Investigation of hidden lipoproteins in *Neisseria meningitidis*

Cansu KARYAL

Kingston University, May 2016

Supervisor: Dr Ruth Griffin

Second Supervisor: Professor Mark Fielder

Acknowledgements

I would like to firstly express sincere gratitude to my supervisor Dr Ruth Griffin for her extraordinary supervision throughout this project. Her expertise within this field has provided the greatest support and guidance one can receive, and I thank her for this. I would also like to show my utmost appreciation to my lovely lab partner and friend, Ronni da Silva for sharing all his experience and for his assistance during this project. I would also like to thank the Faculty of Science, Engineering and Computing at Kingston University for funding this project.

I would like to thank my second supervisor Professor Mark Fielder, not only for his guidance but also for being such a great inspiration and encouraging me to peruse my dream. In addition to this, I would like to thank my amazing friends within the microbiology team (Ezra Rashid, Fredericka Mitchell, Lucky Cullen, Hayley Greenfield and Kelly Robertson) and our safety officer Dr Simon Gould for all their great support.

Finally, I would like to thank all my family members, who have showed great support especially my wonderful parents who have made this journey possible, my partner Mehmet Altay and my best friend Bahar Nurluoz.

Contents

Abstract

Neisseria meningitidis, is a Gram-negative diplococcus responsible for meningitis and septicaemia in adults and children. Serogroup B (MenB) strains account for most cases of invasive meningococcal disease (IMD) in Europe. According to the Meningitis Research Foundation (MRF), in the UK each year, approximately 1,800 MenB cases occur, of which about 10% lead to death. Following great challenges to develop an effective vaccine against MenB strains, two vaccines; Bexsero and Trumenba have been developed. These two vaccines vary in composition however, share a common component, factor H binding protein (fHbp), a well-studied lipoprotein that binds human factor H (hFH). Despite encouraging results, some concerns remain regarding their efficacy to target diverse strains.

The meningococcus expresses a number of virulence factors such as capsule, pili, lipooligosacharides (LOS), and lipoproteins. Bacterial lipoproteins play key roles in maintaining cell wall integrity, promoting adhesion, signal transduction and facilitating nutrient uptake. Importantly lipoproteins provoke host immune responses by the interaction of their lipids with Toll-like receptor 2 and many elicit potent antibody responses. Unsurprisingly, lipoproteins are emerging as promising vaccines. Bacterial lipoproteins are anchored to the inner or outer membrane and are characterised by an N-terminal signal peptide comprising 3 domains ending in a lipobox. The signal peptide signals for transportation from the cytoplasm across the inner membrane to the periplasm where the protein is lipidated and the signal peptide is cleaved. The protein is further acylated before being sorted to the outer membrane.

Based on the physical features of a signal peptide, a predictive algorithm developed was incorporated in the website DOLOP to analyse protein sequences and 69 probable lipoproteins were identified in meningococcal strain, MC58.

Based on the knowledge that the well characterised meningococcal lipoprotein, fHbp was not recognised as a lipoprotein by this tool due to the signal peptide not being located at the N-terminus and rather 40 amino acids downstream of the annotated translation start residue, we hypothesised that other lipoproteins were being missed by annotation of the incorrect start codon. Systematic analysis of each protein sequence from strain MC58 revealed 10 proteins with the signal peptide at the N-terminus. A further 13 putative lipoproteins were found with the signal peptide located up to 30% downstream of the annotated start codon.

Intriguingly, 3 more protein sequences contained the signal peptide in the middle and 15 harboured the signal peptide towards the C-terminus.
In this study, we tested whether the signal peptide could be recognised by the cell if, like fHbp, it is located downstream of the N-terminus and if it is translocated and processed as a lipoprotein, or whether signal peptides must be located at the N-terminus to function.

1. Introduction

1.1. The meningococcus

N. meningitidis is a Gram-negative diplococcus and is the leading cause of bacterial meningitis and septicaemia (Brehony *et al*., 2016). Despite improved prophylactics and therapeutics, meningococcal disease remains a public health threat. Endemics are caused worldwide whereby young children and adolescents are predominately at risk of contracting invasive meningococcal disease (IMD) (Pace and Pollard, 2012). Approximately, 10-30% of the population can become infected, in which the mortality rate can range from 10-15% in developed countries, to more than 20% in developing countries (Stephens and Apicella, 2015). The highest meningococcal cases are seen in the 'African meningitis belt' with incidence of 1,000 cases per 100,000 population during major epidemics (CDC, 2016; Crum-Cianflone and Sullivan, 2016).

The bacterium is human-specific and is carried in the nasopharynx of about 10% of healthy individuals (Public Health England, 2017; Yazdankhah and Caugant, 2004). Bacteria can be transmitted from person to person by droplet infection. Occasionally the bacteria over-power the body immune defences by invading the mucosal epithelium and entering the blood stream, leading to septicaemia. Subsequently, the bacterium may travel to the blood-brain barrier and infect the meninges (Davide *et al*., 2012). Typical symptoms include rash, stiff neck, severe headache and vomiting. As a result of the infection, long-term defects can arise such as neurological disabilities, cognitive impairment and loss of vision or hearing (Pace and Pollard, 2011).

There are 13 serogroups, classified according to their distinct capsular polysaccharide structures (Roupheal and Stephens, 2015; Davide *et al*., 2012). Six serogroups; A, B, C, W-135, X and Y account for most cases of invasive meningococcal disease (Roupheal and Stephens, 2015). Conjugated polysaccharide vaccines have been used to protect against strains of serogroups A, C, Y and W-135, in the form of monovalent (MenA and MenC) and quadrivalent (MenACWY) vaccines (Crum-Cianflone and Sullivan, 2016). However, producing a vaccine against serogroup B has been extremely challenging. According to the European Centre for Disease Prevention and Control, in 2012, serogroup B was responsible for 68% of 3,463 confirmed cases in Europe. A study carried out in

England and Wales from 2011 to 2013 revealed that 78.4% of 2,547 confirmed cases were accounted for by group B strains (Clark *et al*., 2016). Each year approximately 1,800 of MenB cases occur in the UK of which about 10% result in death (MRF, 2017). The presence of polysialic acid found within the serogroup B capsule is highly similar to structures found on neural cells, resulting in this vaccine candidate being poorly immunogenic. Other factors which have hindered the production of a vaccine have been the antigenic variability of serogroup B antigens (Davide *et al*., 2012; Gandhi *et al*., 2016). However, surface exposed lipoprotein, factor H binding protein (fHbp), expressed by strains of all serogroups has emerged as a promising vaccine antigen that can protect from serogroup B disease. FHbp, binds to human factor H (hFH), interfering with the alternative complement pathway activation and therefore, preventing complement-mediated killing. Importantly, fHbp elicits a protective antibody response, as shown by serum bactericidal antibody (SBA) assays (the correlate for protection) (Holst *et al*., 2003). Koeberling *et al*. (2011) showed that a critical level of expression of fHbp was required to elicit a protective SBA response. Given that fHbp is expressed in more than 97% of serogroup B strains (Gandhi *et al*., 2016), this led to the development of two fHbp based vaccines; MenB-fHbp (Trumenba) and 4CMenB (Bexsero). Bexsero was licenced in Europe in 2013 (European Medicine Agency, 2015) and Trumenba was licensed in the US in 2014 for ages 10-25 (Pfizer, 2014). The two vaccines vary in composition. Trumenba is comprised of two fHbp variants, one from subfamily A and another from subfamily B, whereas Bexsero is composed of four antigens; GNA2091 fused to subfamily B fHbp, Neisserial heparin-binding antigen (NHBA) fused to GNA1030, *N. meningitidis* adhesion A (NadA) and outer membrane vesicles (OMV) predominantly containing PorA (major porin) (Gandhi *et al*., 2016). Despite encouraging results, some concerns remain regarding their safety and their ability to target diverse strains. For Tumenba in particular, this limitation can be explained by the fact that not all strains synthesise or export fHbp (McNeil *et al*., 2013). Moreover, the level of expression of fHbp varies between strains (Biagini *et al*., 2016) and within strains due to regulation of its expression by external factors including oxygen, iron availability and temperature (Oriente *et al.,* 2010; Sanders *et al.,* 2012; Loh *et al*., 2016).

1.2. Virulence factors

There are many factors which contribute to the virulence of *N. meningitidis.* The polysaccharide capsule, outer membrane proteins (OMP) including pili, adhesion molecules (Opa and Opc) and porins (PorA and PorB), lipooligosaccharides (LOS) and lipoproteins (Nakayama *et al*., 2012; Rouphael and Stephens, 2015).

1.2.1. Capsule

The major virulence factor in *N. meningitidis* is the polysaccharide capsule (Davide *et al*., 2012). In all meningococcal serogroups, the capsule plays a crucial role in the survival of the bacterium within the blood, by providing protection against host defence mechanisms. The capsule is abundant in sialic acid, a monosaccharide composed of α-2, 8-linked N-acetylneuramic acid (NeuNAc) which interferes with the alternative pathway activation. As a result, antibody-mediated killing and phagocytosis is inhibited (Jarvis and Vedos, 1986; Frosch *et al*., 1988). One other way in which the pathogen avoids detection by the immune system is by capsule switching. This occurs as a result of genetic exchange of genes involved in the biosynthesis of the capsule by transformation and allelic exchange (Swartley *et al*., 1997; Rouphael and Stephens, 2015).

1.2.2. OMP (pili, adhesion molecules, porins)

Initial attachment of the bacterium to the host is vital in colonisation of mucosal membranes of the nasopharynx and is achieved mainly by, Type IV pili (Stephens and McGee, 1981). These flexible filaments project from the outer membrane and aid in invasion by adhering to epithelial cells. By generating a twitching motion, they allow the bacterium to move across epithelial surfaces (Carbonelle *et al.,* 2009; Rouphael and Stephens, 2015). Once colonised, adhesins such as the opacity proteins, Opa and Opc, found on the meningococcal cell surface interact with CD66/CEACAM receptors on the host cell, initiating intimate binding (Virji *et al*., 1993). Colonisation is also facilitated by the minor adhesins; NadA, *Neisseria* hia homologue A (NhhA), adhesion penetration protein (App) and meningococcal serine protease A (MspA). NadA is more abundant in virulent strains than carrier isolates and are hence likely associated with virulence (Capacchi *et al*., 2005). This is also seen with NhhA and App, however, they are less effective at binding to epithelial cells compared to other minor adhesins (Hill *et al*., 2010). Furthermore,

molecules such as iron-binding proteins; transferrins and lactoferrins play an important role in iron sequestration. Iron is an essential component required for the growth of the bacterium. In addition, porins, PorA and PorB, which allow the diffusion of nutrients into the bacterial cell are involved in host cell interactions and generate potent antibody responses (Rouphael and Stephens, 2015).

1.2.3. Lipooligosaccharide (LOS)

Lipooligosaccharide (LOS), also referred to as endotoxin, is a key virulence determinant and provokes an innate immune response. LOS detaches from the bacterial cell wall, circulates within the host blood stream and binds to Toll-like receptors (TLR) found on immune cells such as dendritic cells, macrophages and mast cells. LOS specifically binds TLR4 which leads to cytokine release and can culminate in septic shock (Hou *et al*., 2008; Tang *et al*., 2015).

1.2.4. Lipoproteins

More and more research is highlighting the importance of lipoproteins in IMD. In Gram-negative bacteria, lipoproteins play a major role, not only in virulence such as adhesion, colonisation and activation of the immune system, but in cell wall integrity, cell division, nutrient uptake, signal transduction and antibiotic resistance (Nakayama *et al*., 2012; Zückert, 2014; Chahales and Thanassi, 2015). TLR2 on the surface of monocytes, macrophages, dendritic cells, mast cells and B cells recognise lipoproteins (Kang *et al*., 2009) and activates a cascade of inflammatory cytokines which help to eliminate the pathogen by promoting phagocytosis (Hou *et al*., 2008)

Bacterial lipoproteins are lipidated at their N-terminus and anchor to the outer membrane by their fatty acid moieties (Babu *et al*., 2006). All lipoproteins are synthesized in the cytoplasm (Figure 1.2) as a preprolipoprotein precursor containing a signal peptide at the N-terminus, approximately 20-40 amino acids in length. The signal peptide has three distinct regions (Figure 1.1), consisting of a positively charged n-region, a hydrophobic h-region and a c-region (Heijne, 1989; Babu *et al*., 2006). The amino acid residues in the n-region range from five to seven and contain no less than two positively charged amino acids. The centre h-region, contains between 7 to 22 amino acids, which are mainly hydrophobic (Babu *et al*., 2006). At the C-terminus (c-region), is a characteristic, 4 amino acid motif called the lipobox with a consensus sequence (LVI)(ASTVI)(GAS)C ending in a conserved cysteine (Heijne, 1989; Zückert, 2014). The preprolipoprotein is translocated across the cytoplasmic or inner membrane and into the periplasm by the Sec or Tat pathway (Figure 1.2). A three step, enzymatic process follows for post-translational modification of the preprolipoprotein (Zückert, 2014). The initial step involves the addition of diacylglycerol to the conserved cysteine residue of the lipobox, catalysed by the enzyme preprolipoprotein diacylglyceryl transferase (Lgt) (Sankaran and Wu, 1994). Following this, the signal peptide is cleaved at the amino acid residue immediately upstream of the cysteine residue by prolipoprotein signal peptidase (Lsp or signal peptidase II) generating a diacylated apolipoprotein (Tokunaga *et al*., 1982; Inouye *et al*., 1983). These fatty acids allow the apolipoprotein to remain attached to the inner membrane after signal peptide cleavage (Chahales and Thanassi, 2015). Finally, a third fatty acid is added to the amide group of the cysteine by apolipoprotein N-acyltransferase, Lnt. The resultant mature lipoprotein is then ready to be sorted to the outer membrane.

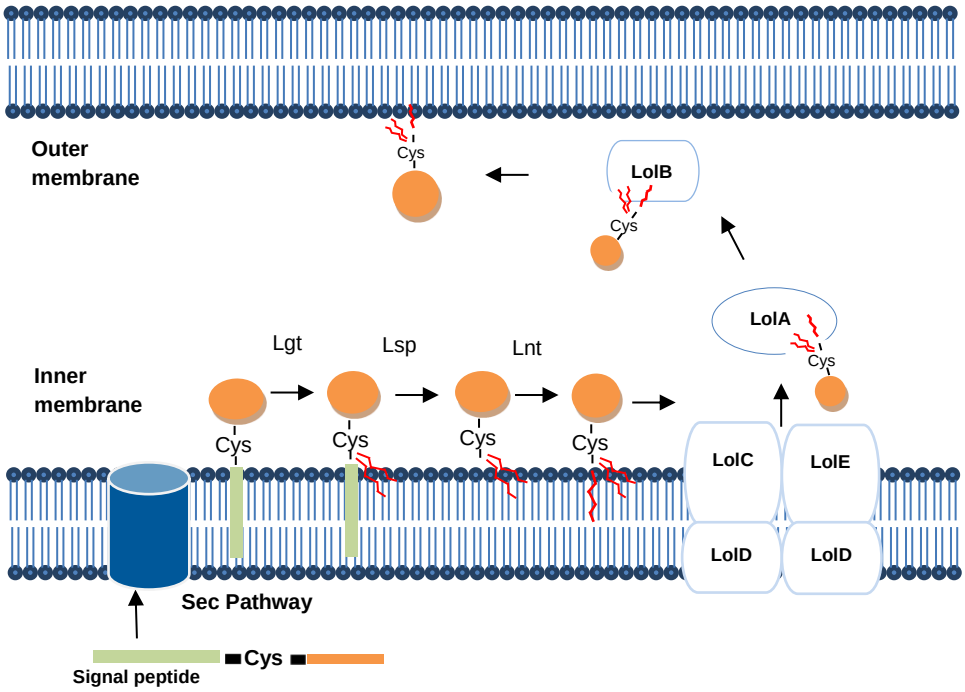


Figure 1.2 Lipoprotein sorting pathway. The above figure demonstrates lipid modification of lipoproteins exported from the cytoplasm to the outer membrane.

Once triacylated, the mature lipoprotein is usually sorted to the outer membrane by the localisation of lipoproteins (Lol) apparatus (Hooda *et al*., 2016). In *Escherichia coli* (*E.coli),* the +2 rule states that the presence of an Aspartate (D) residue at position +2, after the conserved cysteine causes lipoprotein retention within the inner membrane (Yamaguchi *et al*., 1988). Hence, any other amino acid residue at the +2 position signals for the export of the lipoprotein to the outer membrane. This rule may apply to *N. meningitidis*. In *Pseudomonas aeruginosa*, specific residues present at position +3 and +4 have been shown to signal for outer membrane export (Narita and Tokuda, 2006). The Lol system operates using a LolCDE protein complex, an ABC (ATP-binding cassette) transporter, which releases lipoproteins from the inner membrane to the periplasmic chaperone LolA (Matsuyama *et al*., 1995). This action is fuelled by ATP hydrolysis catalysed by ATPase LolD, in which LolA carries the lipoprotein substrate from LolCDE complex, delivering it to the outer membrane receptor LolB-also an outer membrane lipoprotein (Matsuyama *et al.,* 1995; Tanaka *et al*., 2001). Until recently, the final step involved in the export of lipoproteins to the cell surface was unknown. In a study conducted by Hooda *et al.* (2016), it was revealed that *N. meningitidis* utilises a surface lipoprotein assembly modulator (Slam) which flips OMP's such as fHbp to the cell surface.

Many lipoproteins expressed at the bacterial cell surface have been shown to elicit potent and protective antibody responses. It is therefore not surprising that lipoproteins are emerging as promising vaccines. To this end the lipoproteome of *N. meningitidis* was investigated.

Using the lipoprotein prediction tool in the website DOLOP (http://www.mrc-lmb.cam.ac.uk/genomes/dolop/), it was previously reported by Babu *et al*. (2006), that there are 69 lipoproteins in *N. meningitidis* strain MC58. Given the stringency of this feature, this is a conservative estimate. Recently, da Silva *et al*. (2016) showed that the signal peptide of fHbp was 26 amino acids long and positioned 40 amino acids downstream of the methionine that is annotated as the translation start residue (NCBI). Due to the signal peptide not positioned at the N-terminus, fHbp has missed being annotated as a lipoprotein. This led us to suspect that there are other lipoproteins in the genome, annotated with the wrong methionine as the translation start residue that have failed to be recognised as lipoproteins.

In this study, a systematic analysis of each of the 2,119 protein sequences of *N. meningitidis* MC58 in NCBI (https://www.ncbi.nlm.nih.gov/) by the above predictive algorithm led to the identification of 10 more putative lipoproteins containing an N-terminal signal peptide (Table 3.1). A second group of 13 proteins were identified with a signal peptide within the first 30% of the protein sequence (Table 3.2), a third group of 3 proteins contained a signal peptide in the middle of the gene (Table 3.3) and a fourth group of 15 proteins revealed a signal peptide positioned towards the C-terminus (Table 3.4). We questioned whether signal peptides can be recognised if not positioned at the N-terminus. Five of these putative lipoproteins were selected for investigation in this study along with fHbp and NMB1468 as positive controls. The latter was previously experimentally verified as a lipoprotein by Hsu *et al*. (2008).

Using the approach taken by da Silva *et al.* (2016), truncated versions of these lipoproteins were cloned into the *Neisseria* complementation vector and used to transform strains MC58 and MC58Lnt. MC58Lnt contains a transposon in the *lnt* gene and is therefore incapable of triacylating lipoproteins. Resolving these truncated, c-Myc tagged lipoproteins by SDS-PAGE, enables differentiation of the

lost single fatty acid in MC58Lnt and this size difference can be detected by Western blotting with an anti-c-Myc antibody. Therefore, the protein in question is inferred to be a lipoprotein if the migration differs in MC58Lnt compared to MC58.

2. Methods

2.1. Bioinformatics

2.1.1. Identifying putative meningococcal lipoproteins

Using the NCBI database (https://www.ncbi.nlm.nih.gov/), each of the 2,119 protein sequences of *N. meningitidis* MC58 were analysed by the lipoprotein predictive algorithm tool in the DOLOP website (http://www.mrc-lmb.cam.ac.uk/genomes/dolop/) for the presence of a signal peptide.

2.1.2. Exploring hidden putative meningococcal lipoproteins

In order to identify putative lipoproteins with an internal signal peptide all combinations of the lipobox consensus (LVI)(ASTVI)(GAS)C (Table 2.1) were used as the query sequence in a BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) analysis. The full protein sequence was investigated for the presence of the signal peptide adjacent to and upstream of the lipobox. This was then verified by the lipoprotein predictive algorithm function in the DOLOP website (http://www.mrc-lmb.cam.ac.uk/genomes/dolop/). An example is shown for protein NMB0727 (Figure 2.2).

Table 2.1 All possible combinations of the lipobox against the protein sequence of MC58 used as query in a sequence BLASTp analysis.

Lipobox query sequence		
L	**V**	**I**
LAAC	VAAC	IAAC
LAGC	VAGC	IAGC
LASC	VASC	IASC
LSAC	VSAC	ISAC
LSGC	VSGC	ISGC
LSSC	VSSC	ISSC
LTAC	VTAC	ITAC
LTGC	VTGC	ITGC
LTSC	VTSC	ITSC
LVAC	VVAC	IVAC
LVGC	VVGC	IVGC
LVSC	VVSC	IVSC
LIAC	VIAC	IIAC
LIGC	VIGC	IIGC
LISC	VISC	IISC

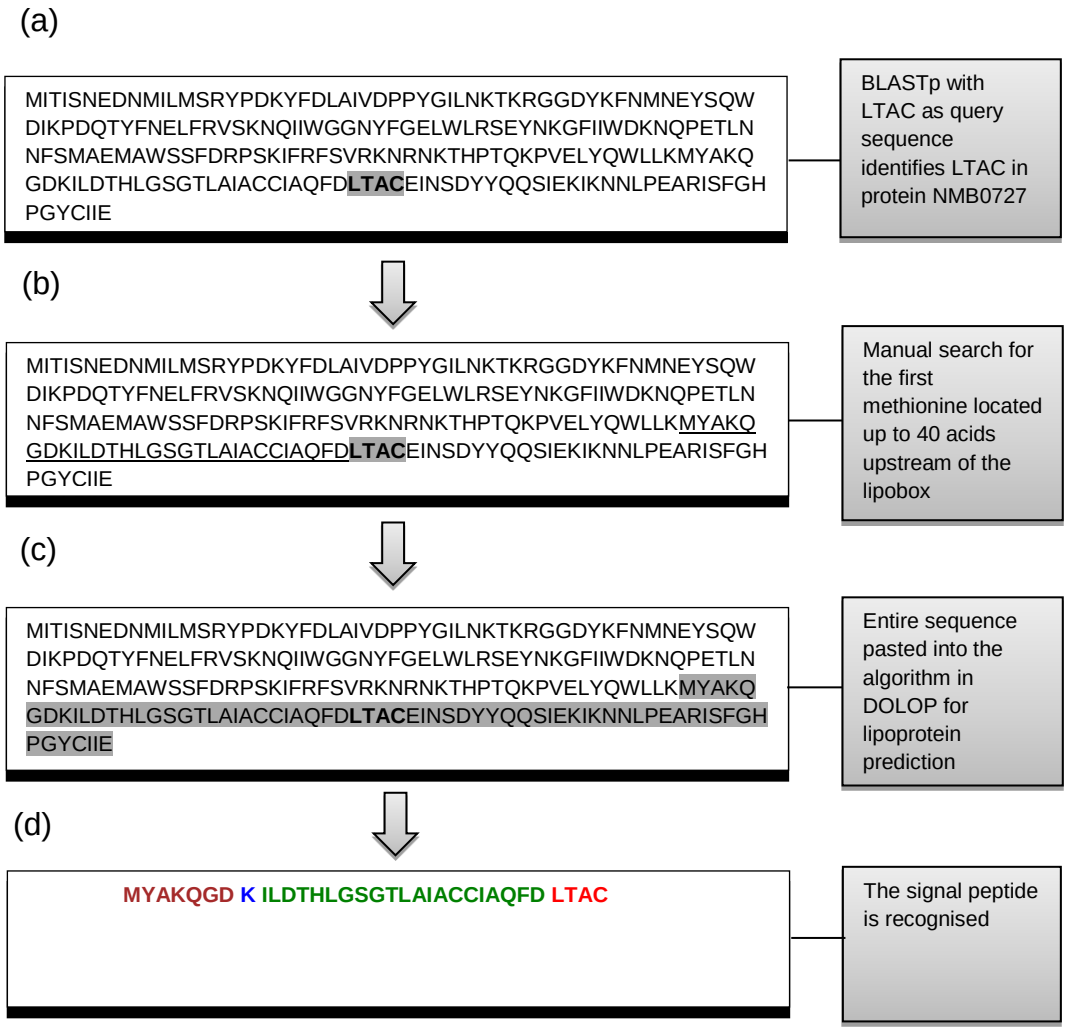


Figure 2.2 Identification of a signal peptide in NMB0727.

2.2. Bacterial strains and growth conditions

N. meningitidis MC58, strain serogroup B: 15:P1.7, 16, ST-74; ET-5, was obtained from LGC Standards and MC58Lnt was kindly provided by Dr Ruth Griffin and Ronni da Silva. All meningococcal strains were grown overnight at 37°C in 5% CO_2 on GC agar plates containing Kellogg's supplements I and II (Kellogg *et al*., 1963). Piliated MC58 and MC58Lnt cells were cultured for transformation by overnight growth at 30°C in 5% CO_2.Transformants were selected on GC agar containing 0.3 µg/ml of erythromycin.

Subcloning Efficiency™ DH5α™ Competent cells (Invitrogen™) were used for cloning and grown at 37°C on Luria-Bertani (LB) agar or in LB broth with shaking at 200 rpm and with 30 µg/ml of kanamycin and 100 µg/ml of ampicillin where appropriate.

Meningococcal cells used for whole cell (WC) lysate preparations were grown overnight on GC agar plates containing 0.5 mM Isopropyl β-1-thiogalactopyranoside (IPTG) at 37°C in 5 % CO_2. The component IPTG is used to enhance the expression of recombinant protein by stimulating transcription of the *lac* operon (Biologicscrop, 2016).

GC broth cultures were obtained by harvesting bacterial cells from overnight grown GC agar plates containing 0.5 mM IPTG and re-suspended in Kellogg's I and II supplemented GC broth. 1 ml of bacterial suspension at optical density (OD) A_{600}1.0 was used to inoculate 20 ml supplemented GC broth in a 250 ml sterilised conical flask. Cells were grown until A_{600} 0.1 was reached at 84 rpm shaking at 37°C in 5 % CO_2.

2.2.1. Glycerol stock of bacterial strains

Bacterial samples were frozen at – 80°C by mixing 500 µl of bacterial broth culture or bacterial suspension in Kellogg's I and II supplemented GC broth with 500 µl 30% (v/v) sterile glycerol.

2.3. Molecular methods for DNA manipulations

2.3.1. Genomic DNA extraction

DNA extraction of meningococcal strains was carried out from a fresh overnight grown plate using the Gentra Puregene Yeast/ Bact. Kit (Qiagen) according to the manufacturer's protocol. Concentration and purity was measured using the NanoVue ™ Plus Spectrophotometer (GE Healthcare Lifesciences).

2.3.2. Plasmid extraction

Plasmid DNA extraction was performed using QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions followed by concentration and purity measurements using NanoVue™ Plus Spectrophotometer. Plasmid DNA was stored at -20 ºC.

2.3.3. Gene clean

Gene clean was performed using PCR Mini elute kit (Qiagen) according to the manufacturer's instructions. All PCR products and restriction digests were purified by gene clean.

2.3.4. PCR reactions and primers

PCR reactions were performed using *Taq* DNA polymerase (Qiagen). The components and reaction conditions listed in Table 2.3 and 2.4 were performed according to manufacturer's protocol. All primers were purchased from Sigma aldrich and are listed in Table 2.6. Annealing positions of each primer are shown in Table 2.5.

Table 2.3 Components of PCR reaction using *Taq* DNA polymerase**.**

Reagents	**Volume (µl)**	
	DNA	**No DNA control**
Taq DNA polymerase	0.5	0.5
QIAGEN® PCR Buffer 10x	10	10
dNTP Mix, 10 mM	2	2
Forward primer 10 µM	5	5
Reverse primer 10 µM	5	5
RNase free water	76.5	77.5
Template DNA	1.5	-
Total Volume	**100**	**100**

Table 2.4 The typical conditions used for PCR.

Step	**Temperature (°C)**	**Time (mins)**	**No. of Cycles**
Initial Denaturation	94	3	1
Denaturing	94	0.5	35
Annealing	-	0.5	
Extension	72	1	
Final Extension	72	10	1

Table 2.5 Annealing position of the forward and reverse primers used to generate tagged proteins. The signal peptide is highlighted in redfor each gene.

