UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Dainis Ivanovs

L-Phenylalanine Biosensor in Saccharomyces cerevisiae

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor(s):
Senior scientist, PhD Petri-Jaan Lahtvee
Research fellow, PhD Nemailla Bonturi

Tartu 2019

L-Phenylalanine Biosensor in Saccharomyces cerevisiae

Abstract:

In vivo biosensors can be used for the detection of intracellular and extracellular stimuli leading to a reporter molecule response. Biosensors are becoming increasingly popular in synthetic biology and metabolic engineering for implementation in synthetic pathways to control the expression of reporter genes in response to specific metabolites or external tors. In vivo biosensors are generally divided into three categories: transcription-based, translation-based and post-translational biosensors. L-phenylalanine is one of the most demanded amino acids and it faces an increasing commercial interest since it is a precursor metabolite for flavonoids and other chemicals including the artificial sweetener aspartame, vanillin, aromadedrin, rutin and many others. Therefore, it is of interest in various tries such as pharmaceutical, cosmetic and biotechnology. L-phenylalanine is one of the aromatic amino acids yielded from the Shikimate pathway, also including tyrosine and tryptophan. In this paper the author constructed a transcription factor-based L-phenylalanine biosensor in the yeast Saccharomyces cerevisiae. The newly constructed strain has two genetic constructs integrated into its genome. First construct uses tional regulatory protein TyrR which forms a dimer only in the presence of L-phenylalanine. The dimer later binds to a TYR strong box sequence in the CYC1p_T1 promoter region of the second construct initiating green fluorescent protein production resulting in a fluorescent signal.

Keywords:

L-phenylalanine, biosensor, Saccharomyces cerevisiae, transcription factor, TyrR ry protein, Shikimate pathway, flavonoids.

CERCS: T490

L-fenüülalaniini biosensor pärmis Saccharomyces cerevisiae

Lühikokkuvõte:

In vivo biosensoreid võimaldavad detekteerida kindlaid rakusiseseid või -väliseid stiimuleid ning reageerida neile mõne soovitud molekulaarse vastusega. Biosensorite

kasutamine on tänu sünteetilise bioloogia ning metaboolse inseneerimise metoodikate arengule pidevalt kasvanud, sest võimaldab kontrollida ning reguleerida soovitud valkude või metaboliitide tasemeid rakus. *In vivo* biosensorid jagatakse tavaliselt kolmeks: (i) transkriptsiooni-põhised, (ii) translatsiooni-põhised ning (iii) translatsiooni-järgsed biosensorid. L-fenüülalaniin on aminohape ning prekursor-molekul näiteks flavanoididele – keemiliste elementide grupile, kuhu kuuluvad ka kunstlik magusaine aspartaam, vaniliin, aromadedriin, rutiin ning paljud teised kemikaalid. Seetõttu on sel molekulile suur kommertsiaalne huvi nii farmaatsias, kosmeetika-tööstuses kui biotehnoloogias. L-fenüülalaniini biosünteesitakse Sikimaadi rajas, mis toodab ka teisi aromaatseid aminohappeid nagu türosiin ja trüptofaan. Käesolevas lõputöös kirjeldatakse transkriptsioonifaktori meetodil põhinevat L-fenüülalaniini biosensori konstrueerimist pärmis *Saccharomyces cerevisiae*. Uudset biosensorit sisaldava pärmi tüve genoomi on lisatud kaks uut kostrukti. Esimene konstrukt kasutab transkriptsiooni regulaator-valku TyrR, mis moodustab dimeerse kompleksi vaid L-fenüülalaniini juuresolekul. Dimeerne TyrR kompleks on aga võimaline seonduma TYR ’strong box’ järjestusega CYCp_T1 promootor-regioonis teise kontruktiga, mis seejärel initsieerib rohelise floresentsvalgu sünteesi, mis ongi disainitud biosensori väljund-signaal.

Võtmesõnad:

L-fenüülalaniini, biosensor, Saccharomyces cerevisiae, transkriptsioon faktor, TyrR reguleeriv valku, Shikimaalne rada, flavonoidid.

CERCS: T490

TABLE OF CONTENTS

TERMS, ABBREVIATIONS AND NOTATIONS

L-Phe – L-phenylalanine aromatic amino acid

L-Tyr – L-tyrosine aromatic amino acid

L-Trp – L-tryptophan aromatic amino acid

LB – lysogeny broth nutritionally rich medium

YPD – yeast extract peptone dextrose medium

DELFT – minimalistic mineral medium

PEP – phosphoenolpyruvate

E4P – erythrose 4-phosphate

DAHP – 3-Deoxy-D-arabino-heptulosonic acid 7-phosphate

UV – ultraviolet electromagnetic radiation

PAL –phenylalanine-ammonia-lyase

C4H –cinnamate 4-hydroxylase

4CL – 4-coumarate: CoA ligase

AD – Alzheimer's disease

NDM1 – New Delhi metallo-β-lactamase-1

BACE1 – β active site-cleavage enzyme-1

TF – transcription factor

TR – transcriptional regulator

RNA – ribonucleic acid

DNA – deoxyribonucleic acid

GFP – green fluorescent protein

Xgal – 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside

lacZ – b-galactosidase

CCM – *cis*,*cis*-muconic acid

HPLC – high pressure liquid chromatography

αCTD – C-terminal domain of alpha subunit of the RNA polymerase

LTTR – LysR-type transcriptional regulator

ATP – adenosine triphosphate

T1 – the first TATA box region of the CYC1p promoter DNA sequence

CYC1p – cytochrome c isoform 1 promoter

CYC1t – cytochrome c isoform 1 terminator

pFCY2 – purine-cytosine permease promoter

tADH1 – alcohol dehydrogenase promoter

TyrR – Tyr tyrosinase regulon

AMP – ampicillin

LEU – amino acid leucine

HIS – histidine

V – volts

mA – miliampers

G – gravitational force of earth

rpm – rotations per minute

min – minutes

sec – seconds

h – hour

bp – DNA base pairs

kB – DNA kilobases (1 kb = 1000 bp)

INTRODUCTION

The aromatic amino acid L-phenylalanine is one of the three amino acids including L-tyrosine and L-tryptophan which is yielded from the shikimate metabolic pathway and is later derived into various flavonoids like naringenin, aspartame and vanillin or other aromatic derivatives by use of the phenylpropanoid metabolic pathway. In order to produce these molecules industrially a number of microorganisms are being used including *Saccharomyces cerevisiae* yeast.

The problem is that there have been no previous biosensors developed for sensing the concentrations of L-phenylalanine in yeast which is a key molecule in production of phenypropanoid pathway products. This means that selection of best phenotypes for the overproduction of L-phenylalanine and thus its downstream dervatives uses expensive, labor intense and time consuming methods such as high pressure liquid chromatography analysis. Luckily there have been previous attempts in constructing biosensors which employ bacterial transcriptional activators in yeast and a functional design template was made. It is also known that L-phenylalanine biosensors in *E. coli* have been made where a functional transcription activating sequence and a transcriptional activator where identified. The transcription factor identified is the TyrR regulon which regulates the aromatic amino acid biosynthesis via the Shikimate pathway. The transcription activating sequence was a so called TyrR strong box which can activate gene transcription in the presence of L-phenylalanine.

The goal of this research was to identify a functional L-phenylalanine biosensor design for *S. cerevisiae* and construct it. It was determined that the biosensor should consist of two genetic constructs. One containing a TyrR regulon capable of dimerising and binding specific transcriptional activation binding sites in the presence of L-phenylalanine under the control of different promoters. The second construct contains a green fluorescent protein coding sequence controlled by a CYC1p promoter with a TyrR strong box integrated in its first TATA box position.

All goals where not achieved but multiple strides were made towards the biosensor assembly in *S.cerevisiae*. The author hopes the current biosensor design can be further tested and implemented for use in the laboratories.

1 LITERATURE REVIEW

1.1 Biosynthesis of L-phenylalanine

The shikimate metabolic pathway (Figure 1.) is only found in microorganisms and plants, but never in mammals which makes this pathway highly interesting for production of herbicides, antibiotics and live vaccines. It links the carbohydrate metabolism with the biosynthesis of aromatic amino acids (L-phenylalanine, L-tryptophan and L-tyrosine) and their secondary metabolites. The common precursor for all aromatic amino acids is chorismate which is produced from central carbon metabolism derivatives phosphoenolpyruvate (PEP) which comes from the glycolysis pathway and erythrose 4-phosphate (E4P) yielded from the pentose phosphate pathway. Chorismate is produced via seven metabolic steps in the shikimate pathway from the precursors. In microorganisms this pathway is regulated by the repression of its initial enzymes and by feedback inhibition. The carbon flow into the shikimate pathway for bacteria is controlled by feedback inhibition of aromatic amino acid-sensitive DAHP synthase isoenzymes which is in charge of the initial PEP and E4P concentration into a heterocyclic compound named 3-Deoxy-D-arabino-heptulosonate7-phosphate also known as DAHP which gets eventually metabolised into chorismate. Since L-phenylalanine and L-tyrosine have similar molecular structures, L-tyrosine has an extra hydroxide group on its hexane ring whereas L-phenylalanine does not, they are both a part of the same branch off point in chorismite metabolism. Chorismate is metabolised into prophenate by chorismite mutase, which is metabolised into either 4-hydroxyphenylpyruvate by prephenate dehydrogenase or into phenylpyruvate by prephenate dehydratase. The aromatic amino acid L-tyrosine is produced from 4-hydroxyphenypyruvate which is catalysed by tyrosine transaminase whereas L-phenylalanine is produced from phenylpyruvate with the help of phenylalanine transaminase (Bentley, 1990), (Herrmann and Weaver, 2002), (Herrmann, 2007), (Gosset, 2009).

