2016

Characterization of a Phosphonate-Specific Cytidylyltransferase

Kissa Batul
Wilfrid Laurier University, kissabatul@gmail.com

Characterization of a Phosphonate-Specific Cytidylyltransferase

By

Kissa Batul

(Honours BSc. Biochemistry and Biotechnology, Wilfrid Laurier University, 2014)

THESIS

Submitted to the Department of Chemistry and Biochemistry

Faculty of Science

in partial fulfillment of the requirements for the

Master of Science in Chemistry and Biochemistry

Wilfrid Laurier University

2016

Abstract

Antibiotic resistance is a major global health concern that requires new therapeutic approaches. Furthermore, a lack of narrow spectrum antibiotics on the market produces unintended consequences with respect to changes in our microbial make up. Phosphonates are reduced versions of phosphates that possess a C-P bond which is more resistant to enzymatic and chemical degradation. The role of phosphonate containing macromolecules (e.g. cell surface polysaccharides) remains enigmatic, however their presence suggests that they may confer an advantage. The biosynthesis of phosphonate-containing macromolecules is unknown, but a pathway is proposed involving aLicC type cytidylyltransferase-catalyzed conjugation to a phosphonate followed by a LicD type phosphotransferase- catalyzed attachment to a macromolecule. This proposed pathway is analogous to the partially characterized phosphocholine (PC) tailoring pathway in teichoic acid biosynthesis in *Streptococcus pneumoniae.* In this study, the LicC homolog PngC from *Atopobium rimae* (Ari-PngC) was purified and its activity was compared to LicC from *S. pneumoniae* (Spn-LicC). Significantly, Ari-PngC preferred the phosphonate substrate 2-aminoethylphosphonate (AEP) over PC, and vice versa for Spn-LicC. Specifically, the K_M and k_{cat} values for Spn-LicC towards PC were 0.020 ±0.011 mM and 1.52 ± 0.243 s^{-1} respectively, yielding k_{cat}/K_M of 77.9 M^{-1} s^{-1}. In contrast, the K_M value for Spn-LicC towards AEP was 0.318 ± 0.126 mM and k_{cat} of 0.722 ± 0.053 s^{-1}, yielding kcat/km of 2.27 M-1 s-1 and revealing a 34-fold preference for PC. The opposite was seen for Ari-PngC, where AEP gave a K_M value of 0.011 ± 0.001 mM , k_{cat} of 2.72 ± 0.079 s^{-1} and k_{cat}/K_M of 239 M^{-1} s^{-1}, while towards PC

Ari-PngC yielded specificity constants of 1.67 and 0.74 when modelled by the Michaelis-Menten or substrate inhibition equation, respectively. Regardless of the correct fit, PngC clearly preferred AEP over PC. Overall, these results establish clear substrate selectivity of phosphonate versus PC tailoring pathways and set the stage for developing narrow spectrum antimicrobials.

Statement of Work

The cloning of PngC from *A. rimae* into pET29 and cloning of PngC from *O. uli* into pET29 was completed by Jayne Kelso and Maddison Bibby. The cloning of LicC from *S.pneumoniae* was completed by Rebecca Sullivan.

Acknowledgements

There are many people without whom I could have not finished my thesis. First and foremost, I would like to give my sincerest acknowledgement to my supervisor, Dr. Geoff Horsman, for his excellent guidance. The work presented in this thesis would have not been possible without his support and encouragement throughout my graduate studies. My thanks also goes to my thesis committee, Dr. Michael Suits, and Dr. Gabriel Moreno Hagelsieb for providing great advice.

I am grateful to all the members of the Horsman lab, past and present for joining me in my lab dance parties and creating an enjoyable research environment. I am very fortunate to have them as my colleagues. I would like to give a special shoutout to Mark Aliwalas for being my partner in crime and joining me in all my singing sessions in lab. All these people have kept me sane during the most challenging times. I love you guys!

I would also like to give a special thanks to Jayne Kelso, Maddison Bibby and Rebecca Sullivan for helping me with my project.

Finally, I would like to thank my family for their endless love, support and understanding.

Table of Contents

List of Tables

List of Figures

1.1 Human Microbiome and Antibiotic Resistance

The human microbiome describes all microorganisms that reside on and within the human body[1]. The list of microorganisms include bacteria, fungi, and viruses. Every individual has a unique microbiome that is acquired from the environment at birth. The microbial profile within the microbiome varies with gender, diet, climate, age and hygiene[1]. Microorganisms within the microbiome have been studied to determine their interaction with human cells and the benefit or harm they might bring to it. Microorganisms, or microbes have been long viewed as things to be gotten rid of, but as researches learn more about microbes, they have discovered the importance of microbes within the body. Although microbes are important in maintaining a healthy body, imbalances to microbial populations can give rise to opportunistic bacteria which can lead to disease[1].

The discovery of antibiotics to treat bacterial infections have dramatically changed human health, and many once deadly infections are now curable[3]. Yet, there has been a rise in bacteria that are no longer effectively killed by antibiotics[3,4]. These bacteria are known as antibiotic resistant, and are a growing problem in medicine around the world. It has been estimated that more than 18,000 hospitalized patients in Canada acquire antimicrobial resistance infections[3,4,5.]. There is also a seven-fold increase in the incidence of Vancomycin resistant Enterococci infections between 2007-2012 in Canada alone[3]. Not only is antibiotic resistance an issue that threatens the human health, it is also the lack of narrow spectrum antibiotics available in the market[4]. Broad spectrum antibiotics disrupt the human microbiota because not only do they kill harmful bacteria, they also kill the surrounding non harmful bacteria. This allows

for bacteria to have an increased selective pressure to develop resistance. New solutions are ungently needed to counteract this loss in antimicrobial efficacy[4].

1.2 Phosphorus Containing Compounds

Phosphorus is essential to all living organisms as primary constituent of DNA , RNA and ATP[6]. It is also a key structural component of phospholipids, phosphoproteins, phosphorylated exopolysaccharides and numerous metabolites. In addition, phosphorus provides buffering capacity, impacts the solubility of organic molecules and can provide a high concentration of negative charge within a given molecular dimension[6]. Phosphorus predominantly exists in the most oxidized state (+5 valence) in the form of inorganic phosphate, phosphate esters and phosphoanhydrides[8]. However, there is a naturally occurring phosphorus compounds called phosphonates which are found in a +3 valence state[8]. In the recent years, there has been a growth in literature on this group of less studied phosphorus compounds. Phosphonates are a chemically stable group of compounds containing a direct carbon-phosphorus (C-P) bond[7]. Since their relatively recent discovery of the first naturally occurring phosphonate, 2-aminoethylphosphonic acid (AEP), in the acid hydrolysate of rumen protozoan in 1959, C-P compounds have been identified in a number of bacteria, archaea, and eukaryotes[9,10,11]. In many living organisms these C-P bond compounds are found as constituents of extracytoplasmic macromolecules such as phospholipids, phosphonoglycans, glycoproteins and lipids[12].

Studies by several research group showed that 2-AEP and derivatives were common constituents in lipids in a number of ciliated protozoa, anemones, corals and mollusks. The lipid backbone, where 2-AEP or derivatives is bound, is usually either a

ceramide as in sphingophospholipids, or a diacylglycerol as in glycerolphosphonolipids[13]. Other phosphonate head group have also been reported such as 1-hydroxy-2-AEP found in a sphingophosphonolipid from *Bdellovibrio stolpii* and N-acyl-AEP and N,N-acylmethyl-AEP found in a sphingophosphonolipid from *Corbicula sandai*[14].

Phosphate is a common modification of polysaccharides, but phosphonate containing polysaccharide examples are not very common. A phosphonate containing polysaccharide or a phosphonoglycan was initially isolated from the plasma membrane of the soil amoeba *Acanthamoeba castellanii* and was shown to be a lipopshosphonoglycan[15]. Among bacteria, *Bacteroides fragilis NCTC 9343* produces a capsular polysaccharide complex (CPC), (See Figure 1) which is directly involved in abscess formation in animal models[16]. A capsular polysaccharide has also been isolated from the outer membrane of the luminal bacterium *Fibrobacter succinogenes S85*[17].

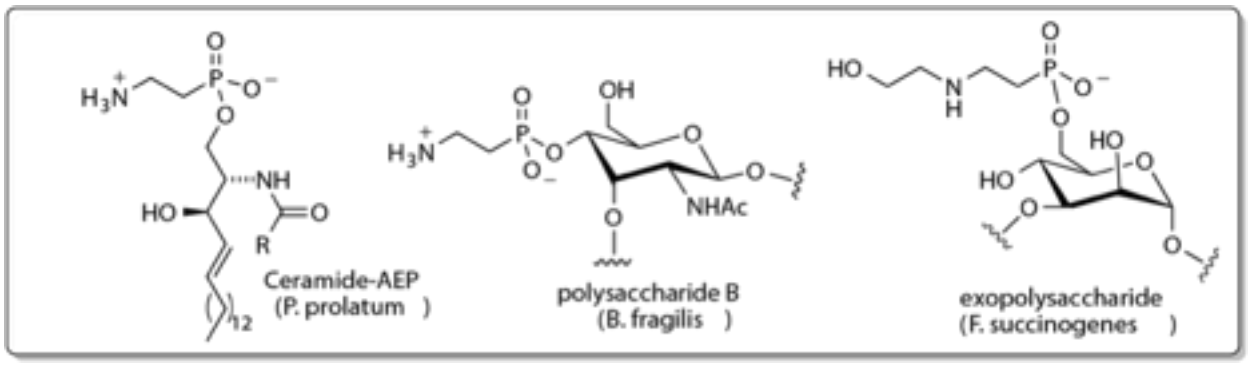


Figure 1. Phosphonate containing macromolecules isolated. Some examples include the phosphonolipid ceramide -AEP from *Pythium prolatum*, the capsular polysaccharide B (PBS) from *Bacteroides fragilis* and an exopolysaccharide from *Fibrobacter succinogenes*[16,17]

The most striking example of phosphonoglycan occurrence is in the freshly laid egg masses of fresh water snail *Helisoma,* where almost 85% of the phosphorus is in the

form of AEP and another unknown phosphonate linked to carbohydrates[18]. Phosphonates are now known to be produced by many primitive life forms, in which they are normally found as a side group on exopolysaccharides and as a feature of the glycan decoration present on glycoproteins.

Sample of phosphonates have been discovered in different sites and depths in the Pacific Ocean, Atlantic Ocean and the North Sea, suggesting their importance in the marine environment. Although phosphate esters represent about 75% of dissolved organic phosphorous in marine environments, phosphonates comprise the remaining 25%[19,20,21] Overall, phosphonates are an important sub-class of phosphorus that may play a critical role in phosphorus metabolism. In addition, the ability to synthesize phosphonates is present and relative common among microbes in both marine and terrestrial environments[21].

The biological function of phosphonate containing macromolecules remains puzzling but their presence suggests that they may confer an advantage. There is evidence suggesting that the more stable C-P bond may allow cell surface macromolecules to persist in challenging host cell environments. Despite the unclear role in virulence and host survival for phosphonate containing macromolecules the potential roles for these hydrolytically stable molecules need to be further studied.

1.3 Properties and Types of Phosphonates

The defining structural motif of phosphonate natural products are the carbon-phosphorus bond. The significant feature of phosphonates is the thermal and chemical stability of the C-P bond. Although the C-P bind is weaker than the carbon - oxygen phosphorus bond , the departing alkoxide of the latter is a better leaving group than the

carbanion of the former in a hydrolysis reaction[7]. Phosphonates are relatively inert and more persistent in threatening environments. This includes exposure to hydrolytic enzymes, phosphorus limited surroundings and chemical treatments such as boiling acids and bases[22]. Some phosphonates act as potent enzyme inhibitors because they are structural mimics of phosphate esters, carboxylic acids ad tetrahedral intermediates and can compete with the latter for binding to enzyme active sites[12,22]. For instance, proteolysis is an important form of cellular regulation, in which peptide cleavage occurs at the carboxyl side of targeted amino acids in order to break proteins down in living organisms. C-P inhibitors are structurally similar and possess higher affinity to enzyme active sites than these phosphate esters and carboxylic acids, allowing them to interfere with these cellular processes[12]. Given the ubiquitous roles of phosphorylated intermediates and carboxylates in cellular biology, C-P bond containing compounds could potentially target a wide variety of cellular pathways[12].

To date, there are approximately 30 known natural C-P compounds making them an under explored area of research despite the impressive range of biological activities exhibited by these molecules[12,22]. C-P compounds are known to play a key role in metabolic and signalling pathway and possess antibacterial, antiviral, anti parasitic and herbicidal properties[11]. Additionally, enzymes found in their metabolic pathways are involved in the production of a number of critical natural products (see Figure 2). These include the approved antibiotics fosfomycin, dehydrophos, and plumbemycin; the widely used herbicides phosphinothricin tripeptide (PTT) and phosphonitricin; the clinical candidates for the treatment of malaria, FR-90098 and fosmidomycin; and antihypertensive peptides such as I5B2 (identical with K-4) and K-26[22].

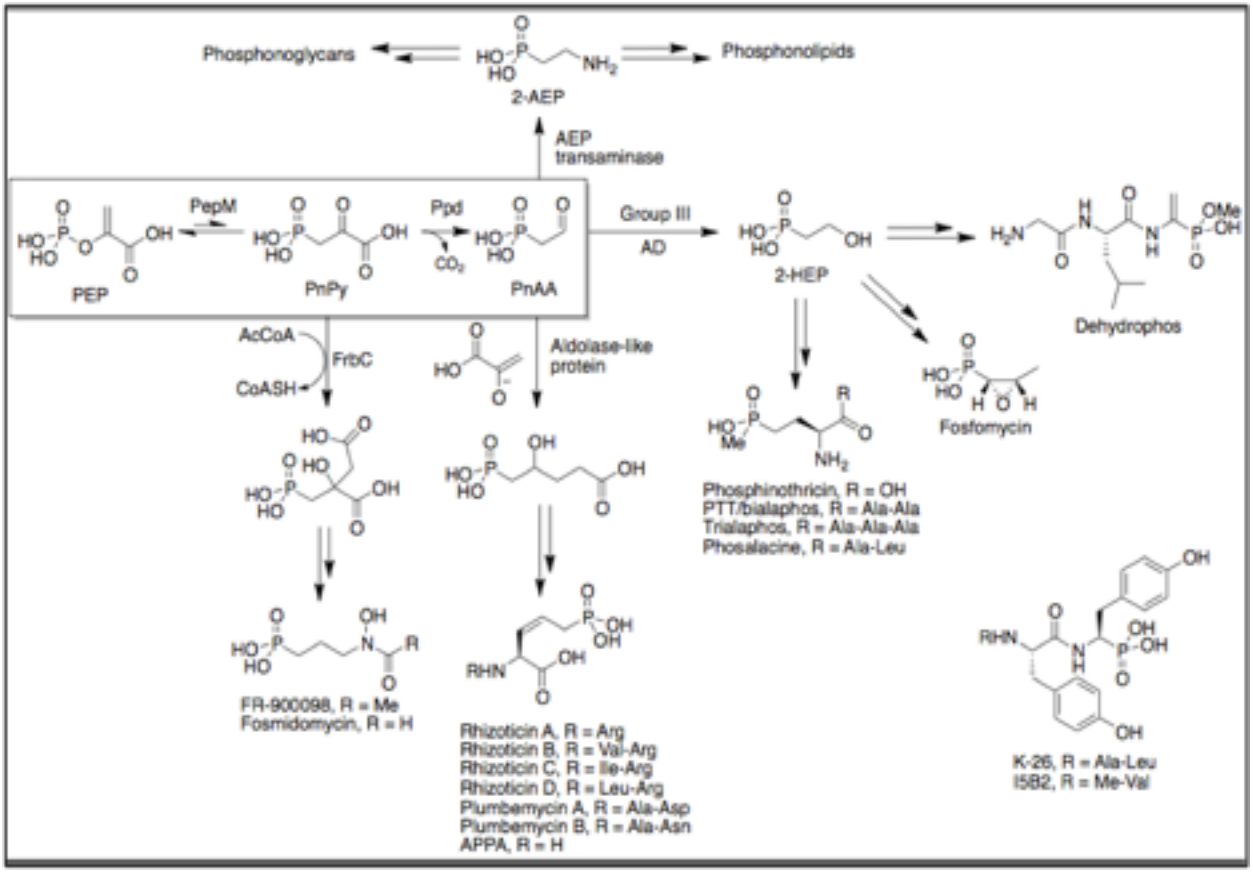


Figure 2. Overview of synthesis of natural products that possess a C-P bond and show antibacterial, antimalarial and herbicidal properties. The boxed reaction is conserved for most known phosphonates[22].

1.4 Phosphonate Biosynthesis

Despite the diversity of natural phosphonates, genetic studies have revealed that many of the early steps of these phosphonate biosynthetic reactions are conserved across different species and pathways[23]. This is additionally displayed by similarities in biosynthetic gene clusters of various bacterial strains.Initial C-P bond synthesis is proposed to occur via intramolecular rearrangement of phosphoenolpyruvate (PEP) to form phosphonopyruvate (PnPy) catalyzed by phosphoenolpyruvate mutase (PEP mutase)[12]. Surprisingly, this step is thermodynamically unfavourable, with an equilibrium

constant of about 500 in the directional of PEP. The removal of PnPy in this reaction is thus important to drive the reaction forward, and this often occurs by decarboxylation of PnPy to form phosphonoacetaldehyde (PnAA)[12]. Several reactions can take to transform PnAA, one of which includes transamination to generate 2-aminoethylphosphonic acid (AEP)[22]. In alternative biosynthetic pathway, NAD(P)H-dependent group III alcohol dehydrogenase reduces PnAA to 2-hydroxyelthyphosphonate (HEP)[22]. The full biosynthetic reaction pathway can be seen in Figure 3.

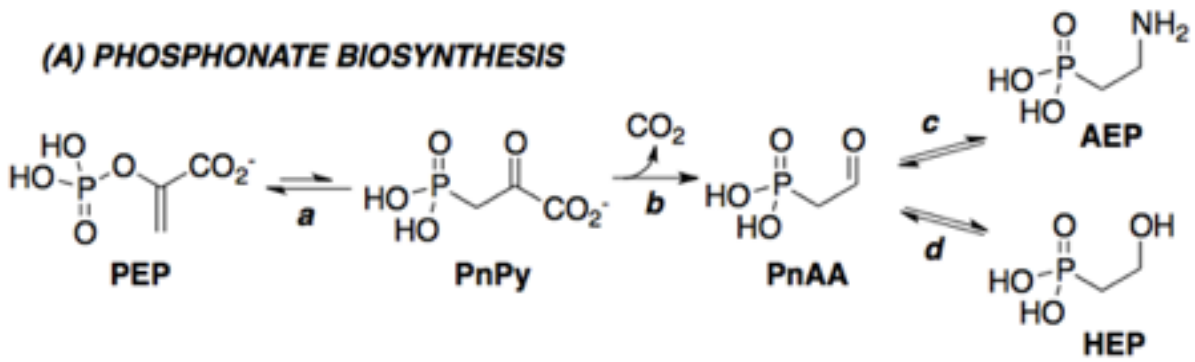


Figure 3. Proposed biosynthetic pathway for phosphonate small molecules. Labeled enzymes in the pathway are as follows: **a**= PEP mutase, **b**= PnPy decarboxylase, **c**= AEP transaminase, **d=** alcohol dehydrogenase

Phylogenetic classification of PEP mutase revealed that there are at least 10 general groupings of gene clusters capable of generating diverse phosphonates and tailored lipids and glycans . In human pathogenic and commensal bacteria, the gene clusters typically encode the AEP biosynthetic enzymes (**a, b,** and **c** or **d** in Fig. 3). Interestingly enough, gene clusters also encode genes for cytidylyltransferase and phosphotransferase/glycosyltransferase, but the purpose of those genes remains unknown.

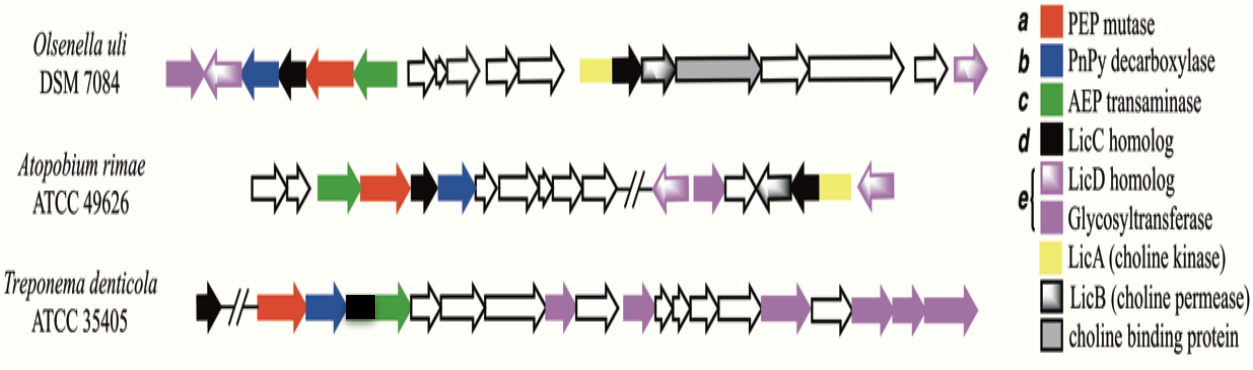


Figure 4. Phosphonate gene clusters from representative bacteria, with genes colour coded according to annotation. Lower case letters beside the colour codes in the legend correspond to the letters that label the enzymatic steps in Figure 3.

1.5 Choline mechanism in *Streptococcus pneumoniae*

Choline metabolism plays a key role in cell separation, transformation, autolysis and pathogenicity of *Streptococcus pneumoniae*[25]. Choline in the form of phosphocholine (P-Cho) is incorporated into the cell wall techie acid (TA) and lipotechoic acid (LTA) (Figure 6)[26.27].

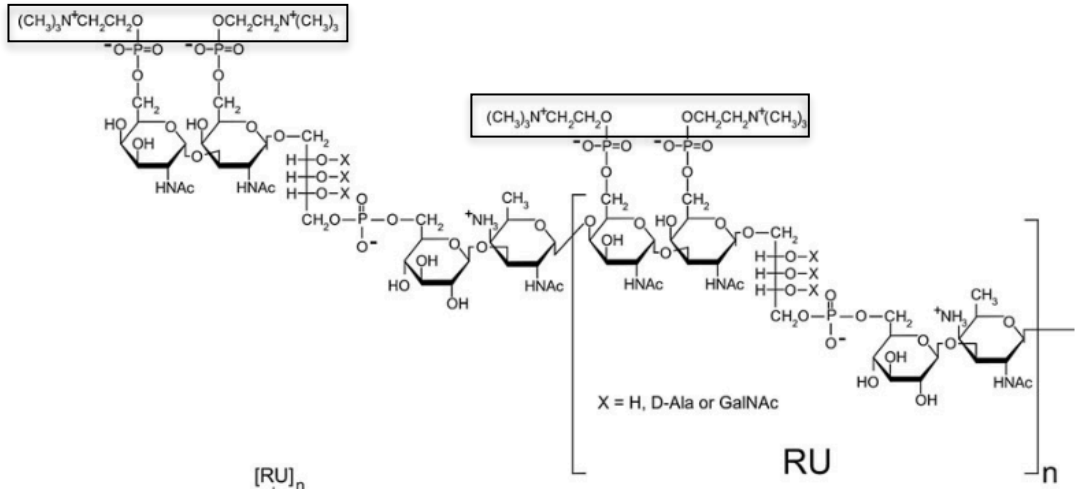


Figure 5. Teichoic acid of *Streptococcus pneumoniae* with P-Cho incorporated into it. P-Cho is highlighted in the box[26].

Cell wall phosphocholine also serve as scaffolds for a group of choline-binding proteins that are exported from cells and attached to the cell surface by their homologous choline-binding domains[27].

The importance of choline in pathogenicity is not confined to *S. pneumoniae,* but also plays a role in other bacteria such as *Hemophilus influenzae, Pseudomonas aeruginosa* and *Neisseria gonorrhoaea*[28] .The pathway for choline metabolism in *S. pneumoniae* consists of a choline transport system, a choline kinase, CTP:phosphocholine cytidylyltransferase (CCT) and a choline phosphotransferase that transfers P-Cho from CDP-Cho to either lipotechoic acid or lipopolysaccharide[28]. The existence of this pathway is supported by the detection of choline kinase and CCT activity in crude extracts of *S. pneumoniae .* Bioinformatic analyses have revealed 5 domains, *licA*, *licB*, *licC*, *licD1*, and l*icD2*[28,29]. These domains share identity with the *H. influenzae lic* loci. The percent similarity can be seen in Figure 7[28,29]. Hypothesis were drawn from bioinformatic analysis and it was concluded that *licA* corresponds to choline kinase[28]. LicB has several predicted transmembrane domains and is thus thought to be a choline transporter[28]. The hydrophilic *licC* gene product is a candidate for the CCT due to the resemblance of its amino terminus to the amino-terminal 60 residues of NTP transferase family members[28]. Lastly, *licD* gene is a candidate for the choline phosphotransferase[28]. The *licC* gene in *S. pneumoniae* and in *H. influenzae* are 37%

identical and 60% similar[28].

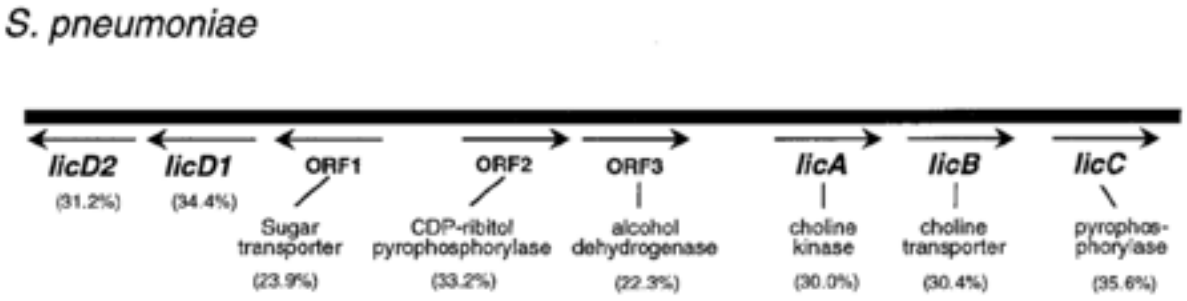


Figure 6. Gene cluster of *S. pneumoniae* and potential function of the gene products based on sequence similarity are listed[28]

It is proposed that choline is acquired from the host and imported to the cell via LicB which is a choline transporter. Once in the cell, the choline kinase LicA, phosphorylates choline to phosphocholine. Phosphocholine is then CDP conjugated via phosphocholine cytidylyltransferase. The CDP conjugated phosphocholine is then attached to a glycan by choline phosphotransferase LicD. Figure 8 shows the proposed pathway for choline tailoring. Although much is not known about this pathway, gene disruptions of either *licC* or *licD* have attenuated the virulence of *S. pneumoniae* in mouse models[29,30].

