NTNU – Trondheim
Norwegian University of
Science and Technology

Modification and Characterization of the Catalytic Subunit, the A-module, of the Mannuronan C-5 Epimerases AlgE4 and AlgE6

Lene Brattsti Dypås

Chemical Engineering and Biotechnology
Submission date: Januar 2015
Supervisor: Gudmund Skjåk-Bræk, IBT
Co-supervisor: Finn Lillelund Aachmann, IBT

Norwegian University of Science and Technology
Department of Biotechnology

Preface

This work was submitted as the final thesis of a five year Master of Chemical Engineering and Biotechnology at the Norwegian University of Science and Technology, and it was written at the Department of Biotechnology. My head supervisor at the department was Professor Gudmund Skjåk-Bræk and I would like to thank him for this opportunity to work as part of an engaging research group, and for all his enthusiasm whenever I had news from the lab. My immediate supervisor was PhD Finn Lillelund Aachmann, and my deepest thanks are due for all his help throughout this process, his engaging lessons on various subjects, and also for trusting that I would figure most of it out on my own.

Special thanks are extended to Gerd Inger Sætrom for our countless discussions of procedures, problems, and results, as well as all other completely unrelated topics. I am also grateful for the always welcoming Wenche Iren Strand, for all her help whenever I stopped by the biopolymer lab, and also for providing me with NMR spectra of "countless" samples. Thanks to Olav Andreas Aarstad for analysing my samples by HPAEC-PAD, and for patiently explaining me the process in detail. Also, thanks to Annalucia Stanisci for greeting me with a smile and advice as I stopped by her office. Finally, thanks to Merethe Christensen Vadseth and Tonje Stavne for guiding me every time I got lost looking for chemicals or equipment.

Lastly, I would like to thank my parents and older brother, for always being equally excited about my work and results as I have been myself, despite not really understanding what I was on about most of the time.

Abstract

Alginate is a family of polysaccharides composed of (1→4)-linked β-D-mannuronic acid (M) and α-L-guluronic acid (G), with a wide variety of applications and even greater prospects for the future. Alginate is first synthesized as poly-β-D-mannuronic acid and the G-residues are introduced by enzymatic C-5 epimerization of individual M-residues, with different enzymes introducing different degrees of epimerization and patterns. The physical properties of alginates are highly dependent on their chemical composition, and their great degree of sequential and compositional variation creates the need for structural control. By expressing mannuronan C-5 epimerases in their native, or even hybrid form, in alginate-producing bacteria, it could be possible to produce alginates with carefully designed properties and a high level of compositional homogeneity. This is especially promising for the strict demands of biomedical applications.

Great efforts have been made to obtain detailed characterization of the C-5 epimerases and their interactions with alginates, yet much is still unclear. In particular, a family of seven extracellular epimerases of *Azotobacter Vinelandii*, AlgE1-7, has been extensively studied. They all consist of two structural domains; the A-module, which has been found to hold the catalytic activity, and the R-module, whose role is not yet defined, but has been suggested as regulatory.

This study, in two parts, aims at further clarifying the role of the R-module, specifically in AlgE4, and to learn more about the structure function relationship in the A-module of AlgE6.

The first part is a comparative study of the A-module in AlgE4 and the full length epimerase. The result of this work supported previous findings of the R-module having a role in lowering the concentration of calcium ions needed for full activity in AlgE4. They also suggested that the R-module facilitates a stable degree of epimerization within a wider calcium concentration range, but that it was not needed to achieve full degree of epimerization. No evidence was found to suggest that the R-module gives rise to product specificity, but it did increase processivity, possibly by

contributing to stronger association with the alginate polymer. Results concerning temperature dependence were inconclusive, and further work is needed.

In the second part, structures in the C-terminal region of the A-module were studied by creating three AlgE6A clones with varying degree of truncation at the C-terminal end. It was hypothesized that removal of the first 49 residues (AlgE6A_336) would give a functional epimerase, removal of 90 residues (AlgE6A_295) would significantly reduce epimerase activity, and removal of 114 residues (AlgE6A_272) would give a non-functional epimerase. Only AlgE6A_336 resulted in introduction of G-residues, and the G-content in its product was minimal. Structures at the C-terminal end were therefore revealed to be more essential for activity than first thought. Results suggested that the loop extending from the β-helix could be especially important, as the side of this loop became exposed in AlgE6A_336, where almost all activity was lost, and then completely removed in AlgE6A_295 along with complete loss of function. It was also noted that removal of the R-module is a less invasive modification as it cuts the protein at a transition between structurally separate domains, while the truncation at the C-terminal cuts within one such defined domain. Structural analysis of the AlgE6A clones would be the next step in this work, to identify whether improper folding could be the reason behind this loss of function.

Sammendrag

Alginat er en familie polysakkarider sammensatt av (1→4)-bundet β-D-mannuronat (M) og α-L-guluronat (G), med et bredt spekter av bruksområder. Alginat syntetiseres først som som ren poly-β-D-mannuronat og G-residuene introduseres senere ved enzymatisk C-5-epimerisering av individuelle M-residuer. De fysiske egenskapene til alginat er i høy grad avhengig av deres kjemiske sammensetning, og grunnet den enorme variasjonen i kjedelengde, G-innhold, og monomersekvens, er det et stort behov for å bedre kunne kontrollere disse parameterene. Ved å uttrykke mannuronan-C-5-epimeraser, og hybrider av de, i alginatproduserende bakterier, kan det være mulig å produsere alginater med nøyaktig planlagte egenskaper og et høyt nivå av homogenitet. Disse "design-alginatene" er spesielt lovende for de høye kvalitetskravene innenfor medisinske bruksområder.

Mye arbeid har allerede blitt gjort for å oppnå detaljert karakterisering av C-5-epimerasene og deres interaksjoner med alginat. En familie av syv ekstracellulære epimeraser fra *Azotobacter vinelandii*, AlgE1-7, har vært særlig omfattende studert, men mye er fremdeles uklart. Disse epimerasene er alle sammensatt av to typer moduler; A-modulen, som inneholder det katalytiske setet, og R-modulen, hvis rolle ikke har blitt definert, men som er foreslått å være regulatorisk.

Denne studien, i to deler, tar sikte på å bedre forstå rollen til R-modulen, spesifikt i AlgE4, og å lære mer om forholdet mellom struktur og funksjon i A-modulen til AlgE6.

Første del er en sammenligning av den enzymatiske aktiviteten til AlgE4 og en modifisert epimerase kun bestående av dens A-modul. Resultatet av dette arbeidet støtter tidligere funn som tilsier at R-modulen senker konsentrasjonen av kalsiumioner nødvendig for full aktivitet. Det ble også funnet at R-modulen legger til rette for en stabil grad av epimerisering innenfor større variasjoner i kalsiumkonsentrasjon, og at R-modulen ikke er nødvendig for å oppnå full epimerisering av mannuronan. Det ble ikke funnet noe som støtter spekulasjoner om at R-modulen påvirker produktspesifisiten når det gjelder formasjon av G- eller MG-blokker, men

det ble vist at den øker prosessiviteten. Det kan tenkes at R-modulen øker prosessivteten ved å bidra til sterkere assosiasjon med alginatpolymeret. Resultater vedrørende temperaturavhengiget var uklare, og videre arbeid er nødvendig.

Andre del av studien undersøker strukturer i C-terminalregionen til A-modulen, ved å lage tre AlgE6A-kloner hvor deler av C-terminalregionen er fjernet. Det ble antatt at det å kutte av de første 49 aminosyreresiduene fra C-terminal ville resulter i en funksjonell epimerase (AlgE6A_ 336), 90 residuer ville redusere epimerase aktiviteten betydelig (AlgE6A_295), mens 114 residuer ville resultere i tap av all aktivitet (AlgE6A_272). Kun AlgE6A_336 resulterte i epimerisering, og G-innholdet i produktet var minimalt. Strukturer i C-terminalregionen ble derfor funnet å være viktigere for aktivitet enn først antatt. Resultater tyder også på at løkken som strekker seg ut fra β-heliksen kan være spesielt viktig, da den ene siden av løkken blir eksponert i AlgE6A_336, hvor nesten all aktivitet går tapt, og i AlgE6A_295 blir den fjernet helt, sammen med den resterende epimerasefunksjonen. Det ble også notert at det å ta bort R-modulen er et mindre invasivt inngrep da proteinet blir kuttet ved en klar overgang i strukturen mellom to separate områder, mens det å kutte av deler av C-terminalen innvolverer å fjerne deler innenfor ett slikt strukturelt avgrenset område. Analyser av strukturen til AlgE6A-klonene vil være et naturlig neste steg i dette arbeidet, for å identifisere om misfolding kan være årsaken bak tapet av funksjon.

Symbols and Abbreviations

ε	Molar Extinction Coefficient [$M^{-1}cm^{-1}$]
G	α-L-guluronic acid
M	β-D-mannuronic acid
A_λ	Absorbance at wavelength λ
c	Concentration [mol/L]
l	Light path length [cm]
Amp	Ampicillin
b-ME	b-Mercaptoethanol
CBD	Chitin Binding Domain
DP	Degree of Polymerization
DTT	1,4-Dithiothreitol
FPLC	Fast Protein Liquid Chromatography
HPAEC-PAD	High-Performance Anion-Exchange Chromatography with Pulsed Amperometric Detection
IMPACT	Intein Mediated Purification with an Affinity Chitin-binding Tag
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LA	Lysogeny Agar
LB	Lysogeny Broth

MCS	Multiple Cloning Site
NMR	Nuclear Magnetic Resonance
poly-M	Mannuronan
RU	Relative Units (amount of enzyme that catalyses the conversion of 1 µmol substrate per minute)
SDS-PAGE	Sodium Dodecyl Sylfate Polyacrylamide Gel Electrophoresis
SOC	Super Optimal Broth with Catabolite Repression
Spc	Spectinomycin
TAE	Tris-Acetate-EDTA
TSP	3-(Trimethylsilyl)propionic-2,2,3,3-d_4 acid
TTHA	Triethylenetetraminehexaacetic acid

Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Alginate

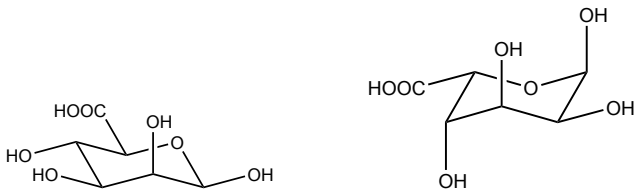
Alginate is a family of unbranched polysaccharides with a wide variety of current and potential future applications. Their stabilizing and viscosifying properties and ability to form hydrogels have made them important in both industry and medicine, where food thickeners, electrochemical biosensors, textile printing, biomaterials, drug delivery, and cell imobilization are only some of the uses for these biopolymers. Alginates are primarily found in the cell walls of brown algae, but they are also produced by some bacteria and have been identified in *Azotobacter vinelandii* and members of the genus *Pseudomonas*.

1.1.1 Structure

Alginate is a linear molecule composed of (1→4)-linked β-D-mannuronic acid (M) and α-L-guluronic acid (G). The arrangement of the monomers is that of a block copolymer with consecutive sequences of G's (G-blocks), consecutive sequences of M's (M-blocks), and alternating sequences of both M and G (MG-blocks) (Haug et al., 1966, 1967). Unlike many other polysaccharides, the composition of alginate does not follow a specific pattern and there are large variations, both in the monomer ratio and in the monomeric sequence. In addition to its heterogeneity in composition, alginate is also polydisperse with a wide range of molecular weight distributions.

G is the C-5 epimer of M and while M exists in the $^4C_1$ conformation, the $^1C_4$ conformation is more energetically favourable for G due to the changed configuration of the carboxyl group on C-5 (Atkins et al., 1970, 1971; Smidsrød et al., 1973). The molecular structure of the monomers and their conformation is presented in Figure 1.1. The hydroxyls at C-1 and C-4 are equatorial in M, but they become axial in G, which leads to important differences in the glycosidic linkage between two monomers. Linkages are di-equatorial (MM) in M-blocks, di-axial (GG) in G-blocks,

and equatorial-axial (MG) or axial equatorial (GM) in MG-blocks. Figure 1.2 shows the structure of the three types of blocks, illustrating the effect produced by the different linkages.



(a) (β-D-mannuronic acid (M) **(b)** α-L-guluronic acid (G)

Figure 1.1: Molecular structure of the alginate monomers (a) β-D-mannuronic acid (M) and (b) α-L-guluronic acid (G) in their respective conformations $^4C_1$ and $^1C_4$.

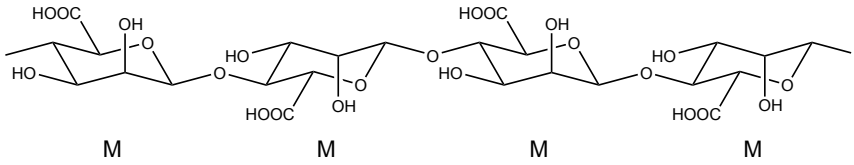
1.1.2 Physical Properties

Solubility

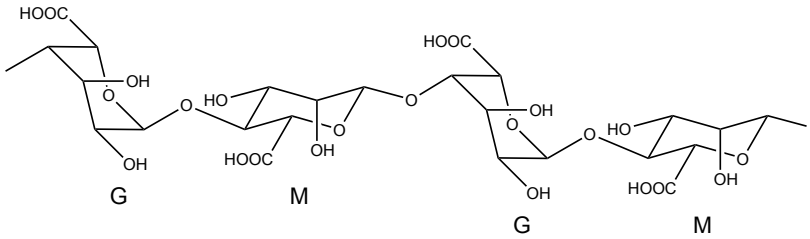
The measured pK_a values of 3.38 for β-D-mannuronic acid and 3.65 for α-L-guluronic acid, means that the carboxylic groups are fully deprotonated at physiological pH, making alginate a polyanion (Haug, 1964). Alginate is therefore soluble at pH>3.5 while an abrupt lowering to pH below 3 will cause precipitation of alginic acid. Decreasing the pH in a controlled manner by the use of a slowly hydrolyzing acid, e.g. GDL (D-glucono-δ-lactone), will instead result in acidic gel formation (Draget et al., 1994).

The polyanion character of alginate causes its physical properties in solution to be highly dependent on the type of neutralizing cation present. Alginate salts of monovalent cations will generally be soluble in water. Counterions dissociate from the polyelectrolyte in solution and the positive contribution to the entropy of mixing caused by this dissociation is the drive behind alginate solubility. This drive from counterion dissociation will decrease with increasing ionic strength, concurrently decreasing alginate solubility, and it is therefore recommended to dissolve alginates in pure water prior to adding salts in high concentrations (Draget et al., 2005). Most divalent metal ions will, however, reduce solubility and cause gel formation (Haug and Smidsrød, 1965).

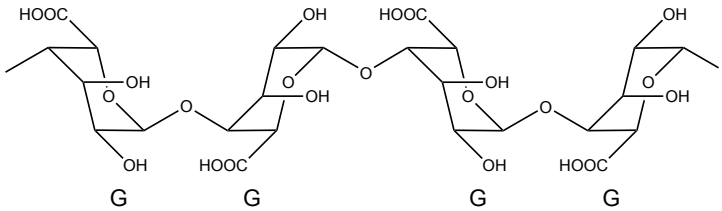
The monomer sequence will also affect the solubility of alginates, and at low pH the solubility increases with the presence of block types in the order G-blocks<M-blocks<MG-blocks (Draget et al., 2005). Solubility of alginates is therefore depen-



(a) M-blocks



(b) MG-blocks



(c) G-blocks

Figure 1.2: Molecular structure of the three different block types found in alginate, (a) M-blocks, (b) MG-blocks, and (c) G-blocks.

dent on a combination of pK_a values, monomer sequence, concentrations of ions and polymer, and even the chain length due to the fact that longer chains have lower entropy, rendering them less soluble.

Viscosity

The intrinsic viscosity [μ] of a polymer is a measure of its ability to increase the viscosity of a specific solvent, and it is strongly dependent on polymer conformation. More extended polymer chains will have higher viscosity, and factors such as pH and the presence of ions greatly influence chain extension. Other factors that affect the viscosity of an alginate solution are alginate concentration, molecular weight, and monomer sequence. Alginates adapt a very extended random coil formation in aqueous solution due to intramolecular electrostatic repulsion and their inherent chain stiffness, and alginate solutions are therefore highly viscous in general (Smidsrød and Haug, 1968b; Smidsrød, 1970).

Gel Formation

Binding of divalent cations, Ca^{2+} in particular, mediates alginate gel formation by cross-linking of polymer chains, and most commercial applications of alginates are based on their unique gel formation characteristics of thermo-irreversibility.

The affinity of alginates towards divalent metal ions has been shown to not be purely electrostatic in nature, as it is also dependent on the polymer sequence (Smidsrød and Haug, 1968a). M- and MG-blocks show little to no preference for the different metal ions, while G-blocks have an increasing binding affinity in the order Mg^{2+}<Ca^{2+}<Sr^{2+}<Ba^{2+} (Haug and Smidsrød, 1965). The binding of calcium ions to G-blocks is a cooperative process where the affinity towards the ion increases with increasing concentrations of the ion (Smidsrød and Haug, 1972). This binding of calcium ions to G-blocks is the basis for alginate gel formation by polymer cross-linking, and it is described by the so called "egg-box" model (Grant et al., 1973). In this model the ion-binding is explained by the conformations of the monomer residues in the polymer chain where the glycosidic linkages in G-blocks cause the sugar rings of the residues to orient close to perpendicular relative to the polymer chain, forming a cavity as seen in Figure 1.2. Figure 1.3 illustrates how calcium ions positioned between G-blocks fit into these cavities and coordinate with hydroxyl and carboxyl groups from both chains, forming junction zones between the chains. Formation of several junction zones gives rise to the continuous network of cross-linked polymer chains and gel formation. The role of MG-blocks in the formation of junction zones has been studied by NMR analysis, and formation of junction zones

between MG-blocks, as well as mixed junction zones between MG and G-blocks, was demonstrated (Donati et al., 2005).

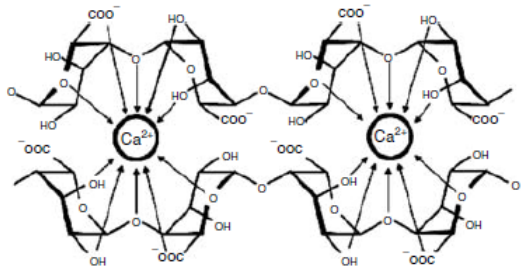


Figure 1.3: Calcium ions coordinated with hydroxyl and carboxyl groups in the cavities of G-blocks, forming a junction zone between two alginate polymer chains in accordance with the "egg-box" model. Illustration from Donati and Paoletti (2009)

1.1.3 Sources

Large quantities of alginate can be found in species of brown algae where it comprises up to 40 % of the dry matter and occurs mainly in the cell walls and intracellular mucilage, providing mechanical strength and flexibility (Painter, 1983). A calcium binding substance produced in the red algae *Serraticardia maxima* has been identified as alginate (Okazaki et al., 1982) and several bacteria produce alginates as an extracellular matrix. In *Azotobacter vinelandii* and *Azotobacter chroococcum* (Gorin and Spencer, 1966; Cote and Krull, 1988), alginate has been found as part of a protective cyst during metabolic dormancy (Sadoff, 1975) and several members of the genus *Pseudomonas* also produce bacterial alginates (Govan et al., 1981), including the pathogen *P. aeruginosa* associated with patients suffering from cystic fibrosis (Iacocca et al., 1963; Linker and Jones, 1966). *P. aeruginosa* produces a thick alginate-containing biofilm that enhances adhesion to solid surfaces and protects it from the hosts immune system and antibiotics, while at the same time contributing to severe deterioration of lung function in the cystic fibrosis patient (Lyczak et al., 2002).

The composition of alginates reflects their function and is highly dependent on the source. In brown algae, alginate composition has been shown to vary depending on season, growth conditions, species, type of tissue, and even the age of the tissue (Andresen et al., 1977; Haug et al., 1974). The G-content increases with age and the variation between sources is greater for old tissues. In contrast with algal alginates, the bacterial alginates are acetylated in the C-2 and/or C-3 position

on M-residues, and the degree of acetylation varies relative to the content of this residue (Skjåk-Bræk et al., 1985). These O-acetyl groups play a role in controlling monomer sequence and chain length as they have been shown to prevent epimerization by the mannuronan C-5 epimerase AlgG and hydrolysis by the lyase AlgL (Franklin and Ohman, 1993; Franklin et al., 1994). Mannuronan C-5 epimerases and alginate lyases are further discussed in Section 1.2 and Section 1.3, respectively.

Variation in alginate composition during the life cycle of *A. vinelandii* is highly representative of how function is reflected in composition (Sadoff, 1975). Alginates involved with the outer layers during encystment have been found to contain a much higher G-content when compared to alginates closer associated with the vegetative cell, and the absence of calcium ions in an encystment medium leads to aborted encystment, indicating the need for introduction of G by the calcium dependent AlgE epimerases (Page and Sadoff, 1975, 1976). The *A. vinelandii* mannuronan 5-C epimerases are further discussed in Section 1.2.1.

1.1.4 Applications

Alginates have already been applied in a wide range of industrial and medical fields, and these widespread uses reflect the heterogeneity in the chemical structure of alginates. Applications of alginate are mostly based on their rheological properties or their unique thermo-irreversible gelling properties. Today, the bulk alginate used for industry purposes is extracted from brown algae and the annual production was in 2009 estimated to 38,000 tonnes (Helgerud et al., 2009). Biomedical applications however usually require a higher level of compositional homogeneity than algal alginates can provide, making bacterial alginates important for future needs. The production of pure mannuronan in an *algG* mutant of *Pseudomonas fluorensence* (Gimmestad et al., 2003) and successful expression of specific and modified epimerases in *E. coli* (Ertesvåg et al., 1994, 1995; Svanem et al., 1999), has opened up for the possibility of producing "tailor-made" alginates with high homogeneity and targeted physical properties by controlling epimerase expression in alginate-synthesising bacteria (Ertesvåg et al., 1996; Mørch et al., 2007).

Many industrial applications make use of the viscosifying, stabilizing and water retaining properties of alginates, and examples include alginates as shear-thinning viscosifyers in textile printing, coating for surface uniformity in paper production, and binding agents in the production of ceramics and welding rods (Onsøyen, 1996). Alginates are also widely used in the food industry as additives, and common applications are as thickeners in jams and different sauces, and as a stabilizer in ice cream to achieve more desirable melting characteristics (Brownlee et al., 2005; Regand and Goff, 2003). Reconstruction of food to make them more appealing is also commonly

achieved with alginates, and examples include onion rings and pimento sections in olives, where alginate facilitates the production of uniform size and consistency (Brownlee et al., 2005).

Treatment of oesophageal reflux is an important pharmaceutical application of alginates, where an alginate solution with sodium bicarbonate forms a carbonated gelatinous foam when it comes in contact with gastric acid (Banning et al., 2002). When reflux occurs, this foam will float on top of the stomach contents as a protective barrier. Alginates are also extensively used as wound dressing in surgery and other clinical wound management (Thomas, 2000), and their properties as hydrogels are being utilized in tissue engineering (Augst et al., 2006).

Techniques have been developed for encapsulation and immobilisation of living cells using calcium-alginate beads and holds exciting prospect for use in cell transplantation (Smidsrød and Skjåk-Bræk, 1990). The purpose of the encapsulating gel is to act as immunoisolation of the transplanted cells, while still allowing for free diffusion of nutrients, signalling molecules and the therapeutics produced by the cell. Several cell systems are of interest in this type of therapy, but treatment of type 1 diabetes by transplantation of insulin-producing Islets of Langerhans is the most extensively studied (Skjåk-Bræk and Espevik, 1996). Alginates rich in mannuronic acid have been found to induce an inflammatory response (Otterlei et al., 1991), and while this poses a challenge for its use as a cell immobilization matrix it has also generated interest in possible uses for the biological activity of the molecule itself (Skjåk-Bræk and Espevik, 1996).

The adjustable rheological properties have also made alginate interesting as a matrix in drug delivery systems, where they can be exploited for controlled diffusion or delayed release of various drugs (Gombotz and Wee, 1998; Shilpa et al., 2003). This concept has been studied for the use of Mn-alginate gels for controlled release of Mn^{2+} for optimized application of manganese-enhanced magnetic resonance imaging (Mn-MRI) (Mørch et al., 2012).

1.2 Mannuronan C-5 Epimerases

Alginate biosynthesis starts with production of the homopolymer mannuronan, and introduction of G residues occurs later by epimerization of M residues in the polymer chain. This epimerization step is, in all studied alginate-producing organisms, catalysed by a family of enzymes called mannuronan C-5 epimerases (Valla et al., 2001). The presence of such an enzyme was first discovered in *A. vinelandii* (Larsen and Haug, 1971; Haug and Larsen, 1971) and soon after in the marine brown algae *Pelvetia canaliculata* (Madgwick et al., 1973). Mannuronan C-5 epimerase activ-

ity has later been reported in several other species of brown algae (Ishikawa and Nisizawa, 1981; Greene and Madgwick, 1986; Nyvall et al., 2003) and in bacterial species of the genus *Pseudomonas* (Franklin et al., 1994; Penaloza-Vazquez et al., 1997; Morea et al., 2001). Seeing as all alginates are produced from homopolymeric mannuronan by enzymatic epimerization, the chemical composition of alginate, and thus its physical properties, is determined by the catalytic properties of the epimerases.

Originally isolated from *P. aeruginosa* (Franklin et al., 1994), the periplasmic and calcium independent mannuronan C-5 epimerase AlgG is a conserved part of the alginate biosynthesis in all known alginate-producing bacteria (Rehm et al., 1996). AlgG, but not its epimerase activity, is needed for alginate polymer formation and an *algG* mutant of *P. fluoresence* produces pure polymannuronan (Gimmestad et al., 2003). In addition to AlgG, *A. vinelandii* also expresses seven extracellular calcium dependent mannuronan C-5 epimerases (AlgE1-AlgE7) that all have similar structures based on two types of modules (Ertesvåg et al., 1994, 1995; Svanem et al., 1999). These epimerases have all been cloned and expressed in *E. coli* and even though they share sequence homology and structure similarity, they exhibit different substrate specificities and introduce different patterns of epimerization. A calcium dependent epimerase with modular structure similar to the AlgE epimerases of *A. vinelandii* and a combined function as an *O*-acetylhydrolase has been isolated from *P. syringae* (Bjerkan et al., 2004b).

Much less is known about algal epimerases, but sequence analyses have indicated that the genome of *Laminaria digitata* encodes at least 21 mannuronan C-5 epimerases, and isolations of six different cDNAs (ManC5-E1 to ManC5-E6) show homology with the bacterial mannuronan C-5 epimerases (Nyvall et al., 2003).

1.2.1 The Extracellular Mannuronan C-5 Epimerases of *A. Vinelandii*

General Structure

AlgE1-AlgE7 are all modular enzymes composed of two different structural subunits, the catalytic A-module and the regulatory R-module, with a C-terminal peptide called the S-motif on the last R-module (Ertesvåg et al., 1994, 1995). AlgE1 contains two A-modules and they have both shown epimerase activity when expressed separately in *E. coli*, while addition of at least one R-module increased reaction rates and separate expression of an R-module was inactive (Ertesvåg and Valla, 1999). This indicates that the A-modules contain the catalytic activity, while the R-modules posses accessory and regulatory properties. The function of the S-motif is unknown (Ertesvåg et al., 1995).

Each epimerase consists of one or two A-modules and one to seven R-modules, and their modular structures, molecular weights, and patterns of epimerization are presented in Table 1.1 (Ertesvåg et al., 1998, 1999). The degree of homology in the sequences of the modules varies between the different epimerases, as well as within a single epimerase (Ertesvåg et al., 1999). AlgE2 and AlgE5 have a high degree of structural homology, as does AlgE1 and AlgE3, except for their C-terminal R-modules. AlgE4 and AlgE6 are also structurally similar with a high degree of homology in their A-module, but not in their R-modules, and AlgE6 contains three R-modules compared to only one in AlgE4.

Table 1.1: The extracellular mannuronan C-5 epimerases AlgE1-7 of *A. vinelandii* and their respective molecular weights, modular structures and enzymatic activities (Ertesvåg et al., 1998, 1999). The long rectangles represent the A-modules, and their sequential homology is reflected by their shading. The small squares represent the R-modules, and their sequential homology is reflected by Greek letters.

