capable of secreting protein with the repeating Gly-X-X triplet amino acid motifs, characteristic of human collagen, therefore the most likely reasons for the absence of secreted protein is either the size of COL1A1 (45kDa, compared to 35.5kDa or that following the overnight induction protocol proteolytic enzymes accumulate in the supernatant and secreted COL1A1 is degraded (degradation was seen following overnight secretion of E2 enzyme in Chapter 3. It would be advantageous to alter the expression protocol for COL1A1 to observe whether secreted protein could be observed. Finally Scl2 is itself relevant in terms of therapeutics (An et al., 2014; Cereceres et al., 2015), therefore this demonstrates an additional use of the FT3SS to secrete an industrial biotechnology relevant protein. As stated previously is would be beneficial to purify collagen like proteins to investigate adhesive and healing properties and also to gather information on yield.

6.5.4. Further improvements for collagen secretion through the FT3SS

As discussed the use of GC-MS would be advantageous in terms of investigating whether intracellular proline is important in the expression of proline rich recombinant proteins. As it is not known if the mutants implemented in the strains which resulted in increased COL1A1 expression resulted in increased proline or whether increased COL1A1 expression was as a

result of another factor. The use of GC-MS could elucidate this and although it could not confirm a direct link, the identification of a correlation for increased proline and increased COL1A1 would be compelling.

In an effort to generate strains which could express (and ultimately) secrete COL1A1, it would be advantageous to produce the full set of mutants which were identified here. As knockout mutations resulted in improved expression it would be interesting to combine all three strategies and also to implement the ones which were not possible (*proB D107N* and Δ*proB* Δ*gdhA*). It would be interesting to see whether *proB D107N* Δ*proB* resulted in an increase in COL1A1, as this strain was reported to produce high intracellular proline levels (Csonka, 1988). Although it is not known if increased proline increased COL1A1 production in this strain other strain improvements may include upregulating expression of the transporters in an effort to improve proline transport into the cell. An ideal target for this would be the tightly regulated PutP transporter. As it is known that PutA downregulates PutP, it would be advantageous to remove this gene and at the same time engineer a promoter replacement for a strong promoter upstream of *putP.* In addition to these strategies use of flux balance analysis would be helpful in both confirming the strategies already implemented and also identifying new routes to improving intracellular proline abundance. Flux balance analysis optimises the flow of metabolites through a system, based on an objective function. In the case of this work the objective function would be: maximisation of intracellular proline. Gene knockouts can be implemented and the change in flux through each reaction can be measured. This can be used to identify novel routes for metabolic network improvement which may not be in the literature and also identify any unforeseen cell consequences of initiating gene knockouts, which may affect cell fitness (objective function: biomass). Models do not necessarily represent the biology and this affects predictability, so care must be taken to ensure the model is as biologically representative as possible. For example the Reed genomic scale model (Reed et al., 2003) allows proline transport into the extracellular space of all three proline transporters, however this is not the case, therefore the model should be adjusted.

In terms of secreting COL1A1, as discussed in the previous section alteration of the expression protocol may result in the detection of COL1A1 in the supernatant. Furthermore integration of the knockout mutants which improve both secretion and intracellular COL1A1 expression could be combined (i.e. ΔCKL Δ*clpX* ΔmotAB Δ*gdhA*), to see if secretion could be achieved. In terms of improving yields it may be beneficial to grow cells in richer media for a more

prolonged period of time. For example Peng et al. (2012) expressed 1g L^{-1} Scl2 when cells were grown to OD_{600} 1.0, whereas around 10g L^{-1} Scl2 was achieved following growth in excess of OD_{600} 100. It would also be interesting to express COL1A1 in cells expressing prolyl-4-hydroxylase and lysyl-hydroxylase-3, to investigate whether this would be an effective means of producing collagen with correct the correct hydroxylation modifications of prolyl and lysyl residues.

Finally as a therapeutic use is the endpoint goal of expressing COL1A1 and to an extent Scl2, it would be beneficial to purify some of this protein and investigate functionality, as this is essential if the protein is to be used in a therapeutic setting. This could involve investigating whether isolated recombinant collagen proteins are able to associate with the extracellular matrix of mammalian cells –as collagen I does *in vivo* or whether it can improve wound healing in an animal model. Although collagen I formation requires COL1A2 in addition to COL1A1 to form a triple helix, COL1A1 will make trimeric helixes; these are resistant to proteolysis, therefore the formation of triple helixes can be assessed by challenging isolated COL1A1 with proteolytic enzymes, if it is resistant this is evidence that triple helixes formed (Buechter et al., 2003; Howes et al., 2014).

Chapter Seven: Discussion and future prospects

7.1. Summary of major findings

The work carried out in this thesis describes attempts at generation of a high capacity FT3SS *E. coli* secretion machine for directed secretion of a number of proteins into the extracellular media of cell cultures. Modification of the FT3SS, coupled with the generation of secretion constructs allowed the potential of this secretion platform to begin to be explored. This secretion platform was developed due to the requirement of industrial biotechnology for processes which are low in both cost and time in terms of growth requirements and reduced downstream processing, yet result in high quality products. Downstream processing can account for up to 80% production costs; continuous culture systems which do not require cell lysis to retrieve product are attractive in terms of cost reduction, therefore secretion into the extracellular media is desirable (Walsh, 2014). Furthermore when extruded from the cell cytoplasm protein product quality can be improved due to reduced proteolysis and correct protein folding (Baeshen et al., 2015). Efforts to secrete protein are usually confined to the periplasm of cells, while the aforementioned beneficial characteristics are achieved here, this can be improved further with secretion to the extracellular space, while at the same time simplifying downstream processing (Choi & Lee, 2004). Secretion into the extracellular space has been achieved in a number or organisms and cell lines (including mammalian cell lines e.g. CHO and HEK, yeasts e.g. *Pichia pastoris* and *Saccharomyces* and bacteria e.g. *Pseudomonas* and *Bacillus*), however the use of *E. coli* would be favoured due to the low growth requirements, high growth rate and industry acceptability (Baeshen et al. 2014; Yoon et al. 2010; Ferrer-Miralles & Villaverde, 2013). The caveat to this is that while secretion to the periplasm has been exploited in *E. coli* the implementation of extracellular secretion is poor. Therefore the basis of this thesis was to generate an *E. coli* based platform to implement extracellular secretion of protein; specifically this concerned the modification of the FT3SS into a high capacity secretion system. Drawing on reports of recombinant protein secretion through various FT3SS mutants (Majander et al. 2005; Singer et al. 2012; Végh et al. 2006), it

was hoped that modifying the native high capacity extracellular secretion system, would result in high yields of secreted recombinant protein of industrial biotechnological relevance. Once applied, efforts were then made to improve the platform secretion system in terms of both improved yield and range of secreted protein products. Strategies to enable the production of this secretion platform drew on traditional microbiology strain improvement methods and synthetic biology concepts. As a prelude to the final discussion the major findings of each chapter of this thesis are summarised here.

7.1.1. Chapter Three: Strain improvements for increased secretion capacity of native substrates through the T3SS

The main focus of this chapter was to identify a platform *E. coli* strain to form the basis of a secretion platform. Once established a protein secretion assay was developed and steps were taken to improve secretion capacity by the implementation of strain improvements. The major findings of this chapter were:

- Using a plasmid based system, the expression and secretion of E2 (the native FT3SS substrate FliC with residues Δ191-280 removed) was investigated to establish that the HAPless secretion platform strain Δ*fliC* Δ*flgKL* (Δ*CKL*) had the capacity to secrete around twice the concentration of E2 protein in comparison to the CAPless strain: Δ*fliCD*. With the use of an E2 protein standard of known concentration, this was equated to 9.19mg L^{-1} secreted protein. While this was reasonably impressive compared to other efforts to secrete protein from the FT3SS, it was hoped that this could be improved to give this protein secretion platform a standing in industrial biotechnology applications.
- The HAPless strain was used herein to investigate both the intracellular and secreted concentrations of E2 protein. A protein secretion assay was developed to measure this effectively and accurately, this was paired with a gene expression assay and growth curves; together this allowed a wealth of information on the expression and secretion capacity, growth phenotype and gene expression profile of cells. Coupled with measurements at different times, this allowed a dynamic and temporal perspective on these factors.

- Gene knockouts were implemented using Lambda-Red recombinase methods, to generate negative control strains, which were either void of FT3SS apparatus or locked in early substrate secretion (E2 is late). With these it was demonstrated that protein observed in the supernatant was a consequence of directed protein secretion and not cell lysis or membrane leakage.
- Gene knockouts were then identified based on information reported in the literature with a view to improving the secretion capacity of the platform secretion strain. These were based on strategies to remove negative feedback of flagella gene expression or to reduce the metabolic burden.
- Not all strategies resulted in increased protein secretion. The presence of an IS5 mutation meant that *lrhA* was already ineffective at regulating flagella gene expression in the wildtype. The removal of *dksA* resulted in decreased protein secretion, presumably due to the fact that it is linked to the stringent response, which was not in effect here as experiments were carried out in minimal media
- The removal of *clpX* resulted in improved secretion of flagella both in strains with functional flagella (improved motility and abundance of flagellin monomers) and in terms of secretion with 23.457 mg L^{-1} E2 being detected in the supernatant of cell culture.
- Further knockout mutations also demonstrated promise in terms of improving secretion capacity, removal of *motAB* resulted in over twice as much E2 protein being secreted into the media.
- It was also observed that protein secretion substrate is not insulated to the effects of the cells, as little E2 was expressed in the Δ*CKL* Δ*fliDST* strain in the absence of the FliC chaperone FliS. This, coupled with the desire to assess the secretion capacity of the platform secretion system with reference to industrially relevant proteins, drove the requirement to secrete non-native proteins in the secretion system.
- Together these results demonstrated the great promise and potential of the FT3SS based secretion platform for extracellular protein secretion.

7.1.2. Chapter Four: Further development and testing of the engineered FT3SS system: non-native protein secretion and assay development

This chapter focused on integrating a protein secretion construct into the secretion platform to aid production of a range of recombinant proteins and processing of secreted product. In this chapter the expression and secretion of the recombinant antibody fragment CH2 was investigated in a number of strains designed to demonstrate increased secretion capacity. Major findings of this chapter included:

- The advent of an inducible secretion construct consisting of FT3SS motifs (the 5' and 3' UTRs and 47 amino acid secretion signal peptide) and features to allow detection, purification and isolation of cargo protein was effective and useful.
- In combination with the modified FT3SS secretion platform strain it was possible to express and secrete the first recombinant protein (the antibody fragment -CH2) in this system.
- True secretion was established by observing negative secretion in several control situations, i.e. strains lacking FT3SS or locked for early substrate specificity or secretion constructs without FT3SS signals.
- The protein secretion assay was modified slightly to enable measurement of intracellular and secreted CH2 protein in a number of cell cultures grown in tandem; this improved experimental output.
- Secretion capacity of CH2 harboured in the secretion construct was found to be around 1.5 times higher in all of the strains which had been designed to exhibit increased secretion: (Δ*CKL*) Δ*motAB,* Δ*flgMN,* Δ*fliDST,* Δ*clpX.*
- With the derivation of a CH2 protein standard of known concentration it was established that the best secreted yield was 219.66µg L^{-1} Δ*CKL* Δ*flgMN*. As a truly secreted protein from *E. coli* this was a reasonable yield, especially given the consideration that growth times and cell densities are much higher in industrial production of recombinant protein, thus skewing results.
- These results highlighted the potential of strain improvement based strategies , it was hoped that these improvements in secretion capacity would cumulate if the knockout mutagenesis strategies for improved secretion were combined. However this would require a more high throughput test

7.1.3: Chapter Five: Development of a high-throughput assay for FT3SS secretion to screen a multitude of secretion strains and plasmids

In this chapter a high throughput protein secretion assay was developed to improve the rate and accuracy at which secretion in the secretion platform could be assessed. This allowed screening of a large number of modifications to both the secretion platform strain and construct to be assessed, with the view of identifying a combination which would result in a high secretion capacity of the modified FT3SS. The main findings of this chapter included:

- Development of an enzyme based secretion assay involved modification of the secretion construct protein cargo to permit the expression and secretion of a recombinant cutinase. The secretion of an active enzyme was of particular interest as it demonstrated that correct protein folding occurred following secretion through the FT3SS.
- Development of a MUB based protein secretion assay was arduous; however rigorous optimisation and use of controls enabled an effective, accurate and high throughput fluorescent plate reader based secretion assay to be deployed.
- A combination of Lambda-Red recombinase and P1 phage transduction methods resulted in a large panel of 15 secretion strains, with the potential to demonstrate much higher secretion capacities than the prototype Δ*CKL* secretion platform strain. This involved the combination of up to three of the knockout mutations which had previously been shown to lead to increased secretion of E2 or CH2 -Δ*motAB,* Δ*flgMN,* Δ*fliDST* and Δ*clpX*.
- In conjunction the prototype secretion construct was modified to investigate whether the removal of any of the components of the *E. coli* derived secretion construct elements (5’ and 3’ UTR and the 47 amino acid secretion signal peptide) would improve expression of secretion. This required generation and testing of nine secretion construct variants.
- Following screening of secretion capacity with the MUB protein secretion assay it was established that secretion capacity (quantified by fluorescence output) was improved around 8 fold in the *ΔCKL ΔclpX ΔmotAB* secretion strain. In addition while a variant of the secretion construct which harboured only the 3’ UTR expressed around 2.2 times more cutinase than the prototype secretion construct, the best output in terms of

improved secretion was observed in the absence of the 3' UTR only –resulting in up to a 15 fold increase in secretion. Combined the 'mark 2' secretion strain and secretion construct, resulted in almost a 25-fold increase in secretion over the original prototype Δ*CKL* strain.

- Together this demonstrated the power of a high throughput secretion assay and that strategies to improve secretion capacity of the FT3SS could be combined. The mark 2 secretor strain alone was also shown to result in increased E2 (3.75 fold) and CH2 (2.56 fold) secretion compared to the prototype secretion strain. If this was combined with the mark 2 secretion construct the scope to secrete industrial biotechnology relevant concentrations of recombinant protein would be large

7.1.4. Chapter Six: 45kDa human collagen expression for FT3SS secretion

In the final chapter an effort was made to showcase the potential of the secretion platform for the production of a large eukaryotic protein with therapeutic applications: the 45kDa human pro α (I) collagen chain. Metabolic engineering was implemented to see if this could increase the yield of recombinant collagen expressed in the secretion platform. Major findings are given below:

- The 45kDa human pro α (I) collagen chain (COL1A1) was expressed from the secretion construct in the original prototype secretion platform strain. Although not secreted, expression of this protein was still extremely impressive, as it is the largest fragment of a collagen I chain to be expressed in *E. coli*.
- Efforts were made to increase the concentration of intracellular proline, in an effort to improve COL1A1 expression. While inducing osmotic shock was not effective, strain improvements were. Guided by the literature and the proline metabolic network, a number of knock out mutants were generated, the removal of glutamate hydrogenase (*gdhA*) from the metabolic network resulted in up to 1.5 times more collagen being expressed, however secreted COL1A1 was not detected.

- Finally to see if the modified FT3SS was capable of secreting a collagen-like protein, Scl2 was expressed in the mark 2 secretion strain; this resulted in a 1.5 fold improvement in comparison to the prototype secretion strain.
- In combination these results show that in the near future a secretion and proline maximised secretion platform may be utilised to secrete full length COL1A1 for therapeutic application.

In addition to the discussion points made in each chapter, the following themes emerged.

7.2. Improved secretion capacity of the FT3SS

The FT3SS had previously been modified to result in secretion of protein into the media with an industrial biotechnology goal in mind, however the work in this thesis is far more extensive in terms of modifying both strain and secretion construct than any work carried out previously (Majander et al. 2005; Singer et al. 2012; Widmaier et al. 2009). While the aforementioned studies investigated different secretion peptide signals, only one study modified a strain to enable increased secretion, which resulted in a 2 fold increase in secreted protein (Singer et al. 2012), a value surpassed in this thesis by over tenfold. Furthermore aside from the Majander et al. (2005) study, which concerned *E. coli,* all of these investigations involved *Salmonella,* therefore work has greatly furthered the understanding of recombinant protein secretion through the FT3SS in an industrially relevant organism. The yield of secreted protein achieved in this work bettered those achieved in these studies and when incubation time and cell density were considered, it represented a substantial improvement. For examples the highest achieved secreted yield was 12mg L^{-1} and involved growing cells for 24 hours as opposed to around 6 hours in this study -although the OD_{600} reached was not reported, it can be assumed that this would exceed the OD_{600} of 1.0-1.5 used in this thesis (Majander et al., 2005).

Although the strain based improvements to secretion capacity were considerable in this study it would be beneficial to continue to investigate if this could be improved further. It is envisioned that this would be facilitated by the MUB protein secretion assay, because the scope for screening mutants once derived is large. These points were made in the discussion of

Chapter 5, but to reiterate, the MUB protein secretion assay could be used to investigate the secretion capacity of a wide range of mutant strains in the Δ*CKL* background. Mutations could be implemented in a direct manor based on ideas from the literature or informed by information derived from modelling or proteomic analysis, which could identify bottlenecks to FT3SS secretion apparatus or recombinant protein production. Additionally use of a transposon library would allow a huge range of mutants to be screened with the view to finding novel candidates for high capacity secretion strains. A modification of the secretion construct could allow multiple traits to be selected for, for example if the secretion construct was induced by expression of *putP,* a high level of cutinase secretion would suggest both increased proline transport and high secretion capacity. This strain could be selected to secrete proline rich proteins, such as collagen. Additionally a library of secretion construct plasmids could be generated to investigate variation in secretion based on modifications to genetic elements such as the promoter and RBS and also the secretion signal or 5'UTR (which was shown here to be important for efficient secretion). As it was observed that the presence and absence of tags (UTR, secretion signal peptide) effected both expression and secretion it would be beneficial to investigate optimum combinations of these to increase secretion. This may involve using sections of these components, varying combinations or utilising other secretion signals i.e. for FlgM. The assay could also be used to investigate optimum expression conditions for increased yields of secreted proteins, with focus on the potential that expression and secretion capacity must be equally matched for optimum secretion. In terms of investigating the most effective cell culture methods for secretion a range of medias, induction and incubation protocols could be assessed –for example growing to high cell density and then inducing, growing at low temperature or growing in bioreactors with aeration (Rosano & Ceccarelli, 2014; Soini et al. 2008). It would be beneficial to obtain use of a parallel bioreactor system so that multiple experimental set ups could be implemented in tandem to investigate these effects, this would have the added benefit of being more akin in terms of reactor design to industrial fermentation methods.

