I want to thank my supervisor Helga Ertesvåg, Marte Maråk and E-Ming Rau for their continued support throughout my thesis period, as well as Simon Blindheim, Vivien Jessica Klein, Lars Lauvstad Sættem and Joakim Tafjord for their proof reading efforts.

Summary

Azotobacter vinelandii has an uncommon ability to enter a dormant state called a cyst. The cell halts most of its metabolic activity and covers itself in two protective layers which improves its resistance to various levels of stress. These layers consist of polysaccharides among other components, many of which are not fully characterized. Two gene clusters containing genes related to polysaccharide production have been found to be upregulated during encystment, and it was speculated that these genes might take part in the encystment process. Two such genes are *Avin05390* and *Avin30120*.

The effect of the genes on encystment was tested by replacing parts of the genes with an antibiotic resistance marker to knock them out. A deletion and a repair plasmid were constructed for both genes with the intention of using a recombination process to inactivate and subsequently repair the genes. The deletion and subsequent repair of the *Avin05390* gene was successful, but sequencing of the utilized deletion plasmid suggested it contained mutations affecting the adjacent gene. Deletion of the *Avin30120* gene was unsuccessful.

Encystment was induced in the newly produced strains as well as the wild type strain. The cells were then dried before germination was induced. The strain with a deleted *A.vinelandii* gene demonstrated growth when cultivated during encystment, proving that the *Avin05390* gene is not essential. Testing encystment for the strains with a deleted and repaired 5390 gene, as well as the wild type, gave results with standard deviations deviated too much to make any conclusions.

Oppsummering

Azotobacter vinelandii har den uvanlige evnen å gå inn i et metabolsk nedregulert stadie kalt en cyste. Cellen stopper de fleste metabolske aktiviteter og dekker seg selv med to beskyttende lag som forbedrer cellens motstandsdyktighet mot forskjellige former av stress. De to lagenes innhold er ikke enda blitt fullt karakterisert, men de består av blant annet polysakkarider. To genklynger, hvis funksjon hovedsakelig er relatert til produksjon av polysakkarider, har blitt observert å være oppregulerte under cystedannelse, hvilket har ført til spekulasjoner om at disse genklyngene tar del i cystedannelsen. Dette har ført til spekulasjoner om at disse genklyngene tar del i cystedannelsen. To gener i disse klyngene er *Avin05390* og *Avin30120*.

Genenes effekt på cystedannelse ble testet ved å fjerne deler av DNA sekvensen for å deaktivere genene. Plasmider ble laget for å slette deler av og deretter å reparere de utvalgte genene via rekombinering. Dette var for å lage bakteriestamme med deaktivere og reaktiverte versjoner av genene. Bakteriestammer med deaktivert og reaktivert *Avin05390* gen ble laget, men sekvensering av plasmidet dannet for å deaktivere genet indikerer at det inneholder mutasjoner som vil påvirke et annet gen i klyngen. Deaktivering av *Avin30120* genet fungerte ikke.

Cystedannelse ble indusert for de produserte *A. vinelandii*-stammene, samt den orginale stammen. Cellene ble deretter tørket før de ble kultivert. *A.vinelandii*-stammen med et deaktivert *Avin05390*-gen overlevde tørking, noe som viser at genet ikke er essensielt for å danne cyster. Testing av cystedannelsen med de forskjellige stammene ga resultater med et for stort standardavvik til å kunne komme med en konklusjon.

Preface

This thesis is a continuation of a project from fall 2018 that aims to determine if the *Avin05390* and *Avin30120* genes affects the *A. vinelandii's* ability to form cysts. It was carried out from April to December 2019 as part of my MTech in Biotechnology studies at NTNU.

Table of Contents

List of Tables

List of Figures

Abbreviations

SSB	Single-strand binding protein
PCR	Polymerase Chain Reaction
KmR	Kanamycin resistance gene
ApR	Apramycin resistance gene
tetA	Tetracyclin resistance gene
tetR	Tetracyclin repressor gene
LB	Luria Bertani
BM	Burk's medium
RO water	water treated by reverse osmosis

Chapter 1

Introduction

1.1 *Azotobacter vinelandii*

A.vinelandii is a gram negative bacterium initially discovered by J.G. Lipman in Vineland, New Jersey in 1903 (Lipman, 1903). It belongs to the *Azotobacter* genus, which itself is part of the *Pseudonomadaceae* family. *A. vinelandii* is rod shaped in its vegetative state with the dimentions of 2x4 μm(Jensen, 1954), while it tends to become peanut shaped during cell division and circular during encystment. The cell is motile due to flagella covering the cell (H.L., 1975). One uncommon trait of the bacteria is that one cell contains up to 80 copies of its genome (Nagpal et al., 1989). *A. vinelandii* is aerobic and encodes several nitrogenases that allows the cell to fixate nitrogen in the air, converting it into ammonium. The nitrogenases are all dependent on a metal-cofactor, containing either molybdenum, vanadium or iron. If molybdenum is present it represses the two other nitrogenases, while vanadium represses the iron nitrogenase. This diatrophic ability makes it a suitable nitrogen fertilizer for use in agriculture. If other suitable nitrogen sources are available, for example ammonia, nitrate, aminoazids or urea, the cell will utilize these instead of fixing nitrogen in the air. Ammonium salts also repress the synthesis of all nitrogenases. Some strains of *A.vinelandii* also have the ability to produce alginate, a substance used in various industries for its gel-forming quality (Noar and Bruno-Bárcena, 2018). *Azotobacter* can utilize a large variety of carbon sources, for example various alcohols, organic acids and saccharides. *A.vinelandii* however, does not utilize tryptone and yeast extract, which means that the bacteria cannot grow on Luria broth (LB). It can however grow on Burk's medium (BM), a growth medium for diazotrophs. The medium contains molybdenum, allowing nitrogenase activity in *A.vinelandii*. The optimal incubation temperature of *A.vinelandii* is 30 °C.

1.1.1 Cyst formation

A. vinelandii cells have the ability to enter a resting stage called a cyst when growth conditions decline. During encystment, the cells will change and become covered by two layers of coating. The inner layer is called intine and the outer is called the exine. Cysts are metabolically dormant and will remain in this state until conditions have improved enough to allow vegetative growth again (H.L., 1975). This improves the cell's resistance to various stresses such as desiccation, ultraviolet(UV) radiation, gamma radiation and sonic radiation (Socolofsky and Wyss, 1962). Encystment also halts the growth of the cell, as well as the majority of its metabolic activity, until growth conditions have improved sufficiently. Upon germination, the cell then breaks out of its cyst layers and regains its vegetative state (H.L., 1975). Cyst formation has been found to be induced when either 1-butanol or β-hydroxybutyrate (BHB) are used as the sole carbon source. This is assumed to be due to these substrates being intermediates of the PHB synthesis, functioning as a signal to start encystment (Lin and Sadoff, 1968).

Encystment starts by changing the shape from rod to circular. The cell loses the flagella on the surface of the cells within the first four hours and finishes any ongoing cell divisions, which results in a lack of mobility. An alginate capsule layer then emerges, covering the cell. Between approximately 6 to 30 hours after cyst inductions, the cells excrete a membrane-like material on the surface which breaks off to form sheets. The sheets will produce layers covering the cell, resulting in the exine layer. 24 hours after the encystment process started, the cells starts to accumulate poly-β-hydroxybutyrate (PHB) in vesicles in the matrix of the cells (H.L., 1975). PHB is a polymer that peaks in concentration during cyst formation and close to 70 % of the dry weight of the cell will consist of PHB containing vesicles after cyst formation. It is assumed that the vesicles serve as polymer storages during encystment (Noar and Bruno-Bárcena, 2018). An apparently empty space was observed between the exine layer and the central body, which starts filling up and increases in size about 36 and 48 hours after encystment. Between the 3rd and the 5th day, the exine developes into a thicker, hard protective layer that covers the cell. The central body has shrunk, while the intine has become much larger and more visually defined. At this point, the encystment is complete (Wyss et al., 1961). Germination, which is the process of cells returning to a vegetative state, is activated when the cells are cultivated in favorable growth conditions on Burk's medium with glucose as a carbon source. The cell swells until it takes up the space of the intine, and the bacteria will eventually break the exine and emerge. germination ends with the bacteria regaining its flagella(H.L., 1975).

Chemical analysis found that the exine layer contains 32% carbohydrates, 28% proteins and 28% lipids, and the last 4% consists of ash, a mixture of calcium, magnesium and and phosphor. The intine layer contained 44% carbohydrates, 36% lipids, 9% proteins and 3% ash (Lin and Sadoff, 1969). Both layers consisted of saccharides such as glucose, mannose, xylose and rhamnose, while glucosamine and galactosamine were found exclusively in the exine (Lin and Sadoff, 1969). It has also been observed that 40% of the exine carbohydrate and 73% of the intine carbohydrate is made up of the uronic acids of alginate Page and Sadoff (1975). The composition of the remaining carbohydrates is not well known. Alginate consists of two uronic acids: mannuronic acid (M) and guluronic acid (G). Regarding the relative content of these acids, the exine seems to be richer in guluronic acid while the intine is richer in mannuronic acid (H.L., 1975). Alginate is initially synthesized as a homopolymeric mannuronan, which can be converted to guluronic acid by mannuronan C-5 epimerases. *A. vinelandii* secrete the seven calcium dependent mannuronan C-5 epimerases AlgE1-7 that are necessary to produce high G-content in the alignates. These epimerases alter the 5-carbon of the mannuronic acids stereochemically to produce the guluronic acids - a crucial enzyme activity - for encystment (Steigedal et al., 2008). Activity of the various epimerases introduces different alginate compositions (Ertesvåg, 2015). Enzyme activity has been found to be dependent on the calcium concentration, and the absence of calcium renders enzyme activity negligible (Larsen and Haug, 1971). The guluronic acid is essential for the gelling properties of alginate, and it has been speculated that G-rich exine layer is created by cross-linking of calcium. Therefore calcium could be vital for the gelation and formation of the exine layer (H.L., 1975). This is supported by the fact that ethylenediaminetetraacetic (EDTA), a substance that creates stable bonds with calcium, in combination with a tris buffer break up the exine layer of the cyst (Socolofsky and Wyss, 1961; Lin and Sadoff, 1969). An omission of calcium in the growth medium will therefore make encystment impossible (H.L., 1975).

1.2 Methods for genome editing used in this work

When studying the genes of an organism and their effects, genetic editing can be utilized to be change the qualities of the genes. This allows observation of the resulting changes in the cells properties. DNA sequences can be manipulated in a variety of ways. Bases of DNA sequences can be altered and sections can be removed or inserted. The process of such changes is called genetic editing, and can be utilized in the study of an organism to alter the properties/qualities of its genes. The editing can be done by transfering a DNA template of the intended

change into the genome with the help of a vector. A common example of such a vector is a plasmid, a cirular DNA that is transferred from bacteria to other cells. Plasmids can be constructed to contain the DNA template of the intended insertion. Their construction requires the production DNA fragments by PCR or cutting DNA with restriction enzymes and combining them by either ligation, TOPO or sequence and ligation independant (SLIC) cloning. Newly produced plasmids may be introduced into the cell by a process known as transformation, in which the bacterial cell transfers foreign DNA through the cell membrane. This is mediated by specialized proteins conserved among various species. Transformations happen irregularly in nature, but the cells transformation efficiency (known as competence) can be improved. Treating cells with metal ions, such as calcium and rubidium, has been proven to increase the chemical competence of the cell. It is believed that this is because of depolarization of the cell membrane, which reduces its repulsion of DNA molecules (Azif et al., 2017). Although the method was not used, the process of electroporation could also have been utilized instead of chemical transformation. Another way to introduce foreign DNA is *conjugation*, which is the transfer of DNA between bacterial cells. Conjugation of plasmids starts by one cell attacking another with a pilus-like entity, connecting the cell's matrixes. The plasmid is nicked on one of its strands, separated and the nicked strand is transferred through the pilus. Following this, the single stranded DNA is then made circular and a complementing strand is synthesised, making the plasmid doublestranded (Llosa et al., 2002). When the plasmids have been transferred into the cells, the target DNA segments in the genome needs to be replaced by the DNA template in the plasmid. The chosen method of introducing the plasmids into the genome is Homologous recombination.

1.2.1 Homologous recombination

Recombination is a process of cross-connecting DNA strands that can be utilized to incorporate foreign DNA into a cell's genome. Homologous recombination is a variation that requires the flanking ends of the template DNA to be homologous to a segment of the genome. The foreign DNA may also be circular DNA with one homologous segment, such as for example plasmids. This process is not fully understood (Kowalczykowski et al., 1994), but a suggested model of the mechanism of plasmid recombination is shown in Figure 1.1. The first step of the recombination process is to turn the double-stranded genome DNA strand into a single-strand DNA that a RecA protein can interact with. This step can happen in several ways depending on the mediating enzymes (e.g. exonucleases or helicases). Single strand binding protein (SSB) molecules then mediates the interaction between the RecA protein and the 3'-strand of the genome DNA. The protein-DNA complex identifies an area within the plasmid containing a homologous section and pairs the prepared 3'-strand with its complementary DNA segment. RecA can then pair the displaced strands with each other, creating a heteroduplex molecule joined together as a Holliday junction (Figure 1.1). The next step is branch migration, which involves moving the junction along the strand, and extending the cross bound segment of the heteroduplex. This is a function of the RecA protein, but can also be mediated by the bidirectional RuvAB complex and RecG helicase. The last step is the resolution of the Holliday junction, in which the RuvC enzyme symmetrically cleaves the junction before a DNA ligase joins the four endings to two separate strands. This can produce two separate results. Firstly, a "patched" resolution separates the plasmid completely from the bacterial genome, leaving a segment of one of its strands behind. Secondly, a "spliced" resolution binds the DNA ends across and incorporates the plasmid into the genome, flanked by the homologous DNA segments (Kowalczykowski et al., 1994).

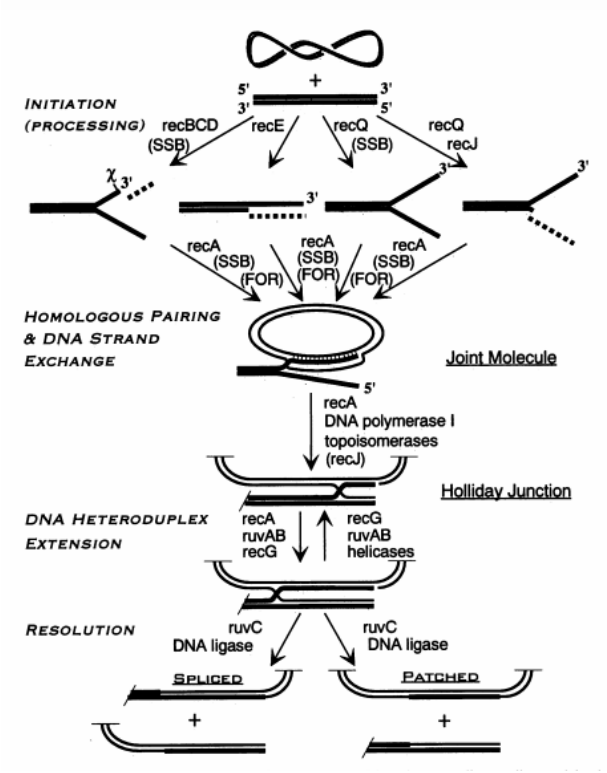


Figure 1.1: A biochemical model for genetic recombination in prokaryotes. The model splits into two separate resolutions of the cross bound intermediates at the last step (Kowalczykowski et al., 1994).

Homologous recombination has been utilized to modify the genome by addition and removal of DNA segments. Deletion can be performed by introducing a plasmid containing segments homologous to the segments flanking the intended deletion. Repairing a deleted gene requires a plasmid containing a wild type version of the gene. Figure 1.2 vizualises the recombination process that can be used to either repair or delete a gene. The first step is to introduce the plasmid into the genome by the spliced resolution of the recombination (step 1). This will result in the wild type and deleted version of the gene flanking the rest of the plasmid in the genome. The next step is a recombination of the homologous segments of the two versions of the genes, which removes the plasmid from the genome (step 2). Which of the two homologous sections flanking the two variations genes is used during the second recombination determines the gene variant that remain in the genome (Pedersen, 2019). To ensure that the recombination process gives the intended insertion, a plasmid has to be constructed properly. The following section will explain qualities required from the plasmids and the genes that will be responsible for them.

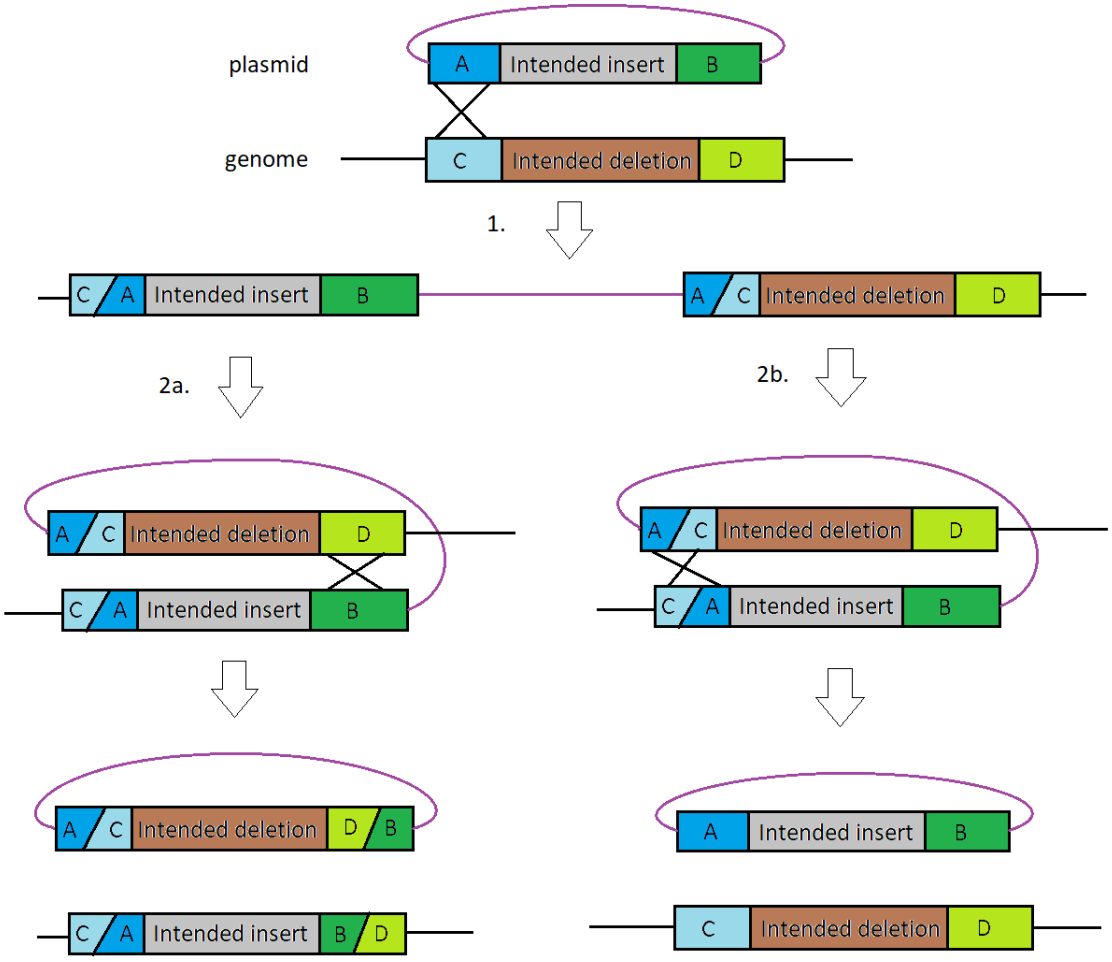


Figure 1.2: A recombination strategy to introduce mutation to genes. This strategy consists of two homologous recombinations as described in step 2b in Section 1.1. Segments A/C and B/D are homologous. Step 1 shows insertion of the plasmid into the genome by ”spliced” recombination. Step 2 can be resolved in two separate alternative recombinations that removes the plasmid from the genome. 2a results in a successful deletion while 2b results in the section reverting to its wild type variant. Copied from (Pedersen, 2019).

1.2.2 Plasmid functions

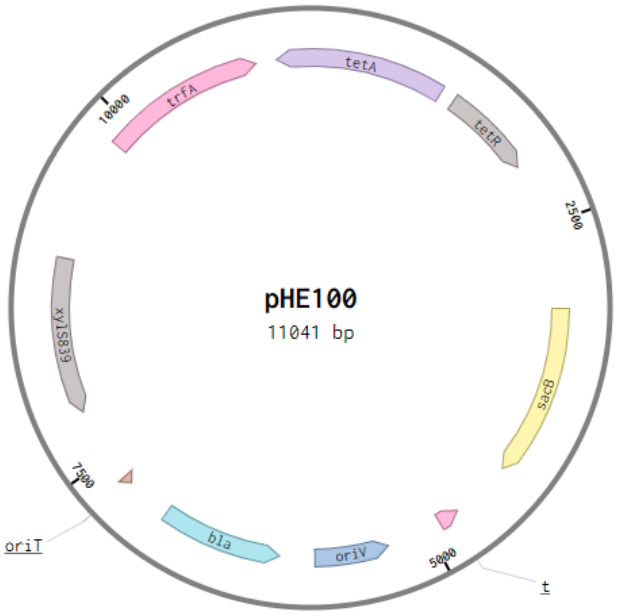


Figure 1.3: Plasmid pHE100.