	[kDa]	**Modular structure**	**Enzymatic activity**
AlgE1	147.2	α β γ ε	GG- and MG-epimerase
AlgE2	103.1	α β γ δ	GG-epimerase
AlgE3	191.0	α β γ α β γ δ	GG- and MG-epimerase
AlgE4	57.7	α	MG-epimerase
AlgE5	103.7	α β γ δ	GG-epimerase
AlgE6	90.2	ε ε ε	GG-epimerase
AlgE7	90.4	ζ η δ	GG-epimerase and lyase

X-ray crystallography studies of the A-module of AlgE4 have revealed its structure as a right-handed parallel β-helix with an amphipathic α-helix near the N-terminal end (Rozeboom et al., 2008). Four parallel β-sheets and 12 complete turns make up the β-helix motif, and the N-terminal end of the helix is "capped" by the amphipathic α-helix. Extending from the β-helix motif is a loop formed by a two-stranded anti-parallel β-sheet, made up of residues 307-309 and 316-318. The catalytically important amino acid residue Asp^{152} (Svanem et al., 2001), is surrounded by loops extending from the surface, forming a positively charged cleft identified as the catalytic site (Rozeboom et al., 2008). In addition to Asp^{152}, three other amino acid residues, Tyr^{149}, His^{154}, and Asp^{178}, have been identified as absolutely essential for epimerase activity. Asp^{178} is conserved throughout all the *A. vinelandii* epimerases,

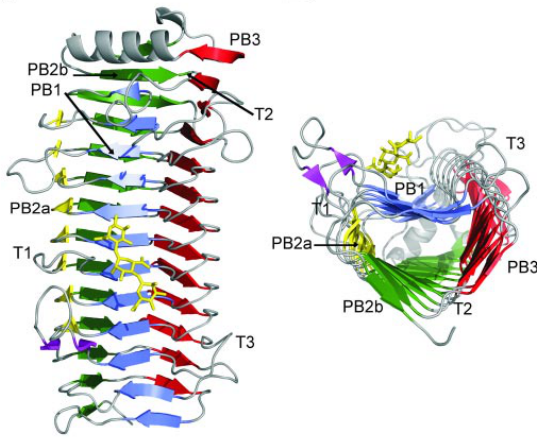
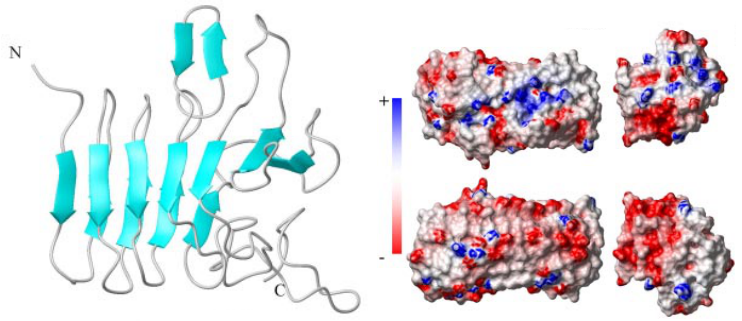


Figure 1.4: Structure of the A-module in AlgE4, as determined by Rozeboom et al. (2008), with a stick representation of a mannuronan trisaccharide bound in the active cleft. PB1 (blue), PB2a (yellow), PB3b (green), and PB3 (red), are the four β-sheets that make ut the β-helix. T1-T3 mark the turns between sheets. Right: front view of the active cleft, Left: view from the C-terminal end. Figure is from Rozeboom et al. (2008).

while the other three residues are conserved in all known mannuronan C-5 epimerases (Rozeboom et al., 2008). The 3D model of the A-module in AlgE4, by Rozeboom et al. (2008), is included in Figure 1.4.

The R-module of AlgE4, with a C-terminal S-motif, was studied by NMR and it was found that the N-terminal region is folded into a highly ordered right-handed parallel β-roll, while the C-terminal region forms a less ordered loop structure (Aachmann et al., 2006). The C-terminal end of the R-module (residues 145-167) has no ordered secondary structure. The NMR structure obtained by this study is included in Figure 1.5a, presented as a ribbon drawing. Overall, the R-module has an elongated structure with a positively charged patch of arginine and lysine residues along the surface of the small groove on the front side of the β-roll, and negatively charged patches of aspartic acid and glutamic acid residues located at the turns of the β-roll. The positively charged patch has been shown to interact strongly with an alginate M_5-pentamer (Aachmann et al., 2006), contrasting previous work where atomic force microscopy showed no direct interaction between the R-module and the alginate polymer (Sletmoen et al., 2004). The AlgE4 A-module also contains a positively charged patch on its surface, positioned on one side of the active site cleft, and the secondary structures of both modules suggests that connecting the C-terminus of



(a) **(b)**

Figure 1.5: (a) NMR structure of the R-module in AlgE4 presented as its secondary structure elements. (b) Electrostatic potential models for the A-module and R-module in AlgE4. Blue: positive potential, Red: negative potential. Both figures are from Aachmann et al. (2006).

the A-module to the N-terminus of the R-module could result in a combined elongated structure with a continuous positive patch for substrate binding (Aachmann et al., 2006). An electrostatic potential model from Aachmann et al. (2006) is included in Figure 1.5b.

All seven AlgE epimerases are dependent on Ca^{2+} and the direct binding of Ca^{2+} has been demonstrated for both modules (Ertesvåg et al., 1998; Ertesvåg and Valla, 1999). The R-module is not required for epimerase activity, but its presence increases the reaction rate and reduces the Ca^{2+} concentration needed for full activity (Ertesvåg and Valla, 1999). The R-module of AlgE4 contains four to seven copies of a nonameric repeat with consensus sequence LXGGAGXDX, a sequence similar to calcium binding motifs previously identified in functionally unrelated proteins in *P. aeruginosa* (Baumann et al., 1993). These nonameric motifs in the R-module are likely involved with calcium binding in a similar manner to that observed for the metalloprotease of *P. aeruginosa* (Aachmann et al., 2006). The crystal structure of the AlgE4 A-module revealed the presence of six Ca^{2+} binding sites (Rozeboom et al., 2008).

Patterns of Epimerization

The seven AlgE epimerases of *A. vinelandii* all have different substrate specificities, distinct patterns of epimerization, and different Ca^{2+} concentrations required for full activity (Ertesvåg et al., 1998; Holtan et al., 2006). Out of all seven AlgE epimerases, AlgE2 and AlgE5 have greatest sequential homology (see modular structure in Ta-

ble 1.1). They have both shown a preference for M residues located next to a G residue and their activity introduces G-blocks, with AlgE2 giving rise to shorter G-blocks when compared to AlgE5 (Ertesvåg et al., 1999; Ramstad et al., 1999). AlgE6 also introduces G-blocks, and has been found capable of producing longer stretches of G residues compared to the products of AlgE2 and AlgE5 (Svanem et al., 1999; Ertesvåg et al., 1999; Holtan et al., 2006). AlgE4 introduces strictly alternating sequences (Ertesvåg et al., 1995; Høidal et al., 1999), while AlgE1 and AlgE3, which both contain two A-modules, introduce MG-blocks as well as G-blocks (Ertesvåg et al., 1998, 1999). This bifunctionality of AlgE1 and AlgE3 can be explained by each of the two separate catalytic A-modules being responsible for one block type (Ertesvåg et al., 1998). AlgE7 stands out from the group by having a combined G-block forming mannuronan C-5 epimerase activity and lyase activity. Both reactions are probably catalysed at the same active site (Svanem et al., 2001). The epimerization patterns are summarised in Table 1.1. Of particular interest is comparing the large difference in the product formation by AlgE4 and AlgE6, producing MG-blocks and G-blocks respectively, alongside their very similar structures.

Reaction Mechanism

The mechanism of mannuronan C-5 epimerization has been proposed as a three step reaction (Gacesa, 1987) and this mechanism has been supported by more recent findings from an X-ray crystallography study of the A-module in AlgE4 (Rozeboom et al., 2008). The proposed mechanism shares the first two steps with the alginate lyases, and this is further described in Section 1.3. The three steps of the proposed reaction mechanism are (Gacesa, 1987):

1. Neutralization of the negative charge on the carboxyl anion.
2. Base-catalysed abstraction of the C-5 proton.
3. Protonation at C-5 on the opposite side of the sugar ring, resulting in C-5 epimerization.

In the first step, Tyr^{149} has been proposed to act as a general base, abstracting the C-5 proton. The side chain of the amino acid residue Arg^{195} has been identified to lie at hydrogen bonding distance to the Tyr^{149} hydroxyl group and could be acting as a proton relay path to the solvent, effectively removing the proton from the active site. The position of His^{154} suggests that it may be the general acid that donates a proton to the opposite side of C-5 in the final step of the epimerization mechanism (Rozeboom et al., 2008). A proposed role of the other two essential amino acid residues, Asp^{152} and Asp^{178}, is in ensuring the neutralization of the carboxyl group

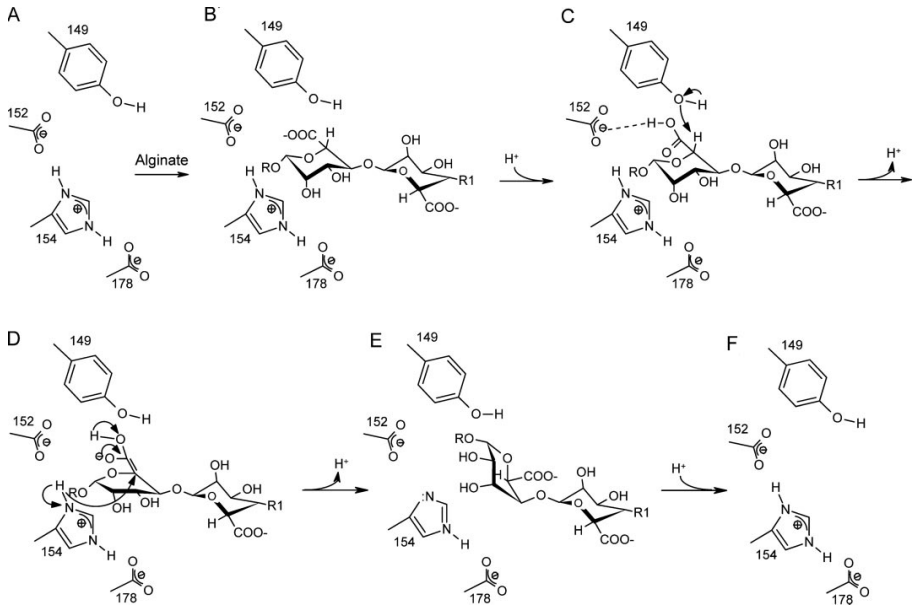


Figure 1.6: Reaction mechanism as proposed by Gacesa (1987), supported by findings in Rozeboom et al. (2008). Figure is from Rozeboom et al. (2008).

for substrate stabilization during the first step (Rozeboom et al., 2008). An overview of the reaction mechanism is included in Figure 1.6.

The C-5 epimerization of mannuronan is considered to be irreversible, as there are currently no reported observations of a decrease in G-content related to incubation with mannuronan C-5 epimerases.

Mode of Action

The process of sequence generation by the action of a mannuronan C-5 epimerase and how it results in different end-products, can be divided into four independent features of the epimerase (Hartmann, 2000):

1. Product specificity
2. Enzyme processivity
3. Enzyme reaction rate

4. Substrate selectivity

As previously mentioned (see Table 1.1), the product specificity of mannuronan C-5 epimerases divides them into three groups based on their product being enriched in G-blocks, MG-blocks, or both. The activity of an epimerase after introduction of the first G-residue can be epimerization of the neighbouring M-residue, epimerization of the second M-residue neighbour, or dissociation from the polymer chain. Each action resulting in different block types.

The processivity of an enzyme describes the average number of repeated reactions it performs between association and dissociation from the substrate. For an enzyme to be processive it must be capable of recognizing and binding to substrate residues around the catalytic site and not just in the catalytic site itself. All the AlgE epimerases of *A. vinelandii* possess this quality and have been found to accomodate several substrate residues in their binding sites. For mannuronan C-5 epimerases this parameter relates to block-elongation, dissociation from the polymer, or change of catalytic activity to formation of a different block type. A reflection of possible differences in processivity can be seen for AlgE2, AlgE5, and AlgE6 where they all have a product specificity of G-block formation, but of different lengths. When considering the processivity of the bifunctional AlgE1 and AlgE3, it is necessary to take into account the individual processivities of each catalytic A-module as well as the probability of a change in catalytic activity from one block type to the other. AlgE4 has been revealed to act by a processive mechanism, sliding along the polymer chain epimerizing every other M residue (Campa et al., 2004).

The reaction rate is a characteristic of each individual enzyme and when several different enzymes compete for one substrate it becomes an important factor for the end-product. This is the case for the AlgE epimerases of *A. vinelandii*, where several different epimerases are present (Ertesvåg et al., 1995; Svanem et al., 1999).

The substrate specificity of mannuronan C-5 epimerases refers to their affinity towards a specific monomer sequence in the polymer chain. The sterical arrangements of the two monomers, discussed in Section 1.1.1, are sufficiently different to provide substrate specificity based on monomer sequence. The mannuronan C-5 epimerases act by a multiple-attack mechanism (Larsen et al., 1986) and each time an epimerase dissociates from an alginate chain it is faced with a multitude of possible substrate sequences for its next attack. Also, as the epimerization proceeds, new sequences are continually formed, providing a new set of possible association sites. Previous studies of AlgE6 revealed that it prefers substrates already containing G-blocks, acting by elongation or condensation of previously formed G-blocks (Holtan et al., 2006). AlgE6 has a minimum substrate of seven monomers and it acts on the fourth residue, suggesting that it contains seven binding sites with catalytic activity

in the fourth site (Holtan et al., 2006). A hexameric oligomer has been revealed as the minimum substrate for AlgE4, and epimerization of the third residue suggested the model of six binding sites with catalytic activity in site number three (Campa et al., 2004).

1.3 Alginate Lyases

Alginate lyases catalyse the degradation of alginates and have been isolated both from alginate-synthesizing organisms and non-alginate-synthesizing organisms, including brown algae (Madgwick et al., 1973), marine invertebrates (Boyen et al., 1990), and a wide range of marine and soil microorganisms (Preiss and Ashwell, 1962a,b; Kennedy et al., 1992; Dunne and Buckmire, 1985). The production of lyases in organisms that do not produce alginate has been linked to the use of depolymerized alginate as a carbon source, and in alginate-synthesizing organisms they appear to play a role both in biosynthesis and degradation of alginate. Marine mussels utilize algal matter as a carbon source and have been found to produce alginate lyases as part of their carbohydrate-degrading enzyme mix (Seiderer et al., 1982). *P. aeruginosa*, associated with cystic fibrosis patients, produces a thick alginate-containing biofilm for better adhesion to surfaces, and the periplasmic lyase AlgL aids in cell detachment from this biofilm to allow colonization of new sites (Boyd and Chakrabarty, 1994). In the biosynthesis of alginates, AlgL has been found to have a role in degrading alginates that fail to be exported out of the cell, preventing accumulation of alginate in the periplasmic space (Bakkevig et al., 2005), and might also have an editing role in controlling chain length (May and Chakrabarty, 1994).

Alginate lyases have different substrate specificities and are classified as poly-G (EC 4.2.2.11) or poly-M (EC 4.2.2.3) specific. This classification refers to their dominant action as almost all classified alginate lyases cleave more than one of the four possible linkages (M-M, M-G, G-M, G-G). The only two known cases of specificity towards only one linkage type are AlxM (Chavagnat et al., 1998) and AlyA5 (Tøndervik et al., 2010) with strong preferences for M-M and G-G respectively. Endolytic cleavage, attacking internal sites of the polymer chain, is displayed by most lyases, but a few exolytic lyases, removing monomers or dimers from the end of the polymer chain, have been found (Brown and Preston III, 1991; Nakada and Sweeny, 1967).

Alginate lyases degrade alginates by cleaving the (1→4) O-glycisidic linkage through β-elimination, yielding oligosaccharide products containing an unsaturated monomer residue at the non-reducing end. A three step mechanism of action has been proposed (Gacesa, 1987, 1992), where the first two steps are shared with mannuronan C-5 epimerases (see Section 1.2.1) and involve neutralization of the negative charge

Table 1.2: Four lyases; AlyA (Østgaard et al., 1993), AlyA5 (Tøndervik et al., 2010), *H. tuberculata* lyase (Heyraud et al., 1996), and AlxM (Chavagnat et al., 1998), commonly used in research that each display one of the four possible substrate specificities. Aly A and the *H. tuberculata* lyase were used in this study.

Lyase (abbreviation)	**Specificity**	**Source**
AlyA (G-lyase)	G-G and M-G	*Klebsiella Pneumoniae*
AlyA5 (GG-lyase)	G-G	Mutant of AlyA
H. tuberculata lyase (M-lyase)	M-M and M-G	*Haliotis tuberculata*
AlxM (MM-lyase)	M-M	(*Photobacterium* sp.(ATCC 43367)

on the carboxyl anion followed by abstraction of the proton on C-5 resulting in a carbanion. The third and final step is β-elimination where a proton is donated to the O-4 leaving group and the glycosidic bond is cleaved with formation of a double bond between C-4 and C-5.

Most applications of alginate lyases lie in research instead of commercial fields. The cleavage action of lyases result in an unsaturated uronic acid residue at the non-reducing end and this residue will act as a chromophore giving rise to absorbance at 230 nm. Depending on the type of lyase or lyases applied, this feature provides a simple method for measuring alginate concentration or epimerase activity (Østgaard et al., 1993; Tøndervik et al., 2013). The study of block-structure in alginates also exploits the different specificities of lyases by degrading the polymers in ways that leave desired block structures intact before analysis by e.g. NMR or HPAEC-PAD (Tøndervik et al., 2010; Aarstad et al., 2011). An epimerase activity assay and G-block distribution analysis by HPAEC-PAD, both applying the action of lyases, are described in Section 3.2.12 and Section 3.2.14, respectively.

Chapter 2. Aim and Strategy

Alginates have a wide variety of applications and their prospects for the future are even greater. The physical properties of alginates are highly dependent on their chemical composition, and the enormous degree of sequential and compositional variation creates the need for structural control. By expressing mannuronan C-5 epimerases in their native, or even hybrid form, in alginate-producing bacteria, it could be possible to produce alginates with carefully designed properties and a high level of compositional homogeneity. This is especially promising for the strict demands of biomedical applications. Great efforts have been made to obtain detailed characterization of the C-5 epimerases and their interactions with alginates, yet much is still unclear. In particular, the seven extracellular AlgE epimerases of *Azotobacter Vinelandii* have been extensively studied, and this study aims at clarifying the role of the R-module, specifically in AlgE4, and to learn more about the structure function relationship in the A-module of AlgE6.

This study was performed as two separate parts, both a continuation of work by Dypås (2014). In the earlier work, the A-module in AlgE6 was found to be adequately active compared to the full length epimerase. It was also suggested that the R-module mostly has a role related to the epimerization pattern, and that this role is highly dependent on the calcium concentration. Due to the high degree of sequential homology between the A-modules in AlgE4 and AlgE6, the first part of this study compares full length AlgE4 to its A-module, as it would be interesting to see if AlgE4A behaves similar to AlgE6A.

The second part of this study is focused on the connection between structure and function within the A-module itself, by attempting to create a smaller epimerase with retained functionality. This was done by creating three AlgE6A clones with varying degree of truncation at their C-terminal end. Not only is this of interest in an effort to elucidate details of the mechanisms behind the epimerase activity in AlgE6, it is also highly relevant for industrial purposes. The possibility of producing a smaller epimerase with retained functionality could have financial impact due to lowered

production cost of the enzyme itself, and by changing the conditions needed for epimerization, as a smaller epimerase will have better access to substrate in a tight space.

Chapter 3. Materials and Methods

3.1 Materials

3.1.1 Alginate

Mannuronan (alginate poly-M) was isolated from an AlgG-negative mutant of *Pseudomonas fluorescens* (Gimmestad et al., 2003), produced at NTNU/SINTEF. Poly-M used in this study had a low intrinsic viscosity of approximately 800 mL/g. Partially acid hydrolyzed poly-M with a DP of ~70 was used in some experiments. DP refers to the degree of polymerization, i.e. the number of monomer residues in an oligomer chain.

3.1.2 Bacterial Strains, Plasmids, and Synthetic Genes

Bacterial Strains and Plasmids

All bacterial strains and plasmids, and their relevant characteristics, are presented in Table 3.1.

The T7 expression strain *E. coli* T7 Express (New England Biolabs Inc.) was used in the production of both epimerases. AlgE4 was expressed in T7 Express with the plasmid pTB182 (In house, un.pub.). AlgE4A was expressed in T7 Express with the plasmid pTB26 (Rozeboom et al., 2008).

Supercompetent *E. coli* DH5α (Invitrogen^{TM}) was used for gene amplification and transformation of ligation products. The T7 expression strain *E. coli* OverExpress c41(DE3) (Lucigen Corp.) was used for production of the modified epimerases with the plasmids pTYB1_272, pTYB1_295, and pTYB1_336, constructed in this study.

Three modified versions of the AlgE6 gene were used in this study, consisting of its first 272, 295, and 336 amino acid residues, starting at the N-terminal of the A-module, resulting in varying degrees of truncation at the A-module's C-terminal. Synthetic genes were ordered from GenScript Inc. (USA), and were delivered as 4 µg

plasmid preparations with each gene as an insert in the pUC57 vector. The synthetic genes were inserted into the plasmid pTYB1 (New England Biolabs Inc.) for expression. More information about the pUC57 constructs and the pTYB1 plasmid can be found in Appendix A and Appendix B, respectively.

3.1.3 Enzymes

Restriction Endonucleases and DNA Ligase

Enzymes used in restriction digestion and ligation:

- NdeI, 20000 U/mL (New England Biolabs Inc.)
- XhoI, 20000 U/mL (New England Biolabs Inc.)
- T4 DNA Ligase, 400000 U/mL (New England Biolabs Inc.)

Epimerase Standards and Lyases

Enzymes for the microassay in Section 3.2.12 were provided by SINTEF:

- AlgE4 standard, freeze dried
- AlgE6 standard, freeze dried
- The G-lyase AlyA from *Klebsiella pneumoniae*

Table 3.1: Bacterial strains and plasmids, and their relevant characteristics for use in this study.

Strain or plasmid	Relevant Characteristics	Ref.
E. coli		
DH5α	Cloning strain with high transformation efficiency	Invitrogen^{TM}
T7 Express	BL21(DE3) derivative, encodes T7 RNA polymerase under control of the *lac*UV5 promoter, deficient in proteases Lon and OmpT	New England BioLabs Inc.
OverExpress c41(DE3)	BL21(DE3) derivative, encodes T7 RNA polymerase under control of the *lac*UV5 promoter, deficient in proteases Lon and OmpT	Lucigen Corp.
Plasmids		
pTYB1	IMPACT^{TM} C-terminal fusion vector with a self-cleavable intein and a chitin binding tag, see Appendix B	New England BioLabs Inc.
pTYB4	IMPACT^{TM} C-terminal fusion vector with a self-cleavable intein and a chitin binding tag (differs from pTYB1 only in sites present in the polylinker)	New England BioLabs Inc.
pTB26	Derivative of pTYB4 with insertion of the sequence encoding the first 378 amino acids of AlgE4	Rozeboom et al. (2008)
pTB182	Derivative of pHH4 (Høidal et al., 1999) where the NcoI-BgIII fragment encoding AlgE4 A-R, the A-module without the S-sequence, is inserted in pTYB4	In house, unpublished
pUC57	High copy plasmid for cloning, see Appendix A	GenScript USA Inc.
pTYB1_272	Derivative of pTYB1 with insertion of the sequence encoding the first 272 amino acids of AlgE4	This study
pTYB1_295	Derivative of pTYB1 with insertion of the sequence encoding the first 295 amino acids of AlgE4	This study
pTYB1_336	Derivative of pTYB1 with insertion of the sequence encoding the first 336 amino acids of AlgE4	This study
pUC57_272	Derivative of pUC57 with insertion of the sequence encoding the first 272 amino acids of AlgE4	This study
pUC57_295	Derivative of pUC57 with insertion of the sequence encoding the first 295 amino acids of AlgE4	This study
pUC57_336	Derivative of pUC57 with insertion of the sequence encoding the first 336 amino acids of AlgE4	This study

3.1.4 Growth Media, Buffers, and Other Solutions/Chemicals

0.8 % Agarose gel with GelGreen (400 mL):
3.2 g Agarose (Lonza, USA)
400 mL 1xTAE
Boiled in microwave before addition of 20 µL GelGreen

Ampicillin (Amp):
100 mg/mL Ampicillin (AppliChem, Germany) stock solution made with deionized water
Sterile filtered
Stored as aliquots of 1 mL at −20 °C
Used in media as 100 µg/mL

Cleavage Buffer (IMPACT^{TM}):
20 mM HEPES (AppliChem, Germany)
5 mM $CaCl_2$*$2H_2O$ (VWR International, Norway)
800 mM NaCl (VWR International, Norway)
Dissolved in deionized water, pH adjusted to 6.9 and filtered using a Stericup filter unit with a 0.22 µm cut-off
50 mM 1,4-Dithiothreitol (DTT) (VWR International)
Made fresh and kept on ice while loading on column

6x DNA Loading Buffer for Agarose Gel Electrophoresis:
0.25 % w/v Bromphenol Blue (Sigma-Aldrich, Germany)
0.25 % w/v Xylene Cyanol FF (Sigma-Aldrich, Germany)
40 % w/v Sucrose (AppliChem, Germany)
Dissolved in deionized water and stored as aliquots of 1 mL at −20 °C

Column Buffer (IMPACT^{TM}):
20 mM HEPES
5 mM $CaCl_2$*$2H_2O$
800 mM NaCl
Dissolved in deionized water, pH adjusted to 6.9 and filtered using a Stericu filter unit with a 0.22 µm cut-off

Deuterium Oxide:
D_2O (C/D/N Isotopes, Chiron AS)

Enzyme Buffer C:
20 mM MOPS (Research Organics/Sigma-Aldrich, Sweden)
3.64 mM $CaCl_2$*$2H_2O$
100 mM NaCl
0.2 mg/mL Spectinomycin (Sigma-Aldrich)
Dissolved in milli-Q water

4x Epimerization Buffer:
200 mM MOPS
10 mM $CaCl_2$*$2H_2O$
300 mM NaCl
Dissolved in milli-Q water and pH adjusted to 6.9 with NaOH

Isopropyl β-D-1-thiogalactopyranoside (IPTG):
0.5 M IPTG stock solution made with deionized water
Sterile filtered
Stored as aliquots of 1 mL at $-20\,^{\circ}C$

Ligation Buffer:
10x (10 mM ATP) T4 DNA Ligase Buffer (New England Biolabs Inc.)

Lysis Buffer (IMPACT^{TM}):
20 mM HEPES 5 mM $CaCl_2$*$2H_2O$
500 mM NaCl
0.05 % $^{v}/_{v}$ Triton X-100 (Sigma-Aldrich)
Dissolved in deionized water, pH adjusted to 6.9 and filtered using a Stericup filter unit with a 0.22 µm cut-off

Lysogeny Agar (LA):
10 g/L Bacto Tryptone (Oxoid, UK)
5 g/L Yeast Extract (Oxoid, UK)
10 g/L NaCl
15 g/L Agar (Oxoid, UK)
Dissolved in deionized water and autoclaved

Lysogeny Broth (LB):
10 g/L Bacto Tryptone
5 g/L Yeast Extract
10 g/L NaCl

Dissolved in deionized water and autoclaved

3xLB:
30 g/L Bacto Tryptone
15 g/L Yeast Extract
10 g/L NaCl
Dissolved in deionized water and autoclaved

Restriction Digestion Buffer:
CutSmart™ (New England Biolabs Inc.)

3x SDS Loading Buffer:
150 mM Tris (Sigma-Aldrich)
377 mg/mL Glycerol (VWR International, Belgium)
60 mg/mL SDS (VWR International, Belgium)
3 mg/mL Bromophenol Blue
46.5 mg/mL 1,4-Dithiothreitol (DTT)

SDS Running Buffer:
20x ClearPAGE™ SDS-reducing Running Buffer (C.B.S. Scientific) diluted to 1x with milli-Q water

SDS Staining Solution:
Instant Blue™ (Expedeon)

SDS Protein Standards:
Precision Plus Protein™ Dual Color standard (Bio-Rad)
Amersham Low Molecular Weight Calibration Kit Protein Mixture Standard (GE Healthcare)
(see Appendix C)

Spectinomycin (Spc):
200 mg/mL Spectinomycin stock solution made with deionized water
Sterile filtered
Stored as aliquots of 1 mL at −20 °C
Used in media as 200 µg/mL

Substrate Solution A:
1.33 mg/mL poly-M (NTNU/SINTEF)

20 mM MOPS
3.64 mM $CaCl_2$*$2H_2O$
100 mM NaCl
Dissolved in deionized water

Super Optimal Broth with Catabolite Repression (SOC):
2 % w/v Tryptone
0.5 % w/v Yeast Extract
8.56 mM NaCl
2.5 mM KCl (Merck, Germany)
10 mM $MgCl_2$*$6H_2O$ (Merck, Germany)
20 mM Glucose, sterile filtered (VWR International, Belgium)
Dissolved in deionized water and autoclaved before addition of glucose
Stored as aliquots of 1 mL at −20 °C

50xTris-Acetate-EDTA (TAE) buffer (1000 mL):
242 g Tris-base (Tris(hydroxymethyl)aminomethane) (VWR International, Belgium)
57.1 mL 100 % Acetid acid
100 mL 0.5 M EDTA, pH 8 (VWR International, Belgium)
Deionized water added to a total volume of 1000 mL and autoclaved

Triethylenetetraminehexaacetic acid (TTHA):
TTHA (Sigma-Aldrich, Germany)

3-(Trimethylsilyl)propionic-2,2,3,3-d_4 acid (TSP):
TSP (Sigma-Aldrich, USA)

3.2 Methods

3.2.1 The IMPACT™ System for Protein Expression and Purification

The IMPACT™ (Intein Mediated Purification with an Affinity Chitin-binding Tag) system was used in this study as it facilitates chromatographic purification of free recombinant proteins directly from crude cell lysates in one single step (Chong et al., 1997; IMPACT™ Manual, 2014). A schematic illustration from the IMPACT™ manual is included in Figure 3.1. The system is based on fusing the protein of interest to a modified protein splicing element, the *Sce* VMA intein from *Saccharomyces cervisiae*, in combination with a chitin binding domain (CBD) from *Bacillus circu-*

lans as an affinity tag. On a chitin resin column, the CBD will adsorb and bind the target-intein-CBD complex strongly to the resin while all unwanted proteins are flushed out. Addition of DTT, b-mercaptoethanol (b-ME), or cysteine, will then induce the *Sce* VMA intein to undergo self-cleavage at the N-terminal peptide linkage, releasing the protein of interest to be eluted as a pure product. Induction of cleavage by DTT, b-ME, or cysteine, has been found effective at low temperatures and within a wide pH range, facilitating purification of the target protein under mild conditions (Chong et al., 1997). 50 mM DTT, as used in this study, is more than enough to reduce all disulfide bridges in a protein and thus might not always classify as mild conditions. However, due to the absence of disulfide bridges in the epimerases, the target proteins in this study are not sensitive to DTT.

