DNA repair during bacterial competence? A potential role for the periplasmic ligase, Lig E.

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Science (Research) in Molecular and Cellular Biology

at

The University of Waikato

by

JOLYN TYN TYN PAN

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2021

Abstract

The rapid rise in antibiotic resistance among pathogenic bacteria calls for better understanding of drivers of resistance and pathways by which it occurs. Identification of such mechanisms could identify novel pathways that can be targeted for therapeutic interventions. In particular, disruption of the DNA repair system, which includes ligases that seal breaks in DNA, would greatly discourage bacterial growth. Recently, an enigmatic ATP-dependent DNA ligase, Lig E, has been discovered in many species of antibiotic-resistant Gram-negative bacteria. This enzyme is not only minimal and compact, but also contains an N (amino)-terminal signal peptide that indicates a likely periplasmic location, as opposed to the cytoplasm, the location of genomic DNA. It is thus hypothesised that Lig E has a potential role in bacterial competence and aids in repairing breaks in damaged DNA obtained from the environment, a role which may enhance the acquisition of antibiotic resistance genes under DNA damaging conditions.

In recent years, there has been growing public concern about the human pathogen *Neisseria gonorrhoeae*, which causes the sexually transmitted disease, gonorrhoea. Previously easily treatable, this disease is now spreading at alarming rates due to the bacterium's rapid acquisition of antibiotic resistance genes, attributable to its ability to take up pieces of DNA without regulation. It is in relation to this natural competence that Lig E is hypothesised to function, not only by accelerating the uptake process, but also by increasing the uptake efficiency. Thus, the aim of this thesis was to identify the cellular location and biological function of Lig E in *N. gonorrhoeae* (Ngo-Lig) by generating *in vivo* mutants, the effects of which were characterised both *in vivo* (growth experiment) and *in vitro* (ligation assays). Steps were also made towards developing an *in vivo* assay to test the role of Ngo-Lig in DNA uptake by optimisation of uptake reporter constructs.

Results obtained over the course of this thesis showed that disruption of *Lig E* in *N. gonorrhoeae* leads to a decrease in the rates of gonococcal growth. From this, the hypothesis of the role of Lig E evolved to also consider its potential role in biofilm formation, which is important for the bacterium's attachment and infection.

Although further research into the biological role of Lig E are necessary, the results collected demonstrated potentially novel pathways involving Lig E that may be targeted by future drug developments to tackle this emerging threat in our community.

Acknowledgments

There are many to whom I owe the utmost gratitude for the completion of my thesis. First of all, to my primary supervisor, Dr. Adele Williamson. Thank you for basically holding my hand the entire journey and for always being ready to offer answers and solutions. This project has meant the world to me and it would not have been possible without you. Thank you also to my co-supervisor, Dr. Joanna Hicks, for helping with all things gonorrhoea related – I can now proudly tell everyone that I have dedicated 1.5 years of my life trying to make my gonorrhoea flourish, although with much difficulty. To Annmaree, Liz and Keely, thank you for helping me get started in the lab and for putting up with my constant questions. I would never have been able to move forward without the three of you. Thank you also to Dr. Vic Arcus, for introducing me and guiding me towards the direction of this project.

To everyone in the entirety of the C2 labs, and especially the Proteins and Microbes lab, thank you for your companionship and kindness, and for accepting my nap-prone presence in the office. I could not have wished for a better family to start my research journey with, despite all the yummy baking-induced weight gain I have obtained. This section would also be grossly incomplete without acknowledging the incredible Dr. Judith Burrows. Judith, thank you for always being so patient with my constant requests and for basically taking care of the entire lab. I can never imagine how the lab would function without you. To the ligase team; Adele, Liz, Ash and Bronwyn, I could not have had better lab bench mates. This sticky enzyme will forever be my number one enzyme.

To Jenn and Robin, thank you for keeping me sane but also driving me crazy by constantly asking how my gonorrhoea is going, out loud, in public. Thank you for putting up with my rants and for distracting me with TV binges and Art Nights. To 4pak, thank you for always checking in whenever I was on the brink of giving up. To mum, and to the craziest siblings I could ask for, JJ, Kor and Gab, thank you for your constant support and for always believing in me when I most definitely

wouldn't. And lastly, to dad, thank you for always believing in my education. You would have been so incredibly proud. In fact, this one is for you.

Table of Contents

List of Figures

List of Tables

List of Equations

Abbreviations

(TEV)	Tobacco Etch Virus protease cleavage site; E-N-L-Y-F-Q-|-Gly/S
3'OH	3'-hydroxy
5'PO$_4$	5'-phosphate
A	Adenine
A230	Absorbance at 230 nm
A260	Absorbance at 260 nm
A280	Absorbance at 280 nm
AD domain	Adenylation domain
ADLs	ATP-dependent ligases
Ame-Lig	Lig E from *Alteromonas mediterranea*
AMP	Adenine monophosphate
APS	Ammonium persulfate
ATP	Adenine triphosphate
attB	Attachment site for bacteria
attL	Left attachment site
attR	Right attachment site
b-ADLs	Bacterial ATP-dependent ligases
b-Lig E	Bacterial Lig E
bp	Base pair
BRCT domain	Breast cancer carboxy terminal domain
BSA	Bovine serum albumin
C	Cytosine
Cal	*Candida albicans*
CDS	Coding DNA sequence
CFU	Colony forming units
ChV	*Chlorella* virus
ChV-Lig	*Chlorella* virus ligase
CO$_2$	Carbon dioxide
C-terminus/C-terminal	Carboxy-terminus/Carboxy-terminal
D	Aspartic acid

DMEM/F12	Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12
DMSO	Dimethylsulfoxide
DNA	Deoxyribose nucleic acid
DSBs	Double-stranded breaks
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
DUS	DNA uptake sequence
E	Glutamic acid
EDTA	Ethylenediaminetetraacetic
Efa	*Enterococcus faecalis*
ESCs	Extended-spectrum cephalosporins
exDNA	Extracellular DNA
F	Phenylalanine
FBS	Foetal bovine serum
G	Guanine
Gly	Glycine
GC	Gonococcal
GCB medium	Gonococcal base medium
GFP	Green fluorescent protein
GGI	Gonococcal genetic island
GTP	Guanosine triphosphate
H_2O	Water
HGT	Horizontal gene transfer
HhH motif	Helix-hairpin-helix motif
Hin-Lig	Lig E from *Haemophilus influenzae*
His	Histidine
His-(TEV)-Lig E	Lig E with an N-terminal hexa-His tag, followed by a TEV protease cleavage site
His-(TEV)-Lig E-GFP	C-terminally GFP-tagged Lig E with an N-terminal hexa-His tag followed by a TEV protease cleavage site
His-(TEV)-Lig E-His	C-terminally His-tagged Lig E with an N-terminal hexa-His tag followed by a TEV protease cleavage site

hr	Hour
I	Isoleucine
IDT^{TM}	Integrated DNA Technologies^{TM}
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Lysine
K_2HPO_4	Dipotassium phosphate
KanaR	Kanamycin resistance gene
kb	Kilobase
KCl	Potassium chloride
kDa	KiloDalton
KH_2PO_4	Monopotassium phosphate
KO	Knock-out
L	Leucine
LB	Lysogeny broth
Lig E-GFP	C-terminally GFP-tagged Lig E
Lig E-His	C-terminally His-tagged Lig E
Lig E-His-KanaR	*Lig E* mutant with a hexahistidine tag and kanamycin resistance gene on the C-terminus
Lig E-KO-KanaR	*Lig E* mutant with a kanamycin resistance gene disrupting the gene
mA	Milliampere
MBP	Maltose binding protein
MDR	Multi-drug resistance
MES	2-(*N*-morpholino)ethanesulfonic acid
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
MIC	Minimum inhibitory concentration
min	Minute
mRNA	Messenger ribonucleic acid
Mtu-Lig	ATP-dependent ligase from *Mycobacterium tuberculosis*
N	Asparagine

NAATs	Nucleic acid amplification tests
NaCl	Sodium chloride
NDLs	β-NAD$^+$-dependent ligases
NEB	New England Biolabs
NGC^{TM}	Next-Generation Chromatography^{TM}
Ngo-Lig	Lig E from *Neisseria gonorrhoeae*
NHEJ	Non-homologous end joining
Nme-Lig	Lig E from *Neisseria meningitidis*
Ntase domain	Nucleotidyltransferase domain
N-terminus/N-terminal	Amino-terminus/Amino-terminal
OB domain	Oligonucleotide-binding domain
OD	Optical density
OD$_{600}$	Optical density at 600 nm
OMP	Outer membrane protein
PCR	Polymerase chain reaction
PE domain	Phosphoesterase domain
pI	Isoelectronic point
PMN	Polymorphonuclear leukocytes
PPI	Pyrophosphate
PPP	Precision Plus Protein^{TM} All Blue Prestained Protein standard
PrimPol domain	DNA primase/polymerase domain
Psy-lig	Lig E from *Psychromonas sp.*
Q	Glutamine
RNA	Ribonucleic acid
Rnl2	T4 RNA ligase 2
rpm	Revolutions per minute
S	Serine
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
Sec	Secretory
secs	Seconds
Soc media	Super optimal broth with catabolite repression media

ssDNA	Single-stranded DNA
STI	Sexually transmitted infection
T	Thymine
T4SS	Type IV secretion system
T7	T7 bacteriophage
T_a	Annealing temperature
TAE	Tris base, acetic acid and EDTA
Tat	Twin-arginine translocation
Tb media	Terrific broth media
TBE	Tris-borate-EDTA
TBS-T	Tris buffered saline-Tween 20
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
TEV	Tobacco Etch Virus protease
Tfi	*Tiedemannia filiformis*
Tfp	Type IV pili
TG-SDS	Tris-glycine-SDS
T_m	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
TSB	Tryptic soy broth
USS	Uptake signal sequence
UV	Ultraviolet
V	Volts
w/v	Weight per volume
WHO	World Health Organisation
WT	Wild-type
XDR	eXtensive Drug Resistance
Y	Tyrosine
β-lac	*β-lactamase* gene
$β-NAD^+$	β-nicotinamide adenine dinucleotide
β-NMN	β-nicotinamide mononucleotide

Chapter One

Introduction

1.1 *Neisseria gonorrhoeae*

Neisseria gonorrhoeae is an obligate pathogen, responsible for the second most common sexually transmitted infection (STI) in the world, gonorrhoea (Lee *et al.*, 2017). With more than 78 million reported infections globally in 2012, an exponential increase from the 30,000 cases reported in 1970, it was, and still is, continually spreading (Tapsall, 2005; Unemo & Shafer, 2011; Cehovin & Lewis, 2017; Lee *et al.*, 2017). In addition, rates are much higher in less developed nations and among individuals of lower socioeconomic groups, for which numbers are often underreported. In fact, the World Health Organisation (WHO) estimates that there are 106 million new cases of gonorrhoea each year (Unemo & Shafer, 2014).

N. gonorrhoea is a Gram-negative proteobacterium with a diplococcal morphology, and is one of two pathogenic species in the genera *Neisseria*, the other being *Neisseria meningitidis* (Unemo & Shafer, 2014; Cehovin & Lewis, 2017). Oftentimes, a diagnosis begins with either identification of the bacterium in polymorphonuclear (PMN) leukocytes from patient urine samples, or via nucleic acid amplification tests (NAATs) (Unemo & Shafer, 2014). In order to live in the relatively hostile human environment where the threat of oxidative stress and neutrophil consumption is high, *N. gonorrhoeae* displays defence mechanisms such as antigenic and phase variations to avoid the immune system by varying the lengths of its hetero- and homopolymeric tracts of tandem repeats (Srikhanta *et al.*, 2010; Unemo & Shafer, 2011). This allows it to alter its gene expression to increase its population diversity. Specifically, *N. gonorrhoeae* resides in the human urogenital epithelia, leading to inflammation that presents as urethritis in men and cervicitis in women (Edwards & Apicella, 2004; Tapsall, 2005; Marri *et al.*, 2010; Dillard, 2011; Cehovin & Lewis, 2017). As a result, patients often observe pain with urination due to urethral inflammation (urethritis), or abnormal bleeding and pain outside of their menstruation period, associated with cervical inflammation (cervicitis). Exacerbating this, infections of the lower genital tract in women may be

asymptomatic (≥50% in women in contrast to ≤10% in men), allowing ascension of the infection onto the upper genetic tract (the uterus, fallopian tubes and ovaries) (Unemo & Shafer, 2014). This is also associated with the bacterium's ability to form biofilms, which aids with cell attachment and leads to increased overall resistance of the bacterial community (Greiner *et al.*, 2005; Steichen *et al.*, 2008). These asymptomatic infections often result in chronic pelvic inflammatory disease present as persistent pain and may cause ectopic pregnancy and infertility due to scarring and blockage of the fallopian tubes (Edwards & Apicella, 2004; Lee *et al.*, 2017). Untreated gonorrhoea in pregnant women may also be transferred to the baby, leading to neonatal conjunctivitis and blindness in the new-born (Dillard, 2011). Apart from the conjunctiva and urogenital epithelia, *N. gonorrhoeae* also infects the pharynx and rectal mucosa of both sexes (Edwards & Apicella, 2004; Tapsall, 2005; Unemo & Shafer, 2011). These infections may be asymptomatic and more difficult to treat, exacerbating existing infections through changes in the genital epithelial cells, including increasing the likelihood of human immunodeficiency virus transmissions by five times compared to a healthy individual (Edwards & Apicella, 2004; Tapsall, 2005; Unemo & Shafer, 2011).

Alongside the steady increase in incidence rate is the accompanying increase in antibiotic resistance, making gonorrhoea difficult to treat and creating a major public health crisis. The move away from penicillin, the first line of defence used against *N. gonorrhoeae* in 1943 (after the brief use of sulphonamides), was caused by the complete bacterial resistance developed after 10-15 years. This initially led to increased dosages used (due to the increase in minimum inhibitory concentrations (MICs)), before their complete inefficiencies after a global spread of the antibiotic resistance gene, *β-lactamase* (Robson & Salit, 1972; Unemo & Shafer, 2011). Soon after, a shift towards the use of spectinomycin and tetracycline followed. However, complete resistance to these was also inevitable after the spread of the *tetM* gene that confers resistance to tetracycline (Unemo & Shafer, 2014). Despite the development and current use of third-generation extended-spectrum cephalosporins (ESCs) with improved activities, like cefixime (oral) and ceftriaxone (injection), antibiotic resistance in *N. gonorrhoeae* is still rapidly evolving (Unemo & Shafer, 2011, 2014). Rather than relying on the historical bacterial classification of multidrug resistance (MDR) (resistance to one category I

drug (either ESCs or spectinomycin), along with two category II drugs (penicillin, fluoroquinolone, azithromycin, aminoglycoside and carbapenem)), it has been necessary in recent years to include another eXtensive Drug Resistance (XDR) category for *N. gonorrhoeae*, which includes resistance to at least two category I and three category II drugs (Lee *et al.*, 2017).

In fact, a 2014-2015 *N. gonorrhoeae* population study in New Zealand demonstrated non-susceptibility of *N. gonorrhoeae* isolates to penicillin (98%; 85.7% intermediate susceptibility, 12.3% fully resistant), ciprofloxacin (32% fully resistant) and tetracycline (68%; 41.7% intermediate susceptibility, 26.6% fully resistant) (Tapsall, 2005; Lee *et al.*, 2017). However, the most concerning results were related to the decreasing susceptibility of isolates towards currently used ESCs, with nearly 4% showing increasing MIC (Lee *et al.*, 2017). Seven strains displayed decreased susceptibility to category I drugs, along with resistance to at least two category II drugs, of which the authors described as 'pre-MDR' (Lee *et al.*, 2017). Although no outright resistance has been detected in New Zealand, one can only predict that the same trends would be observed with the currently used ESCs. For example, in 1988, the level of *N. gonorrhoeae* resistance to penicillin and tetracycline sat at only 0.2% and 3.7% respectively as opposed to the current levels (Brett *et al.*, 1992). Ceftriaxone-resistant gonococcal strains have now been identified in both Japan and Argentina, which raised alarms of a possible spread of this resistant phenotype via human contact (Ohnishi *et al.*, 2011; Unemo & Shafer, 2014; Gianecini *et al.*, 2016). Taking into account the removal of cefixime as a treatment option in some countries, the only available option for many cases is a dual-antibiotic strategy using ceftriaxone and either azithromycin or doxycycline, which may not be applicable in less developed areas where treatments are expensive (Unemo & Shafer, 2014). Thus, more robust treatments against *N. gonorrhoea* must be designed before the development of a completely untreatable superbug and the resultant public health crisis, especially among people with limited access to affordable healthcare.

1.1.1 Natural transformation and competence

1.1.1.1 Bacterial competence

The rapid acquisition of antibiotic resistance genes among *N. gonorrhoeae* can be attributed to its competence. Competence refers to an organism's ability to take up extracellular genetic material or DNA (deoxyribose nucleic acid) and incorporate it into its own genome (Sparling, 1966; Seitz & Blokesch, 2013; Veening & Blokesch, 2017). If occurring naturally for the bacterium, this is termed natural transformation. Compared to the other two forms of DNA exchange via horizontal gene transfer (HGT) (i.e. transduction and conjugation), natural transformation does not require other genetic elements, but rather, relies on the competent state of the acceptor organism (Seitz & Blokesch, 2013; Seitz *et al.*, 2014; Veening & Blokesch, 2017). This is particularly common in many human pathogens like *Haemophilus influenzae* and *Vibrio cholerae* (Hamilton & Dillard, 2006).

There are currently three theories for the role of competence in bacteria. These include the 'DNA for food', 'DNA for genome repair' and 'DNA for evolution' theories. The DNA for food theory stems from the high energy requirement of *de novo* nucleotide synthesis, forcing bacteria to turn to extracellular DNA (exDNA) for nutrition to save energy and resources (Redfield, 1993, 2001). This theory is supported by the frequent uptake of DNA from dead bacteria during natural transformations, which due to the poor quality of DNA and high accumulation of mutations, would not be suitable for either genome repair or evolution, especially considering it was selected against (i.e. indicating a non-adaptive genotype) (Seitz & Blokesch, 2013; Veening & Blokesch, 2017). Furthermore, induction of competence in pathogens like *H. influenzae* normally occurs during their stationary phases where nutrition is low, offering the bacteria a short-term advantage and means of survival during this stage (Redfield, 1993; Seitz & Blokesch, 2013).

The DNA for genome repair theory is based on evidence of exDNA integration into the genome. It is further supported by the presence of specific sequences found in some bacteria like the *Neisseria spp.*, which encourages uptake of DNA from only closely related organisms as such closely related DNA is most suitable for use in homologous recombination-based repair processes (Seitz & Blokesch, 2013).

4

Furthermore, once exposed to DNA damaging agents, bacteria like *Streptococcus pneumoniae* observe an increase in transformation rates through fratricide, which encourages uptake of only species-specific DNA (Gilmore & Haas, 2005; Engelmoer & Rozen, 2011; Seitz & Blokesch, 2013). In fact, both the process of autolysis during the stationary phase in *N. gonorrhoeae* and its Type IV secretion system (T4SS) (encoded on a gonococcal genetic island (GGI) and allows single-stranded DNA (ssDNA) and proteins to be secreted) also provide evidence for the DNA for genome repair theory through release of DNA from neighbouring gonococcal cells (Hamilton *et al.*, 2005; Hamilton & Dillard, 2006; Seitz & Blokesch, 2013).

The last DNA for evolution theory encompasses the act of increasing genetic information to increase fitness and survival of the species, which may explain the rapid rise in antibiotic resistance and mosaicism observed in many bacterial genome (Veening & Blokesch, 2017). Based on this theory, stress-related events are an important trigger for competence, allowing the bacterium to take up genes that increase its chances of survival (Engelmoer & Rozen, 2011; Seitz & Blokesch, 2013). For example, the observation of MDR strains in *Acinetobacter baylyi* straight after DNA uptake strengthens this hypothesis (Perron *et al.*, 2012).

1.1.1.1.1 DNA uptake machinery of Gram-negative bacteria

Regardless of the underlying driver for DNA uptake during competence, there are variations between the different DNA uptake machineries found in bacteria. In particular, differences exist between Gram-positive bacteria that contain a single cell membrane and Gram-negative bacteria which have two membranes (outer and inner) and a peptidoglycan layer that have to be crossed before homologous recombination can occur (Matthey & Blokesch, 2016). To do so, most Gram-negative bacteria express the dynamic Type IV pili (Tfp) fibrous system that is important for transformations (Adams *et al.*, 2019). This system comprises a major protein, PilA, which allows for species-specific recognition (Adams *et al.*, 2019). After being processed in the inner membrane, the PilA protein is assembled and polymerised with other proteins (PilB, C, M, N, O, P, Q, F, T), allowing it to fit into the outer membrane secretin pore (PilQ) where it can bind extracellular double-

stranded DNA (dsDNA) (Matthey & Blokesch, 2016). There are two ATPases that are important for the continual extraction and retraction of the Tfp for DNA uptake. These include PilB (allowing for extraction past the outer membrane) and PilT (allowing for retraction), the latter of which is crucial for DNA uptake (Matthey & Blokesch, 2016; Adams *et al.*, 2019). In the absence of PilT, bacteria often form aggregates to increase the likelihood of Tfp utilisation (Adams *et al.*, 2019).

In concert with the Tfp system is the periplasmic ComEA DNA binding protein, important for stabilising and pulling the DNA through, allowing the DNA to travel past the inner membrane via the ComEC translocator with help from ComF (Provvedi & Dubnau, 1999; Seitz *et al.*, 2014; Matthey & Blokesch, 2016; Adams *et al.*, 2019). After restriction and degradation of one strand, the ssDNA is protected from degradation in the cytoplasm via Ssb, RecA and DprA proteins before recombination into the genome (Koomey & Falkow, 1987; Chaussee & Hill, 1998; Stohl & Seifert, 2001; Hamilton & Dillard, 2006; Hepp *et al.*, 2016; Matthey & Blokesch, 2016). This system is highly conserved among Gram-negative bacteria and through facilitating DNA uptake, underpins the diversity observed among bacteria.

1.1.1.2   Natural transformation in *N. gonorrhoeae*

The panmictic and diverse nature of *N. gonorrhoeae* is associated with its wealth of both mobile genetic elements that allow for intrachromosomal rearrangements, and the GGI (present in 80% of gonococcal strains) that allow for interchromosomal gene transfer. However, variations in its genome can also be attributed to its ability to undergo natural transformation at any stage of its life, with the only requirement being piliation (see section 1.1.1.1.1) (Sparling, 1966; Chen & Gotschlich, 2001; Hamilton & Dillard, 2006; Ramsey *et al.*, 2012; Seitz & Blokesch, 2013; Cehovin & Lewis, 2017). Perhaps the most interesting aspect about its competence is its unregulated nature, which is different from many other naturally competent bacteria like *H. influenzae* or *S. pneumoniae* (Hamilton & Dillard, 2006). This, in addition to its resultant antigenic and phase variation, may be accountable for the rapid rise in antibiotic resistance, making its competence an ideal target against the pathogen (Dillard, 2011; Unemo & Shafer, 2011).

The strong selection pressure present in the *N. gonorrhoeae* habitat such as the presence of macrophages and PMN leukocytes, creates a need for increased uptake of exDNA for optimised fitness, adaptation and virulence, leading to the rise of MDR in the gonococcal population (Marri *et al.*, 2010; Seitz & Blokesch, 2013; Unemo & Shafer, 2014). In particular, *N. gonorrhoeae* coexists with other commensal *Neisseria* species in the pharynx, the latter of which may be frequently subjected to antimicrobials. This condition selects for resistant bacteria and may be followed by subsequent gene transfer onto *N. gonorrhoeae* (Catlin, 1973; Unemo & Shafer, 2014). In fact, the mutated *penA* gene in ESC-resistant *N. gonorrhoeae* is a mosaic gene obtained after recombination with partial *penA* genes from commensal *Neisseria spp.* (Spratt *et al.*, 1992; Unemo & Shafer, 2014). Most interestingly, in 2011, HGT was detected in *N. gonorrhoeae* with the human long interspersed nuclear element L1 gene, showing evidence of potential gene transfer with other types of organisms as well (Anderson & Seifert, 2011).

1.1.1.2.1 DNA uptake in N. gonorrhoeae

There are three main steps to DNA uptake in *N. gonorrhoeae*; 1. donation, 2. binding and uptake, and 3. processing and homologous recombination (Figure 1.1). Specifically, *N. gonorrhoeae* takes up DNA (secreted via the T4SS machinery or autolysis) with a precise 10 bp DNA uptake sequence (DUS) (5'-GCCGTCTGAA-3'), which is found frequently throughout its genome (approximately every 1096 bp, or in total 2000 times in the FA 1090 strain) (Goodman & Scocca, 1988; Hamilton & Dillard, 2006; Spencer-Smith *et al.*, 2016). These sequences are often found as inverted repeats between a specific gene, where it aids with transcription termination through the formation of hairpin structures (Goodman & Scocca, 1988; Hamilton & Dillard, 2006; Ambur *et al.*, 2007; Spencer-Smith *et al.*, 2016). A majority of these DUS sequences (76%) were found to be extended, with an additional AT dinucleotide at the 5' end that increases efficiency of DNA uptake by three-fold (Ambur *et al.*, 2007). The presence of the same DUS sequence among other *Neisseria spp.* also allows uptake of DNA from other *Neisseria spp.*, as evidenced by the development of the aforementioned *penA* gene (Spratt *et al.*, 1992; Hamilton & Dillard, 2006). This system is not limited to *N. gonorrhoeae*, and is

present in *H. influenzae* as well, which relies on its uptake signal sequence (USS), 5'-AAGTGCGGT-3', to discriminate between genus-specific and foreign DNA for incorporation into its genome (Goodman & Scocca, 1988).
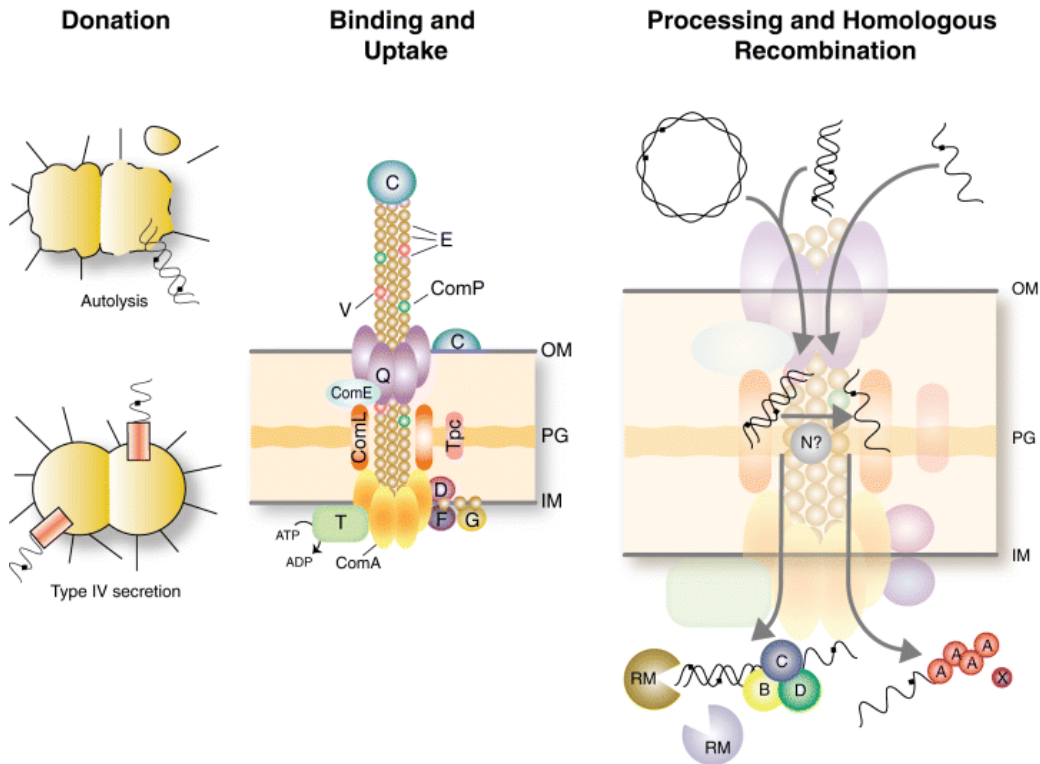
Figure 1.1. The DNA uptake process in *N. gonorrhoeae* consists of three major steps; Donation of DNA (with DNA uptake sequences) through autolysis or the Type IV secretion system; Binding of the DNA onto the Type IV pili system (with Pil proteins E, V, Q, C, T, D, F, G) and uptake of the DNA into the periplasmic/peptidoglycan layer via the periplasmic Com E protein; Processing of the piece of DNA into a single-strand and the subsequent homologous recombination into the genome via a RecA-dependent pathway. Figure from (Hamilton & Dillard, 2006). OM – outer membrane, IM – inner membrane, PG – peptidoglycan.

After binding of non-specific DNA onto the surface of cells, the surface protein, ComP, which recognises the DUS sequence, initiates the second step in gonococcal DNA uptake (Hamilton & Dillard, 2006). This allows for specific uptake of DNA into the periplasm via the Tfp machinery. The ComE protein (a ComEA homologue) binds and pulls the DNA into the periplasm before it is processed (third step), which includes degradation of one strand (Chen & Gotschlich, 2001; Hamilton & Dillard, 2006; Hepp *et al.*, 2016). After travelling past the inner membrane into the cytoplasm, homologous recombination can occur. The processing of DNA in the periplasm is an enigmatic step in DNA uptake as much has yet to be uncovered of

this process. Thus, the function of Lig E, a DNA ligase that contains a putative periplasmic signal peptide, is of much interest, especially in regards to its possible role in bacterial competence.

1.2 DNA ligases

From their discovery in 1967 by the Gellet, Lehman, Richardson and Hurwitz laboratories, the role of DNA ligases in sealing phosphodiester breaks between the 3'-OH (3'-hydroxy) and 5'-PO$_4$ (5'-phosphate) ends of DNA to maintain genomic integrity during repair, replication and recombination has now been solidly established (Zimmerman *et al.*, 1967; Nair *et al.*, 2007; Shuman, 2009; Pergolizzi *et al.*, 2016). As part of the nucleotidyl transferase superfamily, DNA ligases have high structural similarities with other GTP (guanidine triphosphate)-dependent mRNA (messenger ribonucleic acid) capping enzymes and ATP (adenine triphosphate)-dependent RNA (ribonucleic acid) ligases that also fall within this group. These enzymes consist of two core domains; a larger catalytic adenylation (AD) domain (sometimes called the nucleotidyltransferase (NTase) domain) and a smaller oligonucleotide-binding (OB) domain. These are joined together by a flexible linker that facilitates transitions between the open and closed conformations of these enzymes (Figure 1.2) (Subramanya *et al.*, 1996; Doherty & Wigley, 1999; Shuman & Lima, 2004; Nair *et al.*, 2007; Williamson *et al.*, 2016).
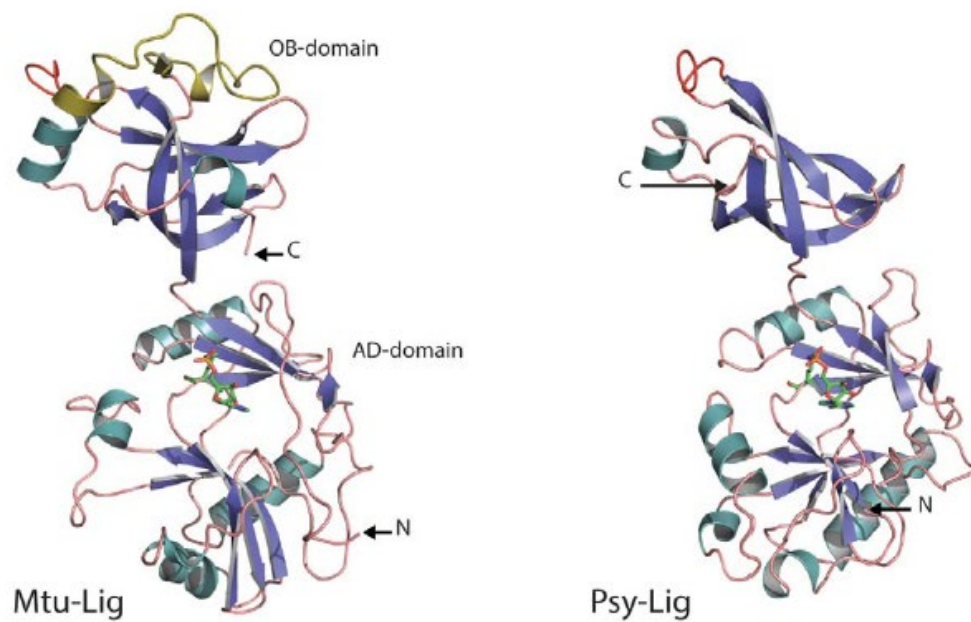
Figure 1.2. Structures of a *Mycobacterium tuberculosis* ATP-dependent ligase (ADL) (Mtu-Lig) and the *Psychromonas* sp. ADL (Psy-Lig) showing the oligo-binding (OB)-domain at the top, and the adenylation (AD)-domain at the bottom. Figure from (Williamson *et al.*, 2016), with the AMP cofactor covalently bound in the active site of the AD-domain. Arrows indicate the respective N (amino)- and C (carboxy)-termini.

The AD domain binds an adenylate donor, the action of which activates the ligase. Within this domain, there are five conserved motifs also present in other nucleotidyl transferase enzymes (these include the K(lysine)-E(glutamic acid)-Y(tyrosine)/F(phenylalanine)-D(aspartic acid)/E-K amino acids and their surrounding environments) (Figure 1.3), which play a role in cofactor binding and catalysis (Subramanya *et al.*, 1996; Doherty & Suh, 2000; Wilkinson *et al.*, 2001; Shuman & Lima, 2004). In contrast, the OB domain functions as an anchor, engaging and positioning the substrate DNA during catalysis (Doherty & Wigley, 1999; Doherty & Suh, 2000; Magnet & Blanchard, 2004). In addition to these two core domains, additional domains attached to the N (amino)- and C (carboxy)-termini may also be present for increased specificity and complete encirclement of the DNA substrate.
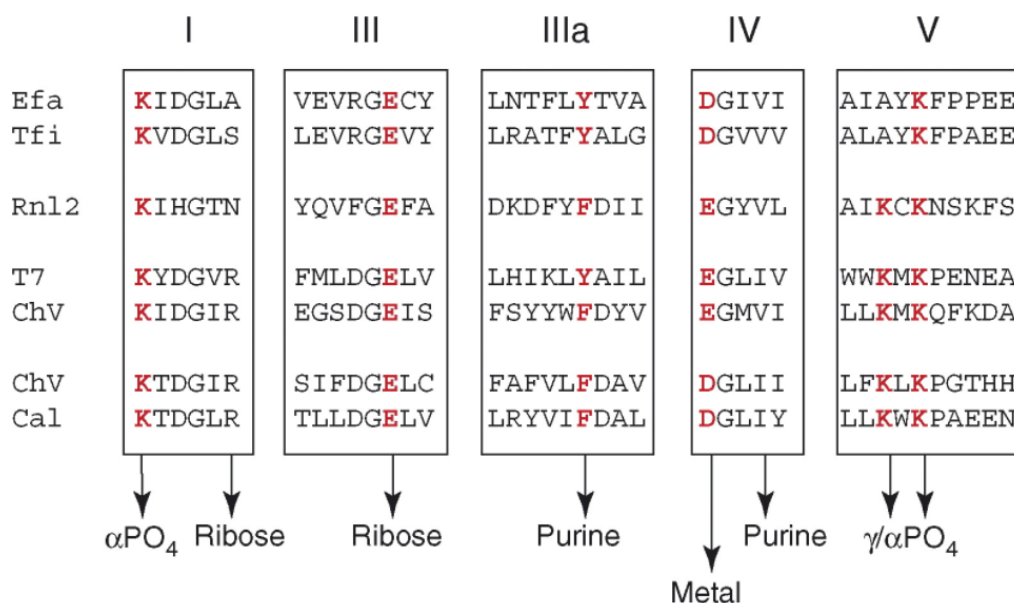
Figure 1.3. Conserved AD domain motifs among nucleotidyl transferase enzymes including NAD⁺-dependent ligases from *Enterococcus faecalis* (Efa) and *Tiedemannia filiformis* (Tfi), ATP-dependent ligases like the T4 RNA ligase 2 (Rnl2), and those from the T7 bacteriophage (T7) and *Chlorella* virus (ChV), and GTP-dependent capping enzymes from ChV and *Candida albicans* (Cal). Figure from (Shuman & Lima, 2004), with contacts to the nucleotide cofactor indicated below the motifs. The conserved amino acids (lysine (K)-glutamic acid (E)-tyrosine (Y)-aspartic acid (D)-K) are shown in red.

The operation of DNA ligases is often described in three 'ping-pong'-like steps (Figure 1.4) (Wilkinson *et al.*, 2001; Shuman & Lima, 2004). In the first step, the adenylate donor binds to the active site of the ligase before nucleophilic attack by a conserved lysine residue onto the α-phosphate of the donor, resulting in a high energy adenylated intermediate with a phosphoamide bond between the lysine and the resultant AMP (adenine monophosphate) molecule (step A in Figure 1.4) (Shuman, 2004; Shuman & Lima, 2004; Tomkinson *et al.*, 2006; Pergolizzi *et al.*, 2016). In the second step, the newly adenylated ligase is subjected to nucleophilic attack by the 5'-PO₄ group of the nicked DNA substrate to form a DNA-adenylate intermediate (step B in Figure 1.4), before a final nucleophilic attack of the substrate 3'-OH group onto the activated 5-PO₄ end (step C in Figure 1.4), which forms the final covalent phosphodiester bond between both ends of the nick, sealing the once fragmented DNA, with release of AMP (Shuman, 2004; Shuman & Lima, 2004; Tomkinson *et al.*, 2006; Pergolizzi *et al.*, 2016; Williamson & Leiros, 2019; Williamson & Leiros, 2020).
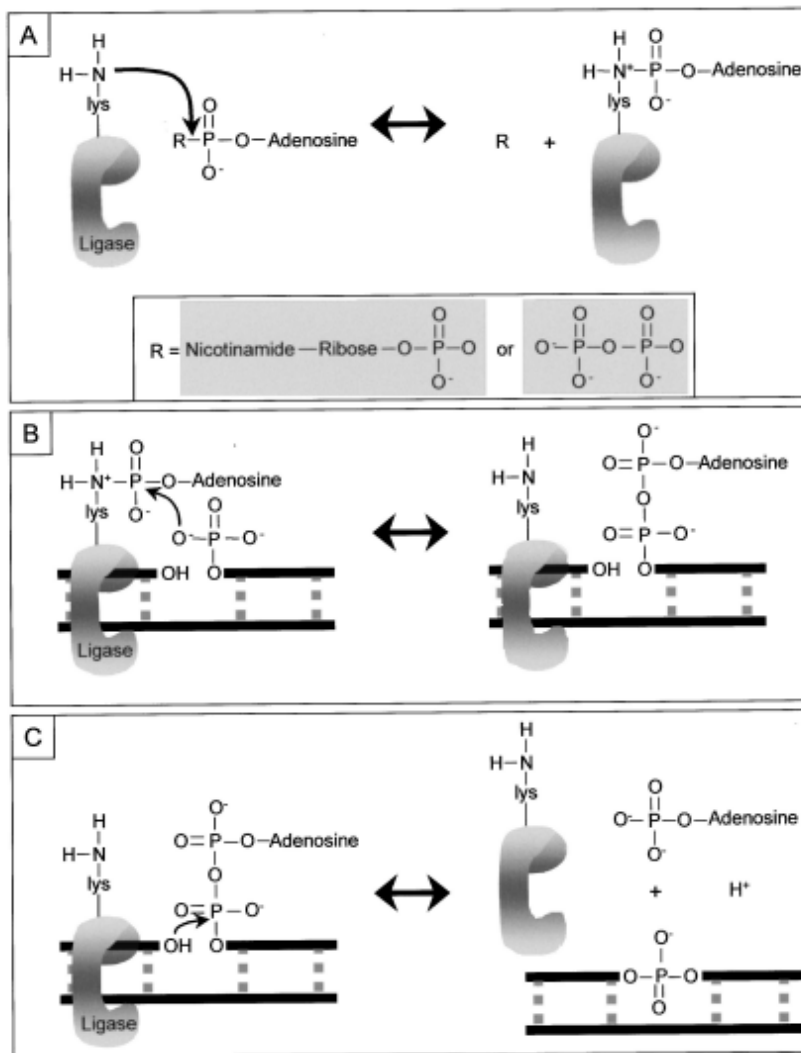
Figure 1.4. DNA ligation occurs in three 'ping-pong'-like steps (A, B and C). In step A, the catalytic lysine (lys) of the ligase is adenylated by an adenylate donor. In step B, the adenyl group is transferred onto the 5'-phosphate end of the nick. In step C, the activated 5'-phosphate end interacts with the 3'-hydroxy end of the nick, forming a phosphodiester bond, which is followed by the release of the adenylate by-product. Figure retrieved from (Wilkinson *et al.*, 2001).

There are two different adenylate donors used by DNA ligases; β-nicotinamide adenine dinucleotide (β-NAD$^+$) and adenine triphosphate (ATP).

1.2.1   β-NAD$^+$- dependent ligases (NDLs)

In all bacteria, NDLs (β-NAD$^+$-dependent ligases) carry out the necessary housekeeping processes like repair and recombination, although they are also found in a few species of archaea and viruses (Wilkinson *et al.*, 2001; Shuman & Lima, 2004; Pergolizzi *et al.*, 2016). The leaving group of β-NAD$^+$ is β-NMN (β-

nicotinamide mononucleotide). These highly conserved ligases, also known as Lig A, consist of five domains (Figure 1.5); domain 1 (PF01653), the AD domain including subdomain 1a that binds the leaving β-NMN group, and subdomain 1b that binds the resultant AMP; domain 2 (PF03120), the OB domain which engages the DNA; domains 3a (PF03119) and 3b (PF14250) which contain a zinc finger and a helix-hairpin-helix (HhH) motif respectively; and domain 4 (PF00533) which contains a breast cancer carboxy terminal (BRCT) domain (Doherty & Suh, 2000; Wilkinson *et al.*, 2001; Magnet & Blanchard, 2004; Shuman & Lima, 2004; Zhu & Shuman, 2007; Pergolizzi *et al.*, 2016). The additional domains (domains 3a, 3b and 4) aid in complete encirclement and engagement with the substrate DNA (Nandakumar *et al.*, 2007).

Protein Family Name	Pfam Domain Architecture	Size of Protein (number of amino acids)
E. coli LigA	PF01653 PF03120 PF03119 PF14520 PF00533	671

Figure 1.5. Pfam domain architecture of the NAD$^+$-dependent Lig A in *Escherichia coli*. Lig A contains three additional domains (PF03119, PF14250 and PF00533) in addition to the core adenylation domain (PF01653) and the oligonucleotide-binding domain (PF03120). Figure from (Pergolizzi *et al.*, 2016).

1.2.2 ATP-dependent ligases (ADLs)

In contrast to bacteria, eukaryotes, archaea, bacteriophage and a majority of eukaryotic viruses rely on ADLs (ATP-dependent ligases) for housekeeping processes (Tomkinson *et al.*, 2006; Williamson *et al.*, 2016; Shi *et al.*, 2018). In addition to the five conserved motifs in the AD domain (Figure 1.3), the OB domain of ADLs also contain a sixth conserved motif important for interacting with the β- and γ-phosphates of ATP that later form the pyrophosphate (PP$_i$) leaving group (Cheng & Shuman, 1997; Odell *et al.*, 2000; Shuman & Lima, 2004; Pergolizzi *et al.*, 2016). However, expression of ADLs in bacteria is not ubiquitous. Differences exist between the structure and function of b-ADLs (bacterial-ADLs), as well as the number of b-ADLs expressed in a species (Subramanya *et al.*, 1996; Williamson & Pedersen, 2014). For example, the most common extension to the AD and OB domains among ADLs is an N-terminal DNA binding domain consisting of alpha

helices that allow complete encirclement of the DNA substrate (Doherty & Suh, 2000; Shi *et al.*, 2018; Chen *et al.*, 2019). It is strongly believed that these ligases were obtained through horizontal gene transfer, and act in bacteria as accessory enzymes in concert with the essential roles of NDLs (Wilkinson *et al.*, 2001; Williamson *et al.*, 2016). Regardless, based on the domain architecture, b-ADLs can be classified into four different groups; Lig B, Lig C, Lig D and Lig E (Figure 1.6) (Pergolizzi *et al.*, 2016; Williamson *et al.*, 2016).
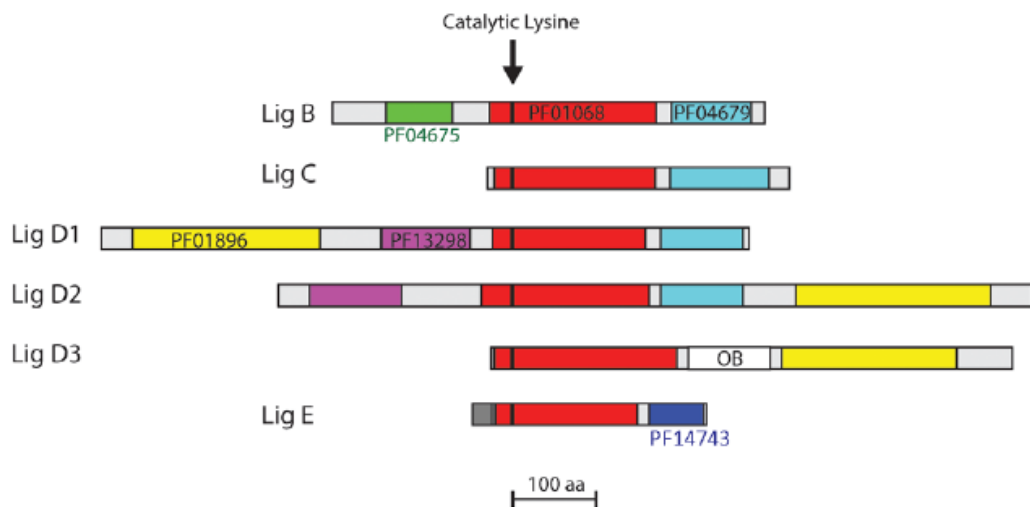
Figure 1.6. Pfam domain architectures of bacterial ATP-dependent ligases (Lig B, C, D and E), with figure from (Williamson *et al.*, 2016). Red (PF01068) illustrates the adenylation domain, and both blue (PF05679; PF14743) and white indicate the oligonucleotide-binding domains. Additional domains are illustrated in green, yellow and purple.

1.2.2.1 Lig B, C, and D

The domain architectures of Lig B, C and D vary greatly. In addition to the core AD (PF01068) and OB (PF04679) domains, Lig B has an additional N-terminal DNA binding domain (PF04675) which allows for nick sealing activity (Gong *et al.*, 2004). Apart from its wide distribution among bacteria, its domain architecture is also found in eukaryotic and archaeal enzymes, making it the largest group of ADLs with this specific domain arrangement (Williamson *et al.*, 2016; Williamson & Leiros, 2020). The gene itself is often found in clusters with a helicase, an endonuclease and an exonuclease (Gong *et al.*, 2004; Williamson *et al.*, 2016). In contrast, Lig C has a minimal structure, consisting only of the core AD and OB

domains (Zhu & Shuman, 2007). Originally, due to the presence of a Ku homologue in Lig C-containing bacteria, it was believed that Lig C acted as a backup mechanism for non-homologous end joining (NHEJ) of double-stranded breaks (DSBs), which Ku proteins are critical for. However, it has since been shown that Lig C is instead responsible for base excision repair during stationary phases, sealing gaps formed by lesion removal (Płociński *et al.*, 2017). Lig C is found in gene clusters with a Prim-Pol (Primase-Polymerase) enzyme, nucleases, and DNA glycosylases, and is present in a few eukaryotes and archaea (Williamson *et al.*, 2016).

On the other hand, Lig D has an additional PrimPol (DNA primase/polymerase) domain (PF01896) and a PE (phosphoesterase) domain (PF13298) encoded in the same polypeptide (Weller *et al.*, 2002). Three arrangements of these domains exist, namely Lig D1, D2 and D3 (Figure 1.6). It is this ligase that is responsible for the aforementioned NHEJ of DSBs through interactions with a Ku homologue protein via the PrimPol domain (Weller *et al.*, 2002; Zhu & Shuman, 2007). The PE domain is important for converting 3'-PO_4 ends resulting from 'dirty' DNA breaks to 3'-OH ends that can be ligated (Zhu & Shuman, 2007). Although Lig D2 may be found in eukaryotes and archaea, Lig D1 and D2 are solely bacterial (Williamson *et al.*, 2016).

Regardless of their different architectures, it is believed that Lig B, Lig C and Lig D were inherited together from a common ancestor of possible archaeal origins, due to their structural similarities and frequent co-occurrence amongst each other (Pergolizzi *et al.*, 2016; Williamson *et al.*, 2016). For example, both Lig C and Lig D2 are often found with Lig B, although some exceptions do exist (Williamson *et al.*, 2016).

1.2.2.2   Lig E

Despite the high diversity of the aforementioned b-ADL architectures, the first b-ADL characterised in 1997 was a small ligase, Lig E, found in *H. influenzae* (hereinafter Hin-Lig) (31 kDa) (Cheng & Shuman, 1997). Although Lig E contains the core AD and OB domains, its OB domain (PF14743) is smaller than those of

other b-ADLs, with fewer helical structures (Williamson *et al.*, 2016). Interestingly, this shorter OB domain is similar to that found in the T7 bacteriophage and the *Chlorella* virus (Subramanya *et al.*, 1996; Odell *et al.*, 2000). Apart from the two core catalytic domains, no other accessory domains are appended to the enzyme, making it extremely compact and minimal (20.7-32.5 kDa) (Magnet & Blanchard, 2004; Williamson *et al.*, 2016). Due to its ability to seal singly-nicked DNA, it is hypothesised that the function of Lig E lies solely on sealing single-stranded nicks (Williamson & Pedersen, 2014).

Owing to its minimal nature, Lig E forms a C-shaped clamp around its substrate, which only partially encircles the DNA, with reliance on the basic residues of the OB domain to form charge-pair interactions that do not require prior-ordering (Magnet & Blanchard, 2004; Williamson *et al.*, 2014; Williamson *et al.*, 2018). In contrast, the classic minimal DNA ligase, the *Chlorella* ligase (hereinafter ChV-Lig) (298 amino acids), has an additional lysine-rich β-hairpin loop (30 amino acids) that extends from the OB domain and provides intrinsic nick sensing (Odell *et al.*, 2000; Nair *et al.*, 2007). In comparison to the bacterial Lig E (b-Lig E), ChV-Lig undergoes an increase in secondary structuring when DNA is bound, whereby its disordered loop becomes structured upon DNA binding (Odell *et al.*, 2000; Nair *et al.*, 2007). Regardless, in both cases, the DNA nick sits above the AMP binding pocket, forming the basis of the Lig E-associated intrinsic nick-sensing abilities (Williamson *et al.*, 2018).

A notable feature of b-Lig E is the lysine rich motif (Gly(glycine)-K-Gly-K-Aromatic) between the 9th and 10th β-strand of the OB domain which replaces the latch found in ChV-Lig, as well as the F-Basic-I(isoleucine)-Gly-S(serine)-Gly-F-x-D motif (Williamson *et al.*, 2018). As of now, three different conformations of Lig E have been characterised via X-ray crystallography; a partially open conformation in Psy-Lig (Lig E from *Psychromonas spp.*), a fully open conformation in Psy-Lig, and a DNA-bound conformation in Ame-Lig (Lig E from *Alteromonas mediterranea*) (Williamson *et al.*, 2018). Despite their differences, all three conformations share the same stabilising non-covalent bonds between the AD and OB domains (Williamson *et al.*, 2018).

1.2.2.2.1 *Phylogenetic analyses of Lig E*

Based on the domain architecture and phylogenetic analyses, Lig E is believed to have been acquired separately from Lig B, C and D, as it forms a different clade from the other b-ADLs (Figure 1.7) (Williamson *et al.*, 2016). This is supported by similarities found between Lig B, C and D with archaeal ADLs, and that between Lig E and bacteriophage ADLs, which suggest a bacteriophage origin of Lig E via horizontal transfer (Williamson *et al.*, 2016). Based on analyses conducted on Hin-Lig, the origins of Lig E is speculated to be quite recent, owing to the high similarities between Hin-Lig and Lig E from *Neisseria spp*, which is absent in other *Haemophilus spp*. like *Haemophilus ducreyi* (Williamson *et al.*, 2016). Additionally, Lig E does not co-occur with any other b-ADLs, and no consistent *Lig E* gene organisation can be identified among different bacteria (Williamson *et al.*, 2016). Most interestingly, it is largely restricted to Gram-negative proteobacteria, including both free-living and commensal pathogens (like *H. influenzae*, a gammaproteobacterium, and *N. gonorrhoeae*, a betaproteobacterium), as well as marine bacteria (Williamson *et al.*, 2016). Of these bacteria, none encode a Ku homologue in their genomes, suggesting that Lig E is not involved in NHEJ repair.
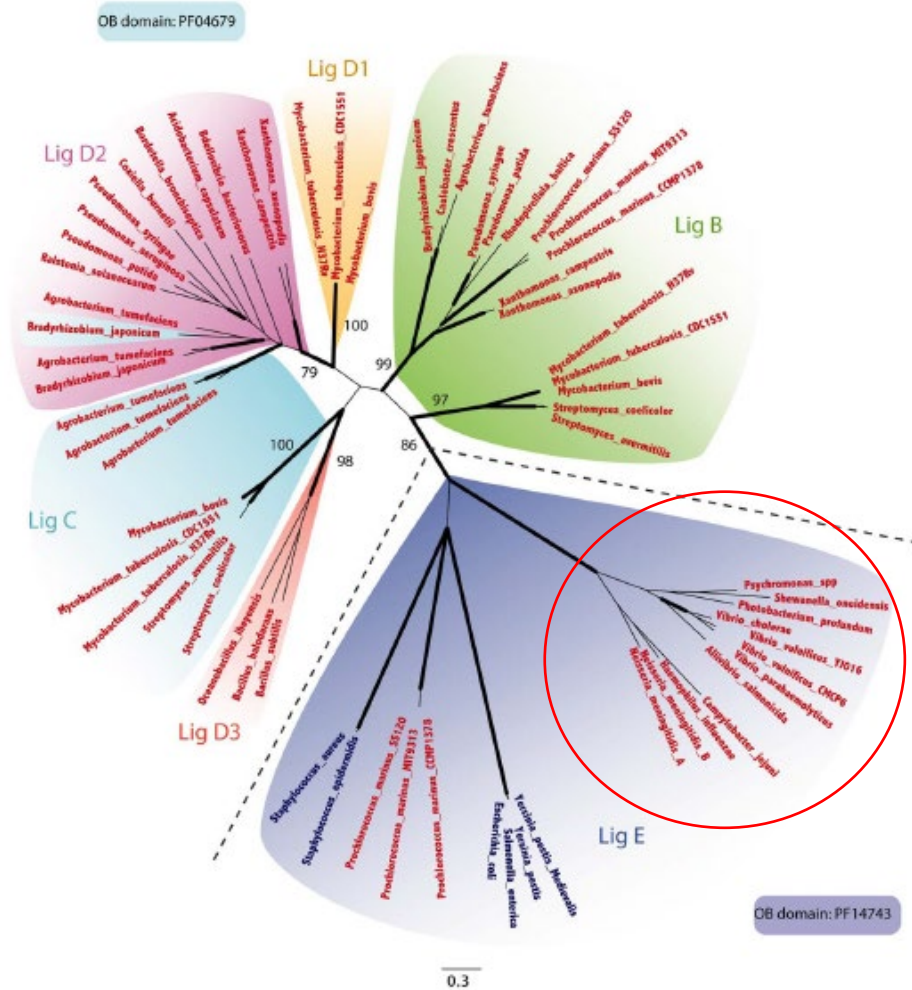
Figure 1.7. Phylogenetic tree of bacterial ATP-dependent ligases based on their domain arrangements and biochemical characteristics, retrieved from (Williamson *et al.*, 2016). Lig E (circled) forms a different clade from the other ATP-dependent ligases. The three left branches in the Lig E clade (uncircled) are no longer classified as Lig E due to the presence of additional DNA binding domains.

1.3 Current hypothesis: A role for Lig E in DNA uptake

Interestingly, Lig E contains a putative 20-40 amino acid-long periplasmic signal peptide on its N-terminal region, the removal of which increases its activity and stability, providing further evidence of its role as a localising sequence (Magnet & Blanchard, 2004; Williamson & Pedersen, 2014). As an organism's genetic material is contained within its cytosol, it is hypothesised that the role of Lig E lies in sealing single-stranded nicks from fragmented exDNA during competence and uptake, which will either increase or replace genetic information (Figure 1.8) (Chaussee & Hill, 1998; Abbasian *et al.*, 2019). This process is believed to occur before the

18

formation of ssDNA and its translocation into the cytoplasm, allowing for longer pieces of DNA to integrate into the genome.
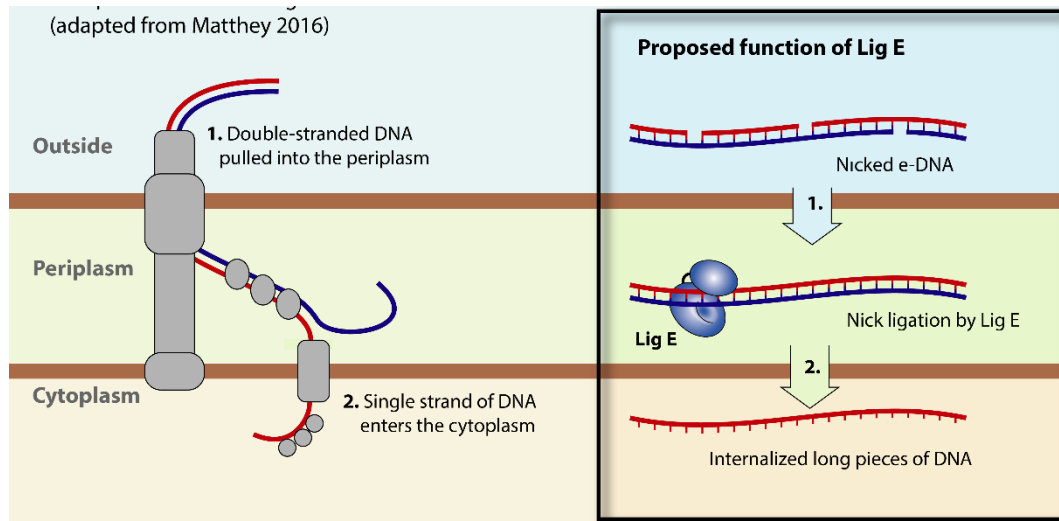
Figure 1.8. Proposed function of Lig E in competence of Gram-negative bacteria (Williamson, 2021, unpublished). It is hypothesised that Lig E seals nicked duplex DNA in the periplasm which increases its integrity for subsequent homologous recombination into the genome.

