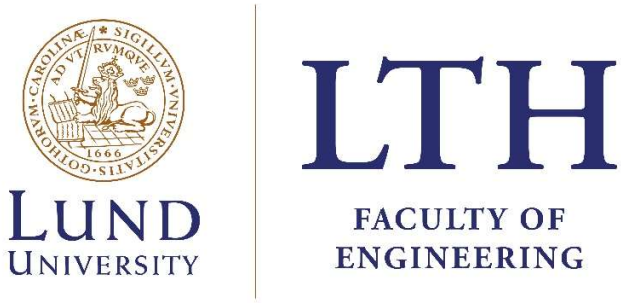


Using CRISPR/Cas9 to regulate enzyme expression levels in the Weimberg pathway for optimal xylose consumption in *Saccharomyces cerevisiae*

Anna Heiling

Supervisors: Celina Borgström, Marie Gorwa Grauslund
Examiner: Catherine Paul

Master Thesis
Applied Microbiology (TMB)
Lund University
2020-06-10

Abstract

Saccharomyces cerevisiae is a commonly used industrial organism which cannot consume xylose naturally. This is a problem since xylose is a major component in renewable resources, for instance lignocellulose. The Weimberg pathway is a pathway that converts xylose into α-ketoglutarate. It has been successfully integrated into *S. cerevisiae*, enabling it to utilize xylose. The successful strain had the genes *xylB*, *xylD*, *xylX* (from *Caulobacter crescentus*) and *ksaD* (from *Corynebacterium glutamicum*) integrated with one copy of *xylB* and four copies of each of the others. In addition, the strain had the iron regulon repressor gene, *FRA2*, deleted to enhance the activity of the XylD enzyme.

In the present master thesis project, the aim was to find the optimal number of copies of the genes *xylB*, *xylD*, *xylX* and *ksaD* as well as what enzyme activities that are needed for the pathway to be functional. This was done by creating a pallet of strains that contained different copy numbers of the genes, using CRISPR/Cas9. Enzyme activity was measured before further analysis of biomass formation and metabolite production.

It was found that TMB CB 016 (*xylB*, *xylD*, 4x(*xylX*, *ksaD*)) had an activity in the coupled assay, measuring the activity of XylD, XylX and KsaD together. However, due to the problems with the assay the results are unreliable and need to be further investigated. The constructed strains did not grow as well as the original strain with the functional Weimberg pathway, nor were any of them able to consume xylose.

Acknowledgements

This master's degree project was conducted at the Applied Microbiology Department at Lund University during the spring semester of 2020.

I would like to thank my supervisors at TMB. Celina Borgström for teaching me how to perform all the experiments in the lab, taking late evening samples, discussing the results with me and answering all of my questions. Marie Gorwa Grauslund for giving me valuable inputs when the experiments did not go as planned.

I would also like to thank Viktor Persson for designing the gRNA plasmid sequences used in both our projects, Henrik Almqvist for performing the UPLC experiments for me and Catherine Paul for taking the time to be my examiner.

Last but not least I would like to thank all the people at TMB for the great times and the laughs.

Table of content

1. Introduction

Nowadays it is of high interest to use sustainable and renewable carbon sources in the industrial production of fine and bulk chemicals. Lignocellulose, which contains a high percentage of xylose, is one of the promising renewable sources (Hongzhang, 2014). However, one of the most used organisms in industrial production, the baker's yeast *Saccharomyces cerevisiae*, does not use xylose naturally (Barnett, 1976).

1.1. The Weimberg pathway

Several engineering strategies have been carried out to enable xylose consumption by *S. cerevisiae*. Most strategies used the xylose reductase/xylitol dehydrogenase (XR/XDH) or xylose isomerase (XI) pathway which leads to the production of pyruvate derivatives, in particular ethanol (Kim, et al., 2013). Using the same pathway for the production of carboxylic acids results in carbon losses. The losses occur both in the cytosol and mitochondria in the form of carbon dioxide. Therefore, it is of interest to introduce the Weimberg pathway into *S. cerevisiae*, which does not result in carbon losses when producing carboxylic acids from xylose (Borgstrom, et al., 2019).

In the Weimberg pathway xylose is converted into α-ketoglutarate in five steps, see figure 1. Oxidation of xylose to xylonolactone, followed by an isomerization to xylonate, dehydration to 2-keto-3-deoxy xylonate, dehydration to α-ketoglutarate semialdehyde and finally oxidation to α-ketoglutarate. It is an oxidative pathway and growth through it is dependent on the tricarboxylic acid cycle (TCA cycle) and respiration. The pathway was first discovered in the 1960s in the bacterium *Pseudomonas fragi* (Weimberg, 1961). Since then the pathway has been found in other organisms, one of them being *Caulobacter crescentus* (Stephens, et al., 2007). The genes coding for the pathway in *C. crescentus* have been introduced into other organisms, for example *Corynebacterium glutamicum* and *S. cerevisiae* (Radek, et al., 2014; Borgstrom, et al., 2019; Wasserstrom, et al., 2018).

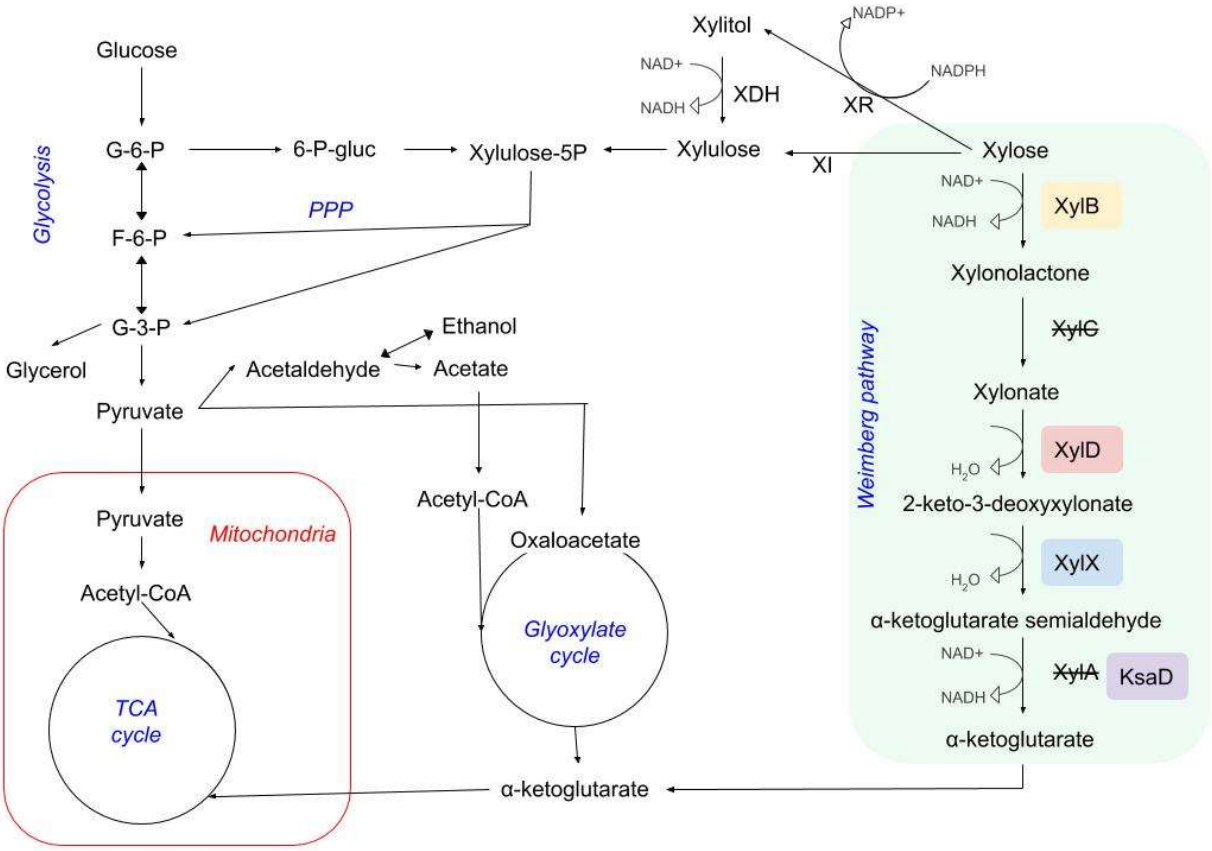


Figure 1. Metabolic map showing how xylose can be consumed in Saccharomyces cerevisiae *via three different pathways. Xylose is converted to glycolytic intermediates through the XR/XDH pathway or the XI pathway and the pentose phosphate pathway (PPP). In the Weimberg pathway xylose is converted to α-ketoglutarate, an intermediate of the tricarboxylic acid cycle (TCA cycle). The enzymes involved in the Weimberg pathway are stated: XylC is crossed over because it can also occur spontaneously and XylA from* C. crescentus *is replaced by KsaD from* C. glutamicum *since it has been shown that it improves the pathway efficiency in* S. cerevisiae *(Nygård, et al., 2014; Borgstrom, et al., 2019).*

When introducing the Weimberg pathway into *C. glutamicum*, the operon *xylXABCD* from *C. crescentus* was expressed, resulting in the organism being able to utilize xylose. The operon contained the genes *xylX*, *xylA*, *xylB*, *xylC* and *xylD*, coding for the enzymes XylX (3-keto-2-deoxy-xylonate dehydratase), XylA (α-ketoglutarate semialdehyde), XylB (xylose dehydrogenase), XylC (xylonolactonase) and XylD (xylonate dehydratase), respectively. The enzymes are involved in the five steps of the Weimberg pathway, converting xylose into α-ketoglutarate, see figure 1 (Radek, et al., 2014). In a later study the gene *ksaD* (coding for KsaD, an analogue to XylA) in *C. glutamicum* was found to have similar activity as XylA, with only 25% sequence similarity but with a high secondary structure resemblance. In addition, the study showed that either XylD or XylX are needed in *C. glutamicum* for the Weimberg pathway to work (Brüsseler, et al., 2019).

In *S. cerevisiae* the upper part of the Weimberg pathway (the conversion of xylose to xylonate) was first successfully introduced by integrating the genes *xylB* and *xylC* from *C. crescentus*,

resulting in accumulation of xylonate. The gene encoding for aldose reductase, *GRE3*, was deleted in order to reduce the competing production of xylitol (Toivari, et al., 2012). Addition of the lower Weimberg pathway did not result in any downstream flux from xylonate. It was found that a strain expressing *xylC* did not grow on glucose supplemented with xylose but a strain without XylC did (Wasserstrom, et al., 2018). Another study has shown that a rapid conversion to xylonate led to a decrease in the internal pH which initiates cell death. It also showed that not expressing XylC led to a slower production of xylonate in *S. cerevisiae*, with the conversion of xylonolactone to xylonate also happening spontaneously (Nygård, et al., 2014; Buchert & Viikari, 1988). Therefore, it was concluded that not expressing XylC was favorable since the strain could still grow on a mix of glucose and xylose but the xylonate production was slower which decreases the level of cell death.

When XylD from *C. crescentus* was introduced into *S. cerevisiae* in a later study, it turned out to have a low activity. This was believed to originate from poor enzyme maturation because of insufficient Fe-S cluster loading (Salusjärvi, et al., 2017). Therefore, the iron regulon repressor *FRA2* was deleted, which was expected to lead to more Fe-S clusters being available in the cytosol for the enzyme. The deletion indeed resulted in a 27-fold increase in the XylD activity (Salusjärvi, et al., 2017).

A recent study showed it possible to functionally introduce the Weimberg pathway into *S. cerevisiae* and thereby enable growth from xylose. The successful strain had *FRA2* deleted and introduced one copy of the gene *xylB* from *C. crescentus*, four copies of the genes *xylD* and *xylX* from *C. crescentus* as well as four copies of *ksaD* from *C. glutamicum*. The study showed a 33-fold increase in enzyme activity when comparing KsaD to XylA and the working strain had a biomass yield in C-mol biomass/ C-mol substrate (xylose+ glucose) of 0.41, calculated from start to end point in a bioreactor cultivation (Borgstrom, et al., 2019). However, it was unclear whether the copy number of each gene and the ratio between the enzyme activities was optimal for growth on xylose.

1.2. CRISPR/Cas9

CRISPR stands for clustered regularly interspaced short palindromic repeats. As the name indicates, the CRISPR loci consists of short repeats which are separated by variable sequences, so called spacers. In 2007, the CRISPR and the associated Cas genes were found to work as an immune system in prokaryotes, providing protection and resistance against bacteriophages. It was found that these spacers derived from the genomic sequence of previous viral encounters (Barrangou, et al., 2007). These integrated spacers prevent the prokaryote from being infected again. The combination of these spacers and the Cas proteins create an antiviral reponse if the virus is encountered again (Brouns, et al., 2008).

After the discovery of this imune system, further advances have been made and the CRISPR/Cas9 system has become a useful genitic engineering tool. In 2013 the system was used in *S. cerevisiae*. The Cas9 protein uses a guide RNA (gRNA) to cut at a specific sequecne in the genome. The double stranded break that occurs can then be repaired by homologous recombination or by non-homolougs end joining, the former is used in *S. cerevisiae* (DiCarlo, et al., 2013).