The T7 Expression System

The T7 expression system is used with the IMPACT^{TM} system for inducible protein production in *E. coli*. T7 RNA polymerase binds to a distinct promoter sequence that is not recognized by *E. coli* RNA polymerase and cloning the target gene such that it is under the control of the T7 promoter will prevent expression during propagation in *E. coli* (Reece, 2004). To make expression of the target gene inducible, a copy of the T7 RNA polymerase gene combined with the *lac* promoter is also introduced.

Several *E. coli* strains carrying a chromosomal copy of T7 RNA polymerase under the control of a *lac* promoter have been developed, including T7 Express and OverExpress c41(DE3), as used in this study. Both strains are derivatives of *E. coli* BL21(DE3) which is deficient in the two proteases Lon and OmpT, and where T7 RNA polymerase is under the control of the *lac*UV5 promoter. These strains are used as hosts for expression of a target gene cloned into a vector whose expression is controlled by the T7 RNA polymerase promoter.

3.2.2 Modification of the AlgE6 Gene

To identify interesting sites for modification of the A-module in AlgE6, the A-module in AlgE4, whose structure has been determined by Rozeboom et al. (2008), was used as a model. The 3D model of AlgE4A was downloaded from the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB code: 2PYG). PyMOL^{TM} (DeLano Scientific LLC) was used to visualize and examine tertiary protein structure alongside its amino acid sequence. A proposed 3D-model of AlgE6A from its amino acid sequence was created using SWISS-MODEL (ExPASy) with AlgE4A as a template. SWISS-MODEL is a system based on protein structure homology modelling, that is, it generates structural models based on the amino acid sequence

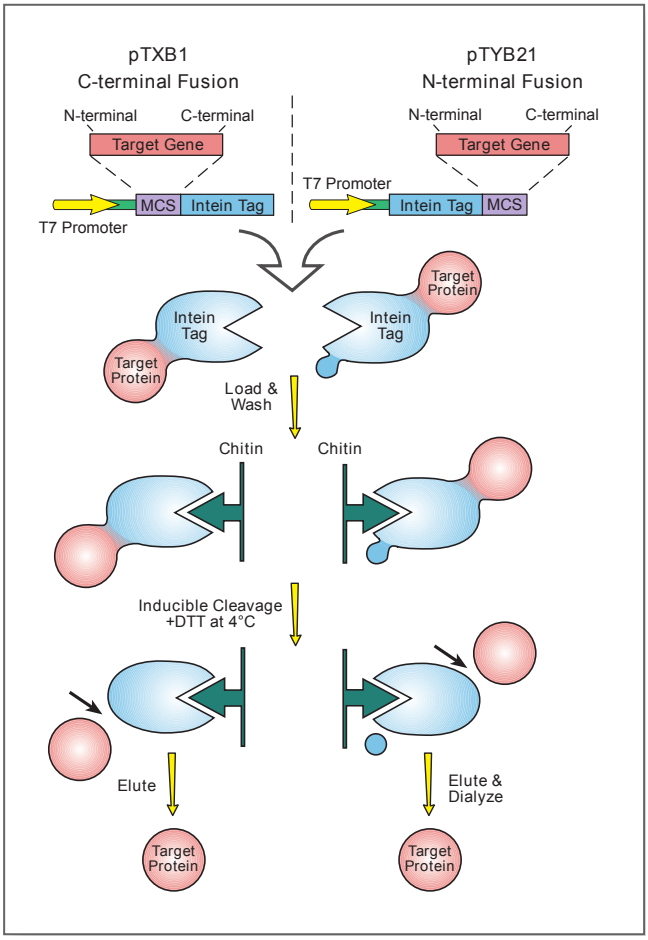


Figure 3.1: Illustration of the IMPACT^{TM} system for expression and purification of free recombinant proteins (IMPACT^{TM} Manual, 2014).

of a protein as compared with experimentally determined structures of other, evolutionarily related, proteins (Arnold et al., 2006; Biasini et al., 2014; Guex et al., 2009; Kiefer et al., 2009).

Synthetic genes were ordered from GenScript Inc. (USA), and were delivered as 4 µg plasmid preparations of the cloning vector pUC57 (see Appendix A for more details). Plasmid preparations were dissolved in 15 µL sterile deionized water, and 1 µL dissolved DNA was used in transformations.

3.2.3 Chemical Transformation

Chemical transformation is a common method to transform *E. coli* with plasmid DNA (Green and Sambrook, 2012). The first step in this process involves incubation of the bacterial cells in a solution containing cations (Aune and Aachmann, 2010). It is thought that the positive ions help mask the negative charges on the naked DNA, thereby condensing the DNA (Aune and Aachmann, 2010), as well as facilitating contact with the cell membrane, as electrostatic repulsion is cancelled out (Weston et al., 1981). Cells treated this way are said to be chemically competent, as they have been made able to take up extracellular DNA. Competent cells in this study have been treated with a solution containing RbCl. The cells are also incubated on ice, forcing the membrane to take on a more regularly packed and stiff structure, which is thought to cause small openings in the membrane (Hanahan, 1983). After this preparation of the cells, they are exposed to a heat-shock that is believed to transiently increase cell wall permeability and allow for uptake of DNA (Cosloy and Oishi, 1973).

All transformations were performed by thawing competent cells (0.1 mL cells in 1.5 mL eppendorf tubes) on ice before adding DNA. After addition of DNA, cells were incubated on ice for 5-30 minutes, heat shocked at 42 °C for 45 seconds, and cooled on ice for 2 minutes. 900 µL prewarmed SOC was added and cells were incubated at 37 °C for 1 hour. 100 µL of transformed cells was spread on LA plates (with 100 µg/mL Amp) and incubated at 37 °C overnight. Cells were plated out in three different concentrations; the reaction concentration, 10x diluted, and concentrated, by pelleting remaining cells at 12 000 g for 5 minutes and removing excess supernatant over 100 µL. Amount of DNA added and incubation time on ice varied depending on bacterial strain and level of difficulty for each transformation.

3.2.4 Amplification and Isolation of Plasmid

DNA from successful transformations was amplified by inoculating 5 mL LB medium (with 100 µg/mL Amp) with one colony, transferred from LA plates using a sterile

toothpick, and incubating at 37 °C overnight. 1.5 mL of culture was transferred to eppendorf tubes, and plasmid DNA was isolated using the Wizard® *Plus* SV Minipreps DNA Purification System (Promega, USA). The miniprep protocol was followed with the addition of one extra washing cycle.

The Wizard® miniprep procedure to isolate plasmid DNA from bacterial cell culture is based on the alkaline lysis method, first developed by Birnboim and Doly (1979), and purification by silica resin (Promega, 2012). Lysis under alkaline conditions will cause the, now exposed, DNA to denature. Neutralization of the solution will allow for the small circular DNA of plasmids to renature, while the longer chromosomal DNA stays denatured and can be precipitated from solution (Green and Sambrook, 2012). High concentrations of chaotropic salts cause DNA to adsorb onto silica (Chen and Thomas Jr, 1980; Marko et al., 1982; Boom et al., 1990), and a column packed with silica resin is therefore used to further purify the plasmid DNA.

If a higher concentration of isolated DNA was needed, this was achieved using the DNA Clean & Concentrator-5 Kit (Zymo Research, Inc.), eluting DNA with 20 µL DNA elution buffer.

3.2.5 Restriction Endonuclease Digestion

Type II restriction endonucleases recognize and cleave DNA, from any source, at specific nucleotide sequences (Green and Sambrook, 2012), and are therefore essential to recombinant DNA technology. A common use for restriction enzymes is to isolate a DNA fragment of interest, e.g. a gene, and inserting it into a suitable vector for cloning. This procedure is often aided by an important feature of many restriction enzymes, namely, making staggered cuts. Staggered cuts in the double-stranded DNA results in complementary single stranded ends on the DNA segments, also called "sticky ends". If the target vector and the DNA containing the fragment to be isolated are both digested with the same restriction enzymes, the ends of the fragment will anneal, or "stick", to the ends of the vector through complementary base pairing. DNA ligase will then be able to covalently join them together by creating a phosphodiester bond (Snustad and Simmons, 2012), as further discussed in Section 3.2.7.

The restriction endonucleases NdeI and XhoI were used, and a simultaneous double digest reaction was set up using CutSmart Buffer (New England Biolabs). Restriction digests were incubated at 37 °C for 1-2 hours. Further information about reaction volume and component ratios are given for each digestion in Table 3.2, presenting restriction digestions to excise target genes from pUC57 and open the expression vector pTYB1, and Table 3.3, presenting restriction digestions to control ligation products.

Table 3.2: Set up of restriction digestions to excise target genes from pUC57, and to open pTYB1 for use as expression vector. pUC57_X: pUC57_272, pUC57_295, or pUC57_336. All volumes are given in µL.

	Restriction Digest - Insert	**Restriction Digest - Vector**
DNA		
-Insert: pUC57_X	6	
-Vector: pTYB1		12
CutSmart buffer	2	2
Sterile deionized water	10	4
NdeI	1	1
XhoI	1	1
Total reaction volume:	20	20

Table 3.3: Set up of restriction digestions to verify successful insertion of *algE6A_272*, *algE6A_295*, and *algE6A_336*, in the expression vector pTYB1. pTYB1_X: pTYB1_272, pTYB1_295, or pTYB1_336. All volumes are give in µL. Digest 1: control after transformation with ligation products, Digest 2: control of ligation product before second attempt at transforming *E. coli* OverExpress c41(DE3) with AlgE6A_272 and AlgE6A_295.

Component	**Digest 1**	**Digest 2**
DNA (pTYB1_X)	20	2
CutSmart buffer	4	1
Sterile deionized water	14	5.5
NdeI	1	0.75
XhoI	1	0.75
Total reaction volume:	40	10

3.2.6 Agarose Gel Electrophoresis

Agarose dissolved in buffer forms a matrix through hydrogen bonding and hydrophobic interactions, creating pores in the gel, and the diameter of the pores depends on the concentration of agarose (Green and Sambrook, 2012). Agarose gel electrophoresis is a method for separating DNA fragments by size, and is usually the method of preference for fragments within the size range 200–20 000 bp (Smith, 1996). The charge per unit length of DNA is a constant, and in an electric field, DNA of any length will therefore migrate towards the positive electrode at the same rate if friction is assumed negligible. Increasing friction by running the DNA fragments through a gel provides separation by size, as small fragments experience less friction compared to larger fragments, and longer DNA fragments also have greater trouble "worming" through the pores without getting tangled up. The migration rate for lin-

ear DNA is inversely proportional to the $\log_{10}$ of the number of base pairs (Helling et al., 1974). Circular DNA, i.e. plasmids, do not follow this relationship as their apparent size in the gel matrix is smaller. After separation on gel, the fragments are visualized under ultraviolet light (UV) by staining with an intercalating fluorescent dye, such as GelGreen which is used in this study.

0.8 % agarose with 20 µL GelGreen was cased in a gel tray with gel comb in place, and left to solidify for 30 minutes before adding 1x TAE. 6x loading buffer was added to restriction digestion reactions before they were loaded on wells. GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific) was used as reference and gels were run at 85–100 V, depending on gel size. DNA fragments were visualized using the ChemiDoc™ XRS+ system with Image Lab™ 4.1 (Bio-Rad).

DNA fragments isolated on gel for use in ligation reactions were extracted with the Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Inc.). The kit protocol was followed with the addition of one extra washing cycle, and DNA was eluted from the column with 30 µL DNA elution buffer.

3.2.7 Ligation

As mentioned in Section 3.2.5, the complementary ends of DNA fragments from a restriction digestion will associate and form hydrogen bonds, but DNA ligase is needed to covalently join the molecules together. The evolutionary role of DNA ligase is to close nicks in double-stranded DNA during DNA replication, as well as repair damaged DNA, by forming a phosphodiester bond (Reece, 2004). When exploiting this process to create recombinant DNA, the energy required to form a phosphodiester bond must be provided as nicotinamide adenine dinucleotide (NAD) or adenosine triphosphate (ATP).

Ligation of gene inserts and the opened pTYB1 vector was performed using different ratios of insert to vector due to unknown concentration of DNA in samples. A control without insert was included to check for re-ligation of the vector. The ligation reaction was set up as presented in Table 3.4, using T4 DNA ligase (New England Biolabs Inc.) and 10x T4 DNA ligase buffer (New England Biolabs Inc.). Samples were incubated overnight at 16 °C.

3.2.8 Production of Epimerases

Control of Protein Expression in AlgE6A Clones

A small control production was performed to verify IPTG inducible expression of target proteins in the AlgE6A clones. 5 mL LB medium (with 200 µg/mL Amp) was inoculated with one colony, transferred with a sterile toothpick from an LA plate

Table 3.4: Ligation reaction to insert target gene in the expression vector pTYB1. Three different ligation reactions were set up with varying insert to vector ratios. A religation control without insert was also included. All volumes are in µL. Insert refers to either *algE6A_272*, *algE6A_295*, or *algE6A_336*, while vector refers to pTYB1.

	Ligation 1	**Ligation 2**	**Ligation 3**	**Control**
Vector	3	3	3	3
Insert	15	5	3	0
T4 DNA Ligase Buffer	2.5	2	2	2
T4 DNA Ligase	1.5	1	1	1
Sterile deionized water	3	9	11	14
Total reaction volume:	25	20	20	20

with the final transformants, and incubated overnight at 37 °C as pre-culture. 20 mL LB medium (with 200 µg/mL Amp) was inoculated with 1 %(v/v) pre-culture and incubated at 37 °C for three hours. Protein production was induced by adding IPTG to a final concentration of 1 mM, and the induced culture was incubated at 37 °C for another three hours. A parallel control culture was also produced, without induction by IPTG.

The cells were harvested by centrifugation at 5,000 *g* on a Sorvall centrifuge for 10 minutes at 4 °C. The supernatant was discarded and the cell pellet was resuspended in 3 mL lysis buffer and lysed by sonication on ice for 5 minutes, using a Branson Sonifier 250 with a 3 mm double stepped microtip, set to output control 4.5 and duty cycle 40 %. 1 mL lysed cells were transferred to an eppendorf tube and centrifuged at 12 000 g for 10 minutes. The supernatant was analysed by SDS-PAGE, as described in Section 3.2.10.

Main Culture for Production of Epimerases

For production of AlgE4 and AlgE4A, 25 mL LB medium (with 200 µg/mL Amp) was inoculated with cells from frozen glycerol stock, and incubated overnight at 37 °C as pre-culture.

For production of AlgE6A_272, AlgE6A_295, and AlgE6A_336, 25 mL LB medium (with 200 µg/mL Amp) was inoculated with one colony, transferred with a sterile toothpick from an LA plate with the final transformants, and incubated overnight at 37 °C as pre-culture.

For all epimerases, 500 mL 3xLB medium (with 200 µg/mL Amp) was inoculated with 1 %(v/v) pre-culture in a 2 L baffled flask and incubated at 37 °C. OD_{600nm} was

monitored on a SpectraMax Plus^{384} absorbance microplate reader (Molecular Devices), and at OD_{600nm} around 0.8-1.0, IPTG was added to a final concentration of 1 mM to induce expression of the target protein. The induced culture was incubated overnight at 16 °C.

The cells were harvested by centrifugation at 5,500 *g* on a Sorvall centrifuge for 5 minutes at 4 °C. The supernatant was discarded and the pellets were stored at −20 °C for further processing to isolate and purify the expressed proteins.

3.2.9 Protein Isolation and Purification

Protein Isolation

The cell pellet was resuspended in 15–20 mL lysis buffer with half a tablet of cOmplete EDTA-free Protease Inhibitor Cocktail (Roche). Resuspended cells were lysed by sonication on ice for 10 minutes, using a Branson Sonifier 250 with a 13 mm tip step horn with a threaded body, set to output control 4.5 and duty cycle 40 %. The lysed cells were centrifuged for 45 minutes at 23,400 *g* and 4 °C. Supernatant was filtered using a Stericup filter unit with a 0.22 µm cut-off.

Protein Purification by FPLC

Fast protein liquid chromatography (FPLC) was first developed by Pharmacia LKB in Sweden in 1982 (Richey, 1982), as an automated high-resolution system for fast and biocompatible purification of proteins. The full range of chromatography modes can be accommodated by the FPLC system, including ion exchange, affinity, and reverse phase (Madadlou et al., 2011).

Proteins in this study were expressed in the IMPACT^{TM} system, as described further in Section 3.2.1, and purified by affinity chromatography performed on an ÄKTAFPLC system (GE Healthcare). The system includes an in-line UV absorption monitor that was used to continuously measure protein concentration in the effluent by the absorption at 280 nm.

A column was prepared with 20 mL Chitin Resin (New England Biolabs), following the product protocol (IMPACT^{TM} Manual, 2014). The filtered lysate was loaded on the column and unwanted proteins were washed out with column buffer until the A_{280nm} reached a stable baseline (>200 mL). Column was then flushed with 60 mL cleavage buffer and left overnight for on-column cleavage. The target protein was eluted with up to 60 mL column buffer in fractions of 10 mL. Column was regenerated by stripping the chitin resin with 0.3 M NaOH and flushing with water and column buffer (or ethanol for storage of column), as described in the product protocol.

3.2.10 SDS-PAGE for Protein Identification and Evaluation of Sample Purity

The migration of proteins during gel electrophoresis is affected by both shape and charge in addition to size, allowing some proteins with large differences in molecular weight to migrate together. Separating polypeptides based on molecular weight alone can be achieved by polyacrylamide gel electrophoresis (PAGE) in the presence of sodium dodecyl sulfate (SDS) (Lodish, 2013). The anionic detergent SDS binds along the backbones of polypeptides, disrupts non-covalent interactions and, in combination with a reducing agent that breaks disulfide bonds, causes the protein to unfold. The amount of SDS interacting with a protein is relatively proportional to chain length (~1.5 SDS molecules per amino acid residue), and the net negative charge it introduces will therefore confer a similar charge to mass ratio on all proteins, leaving molecular weight as the primary determinant of migration rate during gel electrophoresis (Lodish, 2013). Polyacrylamide is the preferred gel for electrophoretic separation of proteins. Compared to agarose gels they have smaller pores and higher resolving power, and their pore size can also be adjusted to best resolve the protein of interest.

12 % ClearPAGE^{TM} SDS precast gels with 12 or 17 wells (C.B.S. Scientific) were used. Samples were prepared in eppendorf tubes with 3x SDS loading buffer and boiled at 95 °C for 5 minutes before they were loaded on the gels. Amersham Low Molecular Weight Calibration Kit Protein Mixture standard (GE Healthcare) and Precision Plus Protein^{TM} Dual Color standard (Bio-Rad) were used as reference (see Appendix C for details). The gels were immersed in SDS running buffer and wells were washed before loading. Electrophoresis was run at 130 V for 85-100 minutes. Gels were stained with Instant Blue^{TM} (Expedeon, Inc.) for one hour on a shaker, and washed in water overnight. Images of gels were obtained using the ChemiDoc^{TM} XRS+ system with Image Lab^{TM} 4.1 (Bio-Rad)

3.2.11 Estimation of Protein Concentration

NanoDrop

NanoDrop spectrophotometers utilize the surface tension properties of liquids to hold microvolume samples between two optical pedestals during measurements. The required sample volume of 1 to 2 µL provides a method for efficient estimation of protein concentration when sample volume is limited. This technology also eliminates the need for cuvettes or capillaries, and removing these contaminants from the process allows for real time change of path length during measurements, permitting a wider range of protein concentrations to be measured without dilution (Desjardins et al., 2009a,b).

The Beer-Lambert Law (Equation (3.1)) states that there is a linear relationship between the absorbance, A, and the concentration, c (mol/L), of a sample (Grimsley and Pace, 2004).

$$A = \varepsilon l c \tag{3.1}$$

ε is the molar extinction coefficient ($M^{-1}cm^{-1}$), an intrinsic property defining how strongly a species absorbs light at a given wavelength, and l is the light path length (cm). The optical absorbance of proteins is commonly measured at 280 nm (A_{280nm}), a method based on the chromophore effect of the amino acid residues tryptophan, tyrosine, phenylalanine, and cysteine (Grimsley and Pace, 2004). NanoDrop uses a standard setting of ε 0.1 % where a solution of 1 mg/mL protein gives rise to an absorbance of 1.0 at 280 nm, with a path length of 1 cm (NanoDrop, 2008). This corresponds to a molar extinction coefficient of 10,000 $M^{-1}cm^{-1}$. The molar extinction coefficient depends on the composition of a protein, as it is related to the absorbance of selected amino acid residues. All the proteins produced in this study have values for ε two to four times higher than the standard setting for NanoDrop, and measured concentrations were therefore normalized relative to their respective coefficients.

SDS-PAGE

The protein concentration was also estimated by relative quantification of protein bands from SDS-PAGE. This method is based on calculating the densities of protein bands on a gel, relative to the density of a chosen reference band. By using a band of known protein quantity as reference, such as a recombinant protein from the standard, concentration can be estimated for a protein of interest. Imaging and analysis of gels was performed using the ChemiDoc™ XRS+ system with the Band Analysis tool in Image Lab™ 4.1 (Bio-Rad).

3.2.12 Epimerase Activity Microassay

Measurements of epimerase activity were performed based on the action of G-Lyase on epimerized poly-M. As described in Section 1.3, cleavage of a glycosidic linkage in alginate by a lyase results in an unsaturated residue chromophore that gives rise to UV absorbance at 230 nm. The G-lyase AlyA of *K. pneumoniae* has MG- and GG-specificity (see Table 1.2) and will therefore introduce one unsaturated residue per G residue, reflecting the degree of epimerization.

The assay was set up on a standard 96-well microplate. 50 µL enzyme buffer C was added to all wells, and different concentrations of enzyme were achieved by adding 50 µL of sample to one column and making 2-fold dilutions from column to

column. 150 µL substrate solution A was added to all wells, and the plate was left for incubation at 37 °C overnight. Absorbance at 230 nm was measured using a SpectraMax Plus^{384} absorbance microplate reader (Molecular Devices). 10 µL G-lyase was added to each well, and the plate was incubated at 25 °C for 4 hours. Absorbance was again measured at 230 nm and epimerase activity was estimated from the difference in absorbance. Enzyme standards AlgE4 or AlgE6 were included in the microassays for reference.

3.2.13 ^{1}H-NMR Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a method of molecular analysis that exploits the magnetic behaviour of atomic nuclei. ^{1}H- and ^{13}C-NMR is commonly used as a method of analysing composition and sequential parameters in alginates (Grasdalen et al., 1979, 1981; Grasdalen, 1983). Due to the sequential heterogeneity of alginates it is necessary to consider block type and block length in addition to the monomer ratio. The parameters required in this study are the average fractions of M and G (F_M and F_G), also called monads which describe the monomer content, while sequence information is described by the frequencies of the four diads (F_{MM}, F_{MG}, F_{GM}, and F_{GG}) and the four G-centred triads (F_{GGG}, F_{GGM}, F_{MGG}, and F_{MGM}). These parameters can then be used to calculate the average G-block length, $N_{G>1}$, from Equation (3.2).

$$N_{G>1} = \frac{F_G - F_{MGM}}{F_{GGG}} \tag{3.2}$$

This information is obtainable by ^{1}H-NMR, whereas ^{13}C-NMR is needed for frequencies of the M-centered triads and calculation of average M-Block length.

Depolymerization by Acid Hydrolysis

Alginates of high molecular weight were partially depolymerized by a two-step acid hydrolysis to lower the viscosity prior to NMR analysis. Acid hydrolysis is performed in two steps because an abrupt lowering of the pH to 3.8, close to the pK_a of alginate, would cause alginate in the sample to precipitate as discussed in Section 1.1.2. The pH of epimerized samples was adjusted to 5.6 and samples were incubated in a water bath at 95 °C for one hour. Samples were cooled down to room temperature using a water bath, pH was adjusted to 3.8, and samples were incubated in a water bath at 95 °C for 50 minutes. Samples were cooled down to room temperature again on water bath and pH was adjusted to 6.9 before samples were freeze-dried.

^{1}H-NMR Sample Preparation and Analysis

Freeze-dried sample was weighed out, transferred to an NMR tube, and dissolved in 600 µL D_2O. 20 µL TTHA (0.3 M, pH 7) and 5 µL TSP (1 %) was also added. D_2O is a common polar solvet used in ^{1}H-NMR to avoid a high background signal from the solvent. The chelator TTHA was added to remove divalent ions, and TSP was used as internal reference. ^{1}H-NMR spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C. Measurements and data analysis was performed by Wenche Iren Strand (Department of Biotechnology, NTNU).

3.2.14 Block Distribution Analysis

The distribution of G-blocks, as well as M-blocks and MG-blocks, in alginates can be studied by combining enzymatic hydrolysis with High-Performance Anion-Exchange Chromatography and Pulsed Amperometric Detection (HPAEC-PAD)(Aarstad et al., 2011). Application of specific alginate degrading enzymes, i.e. lyases, makes it possible to leave one block type intact while degrading the polymer, and the resulting oligosaccharides will therefore hold information about the block length distribution. The mechanisms and specificities of the different lyases, which this method is based on, are described in Section 1.3. HPAEC-PAD has previously been shown to be an effective tool in the analysis of oligosaccharides derived from alginates (Ballance et al., 2005; Campa et al., 2004).

Epimerized samples were degraded with MM-lyase (see Table 1.2), and subsequently analysed by HPAEC-PAD. This procedure was performed by Olav Andreas Aarstad (Department of Biotechnology, NTNU), following the method previously described in Aarstad et al. (2011).

3.2.15 Epimerization of Alginate for Analysis by ^{1}H-NMR or HPAEC-PAD

10–20 mg poly-M was dissolved in MQ-water before adding 4x epimerization buffer to a 1x concentration in the total reaction volume. Water and buffer volume varied depending on the total reaction volume, and is specified for each epimerization in Chapter 4. Amount of enzyme was calculated to get approximately the same molarity for each enzyme, and this estimation was based on the concentration ratios found from SDS-PAGE, or by using freeze-dried enzyme. After addition of enzyme, samples were incubated at 37 °C on a rotator plate for 1-3 nights.

EDTA was added to a final concentration of 4mM to stop the epimerization reaction. Samples were transferred to dialysis tubes (Spectra/Por Dialysis Membrane, MWCO: 3.5 kDa or 12–14 kDa) and dialysed three times against 7 L 50 mM NaCl and four times against 7 L MQ-water.

One epimerization was performed to test calcium dependence. In this case, 4x epimerization buffer without calcium was used, and different concentrations of Ca^{2+} was instead added to each sample.

A second epimerization tested temperature stability, and these samples were incubated at 40 °C, 50 °C, and 60 °C, instead of the optimal 37 °C.

Chapter 4. Results

This study is a continuation of the work performed by Dypås (2014), and it consists of two separate parts. Part one is a comparison of the full length AlgE4 and its A-module AlgE4A, and part two is a study of three truncated versions of the A-module from AlgE6.

4.1 Comparison of Full Length AlgE4 and its A-module AlgE4A

4.1.1 Production and Purification of AlgE4 and AlgE4A

The epimerases AlgE4 and AlgE4A were produced and purified as described in Sections 3.2.8 and 3.2.9, using bacterial strains and plasmids listed in Section 3.1.2. A chromatogram from purification of AlgE4 by FPLC is shown in Figure 4.1. Purification of AlgE4A produced a close to identical chromatogram, however, this was not recorded. The chromatogram presents UV absorption at 280 nm ($A_{280\text{nm}}$) as a function of volume (mL) eluted from the column, and gives an indication of protein concentration in the effluent. Fractions A1-A5 are indicated in red along the x-axis and the protein was mainly eluted in fractions A1 and A2, as measured $A_{280\text{nm}}$ approaches a baseline value after fraction A2. The chromatogram also suggests that fraction A1 had a higher concentration of protein compared to fraction A2.

SDS-PAGE to evaluate fractions from FPLC was performed as described in Section 3.2.10 and the result is presented in Figure 4.2. Molecular weights of AlgE4 and AlgE4A, obtained using ProtParam (ExPASy, Gasteiger et al. (2005)), are 55 707 kDa and 39 798 kDa, respectively. Strong protein bands were present in fractions A1 and A2 for AlgE4 in Figure 4.2a and for AlgE4A in Figure 4.2b. These bands corresponded to molecular weights that were slightly higher than the actual size of the epimerases. There was also a weak band present in fraction A3, and possibly fraction A4, while nothing appeared in fraction A5. Unwanted protein bands, indicative of sample contamination, were absent in all fractions.

