Understanding the Role and Mechanism of the DedA Family of Integral Membrane Proteins in Antimicrobial Resistance

Joseph Carroll

The University of Kent School of Biosciences – August 2018

Word Count: 17,110

Number of Pages: 57

Abstract

Membrane proteins perform critical functions such as signal transduction, ion transport, and drug efflux among many other roles. DedA are a ubiquitous family of integral membrane proteins that have been shown to be important for bacterial viability. In addition, DedA proteins contribute substantially to antimicrobial resistance in clinically relevant pathogens including *Klebsiella pneumoniae*, where a DedA family member was recently shown to be essential for resistance to colistin. Targeting these proteins for inhibition could sensitise bacteria to other antimicrobials, making them easier to kill. However, this quest is hampered by the lack of basic understanding of the structure of DedA, their mechanism of action, and their actual physiological role in the cell. To address the dearth of structural and functional information on the DedA family, a panel of these integral membrane proteins have been cloned into two expression systems pET and pBAD. One particular *E.* coli DedA homologue, YqjA, was then overexpressed, solubilized, and purified. Following this, YqjA was analyzed using size exclusion chromatography. The SEC trace obtained indicated that YqjA may function as both a monomer and a dimer, although further analysis will be required to confirm this. Future work will include x-ray crystallography, cross-linking, and mass spectrometry studies.

Table of Contents

1. Introduction

1.1. Background

Since the development of antimicrobials in the early 20th century by the likes of Alexander Fleming and Ernest Duchesne millions of lives have been saved through their use. However, since then, microorganisms are increasingly developing resistance to these compounds resulting in a decline in efficacy of antimicrobial agents[1]. Antimicrobial resistance now poses a major threat globally, with the World Health Organization (WHO) classing antimicrobial resistance as an emerging global disease[2]. There are several mechanisms through which antimicrobial resistance can occur. One such mechanism involves the upregulation of enzymes which break down or inactivate antimicrobial agents before they can exert there affect. This includes the production of ß-lactamases which inactivate ß-lactams by hydrolysing the ß-lactam amide. Another example are transpeptidases, which covalently bind and catabolize penicillins in *Staphylococcus aureus*[3–5]. Antimicrobial resistance can also occur through the upregulation of efflux pumps which remove the antimicrobial before they can reach their target such as efflux of fluoroquinolones in *S.aureus*[3]. Modification of the target receptor site of the antimicrobial agent can also incur resistance such as in pneumococci, to which a change in penicillin binding protein 2 (PBP2) resulted in penicillin resistance[3,6].

Some pathogens present a greater risk than others, such as *Pseudomonas aeruginosa*, *Staphylococcus aureus, Klebsiella pneumoniae* and *Escherichia coli.* This is due to several factors including the prevalence of resistant strains in a clinical setting and the number and type of antimicrobial agents they carry resistance to, with some acquiring resistance to more common antimicrobials making treatment more challenging. Some strains of pathogens such as S. *aureus and P. aeruginosa* have developed resistance to several antimicrobials, this is called multidrug resistance (MDR). This occurs because of a combination of mechanisms, such as in some *P. aeruginosa* isolates in which efflux pumps and poor outer membrane permeability result in resistance to several antimicrobials including ß-lactams, tetracycline, fluoroquinolones and chloramphenicol[3,7,8]. MDR is particularly problematic as it could mean treatment of MDR clinical isolates requires the use of last resort antimicrobials such as colistin, to which resistance

has begun to emerge in several bacterial species[68], and which has also been shown to possess neurotoxic and nephrotoxic properties[9,10]. Due to the prevalence of antimicrobial resistance, alternative strategies to combat this issue are currently being developed. Current strategies involve either targeting resistance mechanisms or developing novel compounds which will be unaffected by the various resistance mechanisms[5]. The discovery of novel integral membrane proteins which play a role in antimicrobial resistance mechanisms could lead to the development of small molecule inhibitors which target these proteins, perturbing function, which could restore function to some antimicrobials[5].

1.2. DedA Family of Integral Membrane Proteins

DedA are a family of highly conserved integral membrane proteins found in all domains of life and have been shown to be important in bacterial viability[11]. The *ecdedA* gene was first discovered downstream of the hisT operon in 1987[69], and is short for downstream *E. coli* DNA gene A[12]. With homologues present in most sequenced genomes, DedA homologs have been shown to play a wide range of different roles in various organisms, such as in *Klebsiella pneumoniae*, a DedA homologue has been shown to be essential for resistance to colistin[13,14] or in *Borrelia burgdorferi*, in which a DedA homologue, annotated *bb0250*, has been shown to be essential for cell division and envelope integrity[15].

There are currently thousands of *dedA* homologues, each consisting of ~200-250 amino acids, with some organisms possessing as many as eight homologues and others just one. *E.coli* are an example of this, possessing eight DedA homologues including *yqjA*, *yghB*, *dedA*, *ydjX*, *ydjZ*, *yqaA*, *yabI*, and *yohD,* with the best characterised being YqjA and YghB. Phylogenetic analysis of 350 bacterial genomes from the NCBI database using BLAST on the amino acid sequences of all eight *E. coli* DedA family members[16], revealed that 33 bacterial species did not possess a DedA family member[12]. The presence of a DedA homologue was found to be highly variable and didn't seem to correlate with the environment in which the species colonized[12]. Despite the DedA family being ubiquitous, their role remains currently unknown, with the amino acid sequence of all DedA family members failing to resemble any characterised enzyme, transporter, ion channel or signalling protein[17].

Recent structural data has shown that some secondary transporters previously believed to belong to separate families may share an ancestry due to a shared domain in their 3D structure. BetP, a member of the betaine carnitine choline transporter family (BCCT)[18], AdiC from the amino acid-polyamine-organocation (APC) superfamily[19,20], and the bacterial homologue of the neurotransmitter/sodium symporter family (NSS) LeuT[21] have all been shown to share a common structural domain consisting of an inverted 5 transmembrane helical repeat. These two 5 transmembrane helices together form a 10 transmembrane (TM) helix subdomain commonly referred to as the LeuT fold due to it originally being discovered in LeuT, which is thought to play a critical role in substrate recognition and binding[21].

LeuT is a Na^+ coupled symporter, originally discovered in the thermophilic bacterium *Aquifex aeolicus*, with its crystalline structure first being published in 2005[21]. Despite there being little in the way of sequence similarity between the 5TM inverted repeats there is a high amount of structural similarity[22]. LeuT and other related transporters are believed to operate using the rocking bundle mechanism[23]. The rocking bundle is a form of alternating access which involves the rotation of several transmembrane domains which together form a bundle. This movement appears to open and close the cytoplasmic and extracellular permeation pathways[23]. Structural analysis of several DedA homologues, all of which have been predicted to consist of 5 TM domains, showed that they share similar structural domains with the 5TM inverted repeats found in LeuT[16,24]. This could indicate that some DedA family members may function interchangeably as either heterodimers or homodimers[25,26] with a similar mechanism to that of LeuT.

1.3. What is Known About the Structure of DedA?

Currently there is no known structure for any member of the DedA family. Although some structural features, such as the number of predicted transmembrane regions, of certain DedA homologues is known[27], primarily as a result of hydropathy profile alignment. This technique involves analysing the hydropathy profiles of various proteins to assess how closely related they are based on their predicted structure. This technique is particularly useful as the amino acid sequence can vary significantly between homologues due to rapid changes occurring as the proteins evolve. However, due to necessity, the structures typically remain very well conserved[28].

It has been observed that many of DedA homologues consist of between 4-6TM helices, with many possessing 5TM helices resembling those observed in LeuT[29]. This was predicted using transmembrane helices based on the hidden Markov model (TMHMM)[25]. Hydropathy profiles have been shown to be more accurate than amino acid sequence alignment for comparing proteins structures to one another since the hydropathy profiles predict the transmembrane regions of each protein by assessing the hydrophobicity of the amino acids which comprise the protein. This method of transmembrane region prediction is less prone to be affected by mutations since the hydrophobicity of the proteins can remain the same. Therefore, for proteins with a low sequence homology, hydropathy profiles can be a more accurate way of comparing the structures[28].

DedA are an ancient family of integral membrane proteins, as indicated by their presence in all domains of life. Not only this, but in most organisms in which DedA homologues are present, they are typically found in multiples[12]. This means that it is likely that larger and more recently formed protein families such as those which possess a LeuT fold may have evolved from DedA-like protein[25]. In a study by Khafizov et al, they compared LeuT-fold structures and members of the DedA family using a family averaged hydropathy profile alignment. They found that there was a high degree of conservation and variability in the same regions in both families, such as unusual features of the third transmembrane region and canonical hydrophobic peaks for transmembrane regions 2, 4, and 5. This evidence seems to indicate that ancient members of the DedA family may have undergone a gene duplication and inversion to form what is now referred to as the LeuT fold, however more structural evidence will be required to confirm this.

1.4. What is Known About the Function and Physiological Role of DedA?

Although the function of DedA family members currently remains elusive, the pleiotropic phenotype observed in BC202 (W3110 *ΔyqjA::ΔyghB*) could give an indication of some potential functions of this protein family. BC202 is typically used for *E. coli* DedA homologue functional studies, since the only observable phenotype in the single-deletion mutant of *yqjA* is alkaline pH sensitivity, failing to grow above pH 8.5[27]. Simultaneous deletion of the genes encoding *yqjA* and *yghB* results in temperature sensitivity, causing growth retardation at temperatures above 42°C[11]. Other phenotypes include cell division defects, as a result of inefficient export of periplasmic amidases AmiA and AmiC via the twin arginine transport (tat)

pathway, and drug sensitivity[27]. A distinguishable change also occurs in the composition of the phospholipid membrane[24]. In the parent *E. coli* strain W3110, phosphatidylethanolamine (PE) comprises approximately 60% of the total membrane phospholipids. In BC202, when grown at 30°C (the permissive temperature), phospholipids were synthesised normally but the phospholipid composition in the membrane was altered, with PE comprising less than 50% of the total phospholipid composition and higher levels of the acidic phospholipids phosphatidylglycerol (PG) and cardiolipin (CL) were observed[30]. This phenotype is only observed when both *yqjA* and *yghB* are knocked out, meaning they could either be redundant, or potentially function interchangeably as either heterodimers or homodimers[25,26].

Multiple sequence alignments have identified several conserved acidic residues within YqjA and YghB whose role remains unknown[26]. Mutational analyses were conducted on several of these conserved residues to determine what, if any, role they play in the function and oligomeric state of YqjA. It was found that acidic residues glutamic acid 39 (E39) and aspartic acid 51 (D51) mutants, which had been mutated to alanines to impair the negative charge and therefore function, were unable to rescue growth of BC202 at 44°C. This indicates that these two acidic residues are crucial for the function of YghB and YqjA[26]. All other acidic residues, which had also been mutated to alanine's, were proven to be non-essential for the function of YqjA, despite being conserved among several DedA family members[26]. Alanine substitution of both E39 and D51 acidic residues also demonstrated that these residues are non-essential for the interaction between and formation of YqjA homodimers in-vivo[26], with both YqjA & the E39A/D51A mutant shown to be present in different oligomeric states and form homodimers[26]. Disulphide cross-linking also demonstrated that YqjA dimerizes in the absence of YghB[26]. It was also shown that the deletion of E39 and D51 between TM helices 1 and 2 in both YqjA and YghB prevented these two proteins from rescuing growth at 44°C[26]. This is particularly important since membrane embedded acidic residues are a prominent feature of proton dependent transporters, which could indicate that both YqjA, and YghB may function as transmerane transporters[26].

As previously mentioned, *E. coli* possesses 8 DedA homologues, *yqjA*, *yghB*, *EcdedA*, *ydjX*, *ydjZ*, *yqaA*, *yabI*, and *yohD*. It has been shown that recombinant expression of each of these homologues in BC202 brings mixed results, with 4 homologues, YqjA, YghB, YabI, and YohD, capable of restoring growth above 42°C and cell division and the remaining 4, EcDedA, YdjX, YdjZ, and YqaA, unable to do so[11]. This could be an indication that the homologues

present in each of these groups have diverged to have different functions within the cell. The DedA homologues that can restore these functions appear to be located together separately from those which are unable to do so on a phylogenetic tree. Although the function of these proteins is unknown this could indicate that these two protein groups have diverged to have different functions.

Many physiological functions are dependent on a stable pH which is maintained by the PMF. Therefore, any disturbance in the PMF can have disastrous effects on the cell by disrupting the function of various macromolecules including enzymes, and membrane proteins which are responsible for a whole host of cellular functions[31]. An example of functions which are heavily dependent on the PMF include protein translocation[32], ATP dependent drug extrusion systems[33], and drug efflux pumps[34].

It is unclear yet why such a diverse phenotype is observed in BC202. However, it is possible to restore growth above 42°C by lowering the pH of the growth conditions to pH 6.0 or by overexpressing MdfA, a Na^+ K^+/H^+ antiporter from the major facilitator superfamily (MFS)[24]. Some studies have suggested that this might be an indication that YqjA, YghB and potentially other DedA family members, may play a role in maintaining the proton motive force within the cell[12,16,24]. Another factor that may suggest a role for the DedA family in maintaining the PMF is that BC202 is sensitive to biocides exported by Mdfa. Such biocides include methyl viologen, benzalkonium chloride, acriflavine, β-lactam antibiotics, and ethidium bromide. However, BC202 is not sensitive to all biocides meaning its outer membrane remains intact. This means that the biocide sensitivity observed in BC202 could be due to a disturbed PMF. The role of this protein family still remains unknown, and due to its diversity, it is possible that DedA homologues may have evolved different functions, especially since some members of the DedA family have been shown to confer resistance to specific antimicrobials. DedA in *Klebsiella pneumoniae*, has been shown to be essential for resistance to colistin[13], and YqjA in *Salmonella enterica* has been shown to confer resistance to the antimicrobial peptide magainin[13,35]

1.5. Aim of this Thesis

It is estimated that membrane proteins make up 15-30% of open reading frames in sequenced genomes[36]. However, only 1% of all known proteins in the protein data bank are membrane proteins[12]. The sheer size of the gap between these two values highlights how much more information there is to discover. Despite this vast gap in knowledge of this group of proteins, it is clear that they are vital for many cellular functions as indicated by the fact that ~60% of drugs target proteins located on the cell surface[37]. Integral Membrane Proteins are responsible for a multitude of diseases such as hyperkalemic periodic paralysis, long QT syndrome, and cystic fibrosis. Treatment of such diseases involves the use of various small molecule inhibitors (SMIs) which bind and inhibit the function of their target, such as drug efflux pumps, and therefore sensitize multidrug resistant strains to antimicrobials[38–41].

The development of such treatments relies on structure-based drug design which, in turn, requires information about the function and structure of this protein family. In order to acquire this information, *in vitro* structural and functional studies must first be carried out. Examples of techniques used in these studies include X-ray crystallography, size exclusion chromatography with multi-angle light scattering (SEC-MALS), mass spectrometry, nuclear magnetic resonance spectrometry (NMR), thermal shift assays and *in vitro* activity assays. These require the production of enough protein for use in the various biochemical and biophysical assays, especially since membrane proteins are particularly low in abundance, unstable, and require a native lipid environment in order to remain stable[42].

Protein production involves several steps which each require optimisation. Initially the gene encoding the protein of interest is cloned into an expression vector, following which, the expression vector is transformed into a suitable strain of *E. coli*. The protein of interest is then overexpressed, and finally the membrane is extracted and the protein of interest is solubilized and purified. Many of these steps can be optimised in order to increase the amount of purified protein acquired. Among many others, some of the steps which can be optimized include the expression system used, the strain of E.coli, and the method used to solubilize the protein of interest.

In order to ensure that the recombinant expression of the proteins of interest is successful, two different expression systems will be utilised. These are the pET and pBAD expression systems, which are commonly used in recombinant protein expression and possess several different characteristics, making them particularly useful for optimisation of the expression process (fig 1 & 2.). When undertaking recombinant protein expression, the plasmid expression system being utilised will have a substantial impact on the overall process. Most commonly used expression plasmids, including pET and pBAD, consist of various combinations of different replicons, promoters, selection markers, and multiple cloning sites[70]. The replicon controls the plasmid copy number, which determines the number of copies of the recombinant plasmid within the host cell. A high plasmid copy number however, does not necessarily correspond to higher yields of recombinant protein. This is due to the fact that a high number of plasmids within the host cell can result in metabolic stress, which can decrease the growth rate and potentially result in plasmid instability, ultimately lowering the amount of cells undergoing protein synthesis[71]. Typically, pET expression vectors contain the pMB1 origin of replication, which results in approximately 15-60 copies per cell[70]. Whereas pBAD expression vectors tend to have a lower copy number with approximately 10-12 copies per cell[70].

The promoter regions in pET and pBAD expression systems differ significantly as well (fig. 1 & 2). The T7 promoter region present in pET vectors is commonly used since it can result in up to 50% of total cell protein consisting of the protein of interest[72]. In this system, the gene encoding the protein of interest is cloned downstream of a promoter region which can be recognized by phage T7 RNA polymerase which is typically present in the host organism on either a separate plasmid or in the host genome of the strain being used (fig. 1). In the case of this study, the gene encoding T7 RNA polymerase is present in the genome of the strain being used (BL21-AI) under the transcriptional control of a *lac*UV5 promoter[72] (fig. 1). This means that the system can be induced by the addition of the non-hydrolyzable isopropyl-**β**-D-1-thiogalactopyranoside (IPTG). It is possible for small amounts of T7 RNA polymerase to be produced in this system as a result of leaky expression of it's gene. However, this is somewhat reduced by the presence of a *lacO* operon downstream of the T7 promoter, allowing further control of expression (fig. 1). Leaky expression in the T7 system is due to the fact that it relies on negative control of the lac promoter[72] (fig. 1). In pBAD expression vectors, which rely on positive control, there is a lower level of background expression due to the presence of the promoter region, $araP_{BAD}$[73]. AraC, controls this promoter region through both activation and repression (fig. 2). When there is no arabinose inducer present, AraC is able to repress

translation of the gene of interest by binding to two sites in the bacterial DNA[73]. This process forms a loop of DNA which prevents RNA polymerase from binding to the promoter region (fig. 2). Once the inducer, arabinose, is added to the growth media, AraC switches function and instead promotes transcription of the gene of interest by binding to the *ara* promoter region[73] (fig. 2). This system means that expression is only able to take place in the presence of arabinose.

Another difference between the pET and pBAD expression systems is the type of selection marker used. Selection markers are typically added to plasmid expression systems in order to deter the growth of non-plasmid containing cells. In the case of both pET and pBAD the presence of an antibiotic resistance gene allows cells containing a copy of the recombinant plasmid to grow in the presence of certain antibiotics. In the pET expression system, the selective marker kan confers resistance to kanamycin. Whereas in pBAD, the selective marker amp confers resistance to ampicillin.

One highly significant difference between both the pET and pBAD expression systems is the location of the affinity tag. Affinity tags typically involve the expression of a stretch of amino acids which form a peptide tag attached to the protein of interest, forming a chimeric protein. The presence of an affinity tag means that the recombinant protein can be detected using western blotting during expression trials, as well as one-step affinity purification. Affinity tags can however, occasionally interfere with the protein of interests function and structure[74]. Different expression systems are available with affinity tags in different locations, either at the N or C-terminal. The location of the affinity tag can have a significant impact on the expression and purification of the desired protein, since the tag could, at one end of the protein, be buried inside the fold. Whereas at the other end of the protein, the tag could be solvent-accessible, which is preferable. Since there is currently no three-dimensional structure for the proteins in this study, it is impossible to tell which location would be preferable for the affinity tag. In the case of this study, the affinity tag being used is a polyhistidine-tag which is located at the C-terminal end in pBAD constructs, and at the N-terminal end in pET constructs (Table. 2).

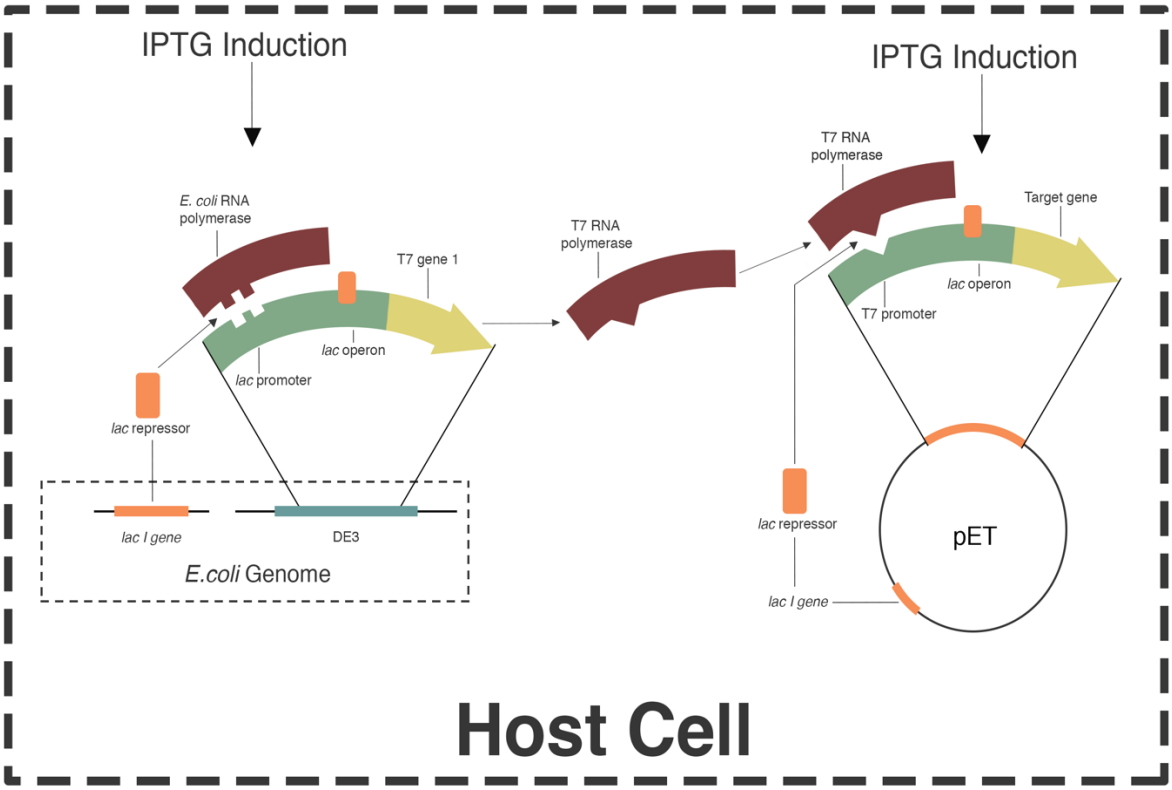


Figure 1. The pET expression system, showing the mechanism of action of the T7 expression system. This figure shows the negative control mechanism by which the addition of IPTG results in expression of the gene of interest. IPTG binds to the lac repressor protein, this results in a loss of inhibition of the lac promoter which results in expression of T7 RNA polymerase in the *E. coli* host genome, which, in turn, results in the expression of the downstream target gene in the pET vector.

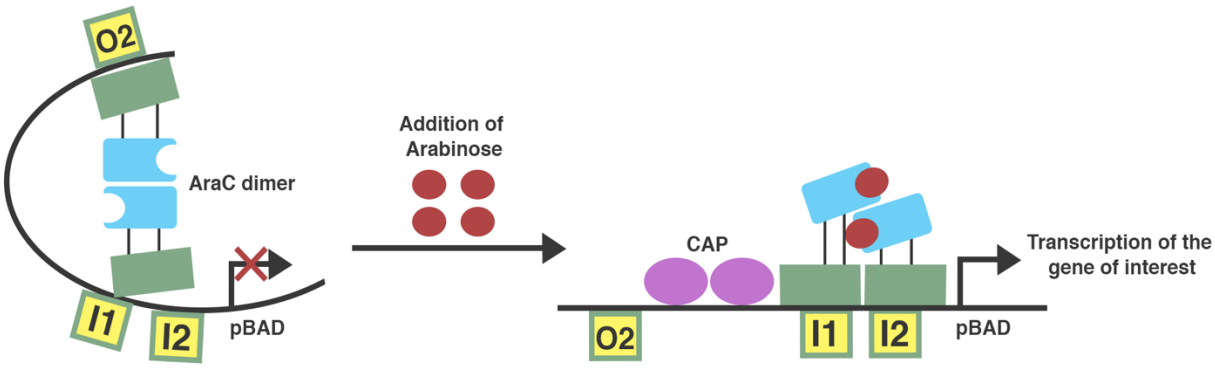


Figure 2. The pBAD expression system. In the absence of arabinose, *Ara*C binds to both O2 and I1 preventing transcription of the gene of interest. Once arabinose has been introduced to the media, it binds to *Ara*C causing it to bind to the *ara* promoter region, I1, and I2, resulting in downstream expression of the gene of interest. Binding of cAMP also forms a complex with cAMP activator protein (CAP) which can also result in the *ara*C promoter binding to I1 and I2.

Once the protein of interest has been successfully overexpressed and membrane extraction has taken place, the protein must be solubilized and purified. Typically this is done using detergents to solubilize the protein, since they are convenient and effective at solubilizing membrane proteins[75]. However, detergents are poor at mimicking the membrane proteins natural environment, and as such, they can destabilize and inactivate membrane protein activity over time[75]. In order to overcome some of the challenges encountered when solubilizing with conventional detergents, a novel technique, using styrene maleic acid (SMA) to extract and purify membrane proteins, has been developed, which negates the use of standard detergents in this process[76]. SMA purification involves the use of self-assembling SMA lipid polymers to extract membrane proteins whilst keeping them in their local lipid environment. This enables the proteins to retain function and stability throughout the purification process and downstream analysis[76]. The alternating hydrophilic nature of maleic acid and hydrophobic styrene result in an amphipathic SMA polymer, capable of inserting into biological membranes, where it can encircle and extract small discs of membrane encapsulating an integral membrane protein[76]. These discs, often referred to as styrene maleic acid lipid polymers (SMALPs) are water soluble, and can therefore be purified using standard affinity chromatography techniques, providing the protein of interest contains a suitable tag. Therefore, SMA solubilization can be used to optimize protein purification without too much adjustment to the protein purification protocol.