It was evident in this thesis that not all proteins are secreted at the same capacity in all strains, therefore using a cutinase only secretion assay to screen for effective secretion, may not result in the best optimisation strategy for every protein. This could potentially be overcome, while still retaining the high throughput nature of the assay by modifying the secretion construct to harbour another enzyme such as sialidase which hydrolyzes terminal sialic acid residues on polysaccharide chains and can therefore be linked to a range of substrates –including

(Chokhawala et al., 2007) or reactive protein such as GFP or tetracysteine, which can be fluorescently labelled following secretion (DeLisa et al., 2002; Haitjema et al., 2014). Additionally different proteins could be secreted which are compatible with the optimised MUB assay, for example lipase (Roberts 1985). Lipase is also compatible with a large range of assays, including the tribuyrin and pNPB assays attempted here and also the precipitation of fatty acids from Tween or olive oil, or the interfacial tension experiments, where the change in shape of an enzyme-droplet is altered from a drop to a tear shape (Gupta et al., 2003; Eom et al., 2005).

It may also be possible to harness the FT3SS to implement strain based improvements to secretion via selective pressure (ideally following a protocol to generate lots of mutants i.e. a transposon library). If mildly toxic proteins were inserted into the secretion construct, cells with higher secretion capacities are more likely to survive. For example investigation into the expression of toxic proteins in *E. coli* such as DNAase and ribonuclease has found that cells can withstand varying amounts of toxins (Saïda et al., 2006; Doherty et al., 1993), therefore this different intracellular concentrations of toxin would have a variable effect on cell survival; if the toxin was linked to a secretion signal and expression is tuned correctly, this could result in a population of cells which survive or die based on their ability to secrete a toxin.

In reality it is likely that a combination of these strategies would be incorporated to result in a strain which is better designed to secrete protein through the FT3SS than the current mark 2 secretor strain, however all the tools are in place to realise this. An improved secretor strain is likely to maintain the Δ*CKL* Δ*clpX* Δ*motAB* background as it is so efficient, but also be 'stripped back' in terms of non-essential, non-detrimental genes such as the class III flagella genes and also harbour some gene knockouts throughout the chromosome, which were found to result in increased secretion capacity during non-directed mutagenesis e.g. a transposon library. This could be teamed with a high performing secretion construct, which would carry a favourable combination of UTR and secretion signal peptide residues to ensure good expression and secretion.

7.3. The modified FT3SS as a tool

The modified FT3SS was developed with the aim of secreting biotechnologically relevant protein; however in carrying out the work findings have also alluded to the novel insights on characteristics of the FT3SS –for example the role of the 5' and 3' UTR and 47 amino acid secretion signal on expression and secretion. With this in mind it would also be possible to utilise the FT3SS secretion platform developed here to further knowledge on the FT3SS. For example knockout mutagenesis of genes followed by measurement of protein secretion could be carried out to investigate the essentiality of components of the FT3SS or the effect of transcriptional regulators or environmental conditions. While the FT3SS is very well studied, as investigation is low throughput, the combinations of gene knockouts are generally logical, however this could be used to investigate a wide range of combinational FT3SS gene knockouts, increasing the potential to discover a new feature, function or interaction of proteins in the FT3SS. Furthermore as investigation into flagella phenotype is usually related to motility on motility agar plates, this may provide a more accurate representation of FT3SS activity of cells in liquid media. In addition to this, the necessity of regions of the 5' and 3' UTRs and secretion signals could also be investigated in a high throughput manner. A 5' UTR harbouring secretion signal without the 47 amino acid secretion signal was not generated in this study and there would be great value in observing the effect of this on secretion, as the role of these elements is under debate (Majander et al. 2005; Végh et al. 2006; Bonifield & Hughes, 2003). Furthermore if the 47 amino acid secretion signal is found to be non-essential a minimal secretion construct could be assembled. The cutinase secretion assay could also be used to investigate the amount of protein secreted without the tag accessory proteins which are harboured in the secretion construct as this may well improve secretion.

7.4. The role of the modified FT3SS in industrial biotechnology

Work carried out in this thesis has demonstrated the suitability of the modified FT3SS for use in industrial biotechnology. It was possible to secrete the nearly native E2 protein to the match the yields seen in other FT3SS examples, the best achieved in the literature was 12 mg L^{-1}

following high density culture, whereas I achieved around 23.5 mg L^{-1}, when cells were grown to an OD_{600} of 1.5. While yields did not match those seen for secreted protein in the expression systems used in industry (e.g. *Pichia* up to 15g L^{-1} rodent collagen fragments (Werten et al., 1999)) or for *E. coli* following secretion into the periplasm via the TAT system (60mg L^{-1} GFP in batch culture or 1g L^{-1} in fed batch high cell density culture (Matos et al., 2012)), there are many benefits to using an extracellular secretion *E. coli* based expression system which make this attractive for industrial use. In comparison to the other *E. coli* extracellular secretion system (T1SS) which has been investigated for industrial biotechnology use, the yields recorded in this work far surpassed them (reported at 2mg L^{-1} (Fernandez & de Lorenzo, 2001; Fernández 2004), demonstrating that the FT3SS is the most amenable secretion system to be utilised for technological purposes. Furthermore the design of the secretion construct means that the secretion platform could be adapted to fit industry specifications, both for the protein secreted, but also for the purification tag or protease cleavage site for example. In fact to achieve a wider view of the range of industrially relevant protein which could be expressed and secreted in this platform, it would be advantageous to approach companies to see if they would collaborate in trying to secrete proteins which they routinely produce intracellularly. Both yield and quality of secreted protein could be assessed. In the near future it would be interesting to try and secrete some more biotechnology and synthetic biology proteins, in terms of biotechnology, the expression of human growth hormone, insulin and ScFvs in a range of expression systems is widely reported, therefore it would be of interest to measure how the FT3SS compares. In terms of a synthetic biology challenge, the expression of spider silk is a popular theme in the field and causes a particular challenge because the desired strength and flexibility of the silk is dependent on correct translation and assembly and lack of degradation –the latter of which could be overcome if secreted to the media. Spider silk is formed of protein subunits which assemble to form a silk thread. Silk monomers have been expressed in *E. coli* from gene constructs which comprise of up to 96 repeats of monomers to enable expression of a silk pseudo-polymer (Xia et al., 2010; Michalczechen-Lacerda et al., 2014). Another strategy is to secrete spider silk as monomers either use them in this form or assemble them into polymers. As monomers have a tendency to self-assemble in cells, secretion would be beneficial and was previously achieved through the *Salmonella* TT3SS to yields of up to 70 no l^{-1} h^{-1} (Widmaier et al., 2009). It would be of interest to investigate what yield of secreted pseudo-polymers or monomers would be achievable in the *E. coli* FT3SS.

In general it would be desirable to gain more information on the quality of secreted protein in terms of fidelity to the anticipated residue sequence –this is relevant to both efficient transcription and translation and also in relation to proteolysis; this could be ascertained with mass spectrometry. Correct folding is required to ensure protein function and also low immunogenicity, therefore structural analysis. This was successfully demonstrated in the MUB cleavage assay which relies on functional enzyme, but could also include investigating whether therapeutic protein is functional. For example whether recombinant collagen can form links between mammalian cell lines with good extracellular membrane expression or its effect on wound healing models. Different culture methods could also result in novel products, for example the co-culture of strains expressing different proteins, to result in a multi protein secreted fraction. For example if cells were co-cultured which expressed either COL1A1 or COL1A2 this may enable the formation of trimeric full length type I collagen triple helixes. Another variation in culture method would be to grow secretion strains in biofilms, so that secreted protein may aggregate as a film or sheet of secreted protein based material, for example if collagen were used- enabling potential applications such as bacterial based production of collagen sheets (even by 3D printing) –previously *Staphylococcus* were 3D printed to produce a material, but it would be advantageous if these cells secreted therapeutic protein (Villar et al., 2013)

Finally in order for this secretion platform to be cost effective in an industrial setting there are two requirements; firstly that yield is increased, while this can be achieved in bench scale experiments by the strain and secretion construct improvements explored here and by different incubation methods, a clear route to improving cell biomass and therefore -in theory- expressed and secreted protein per volume of media, is to grow cells in rich media to a high optical density –this is best achieved by use of fed batch fermentation. Care would have to be taken that cell lysis did not occur, however to a degree as long as it occurred to a low degree, this secretion expression platform could still be cost effective as the positive aspects of extracellular secretion would still largely be at play, however media would require some downstream processing to remove protein contaminants. The alternative method of this secretion platform being economically viable would be to express and secrete a protein which is high value, thus requiring lower concentrations of product.

As discussed above yield can be improved by growing cells at high cell density –this is seen throughout the literature and skews the yields reported when compared to the results

achieved in this body of work. A worked example will be described to demonstrate this; for example secretion of GFP via TAT to the periplasm of *E. coli* in a small scale batch experiment resulted in a 60mg L^{-1} concentration of periplasmic derived protein. However cells were grown to OD_{600} 20. Therefore per Litre of culture, there are 13 or 20 times more *E. coli* present than there are when cells are cultures for E2 or CH2 secretion respectively. Therefore if it is assumed that cells are just as efficient at secreting at high optical densities (there are some caveats to this, including increased likelihood of plasmid loss, pH and oxygen reduction leading to reduced expression (Sivashanmugam et al., 2009)) then the yields will be more impressive than initially reported. For example when calculated per OD unit, the 60mg L^{-1} concentration, equates to 3µg GFP per OD unit (60/1000/20). This is comparable to 15µg E2 per OD unit (23.5 mg L^{-1}/1000/1.5) and 219ng CH2 per OD unit (219 µg L^{-1}/1000/1). Displaying results in this fashion highlights that the yields achieved in this work (in terms of E2) actually surpass those reported in this study. While CH2 yield was still lower that GFP, this data handling does reveal that only 13 (3000/219) times less efficient than that seen in TAT secretion, as opposed to the 250 fold difference (60000/219 estimated previously. In fed batch fermentation the same study reported a 1g L^{-1} yield of GFP extracted from the periplasm, however the OD_{600} exceeded 150, this would relate to 6.66mg GFP per OD unit (1000mg L^{-1}/1000/150). This is a much higher yield than those reported in this work and for GFP grown in small batch fermentation, demonstrating the benefits of supplementing cells with rich media. It is reported that for industry to consider a new platform technology for protein secretion yield should be in the 10s of mg L^{-1} (personal communication, Ian Hodgson, Head of Molecular Biology, FUJIFILM Diosynth Biotechnologies). This was not achieved for the native E2 protein, but not the recombinant CH2 protein in the experiments carried out in this thesis, however the aforementioned calculations show that a weight per volume request is subject to cell density, therefore with the mark 2 secretor strain and improved fermentation procedures, perhaps the FT3SS could be productive enough. Investigation into biotechnological suitability would also require checking that the tags used are compliant with industry –the main driver being cost. Strep and FLAG resins are expensive (although are used), therefore the process may be more cost effective if a non-chromatographic tag was used, such as elastic-like polypeptide, which precipitates when certain temperatures of buffers are applied. Equally, although proteases are utilised to remove tags in industry, the incorporation of a self-cleaving tag may be more cost effective than use of TEV protease (Fong et al., 2010).

7.5. Conclusion

The body of work in this thesis demonstrates the amenability of *E. coli* to both produce and secrete industrial biotechnology relevant protein products. Improvements were made to the secretion platform strain and secretion construct, which resulted in an increased yield of a range of protein products. As discussed, this represents a basic level of improvement in terms of the potential for future increases in secretion capacity. Through this thesis three iterations of the design process were carried out following generation of the initial secretion platform strain: (1) single strategy mutant generation, (2) combination of mutant strategies and (3) modification of the secretion construct. During these three iterations a 25 fold increase in the concentration of secreted protein was achieved. If applied to the highest yield achieved (23.5 mg L^{-1}), this could result in a yield of almost 600mg L^{-1}. As described in the discussion, in comparison to the secretion capacity of other *E. coli* secretion systems the choice of the FT3SS is beneficial. The yields of extracellular protein secretion surpassed that observed through T1SS secretion by far; furthermore depending on how yields are interpreted, the FT3SS is competitive with periplasmic T2SS secretion. With regard to the T2SS, secretion through the FT3SS has the added benefit of resulting in extracellular secretion, which requires less downstream processing and is likely to result in improved product quality. While yields achieved in other expression systems (*Pichia, Bacilus*) may exceed the FT3SS, use of an *E. coli* expression system is also desirable due to the low costs associated with its fermentation. Furthermore while the yields achieved are impressive, it could be further improved upon with the utilisation of a range of tools, relating to techniques from traditional microbiology, synthetic biology and computational biology. With the integration of these techniques I have confidence that the FT3SS secretion platform system could be competitive in the field of industrial biotechnology.

References

Abby, S.S. & Rocha, E.P.C., 2012. The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. M. Achtman, ed. *PLoS genetics*, 8(9), p.e1002983.

Aberg, A., Fernández-Vázquez, J., Cabrer-Panes, J.D., Sánchez, A. & Balsalobre, C., 2009. Similar and divergent effects of ppGpp and DksA deficiencies on transcription in Escherichia coli. *Journal of bacteriology*, 191(10), pp.3226–36.

Abrusci, P., Vergara-Irigaray, M., Johnson, S., Beeby, M.D., Hendrixson, D.R., Roversi, P., Friede, M.E., Deane, J.E., Jensen, G.J., Tang, C.M. & Lea, S.M., 2013. Architecture of the major component of the type III secretion system export apparatus. *Nature structural & molecular biology*, 20(1), pp.99–104.

Acharya, A., Xu, X.-J., Husain-Ponnampalam, R.D., Hoffmann-Benning, S. & Kuo, M.-H., 2005. Production of constitutively acetylated recombinant p53 from yeast and Escherichia coli by tethered catalysis. *Protein Expression and Purification*, 41(2), pp.417–425.

Adams, E. & Frank, L., 1980. Metabolism of proline and the hydroxyprolines. *Annual review of biochemistry*, 49, pp.1005–61.

Adler, J., 1975. Chemotaxis in bacteria. *Annual review of biochemistry*, 44, pp.341–56.

Agapakis, C.M., Ducat, D.C., Boyle, P.M., Wintermute, E.H., Way, J.C. & Silver, P.A., 2010. Insulation of a synthetic hydrogen metabolism circuit in bacteria. *Journal of biological engineering*, 4(1), p.3.

Ahmed Al-Tammar, K., Omar, O., Abdul Murad, A.M. & Abu Bakar, F.D., 2016. Expression and characterization of a cutinase (AnCUT2) from Aspergillus niger. *Open Life Sciences*, 11(1), pp.29–38.

Akeda, Y. & Galán, J.E., 2005. Chaperone release and unfolding of substrates in type III secretion. *Nature*, 437(7060), pp.911–915.

Alcock, F., Baker, M.A.B., Greene, N.P., Palmer, T., Wallace, M.I. & Berks, B.C., 2013. Live cell imaging shows reversible assembly of the TatA component of the twin-arginine protein transport system. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), p.E3650.

Aldridge, C., Poonchareon, K., Saini, S., Ewen, T., Soloyva, A., Rao, C. V, Imada, K., Minamino, T. & Aldridge, P.D., 2010. The interaction dynamics of a negative feedback loop regulates flagellar number in Salmonella enterica serovar Typhimurium. *Molecular microbiology*, 78(6), pp.1416–30.

Aldridge, P., Gnerer, J., Karlinsey, J.E. & Hughes, K.T., 2006. Transcriptional and translational control of the Salmonella fliC gene. *Journal of bacteriology*, 188(12), pp.4487–96.

Aldridge, P., Karlinsey, J. & Hughes, K.T., 2003. The type III secretion chaperone FlgN regulates flagellar assembly via a negative feedback loop containing its chaperone substrates FlgK and FlgL. *Molecular Microbiology*, 49(5), pp.1333–1345.

Aldridge, P.D., Karlinsey, J.E., Aldridge, C., Birchall, C., Thompson, D., Yagasaki, J. & Hughes, K.T., The flagellar-specific transcription factor, σ28 , is the Type III secretion chaperone for the flagellar-specific anti- σ 28 factor FlgM.

Amsler, C.D., Cho, M. & Matsumura, P., 1993. Multiple factors underlying the maximum motility of Escherichia coli as cultures enter post-exponential growth. *Journal of bacteriology*, 175(19), pp.6238–44.

An, B., Abbonante, V., Xu, H., Gavriilidou, D., Yoshizumi, A., Bihan, D., Farndale, R.W., Kaplan, D.L., Balduini, A., Leitinger, B. & Brodsky, B., 2016. Recombinant Collagen Engineered to Bind to Discoidin Domain Receptor Functions as a Receptor Inhibitor. *The Journal of biological chemistry*, 291(9), pp.4343–55.

An, B., Kaplan, D.L. & Brodsky, B., 2014. Engineered recombinant bacterial collagen as an alternative collagen-based biomaterial for tissue engineering. *Frontiers in chemistry*, 2, p.40.

Anderson, D.M., Ramamurthi, K.S., Tam, C. & Schneewind, O., 2002. YopD and LcrH regulate expression of Yersinia enterocolitica YopQ by a posttranscriptional mechanism and bind to yopQ RNA. *Journal of bacteriology*, 184(5), pp.1287–95.

Andrianantoandro, E., Basu, S., Karig, D.K. & Weiss, R., 2006. Synthetic biology: new engineering rules for an emerging discipline. *Molecular systems biology*, 2, p.2006.0028.

Anton, L., Majander, K., Savilahti, H., Laakkkonen, L. & Westerlund-Wikstrom, B., 2010. Two distinct regions in the model protein Peb1 are critical for its heterologous transport out of Escherichia coli. *Microbial Cell Factories*, 9(1), p.97.

Arkin, A., 2008. Setting the standard in synthetic biology. *Nature biotechnology*, 26(7), pp.771–4.

Auvray, F., Ozin, A.J., Claret, L. & Hughes, C., 2002. Intrinsic Membrane Targeting of the Flagellar Export ATPase FliI: Interaction with Acidic Phospholipids and FliH. *Journal of Molecular Biology*, 318(4), pp.941–950.

Auvray, F., Thomas, J., Fraser, G.M. & Hughes, C., 2001. Flagellin polymerisation control by a cytosolic export chaperone. *Journal of Molecular Biology*, 308(2), pp.221–229.

Baeshen, M.N., Al-Hejin, A.M., Bora, R.S., Ahmed, M.M.M., Ramadan, H.A.I., Saini, K.S., Baeshen, N.A. & Redwan, E.M., 2015. Production of Biopharmaceuticals in E. coli: Current Scenario and Future Perspectives. *Journal of microbiology and biotechnology*, 25(7), pp.953–62.

Baker, J.L., Çelik, E. & DeLisa, M.P., 2013. Expanding the glycoengineering toolbox: the rise of bacterial N-linked protein glycosylation. *Trends in Biotechnology*, 31(5), pp.313–323.

Baker, T.A. & Sauer, R.T., 2012. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1823(1), pp.15–28.

Baneyx, F., 1999. Recombinant protein expression in Escherichia coli. *Current Opinion in Biotechnology*, 10(5), pp.411–421.