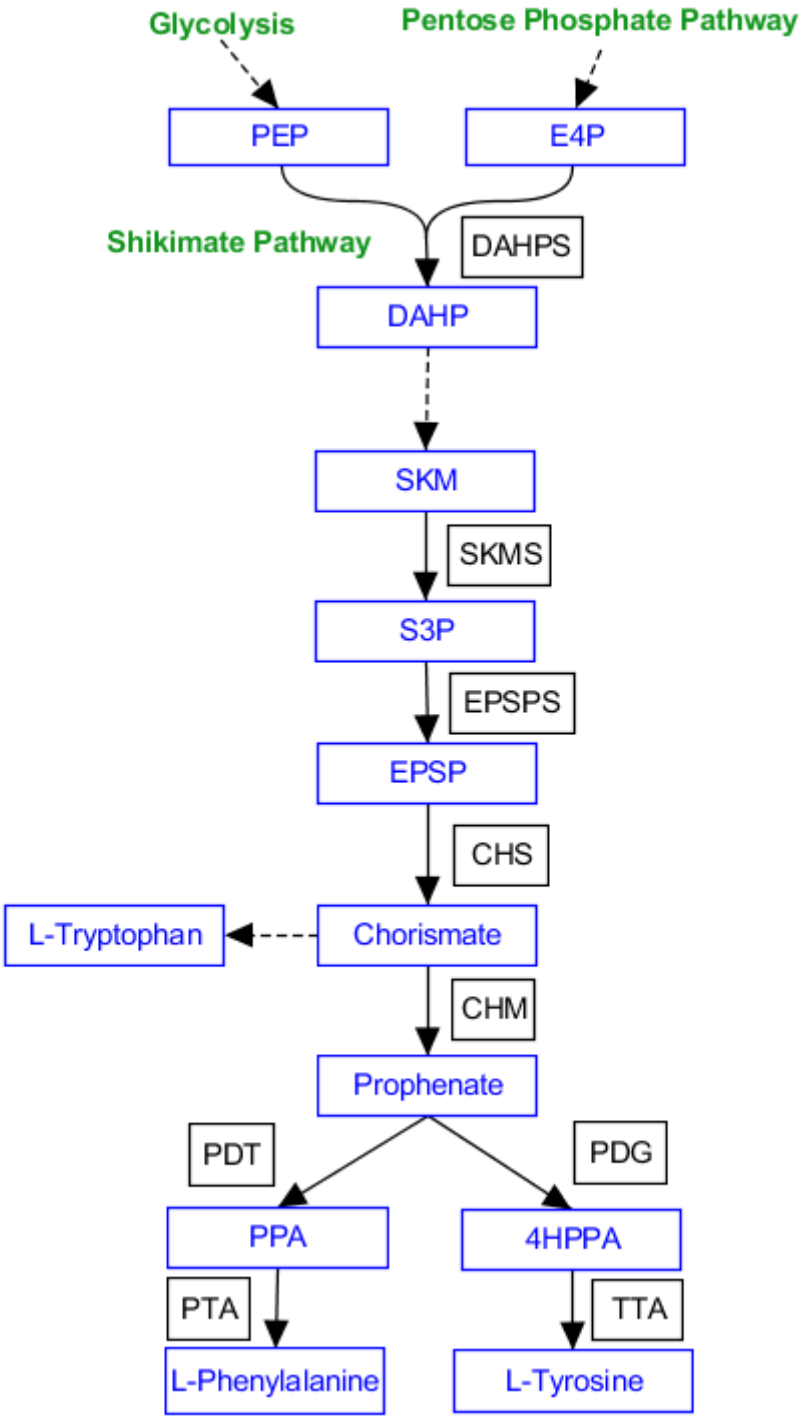


Figure 1. Biosynthesis of L-phenylalanine via Shikimate metabolic pathway in S. cerevisiae.

1.2 L-phenylalanine derived high-value products

The aromatic amino acid L-phenylalanine is a highly demanded amino acid as its derivatives account for 1 billion US dollars in sales (Sprenger, 2007). The derivatives of L-phenylalanine are yielded via the phenylpropanoid pathway (Figure 2.) which supplies

plants with a large variety of phenolic compounds involved in a wide range of functions ranging from plant protection from ultraviolet (UV) radiation, antimicrobial effects, antimicrobial activity and signal transduction. The core phenylpropanoid pathway (PHENYLPROPANOID PATHWAY) degrades phenylalanine into 4-coumaroyl-CoA which is a central precursor for most flavonoids produced from this pathway (Hahlbrock and Scheel, 1989). The degradation happens in three enzymatic reactions. Initially phenylalanine-ammonia-lyase (PAL) catalyses the nonoxidative deamination of L-phenylalanine into cinnamic acid (MacDonald and D'Cunha, 2007). Next cinnamic acid is converted into 4-coumaric acid by cinnamate 4-hydroxylase (C4H) by attaching a hydroxide group to cinnamic acids fourth carbon (4C) of its phenyl group. The last reaction of the core phenylpropanoid pathway forms 4-coumaroyl-CoA from 4-coumaric acid with the help of 4-coumarate: CoA ligase (4CL) which forms a thiol ester in the place of the carboxyl group (Hahlbrock and Scheel, 1989).

The phenylpropanoid Pathway has multiple branching off points. One of which is the branch off point at which uses into 4-coumaric acid and 4-coumaroyl-CoA for lignin production. The enzyme chalcone synthase uses 4-coumaroyl-CoA and malonyl-CoA from the fatty acid biosynthesis pathway and uses the three acetate units of malonyl-CoA to condense the into a new aromatic ring and thus forming naringenin chalcone. Naringenin chalcone is the basis for synthesizing all major flavonoid subgroups flavones, flavanoles, tannins, anthocyanins and isoflavones. (Panche, Diwan and Chandra, 2016), (Weisshaar and Jenkinst, 1998).

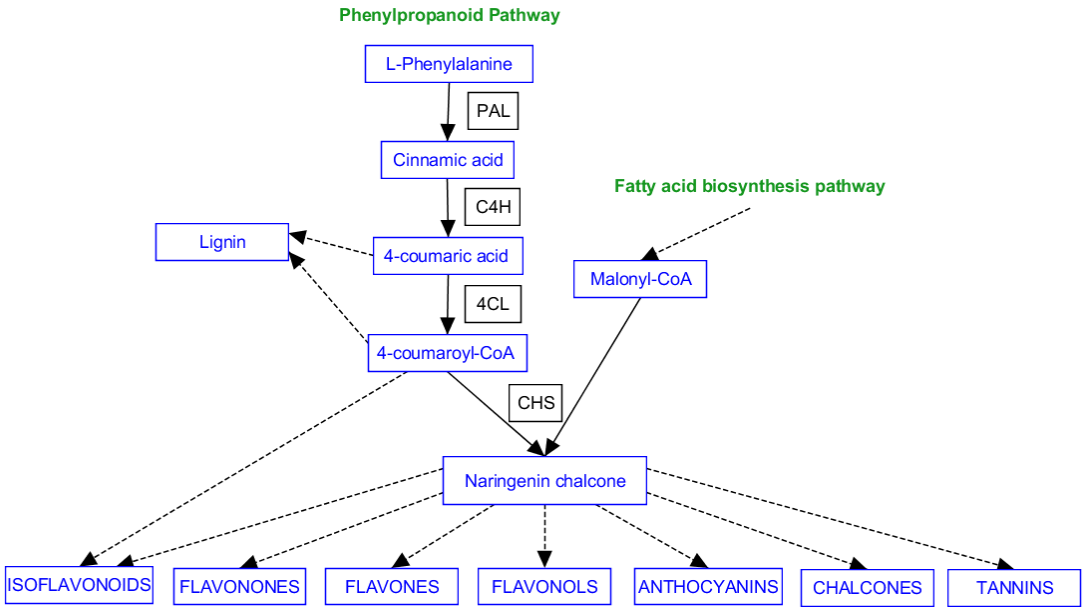


Figure 2. Phenylpropanoid metabolic pathway.

Some particularly interesting compounds for humans include rutin, aromadedrin, ginkgetin, naringenin, quercetin, myricetin and aspartame. Rutin has been shown to have potential in the treatment of diabetes because of its ability to inhibit aldose reductase which is known for causing diabetic complications. Aromadedrin has been found to inhibit lipoxygenase which is involved in causing inflammation. So aromadedrin is being researched as a potent drug for treating inflammation. Ginkgetin has been found to have antiplatelet effects in humans. Naringenin has been found to be a potent antimicrobial inhibitor and is being researched as a potential antibiotic. Quercetin is being researched as a potential antibiotic against superbugs which are New Delhi metallo-β-lactamase-1 (NDM-1) positive and currently have no known antibiotics against them. Myricetin inhibits β active site-cleavage enzyme-1 (BACE-1) activity which is known to be involved in causing Alzheimer's disease (AD) so it is being researched for its potential in AD treatment. (Panche, Diwan and Chandra, 2016) Aspartame is a natural sweetener with roughly the same amount of calories per gram as sucrose, but it is almost 200 times sweeter then sucrose thus the amount required to reach the same level of sweetness is much less which makes it useful as an artificial sweetener with low calories (Suzuki, 2013).

This shows that construction of biosensors for sensing key molecules in these pathways would be beneficial for improving target compound production in microbial strains of interest. The amino acid L-phenylalanine seems to be the best option since it is the intermediate of both metabolic pathways described above and could serve as an optimal target molecule for detection.

1.3 Biosensors

The term "biosensor" describes a wide range of analytical devices containing a biological sensing element with a wide range of applications including drug discovery, environmental monitoring, biomedicine, industrial biotechnology and the list goes on. The first biosensor was created around five decades ago and was designed to measure glucose concentrations in a sample with the help of an immobilized glucose oxidase electrode (Clark and Lyons, 1962). *In vivo* biosensors can generally be divided into three main categories depending on at which stage of the central dogma of molecular biology they operate on whether its transcriptional, translational or the post-translational stage of protein synthesis from DNA. Usual reporters used in *in vivo* biosensors to produce an observable output signal are fluorescence, bioluminescence and colour change (Goers *et al.*, 2013).

Transcription factors function by regulating cellular gene expression depending on specific metabolite concentrations present intra- and extracellularly. These properties have made TF's ideal to be exploited for *in vivo* biosensor construction since they allow for real-time monitoring of production of specific target molecules (Shi, Ang and Zhao, 2018). Such biosensors allow measurement of small deviations in ligand concentrations which lead to large changes in protein abundance thanks to amplification through transcription and thus allows for a wide dynamic range and high sensitivity of the biosensor (Goers *et al.*, 2013).

