Isolation and characterisation of genes encoding biopolymer manufacturing enzymes

by

Telana Rapp

Thesis presented in partial fulfilment of the requirements of the degree Master of Science in Plant Biotechnology at Stellenbosch University

Supervisor: Dr. Jan Bekker

Co-supervisor: Prof. Jens Kossmann

Faculty of Agricultural Sciences

Department Genetics

Institute for Plant Biotechnology

March 2012

Declaration

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

T Rapp

March 2012

ABSTRACT

Biopolymers exhibit the required material properties to replace conventional, non-biodegradable, petroleum-based polymer products. They have a closed carbon cycle, making them carbon neutral and environmentally friendly. Biopolymers are produced from non-toxic substrates during *in vivo* enzymatic reactions. Biosynthesis of the most commercially important biopolymers is too complex to be reproduced in *in vitro* reactions. Identification of the genes responsible for their biosynthesis has been under investigation, with some pathways already elucidated. The genes involved in the biosynthesis of these polymers have been targeted for genetic manipulation to increase productivity, as well as create tailor-made polymers. Novel biopolymers and the genes responsible for their synthesis are of interest for their potential commercial applications. Bacteria produce a wide range of biopolymers and are being implemented as the bio-factories for biopolymer production. They are capable of utilising easily accessible and renewable carbon sources such as sucrose for polymer biosynthesis. Bacteria thus allow for economical production of these environmentally beneficial polymers.

In this study, the gene responsible for the production of an unknown biopolymer from an unknown bacterium was identified. The biopolymer producing bacteria were grown on media enriched with sucrose as carbon source, during an expression library screening in a previous study. Expression library technology was used to search for the gene and it was identified as a 424 amino acid levansucrase which had a 100% homology to *Leuconostoc mesenteroides* M1FT levansucrase (AAT81165.1). Biopolymer analysis revealed that the biopolymer was a levan, a polysaccharide consisting of only fructose molecules with a molecular weight of ± 5 kDa. Analysis of a 516 bp fragment of the 16S rRNA determined that the unknown bacteria were a *Pseudomonas* species.

SAMEVATTING

Bio-polimere besit noodsaaklike materiële eienskappe wat toelaat dat dit konvensionele, nie bio-afbreekbare, petroleum-gebasseerde polimeer produkte kan vervang. Hulle het n geslote koolstof kringloop en is dus koolstof neutraal en omgewingsvriendelik. Bio-polimere word vervaardig van nie-toksiese substrate, gedurende ensiematiese reaksies *in vivo*. Die belangrikste kommersiële bio-polimere se ensiematiese produksie is te kompleks om in 'n *in vitro* reaksie te herproduseer. Ondersoeke tot die identifikasie van die gene wat verantwoordelik is vir die produksie van die polimere is onderweg, en sommige produksie paaie is reeds bekend. Die bekende gene word geteiken vir genetiese manipulasie om hulle produktiwiteit te vermeerder en om unieke polimere te produseer. Unieke bio-polimere en die gene wat vir hul produksie verantwoordelik is, is van belang vir hulle potentiële implimentering in komersiële toepassings. Bakteria produseer 'n verskeidenheid bio-polimere en word as die bio-fabrieke vir polimeerproduksie geimplimenteer. Hulle kan maklik bekombare koolstofbronne, soos sukrose, gebruik om bio-polimere te produseer. Bakteria laat dus die ekonomiese produksie van hierdie omgewingsvriendelike polimere toe.

In hierdie studie word die geen wat verantwoordelik is vir die produksie van 'n onbekende bio-polimeer van 'n onbekende bakteria, geidentifiseer. Die bakteria was gevind op media, wat verryk was met sukrose as koolstofbron, tydens 'n vorige studie, waartydens 'n uitdrukkingsbiblioteek gesif was op hierdie media. Uitdrukkingsbiblioteek tegnologie was gebruik om die geen te vind. Die geen was geidentifiseer as 'n 424 aminosuur, homo-fruktose-polimeer produseerende geen, 'n "levansucrase". Die geen het 'n 100% homologie met die M1FT "levansucrase" geen (AAT81165.1) van *Leuconostoc mesenteroides* gehad. Analise van die bio-polimeer het bepaal dat die polimeer 'n polisakkaried was, wat slegs uit fruktose molekules bestaan het. Die molekulêre gewig van die polimeer was ± 5 kDa. Analise van 'n 516 bp fragment van die 16S rRNS het bepaal dat die bakteria van die *Pseudomonas* spesie afkomstig was.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Isolation and characterisation of genes encoding biopolymer manufacturing enzymes

1 GENERAL INTRODUCTION

The global movement to cleaner and greener living, to decrease the rate at which global warming, pollution and landfill are affecting the world, has spiked the recent interest in the production of environmentally friendly consumer products. Biopolymers are naturally produced macromolecular molecules synthesised by all living organisms. They consist of smaller, linked molecules and are essential as structural, energy storage and protection molecules. They are biodegradable, renewable and have a closed carbon cycle, making them carbon neutral (Koller *et al*., 2010).

Biopolymers are produced from natural, non-toxic substrates, often making them biocompatible. Biocompatibility is the inability of the biopolymer to elicit an immune response, conceptually they can be used in medical applications such as drug delivery, tissue engineering and wound dressing (Koller *et al*., 2010; Rehm, 2010). The ever-growing success in the application of microbially produced biopolymers such as xanthan gum and dextran has increased interest in finding novel biopolymers that could be used for the same, or different and new applications (Sutherland, 1998). The diversity and pliability of biopolymers make them suitable for a wide range of applications in the medical, food and industrial sectors.

Novel biopolymer producing genes can be identified, isolated and characterised from any biological biopolymer producing source. These genes can be introduced into heterologous host bacteria such as *Escherichia coli*, which have already been used for recombinant bacterial biopolymer production (Chen, 2009). The recombinant *E. coli* can then be used as the bio-manufacturers of the novel biopolymer in vast quantities in a time and cost effective manner.

The endless properties biopolymers boast make them appealing in replacing non-renewable and non-biodegradable, oil-based polymer products (Koller *et al*., 2010; Rehm, 2010). This replacement will decrease the tremendous amount of space required to house the non-biodegradable, non-renewable waste being produced by the 7 billion people

9

inhabiting the earth, as well as contributing to the efforts to alleviate the pace of global warming.

2 LITERATURE REVIEW

2.1   BIOPOLYMERS

Biopolymers are macromolecular structures consisting of multiple, linked molecules which are produced during enzymatic reactions in a biological system. They can consist of identical molecule repeats or from different molecules, respectively known as homo- and hetero-biopolymers. They may also be simple single chain or complex branched chain biopolymers. The chemical and structural composition of biopolymers give them their unique characteristics and determines their biological function as well as their material properties (Tao *et al*., 2007; Yang and Zhang, 2009). Medical grade biopolymers are used for *in vivo* medical applications or in close proximity to biological fluids.   Specific bio-relevant analysis (BA) has been established to determine whether biopolymers of interest would be suitable for various medical applications (Jaffe *et al*., 2003).

Biopolymers are produced from non-toxic substrates and are fully biodegradable. They degrade completely to carbon dioxide (CO_2) and water (H_2O) when exposed to any given environment with microbial flora (Koller *et al*., 2010).   The microbial flora produce extracellular polymer hydrolysing enzymes (depolymerases and hydrolases) which break down the biopolymer backbone and produce smaller polymer fragments which can be taken up by the microbial cells as a carbon and energy source (Rehm, 2011).

Various naturally occurring biopolymers and derivatives of these biopolymers exhibit material properties which are capable of filling the requirements necessary for various industrial, medical and food applications (Crini, 2005; Persin *et al*., 2011; Yang and Zhang, 2009). Investigation into economical, mass manufacturing of biopolymers with material properties has been set afoot and various systems have been identified. Post-production modification of agro-industrial produced plant biopolymers with chemical treatment can manipulate these plant biopolymers to the desired chemical composition and structure (Persin *el al*., 2011). Cell free systems are systems in which purified enzymes are used in catalysed reactions to produce the desired polymer (Perugino *et al*., 2004).   Whole cell

systems are systems incorporating living cells as the bio-factories for the production of the desired biopolymer (Ruffing and Chen, 2006).

The non-commercially exploitable agro-industrial derived waste biopolymers could be transformed into polymers with material properties by treating them with enzymes or chemicals to manipulate the polymer to a desired structure or composition. The biopolymers have to be isolated and purified before these treatments can commence. Isolation and purification processes can be extremely time consuming and require the use of hazardous chemicals. The chemical or enzymatic treatment to conform the biopolymer to the desired structure and composition may only affect parts of the polymer which are in direct contact with the modifying agent (Persin *et al.*, 2011). After transformation of the biopolymer, it is necessary to isolate and purify the modified biopolymer. This process of purification, chemical or enzymatic transformation, followed by another purification step, can be time consuming and expensive. Transformation is also dependent on the modifying agents available and creates a limitation in the diversity of the end-product.

In a study by Crini (2005), chemical treatment of chitin, chitosan, starch and cyclodextrin, which was obtained as industrial waste, increased the absorbance and molecular specificity for absorption of these biopolymers. These modified biopolymers could then be used in the purification of toxic waste water by absorbing polluting molecules from the waste water.

In a cell free system, the enzyme or enzymes responsible for the biosynthesis of the biopolymer are used in an *in vitro* controlled and catalysed reaction to produce the polymer of interest (Perugino *et al.*, 2004). Cell free synthesis requires the large scale production of high purity biosynthetic enzymes and in some cases expensive cofactors such as energy and metal cofactors (Perugino *et al.*, 2004). In some instances, more than one enzyme is required to produce the desired polymer and the reactions with the different enzymes will have to be done separately. Separate reactions may be required as the optimal conditions and cofactors for enzyme efficiency may vary. These extra steps add more time consuming elements of isolation and purification. This approach to polymer production is both time consuming and costly and thus not economically viable.

Whole cell systems incorporate living cells or whole organisms for the biosynthesis of the biopolymer (Ruffing and Chen, 2006). The optimal conditions, as well as cofactors required for enzyme activity, are made available by the cell. The use of bacteria is preferred in this system, because they grow at a fast rate and are efficient at producing biopolymers without

hindering bacterial growth. Batch fermentations are preferred for large-scale production of bacterial biosynthesised biopolymer, as they assure reproducibility and supply (Luong *et al.*, 1988). Microbially produced biopolymers offer the possibility of continued biopolymer production in bioreactors. The rate at which the biopolymer is being produced can also be controlled and thus an increase in consumer demand can be supplemented at an acceptable rate. Simple kinetic models have been established that determine biomass, product and substrate evolution time for various well studied exopolysaccharide (EPS) producing batch cultures (Luong *et al.*, 1988).

The appeal of the bacterial whole cell system lies in the turnover time of substrate to biopolymer, as well as the decrease of post-production chemical treatment. Bacteria are also capable of utilising carbon sources unrelated to the substrate required for biopolymer production and synthesise the appropriate substrate, effectively reducing the cost of expensive substrates which would be required in cell free systems. A prime example is that of a glycerol by-product which is used to synthesise substrate by *Pseudomonas oleovorans* to produce a polysaccharide biopolymer (Alves *et al.*, 2011). Recombinant DNA technology can be used to identify possible novel biopolymer synthesising genes in a bacterial system. It is also possible to produce tailor-made biopolymers in bacteria by genetically engineering the genes involved in the biosynthesis pathway. These modifications will equip the bacteria with favourable traits in novel-, tailor-made- and increased biopolymer production. Novel biopolymer producing genes can be derived from various natural biological systems and can be incorporated into bacteria to synthesise these biopolymers at extremely fast rates. Isolation and purification of microbially produced biopolymers in controlled fermentation reactions is easier than for naturally acquired biopolymers.

2.2 MICROBIAL BIOPOLYMERS

Bacteria produce both intracellular and extracellular biopolymers. The intracellular biopolymers are very limited when compared to the wide range of extracellular biopolymers bacteria can produce (Rehm, 2010). These extracellular biopolymers are produced effectively without inhibiting normal cellular growth. Some bacteria are capable of producing vast quantities of extracellular biopolymer, up to 188 g/L of culture (Ruffing and Chen, 2006). Microbial biopolymer production is linked to the primary metabolic pathways as shown in Figure 2.2.1 (page 13).

The biological functions of some of these biopolymers are retained for a wide range of prokaryotes (Anderson *et al*., 1990). Biopolymers can also be specifically produced by certain bacteria and have specific functions (Rehm and Valla, 1997). Many bacterial species are capable of synthesising several biopolymers, such as *Pseudomonas aeroginosa* which is known to produce alginate, rhamnolipids and polyhydroxyalkanoic acids (PHAs) (Pham *et al*., 2004; Ryder *et al*., 2007). There are four major classes of biopolymers which are produced by bacteria: polysaccharides, polyesters, polyamides and inorganic polyanhydrides (Rehm, 2010).
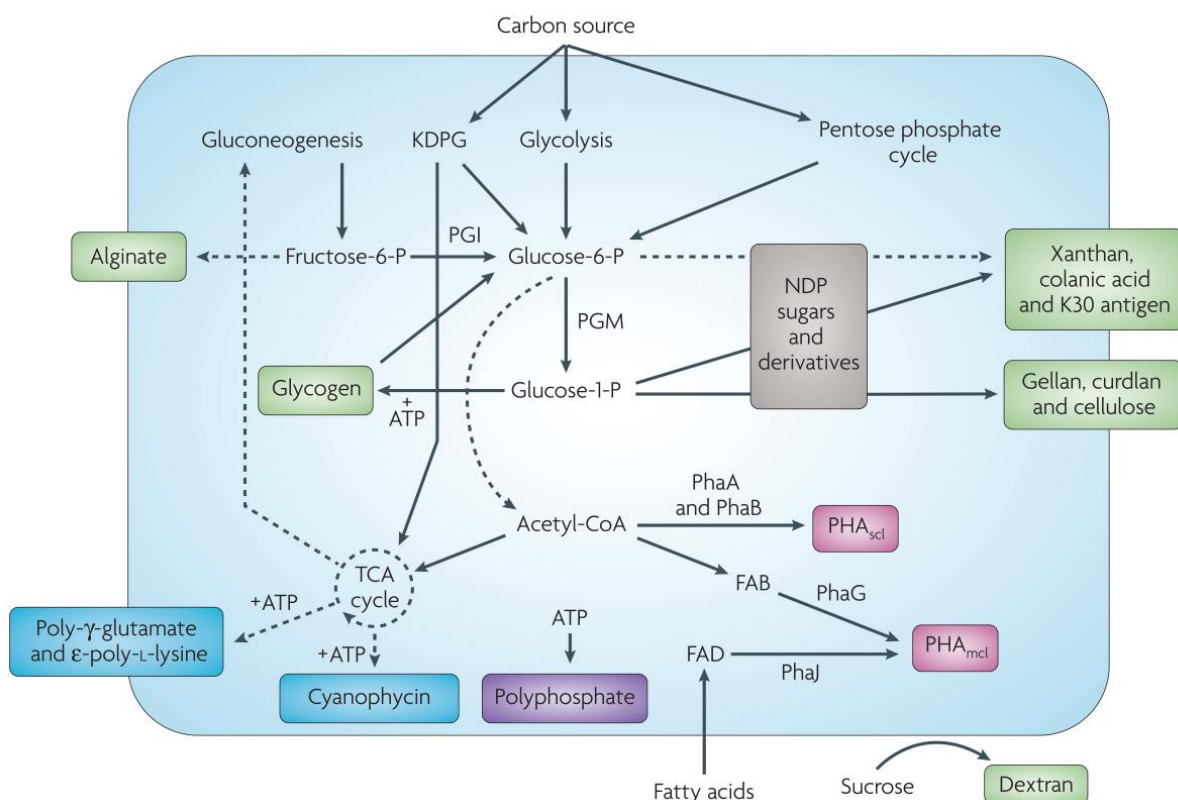
Figure 2.2.1: Schematic representation of bacterial biopolymer biosynthesis pathways from intermediates of primary metabolism (Rehm, 2010).

