Konstrukcija sojeva Saccharomyces cerevisiae za produkciju aminokiselina nalik mikosporinu

Jurić, Vanja

University of Zagreb

Faculty of Science

Department of Biology

Vanja Jurić

CONSTRUCTION OF *Saccharomyces cerevisiae* STRAINS FOR MYCOSPORINE-LIKE AMINO ACIDS PRODUCTION

Graduation thesis

Zagreb, 2019

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Biološki odsjek

Vanja Jurić

KONSTRUKCIJA SOJEVA *Saccharomyces cerevisiae* ZA PRODUKCIJU AMINOKISELINA NALIK MIKOSPORINU

Diplomski rad

Zagreb, 2019.

The experimental work presented in this thesis was carried out at the Novo Nordisk Foundation Center for Biosustainability, Kongens Lyngby, Denmark, under the supervision of Irina Borodina, PhD and Iben Møller-Hansen, PhD, co-supervised by Associate Professor Ivana Ivančić Baće (University of Zagreb). This thesis is submitted for evaluation to the Department of Biology, Faculty of Science, University of Zagreb in order to acquire the Master of Molecular Biology title.

ACKNOWLEDGEMENTS

I would first like to thank my supervisors. To Irina Borodina for giving me the opportunity to join her research group and for all the valuable experience I gained there. To Iben Møller-Hansen for the continuous support of my research, for her positive attitude, motivation, and advice. I deeply appreciate the knowledge they shared with me, and I will try to perform up to their standard in my future work and study. I would also like to thank my co-supervisor Ivana Ivančić Baće for investing her time into helping me out with thesis writing.

I would like to thank my fellow labmates and master's students for their feedback and cooperation, for their friendship and for all the fun we had.

Finally, I must express my gratitude to my family and to my girlfriend for providing me with support and encouragement throughout my study and through the process of researching and writing this thesis.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek Diplomski rad

KONSTRUKCIJA SOJEVA KVASCA *Saccharomyces cerevisiae* ZA PRODUKCIJU AMINOKISELINA NALIK MIKOSPORINU

Vanja Jurić
Rooseveltov trg 6, 1000 Zagreb, Hrvatska

Aminokiseline nalik mikosporinu su spojevi poznati po upijanju UV zračenja te su prisutni u brojnim morskim mikroorganizmima koji luče ove sekundarne metabolite da bi se zaštitili u uvjetima stresa izazvanog UV zračenjem. Ovi spojevi tipično apsorbiraju na valnim duljinama od 310 nm do 360 nm. Njihova fotoprotektivna svojstva čine ih obećavajućim kandidatima za farmaceutske i kozmetičke aplikacije. Cilj ovog istraživanja bila je konstrukcija sojeva kvasca *Saccharomyces cerevisiae* koji će u kromosomima 10 i 11 imati stabilno integrirani heterologni biosintetski put iz cijanobakterije *Anabaena variabilis* za proizvodnju aminokiselina nalik mikosporinu. Konstruirana su tri integrativna plazmida koji sadržavaju gene za biosintetski put iz *A. variabilis* uz odgovarajuće konstitutivne promotore i terminatore. Put je ugrađen u haploidne i diploidne sojeve *S. cerevisiae* te u derivate tih sojeva koji eksprimiraju heterologni put za konzumaciju ksiloze kako bi uz glukozu mogli rasti i na mediju s ksilozom. Ispravnost plazmidnih konstrukata je potvrđena metodom analize kolonija PCR-om te sekvenciranjem. Metodom analize kolonija PCR-om je potvrđena ugradnja vektora u kromosome *S. cerevisiae*, a metodom tekućinske kromatografije visoke djelotvornosti (HPLC) te masenom spektrometrijom (LC-MS) detektirani su proizvedeni spojevi iz puta sinteze aminokiselina nalik mikosporinu u konstruiranim sojevima.

(55 stranice, 16 slika, 6 tablica, 44 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u Središnjoj biološkoj knjižnici
Ključne riječi: aminokiseline nalik mikosporinu, put ksiloza izomeraze, put pentoza fosfata

Voditelj:	Irina Borodina, PhD
Suvoditelj:	izv. prof. dr. sc. Ivana Ivančić Baće
Ocjenitelji:	izv. prof. dr. sc. Ivana Ivančić Baće
	doc. dr. sc. Sunčica Bosak
	doc. dr. sc. Tomislav Ivanković

Rad prihvaćen: 3. listopada 2019.

BASIC DOCUMENTATION CARD

University of Zagreb
Faculty of Science
Department of Biology

Graduation Thesis

CONSTRUCTION OF *Saccharomyces cerevisiae* STRAINS FOR MYCOSPORINE-LIKE AMINO ACIDS PRODUCTION

Vanja Jurić
Rooseveltov trg 6, 1000 Zagreb, Hrvatska

Mycosporine-like amino acids are compounds known for their UV-absorbing properties and are found in numerous marine microorganisms which produce these secondary metabolites as a way of dealing with UV-induced stress. Their absorption maxima typically range from 310 to 360 nm. The aim of this thesis was to construct *Saccharomyces cerevisiae* strains with heterologous biosynthetic mycosporine-like amino acids pathway from cyanobacterium *Anabaena variabilis* stably integrated into chromosomes 10 and 11. Three plasmids containing the pathway from *A. variabilis* with corresponding constitutive promoters and terminators were constructed. The pathway was integrated into haploid and diploid *S. cerevisiae* strains and into their derivatives which express a heterologous pathway for xylose consumption so that the strains would be able to grow on medium supplemented with xylose, in addition to glucose. Plasmid constructs were verified by colony PCR and sequencing. Colony PCR was used to confirm vector integration into *S. cerevisiae* chromosomes and high performance liquid chromatography (HPLC) combined with mass spectrometry (LC-MS) was used to detect the produced compounds from the mycosporine-like amino acids pathway in constructed yeast strains.

(55 pages, 16 figures, 6 tables, 44 references, original in: English)

Thesis deposited in Central Biological Library
Key words: mycosporine-like amino acids, xylose isomerase pathway, pentose phosphate pathway

Supervisor:	Irina Borodina, PhD
Co-supervisor:	dr. sc. Ivana Ivančić Baće, Assoc. Prof.
Reviewers:	dr. sc. Ivana Ivančić Baće, Assoc. Prof.
	doc. dr. sc. Sunčica Bosak
	doc. dr. sc. Tomislav Ivanković

Thesis accepted: October 3rd, 2019

1 INTRODUCTION

1.1 MYCOSPORINE-LIKE AMINO ACIDS

1.1.1 PROPERTIES AND APPLICATIONS

Ultraviolet UV-A (315–400 nm) and UV-B (280–315 nm) radiation is known to cause damage to biological macromolecules (Setlow and Setlow, 1962). For microorganisms exposed to high-intensity light, especially marine microorganisms, an important strategy to deal with UV radiation is the biosynthesis of photoprotective molecules, essentially serving as sunscreens. Among such metabolites, mycosporines and mycosporine-like amino acids (MAAs) are known for their unique chemical structures and their wide phylogenetic distribution, as they can be found in cyanobacteria, fungi and algae. They are characterized by their small molecular weight (< 500 Da), colorlessness, high polarity and water solubility. They were named for their ability to mediate light-induced fungal sporulation, with over 30 of them identified to date (Bandaranayake, 1998).

Mycosporines are composed of a cyclohexenone core linked with a proteogenic or modified amino acid residue, while mycosporine-like amino acids are linked with two such residues. The amino acids are conjugated to the central ring through imine linkages (Sinha et al., 2007). Depending on their molecular structure, their absorption maxima typically range from 310 to 360 nm and their molar extinction coefficients (ε) range from 28,100 to 50,000 M^{-1} cm^{-1} (Vaishampayan et al., 1998; Richardson and Jackson, 2007). The extensive conjugation (i.e., alternating single and double bonds) present in MAA structures facilitates UV absorption. UV-absorbing compounds often generate free oxygen radicals, which is detrimental to the biological system, but MAAs dissipate the excess energy as heat without generating free oxygen radicals (Conde *et al.*, 2004). Figure 1 presents structures of shinorine (Simon et al., 2014) and porphyra-334 (Klisch et al., 2007) as examples of MAAs, along with 4-deoxygadusol (4-DG) which serves as a cyclohexenone core.

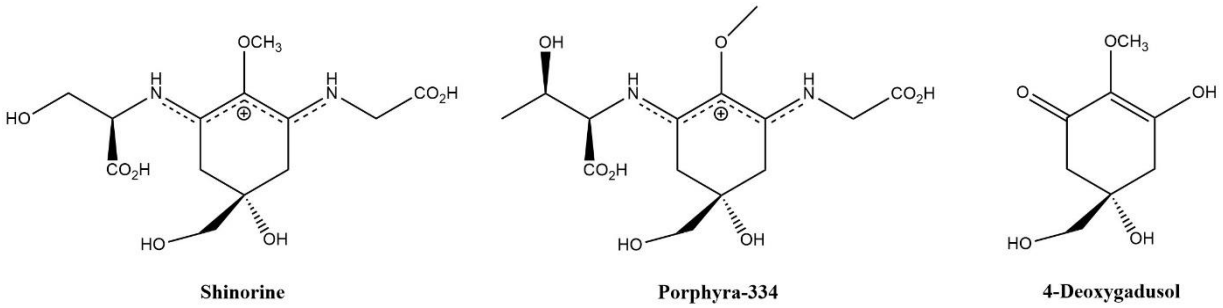

Figure 1. Structures of shinorine, porphyra-334 and 4-deoxygadusol. Shinorine and porphyra-334 are represented as resonance hybrids to showcase the conjugated bonds which facilitate UV absorption. 4-deoxygadusol serves as a core to which amino acids are attached.

Their properties make them promising candidates for pharmaceutical and cosmetic applications and some of them, such as shinorine and porphyra-334 which can be found in cyanobacteria *Anabaena variabilis* and *Porphyra umbilicalis*, have already been commercialized as active ingredients in sunscreen creams (Cardozo et al., 2007).

1.1.2 METABOLIC ROUTES TO MYCOSPORINE-LIKE AMINO ACIDS

1.1.2.1 PENTOSE PHOSPHATE PATHWAY

The pentose phosphate pathway (Fig. 2), also called the phosphogluconate pathway or the hexose monophosphate shunt is a pathway parallel to glycolysis. There are two distinct phases in the pentose phosphate pathway. The first phase is the oxidative phase in which the reducing agent nicotinamide adenine dinucleotide phosphate (NADPH) is generated, and the second phase is the non-oxidative phase in which 5-carbon sugars are synthesized. The oxidative phase is an irreversible process and the produced NADPH is used in reductive biosynthesis in cells, such as fatty acid synthesis. NADPH is also used to protect cells from oxidative stress, since it is a cofactor for glutathione- and thioredoxin-dependent enzymes that defend cells against oxidative damage. The non-oxidative phase is a reversible phase in which a series of sugars is produced: ribose-5-phosphate is used in the synthesis of nucleotides, while erythrose-4-phosphate is used in the synthesis of aromatic amino acids. In yeast, the pentose phosphate pathway takes place in the cytosol (Berg *et al.*, 2002). Sedoheptulose 7-phosphate, one of the intermediates of the pentose phosphate pathway is used as a precursor to a range of mycosporines and mycosporine-like amino acids (Balskus and Walsh, 2010).

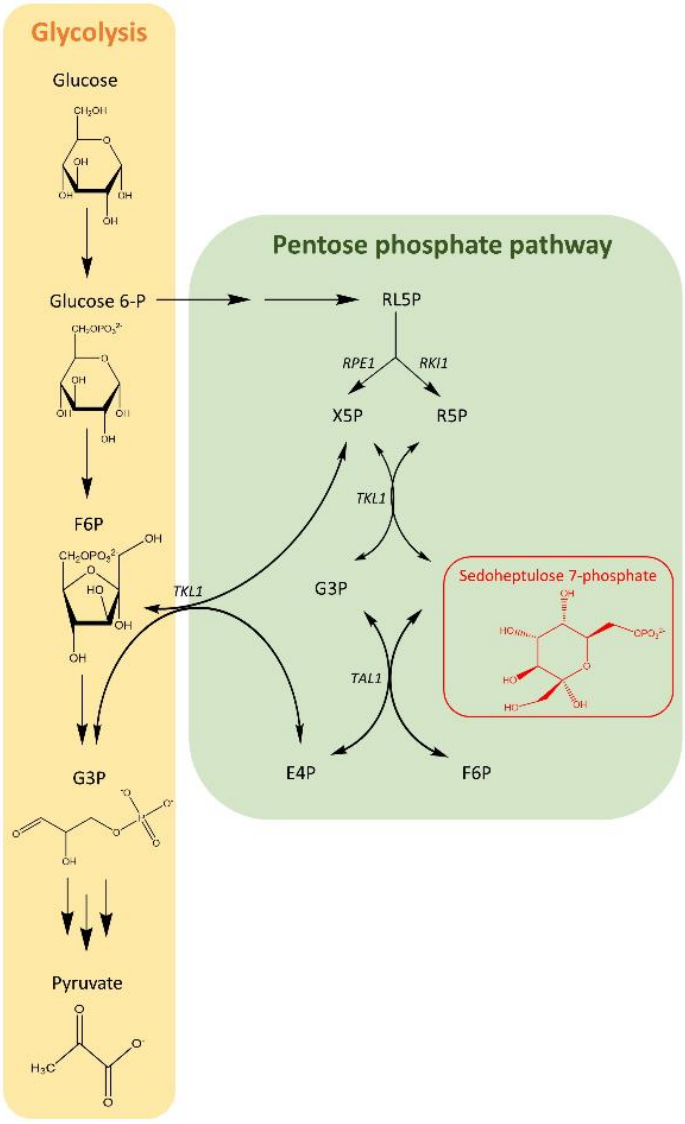

Figure 2. Schematic representation of the non-oxidative phase of the pentose phosphate pathway and glycolysis. Sedoheptulose 7-phosphate, a precursor of mycosporines and mycosporine-like amino acids is shown in red. Multiple arrows represent multiple enzymatic reactions grouped for simplicity. Glucose 6-P; glucose 6-phosphate, F6P; fructose 6-phosphate, G3P; glyceraldehyde 3-phosphate, RL5P; ribulose 5-phosphate, *RPE1*; D-ribulose-5-phosphate-3-epimerase, *RKI1*; ribose-5-phosphate ketol-isomerase, X5P; xylulose 5-phosphate, R5P; ribose 5-phosphate, *TKL1*; transketolase 1, *TAL1*; transaldolase 1, E4P; erythrose 4-phosphate.

In *Saccharomyces cerevisiae*, the pentose phosphate pathway is used in the industrial production of ethanol from xylose. In production based on yeast cell factories, pure glucose is one of the major costs. Xylose serves as a low-cost and abundant material and can be found in agricultural wastes, wood and wastes from pulp and paper industry can be potentially used as a cheaper carbon source, hence lowering the cost of the product (Walfridsson et al., 1995). *S. cerevisiae* does not naturally metabolize xylose, but recombinant strains containing genes from

Pichia stipitis have been designed to metabolize xylose by incorporating it in the pentose phosphate pathway (Jeppsson et al., 2002).

1.1.2.2 MAA BIOSYNTHESIS FROM SEDOHEPTULOSE 7-PHOSPHATE

Two alternative pathways for MAA synthesis have been suggested: one of them uses 3-dehydroquinate (DHQ) as a precursor, an intermediate of the shikimate pathway, while the other pathway uses sedoheptulose 7-phosphate (S7P) as a precursor, an intermediate of the pentose phosphate pathway (Favre-Bonvin et al., 1987). The S7P pathway has been proven to be involved in MAAs mycosporine-glycine and shinorine synthesis based on the biochemical characterization of MAA gene clusters in cyanobacteria *Anabaena variabilis* ATCC 29413 and *Nostoc punctiforme* ATCC 29133 and their successful expression in a heterologous bacterial host. In addition, analogous pathways have been identified in other sequenced organisms (Balskus and Walsh, 2010).

S7P is converted to shinorine in four enzymatic steps (Fig. 3). In the first step, S7P is converted to 2-demethyl-4-deoxygadusol (DDG) by 2-demethyl 4-deoxygadusol synthase (DDGS). DDG is then converted to 4-deoxygadusol (4-DG) via O-methyltransferase (OMT). 4-DG serves as a cyclohexenone core to which amino acids are attached in the last two steps. Glycine is conjugated to 4-DG by ATP-grasp ligase, forming mycosporine-glycine (MG). The pathways in *A. variabilis* and *N. punctiforme* differ in the last step: in *A. variabilis,* nonribosomal peptide synthetase (NPRS)-like enzyme attaches serine to MG, finally forming shinorine, while in *N. punctiforme,* serine is attached to MG via D-ala-D-ala ligase, also forming shinorine (Balskus and Walsh, 2010; Katoch et al., 2016a).

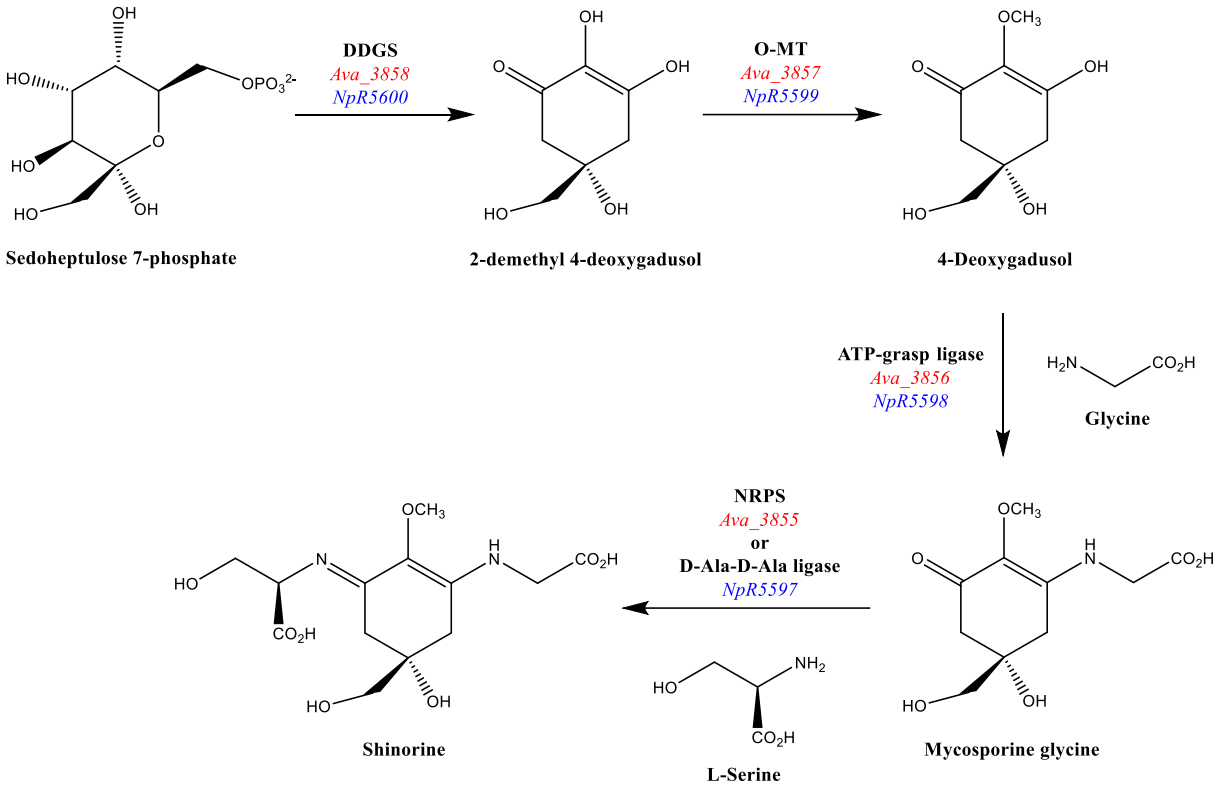

Figure 3. Schematic diagram of shinorine biosynthesis. Sedoheptulose 7-phosphate, a pentose phosphate pathway intermediate, is converted to shinorine with 2-demethyl 4-deoxygadusol, 4-deoxygadusol and mycosporine-glycine as intermediates by sequential catalytic reactions of DDG synthase (DDGS), O-methyl transferase (O-MT), ATP-grasp ligase, and nonribosomal peptides synthetase (NRPS) in *A. variabilis* or D-Ala-D-Ala ligase in *N. punctiforme*. Corresponding genes from *A. variabilis* are shown in red, while genes from *N. punctiforme* are shown in blue.