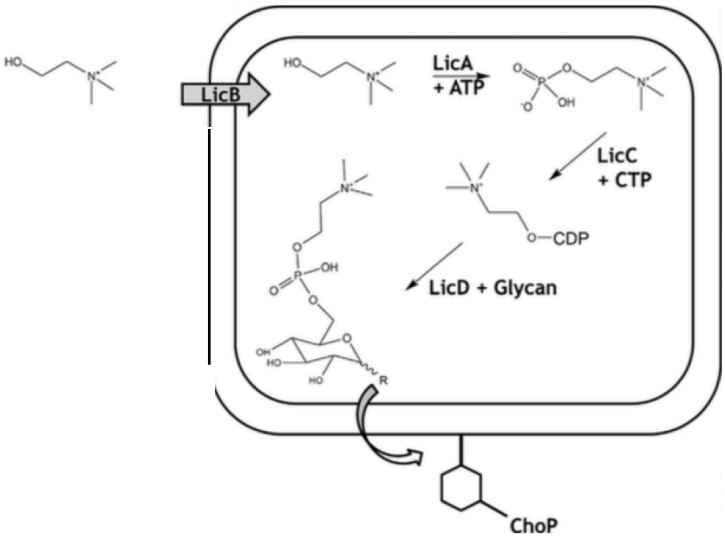


Figure 7. Proposed pathway for choline tailoring[26].

LicC from *S. pneumoniae* is a member of the nucleoside triphosphate transferase superfamily[31], and it catalyzes the transfer of a cytosine monophosphate from CTP to phosphocholine, which in turn forms CDP-choline. The enzyme possess a high degree of selectivity for phosphocholine and CTP as its natural substrates, and the related compounds phosphoethanolamine and ATP as poor substrates[31]. LicC when recently purified, was shown to catalyze cytidylyltransferase reaction. Crystal structure of LicC is a monomer, with an active site containing a deep pocket with negative electrostatic potential[31,33]. A bound magnesium balances the negative charges from the amino acid residues of Asp107, Glu135, Asp192, Glu216 and Asp218[31,33]. The amino acid residues Asp107 and Glu216 are characteristic of the NTP transferase family which further prove that this enzyme resembles the NTP transferases[31,33].

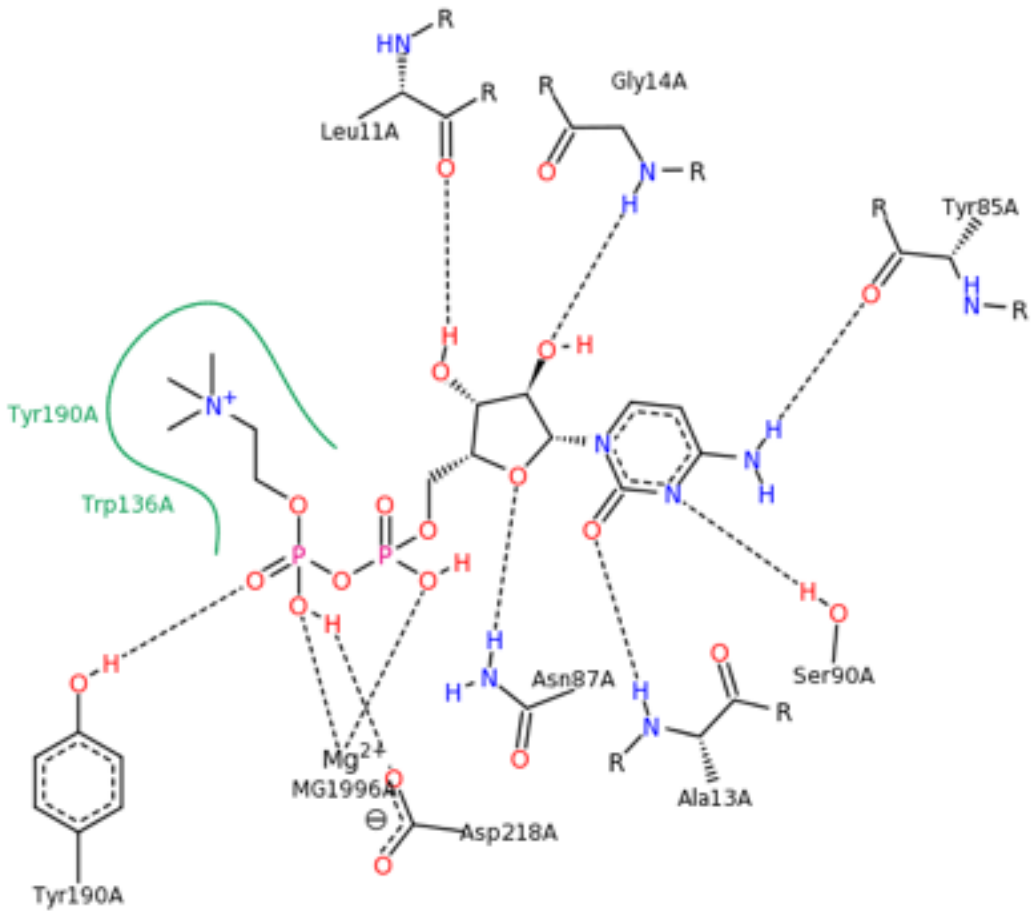


Figure 8. Active site interactions of LicC with phosphocholine, indicating many important active site residues involved in binding. Black dashed lines indicate hydrogen bonds, salt bridges, and metal interactions. Green solid line show hydrophobic interactions and green dashed lines show π-π and π-cation interactions. PDB: 1JYL[31,33]

The final tailoring step as seen in Figure 7, involves the attachment of an activated phosphocholine conjugate onto a macromolecular scaffold using phosphotransferase or glycosyltransferase[26,29]. The LicD1 and LicD2 enzymes have been hypothesized to act as a phosphotransferase in the *S. pneumoniae* tailoring pathway[26,28,29]. The loading of techoic acids with LicD enzymes appear to be a membrane -associated process, with glycolipid anchors and peptidoglycan acting to incorporate the proteins into the cell wall[26,28,29]. Although many of the teichoic acid genes remain to be identifies, the LicD1 enzymes display significant similarity to other phosphotransferases.

Glycosyltransferases are also abundant in gene clusters of *S. pneumoniae* and varies other gene clusters (Figure 4)[26,28,29]. They catalyze the transfer of sugars, using activated donor substrates that contain a substituted phosphate leaving group; acceptor substrates make up a diverse group of substrate alcohols such as proteins, lipids, nucleic acids, antibiotics, and numerous small molecules. The donor sugar substates are most commonly activated nucleoside diphosphate sugars such as UDP and GDP, but may also be nucleoside monophosphate sugars such as CMP nuclei acids, lipid phosphates and unsubstituted phosphates[26]. In the absence of LicD homologs, glycosyltransferases are seen as principal enzymes in glycan tailoring. The exact mechanism involved in LicD enzymology has not been fully characterized.

1.6 Proposed Phosphonate Tailoring Pathway

It is known that choline modifications are involved in virulence, but the role of phosphonate modifications remains unknown. The genes responsible for choline modifications have been identified, but the genes responsible for phosphonate modifications have yet to be identified and biochemically characterized. Interestingly enough, phylogenetic analyses of several opportunistic bacteria carrying genes for phosphonate biosynthesis (Figure 4), show the presence of cytidylyltransferase and glycosyltransferase/phosphotransferase, that share a homology with the genes involved in choline modifications. It is hypothesized that in addition to its role as the biosynthetic entry point for small molecule bioactive natural products such as the antibiotic fosfomycin, PEP mutase also plays a role in decorating cellular polysaccharides and lipids with phosphonate moieties. It is also the introductory enzyme leading to the creation of phosphonates (AEP) that may play a significant role in macromolecular

tailoring. As previously mentioned, the biological function of phosphonate containing macromolecules remains enigmatic, but phosphonate biosynthetic gene clusters are widespread in bacteria and suggest that they possess an advantage (Figure 4). For example, for pathogens, phosphonate modifications would render cell surface macromolecules more resistant to both acid and enzymatic hydrolysis in the lysosome[24]. For commensal obligate anaerobes like *Bacteroides* that are also opportunistic pathogens, phosphonates may be important in the highly anaerobic environment of the human distal intestinal lumen[24].From this information a biosynthetic pathway for macromolecular tailoring can be proposed. It is hypothesized that a small molecule phosphonate such as AEP is attached to macromolecules such as polysaccharides in a two step process (Figure 5). The first step is the activation of the phosphonate (AEP) as a CMP conjugate catalyzed by a cytidylyltransferase. The second step is the attachment to a macromolecule nucleophile for example a sugar hydroxyl group, catalyzed by a phosphotransferase or glycosyltransferase.

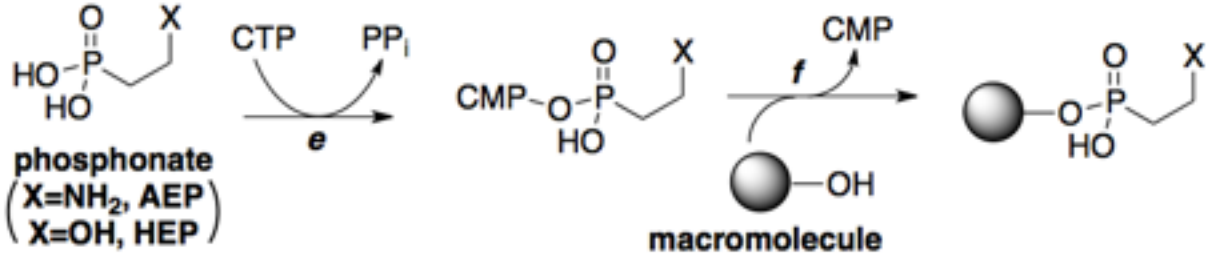


Figure 9. Proposed tailoring steps to add a phosphonate to a macromolecule with the following enzymes: **e**, cytidylyltransferase; **f**, phosphotransferase or glycosyltransferase.

Focusing on the phosphonate tailoring hypothesis, cytidylyltransferase as a homolog for LicC may lead to support of the hypothesis that small molecule phosphonate molecules such as AEP can be used to create phosphonate containing intermediates, which then

are catalyzed by LicD homologs such as phosphotransferases and glycosyltransferases in producing macromolecules important for cell metabolism. This proposed mechanism is analogous to the choline modification in *S. pneumoniae*.

1.7 Phosphonate modifications vs choline modifications

In Figure 4, three different gene clusters of opportunistic bacteria are shown, bacteria such as *A. rimae* and *O. uli* seem to possess both phosphonate biosynthesis genes, and teichoic biosynthesis genes. There are many hypotheses as to why phosphonate containing macromolecules may predominate instead of the phosphocholine containing macromolecules. One hypothesis is that choline has to be acquired from the host, and brought into the cell to be able to be attached to a macromolecule. In the event where choline is scarce, the host may utilize PEP, already present in the cell, convert it to a phosphonate, and attach it to a macromolecule. Since phosphonates are more resistant to degradation and are similar in size and charge to phosphocholine, a bacterial host may be able to keep its pathogenicity by swapping choline and utilizing phosphonate moieties. Another hypothesis may be that since choline binding proteins are recognized by the immune system, phosphonate containing macromolecules can be used to evade the immune system. However, as stated earlier the role of phosphonate tailoring macromolecules in virulence is unknown.

1.7 Hypothesis

It is known that phosphocholine is the substrate that is catalyzed by LicC to be added onto the teichoic acid, but what is unknown, is if LicC homologs from phosphonate containing bacterial strains use AEP to catalyze a similar reaction. A structure for LicC from *S. pneumoniae* has been crystallized and the active site residues are established,

however, the similarity of those active site residues to other cytidylyltransferases has not been identified. Based on the information in the literature, few hypotheses can be made. The first hypothesis is that the active site residues for all cytidylyltransferase are going to be different, to accommodate the substrate that they catalyze. The second hypothesis is that LicC homologs in phosphonate pathways will have greater specificity for its own substrate. LicC homologs from *A. rimae* and *O. uli* will have greater specificity for phosphonate substrate AEP, and LicC from *S. pneumoniae* will have greater specificity for phosphocholine. This being said, it is hypothesized that due to the nature of LicC genes, they will also have some specificity for substates that may resemble their native substrates.

1.8 Objectives:

The aims of this project are as follows:

1. Bioinformatic analysis of LicC from *S. pneumoniae* (Spn-LicC) and PngC from *Atopobium rimae* and *Olsenella uli* (Ari-PngC and Oul-PngC)
2. Functional Characterization of Spn-LicC and Ari-PngC.

2. Materials and Methods

2.1 Materials Used

2.1.1 Cloning, Expression and Purification

The *Atopobium rimae* gDNA was purchased from Cedarlane Labs (Burlington, ON). *Olsenella uli* gDNA was purchased from Biomatik (Cambridge, ON). Synthetic genes for *O.* and *A. rimae* were purchased from Bio Basic (Markham, ON). Primers were purchased from Integrated DNA Technologies (Coralville, IA). PCR buffers and polymerases were purchased from Bio Basic (Markham, ON). A C1000 Touch Thermal Cycler was purchased from BioRad (Mississauga, ON). Wizard SV Gel and PCR Clean -Up System were purchased was purchased from Promega (Madison, WI). Agarose gel electrophoresis was performed using a gel apparatus and power source from VWR (Mississauga, ON). A Gel Doc EZ imager was used for gel imaging and purchased from BioRad (Mississauga, ON). The PureYield Plasmid Miniprep system was purchased from Promega (Madison, WI). DNA sequencingg was completed at the Hospital for Sick Children at the Centre for Applied Genomics (Toronto, ON). The Gem T-Easy Vector System was purchased from Promega (Madison, WI). *Escherichia coli* DH5a and BL21 were purchased from Novagen (Madison, WI). The pET29 vector used for cloning was purchased from Novagen (Madison, WI). NdeI and XhoI restriction enzymes were purchased from ThermoFisher Scientific (Waltham, MA). The Heroes Multifuge X1R centrifuge was purchased from ThermoFisher Scientific (Waltham, MA). High speed centrifuge tubes and low speed 500 mL tubes were purchased from ThermoFisher Scientific (Waltham, MA). The Q-Sonica Q125 sonic processor was used for sonications (Newton, CT). Ultracentrifugal devices were purchased from Pall Corporations (Port

Washington, NY). PD-10 columns were purchased from BioRad (Mississauga, ON). SDS-PAGE gels were cast using TGX Stain Free FastCast Acrylamide Kit 12% purchased from Bio Rad (Mississauga, ON).

2.1.2 MBP-Cloning and Anion Exchange Purification

pMAL vector was purchased from New England Biolabs (Ipswich, MA). XmnI and NotI restriction enzyme was purchased from ThermoFischer Scientific (Waltham, MA). Primers were purchased from Integrated DNA Technologies (Coralville, IA). The FPLC machine was purchased and Enrich Q Resin column were purchased from Bio-Rad (Mississauga, ON). The DynaLoop-90 was used to inject sample and was purchased from Bio-Rad. Tris and NaCl were purchased from Amresco (Cleveland, OH).

2.1.3 Crystal Tray Set-Up

Crystal Gryphon LCP system and the INTELLI-PLATE 96-2 plate from Art Robins Instruments (Sunnyvale, CA) was used to set up crystal tray trials. Plastic cover slip was purchased from VWR (Mississauga, ON). Crystal screen 1&2 and Index were purchased from Hampton Research (Aliso Viejo, CA). Crystal screens MCSG1 and MCSG2 were purchased from Anatrace (Maumee, OH). Crystal trays were incubated at 18 C in a crystallization incubator from Molecular Dimensions (Newmarket, UK). Crystal trays were examined under an Olympus SZX16 light microscope (Tokyo, Japan).

2.1.4 Pyrophosphate Activity Assay

The EnzCheck Pyrophosphate Assay Kit (E-6645) was purchased from ThermoFischer Scientific (Waltham, MA). xMark Microplate Absorbance Spectrophotometer was purchased from Bio-Rad (Mississauga, ON) and was used to analyze the results. CTP was purchased from ThermoScientific (Waltham, MA) and MgCl2 was purchased from

Amresco (Cleveland, OH). Chelax resin was purchased from Sigma Aldrich (St.Louis, MO).

2.2 Methods

2.2.1 Cloning, Expression and Purification

Polymerase Chain Reaction (PCR): Forward and reverse primers were made for *pngC* and *pngD* genes from *A.rimae* and *O.uli.* Primer sequences can be found in Appendix . All of the primers included a 6X His-tag. A 100 µL reaction mixture was then made for each of the species in order to amplify the gene. The reaction mixture consisted: of DMSO, DI water, 10 mM dNTP mixture, 2x Taq buffer, Taq polymerase, forward primer, reverse primer, and the species DNA. The reaction mixture was then aliquoted into eight separate 10 µL fractions in an 8- strip PCR tube strip. The eight wells were then put into a thermal cycler, set to the conditions that can be seen in Appendix. Each of the wells in the PCR tube strip were set to be exposed to a different annealing temperature within the thermal cycle in order to determine which temperature in a specific gradient allowed for the best amplification of the gene. In cases where a temperature gradient was not used and an optimal annealing temperature was already known, the Taq buffer was replaced with a 5x Phusion buffer and Phusion polymerase to replace the Taq polymerase. These results were checked using gel electrophoresis, as described in the next protocol.

Agarose Gel and Gel Extraction: A 1% agarose gel was made using 0.40 g of lab grade agarose, 40 mL of 1X TAE buffer, and 2 µL of 10 mg/mL ethidium bromide. The preparation of 1X TAE buffer is listed in Appendix . After solidification of the prepared

gel, the 10 µL fractions of the PCR reaction mixtures were mixed with about 2 µL of bromophenol blue dye (unless the they were made with the Phusion buffer, which was already dyed green). Each was loaded into a separate well in the gel, along with a final well of about 5 µL of 1Kb plus ladder. Gels were run at 100 V and 200 mA for about 45 minutes. Images were analyzed using ImageLab software and the Bio-Rad "Gel Doc EZ Imager".

Excision of PCR amplifications: Agarose gels were examined under a LabNet UV trans-illuminator and desired bands were cut out using a blade. From these excised gel slices, the DNA was amplified and extracted using the Promega Wizard SV Gel and PCR Clean-Up System.

Clean-up and addition of an A-tail: Phusion polymerase has exonuclease activity that cuts out the overhang of the DNA amplifications. An A-tail was added to the extractions in order to create a new overhang that would be required for complementarity of residues when performing a ligation. This was done by mixing: 20 µL DNA, 1 µL Taq polymerase, 1 µL 10mM dNTP mxture, 2.5 µL 10X Taq buffer, and 1 µL 50mM $MgCl_2$. This reaction mixture was then cleaned up using the Promega Wizard SV Gel and PCR Clean-Up protocol. The cleaned reaction mixture was then run on a 1% agarose gel to ensure that the presence of the appropriately sized DNA bands.

Cloning into pGEM: The pGEM-T Easy vector was used to clone the amplifications to subsequently be used for sequencing. The cleaned gel extract of DNA was ligated into

the pGEM vector using the following reaction mixture: 3 µL DNA, 1 µL pGEM-T Easy vector, 5 µL 2X Rapid Ligation buffer, and 1 µL T4 DNA ligase. Positive controls were created by replacing the DNA with 2 µL Promega control DNA, and negative controls were created by replacing the DNA with 3 µL sterile water.

Transformation of recombinant plasmids into *E. coli* DH5α: *E.coli* DH5α competent cells were prepared as explained in A5 of Appendix A. Aliquots of 150 µL were thawed on ice, and then either the ligation reaction mixture or about 2-4 µL of mini-prepped DNA directly were added to the thawed competent cells. The negative and positive controls were transformed in the same manner as the ligation. The microcentrifuge tubes were then left on ice for 20 minutes and then heat shocked for 2 minutes at 42°C. 1 mL of LB broth was then added to the mixtures, after which they were sealed tightly and placed in a shaker at 37°C for one hour. The transformation mixtures in the tubes were then spun down in a microcentrifuge for one minute, and all but 100 µL of the supernatant was poured off. The pellet formed was resuspended in the remaining supernatant and plated directly on XCI LB-agar plates, containing 100 µg/mL carbenicillin, 0.5 mM IPTG, and 80 µg/mL X-Gal. Plates were then left to incubate overnight (approximately 12 hours) at 37°C. Refer to Appendix for preparation of LB broth and LB-agar plates.

Preparing plasmids for sequencing: A single colony from the transformation plates were picked using a sterile pipette tip and grown in a 3 mL culture containing 6 µL of ampicillin. The cultures were left to grow overnight. The PureYield Plasmid Miniprep System was used to isolate the recombinant pGEM plasmid that was then digested with

NdeI and *XhoI* using the following mixture: 2 µL DNA, 2 µL10X FastDigest green buffer, 1 µL *NdeI*, 1 µL *XhoI*, and 14 µL nuclease-free water. This mixture was incubated at 37 °C for about one hour. To confirm the presence of inserts of the correct size, the mixtures were run on a 1% agarose gel and imaged, as described previously. The gel extraction protocol previously described was used to excise bands of the correct size from the digest. Additionally, if a plasmid of the correct size was identified, about 300 ng of the recombinant pGEM plasmid was added to sterile water to make a solution totalling 7 µL.

Cloning into pET29: Excised gel bands from the digests were ligated with pET29. Preparation of the pET29 construct is referred to in Appendix . The reaction mixture contained: about 3 µL pET29, about 5 µL DNA insert, 1 µL 10X Ligase buffer, 1 µL T4 DNA ligase. The exact amounts of pET29 and DNA insert added were subject to change depending on their relative concentrations; the aim was to add about 3X the concentration of DNA compared to pET29. A negative control was made by replacing the DNA with sterile water. The ligation was refrigerated overnight, or left at room temperature for 1 hour. The ligation reaction mixture was then transformed into *E.coli* DH5α competent cells as described above. However, kanamycin LB-agar plates were used with 50 µg/mL kanamycin. Cultures were grown from the colony as described before and then mini-prepped to prepare the recombinant pET29 plasmids for another transformation into *E.coli* BL21. These cultures were grown with 5 µL of kanamycin per mL of LB broth used.

Inoculation and induction of pET29 plasmids: One 50 mL culture was then grown from a single colony with 250 µL of kanamycin and grown to an OD_{600} of about 0.5. This culture was then aliquoted into 4 separate 1L flasks, each with about 10 mL of the culture and 5 mL of 10 mg/mL kanamycin. The flasks were then left to shake at 37 °C until an OD_{600} of about 0.2 was reached. At this time, the shaker temperature was turned down to about 18 °C and the flasks were left to shake until an OD_{600} of about 0.5 was reached. At this time, 1 mL of the culture was removed and then spun down in a microcentrifuge for about one minute to collect an uninduced pellet. The culture was then induced with 1 mL of 0.1 M IPTG per 1 L of culture. The culture was then left to shake at 18 °C overnight, or approximately 12 hours. Before purification, about 500 µL of culture was spun down in a microcentrifuge for about one minute to collect an induced cells pellet.

Protein purification: The total of 4 L of culture was spun down in a centrifuge at 5000 rpm for 10 min at 4 °C. The supernatant was then discarded and the resulting pellets were resuspended in cell lysis buffer (50 mM $NaPO_4$, 0.3 M NaCl, pH 7.5), using an amount of buffer that was approximately 5% of the original culture volume. Preparation of the buffers can be seen in Appendix B. About 0.2 to 2 mg of both lysozyme and DNaseI was added to the solution and then kept on ice 20-30 minutes. The solution was then sonicated on ice, programmed to six 10 second intervals of sonicating with 50 seconds intervals of cooling periods. The mixture was then centrifuged at 14,000 rpm for 20-30 minutes at 4 °C. 100 µL of the supernatant was then removed and frozen at -20 °C to be analyzed later as the “raw extract” sample. The pellet was also poked, and a small portion of it was saved and labeled “cell debris”. The supernatant was then

collected in one conical tube, and about 1 mL of Ni-NTA slurry was added. The two were then mixed and shaken on ice for 10-20 minutes. The solution was then loaded on an Econo-Pac chromatography column and the flowthrough was collected and set aside. The column was then washed twice with 5 mL of wash buffer (lysis buffer with 20 mM imidazole). Each wash was collected and set aside. An elution buffer (lysis buffer with 250 mM imidazole) was then added to the column, and the elution was collected in an ultrafiltration tubes (MWCO 30 kDa). The elution was concentrated by centrifugation at 5000 rpm for 40 minutes at 4 °C. In the meantime, the "gravity protocol" was started by desalting one PD-10 column. This was done by washing it with equilibration buffer (50 mM $NaPO_4$ buffer at pH 8). The solution present above the filtration membrane was added to the PD10 column and allowed to enter the packed bed. The amount of solution added to the column was topped up to 3 mL using equilibration buffer. The flow through was discarded. Afterward, 4 mL equilibration buffer was added to the PD10 column and the eluent was collected in another ultrafiltration tube. The sample was centrifuged at 4000 rpm for 40 minutes at 4 °C. The sample present above membrane was then flash-frozen into beads using liquid nitrogen. The beads were then left in a -80 °C freezer in a cryotube.

Purification modifications for LicC genes: Phosphate free buffers such as Tris was used for LicC purifications in order to be compatible for the kinetic assay. In addition to changing to phosphate free buffers, 2 mM DTT and 20% Ammonium sulfate were added to the buffers. The buffers and their preparation can be seen in Appendix **.** The purification protocol was the same as described above.

Bradford assay to determine protein concentrations: A Coomassie Blue Bradford reagent was prepared, as seen in Appendix . Six vials were prepared as controls for the Bradford standard curve. The vials' contents can be seen in Appendix**.** Four vials were then prepared for the protein samples. One vial was contained 5 µL flowthrough, one with 5 µL raw extract, and two for the final eluted sample (one containing 5 µL of sample and another with 20 µL of sample). The vials were then topped up with sterile water until a total of 100 µL was reached. 5 mL of Bradford reagent was then added to each of the vials and they were left to react for about 5 minutes. About 1 mL of the contents from each vial was added to separate cuvettes, and their absorbance readings were measured using a spectrophotometer. The concentrations were then calculated by comparing the results of the standard curve to the absorbance readings of the samples. This was done using a template in Excel.