Polyesters such as polyhydroxyalkanoic acids (PHA) are produced in bacteria as an intracellular carbon storage molecule (Rehm, 2010). Accumulation of PHA is highly regulated by the biosynthesis genes during unfavourable growth conditions during which a large supply of a suitable carbon source is available but there is an insufficient amount of other nutrients, such as phosphorous and nitrogen (Anderson *et al*., 1989; Kessler and

Witholt, 2001). Classification of PHAs are achieved by the carbon chain length of their constituents and can be short-chain-length (3-5 carbons), medium-chain-length (6-14 carbons) or long-chain-length (>15 carbons) (Koller *et al.*, 2010).

The PHA is synthesised in an intracellular, insoluble spherical inclusion which has a hydrophobic PHA core and is mainly surrounded by the PHA metabolism enzymes (Grage *et al.*, 2009). Over a 150 different constituents have been discovered, rending the diversity in the material properties of the PHAs tremendous (Rehm, 2010). The crystalinity and melting temperatures of PHAs correlate to those of oil-based plastics and offer a biodegradable and renewable alternative to these oil-based plastics (Koller *et al.*, 2010). These PHAs are also commonly referred to as bacterial bioplastics and are already being produced in large-scale bacterial fermentations which are producing around 100 000 tonnes per year (Chen, 2009).
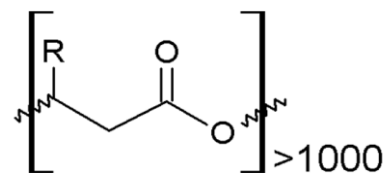
Figure 2.2.2: Chemical structure of polyhyrdoxyalkanoate (PHA). R =alkyl chain of 1-11 carbon atoms (Rehm, 2010).

Polyamides are biopolymers consisting of amino acid subunits which are attached by amide bonds (Oppermann-Sanio and Steinbüchel, 2001). Micro-organisms are known to produce three types of polyamides, which are produced enzymatically and independently of ribosomal protein biosynthesis (Oppermann-Sanio and Steinbüchel, 2001). Poly-γ-glutamate (PGA) and ε-poly-L-lysine (PL) are extracellular polyamides and cyanophycin granule peptide (CGP) is an intracellular polyamide. The polyamides PGA and CGP with chemically reduced arginine content have similar material and chemical properties to chemically synthesised polyacrylates (Oppermann-Sanio and Steinbüchel, 2001; Joentgen *et al.*, 1998). Polyacrylates are used as dispersants, antiscalants or superabsorbers. Bacterial polyamides could be used as an alternative, renewable, non-toxic and biodegradable polymer for the same applications as polyacrylates (Rehm, 2010).
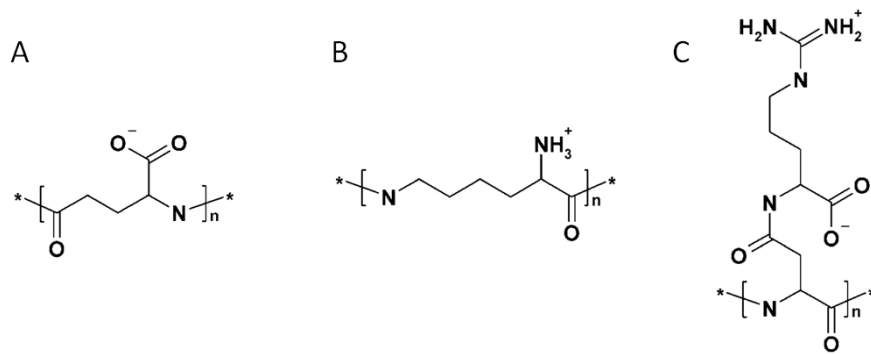
Figure 2.2.3: Chemical structures of known polyamides A: poly-γ-glutamate (PGA), B: ε-poly-L-lysine (PL) and C: cyanophycin granule peptide (CGP) (modified from Oppermann-Sanio and Steinbüchel, 2001).

The only polyanhydride found in living organisms is inorganic polyphosphate, which has various industrial applications, such as the replacement of asbestos, as a flame retardant and as a food flavour enhancer (Rehm, 2010). The commercial production of polyphosphate by bacteria is not economically feasible as polyphosphate is sufficiently produced from the dehydration of naturally occurring rock phosphate (Rehm, 2010). Bacteria which are capable of producing polyphosphate from high exogenous concentrations of phosphate could be incorporated in the treatment of nitrified and waste water treatments to reduce high phosphate levels and produce polyphosphate as a by-product (Rehm, 2010).
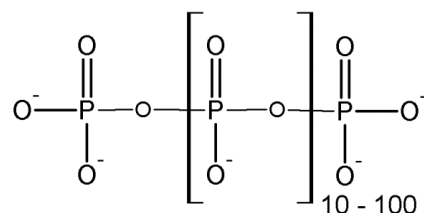
Figure 2.2.4: Chemical structure of polyphosphate (Rehm, 2010).

Bacteria produce a wide range of polysaccharides, which can be subdivided into three types: exo-polysaccharides (EPS), capsular-polysaccharides and intracellular-

polysaccharides (Rehm, 2010). The most commercially applied microbial polysaccharides are EPS, with xanthan being the most important with applications in the industrial, food and pharmaceutical industries. These applications include gelling agent in explosives, foam stabiliser in beer and retarded drug release (Becker *et al.*, 1998). The three types can be further categorised into repeat unit-, repeating- and non-repeating polymers (Becker *et al.*, 1998; Remminghorst and Rehm, 2006; Whitfield, 2006). The biosynthesis of the different types of polysaccharides requires the functions of different genes. These genes are varied for the various polysaccharides and are organised in biosynthesis gene clusters (Rehm, 2010). This study focussed on polysaccharides, which will be discussed in more detail in the next section.

2.3 MICROBIAL POLYSACCHARIDES

The expression of the biosynthesis gene clusters (operons) for polysaccharides is highly regulated (Rehm, 2010, Wilson *et al.*, 2010). Various regulatory elements have been identified and have similarity in a wide range of bacteria. Examples of such extensive regulation processes are two-component signal transduction pathways, which are found in most bacterial and archaeal species (Castañeda *et al.*, 2001). During this process, phosphotransfer between two conserved components regulates transcription. Such systems include the GacS/GacA system (Castañeda *et al.*, 2001). The concentration of cyclic diguanylate (c-di-GMP) regulates the expression of genes involved in EPS production, during increased concentration; c-di-GMP binds to the FleQ transcriptional regulator to alleviate repression of expression (Hickman and Harwood, 2008). Quorum-sensing (QS) regulates gene expression by the production of a higher concentrations of 3-oxo-dodecanoyl (3-O-C$_{12}$) homoserine lactone during increased population density which forms a complex with the transcriptional regulator LasR. LasR activates transcription of various genes, including the expression of the *pel* operon in *Pseudomonas aeruginosa* which encodes the genes required for the synthesis and secretion of a biofilm associated glucose-rich polysaccharide (Sakuragi and Kolter, 2007). Alternative RNA polymerase σ-factors and anti σ-factors, as well as integration host factor (IHF)-dependant processes also exist in the regulation of gene expression for the production of bacterial polysaccharides (Ionescu and Belkin, 2009; Leech *et al.*, 2008; Weber *et al.*, 2006).

Exo-polysaccharides are found on the extracellular regions of the cell and can either be attached to the cell wall or detached (Sutherland, 1998). The polymer plays an important role as matrix components in biofilms and expression of EPS biosynthesis genes are often correlated to the biofilm growth mode (Hall-Stoodley *et al*., 2004). Biofilms are microbial accretions which attach to a surface and are enclosed by a matrix which forms a structurally and dynamically complex biological system (Hall-Stoodley *et al*., 2004). Biofilms help maintain a type of homeostasis which ensures survival of the microbes in diverse and unfavourable environments. This renders the microbes more antibiotic- and host defence resistant, which also improves pathogenesis of biofilm-forming opportunistic pathogens (Hall-Stoodley *et al*., 2004).

Some archaeans such as *Haloferax mediterranei* produce EPS (Parolis *et al*., 1996). The production of EPS gives bacteria an amorphous slimy morphology (Sutherland, 1998). Factors determining the potential applications of EPS are molecular weight, chemical composition, branching and stiffness of the EPS (van Kranenburg *et al*., 2000). The biosynthesis of the two commercially important EPSs, xanthan and dextran, will be discussed as examples of EPS biosynthesis.

Xanthomonas campestris is a plant pathogenic bacterium from the *Pseudomonaceae* family and is responsible for the production of xanthan gum (Becker *et al*., 1998; García-Ochoa *et al*., 2000). Xanthan is a complex repeat unit exopolysaccharide consisting of ᴅ-glucosyl, ᴅ-mannosyl and ᴅ-glucuronyl acid residues in a molar ratio of 2:2:1, with variable proportions of *O*-acetyl and pyruvyl residues (Figure 2.3.1, page 18) (Becker *et al*., 1998). It was discovered in the 1950s and has since become the most extensively studied and commercially important microbial biopolymer (García-Ochoa *et al*., 2000; van Kranenburg *et al*., 1999). The unique rheological property of xanthan has made it applicable in the food, non-food and oil recovery sectors as a thickener, dispersing agent and stabiliser of emulsions and suspensions (Becker *et al*., 1998; García-Ochoa *et al*., 2000). According to a review by Sutherland in 1998, more than 20 000 tonnes of xanthan was being produced per annum. The biosynthesis and secretion of xanthan is very similar to that of O antigens and capsular-polysaccharides (Rehm, 2010; van Kranenburg *et al*., 1999). Biosynthesis of xanthan has been accepted as a model system for bacterial EPS production (Becker *et al*., 1998; van Kranenburg *et al*., 1999).

Figure 2.3.1: Chemical structure of the repeat unit of xanthan (Becker *et al*., 1998).

Xanthan biosynthesis genes are located on a 16 kb *gum* operon which consists of 12 open reading frames *gum*B to *gum*M and 2 promotors, 1 upsteam of *gum*B and the other upstream of *gum*K   (Figure 2.3.2) (Becker *et al*., 1998; Katzen *et* al., 1996; Katzen *et* al., 1998). The fist committed step in the biosynthesis of EPS is the synthesis of the activated polymer precursors and is accomplished by pyrophosphorylases, such as UDP-glucose pyrophosphorylase and dehydrogenases, such as UDP-glucose dehydrogenase (Becker *et al*., 1998; Ruffing and Chen, 2006).

Figure: 2.3.2: Xanthan biosynthesis (*gum*) operon of *Xanthomonas campestris* pv. campestris (modified from Rehm, 2010).

Xanthan biosynthesis occurs in the following five steps; in the first step the activated precursor neucleoside diphosphate sugars (UDP-glucose and GDP-mannose) and neucleoside diphosphate sugar acids (UDP-glucuronic acid) are synthesised. In the

second step, these activated precursors donate their glycosyl residues to a polyprenol phosphate lipid carrier at the cytosolic face of the cell membrane to assemble the pentasaccharide of the repeat unit. In the third step at the lipid-linked pentasaccharide level, acetylation and pyruvylation of the mannose molecules in the repeat unit occurs. During the fourth step the repeat unit is moved to the periplasmic face of the cell membrane, where the polymerisation of the repeat units is catalysed. During the final step, the polymerised product (xanthan) is secreted to the extracellular region (Katzen *et al.*, 1998; Ielpi *et al.*, 1981; Ielpi *et al.*, 1993).
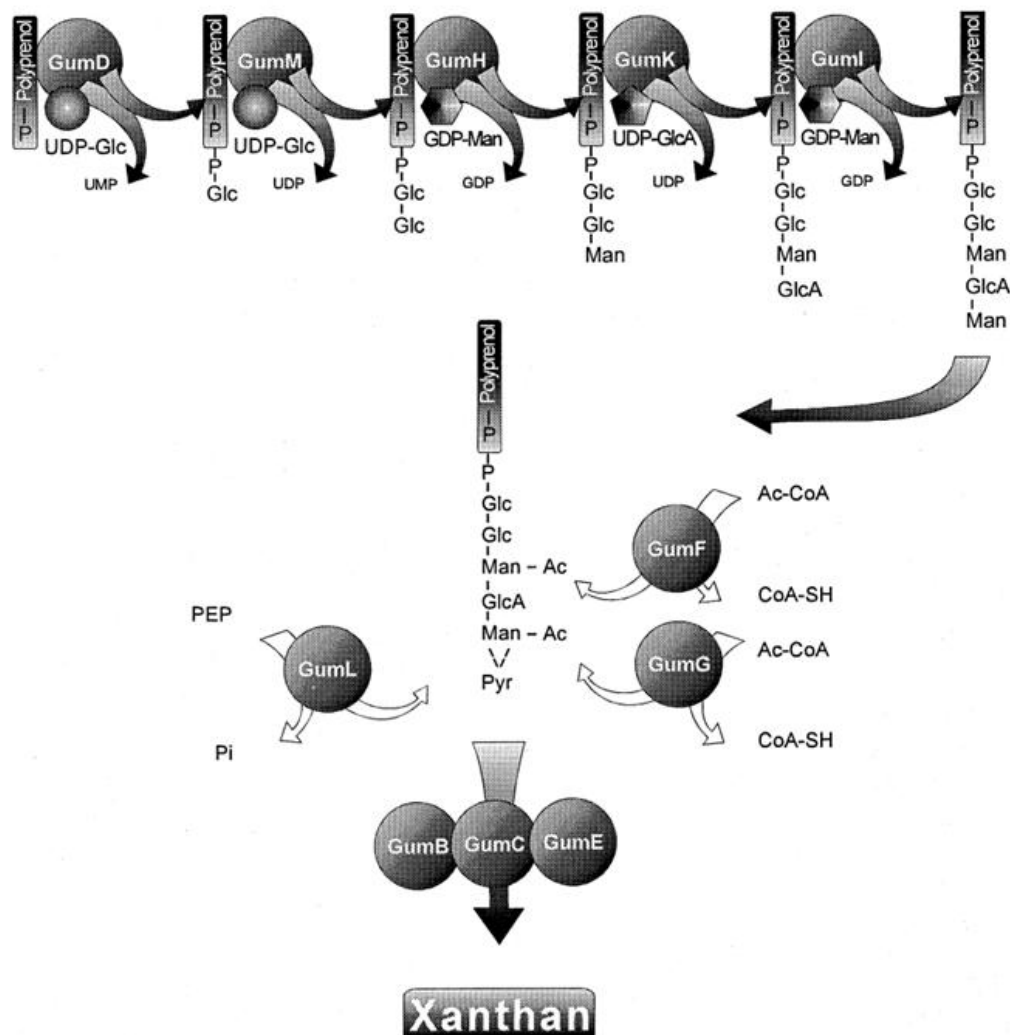

Figure 2.3.3: Biosynthetic pathway of the repeat unit exopolysaccharide, xanthan (modified from Becker *et al.*, 1998).