For example, in a mixed population of bacteria, where some have acquired antibiotic resistance genes, an event like oxidative stress or neutrophil consumption in the human host may kill most of the population but spare some, albeit with some damage to the genomic DNA. The fragmented DNA from the lysed bacteria containing antibiotic resistance genes may then be taken up by the surviving, but damaged bacteria. Lig E may seal the breaks in the antibiotic resistance-containing fragments, allowing integration and acquisition of these genes into the genome. This hypothesis is supported by both the presence of a ComEA homologue in Lig E-containing bacteria that indicates bacterial transformability, as well as evidence of extracellular ATP that may allow Lig E to function in the periplasm (Mempin *et al.*, 2013; Williamson *et al.*, 2018). However, Lig E-containing bacteria are not the only naturally competent bacteria, indicating that this pathway may be an alternative pathway assisting DNA uptake, albeit not essential (Williamson *et al.*, 2016). Thus, Lig E would not be vital for transformation, but its function may be to enhance uptake of long DNA pieces.

The characterisation and study of Lig E would have important implications considering the number of human pathogens that contain the enzyme including *H. influenzae* (Hin-Lig), *N. meningitidis* (Nme-Lig), *N. gonorrhoeae* (Ngo-Lig), *V. cholerae*, and *Campylobacter jejuni*, making it a strong candidate as a therapeutic target against these pathogens (Williamson *et al.*, 2018). However, to date, no *in vivo* experiments have been conducted to confirm either its potential extracellular or periplasmic location, and its potential role of sealing exDNA during DNA competence. In particular, studying this minimal enzyme in the potential superbug, *N. gonorrhoeae*, is not only important for gaining a better understanding of horizontal gene transfer across different organisms, but is also critical for finding new targets against the bacteria before it becomes an untreatable epidemic across the globe.

1.4 Research objectives

The aim of my Master's project was to determine the biological location and function of Lig E in *N. gonorrhoeae* (hereinafter, Ngo-Lig) and to investigate its potential role in bacterial competence. Hence, the following objectives were set:

1) Determine the importance of Ngo-Lig on cell viability and growth by generation of *Lig E* knockouts in *N. gonorrhoeae*.
2) Investigate the cellular location of Ngo-Lig to confirm a potential periplasmic location using fluorescent and/or antigenic tagging.
3) Characterise the DNA ligase activity of Ngo-Lig *in vitro* to provide insight into its biological function.
4) Develop a DNA uptake reporter assay to quantify the role of Lig E in competence.

Chapter Two

Methods

2.1 Construction and evaluation of Ngo-Lig mutants *in vivo*

2.1.1 Manipulation of *N. gonorrhoeae*

2.1.1.1 General *N. gonorrhoeae* culture

Throughout the project, the MS11 *N. gonorrhoeae* strain was used for all gonococcal work. Glycerol stocks (25%) were streaked and cultured onto either solid gonococcal base medium (GCB) agar or liquid Gibco® Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) media and kept at 37 °C with 5% carbon dioxide (CO_2) (Dillard, 2011).

To prepare GCB agar, Difco^{TM} gonococcal (GC) medium base (36.25 g) was mixed with Difco^{TM} granulated agar (1.25 g) in Milli-Q® Ultrapure water (1 L) before autoclaving. BD BBL^{TM} IsoVitalex^{TM} Enrichment (10 mL) and the desired antibiotics (Table 2.1) was added to cooled agar (50-60 °C), before pouring into agar plates (roughly 25 mL per plate). The plates were stored at 4 °C, pre-warmed before use, and kept upside-down after inoculation with culture.

Table 2.1. Concentrations of antibiotics used for general *N. gonorrhoeae* culture. Erythromycin and kanamycin were used for the selection of resistant strains, while vancomycin and amphotericin B were used to prevent contamination.

Antibiotics	Final concentration used (µg/mL)
Erythromycin	10
Kanamycin	50
Vancomycin	2.5
Amphotericin B	1

The DMEM/F-12 media used for liquid culture of MS11 was supplemented with 10% foetal bovine serum (FBS) (Moregate Biotech), as well as the desired antibiotics highlighted in Table 2.1. The media was pre-warmed before inoculation with MS11.

2.1.1.2 Gram-staining of gonococcal cultures

Gram-staining was performed to confirm the presence of the Gram-negative *N. gonorrhoeae* in the cultures (Smith & Hussey, 2005). Resuspended colonies (one colony in 20 µL Milli-Q® Ultrapure water) were smeared onto a microscope slide before passing the slide through a flame. Crystal violet staining reagent (crystal violet (10 g/L)) was added onto the slide for at least one minute. The dye was washed away before the addition of Gram's iodine solution (iodine (3.3 g/L) and potassium iodide (6.7 g/L)) for one minute, which was washed off before the addition of decolourising solution (25% acetone and 75% isopropyl alcohol) for at least 15 seconds. After flooding the slide with safranin solution (Safranin O (4 g/L) in 20% ethanol) for at least one minute, the slide was washed, and the colonies were visualised under a microscope with the oil immersion lens.

2.1.2 Generation of Ngo-Lig mutants in *N. gonorrhoeae*

To evaluate the function and location of Ngo-Lig *in vivo*, three different Ngo-Lig mutants were of interest (these are described in more detail in Chapter Three and shown in comparison to the WT genome in Figure A.1 in the Appendices). These included the Lig E-green fluorescent protein (GFP) construct (Lig E-GFP; Figure A.2), the Lig E-His (histidine)-Kanamycin resistance (KanaR) construct (Lig E-His-KanaR; Figure A.3) and the Lig E-knock-out (KO)-KanaR construct (Lig E-KO-KanaR; Figure A.4). The constructs were ordered from Twist Bioscience and were reconstituted at 10 ng/µL in Tris-Ethylenediaminetetraacetic acid (EDTA) (TE) buffer (Tris (tris(hydroxymethyl)aminomethane) pH 8.0 (10 mM) and EDTA pH 8.0).

2.1.2.1 Spot transformation and securement of Ngo-Lig mutants

Ten ng of the constructs of interest (i.e. 10 µL) were spotted twice onto GCB agar. Single wild-type (WT) piliated MS11 gonococcal colonies (at least 36 hours old) were streaked through the spots. After a 24-hour incubation at 37 °C and 5% CO_2, colonies growing at the spotted locations were re-streaked onto GCB plates with kanamycin for selection. These were screened via colony polymerase chain reactions (PCRs) and sequencing as described below.

Positive transformants were secured as glycerol stocks (25%) by mixing 500 µL of the liquid culture of the transformants (10% FBS DMEM/F-12) with 500 µL of 50% glycerol in cryotubes. These were stored at -80 °C.

2.1.2.2 Validation of gonococcal transformants via colony PCR

Transformations of individual colonies growing on kanamycin-containing GCB plates were verified via colony PCR. A PCR master mix based on the number of colonies to be tested was prepared (Table 2.2) before aliquoting 20 µL of the mix into each PCR tube. Primers used for verification of the different Ngo-Lig mutants are listed in Table 2.3. In all cases, the primers were reconstituted in TE buffer. Single colonies were picked using a sterile pipette tip and streaked onto a kanamycin GCB grid plate before mixing the tip with the respective PCR mixture. The GCB grid plate was left incubating at 37 °C and 5% CO_2.

Table 2.2. The volume of reagents used for individual gonococcal colony PCRs

Reagents	Volume (µL)
5x HOT FIREPol® Blend Master Mix (Solis BioDyne)	4
Forward primer (10 µM)	0.6
Reverse primer (10 µM)	0.6
Milli-Q® Ultrapure water	14.8
Total volume	**20**

Table 2.3. Primer combinations used for Ngo-Lig mutant verification. Sequences and melting temperatures of the primers are given in Table A.1.

Construct/genotype detected	Forward primer	Reverse primer	Annealing temperature (T_a) (°C)
Lig E-GFP	Primer_3_GFP_FD	Primer_2_Rev	55.0
Lig E-His-KanaR	Primer_4_His_FD	Primer_2_Rev	56.1
Lig E-KO-KanaR	Primer_1_KO_FD	Primer_2_Rev	53.8
WT	Primer_1_KO_FD	Primer_2_Rev	53.8

PCR was run on the PCR T100^{TM} Thermal Cycler (Bio-Rad) (as per Table 2.4) with a WT and a water-blank negative control. The products (10 µL) were electrophoresed on 1% agarose gels at 100 V, which were prepared using HyAgarose^{TM} LE Agarose (1g) and Tris base, acetic acid and EDTA (TAE) buffer (100 mL of Tris (50 mM), acetate (20 mM) and EDTA (1 mM)) with either SYBR® SAFE DNA Gel Stain or thiazole orange (10,000 in dimethylsulfoxide (DMSO)) (10 µL). A 1 Kb Plus ladder (7 µL, Invitrogen^{TM}) (Figure D.1) was run at the same time. The agarose gels were visualised using the iBright FL1000 Imaging System (ThermoFisher Scientific).

Table 2.4. Polymerase chain reaction conditions for *N. gonorrhoeae* mutant screening

Temperature	Duration	Repeats
95 °C	15 min	1x
95 °C	20 secs	
2-6 °C below the lowest melting temperature (T$_m$) of the primer set used	30 secs	30x
72 °C	1 min/kilo base (kb) of PCR product	
72 °C	10 min	1x

2.1.2.3   Confirmation of positive gonococcal transformations via sequencing

PCR-verified colonies of the Ngo-Lig transformants were reamplified using the sequencing primers (Primer_seq_Lig-E_Fd and Primer_seq_Lig-E_Rev (annealing temperature (T$_a$) = 53.1 °C) primers (Table A.1)) before verification on a 1% agarose gel and visualisation as per section 2.1.2.2. The products were either directly purified from the PCR mixture (column purification) or the gel (gel purification) before storage at -20 °C.

For column purifications, PCR products were purified after amplification using the High Pure PCR Product Purification Kit (Roche) as per the manufacturer's protocol. Quantification (ng/µL) and purity (A260/A280 and A260/A230) were confirmed via the NanoDrop^{TM} 2000 spectrophotometer (1.5 µL). For gel purifications, the PCR product bands on 1% agarose were visualised under ultraviolet (UV) light and

excised using a scalpel before purification using either the High Pure PCR Product Purification Kit (Roche) or the QIAquick® Gel Extraction Kit (QIAGEN) (15 and 25 µL elution volume respectively). Quantification and purity were confirmed via the NanoDrop^{TM} 2000 spectrophotometer (1.5 µL).

Based on the Nanodrop^{TM} concentrations, the respective volumes of PCR product required was calculated to send for sequencing (2.5 ng of DNA for every 100 bp along with either the forward or reverse sequencing primers (4 pmol)). The sequencing mixture was made up to 20 µL with UltraPure^{TM} DNase/RNase-free distilled water (Invitrogen^{TM}) and sent to Massey Genome Service for sequencing before analysis in Geneious Prime (v. 2020.1) (Kearse *et al.*, 2012).

2.1.3 Growth assay of the Ngo-Lig mutants

A growth experiment was performed with the Lig E-His-KanaR MS11 and Lig E-KO-KanaR MS11 mutants generated in section 2.1.2, as well as the WT MS11. For each condition, three seeder cultures were prepared using 10% FBS DMEM/F-12 (5 mL) from colonies grown on GCB agar (with kanamycin for both Lig E-His-KanaR and Lig E-KO-KanaR) 48 hours prior. Different measurements were recorded at specific times as per Table 2.5.

Table 2.5. Outline of the *N. gonorrhoeae* growth experiment performed to obtain optical density (OD$_{600}$) and colony forming unit (CFU) readings, as well as to isolate supernatant and pellet samples for further analysis

Time (hr)	Measurements		
0	OD$_{600}$		
2	OD$_{600}$	CFU	
5	OD$_{600}$		Supernatant and pellet isolation
13	OD$_{600}$	CFU	
15	OD$_{600}$		Supernatant and pellet isolation
17	OD$_{600}$	CFU	
19	OD$_{600}$		Supernatant and pellet isolation
21	OD$_{600}$	CFU	
23	OD$_{600}$		Supernatant and pellet isolation
25	OD$_{600}$	CFU	
27	OD$_{600}$		Supernatant and pellet isolation
29	OD$_{600}$	CFU	
31	OD$_{600}$		Supernatant and pellet isolation
34	OD$_{600}$	CFU	
37	OD$_{600}$		Supernatant and pellet isolation
39	OD$_{600}$	CFU	
41	OD$_{600}$	CFU	Supernatant and pellet isolation
43	OD$_{600}$	CFU	
45	OD$_{600}$	CFU	Supernatant and pellet isolation
47	OD$_{600}$	CFU	
49	OD$_{600}$	CFU	Supernatant and pellet isolation
51	OD$_{600}$	CFU	
53	OD$_{600}$	CFU	Supernatant and pellet isolation
61	OD$_{600}$	CFU	

The OD$_{600}$ reading for each seeder culture was measured on a Biochrom WPA C08000 Cell Density Meter. Based on the OD$_{600}$ readings, the amount required to inoculate 30 mL of media at an OD$_{600}$ of 0.01 was calculated and the respective amount was added to the 30 mL of media (three repeats for each condition). To measure the amount of colony forming units at the specific time points outlined in

Table 2.5, a 10^{-2}, 10^{-4} and 10^{-6} dilution was prepared for each culture (200 µL final volume). Each dilution mixture was spotted three times (5 µL) onto GCB plates before incubation at 37 °C and 5% CO_2. The number of single visible colonies were counted where possible under a Maggylamp magnifier.

At the specific time points outlined in Table 2.5, supernatant and pellet samples were isolated from the 1 mL sample used for OD_{600} measurements, which was split into two 500 µL samples and centrifuged (10,000 rpm, 10 min). The supernatant from both samples were isolated and frozen at -80 °C for DNA analyses, along with one of the pellet samples for future protein analysis. The remaining pellet was resuspended in TRIzol^{TM} Reagent (1 mL, Invitrogen^{TM}) before freezing at -80 °C for potential transcriptomics analyses in the future.

2.1.4   Extracellular DNA analysis of Ngo-Lig mutants

Analyses of the supernatant samples were performed using the Quant-iT^{TM} PicoGreen^{TM} dsDNA Assay Kit (ThermoFisher Scientific) according to the manufacturer's instructions (Mansfield *et al.*, 1995). A calibration curve was built using lamba DNA prepared via serial dilution in TE buffer (0, 1, 10, 100 and 1000 ng/µL) with a 1:1 ratio of PicoGreen® dsDNA dye (in DMSO) in a total volume of 400 µL. Fluorescence at 500-550 nm was measured using a Hitachi F-7000 Fluorescence Spectrophotometer (photomultiplier tube voltage: 500 V; excitation wavelength: 480 nm; scan speed: 1200 s; excitation slit: 5 nm; emission slit: 5nm; response: 0.5 s).

Stored supernatant samples (500 µL) from hours 5, 23, 31, 45 and 51 of the WT MS11, the Lig E-KO-KanaR mutant and the Lig E-His-KanaR mutant were subjected to the same fluorescence analysis after addition of EDTA (5 mM) and subsequent heat treatment at 95 °C (10 min). A 10% FBS DMEM/F-12 sample served as a blank. The lambda DNA calibration curve was used to calculate the amount of DNA detected in the supernatant samples.

2.1.5 Western blot for detection of the His-tagged Ngo-Lig protein

2.1.5.1 Sample preparation and sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Western blot was performed to quantify the amount of protein of interest in both the supernatant and resuspended pellet samples from the growth experiment. Pellet samples were resuspended in 100 µL lysis buffer (Tris pH 8.0 (50 mM), NaCl (750 mM), MgCl$_2$ (1 mM), glycerol (5%)). Both the supernatant and resuspended pellet samples were prepared with a 1 in 4 dilution of QX4 loading dye (4x) (Tris pH 6.8 (250 mM), glycerol (20%), SDS (4%), mercaptoethanol (10%) and bromophenol blue (0.025% w/v)), which was heated at 95 °C (10 mins). These were run on SDS-PAGE (sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE)) gels (10 µL) comprised of a 5% stacking gel layer on top of a 12% resolving gel layer (Table 2.6). The BenchMark^{TM} His-tagged Protein standard (ThermoFischer Scientific) (7 µL, Figure D.3) was loaded with every run. Gels were run at 80-90 V for the stacking layer and 100-120 V for the resolving layer with Tris-glycine-SDS (TG-SDS) buffer (Tris (25 mM), glycine (250 mM), SDS (0.1% w/v)).

Table 2.6. Reagents required to make five SDS-PAGE gels

	5% Stacking gel (mL)	12% Resolving gel (mL)
Milli-Q® Ultrapure water	8.500	10.050
30% Acrylamide (Bio-Rad)	2.125	12.000
Stacking buffer (Tris pH 6.8, 1.0 M)	1.600	-
Resolving buffer (Tris pH 8.8, 1.5 M)	-	7.500
10% SDS	0.125	0.300
10% ammonium persulfate (APS)	0.063	0.150
Tetramethylethylenediamine (TEMED)	0.0063	0.015
Total	**12.4193**	**30.015**

2.1.5.2   Protein transfer and antibody incubation

Once run, the proteins from the SDS-PAGE gels were transferred onto either a polyvinylidene fluoride (PVDF) transfer membrane (ThermoFisher Scientific; 0.45 μm) or an Amersham™ Protran™ nitrocellulose blotting membrane (Cytiva; 0.45 μm) in a transfer sandwich comprised of a sponge, two filter papers, the membrane and the gel, two additional filter papers and an additional sponge, all of which were equilibrated with transfer buffer (Tris (25 mM), glycine (192 mM), methanol (20%), SDS (0.01%)) for at least five minutes. Transfer was performed in a tank with an ice block at 100 V (one hour) with stirring of the transfer buffer, before staining of the membrane with Ponceau S (Red ponceau (0.2%, Sigma-Aldrich), acetic acid (1%)) for at least five minutes to check protein transfer efficiency. The membrane was washed with Tris buffered saline-Tween 20 (TBS-T) solution (Tris (20 mM), NaCl (0.15 M), Tween-20 (0.1%), pH 7.6) before blocking with the blocking buffer (10% low fat milk in TBS-T) for an hour. Incubation with the desired concentrations of primary antibody (in blocking buffer) was performed overnight at 4 °C, followed by five washes with TBS-T before the final membrane incubation with the desired concentrations of secondary antibody (in blocking buffer). The membrane was washed five times with TBS-T and visualised using the iBright FL1000 Imaging System.

Two primary antibodies and three secondary antibodies were used in this project. The primary antibodies used were an anti-His-tag mouse monoclonal (HIS.H8), sc-57598 igG$_{2b}$ antibody (Santa Cruz Biotechnology; 100 μg/mL) and an anti-*N. gonorrhoeae* monoclonal igG$_{2a}$ antibody (386/418) ab62964 (Abcam; 1mg/mL) against the gonococcal outer membrane protein (OMP). The secondary antibodies used were an anti-mouse m-igGκBP-CFL 647: sc-516179 monoclonal antibody conjugated to the CruzFluor™ 647 fluorophore (Santa Cruz Biotechnology; 200 μg/mL), an anti-mouse m-igGκBP-HRP: sc-516102 monoclonal antibody conjugated to horseradish peroxidase (Santa Cruz Biotechnology; 200 μg/mL) and a goat anti-mouse polyclonal IgG antibody conjugated to horseradish peroxidase ab97023 (Abcam; 1mg/mL). Before visualisation, membranes incubated with the horseradish peroxidase-conjugated secondary antibodies were incubated with a 1:1 ratio of the SuperSignal™ West Femto Luminol/Enhancer Solution (ThermoFisher

Scientific) and the SuperSignal™ West Femto Stable Peroxide Solution (ThermoFisher Scientific) for at least five minutes.

2.2 Recombinant production of Ngo-Lig

2.2.1 General *E. coli* culture

Streaked or inoculated *E. coli* cells were incubated at 37 °C (unless otherwise specified) and always upside-down if on solid agar or with shaking (180-200 rpm) if in liquid media. The DH5α strain was used for plasmid isolation and the BL21(DE3)pLysS and Origami™(DE3) strains were used for protein expression (Passarinha *et al.*, 2006; Dvorak *et al.*, 2015; Tang *et al.*, 2016). The genotypes of the different *E. coli* strains of interest can be found in Table B.1.

Solid lysogeny broth (LB) agar used for *E. coli* growth was prepared using Bacto™ peptone (10 g), yeast extract (5 g, Condalab), sodium chloride or NaCl (10 g) and Bacto™ agar (15 g), made up to 1 L with Milli-Q® Ultrapure water. The media was autoclaved and the desired antibiotics required (Table 2.7) were added, before pouring into agar plates (30 mL per plate) when needed.

Table 2.7. Concentrations of antibiotics used for general E. coli culture for selection of resistant strains

Antibiotics	Final concentration used (μg/mL)
Ampicillin	100
Chloramphenicol	34
Kanamycin	50

The liquid LB media/broth used for *E. coli* was prepared using Bacto™ peptone (10 g), yeast extract (5 g) and NaCl (10 g), made up to 1 L with Milli-Q® Ultrapure water before autoclaving. The desired antibiotics required (Table 2.7) were added on the day of usage.

2.2.1.1 Preparation and transformation of chemically competent E. coli cells

Throughout the project, chemically competent *E. coli* cells were used for transformations. These were prepared by streaking *E. coli* glycerol stocks onto LB agar before incubation overnight (37 °C). A single colony was picked and added onto LB media (10 mL) before incubation overnight (37 °C, 200 rpm). A portion of the seeder culture (1 mL) was added onto warmed LB media (100 mL) and incubated (37 °C, 180-200 rpm) until the optical density (OD) at 600 nm (OD_{600}) readings measured on the Multiskan G0 spectrophotometer (ThermoFisher Scientific) reached 0.5-0.7. The culture was chilled on ice (30 min) before centrifugation (4600 rpm, 20 min, 4 °C) to isolate the pellet, which was resuspended in tryptic soy broth (TSB) media (10 mL) (10% w/v (weight per volume) polyethylene glycol 8000, 10 mM magnesium sulfate ($MgSO_4$), 20 mM magnesium chloride ($MgCl_2$) and 5% DMSO added directly before use). Cells were aliquoted (100 µL samples) and snap-frozen using liquid nitrogen before storage at -80 °C.

To 100 µL of thawed chemically competent DH5α cells, plasmids of interest (5-10 µL) were added before incubation on ice (30 min). The cell mixture was heated (42 °C, 45 sec) and cooled (on ice, 2 min) before the addition of super optimal broth with catabolite repression (SOC) media (1 mL) (Bacto^{TM} Tryptone (20 g), yeast extract (5 g), NaCl (0.585 g), potassium chloride or KCl (0.186 g), $MgSO_4.7H_2O$ (2.46 g) and $MgCl_2.6H_2O$ (20.3 g) made to 1 L with Milli-Q® Ultrapure water) and incubation at 37 °C (1 hr). The cell mixture was spun (13000 rpm, 1 min) and the pellet was resuspended in 200 µL of supernatant before spreading onto LB agar with the desired antibiotics, followed by incubation overnight at 37 °C.

Seeder cultures were prepared by picking an overnight colony and inoculating LB media (5 mL) with the desired antibiotics before incubation overnight (37 °C, 180-200 rpm). Glycerol stocks (25%) were prepared using 500 µL of the culture and an equal volume of 50% glycerol, which were stored at -80 °C.

2.2.1.2 Plasmid purification and confirmation via PCR

After preparation of glycerol stocks the remaining transformed DH5α cells seeder cultures were centrifuged (4600 rpm, 8 min) and purified using the QIAprep Spin

Miniprep Kit (Qiagen) according to the manufacturer's protocol. Concentration and purity of the isolated plasmids were determined by NanoDrop™ before storage of the plasmid at -20 °C. PCR was performed (as per Table 2.4 and Table 2.8) before visualisation of the bands under UV after a 1% agarose run.

Table 2.8. The volume of reagents used for construct polymerase chain reactions with 2 µL of construct DNA

Reagents	Volume (µL)
5x HOT FIREPol® Blend Master Mix	4
Forward primer (10 µM)	0.6
Reverse primer (10 µM)	0.6
Milli-Q® Ultrapure water	12.8
Construct/DNA	2
Total volume	**20**

2.2.2   Recombinant expression of Ngo-Lig

Three Ngo-Lig variants were ordered from Twist Bioscience pre-cloned into pDONR221 donor plasmids (Figure B.1)) between two *attL* (left attachment site) regions. The three Lig E variants that had to be cloned into destination plasmids for protein expression in *E. coli* included His-(TEV)-Lig E-GFP (Figure B.2), His-(TEV)-Lig E (Figure B.3 (a)), and Lig E-His (Figure B.4 (a)). As described in the results of Chapter Four, it was necessary to further modify these donor constructs via restriction digestion and ligation to increase protein yields.

2.2.2.1   Cloning of plasmids via restriction digestion and ligation

Restriction digest was performed on both the His-(TEV)-Lig E and Lig E-His pDONR221 constructs (1 µg plasmid) to produce a pDONR221 His-(TEV)-Lig E-His containing plasmid (Figure B.5 (b)) using EcoRI and BamHI enzymes. Restriction digest mixtures (50 µL) were set up for each construct as per Table 2.9, with the desired restriction digest enzymes added last. The mixtures were incubated at 37 °C while shaking (500 rpm) for two hours. Purple gel loading dye (NEB) (6x, 10 µL) was added onto the 50 µL digest mixtures, before running the whole volume on a 1% agarose gel. The gel was visualised under UV and the fragments of interest

were either excised and purified with the QIAquick® Gel Extraction Kit (Qiagen) (25 µL final elution volume) or the remaining amplification product was purified using the High Pure PCR Product Purification Kit (Roche) with a final elution volume of 15 µL. Concentration and purity were verified via the NanoDrop™ spectrophotometer before storage at -20 °C.

Table 2.9. The composition of a digest mixture prepared for restriction digestion. The restriction digest enzymes were always added last.

Reagents	Amount required
DNA	1 µg
Restriction enzymes	1 µL (10 units)
10x CutSmart™ buffer (New England Biolabs/ NEB)	5 µL
UltraPure™ DNase/RNase-free distilled water (Invitrogen™)	Up to 50 µL final volume

Ligation of the purified digested fragments of interest were performed using a 1:3 ratio of the 'vector' (larger Lig E- His construct (Figure B.4 (c))):'insert'(smaller His-(TEV)-Lig E (Figure B.3 (b)) (30 fmol:90 fmol) based on the expected sizes of 2181 bp for the 'vector' and 992 bp for the 'insert'. The reaction mixture consisted of 10x buffer for T4 DNA ligase with ATP (10 mM) (2 µL, NEB) and T4 DNA ligase (1 µL, Promega, 3U/µL) which was added last. After taking into consideration the amount of purified vector and insert to add based on their expected sizes (Equation 2.1 and Equation 2.2 respectively), the mixture was made up to 20 µL using UltraPure™ DNase/RNase-free distilled water (Invitrogen™). The ligation mixtures were incubated overnight at 16 °C before heat inactivation at 65 °C (10 min) and storage at -20 °C.

Equation 2.1. Calculation of the volume of digested vector or plasmid added during ligation

$$Amount\ of\ vector\ (ng) = \frac{(Number\ of\ bp\ of\ vector) \times 30\ fmol)}{10^6}$$

$$Volume\ of\ vector\ (\mu L) = \frac{Amount\ of\ vector\ (ng)}{Concentration\ of\ vector\ (\frac{ng}{uL})}$$

Equation 2.2. Calculation of the volume of digested insert added during ligation

$$Amount\ of\ insert\ (ng) = \frac{(Number\ of\ bp\ of\ insert) \times 30\ fmol)}{10^6}$$

$$Volume\ of\ insert\ (\mu L) = \frac{Amount\ of\ insert\ (ng)}{Concentration\ of\ insert\ (\frac{ng}{uL})}$$

Isolation of the ligated plasmid was achieved through transformation into chemically competent DH5α cells (selected via kanamycin) before plasmid purification.

2.2.2.2 Gateway cloning/LR recombination to obtain expression plasmids, and transformation of the expression plasmids into *E. coli*

Gateway cloning of the three pDONR221 constructs ordered from Twist Bioscience and the ligated pDONR221 construct was performed with various destination plasmids (pDEST14, pDEST17 and pHMGWA (Figure B.6, Figure B.7 and Figure B.8 respectively)) to create different expression plasmids (Table 2.10).

Table 2.10. Summary of the Gateway cloning of recombinant Lig E protein variants performed throughout the project

Donor plasmid	Destination plasmid	Expected protein size (kDa)
pDONR221 (His-(TEV)-Lig E) (Figure B.3)	pDEST17	32.687 (Figure B.9)
	pHMGWA	73.091 (Figure B.10)
pDONR221 (Lig E-His) (Figure B.4)	pDEST14	29.308 (Figure B.11)
pDONR221 (His-(TEV)-Lig E-GFP) (Figure B.2)	pDEST17	59.379 (Figure B.12)
pDONR221 (His-(TEV)-Lig E-His) (Figure B.5)	pDEST17	32.687 (Figure B.13)
	pHMGWA	73.964 (Figure B.14)

The cloning mixture consisted of the donor plasmid (2 µL, >100 ng/µL), destination plasmid (2 µL) and the Gateway™ LR Clonase™ II (Invitrogen), which was incubated at room temperature overnight. Proteinase K (Invitrogen) (1 µL, 2 µg/µL) was added before a final incubation (37 °C, 10 min).

Transformation was performed onto chemically competent DH5α cells using the whole cloning reaction mixture (5 µL) with selection via ampicillin, before plasmid purification and quantification, and storage of the plasmids at -20 °C. These plasmids were verified using the T7_promoter_primer and T7_terminator_primer primers (Table B.2) for the pDEST17 and pDEST14-cloned constructs, and the MBP_forward_primer and T7_terminator_primer primers (Table B.2) for the pHMGWA-cloned constructs. Positive (using pure destination plasmid) and negative (using Milli-Q® Ultrapure water) samples were prepared for each PCR amplification. Plasmids were verified via a 1% agarose gel.

2.2.3   Small-scale expression testing of recombinant Ngo-Lig variants

The two chemically competent strains used for expression of recombinant Lig E were BL21(DE3)pLysS and Origami™(DE3) cells (Table B.1). These were treated and grown in the same conditions as the DH5α cells.

2.2.3.1 Transformations, induction and harvesting of Ngo-Lig proteins after small-scale expression

Transformations of the purified expression plasmids (5 µL) were performed and plated onto ampicillin and chloramphenicol LB agar for BL21(DE3)pLysS cells, and ampicillin LB agar for Origami™(DE3) cells. Seeder cultures (5 mL) were prepared for each condition before long-term storage at -80 °C via 25% glycerol stocks.

Growth cultures (50 mL) were prepared from each seeder culture (1 mL) using 1% glucose terrific broth (TB) media with the desired antibiotics. TB media was prepared using Bacto™ Tryptone (12 g), yeast extract (24 g, Condalab), glycerol (4 mL) and Milli-Q® Ultrapure water (900 mL), which was autoclaved before the addition of autoclaved phosphate solution (monopotassium phosphate/KH_2PO_4 (2.31 g), dipotassium phosphate/K_2HPO_4 (12.54 g) in Milli-Q® Ultrapure water (100 mL)). The growth cultures were incubated at 37 °C (180-200 rpm) until the OD_{600} measured on the Multiskan G0 spectrophotometer reached 0.3-0.4. The samples were induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) (0.5 mM), before incubation at various conditions which included overnight at 15 °C, overnight at 20 °C or for four hours at 37 °C.

After the desired incubation, the cultures were centrifuged (4600 rpm, 20 min, 4 °C) before resuspension of the pellet in DNase (2 µL, 20 µg/mL) and lysis buffer (2 mL). The samples were sonicated using a 2 mm probe tip (Table 2.11) before further centrifugation (4600 rpm, 20 min, 4 °C) to isolate the soluble (supernatant) proteins from the insoluble (pellet) proteins.

Table 2.11. Sonication conditions for small- and large-scale protein expression

Probe tip diameter (mm)	Amplitude	Process time (min)	Pulse on (s)	Pulse off (s)	Total volume (mL)
2	4	1	1	1	5
6	3	5	1	1	90

2.2.3.2 Analysis of small-scale expression samples

Ni Sepharose^{TM} High Performance nickel resin beads (20 µL, GE Healthcare) were washed with lysis buffer (1 mL) and spun (10,000 rpm, 30 secs) before removal of the supernatant. The protein lysate/supernatant samples obtained after sonication (from section 2.2.3.1) was added to the washed nickel beads and left at room temperature (15 min) before further centrifugation (10,000 rpm, 30 secs) and removal of the supernatant. This was washed with lysis buffer (1 mL) and spun (10,000 rpm, 30 secs) before final removal of the supernatant.

Gel samples of the soluble protein lysate/supernatant and the insoluble protein/pellet samples obtained after sonication (from section 2.2.3.1), as well as the soluble protein-washed nickel resin beads were prepared using the QX4 loading dye as per Table 2.12. A sample of the pellet was resuspended in Milli-Q® Ultrapure water (500 µL). Gel samples were heated at 90 °C for 10 minutes.

Table 2.12. Preparation of gel samples for proteins isolated during small scale expression

Reagents	Supernatant (µL)	Pellet resuspension (µL)	Supernatant washed-nickel beads (µL)
Sample	10	5	20
QX4	10	20	30
Milli-Q® Ultrapure water	20	15	-
Total volume	**40**	**40**	**50**

Gel samples (10 µL) were run on SDS-PAGE gels composed of a 5% stacking gel layer on top of a 12% resolving gel layer (Table 2.6). A Precision Plus Protein^{TM} (PPP) All Blue Prestained standards ladder (7 µL, BioRad, Figure D.2) was loaded with every run. Gels were visualised using the Quick Coomassie Blue Staining method (Kurien & Scofield, 1999) using Fairbanks A solution (Coomassie® Brilliant Blue R-250 (0.05%, PanReac AppliChem), isopropanol (25%) and acetic acid (10%)) which was microwaved with the gel (30 secs) before shaking (15 min).

Destaining was performed using the Fairbanks Destaining solution (acetic acid (10%)) which was microwaved with the stained gel (30 secs) before shaking for 30 minutes. The gels were visualised using the iBright FL1000 Imaging System (ThermoFisher Scientific).

2.2.4 Large-scale expression of recombinant Ngo-Lig proteins

After identification of the optimal expression conditions for each protein (overnight at 15 °C, overnight at 20 °C or for four hours at 37 °C), expression was repeated on a larger scale (1 L). Seeder cultures (30 mL) were prepared using the 25% glycerol stocks of BL21(DE3)pLysS cells in 1% glucose TB media with the desired antibiotics, which were incubated overnight (37 °C, 180-200 rpm). The seeder cultures (30 mL) were used to inoculate 1% glucose TB media (1 L) that were subsequently incubated at 37 °C (180-200 rpm) until the OD_{600} measurement reached 0.3-0.4, before incubation of the cultures at the chosen temperature (either 15 or 20 °C) for one hour. IPTG (0.5 mM) was added and incubation at the chosen temperature was continued overnight. The cultures were centrifuged (4600 rpm, 30 min, 4 °C) to isolate the pellets, which were weighed and frozen at -20 °C.

2.2.5 Protein purification

Protein purification was performed on either the GE Äkta Prime system at 4 °C, or on the Next-Generation Chromatography™ (NGC™) Chromatography system (Bio-Rad) at room temperature. The typical protein purification workflow included an IMAC (immobilised metal affinity chromatography) via a nickel pull-down assay, desalting or buffer exchange, an overnight Tobacco Etch Virus (TEV) protease incubation, a reverse nickel pull-down assay and size exclusion or gel filtration chromatography. The buffers of interest for each procedure are shown in Table 2.13.

Table 2.13. Composition of buffers used for Ngo-Lig protein purification

	Buffer A	Buffer B	Buffer C (low salt)	Buffer C (high salt)	MBP binding buffer	MBP elution buffer
Tris pH 8.0	50 mM	50 mM	50 mM	50 mM	-	-
Tris pH 7.4	-	-	-	-	20 mM	20 mM
NaCl	750 mM	750 mM	100 mM	500 mM	200 mM	200 mM
Imidazole	10 mM	500 mM	-	-	-	-
EDTA	-	-	-	-	1 mM	1 mM
Glycerol	5%	5%	5%	5%	-	-
Dithiothreitol (DTT) – added on the day of use	-	-	1 mM	1 mM	1 mM	1 mM

The frozen cell pellets were resuspended in lysis buffer (5 mL per gram of pellet) and sonicated using a 6 mm probe (Table 2.11) before centrifugation (9000 rpm, 60 min, 4 °C). Gel samples were prepared for both the supernatant and pellet samples as described for small-scale expression (Table 2.12) and run on a 12% SDS-PAGE gel. The supernatant was incubated overnight with ATP (0.1 mM, Sigma) and Benzonase® nuclease (1 µL per gram of pellet, Novagen®) while shaking (4 °C). For each 50 mL of supernatant, 1 mL of buffer B (Table 2.13) was added (10 mM final imidazole concentration) before filtration using a 0.3 µm syringe filter.

2.2.5.1 Nickel pull-down

Nickel pull-down was performed using either a 1 mL (for pellet from a 1 L culture) or a 5 mL (for pellet from three combined 1 L cultures) HisTrap™ HP nickel column (GE Healthcare). The filtered protein lysate was either loaded onto a 50 mL Superloop™ (GE Healthcare) backfilled with buffer A (Table 2.13), which was connected to the HisTrap™ HP column on the GE ÄKTA Prime system, or flowed onto the HisTrap™ HP column via the sample pump on the NGC™ system. The

protein was injected at 2.5 mL/min (maximum pressure: 0.5 MPa) and the flow-through fractions (4 mL) were collected before elution of the protein from the nickel column using a 50 mL gradient of buffer B (Table 2.13) (0-100% at 2.5 mL/min) with a fraction collection size of 2 mL. Gel samples of the fractions collected were prepared (15 µL fraction/sample, 5 µL QX4 loading dye) before analysis via 12% SDS-PAGE gels to identify the protein-containing fractions.

2.2.5.2 Desalt, TEV cleavage and reverse nickel pull-down

Protein of interest containing fractions identified were pooled and concentrated using Amicon® Ultra Centrifugal Filters (10 kDa) via centrifugation (3600 rpm, 4 °C) until a final volume of 5 mL was reached. The concentrated protein was either loaded onto the 50 mL Superloop™ (GE Healthcare) backfilled with buffer C (low salt) (Table 2.13), which was connected to a HiPrep™ 26/10 Desalting column (Cytiva) on the GE ÄKTA Prime system, or flowed onto the HiPrep™ 26/10 Desalting column via direct injection onto a 5 mL loop on the NGC™ system. The protein was run at 3 mL/min (maximum pressure: 0.15 MPa) with a 2 mL fraction collection size.

Fractions identified via the A280 peak were pooled together. Gel samples were prepared before the addition of TEV protease (0.1 mg/mL). The protein samples were aliquoted into 2 mL tubes (1 mL aliquots) and left at room temperature for one hour before shaking overnight at 4 °C. The aliquoted protein was centrifuged (13500 rpm, 10 min, 4 °C) and gel samples were prepared for both the pooled supernatant and any precipitate or pellet that may have formed. The gel samples were run on a 12% SDS-PAGE gel.

The pooled supernatant was injected onto the same HisTrap™ HP column from section 2.2.5.1 using the Superloop™ (backfilled with buffer C)) via the GE ÄKTA Prime system at 2 mL/min (1 mL fraction collection) before elution with a 20 mL gradient of buffer B (0-100%) at 2.5 mL/min. Gel samples were prepared for the fractions collected and run on a 12% SDS-PAGE gel to identify protein-containing fractions.

2.2.5.3 Gel filtration or size exclusion and securement of protein stocks

The pooled protein fractions containing the desired protein were concentrated using the Amicon® Ultra Centrifugal Filter (10 kDa) until a final volume of 2 mL or 5 mL was reached. The protein was loaded onto a HiLoad® 16/600 Superdex® 75 (Cytiva) column either through a standard 2 mL loop on the GE ÄKTA Prime system at 0.5 mL/min (maximum pressure: 0.5 MPa) with 2 mL fraction collection using buffer C (either low salt or high salt (Table 2.13), or through a standard 5 mL loop on the NGC^{TM} system. Gel samples of the fractions collected were prepared and run on a 12% SDS-PAGE gel to identify protein-containing fractions.

Protein-containing fractions confirmed via 12% SDS-PAGE gel runs were concentrated using the Amicon® Ultra Centrifugal Filters (10 kDa) via centrifugation until a final volume of 1 mL was reached. Nanodrop^{TM} readings (A_{280}) were recorded. Half the concentrated protein was stored with equal amounts of 50% glycerol and snap frozen via liquid nitrogen in 200 µL aliquots. The other half was directly snap frozen via liquid nitrogen in 100 µL aliquots. The protein aliquots were stored at -80 °C.

2.2.5.4 Additional purification procedures

Cation exchange was performed after size exclusion for one protein purification (Lig E-His) to separate two proteins that were eluting simultaneously, using a HiTrap® SP HP column (Cytiva). A 1:3 dilution was performed on the pooled protein with 2-(*N*-morpholino)ethanesulfonic acid (MES) pH 6.15 buffer to lower the protein pH and salt content. This was loaded onto the HiTrap® SP HP column using the sample pump on the NGC^{TM}, which was washed with the MES buffer. The column was eluted using a gradient of NaCl (1 M) at 2 mL/min with a fraction collection size of 1 mL. Gel samples were prepared of the fractions collected and verified via a 12% SDS-PAGE gel.

A reverse MBP (maltose binding protein)-trap was performed after a size exclusion for one protein purification (Lig E-His) to separate the uncleaved protein and the cleaved MBP-tag, both containing an MBP-tag from the cleaved protein. A 1:1 dilution was performed on the pooled protein with MBP binding buffer (Table 2.13),

which was filtered using a 0.3 µm syringe filter before loading onto a 50 mL Superloop™ (GE Healthcare), backfilled with MBP binding buffer (Table 2.13). This was connected to an MBPTrap™ HP (Cytiva) (5 mL) column on the GE ÄKTA Prime system. The protein was injected onto the column at 2 mL/min (1 mL fraction collection) before elution via a 25 mL gradient of MBP elution buffer (Table 2.13) (0-100%) at 2.5 mL/min. Gel samples were prepared for the fractions collected and run on a 12% SDS-PAGE gel to identify protein-containing fractions.

2.3 *In vitro* assay of recombinant Ngo-Lig

2.3.1 Determining Ngo-Lig protein concentration

Protein concentration was determined via the Bradford assay method with the Bio-Rad Protein Assay Dye Reagent Concentrate that had been diluted (1:4) with Milli-Q® Ultrapure water (Bradford, 1976). Standards of bovine serum albumin (BSA) protein (1.40, 0.75, 1.00, 0.50 and 0.25 mg/mL) were prepared via serial dilution in buffer C (low salt) (Table 2.13) before mixing the standards (8.3 µL) in diluted Protein Assay Dye Reagent Concentrate (241.7 µL) on a shaker (30 secs). This was performed in triplicate. The mixtures were left at room temperature for 15 minutes before measurement of the absorbance (595 nm) on the Multiskan G0 spectrophotometer. The frozen proteins (25% glycerol stocks) (Lig E-GFP, Lig E, Lig E-His) were subjected to the same treatment to record the absorbance at 595 nm, before comparison to the BSA standard curve to determine the protein concentrations.

2.3.2 Ngo-Lig ligation assay of dsDNA

Based on the protein concentrations calculated for each Lig E variant, the protein aliquots were diluted using buffer C to the lowest concentration (906 nM; Lig E-His). Serial dilution was performed for all three proteins using buffer C. The master mix (Table 2.14) for the ligation assay (with a singly nicked duplex DNA) was prepared before incubation at 90 °C (5 min). For each dilution, 2.5 µL of the protein was added to 22.5 µL of the master mix in triplicates, which was incubated at 25 °C (30 min). The mixture was heated at 95 °C (5 min) before the addition of the quench buffer (5 µL) (95% formamide (9.5 mL), EDTA (100 µL), bromophenol blue (100

µL)) and subsequent heat inactivation at 95 °C (5 min). Samples (10 µL) were loaded onto 20% urea gels (see section 2.3.2.1) which were run at 5 mA per gel in TBE (tris-borate-EDTA) buffer (Tris (89 mM), boric acid (89 mM), EDTA (2 mM)). Fluorescence of the urea gels was visualised using the iBright FL1000 Imaging System (ThermoFisher Scientific) using the 'Fluorescent blots' and Fluorescein' settings and the bands were quantified using ImageJ (Abràmoff *et al.*, 2004).

Table 2.14. Composition of ligation gel-based assay master mix (for 50 individual ligation reactions) for a singly nicked dsDNA assay, with ATP added after cooling of the master mix. Sequences of the ligation constructs can be found in Table B.3.

Reagents	Amount (µL)
Ligation construct 1 (L1, 0.5 µM)	200
Ligation construct 2 (L2, 2.5 µM)	200
Ligation construct 3 (L3, 2.5 µM)	200
$MgCl_2$ (1 M)	12.5
10xT4 buffer (Tris pH 8.0 (500 mM), NaCl (500 mM), DTT (100 mM))	125
Milli-Q® Ultrapure	375
ATP (100 mM)	12.5
Total	**1125.0**

The ligation assay was performed with variations in the pH of the 10x T4 buffer (using Tris buffer for pH 7.1-9.0 or MES buffer for pH 5.5-6.2) to determine the optimal pH for Ngo-Lig ligation, and with variations in the concentrations of ATP. Different combinations of the ligation constructs (L1 and L2) with different complement strands were also prepared to determine the optimal ligation substrate for Ngo-Lig (Table 2.15). Samples of these were treated and analysed as above, with an overnight incubation at 15 °C for blunt-ended and gapped substrates as opposed to the 25 °C incubation (30 min) for single nick, overhang and mismatch substrates.

Table 2.15. Combinations of constructs used to make different dsDNA substrates for gel-based ligation assays. Sequences of the ligation constructs can be found in Table B.3.

Substrates	Ligatable constructs	Complementary constructs
Single nick	L1, L2	L3
Blunt-ended	L1, L2	L6, L7
Overhang	L1, L2	L8, L9
Mismatch	L1, L2	L10
Gapped	L1, L2	L11

2.3.2.1 Urea gels

To prepare two 20% urea gels, urea (8.4 g) (Sigma-Aldrich) was stirred and heated (50 °C) with 10xTBE (2 mL) and 40% acrylamide (10 mL) (Bio-Rad), made up to 20 mL with Milli-Q® Ultrapure water. APS (200 µL) and TEMED (6 µL) were added before pouring the mixture into a gel-setting cast. Urea gels were used within the week of production.

2.4 Construction of a reporter construct for *N. gonorrhoeae* DNA uptake

2.4.1 Cloning of the reporter constructs of interest

Two constructs were used for DNA uptake experiments, one with a *β-lactamase* (*β-lac*) gene (Figure C.2 (a)) and one with a *GFP* gene (Figure C.3 (a)), both of which were *N. gonorrhoeae* optimised (Goodman & Scocca, 1988; Dillard, 2011). The *β-lac* gene construct was ordered from Invitrogen™ with EcoRI and XhoI overhangs on the 5' and 3' ends respectively, while the *GFP* gene construct was ordered from Twist Bioscience and was used in prior MS11 transformations (section 2.1.2). The GFP construct was amplified via PCR with the GFP_amp_Fd and GFP_amp_Rev primers (Table C.1) ($T_a = 54$ °C), which contained EcoRI and XhoI overhangs to add to the gene during PCR. PCR was run as per Table 2.4 before visualisation via 1% agarose under UV.

2.4.1.1 *N. gonorrhoeae* **optimised pMR32 plasmid manipulation**

To optimise uptake of the reporter genes into *N. gonorrhoeae* (between the *trpB* and *igA* genes), both constructs were cloned into the pMR32 plasmid (Figure C.1) (Ramsey *et al.*, 2012). *E. coli* cells (DH5α) expressing the pMR32 plasmid was gifted to Dr. Joanna Hicks by Professor Joseph Dillard.

Seeder cultures of pMR32 expressing DH5α cells were prepared by inoculating kanamycin LB media (5 mL) with the glycerol stock (10 µL), which was incubated overnight (37 °C, 180 rpm) before isolation, quantification and storage of the pMR32 plasmid (see section 2.2.1.2).

2.4.1.2 **Preparing gonococcal DNA uptake reporter constructs (constructs 'A')**

2.4.1.2.1 *Restriction digest and purification*

Restriction digest was performed for both pMR32 and the *β-lac* and *GFP* constructs using the XhoI and EcoRI enzymes (Table 2.16) to insert the reporter constructs into the pMR32 plasmid. This was followed by standard PCR purification and quantification processes as per section 2.1.2.3.

Table 2.16. Components of the restriction digest mixture prepared for each gonococcal DNA uptake reporter construct. The restriction digest enzymes were added last for each mixture.

Reagents	pMR32 plasmid digestion mixture	Construct A (β-lactamase) digestion mixture	Construct B (GFP) digestion mixture
DNA	1 µg	0.75 µg	0.75 µg
XhoI (Roche, 10 U/µL)	1 µL (10 units)	1 µL (10 units)	1 µL (10 units)
EcoRI (NEB, 20000 U/mL)	1 µL (20 units)	1 µL (10 units)	1 µL (10 units)
10x CutSmart^{TM} buffer (NEB)	5 µL	5 µL	5 µL
UltraPure^{TM} DNase/RNase-free distilled water (Invitrogen^{TM})	Up to 50 µL final volume	Up to 50 µL final volume	Up to 50 µL final volume

Ligation of the purified digested constructs was performed using a 1:3 ratio of vector: insert (30 fmol:90 fmol) to form the pMR32:β-lac (Figure C.2 (a)) and pMR32:GFP plasmids (Figure C.3 (a)) based on the expected sizes of 4710 bp for pMR32, 852 bp for the *β-lac* construct and 738 bp for the *GFP* constructs (Equation 2.1 and Equation 2.2).

The ligation mixtures (10 µL) were transformed onto 100 µL of chemically competent DH5α *E. coli* cells as per section 2.2.1.1 before isolation of the ligated plasmids using the QIAprep Spin Miniprep Kit (Qiagen) according to the manufacturer's protocol. The empty pMR32, pMR32:β-lac and pMR32:GFP plasmids were treated in three ways to obtain linear DNA including linearisation using the SphI restriction enzyme (NEB, 10000 U/mL). PCR was also performed using the Fd_whole_substrate and Rev_whole_substrate primers (Table C.1) (T_a = 47 °C) to yield constructs 'A' (Figure C.1 (a), Figure C.2 (b) and Figure C.3 (b) respectively), and the Fd_constructmaking_longerpMR32 and Rev_constructmaking_longerpMR32 primers (Table C.1) (T_a = 51 °C) to yield

longer constructs 'A' (Figure C.2 (c) and Figure C.3 (c)) before visualisation under UV after a 1% agarose run. The bands of interest were removed and gel-purified before storage at -20 °C.

2.4.2 Verification of the integrity of the DNA uptake reporter constructs

To verify the integrity of constructs 'A', and the plasmids themselves, the constructs were sent for Sanger sequencing. Based on the Nanodrop™ concentrations, the respective volumes of PCR products required were calculated to send for sequencing (2.5 ng of DNA for every 100 bp along with either the Fd_whole_substrate, Rev_whole_substrate, Fd_new_ermC_seq or Rev_new_ermC_sew primers (4 pmol)). The sequencing mixtures were made up to 20 µL with UltraPure™ DNase/RNase-free distilled water (Invitrogen™) and sent to Massey Genome Service for sequencing before analysis on Geneious Prime (v. 2020.1) (Kearse *et al.*, 2012).

2.4.3 *N. gonorrhoeae* transformations of the DNA uptake reporter constructs

Spot transformation was performed as per section 2.1.2.1 with different variants of the pMR32, pMR32:β-lac and pMR32:GFP constructs. These included the linearised plasmids (with SphI), constructs 'A' (amplified with the Fd_whole_susbtrate and Rev_whole_substrate primers) and the longer constructs 'A' (amplified with the Fd_constructmaking_longerpMR32 and Rev_constructmaking_longerpMR32 primers), which were later grown on erythromycin GCB plates. Colonies were verified using the specific primers outlined in Table 2.17.

Table 2.17. Primer sets used for verification of pMR32 transformations in *N. gonorrhoeae* via colony PCRs

Construct	Forward primer	Reverse primer	T_a (°C)
Linearised plasmid	Fd_sequencingcheck _longerpMR32	Rev_sequencingcheck _longerpMR32	46.9
Constructs 'A'	Fd_seq_DNA_uptake	Rev_seq_DNA_uptake	52.2
Longer constructs 'A'	Fd_sequencingcheck _longerpMR32	Rev_sequencingcheck _longerpMR32	46.9

Chapter Three

Importance and location of Ngo-Lig

3.1 Introduction

The prediction of a signal peptide on the N-terminus of Lig E (Figure 3.1) posed a question about its function in relation to its location. Analysis of the signal peptide sequence indicates that Lig E is potentially periplasmic, which is interesting considering the cytosolic location of other ADLs (Magnet & Blanchard, 2004). The question of Lig E's function then evolved into one regarding the origin of its substrates for ligation, as bacterial genomic DNA is contained within the cytoplasm, which leaves environmental or extracellular pieces of DNA as possible substrates for Lig E.

Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.9945	0.0014	0.0022	0.002

Figure 3.1. Predicted signal peptide (24 amino acids) of Lig E from MS11 *N. gonorrhoeae* (Ngo-Lig) via the SignalP-5.0 prediction programme, with a predicted 0.99 probability for export via the secretory (Sec) pathway (Nielsen, 2017). The red line indicates the likely signal peptide, the green line indicates the likely cleavage site, and the orange line represents the rest of Ngo-Lig protein sequence.

It would be interesting then, to investigate this putative periplasmic location of Lig E to infer its possible function. Owing to their oxidative environments, the majority

49

of Gram-negative bacterial proteins with disulphide bonds tend to mature in the periplasm (Majoul *et al.*, 1997; Karyolaimos *et al.*, 2019). Analyses of the structure of Lig E from *V. cholerae* and *A. mediterranea* indicate the absence of cysteine, and hence disulphide bonds, in the proteins although the environment of *V. cholerae* in brackish water and *A. mediterranea* in its marine habitat may not be a suitable environment for this bond formation (Majoul *et al.*, 1997; Williamson *et al.*, 2018). However, protein modelling of Lig E from *C. jejuni* and both Nme-Lig and Ngo-Lig indicate at possible disulphide bonds within the protein, which is consistent with their oxidative habitats in the human host (Figure 3.2). Regardless, the signal peptide for periplasmic proteins would be removed by peptidases as it passes through a translocon (Dalbey, 1991; Denks *et al.*, 2014; Karyolaimos *et al.*, 2019). For Lig E, removal of the signal peptide enhances its activity, strongly indicating that it exerts its effects in the periplasm (Magnet & Blanchard, 2004; Williamson & Pedersen, 2014). There are two possible methods of transportation across the membrane; the general secretory (Sec) translocon, and the twin-arginine translocation (Tat) pathway (Denks *et al.*, 2014; Freudl, 2018). Based on the prediction by SignalIP (Figure 3.1), Ngo-Lig highly likely translocates via the Sec pathway in an unfolded nature, again hinting at its possible maturity in the periplasmic space.
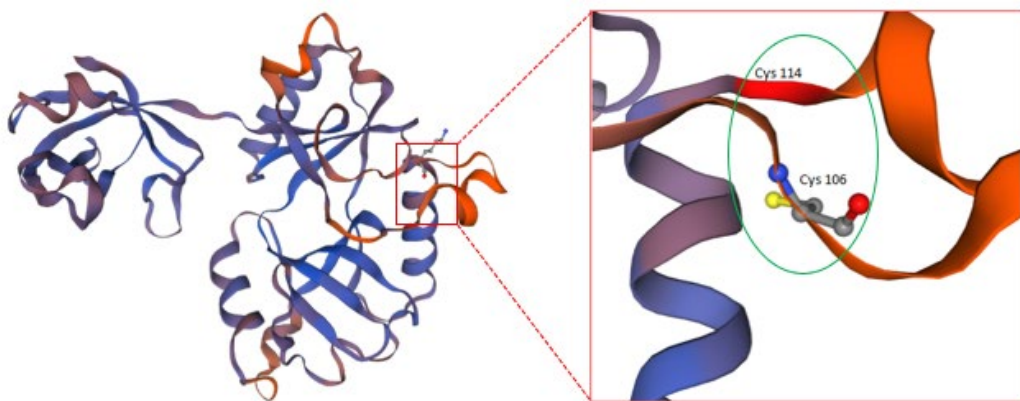
Figure 3.2. Homology model of the Lig E structure from *C. jejuni* aligned to Ame-Lig (6GDR) created via SWISS-MODEL (Waterhouse *et al.*, 2018). Additional features found in this protein are shown in the red square, which may be a candidate location for disulphide bond formation, indicating possible periplasmic maturation and location of this protein.