Barcena Menendez, D., Senthivel, V.R. & Isalan, M., 2015. Sender–receiver systems and applying information theory for quantitative synthetic biology. *Current Opinion in Biotechnology*, 31, pp.101–107.

Barembruch, C. & Hengge, R., 2007. Cellular levels and activity of the flagellar sigma factor FliA of Escherichia coli are controlled by FlgM-modulated proteolysis. *Molecular Microbiology*, 65(1), pp.76–89.

Barker, C.S., Prüss, B.M. & Matsumura, P., 2004. Increased motility of Escherichia coli by insertion sequence element integration into the regulatory region of the flhD operon. *Journal of bacteriology*, 186(22), pp.7529–37.

Baeshen, N.A., Baeshen, M. N., Skeikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S., & Redwan, E. M., 2014. Cell factories for insulin production. *Microbial Cell Factories*, 13(1), p.141.

Bedau, M.A., Parke, E.C., Tangen, U. & Hantsche-Tangen, B., 2009. Social and ethical checkpoints for bottom-up synthetic biology, or protocells. *Systems and synthetic biology*, 3(1-4), pp.65–75.

Belasco, J.G., 2010. All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nature Reviews Molecular Cell Biology*, 11(7), pp.467–478.

Benedict, M.N., Gonnerman, M.C., Metcalf, W.W. & Price, N.D., 2012. Genome-scale metabolic reconstruction and hypothesis testing in the methanogenic archaeon Methanosarcina acetivorans C2A. *Journal of bacteriology*, 194(4), pp.855–65.

Beneyton, T. Wijaya, I. P. M., Postros, P., Najah, M., Leblond, P., Couvent, A., Mayot, E., Griffiths, A. D. & Drevella, A., 2016. High-throughput screening of filamentous fungi using nanoliter-range droplet-based microfluidics. *Scientific Reports*, 6, p.27223.

Bennett, J.C.. & Hughes, C., 2000. From flagellum assembly to virulence: the extended family of type III export chaperones. *Trends in Microbiology*, 8(5), pp.202–204.

Berg, H.C., 2003. The rotary motor of bacterial flagella. *Annual review of biochemistry*, 72, pp.19–54.

Berlec, A. & Štrukelj, B., 2013. Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells. *Journal of Industrial Microbiology & Biotechnology*, 40(3-4), pp.257–274.

Biek, D.P. & Cohen, S.N., 1986. Identification and characterization of recD, a gene affecting plasmid maintenance and recombination in Escherichia coli. *Journal of bacteriology*, 167(2), pp.594–603.

Blair, D.F., 1995. How bacteria sense and swim. *Annual review of microbiology*, 49, pp.489–522.

Blair, D.F. & Berg, H.C., 1990. The MotA protein of E. coli is a proton-conducting component of the flagellar motor. *Cell*, 60(3), pp.439–449.

Blight, M.A. & Holland, I.B., 1994. Heterologous protein secretion and the versatile Escherichia coli haemolysin translocator. *Trends in Biotechnology*, 12(11), pp.450–455.

Bonifield, H.R. & Hughes, K.T., 2003. Flagellar phase variation in Salmonella enterica is mediated by a posttranscriptional control mechanism. *Journal of bacteriology*, 185(12), pp.3567–74.

Breitling, R. & Takano, E., 2015. Synthetic biology advances for pharmaceutical production. *Current Opinion in Biotechnology*, 35, pp.46–51.

Brodsky, B. & Persikov, A. V, 2005. Molecular structure of the collagen triple helix. *Advances in protein chemistry*, 70(null), pp.301–39.

Brooks, S.A., 2004. Appropriate Glycosylation of Recombinant Proteins for Human Use: Implications of Choice of Expression System. *Molecular Biotechnology*, 28(3), pp.241–256.

Brown, J.D., Saini, S., Aldridge, C., Herbert, J., Rao, C. V & Aldridge, P.D., 2008. The rate of protein secretion dictates the temporal dynamics of flagellar gene expression. *Molecular microbiology*, 70(4), pp.924–37.

Buechter, D.D., Paolella, D.N., Leslie, B.S., Brown, M.S., Mehos, K.A. & Gruskin, E.A., 2003. Co-translational incorporation of trans-4-hydroxyproline into recombinant proteins in bacteria. *The Journal of biological chemistry*, 278(1), pp.645–50.

Bushnell, B.D., McWilliams, A.D., Whitener, G.B. & Messer, T.M., 2008. Early clinical experience with collagen nerve tubes in digital nerve repair. *The Journal of hand surgery*, 33(7), pp.1081–7.

Callaway, E., 2014. First synthetic yeast chromosome revealed. *Nature*.

Camberg, J.L., Hoskins, J.R. & Wickner, S., 2011. The Interplay of ClpXP with the Cell Division Machinery in Escherichia coli. *JOURNAL OF BACTERIOLOGY*, 193(8), pp.1911–1918.

Cameron, D.E., Urbach, J.M. & Mekalanos, J.J., 2008. A defined transposon mutant library and its use in identifying motility genes in Vibrio cholerae. *Proceedings of the National Academy of Sciences of the United States of America*, 105(25), pp.8736–41.

Canton, B., Labno, A. & Endy, D., 2008. Refinement and standardization of synthetic biological parts and devices. *Nature biotechnology*, 26(7), pp.787–93.

Cao, E., Chen, Y., Cui, Z. & Foster, P.R., 2003. Effect of freezing and thawing rates on denaturation of proteins in aqueous solutions. *Biotechnology and bioengineering*, 82(6), pp.684–90.

Carlson, R., 2016. Estimating the biotech sector's contribution to the US economy. *Nature biotechnology*, 34(3), pp.247–55.

Carvalho, C.M.L., Aires-Barros, M.R. & Cabral, J.M.S., 1998. Cutinase structure, function and biocatalytic applications. *EJB Electronic Journal of Biotechnology*, 1(3), pp.717–345.

Casadaban, M.J. & Cohen, S.N., 1980. Analysis of gene control signals by DNA fusion and cloning in Escherichia coli. *Journal of Molecular Biology*, 138(2), pp.179–207.

Cascales, E. & Christie, P.J., 2004. Agrobacterium VirB10, an ATP energy sensor required for type IV secretion. *Proceedings of the National Academy of Sciences of the United States of America*, 101(49), pp.17228–33.

Cereceres, S., Touchet, T., Browning, M.B., Smith, C., Rivera, J., Höök, M., Whitfield-Cargile, C., Russell, B. & Cosgriff-Hernandez, E., 2015. Chronic Wound Dressings Based on Collagen-Mimetic Proteins. *Advances in wound care*, 4(8), pp.444–456.

Chaban, B., Hughes, H.V. & Beeby, M., 2015. The flagellum in bacterial pathogens: For motility and a whole lot more. *Seminars in Cell & Developmental Biology*, 46, pp.91–103.

Chandran Darbari, V. & Waksman, G., 2015. Structural Biology of Bacterial Type IV Secretion Systems. *Annual Review of Biochemistry*, 84(1), pp.603–629.

Charriere, G., Bejot, M., Schnitzler, L., Ville, G. & Hartmann, D.J., 1989. Reactions to a bovine collagen implant. *Journal of the American Academy of Dermatology*, 21(6), pp.1203–1208.

Chatzi, K.E., Sardis, M.F., Karamanou, S. & Economou, A., 2013. Breaking on through to the other side: protein export through the bacterial Sec system. *Biochemical Journal*, 449(1).

Chen, M., Costa, F.K., Lindvay, C.R., Han, Y.-P. & Woodley, D.T., 2002. The recombinant expression of full-length type VII collagen and characterization of molecular mechanisms underlying dystrophic epidermolysis bullosa. *The Journal of biological chemistry*, 277(3), pp.2118–24.

Chen, M. T., Lin, S., Shandil, I., Andrews, D., Stadheim, T.A. & Choi, B.-K., 2012. Generation of diploid Pichia pastoris strains by mating and their application for recombinant protein production. *Microbial cell factories*, 11(1), p.91.

Chen, S., Su, L., Chen, J. & Wu, J., 2013. Cutinase: Characteristics, preparation, and application. *Biotechnology Advances*, 31(8), pp.1754–1767.

Chen, S., Tong, X., Woodard, R.W., Du, G., Wu, J. & Chen, J., 2008. Identification and characterization of bacterial cutinase. *The Journal of biological chemistry*, 283(38), pp.25854–62.

Chen, S. M., Tsai, Y. -S., Wu, C. -M., Liao, S. -K., Wu, L. -C., Chang, C. -S., Liu, Y. -H. & Tsai, P. -J., 2010. Streptococcal collagen-like surface protein 1 promotes adhesion to the respiratory epithelial cell. *BMC Microbiology*, 10(1), p.320.

Cheng, L.W. & Schneewind, O., 2000. Type III machines of Gram-negative bacteria: delivering the goods. *Trends in Microbiology*, 8(5), pp.214–220.

Chilcott, G.S. & Hughes, K.T., 2000. Coupling of Flagellar Gene Expression to Flagellar Assembly in Salmonella enterica Serovar Typhimurium and Escherichia coli. *Microbiology and Molecular Biology Reviews*, 64(4), pp.694–708.

Chintalacharuvu, K.R., Yu, L.J., Bhola, N., Kobayashi, K., Fernandez, C.Z. & Morrison, S.L., 2002. Cysteine residues required for the attachment of the light chain in human IgA2. *Journal of immunology (Baltimore, Md. : 1950)*, 169(9), pp.5072–7.

Choi, J.H., Keum, K.C. & Lee, S.Y., 2006. Production of recombinant proteins by high cell density culture of Escherichia coli. *Chemical Engineering Science*, 61(3), pp.876–885.

Choi, J.H. & Lee, S.Y., 2004. Secretory and extracellular production of recombinant proteins using Escherichia coli. *Applied microbiology and biotechnology*, 64(5), pp.625–35.

Chokhawala, H.A., Yu, H. & Chen, X., 2007. High-throughput substrate specificity studies of sialidases by using chemoenzymatically synthesized sialoside libraries. *Chembiochem : a European journal of chemical biology*, 8(2), pp.194–201.

Christie, P.J., Whitaker, N. & González-Rivera, C., 2014. Mechanism and structure of the bacterial type IV secretion systems. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), pp.1578–1591.

Chung, C.W., You, J., Kim, K., Moon, Y., Kim, H. & Ahn, J.H., 2009. Export of recombinant proteins in Escherichia coli using ABC transporter with an attached lipase ABC transporter recognition domain (LARD). *Microbial cell factories*, 8(1), p.11.

Church, G.M., Elowitz, M.B., Smolke, C.D., Voigt, C.A. & Weiss, R., 2014. Realizing the potential of synthetic biology. *Nature reviews. Molecular cell biology*, 15(4), pp.289–94.

Claret, L. & Hughes, C., 2002. Interaction of the Atypical Prokaryotic Transcription Activator FlhD2C2 with Early Promoters of the Flagellar Gene Hierarchy. *Journal of Molecular Biology*, 321(2), pp.185–199.

Clark, A.J., 1998. The mammary gland as a bioreactor: expression, processing, and production of recombinant proteins. *Journal of mammary gland biology and neoplasia*, 3(3), pp.337–50.

Clark, C.C., 1979. The distribution and initial characterization of oligosaccharide units on the COOH-terminal propeptide extensions of the pro-alpha 1 and pro-alpha 2 chains of type I procollagen. *J. Biol. Chem.*, 254(21), pp.10798–10802.

Clarke, M.B. & Sperandio, V., 2005. Transcriptional regulation of flhDC by QseBC and sigma (FliA) in enterohaemorrhagic Escherichia coli. *Molecular microbiology*, 57(6), pp.1734–49.

Cornelis, G.R. & Van Gijsegem, F., 2000. Assembly and function of type III secretory systems. *Annual review of microbiology*, 54, pp.735–74.

Cornelis, P., 2000. Expressing genes in different Escherichia coli compartments. *Current opinion in biotechnology*, 11(5), pp.450–4.

Costa, T.R.D., Felisberto-Rodrigues, C., Meir, A., Prevost, M.S., Redzej, A., Trokter, M. & Waksman, G., 2015. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature reviews. Microbiology*, 13(6), pp.343–59.

Crago, A.M. & Koronakis, V., 1998. Salmonella InvG forms a ring-like multimer that requires the InvH lipoprotein for outer membrane localization. *Molecular microbiology*, 30(1), pp.47–56.

Csonka, L.N., 1988. Regulation of cytoplasmic proline levels in Salmonella typhimurium: effect of osmotic stress on synthesis, degradation, and cellular retention of proline. *Journal of bacteriology*, 170(5), pp.2374–8.

Csonka, L.N., Gelvin, S.B., Goodner, B.W., Orser, C.S., Siemieniak, D. & Slightom, J.L., 1988. Nucleotide sequence of a mutation in the proB gene of Escherichia coli that confers proline overproduction and enhanced tolerance to osmotic stress. *Gene*, 64(2), pp.199–205.

Daniell, S.J., Kocsis, E., Morris, E., Knutton, S., Booy, F.P. & Frankel, G., 2003. 3D structure of EspA filaments from enteropathogenic Escherichia coli. *Molecular Microbiology*, 49(2), pp.301–308.

Datsenko, K.A. & Wanner, B.L., 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), pp.6640–5.

DeFrees, S., Wang, Z. G., Xing, R., Scott, A.E., Wang, J., Zopf, D., Gouty, D.L., Sjoberg, E.R., Panneerselvam, K., Brinkman-Van der Linden, E.C.M., Bayer, R.J., Tarp, M.A. & Clausen, H., 2006. GlycoPEGylation of recombinant therapeutic proteins produced in Escherichia coli. *Glycobiology*, 16(9), pp.833–43.

Degering, C., Eggert, T., Puls, M., Bongaerts, J., Evers, S., Maurer, K.-H. & Jaeger, K. E., 2010. Optimization of Protease Secretion in Bacillus subtilis and Bacillus licheniformis by

Screening of Homologous and Heterologous Signal Peptides. *Applied and Environmental Microbiology*, 76(19), pp.6370–6376.

Delepelaire, P., 2004. Type I secretion in gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1694(1), pp.149–161.

DeLisa, M.P., Samuelson, P., Palmer, T. & Georgiou, G., 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *The Journal of biological chemistry*, 277(33), pp.29825–31.

Denks, K., Vogt, A., Sachelaru, I., Petriman, N.-A., Kudva, R. & Koch, H. G., 2014. The Sec translocon mediated protein transport in prokaryotes and eukaryotes. , 31(2-3), pp.58–84.

Diepold, A. & Wagner, S., 2014. Assembly of the bacterial type III secretion machinery. *FEMS Microbiology Reviews*, 38(4), pp.802–822.

Dobó, J., Varga, J., Sajó, R., Végh, B.M., Gál, P., Závodszky, P. & Vonderviszt, F., 2010. Application of a short, disordered N-terminal flagellin segment, a fully functional flagellar type III export signal, to expression of secreted proteins. *Applied and environmental microbiology*, 76(3), pp.891–9.

Doherty, A.J., Connolly, B.A. & Worrall, A.F., 1993. *Overproduction of the toxic protein, bovine pancreatic DNasel, in Escherichia coli using a tightly controlled T7-promoter-based vector*,

Donath, M.J., Dominguez, M.A. & Withers, S.T., 2011. Development of an automated platform for high-throughput P1-phage transduction of Escherichia coli. *Journal of laboratory automation*, 16(2), pp.141–7.

Donato, G.M. & Kawula, T.H., 1998. Enhanced Binding of Altered H-NS Protein to Flagellar Rotor Protein FliG Causes Increased Flagellar Rotational Speed and Hypermotility in Escherichia coli. *Journal of Biological Chemistry*, 273(37), pp.24030–24036.

Dong, H., Nilsson, L. & Kurland, C., 1995. Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction. *J. Bacteriol.*, 177(6), pp.1497–1504.

Dragosits, M., Nicklas, D. & Tagkopoulos, I., 2012. A synthetic biology approach to self-regulatory recombinant protein production in Escherichia coli. *Journal of biological engineering*, 6(1), p.2.

Dubendorf, J.W. & Studier, F.W., 1991. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of Molecular Biology*, 219(1), pp.45–59.

Dunlap, V.J. & Csonka, L.N., 1985. Osmotic regulation of L-proline transport in Salmonella typhimurium. *Journal of bacteriology*, 163(1), pp.296–304.

Dvorak, P., Chrast, L., Nikel, P. I., Fedr, R., Soucek, K., Sedlackova, M., Chaloupkova, R., de Lorenzo, V., Prokop, Z. & Damborsky, J., 2015. Exacerbation of substrate toxicity by IPTG

in Escherichia coli BL21(DE3) carrying a synthetic metabolic pathway. *Microbial Cell Factories*, 14(1), p.201.

Dyrløv Bendtsen, J., Nielsen, H., von Heijne, G. & Brunak, S., 2004. Improved Prediction of Signal Peptides: SignalP 3.0. *Journal of Molecular Biology*, 340(4), pp.783–795..

Endy, D., 2005. Foundations for engineering biology. *Nature*, 438(7067), pp.449–53.

Eom, G.T., Song, J.K., Ahn, J.H., Seo, Y.S. & Rhee, J.S., 2005. Enhancement of the efficiency of secretion of heterologous lipase in Escherichia coli by directed evolution of the ABC transporter system. *Applied and environmental microbiology*, 71(7), pp.3468–74.

Erhardt, M., Hirano, T., Su, Y., Paul, K., Wee, D.H., Mizuno, S., Aizawa, S.-I. & Hughes, K.T., 2010. The role of the FliK molecular ruler in hook-length control in *Salmonella enterica*. *Molecular Microbiology*, 75(5), pp.1272–1284.

Erhardt, M. & Hughes, K.T., 2010. C-ring requirement in flagellar type III secretion is bypassed by FlhDC upregulation. *Molecular microbiology*, 75(2), pp.376–93.

Erhardt, M., Namba, K. & Hughes, K.T., 2010. Bacterial nanomachines: the flagellum and type III injectisome. *Cold Spring Harbor perspectives in biology*, 2(11), p.a000299.

Erhardt, M., Singer, H. M., Wee, D. H., Keener, J. P. & Hughes, K. T., 2011. An infrequent molecular ruler controls flagellar hook length in Salmonella enterica. *The EMBO journal*, 30(14), pp.2948–61.

Erhardt, M., Mertens, M.E., Fabiani, F.D. & Hughes, K.T., 2014. ATPase-independent type-III protein secretion in Salmonella enterica. *PLoS genetics*, 10(11), p.e1004800.

Eser, M., Masip, L., Kadokura, H., Georgiou, G. & Beckwith, J., 2009. Disulfide bond formation by exported glutaredoxin indicates glutathione's presence in the E. coli periplasm. *Proceedings of the National Academy of Sciences of the United States of America*, 106(5), pp.1572–7.