The biosynthesis of xanthan is depicted in Figure 2.3.3 on page 19. The first glycosyl residue transferred in the biosynthesis of the pentasaccharide is a glucosyl-1-phosphate from UDP-glucose to a polyisoprenol phosphate and this is catalysed by a glycosyltransferase (Becker *et al*., 1998). Glycosyltransferases are enzymes which require nucleotide-sugar precursors to catalyse the transfer of glycosyl residues from a donor molecule to a particular acceptor molecule (Monsan *et al*., 2001). The other gycosyl residues are added accumulatively to the already lipid-linked glucosyl residue. Repeat unit polymerisation occurs at the reducing end of the repeat units, which is similar to the synthesis of O antigen (Ielpi *et al*., 1993; Whitfield, 1995). The repeat unit polymerisation and secretion is mediated by membrane-spanning-multi-protein complexes which require a lipid carrier to necessitate the secretion of the repeat unit polysaccharide across the cytoplasmic membrane (Rehm, 2010).

Dextran is synthesised from sucrose by dextransucrase which is a glucansucrase. Transglycosylases (glycansucrases) are enzymes which catalyse the transfer of a glycosyl residue from a disaccharide sugar molecule by using the energy of the o-sidic bond in the molecule to catalyse the transfer of a glycosyl residue to an acceptor molecule (Monsan *et al*., 2001). Dextran is a homo-polymer consisting of glucose residues with α-(1,6) bonds in the linear chain and either α-(1,3) or α-(1,4) bonds at the branching sites.

The molecular weight of dextran is varied and can range from 10^6 to 10^9 Da. Chemical hydrolysis of high-molecular-weight dextran can produce dextrans with an average molecular weight of 70 kDa which is used for the production of Sephadex®, which facilitates permeation separation chromatography, as well as blood plasma substitutes (Grönwall and Ingelman, 1948; Monsan *et al*., 2001). The expression of dextransucrase is induced by the presence of sucrose. Dextransucrase is secreted and anchored to the cell wall and catalyses the biosynthesis of dextran in the extracellular region (Rehm, 2010).
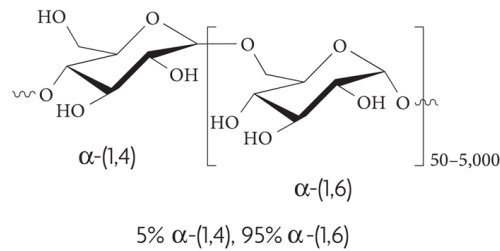
Figure 2.3.4: Chemical structure of the homo-exopolysaccharide dextran (Rehm, 2010).

Capsular polysaccharides (CPSs) are biopolymers which stay attached to the cell surface after they have been secreted. They mainly function as surface antigens and virulence factors. The biosynthesis of CPS has been extensively studied in *E. coli* and other pathogenic bacteria as the CPS enables these pathogens to resist antibiotic treatment and phagocytosis by macrophages (Arakawa *et al.*, 1995; Rahn *et al.*, 1999; Whitfield, 2006). *E. coli* produces two serotype-specific surface polysaccharides, the lipopolysaccharide (LPS) O antigen and the CPS K antigen (Whitfield, 2006). Research done on CPS biosynthesis has helped in understanding the biosynthesis pathway of xanthan as these pathways are very similar and both these types are repeat-unit polysaccharides (Rehm, 2010). There are no commercial applications for CPSs, but due to their role in bacterial virulence, extensive research has gone into these polymers and their biosynthesis to identify targets for treatments of infections of CPS producing bacteria (Øskov *et al.*, 1977; Rehm, 2010; Whitefield, 2006).

Research into the biosynthesis and secretion of CPS shows evidence of *trans*-envelope multiple-protein-complexes which mediates both the synthesis of the CPS as well as the secretion thereof (Whitfield, 2006). Repeat units of CPSs are synthesised in the cytoplasm at the plasma membrane face. The glycosyl residues of activated nucleotide sugars are transferred to a lipid (undecaprenol pyrophosphate) by glycosyltransferases to form a lipid-linked intermediate (Rahn, 1999). Complete repeat units are then transferred across the cytoplasmic membrane to the periplasm where the units are polymerised and secreted to the extracellular region where they remain anchored to the cell membrane. The means by which the CPS is attached to the cell surface has not been completely determined (Whitfield, 2006).

21
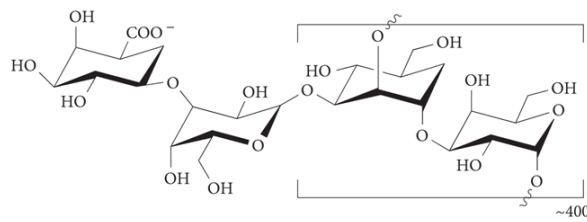
Figure 2.3.5: Chemical structure of the K30 antigen (Rehm, 2010).

Microbes synthesise intracellular structural and storage polysaccharides. These biopolymers are not used for commercial applications. The peptidoglycan known as murein is the structural biopolymer synthesised by most microbes. It is a hetero-polymer consisting of linear glycan chains which are cross linked with short peptides (Vollmer *et al.*, 2008). This biopolymer is found in the periplasm of the cells and assists in defining cellular shape, withstanding turgor and anchoring other essential envelope components such as proteins and teichoic acids (Dramsi *et al.*, 2008; Neuhaus and Baddiley, 2003).

Glycogen is the only storage polysaccharide found in bacteria and archaea. It is synthesised by glycogen synthase and glycogen branching enzyme. Glycogen synthesis occurs when there is efficient carbon available but limitations in other essential nutrient elements. It is a branched homo-polysaccharide consisting of α-(1,4) linked glucose molecules with α-(1,6) linked glucose molecules at the branching points (Figure 2.3.6) (Wilson *et al.*, 2010). The average length of bacterial glycogen branches is between 8 and 12 glucose molecules, whereas the polymer will consist of several thousand units.

Figure 2.3.6: Chemical structure of bacterial glycogen (Rehm, 2010).

2.4 NOVEL AND TAILOR-MADE BIOPOLYMERS

Novel biopolymer producing genes can be isolated from genomic and metagenomic libraries. These libraries can be screened for biopolymer production in non-biopolymer-producing heterologous hosts (Kang *et al*., 2005). *E. coli* has been found to be efficient in the production of various biopolymers which are not indigenous to the bacterium. Recombin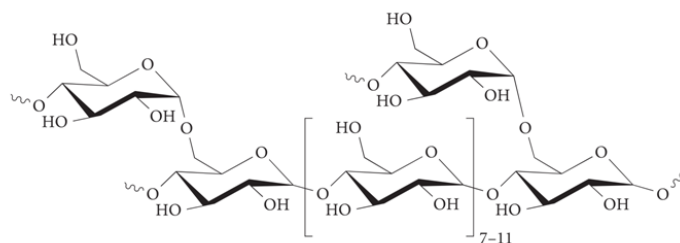ant *E. coli* has already been implemented in large scale commercial fermentations of PHA (Chen, 2009). This organism is able to utilise inexpensive carbon sources to produce the required substrates for biopolymer production. Extensive studies into biopolymer production and substrates required for the various types of biopolymers allows for library screening with specific substrates for different and selective biopolymer production.

The current understanding of biopolymer biosynthesis pathways has been accomplished by genome sequencing, functional genomics as well as cloning and characterisation of the biopolymer biosynthesis genes of commercially relevant biopolymers (Lee, 2006; Vorhölter *et al*., 2008). This knowledge has been used to improve biopolymer production by pathway reconstruction and genetic engineering, these methods also allow for the production of tailor-made biopolymers (Barreras *et al*., 2008; Taguchi *et al*., 2008). Metabolic engineering of bacterial biosynthesis pathways can even lead to the production of novel 'unnatural' biopolymers such as polylactate (PLA) and polythioesters (Doi, 2002; Taguchi *et al*., 2008; Lütke-Eversloh *et al*., 2002). New biopolymer producing strains can be engineered by rational biosynthesis pathway design as well as engineering of key enzymes to desired traits with strategies such as site-directed evolution and random mutagenesis (Taguchi *et al*., 2008).

2.5 HISTORY AND AIM

An amorphous slime producing colony was observed during a previous study conducted in the IPB by Anneke Brand, during which cDNA expression libraries of *Gigartina polycarpa* (red algae) and *Pyura stolonifera* (redbait) were screened for production of galactan and other polysaccharides in *E. coli* on media enriched with sucrose and other disaccharides. The observation was made after the plates were stored at 4 °C for several days. Plasmid DNA was extracted from this colony and after several failed attempts to regain the phenotype in wild type *E. coli* it was assumed that the amorphous slime producing colony was a contaminant.

Intrigued by this biopolymer producing bacteria, it was decided to identify the unknown bacteria, as well as isolate and identify the genes responsible for the biopolymer production and the identification of the biopolymer. Thus the aim of this study was to identify the contaminant using 16S ribosomal RNA (rRNA) analysis, as well as isolating the genes responsible for biopolymer production by means of genomic expression library screening and to identify and characterise the biopolymer being produced.

3 MATERIALS

Chemicals used during this study were obtained from Sigma-Aldrich (Pty) Ltd., South Africa, unless specified otherwise.

All nucleic acid modifying enzymes were from Fermentas and were obtained from Inqaba Biotec™ (Inqaba Biotechnical Industries (Pty) Ltd., South Africa), unless specified otherwise.

All primers used during this study were synthesised by Inqaba Biotec™.

4 METHODS

4.1 ANALYSIS OF ORIGINAL BIOPOLYMER PRODUCING BACTERIA

4.1.1 Genomic DNA isolation

Genomic DNA (gDNA) was isolated from the biopolymer producing bacteria, from which the expression library was created, with a ZM Fungal/Bacterial DNA miniprep™ kit (Zymo Research Corporation, USA catalogue # D6005) as per manufacturer's specifications. Isolated gDNA was stored at -20 °C and kept on ice during further experimental procedures.

4.1.2 16S rRNA analysis

The 16S rRNA molecular work was kindly done by Mr. Charl Marais at the Institute for Plant Biotechnology at Stellenbosch University. The 16S rRNA of the original biopolymer producing bacteria was amplified from the gDNA using the polymerase chain reaction (PCR). A Kapa HiFi™ PCR kit (Kapa Biosystems (Pty) Ltd., South Africa) was used to amplify a ± 500 bp section of the 16S rRNA with degenerate primers E9F: 5'-GAGTTTGATCCTGGCTCAG-3' and U529R: 5'-ACCGCGGCKGCTGGC-3' (Baker *et al.*, 2003). The reaction was set up as per manufacturer's recommendation with 1 µL of the gDNA being used during the PCR. The recommended cycling protocol for Kapa HiFi™ was used, the primers were annealed at 52 °C and extension was allowed for 30 s. Amplification cycling was repeated for 30 cycles.

The PCR amplification was separated and visualised by gel electrophoresis on a 1% agarose gel stained with ethidium bromide (EtBr) under UV illumination. Banding correlating with the 500 bp λ *Pst*I marker region was excised and purified with a GeneJET™ Gel Extraction kit (Thermo Fisher Scientific Inc., USA catalogue # K0692). The purified 16S rRNA amplification product was ligated into pJet1.2/blunt cloning vector as per CloneJet™ PCR Cloning kit (Thermo Fisher Scientific Inc., USA catalogue # K1232) specifications for a blunt end ligation.
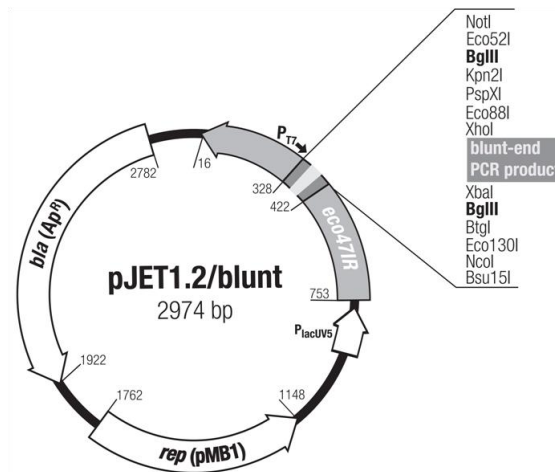
Figure 4.1.1: pJet1.2/blunt cloning vector map (Thermo Fisher Scientific Inc., USA).

Escherichia coli strain DH5α was transformed with the ligation reaction by method of heat shock (HS). Heat shock competent cells were prepared as per Tu *et al*. (2005). Aliquots of 100 µL HS competent cells were snap frozen with liquid nitrogen and stored at -80 °C. During HS transformation, 5 µL of the pJet1.2/blunt ligation was added to 50 µL of heat shock competent cells and placed on ice for 30 min. Cells were then incubated at 42 °C for 45 s and then placed on ice for a further 2 min before adding 1 mL Luria-Bertani (LB, 1% tryptone, 0.5% yeast extract, 1% NaCl (Bertani, 1951)). The transformation culture was incubated at 37 °C with shaking at 200 rpm for 1 hr. A 100 µL aliquot of transformed culture was spread out per plate of selection media consisting of LB agar (LB medium solidified with 1.5% (w/v) bacteriological agar) containing 50 µg/mL ampicillin. Plates were incubated at 37 °C for 12 h. Following incubation, 10 bacterial colonies were selected at random and were used as template for a colony PCR using the pJet1.2 forward primer (5'-CGACTCACTATAGGGAGAGCGGC-3') and pJet1.2 reverse primer (5'-AAGAACATCGATTTTCCATGGCAG-3') to confirm insertion of the 16S rRNA PCR product

into the pJet1.2/blunt vector. A positive colony was then used to inoculate 2 mL LB and incubated overnight at 37 °C with shaking at 200 rpm. Plasmid DNA was extracted from the overnight culture with a GeneJET™ Plasmid Miniprep kit (Thermo Fisher Scientific Inc., USA catalogue # K0503) as per manufacturer's specification. The plasmid DNA was quantified with a NanoDrop ND 1000 (Thermo Fisher Scientific Inc., USA) and sent for sequencing with the pJet forward primer to the Central Analytical Facility (CAF) at Stellenbosch University.

The data received from the sequencing facility was aligned with the National Centre for Biotechnology Information (NCBI) 16S microbial database and was accomplished by using the Basic Local Alignment Search Tool (BLAST, http://www.ncbi.nlm.nih.gov./BLAST/; Altschul *et al*., 1990). The 16S rRNA sequence data and the first 20 sequences showing homology obtained from the NCBI BLAST search was aligned with the ClustalW multiple alignment function in the BioEdit Sequence Alignment Editor program (Hail, 1999). The alignment was used to draw a phylogeny reconstruction tree with the neighbor-joining statistical method and tested with the bootstrap method which was replicated 1000 times in Molecular Evolutionary Genetics Analysis (MEGA) version 5 (Tamura *et al*., 2011). Sequences which were distant from the original bacteria's 16S rRNA sequence were removed from the alignment to optimise the build of the phylogenetic tree.

4.1.3 Antibiotic resistance analysis

The original biopolymer producing bacteria was streaked out on several LB agar plates containing different antibiotics. Kanamycin (35 µg/mL), ampicillin (50 µg/mL), chloramphenicol (34 µg/mL), tetracycline (10 µg/mL) and streptomycin (10 µg/mL) resistance were investigated. The plates were incubated at 37 °C for 48 h.

4.2 EXPRESSION LIBRARY

4.2.1 Expression library construction and vector

The expression library was created by Mr. Marnus Smith at the Institute for Plant Biotechnology at Stellenbosch University. The expression vector pBK-CMV (Agilent Technologies, Stratagene Products Division) was used to create the expression library. This vector has an isopropyl-β-D-thiogalactopyranoside (IPTG) prokaryotic inducible *lac*

promoter, a *LacZ* (β-galactosidase) reporter gene and a kanamycin resistance gene for selection. The *LacZ* reporter gene allows for blue white colony selection of putative vector insertions when grown on media containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal).