To date, several MAAs have been produced by introducing heterologous genes into a range of bacterial hosts. Expression of MAA biosynthetic genes from *Actinosynnema mirum* DSM 43827 in *Streptomyces avermitilis* SUKA22 resulted in 154 mg/L shinorine and 188 mg/L total MAAs, while the expression of the same genes in *Corynebacterium glutamicum* produced only 19 mg/L of shinorine (Tsuge et al., 2018). Expression of shinorine biosynthetic gene cluster from *A. variabilis* in *Escherichia coli* produced a low amount of shinorine of around 0.15 mg/l (Balskus and Walsh, 2010), using sedoheptulose 7-phosphate supplemented in the medium as the substrate. Genes from *Nostoc punctiforme* ATCC 29133 have been used to successfully produce 31 mg/l of shinorine in *Saccharomyces cerevisiae*, using xylose as the substrate (Park et al., 2019).

1.2 USES OF *Saccharomyces cerevisiae* IN MOLECULAR BIOLOGY AND BIOTECHNOLOGY

The yeast *Saccharomyces cerevisiae* has been used as a model organism in molecular biology for decades. Like *Escherichia coli*, it is unicellular and can be grown in culture in large numbers easily. However, since it is a eukaryote, results from studies on *S. cerevisiae* can more easily be applied to other eukaryotes such as humans. Another important distinction of *S. cerevisiae* as a model organism is the ease of genetic manipulation – genes can be easily moved to or removed from yeast cells either on plasmids or they can be integrated into chromosomes. There is a range of selectable markers for use in yeast and standardized plasmids with naturally occurring origins of replication (autonomously replicating sequences – ARS) which were adapted to replicate plasmids in yeast. Furthermore, *S. cerevisiae* uses homologous recombination to readily integrate transformed DNA into its genome (Duina et al., 2014).

In nature, *S. cerevisiae* can readily switch between two mating types: haplotype *MAT*a which can mate with the other haplotype *MAT*α to form diploid *MAT*a/*MAT*α cells. Haploid cells undergo mitotic cell division through budding. Under nutrient-poor conditions, diploid cells undergo meiosis and sporulation, producing four haploid spores which germinate into two *MAT*a and two *MAT*α cells. However, in laboratory strains, mating types are often stable and do not switch due to the absence of a functional *HO* (homothallic switching) endonuclease (Duina et al., 2014).

S. cerevisiae has been widely used in biotechnology. In addition to traditional uses like alcohol fermentation and baking, genetic engineering techniques have been applied to yeast to produce many industrially relevant chemicals and for heterologous expression of proteins. In addition, *S. cerevisiae* is GRAS (Generally Recognized as Safe) approved by the United States Food and Drug Administration ("Microorganisms & Microbial-Derived Ingredients", 2015), which makes products intended for human consumption or application generally easier to approved for use compared to production in other organisms. Genetic engineering of yeast for biosynthesis generally involves the conversion of precursor metabolites into products of interest with several or all of the steps in the pathway catalyzed by heterologous enzymes. The host genome is therefore manipulated to increase the activity of enzymes making the product of interest and to decrease the activity of enzymes in competing pathways (Borodina and Nielsen, 2014). Compared to prokaryotes, the presence of organelles in *S. cerevisiae* can be used to compartmentalize reactions in different environments, therefore avoiding competing pathways

or trapping hydrophobic or toxic products (Lian et al., 2018). Even though some heterologous pathways are hard to replicate in *S. cerevisiae* due to species-specific traits, *S. cerevisiae* is still the host of choice in many cases, due to the readily available genetic engineering tools and vast knowledge about yeast molecular biology (Kavščeк et al., 2015).

In recent years, new methods like Zinc finger nucleases, transcription activator like effector nucleases (TALEN) and the CRISPR-Cas system (clustered regularly interspaced short palindromic repeats) have been used for genome editing. These methods are more efficient, more accurate and faster compared to traditional methods. The CRISPR-Cas system has been used more extensively compared to other methods because of the simplicity of its design and cost-effectiveness. CRISPR-Cas genetic editing utilizes homologous recombination. Compared to other traditional methods using homologous recombination which use selection markers which then have to be removed to reuse in case of integrating additional genes, CRISPR-Cas editing can be done marker-free, as the double-stranded break made by Cas enzyme actually serves as a method of counterselection. The frequency of homologous recombination events is also greatly increased with Cas-induced DNA damage compared to traditional gene targeting approaches. (Kavščeк et al., 2015).

1.3 GENETIC ENGINEERING TOOLS FOR GENE INTEGRATION INTO *Saccharomyces cerevisiae* CHROMOSOMES

1.3.1 USER CLONING

Uracil-specific excision reagent (USER®) cloning method (Bitinaite et al., 2007) is used for directional assembly of one or several DNA fragments and cloning into a vector (Fig. 4). The USER enzyme is a mixture containing uracil DNA glycosylase and DNA glycosylase-lyase Endo VIII. Target DNA and vectors are generated by polymerase chain reaction (PCR) with overhangs of 6-10 bases of homology at their ends. In the homology region, PCR primers contain a single deoxyuracil residue (dU) which is recognized by the USER enzyme. The USER enzyme removes the dU residue and generates DNA fragments with 3’ overhangs. The primer design is such that the generated fragments have complementary 3’ overhangs, enabling directional assembly of DNA. Vectors are also amplified with primers containing dU, thus enabling DNA fragment insertion into linearized vectors with complementary overhangs. In case of hybridization of DNA fragment and destination vector with over 8 nucleotides long

complementary overhangs, it is sufficient to transform the reaction mixture directly into *E. coli* without previous ligation (Nour-Eldin et al., 2006).

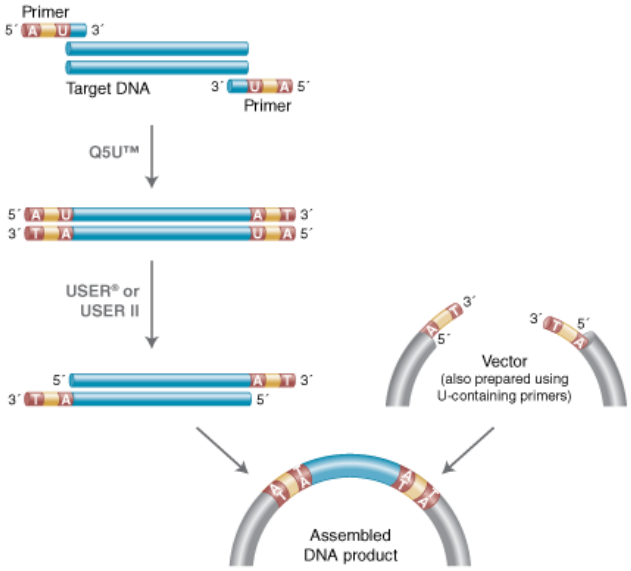

Figure 4. DNA assembly with USER® method. Target DNA is amplified with primers containing a single dU residue which is recognized and removed by the USER enzyme, therefore generating 3' overhangs. Vectors (or additional inserts) are prepared in the same way with dU-containing primers, making the linearized vector ends complementary to the target DNA. Figure taken from Bitinaite et al. (2007).

1.3.2 EASYCLONE VECTOR SET

Genes and promoters are PCR amplified with primers for USER cloning, creating cassettes which consist of two genes with a bidirectional promoter between them or of one gene with one promoter. These cassettes are cloned into a linearized EasyClone vector (Jessop-Fabre et al., 2016). The linearized vector contains terminator sequences at each end, with their ends designed to be complementary to the 3' USER overhangs of the cassettes. The insert/vector mix is transformed into *E. coli*, whose endogenous repair system ligates the inserts and the vector. The vectors are then propagated in and purified from *E. coli*.

EasyClone vectors are designed for the integration of genes into *S. cerevisiae* using CRISPR-Cas9. Each of these vectors contains two sequences (designated as Up and Down regions) homologous to a *S. cerevisiae* chromosome, which after the vectors (containing desired genes) are linearized facilitate homologous recombination and integration into *S. cerevisiae* chromosomes. These integration sites have previously been designed to be chromosomal intergenic stable sites, avoiding risk of chromosomal rearrangements. Each of these vectors has

a corresponding gRNA vector, necessary to guide Cas9 to the target chromosomal sequence to introduce a double stranded break (Stovicek et al., 2015). The technique is explained in detail with specific examples in the Materials and Methods section.

2 PROJECT OBJECTIVES

The goal of this research was to engineer *Saccharomyces cerevisiae* strains with stable integration and overexpression of four codon-optimized genes from cyanobacterium *Anabaena variabilis* ATCC 29413 for the production of mycosporine-like amino acids (MAAs). To do that, the following aims needed to be accomplished:

- Construction of two expression cassettes, with each cassette containing two out of four biosynthetic genes and two strong constitutive promoters, with the cassettes being flanked by regions homologous to *S. cerevisiae* chromosomal sequences
- Construction of a third expression cassette containing four genes and four promoters, by combining the previous two constructed cassettes, with the cassette being flanked by regions homologous to chromosomal sequences from Ethanol Red®, an industrial diploid strain of *S. cerevisiae*
- Assembly of the expression vector in *E. coli* and verification of correct assembly by colony PCR and sequencing
- Transformation of expression vectors into different parental *S. cerevisiae* strains, haploid and diploid
- To increase the MAA substrate sedoheptulose 7-phosphate pool, transformation of expression vectors into parental *S. cerevisiae* strains which heterologously express bacterial genes for xylose consumption
- Integration of expression vectors into *S. cerevisiae* chromosomes using CRISPR-Cas9
- Verification of correct vector integration by colony PCR
- Growth of the constructed strains in media containing glucose and xylose
- Detection of MAAs production using high performance liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS)

3 MATERIALS AND METHODS

3.1 CHEMICALS AND REAGENTS

Table 1. Chemicals and reagents.

Name	Manufacturer
Agarose	Sigma-Aldrich
Ammonium formate	Sigma-Aldrich
Ampicillin	Sigma-Aldrich
Chloroform	Sigma-Aldrich
D-(+)-Glucose	Sigma-Aldrich
D-(+)-Xylose	Sigma-Aldrich
EnPump 200	EnPresso GmbH
G418 solution	Sigma-Aldrich
Helioguard 365	Mibelle Biochemistry
Lithium acetate	Sigma-Aldrich
Milli-Q ultrapure water	Sigma-Aldrich
Nourseothricine sulfate	Sigma-Aldrich
Polyethylene glycol (PEG)	Sigma-Aldrich
Salmon Sperm Deoxyribonucleic acid (SS-DNA)	Sigma-Aldrich
Yeast Nitrogen Base with Amino Acids	Sigma-Aldrich

3.2 COMMERCIAL KITS, ENZYMES, AND BUFFERS

Table 2. Commercial kits, enzymes and buffers.

Name	Manufacturer
10X CutSmart buffer	New England BioLabs
2x OneTaq Master Mix	New England BioLabs
FastDigest 10X buffer	Thermo Fisher Scientific
FastDigest NotI	Thermo Fisher Scientific
FastDigest SfaAI	Thermo Fisher Scientific
GeneRuler 1 kb DNA Ladder	Thermo Fisher Scientific
Nb.BsmI	New England BioLabs
NEB 3.1 buffer	New England BioLabs
NucleoSpin® Gel and PCR Clean-up kit	Macherey-Nagel
NucleoSpin® Plasmid kit	Macherey-Nagel
Phusion U DNA polymerase Master Mix	Thermo Fisher Scientific
USER enzyme	New England BioLabs

3.3 STRAINS USED IN THIS STUDY

For cloning procedures and plasmid propagation, *Escherichia coli* DH5-α competent cells were used. Four strains of *S. cerevisiae* were used as parental strains for the integration of genes from *A. variabilis*. One strain is the haploid CEN.PK113-7D (P. Kötter, Goethe Univerity Frankfurt, Germany). The second strain is its xylose-consuming derivative ST8153 CEN.PK TyXylC1 (I. Borodina, Novo Nordisk Foundation Center for Biosustainability). The other two strains are the diploid industrial Ethanol Red® (S.I. Lesaffre, Fermentis, France) strain and its xylose-consuming derivative ST8155 ER TyXylC1 (I. Borodina, Novo Nordisk Foundation Center for Biosustainability, Denmark). Ethanol Red® is a rapid fermenting, temperature

tolerant strain known for its high yields in industrial production. The diploidy of Ethanol Red® strain ensures that every inserted gene will be present in two copies compared to only one copy in CEN.PK113-7D strain. A flowchart of *S. cerevisiae* strain construction can be seen in Figure 10.

3.4 MEDIA AND GROWTH CONDITIONS

Media composition is shown below in Table 3. *E. coli* cells were grown at 37 °C and 30 °C in Luria-Bertani (LB) medium supplemented with 100 mg/l ampicillin when required. Super Optimal Broth with Catabolite Repression (SOC) liquid medium was used for the recovery of competent *E. coli* after heat shock at 42 °C and transformation.

S. cerevisiae cells were grown at 30 °C in Yeast Extract Peptone Dextrose (YPD) medium, supplemented with 100 mg/l nourseothricin (cloNAT) and 200 mg/l G418 when required for transformant selection. For HPLC measurements, *S. cerevisiae* cells were grown in Synthetic Complete (SC) medium in three different glucose to xylose ratios (20:0, 10:10, 2:18), with the total dextrose concentration being 20 g/L. For LC-MS measurements, *S. cerevisiae* cells were grown in Delft minimal medium containing 7.5 g/l $(NH_4)_2SO_4$, 14.4 g/l KH_2PO_4, 0.5 g/l $MgSO_4 \cdot 7H_2O$, 2 ml/l trace metal solution and 1 ml/l vitamin solution (Verduyn et al., 1990). For pre-culture in Delft minimal medium, 20 g/L glucose was added. For controlled fed-batch fermentation in Delft minimal medium, EnPump 200 was added. EnPump 200 consists of a powdered polysaccharide and an enzyme which releases glucose at a steady rate, therefore simulating fed-batch culture.

Liquid cultures were agitated at 250 rpm. When plates were used for growth, the medium was supplemented with 20 g/L agar.

Table 3. Media composition.

Components	LB	SOC	YPD	SC
Tryptone	10 g/L	20 g/L		
Peptone			20 g/L	
Yeast extract	5 g/L	5 g/L	10 g/L	
NaCl	10 g/L	10 mM		
$MgSO_4$		10 mM		
KCL		2.5 mM		
Dextrose		20 mM (glucose)	20 g/L (glucose)	20 g/L
Yeast Nitrogen Base with Amino Acids				6.7 g/L
Agar	± 20 g/L		± 20 g/L	

3.5 SEQUENCES

Protein sequences of four *A. variabilis* enzymes; Ava_3855 (accession number ABA23460.1), Ava_3856 (ABA23461.1), Ava_3857 (ABA23462.1) and Ava_3858 (ABA23463.1), were found in the National Center for Biotechnology information (NCBI) database. These sequences (see Appendix) were then codon-optimized for *S. cerevisiae* using GeneGenie (Swainston et al., 2014), an online codon-optimization tool. A stop codon was then added to the generated DNA sequences and the full-length biosynthetic coding sequences (CDS) were ordered from GeneArt Gene Synthesis service (Thermo Fisher Scientific, USA). Strong constitutive promoters *TEF1* and *PGK1* obtained on a plasmid template from Verena Siewers (Chalmers University of Technology, Sweden) were used to amplify a bidirectional double promoter, which is also a part of the EasyClone toolbox (Jessop-Fabre et al., 2016).

3.6 PRIMERS AND GENERATION OF PCR FRAGMENTS FOR USER CLONING

Primers for gene and promoter amplification were designed specifically for USER cloning (Fig. 5). Thus, all forward primers intended for cloning genes into Gene 1 position start with the sequence AGTGCAGGU AAAACA ATG… (AGTGCAGGU – USER linker; AAAACA – Kozak sequence; ATG – start codon) at their 5' end. All reverse primers for cloning into Gene 1 position start with the sequence CGTGCGAU TAA… (CGTGCGAU – USER linker; TAA – stop codon) at their 5' end. Likewise, for genes in Gene 2 position, forward primers start with ATCTGTCAU AAAACA ATG and reverse primers start with CACGCGAU TTA at their 5´ ends. After treatment with USER enzyme, these primers enable overhang complementarity and directional assembly of genes and promoters (Fig. 4 and Fig. 5). The Kozak sequence is added to the primers to initiate translation in *S. cerevisiae*.

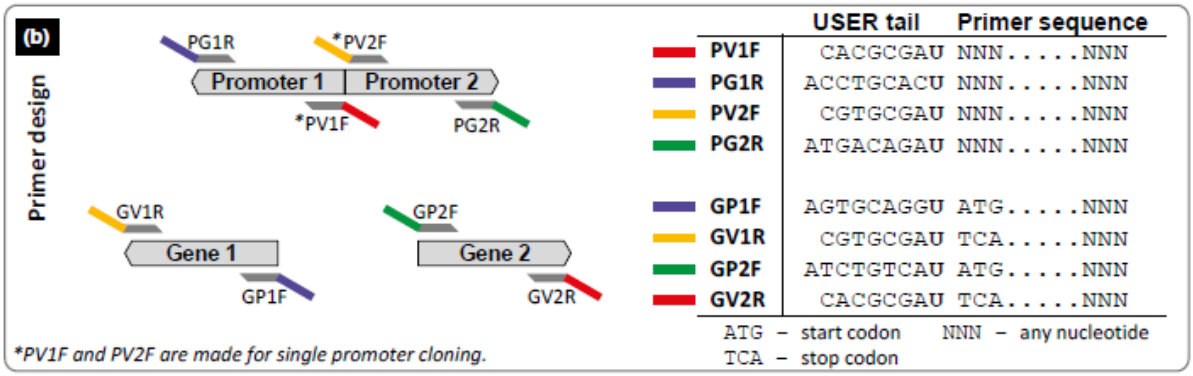

Figure 5. Primer design for USER cloning into EasyClone vectors. Figure taken from Jensen et al. (2014).