Currently, most research that has been carried out on the DedA family has been done *in vivo*. However, this comes with many problems for functional studies as it can be hard distinguishing individual protein function from that of other surrounding membrane proteins within the cell[44]. The aims of this research were as follows:

1. Clone a Panel of DedA genes from various organisms.
2. Assess their function by developing an *in vivo* functional assay.
3. Perform expression trials on clones which produce functional proteins.
4. Overexpress and purify functional proteins using both detergent based and SMA based methods.

5. Characterise the protein by developing an *in vitro* functional assay which would assess internal pH, using a pH sensitive fluorophore such as pyranine, as an indicator of proton movement[45].

2. Materials and Methods

2.1. Growth Conditions

Cells were grown in Lysogeny Broth (LB) medium containing 10 g/L NaCl (Fisher), 10 g/L Tryptone (Oxoid), 5g Yeast Extract l^{-1} (BD Biosciences), with the addition of 1% Bacto agar (Duchefa Biochemie) for solid cultures. All media was supplemented with the following antibiotics where indicated (100µg Ampicillin ml^{-1}, 30µg Kanamycin ml^{-1}). Cultures were also treated with 0.1% L-arabinose (w/v) or 1mM Isopropyl-β-D-thiogalactoside (IPTG) unless indicated otherwise. All cultures were grown at 37°C unless stated otherwise.

2.2.Molecular Weight Marker

Overnight cultures of Strain DH5*α* containing plasmids pPSU1 and pPSU2 individually were grown overnight, from glycerol stocks, under the conditions outlined above, supplemented with ampicillin as indicated. The plasmids pPSU1 and pPSU2 were then isolated using the QIAprep Spin Miniprep Kit (Qiagen). The 100 base pair DNA ladder was then prepared using a previously described protocol[46].

2.3.Clone Development

2.3.1 Restriction Enzyme Based Cloning Into a pBAD Expression Vector.

All *E. coli* genes were amplified from strain W3110 using either colony PCR or genomic DNA isolated from W3110 overnight cultures using GeneJET genomic DNA purification kit (Thermo Scientific). *bbDedA*, *vcDedA* and *saDedA* were amplified from a modified pET vector[47] containing *Borrelia burgdorferi, Vibrio cholerae, and Staphylococcus aureus* DNA synthesised and codon optimised using Invitrogen geneArt (stored at concentration at -20°C). The primers used are shown in Table 1. PCR was completed using KOD Hot Start DNA

Polymerase (Novagen). Following this, confirmation that PCR had worked was conducted by running the PCR products on a 1% agarose gel alongside the Penn State DNA ladder mentioned previously. The PCR products were then purified using the QIAquick PCR Purification Kit (Qiagen). A modified pBAD expression vector was digested using NcoI and KpnI (New England Biolabs), and run on an agarose gel. The vector was then extracted using the QIAquick gel extraction kit (Qiagen). Once purified, the concentration of the vector and amplified gene at an insert to vector molar ratio of 3:1 and 5:1 and ligated using T4 DNA ligase. See heat shock transformation below.

2.3.2 Gibson Assembly Cloning of pBADHis*yghB*.

YghB was amplified from strain W3110 genomic DNA isolated from W3110 overnight cultures using GeneJET genomic DNA purification kit (Thermo Scientific). The vector pBADHis*gltPh* was isolated using the QIAprep Spin Miniprep Kit (Qiagen) from an overnight culture. PCR was completed using KOD Hot Start DNA Polymerase (Novagen). 0.5 µl of Dpn1 was added to PCR products and vectors which were then incubated overnight at 37°C. Both DNA fragments were then joined using Gibson assembly®[48]. 1.25 µl of insert and 1.25µl of vector were then added to 5 µl of Gibson reaction mix taken from a 1.2 ml aliquot which contained 320 µl of 5x Isothermal (ISO) reaction buffer[48], 0.64 µl of 10 Units/µl T5 exonuclease[48], 20 µl of 2 Units/µl Phusion polymerase[48], 160 µl of 40 Units/µl Taq ligase[48], and 700 µl milliQ water. This was done for 60 minutes at 50°C. These ligations were then transformed into competent TOP10 cells. See heat shock transformation.

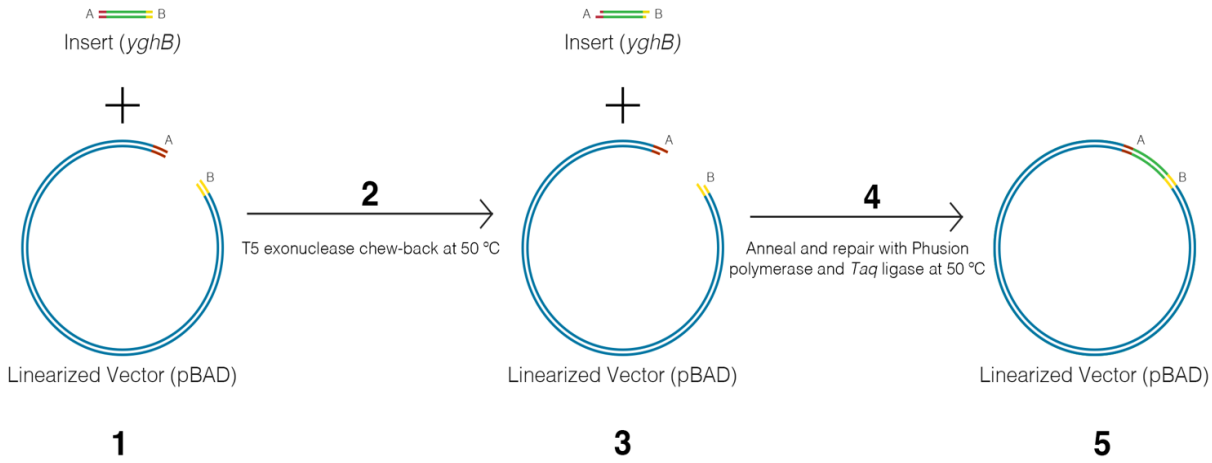


Figure 3. Gibson Assembly Cloning of pBADHis*yghB*. The first step (indicated by 1) involved the amplification of *yghB* from strain W3110 genomic DNA as well as amplification of a linearized pBADHis vector, both with overlapping regions (indicated by A and B). The Remaining steps took place as part of the Gibson Assembly reaction. 1.25 µl of the linearized vector and 1.25 µl of insert were then added to 5 µl of Gibson reaction mix taken from a 1200 µl aliquot containing 320 µl of 5x Isothermal (ISO) reaction buffer[48], 0.64 µl of 10 Units/µl T5 exonuclease[48], 20 µl of 2 Units/µl Phusion polymerase[48], 160 µl of 40 Units/µl Taq ligase[48], and 700 µl milliQ water. Following this the above mixture was incubated for 60 minutes at 50ºC, over which time the T5 exonuclease degrades the linearized DNA of the vector and insert in a 5' to 3' direction, degrading the homologous regions and leaving a 3' overhang. Following this, the insert and vector anneal, Phusion polymerase seals the gaps with nucleotides, and *Taq* ligase then covalently bonds the annealed fragments.

2.3.3 Restriction Enzyme Based Cloning Into a pET Expression Vector.

All *E. coli* genes were amplified from strain W3110 using either colony PCR or genomic DNA isolated from W3110 overnight cultures using GeneJET genomic DNA purification kit (Thermo Scientific). The primers used are shown in Table 1. PCR was completed using KOD Hot Start DNA Polymerase (Novagen). Insert and vector were both digested using AscI and XhoI (New England Biolabs). Following this, confirmation that PCR had worked was conducted by running the PCR products on a 1% agarose gel alongside the Penn State DNA ladder mentioned previously. The PCR products were then purified using the QIAquick PCR Purification Kit (Qiagen). Once purified, the concentration of the vector and insert were then measured using a nanodrop 2000c spectrophotometer (Thermo scientific) and ligated at a ligand to vector molar ratio of 3:1 and 5:1 using T4 DNA ligase with an empty pET/His vector which had also been digested with AscI and XhoI and extracted using the QIAquick Gel Extraction Kit (Qiagen). See heat shock transformation below.

Primer Name	Vector	Primer Sequence
kpDedA forward	pBAD	GAATTAACCATGGATCTGATTCATTTTCTGATTGATTTC
kpDedA reverse	pBAD	ACCAGGGTACCTTTGGCCTGTTTCGCTGCT
vcDedA forward	pET	TCAGGGGCGCGCCATGGAAGCCGCTTTTAG
vcDedA reverse	pET	TGGTGCTCGAGTATTATGCGGTTTCCTTAATCATTTTGG
vcDedA forward	pBAD	TTAACCATGGAAGCCGCTTTTAGCGAATGG
vcDedA reverse	pBAD	ATTAGGGTACCTGCGGTTTCCTTAATCATTTTGG
saDedA forward	pET	TCAGGGGCGCGCCATGGAACAAATTATCACTG
saDedA reverse	pET	TGGTGCTCGAGTATTATTTTACTCTCTTACGTTTGC
saDedA forward	pBAD	TTAACCATGGAACAAATTATCACTGAATTTATTAGCC
saDedA reverse	pBAD	ATTAGGGTACCTTTTACTCTCTTACGTTTGC
paDedA4 forward	pBAD	GAATTAACCATGGATTTTAATCCGATTGATCTGATTC
paDedA4 reverse	pBAD	ACCAGGGTACCATCGCTCTGTGCTTTTGC
paDedA5 forward	pBAD	ATTAACCATGGCGATGCTGCAACAGTTTCTGC
paDedA5 reverse	pBAD	CACCAGGGTACCTTCGCTTTTATCTGCATTACC
bbDedA forward	pBAD	GAATTAACCATGGCGACCAAAATGTATATTAACACCATC
bbDedA reverse	pBAD	ACCAGGGTACCTTTCAGGTTTTTATCCAC
yqjA forward	pET	TCAGGGGCGCGCCATGGAACTTTTGACCCAATTG
yqjA reverse	pET	TGGTGCTCGAGTATTACCCCCGATTTCCATATTTC
yqjA forward	pBAD	GAATTAACCATGGAACTTTTGACCCAATTGCTGC
yqjA reverse	pBAD	ACCAGGGTACCCCCCCCGATTTCCATATTTC
yghB forward	pET	TCAGGGGCGCGCCATGGCTGTTATTCAAGATATC
yghB reverse	pET	TGGTGCTCGAGTATTAGGCGTTACAGTATTTTTTTTTAATCACCAC
yghB GA vector forward	pBAD	GGTACCCTGGTGCCGCGCGG
yghB GA vector reverse	pBAD	CATGGTTAATTCCTCCTGTTAGCCCAAAAAA
yghB GA insert forward	pBAD	TTGGGCTAACAGGAGGAATTAACCATGGCTGTTATTCAAGATATCATC
yghB GA insert reverse	pBAD	GACCGCTGCCGCGCGGCACCAGGGTACCGGCGTTACAGTATTTTTT

Table 1. The sequence of each primer used throughout this research alongside the vector being utilised. All yghB GA primers were used for Gibson Assembly cloning.

2.3.4. Sequencing

The accuracy of the cloned gene sequences were checked by sequencing (GATC) (Appendix fig. 1 - 16).

All sequencing data was analysed using ApE to read ABI sequencing trace files and conduct sequence alignment to make sure the constructs were correct (Appendix fig. 1 – 16).

Strain	Plasmid	Defining Characteristics	Source
TOP10, MCΔΔ	pBADHis*gltPh*	C-Terminal His-Tag; Amp^r	Mulligan, Chris.
	pBADHis*yqjA*		This Study
	pBADHis*yghB*		This Study
	pBADHis*paDedA4*		This Study
	pBADHis*paDedA5*		This Study
	pBADHis*bbDedA*		This Study
	pBADHis*kpDedA*		This Study
	pBADHis*vcDedA*		This Study
	pBADHis*saDedA*		This Study
BL21-AI, MCΔΔ	pETHis*yqjA*	N-Terminal His-Tag; Kan^r	This Study
	pETHis*yghB*		This Study
	pETHis*paDedA4*		Mulligan, Chris.
	pETHis*paDedA5*		Mulligan, Chris
	pETHis*bbDedA*		Mulligan, Chris
	pETHis*kpDedA*		Mulligan, Chris
	pETHis*vcDedA*		This Study
	pETHis*saDedA*		This Study

Table 2. The panel of clones used in this study, their source, and the strains they were transformed into.

2.4. Competent cells

4 ml overnight cultures of TOP10 were prepared from glycerol stocks. 4ml LB was inoculated with *E. coli* TOP10 from a glycerol stock and grown overnight at 37°C in a shaking incubator. 1ml of the overnight culture was used to inoculate four 500 ml conical flasks containing 100 ml LB each and grown at 37°C until and OD_{600} of 0.4 was reached. The cells were incubated on ice for 20 minutes, then harvested by centrifugation at 3000xg for 10 min at 4°C. The cells were resuspended in 7.5ml 0.1 M $CaCl_2$, incubated on ice for 30 min, and harvested by centrifugation using the same conditions as above. The cells were resuspended in 8 ml 0.1 M $CaCl_2$ and 15% (v/v) glycerol, divided into 200 µl aliquots, flash frozen and stored at -80°C.

2.5. Transformation

2.5.1 Freeze Thaw Transformation

5 ml liquid culture of the desired strain was grown overnight per the growth conditions outlined above. Following this, 4.5 ml LB is inoculated with 500 µl of the overnight culture and incubated for 60 minutes at 37°C in a shaking incubator. Cells were then harvested by

centrifugation at 13000 xg for 5 minutes and the pellet was resuspended in 600 µl of cold 50 mM $CaCl_2$. The cell suspension was divided into 100 µl aliquots in sterile eppendorf tubes and kept on ice. 1 µl of plasmid DNA was added to the 100 µl cell suspension and vortexed to mix. The cell suspension was snap frozen on ethanol and dry ice mixture for 1.5 minutes. The sample was then thawed at 37°C in a heating block for 2 minutes. 1 ml of LB was added to the cell suspension and incubated for 60 minutes at 37°C in a shaking incubator. 100 µl of the cell suspension was spread on a selective plate and incubated overnight at 37°C.

2.5.2 Heat Shock Transformation

Chemically competent cells were first removed from -80°C and thawed on ice. 1 µl of plasmid DNA was added to 50 µl competent cell suspension and mixed using the pipette tip. The cell suspension was incubated on ice for 30 minutes. The mixture was then heated at 42°C for 30 seconds and cooled on ice for 2 minutes. 200 µl of SOC medium was then added to the cell suspension and the mixture is incubated at 37°C for 60 minutes. After incubation, 10 µl and 190 µl of the cell suspension were then plated separately onto selective plates and grown overnight according to the conditions described in section 2.1.

2.6. Development of an In-vivo Functional Assay

2.6.1. Temperature Sensitivity Assay

5 ml LB cultures of each strain supplemented with 500 µg ampicillin (shown in Table 6) were grown overnight at 37°C in a shaking incubator. The following day, 500 µl of overnight culture was used to inoculate a further 5 ml of LB media of each strain also supplemented with 500 µg ampicillin and grown for a further 4-6 hours. The optical density (OD) of each culture was assessed by measuring the absorbance at 600 nm using a nanodrop 2000c spectrophotometer (Thermo scientific). The OD_{600} values were then normalised to account for a difference in cell count between cultures. Each culture was serially diluted 1:10 four times so that the final dilution is 1:10000. 5 µl of each culture was spotted onto plates consisting of LB supplemented with 100µg Ampicillin and 0.1% L-arabinose (w/v). A control plate also used to ensure that any perturbed growth was the result of growth at a higher temperature. The plates were incubated overnight at both 30°C, for the control plate, and 44°C for the other.

2.6.2. Ethidium Bromide Sensitivity Assay

5 ml cultures of each strain supplemented with 500 µg ampicillin were grown overnight at 37°C in a shaking incubator. The following day, 500 µl of overnight culture was used to inoculate a further 5 ml culture of each strain, also supplemented with 500 µg ampicillin and grown for a further 4-6 hours. The OD_{600} of each culture was then assessed. The OD_{600} values were then normalized. Each culture was then serially diluted 1:10, four times so that the final dilution was 1:10000. 5 µl of each culture was then spotted onto plates consisting of LB, as described above, supplemented with 100µg Ampicillin, 0.1% L-arabinose (w/v) and 75 µg ml^{-1} ethidium bromide, unless stated otherwise. A control plate also consisting of LB supplemented with 100µg Ampicillin and 0.1% L-arabinose (w/v) was also used to ensure that any perturbed growth was the result of the ethidium bromide. The plates are then incubated for 22-24 hours at 30°C.

2.7. Protein Expression

Colonies of each strain were used to inoculate 5 ml day cultures supplemented with 500 µg of ampicillin for strains containing pBAD expression vectors and 150 µg of kanamycin for strains containing pET expression vectors, which were grown throughout the day at 37°C in a shaking incubator for ~2.5 hours. The OD_{600} was measured of each culture, and 5ml of each culture was used to inoculate a larger culture of each strain also supplemented with ampicillin or kanamycin. Cultures were grown overnight per the growth conditions outlined in section 2.1. The following morning the OD_{600} values were measured for each culture and 20 ml of each culture was used to inoculate a 1 L culture. These cultures were grown per the previously described growth conditions outlined in section 2.1 above for ~ 2.5 hours until an OD_{600} of 0.6-0.8 was reached. Following this, 1 mM IPTG was added to cultures containing a pET construct and 0.1% L-arabinose was added to cultures containing a pBAD construct. The cultures were then incubated under the same growth conditions for a further 2 hours. Cells were harvested by centrifugation for 20 minutes at 4000xg, the supernatant was removed and the pellets were then resuspended in 20 ml lysis buffer per litre of cells harvested, and stored at -20°C.

2.8. Membrane Preparation

Cellular pellets were removed from -20°C storage and thawed at room temperature. Once thawed, the samples were kept on ice. Each sample was then placed into a metal beaker and placed into an ice bath. The samples were then sonicated at 8.2 amplitude for 3 seconds on 7 seconds off for 10 minutes per sample using a Soniprep 150 plus sonicator (MSE). The lysate was then poured into 30ml centrifuge tubes, balanced, and centrifuged at 20,000xg for 20 minutes in an Avanti J-25 centrifuge (Beckman Coulter). The supernatant was then poured into Ti70 or Ti45 ultracentrifuge tubes (Beckman Coulter) and centrifuged at 41000 rpm for 90-120 minutes in an Optima L-90K ultracentrifuge (Beckman Coulter). The supernatant was then poured off and the pellet was resuspended in 2.5ml Buffer A (containing 50 mM Tris pH 8, 100 mM NaCl, and 5% glycerol) per litre of cells centrifuged. The resuspended pellets were then stored overnight at -20°C before being purified the following day.

2.9. Protein Purification

2.9.1. n-Dodecyl β-D-Maltoside (DDM) Solubilisation and Purification

The resuspended pellets were thawed at room temperature. Once thawed, the samples were immediately placed on ice. Membranes were solubilised by adding DDM to a final concentration of 1% (w/v). The samples were then incubated for 40 minutes at 4°C with agitation. The samples were then placed in Ti70 ultracentrifuge tubes (Beckman Coulter) and centrifuged for 30 minutes at 170,000xg in an Optima L-90K ultracentrifuge (Beckman Coulter). Once centrifuged, 2 ml Talon superflow resin containing 50% resin suspended in ethanol was added to the supernatant from each sample and bound for 1 hour at 4°C. The samples were poured through 10ml poly-prep chromatography columns (Bio-rad) and the flow through was collected. The column was washed with wash buffer, containing 20 ml Buffer A (previously described), 50 µl 4 M imidazole pH 8 (10 mM final concentration), and 0.2 ml 10% DDM (w/v) (0.1% final concentration), and collected. Following this, the protein was eluted and collected in 4-10 fractions and either stored overnight at 4°C, run on an SDS PAGE gel or run on a SEC column.

2.9.2. Styrene Maleic Acid (SMA) Solubilisation and Purification

The resuspended pellets were thawed at room temperature. Once thawed, the samples were immediately placed on ice. 3% SMA (w/v) was then added to each sample. The samples were then incubated for 120 minutes at 4°C with agitation. The samples were then placed in Ti70 ultracentrifuge tubes and centrifuged for 30 minutes at 170,000xg in an Optima L-90K ultracentrifuge (Beckman Coulter). Once centrifuged, 2 ml Talon superflow resin containing 50% resin suspended in ethanol was added to the supernatant from each sample and bound for 1 hour at 4°C. The samples were poured through 10ml poly- prep chromatography columns (Bio-rad) and the flow through was collected. The column was washed with wash buffer, containing 20 ml Buffer A (previously described), and 50 µl 4 M imidazole pH 8 (10 mM final concentration), and collected. Following this, the protein was eluted and collected in 4-10 fractions and either stored overnight at 4°C, run on an SDS PAGE gel or run on a SEC column.

2.10. Size Exclusion Chromatography (SEC)

A superdex 200 10/300 GL column was first equilibrated with buffer. The buffers used consisted of 20 mM Tris, pH 8, 100 mM NaCl, 5% Glycerol, and 0.1% DDM (for samples solubilized with DDM) and 20 mM Tris, pH8, 100 mM NaCl, and 5% Glycerol for samples eluted with SMA. Once the column was equilibrated the sample was then run. The sample was taken up into a 1ml syringe, loaded onto the column, and run. After the sample had been run, the column was washed with 50 ml milli Q water followed by 50 ml 20% EtOH which had both been degassed. This column had a void volume of 8.2 ml.

2.11. SDS PAGE Gels

The Resolving component of each SDS PAGE gel was made using 1.4ml milliQ water, 1ml 1.875M tris HCl (pH 8.85), 50 µl 10% SDS, 2.5ml 30% Acrylamide, 5 µl temed, and 5 µl APS (v/v). The Stacking component of each SDS PAGE gel was made using 1.8ml milliQ water, 150 µl 1M tris HCl (pH 6.8), 25 µl 10% SDS, 440 µl 30% acrylamide, 3.5 µl temed, and 20 µl 15% APS (v/v). The protein samples were first thawed or removed from 4°C and 4x SDS sample

buffer was added to each sample. A biorad gel tank was then filled to the correct marker with SDS protein gel running buffer (PGRB). The samples are then loaded onto the gel alongside the protein ladder. Once all the samples have been loaded correctly the lid is placed on the gel tank and plugged into a biorad gel tank and run at 60 mA for 2 gels or 40 mA for one gel. Both are run for 40 minutes.

2.12. Gel Staining

Once the gels had either finished running or, if being used for western blotting, finished transferring, the buffer was poured off and the gel removed. The gel was then placed into a box and stained with Coomassie blue stain containing 50% methanol, 50% acetic acid, and 0.2% Coomassie blue, for approximately 1 hour, unless stated otherwise, at room temperature.

2.13. Western Blotting

Once the gels had finished running, the buffer was poured off and the gel was removed from the tank. Some whatman filter paper (GE life sciences) cut to the same dimensions as the cassette was placed into the cassette alongside some foam pads (Biorad) and soaked in western blotting buffer, as previously described. Some Amersham™ Hybond™ P0.45µm PVDF membrane (GE Healthcare), cut to the same dimensions as the gel was then activated by soaking in methanol for a few minutes. The membrane and gel were then placed into the cassette. The cassette was then placed into an electrophoresis chamber (Biorad) with an ice pack, to prevent overheating, and is filled with western blotting buffer, and run for 60 minutes at 40V for one blot and 80V for two.

Once the transfer was complete the membrane was removed from the cassette. The membrane was then stained with 10 ml Ponceau Red for 15 minutes whilst being shaken on a gyrorocker. The stain was then poured off and the membrane was washed with water. Once washed, the membrane was then blocked with 20 ml blocking solution containing 5% milk powder in 1x TBST for 1 hour at room temperature on a gyrorocker. After 1 hour, the blocking solution was removed, and the membrane was washed with 1x TBST for 15 minutes, followed by 2 more washes for 5 minutes each. The membrane was then placed in a plastic bag with 5 ml primary

hybridization buffer containing 2.5 µl Anti-His [C-Term] primary antibody (Novex), in 1x TBST with 1% milk powder (Sigma Aldrich) and placed in a sealed plastic bag. The membrane was then placed on a shaker overnight at 4°C. The following day, the membrane was removed from the plastic bag, placed in a box, and washed with 1x TBST for 15 minutes followed by 3 more washes, each for 5 minutes at room temperature on a gyro rocker. 20 ml secondary antibody buffer containing 4 µl Stabilized Peroxidase Conjugated Goat Anti-Mouse [H + L] secondary antibody (Invitrogen) in 1x TBST with 1% milk powder (Sigma Aldrich) was then added to the membrane and shaken for no more than 1 hour at room temperature. Following this, the secondary buffer was then poured off and the membrane was washed, again, in 1x TBST for 1x 15-minute wash, followed by 6 more washes, each 5 minutes long. The membrane was then left in 1x TBST until it is ready to be imaged. A SuperSignal West Femto Maximum Sensitivity Substrate ECL kit (Thermo Scientific) was then used per manufacturer's instructions. The membrane was then placed in a Chemidoc XRS+ (Biorad) and the chemi option was selected under blot. The blot was imaged for 300 seconds with an image being taken every 3 seconds. An image was selected. Another image of the membrane was taken on the colorimetric setting, to visualise the molecular weight markers, and the two images were merged.

3. Results

3.1.Establishing a Library of Clones

When beginning to establish a library of clones, both pET and pBAD expression systems were utilised, since both systems are useful for different reasons. For example, the pBAD expression system contains araBAD, a tightly regulated promoter region, which is beneficial as, unlike pET, pBAD will result in less uncontrolled expression, which is particularly problematic for membrane proteins [49]. Overexpression of membrane proteins in the cytoplasmic membrane is required to produce enough protein for structural and functional studies. However, membrane protein overexpression usually results in the accumulation of these proteins in the membrane, which is toxic to the cell and reduces the amount of overall expression of the protein of interest. This is due to Sec translocon capacity, of which there are two effects, the aggregation of protein precursors in the cytoplasm, and shifted, inefficient energy metabolism[50]. Despite the obvious advantages of utilising the pBAD expression system for the production of recombinant

membrane proteins, the pET expression system also has many advantages for the production of recombinant proteins making it one of the most widely used expression systems for the production of recombinant proteins[51]. Some such examples of these advantages include the strong activity from the T7 promoter region which can result in a significant amount of recombinant protein being produced, in some cases over 50% of the total amount of protein produced by the cell can be the recombinant protein[52]. It was yet unclear which expression system would yield a higher amount of protein, since these systems have several major differences. One major difference is the location of his tag, which was previously thought to have a low impact on the structure and function of the protein. However, recent evidence is beginning to show that this is not the case[53]. It has been demonstrated that the location of the his-tag, either N-terminal or C-terminal can have a large impact on the structure and other properties of the protein, such as thermal stability[54,55]. With pET containing an N- terminal tag and pBAD a C-terminal tag it was possible that the location of these different his- tags could impact the expression, structure and function of these proteins[55]. Therefore, various conditions were trialled by cloning each DedA homologue into both expression vectors (Table. 2).