Genomic DNA was isolated from the original biopolymer producing bacteria and partially digested with *Sau*IIIA. The digested gDNA was separated and visualised by gel electrophoresis in a 1% agarose gel stained with EtBr under UV illumination. A segment of the digest containing 4–5 kb fragments correlating with the molecular weight marker was isolated from the agarose gel with a GeneJET™ Gel Extraction kit (Thermo Fisher Scientific Inc., USA catalogue # K0692) as per manufacturer's specification. The digested sticky end of *Sau*IIIA has compatibility with *Bam*HI digested sticky ends. The isolated and purified size fragment of the *Sau*IIIA gDNA digestion was ligated into the *Bam*HI site of the pBK-CMV vector.

4.2.2 Library screening - Bacterial transformation

Bacteria were transformed by method of heat shock (HS). *E. coli* strain DH5α was used during this study. Heat shock competent cells were prepared as per Tu *et al.* (2005). Aliquots of 100 µL HS competent cells were snap frozen with liquid nitrogen and stored at - 80 °C. During HS transformation 1 µL of the pBK-CMV expression library was added to 50 µL of heat shock competent cells and placed on ice for 30 min. Cells were incubated at 42 °C for 45 s and then placed on ice for a further 2 min before adding 1 mL LB. The transformation culture was incubated at 37 °C with shaking at 200 rpm for 1 hr. A 100 µL of transformed culture was spread out per plate of selection media consisting of LB agar containing 35 µg/mL kanamycin, 0.1 µM IPTG and 4% (w/v) sucrose. Plates were incubated at 37 °C for 48 h. Positive clones were identified as colonies producing an amorphous slime.

4.3 LIBRARY POSITIVE CLONE PLASMID DNA ANALYSIS

4.3.1 Plasmid isolation

The positive screen was streaked out on LB agar containing 35 µg/mL kanamycin, 10 µg/mL IPTG and 4% (w/v) sucrose and incubated at 37 °C overnight. A single colony was

picked and used to inoculate 4 mL LB containing 35 µg/mL kanamycin. The culture was incubated at 37 °C with shaking at 200 rpm overnight. Plasmid DNA (pDNA) was extracted from 3 mL of the overnight culture by the method of alkaline lysis. Glycerol stocks were prepared with the remaining culture by adding 850 µL of the bacterial culture to 200 µL 80% glycerol which was previously sterilised in Cryo Tubes™ (Thermo Fisher Scientific Inc., USA catalogue # 377224) by autoclaving at 121 °C with a pressure of 103.5 kPa for 20 min. All glycerol stocks were stored at -80 °C.

Cells in the 3 mL overnight culture were pelleted by centrifugation in a desk top centrifuge (Heraus, Biofuge pico, catalogue # 3325) at 13000 rpm (±16000 x g) for 2 min. The supernatant was discarded and the pellet was re-suspended in 200 µL of a solution containing 50 mM TRIS-HCl pH 8, 10 mM EDTA, 0.1 g/L RNAse A by vortexing. After re-suspension, the culture was lysed by adding 200 µL of a solution containing 200 mM NaOH, 1% (w/v) SDS and mixing gently by inverting the tube four times. After mixing, 200 µL of ice cold solution containing 3 M potassium acetate, pH 5.5 was added and the solution was mixed gently by inverting the tube four times. The solution was then placed on ice for 5 min. The sample was centrifuged at 13000 rpm for 10 min. The upper phase was aspirated and placed in a new 2 mL tube. A 1:1 ratio of a 1:1, phenol: chloroform solution was added to the sample. This solution was mixed by vortexing for 30 s. The sample was then centrifuged for 5 min at 13000 rpm. The upper phase was aspirated into a new 2 mL tube. Three times the volume of the upper phase of ice cold 100% ethanol (EtOH) was added to the sample. It was placed on ice for 10 min before being centrifuged at 13000 rpm for 10 min at 4 °C. The liquid was aspirated and discarded while taking care not to disturb the DNA pellet. The pellet was washed twice with 1 mL 70% EtOH. The sample was allowed to dry on the bench for 30 min. The pellet was then re-suspended in 50 µL MilliQ (MQ) H_2O. All DNA samples were stored at -20 °C and kept on ice during further experimental procedures.

4.3.2 Plasmid re-transformation

Plasmid DNA from the isolated positive clone was used to transform *E. coli* strain DH5α cells. This was done by method of HS, were 1 µL pDNA was used to transform heat shock competent cells as described on page 28. A 100 µL aliquot of the transformation culture was plated out per plate of LB agar containing 35 µg/mL kanamycin and incubated at 37 °C for 48 h.

4.3.3 Restriction enzyme analysis of the positive clone

The insert of the positive clone pDNA was analysed by restriction enzyme digest. The pDNA was separately digested with *Xho*I, *Pst*I, *Hin*dIII as well as double digestions with *Xho*I/*Pst*I and *Hin*dIII/*Pst*I. None of these restriction enzymes has star activity. Optimal digestion conditions and buffers were used as per manufacturer's protocol. Double digest conditions were used as per DoubleDigest™ (Thermo Fisher Scientific Inc., USA) specifications. For each digestion reaction, 200 ng pDNA was used in a 50 µl reaction. Reactions were incubated at 37 °C overnight.

Restriction patterns were separated and visualised with gel electrophoresis in a 1% agarose gel stained with EtBr under UV illumination.

4.3.4 Positive clone insert fragment recombination in pBluescript SK⁻

Another 1 µg of the positive clone pDNA was digested with a combination of *Xho*I and *Pst*I as per DoubleDigest™ (Thermo Fisher Scientific Inc., USA) specifications in a 50 µL reaction. The reaction was incubated at 37 °C overnight and was separated and visualised by gel electrophoresis in a 1% agarose gel stained with EtBr. Both of the insert fragments visualised were isolated from the agarose gel with a GeneJET™ Gel Extraction kit (Thermo Fisher Scientific Inc., USA catalogue # K0692). The two fragments were recombined respectively into two separate pBluescript SK⁻ vectors.

The pBluescript SK⁻ vector was digested with *Xho*I and with a combination of *Pst*I and *Xho*I respectively as per manufacturer's and DoubleDigest™ (Thermo Fisher Scientific Inc., USA) specifications. The reaction volume was 50 µL, 0.5 µg pBluescript SK⁻ was digested and the reaction was incubated at 37 °C overnight. The two digests were cleaned with a GeneJET™ PCR Purification kit (Thermo Fisher Scientific Inc. USA, catalogue # K0702) as per manufacturer's specification. The pBluescript SK⁻ vector digested with *Xho*I was dephosphorylated with FastAP™ Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific Inc., USA, catalogue # EF0651) as per manufacturer's specification and was heat inactivated by incubation at 75 °C for 5 min.

The two insert fragments were ligated into their respective pBluescript SK⁻ vectors with T4 DNA Ligase (Thermo Fisher Scientific Inc., USA, catalogue # EL0011) as per manufacturer's specifications. In the reaction, 5 µL vector DNA and 10 µL insert DNA was used. The ligation reactions were transformed into heat shock competent *E. coli* strain

DH5α cells by method of heat shock transformation as described on page 28. During transformation, 5 µL of the ligation reaction was used. The transformed cells were plated out on LB agar containing 50 µg/mL ampicillin, 0.1 µM IPTG and 20 µg/mL X-gal. Plates were incubated overnight at 37 °C.

The pBluescript SK⁻ vector has an isopropyl-β-D-thiogalactopyranoside (IPTG) prokaryotic inducible *lac* promoter, a *LacZ* (β-galactosidase) reporter gene and an ampicillin resistance gene for selection. When transforments are grown on media containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), blue/white selection of putative vector insertions are possible.
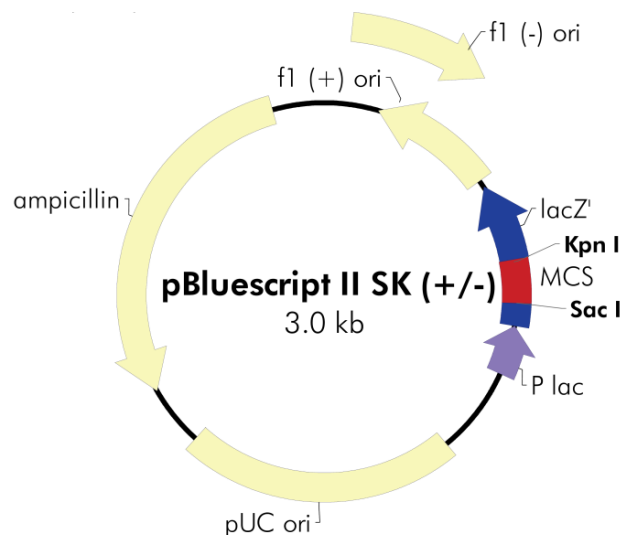
Figure 4.3.1: pBluescript II SK(+/-) vector map (pBluescript II Phagemid Vectors, Instruction Manual, Stratagene).

4.3.5 pBluescript SK⁻ insert analysis

Five white colonies from each pBluescript SK⁻ transformation were streaked out on LB agar plates containing 50 µg/mL ampicillin, 0.1 µM IPTG and 4% (w/v) sucrose. These plates were incubated at 37 °C overnight. The same white colonies were also used to inoculate 4 mL LB which was incubated at 37 °C overnight with shaking at 200 rpm.

Plasmid DNA was isolated from 3 mL of the overnight culture by method of alkaline lysis as described on page 29. The inserts of the pDNA were investigated by restriction enzyme digest with both *Xho*I and *Pst*I, as well as *Xho*I respectively. Reactions were set up as per manufacturers and DoubleDigest™ (Thermo Fisher Scientific Inc., USA) specifications in a 50 µL reaction with 5 µL of pDNA. The reactions were incubated at 37 °C overnight and restriction patterns were separated and visualised with gel electrophoresis in a 1% agarose gel stained with EtBr under UV illumination.

4.3.6 Sequencing

The pBluescript SK⁻ vector containing the *Pst*I, *Xho*I insert fragment of the positive pBK-CMV clone, which produced biopolymer, was cleaned with a GeneJET™ PCR purification kit (Thermo Fisher Scientific Inc., USA, catalogue # K0702) as per manufacturer's specification, before being sent for sequencing. Samples were quantified with a NanoDrop ND 1000 (Thermo Fisher Scientific Inc., USA). Sequencing was carried out at the Central Analytical Facility (CAF) at Stellenbosch University. The first round of sequencing was done with primers T7: 5'-TAATACGACTCACTATAGGG-3' and T3: 5'-ATTAACCCTCACTAAAGGGA-3'. Thereafter, primer walking from the sequence derived with the T7 primer was done with primers TRLVF: 5'-AGAAGGTCAAGCCGCTGTT-3'; TRLVF2: 5'-AGGCAGCGGGTTTGAATAC-3'; TRLVF3: 5'-ATAGGCACTCACCCCACTGT-3' and from the sequence derived from the T3 primer, primer TRLVRR: 5'-ATGCATTACGTGCCCTCAG-3' was used.

The data received from the sequencing facility was manually aligned with the BioEdit Sequence Alignment Editor program (Hail, 1999). The BioEdit Sequence Alignment Editor program was used to search for open reading frame's (ORFs) with the "Find Next ORF" function. Sequence alignment with the National Centre for Biotechnology Information (NCBI) nucleotide collection database was accomplished by using the Basic Local Alignment Search Tool (BLAST, http://www.ncbi.nlm.nih.gov./BLAST/; Altschul *et al.*, 1990).

The nucleotide sequence of the ORF and the three sequences showing high homology obtained from the NCBI BLAST search was aligned with the ClustalW multiple alignment function in the BioEdit Sequence Alignment Editor program (Hail, 1999). The alignment was used to draw a phylogeny reconstruction tree with the neighbor-joining statistical method and tested with the bootstrap method which was replicated 1000 times in Molecular Evolutionary Genetics Analysis (MEGA) version 5 (Tamura *et al.*, 2011).

The molecular weight of the amino acid composition of the ORF was calculated in BioEdit Sequence Alignment Editor program using the "Amino Acid Composition" function. The amino acid sequence was aligned with NCBIs non-redundant protein sequences database with a protein-protein BLAST (blastp). The first 20 sequences showing homology obtained from the NCBI BLAST search were aligned with the ClustalW multiple alignment function in the BioEdit Sequence Alignment Editor program. The alignment was used to draw a phylogeny reconstruction tree with the neighbor-joining statistical method and tested with the bootstrap method which was replicated 1000 times in Molecular Evolutionary Genetics Analysis (MEGA) version 5. Sequences which were distant from the ORF amino acid sequence were removed from the alignment to optimise the build of the phylogenetic tree.

4.4 BIOPOLYMER ISOLATION, PURIFICATION AND ANALYSIS

4.4.1 Biopolymer substrate analysis

The original polymer producing bacteria, as well as the pBK-CMV positive clone and the pBluescript SK⁻ *Pst*I/*Xho*I insert clone, were streaked out on several LB agar plates containing 4% (w/v) of different simple sugars (glucose, fructose, mannose, maltose, galactose, sucrose and lactose) as well as 50 µg/mL ampicillin or 35 µg/mL kanamycin and 0.1 µM IPTG. A control plate containing no sugar was also streaked out. The plates were incubated at 30 °C for 48 h.

4.4.2 Biopolymer isolation

Bacteria were plated onto four plates of LB agar containing 10 µg/mL IPTG, 4% (w/v) sucrose and 50 µg/mL of the respective antibiotic for selection and incubated at 30 °C for 48 h. Cultures were scraped off the plates and pooled into 50 mL tubes and re-suspended in 15 mL MQ H$_2$O by vortexing. The re-suspended culture was centrifuged at 4000 x g for 5 min in a Sorvall® RC5C Plus centrifuge and the supernatant was transferred to a new 50 mL tube. The supernatant was centrifuged again at 4000 x g for 5 min and was placed in a new 50 mL tube. The 50 mL tube containing the supernatant was filled with ice cold 100% EtOH. The samples were incubated at -20 °C for 1 h, after which they were centrifuged at 8000 x g for 15 min to pellet the biopolymer. The supernatant was discarded and the tubes were incubated at 55 °C for 1 h to dry the biopolymer. The dried biopolymer was stored at room temperature for future use.

4.4.3 Biopolymer purification

The dried biopolymer was re-suspended in 15 mL MQ H_2O. SnakeSkin dialysis tubing with a 3.5 kDa molecular weight cut-off (MWCO) (Thermo Fisher Scientific Inc., USA catalogue # 68035) was used to dialyse 5 mL of the dissolved biopolymer solution. Dialysis was performed in 2 L dH_2O containing 0.02% sodium azide (NaN_3) for 16 h at 4 °C with stirring. Dialysed biopolymer was removed from the dialysis tubing and was placed into 2 mL tubes. The biopolymer solution was then allowed to dry in a Genevac® personal evaporator EZ-2 (Genevac Ltd.) and stored at room temperature until further experimental procedures.

4.4.4 Biopolymer size exclusion chromatography

Size exclusion chromatography was done on a Sephacryl™ S-400 HR (GE Healthcare HIPrep™ 26/60 catalogue # 28-9356-05) size exclusion column with an AKTA prime (Amersham pharmaciabiotech) pump. The elution volume was set at 5 mL. The column was equilibrated with one column volume (320 mL) of a 0.05 M sodium phosphate pH 7.0, 0.25 M NaCl solution. Three carbohydrate size standards (1.4 mDa, 270 kDa and 5 kDa) were prepared by dissolving 8 mg of each standard in 2 mL 0.05 M sodium phosphate pH 7.0, 0.25 M NaCl. The standards solution was filtered through a 0.20 μm single use filter (GVS Filter Technology) before being injected into the column. After elution of the standards the column was again washed with one column volume of 0.05 M sodium phosphate pH 7.0, 0.25 M NaCl.