Transcription based biosensor usually employ promoters which are activated by transcription factors or other signalling cascades. Their response time in majority of the cases is measured in hours and the metabolic burden they pose to the cell is comparatively small since the reporter molecules are only expressed in the presence of an inducer (transcription factor), overexpression of the relevant transcription factor is often necessary when using non-native TF's whereas when using endogenous TF's overexpression might be unnecessary. They are quite specific, but such biosensors have some minor issues like the final output does not exactly correlate with the target molecule concentration because of downstream steps in protein synthesis thus the output signal can be linked to gene expression of the target gene. The construction of such biosensors is relatively easy since they can be constructed from pre-existing genetic components and plasmids. In order to properly characterize the constructed biosensor a negative control cell line is highly advisable in order to define the proper functional range and reporter protein characteristics. There are two options for negative control cell lines. First is a wild-type cell line without the construct required for biosensor functionality integrated, but this option provides only absolute measurements of cell growth and expression characteristics at different test conditions. The second option is using a promoter reference standard, which is essentially the same construct, but without the possibility of TF binding to the promoter activating reporter gene transcription which allows for ratiometric relative measurements thus allowing for better biosensor characterization (Goers *et al.*, 2013)

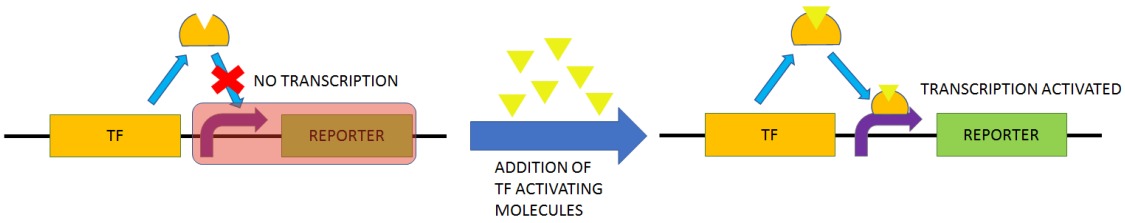


Figure 3. Transcription factor – based biosensor mechanism of action.

Translation based biosensors which are essentially RNA switches, are designed to give a measurable output signal in the presence of the target of interest and generally have faster response times and smaller metabolic burden to the cells in comparison with transcription-based biosensor designs since RNA folding is universal between different organisms. In general, such biosensors have three functional domains where one receives the signal, second processes the signal and third produces a measurable output signal. The signal receiver is an aptamer domain which recognises and specifically binds a target molecule thanks to RNA forming specific secondary and tertiary structures in the presence of the target. The second step (signal processing) happens when the switch domain of the RNA molecule changes the processing intensity, translation rate or stability of the device in response to target binding to the aptamer domain. The third domain opens up and allows for translation of a target reporter protein in the presence of the bound target molecule thus producing a measurable output signal (Goers *et al.*, 2013).

Post-translational biosensors function on the protein basis which are constitutively produced in the cells but are only activated when a target molecule is present thus producing a measurable output signal. Such biosensor designs are very varied in comparison to transcription and translation-based biosensor designs because of the wide protein diversity and function range available in nature, thus an exact design strategy cannot be given since it largely depends on the specific cases of each biosensor and the molecular mechanisms employed, but a good examples are enzymes catalysing target reactions employing one or more secondary substrates to produce an output signal in the presence of a target molecule (Goers *et al.*, 2013).

Fluorescence based reporters are fluorescent proteins such as GFP and many others. They produce a fluorescent output of a predetermined wavelength after being excited by light of a different wavelength thus allowing excited electrons to jump from a higher energy level

to a lower one by emitting light. These proteins offer an output signal with high sensitivity and specificity. By using various fluorescent proteins with different colours it is possible monitor various parameters inside the cell simultaneously (Goers *et al.*, 2013). In vivo fluorescent proteins have been in the forefront of understanding various biological processes such as observing various metabolic pathway functions inside the cells (Kunzelmann, Solscheid and Webb, 2014), imaging of fixed cells using fluorescent-tagged antibodies (Oldach and Zhang, 2014), mitochondrial physiology (De Michele, Carimi and Frommer, 2014) and for visualization of various other celullar processes. Overall fluorescnence has helped visualize most of the experiments involving cells since the first introduction of fluorescent proteins. These methods have helped in developing biosensors which exploit small molecule/metabolite binding transcriptional factors, second messengers and proteins. In such biosensors the concentration of specific metabolites intracellular concentrations which are either directly or indirectly proportional to the GFP (or any other fluorescent protein) fluorescence intensity (Goers *et al.*, 2013).

Bioluminescent proteins, e.g. firefly luciferase, are used similarly as fluorescent proteins, but the mechanism of function differs. Proteins from this class do not emit the light themselves as fluorescent proteins do, but rather they act as enzymes which catalyze chemical reactions leading to a bioluminescent light emission thus the substrate meant for catalysis by these proteins needs to be provided (Goers *et al.*, 2013).

Biosensors that are design to be used for on the field applications the most suitable reporter type is most often colour change which eliminates the need for expensive scientific devices for measuring output signal like in the cases of fluorescence and bioluminescence. Such systems have a wider range in terms of their mechanism of action, but in majority of cases it requires the addition of a substrate which can be chemically converted into a related compund which produces colour. The most often used colour change reporters include the chromogenic compound 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (Xgal) and b-galactosidase (lacZ) gene (Goers *et al.*, 2013).

1.4 Host system

The baker's yeast formally known as *Saccharomyces cerevisiae* is the most widely used microorganism species in the biotechnology industry worldwide and over exceeds the use of other microorganisms in terms of cell weight produced per year by over two folds. Because of the quickly depleting natural non-renewable resources such as petrochemicals and rapidly changing environmental factors thus ever increasing the need for sustainable and

renewable sources yeast has gained renewed interest as production source for various chemicals used in both traditional markets such as food, agriculture, feed and also for the production of specific biocatalysts used in pharmaceutical and fine chemical synthesis industries since the usual chemical synthesis methods traditionally employed in them often require the use of low activity catalysts providing small yields of target compounds and use of chemicals derived from the petrol industry which is a finite source. Whereas naturally produced catalysts can be evolved using various adaptive evolution techniques to cause the microorganism to increase activities of certain enzymes or even evolve new catalytic activities previously not seen in nature if correct constraints have been chosen. The potential of yeast in the biotechnology industry and biochemical synthesis is ever expanding and not yet fully understood (Johnson and Echavarri-Erasun, 2011). Yeast has been used in metabolic engineering from the start since it is a model organism. Studies have shown that yeast is a favourable microorganism for over production of phenylpropanoid pathway derivatives with increased yields in comparison with the native host species of these aromatic molecules such as naringenin (Jiang and Morgan, 2005).

Even though metabolic engineering is becoming more and more valuable it still faces several challenges in order to mature as an area of engineering. This is largely due to the expensive methods employed, like chromatographic methods, to measure cellular compound production during the design process which causes the design-build-test cycles are excruciatingly long and may take up to several years in order to develop a strain with the desired characteristics. In order to counteract this problem, scientists are developing biosensors to get fast and direct readings of specific metabolite concentrations. It has been previously reported that construction of TF based biosensors is a viable option for detection of secondary metabolites from the phenylpropanoid pathway such as vanillin, naringenin, benzoate and various other metabolites (Rogers, Taylor and Church, 2016) it is safe to assume that a similar system can be built for L-phenylalanine detection in yeast since it has not been done before in eukaryotic microorganisms, but only in bacteria such as *E. coli* (Mahr *et al.*, 2016).

There has been a previous research done in the use of prokaryotic small molecule binding transcriptional activators and their use in biosensor design in *S. cerevisiae*. In nature bacterial transcriptional activation can take place when a transcriptional activator binds to a designated operator site in the promoter and by this enhancing the ability to guide RNA polymerase to the promoter and thus activate gene transcription or alternatively by directly in-

teracting with the RNA polymerase itself in order to initiate transcription. In the given research specifically researchers implemented a prokaryotic transcriptional activator from bacteria in order to construct a biosensor for sensing *cis,cis*-muconic acid (CCM) in the budding yeast *S. cerevisiae.* The CCM molecule is an intermediate in the aromatic compound precursor (Skjoedt *et al.*, 2016).

1.5 Biosensor design

A general design for constructing such biosensors was made using a multiparametric engineering strategy. The biosensor design consisted most importantly of a CYC1 promoter which has been reported to be a fitting promoter for introducing other non-native TF binding sites in the yeast genome A BenM sequence, which is a transcriptional regulator activated by CCM and belongs to the prokaryotic family of LysR-type transcriptional regulators (LTTR's). The sequence was introduced in the first so called TATA box (T1) position of the CYC1 promoter yielding a CYC1p_BenM_T1 promoter. The reporter molecule chosen was GFP, so a GFP gene was added after the CYC1p and CYC1p_BenM_T1 promoters yielding a biosensor construct. After performing flow cytometry experiments to measure GFP fluorescence output and concluded that the construct with the CYC1p_BenM_T1 produced a GFP response when CCM was present in the medium compared to the CYC1p promoters which produced only background autofluorescence.

This biosensor design was further tested by integrating various other LTTR family TR's such as FdeR,cPcaQ ArgP and MdcR native to various bacterial species and activated by naringenin, protocathecuic acid, L-arginine and malonic acid, respectively. The DNA sequences for all these TR's where integrated into the T1 location of the CYC1p promoter just as it was done with BenM. The results showed that all TR's introduced in the CYC1p_X_T1, with X representing each of the TR's, activated GFP expression if their respective inducing molecule was present in the medium thus showing that the biosensor design discussed is viable which was further backed up by high pressure liquid chromatography (HPLC) experiments combined with slow cytometry showing a 98% correlation between CCM production titers and GFP fluorescent output (Skjoedt *et al.*, 2016).

The TyrR protein is native to *E. coli* and plays both the role of a transcriptional repressor and transcriptional activator for various genes involved in aromatic amino acid production. The TyrR has at least eight transcriptional units and can interact with all three of the aromatic amino acids yielded from the Shikimate pathway, ATP and the alpha subunit of

RNA polymerase. TyrR protein consists of three functional domains (N-terminal domain, Central domain and a C-terminal domain).

The N-terminal domain plays the main role in gene expression activation, and consists of three (ACT, PAM, DIM) regions. The ACT region is thought to be responsible for the binding of aromatic amino acids that are involved in activation of gene expression. The PAM region is responsible for the binding of αCTD which is important for the initiation of gene transcription. Whereas the DIM is a dimerization motif of the TyrR protein which has a tendency to form dimers when overexpressed (Pittard, Camakaris and Yang, 2005).

The Central domain of the TyrR protein plays no role in the transcriptional activation but is involved in ATP binding and hexamerization. It consists of an ATP binding sites, an ATP-dependent tyrosinase binding site and a hexamerization region. ATP helps to increase TyrR protein affinity by around four times. The ATP-dependent tyrosinase binding site is integral for the tyrosine mediated repression and hexamerization. When TyrR protein is overproduced in the presence of tyrosine and ATP it forms hexamers. The central region also has autokinase, autophosphotase and phosphatase activity although the exact roles these activities play in repression have yet to be found. But it is known that in the presence of either tyrosine or phenylalanine the phosphatase activity is inhibited (Pittard, Camakaris and Yang, 2005).