Analysis of protein expression: Protein samples were run on an SDS-PAGE gel to check for protein expression and appropriately sized protein bands. Samples examined: uninduced cells, induced cells, raw extract, inclusion bodies, flow through, wash #1, wash #2, and the final protein sample. The samples were prepared by adding about 6 µg of sample to microcentrifuge tubes. The amount of solution to add for each sample was determined using the Bradford assay sample concentrations. The solutions were topped up to 16 µL using sterile water, and then 4 µL of 5X laemmli blue buffer was added to each of the samples. For the cell debris, induced and uninduced cells a pipette was used to simply touch the pellets. The pipettes were then dipped into 16 µL of sterile water before adding the 5X buffer. All of the samples were then placed in boiling water

for about 5 minutes. After cooling, the full 20 µL of each sample were inserted into the wells of an SDS-PAGE gel that was prepared using the BioRad TGX Stain-Free FastCast Acrylamide Kit, and soaked in 1X SDS-PAGE buffer. An SDS marker was inserted into the wells at each end. The SDS-PAGE was run at 100V for about 30 minutes, and then 200V for about 45 minutes, or until all of the blue dye had run off the gels. The gel was then imaged using the Gel Doc EZ Imager and "Stain-Free" protocol in ImageLab. Appendix displays the preparation of the SDS- PAGE buffer and molecular weight markers.

2.2.2 MBP Cloning and Anion Exchange Purification

Cloning a MBP tag: *O.uli pngC* and *pngD*were cloned into pMAL vector to include a maltose binding protein (MBP) tag to increase solubility. MBP primers can be seen in Appendix. The DNA was amplified, ligated into pGEM T-easy vector and sent out for sequencing using the protocols described above. Once sequencing confirmed the amplification of the correct insert, the insert was ligated into pMAL vector using XmnI and NotI restriction sites. Since the constructs do not have a His-tag, anion exchange chromatography was used instead.

Anion Exchange Purification:

Buffers to be used for anion exchange were prepared as described in Appendix**.** To prepare the FPLC for purification, all of the lines were transferred from an ethanol solution into their corresponding buffers. The ethanol was rinsed out of the lines beforehand, and then the new solutions or buffers were pulled through the new lines

until there was no air present in the lines. Dynaloop-90 was connected to the system, and rinse with water, and Buffer A. The anion exchange column, was washed with Buffer A, Buffer B and equilibrated with Buffer A. The ChromLab software on the computer was connected to the FPLC, and the template for anion exchange was chosen. Proteins were cloned into pET29 in the manner described above, and purified up until the Raw Extract. Raw Extract was injected into the DynaLoop-90 and was run over the anion exchange column. Fractions of 2 mL samples were eluted and collected. Eluted fractions were collected, spun down in a centrifuge to concentrate. A Bradford assay and SDS-PAGE were performed to determine the concentration, and purity of the proteins.

2.2.3 Crystal Tray Set Up

Preparation of Protein Sample:

Protein samples need to be very pure and concentrated in order to set up crystal trays. Protein samples were run over IMAC chromatography followed by anion exchange chromatography to achieve the purest form of sample. Protein samples to be used were prepared in the same manner as those used to create frozen protein beads. However, the protocol is modified starting at the PD-10 column washes. Instead of washing the column with equilibration buffer the PD-10 column was washed with 20 mL of Buffer A. Buffer A was also used to elute the sample from the PD-10 column. The sample was concentrated, and injected into the FPLC for anion exchange. Once anion exchange was completed, the fractions were pooled, and concentrated down to a total of atleast 10 mg. A minimum volume of 100 uL of sample was required.

Preliminary crystal tray set-up:

A 96 well tray was used to set up crystal trays. This was done using the Crystal Gryphon LPC system. Four different “store bought” crystal trays were used; MCSG 1&2, Index, and crystal screen 1+2 were used. Once the crystal trays were set up, plastic cover slips were then placed on top of the corresponding wells to seal the wells. The sealed wells were left to equilibrate for about 3 days in the 18 C incubator, and then analyzed under a light microscope. The tray was analyzed every three to four days for results.

2.2.4. Pyrophosphate Activity Assay

EnzCheck Pyrophosphate Assay Kit was used to determine the activity of LicC and PngC . This assay is a coupled assay, where the first part of the reaction uses cytidylyltransferase enzyme, and the second part of the reaction uses pyrophosphatase. The reaction is measured at 360 nm using UV-Vis Spectrophotometry. The first reaction that uses LicC releases pyrophosphate, which then is made into two inorganic phosphates by inorganic pyrophosphatase. The inorganic phosphates react with 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) using purine nucleoside phosphorylase (PNP) to get ribose 1-phosphate and 2-amino-6-mercapto-7-methyl-purine, which is then detected at 360 nm (Figure 10).

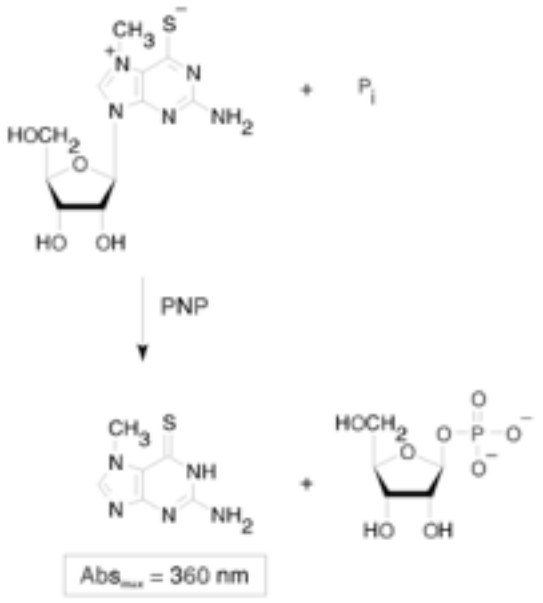


Figure 10. Conversion of 2-amino-6-mercapto-7-methyl-puring ribonucleoside (MESG) to ribose 1- phosphate and 2-amino-6-mercapto-7-methylpurine using ribonucleoside phosphorylase (PNP) where the formation of 2-amino-6-mercapto-7-methylpurine is measured at 360nm.

Each reaction was kept to a volume of 100 uL and performed in triplicate. Protein samples were prepared in the same manner as described above, containing very pure samples as frozen beads. The reaction components were split between 2 wells, one well contained the substrate and enzyme, while the other well contained CTP. Well 1 contained 0.5 uL (1U) PNP, 0.5 uL (0.03U) IPP, 10 uL (0.2 mM) MESG, 2.5 uL 20X reaction buffer, 16 uL (4 mM)CTP, and 20.5 uL of water, making a total of 50 uL. Well 2 contained 0.5 uL (1U) PNP, 0.5 uL (0.03U) IPP, 10 uL (0.2 mM) MESG, 2.5 uL 20X reaction buffer, 12 uL (7 mM) MgCl2, enzyme (depending on concentration of choice), substrate (depending on concentration of choice) and H2O to fill to a final volume of 50 uL. A control with either no enzyme or substrate was made for each trial. To avoid pipetting errors, parent mixes were created for Well 1 and Well 2. If for example, 3 different substrate concentrations were being tested in triplicate, a parent mix would be

made with 9 times the amount of components in each well in a micro centrifuge tube and then be aliquoted out into each well. For example, instead of 0.5 uL of PNP, you now would add 4.5 uL of PNP into the micro centrifuge tube. In the end, each well would get 50 uL of parent mix aliquoted out. Once the parents mix is aliquoted out, the separate wells were incubated for 30 mins at room temperature. After 30 mins, well 2 was combined to well 1, and absorbance was recorded at 360 nm for 30 minutes.The kinetic data was analyzed by excel and plotted using the software R. R uses non linear regressions, as well as 95% confidence intervals to determine Michaelis-Menten and substrate inhibition curves.

3. Results

3.1 Bioinformatic Analysis

LicC from S.pneumoniae has been crystallized previously and based on the crystal structure important residues in the active site were determined. Based on the amino acid sequences of Spn-LicC Ari-PngC and Oul-PngC, structures were created in Phyre2. Phyre2 is a protein fold recognition sever which uses amino acid sequences to predict and analyze protein structure. The structures created by Phyre2 were analyzed to see if the active sites of all the cytidylyltransferaseswere similar. The active site of Spn-LicCwas overlaid with the predicted structure of Ari-PngC.

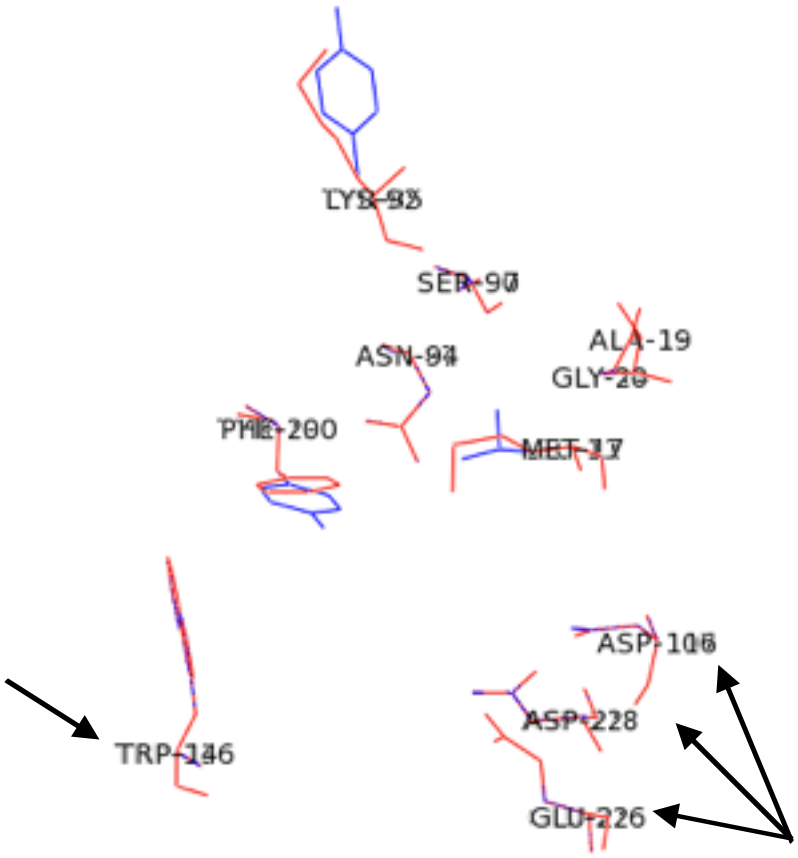


Figure 11. Active sites of Spn-LicC in blue and Ari-PngC in red overlaid. Arrows pointing toward important residues that are conserved in both structures.

From the figure above it can be seen that the active site residues are conserved between both LicC genes. The important residues in Spn-LicC such as Glu 216, and Asp 107 which are characteristic of this class of enzymes are conserved. As well as additional residues such as Trp 136 and Asp 218 are also conserved in Ari-PngC. Since the active sites of both of these enzymes resembles each other, it can be concluded that both of these enzymes behave in a similar fashion when catalyzing a reaction.

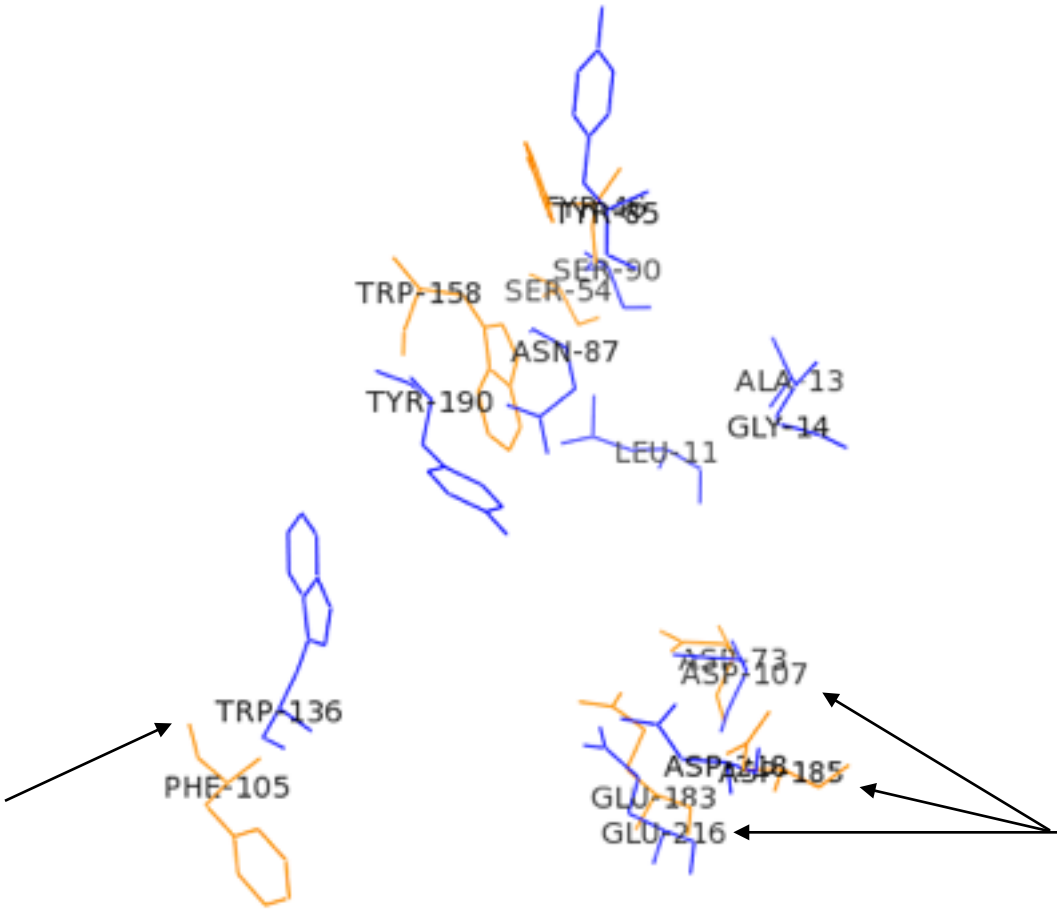


Figure 12. Active sites of Spn-LicC in blue and Oul-PngC in orange overlaid. Arrows pointing toward important residues that are conserved in both structures.

The important residues Asp 107, Glu 216 and Asp 218 are also conserved in the LicC from *O. uli*. Although Trp 136 is changed to Phe 105 in Oul-PngC*,* the residues are similar in structure and may be involved in the same manner. It can also be concluded that Oul-PngC behaves in a similar way as Spn-LicCwhen catalyzing a reaction. Although bioinformatic analysis reveals important residues, crystal structures of both enzymes will need to be obtained and compared to the crystal structure of LicC from *S.pneumoniae.*

3.2 Cloning, Protein Expression and Purification

The genomic DNA for *A.rimae* and *O.uli* were purchased and stored according to manufacturers suggestions. Primers were created (Appendix) for the *pngC* gene for the two strains, and PCR was performed according to the conditions in Appendix. The results of the PCR were visualized on a 1% agarose gel.

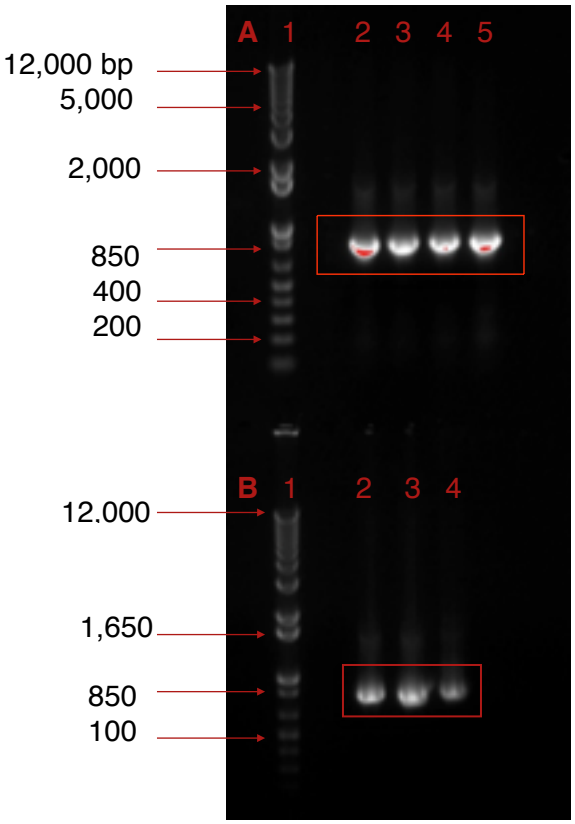


Figure 13. Taq polymerase PCR of Ari-*pngC* with an annealing temperature ranging from 50-65 C. Lanes A1 and Lanes B1 represent the molecular weight ladder. Lanes A2-5 and B2-4 show amplicons at the expected size of 850 base pairs (bp).

The amplification was successful with bands showing up around 850 bp which is the correct size, but it can be seen that the optimal temperature for annealing is at lower temperatures. Since Taq polymerase is more prone to making mistakes and mutations, Phusion PCR at an annealing temperature of 56 C was performed.

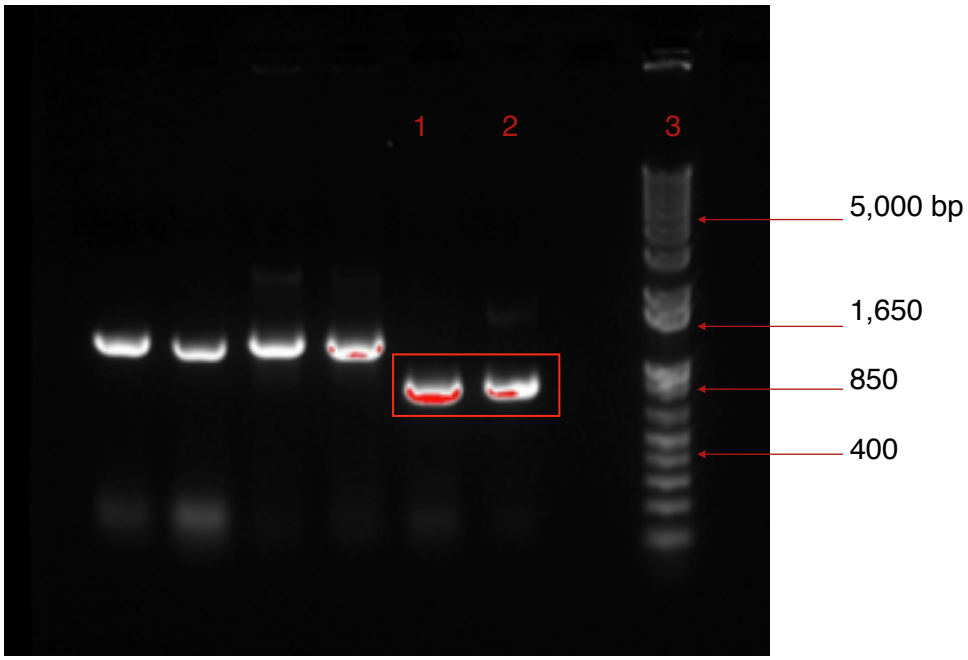


Figure 14. Agarose gel of Phusion PCR of Ari-*pngC* at 56 C. Lanes 1-2 show the amplicons at 850 bp's and lane 3 is the molecular weight ladder. Unlabelled lanes are irrelevant to this research.

Phusion PCR gave great success in amplifying *pngC*. The amplified bands were at the correct size of around 800 bp's which is the expected size of this homolog. Next, the bands in the above figure were extracted, and cleaned using the Promega PCR Clean - Up System. The cleaned up amplicons were run on an agarose gel to confirm the size of the amplicons.

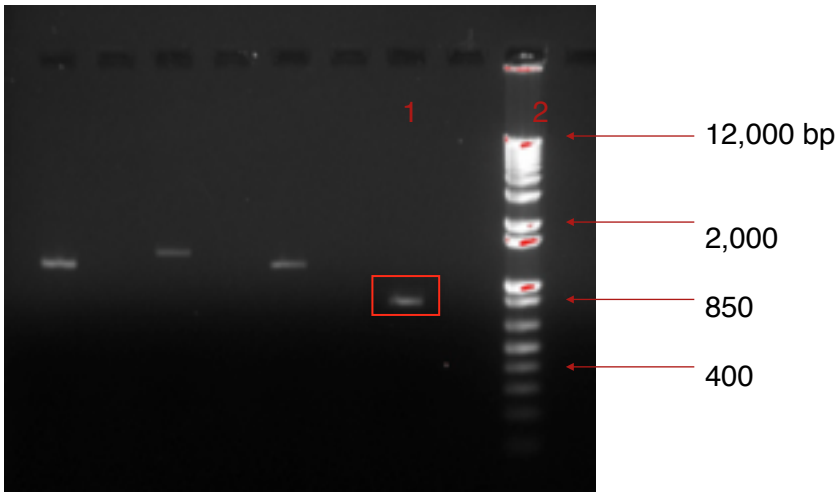


Figure 15. Clean up of Ari-*pngC*after PCR, run on 1% agarose gel. Lane 1 shows the amplicon at around 850 bp, and lane 2 shows the molecular weight ladder. Unlabelled lanes are irrelevant to this research.

The cleaned up amplicons were indeed the correct size. The next steps were to add an Adenine tail to the amplicons, since Phusion polymerase has exonuclease activity that cuts out the overhang of the DNA amplifications. An adenine tail is required to create a new overhang to perform successful ligation into different pGEM vector. Adenine tail was added using the protocol described in materials and methods, and run on a 1% agarose gel.

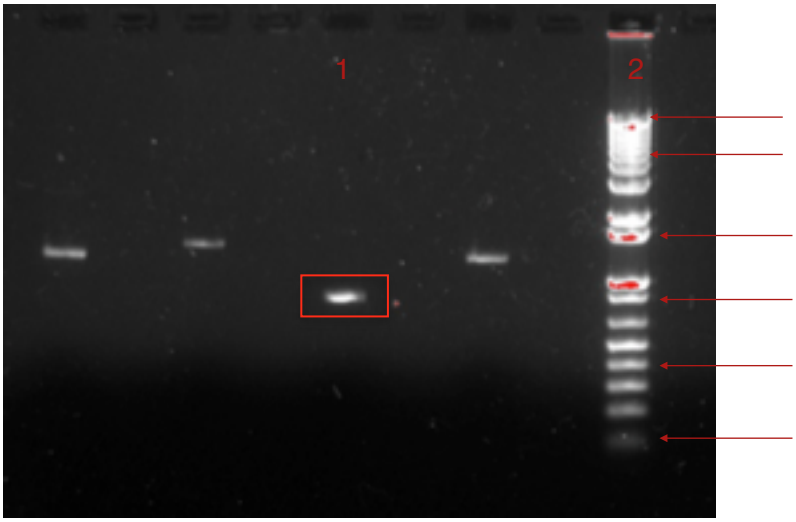


Figure 16. Clean up of Ari-*pngC* after adenine tail addition. Lane 1 shows the product of the clean at around 850 bp, and lane 2 shows the molecular weight ladder. Unlabeled lanes are irrelevant to this research.

The adenine overhang addition was a success, and the band was approximately around 800 bp which is the correct size. The amplicon was ligated into pGEM, and transformed into DH5a competent cells, plated on ICX plate and placed in the incubator at 37 C overnight. A few white colonies were picked, grown and mini-prepped to isolate the recombinant pGEM plasmid. To verify that the ligation into pGEM was successful, a digest was done using EcoRI restriction enzyme.

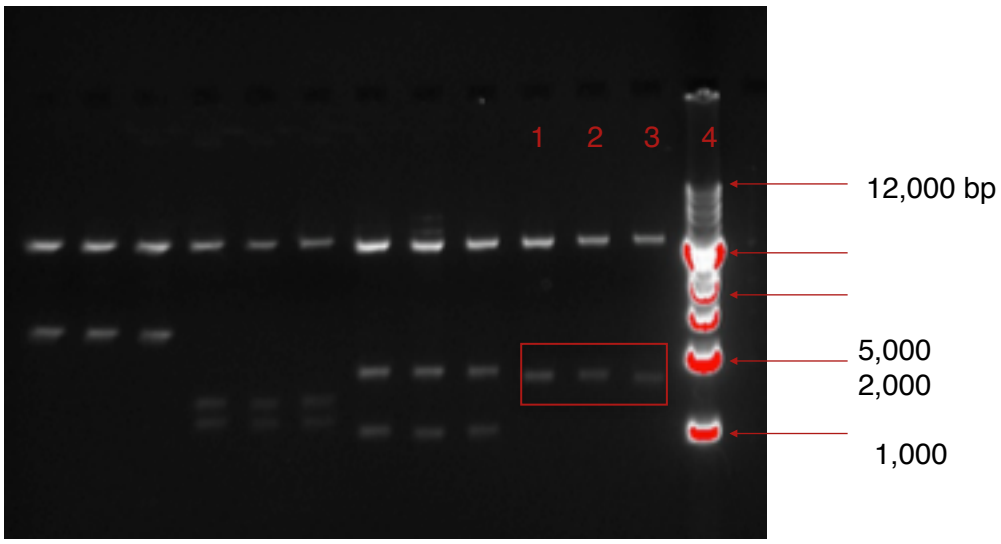


Figure 17. pGEM digest of Ari-*pngC* using EcoRI restriction enzyme. Lanes 1-3 show the digested product at 850 bp and lane 4 shows the molecular weight ladder. Unlabelled lanes are irrelevant to this research.

The pGEM digest was a success as seen in Figure 17, the bands are around 800 bp which is the estimated size of the Ari-*pngC*. The gene is now ready to be ligated into pET29. The Ari-*pngC* gene was cut out of pGEM using NdeI and XhoI restriction enzymes, and ligated into pET29 vector. The ligations were transformed into Dh5a competent cells, and incubated at 37 C overnight. Colonies were picked from the Dh5a plates, grown overnight in LB, and mini-prepped according to the protocol described above. To determine whether Ari-*pngC* was successfully ligated into pET29, a restriction digest was completed using NdeI and XhoI.

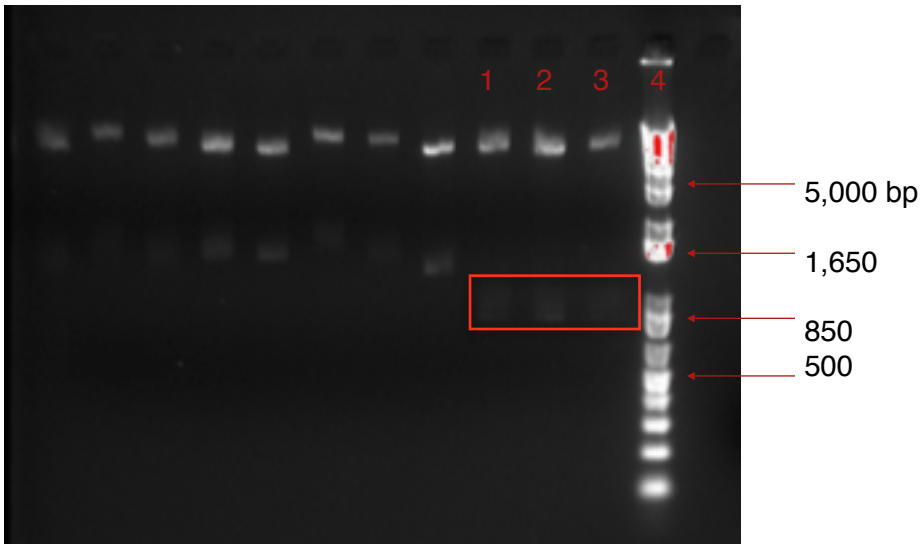


Figure 18. Digests of Ari-*pngC*-pET29 constructs using restriction enzymes NdeI and XhoI. Lanes 1-3 show digested constructs at 850 bp, and lane 4 shows molecular weight ladder. Unlabelled lanes are irrelevant to this research.