All primers were designed in CLC Main Workbench 8.1.2 (Qiagen, Netherlands) and were ordered from Eurofins Genomics (Germany). Working solutions were prepared in Milli-Q ultrapure water (Sigma-Aldrich, USA), diluting a 100 µM stock solution to 10 µM. Both stock and working solutions were stored at -20 °C. Primers used for gene and promoter amplification are listed in Table 4, while primers used for sequencing and colony PCR can be found in the Appendix (Supplementary Table 1).

Table 4. List of primers used for USER cloning. The sequences point from 5' to 3' end. Overhangs containing the USER linkers and the Kozak sequence are written in lowercase and sequences complementary to the gene/promoter targets are written in uppercase. The USER linker is marked red.

Name	Gene or promoter	USER position	Direction	Sequence 5' to 3' end	Fragment size
PR-24535 (Ava_3855_fw)	*Ava_3855*	Gene 1	FW	agtgcagguaaaacaATGCAAACTATCGATTTTAATATAAG	2667 bp
PR-24536 (Ava_3855_rev)	*Ava_3855*	Gene 1	REV	cgtgcgauTTAAGAATTGTTTTCCAGACAG	
PR-24537 (Ava_3856_fw)	*Ava_3856*	Gene 2	FW	atctgtcauaaaacaATGGCACAATCATTACCATTG	1400 bp
PR-24538 (Ava_3856_rev)	*Ava_3856*	Gene 2	REV	cacgcgauTTAATCACCACCTAGTTCAAC	
PR-24479 (Ava_3857_fw)	*Ava_3857*	Gene 1	FW	agtgcagguaaaacaATGACTAATGTTATTGTTCAACC	863 bp
PR-24580 (Ava_3857_rev)	*Ava_3857*	Gene 1	REV	cgtgcgauTTATGGTTGAATTCTACGAATAATG	
PR-24533 (Ava_3858_fw)	*Ava_3858*	Gene 2	FW	atctgtcauaaaacaATGTCTATTGTTCAAGC	1256 bp
PR-24534 (Ava_3858_rev)	*Ava_3858*	Gene 2	REV	cacgcgauTTATTTAACAGAACCAATC	
PTEF1_fw	*pTEF1*	Promoter 1	FW	acctgcacuTTGTAATTAAAACTTAGATTAGATTG	1430 bp
PPGK1_rv	*pPGK1*	Promoter 2	REV	atgacagauTTGTTTTATATTTGTTG	

All genes and promoters were amplified using Phusion U DNA polymerase Master Mix following the manufacturer's instructions. PCR was performed using the ordered biosynthetic genes and the double promoter containing plasmid as templates. The standard PCR program consisted of the following steps: 98 °C for 2 min; 32 cycles of 98 °C for 20 sec, 55 °C for 30 sec, and 72 °C for 30 sec/kb; and 72 °C for 5 min.

PCR amplicons were separated on a 1% (w/v) agarose gel and purified using NucleoSpin® Gel and PCR Clean-up kit. After purification, the concentration of the products was measured on a NanoDrop 8000 (Thermo Fisher Scientific, USA) spectrophotometer.

3.7 VECTORS

EasyClone integrative vectors (Mikkelsen et al., 2012) are designed to integrate at 14 locations on *S. cerevisiae* chromosomes X, XI and XII with no significant impact on growth rates. Each plasmid contains a USER insertion site flanked by a *CYC1* and an *ADH1* terminator sequence. The insertion site can be cut by SfaAI restriction enzyme to linearize the vector and by Nb.BsmI nicking enzyme to nick the linear ends, therefore generating 3' single stranded ends that are complementary to the 3' ends of the genes and promoters amplified with USER primers. The vectors also contain an ampicillin resistance gene, an *E. coli* origin of replication and two NotI restriction enzyme sites. After USER cloning, the vectors can be cut by NotI to generate a linearized fragment containing the inserted genes and promoters flanked by the terminators. This fragment is also flanked by regions homologous to *S. cerevisiae* chromosomes (designated as Up and Down regions in Figure 6 which can be used for CRISPR-Cas9 assisted integration).

A competent *S. cerevisiae* strain is transformed with the linearized integration fragment, an episomal plasmid expressing Cas9 and G418 resistance and with a vector expressing single guide RNA (gRNA) and nourseothricin resistance. The positive transformants are selected on plates containing G418 and nourseothricin. The gRNA binds to the Cas9 endonuclease and is complementary to the target sequence. Cas9 is then guided by the gRNA to the target sequence (one of the *S. cerevisiae* chromosomes). Cas9 introduces a double stranded break, which is then repaired with the introduced linear fragment by the native homologous recombination machinery of the host. The gRNA vector can later be removed by growing the cells in non-selective medium.

The vectors that were used in this project (Fig. 6) are pCfB2903 (XI-2-MarkerFree) and pCfB3035 (X-4-MarkerFree), provided by Dr. Irina Borodina (The Novo Nordisk Center for Biosustainability, Denmark) and which integrate to XI-2 (chromosomal coordinates 91575-92913) and X-4 (236336-237310) loci in *S. cerevisiae*, respectively.

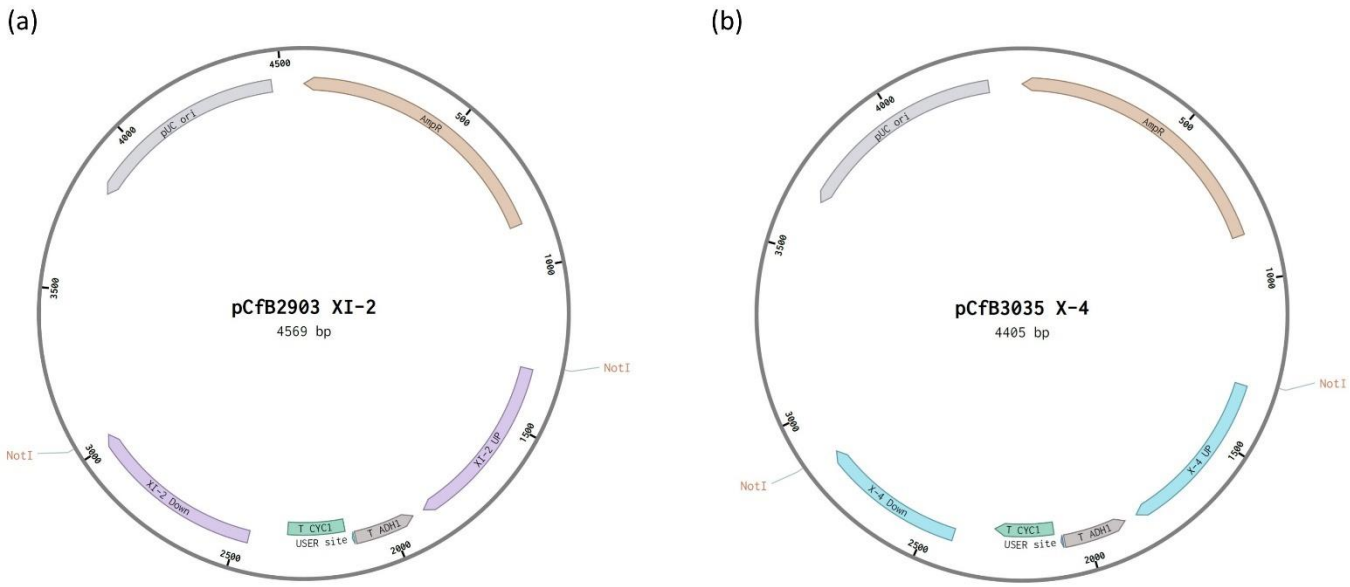

Figure 6. Schematic representations of EasyClone vector backbones. (a) plasmid pCfB2903 XI-2. (b) plasmid pCfB3035 X-4. USER site – cloning site recognized by SfaAI and Nb.BsmI; T ADH1 – *tADH1* terminator; T CYC 1 – *tCYC1* terminator; XI-2 Down and XI-2 Up – sequences homologous to *S. cerevisiae* chromosome XI; X-4 Down and X-4 Up – sequences homologous to *S. cerevisiae* chromosome X; NotI – restriction enzyme NotI site; AmpR – β-lactamase gene for ampicillin resistance; pUC ori – origin of replication for *E. coli*.

In addition to integrative vectors, episomal vector pCfB2312 (TEF1p-Cas9-CYC1t_kanMX) was used for Cas9 expression. Episomal vectors pCfB3044 (p-gRNA XI-2) and pCfB3042 (p-gRNA X-4) were used to express guide RNAs in *S. cerevisiae* strain CEN.PK113-7D and its derivatives, targeting chromosomal sites XI-2 and X-4, respectively. For *S. cerevisiae* strain Ethanol Red® and its derivatives, vector pCfB3589 (EthanolRed gRNA X-4) was used to express guide RNA targeting chromosomal site X-4.

3.7.1 VECTOR PREPARATION

The integrative vectors were first linearized with FastDigest SfaAI restriction endonuclease (20 µg of vector, 20 µl of FastDigest buffer, 5 µl of SfaAI, water to 200 µl, 1 h incubation at 30 °C), then purified from solution using NucleoSpin® Gel and PCR Clean-up kit, eluting with 40 µl of elution buffer and their concentration was measured on NanoDrop 8000. Linearized vectors were then nicked with Nb.BsmI enzyme (40 µl of SfaAI-digested vector, 5 µl of NEB 3.1 buffer, 5 µl of Nb.BsmI, 1 h incubation at 65 °C), separated on a 1% (w/v) agarose gel and purified using NucleoSpin® Gel and PCR Clean-up. The USER-ready vectors were then stored at -20 °C.

3.7.2 USER CLONING AND *E. coli* TRANSFORMATION

To assemble plasmids, 10 µl USER reactions were prepared as following: 1 µl 10X CutSmart buffer, 1 µl of USER enzyme in an insert to vector molar ratio of 3:1. The reactions were incubated under the following conditions: 37 °C for 25 min, 25 °C for 1 h, 20 °C for 10 min and 15 °C for 10 min.

After incubation, the USER reaction mix was used to transform *E. coli* DH5-α chemically competent cells using the following protocol: 10 µl of the USER reaction mix was added to 90 µl of thawed *E. coli* cells and the mixture was left for 10 min on ice. A heat shock was then performed by incubating the cells at 42 °C for 90 s. The cells were then incubated for 2 min on ice and 50 µl of SOC medium was added. The mixture was then left to incubate at 37 °C for 30 min. After incubation, the mixture was plated on selective LB plates supplemented with ampicillin and the cells were grown for 24 h at 30 °C. To quantify the number of false positives, *E. coli* DH5-α cells were also transformed with corresponding SfaAI/Nb.BsmI-treated empty vectors.

To assemble plasmid pCfB9117 (Fig. 8a), PCR amplified genes *Ava_3857* and *Ava_3858* and PCR amplified promoter *pTEF-pPGK* were used as inserts along with SfaAI/Nb.BsmI-treated vector pCfB2903. To assemble plasmid pCfB9118 (Fig. 8b), PCR amplified genes *Ava_3855* and *Ava_3856* and PCR amplified promoter *pTEF-pPGK* were used as inserts along with SfaAI/Nb.BsmI-treated vector pCfB3035.

To assemble plasmid pCfB9219 (Fig. 8c), previously assembled plasmids pCfB9117 and pCfB9118 were joined together by a PCR reaction followed by USER cloning. pCfB9117 and pCfB9118 were used as templates in a PCR reaction as described above with USER primers (Supplementary table 1) to generate a) a product containing gene *Ava_3857*, promoter *pTEF-pPGK*, gene *Ava_3858* and terminator *tCYC1* from plasmid pCfB9117 and b) a product containing terminator *tADH1*, gene *Ava_3855*, promoter *pTEF-pPGK* and gene *Ava_3856* from plasmid pCfB9118. As these PCR products were designed to have complementary USER overhangs, they were joined together with a USER reaction and cloned into SfaAI/Nb.BsmI-treated vector pCfB3035, as described above.

3.7.3 VERIFICATION OF PLASMID ASSEMBLY BY COLONY PCR AND SEQUENCING

To verify that the assembled plasmids contain the inserted sequences, colonies with transformed *E. coli* cells were screened by colony PCR. Colonies were touched with a pipette tip and to each colony biomass the following mix was added: 1 µl of 10 µM forward verification primer, 1 µl of 10 µM reverse verification primer, 5 µl of 2x OneTaq Master Mix and 3 µl of water. The PCR program consisted of the following steps: 94 °C for 2 min; 32 cycles of 94 °C for 20 sec, 50 °C for 30 sec, and 68 °C for 30 sec/kb; and 68 °C for 5 min. The PCR reactions were analyzed on 1% (w/v) agarose gels.

Colonies confirmed to have the inserts were used to inoculate 5 ml of LB broth supplemented with ampicillin and were grown overnight at 30 °C. 500 µl of each overnight culture was mixed with 50% glycerol and stored at -80 °C for later use. The rest of the overnight culture was used to purify the plasmids with NucleoSpin® Plasmid kit.

Plasmids that were confirmed to have the inserts were after purification additionally validated by Sanger sequencing (Eurofins Genomics, Germany) and analyzed with CLC Main Workbench 8.1.2 (Qiagen, Netherlands).

3.8 STRAIN CONSTRUCTION

The first enzyme in the shinorine biosynthesis pathway, 2-demethyl 4-deoxygadusol synthase (DDGS), uses sedoheptulose 7-phosphate (S7P) as a substrate. As S7P is naturally found in yeast as a pentose phosphate pathway intermediate, increasing the carbon flux toward the pentose phosphate pathway might be a way to increase the S7P pool. In turn, DDGS would have more of that substrate available to produce mycosporine-glycine and shinorine.

One way of increasing the flux towards the pentose phosphate pathway is xylose utilization. *S. cerevisiae* does not have the ability to utilize xylose naturally, but strains which heterogously express xylose utilization pathways have been designed. One of two ways to utilize xylose in *S. cerevisiae* is usually used: the first way is the xylose reductase/xylitol dehydrogenase (XR/XDH) pathway from fungi and the other one is the xylose isomerase (XI) pathway from bacteria (Qi et al., 2015). One reported problem with the XR/XDH pathway is cofactor imbalance: XR uses NADPH+ as a cofactor and XDH uses NAD+, thus avoiding cofactor regeneration and generating an imbalance (Li et al., 2016).

To avoid cofactor imbalance, strains expressing the xylose isomerase pathway (XI) were used in this study. These strains express three heterologous genes for xylose utilization: *Pichia stipitis SUT1*, encoding a xylose transporter; *Clostridium phytofermentans XylA*, encoding D-xylose isomerase (XI) and *Pichia stipitis XYL3*, encoding D-xylulokinase (XK). In addition, these strains overexpress two native *S. cerevisiae* genes (*RPE1* and *RKI1*) to direct the carbon flux towards ribulose 5-phosphate (R5P), another pentose phosphate pathway intermediate and a sedoheptulose 7-phosphate precursor.

Figure 7 shows the interconnected pathways involved in the production of MAAs used in this study: glycolysis, the pentose phosphate pathway, heterologous shinorine pathway and heterologous xylose consumption pathway (xylose isomerase pathway).

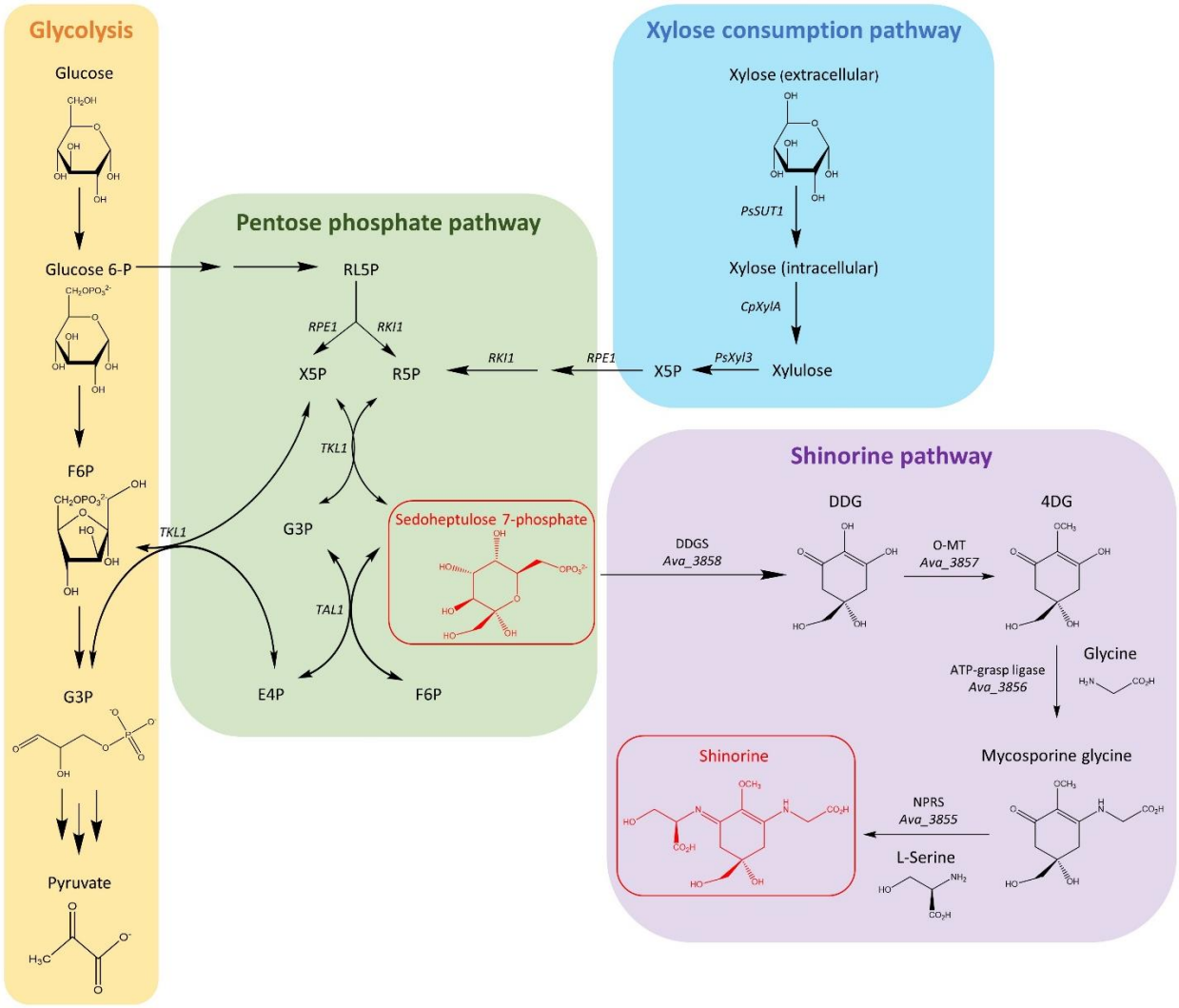

Figure 7. Schematic representation of the interconnected heterologous shinorine biosynthesis pathway from *A. variabilis* and xylose consumption pathway, along with glycolysis and the pentose phosphate pathway. Sedoheptulose 7-phosphate, a precursor of mycosporines and mycosporine-like amino acids is shown in red. Multiple arrows represent multiple enzymatic reactions grouped for simplicity. Glucose 6-P; glucose 6-phosphate, F6P; fructose 6-phosphate, G3P; glyceraldehyde 3-phosphate, RL5P; ribulose 5-phosphate, *RPE1*; D-ribulose-5-phosphate-3-epimerase, *RKI1*; ribose-5-phosphate ketol-isomerase, X5P; xylulose 5-phosphate, R5P; ribose 5-phosphate, *TKL1*; transketolase 1, *TAL1*; transaldolase 1, E4P; erythrose 4-phosphate, DDGS; DDG synthase, O-MT; O-methyl transferase, NPRS; nonribosomal peptides synthetase.