Plasmids pHE100 and pHE257 are derivatives of pHE95(Gimmestad et al., 2009). They contain genes that allow for various methods of positive control of the cells. The properties of these plasmids are useful when performing the recombination process described in Figure 1.2 as it will allow for positive selection throughout the various steps. pHE100 and pHE257 contain genes that ensure resistance to two antibiotics, tetracycline and ampicillin. Different antibiotic genes could be used for positive selection of cells containing the gene throughout the steps of the process, either within the genome or a plasmid. The genes *tetR* and *tetA* are two genes that in concert will result in resistance to the tetracycline antibiotics. As the antibiotic can pass through the phospolipid bilayers (Argast and Beck, 1984), the genes must allow the cell to deal with the increasing concentration of the antibiotic. *tetA* encodes a transmembrane transport protein that exports the antibiotic out of the cell, protecting the cell from tetracycline (Marger and Saier, 1993). However, the protein is toxic in larger quantities, requiring the transcription of the gene to be regulated (Guay and Rothstein, 1993). tetR forms a dimer, which binds to the operator of both of the genes, repressing the transcription of *tetA*. The repressor protein disassociates from the operator as a result of interactions with a tetracycline-metal complex, allowing transcription. This implies that the presence of tetracycline upregulates transcription of the genes, and this interaction induces production of the tetA transport protein in the presence of tetracycline (Hillen and Berens, 1994). The *bla* gene produces a β-lactamase protein, which deactivates ampicillin. β-lactamase binds the antibiotics by noncovalent bindings between a hydroxyl group and the carbonyl oxygen of the antibiotics β-lactam group. The hydroxyl group acylates the β-lactam group by creating a covalent bond to the hydroxyl group. β-lactamase is released by hydrolysis, changing the antibiotic into a serine ester and deactivating it (Livermore, 1995).

sacB encodes a levan sucrase, which in the presence of sucrose produces levan. Many Gram negative bacteria, including *A.vinelandii*, will not survive the production of levan (Gimmestad et al., 2009). *sacB* can therefore be used for counterselection of cells containing the *SacB* gene by growing cells on media with 5% or more sucrose as the carbon source (Gay et al., 1999). This can be used for counterselection to ensure that the plasmid has been successfully removed from the cell and its genome after the second recombination. A transcription cascade has been constructed to control the replication of the plasmid based on the growth medium content. This replication mechanism of the plasmid is based on the R2K plasmid. Replication in the R2K plasmid is initiated when the expressed TrfA protein binds the origin of replication (oriV) a 17 bp repeats in the DNA segment (Perri and Helinski, 1993). Transcription of the *trfA* gene, regulated by a promoter called *Pm*, cannot bind RNA polymerase without the help of a positive regulator. This regulator is the XylS protein, and the specific protein utilized in plasmids pHE100 and pHE257 is the XylS839 protein. XylS is constitutively expressed in low levels, but is only active in the presence of a benzoate substrate, which works as a positive effector for XylS (Marqués et al., 1999). The coactivator utilized in this project was m-toluate. This dependency allows for choosing whether the plasmid is allowed to replicate, ensuring a quick way to eliminate the plasmid from the cell. The benzoate dependence of the transcription cascade controlling *tfrA* gene can be used in concert with antibiotic resistance to ensure that the plasmid DNA is inserted in the cell's genome.

1.3 The objective of the project

1.3.1 Background

Two gene clusters of *A. vinelandii* were found to be upregulated during cyst formation through microarray studies (Ertesvåg H., unpublished). The genes of the two clusters are detailed in tables 1.1 and 1.2, and the information was obtained from the genome sequence with the accession number CP001157 from genbank (Setubal et al., 2009). The genes of both clusters are generally shown to encode for proteins related to the biosynthesis of polysaccharides, for example glycosyltransferases, polysaccharide bioosynthesis proteins and epimerases. The *Avin30120* gene cluster also contained two genes related to the production of exopolysaccharides. One of the genes is *Avin29900*, which encodes a Wzz family protein related to the biosynthesis of branched heteropolysaccharides through the Wzx/Wzy-dependent pathway. The other gene is *Avin29930*, which encodes a ABC transporter protein related to the biosynthesis of linear heteropolysaccharides through the ABC transporter-dependent pathway (Rühmann et al., 2015). This all suggest that the gene cluster containing *Avin30120* produces various exopolysaccharides. The distance between the genes in the gene cluster containing *Avin30120* are sometimes so large that the genes in the cluster might not all be in an operon together. Table 1.2 displays the distance between the genes. Both clusters contain many genes related to polysaccharide synthesis, such as glycosyl transferases and polysaccharide biosynthesis proteins. As the cyst layers have a substantial polysaccharide content and the clusters is upregulated during encystment, the clusters could potentially be involved in biosynthesis of the polysaccharides required to produce the cyst layers. Instead of testing the effect of the entire clusters on encystment, one gene from each cluster was chosen as the focus of this thesis. These were chosen as they are both related to polysaccharide production. The *Avin05390* was chosen as it is believed to encode a biosynthesis protein producing a exopolysaccharide, which specifically has functions outside the cell. The *Avin30120* was chosen as it is believed to encode a glycosyl transferase, an enzyme which elongates glycan chains. Deletion of these genes might therefore affect a cell's ability to become cysts if the polysaccharides expressed by the gene are essential for encystment.

Table 1.1: Genes in *Avin05390* gene cluster. The table contains the gene codes and the suggested gene titles of the genes in the gene cluster that *Avin05390* is included in. The start and end coordinates of the gene is also included. The distance between a gene and the gene listed below is also described. A negative distance shows that the genes overlap, and the value describes the number of bases that the genes overlap with (Setubal et al., 2009).

gene code	suggested title	start coordinate	stop coordinate	distance to the next gene
Avin05280	cellobiose-phosphate cleavage protein	507149	506364	15
Avin05290	concerved hypothetical protein	508273	507134	142
Avin05300	concerved hypothetical protein	510355	508415	-3
Avin05310	transmembrane protein	511653	510352	-3
Avin05320	Polysaccharide deacetylase	512705	511650	10
Avin05330	Glycosyl transferase, family 2 protein	513866	512715	-7
Avin05340	Glycosyl transferase, family 2 protein	514824	513859	-3
Avin05350	O-antigen polymerase, Wzy protein	516224	514821	5
Avin05360	Glycosyl transferase, group 1	517341	516229	-27
Avin05370	lipopolysaccharide biosynthesis protein	518903	517314	3
Avin05380	Polysaccharide export protein	519910	518906	29
Avin05390	Exopolysaccharide biosynthesis protein	520610	519939	43
Avin05400	sugar transferase	521252	520653	125
Avin05410	NAD-dependent epimerase/dehydratase	522498	521377	

Table 1.2: Genes in *Avin30120* gene cluster. The table contains the gene codes and the suggested gene titles of the genes in the gene cluster that *Avin30120* is included in. The start and end coordinates of the genes are also included. The distance between a gene and the gene listed below is also described. A negative distance shows that the genes overlap, and the value describes the number of bases that the genes overlap with (Setubal et al., 2009).

gene code	gene product	start coordinate	stop coordinate	distance to the next gene
Avin29880	Multidrug efflux protein	3089800	3091179	2604
Avin29890	Low molecular weight phosphotyrosine protein phosphatase	3091958	3092404	33
Avin29900	Protein-tyrosine kinase wzz family protein	3092437	3094641	55
wbpO	polysaccharide biosynthesis protein	3094696	3095973	-3
wbpP	NAD-dependent epimerase protein	3095970	3097073	736
Avin29930	ABC transporter protein	3097809	3099650	-3
Avin29940	Phosphoribosyltransferase protein	3099647	3100612	13
Avin29950	concerved hypothetical protein	3100625	3101662	-3
Avin29960	UDP-glucose/GDP-mannose dehydrogenase protein	3101659	3102981	31
Avin29970	Glycosyl transferase family 2 protein	3103012	3103890	159
Avin29980	Glycosyl transferase family 2 protein	3104049	3105026	-3
Avin29990	Glycosyl transferase group 1 family protein	3105023	3106159	-13
Avin30000	Glycosyl transferase group 1 family protein	3106146	3107258	-3
Avin30010	concerved hypothetical protein	3107255	3108505	-3
Avin30020	Glycosyl transferase group 1 family protein	3108502	3109587	14
Avin30030	concerved hypothetical protein	3109601	3110821	-3
Avin30040	Glycosyl transferase group 1 family protein	3110818	3111942	0
Avin30050	Bacterial sugar transferase family protein	3111942	3112538	-3
Avin30060	Trimeric LpxA-like family protein	3112535	3113164	992
Avin30070	DegT/DnrJ/EryC1/StrS aminotransferase family protein	3114156	3115331	325
Avin30080	Polysaccharide biosynthesis	3115656	3117953	183
Avin30090	glycosyl transferase	3118136	3118816	172
Avin30100	glycosyl transferase	3118988	3119998	341
Avin30120	glycosyl transferase family 2 protein	3120339	3121325	474
Avin30130	Polysaccharide export protein	3121799	3122938	

1.3.2 The approach

The objective of this thesis was to determine whether there is a relationship between cyst formation and gene clusters containing the *Avin05390* or the *Avin30120* genes. The general approach was to knock out a gene within the respective clusters separately to test the effect on the cells ability to form cysts. Although this change will not adress the effect of every gene in the cluster, the knockout of one gene would be indicative of the rest, as genes in clusters tend to pertain to the same functions. *Avin05390* was chosen as the focus of its cluster, as it encodes for a protein that directly relates to exopolysaccharides. It is also one of the first genes in its operon, and knock out of this gene will likely hinder expression of the genes further downstream. For the other cluster, *Avin30120* was chosen as it encodes for a glycosyl transferases, an enzymes that is necessary to elongate polysaccharides.

The first step was to attempt to produce separate *A. vinelandii* mutants where a deletion was introduced into the genes. The deletion was then repaired to ensure that any changes to the plasmid is due to the deletion and not to a side effect of the recombination. Encystment was induced in all of the produced mutants strains as well as the wild type strain, before being dried in conditions that would require encystment to survive. The dried strains were cultivated both before and after drying. The resulting colony count was used to quantify the fraction of cells that survives this process. Comparing the strain's ability to form cysts would then show how the deletion affects encystment. This analysis can also help determine whether the recombination process itself had an effect by comparing the repaired strain to its wild type.

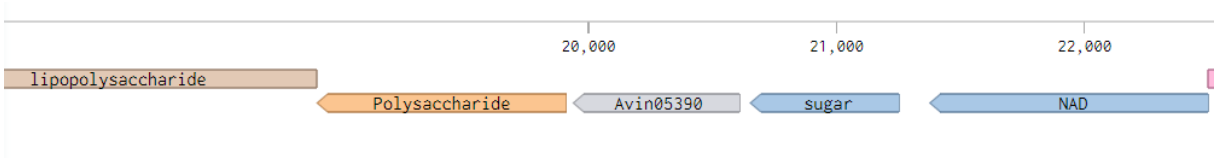


Figure 1.4: An overview of a segment of the gene cluster containing the gene Avin05390.

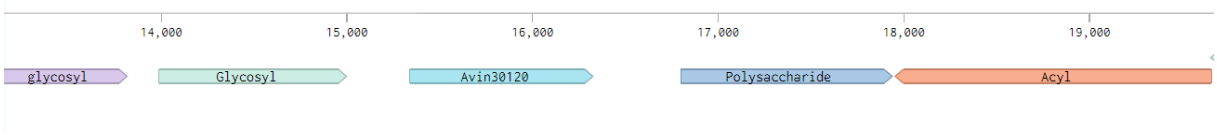


Figure 1.5: An overview of a segment of the gene cluster containing the gene Avin30120.

Chapter 2

Materials and methods

2.1 Mediums and solutions

Luria broth (LB) medium

10.0 g/l tryptone
5.0 g/l yeast extract
5.0 g/l NaCl
RO water
The solution was autoclaved.

If LA plates were produced, LB medium was prepared with 15 g/l agar and stirred thoroughly before autoclaving.

Psi medium

5.0 g/l Yeast extract
20.0 g/l tryptone
5.0 g/l $MgSO_4$

The solution was then pH adjusted to 7.6 by use of KOH before autoclaving.

TFB1

50 mM $MnCl_2 \cdot 4H_2O$
10 mM $CaCl_2 \cdot 2H_2O$
30 mM KAc
100 mM RbCl
15 % v/v glycerol
RO water

pH adjusted to 5.8 with diluted acetic acid.
Sterile filtrated.

TFB2
10 mM MOPS
10 mM RbCl
75 mM $CaCl_2 \cdot 2H_2O$
15 % v/v glycerol
RO water

pH adjusted to 6.5 with diluted acetic acid.
Sterile filtrated.

SOC medium
20.0 g/l tryptone
5.0 g/l yeast extract
5.08 g/l $MgCl_2$
0.5 g/l NaCl
2.5 mM KCl
3.6 g/l glucose
RO water

Sterile filtrated and stored in 1.5 ml Eppendorf tubes at -20 °C.

Burk's medium (BM)
5x Burk's salts:
1.0 g/l $MgSO_4 \cdot 7H_2O$
1.0 g/l NaCl
3.2 g/l K_2HPO_4
0.8 g/l KH_2PO_4
RO water

Separate 1000x stock solutions
15 mg/ml $FeSo_4$
50 mg/ml $CaCl_2$
2.1 mg/ml Na_2MoO_4

Dissolved separately in RO water, all stock solutions were sterile filtered and then stored at -20 °C.

Carbon source
Three separate, differently prepared carbon sources were utilized for this medium through the project.

Glucose
Glucose equivalent to 20 g/l concentration in the final solution was diluted with 50 ml RO water. This solution was then autoclaved.

Sucrose
Sucrose equivalent to 60 g/l concentration in the final solution was diluted with 50 ml RO water. This solution was then autoclaved.

Butanol
A 1-butanol volume equal to 0.2 % of the final concentration was sterile filtrated.

Final preparation
When using glucose or sucrose as the carbon source for the 500 ml medium, 100 ml 5x Burk's salts and 350 ml RO water were added and the solution was autoclaved. After the solution had cooled down sufficiently, 1/1000 of the stock solution of the $FeSo_4$ and $CaCl_2$ stock solutions, 1/100 of the Na_2MoO_4 stock solution and 50 ml of the carbon source solution were added to the 1x Burk's salts solution.

When 1-butanol was used as a carbon source, 399 ml RO water was mixed with 100 ml 5x Burk's salt and the mixture was autoclaved, 1 ml sterile filtered 1-butanol added as the carbon source, along with the salt stock solutions as described above.

If Burk's agar plates were produced, agar equal to 15 g/l of the final concentration was added to the Burk's salt/RO solution before autoclaving.

Agarose gel
400 ml TAE
0.05 μl Gelred or Gelgreen
3.2 g agarose

The solution (initially not containing the Gelred or Gelgreen) was dissolved by boiling it in the microwave, and was cooled before adding the Gelred or Gelgreen. The final solution was stored at 65 °C, and was poured into a form before use, resulting in the solution solidifying into a gel containing wells with decreasing temperatures.

2.2 Plasmids.

Table 2.1: Plasmids utilized throughout the project

Plasmid	Descriptions	Source
pCR-blunt II-TOPO	*LacZα*, *ccdB*, zeocin and kanamycin resistance KanR, pUC Ori	Invitrogen zero blunt TOPO PCR cloning kit kit supplied by termofisher
pIO100	AprR, upstream AmpR, *LacZ*, oriT, *rep*, *aph(3')-lla*, oripBR322, *per*	Onsager, not published)
pHE100	Low copy plasmid,derivative of pHE95. TetR, AprR, *sacB*. Replication initiation based on R2K plasmid, pm controlled *trfA* expression.	(Ertesvåg, not published), (Gimmestad et al., 2009)
pHE257	Low copy plasmid, derivative of pHE100. TetR, AprR, *sacB*. Replication initiation based on R2K plasmid, pm controlled *trfA* expression.	(Ertesvåg, not published)
pJLP003	Derivate of pCR-blunt II-TOPO with a 3.5kb PCR fragment del5390 insert containing AprR, KanR, ZeoR	(Pedersen, 2019)
pJLP004	Derivate of pCR-blunt II-TOPO with a 4kb PCR fragment del30120 insert containing AprR, KanR, ZeoR	(Pedersen, 2019)
pJLP005	Low copy plasmid, derivative of pHE257 with insert of del5390 from pJLP003, AprR, TetR	This project
pJLP006	Low copy plasmid, derivative of pHE257 with insert of del30120 from pJLP004, AprR, TetR	This project
pJLP007	Derivate of pCR-blunt II-TOPO with a 2.6 kb rep5390 inserted, KanR, ZeoR	This project
pJLP008	Derivate of pCR-blunt II-TOPO with a 3 kb with rep30120 insert, KanR, ZeoR	This project
pJLP009	Low copy plasmid, derivative of pHE100 with insert of rep5390 from pJLP007, TetR	This project
pJLP010	Low copy plasmid, derivative of pHE100 with insert of rep30120 from pJLP008, TetR	This project

2.3 Bacterial strain and plasmids

Table 2.2: Bacterial strains utilized throughout the project

bacteria	Descriptions	Source
E. coli DH5α	R6k origin of replication sequence, lysogenized with λpir	(Royo et al., 2005)
E. coli S.17.1	*RecA*, Rp4 transfer function	(Lorenzo et al., 1993)
A. vinelandii ATCC 12518	*A. vinelandii* wild type	Azotobacter vinelandii Lipman (ATCC® 12518™) supplied by ATCC
A. vinelandii ATCC 12518 Δ5390::AprR	*A. vinelandii* containing AprR gene and a deletion in the 5390 gene	This project
A. vinelandii ATCC 12518 rep	repaired derivate of the *A. vinelandii* ATCC 12518 Δ5390::AprR	This project

2.4 Primer sequences

Table 2.3: Primer sequences

Primers	Sequence
opp5390F	CAAGGCCCTTCGGGAATTG
opp5390R	GGGTTTGCCTAGGTAGCATGGTCATGGGCTGACTC
ned5390F	CCATGCTACCTAGGCAAACCCGCTGGGAAAGTC
ned5390R	GCTAGCCGCTGAAGAGCATCAC
sjekk5390F	GAGCTGCCCAACCTGATCAAC
sjekk5390R	GGGCGAATCCATCTCGATGTG
opp30120F	ATCCCGACATGTTCGGTTC
opp30120R	CGGCGTTCCTAGGTAAGGTGCCACGTTGTAGAC
ned30120F	CACCTTACCTAGGAACGCCGGAAAGTCAAAC
ned30120R	AAGGCTCCGGTGAGAATC
sjekk30120F	TCCACGACTCTCGAAGTTG
sjekk30120R	AGTTTCTCAGGGCGAGCAG
30120repF	CGAGCACCTGTCCGATATAGTC
30120repR	CCTGCGGATTCTCGATATAGCG
TOPOHE100F	CATCGCGCGCACCGTACGTCGCTCGGATCCACTAGTAACG
TOPOHE100R	GATGGGTTAAAAAGGATCCACGAATTGGGCCCTCTAGATG
TOPOHE100FAvrII	GACGTCACCGGTTCTAGATAGCTCGGATCCACTAGTAACG
amsekV	CGGACCTTGGAGTTGTCTCTG

2.5 PCR amplification

PCR was utilized to amplify DNA for two main uses: To insert the amplified sequence by cloning, or to verify mutants. PCR amplification of a DNA segment will produce and replicate copies of the given DNA sequence, and the amplification makes visualization and isolation by gel electrophoresis possible. The first step of PCR amplification is to produce a mixture containing the components required for PCR, which includes the template DNA, DNA polymerase, two primers, DNA nucleotides and a buffer. A detailed list of components used is shown in Table 2.4.

Table 2.4: PCR reaction mixture

Component	Quantity
5x NEB Q5 reaction buffer	5 μl
10 mM NEB deoxynukleotid solution mix	0,5 μl
10 μM forward primer	1.25 μl
10 μM reverse primer	1.25 μl
Template DNA	80-100 ng
NEB Q5 high fidelity DNA polymerase	0.25 μl
5x NEB Q5 high GC enhancer	5 μl
RO autoclaved water	Up to 25 μl

This mixture then had its temperature changed over time to create conditions that allow the different steps of the PCR process to occur in 35 cycles of repetition. The temperatures and durations of the different steps are described in Table 2.5. The first step of the PCR process is denaturation, which separates the two parallel strands from each other. The first denaturation is longer, as it denatures the double stranded template DNA containing the sequence intended to be amplified, while later denaturations only separates the produced DNA sequences from the template DNA. After denaturation, annealing is observed when the primers bind to the DNA template. The temperature of this step depends on the primers used. At this point the DNA polymerase starts replicating the DNA downstream of the primer's binding area at a slow rate. The last step changes the temperature to improve the replication rate of the polymerase. This step lasts 30 seconds per kb of the amplified DNA sequence. These three steps result in elongated DNA sequences starting at the sequence complementary to the primer.

Table 2.5: General thermocycling conditions used for a PCR reaction

Step	**reaction**	**Temp (C)**	**Time (seconds)**
1	Initial denaturation	98	30
2	30 cycles	98	10
3		variable	30
4		72	30 sek/kb
5	Final extention	72	120
6	Hold	4	infinite

Some of the produced DNA strands from the preceding PCR cycles can serve as template DNA for the next cycle of the amplification. If the polymerase binds to such a strand during the second cycle by the second primer's complementary sequence, the resulting sequence will be stretching between the sequences complementary to the primers. As the number of produced strands are amplified for every cycle, the primary product of the PCR is the DNA sequence streaching between the primers. The last cycle is extended to ensure that all extensions are completed. The temperature was lowered to 4 °C after the last cycle until the samples were picked up.

Procedure
The components of Table 2.4 were mixed on ice. The last component to the mixture is the DNA polymerase, added just before the PCR was started. The solution produced was then transferred to a PCR machine prepared with the cycle conditions shown in Table 2.5.
The PCR product was stored at -20 °C.

2.6 Benchling

When various enzyme restriction cutting or DNA ligation were planned, the online tool *Benchling* was used to view the plasmids and other DNA sequences in detail. Segments of the DNA sequences, e.g. genes, can be marked with a tag. Cut sizes along the DNA sequences for a large range of restriction enzymes from suppliers such as NEB and TermoFisher allow users to simulate cutting of DNA with the chosen enzymes. These features were utilized to plan cutting plasmids and ligation of fragments in order to produce new plasmids. Additionally, *Benchling* was used to align and compare sequences with each other. Some of the files returned gave detailed curve displaying the confidence of the individual bases of the sequencing. This was used during sequencing of PCR products to determine whether mutations had occurred.