The C-terminal domain consists of two regions (DIM and HXH). The DIM region is involved in dimerization of the TyrR protein. The HXH region is a classical helif-turn-helix DNA-binding motif and plays the main role in the recognition of the so called “TyrR box” DNA sequences which plays an integral part in gene regulation by the TyrR protein (Pittard, Camakaris and Yang, 2005).

The TyrR boxes are DNA sequences that are bound by the TyrR protein. Altogether there have been 17 TyrR boxes identified thus far for the eight characterized members of the regulon. All the TyrR box sequences are related to the $TGTAAAN_6TTTACA$ palindrome. The TyrR boxes are split into weak and strong, where the strong boxes activate gene transcription whereas the weak boxes repress gene transcription. In operons containing these weak and strong boxes near or inside of the promoters provides a sensing mechanism for aromatic amino acids.

In order to initiate gene activation by TyrR four conditions have to be met which include an imperfect promoter able to be activated, TyrR strong box appropriately located in perspective to the promoter, one of the three aromatic amino acids in combination with the

TyrR protein and the αCTD region of RNA polymerase (Pittard, Camakaris and Yang, 2005).

Although TyrR binds effectively to the strong boxes in the absence of aromatic amino acids, it cannot recruit an RNA polymerase without an aromatic amino acid present. If phenylalanine is present and tyrosine is absent in the media the TyrR protein forms dimers and can bind to the TyrR strong box and recruit the αCTD of RNA polymerase thus activating transcription of a gene. Both tyrP and mtr genes are activated by the TyrR protein in the presence of phenylalanine (Pittard, Camakaris and Yang, 2005), suggesting that their corresponding TyrR boxes could be viable mechanisms for biosensor construction.

In the research performed by Regina Mahr and her team they performed the screening of an *E. coli* promoter library to find the best promoter for the construction of a phenylalanine biosensor. They identified a promoter from the mtr gene encoding for L-tryptophan transporter to be suitable for the biosensor design since the mtr is activated by the TyrR protein if phenylalanine is present in the growth medium thus initiating gene transcription. The biosensor construct was designed in a way where the mtr promoter was located downstream from a TyrR strong box which facilitates the TyrR protein binding in the presence of phenylalanine and thus initiating the transcription of the *venus* fluorescent reporter gene (Mahr *et al.*, 2016).

The authors of this paper stress that there are five parameters that require attention when screening for optimal promoter designs for biosensor construction which include, organism adaption to its growth medium, the mode of gene expression activation, uptake and catabolism of target effector molecules, fluorescent protein maturation and carefully chosen composition of the growth medium in order to avoid high autofluorescence. To enrich promoters and get rid of constitutively activated promoters a screening method where cells with high fluorescent output when induced were chosen and choosing promoters with low fluorescent output when target effector is not present. This strategy yielded enriched promoters in the presence of L-phenylalanine. It was shown that the given mtr promoter can be employed for L-phenylalanine biosensor construction (Mahr *et al.*, 2016).

Whereas the screening process suggested for selection of negatively regulated promoters in the presence of the chosen effector. This means that promoter's with low fluorescence when the effector is present are chosen and in the next round of screening promoters exhibiting high fluorescence when the effector is absent where chosen.

The two promoters suggested for L-phenylalanine biosensor construction where suggested mtr and TyrP, but only mtr was chosen because it has been established that the TyrP protein has 10-fold lower gene expression levels in the presence of L-phenylalanine in comparison with the mtr promoter (Mahr *et al.*, 2016).

The L-phenylalanine biosensor in *S. cerevisiae* is based on TF's and will consist of two genetic constructs TF and reporter protein coding sequence. First construct contains TyrR regulon coding sequence which will function as the transcription factor of the biosensor activated by L-phenylalanine. Second construct consists of GFP protein coding sequence regulated by the CYC1p promoter in a similar fashion as done by Mette L Skjoedt and his team (Skjoedt *et al.*, 2016). The CYC1p promoter contains a mtr strong box in its T1 position yielding CYC1p_T1_TyrR promoter (Skjoedt *et al.*, 2016), (Mahr *et al.*, 2016) which is bound by the TyrR regulon dimer formed in the presence of L-phenylalanine and thus initiating transcription of GFP.

2 THE AIMS OF THE THESIS

The general aim of my BSc thesis was to design and construct L-phenylalanine biosensor for the yeast Saccharomyces cerevisiae. The more specific aims of the work are as follows:

- Establishing Gibson assembly methodology in TUIT SynBio laboratory.
- Design a functional biosensor for L-phenylalanine.
- Construct and assemble the biosensor by:
    - Construction of the TyrR regulon containing genetic construct regulated by varying promoters (pFCY2 and CYC1p) and transcription terminator (tADH1).
    - Integration of the TyrR genetic construct into the *S. cerevisiae* genome.
    - Assembly of the CYC1p_T1_TyrR promoter containing the TyrR strong box transcription activation site.
    - Construction of the reporter protein (GFP) construct regulated by two promoters: CYC1p_T1_TyrR and CYC1p as the reference and transcription terminator (tADH1).
    - Integration of the reporter protein genetic construct into the *S. cerevisiae* genome.
- Testing of the newly constructed biosensor containing *S. cerevisiae* strain in 64-well plate reader fluorescence measurement experiments.
- Characterization of the newly constructed *S. cerevisiae* biosensor strain in Bioreactor cultivation experiments.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Recombinant DNA assembly and plasmid construction

Here the author would like to explain the techniques used in the construction of biosynthetic DNA pathways leading to the creation of the L-phenylalanine biosensor in *S. cerevisiae*.

Standard DNA manipulation was carried out using polymerase chain reaction (PCR), DNA purification from the agar, enzymatic digestion of plasmids and plasmid extraction from bacterial cells. This also includes plasmid digestion for integration into the *S. cerevisiae* genome. DNA sequencing services were provided by EBK DNA sekveneerimise tuumiklabor. All the plasmids (pRS303 and pRS306) were kindly provided by LoogLab (University of Tartu, Tartu, Estonia).

First construct containing the TyrR regulon was integrated into the pRS303 plasmid provided by the SynBio Era laboratory of Tartu University. All the genetic parts like the TyrR gene sequence was synthesized by IDT (Integrated DNA Technologies, Coralville, USA). Promoters pFCY2 and CYC1p and tADH1 terminator sequences were taken from iGem (International Genetically Engineered Machinery competition) distribution kit 2017 (iGem, Cambridge, USA). The pRS303 plasmid was digested with FD (FastDigest) EcoRI (FD0274) and FD Xba1 (FD0684) enzymes (Thermo Fisher Scientific, Waltham, USA) at 37°C for 1 h. The promoters (CYC1p and pFCY2), (primers 1-4, Table of primers, Appendix), gene (TyrR) (primers 5 and 6 or 6 and 7, Table of primers, Appendix) and terminators (tADH1), (primers 10 and 11, Table of primers, Appendix)were PCR amplified to introduce homologous (~20 bp) overhang sequences for each of the respective construct parts. The digested plasmids and the PCR amplified products were purified using gel electrophoresis. The samples where loaded in 1% agar gel and electrophorated for 30 min at 125 V and 400 mA. The extracted DNA fragments encased within the agar gel where gel purified using FavorPrep GEL/PCR Purification Mini Kit (FAGCK001-1), (FavorGen Biotech Corp, Vienna, Austria). The whole construct was assembled into the pRS303 plasmid following the Gibson assembly protocol (E5510) provided by NEB (New England BioLabs incorporated, Ipswich, USA) for assembly of four to six DNA fragments and incubated at 50°C for 1 h. The newly assembled construct was transformed into *E. coli* for plasmid am-

plification by initially thawing the competent bacterial cells on ice. 50 µL of thawed cells where mixed with 1 µL of constructed plasmid DNA and left on ice for 30 min. This was followed by 2 min of heat shock at 42°C and again put on ice for 2 min after which cells where mixed with 400 µL of LB media and incubated at 37°C for 30 min. The process was continued by centrifugation of the grown cells at 6000rpm for 1min and the supernatant was removed. Cells were resuspended in the remaining media and plated on LB+AMP plates overnight at 37°C incubation chamber. The selected bacterial colonies were grown overnight at 37°C in the shaking incubation chamber in 5 ml of LB liquid media with 5 µL of AMP antibiotic added to the medium. Plasmid was rescued after overnight bacterial incubation using FavorPrep Plasmid Extraction Mini Kit (FAPDE300), (FavorGen Biotech Corp, Vienna, Austria) according to the protocol provided by the manufacturer. The verification of the assemblies was tested using digestion enzymes FD Xba1 and FD Pst1 (FD0615), (Thermo Fisher Scientific, Waltham, USA) and incubated at 37°C for 1 h after which the digested samples were gel electrophorated. The previously extracted plasmid from the LB+AMP liquid colonies were digested using FD-Pfl23II (FD0854), (Thermo Fisher Scientific, Waltham, USA) at 37°C for 1h to digest the plasmid in the middle of the HIS resistance marker. The digestion enzyme was heat inactivated at 65°C for 5min as described by the provided protocol which came with the enzyme. The digested plasmid transformed into the yeast genome following the protocol described by J.H. Hegemann and S.B. Heick (Radamson, 2018). After incubation at 30°C for 3-5 days the yeast colonies were harvested. The correct integration of the construct was tested using PCR amplification of the construct of interest and via genetic sequencing. This yielded SBY65 and SBY66 yeast strains with pFCY2 and CYC1p promoters for respective yeast strain integrated into the genome. Positive colonies were re-plated on a new YPD+HIS plates and grown at 30°C for two days. Stocks were made by mixing 250 µL of 60% glycerol solution with 750 µL of YPD medium and the subsequent yeast colonies were tested by PCR amplification of the TyrR construct from the genomic DNA. To extract the genomic DNA from yeast cells, a colony was resuspended in 30 µL of NaOH (Sodiem Hydroxide, 20mM). Put in 100°C for 10 min and then placed on ice. After which they were centrifugated at 18 G (gravitational force of earth) for 1min. The supernatant was used for PCR amplification for confirmation (~2 µL per reaction). The positive colonies were suspended in this medium and stored at -80°C freezer.

