Juuso Taskinen

Construction of cloning vectors for Spo20 mediated membrane anchoring of hyaluronic acid metabolic pathway enzymes in *Pichia pastoris*

Helsinki Metropolia University of Applied Sciences

Bachelor of Engineering

Biotechnology and Food Engineering

Bachelor's Thesis

30.5.2017

Metropolia

Author(s) Title Number of Pages Date	Juuso Taskinen Construction of cloning vectors for Spo20 mediated membrane anchoring of hyaluronic acid metabolic pathway enzymes in *Pichia pastoris* 31 pages + 4 appendices 30.5.2017
Degree	Bachelor of Engineering
Degree Programme	Biotechnology and Food engineering
Instructors	Outi Koivistoinen, Senior Scientist Tiina Soininen, Senior Lecturer

This thesis was done in collaboration with VTT's production host engineering research team, and it was part of the academy of Finland funded SynBioPol project. Aim of the SynBioPol project was to investigate different types of scaffolding or anchoring methods to improve the metabolic flux of the hyaluronic acid pathway

Hyaluronic acid is a high molecular weight polymer which consists of repeating units of disaccharides glucuronic acid and N-actylglycosamine linked by β1,3- and β1,4-glycosidic bonds. Hyaluronic acid is synthesized by the transmembrane protein hyaluronan synthase, which is a common enzyme in mammals and amphibians.

The aim of this thesis was to investigate whether or not the *Saccharomyces cerevisiae* sporulation-specific protein 20 could be used as a membrane anchor to improve the metabolic flux of the hyaluronic acid metabolic pathway, by fixing the metabolic enzymes to cell membranes. The localization of the metabolic pathway enzymes was to be confirmed by creating a chimera protein with the sporulation specific protein 20 linked to one of the terminals of the metabolic enzyme and a fluorescent protein linked to the opposite terminal.

Due to time limitations not all goals set for this thesis were met and only the *xhasA2* and *xhasB* genes were inserted into the expression vector pGAPZB. The correct cloning vectors were screened with colony PCR, restriction digestion analysis and finally sequenced to verify the composition of these cloning vectors. The correct cloning vectors were linearized with Bgl-II, transformed into *Pichia pastoris* with electroporation and then screened with colony PCR. On the basis of the colony PCR results, a handful of potential recombinant strains were produced.

Keywords	hyaluronic acid, membrane anchoring, Spo20, *Pichia pastoris*

Tekijä(t) Otsikko Sivumäärä Aika	Juuso Taskinen Kloonausvektorien rakentaminen hyaluronihapon metaboliaentsyymien ankkuroimiseksi solun kalvoille Spo20-proteinin avulla 31 sivua + 4 liitettä 30.5.2017
Tutkinto	Insinööri (AMK)
Koulutusohjelma	Bio- ja elintarviketekniikka
Ohjaajat	Erikoistutkija, Outi Koivistoinen Lehtori, Tiina Soininen

Tämä insinöörityö tehtiin Teknologian tutkimuskeskuksen VTT Oy:n production host engineering -tutkimusryhmälle. Työ tehtiin osana Suomen Akatemian rahoittamaan SynBioPol-projektia, jonka tavoitteena on tutkia erilaisia skaffolding- ja ankkurointitekniikoita metaboliareittien, kuten, hyaluronihapon saantojen parantamiseksi.

Hyaluronihappo on molekyylipainoltaan suurin polymeeri, joka rakentuu glukuronihaposta ja N-asetyyliglukosamiinista, jotka ovat linkittyneet toisiinsa β1,3- ja β1,4-glykosidisidoksin. Hyaluronihappoa syntetisoi transmembraaniproteiini hyaluronihapposyntaasi. Hyaluronihapposyntaasi on yleinen proteiini sekä nisäkkäiden että sammakkoeläimien solukalvoilla.

Tämän insinöörityön tarkoituksena oli tutkia, mikäli *Saccharomyces cerevisiaen* itiöintispesifistä proteiini 20:tä voitaisiin käyttää ankkuroimaan hyaluronihapon metaboliaentsyymejä solun kalvoille. Metaboliaentsyymien lokalisaatio solussa oli tarkoitus tarkistaa rakentamalla fuusioproteiini, jossa metaboliaentsyymin terminaaleihin olisi kiinnitetty itiöintispesifinen proteiini 20 ja vastakkaiseen terminaaliin fluoresoiva proteiini.

Tämä insinöörityö ei saavuttanut kaikkia sille asetettuja tavoitteita, ja vain ekspressiovektorit, johon geenit *xhasA2* ja *xhasB* oli liitetty, ehdittiin rakentaa. Kloonatut plasmidit seulottiin pesäke-PCR:llä, jonka jälkeen niiden rakenne varmistettiin restriktioentsyymianalyysilla sekä sekvensoinnilla. Oikein rakentuneet vektorit linearisoitiin Bgl-II-entsyymillä, jonka jälkeen ne transformoitiin *Pichia pastorikseen* elektroporaatiolla. Transformantit seulottiin pesäke-PCR:llä, jonka avulla löydettiin muutama potentiaalinen rekombinantti *Pichia pastoris* -kanta.

Avainsanat	Hyaluronihappo, solukalvo ankkurointi, Spo20, *Pichia pastoris*

Table of Contents

XhasA2 and *xhasB* colony PCR results

Restriction analysis results for the pGAPZA-insert vectors

Pichia pastoris colony PCR results with the Bgl-II linearized vectors

Abbreviations

FP	Fluorescent protein
gDNA	genomic DNA
GFP	Green fluorescent protein
GlcA	Glucuronic acid
GlcNAc	N-actylglycosamine
GRAS	Generally regarded as safe (microbes)
HA	Hyaluronic acid
HAS	Hyaluronan synthase
MDa	Mega Dalton
MW	Molecular weight
RPM	Revolutions per minute
SNAP-25	Synaptosome-associated protein
SNARE	Soluble NSF attachment protein receptors
Spo20	*Saccharomyces cerevisiae* sporulation specific protein 20
UDP	Uridine diphosphate
VAMP	Vesicle associated membrane protein

1 Introduction

The aim of this thesis was to investigate whether the *Saccharomyces cerevisiae* sporulation specific protein 20 could be used to anchor hyaluronic acid pathway enzymes to the plasma membrane of *Pichia pastoris*, essentially creating a substrate channel for more efficient hyaluronic acid production. Hyaluronic acid pathway enzymes are natively scattered in the cytosol; therefore, many of the precursors needed to produce high quality hyaluronic acid are depleted before the transmembrane protein hyaluronan synthase can turn them into HA. This inefficient precursor utilisation leads to lower molecular weight and lower production titers of hyaluronic acid in recombinant production hosts.

Hyaluronic acid is an interesting molecule due to its versatile physical and biochemical properties as well as high immunocompatibility. Due to these properties, hyaluronic acid has already been used in numerous medical, cosmetic and pharmaceutical applications. The market value of hyaluronic acid was estimated to be 5.32 billion USD in 2012 and is expected to grow up to 9.85 billion USD by 2019 (De Oliveria *et.al.,* 2016).

This thesis was made for VTT's production host engineering research team as part of the Academy of Finland funded SynBioPol project. VTT's production host engineering team provides novel genetic tools accompanied by bioinformatics to efficiently engineer microbes for the production of useful chemicals and biofuels. VTT Technical Research Centre of Finland Ltd is a research and technology company that provides technical and innovative solutions for both the private and the public sector.

This thesis will first discuss the general information about hyaluronic acid and the mechanism behind its production, SNARE proteins, sporulation specific protein 20, briefly discuss fluorescent proteins and their application in live cell imaging and the genetic engineering tools for *Pichia pastoris*. Then, it will move on to highlight the used materials, molecular biology and microbiology methods used in this thesis, to and conclude with the results and the discussion.

1.1 Hyaluronic acid

Hyaluronic acid (HA) (Figure 1) is a linear and rigid biopolymer that belongs to the class of glycosaminoglycans. HA consists of repeating units of disaccharides glucuronic acid (GlcA) and N-actylglycosamine (GlcNAc) linked by β1,3- and β1,4-glycosidic bonds. HA is produced by an transmembrane enzyme called hyaluronan synthase (HAS) that can be found from the plasma membrane of mammal, amphibian and bacterial cells. HA is present in all human organs, and it is highly abundant in connective tissues such as cartilage tissue and responsible for the maintenance of tissue homeostasis (De Oliveria *et.al.,* 2016). HA is also directly involved in several other processes like embryogenesis, inflammation, metastasis, angiogenesis and wound healing. HA is also responsible for the smooth appearance of skin and with aging HA production decreases which results in wrinkles (De Oliveria *et.al.,* 2016).

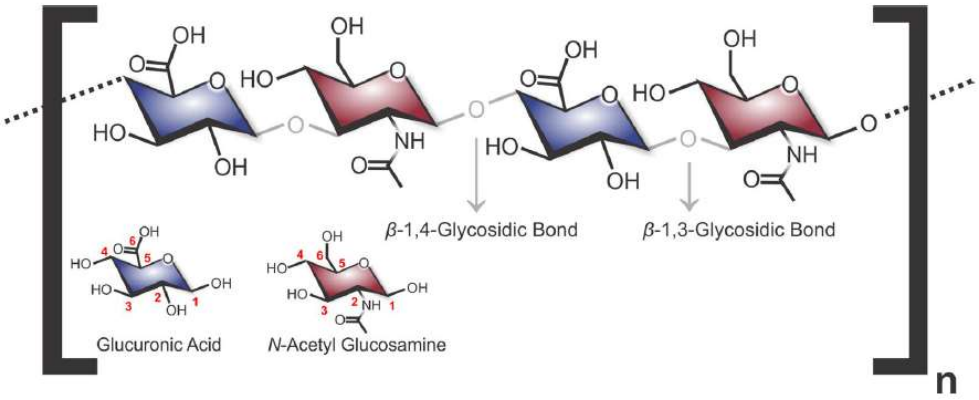

Figure 1. Structure of HA showing the β-1,3 and β-1,4 glycosidic bonds and the precursors glucuronic acid and N-actylglycosamine (De Oliveria *et.al.,* 2016)

1.1.1 Production of HA

Traditionally HA has been extracted chemically from animal tissues like rooster combs, bovine eyes and umbilical cords. This approach has many disadvantages and it can lead to degradation of HA, due to endogenous hyaluronidase enzyme activity and harsh extraction conditions. This method also has high purification cost due to concerns over infectious zoonotic disease and allergic reactions. Today HA is produced mainly through microbial fermentation although this method produces lower molecular weight (MW) HA than what can be obtained from animal tissues (De Oliveria *et.al.,* 2016; Joeng *et.al.,* 2014).

Most naturally occurring microbes that produce HA are pathogenic bacteria such as *Pasteurella multocida* and gram-positive *Streptococcus* group A and C. Some eukaryotes like the yeast *Cryptococcus neoformans* or an algae *Chlorella*.sp can produce HA natively although in the latter case this only happens when *Chorella.*sp is infected with the *Paramecium bursaria* chlorella virus (De Oliveria *et.al.,* 2016). The first wild type microbe used in industrial HA production was *Streptococcus zooepidemicus* in the early 1980s. Although *Streptococcus zooepidemicus* can produce a high molecular weight HA, it also produces potentially toxic and pathogenic side products that need to be purified, raising the purification cost (De Oliveria *et.al.,* 2016). Higher purification cost associated with pathogenic bacteria fermentations has led to the development of several recombinant generally regarded as safe (GRAS) organisms that can produce high yields of HA, with sufficient MW for some medical applications (De Oliveria *et.al.,* 2016; Joeng *et.al.,* 2014). Most recombinant host producing HA are GRAS bacteria that have already been used in industrial application for years. These include bacteria such as*, Lactococcus lactis*, *Corybacterium glutamicum* and *Bacillus subtilis.* Only a few eukaryotic organisms have been genetically engineered to produce HA. These eukaryotes include yeasts like *Saccharomyces cerevisiae*, *Pichia pastoris* and plant cell cultures of *Nicotiana benthamiana* (De Oliveria *et.al.,* 2016; Joeng *et.al.,* 2014).

1.1.2 HA metabolism

The HA metabolic pathway originates from the glycolytic pathway. The precursors for HA are synthesized from deviations of glucose-6-phospate and fructose-6-phosphate by multiple enzymes inside the cytoplasm (Figure 2) (De Oliveria *et.al.,* 2016). The production of HA by the HAS enzyme competes for the HA pathway precursor with other cellular mechanisms like cell wall polysaccharide synthesis, teichoic acid and peptidoglycan synthesise as well as glycolysis (De Oliveria *et.al.,* 2016). This competition between these often essential cellular mechanisms, can greatly reduce the availability of precursors for the HAS enzyme and therefore lead to reduction in HA production and molecular weight (MW) (De Oliveria *et.al.,* 2016). Most of these enzymes are coded by the production hosts native genes that are called *hasC*, *hasB* and *hasE* in this thesis. Since many industrial yeast strains cannot produce the HAS enzyme on their own and genes from other eukaryotes must be transformed into these yeasts in ordered for them to produce HA. These recombinant genes are called *xhasA2* and *xhasB* in this thesis, and they are genes from *Xenopus laevis* (Joeng *et.al.,* 2014).