3.8.1 YEAST TRANSFORMATION AND VECTOR INTEGRATION

3.8.1.1 VECTOR LINEARIZATION

Before transformation into yeast, the constructed integrative vectors were linearized with NotI restriction endonuclease. NotI generates two fragments: the excised DNA fragment targeted for integration, flanked with regions homologous to a yeast chromosome to facilitate recombination, and the 2.8 kb backbone from the parental plasmid.

Glycerol stocks of the verified constructed integrative plasmids were used to inoculate 5 ml of LB broth supplemented with ampicillin and were grown overnight at 30 °C. The plasmids were then purified with NucleoSpin® Plasmid kit, eluting with 30 µl of Milli-Q ultrapure water. A 50 µl reaction was prepared with 0.2 µl FastDigest NotI per 1 µg of plasmid DNA, 5 µL of FastDigest buffer and Milli-Q water to 50 µl. The mixture was incubated at 37 °C for 1 hour. Linearized vectors were then separated from the released 2.8 kb backbones on a 1% (w/v) agarose gel and purified using NucleoSpin® Gel and PCR Clean-up kit. The concentrations of the linearized plasmids were measured with NanoDrop 8000. These steps were repeated until at least 4 µg of each linearized plasmid DNA was collected.

3.8.1.2 COMPETENT CELLS PREPARATION

The parental *S. cerevisiae* strains were inoculated into either 10 ml of YPD (if they do not contain a Cas9 expression vector) or into 10 ml of YPD supplemented with 200 mg/l of antibiotic G418 (if they were previously transformed with a Cas9 expression vector) and grown overnight at 30°C. The titer of the yeast culture was determined and 50 ml of fresh YPD medium was inoculated to give 5×10^6 cells/ml (OD_{600} ~ 0.6) of the culture. The culture was grown for approximately 4 hours until the titer was 2×10^7 cells/ml (OD_{600} ~ 4). Cells were harvested by centrifugation at 3000 × g for 10 min, washed in 25 ml of Milli-Q water, harvested again and resuspended in 1 ml of Milli-Q water. The cell suspension was transferred to a 1.5 ml microcentrifuge tube, spun down for 30 s and the supernatant was discarded. Water was added to a final volume of 1 ml and the cells were resuspended by vortex mixing. For each transformation, one microcentrifuge tube was prepared by pipetting 100 µl of the suspension, centrifuging at 3000 × g for 10 min and removing the supernatant.

3.8.1.3 TRANSFORMATION, INTEGRATION AND SELECTION

Competent *S. cerevisiae* cells were transformed using the Lithium acetate/single-stranded carrier DNA/Polyethylene glycol method (Gietz and Schiestl, 2007). For each transformation, 360 µl of the transformation mix was prepared. The transformation mix required for a single transformation contains 240 µl of 50% (w/v) polyethylene glycol (PEG), 36 µl of 1M lithium acetate, 10 µl of 10 mg/ml SS-carrier DNA (salmon sperm DNA) previously boiled for 3 minutes and kept on ice, 0.5 µg of gRNA helper vector (not added if transforming with Cas9 expression vector), 2-4 µg of linearized integrative vector or Cas9 expression vector and Milli-Q water to 360 µl. The mix was kept on ice until mixed with competent cells. For each transformation, 360 µl of the transformation mix was added on top of the previously prepared competent cell pellet in a microcentrifuge tube and cells were resuspended in the mix. A heat shock was then performed by incubating the tubes at 42 °C for 40 min. The cells were then spun down and the supernatant was removed. The cells were left to incubate in YPD for 2 hours at 30°C with shaking to provide time for expression of the antibiotic resistance genes.

Transformed cells were spun down, resuspended in 200 µl of Milli-Q water and plated on YPD agar plates supplemented with appropriate antibiotics depending on which plasmid was used for the transformation. When transforming with Cas9, plates were supplemented with 200 mg/l G418 as the Cas9 expression vector contains a G418 resistance marker. When transforming with linearized integrative vector and gRNA helper vector, plates were supplemented with 200 mg/l G418 to maintain the Cas9 vector and with 100 mg/l nourseothricin (cloNAT) as the gRNA helper vectors contain a nourseothricin resistance marker.

As the parental strains had to be transformed with the Cas9 expression vector first and as each integrative vector requires its own gRNA helper vector, this transformation and selection process was done in several rounds: Cas9 expression vector transformation, followed by integrative vector and gRNA vector transformation, followed by another round of integrative vector and gRNA vector transformation until the strains were transformed with the complete set of genes for the heterologous pathway and the corresponding gRNAs.

3.8.1.4 VERIFICATION OF CORRECT VECTOR INTEGRATION BY COLONY PCR

To verify that the integrative vectors had integrated into desired sites, a colony PCR is done with the colonies from the plate with transformed cells. Each colony is tested with three primers. The first two primers bind to a particular region on the *S. cerevisiae* chromosome. If there is no

integration in that region, the sequence between those two primers is amplified and the corresponding band can be seen in an agarose gel. The third primer is specific to the integrative vector. If the vector has integrated, the third primer in combination with one of the chromosome-specific primers amplifies the sequence between them and the corresponding band can be seen. In this case, the sequence between the two chromosome-specific primers cannot be amplified in a concentration visible on a gel because the integrated vector separates those chromosomal regions by five or more kilobases, which the DNA polymerase is not able to amplify at the conditions used for colony PCR. The primer combinations are described in Table 5.

Table 5. Primer combinations used for yeast colony PCR. Three primers are used at the same time in one PCR reaction. If the vector has integrated in the chromosomal site, chromosome-specific primer 1 and vector-specific primer are going to amplify the ‘’correct’’ fragment. If there is no integration, the two chromosome-specific primers are going to amplify the ‘’incorrect’’ fragment.

Chromosomal site		**Primer sequence 5' to 3' end**	**Fragment size**	
			Correct	**Incorrect**
X-4	**Chromosome-specific primer 1**	CTCACAAAGGGACGAATCCT	983 bp	1394 bp
	Chromosome-specific primer 2	GACGGTACGTTGACCAGAG		
	Vector-specific primer	GTTGACACTTCTAAATAAGCGAATTTC		
XI-2	**Chromosome-specific primer 1**	GTTTGTAGTTGGCGGTGGAG	963 bp	1573 bp
	Chromosome-specific primer 2	GAGACAAGATGGGGCAAGAC		
	Vector-specific primer	GTTGACACTTCTAAATAAGCGAATTTC		

Several colonies from each transformation plate were picked, resuspended in 5 µl of Milli-Q water and incubated for 10 min at 99 °C. Cells were then spun down and 2 µl of the supernatant was used as the template for PCR. For each colony the following PCR mix was used: 1 µl of 10 µM chromosome-specific primer 1, 1 µl of 10 µM chromosome-specific primer 2, 1 µl of 10 µM vector-specific primer, 5 µl of 2x OneTaq Master Mix and 2 µl of water. The PCR program consisted of the following steps: 94 °C for 1 min; 35 cycles of 94 °C for 20 sec, 50 °C for 30 sec, and 68 °C for 1.5 min; and 68 °C for 5 min. The PCR reactions were analyzed on 1% (w/v) agarose gels.

After confirmation of integration, the gRNA helper vector was removed. As the cells containing the gRNA vector (with nourseothricin resistance) need nourseothricin to keep the selection pressure, the selected transformants were plated on a YPD agar plates supplemented only with 200 mg/L G418 (leaving out the nourseothricin) and incubating overnight at 30 °C. Colonies from those plates were the following day replica plated on a YPD agar plates supplemented with both 200 mg/L G418 and 100 mg/L nourseothricin and incubated overnight at 30 °C. 4. Colonies that have not grown on the replica plate are confirmed to have successfully lost the gRNA helper vector and were used for subsequent rounds of transformation.

3.9 DETECTION OF MYCOSPORINE-LIKE AMINO ACIDS

3.9.1 HIGH PERFORMANCE LIQUID CHROMATOGRAPHY (HPLC) ANALYSIS

Constructed strains with complete MAA pathway were grown in Synthetic Complete (SC) medium in three different glucose to xylose ratios (20:0, 10:10, 2:18), with the total dextrose concentration being 20 g/L. Samples were measured after 24 h, 48 h and 96 h. Medium and cell extracts were analyzed separately. Medium was prepared by spinning down the cells and filtering the supernatant through a 0.22 μm pore size membrane filter. Cell extracts were prepared by centrifuging (3000 × g, 10 min) 4 ml of cell culture, resuspending the pellet in 1 ml of Milli-Q ultrapure water, adding 1.5 ml of chloroform and vortexing for 3 min. The water layer was then separated and filtered through a 0.22 μm pore size membrane filter. Shinorine was detected using the UltiMate 3000 HPLC system (Thermo Fisher Scientific, USA) with Zorbax Eclipse PLUS C18 (5 μm, 4.6 x 250 mm) column. A standard curve was generated using dilutions of Helioguard 365 (*Porphyra umbilicalis* extract with 0.1% MAAs shinorine and porphyra-334 corresponding to 1 g/l, Mibelle Biochemistry). The solvent and mobile phase comprised of acetonitrile in water (5:95 ratio) at a flow rate of 0.5 ml/min. The injection volume was 10 μl. Shinorine was detected by measuring absorbance with UV-vis detector at 334 nm. The column temperature was set at 40 °C.

3.9.2 LIQUID CHROMATOGRAPHY-MASS SPECTROMETRY (LC-MS) ANALYSIS

For LC-MS analysis, constructed strains were cultivated in micro-titer plates. Cells were inoculated from a 400 µl Delft minimal media pre-culture with glucose into a 400 µl Delft minimal media supplemented with EnPump 200 and incubated for 72 h with shaking. Metabolites for analysis were extracted as follows: cultures were supplemented with 100% acetonitrile at a ratio of 1:1, vortexed, then centrifuged at 3000 × g for 5 min. The resulting supernatant was analyzed using LC-MS with the following conditions: LC-MS measurements were done using a Dionex UltiMate 3000 Ultra-High Performance Liquid Chromatography (UHPLC) (Thermo Fisher Scientific, USA) system, connected to an Orbitrap Fusion Mass Spectrometer (Thermo Fisher Scientific, US). The UHPLC was equipped with a ZIC-HILIC column (3 µm, 15 cm x 2.1 mm). The temperature was set to 35 °C with the flow rate of 0.5 mL/min. The system was running an isocratic gradient with a mobile phase consisting of 20% 10 mM ammonium formate (pH 3) and 80% acetonitrile, with 0.1% formic acid. The samples were passed on to the MS equipped with a heated electrospray ionization source (HESI) in positive-ion mode.

4 RESULTS

4.1 INTEGRATIVE PLASMID CONSTRUCTION

In total, three expression cassettes in three plasmids were constructed. The first cassette contains gene *Ava_3857*, a bidirectional strong constitutive promoter *pTEF-pPGK* and gene *Ava_3857* and was cloned into vector pCfB2903 (XI-2-MarkerFree), therefore generating plasmid pCfB9117 (XI-2 Ava_3857<-pTEF-pPGK->Ava_3858) which integrates into *S. cerevisiae* chromosomal site XI-2 (Fig. 8a). The second cassette contains gene *Ava_3855*, a bidirectional promoter *pTEF-pPGK* and gene *Ava_3856* and was cloned into vector pCfB3035 (X-4-MarkerFree), therefore generating plasmid pCfB9118 (X-4 Ava_3855<-pTEF-pPGK->Ava_3856) which integrates into site X-4 (Fig 8b).

Due to the sequence variations between *S. cerevisiae* strains CEN.PK113-7D and Ethanol Red®, the gRNA used for integrating plasmid pCfB9117 into site XI-2 is not able to recognize the target site in the Ethanol Red® strain chromosome, so a third cassette was constructed specifically for the Ethanol Red® strain. This cassette was generated by PCR amplifying cassettes from the other two plasmids, pCfB9117 and pCfB9118, along with their terminator sequences, using USER primers with complementary overhangs. The resulting two PCR products were then joined at their 3' complementary ends and cloned into vector pCfB3035 by USER cloning. The cassette therefore contains all of the biosynthetic genes (*Ava_3855*, promoter *pTEF-pPGK*, *Ava_3856*, terminator *tCYC1* followed by terminator *tADH1*, *Ava_3857*, promoter *pTEF-pPGK*, *Ava_3858*), all cloned into one plasmid, pCfB9219, which integrates into site X-4 (Fig. 8c).

In addition to being able to integrate into *S. cerevisiae* genome, all constructed plasmids are capable of replicating in *E. coli* and they carry a selection marker for *E. coli*.

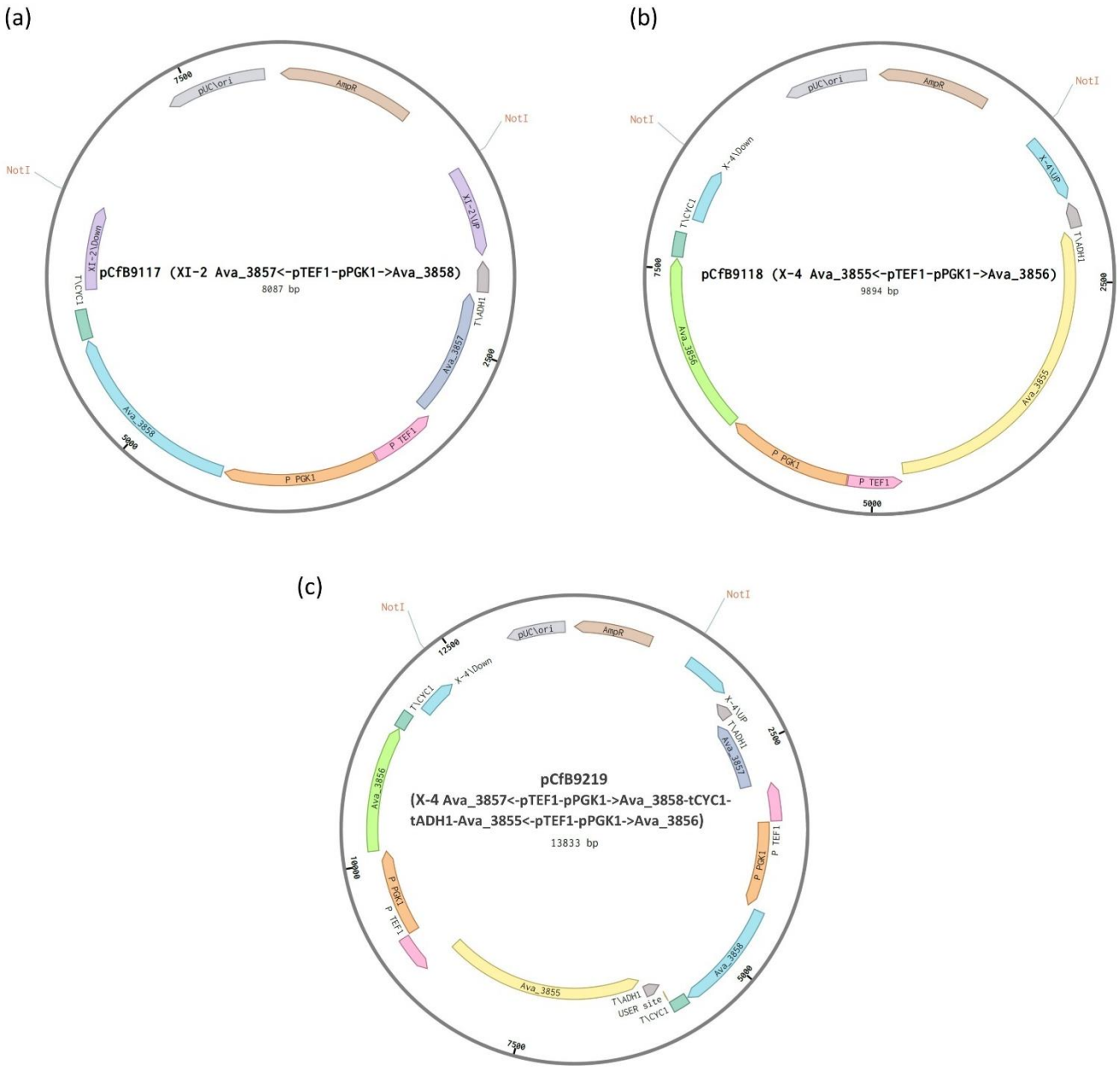

Figure 8. Schematic representation of constructed plasmids. (a) plasmid pCfB9117. (b) plasmid pCfB9118. (c) plasmid pCfB9219. Ava_3858 – *Ava_3858* gene for 2-demethyl 4-deoxygadusol synthase (DDGS); Ava_3857 – *Ava_3857* gene for O-methyltransferase (OMT); Ava_3856 – *Ava_3856* gene for ATP-grasp ligase; Ava_3855 – *Ava_3855* gene for nonribosomal peptide synthetase (NPRS)-like enzyme; P PGK1 – *pPGK1* promoter; P TEF1 – *pTEF1* promoter; T ADH1 – *tADH1* terminator; T CYC 1 – *tCYC1* terminator; XI-2 Down and XI-2 Up – sequences homologous to *S. cerevisiae* chromosome XI; X-4 Down and X-4 Up – sequences homologous to *S. cerevisiae* chromosome X; NotI – restriction enzyme site used for vector linearization; AmpR – β-lactamase gene for ampicillin resistance; pUC ori – origin of replication for *E. coli;* USER site – site used for joining pCfB9117 and pCfB9118 to construct pCfB9219.