In order to introduce the mtr strong TyrR box into the CYC1p promoters T1 position overlapping PCR method was employed. Initially two fragments of CYC1p were ordered from IDT (Integrated DNA Technologies, Coralville, USA) where the promoter was split in two segments. First segment contained the CYC1p fragment from the beginning of the promoter up till the T1 position where the strong box should be further known as CYC1p_T1, (primers 14 and 15, Table of primers, Appendix). The second segment consisted of the CYC1p promoter T1 position with the strong box inside up till the end of the CYC1p promoter DNA sequence further known as T1_TyrR, (primers 16 and 17, Table of primers, Appendix). After PCR amplification of each of these fragments were electrophorated in 1% agarose gel and extracted by using the FavorPrep GEL/PCR Purification Mini Kit (FAGCK001-1), (FavorGen Biotech Corp, Vienna, Austria). Later, overlapping PCR was employed for the assembly of these fragments forming CYC1p_T1_TyrR promoter using the following reaction (Table 1) and PCR program described below (Table 2). The buffer used for the PCR reaction is DreamTaq buffer (10x), (B65), (Thermo Fisher Scientific, Waltham, USA). Correct fragment amplification and assembly was confirmed using gel electrophoresis and DNA sequencing.

Reagents	**1x reaction**
DreamTaq buffer (10x)	25 µL
CYC1p_T1	2 µL
T1_TyrR	1 µL
Destilled water	17 µL
2,5 µL of each primer (primers 14 and 17, Table of primers, Appendix) added into the reaction after the first PCR cycle	

Table 1. Reaction mixture of for the overlapping PCR for CYC1p_T1 and T1_TyrR assembly into CYC1p_T1_TyrR promoter

Program	**Time**	**Temperature**	
Denaturation	3 min	95°C	
Denaturation	30 sec	95°C	35cycles
Annealing	20 sec	52°C	
Extension	1 min	72°C	
Hold	∞	4°C	

Table 2. Overlapping PCR program for assembling CYC1p_T1_TyrR promoter

Second construct containing the CYC1p and CYC1p_T1_TyrR promoter sequences where integrated into the pRS306 plasmid already containing the GFP and CYC1t terminator sequences provided by the SynBio Era laboratory of Tartu University. All promoter (CYC1p_T1_TyrR and CYC1p) sequences were synthesized by IDT (Integrated DNA Technologies, Coralville, USA). The promoters (CYC1p_T1_TyrR and CYC1p), (primers 20 and 21, Table of primers, Appendix) and the plasmid pRS306 (primers 18 and 19, Table of primers, Appendix) where PCR amplified to introduce homologous (~20bp) overhang sequences for each of the respective construct parts. The PCR amplified products where purified from the using gel electrophoresis. The samples were loaded in 1% agar gel and electrophorated for 30 min at 125 V and 400 mA. The extracted DNA fragments encased within the agar gel were gel purified using FavorPrep GEL/PCR Purification Mini Kit (FAGCK001-1), (FavorGen Biotech Corp, Vienna, Austria) for the CYC1p and CYC1p_T1_TyrR sequences and the plasmid using Zymo-Spin 1 DNA extraction columns (C1003-50), (Irvine, USA). The promoter sequences were assembled into the pRS306 plasmid following the Gibson assembly protocol (E5510) provided by NEB (New England BioLabs incorporated, Ipswich, USA) for assembly of two to three DNA fragments and incubated at 50°C for 1 h. The newly assembled construct was transformed into *E. coli* for plasmid amplification by initially thawing the competent bacterial cells on ice. 50 µL of thawed cells where mixed with 1 µL of constructed plasmid DNA and left on ice for 30 min. This was followed by 2min of heat shock at 42°C and again put on ice for 2 min after which cells where mixed with 400µL of LB media and incubated at 37°C for 30 min. The process was continued by centrifugation of the grown cells at 6000 rpm for 1 min and the supernatant was removed. Cells where resuspended in the remaining media and plated on LB+AMP plates overnight in a 37°C incubation chamber. The selected bacterial colonies were grown overnight at 37°C in the shaking incubation chamber in 5 ml of LB liquid media with 5 µL of AMP antibiotic added to the medium. Plasmid was rescued after overnight bacterial incubation using FavorPrep Plasmid Extraction Mini Kit (FAPDE300), (FavorGen Biotech Corp, Vienna, Austria) and the protocol provided with it. The verification of the assembled was tested using colony PCR by amplifying the whole genetic construct (CYC1p/CYC1p_T1_TyrR + GFP + CYC1t). Other half of the colony was grown overnight at 37°C in the shaking incubation chamber in LB 5 ml of liquid media with 5 µL of AMP antibiotic added to the medium. Plasmid was rescued after overnight bacterial incubation using Zymo-Spin 1 DNA extraction columns (C1003-50), (Irvine, USA) and the protocol provided with it. The extracted plasmid was digested using FD-Apa1 (FD1414),

(Thermo Fisher Scientific, Waltham, USA) at 37°C for 1 h. The digestion enzymes where heat inactivated at 65°C for 5min as described by the provided protocol which came with the enzyme. The digested plasmid was transformed into the SBY65 and SBY66 yeast strain genome following the protocol described by J.H. Hegemann and S.B. Heick (Radamson, 2018). After incubation at 30°C for 3-5 days the yeast colonies where harvested. The correct integration of the construct was tested using PCR amplification of the construct of interest and via genetic sequencing. Positive colonies where re-plated on new YPD+URA plates and grown at 30°C for two days. Stocks were made by mixing 250 µL of 60% glycerol solution with 750 µL of YPD medium and the subsequent yeast colonies where suspended in this medium and stored at -80°C freezer.

3.1.2 64-well plate reader experiments

The microplate reader experiment was designed to measure the optimum functional range of the phenylalanine biosensor. Each test condition should be performed in triplicates. The yeast strains with the biosensor constructs were inoculated in DELFT minimalistic mineral media with ranging phenylalanine concentrations (0.2-4 mM) since previous literature has reported a similar functional ranges in case of *E. coli* (Mahr *et al.*, 2016). The conditions were divided as follows:

1) Control strain and functional biosensor strain DELFT+ L-Phe (0.2 mM) media 3x
2) Control strain and functional biosensor strain DELFT+ L-Phe (1 mM) media 3x
3) Control strain and functional biosensor strain DELFT+ L-Phe (2 mM) media 3x
4) Control strain and functional biosensor strain DELFT+ L-Phe (2.5 mM) media 3x
5) Control strain and functional biosensor strain DELFT+ L-Phe (4 mM) media 3x

Multiple control wells were used as follows:

6) Only DELFT media 3x
7) Each yeast strain in pure DELFT media 3x
8) All yeast strains in DELFT media with tryptophan (2 mM) 3x

This was done since L-tryptophan may cause unwanted TyrR dimerization (Pittard, Camakaris and Yang, 2005) and thus initiation of GFP transcription. Pure DELFT media with inoculated yeast cells were used in order to measure background fluorescence and help in determining the biosensor functional range. The optical density (OD) and fluores-

cence where measured each hour for a 24h long period and the data were analysed using Microsoft Excel computer program.

3.2 RESULTS & DISCUSSION

Here the author would like to explain the results yielded after DNA assembly steps and 64-well plate reader experiments just as well discuss the implications of each result.

3.2.1 Genetic construct assembly containing TyrR regulon sequence

The TyrR regulon will function as the sensing element of the biosensor. TyrR in the presence of L-phenylalanine will dimerise and bind to the CYC1p_T1_TYR promoter (Figure 17). To vertify correct DNA assembly enzymathic digestion and gel electrophoresis were employed (Figure 4). Two fragments were expected. The ~5 kb bigger fragment containing the digested pRS303 plasmid with the TyrR construct integrated. The DNA ladder employed is GeneRuler 1 kb Plus DNA Ladder (MAN0013047), (Thermo Fisher Scientific, Waltham, USA) and GeneRuler 50 bp DNA Ladder (SM037), (Thermo Fisher Scientific, Waltham, USA).. Proving that the Gibson assembly (E5510) performed was successful. The ~600bp size difference can be explained by the fact that one of the samples has a CYC1p promoter (290 bp) and the other containing a pFCY2 promoter (900 bp) accordingly.

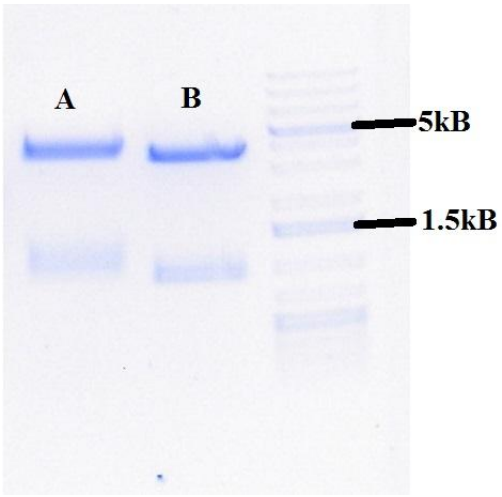


Figure 4. Gel electrophoresis results after construct plasmid digestion wit Xba1 and Pst1 enzymes (~4-4.5 kB) where A – construct containing pFCY2 and B – construct containing CYC1p.

After bacterial transormation was done, yielded colonies where tested by PCR amplifying the TyrR DNA sequence which is ~1500 bp long (Figure 5). The DNA ladder employed is

GeneRuler 1 kb Plus DNA Ladder. The results seem to be positive thus confirming the plasmid has been correctly absorbed by the *E. coli* bacteria. The amplification results showed expected sizes.

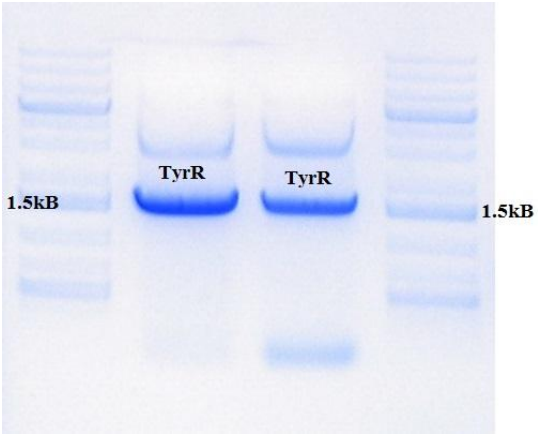


Figure 5. Amplification of the TyrR DNA sequence (~1.5 kB) from transformed *E. coli* bacterial colonies (primers 6 and 7 for CYC1p containing, 5 and 6 for pFCY2 containing construct, Table of primers, Appendix)

Lastly, after yeast transformation, the construct integration was confirmed by isolation of the yeast genomic DNA and PCR amplification of the TyrR gene sequence with the expected size of ~1500 bp (Figure 6). The DNA ladder employed is GeneRuler 1 kb Plus DNA Ladder. The results showed expected sizes.

