Universidade de Aveiro Departamento de Química
2015

Liliana Patrícia dos Reis Teixeira

Produção de soforolípidos com cauda hidrofóbica ramificada

Production of sophorolipds with a branched hydrophobic tail

Universidade de Aveiro Departamento de Química
2015

Liliana Patrícia dos Reis Teixeira

Produção de soforolípidos com cauda hidrofóbica ramificada

Production of sophorolipds with a branched hydrophobic tail

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biotecnologia, realizada sob a orientação científica da Doutora Luísa Serafim, Professor Auxiliar do Departamento de Química da Universidade de Aveiro

Trabalho realizado sob orientação do Professor Doutor Wim Soetaert e Doutora Inge Van Bogaer na universidade de Ghent, ao abrigo do programa de mobilidade Eramus, para obtenção do grau de mestre em Biotecnologia na Universidade de Aveiro.

Jurí

Dr. ir. Inge Van Bogaert

Pós-doutoral com agregação na Universidade de Ghent

Dr. ir. Sofie De Maeseneire

Pós-doutoral com agregação na Universidade de Ghent

Ir. Brecht Vanlerberghe

R&D Manager na Bio Base Europe Pilot Plant

Agradecimentos

Working in InBio was one of the most profitable and enjoyable experiences of my life. I specially thank dr. ir. Inge Van Bogaert for accepting me in the Biosurf group, for all the help, advice, patience and guidance. Special thanks to Prof. dr. ir. Wim Soetaert for allowing me to work and improve myself professionally in InBio. Thank you Lien Saey for always been available and willing to answer my questions and help me in whatever I needed. Also thanks to Kwok Kuen Chow for the guidance. To the full InBio team, and especially to the Biosurf group, it was a pleasure to work with you, learn from your experience and to have made science together. Thanks to my coordinator in Portugal, Prof. dr. ir. Luísa Serafim, for the patience and help though all my Erasmus experience.

There is no one that I thank the most than my parents, António Teixeira and Idalina Teixeira. They have supported me mentally and monetarily throughout the process and they have never felt more proud of me. Without them, my Erasmus experience wouldn't have been possible neither the concretion of my master thesis. Whenever I felt frustrated or happy with my work, they were there to support me. I will never forget their effort and sacrifice. Thanks to all my family that kept in touch despite the distance. To my friends in Portugal, that have never forgotten me and were there for my high and low moments. To my friends in Gent, either the Portuguese or my lab coworkers, thank you for making my life in Belgium more enjoyable. Special thanks to Rita Almeida and José Loureiro, who have always been beside me.

Liliana Teixeira

Palavras-chave

Biosurfactantes, *Starmerella bombicola, Rhodotorula bogoriensis Elizabethkingia meningoseptica, Rhodococcus rhodochrous,* oleate hidratase, Guerbet alcoois.

Resumo

Soforolípidos são um tipo de biosurfactantes produzidos por vários microorganismos, incluindo a levedura *Starmerella bombicola*. As suas principais aplicações devem-se à sua sua atividade emulsificante e propriedades antimicrobianas.

Soforolípidos produzidos por *Starmerella bombicola* são compostos por uma molécula de soforose e um ácido gordo hidroxilado de 16 ou 18 átomos de carbono. Apesar das interessantes caraterísticas dos soforolípidos clássicos produzidos pela levedura referida, há indicações de que os soforolípidos com cauda hidrofóbica ramificada têm maior atividade a temperaturas mais elevadas. Estes soforolípidos são produzidos por *Rhodotorula bogoriensis*, no entanto o rendimento obtido por esta levedura é baixo. Ao produzir estes soforolípidos em *Starmerella bombicola*, o rendimento obtido poderá ser superior. A produção destes soforolípidos irá também permitir o alargamento das aplicações dos biosurfactantes.

Esta tese descreve duas medidas para a produção de soforolípidos com cauda hidrofóbica ramificada: uso de álcoois de Guerbet como substrato para a cultura de *Starmerella bombicola*, e a introdução de genes que serão responsáveis pela hidroxilação no meio da cadeia hidrofóbica dos soforolípidos. Álcoois de Guerbet são moléculas ramificadas, sendo esperado a sua incorporação nas moléculas de soforolípidos. Os genes usados são provenientes de *Elizabethkingia meningoseptica* e *Rhodococcus rhodochrous*, e codificam para uma oleate hidratase, enzima responsável pela conversão do ácido oleico no ácido 10-hidroxistearico.

Keywords

Biosurfactants, *Starmerella bombicola*, *Elizabethkingia meningoseptica, Rhodococcus rhodochrous,* oleate hydratase, Guerbet alcohols

Abstract

Sophorolipids are a type of biosurfactants produced by several microorganisms, including the yeast *Starmerella bombicola.* They find application mainly due to their emulsifying activity and antimicrobial properties.

Sophorolipids produced by *Starmerella bombicola* are composed by one sophorose molecule and a hydroxylated fatty acid of 16 or 18 carbon atoms. One or two acetylations can occur, one on each glucose, and a lactonization as well. Despite the interesting characteristics of classic sophorolipids produced by the referred yeast, there are indications that sophorolipids with branched a hydrophobic tail have higher activity at lower temperatures. These sophorolipids are produced by *Rhodotorula bogoriensis,* however the yields obtained with this yeast are low. By producing these sophorolipids in *Starmerella bombicola*, the yields obtained could be higher. Furthermore, the production of these novel sophorolipids will allow broadening of the biosurfactants applications.

This thesis describes two approaches to produce sophorolipids with a branched hydrophobic tail: use of Guerbet alcohols as substrates for the *Starmerella bombicola* culture, and the introduction of genes that will take care of the in-chain hydroxylation of the sophorolipids. Guerbet alcohols are branched molecules, and are expected to be incorporated in the sophorolipids molecules. The genes to be used are from *Elizabethkingia meningoseptica* and *Rhodococcus rhodochrous* and encode for an oleate hydratase, an enzyme responsible for the conversion of oleic acid into 10-hydroxystearic acid.