1.3. Aim

In the present degree project, the objective was to further optimize the utilization of xylose in *S. cerevisiae* carrying the Weimberg pathway. This was done by regulating the expression level of the genes involved in the Weimberg pathway (*xylB*, *xylD*, *xylX* and *ksaD*) using CRISPR/Cas9. Strains with different copy numbers of the genes were constructed and the enzyme activities were measured. The strains with the best activities were further analyzed for biomass formation and metabolite production. Comparison between the strains were conducted in order (i) to determine whether all heterologous genes were needed in *S. cerevisiae* and (ii) to find the optimal copy number of the genes for the Weimberg pathway to work in *S. cerevisiae*. The impact of deleting the *FRA2* gene was also studied. Consequently, the research questions (RQ) for this degree project were:

RQ1: What are the enzyme activities needed for the Weimberg pathway to work in *Saccharomyces cerevisiae*?
RQ2: How many copies are needed of *xylD*, *xylX* and *ksaD*?

This project was expected to give a better understanding of how *Saccharomyces cerevisiae* can be engineered in order to improve its utilization of xylose. As well as, how the Weimberg pathway can be introduced into an organism to improve its ability to consume xylose. Furthermore, the project contributes to a more sustainable, efficient and profitable use of substrates in any industry using microorganisms for production.

2. Material and method

2.1. Plasmids, strains and media

The plasmids used in this project can be found in table 1, the yeast strains used can be found in table 2 and the primers in appendix table A. *Escherichia coli* transformants were plated on agar plates with LB medium (5 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl) supplemented with 100 mg/L ampicillin. *E. coli* cultivations were incubated in 37 ℃ on a shaking table in 15 mL falcon tubes containing 5 mL LB medium supplemented with 100 mg/L ampicillin. For yeast cultivations, cells were grown at 30℃ on a shaking incubator in shake flasks at 10% of total volume in yeast extract peptone dextrose medium (YPD medium) containing 20 g/L glucose, 20 g/L peptone and 10 g/L yeast extract, supplemented with 200 mg/L of geneticin. Transformed cells, using the CRISPR/Cas9 system, were plated on agar plates with YPD medium supplemented with 200 mg/L geneticin and 100 mg/L nourseothricin to select for the Cas9-kanMX and gRNA plasmids. *Caulobacter crescentus* was cultivated in 250 mL shake flasks at 30℃ on a shaking incubator in 25 mL PYE medium (2 g/L peptone, 1 g/L yeast extract, 0.8 mM $MgSO_4$, 0.5 mM $CaCl_2$) supplemented with or without xylose 2 g/L.

Table 1. Plasmids used in this project, including plasmid name, relevant genotype and reference.

Plasmid name	Relevant genotype	Reference
Cas9-kanMX	*TEF1p-Cas9-CYC1t*_kanMX. Geneticin marker.	(Stovicek, et al., 2015)
pCfB3043	gRNA targeting site XI-1. Nourseotricin marker.	(Jessop-Fabre, et al., 2016)
pCfB3048	gRNA targeting site XII-2. Nourseotricin marker.	(Jessop-Fabre, et al., 2016)
pCfB3052	gRNA targeting site X-4, XI-3 and XII-5. Nourseotricin marker.	(Jessop-Fabre, et al., 2016)
pCfB3053	gRNA targeting site X-2, XI-5 and XII-4. Nourseotricin marker.	(Jessop-Fabre, et al., 2016)
pCfB3043-gRNA2	gRNA targeting site *FRA2*. Nourseotricin marker.	This study
pCfB2899	Flanks for integration site X-2	(Jessop-Fabre, et al., 2016)

pCfB2904	Flanks for integration site XI-3	(Jessop-Fabre, et al., 2016)
pCfB2909	Flanks for integration site XII-5	(Jessop-Fabre, et al., 2016)
pCfB3035	Flanks for integration site X-4	(Jessop-Fabre, et al., 2016)
pCfB3036	Flanks for integration site XI-1	(Jessop-Fabre, et al., 2016)
pCfB3037	Flanks for integration site XI-5	(Jessop-Fabre, et al., 2016)
pCfB3039	Flanks for integration site XII-2	(Jessop-Fabre, et al., 2016)
pCfB3040	Flanks for integration site XII-4	(Jessop-Fabre, et al., 2016)
pCfB3036-*ksaD*	*ksaD*_XI-1	This study
pCfB3035-*ksaD*	*ksaD*_X-4	This study
pCfB2904-*ksaD*	*ksaD*_XI-3	This study
pCfB2909-*ksaD*	*ksaD*_XII-5	This study
pCfB3039-*xylX*	*xylX*_XII-2	This study
pCfB2899-*xylX*	*xylX*_X-2	This study
pCfB3037-*xylX*	*xylX*_XI-5	This study
pCfB3040-*xylX*	*xylX*_XII-4	This study
pCfB2899-*xylD*	*xylD*_X-2	This study
pCfB3037-*xylD*	*xylD*_XI-5	This study
pCfB3040-*xylD*	*xylD*_XII-4	This study
YIpAGS8	*xylD*	(Wasserstrom, et al., 2018)
pRS303N	Nourseotricin marker.	(Taxis & Knop, 2006)

Table 2. The strains used in the project, including their strain name, relevant genotype, encoded enzymes introduced into the yeast and reference.

Strain name	Relevant genotype	Encoded enzymes introduced into yeast	Reference

CEN.PK 113-7D	MATa *URA3 HIS3 LEU2 TRP1 MAL2-8c SUC2*	-	(Entian & Kötter, 2007)
TMB4512	CEN.PK 113-7D, *Δgre3::xylB*	XylB	(Wasserstrom, et al., 2018)
TMB CB 010	TMB4512, XI-1::*ksaD*	XylB, KsaD	This study
TMB CB 011	2, X-4::*ksaD*, XI-3::*ksaD*, XII-5::*ksaD*	XylB, 4xKsaD	This study
TMB CB 012	3, XII-2::*xylX*	XylB, XylX, 4xKsaD	This study
TMB CB 013	4, X-2::*xylX*, XI-5::*xylX*, XII-4::*xylX*	XylB, 4x(XylX, KsaD)	This study
TMB4586	CEN.PK 113-7D, Δ*gre3*::AGS8B, *fra2*::*hphNT1*, *xylA*::*ksaD*	XylB, XylD, XylX, KsaD	(Borgstrom, et al., 2019)
TMB CB 014	TMB4586, X-4::*ksaD*, XI-3::*ksaD*, XII-5::*ksaD*	XylB, XylD, XylX, 4xKsaD	This study
TMB CB 015	7, X-2::*xylD*, XI-5::*xylD*, XII-4::*xylD*	XylB, XylX, 4x(XylD, KsaD)	This study
TMB CB 016	7, X-2::*xylX*, XI-5::*xylX*, XII-4::*xylX*	XylB, XylD, 4x(XylX, KsaD)	This study
TMB CB 017	3, *fra2*::*xylD*	XylB, XylD, 4xKsaD	This study
TMB CB 018	10, X-2::*xylD*, XI-5::*xylD*, XII-4::*xylD*	XylB, 4x(XylD, KsaD)	This study
TMB4590	TMB4586, X-2::*xylD*/*xylX*/*ksaD*, XI-5::*xylD*/*xylX*/*ksaD*, XII-4::LWA46	XylB, 4x(XylD, XylX, KsaD)	(Borgstrom, et al., 2019)

2.2. Amplification PCR and colony PCR

Amplification PCR reactions were performed with a mixture of 5X GC buffer, 10 mM dNTPs, 10 pmol/µL forward primer, 10 pmol/µL reverse primer, a few ng/µL template DNA, polymerase phusion (ThermoScientific, Waltham, MA, USA) and water. The program followed, including temperatures, times and number of cycles can be found in appendix table B. The annealing temperatures (A1 and A2) and the time of the second extension (E) were determined depending on the primers used in the PCR reaction. The reactions run, including the annealing temperatures and the extension time as well as the primers that were used can be found in appendix C.

When colony PCR was done for the yeast transformants, colonies were picked with sterile pipette tips and put into 15 µL of 0.02 M NaOH. The tubes were boiled at 99 °C for 10 min and centrifuged for 30 seconds at 13200 rpm. The PCR mixture contained 10X Dream Taq Buffer, dNTPs 10mM, forward primer 10 pmol/µL, reverse primer 10 pmol/µL, betaine 5M, Dream Taq and water. 23 µL of PCR mixture and 2 µL of the supernatant from a colony tube was used for each PCR reaction. The PCR program followed can be found in appendix table D. The reactions run, including the annealing temperature and extension time as well as primers used, can be found in appendix table E.

2.3. Gel electrophoresis

To verify that the fragment had been successfully amplified in the amplification PCR reaction and to verify the colony PCR, gel electrophoresis was run. A 0.8 % agarose gel stained with GelRed according to the supplier's instructions was used. 5 µL of sample were mixed with 1 µL of loading dye and loaded into the wells of the gel. A GeneRuler was loaded in one well. The gel electrophoresis was run for 30-40 min.

2.4. Plasmid construction

In order to use the CRISPR/Cas9 system (CRISPR toolbox from DTU (Jessop-Fabre, et al., 2016)), a gRNA plasmid and a plasmid (or plasmids) containing donor DNA are needed. In the present degree project, the gRNA plasmids to construct all strains except for strain TMB CB 017 were already available at the department. The plasmids containing the donor DNA were all constructed in the present study.

2.4.1. gRNA plasmid construction

Four different gRNA plasmids for the construction of TMB CB 017 were made in order to increase the chance of one of them working. For TMB CB 017, one copy of *xylD* was to be inserted and *FRA2* deleted. Therefore, the gRNA targeted the *FRA2* sequence in *S. cerevisiae*. The gRNA sequences were designed in Benchling using the CRISPR tool to find sequences that were next to a PAM site and had low off-target rating. Four primers were designed, all containing

a sequence annealing to a plasmid named 3043, but each with a unique sequence annealing to the *FRA2* sequence. PCR with the primers and the plasmid 3043 were done to get a linear fragment. The PCR product was cut with *Dpn*I to get rid of the 3043-template plasmid. *Dpn*I only cuts methylated DNA, the 3043 plasmid is methylated but the PCR fragments are not. PCR purification was made and followed by ligation with T4 DNA Ligase (ThermoScientific, Waltham, MA, USA) to ligate the fragment into plasmids, according to the supplier's instructions. The plasmids were then transformed into *E. coli*, using the Inoue method for ultracompetent cells (Im, et al., 2011) and the transformants were plated. A few of the colonies that grew were selected for overnight cultivation. The next day, plasmid purification was made and the DNA concentration was measured using BioDrop DUO (Biodrop Ltd, Cambridge, England). The plasmids were then used in the yeast transformation.

2.4.2. Donor DNA plasmids

To construct the donor plasmids, the donor DNA was first amplified using amplification PCR. The donor DNA consists of a flank, promoter, ORF (*xylD*, *xylX* or *ksaD*), terminator and another flank. The flanks are matching the part of the genome where the donor DNA are to be inserted. Primers were designed to amplify the fragment. At the end of each primer a restriction enzyme sequence was added as well as a few extra base pairs. The restriction enzymes used were *Sda*I, *Asc*I and *Sac*I. Details on which primers were used in each amplification PCR can be found in appendix table C and all the primers can be found in appendix table A.

The plasmids pCfB3036, pCfB3035, pCfB2904, pCfB2909, pCfB3039, pCfB2899, pCfB3037 and pCfB3040 were used as the plasmids into which the donor DNA was ligated. The PCR amplified donor DNA fragment and the plasmid were cut with the same restriction enzymes. Gel electrophoresis was ran to verify that the plasmids were cut, after which gel extraction was performed using a gel extraction kit (GeneJET Gel Extraction Kit, ThermoScientific, Waltham, MA, USA) to purify the cut vector. The cut donor DNA fragments were ligated into the plasmid vectors and the plasmids were used to transform *E. coli*. Colony PCR was then done on a few of the grown colonies to verify that the donor DNA was present in the plasmid. The positive colonies were grown overnight. The next day the plasmids were purified using a plasmid purification kit (GeneJET Plasmid Miniprep Kit, ThermoScientific, Waltham, MA, USA), the DNA concentration were measured using BioDrop DUO (Biodrop Ltd, Cambridge, England) and the plasmids were sent to be sequenced to make sure that no mutation occurred in the donor DNA.

2.5. Strain construction

For the transformations, the CRISPR/Cas9 system was used and the high efficiency LiAc protocol (Gietz & Schiestl, 2007) was followed, with the exception that 40 µL DMSO was added prior to heat shocking at 42°C for 20 seconds.