Analysis of recombinant Lig E from different species by various groups have revealed its ability to seal single-stranded nicks (Magnet & Blanchard, 2004; Williamson & Pedersen, 2014; Williamson *et al.*, 2018). This, in conjunction with its hypothesised periplasmic location, suggests that Lig E could be involved in an alternative pathway for increasing the information content of exDNA, thus potentially having a role in DNA competence (Magnet & Blanchard, 2004). However, no *in vivo* work has been conducted on Lig E. Hence, the aim of this chapter was to create mutants of the naturally competent *N. gonorrhoeae* for future *in vivo* investigations of the location and function of Ngo-Lig. As part of this chapter, the generation of three different *N. gonorrhoeae* Ngo-Lig mutants was attempted via spot transformation. This was performed using synthetic DNA constructs that contained DUS sequences for optimised gonococcal uptake, and flanking regions (150 bp) with homology to the points of insertion into the *N. gonorrhoeae* genome (Chen & Gotschlich, 2001). The first was a Lig E-GFP mutant (Figure A.2) with a GFP tag on the C-terminus of the gene to track the cellular location of Ngo-Lig via fluorescent microscopy. The second was a Lig E-KO-KanaR mutant (Figure A.4), whereby the Ngo-Lig gene would be interrupted by a kanamycin resistance gene to determine the importance of Ngo-Lig on cell viability. The third was a Lig E-His-KanaR mutant with a hexahistidine-tag on the C-terminus of Ngo-Lig (Figure A.3), as well as a kanamycin resistance gene to not only act as a control for its presence in the Lig E-KO-KanaR mutant, but to also serve as an alternate pathway for investigating the potential periplasmic location of Ngo-Lig via subcellular fractionation (Malherbe *et al.*, 2019). Of the three, only the Lig E-His-KanaR and Lig E-KO-KanaR mutants were successful, which were subjected to a growth experiment to investigate the effects of any changes made on the viability of gonococcal cells.

3.2 Results and discussion

3.2.1 Generation of Ngo-Lig mutants *in vivo*

3.2.1.1 Lig E-His-KanaR and Lig E-KO KanaR mutants

The colony PCR results for both the Lig E-His-KanaR (Figure 3.3 (a)) and Lig E-KO-KanaR (Figure 3.3 (b)) mutants amplified with the Primer_4_His_FD and

Primer_2_Rev primers, and the Primer 1_KO_FD and Primer_2_Rev primers respectively show successful transformation of both constructs into the WT *N. gonorrhoeae*. Based on the primer sets used, the expected product sizes were 1050 bp for the Lig E-His-KanaR mutants (0 bp for the WT) and 1296 bp for the Lig E-KO-KanaR mutants (1115 bp for the WT), which were observed in the respective gels.
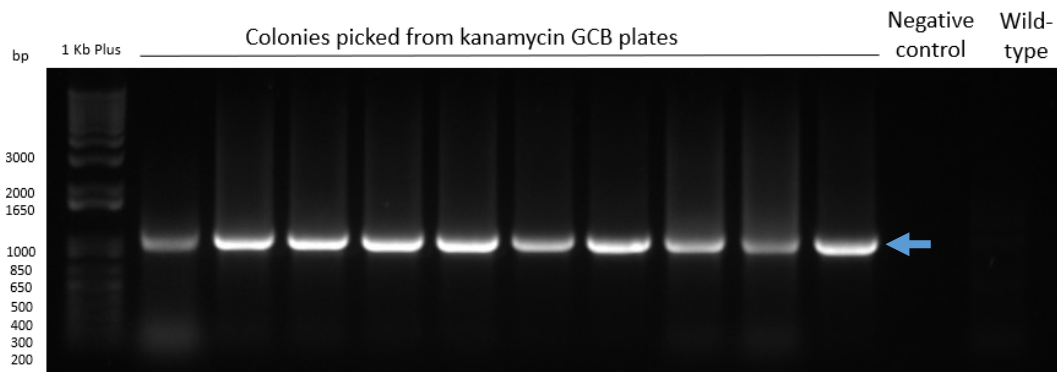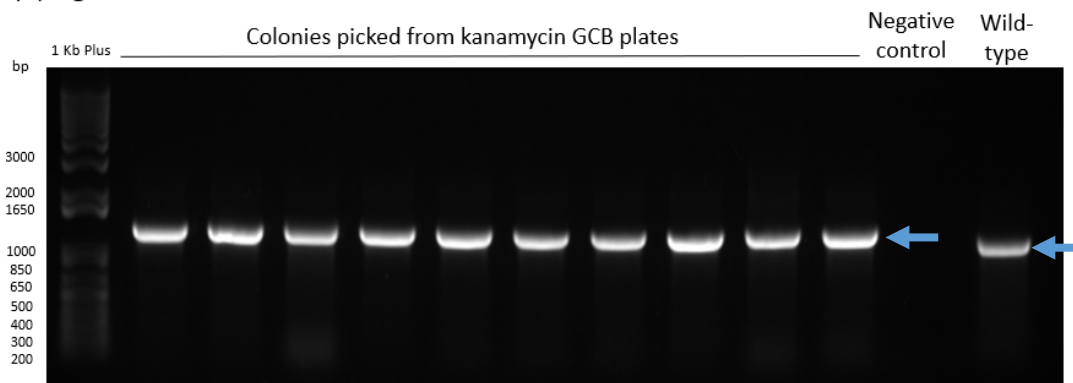
Figure 3.3. Agarose gels (1%) of colony PCR samples after MS11 *N. gonorrhoeae* transformations of (a) Lig E-His-KanaR and (b) Lig E-KO-KanaR. Colonies were picked from kanamycin GCB plates. The blue arrows indicate the bands of interest. Sizes of the 1 Kb Plus ladder used are labelled.

The correct integration of both Lig E-His-KanaR and Lig E-KO-KanaR constructs were verified via Sanger sequencing (Figure 3.4 (a) and (b)). For both mutants, the forward and reverse sequences aligned with their predicted reference sequences and these changes were absent in the WT *N. gonorrhoeae* sequence (Figure 3.4 (c)). Despite the presence of some poor-quality sequencing towards the ends of both forward and reverse sequences (visualised via red boxes), we were confident that

the integration of both constructs had occurred in the right location in relation to the Ngo-Lig gene.

(a) Lig E-His-KanaR
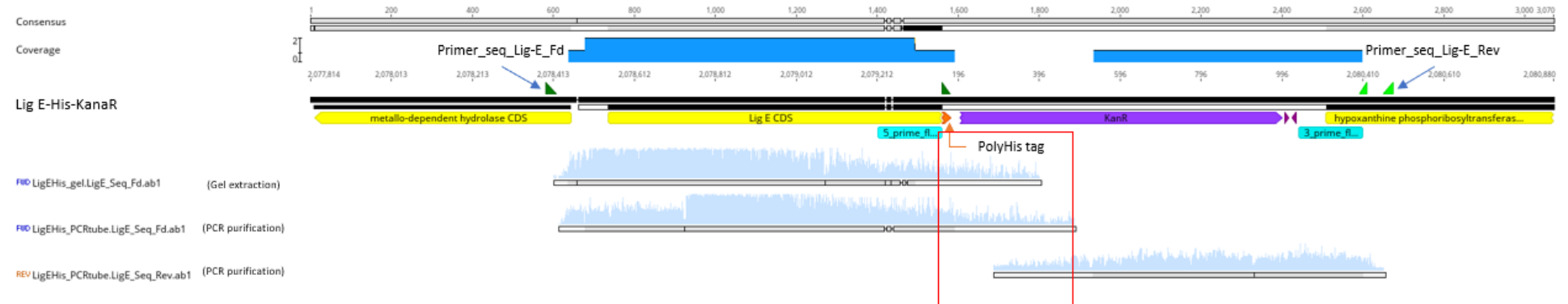
(b) Lig E-KO-KanaR
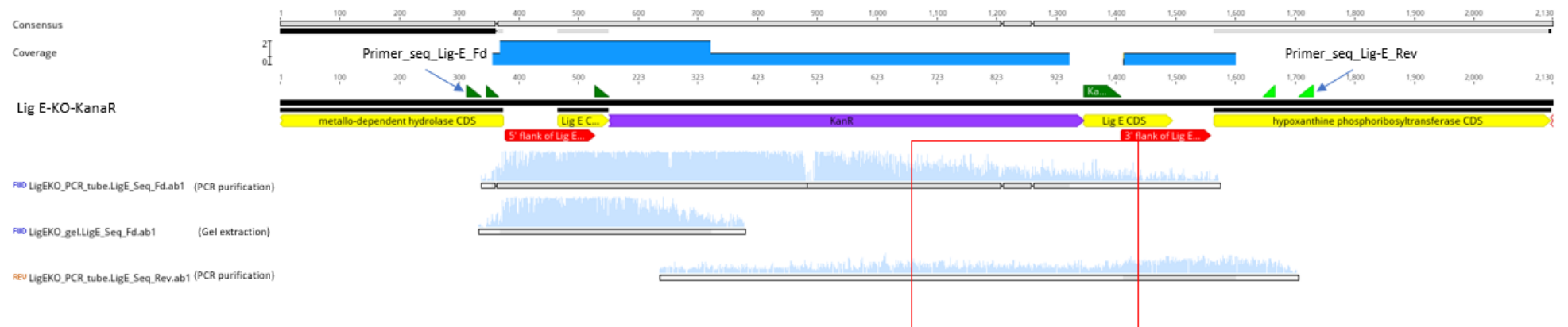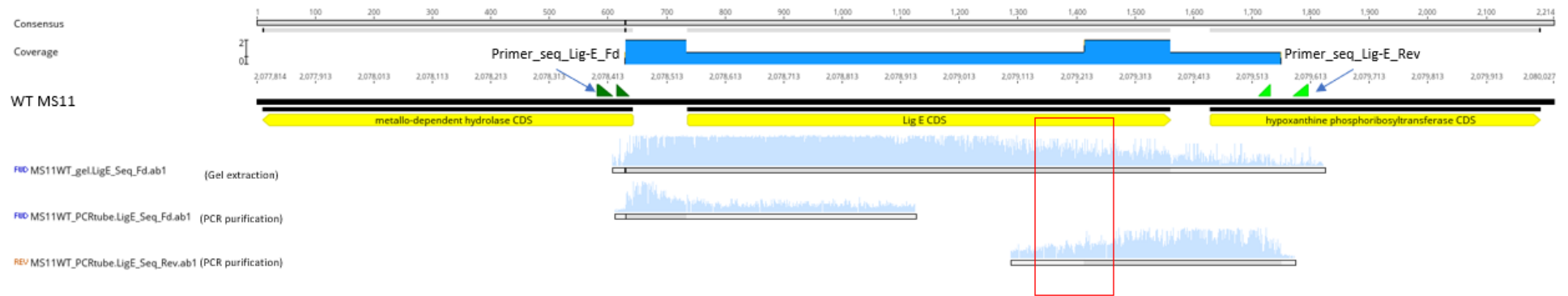
Figure 3.4. Alignment of the intended MS11 *N. gonorrhoeae* **mutant sequences constructed** *in silico* **via Geneious (top for each), aligned with the forward and reverse Sanger sequencing results from the selected mutants (bottom). Both forward and reverse sequences were gel- and PCR-purified to determine which method led to the best sequencing results, although the reverse gel-purified sequences were of poor quality for alignment and were discarded (a) Lig E-His-KanaR mutants (b) Lig E-KO-KanaR mutants (c) Wild-type (WT) MS11. Primers used for sequencing are indicated by the blue arrows. The red boxes represent areas of uncertainty in the sequencing. Fd = forward sequences sequenced with the Primer_seq_LigE_Fd primer, Rev = reverse sequences sequenced with the Primer_seq_LigE_Rev primer.**

.

The successful knock-out of *Lig E* in *N. gonorrhoeae* here strongly indicates that this gene is non-essential for gonococcal survival, as opposed to NDLs like *lig A* in *H. influenzae*, the knock-out of which leads to death of the cells (Preston *et al.*, 1996). Thus, the question of the exact role of Lig E remains. To answer this question *in vivo*, it is essential to first control for any effects of the transformation process performed that may affect results in further experiments via a growth experiment.

3.2.1.2 Lig E-GFP mutants

With transformation of the Lig E-GFP construct, the initial plan for screening of successful gonococcal transformants was to screen for fluorescent colonies, which would also allow for direct determination of Ngo-Lig's cellular location via microscopy, without the extensive western blotting process (Phillips, 2001). However, this method relies on high expression of the GFP protein and would exclude any successful transformants that expresses the protein at low levels. For this reason, colony PCR was also performed after transformation using the Primer_3_GFP_FD and Primer_2_Rev primers. The absence of the expected 872 bp band (Figure 3.5) illustrate the lack of successful transformation. Rather, PCR of these colonies resulted in the same bands as that observed in the WT MS11 *N. gonorrhoeae*, with consistent double bands between 1000 and 1650 bp. However, Primer_3_GFP_FD only binds to the 3' end of the *GFP* gene, which is absent in the WT MS11 genome, so should not result in any amplification in the WT colonies. The absence of any bands in the negative controls rules out any issue with contamination, indicating that non-specific amplifications may have occurred. These non-specific primings are particularly common for repeat sequences and regions with high percentage of GC bases, which sits at 52.43% for the gonococcal genome, making this likely (Kieleczawa, 2006; Hommelsheim *et al.*, 2014; Lu *et al.*, 2019). For example, the 3' sequences of the Primer_3_GFP_FD primer (AACTTTTCACT<u>GGA</u><u>GTT</u><u>GTCCCA</u>) may have bound through strong GC pairings to the upstream metallo-dependent hydrolase gene via the bases present in the gene in the same order (underlined above). This, in combination with the reverse primer, would yield a 1153 bp band, observed in Figure 3.5.
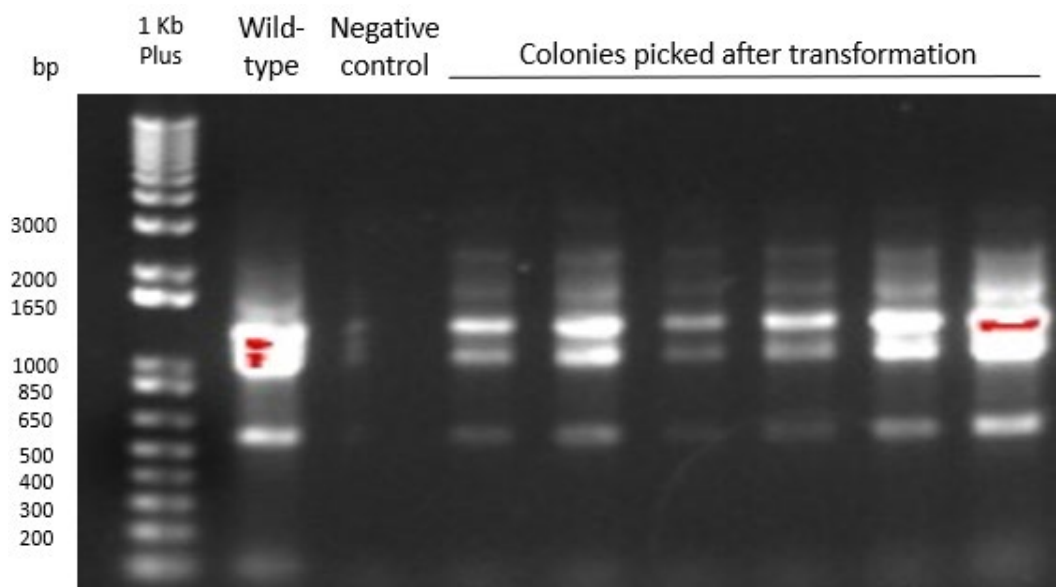
Figure 3.5. Agarose gel (1%) of colony PCR samples after MS11 *N. gonorrhoeae* transformations of the Lig E-GFP construct. Colonies were picked from GCB plates without antibiotics. Sizes of the 1 Kb Plus ladder used are labelled.

It should be noted however, that the main reason for the minimal Lig E-GFP transformation detection is the lack of an antibiotic resistance gene as a selective marker. This selection gene was omitted as the construct was already large and transformants were planned to be selected for via fluorescence. Unfortunately, this led to complications in terms of background WT MS11 colonies that may have been screened instead of the transformants. However, it is also possible that the *GFP* gene is deleterious to the bacterium. Owing to its large size (28 kDa), various groups have noted difficulty of the GFP protein to fold properly in the periplasm of *E. coli*, as multiple GFP-hybrid proteins with a signal peptide (for unfolded export via the Sec pathway) localised in the cytoplasm instead (Feilmeier *et al.*, 2000; Santini *et al.*, 2001; Thomas *et al.*, 2001; Seitz *et al.*, 2014). In contrast, export of folded GFP via the Tat pathway has been successful (Santini *et al.*, 2001; Thomas *et al.*, 2001; Seitz *et al.*, 2014). Based on the prediction by SignalIP, Ngo-Lig is highly likely to be transported via the Sec pathway in an unfolded state (Figure 3.1), suggesting difficulty of GFP to be used for Ngo-Lig tracking with its specific leader sequence. Possible alternatives to this include splitting the GFP protein between the cytoplasm and the periplasm to make translocation easier which can also be used for later interaction studies (Cabantous *et al.*, 2005; Cabantous & Waldo, 2006). Alternatively, a less complicated method is the use of a different fluorescent protein

like mCherry, which although it has a lower fluorescent intensity than GFP, has been proven to work in *Neisseria spp.* for periplasmic proteins (Toddo *et al.*, 2012; Imhaus & Duménil, 2014; Seitz *et al.*, 2014; Pena *et al.*, 2020). However, as the inclusion of the kanamycin resistance gene was highly successful in the other two transformations, the first step would be the addition of this resistance gene in the *GFP* construct to allow for more robust gonococcal transformation selection.

3.2.2 Role of Ngo-Lig on cell viability

After the generation of Ngo-Lig mutants in *N. gonorrhoeae*, a 63-hour growth experiment of the WT MS11 and the Lig E-His-KanaR and Lig E-KO-KanaR mutants was performed. This was done for three reasons: i) to check if Ngo-Lig is essential for normal growth, ii) to control for the effects of both *N. gonorrhoeae* transformations and the kanamycin selection marker and iii) to obtain gonococcal samples for further analysis. During this experiment, the OD_{600} readings were measured every two hours where possible. With every second OD_{600} measurement, CFU counts were recorded to determine whether any increase in OD_{600} readings reflected the number of viable cells. This was alternated with supernatant and pellet isolation for further DNA and protein analyses. To ensure that each of the replicate cultures started with similar cell densities, triplicate seeders (Table A.2) were first prepared before the generation of the growth cultures.

3.2.2.1 N. gonorrhoeae growth experiment: OD_{600} measurements

The OD_{600} readings from the growth experiment with the WT MS11, Lig E-His-KanaR and Lig E-KO-KanaR mutants (Figure 3.6 (a)(b)(c) respectively) each generated distinct growth curves. All three retained the features of a general growth curve with four distinct phases (Hall *et al.*, 2013; Wang *et al.*, 2015): a lag phase (around 0-15 hours) where the OD_{600} readings recorded were below the range of detection as the bacteria started to grow, an exponential or log phase (approximately 15-43 hours) where the OD_{600} readings increased exponentially, a stationary phase (around 43-49 hours) where the OD_{600} readings plateaued as nutrients became a limiting factor, and a subsequent death or lysis phase (up to 63 hours) where the OD_{600} readings decreased as cells started to lyse. It is interesting to note however, that for both the Lig E-His-KanaR and Lig E-KO-KanaR mutants, the growth curve

of one of the three replicates deviated from the others (culture C for Lig E-His-KanaR and culture A for Lig E-KO-KanaR), while all three replicates for the WT MS11 were similar. The resultant average OD_{600} readings taking all three replicates into consideration (Figure A.5) was affected by these deviants, resulting in large standard errors for both the Lig E-His-KanaR and Lig E-KO-KanaR mutants.
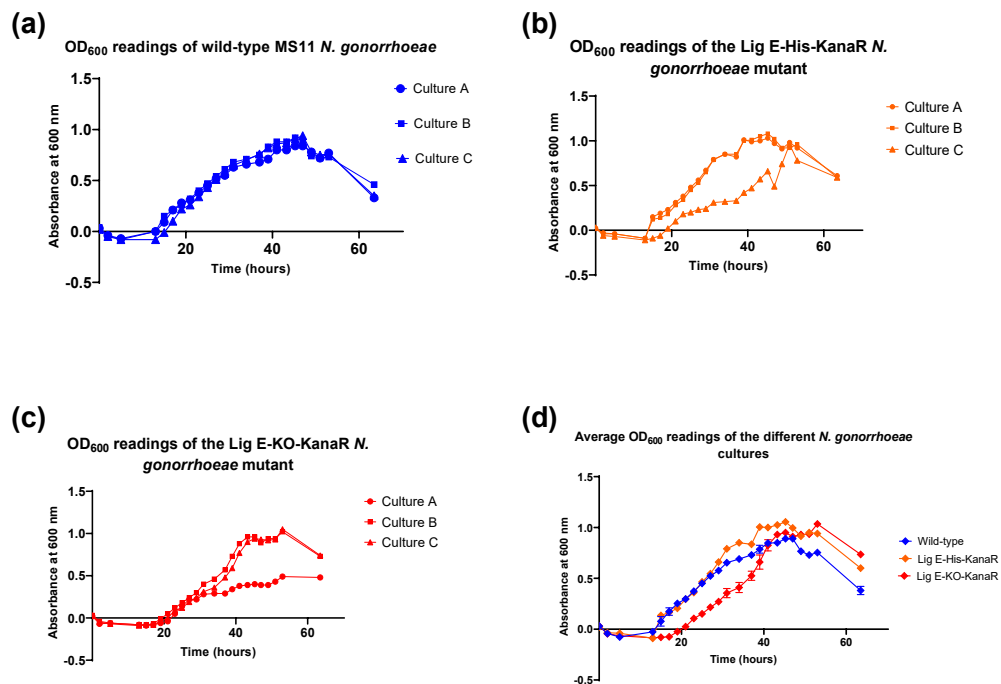
Figure 3.6. OD_{600} readings of each MS11 *N. gonorrhoeae* culture from the growth experiment with three biological replicates and standard errors plotted (a) Replicates of the wild-type MS11 (b) Replicates of the Lig E-His-KanaR mutant (c) Replicates of the Lig E-KO-KanaR mutant (d) Average readings of all three *N. gonorrhoeae* variants with three replicates for WT MS11, two replicates for Lig E-His-KanaR and two replicates for Lig E-KO-KanaR

After further analysis of the data, a decision was made to treat culture C of the Lig E-His-KanaR mutant and culture A of the Lig E-KO-KanaR mutant as outliers, the removal of which generated the average growth curve in Figure 3.6 (d). From this graph, several key differences were observed between the three mutants. In general, the Lig E-KO-KanaR mutants had a longer average lag phase (up to around 20 hours), while the lag phases for both the WT and Lig E-His-KanaR mutant were much shorter (up to around 15 hours). Despite entering the stationary phase at a similar time (approximately 41-43 hours), the OD_{600} readings of the Lig E-KO-KanaR mutant during the exponential phase were not as high as those of the other

two conditions. In fact, the average OD_{600} reading for the Lig E-KO-KanaR mutant during the stationary phase was similar to those of the lysis phase of the WT MS11 and the Lig E-His-KanaR mutant. Furthermore, the lysis phase of the Lig E-KO-KanaR mutant is not as defined as the other two conditions, as a decrease in OD_{600} readings was only observed for a single time point (hour 63) as opposed to a number of decreasing timepoints. At this stage, the OD_{600} readings for the Lig E-KO-KanaR mutants were higher than that of the other two conditions, again hinting at its continuous lag throughout its growth. This lag was also consistently observed in the pilot growth experiments (Figure A.6). However, when taking into consideration the rates of exponential growth of all three conditions between hours 19 to 34 (Figure 3.7), no obvious differences were observed (rate of exponential growth for WT MS11, Lig E-His-KanaR and Lig E-KO-KanaR were 0.03137, 0.04538 and 0.02995 respectively), suggesting similar rates of growth for all cultures. Based on this, it is predicted that the growth trajectory of the Lig E-KO-KanaR mutants were not due to slow growth or the inability to divide, but due to these mutants spending a longer time in the lag phase.
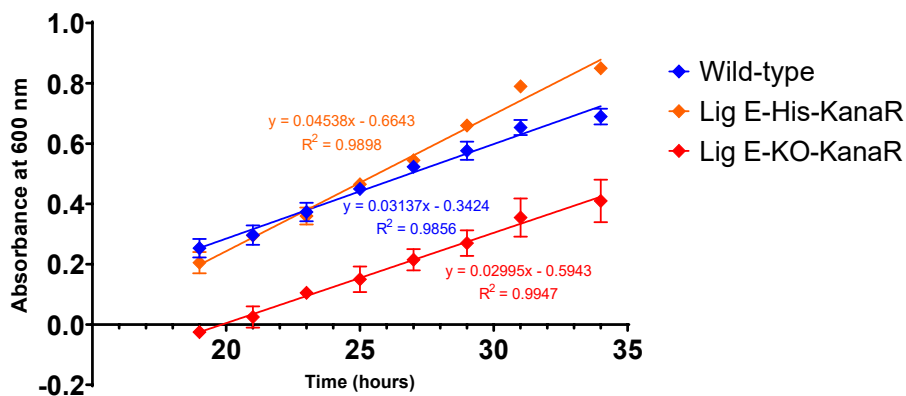
Figure 3.7. Linear range from hours 19 to 34 of average MS11 *N. gonorrhoeae* OD_{600} readings for all three Ngo-Lig variants from Figure 3.6 (d), with standard errors plotted

Despite the similar growth trajectories between the WT MS11 and Lig E-His-KanaR mutant (Figure 3.6 (d)) it was also observed that the Lig E-His-KanaR peak OD_{600} readings (at the late exponential, stationary phase and lysis phases; 27-63 hours) were higher than the peak WT MS11 readings at those timepoints. This was

observed consistently both in the main growth experiment and the pilot experiments (Figure A.6). Regardless, the purpose of the Lig E-His-KanaR mutant was to act as a control for the kanamycin resistance gene present in the Lig E-KO-KanaR mutant. Based on the OD$_{600}$ readings, it can be inferred that neither elimination of the *Lig E* gene, nor insertion of a His-tag on the C-termini of the Ngo-Lig protein, had any adverse effect on the growth of *N. gonorrhoeae*, as both had similar growth trajectories to the WT MS11. In addition, it does not appear that the results observed for the Lig E-KO-KanaR were caused by the kanamycin resistance gene as the His-tagged variant did not show any slowed growth. It is thus possible to conclude that Ngo-Lig is not essential for either gonococcal survival or growth as the Lig E-KO-KanaR mutants were still able to grow, albeit slower. Although OD readings provide a good indication of the growth of cells, results recorded were based on the number of cells in media, regardless of whether they were viable or not (Scott, 2011). Hence, it is also necessary to account for the number of viable cells through CFU readings.

3.2.2.2   N. gonorrhoeae growth experiment: CFU/mL measurements

For all three biological replicates of each variant, three technical replicates from three separate dilutions were created to calculate the number of colonies formed, which resulted in nine technical replicates for each condition (Figure 3.8 (a) for WT MS11, Figure A.7 (a) for Lig E-His-KanaR and Figure A.7 (b) for Lig E-KO-KanaR), comparisons of which can be observed in Figure A.7 (c). As cultures C and A of the Lig E-His-KanaR and Lig E-KO-KanaR mutants were removed from the final average OD$_{600}$ readings (Figure 3.6 (d)), the corresponding CFU/mL readings were also removed from the final CFU analyses for consistency among the growth experiment (i.e. Figure 3.8 (b) for Lig E-His-KanaR, Figure 3.8 (c) for Lig E-KO-KanaR, Figure A.7 (d) for all three conditions), which led to smaller errors and more robust data points. Due to rapid growth of the bacteria after plating onto GCB plates, technical difficulties in colony counting were encountered and some of the raw data were excluded as they were not reliable. These included the reading at hour 49 for the average Lig E-KO-KanaR CFU/mL (indicated via the red arrow in Figure 3.8 (c)), which resulted in the final average CFU/mL graph in Figure 3.8

(d). Another overgrown data point included the reading at hour 13 for the WT MS11 cultures (blue arrows in Figure 3.8 (a)).
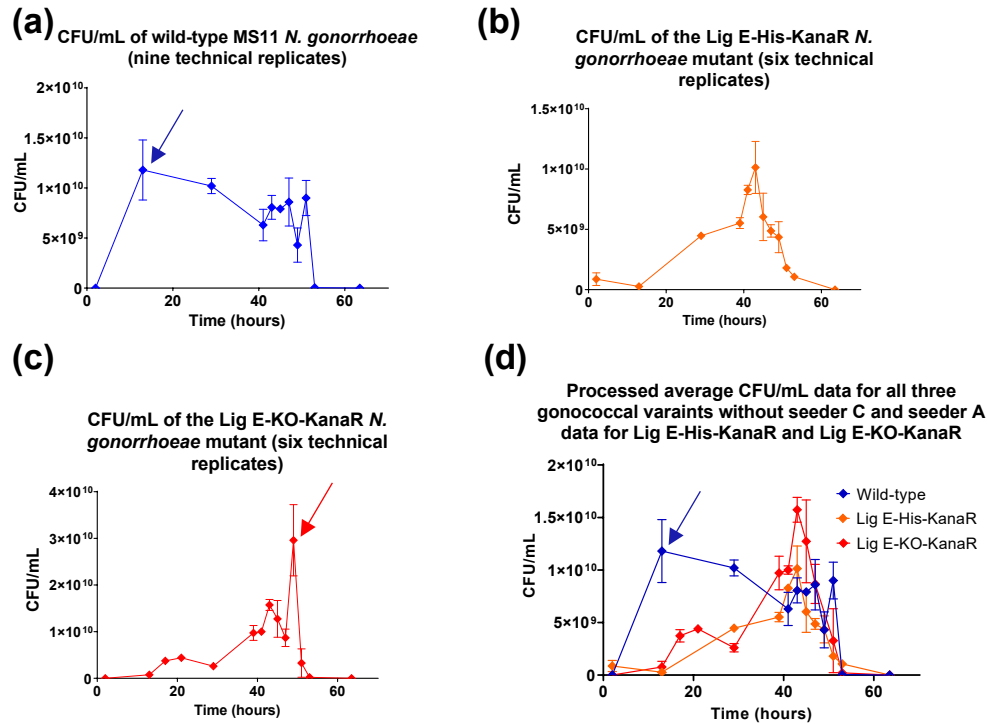
Figure 3.8. Average CFU/mL measurements of the MS11 *N. gonorrhoeae* cultures from the growth experiment with standard errors plotted (a) Average of nine technical replicates for the wild-type variant (b) Average of six technical replicates for the Lig E-His-KanaR variant (culture C was excluded) (c) Average of six technical replicates for the Lig E-KO-KanaR variant (culture A was excluded) (d) Average measurements for all three *N. gonorrhoeae* variants with nine technical replicates for the wild-type cultures, six for Lig E-His-KanaR cultures, and six for Lig E-KO-KanaR cultures, the latter of which did not include the datapoint at hour 49 due to low reliability of the data point. Arrows indicate data points that are discussed in this section. The graphs of the raw data before processing can be found in Figure A.7.

Based on the graph in Figure 3.8 (d), the average CFU/mL readings for both the Lig E-His-KanaR and Lig E-KO-KanaR mutants increased up to approximately hour 43, before a steep decrease. Comparisons of the CFU/mL readings of individual biological replicates (Figure A.8) show similar trajectories as well. When comparing these graphs to their corresponding OD_{600} graphs (Figure 3.6), it can be inferred that the number of viable cells increased up to the stationary point of growth (approximately hour 43) after which it declined due to a decrease in viability. The quantity of viable cells for the two Ngo-Lig mutants were also very similar to each other throughout their growth, indicative of the minimal effects of the

transformation process on the viability of gonococcal cells. In saying the above, the main difference in CFU/mL trajectory was observed for the WT MS11 colonies with a peak at approximately 13 hours. However, as the plates for hour 13 of the WT MS11 were clearly overgrown (indicated via blue arrows), making accurate colony counting extremely difficult, interpretation of this data point has to be treated with caution. By doing so, one could argue that the peak CFU/mL reading occurred at approximately hour 43 which agrees with the corresponding OD_{600} readings, suggesting consistency between the OD_{600} readings and cell viability.

What is most surprising about the data generated in both Figure 3.6 and Figure 3.8 however, is the longer growth phases when compared to other *N. gonorrhoeae* viability curves. For example, work by Kwiatek *et al*. (2014) reported a two-hour lag phase for the WT FA1090 strain, followed by a stationary phase at hour 5. This is in stark contrast to the stationary phase observed at hour 43 in our experiment. Even though the different gonococcal strains used may contribute to the different rates of growth (FA1090 vs MS11), preliminary data published in a thesis showed similar rates of growth for both strains (D'Ambrozio, 2015). Thus, these differences may be attributed to the different growth conditions instead. For example, the Kwiatek group used the recommended GCB broth by Dillard (2011), whereas tissue culture media (DMEM/F-12, 10% FBS) was used in this experiment as it was readily available in our laboratory. Despite the nutritionally rich concoction of DMEM/F-12, some key elements important for *Neisseria spp.* growth were missing, including biotin and some inorganic salts like potassium sulphate, magnesium chloride and ammonium chloride, which may slow growth and explain the 63-hour growth observed in our experiments (Griffin & Rieder, 1957; Catlin, 1973). Hence, any liquid culture of *N. gonorrhoeae* in the future should be performed in a chemically defined media to promote optimum gonococcal growth (Catlin, 1973). Furthermore, other groups have used the BacTiter-Glo[TM] Microbial Cell Viability Assay to measure cell viability where the amount of detected ATP corresponds to the number of viable cells, which offers greater accuracy due to the reduced number of steps that may lead to increased errors (Sule *et al.*, 2008; Kwiatek *et al.*, 2014). This method can be adopted if this experiment is repeated in the future.

3.2.2.3 N. gonorrhoeae growth experiment: External DNA quantification

After determining that *Lig E* is not essential in *N. gonorrhoeae*, the intention was to perform experiments in relation to its function. As the hypothesis of Lig E's function was that it seals breaks in exDNA, the supernatant samples isolated during the growth experiment were analysed with the Quant-iT™ PicoGreen™ dsDNA Assay Kit to quantify the amount of exDNA present in the growth medium. Five different time points were chosen based on the different stages of growth: one during the lag phase (hour 5), two during the exponential phase (hours 23 and 31), one during the stationary phase (hour 45) and one during the lysis phase (hour 52). Due to the differences in OD_{600} readings in cultures C and A of the Lig E-His-KanaR and Lig E-KO-KanaR mutants, the cultures were initially treated as separate to look at variations between different biological replicates (Figure 3.9 (a)(b)(c)).
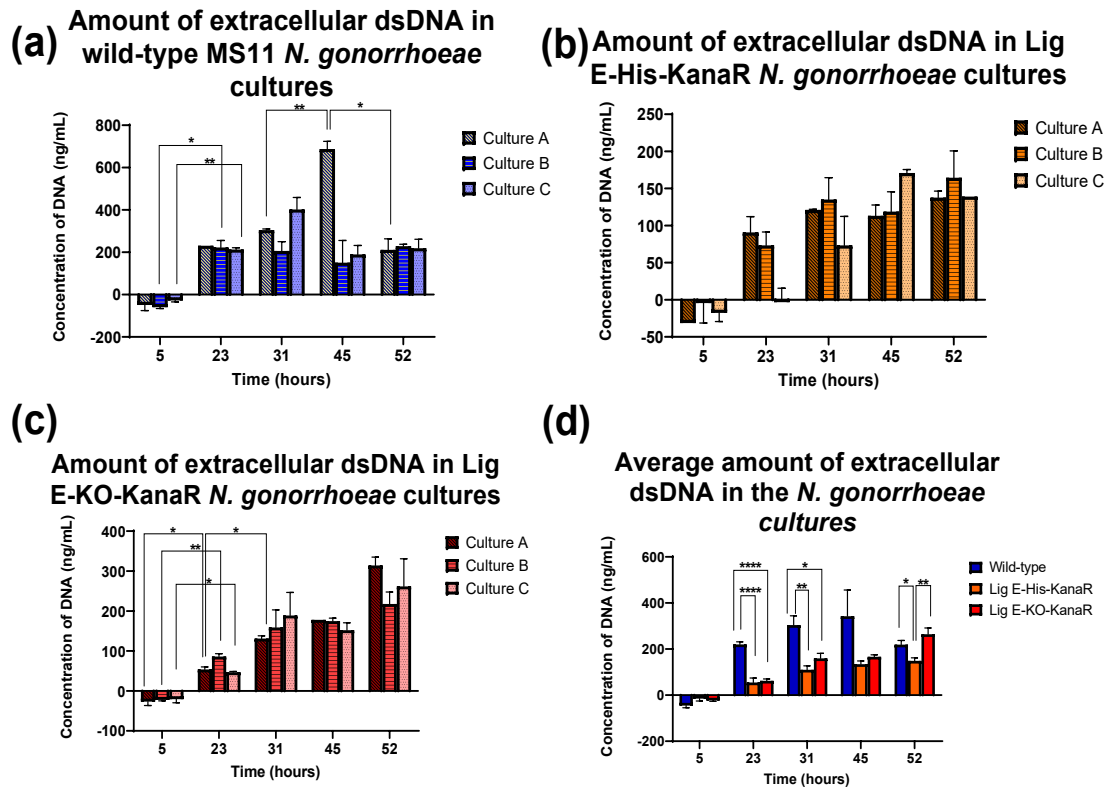
Figure 3.9. Amount of extracellular dsDNA detected in the supernatant samples of MS11 *N. gonorrhoeae* cultures obtained from the growth experiment and quantified using the PicoGreen® dye kit. The standard errors and any significant differences from unpaired two-tailed t-tests are plotted ($p < 0.05$ (*)) (a) Replicates from wild-type MS11 cultures (b) Replicates from Lig E-His-KanaR cultures (c) Replicates from Lig E-KO-KanaR cultures (d) Average results of all three biological replicates for each MS11 *N. gonorrhoeae* variants. Any significant differences between the different time points for the average results of each variant are shown in Figure A.9.

Based on the average results (Figure 3.9 (d)), a general pattern of increasing exDNA in the gonococcal growth media as the time and number of cells increases can be observed, especially during the stationary or lysis phases when compared with the lag phase (hour 5). There were considerable variations between each biological replicate for each condition. For the WT MS11 for example (Figure 3.9 (a)), culture A observed a steep increase in the amount of exDNA during the stationary phase (hour 45) and a sharp decrease during the death or lysis phase (hour 52). This was not observed for the other two WT MS11 cultures, where the stationary phases at hour 45 observed a lower amount of exDNA compared to the exponential phases (hours 23 and 31). Average results for the WT MS11 (Figure A.9 (a)) indicate significant increases in the amount of exDNA only between the lag and exponential phases but not between the stationary and lysis phases.

In contrast to the WT MS11 cultures, the general trend of increasing exDNA for the Lig E-His-KanaR mutant is more prominent, with some variation between the three replicates (Figure 3.9 (b)). As per the WT MS11, average results among the biological replicates of Lig E-His-KanaR (Figure A.9 (b)) also indicate significant increases only between the lag and exponential phases, but not between the stationary and lysis phases, and the amount of exDNA in comparison to the WT MS11 is lower. Interestingly, no major differences were observed between cultures C and A of the Lig E-His-KanaR and Lig E-KO-KanaR mutants and their other two replicates. In fact, cultures A and B of the Lig E-KO-KanaR observed a steady increase in the amount of exDNA as time progressed, even up to hour 52, without the eventual decrease observed in the other two conditions (Figure 3.9 (c)). This may be related to the fact that the Lig E-KO-KanaR growth cultures never really reached a death or lysis phase where one would expect a drop in OD_{600} readings. Thus, one would predict that the amount of exDNA at this time point would be low as well. It is also possible that these mutants were still in the stationary phase at hour 53, and that the lysis phase was not actually observed. It is interesting to note that this mutant was the only condition that observed significant increases of exDNA between the stationary and death phases (Figure A.9 (c)), which may be due to the absence of Lig E in this mutant that may lead to ineffective DNA uptake and a semi constant pool of exDNA towards the end of its growth. However, this does not provide conclusive evidence that Lig E is responsible for the decrease in

exDNA observed for both the WT MS11 and Lig E-His-KanaR mutant during the lysis phase due to the high amount of variation recorded.

It is worth noting that some technical limitations may have impacted the measurement of exDNA. For example, supernatant samples used during this experiment were first heat- and EDTA-treated before analysis with the Quant-iT™ PicoGreen™ dsDNA Assay Kit to allow us to work with them outside of the PC2 containment laboratory. As was evident from their effects on the control lambda bacteriophage DNA (Figure A.10 (c)), this heat- and EDTA-treatment affected the fluorescence recorded, and hence the concentration of DNA calculated. The heating of DNA at 95 °C and subsequent reannealing at room temperature may have had inconsistent reannealing rates, leading to more ssDNA than dsDNA in some samples, and hence a lower fluorescence detected. This could skew quantification with the dye as PicoGreen™ only detects dsDNA (Dragan *et al.*, 2010; Wang *et al.*, 2014). Even though it is believed that Lig E works with dsDNA, the activity of Ngo-Lig on ssDNA has not been investigated. If the experiment was to be repeated, other ssDNA detecting dyes like SYBR-2 could be used in concert with the PicoGreen™ dye to quantify the amount of ssDNA in the environment (Pjura *et al.*, 1987; Günther *et al.*, 2010; Han *et al.*, 2019). In addition to the treatment of samples, two other possible sources of inconsistencies include the growth media and the dye kit used. The growth media used contained FBS, which originates from foetal bovine blood, and any nucleic acid present in FBS is also detected in the supernatant, as observed in the standard curves generated with and without 10% FBS (Figure A.10 (a)). Hence, the final calibration curve used was based on heat- and EDTA-treated lambda DNA in 10% FBS DMEM/F12 to mimic the conditions of samples (actual calibration curve used in Figure A.10 (b)).

Regardless, it is clear that more thorough investigations into the correlation between the stationary and death phases and the depletion of exDNA are required to confidently infer the role of Ngo-Lig on any of the differences observed. It would be interesting to also quantify the amount of extracellular ATP that may act as the necessary cofactor for Ngo-Lig in the periplasm to determine if Ngo-Lig is indeed more active during the lysis phase of growth.

3.2.3   A new hypothesis of Lig E's role in biofilm formation

The initial primary hypothesis of the function of Lig E focussed on Lig E's possible role in competence. However, the results obtained from the growth experiment and dsDNA analysis (sections 3.2.2.1, 3.2.2.2 and 3.2.2.3) suggest additional or alternate roles of this protein. In particular, the longer lag phase for the Lig E-KO-KanaR mutant (Figure 3.6 (d)) begs the question of why the absence of a competence-related protein would affect gonococcal growth in liquid media when no stressors were introduced. A possible technical explanation for the slower Lig E-KO-KanaR growth is the impact of the gene disruption on the genes surrounding *Lig E*. It is possible that *Lig E* is part of a major operon and that deleting this part of the operon may slightly slow gonococcal growth (Lim *et al.*, 2011). In particular, the surrounding genes of *Lig E* in MS11 *N. gonorrhoeae* include a metallo-dependent hydrolase at its 5' end and a hypoxanthine phosphoribosyltransferase at the 3' end, both of which may aid growth through hydrolysis or purine biosynthesis (Shinners & Catlin, 1982). To test this hypothesis in the future, it is possible to mutate the catalytic lysine of Ngo-Lig and compare its growth to the Lig E-KO-KanaR mutant to determine if the slower growth observed here was due to a disruption of the *Lig E* genetic environment or due to disruption of the Lig E protein itself. This mutation, when performed on recombinant Ame-Lig, reduced its DNA and ATP binding abilities, and hence its ligation activity (Williamson *et al.*, 2018). Alternatively, mutants with an inactivated *Lig E* (either through gene disruption or point mutation) can be complemented by a knock-in of *Lig E* and its promoter at a separate neutral position to determine if this restores normal growth.

Although the lag observed may be an artefact of gene disruption, a plausible alternate explanation is a disruption in other biological processes like biofilm formation. Biofilm is a community of bacteria attached to a surface in a matrix of extracellular polymeric substances (Karygianni *et al.*, 2020). As biofilm is important in holding communities of *N. gonorrhoeae* together to increase their chances of survival (through for example, resistance to some antimicrobial drugs and evasion of the host immune response), the disruption of biofilm formation may affect the bacterial growth (Greiner *et al.*, 2005). Further supporting this theory is the large quantity of exDNA of chromosomal origin that is a major constituent of

gonococcal biofilm due to its critical role in structural integrity (Harmsen *et al.*, 2010; Falsetta *et al.*, 2011; Jakubovics *et al.*, 2013; Devaraj *et al.*, 2019). These are obtained either from active secretion via the T4SS system during the lag phase, or via the autolysis of cells that eliminates the need for cell-to-cell contact (Dillard & Seifert, 2001; Hamilton *et al.*, 2005; Salgado-Pabón *et al.*, 2007; Kouzel *et al.*, 2015; Callaghan *et al.*, 2017). Biofilm formation from secreted DNA is aided by the action of endonucleases and deoxyribonucleases (DNases) that decrease biofilm thickness via their cleaving effects on DNA, thus affecting the pathogenicity of *N. gonorrhoea* (Harmsen *et al.*, 2010; Falsetta *et al.*, 2011; Jakubovics *et al.*, 2013; Zweig *et al.*, 2014). Furthermore, secreted chromosomal ssDNA tend to stay in the extracellular environment rather than being taken up immediately (Dillard & Seifert, 2001). Such DNA may act as a substrate for Lig E ligation as it has been shown that shorter pieces of DNA (<500 bp) are unable to restore biofilm formation and attachment of *Listeria spp.*, as opposed to their longer counterparts (>500 bp) (Harmsen *et al.*, 2010; Steichen *et al.*, 2011; Devaraj *et al.*, 2019). In fact, the only significant increases in exDNA between different growth phases were observed for the Lig E-KO-KanaR mutant (Figure A.9 (c)), which was not significant for the other two variants. The involvement of Lig E in biofilm formation may also explain the higher growth rates of the His-tagged variant. In particular, the positive His-tag on the C-terminus of Lig E may increase affinity of the protein towards negative exDNA strands, potentially contributing to more robust and structurally intact biofilms via increased ligation activity of Lig E, leading to its higher growth rates observed during the growth experiment (Figure 3.6 (d)) (Holmberg *et al.*, 2012).

If this biofilm theory holds true, it would have important implications in the pathogenesis of asymptotic women as biofilm formation is critical for attachment onto human urogenital tract cells (Steichen *et al.*, 2008; Steichen *et al.*, 2011). In fact, work by Kwiatek *et al.* in 2014 on generating knock-outs of the gonococcal *dam replacing gene* important for biofilm formation generated growth curves with a longer lag phase and lower number of dispersed cells in culture, synonymous to that of the Lig E-KO-KanaR mutants. Further transcriptomic work with their mutants showed the knock-out effects on genes important for "*translation, DNA repair, membrane biogenesis and energy production*" and decreased ability of the cells to adhere to human epithelial cells (Kwiatek *et al.*, 2014). With transcriptomic

analyses on the supernatant samples collected during the course of this thesis, it will be possible to compare the activated genes for each mutant. Furthermore, one could look at the rates of biofilm formation using a continuous flow chamber and a glass surface for the biofilm to form, in addition to further infection experiments of human cervical cells (Kouzel *et al.*, 2015).

It has to be made clear that determining the exact role of Lig E in relation to biofilm may not be straightforward. Based on the results collected, we postulate three different possible roles of Ngo-Lig. The first, is that Ngo-Lig may be working with the T4SS system, sealing any breaks in DNA to be exported, which agrees with the trend of increasing exDNA as the number of cells increase. What is particularly interesting is the discovery of a nuclease with a Tat dependent signal peptide in the periplasm of *N. gonorrhoeae*, which appears to remodel ssDNA destined for the biofilm (Falsetta *et al.*, 2011; Steichen *et al.*, 2011). It is possible that Ngo-Lig is working in concert with this nuclease in the periplasm for optimal biofilm thickness. Secondly, Ngo-Lig may be working with DNA obtained via autolysis or fratricide of neighbouring gonococcal cells for biofilm formation. As these pieces of DNA would contain DUS sequences, they can be readily taken up by the gonococcal cells. Due to the lysis of cells and the presence of nucleases in the oxidative environment, these pieces of DNA would contain nicks and breaks, of which Ngo-Lig may repair in the periplasm before export into the extracellular space to strengthen the biofilm. This would explain the continuous increase in exDNA even at hour 53 for the Lig E-KO-KanaR mutant, and the ability of the FA1090 strain to still form biofilms even without the T4SS system (Greiner *et al.*, 2005; Zweig *et al.*, 2014). The third possible role of Ngo-Lig relates to the initial hypothesised role of Lig E with DNA competence. This is plausible as the high amount of exDNA in the biofilm may act as a pool for gene transfer or nutrients, increasing the transfer of antibiotic resistance genes and hence the evolution of the multidrug resistant bacterium (Harmsen *et al.*, 2010; Devaraj *et al.*, 2019). In fact, it is plausible that Ngo-Lig is carrying out all three proposed roles in the periplasm; the first and the second proposed role during early stages of growth where secreted ssDNA (retained in blebs, for formation of 3-dimensional biofilm structures) and dsDNA in the biofilm is important, and the third role during later stages of growth when the ssDNA is no longer retained, leaving just the dsDNA in the mature biofilm (Falsetta *et al.*, 2011;

Zweig *et al.*, 2014). It has been shown that DNA transfer of antibiotic resistance genes and hence multidrug resistance, occurs the most frequently in early biofilm formation, which correspond to where exDNA is present at the highest amount (for up to 24 hours) possibly due to the less dense nature of the biofilm (compared to late biofilm formation), allowing DNA to be more motile (Kouzel *et al.*, 2015). Notably, biofilm formation has been documented in other Lig E-containing bacteria like *H. influenzae*, *V. cholerae* and *C. jejuni* as well (Domenech *et al.*, 2016; Karygianni *et al.*, 2020). Although, the connection between Lig E and biofilm formation has not been experimentally investigated our work, the predicted periplasmic location of Lig E makes this theory entirely plausible. Hence, the first step to validating this biofilm theory would be to confirm the potential periplasmic location of Lig E.

3.2.4   Western blot optimisations

Irrespective of the biological function of Ngo-Lig, the cellular location of Ngo-Lig is still of interest. Apart from tracking via fluorescence microscopy, an alternative to elucidating the potential periplasmic location of Ngo-Lig is to carry out subcellular fractionation of the Lig E-His-KanaR cells to isolate the periplasmic proteins, before performing a western blot against the poly-histidine tag on the C-terminus of Ngo-Lig to determine its location using a primary anti-His-tag mouse monoclonal antibody (Ramsey *et al.*, 2012). In addition, a primary *N. gonorrhoeae* monoclonal antibody against the gonococcal OMP would be used to account for any loading variation. Before this can be performed, optimisations of the western blot protocol using recombinant C-terminally His-tagged Ngo-Lig (i.e. Lig E-His) have to be conducted to ensure that we can reliably detect the His-tag. Details of the production of recombinant Lig E-His used here can be found Chapter Four. Optimisations of the western blotting protocol throughout the course of this thesis showed higher detection of Lig E-His with the nitrocellulose membrane over the PVDF membrane (Figure A.11), as well as the use of a 1:500 dilution of both primary antibodies. Initial plans involved the usage of a CruzFluor^{TM} 647 fluorophore-conjugated secondary antibody for detection via fluorescence. However, imaging of the membranes proved difficult due to the high background noise relative to the signal observed (Figure A.12). For this reason, we tested two

new secondary antibodies; a mouse monoclonal and a mouse polyclonal antibody, both conjugated to horseradish peroxidase, which can be detected via chemiluminescence.

Using the His-standard ladder as a positive control, it is evident that horseradish peroxidase-conjugated secondary antibodies have improved detection of our proteins of interest compared to the fluorophore-conjugated counterpart, due to the essential enzymatic reaction in the last step that amplifies any low protein signals (Nakane & Kawaoi, 1974). Furthermore, we also found stronger detection of the proteins of interest using the polyclonal antibody compared to the corresponding monoclonal antibody (Figure 3.10 (a)). This is due to its enhanced recognition of an antigen through binding of different epitopes, increasing detection over monoclonal antibodies that recognise a single epitope (Lipman *et al.*, 2005). Detection of the Lig E-His protein by the polyclonal secondary antibody over the background of other *N. gonorrhoeae* proteins was also tested via addition of the recombinant protein to the gonococcal lysate, followed by western blotting using either anti-His or anti-gonococcal OMP primary antibodies, which yielded positive detection of both proteins (Figure 3.10 (b)). The polyclonal secondary antibody provided good signal to noise ratio for Lig E-His detection (Figure 3.10 (b)), however significant background was observed with the gonococcal OMP. Thus, a future step will include optimisations of OMP detection using the monoclonal secondary antibody or via different dilutions of the polyclonal secondary antibody. Based on the blot in Figure 3.10 however, it was decided that the final optimal condition for this work was a 1:500 dilution of both primary antibodies, and a 1:1000 dilution of the polyclonal mouse secondary antibody on a nitrocellulose membrane.
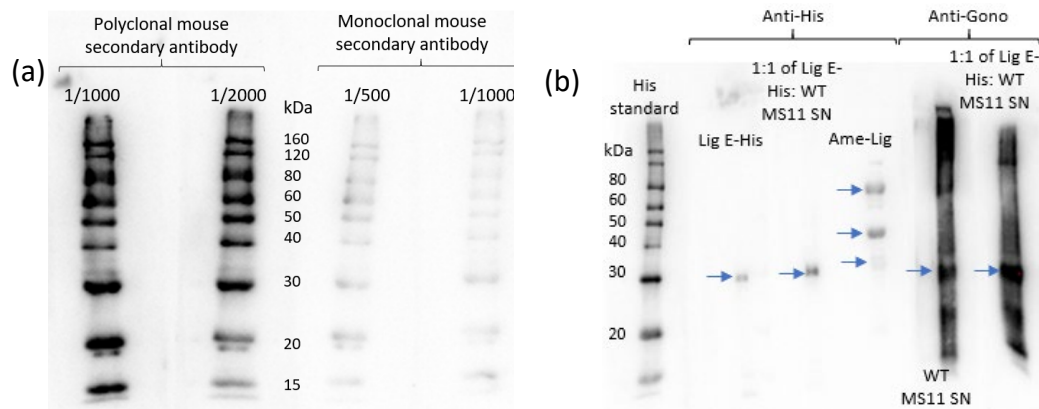
Figure 3.10. Western blot optimisations of mouse secondary antibodies conjugated to horseradish peroxidase (a) Comparisons of the efficacy of polyclonal and monoclonal mouse secondary antibody at different dilutions at detecting the His-standard ladder (b) Use of the polyclonal mouse secondary antibody (1:1000) to detect the proteins of interest (Lig E-His (anti-His antibody, 29 kDa) both alone and in the presence of the gonococcal lysate, a positive control Ame-Lig-His protein produced by another student (anti-His, 32 kDa for the cleaved protein, 44 kDa for the solubility tag and 76 kDa for the uncleaved protein with the solubility tag) and the gonococcal outer membrane protein (anti *N. gonorrhoeae* outer membrane protein antibody, 30 kDa) both alone and in the presence of the recombinant Lig E-His), with arrows pointing at the bands of interest. Note, the approximate molecular weight of the gonococcal OMP is also 30 kDa.

3.2.4.1 Expression of Lig E-His in gonococcal mutant samples

After optimisation of the western blot conditions, the next step was to apply these conditions to the Lig E-His-KanaR mutant pellet samples obtained during the growth experiment. Despite successful transfer of the proteins from the SDS-PAGE gel onto the membrane (Figure A.13 (a)), no His-tagged proteins, and hence no His-tagged Ngo-Lig proteins were detected at any of time points (Figure 3.11). Bands present in the His-standard and the recombinant gonococcal Lig E-His lanes (29 kDa) indicate that the blotting conditions used were sufficient for protein detection and that the issue lay with the resuspended pellet samples instead.
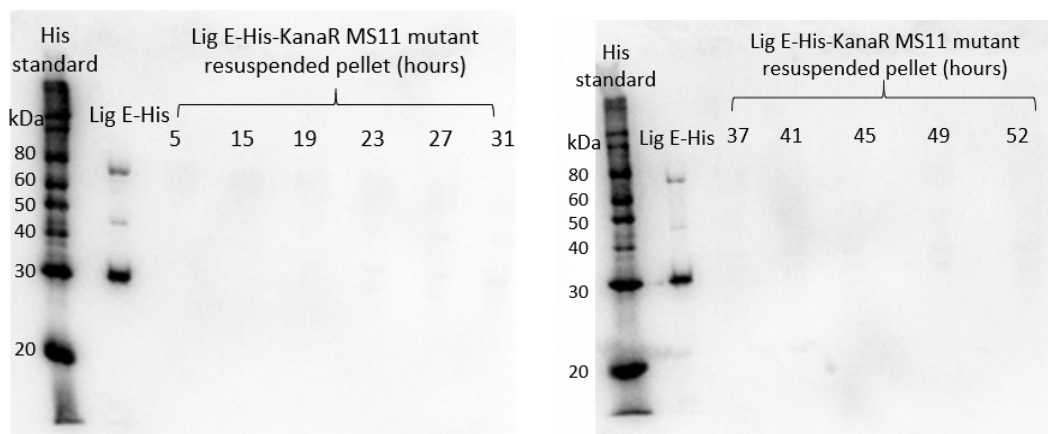
Figure 3.11. Western blot for detection of any His-tagged Ngo-Lig (anti-His antibody, 29 kDa) in the Lig E-His-KanaR mutant pellet samples isolated from the growth experiment. The faint bands above the expected 29 kDa band for the recombinant Lig E-His protein correspond to the MBP-tag and the tagged protein as an artefact from the purification process.