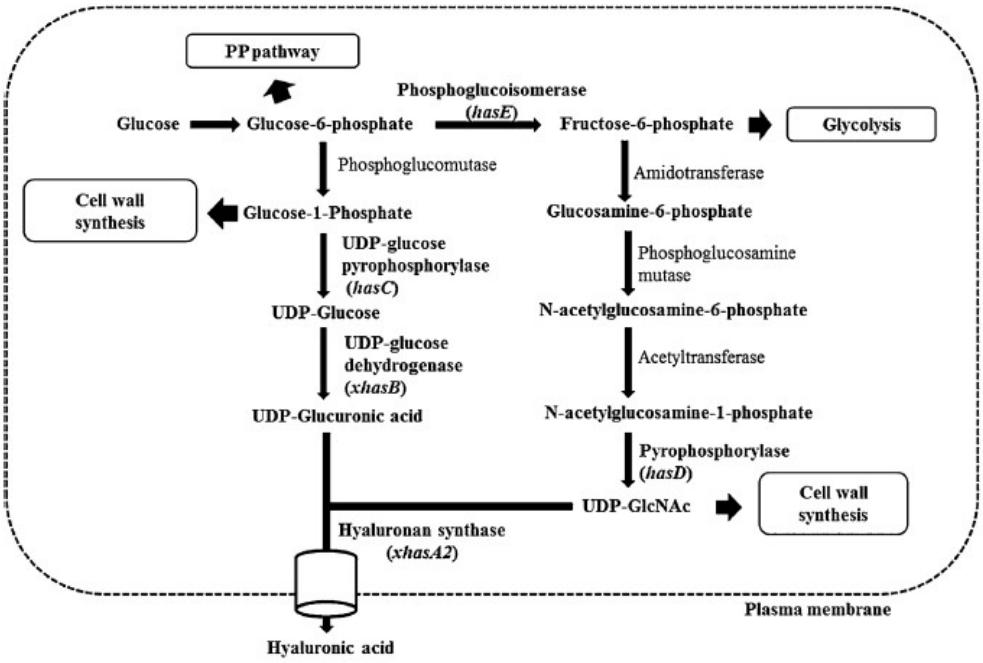

Figure 2. HA pathway in a recombinant *Pichia pastoris. XhasA2* and *xhasB* genes coding the HAS enzyme and UPD-glucose dehydrogenase are transformed from *Xenopus laevis* since *Pichia pastoris* does not produce the HAS enzyme natively (Joeng E *et.al.,* 2014).

The HAS enzyme (*xhasA2*) is a transmembrane protein that catalyses the polymerisation of UDP-GlcA and UDP-GlcNAc intracellularly to form HA that is then secreted to extracellular matrix. The HAS enzyme is a multidomain enzyme that contains various binding sites (De Oliveria *et.al.,* 2016) (Figure 3):

- (1 – 2) UDP-GlcA- and UDP-GlcNAc- binding sites that capture the precursors
- (3 – 4) UDP-GlcNAc β1,4- and UPD-GlcA β1,3-transferase domains that bind the precursors together
- (5 – 6) HA acceptor site that receives the polymer and the excretion site that excretes the polymer out of the cell.

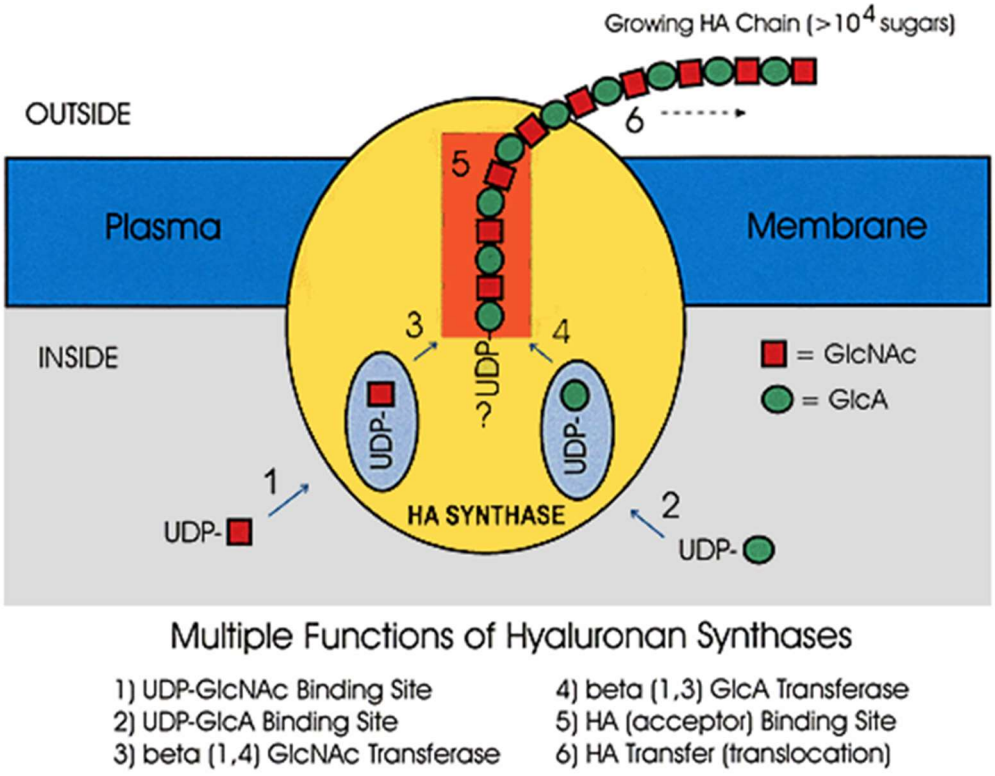

Figure 3. Proposed structure of the HAS enzyme and its various binding domains (Weigel, 2004)

1.1.3 Applications of HA

HA with sufficient MW of more than 2 MDa has a multitude of industrial and medical uses, due to its versatile physiochemical properties such as high hygroscopicity, viscoelasticity, high biocompatibility, non-immunogenicity and non-toxic degradation products (De Oliveria *et.al.,* 2016; Kogan *et.al.,* 2007). The properties of HA are highly dependent on its MW and HA molecules with small MW often have opposing biological function compared to its high molecular weight counterparts (De Oliveria *et.al.,* 2016; Kogan *et.al.,* 2007). The medical industry has a wide range of surgical uses for HA (Kogan *et.al.,* 2007) and these aplications include:

- Viscosurgery where HA is used to protect and provide space between delicate tissues during ophtmalogical surgeries
- Viscoaugmentation where HA is used to fill and augment tissue spaces such as skin, muscles and vocal tissues
- Viscoseparation where HA is used to separate injured or traumatized connective tissue to prevent excessive scaring and adhesion

- Viscosupplementation where HA is used to replace or supplement of tissue fluids
- Viscoprotection where HA acts as a barrier for injured tissues protecting them from dryness, environmental agents and it promotes wound healing

The cosmetic industry also has a variety of uses for HA, and it can be found in many over the counter beauty products, boasting claims such as:

> "Proven anti-aging skincare properties fill-in and diminish the look of fine lines and wrinkles. See a noticeable difference with visibly plumper, softer and smoother skin with this anti-aging hyaluronic acid serum (Amazon, 2017) "

Unfortunately claims such as these are false because HA cannot penetrate the skin due to its high molecular weight (Boss and Meinardi, 2000), but HA can be used as dermal filler if it is injected directly under the skin, for example, to smoothen wrinkles (Kogan *et.al.,* 2007). The drug industry has used HA for drug delivery, and it has been used to make microcapsules and microspheres to deliver compounds, such as plasmids and monoclonal antibodies (Kogan *et.al.,* 2007).

1.2 SNARE proteins as membrane anchors

Soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptors (SNARE proteins) are responsible for vesicle fusion to the plasma membrane and other intracellular membranes. SNAREs can be divided into two groups; v-SNAREs that primarily reside in vesicles and mediate vesicle fusion and t-SNAREs that reside on plasma- and other membranes (Liu S *et.al.,* 2007). Three examples of distinct SNARE proteins can be found in the neuronal SNARE complex where they mediate vesicle fusion to the presynaptic membrane (Dunman and Forte, 2003). These proteins are called VAMP (vesicle associated membrane protein), SNAP-25 (synaptosome-associated protein) and syntaxin 1A. VAMP (sometimes referred to as synaptobrevin) is a single pass transmembrane protein that can be considered as a v-SNARE since it is localized to vesicle membranes. SNAP-25 on the other hand is regarded as a t-SNARE protein and it has two SNARE domains which associate it to the plasma membrane. Also, considers as a t-SNARE protein, syntaxin 1A binds to the plasma membrane with only one transmembrane domain (Dunman and Forte, 2003).

The prevailing theory how the aforementioned proteins form the SNARE complex and mediate membrane fusion is called *zippering*, although this so-called SNARE hypothesis is still debated (Chen Yu *et.al.,* 2001). The SNARE hypothesis suggests that the SNARE proteins bind together in such a way, that v-SNARE starts to coil around the t-SNARE proteins N-terminus towards the C-terminus, pulling the vesicle and the plasma membrane closer to each other and at the same time providing the necessary lateral tension to fuse the membranes together as seen in Figure 4 (Liu *et.al.,* 2007; Dunman and Forte, 2003; Chen Yu *et.al.,* 2001).

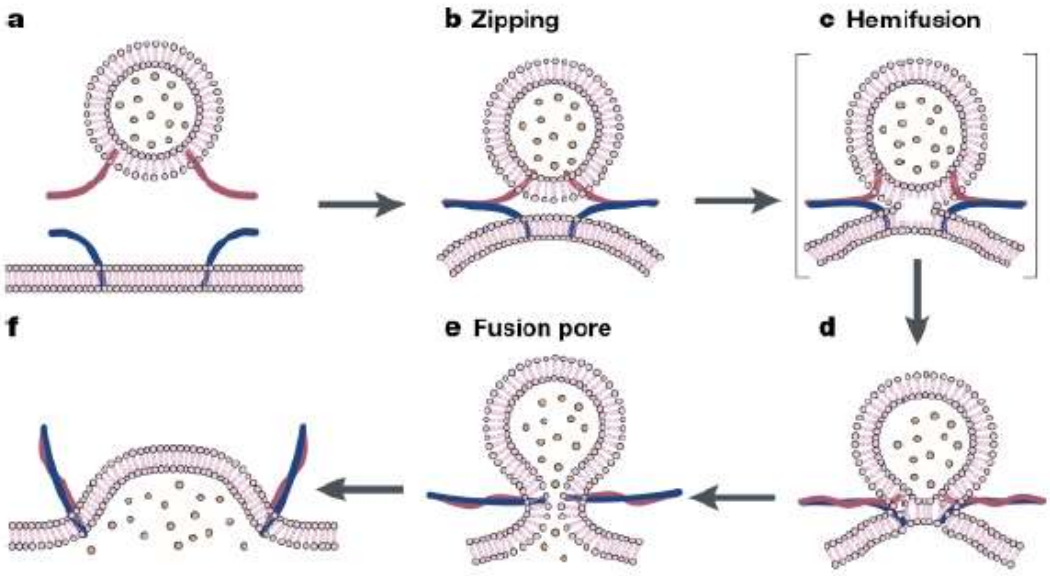

Figure 4. Proposed model of the SNARE hypothesis. A: SNARE proteins not yet connected. B: SNARE complex starts *zippering* from N-terminus towards C-terminus. C: Further *zippering* causes lateral tension and exposes bilayer interior. D: Void space in the membrane junction causes the layers to fuse. E: Lateral tension in the bilayer contact area causes membrane breakdown, forming a fusion pore. F: Fusion pore expands and membrane relaxes (Chen Yu A, 2001).

All the aforementioned SNARE proteins have similar homologs in other eukaryotic organisms like yeasts.