Firstly, for the CRISPR/Cas9 system to work, a Cas9 plasmid was used to transform the background strains. The other strains were constructed from these background strains by transforming them with different gRNA plasmids and donor DNA from plasmids (cut using the restriction enzyme *Not*I). For more detailed info about what plasmids were used see appendix table F. The transformed cells were plated on plates containing YPD, G418 and cloNAT to select for the Cas9 and gRNA plasmids. Two negative controls were made with water and gRNA and a positive control was made with the plasmid PRS303N.

After growth, colony PCR was performed with two to six colonies to verify that the correct gene(s) had been integrated and that they were in the right position. The growth of the constructed strain was compared to the background strain by measuring the OD every hour for eight hours and once again at 24 hours. Pre-cultures were made in YPD and cultivations in 250 mL baffled shake flasks with 25 mL YNB medium 6.7 g/L (including ammonium sulfate 5 g/L), glucose 5 g/L and 25 mM potassium phthalate buffer (pH 5.5). If the growth rate was acceptable the strain was plated on YPD+G418 to lose the gRNA plasmid. Around 50 colonies from this plate were replica plated on YPD+G418+cloNAT and YPD+G418 to verify the loss of the gRNA plasmid. A successfully transformed colony, one that grew only on YPD+G418 and not on YPD+G418+cloNAT, was kept in a 20% glycerol-stock at -80 ℃.

2.6. *In vitro* enzyme activity measurement

A single cell colony was inoculated in 5 mL YPD and incubated in a 30 ℃ shaking incubator overnight. The cells were harvested, washed with 1 mL water and weighed by weighing an empty tube, followed by adding the cells and centrifuging the tube to remove the supernatant after which the tube was weighed once again. Knowing the cell mass, 2.5 μL of Y-PER (Pierce, Rockford, IL, USA) was added per mg of cells to extract the protein, according to the manufacturer's instructions.

The total protein concentration was determined with Bradford assay using bovine serum albumin, BSA, as a standard. The standards were measured in quadruplicates while the samples were measured in biological duplicates and technical duplicates.

For all enzyme activity measurements, a mixture of 0.1 M Tris-HCl pH 8.0, 10 mM $MgCl_2$ and 2 mM NAD^+ were used. All measurements were performed in duplicates, both biological duplicates and technical duplicates. The enzyme activity of XylB, KsaD and XylD+XylX+KsaD was determined by measuring the rate of NADH formation, through optical density at 340 nm using a Multiskan Ascent reader (Thermo Electro Corporation, Finland). For XylB and KsaD, xylose (0.1 M) and glutaraldehyde (50 mM), respectively, were used as a substrate and the NADH formation rate was measured. For the coupled assay (XylD+XylX+KsaD), xylonate was

used as a substrate and pre-studies were made with *Caulobacter crescentus* in order to determine how much xylonate was necessary. It was decided that 80 mM would be used.

2.7. Aerobic cultivation in shake flask

The cultivation in shake flasks were performed in 250 mL baffled Erlenmeyer flasks and incubated at 30 °C in a shaking incubator. The flasks contained 50 mL of defined medium consisting of YNB 6.7 g/L (including ammonium sulfate 5 g/L), glucose 5 g/L, xylose 10 g/L and 25 mM potassium phthalate buffer (pH 5.5). Cells from a pre-culture were used to inoculate the flask to an initial OD_{620nm} of 0.1. The pre-culture was grown over night in a shaking incubator at 30 °C and consisted of one colony picked from a plate that was inoculated in 10 mL YPD in a 50 mL falcon tube. The cells were pelleted by centrifugation at 4000 rpm for 5 minutes, resuspended in water and transferred to a shake flask. Samples were taken at hours 0, 7, 12, 24, 31, 48 and 54. Samples for OD and HPLC analysis of metabolites were taken at all sample points. Viability using flow cytometry was measured at hours 0, 24, 31 and 54. Microscopy was done at hours 7 and 54. The CDW was measured at hour 54.

Xylose cultivations were performed in 500 mL baffled Erlenmeyer flasks, incubated at 30 °C in a shaking incubator. The flasks contained 50 mL of defined medium consisting of YNB 6.7 g/L (including ammonium sulfate 5 g/L), xylose 10 g/L and 25 mM potassium phthalate buffer with pH 5.5. Cells from a pre-culture were used to inoculate a flask to an initial OD_{620nm} of 0.1. The pre-cultures were grown in YPD and the cells were harvested by centrifugation and resuspended in water before inoculation.

2.8. OD measurement and CDW

The optical density, OD, was measured using a spectrophotometer (Ultrospec 2100 Pro spectrophotometer, Amersham Biosciences Corp., USA) at 620 nm. Samples with an absorbance unit over 0.3 were diluted. The measurements were performed in duplicates.

Cellular dry weight, CDW, was determined using filter papers with a pore size of 0.45 µm. The filter papers were dried in a microwave at 350 W for 4 minutes before being weighed. 5 mL of cell sample, from the 54 hour sample in the shake flask experiment, were poured over the filter paper. 15 mL water was used for washing. The papers were then dried in a microwave at 350 W for 8 minutes and left in a desiccator for cooling before being weighed again.

2.9. HPLC and UPLC

Samples of 2 mL were taken from the shake flask experiment and put in 2 mL Eppendorf tubes. The tubes were centrifuged for 2 min and the supernatant was put in new 2 mL Eppendorf tubes that were stored in a freezer until used for analysis.

For the HPLC measurements, a standard series was prepared with glucose, xylose, xylitol, glycerol, acetate and ethanol. The concentrations of each component were the same in each standard and the concentrations prepared were 6 g/L, 4.5 g/L, 3 g/L, 1.5 g/L, 0.75 g/L and 0.375 g/L. The samples were diluted twice by adding 400 µL of sample and 400 µL of water into HPLC vials which were put in a 48-vial plate. 800 µL standards were used and put in vials. Two plates were prepared with the samples and standards and run in the HPLC (Alliance 2695, Waters Milford, MA, USA). A flow rate of 0.6 mL/min and a sample volume of 20 µL were used. The analyzing time was set to 30 min per sample. 5 mM H_2SO_4 were used as mobile phase. An Aminex HPX-87H column (300*7.8mm, 9 µm, BIO-RAD, Hercules, CA, USA) was used at 60 °C.

Xylonate was measured using UPLC (Acquity H-Class, Waters, Milford, MA, USA). A Waters UPLC BEH Amide column (2.1*100mm, 1.7 µm) was used as well as a gradient method specialized for detecting xylonate in the presence of glucose and xylose (Almqvist, et al., 2017).

2.10. Flow cytometry - cell viability

Samples were taken from the shake flask experiment. 300 µL of cell sample with an OD of maximum 1 was mixed in an Eppendorf tube with 300 µL of a mixture of phosphate buffered saline (PBS) and propidium iodide (PI) with 1.32 µg/mL. The tubes were incubated in the dark for 10 minutes before being analyzed with a BD Accuri C6 flow cytometer (Becton-Dickinson) to see the percentage of live and dead cells.

3. Result and discussion

3.1. Constructing a pallet of strains using CRISPR/Cas9

The successful integration of the Weimberg pathway into *Saccharomyces cerevisiae*, enabling it to grow on xylose in the presence of glucose, was obtained in a strain containing one copy of the gene encoding for the enzyme XylB and four copies each of the genes encoding XylD, XylX and KsaD. In addition the *FRA2* gene was deleted in the strain to increase the XylD activity (Borgstrom, et al., 2019). Although this combination has been proven to work in *S. cerevisiae*, new questions arised concerning further optimisation. Are four copies of *xylD*, *xylX* and *ksaD* necessary for the pathway to work? Are both XylD and XylX necessary? The second question is of interest since it recently has been shown that either XylD or XylX are needed in *C. glutamicum* for the Weimberg pathway to work (Brüsseler, et al., 2019).

In an attempt to answer these questions, a pallet of strains containing different copy numbers of the genes encoding for the enzymes involved in the Weimberg pathay was constructed and analysed in the present master thesis project. Two background strains, TMB4512 and TMB4586, were used to construct nine different strains. An overview of the pallet of strains, including the copy numbers of the different genes, can be seen in table 3.

Table 3 . Overview of the strains, showing their TMB number and their gene copy number of xylB, xylD, xylX and ksaD.

	TMB number											
	4512	CB 010	CB 011	CB 012	CB 013	4586	CB 014	CB 015	CB 016	CB 017	CB 018	4590
xylB	X	X	X	X	X	X	X	X	X	X	X	X
xylD +fra2Δ	-	-	-	-	-	X	X	4X	X	X	4X	4X
xylX	-	-	-	X	4X	X	X	X	4X	-	-	4X
ksaD	-	X	4X	4X	4X	X	4X	4X	4X	4X	4X	4X

The strains were constructed using the CRISPR/Cas9 system. For the CRISPR/Cas9 system to work, the organism has to produce the Cas9 protein and have a gRNA and a donor DNA. Therefore, in order to even construct the pallet of strains, both plasmids containing donor DNA and plasmids containing gRNA had to be constructed.

3.1.1. Plasmid construction

A majority of the donor DNA plasmids had already been made and were available at the department with the exception of the *xylD* plasmids that were needed for the construction of TMB CB 015 and CB 018. These were constructed according to the description in the Material and Method part and were cut prior to successfully being used in the transformations.

Regarding the gRNA plasmids, all were available with exception of the gRNA needed for the construction of TMB CB 017. For this, four different gRNA plasmids were created to guide Cas9 in cutting the *FRA2* gene, as detailed in the Material and Method part. Three of the four gRNA plasmids were successfully constructed and were able to create a cut that was repaired with a donor DNA and whose position was verified with a colony PCR. The fourth plasmid did not result in any growth after yeast transformation, which probably is due to inefficient targeting by the gRNA and thus there was no reparation of the cut with the donor DNA. After yeast transformation with the plasmids a colony PCR was made on the transformants. The transformants with the gRNA2 plasmid were found to have a higher number of clones with the correct insert, where the *xylD* cassette (upstream flank, promoter, *xylD*, terminator and downstream flank) had been inserted and the *FRA2* gene deleted, than the other two plasmids. Therefore, it was decided to use the transformants from using the gRNA2 plasmid in the experiments to follow.

3.1.2. Strain construction and validation

When the plasmids were ready, the transformations were performed. First, TMB4512, which only has one copy of *xylB*, was transformed with one copy of *ksaD*, creating strain TMB CB 010. As the copy number of the lower Weimberg pathway has been shown to affect pathway flux (Borgstrom, et al., 2019), three copies of *ksaD* were inserted into TMB CB 010 and TMB4586, creating TMB CB 011 and TMB CB 014, see figure 2. These two strains were two platform strains to which either *xylD* and/or *xylX* could be added with different copy numbers to create a variety of strains. TMB CB 011 was transformed with one copy of *xylD* and one copy of *xylX* to create TMB CB 012 and TMB CB 017 respectively. Three copies of *xylD* were added to TMB CB 012 and TMB CB 014, creating TMB CB 013 and TMB CB 016. Three copies of *xylX* were added to TMB CB 017 and TMB CB 014, creating TMB CB 018 and TMB CB 015.

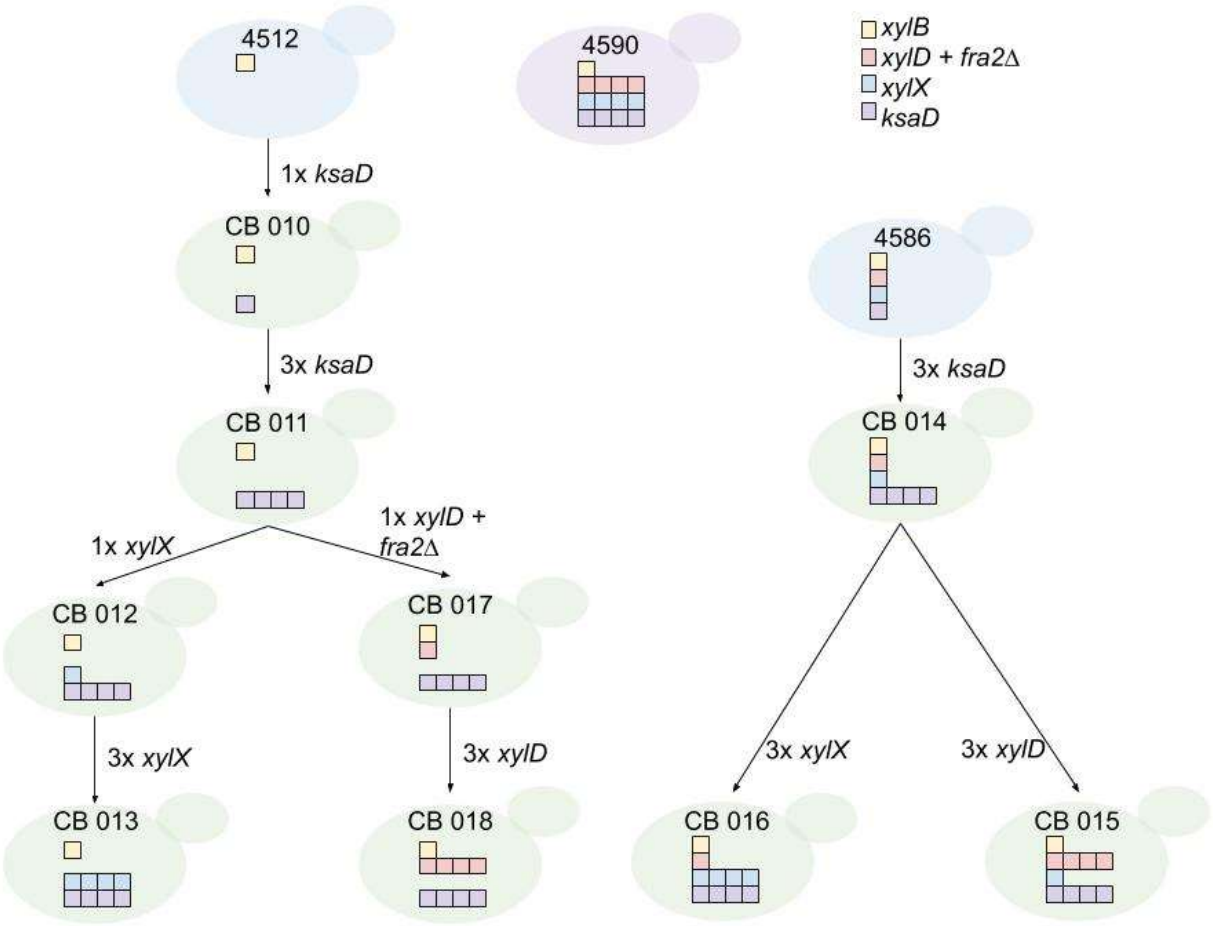


Figure 2. Overview of how the strains were constructed from two background strains. The background strains are shown in blue and the constructed strains in green. TMB4590, containing a functional Weimberg pathway (Borgstrom, et al., 2019), is shown in purple.