There are several possible reasons for the lack of detection of any His-tagged Ngo-Lig in the resuspended cell pellets (Figure 3.11) This includes the possibility that Ngo-Lig may be extracellular rather than periplasmic, and hence would be present in the supernatant, and not the pellet. To determine if Ngo-Lig was extracellularly localised, the western blot was repeated with the supernatant samples from the WT MS11 and both Lig E-His-KanaR and Lig E-KO-KanaR mutants from hours 31 (exponential phase) and 45 (stationary phase) (Figure 3.12). The absence of bands in these samples despite successful protein transfer (Figure A.13 (b)), accompanied by the detection of the recombinant Lig E-His positive control (29 kDa), either indicates that Ngo-Lig was not expressed in the supernatant, or that expression of the protein was below the limit of detection of our methods. This was also observed for the supernatant samples subjected to a nickel pull-down, which was performed in an attempt to enrich for any low levels of His-tagged Ngo-Lig. Based on the results from the growth experiment (section 3.2.2), it was evident that the absence of Ngo-Lig had an impact on the gonococcal phenotype (slower growth, longer lag phase), which indicates that Ngo-Lig does play a role under the conditions performed. It is hence possible that Ngo-Lig exerts its effect in the lag phase, where the number of cells are low, and hence detection of the protein would be difficult. To produce sufficient protein for future Ngo-Lig localisation experiments, the Lig E-His-KanaR mutant will be grown in a condition where the role and activity of Ngo-Lig is believed to be of most importance. This may include a DNA uptake

setting, or a biofilm assay at a high throughput (O'Toole, 2011; Haney *et al.*, 2018). Not only will this allow us to observe the different phenotypes of the Ngo-Lig mutants in relation to its possible function, but it will also enable us check for any expression of Ngo-Lig.
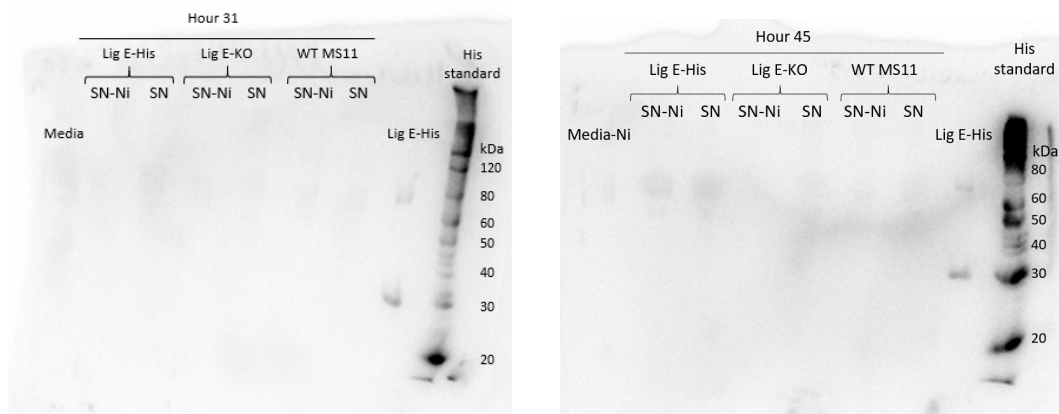
Figure 3.12. Western blot for detection of Lig E-His (anti-His antibody, 29 kDa) in the gonococcal supernatant samples isolated from the growth experiment. SN = supernatant samples; SN-Ni = nickel pull-down of supernatant samples to select for His-containing proteins.

3.3 Conclusion

The aim of this chapter was to generate Ngo-Lig mutants in *N. gonorrhoeae* to elucidate the location and importance of Ngo-Lig *in vivo.* Despite the lack of success in generating a fluorescently-tagged mutant of Ngo-Lig due to the nature and size of the GFP protein, the Lig E-His-KanaR and Lig E-KO-KanaR transformations proved successful, allowing us to generate several growth curves based on the OD_{600} and CFU/mL readings, as well as to collect supernatant and pellet samples for future investigations of the role of Ngo-Lig. The main results obtained from these experiments showed that Ngo-Lig is not essential for *N. gonorrhoeae* survival and growth. However, disruption of its gene in *N. gonorrhoeae* also leads to a longer lag phase and lower OD_{600} readings during growth. From the supernatant samples collected, the amount of exDNA was quantified and a general pattern of increasing exDNA with increasing number of *N. gonorrhoeae* cells was found for the Lig E-KO-KanaR mutant. Although little can be said about the correlation between the disruption of *Lig E* and the amount of DNA in the environment due to the variability between replicates, these results

provide a basis for further experiments regarding the possible substrates for Ngo-Lig activity in the periplasm.

Despite the ease of tracking fluorescently labelled protein to determine the location of Ngo-Lig, it is also possible to perform subcellular fractionation with the Lig E-His-KanaR mutant, followed by immunoblotting to confirm its putative location in the periplasm (Ramsey *et al.*, 2012). Initial analysis of the pellet and supernatant samples of this mutant showed no expression of any His-tagged Ngo-Lig protein, however the expression of the protein may have been below the detection limits of our methods considering the different mutant phenotypes observed throughout this chapter, especially during the lag phases. Thus, the next step would be to carry out a high throughput experiment with various conditions in parallel where Ngo-Lig might be highly expressed.

Based on these results, three different roles of Ngo-Lig are proposed. The first, is that Ngo-Lig works in the periplasm, in concert with a separate periplasmic nuclease, to remodel secreted chromosomal ssDNA for biofilm formation at initial stages of growth. The second is that Ngo-Lig works in the periplasm to seal breaks in DNA obtained from neighbouring lysed cells, again to contribute to biofilm formation at various stages of growth (possibly with both dsDNA and ssDNA). The third, is that Ngo-Lig works in the periplasm to seal breaks in DNA obtained from the biofilm or the environment to increase information content for homologous recombination at later stages of growth. If the role of Ngo-Lig in biofilm formation holds true, this would have important implications for the pathogenicity of *N. gonorrhoeae*, as biofilms are necessary for its infection and attachment to human epithelial cells in the genital tract, allowing for asymptomatic infections to occur and eventual resistance and persistence in the community (Steichen *et al.*, 2008; Falsetta *et al.*, 2011; Kwiatek *et al.*, 2014). It would be interesting then, to check if this is also translated across other Lig E-containing pathogens as the disruption of this critical part of the life cycle could affect their ability to cope with stress and hence offer a potential method to tackle the spread of human pathogens (Steichen *et al.*, 2008).

Although no concrete conclusions can be drawn on the function and location of Ngo-Lig, this chapter forms the basis of future experiments with Ngo-Lig *in vivo* by acting as a steeping stool in terms of the transformation process in *N. gonorrhoeae*, the use of the GFP protein for periplasmic protein tagging, the optimisation of Western blots in relation to our proteins of interest, and the formation of a few potential roles of Ngo-Lig in terms of biofilm formation.

Chapter Four
Activity of Ngo-Lig

4.1 Introduction

From Chapter Three, possible roles of Ngo-Lig with biofilm formation and competence were proposed. However, to better understand its biological function, it is necessary to first comprehend Ngo-Lig's mode of action. Previous work with recombinant Lig E from other bacteria demonstrated their abilities to ligate singly nicked dsDNA, and to a lesser extent, mismatch, overhang and gapped dsDNA (Cheng & Shuman, 1997; Magnet & Blanchard, 2004; Williamson *et al.*, 2014). Although it is predicted that Ngo-Lig functions similarly to other previously characterised Lig E (i.e. Ame-Lig, Vib-Lig, Psy-Lig, Hin-Lig and Nme-Lig) differences in experimental conditions (e.g. retaining the signal peptide on Nme-Lig) necessitate specific studies on Ngo-Lig. It is crucial to elucidate the mechanism and activity of Ngo-Lig via understanding its substrate of choice, to better understand how it works *in vivo* in *N. gonorrhoeae*.

There were two main objectives associated with this chapter. The first was to elucidate the activity of Ngo-Lig and to compare it to previously characterised Lig Es. The second was to ensure that the *in vivo* modifications made to Ngo-Lig in Chapter Three were not negatively affecting its activity and function. To do so, three different variants of the mature *Lig E* gene from *N. gonorrhoeae* (locus:NGFG_01849; protein:EEZ48933.1) without the predicted signal peptide (Figure 3.1) were recombinantly expressed and purified from *E. coli*. These included Lig E-GFP (C-terminally GFP-tagged), Lig E-His (C-terminally His-tagged) and the native Lig E from *N. gonorrhoeae*. Details of the cloning and expression processes performed are given in Chapter Two. Briefly, all genes were *E. coli* codon optimised and pre-cloned between two *attL* sites in pDONR221 Gateway donor plasmids, which were ordered from Twist Bioscience as His-(TEV)-Lig E-GFP, Lig E-His and His-(TEV)-Lig E respectively, where (TEV) symbolises a TEV protease cleavage site (E-N(asparagine)-L(leucine)-Y-F-Q(glutamate)-|-Gly/S). The additional His-tags before the cleavage site were used

for purification via immobilised affinity chromatography. Gateway cloning/LR reactions were performed to sub-clone the constructs of interest from between the *attL* sites of the pDONR221 vectors into the region between two *attR* (right attachment site) sites of destination plasmids, previously occupied by a *ccdB* toxicity gene (Katzen, 2007; Reece-Hoyes & Walhout, 2018). This yielded expression plasmids with the constructs of interest between two *attB* (attachment site for bacteria) sites, ready for transformation into *E. coli* (Reece-Hoyes & Walhout, 2018). The variety of destination plasmids available for Gateway cloning (e.g. pDEST14, pDEST17, pHMGWA) is attributed to the different N- and C-terminal sequences in these plasmids, allowing various tags to be added to the proteins of interest (Reece-Hoyes & Walhout, 2018). For example, the pDEST17 destination plasmid contains a start codon (ATG) and an N-terminal His-tag between the T7 promoter and the first *attR* site. On the other hand, the pDEST14 destination plasmid relies on the native start codon, which must be present in the donor plasmid, and was only present in the pDONR221 Lig E-His plasmid in this case. Thus, cloning into pDEST14 could not be performed for both the His-(TEV)-Lig E and His-(TEV)-Lig E-GFP donor plasmids. As a result, three expression plasmids for the three Ngo-Lig variants were planned at the start of this thesis. The first, a pDEST17 His-(TEV)-Lig E plasmid (Figure B.9), which after TEV cleavage, would allow purification of the native Ngo-Lig protein without the additional His-tag. The second, a pDEST17 His-(TEV)-Lig E-GFP plasmid (Figure B.12), which after TEV cleavage, would allow purification of Lig E-GFP without the additional His-tag. The third, a pDEST14 Lig E-His plasmid for expression of the Lig E-His variant (Figure B.11).

After transformation of the expression plasmids into *E. coli*, small-scale expression (30 mL) was conducted to determine the optimal conditions for protein expression. Ultimately, the selected conditions were transferred onto a larger scale (1 L) that were later purified via methods specific to the differently tagged Ngo-Lig variants to obtain large amounts of the pure proteins of interest. Gel-based ligation assays were then performed to determine the activity of Ngo-Lig. This chapter focuses on the results of the Ngo-Lig DNA ligation assays, as well as the optimisations performed throughout the whole expression-purification process.

4.2 Results and discussion

4.2.1 Ngo-Lig expression plasmid preparation

4.2.1.1 Cloning and small-scale expressions of Ngo-Lig

To generate expression plasmids for protein production, Gateway cloning was conducted between the Lig E-His pDONR221 plasmid and pDEST14, as well as between both the His-(TEV)-Lig E and His-(TEV)-Lig E-GFP pDONR221 plasmids and pDEST17. Amplification of these expression plasmids using the T7_promoter_primer and T7_terminator_primer primers showed successful cloning evidenced by the presence of bands that correspond to the expected 945, 1035 and 1746 bp sizes for the Lig E-His pDEST14 (Figure 4.1 (a)), His-(TEV)-Lig E pDEST17 (Figure 4.1 (b)) and His-(TEV)-Lig E-GFP pDEST17 (Figure 4.1 (b)) expression plasmids respectively. This was further verified via the presence of the expected 1977 and 1866 bp band for the positive pDEST14 and pDEST17 plasmid control amplified with the same primers.
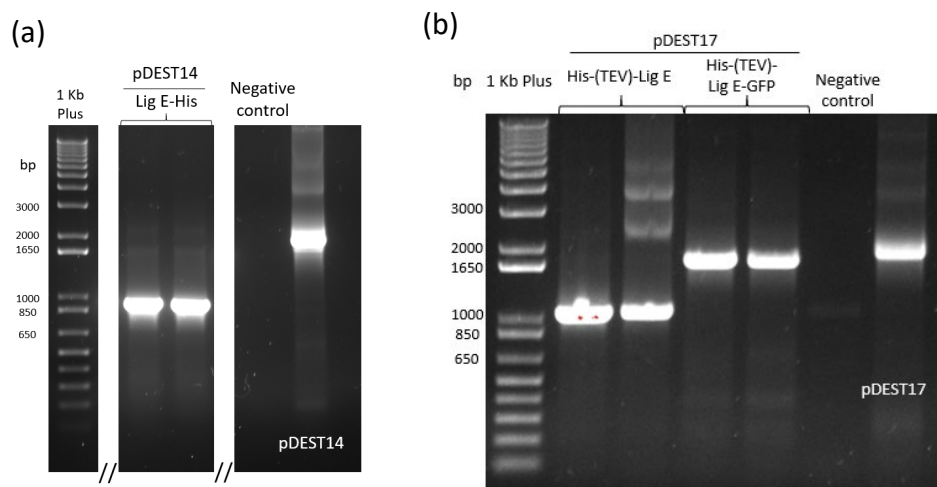
Figure 4.1. Agarose gels (1%) of samples after Gateway cloning between donor and destination plasmids to yield expression plasmids amplified with the T7_promoter_primer and T7_terminator_primer primers (a) pDEST14 Lig E-His, cropped ('//') to show the lanes of interest (b) pDEST17 His-(TEV)-Lig E and pDEST17 His-(TEV)-Lig E-GFP. Sizes of the 1 Kb Plus ladder used are labelled.

Small-scale expression of all three plasmids was performed to determine the rate of expression at two different temperatures (15 and 20 °C overnight) that were previously used to express other Lig E variants, Psy-Lig and Vib-Lig (Williamson & Pedersen, 2014; Williamson *et al.*, 2014). These were tested using

Origami™(DE3) cells that express thioredoxin reductase and glutathione reductase to aid with protein folding. However, this small-scale expression yielded no expression of the His-(TEV)-Lig E or His-(TEV)-Lig E-GFP proteins and very little for the Lig E-His protein (Figure B.15). To further optimise expression, the experiment was repeated with a different expression strain, BL21(DE3)pLysS, which expresses a T7 lysozyme protein that represses leaky expression in the presence of glucose (Grossman *et al.*, 1998). To try and increase the yield of soluble protein production, 1% glucose was added to the media to act as an alternative carbon source for suppression of any leaky Ngo-Lig expression that may lead to their toxicity and degradation (Grossman *et al.*, 1998). Furthermore, as the natural habitat of *N. gonorrhoeae* is the human body, expression at 30 and 37 °C (4 hours) was also of interest. Results from this (Figure 4.2) showed more expression of the insoluble His-(TEV)-Lig E-GFP protein (59.3 kDa) in the pellet sample over its supernatant counterpart, especially at 37 °C. On the other hand, although both insoluble and soluble expressions at 20 and 30 °C for the His-(TEV)-Lig E (32.6 kDa) protein were equal, more expression can be observed in the pellet samples at 37 °C. Both proteins were also detected in the nickel pull-downs. However, this was more prominent at 20 °C over 37 °C, indicating that 20 °C is the optimal condition for obtaining soluble expression of these two proteins.
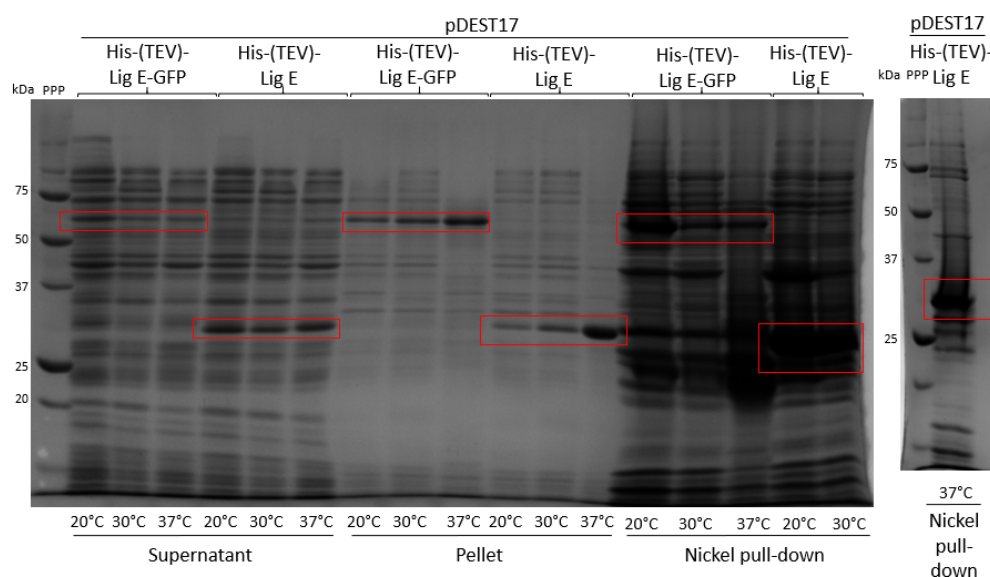
Figure 4.2. SDS-PAGE (12%) of the pDEST17 expression plasmids after small-scale expression in BL21(DE3)pLysS. The red boxes indicate the bands of interest. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.

It is interesting to note the effect of the high temperature (37 °C) at increasing insoluble protein expression. In general, it has been shown that lower temperatures during induction increase the expression of soluble proteins, hence both 15 and 20 °C were chosen as they allow for a suitable environment for proper protein folding (Shirano & Shibata, 1990; Kataeva *et al.*, 2005; Francis & Page, 2010). Based on the results, the high 37 °C temperature greatly affects the folding and solubility of the protein in *E. coli*, possibly due to the enhanced activity of the cell allowing for increased protein aggregation (Vasina & Baneyx, 1997; Francis & Page, 2010). Hence, an induction temperature of 20 °C was used for large-scale expression of these proteins.

Despite the aforementioned optimisations, several growth issues of the BL21(DE3)pLysS cells carrying the Lig E-His pDEST14 expression plasmid were identified. In particular, the OD_{600} readings never reached the expected 0.3-0.4 values (OD_{600} readings stayed at around 0.005). This is in contrast to their effects in DH5α and Origami^{TM}(DE3) cells which were able to grow to a sufficiently high OD for plasmid isolation and induction of protein expression respectively. Initial thoughts leaned toward an issue with the nature of the protein itself. For example, difficulty with expressing Vib-Lig in *E. coli* was previously observed by another group, which they inferred to be toxic due to the possible binding of Lig E to either the intracellular *E. coli* ATP or DNA, affecting their internal repair and replication machineries (Williamson & Pedersen, 2014). Regardless, we have successfully expressed recombinant protein for two versions of Ngo-Lig using pDEST17 plasmids, the His-(TEV)-Lig E and His-(TEV)-Lig E-GFP proteins. To optimise the expression of Lig E-His, a new construct was designed to make it similar to the other two, namely a His-(TEV)-Lig E-His construct, which only after TEV cleavage, allows the isolation of Lig E-His.

4.2.1.2 Making a new His-(TEV)-Lig E-His pDONR221 construct

4.2.1.2.1 *Restriction digestion and ligation*

To produce a His-(TEV)-Lig E-His pDONR221 plasmid, restriction digest was performed using EcoRI and BamHI to isolate a Lig E-His gene fragment from pDONR221 Lig E-His (996 bp), and the His-(TEV) fragment from the pDONR221

His-(TEV)-Lig E plasmid (2181 bp), both of which can be ligated together. This will ensure the removal of the start methionine codon from the Lig E-His fragment that will be replaced with the His-(TEV) construct. An agarose gel run of these fragments showed constructs of the intended sizes highlighted in red (Figure 4.3). After extraction of these bands, ligation was performed to yield pDONR221 His-(TEV)-Lig E-His and the ligation products were used for the next LR recombination step.
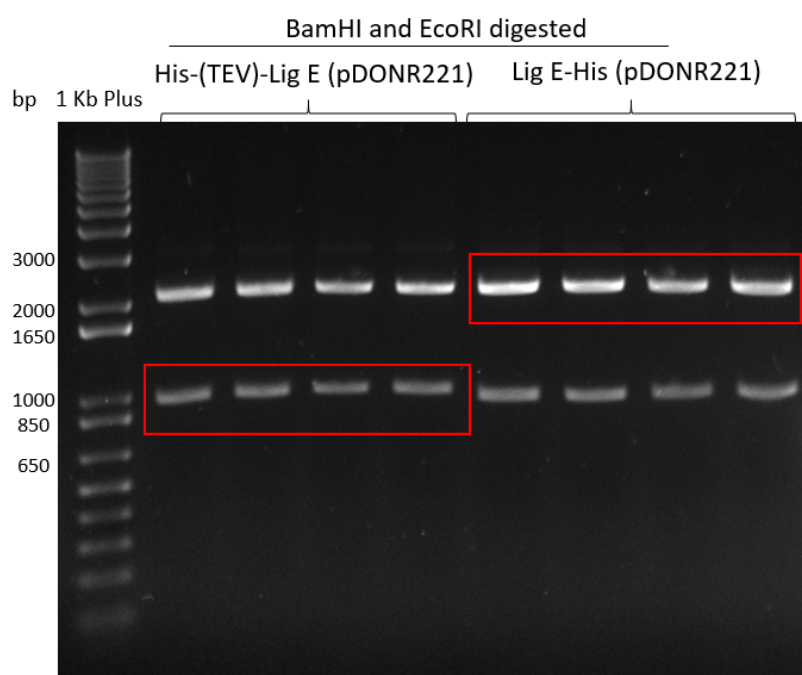
Figure 4.3. Agarose gel (1%) of both His-(TEV)-Lig E and Lig E-His pDONR221 donor plasmids after digestion with BamHI and EcoRI. Sizes of the 1 Kb Plus ladder used are labelled.

4.2.1.2.2 *LR reactions and small-scale expression*

LR recombination was performed between the pDONR221 His-(TEV)-Lig E-His plasmid and two different destination plasmids, pDEST17 and pHMGWA, the latter of which allows an MBP-tag to be added to the N-terminus of the protein, which increases its solubility and allows for better protein separation. Amplification of the expression plasmids with either the MBP_forward_primer or the T7_promoter_primer with the T7_terminator_primer (Figure 4.4) yielded products of the desired sizes (1052 and 2203 bp for the pDEST17 and pHMGWA His-(TEV)-Lig E-His expression plasmids with the T7_promoter_primer; 1009 bp for

the pHMGWA His-(TEV)-Lig E-His expression plasmids with the MBP_forward_primer), which along with the right sizes of the destination plasmids (1896 and 3046 bp for pDEST17 and pHMGWA respectively), suggest both successful ligation and recombination of the donor and destination plasmids.
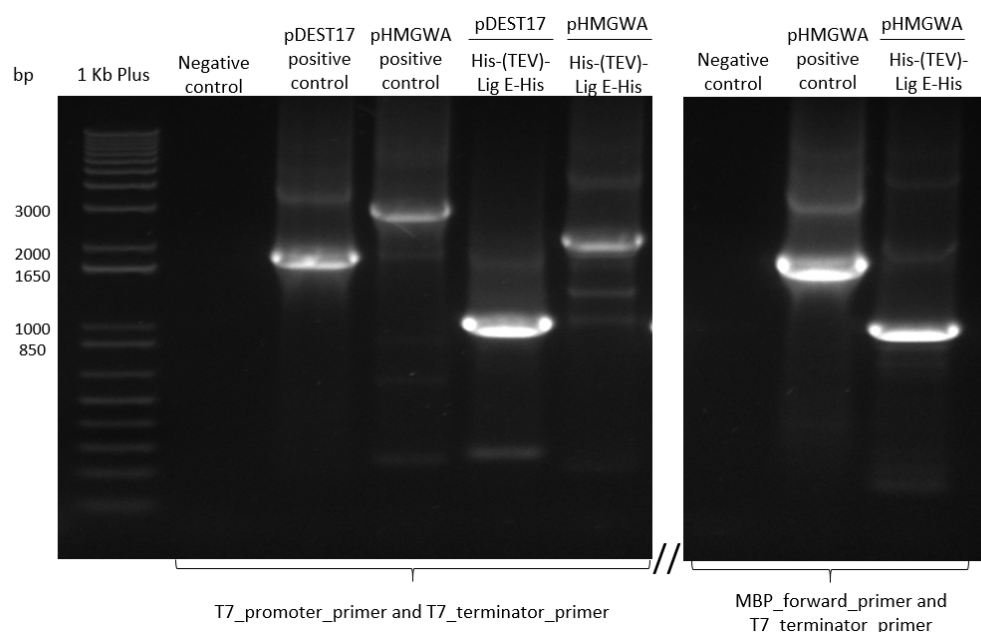
Figure 4.4. Agarose gel (1%) of both pDEST17 and pHMGWA His-(TEV)-Lig E-His expression plasmids, amplified with either the T7_promoter_primer and T7_terminator_primer primers or the MBP_forward_primer and T7_terminator_primer primers. The gel was cropped ('//') to show only the bands of interest. Sizes of the 1 Kb Plus ladder used are labelled.

After isolation of the new His-(TEV)-Lig E-His expression plasmids, small-scale expression was performed in BL21(DE3)pLysS cells with 1% glucose at different temperatures (15, 20 and 37 °C). Results from this (Figure 4.5) show similar levels of soluble and insoluble expression of the His-(TEV)-Lig E-His protein from pDEST17 (33.5 kDa), with the strongest expression at 20 °C in the supernatant. This protein was also visible in the nickel pull-down of the supernatant. In contrast, soluble expression of the same protein from pHMGWA (73.9 kDa) was significantly amplified across all three temperatures, especially in comparison to its insoluble counterpart in the pellet, where expression was low. This was further evidenced by the strong overexpression of the protein in the nickel pull-down, indicating an increase in the protein's solubility.
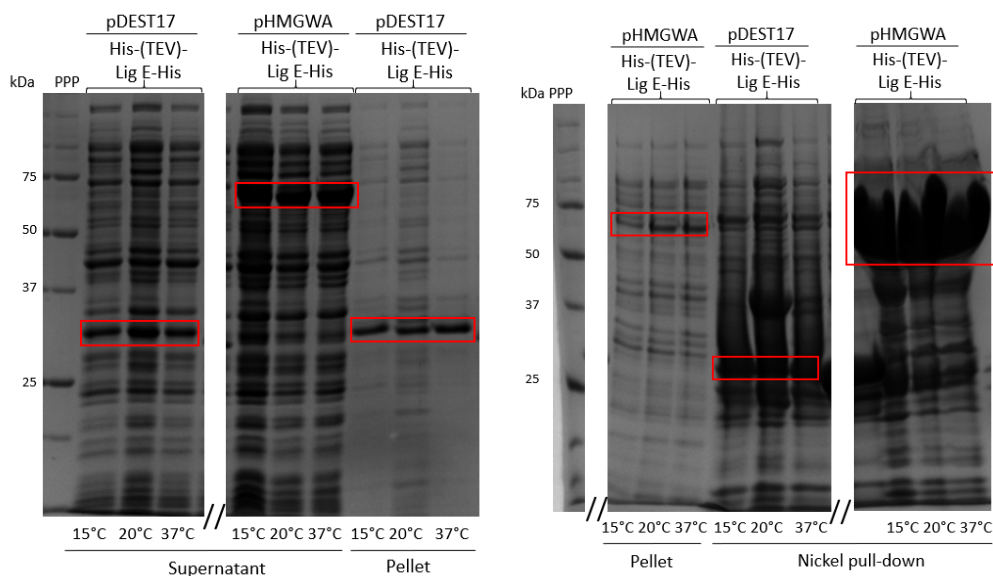
Figure 4.5. SDS-PAGE (12%) of supernatant, pellet and nickel pull-down samples after small-scale expression of pDEST17 His-(TEV)-Lig E-His and pHMGWA His-(TEV)-Lig E-His in BL21(DE3)pLysS. The red boxes indicate the bands of interest. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.

Based on Figure 4.5, it is evident that the use of the pHMGWA plasmid significantly improves protein solubility over the pDEST17 plasmid due to the presence of the large, soluble MBP protein, which not only increases the solubility of the His-(TEV)-Lig E-His protein, but due to its large size is believed to block the recombinant ligase from binding to DNA of the *E. coli* host (Kataeva *et al.*, 2005; Williamson & Pedersen, 2014; Reuten *et al.*, 2016). In particular, the MBP-tag also increases protein stability and facilitates proper folding by preventing aggregation of the proteins of interest (Kataeva *et al.*, 2005; Reuten *et al.*, 2016). Hence, the pHMGWA His-(TEV)-Lig E-His expression plasmid was used for future purification for gel-based assays, which after cleavage of the His-MBP-His-(TEV)-Lig E-His protein, produces Lig E-His.

4.2.1.3   Additional LR reactions and small-scale expressions for future Ngo-Lig ligation assay comparisons

The decision to use the pHMGWA His-(TEV)-Lig E-His plasmid for future Lig E-His purifications may lead to some complications due to the longer nature of the purification process needed to remove the MBP-tag. As a control, LR

recombination was performed between the His-(TEV)-Lig E donor plasmid and pHMGWA. Confirmation via the T7_promoter_primer and T7_terminator_primer primers showed successful cloning via the presence of the expected 2186 bp band (Figure B.16 (a)) and the presence of the expected 3046 bp band for the pHMGWA positive control. Small-scale expression (Figure B.16 (b)) showed slightly higher expression of the protein (73.0 kDa) in the supernatant compared to the pellet, with no specific preference for the induction temperature, and the presence of strong bands of the same size in the nickel pull-down indicate strong soluble expression of the protein.

In summary, five different expression plasmids of the Ngo-Lig variants were generated during the course of this thesis (Table 4.1). Of these, pDEST17 His-(TEV)-Lig E was used to obtain the native Ngo-Lig protein, pDEST17 His-(TEV)-Lig E-GFP was used to obtain Lig E-GFP and pHMGWA His-(TEV)-Lig E-His was used to obtain Lig E-His for gel-based assays.

Table 4.1. Expression plasmids of different Ngo-Lig variants of which BL21(DE3)pLysS stocks have been obtained. Proteins that were later purified are indicated in bold.

Protein of interest	Donor plasmid (pDONR221)	Destination plasmid	Purpose
Lig E	**His-(TEV)-Lig E (Figure B.3)**	**pDEST17**	**To obtain Lig E for assays**
		pHMGWA	**To act as a control for pHMGWA purification processes**
Lig E-GFP	**His-(TEV)-Lig E-GFP (Figure B.2)**	**pDEST17**	**To obtain Lig E-GFP for assays**
Lig E-His	**His-(TEV)-Lig E-His (Figure B.5)**	**pHMGWA**	**To obtain Lig E-His for assays**
		pDEST17	Stocks kept for future use; expression lower than pHMGWA His-(TEV)-Lig E-His

4.2.2 Protein purification of the different Ngo-Lig variants

The first step of the general protein purification workflow followed throughout the project was an immobilised metal affinity chromatography (IMAC). This was performed using a nickel column (Histrap) to isolate polyHis-containing proteins that interact with the column and were eluted using a high salt buffer (buffer B, 750 mM NaCl) (Sulkowski, 1985). Collection of the His-containing proteins was followed by a desalting procedure to remove imidazole and to lower the amount of salt present, allowing optimal TEV protease cleavage as this enzyme is hindered by the presence of high salt (Porath & Flodin, 1959; Dougherty *et al.*, 1989; Nallamsetty *et al.*, 2004). In most cases, a reverse IMAC/nickel procedure (reverse-Histrap) was performed after TEV cleavage to isolate the cleaved protein without the His-tag, followed by size exclusion to remove any aggregated proteins or other minor contaminants (Barth *et al.*, 1994). This process has been used for other Lig E purifications from other species (Williamson & Pedersen, 2014), and is illustrated by the first purification performed here to isolate Lig E-GFP away from the expressed His-(TEV)-Lig E-GFP protein (Figure 4.6).
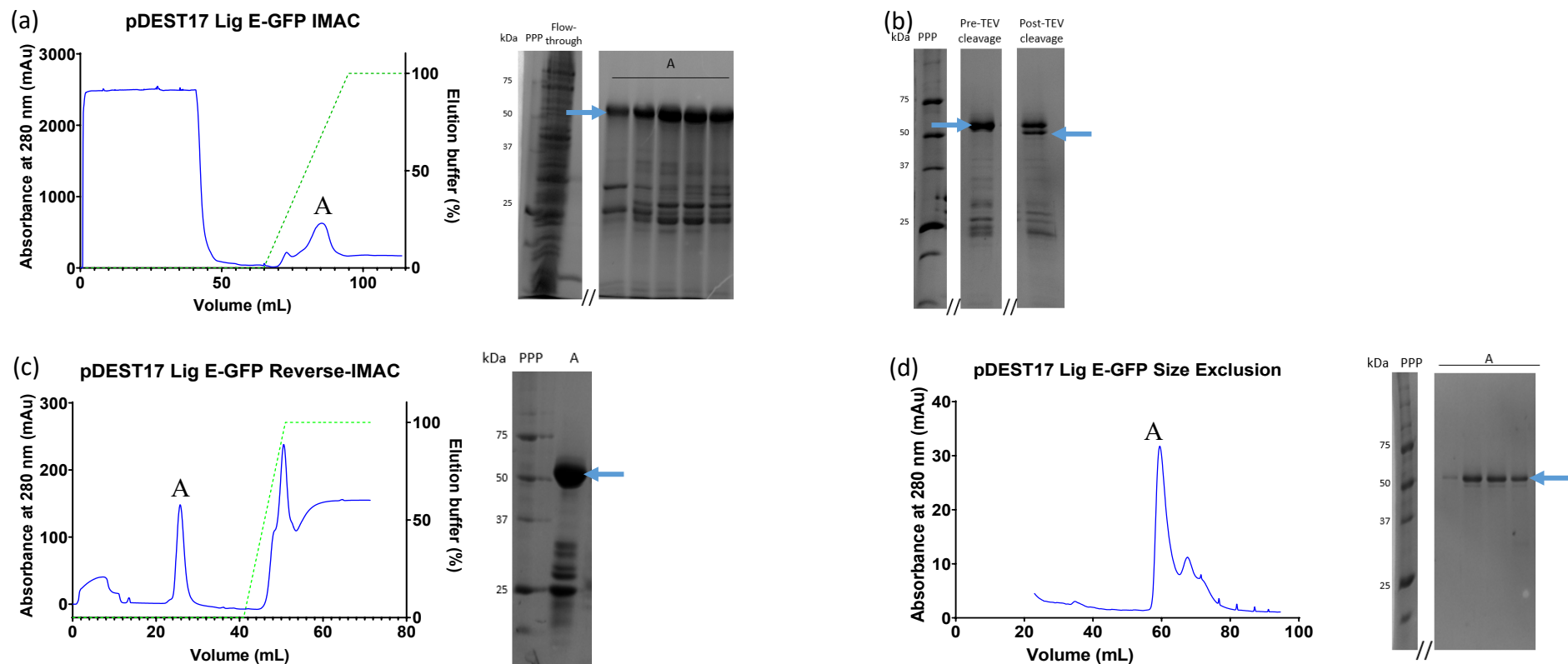
Figure 4.6. Chromatograms and SDS-PAGE gels (12%) from the His-(TEV)-Lig E-GFP purification to isolate Lig E-GFP (a) IMAC via a nickel column (b) TEV cleavage after desalting to lower the salt concentration (c) Reverse-IMAC via a nickel column (d) Size exclusion. Gels are cropped ('//') to show only the bands of interest which are indicated via blue arrows. The letter 'A' is used to show the peaks of interest on the chromatograms and their corresponding visualisations on the gels. UV intensity or absorbance at 280 nm (mAu) is represented in blue and the amount of elution buffer (%) is represented in green. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.

From the nickel pull-down of the Lig E-GFP purification (Figure 4.6 (a)), the His-(TEV)-Lig E-GFP containing fractions that eluted with 30% buffer B (59.3 kDa; peak B) were isolated and cleaved with TEV after a desalting procedure (Figure 4.6 (b)). Despite the low cleavage efficiency of TEV (less than 50% based on the intensity of the cleaved and uncleaved proteins (cleaved protein: 55.1 kDa)), sufficient cleaved protein was isolated for the reverse-His procedure (Figure 4.6 (c)). This was efficient at isolating Lig E-GFP (peak A) away from His-(TEV)-Lig E-GFP and the cleaved His-tag that bound to the column before their elution with buffer B. A final size exclusion procedure (Figure 4.6 (d)) to isolate the Lig E-GFP protein (peak A; 55.1 kDa) proved efficient at removing the lower molecular weight contaminants that were present in the prior reverse-His procedure.

Similar purification processes were repeated for the other three proteins (His-(TEV)-Lig E and MBP-His-(TEV)-Lig E to isolate the native Lig E, and MBP-His-(TEV)-Lig E-His to isolate Lig E-His). However, as detailed in the following sub-sections, there were variations specific to each protein or construct. For the native Lig E (28.4 kDa) purified from either constructs, the cleaved protein after a reverse-His trap (gel: Figure 4.7 (a)(b); chromatogram: Figure B.17 (a)(b)) was directly frozen due to frequent loss of the protein during size exclusion, while for Lig E-His (29.2 kDa), the cleaved protein was frozen for future assays directly after an MBP-trap to remove the MBP-tag (gel: Figure 4.7Figure B.17 (c); chromatogram: Figure B.17 (c)).
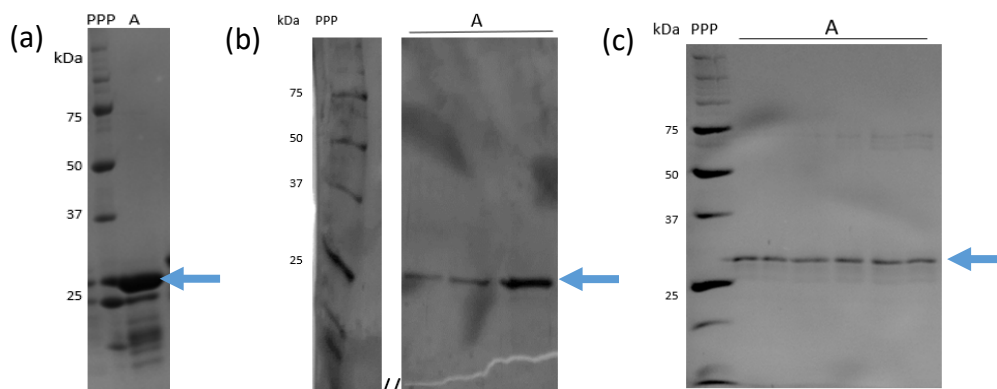
Figure 4.7. SDS-PAGE (12%) of proteins from the final purification procedures performed before freezing (a) Frozen fractions of Lig E from a His-(TEV)-Lig E purification after a reverse-IMAC (b) Frozen fractions of Lig E from an MBP-His-(TEV)-Lig E purification after a reverse-IMAC, the gel of which was cropped ('//') to show only the lanes of interest (c) Frozen fractions from Lig E-His from a His-MBP-His-(TEV)-Lig E-His purification after an MBP-trap. Blue arrows indicate the bands of interest. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled. Chromatograms corresponding to these gels can be found in Figure B.17.

4.2.2.1 Native Lig E purification

Despite the straightforward purification of Lig E-GFP, the purification of Lig E from the His-(TEV)-Lig E protein was performed several times due to the multiple difficulties encountered. For example, the IMAC produced a broad elution peak of the protein with buffer B (approximately 25%), even eluting over 196 mL from a 50 mL sample (Figure B.18 (a)). Thus, significantly higher imidazole is recommended for future IMAC purifications for Lig E. Another consistent issue observed was the elution of the pre-cleaved tagged Lig E protein with, or after the conductivity peak during the desalting procedure (Figure B.18 (b)). The co-elution of the salt led to inefficient TEV cleavage of the protein (peak B; uncleaved vs cleaved protein: 32.6 vs 28.4 kDa) as opposed to that (peak A) which eluted before the salts. However, the intensity of peak A was low (50 mAu), and hence the amount of cleaved protein was not of significant quantity for some downstream processes like crystallisation of the native protein. Furthermore, on several occasions, the protease did not perform as expected and incubation of the protein with TEV observed no cleavage.

As mentioned earlier, size exclusions of Lig E after a reverse-His procedure resulted in no peaks at all. Hence, the protein from the reverse-His (peak A in Figure

B.17 (a)) was isolated and frozen to prevent any further lost of the native Lig E, despite the additional lower molecular weight protein bands observed via gel electrophoresis. The loss of Lig E during size exclusion may be associated with a loss in stability of the protein throughout the purification process, leading to degradation or precipitation of the protein. Although freezing the cleaved protein from reverse-His secured protein stocks for future work, caution has to be placed on the lower molecular weight contaminants that were not separated out via size exclusion.

4.2.2.2 Lig E-His purification

An IMAC, desalt and TEV cleavage of the His-MBP-His-(TEV)-Lig E-His protein (uncleaved and cleaved Lig E-His: 73.9 and 29.2 kDa; cleaved MBP-tag: 44.7 kDa) (results not attached) saw successful isolation of Lig E-His due largely to the increased solubility from the MBP-tag. In the case of Lig E-His, a size exclusion was performed directly after the TEV cleavage instead of a reverse-His procedure due to the presence of a His-tag on the cleaved protein (Lig E-His), the uncleaved protein (His-MBP-His-(TEV)-Lig E-His) and the cleaved tag (His-MBP-His), that would allow all three versions of the protein to interact with the nickel column.

Initial size exclusion performed with this protein on the NGC system (Figure B.19 (a)) showed efficient removal of the MBP-tag away from the cleaved protein (peak B vs peak A). However, the cleaved and uncleaved proteins eluted together, which was unable to be separated via a cation exchange (Figure B.19 (b)). This cation exchange was performed using a negatively charged resin and MES buffer at pH 6.15 to allow the proteins to be positively charged (isoelectronic point (pI) uncleaved: 8.89, pI cleaved: 10) and elute at different times. A size exclusion with higher salt content (500 mM NaCl) (Figure B.19 (c)) to neutralise any surface charges that may allow the proteins to associate with each other was also inefficient at separating the cleaved and uncleaved proteins.

In an attempt to isolate the protein via a different method, a reverse MBP-trap was performed (Figure B.17 (c)), where the uncleaved protein and cleaved tag containing the MBP-tag would bind to the column, and the cleaved would

elute first. Initial fractions from this process contained only the cleaved protein, as desired, which were immediately frozen and kept for ligation assays. However, there were still a considerable number of fractions from this process in which both versions of Lig E-His eluted together, warranting future investigations of the possible multimeric nature of Lig E.

4.2.3 Ngo-Lig activity assays

4.2.3.1 Effect of C-terminus tags on Ngo-Lig activity

In Chapter Three, a His- and GFP-tag were designed for *in vivo* insertion at the C-terminus of *Lig E* in *N. gonorrhoeae* with the intention of using these mutants to track the cellular location of Ngo-Lig, although the latter observed no positive results. To determine the effects of these tags on the activity of Ngo-Lig, gel-based DNA ligation assays, established by other DNA ligase researchers, were performed using the purified Ngo-Lig proteins from this chapter (Cheng & Shuman, 1997; Tang *et al.*, 2003; Magnet & Blanchard, 2004; Wang *et al.*, 2013; Williamson & Pedersen, 2014). In this assay, a 5' oligonucleotide sequence containing a fluorescent 6-fluoroscein moiety (in this case, the 20-mer, L1) would be ligated to a 3' phosphorylated oligonucleotide sequence (the 20-mer, L2) after the two had been annealed onto the complementary template sequence (the 40-mer, L3), serving as a singly nicked dsDNA template to assess Ngo-Lig activity after separation on a denaturing urea gel (Figure 4.8) (Williamson *et al.*, 2014). Successful ligation of these constructs would yield a 40-nucleotide double-stranded product.
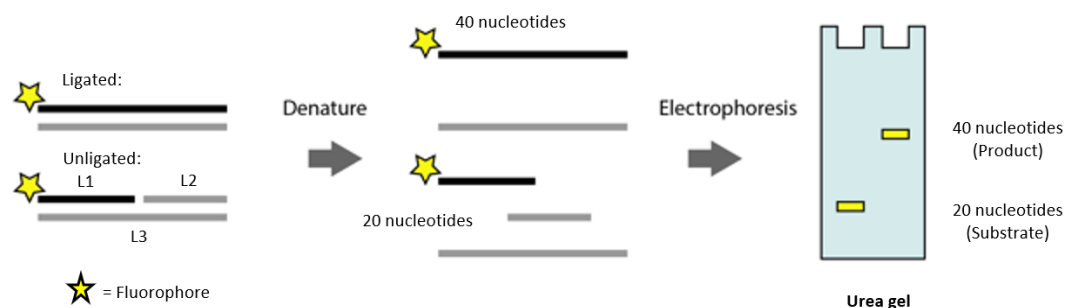
Figure 4.8. Scheme of the gel-based ligation assay procedure performed using a singly-nicked double-stranded substrate with a nick between the L1 and L2 constructs

Due to the different purification outcomes for each Ngo-Lig protein variant, there were considerable differences in the relative concentrations of frozen protein stocks available for assays (Figure B.20), with the native Ngo-Lig protein (from pDEST17) having the highest protein concentration due to the shorter purification process performed (frozen from reverse-His). Quantification of the frozen protein stocks was previously performed via their absorbance at 280 nm (using the NanoDrop™ machine) (Table B.4). However, as the excitation wavelength range for GFP overlaps with this (maximum at 390 nm, with a second peak at 250-300 nm), the Bradford method of quantification which measures protein absorbance at 595 nm was used for quantification instead (Bradford, 1976; Yakhnin *et al.*, 1999). Based on the equation of the BSA standard curve (Figure B.21) and the Bradford absorbances of the frozen proteins (Table B.5), the concentrations of all three proteins were calculated (Table B.6) and used in the subsequent gel-based ligation assay analysis (Figure 4.9).
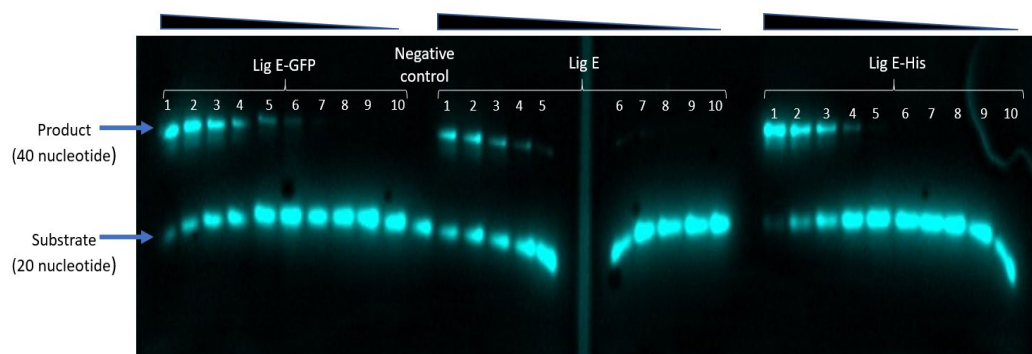
Figure 4.9. Denaturing urea gels (20%) of different Ngo-Lig variants after gel-based activity assays, where the higher band represents the fluorescent products and the lower band represents the fluorescent substrates. The concentrations of proteins for each variant in ascending order (indicated by the thickness of the triangle) were: Lig E-GFP (1 to 10) - 93.7, 74.6, 62.1, 46.6, 34.5, 26.6, 18.6, 9.3, 3.7, 1.8 μM; Lig E (1 to 10)- 82.1, 65.7, 54.7, 41.0, 30.4, 23.4, 16.4, 8.2, 3.2, 1.6 μM; Lig E-His (1 to 10) - 90.6, 72.4, 60.4, 45, 33.5, 25.8, 18.1, 9.0, 3.6, 1.8 μM. The inconsistencies in gel running at the ends of the gels (smileys") are an artefact of the edge wells and do not indicate a change in DNA size.

As expected, the higher the concentration of each Ngo-Lig variant, the higher the ligation activity, indicated by the increased amount of ligation product (Figure 4.9). Likewise, the lower concentrations of proteins for all three Ngo-Lig variants had lower ligation activities, characterised by the intensity of the lower bands which correspond to the smaller DNA products (unligated L1) that is also the only band

observed in the negative control. As it was difficult to observe any significant differences between the three protein variants by eye, this assay was performed in replicates of six and the bands were quantified in ImageJ to calculate the ratio of ligated products (higher band) over the total fluorescence detected (ligation efficacy (%)) (Figure 4.10).
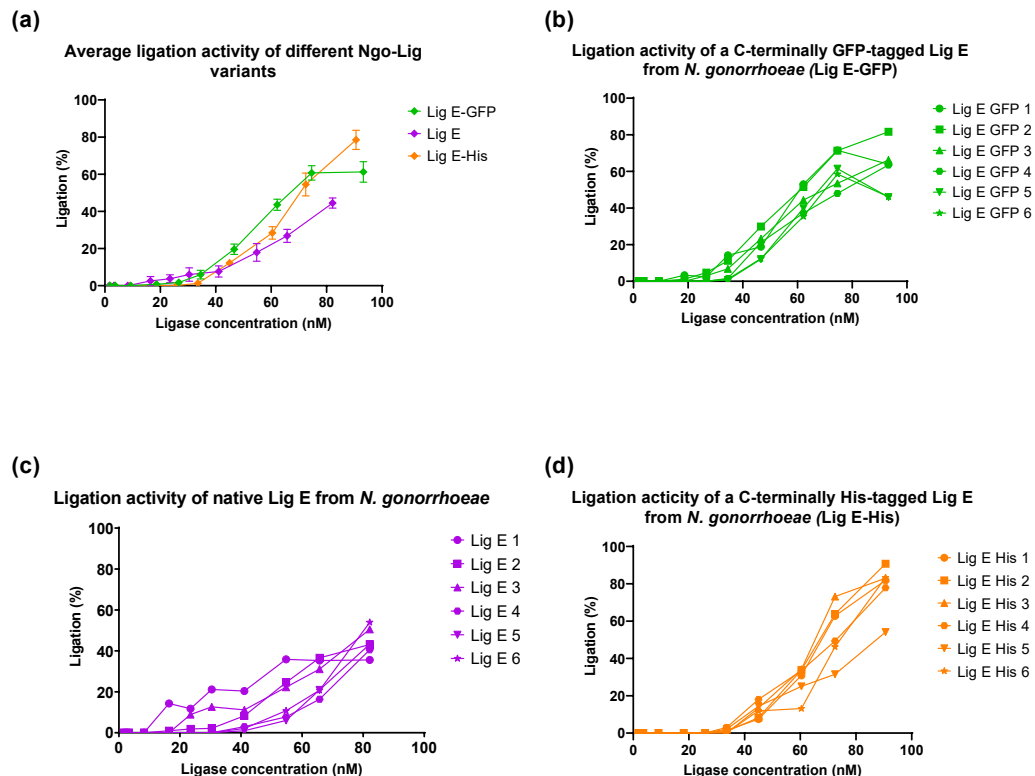
Figure 4.10. Quantification of the ligation activity of different Ngo-Lig variants (with six replicates each) via ImageJ, with standard errors plotted (a) Average results of all three proteins variants (b) Individual replicates of Lig E-GFP (c) Individual replicates of Lig E (d) Individual replicates of Lig E-His

Based on Figure 4.10 (a), a general trend of increasing ligation activity as the concentration of protein increases can be observed, with a linear range for all three variants (Figure 4.10 (b-d)) between 40-90 nM of protein. What was particularly interesting however, was the lower rates of ligation for the native Ngo-Lig protein compared to the other two-tagged variants. For example, based on the trend in Figure 4.10 (a), at a concentration of 80 nM enzyme, the ligation activity of the native Ngo-Lig is predicted to be approximately 40%, while the efficiencies of Lig E-GFP and Lig E-His would be around 60% and 70% respectively. Possible reasons

for this include the Bradford method of protein quantification used, which has low detecting limits and is dependent on the specific composition of the protein, especially in relation to its lysine and arginine composition (Bradford, 1976; Noble *et al.*, 2007). On the other hand, the purification processes used for all three Ngo-Lig variants also differed from each other. To determine if either the purification method or the origin of the protein had any effects on the ligation activity, the assay was repeated with the native Ngo-Lig protein purified from the pHMGWA plasmid (Figure 4.11).

Comparisons of the ligation activities of different Ngo-Lig variants obtained through different purification methods
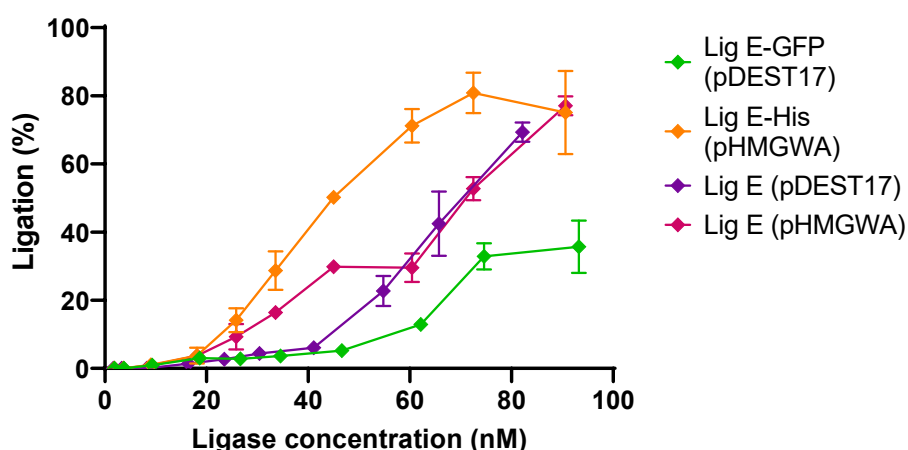
Figure 4.11. Average results of the ligation activity of different Ngo-Lig variants from different destination plasmids (with three replicates), quantified via ImageJ

Based on Figure 4.11, it is clear that the two native Ngo-Lig proteins from different sources had very similar ligation efficiencies to each other (little batch to batch variation), while the Lig E-His variant still had a higher activity. In contrast to the earlier graph in Figure 4.10, the Lig E-GFP variant had the lowest efficiency, which is likely due to its degradation and decrease in protein stability over time in the freezer (-20 °C) that underwent multiple freeze-thaw cycles compared to the fresh aliquots used for the other three proteins. As this indicates that the protein does not tolerate multiple freeze-thaw cycles, the stability of the three Ngo-Lig variants should be examined in the future. Although the native Ngo-Lig protein derived from the pHMGWA plasmid was used as a control for the extensive purification process for Lig E-His, it must be noted that the purification processes for both were not the

same. For Lig E-His, the purification involved an IMAC, a desalt, TEV cleavage, a high salt size exclusion and an MBP-trap while Ngo-Lig derived from the same plasmid underwent a purification process very similar to Ngo-Lig from pDEST17, with an IMAC, a desalt, TEV cleavage and a reverse-His trap. Perhaps a better explanation for the low activity of both native Ngo-Lig proteins relative to the His-tagged variant might be the underestimation of the concentration of Ngo-Lig, owing to the low molecular weight impurities from the reverse-His step that may have affected protein quantification (Figure 4.7). It is also possible that *E. coli* DNA may have been present before purification and was interfering with the activity assays (Williamson & Pedersen, 2014).

Despite this, the main aim of this experiment was to determine if the His- and GFP-tags had any effects on the activity of Ngo-Lig. Due to its small size, the His-tag was predicted to have very little effect on the function of the protein. However, there are discrepancies in literature about the effect of this tag on various proteins, with some groups finding a decrease in thermal stabilities and catalytic activities of their His-tagged proteins, and others finding increased stabilities and activities of their His-tagged proteins, although only in the presence of polar organic solvents, which was not included in our purification buffers (Booth *et al.*, 2018; de Almeida *et al.*, 2018; Meng *et al.*, 2020). However, these results are very dependent on the specific proteins and conditions used during the experiment. In our case for example, it is possible that the positive His-tag had slightly increased the affinity of Lig E-His to the negative DNA strands, which also aligns with the slightly higher rates of growth for the Lig E-His-KanaR mutant over the WT *N. gonorrhoeae* observed in Chapter Three.

On the other hand, the effects of a C-terminal GFP-tag on decreased protein function and stability is often associated with its large size, causing steric hindrance and interfering with the protein's ability to fold properly and hence interact with other proteins (Zhang *et al.*, 2014; Sokolovski *et al.*, 2015). In this case however, the activity of Lig E-GFP seemed similar to that of Lig E-His *in vitro* (Figure 4.10 (a)) although its effects *in vivo* were not observed due to lack of successful MS11 transformations. Despite this, the results from this experiment show that the C-terminal His-tag added *in vivo* onto Ngo-Lig in Chapter Three, and any future work

with the GFP mutant, do not seem to be affecting the ligating activity of Ngo-Lig *in vitro*.