Gene	Primer design using genomic sequence
fHbp	TTAATTAAATGCCGTCTGAACCGCCGTTCGGACGACATTTGATTTTTGCTTCTTTGACCTGCCTCAT TGATGCGGTATGCAAAAAAAGATACCATAACCAAAATGTTTATATATTATCTATTCTGCGTATGACTA GGAGTAAACCTGTGAATCGAACTGCCTTCTGCTGCCTTTCTCTGACCACTGCCCTGATTCTGACCG CCTGCAGCAGCGGAGGGGGTGGTGTCGCCGCCGACATCGGTGCGGGGCTTGCCGATGCACTAA CCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGATCAGTCCGTCAGGAAA AACGAGAAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGAAACGGTGACAGCCTCAA TACGGGCAAATTGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGACG GGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACGAACAGAAACTGATTAGCGAA GAAGACCTGTAGGTTTAAAC
NMB1468	TTAATTAAATGCAAAAAATTATTGATTGCCGCAATGATGGCGGCTGCCTTGGCAGCTTGTTCGCAAGA AGCCAAACAGGAGGTTAAGGAAGCGGTTCAAGCCGTTGAGTCCGATGTTAAAGACACTGCGGCTT CTGCCGCCGAGTCTGCCGCTTCTGCCGTCGAAGAAGCGAAAGACCAAGTCAAAGATGCTGCGGCT GATGCAAAGGCAAGTGCCGAGGAAGCTGTAACTGAAGCCAAAGAAGCTGTAACTGAAGCAGCTAA AGATACTTTGAACAAAGCTGCCGACGCGACTCAGGAAGCGGCAGACAAAATGAAAGATGCCGCCA AAGAACAGAAACTGATTAGCGAAGAAGACCTGTAGGTTTAAAC
NMB0727	TTAATTAAATGATAACTATTTCAAATGAAGATAACATGATCTTAATGTCTCGGTATCCTGACAAGTAT TTTGATTTGGCAATTGTAGATCCTCCTTATGGGATTTTGAATAAAACTAAACGTGGTGGTGATTATAA ATTCAATATGAATGAATACTCACAATGGGATATTAAGCCAGACCAAACTTACTTTAATGAATTATTTC GCGTGTCAAAAAATCAAATTATTTGGGGTGGGAATTATTTTGGCGAGTTATGGTTGAGAAGTGAATA TAATAAAGGATTTATTATTTGGGATAAGAATCAACCAGAGACATTAAATAATTTTTCTATGGCGGAAA TGGCTTGGTCGTCATTCGATAGGCCATCTAAAATTTTCCGGTTTAGTGTGCGGAAAAATCGTAATAA AACTCACCCAACACAAAAACCAGTCGAATTATATCAGTGGTTGTTAAAAATGTATGCAAAGCAGGGT GATAAGATTTTAGATACACATTTAGGAAGTGGAACTCTTGCTATTGCATGCTGCATTGCACAGTTTG ATTTGACAGCTTGTGAAATCAATTCCGATTATTACCAACAATCGATTGAGAAAATAAAAAATAATTTA CCTGAAGCTAGAATCAGTTTTGGGCATCCAGGTTATTGTATTATTGAAGAACAGAAACTGATTAGCG AAGAAGACCTGTAGGTTTAAAC
NMB0949	TTAATTAAATGGTAGAACGTAAATTGACCGGTGCCCATTACGGTTTGCGCGATTGGGTGATGCAAC GTGCGACTGCGGTTATTATGTTGATTTATACCGTTGCACTTTTAGTGGTTCTATTTTCCCTGCCTAA AGAATATTCGGCATGGCAGGCATTTTTTAGTCAAACTTGGGTAAAAGTATTTACCCAAGTGAGCTTC ATCGCCGTATTCTTGCACGCTTGGGTGGGTATCCGCGATTTGTGGATGGACTATATCAAACCCTTC GGCGTGCGTTTGTTTTTGCAGGTTGCCACCATCGTTTGGCTGGTCGGCTGTCTCGTGTATTCAGTT AAAGTGATTTGGGGGGAACAGAAACTGATTAGCGAAGAAGACCTGTAGGTTTAAAC
NMB1447	TTAATTAAATGATGAAACTCAATCCCCAACAGCTCGAAGCCGTCCGCTACCTCGGCGGCCCACTGC TCGTCCTTGCCGGTGCAGGCAGCGGCAAAACCGGCGTGATTACTCAAAAAATTAAGCATTTGATTG TCAATGTCGGCTACCTGCCGCATACCGTTGCCGCAATTACCTTTACCAACAAAGCCGCTGCGGAAA TGCAGGAGCGCGTTGCCAAAATGCTGCCCAAACCGCAAACGCGCGGGCTGACGATTTGCACGTTC CACTCTTTGGGCATGAAGATTCTGCGCGAAGAGGCGAACCATATTGGTTACAAAAAAAACTTCTCC ATTCTCGATTCTACCGACAGCGCGAAAATCATCGGCGAACTCTTAGGCGGTACGGGCAAAGAAGC CGTATTCAAGGCGCAGCACCAGATTTCCTTGTGGAAAAACGATTTAAAAACGCCTGAAGATGTCGT TCAGACGGCATCGAACATTTGGGAACAACAAACCGCACGCGTGTATGCGAGCTATCAGGAAACCT TACAAAGCTATCAGGCAGTGGACTTCGACGACTTAATCCGCCTGCCTGCCGTGCTGTTGCAGCAA AACAGCGAAGTGCGCAACAAATGGCAGCGGCGGCTGCGTTATCTGTTGGTTGACGAATGCCAAGA TACGAATACCTGCCAATTTACGTTGATGAAGCTGCTGACCGGCGCGGAAGGTATGTTTACCGCCGT CGGCGACGACGACCAGTCCATCTACGCATGGCGCGGTGCGAACATGGAAAACCTGCGTAAAATG CAGGAAAACTATCCGCAGATGAAGGTCATCAAACTGGAGCAAAACTACCGCTCCACCGCGCGGAT TCTCAAAATCGCCAACAAAGTCATCGAAAACAACCCCAAGCTGTTTACCAAAAAACTTTGGTCGCAA TTGGGCGAAGGCGAGCCGGTCAAAGTCGTTGCCTGCCAAAACGAGCAACACGAAGCCGACTGGG TCGTCAGCCAAATCGTCAAACAAAAACTCATCGGCGGCGACAAAACCCAATATGCCGATTTCGCCG TGTTATACCGGGGAAAGCATCAGGCGAGGATTTTCGAGGAAGCATTGCGCGGCGCGCGCATCCC CTACCAGCTCTCCGGCGGACAAAGCTTTTTCGACAAAGCCGAAATCAAAGACGTGTTGTCTTATGT GCGGCTGCTTGCCAACCCCAACGACGATCCCGCCTTTCTGCGTGCCGTTACCACGCCCAAACGCG GCATCGGCGATGTCACGCTGGGCAAGCTCAACACTTACGCGCACGAACACGAATGCAGCCTGTAT GAAGCCGCGCAAAACGAAGAAGCCCTTGCCACGCTGAACAATACCAACCGCCAACACCTGCAAAC CTTTATGGATATGTTCGTCAGCTACCTCGCCAAAGCCGAAACCAGCGAAGCGGGCGAGTTCATCA ACAGCCTGCTCGAAGAAATCGACTATGAAAACCATTTGATGCAAAACGAAGAAGGCAAAGCCGGC GAAATCAAATGGCGCAACGTCGGCGATTTGGTATCATGGTTTGCGCGAAAAGGCGGGGAAGACGG CAAAAACATCATCGAACTCGCCCAAACCGTCGCCTTGATGACGCTTTTGGAAGGAAAAGACGAAGA AGAAACCGATGCCGTCTCGCTATCCACGCTACACGCCGCCAAAGGTTTGGAGTATCCGTATGTTTT CCTTGTCGGTTGCGAAGAAGGCGTTTTGCCGCACAACGACAGTATCGAAGAGGGCAACGTCGAAG

	AAGAACGCCGCCTGATGTACGTCGGCATCACCCGCGCCAAACGCCAACTCACACTGACCCACTGC GTCAAACGCAAAAAACAAGGCACATGGCAGTTCCCCGAACCCAGCCGATTCATAGACGAAATGCC GCAGGAAGATTTGAAAATCCTGGGGCGCAAAGGCGGCGAACCGATTGTCAGCAAAGAAGAAGGC AGACGCAACCTTGCCGATATAATCGGAAGGCTCGACAACCTAAAAAAAAAGCGGCGCGGCGGATGA ACAGAAACTGATTAGCGAAGAAGACCTGTAGGTTTAAAC
NMB1564	TTAATTAAATGCAGGTTACATCAAAATGGATAGACGGGATGTGTTTTGTCGGCACGACGGAAGGCG GGCACAGCGTCGTTATGGAGGGGTCGGCGGCAGAAGGTAAGGCTAAGCGCGGGCCCAGCCCTTT GGAAATGCTGCTGTTGGGCGTGGCGGGCTGTTCGAGCATCGATGTGGTGATGATTGCCGAAAAAC AGCGTCAGAAAGTGACTGACTGCCGTGCGACGGTTACGGCGAAACGGGCGGACGATGCGCCGCG CGTGTTTACCGAAATCCACATCCATTTCAAAGTATTCGGGCATGATTTGAAAGAATCGGCCATTGAG CGCGCCGTTCAGATGTCTGCCGAAAAATACTGTTCGGCTTCGATTATGTTGGGCAAAGCGGCAAA GATTACCCACAGTTTTGAAATTGCCGGGGCAGATAAAGAACAGAAACTGATTAGCGAAGAAGACCT GTAGGTTTAAAC
NMB1566	TTAATTAAATCAAAAACATCGTCATCCTGATTTCTGGACGCGGCAGCAATATGCAGGCAATCGTCAA TGCCGCCATTCACAACGTCCGCATTGCCGCCGTGTTGAGCAACAGCGAAACGGCTGCCGGTTTGC AATGGGCGGCCGAACGCGGCATCCCGACCGATAGCCTGAATCATAAAAACTTTACATCCCGGCTT GCCTTCGATACCGCCATGATGGAGAAAATCGACGCATATCAACCCGACTTGGTGGTTTTGGCAGG TTTTATGCGGATTCTGACCCCCGAGTTTTGCGCCCGTTACGAAGGCAGGCTGATGAACATTCACCC GTCCATCCTTCCCTCGTTTACCGGACTTCATACGCACGAACGCGCTTTGGAGGCGGGCTGCCGCG TTGCCGGCTGGCACCATCCATTTCGTTACTGCCGAACTGGATTGCGGCCCGATTGTATCGCAAGGG GTTGTGCCGATACTCGACGGCGATACGGCAGACGATATTGCCGCACGGGTTTTGGCTGTCGAGCA TAAACTTTATCCGAAAGCCGTTGCCGATTTTGCCGCCGGCCGCCTGATTATTGAGGGAAACCGCGT CAGAAATTCGGAAAACGCCGATGCCGCCCGTTTTCTGACGGCGGAACAGAAACTGATTAGCGAAG AAGACCTGTAGGTTTAAAC
His-NMB0949	AACAGGATCCATGCACCACCACCACCACCACGTAGAACGTAAATTGACCGGTGCCCATTACGGTTT GCGCGATTGGGTGATGCAACGTGCGACTGCGGTTATTATGTTGATTTATACCGTTGCACTTTTAGT GGTTCTATTTTCCCTGCCTAAAGAATATTCGGCATGGCAGGCATTTTTTAGTCAAACTTGGGTAAAA GTATTTACCCAAGTGAGCTTCATCGCCGTATTCTTGCACGCTTGGGTGGGTATCCGCGATTTGTGG ATGGACTATATCAAACCCTTCGGCGTGCGTTTGTTTTTGCAGGTTGCCACCATCGTTTGGCTGGTC GGCTGTCTCGTGTATTCAGTTAAAGTGATTTGGGGGGAATTC
His-NMB0949-GFP	TTAATTAAATGCACCACCACCACCACCACGTAGAACGTAAATTGACCGGTGCCCATTACGGTTTGC GCGATTGGGTGATGCAACGTGCGACTGCGGTTATTATGTTGATTTATACCGTTGCACTTTTAGTGG TTCTATTTTCCCTGCCTAAAGAATATTCGGCATGGCAGGCATTTTTTAGTCAAACTTGGGTAAAAGT ATTTACCCAAGTGAGCTTCATCGCCGTATTCTTGCACGCTTGGGTGGGTATCCGCGATTTGTGGAT GGACTATATCAAACCCTTCGGCGTGCGTTTGTTTTTGCAGGTTGCCACCATCGTTTGGCTGGTCGG CTGTCTCGTGTATTCAGTTAAAGTGATTTGGGGGGATGGTGAGCAAGGGCGAGGAGCTGTTCACC GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCG GCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAA GCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTGACCTACGGCGTGCAGTGCTTCGCCCGC TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGA GCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGC GACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGG GGCACAAGCTGGAGTACAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGAAGAAC GGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACC ACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAG CACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTC GTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACTCGAGAAGCTTGATCC GGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCAT AACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGA GTTTAAAC

Key

=	**c-Myc tag**	=	**Direction of Forward primer**
=	**Restriction sites**	=	**Direction of Reverse primer**
=	**Annealing region of primers**	=	**Start codon**
=	**His tag**	=	**Signal peptide**
=	**GFP**		

Table 2.6 Full-length primer sequences used for PCR amplifications. Restriction sites are underlined. c-Myc and His tags and GFP reporter are in bold. Annealing temperatures are also indicated.

Gene	Primer name	Sequence (5'-3')	Annealing temp. (°C)
fHbp	PacfHbpFor	gaccTTAATTAAATGCCGTCTGAACCGCCGTTCGGA	49.9
	PmefHbpMycRev	acgtGTTTAAA**CCTACAGGTCTTCTTCGCTAATCAGTTTC TGTTC**GTTTGTATACTTGGAACTC	
NMB1468	PacNMB1468For	gaccTTAATTAAATGAAAAAATTATTGATTGCCGCAA	51.8
	PmeNMB1468Myc Rev	acgtGTTTAAA**CCTACAGGTCTTCTTCGCTAATCAGTTTC TGTTC**TTTGGCGGCATCTTTCAT	
NMB0727	PacNMB0727For	gaccTTAATTAAATGATAACTATTTCAAATGA	40.0
	PmeNMB0727Myc Rev	acgtGTTTAAAC**CTACAGGTCTTCTTCGCTAATCAGTTTC** TGTTCTTCAATAATACAATAACCT	
NMB0949	PacNMB0949For	gaccTTAATTAAATGGTAGAACGTAAATTGACCGGT	50.0
	PmeNMB0949Myc Rev	acgtGTTTAAAC**CTACAGGTCTTCTTCGCTAATCAGTTTC TGTTC**CCCCCAAATCACTTTAACT	
NMB1447	PacNMB1447For	gaccTTAATTAAATGATGAAACTCAATCCCCAACAG	54.9
	PmeNMB1447Myc Rev	acgtGTTTAAAC**CTACAGGTCTTCTTCGCTAATCAGTTTC TGTTC**ATCCGCCGCGCCGCTTTTT	
NMB1564	PacNMB1564For	gaccTTAATTAAATGCAGGTTACATCAAAATG	42.0
	PmeNMB1564Myc Rev	acgtGTTTAAA**CCTACAGGTCTTCTTCGCTAATCAGTTTC TGTTC**TTTATCTGCCCCGGCAATT	
NMB1566	PacNMB1566For	gaccTTAATTAAATGAAAAACATCGTCATCCTGA	51.7
	PmeNMB1566Myc Rev	acgtGTTTAAAC**CTACAGGTCTTCTTCGCTAATCAGTTTC TGTTC**CGCCGTCAGAAAACGGGCG	
Primers used for His-NMB0949-GFP amplification			
NMB0949	BamHIS0949For	aacaGGATCCATG**CACCACCACCACCACCAC**GTAGAAC GTAAATTGACCGGT	50
	EcoRI0946Rev	agctGAATTCCCCCCAAATCACTTTAACTG	
His-NMB0949-GFP	PacHis0949 For	agctTTAATTAAATG**CACCACCACCACCACCAC**GTAGAAC GTA	40-45

	PmelGFP Rev	agctGTTTAAAC**TCCGGATATAGTTCCTCCTTTCAG**	

2.3.5. Agarose gel electrophoresis

PCR products were visualised using 1% (w/v) agarose gel stained with 1% (v/v) SYBR Safe (Invitrogen™) in 1X TBE diluted from 10X TBE (800 ml H_2O, 106 g Tris base, 55 g Boric acid, 7.5 g EDTA disodium salt at pH 8). 1kb Plus DNA ladder (Invitrogen™) was used to verify band sizes according to base pairs (bp). Briefly, 1 µl 1kb Plus DNA ladder was mixed with 1 µl 5X GelPilot DNA loading dye (Qiagen) and 3 µl RNase free water and annotated as L in all agarose gel electrophoresis figures. Mini-sub Cell GT cell (Bio-RAD) electrophoresis chamber was used to run agarose gels at 70 volts for 40 minutes. Each well was loaded with 5 µl of sample which included 4 µl PCR product and 1 µl 5X loading dye.

2.3.6. Restriction digest

Restriction digests were performed using enzymes purchased from New England Biolabs using the components listed in Table 2.7 and 2.8 at 37°C with 1 hour incubation for each digest.

Table 2.7 Components of a typical restriction digest of plasmid DNA.

Components	Volume (µl)
Plasmid DNA	25
Restriction enzyme	5
CutSmart buffer (10 X)	10
RNase free water	60

Table 2.8 Components of a typical restriction digest of PCR product.

Components	Volume (µl)
PCR product	50
Restriction enzyme	5
CutSmart buffer (10 X)	10
RNase free water	35

2.3.7. Dephosphorylation of plasmid vector

Following restriction digest, the plasmid vector was dephosphorylated using Alkaline Phosphatase, Calf Intestinal (CIP) enzyme (Invitrogen™). The enzyme catalyses the removal of 5' phosphate in order to prevent self-ligation of the vector. The reaction was performed using 50 µl purified vector, 10 µl 10X CutSmart buffer (New England Biolabs), 38 µl H_2O and 2 µl CIP, incubated at 37ºC for 30 minutes. A further 2 µl of CIP was added and further incubated at 37 ºC for another 30 minutes.

The dephosphorylated vector was purified (section 2.3.3) and eluted in 12 µl elution buffer.

2.3.8. Ligation reaction

The following formula was used to determine the amount of insert DNA and plasmid vector DNA required for each ligation reaction:

$$ng\ of\ insert\ required\ = \frac{ng\ of v\ ector\ x\ kb\ size\ of\ insert}{kb\ size\ of\ vector} x\ molar\ ratio\ of\ \frac{insert}{vector}$$

A 1:6 vector to insert molar ratio was used for all ligation reactions. A plasmid concentration of 50 ng was used in a total volume of 10 µl, containing; 1 µl T4 DNA ligase (Promega), 1 µl 10X T4 ligase buffer (Promega), and RNase free water. Each reaction included a negative control of vector-only and no insert DNA, in order to verify any self-ligation. The reaction was incubated overnight at 4°C.

2.4. Transforming *E.coli*

2.4.1. Transforming Subcloning Efficiency™ DH5α™ Competent cells

Each ligation reaction was used to transform subcloning Efficiency™ DH5α™ Competent cells (Invitrogen™) using the heat shock method. A total of 2 µl of each ligation reaction, alongside the vector-only negative control was added to 50 µl competent cells. Cells were gently mixed then incubated on ice for 30 minutes. Heat shock was applied by incubating at 42ºC for 30 seconds and immediately transferring onto ice. All contents were then transferred into a 10 ml falcon tube containing 950 µl LB broth. Cells were left to shake at 225 rpm for 1 hour at 37ºC. Following incubation, 100 µl of this “dilute” cell culture was plated onto pre-warmed LB plates containing 30 µg/ml kanamycin when pGCC4 (Addgene) plasmid was used and 100 µg/ml ampicillin when pRSET-EmGFP (Invitrogen™) plasmid was used. The remaining contents were spun down for 1 minute at 6000 g and with the removal of most of the supernatant, the pellet was re-suspended in the remaining 100 µl. This “neat” suspension was plated onto antibiotic plates and incubated overnight at 37°C.

2.4.2. Screening of transformed cells

Transformants were selected from either dilute or neat positive control plates following verification of no growth on negative control plates. The number of colonies varied from each transformation. In each case, several colonies were isolated and used to inoculate LB broth (section 2.2) ready for plasmid DNA extraction (section 2.3.2).

2.4.3. Verifying *E. coli* transformants by PCR and DNA sequencing

Transformants were verified by PCR using primers specific for the pGCC4 vector (Table 2.9). Transformants were verified according to the size of PCR product. The pGCC4 vector backbone was used as a positive control and no DNA as a negative control for the PCR. The forward pGCC4 primer anneals up to 116 bp upstream of the *P*acI restriction site and the reverse primer anneals up to 91 bp downstream of the *P*meI restriction site. Agarose gel electrophoresis was used to visualise PCR products as described in section 2.3.5. Once the correct band size was confirmed, plasmid DNA extractions were prepared and sent off for sequencing to Eurofins Genomics or Genewiz in the volumes and concentrations shown in Table 2.9. The pGCC4 primers used for PCR were also used for DNA sequencing (Table 2.9).

Table 2.9 Primers used for PCR and DNA sequencing to verify transformants. The concentrations and volumes required for sequencing by Eurofins Genomics or Genewiz are shown.

	Eurofins Genomics		**Genewiz**	
Sample	Concentration	Sample volume (µl)	Concentration	Sample volume (µl)
Plasmid DNA	50-100 ng/µl	15	100 ng/µl	20
Forward primer	10 µM	30	5 µM	10
Reverse primer	10 µM	30	5 µM	10
pGCC4 primers		**Sequence (5'-3')**		**Annealing temp. (°C)**
Forward pGCC4 primer		AGACATCCACCAAACCATCC		50
Reverse pGCC4 primers		TGCTTCCGGGTGTTGTGTGG		

2.5. Transforming *N. meningitidis*

2.5.1. Transforming MC58 and MC58Lnt

Following plasmid extraction, all plasmid constructs (pGCC4-fHbp, pGCC4-NMB1468, pGCC4-NMB0727, pGCC4-NMB0949, pGCC4-NMB1477, pGCC4-NMB1564, pGCC4-NMB1566 and pGCC4-His-NMB0949-GFP) were individually transformed into both MC58 and MC58Lnt in order to express the tagged putative lipoproteins. Piliated cells from each strain were streaked onto pre-warmed GC plates lacking antibiotic with two circles of 1 cm diameter marked under the plate; one for DNA and one for the no DNA negative control. Transformation media was prepared by mixing 10 ml GC broth, 100 µl Kellogg's I, 10 µl Kellogg's II and 100 µl 1M $MgSO_4$ which was then filter sterilised. 10 µl plasmid DNA was mixed with 10 µl transformation media and 15 µl of this mixture was used to spot onto the streaked agar within the circle labelled DNA. Likewise, 15 µl of transformation media alone was spotted onto the circle labelled no-DNA. Each plate was left facing up until all media was adsorbed. Plates were inverted and incubated overnight at 37°C in 5% CO_2. 1 ml of filter sterilised supplemented GC broth was placed into 2 sterile 1.5 ml eppendorfs, one labelled DNA and one no DNA. A 10 µl sterile loop was used to gently scope bacterial cells from the circle containing DNA which was re-suspended in the eppendorf labelled DNA and likewise for no DNA. 100 µl were plated onto 10 pre-warmed GC plates containing 0.3 µg/ml erythromycin. 100 µl no DNA were also plated onto an erythromycin plate and incubated at 37°C in 5% CO_2. Plates were screened the next day for colonies on the DNA plates and the 'no DNA' plates were checked for no growth. Individual colonies were then picked and re-streaked onto erythromycin containing plates.

2.5.2. Verification of meningococcal transformants

PCR was used to verify MC58 and MC58Lnt strains transformed with each construct using the pGCC4 primers shown in Table 2.9 and reagents and conditions stated in Table 2.3 and 2.4. For each construct, one or two transformant colonies from each strain of MC58 and MC58Lnt were selected and genomic DNA was extracted (section 2.3.1), followed by PCR amplification (section 2.3.4) and agarose gel electrophoresis (section 2.3.5).

2.6. Detection of tagged proteins by Sodium Dodecyl Sulphate-Polyacrylamide Gel electrophoresis (SDS-PAGE) and Western blotting

2.6.1. SDS-PAGE

It was shown previously in MC58Lnt that the failure to triacylate lipoproteins resulted in the accumulation of diacylated lipoprotein which are subsequently down regulated or proteolysed (da Silva *et al*., 2016). This occurs progressively during bacterial growth such that at A_{600} 1.0 all the diacylated lipoprotein disappears (unpublished). Therefore, all c-Myc tagged transformants as well as negative controls were harvested at A_{600} 0.1 using broth cultures (section 2.2) unless where stated. Cells were harvested at A_{600} 0.1, 600 µl of bacterial suspension were transferred to a sterile 1.5 ml eppendorf and centrifuged at 14500 g for 10 minutes. The supernatant was removed by pouring and pipetting. The pellet formed was re-suspended in 300 µl of 1X Laemmli sample buffer (2-mercaptoethanol, bromophenol blue, glycerol, SDS, Tris-HCl) by pipetting and heated to 95ºC for 5 minute. For MC58 and MC58Lnt His-NMB0949-GFP transformants bacterial cells were grown on GC plates (section 2.2) instead of GC broth and bacterial cells were re-suspended in phosphate buffered saline (PBS) to an OD of A_{600} 1.0. 600 µl of this cell suspension was centrifuged at 1400 g for 10 minutes and the pellet formed was re-suspended in 300 µl of 1X Laemmli sample buffer by pipetting and heated to 95 ºC for 5 minutes.

Cell lysates were fractionated by 16% (w/v) SDS-PAGE (National Diagnostics). 25 µl of each cell lysate, including positive and negative controls were loaded alongside 7.5 µl of SeeBlue2® Pre-stained protein standard (Invitrogen™). Gels were run using 1X Tris-Glycine-SDS PAGE running buffer containing 0.025 M Tris Base, 0.192 M Glycine and 0.1 % (w/v) SDS (National Diagnostics) in the Mini-PROTEAN® Tetra Cell Systems (BIO-RAD) tank, at 150 volts for 1 hour and 30 minutes. Ice packs were placed around the tank to avoid overheating and distortion of bands.

2.6.2. Western blotting

The gels were transferred to a 0.22 µM PVDF membrane (GE Healthcare Life Sciences) using the TE 7 7 PWR Hoefer transblotter and a semi-dry approach. Each gel was run for 2 hours and 20 minutes at 42 mA. The PVDF membrane was briefly placed in methanol for activation and then soaked in transfer buffer

containing 48 M Tris base and 39 M Glycine at pH 8.3 with 70 % methanol. Six pieces of Whatmann paper (GE Healthcare Life Sciences) were soaked in the transfer buffer and 3 of these were stacked, then placed on the transblotter, followed by the PVDF membrane with the gel aligned above. The other 3 Whatmann papers were placed above this. Once transferred, the membranes were blocked for 2 hours on a shaking platform at room temperature. Blocking solutions were made in Tris-buffered saline (TBS) containing 0.1% (v/v) Tween-20 (TBST) using either bovine albumin serum (BSA) or dry milk (Table 2.10). The membranes were then incubated overnight at 4°C with the appropriate antibodies diluted in TBST using either BSA or dry milk, as shown in Table 2.10 for each transformant. The membranes were then washed for 2 hours in TBST by changing the solution every 20 minutes. Following this, the membranes were incubated at room temperature for 1 hour with the appropriate antibody diluted in TBST containing either BSA or dry milk. The membranes were then washed in TBST for 1 hour with 15 minute intervals of changing solution. The protein bands were detected using Amersham ECL start Western Blotting Detection Reagent (GE Healthcare Life Sciences). The expected molecular weights were calculated using Expasy Compute pI/Mw (http://web.expasy.org/compute_pi/).

Table 2.10 Western blot conditions for detection of recombinant MC58 and MC58Lnt proteins.

	Antibodies and dilutions used for tagged transformants		
Incubation Step	**c-Myc**	**His**	**GFP**
Blocking	5% Milk	3% BSA	3% BSA
Primary antibody	anti-c-Myc (Abcam) diluted 1:2000 in 1% milk	anti-His (Cell signalling Technology) diluted 1:1000 in 1% BSA	anti-GFP (eBiosciences™) diluted 1:1000 in 1% BSA
Secondary antibody	anti-mouse HPR-conjugated secondary (New England Biolabs) diluted 1:1000 in 1% milk	anti-rabbit IgG, HRP-linked secondary antibody (Cell Signalling Technology) diluted 1:1000 in 1% BSA	anti-mouse HPR-conjugated secondary (New England Biolabs) diluted 1:1000 in 1% BSA

2.7. Immunofluorescence microscopy

NMB1468 c-Myc recombinant strains of MC58 and MC58Lnt were harvested at A_{600} 0.1 from broth culture (section 2.2). 50 µl of bacterial culture were spotted onto circular glass cover-slips of 13 mm diameter and left to dry. The coverslips were placed in 24 well plates and fixed with 300 µl PBS containing 4% (v/v) paraformaldehyde for 20 minutes. The coverslips were then washed with 1 ml PBS and then blocked for 30 minutes with 500 µl PBS containing 1% (w/v) BSA. Blocking reagent was removed by washing with 1 ml PBS followed by overnight incubation at 4ºC with 500 µl PBS containing 1% (w/v) BSA and 5 µg/ml anti-c-Myc antibody (Abcam). The primary antibody was removed and the coverslips were washed 3 times using 1 ml PBS, followed by 1 hour incubation with 200 µl PBS containing 1% (w/v) BSA with Alexa Fluor 555 labelled donkey anti-mouse IgG (Abcam) secondary antibody (1:500) and then FITC-labelled rabbit polyclonal IgG (1:500) (Abcam) raised against whole cell *N. meningitidis,* with 1 ml washes in between each antibody. The coverslips were washed 3 times, dipped in ionised water and mounted onto a glass slide (cells facing down) by adding a drop of fluoroshield mounting medium containing DAPI (Abcam). The slides were left to dry and the coverslips were sealed gently using clear nail polish.

3. Results

3.1. Bioinformatics

3.1.1. Lipoproteins predicted by an algorithm tool in DOLOP

Babu *et al*. (2006) previously reported that 69 out of 2079 protein sequences in *N. meningitidis* MC58, i.e. 3.3 % of its genome contained a signal peptide at the N-terminus. In this study, systematic analysis of each protein sequence (2119) of MC58 revealed a further 10 proteins with a signal peptide at the N-terminus (highlighted in blue in Table 3.1) i.e. 3.7% of the genome.

Table 3.1 Protein sequences of *N. meningitidis* MC58 predicted as lipoproteins by the algorithm in the DOLOP website (http://www.mrc-lmb.cam.ac.uk/genomes/dolop/).