Five mg of the dried, purified biopolymer was dissolved in 2 ml 0.05 M sodium phosphate pH 7.0, 0.25 M NaCl. The biopolymer solution was filtered through a 0.20 μm single use filter (GVS Filter Technology) before being injected into the column. After elution of the biopolymer the column was washed with one column volume of 0.05 M sodium phosphate pH 7.0, 0.25 M NaCl.

Both the standards and the biopolymer elution fractions were analysed by transferring 25 μL of each elution volume in elution order to a 96 well ELISA plate (Greiner Bio-One). This was performed in triplicate on the same plate. Twenty-five μL 5% phenol was added to each sample and agitated for 1 min and 125 μL 96% (v/v) H_2SO_4 was added to each sample (modified from Fournier, 2001). The change in absorbance of the samples was measured in a μQuant (Bio-Tek Instruments Inc., Analytical and Diagnostic Products 2000) plate reader at 492 nm and annotated with KC4™ version 2.7, rev 8 (Kinetical for Windows)

software. Averages of the triplicate samples were determined and used to plot a size to elution volume graph in Excel (Microsoft® Office, 2007).

4.4.5 Biopolymer hydrolysis

The biopolymers were hydrolysed using sulphuric acid (H_2SO_4). Five mg dry purified polymer was weighed and placed in a 2 mL screw cap tube. A 100 µL of 72% (v/v) H_2SO_4 was added and the sample was incubated at 30 °C for 45 min. The sample was then diluted to 4% (v/v) H_2SO_4 by adding 1.8 mL MQ H_2O. After dilution, the sample was autoclaved at 121 °C with a pressure of 103.5 kPa for 20 min. The sample was allowed to cool before adding 40 µL of a 0.05 g/mL solution bromophenol blue (3',3",5',5"-tetrabromophenolsulfonphthalein) and transferring the sample to a 15 mL tube in which 6 mL 0.18 M barium hydroxide [$Ba(OH)_2$] was added. The samples were yellow in colour and barium carbonate ($BaCO_3$) was added bit by bit during continued vortexing until the solution became green. The sample was then centrifuged at 4500 x g and the supernatant aspirated into a new 15 mL tube without disturbing the pellet (insoluble $BaSO_4$). The centrifugation and aspiration steps were repeated until no precipitate pellet was visible. The sample was dried in a Genevac® personal evaporator EZ-2.

The dry sample was dissolved in 1 mL 100% GC-MS grade methanol (MeOH) and filtered through a 0.20 µm single use filter (GVS Filter Technology). The filted sample was dried in the Genevac® and then re-dissolved in 1 mL 100% GC-MS grade MeOH and dried again. This was repeated three times. The dried sample was stored at room temperature until further experimental procedures.

4.4.6 Gas chromatography – mass spectrometry (GC-MS) analysis

The dried hydrolysed biopolymer as well as glucose, fructose and mannose standards were derivatised for GC-MS analysis. Sugar standards were prepared by dissolving 10 mg of the respective sugars in 1 mL 100% GC-MS grade MeOH. Only 10 µL (100 µg) of the dissolved sugars were placed in new 2 mL tubes. The 5 mg dried hydrolysed biopolymer samples were dissolved in 1 mL GC-MS grade MeOH and 200 µL (1 mg), was placed into new 2 mL tubes. The new 2 mL tubes containing the sugar standards and biopolymer samples each had 2 µL of a 2 mg/mL myo-inositol solution added to them for an internal standard. These samples were then dried in the Genevac®.

Derivatisation was accomplished by adding 100 µL of a 1:3:9 ratio solution of chlorotrimethylsilane: hexamethyldisilazane: pyridine to each sample and incubating the samples at 37 °C for 45 min. After incubation, an additional 100 µL of pyridine was added to each sample. The samples were centrifuged at 13000 rpm for 3 min. Without disturbing the pellet, the supernatant of each sample was aspirated and placed into pre-marked GC-MS glass vials which were sealed and sent for analysis at the Central Analytical Facility (CAF) at Stellenbosch University. GS-MS data was analysed with the Automated Mass Spectral Deconvolution and Identification System (ADMIS) (Stein, 1999).

4.4.7 Nuclear magnetic resonance (NMR) analysis

Biopolymer samples were prepared for NMR analysis by dissolving 10 mg purified polymer in 1 mL deuterium monoxide (D_2O) and adding tetramethylsilane (TMS) to a final concentration of 4.5 mM. A final volume of 750 µL of this solution was placed in disposable NMR tubes and sent for analysis at the Central Analytical Facility (CAF) at Stellenbosch University.

A 600 MHz Varian INOVA NMR was used to create the proton (^{1}H) and carbon (^{13}C) as well as the 2D Heteronuclear Single Quantum Coherence (HSQC) spectra of the biopolymers. The NMR spectra were analysed with the statistical software program R version 2.13.2 (Schwartz *et al*., 2008) using the package rNMR version 1.1.7 (Lewis *et al*., 2009), as well as the Advanced Chemistry Development (ACD)- ACD/NMR Processor Academic Edition version 12.01 (Advanced Chemistry Development, Inc., Canada, USA). The 2D HSQC spectra of the biopolymer samples were converted to ucsf format and imported into rNMR package with the statistical software program R. The spectra were overlaid and regions of interest (ROIs) were automatically generated with default parameters and a strong noise filter applied. All the ROIs in all the files were used to produce the ROI table in a multiplot. The ^{13}C and ^{1}H NMR data were opened separately in the 1D NMR Processor of ACD/NMR Processor. All three the samples per data set were opened on the same window with the add mode. The offset of the phasefiles was set manually with the manual offset function.

5 RESULTS

5.1 ORIGINAL BIOPOLYMER PRODUCING BACTERIA

The ± 500 bp fragment of the 16S rRNA was successfully amplified from the gDNA of the original biopolymer producing bacteria and inserted into the pJet1.2/blunt vector. Sequencing data produced a 516 bp sequence (see appendix for sequence, page 61). Alignment of the sequence with NCBI's 16S microbial database showed that the sequence had high homology to a wide range of *Pseudomonas* species. The sequences of several *Pseudomonas* species from the NCBI database search were aligned with the 16S rRNA sequence of the original biopolymer producing bacteria (Figure 5.1.1, page 38). A phylogenetic neighbor-joining tree was drawn with the aligned 16S rRNA sequences which determined the evolutionary distances between the sequences (Figure 5.1.2, page 38).

The original biopolymer producing bacteria had antibiotic resistance to ampicillin, kanamycin, chloramphenicol and streptomycin but not to tetracycline (Figure 5.1.3, page 39). This multiple antibiotic resistances are characteristic of many *Pseudomonas* species. Biopolymer production was only visualised on media containing sucrose as carbon source. The bacteria did not produce a biopolymer on any other carbon source.
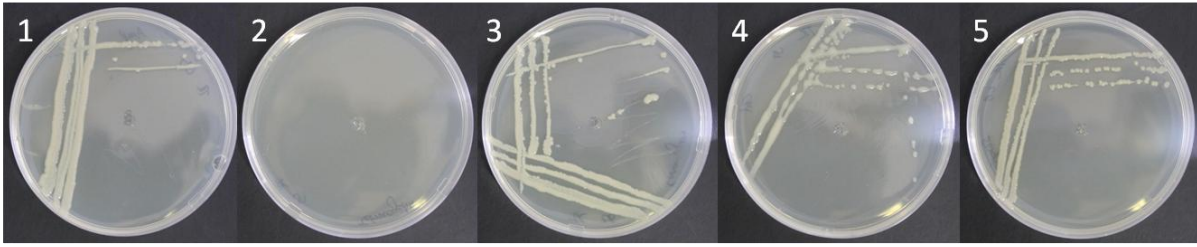
Figure 5.1.1: BioEdit sequence alignment of *T. original* bacteria 16S rRNA sequence and the NCBI BLAST hits.

Figure 5.1.2: Phylogenetic neighbor-joining tree of *T. original* bacteria 16S rRNA

38

Figure 5.1.3: The polymer producing bacteria plated out on LB media containing different antibiotics: 1: Ampicillin (50 µg/mL), 2: Tetracycline (10 µg/mL), 3: Kanamycin (35 µg/mL), 4: Chloramphenicol (34 µg/mL), 5: Streptomycin (10 µg/mL).

5.2 POSITIVE CLONE

A colony producing an amorphous slime was visualised during library screening of the pBK-CMV library after 2 days incubation at 37 °C (Figure 5.2.1). The positive clone was named TPS1. Plasmid DNA was isolated from the TPS1 clone and the phenotype was retained after re-transformation into wild type *E. coli* strain DH5α.

Figure 5.2.1: Amorphous slime phenotype of the positive clone TPS1.

The pDNA was digested with both *Xho*I and *Pst*I, which have restriction sites located upstream and downstream respectively of the *Bam*HI restriction site into which the library fragments were inserted in the multiple cloning site of pBK-CMV (Figure 5.2.2, page 40). It was discovered that the TPS1 insert has an internal *Xho*I restriction site producing 2 insert fragments, a *Xho*I/*Xho*I (XX) fragment of ± 2.4 kb and a *Pst*I /*Xho*I (PX) fragment of ± 2.7 kb (Figure 5.2.3, page 40).
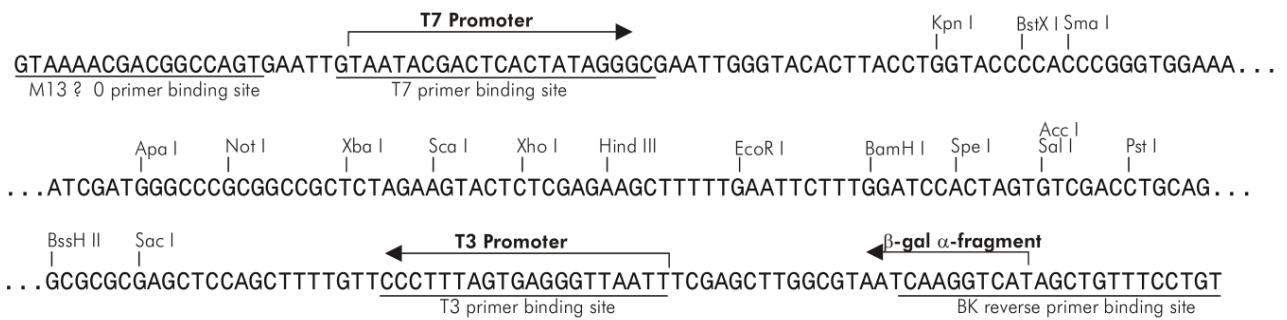
Figure 5.2.2: pBK-CMV multiple cloning site (pBK-CMV Phagemid vector, Instruction Manual, Stratagene).
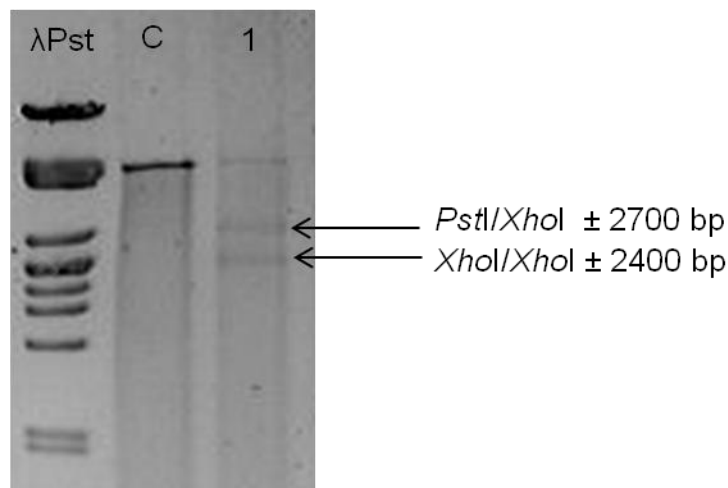

Figure 5.2.3: Restriction digests of C: pBK-CMV empty vector and 1: positive clone, TPS1, with both *Pst*I and *Xho*I.

Both the XX and PX insert fragments of the TPS1 digest were isolated and purified from the agarose gel and inserted separately into pBluescript SK⁻ vectors. Of the putatively transformed white colonies, 5 colonies of each insert fragment containing vectors were streaked out on selection media containing sucrose. After 2 days incubation at 37 °C, one of the colonies containing the PX insert was producing an amorphous slime which was similar to that of the TPS1 positive clone. This clone was named PX5. Plasmid DNA was isolated from the PX5 clone and the phenotype was retained after re-transformation into

40

wild type *E. coli* strain DH5α. Restriction digest analysis of the PX5 clone with both *Pst*I and *Xho*I restriction enzymes confirmed the ± 2.7 kb insert into the pBluescript SK⁻ vector (Figure 5.2.4).
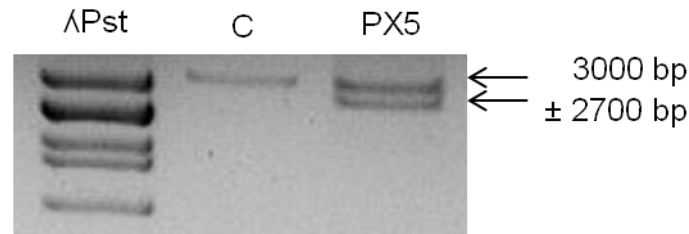

Figure 5.2.4: Restriction digests of C: empty pBluescript SK⁻ vector and PX5 with both *Pst*I and *Xho*I.

The PX5 clone was chosen for sequencing as it still produced the biopolymer and had a smaller insert size. Sequencing of the entire insert was accomplished by primer walking from both ends of the PX5 clone insert. The sequence data from the PX5 insert was manually aligned and a 77 bp overlap was obtained with the sequences from both sides of the insert, producing a 2639 bp sequence.

Sequence alignment of the PX5 insert sequence with the NCBI nucleotide collection database showed that the insert had highest homology in the first 1537 bp to *Leuconostoc mesenteroides* (AY665464.1), 99% and a 485 bp fragment downstream with *Pseudomonas fluorescens* strain SS101 (EU199081.2), 83% (Figure 5.2.5, page 42). The first 1433 bp of the insert also had 89% homology *Pseudomonas aurantiaca* (AF30613.1), 87% homology to *Pseudomonas brassicacearum* subsp. *brassicacearum* (CP002585) and 87% homology to *Pseudomonas fluorescens* F113 (CP003150.1). Sequences fragments in the downstream 485 bp sequence also had 93% homology to *Pseudomonas fluorescens* SBW25 (AM181176.4), 87% homology to *Pseudomonas brassicacearum* subsp. *brassicacearum* NFM421 (CP002585.1), 87% homology to *Pseudomonas fluorescens* F113 (CP003150.1) and 80% *Pseudomonas fluorescens* SS101 (EU199080.2).
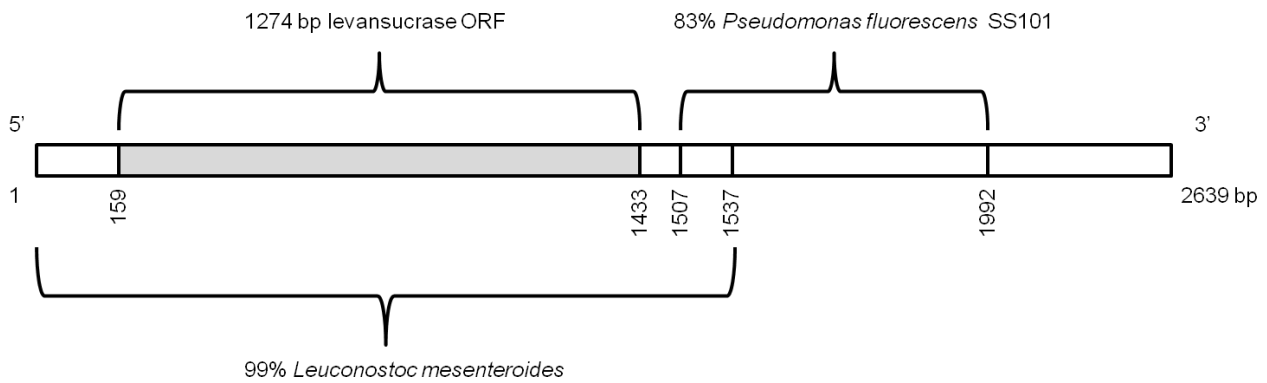
Figure 5.2.5: Schematic representation of sequence homology and ORF of the PX5 insert.