2.7 Cutting with restriction enzymes

To separate DNA fragments from a plasmid, restriction enzymes were utilized to cut the plasmids at intended locations. Restriction enzymes can recognize specific DNA sequences and separate the DNA strand at the recognized DNA sequence. Different restriction enzymes cut at different known sequences, allowing isolation of specific parts of plasmids. The different fragments were separated by gel electrophoresis. This separation had primarily two functions: The first is to isolate a wanted DNA fragment for ligation with other fragments to produce new plasmids. The second function is to control whether the plasmid is the intended product of a ligation or one of the parent plasmids of the ligation.

Procedure

- A solution was made containing the following components:

  1 - 17 (μl) solution containing DNA
  0.5 (μl) of each restriction enzyme
  2 (μl) restriction enzyme buffer
  DNase free water up to 20 μl

- Restriction enzymes was the last component added. The volume of the added DNA solution was dependent on the concentration of the DNA.

- Different restriction enzymes requires different restriction enzyme buffers. The mixture was incubated at the incubation temperature best suited for the used restriction enzymes. The incubation lasted for between an hour and a day.

Several restriction enzymes would sometimes have to be utilized in the same reaction in order to get the desired DNA fragments. When these restriction enzymes share a preferred buffer and activation temperature, the enzymes were added simultaneously as described above. When the enzymes differ, the preferred buffer with the lower NaCl concentration and the corresponding restriction was utilized. After incubation, the NaCl concentration and/or the temperature was changed to the other enzyme's optimal incubation temperature to ensure the optimal environment for the added restriction enzymes reaction before continued incubation.

2.8 Separation of DNA fragments by gel electrophoresis

When PCR reactions or reaction enzyme cuttings were performed, one or more DNA fragments of various sizes were produced. These fragments were separated by agarose gel electrophoresis. Solutions containing the DNA fragments were placed in the wells of an agarose gel in a Tris-Acetate-EDTA (TAE) buffer before an electric current was introduced across the gel, moving the fragments through the gel. The agarose gel contains canals of various sizes that the molecules travel through, and the smaller DNA strands would be able to move through more canals than larger molecules. This would allow the smaller fragments to move more effectively and faster than larger fragments, separating the fragments. The gel contains either gelred or gelgreen, a flourescent molecule which binds DNA fragments and allows UV lights to visualize the fragments. DNA ladders were added in another well to allow for comparison to determine the unknown fragments sizes. The ladders used were made by cutting Lambda DNA with PstI and HindIII, and the ladders are displayed in appendix 5. A Biorad chemidoc XRS+ controlled with a Biorad Image software uses UV light to take a picture of the visualized DNA fragments.

Gel electrophoresis of cut plasmids were used to verify whether ligation produced the plasmids intended and to show the resulting fragments of a PCR process. The process could also be used to isolate DNA fragments from cut plasmids by excising the bands from the gel. Imaging with the Chemidoc was performed with a xcitaBlue conversion screen under the gel when a DNA fragment was isolated.

Procedure

- A solution with a final volume of 10 μl was made containing a DNA-binding buffer, and the DNA fragments was produced. Sterile RO water was used to increase the volume to its final volume. If the solution contained PCR products, 3 μl of the PCR product was added.
- A prepared gel was transferred to a container with a TAE buffer that covers the gel. The prepared solutions were transferred to the wells. A lid connected to a gel electrophoresis device was put on the container and the device was set to apply a constant electric current over the gel for the duration required for the buffer to traverse the gel.
- The gel was transferred to the chemidoc after the gel electrophoresis was finished, and a picture was taken of the gel when exposed to UV light.

  When planning to excise the gel containing the DNA, a xcitaBlue conversion screen was placed in the Chemidoc under the gel to protect the DNA from degradation due to UV lights.

  Excision of the DNA fragments from the gel was achieved by cutting the gel on a cutting board placed on the conversion screen. The gel was exposed to UV light in order to be able to see the DNA, and appropriate protective equipment was utilized.

2.9 Extraction of DNA fragments in agarose gel

To use DNA fragments for further work, the fragment would have to be isolated from the gel. This would require dissolving the gel containing the fragments. Later, contaminants such as the agarose gel residues, enzymes and more would have to be removed, effectively done by funneling the solution through a membranes that bind DNA only and not the contaminants. This was accomplished by using a Monarch gel extraction kit. All centrifugations in the following steps were preformed at 16,000 x g for 1 minute.

Procedure

- The gel was transferred to an eppendorf tube, and a Monarch dissolving buffer was added in a ratio of 4 ml per gram gel. The gel was kept at 53 °C while the gel dissolved.
- After the gel was dissolved, the solution was transferred to a spin column in a collection tube and centrifuged in order to move the solution through the filter of the spin column, binding the DNA in the filter. The flow through was discarded and the spin column was reinserted.
- 200 μl Monarch DNA wash buffer was added to the spin column and the column was centrifuged. This was repeated once more.
- The spin column was then transferred to an eppendorf tube.
- 16 μl Monarch DNA elution buffer was added to the spin column and the column was centrifuged. The buffer has a pH above 7, which releases the DNA from the column's membrane. The solution in the eppendorf tube contained purified DNA and was stored at -20 °C.

2.10 Measuring DNA concentration in a solution

Knowing the concentration of DNA was necessary to perform certain processes, e.g. during SLIC cloning and ligation, to achieve the right ratio of DNA fragments intended to be combined. To determine the DNA concentration of a given solution a Nanodrop One Microvolume UV-Vis spectrophotometer was used per the producers guideline. This equipment utilized the principle of the Beer-Lambert equation $c = A/\epsilon \cdot b$ to measure the solutions absorbance of UV: A is the difference in absorption of UV measured in the sample compared to a blank sample, b is the path length between the UV emitter and the sensor registering the adsorption and ϵ is the absorbidity coefficient determined by the substance measured. The absorptidity coefficient is specified ahead of measurement and the length is determined by the equipment itself. These measurements are performed with UV wavelengths of 260 nm, 230 nm and 280 nm. The equipment calculates the ratio of the measurements of 260 nm in relation to the other two, indicating the purity of the sample. A pure sample of double stranded DNA will have a A260/A230 ratio between 1.8-2.2 and a A260/A280 ratio below 1.8.

Procedure
The Nanodrop is set to measure double stranded DNA. 1 μl of the blank is loaded onto the dock and the Nanodrop is calibrated. The dock is cleaned and 1 μl of the sample is added. The sample is then measured.

2.11 TOPO cloning

The first step of plasmid construction is to insert the section intended for recombination to a plasmid. The procedure was done with aZero Blunt™ TOPO™ PCR Cloning Kit supplied by termofisher following the producers guidelines. This was achieved by performing a TOPO cloning, allowing for insertion of an isolated PCR products with blunt ends into a PCR-blunt-II-TOPO plasmid vector. This plasmid is cleaved with Topoisomerase I from the *Vaccinia* virus, resulting in a single cut at the end of a 5'-CCCTT DNA sequence. The DNA sequence intended to be inserted can then attack the blunt end of the vector, resulting in insertion. Whether the topoisomerase is involved in this step is unclear (Shuman, 1991). Plasmid vectors supplied by the kit had already been cut and were prepared to accept an insert.

The prepared plasmid could either accept the insert or bind its end, reverting to its original end. To avoid this, the cut sequence is only found in the middle of a LacZα fused with a *ccdB* gene. If the insert is successfully integrated,

the gene will be separated in the middle and transcription will be impossible. If the plasmid reverts to its original form, the *ccdB* gene is once more transcribed. The transcription of this gene will lead to lethality for *E.coli* cells. The plasmid also contains a kanamycin resistance gene, allowing for positive selection of cells with recombinant plasmids on LA-plates containing kanamycin.

Procedure

Before the TOPO cloning could be started, a PCR amplification was performed, amplifying a section within the bacterial genome of *A. vinelandii*. The PCR product was then isolated and purified by gel electrophoresis. The following solution was made and gently mixed in an eppendorf tube:

2 μl PCR product
0.5 μl salt solution
0.5 μl PCR II-blunt-TOPO
RO water up to 5 μl

The solution was incubated for 5 minutes at room temperature to allow the PCR product to be incorporated into the vector. The solution was then transformed into an *E.coli* DH5 α as described in Section 2.15.

2.12 Sequence and ligation-independent cloning

Sequence and ligation-indepentent cloning (SLIC) is a technique alongside TOPO cloning and ligation, utilized in this project to produce plasmids that allow insertion of PCR products into vectors with homologous sections of 15 bp or larger at the ends. Before SLIC could begin, the insert was produced by PCR and the vector was cut to have two ends homologous to the ends of the insert. SLIC is started by creating a solution containing the linearized vector, a PCR insert and T4 polymerase. T4 DNA polymerase will then activate its exonuclease ability to degrade the 3'-ends of both the insert and the vector. This creates overhangs on both the insert and vector that complement each other. This is followed by incubation on ice to allow annealing of the insert and vector. The two blunt ends will bind to each other and the T4 polymerase would patch up the holes in the newly annealed DNA strands (Jeong et al., 2012).

Procedure

- The vector was digested with restriction enzymes. The DNA was separated by gel electrophoresis and purified. The elution of the purification was done with 10 mM TrisCl at pH 8.0-8.5.
- The gene intended for insertion was amplified by PCR with a 15 bp or larger end homologous to the vector end. The results were controlled by gel electrophoresis and the product was purified. The elution during the purification was done with 10 mM TrisCl at pH 8.0-8.5.
- A solution was produced in an eppendorf tube. This solution had a total volume of 10 μl and contained 1X NEB buffer 2.1 and a given concentration of the vector and insert with a molar ratio between 1:2 and 1:4.
- 0.5 μl T4 polymerase and the solution was incubated for 2.5 minutes.
- The solution was put on ice for 10 minutes.
- The ligation was stored at -20 °C.
- The solution was transformed into cells as described in Section 2.15.

2.13 DNA ligation

Ligation is a method of binding two ends of linear DNA strands to each other, making them one combined strand. This requires the ends of the DNA strands to complement each other, either as blunt ends or sticky ends where the unpaired base overhangs can pair up. This can also bind the ends of a single DNA strand. The process can be used to produce new plasmids from two different linear DNA fragments isolated from cut plasmids.

- A solution was made in a eppendorf tube as described below: 2 μl 10x buffer for T4 DNA ligase

  Solutions containing DNA fragments (maximum 17.5 μl)
  Enough RO water to make the total volume 19.5 μl
- The DNA vector and insert was added in a molar ratio of 1:3.
- 0.5 μl T4 DNA ligase was added to to the eppendorf tube and the tube was directly transferred to a water bath at 37 °C.
- The ligation was stored at -20 °C.
- The solution was transformed into cells as described in Section 2.15.

2.14 Production of competent cell stocks

Some bacteria species have the ability to transfer extracellular plasmids by transferring the plasmid past the cell wall. *E.coli* can naturally do so infrequently, but can also made competent and be stored as cell stocks at -80 °C. This preparation makes the cell chemically potent, and involves exposing the cells to calcium chloride and mono- or divalent metal ions such as rubidium, potassium, manganese or potassium to improve the cell's ability to adsorb plasmids. The negative charge of both the cell membrane and foreign DNA can result in repulsion of the DNA, and the calcium chloride treatment is believed to bind to both and neutralize their charge. Treating the cells with rubidium along calcium has been shown to improve the yield of transformants compared to treatment with calcium alone (Azif et al., 2017).

Procedure

- *E.coli* cells from a stock culture was inoculated in 10 ml of LB medium, shaken at 37 °C over night.
- 2 ml of the culture was transferred to 200 ml Psi-medium in an Erlenmeyer flask, and the medium was inoculated at 37 °C.
- The cell concentration of the medium was continually controlled by transferring 1 ml to a cyvette and measuring the optical density (OD) with a spectrophotometer. OD is a measure of the light scattering of a light source passing through a solution. The spectrophotometer emits light through the medium and a photometric cell on the other side translates the incoming light into a signal. This allows determination of the medium's light scattering properties, which can be compared to the same medium without cells to determine the cell concentration. The wavelength of the light source was 600 nm.
- When the OD of the solution was found to be between 0.40-0.43, the flask containing the Psi was put on ice for 15 minutes. During this time the TFB1 and TFB2 were also put on ice.
- The medium was transferred to two separate sentrifuge cups and centrifuged at 4000 rpm and 4 °C for 5 minutes. The supernatant was removed.
- 80 ml TFB1 buffer was transferred to the centrifuge cup and the pellet was resuspended. The cups were sentrifuged with the same setting as before. The supernatant was removed.
- 6 ml of TFB2 buffer was added to the centrifuge cup and the pellet was resuspended.
- The solution was transferred to the cold eppendorf tubes, 100 μl per tube.
- The tubes containing the cells were placed in a container with liquid nitrogen for 5-10 seconds. The cells were stored at -80 °C.

2.15 Transformation

The cells prepared in Section 2.14 were later transformed with plasmids and cultivated to replicate the plasmids produced. When the plasmids are transformed into these cells, the qualities of the plasmids can augment the properties of the cells, e.g. giving the cells antibiotic resistance. Transformation was used in this project to introduce plasmids produced through ligation. The produced plasmid was replicated by growing the cells in a medium containing antibiotics that the plasmid encodes resistance to when necessary (Azif et al., 2017).

Procedure

- 100 μl of stock solution of chemically competent E.coli cells were put on ice to thaw.
- When the cells were thawed, 10 μl of ligated plasmids were added to the tube and the tube was incubated on ice for 30 minutes.
- After incubation, the cell stock was transferred to a water bath set at 37 °C for 45 seconds and then back on ice for 2 minutes.
- 900 μl SOC medium preheated to 37 °C was transferred to the eppendorf tube containing cells. The closed tubes were incubated shaking at 37 °C for an hour.
- Agar plates were made and prepared with the relevant medium and antibiotics. The plates were tested with cells without antibiotic resistance to test the effect of the antibiotics.
- 50 and 100 μl of the cell solution were plated out. The rest of the solution was centrifuged, 700 μl of the solution was pipetted out, and the cell pellet was resuspended in the remaining 150 ml and plated out on the last plate.
- The plates were incubated at 37 °C over night.

2.16 Storing bacterial strains

The newly produced bacterial strains or bacterial strains containing newly made plasmids was stored for later use by storing a mixture of 400 μl 60 % glycerol and 800 μl of the culture containing the bacterial strain at -80 °C.

2.17 Isolation of plasmid DNA

Plasmids have to be controlled when produced to ensure that the process was successful. This could either be done by PCR, or by restriction enzyme cutting and examining the resulting fragments. During this project, the produced plasmids were checked with restriction enzymes. This would require the plasmids to be isolated from the cells they are contained in. A Monarch plasmid miniprep kit (New England Biolabs) was used to lyse the cells and purify the solution from cell debris. A description of how the kit was utilized to isolate the plasmids is presented below.

Procedure

- An overnight culture was transferred an eppendorf tube and centrifuged to pellet the cells. The supernatant was removed.
- 1.5 ml was pelleted for high copy plasmids.
- 5 ml was pelleted for low copy plasmids.

- Centrifugations in this procedure were preformed at 16,00 x g for 1 minute unless otherwise specified.
- The cell pellet was resuspended by 200 μl Monarch plasmid resuspension buffer B1.
- 200 μl Monarch plasmid lysis buffer B2 was added to the solution to lyse the cells. The solution is alkaline and increases the pH to a point at which the cell wall breaks down. The tube was closed and carefully inverted until the solution had a uniform purple color. At that point the solution was incubated at room temperature for 1 minutes.
- 400 μl Monarch plasmid neutralization buffer B3 was added to the solution, which reduced the pH and neutralized the lysate activity. This buffer also contains RNase A that would break down RNA. The tube was inverted carefully until the solution became uniformly yellow. The mixture was incubated in room temperature for 2 minutes and then centrifuged for 5 minutes.
- The supernatant was transferred without dislodging the pellet to a spin column in a collection tube. The columns were centrifuged to move the solution through the filter of the spin column to bind the DNA. The flow through was discarded.
- 200 μl Monarch plasmid wash buffer 1 was added to the spin column and the column was centrifuged.
- 400 μl Monarch plasmid wash buffer 2 was added to the spin column and the column was centrifuged.
- The spin column was transferred to a clean eppendorf tube. 50 μl Monarch plasmid elution buffer was added to the column and the column was centrifuged. This results in a solution containing plasmids.
- The solution was stored at -20 °C.

2.18 DNA sequencing

After PCR has been performed and a DNA sequence is produced, PCR product must be sequenced to ensure that the product does not contain a mutation. This was done after the DNA sequence was inserted into a plasmid, by either SLIC or TOPO cloning. The sequencing was done by the company Eurofins GATC Biotech using the Sanger sequencing method. Sanger sequencing is performed by using DNA-polymerase reactions to elongate a primer to create a DNA sequence complementing the target sequence. The DNA polymerase process consists of cycles of annealings, elongations and denaturations as in the PCR process described in Section 2.5, but was performed with only one primer and therefore did not result in amplification of spesific segments. Another important difference to the PCR process is the addition of an altered variation of one of the nucleic acids. This altered version of the nucleic acids have no hydroxide group on the 3-carbon of the acid, and will consequently not allow another nucleic acid to bind to it. When the elongation of the primers during PCR utilizes one of these altered acids, the elongation will stop as it will no longer be possible to add a nucleic acid on the end. This will result in DNA sequences of different lengths that all end on the same altered nucleic acids. This process is repeated with added altered nucleic acids of all four acids separately. Gel electrophoresis is used to separate and visualize the four resulting PCR products. As the length of the product is determined by the sequence, the visualized DNA fragments can be used to determine the DNA sequences (Sanger et al., 1977).

Procedure A mixture was produced for every primer used for sequencing as described below:
400-500 ng DNA
5 μl 5 pmol/μl primer
RO water up to 10 μl

The produced mixtures were sent to Eurofins GATC Biotech for sequencing. The results were returned in computer files. These files contained the resulting sequence as a quantitative measure of signals suggesting a base in each spot of the sequence. The sequences were vizualized with Benchling and analyzed.

2.19 Conjugation and recombination

The method of introducing a mutation to the *A. vinelandii* cells in this work was to transfer a plasmid to the bacteria by conjugation before allowing and selecting for recombination. This permits transfer of a DNA segment. The conjugation and recombination is described further in Section 1.2. Utilizing the plasmid's anitibiotic resistance gene, its dependence on m-toluate for replication and the toxic effect of levan production combined with sucrose on *A. vinelandii* allows for selection throughout the process.

Procedure

- Precultures of *A. vinelandii* and *E. coli* containing the plasmid were prepared. *A. vinelandii* was cultivated for four days. *E. coli* was cultivated overnight.
- 100 μl of the precultures were transferred to new cultures and these were cultivated. *A. vinelandii* was cultivated overnight. *E. coli* was cultivated for 2 hours.
- 4-6 ml of each culture were transferred to, pelleted and had its supernatant removed from the same eppendorf tube. The pellet was resuspended in 100 μl sterile RO water and plated out on a LA plate containing m-toluate and left to incubate overnight. This is the step in which the plasmids are transferred to the *A. vinelandii* from the *E. coli* by conjugation.
- The cells on the plates were transferred to an eppendorf tube and resuspended in BM. A series of dilutions were created and all solutions were plated out on Burk's agar plates, with antibiotics that the plasmid contains a resistance gene against and m-toluate. As Burk's medium contains no nitrogen source, it ensures that only *A. vinelandii* containing the plasmid can grow.
- Four colonies growing on the Burk's plates were transferred to separate 10 ml Burk's medium in 125 mj flasks containing the relevant antibiotic. The reason four parallels were used instead of just one is to ensure that at least one resulted in growth. In this step, only *A. vinelandii* that have incorporated the plasmid into its genome through recombination can grow, as the lack of m-toluate will not allow the plasmid to replicate. The four parallels were carried out through the next step.
- 100 μl of the culture was transferred to 10 ml of fresh Burk's medium containing antibiotic and incubated. This step was repeated 4 times. This was to allow selection for the genomic copy of the approximately 80 copies in the cells that contain the antibiotic resistance genes. As this is a step that one might have to restart from, frozen bacteria stocks were made to allow this.
- 100 μl of the last culture was transferred to 10 ml of fresh Burk's medium without antibiotics and incubated. This allows the plasmid to be removed from the bacteria by a second recombination process.
- 50 μl of the solution was transferred to Burk's agar plates with 6 % sucrose as its only carbon source and incubated. Due to the *sacB* gene in the plasmids, any bacteria that has not removed the plasmid from its genome will transcribe the levansucrase enzyme, which produces levan in the presence of sucrose. As the production of levan results in the death of *A. vinelandii* cells, this will kill the cells still containing the plasmid. When the inserted DNA sequence contained an antibiotic resistance gene, the relevant antibiotic

was added to the plates. This allowed selection for cells that had removed the plasmid while retaining the insert containing the resistance gene.

- When the DNA fragment that targeted for removal contained an antibiotic resistance gene that the insert did not, a negative selection step was required. Colonies isolated from the sucrose plates were transferred to two separate Burk's plates - one with and one without the antibiotic relevant to the plasmid. Colonies that could not survive the antibiotic plates were utilized in the next step. This allowed selection of colonies that had successfully removed the plasmid containing the resistance gene.
- Colonies from the preceding step were tested by utilizing colony PCR and inspecting the product through gel electrophoresis to determine whether the recombination had been successful. As the replaced DNA sequence and its replacement had different sizes, a PCR process would be able to determine which segments were contained in the a colony's genome copies. The process was successful if only the insert was amplified. If both DNA segments were still present, the selection was not completely successful and the process had to be restarted at the selection step.