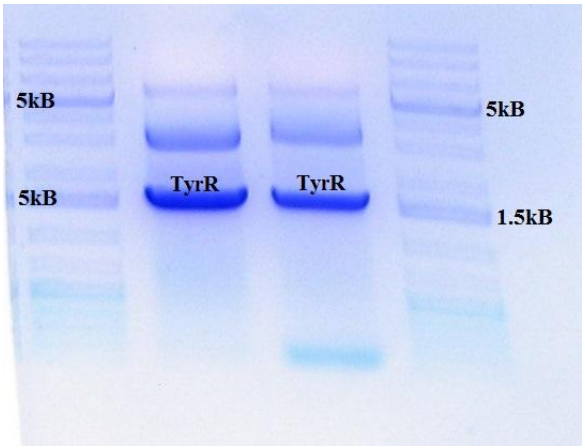


Figure 6. PCR amplified TyrR DNA sequence from yeast genomic DNA first well has the construct containing CYC1p and second has pFCY2(primers 6 and 7 for CYC1p containing, 5 and 6 for pF-CY2 containing construct, Table of primers, Appendix).

3.2.2 CYC1p_T1_TyrR promoter assembly

To assemble the CYC1p_T1_TyrR promoter each of the seperately PCR amplified DNA sequences (CYC1p_T1 ~142 bp and T1_TyrR ~212 bp) where electrophorated in 1% agarose gel (Figure 7). The DNA ladder employed was. The DNA ladder employed is GeneRuler 50 bp DNA Ladder.

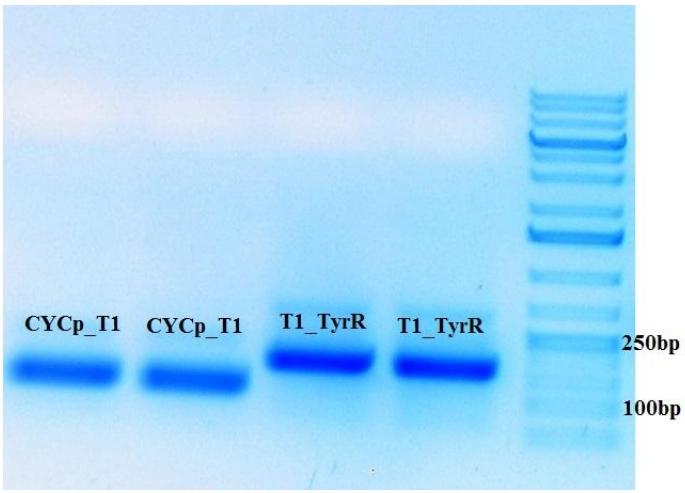


Figure 7. PCR amplified fragments of CYC1p_T1 (~142 bp), (primers 14 and 15, Table of primers, Appendix) and T1_TyrR (~212 bp), (primers 16 and 17, Table of primers, Appendix).

The gel extraction was followed by overlapping PCR to assemble the CYC1p_T 1 and T1_TyrR fragments into the CYC1p_T1_TyrR promoter (~300 bp), (Figure 8). The PCR products were electrophorated and later verified using DNA sequencing which confirmed succesfull promoter construction.

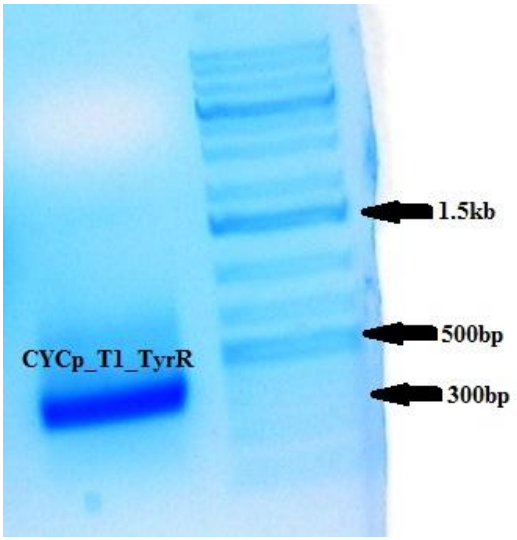


Figure 8. Gel picture of the overlapping PCR, (primers 14 and 17, Table of primers, Appendix sample reulting in the CYC1p_T1_TyrR promoter.

3.2.3 Genetic construct assembly containing the GFP reporter protein

Initially the author tried assembling the individual CYC1p_T1_TyrR/CYC1p promoter (~300 bp), (primers 3 and 4, Table of primers, Appendix) with GFP gene (720 bp), (primers 12 and 13, Table of primers, Appendix), tADH1 terminator (350 bp), (primers 8 and 9, Table of primers, Appendix) and the pRS304 plasmid (4443 bp), digested with FD EcoR1 and FD Bcu1, parts. The Gibson Assembly was done following the same guidlines as with the TyrR genetic construct for assembly of four to six parts. The parts were initially PCR amplified to introduce homologous regions for each of the respective parts to assemble and gel purified (Figure 9) including pRS304 digested plasmid (Figure 10). The DNA ladder employed is GeneRuler 1 kb Plus DNA Ladder. The fragments were within the expected size range. Confirming that the PCR amlification worked and homologous sticky ends should have been introduced.

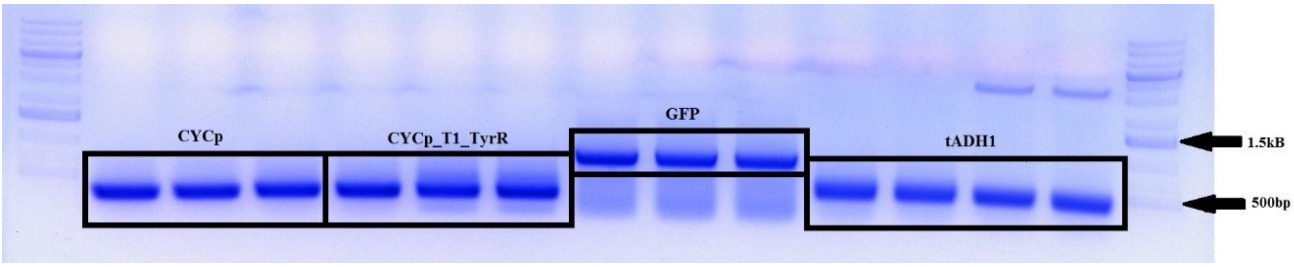


Figure 9. CYC1p construct fragment PCR amplification results.

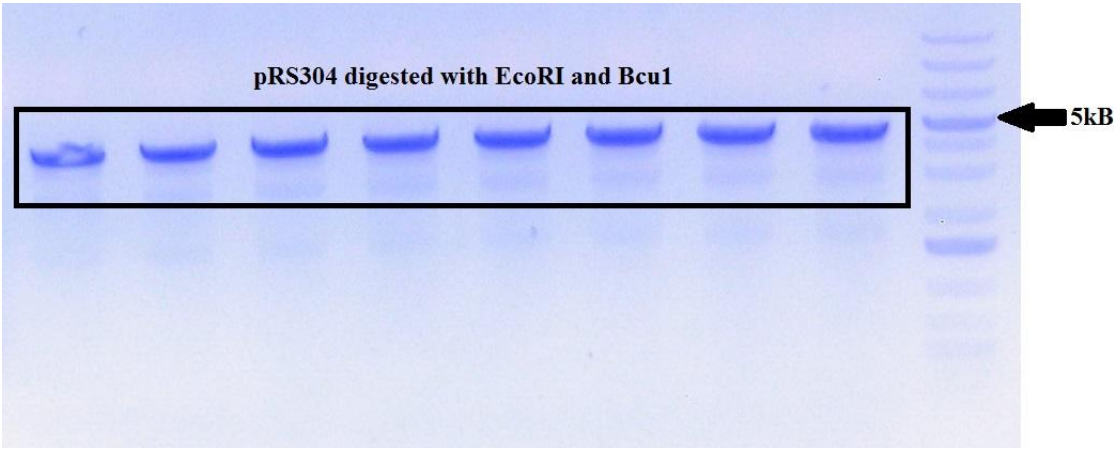


Figure 10. Plasmid's pRS304 digestion with EcoR1 and Bcu1.

After the bacterial transformation colonies which grew where checked for whole construct (CYC1p/CYC1p_T1_TyrR+GFP+tADH1, ~1350 bp) correct assembly via colony PCR (primers 3 and 9, Table of primers, Appendix), (Figure 11).

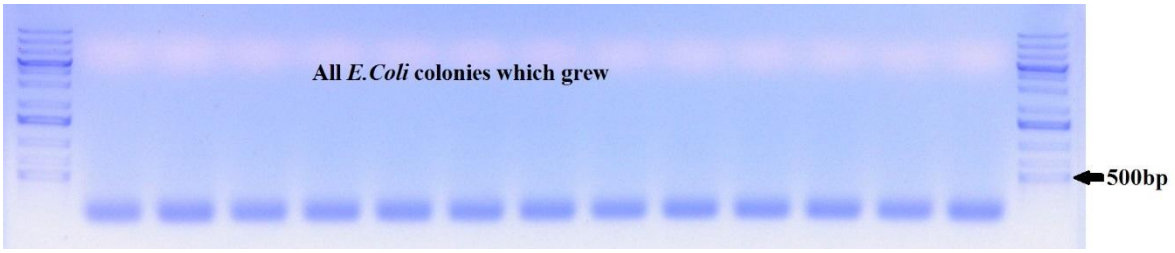


Figure 11. Colony PCR to check for whole construct assembly.

Aftter receiving the author troubleshoot, the reasons for the failure and PCR amplified GFP gene (720 bp) from its original source (Figure 12) and sent the purified GFP samples for DNA sequencing. The results showed that the GFP DNA sequence had undergon genetic mutations and the DNA sequence had frame shifted and thus rendered the GFP protein unfunctional.

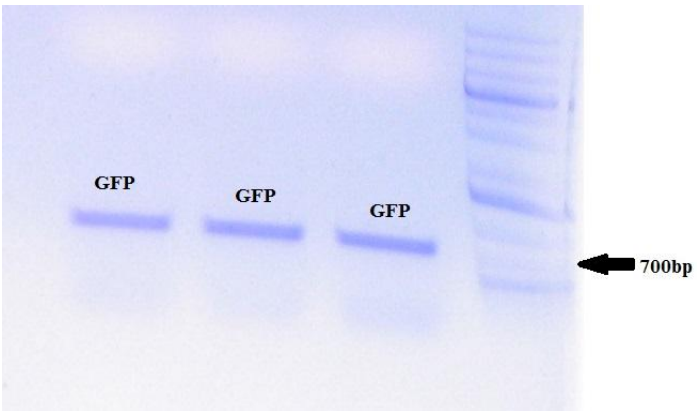


Figure 12. GFP (720 bp) gene amplification PCR result (primers 12 and 13, Table of primers, Appendix)

After learning that the current GFP used is unfunctional the author opted for a different approach for CYC1p and CYC1p_T1_TyrR fragment integration into a plasmid so a construct which was already sequenced and confirmed was provided by the SynBio Era laboratory. It already consisted of pGAL1+GFP+tADH1+pRS306 construct.