Several DedA homologues from various organisms such as *E. coli*, *K. pneumoniae*, *P. aeruginosa*, and *B. burgdorferi* were selected for this research, since some of these proteins have been shown to restore growth in BC202, meaning testing whether these proteins are functional would be straightforward. These genes were then cloned using restriction enzyme-based cloning and Gibson assembly into both pET and pBAD expression systems. This was done in the hope that one of these expression systems would be capable of sufficiently expressing the protein under the growth conditions outlined in section 2.1. The T7 polymerase pET expression system is one of the most widely used expression systems with >90% of protein expression systems in the protein data bank consisting of pET constructs[56]. The pBAD expression system, which, due to the positive control of the araBAD promoter region, as opposed to negative control of expression of pET, can be particularly useful for the expression of membrane proteins. This is because of the graded expression which allows overexpression to a level that becomes it does not become toxic to the cell due to inclusion body formation and protein aggregation[56].

Restriction enzyme-based cloning was used to insert all DedA genes (Table. 2) excluding pETHis*bbDedA*, pETHis*kpDedA*, pETHis*paDedA4*, and pETHis*paDedA5*, which had already been established, into both modified pBAD and modified pET vectors[47]. The first step in restriction enzyme-based cloning requires the amplification of the gene of interest from a template strand of DNA such as genomic DNA. Utilising a technique called polymerase chain reaction (PCR), the gene of interest will be replicated. After this, each PCR product run on an agarose gel alongside a molecular weight marker to confirm whether the PCR had correctly amplified the gene of interest as indicated by clear, defined, bands at the correct molecular weight for the gene. This is particularly important as PCR often requires optimisation by changing factors such as the annealing temperature. Each amplified gene product was run on an agarose gel to confirm that PCR had worked (Fig. 4.).

Restriction endonucleases AscI and XhoI were then used to cleave the ends of the gene of interest and the modified pET vector at the corresponding regions in the multi-cloning site (MCS). This same process was repeated using NcoI and KpnI for the gene of interest and modified pBAD vector at the corresponding regions in the MCS. This creates complementary ends which can be joined together by T4 DNA ligase. However, before ligation can occur, the restriction endonuclease digested vector is run on an agarose gel to separate out the vector from the region that has been removed. The band containing the digested vector is then removed using a technique called gel extraction which purifies the digested vector from the gel.

Alongside this, a technique called PCR purification is used to remove the components from the PCR mixture such as restriction enzymes, buffers, nucleotides, and primers. Once this has been completed, T4 DNA ligase alongside the vector and insert are mixed at a specific molar ratio, in this case a vector to insert ratio of 1:3 and 1:5 were used to ensure that the ligation worked. However, this process required a significant amount of optimisation to function correctly. This could have been caused by several factors such as a lack of sufficient insert or vector, since a large amount of each is lost during PCR clean-up and gel extraction, incorrect molar ratio, inappropriate buffer, or UV damaged vector due to the length of time spent on gel extraction.

The genes of interest were inserted in-frame with an N-terminal his-tag into a modified pET vector[47] as well as being inserted in-frame with a C-terminal his-tag into a modified pBAD

vector. Both modified pET and pBAD vectors used for cloning had previously been used to overexpress other membrane proteins.

The gene encoding *yghB* was found to possess a NcoI cut site, precluded the use of restriction enzyme-based cloning since this would have cleaved the cut site within the gene. However, the use of a novel method for molecular cloning called the Gibson assembly method would allow the in-frame insertion of *yghB* into the modified pBAD with a C-terminal his-tag without the use of restriction endonucleases (Fig. 3.).

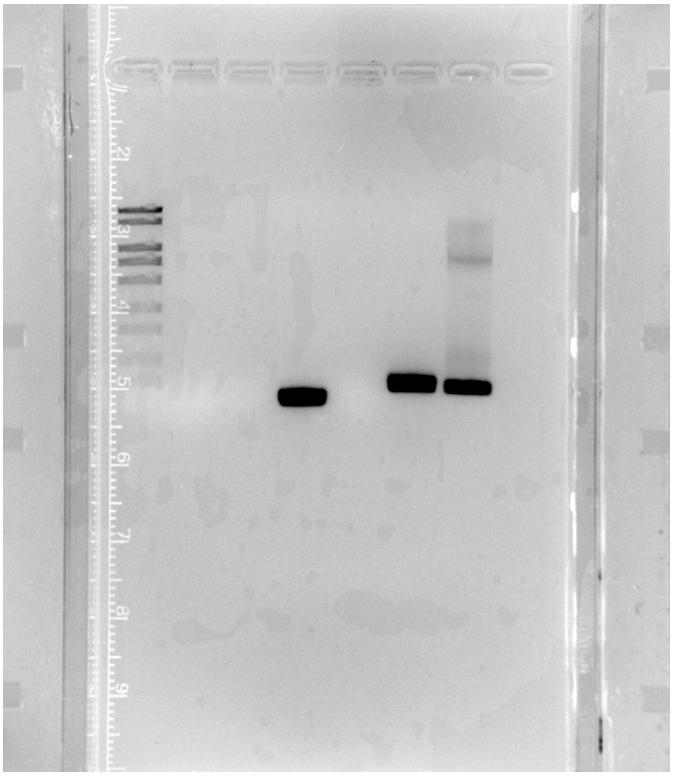


Figure 4. This is an example of an agarose gel with PCR products run on it, and demonstrates what is expected to be observed when PCR is successful.

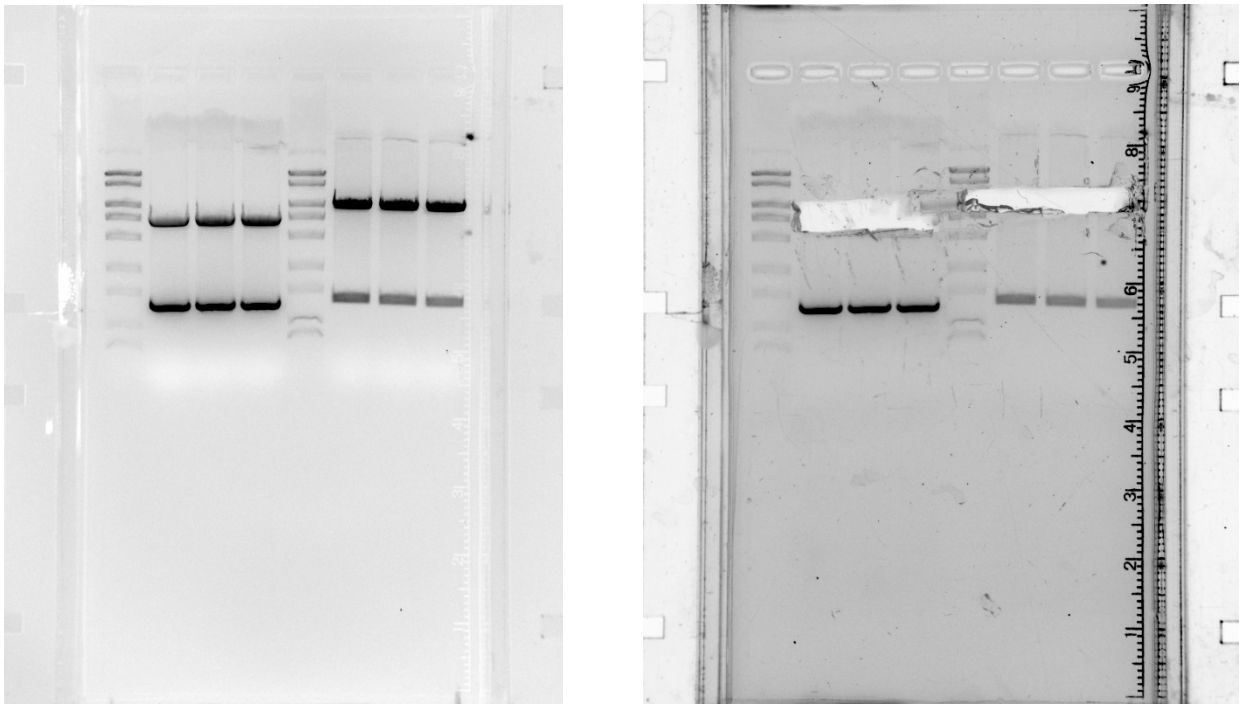


Figure. 5. Bands excised in the process of gel extraction. As shown, bands are excised by cutting them from the gel. This can be an inefficient process.

3.2. Development of an *In-vivo* Functional Assay

Once all the genes had been cloned into a modified pBAD expression vector and sequenced, it was necessary to find a way of measuring whether they were able to produce functional protein, since this was necessary for large-scale expression and purification. This was made possible by the temperature sensitive phenotype observed in the double knockout strain *E. coli* MC4100 *ΔyqjA::ΔyghB* (MCΔΔ)[26]. Since the observed phenotype, when the double knockout strain is grown at an elevated temperature of 44°C, is a complete lack of growth due to unknown factors[26]. It has been shown that it is possible to restore growth at above 42°C in this strain when certain DedA homologues are overexpressed including several of the genes in this study, such as *yqjA*[27], *yghB*[27], *bbDedA*[15], *paDedA4*, and *paDedA5*. Therefore, by transforming the double knockout with each construct (Table 2.) and overexpressing the protein of interest it was possible to assess whether the protein is functional. To achieve this, an assay using the temperature sensitive phenotype observed in MCΔΔ was utilised to gauge heterologously expressed protein function.

In this assay, a control plasmid containing the protein, GltPh, was utilised to control for any affect on the restoration of growth above 44ºC which the presence of any overexpressed recombinant membrane protein might have had on the growth of the double knockout strain, MCΔΔ. GltPh is a glutamate transporter isolated from *Pyrococcus horikoshii* which transports aspartate, along with three Na^+ ions, into the cytoplasm[77]. Alongside this, GltPh also displays stoichiometrically uncoupled Cl^- conductance[77]. When establishing the aforementioned *in vivo* functional assay, a pBAD plasmid containing *gltPh*, denoted pBADHis*gltPh*, (Table 2.) was transformed into MC4100 wt and MCΔΔ. This

3.2.1.Temperature Sensitivity Assay

Since previous studies have shown that it is possible to restore growth at an elevated temperature (44ºC) in the MCΔΔ by inducing an expression vector containing various DedA homologues[26]. It was possible to exploit this phenotype and develop a method to assess whether the proteins, when overexpressed, are functional. Various DedA homologues were overexpressed using a modified pBAD expression vector with a C-terminal his-tag in the strain MCΔΔ alongside both MC4100 transformed with pBADHis*gltPh* and MCΔΔ transformed with pBADHis*gltPh*. A pBAD expression vector was used since pET vector expression requires the use of a strain lysogenized by a DE3 phage fragment56 such as BL21(DE3). Following transformation, the cells were then serially diluted and spotted on select plates containing L-arabinose to induce expression. The plates were then be incubated at an elevated temperature of 44ºC as well as 30ºC.

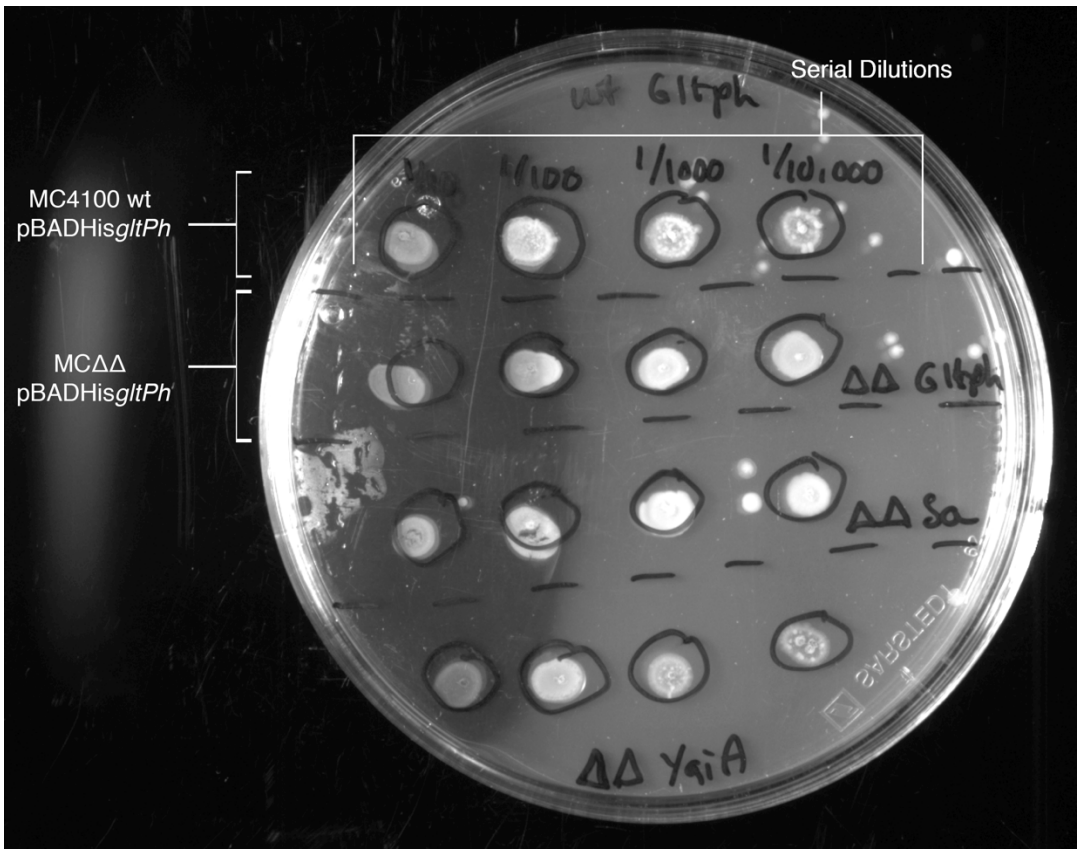


Figure 6. An LB Plate from the temperature sensitivity assays showing the continued growth of the double knockout strain MCΔΔ transformed with pBADHis*gltPh*, denoted MCΔΔ pBADHis*gltPh*, above 44ºC. Before spotting, each of the samples OD values were initially taken, after which they were normalized. Moving rightwards, each spot was serially diluted ten-fold further than the previous sample before 5 µl of which was spotted.

As shown in figure 6, growth at 44ºC presented many problems including the fact that strain MCΔΔ continued to grow at elevated temperatures despite previous studies having shown this strain to stop growing at 44ºC. Other technical issues encountered include the agar drying out and melting making it much harder to interpret the results. To combat this, the concentration of agar was increased from 1% to 5% in 1% increments (fig. 9). However, although increasing the agar concentration did decrease the amount of drying and melting occurring in the plates, it also reduced visibility of the results and did not correct the issue of growth of MCΔΔ above 44ºC.

One possible explanation for the lack of temperature sensitivity of the MCΔΔ strain is that the *E. coli* strain utilised in this study was not a double-deletion mutant of the genes encoding for *yqjA* and *yghB*. To test this possibility and ensure that both genes were knocked out in the strain primers which bound 500 bp upstream and downstream of both the *yqjA* and *yghB* genes in the MC4100 genome were designed (fig. 7 & 8.). Genomic DNA (gDNA) extraction was used to

purify the genomic DNA from both cultures. PCR was then used to amplify the region of interest with the upstream and downstream primers. As shown below, if both *yqjA* and *yghB* are knocked out then the PCR product will be 1000 base pairs long. However, if the strain is not a double- deletion mutant for *yqjA* and *yghB*, then the PCR products will be approximately 1600 bp long. The PCR product size was determined by running them on an agarose gel alongside a molecular weight marker which would give an indication of the size of each band produced. Despite the continued growth at 44°C the MCΔΔ strain being used in this study was in fact the double-deletion mutant (fig 8.). This was particularly unusual, and the cause of this continued growth is unknown. Therefore, to develop an alternative *in vivo* functional assay, the increased sensitivity to xenobiotics such as ethidium bromide (EtBr) was exploited, as observed in MCΔΔ.

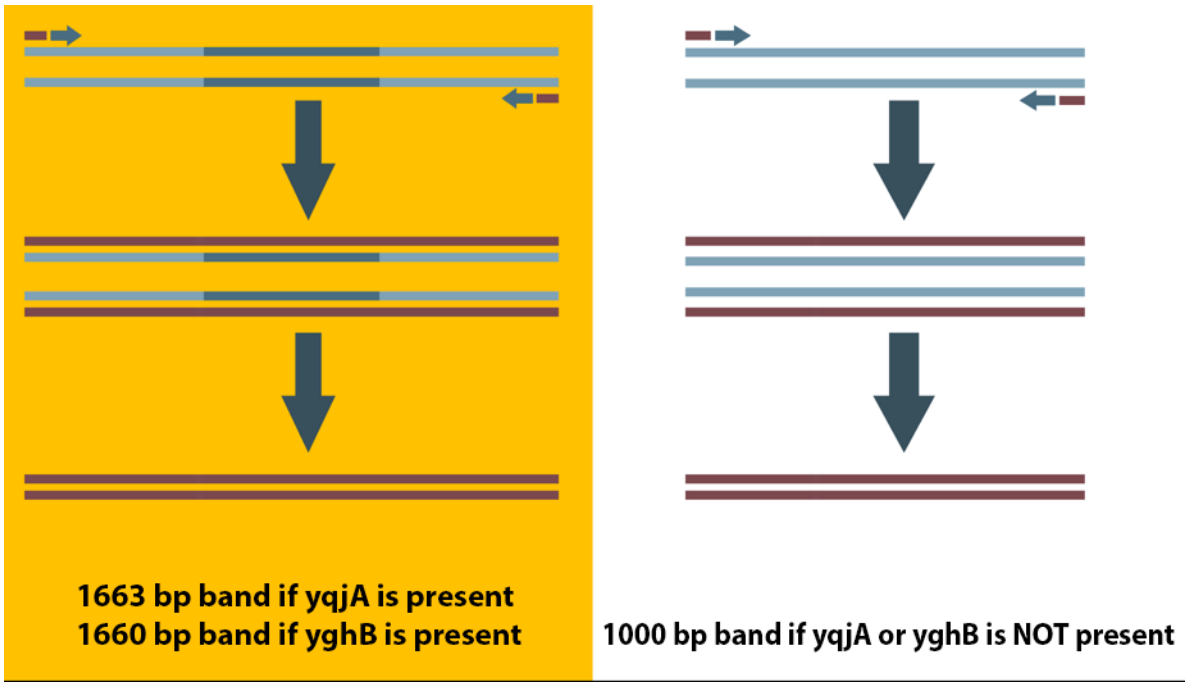


Figure 7. How PCR was utilised to check if the double knockout strain is as expected. As shown the primers will bind 500 bp upstream and downstream of the gene of interest (shown as a darker region). These will then elongate and form complementary fragments. If the gene of interest (*yqjA* and *yghB*) is present, then the band will be 1000 bp long, if not it will be 1663 bp for *yqjA* and 1660 bp for *yghB*. This can be confirmed by running the PCR products on an agarose gel alongside a molecular weight marker to confirm the size of the bands.

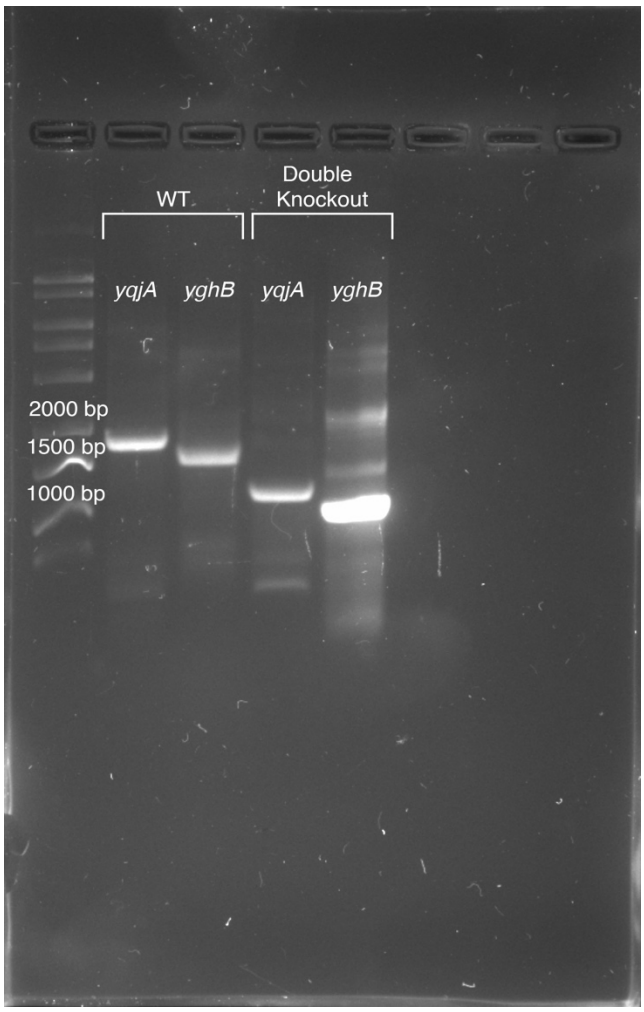


Figure 8. An agarose gel containing the PCR products from the PCR to confirm the double-deletion mutant phenotype. As observed, the double knockout is confirmed here by the presence of two bands of approximately 1000 base pairs in length. The wild type was used as a reference, as shown the bands produced by the presence of *yqjA* and *yghB* are approximately 1600 base pair in length.

3.2.2. Ethidium Bromide Sensitivity Assay.

The minimum inhibitory concentration (MIC) of ethidium bromide BC202 (W3110 *ΔyqjA::ΔyghB*), is 16-fold lower than that of the parent strain W3110[16]. Overexpression of YqjA in BC202 can restore the MIC to normal levels[16]. By growing MCΔΔ on medium containing ethidium bromide at a concentration above the MIC for MCΔΔ, and below that of the wild type. It is possible to assess growth of MCΔΔ, when overexpressing the protein of interest, as an indication of whether the proteins are functional, by using an *in vitro* assay, which utilised restoration of growth on plates containing EtBr as a way of assessing heterologously expressed protein function. This involved spotting serially diluted cultures of MC4100 transformed with pBADHis*gltPh*, MCΔΔ transformed with pBADHis*gltPh*, and

MCΔΔ containing each gene of interest (GOI) on a pBAD expression vector, onto selective plates which also contained ethidium bromide.

To determine the concentration of EtBr that would give the clearest difference in growth between the wild type and double knockout, a range of EtBr concentrations were screened, ranging from 25µg ml^{-1} to 75µg ml^{-1}. Initially, 25µg ml^{-1} appeared to be sufficient to halt growth of the knockout strain (fig. 9.). However, it became apparent when conducting repeat assays that a higher concentration of ethidium bromide was required to completely prevent growth, therefore the concentration was increased to 75 µg ml^{-1} EtBr. This new assay produced much clearer results (fig 10.), with all constructs being shown to be capable of restoring growth in the double knockout strain. Some constructs appeared to rescue better than others, however this could be due to the expression conditions (e.g. inducer concentration). Since overexpression of this panel is capable of restoring growth in the double knockout strain, this indicates that the proteins of interest are being produced and are functional in the membrane.

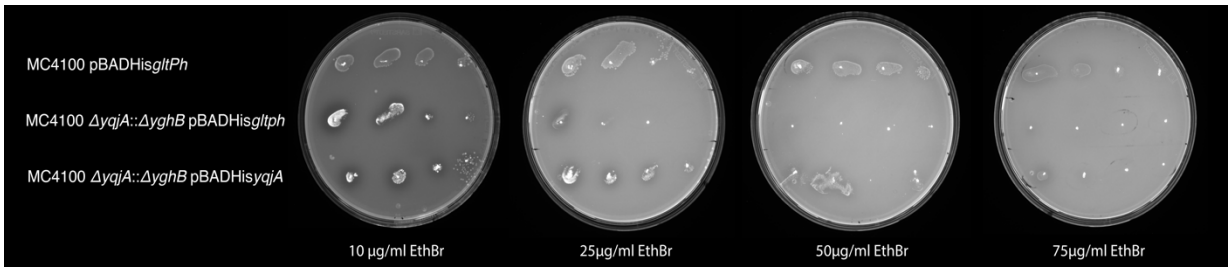


Figure 9. The determination of the correct ethidium bromide concentration to use for the ethidium bromide sensitivity assays. Each colony moving rightwards is a further 1/10 serial dilution. The initial spot is the normalized culture, undiluted. As shown, there is a distinct lack of colonies for concentrations above 75 µg ml^{-1}. This was not the case when continuing with the ethidium bromide assays and so 75 µg ml^{-1} EtBr was used, since it partially inhibited growth, but not completely. The top row consists of MC4100 transformed with the control plasmid, pBADHis*gltPh*. The middle row consists of MCΔΔ transformed with the control plasmid pBADHis*gltPh*. The bottom row consists of MCΔΔ transformed with pBADHis*yqjA*.