Eskelin, K., Ritala, A., Suntio, T., Blumer, S., Holkeri, H., Wahlström, E.H., Baez, J., Mäkinen, K. & Maria, N.A., 2009. Production of a recombinant full-length collagen type I alpha-1 and of a 45-kDa collagen type I alpha-1 fragment in barley seeds. *Plant biotechnology journal*, 7(7), pp.657–72.

Evans, L.D.B., Hughes, C. & Fraser, G. M., 2014. Building a flagellum outside the bacterial cell. *Trends in microbiology*, 22(10), pp.566–72.

Evans, L.D.B., Poulter, S., Terentjev, E.M., Hughes, C. & Fraser, G.M., 2013. A chain mechanism for flagellum growth. *Nature*, 504(7479), pp.287–90.

Evans, L.D.B., Stafford, G.P., Ahmed, S., Fraser, G.M. & Hughes, C., 2006. An escort mechanism for cycling of export chaperones during flagellum assembly. *Proceedings of the National Academy of Sciences of the United States of America*, 103(46), pp.17474–9.

Fahrner, K.A. & Berg, H.C., 2015. Mutations That Stimulate flhDC Expression in Escherichia coli K-12. *Journal of bacteriology*, 197(19), pp.3087–96.

Falcioni, F., Blank, L.M., Frick, O., Karau, A., Bühler, B. & Schmid, A., 2013. Proline availability regulates proline-4-hydroxylase synthesis and substrate uptake in proline-hydroxylating recombinant Escherichia coli. *Applied and environmental microbiology*, 79(9), pp.3091–100.

Ferdous, Z. & Grande-Allen, K.J., 2007. Utility and control of proteoglycans in tissue engineering. *Tissue engineering*, 13(8), pp.1893–904.

Fernández, L.A., 2004. Prokaryotic expression of antibodies and affibodies. *Current Opinion in Biotechnology*, 15(4), pp.364–373.

Fernandez, L.A. & de Lorenzo, V., 2001. Formation of disulphide bonds during secretion of proteins through the periplasmic-independent type I pathway. *Molecular Microbiology*, 40(2), pp.332–346.

Ferrer-Miralles, N. & Villaverde, A., 2013. Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microbial cell factories*, 12, p.113.

Fichard, A., Tillet, E., Delacoux, F., Garrone, R. & Ruggiero, F., 1997. Human recombinant alpha1(V) collagen chain. Homotrimeric assembly and subsequent processing. *The Journal of biological chemistry*, 272(48), pp.30083–7.

Fidan, O. & Xhan, J., 2015. Recent advances in engineering yeast for pharmaceutical protein production. *RSC Adv.*, 5(105), pp.86665–86674.

Fink, D.W. & Koehler, W.R., 1970. pH Effects on fluorescence of umbelliferone. *Analytical Chemistry*, 42(9), pp.990–993.

Fior, S., Vianelli, A. & Gerola, P.D., 2009. A novel method for fluorometric continuous measurement of β-glucuroanidase (GUS) activity using 4-methyl-umbelliferyl-β-d-glucuronide (MUG) as substrate. *Plant Science*, 176(1), pp.130–135.

Fitzgerald, D.M., Bonocora, R. P. & Wade, J. T., 2014. Comprehensive Mapping of the *Escherichia coli* Flagellar Regulatory Network L. Søgaard-Andersen, ed. *PLoS Genetics*, 10(10), p.e1004649.

Flynn, J.M., Neher, S.B., Kim, Y. I., Sauer, R.T. & Baker, T.A., 2003. Proteomic Discovery of Cellular Substrates of the ClpXP Protease Reveals Five Classes of ClpX-Recognition Signals. *Molecular Cell*, 11(3), pp.671–683.

Fojan, P., Jonson, P.H., Petersen, M.T.. & Petersen, S.B., 2000. What distinguishes an esterase from a lipase: A novel structural approach. *Biochimie*, 82(11), pp.1033–1041.

Fong, B.A., Wu, W.Y. & Wood, D.W., 2010. The potential role of self-cleaving purification tags in commercial-scale processes. *Trends in Biotechnology*, 28(5), pp.272–279.

Francez-Charlot, A., Laugel, B., Van Gemert, A., Dubarry, N., Wiorowski, F., Castanié-Cornet, M.-P., Gutierrez, C. & Cam, K., 2004. RcsCDB His-Asp phosphorelay system negatively regulates the flhDC operon in Escherichia coli. *Molecular Microbiology*, 49(3), pp.823–832.

Fraser, G.M., Bennett, J.C.Q. & Hughes, C., 1999. Substrate-specific binding of hook-associated proteins by FlgN and FliT, putative chaperones for flagellum assembly. *Molecular Microbiology*, 32(3), pp.569–580.

Fraser, G.M., Hirano, T., Ferris, H.U., Devgan, L.L., Kihara, M. & Macnab, R.M., 2003. Substrate specificity of type III flagellar protein export in Salmonella is controlled by subdomain interactions in FlhB. *Molecular Microbiology*, 48(4), pp.1043–1057.

Frenzel, A., Hust, M. & Schirrmann, T., 2013. Expression of recombinant antibodies. *Frontiers in immunology*, 4, p.217.

Frey, S. & Görlich, D., 2014. A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. *Journal of Chromatography A*, 1337, pp.95–105.

Friess, W., 1998. Collagen – biomaterial for drug delivery1Dedicated to Professor Dr. Eberhard Nürnberg, Friedrich-Alexander-Universität Erlangen-Nürnberg, on the occasion of his 70th birthday.1. *European Journal of Pharmaceutics and Biopharmaceutics*, 45(2), pp.113–136.

Fronzes, R., Christie, P.J. & Waksman, G., 2009. The structural biology of type IV secretion systems. *Nature reviews. Microbiology*, 7(10), pp.703–14.

Fukuda, K., Hori, H., Utani, A., Burbelo, P.D. & Yamada, Y., 1997. Formation of recombinant triple-helical [alpha 1(IV)]2 alpha 2(IV) collagen molecules in CHO cells. *Biochemical and biophysical research communications*, 231(1), pp.178–82.

Furukawa, Y., Inoue, Y., Sakaguchi, A., Mori, Y., Fukumura, T., Miyata, T., Namba, K. & Minamino, T., 2016. Structural stability of flagellin subunit affects the rate of flagellin export in the absence of FliS chaperone. *Molecular Microbiology*.

Galán, J.E. & Wolf-Watz, H., 2006. Protein delivery into eukaryotic cells by type III secretion machines. *Nature*, 444(7119), pp.567–73.

Galeva, A., Moroz, N., Yoon, Y.H., Hughes, K.T., Samatey, F.A. & Kostyukova, A.S., 2014. Bacterial flagellin-specific chaperone FliS interacts with anti-sigma factor FlgM. *Journal of bacteriology*, 196(6), pp.1215–21.

Galperin M, Dibrov, P.A. & Glagolev, A.N., 1982. delta mu H+ is required for flagellar growth in Escherichia coli. *FEBS letters*, 143(2), pp.319–22.

Gatherer, D., 2010. So what do we really mean when we say that systems biology is holistic? *BMC systems biology*, 4, p.22.

Gauger, E.J., Leatham, M.P., Mercado-Lubo, R., Laux, D.C., Conway, T. & Cohen, P.S., 2007. Role of motility and the flhDC Operon in Escherichia coli MG1655 colonization of the mouse intestine. *Infection and immunity*, 75(7), pp.3315–24.

Gehlsen, K.R., Gong, R., Bramhill, D., Wiersma, D.A., Kirkpatrick, S.A., Wang, Y., Feng, Y. & Dimitrov, D.S., 2012. Pharmacokinetics of engineered human monomeric and dimeric CH2 domains. *mAbs*, 4(4), pp.466–474.

Gelse, K., Pöschl, E. & Aigner, T., 2003. Collagens--structure, function, and biosynthesis. *Advanced drug delivery reviews*, 55(12), pp.1531–46.

Georgiou, G. & Segatori, L., 2005. Preparative expression of secreted proteins in bacteria: status report and future prospects. *Current opinion in biotechnology*, 16(5), pp.538–45.

Gerber, B.R., Asakura, S. & Oosawa, F., 1973. Effect of temperature on the in vitro assembly of bacterial flagella. *Journal of Molecular Biology*, 74(4), pp.467–487.

Gilda, J.E., Ghosh, R., Cheah, J.X., West, T.M., Bodine, S.C. & Gomes, A. V, 2015. Western Blotting Inaccuracies with Unverified Antibodies: Need for a Western Blotting Minimal Reporting Standard (WBMRS). *PloS one*, 10(8), p.e0135392.

Gimpel, J.A., Hyun, J.S., Schoepp, N.G. & Mayfield, S.P., 2015. Production of recombinant proteins in microalgae at pilot greenhouse scale. *Biotechnology and Bioengineering*, 112(2), pp.339–345.

Girgis, H.S., Liu, Y., Ryu, W.S. & Tavazoie, S., 2007. A comprehensive genetic characterization of bacterial motility. *PLoS genetics*, 3(9), pp.1644–60.

Glick, B.R., 1995. Metabolic load and heterologous gene expression. *Biotechnology Advances*, 13(2), pp.247–261.

Glowacki, J. & Mizuno, S., 2008. Collagen scaffolds for tissue engineering. *Biopolymers*, 89(5), pp.338–44.

Goldberg, I., Salerno, A.J., Patterson, T. & Williams, J.I., 1989. Cloning and expression of a collagen-analog-encoding synthetic gene in Escherichia coli. *Gene*, 80(2), pp.305–314.

Gophna, U., Ron, E.Z. & Graur, D., 2003. Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. *Gene*, 312, pp.151–63.

Gottesman, S., 1996. Proteases and their targets in *Escherichia coli* . *Annual Review of Genetics*, 30(1), pp.465–506.

Grebe, T.W. & Stock, J., 1998. Bacterial chemotaxis: The five sensors of a bacterium. *Current Biology*, 8(5), pp.R154–R157.

Gibson, D.G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. L., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkock, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-

Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A, Smith, H. O. & Venter, J.C., 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. *Science (New York, N.Y.)*, 329(5987), pp.52–6.

Greer-Phillips, S.E., Alexandre, G., Taylor, B.L. & Zhulin, I.B., 2003. Aer and Tsr guide Escherichia coli in spatial gradients of oxidizable substrates. *Microbiology*, 149(9), pp.2661–2667.

Griswold, K.E., Mahmood, N.A., Iverson, B.L. & Georgiou, G., 2003. Effects of codon usage versus putative 5'-mRNA structure on the expression of Fusarium solani cutinase in the Escherichia coli cytoplasm. *Protein Expression and Purification*, 27(1), pp.134–142.

Guo, S., Alshamy, I., Hughes, K.T. & Chevance, F.F. V, 2014. Analysis of factors that affect FlgM-dependent type III secretion for protein purification with Salmonella enterica serovar Typhimurium. *Journal of bacteriology*, 196(13), pp.2333–47.

Gupta, P., Ghosalkar, A., Mishra, S. & Chaudhuri, T.K., 2009. Enhancement of over expression and chaperone assisted yield of folded recombinant aconitase in Escherichia coli in bioreactor cultures. *Journal of Bioscience and Bioengineering*, 107(2), pp.102–107.

Gupta, R., Rathi, P., Gupta, N. & Bradoo, S., 2003. Lipase assays for conventional and molecular screening: an overview. *Biotechnology and Applied Biochemistry*, 37(1), p.63.

Gurramkonda, C., Polez, S., Skoko, N., Adnan, A., Gäbel, T., Chugh, D., Swaminathan, S., Khanna, N., Tisminetzky, S. & Rinas, U., 2010. Application of simple fed-batch technique to high-level secretory production of insulin precursor using Pichia pastoris with subsequent purification and conversion to human insulin. *Microbial cell factories*, 9, p.31.

Gustafsson, C., Govindarajan, S. & Minshull, J., 2004. Codon bias and heterologous protein expression. *Trends in Biotechnology*, 22(7), pp.346–353.

Hahn, H.P. & Specht, B.-U., 2003. Secretory delivery of recombinant proteins in attenuated *Salmonella* strains: potential and limitations of Type I protein transporters. *FEMS Immunology & Medical Microbiology*, 37(2-3), pp.87–98.

Haitjema, C.H., Boock, J.T., Natarajan, A., Dominguez, M.A., Gardner, J.G., Keating, D.H., Withers, S.T. & DeLisa, M.P., 2014. Universal genetic assay for engineering extracellular protein expression. *ACS synthetic biology*, 3(2), pp.74–82.

Hall, B.G., Acar, H., Nandipati, A. & Barlow, M., 2014. Growth rates made easy. *Molecular biology and evolution*, 31(1), pp.232–8.

Hannig, G. & Makrides, S.C., 1998. Strategies for optimizing heterologous protein expression in Escherichia coli. *Trends in Biotechnology*, 16(2), pp.54–60.

Heinemann, M. & Panke, S., 2006. Synthetic biology--putting engineering into biology. *Bioinformatics (Oxford, England)*, 22(22), pp.2790–9.

Helaakoski, T., 1996. Characterization of Human Type III Collagen Expressed in a Baculovirus System. *Journal of Biological Chemistry*, 271(20), pp.11988–11995.

Hellwig, S., Drossard, J., Twyman, R.M. & Fischer, R., 2004. Plant cell cultures for the production of recombinant proteins. *Nature biotechnology*, 22(11), pp.1415–22.

Henderson, G.E., Isett, K.D. & Gerngross, T.U., 2011. Site-specific modification of recombinant proteins: a novel platform for modifying glycoproteins expressed in E. coli. *Bioconjugate chemistry*, 22(5), pp.903–12.

Henderson, I.R., Navarro-Garcia, F., Desvaux, M., Fernandez, R.C. & Ala'Aldeen, D., 2004. Type V protein secretion pathway: the autotransporter story. *Microbiology and molecular biology reviews : MMBR*, 68(4), pp.692–744.

Heng, C., Chen, Z., Du, L. & Lu, F., 2005. Expression and secretion of an acid-stable alpha-amylase gene in Bacillus Subtilis by SacB promoter and signal peptide. *Biotechnology letters*, 27(21), pp.1731–7.

Herlihey, F.A., Moynihan, P.J. & Clarke, A.J., 2014. The essential protein for bacterial flagella formation FlgJ functions as a β-N-acetylglucosaminidase. *The Journal of biological chemistry*, 289(45), pp.31029–42.

Hirano, T., Yamaguchi, S., Oosawa, K. & Aizawa, S., 1994. Roles of FliK and FlhB in determination of flagellar hook length in Salmonella typhimurium. *Journal of bacteriology*, 176(17), pp.5439–49.

Ho, B.T., Dong, T.G. & Mekalanos, J.J., 2014. A view to a kill: the bacterial type VI secretion system. *Cell host & microbe*, 15(1), pp.9–21.

Homma, M., DeRosier, D.J. & Macnab, R.M., 1990. Flagellar hook and hook-associated proteins of Salmonella typhimurium and their relationship to other axial components of the flagellum. *Journal of molecular biology*, 213(4), pp.819–32.

Homma, M., Fujita, H., Yamaguchi, S. & Iono, T., 1984. Excretion of unassembled flagellin by Salmonella typhimurium mutants deficient in hook-associated proteins. *Journal of Bacteriology*, 159(3), p.1056.

Hosking, E.R., Vogt, C., Bakker, E.P. & Manson, M.D., 2006. The Escherichia coli MotAB proton channel unplugged. *Journal of molecular biology*, 364(5), pp.921–37.

Hou, Y.,Guey, L. T., Wu, T., Gau, R., Cogan, J., Wang, X., Hong, E., Ning, W. V., Keene, D., Liu, N., Huang, y., Kaftan, C., Tangarone, B., Quinones-Garcia, I., Uitto, J., Francone, O. L, Woodley, D. T. & Chen, M., 2015. Intravenously Administered Recombinant Human Type VII Collagen Derived from Chinese Hamster Ovary Cells Reverses the Disease Phenotype in Recessive Dystrophic Epidermolysis Bullosa Mice. *Journal of Investigative Dermatology*, 135(12), pp.3060–3067.

Howes, J. M., Bihan, D., Slatter, D.A., Hamaia, S.W., Packman, L.C., Knauper, V., Visse, R. & Farndale, R.W., 2014. The recognition of collagen and triple-helical toolkit peptides by MMP-13: sequence specificity for binding and cleavage. *The Journal of biological chemistry*, 289(35), pp.24091–101.

Hu, Y., Wang, Y., Ding, L., Lu, P., Atkinson, S. & Chen, S., 2009. Positive regulation of flhDC expression by OmpR in Yersinia pseudotuberculosis. *Microbiology*, 155(11), pp.3622–3631.

Huang, H.C., Sherman, M.Y., Kandror, O. & Goldberg, A.L., 2001. The molecular chaperone DnaJ is required for the degradation of a soluble abnormal protein in Escherichia coli. *The Journal of biological chemistry*, 276(6), pp.3920–8.

Hughes, K.T., Gillen, K.L., Semon, M.J. & Karlinsey, J.E., 1993. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. *Science (New York, N.Y.)*, 262(5137), pp.1277–80.

Hunt, I., 2005. From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein expression and purification*, 40(1), pp.1–22.

Hutchison, C.A., Chuang, R. -Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z. -Q., Richter, R. A., Strychalski, E. A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G. & Venter, J.C., 2016. Design and synthesis of a minimal bacterial genome. *Science*, 351(6280).

Ikeda, T., Homma, M., Iino, T., Asakura, S. & Kamiya, R., 1987. Localization and stoichiometry of hook-associated proteins within Salmonella typhimurium flagella. *Journal of bacteriology*, 169(3), pp.1168–73.

Ikemura, T., 1981. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the E. coli translational system. *Journal of Molecular Biology*, 151(3), pp.389–409.

Ikeno, S. & Haruyama, T., 2013. Boost Protein Expression through Co-Expression of LEA-Like Peptide in Escherichia coli A. Mitraki, ed. *PLoS ONE*, 8(12), p.e82824.

Jalalirad, R., 2013. Production of antibody fragment (Fab) throughout Escherichia coli fed-batch fermentation process: Changes in titre, location and form of product. *Electronic Journal of Biotechnology*, 16(3).

Jarvis, K.G., Girón, J.A., Jerse, A.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. *Proceedings of the National Academy of Sciences of the United States of America*, 92(17), pp.7996–8000.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. & Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N.Y.)*, 337(6096), pp.816–21.

Jonasson, P., Liljeqvist, S., Nygren, P.-A. & Ståhl, S., 2002. Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli. *Biotechnology and applied biochemistry*, 35(Pt 2), pp.91–105.

Jones, K.L., Kim, S.-W. & Keasling, J.., 2000. Low-Copy Plasmids can Perform as Well as or Better Than High-Copy Plasmids for Metabolic Engineering of Bacteria. *Metabolic Engineering*, 2(4), pp.328–338.