Even though the bands are very faint in Figure 18, it is apparent that the digest was successful, and that Ari-*pngC* gene was successfully ligated into pET29. The next step was to express and purify Ari-PngC protein. Ari-*pngC*-pET29 construct was transformed into BL21 competent cells, and expressed using IPTG. The construct was purified using immobilized metal affinity chromatography, and run on a SDS-PAGE gel to see if the correct protein was expressed, and purified.

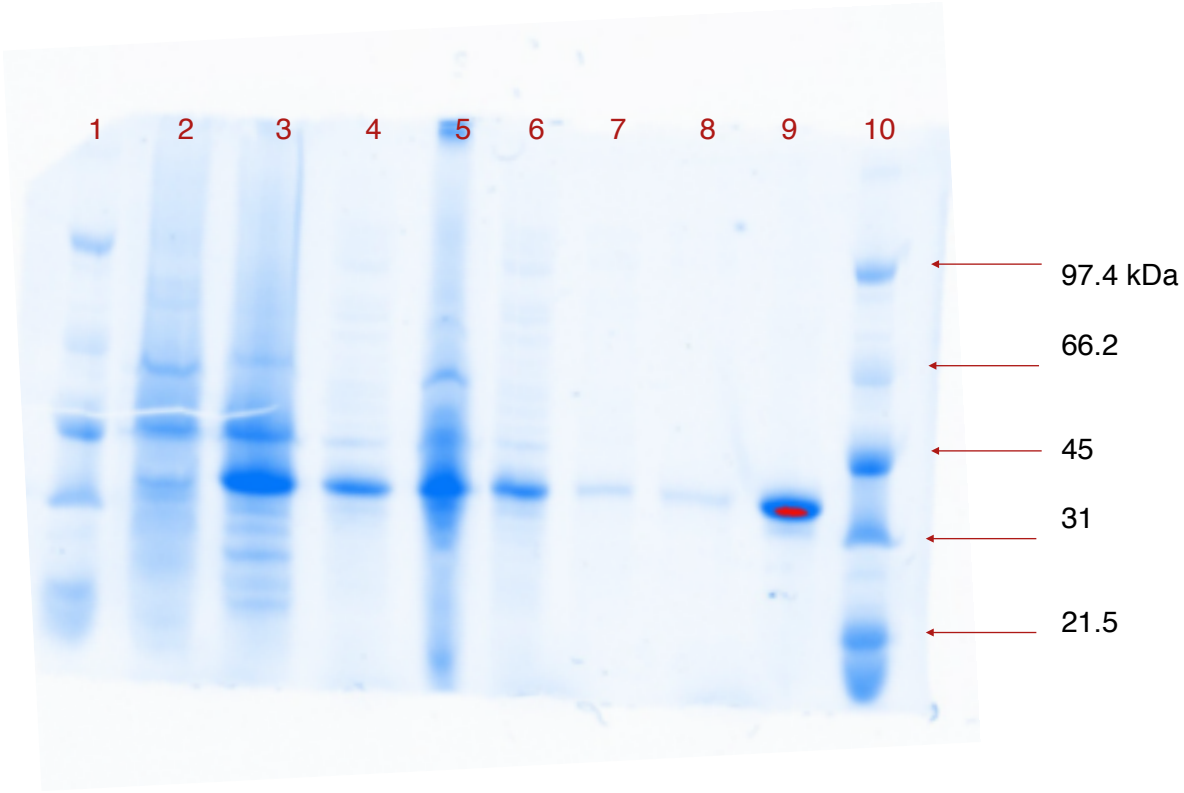


Figure 19. SDS-PAGE for Ari-PngC. Lane 1 and 10 show molecular weight ladder. Lane 2 shows uninduced fraction, lane 3 shows induced fraction, lane 4 shows raw extract, lane 5 shows cell debris, lane 6 shows flow through, lane 7 shows wash 1, lane 8 shows wash 2 and lane 9 shows pure protein at around 31 kDa.

Ari-PngC was expressed and purified, and can be seen on the SDS PAGE gel at around 31 kDa which is the correct size for the gene. It can also be seen from the gel that the protein is pure, and can be used for further analysis. After performing a bradford assay, the concentration of the gene was 2.4 mg/mL from a 600 mL purification.

Preliminary crystal trays was set up for Ari-PngC using the MCSG1, MCSG2, Crystal screen 1&2, and Index but no hit was obtained.

The next bacterial strain with LicC and LicD homologs to express and purify was *O.uli.* Instead of cloning from genomic DNA, synthetic genes for Oul-*pngC* and Oul-*pngD* were purchased. These genes can be ligated into pET29 without any additional cloning. Both Oul-*pngC* and Oul-*pngD* constructs came in a PBK vector that has a

kanamycin resistance, and NdeI and XhoI cut sites. A restriction digest with NdeI and XhoI was done on OuI-*pngC*-PBK and OuI-*pngD*-PBK constructs and run on a 1% agarose gel.

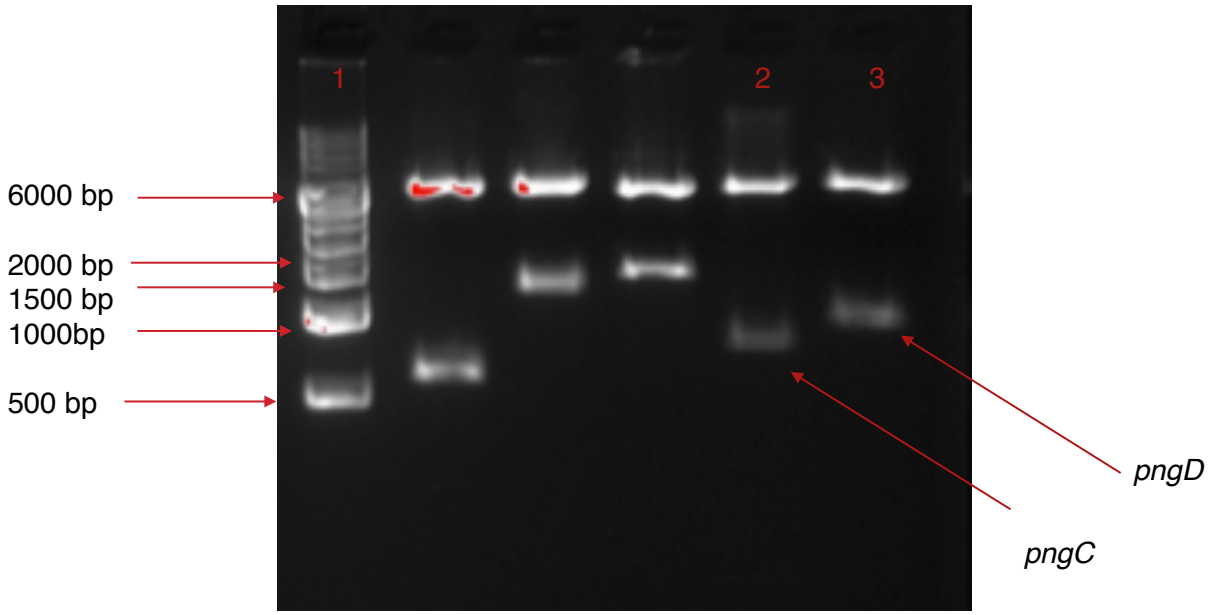


Figure 20. Restrict digest of OuI-p*ngC*-PBK and OuI-p*ngD*-PBK constructswith NdeI and XhoI. Lane 1 shows the molecular weight ladder. Lane 2 shows digested LicC at 790 bp, and lane 3 shows digested LicD at 930 bp. Unlabelled lanes are irrelevant to this research.

From the gel it can be seen that OuI-*pngC* and OuI-*pngD* constructs are 790 bp and 930 bp respectively which confirms that they are at a correct size. Both genes were extracted from the gel,and cleaned. The cleaned genes were run on a 1% agarose gel to confirm the concentration and presence of the band.

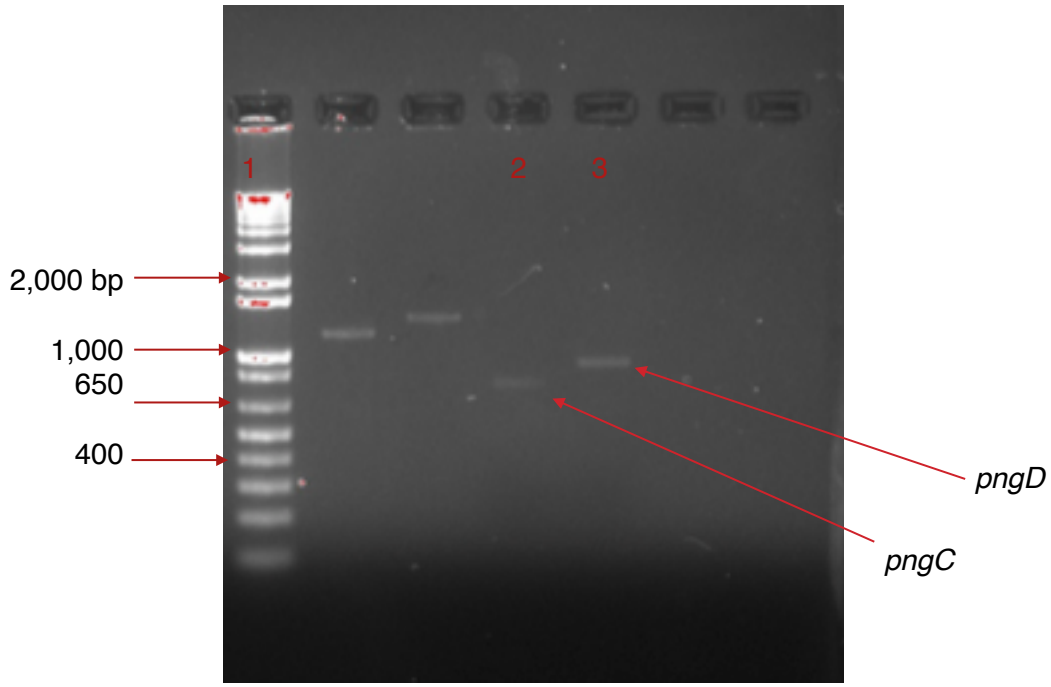


Figure 21. Cleaned gel extracts of Oul-P*ngC* and Oul-P*ngD*. Lane 1 shows molecular weight ladder. Lane 2 shows cleaned LicC product at 790 bp and lane 3 shows cleaned LicD product at 930 bp. Unlabelled lanes are irrelevant to this research.

Both of the genes were cleaned successfully, and were ligated into pET29. The ligations were transformed into Dh5a competent cells, grown and mini-prepped. To check if the genes were successfully ligated, restriction digests were done with NdeI and XhoI.

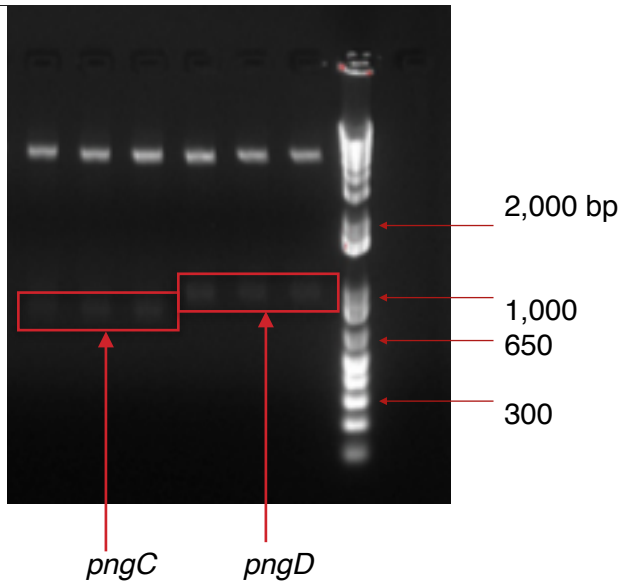


Figure 22. OuI-*pngC*pET29 and OuI-*pngD*pET29 construct digests with NdeI and XhoI restriction enzymes. Lanes 1-3 show LicC construct digest at 790 bp and Lanes 4-6 show LicD construct digest at 930 bp. Lane 7 shows molecular weight ladder.

It can be concluded from Figure 21 that the genes were successfully ligated into pET29. The genes were then transformed into BL21 competent cells to be expressed and purified using IMAC. The purified genes were run on an SDS -PAGE to see purity of the proteins.

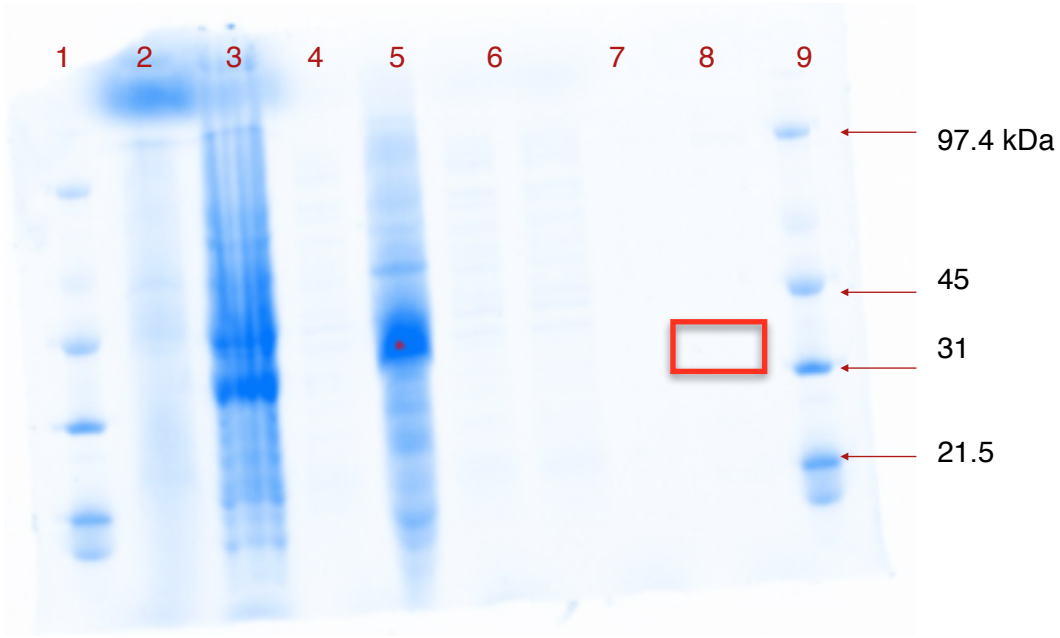


Figure 23. SDS-PAGE for Oul-PngC. In order from left to right, lane 1 and 9 are molecular weight ladders, lane 2 is uninduced fraction, lane 3 is induced fraction, lane 4 is raw extract, lane 5 is cell debris, lane 6 is wash 1, lane 7 is wash 2, and lane 8 is final eluted protein. The expected size of the protein is around 31 kDa.

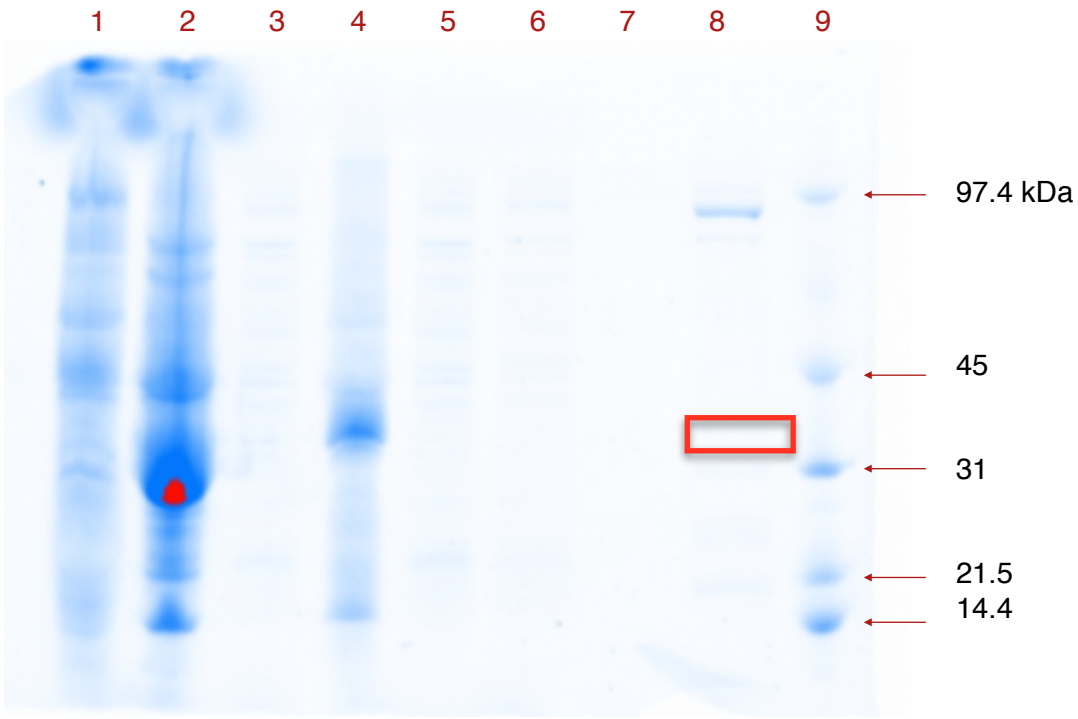


Figure 24. SDS PAGE for Oul-PngD. In order from left to right, lane 1 is uninduced fraction, lane 2 is induced fraction, lane 3 is raw extract, lane 4 is cell debris, lane 5 is flow through, lane 6 is wash 1, lane 7 is wash 2 and lane 8 is final eluted protein and lane 9 is molecular weight ladder. The expected size of the protein is around 31 kDa

It can be seen from the gels above that both Oul-*pngC*-pET29 and Oul-*pngD*-pET29 constructs are expressed but not soluble. Most of the protein can be seen in the cell debris which mostly comprises of insoluble fractions. From Bradford assays, the concentration of both proteins was 0 mg/mL which further proves that none of the proteins were soluble.

The next steps were to make soluble versions of Oul-PngC and Oul-PngD, and to do this pMAL vector was used. pMAL vector possess a maltose binding protein tags which helps solubilize proteins. Primers were made (Appendix) for the pMAL cloning with XmnI and NotI cut sites. Phusion PCR was performed with the new primers on Oul-*pngC*-pET29 and Oul-*pngD*-pET29 constructs, to introduce the MBP tag onto the gene. The products of PCR were run on a 1% agarose gel.

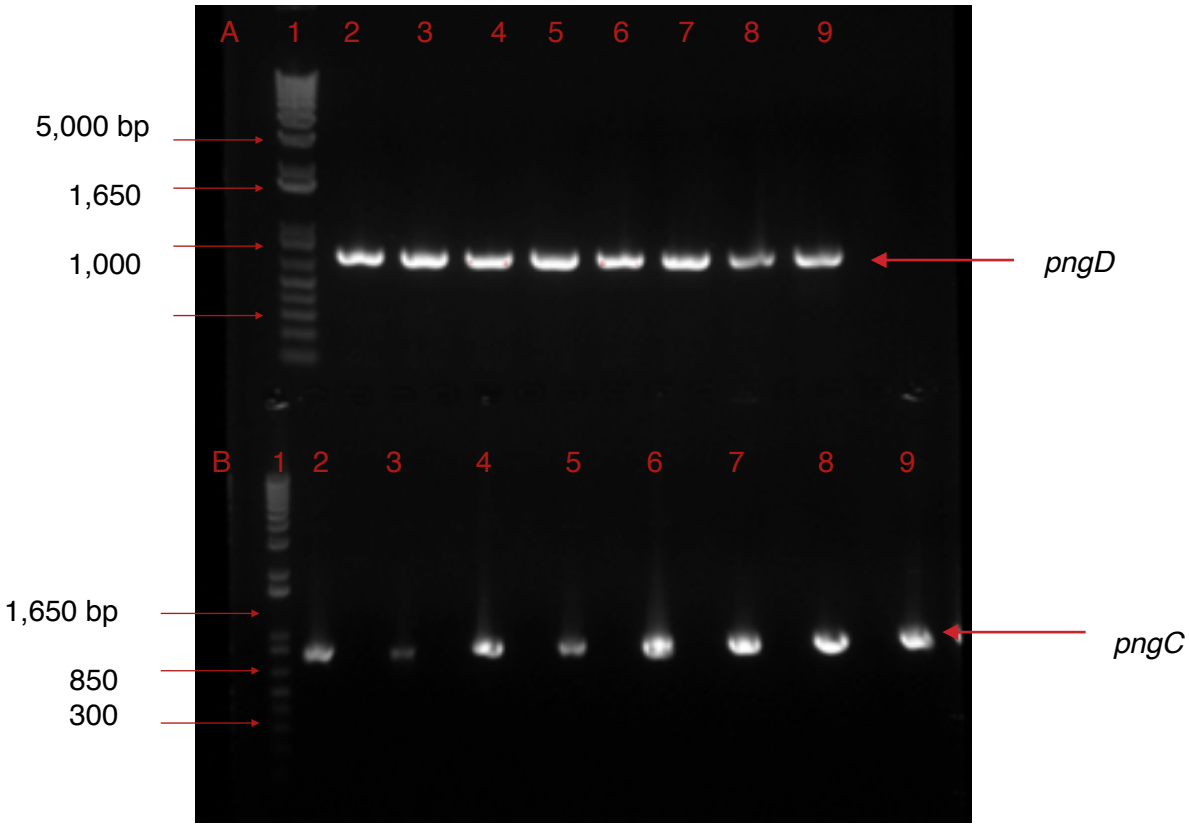


Figure 25. Phusion PCR of Oul-*pngC* and Oul-*pngD* for pMAL cloning. Gel A shows *pngC* amplicons in lanes 2-9, with molecular weight ladder in lane 1. Gel B shows amplicons of *pngD* in lane 2-9, with molecular weight ladder in lane 1. The expected size *pngC* is around 790 bp and *pngD* is around 930 bp.

The PCR was successful, and the bands were extracted from the gel and cleaned using Promega Wizard Clean Up System. The cleaned up amplicons were run on a 1% agarose gel.

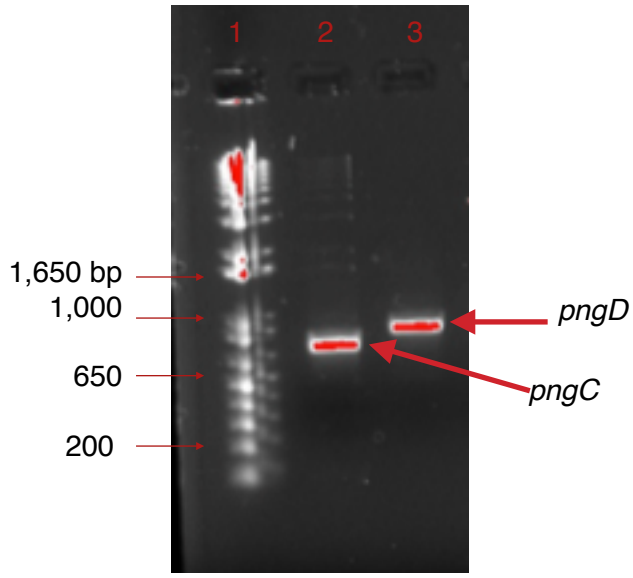


Figure 26. Gel extract cleans of Oul-*pngC* and Oul-*pngD* amplicons. Oul-*pngC* clean is in lane 2 at 790 bp and Oul-*pngD* clean is in lane 3 at 930 bp. Lane 1 shows the molecular weight ladder.

The amplicons were the correct size, and are ready for the addition of adenine overhang to ligate into pGEM. Adenine overhang was added and genes were run on a 1% agarose gel to confirm size.

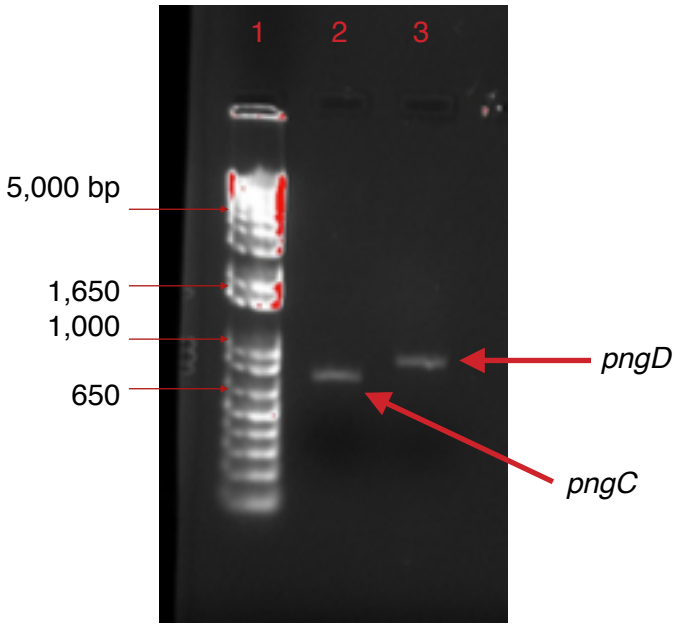


Figure 27. Adenine overhang addition to Oul-*pngC* and Oul-*pngD*. Oul-*pngC* is in lane 2 at 790 bp and Oul-*pngD* is in lane 3 at 930 bp. Molecular weight ladder is in lane 1.

The agarose gel shows the genes to be at the correct size. The genes were then ligated into pGEM and sent out for sequencing. The sequencing results came out negative, and it was determined that the primers were made incorrectly. New primers were made with NdeI and EcoRI as the new cut sites and Phusion PCR was done on OuI-*pngC*-pET29 and OuI-*pngD*-pET29 constructs. PCR was successful, the genes were cleaned, and ligated into pGEM successfully. The new constructs were sent out for sequencing and the results came out positive. The constructs were cut out of pGEM, and ligated into pMAL using NdeI and EcoRI cut sites. The ligations were transformed into Dh5a competent cells. Few colonies were picked, grown and mini-prepped. To confirm whether the correct inserts were ligated into pMAL successfully, a double digest with NdeI and EcoRI was done.