NMB no.	Function predicted in NCBI	Size of protein (no. of amino acids)	Sequence of the signal peptide	Amino acid at +2
NMB0032	Hypothetical protein	175	MEM K QMLLAVGVVAV LAGC	G
NMB0033	Putative membrane-bound lytic mureintransglycosylase A	441	MKKYLF R AALYGIAAAI LAAC	Q
NMB0035	Hypothetical protein	388	MR K FNLTALSVMLALG LTAC	Q
NMB0039	Hypothetical protein	90	MR K TFLFLTAAAAL LSGC	A
NMB0054	Hypothetical protein	135	MEIRAI K YTAMAALLAFTVAGC	R
NMB0071	ctrA-Capsule polysaccharide export outer membrane protein	391	MFKVKFYI R HAVLLLCGSL IVGC	S
NMB0086	Hypothetical protein	338	MYR K LIALPFALL LAAC	G
NMB0092	Hypothetical protein	75	MV R FFVLSFLTLINLCS LSAC	N
NMB0204	Lipoprotein	125	MN K TLILALSALLG LAAC	S
NMB0278	dsbA-1-Thiol:disulfide interchange protein DsbA	232	MKS R HLALGVAALFA LAAC	D
NMB0294	dsbA-2-Thio:disulphide interchange protein DsbA	231	MKL K TLALTSLTLLA LAAC	S
NMB0374	MafB-like protein	467	M K PLRRLTNL LAAC	A
NMB0375	mafA-1-adhesinMafA	313	M K TLLLLIPLV LTAC	G
NMB0430	prpB 2-methylisocitrate lyase	292	MMSQHSAGARFRQAV K ESNPLA V AGC	V
NMB0532	htrA-protease Do	499	MFK K YQYLALAALCAAS LAGC	D
NMB0580	Protein disulfide isomerase NosL	164	MK K TLLAIVAVSA LSAC	R
NMB0623	potD-2 spermidine/putrescine ABC transporter substrate-binding protein	379	MK K TLVAAAILSLA LTAC	G
NMB0652	mafA-2 adhesinMafA	313	MK TLLLLIPLV LTAC	G
NMB0653	MafB-like protein	422	M K PLRRLTNL LAAC	A
NMB0703	comL-Competence lipoprotein	267	MK K ILLTVSLGLA LSAC	A
NMB0752	bacterioferritin-associated ferredoxin	66	MFVCICNAVTDHQI K ETIAAGATTMGDLQSQLG VASC	C

NMB0787	amino acid ABC transporter substrate-binding protein	275	MMLK K FVLGGIAALV LAAC	G
NMB0844	Hypothetical protein	107	M KKCILGI LTAC	A
NMB0873	outer membrane lipoprotein LolB	193	M K HTVSASVILL LTAC	A
NMB0923	cytochrome c	152	M K TQISLAAAAITLL LSAC	G
NMB0938	Hypothetical protein	278	MKN K TSSLLLWLTAIM LTAC	S
NMB1010	Hypothetical protein	187	M K ILALLIAATCA LSAC	G
NMB1017	sbp sulfate ABC transporter substrate-binding protein	351	M K TYAPALYTAAL LTAC	S
NMB1035	Hypothetical protein	84	MN K LFITALSALA LSAC	A
NMB1057	ggt gamma-glutamyltranspeptidase	606	MPCMNHQSNSGEGVLVA K TYLLTALIMSMT ISGC	Q
NMB1060	fbp fructose-1,6-bisphosphatase	324	MDTLT R FLPEHLQQNQLPEALGGVLLSV VSAC	T
NMB1084	Hypothetical protein	173	MECHADVWHFILEKNMK K FYFVLLALG LAAC	G
NMB1087	Hypothetical protein	101	MSLLKTV K MQAAVALTALALTAC	S
NMB1107	Hypothetical protein	200	MNMK K LISAICVSIV LSAC	N
NMB1124	Hypothetical protein	215	M K PLILGLAAVLA LSAC	Q
NMB1125	Hypothetical protein	123	MMNPKTLS R LSLCAAVLA LTAC	G
NMB1126	Hypothetical protein	223	M K TVSTAVVLAAAAVS LTGC	A
NMB1162	Hypothetical protein	215	M K PLILGLAAVLA LSAC	Q
NMB1163	Hypothetical protein	123	MMNPKTLS R LSLCAAVLA LTAC	G
NMB1164	Hypothetical protein	223	M K TVSTAVVLAAAAVS LTGC	A
NMB1211	Hypothetical protein	80	M K YIVSISLAMG LAAC	S
NMB1212	Hypothetical protein	112	M K YIVSISLAMG LAAC	S
NMB1213	Lipoprotein	120	M K YIVSISLAMG LAAC	S
NMB1279	Membrane-bound lytic mureintransglycosylase B	369	MKKR K ILPLAICLAA LSAC	T
NMB1335	Hypothetical protein	186	MN R LLLLSAAVL LTAC	G
NMB1369	Hypothetical protein	184	MK K IIASALIATFA LAAC	Q
NMB1410	Hypothetical protein	179	MDFLEIFIMSAFR K ILLIISCLL IASC	L
NMB1433	Hypothetical protein	177	MFPPD K TLFLCLSALL LASC	G
NMB1468	Hypothetical protein	107	MK K LLIAAMMAAA LAAC	S
NMB1470	Hypothetical protein	181	ML K TSFAVLGGCLL LAAC	G
NMB1523	Hypothetical protein	98	MK K SLFAAALLSLV LAAC	G
NMB1533	Outer membrane protein	183	M K AYLALISAAVIG LAAC	S
NMB1541	Lactoferrin-binding protein	737	MC K PNYGGIVLLPLL LASC	I
NMB1567	Macrophage infectivity potentiator	272	MNTIF K ISALTLSAALA LSAC	G
NMB1578	Hypothetical protein	217	MFSVP R SFLPGVFVLAA LAAC	K
NMB1592	Lipoprotein	162	MK K YLIPLSIAAV LSGC	Q
NMB1594	Spermidine/putrescine ABC transporter substrate-binding protein	376	MT K HLPLAVLTALL LAAC	G
NMB1612	Amino acid ABC transporter substrate-binding protein	268	MNMK K WIAAALACSALA LSAC	G
NMB1623	Pan1 major anaerobically induced outer membrane protein	390	MK R QALAAMIASLFA LAAC	G
NMB1672	Hypothetical protein	172	M R LFPIAAALS LAAC	G
NMB1674	GDSL lipase	213	MPSEKPMNR R TFLLGAGALLLTAC	G
NMB1714	mtrE-Multidrup efflux pump channel protein	467	MDTTL K TTLTSVAAAFA LSAC	T
NMB1716	mtrC-Membrane fusion protein	412	MAFYAFKAM R AAALAAAVALV LS SC	G
NMB1764	Hypothetical protein	104	MK K TLSNLVLISFCSTM LTAC	P
NMB1765	Hypothetical protein	99	MKKKLS K YSLFLSSVFC LTAC	A
NMB1785	Hypothetical protein	79	MRDSMKNW K QFTFFVIL VIAC	Y
NMB1811	pilP Tfp pilus assembly protein PilP	181	M K HYALLISFLA LSAC	S
NMB1880	Hypothetical protein	321	MKP R FYWAACAVL LTAC	S
NMB898	Lipoprotein	171	MKIKQIV K PGLAVLAAGV LSAC	A
NMB1946	Outer membrane protein	287	MKTFF K TLSAAALALI LAAC	G
NMB1949	soluble lytic mureintransglycosylase	616	MYLPSM K HSLPLLAALV LAAC	S
NMB1969	Serotype-1-specific antigen	1082	MRTTPTFPTKTF K PTAMALAVATT LSAC	L
NMB1977	Hypothetical protein	56	M K YGVFFAAATALL LSAC	G
NMB1991	Iron (III) ABC transporter permease	324	MPSEKNIGFMAGSS R PLWVAFALL LVSC	V

NMB2002	Hypothetical protein	72	MSMPEMP K WYDDDGQ IVSC	T
NMB2091	Hemolysin	202	MKPKPHTV R TLIAAIFSLA LSGC	V
NMB2132	Transferrin-binding protein-like protein	488	MFK R SVIAMACIFA LSAC	G
NMB2139	Hypothetical protein	297	MVTFSKI R PLLAIAAAAL LAAC	G
NMB2147	Hypothetical protein	140	M R PIFLSFVLFPIL ITAC	S

In this table and in the subsequent tables (Table 3.2, 3.3 and 3.4) the assigned function and length of each putative lipoprotein is shown. The sequence of the signal peptides with the tripartite components are colour coded, with the n-region in brown, the h-region in green and the c-region (lipobox) in red. The residue at the +2 position is also shown.

Of the 10 novel probable lipoproteins identified, 3 have been annotated as hypothetical proteins (NMB1084, NMB1410, NMB1523), 3 as metabolic enzymes (NMB0430, NMB1057, NMB1060), 2 as proteins associated with iron transport and storage, 1 as an antigen and 1 as a lipoprotein. This is typical of the functional diversity represented by the remaining 69 probable lipoprotein reported by Babu *et al.* (2006). Other than NMB1592 annotated as a lipoprotein, the remaining 9 have not been previously reported to be lipoproteins in the literature and this knowledge sheds new light on their function.

According to the +2 rule, only two proteins, NMB0278 which encodes for DsbA and NMB0532 which encodes for HtrA are predicted as inner-membrane proteins due to the aspartate residue (D) present after the conserved cysteine. If this rule applies to *N. meningitidis,* this would indicate that all the other predicted proteins are outer membrane proteins.

As seen in the above table, of the 79 putative lipoproteins, 37 have been previously annotated as hypothetical proteins and therefore have no assigned function. This includes NMB1468, which was previously experimentally confirmed as a lipoprotein (Hou *et al*., 2008).

3.1.2. Investigating hidden lipoproteins

3.1.2.1. Lipoproteins like fHbp with the signal peptide located within the first 30% of amino acids from the start residue

Since it was previously shown in MC58 that NMB1870 encoding fHbp, the very well characterised lipoprotein contains its signal peptide 40 amino acids

downstream of the predicted translation start residue (da Silva *et al*., 2016), we speculated there may be more lipoproteins present in the genome that have been assigned the wrong start codon and have therefore been missed as putative lipoproteins.

Systematic analysis of each protein sequence of MC58 revealed a further 13 proteins with a signal peptide located within the first 30% of the predicted translation start codon (Table 3.2). NMB0798 and NMB2064 contained a signal peptide of 2 possible lengths and NMB1538 contained a signal peptide of 3 possible lengths.

Table 3.2 Protein sequences of MC58 with the signal peptide located up to 30% from the predicted start residue.

NMB no.	Function predicted in NCBI	Size of protein (no. of amino acids)	Protein sequence	Amino acid at +2
NMB1870	fHbp Annotated As Hypothetial protein	320	MPSEPPFGRHLIFASLTCLIDAVCKKRYHNQNVYILSILRM TRSKPVN R TAFCCLSLTTALI LTACCSSGGGGVAADIGA GLADALTAPLDHKDKGLQSLTLDQSVRKNEKLKLAAQGAE KTYGNGDSLNTGKLKNDKVSRFDFIRQIEVDGQLITLESGE FQVYKQSHSALTAFQTEQIQDSEHSGKMVAKRQFRIGDIA GEHTSFDKLPEGGRATYRGTAFGSDDAGGKLTYTIDFAA KQGNGKIEHLKSPELNVDLAAADIKPDGKRHAVISGSVLY NQAEKGSYSLGIFGGKAQEVAGSAEVKTVNGIRHIGLAAK Q	S
NMB0462	potD-1 spermidine/putrescine ABC transporter substrate-binding protein	459	MQAFSLYPPVGHPDSAKKRQNRADVFLPFWKQDNFLGK SVQYRFEFAQIYFTMPKPCAGVTMGRGDFFFSNLFHQES ARMK K SVLAVLAALS LAACGGSEKNAVQPQADAASAA NAEAAATDTLNIYNWSNYVDESTVEDFKKANNLKLTYDLY ENNETLEAKMLTGKSGYDLVVPGIAFLPRQIEAGAYQKVN KDLIPNYKNIDPELLKMLETADPGNQYAVPYFSGVNTIAITA KGKELLGGKLPENGWDLLFKPEYTHKLKSCGIALWDTPSE MFPILLNYLGKDPKGSNPEDLKAAAEVLKSIRPDVKRFSPS IIDELARGDICLAAGNGGDLNLAKARSEEVKNNVGIEVLTP KGMGFWIESWLIPADAKNVANAHKYINYTLDPEIAAKNGIA VTFAPASKPAREKMPAELVNTRSIFPNEQDMKDGFVMPQ MSTDAKKLSVSLWQKIKVGTN	G
NMB0530	beta-hexosaminidasev	361	MTVPHIPRGPVMADIAAFRLTEEEKQRLLDPAVGGIILFRR NFQNIEQLKTLTAEIKALRTPELIIAVDHEGGRVQRFIEGFT RLPAMSTLGEIWD K DGASAAETAAGQVGRVLATE LSA CGIDLSFTPVLDLDWGNCPVIGNRSFHRNPEAVARLALAL QKGLTKGGMKSCGKHFPGHGFVEGDSHLVLPEDWRSLS ELETADLAPFRIMSREGMAAVMPAHVVYPQVDTKPAGFS EIWLKQILRRDIGFKGVIFSDDLTMEGACGAGGIKERARIS FEAGCDIVLVCNRPDLVDELREDFRIPDNPTLAQRWQYM ANTLGSAAAQAVMQTADFQAAQAFVAGLASPQDTAGGV KVGEAF	G
NMB0751	XerD Tyrosine recombinase	291	MEEGLIDRLLETLWLDRRLSQNTLNGYRRDLEKIARRLSQ SGRMLKDADEADLAAAVYVDGEQ R SSQARA LSACKR LYIWMEREGIRTDNPTRLLKPPKIDKNIPTLITEQQISRLLAA PDTDTPHGLRDKALLELMYATGLRVSEAVGLNFGNVDLD RGCITALGKGDKQRMVPMGQESAYWVERYYTEARPLLLK GRNCDALFVSQKKTGISRQLAWMIVKEYASQAGIGHISPH SLRHAFATHLVRHGLDLRVVQDMLGHADLNTTQIYTHVAN VWLQGVVKEHHSRN	K
NMB0798	FtsH, cell division protein	655	MGNTFKSILVWVALGIGLMAAFNALDGKKEDNGQIEYSQF IQQVNNGEVSGVNIEGSVVSGYLIKGERTDKSTFFTNAPL DDNLIKTLLDKNVRVKVTPEEKPSALAALFYSLLPVLLIGA WFYFMRMQTGGGGKGGAFSFGKSRARLLD K DANKVT	D

			FAD VAGCDEAKEEVQEIVDYLKAPNRYQSLGGRVPRGIL LAGSPGTGKTLLAKAIAGEAGVPFFSISGSDFVEMFVGVG ASRVRDMFEQAKKNAPCIIFIDEIDAVGRQRGAGLGGGND EREQTLNQLLVEMDGFESNQTVIVIAATNRPDVLDPALQR PGRFDRQVVVPLPDIRGREQILNVHSKKVPLDESVDLLSL ARGTPGFSGADLANLVNEAALFAGRRNKVKVDQSDFEDA KDKIYMGPERRSMVMHEDEKRATAYHESGHAIVAESLPF TDPVHKVTIMPRGRALGLTWQLPERDRISMYKDQMLSQL SILFGGRIAEDIFVGRISTGASNDFERATQMAREMVTRYG MSDKMGVMVYAENEGEVFLGRSVTRSQNISEKTQQDIDA EIRRILDEQYQVAYKILDENRDKMETMCKALMEWETIDRD QVLEIMAGKQPSPPKDYSHNLRENADAAEDNAPHAPTRE ETEAPAPADTASTESEQQPENKA	
NMB0928	Hypothetical protein	398	MPSEPFGRHNATNTLISITQDDTMTHI K PVIAALALIG LA ACSGSKTEQPKLDYQSRSHRLIKLEVPPDLNNPDQGNLY RLPAGSGAVRASDLEKRRTPAVQQPADAEVLKSVKGVRL ERDGSQRWLVVDGKSPAEIWPLLKAFWQENGFDIKSEEP AIGQMETEWAENRAKIPQDSLRRLFDKVGLGGIYSTGERD KFIVRIEQGKNGVSDIFFAHKAMKEVYGGKDKDTTVWQPS PSDPNLEAAFLTRFMQYLGVDGQQAENASAKKPTLPAAN EMARIEGKSLIVFGDYGRNWRRTVLALDRIGLTVVGQNTE RHAFLVQKAPNESNAVTEQKPGLFKRLLGKGKAEKPAEQ PELIVYAEPVANGSRIVLLNKDGSAYAGKDASALLGKLHSE LR	S
NMB0982	Chloride channel protein	380	MHFIQHTAYGYGADGVYTSFREGVAQASGMRRVAVLTLC GAVAGSGWWLLKRFGKPQIEIKAALKQPLQGLPFLTTVFH VLLQIITVGLGSPLGREVAPREMTAAFAFAGGK R LGLDE GEMRL LIACASGAGLAAVYNVPLASTLFILEAMLGVWTQ QAVAAALLTSVIATAVARIGLGDVQQYHPANLTVNTSLLWF SAVIGPILGVAAVFFQRTAQKFPFIKRDNIKIIPLAVCMFALI GVISVWFPEILGNGKAGNQLTFGGLTDWQHSLGLTAVKW LVVLMALAVGAYGGLITPSMMLGSTIAFAAATAWNSVFPE MSSESAAIVGAAVFLGVSLKMPLTAIAFILELTYAPVALLMP LCTGMAGAVWVAKKMGFK	A
NMB1206	bfrB Multispecies: bacterioferritin	157	MKGDRLVIRELNKNLGLLLVTINQYFLHARILKNWGFEELG EHFFKQSIVEMKAADDLIERILFLEGLPNLQELG K LLIGE STEE IIACDLTKEQEKHEALLAAIATAEAQQDYVSRDLLEK QKDTNEEHIDWLETQQELIGKIGLPNYLQTAAQED	G
NMB1269	Hypothetical protein	365	MNQTFTLPDTRPYPQNPIKNHLLLNAYQLAHNSSQASRKL SSGQLQTEIRGMLEQNHYINLSLALTMSPDAGTYAALLSS VNAVLDCE K EGEVQWFALPVVL VSGCKKERAIEMKLPT EALFACLQNYPHLRALTQETQWLPYLVHSSDLSAVAPDE WWRAKQNTEAAAQHLRRFAPRPLLLPEGQSVHVVYALG FGSGKVQTALGQNLLQAGLPLMQVWQENLASEGVTLFAN PLSPDSPVRALSDGSHTRQRMAMDVFAANAIRAVRMQSP RVGVVAAAKAGGQILFGFNATDGAFEVVPQVFSWQLSFT DNIAVIQQNFLDLMAECRVEHVYLLHNPLSAGEQESIPSYA EALKREGHNPFF	K
NMB1324	trxB Thioredoxin reductase	218	MSQHRKLIILGSGPAGYTAAVYAARANLNPVIITGIAQGGQ LMTTTEVDNWPADADGVQGPELMARFLAHAERFGTEIIFD QINAVDLQKRPFTLKGDMGEYTCDALIVATGASA K YLGL PSEEAFAGKG VSACATCDGFFYKNQDVAVVGGGNTAVE EALYLANIAKTVTLIHRRSEFRAEKIMIDKLMKRVEEGKIILK LESNLQEVLGDDRGVNGALLKNNDGSEQQIAVSGIFIAIGH KPNTDIFKGQLEMDEAGYLKTKGGTADNVGATNIEGVWA AGDVKDHTYRQAITSAASGCQAALDAERWLGSQNI	A
NMB1538	RpoD RNA polymerase sigma factor RpoD	642	MSRNQNHEEYQDDTRPLSIEEQRARLRQLIIMGKERGYIT YSEINDALPDDMSDADQIDNIVSMISGLGIQVTEHAPDAED ILLSDNAAVTDDDAVEEAEAALSSADSEFGRTTDPVR*MYM* REMGQVDLLTREDEIIIAK K IENALKNMVQA ISACPGSIA EILELIEKIRKDEIRVDEVVEAIIDPNEVLLNELGLGHLETTA PEKPSNDNSDENEDDEESEEDADEISAANLAELKQKVIGH FAQIEKDYKKMIGRLEKHHSRHKDYLAYRDAIANKLLEVRF ATRQIDSLSSSLRGKVENIRKLEREIRDICLDRVHMERDYFI QNFLPEITNLEWIEEEIAKGRVWSDALDRFRHAILEKQTEL ADMEKETRISIEELKEINKNMVSSEKETAAAKQEMIQANLR LVISIAKKYTNRGLQFLDLIQEGNIGLMKAVDKFEYRRGYK FSTYATWWIRQAITRSIADQARTIRIPVHMIETINKMNRISR QHLQETGEEPDSAKLAELMQMPEDKIRKIMKIAKEPISMET PIGDDDDSHLGDFIEDANNVAPADAAMYTSLHEVTKEILES LTPREAKVLRMRFGIDMNTDHTLEEVGRQFDVTRERIRQI EAKALRKLRHPTRSDRLRSFLDSEDSKL	P
NMB1564	Hypothetical protein	140	MQVTSKWIDGMCFVGTTEGGHSVV MEGSAAEGKAK R GPSPLEMLLLG VAGCSSIDVVMIAEKQRQKVTDCRATVT AKRADDAPRVFTEIHIHFKVFGHDLKESAIERAVQMSAEKY CSASIMLGKAAKITHSFEIAGADK	T

NMB2029	thrB Homoserine kinase	305	MSVYTSVSDDEMRGFLSGYDLGEFVSLQGIAQGITNSNYF LTTTSGRYVLTVFEVLKQEELPFFLELNRHLSMKGVAVAA PVARKDG R LDSVLAGKPAC LVACLKGSDTALPTAEQC FHTGAMLAKMHLAAADFPLEMENPRYNAWWTEACARLL PVLSQDDAALLCSEIDALKDNLGNHLPSGIIHADLFKDNVL LDGGQVSGFIDFYYACRGNFMYDLAIAVNDWARTADNKL DEALKKAFIGGYEGVRPLSAEEKAYFPTAQRAGCIRFWVS RLLDFHFPQAGEMTFIKDPNAFRNLLLSLG	I
NMB2064	Hypothetical protein	462	MNAVVVAVIVMLVLSLSRVHVVLSLTVGAFVGGAVAGMPL QNIADAAGQVSQAGIIPVFNKGLEGGAKIALSYAMLGAFA MAITHSGLPQQLAGAVVRKLNRGGMPDSVRSGEGAVKW LLLSIILVMGMMSQNIIPIHIAFIPMIVPPLLLVFN R LKIDRR LIACVITFGLVTTYMFLPYGFGAIFLNEILLGNIHSAAPQLDV KNINVMAAMAIPALGMLAGLLLAFVHYRKPRLYQSNNADT AGNADAANRPQPSAYRSLAAAVAIAVCFAIQLMYEDSLVL GAMLGFAVFMMLGVINRDKANDVFGEGIKMMAMVGFIMI AAQGFAAVMNATGHIQPLVESSMAIFGNSKGMAALAMLV VGLLVTMGIGSSFSTLPIIAAIYVPLCVGLGFSPLATVAIVGT AGALGDAGSPASDSTLGPTMGLNADGQHDHIRDSVIPTFI HYNIPLLIAGWIAAMVL	V

The protein sequence immediately downstream of the signal peptide is highlighted. Signal peptides with different lengths are underlined and shown in italics. Proteins selected for study are highlighted in yellow.

Within this group of predicted protein sequences, 4 have been previously annotated as hypothetical proteins with no assigned function. According to the above data only 1 protein, NMB0798 which encodes for FtsH contains a D residue at position +2.

3.1.2.2. Lipoproteins with signal peptide located in the middle of the protein

We furthered searched for additional signal peptides located near the middle of the protein.

Table 3.3 Protein sequences of MC58 with signal peptide located near the middle of the protein.

NMB no.	Function predicted in NCBI	Size of protein (no. of amino acids)	Protein sequence	Amino acid at +2
NMB0618	ppsA phosphoenolpyruvate synthase	794	MADNYVIWFENLRMTDVERVGGKNASLGEMISQLTE KGVRVPGGFATTAEAYRAFLAHNGLSERISAALAKLD VEDVAELARVGKEIRQWILDTPFPEQLDAEIEAAWNK MVADAGGADISVAVRSSATAEDLPDASFAGQQETFLN INGLDNVKEAMHHVFASLYNDRAISYRVHKGFEHDIVA LSAGVQRMVRSDSGASGVMFTLDTESGYDQVVFVTS SYGLGENVVQGAVNPDEFYVFKPTLKAGKPAILRKTM GSKHIKMIFTDKAEAGKSVTNVDVPEEDRNRFSITDEE ITELAHYALTIEKHYGRPMDIEWGRDGLDGKLYILQAR PETVKSQEEGNRNLRRFAINGDKTVLCEGRAIGQKVG QGKVRLIKDASEMDSVEAGDVLVTDMTDPDWEPV**MK RASAIVTNRGGRTCHAAIIA R ELGIPA VVGC**GNATE LLKNGQEVTVSCAEGDTGFIYAGLLDVQITDVALDNM PKAPVKVMMNVGNPELAFSFANLPSEGIGLARMEFIIN RQIGIHPKALLEFDKQDDELKAEITRRIAGYASPVDFYV DKIAEGVATLAASVYPRKTIVRMSDFKSNEYANLVGG NVYEPHEENPMLGFRGAARYVADNFKDCFALECKAL KRVRDEMGLTNVEIMIPFVRTLGEAEAVVKALKENGLE RGKNGLRLIMMCELPSNAVLAEQFLQYFDGFSIGSND MTQLTLGLDRDSGLVSESFDERNPAVKVMLHLAISAC RKQNKYVGICGQGPSDHPDFAKWLVEEGIESVSLNP DTVIETWLYLANELNK	G
NMB1130	Squalence synthase HpnD	290	MKGLDYCRQKAEESRSSFLSGFRFLTQEKRDAVTVLY AFCRELDDVVDECSNPDVAQATLNWWRGDLDKVFG GAMPEHPVNQALRQVKETFKLPKYELEALIDGMQ**MD LVQARYGSFEEL K LYCHRVAG VVGC**LIARILGFSDD QTLEYADKMGLALQLTNIIRDVGEDARRGRIYLPMEE MRRFDVPASVILQCSPTGNFAELMAFQIKRARETYRE AVSLLPDADKKAQKVGLVMAAVYYELLNEIDRDGAQN VLKYKIALPSPRKKRIALKTWLFGFKPRPGTPERA	L
NMB1566	purN phosphoribosylglycinamidetransformylase	208	MKNIVILISGRGSNMQAIVNAAIHNVRIAAVLSNSETAA GLQWAAERGIPTDSLNHKNFTSRLAFDTAMM EKIDAYQPDLVVLAGFMRILTPEFCARYEGRL**MNIHPSI LPSFTGLHTHE R ALEAGCR VAGC**TIHFVTAELDCGP IVSQGVVPILDGDTADDIAARVLAVEHKLYPKAVADFA AGRLIIEGNRVRNSENADAARFLTA	T

The protein sequence immediately downstream of the signal peptide is highlighted. Proteins selected for study are highlighted in yellow.

3.1.2.3. Lipoproteins with signal peptide located towards the C-terminus of the protein

15 protein sequences of MC58 with the signal peptide located towards the c-terminus were identified. NMB1151 contained a signal peptide of 2 possible lengths and NMB1996 contained a signal peptide of 3 possible different lengths.

Table 3.4 Protein sequences of MC58 with signal peptide located towards the c-terminus.

NMB no.	Function predicted in NCBI	Size of protein (no. of amino acids)	Protein sequence	Amino acid at +2
NMB0111	fmt methionyl-tRNAformyltransferase	308	MKVIFAGTPDFAAAALRAVAAAGFEIPLVLTQPDRPK GRGMQLTAPPVKQAALELGLRVEQPEKLRNNAEALQ MLKEVEADVMVVAAYGLILPQEVLDTPKHGCLNIHAS LLPRWRGAAPIQRAIEAGDAETGVCIMQMDIGLDTGD VVSEHRYAIQPTDTANEVHDALMEIGAAAVVADLQQL QSKGRLNAVKQPEEGVTYAQKLSKEEARIDWSKSAA VIERKIRAFNPVPAAWVEYQGKP**MKIR R AEVVAQQ GAAGEVLSCSADGL VVAC**GENALKITELQPAGGRR MNIAAFAAGRHIEAGAKL	G
NMB0382	rmpM-Outer membrane protein	242	MTKQLKLSALFVALLASGTAVAGEASVQGYTV GQSNEIVRNNYGECWKNAYFDKASQGRVECGDAVA APEPEPEPEPAPAPVVVVEQAPQYVDETISLSAKTLF GFDKDSLRAEAQDNLKVLAQRLSRTNVQSVRVEGHT DFMGSDKYNQALSERRAYVVANNLVSNGVPVSRISA VGLGESQAQ**MTQVCEAEVAKLGAKVS K AKKREA LIAC**IEPDRRVDVKIRSIVTRQVVPAHNHHQH	I
NMB0400	Transposase, truncated	190	MPYYLYCLRLRRLVLIFVNPLYLHQFHETQVGSVKQLI AHFDRLIDELDKQIDDHTHTHFDGKAQVAEQI KGIGSITTATLMAMLPELRRLSHKRIAGLAGIAPHPRE SGETKFKSRCFGGRSAVRKALY**MATVAATRFEPLIR DFHQ R PLSEGKPYKVA VTAC**MRKLLTISNARMRD YFAENDTAENGI	M
NMB0727	DNA modification methylase	216	MITISNEDNMILMSRYPDKYFDLAIVDPPYGILNKTKR GGDYKFNMNEYSQWDIKPDQTYFNELFRVSKN QIIWGGNYFGELWLRSEYNKGFIIWDKNQPETLNNFS MAEMAWSSFDRPSKIFRFSVRKNRNKTHPTQKP VELYQWLLK**MYAKQGD K ILDTHLGSGTLAIACCIAQ FD LTAC**EINSDYYQQSIEKIKNNLPEARISFGHP GYCIIE	E
NMB0949	sdhD Succinate dehydrogenase, hydrophobic membrane anchor protein	113	MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVL FSLPKEYSAWQAFFSQTWVKVFTQVSFIAVFLHAWV GIRDLW**MDYIKPFGV R LFLQVATIVW LVGC**LVYSV KVIWG	L
NMB0998	Oxidoreductase	1277	MTTTTAPQRIREIPYNYTSYTDREIVIRLLGDEAWQIL QDLRGQRKTGRSARMLFEVLGDIWVVVRNPYLVDDL LEHPKRRAALVREMRHRLNEIRKRRDDNRQVDVLVA AAEKAVERFDSSFDETSQKRRQILERLSKITKPHNIMF DGLARVTHVTDATDWRVEYPFVVVNPDTEAEIAPLV RALIELDLVIIPRGGGTGYTGGAIPLDANSAVINTEKLD KHRGVEYVELAGLDGKHPIIRCGAGVVTRRVEETAH QAGLVFAVDPTSADASCVGGNVAMNAGGKKAVLWG TALDNLAYWNMVNPQGEWLRIERVRHNFGKIHDEET AVFDVHTLDSDGINIVKTERLEIPGHKFRKVGLGKDVT DKFLSGLPGVQKEGTDGIITSVAFVLHKMPKYTRTVC MEFFGTVATATPSIVEIRDFLLAHESVRLAGLEHLDW RYVRAVGYATKAAGKGRPKMVLLADVVSDDEAAVE AAAEHICELARARDGEGFIAVSPEARKTFWLDRSRTA AIAKHTNAFKINEDVVIPLERLGEYSDGIERINIELSIQN KLKLCAALEQYLSGKLPIDKMGTDLPTAELLGERGKH ALAHVSAVKTRWEWLLANLDTPLADYKARYGAAVHA APEAKNNESCFIAFRDFRLRVSVKADVMKPLSEIFSG KTDTKIIQGLGKIHAKTVRSRVFVALHMHAGDGNVHT NIPVNSDDAEMLQTAYRSVERIMKIARSLNGVISGEH GIGITKLEFLSDEEMQPFWDYKNQVDPKHTFNRHKL MKGSDLRNAYTPSFELLGAESLIMEKSNLGTIADSVK DCLRCGKCKPVCSTHVPRANLLYSPRNKILGVGLLIE AFLYEEQTRRGVSIKHFEELMDIGDHCTVCHRCVKP CPVNIDFGDVTVAVRNYLADSGHKRFAPAAAMGMAF LNATGPKTIKALRAAMIQIGFPAQNFAYKIGKLLPIGTK KQKAEPKATVGKAPIKEQVIHFINRPLPKNVPAKTPRS LLGIEDGKSIPIIRNPAAPEDAEAVFYFPGCGSERLFS QIGLAVQAMLWHVGVQTVLPPGYMCCGYPQDAGG NKAKAEEMSTNNRVAFHRMANTLNYLDIKTVVVSCG TCYDQLEKYRFEEIFPGCRIIDIHEYLLEKGVKLDGVK GQQYLYHDPCHTPIKTMNATQMASSLMGQKVVLSD	P