4.1.1 VERIFICATION OF PLASMID ASSEMBLY

Correct assembly of the PCR amplified genes and promoters (Table 4) into EasyClone vector backbones (Fig. 6) by USER cloning was validated by *E. coli* colony PCR (Fig. 9 and 10) and Sanger sequencing. As the colony PCR was done using primers TEF1_test_rv/PGK_test-fw and ADH1_test_fw/CYC1_test_rv (details in Supplementary table 1), which bind to the end of the promotor and the beginning of the terminator, respectively, therefore amplifying the gene between them, only plasmids containing the inserted genes and inserted promoters were able to be amplified.

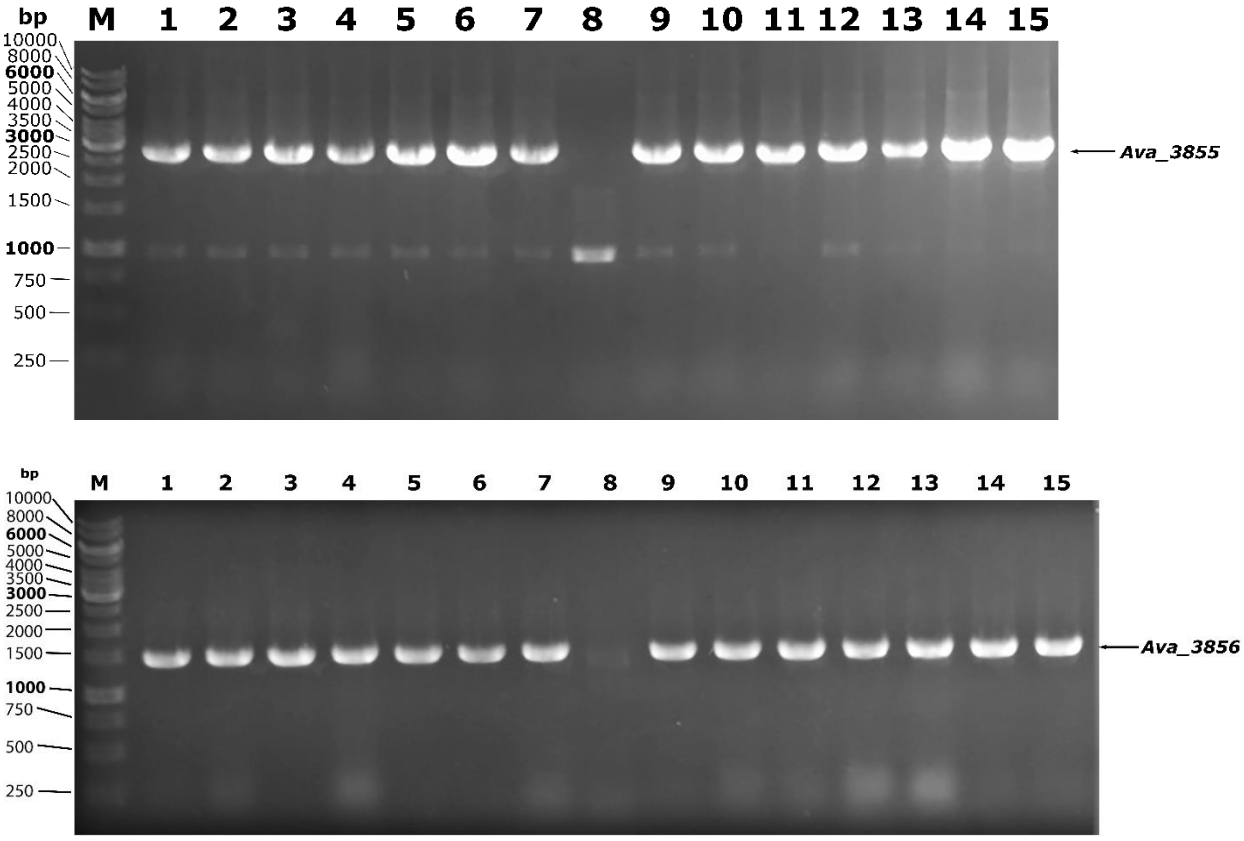

Figure 9. 1% agarose gels showing colonies transformed with plasmid pCfB9118 amplified by colony PCR. M - GeneRuler 1kb DNA ladder; 1-15 – bacterial colonies transformed with plasmid pCfB9118. Positive colonies show a band of around 2700 bp for gene *Ava_3855* and 1400 bp for gene *Ava_3856.* Colonies 1, 2, 14 and 15 were additionally validated by Sanger sequencing. Colony 8 does not show the 2700 bp band but does show a 1000 bp band present in all colonies, possibly representing sequences amplified from genomic DNA. Inkscape was used to adjust the contrast and exposition and to label the samples.

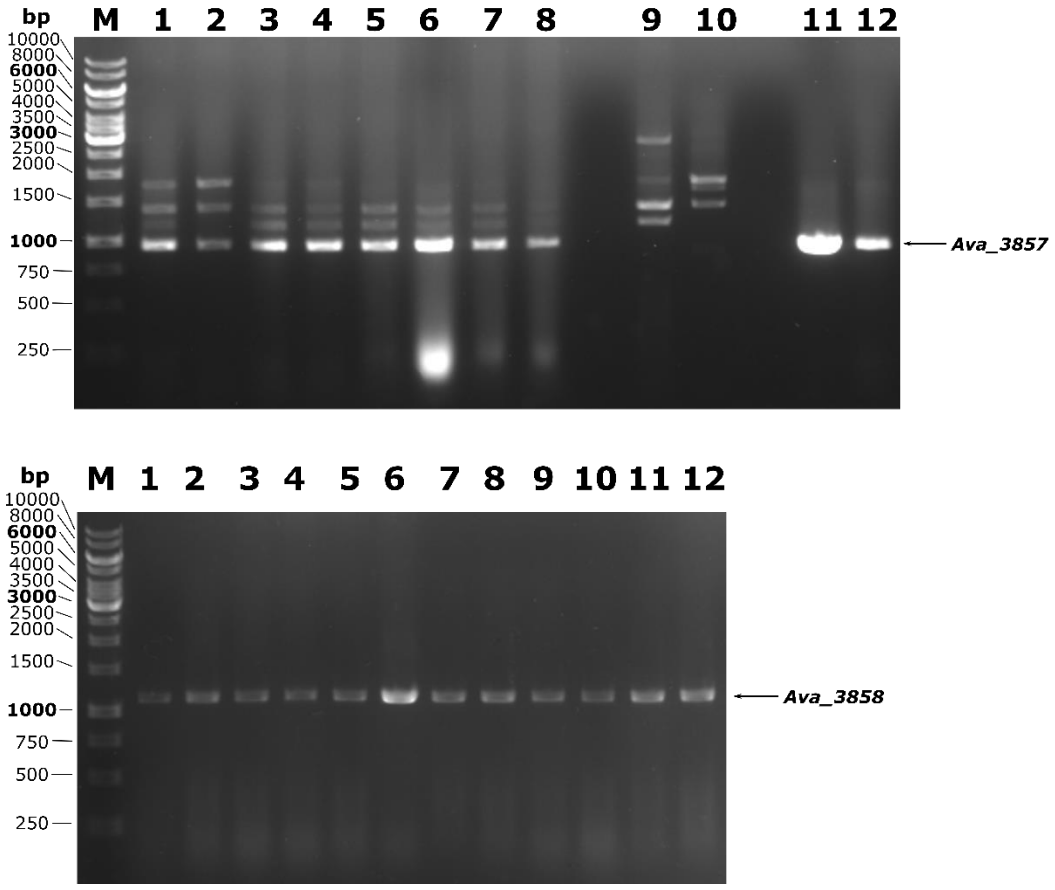

Figure 10. 1% agarose gels showing colonies transformed with plasmid pCfB9117 amplified by colony PCR. M - GeneRuler 1kb DNA ladder; 1-12 – bacterial colonies transformed with plasmid pCfB9117. Positive colonies show a band of around 900 bp for gene *Ava_3857* and 1250 bp for gene *Ava_3858.* After amplification for *Ava_3857*, all colonies except for 9 and 10 show the expected 900 bp band but only colonies 11 and 12 were additionally validated by sequencing because all other colonies show unexpected additional background bands. All colonies are positive for *Ava_3858.* Inkscape was used to adjust the contrast and exposition and to label the samples.

Colonies that were confirmed positive with colony PCR were additionally verified by Sanger sequencing. These were also used as PCR templates and their products were joined together via USER reaction to assemble plasmid pCfB9219. Those that did not have point mutations (plasmid pCfB9117 colony 2 and plasmid pCfB9118 colony 11) were linearized with NotI and used for yeast transformation.

4.2 STRAIN CONSTRUCTION

In total, four strains were constructed to express the MAA synthesis pathway. Two of those strains (ST9377 and ST9417) grow on glucose, while the other two (ST9376 and ST9418) are able to utilize xylose in addition to glucose via the xylose isomerase pathway. Strains ST9377 and ST9376 are haploid and are derived from CEN.PK113-7D parental strain, while strains ST9417 and ST9418 are diploid and are derived from Ethanol Red® parental strain. Ethanol Red® is a rapid fermenting, temperature tolerant strain known for its high yields in industrial production. In addition, the diploidy of Ethanol Red® strain ensures that every inserted gene will be present in two copies compared to only one copy in CEN.PK113-7D strain. The strain construction strategy is summarized in Figure 11.

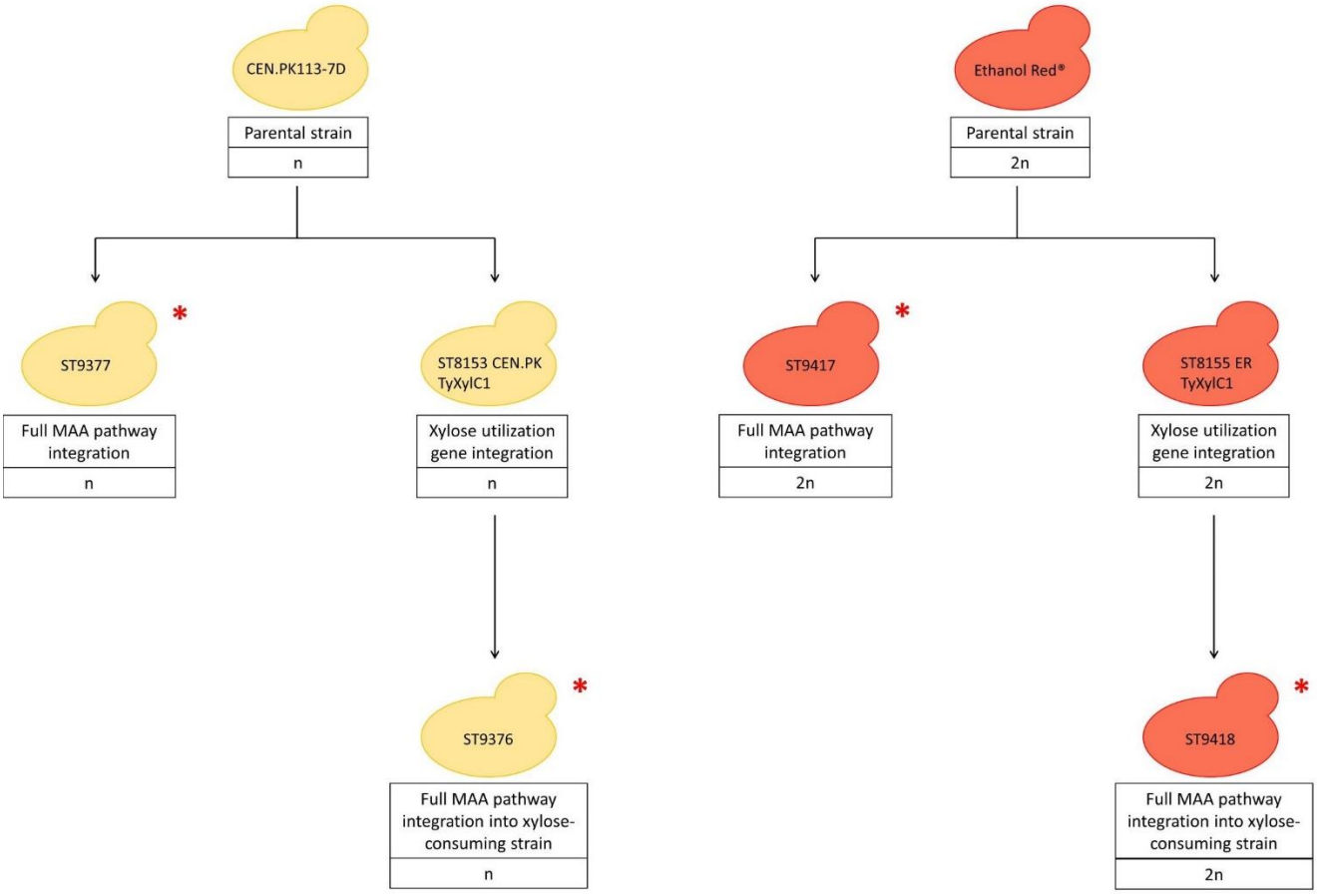

Figure 11. A flowchart of *S. cerevisiae* strains generated in this study. Four constructed strains (ST3377, ST9376, ST9417, ST9418) are marked with red asterisks (*). Haploid strains are marked with "n", while diploid strains are marked with "2n". As the transformation and integration process required multiple rounds, several intermediate strains with partial MAA pathway integration were constructed and are not shown in this flowchart.

4.2.1 VECTOR LINEARIZATION

Plasmids isolated from colonies that had the correct assembly of genes and promoters confirmed by PCR and sequencing were digested with NotI restriction enzyme. NotI digestion of integrative vectors (Fig. 12) generates two fragments, the first of which is always an approx. 3 kb vector backbone, while the other fragment is, depending on the plasmid:

- a 7.1 kb fragment from plasmid pCfB9118 which contains the X-4 Up homologous region, *tADH* terminator, *Ava_3855* gene, *pTEF1* and *pPGK1* promoters, *Ava_3856* gene, *tCYC* terminator and the X-4 Down homologous region
- a 5.3 kb fragment from pCfB9117 which contains the XI-2 Up homologous region, *tADH* terminator, *Ava_3857* gene, *pTEF1* and *pPGK1* promoters, *Ava_3858* gene, *tCYC* terminator and the XI-2 Down homologous region
- or an 11 kb fragment from pCfB9219 which contains the X-4 Up homologous region, *tADH* terminator, *Ava_3857* gene, *pTEF1* and *pPGK1* promoters, *Ava_3858* gene, *tCYC* terminator, *tADH* terminator, *Ava_3855* gene, *pTEF1* and *pPGK1* promoters, *Ava_3856* gene, *tCYC* terminator and the X-4 Down homologous region

A vector without an insert, i.e. the parental vector which was here used as a negative control, generates an 1.7-1.8 kb fragment (homologous sequences and terminators) along with the 3 kb backbone.

In addition to being necessary for yeast transformation, NotI digestion also serves as an additional verification of the correct plasmid assembly, as the size of the generated fragments can be inferred from sizes of the integrated genes and promoters plus the terminators and homologous sequences from parent vectors.

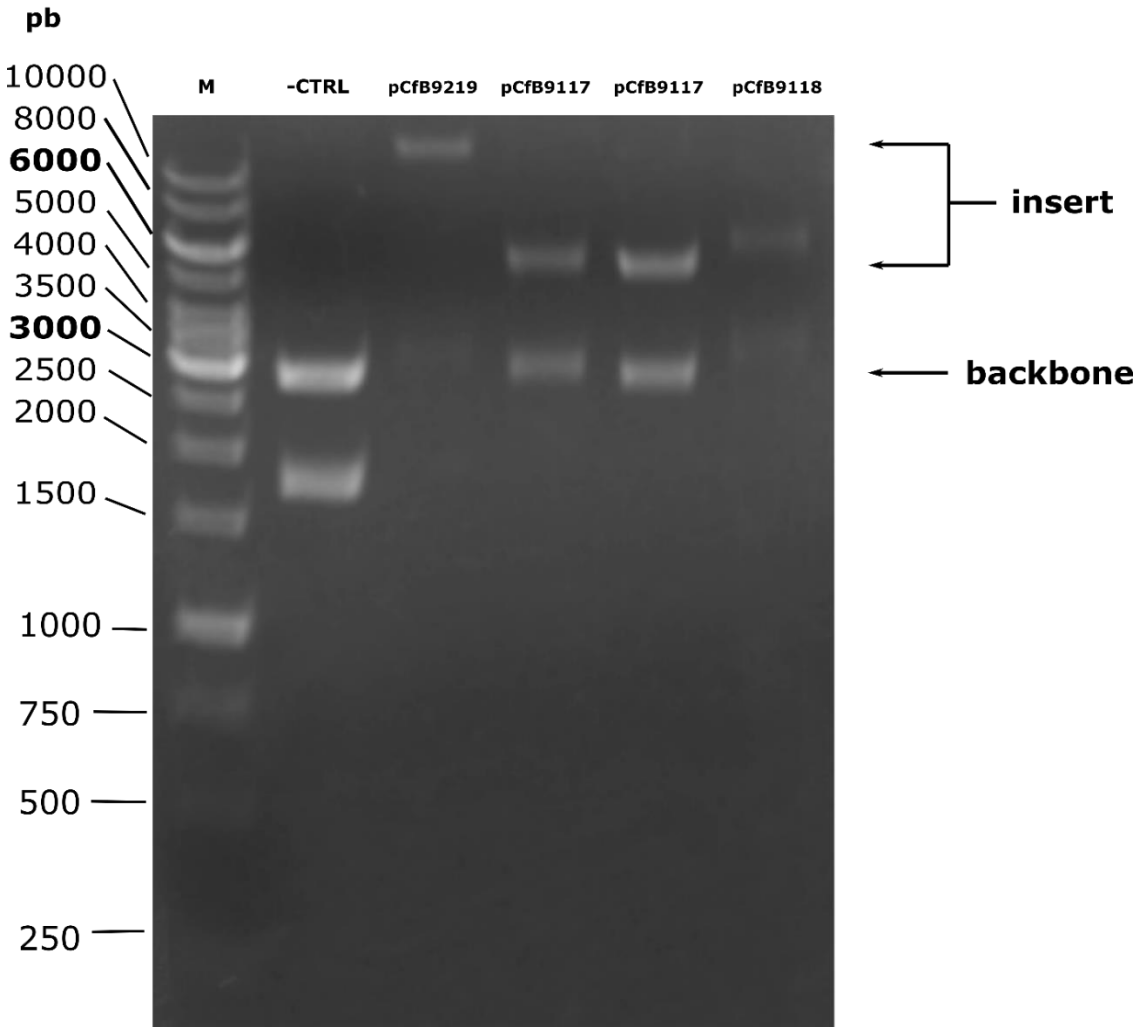

Figure 12. 1 % agarose gel showing constructed integrative vectors digested by NotI restriction enzyme. M - GeneRuler 1kb DNA ladder; -CTRL – negative control (parental vector pCfB2903 without an insert); pCfB9219, pCfB9117 and pCfB9118 – constructed integrative vectors. All plasmids show a band of ~ 3 kb representing the backbone. The negative control shows a 1.8 kb band representing the homologous sequences plus terminators without an insert. All plasmids were successfully digested as they show the bands corresponding to the total size of the inserts. Two samples of pCfB9117 come from two different *E. coli* colonies.