4.2.3.2 Optimal conditions for Ngo-Lig ligation activity

4.2.3.2.1 *Optimal ATP concentration and pH for Ngo-Lig activity*

As Ngo-Lig is an ADL, ATP is required to activate the ligase in the first step of ligation. To characterise its dependence on ATP as a cofactor, the gel-based activity assay was repeated with Ngo-Lig with increasing concentrations of ATP (0, 0.3, 0.5, 0.75, 1.0, 2.5, 5 and 7.5 mM), where the volume of ATP added was kept constant. Based on Figure B.22, ligation activity was detected even at 0 mM ATP (33%). It has to be noted however, that a proportion of the enzyme was already pre-adenylated from the overnight incubation of the protein with ATP (0.1 mM) before purification (Williamson & Pedersen, 2014). By normalising the results to the ligation efficiency at 0 mM ATP (Figure 4.12 (a)), it is clear that the addition of ATP significantly increases the ligation activity of Ngo-Lig by approximately 35%. In fact, this increase in activity was observed with increasing ATP concentration up to 1.75 mM ATP, before this effect plateaued. Throughout the project, an ATP concentration of 1 mM had been used for the ligation assays based on previous work by Cheng and Shuman (1997), although the results obtained here show that increasing the concentration of ATP to up to 2 mM may lead to even greater ligation efficiencies. It would be interesting to compare these results to the amount of extracellular ATP detected in the growth media that may act as a source of cofactor for Ngo-Lig to deduce when Ngo-Lig is most active *in vivo*.
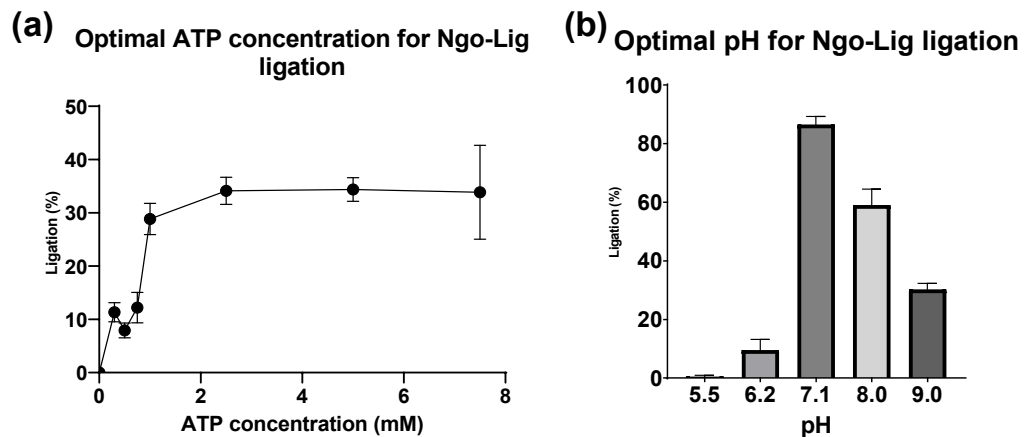
Figure 4.12. Optimal conditions for Ngo-Lig ligation of a singly-nicked dsDNA with standard errors plotted (a) Optimal ATP concentration corrected for the activity at 0 mM ATP (b) Optimal pH

Another factor that must be taken into account is the pH environment for Ngo-Lig to operate. Based on work by other groups, Tris buffer at pH 8.0 was used for prior Lig E ligation assays (Cheng & Shuman, 1997; Williamson *et al.*, 2014). To assess the impact of pH on ligation activity, MES buffer was used for the lower pH ranges (5.5 and 6.2), while Tris buffer was used for the higher pH ranges (7.1, 8.0, 9.0) due to the pH limits of both buffers. Results from this experiment (Figure 4.12 (b)) show that a pH of 7.1 brings about the highest ligation efficiency of Ngo-Lig at 86%, as opposed to the previously used pH 8.0 at 60%. This aligns with the slightly alkaline intracellular pH of human cells that *N. gonorrhoeae* infects (7.0-7.4) and the optimal pH of other ligases (Madshus, 1988). For example, the optimal pH of the standard ATP-dependent T4 DNA ligase used in the laboratory ranges between 7.0-7.8, while that of the NAD^+-dependent thermophilic HB8 DNA ligase is between 7.4-7.6 (Weiss & Richardson, 1967; Takahashi *et al.*, 1984). As the theoretical pI of Ngo-Lig is 10.0, a pH of approximately 7.1 might allow the protein to be sufficiently positively charged to attract the negatively-charged dsDNA substrates. During the growth experiment, changes in media colour, which contained a pH indicator, were indicative of a change in pH as the number of gonococcal cells increased (especially during the death and lysis phases), although this was not quantitatively recorded at that time. Results from this change in pH would also be useful in determining when Ngo-Lig is most active in addition to the amount of extracellular ATP in the supernatant.

4.2.3.2.2 *Optimal substrate for maximum Ngo-Lig activity*

To better understand the mechanism of how Ngo-Lig operates, the activity assay was repeated with five different dsDNA substrates (Figure 4.13 (a)). Although other groups have worked with shorter incubation times for Lig E ligation assays (e.g. 5-10 min for nicked, cohesive and mismatched and 30 min for gapped substrates), 30 minutes was the minimum incubation time performed here due to the low concentrations of purified proteins obtained (Cheng & Shuman, 1997; Williamson *et al.*, 2014). The first substrate used was the standard centrally placed singly-nicked dsDNA (40-mer) used in prior assays. Based on the results in Figure 4.13 (b), Ngo-Lig is able to ligate this singly nicked duplex DNA substrate the most efficiently (approximately 70% efficacy). The second substrate used was one with an overhang or cohesive nick of which Ngo-Lig had approximately 23% ligation efficiency, which is higher than the 4% ligation activity of Ngo-Lig with the third substrate that had a mismatch. In comparison, Ngo-Lig demonstrated no activity with the fourth and fifth blunt-ended and gapped substrates.
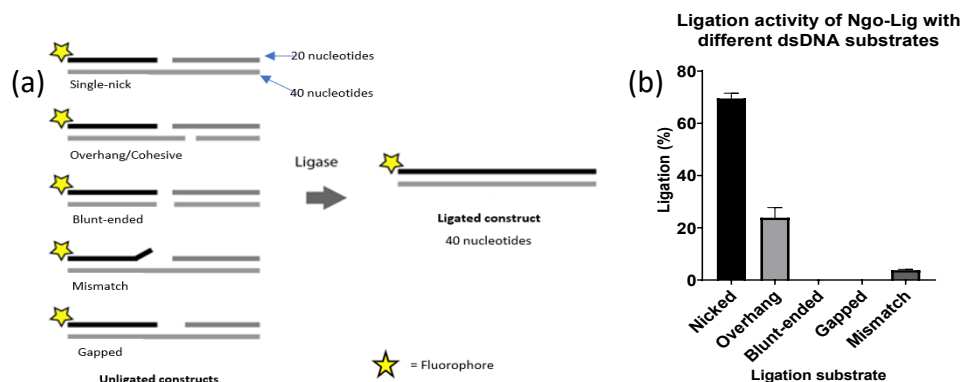
Figure 4.13. Ligation activity of Ngo-Lig with different dsDNA substrates (a) Illustration of the five substrates used in the experiment to yield 40-nucleotide double-stranded substrates (b) Results from the ligation of Ngo-Lig with the five different substrates of interest with standard errors plotted

Results from this experiment very closely agree with the results from other Lig E ligation activities in the literature. In particular, all are capable of sealing a singly-nicked duplex DNA with high efficiencies (this includes Hin-Lig (77% efficacy), Nme-Lig, Psy-Lig, Vib-Lig (80% efficacy) and Ame-Lig) (Cheng & Shuman, 1997; Magnet & Blanchard, 2004; Williamson & Pedersen, 2014; Williamson *et al.*, 2014; Williamson *et al.*, 2018). On the other hand, as the overhang or cohesive substrate

is similar to a singly-nicked substrate, albeit with two nicks, it comes to no surprise that Ngo-Lig is able to seal these nicks as well, although efficiency was much lower than that of the first substrate. This has been shown to be effective in Psy-Lig and Ame-Lig as well (Williamson *et al.*, 2014; Williamson *et al.*, 2018). In contrast, the low ligation activity of Ngo-Lig with the mismatch substrate may be associated with the distortion of the mismatched nucleotide that prevents tight interactions of the ligase with the substrate. This lower efficiency was also observed with Psy-Lig and Ame-Lig but was absent for Nme-Lig, although the presence of the signal peptide on Nme-Lig during these assays may have affected binding of the enzyme to the DNA (note, this reduces ability of Vib-Lig to seal nicks by 2-fold) (Magnet & Blanchard, 2004; Williamson & Pedersen, 2014; Williamson *et al.*, 2014; Williamson *et al.*, 2018).

On the other hand, Ngo-Lig showed no activity with either blunt-ended or gapped substrates, which was also observed for Nme-Lig (Magnet & Blanchard, 2004). What is interesting however, was the low, but detectable activity of Psy-Lig on both these substrates, albeit with higher enzyme concentrations and longer incubation times, and that of Hin-Lig and Ame-Lig with one-nucleotide gapped substrates (Cheng & Shuman, 1997; Williamson *et al.*, 2014; Williamson *et al.*, 2018). Perhaps the reason for this is the minimal structure of Lig E that only allows partial encirclement and limited interactions with the duplex DNA, compared to the complete encirclement in for example, the T4 ligase that is capable of ligating these blunt-ended substrates (Magnet & Blanchard, 2004). On the other hand, the lack of activity of Ngo-Lig with a gapped substrate may also be indicative of its interactions with other proteins that would first fill the gap with the respective base, which would align with the difficulty of other ligases to seal gapped substrates. In contrast to blunt-ended substrates, the activities of Ame-, Hin- and Psy-Lig with gapped substrates show that despite their minimal structures, the partial encirclement of DNA in a C-shaped clamp may allow efficient ligation and twisting of the enzyme compared to the rigid complete encirclement of other ligases that do not allow the 3'-OH sites to interact with the ligase active site (Williamson *et al.*, 2018).

Considering their similarities and evolutionary closeness, it comes to no surprise that the Nme-Lig and Ngo-Lig enzymes show similar ligation activity for singly 5'

phosphorylated nicks of dsDNA. However, differences between Ngo- and Nme-Lig activity with other minimal Lig E like Psy- and Ame-Lig warrants further investigations, especially in relation to their gapped substrate ligation activities. A potential way to do so is to compare their crystal structures to determine interactions at various locations close to the active site that allow the ligase to ligate various substrates, which with further purifications of the native Ngo-Lig to a higher concentration, would be possible. Regardless, results from this experiment, when taking into consideration its potential cellular location, may indicate that Ngo-Lig is sealing nicks in dsDNA rather than joining random pieces of DNA found in the environment that are likely to be blunt-ended. Although the origins for Ngo-Lig substrates are still unknown, these results show that in relation to the next chapter (Chapter Five), it is possible to create overhang substrates for DNA uptake experiments of Ngo-Lig using restriction enzymes, which are easier to prepare (with approximately 4 bp overhangs) compared to a nicked substrate. This will allow investigations onto the role of Ngo-Lig in increasing uptake of antibiotic resistance genes *in vivo*.

4.3 Conclusion

The main purpose of this chapter was to not only express, but also to purify and obtain stocks of the recombinant native Lig E protein from *N. gonorrhoeae*, as well as both Lig E-GFP and Lig E-His. This was to determine the effects of the C-terminal His and GFP-tags on the activity of Ngo-Lig, the addition of which were attempted *in vivo*, and to compare the activity of Ngo-Lig to other recombinant Lig E. Based on the results discussed and presented, these tags did not appear to present any obvious negative effects on the activity of Ngo-Lig. Rather, there was a slight increase in the ligation activity of Lig E-His over the other two variants, which is interesting considering the increase in OD_{600} readings of the Lig E-His-KanaR mutant over the WT *N. gonorrhoeae*. If these differences prove significant, it will indicate that the increased ligation caused by the His-mutant is beneficial for growth, leading to increased cell numbers, and thus hinting at the role of Ngo-Lig in aiding with gonococcal growth (i.e. through biofilm formation). Regardless, we can be confident that any future work with a Lig E-GFP and a Lig E-His mutant *in vivo*

would not, and did not, significantly affect the activity of Ngo-Lig and hence, the results obtained in Chapter Three.

Throughout this chapter, several optimisation steps had to be carried out in terms of both the expression and purification of the three protein variants. In particular, there was difficulty in expressing Lig E-His in pDEST14. Although expression after generation of the His-(TEV)-Lig E-His construct was successful, the addition of a large soluble MBP-tag from the pHMGWA plasmid proved the most efficient at increasing soluble Lig E expression. What was also interesting during the purification process of Lig E-His, was the consistent interaction of cleaved and uncleaved protein, indicating at a possible multimeric nature of the protein. However, previous work with other Lig E proteins hint at a monomeric nature of Lig E. It would be interesting to elucidate this property of Ngo-Lig to better understand its possible interactions with other proteins.

As a result of successful expression and purifications, other ligation assays were able to be performed with the recombinant Ngo-Lig proteins. From here, it was shown that Ngo-Lig is efficient at sealing singly nicked duplex DNA, and to a lesser extent, cohesive and mismatched substrates. Earlier predictions of Ngo-Lig's ability to ligate blunt-ended substrates due to its minimal structure were proven unlikely as no ligation activity was detected. Furthermore, the potential of Ngo-Lig on ligating ssDNA and RNA (the latter of which is evidenced by the activity of the minimal *Chlorella* virus ADL on DNA-RNA hybrids) were not investigated (Lohman *et al.*, 2014). Regardless, results from this chapter opened new avenues for future work with the frozen proteins, which include crystallisation to elucidate the structure of Ngo-Lig to determine any differences between this ligase and those from other bacteria. Apart from allowing for optimal future expression and purification of Ngo-Lig, it also permits for further investigations into DNA uptake *in vivo* as evident by the ability of Ngo-Lig to seal cohesive substrates, thus allowing us to better understand the role of this enigmatic gonococcal enzyme.

Chapter Five

Construction of a reporter construct for DNA uptake in *N. gonorrhoeae*

5.1 Introduction

The discovery of an N-terminal signal peptide that indicates periplasmic translocation of Ngo-Lig, in addition to its ability to seal nicks in dsDNA (Chapter Four) was what led to the hypothesised role of Ngo-Lig in HGT or competence. In particular, it is suggested that Ngo-Lig repairs nicks in exDNA in the periplasm, which increases the information content of that piece of DNA, leading to the integration of that DNA into the genome via homologous recombination, and persistence of multiple antibiotic resistance genes within the community (Magnet & Blanchard, 2004). This hypothesis is supported by the presence of the protein in various naturally competent human pathogens like *H. influenzae* and *V. cholerae*, for which HGT is important to achieve genotypic variation to evade the host immune response (Patel *et al.*, 2011; Low *et al.*, 2014). As antibiotic resistance is accelerated if multiple acquisitions of the gene have occurred in the community, Ngo-Lig's role may be crucial at ensuring the integrity of this piece of DNA that would later be incorporated into the genome (Patel *et al.*, 2011). If this theory holds true, considering the rapid rise of antibiotic resistance amongst *N. gonorrhoeae* rendering it a potential superbug, Ngo-Lig would serve as an ideal target to reduce the spread of this bacteria in the community.

Despite prior discussions of Ngo-Lig's possible role in biofilm formation in the previous chapters (Chapters Three and Four), the broad aim of this chapter was to test the original hypothesis *in vivo*. In particular, comparisons of the uptake of nicked DNA substrates between the WT *N. gonorrhoeae* and the Lig E-KO-KanaR mutant obtained in Chapter Three was of interest to determine any differences in the rates of transformation once the gene expressing Ngo-Lig is disrupted. There were two constructs designed for uptake in *N. gonorrhoeae*; a *GFP* gene where transformation can be monitored via fluorescence, and a *β-lac* gene that confers resistance to ampicillin to determine the efficacy of transformation. The two genes

of interest were cloned into a pMR32 plasmid that contains regions of homology (500 bp) with the MS11 *N. gonorrhoeae* genome to allow insertion of the genes of interest into the intergenic region between the gonococcal *igA* and *trpB* genes (Ramsey *et al.*, 2012). This insertion should not interfere with normal gonococcal functioning as these loci have been used for gonococcal complementation in the past without affecting their infection abilities (Johannsen *et al.*, 1999; Wolfgang *et al.*, 2000; Ramsey *et al.*, 2012). The amplified DNA construct was used for spot transformation with WT *N. gonorrhoeae* and analysed via resistance to erythromycin. Hence, this chapter is focussed on the design and optimisation of these *GFP* and *β-lac* constructs for future DNA uptake experiments, as well as pilot transformations of these constructs into WT MS11.

5.2 Results and Discussion

5.2.1 Generation of reporter DNA uptake constructs

The two constructs that had to be cloned into the pMR32 plasmid were a *β-lac* gene ordered as a gene block with an EcoRI restriction site at the 5' end and a XhoI site at the 3' end (Figure C.2 (a)), and a *GFP* gene. As a *GFP* gene was already available from transformations in Chapter Three (Figure A.2), the same construct was used in this experiment. For cloning into pMR32, new primers with an EcoRI site on the 5' forward primer (GFP_Fd_amp) and a XhoI site on the 3' reverse primer (GFP_rev_amp) were used to amplify this GFP construct, which was verified via a 1% agarose gel run (Figure 5.1) (737 bp product, as observed), the bands of which were extracted and purified.
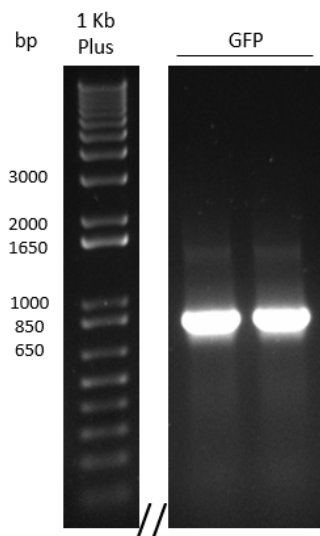
Figure 5.1. Agarose gel (1%) of the *GFP* construct after addition of the restriction enzyme sites, EcoRI and XhoI via the GFP_amp_Fd and GFP_amp_Rev primers. The gel was cropped ('//') to show only the lanes of interest. Sizes of the 1 Kb Plus ladder used are labelled.

Three different constructs were then digested with EcoRI and XhoI; the pMR32 vector plasmid isolated from *E. coli* stocks (digestion of which occurs in between the *igA* and *trpB* flanks, after *ermC*, at a 'gene insertion site' as in Figure C.1), the *β-lac* gene block construct and the *GFP* construct from Figure 5.1. The latter two were ligated onto the digested pMR32 vector in the 'gene insertion site' yielding a pMR32:β-lac and pMR32:GFP plasmid (Figure C.2 (a) and Figure C.3 (a) respectively). Amplification of both new plasmids with the Fd_whole_substrate and Rev_whole_substrate primers (products of which are termed constructs 'A') that bound to the *trpB* and *igA* genes respectively, yielded products of slightly higher band sizes than expected when run on an agarose gel (Figure 5.2 (a)), with both bands resolving around 3000 bp (predicted: 2632 and 2438 bp for the pMR32:β-lac and pMR32:GFP plasmids respectively). However, the band for pMR32:β-lac was still slightly higher than that for pMR32:GFP which was expected. Digestion of these plasmids with the EcoRI and XhoI restriction enzymes used earlier showed bands of similar expected sizes (871 and 4710 bp for pMR32:β-lac, 727 and 4710 bp for pMR32:GFP) (Figure 5.2 (b)), although once again, the accuracy of the band sizes could not be determined with absolute certainty.
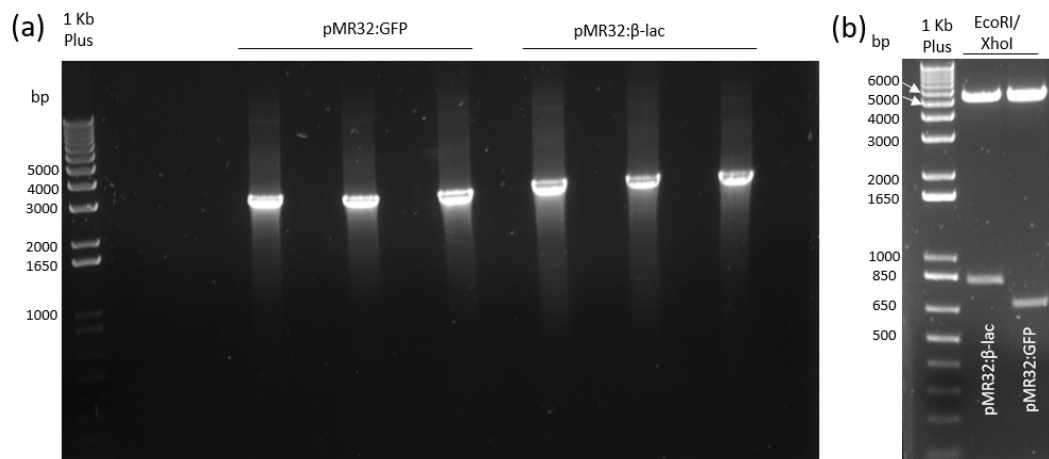
Figure 5.2. Agarose gels (1%) of the ligated pMR32:GFP and pMR32:β-lac plasmids (a) Plasmids amplified with the Fd_whole_substrate and Rev_whole_substrate primers, the products of which are referred to as constructs 'A' (b) Plasmids digested with the EcoRI and XhoI enzymes to confirm ligation efficacy. Sizes of the 1 Kb Plus ladder used are labelled.

Due to the slight differences between the expected sizes and the apparent sizes on the gel (Figure 5.2 (a)), these PCR products were extracted and purified from the gel and sent for Sanger sequencing to check the identity and integrity of these constructs. The sequencing results showed that the *β-lac* gene had been successfully cloned into the pMR32 plasmid, as most of the gene aligned to the expected sequence (Figure 5.3 (a)). On the other hand, only part of the 3' end of the *GFP* gene in the pMR32:GFP construct aligned to the reference sequence as the rest of the sequences were of poor quality (Figure 5.3 (b); highlighted in red). Regardless, it was believed that succesful ligation of the two genes into pMR32 had occurred. Possible explanations for prior variations in sizes when run on an agarose gel may be associated with the possible instability of the ligated plasmids in *E. coli*, leading to slight rearrangements of the plasmid at other locations (Hashem *et al.*, 2002; Chen *et al.*, 2017). This is plausible considering the nature of the plasmid and its optimisation for *N. gonorrhoea* codon usage, which due to its difference to *E. coli*, may be toxic to the bacterium (Kimelman *et al.*, 2012).
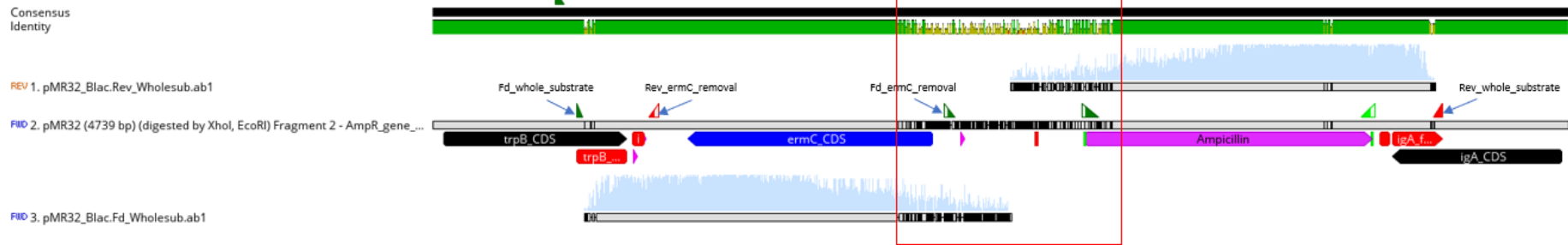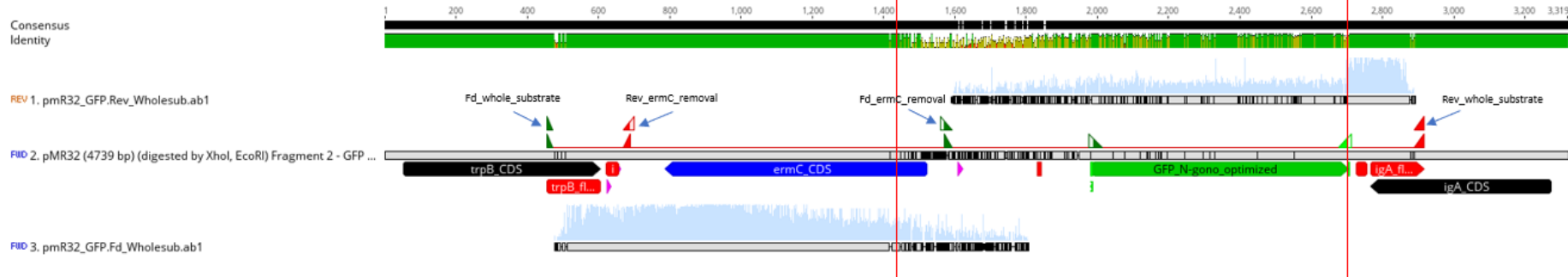
Figure 5.3. Alignment of the intended ligated pMR32 constructs 'A' sequences constructed in silico via Geneious (middle for each) with the forward (Fwd, Fd_whole_substrate primer) (bottom) and reverse (Rev, Rev_whole_substrate primer) (top) Sanger sequencing results from the constructs (a) pMR32:β-lac (b) pMR32:GFP. The red boxes highlight regions of poor sequencing quality and the blue arrows indicate primer-binding regions.

Regardless, it was anticipated that these constructs 'A' which consisted of a 150 bp homology region to the *trp B* gene, an *ermC* gene, the genes of interest under the constitutive gonococcal *opaB* promoter, and a 150 bp homology region to the *igA* gene, could be transformed into *N. gonorrhoeae* and selected for via resistance to erythromycin. This was performed for three reasons; to confirm if the constructs would integrate into the intergenic region between *trpB* and *igA* of the genome, to determine the concentration of ampicillin that the *β-lac* gene confers resistance to for future DNA uptake experiments with nicked substrates, and to determine the correct settings for GFP expression detection if the transformations were successful. In the planned future experiment to determine the effect of a disruption of the Ngo-Lig gene on gonococcal uptake, selection of transformants via erythromycin resistance would not be performed, however this was a convenient tool at this stage to check that everything was working as expected.

5.2.2 *N. gonorrhoeae* DNA uptake experiments of the reporter constructs

Spot transformation was performed using the purified pMR32 constructs 'A' from Figure 5.2 (a) with WT MS11 *N. gonorrhoeae*. Colonies were selected for using erythromycin and verified via the Fd_seq_DNA_uptake and Rev_seq_DNA_uptake primers (Table C.1) that bound to regions slightly upstream and downstream of the Fd_whole_substrate and Rev_whole_substrate primers used to amplify the constructs. The expected sizes for the WT MS11, the pMR32:β-lac 'A' and the pMR32:GFP 'A' mutants with these primers were 508, 2729 and 2585 bp respectively. However, for all colonies tested (79 and 65 colonies for pMR32:β-lac 'A' (Figure 5.4 (a)) and pMR32:GFP 'A' (Figure 5.4 (b)) respectively) either no bands appeared, or the 508 bp band suggest selection of the WT variant. No bands were observed in the negative controls indicating no contamination of the PCR mix. Collectively these results demonstrate the lack of successful integration of these genes into the intergenic region between *trpB* and *igA*. The lack of transformation overall was also verified using internal primers within the construct to check if the genes had integrated elsewhere in the genome (Figure C.4), for which no expected bands were observed. Considering the lack of genes conferring resistance to erythromycin in MS11, it is thought that the colonies selected from the

erythromycin plates were contaminants, and pieces of DNA from lysed WT cells were detected instead, leading to the WT band in these colony amplifications.

(a) pMR32:β-lac 'A' transformation
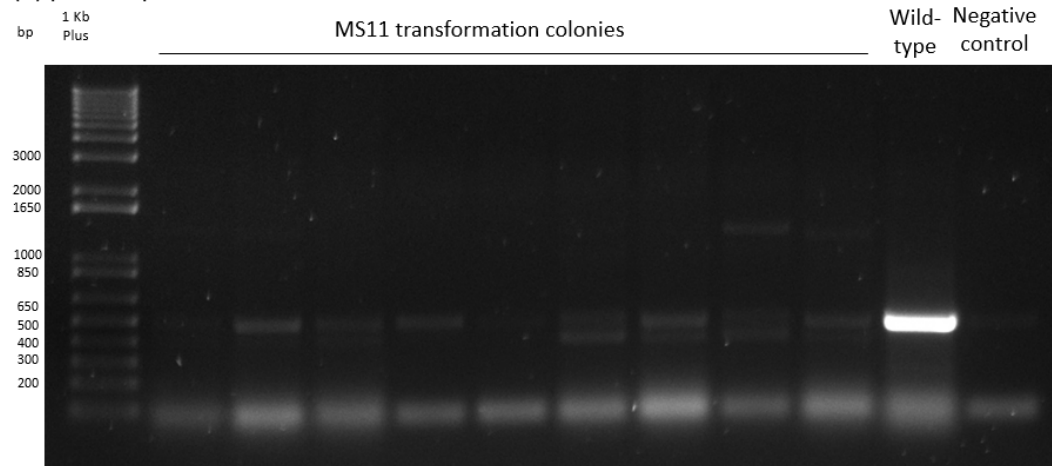
(b) pMR32:GFP 'A' transformation
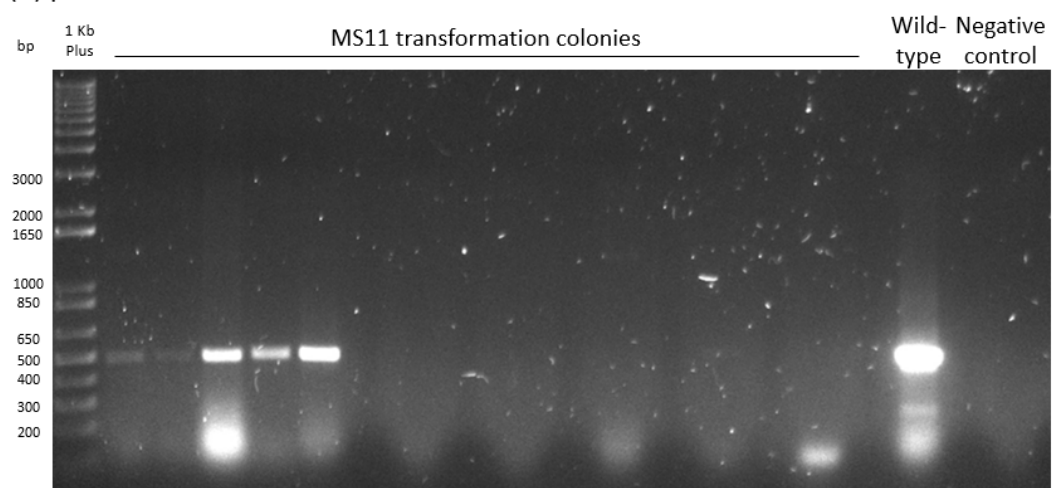
Figure 5.4. Agarose gels (1%) of colony polymerase chain reactions after MS11 *N. gonorrhoeae* transformations of (a) pMR32:β-lac construct 'A' and (b) pMR32:GFP construct 'A'. Colonies were picked from erythromycin GCB plates. Sizes of the 1 Kb Plus ladder used are labelled.

The presence of contaminants would greatly complicate detection of successful transformations due to the high background of contaminants that may be chosen instead. Thus, a Gram-stain was performed on both the WT *N. gonorrhoeae* and the colonies growing on the erythromycin GCB plates. Based on these results, it is clear that the WT gonococcal colonies used for transformation was a Gram-negative diplococcus (purple diplococcus) (Figure 5.5 (b)) which correlates to the morphology of *N. gonorrhoeae*. However, the characteristics of the colonies that

grew on the erythromycin plates were unclear (Figure 5.5 (a)), as the shape and colour of the organisms were not defined for a majority of the colonies. Furthermore, other contaminants (coccus or rod shaped) can be observed in both the selected colonies. As a result of this, multiple streakings of the WT MS11 *N. gonorrhoeae* glycerol stocks were performed to obtain a pure WT colony for future use.
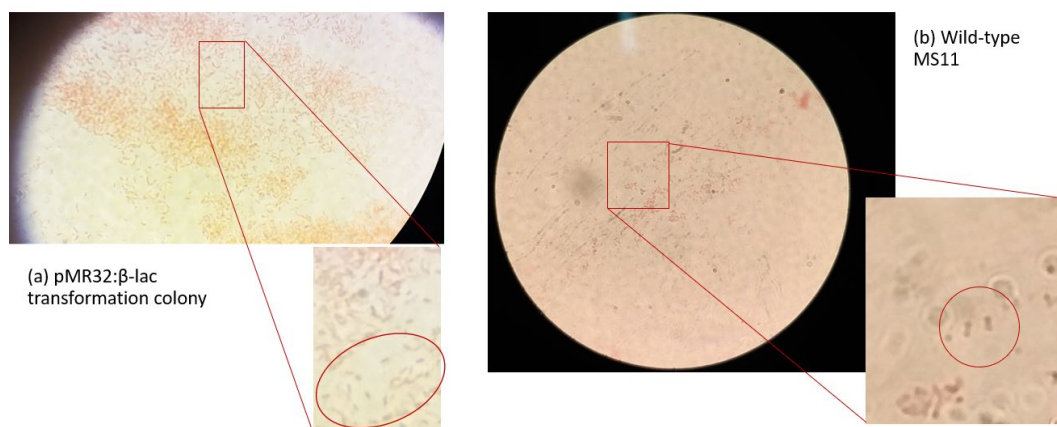
Figure 5.5. Gram-stain of colonies cultured during MS11 *N. gonorrhoeae* manipulation to check for bacterial morphology (a) Colony picked from the erythromycin GCB plate after a pMR32:β-lac construct 'A' transformation (b) Wild-type MS11 grown from glycerol stocks

5.2.3 Increasing homology:insert ratio of the reporter DNA uptake constructs

Despite optimisations of the purity of WT MS11 stocks, transformation with the new stocks yielded similar results to before (Figure 5.4). In an attempt to increase transformation efficacies, new strategies were developed to increase the length of genome homology in relation to the insert length of interest. Previously, in Chapter Three, approximately 150 bp of homology at both 5' and 3' ends of the construct were sufficient for successful integration of the Lig E-His-KanaR and Lig E-KO-KanaR constructs into the gonococcal genome. However, the insertion sequences were shorter than those in this chapter (156-8 bp homology: 880 bp insert for Lig E His-KanaR; 150 bp homology: 859 bp insert for Lig E KO-KanaR; 150 bp homology: 2296 bp insert for pMR32:β-lac 'A'; 150 bp homology: 2152 bp for pMR32:GFP 'A'). It is highly likely that for both pMR32 'A' constructs, the sizes of the insertion sequences were too large in relation to the homologous regions for successful transformation and integration into the genome, especially in

comparison to the recommended 500-1000 flanking regions used by other groups in the past (Dillard, 2011; Ramsey *et al.*, 2012). Two different strategies were adopted to increase the homology:insert ratio in this project. These included creating longer amplified constructs 'A' with new primers to allow for longer homology regions, and linearising the plasmids to use the full 500 bp homology regions for both *igA* and *trpB* during transformations.

5.2.3.1 Creating longer DNA uptake constructs 'A' to increase the length of homology with the gonococcal genome

As part of the first strategy to increase the homology regions of constructs 'A', new primers that bound further out from the *igA* and *trpB* flanking regions of pMR32 were designed to create a 475 bp flanking region to *trpB* (Fd_constructmaking_longerpMR32 primer) and a 497 bp flanking region to *igA* (Rev_constructmaking_longerpMR32 primer), just slightly less than the full 500 bp flanking regions from the pMR32 plasmid (Figure C.1). New verification primers were also designed outside these flanking regions that only bound to the genomic DNA to check for successful transformants (Fd_ sequencingcheck_longerpMR32 and Rev_sequencingcheck_longerpMR32 (Table C.1) (Figure 5.6).
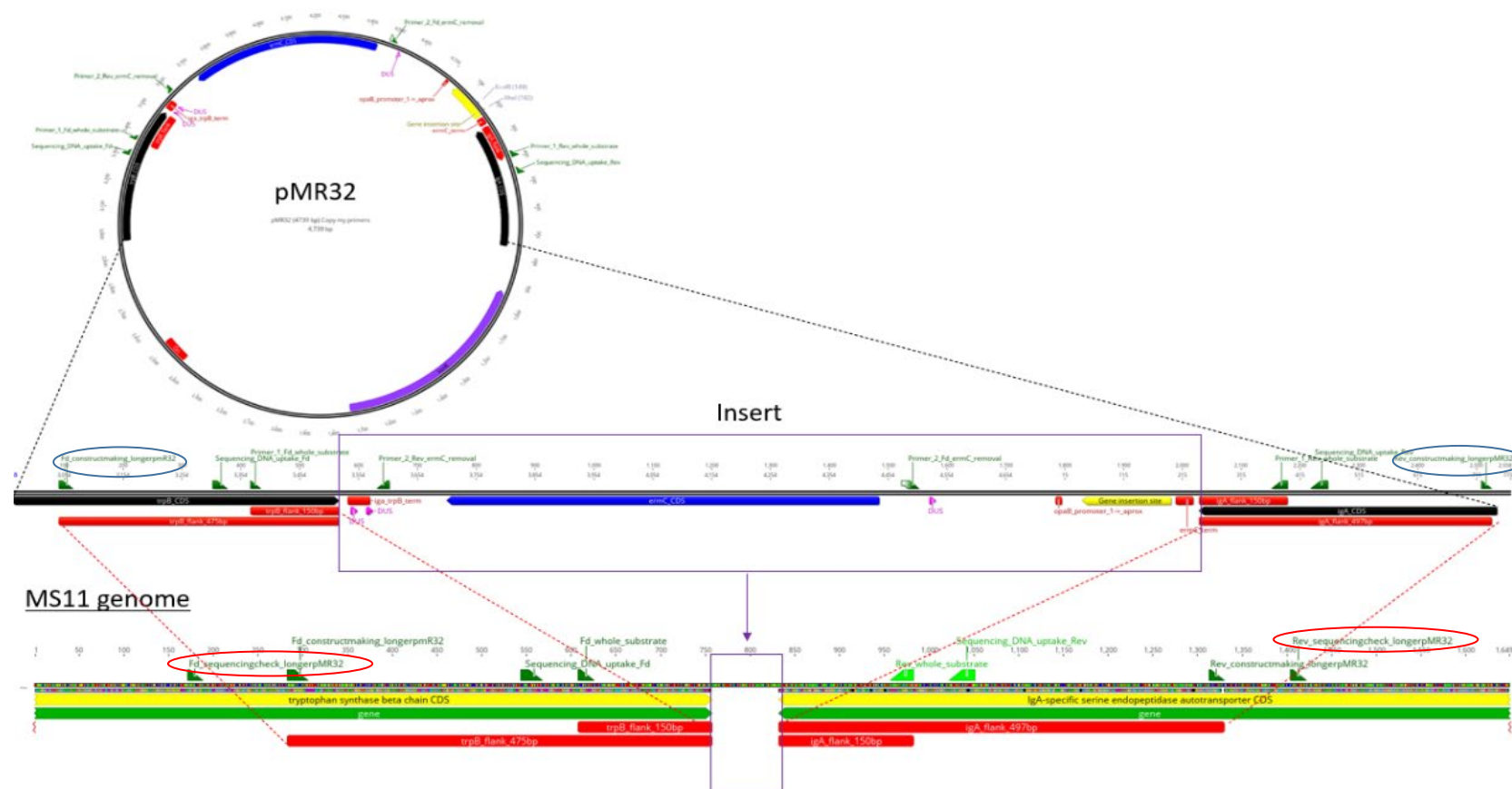
Figure 5.6. Design of the new construct making primers circled in blue (Fd_constructmaking_longerpmr32 and Rev_constructmaking_longerpmr32 primers) and verification primers circled in red (Fd_sequencingcheck_longerpMR32 and Rev_sequencingcheck_longerpMR32) via Geneious to create new pMR32 transformation constructs (longer constructs 'A') with a 475 and 497 bp flanking region to the MS11 *N. gonorrhoeae trpB* and *igA* genes respectively (red dashes) as opposed to the prior 150 bp flanks (constructs 'A'). The purple box represents the insertion sequence.

Amplification of the empty pMR32, pMR32:β-lac and pMR32:GFP plasmids with the new Fd_constructmaking_longerpMR32 and Rev_constructmaking_longerpMR32 primers to create longer constructs 'A' yielded products slightly higher than the expected 2434, 3268 and 3124 bp respectively (Figure 5.7). Similar to the pattern of higher bands observed earlier, despite the accurate sequencing (Figure 5.2, Figure 5.3), it was assumed that these sequences were also as expected and may have been too large for accurate resolution on the agarose gels.
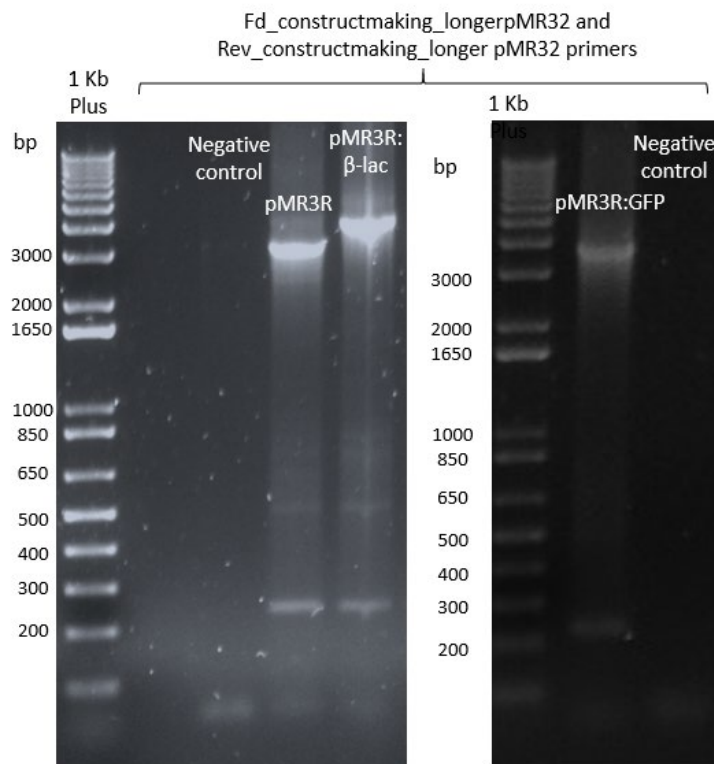
Figure 5.7. Agarose gels (1%) of amplified constructs after the generation of longer constructs 'A' for pMR32, pMR32:β-lac and pMR32:GFP to create constructs with a 475 bp and a 497 bp homology to the gonococcal *trpB* and *igA* genes to increase transformation efficiencies. Sizes of the 1 Kb Plus ladder used are labelled.

Thus, *N. gonorrhoeae* transformations were performed using these longer constructs 'A'. However, consistent growth on the erythromycin plates without any PCR-postive transformants pointed at an underlying contamination issue separate from the earlier contamination of the WT stocks. Repeated transformations of these longer constructs 'A' alongside transformation with the PCR purification elution

buffer found that the elution buffer was the source of contamination. The subsequent treatment of these constructs at 95 °C for 10 minutes before reannealing overnight was successful at eliminating this source of contamination. However, *N. gonorrhoeae* transformations with these treated constructs consistently saw no colonies growing on the erythromycin plates, indicating a lack of successful transformations. Although this may be attributed to the lack of successful reannealing overnight, the higher-than-expected bands from the agarose gels may also point to an underlying issue with the constructs themselves.

5.2.3.2 Linearising the pMR32 plasmid to increase the length of homology to the gonococcal genome

Another option for increasing the homology:insert ratio was to use the whole plasmid for transformation to take advantage of the full 500 bp *igA* and *trpB* homology regions present, which was shown to be successful in *N. gonorrhoeae* in the past (Ramsey *et al.*, 2012). These plasmids must be linearised due to the presence of gonococcal nucleases that would nick the plasmids, making them single-stranded and only allowing for single crossover recombinations (1000-fold decrease in efficiency than double crossover recombinations) (Sox *et al.*, 1979; Ramsey *et al.*, 2012). Although other groups have used both the NheI and PciI sites that occur after the kanamycin resistance gene to do so, these restriction enzymes were not available in our laboratory (Ramsey *et al.*, 2012). Hence the SphI site was chosen as it cleaves close to the NheI and PciI sites in the middle of the kanamycin resistance gene and it only appears once in all three versions of the plasmid (pMR32, pMR32:β-lac and pMR32:GFP). Both pMR32 and pMR32:β-lac were digested with SphI, which was expected to yield a 4739 and 5573 bp linear construct respectively. When run on a gel (Figure 5.8) the undigested circular products resolved slightly lower than expected, which was consistent with them being in the supercoiled configuration as isolated from *E. coli*. This is in comparison to the bands after digestion which ran as predicted, despite the poor running of the digested pMR32:β-lac construct. It is also worth noting that the empty pMR32 plasmid which was amplified with the Fd_whole_substrate and Rev_whole_subtrate primers (expected size: 1762 bp) to yield its own construct 'A', was also run on the same gel in Figure 5.8, with a band much higher than expected (>2000 bp). This will be addressed in

a later section (section 5.2.4). Although digestion was not complete due to the presence of bands that correspond to the undigested plasmid, the linearised pMR32 and pMR32:β-lac constructs were used for further WT *N. gonorrhoeae* spot transformations.
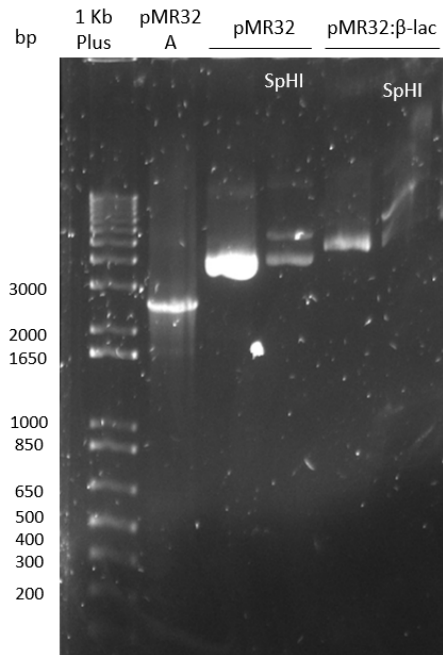
Figure 5.8. Agarose gel (1%) of the pMR32 and pMR32:β-lac plasmids before and after linearisation with SphI. The amplified pMR32 construct 'A' was run at the same time. Sizes of the 1 Kb Plus ladder used are labelled.

Colony PCR of the gonococcal transformants of these linearised plasmids showed the absence of the expected 2636 and 3471 bp bands for pMR32 and pMR32:β-lac when verified using the new Fd_sequencingcheck_longerpMR32 and Rev_sequencingcheck_longerpMR32 primers that sit outside the flanking regions (Figure 5.9 (a)). The presence of the expected 1242 bp band in the WT lane and the absence of bands in the negative control indicate no contamination of the PCR mixtures. This was also observed with the pMR32 construct 'A' of which the expected 1894 band was absent for all colonies, despite the presence of the expected 508 bp band for the WT positive control (Figure 5.9 (c)). These transformations were performed alongside a PilE_KO construct that disrupts the *pilE* gene in *N. gonorrhoeae* via insertion of a kanamycin resistance gene, which was used in prior *N. gonorrhoeae* transformations by other students in this laboratory. Although only one in five colonies that grew was successfully transformed with the PilE_KO

construct, the presence of the expected 1199 bp band in one of the PilE_KO colonies picked (Figure 5.9 (b)) indicates transformability of the WT MS11 stocks used and points at a separate underlying issue in relation to the pMR32 constructs designed. It has to be noted however, that although transformations of *N. gonorrhoea* with linearised plasmids have been successful in the past, issues with methylation due to the preparation of the plasmid from *E. coli* may interfere with the gonococcal transformation process (Russell & Zinder, 1987; Ramsey *et al.*, 2012; Marinus & Løbner-Olesen, 2014). Although this might explain the lack of transformation of the linearised plasmid, the lack of transformation in the amplified PCR constructs indicates that more work must be conducted on characterising any issues associated with the pMR32 plasmid itself.
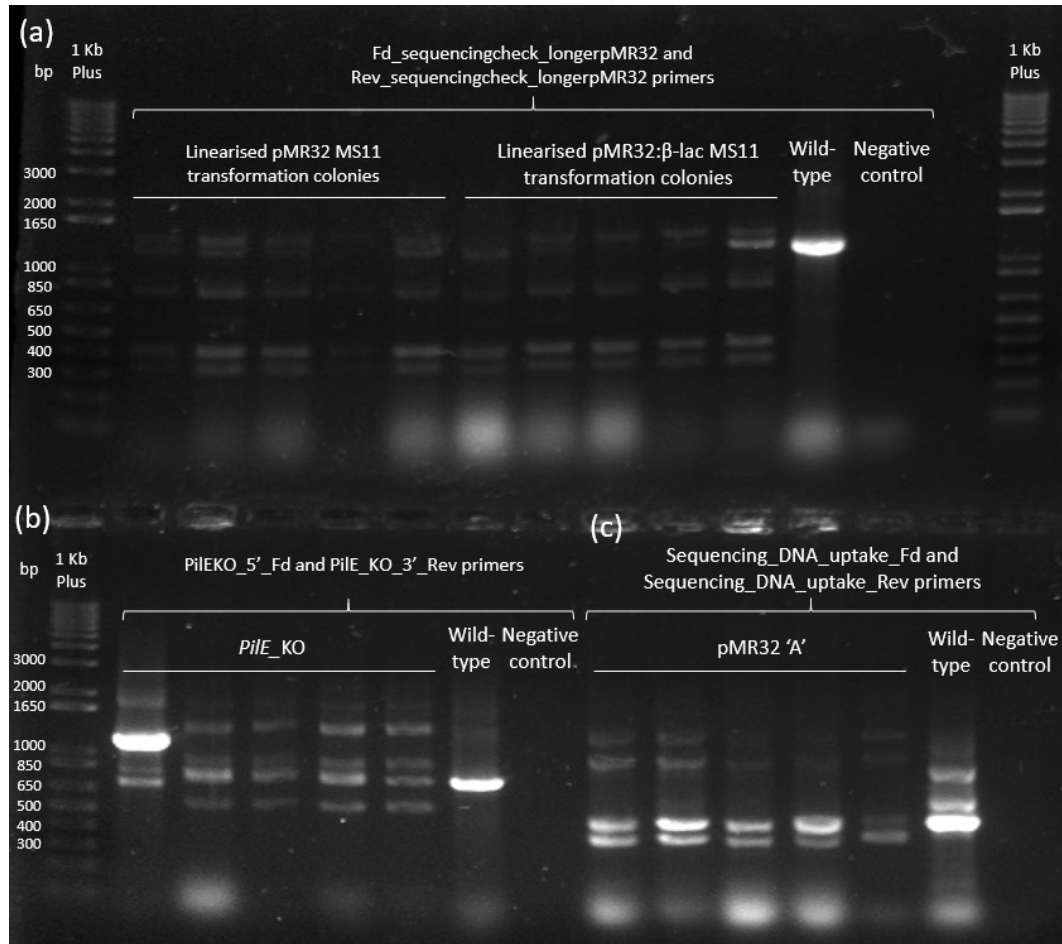
Figure 5.9. Agarose gel (1%) of colony PCRs performed after transformation of MS11 *N. gonorrhoeae* with (a) the linearised pMR32 and pMR32:β-lac plasmids (b) the PilE_KO construct used as a transformation control and (c) the amplified pMR32 construct 'A'. Verification primers for each construct are indicated on the gel. Sizes of the 1 Kb Plus ladder used are labelled.

5.2.4 Characterisation and validation of the pMR32 plasmid

The consistently higher-than-expected bands in the pMR32:β-lac and pMR32:GFP constructs (Figure 5.2, Figure 5.7), which was later confirmed via the blank pMR32 construct 'A' (Figure 5.8), may be responsible for the consistent lack of successful gonococcal transformations throughout this chapter. Although the sequencing results of the pMR32:β-lac and pMR32:GFP constructs 'A' indicated that the sequences of the constructs were as expected, a majority of the sequences were of poor quality, especially towards the end of the constructs. Hence, the blank pMR32 construct 'A' was sent for sequencing using the original Sequencing_DNA_uptake_Fd and Sequencing_DNA_uptake_Rev primers (Figure 5.10 (a)), which showed good alignment of the sequences with the reference sequence, although the part highlighted in the red box were of poor quality. As this part corresponds to the *ermC* gene, there was concern that any disruption in this area may have led to the lack of growth of any successful transformants on the erythromycin plates if the *ermC* gene was not expressed properly.

(a)
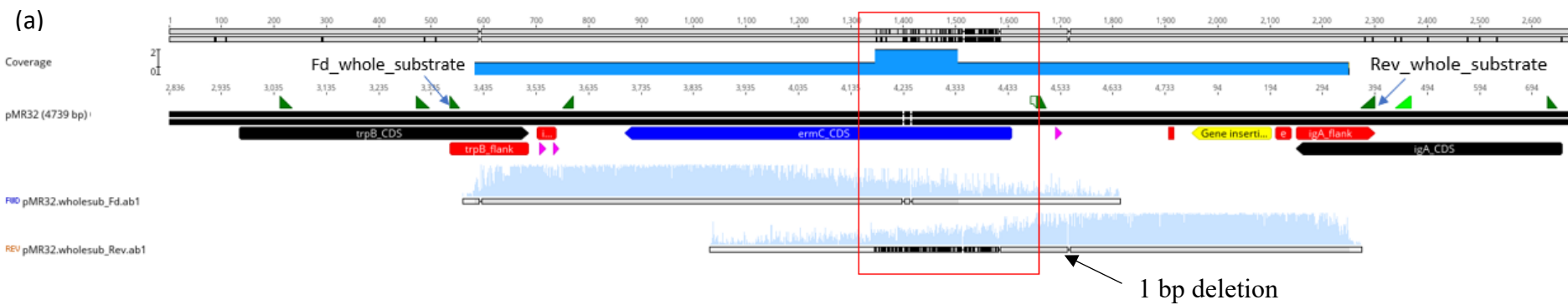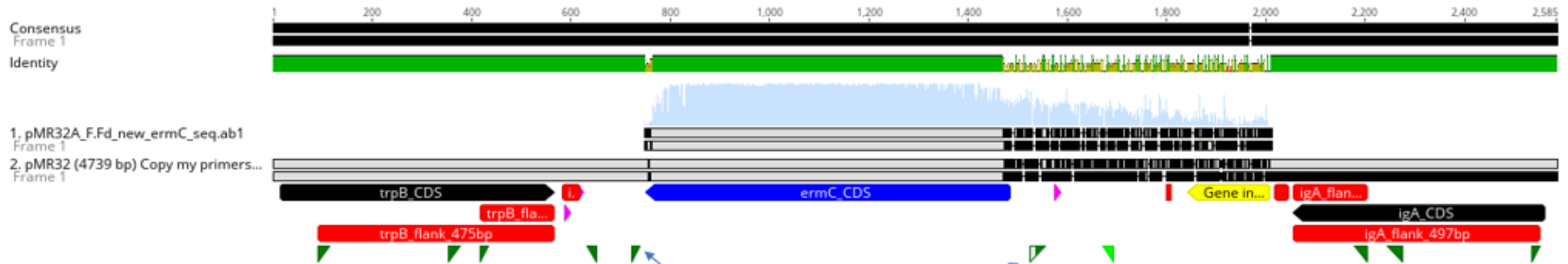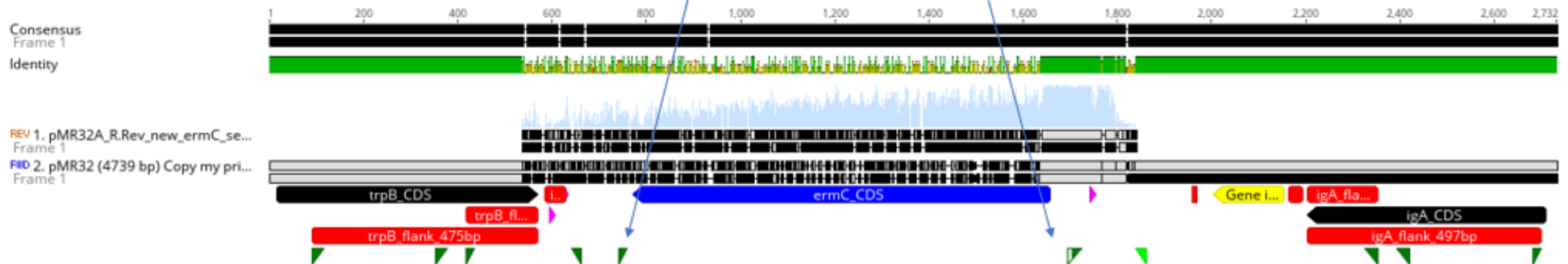
Figure 5.10. Alignment of the reference blank pMR32 construct 'A' sequence constructed in silico via Geneious (top for (a) and bottom for (b)) with the forward (Fwd) and reverse (Rev) Sanger sequencing results from the constructs (bottom) (a) Sequencing with the Rev_whole_substrate and Forward_whole_substrate primers to check construct integrity (b) Sequencing with Fd_new_ermC_seq and Rev_new_ermC_seq primers to check *ermC* integrity. The red box indicates regions of poor sequencing quality while the blue arrows indicate primer-binding regions.

New primers (Fd_new_ermC_seq and Rev_new_ermC_seq primers) were thus designed to check the integrity of the *ermC* gene specifically via sequencing (Figure 5.10 (b)). Results from this suggest integrity of the *ermC* gene. In fact, the only abnormality detected throughout the pMR32 construct 'A' was the 1 bp deletion of an adenine (Figure 5.10 (a)) downstream of a DUS sequence, and upstream of the *opaB* promoter. As both the *ermC* gene and its promoters seem intact, it is unlikely that this 1 bp deletion could affect the expression especially considering the consistently larger PCR products obtained for this construct. Hence, more verification work must be performed with this plasmid in the future to ensure its integrity.

5.2.5 Troubleshooting the DNA uptake construct design and pilot *N. gonorrhoeae* transformations

Throughout the DNA uptake experiment, several complications were encountered. The first was contamination from various sources including the WT *N. gonorrhoea* stock, and the PCR purification kit itself. Although these hindered initial transformant selections, it allowed for the root issue of contamination to be dealt with. On the other hand, several unresolved complications associated with the pMR32 plasmid itself were identified including the consistently higher-than-expected bands despite sequencing results that agree with the reference sequences, hinting at a possible contamination issue with the pMR32 stocks that was unable to be detected via sequencing.