Jong, W.S.P., Soprova, Z., de Punder, K., ten Hagen-Jongman, C.M., Wagner, S., Wickström, D., de Gier, J.-W., Andersen, P., van der Wel, N.N. & Luirink, J., 2012. A structurally informed autotransporter platform for efficient heterologous protein secretion and display. *Microbial cell factories*, 11(1), p.85.

Juhas, M., Eberl, L. & Glass, J.I., 2011. Essence of life: essential genes of minimal genomes. *Trends in Cell Biology*, 21(10), pp.562–568.

Jullesson, D., David, F., Pfleger, B. & Nielsen, J., 2015. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnology Advances*, 33(7), pp.1395–1402.

Jürgen, B., Breitenstein, A., Urlacher, V., Buttner, K., Lin, H., Hecker, M., Schweder, T. & Neubauer, P., 2010. Quality control of inclusion bodies in Escherichia coli. *Microbial Cell Factories*, 9(1), p.41.

Kadler, K.E., Baldock, C., Bella, J. & Boot-Handford, R.P., 2007. Collagens at a glance. *Journal of cell science*, 120(Pt 12), pp.1955–8.

Kadler, K.E., Holmes, D.F., Trotter, J.A. & Chapman, J.A., 1996. Collagen fibril formation. *The Biochemical journal*, 316 (Pt 1, pp.1–11.

Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M.G. & Alon, U., 2001. Ordering Genes in a Flagella Pathway by Analysis of Expression Kinetics from Living Bacteria. *Science*, 292(5524).

Kanonenberg, K., Schwarz, C.K.W. & Schmitt, L., 2013. Type I secretion systems – a story of appendices. *Research in Microbiology*, 164(6), pp.596–604.

Karlinsey, J.E., Lonner, J., Brown, K.L. & Hughes, K.T., 2000. Translation/Secretion Coupling by Type III Secretion Systems. *Cell*, 102(4), pp.487–497.

Karlinsey, J.E., Tanaka, S., Bettenworth, V., Yamaguchi, S., Boos, W., Aizawa, S.-I. & Hughes, K.T., 2000. Completion of the hook-basal body complex of the Salmonella typhimurium flagellum is coupled to FlgM secretion and fliC transcription. *Molecular Microbiology*, 37(5), pp.1220–1231.

Kassner, A., Tiedemann, K., Notbohm, H., Ludwig, T., Mörgelin, M., Reinhardt, D.P., Chu, M.-L., Bruckner, P. & Grässel, S., 2004. Molecular structure and interaction of recombinant human type XVI collagen. *Journal of molecular biology*, 339(4), pp.835–53.

Kato, S., Okamoto, M. & Asakura, S., 1984. Polymorphic transition of the flagellar polyhook from Escherichia coli and Salmonella typhimurium. *Journal of Molecular Biology*, 173(4), pp.463–476.

Kawamoto, A., Morimoto, Y. V, Miyata, T., Minamino, T., Hughes, K.T., Kato, T. & Namba, K., 2013. Common and distinct structural features of Salmonella injectisome and flagellar basal body. *Scientific reports*, 3, p.3369.

Kelly, J.R., Rubin, A.J., Davis, J.H., Ajo-Franklin, C.M., Cumbers, J., Czar, M.J., de Mora, K., Glieberman, A.L., Monie, D.D. & Endy, D., 2009. Measuring the activity of BioBrick promoters using an in vivo reference standard. *Journal of biological engineering*, 3(1), p.4.

Khokhlatchev, A., Xu, S., English, J., Wu, P., Schaefer, E. & Cobb, M.H., 1997. Reconstitution of Mitogen-activated Protein Kinase Phosphorylation Cascades in Bacteria: efficient synthesis of active protein kinases. *J. Biol. Chem.*, 272(17), pp.11057–11062.

Kim, J., Webb, A. M., Kershner, J. P., Blaskowski, S. & Copley, S. D., 2014. A versatile and highly efficient method for scarless genome editing in Escherichia coli and Salmonella enterica. *BMC Biotechnology*, 14(1), p.84.

Kim, J.Y., Kim, Y.-G. & Lee, G.M., 2012. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Applied Microbiology and Biotechnology*, 93(3), pp.917–930.

Kinoshita, M., Hara, N., Imada, K., Namba, K. & Minamino, T., 2013. Interactions of bacterial flagellar chaperone-substrate complexes with FlhA contribute to co-ordinating assembly of the flagellar filament. *Molecular microbiology*, 90(6), pp.1249–61.

Kitagawa, R., Takaya, A. & Yamamoto, T., 2011. Dual regulatory pathways of flagellar gene expression by ClpXP protease in enterohaemorrhagic Escherichia coli. *Microbiology (Reading, England)*, 157(Pt 11), pp.3094–103.

Ko, M. & Park, C., 2000. H-NS-Dependent Regulation of Flagellar Synthesis Is Mediated by a LysR Family Protein. *Journal of Bacteriology*, 182(16), pp.4670–4672.

Kojima, C., Suehiro, T., Watanabe, K., Ogawa, M., Fukuhara, A., Nishisaka, E., Harada, A., Kono, K., Inui, T. & Magata, Y., 2013. Doxorubicin-conjugated dendrimer/collagen hybrid gels for metastasis-associated drug delivery systems. *Acta biomaterialia*, 9(3), pp.5673–80.

Koronakis, V., Eswaran, J. & Hughes, C., 2004. Structure and function of TolC: the bacterial exit duct for proteins and drugs. *Annual review of biochemistry*, 73, pp.467–89.

Korotkov, K. V., Sandkvist, M. & Hol, W.G.J., 2012. The type II secretion system: biogenesis, molecular architecture and mechanism. *Nature Reviews Microbiology*, 10(5), p.336.

Kosarewicz, A., Königsmaier, L. & Marlovits, T.C., 2012. The blueprint of the type-3 injectisome. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1592).

Koster, M., Bitter, W., de Cock, H., Allaoui, A., Cornelis, G.R. & Tommassen, J., 1997. The outer membrane component, YscC, of the Yop secretion machinery of *Yersinia enterocolitica* forms a ring-shaped multimeric complex. *Molecular Microbiology*, 26(04), pp.789–797.

Kranen, E., Detzel, C., Weber, T. & Jose, J., 2014. Autodisplay for the co-expression of lipase and foldase on the surface of E. coli: washing with designer bugs. *Microbial cell factories*, 13, p.19.

Kumar, R. & Shimizu, K., 2010. Metabolic regulation of Escherichia coli and its gdhA, glnL, gltB, D mutants under different carbon and nitrogen limitations in the continuous culture. *Microbial Cell Factories*, 9(1), p.8.

Kutsukake, K., Suzuki, T., Yamaguchi, S. & Iino, T., 1979. Role of gene flaFV on flagellar hook formation in Salmonella typhimurium. *Journal of bacteriology*, 140(1), pp.267–75.

Kutsukake, K., 1994. Excretion of the anti-sigma factor through a flagellar substructure couples flagellar gene expression with flagellar assembly in Salmonella typhimurium. *Molecular and General Genetics MGG*, 243(6), pp.605–612.

Kuwajima, G., Kawagishi, I., Homma, M., Asaka, J., Kondo, E. & Macnab, R.M., 1989. Export of an N-terminal fragment of Escherichia coli flagellin by a flagellum-specific pathway. *Proceedings of the National Academy of Sciences of the United States of America*, 86(13), pp.4953–7.

Kwok, R., 2010. Five hard truths for synthetic biology. *Nature*, 463(7279), pp.288–90.

Kwon, M.-A., Kim, H.S., Yang, T.H., Song, B.K. & Song, J.K., 2009. High-level expression and characterization of Fusarium solani cutinase in Pichia pastoris. *Protein Expression and Purification*, 68, pp.104–109.

Lara-Tejero, M., Kato, J., Wagner, S., Liu, X. & Galán, J.E., 2011. A sorting platform determines the order of protein secretion in bacterial type III systems. *Science (New York, N.Y.)*, 331(6021), pp.1188–91.

Lasch, P., Schmitt, J., Beekes, M., Udelhoven, T., Eiden, M., Fabian, H., Petrich, W. & Naumann, D., 2003. Antemortem identification of bovine spongiform encephalopathy from serum using infrared spectroscopy. *Analytical chemistry*, 75(23), pp.6673–8.

Lee, C.C., Wong, D.W. & Robertson, G.H., 2001. An E. coli expression system for the extracellular secretion of barley alpha-amylase. *Journal of protein chemistry*, 20(3), pp.233–7.

Lee, C.H., Singla, A. & Lee, Y., 2001. Biomedical applications of collagen. *International Journal of Pharmaceutics*, 221(1-2), pp.1–22.

Lee, D.J., Bingle, L.E.H., Heurlier, K., Pallen, M.J., Penn, C.W., Busby, S.J.W. & Hobman, J.L., 2009. Gene doctoring: a method for recombineering in laboratory and pathogenic Escherichia coli strains. *BMC microbiology*, 9(1), p.252.

Lee, S.H. & Galán, J.E., 2004. Salmonella type III secretion-associated chaperones confer secretion-pathway specificity. *Molecular Microbiology*, 51(2), pp.483–495.

Le Fourn, V., Girod, P. A., Buceta, M., Regamey, A. & Mermod, N., 2014. CHO cell engineering to prevent polypeptide aggregation and improve therapeutic protein secretion. *Metabolic Engineering*, 21, pp.91–102.

Lehnen, D., Blumer, C., Polen, T., Wackwitz, B., Wendisch, V.F. & Unden, G., 2002. LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in Escherichia coli. *Molecular Microbiology*, 45(2), pp.521–532.

Lehti, T.A., Bauchart, P., Dobrindt, U., Korhonen, T.K. & Westerlund-Wikström, B., 2012. The fimbriae activator MatA switches off motility in Escherichia coli by repression of the flagellar master operon flhDC. *Microbiology (Reading, England)*, 158(Pt 6), pp.1444–55.

Lemke, J.J., Durfee, T. & Gourse, R.L., 2009. DksA and ppGpp directly regulate transcription of the Escherichia coli flagellar cascade. *Molecular microbiology*, 74(6), pp.1368–79.

Lenski, R.E. & Travisano, M., 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proceedings of the National Academy of Sciences of the United States of America*, 91(15), pp.6808–14.

Li, C., Louise, C.J., Shi, W. & Adler, J., 1993. Adverse conditions which cause lack of flagella in Escherichia coli. *J. Bacteriol.*, 175(8), pp.2229–2235.

Lies, M. & Maurizi, M.R., 2008. Turnover of endogenous SsrA-tagged proteins mediated by ATP-dependent proteases in Escherichia coli. *The Journal of biological chemistry*, 283(34), pp.22918–29.

Lindsley, C.W., 2015. 2014 Global Prescription Medication Statistics: Strong Growth and CNS Well Represented. *ACS Chemical Neuroscience*, 6(4), pp.505–506.

Ling, M., Allen, S.W. & Wood, J.M., 1994. Sequence analysis identifies the proline dehydrogenase and delta 1-pyrroline-5-carboxylate dehydrogenase domains of the multifunctional Escherichia coli PutA protein. *Journal of molecular biology*, 243(5), pp.950–6.

Lino, T., 1974. Assembly of salmonella flagellin in vitro and in vivo. *Journal of Supramolecular Structure*, 2(2-4), pp.372–384.

Liu, L., Liu, Y., Shin, H., Chen, R.R., Wang, N.S., Li, J., Du, G. & Chen, J., 2013. Developing Bacillus spp. as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology. *Applied Microbiology and Biotechnology*, 97(14), pp.6113–6127.

Liu, R. & Ochman, H., 2007. Stepwise formation of the bacterial flagellar system. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17), pp.7116–21.

Liu, Z., Gosser, Y., Baker, P.J., Ravee, Y., Lu, Z., Alemu, G., Li, H., Butterfoss, G.L., Kong, X.-P., Gross, R. & Montclare, J.K., 2009. Structural and Functional Studies of *Aspergillus oryzae* Cutinase: Enhanced Thermostability and Hydrolytic Activity of Synthetic Ester and Polyester Degradation. *Journal of the American Chemical Society*, 131(43), pp.15711–15716.

Lizak, C., Fan, Y.-Y., Weber, T.C. & Aebi, M., 2011. N-Linked glycosylation of antibody fragments in Escherichia coli. *Bioconjugate chemistry*, 22(3), pp.488–96.

Lloyd, S.A., Forsberg, Å., Wolf-Watz, H. & Francis, M.S., 2001. Targeting exported substrates to the Yersinia TTSS: different functions for different signals? *Trends in Microbiology*, 9(8), pp.367–371.

Lobstein, J., Emrich, C.A., Jeans, C., Faulkner, M., Riggs, P. & Berkmen, M., 2012. SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial cell factories*, 11, p.56.

Love, K.R., Politano, T.J., Panagiotou, V., Jiang, B., Stadheim, T.A. & Love, J.C., 2012. Systematic single-cell analysis of Pichia pastoris reveals secretory capacity limits productivity. C. V. Rao, ed. *PloS one*, 7(6), p.e37915.

Lowe, C.R., Lowe, A.R. & Gupta, G., 2001. New developments in affinity chromatography with potential application in the production of biopharmaceuticals. *Journal of Biochemical and Biophysical Methods*, 49(1), pp.561–574.

Lubas, W.A., 2000. Functional Expression of O-linked GlcNAc Transferase. Domain structure and substrate specificity. *Journal of Biological Chemistry*, 275(15), pp.10983–10988.

Lucchetti-Miganeh, C., Burrowes, E., Baysse, C. & Ermel, G., 2008. The post-transcriptional regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages of infection in animal hosts. *Microbiology (Reading, England)*, 154(Pt 1), pp.16–29.

Lynn, A.K., Yannas, I. V & Bonfield, W., 2004. Antigenicity and immunogenicity of collagen. *Journal of biomedical materials research. Part B, Applied biomaterials*, 71(2), pp.343–54.

Macnab, R.M., 2003. How bacteria assemble flagella. *Annual review of microbiology*, 57, pp.77–100.

Macnab, R.M., 2004. Type III flagellar protein export and flagellar assembly. *Biochimica et biophysica acta*, 1694(1-3), pp.207–17.

Magnusson, L.U., Gummesson, B., Joksimović, P., Farewell, A. & Nyström, T., 2007. Identical, independent, and opposing roles of ppGpp and DksA in Escherichia coli. *Journal of bacteriology*, 189(14), pp.5193–202.

Majander, K., Anton, L., Antikainen, J., Lång, H., Brummer, M., Korhonen, T.K. & Westerlund-Wikström, B., 2005. Extracellular secretion of polypeptides using a modified Escherichia coli flagellar secretion apparatus. *Nature biotechnology*, 23(4), pp.475–81.

Makrides, S., 1996. Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiol. Rev.*, 60(3), pp.512–538.

Marmiesse,L., Peyraud,R & Cottret,L., 2015. FlexFlux: combining metabolic flux and regulatory network analyses. *BMC Systems Biology*, 9(1), p.93.

Matos, C.F.R.O., Branston, S.D., Albiniak, A., Dhanoya, A., Freedman, R.B., Keshavarz-Moore, E. & Robinson, C., 2012. High-yield export of a native heterologous protein to the periplasm by the tat translocation pathway in Escherichia coli. *Biotechnology and Bioengineering*, 109(10), pp.2533–2542.

Matos, C.F.R.O., Robinson, C., Alanen, H.I., Prus, P., Uchida, Y., Ruddock, L.W., Freedman, R.B. & Keshavarz-Moore, E., 2014. Efficient export of prefolded, disulfide-bonded recombinant proteins to the periplasm by the Tat pathway in Escherichia coli CyDisCo strains. *Biotechnology Progress*, 30(2), pp.281–290.

Maurer, L.M., Yohannes, E., Bondurant, S.S., Radmacher, M. & Slonczewski, J.L., 2005. pH regulates genes for flagellar motility, catabolism, and oxidative stress in Escherichia coli K-12. *Journal of bacteriology*, 187(1), pp.304–19.

Mayfield, S., 2013. The Green Revolution 2.0: the potential of algae for the production of biofuels and bioproducts [1]. *Genome*, 56(10), pp.551–555.

Mazzocchi, F., Complexity and the reductionism-holism debate in systems biology. *Wiley interdisciplinary reviews. Systems biology and medicine*, 4(5), pp.413–27.

McCormick, K. & Kautto, N., 2013. The Bioeconomy in Europe: An Overview. *Sustainability*, 5(6), pp.2589–2608.

Mellies, J., Brems, R. & Villarejo, M., 1994. The Escherichia coli proU promoter element and its contribution to osmotically signaled transcription activation. *Journal of bacteriology*, 176(12), pp.3638–45.

Mergulhão, F.J.M., Monteiro, G.A., Cabral, J. M. S. & Taipa, M. A. 2004.Design of Bacterial Vector Systems for the Production of Recombinant Proteins in Escherichia coli. *Journal of Microbiology and Biotechnology*, 14(1), pp.1–14.

Mergulhão, F.J.M., Summers, D.K. & Monteiro, G.A., 2005. Recombinant protein secretion in Escherichia coli. *Biotechnology advances*, 23(3), pp.177–202.

Michalczechen-Lacerda, V., Tokareva, O., Bastos, A. de R., da Silva, M., Vianna, G., Murad, A., Kaplan, D. & Rech, E., 2014. Synthetic biology increases efficiency of Escherichia coli to produce Parawixia bistriata spider silk protein. *BMC Proceedings*, 8(Suppl 4), p.P231.

Mijakovic, I., Petranovic, D., Macek, B., Cepo, T., Mann, M., Davies, J., Jensen, P.R. & Vujaklija, D., 2006. Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic acids research*, 34(5), pp.1588–96.

Minamino, T. & Macnab, R.M., 2000. Domain structure of Salmonella FlhB, a flagellar export component responsible for substrate specificity switching. *Journal of bacteriology*, 182(17), pp.4906–14.

Minamino, T. & Macnab, R.M., 2000. FliH, a soluble component of the type III flagellar export apparatus of Salmonella, forms a complex with FliI and inhibits its ATPase activity. *Molecular Microbiology*, 37(6), pp.1494–1503.

Minamino, T. & Namba, K., 2008. Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. *Nature*, 451(7177), pp.485–488.

Minamino, T., 2014. Protein export through the bacterial flagellar type III export pathway. *Biochimica et biophysica acta*, 1843(8), pp.1642–8.

Minamino, T., Morimoto, Y. V, Hara, N., Aldridge, P.D. & Namba, K., 2016. The Bacterial Flagellar Type III Export Gate Complex Is a Dual Fuel Engine That Can Use Both H+ and Na+ for Flagellar Protein Export. *PLoS pathogens*, 12(3), p.e1005495.

Mironova, R., Niwa, T., Handzhiyski, Y., Sredovska, A. & Ivanov, I., 2005. Evidence for non-enzymatic glycosylation of Escherichia coli chromosomal DNA. *Molecular microbiology*, 55(6), pp.1801–11.