1.3 *Saccharomyces cerevisiae* sporulation specific protein 20 (Spo20)

Spo20 is a SNAP25 family protein, and, in *Saccharomyces cerevisiae,* it is an essential factor in vesicle fusion into the forespore membrane. Spo20 is also critical to the formation of the forespore membrane and *spo20*$^{+}$ gene deficient yeasts are not capable of

forming spores or viable cells (Nakase *et.al.,* 2001; Nakanishi et.al. 2004; Neiman *et.al.,* 2000) Spo20 translocates to the forespore membrane on the completion, of meiosis and, has been reported to localize to the nucleus when cells are nitrogen starved, furthermore this nucleus localization persist through out conjugation and meiosis (Nakase *et.al.,* 2001; Nakanishi *et.al.,* 2004; Neiman *et.al.,* 2000)

In a 2000 study by Neiman, et al (Neiman *et.al.,* 2000) provided proof for a regulatory N-terminal region in Spo20 that is required for successful function of the protein during sporulation. A later 2004 study by Nakanishi, et al (Nakanishi *et.al.,* 2004), suggests that Spo20 has an amphipathic helix close to its N-terminal that binds to acidic phospholipids and the most likely ligand for this helix is phosphatidic acid (Nakanishi *et.al.,* 2004). Since the discovery of this amphipathic region of the Spo20, many studies have used this region as a phosphatidic acid biosensor for numerous applications (Liu *et.al.,* 2007; Ferraz-Nogueira et.al, 2014; Horchani et.al, 2014). A 2014 study by Horachin et al (Horchani et.al, 2014), suggested that the amphipathic helix does not bind specifically to phosphatidic acid alone, but rather it has non-specific interactions with membranes containing anionic lipids.

Due to its ability to bind to many of the yeast cells membranes that contain anionic lipids, the amphipathic region of the Spo20 seems a suitable candidate to investigate whether or not it can be used as a part of a chimera protein to anchor the HA pathway enzymes to cell membranes. This membrane anchoring would bring all the necessary enzymes and precursors closer to the HAS enzyme that is already bound to the plasma membrane, essentially creating a more efficient substrate channel and therefore raising MW and production titers.

1.4 Fluorescent proteins and their applications in live cell imaging

The Fluorescent proteins (FP) are proteins, which fluoresces when excited with UV-light. Excitation of the chromophores electrons forces them to jump to a higher orbital and when these electrons decay back to their original orbitals, the released energy is emitted as light, in wavelength unique to that chromophore. The light emitted can be detected by photo sensors and this technique is used, for example, in immunofluorescence assays.

Fluorescent proteins are highly conserved in size and all the green fluorescent protein family proteins (GFP) whether they are genetically manipulated or from different species, are 25 KDa in size (Kremers *et.al.,* 2011). All FPs also share highly conserved amino acids and without them the central chromophore does not form. Some of the examples of these conserved amino acids are glycines 20/33/67/191, tyrosine 66 and phenylalanine 130 (Kremers J *et.al.,* 2011). FPs and GFPs form a barrel like molecule from eleven β-sheets that surround the central α-helix. This protein structure protects the α-helix and the chromophore attached to it from solvents, thus the molecule is very stable (Chudakov *et.al.,* 2010, p1105; Kremers *et.al.,* 2011). Although all FPs have many conserved amino acids, the rest of the amino acid chain can be modified to produce a plethora of FPs with different emission colours and physical properties (Kermes *et.al.,* 2011)

FPs can be used in numerous applications ranging from DNA, RNA, protein, cell and tissue labelling to promoter tracking and studying protein-protein interactions (Chudakov *et.al.,* 2010, p1129-1137). In this thesis FPs are used as protein labels to track the localization of the HA metabolic pathways enzyme Spo20 chimera protein in the cell.

1.5 Recombinant protein production in *Pichia pastoris*

Pichia pastoris is a methylotrophic yeast that can use methanol as its sole carbon source. It has two strong methanol inducible alcohol oxidase promotors AOX1 and AOX2 that are repressed by glucose. The Philips Petroleum Company developed the first *Pichia pastoris* expression system in the 1980s (Gregg, 2017) and the recent release of their expression system has led to a wealth of genetic information, and expression vectors for *Pichia pastoris* are readily available from Invitrogen. *Pichia pastoris* like *Saccharomyces cerevisiae* is considered a model organism and the whole genome of *Pichia pastoris* was sequenced in 2009 (De Schutter *et.al.,* 2009).

A typical genetic engineering workflow begins with the selection of an expression system in other words the host organism. After the host organism has been chosen, the amplification of the gene or the genes of interest with PCR follows. This of course requires enough sequence information about the genes that the right kind of PCR primers can be designed. After the genes of interest have been isolated, they can be inserted into the vector. This happens by identifying the right restriction sites in the vector

that correspond to the inserts restriction sites and choosing the restriction enzymes accordingly.

Inserts are usually designed so that they attach to a core promoter that regulates the expression of the insert. In *Pichia pastoris* the usual core promotors are the AOX1 and AOX2 promotors or the constitutive glyceraldehyde-3-phosphate dehydrogenase promotor also known as the GAP promotor (Li *et.al.,* 2007). When the vector has been assembled, it is transformed by electroporation or chemical transformation into *E.coli* for cloning, then purified and digested with restriction enzymes to see if the vector was assembled correctly or possibly sequenced, to make sure the insert is in the correct frame and that there are no minor assembly errors that are non-detectable by digestions. Correct vectors are then linearized with a restriction enzyme and transformed usually by lithium acetate transformation or by electroporation into yeast cells. In the case of *Pichia pastoris* electroporation is the preferred transformation method (Çelikr *et.al.,* 2011; Li *et.al.,* 2007). Transformed colonies can then be identified by a selective marker that is usually an antibiotic resistance gene or an auxotrophic marker such as LEU3 that is a transcriptional regulator of leucine metabolism. Auxotrophic markers are genes that inactivate the organisms´ ability to produce an organic compound that is essential to its growth. Therefore, auxotrophic organisms cannot grow on media that does not have this essential compound added.

1.6 Vector design

Many *Pichia pastoris* expression vectors share an expression cassette that is composed of an inducible or constitutive promotor (usually AOX1), a transcriptional terminator (usually AOX1 terminator) that has a polyadenylation site, and a multiple cloning site in between the promotor and terminator for the insertion of recombinant gene. These features ensure that the mRNA is familiar to the yeasts cellular machinery and therefore it does not affect the stability or translational efficiency of the mRNA (Li *et.al.,* 2007). Vectors might also have multiple reporter genes such as FP tags that can be fused to the N- or C-terminal of the recombinant protein. For example, in the pGAPZ-plasmid (Figure 5) these protein tags are coded by the c-myc epitope and the 6xHis tag. The c-myc epitope tag can be used as an antibody tag if there are no antibodies for the protein that is being expressed from the vector, it can also be used to purify the protein by affinity chromatography. The 6xHistadine tag can be used to purify the ex-

pressed proteins by affinity chromatography (Life Technologies Corporation, 2010). A vector plasmid also includes multiple other variable features such as an *E.coli* origin of replication for the maintenance of vector plasmids in *E.coli* glycerol stocks, multiple cloning site, terminators and promotors that drive the selection marker expression as well as a the selective marker for both *E.coli* and *Pichia pastoris*.

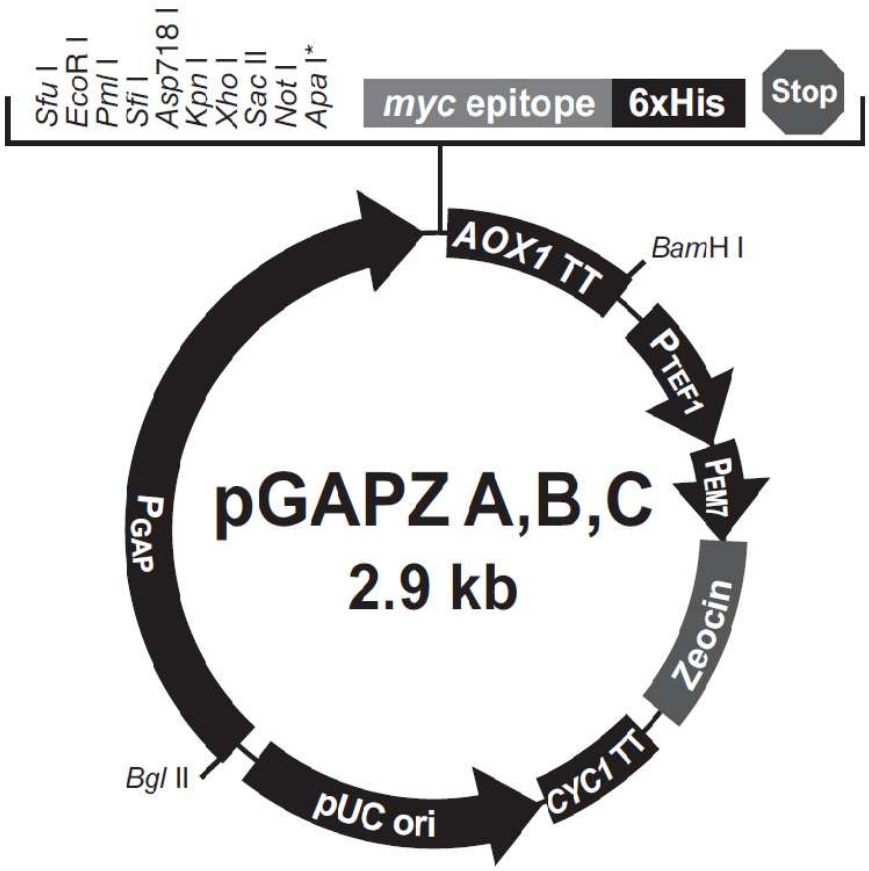

Figure 5. pGAPZ plasmid map. pGAPZ includes three different constitutive promotors; GAP promotor for gene insert expression, TEF1 for Zeocin resistance in *Pichia.p* and EM7 for Zeocin resistance in *E.coli*. The plasmid also includes two terminators AOX1 terminator for the gene insert and CYC1 terminator of Zeocin resistance. Other notable features are the pUC *E.coli* origin of replication, multiple cloning site and two protein tags c-myc epitope and 6xhis tag (Life Technologies Corporation, 2010).

Using the features in the pGAPZ plasmid a chimera protein can be created. In this thesis the chimera proteins are comprised of the FP tag coded to one of the terminals of the HA metabolic enzyme, one of the metabolic pathway enzymes and finally the amphipathic region of the Spo20 coded to the opposite terminal of the metabolic pathway enzyme. pGAPZ can also be used as an integration vector to *Pichia pastoris*. This is possible by linearizing the plasmid at the GAP promotor after which it can be integrated

to the GAP locus by homologous recombination (Life Technologies Corporation, 2010, p3).

1.7 Vector construction

The expression cassettes can be built in the following fashion. First the Spo20 amphipathic region needs to be amplified from the gDNA of *Saccharomyces cerevisiae* with PCR. This can be accomplished by creating a forward primer with one the following restriction sites: EcoR-I, Pml-I or Sfl-I. The reverse primer should include a Xho-I restriction site so the Spo20 gene can be ligated in frame with the metabolic enzyme gene. A GFP gene can be ligated in frame with the metabolic enzyme gene using primers with Xba-I sites designed into them. Examples of these primers can be found from Table 6.

Table 1. Examples of primers to amplify and add restriction sites to the Spo20 amphipathic region and a GFP. Primers generated with Geneious 8.1.9

Gene	PCR product size	Primer direction and restriction site	Sequence	TM °C
GFP	729 bp	Forward, Eco-RI	GAATTCATGAGTAAAGGAGAAGAACT	50.7
		Reverse, Xho-I	CTCGAGTTATTTGTATAGTTCATCCATG	49.4
Spo20	134 bp	Forward, Xba-I	TCTAGAATGGATAATT-GTTCTGGCTCGA	57.6
		Reverse, Xba-I	TCTAGAGCTGGTTTTTGTCGCACATC	58.6

This approach produces a chimera protein with the Spo20 amphipathic region coded to the C-terminal of the protein and a GFP coded to the N-terminal. The order can be reversed by designing the PCR primers with Xba-I and EcoR-I for the GFP gene and Xho-I sites for the Spo20 gene. Once a cloning vector containing the combinations of a FP, a metabolic enzyme (xhasA2 for example) and the Spo20 helix has been assembled into the pGAPZB plasmid, it can be transformed into *E.coli*. These transformants can be screened with colony PCR for the correctly assembled plasmids. Once these plasmids have been isolated a restriction analysis can be performed to screen the correctly assembled plasmids. These plasmids can then be sequenced and the construction of the expression cassette can begin.