When the transformations were done, colony PCR was used to confirm that the genes were introduced and integrated in the right position. In the colony PCR one primer was annealing to the gene that was supposed to have been integrated and a second primer was annealing to the part of the genome where it was supposed to be integrated. This means that there will be a PCR product of the correct length only if the genes are integrated in the correct place.

Figure 3 shows a representative result of colony PCR from TMB CB 011 and TMB CB 014. Two colonies of each strain were tested and since three genes were integrated, there are three assays per colony. The expected amplification length was approximately 1500 base pairs for each assay and as can be seen in figure 3, all assayed colonies were found to be positive. For the rest of the constructed strains the integration was successful as well with exception of TMB CB 015 and TMB CB 018. For TMB CB 015 four colonies were tested and for TMB CB 018 six colonies. For these strains, only one assay worked, namely that of integration in XII-4 (plasmid pCfB3040-

xylD). Due to time constraints, one colony of each strain was chosen at random and used in subsequent analysis.

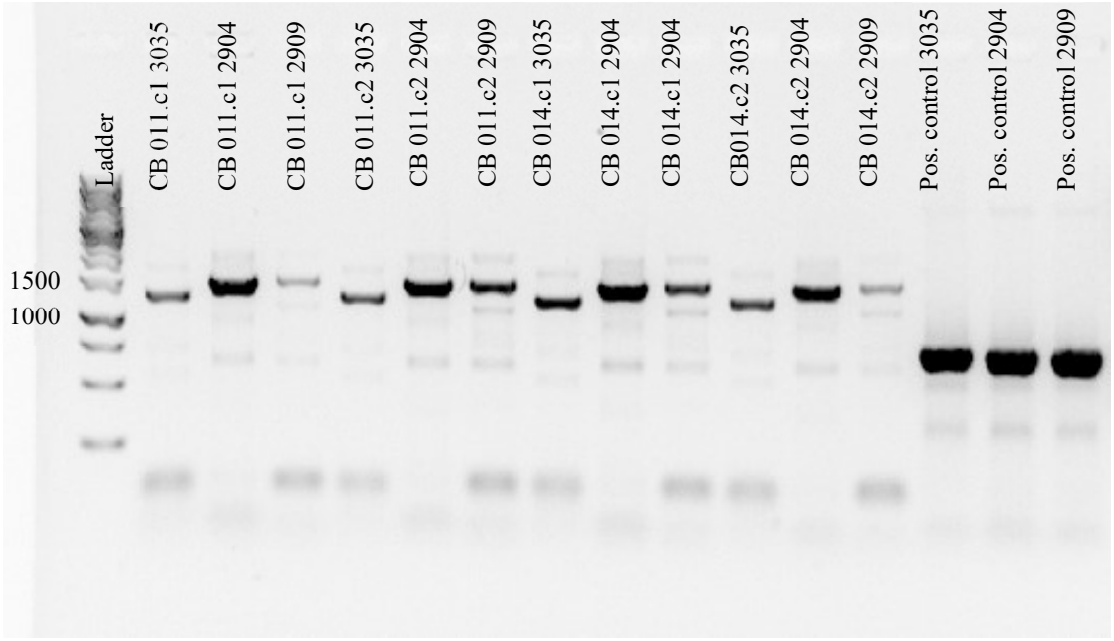


Figure 3. Agarose gel showing the results from the colony PCR of TMB CB 011 (two clones) and TMB CB 014 (two clones). Includes three positive controls for the PCR technique, i.e. using one primer annealing to the gene to be integrated and another annealing to the plasmid, one for each gene integrated.

3.2. XylB and KsaD enzyme activity

Once the strains had been constructed, enzyme activity measurements were done to compare the strains. Measuring the activity of XylB (xylose dehydrogenase) and KsaD (α-ketoglutarate semialdehyde dehydrogenase) had previously been done in *S. cerevisiae* and therefore, a working assay already existed. The assays were first controlled by measuring the activity of CEN.PK 113-7D (carrying none of the Weimberg genes), TMB4512 (*xylB*), TMB4586 (*xylB*, *xylD*, *xylX* and *ksaD*) and TMB4590 (*xylB*, 4x(*xylD*, *xylX*, *ksaD*)). This resulted in no activity for the CEN.PK 113-7D strain, only XylB activity for 4512 and both XylB and KsaD activity for 4586 and 4590, see figure 4. In addition, the KsaD activity for 4590 was around four times higher than for 4586. These two assays were verified to work and therefore the activity of XylB and KsaD were measured in the same way for the strains TMB CB 010, CB 011 and CB 014.

These three strains were chosen since they were the ones transformed with *ksaD* genes and were the parents of all the other strains. Since the later ones had the same *xylB* and *ksaD* copy number it was found unnecessary to measure the KsaD activity for them too, as it theoretically should be conserved. For the three strains, the activity of XylB was calculated to be between 0.18 and 0.21 $\mu mol*min^{-1}*mg^{-1}$. When comparing XylB enzyme activity for the three constructed strains to the background strains, TMB4512 and TMB4586, it was found that 4586 had a lower activity for

XylB, which was not expected. TMB4512, however, had a similar XylB level as the three strains (figure 4).

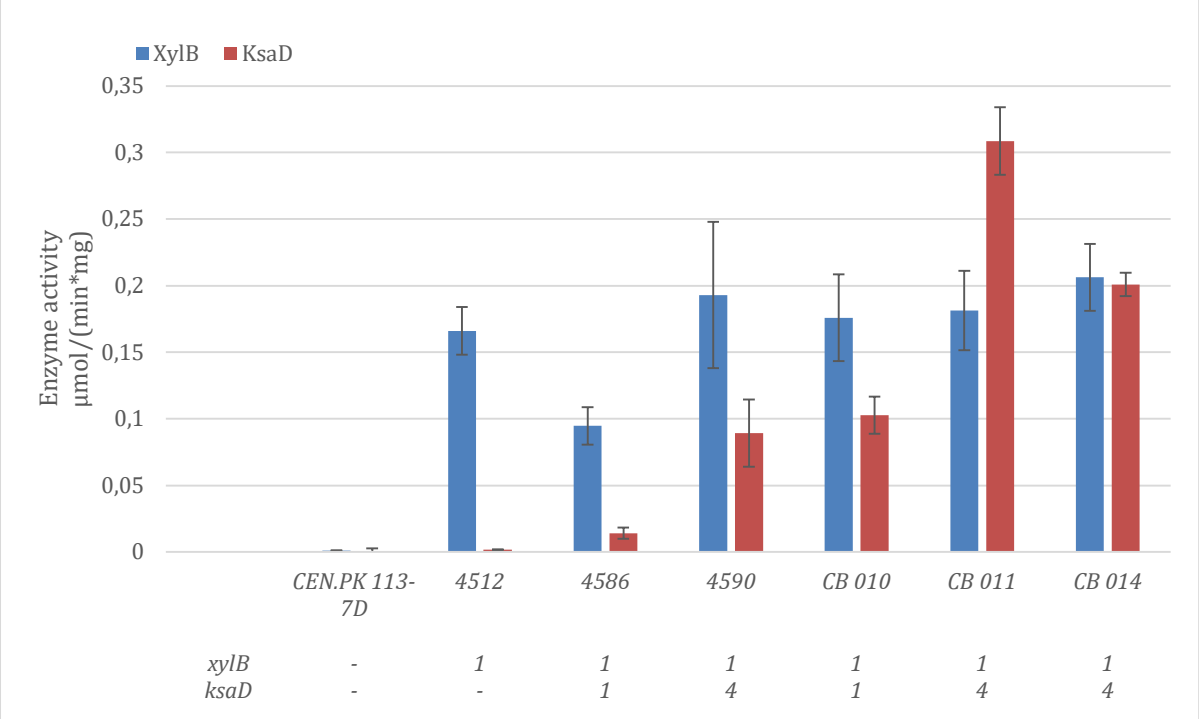


Figure 4. Enzyme activity of XylB (blue bars) and KsaD (red bars) in the control strains CEN. PK 113-7D, TMB4512, TMB4586 and TMB4590 as well as the constructed strains TMB CB 010, 011 and 014.

The KsaD varied more between the three constructed strains, from TMB CB 010 (one gene copy) having the lowest activity of 0.10 µmol*min^{-1}*mg^{-1} to TMB CB 011 (four gene copies), having an activity of 0.31 µmol*min^{-1}*mg^{-1}, see figure 4. Although the constructed strains with four copies of the *ksaD* gene, TMB CB 011 and TMB CB 014, had a higher activity than TMB CB 010, which only has one copy, it was not four times higher. In contrast, when measuring the backgrounds strains (while testing the assay), the activity of KsaD in TMB4590 (4x*ksaD*) was four times higher than TMB4586 (1x*ksaD*). Surprisingly enough, looking at figure 4 the KsaD activity in TMB4590 was almost the same as for TMB CB 010 even though 4590 has four copies of *ksaD* and TMB CB 010 only has one. This could be due to that the genes were not integrated in the same way in the two strains. In 4590 three copies were integrated at one integration site, whereas in the constructed strains from the present project, each gene was integrated at their own integration site.

3.3. Combined enzyme activity assay for XylD/XylX/KsaD

To further compare the strain enzyme activity, it was decided to conduct a coupled assay, measuring the combined activity of XylD, XylX and KsaD. It is not possible to measure the activity of XylD or XylX alone since the product of the dehydration reaction is not available.

However, the combined activity could possibly be determined by using xylonate as a substrate and measuring the formation of NADH, which is produced in the last reaction catalyzed by KsaD (figure 1). Since this coupled assay is not commonly used for *S. cerevisiae* it was decided to first test it with the bacteria *Caulobacter crescentus*, which has the Weimberg pathway naturally, to try to find which concentration of xylonate was needed for the assay to be functional.

C. crescentus was cultivated in shake flasks but no activity was found in the cell extract. Xylose was added to the shake flasks cultivation to induce the expression of the *xyl* operon before harvesting. However, there was background activity and upon adding the substrate (10 mM, 20 mM or 40 mM xylonate) the activity stopped (data not shown). In a second experiment, a xylonate concentration of 20 mM and 80 mM was tried and cell extract from TMB4590 was also tested. To prove that KsaD activity was present in *C. crescentus*, glutaraldehyde was also used as a substrate. This resulted once again in background activity which was subsequently subtracted from the enzyme activity that was calculated after adding the substrate. The results, shown in figure 5, showed that *C. crescentus* indeed had a KsaD activity at around 0.044 µmol*min^{-1}*mg^{-1} but the coupled assay did not work and resulted in a very low activity for *C. crescentus*. Surprisingly, 4590 showed an activity around 0.025 µmol*min^{-1}*mg^{-1} for XylD+XylX+KsaD, from using 80 mM xylonate as a substrate. Using 20 mM xylonate did not result in any activity for 4590.