2.20 Cyst induction and testing

To test the cell's ability to form a cyst after deletion of genes, cyst formation has to be manually induced. *A. vinelandii* grown with 1-butanol, a precursor molecule to the lipids stored in globules as cysts, has been shown to result in cyst formation (Stevenson and Socolofsky, 1966). As *A. vinelandii* halts metabolic processes during encystment, some of the cysts would be able to survive over time without nutrition. The cells were thus dried for a week without nutrition. Comparing the cell counts of bacteria before and after drying allows for quantification of surviving cells. The following outlines the procedure for testing the cells ability to form cysts (Gimmestad et al., 2009).

- The *A. vinelandii* strain was cultivated in 10 ml of Burk's medium.
- 100 μl of the culture was transferred to a Burk's agar plate utilizing 0.2 % 1-butanol as a carbon source and cultivated for 120 hours in order to induce cyst formation.
- Cells from the plates were transferred and resuspended in 100 μl 1x burks buffer. 10 μl of the cell suspension was transferred to an eppendorf tube. This tube was left in a sterile plate casing with its lid open at 30 °C for a week.

  A dilution series of the cell suspension was produced and 50 μl of some of the dilutions were put on BM plates. The resulting cell cultures were counted.
- After a week, the cells in the eppendorf tube were resuspended in 100 μl and a dilution series was produced. 50 μl of the dilutions were plated out on plates of Burk's medium. The emerging colonies were counted.

As the results of this process provides the basis of the project's discussion and conclusion, the process was run with several parallels. As such, 3 biological parallels were split into 2 technical parallels to allow for more inference. Biological parallels were made at the start of the process with several cultivations of the bacteria in separate containers of Burk's medium. The technical parallels were introduced by plating out the liquid cultures on two separate plates with 1-butanol, continuing in parallel throughout the process. Three concentrations of the dilution series from both before and after the cells being dried were plated out for testing. The cell colonies that developed were were counted before and after drying to quantify surviving cells.

Chapter 3

Results

3.1 Plasmid construction

To test the effects of *Avin05390* and *Avin30120* on encystment, the genes must be knocked out by deletion. To ensure that any changes in functionality is due to the deletion, and not due to an unintended change caused by the recombination process itself, the gene is required to be repaired as well. To be able to delete and repair the *Avin05390* and *Avin30120* genes, a separate plasmid had to be made for all four processes. These plasmids must contain a homologous segments to allow the recombination to occur. The deletion plasmids require sections homologous to the segment intended to be altered, either lacking the segment intended for removal or having it substituted by a marker used for selection. Such a DNA segment was produced by amplifying the segments flanking the intended deletion with PCR before combining them by another PCR process. The process of producing this DNA template is described in figure 3.1.

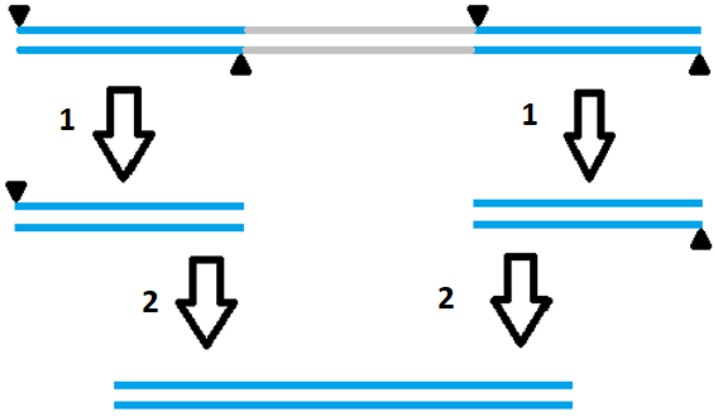


Figure 3.1: Deletion construction strategy. The first step was to amplify the segments flanking the segments intended to be deleted. Arrows in the figure indicate segments complemented by the employed primers. The second step was to join the two fragments as one DNA fragments by utilizing PCR. The PCR process was performed with the two fragments and two primers complementing the area indicated by arrows on the separate fragments. The elongation bound the two fragments, producing a DNA template containing a deletion.

An antibiotic resistance gene can be inserted as a marker by ligation at the deletion location to simplify the selection during the second recombination stage. In a previous project, such homologous segments with a *AprR* marker were produced and inserted into TOPO plasmids. The resulting plasmids are denoted pJLP003 and pJLP004,

respectively, for Avin05390 and Avin30120 deletion (Pedersen, 2019). The homologous section of the repair DNA must be a copy of the original gene and its flanking DNA sections, and is to be produced by PCR amplifications. To be able to insert it into a plasmid, the DNA is inserted into a vector by TOPO cloning and transferred to another vector. The vector in question was derived from pHE100 and pHE257. These plasmids were constructed by binding DNA fragments together by ligation, TOPO cloning or SLIC, to produce circular DNA. These fragments are either PCR products or fragments from plasmids cut with restriction enzymes isolated by gel electrophoresis. The PCR products used throughout the construction of plasmids can potentially have contained mutations. As the mutations inserted by recombinations can have very notable effects on the cells, sequencing the PCR products inserted into the plasmid is necessary. The sequencings discussed in the following sequences are displayed in Appendix 5.

3.1.1 Construction of a plasmid for use in deletion of *Avin05390*

The homologous segment containing the deletion intended for *Avin05390* was produced by PCR amplification from plasmid pJLP003. The pJLP003 is a TOPO plasmid containing a deleted variation of the gene *Avin05390*, flanked with the required homologous segment and containing a *aprR* gene. Subsequently, the PCR product would have to be inserted into a pHE257 derived vector (Figure 3.2).

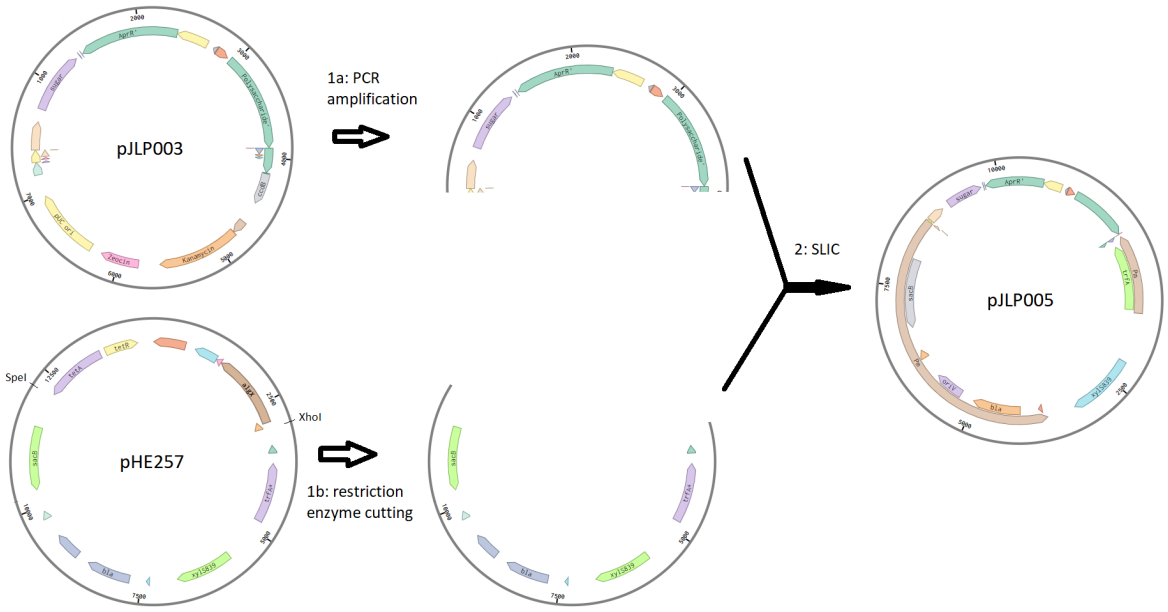
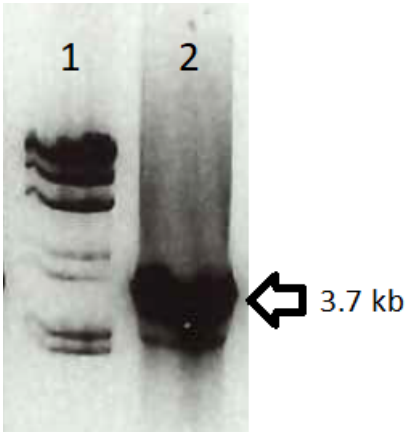


Figure 3.2: Construction strategy of the *Avino5390* deletion plasmid. 1a is the isolation of the insert from pJLP003 by PCR amplification. 1b is the isolation of the vector from pHE257 by restriction enzyme cutting with SpeI and XhoI. The last step is inserting the PCR product into the vector by SLIC.

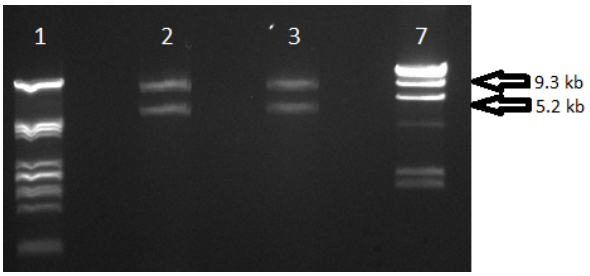
The inserted PCR product was made using pJLP003 as the template DNA and the primers TOPOHE100F and TOPOHE100R, producing a 3.7 kb DNA fragment (Figure 3.3a). The vector was made by cutting pHE257 with SpeI and XhoI, resulting in two fragments of sizes 9.3 and 5.2 kb (Figure 3.3b). The 9.3 kb DNA fragment was isolated for cloning.

The PCR fragment was inserted into the vector by SLIC cloning and the product was transformed into *E.coli S.17.1*. Apramycin was used during this process for positive selection. The plasmid isolated from two of the

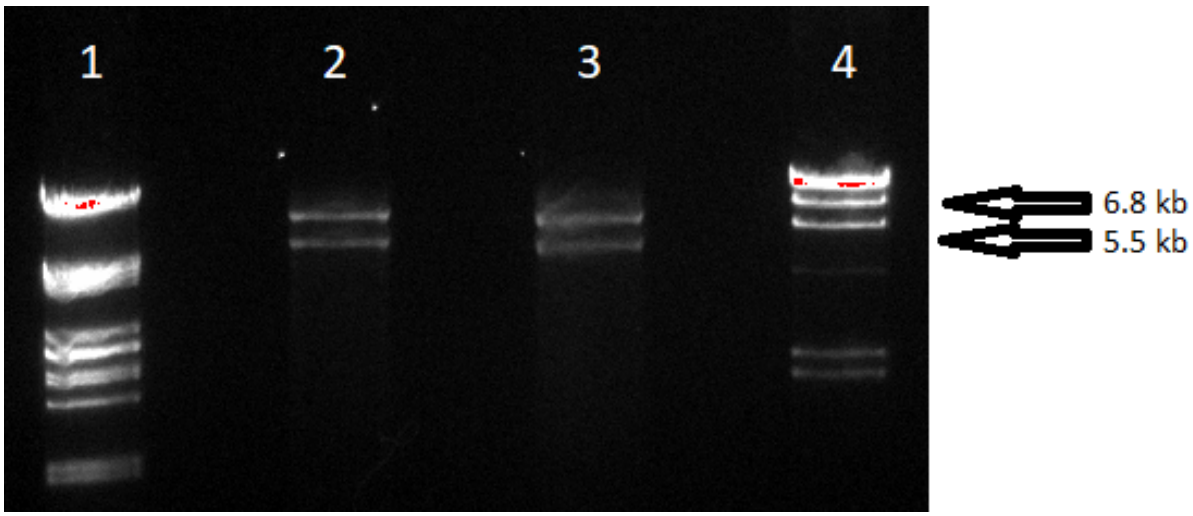
resulting colonies were tested by cutting them using AfeI and ApaLI. Successful ligation should result in two fragments of sizes 6.8 and 5.5 kb, which was the observed result for both transformants (Figure 3.3c).



(a) The pJLP003 PCR product in lane 2. Lane 1 shows a HindIII λ ladder.



(b) pHE257 cut by SpeI and XhoI. The cut plasmids are shown in lanes 2 and 3. Lane 1 and 4 show colony PstI and HindIII λ ladders, respectively.

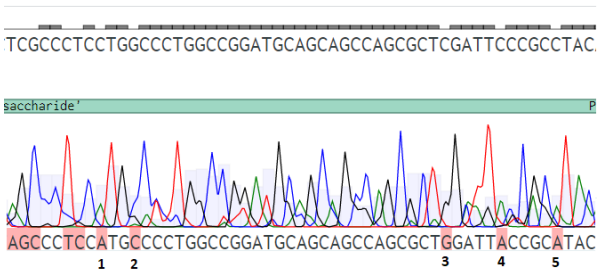


(c) Testing of the transformants by cutting with AfeI/ApaLI. The cut transformants are shown in lane 2 and 3. Lane 1 and 4 show colony PstI and HindIII λ ladders, respectively.

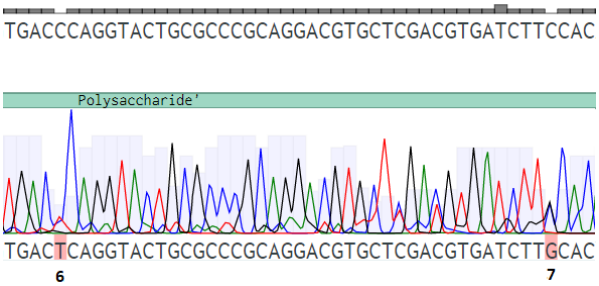
Figure 3.3: Construction of a *Avin05390* deletion plasmid.

The two plasmids were sent for sequencing using the primers opp5390F, sjekk5390F, ned5390R, AmsekV and Del5390. Sequencing of one of the plasmids using opp5390F did not result in a sequence after repeated attempts, leaving larger sequences of the plasmid unchecked. When sequencing the other plasmid, sjekk5390R was also used to sequence the gene. The sequencing covered the majority of the homologous section, but a remaining end segment one of the deleted gene was not sequenced. The sequencing reported 7 mutations that could not be accounted for by overlapping sequencing. These were found in *Avin5380*, a gene adjacent to the. These mutations are displayed in Figure 3.4.

The mutation designated 1 in Figure 3.4a is a frameshift mutation that will change the encoded protein completely, most likely changing its effect entirely. The 6 other mutations were all substitution mutations. Mutation 2 would change amino acid 13 from an alanine to a proline. Mutation 3 would change amino acid 21 from a serine to a tryptophan. Mutation 4 would change amino acid 23 from a proline to a threonine. Mutations 5 and 6 are silent mutations. The last mutation 7 would change amino acid 56 from a phenylalanine to a leucine.



(a) Visualization of pJLP005 sequencing using ned5390R. The upper sequence is the template that the sequencing is compared to. The lower of the sequences are the result of the sequencing, vizualized by Benchling. The red letters marked by numbers are mutations reported by the sequencing. Due to mutation 1 being a frameshift mutation not recognized by the program, the letters marked red are considered mutations. The curve displayed above the sequencing is a display of confidence in the sequencing calculated by Benchling.



(b) Visualization of pJLP005 sequencing using ned5390R. The upper sequence is the template that the sequencing is compared to. The lower of the sequences are the result of the sequencing, vizualized by Benchling. The red letters marked by numbers are mutations reported by the sequencing. The curve displayed above the sequencing is a display of confidence in the sequencing calculated by Benchling.

Figure 3.4: sequencing of the pJLP005 plasmid.

3.1.2 Construction of a plasmid for use in deletion of *Avin30120*

Much like the deletion of the *Avin05390* gene, the deletion of *Avin30120* required the deletion plasmid to contain sections homologous with a part of the gene cluster of *Avin30120*, but with a deleted variation of the gene. The pJLP004 plasmid contains a deleted variation of this gene with an inserted *aprR* gene. The homologous fragment was cut out from pJLP004 and inserted into a vector derived from pHE257 (Figure 3.5).

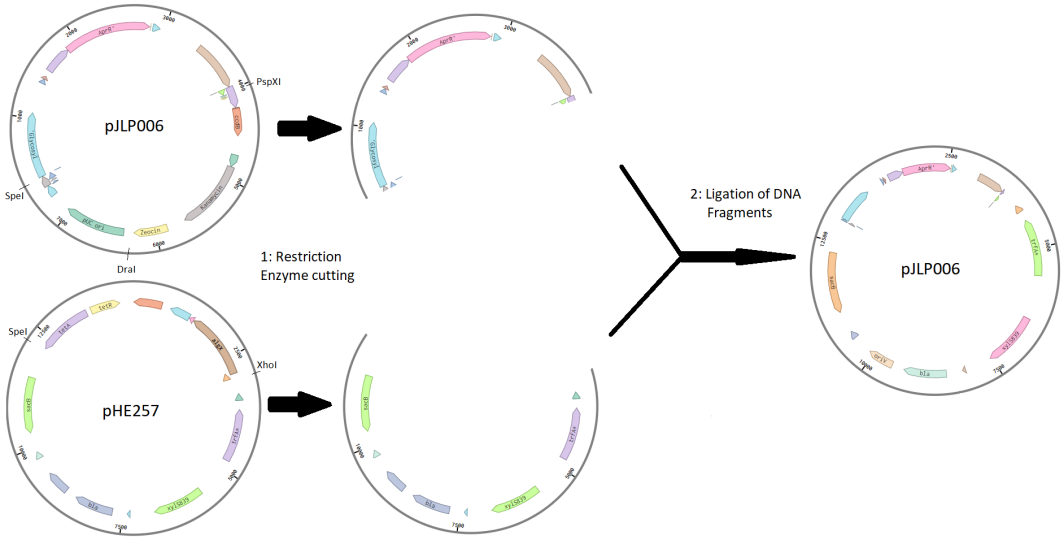
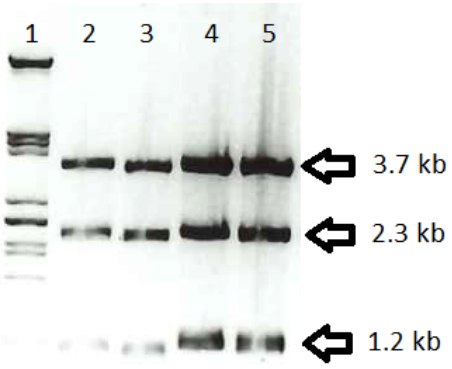
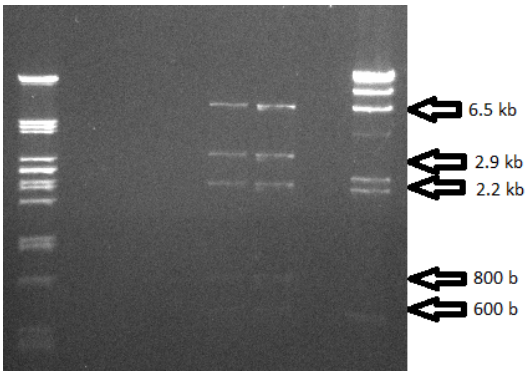


Figure 3.5: Construction strategy of the 30120 deletion plasmid. 1a is the isolation of the insert from pJLP004 by cutting it with SpeI, PspXI and DraI. 1b is the isolation of the vector from pHE257 by cutting it with SpeI and XhoI. The last step is ligation of the two fragments.

The pJLP004 plasmid was cut using SpeI, PspXI and DraI, resulting in 3 segments of sizes 3.7, 2.3 and 1.2 kb (Figure 3.6a). The vector was a pHE257 plasmid cut with SpeI and XhoI, isolating the 9.3 kb fragment (Figure 3.3b) for use in ligation. The ligation was transformed into *E.coli S.17.1*. 2 colonies had their plasmids isolated and tested by cutting it with NsiI. The plasmid construction seems to have been successful, as the result of this cutting consisted of the expected fragments of sizes 6.5, 2.9, 2.2, 0.8, 0.6 and less than 0.1 kb (Figure 3.6b). The plasmid was denoted as pJLP006 and was used for deletion of gene 30120.



(a) pJLP004 cut with SpeI, PspXI and DraI separated in lane 4 and 5. Lane 1 shows a PstI λ ladder.



(b) pJLP004/pHE257 ligation products cut with NsiL shown in lane 2 and 3. Lane 1 and 4 shows colony PstI and HindIII λ ladders, respectively.

Figure 3.6: Construction of *Avin30120* deletion plasmid.

3.1.3 Construction of plasmids for use in repair of *Avin05390*

The plasmid required to repair the *Avin05390* gene had to contain the gene in question as well as the homologous segments flanking the deletion in the mutant around it. This was done by amplifying this section and insert it into a plasmid. The fragment was then transferred to a pHE100 derived vector (Figure 3.7).

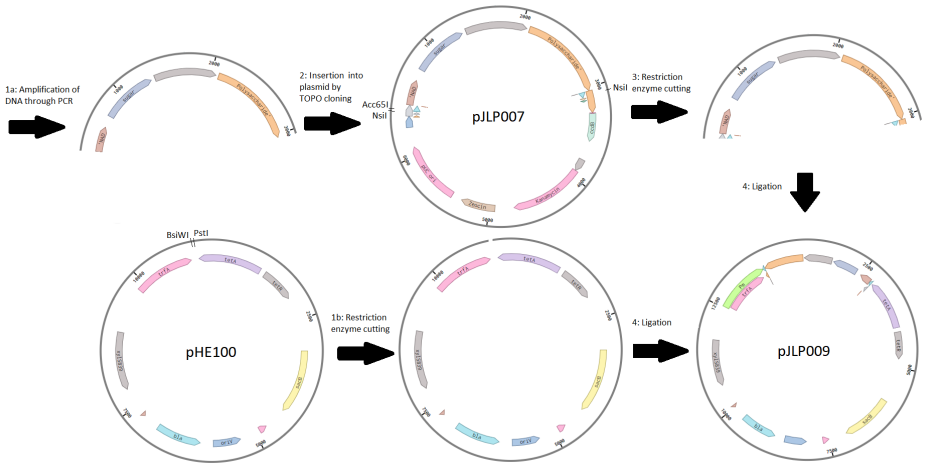


Figure 3.7: Construction strategy of the *Avino5390* repair plasmid. The first step was producing the repaired gene with PCR amplification (Step 1a). The amplified DNA sequences was inserted into a plasmid by TOPO cloning (Step 2). The plasmid was cut (Step 4) and transferred a pHE100 derived vector (Step 4). The vector was prepared by cutting it (Step 1b).