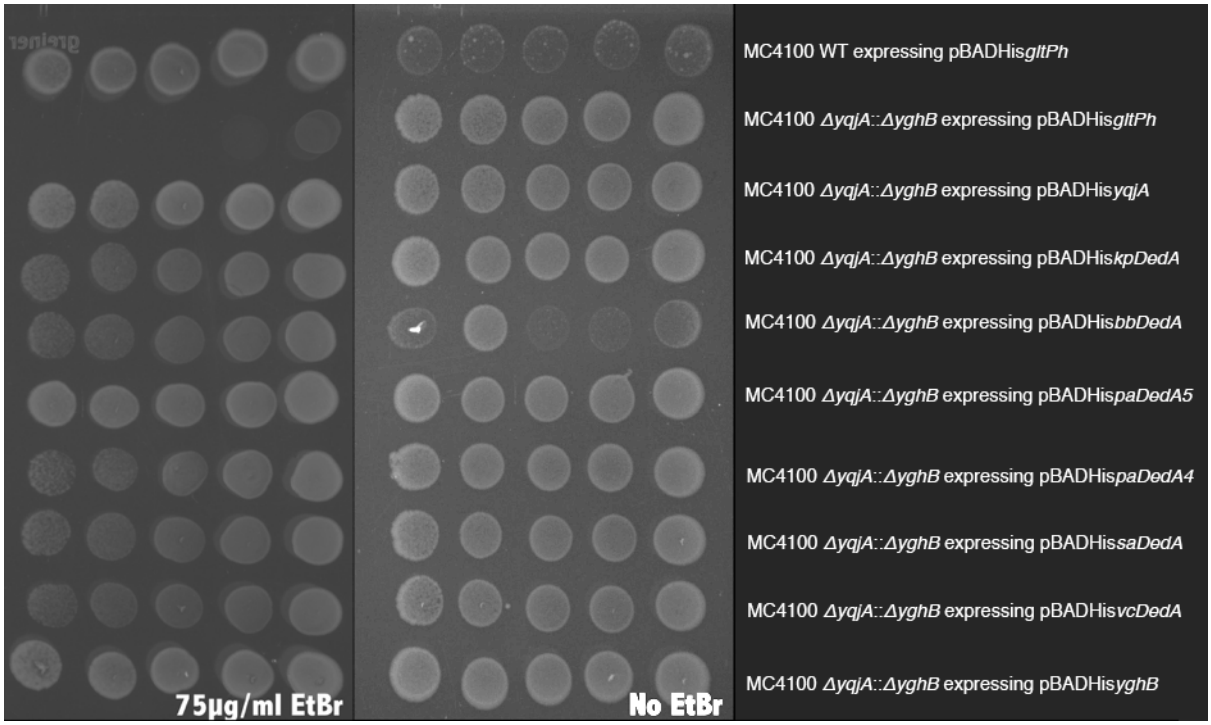


Figure 10. The ethidium bromide rescue assay alongside a control plate which did not contain ethidium bromide. The initial colony consists of the culture which had been normalized but not diluted at this point. Moving leftwards, each spot is a 5 µl, 1/10 serial dilution, of the previous spot. As shown, MCΔΔ transformed with the control plasmid pBADHisgltPh was unable to grow on the plate containing 75 µg ml^{-1} ethidium bromide. Also shown is the restored growth of all MCΔΔ strains transformed with the panel of clones.

3.3. Expression Trials

Since the heterologously expressed proteins of interest had now been shown to be functional, expression trials were now required to determine whether the proteins of interest could be readily expressed, under the previously described growth conditions outlined above, with a high enough yield to enable biochemical and biophysical characterisation. Expression trials were conducted by initially inoculating 1 litre cultures to a starting OD_{600} of 0.05. These were then grown until an OD_{600} of 0.6-0.8 is reached, at which point 0.5mM IPTG was added to strains containing a pBAD expression vector and 0.1% L-arabinose to those containing a pET expression vector to induce expression. Multi-drug export protein A (MepA)[58], was used as a positive control for the expression, since this protein has previously been shown to express under the conditions outlined in section 2.1. The cultures were then incubated for a further 2 hours at 37°C before being centrifuged and the membranes harvested. The membranes were then run on SDS-PAGE gels alongside a molecular weight marker. Western blotting was used

to analyse expression of the proteins (fig. 11.). This was done by transferring the proteins onto polyvinylidene difluoride (PVDF) membrane and using tetra-his antibody to detect the his-tagged proteins of interest. However, several issues were encountered when western blotting which resulted in a lack of bands on the final blot. This could have been due to the primary and secondary antibodies or enhanced chemiluminescence (ECL) kit being used to visualise the blots. This meant that not all proteins of interest were shown to express. Several of the proteins of interest, however, were shown to express in both pBAD and pET vectors (fig. 11.) this is shown by the appearance of bands on the blots at the correct place indicating the proteins are the expected size which was approximately 25 kilodaltons.

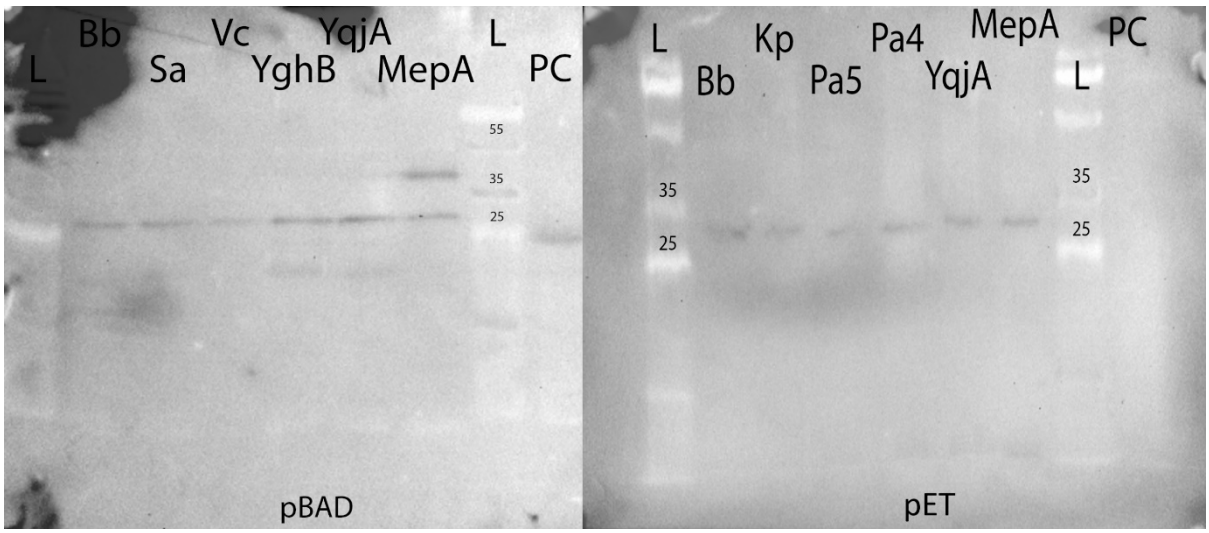


Figure 11 – Western Blots from the expression trials of several pET and pBAD clones. Each was shown here to express, as indicated by the presence of a band at the expected MW of ~25 kilodaltons. L denotes the molecular weight marker used, with the relevant MW bands labelled as 55, 35, and 25 KDa. MepA, shown here, was used to control for expression since it has been shown to express under the conditions used here. PC denotes a control used for the western blots, however the positive control used for pET was different to that used in pBAD. The positive control used for pET didn't appear to give a band, however the reason for this is unknown. Bb, Sa, Vc, YghB, YqjA, Pa4, Pa5, and Kp denote the various DedA homologues being expressed.

3.4. Overexpression and Purification

Although it was not possible to express all proteins of interest without further optimisation, the majority were still shown to express (fig 11.). However, due to time constraints it was not possible to overexpress each protein of interest, this meant having to focus on those which would likely yield the most results. For this reason, the overexpression and purification of YqjA was of primary interest, since it has previously been shown that YqjA can be overexpressed and purified in *E. coli*[26]. The expression trials had already demonstrated that it was possible to

overexpress YqjA when grown at 37°C and induced with 0.5mM IPTG for strains containing a pBAD expression vector and 0.1% L-arabinose for those containing a pET expression vector when the cultures reach an OD_{600} of 0.6-0.8. To yield enough of the protein to enable biochemical and biophysical characterisation, larger scale cultures of up to 5 litres were required. Large scale cultures, between 3-5 litres of Bl21-AI containing *yqjA* in a modified pBAD expression vector with a C-terminal his-tag as well as in a modified pET expression vector with an N-terminal his-tag were grown per the growth conditions outlined in section 2.1. These were induced with 0.5mM IPTG for strains containing a pBAD expression vector and 0.1% L-arabinose for those containing a pET expression vector. Following this the cultures were incubated for 2 hours at 37°C. The cultures were then centrifuged and resuspended in lysis buffer before being lysed by sonication. Once lysed, the cell lysate was centrifuged at 20,000xg to clarify the lysate which was then centrifuged again at 170,000xg to pellet the membranes. The membranes were resuspended in Buffer A and stored at -20°C. YqjA was purified by thawing the membranes, solubilizing in either DDM or SMA, centrifugation of the solubilized sample. This is then followed by incubation with TALON metal affinity resin which binds to the polyhistidine-tag attached to YqjA. The solution containing the resin-bound yqjA is then poured onto a gravity flow column and wash buffer is poured through. Following this, elution buffer is then poured through the column several times and collected, to ensure that all purified protein previously bound to the resin has been eluted (fig. 12.). After purification, several techniques were used to analyse the purified protein such as being run on SDS-PAGE gels, followed by western blots, as well as size-exclusion chromatography. Prior to analysis, the concentration, as indicated by the absorbance at 280 nm, was measured using a nanodrop spectrophotometer. Each elution fraction from the purified protein sample as well as the initial flow through and wash samples were run on an SDS-PAGE gel, alongside a molecular weight marker. It was confirmed that yqjA had been purified, as indicated by the presence of large, defined bands at the expected size on the gels (fig. 13.).

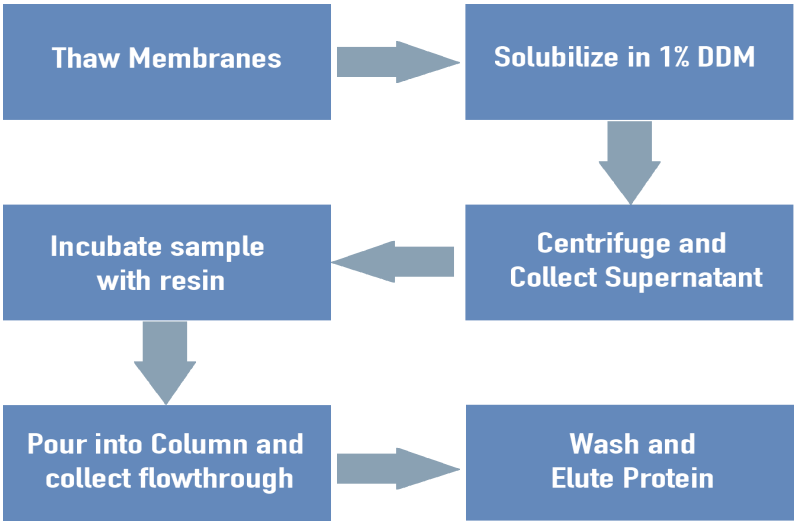


Figure 12. A flow through diagram showing the process of protein purification from the Membrane pellet stage through to the wash and elution of the protein.

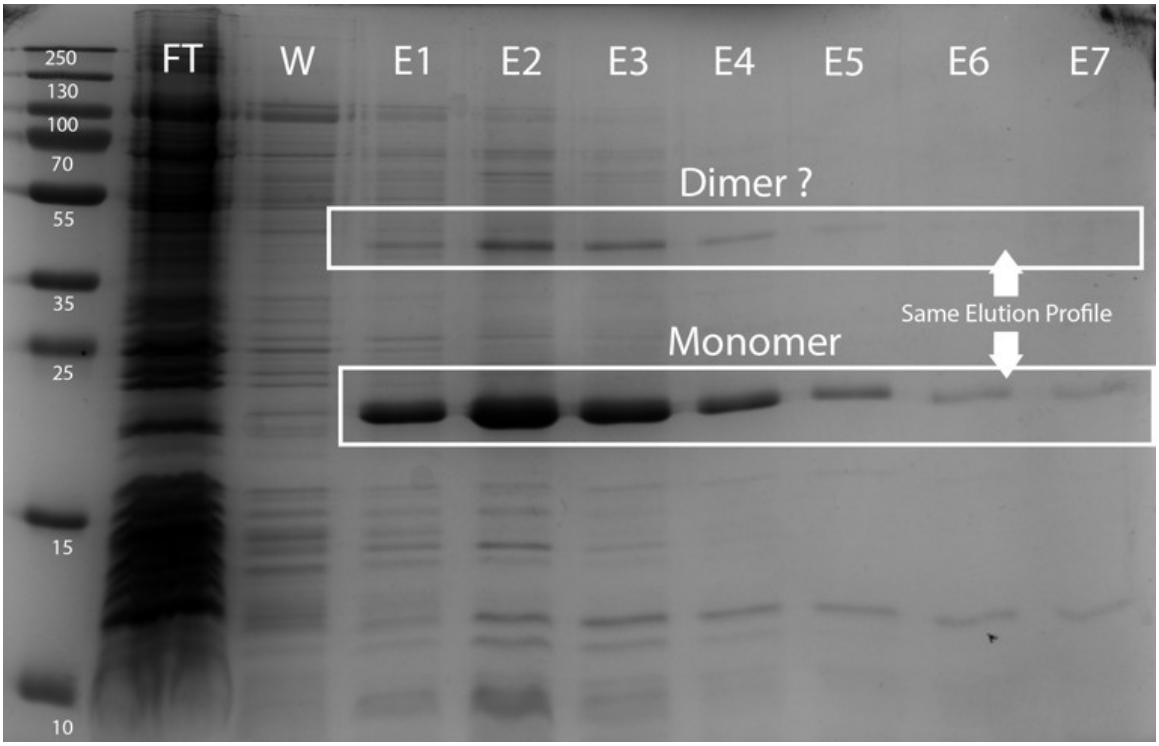


Figure 13. An SDS-PAGE gel of YqjA after the initial purification. The protein was solubilized with DDM. As shown within the white boxes, a large amount of protein was eluted in fractions E1-E7. There appears to be a higher MW band which could potentially be a dimeric form of this protein, however it's not possible to say for sure whether this is the case. FT denotes the initial flow through after resin binding. W denotes the wash before eluting.

3.5. Size-Exclusion Chromatography

Size-exclusion chromatography is an analytical technique used for the fractionation of compounds such as proteins, polysaccharides, and other small molecules in an aqueous solution. It can also be used to remove large protein complexes, as well as smaller contaminants with larger sized molecular species eluting before smaller ones[59]. SEC can be used as a way of separating and quantifying various aggregates and potentially different molecular arrangements of the protein of interest using UV detection[59]. However, the protein sample, when run on the SEC column will be heavily diluted, which could influence the equilibrium between oligomeric states[59]. There are two states in the SEC column, stationary phase, which consists of a porous matrix of beads with a defined pore size range (exclusion limit), and the mobile phase which consists of a buffer which flows through the matrix. Larger complexes which are unable to move in between the pores in the matrix remain in mobile phase and therefore move through the column and elute first. Smaller proteins and other compounds which can move through the pores in the matrix enter the stationary phase and therefore move through the column more slowly, eluting later. Complexes which are larger than the exclusion limit of the column will elute first and produce a peak, this is known as the void volume. As samples move through the column a trace is produced showing several conditions, including the column pressure, as well as the UV absorbance which gives an indication of the protein concentration. As the samples are eluted from the column, they are collected into separate fractions which can then be analysed by running them on an SDS-PAGE gel to confirm that the peak observed in the trace corresponds to the protein of interest.

Several of the elution fractions, E2-E7, with a total volume of ~3.5 ml from the initial, large scale purification of YqjA, were concentrated using a protein concentrator with a molecular weight cut-off of 10 KDa to a final volume of 0.5 ml. The concentrated protein sample was then run on a superdex 200 increase 10/300gl size exclusion column which had been equilibrated with buffer containing the detergent n-decyl-ß-D- maltopyranoside (DM). The SEC trace shown in figure 14 appears to indicate the presence of a dimer, as indicated by the presence of a shoulder on the main peak. This is also indicated by the presence of separate bands at the expected size of a YqjA dimer (fig. 15.). The bands present in the SDS-PAGE gel

from the SEC fractions (fig. 15) also seem to show the presence of an oligomer, however this is only speculative.

The amount of protein observed in the SEC trace, as indicated by the peak in the UV absorbance trace, was substantially lower than expected when compared to the amount of protein present after purification (fig. 13 & 14.). This could potentially have been due to secondary interactions between the protein and the column resulting in protein being lost to the column[60], or a change in detergent between the sample and the detergent used in the buffer. Therefore, to determine whether a change in detergent was the cause of the low absorbance reading, more YqjA was purified. This time YqjA was solubilized using SMA, since, as previously mentioned, SMA is an alternative, detergent-free, way to purify the protein of interest. Alternatively, YqjA was also solubilized using DDM. However this time the DDM solubilized samples were run on the SEC column using a buffer that also contained DDM rather than DM, as was previously used, in order to prevent any potential interactions.

The purified YqjA samples which had been solubilized using SMA were initially run after equilibrating the column with a non-detergent containing buffer. Unexpectedly, the SEC trace appeared to show a highly concentrated peak at the void volume of the column (fig. 16). This could be an indication that the protein formed a complex with the SMA which was larger than the exclusion limit of the column, causing it to elute in the void volume. However, this is also speculative.

The second YqjA sample solubilized in DDM which was run on the column once it had been equilibrated using buffer containing DDM as opposed to DM (fig. 17). The trace observed, however, produced a substantially lower peak in absorbance than the previous two runs. However, this could be due to a lower yield of purified protein since only one elution fraction, E2, was run on the column, as opposed to the previous sample, which consisted of several elution fractions which had been concentrated.

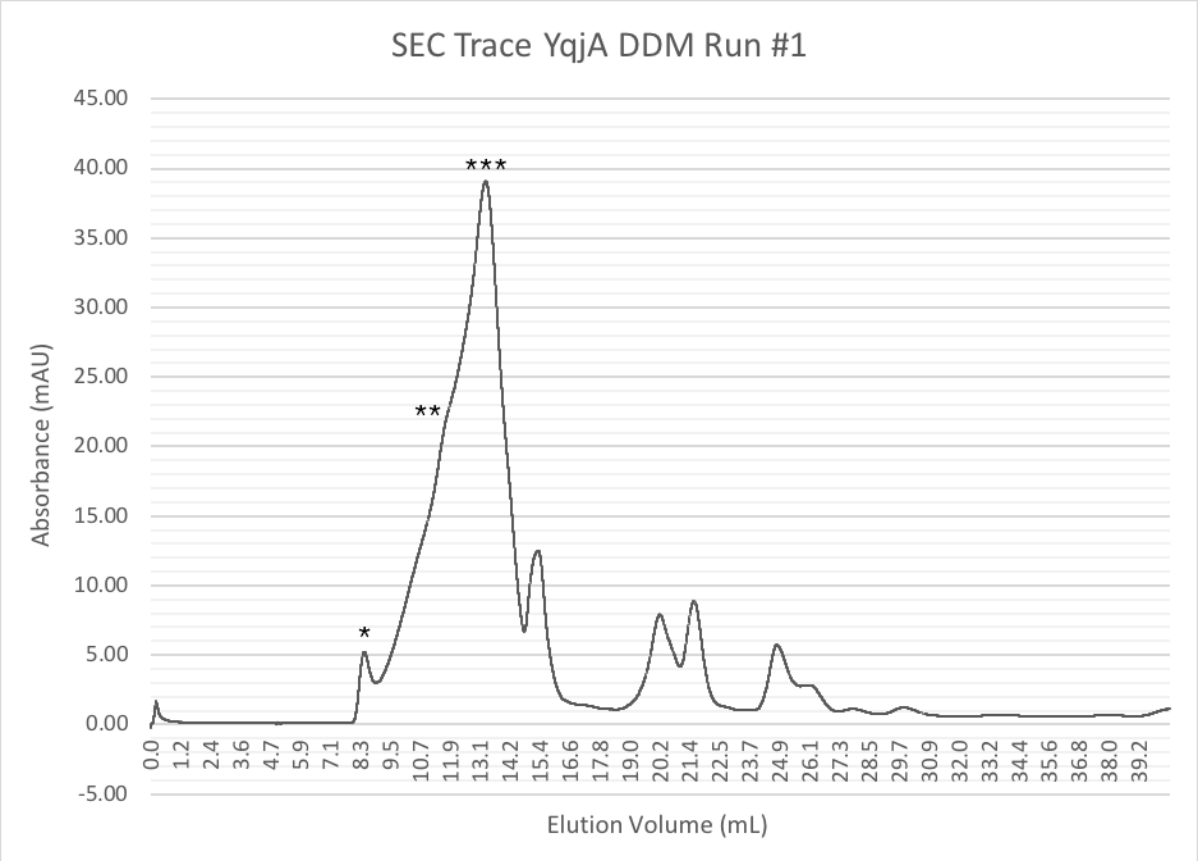


Figure 14. The SEC trace from the initial run of YqjA which had been solubilized using DDM. The trace shows the UV absorbance which gives an indication of the protein concentration. * represents the void volume of the column. ** represents a possible dimer eluting, giving a shoulder on the main peak. *** represents the protein of interest eluting. The peak from the protein of interest shown here was substantially smaller than expected after observing the amount of protein on the SDS-PAGE gel. Throughout this run the column pressure steadily increased. This could be an indication that protein is being lost to the column. This column was calibrated with buffer containing 0.1% DM.

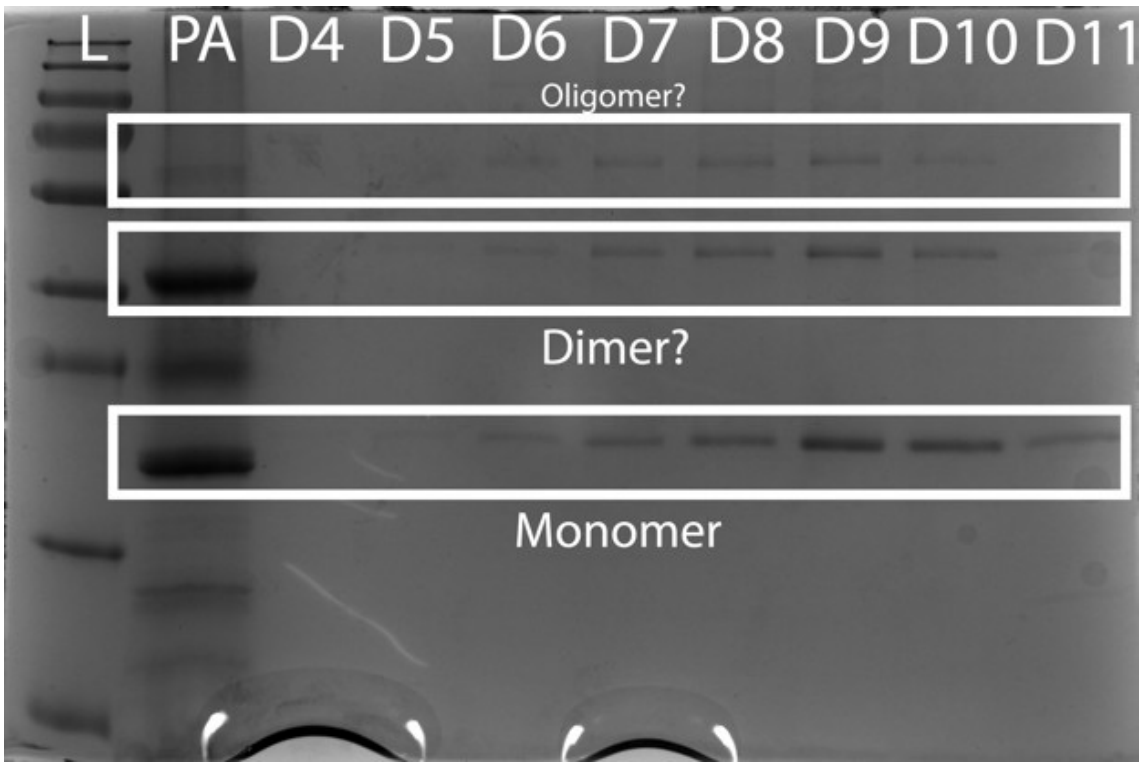


Figure 15. An SDS PAGE Gel showing samples taken from the SEC fractions shown above at the peak. PA represents the pre AKTA sample which consists of elution fractions E2-E7 after concentration. L denotes the molecular weight marker used in this gel to confirm the size of the bands. D4-D11 represents the fractions from the peak in the above SEC trace.