The expression cassettes can be built in the same vector (pGAPZB) using the Bgl-II and Bam-HI restriction sites. One of the vectors can be linearized with Bam-HI and another vector can be digested with both Bam-HI and Bgl-II After the digestion the vectors can be ligated with T4-ligase to form a vector with the pGAPZB backbone, followed by the pGAP promotor, Spo20 helix-*xhasA2*-FP, AOXI terminator, pGAP promotor, Spo20 helix-*xhasB*-FP, AOXI terminator (Figure 6). This vector can then be linearized again with Bam-HI and another vector digested by Bam-HI and Bgl-II can be ligated into that vector. This would create an expression cassette with the genes Spo20 helix-*xhasA2*-FP, Spo20 helix-*xhasB*-FP, Spo20 helix-hasC-FP etc. (Figure 6). This process can be repeated until an expression cassette with all the chimera proteins coded into it is created

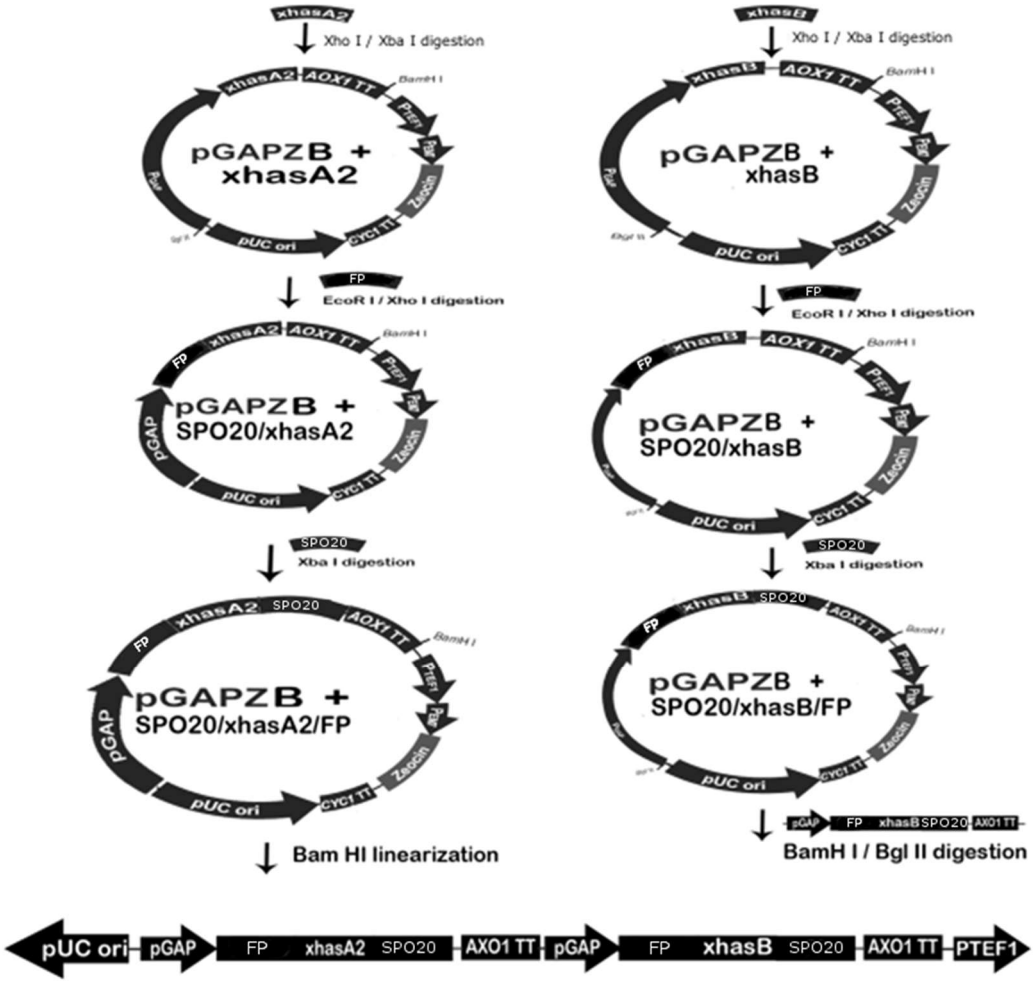

Figure 6. Construction process of the chimera protein expression cassette.

The resulting expression cassette (Figure 7) can then be transformed into *E.coli*, screened with colony PCR, restriction analysis and finally sequenced. Once the right expression vector is found, it can be transformed into *Pichia pastoris.*

Figure 7. Complete expression cassette including all the necessary recombinant genes.

The transformed colonies can be screened with colony PCR, sequenced and finally imaged with fluorescent microscopy to verify the localization of each metabolic enzyme. Finally, the expression of the genes can be measured with qPCR.

2 Materials and methods

The genetic engineering of *Pichia pastoris* GS115 (Invitrogen) was started with the amplification of the genes *xhasA2*, *xhasB*, *hasC*, *hasD* and *hasE* by KAPA Hifi PCR (Kapa Biosystems). After the successful isolation of the genes they and the pGAPZB plasmid (Invitrogen) were digested with Xho-I and Xba-I. After the digestions, the genes were inserted into pGAPZB using T4 ligase (New England Biolabs) and then cloned into a TOP10 *E.coli* strain. Cloned *E.coli* were screened with colony PCR using DyNAzyme II (Thermo Fisher) and the potentially correct plasmids were extracted using GeneJET miniprep (Thermo Fisher) column extraction, and then digested with Xho-I and Xba-I. Digested constructs were screened with agarose gel electrophoresis and sent to Seqlab Sequence Laboratories GmbH (Germany) for sequencing. After sequencing, the correct constructs were linearized using Bgl-II and then transformed to *Pichia pastoris* using electroporation. Transformed *Pichia pastoris* colonies were screened with DyNAzyme II colony PCR.

2.1 Phenol/chloroform purification of *Pichia pastoris* genomic DNA (gDNA)

Purification of *Pichia pastoris* GS115 gDNA was done with phenol/chloroform purification as follows. A 10 mL overnight culture of GS115 was grown in YPD medium in a 250 rpm shaker at 30°C and the cells were collected next morning by centrifuging 5 minutes at 3000 rpm. Supernatant was decanted off and the cell pellet was suspended into 1 mL of sterile deionised water and transferred to a sterile 1.5 mL Eppendorf tube. Cells were pelleted by centrifuging at 13500 rpm (table top centrifuge) for 30 seconds and the supernatant was decanted off. The pellet was suspended into the remaining liquid by pipetting. A volume of 0.2 mL of 2 % Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCL (pH 8), 1 mM Na_2 EDTA was added to the cell suspension. 0.2 mL of phenol:chloroform:isoamyl alcohol in the ratios of 25:24:1 was added with 0,3 g of acid washed glass beads. The sample was then vortexed in a Precellys 24 (Bertin technologies) vortexer in 30 second pulses for total of 4 minutes. The sample was kept on ice between pulses. After vortexing 0.2 mL of TE (pH 8) was added and then centrifuged for 5 minutes at 13500 rpm. The upper water phase from the tube was transferred to a new tube and 0.2 ml of phenol:chloroform:isoamyl alcohol (25:24:1) was added. The sample was inverted a few times and then centrifuged for 5 minutes at 13500 rpm. The upper water phase was transferred to a new tube and 1 mL of ice gold 94 % ethanol was added and the mixture was mixed by inverting the tube. After mixing the sample was centrifuged for 20 minutes at 14000 rpm at 4 °C. The supernatant was removed with a pipette and the DNA was suspended into 0.4 mL of TE (pH 8). 1.5 µL of 20 mg/ml RNase was added and the mixture was incubated for 5 minutes at 37 °C. 1 mL of ice cold 94 % ethanol and 4 µL of 10 M ammonium acetate were added and mixed by inverting the sample. After mixing the sample was centrifuged for 20 minutes at 14000 rpm at 4 °C and the supernatant was removed by pipetting. The DNA pellet was washed with 0.3 mL of 70 % ethanol and centrifuged for 20 minutes at 13500 rpm. The ethanol was decanted of and the pellet was left to dry at room temperature for 4 hours. The pellet was suspended into 50 µL of TE (pH 8) and stored at -20 °C

2.2 PCR and agarose gel electrophoresis

KAPA Hifi DNA polymerase (Kapa Biosystems) was used to amplify the genes *hasD*, *hasC* and *hasE* from the gDNA of *Pichia pastoris* strain GS115 (Invitrogen), and genes *xhasA2* and *xhasB* were isolated from plasmids constructed by MSc. Pandeeswari

Jeeva. KAPA HiFi PCR primers (Integrated DNA Technologies) used in this thesis can be found in Table 1.

Table 2.Kapa HiFi PCR primers, their products sizes, primer directions and restriction sites

Gene	Template	PCR product size	Primer	Sequence	Primer direction and restriction site	TM °C
hasC	gDNA P. pastoris GS115	1512bp	oSBP-072	ACTACTCGA-GAAAAATGTCTGCTTACCAGTCC	Forward XhoI	60.8
			oSBP-073	ACTATCTAGATTAG-TGCTCCAAGA-TAGTCAAGTTTC	Reverse XbaI	58.5
hasD	gDNA P. pastoris GS115	1433bp	oSBP-074	ACTACTCGA-GAAAAATGTCTTT-GGAGCAGTAC	Forward XhoI	59.6
			oSBP-075	ACTATCTAGATTAGA-TAACGTCAC-CGTTCTTGAA	Reverse XbaI	58.2
hasE	gDNA P. pastoris GS115	1683bp	oSBP-076	ACTACTCGA-GAAAAATGCCATCTTTGTTG	Forward XhoI	57.6
			oSBP-077	ACTATCTA-GATTAGGCCCAA-GCCTTGAAC	Reverse XbaI	59.5
xhasA2	Plasmid	~1700bp	oSBP-078	ACTACTCGA-GAAAAATGCAC-TGCGAGAGGTTCATC	Forward XhoI	63.5
			oSBP-079	ACTATCTAGAT-TATGCTAG-CACAAGGTCATGC	Reverse XbaI	59.3
xhasB	Plasmid	~1500bp	oSBP-080	ACTACTCGA-GAAAAATGTTTCAGATCAAAAAAATATGTTGC	Forward XhoI	60.5
			oSBP-081	ACTATCTAGAT-TA-TACCCTCTGCTTTTTATGGGG	Reverse XbaI	60.8

PCR programmes and reaction mixes were generated according to manufacturer's instructions (Kapa Biosystems, 2016). To optimize the PCR and reduce nonspecific PCR products the annealing temperature was varied by ± 2 °C. An example of a KAPA Hifi PCR programme can be found in Table 2.

Table 3.Example of a PCR programme for KAPA HiFi 2x ready mix

Step	°C	Time
1 Initial denaturation	95	3 min
2 Denaturation	98	20 sec
3 Annealing	60	15 sec
4 Extension	72	30-60 sec/kb
5 Cycling	go to step 2	30 x
6 Final extension	72	60 sec/kb
7 Storage temperature	4	forever

All agarose gel electrophoreses were done on a 1% gel with added ethidium bromide and the voltage used was 150 V. The volume of 5X loading dye used was 20% of the total volume loaded into each well and the standard used was GeneRuler 1kb (Thermo Fisher). The amplified genes were purified from the agarose gel with the QIAquick Gel Extraction kit (QIAGEN).

2.3 Restriction enzyme digestions and T4 ligase ligations

Restriction enzyme digestions with Xba-I and Xho-I were done according to manufacturer's instructions (Thermo Fisher, 2012) with one exception, the incubation times for gel extracted and purified genes was 1 hour at 37 °C and 6 hours for pGAPZB at 37 °C. The purification of the *has* gene restriction products were done with the QIAquick PCR Purification kit (QIAGEN) or and the restriction product of the pGAPZB was purified from an agarose gel using the QIAquick Gel Extraction kit (QIAGEN). Ligations of the purified genes to the linearized pGAPZB were done with T4 ligase (New England Biolabs). Molar ratio of vector to insert used was 1:3 and the ligations were incubated overnight at 4 °C.

2.4 *E.coli* electroporation

E.coli electroporation was done with TOP10 electrocompetent *E.coli*. The cells were thawed on ice for ~15 minutes and 2 µL of the ligated construct (for example pGAPZB-*xhasA2*) was added to 50 µL of competent cells. The mixture was gently mixed with a pipette and then moved to a cold 0.2 cm electroporation cuvette. Parameters used for the pulse were as follows:

- Voltage 2.5kV
- Capacitor 25μF
- Resistor 200Ω

After the 2.5kV pulse was given with Bio-Rad GenePulser II, 750 μL of low salt LB was added immediately after to the cuvette. The mixture was mixed by pipetting gently and then transferred to a sterile 15 mL falcon tube. The tubes were incubated at 37 °C, 230 rpm for 30 - 60 minutes, then 200 μL and 20 μL of the transformed cells were plated on separate low salt LB plates containing 25 μg/ml Zeocin. Plates were incubated over night at 37 °C or at 24 °C over the weekend.