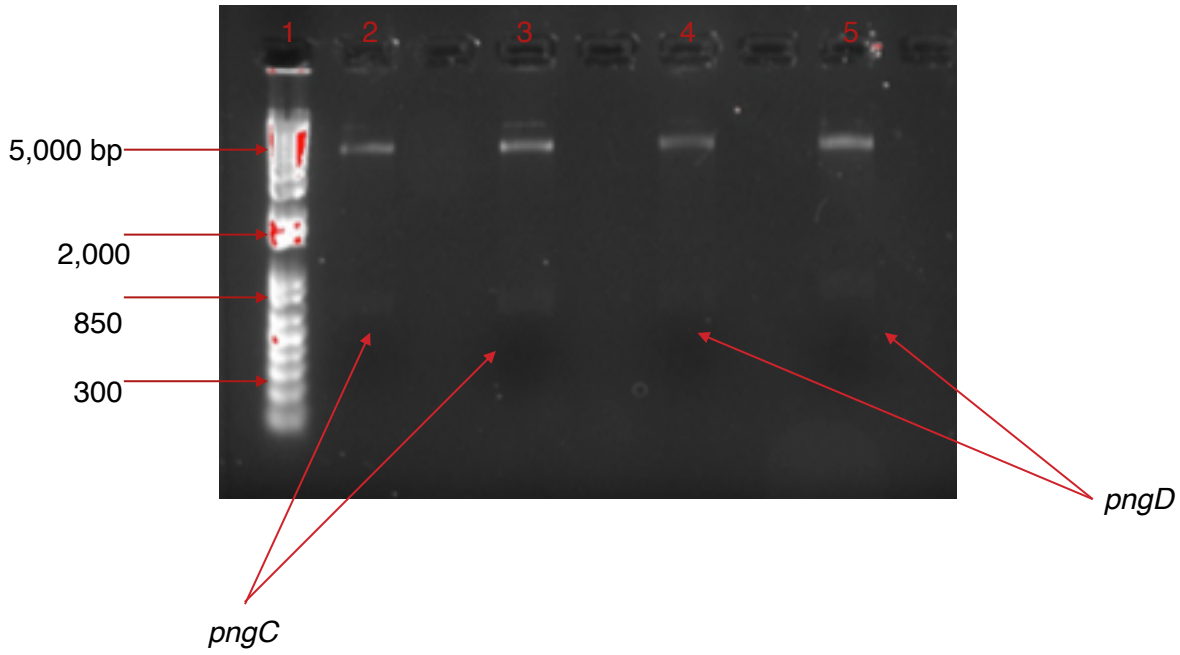


Figure 28.Double digest of OuI-*pngC*-pMAL and OuI-*pngD*-pMAL constructs with NdeI and EcoRI. LicC is in lane 2-3 at 790 bp, LicD is in lane 4-5 at 930 bp and molecular weight ladder is in lane 1.

Since the ligation into pMAL was successful, the next step was to express and purify the construct. Both constructs were expressed in 1 L of LB culture and purified using anion exchange since pMAL does not have a His-tag.

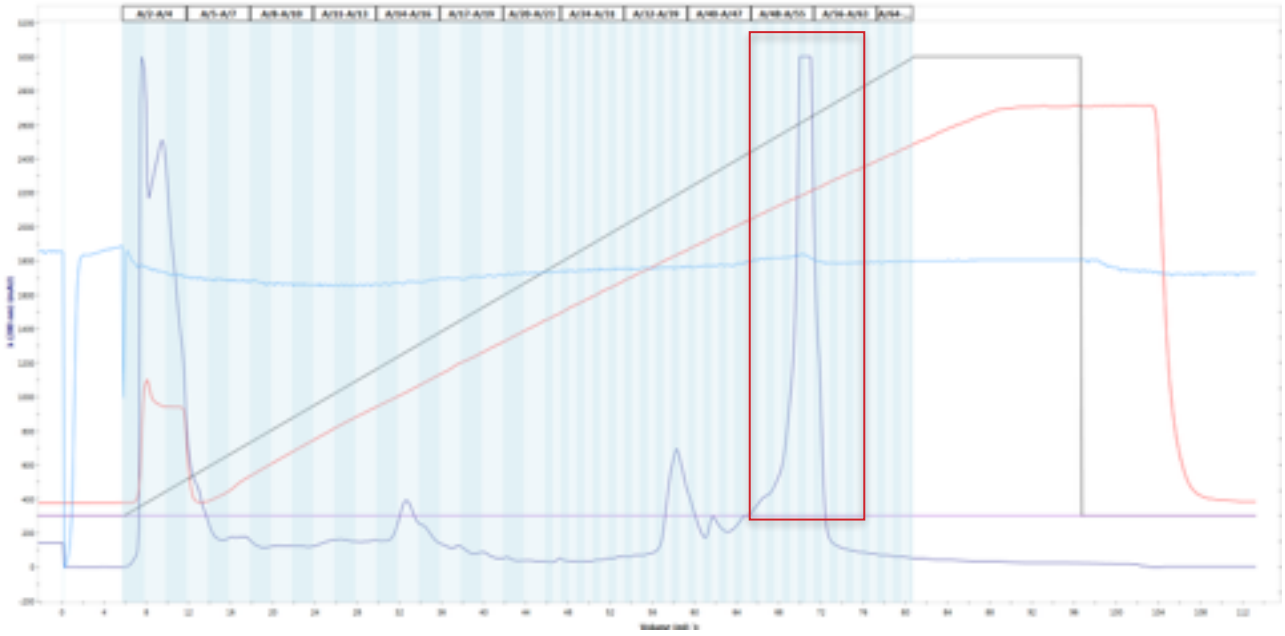


Figure 29. FPLC chromatogram of Oul-PngC-pMAL .

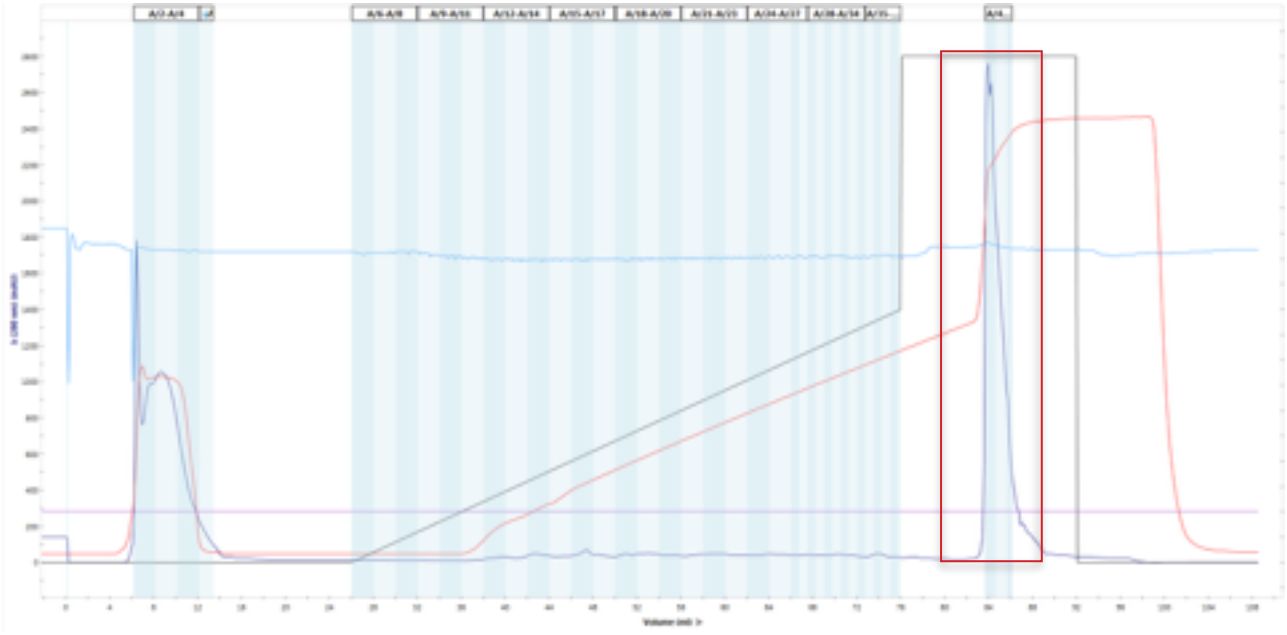


Figure 30. FPLC chromatogram of Oul-PngD-pMAL.

The protein fractions boxed in Figure 29 and 30 were collected, concentrated and run on SDS-PAGE to determine the purity of the proteins.

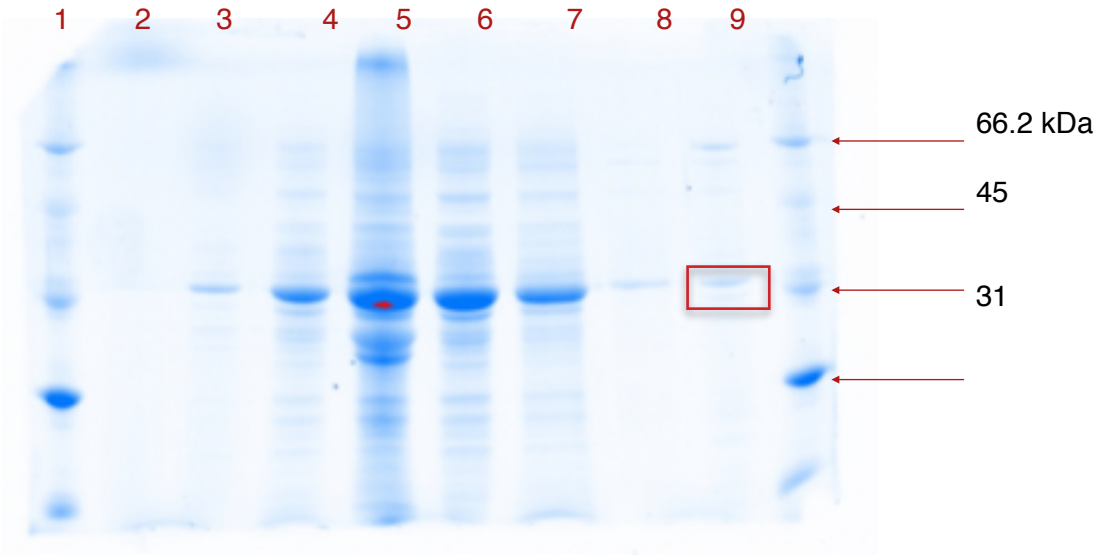


Figure 31. SDS-PAGE for Oul-PngCpMAL. In order from left to right, lanes 1 and 10 are molecular weight ladder, lane 2 is uninduced fraction, lane 3 is induced fraction, lane 4 is raw extract, lane 5 is cell debris, lane 6 is flow through, lane 7 is wash 1, lane 8 is wash 2 and lane 9 is final eluted protein. The expected size of this construct is approximately 90 kDa.

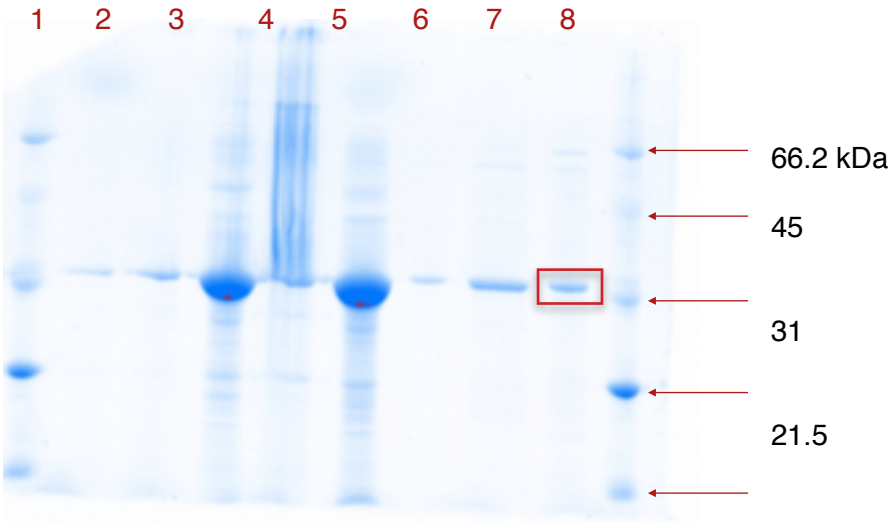


Figure 32. SDS-PAGE for Oul-PngD-pMAL. In order from left to right, the wells contained the following: SDS molecular weight marker, uninduced fraction, induced fraction, raw extract, cellular debris, flow through, wash 1, wash 2, final eluted protein and another SDS molecular weight marker.

Oul-PngC-pMAL and Oul-PngD-pMAL constructs should be around 80 kDa but from both gels it can be seen that both the constructs are at around 31 kDa. It seems like the MBP tag was not incorporated to the genes properly. Since Ari-PngCwas already soluble, it was chosen to be pursued.

To be able to determine whether Ari-PngC behaves in a similar fashion as the Spn-LicC, the *licC* gene was cloned and purified. Spn-*licC* was purchased from BioBasic in a pUC57 vector, with NdeI and XhoI cut sites and a kanamycin resistance. The gene was cut out of pUC57 vector using NdeI and XhoI restriction enzymes, and ligated into pET29 vector using protocols from the methods section. The product of ligation was subjected to another restriction digest with NdeI and XhoI restriction enzymes to verify the success/failure of the ligation. The digests were run on a 1% agarose gel.

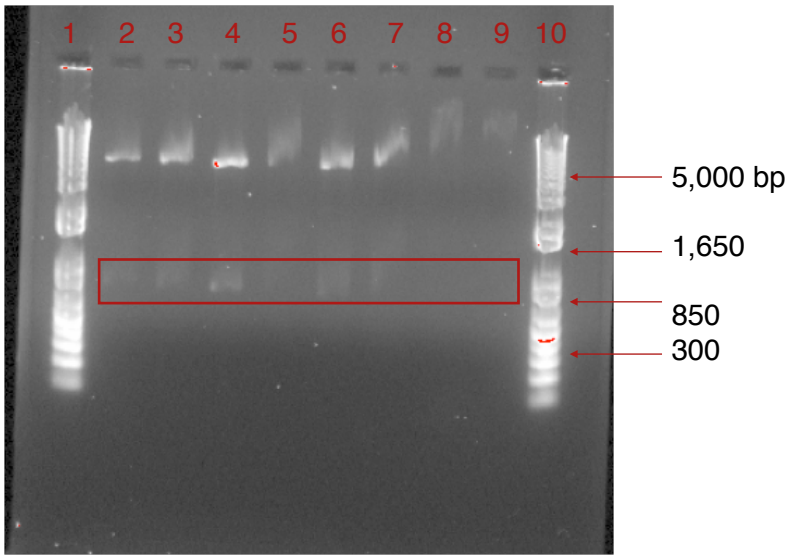


Figure 33. Double digest using NdeI and XhoI of Spn-*licC*-pET29. Lanes 1 and 10 show molecular weight ladder, and lanes 2-9 represent different colonies of the ligated product.

It can be seen from Figure 33, that the ligation was successful. The top band represents pET29 band which is at the expected size of 5000 bp, and the bottom band represents *licC* gene which is at the expected size of 774 bp. The gene was then purified using IMAC, and purified gene was run on a SDS-PAGE gel to determine purity.

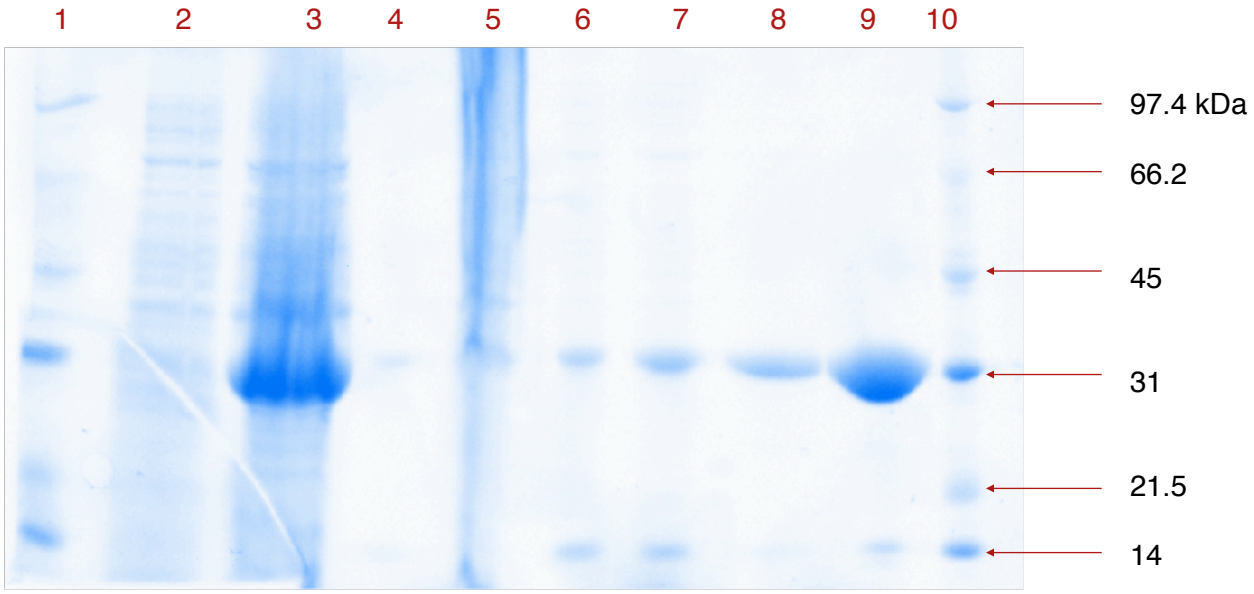


Figure 34. SDS-PAGE gel of Spn-LicC after IMAC purification. In order from left to right, lane 1 and 10 are molecular weight ladder, lane 2 is uninduced fraction, lane 3 is induced fraction, lane 4 is raw extract, lane 5 is cell debris, lane 6 is flow through, lane 7 is wash 1, lane 8 is wash 2 and lane 9 is final eluted protein. The expected size of protein is 31 kDa.

3.3 Pyrophosphate Assay

Having a soluble Spn-LicC and Ari-PngC, the next step was to determine activity of those genes for their respective substrates, and each others substrates. To determine activity, an EnzCheck Pyrophosphate Assay Kit was used. When the cytidylyltransferase enzyme catalyzes the reaction, it releases pyrophosphate which through this kit can be detected. The amount of pyrophosphate released will dictate the rate at which the cytidylyltransferase can catalyze the reaction. In using this kit, there

are several factors to be considered; one is that all the components being used have to be phosphate free, second is that calcium inhibits the enzyme and therefore components have to be calcium free. Another factor to consider when trying to determine the activity of the enzyme, an enzyme-activity test needs to be completed to determine at which concentration the enzyme works best.

The first step of determining the activity of the cytidylyltransferase enzymes was to re-purify the proteins using phosphate free buffers. As well as to help with stability, DTT and Ammonium sulphate were also added to the buffers (Appendix). Both Spn-LicC and Ari-PngC were purified using Tris buffer, and run over SDS-PAGE to determine purity.

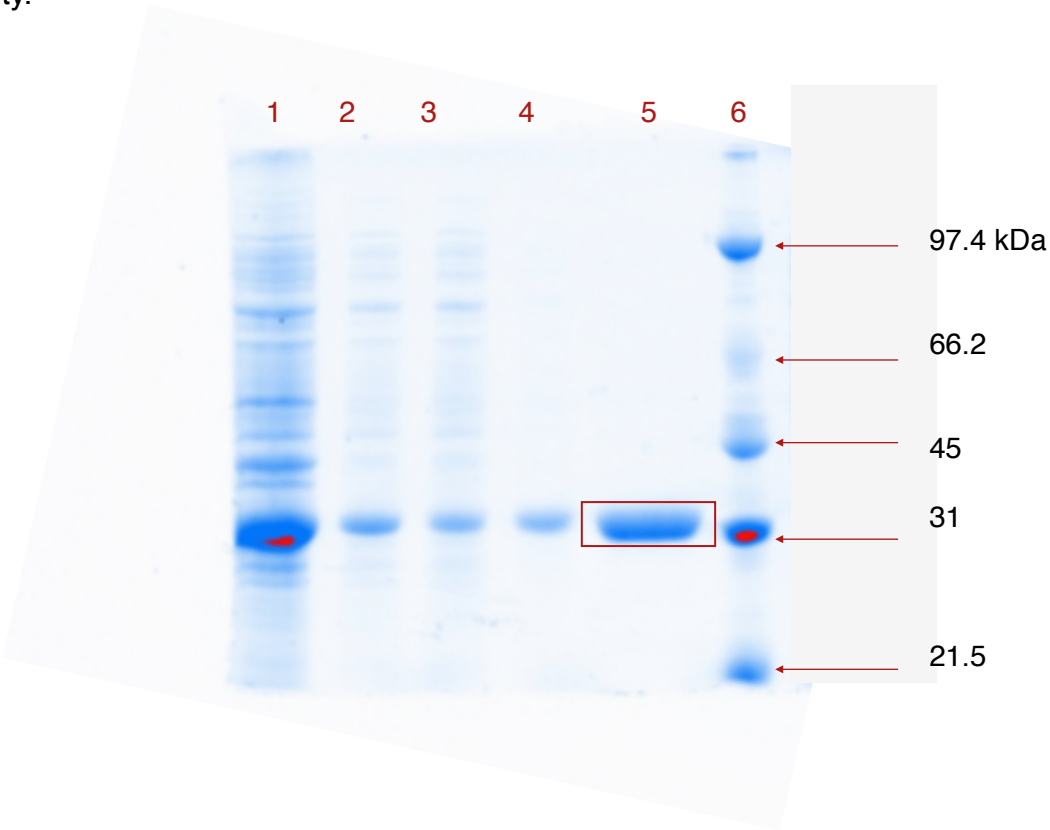


Figure 35. SDS-PAGE gel of Spn-LicC after IMAC purification with Tris buffer. In order from left to right, lane 1 is raw extract, lane 2 is flow through, lane 3 is wash 1, lane 4 is wash 2, lane 5 is final eluted protein at around 31 kDa, and lane 6 is molecular weight ladder.

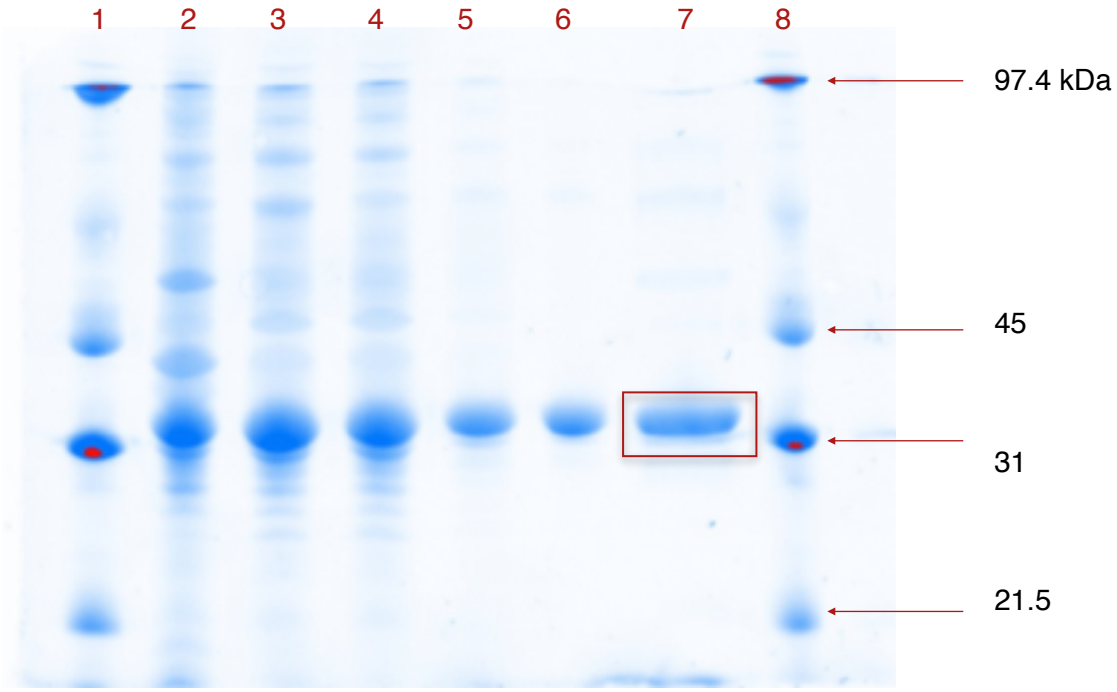


Figure 36. SDS-PAGE gel of Ari-PngC after IMAC purification with Tris buffers. In order from left to right, the wells contained the following: SDS molecular weight marker, cellular debris, raw extract, flow through, wash 1, wash 2, final eluted protein and another SDS molecular weight marker.

Both Spn-LicC and Ari-PngC are pure and are at the correct sizes on the SDS-PAGE gels. The next step was to set up the reactions and determine the activity of enzymes. The first to have its activity determined was Spn-LicC. While carrying out the experiments, it was observed that as you increased the phosphocholine concentration which is the preferred substrate, the rate of reaction decreased. It was discovered that phosphocholine is commercially available only as the calcium salt, and it is previously established that calcium inhibits this enzyme. A way to overcome that was to use Chelax resin to strip any calcium from phosphocholine prior to using it. Methods were established and can be found in the Appendix. Once the phosphocholine was clean of

any potential calcium, a test run was performed using non chelaxed phosphocholine and chelaxed phosphocholine.