Furthermore, when creating the amplified constructs, the initial 150 bp region of homology was small compared to the suggested 500-1000 bp (or at least 300 bp) for double cross over recombination, especially in comparison to the >2000 bp sequence of interest (Hamilton *et al.*, 2001; Dillard, 2011). In addition to the two attempted methods to increase the homology:insert ratio (including longer homology in constructs 'A' and linearising the plasmids) the removal of the *ermC* gene (via restriction enzyme digestion and ligation) is an alternative method to both increase the homology:insert ratio and to decrease the overall length of insert. Furthermore, a two-step transformation process was also planned for which two separate gonococcal transformations have to be performed. The first transformation

would focus on the pMR32 construct 'A' to allow for the *ermC* gene to integrate into the gonococcal genome. After this, a second transformation with either the pMR32:β-lac or pMR32:GFP constructs 'A' can be conducted with this *ermC* mutant, which would have greater chances of integration due to their additional regions of homology with the integrated *ermC* gene in the gonococcal genome (Ramsey *et al.*, 2012). However, this was not performed as the original pMR32 construct 'A' transformation did not work. Although none of these seemed to transform in *N. gonorrhoeae*, the co-occurrence of contamination on the erythromycin plates creates uncertainty as to the roots of the lack of transformation detection. In addition to this, there was concern of the actual sizes of the constructs which may reduce uptake frequency (Dillard, 2011). However, success of gonococcal transformation with linearised plasmids and constructs as large as 6000 bp have been reported, suggesting that the size of the construct should not be a major issue as long as the homology regions are of substantial lengths (Boyle-Vavra & Seifert, 1993; Ramsey *et al.*, 2012).

The use of the *ermC* gene to select for successful transformants was based on the presence of this gene in between the flanking regions of the pMR32 plasmid designed by Ramsey *et al.* (2012). Prior work by Dillard in 2011 showed resistance of the MS11 gonococcal strain to erythromycin at 10 μg/mL, which was also used in this experiment. However, one should be aware of the differences in resistance found between different gonococcal strains, with the FA 1090 strain showing resistance to erythromycin at concentrations as low as 2 μg/mL (Dillard, 2011). Hence, the concentration of erythromycin used in our experiments must be further optimised as it is possible that slight mutations had occurred in our WT MS11 *N. gonorrhoeae* strain. In fact, work by another group was able to select for MS11 transformants using only 2.5 μg/mL of erythromycin (Kouzel *et al.*, 2015). An alternative method that is currently being designed in our laboratory is the use of a different antibiotic resistance gene, namely the kanamycin resistance gene. Not only has this been proven to work with MS11 *N. gonorrhoeae* transformations in Chapter Three, but this resistance gene has also been efficient at selecting for PilE_KO transformants in this chapter. As noted earlier, a selection marker is included at this stage in order to reliably detect expression of the reporter genes. Constructs without a selection marker would be used in the eventual assay, thus,

the presence of the kanamycin resistance gene in the *N. gonorrhoeae* mutants would not be an issue.

Within the pMR32 plasmid lies a strong, constitutive *opaB* promoter derived from the *opaB* gene from *N. gonorrhoeae* (FA 1090), which should allow easy detection of the genes of interest as they would be constantly expressed at high levels (Ramsey *et al.*, 2012; Ball & Criss, 2013). However, it is possible that the use of this promoter led to overexpression of both the *GFP* and *β-lac* genes, which, as they are not essential for gonococcal survival and function, may have been toxic to the cells. Perhaps a better solution to this is to use inducible promoters to regulate gene expression like the *lac* promoter, which was not only was used in the BL21(DE3)pLysS cells in Chapter Four (inducible by IPTG) but has also been used extensively in previous gonococcal work, leading to a 123-fold increase in expression (Seifert, 1997; Ramsey *et al.*, 2012). Alternatively, one could also use a tetracycline-inducible promoter (inducible by anhydrotetracycline) that allows for low expression of the genes of interest in *N. gonorrhoeae* without affecting gonococcal growth (Ramsey *et al.*, 2012).

Although the plan for this chapter included transformation of the *GFP* gene for quick and easy identification of the transformation success via fluorescence, discussion with other gonococcal-focused groups indicated that the bacterium seemed to lose this gene more so than other genes (data not published; Dillard, 2020, personal communication). This is indicative of the possible toxicity of GFP in *N. gonorrhoeae*. In fact, literature search of the use of GFP as a reporter construct in *N. gonorrhoeae* yielded low results, although one group that had been successful utilised an inducible *lac* promoter for *GFP* expression (Christodoulides *et al.*, 2000; van der Ende *et al.*, 2000; Kouzel *et al.*, 2015). Regardless, as discussed in Chapter Three, the possible toxicity of *GFP* in *N. gonorrhoeae* strongly supports the move to a different fluorescent protein like mCherry, which have been successfully used in *N. gonorrhoeae* in the past (Seitz *et al.*, 2014; Kouzel *et al.*, 2015). Furthermore, as with the GFP protein, issues with folding and transport of the β-lactamase protein to the periplasm may come into play, as the protein only functions in the periplasm (Seitz *et al.*, 2014). Hence, it is also possible to use a different antibiotic resistance gene for future DNA uptake experiments.

5.3 Conclusion

The aim of this chapter was to design and optimise constructs to investigate the role of Ngo-Lig in DNA uptake *in vivo*, as well as to perform pilot transformations of these constructs with WT MS11 *N. gonorrhoeae* to optimise both the ampicillin concentrations and fluorescent settings for GFP detection. In particular, the pMR32 plasmid designed by Ramsey *et al*. (2012) was used to insert the genes of interest in an intergenic region between the *igA* and *trpB* loci to minimise disruption of the gonococcal phenotype. However, several issues with the constructs used for transformation were identified, including the lengths of the constructs and the low homology:insert ratio. Despite attempts to optimise these, underlying issues in relation to the possible protein overexpressions due to the use of the strong, constitutive *opaB* promoter, and complications observed in relation to erythromycin resistance gene, were present.

Despite the lack of successful transformations, this chapter served as a critical pilot study for future DNA uptake experiments using pMR32 in *N. gonorrhoeae*. From here, various methods to increase the construct homology length have been attempted and can be used for future transformations now that contamination has been eradicated from our setup. Furthermore, this chapter points at the difficulty of *GFP* to be expressed in *N. gonorrhoeae* which was also explored in Chapter Three. With further optimisations of the pMR32 constructs, transformations of these reporter constructs can be performed to elucidate the role of Ngo-Lig with DNA uptake in the near future.

Chapter Six

Conclusion and future research

The discovery of the first b-ADL identified in *H. influenzae* was puzzling in terms of its exact purpose when one considers the co-occurrence of essential bacterial NDLs. Since then, a variety of b-ADLs have been elucidated, allowing researchers to label these b-ADLs as accessory enzymes, aiding in repair and replication through different specific mechanisms. Arguably, the most enigmatic of b-ADLs has to be Lig E. Not only is it the most minimal in terms of structure, this enzyme also contains an N-terminal signal peptide, which is strongly suggestive of its translocation outside the cell where it joins DNA without any accessory domains. Due to this, we hypothesised that Lig E localises in the periplasm where it aids in sealing nicks found in exDNA to optimise the integrity of the DNA for homologous recombination of new genes into the genome.

Although a wealth of information on the *in vitro* activities of Lig E are available (Cheng & Shuman, 1997; Magnet & Blanchard, 2004; Williamson *et al.*, 2014; Williamson *et al.*, 2018), current information of its function and location *in vivo* is limited. The purpose of this thesis was to fill these gaps in knowledge to understand this protein. Specifically, focus was placed on the naturally competent bacterium, *N. gonorrhoeae*, an increasingly resistant human pathogen that causes the sexually transmitted disease, gonorrhoea, which is the root of a growing public health crisis. What is most interesting is the presence of Lig E in the genomes of other human pathogens as well like *H. influenzae, C. jejuni, N. meningitidis* and *V. cholerae*. It is hoped that by elucidating the function of this protein, we can target its pathway as an alternative treatment against many human pathogens. In particular, this thesis focused on the generation of Ngo-Lig mutants and the characterisation of these mutants *in vivo* and *in vitro*, as well as the optimisation of DNA uptake constructs to test the original hypothesis *in vivo* in the future.

As part of Chapter Three, two Lig E mutants were successfully generated in *N. gonorrhoeae*; a Lig E-His-KanaR mutant, and a Lig E-KO-KanaR mutant. To characterise the effects of these mutations on gonococcal growth and survival, a

growth experiment with these mutants was performed with WT MS11 *N. gonorrhoeae* acting as a control. In particular, slower gonococcal growth was observed after Ngo-Lig was disrupted. This very distinct phenotype for the Lig E-KO-KanaR mutant, coupled with its significant increase in the amount of exDNA over time (which was not significant for the other two variants) was puzzling in relation to our initial hypothesis; why would a disruption of a competence-related protein slow bacterial growth in its normal growth conditions where no stressors were introduced? Thus, a new hypothesis was formed on the function of Ngo-Lig. In particular, we suggest that Ngo-Lig may be important for biofilm formation, an extracellular matrix composed of pieces of DNA that aid with attachment and infection of gonococcal cells (Greiner *et al.*, 2005; Steichen *et al.*, 2008), and that the disruption of this may have led to the lag of growth observed through interruption of cellular aggregation and interactions (Kwiatek *et al.*, 2014). A role for Ngo-Lig in this process is plausible considering the multiple mechanisms available for *N. gonorrhoeae* to excrete pieces of DNA into the environment (i.e. autolysis, the T4SS system).

To further investigate the function of Ngo-Lig, including a potential role in biofilm formation, future experiments that utilise the mutants produced in this thesis have been planned. The first is an infection experiment with a human cervical cell line to compare infection rates of the Lig E-KO-KanaR mutant and WT *N. gonorrhoeae*, as biofilms are critical for gonococcal attachment to human cells (Spence *et al.*, 1997). This is particularly important considering the pathogenic nature of *N. gonorrhoeae* and would probe a role for Lig E in infection progression in general. Secondly, transcriptomic analysis of the samples collected during the growth experiment would allow us to observe any differences in gene expression that occurs when *Lig E* is disrupted to determine if it is part of a major operon (Kwiatek *et al.*, 2014). As DNase degrades the exDNA and hence affects the biofilm, a repeat of the growth experiment with special attention to the conditions of growth at various time points will be indicative of phases where DNase is inhibited (i.e. high salt, low pH and low calcium conditions), allowing us to infer on the stages where biofilms are being formed and to determine if there is any correlation of this with expression of *Lig E* (Jakubovics *et al.*, 2013). Finally, biofilm formation on a glass surface can be compared between the WT *N. gonorrhoeae* and the Lig E-KO-

KanaR mutant, which can be combined with quantification of the amount of exDNA in the biofilm via fluorescent dyes and immunostaining (e.g. DDAO fluorescent dyes) (Jakubovics *et al.*, 2013; Kouzel *et al.*, 2015; Domenech *et al.*, 2016). If Ngo-Lig is truly important for biofilm formation, this would be particularly useful to identify conditions that lead to higher expression of Lig E-His for immunoblotting so as to confirm its potential periplasmic location.

Chapter Four of this thesis focussed on the production of recombinant proteins to determine the validity of any differences found in the *in vivo* experiments, especially any potential effects of additional tags. Perhaps the most interesting observation was the increase in ligation activity of the Lig E-His variant, suggesting that the addition of the His-tag provided an advantage to the Lig E-His-KanaR *N. gonorrhoeae* mutant regardless of the function of Ngo-Lig. Regardless, many unanswered questions from this chapter remained due to the time constraints of this thesis. This included the activity of Ngo-Lig on other substrates that are less frequently tested *in vitro*, like dsDNA of longer lengths (>500 bp) and ssDNA, both of which are present in gonococcal biofilms, or other optimal conditions for Ngo-Lig activity like the temperature of ligation or the metal cofactor, which would be useful when comparing with the *in vivo* gonococcal conditions for ligation. Furthermore, to understand any differences in ligation activity of Ngo-Lig with Lig E from other organisms (Psy-Lig, Ame-Lig, Hin-Lig, Nme-Lig), it would be critical to note any differences in their interactions and structures. As a substantial amount of this chapter explored the optimisations of both protein expression and purification, purification of Ngo-Lig using these optimised methods to provide a higher protein yield will allow us to do so in the near future. Two new analyses are also planned. First, an analytical size exclusion to confirm if Ngo-Lig exists as a monomeric protein, and second, a pull-down assay using affinity chromatography to check if Ngo-Lig is interacting with any other proteins in *N. gonorrhoeae*. This second experiment will focus especially on the potentially periplasmic nuclease, which would be expressed using the recombination, expression and purification protocols optimised in this chapter.

Regarding the original hypothesis of Ngo-Lig's role in DNA competence, Chapter Five was focused on the optimisations of DNA uptake reporter constructs (*GFP*, *β-*

lactamase) for future uptake experiments with nicked dsDNA substrates between the WT *N. gonorrhoeae* and the Lig E-KO-KanaR mutant under DNA damaging conditions. Progress from this chapter was constantly held back by contamination with reagents and the gonococcal stocks. Despite the range of optimisations performed to increase the homology:insert ratio, there seemed to be an underlying issue with the pMR32 plasmid itself, which prevented successful integration of the constructs into the genome. Currently, a new DNA uptake construct is being designed with the commonly utilised kanamycin resistance gene in place of the *ermC* gene. Other avenues for future research include the use of an alternate fluorescent protein, although the most interesting of all would be the effects of inducible promoters on the expression of the reporter genes of interest. Furthermore, this chapter reinforced the results from Chapter Three in terms of the potential toxicity of *GFP* in *N. gonorrhoeae* cells, not only because of its issues with folding, but also due to the nature of the gene itself. As a result of this, prior plans of tracking the location of Ngo-Lig and of fluorescently checking for successful integration of the gene into the genome have been on hold. Instead, an alternative fluorescent tag, mCherry, will be used.

Despite the many questions that remain on the specifics of the function and location of Ngo-Lig, results from this thesis serve as a critical stepping-stone for future characterisations of this ligase. Not only did it open a new discussion in terms of Lig E's role in biofilm formation, but the many optimisations performed in terms of DNA uptake construct making, transformation and protein purification will allow us to answer future questions with more ease. To our knowledge, this is the first example of *in vivo* Lig E characterisation in any bacterium. Considering the naturally competent nature of *N. gonorrhoeae* allowing them to integrate newly acquired antibiotic resistance genes, it is no wonder that we are quickly running out of new targets against this growing superbug. Thus, this area of research is critical to allow us to target the possible roots of its gene acquisition. It is our strong beliefs, that with more work to elucidate the function, regulation and location of Ngo-Lig, we will not only be able to target *N. gonorrhoeae* in the future, but also other Lig E-expressing human pathogens, bringing us one step closer to improving the health of the community as whole.

Appendices

Appendix A: Construction and evaluation of Ngo-Lig mutants *in vivo* (Chapter Three)

Constructs for generation of Ngo-Lig mutants in *N. gonorrhoeae*

Figure A.1. Genomic context of the Lig E coding DNA sequence (CDS) in wild-type MS11 *N. gonorrhoeae* visualised via Geneious. Any primers from Table A.1 that bind to the genome are labelled.
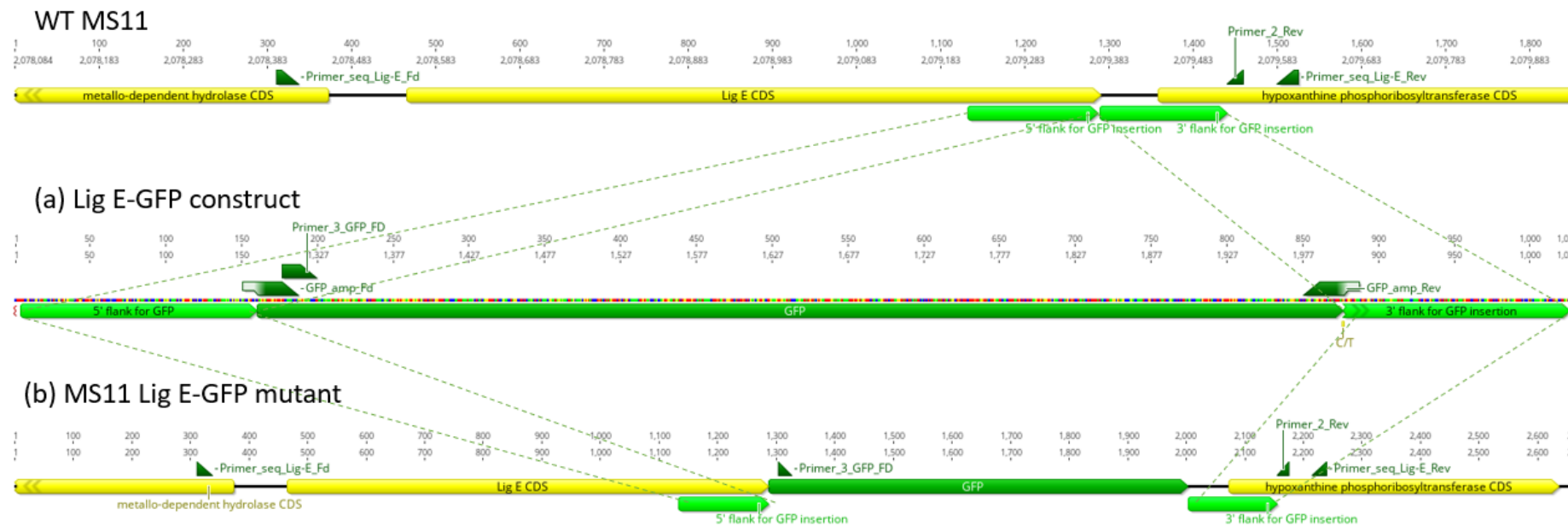
Lig E-GFP construct sequence (5' to 3'):

CAAAAACCGACACGGCGAATTCCGCATCGGCAGCGGTTTCAAGGACAAAGACCGCGACAACCCGCCCAAAATCGGCACGCTGATTACCTACCGTTACCGTGGCTTTACGCGGAAAGGCACGCCGAAATTTGC
CACATTTGTGCGCGTGCGTACCGACCGCATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGT
GATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCATGGCCAACACTTGTCACTACTTTCGGTTATGGTGTTCAATGCTTTGCGAGATACCCAGATCATATGAAACA
GCATGACTTTTTCAAGAGTGCCATGCCTGAAGGTTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAG
TTAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGA
CACAACATTGAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAACG
AAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAATAGCAGGACAGTTCAGACGGCATCCGATACTTTGGTTTATAATTTTCCTTTTACCGA
CCGATTCCGACATATGACCGATTTAGAAACCAAACGCCTTGAAACACAGGCGATGCTTGAAAACGCCGATCTTTTGTTCGACCAAGGCCAATGCCGTG

Figure A.2. Lig E-GFP MS11 *N. gonorrhoeae* construct sequence and design illustrated via Geneious (a) Lig E-GFP construct ordered from Twist Bioscience with homology to the gonococcal genome (b) Genomic context of the *N. gonorrhoeae* Lig E coding DNA sequence (CDS) after transformation with Lig E-GFP

Figure A.3. Lig E-His-KanaR *N. gonorrhoeae* construct sequence and design illustrated via Geneious (a) Lig E-His-KanaR construct ordered from Twist Bioscience with homology to the gonococcal genome (b) Genomic context of the *N. gonorrhoeae* Lig E coding DNA sequence (CDS) after transformation with Lig E-His-KanaR
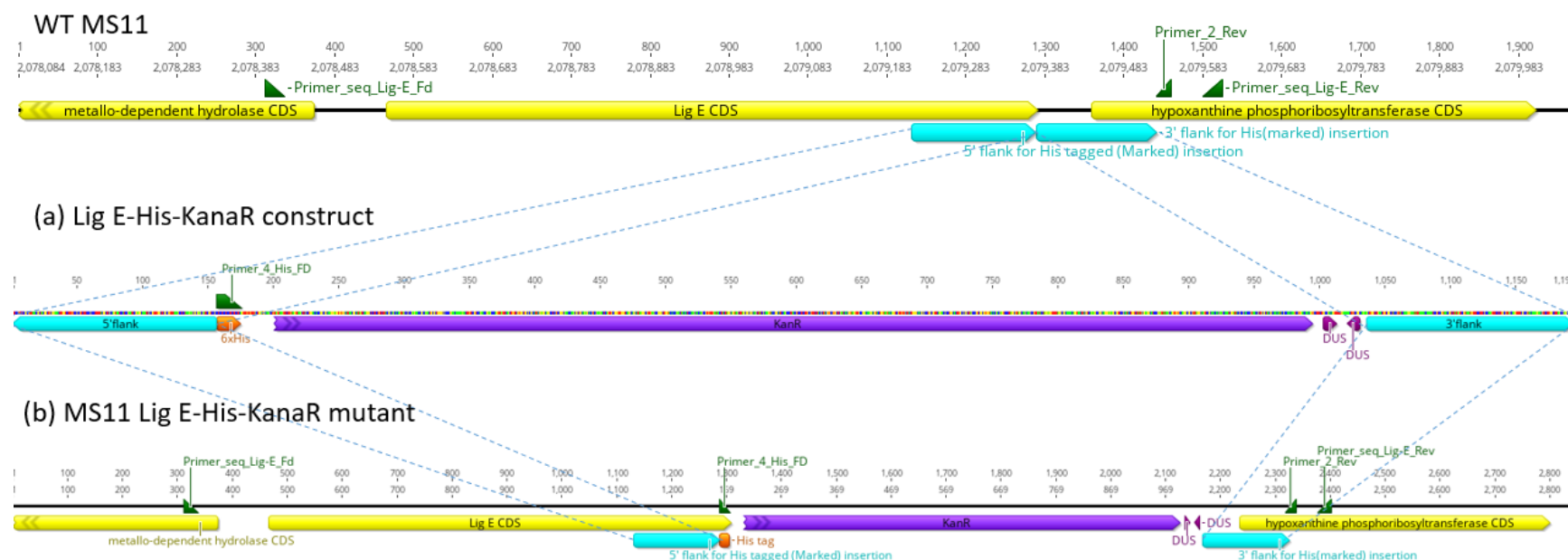
Lig E-His-KanaR construct sequence (5' to 3'):

```
AACCGACACGGCGAATTCCGCATCGGCAGCGGTTTCAAGGACAAAGACCGCGACAACCCGCCCAAAATCGGCACGCTGATTACCTACCGTTACCGTGGCTTTACGCGGAAAGGCACGCCGAAATTTGCCACA
TTTGTGCGCGTGCGTACCGACCGCCACCACCACCACCACCACTGAAAGAGAAAGCAGGTAGCTTGCAGATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCT
ATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCG
GCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCC
GAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATC
AGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGG
AAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTT
TACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCAAACAGGCCGTCTGAAAAAGGTTTTTCAGACGGCCTGTTGCAGGACAGTTCAGACGGCATCCGATA
CTTTGGTTTATAATTTTCCTTTTACCGACCGATTCCGACATATGACCGATTTAGAAACCAAACGCCTTGAAACACAGGCGATGCTTGAAAACGCCGATCTTTTGTTCGACCAAGGCCAATGCCGTGCCGCA
```
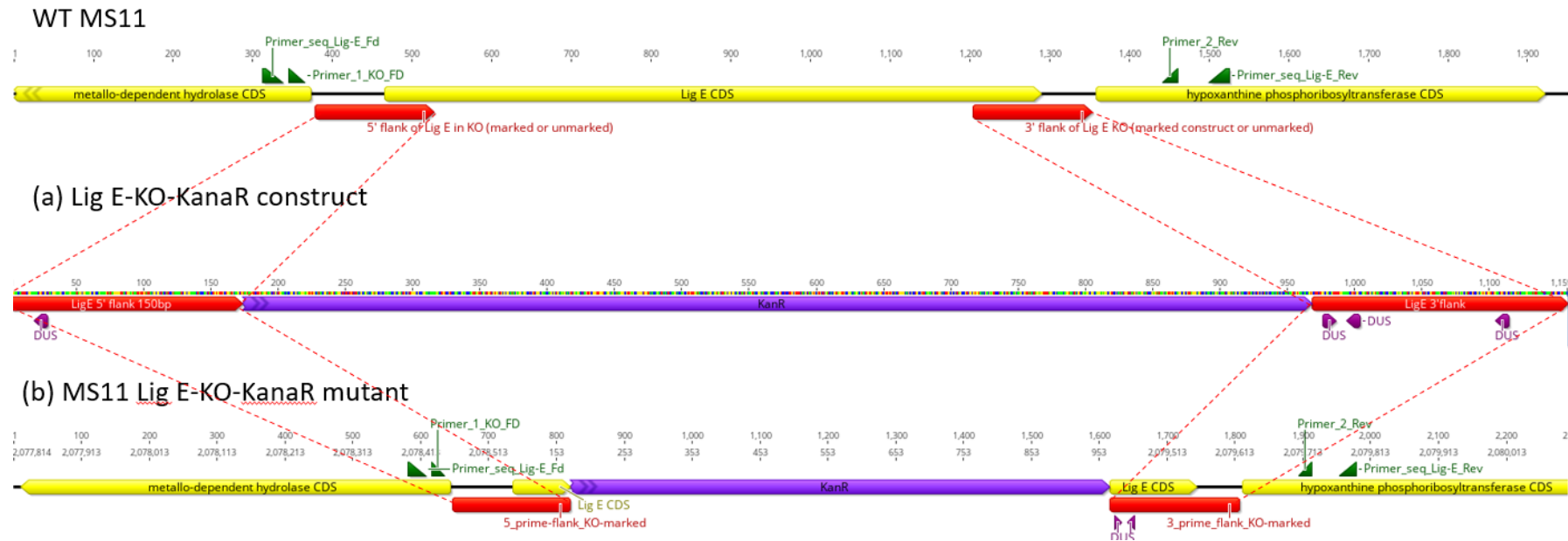
Lig E-KO-KanaR construct sequence (5' to 3'):

GTGTTCCTAAAACGGTTTTTCAGACGGCATCGGGTTTGCCGTTTGTTTTATGGCGGTTTGCCGCCTGTTTTGATATTGGGGGAATCAGGTGATTAAGAAGATAATCGGCGGCATCATACCGATTTTTACGGCGG
TTTTCATCCCTGCATCAAGAGAAAGCAGGTAGCTTGCAGATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCT
CTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCC
TTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATG
CGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGG
CTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACT
GTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCG
CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCAAACAGGCCGTCTGAAAAAGGTTTTTCAGACGGCCTGTTACGCTGATTACCTACCGTTACCGTGGCTTTACGCGGAAAGGCACGCCGAAATTTGC
CACATTTGTGCGCGTGCGTACCGACCGCTGAGCAGGACAGTTCAGACGGCATCCGATACTTTGGTTTATAATTTTCCTTTTACCGACCGATTCC

Figure A.4. Lig E-KO-KanaR MS11 _N. gonorrhoeae_ construct sequence and design illustrated via Geneious (a) Lig E-KO-KanaR construct ordered from Twist Bioscience with homology to the gonococcal genome (b) Genomic context of the _N. gonorrhoeae_ Lig E coding DNA sequence (CDS) after transformation with Lig E-KO-KanaR

Primers for validation of Ngo-Lig mutants

Table A.1. Primers (ordered from Integrated DNA Technologies (IDT[TM]) and Invitrogen[TM] used for MS11 *N. gonorrhoeae* Lig E mutant verifications

Primer name	Direction	Composition (5' to 3')	T_m (°C)
Primer_1_KO_FD	Forward	CGGGGAGAAT TTCGTAACGT	54.9
Primer_2_Rev	Reverse	CACTTTTTGC AGTGCGGCA	57.3
Primer_3_GFP_FD	Forward	AACTTTTCACTG GAGTTGTCCCA	56.7
Primer_4_His_FD	Forward	CACCACCAC CACCACCACT	60.2
Primer_seq_Lig-E_Fd	Forward	CAAATCAGGGTGC AGTTTTGGCGGAA	54.0
Primer_seq_Lig-E_Rev	Reverse	GCCCCGCCCATCA CGGGCAGGAGCA	64.0

Raw and additional results/data from determining the importance and location of Ngo-Lig

Table A.2. Initial OD$_{600}$ readings of both seeder and final growth cultures for each replicate before the growth experiment

Replicate	OD$_{600}$ readings of seeder culture	Amount added to create a 30 mL growth culture (mL)	OD$_{600}$ readings of growth culture at hour 0
WT MS11			
A	0.50	0.600	0.03
B	0.58	0.517	0.04
C	0.47	0.638	0.02
Lig E-His-KanaR MS11			
A	0.38	0.789	0.02
B	0.34	0.882	0.02
C	0.35	0.857	0.03
Lig E-KO-KanaR MS11			
A	0.28	1.071	0.02
B	0.30	1.000	0.03
C	0.30	1.000	0.03

Additional data from the growth experiment: OD$_{600}$ measurements

Average OD$_{600}$ readings of the different *N. gonorrhoeae* cultures with three biological replicates each
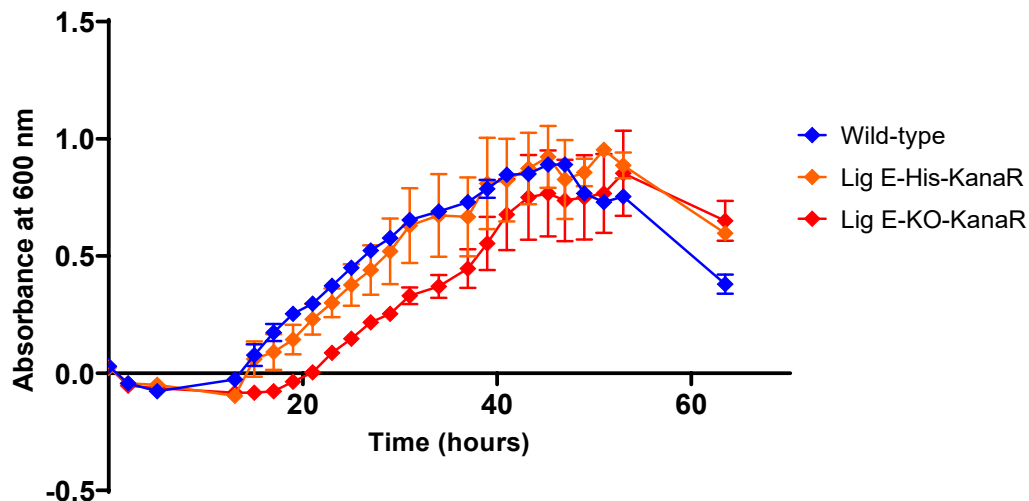
Figure A.5. Average OD$_{600}$ readings of each MS11 *N. gonorrhoeae* culture with three biological replicates for each gonococcal variant. The standard errors are plotted in the graph.

Pilot *N. gonorrhoeae* growth experiment with one replicate per gonococcal variant
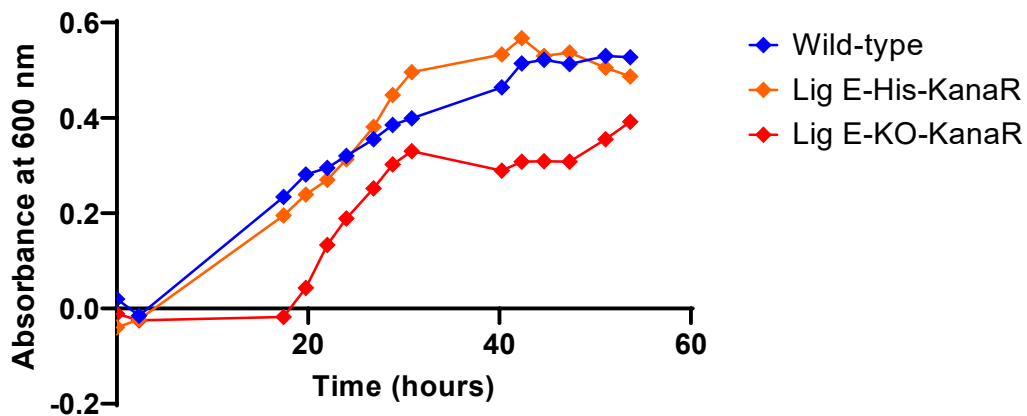
Figure A.6. OD$_{600}$ readings from the pilot growth experiment of MS11 *N. gonorrhoeae* cultures with one replicate each for the wild-type *N. gonorrhoeae*, the Lig E-His-KanaR mutant and the Lig E-KO-KanaR mutant
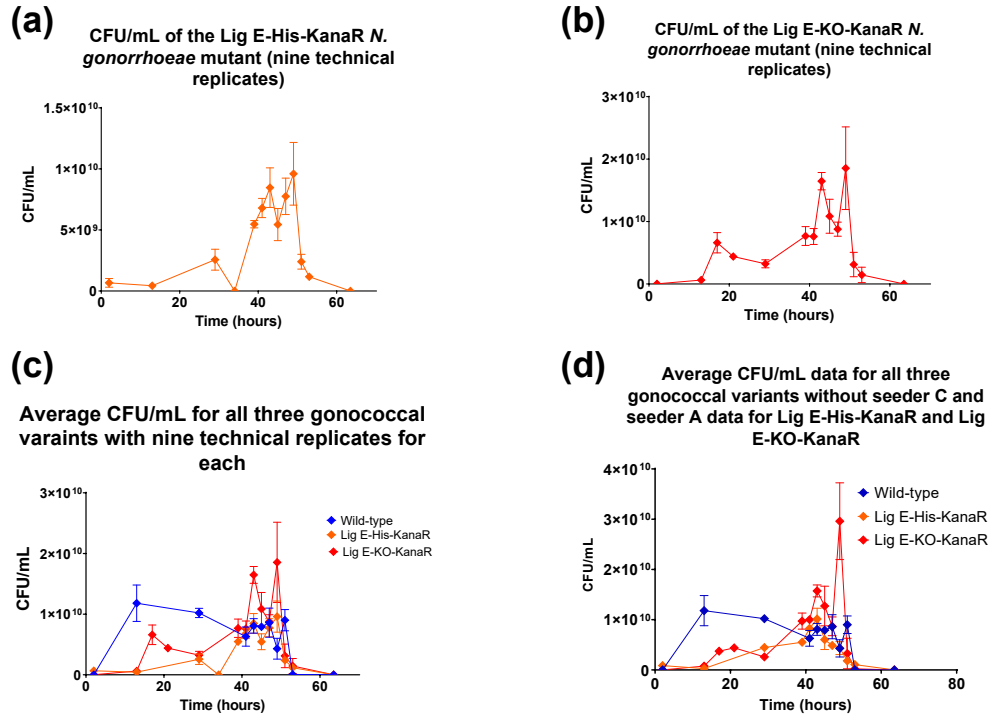
Figure A.7. CFU/mL measurements of MS11 *N. gonorrhoeae* cultures from the growth experiment with standard errors plotted (a) Average of nine technical replicates for the Lig E-His-KanaR cultures (b) Average of nine technical replicates for the Lig E-KO-KanaR cultures (c) Average CFU/mL for all three Lig E *N. gonorrhoeae* variants with nine technical replicates each (d) Average CFU/mL for all three Lig E *N. gonorrhoeae* variants with nine technical replicates for the wild-type MS11 *N. gonorrhoeae*, six for the Lig E-His-KanaR mutant (without seeder C) and six for the Lig E-KO-KanaR mutant (without seeder A)
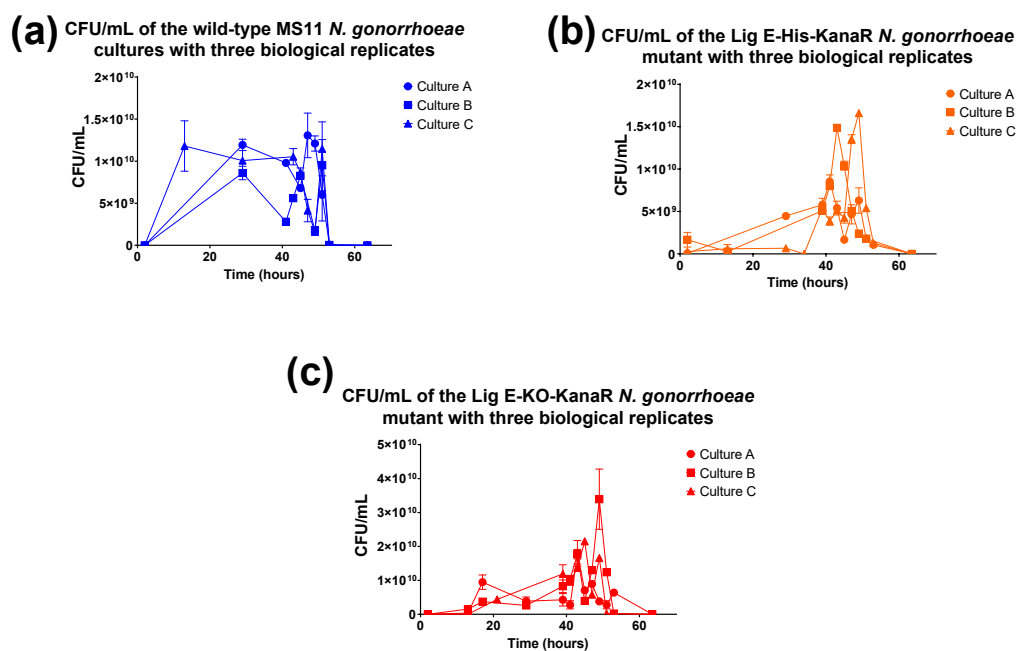
Figure A.8. CFU/mL measurements of each biological replicate for the MS11 *N. gonorrhoeae* variants from the growth experiment, with standard errors plotted (a) Wild-type MS11 *N. gonorrhoeae* cultures (b) Lig E-His-KanaR mutant cultures (c) Lig E-KO-KanaR mutant cultures

Additional data from the growth experiment: exDNA quantification

(a) Average amount of extracellular dsDNA in the wild-type MS11 *N. gonorrhoeae* cultures

(b) Average amount of extracellular dsDNA in the LigE-His-KanaR *N. gonorrhoeae* cultures

(c) Average amount of extracellular dsDNA in the Lig E-KO-KanaR *N. gonorrhoeae* cultures
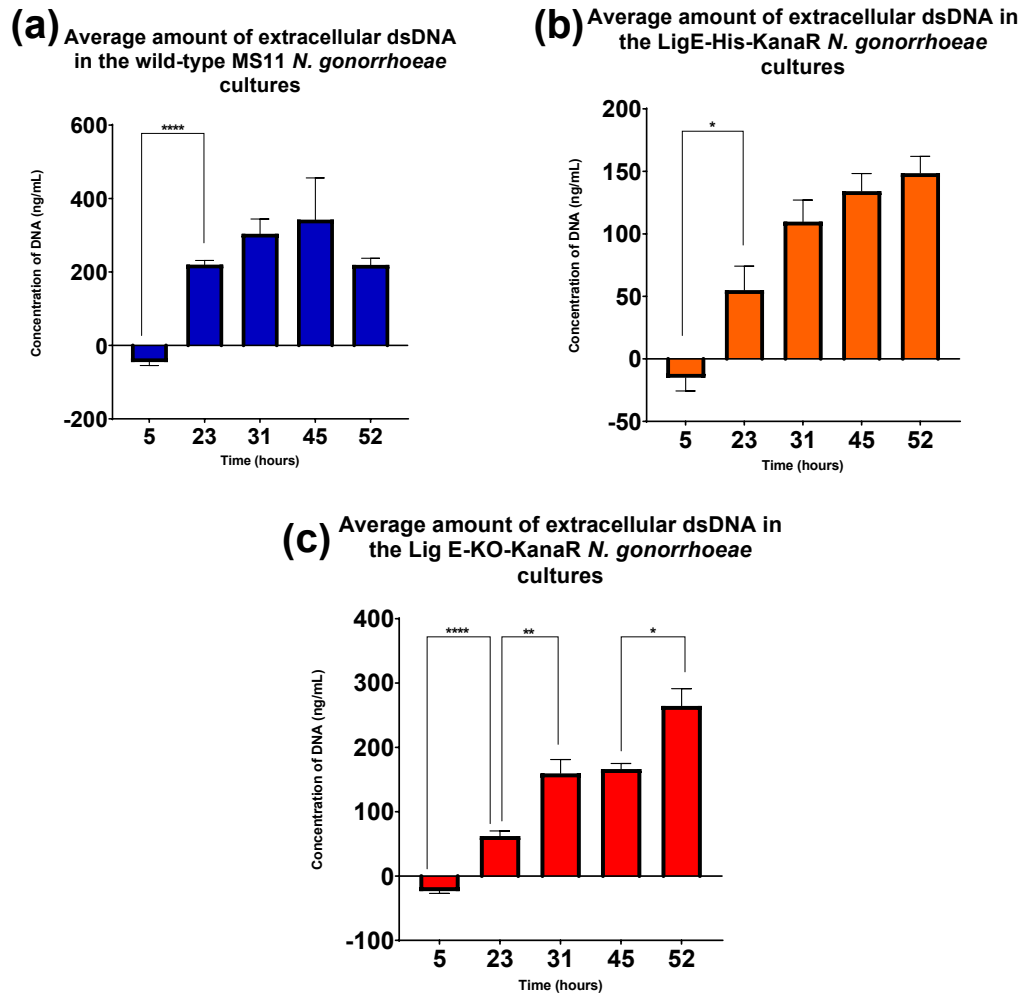
Figure A.9. Average amount of extracellular double-stranded DNA in the MS11 *N. gonorrhoeae* supernatant samples obtained from the growth experiment, with three biological replicates for each condition (a) Wild-type MS11 cultures (b) Lig E-His-KanaR cultures (c) Lig E-KO-KanaR cultures. Quantification of the amount of dsDNA was performed using the PicoGreen® dye kit. The standard errors and significant data (p < 0.05 (*)) from unpaired two-tailed t-tests are plotted for each graph.

(a) dsDNA calibration curve with untreated lambda DNA

- 10% FBS DMEM F12
- DMEM F12

$y = 0.7248x + 31.673$
$R^2 = 0.9939$

$y = 0.5969x + 14.49$
$R^2 = 0.9952$

(b) dsDNA calibration curve with treated lambda DNA in 10% FBS DMEM F12

$y = 0.5718x + 58.88$
$R^2 = 0.9791$

(c) Effect of heat and EDTA treatment on DNA samples

- Treated lamda DNA
- Untreated lamda DNA
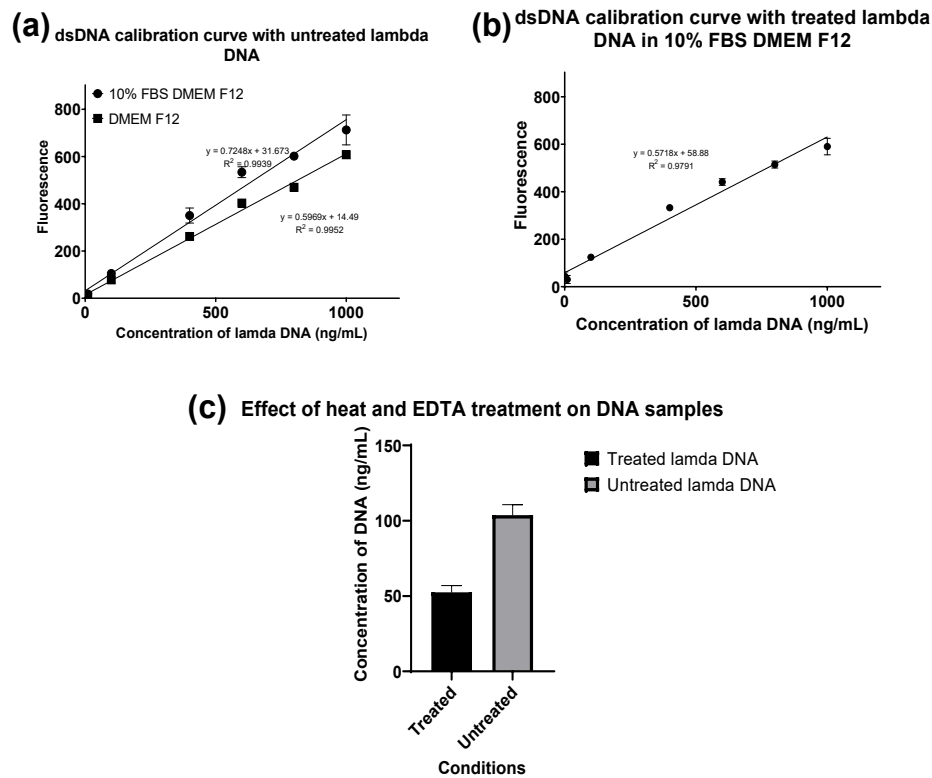
Figure A.10. The effect of foetal bovine serum (FBS) and both heat and EDTA treatment on fluorescence intensity during dsDNA quantification (a) Calibration curve of untreated lambda DNA in DMEM/F-12 media, in the presence and absence of 10% FBS (b) Calibration curve of treated lamda DNA in DMEM/F-12 media in the presence of 10% FBS, which was used for final DNA quantification of supernatant samples (c) Effect of heat and EDTA treatment on 100 ng/mL lamda DNA in 10% FBS DMEM/F-12

Additional data from the western blot for Lig E-His detection
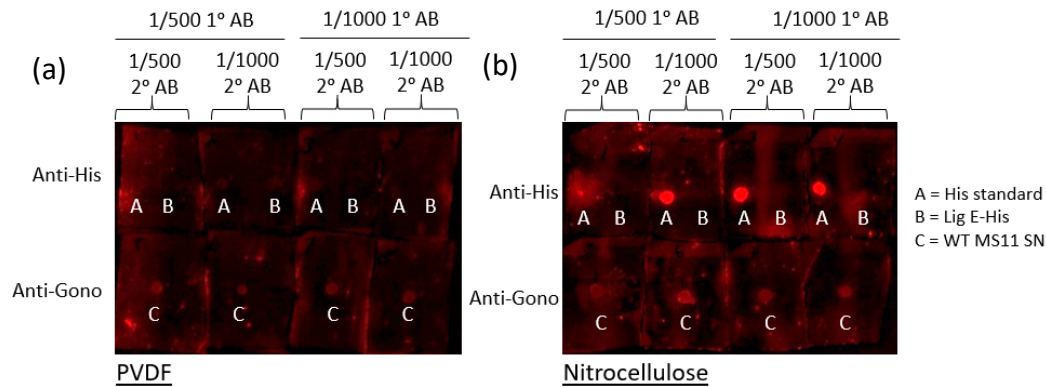
Figure A.11. Dot plot experiment on (a) polyvinylidene fluoride (PVDF) and (b) nitrocellulose membranes. The His standard (A) and Lig E-His (B) samples were incubated with the mouse primary anti-His antibody, while the wild-type MS11 *N. gonorrhoeae* supernatant samples from the growth experiment (C) were incubated with the mouse primary anti-*N. gonorrhoeae* outer membrane protein antibody. Either a 1:500 or 1:1000 dilution was used for both primary antibodies (1° AB). All were incubated with the same secondary AlexaFluor™ 647-conjugated mouse secondary antibody (2°AB) at either 1:500 or 1:1000 dilutions.
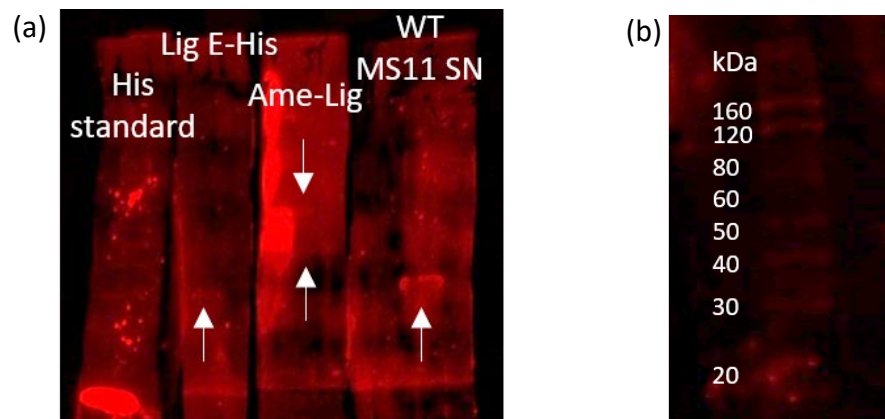
Figure A.12. Optimisation of western blot conditions using the fluorescent AlexaFluor™ 647-conjugated mouse secondary antibody (a) His standard ladder visualised in the presence of the two proteins of interest (Lig E-His: 29 kDa, Ame-Lig: 31 kDa) probed with the anti-His primary antibody, and a wild-type MS11 *N. gonorrhoeae* supernatant (SN) sample from the growth experiment probed with the anti-*N. gonorrhoeae* outer membrane antibody (30 kDa), with arrows pointing at the bands of interest (b) His standard visualised alone, probed with the anti-His antibody
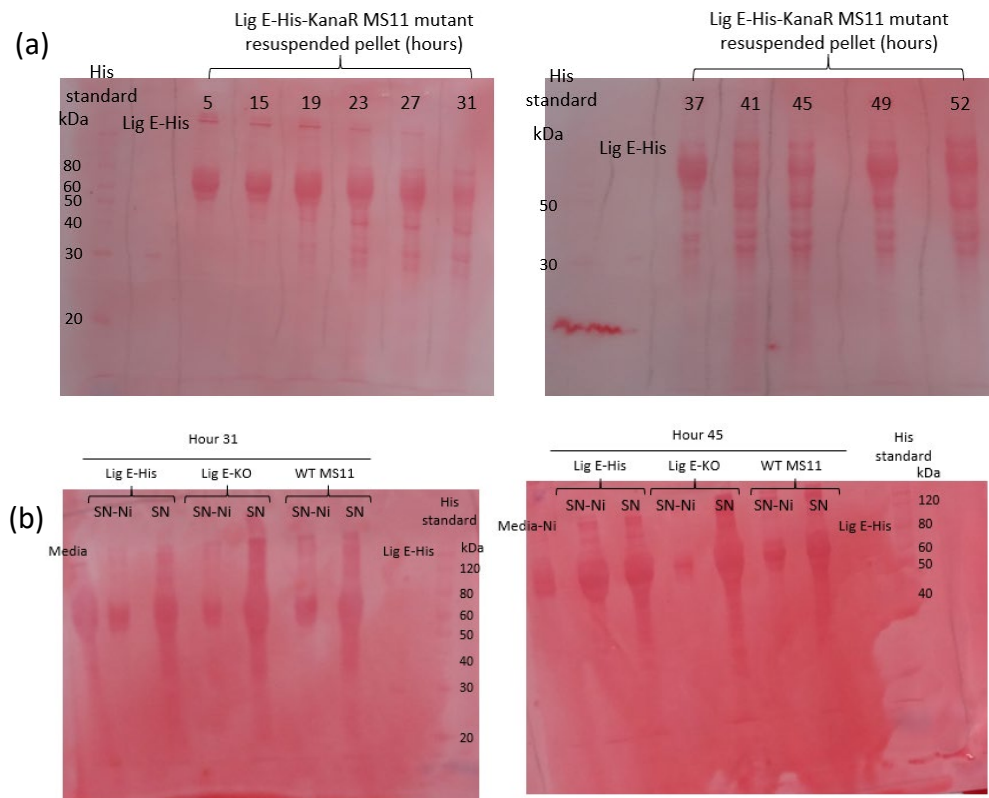
Figure A.13. Transfer of the *N. gonorrhoeae* growth experiment sample proteins from 12% SDS-PAGE gels onto nitrocellulose membranes (western blot), visualised with Ponceau S with a Lig E-His recombinant protein control (29 kDa) (a) Lig E-His-KanaR resuspended pellet samples (b) Wild-type MS11 *N. gonorrhoeae*, Lig E-His-KanaR and Lig E-KO-KanaR supernatant samples. SN = supernatant samples; SN-Ni = nickel pull-down of supernatant samples to select for His-containing proteins.