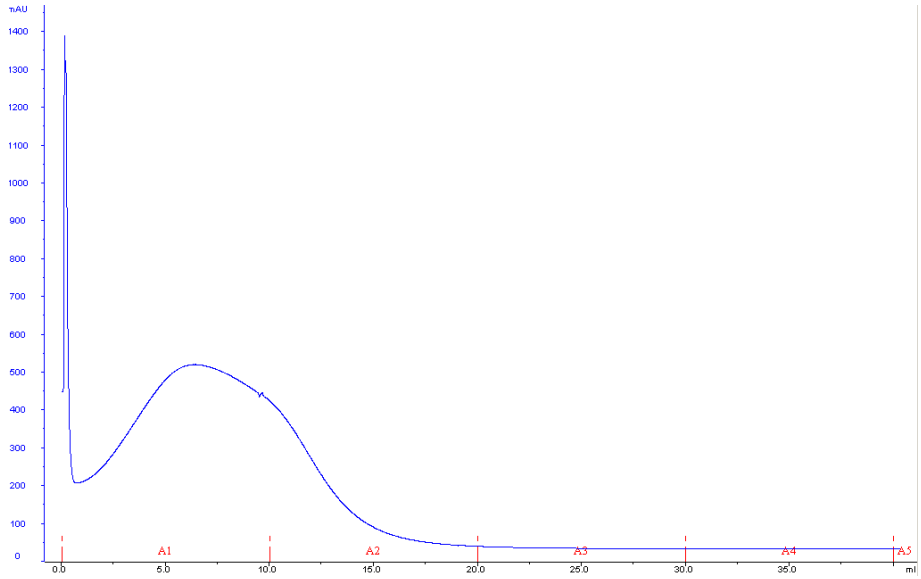


Figure 4.1: Chromatogram from purification of AlgE4 by FPLC on a chitin resin column. UV absorption, $A_{280\text{nm}}$, as an indication of protein concentration in the effluent, is given as a function of volume eluted from the column, in mL. Fractions A1-A5 are indicated in red.

4.1.2 Estimation of Protein Concentration

NanoDrop

Protein concentration of FPLC fractions AlgE4 A1 and A2, and AlgE4A A1 and A2, was estimated by measuring $A_{280\text{nm}}$ on a NanoDrop^{TM} ND-1000 UV-Vis spectrophotometer (see Section 3.2.11). Molar extinction coefficients of 38 280 M^{-1} cm^{-1} for AlgE4 and 33 810 M^{-1} cm^{-1} for AlgE4A were obtained using ProtParam (ExPASy, Gasteiger et al. (2005)). The protein concentrations obtained were normalized relative to their respective molar extinction coefficients, and with respect to the standard, the absorbance from 1 mg/mL will be approximately 3.8 times higher for AlgE4 and 3.4 times higher for AlgE4A. Estimated protein concentrations are presented in Table 4.1. Fraction A1 had the highest concentration for AlgE4. Fractions A1 and A2 for AlgE4A were estimated to have the same prtoein concentration, and this was higher than that estimated for AlgE4 A1.

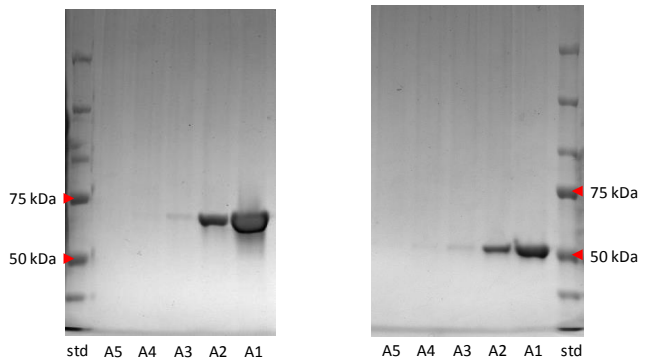


Figure 4.2: Gel images from SDS-PAGE to evaluate fractions from purification of (a) AlgE4 and (b) AlgE4A by FPLC on a chitin resin column. 15 µL FPLC-fraction was prepared with 5 µL 3x SDS loading buffer, and 20 µL was loaded on each well. A1-A5: Fractions A1-A5 from FPLC purification of AlgE4 on gel (a) and AlgE4A on gel (b), std: 7 µL Precision Plus Protein^{TM} Dual Color standard (Bio-Rad). Side labels identify molecular weights, in kDa, of specific proteins in the standard.

SDS-PAGE

Protein concentration in fractions AlgE4 A1 and AlgE4A A1 was also estimated by relative quantification of protein bands on gel from SDS-PAGE, as described in Section 3.2.11, and the results are presented in Table 4.2. Gel image, density analysis from ImageLab, and the necessary calculations, can be found in Appendix D. In strong contrast to estimations from NanoDrop, SDS-PAGE resulted in twice as high concentration for AlgE4 A1 compared to AlgE4A A1.

4.1.3 Epimerase Activity Microassay

Epimerase activity of fractions A1 and A2 for both AlgE4 and AlgE4A was tested in a microassay based on the activity of G-lyase as described in Section 3.2.12. The difference in absorbance at 230 nm ($\Delta A_{230\text{nm}}$) before and after incubation with G-lyase was used to estimate epimerase activity, and the result is presented in Table 4.3. The activity is presented in relative units (RU) per mL, where RU is defined as the amount of enzyme needed to catalyse the conversion of 1 µmol substrate per minute. Estimations of protein concentration from NanoDrop and SDS-PAGE, presented in Section 4.1.2, were used to calculate the activity in RU per mole.

Fraction A1 had highest activity for both AlgE4 and AlgE4A, and was selected for

Table 4.1: Protein concentrations for FPLC fractions AlgE4 A1 and A2, and AlgE4A A1 and A2, estimated by NanoDrop. Concentrations are given by mass (mg/mL), normalized relative to their molar extinction coefficients ε (mg/mL), and as molarity (M) after normalizing to ε.

	Mass [mg/mL]	**Normalized to ε** [mg/mL]	**Molar** [M]
AlgE4 A1	1.40	0.366	6.57×10^{-6}
AlgE4 A2	0.33	0.086	1.54×10^{-6}
AlgE4A A1	1.99	0.589	1.48×10^{-5}
AlgE4A A2	2.01	0.594	1.49×10^{-5}

Table 4.2: Protein concentrations estimated by relative quantification of protein bands from SDS-PAGE. Concentration was estimated for FPLC fractions AlgE4 A1 and AlgE4A A1. Concentration is given by mass [mg/mL] and molarity [M].

	Mass [mg/mL]	**Molar** [M]
AlgE4 A1	1.009	1.81×10^{-5}
AlgE4A A1	0.477	1.20×10^{-5}

use in further study. The mole-based activity calculated from NanoDrop concentrations indicated higher activity for AlgE4, while concentrations estimated from SDS-PAGE resulted in almost identical activity of 1.26×10^{9} and 1.30×10^{9} RU/mole for AlgE4 A1 and AlgE4A A1, respectively.

Table 4.3: Estimations of epimerase activity obtained from a microassay based on the activity of G-lyase on epimerized poly-M. G-lyase activity was calculated from measurements of ΔA_{230}. Activity based on protein concentration estimated from NanoDrop is presented first, followed by activity based on protein concentration estimated from SDS-PAGE. Concentration is given in mol/mL, and activity is given in RU/mL and RU/mol.

Nanodrop

	Concentration [mol/mL]	**Activity** [RU/mL]	**Activity** [RU/mol]
AlgE4 A1	6.57×10^{-9}	22.89	3.49×10^{9}
AlgE4 A2	1.55×10^{-9}	5.07	3.28×10^{9}
AlgE4A A1	1.48×10^{-8}	15.61	1.06×10^{9}
AlgE4A A2	1.49×10^{-8}	0	0

SDS-PAGE

	Concentration [mol/mL]	**Activity** [RU/mL]	**Activity** [RU/mol]
AlgE4 A1	1.81×10^{-8}	22.89	1.26×10^{9}
AlgE4A A1	1.20×10^{-8}	15.61	1.30×10^{9}

4.1.4 Calcium Dependence Activity Assay for Analysis by ^{1}H-NMR and HPAEC-PAD

All seven of the extracellular mannuronan C-5 epimerases of *A. vinelandii* are dependent on Ca^{2+} ions, and their direct binding of Ca^{2+} ions has previously been identified (Ertesvåg et al., 1998; Ertesvåg and Valla, 1999). Multiple Ca^{2+} binding sites have also been identified in the R-modules of AlgE epimerases and could be of importance to the mechanism of epimerization and mode of action (Aachmann et al., 2006). Calcium dependence in AlgE4 and AlgE4A was therefore compared to study what effect the R-module has on epimerase activity and epimerization pattern in relation to the calcium concentration. Poly-M was epimerized with AlgE4 and AlgE4A at varying Ca^{2+} concentrations in the range 1–8 mM, and the product was analysed by ^{1}H-NMR, and selected samples were also analysed by HPAEC-PAD. Procedures for epimerization, sample preparation, ^{1}H-NMR, and HPAEC-PAD, are described in Sections 3.2.13 to 3.2.15. Each sample was prepared with 20 mg poly-M in a total reaction volume of 8 mL. 100 µL AlgE4 or AlgE4A was added to each sample before they were incubated at 37 °C for three nights.

^{1}H-NMR spectra of the epimerized poly-M revealed that all samples had been

fully epimerized and a second calcium dependence assay in the range 0.1–8 mM was therefore performed with less enzyme and incubation just overnight. Freeze dried AlgE4 and AlgE4A was used and added to an epimerase to alginate ratio of approximately 1:400.

^{1}H-NMR

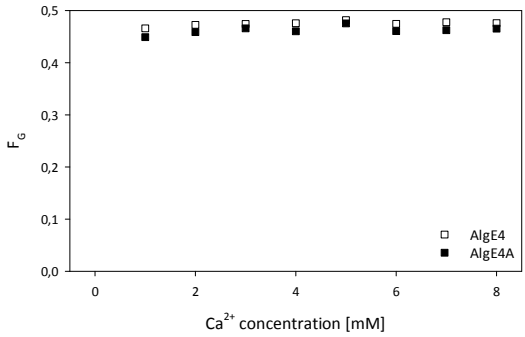
F_G for the epimerized poly-M was calculated from ^{1}H-NMR spectra (spectra included in Appendix E) and is presented in Figure 4.3 as a function of Ca^{2+} concentration. The first assay resulted in all samples being fully epimerized, as seen in Figure 4.3a, with no significant difference when comparing activity by AlgE4 to AlgE4A. F_G was slightly smaller from epimerization with AlgE4A compared to AlgE4, but the difference is minimal. In the second assay however, with less enzyme and shorter incubation time, there was a clear difference between the product of AlgE4 and the product of AlgE4A, presented in Figure 4.3b. AlgE4 and AlgE4A both had little to no activity at 0.1 mM Ca^{2+} (< 1 % G), but the activity of AlgE4 climbed quickly, reaching peak activity at 2 mM Ca^{2+} and maintaining this level of activity throughout the rest of the concentration range (6 mM Ca^{2+}, due to a bad recording for the last sample point). The activity of AlgE4A climbed steadily up to the 3.5 mM sample point, after which it started to level out before reaching peak activity only around 6–8 mM Ca^{2+}. G-content introduced by AlgE4A was consistently lower than that introduced by AlgE4, and there was no sign of G-blocks in the spectra.

HPAEC-PAD

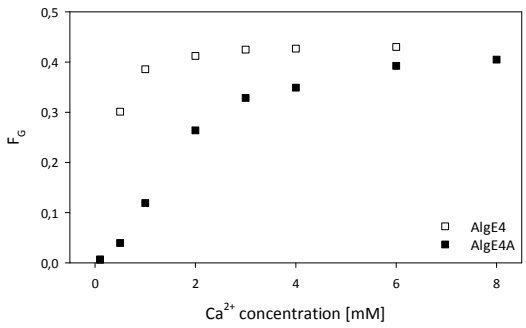
^{1}H-NMR from the second epimerization showed that poly-M epimerized by AlgE4 at 0.5 mM Ca^{2+} and by AlgE4A at 3 mM Ca^{2+}, had very similar G-content, making them good candidates for comparison of block structure. Samples of poly-M epimerized by AlgE4 and AlgE4A at 0.5 and 3 mM Ca^{2+} was analysed by HPAEC-PAD as described in Section 3.2.14, and the result is presented in Figure 4.4 as superimposed chromatograms. The epimerized poly-M was degraded with MM-lyase before analysis to obtain intact MG-block oligomers. Peaks were assigned by including two standards as reference, poly-M and poly-MG, both partially degraded by M-lyase, respectively resulting in M-block and MG-block oligomers of various lengths.

The chromatograms of poly-M epimerized by AlgE4 at 0.5 mM Ca^{2+} and by AlgE4A at 3 mM Ca^{2+} show that the general pattern of epimerization was somewhat similar, with introduction of long MG-blocks, which is the characteristic epimerization pattern for AlgE4. However, the introduced MG-blocks were shorter for AlgE4A compared to AlgE4. This can be seen from the last peak in the chromatogram being

sharper and taller for AlgE4A, indicating a larger portion of the G-content giving rise to shorter MG-blocks, and by the decline of this peak reaching baseline faster than for AlgE4, indicating the presence of longer MG-blocks in the AlgE4 product.



(a)



(b)

Figure 4.3: F_G calculated from ^{1}H-NMR spectra of poly-M epimerized by AlgE4 and AlgE4A at different calcium concentrations in the ranges (a) 1–8 mM and (b) 0.1–8 mM. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.

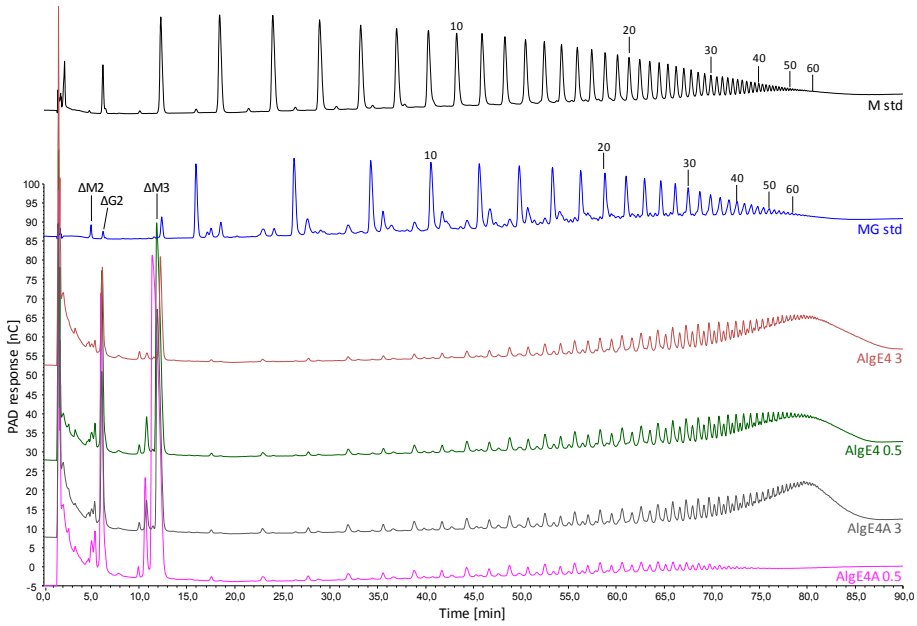


Figure 4.4: Superimposed HPAEC-PAD chromatograms from analysis of poly-M after epimerization with AlgE4 and AlgE4A in the presence of 0.5 and 3 mM Ca^{2+}. The epimerized poly-M was degraded with MM-lyase to obtain intact MG-block oligomers. Standards included in the chromatogram are poly-M (M std) and poly-MG (MG std), both degraded with M-lyase. Peak labels refer to the assigned identity of oligosaccharides responsible for the peak, with numbers referring to block length. Chromatograms were obtained from separation of oligomers on a Dionex IonPac AS4A column with a linear gradient of 8.75 mM sodiumacetate/min in 100 mM sodiumhydroxide, at a flow rate of 1 mL/min (Aarstad et al., 2011).

4.1.5 Temperature Stability Activity Assay for Analysis by ^{1}H-NMR

The activity of enzymes is highly dependent on temperature, as temperature affects the rate of all reactions and also the structural integrity of proteins. Higher temperatures leads to greater kinetic energy and increases the probability of productive interaction between substrate and enzyme, but at the same time greater kinetic energy could destabilize the bonds and forces responsible for maintaining native protein structure (Hardin et al., 2012). Increasing the temperature beyond a certain point will therefore cause the enzyme to denature and loose its activity.

This dependence on temperature is naturally closely linked to protein structure, and it is therefore interesting to see what effect the R-module has on maintaining

epimerase activity in the A-module of AlgE4 as temperature increases. The optimal temperature for epimerization of alginate with AlgE4 is close to 37 °C, and the activity is known to be maintained up to temperatures as high as 50 °C with a strong decrease closer to 60 °C (Høidal et al., 1999). Epimerization of poly-M with AlgE4 and AlgE4A was performed at 40 °C, 50 °C, and 60 °C, and the product was analysed by ^{1}H-NMR. Procedures for epimerization, sample preparation, and ^{1}H-NMR, are described in Sections 3.2.13 and 3.2.15. Each sample was prepared with 10 mg partially hydrolysed poly-M (DP~70) and a total reaction volume of 4 mL. Freeze dried AlgE4 and AlgE4A was used, and added to an epimerase to alginate ratio of approximately 1:200 before samples were incubated at their respective temperatures overnight. Acid hydrolysis before ^{1}H-NMR was not needed due to the already low DP.

F_G for the epimerized poly-M was calculated from ^{1}H-NMR spectra (spectra are included in Appendix F) and is presented in Figure 4.5 as a function of temperature. Epimerase activity was significantly lower for AlgE4A at all temperatures. There was little difference in activity from 40 °C to 50 °C, with a slight increase for AlgE4 and a slight decrease for AlgE4A. At 60 °C, no sign of activity could be observed for AlgE4A, while AlgE4 still exhibited low level activity.

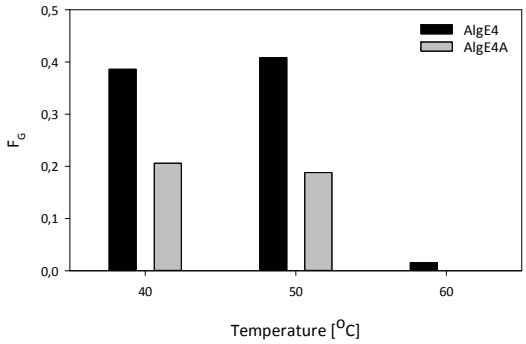


Figure 4.5: F_G calculated from ^{1}H-NMR spectra of poly-M epimerized by AlgE4 and AlgE4A at 40 °C, 50 °C, and 60 °C. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.

4.2 Truncated Versions of the A-module from AlgE6

4.2.1 Modification of the AlgE6A Gene

Tertiary structure of the A-module from AlgE4, determined by Rozeboom et al. (2008), was visualized in PyMOL^{TM} (DeLano Scientific LLC) and is presented in Figure 4.6. The figure includes the four amino acids, Tyr^{149}, Asp^{152}, His^{154}, and Asp^{178}, that have been identified as absolutely essential for epimerase activity and are suggested to have a role in catalysis, as discussed in Section 1.2.1. The N-terminal region of AlgE4A consists of a more complex and unique tertiary structure compared to the C-terminal region which consists of similar repetitive sections. It was therefore decided to shorten the peptide by starting at the C-terminal, leaving the other end intact. Interesting sites for modification were identified by focusing on areas without ordered secondary structure and then looking at the amino acid residues in these areas to avoid creating a hydrophobe C-terminal. In addition, it was considered that the A-module is a right-handed β-helix with 12 complete turns that form four parallel β-sheets. By choosing Ser^{272}, Gln^{295}, and Ser^{336}, as C-terminal residues, all three modified proteins would complete a full turn of the β-helix before the cut-site. This was suggested to have less of a destabilizing effect on tertiary structure, compared to creating a C-terminal in the middle of a turn, due to each turn adding one more β-strand to each of the four β-sheets.

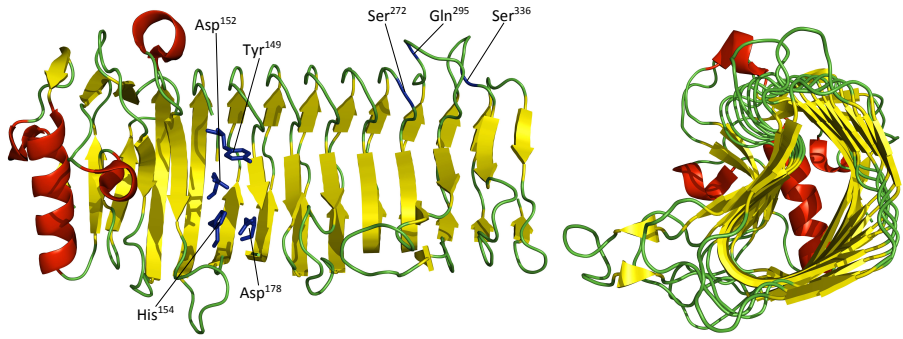


Figure 4.6: Tertiary structure of the A-module in AlgE4, determined by Rozeboom et al. (2008), visualized in PyMOL^{TM} (DeLano Scientific LLC). (a) Secondary structure elements making up the β-helix, with the α-helix cap at the N-terminal. Amino acid residues Tyr^{149}, Asp^{152}, His^{154}, and Asp^{178}, are included in blue stick representation. Ser^{272}, Gln^{295}, and Ser^{336}, are sites for modification. (b) Same as (a), seen from the C-terminal end. Red: α-helices, Yellow: β-strands.

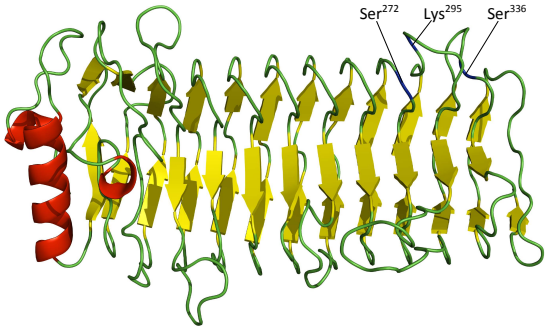
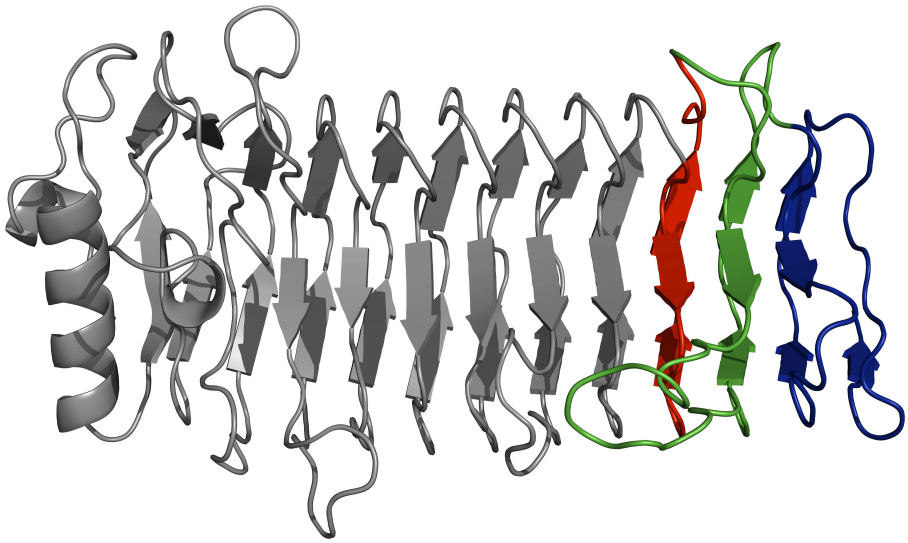


Figure 4.7: Tertiary structure of the A-module in AlgE6, created using SWISS-MODEL (ExPASy) with AlgE4A as a template, visualized in PyMOL^{TM} (DeLano Scientific LLC) as its secondary structure elements. Amino acid residues Ser^{272}, Lys^{295}, and Ser^{336}, indicate sites for modification. Red: α-helices, Yellow: β-strands.

A proposed 3D-model of AlgE6A was created from its amino acid sequence, using SWISS-MODEL (ExPASy) with AlgE4A as a template. This model was visualized in PyMOL^{TM} (DeLano Scientific LLC) and is included in Figure 4.7. Amino acid residues corresponding to the proposed sites for modification in AlgE4A were identified as Ser^{272}, Lys^{295}, and Ser^{336}. From this it was decided to produce and test three modified proteins, named AlgE6A_336, AlgE6A_295, and AlgE6A_272, by removing 49, 90, and 114 amino acid residues, respectively, from the C-terminal end of the A-module. Synthetic genes were ordered from GenScript (USA), as plasmid constructs of the cloning vector pUC57. Further details about the plasmid constructs are included in Appendix A. To better visualize the changes that were made to AlgE6A, Figure 4.8 presents tertiary structure and amino acid sequence colour coded to represent how much of the peptide was removed in each case.



(a)

N-terminal

MDYNVKDFGALGDGVSDDRVAIQAAIDAAHAAGGGTVYLPPGEYRVSAAGEPSDGCLTLR

DNVYLAGAGMGQTVIKLVDGSAQKITGIVRSPFGEETSNFGMRDLTLDGNRANTVDKVDG

WFNGYAPGQPGADRNVTIERVEVREMSGYGFDPHEQTINLVLRDSVAHHNGLDGFVADYQ

IGGTFENNVAYANDRHGFNIVTSTNDFVMRNNVAYGNGGNGLVVQRGSENLAHPENILID

GGSYYDNGLEGVLVKMSNNVTVQNADIHGNGSSGVRVYGAQGVQILGNQIHDNAKTAVAP

EVLLQSYDDTLGVSGNYYTTLNTRVEGNTITGSANSTYGVQERNDGTDFSSLVGNTINGV

QEAAHLYGPNSTVSGTVSAPPQGTD

C-terminal

(b)

Figure 4.8: (a) Tertiary structure for the A-module in AlgE6 colour coded to modifications. Consecutive removal of each coloured part results in AlgE6A_336 (blue), AlgE6A_295 (green), and AlgE6A_272 (red). (b) Amino acid sequence of AlgE6A with colour codes corresponding to those in (a).

4.2.2 Amplification and Isolation of Insert and Vector

Amplification of Target Genes and Production of the Expression Vector, pTYB1

E. coli DH5α was transformed with plasmids pUC57_272, pUC57_295, and pUC57_336, following the protocol in Section 3.2.3. Incubation time on ice after addition of DNA was only 5 minutes, due to this being a relatively easy transformation of supercompetent DH5α with concentrated plasmid DNA. Transformation resulted in almost overgrown plates, and these successful transformants were used to amplify and isolate the plasmid DNA by following the protocol in Section 3.2.4.

The plasmid pTYB1, to be used as expression vector, was produced and isolated by following the protocol in Section 3.2.4, with the exception that the LB medium was inoculated with cells from frozen glycerol stock, not colonies from LA plates.

Restriction Digestion and Separation of DNA Fragments by Agarose Gel Electrophoresis

To excise the target genes from pUC57 and insert them in the expression vector pTYB1, the isolated DNA was digested with the restriction endonucleases NdeI and XhoI, as described in Section 3.2.5. This would open the pTYB1 vector and excise target genes from pUC57, while at the same time creating complementary "sticky" ends between vector and target genes. Simulations of the restriction digestions in Clone Manager Professional version 6.00 (Sci-Ed Software) were used to obtain lengths of the resulting DNA fragments, presented in Table 4.4

To identify and isolate the excised target genes and the opened vector, DNA fragments in the restriction digests were separated by agarose gel electrophoresis as described in Section 3.2.6. Undigested pTYB1 was included as a control. An image of the visualized DNA fragments after separation on agarose gel is presented in Figure 4.9.

All restriction digestions to excise target genes from pUC57 gave rise to three DNA bands each. One band appeared in line with the 250 bp marker and another was in line with the 2500 bp marker, corresponding to the expected DNA fragments of 249 bp and 2468 bp, respectively. The third band from each digestion corresponded with the expected lengths for their respective target genes, *algE6A_272* (818 bp), *algE6A_295* (887 bp), and *algE6A_336* (1010 bp), and these were isolated from the gel as described in Section 3.2.6. The restriction digestion to open pTYB1 gave rise to a single band close to the 8000 bp marker, which corresponds to the opened vector having a length of 7439 bp, and this DNA was therefore isolated from the gel. The second DNA fragment from this digestion can not be seen on the gel, but given its small size of only 38 bp, it was not expected to give rise to a band. Undigested

Table 4.4: Expected lengths, in bp, of DNA fragments resulting from digestion of pTYB1, pUC57_272, pUC57_295, and pUC57_336, with NdeI and XhoI. DNA fragments of interest are labelled as "opened vector" or "gene". The restriction digestions were simulated in Clone Manager Professional version 6.00 (Sci-Ed Software) to obtain lengths of the resulting DNA fragments.

Plasmid	Fragment	Length [bp]
pTYB1	1	38
	2 (opened vector)	7439
pUC57_272	1	249
	2 (gene)	818
	3	2468
pUC57_295	1	249
	2 (gene)	887
	3	2468
pUC57_336	1	249
	2 (gene)	1010
	3	2468

pTYB1 gave rise to a band close to the 3500 bp marker, and further confirms that the band around 8000 bp is the opened vector, as undigested plasmids have a smaller apparent size on agarose gel and should migrate at a faster rate.

4.2.3 Cloning of Target Genes Into the Expression Vector pTYB1

Ligation and Transformation

Ligation was performed as described in Section 3.2.7 to insert *algE6A_272*, *algE6A_295*, and *algE6A_336*, into their expression vector pTYB1. The DNA isolated from restriction digestions in Section 4.2.2 was used.