4.2.2 VERIFICATION OF CORRECT VECTOR INTEGRATION

S. cerevisiae strains (Fig. 11) transformed with the linearized integrative vectors were tested with colony PCR to verify that the vectors had integrated in the correct chromosomal sites.

Three primers (Table 5) were used at the same time for each PCR reaction. If the vector has integrated in the chromosomal site X-4, chromosome-specific primer 1 and vector-specific primer are going to amplify a 1 kb fragment. If there is no integration, the vector-specific primer

will not have a sequence to bind and the two chromosome-specific primers are going to amplify a 1.4 kb fragment. If the vector has integrated in the chromosomal site XI-2, chromosome-specific primer 1 and vector-specific primer are going to amplify a 1 kb fragment. If there is no integration, the two chromosome-specific primers are going to amplify a 1.6 kb fragment. Figure 13 shows that the integrative vectors have been successfully integrated into all parental strains.

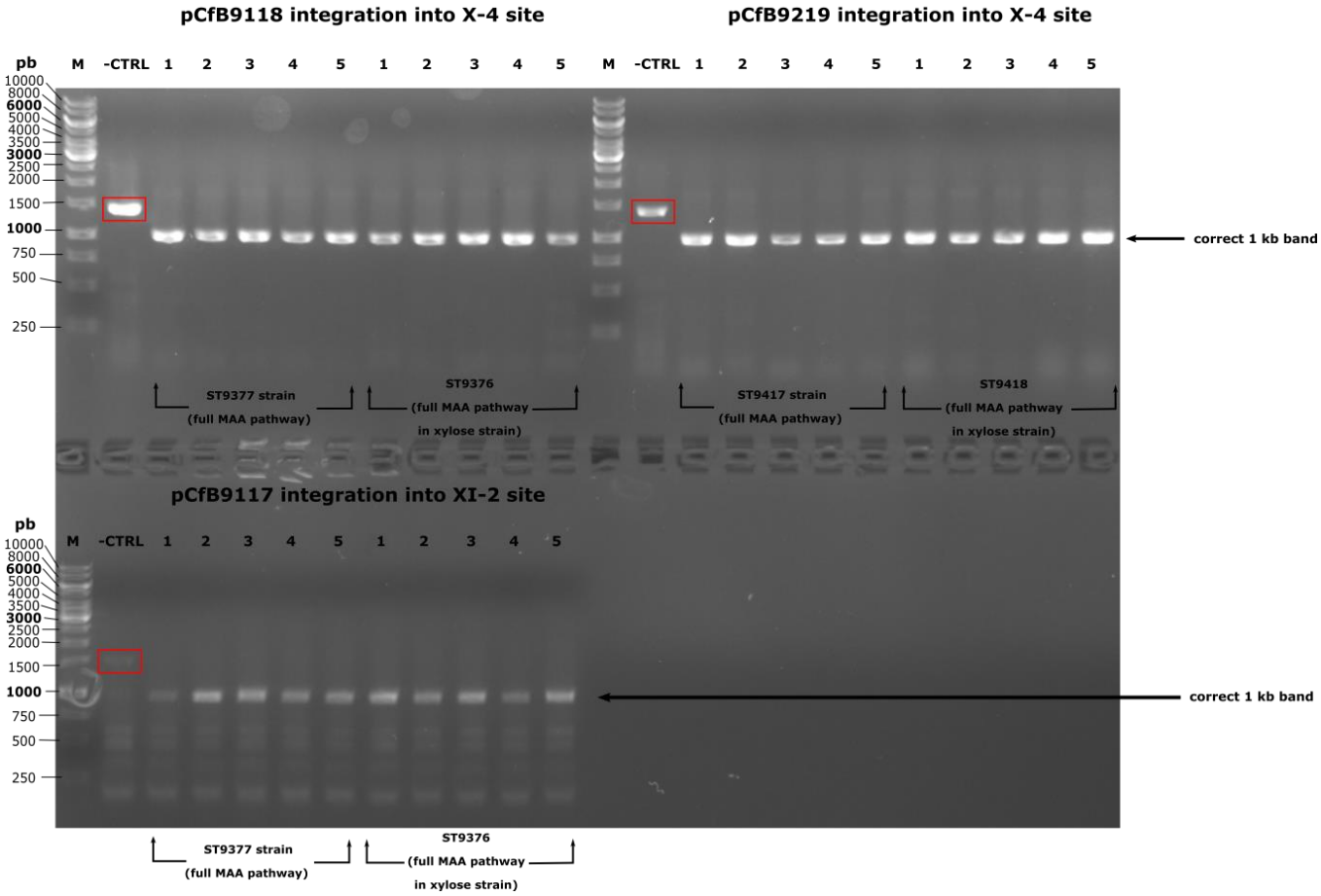

Figure 13. 1 % agarose gel showing constructed *S. cerevisiae* strains verified for integration by colony PCR. M - GeneRuler 1kb DNA ladder; -CTRL – negative control (parental CEN.PK113-7D or Ethanol Red® strain without integration); 1-5 – yeast colonies that were tested. All tested colonies show a 1 kb band amplified by a chromosome-specific and vector-specific primer, signifying correct vector integration. Negative control (parental strains) shows a 1.4 kb band (red squares in the upper gel) for the X-4 chromosomal site and a 1.6 kb band (red square in the lower gel), signifying that there was no integration.

4.3 MYCOSPORINE-LIKE AMINO ACIDS PRODUCTION

The lowest concentration of the shinorine standard that was able to be detected with HPLC was 2.5 mg/l. Compared to the standard, HPLC did not detect peaks corresponding to shinorine

(which absorbs UV at 334 nm) in any of the samples in three different glucose to xylose ratios (20:0, 10:10, 2:18) after 24, 48 or 96 hours. Compared to the parental strains, the constructed strains did not show any additional peaks. An example of HPLC results is shown in Figure 14. As the standard did not contain mycosporine-glycine, which is the other MAA whose production was possible with the inserted genes and is also a precursor of shinorine, HPLC measurement of potentially produced mycosporine-glycine was not possible.

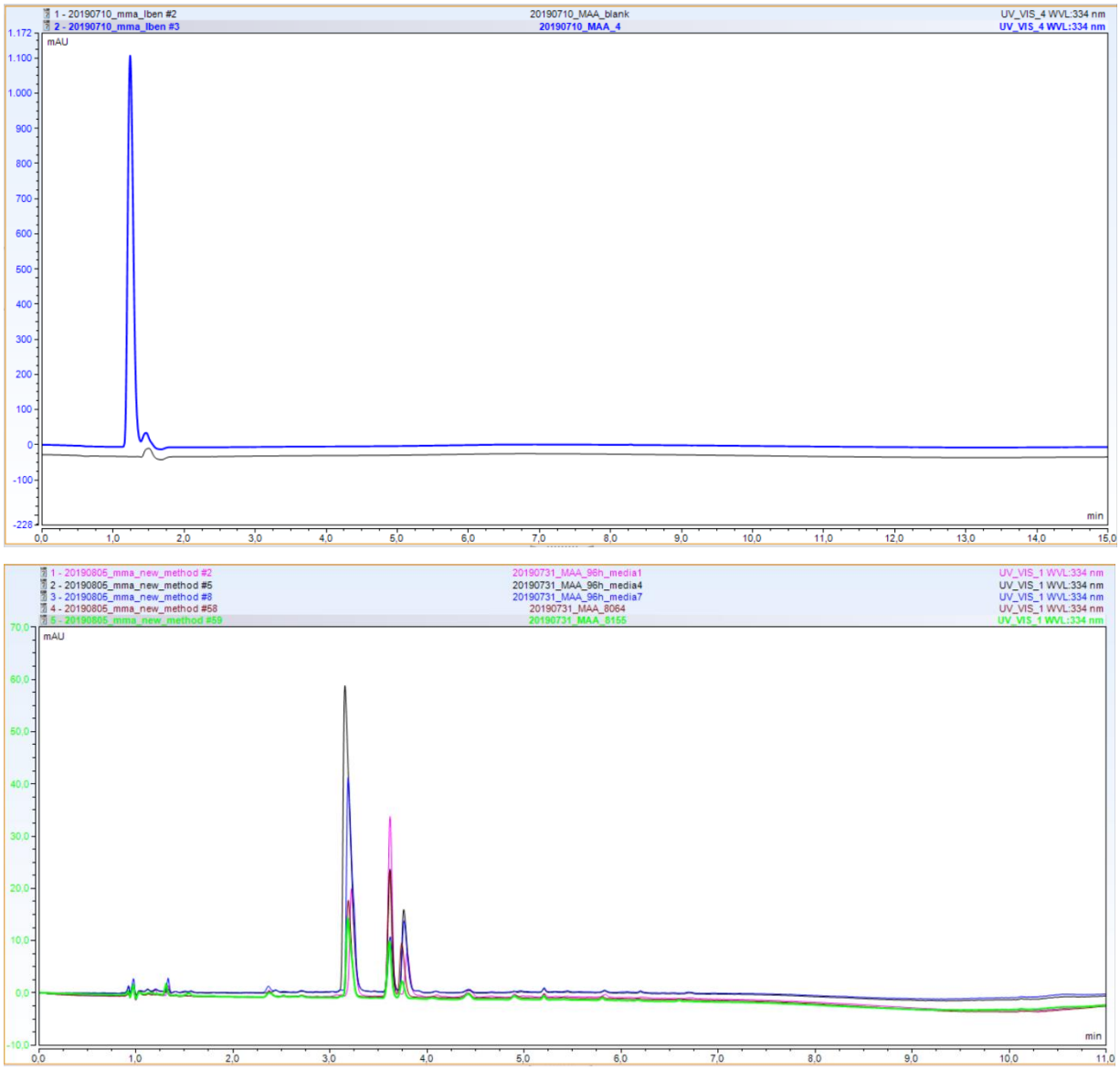

Figure 14. An example of HPLC analysis. Upper graph shows the 334 nm shinorine peak in the 25 mg/l standard (blue). Lower graph is an example of HPLC measurements of strains after 96 h of growth in 20 g/l glucose (pink), 10:10 g/l glucose to xylose (black), 2:18 glucose to xylose (blue) compared to their parental strains grown in 20 g/l glucose (brown) and in 10:10 g/l glucose to xylose (green). The constructed strains do not show the 334 nm shinorine peak or any additional peaks compared to their parental strains.

LC-MS, which can detect compounds in lower concentrations compared to HPLC was not able to detect shinorine either (Fig. 15). However, LC-MS measurements detected production of mycosporine-glycine in all constructed strains. Data from papers which characterized mycosporine-glycine and described its peak patterns were used as reference (Balskus and Walsh, 2010; Wada et al., 2015). As the standard did not contain mycosporine-glycine, it was not possible to estimate the produced concentration. The presence of peaks (Fig. 16) corresponding to mycosporine-glycine, a precursor of shinorine, in the LC-MS measurements shows that the integrated MAA pathway is active in the constructed strains up until the last step which is the attachment of serine to mycosporine-glycine to form shinorine, catalyzed by a nonribosomal peptide synthetase (NPRS).

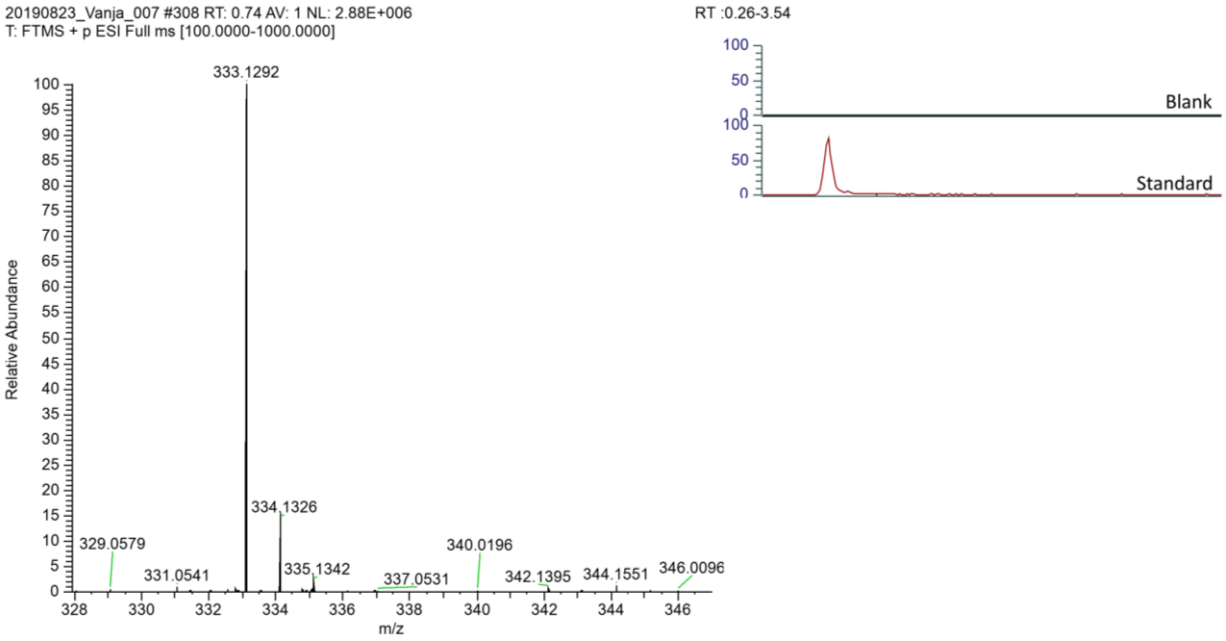

Figure 15. LC-MS analysis of shinorine. Left figure shows the mass spectrum of the analyzed standard – the peaks are amounts of ions produced in relation to the amount of the most abundant ion (the base peak, shinorine – 333.1292). Right figure is a comparison of the standard to the blank. Compared to the blank, a peak representing a compound with a mass of around 333 g/mol can be seen in the standard sample LC-MS analysis. This mass corresponds to the mass of shinorine which was not detected in any of the samples from the constructed strains.

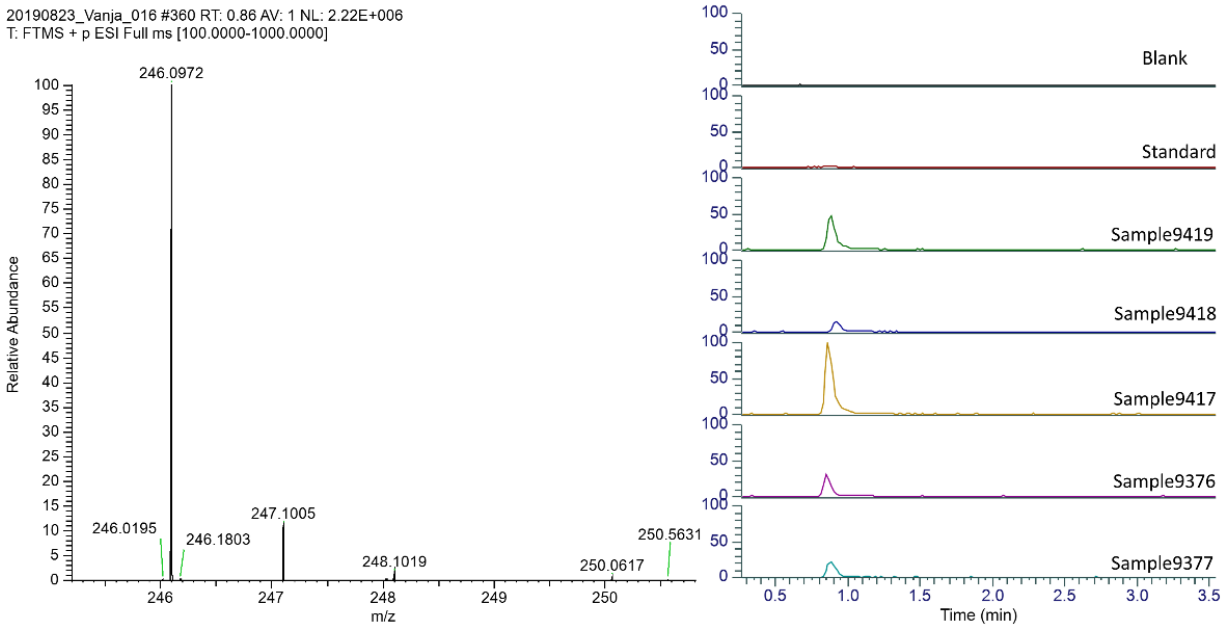

Figure 16. LC-MS analysis of mycosporine-glycine. Left figure shows the mass spectrum of sample 9417 – the peaks are amounts of ions produced in relation to the amount of the most abundant ion (the base peak, mycosporine-glycine – 246.0972). Right figure is a comparison of all samples to the standard and the blank. All samples from constructed strains (ST9419, ST9418, ST9417, ST9376, ST9377) show a 246 g/mol peak corresponding to the mass of mycosporine-glycine. If a standard of known concentration was available, the area under the peak curve could be used to calculate the concentration in each strain. However, the only available standard for this thesis was the shinorine standard (mycosporine-glycine concentration there is unknown and negligible as the standard is mostly comprised of shinorine which uses up mycosporine-glycine as a precursor). Without a proper mycosporine-glycine standard of known concentration, this analysis only serves as a qualitative measurement of mycosporine-glycine production.

5 DISCUSSION

Mycosporines and mycosporine-like amino acids (MAAs) are secondary metabolites produced by microorganisms exposed to high-intensity light. These compounds have photoprotective properties, enabling them to absorb UV-A and UV-B radiation, essentially serving as "microbial sunscreens". These properties make them promising candidates for pharmaceutical and cosmetic applications and some of them have already been commercialized as active ingredients in sunscreen creams. To date, several MAAs have been produced by introducing heterologous genes into a range of bacterial hosts. Expression of shinorine biosynthetic gene cluster from *A. variabilis* in *Escherichia coli* produced a low amount of shinorine of around 0.15 mg/l (Balskus and Walsh, 2010), using sedoheptulose 7-phosphate supplemented in the medium as the substrate. This thesis was intended to be the first report of MAA production in *S. cerevisiae*, but halfway through these experiments, a paper by Park et al. (2019) was published, describing the use of genes from *Nostoc punctiforme* ATCC 29133 to produce 31 mg/l of shinorine in *Saccharomyces cerevisiae*, using xylose as the substrate. Though unfortunate for this thesis, the Park et al. (2019) paper still served as a valid reference point for the remainder of this thesis. If that information had been available earlier, different approaches for this thesis might have been taken.