Table of contents

Abbreviation

A – Amperes
A – Adenine
adh – alcohol dehydrogenase
Ahd1 – alpha hydroxylase
Amp – ampicillin
at - acetyltransferase
BLAST – Basic Local Alignment Search Tool
bp – base pairs
ºC – Celsius
C – Carbon
C – Cytosine
Ca – Calcium
Cl – Chlorine
CMC – critical micelle concentration
CSM – complete supplement mixture
CFU – colony forming units
dTDP – deoxythymidine diphosphate
DMF – dimethyl formamide
DMSO – dimethyl sulfoxide
DNA - deoxyribonucleic
dNTPs – deoxynucleotide
DTT – dithiothreitol
EDTA – ethylenediamide tetraacetic acid
ELSD – evaporative light scattering detection
ExPASy – Expert Protein Analysis System
FabI – enoyl-ACP reductase
FAD – flavin adenine dinucleotide
Fe - iron
FMN – flavin mononucleotide
G – Guanine
g - grams
G16 – 2-hexyl-1-decanol
G20 – 2-octyl-1-dodecanol
G24 – 2-decyl-1-tetradecanol
gDNA – genomic DNA
H – Hydrogen
h – hour
HPLC – high performance liquid chromatography
IPTG – isopropyl β-D-1 thiogalactopyzanoside
K – Potassium
l - liter
LB – luria bertani
LC-MS – liquid chromatography mass spectroscopy
M - molar
MDR – multidrug resistance transporter
MEL's – mannosyl erythritol lipids
Mg – Magnesium
min - minutes

MQ – Milli-Q
N - Nitrogen
Na – Sodium
NADH – nicotinamide adenine dinucleotide phosphate
O – Oxygen
OD – optical density
P – Phophorus
p - promoter
PCR – polymerase chain reaction
PMSF – phenylmethanesulfonylfluoride
S – sulfer
SD – synthetic dextrose
SHP – sequencial hybrid primer
T – temperature
T – thymine
t - terminator
TAE – Tris-acetate-EDTA
TBE – Tris-borate-EDTA
Uat – acytil transferase
Ugt – glucosyltransferase
UV – Ultraviolet
V – volume
V – volts
WT – wild type
X-GAL – 5-bromo-4-chloro-3 indolyl
YPD – yeast extract-peptone-dextrose

Figures list

Table list

Literature review

1. Surfactants

Surfactants are compounds that have the ability to decrease surface tension between two phases. They are amphiphilic molecules consisting of a hydrophilic head (carbohydrate, alcohol alkoxylates, amino acids, carboxylates, sulphates, sulphonates our phosphates) and a hydrophobic tail (hydrocarbon chain) (Hommel & Ratledge 1993). Surfactant activity is evaluated by their capacity of decreasing surface tension between water and air (72.80 mN/m), its critical micelle concentration (CMC) and its hydrophilic-lipophilic balance (Scamehorn *et al.*, 2004).

Surfactants' production worldwide is more than 13 million ton per year and they are considered one of the most important classes of industrial chemicals (Levinson, 2009). These compounds have a wide variety of applications: they can act as detergents, wetting agents, emulsifiers, foaming agents and dispersants (Reis *et al*, 2013). Surfactants are present on a large number of products that we use on daily basis, mainly in household and laundry detergents (more than half of the annual production) (Van Bogaert *et al*., 2011).

The majority of these compounds are obtained from oleochemical and petrochemical resources, which are often associated with ecological problems due to their environment toxicity, bio-accumulation and biodegradability (Ivanković & Hrenović, 2010). Due to this issue, the interest in the production of biosurfactants is growing.

2. Biosurfactants

Biosurfactants produced by micro-organisms have high economic and environmental interest, since renewable sources are used for their production, they have low ecotoxicity, have mild production conditions and have easy biodegradability (Van Bogaert *et al.*, 2011). They are a group of microbial surface-active compounds, which are classified according to their chemical structure and microbial origin. (Banat *et al.,* 2010).

The main application of biosurfactants is in cleaning products, where they can substitute the classic surfactants. They also have environmental applications in remediation and soil washing, although limited. They are used in oil recovery and

processing. They have application in medicine and pharmaceutical industry, which take advantage of their antimicrobial activity. They have application in gene therapy, immunotherapy and medical insertion due to their anti-adhesive activity and immunemodulating properties (Banat *et al.*, 2010).

Despite the fact that biosurfactants show better properties that their chemical analogous, including better stability at different values of pH, temperature and salinity, and their environmental advantages, their introduction in market is not fully accomplished. The main reason is their production cost, which is not competitive enough with the prices of the surfactants (Banat *et al.*, 2010). By developing bioprocesses which have higher efficiency, use of cheap resources, and construction of overproducing strains, this problem can be overcome (Raza *et al.*, 2009; Thavasi *et al.,* 2007, 2008).

2.1 Classes of Biosurfactants

Biosurfactants can be subdivided in two main classes: glycolipids, which include rhamnolipids, sophorolipids, cellobiose and mannosylerythritol; and lipopeptides.

Rhamnolipids (figure 1) are one of the most studied classes of biosurfactants, having applications in soil remediation and pest control (Mulligan & Wang, 2006; Sha *et al.*, 2012). The most important producing organism is *Pseudomonas aeruginosa,* obtaining yields of 100g/L (Giani *et al.*, 1997). This organisms has the disadvantage of being pathogenic. Rhamnolipis are constituted by one or two L-rhamnose molecules and a mono or dimer fatty acid. The biosynthesis of rhamnolipids starts by the production of deoxythymidine diphosphate (dTDP)-L-rhamnose, in which a glucose-1-phosphate is converted into a rhamnose. This reaction is mediated by the rmlABCD operon. Secondly, a β-hydroxy fatty acid is coupled to an acyl carrier protein, by RhlA, forming a 3-(3-hydroxyalkanoyloxy) alkanoic acid. The hydrophilic and the hydrophobic moieties are connected, mediated by RhlB or FhlC, forming a monorhamnolipid or a dirhamnolipid, respectively (Abdel-Mawgou *et al*., 2010).