Appendix B: Recombinant production of Ngo-Lig (Chapter Four)

E. coli strains of interest

Table B.1. Genotypes of the different E. coli strains used for recombinant Ngo-Lig protein production

E. coli strain	Genotype	Antibiotic resistance	Purpose
DH5α	F⁻ *endA1 glnV44 thi-1 recA1 relA1 glyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17 (r$_K$⁻m$_K$⁺), λ⁻	-	Plasmid isolation
BL21(DE3)pLysS	F⁻ ompT gal dcm lon hsdS$_B$(r$_B$⁻m$_B$⁻) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB⁺]$_{K-12}$(λ^S) pLysS[T7p20 ori$_{p15A}$](Cm^R)	Chloramphenicol	Protein expression
Origami™(DE3)	Δ(*ara-leu*)7697 Δ*lacX₇₄* Δ*phoA Pvu*II *phoR araD₁₃₉ ahpC gale galK rpsL* F'[*lac⁺ lacI�q pro*] (DE3) *gor522*∷Tn*10 trxB* (Str^R, Tet^R)	-	Protein expression

Plasmid constructs of interest

Donor plasmids
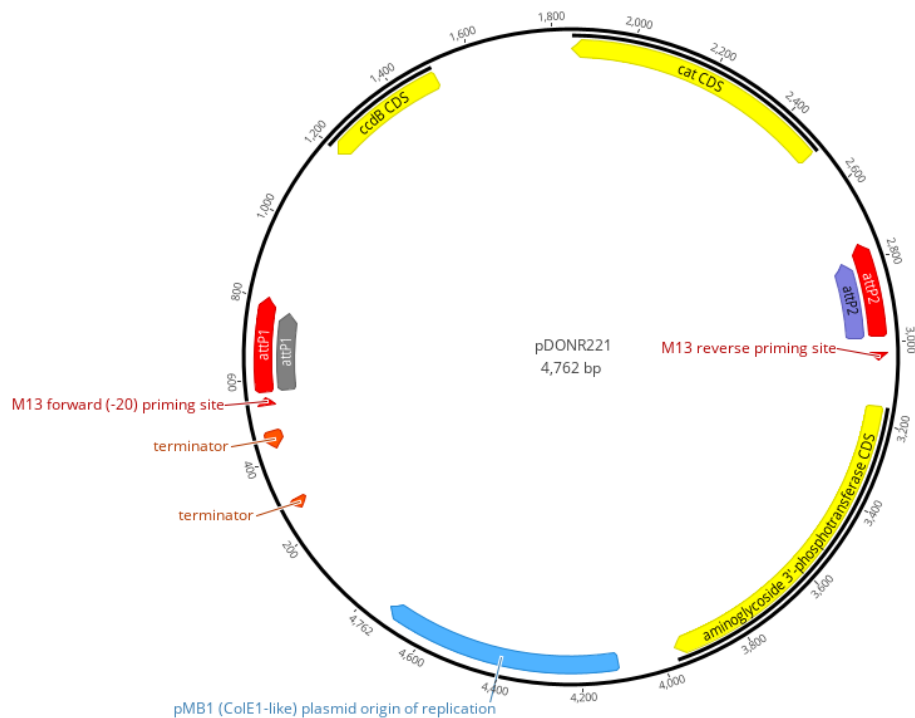
Figure B.1. pDONR221 donor plasmid visualised via Geneious. Constructs for recombinant protein expression were ordered from Twist Bioscience in these plasmids.
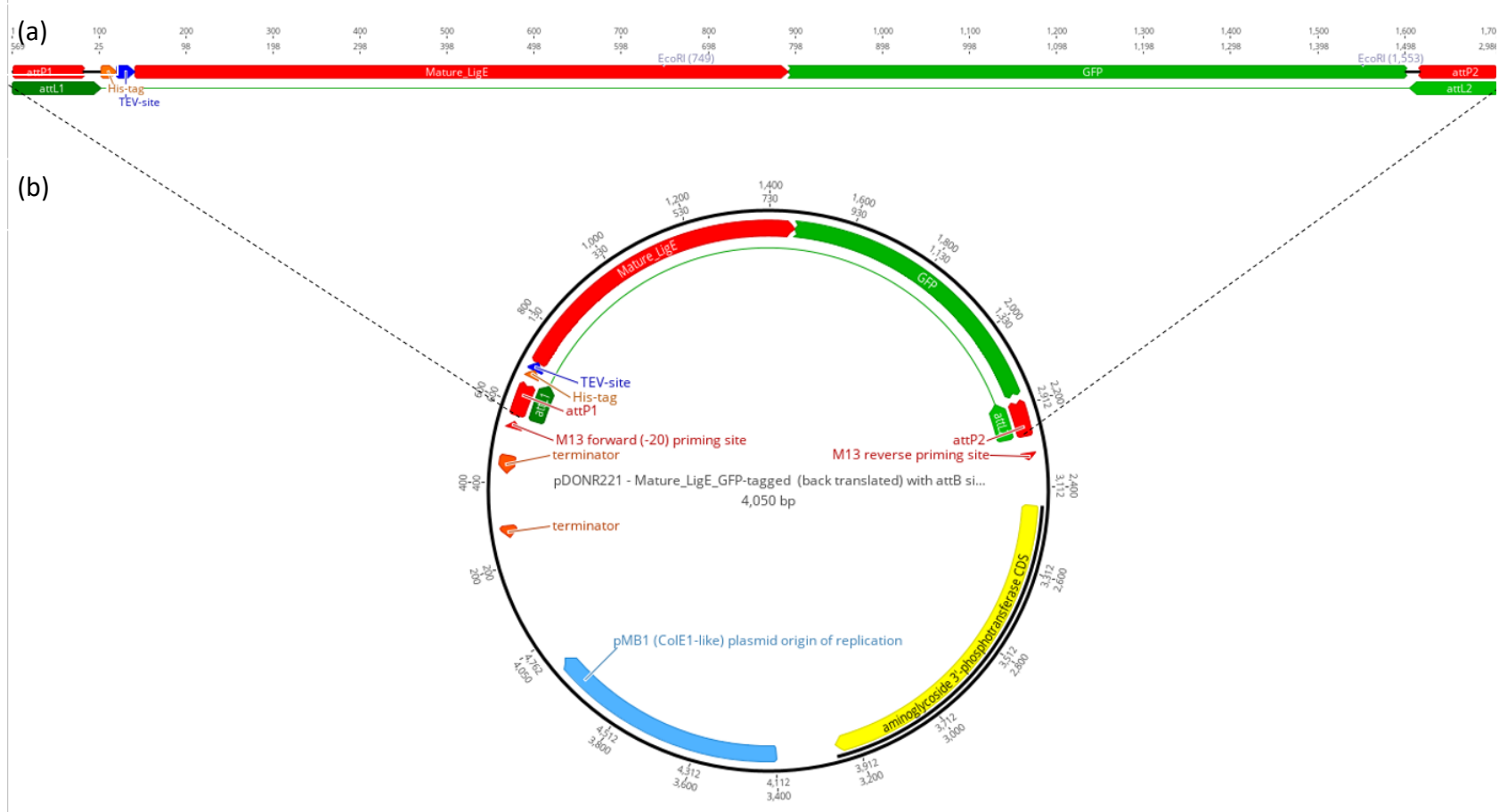
Figure B.2. pDONR221 His-(TEV)-Lig E-GFP plasmid visualised via Geneious (a) Construct of interest (C-terminally GFP-tagged Lig E with an N-terminal His-tag and TEV cleavage site) to be cloned into a destination plasmid via Gateway cloning (b) pDONR221 plasmid with the construct of interest between two *attL* sites shown in green
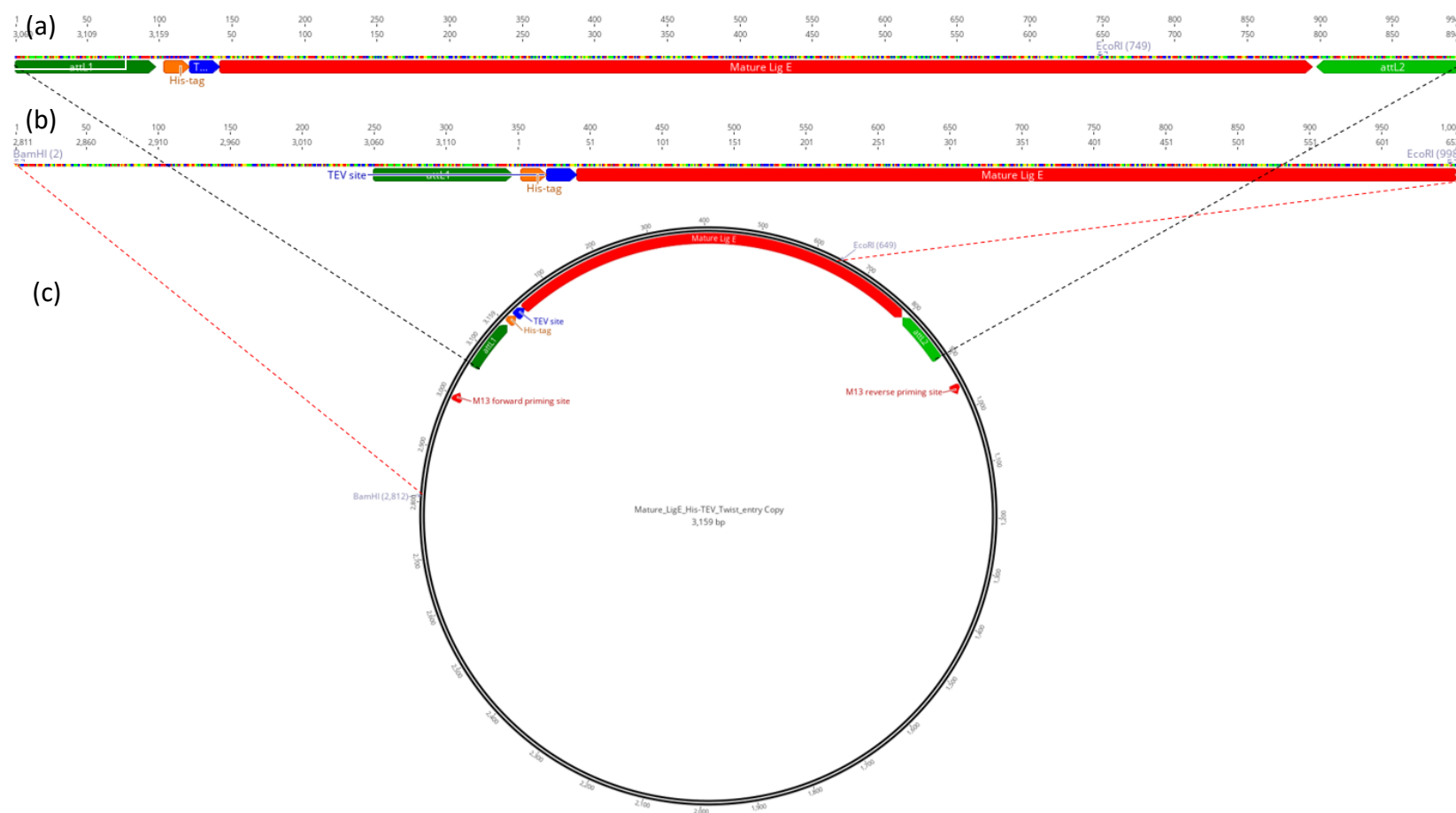
Figure B.3. pDONR221 His-(TEV)-Lig E plasmid visualised via Geneious (a) Construct of interest (Lig E with an N-terminal His-tag and a TEV cleavage site) to be cloned into a destination plasmid via Gateway cloning (b) EcoRI and BamHI-digested "insert" construct to form a new pDONR221 His-(TEV)-Lig E-His plasmid (c) pDONR221 plasmid with the construct of interest between two *attL* sites shown in green.
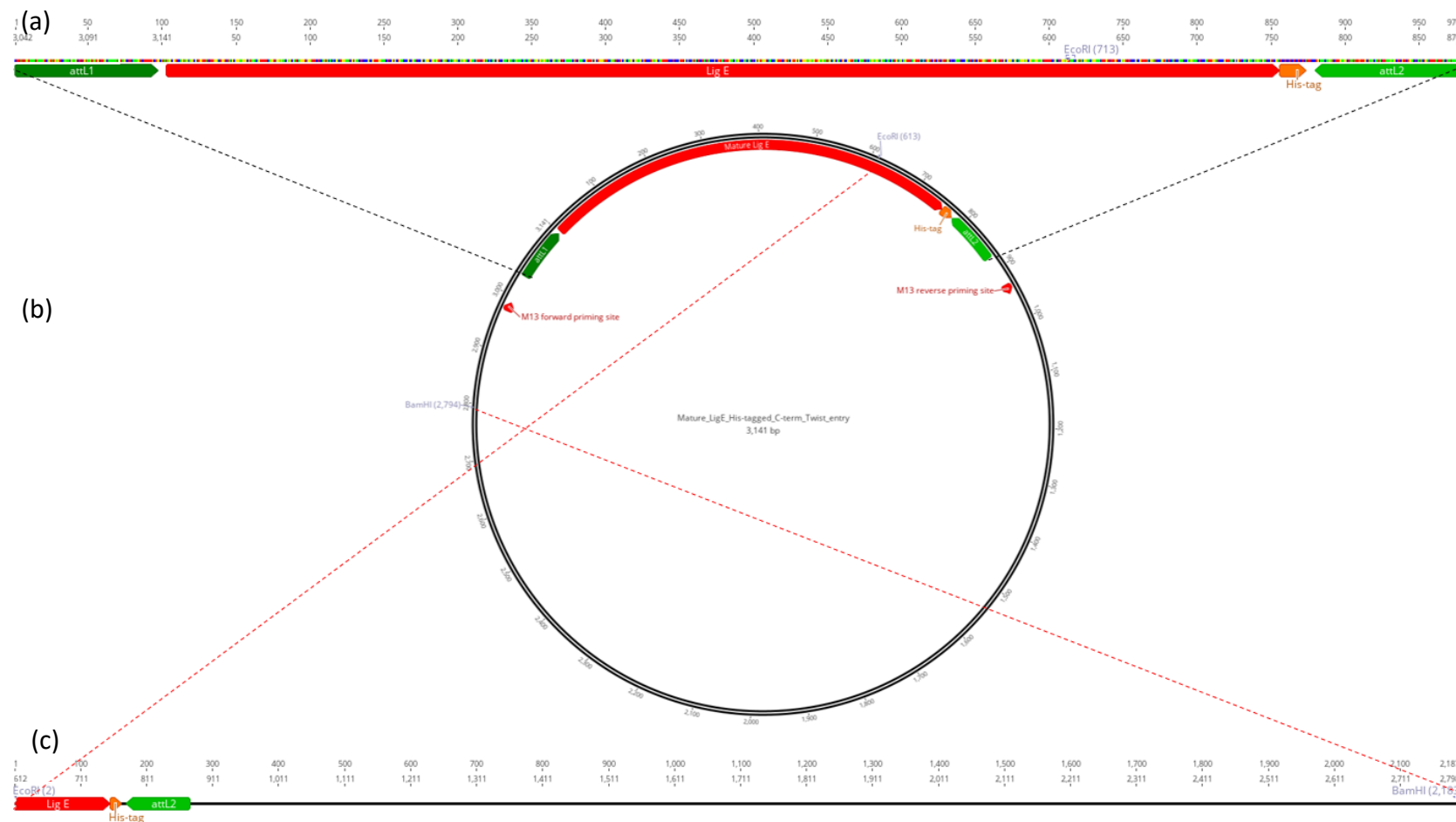
Figure B.4. pDONR221 Lig E-His plasmid visualised via Geneious (a) Construct of interest (C-terminally His-tagged Lig E) to be cloned into a destination plasmid via Gateway cloning (b) pDONR221 plasmid with the construct of interest between two *attL* sites shown in green (c) EcoRI and BamHI-digested "vector" construct to form a new pDONR221 His-(TEV)-Lig E-His plasmid
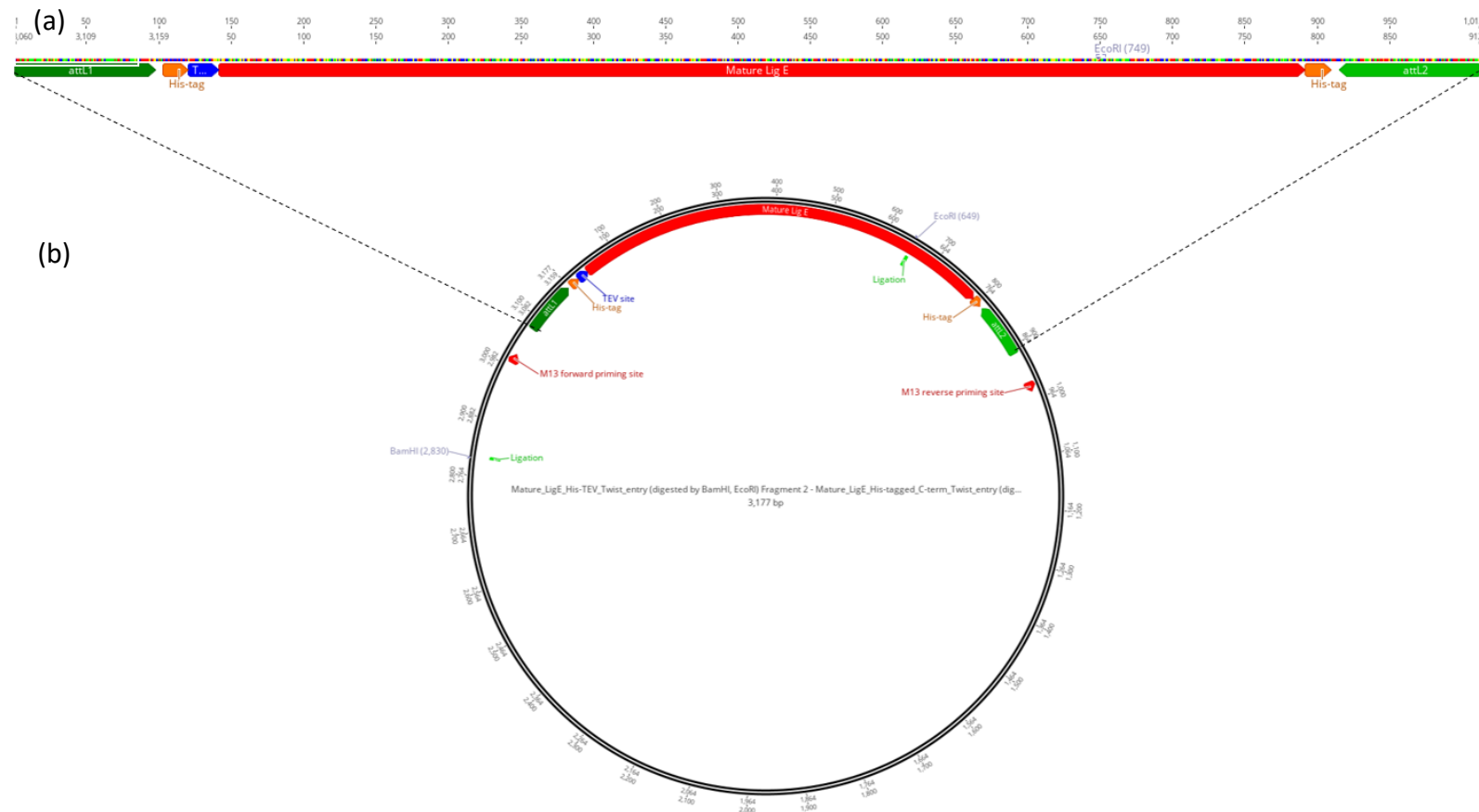
Figure B.5. pDONR221 His-(TEV)-Lig E-His plasmid formed after restriction digestion and ligation of pDONR221 His-(TEV)-Lig E (Figure B.3 (b)) and pDONOR221 Lig E-His (Figure B.4 (c)) plasmids, visualised via Geneious (a) Construct of interest (C-terminally His-tagged Lig E with an N-terminal His-tag and a TEV cleavage site) to be cloned into a destination plasmid after Gateway cloning (b) pDONR221 plasmid with the construct of interest between two *attL* sites

Destination plasmids
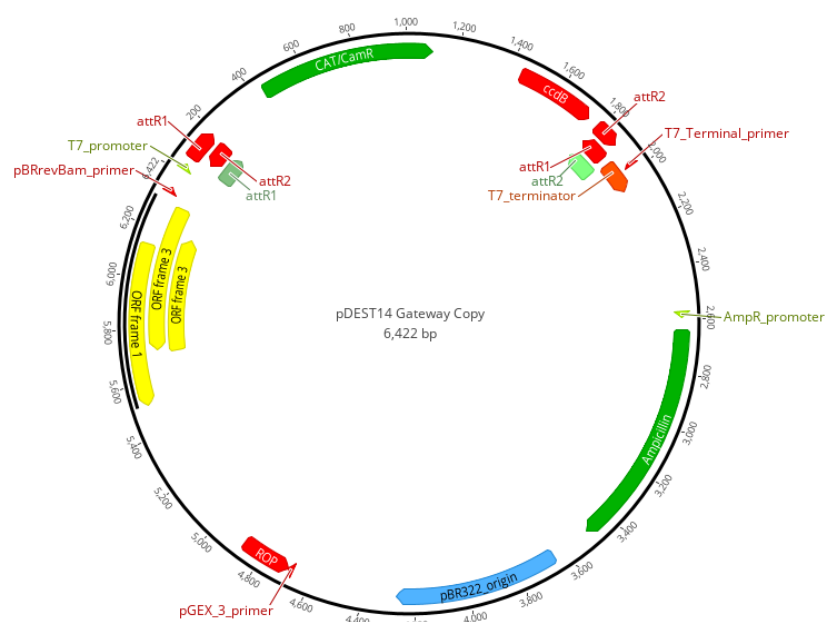
Figure B.6. pDEST14 destination plasmid visualised via Geneious. The area between the two labelled *attR* sites shown in red including the *ccdB* toxicity gene were swapped for the constructs of interest between two *attL* sites in the donor plasmid after Gateway cloning.
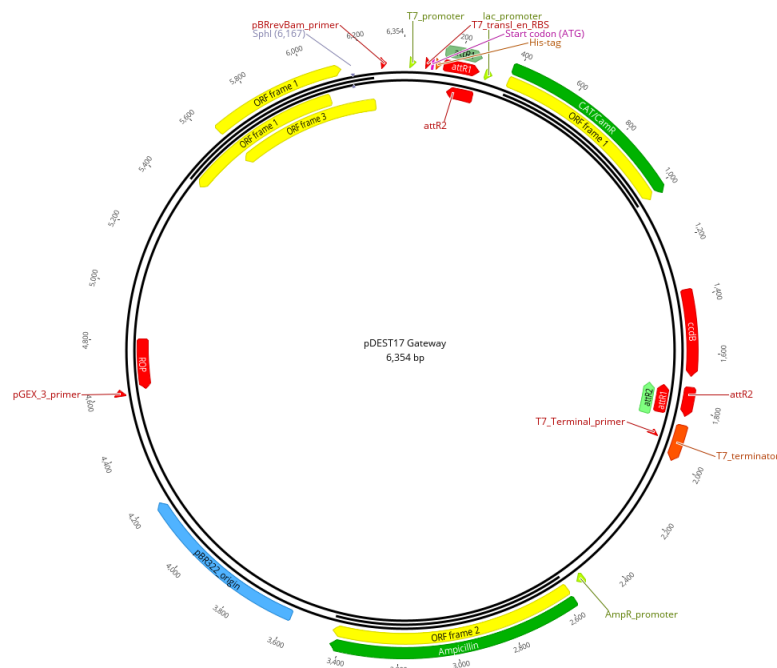
Figure B.7. pDEST17 destination plasmid with a start codon (pink) and a poly-histidine tag (orange) after the T7 promoter, visualised via Geneious. The area between the two labelled *attR* sites shown in red including the *ccdB* toxicity gene were swapped for the constructs of interest between two *attL* sites in the donor plasmid after Gateway cloning.
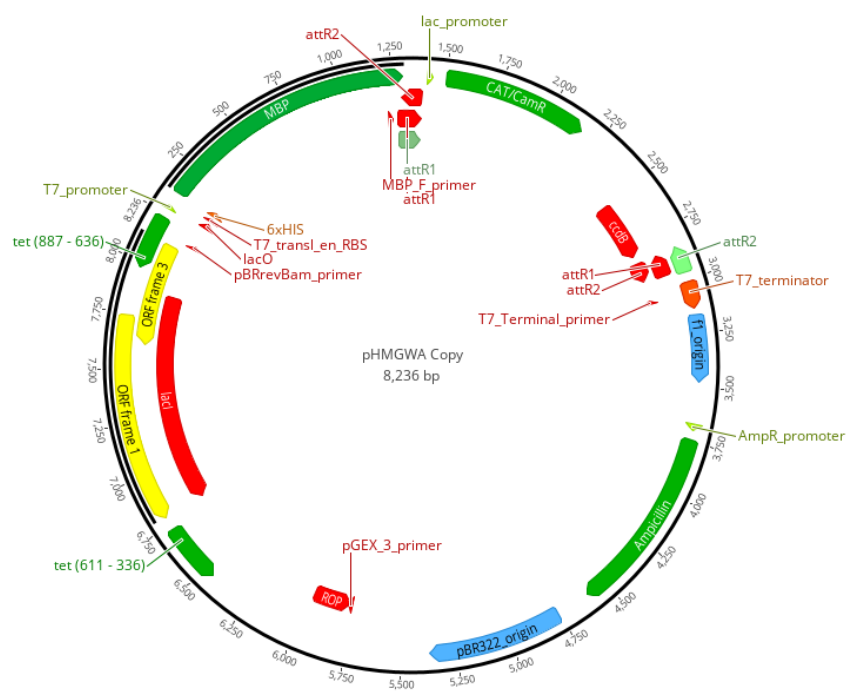
Figure B.8. pHMGWA destination plasmid with a gene encoding the maltose binding protein after the T7 promoter, visualised via Geneious. The area between the two labelled *attR* sites shown in red including the *ccdB* toxicity gene were swapped for the constructs of interest between two *attL* sites in the donor plasmid after Gateway cloning.

147

Expression plasmids
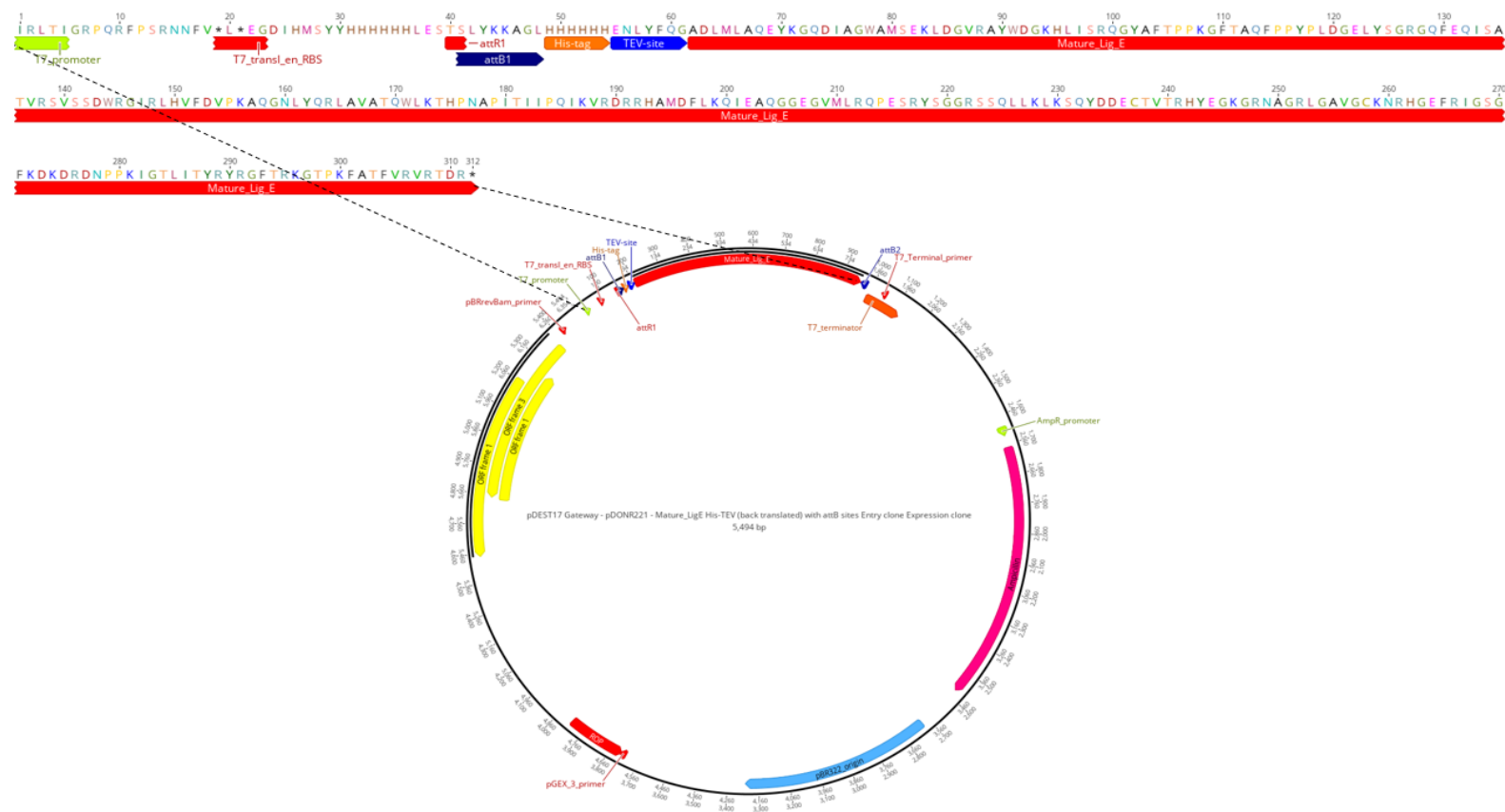
Figure B.9. His-(TEV)-Lig E pDEST17 expression plasmid illustrated via Geneious, which after expression and TEV cleavage will yield the native Lig E protein. The amino acid sequence of the expressed protein from this plasmid is shown in the zoomed in section. The *attB* sites shown in blue indicate sites of insertion after cloning of the donor plasmid construct of interest into the destination plasmid.
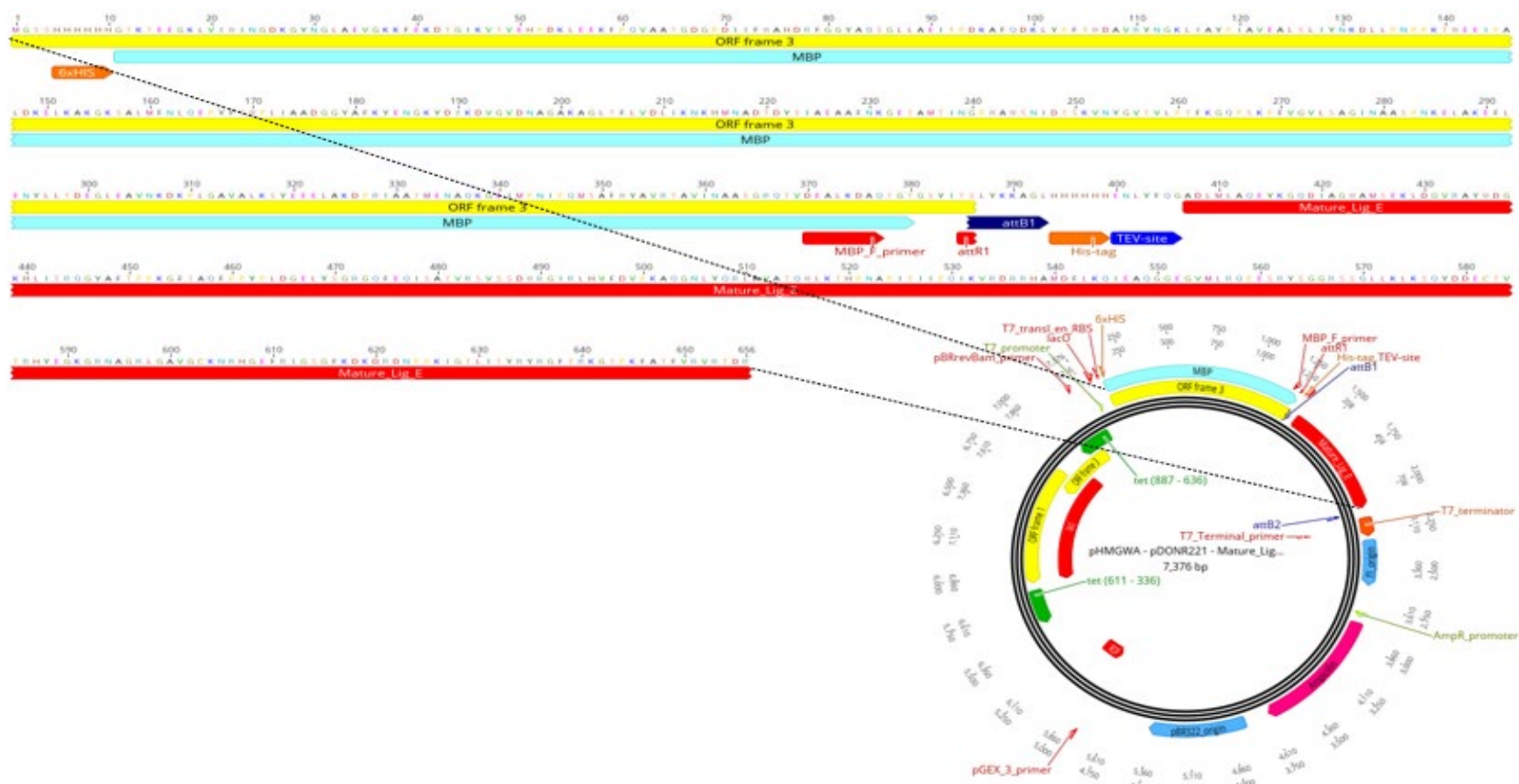
Figure B.10. His-(TEV)-Lig E pHMGWA expression plasmid illustrated via Geneious, which after expression and TEV cleavage will yield the native Lig E protein. The amino acid sequence of the expressed protein from this plasmid is shown in the zoomed in section. The *attB* sites shown in blue indicate sites of insertion after cloning of the donor plasmid construct of interest into the destination plasmid.

Figure B.11. Lig E-His pDEST14 expression plasmid illustrated via Geneious, which after expression will yield the C-terminally His-tagged Lig E protein. The amino acid sequence of the expressed protein from this plasmid is shown in the zoomed in section. The *attB* sites shown in blue indicate sites of insertion after cloning of the donor plasmid construct of interest into the destination plasmid.

Figure B.12. His-(TEV)-Lig E-GFP pDEST17 expression plasmid illustrated via Geneious, which after expression and TEV cleavage will yield the C-terminally GFP-tagged Lig E protein. The amino acid sequence of the expressed protein from this plasmid is shown in the zoomed in section. The *attB* sites shown in blue indicate sites of insertion after cloning of the donor plasmid construct of interest into the destination plasmid.
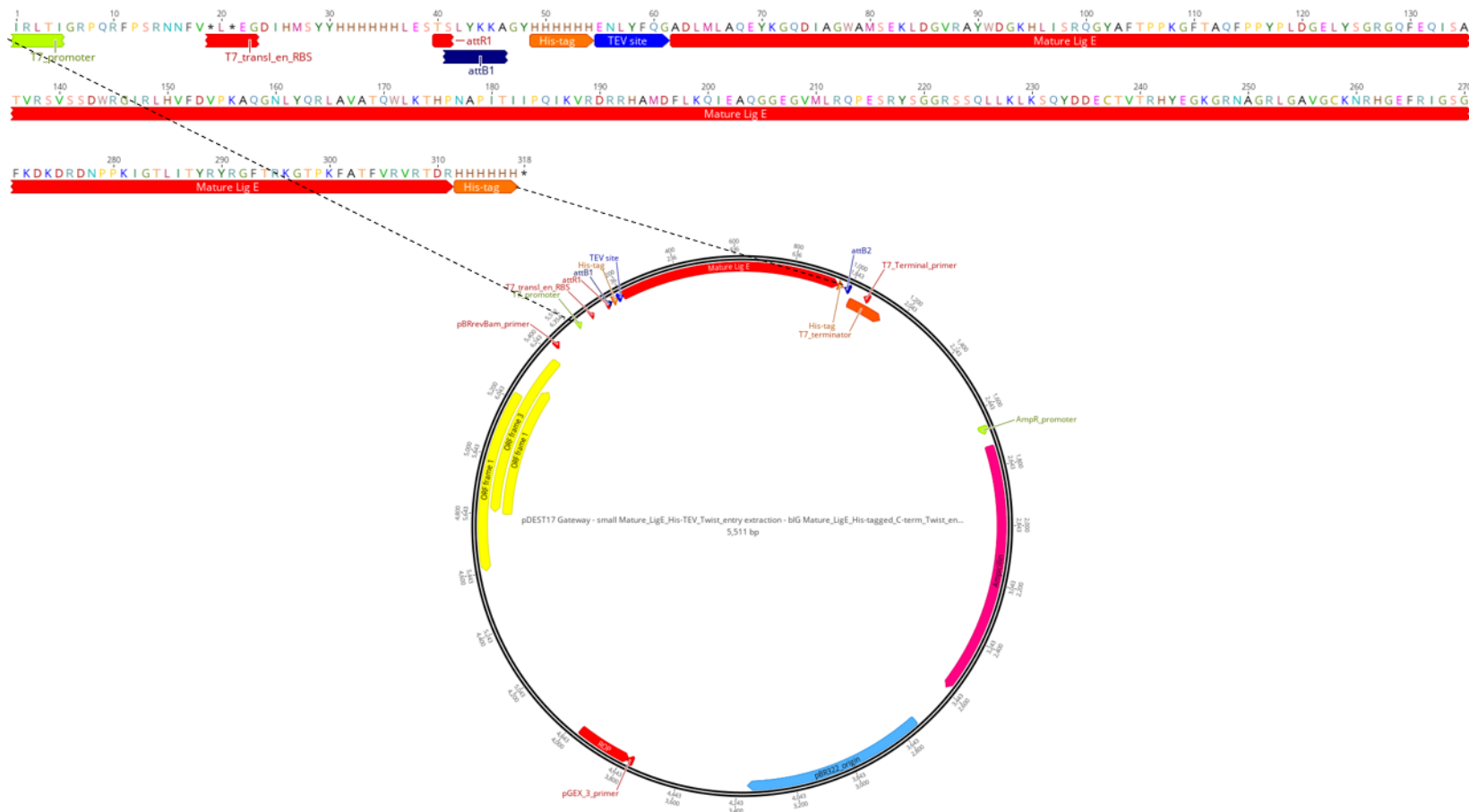
Figure B.13. His-(TEV)-Lig E-His pDEST17 expression plasmid illustrated via Geneious, which after expression and TEV cleavage will yield the C-terminally His-tagged Lig E protein. The amino acid sequence of the expressed protein from this plasmid is shown in the zoomed in section. The *attB* sites shown in blue indicate sites of insertion after cloning of the donor plasmid construct of interest into the destination plasmid.
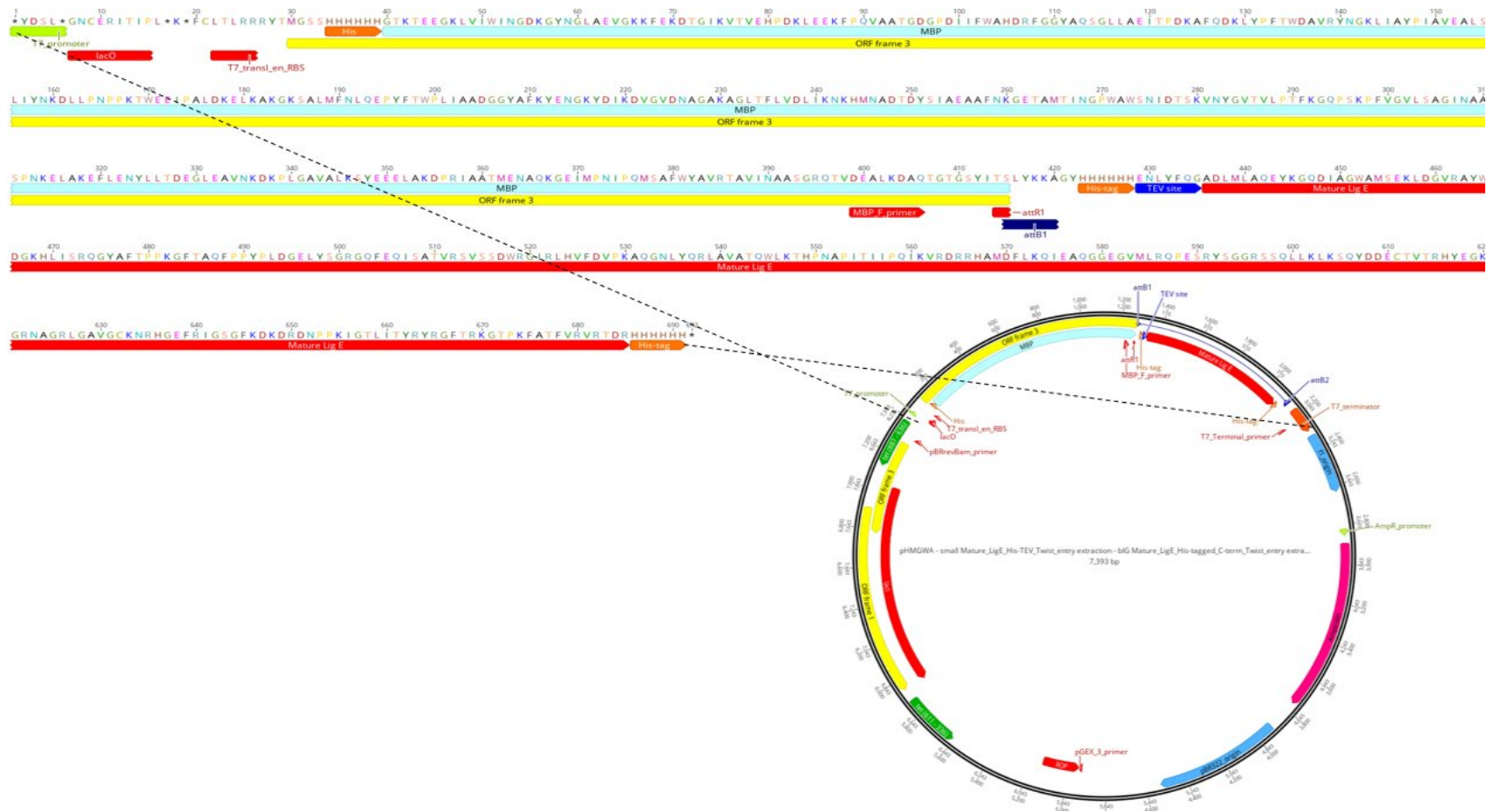
Figure B.14. His-(TEV)-Lig E-His pHMGWA expression plasmid illustrated via Geneious, which after expression and TEV cleavage will yield the C-terminally His-tagged Lig E protein. The amino acid sequence of the expressed protein from this plasmid is shown in the zoomed in section. The *attB* sites shown in blue indicate sites of insertion after cloning of the donor plasmid construct of interest into the destination plasmid.

153

Primers used for validation of the expression plasmids produced

Table B.2. Primers used for destination and expression plasmid verification

Primer name	Direction	Composition (5' to 3')	T_m (°C)
T7_promoter_primer	Forward	TAATACGACTCACTATAGGG	54.0
T7_terminator_primer	Reverse	CCGCTGAGCAATAACTAGC	56.7
MBP_forward_primer	Forward	GATGAAGCCCTG AAAGACGCGCAG	56.0

Gel-based ligation assay constructs

Table B.3. Composition of gel-based assay ligation constructs. L1 and L2 were used as the ligation constructs while the other constructs were used as complement constructs.

Construct	Composition
L1	5'-(6-carboxyfluorescein) AGGCCATGGCTGATATCGCA-3'
L2	5'-(phosphate) TAGGCATTCGAGCTCCGTCG-3'
L3	5'- CGACGGAGCTCGAATGCCTAT GCGATATCGGCCATGGCCT -3'
L6	5'-CGACGGAGCTCGAATGCCTA-3'
L7	5'-(phosphate) TGCGATATCAGCCATGGCCT-3'
L8	5'-(phosphate) ATATCAGCCATGGCCT-3'
L9	5'-CGACGGAGCTCGAATGCCTATGCG-3'
L10	5'-CGACGGAGCTCGAATGCCT ACGCGATATCAGCCATGGCCT-3'
L11	5'-CGACGGAGCTCGAATGCCTA GTGCGATATCAGCCATGGCCT-3'

Raw and additional results/data from production and assessment of recombinant Ngo-Lig

Additional results from cloning and small-scale expression of Ngo-Lig expression plasmids
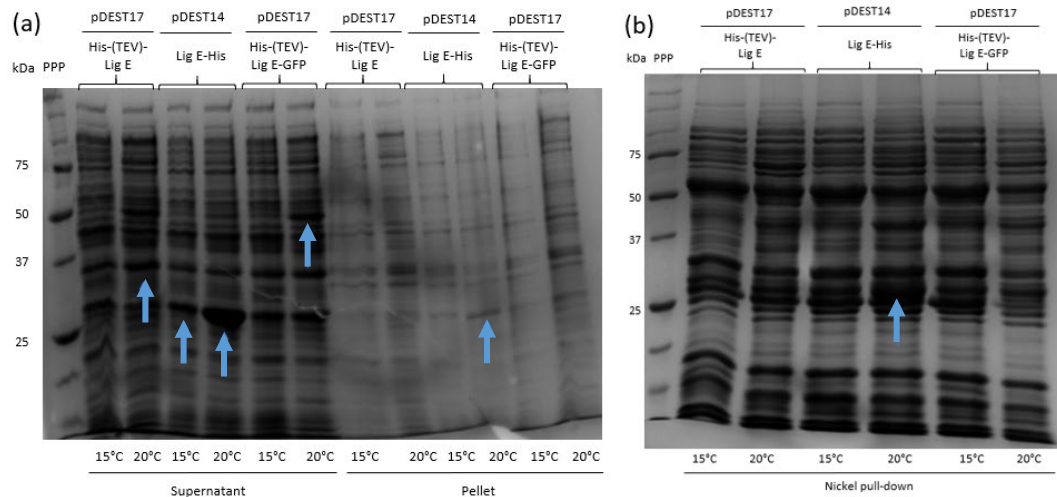
Figure B.15. SDS-PAGE (12%) of pDEST14 and pDEST17 expression plasmids after small-scale expression in Origami™(DE3) cells. The blue arrows indicate the bands of interest. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.
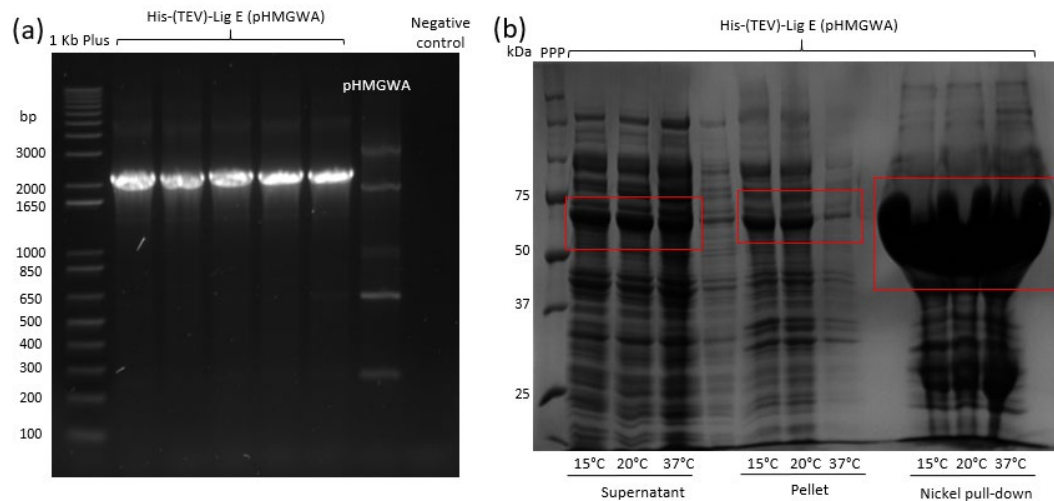
Figure B.16. Expression plasmid produced after Gateway cloning of the pDONR221 His-(TEV)-Lig E plasmid and the pHMGWA destination plasmid (a) Agarose gel (1%) of the expression plasmid to confirm plasmid integrity after amplification with the T7_promoter_primer and T7_terminator_primer primers (b) SDS-PAGE (12%) after small scale expression of the expression plasmid in BL21(DE3)pLysS cells, with red boxes indicating the bands of interest. Sizes of the 1 Kb Plus ladder and the molecular weights of the Precision Plus Protein Standard (PPP) used are labelled.

Additional Ngo-Lig protein purifications
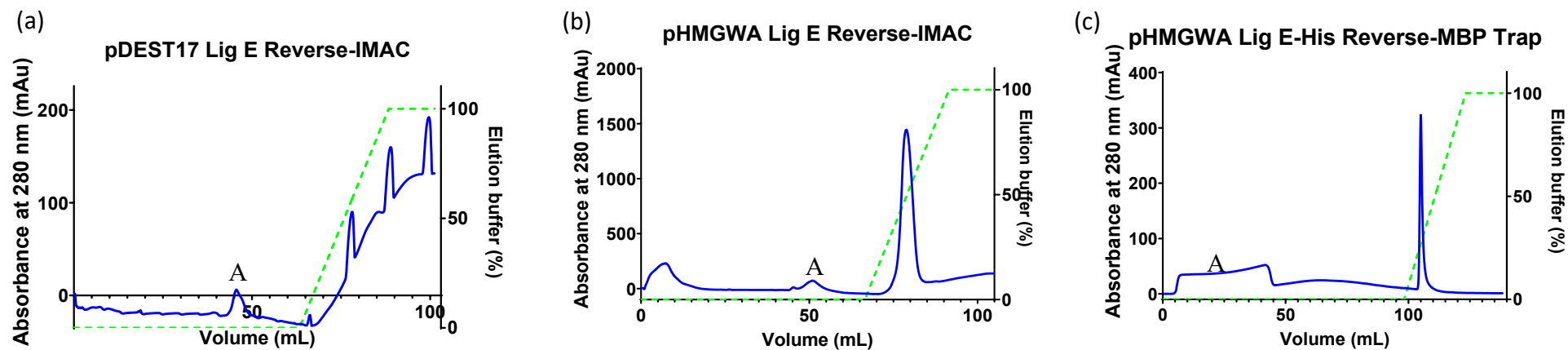
Figure B.17. Chromatograms of the final purification procedures performed on the different Lig E proteins before freezing (a) Reverse-IMAC via a nickel column of Lig E to separate the cleaved protein from the uncleaved His-(TEV)-Lig E (b) Reverse-IMAC via a nickel column of Lig E to separate the cleaved protein from the uncleaved MBP-His-(TEV)-Lig E (c) Reverse-MBP trap of Lig E-His to separate the cleaved protein from the uncleaved MBP-His-(TEV)-Lig E-His. The letter 'A' is used to show the peaks of interest on the chromatogram and their corresponding resolution on the gels in Figure 4.7. UV intensity or absorbance at 280 nm (mAu) is represented in blue and the amount of elution buffer (%) is represented in green.
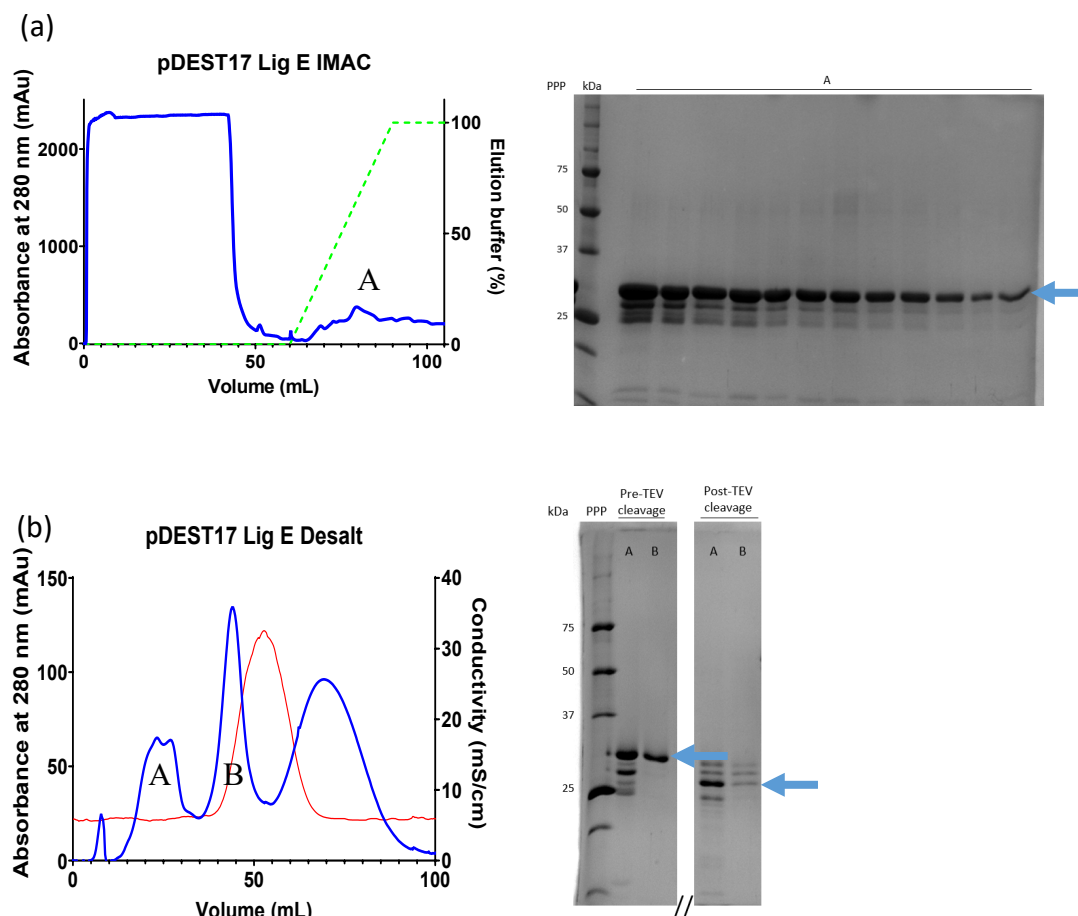
Figure B.18. Chromatogram and SDS-PAGE gels (12%) from the His-(TEV)-Lig E purification to isolate the native *N. gonorrhoeae* Lig E protein (a) IMAC via a nickel column (b) TEV cleavage after desalting to lower the salt concentration, with a cropped ('//') gel to show only the lanes of interest. Bands of interest are indicated via blue arrows. The letters 'A' and 'B' are used to show the peaks of interest on the chromatogram and their corresponding resolution on the gels. UV intensity or absorbance at 280 nm (mAu) is represented in blue, the conductivity (mS/cm) is shown in red and the amount of elution buffer (%) is represented in green. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.
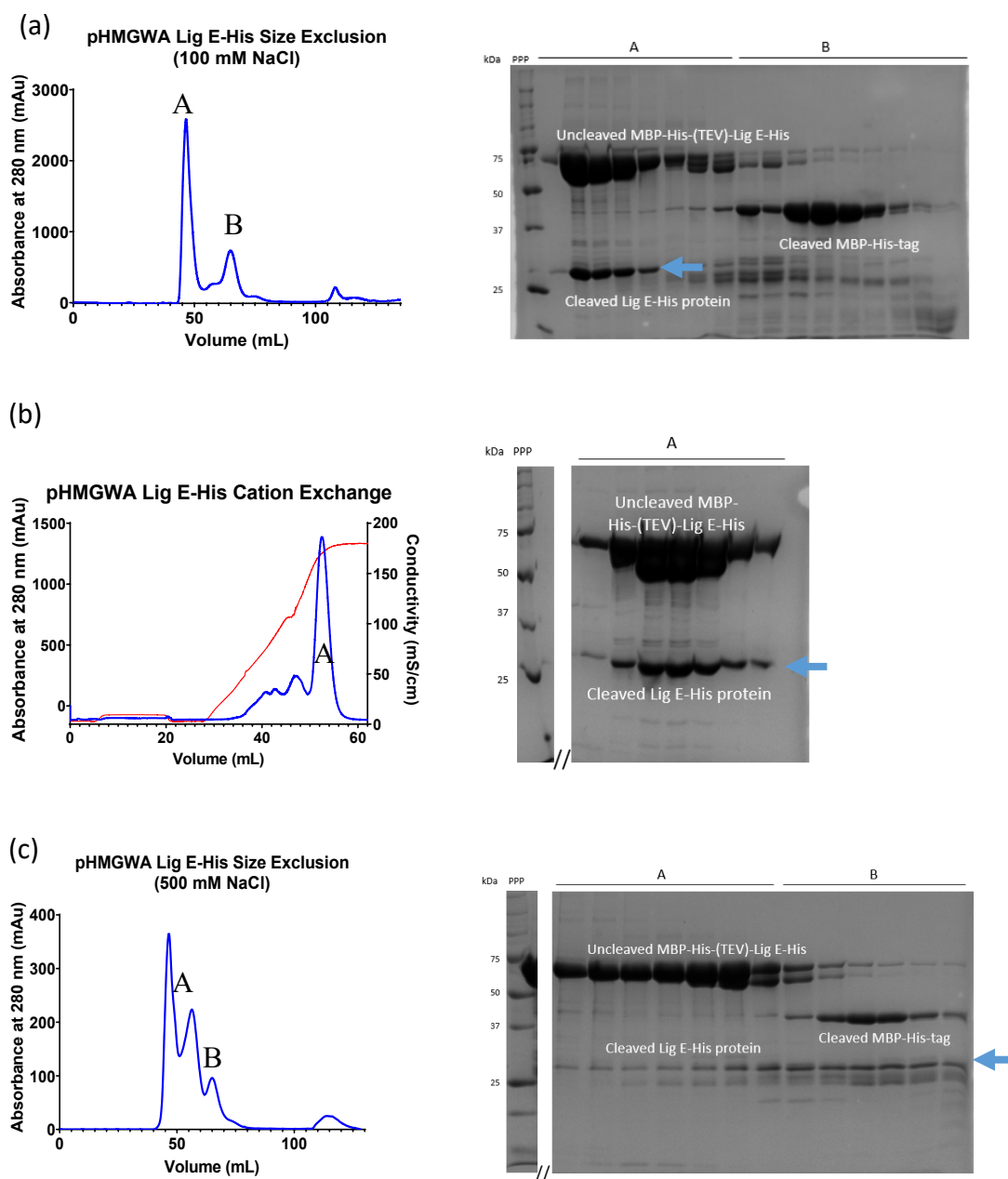
Figure B.19. Chromatograms and SDS-PAGE (12%) from the His-MBP-His-(TEV)-Lig E-His purification to isolate Lig E-His (a) Standard size exclusion with 100 mM NaCl buffer C (b) Cation exchange with buffers at pH 6.15 (c) Size exclusion with buffer C at a higher NaCl concentration (500 mM NaCl). Gels are cropped ('//') to show only the lanes of interest with the bands of interest indicated via blue arrows. The letters 'A' and 'B' are used to show the peaks of interest on the chromatogram and their corresponding resolution on the gels. UV intensity or absorbance at 280 nm (mAu) is represented in blue and the conductivity (mS/cm) is shown in red. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.

Additional results from Ngo-Lig protein quantification
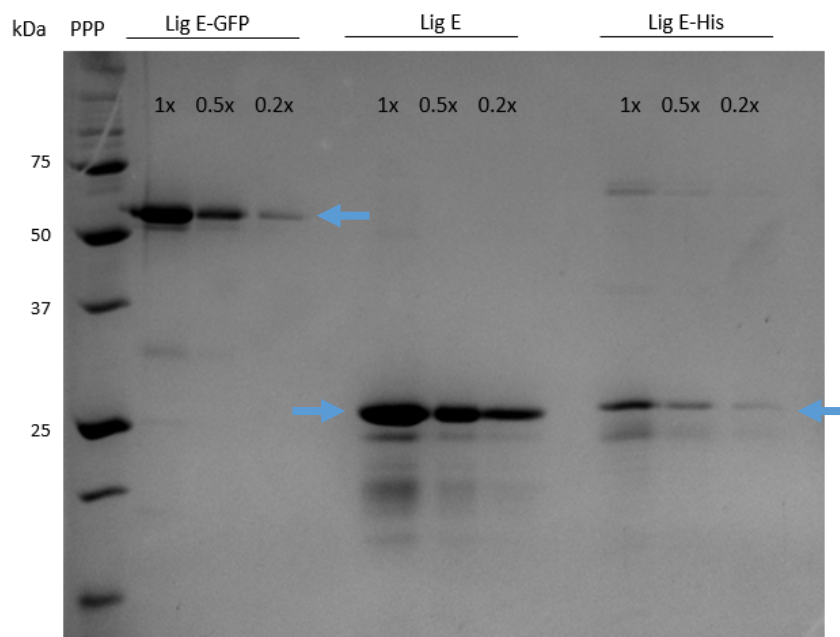
Figure B.20. SDS-PAGE (12%) of frozen Ngo-Lig protein stocks at a 1x, 0.5x and 0.2x dilution to determine the relative protein concentrations. The blue arrows indicate the expected protein sizes for each variant. Molecular weights (kDa) of the Precision Plus Protein Standard (PPP) used are labelled.

Table B.4. Calculated concentrations of frozen Ngo-Lig proteins based on the absorbance at 280 nm (via Nanodrop readings). The concentrations were calculated based on the A=εcl equation where A=absorbance, ε=extinction coefficient, c=concentration of protein and l=length of cuvette (1 cm).

Protein	Lig E-GFP	Lig E	Lig E-His
Nanodrop absorbance (A) at 280 nm	-	1.086	0.402
Extinction coefficient (ε) (M^{-1} cm^{-1}) at 280 nm (assuming all pairs of cysteines form disulphide bonds)	60530	38515	38515
Concentration of protein (μM) (c=A/ε)	-	28.19	10.43
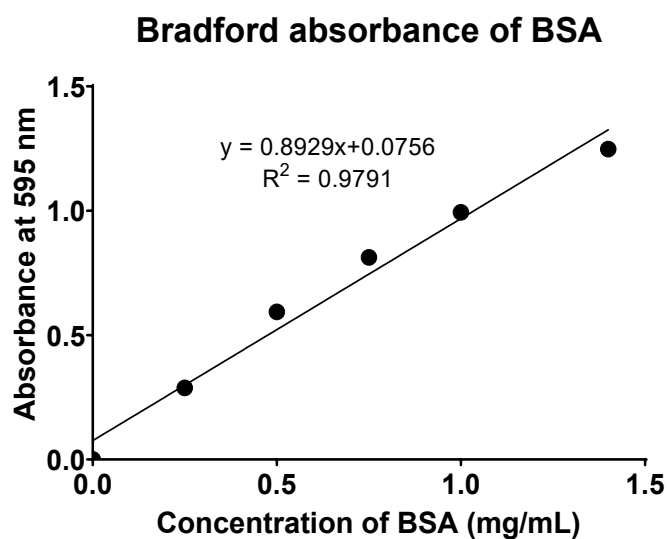
Figure B.21. Standard curve of the Bradford protein quantification assay constructed using the absorbance at 595 nm of different concentrations of bovine serum albumin (BSA) proteins

Table B.5. Raw Bradford absorbance data of frozen Ngo-Lig proteins at 595 nm. Concentrations of the frozen proteins were calculated based on these absorbances and the standard curve in Figure B.21.