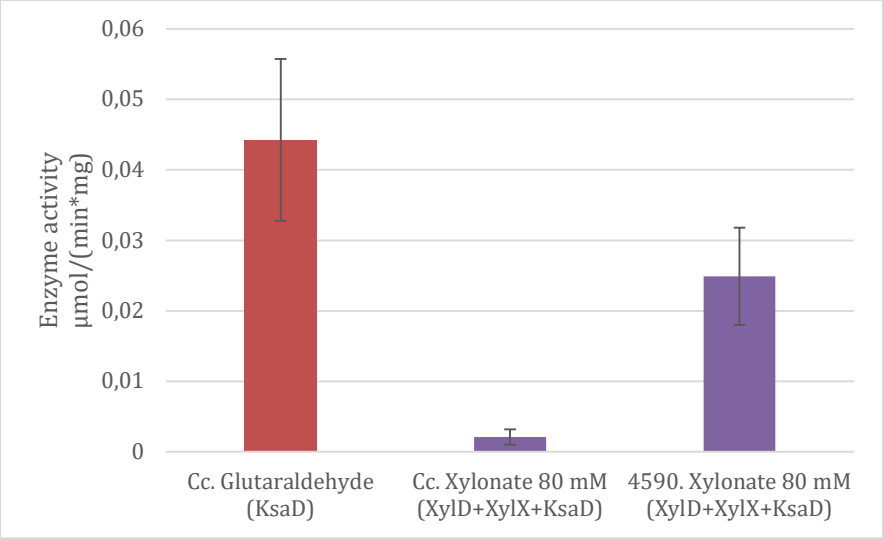


Figure 5. The enzyme activity from C. crescentus *and 4590 when testing the coupled assay for measuring the enzyme activity of XylD, XylX and KsaD together. The coupled assay for* C. crescentus *and TMB4590 is shown in purple bars. For control the KsaD activity in* C. crescentus *was measured, shown in red.*

It was unexpected that the coupled assay did not work in *C. crescentus* in these set of experiments as it has worked before, using the same concentration of xylonate (Almqvist, et al., 2018). One possible explanation may be the conditions under which the cells were grown prior to harvesting. When the assay worked the cells were harvested from a bioreactor, whereas in the

present study, the cells were inoculated in shake flasks from frozen stock and harvested from the shake flasks.

As the coupled assay worked for 4590 with 80 mM xylonate, it was decided to use this concentration and measure the activity in the constructed strains with 4590 as reference (figure 6). All constructed strains were measured with exception of TMB CB 010 since it is unlikely to have had any activity as it only carries one copy of *ksaD* and none of *xylD* and *xylX*.

In the first set of experiments, it was first found that TMB CB 016 (1x *xylD*, 4x *xylX*, 4x *ksaD*) had a distinctively higher combined enzyme activity than the rest of the assessed strains, 0.012 µmol*min^{-1}*mg^{-1} compared to not more than 0.003 µmol*min^{-1}*mg^{-1}. TMB CB 016 even had a higher activity than the control TMB4590 (4x *xylD*, 4x *xylX*, 4x *ksaD*). This suggested that four copies of *xylX* together with one copy of *xylD* and Δ*fra2* could give the most effective combined activity.

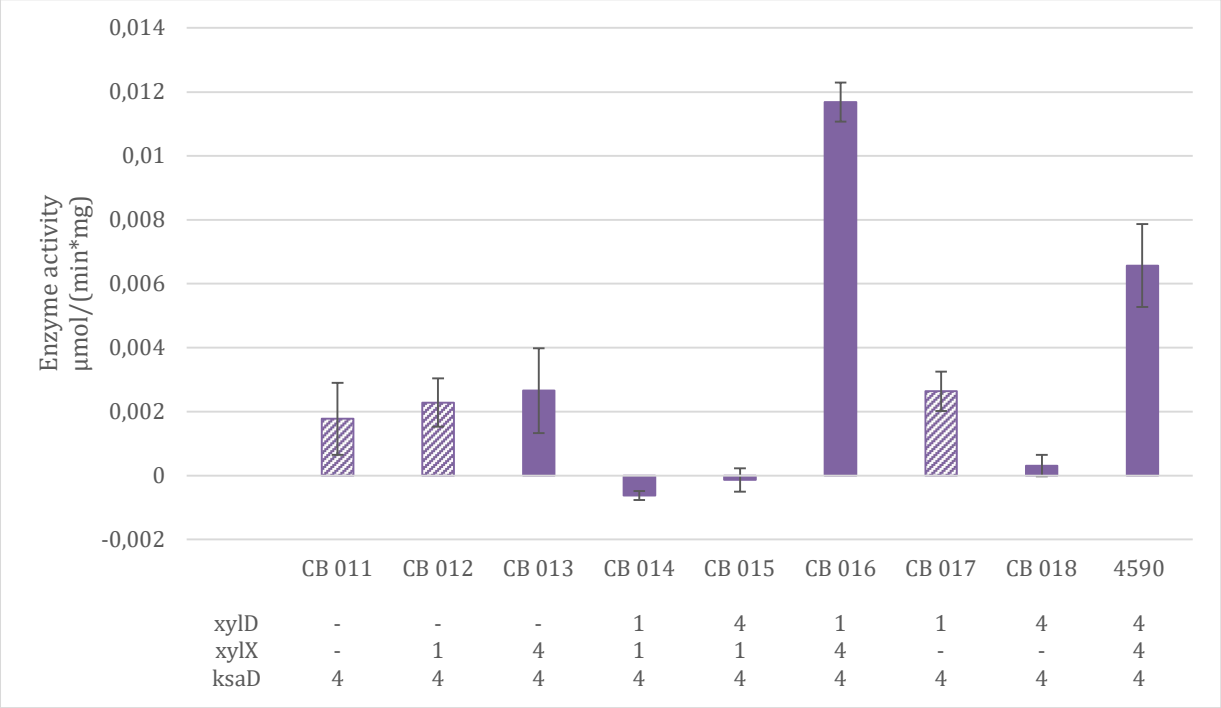


Figure 6. Results from the coupled assay, showing the enzyme activity of XylD+XylX+KsaD of the constructed strains, as well as 4590 as reference. The strains with striped bars were measured on one day and the strains with full bars were measured on another day.

The coupled assay results were found to differ between biological replicates. For instance, for TMB CB 016, one biological replicate displayed high enzyme activity whereas the second replicate showed none. The activity also varied from experiment to experiment for other strains, for instance in experiments where strains were divided into two separate days (one day measuring TMB CB 011, 012, 017 and TMB4590 and another day measuring TMB CB 013, 014, 015, 016, 018 and TMB4590). For example 4590 had an activity at around 0.025 µmol*min^{-1}*mg^{-1} in the

experiment with *C. crescentus*, no activity in the first experiment with the new strains and around 0.007 µmol*min^{-1}*mg^{-1} in the second experiment with the new strains.

For comparison, the cell extracts were used to look at the activity of KsaD (figure 7). The results suggested that the problems originated in the coupled assay and not in the cell samples as all strains showed activities ranging from 0.33 µmol*min^{-1}*mg^{-1} to 0.86 µmol*min^{-1}*mg^{-1}. These values differ from the original measurement showed earlier where the highest value was about 0.31 µmol/min^{-1}*mg^{-1}. This might be due to not reaching saturation when doing the earlier measurements, were the cell extract samples were not as dilutes as in this experiment. Normally that is a requirement when doing enzymatic assays but since in this case the assays are performed only to see a potential difference between the strains, not to get the precise number, it is accepted.

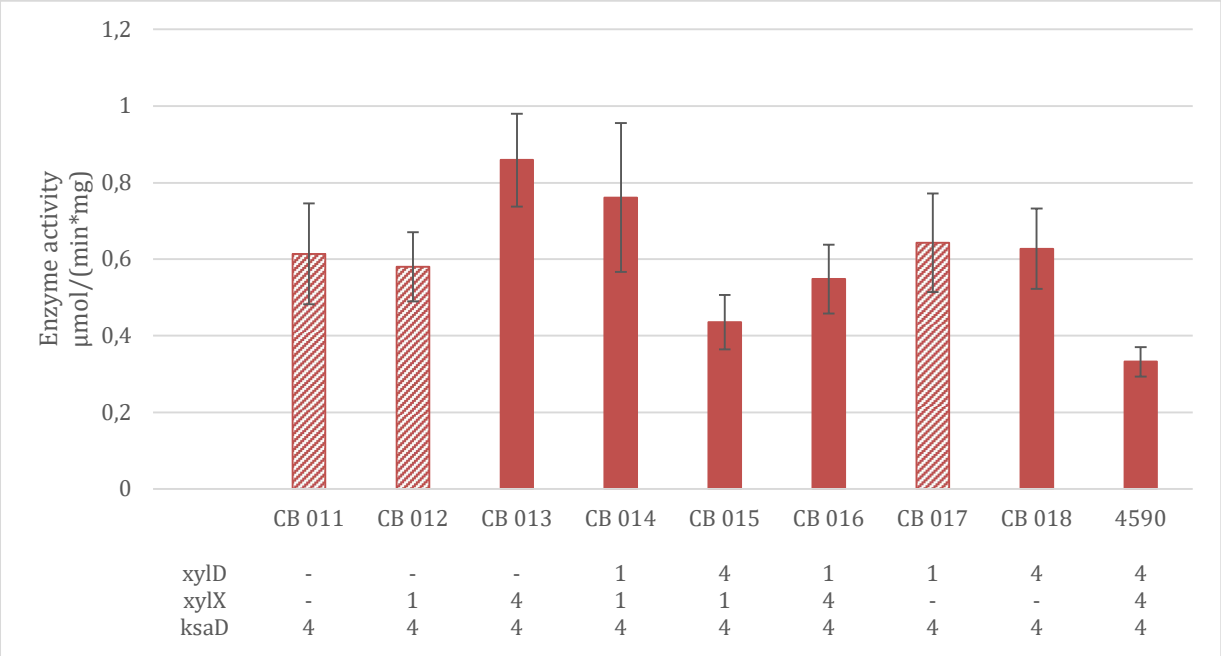


Figure 7. Results from the control of KsaD activity with glutaraldehyde as a substrate run together with the coupled assay of the constructed strains and 4590. The strains with striped bars were measured on one day and the strains with full bars were measured on another day.

In summary, it can be concluded that the coupled assay was not reliable. Despite this, the activity for TMB CB 0016 might be higher than for the other strains, indicating that three copies of *xylX* and one copy of *xylD* + *fra2*Δ might be important and possibly better than 4590 (*xylB*, 4x(*xylD*, *xylX*, *ksaD*)).

3.4. Growth from xylose as sole carbon source

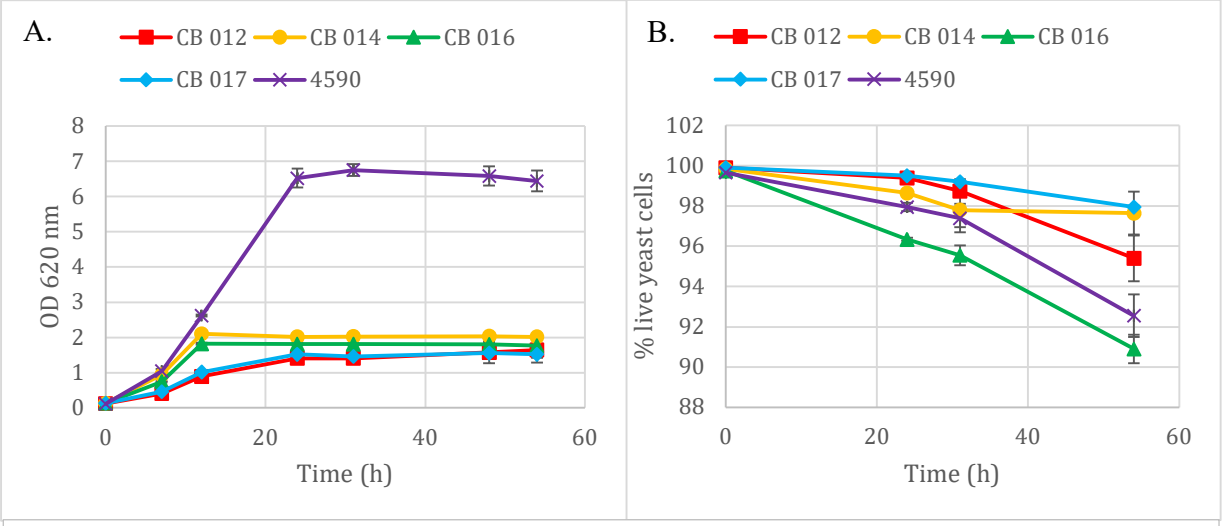
To test if any of the constructed strains could grow on xylose as sole carbon source, Falcon tubes were prepared with YNB defined medium, potassium phthalate buffered to pH 5.5 and supplemented with 10 g/L xylose. The tubes were inoculated to an initial OD of 0.1 with cells from a pre-culture of rich YPD medium (20 g/L glucose) and were incubated in a shaking

incubator at 30 °C. After two days the OD was measured, with only TMB CB 013 showing an increase in OD to around 0.6. One reason for why TMB CB 013 had grown could be that a small volume of YPD had also been transferred to the Falcon tubes during inoculation. To investigate this, 1 ml from the TMB CB 013 Falcon tube was transferred to a new tube with xylose, YNB and potassium phthalate. The OD was measured again for the original Falcon tubes and the new TMB CB 013 Falcon tubes after three days. TMB CB 013 OD remained unchanged, confirming the lack of growth on xylose. In parallel, an increase in OD was observed for TMB CB 0016 and 017. A possible reason for the growth in TMB CB 016 and TMB CB 017 could be due to a contamination from opening the tubes when taking the first sample, although this was done in a sterile hood. Therefore, it was decided to study the strains in shake flasks with only xylose. Measuring the OD after twelve days showed no increase in OD for any of the strains. TMB4590 and TMB CB 013 stayed at the same OD, around 0.1. The other strains all decreased in OD.