			RCCGESGMFAVKRPDIATQVKFRKQEEIEKNL K EL PQGEPVKM LTSCPACLQGLSRYADDNNMPADYIVIE MAKYILGENWLDEFVKKANNGGVEKVLL	
NMB1047	Hypothetical protein	28	MNKTLSILPVAILLGGCAAGGGNTFGSLDGGTGMGG SIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQAE WENKICACVAQEAPERMLAPST R NQALAALTAKT VSACFKHLYR	F
NMB1151	cysI-1 Sulfate reductase subunit beta	589	MTVQTKTKGLAWQEKPLSDNERLKTESNFLRGTILD DLKDPLTGGFKGDNFQLIRFHGMYEQDDRDIRAERA EAKLEPLKFMLLRCRLPGGIIKPSQWIELDKFARENS HYRSIRLTNRQTFQFHGVPKAKLQTMHRLLHKLGLD SIATAADMNRNVLCTSNPIESELHRQAYEYAKKISEHL LPRTRGYLDVWVDGKKVQSSDDFLQEDEPILGKTYL PRKFKTAVVIPPLNDVDCYGNDLDFVAVSDGNGQLA GFNVLAGGGLSMEHGNTKTYPNISLELGFVPPEHAL KAAEAVVTTQRDFGNRSDRKNARTRYTIQNMGLDNF RAEVERRMGMPFEPVRPFKFTGRGDRIGWVKGIDG NWHLTLFIESGRLVDEGGKQLLTGVLEIAKIHKGDFRI TANQNLIVANVAEADKAKIEEFARTYGLIRNDVSKLRE NAMSCVSFPTCPLAMAEAERVLPDFIGELDKIMA K HGTSDDYIVTR ITGCPNGCGRAMLAEIGLVGKAVGR YNLHIGGDREGVRIPRLYKENITLPEILAELDDLIGKW AAERNIGEGFGDFAIRTGIVKPVLNAPVDFWDASKAV AIARA	P
NMB1362	Excinuclease ABC subunit C	617	MNKETRFPEHFDIPLFLKNLPNLPGVYRFFNESGNVL YVGKAVNLKRRVSGYFQKNDHSPRIALMVKQVHHIE TTITRSESEALILENNFIKALSPKYNILFRDDKSYPYLM LSGHQYPQMAYYRGTLKKPNQYFGPYPNSNAVRDSI QVLQKVFMLRTCEDSVFEHRDRPCLLYQIKRCTAPC VGHISEEDYRDSVREAATFLNGKTDELTRTLQHKMQ TAAANLQFEEAARYRDQIQALGIMQSNQFIDSKNPNN PNDIDLLALAVSDGLVCVHWVSIRGGRHVGDKSFFP DTKNDPEPNGQDYAEAFVAQHYLGKSKPDIIISNFPV PDALKEALEGEHGKQMQFVTKTIGERKVRLKMAEQN AQMAIAQRRLQQSSQQHRIDELAKILGMDSDGLN R LECFDISHTQGEAT IASCVVYDEQNIQPSQYRRYNIT TAKPGDDYAAMREVLTRRYGKMQEAEANGETVKWP DAVLIDGGKGQIGVAVSVWEELGLHIPLVGIAKGPER KAGMEELILPFTGEVFRLPPNSPALHLLQTVRDESHR FAITGHRKKRDKARVTSSLSDIPGVGSKRRQALLTRF GGLRGVIAASREDLEKVEGISKALAETIYNHLH	V
NMB1447	Rep ATP-dependent DNA helicase	671	MMKLNPQQLEAVRYLGGPLLVLAGAGSGKTGVITQK IKHLIVNVGYLPHTVAAITFTNKAAAEMQERVAKMLPK PQTRGLTICTFHSLGMKILREEANHIGYKKNFSILDST DSAKIIGELLGGTGKEAVFKAQHQISLWKNDLKTPED VVQTASNIWEQQTARVYASYQETLQSYQAVDFDDLI RLPAVLLQQNSEVRNKWQRRLRYLLVDECQDTNTC QFTLMKLLTGAEGMFTAVGDDDQSIYAWRGANMEN LRKMQENYPQMKVIKLEQNYRSTARILKIANKVIENNP KLFTKKLWSQLGEGEPVKVVACQNEQHEADWVVSQ IVKQKLIGGDKTQYADFAVLYRGKHQARIFEEALRGA RIPYQLSGGQSFFDKAEIKDVLSYVRLLANPNDDPAF LRAVTTPKRGIGDVTLGKLNTYAHEHECSLYEAAQNE EALATLNNTNRQHLQTFMDMFVSYLAKAETSEAGEFI NSLLEEIDYENHLMQNEEGKAGEIKWRNVGDLVSWF ARKGGEDGKNIIELAQTVALMTLLEGKDEEETDAVSL STLHAA K GLEYPYVF LVGCEEGVLPHNDSIEEGNV EEERRLMYVGITRAKRQLTLTHCVKRKKQGTWQFPE PSRFIDEMPQEDLKILGRKGGEPIVSKEEGRRNLADII GRLDNLKKSGAAD	E
NMB1572	acnB Aconitate hydratase B	861	MLEAYRKAAAERAALGIPALPLNAQQTADLVELLKSP PAGEGEFLVELLAHRVPPGVDDAAKVKASFLAAVAE GSASSPLISPEYATELLGTMLGGYNIHALIELLDDDKL ASIAAKGLKHTLLMFDSFHDVQEKAEKGNKYAQEVL QSWADAEWFASRAKVPEKITVTVFKVDGETNTDDLS PAPDAWSRPDIPLHALAMLKNPRDGITPDKPGEVGPI KLLEELKAKGHPVAYVGDVVGTGSSRKSATNSVIWH TGEDIPFVPNKRFGGVCLGGKIAPIFFNTQEDSGALPI EVDVSALKMGDVVDILPYEGKIVKNGETVAEFELKSQ VLLDEVQAGGRINLIIGRGLTAKAREALKLPASTAFRL PQAPAESKAGFTLAQKMVGRACGLPEGQGVRPGTY CEPRMTTVGSQDTTGPMTRDELKDLACLGFSADMV MQSFCHTAAYPKPVDVKTHKELPAFISTRGGVSLRP GDGVIHSWLNRLLLPDTVGTGGDSHTRFPIGISFPAG	P

			SGLVAFAAATGVMPLDMPESVLVRFSGKLQPGVTLR DLVNAIPLYAIKQGLLTVAKAGKKNIFSGRILEIEGLPD LKVEQAFELTDASAERSAAGCTVKLNKEPIIEYMKSN VVLMKNMIANGYQDPRTLERRIKAMEKWLANPELLE AD K DAEYAAVIEINMDDIKEP IIACPNDPDDVCFMS ERSGTKIDEVFIGSCMTNIGHFRAASKLLEGKADTPV RLWIAPPTKMDAKQLSDEGHYGVLGRAGARMEMPG CSLCMGNQAQVREGATVMSTSTRNFPNRLGKNTFV YLGSAELAAICSKLGKIPTVEEYQANIGIINEQGDKIYR YMNFNEIDSYNEVAETVNV	
NMB1602	Transposae	372	MLIHYIDIAKRNFVIAVSSLSKTKTETNNPKGIAHTIEYL KKHKVALVVTESTGGLEIPAAKAIHRAGIA VIIANPRQTHQFAQSQSLTKTDAKDAKMPAFFAQMK AQKEDWQTMPYHPPTEAEEVLEALVNRRNQSADMR TAEKNRLHQVHETQVGSVKQLIAHFDRLIDESDKQID DHTHTHFDGKAQVAEQIKGIGSITTATLMAMLPELGR LSHKRIASLVGIAPHPRKSGEAKFKSRCFGGRSAVLK ALYMATVAATRFEPLIRDFHQ R PLSEGKPYKVA VT ACMRKLLETFAKFLSLTTTEIPTQVFGCFRPKYRLILP KHPLNPPRTPDNQASGLPFRRQRAHLACWRLSTGS NTSPSDGFAHSL	M
NMB1643	InfB Translation initiation factor IF-2	962	MSNTTVEQFAAELKRPVEDLLKQLKEAGVSKNSGSD SLTLDDKQLLNAYLTKKNGSNSSTISIRRTKTEVSTVD GVKVETRKRGRTVKIPSAEELAAQVKAAQTQAAPVR PEQTAEDAAKARAEAAARAEARAKAEAEAAKLKAAK AGNKAKPAAQKPTEAKAETAPVAAETKPAEESKAEK AQADKMPSEKPAEPKEKAAKPKHERNGKGKDAKKP AKPAAPAVPQPVVSAEEQAQRDEEARRAAALRAHQ EALLKEKQERQARREAMKQQAEQQAKAAQEAKTGR QRPAKPAEKPQAAAPAVENKPVNPAKAKKEDRRNR DDEGQGRNAKGKGGKGGRDRNNARNGDDERVRG GKKGKKLKLEPNQHAFQAPTEPVVHEVLVPETITVAD LAHKMAVKGVEVVKALMKMGMMVTINQSIDQDTALI VVEELGHIGKPAAADDPEAFLDEGAEAVEAEALPRPP VVTVMGHVDHGKTSLLDYIRRTKVVQGEAGGITQHIG AYHVETPRGVITFLDTPGHEAFTAMRARGAKATDIVIL VVAADDGVMPQTIEAIAHAKAAGVPMVVAVNKIDKEA ANPERIRQELTAHEVVPDEWGGDVQFIDVSAKKGLNI DALLEAVLLEAEVLEL TAPVDAPAKGIIVEARLDKGRGAVATLLVQSGTLKKG DMLLAGTAFGKIRAMVDENGKSITEAGPSIPVEILGLS DVPNAGEDAMVLADEKKAREIALFRQGKYRDVRLAK QQAAKLENMFNNMGETQAQSLSVIIKADVQGSYEAL AGSLKKLSTDEVKVNVLHSGVGGITESDVNLAIASGA FIIGFNVRADASSRKLAENENVEIRYYNIIYDAINDVKA AMSGMLSPEEKEQVTGTVEI R QVISVSKVGN IAGC MVTDGVVKRDSHVRLIRNNVVIHTGELASLKRYKDDV KEVRMGFECGLMLKGYNEIMEGDQLECFDIVEVARS L	M
NMB1684	serS serine—tRNA ligase	431	MLDIQLLRSNTAAVAERLARRGYDFDTARFDTLEERR KSVQVKTEELQASRNSISKQIGALKGQGKHEEAQAA MNQVAQIKTDLEQAAADLDAVQKELDAWLLSIPNLPH ESVPAGKDETENVEVRKVGTPREFDFEIKDHVDLGE PLGLDFEGGAKLSGARFTVMRGQIARLHRALAQFML DTHTLQHGYTEHYTPYIVDDTTLQGTGQLPKFAEDLF HVTRGGDETKTTQYLIPTAEVTLTNTVADSIIPSEQLP LKLTAHSPCFRSEAGSYGKDTRGLIRQHQFDKVEMV QIVHPEKSYETLEEMVGHAENILKALELPYRVITLCTG DMGFGAA K TYDLEVWVPAQNTYRE ISSCSNCEDF QARRLKARFKDENGKNRLVHTLNGSGLAVGRTLVAV LENHQNADGSINIPAALQPYMGGVAKLEVK	S
NMB1996	purl phosphoribosylformylglyvinamidine synthase	1320	MSVVLPLRGVTALSDFRVEKLLQKAAALGLPEVKLSS EFWYFVGSEKALDAATVEKLQALLAAQSVEQTP KAREGLHLFLVTPRLGTISPWASKATNIAENCGLAGIE RIERGMAVWLEGRLNDEQKQQWAALLHDRMTE SVLPDFQTASKLFHHLESETFSGVDVLGGGKEALVK ANTEMGLALSADEIDYLVENYQALQRNPSDVELM MFAQANSEHCRHKIFNADFILNGEKQPKSLFGMIRDT HNAHPEGTVVAYKDNSSVIEGAKIERFYPNAAE NQGYRFHEEDTHIIMKVETHNHPTAIAPFAGAATGAG GEIRDEGATGKGSRPKAGLTGFTVSNLNIPDLK QPWEQDYGKPEHISSPLDIMIEGPIGGAAFNNEFGRP NLLGYFRTFEEKFDGQVRGYHKPIMIAGGLGSI QAQQTHKDEIPEGALLIQLGGPGMLIGLGGGAASSM DTGTNDASLDFNSVQRGNPEIERRAQEVIDRCWQ	G

			LGGKNPIISIHDVGAGGLSNAFPELVNDARRGAVFKL REVPLEEHGLNPLQIWCNESQERYVLSILEKDL DAFRAICERERCPFAVVGTATDDGHLKVRDDLFANN PVDLPLNVLLGKLPKTTRTDKTVAPSKKPFHAGD IDITEAAYRVLRLPAVAAKNFLITIGDRSVGGLTHRDQ MVGKYQTPVADCAVTMMGFNTYRGEAMSMGEK PTVALFDAPASGRMCVGEAITNIAAVNIGDIGNIKLSA NWMAACGNEGEDEKLYRTVEAVSKACQALDLS IPVGKDSLSMKTVWQDGEEKKSVVSPLSLIISAFAPV KDVRKTVTPELKNVEDSVLLFVDLGFGKARMGG SAFGQVYNNMSGDAPDLDDTGRLKAFYSVIQQLVAE NKLLAYHDRSDGGLFAVLVEMAFAGRCGLDIDLN LLLAQTFITNHTALSQSLRTEEVKALAEWQETIARTLF NEELGAVIQVRKQDVADIINLFYQQQLHHNVF EIGTLTDENTLIIRDGQTHLISDNLIKLQQTWQETSHQI QRLRDNPACADSEFALIGDNERSALFADVKFDVNEDI AAPFINSGAKPKIAILREQGVNGQIE***MAAAFTRAGFD*** ***AYDVH***MSDLMAG R IHLADFKM LAACGGFSYGDVL GAGEGWAKSILFHPALRDQFAAFFADPDTLTLGVCN GCQMVSNLAEIIPGTAGWPKFKRNLSE QFEARLSMVHVPKSASLILNEMQGSSLPVVVSHGEG RADFALHGGNISADLGIALQYIDGQNQVTQTYPL NPNGSPQGIAGVTNADGRITIMMPHPERVYRAAQMS WKPEGWTELSGWYRLFAGARKALG	

The protein sequence immediately downstream of the signal peptide is highlighted. Proteins selected for study are highlighted in yellow.

In the above table, one protein has been previously annotated as a hypothetical protein and therefore have no assigned function.

3.2. Testing if signal peptides located internally in the protein are recognised for translocation, lipid modification and cleavage.

Since preprolipoproteins are translocated by the Sec translocon as linear proteins to which SecA or SecB proteins bind (Auclair *et al*., 2012), in theory it should be possible for preprolipoproteins with signal peptides located at any region of the protein to be bound by these proteins for transport across the inner membrane for subsequent lipid modification and cleavage of the signal peptide. By simply tagging a c-Myc epitope to the C-terminus of the protein, and expressing this in MC58, the size of the tagged protein can be determined by Western blotting with an anti-c-Myc antibody. If cleavage of the signal peptide occurs, this will be reflected by the reduced size of the tagged protein. In addition, upon expressing the tagged protein in MC58Lnt, the subtle mobility difference due to loss of the third fatty acid can be visualised.

Five proteins were chosen for this study; one with the signal peptide located within the first 30% of the protein from the predicted start residue (NMB1564) (Table 3.2), one with the signal peptide located in the middle of the protein (NMB1566) (Table 3.3) and three with the signal peptide located towards the C-terminus of the

protein (NMB0727, NMB0949 and NMB1447) (Table 3.4). The previously characterised lipoprotein, NMB1468 (Table 3.1) (Hsu *et al*., 2008) and fHbp (Table 3.2) were chosen as positive controls.

3.2.1. Cloning c-Myc tagged gene encoding putative lipoproteins in pGCC4

3.2.1.1. Preparation of insert DNA

Genomic DNA from *N. meningitidis* MC58 was used as template to PCR amplify the 7 selected genes using the primers listed in Table 2.6. A *P*acI (TTAATTAA) restriction site was included in the forward primer and a *P*meI (GTTTAAAC) restriction site was incorporated in the reverse primer downstream of a c-Myc epitope (CTACAGGTCTTCTTCGCTAATCAGTTTCTGTTC) (Figure 3.5). The correct sized band was verified by PCR and agarose gel electrophoresis (section 2.3.4 and 2.3.5) and gene clean was performed (section 2.3.3). A double digest using the enzymes *P*acI and *P*meI was performed on all PCR products as described in section 2.3.6 followed by gene clean (section 2.3.3)

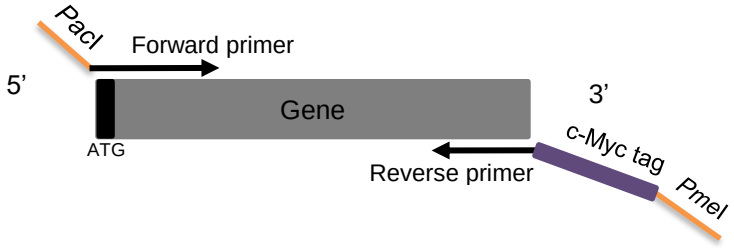


Figure 3.5 Gene encoding putative lipoprotein amplified using forward and reverse primers with the restriction sites *Pac*I and *Pme*I and a c-Myc encoding epitope at 3' end of the gene.

3.2.1.2. Preparation of vector DNA

The pGCC4 vector (Figure 3.6) was digested with *P*acI and *P*meI (section 2.3.6) and then dephosphorylated (section 2.3.7). Ligations of PCR products to pGCC4 were performed, as described in section 2.3.8. The ligation reactions were used to transform DH5α^{TM} cells using the heat shock method outlined in section 2.4.1 and colonies were verified by PCR as stated in section 2.4.3. Each plasmid clone was further verified by sequencing as outlined in section 2.4.3. The sequencing data

confirmed the correct sequence for each clone (Appendix 2). Once colonies with correct band size were verified, each plasmid construct (Figure 3.7) was transformed into strains MC58 and MC58Lnt as described in section 2.5.1.

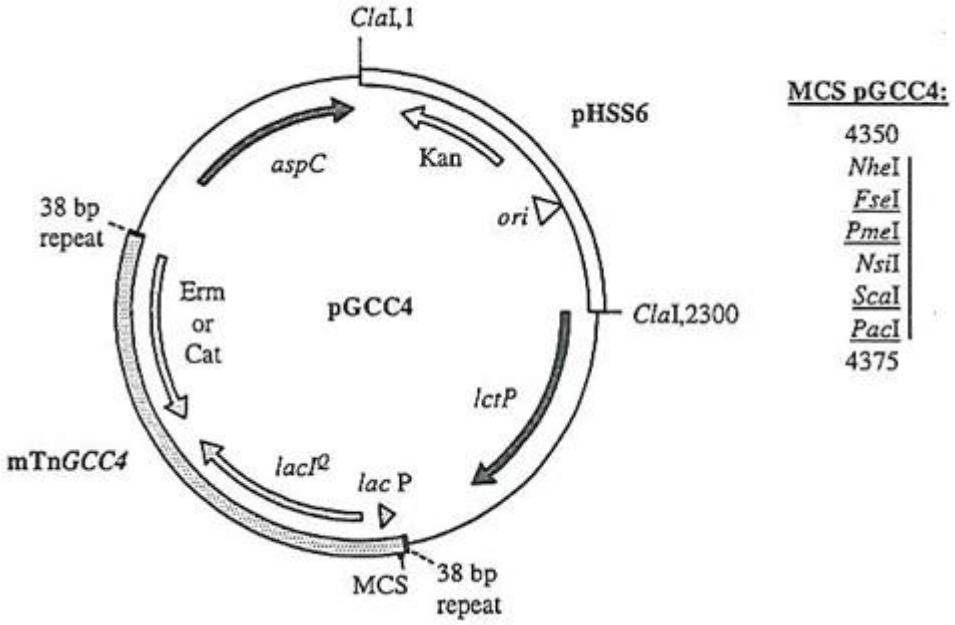


Figure 3.6 Map of *Neisseria* complementing vector, pGCC4 (https://www.addgene.org/37058/). The above map shows the *aspC* and *lctP* genes which flank the antibiotic resistance marker and multiple cloning site (MCS). The gene of interest is cloned into the MCS downstream of the IPTG- inducible *lac*Z promotor. Following transformation into *N. meningitidis,* homologous recombination occurs between the *aspC* and *lctP* genes permitting integration of the cloned gene in the intergenic sequence between these 2 genes without causing polar effects (Mehr and Seifert, 1998).

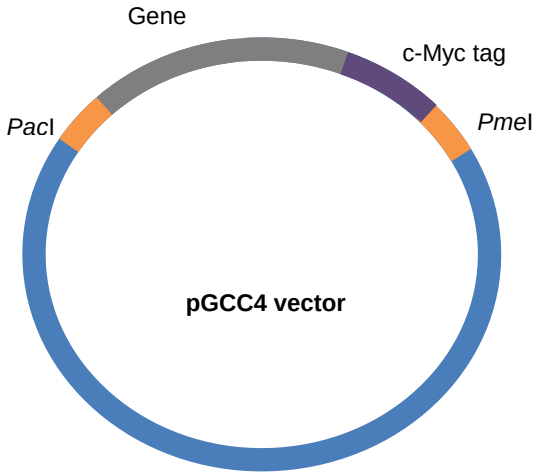


Figure 3.7 Map of recombinant plasmid construct (not drawn to scale). The diagram above shows the gene encoding for putative lipoprotein fused with a c-Myc tag at the 3' end cloned into restrictions site *Pac*I and *Pme*I of pGCC4 plasmid vector.

3.2.1.3. Verification of cloned c-Myc tagged proteins

Each plasmid construct was verified by PCR using pGCC4 primers (Table 2.9). The PCR products were visualised by agarose gel electrophoresis (section 2.3.5). The expected band size was observed for each PCR product (Figure 3.8) and the sequence was confirmed by DNA sequencing, in both directions using the pGCC4 primers (section 2.4.3) (Appendix 2).

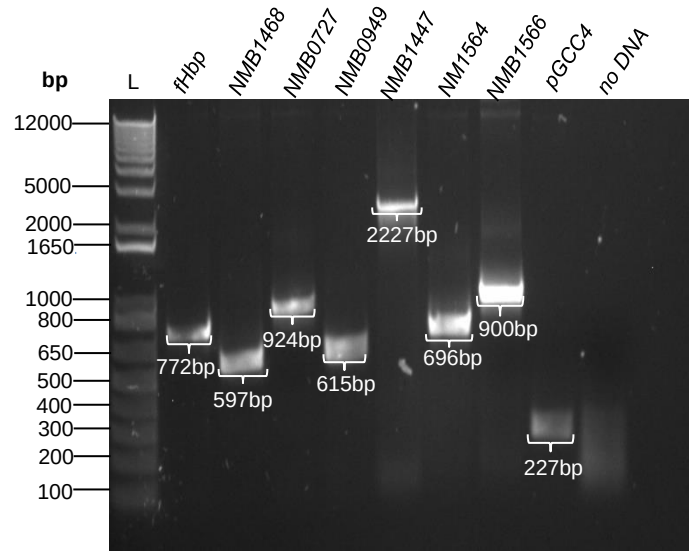


Figure 3.8 PCR products of recombinant plasmids.

3.2.1.4. Verification of c-Myc tagged recombinant strains of MC58 and MC58Lnt

For each recombinant meningococcal strain generated by transformation with pGCC4 containing c-Myc tagged fHbp, NMB1468, NMB0727, NMB0949, NMB1447, NMB1564, NMB1566, 1 colony was selected. Following genomic extraction, PCR was conducted using the pGCC4 primers (Table 2.9) and the PCR product was visualised using agarose gel electrophoresis (section 2.3.5). The expected band size for each recombinant clone of strain MC58 (Figure 3.9a) and MC58Lnt (Figure3.9b) was observed.

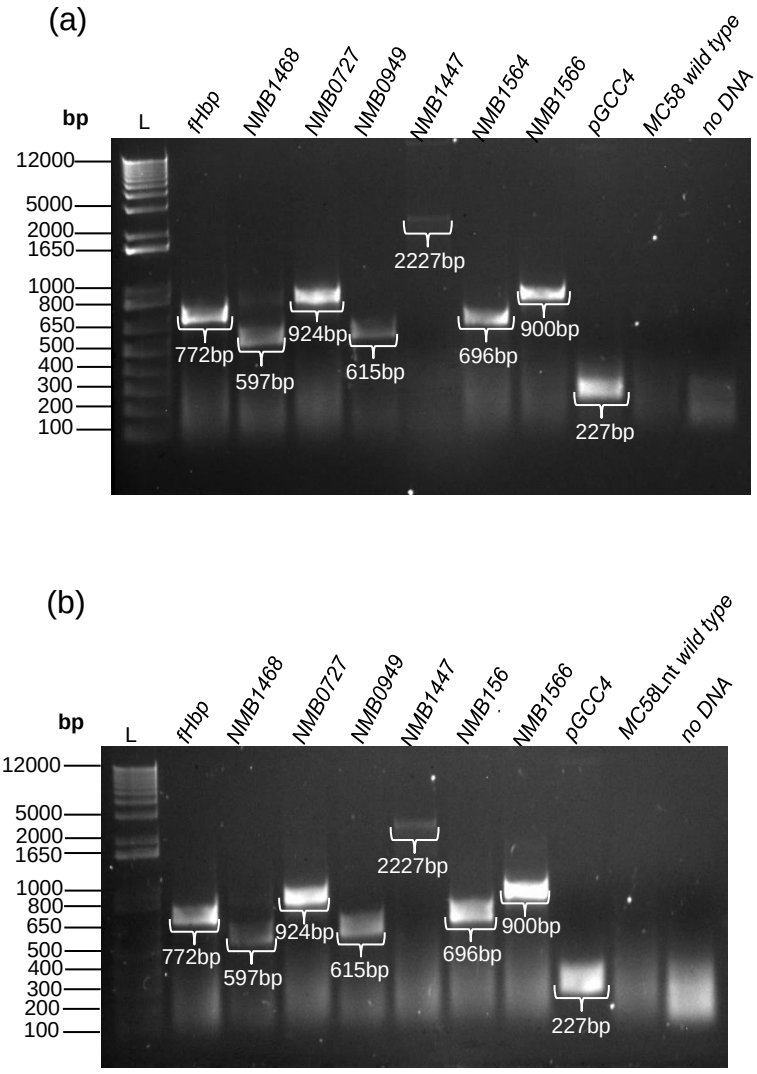


Figure 3.9 Agarose gel electrophoresis of PCR products of recombinant strains of MC58 (a) and MC58Lnt (b).

DNA sequencing with pGCC4 forward and reverse primers confirmed the expected DNA sequence (Appendix 2).

3.2.2. Investigation of expression c-Myc tagged NMB1468 in MC58 and MC58Lnt

Whole cell lysates of broth cultures A_{600} 0.1 of recombinant strains of MC58 and MC58Lnt were fractionated by 16% (w/v) SDS-PAGE, transferred to a PVDF membrane and probed with an anti-c-Myc antibody. Non-transformed strains, MC58 and MC58Lnt were used as negative controls for all Western blots.

The c-Myc tagged protein, NMB1468 (positive control) has a molecular weight of 12.9 kDa and once cleaved, 10.3 kDa. A band of approximately 10 kDa was observed for both recombinant strains and recombinant MC58Lnt showed a band with reduced intensity (Figure 3.10). c-Myc tagged NMB1468 of MC58Lnt migrated slightly further in comparison to MC58 as shown in Figure 3.10. These results support previous observations for His tagged fHbp recombinant strains of MC58 and MC58Lnt from which acylation by Lnt is inferred (da Silva *et al*., 2016).

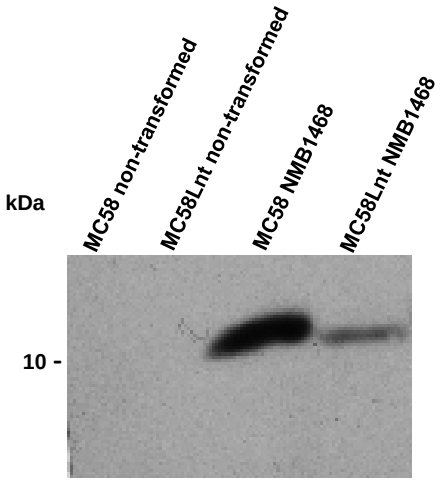


Figure 3.10 Western blot of whole cell lysates of c-Myc tagged NMB1468 recombinant strains of MC58 and MC58Lnt and non-transformed negative controls with anti-c-Myc antibody.

3.2.3. Immunofluorescence microscopy of c-Myc tagged NMB1468 in MC58 and MC58Lnt

In order to verify cell surface expression of c-Myc tagged NMB1468 in MC58 and MC58Lnt, immunofluorescent microscopy of whole cells was performed. Images were captured with the Nikon ECLIPSE 80*i* Microscope using the appropriate filter with 20X magnification (Figure 3.11).

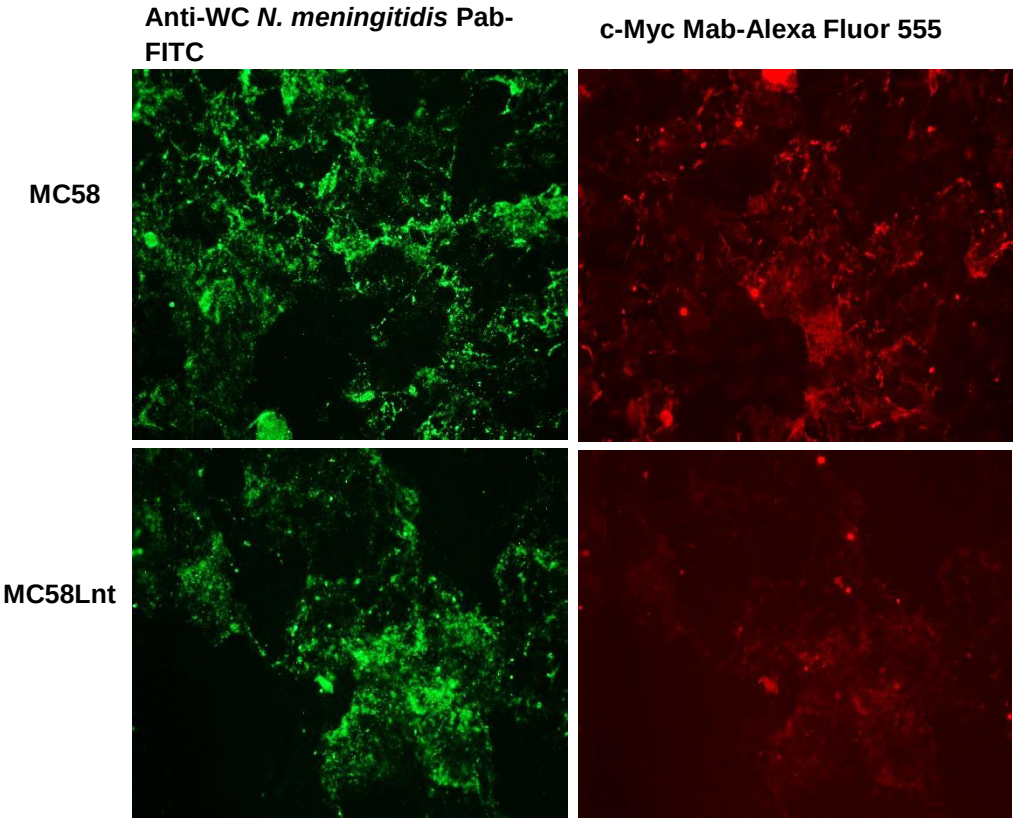


Figure 3.11 Immunofluorescence microscopy of whole cells of strain MC58 and MC58Lnt expressing c-Myc tagged NMB1468.

To confirm the presence of meningococcal cells, cells of both MC58 and MC58Lnt were incubated with FITC-labelled rabbit polyclonal IgG raised against WC *N. meningitidis* as shown in Figure 3.11 (Left and right panel). The cell surface expression of c-Myc tagged recombinant NMB1468 in both strains was compared by also using anti-c-Myc antibody (right panel). MC58Lnt cells showed reduced level of expression of c-Myc tagged NMB1468 compared to cells of strain MC58.