2.5 *E.coli* colony PCR

Transformed *E.coli* colonies were screened using *E.coli* colony PCR according to manufacturer's instructions (Thermo Fisher, DyNAzyme II) with three exceptions, the reaction volume was scaled down to 30 μL, the initial denaturation time was increased to 3 minutes (lysis of *E.coli* cells) and the template used was a pipette tip full of *E.coli* colony that was mashed to the bottom of a 96-well plate well and the amount of sterile water was adjusted accordingly. The same pipette tip was used to make a streak on a fresh low salt LB plate containing 25 μg/mL zeocin and the plate was incubated at 37 °C overnight. The primers used for *E.coli* colony PCR can be found in Table 3

Table 4. *E.coli* colony PCR primers.

Gene name	Template	PCR product size	Primer	Sequence	TM °C
xhasA2	xhasA2+pGAPZB	675bp	oSBP 14 foward	GTACACAACCGAAGTAAGACTG	57.9
			oSBP 15 reverse	GGCACTACACCTTATCATTCAG	60.1
xhasB	xhasB+pGAPZB	825bp	oSBP 17 foward	AGGCACGTATTAATGCTTGGA	63.2
			oSBP 19 reverse	TGGGTCATGCTTTCAGAAGG	65.1

Volume of 7,5 μL of loading dye was added to the PCR products and the PCR products were loaded onto a 1% agarose gel and voltage of 150 V was applied for roughly 1

hour. Suitable candidates for plasmid purification were chosen according to the PCR product size and 2.5 mL of fresh low salt LB media containing 25 µg/mL zeocin was inoculated with the *E.coli* from the streaks. The *E.coli* were grown overnight at 37 °C.

2.6 Restriction enzyme analysis of purified cloned plasmids

The cloned plasmids chosen with *E.coli* colony PCR were purified using the GeneJET miniprep plasmid purification kit (Thermo Fisher) and restriction digested as follows. Into a 1.5 mL tube 5 µL of cold sterile water was added followed by 1.4 µL of 10X FastDigest buffer (Thermo Fisher) as well as 0.3 µL (2 U/µL) of both Xba-I and Xho-I (Thermo Fisher). The mixture was mixed by pipetting and briefly spun down with a table top centrifuge. A volume of 7 µL of the purified cloned plasmid was added and the mixture was incubated for 1 hour at 37 °C. After the incubation 3.5 µL of loading dye was added and the mixture was loaded onto a 1% agarose gel. A voltage of 150 V was applied for roughly 1 hour. The correctly ligated plasmids were identified from the gel and sent for sequencing.

2.7 Sequencing

Successfully constructed plasmids were sent to Seqlab Sequence Laboratories GmbH (Germany). The primers used for sequencing can be found in Table 4. The sequencing solutions were done according to Seqlabs instructions (SEQLAB Sequence Laboratories, 2013)

Table 5.Sequence primers for pGAPZB *xhas* constructs

Primer	Sequence	TM °C
oSBP 082	CTAGCAAGACCGGTCTTCTC	55.5
oSBP 083	GGTTTCTTCTGACCCAAAGAC	54.7

After the confirmation that the constructs were successfully ligated and the insert was in the correct frame, the plasmids were linearized with Blg-II for transformation by electroporation into *Pichia pastoris.*

2.8 BgI-II linearization and *Pichia pastoris* electroporation

BgI-II linearization of the plasmid DNA was done according to manufacturer's instructions (Thermo Fisher 2012). The amount of DNA digested was between 4 µg and 10 µg and the digestion time was 2 hours at 37 °C. The digestion products were purified with QIAquick PCR Purification kit (QIAGEN) and then ethanol precipitated with 1/10 volume of 3 M sodium acetate, 2.5 volumes of 96 % ethanol. The mixture was incubated at - 20 °C for 20 minutes and then centrifuged for 10 minutes at 14000 rpm at 4 °C. The supernatant was pipetted off and the pellet was washed with one volume of 70 % ethanol. After the addition of 70 % ethanol, the mix was centrifuged for 5 minutes at 14000 rpm at 4 °C. The pelleted DNA was suspended in cold sterile water.

The electrocompetent cells were made from a glycerol stock of *P.pastoris* strain GS115 according to manufacturer's instructions (Life Technologies Corporation, 2010, p17). The electroporation to *P.pastoris*, with the combinations of cloning vectors pGAPZB-*xhasA2*, -*xhasB* as well as -*xhasA2 and -xhasB*, were done according to manufacturer's instructions (Life Technologies Corporation, 2010, p17; Bio-Rad Laboratories, p10). Transformants were plated on YPDS plates containing 100 µg/ml Zeocin and grown at 30 °C for four days.

2.9 *Pichia pastoris* colony PCR

Transformed colonies were screened using colony PCR (DyNAzyme II). Streaks from the transformed colonies of *Pichia pastoris* were made on a fresh YPDS plates containing 100 µg/ml Zeocin and incubated overnight at 30 °C. These streaks were screened using the following protocol for the correct integration of the *xhas* genes into the GAP locus. Small amount of the yeast colony was suspended into 30 µL of Zymolyase (1 mg/ml Zymolyase in 1 M sorbitol) in a 96 well plate well. The suspension was incubated for 1.5 hours at 37 °C and 5 µL was used as a template for DyNaZyme II PCR which was done according to manufacturer's instructions (Thermo Fisher, DyNAzyme II). Primers used in the *Pichia Pastoris* colony PCR can be found in Table 5.

Table 6.Primers used for *Pichia Pastoris* colony PCR

Gene name	Template	PCR product size	Primer	Sequence	TM °C
xhasA2	pGAPZB+xhasA2 integrated	432bp	oSBP 082 forward	CTAGCAAGAC-CGGTCTTCTC	55.5
			oSBP 016 reverse	GCAC-TTTCTCAGGTAG-TCCTC	59.5
xhasB	pGAPZB+xhasB integrated	392bp	oSBP 082 forward	CTAGCAAGAC-CGGTCTTCTC	55.5
			oSBP 018 reverse	GCTGATAAACAC-TAAATCCGCT	61.4

A volume of 12.5 µL of 5X loading dye was added to the PCR products and then loaded into a 1% agarose gel. A voltage of 150 V was applied for roughly 1½ – 2 hours.

3 Results

Kapa Hifi PCR

Two specific PCR products were produced from the plasmids containing the *xhasA2* and *xhasB* genes and no specific products were produced from the gDNA of *P.pastoris* GS115. The annealing temperature was varied by ± 2 °C, and the template DNA concentration was varied from either ¼ or 1 (1 – 10 ng/µL) times the recommended amount. Neither the DNA concentration nor the varied annealing temperature, affected the PCR products that were amplified from the gDNA (figure 8).

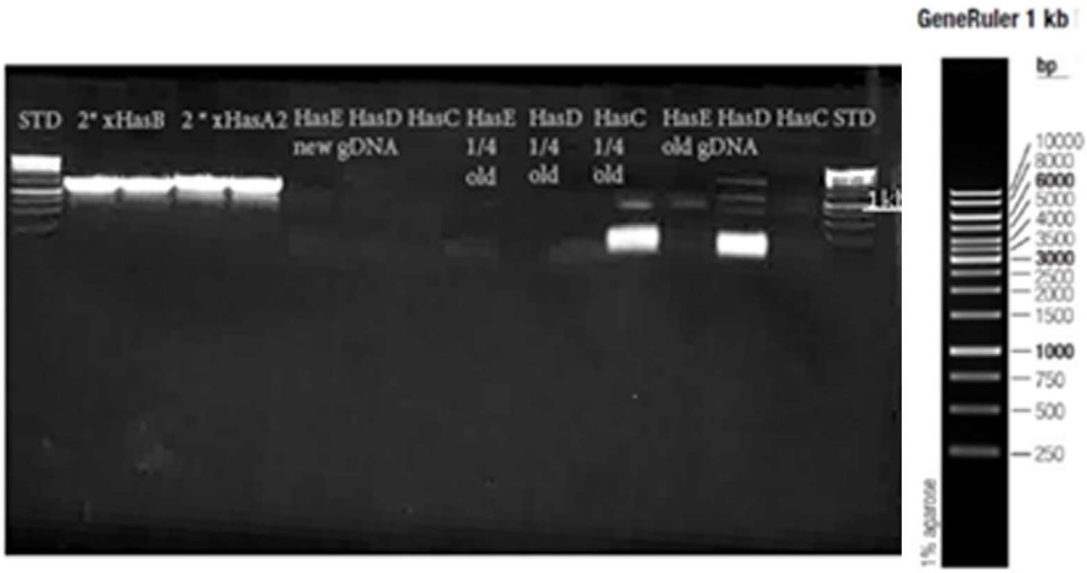

Figure 8. KapaHifi PCR products at an annealing temperature of 60 °C and a varying template concentration.

No correct sized PCR products (~ 1.4 kb – 1.7 kb) were produced at any annealing temperature or template concentrations from the gDNA (Appendix 1). *XhasA2* and *xhasB* PCR products were produced at an annealing temperature of 60 °C and the recommended template concentration (Figure 8).

E.coli colony PCR

E.coli colony PCR produced specific products only from transformants which had the pGAPZA-*xhasA2* and pGAPZB-*xhasA2* cloning vectors and no specific products were produced from the *xhasB* cloning vector even though the number of amplification cycles were increased from 30 to 35 and the annealing temperature was changed from 50 °C to 55 °C. The Colony PCR produced specific products from the wrong pGAPZA-*xhasA2* cloning vector transformants (Appendix 2, Figure 2.1 A), Once the correct cloning vector pGAPZB-*xhasA2* was transformed into the *E.coli,* the correct PCR products were produced from the pGAPZB-*xhasA2* transformants (Figure 9).

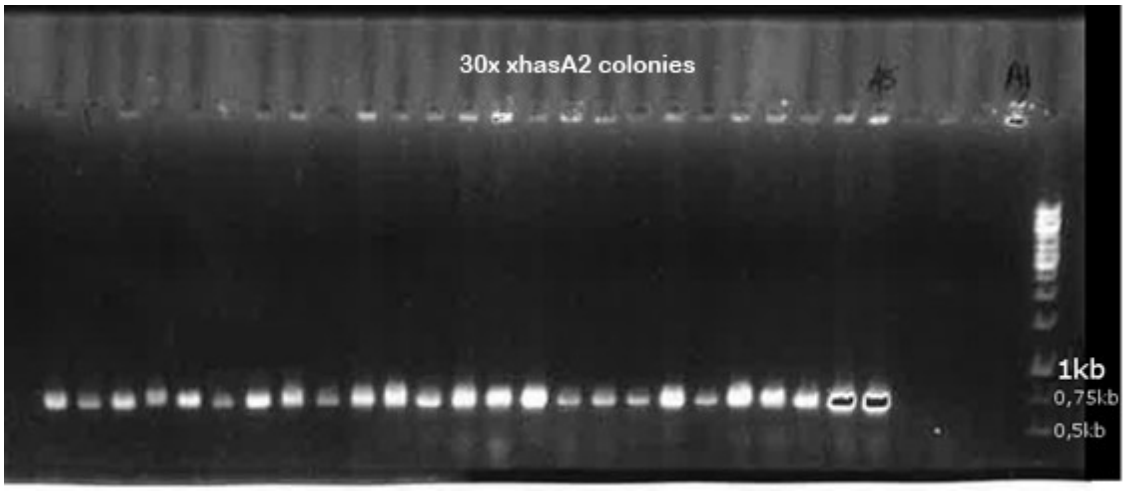

Figure 9. *E.coli* colony PCR where pGAPZB *xhasA2* transformants are producing correct sized PCR products (~ 700 bp). Annealing temperature 50 °C and amplification of 30 cycles

No specific PCR products were produced from either of the cloning vectors pGAPZA-*xhasB* or pGAPZB-*xhasB* (Appendix 2, Figures 2.1 B - D). The annealing temperature (50 - 55 °C) and the amplification cycles (30 - 35 cycles) were varied, but this approach did not produce any specific PCR products of the correct size (~ 800 bp).

Restriction analysis of the cloning vectors

Initial unsuccessful ligations with the Xba-I and Xho-I digested products were due to the wrong cloning vector pGAPZA used. Multiple troubleshooting efforts were made to deduce the problem with this ligation. The digestion time for the pGAPZA was increased from 1 hour to 6 hours at 37 °C; this approach did not produce any correct transformants (Appendix 3, Figure 3.1). Ligation time was then increased from 1 hour at room temperature to overnight at 4°C with no significant improvement (Appendix 3, figure 3.2). After these results, the pGAPZA vector was changed to the pGAPZB vector. Once the pGAPZB vector was purified correctly and the concentration of the vector was high enough in the ligation mixture (~ 25 - 50 ng/µL), the ligation produced correct transformants (Figure 10).