3.5. Poor growth of constructed strains not correlated to viability

Due to none of the strains being able to grow on xylose as the sole carbon source, a growth in shake flasks under controlled conditions experiment were conducted, using xylose and glucose as carbon source, in order to initiate growth on glucose and favor xylose consumption. In this experiment five strains were chosen, all having one copy of *xylB* and four copies of *ksaD*, but with varying copy number of *xylD* and *xylX*: TMB CB 012, 014, 016, 017 and 4590. Strain CB 016 (1x *xylD*, 4x *xylX*) was chosen due to its activity in the coupled enzyme activity assay and CB 014 (1x *xylD*, 1x *xylX*) was chosen to see any potential difference of copy number of *xylX*. CB 012 (1x *xylX*) and CB 017 (1x *xylD*) were chosen to possibly be able tot tell is it is enough for the Weimberg pathway to work with either XylD or XylX. 4590 was chosen to have as a reference.

OD was measured at all samples point to see how well the strain grew. A big difference in growth was observed between the constructed strains and TMB4590, see figure 8 A. 4590 reached an OD of around 6.5 whereas all the constructed strains had an OD of less than 2. The CDW was measured at 54 hours and showed a similar result. 4590 had a CDW of 2.33 g/L and the constructed strains TMB CB 012, 014, 016 and 017 had a CDW of 0.41 g/L, 0.50 g/L, 0.49 g/L and 0.43 g/L, respectively. TMB CB 014 and TMB CB 016 seem to have a similar growth and CDW, as did TMB CB 012 and TMB CB 017. The viability was measured at 0, 24, 31 and 54 hours. One theory was that the poor growth in the constructed strains might have been due to a low viability in these strains. However, strain 4590 had the second lowest viability (92.6%) and TMB CB 017, with the highest viability (98.0%), was one of the strains with the poorest growth, see figure 8 B.



	CB 012	CB 014	CB 016	CB 017	4590
xylB	1	1	1	1	1
xylD + fra2Δ	-	1	1	1	4
xylX	1	1	4	-	4
ksaD	4	4	4	4	4

Figure 8. Results in OD (A) and viability in % live yeast cells (B) for TMB CB 012 (red squares), CB 014 (yellow circles), CB 016 (green triangles), CB 017 (blue rhombs) and 4590 (purple crosses) during growth under controlled conditions in shake flasks with glucose and xylose.

To see if the cells were stressed during the growth and if this could be a possible explanation for the poor growth, cell morphology was investigated. At time 7 hours, it was assumed that none of the strains were stressed. Therefore, only TMB CB 012 was observed as reference, showing healthy cells with most of them budding, see figure 9 A. At time 54, TMB CB 016 and TMB CB 017 were looked at since they were the strains with lowest and highest viability, pictures of these cells are found in figure 9 B and C.

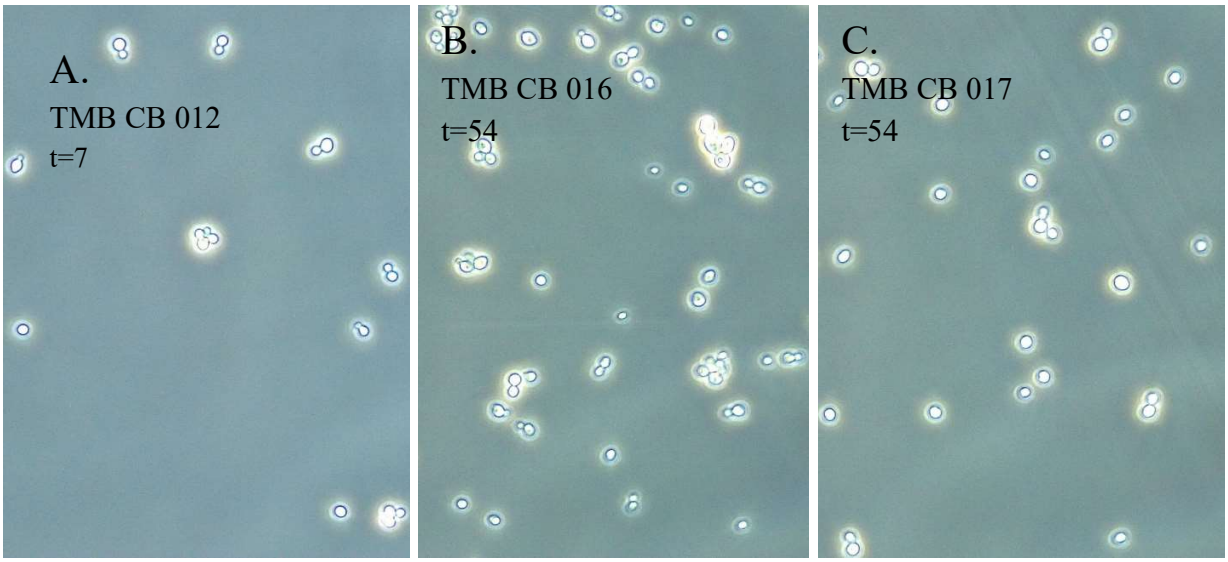


Figure 9. Pictures of cells looked at in microscope. The cells were grown under controlled conditions in shake flasks. A shows cells from strain TMB CB 012 at hour 7. B shows cells from TMB CB 016 at hour 54. C shows cells from TMB CB 017 at hour 54.

Looking at TMB CB 016, the strain with the lowest viability, there was a clear difference compared to the cells of CB 012. Some of the cells were clustering together, which can be a sign of stress and there were not as many cells budding. TMB CB 017 had the highest viability and in this case, the cells looked rather similar to TMB CB 012 with the main difference that most of the cells were now not budding. From these pictures, there seemed to be a correlation between viability and cell clustering. However, this appears to not be the reason for the impaired growth pattern seen in the constructed strains.

3.6. Strain re-validation

In the search for finding the reason behind the impaired growth of the constructed strains, another set of growth experiments were conducted on glucose. The objective was to verify that the inserted genes had no negative impact on growth. The constructed strains TMB CB 010, 011 and 014 were compared to their background strains TMB4512 and TMB4586. For TMB CB 010, three transformation clones were tested to see any possible differences between colonies of the same strain. The OD was measured once an hour for eight hours and then again after 24 hours. Similar results were obtained for all constructed strains. During the first eight hours, the constructed strains grew similar to the background strains. However, at 24 hours cell growth basically stopped whereas the OD of the background strains kept increasing, see figure 10. This was observed for both TMB CB 010 and TMB CB 011, which originated from TMB4512 and for TMB CB 014, which originated from TMB4586. This could suggest that the problem lies in the transformations rather than in one of the background strains.

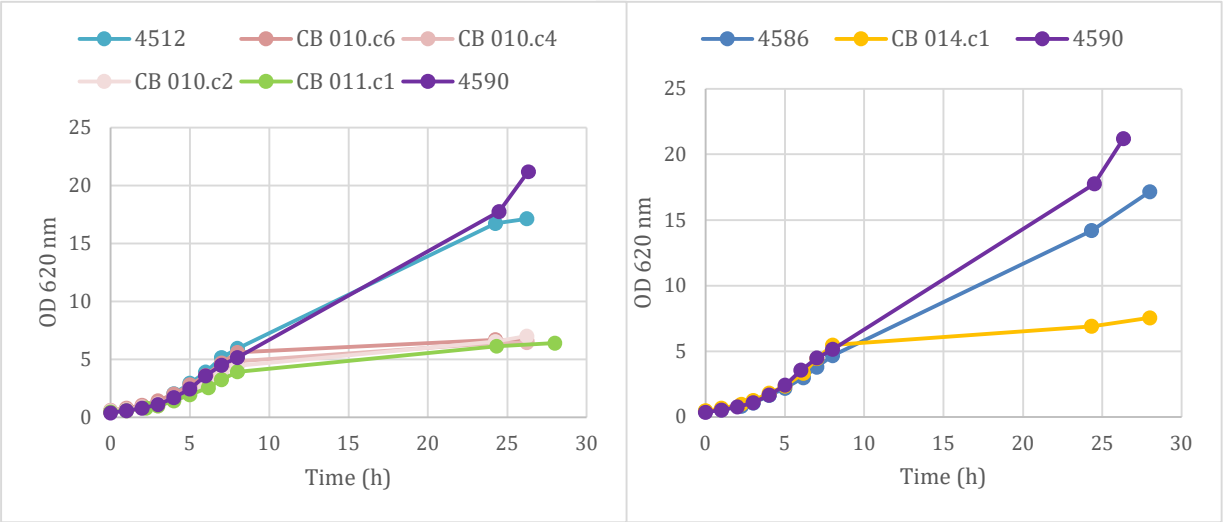


Figure 10. Results from OD measured once an hour for eight hours and again after 24 hours. Left picture shows the background strain TMB4512 and strains constructed from this, three clones of TMB CB 010 and one of TMB CB 011. The right figure shows the background strain TMB4586 and the constructed strain created from this, TMB CB 014. TMB4590 is included in both graphs as a reference.

TMB CB 010 was constructed by integration of one copy of *ksaD* whereas TMB CB 014 was constructed by introducing three additional copies of *ksaD* to TMB4586. TMB4586 showed no problem with growth despite having one copy of *ksaD*. However, this integration had been made using another method than in the present degree project. As mentioned before, in 4590 several genes were integrated at one integration site whereas in the present project each gene was integrated at their own integration site. In addition, in 4586 only the ksaD gene was integrated using CRISPR/Cas9, the rest were integrated using homologous recombination (Borgstrom, et al., 2019; Wasserstrom, et al., 2018). Therefore, we hypothesize that growth might have been affected by the way the genes were integrated in the present degree project. Since the results for all three colonies of TMB CB 010 were very much the same, it seems unlikely that the impaired growth was due to a bad clone, but rather something to the transformation and/or selection.

Since the growth problem can be traced back to the first transformation, the problem solving has to start here. In order to get better growth more clones should be tested in order to find a potentially better one. In addition, the transformation could be done again in an attempt to find a better clone, but it might be better to change the integration sites. It has been shown that different integration sites result in different expression levels due to that the transcription levels differ between the regions of the chromosomes. A lower expression level will result in a lower activity (Flagfeldt, et al., 2009). In addition, any genes already present at the integration site might be affected upon integrating a new one. That might have been the problem in this case, that the integration of the genes disturbed other genes, resulting in impaired growth. However, it did not seem to be any such genes at the used integration sites.

3.7. No xylose consumption or ethanol re-uptake in the constructed strains

In agreement with the growth pattern, clear differences in metabolite profile were observed between the constructed strains and the background strains in the experiment of glucose-xylose mixtures. None of the constructed strains consumed any xylose whereas TMB4590 consumed 2.87 g/L xylose after 54 hours, see figure 11.

All the strains consumed all the glucose, TMB CB 014, TMB CB 016 and TMB4590 after 12 hours and TMB CB 012 and TMB CB 017 after 24 hours. Considering the CDW in combination with calculating the yield in biomass from the substrate consumed, it is clear that TMB4590 produced more biomass per C-mol substrate consumed. TMB4590 had a biomass yield of 35.9% compared to 10.0%, 12.0%, 11.8% and 10.5% for TMB CB 012, 014, 016 and 017, respectively. TMB CB 014 and TMB CB 016 had a slightly higher yield than TMB CB 012 and TMB CB 017 and they also consumed the glucose faster.