The insert contained an open reading frame (ORF) of 1274 bp consisting of 424 amino acids encoding a levansucrase. The calculated molecular weight of the levansucrase was 47.1 kDa. The nucleotide sequence of the levansucrase had 100% homology to the *Leuconostoc mesenteroides* levansucrase gene (*m1ft*) (AY665464.1) and 90% identity to *Pseudomonas chlororaphis* subsp. *aurantiaca* levansucrase gene (*lscA*) (AF306513.1). The nucleotide sequences which aligned to the NCBI database were re-aligned to the PX5 levansucrase nucleotide sequence in BioEdit (Figure 5.2.6, page 43) and were used to draw a phylogenetic neighbor-joining tree (Figure 5.2.7, page 43). The amino acid sequence of the PX5 levansucrase had 100% homology to *Leuconostoc mesenteroides* levansucrase (AAT81165.1), 96% homology to *Pseudomonas chlororaphis* subsp. *aurantiaca* levansucrase (AF306513.1) and 93% to *Pseudomonas fluorescens* F113 M1ft levansucrase (AEV65175.1). The amino acid sequences which aligned to the NCBI database were re-aligned to the PX5 levansucrase amino acid sequence in BioEdit (Figure 5.2.8, page 44) and was used to draw a phylogenetic neighbor-joining tree (Figure 5.2.9, page 44). The PX5 levansucrase has a specific hit to glycosyl hydrolase (GH) family 68 and falls under the GH43_62_32_68 superfamily.
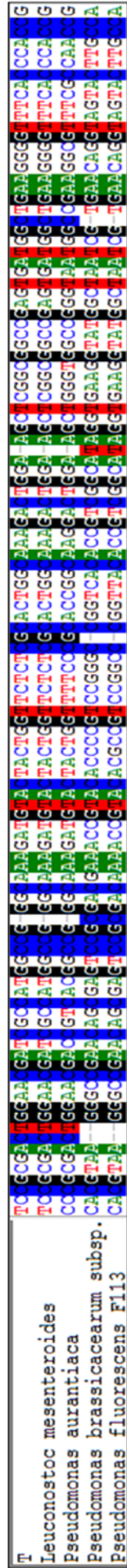
Figure 5.2.6: BioEdit sequence alignment of T. levansucrase nucleotide sequence in the PX5 clone and the NCBI BLAST hits.
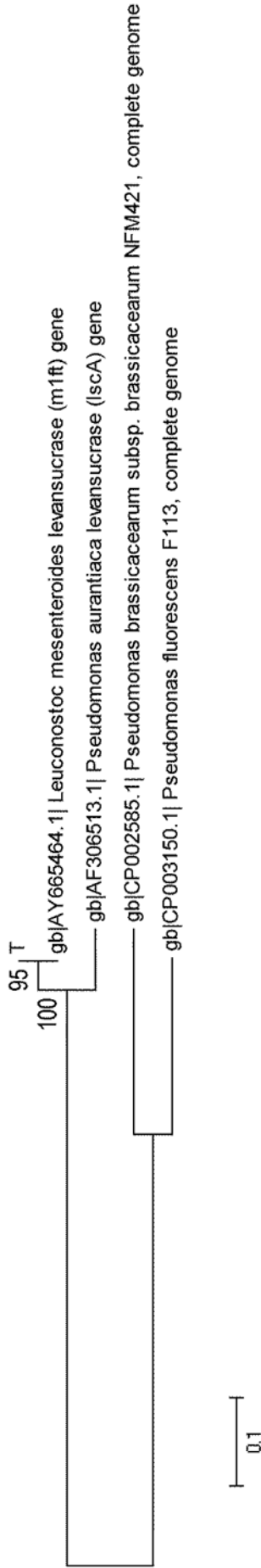

Figure 5.2.7: Phylogenetic neighbor-joining tree of T. levansucrase nucleotide sequence in the PX5 clone.

43

Figure 5.2.8: BioEdit sequence alignment of T. levansucrase amino acid sequence of the PX5 clone and the NCBI BLAST hits.
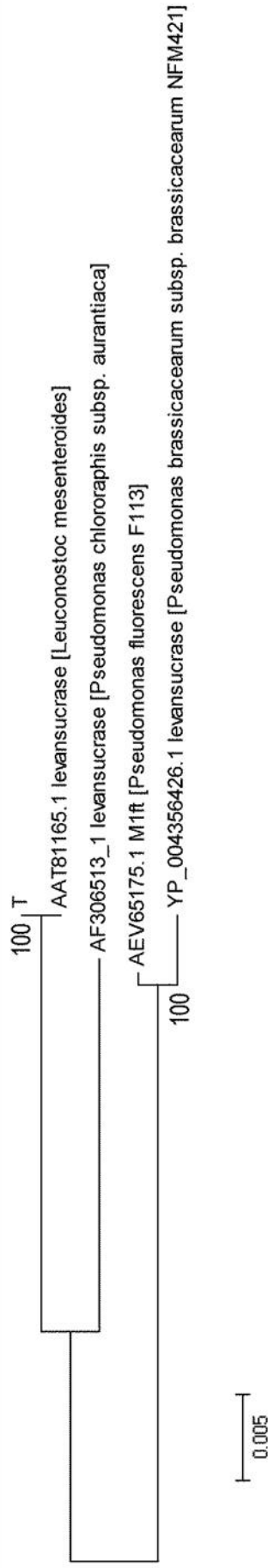

Figure 5.2.9: Phylogenetic neighbor-joining tree of T. levansucrase amino acid sequence of the PX5 clone.

44

5.3 BIOPOLYMER CHARACTERISATION

The original biopolymer producing bacteria had an amorphous slime phenotype which is characteristic of exopolysaccharide (EPS) producing bacteria (Sutherland, 1998). Size exclusion chromatography determined that the EPS from the original biopolymer producing bacteria had a molecular weight of approximately 5 kDa (Figure 5.3.1). The molecular weight correlates to the molecular weight of ± 27 sugar molecules with a $C_6H_{12}O_6$ chemical composition. Production of EPS by the TPS1 positive clone and the PX5 clone were only visualised on media containing sucrose as carbon source.
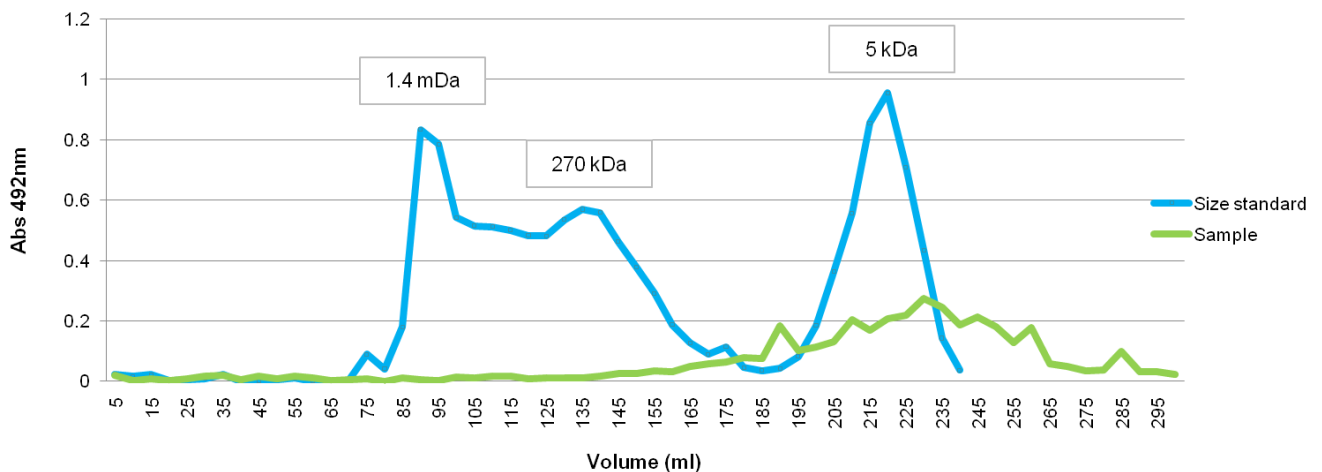
Figure 5.3.1: Size exclusion chromatography graph of size standards and original bacteria's biopolymer on a Sephacryl™ S-400 HR column.

The EPS of the original biopolymer producing bacteria (T), positive clone in pBK-CMV (TPS1) and *Pst*I/*Xho*I fragment of TPS1 in pBluescript SK⁻ (PX5) were isolated and purified for analysis by GC-MS and NMR. The three EPSs were completely hydrolysed and derivatised for GC-MS analysis. GC-MS total ion chromatogram (TIC) data analysis of the three EPS determined that they were the same and correlated to the GC-MS TIC of the fructose standard (Figure 5.3.2, page 46).
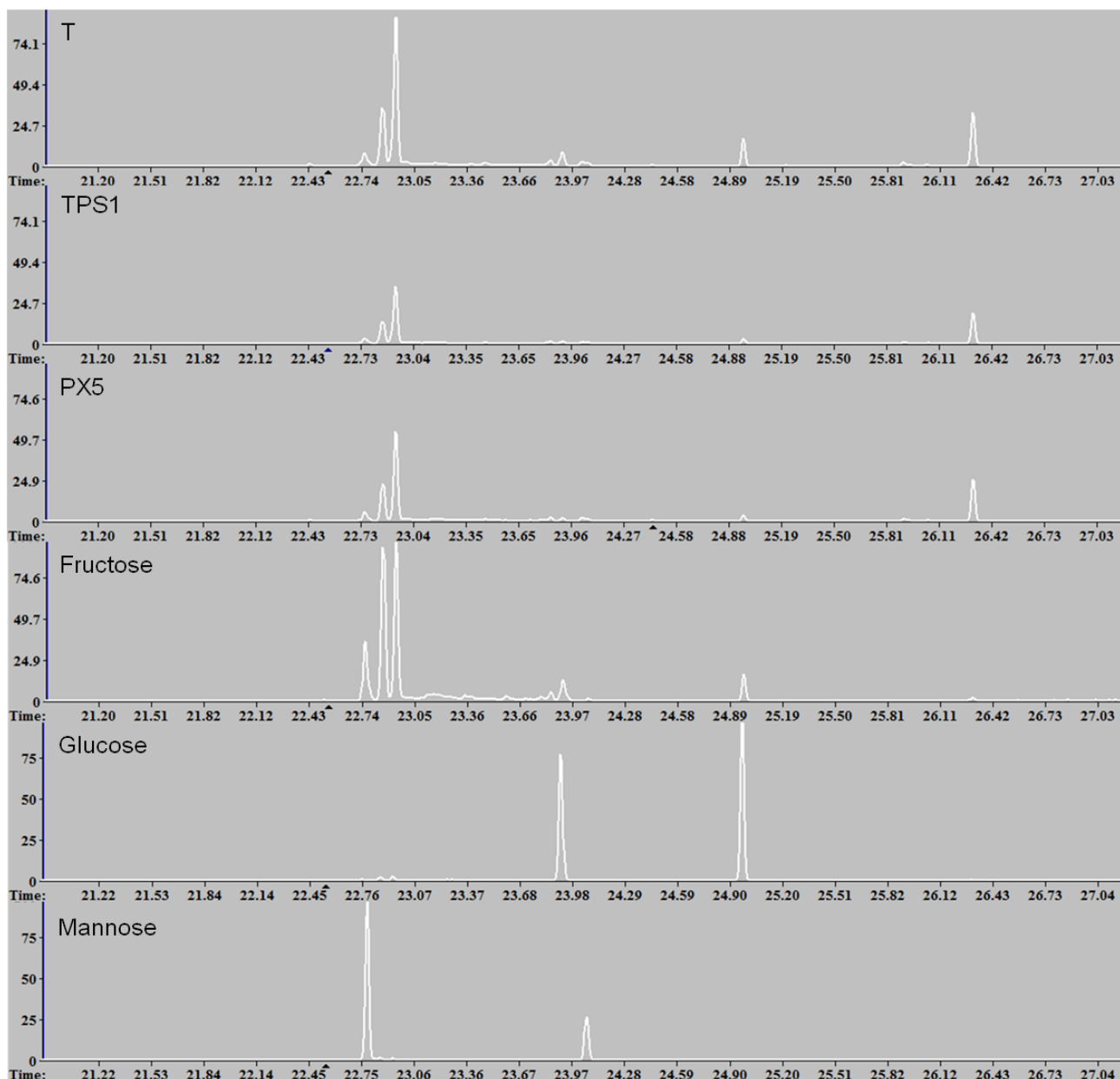
Figure 5.3.2: GC-MS TIC chromatogram analysis of the derivatised hydrolysed polymers, T: original bacteria EPS, TPS1: positive clone in pBK-CMV EPS, PX5: *Pst*I/*Xho*I positive clone fragment in pBluescript SK⁻ EPS as well as fructose, glucose and mannose standards. The peak at retention time 26.30 is of the internal myo-inositol standard.

The ^{13}C NMR analysis of the three EPSs: T, TPS1 and PX5 were done at 150 MHz. The ^{13}C spectra of the three EPS were the same (Figure 5.3.3, page 48). The ^{13}C spectra were correlated to published ^{13}C data for levan and it was determined that they were similar (Table 5.3.1, page 47). The ^{1}H NMR analysis of the three EPSs determined that they were

the same on the proton level (Figure 5.3.4, page 49). The regions of interest (ROIs) in the 2D Heteronuclear Single Quantum Coherence (HSQC) spectra data from the three EPSs were the same (Figure 5.3.5, page 49).

Table 5.3.1: ^{13}C-NMR spectra peaks of EPS: EPS of this study, levan and M1FT EPS: levan produced by *Leuconostoc mesenteroides* levansucrase.