The promoters CYC1p (~280 bp) and CYC1p_T1_TYR (~320 bp), (Figure 13) and the plasmid GFP+CYC1t+pRS306 (~5350 bp), (Figure 14) were PCR amlified to introduce homologous overhangs for Gibson assembly. The PCR products where electrophorated and later extracted from the 1% agarose gel. The DNA ladder employed is GeneRuler 1 kb Plus DNA Ladder. The fragments had expected sizes.

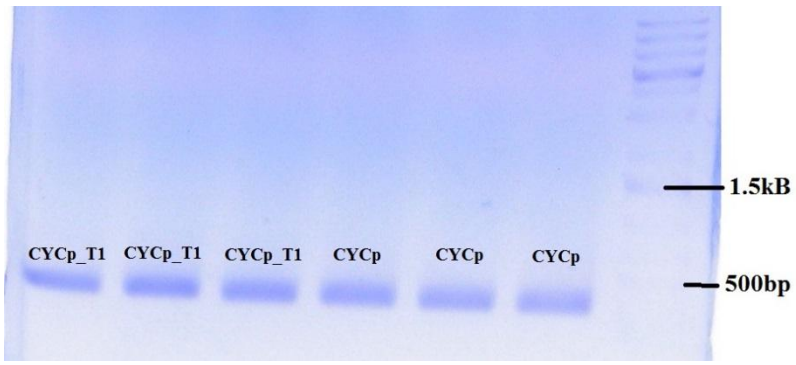
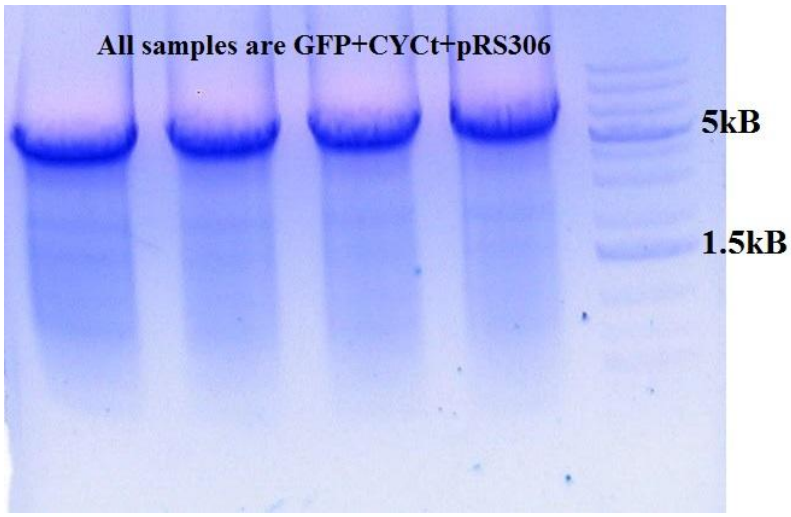


Figure 13. PCR amplified CYC1p (~290 bp) and CYC1p_T1_TYR fragments (~320 bp), (primers 20 and 21, Table of primers, Appendix).



Picture 14. PCR amplified GFP+CYC1t+pRS306 plasmid (~5350bp), (primers 20 and 22, Table of primers, Appendix).

These samples served as the basis for the Gibson assemblies performed. In order to increase the probability of a succesfull reaction the pRS306 plasmid was purified using Zymo-Spin 1 DNA extraction columns (C1003-50), (Irvine, USA) which has been shown to produce more highly concentrated DNA samples in comparison with the traditional columns used which allowed for optimum Gibson assembly conditions. They also served to verify that the PCR amplification of the constructs was correct.

After Gibson assembly and bacterial transformation the correct assembly (Figure 15) was verified using colony PCR to amplify the whole construct (CYC1p/CYC1p_T1_TRY+GFP+CYC1t) which after electrophorated yielded fragments with the expected sizes (~1405 bp). The DNA ladder employed is GeneRuler 1 kb Plus DNA Ladder.

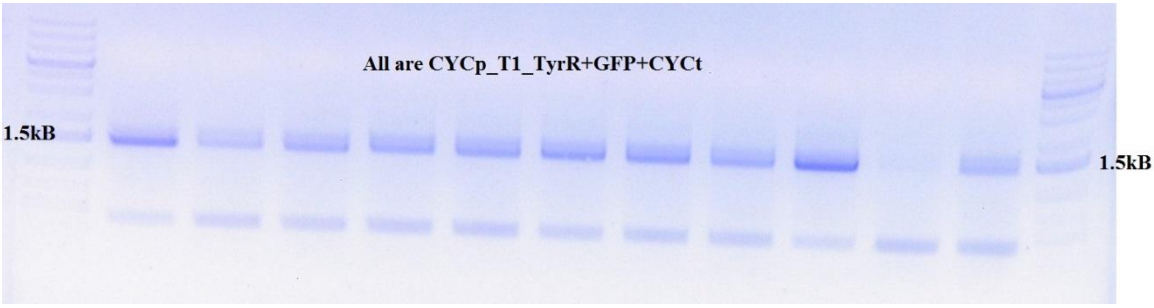


Figure 15. Colony PCR results for amplification of CYC1p/CYC1p_T1_TRY+GFP+CYC1t constructs (~1405 bp) confirming assembly (primers 19 and 20, Table of primers, Appendix).

These results confirmed that the Gibson assembly was succesful and the author can extract the plasmids from the *E. coli* bacterial cultures.

For integration into the *S. cerevisiae* genome the plasmid containing the aforementioned construct was digested with the FD Apa1 enzyme (~5.6 kB) and vertified using electrophoration (Figure 16). The DNA ladder employed is GeneRuler 1 kb Plus DNA Ladder. The results showed the expected DNA sizes except for the last one which is the same plasmid, but undigested funcioning as a control well.

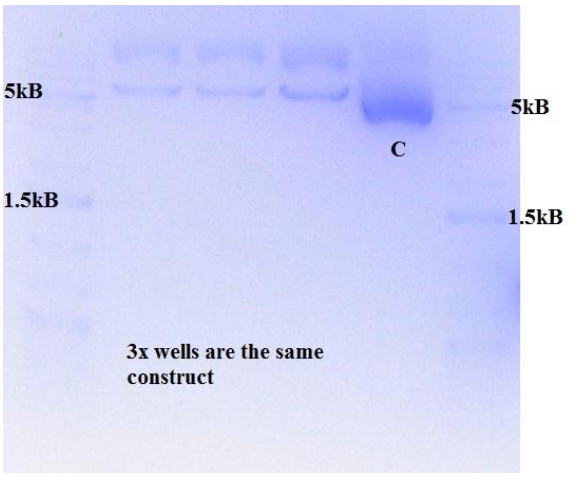


Figure 16. Digested CYC1p/CYC1p_T1_TRY+GFP+CYC1t+pRS306 constructs using FD Apa1 enzyme. Letter C - undigested plasmid as control.

Overall the current results are promising, but given the current time limitations the L-phenylalanine biosensor will have to be tested and charcterized after the completion of the authors bachelor's program, which is a pitty. However, a design ofr the biosensor together with significant progress towards the construction of the sensor has been made (Figure 17). The L-phenylalanine biosensor consists of two parts – the transcription factor and the reporter gene (GFP) construct. After TyrR DNA sequence into the TyrR protein it can

recognize L-phenylalanine in the intracelullar space and dimerize. The dimerized TyrR binds to the TyrR strong box also known as mtr strong box (Mahr *et al.*, 2016) embedded within the T1 position of the CYC1p_T1_TyrR promoter and now can recruit the αCTD subunit of RNA polymerase and thus activate GFP gene transcription and produce a green fluorescent output.

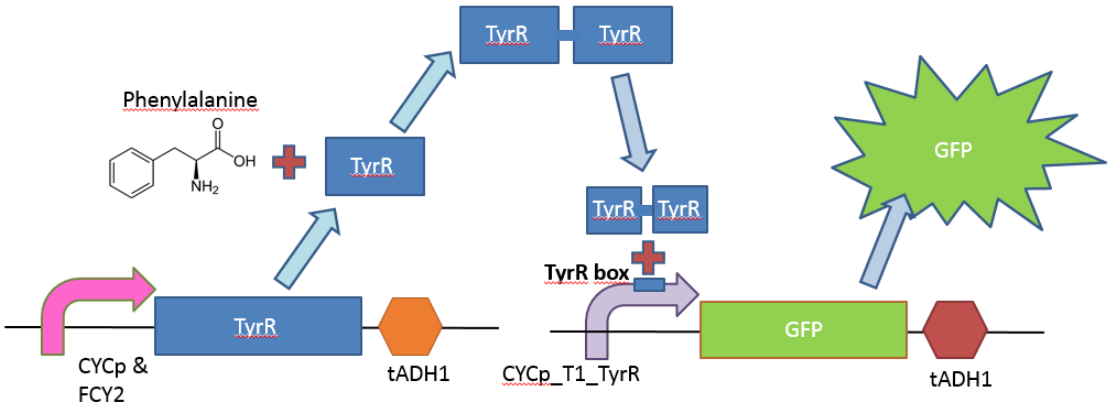


Figure 17. L-phenylalanine biosensor design consisting of a TyrR transcription factor construct and a reporter construct employing the CYC1p_T1_TyrR promoter capable of being bound by a dimerized TyrR protein in the presence of L-phenylalanine and initiating GFP transcription.

After succesful biosensor characterization many of the hypotheses descibed below can be verified or disproven. This will provide valuable insight for further various yeast based biosensor construction attempts. Especially for the use in selecting strains with upregulated aromatic amino acid or their derivative production.

By using the general biosensor design described in this thesis poses two limiting factors. First being whether or not the transcription factor of choice is native to the host organism and thus largely likely that functional or to some other microorganism. Using prokaryotic microorganism transcription factors have to be tested in yeast to determine wheather or not the biosensor is functional since its not known wether or not they undergo any post-translational modifications in eukaryotic cells thus rendering them unfunctional or distorting their functionality otherwise. The only family of prokaryotic transcriptional factors sure to function in yeast are LysR-type transcriptional regulators (LTTR's) which has been previously proven to work (Skjoedt *et al.*, 2016).

In order to make the biosensor useful it can be tested in other yeast strains designed for increased L-phenylalanine or its derivative production. The use of the biosensor could be widely used for adaptive laboratory evolution of such strains thus eliminating the need for alternative and expensive screening methods such as chromatography. The price, time and manpower for strain selection and evolution for increased target compound production would be easier then ever. The author suggests the biosensor screening method in the future to make the process more cost efficient and yielding.