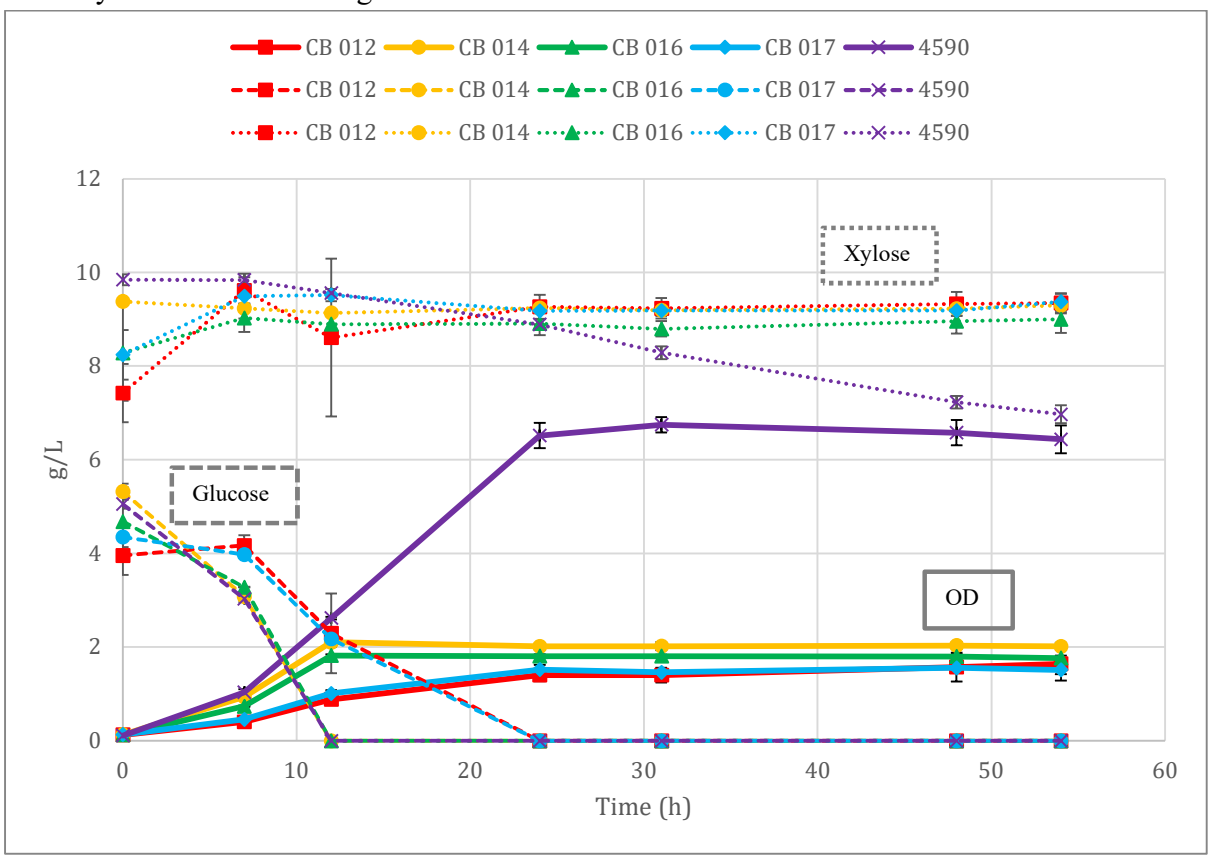


Figure 11. Results from the growth under controlled conditions in shake flasks for TMB CB 012 (red squares), TMB CB 014 (yellow circles), TMB CB 016 (green triangles), TMB CB 0017 (blue rhombs) and TMB4590 (purple crosses). OD is illustrated in full lines, glucose concentration in striped lines and xylose concentration in dotted lines.

For all four constructed strains, the OD stopped increasing at the time glucose ran out, see figure 11. In contrast, TMB4590 OD continued increasing despite all the glucose having been

consumed, which could be because 4590 consumes xylose as well. Looking at the change in ethanol concentration over the cultivation there is another clue as well.

All strains produced ethanol from the very start. At hour 12 when the glucose had run out for TMB4590, the ethanol concentration decreased until hour 24 where it went down to 0 g/L (figure 12). At the very same hour the OD stops increasing. This indicates that the cells produce ethanol from the glucose and when the glucose runs out, they use the produced ethanol as carbon source to build biomass until ethanol runs out as well, a phenomenon called the diauxic shift.

The constructed strains produced ethanol to around the same concentration as TMB4590. However, once glucose ran out, the ethanol concentration decreased slowly, see figure 12. The decrease could originate from a significantly lower consumption of ethanol but it is more likely that the ethanol evaporated. Altogether these results indicate that the growth problem might be due to the strains not consuming ethanol to the same extent as TMB4590.

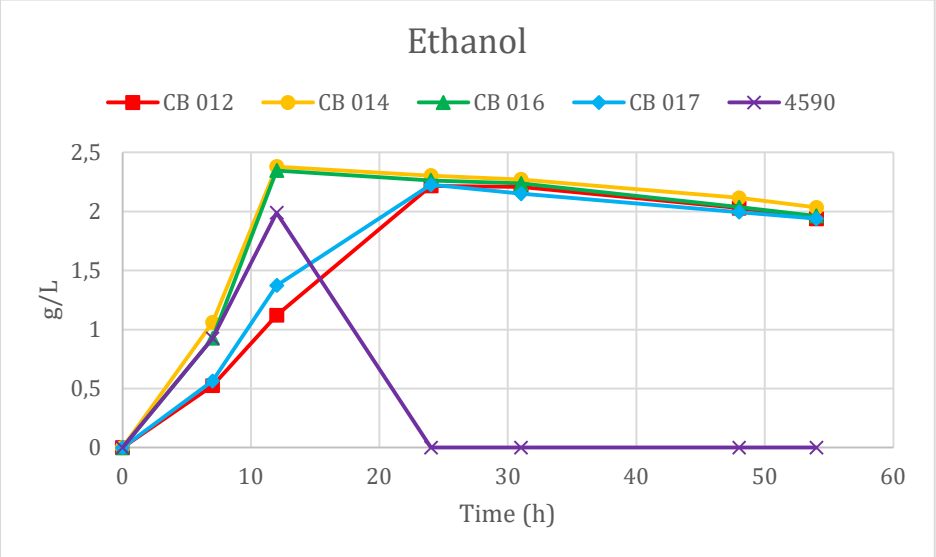


Figure 12. Results from HPLC analysis of samples taken from the growth in shake flasks under controlled conditions experiment, showing the ethanol concentration. TMB CB 012 (red squares), TMB CB 014 (yellow circles), TMB CB 016 (green triangles), TMB CB 017 (blue rhombs) and TMB4590 (purple crosses).

Another interesting discovery was made from the metabolite analysis, all the constructed strains produced more glycerol than TMB4590 (figure 13A). After 54 hours, TMB4590 did not have any glycerol production while the constructed strains reached 0.6 g/L glycerol, see figure 13 A. In addition, 4590 produced a small amount of xylitol while the constructed strains did not, see figure 13 B. In TMB4590 the *GRE3* gene, that encodes an unspecific xylose reductase, was deleted to prevent the production of xylitol. Still, a small amount of xylitol was observed, which can be attributed to the activity of other unspecific reductases. However, xylitol was not observed in the constructed strains that also have the *GRE3* gene deleted. Instead glycerol production is observed.

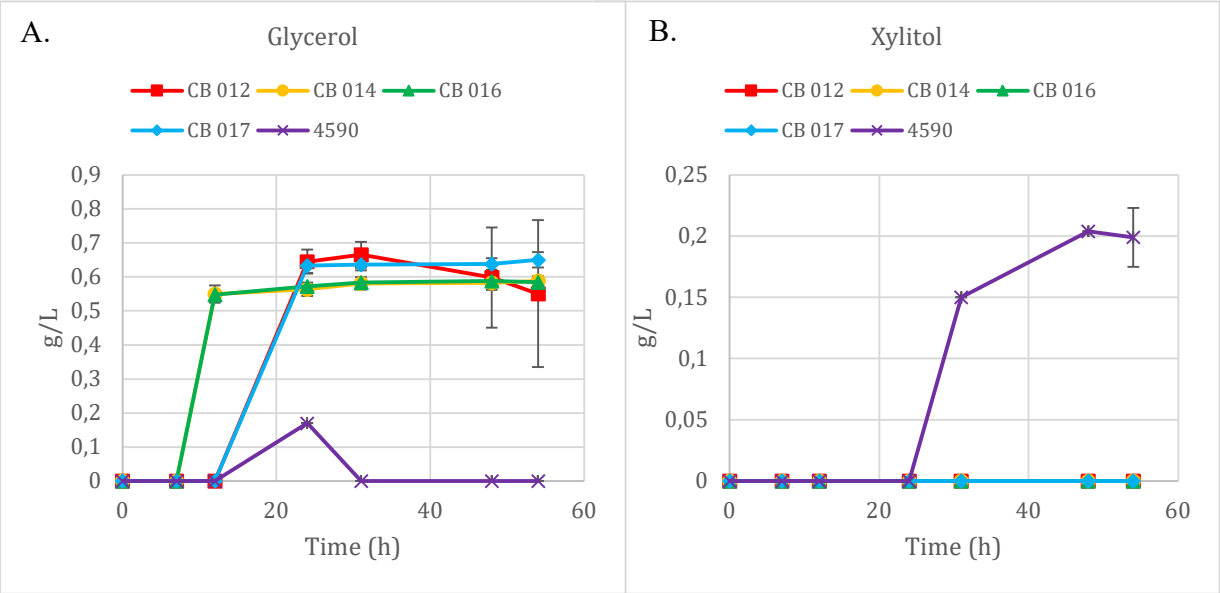


Figure 13. Results from the growth in shake flasks under controlled conditions experiment showing glycerol concentration (A) and xylitol concentration (B). TMB CB 012 (red squares), CB 014 (yellow circles),CB 016 (green triangles), CB 017 (blue rhombs) and TMB4590 (purple crosses).

Overall, the metabolite profile of TMB CB 012 and TMB CB 017 were similar to each other, the same was observed between TMB CB 014 and TMB CB 016. This is probably due to them originating from the same background strain TMB4568 and TMB4512, respectively.

The results all point to that the constructed strains have problems with respiration and that the problems arise already after the first transformation. There might be a problem with the integration sites, but there did not seem to be any genes involved in respiration at the integration sites used. A problem in respiration would impact on the Weimberg pathway as it is dependent on respiration. Looking at the xylonate production, 4590 produced around 3.5 g/L after 54 hours but the constructed strains had at this point no xylonate, see appendix figure A. The lack of xylonate indicates that no xylose is converted in the Weimberg pathway. However, it is difficult to say if that is due to the respiration problem or the gene copy number.

5. Conclusion

In the present master thesis project, a set of strains combining different levels of the Weimberg pathway enzymes were successfully constructed. It was found that TMB CB 016 (*xylB*, *xylD*, 4x(*xylX*, *ksaD*)) had an activity in the coupled assay, measuring the activity of XylD, XylX and KsaD together. However, due to the problems with the assay the results are unreliable and need to be further investigated.

None of the constructed strains grew as good as TMB4590 on glucose, nor as well as the background strains from which they were created. Also, none of the constructed strains were able to consume xylose. The fact that the constructed strains did not assimilate ethanol, produce glycerol nor produce xylitol, point towards a problem in respiration, which would also affect the biomass formation and growth. However, since the Weimberg pathway is dependent on respiration, it is impossible to say if the constructed strains did not consume xylose due to the copy numbers of the genes or because of the respiratory problem.

Regarding the enzyme activities needed for the pathway to work (RQ1) it is still too early to give the optimal number from the constructed strains since none of them consumed xylose. When it comes to RQ2, for now, since only TMB4590 has been able to utilize xylose, all four enzymes in the Weimberg pathway, with one copy of *xylB* and four copies of *xylD*, *xylX* and *ksaD*, are necessary for it to work. Further research must be done to fully answer the research questions, starting with creating strains without impaired growth. For instance, by re-doing the transformations and/or measuring the OD of more clones after transformation to try to find one with better growth. Another alternative is to use different integration sites for the genes, even though it appeared that none of the integration sites contained any genes involved in respiration. One could also try to use the same technique as when creating TMB4590, to integrate several genes at the same site.

References

Almqvist, H. et al., 2018. Characterization of the Weimberg pathway in Caulobacter crescentus. *Fermentation,* 4(44).

Almqvist, H., Sandahl, M. & G, L., 2017. A rapid method for analysis of fermentatively produced D-xylonate using ultra-high performance liquid chromatography and evaporative light scattering detection. *Bioscience, Biotechnology, and Biochemistry,* 81(6).

Barnett, J., 1976. The utilization of sugars by yeast. *Advances in carbohydrate chemistry and biochemistry,* Volume 32, pp. 125-234.

Barrangou, R. et al., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* Volume 315, pp. 1709-1712.

Borgstrom, C. et al., 2019. Identification of modifications producing growth on xylose in recombinant Saccharomyces cerevisiae strains carrying the Weimberg pathway. *Metabolic Engineering,* Volume 55, pp. 1-11.

Brouns, S. J. J. et al., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science,* Volume 321, pp. 960-963.

Brüsseler, C., Späth, A., Sokolowsky, S. & Marienhagen, J., 2019. Alone at last! - Heterologous expression of a single gene is sufficient for establishing the five-step Weimberg pathway in Corynebacterium glutamicum. *Metabolic Engineering Communications,* Volume 9.

Buchert, J. & Viikari, L., 1988. The role of xylonolacetone in xylonic acid production by Pseudomonas fragi. *Applied Microbial Biotechnology,* Volume 27, pp. 333-336.

DiCarlo, J. E. et al., 2013. Genome engineering in Saccharomyces cerevisiar using CRISPR-Cas systems. *Nucleic Acid Research,* 41(7), pp. 4336-4343.

Entian, K.-D. & Kötter, P., 2007. Yeast genetic strain and plasmid collections. *Methods in Microbiology,* Volume 36, pp. 629-666.

Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J., 2009. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast,* 26(10).

Gietz, R. D. & Schiestl, R. H., 2007. High-efficiency yeast transformation using the liAc/SS carrier DNA/PEG method. *Nature Protocols,* Volume 2, pp. 31-34.

Hongzhang, C., 2014. *Biotechnology of Lignocellulose.* Netherlands: Springer.

Im, H., Sambrook, J. & Russell, D. W., 2011. The inoue method for preparation and transformation of competent E. coli: "Ultra component" cells. *Bio-protocol,* 1(20).

Jessop-Fabre, M. et al., 2016. EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae vis CRISPR-Cas9. *Biotechnology Journal,* 11(8), pp. 1110-1117.

Kim, S., Park, Y.-C., Jin, Y.-S. & Seo, J.-H., 2013. Strain engineering of Saccharomyces cerevisiae for enhanced xylose metabolism. *Biotechnology Advances,* 31(6), pp. 851-861.

Nygård, Y. et al., 2014. Single cell and in vivo analyses elucidate the effect of xylC lactonase during production of D-xylonate in Saccharomyces cerevisiae. *Metabolic Engineering,* Volume 25, pp. 238-247.

Radek, A. et al., 2014. Engineering of Corynebacterium glutamicum for minimized carbon loss during utilization of D-xylose containing substrates. *Journal of Biotechnology,* Volume 192, pp. 156-160.
Salusjärvi, L. et al., 2017. Production of ethylene glycol or glycolic acid from D-xylose in Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology,* Volume 101, pp. 8151-8163.
Stephens, C. et al., 2007. Genetic analysis of a novel pathway for D-xylose metabolism in Caulobacter crescentus. *Journal of Bacteriology,* Volume 189, pp. 2181-2185.
Stovicek, V., Borodina, I. & Forster, J., 2015. CRISPR-Cas system enables fast ans simple genome editing of industrial Saccharamyces cerevisiae strains. *Metabolic Engineering Communications,* Volume 2, pp. 13-22.
Taxis, C. & Knop, M., 2006. System of centromeric, episomal, and integrative vectors based on drug resistance markers for Saccharomyces cerevisiae. *BioTechniques,* 40(1).
Toivari, M. et al., 2012. Metabolic engineering of Saccharomyces cerevisiae for bioconversion of D-xylose to D-xylonate. *Metabolic Engineering,* Volume 14, pp. 427-436.
Wasserstrom, L. et al., 2018. Exploring D-xylose oxidation in Saccharomyces cerecisiae through the Weimberg pathway. *AMB Express,* 8(33).
Weimberg, R., 1961. Pentose oxidation by Pseudomonas fragi. *Biological Chemistry,* Volume 236, pp. 629-635.

Appendix

Table A. Primers used in the current degree project. Annealing sequences are underlined and an eventual restriction enzyme sequence is highlighted in italics.

Primer name	Forward or reverse	Primer sequence, 5' to 3'	T_M (T_M with restriction enzyme seq)	Restriction enzyme
5_outside_F1	F	TCCGACTGTGTATTGGAATAAGTTTTTCGGTGTTATA TATATACATATAT TGTAGTCGTGCAATGTATGAC	55.9 (73.4)	-
5B_outside_R1	R	TTCAATGCCGTCTCCTTTCGAAACTTAATAATAAAAC AATATCATCCTTT GTATTGATAATGATAAACTCGGTG	55.9 (73.3)	-
6_XylD_F	F	TAGAT *GGCGCGCC* TGTAGTCGTGCAATGTATG	52.4 (70.8)	*Asc*I
6_XylD_SacI_F	F	CTCGCT *GAGCTC* TGTAGTCGTGCAATGTATG	52.4 (69.5)	*Sac*I
6_XylD_R	R	TAAG *CCTGCAGG* AATGATAAACTCGGTGCTGG	55.2 (69.5)	*Sda*I
FRA2_gRNA_1_f	F	TAGCTGCAAGTGCTACACTG GTTTTAGAGCTAGAAATAGCAAG	55.3 (71.3)	-
FRA2_gRNA_2_f	F	TCAGTCGTTTGATATTGTGG GTTTTAGAGCTAGAAATAGCAAG	55.3 (69.4)	-
FRA2_gRNA_3_f	F	ACAATGATATGGTAAACTTG GTTTTAGAGCTAGAAATAGCAAG	55.3 (67.5)	-
FRA2_gRNA_4_f	F	TGGCTGATTCTATCCTCTCG GTTTTAGAGCTAGAAATAGCAAG	55.3 (71.3)	-
LWA_103	R	GATCATTTATCTTTCACTGC	51.1	-
cg0535_int_f	F	CGCTGCTTCCTTGGTTG	55.2	-
XylX_int_f	F	AGTCAGGGTATCCACTCC	56.0	-
FBAIt_SdaI_r	R	TAT*C CTGCAGG* CTCCCACAAAGAACTC	49.2 (66.5)	*Sda*I
3035_int_DS	R	GCTCATCACGGAACTGTAAC	57.3	-
2904_int_DS	R	AGAATTGGCGACACATTGAG	55.2	-
2909_int_DS	R	GTAGATGTAATCAATGAAGCGG	56.5	-
3039_int_DS	R	CGGTTGACCATAGTATTCACC	57.9	-
2899_int_DS	R	ACGCCACCAAGACAATATCC	57.3	-
3037_int_DS	R	CACAGTGTACCAACGATTCG	57.3	-
3040_int_DS	R	GTGAAATCTCTTTGCGGTAG	55.2	-
333	R	GCTGTTGATGATAGTACTGC	55.2	-
334	F	AGGTGAAAGAATTGAAAAGG	51.1	-

Table B. Amplification PCR program. A1 is the lowest T_M of the primers used without the restriction enzyme sequence plus 3 ℃, A2 is the lowest T_M of the primers used with the restriction enzyme sequence plus 3 ℃ and E is 30 s per 1000 bp of the fragment produced. The annealing temperatures, A1 and A2, and the time used in the second extension, E, were determined based on the primers used in the PCR reaction.

Step	℃	Time	Cycles
Initial denaturation	98	30 s	1
Denaturation	98	10 s	5
Annealing	A1	15 s	5
Extension	72	90 s	5
Initial denaturation	98	30 s	1
Denaturation	98	10 s	25
Annealing	A2	15 s	25
Extension	72	E s	25
Final extension	72	10 min	1
Hold	4		

Table C. Shows the performed amplification PCR reactions, including which primers used, approximate length of PCR product, annealing temperature and extension time used.

PCR product	Primers	Length of PCR product	A1 (°C)	A2 (°C)	E1 (sec)
XylD 6	6_XylD_F 6_XylD_R	~3000 bp	55.4	72	100
XylD 6 SdaI/SacI	6_XylD_SacI_F 6_XylD_R	~3000 bp	52.4	72	105
gRNA FRA2 plasmids	FRA2_gRNA_1_f FRA2_gRNA_2_f FRA2_gRNA_3_f FRA2_gRNA_4_f LWA_103	~5300 bp	54.1	-	160
XylD with FRA2 US and FRA2 DS	5_outside_F1 5B_outside_R1	~3100 bp	58.9	72	90

Table D. Colony PCR program. A3 is the lowest T_M of the primers used minus 5 °C and E2 is 1 min per 1000 bp of the fragment produced. A3 and E2 were determined based on the primers used in the PCR reaction.

Step	°C	Time	Cycles
Initial denaturation	94	5 min	1
Denaturation	95	10 s	30
Annealing	A3	30 s	30
Extension	72	E2	30
Final extension	72	10 min	1
Hold	4		

Table E. Shows the performed colony PCR reactions, including which primers were used, approximate length of PCR product, annealing temperature and extension time used.

Strain controlled	Primers	Length of PCR product(s)	A3 (°C)	E2 (sec)
TMB CB 010	cg0535_int_f 3035_int_DS FBAIt_SdaI_r	~1500 bp	49.1	45
TMB CB 011 TMB CB 014	cg0535_int_f 3035_int_DS 2904_int_DS 2909_int_DS FBAIt_SdaI_r	~1250 bp, ~1450 bp & ~1500 bp	50.2	90
Constructed plasmids 2899-XylD & 3037-XylD	6_XylD_F 6_XylD_R	~3000 bp	64.5	180
gRNA3+XylD (TMB CB 017) gRNA4+XylD (TMB CB 017)	334 333 6_XylD_F	~3400 bp & ~750 bp (neg. control)	48.1	204
gRNA2+XylD (TMB CB 017) TMB CB 012	334 333 6_XylD_F XylX_int_f 3039_int_DS FBAIt_SdaI_r	~3400 bp, ~750 bp (neg. control), ~1500 bp.	46.1	204
TMB CB 013 TMB CB 015 TMB CB 016	XylX_int_f 6_XylD_F 2899_int_DS	~1150 bp, ~1200 bp, ~1250 bp,	47.4	216

TMB CB 018	3037_int_DS 3040_int_DS FBAIt_SdaI_r	~3500 bp, ~3600 bp.		

Table F. The plasmids used to construct each strain. The plasmids containing the donor DNA and the plasmids with the gRNA.

TMB number	Donor DNA	Plasmids with gRNA
CB 010	pCfB3036-*ksaD*	pCfB3043
CB 011	pCfB3035-*ksaD* + pCfB2904-*ksaD* + pCfB2909-*ksaD*	pCfB3052
CB 012	pCfB3039-*xylX*	pCfB3048
CB 013	pCfB2899-*xylX* + pCfB3037-*xylX* + pCfB3040-*xylX*	pCfB3053
CB 014	pCfB3035-*ksaD* + pCfB2904-*ksaD* + pCfB2909-*ksaD*	pCfB3052
CB 015	pCfB2899-*xylD* + pCfB3037-*xylD* + pCfB3040-*xylD*	pCfB3053
CB 016	pCfB2899-*xylX* + pCfB3037-*xylX* + pCfB3040-*xylX*	pCfB3053
CB 017	*xylD*	gRNA2
CB 018	pCfB2899-*xylD* + pCfB3037-*xylD* + pCfB3040-*xylD*	pCfB3053

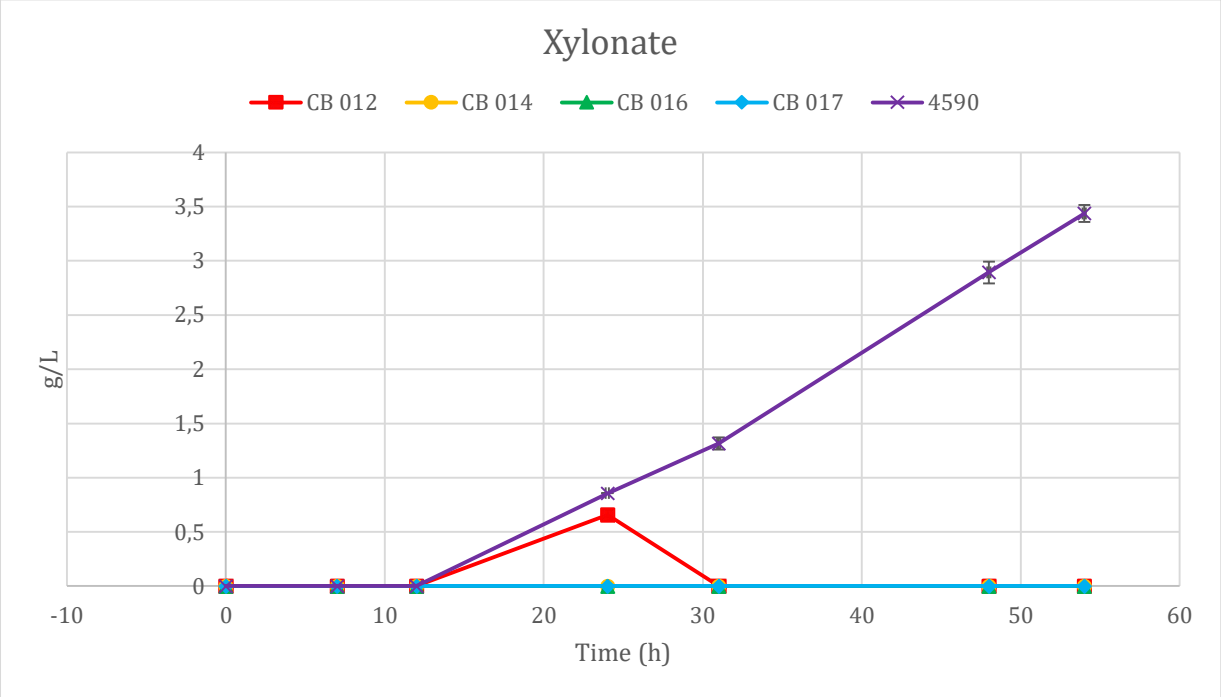


Figure A. Results from the growth in shake flasks under controlled conditions experiment, showing xylonate concentration. TMB CB 012 (red squares), CB 014 (yellow circles), CB 016 (green triangles), CB 017 (blue rhombs) and TMB4590 (purple crosses).