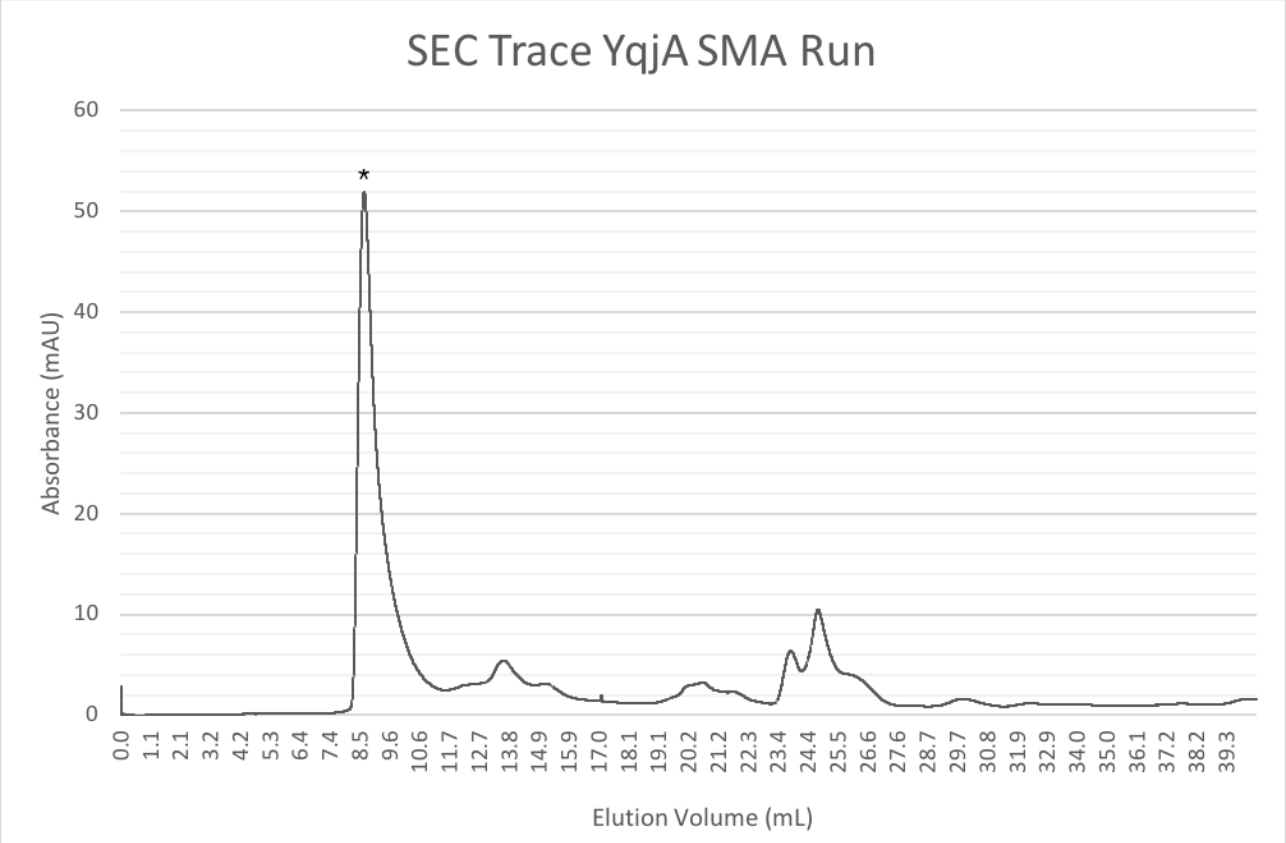


Figure 16. The SEC trace from the second run of YqjA which had been solubilized using SMA. The trace shows the UV absorbance which gives an indication of the protein concentration. * represents the peak in absorbance caused by the protein of interest eluting. The peak shown here was again, substantially smaller than expected. The void volume of this column is 8.2 mL, meaning the peak given in the above trace could be an indication that the protein is eluting in the void volume.

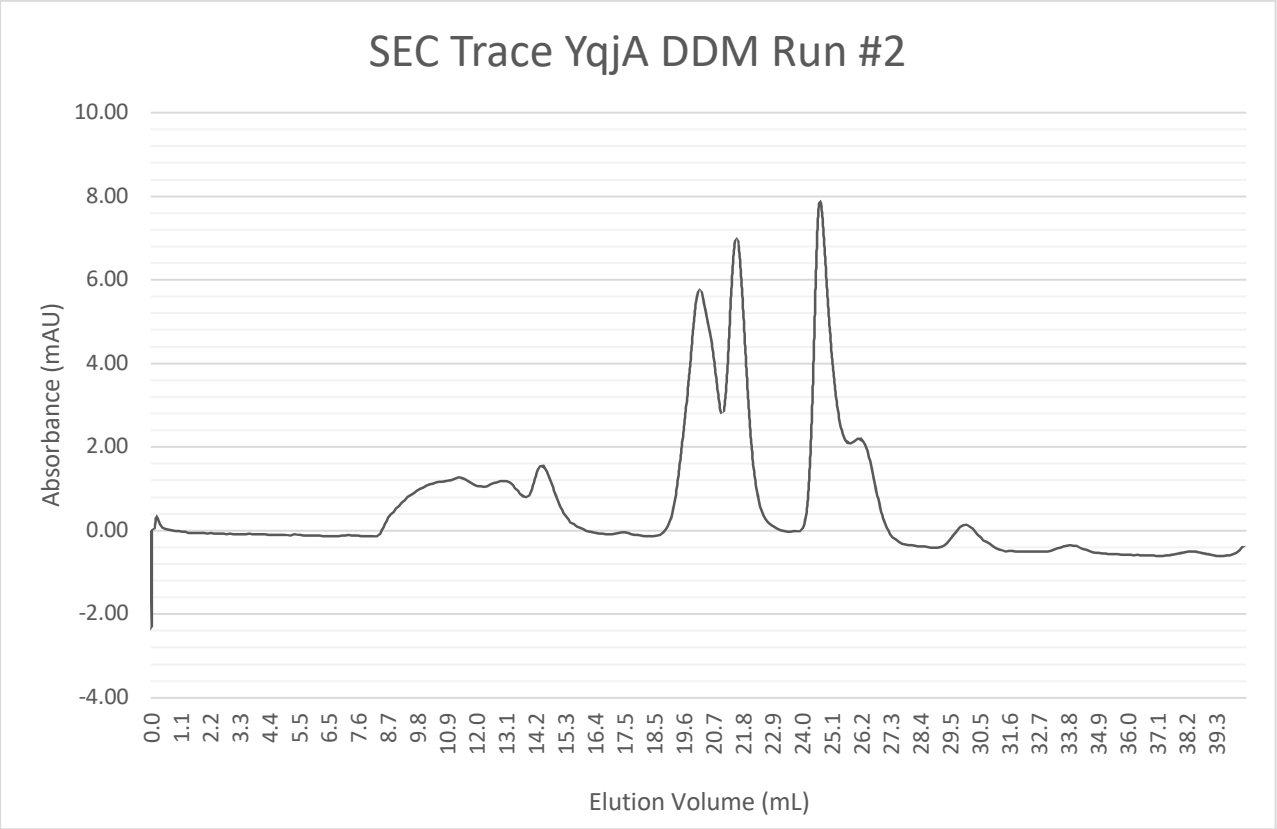


Figure 17. The SEC trace from the final run of YqjA which had been solubilized using DDM. This column was equilibrated with buffer containing 0.1% DDM. The trace shows the UV absorbance which gives an indication of the protein concentration. Virtually no peak was observed in this trace, possibly due to errors in the purification process leading to a loss of protein. This could be an indication that protein is being lost to the column.

4. Discussion

DedA are a ubiquitous family of integral membrane proteins that play a key role in bacterial viability[11]. Knocking out members of the DedA family such as the removal of YqjA and YghB of *E. coli* has shown what a key role these proteins play within their hosts[11]. For example, the removal of DedA from *Klebsiella pneumoniae* was shown to be essential for colistin resistance[13]. The structure, function, and physiological role of this protein family currently remains elusive, although several *in vivo* studies have begun to uncover features and potential functions. Hydropathy profile alignments have revealed that the leuT fold may have evolved from an ancient member of the DedA family[26]. This has led to a hypothesis that YqjA and YghB may in fact operate as hetero- or homodimers in E. coli although this is highly speculative[26]. Due to their ubiquitous nature, it has been proposed that members of this family may have diverged to develop different functions from one another[11]. For example, it has been shown that the expression of 4 *E. coli* DedA homologues, YqjA, YghB, YabI, and YohD, are capable of restoring growth at 42°C, while the remaining 4 homologues, EcDedA, YdjX, YdjZ, and YqaA, are unable to do so[11]. These two groups have also shown to be phylogenetically different from one another[11].

The current school of thought is that at least some members of the DedA family might be responsible for maintaining the PMF[16]. This can be explained by some of the phenotypes observed in BC202 which exhibits temperature sensitivity which can be rescued by growth at a lower pH (pH 6.0) as well as cell division defects which are the result of inefficient export of periplasmic amidases AmiA and AmiC by the tat pathway. Ineffective tat export is likely also due to a disturbance in the PMF. A notable difference in phospholipid composition is also observed in this mutant, with a higher number of acidic phospholipids phosphatidylglycerol (PG) and cardiolipin (CL) as well as a decrease from 60% to <50% of phosphatidylethanolamine from the wild type strain. Mdfa, a Na^+ K^+/H^+ antiporter from the major facilitator superfamily (MFS)[24], is also capable of restoring normal growth to the cell, which could indicate that members of the DedA family may function as transporters[2627]. These factors seem to indicate that members of the DedA family may play a role in maintaining the PMF, potentially as membrane transporters.

The aim of this study was to clone a panel of DedA genes from various organisms, assess their function by developing an *in vivo* functional assay, and perform expression trials on the clones which are shown to produce functional proteins. Following this the aim was to overexpress and purify the functional proteins and finally characterise them by developing an *in vitro* functional assay.

4.2. Establishing a Library of Clones

Several DedA homologues from various organisms such as *E. coli*, *K. pneumoniae*, *P. aeruginosa*, and *B. burgdorferi* were used in this research. These homologues were selected for their ability to restore growth at an elevated temperature in the mutant strain BC202. Restriction enzyme-based cloning was used for the majority of this research. This involved replicating the gene of interest using PCR. Following this, restriction endonucleases AscI and XhoI were used to cleave the gene of interest and a modified pET expression vector with an N-terminal polyhistidine-tag. The gene of interest was then ligated into the vector using T4 DNA ligase. This method of cloning can be simple and straightforward. However, several issues arose including unsuccessful PCR attempts, inefficient ligation, which were likely due to either an inappropriate vector to insert molar ratio, or a lack of vector and insert DNA which could have been loss during gel extraction and PCR purification. This can be seen in the decrease in concentrations of each sample following PCR, PCR purification and gel extraction. Optimisation of PCR conditions such as by changing the annealing temperature or elongation time corrected any issues with this process.

An alternative to restriction enzyme-based cloning is a newly developed method called Gibson Assembly. This method was used for cloning *yghB* into pBAD due to the presence of a kpnI cut site within *yghB* preventing the use of conventional cloning with this restriction endonuclease. In order to circumvent this issue, a technique referred to as Gibson Assembly cloning was used. This method involved replicating both *yghB* from genomic DNA isolated from strain W3110, and a linearized recipient pBADHis plasmid using primers designed with an overlapping region (fig. 3.). The vector and insert are then mixed together with Gibson master mix, containing Isothermal (ISO) reaction buffer[48], T5 exonuclease[48], Phusion polymerase[48], Taq ligase[48], and milliQ water. This was done for 60 minutes at 50°C, during

which time T5 exonuclease chews-back the overlapping regions, *yghB* anneals to the pBADHis vector, and Phusion polymerase and *Taq* ligase seal the nicks. The newly formed vector is then transformed into the desired strain.

Some of the cloning issues highlighted above such as a loss of DNA due to PCR purification, and gel extraction, could be circumvented with the use of Gibson assembly for all clones rather than solely for cloning *yghB* into pBAD. This is because the Gibson assembly system is simple and efficient making this system beneficial for a high throughput approach to cloning[48]. The use of such method could have potentially accelerated the cloning process, providing more time for downstream analysis. Another alternative to restriction enzyme-based cloning, called In-fusion cloning, which involves inserting the gene of interest into an expression vector, such as a GFP-fusion vector. The use of a GFP-fusion vector to make membrane protein-GFP fusions can accelerate the process of optimizing expression and purification of membrane proteins, since it allows visualization of the membrane proteins of interest at any stage during this process[61]. In-fusion cloning is similar to Gibson assembly since both methods involve ligating at least two, linearized, DNA fragments. However, with In-fusion cloning, the vector is linearized using a single restriction endonuclease at the site intended for the gene of interest. In-fusion cloning also requires primers for the gene of interest to have an overlapping region, and utilises exonuclease to remove 5' nucleotides before annealing the gene of interest with the GFP-fusion vector, pET28a (+)-derived GFP-8His fusion vector[61,62].

4.3. Temperature sensitivity assay

Once a library of clones had been established, to ensure that they were capable of expressing functional protein, an *in vivo* functional assay to assess the whether the heterologously expressed protein of interest was functional was developed. This involved transforming each oftheexpression vector clones into strain MCΔΔ which is then be grown as an overnight culture, alongside MC4100 transformed with pBADHis*gltPh* and MCΔΔ transformed with a modified pBAD plasmid, pBADHis*gltph*, containing the gene encoding Gltph, a glutamate transporter[63]. This was used to control for any non-specific effect that the expression of membrane protein could cause. The development of this assay was particularly useful as it was relatively high throughput and simple. However, several issues were encountered a couple of issues with this assay. The first of which was the drying of plates due to the high incubation temperature. To

counteract this issue, agarose concentrations of 1-5% were tested to determine whether this could correct the issue, however, this decreased visibility making results harder to interpret.

The second issue encountered was the continued growth of MCΔΔ pBADHis*gltPh* even though this MCΔΔ had previously been shown to halt growth[26] and it was unlikely that GltPh was having any effect on this. A potential explanation for this issue is thattheMCΔΔ strain is mislabelled and is not a double-deletion mutant of *yqjA* and *yghB*. To test this, two 500 base-pair primers which would bind upstream and downstream of *yqjA* and *yghB* in the MC4100 genome (Fig. 7.) were designed. Overnight cultures of MC4100 and MCΔΔ were grown and gDNA extraction was used to purify the genomic DNA from each strain. PCR was then conducted, using the primers we designed and gDNA from MC4100 and MCΔΔ to amplify the selected regions. The presence of *yqjA* and *yghB* would generate a 1600 base-pair fragment, and the deletion would produce a 1000 base-pair fragment. The PCR products were run on an agarose gel alongside a molecular weight marker to determine the size of each band. This confirmed that the MCΔΔ strain being used was the double-deletion mutant for both *yqjA* and *yghB* (fig. 8.). The cause of the continued growth of MCΔΔ above 44°C remains unknown.

4.4. Ethidium bromide sensitivity assays

An alternative approach to this assay was to utilize an alternative way to halt growth in MCΔΔ which could be restored by the expression oftheproteins of interest. The MIC of ethidium bromide BC202 (W3110 *ΔyqjA::ΔyghB*), is 16-fold lower than that of the parent strain W3110[16] and overexpression of YqjA in BC202 can restore the MIC to normal levels[16]. Therefore, an alternative method to analyse function intheproteins of interest utilising the significant difference in ethidium bromide MIC was developed. This involved growing overnight cultures of MC4100, MCΔΔ transformed with a modified pBAD plasmid, pBADHis*gltph*. These cultures were then serially diluted and spotted onto selective plates which also contained ethidium bromide at a concentration above the MIC of MCΔΔ but below the MIC of MC4100. This way, if the proteins of interest are expressed and functional, growth will be restored to MCΔΔ. This approach proved much more effective than the temperature assays since growth of MCΔΔ was inhibited. Several concentrations of ethidium bromide ranging from 25-75 μg ml^{-1} were trialled (fig. 9.). The final concentration of ethidium bromide used in the assay was 75 μg

ml^{-1} since this inhibited growth of all dilutions of MCΔΔ excluding the first spot. Partial lack of growth was ideal since without any growth it would be impossible to determine whether complete lack of growth of MCΔΔ is due to other factors.

4.4.2. All Genes Rescue Restore Growth on Ethidium Bromide

The MIC of ethidium bromide in BC202 is 25 μg ml^{-1} compared to the MIC of the parent strain, W3110, of 400 μg ml^{-1}, which is 16-fold higher. Cultures of each MCΔΔ strain transformed withtheclones, MC4100 transformed with pBADHis*gltPh*, and MCΔΔ transformed with pBADHis*gltPh* were grown overnight. These cultures were then serially diluted and spotted onto selective plates that also contained 75 μg ml^{-1} ethidium bromide. These were then incubated for 22-24 hours at 30°C. All transformants grew (fig 10.), indicating that each gene of interest in this study is expressed and functional in the membrane and capable of correcting the ethidium bromide sensitive phenotype.

4.4.3. Is There a Link Between Temperature Sensitivity and Ethidium Bromide Sensitivity?

The homologue of DedA present in *Klebsiella pneumoniae*, KpDedA, was also able to restore growth in this rescue assay (fig. 10.). This is particularly interesting since KpDedA has a 92% amino acid sequence identity to EcDedA[13], a DedA homologue which is unable to restore growth to BC202 at 42°C. EcDedA, as previously mentioned, is one of eight homologues expressed in *E. coli*. As with all DedA family proteins, not much is known about EcDedA. However, one study showed in a phylogenetic analysis of the DedA homologues in *E. coli*, that it may, alongside three other *E. coli* DedA family members, YdjX, YdjZ, and YqaA, have evolved different functions from the other four *E. coli* DedA homologues, YqjA, YghB, YabI, and YohD, which can correct the temperature sensitive phenotype in BC202. This finding could suggest that the temperature sensitivity phenotype and the increased sensitivity to xenobiotics such as ethidium bromide observed in the double-deletion mutant may be the result of separate mechanisms and therefore unrelated to one another. Another possibility is that the KpDedA possesses an essential acidic residue in the non-conserved region with EcDedA which is critical for the ability of this protein to be able to rescue the ethidium bromide sensitive phenotype. One way to determine whether the temperature and ethidium bromide sensitive

phenotypes are linked is by assessing whether KpDedA can restore growth to MCΔΔ when grown at 44°C. However, the continued growth of MCΔΔ on solid media hindered this. A potential alternative to this method which might be able to overcome this issue would be to use liquid cultures and assess growth by taking OD_{600} readings regularly.

4.4. Optimization of Protein Expression and Purification

Recombinant expression and purification of membrane proteins can be a challenging process with many potential bottlenecks which are result of factors such as secondary structure stability, target protein toxicity, codon usage, and proteolytic degradation[64,65,53]. Optimisation of protein expression could have been implemented to maximise heterologously expressed protein yield. This could potentially have been achieved by altering the growth conditions such as incubation temperature, media components, and inducer concentration[57].

Membrane proteins, when expressed, are embedded within the lipid bilayer, the composition of which is highly variable, with many different phospholipids available. The type, and characteristics of the lipids within the membrane proteins native environment has a large impact on the function of the protein, as well as its structure and oligomeric state[66]. Solubilisation of membrane proteins can be carried out using various detergents and other compounds[67]. The choice of detergent used for solubilisation is critical in ensuring that the highest yield of homogenous, stable protein is obtained[67]. Often, various detergents will be screened to find the optimal type and concentration to use[67]. The detergent, n-Dodecyl β-D-Maltoside (DDM) is commonly used to solubilize membrane proteins due it's low cost and the fact that it produces stable solubilized proteins[42]. A downside to the use of detergents for membrane protein solubilisation is that the protein is removed from its native environment which can cause a change in the structure and oligomeric state which is undesirable for use in structural studies[67]. However, an alternative method using a styrene maleic acid copolymer lipid particles (SMALPs), can also be used for the solubilisation of membrane proteins without the use of detergent and whilst maintaining the protein in their native state[43]. SMALPs are a relatively new development in the field of membrane protein purification. Due to the hydrophobicity of styrene and hydrophilic nature of maleic acid, the SMA copolymers can integrate into cell membranes[43] where they extract the integral membrane protein in the form of a small disc. This disc contains the protein of interest, surrounded by native lipids[43]. The

SMA nanodiscs are water-soluble and when used with a suitable affinity tagged protein, can be used with standard chromatographic methods[43]. Although SMALPs were used for solubilisation oftheprotein of interest briefly, a large amount of optimization would be required to get this technique to work successfully withtheprotein. However, the use of this method over detergents could be beneficial, particularly for use in structural studies since the proteins are more likely to maintain their native structure and oligomeric state than when solubilized using detergents[43].

4.5. Size Exclusion Chromatography (SEC) Analysis

Size exclusion chromatography is an analytical technique used for the fractionation compounds such as proteins, polysaccharides, and other small molecules in an aqueous solution[59]. SEC can be used as a way of separating and quantifying various aggregates of the protein of interest by measuring the UV absorbance[59]. When samples are run on a size exclusion column they can enter two different states, stationary phase refers to a porous matrix of beads with pre-determined pore sizes, this is referred to as the exclusion limit, and mobile phase refers to the buffer which runs through the column. Protein samples will run through the column, with larger aggregates, which are unable to enter the porous matrix remaining in mobile phase and eluting first, followed by smaller aggregates, which will move through the matrix and enter mobile phase. As the samples elute from the column they will produce a trace which provides information about the concentration of protein.

Once yqjA had been overexpressed and purified, the samples were run on the size exclusion column. The initial purification, in which YqjA was solubilized using DDM, produced a significant amount of protein (fig. 12.). Elution fractions E2-E7 were concentrated down to a 0.5ml sample which was then run on the SEC column (fig. 14). As the samples eluted from the column they were collected into fractions which were then run on an SDS-PAGE gel alongside a molecular weight marker of pre-determined size to confirm the presence of YqjA at the location of the peaks observed in the SEC trace (fig. 14 & 15). The initial run produced a significantly smaller absorbance peak than expected, since the size of this peak corresponds to the concentration of protein in the sample. This was unexpected due to the amount of protein purified (fig 15.). However, there was a substantial increase in the pre-column pressure throughout, which could be an indication that some protein was being lost to the column. This

could potentially have been due to secondary interactions between the protein and the column. To confirm if this was the case, a larger amount of YqjA was purified, solubilizing with both DDM and SMA, individually.

The SMA solubilized sample was initially run on the column following equilibration with a buffer which didn't contain detergent (fig. 16). The trace produced appears to show a highly concentrated peak at the void volume of the column. This could be an indication that YqjA and the SMA formed a complex which was larger than the exclusion limit of the column, causing the complex to elute in the void volume, however this is completely speculative. Perhaps further optimization of the SMA solubilization process would have yielded higher quality results.

The second sample, solubilized with DDM, was then run on the column having equilibrated the column with a buffer which also contained DDM. The absorbance peak for this sample was significantly lower than expected (fig. 17.) when compared with the previous DDM solubilized sample. The cause of the low absorbance readings for this protein sample is unknown. However, it's possible that, as previously mentioned, some protein may be forming secondary interactions with the column. Optimization of this process, by changing the concentration and type of detergent used could perhaps have improved results.

4.6. Future Work

The function of DedA family members currently remains elusive, which is in-part due to a lack of structural information[26]. In-vitro studies can be particularly useful for functional studies of membrane proteins, because, although the study of membrane proteins in their native environment is preferred, *in vivo* studies can make functional data hard to interpret. This is due to interference with other membrane proteins and cellular activities[44,67]. The overall aim of this research was to express and purify several DedA family members for structural and functional characterisation. This was achieved by first creating a panel of clones consisting of several DedA family members which had been shown to restore growth in the double-deletion mutant BC202 at 42°C. These clones were then transformed into various strains including MCΔΔ, TOP10, and BL21-AI. An *in vitro* functional assay was developed in order to determine

whether these proteins could express and function in the membrane. Once this had been established, they were then transformed into BL21-AI and TOP10 for expression and purification. Only one particular DedA family member, an *E. coli* homologue called YqjA would readily overexpress and purify. This purified protein sample was then analysed using size exclusion chromatography as a way of separating and quantifying the protein of interest using UV detection[59] (fig. 14-17).

Further work must be done to further investigate the structure and function of this protein family. The aim of this work should be to produce a high enough yield of these proteins for further biochemical and biophysical characterisation. One such method would involve developing an *in vitro* functional assay which could be used to assess internal pH, using a pH sensitive fluorophore such as pyranine, as an indicator of proton movement[45]. This would be particularly useful since it is best to conduct functional studies of proteins in their native environment[45]. However, the native environment of the cell can make it hard to control factors such as proton movement and can make interpreting data difficult due to activity of other cellular proteins[45]. This method could provide significant information about the potential function of this family of proteins. Current hypotheses suggest that members of the DedA family may be responsible for maintenance of the PMF[24,26] and could function as membrane transporters[26]. This could mean that a functional assay such as this might elucidate the role of this protein family in maintenance of the PMF and potential substrates if they do function as membrane transporters.

Further study would also involve the structural analysis of this panel of DedA family proteins using a variety of techniques. One such technique, called X-ray crystallography, is commonly used to elucidate the structure of proteins and other macromolecules. X-ray crystallography involves several stages. The protein of interest is initially purified, following which sufficient quality protein crystals must be grown. This process involves exposing the purified protein sample to a number of different conditions in order to coax it out of solution and crystallise. However, there are many issues which can occur at this stage, resulting in various, non-desirable outcomes. These include the formation of poor quality crystals, large crystals which fail to diffract, or an amorphous precipitate, as well as no crystal formation[78]. Some of the conditions which have a bearing on whether or not protein crystallization occurs include the buffers used, the pH, the temperature, and protein concentration, as well as many other variables[78, 79]. Once the protein has formed a crystal, it is placed and adjusted to the correct angle on a goniometer, a device which

allows an object to be rotated to a specific angle[80]. Following this, a targeted beam, consisting of one particular X-ray wavelength, which is made up from accelerating electrons, is focused on the protein crystal as it rotates[80]. As a result, the beam is diffracted at various angles, forming a diffraction pattern as it strikes the sensor or film[80]. From this an electron density map is produced. The electron density map can give an indication of the position of the atoms within the protein crystal, as well as the number and type of bonds formed, as well as various other pieces of information which allow the structure of the protein to be determined[80].

Future work would also involve further structural analysis to determine whether or not YqjA and YghB are capable of interchangeably forming hetero- and homodimers. This is of particular interest since YqjA has been shown to form homodimers in-vivo in the absence of YghB[25, 26]. Several strategies can be employed in order to achieve this. One such technique would involve co-expressing YqjA and YghB with different affinity tags in a single strain of *E. coli*. Following this, formaldehyde is added to the media. The formaldehyde then permeates the cell membranes and generates cross-links between any proteins which are interacting. If YqjA and YghB form heterodimers, then the formaldehyde will covalently cross-link the two proteins[81]. The two cross-linked proteins can then be purified separately using a suitable form of affinity purification for the affinity tag present on each of the proteins, such as immobilized metal ion affinity chromatography. Once the proteins have been purified, they can then be analysed using several techniques, such as by running them on an SDS PAGE gel and observing the size of band formed to see if it matches the expected size. Following which the protein samples can be concentrated and run on a SEC column, and the peaks analysed. Alternatively, the purified proteins can be separated using one-dimensional gel electrophoresis, following which the proteins can be analysed using tandem mass spectrometry[81].

5. Conclusion

The aim of this research was to clone a panel of DedA homologues from various organisms, including *E. coli, S. aureus, V. cholerae, P. aeruginosa,* and *B. burgdorferi.* Once a panel of clones had been established, an *in vivo* functional assay which could assess whether the proteins of interest would express and function correctly was developed. Expression trials were then conducted in order to determine which proteins would readily express under the growth conditions outlined in section 2.7. Several of the proteins of interest readily expressed (fig. 11), however YqjA in particular was selected for purification and further downstream analysis via size exclusion chromatography.