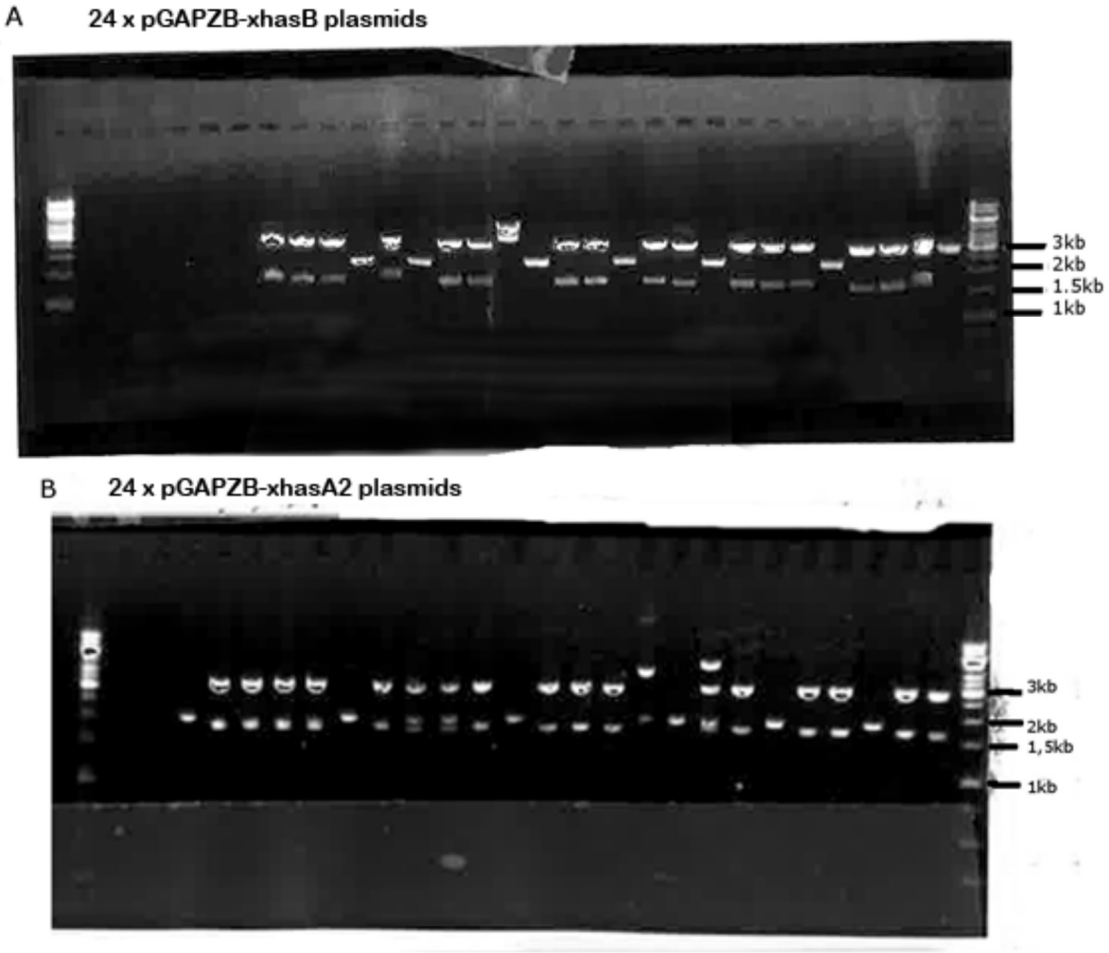

Figure 10. A) Restriction digestion analysis of the pGAPZB-*xhasB* plasmids. Most plasmids are digested to the correct product sizes of 1.5 kb and 2.9 kb. B) Restriction digestion analysis of the pGAPZB-*xhasA2* plasmids. Most plasmids are digested to the correct product sizes of 1.7 kb and 2.9 kb

Four potential pGAPZB-*xhasA2* and three potential pGAPZB-*xhasB* plasmids were sent for sequencing. All the plasmids were assembled correctly, and the sequenced plasmids matched that of the plasmids constructed *in silico* (Genious 8.1.9). The sequenced plasmids were then linearized for *Pichia pastoris* electroporation transformations.

Pichia pastoris colony PCR

Transformations done with Bsp-HI linearized vectors did not produce any correct transformants (Appendix 4, figure 4.1) This is because Bsp-HI cuts in the insert as well as the GAP promotor so even though colonies can be seen on selective plates there is no integrated pGAPZB-*xhasA2/B* vector at the GAP locus. The digestion enzyme was

changed from Bsp-HI to Bgl-II to avoid digesting the *xhasA2* insert into two pieces. After the successful transformation of the linearized vector, there was one colony on the *xhasA2* plates, one on *xhasA2* and *xhasB* plates and a total of eight colonies on the *xhasB* plates. The colony PCR from these colonies produced the correct sized PCR products from all the colonies and the negative control. Since the negative control was producing PCR products even though it should not, two more colony PCR reactions were made with new templates (Figure 11).

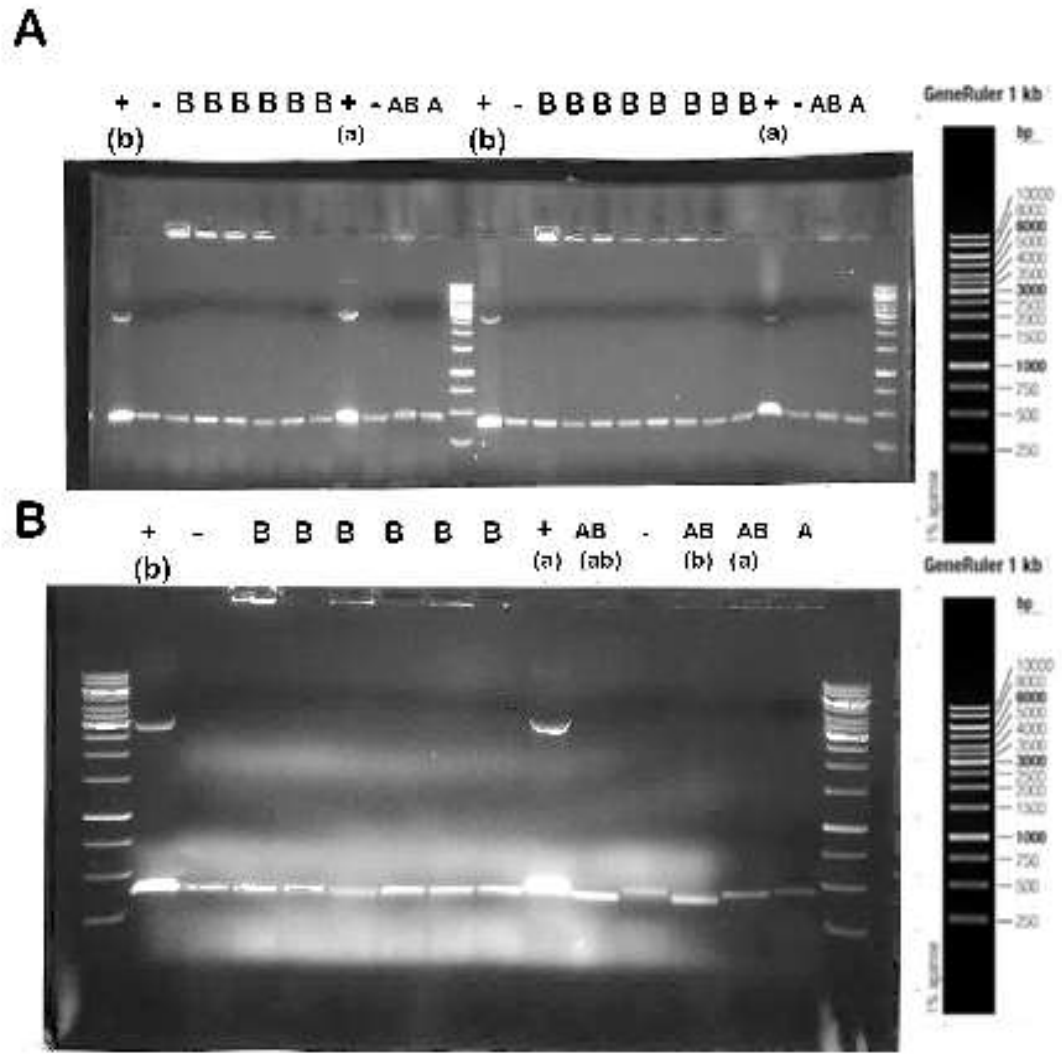

Figure 11. A) Amplification cycles of 35 and 30 with new templates. B) Amplification cycles 35, with new templates. Negative controls producing the correct sized PCR product with both primer pairs. A = colonies transformed with the pGAPZB-*xhasA2 vector*. B = colonies transformed with the pGAPZB-*xhasB* vector. AB = colonies transformed with both vectors. (a) = primers used to screen for the pGAPZB-*xhasA2* vector. (b) = primers used to screen for the pGAPZB-*xhasB* vector. (ab) = both primer pairs.

In Figure 11 we can see that the negative controls are still producing PCR products of the correct size from the new templates.

4 Discussion

Due to the time constraint for the completion of this thesis, no correct constructs with the right combinations of x*hasA2*, *xhasB, hasC*, *hasD*, *hasE*, FPs and *Spo20* were produced. Instead a handful of potential strains of *Pichia pastoris* with the *xhasA2* and *xhasB* integrated to the GAP promoter locus were produced.

KAPA HiFi PCR

The reason why no specific products were produced from the gDNA of *P.pastoris* is probably due to primer design or the template gDNA being used. Non-specific primer binding seems to be the more probable explanation since for example; in figure B (Appendix 1) the primers for *hasD* are producing five different products instead of one.

E.coli colony PCR

One explanation for the correct size PCR products been produced from the pGAPZA transformants, could be that the insert DNA left in the ligation mix has also been transformed into the *E.coli* (Appendix 2, Figure 2.1 A). When the cells lysate, the DNA is released and amplified with PCR producing the correct PCR products. In figure 2.1 A (Appendix 2) the PCR products (~700bp) from the correct pGAPZB-xhasA2 vector can be seen. The difference in the brightness and hence the amount of PCR product can also be seen. Therefore, the assumption that the PCR products been produced in Figure 2.1 A (Appendix 2) are from the insert DNA transformed into to the *E.coli,* could be an explanation for these results. The PCR products been produced in Figure 2.1 A (Appendix 2) could also be due to non-specific primer binding to genomic homologue with a similar sequence to the *xhasA2* gene. Unfortunately, these results do not confirm these assumptions and sequencing of these unspecific PCR products would be needed to confirm them.

The lack of correct PCR products been produced for either the wrong pGAPZA *xhasB* cloning vector or the correct pGAPZB *xhasB* cloning vector (Appendix 2) is most likely

due to primer design. This assumption is supported by the fact that the primers for *xhasB* are producing PCR products of the wrong size as can be seen in Figure 2.1 D (appendix 2).

Vector construction

The results shown in the Appendix 3 are due to an unfortunate mistake, trying to insert the genes *xhasA2* and *xhasB* to the cloning vector pGAPZA. The pGAPZA plasmid does not have the Xba-I restriction site so inserting the *xhasA2* and *xhasB* genes digested with both Xba-I and Xho-I is highly unlikely and the plasmid is more likely to ligate back into itself. After the discovery of the unfortunate mistake made with the wrong cloning vector pGAPZA, some problems with the ligations of the insert into the pGAPZB vector arose. Since the restriction digested pGAPZB vector was not purified on an agarose gel, the small fragment that was cleaved from the vector with Xba-I and Xho-I remained in the ligation mixture. Ligation of this fragment with, the cloning vector is more likely to happen than with the *xhas* genes; therefore, the chances of the insert ligating with the vector reduce significantly, producing vector only transformants. Once these mistakes were corrected, most plasmids analysed with restriction analysis produced that correct sized 2.9 kb vector backbone and either 1.7 kb fragment from the *xhasA2* gene (Figure 10) or 1.5 kb fragment from the *xhasB* gene (Figure 10)

Pichia pastoris colony PCR

The probable explanation for the PCR products been produced from the negative controls is a contamination of both the negative control templates. The template AB is producing different sized products with different primer pairs which suggest that the strain might have both expression cassettes integrated into the GAP locus. The PCR products being produce from the *xhasA2* and *xhasB* templates are roughly the correct size, so it's reasonable to suggest that these expressions cassettes might have also been integrated. These results are somewhat questionable considering that the negative controls are also producing products, and to confirm the integration, another colony PCR could have been made with other primer pairs. The strains could also have been screened with qPCR to confirm the expression of the *xhas* genes.