E. coli DH5α was transformed with the ligation products following the protocol in Section 3.2.3, and the entire volume from each ligation was added to the competent cells. Incubation time on ice after addition of DNA was 5 minutes due to the relative ease of transforming supercompetent DH5α. After incubation, the colonies were counted and the result is presented in Table 4.5 along with a calculation of the ratio between transformants from religation and transformants from vector + insert (religation percentage). LA plates from ligation mix 1 and 2, plated out at the same cell concentration as the transformation reactions, gave the best result, and only

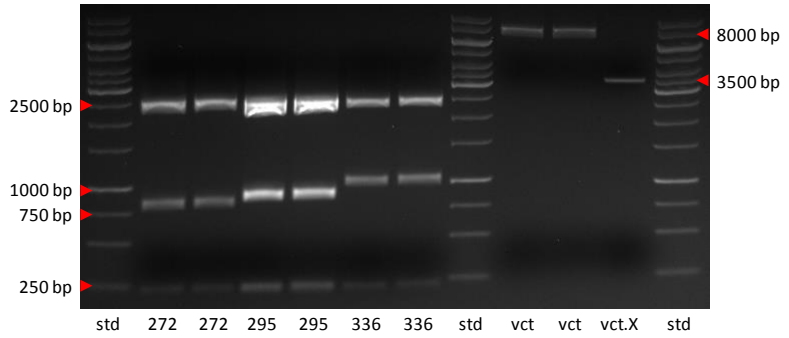


Figure 4.9: Visualized DNA fragments on agarose gel, from restriction endonuclease digestions to isolate target genes and opened vector. std: 0.75 µL GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific), 272: restriction digest of pUC57_272, 295: restriction digest of pUC57_295, 336: restriction digest of pUC57_336, vct: restriction digest of pTYB1, vct_X: undigested pTYB1. Side labels identify the lengths, in bp, of specific fragments in the DNA ladder.

these were counted. The LA plates with high cell concentration were overgrown and there were no colonies on the religation plate with a 10x dilution of cells, so counting these plates for estimation of religation percentage was redundant. The background from religation was estimated to ~2 % for all plates and clearly indicates insertion of the target genes.

Table 4.5: Cell count from transformation of *E. coli* DH5α after ligation of insert (*algE6A_272*, *algE6A_295*, or *algE6A_336*) and vector (pTYB1). The ratio between transformants from religation and transformants from vector + insert is presented as religation percentage.

	Cell count	**Religation percentage**
pTYB1, religation	8	N/A
pTYB1 + *algE6A_272* ligation 1	424	1.9 %
pTYB1 + *algE6A_272* ligation 2	288	2.8 %
pTYB1 + *algE6A_295* ligation 1	592	1.4 %
pTYB1 + *algE6A_295* ligation 2	373	2.1 %
pTYB1 + *algE6A_336* ligation 1	414	1.9 %
pTYB1 + *algE6A_336* ligation 2	331	2.4 %

Restriction Digestion to Control Ligation Product

A restriction digestion was performed to verify ligation of target gene and vector. Plasmid DNA from transformants was amplified and isolated by following the protocol in Section 3.2.4. The isolated plasmids were digested with the restriction endonucleases NdeI and XhoI by following the protocol in Section 3.2.5. DNA fragments from the restriction digests were separated by agarose gel electrophoresis as described in Section 3.2.6. An image of the visualized DNA fragments after separation on agarose gel is presented in Figure 4.10.

All three restriction digestions gave rise to a DNA band corresponding to their respective gene inserts, *algE6A_272* (818 bp), *algE6A_295* (887 bp), and *algE6A_336* (1010 bp). Digestions of pTYB1_272 and pTYB1_336 also resulted in a clear band close to the 8000 bp marker, which is where the opened vector of length 7439 bp was expected to appear. This band is missing from the digestion of pTYB1_295, and a thin smear could instead be observed below the 8000 bp marker (not visible in the gel picture, but was possible to observe in Image Lab^{TM}). This was suspected to stem from star activity of the restriction enzymes, or other contaminations in the restriction digestion, such as nucleases, and the ligation was evaluated as successful.

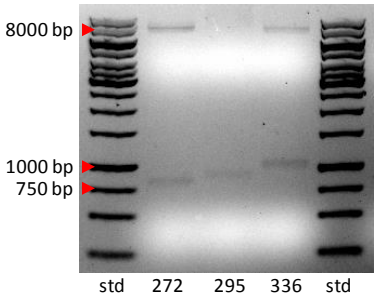


Figure 4.10: Visualized DNA fragments on agarose gel, from restriction digestion with NdeI and XhoI to verify successful insertion of *algE6A_272*, *algE6A_295*, and *algE6A_336*, in the expression vector pTYB1. std: 0.75 µL GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific), 272: restriction digest of pTYB1_272, 295: restriction digest of pTYB1_295, 336: restriction digest of pTYB1_336. Side labels identify the lengths, in bp, of specific fragments in the DNA ladder.

4.2.4 Transformation of the Expression Strain

E. coli OverExpress c41(DE3) was transformed with the plasmids, pTYB1_272, pTYB1_295, and pTYB1_336, (cloned in Section 4.2.3) for expression of the target genes. The transformation was performed as described in Section 3.2.3. Incubation time on ice

after addition of DNA was 30 minutes, due to lower competence of the OverExpress c41(DE3) cells compared to DH5α. Transformed cells were not plated out in a 10x dilution.

Cells transformed with pTYB1_272 resulted in five colonies, pTYB1_295 resulted in zero colonies, and pTYB1_336 resulted in 28 colonies (on plates with high cell concentration). Transformation with pTYB1_336 was evaluated as successful and ready for protein production. Transformation with pTYB1_272 was evaluated as successful, but since the process needed to be repeated for pTYB1_295, it was decided to do the same with pTYB1_272 to obtain better transformation results.

Transformation with OverExpress c41(DE3) can be difficult, so it it was decided to try again with a solution of more concentrated plasmid DNA. LA plates with *E. coli* DH5α transformed with the ligation products pTYB1_272 and pTYB1_295 in Section 3.2.7 were used to amplify and isolate the plasmids, as described in Section 3.2.4. To obtain a higher concentration of plasmid DNA, the last step with the DNA Clean & Concentrator-5 kit was included. DNA was isolated from three different colonies for both plasmids, and to confirm isolation of ligated plasmids resulting from gene insert, a restriction digestion with NdeI and XhoI was performed following the protocol in Section 3.2.5. DNA fragments in the restriction digests were separated by agarose gel electrophoresis as described in Section 3.2.6.

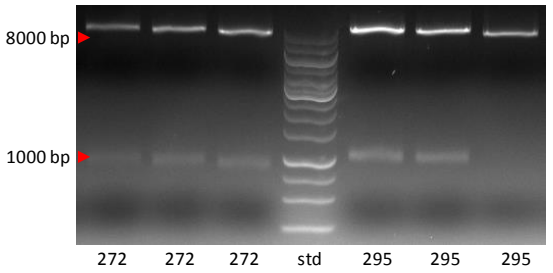


Figure 4.11: Visualized DNA fragments on agarose gel, from restriction digestion to control successful insertion of *algE6A_272* and *algE6A_295* in the expression vector pTYB1. std: 0.75 µL GeneRuler 1 kb DNA Ladder (Thermo Fisher Scientific), 272: restriction digest of pTYB1_272, 295: restriction digest of pTYB1_295. Side labels identify the lengths, in bp, of specific fragments in the DNA ladder.

An image of the visualized DNA fragments after separation on agarose gel is presented in Figure 4.11. All six digestions gave rise to a DNA band close to the 8000 bp marker, and this DNA fragment can be identified as the opened vector pTYB1 (7439 bp). The three digests of pTYB1_272, and digests 1 and 2 of pTYB1_295, all gave rise to a second band close to the 1000 bp marker, the expected area for their respective

target genes, *algE6A_272* (818 bp) and *algE6A_295* (887 bp). Compared to the DNA ladder, the bands lie slightly higher than the expected position of the target gene fragments. However, the *algE6A_272* band lies lower than the *algE6A_295* band and there are no extra bands present, so the plasmids were evaluated to contain the expected gene inserts. Digest 3 of pTYB1_295 did not give rise to a second band, which indicates transformation with religated vector. pTYB1_272, verified in control digest 3, and pTYB1_295, verified in control digest 2, were chosen for transformation of the expression strain.

E. coli OverExpress c41(DE3) was transformed with pTYB1_272 and pTYB1_295 following the protocol in Section 3.2.3. Transformations were performed using 1, 3 and 6 µL DNA. The first transformations of OverExpress c41(DE3) were performed using 30 µL DNA thought to be of very low concentration, and smaller volumes with higher concentration could increase transformation efficiency. Incubation time on ice after addition of DNA was 30 minutes, and transformed cells were not plated out in a 10x dilution. The transformation was evaluated as successful, with several distinct colonies appearing after incubation. A cell count of plates with the same cell concentration as the transformation reaction is presented in Table 4.6.

Table 4.6: Cell count from transformation of OverExpress c41(DE3) with pTYB1_272 and pTYB1_295.

Plasmid	Cell count
pTYB1_272 1 µL	22
pTYB1_272 3 µL	29
pTYB1_272 6 µL	106
pTYB1_295 1 µL	14
pTYB1_295 3 µL	30
pTYB1_295 6 µL	36

4.2.5 Production and Purification of AlgE6A_272, AlgE6A_295, and AlgE6A_336

Small Scale Production to Verify Protein Expression

A small scale production was performed as described in Section 3.2.8 to verify IPTG inducible expression of AlgE6A_272, AlgE6A_295, and AlgE6A_336 in the final clones. Cell lysates from both induced and uninduced cultures were analysed by SDS-PAGE as described in Section 3.2.10, and an image of the gel is presented in Figure 4.12. Molecular weights of the recombinant proteins were obtained using ProtParam (ExPASy, Gasteiger et al. (2005)) and are listed in Table 4.7 (molar extinction coefficients are included as they will be used in Section 4.2.6). Also included in the table are the

expected molecular weights of the proteins in their protein-intein-CBD complexes. The small scale control productions were not purified on a resin column by FPLC, and have therefore not been through DTT induced intein-mediated cleavage where the target protein is separated from the complex. ~56 kDa was used as the added weight of the intein and CBD (Chong et al., 1998).

All induced cultures gave rise to a strong protein band that was not present in the uninduced cultures, indicated with red arrows in Figure 4.12. These bands also all lie slightly above the 75 kDa marker on the protein standard, in the 85 000–90 000 kDa area, as expected for the protein-intein-CBD complexes.

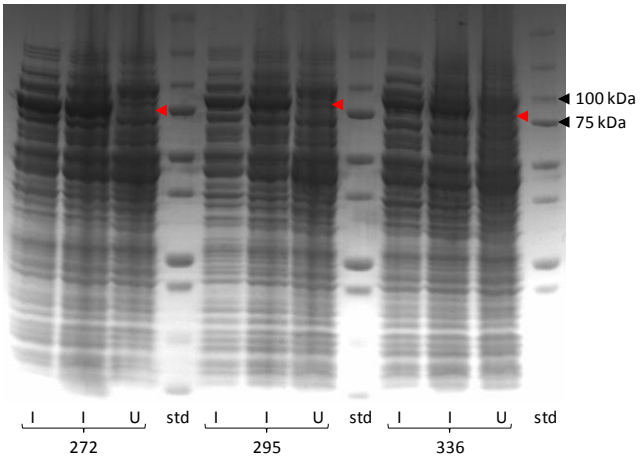


Figure 4.12: Image of SDS-PAGE gel with crude cell extract from induced and uninduced cultures. 10 µL and 20 µL cell lysate was prepared with 5 µL and 10 µL 3x SDS loading buffer, respectively, and the total volumes were loaded on one well each. 20 µL uninduced cell lysate was prepared with 10 µL 3x SDS loading buffer, and the total volume loaded on one well. std: 5 µL Precision Plus Protein^{TM} Dual Color standard (Bio-Rad), I: induced culture, U: uninduced culture, 272: AlgE6A_272, 295: AlgE6A_295, 336: AlgE6A_336. Side labels identify the molecular weights, in kDa, of specific recombinant proteins in the standard.

Main Production

The epimerases AlgE6A_272, AlgE6A_295, and AlgE6A_336, were produced and purified as described in Section 3.2.8 and Section 3.2.9, using the final clones from Section 4.2.4. Superimposed chromatograms from purification by FPLC are shown in Figure 4.13. The chromatograms present UV absorption at 280 nm ($A_{280\text{nm}}$) as

Table 4.7: Molecular weights and molar extinction coefficients, obtained by using ProtParam (ExPASy, Gasteiger et al. (2005)), for AlgE6A_272, AlgE6A_295, and AlgE6A_336, and molecular weights of their respective target protein-intein-CBD complexes. ~56 kDa was used as the added weight of the intein and CBD (Chong et al., 1998). Molecular weights are given in Da, and molar extinction coefficients are given in $M^{-1}cm^{-1}$.

Protein	Molecular weight [Da]	Molecular weight in complex [Da]	Molar extinction coefficient [$M^{-1}cm^{-1}$]
AlgE6A_272	28848	84848	21890
AlgE6A_295	31254	87254	23380
AlgE6A_336	35530	91530	27850

a function of volume (mL) eluted from the column, giving an indication of protein concentration in the effluent. Fractions A1-A6 are indicated along the x-axis and all three proteins were mainly eluted in their respective A1-A3 fractions, as measured A_{280nm} approaches a baseline value after fraction A3. The areas under the curves in the chromatogram also suggest that the A2 fractions have the highest protein concentration. Fractions A4-A6 were discarded.

SDS-PAGE to evaluate fractions from FPLC was performed as described in Section 3.2.10 and the result is presented in Figure 4.14. Molecular weights for AlgE6A_272, AlgE6A_295, and AlgE6A_336, can be found in Table 4.7. Fractions A2 and A3 gave rise to strong protein bands for all three epimerases, while only a weak band appeared for the A1 fractions which were discarded. All protein bands corresponded to the expected location of their respective epimerases, with molecular weights of ~28.8 (AlgE6A_272), ~31.3 (AlgE6A_295), and ~35.5 kDa (AlgE6A_336). Unwanted protein bands, indicative of sample contamination, was absent in all fractions.

4.2.6 Estimation of Protein Concentration

NanoDrop

Protein concentration was estimated for all three epimerases in fractions A2 and A3 from FPLC by measuring A_{280nm} on a NanoDrop^{TM} ND-1000 UV-Vis spectrophotometer (see Section 3.2.11). Molar extinction coefficients are listed in Table 4.7. The protein concentrations obtained were normalized relative to their respective molar extinction coefficients, and with respect to the standard, the absorbance from 1 mg/mL will be approximately 2.2 times higher for AlgE6A_272, 2.3 times higher for AlgE6A_295, and 2.8 times higher for AlgE6A_336. Estimated protein concentrations are presented in Table 4.8. AlgE6A_272 was most concentrated, while AlgE6A_336

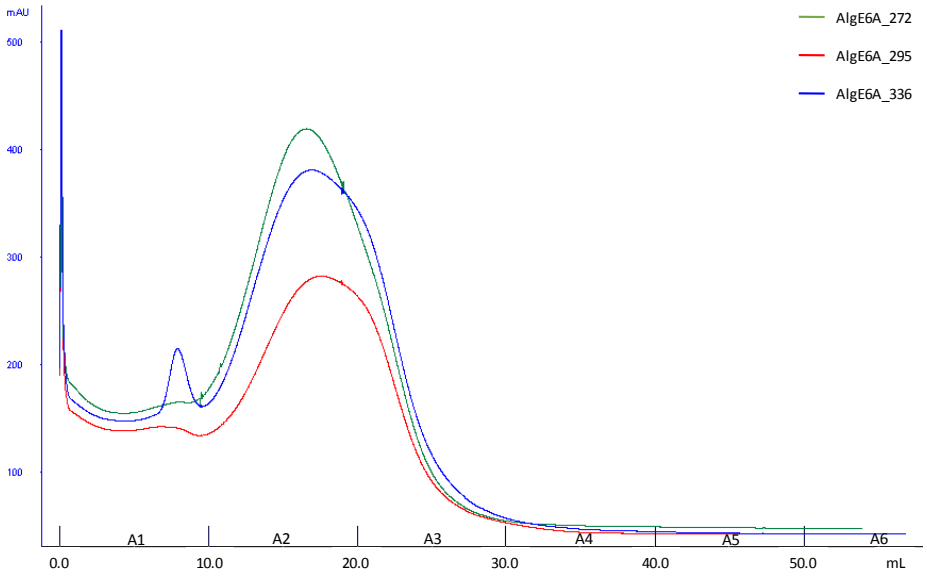


Figure 4.13: Superimposed chromatograms from purification of AlgE6A_272 (green), AlgE6A_295 (red), and AlgE6A_336 (blue), by FPLC on a chitin resin column. UV absorption, $A_{280\text{nm}}$, as an indication of protein concentration in the effluent, is given as a function of volume eluted from the column, in mL. Fractions A1-A6 are indicated along the x-axis.

had the lowest concentration. All three proteins were of higher concentration in their respective A2 fraction.

SDS-PAGE

Protein concentration was also estimated by relative quantification of protein bands on gel from SDS-PAGE, as described in Section 3.2.11, and the results are presented in Table 4.9. Gel image, density analysis from ImageLab^{TM}, and the necessary calculations, can be found in Appendix D. No concentration was obtained for AlgE6A_295 A3 due to problems with Image Lab, this is further explained in the appendix. Consistent with estimations from NanoDrop, AlgE6A_272 was most concentrated, but the lowest concentration was now estimated for AlgE6A_295, not AlgE6A_336. All three proteins were of higher concentration in their respective A2 fraction. Estimation by SDS-PAGE resulted in much lower concentrations overall, roughly 5 to 10 times lower compared to results from NanoDrop.

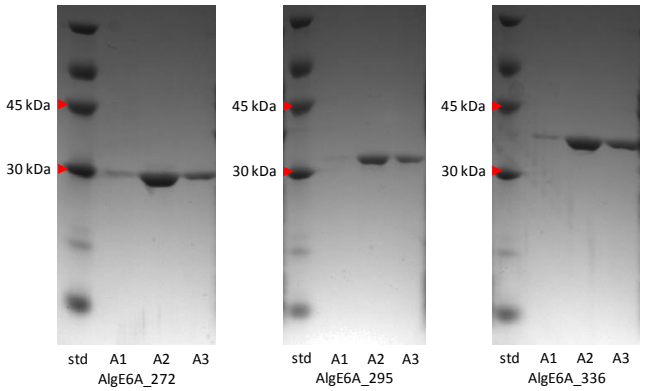


Figure 4.14: Gel images from SDS-PAGE to evaluate fractions from purification of AlgE6A_272, AlgE6A_295, and AlgE6A_336 by FPLC on a chitin resin column. 5 µL FPLC-fraction was prepared with 5 µL 3x SDS loading buffer, and 10 µL was loaded on a well. A1-A3: Fractions A1-A3 from FPLC purification, std: 3 µL Amersham Low Molecular Weight Calibration Kit Protein Mixture standard (GE Healthcare). Side labels identify molecular weights, in kDa, of specific proteins in the standard.

Table 4.8: Protein concentrations estimated for AlgE6A_272, AlgE6A_295, and AlgE6A_336, by NanoDrop. Protein concentration was estimated for FPLC fractions A2 and A3, and are given by mass (mg/mL), normalized relative to their molar extinction coefficients ε (mg/mL), and as molarity (M) after normalizing to ε.

Protein	**Fraction**	**Mass** [mg/mL]	**Normalised to ε** [mg/mL]	**Molar** [M]
A2	**AlgE6A_272**	2.15	0.982	3.40×10^{-5}
	AlgE6A_295	1.67	0.714	2.29×10^{-5}
	AlgE6A_336	1.33	0.478	1.34×10^{-5}
A3	**AlgE6A_272**	0.57	0.260	9.03×10^{-6}
	AlgE6A_295	0.38	0.163	5.20×10^{-6}
	AlgE6A_336	0.44	0.158	4.45×10^{-6}

4.2.7 Epimerase Activity Microassay

Epimerase activity of fractions A1, A2, and A3, was tested for all three proteins in a microassay based on the activity of G-lyase as described in Section 3.2.12. The difference in absorbance at 230 nm ($\Delta A_{230\text{nm}}$) before and after incubation with G-lyase was used to estimate epimerase activity. Results for AlgE6A_336 is presented

Table 4.9: Protein concentrations estimated for AlgE6A_272, AlgE6A_295, and AlgE6A_336, by relative quantification of protein bands from SDS-PAGE. Concentration was estimated for their FPLC fractions A2 and A3. No concentration was estimated for AlgE6A_295 A3, see Appendix D for information. Concentration is given by mass [mg/mL] and molarity [M].

Fraction	**Protein**	**Mass** [mg/mL]	**Molar** [M]
A2	**AlgE6A_272**	0.130	4.49×10^{-6}
	AlgE6A_295	0.066	2.11×10^{-6}
	AlgE6A_336	0.089	2.49×10^{-6}
A3	**AlgE6A_272**	0.036	1.26×10^{-6}
	AlgE6A_295	N/A	N/A
	AlgE6A_336	0.038	1.06×10^{-6}

in Table 4.10, while equivalent tables for AlgE6A_272 and AlgE6A_295 can be found in Appendix G. Protein concentration for the AlgE6 standard added to the first row was 0.2 mg/mL, while for AlgE6A_272, AlgE6A_295, and AlgE6A_336 it was the same as the concentrations in each fraction. The grey cells indicate results from wells containing the AlgE6 standard, while the yellow cells indicate wells with the highest concentration of the protein to be tested.

In all three assays, activity could clearly be observed for the AlgE6 standard, and the activity steadily decreased with increasing dilution, indicating that the microassays worked as expected. However, no such observations of activity resulted from measurements of wells containing AlgE6A_272 or AlgE6A_295. Results for AlgE6A_336 indicated a very low level of epimerase activity in fractions A2 and A3, as measured $\Delta A_{230\text{nm}}$ in the yellow cells was noticeably higher for A2 and A3, and larger than that of the reference wells.

4.2.8 Epimerase Activity Assay for Analysis by ^{1}H-NMR

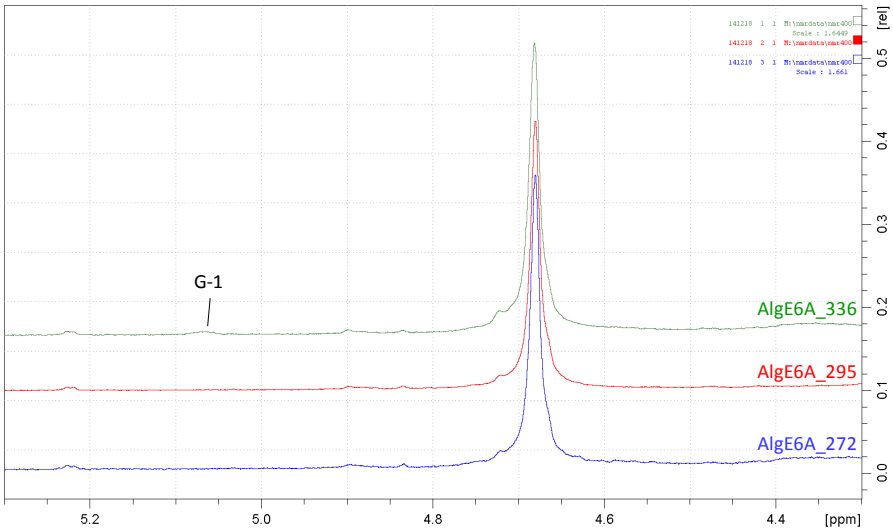
The microassays in Section 4.2.7 indicated that AlgE6A_336 could have low levels of epimerase activity, while AlgE6A_272 and AlgE6A_295 were non-functional. For a more definitive evaluation of epimerase activity in the modified epimerases, poly-M was epimerized with AlgE6A_272, AlgE6A_295, and AlgE6A_336, and analysed by ^{1}H-NMR. Procedures for epimerization, sample preparation, and ^{1}H-NMR, are described in Sections 3.2.13 and 3.2.15. Due to the high concentration of enzyme in the microassay resulting in hardly any activity, it was decided to exaggerate the conditions for epimerization by adding 8 mL AlgE6A_272, AlgE6A_295, or AlgE6A_336, to

Table 4.10: Result of epimerase activity microassay for AlgE6A_336, fractions A1-A3, presented as the difference in absorbance at 230 nm ($\Delta A_{230\text{nm}}$) before and after incubation with G-lyase. Std: AlgE6 standard used as reference, (ref.): reference wells without enzyme, A1-A3: fractions A1-A3 from purification by FPLC. Grey cells indicate wells containing the AlgE6 standard. Yellow cells indicate wells with the highest concentration of AlgE6A_336 A1-A3. The first column indicates level of enzyme dilution for each row.

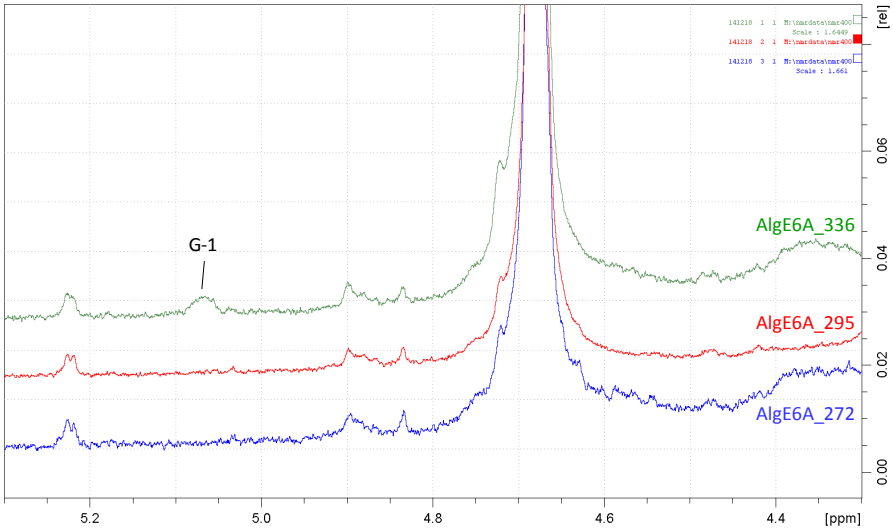
				AlgE6A_336				
Dilution	**Std**	**A1**	**A1**	**A2**	**Std**	**A2**	**A3**	**A3**
(ref.)	0.049	0.049	0.031	0.024	0.035	0.022	0.056	0.048
2x	1.802	0.014	0.010	0.148	1.838	0.152	0.101	0.120
4x	1.745	0.032	0.044	0.118	1.746	0.127	0.096	0.096
8x	1.583	0.039	0.057	0.108	1.593	0.096	0.084	0.092
16x	0.992	0.046	0.057	0.087	1.089	0.079	0.075	0.080
32x	0.425	0.052	0.066	0.077	0.500	0.069	0.069	0.070
64x	0.138	0.057	0.061	0.067	0.176	0.069	0.072	0.073
128x	0.075	0.055	0.059	0.070	0.110	0.063	0.064	0.069
256x	0.068	0.054	0.055	0.052	0.078	0.068	0.063	0.068
512x	0.061	0.053	0.054	0.072	0.069	0.052	0.060	0.053
1024x	0.055	0.055	0.053	0.049	0.052	0.076	0.064	0.049
2048x	0.066	0.073	0.047	0.053	0.046	0.039	0.056	0.067

each sample, and incubate for three nights at 37 °C. Each sample was prepared with 20 mg partially hydrolysed poly-M (DP~70) and a total reaction volume of 14 mL. Acid hydrolysis before ^{1}H-NMR was not needed due to the poly-M already being partially hydrolyzed.

The results are presented in Figure 4.15 as superimposed spectra from ^{1}H-NMR recordings. Poly-M epimerized by AlgE6A_336 gave rise to a weak signal at the G-1 marker, that is barely visible in Figure 4.15a, and a closer view is therefore included in Figure 4.15b. Recordings of AlgE6A_272 and AlgE6A_295 epimerized poly-M showed no indication of G-content in the product.



(a)



(b)

Figure 4.15: Superimposed spectra from ^{1}H-NMR of poly-M after epimerization with AlgE6A_272 (blue), AlgE6A_295 (red), and AlgE6A_336 (green). G-1 assigns the expected position of a signal resulting from G-residues. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.

Chapter 5. Discussion

5.1 Comparative Study of Full Length AlgE4 and its A-module AlgE4A

Earlier work by Dypås (2014) found that the A-module in AlgE6 was adequately active compared to the full length epimerase, and also suggested that the role of the R-module is mostly related to the epimerization pattern and is highly dependent on the calcium concentration. Due to the high degree of sequential homology between the A-modules in AlgE4 and AlgE6, this first part of the study compared full length AlgE4 to its A-module, as it would be interesting to see if AlgE4A behaves similar to AlgE6A.