A PCR was performed by using the primers rep5390F and rep5390R and the isolated genome of *A.vinelandii* as the DNA template. This produced a 2.7 kb DNA segment (Figure 3.8).
The DNA segment was inserted into the TOPO vector by TOPO cloning, and the plasmid was transformed into *E. coli DH5*α as described in Section 2.11. Four resulting colonies had their plasmids isolated and cut by NheI and DraI. All four insertions were successful, resulting in two pairs of cloning products with opposite orientation. One of the two plasmids that resulted in fragments of sizes 3.9 and 2.3 kb were used further (Figure 3.9).
The two plasmids were sequenced using primers sjekk5390F, sjekk5390R, M13F and M13R. One of the plasmids was found to have no mutations compared to the original DNA sequence, while the other plasmids contained two mutations. The correct plasmid was designated pJLP007 and the insert was moved to a pHE100 derived vector.

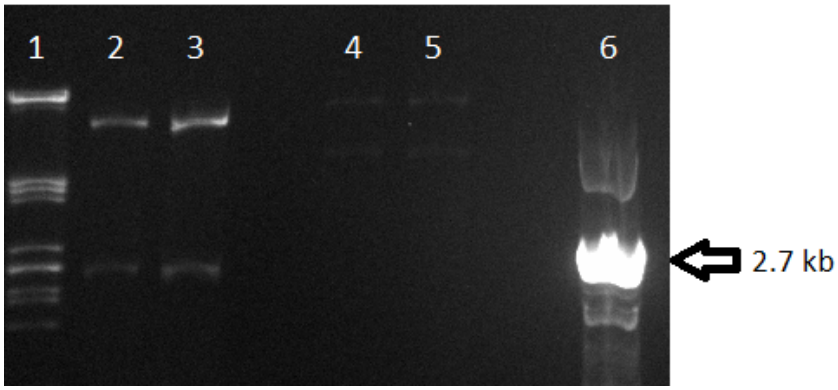


Figure 3.8: The PCR segment produced using the primers rep5390F and rep5390R and the isolated genome of *A. vinelandii* as a DNA template shown in lane 6. Lane 1 shows a PstI λ ladder.

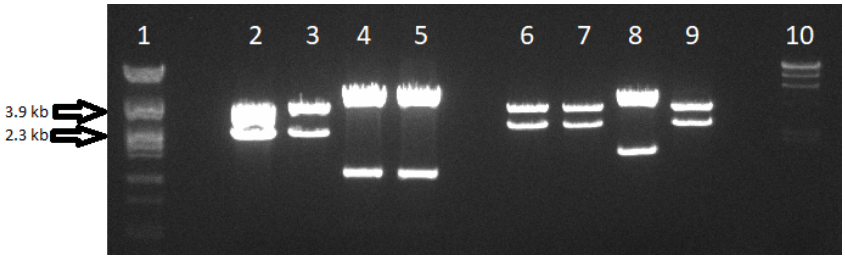


Figure 3.9: TOPO plasmid with inserted rep5390 PCR product cut with NheI and DraI are shown in lanes 2-5. Lane 1 and 10 show colony PstI and HindIII λ ladders, respectively.

pJLP007 was cut with NsiI and Acc65I, producing three fragment of sizes 3.4, 2.8 and less than 0.1 kb (Figure 3.10). Plasmid pHE100 was cut with PstI and BsiWI, producing two fragments of sizes 11 kb and less than 0.1 kb (Figure 3.10).

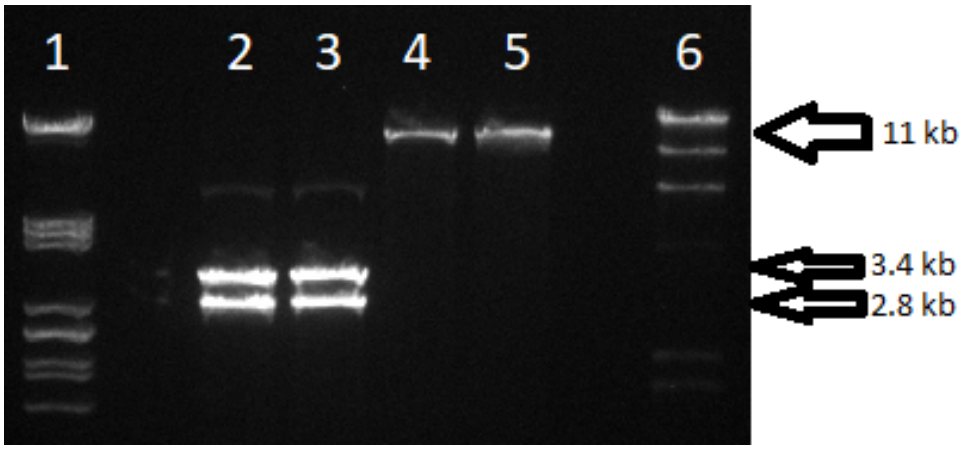


Figure 3.10: pJLP007 cut with Acc65I/NsiI and pHE100 cut with PstI/BSiWI. The cut plasmids are shown in lane 2 and 3. pHE100 cut with PstI and BSiWI is shown in lane 4 and 5. Lane 1 and 6 show colony PstI and HindIII λ ladders, respectively.

The 2.8 kb and 11 kb band of the respective plasmids were isolated, ligated and transformed into an *E. coli S.17.1*. Three colonies were cultivated over night and had their plasmids isolated. The plasmids were cut with NheI and ApaLI, resulting in two band of sizes 9.5 and 4.3 kb (Figure 3.11).
The lower of the two bands of the cut plasmids traveled slightly shorter than the 4.3 kb band of the HindIII DNA ladder. As bands in different lanes occasionally travel at different speeds, it is possible that these bands still represent the intended DNA fragments. The concentration of the PstI lambda ladder seems to have been rather

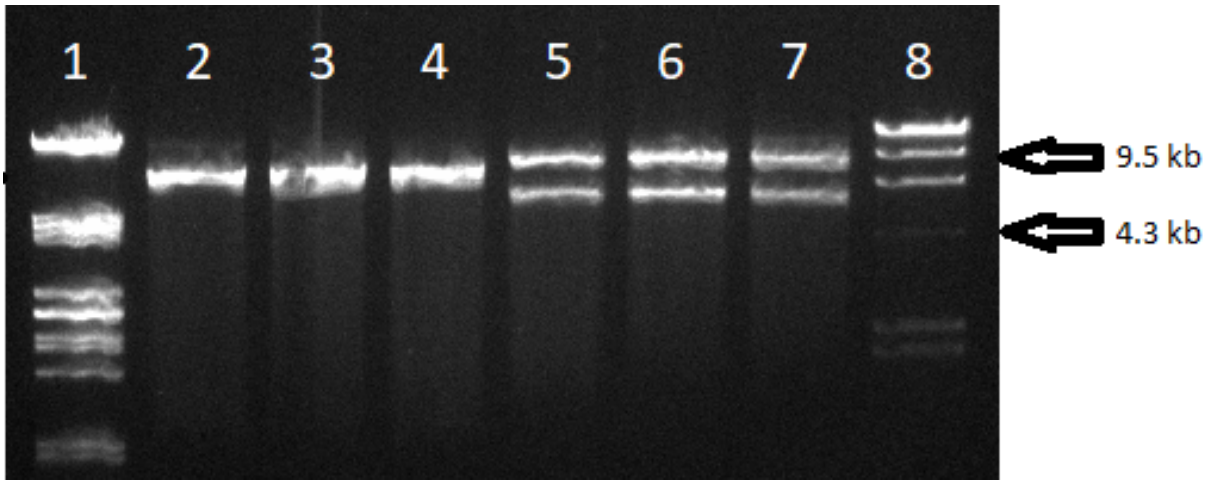


Figure 3.11: pJLP007/pHE100 ligation cut with NheI and ApaLI. Lanes 5-7 show the cut ligation, and Lane 1 and 8 show colony PstI and HindIII λ ladders, respectively.

high, and might therefore have moved across the gel out of sync with the DNA fragments of the other wells. The plasmid is not likely to be the parent plasmid, as cutting these with NheI and ApaLI would result in four bands of sizes 3, 1.8, 0.5 and 0.4 kb for pJLP007 or one 11 kb band for pHE100. The later use of the plasmid detailed in Section 3.4 also suggests that the plasmid was properly produced. This makes the fragment likely to represent a successful ligation. The plasmid corresponding to lane 5 was stored at -80 °C and utilized for repairing the *Avino5390* gene.

3.1.4 Construction of plasmids for use in repair of *Avin30120*

The plasmid needed to repair the *Avin30120* gene had to be created with a homologous section containing the gene. This DNA fragment was produced by PCR amplification, inserted into a TOPO plasmid before being transferred to a pHE100 derived vector (Figure 3.12).

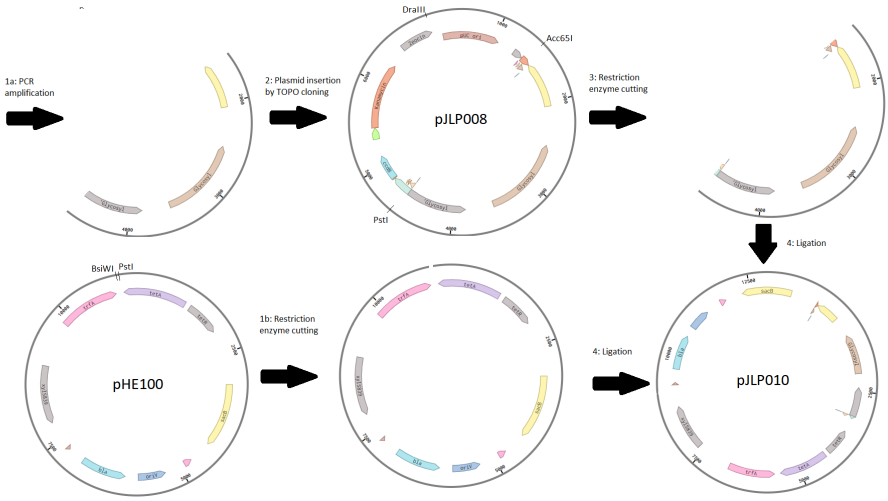
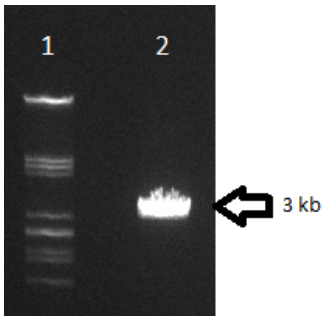
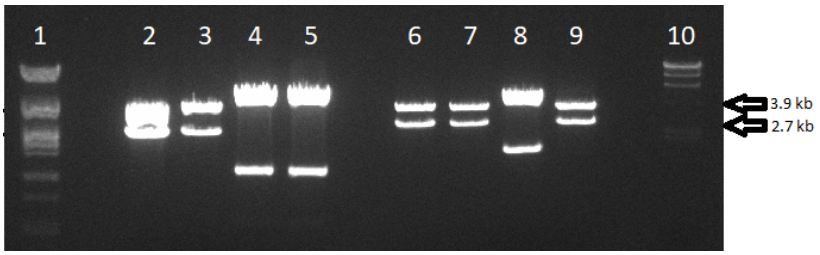


Figure 3.12: Construction strategy for the 30120 repair plasmid. The first step was producing the repaired gene with PCR amplification (Step 1a). The amplified DNA sequences were then inserted into a plasmid by TOPO cloning (Step 2). The plasmid was cut (Step 3) and one of its fragments was transferred to a pHE100 derived vector (Step 4). The vector was prepared by cutting it (Step 1b).

This homologous section was created by PCR and inserted into a TOPO vector by TOPO cloning. The section was thereafter cut out and transferred to the pHE257 derived vector. The PCR process was initialized using the genome of *A.vinelandii* as the template, using the primers opp30120F and ned30120R. This produced a fragment of about 3 kb size (Figure 3.13a), which was inserted into a TOPO vector and transformed into *E. coli DH5*α as described in Section 2.11. Four colonies of the cells transformed with the resulting plasmid were isolated and tested by cutting with NdeI and DraI. The PCR insertion was successful in all of the plasmids, three of these resulting in two fragments of 3.9 and 2.7 kb. The last plasmid resulted in insertion in the opposite orientation as shown in Figure 3.13b.



(a) The Rep30120 PCR product. The PCR segment produced is shown in lane 2. Lane 1 shows a PstI λ ladder.



(b) TOPO cloning product of the 30120 PCR product cut with NdeI and DraI. Cut plasmids are shown in lanes 6-9. Lane 1 and 10 show colony PstI and HindIII λ ladders, respectively.

Figure 3.13: Construction of *Avin30120* repair plasmid

Two of the plasmids were sequenced using the primers sjekk30120F, sjekk30120R, M13F and M13R. Several mutations were found in one of the plasmids. The other plasmid showed one mutation in base 927 (out of 987); a cytosine base had been replaced by a thymine base. When checking the curve of the four sequencing files, all of them gave either equal levels for T and C, or too low levels to conclude whether or not the mutations occured. This mutation would also be silent, as the TTC and TTT codons that the base is involved with both encode for phenylalanine. The latter of these two plasmids were denoted as pJLP008 and used for construction of pJLP010.

pJLP008 was cut with Acc65I, PstI and DraIII, resulting in three bands of sizes 3.1, 2.2 and 1.2 kb (Figure 3.14). The 3.1 kb fragment was isolated from the gel. pHE100 was cut with PstI and BSiWI, resulting in an 11 kb fragment and a fragment of less than 0.1 kb. The 11 kb fragment was isolated (Figure 3.10).

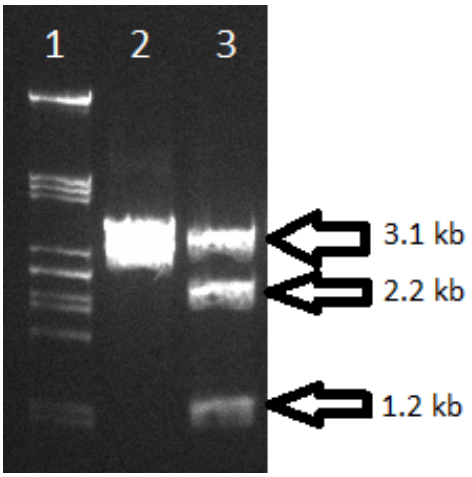


Figure 3.14: pJLP008 cut with Acc65I, PstI and DraIII. The cut plasmid is shown in lane 3, and lane 1 shows a PstI ladder.

The two isolated fragments were ligated and transformed into *E. coli S.17.1*. The plasmids from four colonies of these cells were cut with AfeI and ApaLI to determine whether ligation was successful. Three of the four ligations successfully produced the intended plasmid, as the cutting resulting in two expected fragments of sizes 9 and 5.2 kb. One of three successfully ligated plasmids was designated pJLP010 and stored.

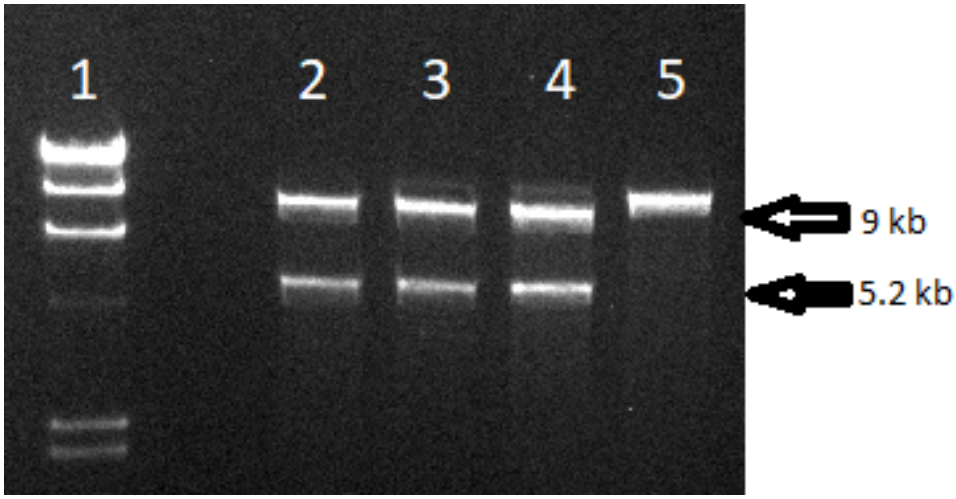


Figure 3.15: pJLP008/pHE100 ligations cut with Afei and ApaLI. Cut plasmids are shown in lanes 2-5. Lane 1 shows a HindIII λ ladder.

3.2 Creation of a *Avin05390* deletion mutant.

The process of conjugation and recombination was utilized, as described in Section 2.19, to replace a section of the *Avin05390* gene in *A. vinelandii* ATCC 12518 wildtype with a section in the pJLP005 plasmid. This recombination was performed with the intention to replace a segment in the middle of the *Avin05390* gene with a DNA segment containing a *AprR* gene. As such, apramycin was utilized to select for cells that accepted the plasmid and for the remainder of the process. Four colonies from the sucrose plates at the end of the procedure were isolated and tested through colony PCR utilizing the primers Sjekk5390F and Sjekk5390R. PCR amplification of the original *Avin05390* fragment would result in 2.1 kb DNA segments while amplification of the altered *Avin05390* gene would result in 1.2 kb DNA fragments. The product of the colony PCR was separated and visualized by gel electrophoresis (Figure 3.16). Colony PCR resulted in a 2.1 kb mutant band with no wildtype band, which shows that the deletion of the *Avin05390* gene was successful. The strain corresponding to lane 3 was stored at -80 °C and denoted *A. vinelandii* ATCC 12518 Δ5390::Apr.

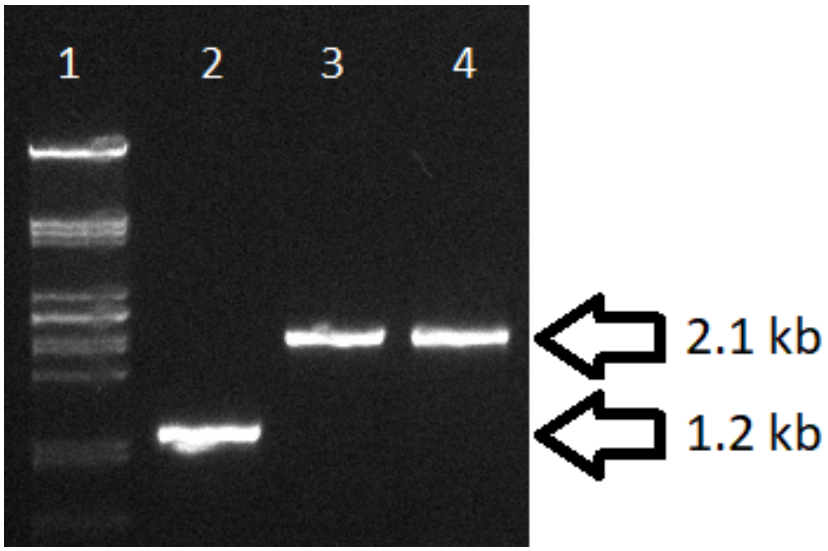


Figure 3.16: Colony PCR of isolated colonies at the end of the *Avin05390* deletion process utilizing primers Sjekk5390F and Sjekk5390R. Lane 2 shows PCR product using the genome of *A. vinelandii* ATCC 12518 and lane 3 and 4 shows PCR product using the genome of *A. vinelandii* ATCC 12518 Δ5390::Apr. Lane 1 shows a PstI λ ladder.

3.3 Creation of a *Avin30120* deletion mutant.

The conjugation and recombination process were performed with plasmid pJLP006, as described in Section 2.19, in an attempt to introduce a deletion in the *Avin30120* gene. This recombination was intended to replace a segment of the *Avin30120* gene with a *AprR* gene. Apramycin was utilized for selection of plasmid-containing A. vinelandii cells until the end of the process. When m-toluate was no longer added to the medium used to cultivate the four parallels, two of the parallels stopped growing after a few days. This seems to suggest that the cells in the two parallels that didn't survive did not retain the *aprR* gene or that inactivation of the *Avin30120* gene resulted in cell death. The two parallels that continued to grow were cultivated on sucrose BM plates at the end of the process. Colonies from these plates were isolated and checked by colony PCR. The colony products were visualized by gel electrophoresis (Figure 3.17). The PCR was performed on each colony with two separate primer pairs, the first pair was opp30120F and ned30120R and the second was rep30120F and rep30120R. In both cases, the deleted gene result would result in bands corresponding to a DNA fragments of 2.2 kb, while the wild type gene would result in bands corresponding to a DNA fragments of 3 kb.

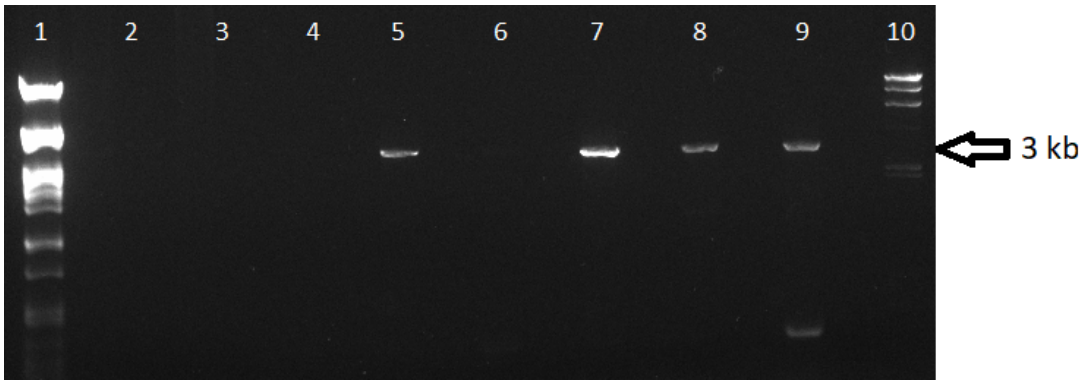


Figure 3.17: Colony PCR used to verify *Avin30120* deletion. The PCR product utilizing primers Opp30120F and Ned30120R are shown in lane 2-5. PCR product utilizing primers Opp30120F and Ned30120R are shown in lane 6-9. lane 1 and 10 shows a PstI and HindIII λ ladder respectively.