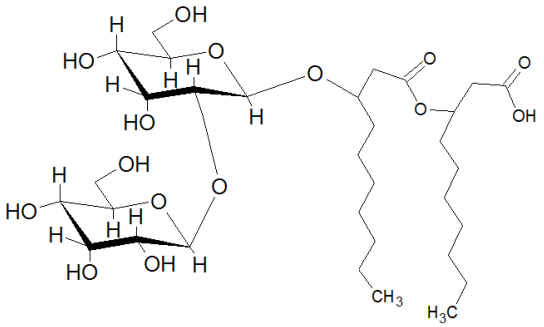


Figure 1 Structure of a dirhamnolipid a-L-rhamnopyranosyl-a-L-rhanopyranosyl-b-hydroxydecanoyl-b-hydroxydecanoat

Cellobiose lipids (figure 2) are used in the formulation of fungicides and are constituted by a cellobiose molecule and a hydroxylated palmitic acid. The biosynthesis of cellobiose lipids starts by the hydroxylation of the terminal carbon of palmitic acid by the enzyme Cyp1. By the action of Cyp2, a second hydroxyl moiety is introduced at the ω-1 position of the palmitic acid. Finally, a glucosyltransferase (Ugt1) couples the cellobiose to the dihydroxy palmitic acid. The enzymes acytil transfersae (Uat1) and acid alpha hydroxylase (Ahd1) can induce further modifications (Teichmann *et al.*, 2011). *Cryptococcus humicola* produces these surfactants in yields of 13.1g/L (Morita *et al.*, 2011).

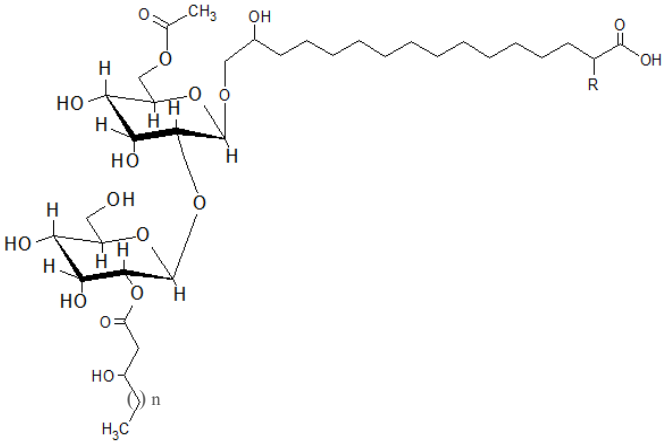


Figure 2 Structure of a cellobiose lipid. R=H or OH, n=2 or 4

Ustilago maydis produces two types of glycolipis (up to 30g/L) under nitrogen starvation: cellobiose lipids and mannosyl erythritol lipids (MEL's). Engineering work can be done to induce the production of only one type of glycolipid. Inhibiting the mannosyltransfersase emt1 blocks the production of MEL's, and by deleting the Cyp1 gene the cellobiose lipids are not synthetized (Spoeckner *et al.*, 1999).

Mannosylerythritol lipids (figure 3) are composed by a 4-O-β-D-mannopyranosyl-meso-erythritol which, depending on the producing organism and

substrate, has different number of acetylations patterns (Morita T *et al*, 2008). MEL's have applications in pharmaceuticals due to their self-assembling properties (Inoh *et al.*, 2011; Morita *et al.*, 2013; Worakitkanchanakul *et al.*, 2008). *Ustilago maydis* produces MEL's in yields of 25.1g/L, *Ps antarctica* produces more than 40g/L (Konishi et al., 2008), and *Pseudozyma parantarctica* higher than 100g/L (Morita et al., 2008).

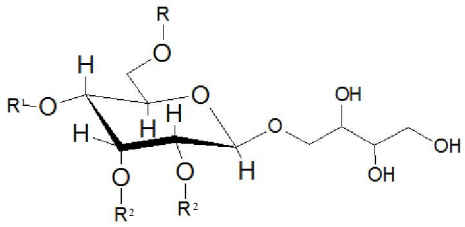


Figure 3 Structures of a mannosylerythritol lipid. R^1 = H or $COCH_3$, R^2 = H or C2-C18 fatty acid

Lipopeptides (figure 4) are constituted of a linear or circular oligopeptide, which number and type of aminoacids varies. They are produced by several species, such as *Aspergillus*, *Bacillus* and *Pseudomonas*. Their importance relays on their antimicrobial activity (Strieker *et al*., 2009). Surfactin, an example of lipopetides, is produced in a concentration of 3.6g/L (Yeh *et al.*, 2005).

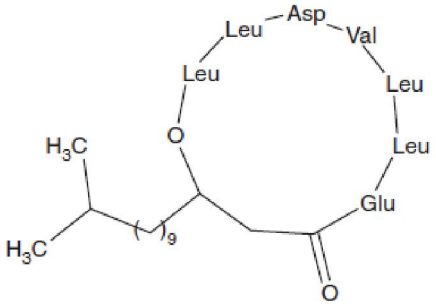


Figure 4 Structure of a lipopeptide (surfactin)

3. Sophorolipids

Sophorolipids have been known for 50 years (Gorin *et al.* 1961), but have recently gain notoriety due to their wide range of applications, their biodegradability, low ecotoxicity, and use of renewable resources in the fermentation process. Sophorolipids are produced by several yeast strains, but the most studied one is *Starmerella bombicola*. A sophorolipid molecule is constituted of a sophorose molecule (2-O-β-D-glucopyranosyl-D-gluco-pyranose) which is connected with a terminal or subterminal hydroxylated fatty acid, and can have one or two acetylations and one lactonization (figure 5).