Mizuno, S., Amida, H., Kobayashi, N., Aizawa, S.-I. & Tate, S.-I., 2011. The NMR structure of FliK, the trigger for the switch of substrate specificity in the flagellar type III secretion apparatus. *Journal of molecular biology*, 409(4), pp.558–73.

Munera, D., Crepin, V.F., Marches, O. & Frankel, G., 2010. N-terminal type III secretion signal of enteropathogenic Escherichia coli translocator proteins. *Journal of bacteriology*, 192(13), pp.3534–9.

Murata, T., Shinozuka, Y., Obata, Y. & Yokoyama, K.K., 2008. Phosphorylation of two eukaryotic transcription factors, Jun dimerization protein 2 and activation transcription factor 2, in Escherichia coli by Jun N-terminal kinase 1. *Analytical Biochemistry*, 376(1), pp.115–121.

Murphy, K.C., 1998. Use of Bacteriophage lambda Recombination Functions To Promote Gene Replacement in Escherichia coli. *J. Bacteriol.*, 180(8), pp.2063–2071.

Mytelka, D. & Chamberlin, M., 1996. Escherichia coli fliAZY operon. *J. Bacteriol.*, 178(1), pp.24–34.

Naas, T., Blot, M., Fitch, W.M. & Arber, W., 1994. Insertion Sequence-Related Genetic Variation in Resting Escherichia coli K-12. *Genetics*, 136(3), pp.721–730.

Nakao, T., Yamato, I. & Anraku, Y., 1988. Mapping of the multiple regulatory sites for putP and putA expression in the putC region of Escherichia coli. *Molecular & general genetics : MGG*, 214(3), pp.379–88.

Nambu, T., Minamino, T., Macnab, R.M. & Kutsukake, K., 1999. Peptidoglycan-hydrolyzing activity of the FlgJ protein, essential for flagellar rod formation in Salmonella typhimurium. *Journal of bacteriology*, 181(5), pp.1555–61.

Narayanan, N., Khan, M. & Chou, C.P., 2010. Enhancing functional expression of heterologous lipase B in Escherichia coli by extracellular secretion. *Journal of industrial microbiology & biotechnology*, 37(4), pp.349–61.

Neubauer, A., Neubauer, P. & Myllyharju, J., 2005. High-level production of human collagen prolyl 4-hydroxylase in Escherichia coli. *Matrix biology : journal of the International Society for Matrix Biology*, 24(1), pp.59–68.

Ni, Y. & Chen, R., 2009. Extracellular recombinant protein production from Escherichia coli. *Biotechnology letters*, 31(11), pp.1661–70.

Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein engineering*, 10(1), pp.1–6.

Nivaskumar, M., Bouvier, G., Campos, M, Nadeau, N., Yu, X., Egelman, E.H., Nigeles, M., & Francetic, O. 2007. Distinct Docking and Stabilization Steps of the Pseudopilus Conformational Transition Path Suggest Rotational Assembly of Type IV Pilus-like Fibers. *Structure*, 22(5), pp.685–696

Nivaskumar, M. & Francetic, O., 2014. Type II secretion system: A magic beanstalk or a protein escalator. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), pp.1568–1577.

Nokelainen, M., Helaakoski, T., Myllyharju, J., Notbohm, H., Pihlajaniemi, T., Fietzek, P.P. & Kivirikko, K.I., 1998. Expression and characterization of recombinant human type II collagens with low and high contents of hydroxylysine and its glycosylated forms. *Matrix Biology*, 16(6), pp.329–338.

Nokelainen, M., Tu, H., Vuorela, A., Notbohm, H., Kivirikko, K.I. & Myllyharju, J., 2001. High-level production of human type I collagen in the yeast Pichia pastoris. *Yeast (Chichester, England)*, 18(9), pp.797–806.

Notbohm, H., 1999. Recombinant Human Type II Collagens with Low and High Levels of Hydroxylysine and Its Glycosylated Forms Show Marked Differences in Fibrillogenesis in Vitro. *Journal of Biological Chemistry*, 274(13), pp.8988–8992.

Nuutila, K., Peura, M., Suomela, S., Hukkanen, M., Siltanen, A., Harjula, A., Vuola, J. & Kankuri, E., 2015. Recombinant human collagen III gel for transplantation of autologous skin cells in porcine full-thickness wounds. *Journal of Tissue Engineering and Regenerative Medicine*, 9(12), pp.1386–1393.

Ohnishi, K., Ohto, Y., Aizawa, S., Macnab, R.M. & Iino, T., 1994. FlgD is a scaffolding protein needed for flagellar hook assembly in Salmonella typhimurium. *Journal of bacteriology*, 176(8), pp.2272–81.

Olmos-Soto, J. & Contreras-Flores, R., 2003. Genetic system constructed to overproduce and secrete proinsulin in Bacillus subtilis. *Applied Microbiology and Biotechnology*, 62(4), pp.369–373.

Olsen, D., 2003. Recombinant collagen and gelatin for drug delivery. *Advanced Drug Delivery Reviews*, 55(12), pp.1547–1567.

Orth, J.D., Thiele, I. & Palsson, B.Ø., 2010. What is flux balance analysis? *Nature biotechnology*, 28(3), pp.245–8.

Oshima, T., Ishikawa, S., Kurokawa, K., Aiba, H. & Ogasawara, N., 2006. Escherichia coli histone-like protein H-NS preferentially binds to horizontally acquired DNA in association with RNA polymerase. *DNA research : an international journal for rapid publication of reports on genes and genomes*, 13(4), pp.141–53.

Ozin, A.J., Claret, L., Auvray, F. & Hughes, C., 2003. The FliS chaperone selectively binds the disordered flagellin C-terminal D0 domain central to polymerisation.

Paddon, C.J., Westfall, P. J., Pitera, D. J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M. D., Tai, A., Main, A., Eng, D., Polichuk, D. R., Teoh, K. H., Reed, D. W., Treynor, T., Lenihan, J., Jiang, H., Fleck, M., Bajad, S., Dang, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, S., Fickes, S & Galazzo, J,., Gaucher, S. P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsurata, H., Vazques, R., Westbalde, L. F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P. S., Keasling, J. D., Reiling, K. K., Renninger, N. S. & Newman, J. D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, 496(7446), pp.528–32.

Paes, L.S., MAntilla, B. S., Zimbres, F. M., Pral, E. M. F., de Melo, P. D., Tahara, E. B., Kowaltowski, A. J., Elias, M. C. & Silber, A. M., 2013. Proline Dehydrogenase Regulates Redox State and Respiratory Metabolism in *Trypanosoma cruz*i. *PLoS ONE*, 8(7), p.e69419.

Pandhal, J., Desai, P., Walpole, C., Doroudi, L., Malyshev, D. & Wright, P.C., 2012. *Systematic metabolic engineering for improvement of glycosylation efficiency in Escherichia coli*,

Panduranga Rao, K., 1996. Recent developments of collagen-based materials for medical applications and drug delivery systems. *Journal of Biomaterials Science, Polymer Edition*, 7(7), pp.623–645.

Parkinson, J.S., 1978. Complementation analysis and deletion mapping of Escherichia coli mutants defective in chemotaxis. *J. Bacteriol.*, 135(1), pp.45–53.

Patel, R., Smith, S.M. & Robinson, C., 2014. Protein transport by the bacterial Tat pathway. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), pp.1620–1628.

Patino, M.G., Neiders, M.E., Andreana, S., Noble, B. & Cohen, R.E., 2002. Collagen as an implantable material in medicine and dentistry. *The Journal of oral implantology*, 28(5), pp.220–5.

Patten, C.L., Kirchhof, M.G., Schertzberg, M.R., Morton, R.A. & Schellhorn, H.E., 2004. Microarray analysis of RpoS-mediated gene expression in Escherichia coli K-12. *Molecular Genetics and Genomics*, 272(5), pp.580–591.

Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W. & Gourse, R.L., 2004. DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell*, 118(3), pp.311–22.

Peng, Y.Y., Howell, L., Stoichevska, V., Werkmeister, J.A., Dumsday, G.J. & Ramshaw, J.A.M., 2012. Towards scalable production of a collagen-like protein from Streptococcus pyogenes for biomedical applications. *Microbial cell factories*, 11, p.146.

Peplow, M., 2013. Malaria drug made in yeast causes market ferment. *Nature*, 494(7436), pp.160–1.

Peplow, M., 2016. Synthetic biology's first malaria drug meets market resistance. *Nature*, 530(7591), pp.389–390.

Perdivara, I., Perera, L., Sricholpech, M., Terajima, M., Pleshko, N., Yamauchi, M. & Tomer, K.B., 2013. Unusual fragmentation pathways in collagen glycopeptides. *Journal of the American Society for Mass Spectrometry*, 24(7), pp.1072–81.

Pesavento, C. & Hengge, R., 2012. The global repressor FliZ antagonizes gene expression by σS-containing RNA polymerase due to overlapping DNA binding specificity. *Nucleic acids research*, 40(11), pp.4783–93.

Petersen, S.B., Fojan, P., Petersen, E.I. & Petersen, M.T.N., 2001. The Thermal Stability of the Fusarium solani pisi Cutinase as a Function of pH. *Journal of biomedicine & biotechnology*, 1(2), pp.62–69.

Petsch, D. & Anspach, F.B., 2000. Endotoxin removal from protein solutions. *Journal of biotechnology*, 76(2-3), pp.97–119.

Pinkas, D.M., Ding, S., Raines, R.T. & Barron, A.E., 2011. Tunable, post-translational hydroxylation of collagen Domains in Escherichia coli. *ACS chemical biology*, 6(4), pp.320–4.

Prabakaran, P., Vu, B.K., Gan, J., Feng, Y., Dimitrov, D.S. & Ji, X., 2008. Structure of an isolated unglycosylated antibody C(H)2 domain. *Acta crystallographica. Section D, Biological crystallography*, 64(Pt 10), pp.1062–7.

Prinz, W.A., Aslund, F., Holmgren, A. & Beckwith, J., 1997. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm. *The Journal of biological chemistry*, 272(25), pp.15661–7.

Prüss, B.M. & Matsumura, P., 1996. A regulator of the flagellar regulon of Escherichia coli, flhD, also affects cell division. *Journal of bacteriology*, 178(3), pp.668–74.

Prüß, B.M., Liu, X., Hendrickson, W. & Matsumura, P., 2001. FlhD/FlhC-regulated promoters analyzed by gene array and lacZ gene fusions. *FEMS Microbiology Letters*, 197(1).

Prüss, B.M., Campbell, J.W., Van Dyk, T.K., Zhu, C., Kogan, Y. & Matsumura, P., 2003. FlhD/FlhC is a regulator of anaerobic respiration and the Entner-Doudoroff pathway through induction of the methyl-accepting chemotaxis protein Aer. *Journal of bacteriology*, 185(2), pp.534–43.

Purnick, P.E.M. & Weiss, R., 2009. The second wave of synthetic biology: from modules to systems. *Nature reviews. Molecular cell biology*, 10(6), pp.410–22.

Ramshaw, J.A., Shah, N.K. & Brodsky, B., 1998. Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides. *Journal of structural biology*, 122(1-2), pp.86–91.

Reed, J.L., Vo, T.D., Schilling, C.H. & Palsson, B.O., 2003. An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). *Genome biology*, 4(9), p.R54.

Retallack, D.M., Jin, H. & Chew, L., 2012. Reliable protein production in a Pseudomonas fluorescens expression system. *Protein Expression and Purification*, 81(2), pp.157–165.

Roberts, I.M., 1985. Hydrolysis of 4-methylumbelliferyl butyrate: A convenient and sensitive fluorescent assay for lipase activity. *Lipids*, 20(4), pp.243–247.

Roberts, M.A.J., Cranenburgh, R.M., Stevens, M.P. & Oyston, P.C.F., 2013. Synthetic Biology. Biology by design. *Microbiology (Reading, England)*, 159(Pt_7), pp.1219–1220.

Roblin, P., Dewitte, F., Villeret, V., Biondi, E.G. & Bompard, C., 2015. A Salmonella type three secretion effector/chaperone complex adopts a hexameric ring-like structure. *Journal of bacteriology*, 197(4), pp.688–98.

Rodríguez-Limas, W.A., Tannenbaum, V. & Tyo, K.E.J., 2015. Blocking endocytotic mechanisms to improve heterologous protein titers in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 112(2), pp.376–385.

Rosano, G.L. & Ceccarelli, E.A., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in microbiology*, 5, p.172.

Rudolph, R. & Lilie, H., 1996. In vitro folding of inclusion body proteins. *FASEB J*, 10(1), pp.49–56.

Rushlow, K.E., Deutch, A.H. & Smith, C.J., 1985. *Identification of a mutation that relieves gamma-glutamyi kinase from allosteric feedback inhibition by proline*,

Rutherford, S.T., Villers, C.L., Lee, J.-H., Ross, W. & Gourse, R.L., 2009. Allosteric control of Escherichia coli rRNA promoter complexes by DksA. *Genes & Development*, 23(2), pp.236–248.

Rutschmann, C., Baumann, S., Cabalzar, J., Luther, K.B. & Hennet, T., 2014. Recombinant expression of hydroxylated human collagen in Escherichia coli. *Applied Microbiology and Biotechnology*, 98(10), pp.4445–4455.

Sahdev, S., Khattar, S.K. & Saini, K.S., 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and cellular biochemistry*, 307(1-2), pp.249–64.

Saïda, F., Uzan, M., Odaert, B. & Bontems, F., 2006. Expression of highly toxic genes in E. coli: special strategies and genetic tools. *Current protein & peptide science*, 7(1), pp.47–56.

Saito, T., Ueno, T., Kubori, T., Yamaguchi, S., Iino, T. & Aizawa, S.-I., 1998. Flagellar filament elongation can be impaired by mutations in the hook protein FlgE of Salmonella typhimurium : a possible role of the hook as a passage for the anti-sigma factor FlgM. *Molecular Microbiology*, 27(6), pp.1129–1139.

Samatey, F.A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M. & Namba, K., 2001. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature*, 410(6826), pp.331–7.

Sandkvist, M., 2001. Biology of type II secretion. *Molecular Microbiology*, 40(2), pp.271–283.

Saraswat, M., Musante, L., Ravidá, A., Shortt, B., Byrne, B., Holthofer, H., Saraswat, M., Musante, L., Ravidá, A., Shortt, B., Byrne, B., Holthofer, H., Byrne, B. & Holthofer, H., 2013. Preparative purification of recombinant proteins: current status and future trends. *BioMed research international*, 2013, p.312709.

Schädlich, L., Senger, T., Kirschning, C.J., Müller, M. & Gissmann, L., 2009. Refining HPV 16 L1 purification from E. coli: reducing endotoxin contaminations and their impact on immunogenicity. *Vaccine*, 27(10), pp.1511–22.

Schellekens, H., 2002. Immunogenicity of therapeutic proteins: Clinical implications and future prospects. *Clinical Therapeutics*, 24(11), pp.1720–1740.

Schlegel, S., Rujas, E., Ytterberg, A.J., Zubarev, R.A., Luirink, J. & de Gier, J.-W., 2013. Optimizing heterologous protein production in the periplasm of E. coli by regulating gene expression levels. *Microbial cell factories*, 12(1), p.24.

Schulz, R.M. & Bader, A., 2007. Cartilage tissue engineering and bioreactor systems for the cultivation and stimulation of chondrocytes. *European biophysics journal : EBJ*, 36(4-5), pp.539–68.

Schweder, T., Lee, K.H., Lomovskaya, O. & Matin, A., 1996. Regulation of Escherichia coli starvation sigma factor (sigma s) by ClpXP protease. *Journal of bacteriology*, 178(2), pp.470–6.

Service, R.F., 2002. Mammalian Cells Spin a Spidery New Yarn. *Science*, 295(5554).

Setina, C.M., Haase, J.P. & Glatz, C.E., 2016. Process integration for recovery of recombinant collagen type I α1 from corn seed. *Biotechnology Progress*, 32(1), pp.98–107.

Shiloach, J. & Fass, R., 2005. Growing E. coli to high cell density—A historical perspective on method development. *Biotechnology Advances*, 23(5), pp.345–357.

Shin, S. & Park, C., 1995. Modulation of flagellar expression in Escherichia coli by acetyl phosphate and the osmoregulator OmpR. *Journal of bacteriology*, 177(16), pp.4696–702.

Shneider, M.M., Buth, S.A., Ho, B.T., Basler, M., Mekalanos, J.J. & Leiman, P.G., 2013. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*, 500(7462), pp.350–3.

Shoulders, M.D. & Raines, R.T., 2009. Collagen structure and stability. *Annual review of biochemistry*, 78, pp.929–58.

Singer, H.M., Erhardt, M., Steiner, A.M., Zhang, M.-M., Yoshikami, D., Bulaj, G., Olivera, B.M. & Hughes, K.T., 2012. Selective purification of recombinant neuroactive peptides using the flagellar type III secretion system. *mBio*, 3(3), pp.e00115–12–.

Singer, H.M., Erhardt, M. & Hughes, K.T., 2013. RflM functions as a transcriptional repressor in the autogenous control of the Salmonella Flagellar master operon flhDC. *Journal of bacteriology*, 195(18), pp.4274–82.

Singer, H.M., Erhardt, M. & Hughes, K.T., 2014. Comparative analysis of the secretion capability of early and late flagellar type III secretion substrates. *Molecular Microbiology*, 93(3), pp.505–520.

Singh, S.M. & Panda, A.K., 2005. Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering*, 99(4), pp.303–310.

Siuti, P., Yazbek, J. & Lu, T.K., 2013. Synthetic circuits integrating logic and memory in living cells. *Nature biotechnology*, 31(5), pp.448–52.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J. & Li, Q., 2009. Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein science : a publication of the Protein Society*, 18(5), pp.936–48.

Stringer, A.M., Singh, N., Yermakova, A., Petrone, B. L., Amarasinghe, J. J., Reyes-Diaz, L., Mantis, N. J. & Wade, J. T., 2012. FRUIT, a Scar-Free System for Targeted Chromosomal Mutagenesis, Epitope Tagging, and Promoter Replacement in Escherichia coli and Salmonella enterica M. A. Webber, ed. *PLoS ONE*, 7(9), p.e44841.

Sleight, S.C., Bartley, B.A., Lieviant, J.A. & Sauro, H.M., 2010. Designing and engineering evolutionary robust genetic circuits. *Journal of biological engineering*, 4(1), p.12.

Soares, C.R.J., Gomide, F.I.C., Ueda, E.K.M. & Bartolini, P., 2003. Periplasmic expression of human growth hormone via plasmid vectors containing the lambdaPL promoter: use of HPLC for product quantification. *Protein engineering*, 16(12), pp.1131–8.

Soini, J., Ukkonen, K. & Neubeuer, P., 2008. High cell density media for Escherichia coli are generally designed for aerobic cultivations – consequences for large-scale bioprocesses and shake flask cultures. *Microbial Cell Factories*, 7(1), p.26.