It is also of interest to consider the fact that the CYC1p promoter employed in construction consists of two TATA box regions where the same or differing transcription factor binding sites can be employed thus allowing the construction of biosensors which function based on various logic gate systems and would allow for more specific conditions to be observed in very specific celullar stimuli. The results could still be measured using the fluorescent output at various induction conditions.

Overall transcription factor based biosensors have a lot of potential for both research and industrial use as long as the person understands the concepts necessary for its construction and has the neccesarry facilities for its development.

SUMMARY

Over the time of developing this project the author has learnt a multitude of genetic manipulation techniques and developed a thorough understanding of experimental process planning as well as deepened the knowledge about the mechanisms involved in metabolic pathways and cell survival.

The Gibson assembly method was succesfully established in the SynBio Era laboratory and is currently being employed in a multitude of projects being developed there.

A functional biosensor design has been succesfully developed based on previously done research. The design has been described in more detail in the „Results & Discussion“ part of this thesis paper.

The genetic pathway for TyrR regulon prouction in yeast has been succesfully developed and integrated within the *S. cerevisiae* genome. It has been confirmed by growing on selection plates containing the respective marker (LEU) and by DNA sequencing.

The CYC1p_T1_TyrR promoter has been assembled and verified to contain the TyrR strong box for TyrR regulon binding in its first TATA box via DNA sequencing. The constructed promoter has been succesfully integrated into the genetic construct containing the GFP coding sequence in parallel with the usual CYC1p promoter used as the control for measurments. The constructed plasmid has been verified using amplification PCR and DNA sequencing techniques. The integration of these sequence into the *S. cerevisiae* genome is still undergoing.

The characterization and functional range determination of the biosensor is still yet to be done in the future in both 64-well plate reader experiments and bio-reactor cultivation experiments.

In retrospect the project had its hurdles and problems, but with the newly developed troubleshooting capabilites these problems have been identified and succesfully overcome. All the goals could not be acieved because of time constraints, but the potential of the biosensor use in close future is likely as long as the finishing steps of characterization are performed.

REFERENCES

Bentley, R. (1990) 'Metabolic Tree with Many Branches', *Methods*, 25, pp. 307–383.

Clark, L. C. and Lyons, C. (1962) 'Electrode Systems for Continuous Monitoring in Cardiovascular Surgery', *Annals of the New York Academy of Sciences*, 102(1), pp. 29–45. doi: 10.1111/j.1749-6632.1962.tb13623.x.

Goers, L. *et al.* (2013) *Engineering Microbial Biosensors*. 1st edn, *Methods in Microbiology*. 1st edn. Copyright © 2013 Elsevier Ltd. All rights reserved. doi: 10.1016/B978-0-12-417029-2.00005-4.

Gosset, G. (2009) 'Production of aromatic compounds in bacteria', *Current Opinion in Biotechnology*, 20(6), pp. 651–658. doi: 10.1016/j.copbio.2009.09.012.

Hahlbrock, K. and Scheel, D. (1989) 'Biology of Phenylpropanoid Metabolism !', *Annual Review of plANT PHYSIOLOGY*, 40, pp. 347–369.

Herrmann, K. M. (2007) 'The Shikimate Pathway: Early Steps in the Biosynthesis of Aromatic Compounds', *The Plant Cell*, 7(7), p. 907. doi: 10.2307/3870046.

Herrmann, K. M. and Weaver, L. M. (2002) 'the Shikimate Pathway', *Annual Review of Plant Physiology and Plant Molecular Biology*, 50(1), pp. 473–503. doi: 10.1146/annurev.arplant.50.1.473.

Jiang, H. X. and Morgan, J. A. (2005) 'Metabolic engineering of the phenylpropanoid pathway in yeast for the synthesis of flavones.', *Abstracts of Papers of the American Chemical Society*, 229(6), pp. U180–U180. doi: 10.1128/AEM.71.6.2962.

Johnson, E. A. and Echavarri-Erasun, C. (2011) *Yeast biotechnology*, *The Yeasts*. Elsevier B.V. doi: 10.1016/B978-0-444-52149-1.00003-3.

Kunzelmann, S., Solscheid, C. and Webb, M. R. (2014) 'Fluorescent Methods for Molecular Motors', 105. doi: 10.1007/978-3-0348-0856-9.

MacDonald, M. J. and D'Cunha, G. B. (2007) 'Erratum: A modern view of phenylalanine ammonia lyase', *Biochemistry and Cell Biology*, 85(6), pp. 759–759. doi: 10.1139/o07-147.

Mahr, R. *et al.* (2016) 'Screening of an Escherichia coli promoter library for a phenylalanine biosensor', *Applied Microbiology and Biotechnology*. Applied Microbiology and Biotechnology, 100(15), pp. 6739–6753. doi: 10.1007/s00253-016-7575-8.

De Michele, R., Carimi, F. and Frommer, W. B. (2014) 'Mitochondrial biosensors', *International Journal of Biochemistry and Cell Biology*. Elsevier Ltd, 48(1), pp. 39–44. doi: 10.1016/j.biocel.2013.12.014.

Oldach, L. and Zhang, J. (2014) 'Genetically encoded fluorescent biosensors for live-cell visualization of protein phosphorylation', *Chemistry and Biology*. Elsevier Ltd, 21(2), pp. 186–197. doi: 10.1016/j.chembiol.2013.12.012.

Panche, A. N., Diwan, A. D. and Chandra, S. R. (2016) 'Flavonoids: an overview', *Journal of Nutritional Science*, 5. doi: 10.1017/jns.2016.41.

Pittard, J., Camakaris, H. and Yang, J. (2005) 'The TyrR regulon', *Molecular Microbiology*, 55(1), pp. 16–26. doi: 10.1111/j.1365-2958.2004.04385.x.

Radamson, H. H. (2018) 'Strain engineering', *CMOS Past, Present and Future*, 765, pp. 41–67. doi: 10.1016/b978-0-08-102139-2.00003-3.

Rogers, J. K., Taylor, N. D. and Church, G. M. (2016) 'Biosensor-based engineering of biosynthetic pathways', *Current Opinion in Biotechnology*. Elsevier Ltd, 42, pp. 84–91. doi: 10.1016/j.copbio.2016.03.005.

Shi, S., Ang, E. L. and Zhao, H. (2018) 'In vivo biosensors: mechanisms, development, and applications', *Journal of Industrial Microbiology & Biotechnology*. Springer Berlin Heidelberg, (0123456789). doi: 10.1007/s10295-018-2004-x.

Skjoedt, M. L. *et al.* (2016) 'Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast', *Nature Chemical Biology*, 12(11), pp. 951–958. doi: 10.1038/nchembio.2177.

Sprenger, G. A. (2007) 'From scratch to value: Engineering Escherichia coli wild type cells to the production of L-phenylalanine and other fine chemicals derived from chorismate', *Applied Microbiology and Biotechnology*, 75(4), pp. 739–749. doi: 10.1007/s00253-007-0931-y.

Suzuki, H. (2013) *Microbial production of amino acids and their derivatives for use in foods, nutraceuticals and medications*, *Microbial Production of Food Ingredients, Enzymes and Nutraceuticals*. Woodhead Publishing Limited. doi: 10.1533/9780857093547.2.385.

Weisshaar, B. and Jenkinst, G. I. (1998) 'Phenylpropanoid biosynthesis and its regulation', *Current Opinion in Plant Biology*, 1(3), pp. 251–257. doi: 10.1016/S1369-5266(98)80113-

1.

Appendix

I. Table of primers

1	pFCY2_F	CGCGGTGGCGGCCGCTCTAGCTAGAGATCTAACACTGATGAAAAG
2	pFCY2_R	CCTATCCTCTATTATAACGATTTATAA-GAATGCGTCTGGAAGTCTTTTG
3	CYC_amp_F	CGC GGT GGC GGC CGC TCT AGA GCG TTG GTT GGT GGA TCA A
4	CYC_amp_R	TCC TCG CCC TTG CTC ACC ATT TAG TGT GTG TAT TTG TGT TTG C
5	TyrR_FCY_F	TCTATTATAACGATTTATAAGAATGCGTCTGGAAGTCTTTTG
6	TyrR_R	AGTCAGAAGAAGAACGAAGAAGCTTTGGACTTCTTCGCCA
7	TyrR_CYC_F	AACACAAATACACACACTAAATGCGTCTGGAAGTCTTTTG
8	tADH1_CYC_F	TGGACGAGCTGTATAAGTAAAGCTTTGGACTTCTTCGCCA
9	tADH1_CYC_R	CGAAATTCCCCTACCCTATGAATTCGATATCAAGCTTATC
10	tADH1_TYR_F	AGTCAGAAGAAGAACGAAGAAGCTTTGGACTTCTTCGCCA

11	tADH1_TYR_R	ATGTACTAGTAGCGGCCGCTAATTCGATATCAAGCTTATC
120	GFP_F	AAC ACA AAT ACA CAC ACT AAA TGG TGA GCA AGG GCG AGG A
13	GFP_R	TGG CGA AGA AGT CCA AAG CTT TAC TTA TAC AGC TCG TCC ATG CC
14	CYCp_T1_F	ATATATGAATTCGCGGCCGCTTCTAGAGAGCGTTGGTTGGTGGATC
15	CYCp_T1_R	TATTTTACATATGTGTCAGCACTAAAGTTGCCTGGCC
16	T1_TYR_F	TGTAAAATACAGGCATATATATATGTGTGCGACGACACATGATC
17	T1_TYR_R	TATAAACTGCAGCGGCCGCTACTAGTATTAGTGTGTGTATTTGTGT
18	pRS306_F	GGCGAATTGGAGCTCTAGTAAGCGTTGGTTGGTGGATCAA
19	pRS306_R	TCTAGAATCCGGGGTTTTTTTTAGTGTGTGTATTTGTGTTTGC
20	306_CYC_F	AACACAAATACACACACTAAAAAAAACCCCGGATTCTAGA
21	306_CYC_R	TTGATCCACCAACCAACGCTTACTAGAGCTCCAATTCGCC
22	M13_R	CAGGAAACAGCTATGAC

NON-EXCLUSIVE LICENCE TO REPRODUCE THESIS AND MAKE THESIS PUBLIC

I, Dainis Ivanovs

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

L-Phenylalanine biosensor in *Saccharomyces cerevisiae*

supervised by Nemailla Bonturi and Petri-Jaan Lahtvee

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Dainis Ivanovs

20/05/2019