Carbon atom	Chemical Shift (ppm)		
	EPS	Levan[a]	M1FT EPS[b]
C1	**62.77**	60.70	60.50
C2	**107.05**	104.20	109.10
C3	**79.16**	76.30	76.50
C4	**78.06**	75.20	75.40
C5	**83.14**	80.30	80.70
C6	**62.77**	63.40	63.60

[a] Kang *et al.*, 2005; [b] Shimamura *et al.*, 1987
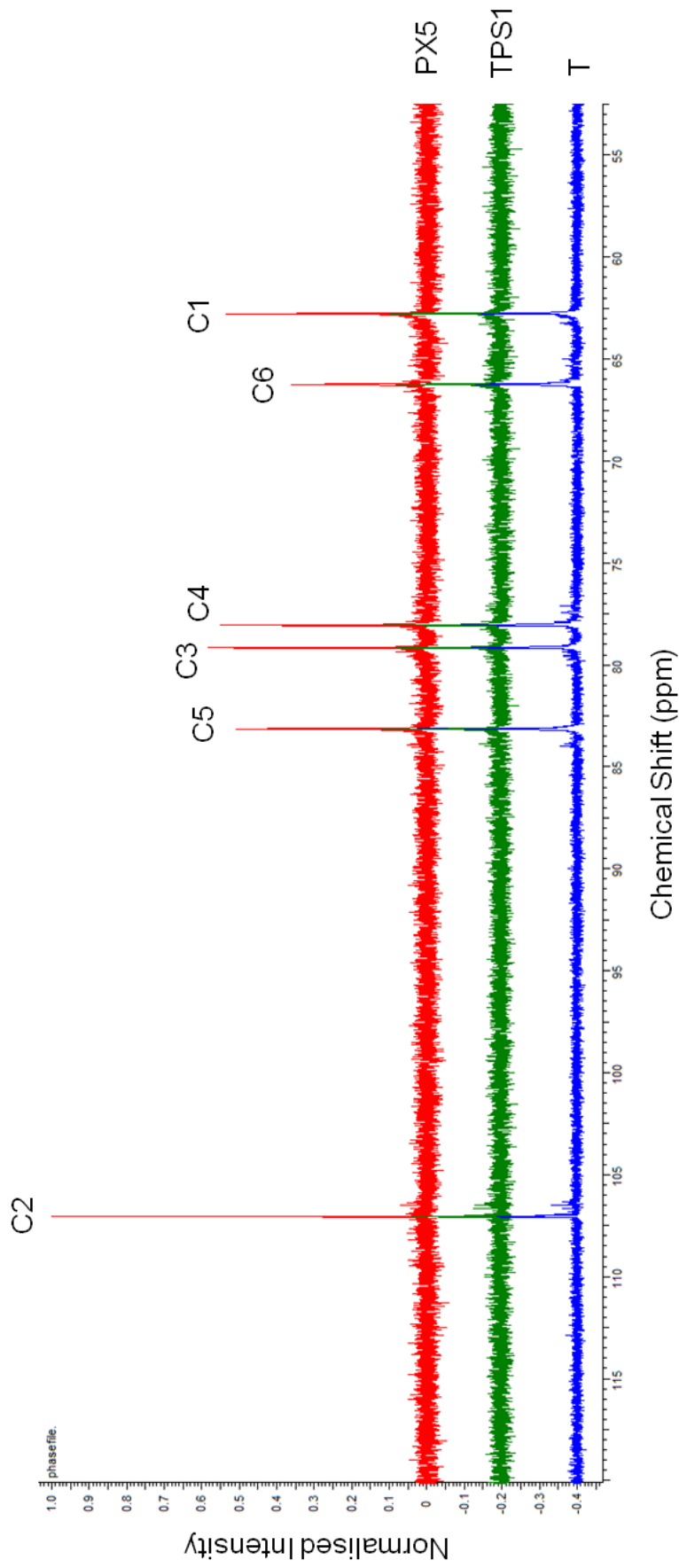
Figure 5.3.3: ^{13}C NMR spectra of T: original bacteria EPS, TPS1: positive clone in pBK-CMV EPS and PX5: *PstI/XhoI* fragment of positive clone in pBluescript SK⁻ EPS.

Figure 5.3.4: ^{1}H NMR spectra of T: original biopolymer producing bacteria EPS, TPS1: positive clone in pBK-CMV and PX5: *Pst*I/*Xho*I fragment of positive clone in pBluescript SK[-] EPS.

Figure 5.3.5: Regions of interest (ROI) table of the 2D Heteronuclear Single Quantum Coherence (HSQC) NMR of T: original biopolym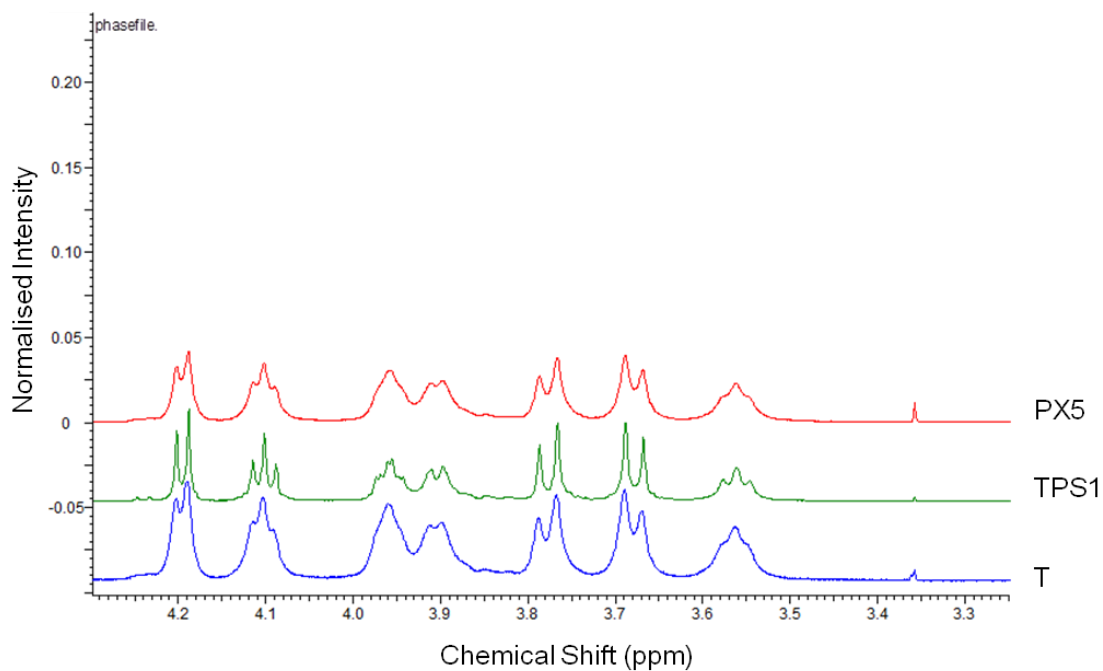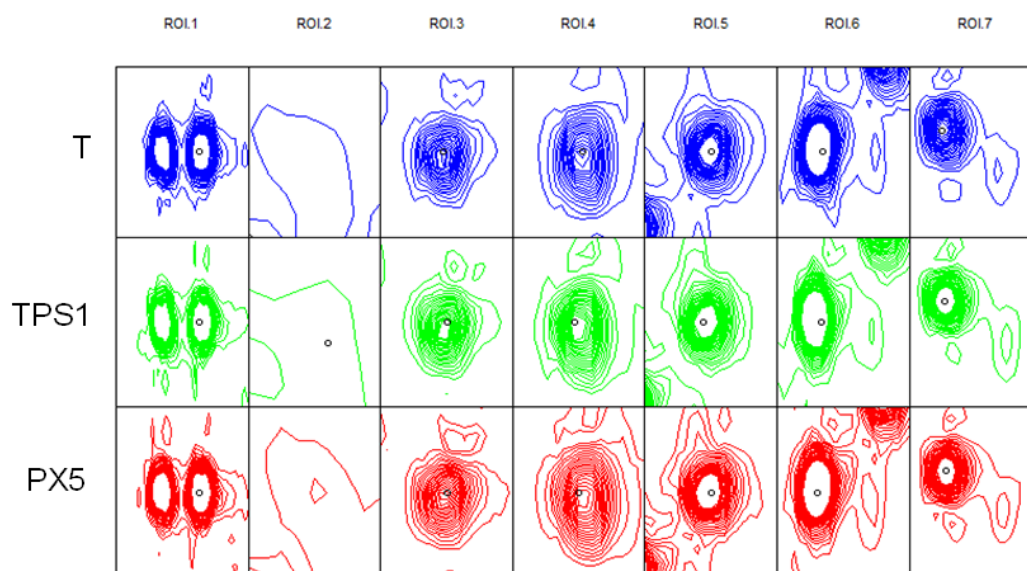er producing bacteria EPS, TPS1: positive clone in pBK-CMV and PX5: *Pst*I/*Xho*I fragment of positive clone in pBluescript SK[-] EPS.

6 DISCUSSION

The 16S rRNA data of the original biopolymer producing bacteria revealed that the bacteria were from the *Pseudomonas* species. The exact identity could only be determined to the genus level and this could possibly be due to the sequencing of only a short fragment of a highly variable area in the 16S rRNA gene (Baker *et al.*, 2003). This highly variable region had high homology with a wide range of *Pseudomonas* species. The phylogenetic tree in Figure 5.1.2 on page 38 indicates the evolutionary similarities between the different *Pseudomonas* 16S rRNA sequences. There are many highly variable regions in the 16S rRNA and the sequencing of the complete 16S rRNA may have been able to identify the exact strain of this *Pseudomonas* bacteria. The identification of the exact *Pseudomonas* strain would have been carried out if the isolated polymer producing gene was novel. The 16S rRNA was only analysed after the genomic expression library was produced and screening of the library had already commenced.

Multiple antibiotic resistances are characteristic of many *Pseudomonas* species as they have multiple antibiotic resistance mechanisms (Livermore, 2011). The production of various biopolymers by different *Pseudomonas* species has been well studied and characterised. Some strains such as *Pseudomonas aeroginosa* are known to produce multiple biopolymers (Ryder *et al.*, 2007; Pham *et al.*, 2004). Biopolymer production was dependant on the availability of sucrose in the culture media, indicating that sucrose was necessary for biopolymer biosynthesis. It was expected that the biopolymer would have been an exopolysaccharide (EPS) as the bacteria had an amorphous slime phenotype which is characteristic of EPS producing bacteria (Sutherland, 1998).

The screening of the gDNA expression library on sucrose-enriched media in a heterologous system using *E. coli* was an effective method for the isolation of the gene responsible for the EPS biosynthesis. The positive clone TPS1 had an amorphous slime phenotype and had an insert size of ± 5.1 kb, which correlates with the size fraction of the digested gDNA used for the production of the expression library. The internal *Xho*I restriction site in the TPS1 clone allowed for the production of two insert fragments when digested with both *Xho*I and *Pst*I. The two fragments, XX (*Xho*I/*Xho*I) and PX (*Pst*I/*Xho*I) were inserted separately into pBluescript SK⁻ expression vectors. The clone containing the PX fragment, PX5, had an amorphous slime phenotype when cultured on media containing sucrose as

carbon source, which was similar to that of the TPS1 clone, indicating that the EPS biosynthesis gene was present in the ± 2.7 kb insert of the PX5 clone.

The PX5 clone was chosen for sequencing as this clone produced the EPS and had an insert size which was considerably smaller than the TPS1 clone.   Sequence data determined that the PX5 insert was 2639 bp in size.   Alignment of the sequence with the NCBI nucleotide collection database revealed that the first 1537 bp of the insert had the highest percentage homology to *Leuconostoc mesenteroides* and the highest percentage homology in a 485 bp fragment downstream to *Pseudomonas fluorescens* strain SS101 (Figure 5.2.5, page 42).   The PX5 insert sequence contained a 1275 bp open reading frame (ORF) which encoded a 424 amino acid levansucrase which has a calculated molecular weight of 47.1 kDa.

Levansucrase, also known as a fructansucrase, is a glycansucrase which catalyses the production of levan.   Levansucrase catalyses the transfer of the fructocyl residue from a sucrose molecule to another fructocyl residue or to a fructan chain.   Levan is a homo-polysaccharide consisting of only fructose molecules and is also known as a fructan.   The fructose molecules are linked by β-(2,6) bonds (Figure 6.1) and some levans contain branches, which are formed by β-(2,1) bonds (Han and Clarke, 1990; Monsan *et al.*, 2001). Levansucrase is an extracellular protein which is secreted from the cell and synthesises levan in the extracellular regions of the cell (Monsan *et al.*, 2001).
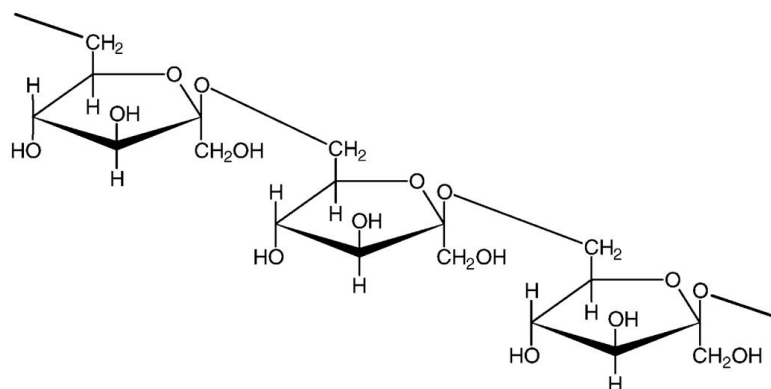

Figure 6.1: Chemical structure of a linear levan (Monsan *et al.*, 2001).

The PX5 levansucrase had specific homology to the glycosyl hydrolase (GH) family 68 (cd08997) which consists of frucosyltransferases (FTFs), including levansucrase (EC 2.4.1.10), β-fructofuranosidase (EC 3.2.1.26) and inulosucrase (EC 2.4.1.9) and falls under the GH43_62_32_68 superfamily (cd08772). It had a 100% homology to *L. mesenteroides* M1FT levansucrase in the nucleotide and amino acid sequences. This could have been due to horizontal gene transfer but it is highly unlikely as there was no sequence homology to *L. mesenteroides* in the sequence downstream of the levansucrase ORF in the PX5 clone insert. In a study by Morales-Arrieta *et al.* (2006) it was speculated that the M1FT levansucrase from *L. mesenteroides* isolated and characterised by Kang *et al.* (2005) was most probably due to genomic DNA contamination, as there was no ORF in the *L. mesenteroides* genome similar to the one reported for the M1FT gene. From the 16S rRNA data obtained in this study it is possible that this gene is in fact a mislabelled *Pseudomonas* gene. Because the PX5 levansucrase has a 100% homology to the levansucrase which was well characterised by Kang *et al.* (2005), no further analyses were planned for this levansucrase.

The EPS of the original biopolymer producing bacteria as well as the EPS produced by the positive clones, TPS1 and PX5 were analysed and it was determined that they were the same. Complete hydrolysis of the EPSs was accomplished and GC-MS analysis of the hydrolysed EPSs confirmed that there was only fructose molecules present in the EPS as the total ion chromatograms (TICs) of the EPSs correlated to the TIC of the fructose standard. This was expected, as the sequence data of the PX5 clone encoded a poly-fructose manufacturing gene. The levan produced by the original biopolymer producing bacteria had a molecular weight of ± 5 kDa which correlated to the molecular weight of ± 27 fructose molecules. The size is relatively small as microbial levans can contain as many as 3 million fructose molecules (French, 1989). This could be due to a short incubation period of the bacteria during polymer biosynthesis, as well as the depletion of the sucrose substrate in the culture medium.

NMR analysis of the three levans also determined that they were the same on the ^{13}C, ^{1}H and 2D HSQC level. The ^{13}C NMR spectra peaks were correlated to published levan ^{13}C NMR spectra peaks in Table 5.3.1 on page 47. The peaks generated in this study were close to, but not at the exact positions of the published data. The distances between the peaks correlated to the distances in the published data. The spectra could have shifted due to a difference in pH from the published data as the pH of the levans in this study was not

standardised before NMR analysis. The carbon numbers were assigned to the peaks in correlation with the published data (Kang *et al*., 2005; Shimamura *et al*., 1987).

Unfortunately the PX5 levansucrase was not a novel biosynthesis gene. The experimental design and method in which the gene was isolated and identified was effective and could be incorporated into similar studies in the future.