Sørensen, H.P. & Mortensen, K.K., 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microbial cell factories*, 4(1), p.1.

Soutourina, O., Kolb, A., Krin, E., Laurent-Winter, C., Rimsky, S., Danchin, A. & Bertin, P., 1999. Multiple control of flagellum biosynthesis in Escherichia coli: role of H-NS protein and the cyclic AMP-catabolite activator protein complex in transcription of the flhDC master operon. *Journal of bacteriology*, 181(24), pp.7500–8.

Soutourina, O.A. & Bertin, P.N., 2003. Regulation cascade of flagellar expression in Gram-negative bacteria. *FEMS Microbiology Reviews*, 27(4).

Sperandio, V., Torres, A.G. & Kaper, J.B., 2002. Quorum sensing *Escherichia coli* regulators B and C (QseBC): a novel two-component regulatory system involved in the regulation of flagella and motility by quorum sensing in *E. coli*. *Molecular Microbiology*, 43(3), pp.809–821.

Stafford, G.P., Ogi, T. & Hughes, C., 2005. Binding and transcriptional activation of non-flagellar genes by the Escherichia coli flagellar master regulator FlhD2C2. *Microbiology (Reading, England)*, 151(Pt 6), pp.1779–88.

Stafford, G.P., Evans, L.D.B., Krumscheid, R., Dhillon, P., Fraser, G.M. & Hughes, C., 2007. *Sorting of Early and Late Flagellar Subunits After Docking at the Membrane ATPase of the Type III Export Pathway*,

Steen, P. Van den, Rudd, P.M., Dwek, R.A. & Opdenakker, G., 2008. Concepts and Principles of O-Linked Glycosylation.

Stein, H., Wilensky, M., Tsafrir, Y., Rosenthal, M., Amir, R., Avraham, T., Ofir, K., Dgany, O., Yayon, A. & Shoseyov, O., 2009. Production of bioactive, post-translationally modified, heterotrimeric, human recombinant type-I collagen in transgenic tobacco. *Biomacromolecules*, 10(9), pp.2640–5.

Szczebara, F.M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., Blanchard, S., Groisillier, A., Testet, E., Costaglioli, P., Cauet, G., Degryse, E., Balbuena, D., Winter, J., Achstetter, T., Spagnoli, R., Pompon, D. & Dumas, B., 2003. Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nature biotechnology*, 21(2), pp.143–9.

Takaya, A., Matsui, M., Tomoyasu, T., Kaya, M. & Yamamoto, T., 2006. The DnaK chaperone machinery converts the native FlhD2C2 hetero-tetramer into a functional transcriptional regulator of flagellar regulon expression in Salmonella. *Molecular Microbiology*, 59(4), pp.1327–1340.

Takaya, A., Erhardt, M., Karata, K., Winterberg, K., Yamamoto, T. & Hughes, K.T., 2012. YdiV: a dual function protein that targets FlhDC for ClpXP-dependent degradation by promoting release of DNA-bound FlhDC complex. *Molecular microbiology*, 83(6), pp.1268–84.

Takors, R. & de Lorenzo, V., 2016. Editorial overview: Microbial systems biology: systems biology prepares the ground for successful synthetic biology. *Current opinion in microbiology*.

Tans-Kersten, J., Brown, D. & Allen, C., 2004. Swimming Motility, a Virulence Trait of *Ralstonia solanacearum* , Is Regulated by FlhDC and the Plant Host Environment. *Molecular Plant-Microbe Interactions*, 17(6), pp.686–695.

Taylor, S.C., Berkelman, T., Yadav, G. & Hammond, M., 2013. A defined methodology for reliable quantification of Western blot data. *Molecular biotechnology*, 55(3), pp.217–26.

Temme, K., Zhao, D. & Voigt, C.A., 2012. Refactoring the nitrogen fixation gene cluster from Klebsiella oxytoca. *Proceedings of the National Academy of Sciences of the United States of America*, 109(18), pp.7085–90.

Thomas, J., Stafford, G.P. & Hughes, C., 2004. Docking of cytosolic chaperone-substrate complexes at the membrane ATPase during flagellar type III protein export. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11), pp.3945–50.

Thomas, S., Holland, I.B. & Schmitt, L., 2014. The Type 1 secretion pathway — The hemolysin system and beyond. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), pp.1629–1641.

Tillet, E., Wiedemann, H., Golbik, R., Pan, T.-C., Zhang, R.-Z., Mann, K., Chu, M.-L. & Timpl, R., 1994. Recombinant expression and structural and binding properties of alpha1(VI) and alpha2(VI) chains of human collagen type VI. *European Journal of Biochemistry*, 221(1), pp.177–187.

Toman, P.D., Chisholm, G., McMullin, H., Giere, L.M., Olsen, D.R., Kovach, R.J., Leigh, S.D., Fong, B.E., Chang, R., Daniels, G.A., Berg, R.A. & Hitzeman, R.A., 2000. Production of recombinant human type I procollagen trimers using a four-gene expression system in the yeast Saccharomyces cerevisiae. *The Journal of biological chemistry*, 275(30), pp.23303–9.

Tomoyasu, T., Ohkishi, T., Ukyo, Y., Tokumitsu, A., Takaya, A., Suzuki, M., Sekiya, K., Matsui, H., Kutsukake, K. & Yamamoto, T., 2002. The ClpXP ATP-Dependent Protease Regulates Flagellum Synthesis in Salmonella enterica Serovar Typhimurium. *Journal of Bacteriology*, 184(3), pp.645–653.

Tomoyasu, T., Takaya, A., Isogai, E. & Yamamoto, T., 2003. Turnover of FlhD and FlhC, master regulator proteins for Salmonella flagellum biogenesis, by the ATP-dependent ClpXP protease. *Molecular Microbiology*, 48(2), pp.443–452.

Tseng, T.-T., Tyler, B.M. & Setubal, J.C., 2009. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC microbiology*, 9 Suppl 1(Suppl 1), p.S2.

Turner, L., Stern, A.S. & Berg, H.C., 2012. Growth of flagellar filaments of Escherichia coli is independent of filament length. *Journal of bacteriology*, 194(10), pp.2437–42.

Van der Wal, F.J., Koningstein, G., ten Hagen, C.M., Oudega, B. & Luirink, J., 1998. Optimization of bacteriocin release protein (BRP)-mediated protein release by Escherichia coli: random mutagenesis of the pCloDF13-derived BRP gene to uncouple lethality and quasi-lysis from protein release. *Applied and environmental microbiology*, 64(2), pp.392–8.

Van Susante, J.L.., Pieper, J., Buma, P., van Kuppevelt, T.H., van Beuningen, H., van der Kraan, P.M., Veerkamp, J.H., van den Berg, W.B. & Veth, R.P.., 2001. Linkage of chondroitin-sulfate to type I collagen scaffolds stimulates the bioactivity of seeded chondrocytes in vitro. *Biomaterials*, 22(17), pp.2359–2369.

Van Ulsen, P., Rahman, S. ur, Jong, W.S.P., Daleke-Schermerhorn, M.H. & Luirink, J., 2014. Type V secretion: From biogenesis to biotechnology. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1843(8), pp.1592–1611.

Vaneechoutte, M., Verschraegen, G., Claeys, G. & Flamen, P., 1988. Rapid identification of Branhamella catarrhalis with 4-methylumbelliferyl butyrate. *Journal of clinical microbiology*, 26(6), pp.1227–8.

Végh, B.M., Gál, P., Dobó, J., Závodszky, P. & Vonderviszt, F., 2006. Localization of the flagellum-specific secretion signal in Salmonella flagellin. *Biochemical and Biophysical Research Communications*, 345(1), pp.93–98.

Villar, G., Graham, A.D. & Bayley, H., 2013. A Tissue-Like Printed Material. *Science*, 340(6128).

Vonderviszt, F., Ishima, R., Akasaka, K. & Aizawa, S.-I., 1992. *Terminal disorder: A common structural feature of the axial proteins of bacterial flagellum?*,

Vuorela, A., Myllyharju, J., Nissi, R., Pihlajaniemi, T. & Kivirikko, K.I., 1997. Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast pichia pastoris: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase. *The EMBO journal*, 16(22), pp.6702–12.

Wacker, M., Linton, D., Hitchen, P.G., Nita-Lazar, M., Haslam, S.M., North, S.J., Panico, M., Morris, H.R., Dell, A., Wren, B.W. & Aebi, M., 2002. N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli. *Science (New York, N.Y.)*, 298(5599), pp.1790–3.

Wada, T., Hatamoto, Y. & Kutsukake, K., 2012. Functional and expressional analyses of the anti-FlhD4C2 factor gene ydiV in Escherichia coli. *Microbiology (Reading, England)*, 158(Pt 6), pp.1533–42.

Walsh, G., 2014. Biopharmaceutical benchmarks 2014. *Nature Biotechnology*, 32(10), pp.992–1000.

Wandersman, C. & Coulthurst, S.J., 2013. The Type VI secretion system – a widespread and versatile cell targeting system. *Research in Microbiology*, 164(6), pp.640–654.

Wang, Q., Zhao, Y., McClelland, M. & Harshey, R.M., 2007. The RcsCDB signaling system and swarming motility in Salmonella enterica serovar typhimurium: dual regulation of flagellar and SPI-2 virulence genes. *Journal of bacteriology*, 189(23), pp.8447–57.

Wang, X. & Wood, T.K., 2011. IS5 inserts upstream of the master motility operon flhDC in a quasi-Lamarckian way. *The ISME journal*, 5(9), pp.1517–25.

Wang, Y., Huang, H., Sun, M., Zhang, Q. & Guo, D., 2012. T3DB: an integrated database for bacterial type III secretion system. *BMC bioinformatics*, 13(1), p.66.

Warming, S., Costantino, N., Court, D.L., Jenkins, N.A. & Copeland, N.G., 2005. Simple and highly efficient BAC recombineering using galK selection. *Nucleic acids research*, 33(4), p.e36.

Waugh, D.S., 2011. An overview of enzymatic reagents for the removal of affinity tags. *Protein expression and purification*, 80(2), pp.283–93.

Wee, D.H. & Hughes, K.T., 2015. Molecular ruler determines needle length for the Salmonella Spi-1 injectisome. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13), pp.4098–103.

Weerasuriya, S., Schneider, B.M. & Manson, M.D., 1998. Chimeric chemoreceptors in Escherichia coli: signaling properties of Tar-Tap and Tap-Tar hybrids. *Journal of bacteriology*, 180(4), pp.914–20.

Wei, B.L., Brun-Zinkernagel, A.M., Simecka, J.W., Prüss, B.M., Babitzke, P. & Romeo, T., 2001. Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli. *Molecular microbiology*, 40(1), pp.245–56.

Weichart, D., Querfurth, N., Dreger, M. & Hengge-Aronis, R., 2003. Global Role for ClpP-Containing Proteases in Stationary-Phase Adaptation of Escherichia coli. *Journal of Bacteriology*, 185(1), pp.115–125.

Weinacker, D., Rabert, C., Zepeda, A.B., Figueroa, C.A., Pessoa, A. & Farías, J.G., 2013. Applications of recombinant Pichia pastoris in the healthcare industry. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*, 44(4), pp.1043–8.

Weis, R., Luiten, R., Skranc, W., Schwab, H., Wubbolts, M. & Glieder, A., 2004. Reliable high-throughput screening with by limiting yeast cell death phenomena. *FEMS Yeast Research*, 5(2), pp.179–189.

Werten, M.W.T., van den Bosch, T.J., Wind, R.D., Mooibroek, H. & de Wolf, F.A., 1999. High-yield secretion of recombinant gelatins byPichia pastoris. *Yeast*, 15(11), pp.1087–1096.

Widmaier, D.M., Tullman-Ercek, D., Mirsky, E.A., Hill, R., Govindarajan, S., Minshull, J. & Voigt, C.A., 2009. Engineering the Salmonella type III secretion system to export spider silk monomers. *Molecular systems biology*, 5, p.309.

Williams, A.W., Yamaguchi, S., Togashi, F., Aizawa, S.I., Kawagishi, I. & Macnab, R.M., 1996. Mutations in fliK and flhB affecting flagellar hook and filament assembly in Salmonella typhimurium. *Journal of bacteriology*, 178(10), pp.2960–70.

Winter, J.M. & Tang, Y., 2012. Synthetic biological approaches to natural product biosynthesis. *Current Opinion in Biotechnology*, 23(5), pp.736–743.

Wood, J.M., 1988. Proline porters effect the utilization of proline as nutrient or osmoprotectant for bacteria. *The Journal of Membrane Biology*, 106(3), pp.183–202.

Wood, J.M., Bremer, E., Csonka, L.N., Kraemer, R., Poolman, B., van der Heide, T. & Smith, L.T., 2001. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comparative biochemistry and physiology. Part A, Molecular & integrative physiology*, 130(3), pp.437–60.

Wozniak, C.E., Lee, C. & Hughes, K.T., 2009. T-POP array identifies EcnR and PefI-SrgD as novel regulators of flagellar gene expression. *Journal of bacteriology*, 191(5), pp.1498–508.

Xia, X.-X., Qian, Z.-G., Ki, C.S., Park, Y.H., Kaplan, D.L. & Lee, S.Y., 2010. Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. *Proceedings of the National Academy of Sciences of the United States of America*, 107(32), pp.14059–63.

Xu, J., Wang, L.N., Zhu, C.H., Fan, D.D., Ma, X.X., Mi, Y. & Xing, J.Y., 2015. Co-expression of recombinant human prolyl with human collagen α 1 (III) chains in two yeast systems. *Letters in Applied Microbiology*, 61(3), pp.259–266.

Xu, R., Luo, Y.-E., Fan, D.-D., Guo, L., Xi, J.-F., Mi, Y. & Ma, P., Improving the production of human-like collagen by pulse-feeding glucose during the fed-batch culture of recombinant Escherichia coli. *Biotechnology and applied biochemistry*, 59(5), pp.330–7.

Xu, X., Gan, Q., Clough, R.C., Pappu, K.M., Howard, J.A., Baez, J.A. & Wang, K., 2011. Hydroxylation of recombinant human collagen type I alpha 1 in transgenic maize co-expressed with a recombinant human prolyl 4-hydroxylase. *BMC biotechnology*, 11(1), p.69.

Yamamoto, S. & Kutsukake, K., 2006. FliT acts as an anti-FlhD2C2 factor in the transcriptional control of the flagellar regulon in Salmonella enterica serovar typhimurium. *Journal of bacteriology*, 188(18), pp.6703–8.

Yanagihara, S., Iyoda, S., Ohnishi, K., Iino, T. & Kutsukake, K., 1999. Structure and transcriptional control of the flagellar master operon of Salmonella typhimurium. *Genes & genetic systems*, 74(3), pp.105–11.

Yang, S., Xu, H., Yan, Q., Liu, Y., Zhou, P. & Jiang, Z., 2013. A low molecular mass cutinase of Thielavia terrestris efficiently hydrolyzes poly(esters). *Journal of industrial microbiology & biotechnology*, 40(2), pp.217–26.

Ying, T., Gong, R., Ju, T.W., Prabakaran, P. & Dimitrov, D.S., 2014. Engineered Fc based antibody domains and fragments as novel scaffolds. *Biochimica et biophysica acta*, 1844(11), pp.1977–1982.

Yokoseki, T., Iino, T. & Kutsukake, K., 1996. Negative regulation by fliD, fliS, and fliT of the export of the flagellum-specific anti-sigma factor, FlgM, in Salmonella typhimurium. *J. Bacteriol.*, 178(3), pp.899–901.

Yokoseki, T., Kutsukake, K., Ohnishi, K. & Iino, T., 1995. Functional analysis of the flagellar genes in the fliD operon of Salmonella typhimurium. *Microbiology (Reading, England)*, 141 (Pt 7, pp.1715–22.

Yonekura, K., Maki, S., Morgan, D.G., DeRosier, D.J., Vonderviszt, F., Imada, K. & Namba, K., 2000. The bacterial flagellar cap as the rotary promoter of flagellin self-assembly. *Science (New York, N.Y.)*, 290(5499), pp.2148–52.

Yonekura, K., Maki-Yonekura, S. & Namba, K., 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature*, 424(6949), pp.643–50.

Yoon, S.H., Kim, S.K. & Kim, J.F., 2010. Secretory production of recombinant proteins in Escherichia coli. *Recent patents on biotechnology*, 4(1), pp.23–9.

Young, E. & Alper, H., 2010. Synthetic Biology: Tools to Design, Build, and Optimize Cellular Processes.

Yu, Z., An, B., Ramshaw, J.A.M. & Brodsky, B., 2014. Bacterial collagen-like proteins that form triple-helical structures. *Journal of Structural Biology*, 186(3), pp.451–461.

Yue, B.G., Ajuh, P., Akusjärvi, G., Lamond, A.I. & Kreivi, J.P., 2000. Functional coexpression of serine protein kinase SRPK1 and its substrate ASF/SF2 in Escherichia coli. *Nucleic acids research*, 28(5), p.E14.

Zechner, E.L., Lang, S. & Schildbach, J.F., 2012. Assembly and mechanisms of bacterial type IV secretion machines. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 367(1592), pp.1073–87.

Zhang, C., Baez, J. & Glatz, C.E., 2009. Purification and characterization of a 44-kDa recombinant collagen I alpha 1 fragment from corn grain. *Journal of agricultural and food chemistry*, 57(3), pp.880–7.

Zhang, F., Yin, Y., Arrowsmith, C.H. & Ling, V., 1995. Secretion and Circular Dichroism Analysis of the C-Terminal Signal Peptides of HlyA and LktA. *Biochemistry*, 34(13), pp.4193–4201.

Zhao, K., Liu, M. & Burgess, R.R., 2007. Adaptation in bacterial flagellar and motility systems: from regulon members to “foraging”-like behavior in E. coli. *Nucleic acids research*, 35(13), pp.4441–52.

Zhao, X., Li, G., Liang, S., Zhao, X., Li, G. & Liang, S., 2013. Several affinity tags commonly used in chromatographic purification. *Journal of analytical methods in chemistry*, 2013, p.581093.

Zhi, H., Wang, J., Wang, S., Wei, Y., Zhi, H., Wang, J., Wang, S. & Wei, Y., 2013. Fluorescent Properties of Hymecromone and Fluorimetric Analysis of Hymecromone in Compound Dantong Capsule. *Journal of Spectroscopy*, 2013, pp.1–9.

Zhou, J., Sharp, L.L., Tang, H.L., Lloyd, S.A., Billings, S., Braun, T.F. & Blair, D.F., 1998. Function of Protonatable Residues in the Flagellar Motor of Escherichia coli: a Critical Role for Asp 32 of MotB. *Journal of Bacteriology*, 180(10), pp.2729–2735.

Zucca, S., Pasotti, L., Politi, N., Cusella De Angelis, M.G. & Magni, P., 2013. A standard vector for the chromosomal integration and characterization of BioBrick™ parts in Escherichia coli. *Journal of biological engineering*, 7(1), p.12.