5.1.1 Evaluation of AlgE4 and AlgE4A Production

AlgE4 and AlgE4A were produced in *E. coli* T7 Express with the plasmids pTB182 and pTB26, respectively, using the IMPACT^{TM} system for expression an purification. After purification on a chitin resin column, the FPLC fractions were evaluated by SDS-PAGE. For both AlgE4 and AlgE4A, two fractions with high protein concentration were identified, A1 and A2, and none of the fractions gave rise to more than one protein band. Molecular weights for AlgE4 and AlgE4A are 55 707 kDa and 39 798 kDa, respectively, however, purified AlgE4 gave rise to a band indicating a weight of ~70 kDa, and the protein band from purified AlgE4A appeared in line with the 50 kDa marker. Aberrant migration during SDS-PAGE has been observed for epimerases previously (Ertesvåg et al., 1994), and the SDS-PAGE results are therefore not interpreted as a fault in production. The size difference identifies AlgE4 as the larger protein, and the absence of contaminants verifies successful use of the IMPACT^{TM} system. As AlgE4 and AlgE4A are expressed in the protein-intein-CBD complex, they are the only proteins that could be eluted at both high concentration and purity after on-column cleavage.

Protein concentration was estimated with a NanoDrop spectrophotometer using the $A_{280\text{nm}}$-method, and by relative quantification of protein bands from SDS-

PAGE. These two methods gave different results, which stems from neither being highly accurate in this case. Measurements of concentration using NanoDrop could have been strongly affected by solvent in the samples. DTT, used in purification by FPLC, affects absorbance readings at 280 nm and since the DTT concentration is unknown in the different fractions, it was not possible to produce an accurate reference (blank) reading. In addition to this, NanoDrop measurements depend on the molar extinction coefficient of the protein that is to be measured. ProtParam (ExPASy) was used to calculate the coefficients, but this tool does not consider the 3D-structure of a protein, only its linear sequence (Gasteiger et al., 2005). For a more accurate measurement by NanoDrop in relation to ε, proteins need to be denatured before measuring absorbance, which was not done in this case due to NanoDrop primarily being used as a quick confirmation of protein presence in the samples.

SDS-PAGE is more accurate in this case, as it gives a visual indication of the relative protein concentration between samples. For example, estimation by NanoDrop resulted in the same concentration for AlgE4A A1 and A2, while SDS-PAGE clearly indicates less protein in A2. NanoDrop also resulted in higher concentration of protein in AlgE4A samples, compared to AlgE4, even though the gel images in Figure 4.2 reveal the opposite. The values estimated by SDS-PAGE quantification gave a more reliable concentration ratio between samples, however, the method was not optimized and the results were used as a guide, not an absolute. Possible steps for optimization could have been the use of a densitometer for imaging, which gives better resolution with Coomassie Blue, and by finding what amount of sample resulted in the clearest bands.

Colorimetric assays based on proteins reacting quantitatively with dyes or metal ions, such as the Bradford or Lowry assays, could have been used as supplemental methods for more accurate estimation of protein concentration, if needed. Particularly the Bradford assay could have been a good supplement to NanoDrop and SDS-PAGE, as it is compatible with low concentrations of DTT (<5 mM) (Green and Sambrook, 2012), and a round of dialysis would be enough to adequately lower DTT concentration in samples.

5.1.2 Epimerase Activity Microassay

Epimerase activity was measured in a microvolume assay by utilizing the action of G-lyase that creates an unsaturated residue, acting as a cromophore, by cleavage of G-G or G-M linkages. The measured $\Delta A_{280\text{nm}}$ clearly indicated higher activity in the A1 FPLC fractions, which also had highest protein concentration, and these were selected for further study. Activity estimated per mole of protein resulted in AlgE4 having higher activity than AlgE4A when based on estimations from NanoDrop. Estima-

tions from SDS-PAGE on the other hand, resulted in close to identical activity in both proteins. Due to inaccuracies in both the microassay and the estimated concentrations, it is difficult to quantify this activity. However, considering that the SDS-PAGE method gave a better indication of the concentration ratio, it is more likely that the activity difference was smaller than that suggested by NanoDrop. This method was mainly used to verify activity in epimerases after production, and to choose fractions for use in analysis on a larger scale.

5.1.3 Effect of the R-module on Calcium Dependence in Epimerase Activity

The R-module's effect on the mode of action in relation to calcium, was studied by a calcium concentration assay with analysis by ^{1}H-NMR to obtain chemical composition. ^{1}H-NMR spectra revealed a clear difference in the level of epimerization in poly-M epimerized by AlgE4 and that epimerized by AlgE4A. AlgE4 and AlgE4A both had little to no activity at the lowest Ca^{2+} concentration, but while AlgE4 reached peak activity at 2 mM Ca^{2+}, the highest G-content introduced by AlgE4A was measured at 8 mM (only slightly lower at 6 mM). The G-content at all sample points was also consistently lower from epimerization with AlgE4A, and the use of freeze dried enzyme for better control of concentration, suggests that this reflects a real difference in activity. The results indicate that the R-module in AlgE4 lowers the Ca^{2+} concentration needed for full activity, consistent with previous findings (Ertesvåg and Valla, 1999), and also facilitates a stable degree of epimerization over a wider concentration range. This dependency was not observed in the previous work with AlgE6, where the A-module and the full length epimerase appeared to have the same narrow concentration optimum and a comparable degree of epimerization at all sample points (Dypås, 2014).

The first calcium dependence assay resulted in complete epimerization of all samples, and the results could not be used to evaluate the R-module's role in relation to calcium. Still, it is worth noting that this result did show that the A-module in AlgE4 is capable of epimerization to the same degree as full length AlgE4. Although G-content in AlgE4A epimerized poly-M is slightly lower throughout, the difference in F_G seen at 5 mM is so small that it holds no real significance.

5.1.4 Effect of the R-module on the Epimerization Pattern of AlgE4

Due to the high degree of sequential homology in the A-modules of AlgE4 and AlgE6, despite the large difference in their product formation, it has been speculated that the R-modules are important for this difference in function. This speculation was not reflected in the ^{1}H-NMR spectra, as they did not indicate G-block introduction

in any of the samples, and clearly show that the block type is highly dependent on the A-modules alone. Product specificity stemming from the environment in the active site has previously been proposed by Rozeboom et al. (2008).

The previous study on AlgE6 suggested that the R-module does have an effect on block length (Dypås, 2014), and samples epimerized by AlgE4 and AlgE4A to approximately the same G-content (and not to full epimerization), were analysed by HPAEC-PAD to obtain information about block length. The characteristic pattern for AlgE4 with introduction of mainly long MG-blocks was observed for the A-module as well, but they were noticeably shorter when compared to the AlgE4 product. This suggests that the R-module increases the level of processivity in AlgE4, possibly by contributing to a stronger association with the alginate polymer as the enzyme moves along the chain. Specific and strong interactions between the R-module of AlgE4 and a pentameric mannuronan (M_5), have been demonstrated (Aachmann et al., 2006), supporting these findings.

Processivity has previously been proposed to ultimately stem from structures in the A-module, as a hybrid enzyme of AlgE2 with residues 215-262 from AlgE4A displayed sings of processivity, despite AlgE2 otherwise being non-processive (Bjerkan et al., 2004a). Rozeboom et al. (2008) later suggested that two "clamp" structures, enclosing the polymer in the active cleft, are involved in making AlgE4 a processive enzyme. This was backed up by half of the second clamp being part of the section transferred to the hybrid AlgE2. Epimerases introducing MG- or G-blocks also seems to be linked to whether it is processive or works by a preferred attack mechanism, respectively (Bjerkan et al., 2004a). Processive formation of G-blocks would involve a 180° rotation of the enzyme relative to the polymer chain, whereas processive formation of MG-blocks only needs to keep moving along the polymer two residues at a time.

5.1.5 Effect of the R-module on Temperature Stability

Enzyme activity is highly dependent on temperature, much due to its effect on structural integrity. The large structural change caused by removing the R-module could destabilize the protein, and epimerization of poly-M was therefore performed at 40 °C, 50 °C, and 60 °C, instead of the optimal 37 °C. Results from ^{1}H-NMR spectra showed that at 40 °C and 50 °C, the level of epimerization for AlgE4A was approximately half of that for AlgE4. This was reflected in the calcium dependence assay, where activity was significantly lower for AlgE4A at the optimal Ca^{2+} concentration for AlgE4. From 40 °C to 50 °C there was only a slight increase in AlgE4 activity and a slight decrease in AlgE4A activity. This difference is too small to hold much significance, especially when considering that the increase for AlgE4 is known to be a

random variation, as its optimum is 37 °C. At 60 °C there was no indication of AlgE4A activity, and AlgE4 had only introduced a very low G-content. Whether the complete lack of activity in AlgE4A is a reflection of lower temperature stability, or if it is simply caused by its baseline activity being half of that for AlgE4, can not be concluded from this experiment alone. Testing the epimerases at temperatures between 50 °C and 60 °C, could be of interest in future work to possibly identify a difference in temperature stability. Also, if excess enzyme was used so that all samples are fully epimerized at low temperatures, it could be possible to identify a point for complete loss of function. Another possibility is to increase the calcium concentration, as AlgE4 and AlgE4A exhibited a similar degree of epimerization above the AlgE4 optimum. However, high concentrations of Ca^{2+} ions have been shown to stabilize the structure of the R-module (Aachmann et al., 2006), and increasing the concentration could have different stabilizing effects on AlgE4 and AlgE4A, creating a new unknown factor.

5.2 Truncated Versions of the A-module from AlgE6

As a continuation of the work performed by Dypås (2014), the second part of this study focused on the connection between structure and function within the A-module itself. This was done by creating AlgE6A clones with varying degree of truncation at their C-terminal end. Not only is this of interest in an effort to elucidate details of the mechanisms behind the epimerase activity in AlgE6, it is also highly relevant for industrial purposes. The possibility of producing a smaller epimerase with retained functionality could have financial impact due to lowered production cost of the enzyme itself, and by changing the conditions needed for epimerization, as a smaller epimerase will have better access to substrate in a tight space.

5.2.1 Modification of the AlgE6 Gene

The A-module in AlgE4, determined by Rozeboom et al. (2008), was used as a model when visualizing the A-module in AlgE6. The amino acid residues Ser^{272}, Gln^{295}, and Ser^{336}, in AlgE4A, were chosen as C-terminal ends for three truncated versions of the A-module. Due to the C-terminal region consisting of similar repetitive sections, it was hypothesized that changes in this region would be less disruptive compared to changes in the more complex and unique structure of the N-terminal. The specific residues were chosen from considerations of avoiding changes to stretches with ordered secondary structure, avoiding creation of a hydrophobe C-terminal, and ending the polypeptide after a full turn of the β-helix. Corresponding amino acid residues in the A-module of AlgE6 were identified as Ser^{272}, Lys^{295}, and Ser^{336}, and it was decided to produce three modified proteins by removing 49 (AlgE6A_336),

90 (AlgE6A_295), and 114 (AlgE6A_272) residues from the C-terminal end. It was hypothesized that AlgE6A_336 would still be functional as these 49 residues appear to mainly be an extension of structure with a stabilizing effect, but not essential for epimerization. AlgE6A_295 involved removal of the loop extending from the β-helix, and this was suspected to have a significant impact on epimerase activity, leading to low level functionality of the enzyme. The smallest protein, AlgE6A_272, was not expected to display epimerase activity, as the complete removal of more than one fourth of the polypeptide is relatively drastic and was suspected to negatively affect protein integrity.

5.2.2 Evaluation of the Cloning Process

To amplify the target gene, *E. coli* DH5α was transformed with plasmids pUC57_272, pUC57_295, and pUC57_336, and transformants were cultured for isolation of plasmid DNA. This was also done to produce pTYB1, but with previously transformed cells from frozen glycerol stock. The isolated plasmid DNA was digested with restriction endonucleases to excise the target genes from pUC57 and insert them in the expression vector pTYB1. The excised target genes and the opened vector were separated by agarose gel electrophoresis. The restriction digestions were visualized in Clone Manager Professional v.6 (Sci-Ed Software), and all the resulting DNA fragments appeared on the gel as expected relative to the DNA ladder. The only fragment that could not be observed was the 38 bp fragment excised to open the vector, but due to its small size this was not surprising. A 38 bp fragment would likely migrate off the gel, and if not, there would probably be too little DNA to create a discernible band. Target genes and opened vector were therefore isolated from the gel for ligation.

E. coli DH5α was transformed with ligated vector and insert, and with a control for religated vector. Results from transformation gave a clear indication of successful ligation, as the background from religation was estimated to be ~2 % for all inserts. A restriction digestion was also preformed to control the product, and the DNA fragments observed after agarose gel electrophoresis corresponded to the length of the opened vector, and that of each respective insert. Digestion of pTYB1_295 did not give rise to a DNA band corresponding to the vector, instead only a weak smear was observed down the lane. This was most likely due to a contamination with nucleases or from star activity of the restriction enzymes. The gene insert band from this digestion was also weaker compared to bands for the other two gene inserts, which could be for the same reasons that explain the missing vector. Based on transformation results together with restriction digestions, it was evaluated that the target genes had been successfully inserted in pTYB1.

The expression strain *E. coli* OverExpress c41(DE3) was transformed with the cloned plasmids, pTYB1_272, pTYB1_295, and pTYB1_336. This transformation only yielded a good number of transformants for pTYB1_336. Transformation with pTYB1-_272 resulted in 5 colonies, while no colonies were observed from transformation with pTYB1_295. In an effort to obtain better results it was decided to repeat transformation with pTYB1_272 alongside the second attempt with pTYB1_295.

To increase transformation efficiency, an additional step to further clean and concentrate DNA, was added when producing and isolating pTYB1_272 and pTYB1 _295 the second time. Control restriction digestions were performed and all three isolates of pTYB1_272 gave rise to both vector and gene insert on agarose gel, while two isolates of pTYB1_295 gave corresponding results. For the third pTYB1_295 isolate, only the opened vector could be identified and this was likely a result of transformation with religated vector. The DNA bands identifying the gene inserts appeared slightly higher on the gel than expected from their lengths. However, the DNA ladder resulted in curved bands, indicating anomalies in its migration. Another observation is that all DNA fragments identified as vector or target gene were positioned in relation to each other exactly as their lengths would predict, as they were all displaced similarly in relation to the DNA ladder. The plasmids were evaluated as successful ligation of vector and target gene.

Transformation of *E. coli* OverExpress c41(DE3) with more concentrated DNA resulted in plates with 30-100 distinct colonies, and all three clones were therefore ready for expression.

5.2.3 Evaluation of AlgE6A_272, AlgE6A_295, and AlgE6A_336 Production

AlgE6A_272, AlgE6A_295, and AlgE6A_336, were produced in their respective clones, discussed in Section 5.2.2, using the IMPACT^{TM} system for expression an purification. Before producing the proteins for activity testing, a small scale production was performed to verify IPTG inducible expression in the clones. Cell lysate from induced and uninduced culture was evaluated by SDS-PAGE, and all three clones gave rise to one strong protein band from their induced cultures that was missing from incubation without induction. These bands also corresponded to expected molecular weights of their respective protein-intein-CBD complexes, further confirming that expression functioned as expected.

After purification of the main production on a chitin resin column, the FPLC fractions were evaluated by SDS-PAGE. For all three proteins, the highest concentrations were identified in fractions A2 and A3, and none of the fractions gave rise to more than one protein band. The protein bands were all located in line with the reference at their respective molecular weights. Pure samples with relatively high

concentration of target protein verifies successful expression using the IMPACT^{TM} system.

Protein concentration was estimated with a NanoDrop spectrophotometer using the A_{280nm}-method, and by relative quantification of protein bands from SDS-PAGE. Both methods resulted in highest protein concentration for AlgE6A_272, and all A2 fractions were more concentrated compared to their respective A3 fraction. NanoDrop identified AlgE6A_336 as the protein of lowest concentration, while SDS-PAGE resulted in lowest concentration for AlgE6A_295. Most striking was that estimation from SDS-PAGE gave values roughly 5-10 times lower than those estimated by NanoDrop. Issues with the two methods have already been discussed for AlgE4 in Section 5.1.1, and the same applies here. NanoDrop was mainly used to quickly verify presence of protein, and SDS-PAGE was used to estimate a concentration ratio between samples. More accurate information about the concentration was not needed for the purposes of this study.

5.2.4 Epimerase Activity Microassay

Epimerase activity in AlgE6A_272, AlgE6A_295, and AlgE6A_336 was first measured in a microvolume assay based on the action of G-lyase. No activity could be observed for AlgE6A_272 or AlgE6A_295 from the measured ΔA_{280nm}, but values for AlgE6A_336 indicated the presence of low level activity. Wells containing the AlgE6 standard behaved as expected, confirming that functionality flaws in the assays did not cause the absence of activity. Wells were loaded with excess enzyme, and this also rules out lack of enzyme as the cause. The microassay results therefore suggested that AlgE6A_272 and AlgE6A_295 were non-functional, and that AlgE6A_336 was able to epimerize poly-M, but just barely.

5.2.5 Epimerase Activity Analysed by ^{1}H-NMR

To better evaluate activity in the modified epimerases, their product was also analysed by ^{1}H-NMR. Reaction conditions were exaggerated by addition of excess enzyme and incubation for three nights. The resulting ^{1}H-NMR spectra were consistent with findings from the microassay, suggesting a very small G-content had been introduced by AlgE6A_336, while there was no sign of functionality in AlgE6A_272 and AlgE6A_295.

5.2.6 Evaluation of Loss of Function from C-terminal Truncation

It is clear that structures at the C-terminal end of AlgE6A are more important for epimerase functionality than first thought. AlgE6A_336, which is missing 49 amino

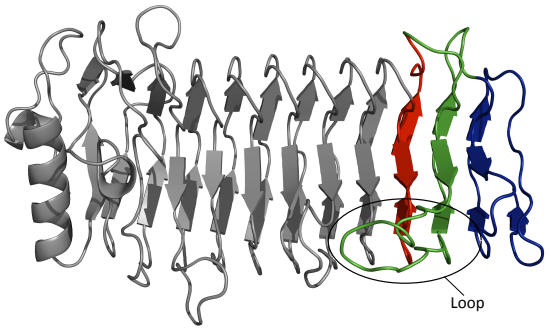


Figure 5.1: Tertiary structure for the A-module in AlgE6 colour coded to modifications. Consecutive removal of each coloured part results in AlgE6A_336 (blue), AlgE6A_295 (green), and AlgE6A_272 (red). Reproduced from Figure 4.8a for clarity, indicating position of the discussed loop extending from the β-helix.

acid residues, retained the ability to epimerize poly-M, but this part of the structure is still needed for any significant level of activity. The removal of 41 more residues in AlgE6A_295 led to all function being lost, revealing that part of this 41 residue peptide is absolutely essential for activity. The most striking structure in this area of the A-module in AlgE4 is the loop made by a two-stranded anti-parallel β-sheet of residues 307-309 and 316-318, that extends from the β-helix. The colour coded model of AlgE6A from Section 4.2.1 is reproduced in Figure 5.1 for clarity. This proposed model of AlgE6A does not contain the anti-parallel β-sheet, but as it is only a homology model with AlgE4A as a template, this is not an important distinction for this discussion. It could be of importance that the side of this loop becomes exposed in AlgE6A_336, leading to changes in its folding, and then, in AlgE6A_295, it is removed completely.

Removal of the entire R-module, or modules, might seem like a bigger invasion of the protein compared to the removal of just 49 more residues, so why are these 49 residues detrimental to functionality? This can be explained by the A- and R-modules being separate and distinct domains. Removal of the R-module ends the protein at a clear transition site between structural domains, which is less invasive and very different from making changes within a domain, such as was done to create the AlgE6A clones.

Structure analysis of the AlgE6A clones was not performed in this study, and whether the loss of function is caused by incorrect folding or by the removal of essential structures, cannot be argued. This would be the natural next step, to elucidate whether the C-terminal region contains residues and structures essential for

function or if it is important for protein folding and stability. NMR spectroscopy structure studies would be beneficial for this, as, unlike X-ray crystallography, it does not require the formation of crystal structure. By comparing NMR spectra of the AlgE6A clones with NMR spectra of the original A-module, observations of large differences would suggest that improper folding of the clones was involved in the loss of function.

5.3 Future Work

Some interesting aspects for future research have already been mentioned, such as continued work on temperature dependence in AlgE4 and AlgE4A, and structure analysis of the AlgE6A clones in an attempt to identify a reason behind the loss of function. Temperature dependence studies were not performed on AlgE6 and AlgE6A in the work by Dypås (2014), and including these epimerases in any future work with AlgE4 and AlgE4A could be beneficial. Clarifying the loss of function in the AlgE6A clones could also be done through further cloning studies, where smaller stretches of amino acid residues are removed, not necessarily starting at the C-terminal. Some work on this is already in progress, including an AlgE4/AlgE6 hybrid where the loop extending from the β-helix has been removed. Results so far are consistent with findings in this study (Aachmann, F. L., personal communication, December 11, 2014), indicating that this loop does have an essential role in epimerase activity.

Further, if structure studies are performed and do indicate proper folding of the AlgE6A clones, work could be done to identify another reason for loss of function, the most obvious being their ability to bind alginate. Analysis of strength and specificity in association between the clones and a mannuronan oligomer would therefore be a natural next step after structure studies.

Within this family of mannuronan C-5 epimerases from *A. vinelandii*, there are always interesting possibilities for future research. In learning more about one, further understanding is also gained concerning all seven, as the reasons behind their differences are clarified. Studies similar to this one could therefore be, and many have been, performed for any of the AlgE epimerases.

Chapter 6. In Conclusion

The aim of this study was to further identify the role of the R-module, specifically in AlgE4, and to learn more about the structure function relationship in the A-module of AlgE6. Work on the A-module in AlgE6, also aimed at creating a smaller epimerase with intact functionality.

A comparative study of the A-module in AlgE4 and the full length epimerase, first of all supported previous studies where the R-module has been found to lower the Ca^{2+} concentration needed for full activity in AlgE4. The R-module also facilitated a stable degree of epimerization over a wider range of concentrations. AlgE4A was able to epimerize poly-M to the same degree as AlgE4, and the R-module is therefore not needed to achieve full epimerization. No G-block introduction was observed for the A-module, and any speculations of product specificity stemming from the R-module were disproved. There was however a decrease in processivity from the removal of the R-module, and it is thought that the R-module could possibly contribute to a stronger association with the alginate polymer, preventing early dissociation. Results from a temperature dependence were inconclusive, and further work is needed.

Structures in the C-terminal region of the A-module were studied by creating three AlgE6A clones with varying degree of truncation at the C-terminal end. It was hypothesized that removal of the first 49 residues (AlgE6A_336) would give a functional epimerase, removal of 90 residues (AlgE6A_295) would significantly reduce epimerase activity, and removal of 114 residues (AlgE6A_272) would give a non-functional epimerase. Only AlgE6A_336 resulted in introduction of G-residues, and the G-content in its product was minimal. Structures at the C-terminal end were therefore revealed to be more essential for activity than first thought. Results suggested that the loop extending from the β-helix could be especially important, as the side of this loop became exposed in AlgE6A_336, where almost all activity was lost, and then completely removed in AlgE6A_295 along with complete loss of function.

The big difference in the modifications of the A-module in AlgE6, and the re-

moval of the R-module, is that the latter cuts the epimerase at a clear point of transition between structurally separate domains, while truncation at the C-terminal removes parts within one such defined domain. Protein modifications performed in the second part of this study were therefore relatively invasive, as reflected in results showing that they were all detrimental to functionality.

Bibliography

Aachmann, F. L., Svanem, B. I., Guntert, P., Petersen, S. B., Valla, S., and Wimmer, R. (2006). NMR structure of the R-module: a parallel β-roll subunit from an azotobacter vinelandii mannuronan c-5 epimerase. *J Biol Chem*, 281(11):7350–6.

Aarstad, O. A., Tøndervik, A., Sletta, H., and Skjåk-Bræk, G. (2011). Alginate sequencing: an analysis of block distribution in alginates using specific alginate degrading enzymes. *Biomacromolecules*, 13(1):106–116.

Andresen, I., Skipnes, O., Smidsrød, O., Østgaard, K., and Hemmer Per, C. H. R. (1977). *Some Biological Functions of Matrix Components in Benthic Algae in Relation to Their Chemistry and the Composition of Seawater*, volume 48 of *ACS Symposium Series*, book section 24, pages 361–381. American Chemical Society. variation of alginate composition in brown algea.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*, 22(2):195–201.

Atkins, E. D. T., Mackie, W., Parker, K. D., and Smolko, E. E. (1971). Crystalline structures of poly-D-mannuronic and poly-L-guluronic acids. *Journal of Polymer Science Part B: Polymer Letters*, 9(4):311–316.

Atkins, E. D. T., Mackie, W., and Smolko, E. E. (1970). Crystalline structures of alginic acids. *Nature*, 225:626–628.

Augst, A. D., Kong, H. J., and Mooney, D. J. (2006). Alginate hydrogels as biomaterials. *Macromolecular bioscience*, 6(8):623–633.

Aune, T. E. V. and Aachmann, F. L. (2010). Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. *Applied Microbiology and Biotechnology*, 85(5):1301–1313.

Bakkevig, K., Sletta, H., Gimmestad, M., Aune, R., Ertesvåg, H., Degnes, K., Christensen, B. E., Ellingsen, T. E., and Valla, S. (2005). Role of the pseudomonas fluorescens alginate lyase (AlgL) in clearing the periplasm of alginates not exported to the extracellular environment. *Journal of bacteriology*, 187(24):8375–8384.

Ballance, S., Holtan, S., Aarstad, O. A., Sikorski, P., Skjåk-Bræk, G., and Christensen, B. E. (2005). Application of high-performance anion-exchange chromatography with pulsed amperometric detection and statistical analysis to study oligosaccharide distributions – a complementary method to investigate the structure and some properties of alginates. *Journal of Chromatography A*, 1093(1–2):59–68.

Banning, D., Dettmar, P., Jolliffe, I., Hampson, F., Onsoyen, E., Field, P., Craig, D., and Kristensen, A. (2002). Forming a protective coating on gastrointestinal mucosal tissue for treating reflux oesophagitis, gastritis, dyspepsia or peptic ulceration. Patent US6395307 B1.

Baumann, U., Wu, S., Flaherty, K. M., and McKay, D. B. (1993). Three-dimensional structure of the alkaline protease of pseudomonas aeruginosa: a two-domain protein with a calcium binding parallel β-roll motif. *The EMBO journal*, 12(9):3357.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., and Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*, 42(Web Server issue):W252–8.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res*, 7(6):1513–23.

Bjerkan, T., Lillehov, B., Strand, W., Skjåk-Bræk, G., Ertesvåg, H., and Valla, S. (2004a). Construction and analyses of hybrid azotobacter vinelandii mannuronan C-5 epimerases with new epimerization pattern characteristics. *Biochem. J*, 381:813–821.

Bjerkan, T. M., Bender, C. L., Ertesvåg, H., Drabløs, F., Fakhr, M. K., Preston, L. A., Skjåk-Bræk, G., and Valla, S. (2004b). The pseudomonas syringae genome encodes a combined mannuronan C-5-epimerase and O-acetylhydrolase, which strongly enhances the predicted gel-forming properties of alginates. *Journal of Biological Chemistry*, 279(28):28920–28929.

Boom, R., Sol, C., Salimans, M., Jansen, C., Wertheim-van Dillen, P., and Van der Noordaa, J. (1990). Rapid and simple method for purification of nucleic acids. *Journal of clinical microbiology*, 28(3):495–503.

Boyd, A. and Chakrabarty, A. M. (1994). Role of alginate lyase in cell detachment of pseudomonas aeruginosa. *Applied and environmental microbiology*, 60(7):2355–2359.

Boyen, C., Kloareg, B., Polne-Fuller, M., and Gibor, A. (1990). Preparation of alginate lyases from marine molluscs for protoplast isolation in brown algae. *Phycologia*, 29(2):173–181.

Brown, B. J. and Preston III, J. F. (1991). L-guluronan-specific alginate lyase from a marine bacterium associated with sargassum. *Carbohydrate research*, 211(1):91–102.

Brownlee, I. A., Allen, A., Pearson, J. P., Dettmar, P. W., Havler, M. E., Atherton, M. R., and Onsoyen, E. (2005). Alginate as a source of dietary fiber. *Crit Rev Food Sci Nutr*, 45(6):497–510.

Campa, C., Oust, A., Skjåk-Bræk, G., Paulsen, B. S., Paoletti, S., Christensen, B. E., and Ballance, S. (2004). Determination of average degree of polymerisation and distribution of oligosaccharides in a partially acid-hydrolysed homopolysaccharide: a comparison of four experimental methods applied to mannuronan. *Journal of Chromatography A*, 1026(1):271–281.

Chavagnat, F., Heyraud, A., Colin-Morel, P., Guinand, M., and Wallach, J. (1998). Catalytic properties and specificity of a recombinant, overexpressed D-mannuronate lyase. *Carbohydr Res*, 308(3-4):409–15.

Chen, C. W. and Thomas Jr, C. A. (1980). Recovery of DNA segments from agarose gels. *Analytical biochemistry*, 101(2):339–341.

Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R. B., and Hirvonen, C. A. (1997). Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene*, 192(2):271–281.

Chong, S., Williams, K. S., Wotkowicz, C., and Xu, M. Q. (1998). Modulation of protein splicing of the saccharomyces cerevisiae vacuolar membrane ATPase intein. *J Biol Chem*, 273(17):10567–77.

Cosloy, S. D. and Oishi, M. (1973). The nature of the transformation process in escherichia coli K12. *Molecular and General Genetics MGG*, 124(1):1–10.

Cote, G. L. and Krull, L. H. (1988). Characterization of the exocellular polysaccharides from azotobacter chroococcum. *Carbohydrate research*, 181:143–152.

Desjardins, P., Hansen, J. B., and Allen, M. (2009a). Microvolume protein concentration determination using the NanoDrop 2000c spectrophotometer. *J Vis Exp*, (33).