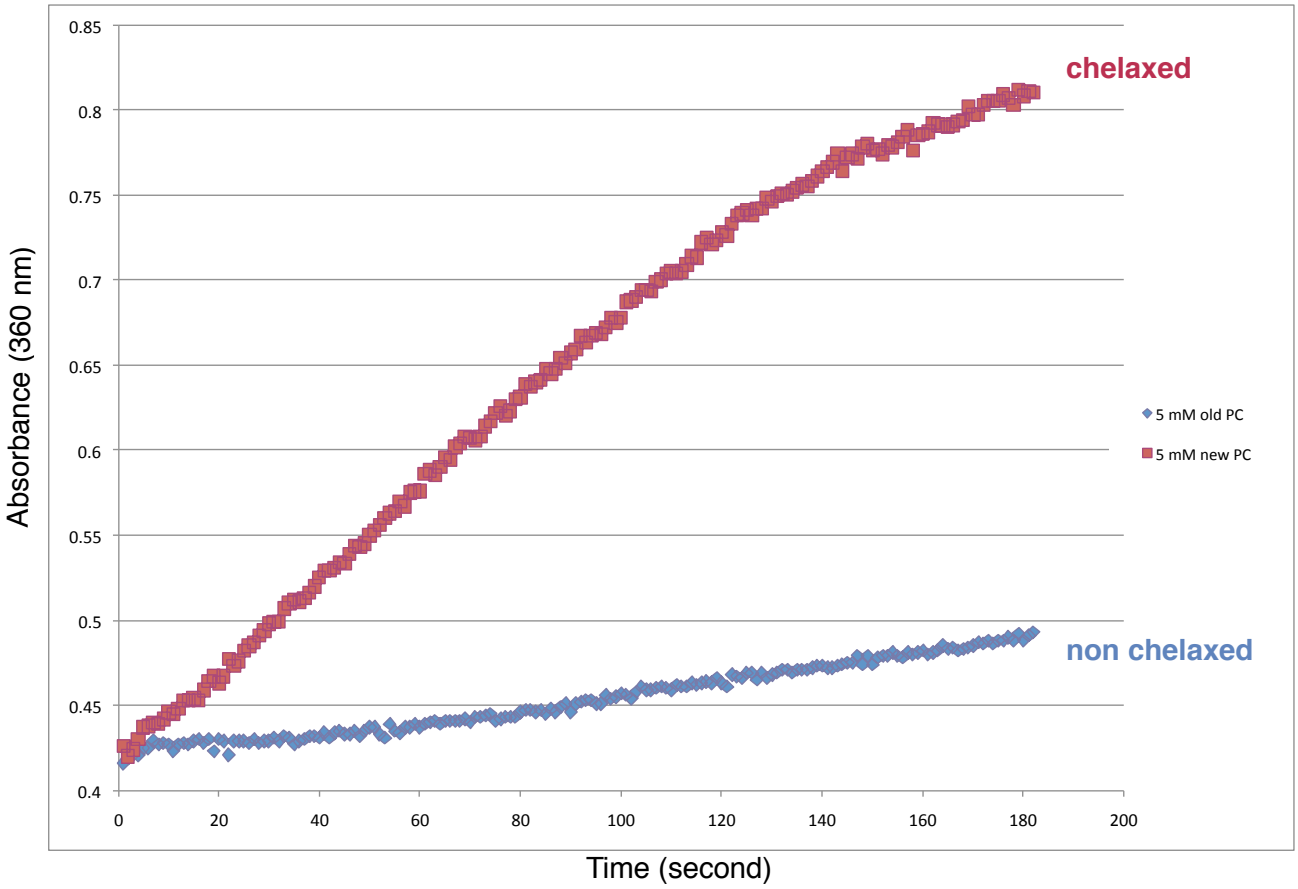


Figure 37. Pyrophosphate release from Spn-LicC enzyme with 5 mM chelaxed and non chelaxed phosphocholine.

From the figure above it can be seen that the absorbance is higher in the chelaxed phosphocholine than the non chelaxed phosphocholine. From this test we can conclude that the chelax protocol was successful, and the removal of calcium reversed the calcium inhibition.

After preliminary runs of both enzymes, the next step was to determine which concentration of enzyme would be best to use with this pyrophosphate assay. An

enzyme-activity test was completed, where different concentrations of enzyme were used with 0.1 mM substrate. An enzyme activity assay for LicC with 0.1 mM phosphocholine included a range of enzyme concentrations of 0.1 nM to 1000 nM. From the graph in Figure 38 it can be seen that the rate increases as the enzyme concentration decreases, and rate decreases as enzyme concentration increases. This can also be seen in the graph in Figure 39, where LicC from *A. rimae* is tested with 0.1 mM AEP and range of enzyme concentrations of 0.1 nM to 100 nM. It appears that these enzymes may be working in a monomeric and dimeric forms depending on the enzyme concentration.

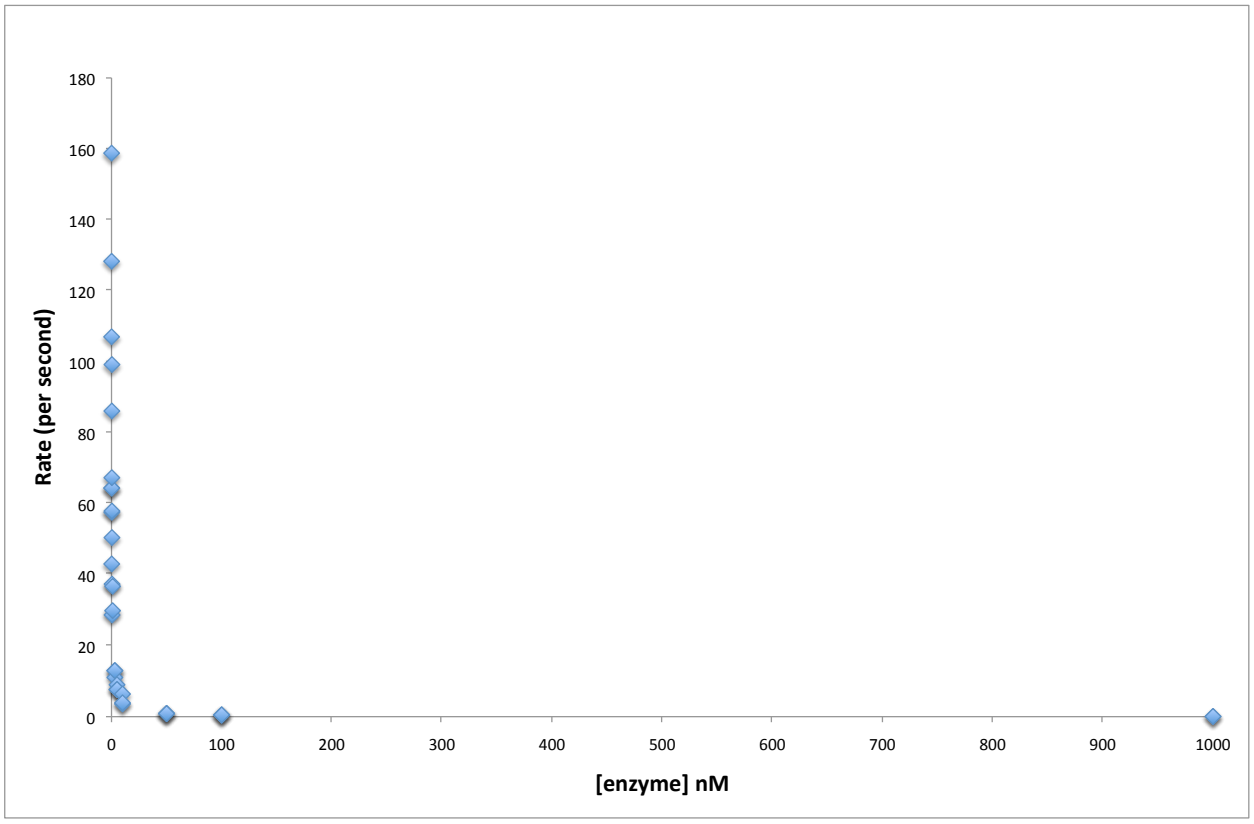


Figure 38. Spn-LicC activity assay with 0.1 mM phosphocholine and various concentrations of enzyme.

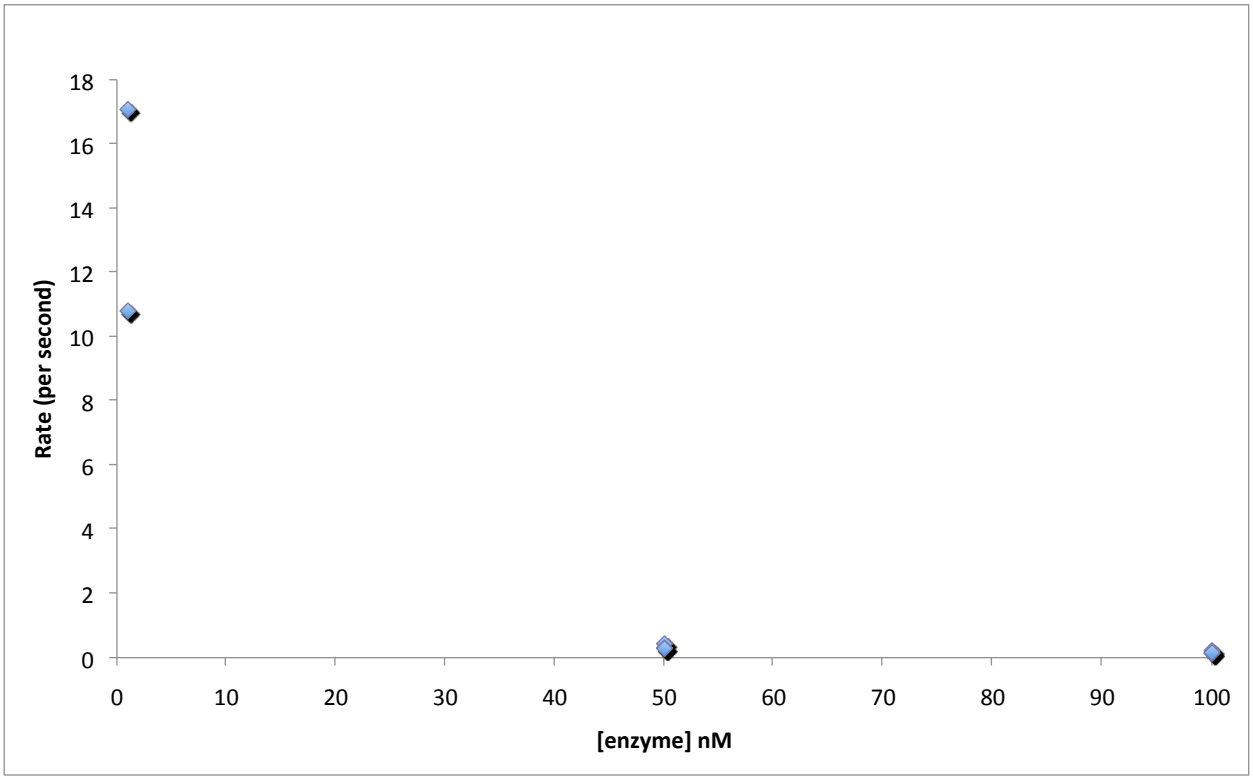


Figure 39. Ari-Png activity assay with 0.1 mM AEP and various concentrations of enzyme.

From these enzyme activity test, 50 nM of enzyme was picked to be used to get full kinetic data. This concentration of enzyme was picked because it was in the middle of the ranges of enzyme concentrations used, and gave activity that could be easily detected and measured.

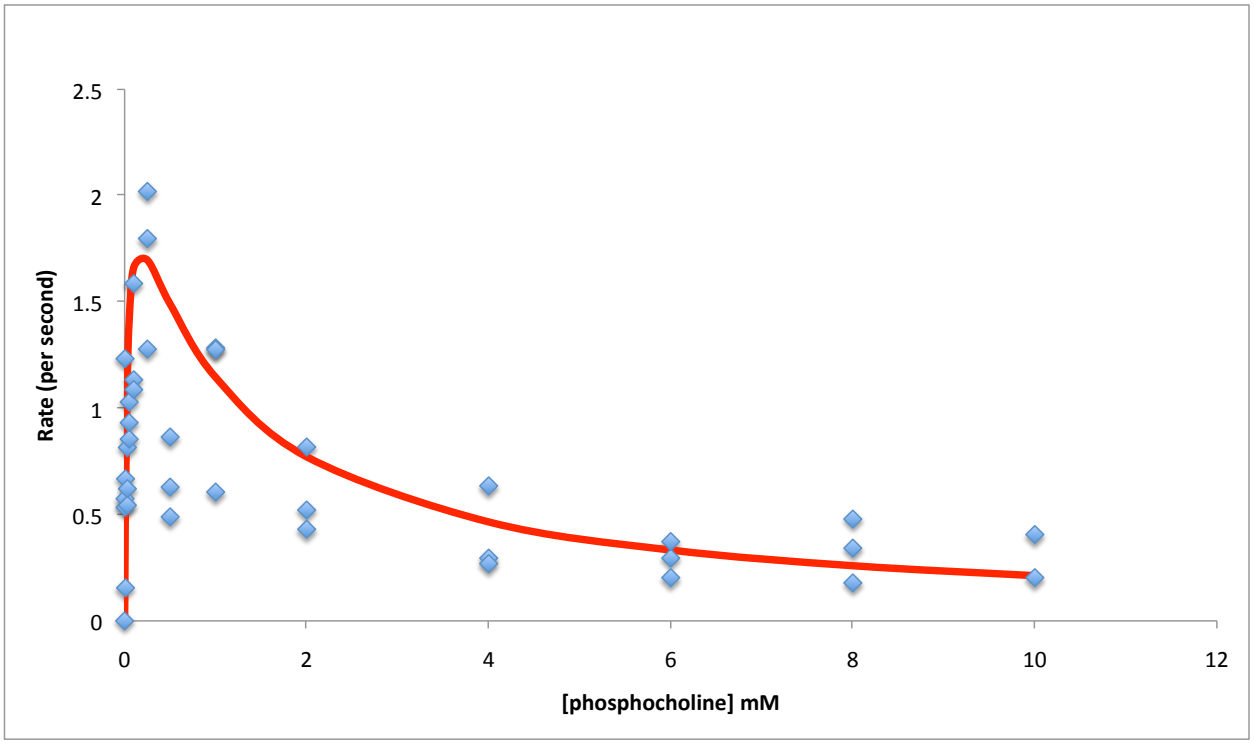


Figure 40. Michaelis -Menten fit for Spn-LicC with phosphocholine.