5 Conclusions

Due to the time constraints set for this thesis, all of its originals goals were not met and this is mostly due to the unfortunate mistake made with the pGAPZA cloning vector. The troubleshooting of the ligations into this vector took most of the time allotted; therefore, not enough time was left, to begin the production of the chimera proteins. The minor mistakes made in the restriction digestion of the cloning vector pGAPZB and the troubleshooting for this ligation also took more time than anticipated. In hindsight, a more detailed plan for the execution of the plasmid construction phase of this thesis would have probably yielded better results, but the initial goal would probably not have been met. The possibility of successful integration of all the expression cassette combinations to *P.pastoris* does give some value to the SynBioPol project. If the integration has happened correctly, these strains can be used as a reference to compere the expression of these *native* proteins to the strain that expresses the chimera proteins.

References

Bio-Rad Laboratories, California 94547 USA. Gene Pulser® Electroprotocol species list. Available from http://www.bio-rad.com/webroot/web/pdf/lsr/literature/D035553_protocols_fungal_plant_other.pdf [Accessed 5.4.2017]

Boss J.D, Meinardi M. 2000. The 500 Dalton rule for the skin penetration of chemical compounds and drugs. Experimental Dermatology, volume 3, issue 3, pages 165 – 169. Available from http://onlinelibrary.wiley.com/doi/10.1034/j.1600-0625.2000.009003165.x/abstract;jsessionid=E87559ECCC54F007241E88BD5905D6EB.f01t02 [Accessed 24.1.2017]

Çelik E, Çalık P. 2011. Production of recombinant proteins by yeast cells. Biotechnology Advances, Volume 30, Issue 5, Pages 1108–1118. Available from http://www.sciencedirect.com/science/article/pii/S0734975011001649 [Accessed 15.2.2017]

Chen Yu A, Scheller Richard H. 2001. SNARE-mediated membrane fusion. Nature Reviews Molecular Cell Biology 2, Pages 98-106. Available from http://www.nature.com/nrm/journal/v2/n2/full/nrm0201_098a.html [Accessed 31.1.2017]

Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. 2010. Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. Physiological Reviews, Volume. 90, No. 3, pages 1103-1163. Available from http://physrev.physiology.org/content/90/3/1103 [Accessed 13.4.2017]

De Oliveria J, Carvalho L, Gomes A, Queiroz L, Magalhães B, Parachin N. 2016. Genetic basis of hyper production of hyaluronic acid in neutral and engineered microorganisms. Microbial Cell Factories 15:119. Available from <http://microbialcellfactories.biomedcentral.com/articles/10.1186/s12934-016-0517-4 [Accessed 20.1.2017]

De Schutter K, Lin YC, Tiels P, Annelies Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, Van de Peer Y,Callewaert N. 2009. Genome sequence of the recombinant protein production host Pichia pastoris. Nature Biotechnology, Volume 27, Pages 561-566, Available from http://www.nature.com/nbt/journal/v27/n6/full/nbt.1544.html [Accessed 15.2.2017]

Duman J.G, Forte J.G. 2003. What is the role of SNARE proteins in membrane fusion? American Journal of Physiology Cell Physiology, Volume 285, Issue 2; C237–C249. Available from http://ajpcell.physiology.org/content/285/2/C237.long#sec-2 [Accessed 31.1.2017]

Ferraz-Nogueira JP, Dı´ez-Guerra FJ, Llopis J. 2014. Visualization of Phosphatidic Acid Fluctuations in the Plasma Membrane of Living Cells. PLoS ONE 9(7), Availa-

ble from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0102526 [Accessed 8.2.2017]

Horchani H, de Saint-Jean M, Barelli H, Antonny B. 2014. Interaction of the Spo20 Membrane-Sensor Motif with Phosphatidic Acid and Other Anionic Lipids, and Influence of the Membrane Environment. PLoS ONE 9(11), Available from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0113484 [Accessed 8.2.2017]

James Cregg, Ph.D. The Pichia System [Internet]. Keck Graduate Institute, Claremont, California. Pichia.com. Available from http://www.pichia.com/science-center/james-m-cregg-ph-d/ [Accessed 14.2.2017]

Joeng E, Shim W, Kim JH. 2014. Metabolic engineering of *Pichia pastoris* for production of hyaluronic acid with high molecular weight. Journal of Biotechnology, volume 185, pages 28 - 36. Available from http://www.sciencedirect.com/science/article/pii/S0168165614002648 [Accessed 20.1.2017]

Kapa Biosystems, Wilmington Massachusetts. 2016. KAPA HiFi HotStart ReadyMix PCR Kit [Technical Data Sheet]. Available from https://www.kapabiosystems.com/assets/KAPA_HiFi_HotStart_ReadyMix_TDS.pdf [Accessed 28.3.2017]

Kogan G, Šoltés L., Stern R. et al. 2007. Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. Biotechnology Letters, volume 29, pages 17- 25. Available from <http://link.springer.com/article/10.1007%2Fs10529-006-9219-z [Accessed 20.1.2017]

Kremers J, Gilbert SG, Cranfill PJ, Davidson MW, Piston DW. 2011. Fluorescent proteins at a glance. Journal of Cell Science, volume 124, pages 157-160. Available from http://jcs.biologists.org/content/124/2/157#sec-2 [Accessed 13.4.2017]

Li P, Anumanthan A, Gao X.G, Ilangovan K, Suzara V.V, Düzgüneş N, Renugopalakrishnan V. 2007. Expression of Recombinant Proteins in *Pichia Pastoris.* Applied Biochemistry and Biotechnology, Volmue 142, Issue 2, Pages 105-124. Available from http://link.springer.com/article/10.1007/s12010-007-0003-x [Accessed 15.2.2017]

Life Technologies Corporation, Carlsbad, California. 2010. pGAPZ A, B, and C, pGAPZα A, B, and C Pichia expression vectors for constitutive expression and purification of recombinant proteins [User's manual]. Available from https://tools.thermofisher.com/content/sfs/manuals/pgapz_man.pdf [Accessed 22.2.2017]

Liu S, Wilson K.A, Rice-Stitt T, Neiman AM, McNew JA. 2007. *In vitro* Fusion Catalysed by the Sporulation-Specific t-SNARE Light-Chain Spo20p is Stimulated by Phosphatidic Acid. Traffic, Volume 8, Issue 11, Pages 1630 – 1643. Available from

http://onlinelibrary.wiley.com/doi/10.1111/j.1600-0854.2007.00628.x/abstract [Accessed 31.1.2017]

Nakanishi H, Santos P, Neiman A.M. 2004 Positive and Negative Regulation of a SNARE Protein by Control of Intracellular Localization, Molecular Biology of the Cell, Volume 15, Pages 1802-1815. Available from http://www.molbiolcell.org/content/15/4/1802.long [Accessed 1.2.2017]

Nakase Y, Nakamura T, Hirata A, Routt Sheri M, Skinner H.B, Bankaitis V.A, Shimoda C. 2001. The Schizosaccharomyces pombe spo20^{+} Gene Encoding a Homologue of Saccharomyces cerevisiae Sec14 Plays an Important Role in Forespore Membrane Formation. Molecular Biology of the Cell, Vol 12, Pages 901-917. Available from http://www.molbiolcell.org/content/12/4/901.full.pdf [Accessed 1.2.2017]

Neiman A.M, Katz L, Brennwald P.J. 2000. Identification of Domains Required for Developmentally Regulated SNARE Function in *Saccharomyces cerevisiae* Genetics, Volume 155, No. 4, Pages 1643-1655. Available from http://www.genetics.org/content/155/4/1643.short [Accessed 1.2.2017]

SEQLAB Sequence Laboratories Göttingen GmbH. 2013. Barcode Economy Run Service sample preparation. Available from http://www.seqlab.de/index.php?id=barcode-economy-run&L=1 [Accessed 5.4.2017]

Thermo Fisher Scientific Inc. 2012. Fast Digestion of DNA [Protocol]. Available from https://tools.thermofisher.com/content/sfs/manuals/MAN0012413_Fast_Digestion_DNA_UG.pdf [Accessed 28.3.2017]

Thermo Fisher Scientific Inc. 2012. Thermo Scientific DyNAzyme II DNA Polymerase [Technical Data Sheet]. Available from https://tools.thermofisher.com/content/sfs/manuals/MAN0012770_DyNAzyme_II_DNAPolymerase_250U_UG.pdf [Accessed 28.3.2017]

Weigel P.H. 1998. Bacterial Hyaluronan Synthase - an update [Internet]. GlycoForum; August 15 1998 [update April 15.2004; cited February 2.2017]. Available from http://glycoforum.gr.jp/science/hyaluronan/HA06/HA06E.html [Accessed 20.1.2017]

https://www.amazon.com/Hyaluronic-Pure-Highest-Anti-Aging-Moisturizer-Non-greasy/dp/B0090UJFYI?th=1 [Accessed 24.1.2017]

KAPA HiFi PCR results

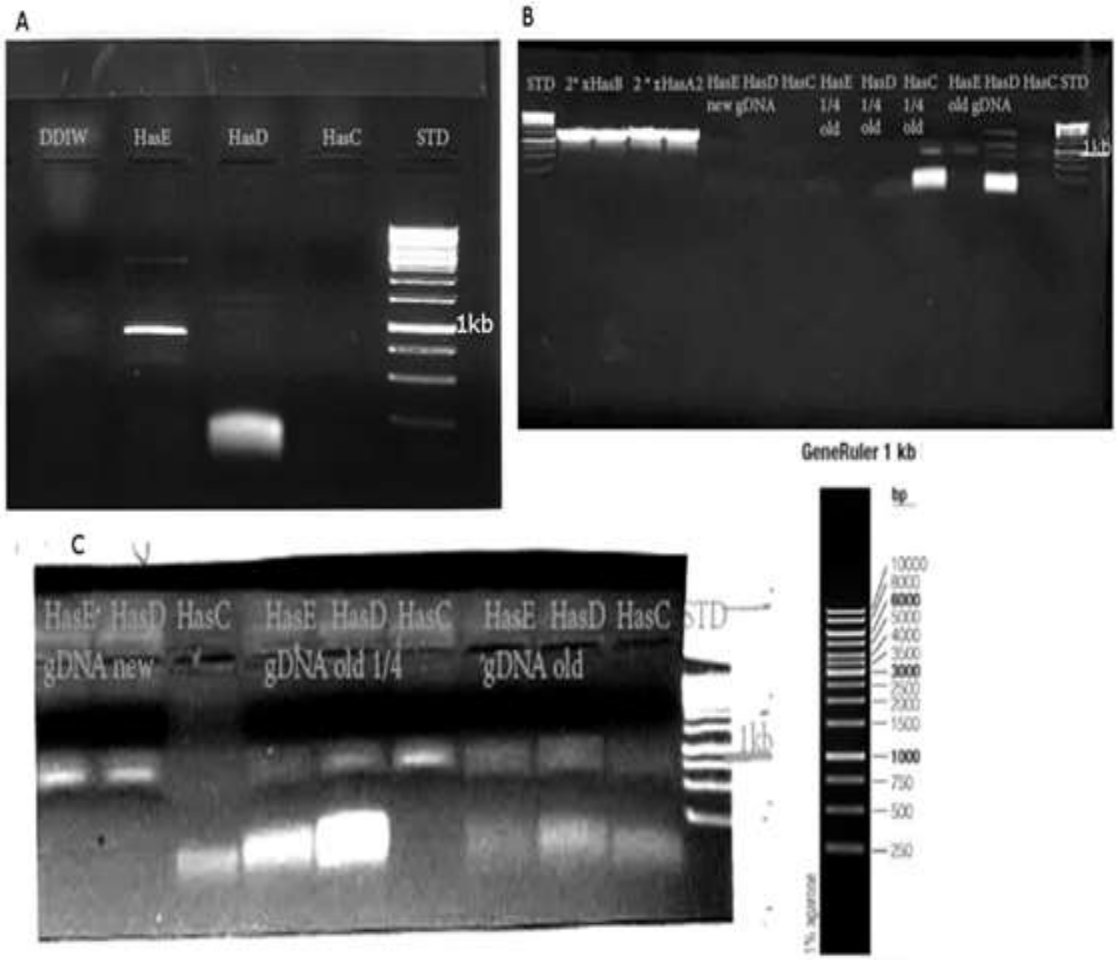

Figure 1.1 A) Kapa Hifi PCR products at an annealing temperature of 58 °C. B) Kapa Hifi PCR products at an annealing temperature of 60 °C and varying template concentration. C) Kapa Hifi PCR products at an annealing temperature of 62 °C and varying template concentration.