Only four of the PCR reactions resulted in bands, and as PCR products of well 5 and 9 correlate to the same colony tested, only three colonies were tested. All these contained one band correlating to a DNA fragment of 3 kb (Figure 3.17). This suggests that the deletion was not successful, and that the *AprR* gene was not inserted, leaving the bacteria vulnerable to antibiotics. The other two parallels that continued to grow did so despite the intended insertion of an *AprR* gene failing. This could be due to ineffective antibiotics, but that seems unlikely as two parallels were unable to grow in the medium containing the antibiotic. Further work towards deletion of the *Avin30120* gene ended here.

3.4 Repair of the *Avin05390* gene deletion

The next step for testing of the *Avin05390* gene was to repair the mutated gene by utilizing recombination as described in Section 2.19. In order to revert the *A. vinelandii* ATCC 12518 Δ5390::Apr back to its wildtype state, the pJLP009 plasmid was conjugated into the bacteria and recombination was used to reinsert the DNA segment removed from the gene and to remove the Apr resistance gene. The antibiotic used throughout this process was tetracyclin. As the second recombination was intended to remove genes related to both apramycin and tetracylin resistance, these antibiotic were excluded from that point. In the final step of the recombination process, apramycin was utilized to determine the colony's resistance to antibiotics. As for the deletion, four colonies were tested with colony PCR using primers Sjekk5390F and Sjekk5390R. A DNA segment produced by amplification of the repaired version of the gene would be 1.2 kb, while segments from amplification of the deleted version of the gene would be a 2.1 kb. The PCR products were separated by gel electrophoresis and is shown in Figure 3.18.

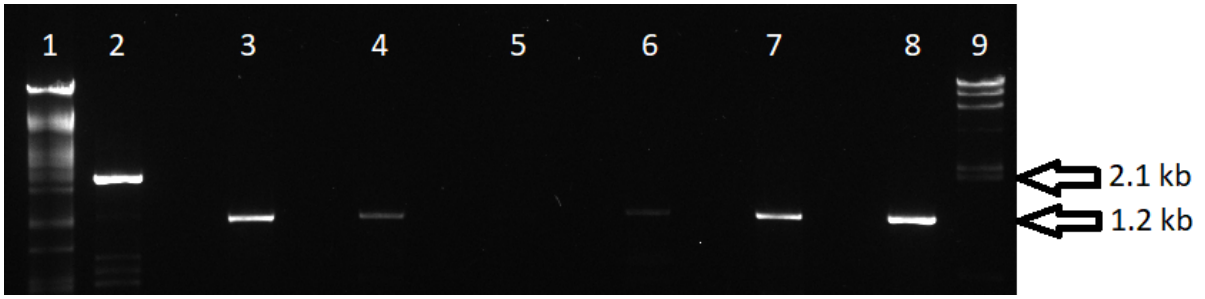


Figure 3.18: Colony PCR of isolated colonies at the end of the *Avin05390* repair process utilizing primers Sjekk5390F and Sjekk5390R. "Lane 2-8 shows the PCR product of *A. vinelandii* ATCC 12518 Δ5390::Apr. Lane 1 and 9 show a PstI and HindIII λ ladder, respectively.

The bands in lane 2 to 8 in Figure 3.18 represent the PCR product of the isolated colonies. Lane 3, 4, 6, 7 and 8 show a single band correlating to segments of 1.2 kb size. This shows that the reparation of the gene was successful in these colonies. Lane 2 shows a band correlating to segments of size 2.1 kb, showing that the deletion is still present in the genome. The colony represented by lane 3 was stored at -80 °C, named *A. vinelandii* JLP 12518 and used for cyst formation testing.

3.5 Testing the cyst formation ability of *A. vinelandii*

The last step of the work was to test the various strains' encystment ability. The general method was to grow cells taken from the same cell suspension before and after the cells were exposed to unfavorable growth conditions. This would then provide a basis for comparison. *A. vinelandii* ATCC 12518, *A. vinelandii* ATCC 12518 Δ5390::Apr and *A. vinelandii* ATCC 12518 rep were put through the cyst induction test as described in Section 2.20. Despite performing the process in a sterile manner, contamination was found on the plates. It was, however, possible to distinguish the contaminated from the non-contaminated colonies due to differences in size, color and texture. The cell colonies after plating the dilution series were counted, and the results can be found in Appendix 5. The number of counted colonies for the plates were used to calculate the number of living cells in the initial cell suspension and the resuspension of the dried cells. This calculation is described in Appendix 5. The calculated cell counts are shown in Table 3.1. This is done to show the approximate amount of cells scraped off the plates and the approximate number of the cells that would survive after the drying process, in order to evaluate how well the strain's ability to survive unfavorable growth conditions through encystment. Table 3.1 shows the results for all three strains as well as the fraction of surviving cells when comparing the cell counts before and after drying the biological parallells. The cell counts of the biological parallels were averages of the largest technical parallel counts. Values used are specified in Appendix 5. Parallels 4-6 of the *A. vinelandii* ATCC 12518 Δ5390::Apr were dried for two days more than the other parallels. Table 3.1 shows the fractions of cells that can survive being dried in unfavorable growth conditions requiring encystment to survive, indicating the cells encystment capability. As the values of the table are made using averages of the technical parallels, the standard deviation of the calculated fraction is also included to determine whether the results are reliable.

Table 3.1: Fraction of surviving cells after encystment. The letter denotes the strain type as the following: D is *A. vinelandii* ATCC 12518 Δ5390::Apr, R is *A. vinelandii* ATCC 12518 rep and W is *A. vinelandii* ATCC 12518. The number denotes the biological parallel. The calculation of the values were calculated as described in Appendix 5.

Biological parallels	Cell count before drying	Cell count after drying	Fraction of surviving cells	Standard deviation
D.1	3.2×10^{10}	9.1×10^{7}	2.83×10^{-3}	1.78×10^{-4}
D.2	7.9×10^{10}	2.3×10^{8}	2.90×10^{-3}	3.39×10^{-3}
D.3	1.2×10^{11}	4.0×10^{6}	3.28×10^{-5}	1.03×10^{-5}
D.4	1.6×10^{11}	5.0×10^{1}	3.16×10^{-10}	2.11×10^{-10}
D.5	7.8×10^{14}	2.2×10^{9}	2.82×10^{-6}	3.98×10^{-6}
D.6	6.9×10^{11}	0	0	-
R.1	1.4×10^{11}	2.0×10^{1}	1.46×10^{-10}	1.6×10^{-10}
R.2	4.7×10^{11}	7.0×10^{4}	1.49×10^{-7}	1.64×10^{-7}
R.3	4.5×10^{11}	6.48×10^{3}	1.45×10^{-8}	1.93×10^{-8}
W.1	1.1×10^{10}	2.0×10^{1}	3.75×10^{-9}	1.90×10^{-9}
W.2	6.9×10^{10}	2.2×10^{4}	3.13×10^{-7}	2.11×10^{-7}
W.3	4.1×10^{11}	8.9×10^{0}	2.21×10^{-8}	1.49×10^{-8}

Chapter 4

Discussion

A deletion was separately attempted to be introduced and then repaired into the *Avin05390* and *Avin30120* with plasmids produced for the purpose. The deletion and repair process was successful, but the deletion might have resulted in unintentional changes. The construction of the *Avin05390* deletion and repair plasmids resulted in plasmids containing the genes and homologous sections as intended, but sequencing of the deletion plasmid suggested that the cloned flanking regions contained several mutations in the *Avin05380* gene that could affect qualities of the protein sequences. These mutations include 4 substitution mutations that change the translated amino acids, which could alter the qualities of the expressed products. Another mutation found was a frameshift mutation early in the sequence, which completely changes the resulting translation and the properties of the products. When looking at the curve depicting the confidence of the sequencing in Figures 3.4a and 3.4b, they were rather short when compared to other amino acids in the sequence. Due to the low confidence of the sequencing, it cannot be confirmed that these mutations are present. Hence, any changes of the cell's phenotype can not be confidently attributed to the intended deletion. I.e. it is possible that the apparent mutations are due to faulty base calls. There is also a chance that the template used for comparison is faulty and that the mutation is actually the correct sequence. This seems unlikely, however, as the sequencing of pJLP007, which contains the same segment, showed no mutations. The sequencing also did not cover a residual DNA segment of the *Avin5390* gene, meaning there could be undiscovered mutations there. As the insertion is intended to inactivate the gene, mutations would not have any effect, but if these mutations remain past the reparations of the gene it could affect the gene's properties. As this segment is directly adjacent to the deletions, the repair is very likely to replace this segment with the equivalent found in the repair plasmid. Also, as the unconfirmed mutations in *Avin05380* are within 350 bases of the deletion and the homologous segment is over 1.2 kb, it is likely that the mutations were removed when deleted segments were repaired. To be able to attribute any changes in the cell's properties to deletion of the gene and not as a byproduct of unwanted mutation, the strains must be tested to determine whether the mutations are present. This could be done by amplifyng the segment of the genome potentially containing mutations in both strains with PCR and sequencing the PCR product. If the experiment were to be repeated, the recombination should be delayed until the homologous segments of the utilized plasmids have been confirmed to be free of mutations.

When attempting to promote deletion of the *Avin30120* gene in four parallel solutions, two of the cultures stopped growing shortly after m-toluate was excluded from the medium, and the two that survived where found to not contain the deletion. This means that the bacteria in the two surviving parallels retained their original genes. As the gene was not present where it was intended, the *AprR* gene was not inserted and the cell should have been voulnerable apramycin. This may suggests developed an alternate method of resistance, or that the cells by unexpected means managed to incorparate the antibiotic resistance gene. The antibiotic might had become

degraded to a point it was rendered inert, but that is unlikely as two parallels did not survive its application, and multiple antibiotic stocks were utilized. As such, deletion of the *Avin30120* gene was unsuccessful. Since none of the parallels survived the recombination process by replacing the *Avin30120* segment with an antibiotic resistance gene, the gene is potentially essential to survival when the cell is grown in Burk's medium.

The deletion process of the *Avin30120* could be repeated in another medium to see if the cells can survive the changes when grown in media containing other substrates. If that is the case then the difference in substrates between the media could be used to investigate the effect of specific deletion on metabolism in specific mediums. There are also many other genes in the specific gene cluster which are related to polysaccharide synthesis, including 8 other glycosyl transferases. Attempting to delete other genes in the cluster might elucidate whether the cluster is essential to survival.

When testing the encystment ability of the *A. vinelandii* ATCC 12518 Δ5390::Apr, five of the six biological parallels grew when cultivated after drying. The survival of the cells suggests that the *Avin5390* gene is not essential for the encystment process. When comparing the fractions of biological parallels of the *A. vinelandii* ATCC 12518 Δ5390::Apr, parallels dried for 7 days (parallels 1-3) showed higher fractions than the parallels dried for 9 days. This difference suggests that 7 days is perhaps not long enough to ensure that only encysted cells survive. The standard deviation of the fraction is far too high to assert anything conclusive. The resulting colony count also shows that the intended dilutions were not achieved. For example, when the technical parallel D.6.1 (Table 5.1) was diluted and the dilutions were plated out, the incubated 10^{-4} dilution resulted in 317 colonies while the 10^{-6} dilution resulted in 41 colonies. This 100x dilution resulted in about a 10x difference. If the dilutions had been made in increments of 10^{-1}, perhaps the concentration of the cells would more closely reflect their dilution. The number of usable parallels is also low for the individual technical parallels. This could have been avoided by incubating the dilutions in increments of 10^{-1}, as that may lead to more plates containing countable colonies. The calculated cell count for some of the parallels, such as the cell count after drying for the biological parallels D.4, R.1 and W.1, could only be performed with colony counts of less than 10 colonies. such low cell counts are inherently imprecise, reducing the certainty of the results.

It is recommended for future work that the encystment test for the strains is repeated with dilutions of smaller increments plated out. This should create a more detailed curve that more accurately determines the fraction of surviving cells. The cells should be dried for a longer period to ensure that only cysts survive. Further improvements on the method of determining whether the surviving cells remained due to encystment or not may also be be useful to achieve more definitive results definitive results. There are also more genes in the clusters related to biosynthesis of polysaccharides that are suitable targets for testing. Bioinformatic studies of the genes of the two clusters might also expand the understanding of the genes' functions within the cell. The mutant strain can be utilized to determine what kind of polysaccharides is produced by *Avin05390*, utilizing methods such as carbohydrate fingerprinting to analyze the difference in polysaccharide production between strains (Rühmann et al., 2015). Little is known about the polysaccharides made by the two clusters, except that they are not alginates, and fingerprinting can be used to study this. Fingerprinting could also be used to ascertain which polysaccharides the mutant is unable to synthesize when compared to the wild type, and whether these polysaccharides are related to encystment. Very few colonies were tested in/when using colony PCR to determine whether deletion of *Avin30120* was successful. Such a small amount of tested colonies are not enough to ensure that the . The stored liquid culture could be cultivated and more colonies could be tested.

Chapter 5

Conclusion

Production of the deletion and repair plasmids for both *Avin05390* and *Avin30120* was successful, but a small section of the deletion plasmid for *Avin05390* was not sequenced. Deletion and repair of the *Avin05390* gene was successful, but whether the insert contained unintentional mutations in a residual segment of the *Avin05390* gene and an adjacent polysaccharide export protein gene is unknown. This would need to be verified by amplification and sequencing of the affected segment in the mutant and the repaired strain before any observations could be attributed to the *Avin05390* deletion. Testing the two produced strains' encystment capabilities compared to the wild type strain suggests that deletion of *Avin05390* does not does not disable the cell's ability to perform encystment. A more detailed version of the test would have to be performed to determine what effect the deletion has. The *A. vinelandii* cell did not survive the attempted deletion of the *Avin030120* gene, suggesting that the gene is essential for cell growth in Burk's medium.

Appendix A: Counted colonies and method of comparison

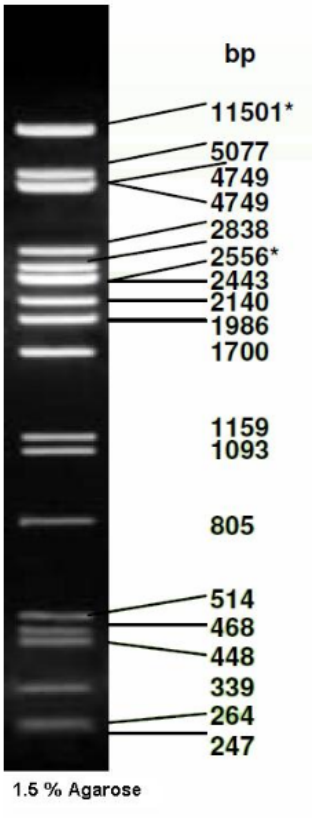
In this appendix, the detailed count from the encystment testing is shown in Table 5.1. A plate count of *x* indicates that the plates in question was overgrown. The cells scraped of an 1-butanol plate were suspended in 100 μl Burks medium. 10 μl of the initial cell suspension was dried and later resuspended in 100 μl 1x burks medium. Both the initial suspension and the resuspension of the dried cells were diluted multiple times to produce different 1 ml solutions of known concentrations. 50 μl of some of these solutions were plated out and the resulting colonies were counted. When quantifying the amount of living cells before drying, the colony count of the plates were used to calculate the cell count in the initial cell suspension. When quantifying the number of surviving cells after drying, the colony count of the plates were used to calculate the cell count in the post-drying cell suspension. Since this latter suspension was 10% of the initial suspension, this calculated cell count after drying was upscaled by multiplying the cell count by 10. The colony counts for all the strains, as well as the calculated cell count, is displayed in Table 5.1, and the bold values are those used in calculating the cell count. When determining the fraction of cells surviving drought, the upscaled cell count calculated after drying is divided by the cell count of the initial cell suspension. The resulting fractions are shown in Table 3.1.

Table 5.1: Cell colony count of the three *A. vinelandii* strains. Three biological parallels split into two technical parallels were plated out before and after 7 days of drying. The first row indicates whether the counts below were made before or after drying, while the second row of the tables indicates the dilution of the solution plated out. The numbers used in the calculation are highlighted in bold. The first column shows the designation of the plate parallel. The samples are denoted by as a.b.c, where a is a letter denoting the given strain, and b and c is indicate the biological and technical parallels. The prefix D denotes a strain containing a deletion, R a strain where the deletion is repaired and W denotes the wild type strain. Colony counts marked *x* "were plates overgrown to a point making counting impossible.

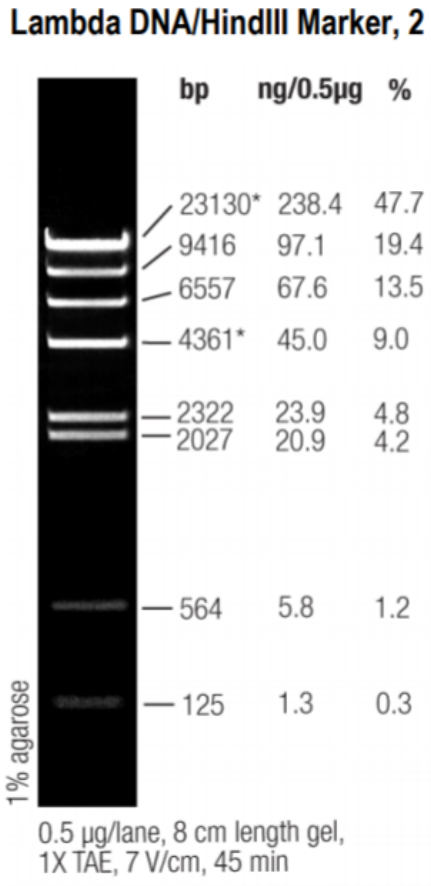
	before drying					after drying			
	10^{-3}	10^{-4}	10^{-6}	10^{-9}	calculated cell count of initial suspension	10^{0}	10^{-3}	10^{-6}	calculated cell count of dried cells
D.1.1	x		**17**	0	3.4×10^{10}	x	**456**	0	9.1×10^{6}
D.1.2	x		**15**	0	3.0×10^{10}	x	**451**	0	9.0×10^{6}
D.2.1	x		**31**	0	6.2×10^{10}	x	**53**	0	1.1×10^{6}
D.2.2	x		**48**	0	9.6×10^{10}	x	**385**	3	7.7×10^{6}
D.3.1	x		**42**	0	8.4×10^{10}	x	**19**	0	3.8×10^{5}
D.3.2	x		**80**	5	1.6×10^{11}	x	**21**	0	4.2×10^{5}
D.4.1		x	**99**		1.1×10^{11}	**2**	0	0	2
D.4.2		x	**38**		2.0×10^{11}	**0**	0	0	0
D.5.1		x	**341**		1.2×10^{12}	x	**7**	0	4.4×10^{8}
D.5.2		x	**128**		1.6×10^{15}	**0**	0	0	8.0×10^{4}
D.6.1		**317**	41		6.4×10^{11}	**0**	0	0	0
D.6.2		x	**130**		7.4×10^{11}	**295**	1	0	0
R.1.1		**486**	22		2.0×10^{11}	**1**	0	0	2
R.1.2		**582**	12		7.6×10^{10}	**3**	0	0	6
R.2.1	x		**38**	0	6.8×10^{11}	**242**	1	0	1.4×10^{4}
R.2.2	x		**31**	0	2.6×10^{11}	x	**2**	0	0
R.3.1	x		**339**	1	6.3×10^{11}	x	**0**	0	0
R.3.2	x		**66**	0	2.6×10^{11}	**447**	0	0	590
W.1.1	x		**56**	0	9.7×10^{9}	**1**	0	0	2
W.1.2	x		**102**	0	1.2×10^{10}	**4**	0	0	6
W.2.1	x		**576**	0	7.6×10^{10}	x	x	**22**	484
W.2.2	x		x	**780**	6.2×10^{10}	x	**4**	0	4.0×10^{3}
W.3.1	x		**318**	15	6.8×10^{11}	x	**0**	0	0
W.3.2	x		**370**	21	1.3×10^{11}	**447**	0	0	894

Appendix B: DNA ladders

DNA ladders are solutions containing DNA fragments of known sizes. The two ladders used in this thesis were made by cutting Lambda DNA supplied by Termofisher with restriction enzymes PstI and HindIII respectively. When utilizing gel electrophoresis to vizualize DNA fragments, the ladder is run in an adjacent well. This enables comparison to fragments of unknown sizes to the fragments of the ladder. The ladders are displayed in figures 5.1a and 5.1b respectively.