Desjardins, P., Hansen, J. B., and Allen, M. (2009b). Microvolume spectrophotometric and fluorometric determination of protein concentration. *Curr Protoc Protein Sci*, Chapter 3:Unit 3.10.

Donati, I., Holtan, S., Morch, Y. A., Borgogna, M., Dentini, M., and Skjåk-Bræk, G. (2005). New hypothesis on the role of alternating sequences in calcium-alginate gels. *Biomacromolecules*, 6(2):1031–40.

Donati, I. and Paoletti, S. (2009). *Material properties of alginates*. Alginates: biology and applications. Springer.

Draget, K. I., Skjåk Bræk, G., and Smidsrød, O. (1994). Alginic acid gels: the effect of alginate chemical composition and molecular weight. *Carbohydrate Polymers*, 25(1):31–38.

Draget, K. I., Smidsrød, O., and Skjåk-Bræk, G. (2005). *Alginates from Algae*. Wiley-VCH Verlag GmbH & Co. KGaA.

Dubendorf, J. W. and Studier, F. W. (1991). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *Journal of molecular biology*, 219(1):45–59.

Dunne, W. M., J. and Buckmire, F. L. (1985). Partial purification and characterization of a polymannuronic acid depolymerase produced by a mucoid strain of pseudomonas aeruginosa isolated from a patient with cystic fibrosis. *Appl Environ Microbiol*, 50(3):562–7.

Dypås, L. B. (2014). Comparative study of the full length mannuronan C-5 epimerase AlgE6 and its catalytic subunit the A-module. Unpublished Project Paper, Norwegian University of Science and Technology, Department of Biotechnology.

Ertesvåg, H., Doseth, B., Larsen, B., Skjåk-Bræk, G., and Valla, S. (1994). Cloning and expression of an azotobacter vinelandii mannuronan C-5-epimerase gene. *Journal of bacteriology*, 176(10):2846–2853.

Ertesvåg, H., Høidal, H. K., Hals, I. K., Rian, A., Doseth, B., and Valla, S. (1995). A family of modular type mannuronan C-5-epimerase genes controls alginate structure in azotobacter vinelandii. *Mol Microbiol*, 16(4):719–31.

Ertesvåg, H., Høidal, H. K., Schjerven, H., Svanem, B. I. G., and Valla, S. (1999). Mannuronan C-5-epimerases and their application for in vitro and in vivo design of new alginates useful in biotechnology. *Metabolic engineering*, 1(3):262–269.

Ertesvåg, H., Høidal, H. K., Skjåk-Bræk, G., and Valla, S. (1998). The azotobacter vinelandii mannuronan C-5-epimerase AlgE1 consists of two separate catalytic domains. *Journal of Biological Chemistry*, 273(47):30927–30932.

Ertesvåg, H. and Valla, S. (1999). The A-modules of the azotobacter vinelandii mannuronan-C-5-epimerase AlgE1 are sufficient for both epimerization and binding of ca^{2+}. *Journal of bacteriology*, 181(10):3033–3038.

Ertesvåg, H., Valla, S., and Skjåk-Bræk, G. (1996). Genetics and biosynthesis of alginates. *Carbohydrates in Europe*, 14:14–18.

Franklin, M. J., Chitnis, C. E., Gacesa, P., Sonesson, A., White, D. C., and Ohman, D. E. (1994). Pseudomonas aeruginosa AlgG is a polymer level alginate C5-mannuronan epimerase. *Journal of bacteriology*, 176(7):1821–1830.

Franklin, M. J. and Ohman, D. E. (1993). Identification of AlgF in the alginate biosynthetic gene cluster of pseudomonas aeruginosa which is required for alginate acetylation. *J Bacteriol*, 175(16):5057–65.

Gacesa, P. (1987). Alginate-modifying enzymes: a proposed unified mechanism of action for the lyases and epimerases. *Febs Letters*, 212(2):199–202.

Gacesa, P. (1992). Enzymic degradation of alginates. *Int J Biochem*, 24(4):545–52.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005). *Protein identification and analysis tools on the ExPASy server*, pages 571–607. Springer.

GenScript (web). pUC57 - simple vector map http://www.genscript.com/gsfiles/techfiles/genscript_puc57-simple_plasmid_map.pdf [downloaded 20.01.15] GenScript Inc. (USA).

Gimmestad, M., Sletta, H., Ertesvåg, H., Bakkevig, K., Jain, S., Suh, S., Skjåk-Bræk, G., Ellingsen, T. E., Ohman, D. E., and Valla, S. (2003). The pseudomonas fluorescens AlgG protein, but not its mannuronan C-5-epimerase activity, is needed for alginate polymer formation. *Journal of bacteriology*, 185(12):3515–3523.

Gombotz, W. R. and Wee, S. (1998). Protein release from alginate matrices. *Advanced Drug Delivery Reviews*, 31(3):267–285.

Gorin, P. A. J. and Spencer, J. F. T. (1966). Exocellular alginic acid from azotobacter vinelandii. *Canadian Journal of chemistry*, 44(9):993–998.

Govan, J. R. W., Fyfe, J. A. M., and Jarman, T. R. (1981). Isolation of alginate-producing mutants of pseudomonas fluorescens, pseudomonas putida and pseudomonas mendocina. *Journal of general microbiology*, 125(1):217–220.

Grant, G. T., Morris, E. R., Rees, D. A., Smith, P. J. C., and Thom, D. (1973). Biological interactions between polysaccharides and divalent cations: The egg-box model. *FEBS Letters*, 32(1):195–198.

Grasdalen, H. (1983). High-field, 1H-NMR spectroscopy of alginate: sequential structure and linkage conformations. *Carbohydrate Research*, 118:255–260.

Grasdalen, H., Larsen, B., and Smidsrød, O. (1979). A pmr study of the composition and sequence of uronate residues in alginates. *Carbohydrate Research*, 68(1):23–31.

Grasdalen, H., Larsen, B., and Smidsrød, O. (1981). ^{13}C-NMR studies of monomeric composition and sequence in alginate. *Carbohydrate Research*, 89(2):179–191.

Green, M. R. and Sambrook, J. (2012). *Molecular cloning: a laboratory manual*. CSH Press, Cold Spring Harbor, N.Y.

Greene, A. and Madgwick, J. (1986). Alginate-modifying enzymes in australian marine algae. *Botanica marina*, 29(4):329–334.

Grimsley, G. R. and Pace, C. N. (2004). Spectrophotometric determination of protein concentration. *Current protocols in protein science*, pages 3.1. 1–3.1. 9.

Guex, N., Peitsch, M. C., and Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*, 30 Suppl 1:S162–73.

Hanahan, D. (1983). Studies on transformation of escherichia coli with plasmids. *Journal of molecular biology*, 166(4):557–580.

Hardin, J., Bertoni, G., Kleinsmith, L. J., and Becker, W. M. (2012). *Becker's world of the cell*. Benjamin Cummings, Boston, MA [etc.].

Hartmann, M. (2000). *Enzymatic tailoring of alginate using mannuronan C-5-epimerases*, volume 2000:119. NTH., Trondheim. Avhandling (dr. ing.) - Norges teknisk-naturvitenskapelige universitet, 2000.

Haug, A. (1964). Composition and properties of alginates. *Rapport (Norsk institutt for tang- og tareforskning)*, 30.

Haug, A. and Larsen, B. (1971). Biosynthesis of alginate: Part II. polymannuronic acid C-5-epimerase from azotobacter vinelandii (lipman). *Carbohydrate research*, 17(2):297–308.

Haug, A., Larsen, B., and Smidsrød, O. (1966). A study of the constitution of alginic acid by partial acid hydrolysis. *Acta Chem Scand*, 20(1):183–190.

Haug, A., Larsen, B., and Smidsrød, O. (1967). Studies on the sequence of uronic acid residues in alginic acid. *Acta Chem Scand*, 21:691–704.

Haug, A., Larsen, B., and Smidsrød, O. (1974). Uronic acid sequence in alginate from different sources. *Carbohydrate Research*, 32(2):217–225.

Haug, A. and Smidsrød, O. (1965). The effect of divalent metals on the properties of alginate solutions. *Acta Chem. Scand*, 19(2).

Helgerud, T., Gåserød, O., Fjæreide, T., Andersen, P. O., and Larsen, C. K. (2009). *Alginates*, pages 50–72. Wiley-Blackwell.

Helling, R. B., Goodman, H. M., and Boyer, H. W. (1974). Analysis of endonuclease R-EcoRI fragments of DNA from lambdoid bacteriophages and other viruses by agarose-gel electrophoresis. *Journal of virology*, 14(5):1235–1244.

Heyraud, A., Colin-Morel, P., Girond, S., Richard, C., and Kloareg, B. (1996). HPLC analysis of saturated or unsaturated oligoguluronates and oligomannuronates. application to the determination of the action pattern of haliotis tuberculata alginate lyase. *Carbohydrate Research*, 291(0):115–126.

Høidal, H. K., Ertesvåg, H., Skjåk-Bræk, G., Stokke, B. T., and Valla, S. (1999). The recombinant azotobacter vinelandii mannuronan C-5-epimerase AlgE4 epimerizes alginate by a nonrandom attack mechanism. *Journal of Biological Chemistry*, 274(18):12316–12322.

Holtan, S., Bruheim, P., and Skjåk-Bræk, G. (2006). Mode of action and subsite studies of the guluronan block-forming mannuronan C-5 epimerases AlgE1 and AlgE6. *Biochem. J*, 395:319–329.

Iacocca, V. F., Sibinga, M. S., and Barbero, G. J. (1963). Respiratory tract bacteriology in cystic fibrosis. *Archives of Pediatrics & Adolescent Medicine*, 106(3):315.

IMPACT^{TM} Manual, N. (2014). Impact^{tm} kit instruction manual, version 3.0, new england biolabs inc. (usa).

Ishikawa, M. and Nisizawa, K. (1981). Polymannuronic acid C-5-epimerase activities in several brown algae and its localization in frond. *Bulletin of the Japanese Society of Scientific Fisheries (Japan)*.

Kennedy, L., Mcdowell, K., and Sutherland, I. W. (1992). Alginases from azotobacter species. *Journal of general microbiology*, 138(11):2465–2471.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-MODEL repository and associated resources. *Nucleic Acids Res*, 37(Database issue):D387–92.

Larsen, B. and Haug, A. (1971). Biosynthesis of alginate: Part I. composition and structure of alginate produced by azotobacter vinelandii (lipman). *Carbohydrate research*, 17(2):287–296.

Larsen, B., Skjåk-Bræk, G., and Painter, T. (1986). Action pattern of mannuronan C-5-epimerase: generation of block-copolymeric structures in alginates by a multiple-attack mechanism. *Carbohydrate Research*, 146(2):342–345.

Linker, A. and Jones, R. S. (1966). A new polysaccharide resembling alginic acid isolated from pseudomonads. *Journal of Biological Chemistry*, 241(16):3845–3851.

Lodish, H. (2013). *Molecular cell biology*. Freeman, New York.

Lyczak, J. B., Cannon, C. L., and Pier, G. B. (2002). Lung infections associated with cystic fibrosis. *Clinical microbiology reviews*, 15(2):194–222.

Madadlou, A., O'Sullivan, S., and Sheehan, D. (2011). Fast protein liquid chromatography. *Methods Mol Biol*, 681:439–47.

Madgwick, J., Haug, A., and Larsen, B. (1973). Polymannuronic acid C-5-epimerase from the marine alga pelvetia canaliculata. *Acta Chem Scand*, 27(9):3592–4.

Marko, M., Chipperfield, R., and Birnboim, H. (1982). A procedure for the large-scale isolation of highly purified plasmid DNA using alkaline extraction and binding to glass powder. *Analytical biochemistry*, 121(2):382–387.

May, T. B. and Chakrabarty, A. (1994). Pseudomonas aeruginosa: genes and enzymes of alginate synthesis. *Trends in microbiology*, 2(5):151–157.

Morea, A., Mathee, K., Franklin, M. J., Giacomini, A., O'Regan, M., and Ohman, D. E. (2001). Characterization of AlgG encoding C5-epimerase in the alginate biosynthetic gene cluster of pseudomonas fluorescens. *Gene*, 278(1):107–114.

Mørch, Y. A., Donati, I., Strand, B. L., and Skjåk-Bræk, G. (2007). Molecular engineering as an approach to design new functional properties of alginate. *Biomacromolecules*, 8(9):2809–14.

Mørch, Y. A., Sandvig, I., Olsen, O., Donati, I., Thuen, M., Skjåk-Bræk, G., Haraldseth, O., and Brekken, C. (2012). Mn-alginate gels as a novel system for controlled release of Mn^{2+} in manganese-enhanced MRI. *Contrast Media Mol Imaging*, 7(2):265–75.

Nakada, H. I. and Sweeny, P. C. (1967). Alginic acid degradation by eliminases from abalone hepatopancreas. *Journal of Biological Chemistry*, 242(5):845–851.

NanoDrop (2008). Technical bulletin T010 protein measurements, NanoDrop 1000 & 8000 rev 4/08. *Thermo Fisher Scientific*.

Nyvall, P., Corre, E., Boisset, C., Barbeyron, T., Rousvoal, S., Scornet, D., Kloareg, B., and Boyen, C. (2003). Characterization of mannuronan C-5-epimerase genes from the brown alga laminaria digitata. *Plant physiology*, 133(2):726–735.

Okazaki, M., Furuya, K., Tsukayama, K., and Nisizawa, K. (1982). Isolation and identification of alginic acid from a calcareous red alga serraticardia maxima. *Botanica marina*, 25(3):123–132.

Onsøyen, E. (1996). Commercial applications of alginates. *Carbohydr. Eur*, 14:26–31.

Otterlei, M., Østgaard, K., Skjåk-Bræk, G., Smidsrød, O., Soon-Shiong, P., and Espevik, T. (1991). Induction of cytokine production from human monocytes stimulated with alginate. *Journal of Immunotherapy*, 10(4):286–291.

Page, W. and Sadoff, H. (1975). Relationship between calcium and uroinic acids in the encystment of azotobacter vinelandii. *Journal of bacteriology*, 122(1):145–151.

Page, W. and Sadoff, H. (1976). Physiological factors affecting transformation of azotobacter vinelandii. *Journal of bacteriology*, 125(3):1080–1087.

Painter, T. J. (1983). Algal polysaccharides. *The polysaccharides*, 2:195–285.

Penaloza-Vazquez, A., Kidambi, S. P., Chakrabarty, A. M., and Bender, C. L. (1997). Characterization of the alginate biosynthetic gene cluster in pseudomonas syringae pv. syringae. *Journal of bacteriology*, 179(14):4464–4472.

Preiss, J. and Ashwell, G. (1962a). Alginic acid metabolism in bacteria i. enzymatic formation of unsaturated oligosaccharides and 4-deoxy-l-erythro-5-hexoseulose uronic acid. *Journal of Biological Chemistry*, 237(2):309–316.

Preiss, J. and Ashwell, G. (1962b). Alginic acid metabolism in bacteria II. the enzymatic reduction of 4-deoxy-L-erythro-5-hexoseulose uronic acid to 2-keto-3-deoxy-D-gluconic acid. *Journal of Biological Chemistry*, 237(2):317–321.

Promega (2012). Promega protocols & applications guide, dna purification, www.promega.com, promega coorporation (usa).

Ramstad, M. V., Ellingsen, T. E., Josefsen, K. D., Høidal, H. K., Valla, S., Skjåk-Bræk, G., and Levine, D. W. (1999). Properties and action pattern of the recombinant mannuronan C-5-epimerase AlgE2. *Enzyme and Microbial Technology*, 24(10):636–646.

Reece, R. J. (2004). *Analysis of genes and genomes*. John Wiley & Sons.

Regand, A. and Goff, H. D. (2003). Structure and ice recrystallization in frozen stabilized ice cream model systems. *Food Hydrocolloids*, 17(1):95–102.

Rehm, B. H., Ertesvåg, H., and Valla, S. (1996). A new azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is part of an alg gene cluster physically organized in a manner similar to that in pseudomonas aeruginosa. *Journal of bacteriology*, 178(20):5884–5889.

Richey, J. (1982). textFPLC: a comprehensive separation technique for biopolymers. *Am. Lab.*, 14:104–129.

Rozeboom, H. J., Bjerkan, T. M., Kalk, K. H., Ertesvåg, H., Holtan, S., Aachmann, F. L., Valla, S., and Dijkstra, B. W. (2008). Structural and mutational characterization of the catalytic A-module of the mannuronan C-5-epimerase AlgE4 from azotobacter vinelandii. *Journal of Biological Chemistry*, 283(35):23819–23828.

Sadoff, H. L. (1975). Encystment and germination in azotobacter vinelandii. *Bacteriological reviews*, 39(4):516.

Seiderer, L., Newell, R., and Cook, P. (1982). Quantitative significance of style enzymes from two marine mussels (choromytilus meridionalis krauss and perna perna linnaeus) in relation to diet. *Marine biology letters*, 3:257–271.

Shilpa, A., Agrawal, S., and Ray, A. R. (2003). Controlled delivery of drugs from alginate matrix. *Journal of Macromolecular Science, Part C: Polymer Reviews*, 43(2):187–221.

Skjåk-Bræk, G. and Espevik, T. (1996). Application of alginate gels in biotechnology and biomedicine. *Carbohydrates in Europe*, 14:19–25.

Skjåk-Bræk, G., Larsen, B., and Grasdalen, H. (1985). The role of o-acetyl groups in the biosynthesis of alginate by azotobacter vinelandii. *Carbohydrate Research*, 145(1):169–174.

Sletmoen, M., Skjåk-Bræk, G., and Stokke, B. T. (2004). Single-molecular pair unbinding studies of mannuronan C-5 epimerase AlgE4 and its polymer substrate. *Biomacromolecules*, 5(4):1288–1295.

Smidsrød, O. (1970). Solution properties of alginate. *Carbohydrate Research*, 13(3):359–372.

Smidsrød, O., Glover, R. M., and Whittington, S. G. (1973). The relative extension of alginates having different chemical composition. *Carbohydrate Research*, 27(1):107–118.

Smidsrød, O. and Haug, A. (1968a). Dependence upon uronic acid composition of some ion-exchange properties of alginates. *Acta Chem. Scand*, 22(6):1989–1997.

Smidsrød, O. and Haug, A. (1968b). A light scattering study of alginate. *Acta Chem. Scand*, 22(3).

Smidsrød, O. and Haug, A. (1972). Dependence upon the gel-sol state of the ion-exchange properties of alginates. *Acta Chem Scand*, 26(5):2063–74.

Smidsrød, O. and Skjåk-Bræk, G. (1990). Alginate as immobilization matrix for cells. *Trends in Biotechnology*, 8(0):71–78.

Smith, D. R. (1996). *Agarose Gel Electrophoresis*, volume 58 of *Methods in Molecular Biology™*, book section 3, pages 17–21. Humana Press.

Snustad, D. P. and Simmons, M. J. (2012). *Genetics*. Wiley, Singapore. 6th ed.

Østgaard, K., Knutsen, S. H., Dyrset, N., and Aasen, I. M. (1993). Production and characterization of guluronate lyase from klebsiella pneumoniae for applications in seaweed biotechnology. *Enzyme and Microbial Technology*, 15(9):756–763.

Svanem, B. I., Skjåk-Bræk, G., Ertesvåg, H., and Valla, S. (1999). Cloning and expression of three new aazotobacter vinelandii genes closely related to a previously described gene family encoding mannuronan C-5-epimerases. *J Bacteriol*, 181(1):68–77.

Svanem, B. I., Strand, W. I., Ertesvåg, H., Skjåk-Bræk, G., Hartmann, M., Barbeyron, T., and Valla, S. (2001). The catalytic activities of the bifunctional azotobacter vinelandii mannuronan C-5-epimerase and alginate lyase AlgE7 probably originate from the same active site in the enzyme. *J Biol Chem*, 276(34):31542–50.

NEB Technical Support (web). Restriction maps - pTYB1 https://www.neb.com/tools-and-resources/interactive-tools/dna-sequences-and-maps-tool [downloaded 20.01.15] new england biolabs inc. (usa).

Thomas, S. (2000). Alginate dressings in surgery and wound management–part 1. *J Wound Care*, 9(2):56–60.

Tøndervik, A., Klinkenberg, G., Aachmann, F. L., Svanem, B. I. G., Ertesvåg, H., Ellingsen, T. E., Valla, S., Skjåk-Bræk, G., and Sletta, H. (2013). Mannuronan C-5 epimerases suited for tailoring of specific alginate structures obtained by high-throughput screening of an epimerase mutant library. *Biomacromolecules*, 14(8):2657–2666.

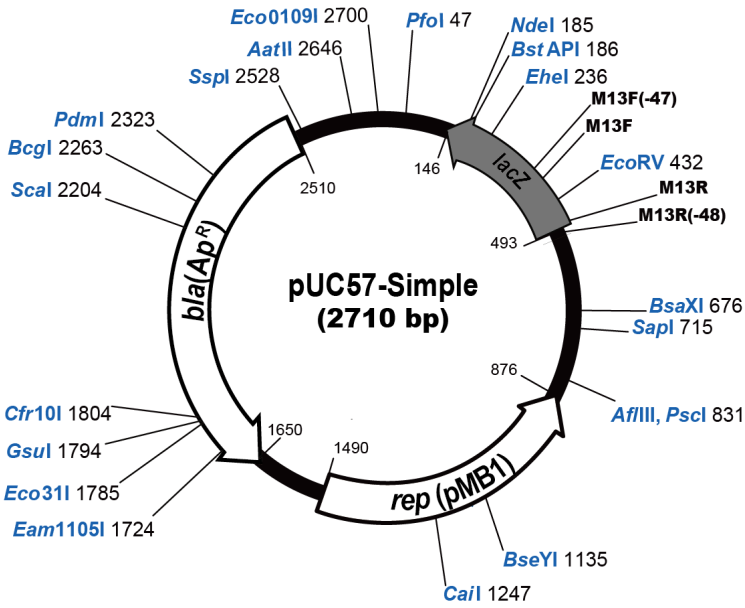
Tøndervik, A., Klinkenberg, G., Aarstad, O. A., Drabløs, F., Ertesvåg, H., Ellingsen, T. E., Skjåk-Bræk, G., Valla, S., and Sletta, H. (2010). Isolation of mutant alginate lyases with cleavage specificity for di-guluronic acid linkages. *Journal of biological chemistry*, 285(46):35284–35292.

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

Weston, A., Brown, M. G., Perkins, H. R., Saunders, J. R., and Humphreys, G. O. (1981). Transformation of escherichia coli with plasmid deoxyribonucleic acid: calcium-induced binding of deoxyribonucleic acid to whole cells and to isolated membrane fractions. *J Bacteriol*, 145(2):780–7.

Appendix A. Synthetic Genes from GenScript (USA) Inc.

Synthetic genes were ordered from GenScript (USA) and were delivered as plasmid constructs of the cloning vector pUC57. A gene map of pUC57 is included in Figure A.1. Restriction sites for NdeI and XhoI were included at either end of the target genes, and then inserted in pUC57 in the Multiple Cloning Site (MCS) by EcoRV (position indicated on the map). The constructs are referred to as pUC57_272, pUC57_295, and pUC57_336, as they, respectively, contain gene insertions coding for the 272, 295, and 336 first amino acids of the A-module in AlgE6, starting at its N-terminal.



Multiple Cloning Sites:

M13F (-47) M13F

5′ C GCC AGG GTT TTC CCA GTC ACG ACG TTG TAA AAC GAC GGC CAG TGA ATT GGA GAT CGG TAC TTC GCG AAT GCG

3′ G CGG TCC CAA AAG GGT CAG TGC TGC AAC ATT TTG CTG CCG GTC ACT TAA CCT CTA GCC ATG AAG CGC TTA CGC

LacZ ← Asn Glu Trp Asp Arg Arg Gln Leu Val Val Ala Leu Ser Asn Ser Ser Pro Val Glu Arg Ile Cys

EcoRV

TCG A**GA TAT C**GG ATG CCG GGA CCG ACG AGT GCA GAG GCG TGC AAG CGA GCT TGG CGT AAT CAT GGT CAT AGC TGT

AGC TCT ATA GCC TAC GGC CCT GGC TGC TCA CGT CTC CGC ACG TTC GCT CGA ACC GCA TTA GTA CCA GTA TCG ACA

Arg Ser Ile Pro Asp Arg Ala Arg Arg Ser Cys Leu Gly Ala His Leu Ser Pro Thr Ile Met Thr Met

M13R

TTC CTG TGT GAA ATT GTT ATC CGC T 3′

AAG GAC ACA CTT TAA CAA TAG GCG A 5′

M13R (-48)

Figure A.1: Gene map for the cloning vector pUC57 (GenScript, web).

Appendix B. pTYB1 - Gene Map

A gene map for the expression vector pTYB1 is included in Figure B.1. Target genes were inserted in pTYB1 for expression, and the vector was opened within the Multiple Cloning Site (MCS) with the restriction endonucleases NdeI and XhoI. The MCS is positioned for translational fusion of the C-terminal of the target gene to the *Sce* VMA intein tag and the chitin binding domain (CBD) (Chong et al., 1997). Transcription is under the control of the T7 promoter, just upstream of the *lac* repressor (Dubendorf and Studier, 1991). The *lac* repressor minimizes basal expression from the T7 promoter.

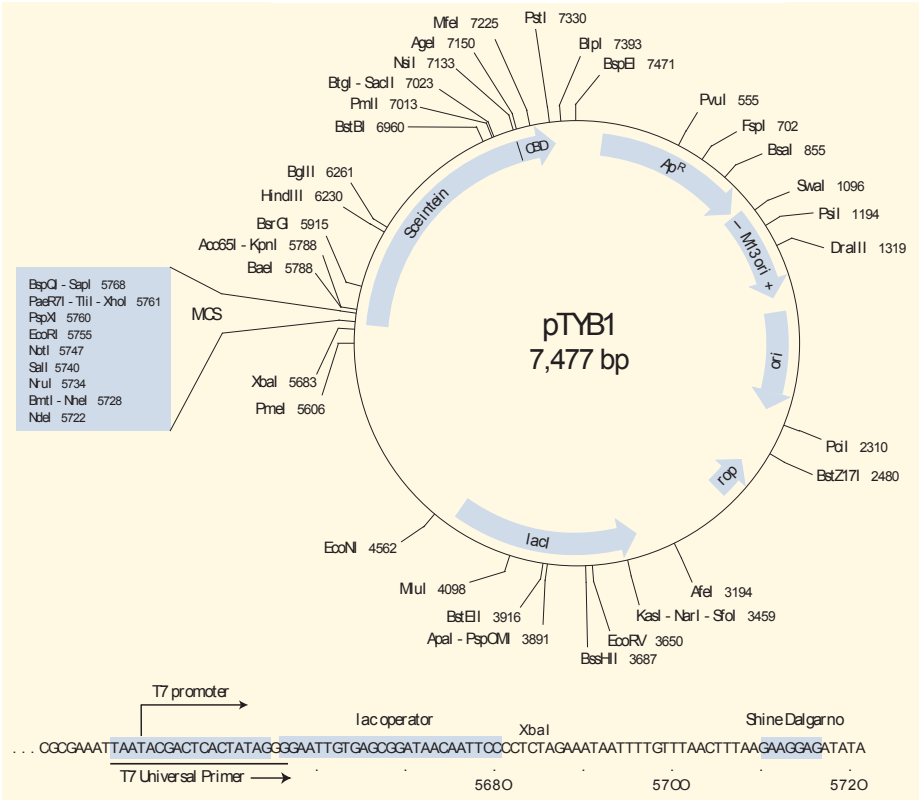


Figure B.1: Gene map for the expression vector pTYB1 (NEB Technical Support, web). Restriction sites for NdeI and XhoI are indicated within the MCS.

Appendix C. Protein Standards Used in SDS-PAGE

C.1 Precision Plus Protein^{TM} Dual Color Standard (Bio-Rad)

Precision Plus Protein^{TM} Dual Color standard (Bio-Rad) was used as standard in SDS-PAGE. The standard is a mixture of 10 recombinant proteins of molecular weights in the range 10–250 kDa with two of the bands stained pink for reference, as shown in Figure C.1. The amount of each recombinant protein is included in the figure as ng/10 µL.

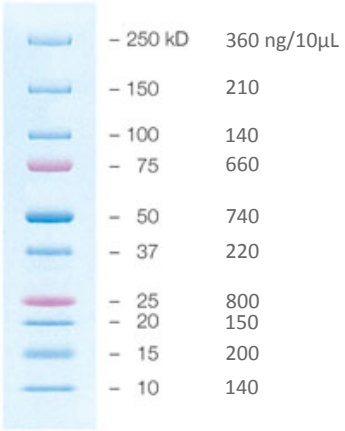


Figure C.1: Molecular weights, in kDa, and concentrations, in ng/10 µL, of the 10 recombinant proteins in Precision Plus Protein^{TM} Dual Color standard (Bio-Rad), with 8 blue-stained bands and two pink-stained reference bands (25 and 75 kDa).

C.2 Amersham Low Molecular Weight Calibration Kit Protein Mixture Standard (GE Healthcare)

Amersham Low Molecular Weight Calibration Kit for SDS Electrophoresis (GE Healthcare) contains a Protein Mixture Standard that also was used as standard in SDS-PAGE. The standard is a mix of six highly purified and well-characterized proteins of molecular weights in the range 14.4–97.0 kDa. Proteins in the standard are listed in Table C.1, along with their molecular weight and concentration.