Appendix

Appendix 1: Flagella type III secretion genes proteins.

Classified into operon class, cellular location, function, stoichiometry, assembly method and size of protein. Collated from Macnab 2003; Berg 2003; Stafford & Hughes 2007; Raha et al. 1994; Pesavento & Hengge 2012; Mytelka and Chamberlin, 1996

Protein	Operon Class	Cellular Location	Function	Stoichiometry	Size (kDa)
FlhA	2	Centre of MS ring	Protein export component	~2	75
FlhB	2	Centre of MS ring	Protein export component	~2	42
FlhC	1	Cytoplasm	Master regulator of class 2 operons		22
FlhD	1	Cytoplasm	Master regulator of class 2 operons		14
FlhE	2	Cell envelope protein	Role in swarming motility		12
FliA	2	Cytoplasm	Sigma factor 28, for class 3 operons		27
FliC	3	Cell exterior	Filament (flagellin)	20000-30000	53
FliD	3	Cell exterior	Filament cap	10	50

FliE	2	Periplasm	Rod MS-ring junction, export gate	~9	11
FliF	2	Cytoplasmic membrane	MS-ring, mount for rotor switch and rod, housing for export apparatus	26	61
FliG	2	Peripheral	C ring, rotor/switch protein, torque generation, binds MS ring and MotA	26	37
FliH	2	Cytoplasm	Negative regulator of FliI, Protein export, ATPase complex. Assosiates with FliM and FlhA. Assosiates with chaperones		26
FliI	2	Cytoplasm	Drives type III protein export. ATPase –unfolds late secretion substrates		49
FliJ	2	Cytoplasm	ATPase complex. Preference for rod and hook protein		17
FliK	2	Cell exterior	Hook-length control		39
FliL	2		Assosiates with MotAB complex, MS ring (FliF) and C ring (FliG) weakly		17
FliM	2	Peripheral	C ring, rotor/switch protein	37	38
FliN	2	Peripheral	C ring, Switch component	110	38

FliO	2	Centre of MS ring	Export component	~1	11
FliP	2	Centre of MS ring	Export component	~4	27
FliQ	2	Centre of MS ring	Export component	~1	10
FliR	2	Centre of MS ring	Export component	~1	29
FliS	3	Cytoplasm	FliC chaperone. Stabilises FlgM		15
FliT	3	Cytoplasm	FliD chaperone		14
FliY	2	Cytoplasm	Regulates class III transcription (?)		
FliZ	2	Cytoplasm	σ^{70} antagonist		
FlgA	2	Periplasm	Chaperone for P ring protein		
FlgB	2	Periplasm	Rod protein	7	
FlgC	2	Periplasm	Rod protein	6	
FlgD	2	Cell exterior	Hook capping protein	~5	

FlgE	2	Cell exterior	Hook protein	132	
FlgF	2	Periplasm	Rod protein	6	
FlgG	2	Periplasm	Distal rod protein	26	
FlgH	2	Outer membrane	L-ring protein	28	
FlgI	2	Periplasm	P-ring protein	24	
FlgJ	2	Periplasm	Rod capping protein, degrades PG layer	5?	
FlgK	3	Cell exterior	HAP1: hook filament junction protein	13	
FlgL	3	Cell exterior	HAP2: hook filament junction protein	~10	
FlgM	3	Cytoplasm	Anti - sigma factor 28		
FlgN	3	Cytoplasm	FlgK, FlgL specific chaperone. Transcriptional activator of FlgM		
MotA	3	Cytoplasmic membrane	Force generation, stator protein	~32	32
MotB	3	Cytoplasmic membrane	Force generation, stator protein, converts proton energy into torque	~16	34

Appendix 2: PCR primers used in the study:

SDM: site directed mutagenesis

Primer	Nucleotide Sequence	Function
proB KO FOR	GCTAAAACGTTCGTTTGATATCATTTTTCCTAAAATTGAATGGCAGAGAATC GTGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
proB KO REV	GTTGCGCTAATTTATACGAGGCTTGCTTCGCGGCAATGCCCATTTGTTCCAG CATATGAATATCCTCCTTA	Construction of gene knockout cassettes
proB KO CONFO	TGATTTTAATTAACGCGCGAATATTCAGCGGG	Conformation of gene knockout
proB US FOR	CGTTTGATATCATTTTTCCTAAAATTGAATGGCAGAGAATC	Construction of cassettes for SDM
D107N REV 1A	GCGTTCACGGTCTTCCATATTAGCACGGGTCAGCAGC	Construction of cassettes for SDM
D107N FOR 2A	GCTGCTGACCCGTGCTAATATGGAAGACCGTGAACGC	Construction of cassettes for SDM
E143A REV 1B	CGTTATCGCCGACCTTAATCGCTGCCGTAGCGACAGCATCG	Construction of cassettes for SDM
E143A FOR 2B	CGATGCTGTCGCTACGGCAGCGATTAAGGTCGGCGATAACG	Construction of cassettes for SDM

proB STOP FOR	GCCGTTCACCGTGATGACATGATTACCCGTTCTGCTTCGAAGTTCC	Construction of cassettes for SDM
proB STOP REV	GCCGTTCACCGTGATGACATGATTACCCGTTAAGTGTAGGCTGGAGCTGCTTCGAAGTTCC	Construction of cassettes for SDM
K1	CAGTCATAGCCGAATAGCCT	Conformation of gene knockout
K2	CGGTGCCCTGAATGAACTGC	Conformation of gene knockout
C1	TTATACGCAAGGCGACAAGG	Conformation of gene knockout
C2	GATCTTCCGTCACAGGTAGG	Conformation of gene knockout
dksA KO FOR	TTTTTCCCCCGAACATGGGGATCGATAGTGCGTGTTAAGGAGAAGCAACGTGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
dksA KO REV	CAAAAAATAAGGCGGGAGCATTTCCCGCCTGTGGTAAACGTGATGGAACGGCTGTAACATATGAATATCCTCCTTA	Construction of gene knockout cassettes
dksA KO CONFO	TTGTCAGAAGCTCAACAGAGGGGGGTAGAAATT	Conformation of gene knockout
lrhA KO FOR	TGTTTGTGTGTGCACAGCATTAACCAGCTCAGTATGAGCCGCCAGTAAGTGATAATATTGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
lrhA KO REV	AAACACCCCCAGCGGCTCGTTTTTTACACTATTGTCTCAGGAATTATCTATC	Construction of gene knockout cassettes

	GTCCGTCGATTCCGGGGATCCGTCGACC	
lrhA KO CONFO	GAGTGAGTAACTTAATAGATATGCGATCCCTGAA	Conformation of gene knockout
motAB KO FOR	GCCTGACGACTGAACATCCTGTCATGGTCAACAGTGGAAGGATGATGTCG TGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
motAB KO REV	TTCATCAAAAAATGTCTGATAAAAATCGCTTATATCCATGCTCACGCTGCAT ATGAATATCCTCCTTA	Construction of gene knockout cassettes
motAB KO CONFO	ATTTTATTACCCACGCTCACCAGCCTGTTGGC	Conformation of gene knockout
flgMN KO FOR	CGATAAATAAGCAACACATGATAAAAGCGCCCTCAATGAGGAATAAACCG TGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
flgMN KO REV	GACGGTGTAACAATGCATTCCGGCCTGCAGTGCAGGCCGGAGATAATCTC ATATGAATATCCTCCTTA	Construction of gene knockout cassettes
flgMN KO CONFO	TAACCCAGTTTCGCCAGGCATGGCGGGTAAAA	Conformation of gene knockout
flgDE KO F	AGCATGATGCTGAAAACCCTTACGCTCGGTCAATAAAGGAGAAAGCTGTG TAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
flgDE KO R	TCGCGGTATAAATTGCGTGATCCATTGAGCTATCCCGTCAGCGACATATGA ATATCCTCCTTA	Construction of gene knockout cassettes

flgDE KO CONFO	ATTCCAGGTTAACGCTGCACCAGGTGCTGCGACAGGCGG	Conformation of gene knockout
gdhA KO F	CAACATAAGCACAATCGTATTAATATATAAGGGTTTTATATCTGTGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
gdhA KO R	GCCCATTTGTAGGCCTGATAAGCGTAGCGCCATCAGGCATTTACAACATTCCGGGGATCCGTCGAC	Construction of gene knockout cassettes
gdhA KO CONFO	GCAATTAATGGAAATGCTAATACTACGGCGAACAATGC	Conformation of gene knockout
fliDST KO FOR	CGATAACCCCGGTATTCGTTTTACGTGTCGAAAGATAAAAGGAAATCGCGTGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
fliDST KO REV	GAAGCGTAGCCGTAATCGGATTATTCGCGAGCCATCGACTCATTCAGATATTCCGGGGATCCGTCGAC	Construction of gene knockout cassettes
fliDST KO CONFO	TGACTTGTGCCATGATTCGTTATCCTATATTG	Conformation of gene knockout
putA KO FOR	TACCAACGTGTTTCAACGTTGTAGTACCTATAAAGTGCTATTGCAATTCAACGTGGAAAGTGAACAACAGGAGTAATGGCGTGTAGGCTGGAGCTGCTTC	Construction of gene knockout cassettes
putA KO REV	ACAATTGACGGCGTCTCTTTTTCAGACTCAAT	Construction of gene knockout cassettes

putA KO CONFO	AAAACGTGCAGCGGATTATTATTGAGGGGATTCGACCACGTTAAAGTACG GCCTCCTCCAACATTGTAGGAGGCCGATGGACACATATGAATATCCTCCTT	Conformation of gene knockout
pJex ins F	CGAGCGTCTGTCTTCTGGCTTGCGTATTAACAG	Colony PCR and sequencing of inserts in pJexpress
pJex seq F	AACGGTTTCCCTCTAGAAATAATTTTGTTT	Colony PCR and sequencing of inserts in pJexpress
pJex seq R	TTCGCCCGGGCTAATTATGGGGTGTCGCCC	Colony PCR and sequencing of inserts in pJexpress
pJex seq alt F	ATTCCACAACGGTTTCCCTCTAGAAATAATTTTG	Colony PCR and sequencing of inserts in pJexpress
5’ and 47aa present (5’ UTR 47aa F)	CGGTACCATATGGCGGGAATAAGGGGCAGAG	Construction of secretion construct variants
No 5’, 47aa present (47aa F)	CGGTGACATATGAAAGAGGAGAAATAGTCCATGGCACAAGTCATTAATAC CAACAGCCTCTC	Construction of secretion construct variants
No 5’ or 47aa	TTAACTTCATATGCTCGAGAGGATAACGAATCATGGAGAATCTGTATTTTC AGGGCGAATCC	Construction of secretion construct variants
coli 26-47 F	CGGTACCATATGAAAGAGGAGAAATAGTCCATGTCGAGTTCTATCGAGCG TCTGTCTTCTGGCTTGCG	Construction of secretion construct variants
sal 26-47 F	CGGTACCATATGAAAGAGGAGAAATAGTCCATGGGAACGGCAATCGAGC GTCTGTCTTCTGGC	Construction of secretion construct variants

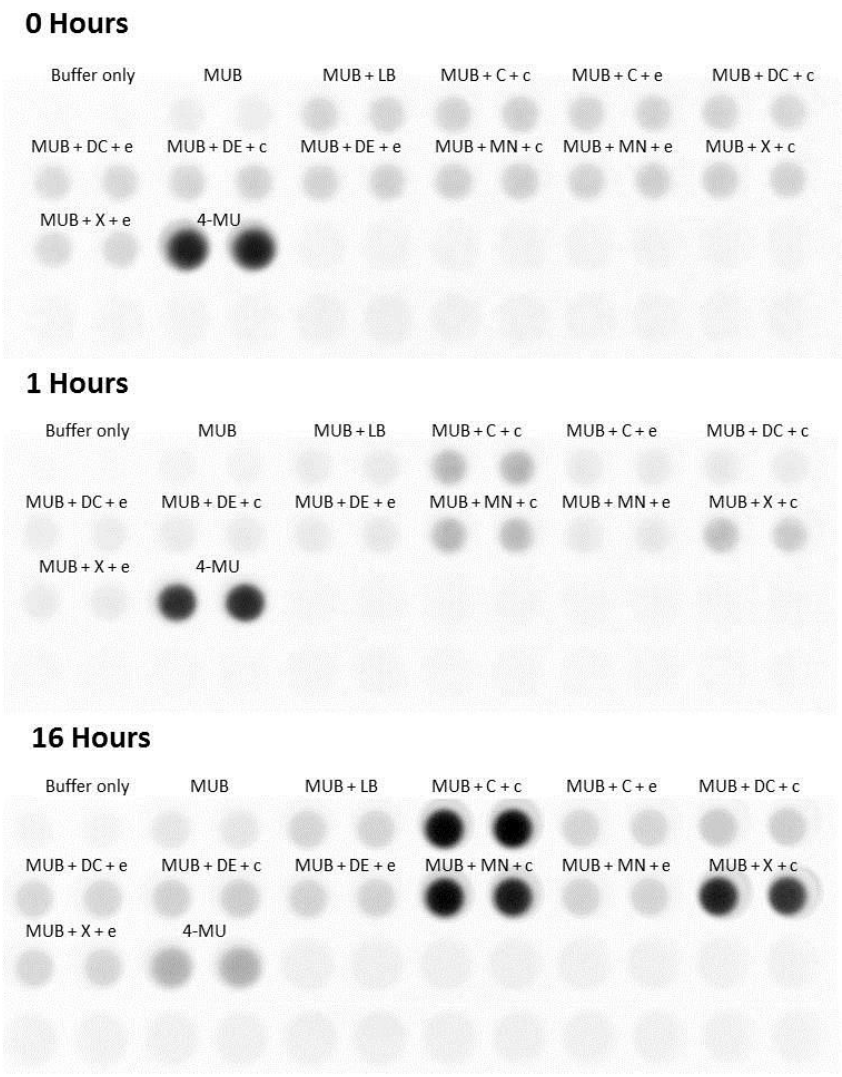
3’ UTR present (3’ UTR R)	TACAGAGGATCCCACGATAAACAGCCCTGCGTT	Construction of secretion construct variants
3’ UTR absent (no 3’)	CGCAGTTTGAGAAATCTAGAAAGCTTTAAGGATCCCGTTAG	Construction of secretion construct variants
coli 26-47 F	CGGTACCATATGAAAGAGGAGAAATAGTCCATGTCGAGTTCTATCGAGCG TCTGTCTTCTGGCTTGCG	Construction of secretion construct variants
sal 26-47 F	CGGTACCATATGAAAGAGGAGAAATAGTCCATGGGAACGGCAATCGAGC GTCTGTCTTCTGGC	Construction of secretion construct variants
IS5 F	CATCGAGCCGTTTTACCCCAAGGCTGGTAATGGC	PCR and sequencing conformation of the IS5 element
IS5 R	GAGATCTCTCCCACTGACGTATCATTTGGTCC	PCR and sequencing conformation of the IS5 element
flhD R	GGAATCTTGCGTCAACTGAGTAATCGTCTGGTGG	PCR and sequencing conformation of the IS5 element
IS5 top F	CCTGTTTCATTTTTGCTTGCTAGCGTAGCG	PCR and sequencing conformation of the IS5 element
IS5 half F	TACAAGGGGTTGCTGAAAAACGATAACCAACTGGCG	PCR and sequencing conformation of the IS5 element
flhD R shifted	CGGAAGTGACAAACCAGTTGATTGGTTTCTGCCAGC	PCR and sequencing conformation of the IS5 element
pJET F	CGACTCACTATAGGGAGAGCGGC	Colony PCR and sequencing of inserts in pJET2.1
pJET R	AAGAACATCGATTTTCCATGGCAG	Colony PCR and sequencing of inserts in pJET2.1

Appendix 3**:** Templates, primers and strains for gene deletions

Gene deletion(s)	**Template DNA**	**Forward primer**	**Reverse primer**	**Resistance cassette**	**Confirm deletion with k1/k2 c1/c2 and:**
proB	KD4	proB KO FOR	proB KO REV	Km	proB KO CONFO
dksA	KD4	dksA KO FOR	dksA KO REV	Km	dksA KO CONFO
lrhA	KD13	lrha KO FOR	lrhA KO REV	Km	lrhA KO CONFO
putA	KD13	putA KO FOR	putA KO REV	Km	putA KO CONFO
fliCD	KD3	fliCD KO FOR	fliCD KO REV	Cm	fliCD KO CONFO
flgMN	KD4	flgMN KO FOR	flgMN KO REV	Km	flgMN KO CONFO
motAB	KD3	motAB KO FOR	motAB KO REV	Cm	motAB KO CONFO
fliDST	KD13	fliDST KO FOR	fliDST KO REV	Km	fliDST KO CONFO
gdhA	KD13	gdhA KO FOR	gdhA KO REV	Km	gdhA KO OCNFO
flgDE	KD4	flgDE KO FOR	flgDE KO REV	Km	flgDE KO CONFO

Appendix 4 Templates and primers to implement *proB* gene substitutions by overlap PCR

PCR Reaction	**Template**	**Forward Primer**	**Reverse Primer**
1A	MC1000 ΔCKL chromosomal DNA	proB US FOR	D107N REV
1B	MC1000 ΔCKL chromosomal DNA	proB US FOR	E143A REV
2A	MC1000 ΔCKL chromosomal DNA	D107N FOR	proB STOP REV
2B	MC1000 ΔCKL chromosomal DNA	E143A FOR	proB STOP REV
3	KD3	proB STOP FOR	proB KO REV
4A	PCR products from reactions 1A, 2A, 3	proB US FOR	proB KO REV
4B	PCR products from reactions 1B, 2B, 3	proB US FOR	proB US REV



Appendix 5: Images of the fluorescence output of reaction mixtures following incubation of MUB with the secreted fraction of *CKL, ΔflhDC, ΔflgDE, ΔflgMN* or *ΔclpX* cells secreting empty or cutinase harbouring secretion construct.

E.coli ΔCKL (C)*, ΔflhDC* (DC)*, ΔflgDE* (DE)*, ΔflgMN* (MN) or *ΔclpX* (X) containing either the plasmid pJexpress-FliC-cutinase-Flag-Strep (c) or pJexpress-FliC-empty-Flag-Strep (e) were induced with 0.05mM IPTG (or 1mM for *ΔflhDC* and *ΔflgDE*) and grown in LB at 37°C with 180rpm agitation. Cells were harvested when cell culture reached OD_{600} 1.0. The supernatant was obtained by centrifugation. Either LB media or cell supernatant (40µL) was added to 160µL 100µM MUB solution (0.05M phosphate citrate buffer, pH 5) and added to a 96 well plate. Following incubation at 30°C fluorescence was measured capturing an image during excitation by UV light in a G:BOX. The image was inverted. One biological replicate, two technical replicates for each reaction mixture combination. Images captured following 0, 1 and 16 Hours incubation