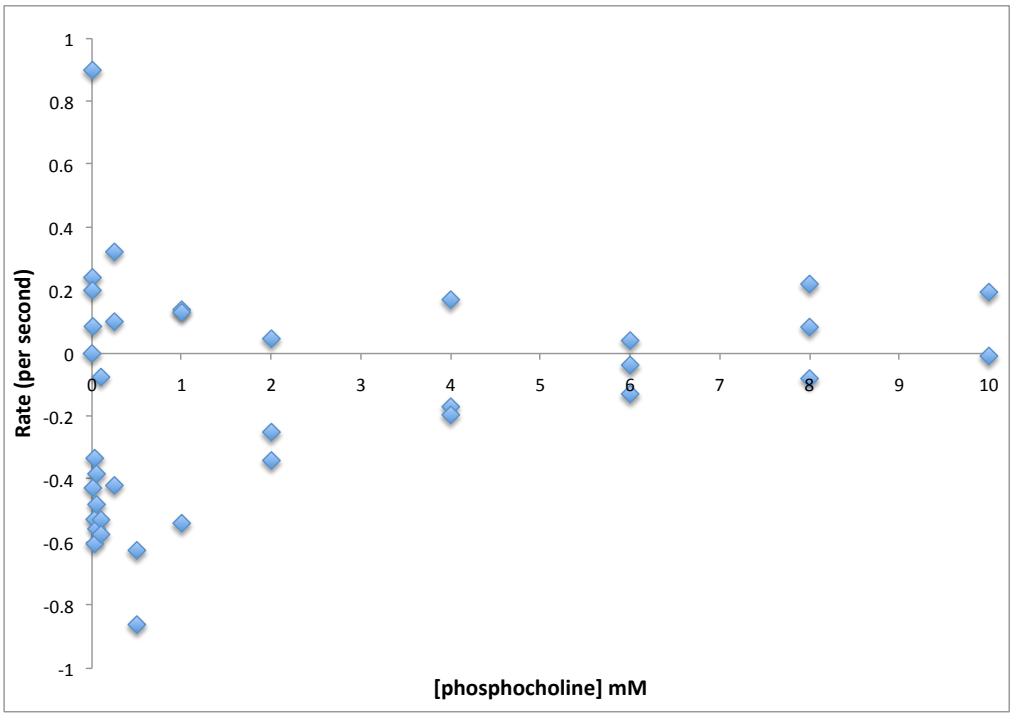


Figure 41. Residual plot for *Spn-LicC* with phosphocholine

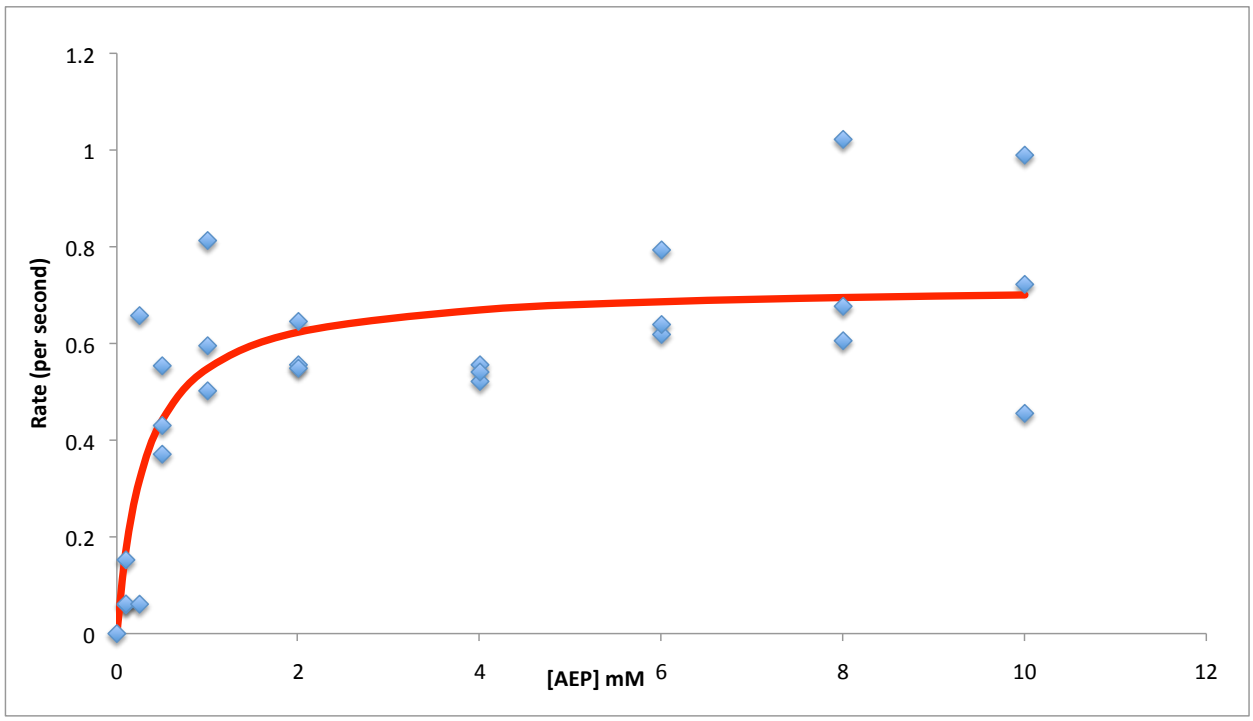


Figure 42. Michaelis-Menten fit for Spn-LicC with AEP

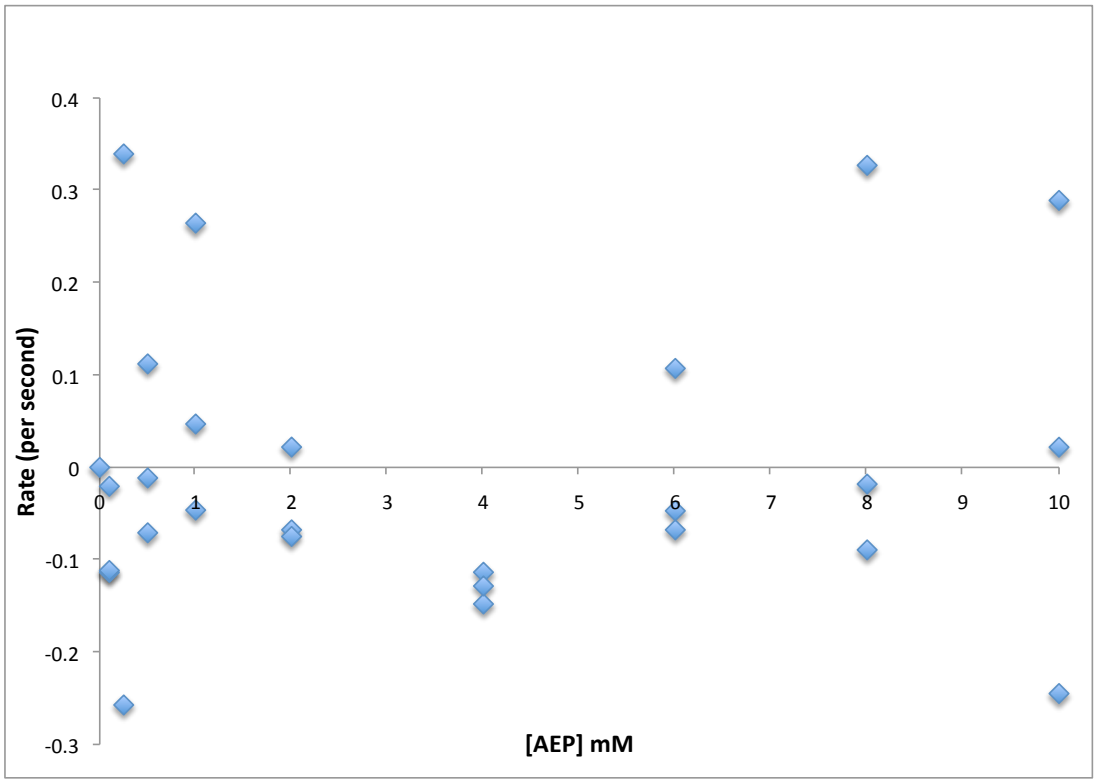


Figure 43. Residual plot for Spn-LicC with AEP

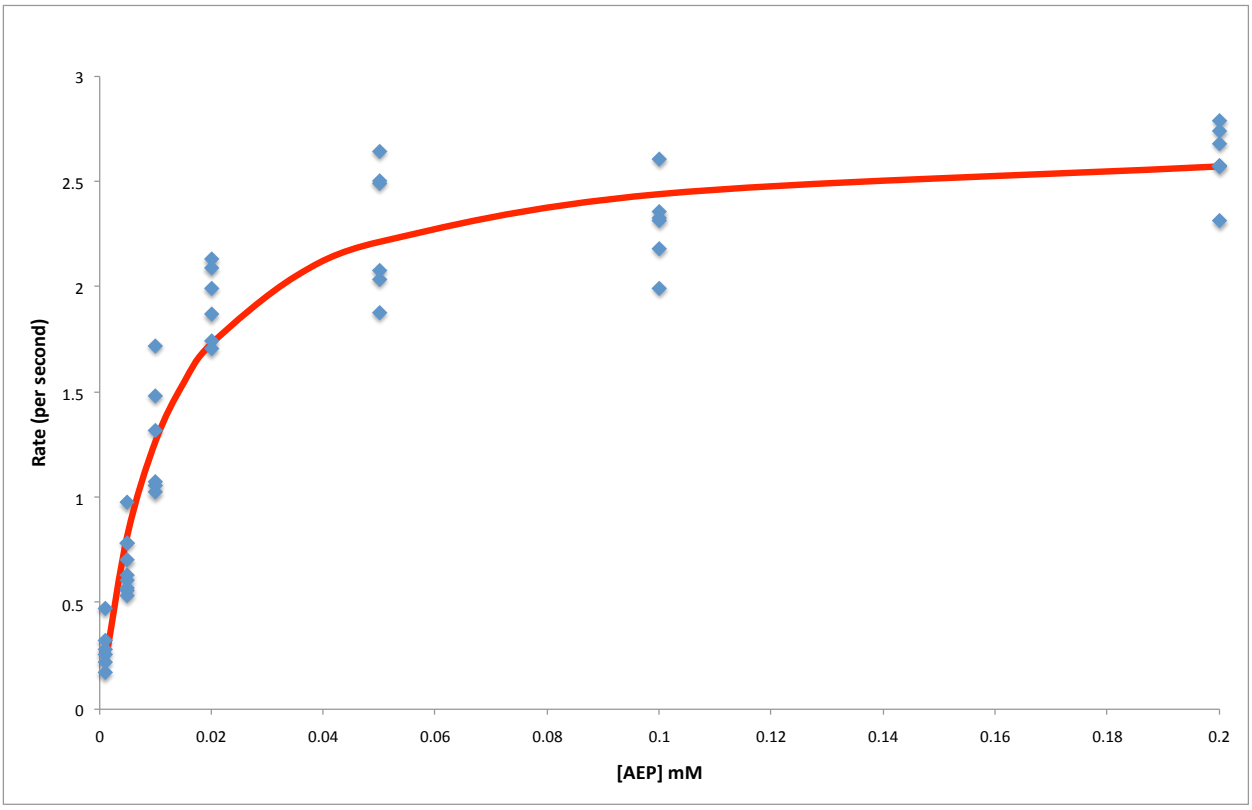


Figure 44. Michaelis -Menten fit for Ari-PngC with AEP

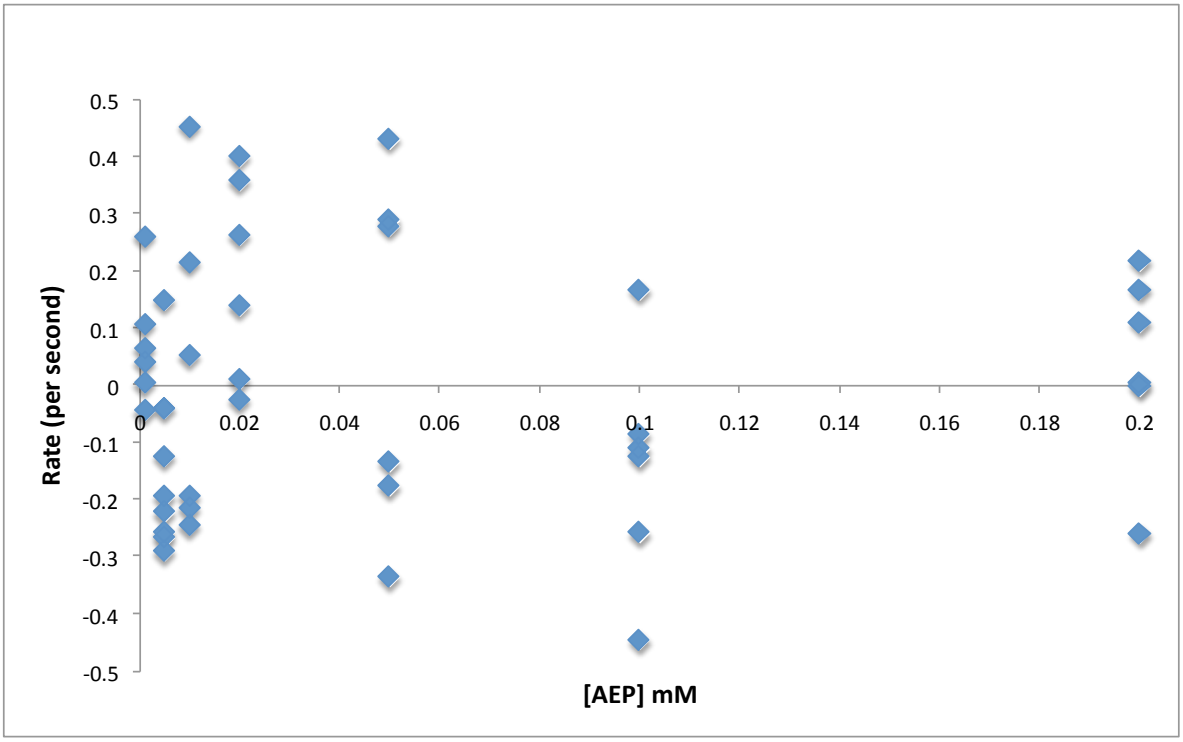


Figure 45. Residual plot for Ari-PngC with AEP

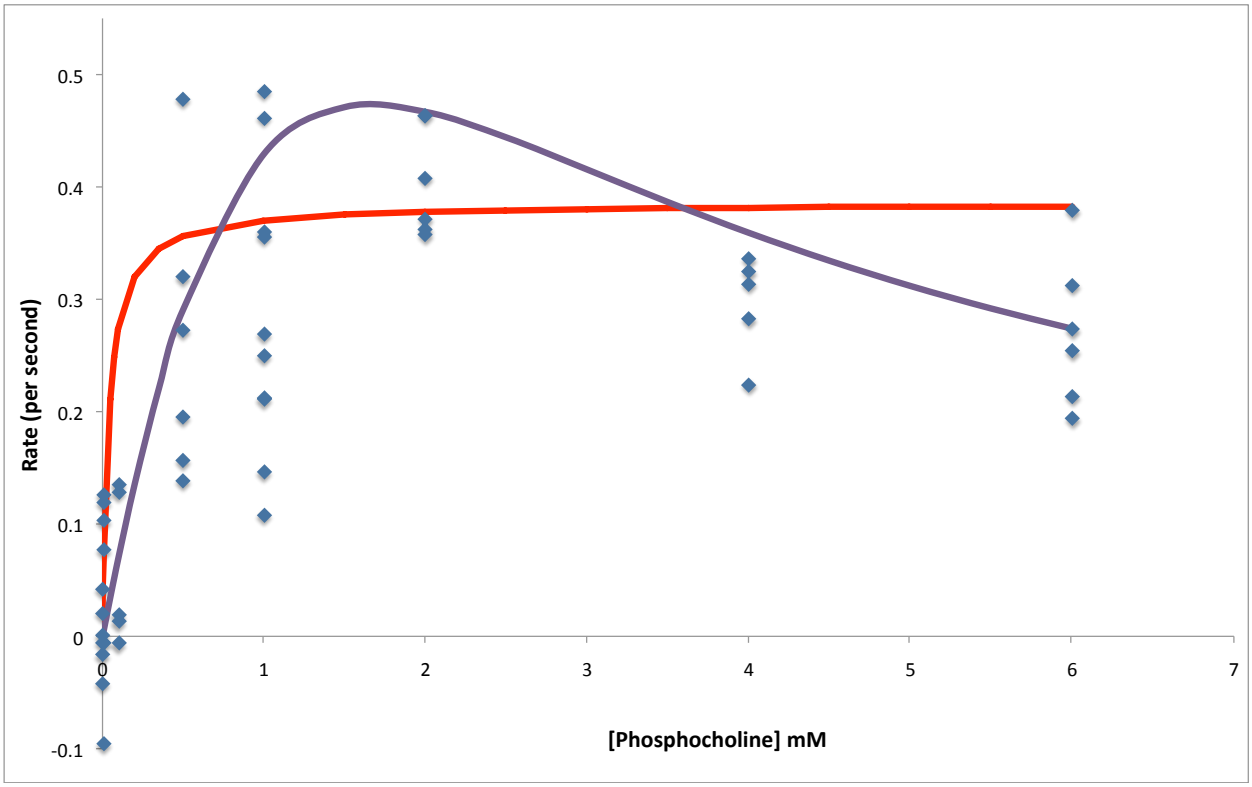


Figure 46. Michaelis - Menten fit (red) and substrate inhibition fit (purple) for Ari-PngC with phosphocholine.

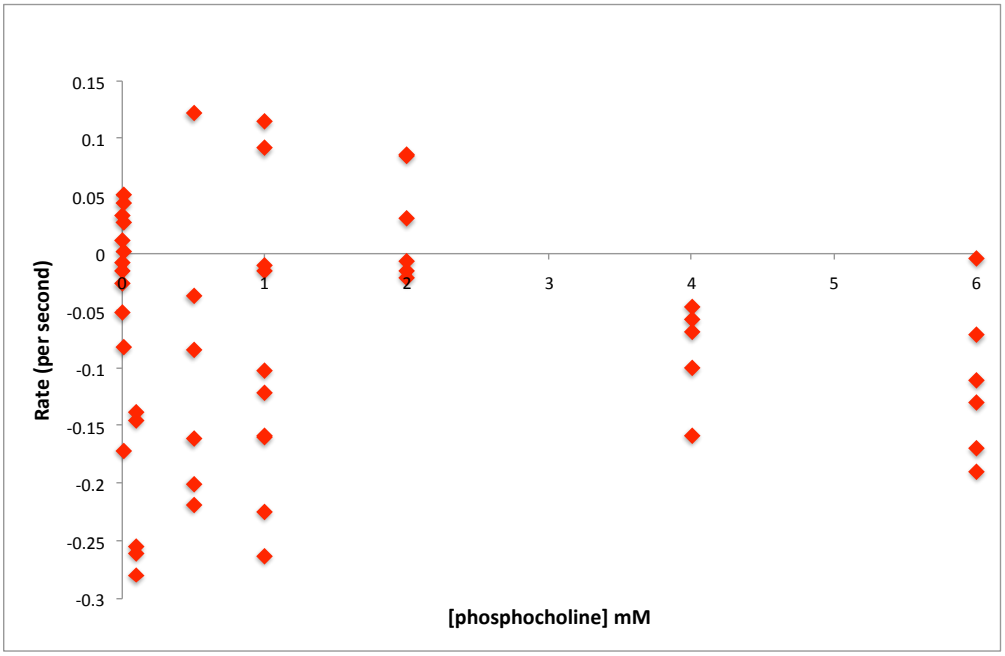


Figure 47. Residual plot for Ari-PngC with phosphocholine (Michaelis-Menten curve)

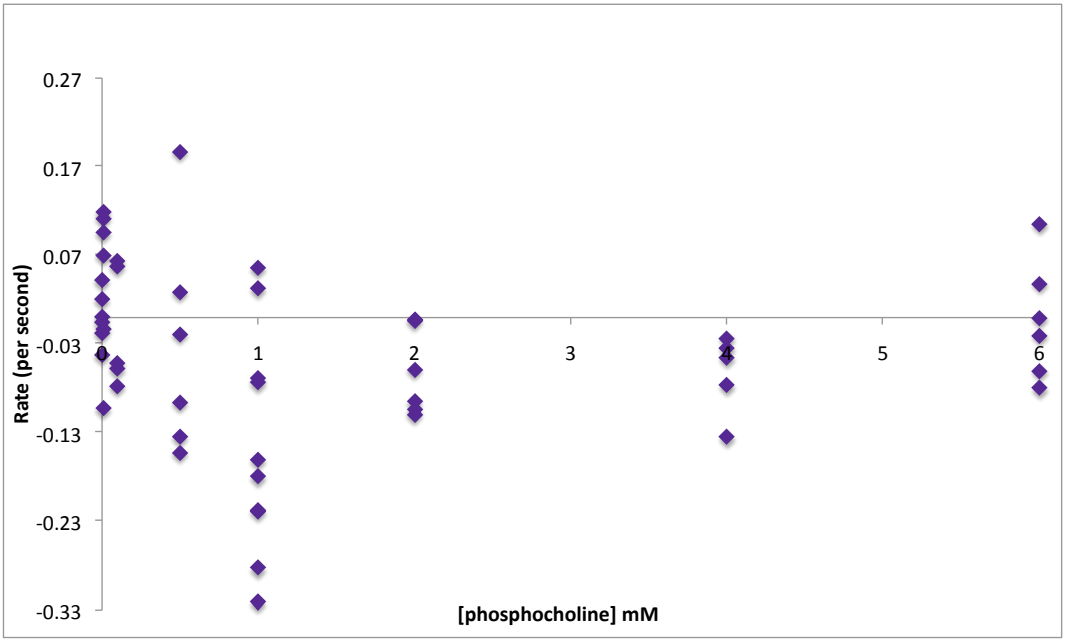


Figure 48. Residual plot for Ari-PngC with phosphocholine (Substrate inhibition)

Table 1. Summary of kinetic parameters for enzymes

Enzymes	Substrate	K_m (mM)	k_{cat} (per second)	K_S (mM)	k_{cat}/K_m (/s*mM)	Substrate Preference
Ari-PngC	PC (MM)	0.026 ± 0.097	0.344 ± 0.028	N/A	1.67	143 fold for AEP
	PC (Inhibition)	1.000 ± 0.826	0.717 ± 0.344	3.99 ± 3.52	0.74	322 fold for AEP
	AEP	0.011 ± 0.001	2.72 ± 0.079	N/A	239	
Spn-LicC	PC	0.020 ± 0.011	1.52 ± 0.243	1.53 ± 0.675	77.9	34 fold for PC
	AEP	0.318 ± 0.126	0.722 ± 0.053	N/A	2.27	

4. Discussion:

The purpose of this project was to functionally characterize cytidylyltransferases from *A. rimae* and *O. uli* to gain insight into the mechanism behind phosphonate containing macromolecules. The first step to characterizing these genes was to use bioinformatic tools such as Phyre2 to predict the preliminary structure. Phyre2 generated predicted structures based on amino acid sequences from the genes. The predicted structures of Oul-PngC and Ari-PngC were compared with the previously characterized Spn-LicC. The active site of Spn-LicC was overlaid with the predicted active sites of Oul-PngC and Ari-PngC to see the similarities and differences. In Figure 11, Spn-LicC in blue is overlaid with Ari-PngC. Important active site residues necessary for catalysis in LicC are Asp 107, Glu 226, Asp 218 and Trp 136[31,33]. In Figure 11, it can be seen that all of these important residues are present in the Ari-PngC as well. There are however, subtle differences in the active sites such as the distance of the residues from each other, which would dictate how well a substrate can bind in the active site to be catalyzed efficiently. Active sites of Oul-PngC was also compared with Spn-LicC in Figure 12. Important residue Trp 136 is switched to Phe 105, but all the remaining important active site residues are conserved. These differences in the active sites point to this enzyme having different substrate specificity.

The second step to characterize these genes was to clone, express and purify them to study their activity against various substrates. All three cytidylyltransferase genes were cloned and expressed in pET29, but only genes from *S. pneumoniae* and *A. rimae* were able to be purified successfully. Once these genes were purified, a pyrophosphate assay was used to determine their activity. The software R was used to

obtained the kinetic parameters. The first step to determining activity was to determine at which concentration these enzymes worked best, and to do so an enzyme -activity assay was performed. In this assay, different concentrations of enzyme were used with 0.1 mM of respective substrate and rate was monitored. From Figure 38 and 39 it can be seen that at lower concentrations of enzyme the rate of reaction is higher. This was an odd behaviour exhibited by the enzymes, however it this type of relationship between enzyme concentration and rate is precedented. A well studied example of this type of relationship is the D-amino acid oxidase[34,35]. D-amino acid oxidase exists in an equilibrium state of the monomer and dimer at appropriate concentrations; however after more analysis it was seen that at lower enzyme concentrations, D-amino acid oxidase shifts to favour the monomer state, giving higher rate values[34,35]. The opposite effect was seen with higher enzyme concentrations, where the rate decreased, and the equilibrium was shifted to favour the dimer state.[34,35]. This type of relationship is not exclusive to D-amino acid oxidase, but also exists in purine nucleoside phosphorylase[36], isocitrate dehydrogenase[37], UMP/CMP kinase[38] and many more. Although the structure of Spn-LicC is classified as a dimer, under these assay conditions, it demonstrates an oligomeric properties. This is also observed for Ari-PngC. This type of unorthodox relationship is interesting, and poses the question of whether this enzyme has just a monomer/dimer form or does it have the capabilities to exists in many oligomeric forms. The information obtained from these enzyme- activity assays was used to determine a concentration at which kinetic parameters were calculated. The kinetic parameters and their summary can be found in Table 1. for both Spn-LicC and Ari-PngC.

The first enzyme to be characterized by this pyrophosphate assay was Spn-LicC This enzyme has been previously biochemically characterized and in literature has a K_m ranging from 0.083 mM to 0.39 mM for its substrate phosphocholine, with a k_{cat} ranging from 3.0/s to 17.5/s[28,31]. There is a discrepancy in the values obtained from this assay and those available in literature which could be explained by the oligomeric forms of this enzyme as mentioned previously. The K_m and K_{cat} for LicC with phosphocholine (PC) were 0.020 ± 0.011 mM and 1.52 ± 0.243 /s respectively. Instead of a classic Michaelis - Menten curve, our data shows a substrate inhibition curve between LicC and phosphocholine, and because of this the K_s value was determined to be 1.53 ± 0.675 mM. From literature it is known that LicC is strongly inhibited by calcium, and phosphocholine when bought commercially is made with calcium. Even though best efforts were made to remove calcium from phosphocholine through Chelax, some calcium may have been left behind causing LicC to be inhibited at higher concentrations of phosphocholine. Further methods will need to be established to measure the calcium concentration in phosphocholine before performing enzymatic assays. Another theory to support substrate inhibition for LicC with phosphocholine can be modelled to a classic substrate inhibition model. Enzyme inhibition is one of the common derivations from Michaelis- Menten kinetics, and means that the velocity curve of a reaction rises to maximum as substrate concentration increases and then descends either to complete inhibitions or to a partial inhibition[39]. This happens when two substrate molecules bind to the enzyme at the active sites and non-catalytic inhibitory sites[39]. In our case, at higher concentrations of phosphocholine, it could be binding to one enzyme active site, rather than binding to various different enzyme active sites, causing the velocity of the

curve to descend. Having kinetic parameters of LicC with phosphocholine, the next step was to see if this enzyme has any specificity for other substrates. The substrate of choice was 2-aminoethlyphosphonic acid (AEP) which is the native substrate proposed for phosphonate tailoring (Figure 42) . The K_m value for LicC with AEP was 0.318 ± 0.126 mM and k_{cat} of 0.722 ± 0.053 /s. The kinetic parameters for AEP are significantly lower than phosphocholine. It was calculated that Spn-LicC has a 34 fold greater preference for phosphocholine than AEP.

The second enzyme to be characterized was Ari-PngC. This enzyme was characterized with AEP, its native substrate and phosphocholine to determine substrate specificity of the homolog. When tested with AEP, this enzyme exhibits a classic Michaelis Menten fit (Figure 44), with with K_m value of 0.011 ± 0.001 mM and a k_{cat} of 2.72 ± 0.079/s. When this enzyme was tested with phosphocholine, which is a similar looking substrate to AEP, the data was inconclusive. The data set could be modelled to both a classic Michaelis- Menten and a substrate inhibition curve, however the kinetic parameters that are modelled to Michaelis Mentin curve are statistically more significant. The residual plots for Ari-PngC with phosphocholine show high variability. Since phosphocholine was used and we know that Spn-LicC is inhibited by calcium, this could be used to explain why it is difficult to determine whether there is a substrate inhibition curve or a Michaelis - Menten curve. Comparing the kinetic parameters obtained with both substrates, it can be seen that AEP is the preferred substrate of choice, with a 143-322 fold greater specificity.

From the results obtained in our assay, Spn-LicC had a k_{cat} of 1.52 /s with its native substrate phosphocholine, while Ari-PngC has a k_{cat} of 2.72 /s for its native substrate AEP. It can be seen that there is a higher turnover number with Ari-PngC from the phosphonate biosynthesis pathway than the choline pathway. Phosphonates are considered to be ancient molecules and this hypothesis is supported by their presence founded on the Murchison meteorite[40]. A theory to support the observation that Ari-PngC found from the phosphonate pathway is more active that Spn-LicC from choline pathway, can be attributed to the fact that Spn-LicC could have evolved from Ari-PngC. Although this hypothesis is just a hypothesis, more study into phosphonate compounds needs to be done, in order to show any evolutionary relationship between these enzymes.

The results of the bioinformatic and kinetic data together prove the hypothesis to be valid. After comparing the active sites of cytidylyltransferase genes, it was apparent that there were subtle differences in the residues surround the active site, and that would be most likely to accommodate different substates based on specificity. This was further confirmed by the kinetic data of Spn-LicC and Ari-PngC, where each enzyme has a preferred substrate. Spn-LicC greatly preferred phosphocholine as its substrate, where as Ari-PngC greatly prefers its substrate AEP. It is hypothesized that cytidylyltransferase genes are used to tailor macromolecules with different moieties, such as phosphocholine or small molecule phosphonates, which are then loaded onto cell wall surfaces. Many antibiotics available in the market target cell wall synthesis, but are broad spectrum and destroy microbes that may not be involved in pathogenicity. This increases the selective pressure for the pathogens that survive the antibiotic

attack, and allows them to replicate to become the dominant type of microbial population, hence causing a greater resistance. Since each of the opportunistic bacteria studied in this project possesses either phosphonate gene cluster, teichoic acid biosynthesis gene cluster or both, antibiotics can be designed to target these pathways. From the data presented in this project, it is evident that each cytidylyltransferase gene has greater specificity for their native substrate. Using this knowledge, narrow spectrum antibiotics can be designed to target specific cytidylyltransferase genes from either phosphonate or teichoic acid synthesis pathway. Inhibiting this cytidylyltransferase gene would stop macromolecules from being loaded onto cell wall surfaces, and may aid in the loss of virulence. This type of antibiotic would be beneficial because it will be very narrow spectrum, ensuring that no other microbe is being destroyed, hence removing any selective pressure to cause resistance.

5. Future Directions and Conclusion

Many questions were answered within this project, but many remain unknown. Additional experiments need to be performed in order to gain further insight into cytidylyltransferase and their function. For example, from enzyme-activity tests of each enzyme it was determined that they may have more than just a dimer function. Experiments such as dynamic light scattering (DLS) can be completed to confirm if these enzyme have more oligomeric forms. In addition to DLS, crystal structure of these enzymes need to be solved to determine how exactly the enzyme functions, and what within the active site allows for specificity of certain substrates more than others. Lastly, it is necessary to determine if the product from enzyme catalysis is what we predict it to be. This can be accomplished by HPLC and mass spectrometry.

In conclusion, there is a urgent need for new antimicrobials to overcome antimicrobial resistance. Phosphonates are a class of natural products that have a C-P bond that is more resistant to degradation. Phylogenic analyses of opportunistic bacteria have revealed a presence of genes necessary for phosphonate biosynthesis, as well as genes that are hypothesized to be involved in macromolecular tailoring. The genes involved in macromolecular tailoring such as PngC and PngD homolog resemble the genes involved in the tailoring of teichoic acids in *S. pneumoniae.* After further bioinformatic analysis it was revealed that Spn-LicC resembles PngC from *A. rimae* and *O. uli.* It was also revealed through kinetic analyses Spn-LicC has greater specificity for phosphocholine, but still has some specificity for AEP, and Ari-PngC has greater specificity for AEP but can still catalyze phosphocholine. If these findings are true, then this information can be used to design therapeutics that would target the *licC* locus,

which would disrupt the tailoring of macromolecules without harming the surrounding environment, and in turn lower any selective pressure on microbes from developing resistance.

6. References:

1. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett, Knight R, Gordon JI. (2007) The human microbiome project: exploring the microbial part of ourselves in a changing world. Nature. 449:804-810.

2. Blaser, MJ. (2014). MIssing microbes: how the overuse of antibiotics is fueling the modern plagues. Macmillan.

3. Government of Canada. Antimicrobial Resistance and Use in Canada. http://www.healthycanadians.gc.ca/drugs-products-medicaments-produits/antibiotic-resistance-antibiotique/antimicrobial-framework-cadre-antimicrobiens-eng.php. Accessed Oct 27, 2014.

4 Huttenhower C and Gevers D. (2012) Structure, function and diversity of the healthy human microbiome. *Nature.*486: 207–214.

5. Van Saene R, Fairclough S, Petros A. (1990) Broad- and narrow-spectrum antibiotics: a different approach. *ClinMicrobiol Infect.* 4:56-57.

6. Smil, V. (2000) Phosphorus in the environment: Natural flows and human interference. *Annu. Rev. Energy Environ. 25*, 53-88.

7. Morgan, E. (1990). Vogel's textbook of practical organic chemistry. 5th edn. *Endeavour*. doi:10.1016/0160-9327(90)90017-L

8. Ju, K. S., J. R. Doroghazi, and W. W. Metcalf. 2013. Genomics-enabled discovery of phosphonate natural products and their biosynthetic pathways. Journal of Industrial Microbiology and Biotechnology 24:24.

9. Metcalf, W. W., B. M. Griffin, R. M. Cicchillo, J. Gao, S. C. Janga, H. A. Cooke, B. T. Circello, B. S. Evans, W. Martens-Habbena, D. A. Stahl, and W. A. van der Donk. 2012. Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. Science 337:1104-1107.

10. Horiguchi, M., and M. Kandatsu. 1959. Isolation of 2-aminoethane phosphonic acid from rumen protozoa. Nature 184**:**901-902.

11. Yu, X.; Doroghazi, J. R.; Janga, S. C.; Zhang, J. K.; Circello, B.; Griffin, B. M.; et al. (2013) Diversity and abundance of phosphonate biosynthetic genes in nature. *PNAS. 110*, 20759- 20764.

12.Metcalf, W. W.; van der Donk, W. A.(2009) Biosynthesis of phosphonic and phsosphinic acid natural products. *Annu. Rev. Biochem. 78*, 65-94.

13. Hilderbrand, R. L., and T. O. Henderson. 1983. Phosphonic acids in nature, p. 5-30. *In* R. L. Hilderbrand (ed.), The Role of Phosphonates in Living Systems. CRC Press, Boca Raton.

14. Hori, T., and I. Arakawa. 1969. Isolation and characterization of new sphingolipids containing *N,N*-acylmethylphosphonic acid and *N*-aclaminoethylphosphonic acid from mussel, *Corbicula sandai*. Biochimica et Biophysica Acta 176**:**898-900.

15. Korn, E. D., D. G. Dearborn, and P. L. Wright. 1974. Lipophosphonoglycan of the plasma membrance of *Acanthamoeba castellanii*. Isolation from whole amoebae and identification of the water-soluble products of acid hydrolysis. Journal of Biological Chemistry 249:3335-3341.

16. Onderdonk, A. B., D. L. Kasper, R. L. Cisneros, and J. G. Bartlett. (1977) The capsular polysaccharide of *Bacteroides fragilis* as a virulence factor: comparison of the pathogenic potential of encapsulated and unencapsulated strains. The Journal of Infectious Diseases 136:82-89.

17. Vinogradov, E., E. E. Egbosimba, M. B. Perry, J. S. Lam, and C. W. Forsberg(2001)Structural analysis of the carbohydrate components of the outer membrane of the lipopolysaccharide-lacking cellulolytic ruminal bacterium *Fibrobacter succinogenes* S85. European Journal of Biochemistry 268**:**3566-3576.

18. Miceli, M. V., T. O. Henderson, and T. C. Myers. 1980. 2-aminoethylphosphonic acid metabolism during embryonic development of the planorbid snail *Helisoma*. Science 209**:**1245-1247.

19. Dyhrman, S. T., J. W. Ammerman, and B. A. S. Van Mooy. **(**2007) Microbes and the marine phosphorus cycle. Oceanography 20:110-116.

20. Clark, L.; Ingall, E. D.; Benner, R. Composition and cycling of marine organic phosphorus. *Limnol. Oceanogr.* 2001, *46*, 309-320.

21. Quin, L. D. The presence of compounds with a carbon-phosphorus bond in some marine invertebrates. *Biochem.* 1965*, 4*, 324-330.

22. Peck, S. C.; Gao, J.; van der Donk, W. A.(2012) Discovery and biosynthesis of phosphonate and phosphinate natural products. *Method. Enzymol. 516*, 101-123.

23. McGrath JW, Chin JP, Quinn JP (2013) Organophosphonates revealed: new insights into the microbial metabolism of ancient molecules. *Nat. Rev. Microbiol.* 11, 412-419.

24. Hooper LV, Midtvedt T, Gordon JI.(2002) How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu. Rev. Nutr. 22, 283-307.

25. Kharat, A. S., and Tomasz, A. (2006) Drastic reduction in the virulence of Streptococcus pneumoniae expressing type 2 capsular polysaccharide but lacking choline residues in the cell wall. *Mol. Microbiol.* 60, 93–107

26. Young, N. M., Foote, S. J., and Wakarchuk, W. W. (2013) Review of phosphocholine substituents on bacterial pathogen glycans: synthesis, structures and interactions with host proteins. *Mol. Immunol.* 56, 563–573

27. Weidenmaier, C., and Peschel, A. (2008) Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat. Rev. Microbiol.* 6, 276–287

28 Rock, C. O., Heath, R. J., Park, H. W., & Jackowski, S. (2001). The licC Gene of Streptococcus pneumoniae Encodes a CTP: Phosphocholine Cytidylyltransferase. *Journal of bacteriology*, *183*(16), 4927-4931.

29. Zhang, J. R., Idanpaan-Heikkila, I., Fischer, W., and Tuomanen, E. I. (1999) Pneumococcal licD2 gene is involved in phosphorylcholine metabolism. *Mol. Microbiol.* 31, 1477–1488

30. Zeng, X.-F., Ma, Y., Yang, L., Zhou, L., Xin, Y., Chang, L., Zhang, J.-R., and Hao, X. (2014) A C-terminal truncated mutation of licC attenuates the virulence of Streptococcus pneumoniae. *Res. Microbiol.* 165, 630–638

31. Kwak B-Y, Zhang Y-M, Yun M, Heath RJ, Rock CO, Jackowski S, Park H-W (2002). Structure and mechanism of CTP:phosphocholine cytidyltransferase (LicC) from Streptococcus pneumoniae. J. Biol. Chem. 277, 4343-4350.

32. Jackowski S and Fagone P. CTP:Phosphocholine Cytidylyltransferase: Paving the Way from Gene to Membrane (2005). The Journal of Biological Chemistry 280:853-856.