YqjA was initially solubilized with DDM and purified. The purified samples were run on a gel to confirm whether the protein was present (fig. 13.). Following this, several elution fractions, E2-E7, were concentrated to a final volume of 0.5ml and the sample was analysed using size exclusion chromatography. However, the peak observed on the SEC trace was significantly lower than expected when compared with the amount of protein present (fig 13 & 14). Throughout the run there was an increase in pre-column pressure which could indicate that this low protein concentration could be due to a loss of protein to the column. The peak observed appeared to indicate the presence of both a monomer and dimer (fig. 14). The low yield could also have been due to interactions between the DDM used to solubilize YqjA and the DM present in the buffer used to equilibrate the SEC column.

In order to determine the cause of the decreased protein concentration, a larger amount of YqjA was purified which was solubilized using both SMA and DDM, individually. These samples were then run on the size exclusion column.

The SEC trace observed when the SMA sample was run appeared to show a highly concentrated peak at the void volume (fig. 16), potentially indicating that the SMA may have formed a large complex with YqjA above the exclusion limit of the column, resulting in the complex eluting in the void volume, although this is completely speculative. The second SEC trace, in which the sample was solubilized using DDM was run on the column equilibrated using buffer containing DDM appeared to give an insignificant peak (fig. 17). The absorbance peak for this sample was significantly lower than expected when compared with the previous

DDM solubilized sample. The cause of the low absorbance readings for this protein sample is unknown. However, it's possible that, as previously mentioned, some protein may be forming secondary interactions with the column. Optimization of this process, by changing the concentration and type of detergent used could perhaps have improved results.

Finally, further biochemical and biophysical characterisation will be required to uncover the structure and function of this elusive protein family. The development of an *in vitro* functional assay to further study the role of this protein in its native environment could potentially uncover the role of this protein family. Other potential future work includes the use of techniques such as x-ray crystallography to determine the structure of these proteins, as well as formaldehyde cross-linking studies followed by tandem mass spectrometry to further investigate the formation of YqjA and YghB hetero- and homodimers.

6. Acknowledgments

I would like to thank Dr. Christopher Mulligan for providing extensive support, supervision, and help throughout this research, as well as C. Sampson and M.Stewart for their support within the laboratory. I would also like to thank K.M. Frain, A. Jones, A.R. Peswani, D.M.J. Mermans, I. Guerrero-Montero, S.L. Bischoff, and E.H. Nemoto-Smith for their help throughout this research. Thanks to the New York Consortium on Membrane Protein Structure for providing the modified pET plasmid used in this research[47].

7. References

1. Cantas, L. *et al.* A brief multi-disciplinary review on antimicrobial resistance in medicine and its linkage to the global environmental microbiota. *Frontiers in Microbiology* **4,** 96 (2013).
2. Roca, I. *et al.* The global threat of antimicrobial resistance: science for intervention. *New Microbes New Infect.* **6,** 22–29
3. Tenover, F. C. Mechanisms of antimicrobial resistance in bacteria. *Am. J. Infect. Control* **34,** 3–10 (2006).
4. Frere, J. Beta-lactamases and bacterial resistance to antibiotics. *Mol. Microbiol.* **16,** 385–395 (1995).
5. Poole, K. Overcoming antimicrobial resistance by targeting resistance mechanisms. *J. Pharm. Pharmacol.* **53,** 283–294 (2001).
6. Leski, T. A. & Tomasz, A. Role of Penicillin-Binding Protein 2 (PBP2) in the Antibiotic Susceptibility and Cell Wall Cross-Linking of *Staphylococcus aureus* : Evidence for the Cooperative Functioning of PBP2 ,. *J. Bacteriol.* **187,** 1815–1824 (2005).
7. Schweizer, H. P. Efflux as a mechanism of resistance to antimicrobials in *Pseudomonas aeruginosa* and related bacteria: unanswered questions. *Genet. Mol. Res.* **2,** 48–62 (2003).
8. Srikumar, K. P. and R. Multidrug Efflux in *Pseudomonas aeruginosa* Components, Mechanisms and Clinical Significance. *Current Topics in Medicinal Chemistry* **1,** 59–71 (2001).
9. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16,** 161–168 (2016).
10. Li, J. *et al.* Colistin: the re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections. *Lancet Infect. Dis.* **6,** 589–601 (2006).
11. Boughner, L. A. & Doerrler, W. T. Multiple deletions reveal the essentiality of the DedA membrane protein family in *Escherichia coli*. *Microbiology* **158,** 1162–1171 (2012).
12. Doerrler, W. T., Sikdar, R., Kumar, S. & Boughner, L. A. New functions for the ancient DedA membrane protein family. *J. Bacteriol.* **195,** 3–11 (2013).
13. Jana, B. *et al.* The secondary resistome of multidrug-resistant *Klebsiella pneumoniae*.

Sci. Rep. **7,** 1–10 (2017).

14. Falagas, M. E. & Kasiakou, S. K. Colistin: The Revival of Polymyxins for the Management of Multidrug-Resistant Gram-Negative Bacterial Infections. *Rev. Anti-Infective Agents* **40,** 1333–1342 (2018).
15. Liang, F. T. *et al.* BB0250 of *Borrelia burgdorferi* is a Conserved and Essential Inner Membrane Protein Required for Cell Division. *J. Bacteriol.* **192,** 6105–6115 (2010).
16. Kumar, S. & Doerrler, W. T. Members of the conserved DedA family are likely membrane transporters and are required for drug resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **58,** 923–930 (2014).
17. Sikdar, R. & Doerrler, W. T. Inefficient Tat-dependent export of periplasmic amidases in an *Escherichia coli* strain with mutations in two DedA family genes. *J. Bacteriol.* **192,** 807–818 (2010).
18. Ressl, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V. & Ziegler, C. Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature* **458,** 47-52 (2009).
19. Shaffer, P. L., Goehring, A., Shankaranarayanan, A. & Gouaux, E. Structure and mechanism of a Na^+ independent amino acid transporter. *Science* **325,** 1010– 1014 (2010).
20. Fang, Y. *et al.* Structure of a Prokaryotic Virtual Proton Pump at 3.2 Å Resolution. *Nature* **460,** 1040–1043 (2010).
21. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* **437,** 215-223 (2005).
22. Kazmier, K., Claxton, D. P. & Mchaourab, H. S. Alternating access mechanisms of LeuT-fold transporters: trailblazing towards the promised energy landscapes. *Curr. Opin. Struct. Biol.* **45,** 100–108 (2017).
23. Forrest, L. R. *et al.* Mechanism for alternating access in neurotransmitter transporters. *Proc. Natl. Acad. Sci.* **105,** 10338–10343 (2008).
24. Sikdar, R., Simmons, A. R. & Doerrler, W. T. Multiple envelope stress response pathways are activated in an *Escherichia coli* strain with mutations in two members of the Deda membrane protein family. *J. Bacteriol.* **195,** 12–24 (2013).

25. Khafizov, K., Staritzbichler, R., Stamm, M. & Forrest, L. R. A study of the evolution of inverted-topology repeats from LeuT-fold transporters using alignMe. *Biochemistry* **49,** 10702–10713 (2010).

26. Keller, R., Schleppi, N., Weikum, J. & Schneider, D. Mutational analyses of YqjA, a Tvp38/DedA protein of *E. coli*. *FEBS Lett.* **589,** 842–848 (2015).
27. Kumar, S., Bradley, C. L., Mukashyaka, P. & Doerrler, W. T. Identification of essential arginine residues of *Escherichia coli* DedA/Tvp38 family membrane proteins YqjA and YghB. *FEMS Microbiol. Lett.* **363,** 133 (2016).
28. Lolkema, J. S. & Slotboom, D. Hydropathy profile alignment : a tool to search for structural homologues of membrane proteins. *FEMS Microbiol. Rev.* **22,** 305 - 322 (1998).
29. Keller, R., Ziegler, C. & Schneider, D. When two turn into one: Evolution of membrane transporters from half modules. *Biol. Chem.* **395,** 1379–1388 (2014).
30. Thompkins, K., Chattopadhyay, B., Xiao, Y., Henk, M. C. & Doerrler, W. T. Temperature sensitivity and cell division defects in an *Escherichia coli* strain with mutations in *yghB* and *yqjA*, encoding related and conserved inner membrane proteins. *J. Bacteriol.* **190,** 4489–4500 (2008).
31. Luminy, A. De & Cedex, M. NIH Public Access. **5,** 135–146 (2011).
32. Yamada, H. & Mizushima, S. Proton Motive Force-dependent and -independent Protein Translocation Revealed by an Efficient *in Vitro* Assay System of *Escherichia coli*. *J. Biol. Chem.* **264,** 1723–1728 (1989).
33. Bolhuis, H., Molenaar, D., Poelarends, G. & Veen, H. W. Van. Proton Motive Force-Driven and ATP-Dependent Drug Extrusion Systems in Multidrug-Resistant *Lactococcus lactis*. *J. Bacteriol.* **176,** 6957–6964 (1994).
34. Fluman, N. & Bibi, E. Bacterial multidrug transport through the lens of the major facilitator superfamily. *Biochim. Biophys. Acta - Proteins Proteomics* **1794,** 738–747 (2009).
35. Shi, Y., Cromie, M. J., Hsu, F., Turk, J. & Groisman, E. A. PhoP-regulated *Salmonella* resistance to the antimicrobial peptides magainin 2 and polymyxin B. *Mol. Microbiol.* **53,** 229–241 (2004).
36. Rawlings, A. E. Membrane proteins : always an insoluble problem? **44,** 790–795 (2016).

37. Overington, J. P., Al-lazikani, B. & Hopkins, A. L. How many drug targets are there? **5,** 993–997 (2006).
38. Cavanagh, J. *et al.* Re-sensitizing Multidrug Resistant Bacteria to Antibiotics by Targeting Bacterial Response Regulators : Characterization and Comparison of Interactions between 2-Aminoimidazoles and the Response Regulators BfmR from *Acinetobacter baumannii* and QseB from F. *Front. Mol. Biosci.* **5,** 1–12 (2018).
39. Chiang, C.-Y. *et al.* Mitigating the Impact of Antibacterial Drug Resistance and Challenges. *MBio* **9,** 1–14 (2018).
40. Wambaugh, M. A., Shakya, V. P. S., Lewis, A. J., Mulvey, M. A. & Brown, J. C. S. High-throughput identification and rational design of synergistic small-molecule pairs for combating and bypassing antibiotic resistance. *PLoS Biol.* **15,** 1–33 (2017).
41. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3,** 301 - 317 (2004).
42. Carpenter, E. P., Beis, K., Cameron, A. D. & Iwata, S. Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.* **18,** 581–586 (2008).
43. Postis, V. *et al.* The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *BBA - Biomembr.* **1848,** 496–501 (2015).
44. Piehler, J. New methodologies for measuring protein interactions *in vivo* and *in vitro*. *Curr. Opin. Struct. Biol.* **15,** 4–14 (2005).
45. Parker, J. L., Mindell, J. A. & Newstead, S. Thermodynamic evidence for a dual transport mechanism in a POT peptide transporter. *Elife* **3,** 1–13 (2014).
46. Henrici, R. C., Pecen, T. J., Johnston, J. L. & Tan, S. The pPSU Plasmids for Generating DNA Molecular Weight Markers. *Sci. Rep.* 1–9 (2017). doi:10.1038/s41598-017-02693-1
47. Love, J. *et al.* The New York Consortium on Membrane Protein Structure (NYCOMPS): a high-throughput platform for structural genomics of integral membrane proteins. *J. Struct. Funct. Genomics* **11,** 191–199 (2010).
48. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6,** 12–17 (2009).
49. Széliová, D., Krahulec, J., Safránek, M. & Tur, J. Modulation of heterologous expression from P BAD promoter in *Escherichia coli* production strains. **236,** 1–9 (2016).

50. Wagner, S. *et al.* Consequences of Membrane Protein Overexpression in *Escherichia coli*. *Mol. Cell. proteomics* **6.9,** 1527–1550 (2007).
51. Jia, B. & Jeon, C. O. High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open Biol.* **6**, 160-196 (2016).
52. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189,** 113–130 (1986).
53. Hannig, G. & Makrides, S. C. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends Biotechnol.* **16,** 54–60 (1998).
54. Zhao, D. & Huang, Z. Effect of His-Tag on Expression, Purification, and Structure of Zinc Finger Protein, ZNF191(243-368). *Bioinorg. Chem. Appl.* **2016**, 820-854 (2016).
55. Booth, W. T. *et al.* Impact of an N‑terminal Polyhistidine Tag on Protein Thermal Stability. *ACS Omega* **3,** 760–768 (2018).
56. Sørensen, H. P. & Mortensen, K. K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.* **115,** 113–128 (2005).
57. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli* : advances and challenges. *Front. Microbiol.* **5,** 1–17 (2014).
58. Keck, W., Van Leeuwen, A. M., Huber, M. & Goodell, E. W. Cloning and characterization of mepA, the structural gene of the penicillin-insensitive murein endopeptidase from *Escherichia coli*. *Mol. Microbiol.* **4,** 209–219 (2018).
59. Raynal, B., Lenormand, P., Baron, B., Hoos, S. & England, P. Quality assessment and optimization of purified protein samples: why and how? *Microb. Cell Fact.* **13,** 1–10 (2014).
60. Bouvier, E. S. P. & Koza, S. M. Advances in size-exclusion separations of proteins and polymers by UHPLC. *Trends Anal. Chem.* **63,** 85–94 (2014).
61. Drew, D., Lerch, M., Kunji, E., Slotboom, D. & Gier, J. De. Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods* **3,** 303-313
(2006).
62. Festa, F., Steel, J., Bian, X. & Labaer, J. High-Throughput Cloning and Expression Library Creation for Functional Proteomics. *Proteomics* **13,** 1381–1399 (2013).
63. Crisman, T. J., Qu, S., Kanner, B. I. & Forrest, L. R. Inward-facing conformation of glutamate transporters as revealed by their inverted-topology structural repeats. *Proc. Natl. Acad. Sci.* **106,** 20752–20757 (2009).
64. Yokoyama, S. Protein expression systems for structural genomics and proteomics.

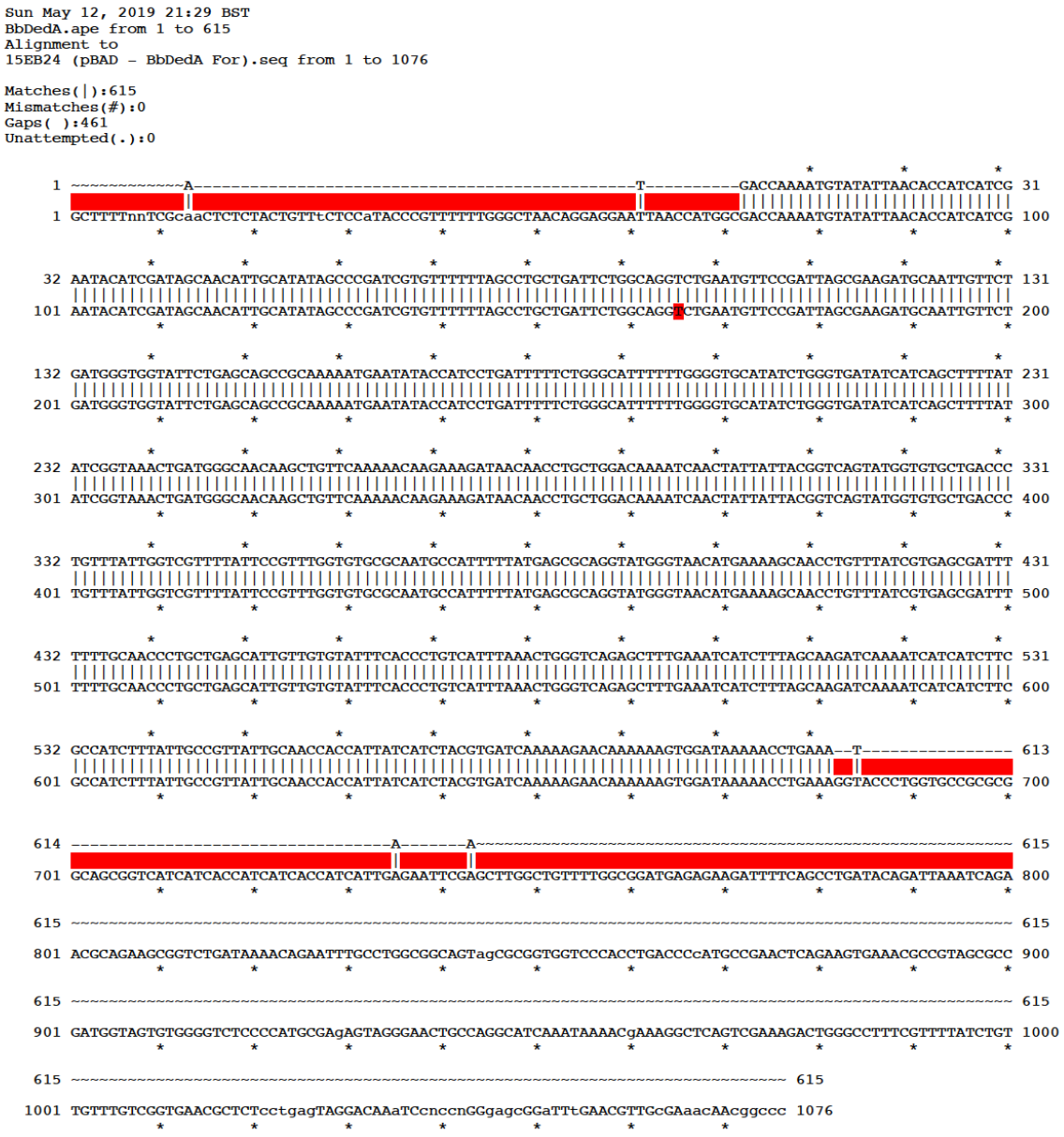
Curr. Opin. Chem. Biol. **7,** 39–43 (2003).

65. Sørensen, H. P., Sperling-petersen, H. U. & Mortensen, K. K. Production of recombinant thermostable proteins expressed in *Escherichia coli*: completion of protein synthesis is the bottleneck. *J. Chromatogr. B* **786,** 207–214 (2003).
66. Shen, H., Lithgow, T. & Martin, L. L. Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities. *PLoS Comput. Biol.* 1589–1607 (2013). doi:10.3390/ijms14011589
67. Liguo, W. & Lige, T. Membrane protein reconstitution for functional and structural studies. *Sci. China Life Sci.* **58,** 66–74 (2015).
68. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet. Infect. Dis.* **16**, 161–8 (2016).
69. Rubin, B. C. *et al.* Structural features of the hisT operon of *Escherichia coli K-12*. *Nucleic Acids Res.* **13**, 5297–5315 (1985).
70. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, 172 (2014).
71. Hoffmann, F. & Rinas, U. Stress Induced by Recombinant Protein Production in Escherichia coli BT - Physiological Stress Responses in Bioprocesses; 73–92 (Springer Berlin Heidelberg, 2004). doi:10.1007/b93994
72. Studier, F. W. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* **219**, 37–44 (1991).
73. Lobell, R. B. & Schleif, R. F. DNA looping and unlooping by AraC protein. *Science (80-.).* **250**, 528 LP – 532 (1990).
74. Schmitt, J., Hess, H. & Stunnenberg, H. G. Affinity purification of histidine-tagged proteins. *Mol. Biol. Rep.* **18**, 223–230 (1993).
75. le Maire, M., Champeil, P. & Møller, J. V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta - Biomembr.* **1508**, 86–111 (2000).
76. Postis, V. *et al.* The use of SMALPs as a novel membrane protein scaffold for structure study by negative stain electron microscopy. *Biochim. Biophys. Acta - Biomembr.* **1848**, 496–501 (2015).
77. Ji, Y., Postis, V. L. G., Wang, Y., Bartlam, M. & Goldman, A. Transport mechanism of a glutamate transporter homologue GltPh. *Biochem. Soc. Trans.* **44**, 898–904 (2016).

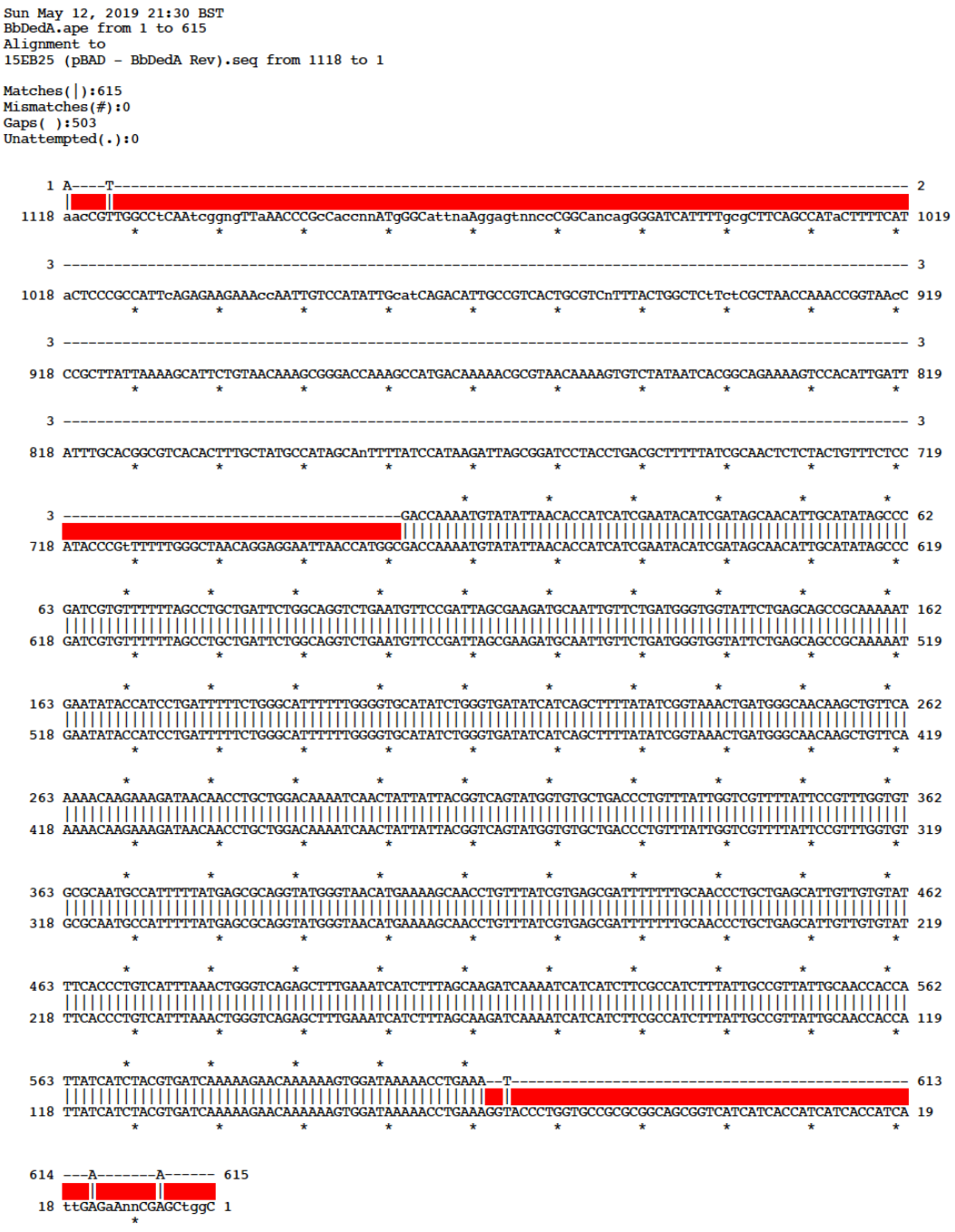
78. Chayen, N. E. & Saridakis, E. Protein crystallization: From purified protein to diffraction-quality crystal. *Nat. Methods* **5**, 147–153 (2008).
79. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **70**, 2–20 (2014).
80. Smyth, M. S. & Martin, J. H. J. X Ray crystallography. *Mol Pathol* **53**, 8–14 (2000).
81. Sutherland, B. W., Toews, J. & Kast, J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein–protein interactions. *J. Mass Spectrom.* **43**, 699–715 (2008).

8. Appendix

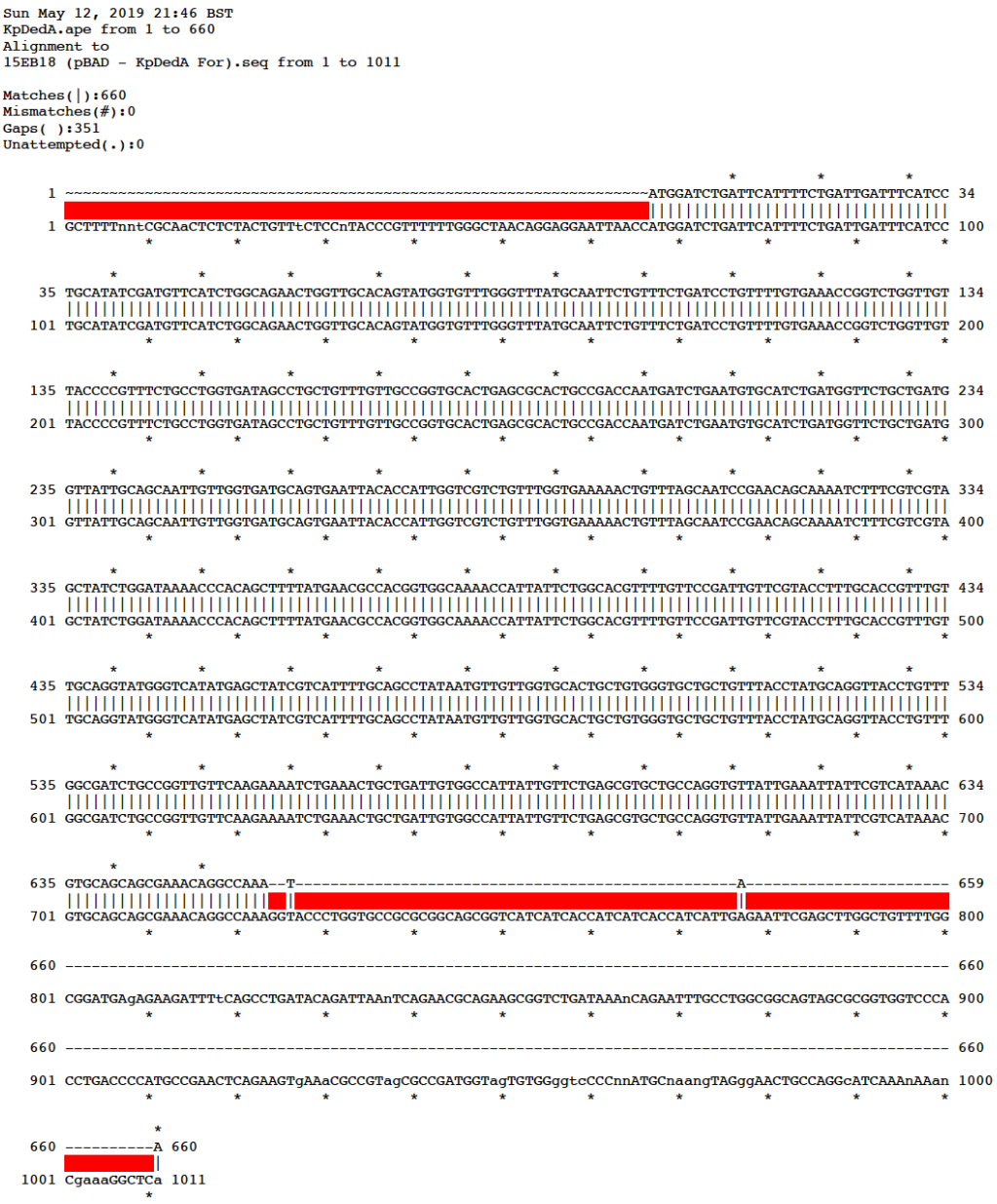
Included below is the sequencing data for each of the constructs included in this study, excluding pETHis*saDedA*, pETHis*vcDedA*, and pBADHis*vcDedA*.