The aim of this thesis was to construct haploid and diploid *S. cerevisiae* strains capable of producing MAAs from glucose and xylose, by stably integrating and overexpressing heterologous genes from *Anabaena variabilis* ATCC 29413. This was done by successfully constructing and integrating two plasmids (pCfB9117 and pCfB9118), together containing the whole MAA biosynthesis pathway, into haploid CEN.PK113-7D strain and its xylose-consuming derivative. Plasmid pCfB9129, containing the whole MAA pathway, was also constructed and successfully integrated into diploid Ethanol Red® strain and its xylose-consuming derivative.

According to several studies about MAA synthesis, the steps for the production of mycosporine-glycine is conserved in many MAA-producing organisms, including *A. variabilis* ATCC 29413 (Balskus and Walsh, 2010) whose genes were used in this thesis, *Aphanothece halophytica* (Waditee-Sirisattha et al., 2014), *Actinosynnema mirum* DSM 43827 (Kiyoko T. Miyamoto et al., 2014) and in *Nostoc punctiforme* ATCC 29133, whose genes were successfully

used to produce shinorine in *S. cerevisiae* which was, up until this thesis, the only reported instance of MAA production in *S. cerevisiae* (Park et al., 2019).

Different MAAs are obtained by attaching different amino acids to a carbon atom in position 1 in mycosporine-glycine. Shinorine contains a serine residue at this position. Genes from *Nostoc punctiforme* ATCC 29133 have been used to produce shinorine in *Saccharomyces cerevisiae*, using xylose as the substrate. The pathways in *A. variabilis* and *N. punctiforme* differ in the last step: in *A. variabilis*, nonribosomal peptide synthetase (NPRS)-like enzyme attaches serine to MG, finally forming shinorine, while in *N. punctiforme*, serine is attached to MG via D-ala-D-ala ligase, also forming shinorine (Fig. 3).

The Park et al. (2019) paper reported higher MAA yields after expressing *N. punctiforme* genes in strains heterologously expressing the xylose reductase/xylitol dehydrogenase (XR/XDH) pathway from yeast *P. stipitis*, compared to the strains growing on glucose. They chose XR/XDH pathway over xylose isomerase pathway (XI) as using the XR/XDR pathway in *S. cerevisiae* has been shown to result in faster ethanol production (Karhumaa et al., 2007). However, the difference in cofactor preference between the mainly NADPH-dependent XR and the strictly NAD+-dependent XDH has been shown to cause cofactor imbalance, therefore limiting the flux from xylose to xylulose and causing excretion of xylitol, an intermediary in xylose to xylulose conversion in XR/XDH pathway (Kötter and Ciriacy, 1993). In this thesis, another approach was used: strains expressing the bacterial xylose isomerase pathway (XI) were used to avoid cofactor imbalance and xylitol accumulation, as in this pathway D-xylose is converted directly into utilizable D-xylulose without a redox reaction being involved (Fig. 7).

While the strains used in this thesis did grow on xylose and glucose in different ratios without any noticeable growth defects, production of shinorine was not detected. However, mycosporine-glycine was successfully produced in all constructed strains (Fig. 16). Without a proper mycosporine-glycine standard of known concentration, this analysis only serves as a qualitative measurement of mycosporine-glycine production. The detection of mycosporine-glycine suggests that only the last step in shinorine biosynthesis (attachment of serine to mycosporine-glycine to form shinorine) was not active. The gene responsible for the last step, *Ava_3855* from *A. variabilis* encoding a nonribosomal peptide synthetase (NPRS) was successfully integrated into *S. cerevisiae*. A potential reason for the nonactivity of this gene could be codon optimization, but as the rest of the genes from *A. variabilis* were active in *S.*

cerevisiae and the same codon optimization tool was used for all of them, this is unlikely. As the correct sequence of the gene was confirmed by sequencing, the integration was confirmed by colony PCR, and the same strong constitutive promoters were used for all the integrated genes, the problem might be that this enzyme is simply unstable or not active in a heterologous eukaryotic host. In any case, the expression levels of these integrated genes should be further investigated (e.g. by reverse transcription quantitative polymerase chain reaction (RT-qPCR) or RNA sequencing) to see if the enzyme is actively expressed and the enzyme should be purified and detected (e.g. by a Western blot) to see if it is present in the cytoplasm. If it is expressed and present in the cytoplasm, but there is no shinorine production, the conclusion would be that it is not active.

One possible solution for this problem is the use of another enzyme for the last step. Unlike in *A. variabilis* where this step is catalyzed by a a nonribosomal peptide synthetase, in *A. halophytica*, *A. mirum* and *N. punctiforme* this step is catalyzed by another enzyme, a D-Ala-D-Ala ligase (Gao and Garcia-Pichel, 2011; Miyamoto *et al.*, 2014; Waditee-Sirisattha *et al.*, 2014). Moreover, the D-Ala-D-Ala ligase homologue from *N. punctiforme* has already been successfully used to produce shinorine from mycosporine-glycine in *S. cerevisiae* (Park et al., 2019). D-Ala-D-Ala ligases are responsible for much of the diversity in MAAs, as they can attach different amino acids to mycosporine-glycine. In *A. mirum*, they attach alanine to mycosporine-glycine to form MAA mycosporine-glycine-alanine, while in *A. halophytica* they attach a second glycine to mycosporine-glycine to form mycosporine-2-glycine (Katoch et al., 2016b). Therefore, in addition to using a D-Ala-D-Ala ligase to produce shinorine, choosing a right D-Ala-D-Ala ligase could also enable the production of other MAAs.

Engineering a cell factory involves several rounds of the so-called design-build-test-learn (DBTL) cycle. In the DBTL cycle, a certain metabolic design is selected (e.g. the pathways and host); DNA assembly and genetic engineering methods are then used to build and improve the pathways in a host organism; the engineered organisms are then tested for target molecule production, transcripts, proteins, and metabolites; the knowledge gained by implementing a metabolic design is then used to analyze the dana and learn, therefore tying improved phenotypes to the designed genotypes and enabling iterations of the DBTL cycle (Liu et al., 2015; Nielsen and Keasling, 2016).

Only one round of the DBLT cycle was done in this experiment. In subsequent iterations of the cycle, several routes could be taken. One route would involve further engineering of the strains. As mentioned above, that could be using a gene coding for a D-Ala-D-Ala ligase (from *N. punctiforme* of from another organism, depending on which MAA is intended to be produced). A quick BLAST search of MAA cluster proteins towards Saccharomyces cerevisiae genome revealed that there is a DDG synthase (3-dehydroquinone synthase) homologue in *S. cerevisiae*, an enzyme called ARO1/YDR127W. In case this enzyme is competing with DDG synthase, it would be possible to do a knock-out of this gene, which is non-essential according to the Yeast Genome Database, and then test the strains for the production of MAAs. The Park et al. (2018) paper disrupted a different competing pathway by deleting the *TAL1* gene encoding the transaldolase involved in transaldolase reaction between sedoheptulose 7-phosphate (S7P) and glyceraldehyde 3-phosphate in the pentose phosphate pathway (PPP), therefore leaving more S7P available for shinorine biosynthesis, but also showing a severe growth defect in xylose medium because the deletion disrupted the PPP. This defect was then balanced by overexpressing the transcriptional factor Stb5 and transketolase Tkl1 from the PPP (Fig. 7). Another possible route would be fermentation optimization by, for example, optimizing the culture conditions by testing media with different (or different ratios of) carbon sources and supplements.

6 CONCLUSIONS

1) Three integrative vectors (pCfB9117, pCfB9118 and pCfB9219) containing codon-optimized genes for MAA shinorine biosynthesis (*Ava_3855*, *Ava_3856*, *Ava_3857* and *Ava_3858*) from cyanobacterium *Anabaena variabilis* ATCC 29413 under the control of constitutive promoters *TEF1* and *PGK1* were successfully constructed.
2) Constructed vectors were stably integrated into chromosomes of four *Saccharomyces cerevisiae* strains: haploid CEN.PK113-7D strain and its derivative expressing the xylose isomerase pathway and diploid Ethanol Red® strain and its derivative expressing the xylose isomerase pathway.
3) Constructed strains grew in media containing different ratios of glucose and xylose, but did not produce any detectable shinorine.
4) All constructed strains produced mycosporine-glycine, a precursor to shinorine, suggesting that the gene *Ava_3855*, responsible for the last step in shinorine biosynthesis (attachment of serine to mycosporine-glycine to form shinorine) was not active in *Saccharomyces cerevisiae.*

7 APPENDIX

7.1 SUPPLEMENTARY TABLES

Supplementary table 1. List of primers used for sequencing, *E. coli* colony PCR and integrative vector pCfB9129 assembly. Overhangs containing the USER linkers are written in lowercase and sequences complementary to the target sequence are written in uppercase. USER linker is marked red.

Name	Target sequence	Notes	Direction	Sequence 5' to 3'
PR-24672 (Seq_Ava_3855_TEF1PGK1_Ava_3856)_1_fw	*Ava_3855/3856*	Sequencing primer for plasmids containing *Ava_3855* and *Ava_3856*	FW	CCTGCGGTTGGATTGGTA
PR-24673 (Seq_Ava_3855_TEF1PGK1_Ava_3856)_2_fw	*Ava_3855/3856*	Sequencing primer for plasmids containing Ava_3855 and Ava_3856	FW	ATTTCACCAGTAACTCCG
PR-24674 (Seq_Ava_3855_TEF1PGK1_Ava_3856)_3_fw	*Ava_3855/3856*	Sequencing primer for plasmids containing Ava_3855 and Ava_3856	FW	GTTCTTGTGGGTATTCTG
PR-24675 (Seq_Ava_3855_TEF1PGK1_Ava_3856)_1_rev	*Ava_3855/3856*	Sequencing primer for plasmids containing Ava_3855 and Ava_3856	REV	GGAACCGATAAACCTAAAC
PR-24676 (Seq_Ava_3855_TEF1PGK1_Ava_3856)_2_rev	*Ava_3855/3856*	Sequencing primer for plasmids containing Ava_3855 and Ava_3856	REV	AAGAGTGTGTGAGAACAG
PR-24677 (Seq_Ava_3855_TEF1PGK1_Ava_3856)_3_rev	*Ava_3855/3856*	Sequencing primer for plasmids containing Ava_3855 and Ava_3856	REV	TTTCCCCTCTTTCTTCCT
PR-24678 (Seq_Ava_3857_PTEF1PGK1_Ava_3858_1_fw)	*Ava_3857/3858*	Sequencing primer for plasmids containing Ava_3857 and Ava_3858	FW	AATGGAGTGACGGGTCTA
PR-24679 (Seq_Ava_3857_PTEF1PGK1_Ava_3858_2_fw)	*Ava_3857/3858*	Sequencing primer for plasmids containing Ava_3857 and Ava_3858	FW	CGTGAGTAAGGAAAGAGT
PR-24680 (Seq_Ava_3857_PTEF1PGK1_Ava_3858_1_rev)	*Ava_3857/3858*	Sequencing primer for plasmids containing Ava_3857 and Ava_3858	REV	TTCTAACTTGGTCTGTCGG
PR-24681 (Seq_Ava_3857_PTEF1PGK1_Ava_3858_2_rev)	*Ava_3857/3858*	Sequencing primer for plasmids containing Ava_3857 and Ava_3858	REV	TCTACTAAACTAAACCACCCCC
ADH1_test_fw	*ADH1* terminator	Colony PCR primer - with TEF1_test_rv (ID339) can be used to test cloning into position "gene 1" between *ADH1* and *TEF1*	FW	GAAATTCGCTTATTTAGAAGTGTC
TEF1_test_rv (ID339)	*TEF1* promoter	Colony PCR primer - with ADH1_test_fw can be used to test cloning into position "gene 1" between *ADH1* and *TEF1*	REV	GCTCATTAGAAAGAAAGCATAGC
CYC1_test_rv	*CYC1* terminator	Colony PCR primer - with PGK1_test_fw can be used to test cloning into position "gene 2" between *CYC1* and *PGK1*	REV	CTCCTTCCTTTTCGGTTAGAG
PGK1_test_fw	*PGK1* promoter	Colony PCR primer - with CYC1_test_rv can be used to test cloning into position "gene 2" between *CYC1* and *PGK1*	FW	TACAGATCATCAAGGAAGTAATTATC
PR-24895 (tADH1 USER fw)	Constructed plasmid pCfB9118	For pCfB9219 assembly: amplifies tADH1 terminator and adds USER overhang compatible with ID402pintRevU. Used in combination with PR-24538 (Ava_3856_rev)	FW	atcgcguGAGCGACCTCATGCTATACC
ID402pintRevU	Constructed plasmid pCfB9117	For pCfB9219 assembly: amplifies tCYC1 terminator and adds USER overhang - compatible to PR-24895. Used in combination with PR-24580 (Ava_3857_rev) primer	FW	acgcgauCTTCGAGCGTCCCAAAACC

7.2 CODON-OPTIMIZED SYNTHETIC GENES FOR *Saccharomyces cerevisiae*:

Ava_3855 (nonribosomal peptides synthetase) from *Anabaena variabilis* (*Trichormus variabilis* ATCC 29413):

ATGCAAACTATCGATTTTAATATAAGAAAGTTATTAGTTGAATGGAACGCTACTC
ACAGAGATTATGATTTATCTCAATCCCTTCACGAATTAATTGTTGCACAAGTCGAA
AGGACTCCAGAAGCAATTGCTGTAACCTTTGACAAACAACAATTGACTTATCAAG
AATTGAATCATAAAGCCAATCAATTGGGTCATTATCTACAGACTTTGGGTGTTCAA
CCAGAAACACTAGTTGGTGTATGCTTAGAAAGATCTTTAGAGATGGTTATTTGCTT
ATTGGGTATTTTGAAAGCTGGTGGTGCCTACGTCCCAATAGATCCAGAATACCCA
CAAGAACGTATCGCTTATATGTTGGAAGATTCTCAAGTTAAGGTTTTGTTGACGCA
GGAAAAGTTATTAAATCAAATTCCACATCACCAAGCACAAACAATTTGTGTAGAT
AGAGAATGGGAAAAGATTTCTACTCAAGCTAATACAAACCCGAAATCTAACATTA

AGACCGATAATCTGGCTTATGTAATATACACATCAGGCTCAACCGGTAAACCTAA
AGGTGCTATGAATACTCATAAGGGTATATGTAATAGATTACTCTGGATGCAAGAA
GCTTATCAAATCGATAGCACGGATAGCATTTTACAAAAGACACCATTTTCTTTTGA
TGTCTCAGTTTGGGAATTCTTTTGGACCTTGTTGACGGGTGCTAGATTGGTTATTG
CAAAACCAGGCGGACATAAAGACTCTGCATATCTAATAGATTTGATTACGCAGGA
GCAGATTACCACATTACATTTTGTGCCATCTATGCTCCAAGTCTTCCTACAGAATA
GACATGTGAGCAAATGTTCATCTTTAAAACGAGTTATTTGTAGTGGTGAAGCTTTA
TCAATTGATTTACAAAATCGGTTTTTCCAACATTTGCAGTGCGAGTTACATAATTT
ATATGGTCCAACGGAAGCTGCAATCGATGTTACTTTCTGGCAATGCCGCAAGGAT
TCGAATCTTAAATCTGTTCCTATTGGTAGACCTATTGCTAATACACAAATTTATAT
TTTAGATGCGGATTTGCAACCTGTTAATATCGGAGTTACTGGTGAAATTTATATTG
GTGGTGTTGGTGTGGCAAGAGGTTACTTGAATAAGGAAGAATTAACCAAAGAAA
AATTCATTATTAATCCATTCCCAAACTCTGAATTTAAGAGATTGTACAAAACTGGT
GATTTGGCCAGGTATTTGCCAGATGGTAATATTGAATACTTAGGTAGAACGGATT
ACCAAGTGAAAATCAGAGGTTATAGAATTGAAATTGGCGAAATTGAGAACGTTTT
GTCTTCTCATCCACAAGTTAGAGAAGCCGTAGTGATAGCAAGAGATGATAACGCT
CAGGAAAAACAAATTATTGCTTACATTACTTACAATAGTATCAAGCCACAATTGG
ATAATTTGAGAGATTTTCTAAAAGCTCGTCTTCCAGATTTCATGATTCCAGCAGCA
TTTGTTATGTTGGAGCATTTACCTCTGACACCATCCGGCAAAGTCGATAGAAAAG
CATTGCCTAAACCAGATTTATTTAATTATTCTGAACATAATTCTTATGTGGCCCCA
CGTAATGAAGTCGAAGAAAAATTGGTTCAGATTTGGTCAAACATTCTTCATTTGCC
TAAGGTTGGTGTCACCGAAAATTTCTTTGCTATCGGCGGCAATTCGTTGAAGGCAT
TACATTTAATCTCTCAAATCGAAGAATTGTTCGCAAAAGAGATTTCTCTTGCCACG
TTATTGACTAATCCGGTTATCGCTGATCTAGCTAAGGTGATTCAAGCAAACAATCA
AATTCATAATTCTCCTTTGGTACCAATCCAACCGCAGGGTAAGCAACAACCATTCT
TCTGCATACATCCAGCTGGAGGCCATGTTTTGTGTTATTTTAAATTGGCTCAGTAC
ATCGGGACTGATCAGCCATTTTATGGTTTGCAAGCCCAGGGTTTTTATGGTGATGA
AGCACCATTGACGCGGGTCGAAGATATGGCATCATTATATGTAAAGACGATAAGA
GAGTTTCAACCACAAGGCCCTTATCGTGTTGGTGGTTGGAGTTTTGGCGGTGTGGT
TGCATATGAAGTCGCTCAACAATTGCATAGACAAGGCCAGGAAGTTTCATTGTTA
GCTATTCTAGATTCGTACGTTCCTATACTCTTGGACAAACAAAAACCAATCGATGA
CGTATATTTAGTGGGTGTACTTTCGAGGGTTTTTGGGGGAATGTTTGGTCAAGATA
ATTTAGTTACTCCTGAAGAAATAGAGAACTTGACGGTTGAGGAGAAAATTAATTA

TATTATTGACAAAGCTCGGTCAGCACGCATTTTTCCACCTGGTGTCGAACGACAA
AATAACAGGAGGATTTTGGATGTGTTAGTTGGTACATTGAAAGCAACTTATAGTT
ACATAAGGCAGCCATACCCGGGTAAAGTAACTGTTTTTCGTGCACGGGAAAAACA
CATTATGGCTCCAGATCCAACATTAGTCTGGGTCGAGTTGTTCTCTGTGATGGCTG
CCCAGGAGATTAAAATCATTGACGTGCCAGGTAATCATTATTCCTTTGTGTTGGAG
CCACACGTACAAGTTTTGGCACAGCGATTGCAAGACTGTCTGGAAAACAATTCTT
AA

Ava_3856 (ATP-grasp ligase) from *Anabaena variabilis* (*Trichormus variabilis* ATCC 29413):