This provides evidence that the disruption in the *lnt* gene results in reduced exportation of diacylated NMB1468 to the cell surface as previously demonstrated for His-tagged fHbp (da Silva *et al*., 2016).

3.2.4. Investigation of expression of the other c-Myc tagged putative lipoproteins in MC58 and MC58Lnt

Whole cell lysates of recombinant strains of MC58 and MC58Lnt were fractionated by 16% (w/v) SDS-PAGE, transferred to a PVDF membrane and probed with anti-c-Myc antibody. Non-transformed strains, MC58 and MC58Lnt were used as negative controls for all Western blots.

The predicted molecular weight (Mw) of both the full length c-Myc tagged proteins and of the cleaved portion (from the cysteine at the lipobox, to the end of the c-Myc epitope) are shown in Table 3.12.

Table 3.12 Expected molecular weight of c-Myc tagged proteins in the absence and presence of cleavage of the signal peptide, as predicted by ExPASy Compute pi/Mw (section 2.6.2).

NMB no.	Expected Mw of full protein sequence (Da)	kDa	Expected Mw if cleaved (Da)	kDa
fHbp	35471.15	**35.5**	28149.38	**28.0**
NMB0949	14334.10	**14.3**	2451.86	**2.4**
NMB0727	26583.23	**26.6**	5258.86	**5.4**
NMB1564	17839.24	**17.8**	11123.77	**11.1**
NMB1477	76801.51	**76.8**	13042.78	**13.0**
NMB1566	23636.10	**23.6**	9348.57	**9.3**

MC58 and MC58Lnt expressing c-Myc tagged fHbp were used as a second positive control, but expression of recombinant fHbp was not detected by Western blotting with anti-c-Myc antibody. Western blots for both MC58 and MC58Lnt strains expressing c-Myc tagged recombinant, NMB0727, NMB1447, NMB1564 and NMB1566 showed bands corresponding in size to that expected for full-length proteins (Figure 3.13a). This suggests that signal peptide cleavage has not occurred.

The expression of c-Myc tagged NMB0949 was not observed in the Western blot (Figure 3.13b). For this recombinant protein, once the signal peptide is cleaved, the resulting protein would give a molecular weight of 2.4 kDa which would be too small for detection by Western blotting. As this recombinant protein was not detected, we predicted that either this is protein was not being expressed or that the signal peptide was being cleaved at the C-terminus.

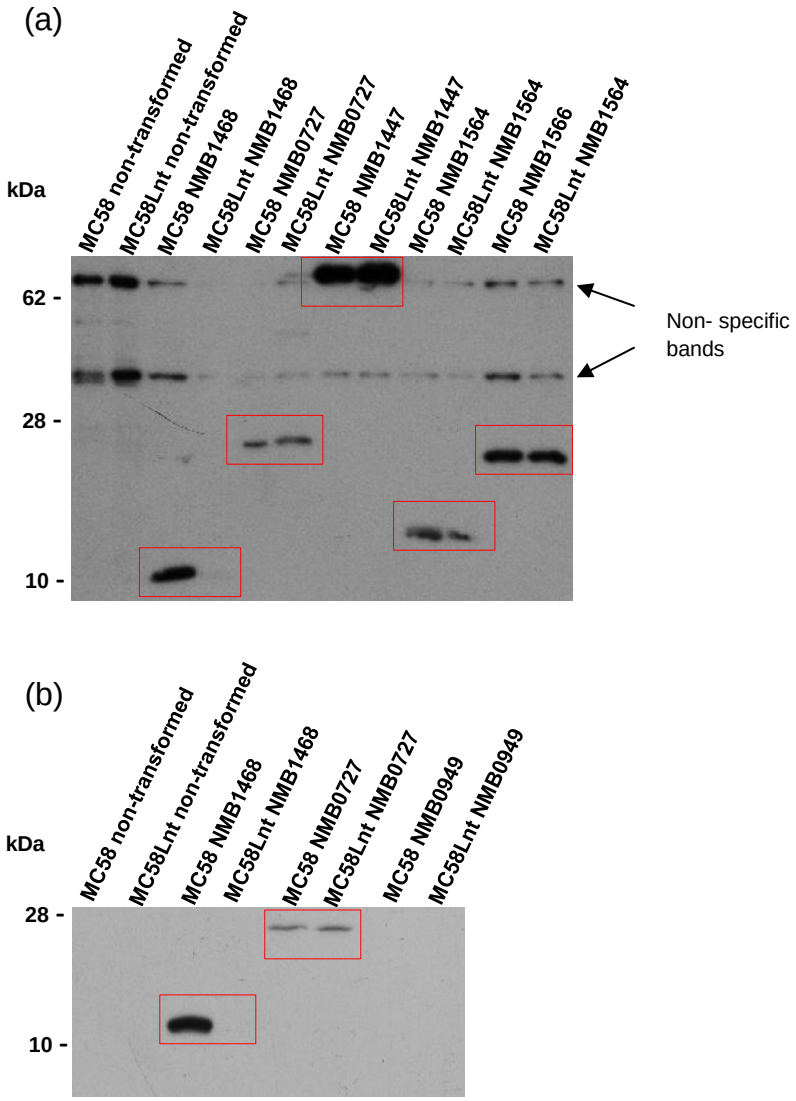


Figure 3.13 Western blot of whole cell lysates of recombinant MC58 and MC58Lnt proteins (a) NMB0727, NMB1447, NMB1564 and NMB1566 with anti-c-Myc antibody. The blot includes non-transformed negative controls and recombinant protein NMB1468. Figure (b) Western blot of recombinant proteins NMB0727 and NMB0949.

3.3. Further investigation of expression and processing of c-Myc NMB0949 in MC58 and MC58Lnt

In order to increase the size of the potentially cleaved C-terminal domain of NMB0949 to allow its detection by Western blot, a larger tag was fused to the end of this gene and that is the GFP reporter gene. In addition, an N-terminal tag was incorporated to allow detection of both portions of the protein. Specifically, a hexa-histidine (His) tag was incorporated at to the 5' end of NMB0949 gene and the GFP encoding gene was fused to the 3' end of the gene by cloning NMB0949 into plasmid pRSET-EmGFP. If processed and cleavage occurred at the lipobox of the signal peptide, this would generate 2 products: the N-terminal portion being 107 amino acids in length (with expected molecular weight of 12.7 kDa as predicted by ExPASy) and the C-terminal portion being 250 amino acids long (with expected molecular weight of 29.0 kDa as predicted by ExPASy). These would be detected by anti-His and anti-GFP antibodies respectively (Figure 3.14).

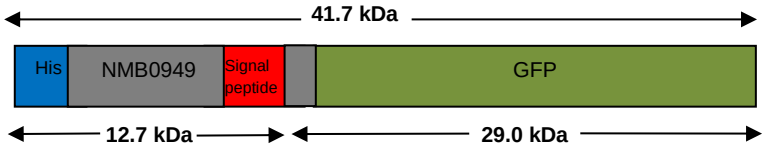


Figure 3.14 Systematic diagram of His-NMB0949-GFP protein (not drawn to scale). Diagram shows the molecular weight of the fusion protein and the molecular weight of the two products generated if the signal peptide is cleaved.

3.3.1. Cloning of N-terminal His tagged, C-terminal GFP tagged NMB0949 into pGCC4 (pGCC4-His-NMB0949-GFP)

3.3.1.1. First step cloning of N-terminal His tagged NMB0949 into pRSET-EmGFP vector

3.3.1.1.1. Preparation of insert DNA

The His-NMB0949 region was PCR amplified using genomic DNA of strain MC58 (section 2.3.4) and using the primers listed in Table 2.6. Primers were designed by incorporating a *B*amHI restriction site (GGATCC) in the forward primer followed by

the initiation codon and the His epitope (CACCACCACCACCACCAC). The reverse primer was designed with the restriction site *E*coRI (GAATTC) immediately downstream of the gene of interest and in place of the stop codon of NMB0949 (Figure 3.15). PCR products were visualised by agarose gel electrophoresis (section 2.3.5) (Figure 3.18). The expected band size of 1,100 bp was observed and the PCR product was gene cleaned (section 2.3.3). Restriction digest using the enzymes *B*amHI and *E*coRI was performed on the PCR product as outlined in section 2.3.6, followed by gene clean (section 2.3.3).

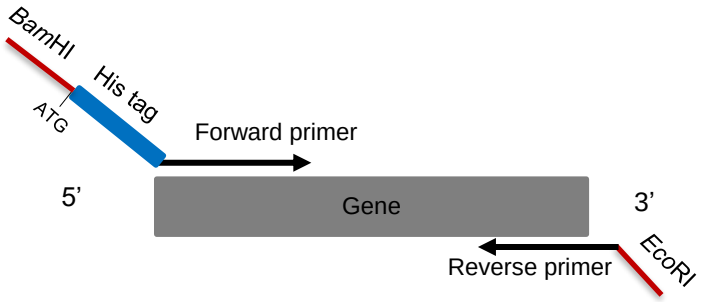


Figure 3.15 The NMB0949 gene amplified using forward and reverse primers with the restriction sites *Bam*HI and *Eco*RI respectively and a His encoding epitope at the 5' end of the gene.

3.3.1.1.2. Preparation of vector DNA

The pRSET-EmGFP plasmid (Figure 3.16) was digested using the enzymes *B*amHI and *E*coRI and dephosphorylated (section 2.3.6 and 2.3.7). The digested PCR products and plasmid vector were ligated (section 2.3.8) and the resulting ligation reaction was used to transform into DH5α™ cells (section 2.4.1) (Figure 3.17). Transformants were isolated following selection on ampicillin containing plates (section 2.4.2) and plasmid DNA was extracted as described in section 2.3.2.

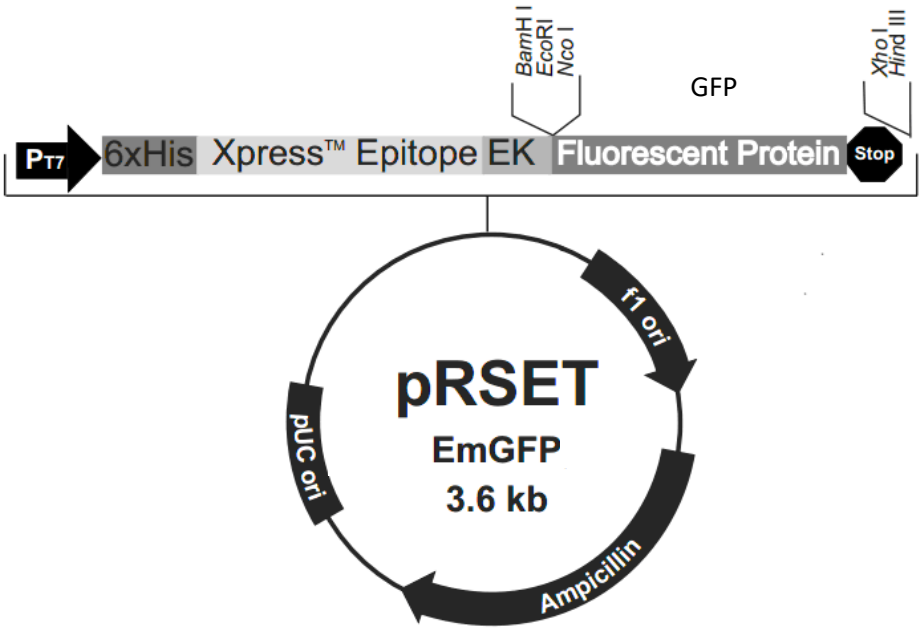


Figure 3.16 Map of pRSET-EmGFP plasmid vector (Invitrogen™). DNA fragments were cloned into *B*amHI and *E*coRI sites (https://tools.thermofisher.com/content/sfs/vectors/prsetfp_map.pdf).

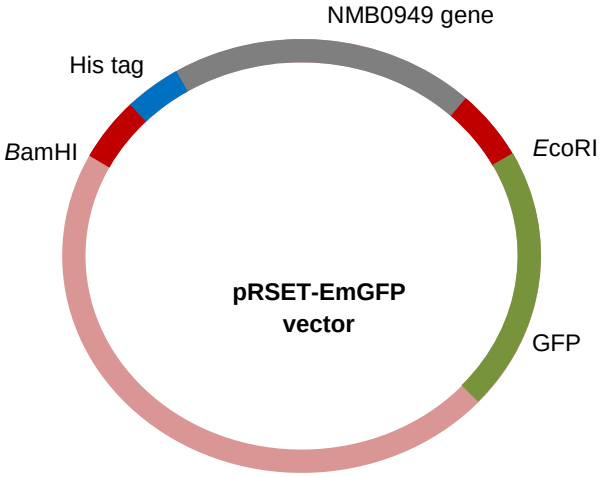


Figure 3.17 Map of recombinant plasmid construct (not drawn to scale). The diagram above shows the gene encoding NMB0949 fused with a His epitope at the 5' end, cloned into restrictions site *B*amHI and *E*coRI of pRSET-EmGFP plasmid vector.

3.3.1.1.3. Verification of pRSET-EmGFP vector His-NMB0949 clones

To confirm the presence of the cloned insert in pRSET-EmGFP, 2 transformant colonies were isolated. Following plasmid DNA extraction (section 2.3.2) PCR was performed with primers that anneal at the 5' and 3' end of the His-NMB0949-GFP (Table 2.6). The correct sized band of 1,100 bp was generated from each transformant (Figure 3.18).

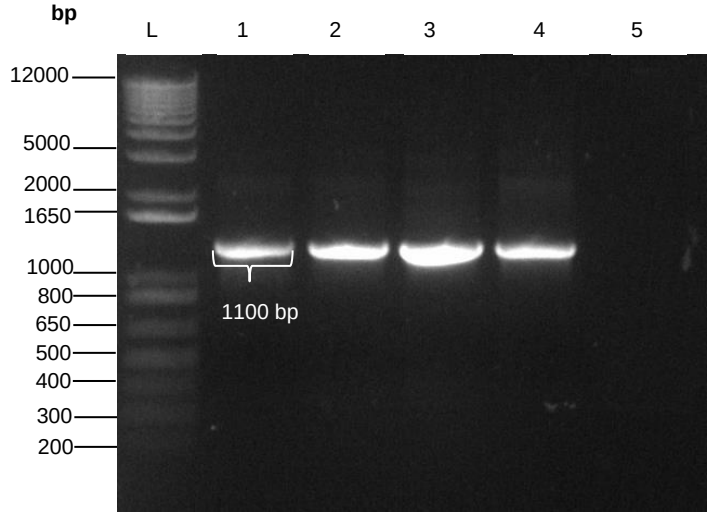


Figure 3.18 Verification of pRSET-EmGFP-His-NMB0949 clones by PCR amplification. Lanes 1 and 2, clone 1 and 2 respectively, with PCR annealing temperature of 40°C, lane 3 and 4, clone 1 and 2 respectively with PCR annealing temperature of 45°C. Lane 5, no DNA control PCR.

3.3.1.2. Second step cloning of N-terminal His tagged, C-terminal GFP tagged NMB0949 into pGCC4 vector

The His-NMB0949-GFP fragment was cloned into the *Pac*I-*Pme*I sites of pGCC4. Briefly, plasmid DNA of pRSET-EmGFP-His-NMB0949 was used as template in PCR with forward primer, incorporating a *P*acI site, that annealed to the 5' end of the His-NMB0949-GFP fragment and with reverse primer incorporating a *P*meI site that anneals to the 3' end of the His-NMB0949-GFP fragment. The PCR product

generated was digested with *Pac*I and *Pme*I then cloned into the *P*acI, *P*meI cut pGCC4 (section 2.3) (Figure 3.19).

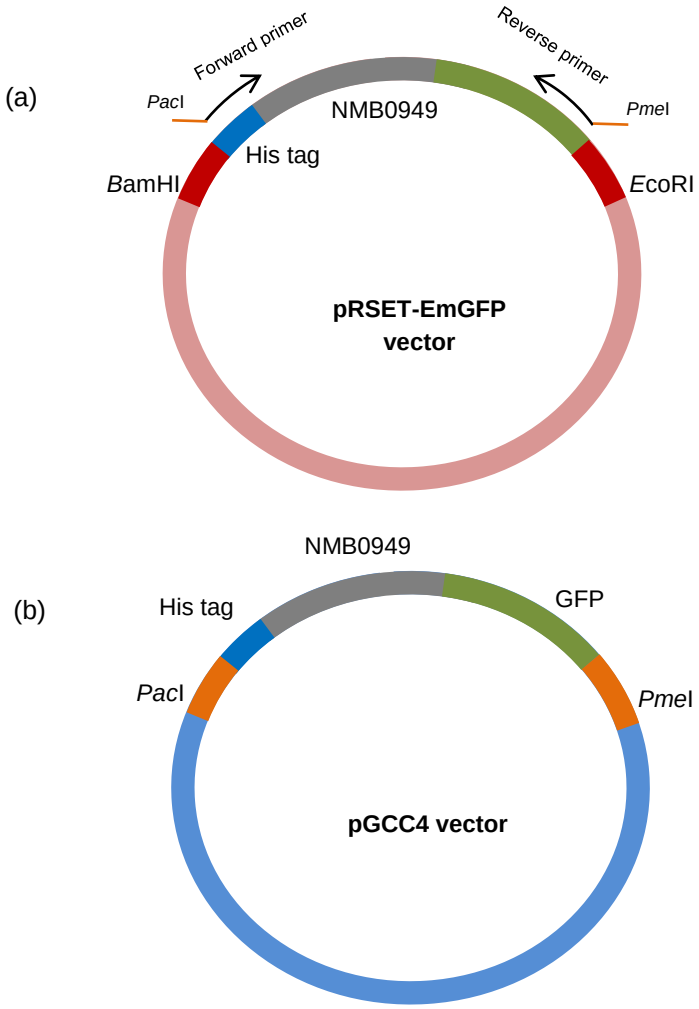


Figure 3.19 Map of recombinant plasmid constructs. The diagram above (a) shows the pRSET-EmGFP-His-NMB0949 plasmid DNA used as template to amplify the His-NMB0949-GFP fragment with primers incorporating with *Pac*I and *Pme*I restriction sites and (b) recombinant plasmid pGCC4 with the insert His-NMB0949-GFP cloned into the *Pac*I and *Pme*I restriction sites.

Following transformation of DH5α^{TM} cells, transformants were selected by growth on erythromycin plates.

3.3.1.2.1. Verification of pGCC4-His-NMB0949-GFP clones

Nine transformant colonies were isolated and plasmid DNA extracted (section 2.3.2) and used as a template for PCR amplification with pGCC4 specific primers (Table 2.9). The expected band size for 4 of these clones was observed by gel

electrophoresis (section 2.3.5) as shown in Figure 3.20. Plasmid 1 was sequenced using pGCC4 primers (Table 2.9) as previously described in section 2.4.3.

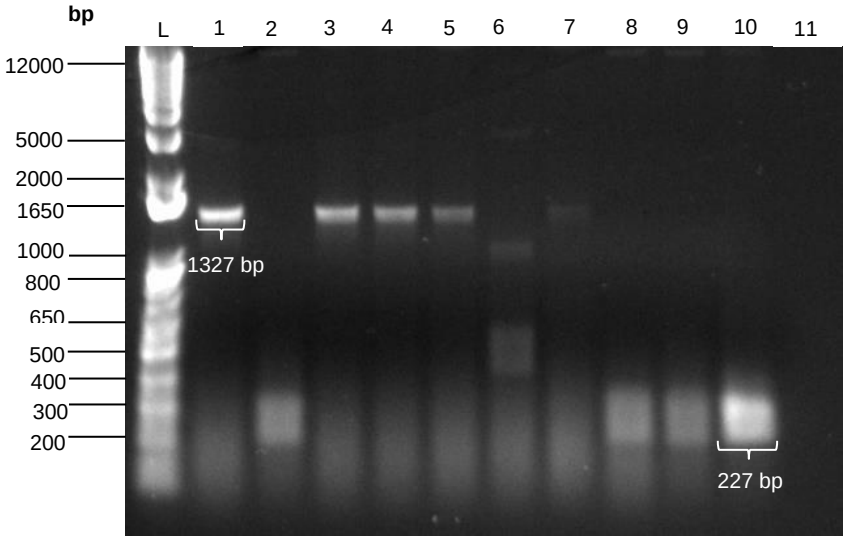


Figure 3.20 PCR products from 9 pGCC4-His-NMB0949-GFP clones using pGCC4 specific primers (lane 1-9). Lane 10 and 11, pGCC4 empty vector and no DNA negative controls respectively.

Following verification by DNA sequencing (Appendix 2), plasmid DNA of clone 1 of pGCC4-His-NMB0949-GFP was used to transform meningococcal strains MC58 and Mc58Lnt and 2 transformants from each strain were selected following growth on erythromycin plates.

3.3.1.3. Verification of MC58 and MC58Lnt transformed with pGCC4-His-NMB0949-GFP

Genomic DNA of these transformants was verified by PCR (section 2.3.4) using pGCC4 primers (Table 2.9).The expected band size of 1,327 bp was observed following agarose gel electrophoresis (section 2.3.5) (Figure 3.21).DNA sequencing with pGCC4 forward and reverse primers confirmed the expected sequence (Appendix 2).

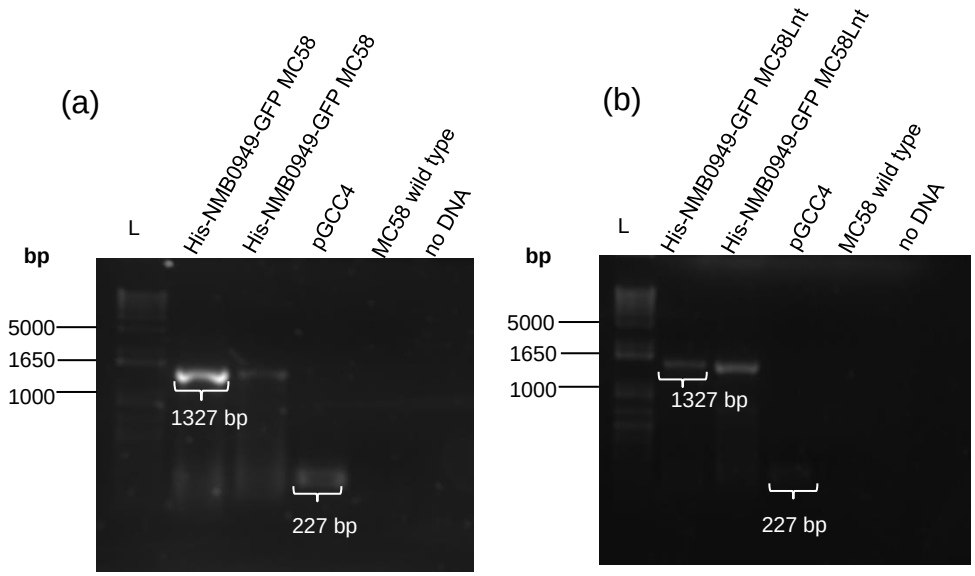


Figure 3.21 PCR products of recombinant clones of MC58 (a) and MC58Lnt (b).

3.3.1.4. Immunofluorescence microscopy of pGCC4-His-NMB0949-GFP in MC58 and MC58Lnt

In order to verify if GFP is expressed in the transformant strains, cells were streaked onto a glass slide and viewed under the Nikon ECLIPSE 80*i* Microscope using the appropriate filter with 10X magnification, for green fluorescence (Figure 3.22).

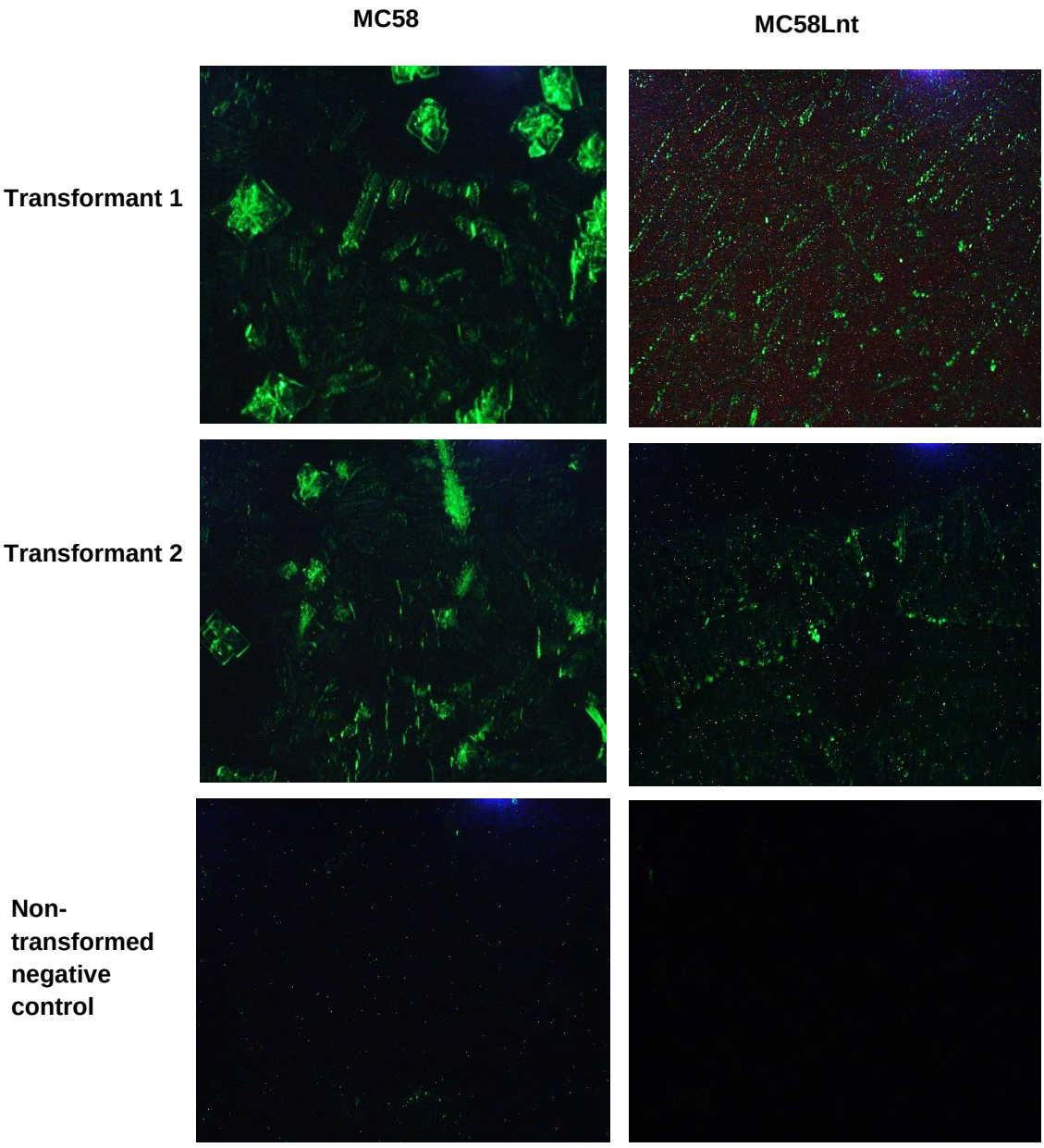


Figure 3.22 Fluorescence of GFP tagged recombinant meningococcal cells.

Green fluorescence was detected in both transformant strains of MC58 and MC58Lnt strains. This indicates that GFP was successfully expressed by these strains. No expression was observed in either of the non-transformed negative controls.

3.3.2. Investigation of expression of His-NMB0949-GFP in MC58 and MC58Lnt

Whole cell lysates of recombinant strains of MC58 and MC58Lnt were fractionated by 16% (w/v) SDS-PAGE as previously described in section 2.6.1 and transferred to two separate PVDF membranes, one probed with anti-His antibody and the other with anti-GFP antibody (section 2.6.2)**.** Non-transformed strains, MC58 and MC58Lnt were used as negative controls for all Western blots.

Both Western blots displayed numerous non-specific bands. The expected molecular weight of the non-cleaved fusion protein of 41.7 kDa was not observed in either blot. However, the expected molecular weight of the 2 products (Figure 3.14) generated following cleavage: i.e. 12.7 kDa His fragment and 29.0 kDa GFP fragment were faintly detected as indicated by the arrows in Figure 3.23 (a) and (b) respectively. These bands were absent in the non-transformed negative control strains. Overall, the western blot results are inconclusive and require further investigation.

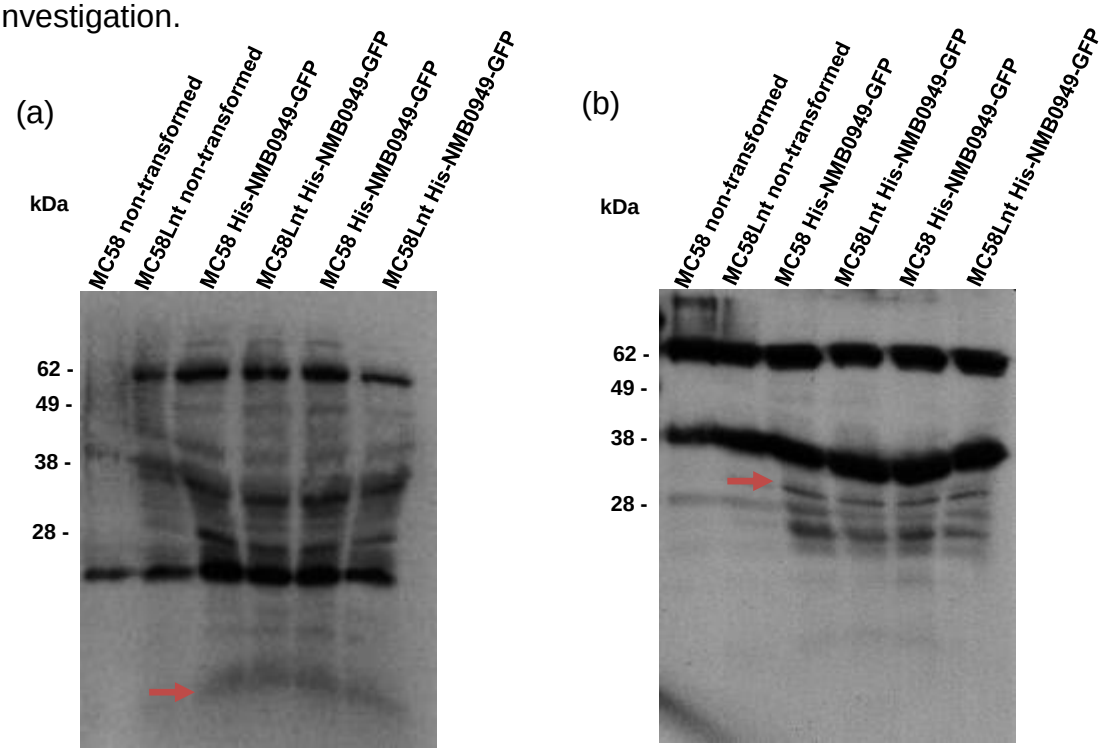


Figure 3.23 Western blot of whole cell lysates of recombinant strains of MC58 and MC58Lnt probed with anti-His (a) and anti-GFP (b) antibodies.