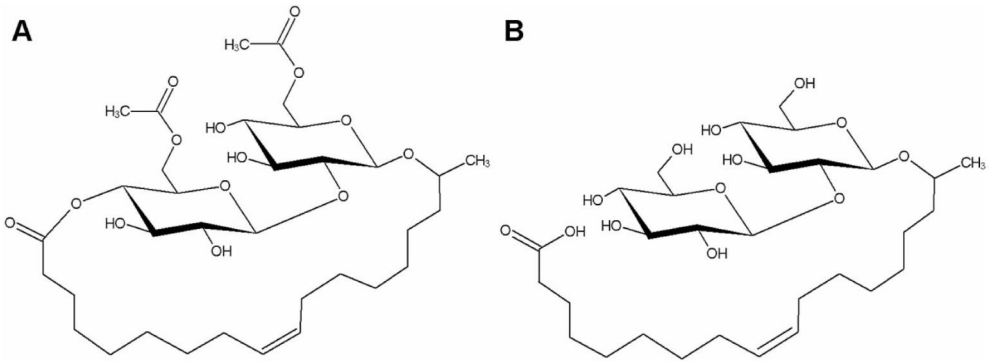


Figure 5 Sophorolipid molecules produced by *Starmerella bombicola.* (a) diacetylated lactonic sophorolipid, (b) non-acetylated open-chain sophorolipid

3.1 Producing organisms

Starmerella bombicola (former *Candida bombicola*) was the third discovered sophorolipid producing strain by Tulloch *et al*. (initially named *Torulopsis bombicola*). *Starmerella bombicola* was firstly isolated from bumblebee honey (Spencer *et al.*, 1970). It is suggested that a symbioses relationship occurs between the bees and the yeast (Rosa *et al.,* 2003).

S. bombicola is a non-pathogenic species and it is known for producing high amounts of sophorolipids (over 400g/L) in low cost fermentative medium and 70% substrate conversion can be achieved (Daniel *et al.*, 1998). The strain ATCC22214 is the most used one, since it is the most efficient sophorolipid producer. This yeast has also shown to be promising for the production of tailor-made sophorolipids and other biomolecules (Roelants *et al.*, 2013).

Sophorolipids are considered secondary metabolites, since they are produced during the stationary phase and they aren't crucial for the cell viability. Synthesis starts at the beginning of the stationary phase and is probably related to nitrogen starvation. Although the physiological role of sophorolipids it still is not clear, they probably act as extracellular carbon storage for the yeast. This way, the carbon source is less available for the development of other microorganisms.Furthermore, sophorolipids also have antimicrobial activity, which contributes to the niche protection (Davila *et al.*, 1992).

Another important sophorolipid producer is *Candida apicola*, which was the first discovered sophorolipid producer (Gorin *et al.* 1961). This microorganism produces the same molecules and with comparable yields as *S. bombicola*. However, it's not as extensively used. Recently, other producing microorganisms have been discovered. *Candida batistae* produces the same sophorolipid molecules as *S.*

bombicola, but the acidic forms are more common than the lactonic ones, the terminal hydroxylation of the fatty acid is preferred and the yields obtained are 6g/L (Konishi *et al.*, 2008). Other sophorolipid producing strains belonging to the *S. bombicola* clade were identified: *Candida riodocensis*, *Candida stellata* and *Candida sp.* Y-27208 (Kurtzman *et al.*, 2010). *Wickerhamiella domercqiae*, a more phylogenetic distant species, produces similar molecules (Chen *et al*. 2006). *Rhodotorula bogoriensis*, another sophorolipid producer, is going to be discuss later on this thesis.

3.2 Structure and properties of sophorolipids

Sophorolipids produced by *S. bombicola* are composed by a sophorose molecule and a hydroxylated fatty acid. However, the referred yeast produces a mixture of sophorolipid molecules, which differ in: the acetylation and lactonization patterns, the incorporation of a fatty acid of 16 or 18 carbon atoms, number of unsaturated bonds in the hydroxylated fatty acid, and the position of the hydroxyl group (terminal or subterminal) (Asmer *et al.*, 1988; Davila *et al.*, 1993). The abundance of lactonic and acetylated forms depends on the strain used and the conditions. Fatty acids with the hydroxylation in the subterminal are more abundant.

The differences between the molecules cause variation in the physicochemical properties. Lactonic sophorolipids display better surface tension, antimicrobial activity and they are more readily isolated, due to the fact that lactonization reduces the rotational freedom of the sophorolipid, enhancing the formation of crystals. The acetylated sophorolipids have a better foam production, solubility and have better antiviral and cytokine stimulating properties (Shah *et al.*, 2005). Sophorolipids are active at a wide temperature range (Nguyen *et al.*, 2010) and salt concentrations (Hirata et al., 2009), but they are instable at high pH values. Above pH 8, irreversible hydrolysis of the acetyl groups and ester bonds occur, and at pH 6 to 7 the sophorolipid molecules dissolve perfectly (Soliance. Product guide, 2004). Sophorolipids dissolve in polar solvents, as ethanol (Hu *et al.*, 2001). Sophorolipids lower surface tension from 72.80mN/m to 30mN/m in water, and have a critical micelle concentration between 11mg/L and 250mg/L. These values slightly differ according to the structure of the sophorolipid molecules (Develter *et al.*, 2010).

3.3 Biosynthesis of Sophorolipids

The sophorolipid biosynthesis (figure 6) starts with the hydroxylation of the terminal (ω) or the subterminal (ω-1) position of a fatty acid, which usually has 16 or 18 carbon atoms. The reaction is mediated by the cytochrome P450 monooxygenase enzyme CYP52M1. Due to the specificity of CYP52M1, fatty acids different than 16 and 18 carbon atoms will be elongated or β-oxidated (Felse *et al.*, 2007).

The enzyme cytochrome P450 oxidoreductase is an electron-donor. The lavin adeninde dinucleotide (FAD) acquires electrons from the reduction of nicotinamide adenine dinucleotide phosphate (NADH) and transfers them to the flavin mononucleotide (FMN). FMN interacts with a cytochrome P450 enzyme and transfers the electrons to its heme iron site, allowing the catalysis to occur (Huang *et al.*, 2008).