Table C.1: Molecular weights, in kDa, and concentration, in µg/200 µL, for the six proteins in the Protein Mixture Standard from the Amersham Low Molecular Weight Calibration Kit for SDS Electrophoresis (GE Healthcare).

Protein (source)	**Molecular Weight** [kDa]	**Concentration** [µg/200 µL]
Phosphorylase b (rabbit muscle)	97.0	67
Albumin (bovine serum)	66.0	83
Ovalbumin (chicken egg white)	45.0	147
Carbonic anhydrase (bovine erythrocyte)	30.0	83
Trypsin inhibitor (soybean)	20.1	80
α-Lactalbumin (bovine milk)	14.4	116

Appendix D. Estimation of Protein Concentration by SDS-PAGE

Protein concentration was estimated by relative quantification of protein bands from SDS-PAGE by using the known protein quantity in the standard as reference. Imaging and analysis of gels was performed using the ChemiDoc^{TM} XRS+ system with the Band Analysis tool in Image Lab^{TM} 4.1 (Bio-Rad). Image Lab^{TM} estimates the volume of each band based on its intensity, which is relatively proportional to protein mass present in the band. Amount of protein is known for all bands in the standard, and by dividing the band intensity with the protein mass present, intensity per mass (e.g. Int/ng) is found. This relationship is then used to find the unknown amount of other proteins from the intensities of their bands.

For all samples, several wells on an SDS-PAGE gel were loaded with different volumes of the same sample, and three or five wells were loaded with standard. The amount of protein present in each sample band was calculated relative to several reference bands from the standard, and then an average value was found for the concentration in each sample.

As an example, Figure D.1 presents a gel image with lanes and bands identified in Image Lab, and lane profile for lane 1 is presented in Figure D.2. This highlights the issue with subtraction of lane background, as seen for the first three identified bands in lane 1. If these bands are included in the calculations they will lead to an underestimation of protein concentration, as they are bands from the standard. If it had been sample bands they would have resulted in an overestimation of protein concentration. All bands where lane background was not sufficiently subtracted, such as for these three, were therefore excluded from calculations.

Gel images and the corresponding intensities calculated by the Band Analysis tool in Image Lab are presented for A2 fractions of AlgE6A clones in Figure D.1 and Table D.1, for A3 fractions of AlgE6A clones in Figure D.3 and Table D.2, and for AlgE4 A1 and AlgE4A A1 in Figure D.4 and Table D.3. Amount of sample loaded on each well is included in the tables. Yellow cells in the tables indicate bands without suf-

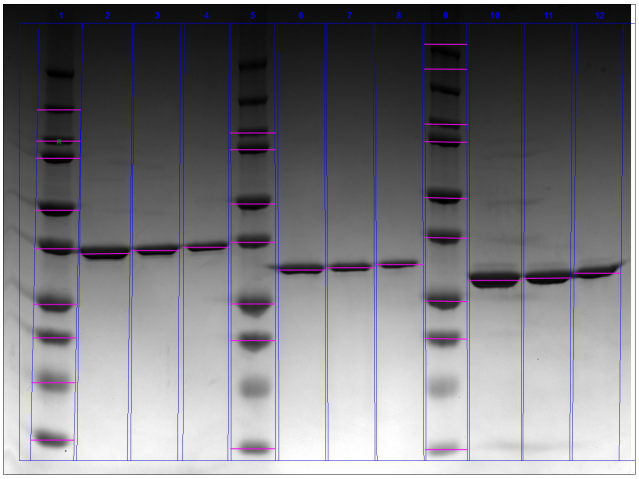
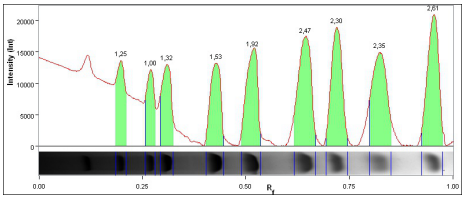


Figure D.1: Gel image from SDS-PAGE to estimate protein concentration in A2 fractions from purification of AlgE6A_272, AlgE6A_295, and AlgE6A_336 by FPLC. Lanes and protein bands analysed by the Band Analysis tool in Image Lab^{TM} 4.1 are indicated. Contents in lanes 1, 5, and 9: Precision Plus Protein^{TM} Dual Color standard (Bio-Rad), Lanes 2-4: AlgE6A_336 A2, Lanes 6-8: AlgE6A_295 A2, Lanes 10-12: AlgE6A_272 A2.

ficient subtraction of lane background, and these were excluded from calculations. As seen in Table D.2, this resulted in no estimation of concentration in AlgE6A_295 A3. Information about concentration of each protein in the standards is included in Appendix C. The results of the calculations are presented in Sections 4.1.2 and 4.2.6, along with molecular weights used to calculate molar concentration.



Band No.	Band Label	Relative Front	Volume (Int)	Rel. Quant.	Band %	Lane %
1		0,199	29 846 706	1,25	7,4	4,3
2		0,270	23 926 408	1,00	6,0	3,5
3		0,309	31 601 727	1,32	7,9	4,6
4		0,428	36 678 804	1,53	9,2	5,3
5		0,516	45 989 446	1,92	11,5	6,7
6		0,643	59 211 904	2,47	14,8	8,6
7		0,720	54 959 909	2,30	13,7	8,0
8		0,822	56 169 693	2,35	14,0	8,2
9		0,953	62 452 480	2,61	15,6	9,1

Figure D.2: Example of a lane profile generated by the Band Analysis tool in Image Lab^{TM} 4.1. Lane profile is for lane 1 on the gel presented in Figure D.1, and presents the intensity of each band in the lane.

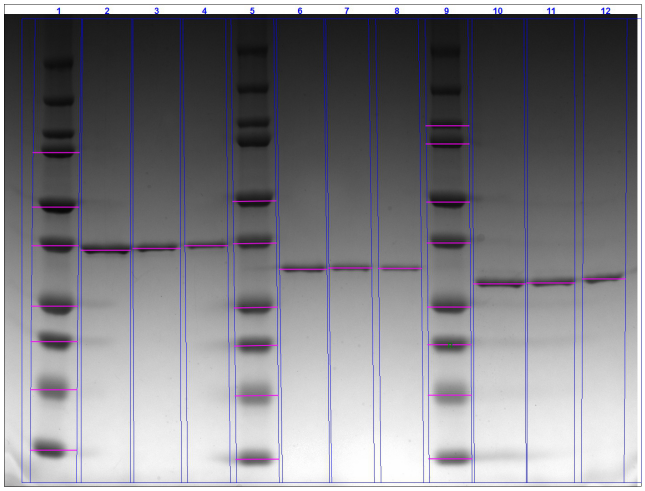


Figure D.3: Gel image from SDS-PAGE to estimate protein concentration in A3 fractions from purification of AlgE6A_272, AlgE6A_295, and AlgE6A_336 by FPLC. Lanes and protein bands analysed by the Band Analysis tool in Image Lab^{TM} 4.1 are indicated. Contents in lanes 1, 5, and 9: Precision Plus Protein^{TM} Dual Color standard (Bio-Rad), Lanes 2-4: AlgE6A_336 A3, Lanes 6-8: AlgE6A_295 A3, Lanes 10-12: AlgE6A_272 A3.

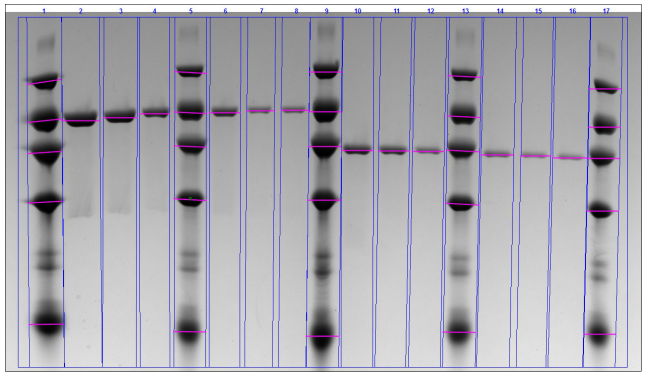


Figure D.4: Gel image from SDS-PAGE to estimate protein concentration in A1 fractions from purification of AlgE4 and AlgE4A by FPLC. Lanes and protein bands analysed by the Band Analysis tool in Image Lab^{TM} 4.1 are indicated. Contents in lanes 1, 5, 9, 13, and 17: Amersham Low Molecular Weight Calibration Kit Protein Mixture standard (GE Healthcare), Lanes 2-4 and 6-8: AlgE4 A1, Lanes 10-12 and 14-16: AlgE4A A1.

Table D.1: Intensities of protein bands on SDS-PAGE gel of A2 fractions for the AlgE6A clones, calculated by the Band Analysis tool in Image Lab^{TM} 4.1. The corresponding gel image is presented in Figure D.1. Yellow cell indicate bands without sufficient subtraction of lane background, and these were excluded from calculations. std: 4 µL Precision Plus Protein^{TM} Dual Color standard (Bio-Rad), 272: AlgE6A_272, 295: AlgE6A_295, 336: AlgE6A_336.

Sample	**Band**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
std		29846706	23926408	31601727	36678804	45989446	59211904	54959909	56169693	62452480
336 5 µL	68845476									
336 3 µL	49605885									
336 2 µL	38610179									
std			26716833	33550011	33684453	35344188	61050429	57824811	40914126	
295 5 µL	53012295									
295 3 µL	42141750									
295 2 µL	102269265									
std			25899105	35596215	34048974	36631863	61877364	53393253	29078682	
272 5 µL	98287488									
272 3 µL	72374925									
272 2 µL	57597176									

Table D.2: Intensities of protein bands on SDS-PAGE gel of A3 fractions for the AlgE6A clones, calculated by the Band Analysis tool in Image Lab^{TM} 4.1. The corresponding gel image is presented in Figure D.3. Yellow cell indicate bands without sufficient subtraction of lane background, and these were excluded from calculations. std: 4 µL Precision Plus Protein^{TM} Dual Color standard (Bio-Rad), 272: AlgE6A_272, 295: AlgE6A_295, 336: AlgE6A_336. Band 2 is excluded from the table as no intensities were estimated for this band number.

Lane	**Band**								
	1	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Std			39 080 076	37 873 314	38 137 902	58 016 988	61 693 068	63 126 984	68 609 892
336 5 µL	43 066 620								
336 3 µL	103 777 232								
336 2 µL	80 870 935								
Std				33 006 800	31 918 765	47 943 935	58 154 155	61 992 535	63 780 150
295 5 µL	101 573 264								
295 3 µL	81 742 275								
295 2 µL	69 333 755								
Std		12 477 015	20 457 680	33 886 690	33 594 945	48 440 595	57 764 180	60 579 885	64 722 930
272 5 µL	35 891 760								
272 3 µL	28 382 130								
272 2 µL	89 887 765								

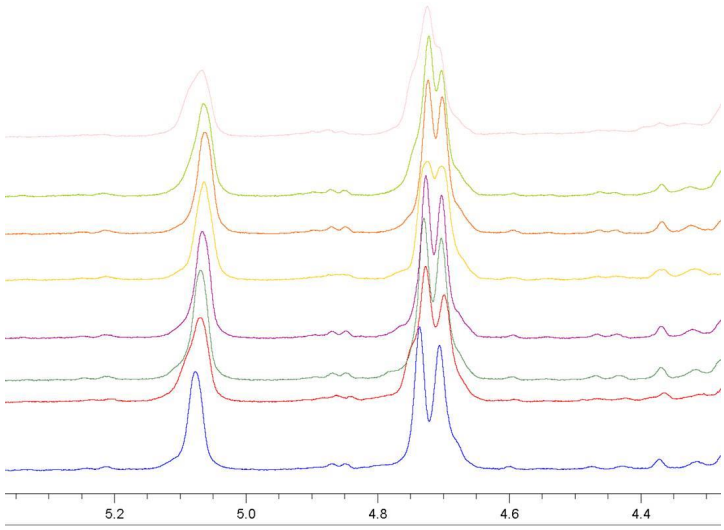
Table D.3: Intensities of protein bands on SDS-PAGE gel of A1 fractions for AlgE4 and AlgE4A, calculated by the Band Analysis tool in Image Lab^{TM} 4.1. The corresponding gel image is presented in **??**. Yellow cell indicate bands without sufficient subtraction of lane background, and these were excluded from calculations. std: Amersham Low Molecular Weight Calibration Kit Protein Mixture standard (GE Healthcare). Band 5 is excluded from the table as no intensities were estimated for this band number.

Sample	**Band**				
	1	**2**	**3**	**4**	**6**
std 7 µL					
AlgE4 3 µL	58183895				
AlgE4 2 µL	44915904				
AlgE4 1 µL	25148958				
std 4 µL	32248320	50819145	47800480	48000940	89332360
AlgE4 1 µL	23017560				
AlgE4 0.5 µL	10268328				
AlgE4 0.5 µL	26435864				
std 7 µL	42848443	58420382	56825447	50415557	99984502
AlgE4A 3 µL	25673856				
AlgE4A 2 µL	22337154				
AlgE4A 1 µL	11638912				
std 4 µL	32027375	51196643	48568903	50277253	32267040
AlgE4A 1 µL	29844145				
AlgE4A 0.5 µL	21712597				
AlgE4A 0.5 µL	23019691				
std 3 µL	29200640	44384512	44159872	46091008	69357053

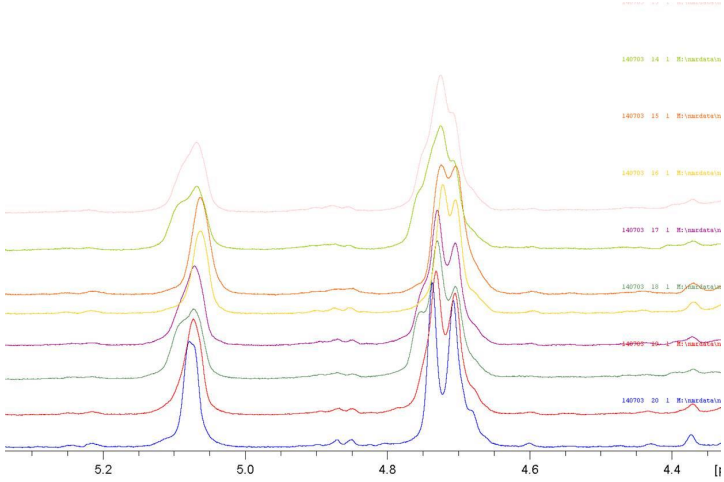
Appendix E. ^{1}H-NMR Spectra of Poly-M Epimerized with AlgE4 and AlgE4A and Variation in Calcium Concentration

Poly-M was epimerized with AlgE4 and AlgE4A at varying Ca^{2+} concentrations in the range 1–8 mM, and the product was analysed by ^{1}H-NMR. The results are presented in Figure E.1 as superimposed ^{1}H-NMR spectra.

A second epimerization was performed with less enzyme and shorter incubation time, with Ca^{2+} concentrations in the range 0.1–8 mM. The resulting ^{1}H-NMR spectra are presented for AlgE4 in Figures E.2 and E.3, and for AlgE4A in Figures E.4 and E.5.

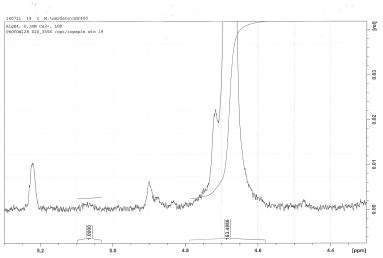


(a)

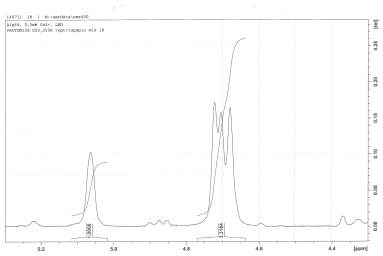


(b)

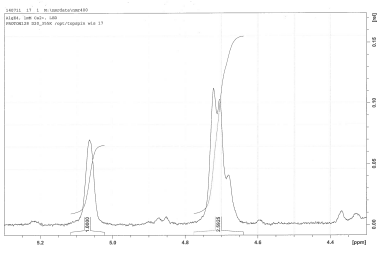
Figure E.1: Superimposed ^{1}H-NMR spectra of poly-M epimerized with (a) AlgE4 and (b) AlgE4A at varying Ca^{2+} concentrations in the range 1–8 mM. In both figures, the top spectre is poly-M epimerized at 1 mM Ca^{2+}, and the Ca^{2+} concentration increases in increments of 1 mM for each spectre towards the bottom. G-residues give rise to a signal at the position of the first peak in each spectre, and M-residues give rise to a signal at the position of the second peak. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.



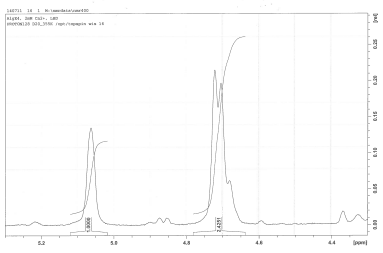
(a) AlgE4 - 0.1 mM Ca^{2+}



(b) AlgE4 - 0.5 mM Ca^{2+}

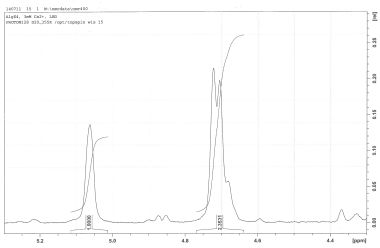


(c) AlgE4 - 1 mM Ca^{2+}

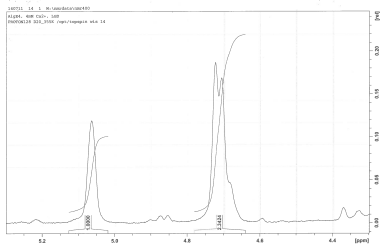


(d) AlgE4 - 2 mM Ca^{2+}

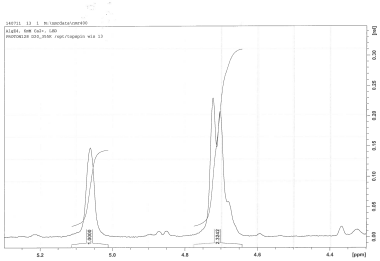
Figure E.2: Integrated ^{1}H-NMR spectra of poly-M epimerized with AlgE4 at Ca^{2+} concentrations in the range 0.1–2 mM. G-residues give rise to a signal at the position of the first integrated peak in each spectre, and M-residues give rise to a signal at the position of the second integrated peak. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.



(a) AlgE4 - 3 mM Ca^{2+}

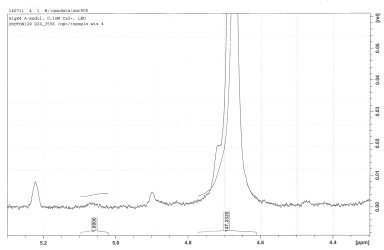


(b) AlgE4 - 4 mM Ca^{2+}

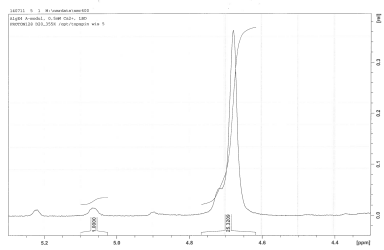


(c) AlgE4 - 6 mM Ca^{2+}

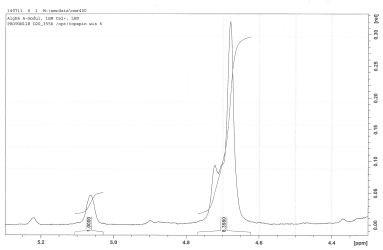
Figure E.3: Integrated ^{1}H-NMR spectra of poly-M epimerized with AlgE4 at Ca^{2+} concentrations in the range 3–6 mM. G-residues give rise to a signal at the position of the first integrated peak in each spectre, and M-residues give rise to a signal at the position of the second integrated peak. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.



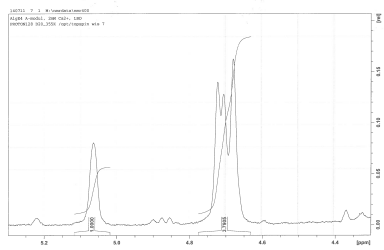
(a) AlgE4A - 0.1 mM Ca^{2+}



(b) AlgE4A - 0.5 mM Ca^{2+}

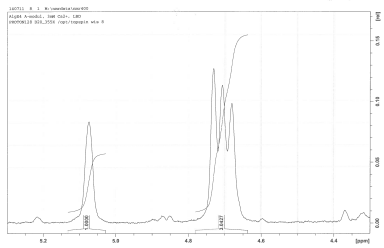


(c) AlgE4A - 1 mM Ca^{2+}

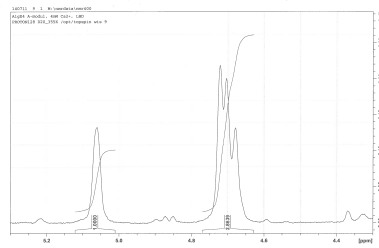


(d) AlgE4A - 2 mM Ca^{2+}

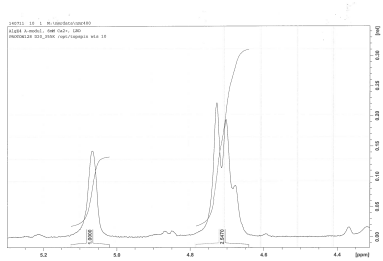
Figure E.4: Integrated ^{1}H-NMR spectra of poly-M epimerized with AlgE4A at Ca^{2+} concentrations in the range 0.1–2 mM. G-residues give rise to a signal at the position of the first integrated peak in each spectre, and M-residues give rise to a signal at the position of the second integrated peak. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.



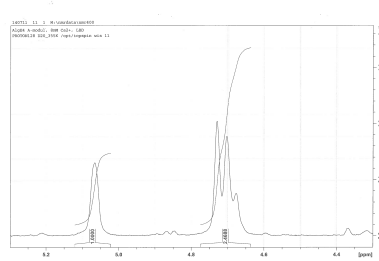
(a) AlgE4A - 3 mM Ca^{2+}



(b) AlgE4A - 4 mM Ca^{2+}



(c) AlgE4A - 6 mM Ca^{2+}

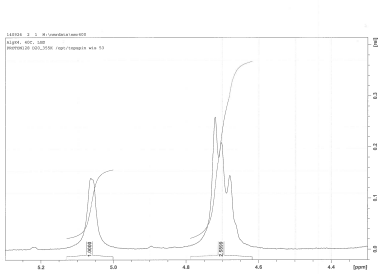


(d) AlgE4A - 8 mM Ca^{2+}

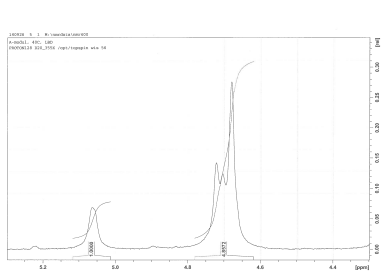
Figure E.5: Integrated ^{1}H-NMR spectra of poly-M epimerized with AlgE4A at Ca^{2+} concentrations in the range 3–8 mM. G-residues give rise to a signal at the position of the first integrated peak in each spectre, and M-residues give rise to a signal at the position of the second integrated peak. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.

Appendix F. ^{1}H-NMR Spectra of Poly-M Epimerized with AlgE4 and AlgE4A at Different Temperatures

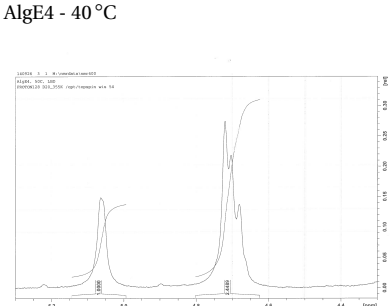
Epimerization of poly-M with AlgE4 and AlgE4A was performed at 40 °C, 50 °C, and 60 °C, and the product was analysed by ^{1}H-NMR. The resulting ^{1}H-NMR spectra are included in Figure F.1.



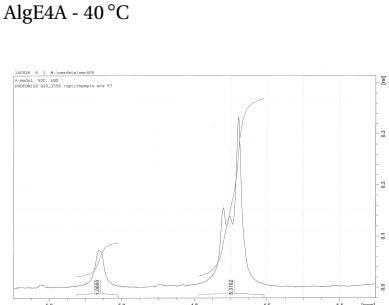
(a) AlgE4 - 40 °C



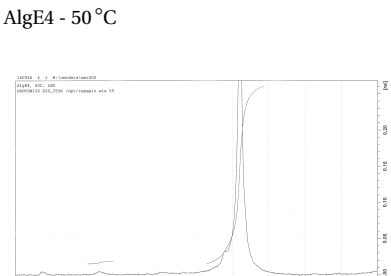
(b) AlgE4A - 40 °C



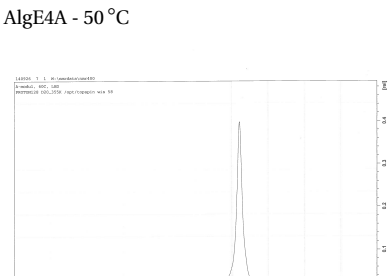
(c) AlgE4 - 50 °C



(d) AlgE4A - 50 °C



(e) AlgE4 - 60 °C



(f) AlgE4A - 60 °C

Figure F.1: Integrated ^{1}H-NMR spectra of poly-M epimerized by AlgE4 and AlgE4A at 40 °C, 50 °C, and 60 °C. G-residues give rise to a signal at the position of the first integrated peak in each spectre, and M-residues give rise to a signal at the position of the second integrated peak. Spectra were recorded on a BRUKER Avance DPX 400 MHz with 5 mm z-grad DUL probe at 95 °C.

Appendix G. Epimerase Activity Microassays of AlgE6A_272 and AlgE6A_295

Epimerase activity in fractions A1, A2, and A3, was tested for all three AlgE6A clones in a microassay based on the activity of G-lyase as described in Section 3.2.12. The difference in absorbance at 230 nm ($\Delta A_{230\text{nm}}$) before and after incubation with G-lyase was used to estimate epimerase activity. The result for AlgE6A_336 was presented in Section 4.2.7, while the results for AlgE6A_272 and AlgE6A_295 are presented in Table G.1 and Table G.2, respectively.

Table G.1: Result of epimerase activity microassay for AlgE6A_272, fractions A1-A3, presented as the difference in absorbance at 230 nm ($\Delta A_{230\text{nm}}$) before and after incubation with G-lyase. Std: AlgE6 standard used as reference, (ref.): reference wells without enzyme, A1-A3: fractions A1-A3 from purification by FPLC. Grey cells indicate wells containing the AlgE6 standard. Yellow cells indicate wells with the highest concentration of AlgE6A_272 A1-A3. The first column indicates level of enzyme dilution for each row.

	AlgE6A_272							
Dilution	**Std**	**A1**	**A1**	**A2**	**Std**	**A2**	**A3**	**A3**
(ref.)	0.055	0.065	0.063	0.039	0.072	0.063	0.066	0.060
2x	1.665	0.035	0.030	0.028	1.578	0.041	0.069	0.058
4x	1.460	0.047	0.050	0.046	1.447	0.061	0.056	0.055
8x	1.362	0.053	0.053	0.051	1.318	0.055	0.067	0.065
16x	1.024	0.054	0.049	0.067	0.961	0.054	0.075	0.063
32x	0.573	0.059	0.055	0.063	0.494	0.071	0.047	0.064
64x	0.238	0.053	0.065	0.070	0.196	0.046	0.036	0.037
128x	0.098	0.027	0.033	0.046	0.090	0.033	0.037	0.015
256x	0.057	0.027	0.021	0.036	0.055	0.044	0.027	0.033
512x	0.033	0.021	0.032	0.040	0.036	0.017	0.018	0.024
1024x	0.035	0.025	0.022	0.037	0.038	0.027	0.025	0.028
2048x	0.055	0.045	0.053	0.038	0.030	0.033	0.037	0.036

Table G.2: Result of epimerase activity microassay for AlgE6A_295, fractions A1-A3, presented as the difference in absorbance at 230 nm ($\Delta A_{230\text{nm}}$) before and after incubation with G-lyase. Std: AlgE6 standard used as reference, (ref.): reference wells without enzyme, A1-A3: fractions A1-A3 from purification by FPLC. Grey cells indicate wells containing the AlgE6 standard. Yellow cells indicate wells with the highest concentration of AlgE6A_295 A1-A3. The first column indicates level of enzyme dilution for each row.

				AlgE6A_295				
Dilution	**Std**	**A1**	**A1**	**A2**	**Std**	**A2**	**A3**	**A3**
(ref.)	0.010	0.016	0.002	0.007	0.023	0.000	0.006	0.004
2x	1.788	0.002	0.008	0.024	1.863	0.030	0.048	0.054
4x	1.688	0.022	0.029	0.044	1.727	0.047	0.067	0.060
8x	1.569	0.057	0.047	0.045	1.600	0.050	0.062	0.063
16x	0.942	0.072	0.057	0.037	0.974	0.051	0.062	0.060
32x	0.416	0.069	0.049	0.049	0.434	0.049	0.069	0.061
64x	0.174	0.079	0.073	0.049	0.174	0.062	0.064	0.057
128x	0.108	0.049	0.057	0.046	0.086	0.046	0.062	0.048
256x	0.069	0.054	0.051	0.069	0.069	0.038	0.054	0.043
512x	0.052	0.048	0.054	0.041	0.072	0.046	0.041	0.042
1024x	0.043	0.037	0.034	0.043	0.042	0.035	0.035	0.037
2048x	0.029	0.031	0.047	0.050	0.033	0.039	0.035	0.043