4. Discussion

The last 4 years have seen a breakthrough in prophylactics for meningococcal serogroup B disease with the availability of both Bexsero and Trumenba (CDC, 2016). However, for both vaccines the need remains to increase the breadth of strains these vaccines can target (Lucidarme *et al.,* 2011; McNeil *et al*., 2013; Biagini *et al*., 2016). Lipoproteins are emerging as promising vaccines against bacterial diseases (Barnett *et al*., 2009) and more and more lipoproteins are being identified such that the lipoproteome of bacteria is larger than originally estimated from genome annotations. Identifying surface exposed lipoproteins that elicit potent protective antibodies could be key in developing novel or improved vaccines. To this end the lipoproteome of *N. meningitidis* was investigated.

The vaccine antigen fHbp is a lipoprotein expressed by most meningococcal strains and binds hFH, which allows the bacterium to evade the host immune system due to the down regulation of the alternative complement pathway (Murthy *et al*., 2009). FHbp of MC58 (NMB1870) is annotated in NCBI as a hypothetical protein 320 amino acids long. Using the predictive algorithm tool in the DOLOP website, the fHbp signal peptide is 26 amino acids long (**MTRSKPVN R TAFCCLSLTTALI LTAC**) and is positioned 40 amino acids downstream of the annotated translation start residue (Table 3.2) (generating a protein of 255 amino acids following cleavage of this signal peptide (da Silva *et al*., 2016). This observation, led us to speculate that there may be other lipoproteins that have been missed in MC58 due to the incorrectly annotated start residue.

In this study, we have used the predictive algorithm tool in the DOLOP website to further investigate the number of lipoproteins predicted in the MC58 genome. The first set of lipoproteins identified (Table 3.1) was obtained by screening all 2,119 protein sequences of MC58, in the NCBI database. Previously, Babu *et al.* (2006) identified 69 lipoproteins out of the 2,079 proteins that they analysed within MC58. In this study, 79 lipoproteins were found. This represents approximately 3.7% of the MC58 proteome. The 10 additional probable lipoproteins identified range in their functionality as annotated in the NCBI database with just one of these, NMB1592, annotated as a lipoprotein. Babu *et al*. (2002) have previously recorded different classifications of lipoproteins. As seen with this study, some of these classifications include enzymes, antigens, structural proteins and hypothetical proteins. The current literature reveals that these lipoproteins have not been

characterised or recognised as lipoproteins before. Three of these probable lipoproteins NMB0430, NMB1057, NMB1060 have been assigned functions in different metabolic pathways and encode for the enzymes 2-methylisocitrate lyase (prpB*),* gamma-glutamyltranspeptidase (ggt*)* and fructose -1, 6-bisphosphatase (fbp*)* respectively. The lipoprotein potential of these 3 annotated proteins may contribute to membrane anchorage either on the cytoplasmic membrane or the outer membrane and therefore aids their function. NMB1057 (*ggt*) has been previously suggested as a surface exposed antigen, however this also requires further investigation (Christodoulides, 2014). The 3 probable hypothetical proteins NMB1084, NMB1410, NMB1523 are of great interest as these proteins remain a mystery with no assigned function. According to the +2 rule, these probable hypothetical proteins are associated with the outer membrane. Moreover, 2 of the probable lipoproteins NMB0725 and NMB1991 are annotated as iron associated proteins. NMB0725 is a bacterioferritin-associated ferrodoxin which is thought to be involved in iron storage and NMB1991, also a membrane enzyme, is a transporter of iron (Garmory *et al*., 2004; Tordello *et* al., 2012). Finally, NMB1969 encodes for serotype-1-specific antigen which is an OMP and NMB1594 which is annotated as a lipoprotein which requires further experimental verification.

Of the 79 probably lipoproteins now identified, as many as 37 are annotated as hypothetical proteins in NCBI including the characterised lipoprotein, NMB1468 and fHbp (NMB1870) (Ferrari *et al.,* 2006; Hsu *et al*., 2008). The need remains to experimentally confirm that these hypothetical proteins are lipoproteins and to test the significance of their lipid moieties in their functional or structural roles.

In order to identify any putative lipoproteins, like that of fHbp with the signal peptide located downstream of the predicted start residue, all possible combinations of the lipobox were used as query sequence in a BLASTp analysis and signal peptides were then manually searched for. Thirty-one additional putative lipoproteins were identified containing signal peptides and these were grouped according to their position. The first group of 13 proteins contained signal peptides within the first 30% of the protein sequence from the annotated start residue (Table 3.2). Two of these proteins, NMB0798 and NMB2064, contained a signal peptide of 2 possible different lengths and NMB1538, contained a signal peptide of 3 possible different lengths (Table 3.2). The second group of 3 proteins, contained signal peptides in the middle of the protein sequence (Table 3.3).

Finally, the last group of 15 proteins revealed signal peptides towards the C-terminus of the protein. Amongst these, NMB1151 contained a signal peptide of 2 possible different lengths and NMB1996 carried a signal peptide of 3 possible different lengths (Table 3.4).

To test if these internally positioned signal peptides are functional, five proteins; NMB0727, NMB0949, NMB1447, NMB1564, NMB1566 were chosen for this study as well as fHbp and NMB1468 lipoproteins as positive controls. Selections were made based on the position of the signal peptide. NMB1564 like fHbp has a signal peptide located within the first 30% of the protein sequence, specifically 24 amino acids downstream of the predicted translation start residue and is annotated as a hypothetical protein. NMB1566 contains a signal peptide in the middle of the protein sequence and NMB0727, NMB0949 and NMB1447 have signal peptides located towards the C-terminus. Up until now, signal peptides have only been found at the N-terminus of proteins and studies experimentally testing the function of signal peptides have focused solely on these lipoproteins with the signal peptide conventionally positioned at the N-terminus (Perlman and Halvorson, 1983; Hayashi and Wu, 1990). The presence of internal signal peptides raises two questions. Firstly, have these proteins been annotated with the wrong methionine as the translation start residue, as with fHbp? This in particular may apply to those proteins with signal peptides located within the first 30% of the protein from the annotated start codon. Secondly, can signal peptides be recognised and cleaved if not located at the N-terminus of the protein? In addition, we investigated whether these proteins are lipidated by the enzyme Lnt.

It has been shown that the *lnt* gene is responsible for the addition of a third fatty acid to fHbp (da Silva *et al.,* 2016). A mutation in the *lnt* gene is not lethal in *N. meningitidis* (da Silva *et al*., 2016*)* or in certain other bacteria including *Francisella novicida, Francisella tularensis and Neisseria gonorrhoea (*LoVullo *et al*., 2015*)* unlike in *E. coli* (Chahales and Thanassi, 2015). MC58 with a mutated *lnt* gene can sort diacylated fHbp to the outer membrane for export to the cell surface, albeit inefficiently. Unexpectedly the level of fHbp expression in MC58Lnt strains was decreased by 10-fold likely due to proteolysis of the diacylated lipoproteins that do not get sorted to the outer membrane (da Silva *et al.,* 2016).

All proteins selected for this study were tagged at the C-terminus with a c-Myc epitope and transformed into strains MC58 and MC58Lnt. If the signal peptide is

cleaved, this will be reflected by the change in its molecular weight. There will also be a mobility difference in the recombinant protein expressed in MC58 and MC58Lnt due to their tri- and di-acylated forms respectively. In addition, the diacylated lipoprotein in strain MC58Lnt should show reduced band intensity in comparison to MC58.

In this study, we showed firstly for the positive control c-Myc tagged protein NMB1468, a mobility difference in strain MC58Lnt and a dramatic reduction in quantity of protein (Figure 3.10) as observed for recombinant fHbp in MC58Lnt (da Silva *et al*., 2016). From this we can infer that NMB1468 is triacylated in MC58 by Lnt and is diacylated in MC58Lnt. da Silva *et al*. (2016) speculated that proteolysis is occurring due to the build-up of diacylated lipoproteins that are not sorted to the outer membrane hence causing envelope stress. In response to over-expressed or misfolded proteins in the periplasm which causes envelope stress, two pathways are employed by *E. coli*, σ^E and Cpx pathways (McBroom and Kuehn, 2007). We predict that one or more of these pathways is employed by *N. meninigitidis.* Generally, σ^E is activated in response to misfolded proteins during OMP biogenesis and the Cpx, a two component pathway, is activated by over-expressed proteins accumulating in the cell envelope (Ravio *et al*., 2013). Both pathways result in expression of proteases which then degrade the misfolded/accumulated proteins in the periplasm.

Cell surface expression of the c-Myc tagged lipoprotein NMB1468, was observed in MC58 by immunofluorescence microscopy (Figure 3.11) and this was dramatically reduced in MC58Lnt. This supports previous immunofluorescence microscopy observations for recombinant fHbp expressed in MC58Lnt which inefficiently exports diacylated fHbp to the cell surface (da Silva *et* al., 2016).

As with NMB1468, the c-Myc epitope was tagged to each of the 5 proteins selected for this study. The predicted change in molecular weight of each protein if the signal peptide is cleaved was determined by ExPASy (http://web.expasy.org/compute_pi/). Following Western blotting with anti-c-Myc antibody, it was found that four of these proteins, NMB0727, NMB1447, NMB1564 and NMB1566 were not cleaved and rather expressed the full size protein.

Bacterial lipoproteins are transported in a linear form across the cytoplasmic membrane by the Sec pathway (Auclair *et al*., 2012). In Gram-negative bacteria,

the preprolipoprotein synthesized in the cytoplasm is generally recognised by SecB (chaperone) which prevents folding and delivers the preprolipoprotein to SecA which provides the energy for the protein to be transported across the cytoplasmic membrane (du Plessis *et al*., 2011). Once transferred across the membrane, the conserved cysteine is acylated by the enzyme Lgt which enables protein anchorage to the membrane and then the signal peptide is cleaved by the enzyme Lsp. With no signal peptide cleavage occurring in these 4 proteins, either SecB does not recognise the signal peptide in the cytoplasm, as it is not positioned at the N-terminus or once the preprolipoprotein is translocated and folding begins in the cytoplasmic membrane, the signal peptide is no longer accessible to Lsp, preventing cleavage. However, if the signal peptide is not accessible to Lsp, this would result in the build-up of proteins in the periplasm and we would expect this to result in periplasmic proteolysis and dramatic reduction in the level of protein detected. Given that we observed strong protein expression from our Western blot analysis, we favour the first hypothesis that SecB is not recognising the signal peptide and that the proteins remain in the cytoplasm. Lipoproteins studied to date contain a signal peptide at the N-terminus and most studies have focused on determining which amino acids in the signal peptide are key in cleavage. A signal peptide with a hydrophobic region abundant in alanine residue has been associated with efficient cleavage (de Souza *et al*., 2011).

Interestingly, the expression of NMB0949 in MC58 and MC58Lnt, was not detected by Western blotting when probed with anti-c-Myc antibody (Figure 3.13b). The full protein sequence of NMB0949 in the NCBI database gives a protein of expected molecular weight of 14.3 kDa including the c-Myc tag. If the signal peptide located towards the C-terminus is cleaved, the predicted molecular weight of the cleaved protein is 2.4 kDa. This would be too small to be detected by Western blotting. This led us to speculate that the signal peptide is being cleaved in NMB0949. In order to investigate this further the C-terminus was fused with GFP which is approximately 27.0 kDa. The cleaved lipoprotein fused to GFP would generate a product of about 29.0 kDa which could be detectable by Western blotting. In addition, to detect expression of the N-terminal region, we fused the start of the protein with a His tag which would generate a peptide of almost 13.0 kDa if cleavage occurs. For the non-cleaved form, when probed with anti-His or anti-GFP antibody, the expected molecular weight is about 42.0 kDa. Unfortunately, for both membranes, there was a high degree of non-specific

binding making it difficult to interpret the data clearly. For the membrane probed with the anti-His antibody, diffuse bands around the expected size for the cleaved protein were observed for the recombinant strains (Figure 3.23a). Likewise, for the membrane probed with anti-GFP antibody, faint bands were observed of the size expected following successful cleavage of the signal peptide (Figure 3.23b). It was confirmed by immunofluorescence microscopy of whole cells of both recombinant strains that GFP was indeed expressed by these strains (Figure 3.22). Overall, despite there being some indication that the signal peptide is cleaved, our results are inconclusive. Due to restricted time, further optimisation was not possible.

The predictive algorithm tool used in this study successfully enabled the identification of 10 more probable lipoproteins in MC58 with the signal peptide with conventional localisation of the signal peptide at the N-terminus. These are not known in the literature to be lipoproteins (except for NMB1594 that was annotated as a lipoprotein) and shed important light on these proteins. Thirteen proteins were identified with the signal peptide located downstream of the predicted translation start residue according to the annotated genome. Whilst one of these proteins that was investigated experimentally in this study demonstrated no cleavage of the signal peptide, given that a precedent has been set by fHbp for successful recognition of its downstream, somewhat “internal” signal peptide, the remaining 12 proteins need to be individually tested.

There are a number of limitations with programs that annotate bacterial genomes, in particular in identifying the correct start codon (Goal *et al.*, 2013). Genome annotation programs, in addition to searching for homologous sequences, rely on signal sensors to identify functional sites including promoters, ribosomal binding site, start and stop codons and transcription terminators (Mathé *et* al., 2002; Rust *et al.,* 2002). However, bacterial genes are often organised in operons on the chromosome and these clusters of genes are transcribed by a single promotor positioned at around -10 base pairs upstream of the start codon of the first gene (Trun and Trempy, 2009). Without promoter cues for the downstream genes in the operon, it is difficult to predict which methionine or valine is employed as the translation start residue. For this reason, in the case of fHbp and likely many other proteins, the incorrect amino acid has been predicted to be the translation start residue. This study highlights the importance of experimentally testing each

putative lipoprotein with downstream signal peptide (as opposed to N-terminal signal peptide).

From our pilot study, we could confirm the importance of the signal peptide being positioned at the N-terminus to function successfully that is to be recognised by the Sec apparatus and permit translocation to the periplasm for lipidation and cleavage prior to export to the cell surface in the case of most lipoproteins. If the +2 rule of *E. coli* applies to the meningococcus, from our study, all putative lipoproteins other than HtrA (NMB0532) and DsbA (NMB0278) which have an aspartate residue at this position, get sorted to the outer membrane. Further validation of this rule is required for *N. meningitidis* as it has recently been suggested by Hooda *et al*. (2017) that *N. meninigitidis* strains do not follow the +2 rule as based on previous data they predict HtrA cannot be the only lipoprotein not exported to the cell surface. Furthermore, as shown by Hooda *et al.* (2016), it is likely that many of these probable lipoproteins are flipped by Slam to be exposed at the cell surface like fHbp, LbpB and TbpB, and this can be readily tested by immunofluorescence microscopy as shown in our study for the lipoprotein NMB1468.

Importantly investigating the ability of these previously unknown lipoproteins to engage the immune system and generate a potent and protective antibody response could provide new leads for an improved meningococcal serogroup B vaccine.

5. References

Addgene (2017) *pGCC4 plasmid vector.* Available at https://www.addgene.org/37058/ (Accessed 1 April 2017).

Auclair, S. M., Bhanu, M. K., & Kendall, D. A. (2012). Signal peptidase I: Cleaving the way to mature proteins. *Protein Science*, 21(1), 13-25.

Babu, M.M., & Sankaran, K. (2002). DOLOP--database of bacterial lipoproteins. *Bioinformatics (Oxford, England),* 18(4), 641-643.

Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L., et al. (2006). A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *Journal of Bacteriology*, 188(8), 2761-2773.

Barrett, A. D., & Stanberry, L. R. (2009). Vaccines for biodefense and emerging and neglected diseases Academic Press.

Biagini, M., Spinsanti, M., De Angelis, G., Tomei, S., Ferlenghi, I., Scarselli, M., et al. (2016). Expression of factor H binding protein in meningococcal strains can vary at least 15-fold and is genetically determined. *Proceedings of the National Academy of Sciences of the United States of America*, 113(10), 2714-2719.

BiologicsCrop (2016) *BiologicsCrop.* Available at http://www.biologicscorp.com/blog/iptg-induction-protein-expression/#.WPdbOvnyu1s (Accessed 5 March 2016).

BLASTp (2016) *National Centre for Biotechnology Information.* Available at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (Accessed 10 November 2016).

Brehony, C., Hill, D. M., Lucidarme, J., Borrow, R., & Maiden, M. C. (2015). Meningococcal vaccine antigen diversity in global databases. *Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, 20(49), 10.2807/1560-7917.ES.2015.20.49.30084.

Capecchi, B., Adu-Bobie, J., Di Marcello, F., Ciucchi, L., Masignani, V., Taddei, A., et al. (2005). Neisseria meningitidis NadA is a new invasin which promotes bacterial adhesion to and penetration into human epithelial cells. *Molecular Microbiology*, 55(3), 687-698.

Carbonnelle, E., Hill, D. J., Morand, P., Griffiths, N. J., Bourdoulous, S., Murillo, I., et al. (2009). Meningococcal interactions with the host. *Vaccine*, 27, B78-B89.

CDC (2016) *Centres for Disease Control and Prevention.* Available at https://www.cdc.gov/vaccines/hcp/vis/vis-statements/mening-serogroup.html (Accessed on 10 March 2017).

Chahales, P., & Thanassi, D. G. (2015). A more flexible lipoprotein sorting pathway. *Journal of Bacteriology*, 197(10), 1702-1704.

Clark, S. A., Lekshmi, A., Lucidarme, J., Hao, L., Tsao, H., Lee-Jones, L., et al. (2016). Differences between culture & non-culture confirmed invasive meningococci with a focus on factor H-binding protein distribution. *Journal of Infection*, 73(1), 63-70.

Christodoulides, M. (2014). Neisseria proteomics for antigen discovery and vaccine development. *Expert Review of Proteomics, 11*(5), 573-591.

Crum-Cianflone, N., & Sullivan, E. (2016). Meningococcal vaccinations. *Infectious Diseases and Therapy*, 5(2), 89-112.

De Souza, G. A., Leversen, N. A., Målen, H., & Wiker, H. G. (2011). Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway. *Journal of Proteomics*, 75(2), 502-510.

DLOLP (2016) *A Database of Bacterial Lipoproteins.* Available at http://www.mrc-lmb.cam.ac.uk/genomes/dolop/ (Accessed 12 November 2016).

Du Plessis, D. J., Nouwen, N., & Driessen, A. J. (2011). The sec translocase. *Biochimica Et Biophysica Acta (BBA)-Biomembranes*, 1808(3), 851-865.

Ecdc (2012) *European Centre for Disease Prevention and Control*. Available at http://ecdc.europa.eu/en/publications/Publications/Surveillance%20of%20IBD%20in%20Europe%202012.pdf (Accessed 21 March 2016).

EMA (2015) *European Medicines Agency Science Medicines Health.* Available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Summary_for_the_public/human/002333/WC500137857.pdf (Accessed 21 December 2016)

ExPASy (2017) *Bioinformatics Resources Portal.* Available at http://web.expasy.org/compute_pi/ (Accessed 20 January 2017).

Feavers, I. M., Suker, J., McKenna, A. J., Heath, A. B., & Maiden, M. C. (1992). Molecular analysis of the serotyping antigens of neisseria meningitidis. *Infection and Immunity,* 60(9), 3620-3629.

Frosch, M., Weisgerber, C., & Meyer, T. F. (1989). Molecular characterization and expression in escherichia coli of the gene complex encoding the polysaccharide capsule of neisseria meningitidis group B. *Proceedings of the National Academy of Sciences of the United States of America*, 86(5), 1669-1673.

Gandhi, A., Balmer, P., & York, L. J. (2016). Characteristics of a new meningococcal serogroup B vaccine, bivalent rLP2086 (MenB-FHbp; trumenba®). *Postgraduate Medicine*, 128(6), 548-556.

Goel, N., Singh, S., & Aseri, T. C. (2013). A review of soft computing techniques for gene prediction. *ISRN Genomics*, 2013.

Hayashi, S., & Wu, H. C. (1990). Lipoproteins in bacteria. *Journal of Bioenergetics and Biomembranes*, 22(3), 451-471.

Hill, D. J., Griffiths, N. J., Borodina, E., & Virji, M. (2010). Cellular and molecular biology of neisseria meningitidis colonization and invasive disease. *Clinical Science* (London, England : 1979), 118(9), 547-564.

Holst, J., Feiring, B., Fuglesang, J., Høiby, E., Nøkleby, H., Aaberge, I., et al. (2003). Serum bactericidal activity correlates with the vaccine efficacy of outer membrane vesicle vaccines against neisseria meningitidis serogroup B disease. *Vaccine*, 21(7), 734-737.

Hooda, Y., Lai, C. C., Judd, A., Buckwalter, C. M., Shin, H. E., Gray-Owen, S. D., et al. (2016). Slam is an outer membrane protein that is required for the surface display of lipidated virulence factors in neisseria. *Nature Microbiology*, 1, 16009.

Hooda, Y., Shin, H. E., Bateman, T. J., & Moraes, T. F. (2017). Neisserial surface lipoproteins: Structure, function and biogenesis. *Pathogens and Disease, 75*(2), 10.1093/femspd/ftx010.

Hou, B., Reizis, B., & DeFranco, A. L. (2008). Toll-like receptors activate innate and adaptive immunity by using dendritic cell-intrinsic and -extrinsic mechanisms. *Immunity*, 29(2), 272-282.

Hsu, C., Lin, W., Li, J., Liu, Y., Tseng, Y., Chang, C., et al. (2008). Immunoproteomic identification of the hypothetical protein NMB1468 as a novel lipoprotein ubiquitous in neisseria meningitidis with vaccine potential. *Proteomics*, 8(10), 2115-2125.

Inouye, S., Franceschini, T., Sato, M., Itakura, K., & Inouye, M. (1983). Prolipoprotein signal peptidase of escherichia coli requires a cysteine residue at the cleavage site. *The EMBO Journal*, 2(1), 87-91.

Jarvis, G. A., & Vedros, N. A. (1987). Sialic acid of group B neisseria meningitidis regulates alternative complement pathway activation. *Infection and Immunity*, 55(1), 174-180.

Kang, J. Y., Nan, X., Jin, M. S., Youn, S., Ryu, Y. H., Mah, S., et al. (2009). Recognition of lipopeptide patterns by toll-like receptor 2-toll-like receptor 6 heterodimer. *Immunity*, 31(6), 873-884.

Kellog, D. S.,Jr, Peacock, W. L.,Jr, Deacon, W. E., Brown, L., & Pirckel, D. I. (1963). Neisseria gonorrhoeae. I. virulence genetically linked to clonal variation. *Journal of Bacteriology*, 85, 1274-1279.

Koeberling, O., Delany, I., & Granoff, D. M. (2011). A critical threshold of meningococcal factor H binding protein expression is required for increased breadth of protective antibodies elicited by native outer membrane vesicle vaccines. *Clinical and Vaccine Immunology* : CVI, 18(5), 736-742.

Loh, E., Lavender, H., Tan, F., Tracy, A., & Tang, C. M. (2016). Thermoregulation of meningococcal fHbp, an important virulence factor and vaccine antigen, is mediated by anti-ribosomal binding site sequences in the open reading frame. *PLoS Pathog*, 12(8), e1005794.

Lucidarme, J., Tan, L., Exley, R. M., Findlow, J., Borrow, R., & Tang, C. M. (2011). Characterization of neisseria meningitidis isolates that do not express the virulence factor and vaccine antigen factor H binding protein. *Clinical and Vaccine Immunology* : CVI, 18(6), 1002-1014.

Mathe, C., Sagot, M. F., Schiex, T., & Rouze, P. (2002). Current methods of gene prediction, their strengths and weaknesses. *Nucleic Acids Research*, 30(19), 4103-4117.

Matsuyama, S., Tajima, T., & Tokuda, H. (1995). A novel periplasmic carrier protein involved in the sorting and transport of escherichia coli lipoproteins destined for the outer membrane. *The EMBO Journal*, 14(14), 3365-3372.

McBroom, A. J., & Kuehn, M. J. (2007). Release of outer membrane vesicles by Gram-negative bacteria is a novel envelope stress response. *Molecular Microbiology*, 63(2), 545-558.

McNeil, L. K., Zagursky, R. J., Lin, S. L., Murphy, E., Zlotnick, G. W., Hoiseth, S. K., et al. (2013). Role of factor H binding protein in neisseria meningitidis virulence and its potential as a vaccine candidate to broadly protect against meningococcal disease. *Microbiology and Molecular Biology Reviews: MMBR*, 77(2), 234-252.

MRF (2017) *Meningitidis Research Foundation*. Available at http://www.meningitis.org/menb (Accessed 28 March 2017).

Murphy, E., Andrew, L., Lee, K. L., Dilts, D. A., Nunez, L., Fink, P. S., et al. (2009). Sequence diversity of the factor H binding protein vaccine candidate in epidemiologically relevant strains of serogroup B neisseria meningitidis. *The Journal of Infectious Diseases*, 200(3), 379-389.

Nakayama, H., Kurokawa, K., & Lee, B. L. (2012). Lipoproteins in bacteria: Structures and biosynthetic pathways. *FEBS Journal*, 279(23), 4247-4268.

Narita, S., & Tokuda, H. (2007). Amino acids at positions 3 and 4 determine the membrane specificity of pseudomonas aeruginosa lipoproteins. *The Journal of Biological Chemistry*, 282(18), 13372-13378.

NCBI (2016) *National Centre for Biotechnology Information.* Available at https://www.ncbi.nlm.nih.gov/ (Accessed 10 November 2016).

Omotajo, D., Tate, T., Cho, H., & Choudhary, M. (2015). Distribution and diversity of ribosome binding sites in prokaryotic genomes. *BMC Genomics*, 16(1), 604.

Oriente, F., Scarlato, V., & Delany, I. (2010). Expression of factor H binding protein of meningococcus responds to oxygen limitation through a dedicated FNR-regulated promoter. *Journal of Bacteriology*, 192(3), 691-701.

Oxford Genetics Biology Engineered (2016) *Oxford Genetics Biology Engineered.* Available at https://www.oxfordgenetics.com/SiteContent/TeamResources/bacterial-promoter-information (Accessed 20 April 2017).

Pace, D., & Pollard, A. J. (2012). Meningococcal disease: Clinical presentation and sequelae. *Vaccine*, 30, B3-B9.

Perlman, D., & Halvorson, H. O. (1983). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *Journal of Molecular Biology*, 167(2), 391-409.

Pfizer (2014) *Pfizer.* Available at http://www.pfizer.com/news/press-release/press-release-detail/trumenba_meningococcal_group_b_vaccine_is_now_available (Accessed 20 December 2016).

PHE (2017) *Public Health England.* Available at https://www.gov.uk/government/collections/meningococcal-disease-guidance-data-and-analysis (Accessed 24 March).

QIAGEN (2017) *Gentra Purgene Yeast/Bact. Kit.* Available at https://www.qiagen.com/us/shop/sample-technologies/dna/genomic-dna/gentra-puregene-yeastbact-kit/#orderinginformation (Accessed 14 March 2017).

QIAGEN (2017) *MiniElute PCR Purification Kit.* Available at https://www.qiagen.com/gb/shop/sample-technologies/dna/genomic-dna/minelute-pcr-purification-kit/#orderinginformation (Accessed 14 March 2017).

QIAGEN (2017) *QIAprep Spin Miniprep Kit.* Available at https://www.qiagen.com/us/shop/sample-technologies/dna/plasmid-dna/qiaprep-spin-miniprep-kit/#orderinginformation (Accessed 14 March 2017).

Raivio, T. L., Leblanc, S. K., & Price, N. L. (2013). The escherichia coli cpx envelope stress response regulates genes of diverse function that impact antibiotic resistance and membrane integrity. *Journal of Bacteriology*, 195(12), 2755-2767.

Rouphael, N. G., & Stephens, D. S. (2012). Neisseria meningitidis: Biology, microbiology, and epidemiology. Neisseria Meningitidis: *Advanced Methods and Protocols*, 1-20.

Rust, A. G., Mongin, E., & Birney, E. (2002). Genome annotation techniques: New approaches and challenges. *Drug Discovery Today*, 7(11), S70-S76.

Sanders, H., Brehony, C., Maiden, M. C., Vipond, C., & Feavers, I. M. (2012). The effect of iron availability on transcription of the neisseria meningitidis fHbp gene varies among clonal complexes. *Microbiology*, 158(4), 869-876.

Sankaran, K., & Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. transfer of diacylglyceryl moiety from phosphatidylglycerol. *The Journal of Biological Chemistry*, 269(31), 19701-19706.

Serruto, D., Bottomley, M. J., Ram, S., Giuliani, M. M., & Rappuoli, R. (2012). The new multicomponent vaccine against meningococcal serogroup B, 4CMenB: Immunological, functional and structural characterization of the antigens. *Vaccine*, 30, B87-B97.

Silva, R., Churchward, C., Karlyshev, A., Eleftheriadou, O., Snabaitis, A., Longman, M., et al. (2016). The role of apolipoprotein N-acyl transferase, lnt, in the lipidation of factor H binding protein of neisseria meningitidis strain MC58 and its potential as a drug target. *British Journal of Pharmacology*.

Stephens DS, Apicella MA. Neisseria meningitidis. In: Mandell, Douglas, and Bennett's principles and practice of infectious diseases. 8th ed. Philadelphia, PA *Elsevier Saunders Inc*.; 2015. p. 2425-45.

Stephens, D. S., & McGee, Z. A. (1981). Attachment of neisseria meningitidis to human mucosal surfaces: Influence of pili and type of receptor cell. *The Journal of Infectious Diseases*, 143(4), 525-532.

Swartley, J. S., Marfin, A. A., Edupuganti, S., Liu, L. J., Cieslak, P., Perkins, B., et al. (1997). Capsule switching of neisseria meningitidis. *Proceedings of the National Academy of Sciences of the United States of America*, 94(1), 271-276.

Tanaka, K., Matsuyama, S. I., & Tokuda, H. (2001). Deletion of lolB, encoding an outer membrane lipoprotein, is lethal for escherichia coli and causes accumulation of lipoprotein localization intermediates in the periplasm. *Journal of Bacteriology*, 183(22), 6538-6542.

Tang, M., Diao, J., Gu, H., Khatri, I., Zhao, J., & Cattral, M. S. (2015). Toll-like receptor 2 activation promotes tumor dendritic cell dysfunction by regulating IL-6 and IL-10 receptor signaling. *Cell Reports*, 13(12), 2851-2864.

Tautz, D. (2009). Polycistronic peptide coding genes in eukaryotes--how widespread are they? *Briefings in Functional Genomics & Proteomics*, 8(1), 68-74.

Tinsley, C. R., Voulhoux, R., Beretti, J. L., Tommassen, J., & Nassif, X. (2004). Three homologues, including two membrane-bound proteins, of the disulfide oxidoreductase DsbA in neisseria meningitidis: Effects on bacterial growth and biogenesis of functional type IV pili. *The Journal of Biological Chemistry, 279*(26), 27078-27087.

ThermoFisher SCIENTIFIC (2017) *pRSET-EmGFP Bacterial Expression Vector.* Available at (Accessed 14 April).

Tokunaga, M., Tokunaga, H., & Wu, H. C. (1982). Post-translational modification and processing of escherichia coli prolipoprotein in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 79(7), 2255-2259.

Del Tordello, E., Bottini, S., Muzzi, A., & Serruto, D. (2012). Analysis of the regulated transcriptome of neisseria meningitidis in human blood using a tiling array. *Journal of Bacteriology, 194*(22), 6217-6232.