ATGGCACAATCATTACCATTGTCTTCTGCTCCAGCTACTCCATCTCTGCCTTCCCA
AACCAAAATTGCTGCTATCATCCAAAATATCTGCACCTTAGCTTTATTGCTCTTAG
CTTTGCCGATTAATGCTACTATTGTTTTTATTTCTTTGTTGGTTTTCAGACCACAAA
AGGTTAAAGCCGCCAATCCACAAACTATTTTGATTTCTGGTGGTAAGATGACAAA
AGCTTTGCAATTGGCGAGATCTTTCCATGCTGCTGGTCATAGAGTGGTTCTAGTCG
AAACTCATAAATATTGGTTGACAGGTCATAGATTTAGCCAAGCCGTTGATAAATT
TTATACTGTTCCAGCCCCACAAGATAATCCACAAGCCTACATCCAGGCTTTGGTTG
ATATTGTTAAACAAGAAAACATCGATGTTTACATTCCAGTTACATCTCCAGTCGGC
TCCTATTACGATTCTTTGGCTAAACCAGAACTTTCCCATTATTGTGAAGTTTTTCAT
TTTGATGCTGATATAACTCAAATGTTAGATGATAAATTTGCCCTGACTCAAAAAGC
CAGAAGTTTAGGTTTATCGGTTCCTAAATCTTTCAAAATTACTTCTCCGGAACAAG
TGATTAATTTTGATTTTTCAGGTGAGACTAGGAAGTATATTTTAAAATCAATTCCA
TATGATTCCGTTAGAAGATTAGATTTGACTAAATTGCCATGTGCAACCCCAGAAG
AAACTGCTGCTTTTGTTAGGTCCCTGCCAATTACTCCAGAAAAACCATGGATCATG
CAGGAATTCATCCCAGGTAAAGAATTTTGTACTCATTCTACAGTCAGAAATGGTG
AATTGAGATTACATTGTTGTTGCGAATCTTCTGCTTTTCAGGTTAACTACGAGAAT
GTCAATAATCCACAAATTACTGAATGGGTCCAACATTTCGTGAAAGAGTTGAAGT
TGACTGGCCAAATTTCATTTGATTTTATTCAAGCTGAGGATGGTACTGTTTATGCT
ATTGAATGTAATCCACGGACTCATTCTGCTATTACTACATTTTATGACCACCCACA
AGTTGCTGAAGCTTACTTGTCACAAGCTCCAACAACTGAAACTATTCAACCATTG
ACTACTTCAAAACCCACCTACTGGACTTATCATGAAGTCTGGAGATTGACAGGTA
TCAGGTCTTTTACTCAATTGCAAAGATGGTTGGGTAATATCTGGAGAGGTACAGA
TGCTATTTATCAACCTGATGATCCATTGCCGTTTTTGATGGTCCATCATTGGCAAA

TTCCATTATTGCTGTTGAACAACTTGAGAAGATTAAAAGGCTGGACAAGAATCGA
TTTTAATATTGGTAAATTGGTTGAACTAGGTGGTGATTAA

Ava_3857 (O-methyl transferase) from *Anabaena variabilis* (*Trichormus variabilis* ATCC 29413):

ATGACTAATGTTATTGTTCAACCAACTGCTAGACCCGTCACTCCATTAGGTATTTT
AACTAAACAATTGGAAGCTATTGTTCAAGAAGTTAAACAACATCCCGATTTGCCA
GGTGAATTAATTGCTAATATTCATCAAGCTTGGCGCTTAGCTGCTGGTATTGATCC
ATATTTGGAAGAATGCACCACTCCAGAATCTCCAGAATTGGCTGCTCTGGCTAAA
ACTACCGCTACTGAAGCTTGGGGTGAACACTTTCACGGTGGTACTACTGTTAGAC
CATTGGAACAGGAAATGTTGTCTGGTCATATTGAAGGTCAAACTTTAAAAATGTT
TGTTCATATGACCAAGGCTAAAAAGGTTTTGGAAATTGGTATGTTCACTGGTTATT
CTGCTTTAGCTATGGCTGAAGCTTTACCAGAGGACGGTTTGTTGGTTGCTTGTGAA
GTTGATCCCTATGCCGCCGAAATTGGTCAAAAAGCTTTTCAACAATCTCCCCACGG
TGGCAAAATTAGAGTTGAATTGGATGCAGCTTTAGCTACTTTGGATAAGTTAGCT
GAAGCTGGCGAGTCTTTCGATTTAGTTTTTATTGATGCTGACAAAAAGGAGTATGT
TGCCTATTTTCATAAATTGTTGGGTTCTTCTTTGCTCGCTCCAGATGGTTTTATTTG
TGTCGATAATACTTTGTTGCAAGGTGAAGTTTATTTACCAGCTGAGGAAAGATCTG
TTAATGGTGAAGCTATCGCTCAGTTCAATCATACCGTTGCTATTGATCCAAGAGTC
GAACAGGTTTTATTACCATTGAGAGATGGGCTGACCATTATTCGTAGAATTCAAC
CATAA

Ava_3858 (demethyl 4-deoxygadusol synthase) from *Anabaena variabilis* (*Trichormus variabilis* ATCC 29413):

ATGTCTATTGTTCAAGCTAAATTTGAAGCTAAGGAAACTTCTTTTCATGTGGAAGG
TTATGAAAAGATCGAGTATGATCTGGTTTATGTTGATGGTATTTTTGAAATTCAGA
ATAGTGCTTTAGCCGATGTCTATCAAGGTTTTGGTAGATGTTTGGCCATTGTTGAC
GCTAATGTTTCTAGATTATATGGTAACCAAATTCAAGCCTATTTTCAATACTATGG
CATTGAGTTGAGATTATTTCCAATTACGATTACAGAACCAGATAAAACAATTCAA
ACCTTTGAAAGGGTTATTGATGTTTTTGCTGATTTTAAGCTGGTGAGGAAAGAACC
AGTTTTAGTTGTCGGTGGTGGTTTGATTACTGATGTTGTTGGTTTTGCTTGCTCCAC
TTATAGAAGATCTTCTAATTATATTAGAATCCCAACTACTTTGATAGGTTTGATTG
ACGCCTCTGTCGCTATTAAAGTTGCTGTTAATCATAGGAAATTGAAGAACCGTTTA

GGTGCTTATCATGCCTCAAGGAAGGTCTTTTTGGATTTTTCTTTGTTGAGAACCTT
GCCGACAGACCAAGTTAGAAATGGTATGGCTGAATTAGTTAAAATTGCTGTCGTT
GCCCATCAAGAAGTTTTTGAATTGCTAGAAAAATATGGTGAAGAGTTGCTGAGGA
CGCATTTTGGTAACATCGATGCTACTCCAGAAATTAAAGAAATTGCCCATAGATT
AACTTACAAAGCAATTCATAAAATGTTGGAATTAGAAGTTCCAAATTTGCATGAA
CTAGACTTAGATAGAGTTATTGCTTACGGTCATACTTGGTCTCCAACTTTGGAGCT
CGCTCCAAGATTGCCAATGTTTCACGGTCATGCCGTTAATGTTGATATGGCTTTTT
CTGCAACTATTGCTGCCAGAAGAGGTTATATTACTATTGCTGAAAGGGACAGGAT
TTTGGGTTTGATGTCTAGAGTCGGTCTGTCCTTGGATCATCCAATGTTGGATATTG
ACATCTTGTGGAGGGGTACTGAATCTATTACTTTAACCAGGGACGGGTTGTTAAG
AGCTGCTATGCCAAAACCGATCGGTGATTGTGTTTTTGTTAATGATTTGACTAGGG
AAGAGTTAGCCGCTGCTTTGGCTGATCATAAAGAACTGTGCACTTCTTACCCAAG
AGGTGGTGAAGGTGTTGACGTTTACCCAGTTTATCAAAAAGAATTGATTGGTTCT
GTTAAATAA

8 LITERATURE

Balskus, E.P. and Walsh, C.T. 2010. The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria. *Science*. **329**(5999), pp.1653–1656.

Bandaranayake, W.M. 1998. Mycosporines: Are they nature's sunscreens? *Natural Product Reports*. **15**(2), pp.159–172.

Berg, J., Tymoczko, J. and Stryer, L. 2002. the Pentose Phosphate Pathway Generates NADPH and Synthesizes Five-Carbon Sugars. *Biochemistry. 5th edition*.

Bitinaite, J., Rubino, M., Varma, K.H., Schildkraut, I., Vaisvila, R. and Vaiskunaite, R. 2007. USER™ friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Research*. **35**(6), pp.1992–2002.

Borodina, I. and Nielsen, J. 2014. Advances in metabolic engineering of yeast Saccharomyces cerevisiae for production of chemicals. *Biotechnology Journal*. **9**(5), pp.609–620.

Cardozo, K.H.M., Guaratini, T., Barros, M.P., Falcão, V.R., Tonon, A.P., Lopes, N.P., Campos, S., Torres, M.A., Souza, A.O., Colepicolo, P. and Pinto, E. 2007. Metabolites from algae with economical impact. *Comparative Biochemistry and Physiology - C Toxicology and Pharmacology*. **146**(1-2 SPEC. ISS.), pp.60–78.

Conde, F.R., Churio, M.S. and Previtali, C.M. 2004. The deactivation pathways of the excited-states of the mycosporine-like amino acids shinorine and porphyra-334 in aqueous solution. *Photochemical and Photobiological Sciences*. **3**(10), pp.960–967.

Duina, A.A., Miller, M.E. and Keeney, J.B. 2014. Budding yeast for budding geneticists: A primer on the Saccharomyces cerevisiae model system. *Genetics*. **197**(1), pp.33–48.

Favre-Bonvin, J., Bernillon, J., Salin, N. and Arpin, N. 1987. Biosynthesis of mycosporines: Mycosporine glutaminol in Trichothecium roseum. *Phytochemistry*. **26**(9), pp.2509–2514.

Gao, Q. and Garcia-Pichel, F. 2011. An ATP-Grasp Ligase Involved in the Last Biosynthetic Step of the Iminomycosporine Shinorine in Nostoc punctiforme ATCC 29133. *Journal of Bacteriology*. **193**(21), pp.5923–5928.

Gietz, R.D. and Schiestl, R.H. 2007. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*. **2**(1), pp.31–34.

Jensen, N.B., Strucko, T., Kildegaard, K.R., David, F., Maury, J., Mortensen, U.H., Forster, J., Nielsen, J. and Borodina, I. 2014. EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. *FEMS yeast research*. **14**(2), pp.238–48.

Jeppsson, M., Johansson, B., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.F. 2002. Reduced oxidative pentose phosphate pathway flux in recombinant xylose-utilizing Saccharomyces cerevisiae strains improves the ethanol yield from xylose. *Applied and Environmental Microbiology*. **68**(4), pp.1604–1609.

Jessop-Fabre, M.M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M.K., Keasling, J.D. and Borodina, I. 2016. EasyClone-MarkerFree: A vector toolkit for marker-less integration of genes into Saccharomyces cerevisiae via CRISPR-Cas9. *Biotechnology Journal*. **11**(8),

pp.1110–1117.

Karhumaa, K., Garcia Sanchez, R., Hahn-Hägerdal, B. and Gorwa-Grauslund, M.-F. 2007. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant Saccharomyces cerevisiae. *Microbial cell factories*. **6**, p.5.

Katoch, M., Mazmouz, R., Chau, R., Pearson, L.A., Pickford, R. and Neilan, B.A. 2016a. Heterologous production of cyanobacterial mycosporine-like amino acids mycosporine-ornithine and mycosporine-lysine in Escherichia coli. *Applied and Environmental Microbiology*. **82**(20), pp.6167–6173.

Katoch, M., Mazmouz, R., Chau, R., Pearson, L.A., Pickford, R. and Neilan, B.A. 2016b. Heterologous Production of Cyanobacterial Mycosporine-Like Amino Acids Mycosporine-Ornithine and Mycosporine-Lysine in Escherichia coli. *Applied and Environmental Microbiology*. **82**(20), p.6167.

Kavšček, M., Stražar, M., Curk, T., Natter, K. and Petrovič, U. 2015. Yeast as a cell factory: Current state and perspectives. *Microbial Cell Factories*. **14**(1).

Klisch, M., Richter, P., Puchta, R., Häder, D.P. and Bauer, W. 2007. The stereostructure of porphyra-334: An experimental and calculational NMR investigation. Evidence for an efficient 'proton sponge'. *Helvetica Chimica Acta*. **90**(3), pp.488–511.

Kötter, P. and Ciriacy, M. 1993. Xylose fermentation by Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology*. **38**(6), pp.776–783.

Li, X., Park, A., Estrela, R., Kim, S.-R., Jin, Y.-S. and Cate, J.H.D. 2016. Comparison of xylose fermentation by two high-performance engineered strains of Saccharomyces cerevisiae. *Biotechnology Reports*. **9**, pp.53–56.

Lian, J., Mishra, S. and Zhao, H. 2018. Recent advances in metabolic engineering of Saccharomyces cerevisiae: New tools and their applications. *Metabolic Engineering*. **50**, pp.85–108.

Liu, R., Bassalo, M.C., Zeitoun, R.I. and Gill, R.T. 2015. Genome scale engineering techniques for metabolic engineering. *Metabolic Engineering*. **32**, pp.143–154.

Microorganisms & microbial-derived ingredients used in food (partial list). *FDA (Aug. 21, 2013).*, https://www.fda.gov/Food/IngredientsPackagingLabel

Mikkelsen, M.D., Buron, L.D., Salomonsen, B., Olsen, C.E., Hansen, B.G., Mortensen, U.H. and Halkier, B.A. 2012. Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metabolic Engineering*. **14**(2), pp.104–111.

Miyamoto, Kiyoko T, Komatsu, M. and Ikeda, H. 2014. Discovery of gene cluster for mycosporine-like amino acid biosynthesis from Actinomycetales microorganisms and production of a novel mycosporine-like amino acid by heterologous expression. *Applied and environmental microbiology*. **80**(16), pp.5028–36.

Miyamoto, Kiyoko T., Komatsu, M. and Ikeda, H. 2014. Discovery of Gene Cluster for Mycosporine-Like Amino Acid Biosynthesis from Actinomycetales Microorganisms and

Production of a Novel Mycosporine-Like Amino Acid by Heterologous Expression H. Nojiri, ed. *Applied and Environmental Microbiology*. **80**(16), pp.5028–5036.

Nielsen, J. and Keasling, J.D. 2016. Engineering Cellular Metabolism. *Cell*. **164**(6), pp.1185–1197.

Nour-Eldin, H.H., Hansen, B.G., Nørholm, M.H.H., Jensen, J.K. and Halkier, B.A. 2006. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research*. **34**(18), p.e122.

Park, S.H., Lee, K., Jang, J.W. and Hahn, J.S. 2019. Metabolic Engineering of Saccharomyces cerevisiae for Production of Shinorine, a Sunscreen Material, from Xylose. *ACS Synthetic Biology*. **8**, pp.346–357.

Qi, X., Zha, J., Liu, G.G., Zhang, W., Li, B.Z. and Yuan, Y.J. 2015. Heterologous xylose isomerase pathway and evolutionary engineering improve xylose utilization in Saccharomyces cerevisiae. *Frontiers in Microbiology*. **6**(OCT).

Richardson, T.L. and Jackson, G.A. 2007. Small phytoplankton and carbon export from the surface ocean. *Science*. **315**(5813), pp.838–840.

Setlow, R.B. and Setlow, J.K. 1962. Evidence that ultraviolet-induced thymine dimers in DNA cause biological damage. *Proceedings of the National Academy of Sciences of the United States of America*. **48**(7), pp.1250–1257.

Simon, G., Kervarec, N. and Cérantola, S. 2014. NMR to Elucidate Structures *In*: *Outstanding Marine Molecules: Chemistry, Biology, Analysis*. Wiley Blackwell, pp.387–430.

Sinha, R.P., Singh, S.P. and Häder, D.P. 2007. Database on mycosporines and mycosporine-like amino acids (MAAs) in fungi, cyanobacteria, macroalgae, phytoplankton and animals. *Journal of Photochemistry and Photobiology B: Biology*. **89**(1), pp.29–35.

Swainston, N., Currin, A., Day, P.J. and Kell, D.B. 2014. GeneGenie: optimized oligomer design for directed evolution. *Nucleic acids research*. **42**(Web Server issue), pp.W395-400.

Tsuge, Y., Kawaguchi, H., Yamamoto, S., Nishigami, Y., Sota, M., Ogino, C. and Kondo, A. 2018. Metabolic engineering of *Corynebacterium glutamicum* for production of sunscreen shinorine. *Bioscience, Biotechnology, and Biochemistry*. **82**(7), pp.1252–1259.

Vaishampayan, A., Sinha, R.P. and Häder, D.P. 1998. Use of genetically improved nitrogen-fixing cyanobacteria in rice paddy fields: Prospects as a source material for engineering herbicide sensitivity and resistance in plants. *Botanica Acta*. **111**(3), pp.176–190.

Verduyn, C., Postma, E., Scheffers, W.A. and van Dijken, J.P. 1990. Physiology of Saccharomyces Cerevisiae in Anaerobic Glucose-Limited Chemostat Culturesx. *Journal of General Microbiology*. **136**(3), pp.395–403.

Wada, N., Sakamoto, T. and Matsugo, S. 2015. Mycosporine-Like Amino Acids and Their Derivatives as Natural Antioxidants. *Antioxidants*. **4**(3), pp.603–646.

Waditee-Sirisattha, R., Kageyama, H., Sopun, W., Tanaka, Y. and Takabe, T. 2014. Identification and upregulation of biosynthetic genes required for accumulation of

Mycosporine-2-glycine under salt stress conditions in the halotolerant cyanobacterium Aphanothece halophytica. *Applied and environmental microbiology*. **80**(5), pp.1763–9.

Walfridsson, M., Hallborn, J., Penttilä, M., Keränen, S. and Hahn-Hägerdal, B. 1995. Xylose-metabolizing Saccharomyces cerevisiae strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Applied and environmental microbiology*. **61**(12), pp.4184–90.

9 CURRICULUM VITAE

Personal information

Name:	Vanja Jurić
Contact:	vanjuric0@gmail.com

Education

2017 – 2019	Graduate program in Molecular Biology, University of Zagreb, Faculty of Science
01/2018 – 07/2018	Erasmus+ student mobility program, University of Natural Resources and Life Sciences, Vienna
2014 – 2017	Undergraduate program in Molecular Biology, University of Zagreb, Faculty of Science

Experience

03/2019 – 09/2019	Master thesis student at The Novo Nordisk Foundation Center for Biosustainability, Denmark, Yeast Metabolic Engineering group, supervised by Irina Borodina, PhD and Iben Møller-Hansen, PhD • Erasmus + traineeship
06/2017 – 09/2017	Internship at Paul Scherrer Institute, Switzerland, Laboratory of Biomolecular Research, supervised by Dr. Philipp Berger • International Association for The Exchange of Students for Technical Experience (IAESTE) internship
02/2017 – 06/2017	Internship at University of Zagreb, Croatia, Faculty of Science, Division of Molecular Biology, supervised by Associate Professor Ivana Ivančić Baće

Languages

Croatian:	Native proficiency
English:	Cambridge English: Advanced certificate, 2015, C1 level