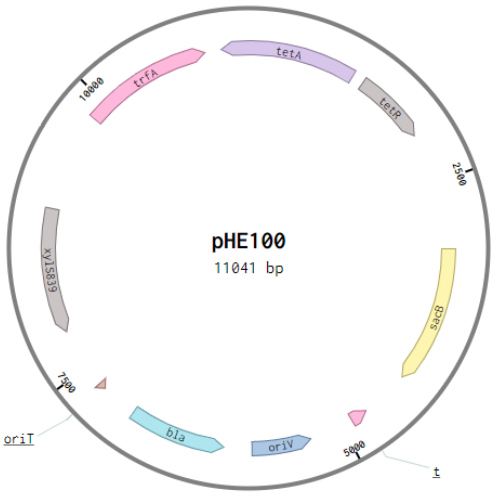
(a) PstI Lambda ladder.(www.geneon.net)



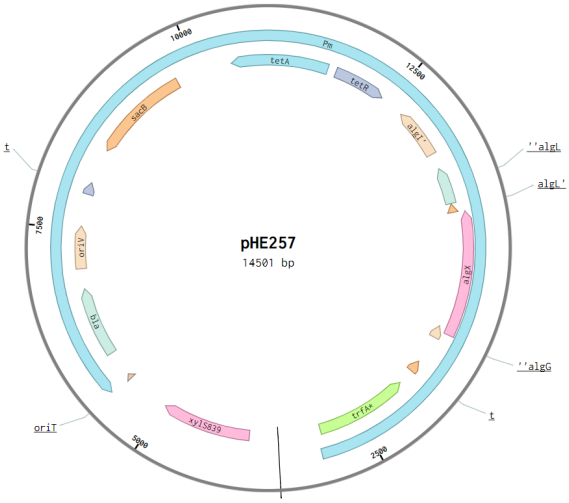
(b) HindIII Lambda ladder. (www.termofisher.com)

Figure 5.1: Construction of *Avin30120* repair plasmid

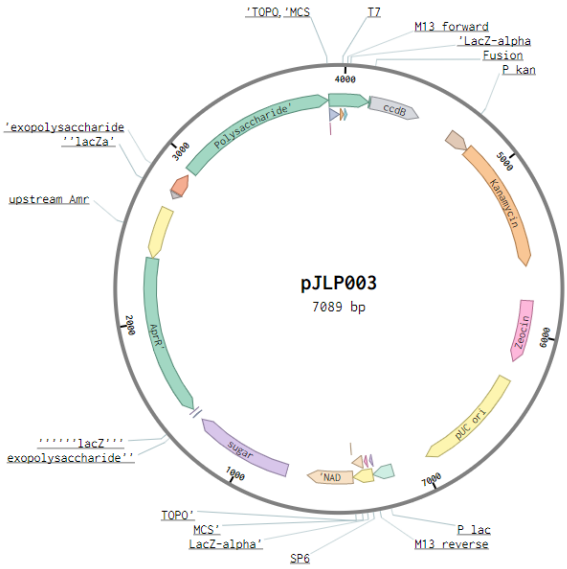
Appendix C: Plasmids used in the project



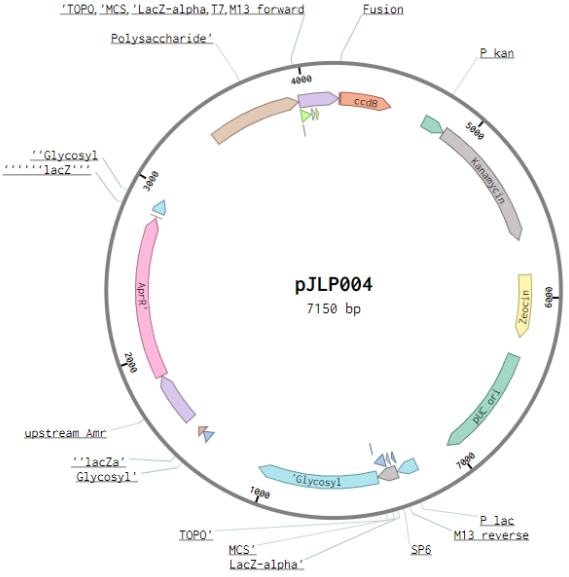
(a) pHE100



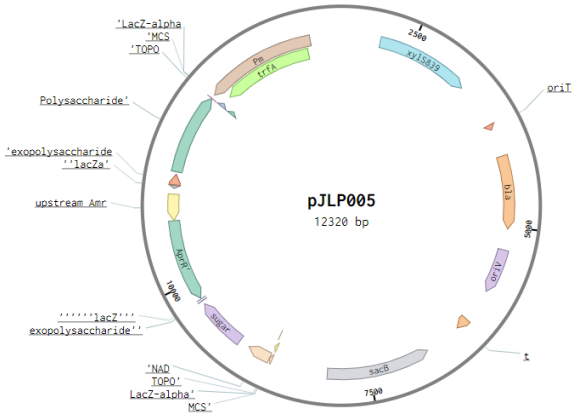
(b) pHE257



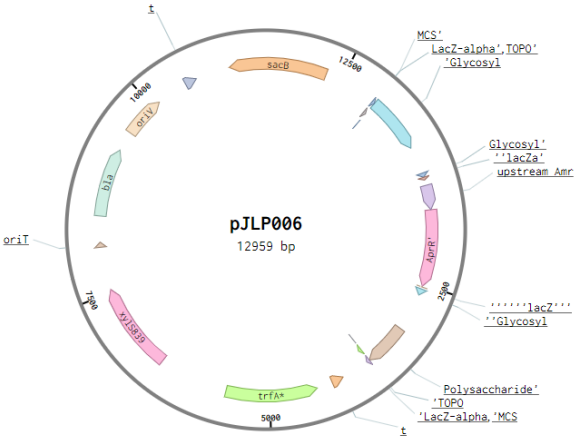
(c) pJLP003



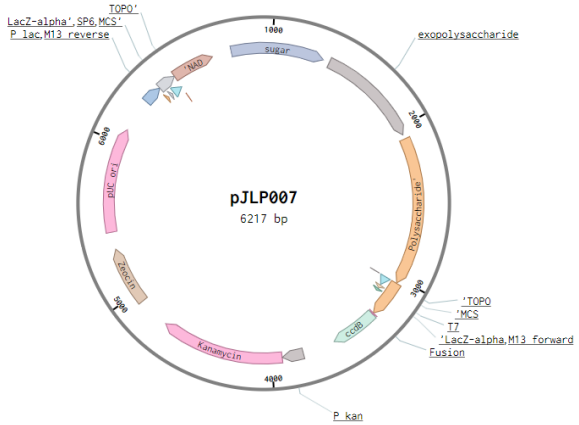
(d) pJLP004



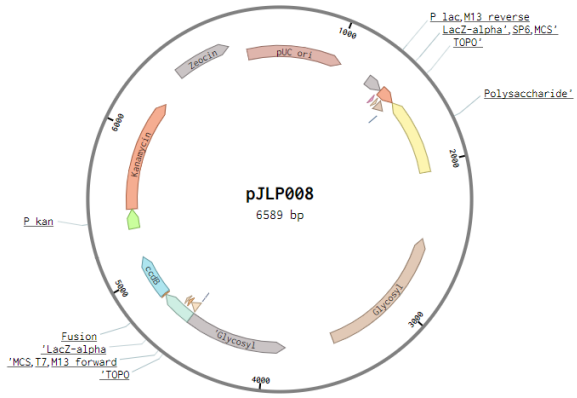
(a) pJLP005



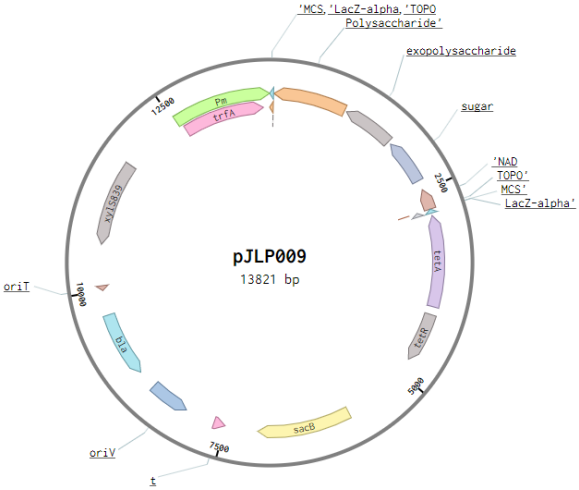
(b) pJLP006



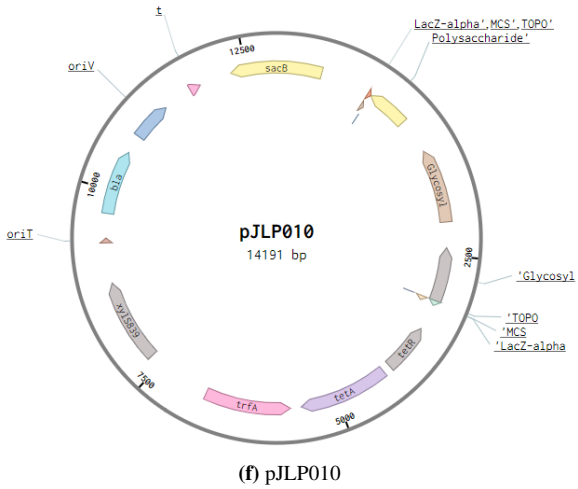
(c) pJLP007



(d) pJLP008



(e) pJLP009



(f) pJLP010

Appendix D: Sequencing

This appendix shows a detailed alignment of the sequencing results of the PCR products used to produce the utilized *Avin05390* deletion plasmid, *Avin05390* repair plasmid and *Avin30120* repair plasmid. The primer's name and related sequence are only displayed where they overlap with the template of the wanted plasmids. Dots indicate where the primer contains the same base as the template, while hyphens indicate where the sequence has no base. The end of the sequencing results are displayed with repeated hyphens that proceed past the end of the hyphen. The *Avin05390* deletion plasmid contains a 870 base unsequenced area between the sequencing results of primers ned5390R and sjekk5390F which is not displayed in the sequencing.

Multiple alignment of the homologous segment of the wanted *avin05390* deletion plasmid and the sequensing results

```
plasmid    CAAGGCCCTTCGGGAATTGCTGCCTTTGCAGCCCGGGGACGTGCTCGACACTTGTGCCGACGTCTCGGAATTGCA
del5390    ............................................................................
opp5390F   ----------------------------------------...........................

plasmid    GCGGGACACCGGCTTTCAGCCCCAGGTAAGTCTCGAAGAAGGCCTCGCCCGATTCGTATCCTGGTTTTGCCACTA
del5390    ............................................................................
opp5390F   ................................T...........................................

plasmid    CTACAGGATCGACAGCCCCGATGCCCCGGCCATTTCATCCGCCGCCGCGGACGGGGAGCTGTCGCGTGAGGTAAT
del5390    ............................................................................
opp5390F   ............................................................................

plasmid    GCGCAATGAATGCCCCGGGAAACTCGTTTCAACAGCCCAGCTTCGTTAAGGACGGCCATTACACGCTCAAGGAAA
del5390    ............................................................................
opp5390F   ............................................................................

plasmid    AACGTACCGATCCCTCGCAGCGGGAGAAGATCCAGAACGCCCTCTGGGTCCACGATCGCGGCTGGTTCAGCGGCCG
del5390     ...........................................................................
opp5390F    ...........................................................................

plasmid    CCAGGGCGGCAGCCCCTGGAGCCTGTCCCGCACCAAGCGCCTCGGCGAGGCCTGCGCCGCCCTGCTGCTGCTGCT
del5390    ............................................................................
opp5390F   ............................................................................

plasmid    CCTGCTCTCCCCCCTGCTGCTCGGCATCGCCCTGCTGATCAGGCTGGACAGCTCCGGTCCCGTGCTGTTCGTCCA
del5390    ............................................................................
opp5390F   ............................................................................

plasmid    GCGGCGCACCGGCTTTCGCGGGCGCCGCTTCGGCATGTACAAGTTCCGCAG-CATGGTGGCCGACGCCG-AGGCC
del5390    ........T...................................C.................A.....
opp5390F   ............................................................................

plasmid    CTGAAGGACTCGCTGCGCCACCTGAACAAGCACGGCGCCGAAGCCATCGACTTCAAGATCGACGCCGACCCGCGG
del5390    ............................................................A.....
opp5390F   ............................................................................
```

```
plasmid     GTGACCGCGGTCGGCCGCTGGCTGCGCCGTACCAGCCTGGACGAGCTGCCCAACCTGATCAACGTGGTGCGCGGC
del5390     ................................-.......A.A......-.......----------------
opp5390F    ....................................................................------

plasmid     GAGATGCGCCTGGTCGGGCCGCGCCCCACCTCGTTCAATGCCAGCCGCTACCACGAGCACCACCTGGGCCGGCTG
opp5390F    ....................................................................------
Sjekk5390F  -------------------------..........-....................................
AmsekV      ------------------------------------------------------------......A.T

plasmid     AG---CATCTATCC-GGGCATGA--CCGGCCTCTGGCAGATCTCC--GGACGCAGC-GACGTCGATTTCGA--CG
opp5390F    ..---.........-........--..................--.........-.............--..
Sjekk5390F  ..---.........-........--..................--.........-.............--..
AmsekV      G.AGC.........T........AC.G...TC......A..TCT.CG.A.......G...........T.GACGA

plasmid     ACCGGGTCGAACTCG-ACATGGACTACATCCGCCGTCAGGGCCCCCTGCTGGACCTGTGGATTCTGCTG-AAAAC
opp5390F    ..............-.....................................................A.....
Sjekk5390F  ..............-.....................................................-.....
AmsekV      C......A.....T.T.......T.....................................................-.....

plasmid     CCCCGTCGAGGTGTTCCACGGCCACGGCGCCGCTTGAACCCTCATACCCCCAGGCAGAGCACGCATCAGGAGTCA
opp5390F    ....................................................................------
Sjekk5390F  ....................................................................------
AmsekV      ....................................................................------

plasmid     GCCCATGACCATGCTAC----TTCTAGAGCGGCCGAGCCTAGGTCAGGCTGCGCAACTGTTGGGAAGGGCGATCG
Sjekk5390F  .................CTAG.....................................................
AmsekV      .................CTAG.....................................................
plasmid     GTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCA
Sjekk5390F  ....................................................................------
AmsekV      ....................................................................------

plasmid     GGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTAATTCGAGCTCAGCCAATCGACTGGCGAGCG
Sjekk5390F  ....................................................................------
AmsekV      ....................................................................------

plasmid     GCATCGCATTCTTCGCATCCCGCCTTCTGGCGGATGCAGGAAGATCAACGGATCTCGGCCCAGTTGACCCAGGGC
Sjekk5390F  ......................-.....................................................
AmsekV      ......................-.....................................................

plasmid     TGTCGCCACAATGTCGCGGGAGCGGATCAACCGAGCAAAGGCATGACCGACTGGACCTTCCTTCTGAAGGCTCTT
Sjekk5390F  ....................................................................------
AmsekV      ..........................................................A.....------

plasmid     CTCCTTGAGCCACCTGTCCGCCAAGGCAAAGCGCTCACAGCAGTGGTCATTCTCGAGATAATCGACGCGTACCAA
Sjekk5390F  ....................................................................------
AmsekV      ....G.....................................................A.....------

plasmid     CTTGCCATCCTGAAGAATGGTGCAGTGTCTCGGCACCCCATAGGGAACCTTTGCCATCAACTCGGCAAGATGCAG
Sjekk5390F  ....................................................................------
AmsekV      ..G.G.....G.............................................G.............G....
```

```
plasmid     CGTCGTGTTGGCATCGTGTCCCACGCCGAGGAGAAGTACCTGCCCATCGAGTTCATGGACACGGGCGACCGGGCT
Sjekk5390F  ............................................................................
AmsekV      ....T...A....G....A.....................A..............----------------

plasmid     TGCAGGCGAGTGAGGTGGCAGGGGCAATGGATCAGAGATGATCTGCTCTGCCTGTGGCCCCGCTGCCGCAAAGGC
Sjekk5390F  ............................................................................

plasmid     AAATGGATGGGCGCTGCGCTTTACATTTGGCAGGCGCCAGAATGTGTCAGAGACAACTCCAAGGTCCGGTGTAAC
Sjekk5390F  .....................................T..........A.A.......T...TC.G.-------

plasmid     TCGCCCTCC-TGGCCCTGGCCGGATGCAGCAGCCAGCGCTCGATTCCCGCCTACATCCTCACCGCCCCCGAGGAC
Ned5390R    .A.......A.C............................G....A....A.....................

plasmid     CAGGCCACGCACAGCGAACTGATCCCGGTGACCCAGGTACTGCGCCCGCAGGACGTGCTCGACGTGATCTTCCAC
Ned5390R    ...............................T.....................................G...

plasmid     ATCGAGATGGATTCGCCCAACGCCTACCGCATCCAGCCGGGCGACCAACTGGAGATCCTGTTCGCCACCGCCAAG
Ned5390R    ............................................................................

plasmid     GGCCTGACCGGCATCAAGACGGTGATGCCCGACGGCACCGTCAATCTGGAATACGTCGGCAGCCTCCAGGTCGCC
Ned5390R    ............................................................................

plasmid     GGGCTCACCGTCGACGAGGCCCAGGCGCTGATGACCGAGCGCTACAAGGACACCCTGCGCACACCGCTGATCACC
Ned5390R    ............................................................................

plasmid     GTCTCGGTGGCCAAGGCGCAGACCCGCGCCCAGAACCTGCGCGACACCCTGTTCAGCCCGAACACCGGCATGAGC
Ned5390R    ............................................................................

plasmid     CGGGAAATCACCGTCGGCGCCGACGGCCGCGCGACCTTCCCGATGCTCGGCAGCATCCGCCTCGGCGGCATGAAC
Ned5390R    ............................................................................

plasmid     AGCGACGAACTCGAAAAACTCCTGAACGAGCGCTACCGCGAGGAAGTCGGGCCGATCCAGGTCGACGTGCTGCTC
Ned5390R    ............................................................................

plasmid     AAGTCCACCGCCGCCAACGAGGTCTACATACTCGGCGAGGTCGGCCAGCCCGGCGCCTATCCGGTGCGCCGGCCG
Ned5390R    ............................................................................

plasmid     ATCTCGGTGCTCGAGGCGCTGACCCTGGCGCGCGGCTACAACCCCGTCACCGCCGACCTGGAATCGGTGCTGATC
Ned5390R    ............................................................................

plasmid     CTGCATCGCGACGGCGAGCGGGTGGTGTCGCGCACCTACGACATGGAGGCGTTGCTCGACAACGACGCCGACAGC
Ned5390R    ............................................................................

plasmid     GTCGCCTACCTGCAGCCGGACGACCTGCTCTTCGTGCCGCGCAGCAGCCTCGCCGGCGCCGGCGACACCATGCGC
Ned5390R    ..............................................-------------------------
```