The second step is the addition of two glucose units, in a stepwise manner, mediated by two UDP-glucose dependent glucosyltransferases (UgtA1 and UgtB1, respectively). The first glucose unit is glycosidically added (in the position C1') to the hydroxyl group of the fatty acid. The second glucose is then glycosidically added to the C2' position of the first glucose unit (Saerens *et al*., 2011a,b). Thereby, the hydroxylated fatty acid is linked to the sophorose molecule by an ether bound, which gives stability to the molecule by turning it less prone to hydrolysis.

The product can be modified by one or two acetylations and a lactonization, by an acetyl-coenzyme A depended acetyltransferase and an extracellular lactone esterase, respectively. The acetylation occurs in the 6'- and/or 6"-position (Esders *et al.*, 1972), and the lactonization occurs between the carboxyl group of the hydoxylate fatty acid and the 4''-hydroxyl group of the sophorose molecule (Asmer *et al.*, 1988). The transport of sophorolipid molecules to the cell exterior is mediated by a multidrug resistance transporter (MDR), which is a membrane integrated protein in the *S. bombicola* cells. The lactonization occurs extracellular. The majority of sophorolipids are modified by acetylations and/or lactonization.

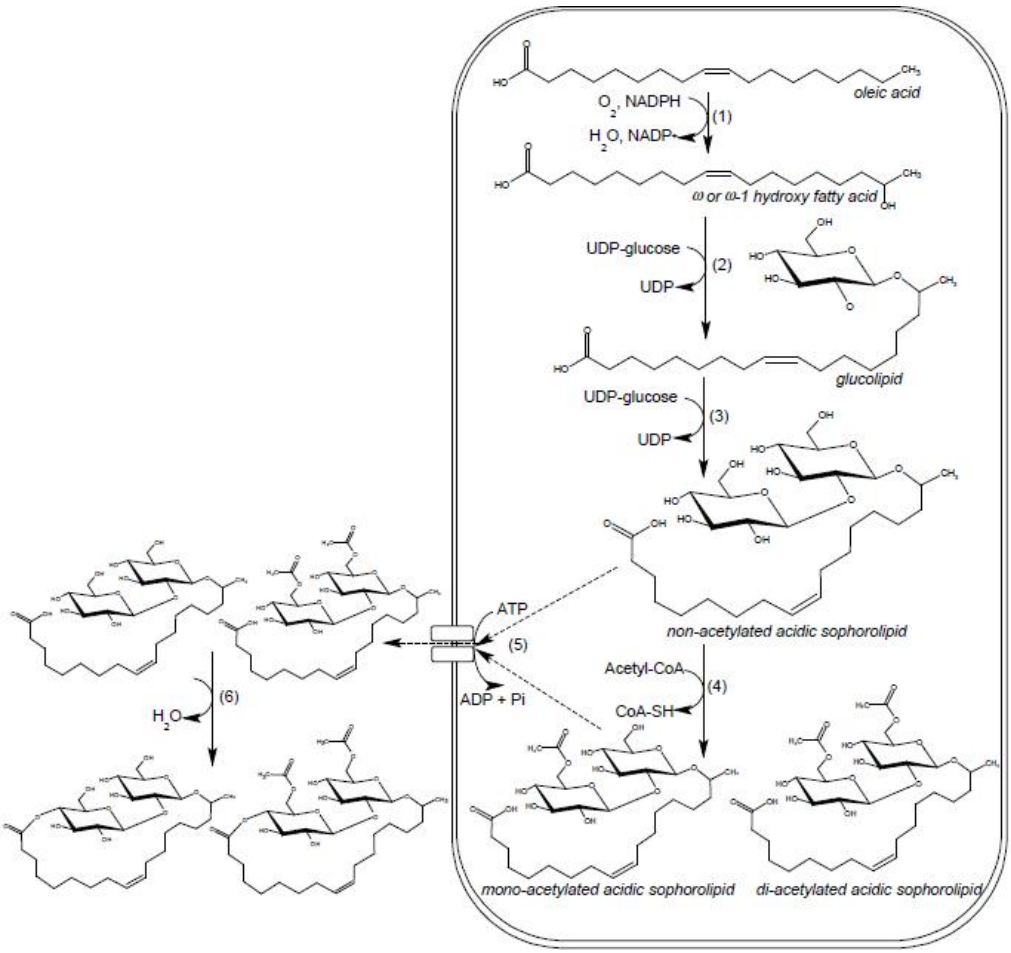


Figure 6 Proposed sophorolipid biosynthetic pathway. (1), cytochrome P450 monooxygenase, (2) glucosyl-transferase I, (3) glucosyltransferase II, (4) acetyltransferase, (5) MDR protein, (6) lactone esterase.

S. bombicola has a sophorolipid gene cluster of 11000bp (figure 7). This cluster is constituted by the genes responsible for the codification of the two gluosyltransfersases (Ugtb1 and Ugtb2), a transporter (MDR), an acetyltransferase (at), and the cytochrome P450 monoxygenase gene (CYP52M1). These genes are flanked by a putative alcohol dehydrogenase (adh) and a gene of unknown function (Van Bogaert et al., 2013). All five genes are intronless and co-regulated. The cluster is present in a subtelomere, which is a region prone to variation and gene transfer, and where chromatin remodeling and epigenetics mediate the transcription of genes. This region is ideal for co-regulated expression of large clusters. (Palmer & Keller, 2010).

The enzyme lactone esterase, responsible for the lactonization of the sophorolipid, is not located in the gene cluster.