33. Protein Data Bank. Catalytic Mechanism of CTP:phosphocholine Cytidylyltransferase from Streptococcus pneumoniae (LicC). ID: IJYL

34. Shiga, K., & Shiga, T. (1972). The kinetic features of monomers and dimers in high- and low-temperature conformational states of D-amino acid oxidase. Biochimica et Biophysica Acta (BBA)-Protein Structure, 263(2), 294-303.

35. Kunio, Yagi., Sugiura, N., Ohama, H., & Ohishi, N. (1973). Structure and Function of D-Amino Acid Oxidase VI. Relation between the Quaternary Structure and the Catalytic Activity. Journal of biochemistry, 73(5), 909-914.

36. Ropp, P. A., and Traut, T. W. (1991) Purine nucleoside phosphorylase. Allosteric regulation of a dissociating enzyme. *J. Biol. Chem. 266*, 7682–7687.

37. LéJohn, H. B., McCrea, B. E., Suzuki, I., and Jackson, S. (1969) Association-dissociation reactions of mitochondrial isocitric dehydrogenase induced by protons and various ligands. *J. Biol. Chem. 244*, 2484–2493.

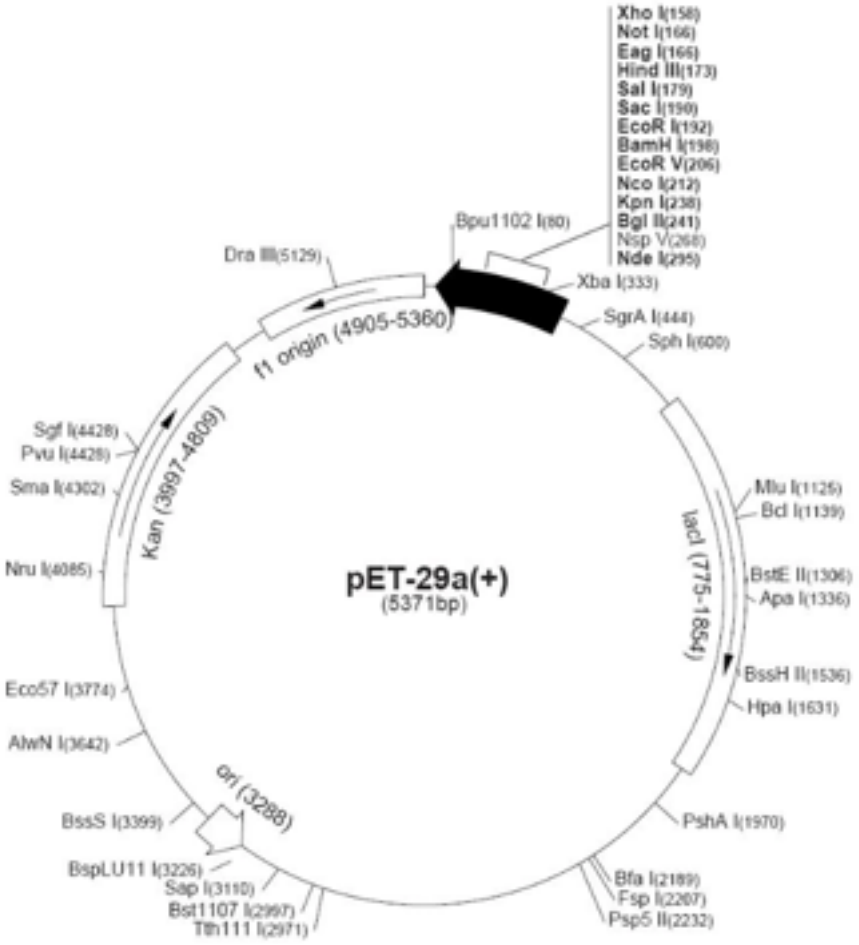
38. Maness, P., and Orengo, A. (1976) Activation of rat liver pyrimidine nucleoside monophosphate kinase. *Biochim. Biophys. Acta 429*, 182–190.

39. Yoshino, M., & Murakami, K. (2015). Analysis of the substrate inhibition of complete and partial types. *SpringerPlus*, *4*, 292.

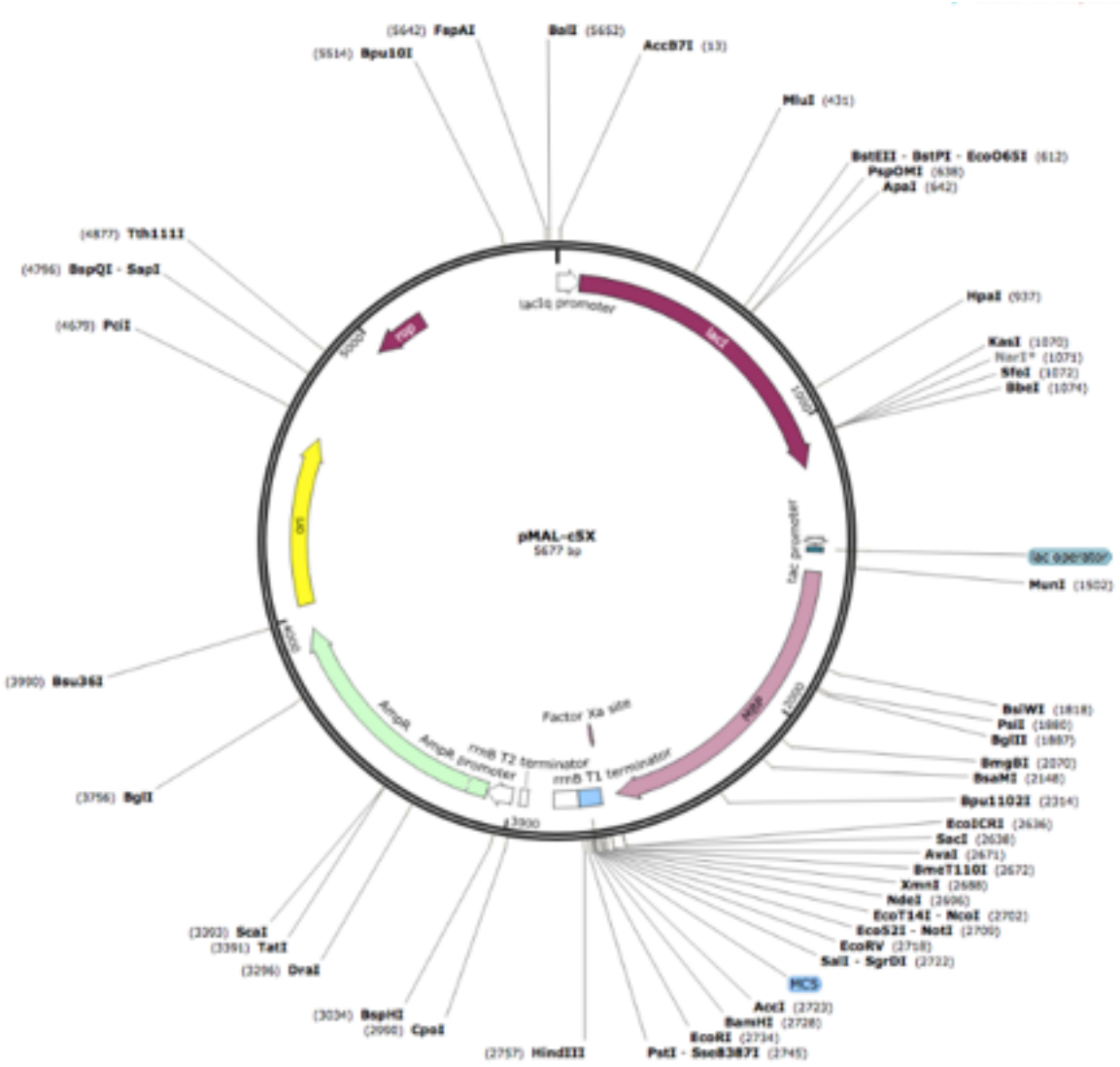
40. Cooper, G. W., Onwo, W. M., & Cronin, J. R. (1992). Alkyl phosphonic acids and sulfonic acids in the Murchison meteorite. *Geochimica et cosmochimica acta*, *56*(11), 4109-4115.

Appendix: Additional Information

A1. Vectors



pET29 vector



pMAL Vector

A2. Primer Sequences

Gene	Sequence	Tm (C)	Restriction Sites	Resistance Marker
LicC - A.rimae	Forward: 5'-CATATGCATCATCATCACCATCATAGCAGCGGTGTTGATCTGGGCACCGAAAATCTGTATTTTCAGAGC ATGGCTTGTGTAAAAGGTTCAAATG-3' Reverse: *5'-CTCGAG* CTAACTCCTCAAAAGATATACGTAG-3'	F: 68 R: 65	NdeI and XhoI	Kanamycin

Gene	Sequence	Tm (C)	Restriction Sites	Resistance Marker
LicC - O.uli (MBP)	Forward: 5'-GAAGGATTTC ATG GCA GCA GGT CTG GGT AC-3' Reverse: 5'-GCGGCCG TTA ACC TTC ATC AAT ATT GCT ACG-3'	F: 64 R: 64	NotI and XmnI	Carbenicillin
LicD - O. uli (MBP)	Forward: 5'-GAAGGATTTC ATG CGT GAA TAT GAT GCA GAA AC-3' Reverse: 5'-GCGGCCG TTA ATA GGC ACC AAA ATC CAG AC-3'	F: 64 R: 64	NotI and XmnI	Carbenicillin

A3. Construct Sequences

LicC - S.pneumoniae (pET29)

Restriction Sites: NdeI and XhoI

Amino Acid Sequence:
MGSSHHHHHHSSGLVPRGSHMKEIRVKAIILAAGLGTRLRPLTENTPKALVQVNQKPLI
EYQIEFLKEKGINDIIIIVGYLKEQFDYLKEKYGVRLVFNDKYADYNNFYSLYLVKEELA
NSYVIDADNYLFKNMFRNDLTRSTYFSVYREDCTNEWFLVYGDDYKVQDIIVDSKAGRI
LSGVSFWDAPTAEKIVSFIDKAYASGEFVDLYWDNMVKDNIKELDVYVEELEGNSIYEID
SVQDYRKLEEILKNEN

DNA Sequence:
CATATGGGTTCTTCTCACCACCACCACCACCACTCTTCTGGTCTGGTTCCGCGTGG
TTCTCACATGAAAGAAATCCGTGTTAAAGCGATCATCCTGGCGGCGGGTCTGGGTA
CCCGTCTGCGTCCGCTGACCGAAAACACCCCGAAAGCGCTGGTTCAGGTTAACCA
GAAACCGCTGATCGAATACCAGATCGAATTCCTGAAAGAAAAAGGTATCAACGACAT
CATCATCATCGTTGGTTACCTGAAAGAACAGTTCGACTACCTGAAAGAAAAATACGG
TGTTCGTCTGGTTTTCAACGACAAATACGCGGACTACAACAACTTCTACTCTCTGTA
CCTGGTTAAAGAAGAACTGGCGAACTCTTACGTTATCGACGCGGACAACTACCTGT
TCAAAAACATGTTCCGTAACGACCTGACCCGTTCTACCTACTTCTCTGTTTACCGTG
AAGACTGCACCAACGAATGGTTCCTGGTTTACGGTGACGACTACAAAGTTCAGGAC
ATCATCGTTGACTCTAAAGCGGGTCGTATCCTGTCTGGTGTTTCTTTCTGGGACGC
GCCGACCGCGGAAAAAATCGTTTCTTTCATCGACAAAGCGTACGCGTCTGGTGAAT
TCGTTGACCTGTACTGGGACAACATGGTTAAAGACAACATCAAAGAACTGGACGTTT
ACGTTGAAGAACTGGAAGGTAACTCTATCTACGAAATCGACTCTGTTCAGGACTACC
GTAAACTGGAAGAAATCCTGAAAAACGAAAACTAACTCGAG

LicC - A. rimae (pET29)

Restriction Sites: NdeI and XhoI

Amino Acid Sequence:
CATATGCATCATCATCACCATCATAGCAGCGGTGTTGATCTGGGCACCGAAAATCTG
TATTTTCAGAGCATGGCTTGTGTAAAAGGTTCAAATGCCGAGAAGACCAATGCGATT
ATAATGGCGGCGGGATTGGGAACGCGCATGGCACCTCTCACCAAAACCACACCTAA
GCCGTTGATTTCGGTCAACGGAACGCCGATGATTGAAACGGTCATCAATGCGCTGG
TTACGGCCGGCGTTGAGAGGATCTCTGTGGTTGTCGGGTATCTCAAAGAGCAGTTT
TGCTATCTTGAGGAGCGCTATCCGGCCGTAGTGCTTGTTGAGAACACGGAGTATTT
GGAGAAGAATAACATCTCATCAATATATGCCGCCGTTGACGTTCTTGAGCAGGGGG
CTACGTTTATTTGCGAGGCTGACCTGGTTATTTCAGATGAGCATATTTTCCAACCACG
GCCTTCTCGGTCGTGCTATTTTGGCCGTAAGTTTTCCGGCCATACGGGCGACTGGG
TGTTTGACCTTGACGATTCGGGAAAGATTGTCCGAATCGGCAAAGGCGGCAGCGAT
ACGTATGCCATGGTGGGACTGTCGTACTTCTCGGCACCGGATGCAAAGCGTTTGGC
ACGGTTTATGCATGATGCCTACAAAGAGACCGGCCACGAGCAGCTCTTTTGGGATG
ATGTGGTGAATAACCATATTGCCGAATTAGATCTTTCAATTCACCCCGTTGAGGCACA
GCAAATTGCGGAGCTTGATAGCGTTGCGGAATTGGCGGCGTTTGACCATGGCTACG
TATATCTTTTGAGGAGTTAG

DNA Sequence:
MACVKGSNAEKTNAIIMAAGLGTRMAPLTKTTPKPLISVNGTPMIETVINALVTAGVERI
SVVVGYLKEQFCYLEERYPAVVLVENTEYLEKNNISSIYAAVDVLEQGATFICEADLVIS
DEHIFQPRPSRSCYFGRKFSGHTGDWVFDLDDSGKIVRIGKGGSDTYAMVGLSYFSAP
DAKRLARFMHDAYKETGHEQLFWDDVVNNHIAELDLSIHPVEAQQIAELDSVAELAAFD
HGYVYLLRS

LicC - O. uli (pET29)/ (pMAL)

Restriction Sites (pET29): NdeI and XhoI
Restriction Sites (pMAL): NotI and XmnI

Amino Acid Sequence:
CATATGCATCATCATCACCATCATAGCAGCGGTGTTGATCTGGGCACCGAAAA
TCTGTATTTTCAGAGCATGGCAGCAGGTCTGGGTACACGTATGGCACCGCTG
ACCCAGATGACCCCGAAACCGCTGATTCGTGTTAATGGCACCCCGATGATTG
AAAGCGTTATTAATGCACTGGAAGCAGCCGGTGTGGCATGTATTTATGTTGTT
GTTGGTTATCTGAAAGATCAGTTCCGCTATCTGGAAGAACGTTATGGTCCGG
TTAAACTGATTGAAAATCCGGAATATCTGAGCAAAAACAACATCAGCAGCATT
TATGCAGCCATTGAAGTTCTGGAACAGGGTGATGCATTTATTTGTGAAGCAGA
TCTGGTTGTTAGCGAAAGCGGTATTTTTAGCGATCTGCCGAGCAAAAGCTGT
TATTTTGGTCGTATGACCGAAGGCTATACAGATGATTGGGCATTTGATCTGGA
TACCACCGGTCGTATTACCCGTGTTGGTAAAGGTGCAGTTAATAGCTATGCAA
TGGTTGGCATCAGCTTTTTCAAAGGTGGTGATGCAGCACATCTGGCACGTTG

TATTCGTAGCGCATATACCCAAGAAGGTCACGAACGTCTGTTTTGGGATGATA
TTGTGAACCAGCACATCGATGAATTCGATCTGCGTATTCGTCCGGTTCTGGA
AGGTCAGGTTGTGGAACTGGATACAATTGAAGAACTGGCAGCCTTTGATCCG
AGCTATAGCGATCAGCTGCGTAGCAATATTGATGAAGGTTAACTCGAG

DNA Sequence:
MIESVINALEAAGVACIYVVVGYLKDQFRYLEERYGPVKLIENPEYLSKNNISSIYAAIE
VLEQGDAFICEADLVVSESGIFSDLPSKSCYFGRMTEGYTDDWAFDLDTTGRITRVGK
GAVNSYAMVGISFFKGGDAAHLARCIRSAYTQEGHERLFWDDIVNQHIDEFDLRIRPVL
EGQVVELDTIEELAAFDPSYSDQLRSNIDEG

LicD - O.uli (pMAL)

Restriction Sites: NotI and XmnI

Amino Acid Sequence:
CATATGCATCATCATCACCATCATAGCAGCGGTGTTGATCTGGGCACCGAAAA
TCTGTATTTTCAGAGCATGCGTGAATATGATGCAGAAACCCTGCGTCATGTTC
AGCAGTGTGAACTGAAAATTCTGAAAGATGTTGCCCGTATTTGTGATGGTCA
GGGTCTGACCTATTTTGGTCTGGCAGGCACCGGTATTGGTGCAATTCGTCAT
AAAGGTTTTATTCCGTGGGATGATGATATTGATATTGGTATGCCTGCACGTGAT
CTGGAACAGCTGGTTAAAATCATTCGTGAAGAACATGCAGGCACCTATGATG
TTATTAACGCCGATATCGATAGCAATTATCCGCTGGCAACCACCCGTATTATGC
TGAAAGGCACCCAGTTTTGTGAAGAAACACTGAGCGAACTGCCGCTGGATC
TGGGTATTTTTCTGGATATGTATGCCTTTGATAACGTGGCCGATGATGAAAAT
GCATATCGTAAACAGGCATGGGATGCATGGTTTTGGGCACATATTCGTATTCT
GATTAGCGTTAGCCATCCGGTTATTCAGGTTCGTGGTTGGCGTGGTCTGCTG
CTGCGTTTTGCATGTGCGGGTGCACATGCCTTTGCACGTATTCTGCGTATTA
GTCCGGAATTTGCCTATCGTAAAGAACGTGAAGCACGTCGTCGTTATGCAAA
TGAAGCGACCAGCCGTATTGGTTATCTGTGTGATACCAATCGTTTTACCCAGA
CCTATGCCTGGGATGATATTAAACCGTTTCTGGACCTGGATTTCGAGGATATG
AAACTGCATTTTCCGCGTGAAATTGATGCAATGCTGCGTGAAATGTTTGGCG
ATTATATGCAGCTGCCTCCGGTTGAAAAACGCAAAAATCATTTTCCGGCACGT
CTGGATTTTGGTGCCTATTAACTCGAG

DNA Sequence:
MREYDAETLRHVQQCELKILKDVARICDGQGLTYFGLAGTGIGAIRHKGFIPWDDDIDIG
MPARDLEQLVKIIREEHAGTYDVINADIDSNYPLATTRIMLKGTQFCEETLSELPLDLGI
FLDMYAFDNVADDENAYRKQAWDAWFWAHIRILISVSHPVIQVRGWRGLLLRFACAGA
HAFARILRISPEFAYRKEREARRRYANEATSRIGYLCDTNRFTQTYAWDDIKPFLDLDFE
DMKLHFPREIDAMLREMFGDYMQLPPVEKRKNHFPARLDFGAY

A4. PCR Conditions

Step	Temperature (C)	Time (min:sec)
1	94	2:00
2	94	0:15
3	Gradient 50-65	0:10
4	72	1:10, return to step 2*30
5	72	5:00
6	4	Infinite hold

A5. PCR Conditions with a set temperature

Step	Temperature (C)	Time (min:sec)
1	94	2:00
2	94	0:15
3	Desired temperature (eg 65)	0:10
4	72	1:10, return to step 2*30
5	72	5:00
6	4	Infinite hold

A6. TAE Buffer Preparation

To make the 25X TAE stock solution:

- Add 141g of Tris base, 28.55 mL glacial acetic acid and 9.3g of EDTA to about 900mL of
  distilled water
- Adjust to 1L with additional distilled water

To make the 1X TAE buffer used in the gel electrophoresis apparatus:

- Add 40mL of the 10X TAE buffer stock to a glass container
- Fill the glass container up to 1L with distilled water

A7. Competent cell protocol

-Streak a glycerol freezer stock of *E.coli* DH5α or BL21 onto LB-agar plates and grow overnight at 37^{o}C

- A single colony was picked and used to inoculate 50mL of LB
- Culture allowed to grow at 37 ^{o}C until OD_{600} reached approximately 0.5
- The culture was transferred to 2 ice-cold sterile 50mL conical tubes which were left to sit for 10 minutes
- The cultures were then centrifuged at 4^{o}C for 10 min at 2700xg
- Supernatant was poured off and the cells were resuspended in 10 mL of 0.1mM ice-cold $CaCl_2$
- Samples were centrifuged once more using the previous two steps
- Supernatant was poured off and the pellets were resuspended in 2mL of 0.1mM $CaCl_2$ with 15% (w/w) glycerol
- The solution was dispensed into 150 uL aliquots
- The aliquots were flash frozen using liquid nitrogen and stored at -80^{o}C

A8. LB and LB-agar

Preparation of LB media:

- Add 2.5 g of LB broth to a glass jar
- Top up to 100mL with distilled water
- Autoclave

Preparation of LB-agar:

- Add 2.5g of LB broth to a glass jar
- Add 1.5g of agar
- Top up to 100mL with distilled water
- Autoclave

A9. Preparation of pET29 extracts for ligation

- Streak glycerol freezer stock of *pET29 sghf* onto LB-agar plates
- Incubate at 37^{o}C overnight
- Pick a single colony into 3 mL and allow to grow at 37^{o}C in LB media with kanamycin overnight
- Mini-prep the culture
- Perform a restriction digest of the resulting DNA sample and run on a gel
- Cut out the top pET29 band (about 5 kb)
- Perform a gel extract

A10. Phosphate Buffers for Purification

Cell Lysis: 50 mM NaPO4 + 300 mM NaCl pH 7.4

- Add 900 mL of Milli-Q water to beaker
- Add 3.87 mL of 1M Na2HPO4
- Add 1.13 mL of 1M NaH2PO4
- Add 17.53 g of NaCl
- Adjust pH to 7.4 with HCl
- Top up to 1L with Milli-Q water

Wash Buffer: 50 mM NaPO4, 20 mM imidazole pH 7.4

- Add 200 mL of Milli-Q water to beaker
- Add 0.97 mL of 1 M Na2HPO4
- Add 0.28 mL of 1M NaH2PO4
- Add 0.34 g of imidazole
- Adjust pH to 7.4 with HCl
- Top up to 250 mL with Milli-Q water

Elution Buffer: 50 mM NaPO4, 250 mM imidazole pH 7.4

- Add 200 mL of Milli-Q water to beaker
- Add 0.97 mL of 1 M Na2HPO4
- Add 0.28 mL of 1M NaH2PO4
- Add 4.25 g of imidazole
- Adjust pH to 7.4 with HCl
- Top up to 250 mL with Milli-Q water

Desalting Buffer: 50 mM NaPO4 pH 8

- Add 900 mL of Milli-Q water to beaker
- Add 3.87 mL of 1M Na2HPO4
- Add 1.13 mL of 1M NaH2PO4
- Adjust pH to 8
- Top up to 1L with Milli-Q water

A11. Tris Buffers for Purification (DTT+AmSO4)

Cell Lysis: 10 mM Tris, 100 mM NaCl, 2 mM DTT pH 7.5

- Add 80 mL of Milli-Q water to beaker
- Add 0.121 g Tris
- Add 0.584 g NaCl
- Add 0.031 g DTT
- Adjust pH with HCl to 7.5
- Top to 100 mL with Milli-Q water

Wash Buffer: 10 mM Tris, 0.02 M imidazole pH 7.5

- Add 80 mL of Milli-Q water to beaker
- Add 0.121 g Tris
- Add 0.136 g imidazole
- Adjust pH with HCl to 7.5

- Top to 100 mL with Milli-Q water

Elution Buffer: 10 mM Tris, 0.25 M imidazole pH 7.5

- Add 80 mL of Milli-Q water to beaker
- Add 0.121 g Tris
- Add 1.70 g imidazole
- Adjust pH with HCl to 7.5
- Top to 100 mL with Milli-Q water

Desalting Buffer: 10 mM Tris, 100 mM NaCl, 2 mM DTT and 25% (1M) AmSO4

- Add 80 mL of Milli-Q water to beaker
- Add 0.121 g Tris
- Add 0.584 g NaCl
- Add 0.031 g DTT
- Add 13.21 g of AmSO4
- Adjust pH with HCl to 7.5
- Top to 100 mL with Milli-Q water

A12. FPLC Buffers

Buffer A:

- Add 1.211g of tris to a beaker
- Add 400mL of Milli-Q water
- Adjust the pH with HCl to 7.5
- Top up to 500 mL with Milli-Q water

Buffer B:

- Add 0.6055g Tris to a beaker
- Add 14.39g of NaCl
- Add about 200 mL of Milli-Q water

- Adjust the pH to 7.5 using HCl
- Top up to 250 mL with Milli-Q water

A13. SDS PAGE Buffer

Preparation of SDS 10X Running Buffer:

- Add 144g glycine to a flask
- Add 30.2g Tris base
- Dissolve together in 800mL of distilled water
- Add 10g SDS and mix
- Top up to 1L with Milli-Q water

Preparation of SDS 1X Running Buffer:

- Add 100mL of the 10X stock SDS buffer to a 1L glass container
- Top up to 1L with Milli-Q water

A14. Bradford Reagent

To prepare 250mL of Bradford reagent:

- Add 25 mg of Coomassie blue to a beaker
- Add 12.5% mL of 95% ethanol
- Add 25 mL of phosphoric acid
- Top up to 250mL with Milli-Q water

A15. Bradford assay controls to create the standard curve

BSA (bovine serum albumin) = control protein at concentration of 2 mg/mL

- Vial 1- 0 ug BSA (reference for the spec)
- Vial 2- 0 ug BSA
- Vial 3- 10 ug BSA
- Vial 4- 20 ug BSA

- Vial 5- 30 ug BSA
- Vial 6- 40 ug BSA
- Sample vials contained 5 ng of sample
- All vials were topped up to 100 mL with sterile water
- 5mL of Bradford reagent was then added to each vial
- Let sit for about 5 minutes
- Take an absorbance reading at OD595 for each cuvette and create a standard curve

A16. Chelax Protocol

- Make 100 mM of Phosphocholine in 10 mL of Milli-Q water
- Make Chelax resin
- 2.68 g of chelax in 6.58 mL of Milli-Q water
- Add resin to column and allow to run through
- Add 10 mL of phosphocholine to the column, and collect the flow through
- Add 0.5 g of addition Chelax to the flow-through collect and stir gently for 1 hour
- After 1 hour, filter the sample and aliquot into small portions