Trun, N., & Trempy, J. (2009). Fundamental bacterial genetics John Wiley & Sons.

Virji, M., Makepeace, K., Ferguson, D. J., Achtman, M., & Moxon, E. R. (1993). Meningococcal opa and opc proteins: Their role in colonization and invasion of human epithelial and endothelial cells. *Molecular Microbiology*, 10(3), 499-510.

von Heijne, G. (1989). The structure of signal peptides from bacterial lipoproteins. Protein Engineering, 2(7), 531-534.

Yamaguchi, K., Yu, F., & Inouye, M. (1988). A single amino acid determinant of the membrane localization of lipoproteins in E. coli. *Cell*, 53(3), 423-432.

Yazdankhah, S. P., & Caugant, D. A. (2004). Neisseria meningitidis: An overview of the carriage state. *Journal of Medical Microbiology*, 53(9), 821-832.

Zückert, W. R. (2014). Secretion of bacterial lipoproteins: Through the cytoplasmic membrane, the periplasm and beyond. Biochimica Et Biophysica Acta (BBA)-*Molecular Cell Research*, 1843(8), 1509-1516.

6. Appendix

6.1. Appendix 1

6.1.1. Abbreviations

Apolipoprotein N-acyltransferase (Lnt)

Factor H binding protein (fHbp)

Human factor H (hFH)

Invasive meningococcal disease (IMD)

Isopropyl β-1- thiogalactopyranoside (IPTG)

Lipooligosacharrides (LOS)

Localisation of lipoproteins (Lol)

Luria-Bertani (LB)

Optical density (OD)

Outer membrane proteins (OMP)

Preprolipoprotein diacylglyceryl transferase (Lgt)

Preprolipoprotein diacylglyceryl transferase (Lgt)

Serogroup B (MenB)

Serum bactericidal antibody (SBA)

Surface lipoprotein assembly modulator (Slam)

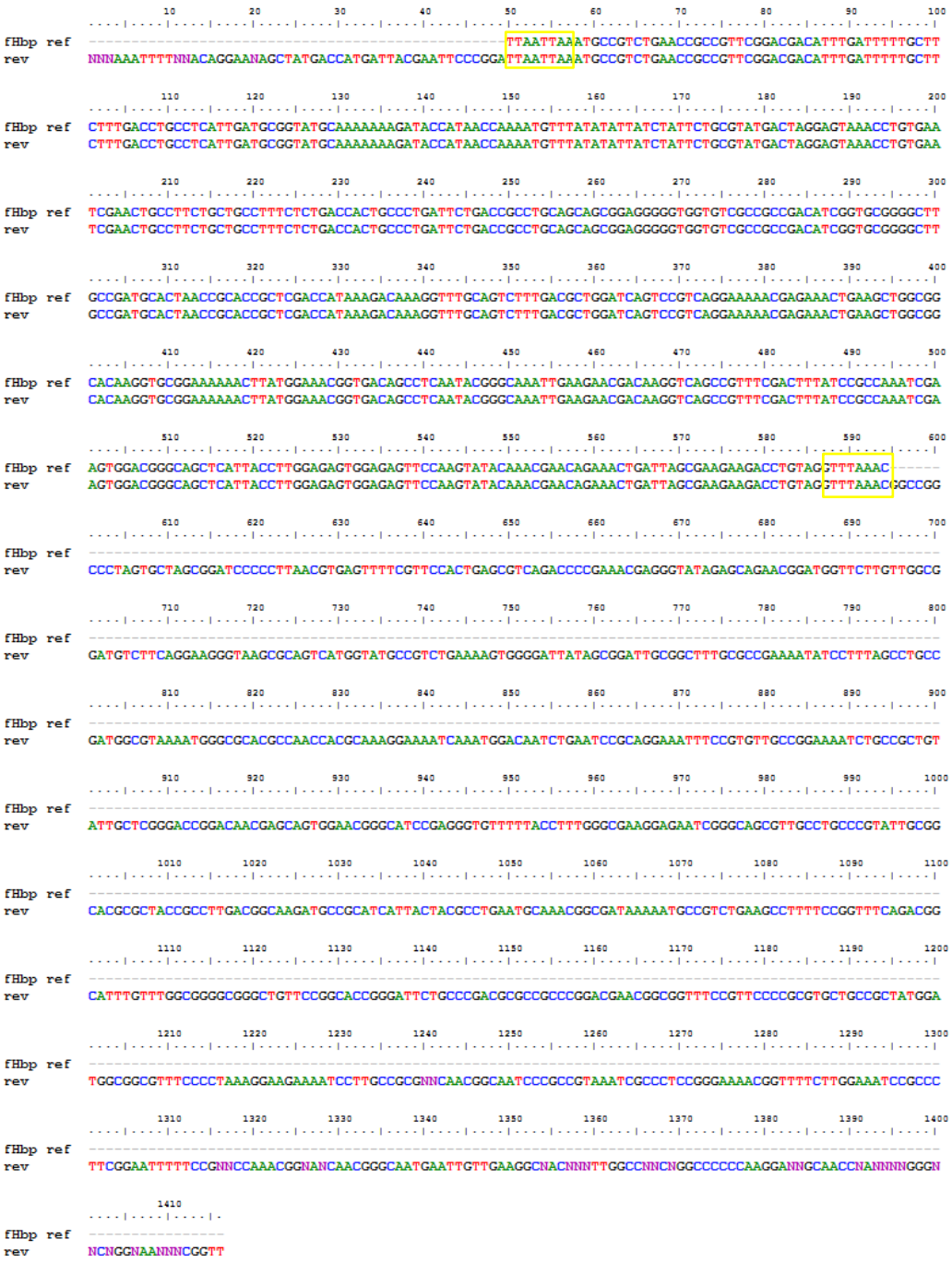
Toll-like receptors (TLR)

Whole cell (WC)

6.2. Appendix 2

6.2.1. Sequencing alignment using the tool BioEdit

For each recombinant plasmid DNA sequence, the forward and reverse sequencing with pGCC4 primers (Eurofins Genomics and Genewiz) were aligned using the tool BioEdit (Figure 6.1- 6.8).


fHbp ref
rev

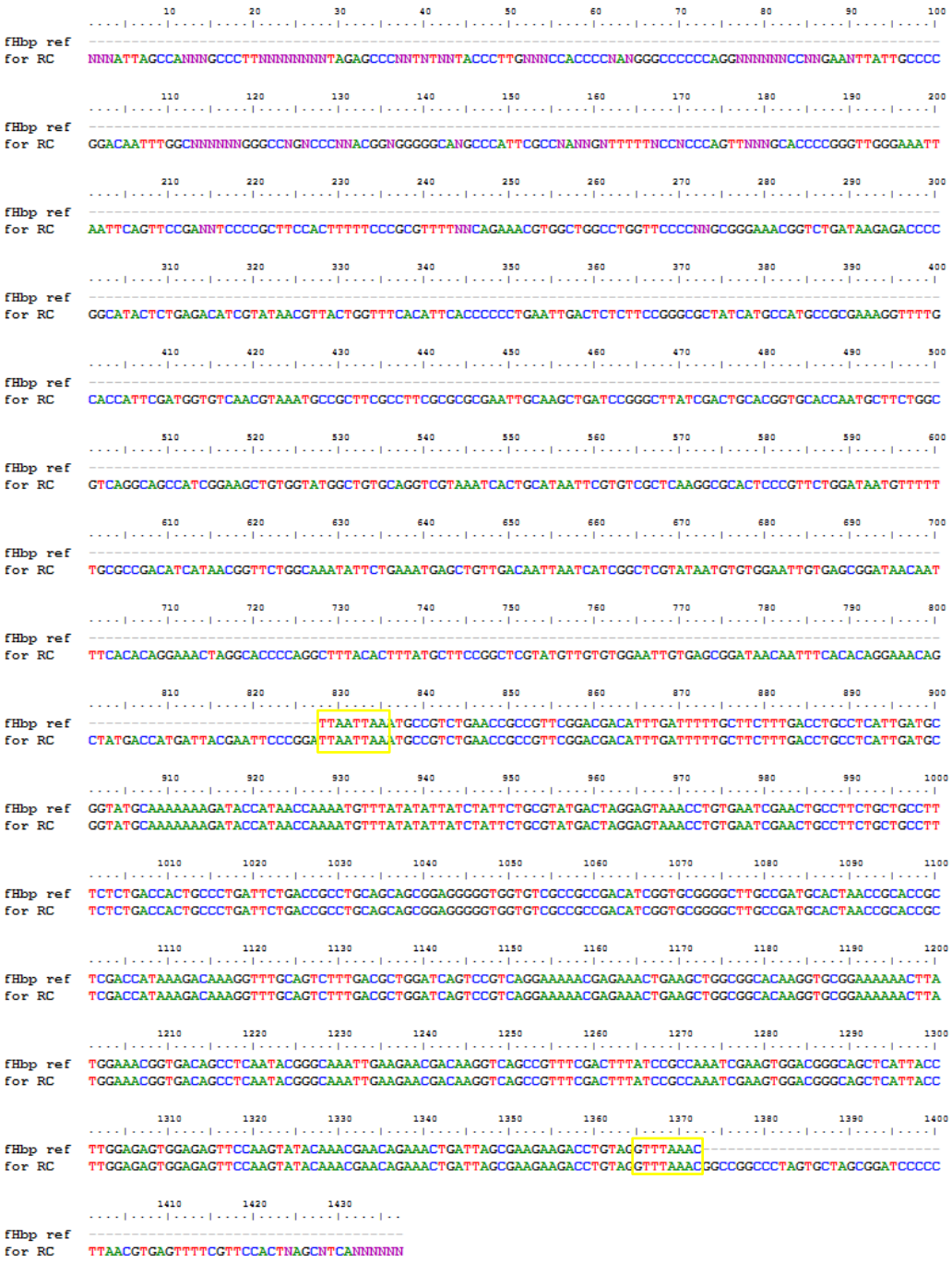
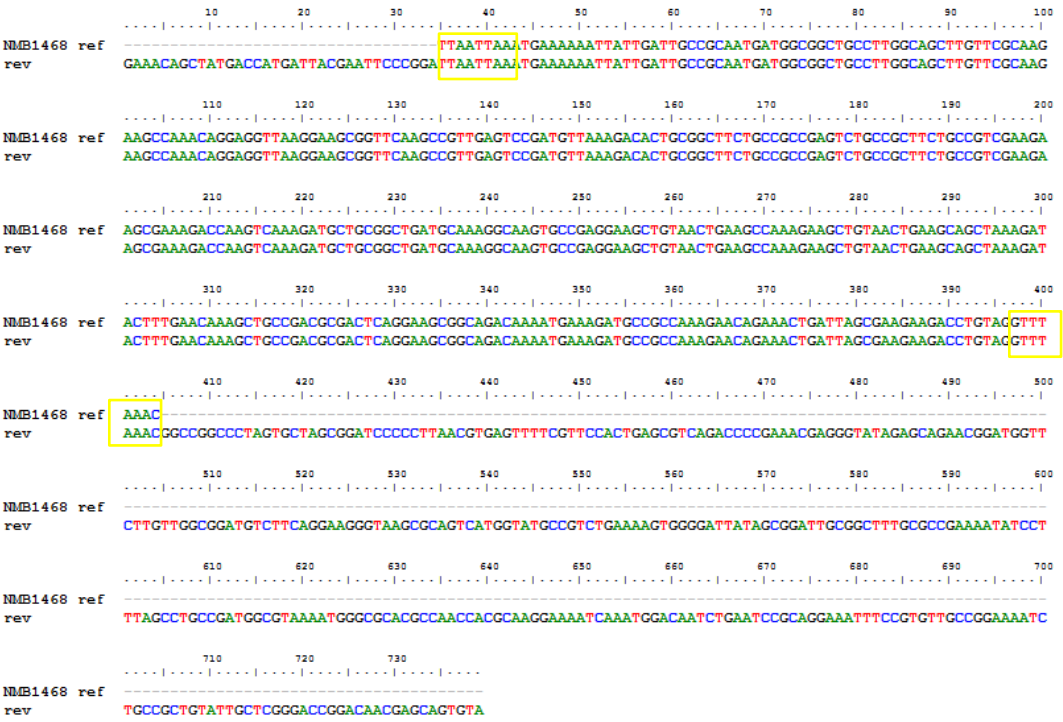


Figure 6.1 Alignment of the DNA sequence of c-Myc tagged *fHbp* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


NMB1468 ref
rev

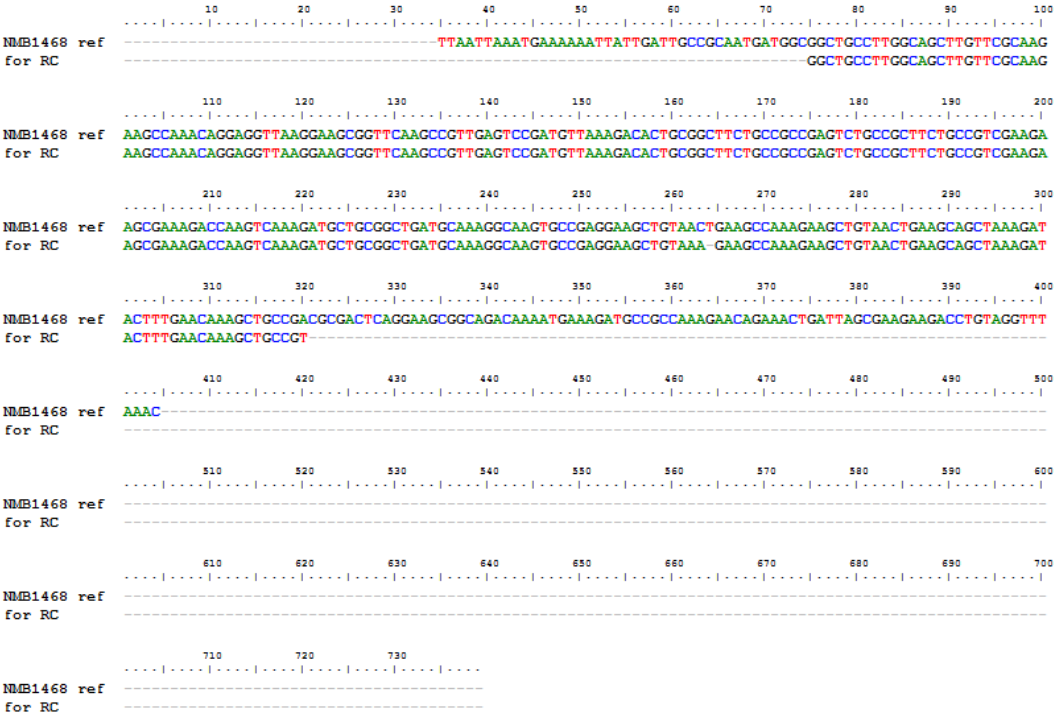
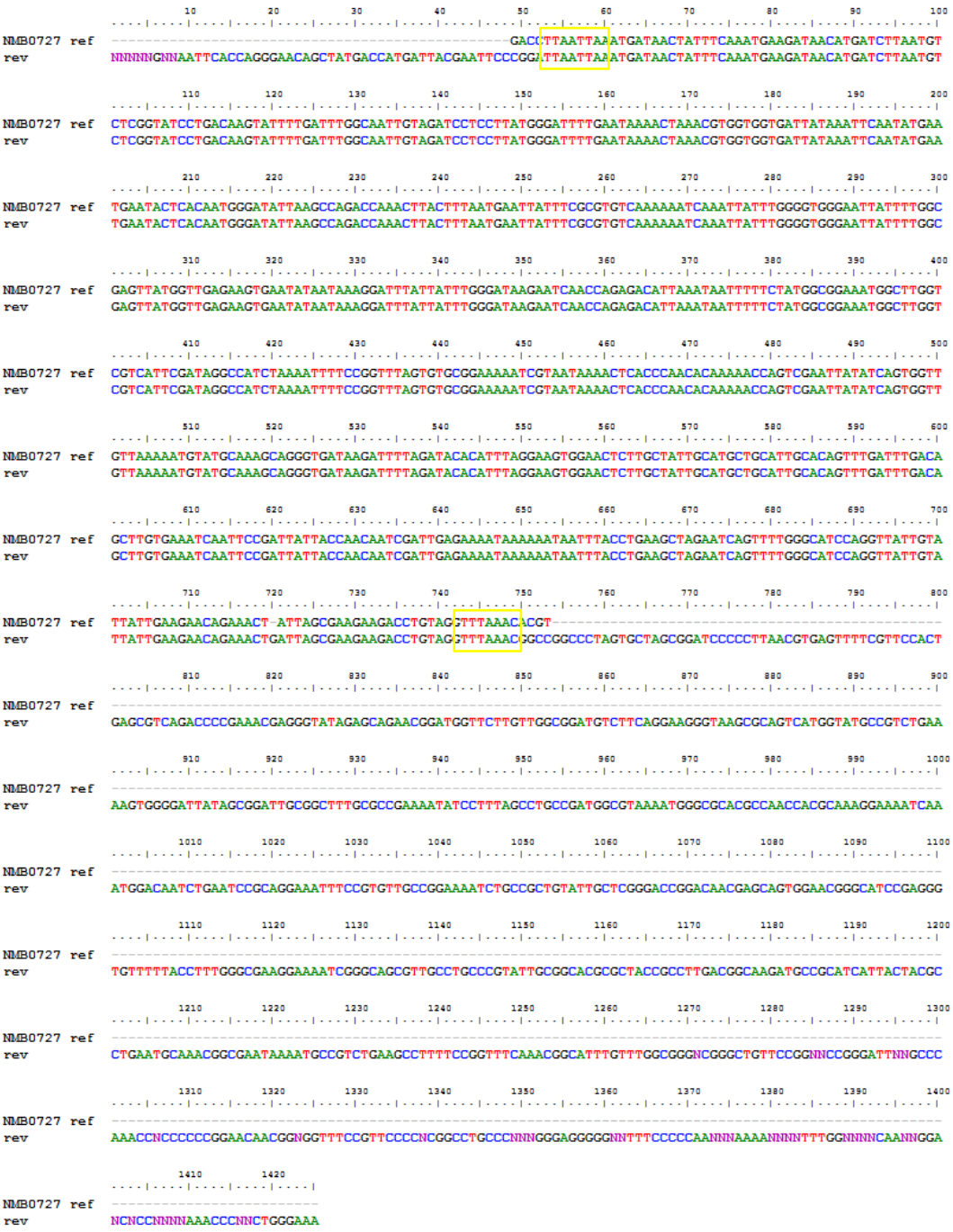


Figure 6.2 Alignment of the DNA sequence of c-Myc tagged *NMB1468* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


NMB0727 ref
rev

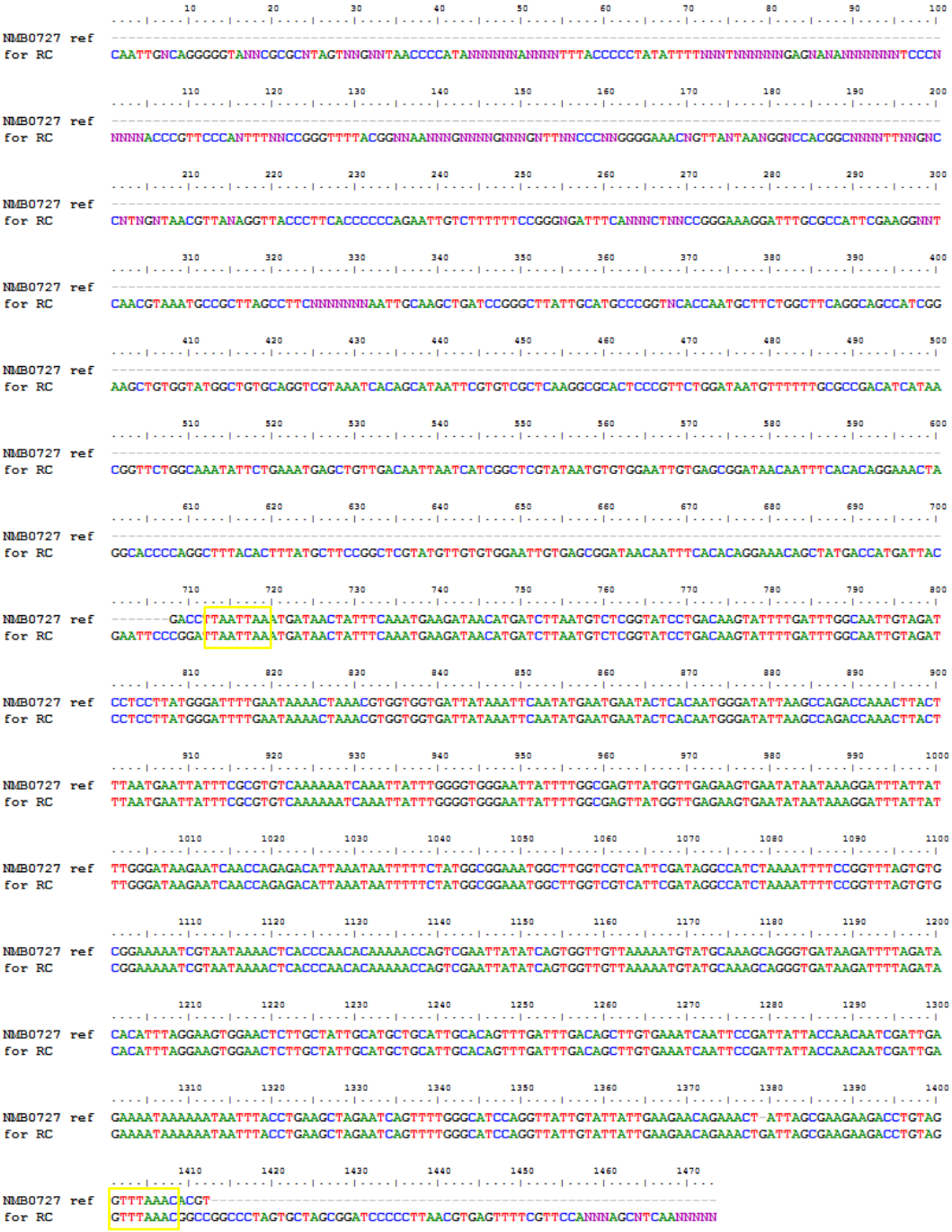
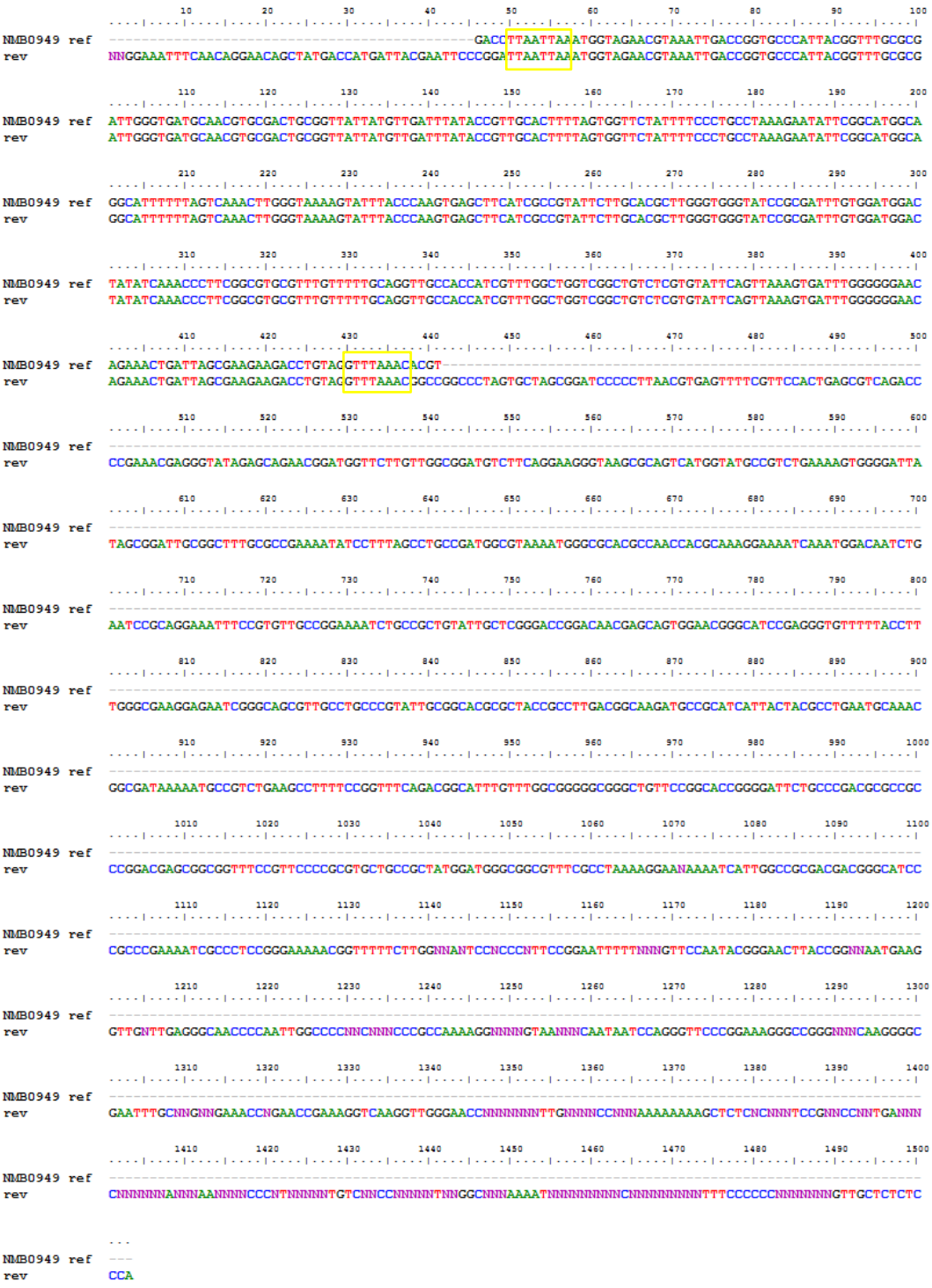


Figure 6.3 Alignment of the DNA sequence of c-Myc tagged *NMB0727* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


NMB0949 ref
rev

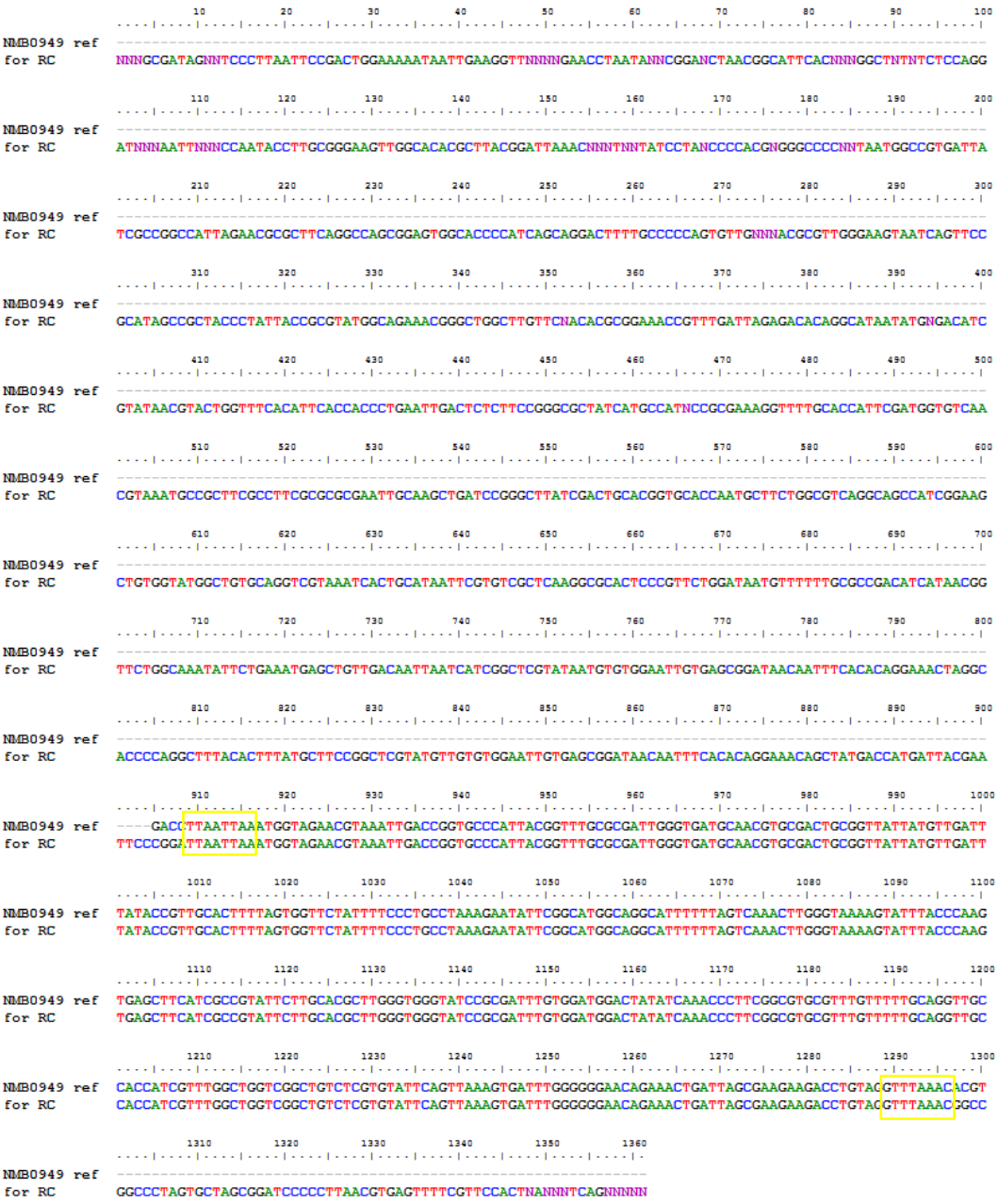
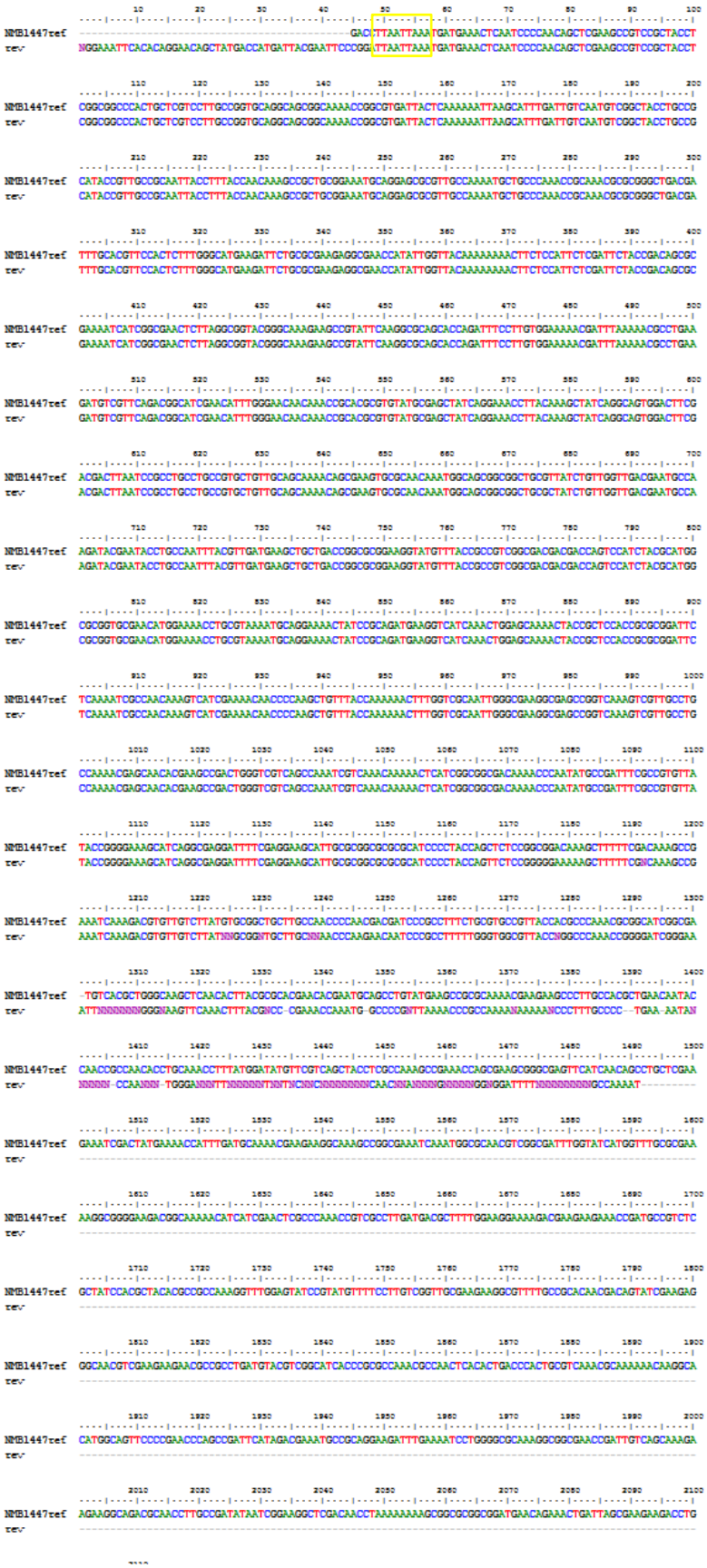