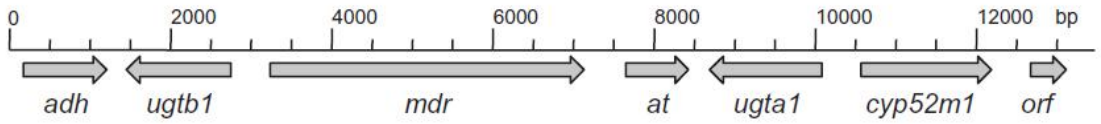


Figure 7 Sophorolipid gene cluster of *Starmerella bombicola*. The cluster is constituted by an acetyltransferase (at), two gluosyltransfersases (ugta1 and ugtb 1), a transporter (mdr), the cytochrome monoxygenase (cyp52m1) function (Van Bogaert *et al.*, 2013).

The optimal growth temperature for *S. bombicola* is at 28.8°C and the optimal temperature for the sophorolipid production is at 21°C. However the fermentation process is normally performed at 25°C or 30°C, since the production at 21°C is difficult and costly (Gobbert *et al.* 1984). The optimal aeration for the production of sophorolipids is between 50 and 80mMO2/Lh. Good aeration conditions are crucial during the exponential phases (Guilmanov *et al.* 2002). The optimal pH value is 3.5 (Gobbert *et al.*, 1984).

The medium used for the production of sophorolipids usually contains a hydrophilic and a hydrophobic carbon sources, a nitrogen source and some minerals in small amounts (Mg2+, Fe3+, Ca2+, Zn2+ and Na+) (Davila *et al.*, 1997).

The hydrophilic source is usually glucose (Klekner *et al.*, 1991). Other substrates, including low cost mediums, can also be used, but usually lower yields are obtained. Oils, fatty acids, and respective alkanes and esters, can be used as hydrophobic carbon source. The rate and position of the hydroxylation of a fatty acid is determined by its chain length, which influences its incorporation into the sophorolipid molecule. The best sophorolipid production yields are obtained when using oleic acid (C18:1). Rapeseed oil is frequently used, due to its low cost and high content of oleic acid (Davila *et al.*, 1994). Usually the hydrophobic source is added in fed-batch (Davila *et al.*, 1992). In downstream process at laboratory scale, polar solvents as ethanol are used to extract sophorolipids from the culture broth. At industrial scale, when dealing with high volumes and yields, centrifugation or decantation can be used (Inoue *et al.*, 1980). Techniques as chromatographic purification with silica gel or preparative reversed phase columns allow the purification of specific sophorolipid structures (Lin, 1996).

3.4 Applications of Sophorolipids

The most important feature of sophorolipids is their surfactant activity, because of this they have applications in house cleaning, cosmetics, pharmaceutical, food, petroleum industries, in remediation and in nanotechnology.

Sophorolipids are commonly used in formulation of cleaning and laundry agents, due to their emulsifying activity (Develter *et al.*, 2007; Hall *et al.*, 1996). They can potentially be used in food industry to increase the quality of wheat flour products (Akari & Akari, 1987).

They are used in cosmetic hygienic and pharmaco-dermatological formulations, taking advantage of their surfactant and antimicrobial activity. They have antibacterial (*Bacillus subtilis, Staphylococcus epidermidis, Staphylococcus aureus, Streptococcus faecium, Propionibacterium acnes* and *Corynebacterium xerosis*) (Lang *et al.*, 1989), antifungal (*Phytophthora* sp. and *Pythium* sp.) (Yoo *et al.*, 2005) and anti- human immunodeficiency virus activity (Gross *et al.*, 2007). Sophorolipids also appear to have anticancer activity, by triggering cell differentiation by interactions with the cellular membrane (Isoda *et al.*, 1997). They have cytotoxic effects in human cancer cells (Chen *et al.*, 2006) and are also modulators of immune response (Sleiman *et al.*, 2009). Due to all these properties, sophorolipids have several potential medical applications and are used in cosmetic formulations.

In petroleum industries, sophorolipids can be used in second oil recovery, and in removing hydrocarbons from drill materials and their regeneration from dregs and muds. (Baviere *et al.*, 1994; Marchal *et al.*, 1999; Pesce, 2002). Sophorolipids can also be used in water and soil decontamination (Ducreux *et al.*, 1997). Recently, nanotechnology applications were discovered for sophorolipids. The nanoparticles need to be aggregated with an agent to allow their dispersion in organic solvents and water. Sophorolipids can act as the referred agent, in which the hydrophobic tail will attach to the particle. They also act as structure directing agents for the nanoparticles (Kasture *et al.*, 2007).

4. Sophorolipids with a branched hydrophobic tail

In this thesis, the aim is to produce sophorolipids with a branched hydrophobic tail, using *S. bombicola* as producing organism. The production of such molecules have low yields by their natural producer, *Rhodotorula bogoriensis*. This problem can be

bypassed by inducing *S. bombicola* to produce sophorolipids with branched hydrophobic tail, making it easier to access their properties and determine their possible applications. This can be achieved by using unconventional hydrophobic carbon sources or by genetic engineering the strain. These new-to-nature molecules are characterized by having a hydroxyl group in the middle of the chain.

4.1 Modified sophorolipids

Sophorolipids which are different from the ones synthetized naturally are called new-to-nature sophorolipids. Modifications in the sophorolipids molecules can be achieved by altering the cultures conditions or by genetic engineering the producing strain.

In order to enhance *S. bombicola* to produce new-to-nature sophorolipids, by adding unconventional substrates to the fermentation medium, two strategies can be used. The first one requires the use of hydroxylated substrates, and the second one the use of hydrophobic substrates with a stearic acid-like structure or chain length. This way, the substrate specificity of cytochrome P450 can be circumvented (Van Bogaert *et al.*, 2011). However, one of the main problems in producing new-to-nature sophorolipids is the fact that the obtained product of the fermentation process is a mixture of these molecules and the classic sophorolipids (Van Bogaert *et al.*, 2008).