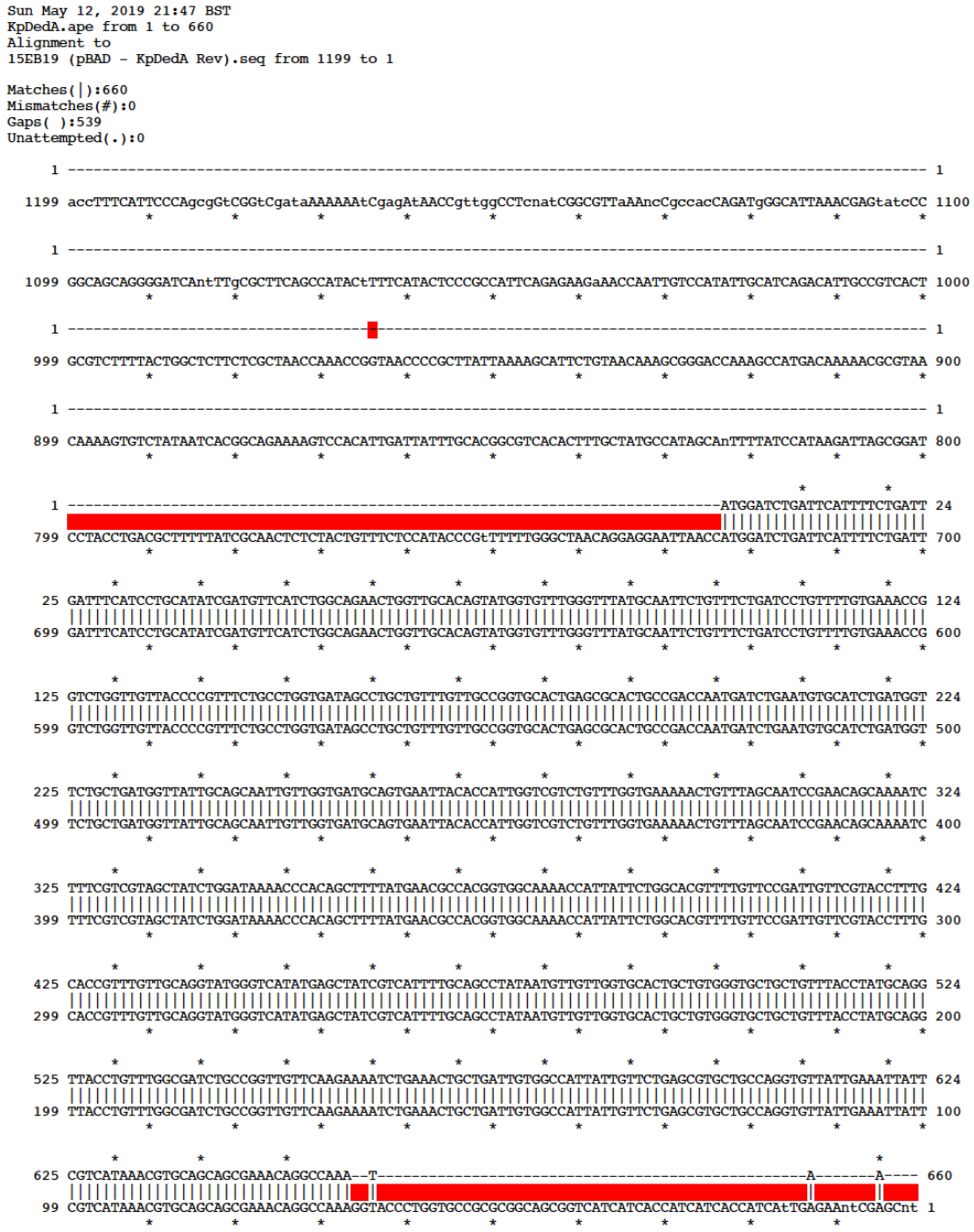
Appendix Figure 1: Alignment of forward sequencing for pBADHis*bbDedA* to the desired construct.



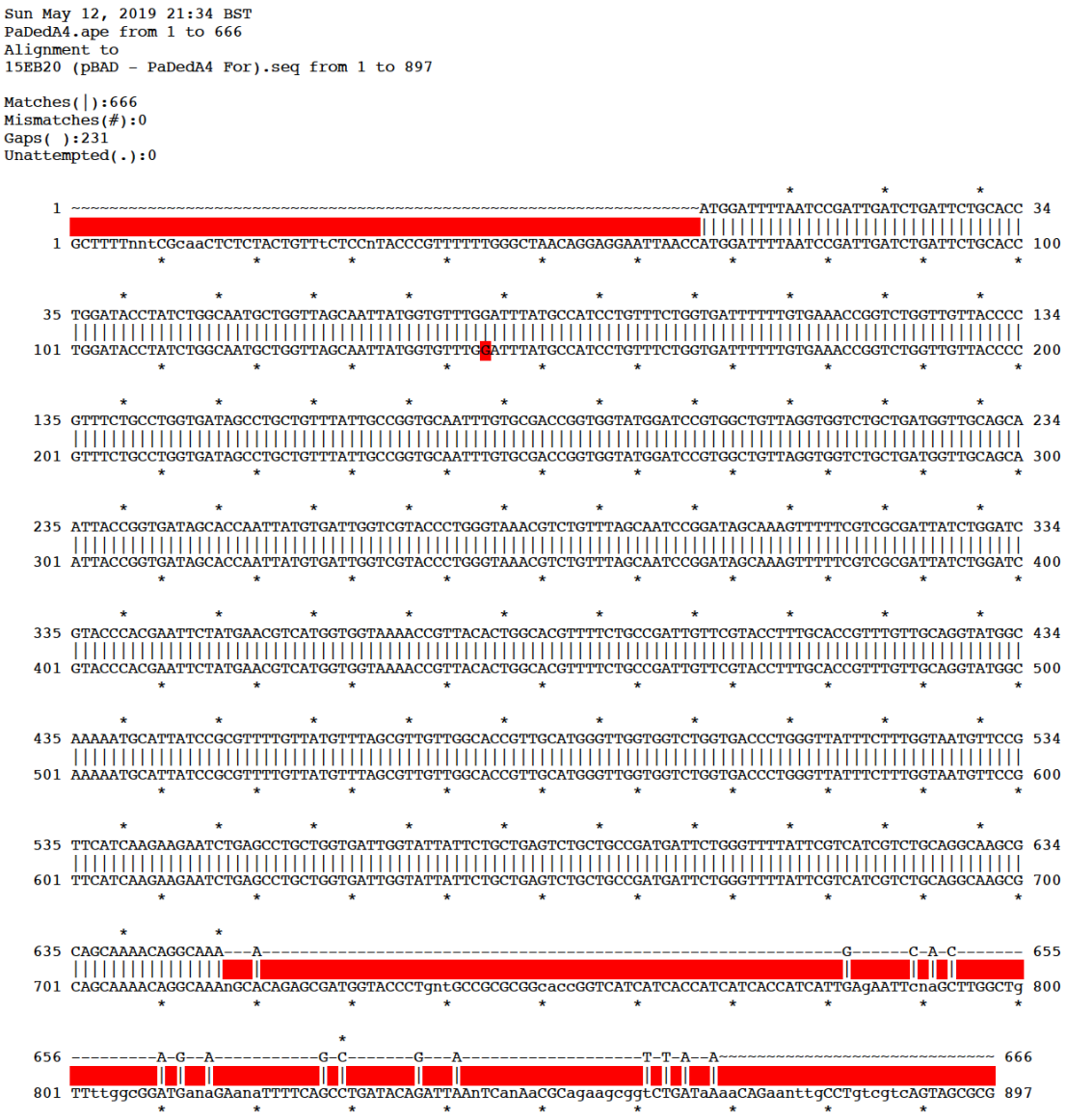
Appendix Figure 2: Alignment of reverse sequencing for pBADHis*bbDedA* to the desired construct.



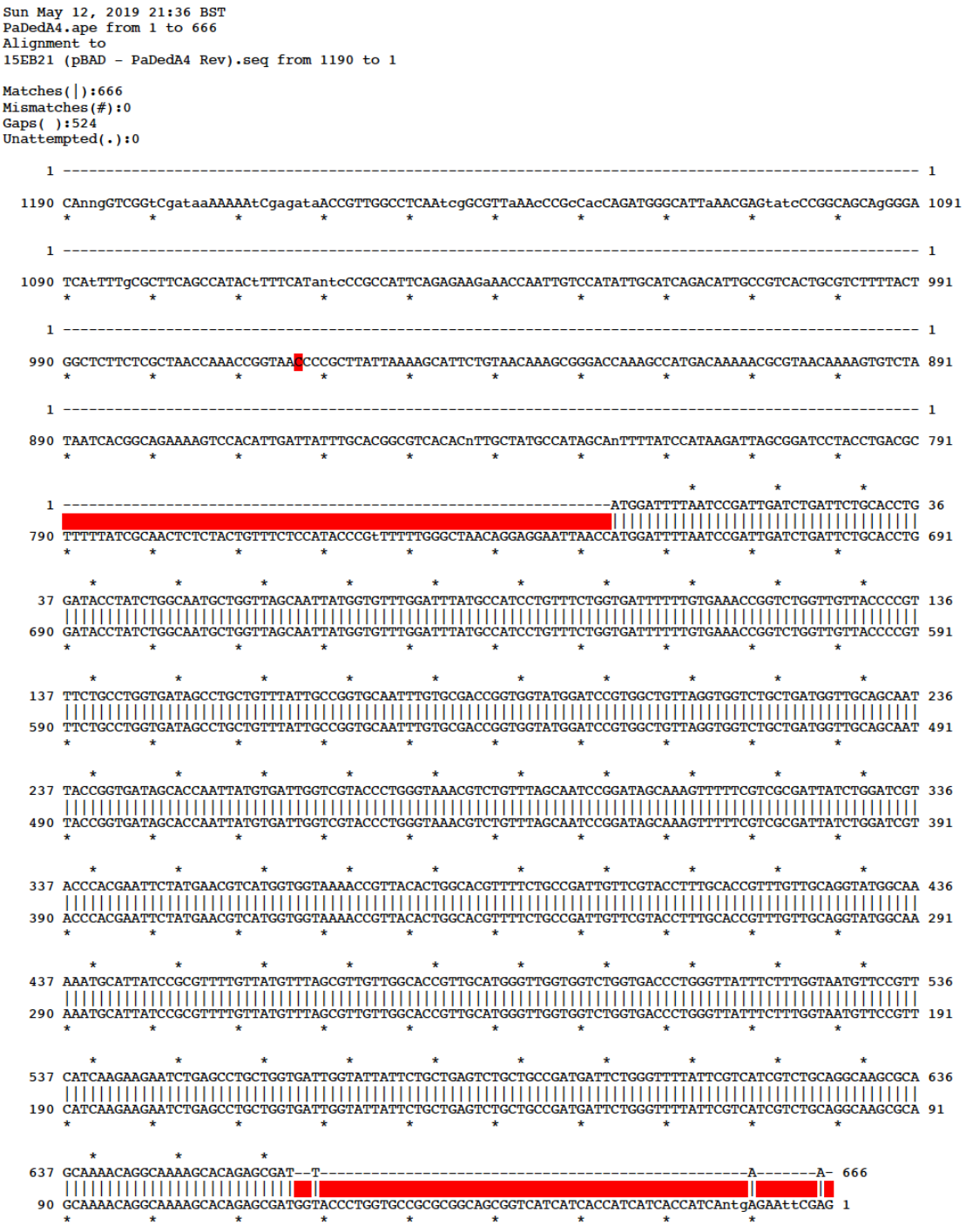
Appendix Figure 3: Alignment of forward sequencing for pBADHis*kpDedA* to the desired construct.



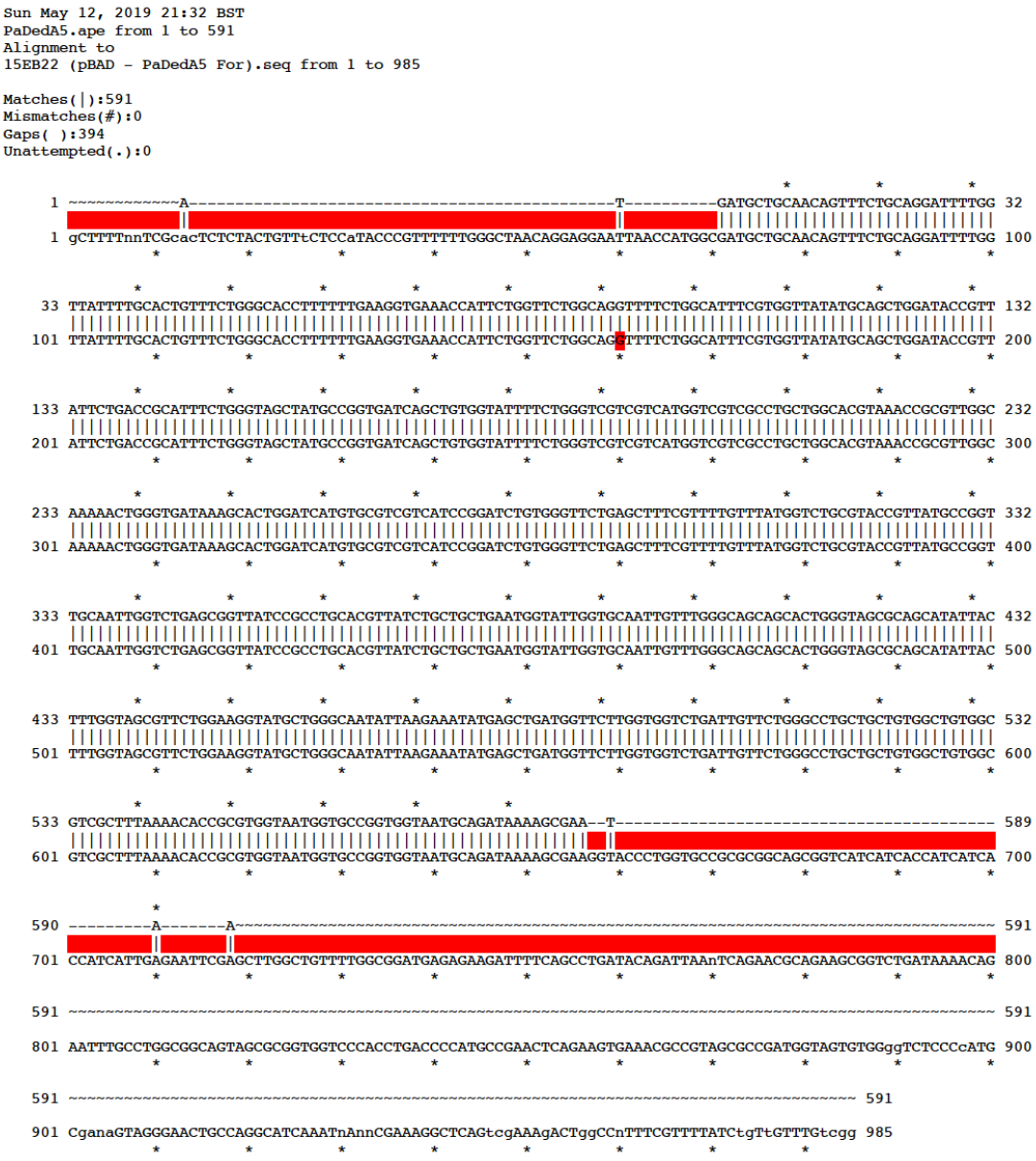
Appendix Figure 4: Alignment of reverse sequencing for pBADHis*kpDedA* to the desired construct.



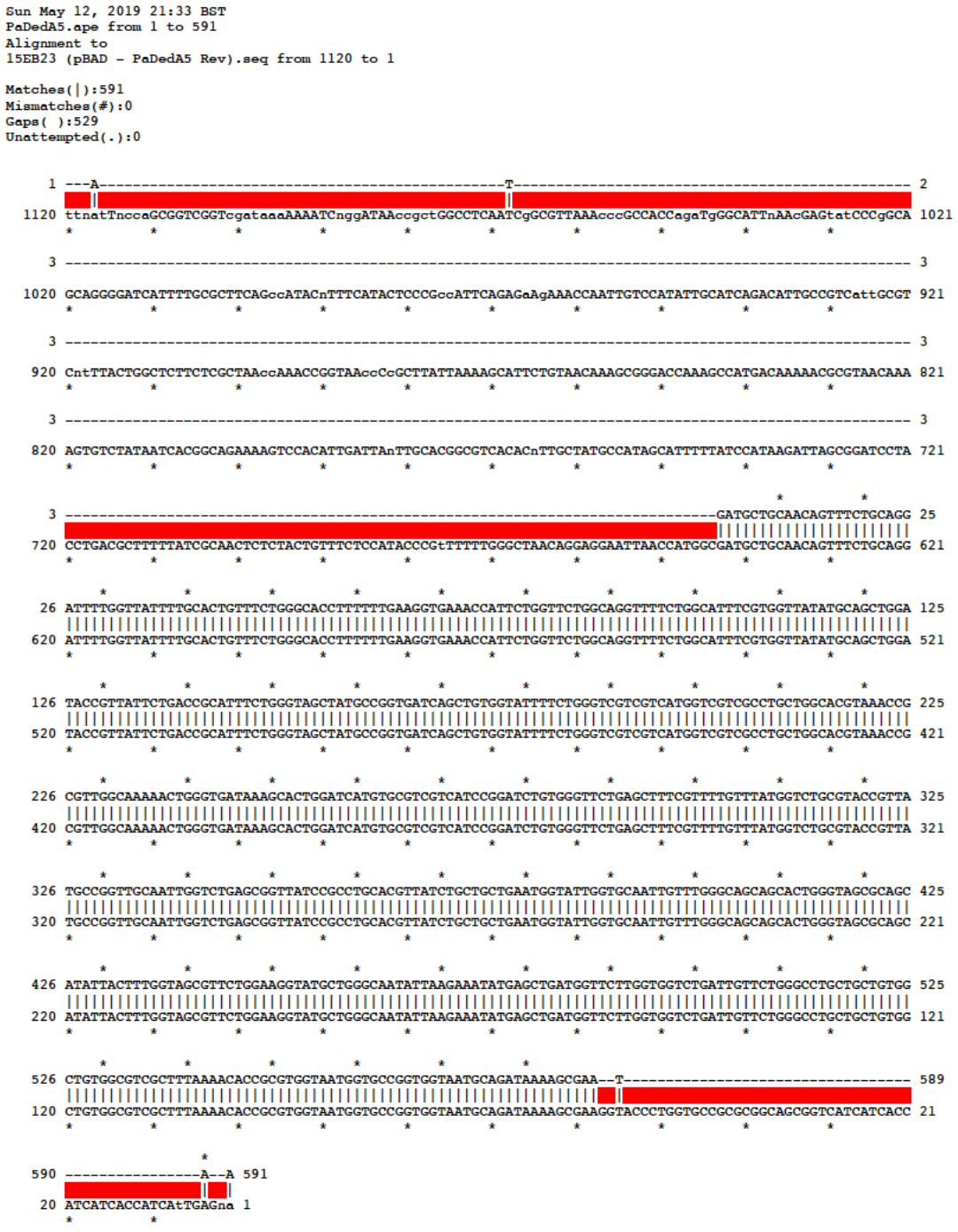
Appendix Figure 5: Alignment of forward sequencing for pBADHis*paDedA4* to the desired construct.



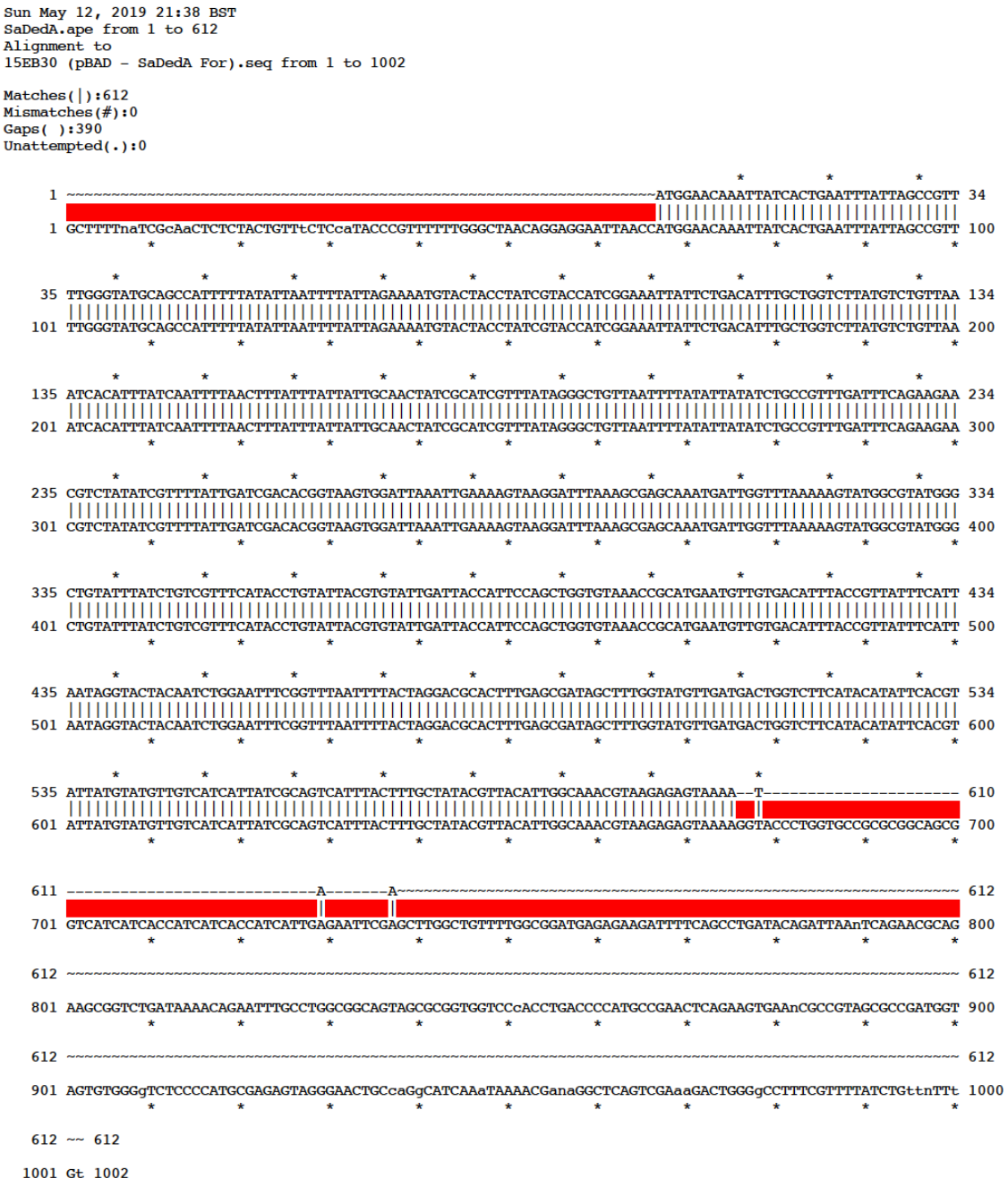
Appendix Figure 6: Alignment of reverse sequencing for pBADHis*paDedA4* to the desired construct.