Figure 6.4 Alignment of the DNA sequence of c-Myc tagged *NMB0949* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


NMB1447ref
tev

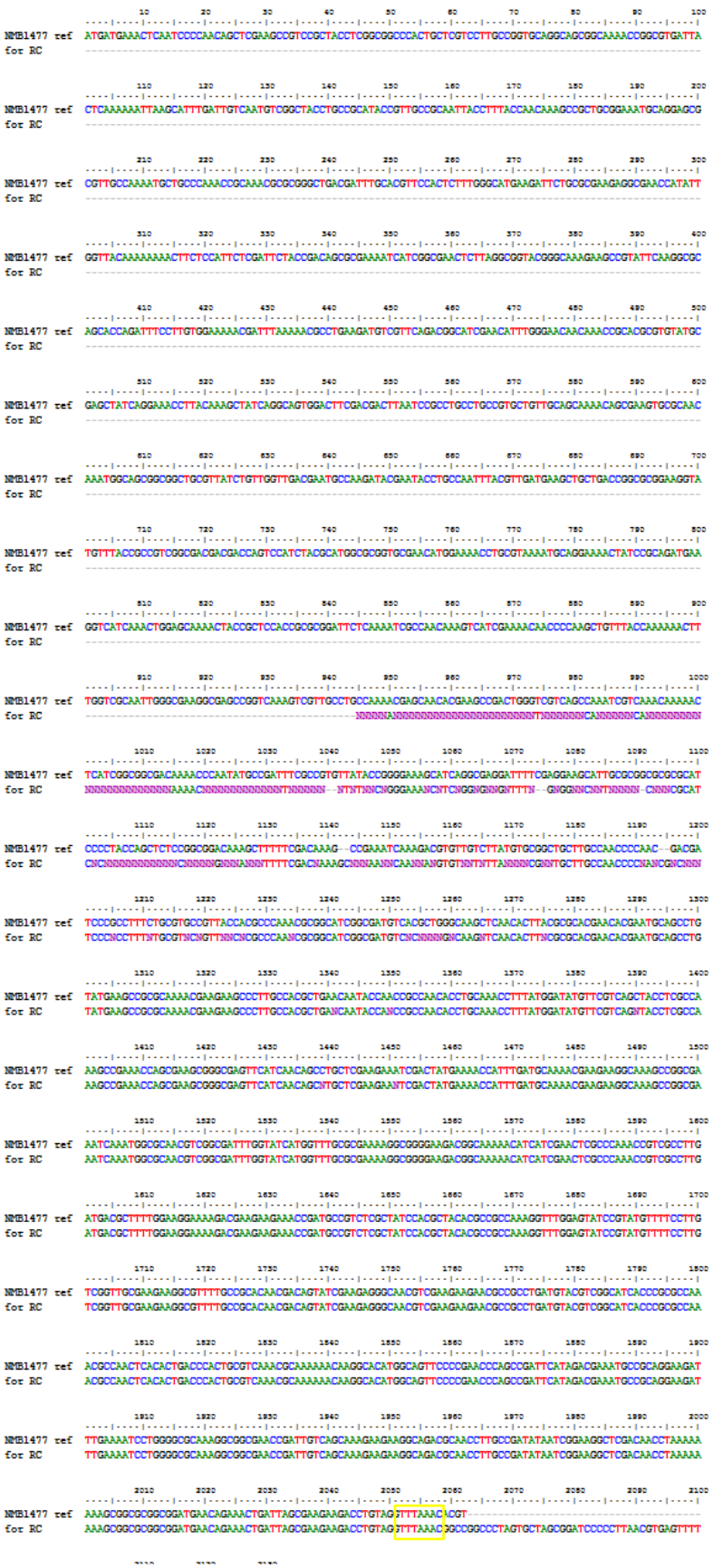
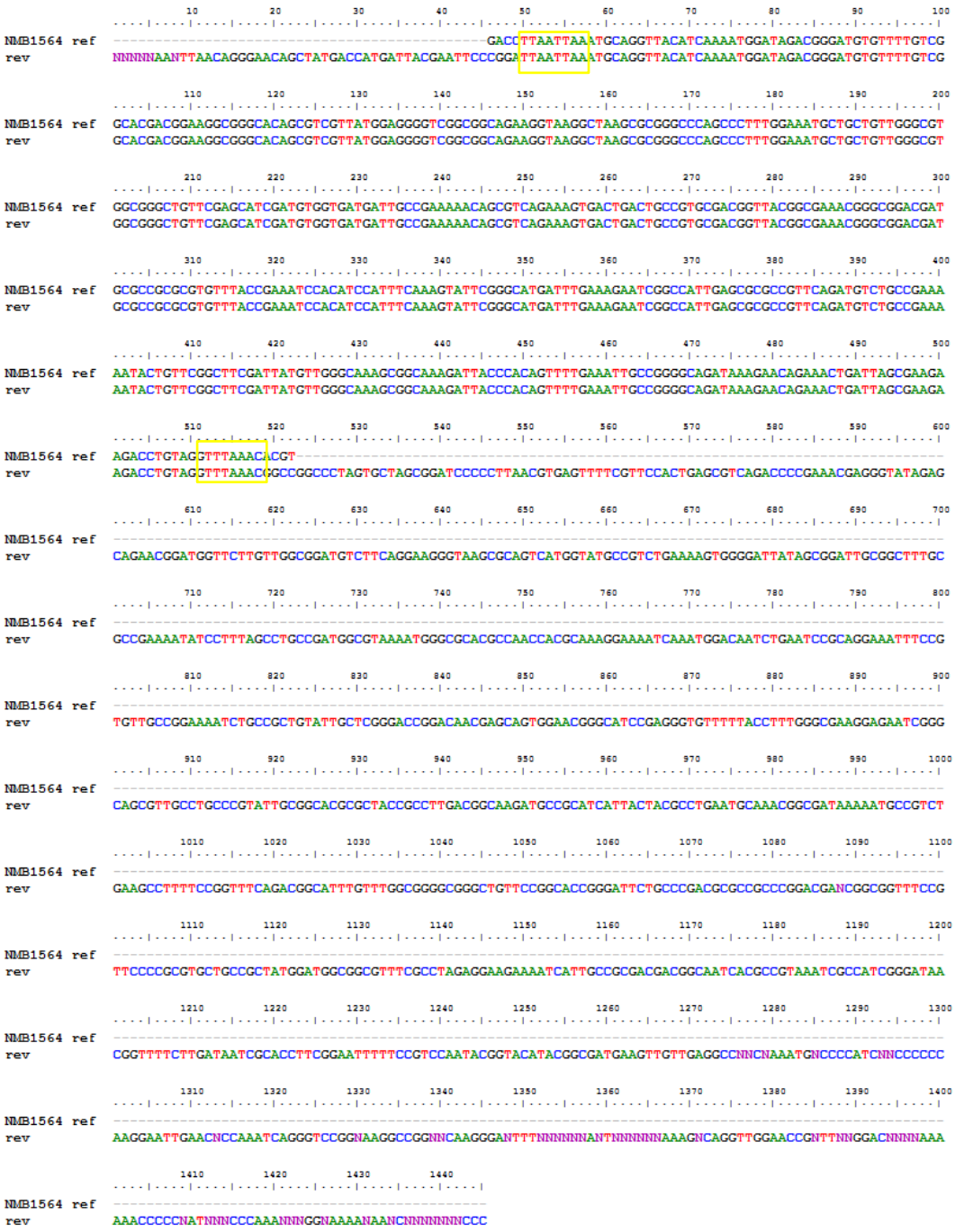


Figure 6.5 Alignment of the DNA sequence of c-Myc tagged *NMB1447* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


NMB1564 ref
rev

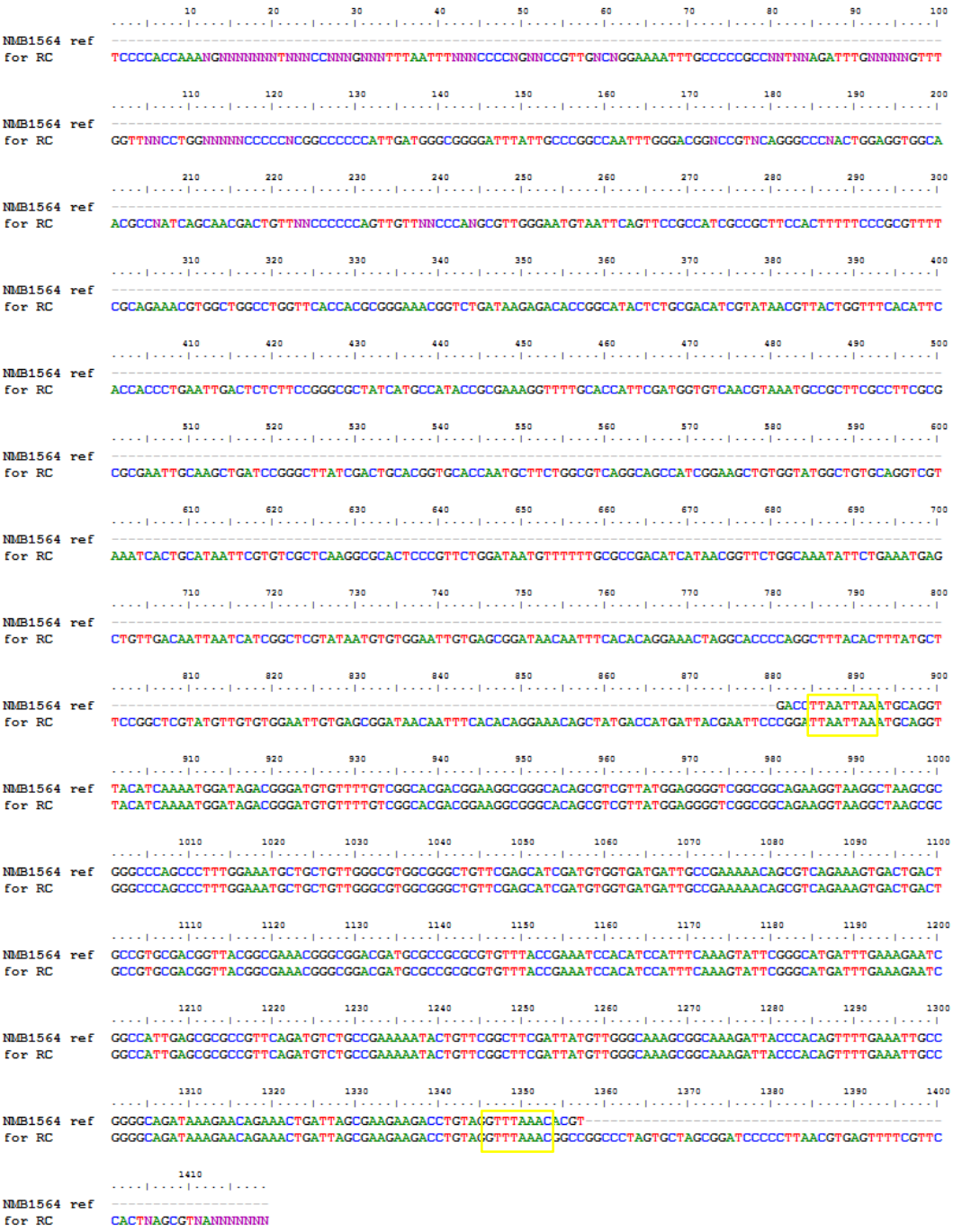
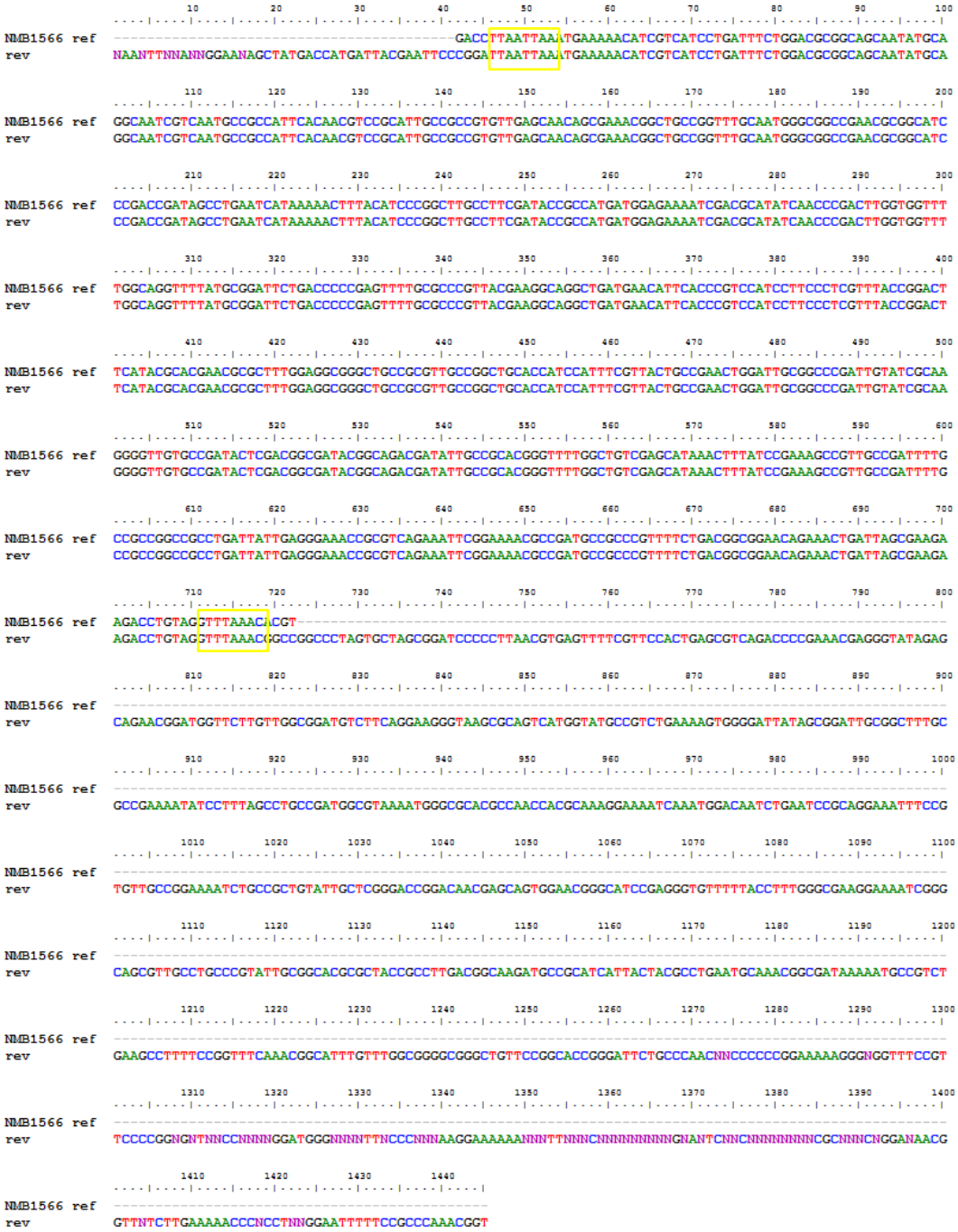


Figure 6.6 Alignment of the DNA sequence of c-Myc tagged *NMB1564* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


NMB1566 ref
rev

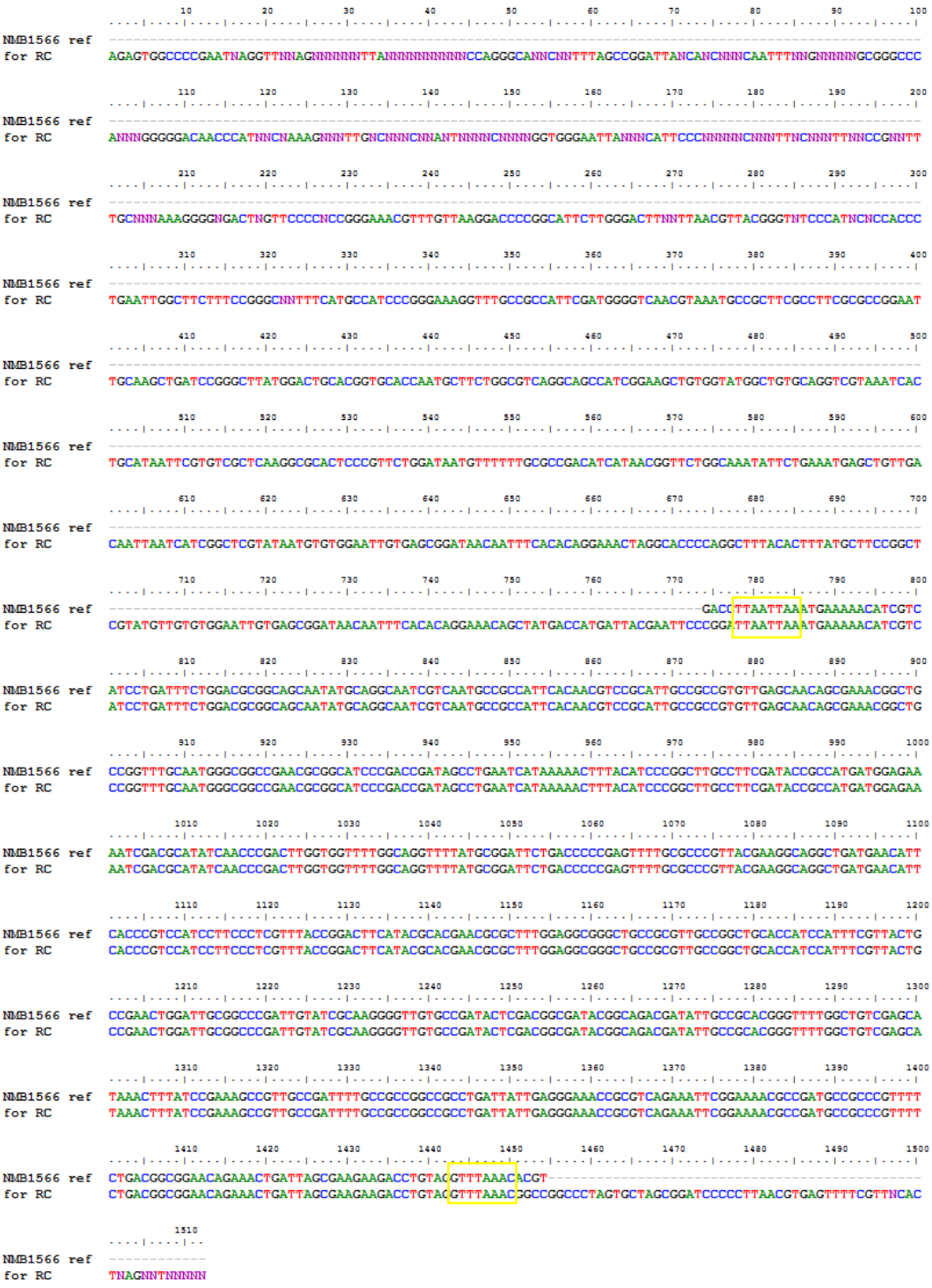
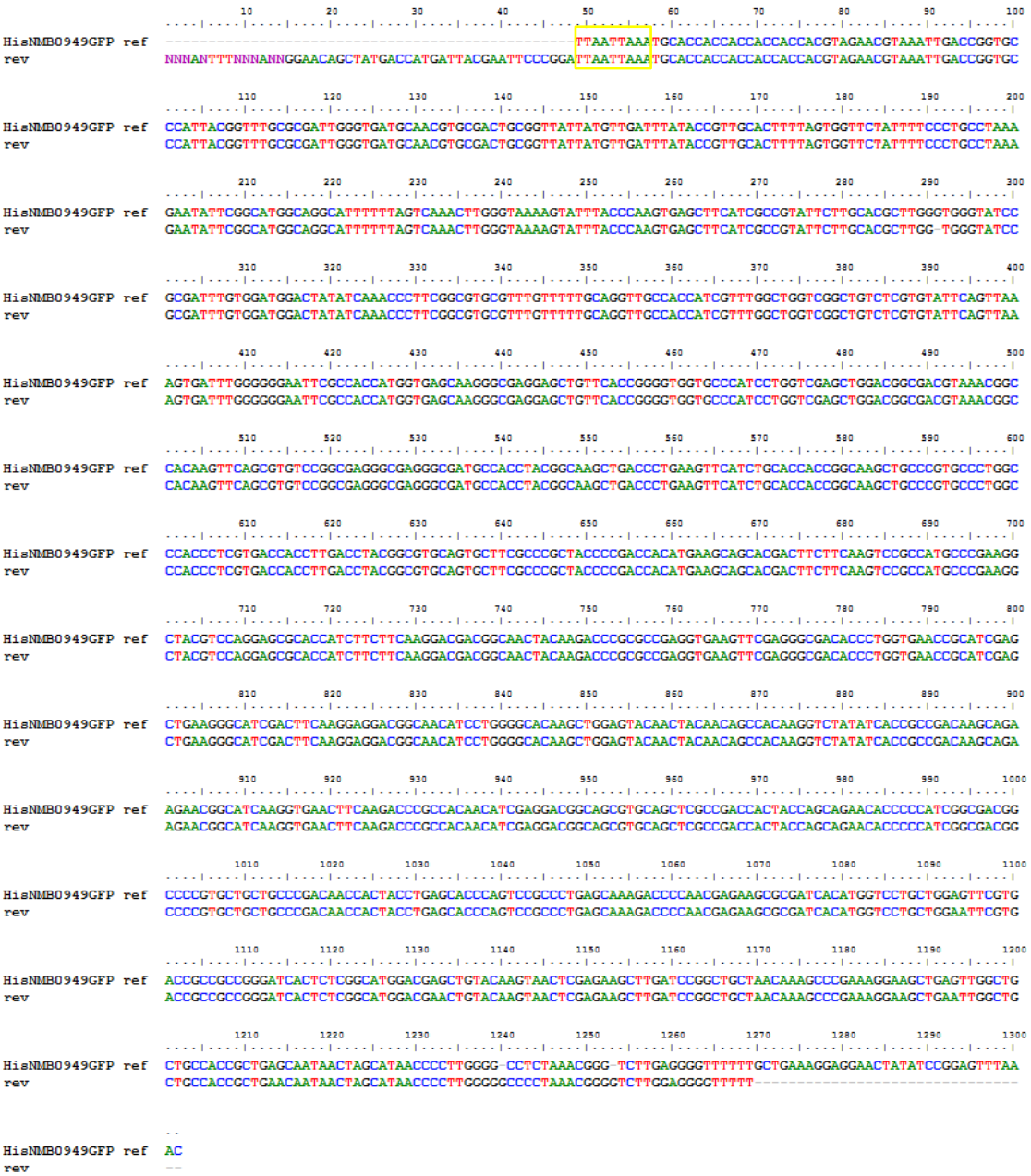


Figure 6.7 Alignment of the DNA sequence of c-Myc tagged *NMB1566* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.


HisNMB0949GFP ref
rev
TTAATTAAATGCACCACCACCACCACCACGTAGAACGTAAATTGACCGGTGC
NNNANTTTNNNANNGGAACAGCTATGACCATGATTACGAATTCCCGGATTAATTAAATGCACCACCACCACCACCACGTAGAACGTAAATTGACCGGTGC
CCATTACGGTTTGCGCGATTGGGTGATGCAACGTGCGACTGCGGTTATTATGTTGATTTATACCGTTGCACTTTTAGTGGTTCTATTTTCCCTGCCTAAA
CCATTACGGTTTGCGCGATTGGGTGATGCAACGTGCGACTGCGGTTATTATGTTGATTTATACCGTTGCACTTTTAGTGGTTCTATTTTCCCTGCCTAAA
GAATATTCGGCATGGCAGGCATTTTTTAGTCAAACTTGGGTAAAAGTATTTACCCAAGTGAGCTTCATCGCCGTATTCTTGCACGCTTGGGTGGGTATCC
GAATATTCGGCATGGCAGGCATTTTTTAGTCAAACTTGGGTAAAAGTATTTACCCAAGTGAGCTTCATCGCCGTATTCTTGCACGCTTGG-TGGGTATCC
GCGATTTGTGGATGGACTATATCAAACCCTTCGGCGTGCGTTTGTTTTTGCAGGTTGCCACCATCGTTTGGCTGGTCGGCTGTCTCGTGTATTCAGTTAA
GCGATTTGTGGATGGACTATATCAAACCCTTCGGCGTGCGTTTGTTTTTGCAGGTTGCCACCATCGTTTGGCTGGTCGGCTGTCTCGTGTATTCAGTTAA
AGTGATTTGGGGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC
AGTGATTTGGGGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC
CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGC
CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGC
CCACCCTCGTGACCACCTTGACCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG
CCACCCTCGTGACCACCTTGACCTACGGCGTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG
CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAG
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGA
CTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAGGTCTATATCACCGCCGACAAGCAGA
AGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGG
AGAACGGCATCAAGGTGAACTTCAAGACCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGG
CCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTG
CCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAATTCGTG
ACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAACTCGAGAAGCTTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG
ACCGCCGCCGGGATCACTCTCGGCATGGACGAACTGTACAAGTAACTCGAGAAGCTTGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAATTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGG-CCTCTAAACGGG-TCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGAGTTTAA
CTGCCACCGCTGAACAATAACTAGCATAACCCCTTGGGGGCCCCTAAACGGGGTCTTGGAGGGGTTTTT
AC
AC

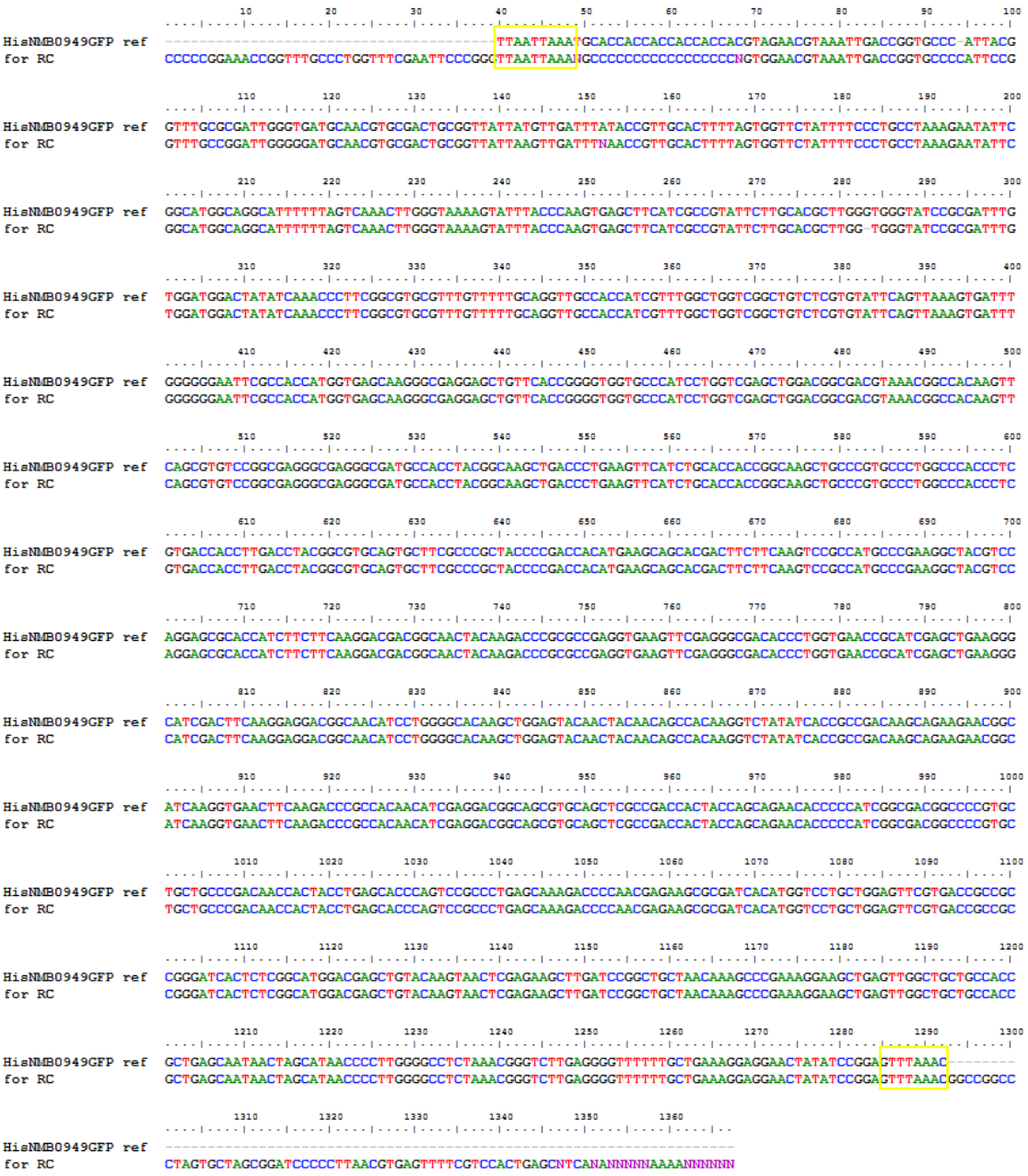


Figure 6.8 Alignment of the DNA sequence of His tagged, GFP reporter containing *His-NMB0949-GFP* obtained from forward (reverse complement) and reverse sequencing with pGCC4 primers. Restriction sites are highlighted in yellow boxes.