7 CONCLUSION

The unknown bacteria were from the *Pseudomonas* species and the specific strain could possibly be identified with analysis of the complete 16S rRNA gene. It was discovered that a levansucrase was producing a levan exopolysaccharide. Unfortunately the gene identified in this study is not a novel levan or novel polysaccharide producing gene. No further analyses are planned for this study, as there are no future applications planned for this levansucrase. In future the bacterial species, as well as the biopolymer of interest, should first be analysed in depth to determine novelty before pursuing a search for the biosynthesis genes.

9 REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology **215**: 403-410

Alves VD, Ferreira AR, Costa N, Freitas F, Reis MAM, Coelhoso IM (2011) Characterization of biodegradable films from extracellular polysaccharide produced by *Pseudomonas oleovorans* grown on glycerol by product. Carbohydrate Polymers **83**: 1582-1590

Anderson AJ, Haywood GW, Dawes EA (1990) Biosynthesis and composition of bacterial poly(hydroxyalkanoates). International Journal of Biological Macromolecules **12**: 102-105

Arakawa Y, Wacharotayankun R, Nagatsuka T, Ito H, Kato N, Ohta M (1995) Genomic organization of *Klebsiella pneumonia cps* region responsible for sterotype K2 capsular polysaccharide synthesis in the virulent strain chedid. Journal of Bacteriolology **177(7)**: 1788-1796

Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S primers (Review article). Journal of Microbiological Methods **55(3)**: 541-555

Barreras M, Salinas SR, Abdian PL, Kampel MA, Ielpi L (2008) Structure and mechanism of GumK, a membrane-associated glucuronosyltransferase. The Journal of Biological Chemistry **283(36)**: 25027-25035

Becker A, Katzen F, Pühler A, Ielpi L (1998) Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Applied Microbiology and Biotechnology **50**: 145-152

Bertani G (1951) Studies on lysogenesis: The mode of phage liberation by lysogenic *Escherichia coli*. Journal of Bacteriology **62**:293-300

Castañeda M, Sánchez J, Moreno S, Núñez C, Epsín G (2001) The global regulators GacA and σ^S form part of a cascade that controls alginate production in *Azotobacter vinelandii*. Journal of Batreriology **183(23)**: 6787-6793

Chen G (2009) A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. Chemical Society Reviews **38**: 2434-2446

Crage K, Jahns AC, Parlane N, Palanisamy R, Rasiah IA, Atwood JA, Rehm BHA (2009) Bacterial polyhydroxyalkanoate granules: biogenesis, structure, and potential use as nano-/micro-beads in biotechnological and biomedical applications. Biomacromolecules **10**: 660-669

Crini G (2005) Recent developments in polysaccharide-based materials used as absorbents in waste water treatment. Progress in Polymer Science **30**: 38-70

Doi Y (2002) Biotechnology: unnatural biopolymers. Nature Materials Letters **783(1)**: 207-208

Dramsi S, Magnet S, Davison S, Arthur M (2008) Covalent attachment of proteins to peptidoglycan. Federation of European Microbiological Societies Microbiology review **32**: 307-320

Fournier E (2001) Colorimetric quantification of carbohydrates. Current Protocols in Food Analytical Chemistry E1.1.1-E1.1.8

French AD (1989) Chemical and physical properties of fructans. Journal of Plant Physiology **134(2)**: 125-136

Gracía-Orchoa F, Santos VE, Casas JA, Gómez E (2000) Xanthan gum: production, recovery, and properties. Biotechnology Advances **18**: 549-579

Grage K, Jahns AC, Parlane N, Palanisamy R, Rasiah IA, Atwood JA, Rehm BHA (2009) Bacterial polyhydroxyalkanoate granules: biogenesis, structure, and potential use as nano-/micro-beads in biotechnological and biomedical applications. Biomacromolecules **10**: 660-669

Grönwall AJT, Ingelman BGA (1948) Manufacture of infusion and injection fluids. US patent 2437518 serial no. 536376

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series **41**: 95-98

Hall-Stoodley L, Costerton JW, Stoodley P (2004) Bacterial biofilms: from the natural environment to infectious diseases. Microbiology **2**: 95-107

Han YW, Clarke MA (1990) Production and characterization of microbial levan. Journal of Agricultural and Food Chemistry **38**: 393-396

Hickman JW, Harwood CS (2008) Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. Molecular Microbiology **69(2)**: 376-389

Ielpi L, Couso RO, Dankert MA (1981) Lipid-linked intermediates in the biosynthesis of xanthan gum. FEBS Letters **130(2)**: 253-256

Ielpi L, Couso RO, Dankert MA (1993) Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris.* Journal of Bacteriology **175(9)**: 2490-2500

Ionescu M, Belkin S (2009) Overproduction of exopolysaccharides by *Escherichia coli* K-12 *rpoS* mutant in response to osmotic stress. Applied and Environmental microbiology **75(2)**: 483-492

Jaffe M, Ophir Z, Pai V (2003) Biorelevant characterization of biopolymers. Thermochimica Acta **396**: 141-152

Joentgen W, Groth T, Steinbüchel A, Hai T, Oppermann FB (1998) Polyasparaginic acid homopolymers and copolymers: biotechnical production and use thereof. International patent application WO 98/39090

Kang HK, Seo MY, Seo ES, Kim D, Chung SY, Kimura A, Day DF, Robyt JF (2005) Cloning and expression of levansucrase from *Leuconostoc mesenteroides* B-512 FMC in *Escherichia coli.* Biochimica et Biophysica Acta **1727**: 5-15

Katzen F, Becker A, Zorreguieta A, Pühler A, Ielpi L (1996) Promotor analysis of the *Xanthomonas campestris* pv. campestris *gum* operon directing biosynthesis of the xanthan polysaccharide. Journal of Bacteriology **178(14)**: 4313-4318

Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Pühler A, Ielpi L (1998) *Xanthomonas campestris* pv. campestris *gum* mutants: effects on xanthan biosynthesis and plant virulence. Journal of Bacteriology **180(7)**: 1607-1617

Kessler B, Witholt B (2001) Factors involved in the regulatory network of polyhydroxyalkanoate metabolism. Journal of Biotechnology **86**: 97-104

Koller M, Salerno A, Dias M, Reiterer A, Braunegg G (2010) Modern biotechnological polymer synthesis: a review. Food Technology Biotechnology **48(3)**: 255-269

Lee SY (2006) Deciphering bioplastic production. Nature Biotechnology **24(10)**: 1227-1229

Leech AJ, Sprinkle A, Wood L, Wozniak DJ, Ohman DE (2008) The NtrC family regulator AlgB, which controls alginate biosynthesis in mucoid *Pseudomonas aeruginosa*, binds directly to the *algD* promoter. Journal of Bacteriology **190(2)**: 581-589

Lewis IA, Schommer SC, Markley JL (2009) rNMR: open source software for identifying and quantifying metabolites in NMR spectra. Magnetic Resonance in Chemistry **47**: S123-S126

Livermore DM (2011) Multiple mechanisms of antimicrobial resistance in *Pseudomonas aeruginosa*: Our worst nightmare? Clinical Infectious Diseases **34**: 634-640

Luong JHT, Mulchandani A, Leduy A (1988) Kinetics of biopolymer synthesis: a revisit. Enzyme Microbiology Technology **10**: 326-332

Lütke-Eversloh T, Fischer A, Remminghorst U, Kawada J, Marchessault RH, Bögershausen A, Kalwei M, Eckert H, Reichelt R, Liu S, Steinbüchel A (2002) Biosynthesis of novel thermoplastic polythioesters by engineering *Escherichia coli*. Nature Materials Letter **773(1)**: 236-240

Monsan P, Bozannet S, Albenne C, Joucla G, Willemot R, Remaud-Siméon M (2001) Hompolysaccharides from lactic acid bacteria. International Dairy Journal **11**: 675-685

Morales-Arrieta S, Rodríguez ME, Segovia L, López-Munguía A, Olvera-Carranza C (2006) Identification and functional characterization of *levS*, a gene encoding for a levansucrase from *Leuconostoc mesenteroides* NRRL B-512 F. Gene **376**: 59-67

Neuhaus FC, Baddiley J (2003) A continuum of anionic charge: structures and functions of ᴅ-alanyl-teichoic acids in gram-positive bacteria. Microbiology and Molecular Biology Reviews **67(4)**: 686-723

Oppermann-Sanio F B, Steinbüchel A (2001) Occurrence, function and biosynthesis of polyamides in microorganisms and biotechnological production. Naturwissenschaften **89**: 11-22

Ørskov I, Ørskov F, Jann B, Jann K (1977) Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriological reviews **41(3)**: 667-710

Parolis H, Parolis LAS, Boán IF, Rodríguez-Valera F, Widmalm G, Manca MC, Jansson P, Sutherland IW (1996) The structure of the exopolysaccharide produced by the halophilic Archeon *Haloferax mediterranei* strain R4 (ATCC 33500). Carbohydrate Research **295**: 147-156

Persin Z, Stana-Kleinschek K, Foster TJ, van Dam JEG, Boeriu CG, Navard P (2011) Challenges and opportunities in the polysaccharides research and technology: The EPNOE views for the next decade in the areas of materials, food and health care. Carbohydrate Polymers **84**: 22-32

Perugino G, Trincone A, Rossi M, Moracci M (2004) Oligosaccharide synthesis by glycosynthases. Trends in Biotechnology **22(1)**: 31-37

Pham TH, Webb JS, Rehm BHA (2004) The role of polyhydroxyalkanoate biosynthesis by *Pseudomonas aeruginosa* in rhamnolipid and alginate production as well as stress tolerance and biofilm formation. Microbiology **150**: 3405-3413

Rahn A, Drummelsmith J, Whitfield C (1999) Conserved organization in the *cps* gene cluster for expression of *Escherichia coli* group 1 K antigens: relationship to the colonic acid biosynthesis locus and the *cps* genes from *Klebsiella pneumonia*. Journal of Bacteriology **181(7)**: 2307-2313

Rehm BHA (2010) Bacterial polymers: biosynthesis, modification and applications. Applied and Industrial Microbiology **8**: 578-592

Rehm BHA, Valla S (1997) Bacterial alginates: biosynthesis and applications. Applied Microbiology and Biotechnology **48**: 281-288

Ruffing A, Chen RR (2006) Metabolic engineering of microbes for oligosaccharide and polysaccharide synthesis. Microbial Cell Factories **5**: 25

Remminghorst U, Rehm BHA (2006) Bacterial alginates: from biosynthesis to applications. Biotechnology Letter **28**: 1701-1712

Ryder C, Byrd M, Wozniak DJ (2007) Role of polysaccharides in *Pseudomonas aeruginosa* biofilm development. Current Opinion in Microbiology **10**: 644-648

Sakuragi Y, Kolter R (2007) Quorum-sensing regulation of the biofilm matrix genes (*pel*) of *Pseudomonas aeruginosa*. Journal of Bacteriology **189(14)**: 5383-5386

Schwartz M, Harrell F Jr, Rossini A, Francis I (2008) R: Regulatory compliance and validation issues a guidance document for the use of R in regulated clinical trial environments. The R Foundation for Statistical Computing, c/o Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Augasse 2-6

Sheng F, Jia X, Yep A, Preiss J, Geiger JH (2009) The crystal structures of the open and catalytically competent closed conformation of *Escherichia coli* glycogen synthase. The Journal of Biological Chemistry **284**: 17796-17807

Shimamura A, Tsuboi K, Nagase NT, Ito M, Tsumori H, Mukasa H (1987) Structural determination of D-fructans from *Streptococcus mutans*, serotype d, c, e, and f strains by ^{13}C-NMR spectroscopy. Carbohydrate Research **165**: 150-154

Stein SE (1999) An Integrated method of spectrum extraction and compound identification from GCMS. Journal of the American Society of Mass Spectrometry **10**: 770-781

Sutherland IW (1998) Novel and established applications of microbial polysaccharides. Tibtech **16**: 41-46

Taguchi S, Yamada M, Matsumoto K, Tajima K, Satoh Y, Munekata M, Ohno K, Kohda K, Shimamura T, Kambe H, Obata S (2008) A microbial factory for lactate-based polyesters using a lactate-polymerizing enzyme. Proceedings of the National Academy of Sciences **105(45)**: 17323-17327

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology Evolution **28(10)**: 2731-2739

Tao Y, Zhang L, Yan F, Wu X (2007) Chain conformation of water-insoluble hyperbranched polysaccharide from fungus. Biomacromolecules **8**: 2321-2328

Tu Z, He G, Li KX, Chen MJ, Chang J, Chen L, Yao Q, Liu DP, Ye H, Shi J, Wu X (2005) An improved system for competent cell preparation on high efficiency plasmid transformation using different *Escherichia coli* strains. Electronic Journal of Biotechnology **8**: 113-120

Weber H, Pesavento C, Possling A, Tischendorf G, Hengge R (2006) Cyclic-di-GMP-mediated signalling within the σ^S network of *Escherichia coli*. Molecular Microbiology doi: 10.1111/j.1365-2958.2006.05440.x

Whitfield C (1995) Biosynthesis of lipopolysaccharides O antigens. Trends in Microbiology **3**: 178-185

Whitfield C (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. The Annual Review of Biochemistry **75**: 39-68

Wilson WA, Roach PJ, Montero M, Baroja-Fernández E, Muñoz FJ, Eydallin G, Viale AM, Pozuetta-Romero J (2010) Regulation of glycogen metabolism in yeast and bacteria. Federation of European Microbiological Societies Microbiology Review **34**: 952-985

Valla S, Li J, Ertesvãg H, Barbeyron T, Lindahl U (2001) Hexuronyl C5-epimerases in alginate and glycosaminoglycan biosynthesis. Biochimie **83(8)**: 819-830

van Kranenburg R, Boels IC, Kleerebezem M, de Vos WM (1999) Genetics and engineering of microbial exopolysaccharides for food: approaches for the production of existing and novel polysaccharides. Current Opinion in Biotechnology **10**: 498-504

Vollmer W, Blanot D, de Pedro MA (2008) Peptidoglycan structure and architecture. Federation of European Microbiological Societies Microbiology review **32**: 149-167

Vorhölter F, Schneiker S, Goesmann A, Krause L, Bekel T, Kaiser O, Linke B, Patschkowski T, Rückert C, Schmid J, Sidhu VK, Sieber V, Tauch A, Watt SA, Weisshaar B, Becker A, Niehaus K, Pühler A (2008) The genome of *Xantomonas camprestis* pv. camprastis B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. Journal of Biotechnology **134(1-2)**: 33-45

Yang L, Zhang L (2009) Chemical structural and chain conformational characterization of some bioactive polysaccharides isolated from natural sources. Carbohydrate Polymers **76**: 349-361

APPENDIX

516 bp sequence of a highly variable region in the 16s rRNA

5'-
GAGTTTGATCCTGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGA
GCGGTAGAGAGAAGCTTGCTTCTCTTGAGAGCGGCGGACGGGTGAGTAATGCCTAGG
AATCTGCCTGGTAGTGGGGGATAACGTTCGGAAACGGACGCTAATACCGCATACGTC
CTACGGGAGAAAGCAGGGGACCTTCGGGCCTTGCGCTATCAGATGAGCCTAGGTCGG
ATTAGCTAGTTGGTGAGGTAATGGCTCACCAAGGCGACGATCCGTAACTGGTCTGAGA
GGATGATCAGTCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCAG
TGGGGAATATTGGACAATGGGCGAAAGCCTGATCCAGCCATGCCGCGTGTGTGAAGA
AGGTCTTCGGATTGTAAAGCACTTTAAGTTGGGAGGAAGGGTTGTAGATTAATACTCT
GCAATTTTGACGTTACCGACAGAATAAGCACCGGCTAACTCTGTGCCAGCCGCCGCG
G -3'