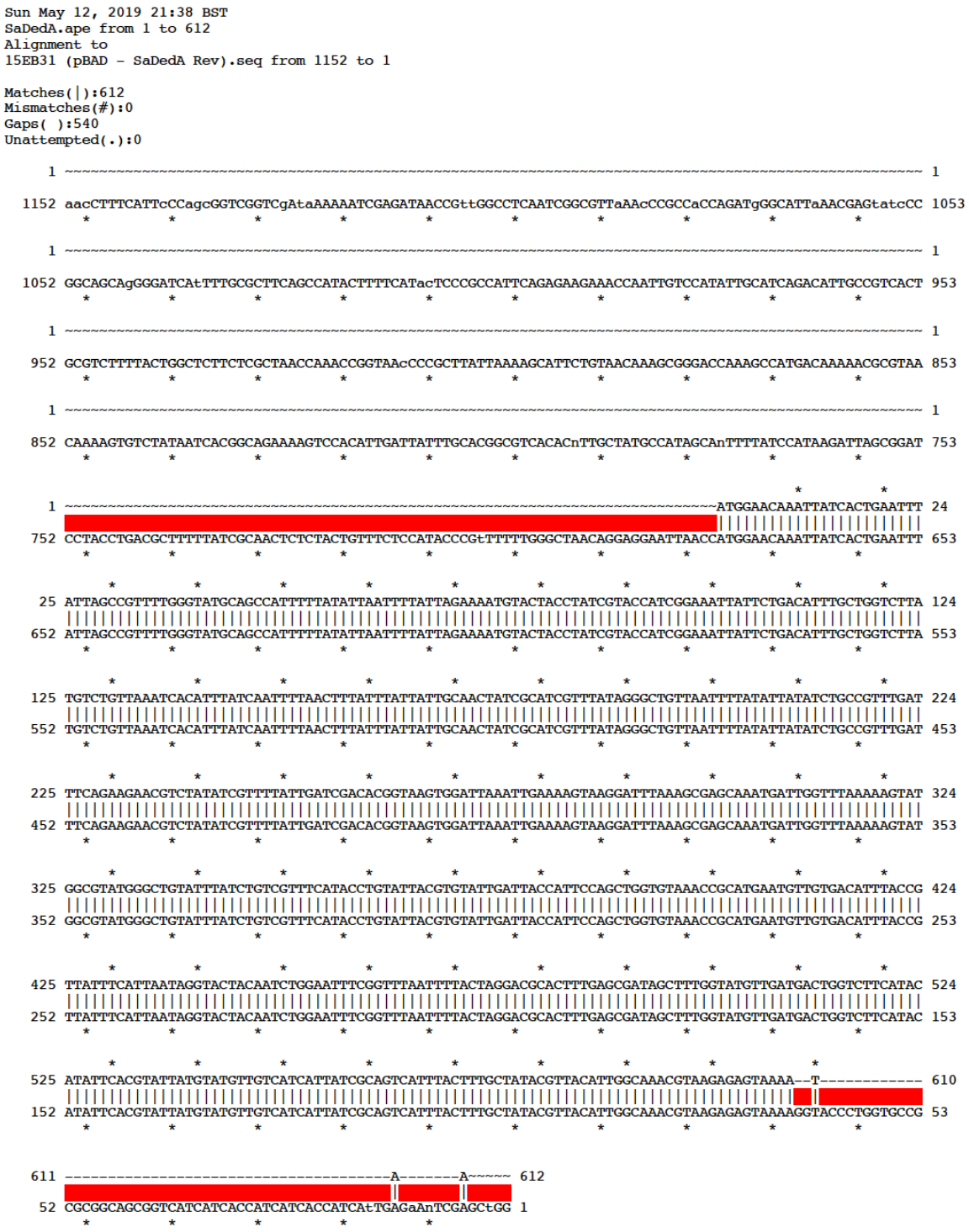
Appendix Figure 7: Alignment of forward sequencing for pBADHis*paDedA5* to the desired construct.



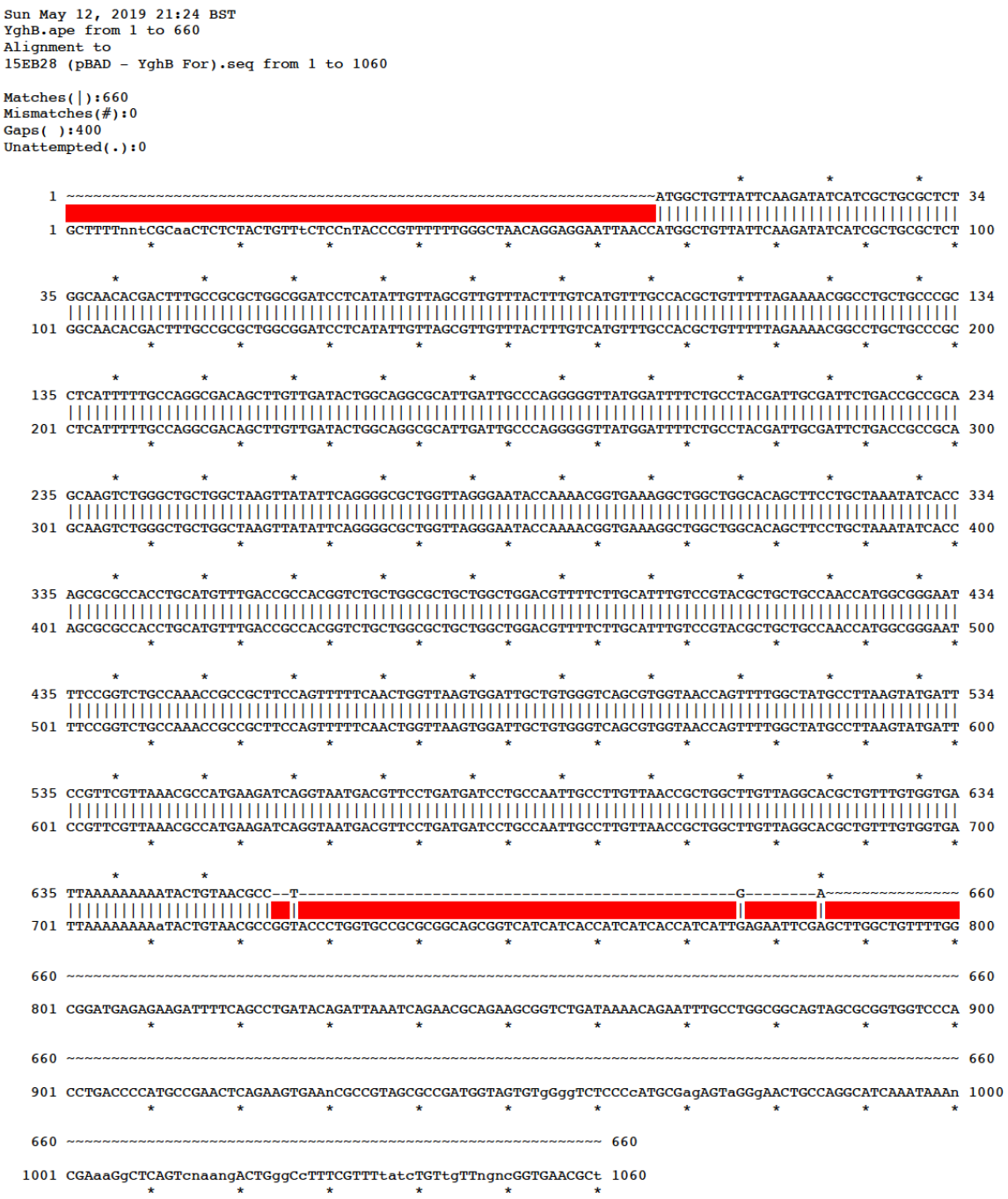
Appendix Figure 8: Alignment of reverse sequencing for pBADHis*paDedA5* to the desired construct.



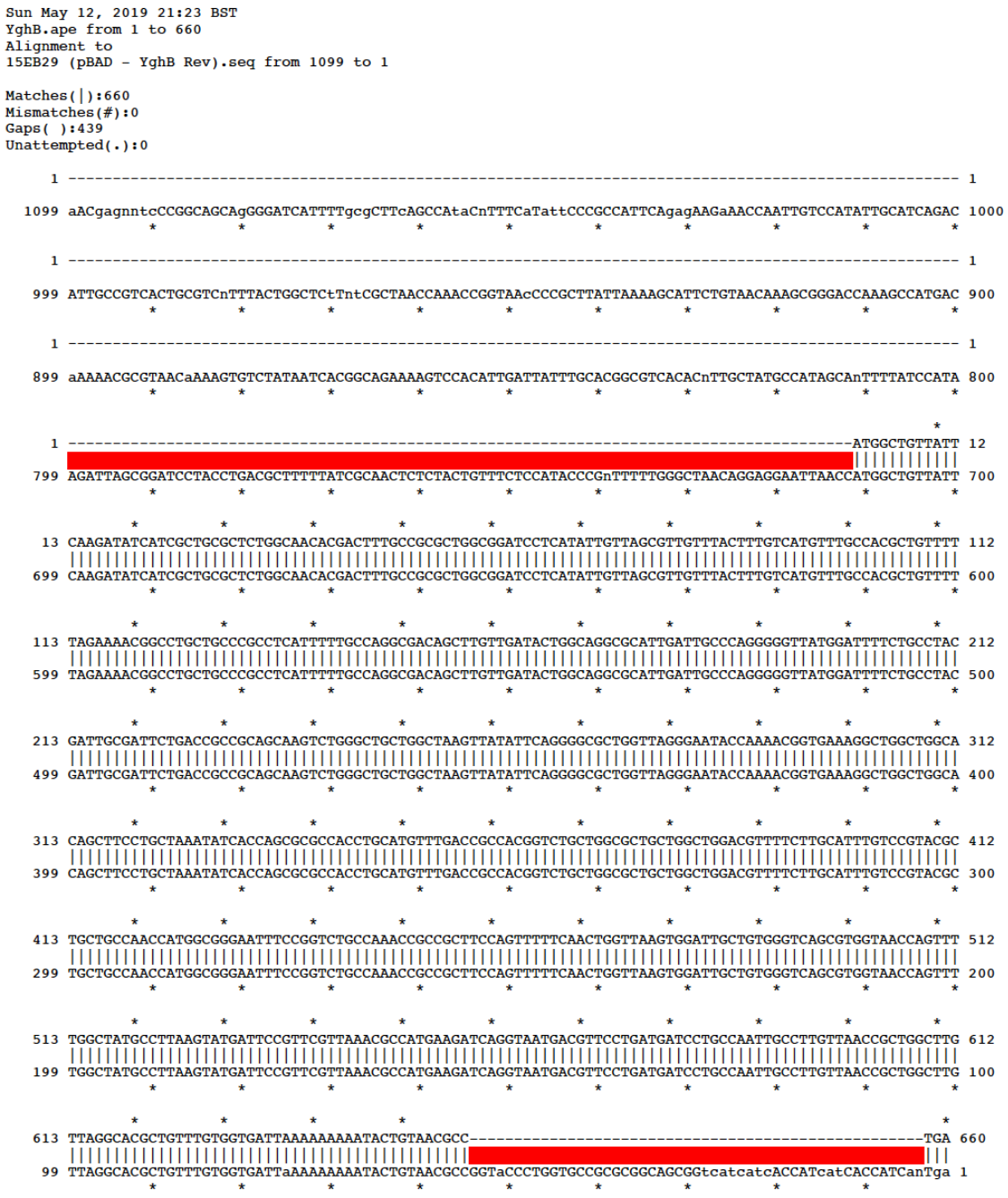
Appendix Figure 9: Alignment of forward sequencing for pBADHis*saDedA* to the desired construct.



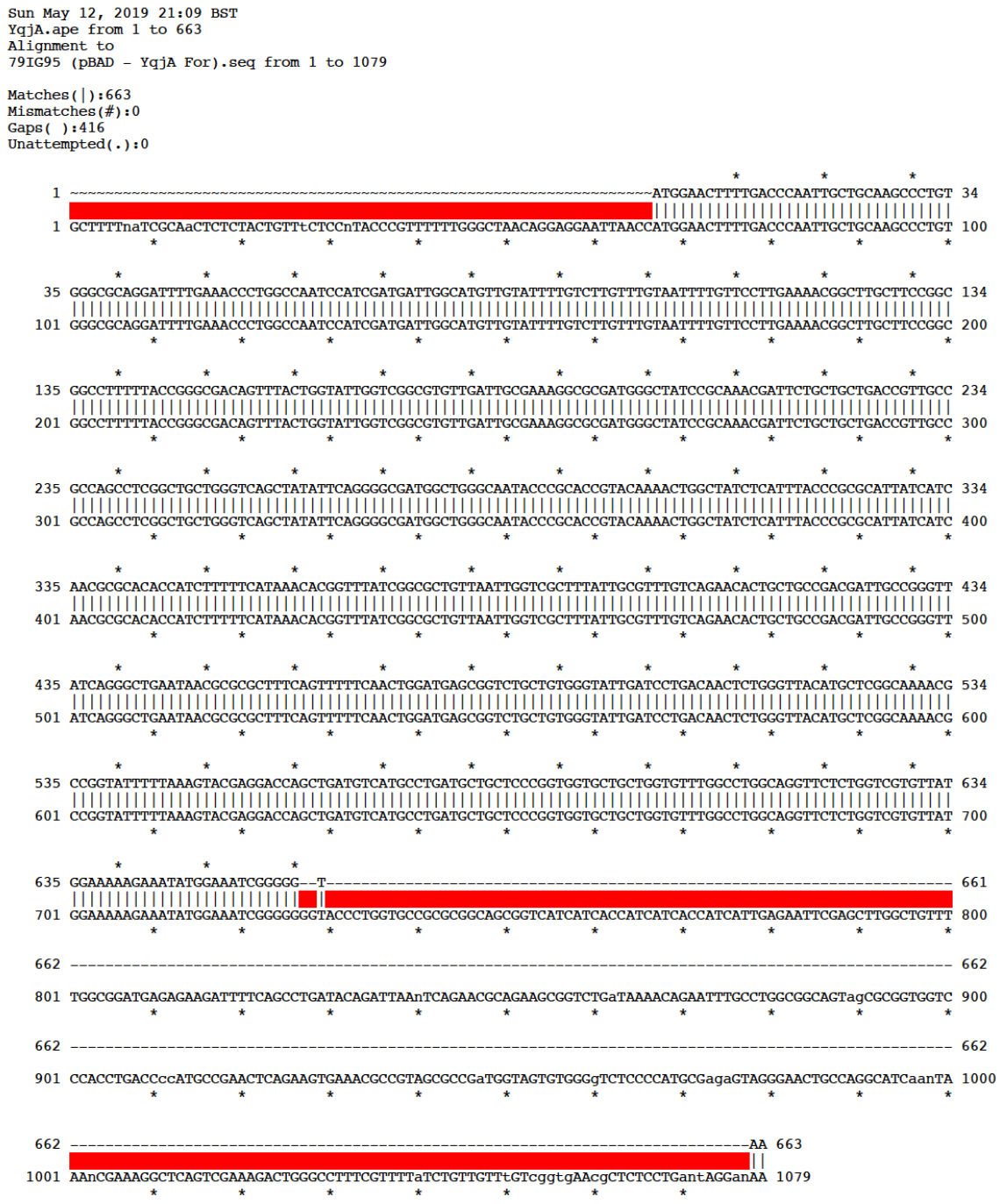
Appendix Figure 10: Alignment of reverse sequencing for pBADHis*saDedA* to the desired construct.



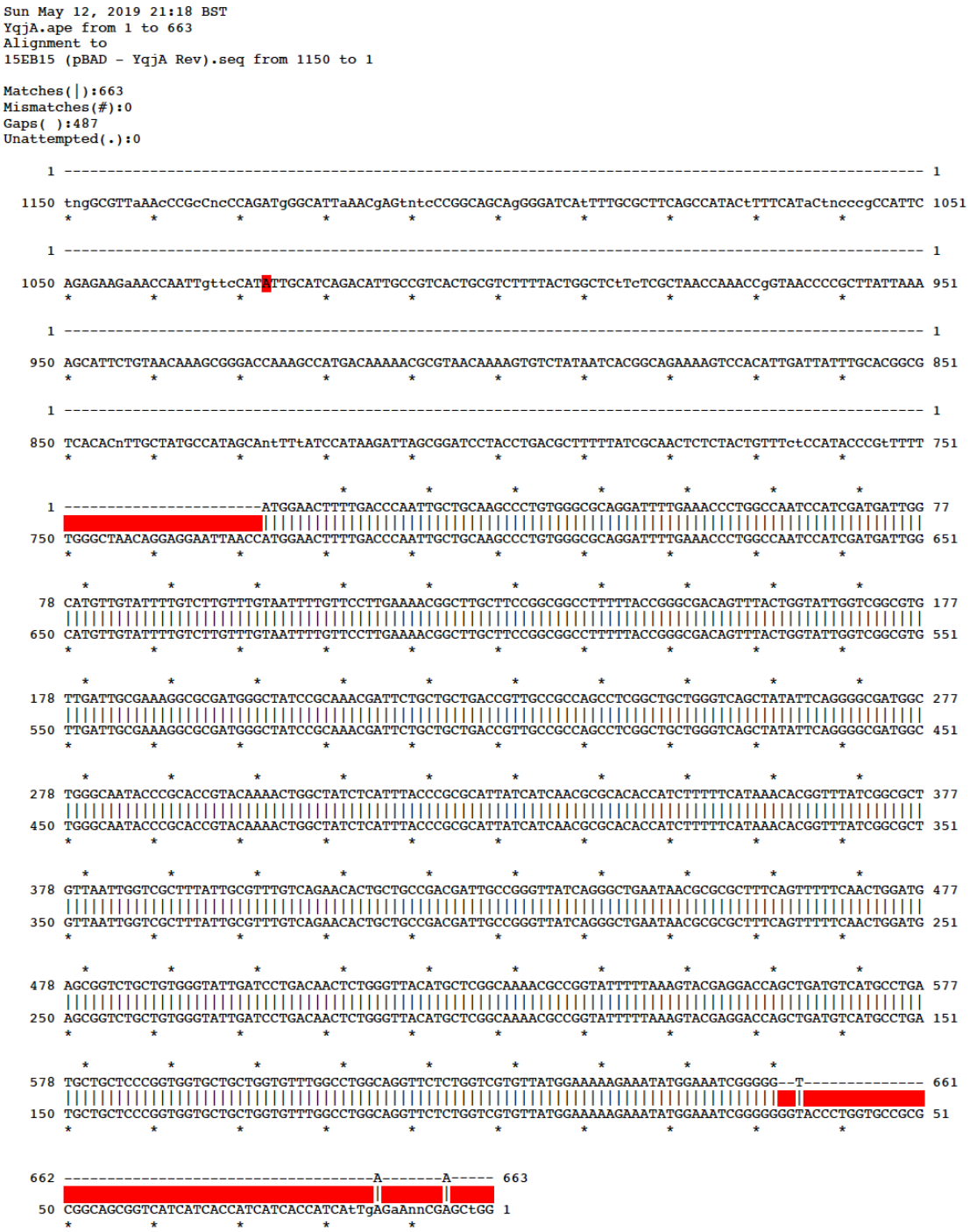
Appendix Figure 11: Alignment of forward sequencing for pBADHis*yghB* to the desired construct.



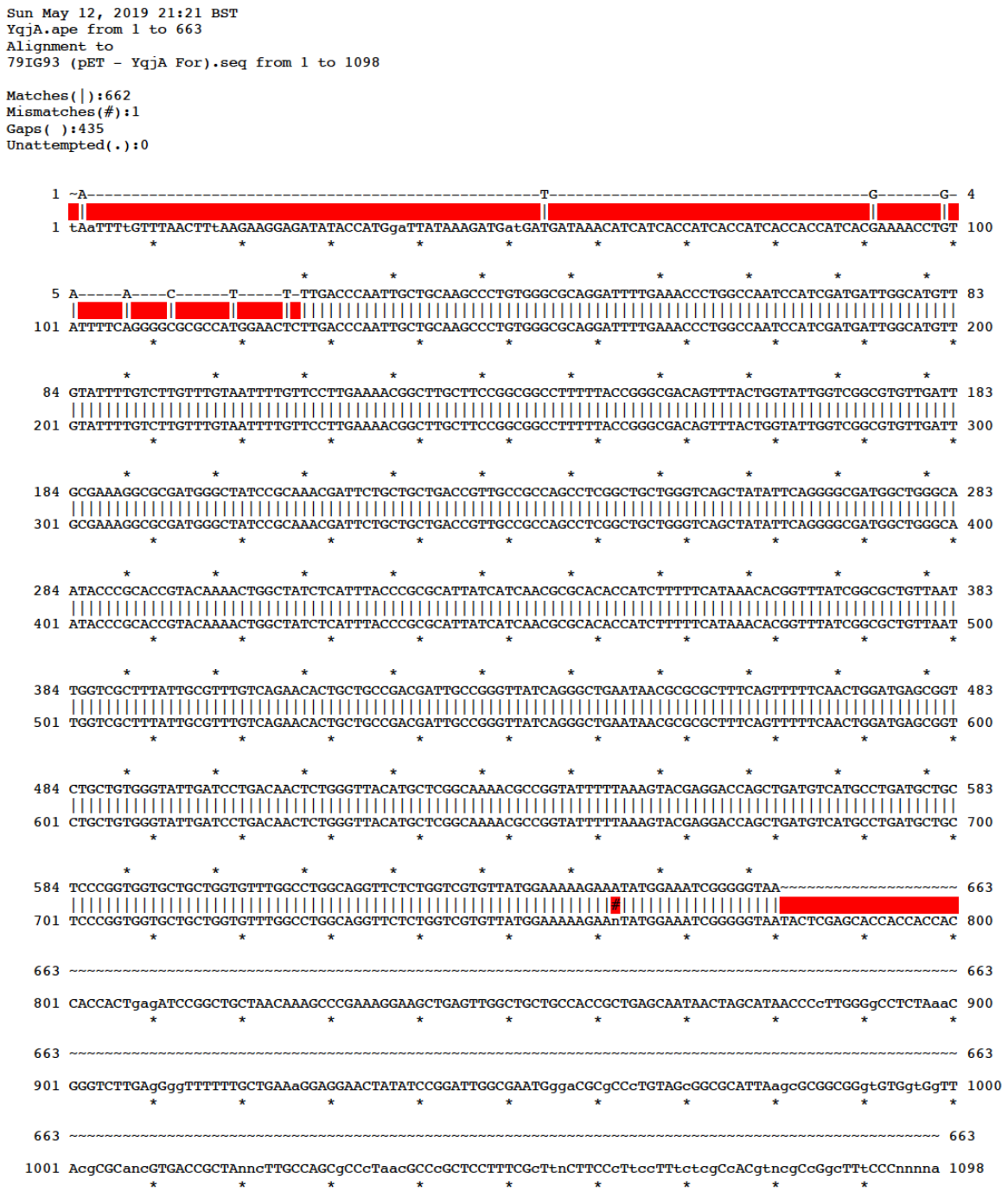
Appendix Figure 12: Alignment of reverse sequencing for pBADHis*yghB* to the desired construct.



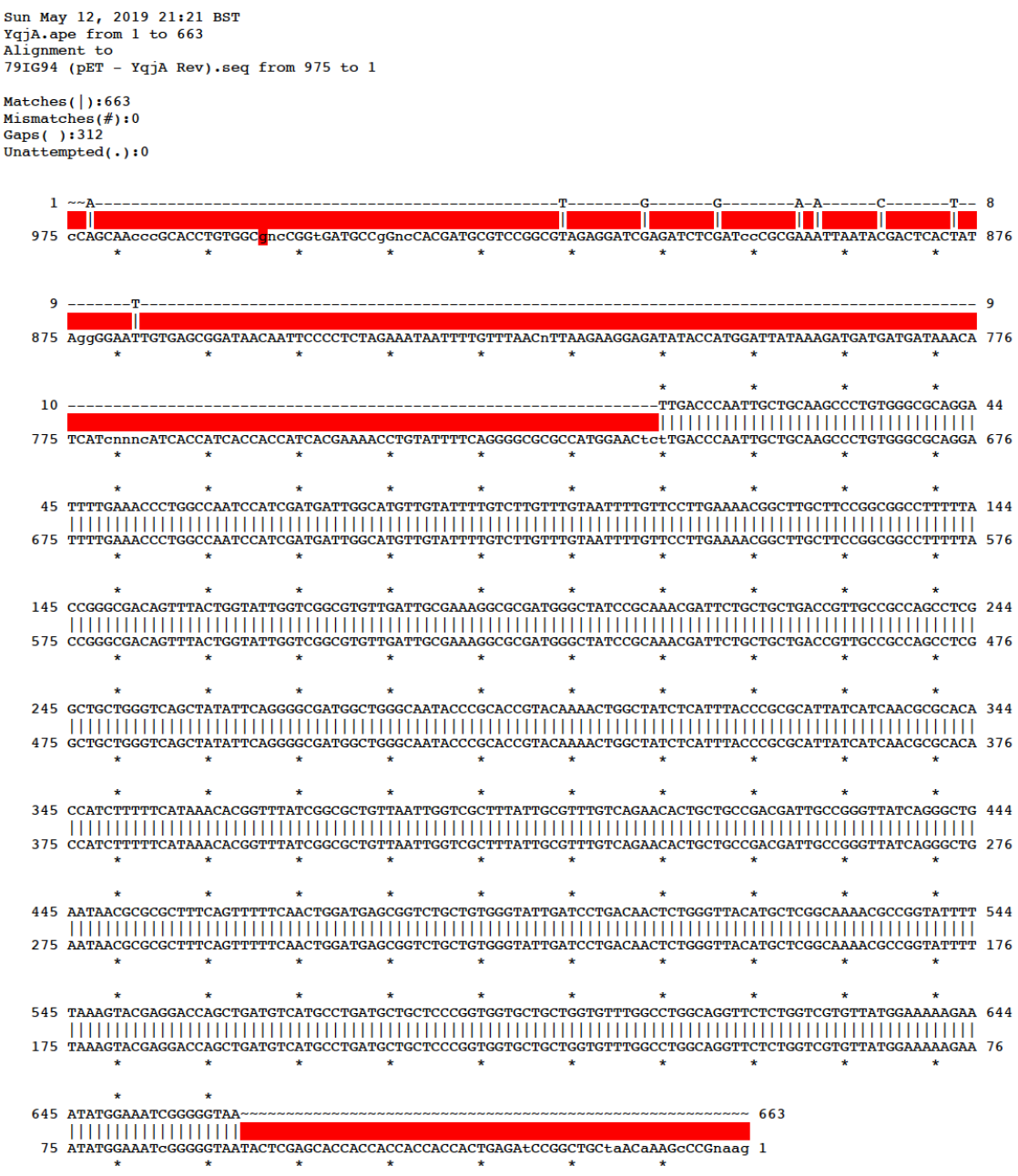
Appendix Figure 13: Alignment of forward sequencing for pBADHis*yqjA* to the desired construct.



Appendix Figure 14: Alignment of reverse sequencing for pBADHis*yqjA* to the desired construct.



Appendix Figure 15: Alignment of forward sequencing for pETHis*yqjA* to the desired construct.



Appendix Figure 16: Alignment of reverse sequencing for pETHis*yqjA* to the desired construct.