Multiple alignment of the homologous segment of the wanted *avin05390* repair plasmid and the sequensing results

```
plasmid     CTATTTAGGTGACACTATAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGC
MI3R        .ACA........-............................................................

plasmid     CGCCAGTGTGCTGGAATTCGCCCTTCAAGGCCCTTCGGGAATTGCTGCCTTTGCAGCCCGGGGACGTGCTCGACA
MI3R        ............................................................................

plasmid     CTTGTGCCGACGTCTCGGAATTGCAGCGGGACACCGGCTTTCAGCCCCAGGTAAGTCTCGAAGAAGGCCTCGCCC
MI3R        ............................................................................

plasmid     GATTCGTATCCTGGTTTTGCCACTACTACAGGATCGACAGCCCCGATGCCCCGGCCATTTCATCCGCCGCCGCGG
MI3R        ............................................................................

plasmid     ACGGGGAGCTGTCGCGTGAGGTAATGCGCAATGAATGCCCCGGGAAACTCGTTTCAACAGCCCAGCTTCGTTAAG
MI3R        ............................................................................

plasmid     GACGGCCATTACACGCTCAAGGAAAAACGTACCGATCCCTCGCAGCGGGAGAAGATCCAGAACGCCCTCTGGGTC
MI3R        ............................................................................

plasmid     CACGATCGCGGCTGGTTCAGCGGCCGCCAGGGCGGCAGCCCCTGGAGCCTGTCCCGCACCAAGCGCCTCGGCGAG
MI3R        ............................................................................

plasmid     GCCTGCGCCGCCCTGCTGCTGCTGCTCCTGCTCTCCCCCCTGCTGCTCGGCATCGCCCTGCTGATCAGGCTGGAC
MI3R        ............................................................................

plasmid     AGCTCCGGTCCCGTGCTGTTCGTCCAGCGGCGCACCGGCTTTCGCGGGCGCCGCTTCGGCATGTACAAGTTCCGC
MI3R        ............................................................................

plasmid     AGCATGGTGGCCGACGCCGAGGCCCTGAAGGACTCGCTGCGCCACCTGAACAAGCACGGCGCCGAAGCCATCGAC
MI3R        ............................................................................

plasmid     TTCAAGATCGACGCCGACCCGCGGGTGACCGCGGTCGGCCGCTGGCTGCGCCGTACCAGCCTGGACGAGCTGCCC
MI3R        ............................................................................

plasmid      AACCTGATCAACGTGGTGCGCGGCGAGATGCGCCTGGTCGGGCCGCGCCCCACCTCGTTCAATGCCAGCCG
MI3R         ...........................................................................
Sjekk 30120F -----------------------------.................-.....................

plasmid     CTACCACGAGCACCACCTGGGCCGGCTGAGCATCTATCCGGGCATGACCGGCCTCTGGCAGATCTCCGGAC
MI3R        ...........................................................................
Sjekk30120F ...........................................................................

plasmid      GCAGCGACGTCGATTTCGACGACCGGGTCGAACTCGACATGGACTACATCCGCCGTCAGGGCCCCCTGCTG
MI3R         ...........................................................................
Sjekk30120F  ...........................................................................
Sjekk30120R  ....G..A..T............-.....G.....................................

plasmid      GACCTGTGGATTCTGCTGAAAACCCCCGTCGAGGTGTTCCACGGCCACGGCGCCGCTTGAACCCTCATACC
MI3R         ............................A..........................................
Sjekk30120F  ...........................................................................
Sjekk30120R  ...........................................................................
```

```
plasmid     CCCAGGCAGAGCACGCATCAGGAGTCAGCCCATGACCATGCTCCGTTCCGAGCAGACCACGCTATTGCTGG
MI3R        ......----------------------------------------------------------------
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     AGCAGGACCGCCGCGTTCTCAACCCCAGCGAAGTCAACCTGGCGGCCACCCTGCTCGACCAGAACTGCCGC
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     ACCCTGCTGCTGATCGCCCCGACCAGCGGCTGTGGCACCACCACCTGCGCCCTGAGCATGGCCCGGCAACT
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     GGCCGGCTCGGTCAAGGGCAAGCTCCTGCTGGTCGATGCCGGCCCCTCGGCCACCGGCCTCAGCTCGCGCC
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     TGGGCATGGCGGCCGACCGCGGCCTGTTCGAGCTGCTGCAGAGCCAGCATCCGGAGGGCGAACTCCGCCAC
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     TGCATCCAGCGCCATCCCGAACTGCCCTTCGATCTGCTGCCGCTCGGCCAGCCTTCGGTGGCGGCCGGCCG
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     CCTCACCGCCGAGGACCTGCAGTACCTGTTCGACTGCCTGGCGGCCCGCTACCGCTTCGTGGTGATCGACG
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     CCGAGGCCGTCTACGGCGGCAGCAGCGGCCTGATCCTGGCGGCCATGGCCGACGCCGTGGCGCTGGTGGTG
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     CGCGCCGAGGAAACCCGCTGGGAAGTCGCCCAGGCCGCCGTGCAGCGCCTGCGCCAGGCCAACGCCAGGCT
Sjekk30120F ................................................................
Sjekk30120R ................................................................

plasmid     GCTCGGCAGCGTGCTCAACGCCCGCCGGCTCTACCTGCCGAAATGGCTCTACCGGCTGCTCTGATCGGATC
Sjekk30120F ................................................................
Sjekk30120R ................................................................
MI3F       --------------------------------.............T.....................

plasmid     GCCCACAGGACGGACGCCACCATGAACAAGCCCCTGCTTTTCCTGCTCGCCCTCCTGGCCCTGGCCGGATG
Sjekk30120F ................................................................
Sjekk30120R ................................................................
MI3F       ..T.........................................................T...................

plasmid     CAGCAGCCAGCGCTCGATTCCCGCCTACATCCTCACCGCCCCCGAGGACCAGGCCACGCACAGCGAACTGA
Sjekk30120F ................................................................
Sjekk30120R ................................................................
MI3F        ................................................................
```

```
plasmid      TCCCGGTGACCCAGGTACTGCGCCCGCAGGACGTGCTCGACGTGATCTTCCACATCGAGATGGATTCGCCC
Sjekk30120F  ..................................................-.....C--------------
Sjekk30120R  ............................ACGT..T-----------------------------------
MI3F         ......................................................................

plasmid   AACGCCTACCGCATCCAGCCGGGCGACCAACTGGAGATCCTGTTCGCCACCGCCAAGGGCCTGACCGGCATCAAG
MI3F      ......................................................................

plasmid   ACGGTGATGCCCGACGGCACCGTCAATCTGGAATACGTCGGCAGCCTCCAGGTCGCCGGGCTCACCGTCGACGAG
MI3F      ......................................................................

plasmid   GCCCAGGCGCTGATGACCGAGCGCTACAAGGACACCCTGCGCACACCGCTGATCACCGTCTCGGTGGCCAAGGCG
MI3F      ......................................................................

plasmid   CAGACCCGCGCCCAGAACCTGCGCGACACCCTGTTCAGCCCGAACACCGGCATGAGCCGGGAAATCACCGTCGGC
MI3F      ......................................................................

plasmid   GCCGACGGCCGCGCGACCTTCCCGATGCTCGGCAGCATCCGCCTCGGCGGCATGAACAGCGACGAACTCGAAAAA
MI3F      ......................................................................

plasmid   CTCCTGAACGAGCGCTACCGCGAGGAAGTCGGGCCGATCCAGGTCGACGTGCTGCTCAAGTCCACCGCCGCCAAC
MI3F      ......................................................................

plasmid   GAGGTCTACATACTCGGCGAGGTCGGCCAGCCCGGCGCCTATCCGGTGCGCCGGCCGATCTCGGTGCTCGAGGCG
MI3F      ......................................................................

plasmid   CTGACCCTGGCGCGCGGCTACAACCCCGTCACCGCCGACCTGGAATCGGTGCTGATCCTGCATCGCGACGGCGAG
MI3F      ......................................................................

plasmid   CGGGTGGTGTCGCGCACCTACGACATGGAGGCGTTGCTCGACAACGACGCCGACAGCGTCGCCTACCTGCAGCCG
MI3F      ......................................................................

plasmid   GACGACCTGCTCTTCGTGCCGCGCAGCAGCCTCGCCGGCGCCGGCGACACCATGCGCCAACTGGCCGACGTGATG
MI3F      ......................................................................

plasmid   CTCTTCAGCGGCTAGCAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGG
MI3F      ......................................................................

plasmid   GCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCC
MI3F      ..............AGA.------------------------------------------------------
```

Multiple alignment of the homologous segment of the wanted *avin30120* repair plasmid and the sequensing results

```
plasmid   TTTAGGTGACACTATAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCA
MI3R      A........-.............................................................

plasmid   GTGTGCTGGAATTCGCCCTTAAGGCTCCGGTGAGAATCACCCGGCGGCCGCCGAAACGGTTCACGCCGACATCCACC
MI3R      ............................................................................

plasmid     TGCGGATTCTCGATATAGCGGGCCAGGCGCCCGGCGATGTCGTTGCGCAATTGCTCGATGGTCAGGCCGGCGGCC
MI3R        ..........................................................................
rep30120R   ----------------------------------.....................................

plasmid     TGGATCTCGCCGATATAGGGGTAGAACAGGCGGCCGTCGGGGCGCACTGGGCGGCCGCTGGCTTCGGACGACTGC
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     TGCACGCCGGCCGGGTTGGTCAGTTCCGGATGATCCCAGACGGTGATATTGAGCACGTCGCCGACGCCGATCAGA
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     TAGGGGCCGGGGGTGTAGCCGGTCAGCTCGGGCGGCAGTTTCTCGACGACGCGGGTGGCTTCCTGGGTCGCCACC
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     AGTCGGGTGGTGATGGGCACCAGCTCGAAGCGCGGGCTTTCCGGCGGGCCCTTGCCGCGGAGTTCGGAGGTGTCG
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     AGGTAGGGGCCGGGGGAGAACATGCAGCCCTGGAGGAGGGCGGCTACCAGAGCGAGGATGCGTGTACGGCGCGCT
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     GCCGCGCTGCCGCTGTGCGGGACGGACAAGGCGAATCCCTCCATGGGCCTTTGGGGCGCTGGAATGACGATCCGT
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     TTTCAGGCGATCGACGGGGAGCCGGAAGCGTTTACCGATGGCCAGGGCGCTAATGGGGATTCCCGCTGCCGGAGG
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     GGGCGGGGTTCCCGTGGCTGACGGCGTGGGCCGGAGGGCTGGGGGCTGGTGCTAGTGTCTCGTCAGCGCTGGTCC
MI3R        ..........................................................................
rep30120R   ..........................................................................

plasmid     GGCCGGGCTTGCTCGGCGACGGGAGCGGAGGGAATGCTTTCATGCTTCAGGCTCCATCTC-GGGAAAAGGAGAGC
MI3R        ...........................................................G..............
rep30120R   ...........................................................-..............

plasmid     CCTTCACACAACATGGGCTATCAAAGGCTAGGTTGTGTCGCTGCCGGCAGCCAAGCGGCGGGACGATTTTCTCCT
MI3R        ..........................................................................
rep30120R   ..........................................................................
```

```
plasmid      GGCGGCCGGACAGGCGTGCGCCTGCGAAGTGGCCGATCCGGCAGCCGGACCATCCGGAGCGGCGCCATGTCGGCT
MI3R         ...........................................................................
rep30120R    ...........................................................................

plasmid      CCGAGGTGTGCGTGATGGTCGGTCGCGGCGACAGCGCAGGACATGGCTCGGTAGGCCATCGATCCCCGCTACATG
MI3R           ..-.A.........-.........CG...-------------------------------------------
rep30120R      ...........................................................................

plasmid        TGCAGACTTTTCCTGAGGGAGTTTCTCAGGGCGAGCAGTTCGGGAAAGACGGAGAACAGGATGCGTTTGACTT
rep30120R      ..................................................A........................
sjekk30120F    -------------------------------------------------.GAGGG.ACCG.A.CG.TGA..
sjekk30120R    ------------------------------------------------GA..A.....................

plasmid        TCCGGCGTTCGATTTTCGTCCTTGCGAACAGCATCACCAGTCTGCCCCTCACGTGCAGGGGGAGTCTGCCCGC
rep30120R      ..................................................C....------------
sjekk30120F    .........GAT...---GT.C.........-......G--T.TG....T...G...-..............
sjekk30120R    ...........................................................................

plasmid        GTAGGTCAGCTTGCAGGTTTCCAGGGCGCGGAGGAGGTTGAGCTGGACGAACTGGTTCAAGGCCGCTCCTGCT
sjekk30120F    ..............-............................................................
sjekk30120R    ...........................................................................

plasmid        TCGCTGTTTTCCGGGGGCAGCTTGATCTTCAGGAGTTCCTGCGCGACGACGAGATAGCCGTTTATATGCCTTT
sjekk30120F    ...........................................................................
sjekk30120R    ...........................................................................

plasmid        TCTGCTTCTTCTGCGTCATGATCGACTCGGGTCGGACACGCCGGTTGTAAAGCCTGTCCGGAATGCAGGCGAT
sjekk30120F    ...........................................................................
sjekk30120R    ...........................................................................

plasmid        CCTGGCATGTTCCTTTTCGAGGAGAAGTCTGGTCGTGAACAGATTGTCCTCGTGGACTATCTCCGGATAGAACCT
sjekk30120F      .........................................................................
sjekk30120R      .........................................................................

plasmid        CAGGCCTTCGAATTCCTTTCTCCGGTACATGTAGAGATAGGGCATCATCAGATAGCTCTTGAGCCGGATCGACTG
sjekk30120F      .........................................................................
sjekk30120R      .........................................................................

plasmid        AATGAAGAAGGTATGTGCCGGAATGTAGGCGCCGTGCAGGTCTGGGACTCTTTCGCCTCTGAAGTTTTCTCTCAG
sjekk30120F     ..........................................................................
sjekk30120R    ...........................................................................

plasmid        GCTCTCGTCCACTCCATCGAAGAAGACAGTCGATGCAAAGAAGACCATGTCCACCTGCTGGCGCCCGAGAATTTC
sjekk30120F    ...........................................................................
sjekk30120R    ...........................................................................

plasmid        CACGCAGGTGGATATGGCCTGTCTTTCCAGAAAGTCATCGCTGTCGACGAACAGAATGTATTCGCCTGTGGCCAA
sjekk30120F    ...........................................................................
sjekk30120R    ...........................................................................
```

```
plasmid      TTCGAGGCCGGTGTTTCTGGCGATCGATACGCCCTGGTTTTTTGTGCTCCTGACCAGCAGGTTGCCGTGCCTGGC
sjekk30120F  ............................................................................
sjekk30120R  ............................................................................

plasmid      ACGGAATGCCTCGAGTATTTCCAGGCTGGTATCCGTGGAGCCGTCATCGACGGCGATGA-TCTCCAGGTTGTCGT
sjekk30120F  ............................................................................
sjekk30120R  ............................................................................
MI3F         ------------------------------------------------.........T...............

plasmid      AGTCCTGGTTGACGGCCGAA-TCCAGGCATTGCGCAAGAT-AAGGTGCCACGTTGTAGACGGGAATGATGACGCT
sjekk30120F  ....................-...................-................................
sjekk30120R  .......T............-...................-................................
MI3F         ....................T...........G.......A........C.........................

plasmid      GACCTTTTCCGGAGCCATGGT-GGTTTCCTTGTGCCAAGAATGAAGTGGGATATCAGGGTGCGCCTGCCTGGTTG
sjekk30120F  ............................................................................
sjekk30120R  ...................C.G.....................................................T...
MI3F         ............................................................................

plasmid      ATTCAGGCATTGAAGAACGGCACCGGTATCGACGGCGGCAGGCTGACAACATGAACTTCACGAGGGCACTTCCAA
sjekk30120F  ............................................................................
sjekk30120R  .A.....-....G......------------------------------------------------------
MI3F         ............................................................................

plasmid      ATACTCGCCTCCGAGCGTGCCCGGAACAGGCCGCGCAGGCTTTCGGTGCCATGATCGGAACGAAGGGTTTTGGAA
sjekk30120F  ............................................................................
sjekk30120R  ------------------------------------------------------------------------
MI3F         ............................................................................

plasmid      GTTCCCTTACAGGGATATCCATGGAGGCGGACAGGATAATGCATTGGAAAGCCGTAAATGAAATCAGGCTCGGAT
sjekk30120F  .-............------------------------------------------------------------
sjekk30120R  ------------------------------------------------------------------------
MI3F         ............................................................................

plasmid      GTTTTGCCGAGACGATCTTTGCTTTTCCATGGCCAATCAACTTCGAGAGTCGTGGAGGTCTATCTGGGAGTTTTC
sjekk30120F  ............................................................................
sjekk30120R  ------------------------------------------------------------------------
MI3F         ............................................................................

plasmid   AGGATCTCCATGGATTGCAATGTTCTAATAAGCTTTCTGGTGATTTTCTTCATGATGGGCTGTTCGACGGGAGGC
MI3F      ..........................................................................

plasmid   TTGCGCTCGGGGCGTGCGGCAAGCAGGGCAAGCCCATGTGGGTAGTGTCGGTATAGCATGTTCGTGGCCTGTTGG
MI3F      ..........................................................................

plasmid   AGTTCCTCCCTGTATTCGGAAAGCTCTTCAAGGTTGTTGATGTTGTAATTCCCCCATCTCAGTTGGGTTCTGTAC
MI3F      ..........................................................................

plasmid   ATTTCGGTAATATTTTCCAGCGCACCGGCGCCCAGCTCGGAAAATCTGCTGACGGCGACGGCATAGGCTCTATAG
MI3F      ..........................................................................
```

```
plasmid  GTCTCCACATATGCCTTGGGGCGACCAGTGATCGACGTGGCGCTGGTTCGGTTGTGGGTATAGTCGTGAAGAATA
MI3F     ........................................................................

plasmid  TCTTTCATGAGGTAAATCGTGGACGCCCGTGAAAAGCAGCCTATGTTCAGCAGCAGGTCTTCATTGAGATTCTGT
MI3F     ........................................................................

plasmid  CGCCAGGGAAAGTCCATTTCCTTGAACGGTAGGATGATTTCCCTTCGATAGAGCTTGTTCCACAGCATTCCGGTT
MI3F     ........................................................................

plasmid  CCGAATTCGAAGCGGCAAAACCTGCCAAATACATCTGTACTTATTTCGGTGTTCTCAACGAAACTGATTTTCGTG
MI3F     ........................................................................

plasmid  CCCAGGCGTTCTCCAGACTCGGCTACGCGATATGAGCCACACACGACTATATCGGACAGGTGCTCGCTGGCCACC
MI3F     ........................................................................

plasmid  TCATGCATCGAACCGAACATGTCGGGATAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCA
MI3F     ........................................................................

plasmid  TGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGA
MI3F     ........................................................................
```

Bibliography

Argast, M., Beck, C., 1984. Tetracycline diffusion through phospholipid bilayers and binding to phospholipids. Antimicrobial agents and chemotherapy 26, 263–265.

Azif, A., Mohsin, H., Tanvir, R., Rehman, Y., 2017. Revisiting the mechanisms involved in calcium chloride induced bacterial transformation. Frontiers in microbiology 8, 2169.

Ertesvåg, H., 2015. Alginate-modifying enzymes: Biological roles and biotechnological uses. Front Microbiol 6, 523.

Gay, P., le Coq D., Berkelman, T., Kado, C.I., 1999. Positive selection procedure for entrapment of insertion sequence elements in gram-negative bacteria. J Bacteriol 31, 1105–1113.

Gimmestad, M., Ertesvåg, H., Heggeset, T.M.B., Aarstad, O., Svanem, B.I.G., Valla, S., 2009. Characterization of three new *Azotobacter vinelandii* alginate lyases, one of which is involved in cyst germination. J. Bacteriol 191, 4845–4853.

Guay, G., Rothstein, D., 1993. Expression of the tetk gene form *Staphylococcus aureus* in *Eschericha coli*: comparison of substrate spesificities of teta(b), teta(c), and tetk efflux protein. Antimicrobial Agents and Chemotherapy 37, 191–198.

Hillen, W., Berens, C., 1994. Mechanisms underlying expression of tn10 encoded tetracycline resistance. annual review of microbiology 48, 345–369.

H.L., S., 1975. Encystment and germination in *Azotobacter vinelandii*. Bacteriology reviews 39, 516–539.

Jensen, H., 1954. the *Azotobacteriaceae*. Bacteriol rev 18, 195–214.

Jeong, J.Y., Yim, H.S., Ryu, J.Y., Lee, H.S., Lee, J.H., Seen, D.S., Kang, S.G., 2012. One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies. Applied and Environmental Microbiology 78, 5440–5443.

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D., Rehrauer, W.M., 1994. Biochemistry of homologous recombination in *Escherichia coli*.

Larsen, B., Haug, A., 1971. Biosynthesis of alginate: Part ii. polymannuronic acid c-5-epimerase from *Azotobacter vinelandii*(lipman). carbohydrate research 17, 297–308.

Lin, L.P., Sadoff, H.L., 1968. Encystment and polymer production by *azotobacter vinelandii* in the presence of β-hydroxybutyrate. J. Bacteriol 95, 2336–2343.

Lin, L.P., Sadoff, H.L., 1969. Chemical composition of *Azotobacter vinelandii* cysts. J. Bacteriol 100, 480–486.

Lipman, J.G., 1903. experiments on the transformation and fixation of nitrogen by bacteria. Report of the New Jersey agricultural experiment station 24, 217–285.

Livermore, D.M., 1995. β-lactamase in laboratory and clinical resistance. clinical microbiology review 8, 557–584.

Llosa, M., gomis rüth, F.X., Col, l.M., Cruz, F.d., 2002. Bacterial conjugation: a two-step mechanism for DNA transport. Molecular Microbiology.

Lorenzo, V., Eltis, L., Kessler, B. Timmis, K., 1993. Analysis of pseudomonas gene products using lacp/ptrp-lac plasmids and transposons that confer conditional phenotypes. gene 123, 17–24.

Marger, M., Saier, M., 1993. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends in biochemical sciences 18, 13–20.

Marqués, S., Manzanera, M., González-Pérez, M.M., Gallegos, M.T., Ramos, J.L., 1999. The xyls-dependent pm promoter is transcribed in vivo by rna polymerase with σ32 or σ38 depending on the growth phase. Molecular Microbiology 31, 1105–1113.

Nagpal, P., Jafri, S., Reddy, M.A., 1989. Multiple chromosomes of *Azotobacter*. J. Bacteriol 171, 3133 – 3138.

Noar, J.D., Bruno-Bárcena, J.M., 2018. *Azotobacter vinelandii*: the source of 100 years of discoveries and many more to come. Microbiol 64, 421–436.

Page, W., Sadoff, H., 1975. Relationship between calcium and uronic acids in the encystment of *Azotobacter vinelandii*. J. Bacteriol 122, 145–150.

Pedersen, J.L., 2019. Characterization of avin05390 and avin30120 effect on cyst formation in *Azotobacter vinelandii*. TBT4500, NTNU .

Perri, S., Helinski, D., 1993. Dna sequence requirements for interactions of the rk2 replication initiation protein with plasmid origin repeats. j.biol.chem 268, 3662–3669.

Royo, J., Moreno-Ruiz, E., Cebolla, A.S.E., 2005. Stable long-term indigo production by overexpression of dioxygenase genes using a chromosomal integrated cascade expression circuit. Journal of biotechnology 116, 113.124.

Rühmann, B., Schmid, J., Sieber, V., 2015. High throughput exopolysaccharide screening platform: from strain cultivation to monosaccharide composition and carbohydrate fingerprinting in one day. carbohydr polym. 122, 210–220.

Sanger, F., Nicklen, S., Coulson, A., 1977. Dna sequencing with chain-terminating inhibitors. proc.natl.acad.sci 74, 5463–5467.

Setubal, J.C., dos Santos, P., Goldman, B.S., Ertesvåg, H., Espin, G., Rubio, L.M., Valla, S., Almeida, N.F., Balasubramanian, D., Cromes, L., Curatti, L., Du, Z., Godsy, E., Goodner, B., Hellner-Burris, K., Hernandez, J.A., Houmiel, K., Imperial, J., Kennedy, C., Larson, T.J., Latreille, P., Ligon, L.S., Lu, J., Mærk, M., Miller, N.M., Norton, S., O'Carroll, I.P., Paulsen, I., Raulfs, E.C., Roemer, R., Rosser, J., Segura, D., Slater, S., Stricklin, S.L., Studholme, D.J., Sun, J., Viana, C.J., Wallin, E., Wang, B., Wheeler, C., Zhu, H., Dean, D.R., Dixon, R., Wood, D., 2009. Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes. j bacteriol 191.

Shuman, S., 1991. Recombination mediated by vaccinia virus dna topoisomerase i in *Escherichia coli* is sequence specific. Proceedings of the National Academy of Sciences 88, 10104–10108.

Socolofsky, M., Wyss, O., 1961. Cysts of azotobacter. J Bacteriol 81, 946 – 954.

Socolofsky, M., Wyss, O., 1962. resistance of the *Azotobacter* cysts. J. Bacteriol 84, 119–124.

Steigedal, M., Sletta, H., Maerk, M., Christensen, B., Bjerkan, T.E., Ellingsen, T., Espin, G., Ertesvåg, H., Valla, S., 2008. *Azotobacter vinelandii* alge mannuronan c-5-epimerase family is essential for the in vivo control of alginate monomer composition and for functional cyst formation. environ. microbiol. 10, 1760–1770.

Stevenson, L., Socolofsky, M., 1966. Cyst formation and poly-β-hydroxybutyric acid accumulation in *Azotobacter*. J. Bacteriol 91, 304–310.

Wyss, O., Neumann, M.G., Socolofsky, M.D., 1961. Development and germination of the *Azotobacter* cyst. The Journal of Cell Biology 10, 555–565.