Palme *et al.* used as substrate a mixture of branched fatty alcohol 2-hexyl-1-decanol (Rilanit G16), and managed to identify a 1-O-β-glucopyranosyl-2-hexyldecanol (figure 8,A). With the incorporation of the branched fatty acids, the hydroxylation is in-chain. However, only one glucose molecule was incorporated. Total yields of 6g/L were obtained. The whole sophorolipid mixture obtained showed efficient low foaming and better to decrease the contact angle at lower concentrations than Simulsol AS 48 (Palme *et al.*, 2010).

Schofield *et al*. (2013) developed a method of producing diamide derivates by transalkylidenation and amidation of lactonic sophorolipids. The obtained molecules have a branched structure (figure 8,B).

Publications of Bhangale *et al.* (2014) and Bajaj *et al.* (2015) report the production of sophorolipids with a branched hydrophobic tail using castor oil as hydrophobic substrate for cultures of *S. bombicola*. Castor oil has ricinoleic acid as constitute, molecule which has an in-chain free hydroxyl group. In Bhangale *et al.*

(2014) study there was production of a sophorolipid with in-chain hydroxylation and lactonization, and in Bajaj *et al.* (2015) study, a branched sophorolipid with in-chain hidroxylation was reported (figure 8,C). However, in this last paper, the researchers present specific molecular masses for both molecules but don't specify the number of acetylations. Both researches used the same substrates for the *S. bombicola* culture but they have obtained different products, making the results dough full.

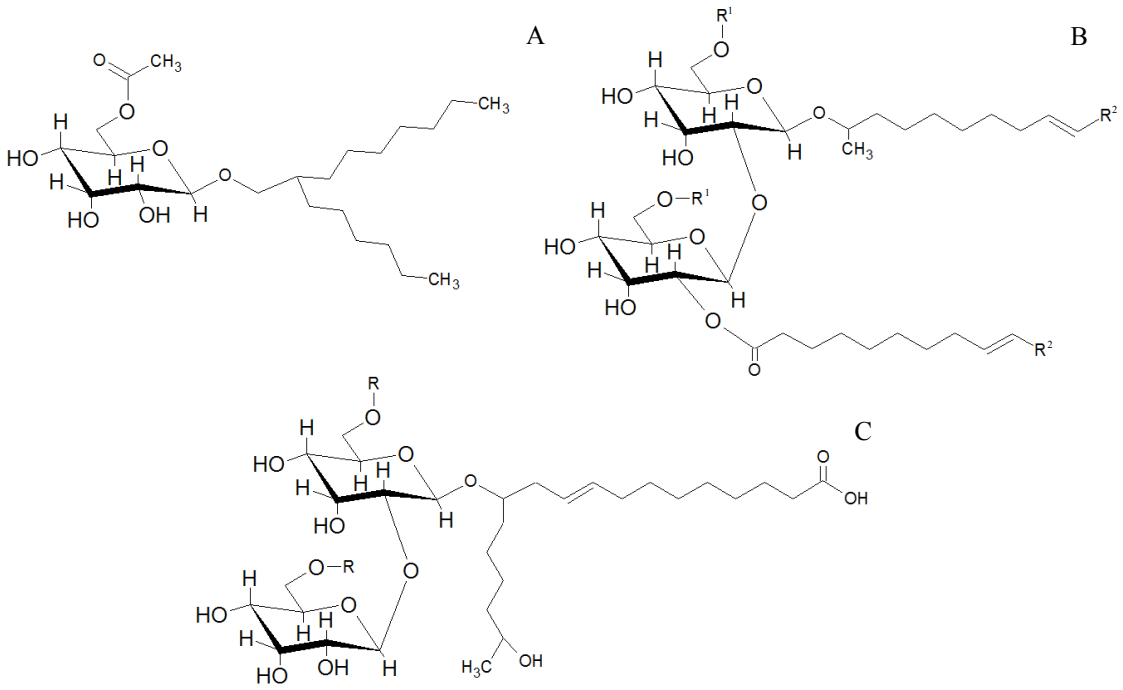


Figure 8 Structures of sophorolipid with a branched hydrophobic tail. Molecule A is 1-O-β-glucopyranosyl-2-hexyldecanol developed by Palme *et al*. Molecule B is diamide derivate developed by Schofield *et al*. R^1= H or Ac, R^2= H, alkyl, aryl, heterocyclic, cationic, anionic groups. Molecule C is a sophorolipid with a branched hydrophobic tail (without acetylation or with one or two- R) developed by Bajaj *et al.*

4.2 Sophorolipids with branched hydrophobic tail produced by *Rhodotorula bogoriensis*

Rhodotorula bogoriensis is a basidiomycetous yeast that produces sophorolipids with a branched hydrophobic tail (all the other producing yeasts are ascomycetes). Sophorolipids produced by *R. bogoriensis* (figure 9) normally contain a docosanoic acid (fatty acid with 22 carbons), they are never lactonized and the hydroxyl group is at C13 position, which causes the formation of a branched hydrophobic tail. The principal component of the sophorolipids mixture, produced by the referred yeast, is a 13-[2-*O*-β-D-glucopyranosyl-β-D-glucopyranosyloxy]-docosanoic acid 6',6''- diacetate glycolipid (Tulloch *et al.*, 1968).

It is difficult to determine the physiochemical properties of these sophorolipids since their production by the referred yeast is low. By enhancing their production in *S. bombicola*, better yields could be obtained. The sophorolipids yields obtained by *R. bogoriensis* with the use of rapessed oil were 1.26g/L and without oil 0.33g/L. The yield obtained when meadowfoam oil was used was 0.77g/L, which is a low value since this oil is enriched in fatty acids with 22 and 20 carbon atoms (Zhang *et al.*, 2011).

The biosynthesis of the mentioned sophorolipids starts by binding of the sophorose molecule to a hydroxyl docosanoic acid, followed by the acetylation of the sophorose molecule in the positions 6' and 6'' (Ribeiro *et al.*, 2012).