xhasA2 and xhasB colony PCR results

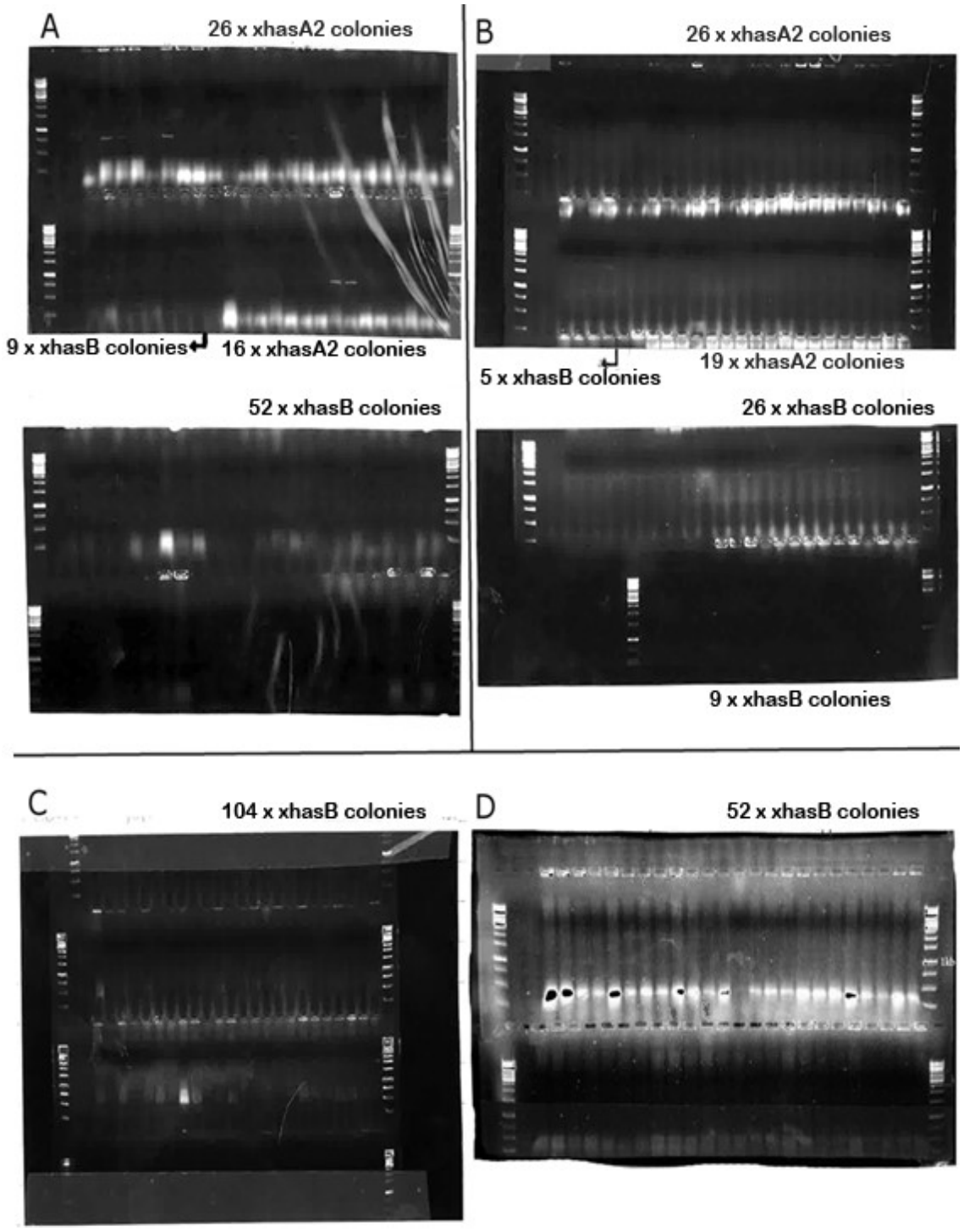

Figure 2.1 Figure A shows the pGAPZA+*xhasA2/B* products. Figure B shows the column purified pGAPZB+*xhasA/B* vector with digestion time of 6 hours, ligation time of 3h and annealing temperature of 50°C. Figure C shows the gel purified pGAPZB+*xhasB* vector with digestion time of 6 hours, ligation time of overnight at 4°C and annealing temperature of 50°C. In figure D the reaction conditions are exactly the same as in figure C but the annealing temperature has been raised from 50°C to 55°C

Restriction analysis results for the pGAPZA-insert vectors

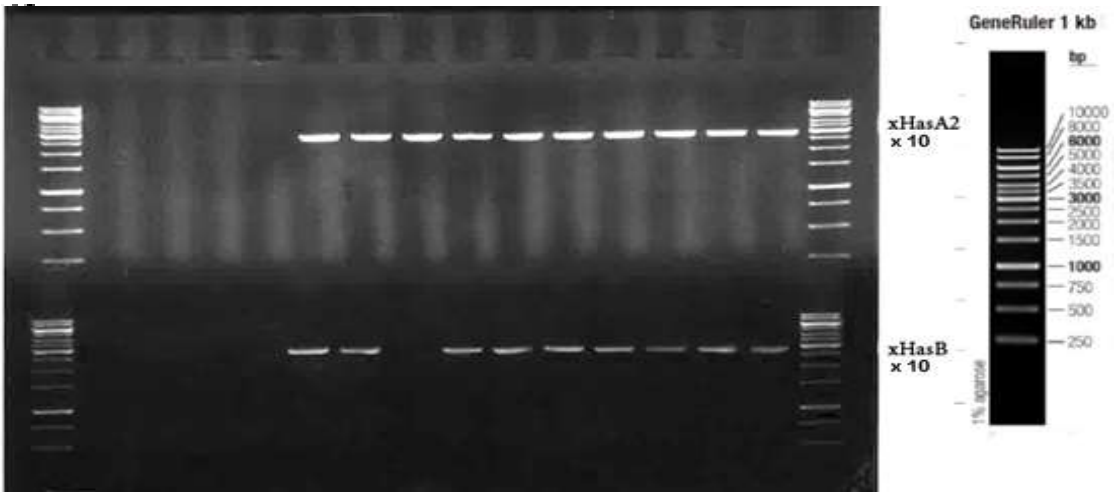

Figure 3.1 Restriction analysis of pGAPZA-*xhasA2/B* transformants. pGAPZA was digested for 6 hours at 37°C then ligated with inserts for 1 hour at RT. Only the linearized vector pGAPZA (2,9 kb) can be seen

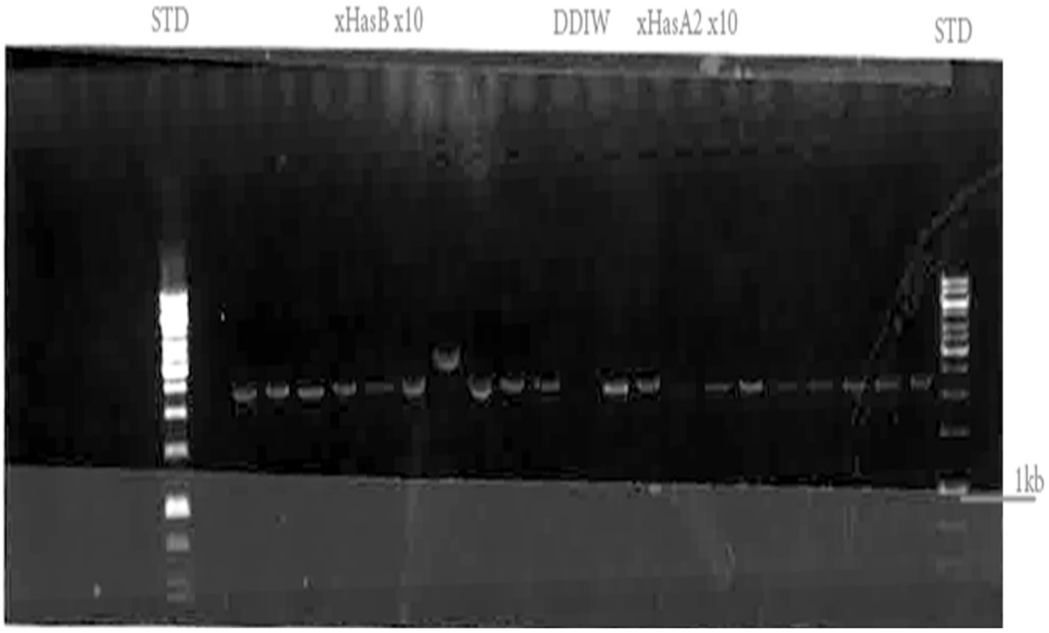

Figure 4.1 Restriction analysis of pGAPZA transformants. Ligation time of pGAPZA and inserts increased to overnight at 4°C

Pichia pastoris colony PCR results with the Bgl-II linearized vectors

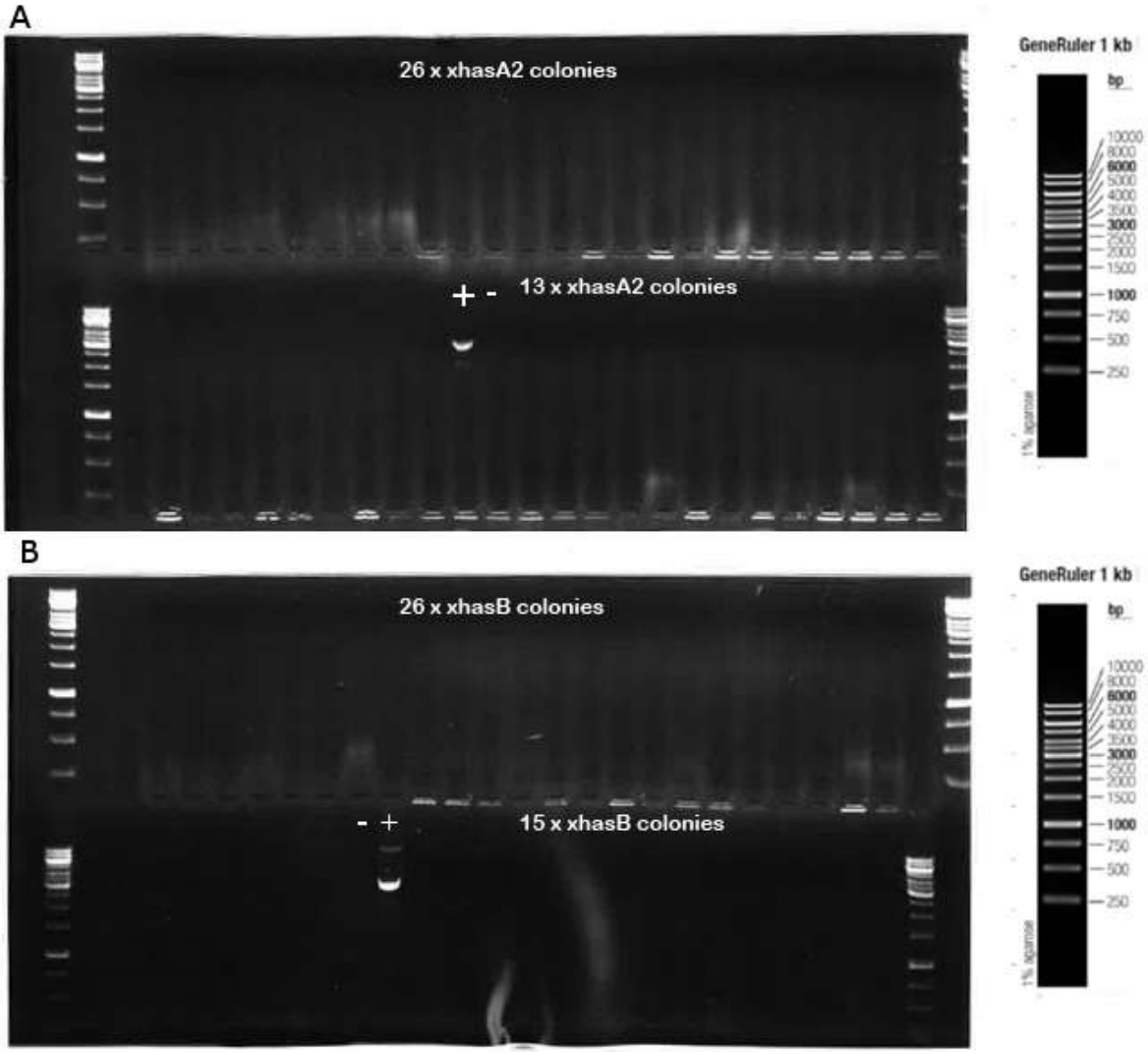

Figure 5.1 A) Colony PCR results of the Bsp-HI linearized pGAPZB-*xhasA2* plasmid. B) Colony PCR results of the Bsp-HI linearized pGAPZB-*xhasB* plasmid. No visible products can be seen except for the positive plasmid control (+). The negative control (-) of an untransformed GS115 colony did not produce any products.