Protein	Lig E	Lig E-GFP	Lig E-His	Blank
Replicate 1	0.8544	0.5891	0.4971	0.3748
Replicate 2	0.8617	0.6643	0.4862	0.3648
Replicate 3	0.881	0.6461	0.5048	0.3778
Average	0.8657	0.6332	0.496	0.3728
Average-blank	0.492	0.260	0.123	0
Bradford concentration (mg/mL)	**0.467**	**0.206**	**0.083**	-

Table B.6. Final adjusted concentrations of the frozen Ngo-Lig proteins based on the Bradford concentrations in Table B.5 and the molecular weight of the proteins

	Lig E	Lig E-GFP	Lig E-His
Molecular weight of protein (g/mol)	28411	55102	29234
Bradford concentration (mg/mL) from Table B.5	0.467	0.206	0.053
Concentration of 25% glycerol stock (μM)	16.43	3.73	1.812
Final concentration of protein (μM)	8.430	1.865	0.906

Optimal ATP concentration for Ngo-Lig ligation (before normalisation)
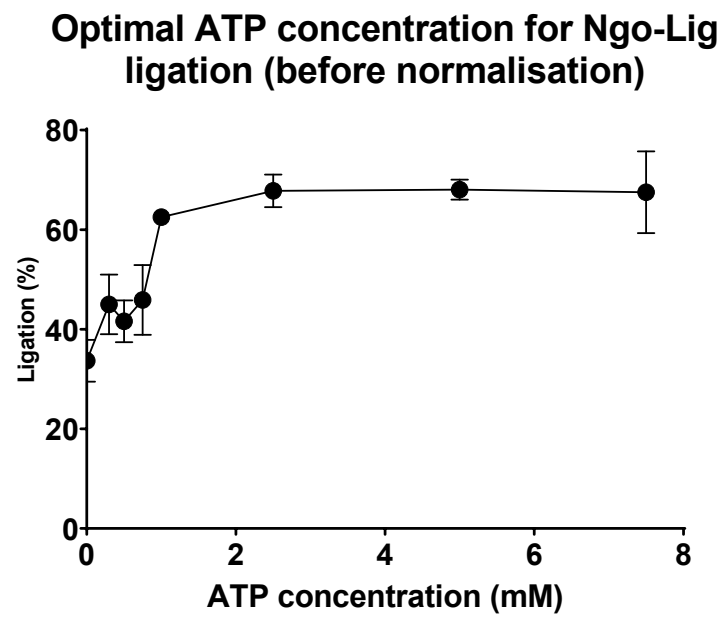
Figure B.22. Ligation activity of Ngo-Lig with different ATP concentrations before adjustment to the ligation activity at 0 mM ATP

Appendix C: Construction of reporter constructs for *N. gonorrhoeae* DNA uptake (Chapter Five)

Reporter constructs designed for gonococcal uptake
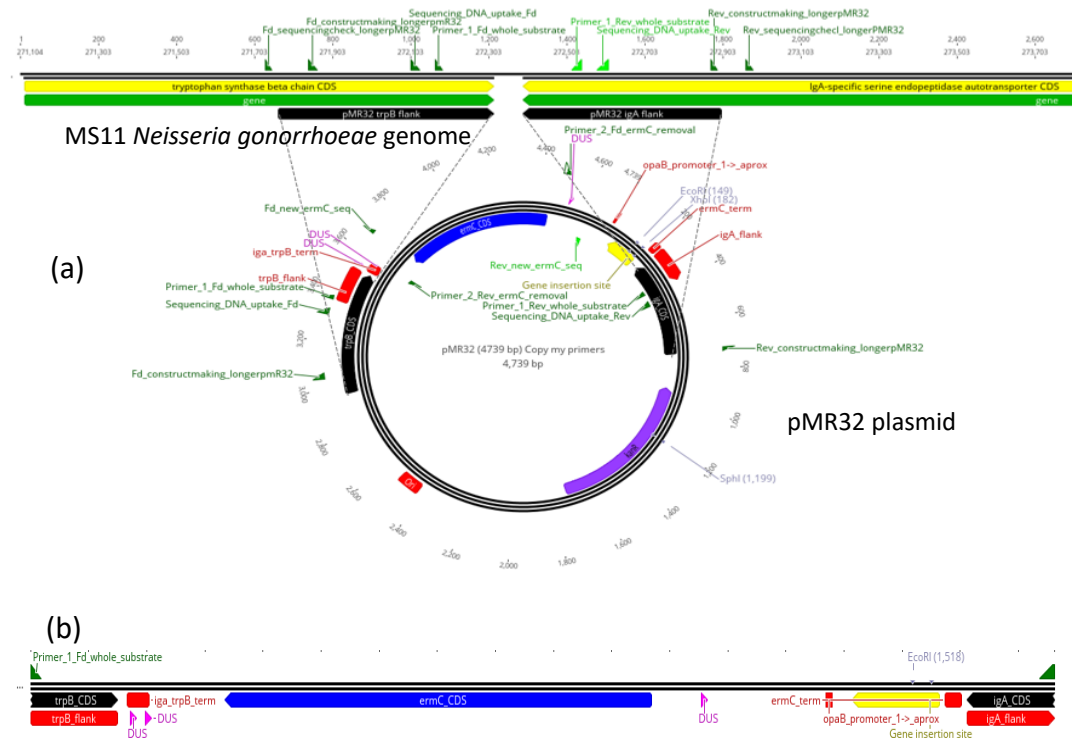
Figure C.1. pMR32 plasmid used for the DNA uptake experiments for MS11 *N. gonorrhoeae* intergenic gene insertion, with 500 bp flanking regions to the *trpB* (tryptophan synthase beta chain) and *igA* (igA-specific endopeptidase autotransporter) genes of the MS11 *N. gonorrhoeae* genome (homology shown in black), illustrated via Geneious. Any primers from Table C.1 that bind to either the pMR32 plasmid or the MS11 *N. gonorrhoeae* igA or trpB gene are shown, as well as the XhoI, EcoRI and SphI restriction sites (a) pMR32 plasmid (b) pMR32 construct 'A' amplified with the Fd_whole_susbtrate and Rev_whole_substrate primers.
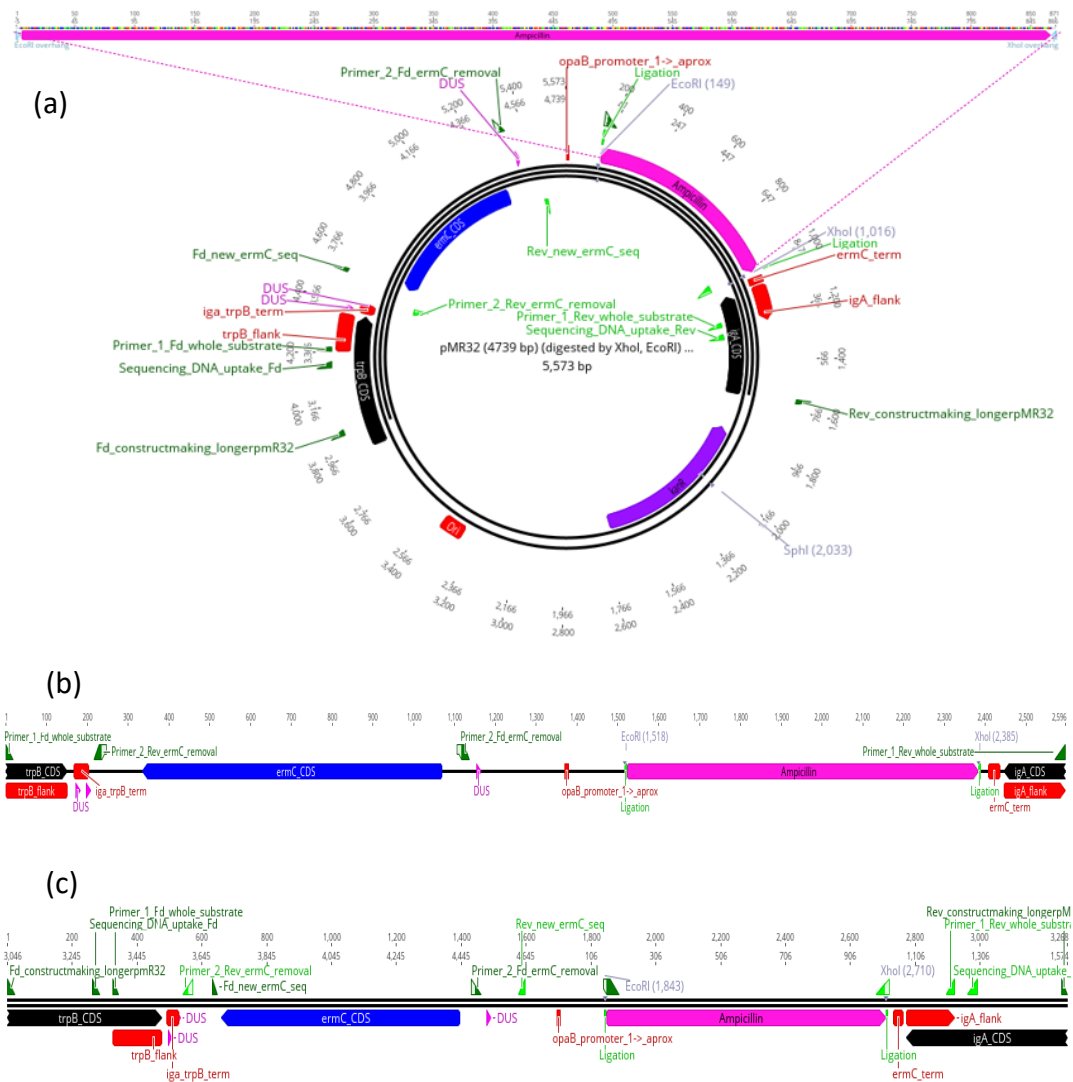
Figure C.2. pMR32:β-lactamase/pMR32:β-lac plasmid used for the DNA uptake experiments for MS11 *N. gonorrhoeae* intergenic gene insertion, with 500 bp flanking regions to the *trpB* (tryptophan synthase beta chain) and *igA* (igA-specific endopeptidase autotransporter) genes of the MS11 *N. gonorrhoeae* genome (homology shown in black), illustrated via Geneious. Any primers from Table C.1 that bind to the constructs are shown, as well as the XhoI, EcoRI and SphI restriction sites (a) *β-lactamase* construct (zoomed in), ligated onto the pMR32 plasmid via EcoRI and XhoI sites (b) pMR32:β-lac construct 'A' amplified with the Fd_whole_substrate and Rev_whole_substrate primers (c) pMR32:β-lac longer construct 'A' amplified with the Fd_constructmaking_longerpMR32 and Rev_constructmaking_longerpMR32 primers.
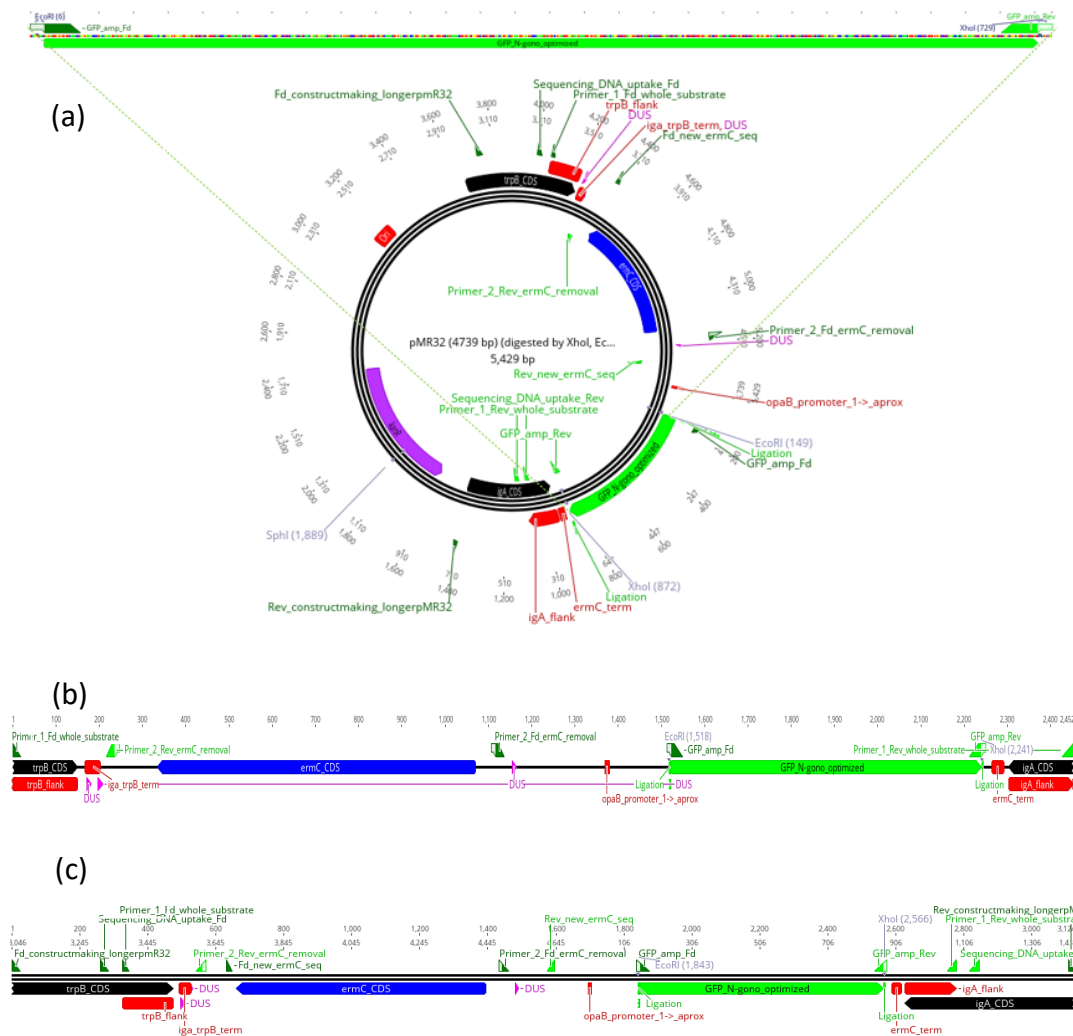
Figure C.3. pMR32:GFP plasmid used for the DNA uptake experiments for MS11 *N. gonorrhoeae* intergenic gene insertion, with 500 bp flanking regions to the *trpB* (tryptophan synthase beta chain) and *igA* (igA-specific endopeptidase autotransporter) genes of the MS11 *N. gonorrhoeae* genome (homology shown in black), illustrated via Geneious. Any primers from Table C.1 that bind to the constructs are shown, as well as the XhoI, EcoRI and SphI restriction sites (a) *GFP* construct (zoomed in), ligated onto the pMR32 plasmid via EcoRI and XhoI sites (b) pMR32:GFP construct 'A' amplified with the Fd_whole_substrate and Rev_whole_substrate primers (c) pMR32:GFP longer construct 'A' amplified with the Fd_constructmaking_longerpMR32 and Rev_constructmaking_longerpMR32 primers.

Primers used for construction of reporter DNA uptake constructs and for verification of construct integration after *N. gonorrhoeae* transformations

Table C.1. Primers ordered from Invitrogen™ used for construct making and transformation verification for the MS11 *N. gonorrhoeae* DNA uptake experiment

Primer name	Direction	Composition (5' to 3')	T$_m$ (°C)
GFP_amp_Fd	Forward	EcoRI ACTGG|AATT|CATGAGTAA AGGAGAAGAACTTTTCACT *GFP* 5' binding region	56.0
GFP_amp_Rev	Reverse	XhoI ACTGC|TCGA|GTTATTTGT ATAGTTCATCCATGCCATG *GFP* 3' binding region	58.0
Fd_whole_substrate	Forward	ATCCCCGCGCTCGAATCC	50.0
Rev_whole_substrate	Reverse	GCCTATAAAGCAG ATAATCAACAGCA	50.0
Fd_constructmaking _longerpMR32	Forward	CGGTTTGTTCTA TCCCTATATCG	52.6
Rev_constuctmaking _longerpMR32	Reverse	GCCCGCACCAATATCGC	57.3
Sequencing_DNA_uptake _Fd	Forward	CACCGTTGCCAA AGATGACGAAGC	54.0
Sequencing_DNA_uptake _Rev	Reverse	CCCTTGTTATCTGCA GCTTACTTTGCCAA	55.0
Fd_sequencingcheck _longerpMR32	Forward	GCGTGATTTCCAATGCG	52.3
Rev_sequencingcheck _longerpMR32	Reverse	GCCTGATCAAAAGTATGC	48.5
Fd_new_ermC_seq	Forward	CGAGATTTTCAGGAGC	46.9
Rev_new_ermC_seq	Reverse	CGTACCATTTG TACTGTCTGC	53.5

Additional results/data from transformation of the reporter DNA uptake construct in *N. gonorrhoeae*
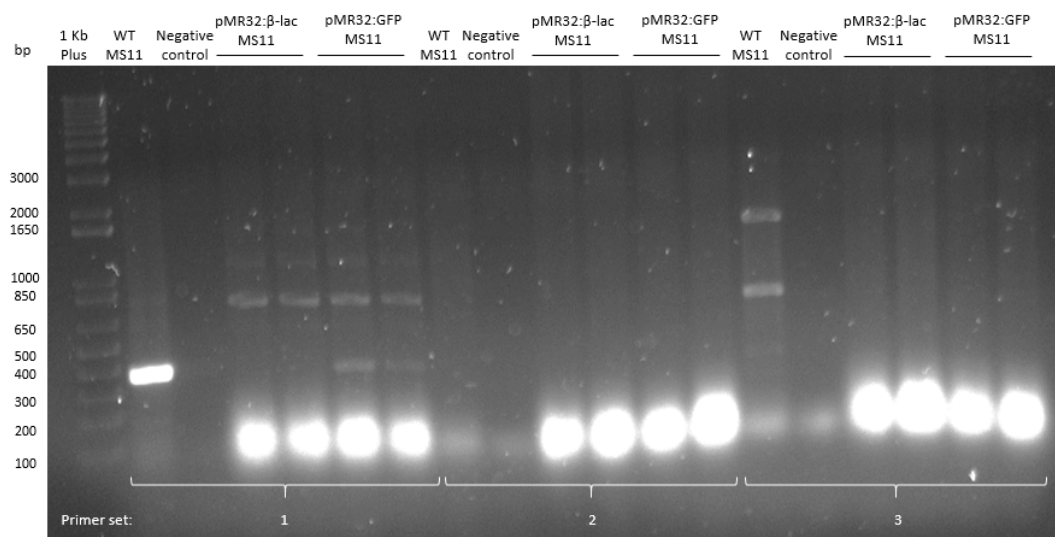
Figure C.4. Agarose gel (1%) of MS11 *N. gonorrhoeae* colonies after transformations with the pMR32:β-lac and pMR32:GFP constructs 'A'. Three different internal primer sets were used to check if the constructs had integrated elsewhere in the genome. The expected sizes for the wild-type (WT) MS11 *N. gonorrhoeae*, pMR32:β-lac and pMR32:GFP constructs 'A' transformants with primer set 1 (Fd_whole_substrate and Rev_whole_susbtrate) were 375, 2632 and 2438 bp respectively, 0, 0 and 737 bp respectively with primer set 2 (GFP_amp_Fd and GFP_amp_Rev) and 0, 1490 and 1346 bp respectively with primer set 3 (Fd_ermC_removal and Rev_whole_substrate). Sizes of the 1 Kb Plus ladder used are labelled.

Appendix D: Standards/ladders
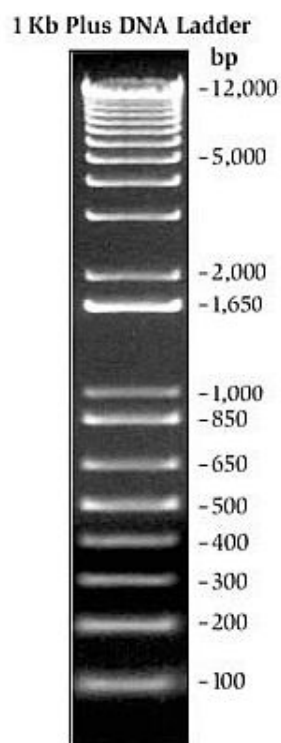
Figure D.11. Kb Plus DNA ladder (Invitrogen™) used for agarose gel runs
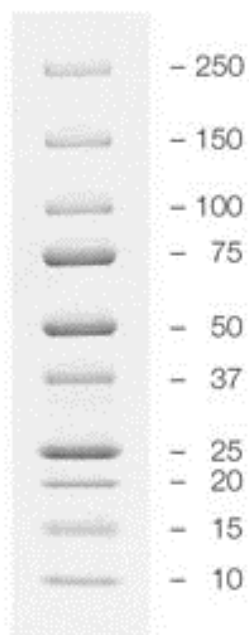
Figure D.2. Precision Plus Protein™ All Blue Prestained Protein standard (BioRad) (annotated as PPP in all SDS-PAGE gels), used for protein gel electrophoreses. Numbers correspond to the protein sizes in kDa.
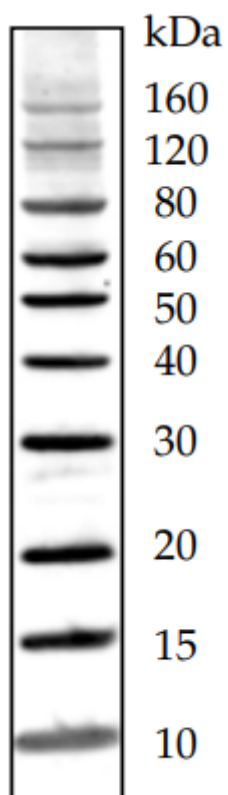
Figure D.3. BenchMark[TM] His-tagged Protein standard (ThermoFischer Scientific) used for western blots, which can be tagged with anti-His primary antibodies

References

Abbasian, B., Shair, A., O'Gorman, D. B., Pena-Diaz, A. M., Brennan, L., Engelbrecht, K., Koenig, D. W., Reid, G., & Burton, J. P. (2019). Potential role of extracellular ATP released by bacteria in bladder infection and contractility. *mSphere, 4*(5).

Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotonics International, 11*(7), 36-42.

Adams, D. W., Stutzmann, S., Stoudmann, C., & Blokesch, M. (2019). DNA-uptake pili of Vibrio cholerae are required for chitin colonization and capable of kin recognition via sequence-specific self-interaction. *Nature Microbiology, 4*(9), 1545-1557.

Ambur, O. H., Frye, S. A., & Tønjum, T. (2007). New functional identity for the DNA uptake sequence in transformation and its presence in transcriptional terminators. *Journal of Bacteriology, 189*(5), 2077-85.

Anderson, M. T., & Seifert, H. S. (2011). Opportunity and means: horizontal gene transfer from the human host to a bacterial pathogen. *mBio, 2*(1), e00005-e11.

Ball, L. M., & Criss, A. K. (2013). Constitutively Opa-expressing and Opa-deficient neisseria gonorrhoeae strains differentially stimulate and survive exposure to human neutrophils. *Journal of Bacteriology, 195*(13), 2982-2990.

Barth, H. G., Jackson, C., & Boyes, B. E. (1994). Size exclusion chromatography. *Analytical Chemistry, 66*(12), 595-620.

Booth, W. T., Schlachter, C. R., Pote, S., Ussin, N., Mank, N. J., Klapper, V., Offermann, L. R., Tang, C., Hurlburt, B. K., & Chruszcz, M. (2018). Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega, 3*(1), 760-768.

Boyle-Vavra, S., & Seifert, H. S. (1993). Shuttle mutagenesis: two mini-transposons for gene mapping and for lacZ transcriptional fusions in Neisseria gonorrhoeae. *Gene, 129*(1), 51-7.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry, 72*(1), 248-254.

Brett, M. S., Davies, H. G., Blockley, J. R., & Heffernan, H. M. (1992). Antibiotic susceptibilities, serotypes and auxotypes of Neisseria gonorrhoeae isolated in New Zealand. *Genitourinary Medicine, 68*(5), 321.

Cabantous, S., Terwilliger, T. C., & Waldo, G. S. (2005). Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology, 23*(1), 102-107.

Cabantous, S., & Waldo, G. S. (2006). In vivo and in vitro protein solubility assays using split GFP. *Nature Methods, 3*(10), 845-854.

Callaghan, M. M., Heilers, J. H., van der Does, C., & Dillard, J. P. (2017). Secretion of chromosomal DNA by the Neisseria gonorrhoeae type IV secretion system. *Current Topics in Microbiology and Immunology, 413*, 323-345.

Catlin, B. W. (1973). Nutritional profiles of Neisseria gonorrhoeae, Neisseria meningitidis, and Neisseria lactamica in chemically defined media and the use of growth requirements for gonococcal typing. *The Journal of Infectious Diseases, 128*(2), 178-94.

Cehovin, A., & Lewis, S. B. (2017). Mobile genetic elements in Neisseria gonorrhoeae: movement for change. *Pathogens and Disease, 75*(6).

Chaussee, M. S., & Hill, S. A. (1998). Formation of single-stranded DNA during DNA transformation of Neisseria gonorrhoeae. *Journal of Bacteriology, 180*(19), 5117-5122.

Chen, I., & Gotschlich, E. C. (2001). ComE, a competence protein from Neisseria gonorrhoeae with DNA-binding activity. *Journal of Bacteriology, 183*(10), 3160-8.

Chen, S., Larsson, M., Robinson, R. C., & Chen, S. L. (2017). Direct and convenient measurement of plasmid stability in lab and clinical isolates of E. coli. *Scientific Reports, 7*(1), 4788.

Chen, Y., Liu, H., Yang, C., Gao, Y., Yu, X., Chen, X., Cui, R., Zheng, L., Li, S., Li, X., Ma, J., Huang, Z., Li, J., & Gan, J. (2019). Structure of the error-prone DNA ligase of African swine fever virus identifies critical active site residues. *Nature Communications, 10*(1), 387.

Cheng, C., & Shuman, S. (1997). Characterization of an ATP-dependent DNA ligase encoded by Haemophilus influenzae. *Nucleic Acids Research, 25*(7), 1369-1374.

Christodoulides, M., Everson, J. S., Liu, B. L., Lambden, P. R., Watt, P. J., Thomas, E. J., & Heckels, J. E. (2000). Interaction of primary human endometrial cells with Neisseria gonorrhoeae expressing green fluorescent protein. *Molecular Microbiology, 35*(1), 32-43.

D'Ambrozio, J. A. (2015). *Insights into the enhanced in vivo fitness of Neisseria gonorrhoeae driven by a fluoroquinolone resistance-conferring mutant DNA gyrase.* Doctoral thesis, Uniformed Services University of the Health Sciences.

Dalbey, R. E. (1991). Leader peptidase. *Molecular Microbiology, 5*(12), 2855-2860.

de Almeida, J. M., Moure, V. R., Müller-Santos, M., de Souza, E. M., Pedrosa, F. O., Mitchell, D. A., & Krieger, N. (2018). Tailoring recombinant lipases: keeping the His-tag favors esterification reactions, removing it favors hydrolysis reactions. *Scientific Reports, 8*(1), 10000.

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. *Molecular Membrane Biology, 31*(2-3), 58-84.

Devaraj, A., Buzzo, J. R., Mashburn-Warren, L., Gloag, E. S., Novotny, L. A., Stoodley, P., Bakaletz, L. O., & Goodman, S. D. (2019). The extracellular DNA lattice of bacterial biofilms is structurally related to Holliday junction recombination intermediates. *Proceedings of the National Academy of Sciences, 116*(50), 25068-25077.

Dillard, J. P. (2011). Genetic manipulation of Neisseria gonorrhoeae. *Current Protocols in Microbiology, Chapter 4*, Unit4A.2-Unit4A.2.

Dillard, J. P., & Seifert, H. S. (2001). A variable genetic island specific for Neisseria gonorrhoeae is involved in providing DNA for natural transformation and is found more often in disseminated infection isolates. *Molecular Microbiology, 41*(1), 263-77.

Doherty, A. J., & Suh, S. W. (2000). Structural and mechanistic conservation in DNA ligases. *Nucleic Acids Research, 28*(21), 4051-8.

Doherty, A. J., & Wigley, D. B. (1999). Functional domains of an ATP-dependent DNA ligase. *Journal of Molecular Biology, 285*(1), 63-71.

Domenech, M., Pedrero-Vega, E., Prieto, A., & García, E. (2016). Evidence of the presence of nucleic acids and β-glucan in the matrix of non-typeable Haemophilus influenzae in vitro biofilms. *Scientific Reports, 6*(1), 36424.

Dougherty, W. G., Cary, S. M., & Dawn Parks, T. (1989). Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology, 171*(2), 356-364.

Dragan, A. I., Casas-Finet, J. R., Bishop, E. S., Strouse, R. J., Schenerman, M. A., & Geddes, C. D. (2010). Characterization of PicoGreen interaction with dsDNA and the origin of its fluorescence enhancement upon binding. *Biophysical Journal, 99*(9), 3010-3019.

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), 201.

Edwards, J. L., & Apicella, M. A. (2004). The molecular mechanisms used by Neisseria gonorrhoeae to initiate infection differ between men and women. *Clinical Microbiology Reviews, 17*(4), 965-81, table of contents.

Engelmoer, D. J. P., & Rozen, D. E. (2011). Competence increases survival during stress in Streptococcus pneumoniae. *Evolution, 65*(12), 3475-3485.

Falsetta, M. L., Steichen, C. T., McEwan, A. G., Cho, C., Ketterer, M., Shao, J., Hunt, J., Jennings, M. P., & Apicella, M. A. (2011). The composition and

metabolic phenotype of Neisseria gonorrhoeae biofilms. *Frontiers in Microbiology, 2*, 75-75.

Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H., & Phillips, G. J. (2000). Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. *Journal of Bacteriology, 182*(14), 4068-76.

Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in E. coli. *Current Protocols in Protein Science, 61*(1), 05.24.01-05.24.29.

Freudl, R. (2018). Signal peptides for recombinant protein secretion in bacterial expression systems. *Microbial Cell Factories, 17*(1), 52.

Gianecini, R., Oviedo, C., Stafforini, G., & Galarza, P. (2016). Neisseria gonorrhoeae resistant to ceftriaxone and cefixime, Argentina. *Emerging Infectious Diseases, 22*(6), 1139-1141.

Gilmore, M. S., & Haas, W. (2005). The selective advantage of microbial fratricide. *Proceedings of the National Academy of Sciences of the United States of America, 102*(24), 8401.

Gong, C., Martins, A., Bongiorno, P., Glickman, M., & Shuman, S. (2004). Biochemical and genetic analysis of the four DNA ligases of mycobacteria. *Journal of Biological Chemistry, 279*(20), 20594-606.

Goodman, S. D., & Scocca, J. J. (1988). Identification and arrangement of the DNA sequence recognized in specific transformation of Neisseria gonorrhoeae. *Proceedings of the National Academy of Sciences of the United States of America, 85*(18), 6982-6986.

Greiner, L. L., Edwards, J. L., Shao, J., Rabinak, C., Entz, D., & Apicella, M. A. (2005). Biofilm formation by Neisseria gonorrhoeae. *Infection and Immunity, 73*(4), 1964-1970.

Griffin, P. J., & Rieder, S. V. (1957). A study on the growth requirements of Neisseria gonorrhoeae and its clinical application. *Yale Journal of Biology and Medicine, 29*(6), 613-21.

Grossman, T. H., Kawasaki, E. S., Punreddy, S. R., & Osburne, M. S. (1998). Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene, 209*(1-2), 95-103.

Günther, K., Mertig, M., & Seidel, R. (2010). Mechanical and structural properties of YOYO-1 complexed DNA. *Nucleic Acids Research, 38*(19), 6526-6532.

Hall, B., Acar Kirit, H., Nandipati, A., & Barlow, M. (2013). Growth rates made easy. *Molecular Biology and Evolution, 31*.

Hamilton, H. L., & Dillard, J. P. (2006). Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Molecular Microbiology, 59*(2), 376-385.

Hamilton, H. L., Domínguez, N. M., Schwartz, K. J., Hackett, K. T., & Dillard, J. P. (2005). Neisseria gonorrhoeae secretes chromosomal DNA via a novel type IV secretion system. *Molecular Microbiology, 55*(6), 1704-21.

Hamilton, H. L., Schwartz, K. J., & Dillard, J. P. (2001). Insertion-duplication mutagenesis of neisseria: use in characterization of DNA transfer genes in the gonococcal genetic island. *Journal of Bacteriology, 183*(16), 4718-4726.

Han, X., Wang, E., Cui, Y., Lin, Y., Chen, H., An, R., Liang, X., & Komiyama, M. (2019). The staining efficiency of cyanine dyes for single-stranded DNA is enormously dependent on nucleotide composition. *Electrophoresis, 40*(12-13), 1708-1714.

Haney, E. F., Trimble, M. J., Cheng, J. T., Vallé, Q., & Hancock, R. E. W. (2018). Critical assessment of methods to quantify biofilm growth and evaluate antibiofilm activity of host defence peptides. *Biomolecules, 8*(2), 29.

Harmsen, M., Lappann, M., Knøchel, S., & Molin, S. (2010). Role of extracellular DNA during biofilm formation by *Listeria monocytogenes*. *Applied and Environmental Microbiology, 76*(7), 2271-2279.

Hashem, V. I., Klysik, E. A., Rosche, W. A., & Sinden, R. R. (2002). Instability of repeated DNAs during transformation in Escherichia coli. *Mutation Research, 502*(1-2), 39-46.

Hepp, C., Gangel, H., Henseler, K., Günther, N., & Maier, B. (2016). Single-stranded DNA uptake during gonococcal transformation. *Journal of Bacteriology, 198*(18), 2515-2523.

Holmberg, M., Hansen, T. S., Lind, J. U., & Hjortø, G. M. (2012). Increased adsorption of histidine-tagged proteins onto tissue culture polystyrene. *Colloids and Surfaces B: Biointerfaces, 92*, 286-92.

Hommelsheim, C. M., Frantzeskakis, L., Huang, M., & Ülker, B. (2014). PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications. *Scientific Reports, 4*(1), 5052.

Imhaus, A.-F., & Duménil, G. (2014). The number of Neisseria meningitidis type IV pili determines host cell interaction. *The EMBO Journal, 33*(16), 1767-1783.

Jakubovics, N. S., Shields, R. C., Rajarajan, N., & Burgess, J. G. (2013). Life after death: the critical role of extracellular DNA in microbial biofilms. *Letters in Applied Microbiology, 57*(6), 467-475.

Johannsen, D. B., Johnston, D. M., Koymen, H. O., Cohen, M. S., & Cannon, J. G. (1999). A *Neisseria gonorrhoeae* immunoglobulin A1 protease mutant is infectious in the human challenge model of urethral infection. *Infection and Immunity, 67*(6), 3009.

Karygianni, L., Ren, Z., Koo, H., & Thurnheer, T. (2020). Biofilm matrixome: extracellular components in structured microbial communities. *Trends in Microbiology, 28*(8), 668-681.

Karyolaimos, A., Ampah-Korsah, H., Hillenaar, T., Mestre Borras, A., Dolata, K. M., Sievers, S., Riedel, K., Daniels, R., & de Gier, J.-W. (2019). Enhancing recombinant protein yields in the E. coli periplasm by combining signal peptide and production rate screening. *Frontiers in Microbiology, 10*(1511).

Kataeva, I., Chang, J., Xu, H., Luan, C. H., Zhou, J., Uversky, V. N., Lin, D., Horanyi, P., Liu, Z. J., Ljungdahl, L. G., Rose, J., Luo, M., & Wang, B. C. (2005). Improving solubility of Shewanella oneidensis MR-1 and Clostridium thermocellum JW-20 proteins expressed into Esherichia coli. *Journal of Proteome Research, 4*(6), 1942-51.

Katzen, F. (2007). Gateway® recombinational cloning: a biological operating system. *Expert Opinion on Drug Discovery, 2*(4), 571-589.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-9.

Kieleczawa, J. (2006). Fundamentals of sequencing of difficult templates--an overview. *Journal of Biomolecular Techniques, 17*(3), 207-217.

Kimelman, A., Levy, A., Sberro, H., Kidron, S., Leavitt, A., Amitai, G., Yoder-Himes, D. R., Wurtzel, O., Zhu, Y., Rubin, E. M., & Sorek, R. (2012). A vast collection of microbial genes that are toxic to bacteria. *Genome Research, 22*(4), 802-809.

Koomey, J. M., & Falkow, S. (1987). Cloning of the recA gene of Neisseria gonorrhoeae and construction of gonococcal recA mutants. *Journal of Bacteriology, 169*(2), 790-5.

Kouzel, N., Oldewurtel, E. R., & Maier, B. (2015). Gene transfer efficiency in gonococcal biofilms: role of biofilm age, architecture, and pilin antigenic variation. *Journal of Bacteriology, 197*(14), 2422-2431.

Kurien, B., & Scofield, R. (1999). Heat mediated quick Coomassie blue protein staining and destaining of SDS-PAGE gels. *Indian Journal of Biochemistry & Biophysics, 35*, 385-9.

Kwiatek, A., Bacal, P., Wasiluk, A., Trybunko, A., & Adamczyk-Poplawska, M. (2014). The dam replacing gene product enhances Neisseria gonorrhoeae FA1090 viability and biofilm formation. *Frontiers in Microbiology, 5*(712).

Lee, R. S., Seemann, T., Heffernan, H., Kwong, J. C., Gonçalves da Silva, A., Carter, G. P., Woodhouse, R., Dyet, K. H., Bulach, D. M., Stinear, T. P., Howden, B. P., & Williamson, D. A. (2017). Genomic epidemiology and

antimicrobial resistance of Neisseria gonorrhoeae in New Zealand. *Journal of Antimicrobial Chemotherapy, 73*(2), 353-364.

Lim, H. N., Lee, Y., & Hussein, R. (2011). Fundamental relationship between operon organization and gene expression. *Proceedings of the National Academy of Sciences, 108*(26), 10626-10631.

Lipman, N. S., Jackson, L. R., Trudel, L. J., & Weis-Garcia, F. (2005). Monoclonal versus polyclonal antibodies: distinguishing characteristics, applications, and information resources. *The ILAR Journal, 46*(3), 258-268.

Lohman, G. J., Zhang, Y., Zhelkovsky, A. M., Cantor, E. J., & Evans, T. C., Jr. (2014). Efficient DNA ligation in DNA-RNA hybrid helices by Chlorella virus DNA ligase. *Nucleic Acids Research, 42*(3), 1831-44.

Low, N., Unemo, M., Skov Jensen, J., Breuer, J., & Stephenson, J. M. (2014). Molecular diagnostics for gonorrhoea: implications for antimicrobial resistance and the threat of untreatable gonorrhoea. *PLoS Medicine, 11*(2), e1001598.

Lu, Q.-F., Cao, D.-M., Su, L.-L., Li, S.-B., Ye, G.-B., Zhu, X.-Y., & Wang, J.-P. (2019). Genus-wide comparative genomics analysis of *Neisseria* to identify new genes associated with pathogenicity and niche adaptation of *Neisseria* pathogens. *International Journal of Genomics, 2019*, 6015730.

Madshus, I. H. (1988). Regulation of intracellular pH in eukaryotic cells. *The Biochemical Journal, 250*(1), 1-8.

Magnet, S., & Blanchard, J. S. (2004). Mechanistic and kinetic study of the ATP-dependent DNA ligase of Neisseria meningitidis. *Biochemistry, 43*(3), 710-717.

Majoul, I., Ferrari, D., & Söling, H.-D. (1997). Reduction of protein disulfide bonds in an oxidizing environment. *FEBS Letters, 401*(2-3), 104-108.

Malherbe, G., Humphreys, D. P., & Davé, E. (2019). A robust fractionation method for protein subcellular localization studies in Escherichia coli. *BioTechniques, 66*(4), 171-178.

Mansfield, E. S., Worley, J. M., McKenzie, S. E., Surrey, S., Rappaport, E., & Fortina, P. (1995). Nucleic acid detection using non-radioactive labelling methods. *Molecular and Cellular Probes, 9*(3), 145-156.

Marinus, M. G., & Løbner-Olesen, A. (2014). DNA methylation. *EcoSal Plus, 6*(1).

Marri, P. R., Paniscus, M., Weyand, N. J., Rendón, M. A., Calton, C. M., Hernández, D. R., Higashi, D. L., Sodergren, E., Weinstock, G. M., Rounsley, S. D., & So, M. (2010). Genome sequencing reveals widespread virulence gene exchange among human Neisseria species. *PLoS One, 5*(7), e11835.

Matthey, N., & Blokesch, M. (2016). The DNA-uptake process of naturally competent *Vibrio cholerae*. *Trends in Microbiology, 24*(2), 98-110.

Mempin, R., Tran, H., Chen, C., Gong, H., Kim Ho, K., & Lu, S. (2013). Release of extracellular ATP by bacteria during growth. *BMC Microbiology, 13*, 301.

Meng, L., Liu, Y., Yin, X., Zhou, H., Wu, J., Wu, M., & Yang, L. (2020). Effects of his-tag on catalytic activity and enantioselectivity of recombinant transaminases. *Applied Biochemistry and Biotechnology, 190*(3), 880-895.

Nair, P. A., Nandakumar, J., Smith, P., Odell, M., Lima, C. D., & Shuman, S. (2007). Structural basis for nick recognition by a minimal pluripotent DNA ligase. *Nature Structural & Molecular Biology, 14*(8), 770-8.

Nakane, P. K., & Kawaoi, A. (1974). Peroxidase-labeled antibody. A new method of conjugation. *Journal of Histochemistry & Cytochemistry, 22*(12), 1084-91.

Nallamsetty, S., Kapust, R. B., Tözsér, J., Cherry, S., Tropea, J. E., Copeland, T. D., & Waugh, D. S. (2004). Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expression and Purification, 38*(1), 108-15.

Nandakumar, J., Nair, P. A., & Shuman, S. (2007). Last stop on the road to repair: structure of *E. coli* DNA ligase bound to nicked DNA-adenylate. *Molecular Cell, 26*(2), 257-271.

Nielsen, H. (2017). Predicting secretory proteins with SignalP. In D. Kihara (Ed.), *Protein function prediction: methods and protocols* (pp. 59-73). New York, NY: Springer New York.

Noble, J. E., Knight, A. E., Reason, A. J., Di Matola, A., & Bailey, M. J. A. (2007). A comparison of protein quantitation assays for biopharmaceutical applications. *Molecular Biotechnology, 37*(2), 99-111.

O'Toole, G. A. (2011). Microtiter dish biofilm formation assay. *Journal of Visualized Experiments : JoVE*(47), 2437.

Odell, M., Sriskanda, V., Shuman, S., & Nikolov, D. B. (2000). Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining. *Molecular Cell, 6*(5), 1183-93.

Ohnishi, M., Saika, T., Hoshina, S., Iwasaku, K., Nakayama, S., Watanabe, H., & Kitawaki, J. (2011). Ceftriaxone-resistant Neisseria gonorrhoeae, Japan. *Emerging Infectious Diseases, 17*(1), 148-9.

Passarinha, L., Diogo, M., Queiroz, J., Monteiro, G., Fonseca, L., & Prazeres, D. (2006). Production of CoIE1 type plasmid by Escherichia coli DH5alpha cultured under nonselective conditions. *Journal of Microbiology and Biotechnology, 16*(1), 20.

Patel, A. L., Chaudhry, U., Sachdev, D., Sachdeva, P. N., Bala, M., & Saluja, D. (2011). An insight into the drug resistance profile & mechanism of drug resistance in Neisseria gonorrhoeae. *The Indian Journal of Medical Research, 134*(4), 419-431.

Pena, M. M., Teper, D., Ferreira, H., Wang, N., Sato, K. U., Ferro, M. I. T., & Ferro, J. A. (2020). mCherry fusions enable the subcellular localization of periplasmic and cytoplasmic proteins in Xanthomonas sp. *PLoS One, 15*(7), e0236185.

Pergolizzi, G., Wagner, G. K., & Bowater, R. P. (2016). Biochemical and structural characterisation of DNA ligases from bacteria and archaea. *Bioscience Reports, 36*(5), 00391-00391.

Perron, G. G., Lee, A. E., Wang, Y., Huang, W. E., & Barraclough, T. G. (2012). Bacterial recombination promotes the evolution of multi-drug-resistance in functionally diverse populations. *Proceedings of the Royal Society B: Biological Sciences, 279*(1733), 1477-84.

Phillips, G. J. (2001). Green fluorescent protein – a bright idea for the study of bacterial protein localization. *FEMS Microbiology Letters, 204*(1), 9-18.

Pjura, P. E., Grzeskowiak, K., & Dickerson, R. E. (1987). Binding of Hoechst 33258 to the minor groove of B-DNA. *Journal of Molecular Biology, 197*(2), 257-271.

Płociński, P., Brissett, N. C., Bianchi, J., Brzostek, A., Korycka-Machała, M., Dziembowski, A., Dziadek, J., & Doherty, A. J. (2017). DNA Ligase C and Prim-PolC participate in base excision repair in mycobacteria. *Nature Communications, 8*(1), 1251.

Porath, J., & Flodin, P. E. R. (1959). Gel filtration: a method for desalting and group separation. *Nature, 183*(4676), 1657-1659.

Preston, A., Maskell, D., Johnson, A., & Moxon, E. R. (1996). Altered lipopolysaccharide characteristic of the I69 phenotype in Haemophilus influenzae results from mutations in a novel gene, isn. *Journal of Bacteriology, 178*(2), 396-402.

Provvedi, R., & Dubnau, D. (1999). ComEA is a DNA receptor for transformation of competent Bacillus subtilis. *Molecular Microbiology, 31*(1), 271-80.

Ramsey, M. E., Hackett, K. T., Kotha, C., & Dillard, J. P. (2012). New complementation constructs for inducible and constitutive gene expression in Neisseria gonorrhoeae and Neisseria meningitidis. *Applied and Environmental Microbiology, 78*(9), 3068-78.

Redfield, R. J. (1993). Genes for breakfast: the have-your-cake-and-eat-it-too of bacterial transformation. *Journal of Heredity, 84*(5), 400-4.

Redfield, R. J. (2001). Do bacteria have sex? *Nature Reviews Genetics, 2*(8), 634-9.

Reece-Hoyes, J. S., & Walhout, A. J. M. (2018). Gateway recombinational cloning. *Cold Spring Harbor Protocols, 2018*(1), pdb.top094912-pdb.top094912.

Reuten, R., Nikodemus, D., Oliveira, M. B., Patel, T. R., Brachvogel, B., Breloy, I., Stetefeld, J., & Koch, M. (2016). Maltose-binding protein (MBP), a secretion-enhancing tag for mammalian protein expression systems. *PloS One, 11*(3), e0152386-e0152386.

Robson, H. G., & Salit, I. E. (1972). Susceptibility of Neisseria gonorrhoeae to seven antibiotics in vitro. *Canadian Medical Association Journal, 107*(10), 959-962.

Russell, D. W., & Zinder, N. D. (1987). Hemimethylation prevents DNA replication in E. coli. *Cell, 50*(7), 1071-1079.

Salgado-Pabón, W., Jain, S., Turner, N., van der Does, C., & Dillard, J. P. (2007). A novel relaxase homologue is involved in chromosomal DNA processing for type IV secretion in Neisseria gonorrhoeae. *Molecular Microbiology, 66*(4), 930-47.

Santini, C. L., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C., & Wu, L. F. (2001). Translocation of jellyfish green fluorescent protein via the Tat system of Escherichia coli and change of its periplasmic localization in response to osmotic up-shock. *Journal of Biological Chemistry, 276*(11), 8159-64.

Scott, S. (2011). Measurement of microbial cells by optical density. *Journal of Validation Technology, 17*(1), 46.

Seifert, H. S. (1997). Insertionally inactivated and inducible recA alleles for use in Neisseria. *Gene, 188*(2), 215-220.

Seitz, P., & Blokesch, M. (2013). Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiology Reviews, 37*(3), 336-363.

Seitz, P., Pezeshgi Modarres, H., Borgeaud, S., Bulushev, R. D., Steinbock, L. J., Radenovic, A., Dal Peraro, M., & Blokesch, M. (2014). ComEA is essential for the transfer of external DNA into the periplasm in naturally transformable Vibrio cholerae cells. *PLoS Genetics, 10*(1), e1004066-e1004066.

Shi, K., Bohl, T. E., Park, J., Zasada, A., Malik, S., Banerjee, S., Tran, V., Li, N., Yin, Z., Kurniawan, F., Orellana, K., & Aihara, H. (2018). T4 DNA ligase structure reveals a prototypical ATP-dependent ligase with a unique mode of sliding clamp interaction. *Nucleic Acids Research, 46*(19), 10474-10488.

Shinners, E. N., & Catlin, B. W. (1982). Arginine and pyrimidine biosynthetic defects in Neisseria gonorrhoeae strains isolated from patients. *Journal of Bacteriology, 151*(1), 295-302.

Shirano, Y., & Shibata, D. (1990). Low temperature cultivation of Escherichia coli carrying a rice lipoxygenase L-2 cDNA produces a soluble and active enzyme at a high level. *FEBS Letters, 271*(1-2), 128-30.

Shuman, S. (2004). NAD+ specificity of bacterial DNA ligase revealed. *Structure, 12*(8), 1335-1336.

Shuman, S. (2009). DNA ligases: progress and prospects. *The Journal of Biological Chemistry, 284*(26), 17365-17369.

Shuman, S., & Lima, C. D. (2004). The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases. *Current Opinion in Structural Biology, 14*(6), 757-64.

Smith, A. C., & Hussey, M. A. (2005). Gram stain protocols. *American Society for Microbiology, 1*, 14.

Sokolovski, M., Bhattacherjee, A., Kessler, N., Levy, Y., & Horovitz, A. (2015). Thermodynamic protein destabilization by GFP tagging: a case of interdomain allostery. *Biophysical Journal, 109*(6), 1157-62.

Sox, T. E., Mohammed, W., & Sparling, P. F. (1979). Transformation-derived Neisseria gonorrhoeae plasmids with altered structure and function. *Journal of Bacteriology, 138*(2), 510-8.

Sparling, P. F. (1966). Genetic transformation of Neisseria gonorrhoeae to streptomycin resistance. *Journal of Bacteriology, 92*(5), 1364-1371.

Spence, J. M., Chen, J. C., & Clark, V. L. (1997). A proposed role for the lutropin receptor in contact-inducible gonococcal invasion of Hec1B cells. *Infection and Immunity, 65*(9), 3736-3742.

Spencer-Smith, R., Roberts, S., Gurung, N., & Snyder, L. A. S. (2016). DNA uptake sequences in Neisseria gonorrhoeae as intrinsic transcriptional terminators and markers of horizontal gene transfer. *Microbial Genomics, 2*(8), e000069-e000069.

Spratt, B. G., Bowler, L. D., Zhang, Q. Y., Zhou, J., & Smith, J. M. (1992). Role of interspecies transfer of chromosomal genes in the evolution of penicillin resistance in pathogenic and commensal Neisseria species. *Journal of Molecular Evolution, 34*(2), 115-25.

Srikhanta, Y. N., Fox, K. L., & Jennings, M. P. (2010). The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. *Nature Reviews Microbiology, 8*(3), 196-206.

Steichen, C. T., Cho, C., Shao, J. Q., & Apicella, M. A. (2011). The *Neisseria gonorrhoeae* biofilm matrix contains DNA, and an endogenous nuclease controls its incorporation. *Infection and Immunity, 79*(4), 1504-1511.

Steichen, C. T., Shao, J. Q., Ketterer, M. R., & Apicella, M. A. (2008). Gonococcal cervicitis: a role for biofilm in pathogenesis. *The Journal of Infectious Diseases, 198*(12), 1856-1861.

Stohl, E. A., & Seifert, H. S. (2001). The recX gene potentiates homologous recombination in Neisseria gonorrhoeae. *Molecular Microbiology, 40*(6), 1301-10.

Subramanya, H. S., Doherty, A. J., Ashford, S. R., & Wigley, D. B. (1996). Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. *Cell, 85*(4), 607-15.

Sule, P., Wadhawan, T., Wolfe, A., & Pruess, B. (2008). Use of the BacTiter-Glo microbial cell viability assay to study bacterial attachment in biofilm formation. *Promega Notes, 99*, 19-21.

Sulkowski, E. (1985). Purification of proteins by IMAC. *Trends in Biotechnology, 3*(1), 1-7.

Takahashi, M., Yamaguchi, E., & Uchida, T. (1984). Thermophilic DNA ligase. Purification and properties of the enzyme from Thermus thermophilus HB8. *Journal of Biological Chemistry, 259*(16), 10041-7.

Tang, Y., Yang, X., Hang, B., Li, J., Huang, L., Huang, F., & Xu, Z. (2016). Efficient production of hydroxylated human-like collagen via the co-Expression of three key genes in Escherichia coli Origami (DE3). *Applied Biochemistry and Biotechnology, 178*(7), 1458-1470.

Tang, Z., Wang, K., Tan, W., Li, J., Liu, L., Guo, Q., Meng, X., Ma, C., & Huang, S. (2003). Real-time monitoring of nucleic acid ligation in homogenous solutions using molecular beacons. *Nucleic Acids Research, 31*(23), e148-e148.

Tapsall, J. W. (2005). Antibiotic resistance in Neisseria gonorrhoeae. *Clinical Infectious Diseases, 41*(Supplement_4), S263-S268.

Thomas, J. D., Daniel, R. A., Errington, J., & Robinson, C. (2001). Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli. *Molecular Microbiology, 39*(1), 47-53.

Toddo, S., Söderström, B., Palombo, I., von Heijne, G., Nørholm, M. H. H., & Daley, D. O. (2012). Application of split-green fluorescent protein for topology mapping membrane proteins in Escherichia coli. *Protein Science, 21*(10), 1571-1576.

Tomkinson, A., Vijayakumar, S., Pascal, J., & Ellenberger, T. (2006). DNA ligases: structure, reaction mechanism, and function. *Chemical Reviews, 106*(2), 687-699.

Unemo, M., & Shafer, W. M. (2011). Antibiotic resistance in Neisseria gonorrhoeae: origin, evolution, and lessons learned for the future. *Annals of the New York Academy of Sciences, 1230*, E19-E28.

Unemo, M., & Shafer, W. M. (2014). Antimicrobial resistance in Neisseria gonorrhoeae in the 21st century: past, evolution, and future. *Clinical Microbiology Reviews, 27*(3), 587-613.

van der Ende, A., Hopman, C. T., & Dankert, J. (2000). Multiple mechanisms of phase variation of PorA in Neisseria meningitidis. *Infection and Immunology, 68*(12), 6685-90.

Vasina, J. A., & Baneyx, F. (1997). Expression of aggregation-prone recombinant proteins at low temperatures: a comparative study of the Escherichia coli cspA and tac promoter systems. *Protein Expression and Purification, 9*(2), 211-8.

Veening, J. W., & Blokesch, M. (2017). Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nature Reviews Microbiology, 15*(10), 621-629.

Wang, L., Fan, D., Chen, W., & Terentjev, E. M. (2015). Bacterial growth, detachment and cell size control on polyethylene terephthalate surfaces. *Scientific Reports, 5*(1), 15159.

Wang, X., Lim, H. J., & Son, A. (2014). Characterization of denaturation and renaturation of DNA for DNA hybridization. *Environmental Health and Toxicology, 29*, e2014007-e2014007.

Wang, Y., Xie, J.-J., Han, Z., Liu, J.-H., & Liu, X.-P. (2013). Expression, purification and biochemical characterization of Methanocaldococcus jannaschii DNA ligase. *Protein Expression and Purification, 87*(2), 79-86.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research, 46*(W1), W296-W303.

Weiss, B., & Richardson, C. C. (1967). Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from Escherichia coli infected with T4 bacteriophage. *Proceedings of the National Academy of Sciences of the United States of America, 57*(4), 1021-1028.

Weller, G. R., Kysela, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day, J. P., Wilkinson, A., d'Adda di Fagagna, F., Devine, K. M., Bowater, R. P., Jeggo, P. A., Jackson, S. P., & Doherty, A. J. (2002). Identification of a DNA nonhomologous end-joining complex in bacteria. *Science, 297*(5587), 1686-9.

Wilkinson, A., Day, J., & Bowater, R. (2001). Bacterial DNA ligases. *Molecular Microbiology, 40*(6), 1241-1248.

Williamson, A., Grgic, M., & Leiros, H. S. (2018). DNA binding with a minimal scaffold: structure-function analysis of Lig E DNA ligases. *Nucleic Acids Research, 46*(16), 8616-8629.

Williamson, A., Hjerde, E., & Kahlke, T. (2016). Analysis of the distribution and evolution of the ATP-dependent DNA ligases of bacteria delineates a distinct phylogenetic group 'Lig E'. *Molecular Microbiology, 99*(2), 274-90.

Williamson, A., & Leiros, H.-K. S. (2020). Structural insight into DNA joining: from conserved mechanisms to diverse scaffolds. *Nucleic Acids Research, 48*(15), 8225-8242.

Williamson, A., & Leiros, H. S. (2019). Structural intermediates of a DNA-ligase complex illuminate the role of the catalytic metal ion and mechanism of phosphodiester bond formation. *Nucleic Acids Research, 47*(14), 7147-7162.

Williamson, A., & Pedersen, H. (2014). Recombinant expression and purification of an ATP-dependent DNA ligase from Aliivibrio salmonicida. *Protein Expression and Purification, 97*, 29-36.

Williamson, A., Rothweiler, U., & Leiros, H. K. (2014). Enzyme-adenylate structure of a bacterial ATP-dependent DNA ligase with a minimized DNA-binding surface. *Acta Crystallographica Section D: Biological Crystallography, 70*(Pt 11), 3043-56.

Wolfgang, M., van Putten, J. P., Hayes, S. F., Dorward, D., & Koomey, M. (2000). Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for bacterial pili. *The EMBO Journal, 19*(23), 6408-6418.

Yakhnin, A., Vinokurov, L., Surin, A., & Alakhov, Y. (1999). Green fluorescent protein purification by organic extraction. *Protein Expression and Purification, 14*, 382-6.

Zhang, F., Moniz, H. A., Walcott, B., Moremen, K. W., Wang, L., & Linhardt, R. J. (2014). Probing the impact of GFP tagging on Robo1-heparin interaction. *Glycoconjugate Journal, 31*(4), 299-307.

Zhu, H., & Shuman, S. (2007). Characterization of Agrobacterium tumefaciens DNA ligases C and D. *Nucleic Acids Research, 35*(11), 3631-3645.

Zimmerman, S. B., Little, J. W., Oshinsky, C. K., & Gellert, M. (1967). Enzymatic joining of DNA strands: a novel reaction of diphosphopyridine nucleotide. *Proceedings of the National Academy of Sciences of the United States of America, 57*(6), 1841-1848.

Zweig, M., Schork, S., Koerdt, A., Siewering, K., Sternberg, C., Thormann, K., Albers, S. V., Molin, S., & van der Does, C. (2014). Secreted single-stranded DNA is involved in the initial phase of biofilm formation by Neisseria gonorrhoeae. *Environmental Microbiology, 16*(4), 1040-52.