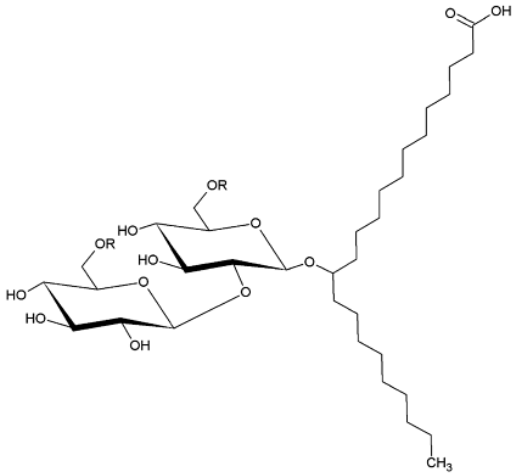


Figure 9 Structure of a sophorolipid with a branched hydrophobic tail produced by *R. bogoriensis.* R=H or $COCH_3$

4.3 Oleate hydratase

Oleate hydratase is an enzyme responsible for the conversion of oleic acid into 10-hydroxystearic acid by the stereospecific hydration of its 9,10 double bond (figure 10). It was first discovered by Wallen *et al.*, in *Pseudomonas* sp. strain 3266, but was only recently isolated and biochemical characterized (Bevers *et al.*, 2009). This enzyme is common in many microorganisms, such as *Elizabethkingia meningoseptica* and *Rhodococcus rhodochrous*.

Oleic acid is important to many species of microorganisms, since it is convertible to a wide range of valuable derivates. Oleic acid can undergo hydration or epoxidation of its double bond, reduction of the carboxylic acid, hydroxylation at the saturated carbon chain and degradative reactions (Saleh *et al.*, 1992). However, oleic acid is toxic to some species of bacteria (Raychowdhury *et al.*, 1985) by disrupting their cell membrane (Freenway & Dyke, 1979) and by inhibiting the activity of enoyl-ACP

reductase (FabI) (Zheng *et al.*, 2005). For these reasons, the conversion of oleic acid is considered a detoxification and a survival mechanism for the microorganism when present in fatty acid-rich environments (Joo *et al*., 2012). Oleate hydratase also plays part in the virulence of some microorganisms, as *Streptococcus pyogenes* M49 (Volkov *et al.*, 2010).

10-hydroxystearic acid is produced by *Pseudomonas* during the stationary phase and it is not further metabolized. However, some species of microorganisms further oxidize it to 10-ketostearic acid (Bevers *et al.*, 2009).

Oleate hydratase can be used to produce an alkene by dehydration of an alcohol (Marliere 2011) and sebacic acid (Piatesi *et al.*, 2014). This enzyme can be used to take care of in-chain hydroxylation of a fatty acid, and that way produce sophorolipids with a branched hydrophobic tail using *S. bombicola*.

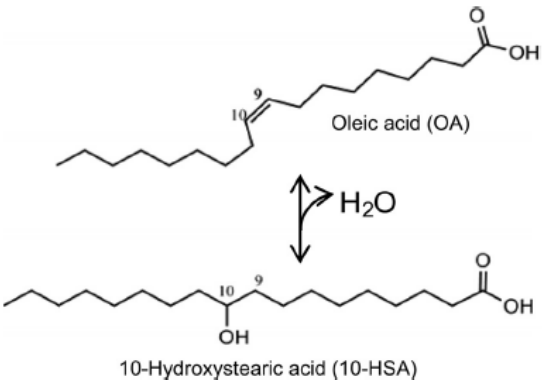


Figure 10 Conversion of oleic acid into 10-hydroxystearic acid by oleate hydratase

Materials and Methods

1. Medium preparation and strains

1.1 Lang medium

Liquid medium used for the production of sophorolipids by *Starmerella bombicola*. Lang-medium is constituted of 4g/L of yeast extract, 5g/L of 3Nacitraat·2H2O, 1.5g/L of NH4Cl, 1g/L of KH2PO4, 0.16g/L of K2HPO4, 0.7g/L of MgSO4·7H2O, 0.5g/L of NaCl and finally 0.27g/L of CaCl2·2H2O. The pH is set at 5.8, with the use of a pH meter (Lang *et al.,* 2000).

1.2 3C medium

Solid medium used for the growth of *Starmerella bombicola*. 3C medium is constituted of 10g/L of yeast extract, 1g/L of ureum, 110g/L of glucose and 20g/L of agar.

1.3 Synthetic dextrose (SD) medium

Medium used for the growth of *Starmerella bombicola* competent cells after transformation. SD medium contains 6.7g/L of yeast nitrogen base without amino acids, 22g/L of glucose and 20g/L of agar in the case of solid medium. The plates used were Ura negative, so a complete supplement mixture (CSM) stock solution (10x) was made and 25mL added to the composition of the medium.

1.4 Yeast extract-peptone-dextrose (YPD) broth

Medium used for maintaining and propagating *Starmerella bombicola* cells. YPD is constituted of 10g/L of yeast extract, 20g/L of peptone, 20g/L of dextrose and 15g/L of agar in the case of solid medium (yeast extract-peptone-dextrose agar – YPD).

1.5 Luria Bertani (LB) medium

Liquid and solid medium used for the growth of *Escherichia coli*. LB medium is constituted of 10g/L of Tryptone, 5g/L of yeast extract and 5g/L of NaCl. For solid medium, 15g/L of agar is used (Bertani, 1951).For selection with ampicillin, 100mg/mL is used. For blue and white screening of colonies, the following reagents are used: 3%

of X-gal (5-bromo-4-Chloro-3indolyl) in DMF (dimethyl formamide), 100mM of IPTG (isopropyl β-D-1 thiogalactopyzanoside) and 0.1g/mL of ampicillin.

1.6 Gym medium

Liquid and solid medium used for the growth of *Rhodococcus rhodochrous*. Gym medium is constituted of 4g/L of glucose, 4g/L of yeast extract and 10g/L of malt extract. For solid medium, 2g/L of CaCO3 and 12g/L of agar are used. The pH is set at 7.2.

2. Production of sophorolipids using Guerbet alcohols

2.1 Strains and culture conditions

The yeast *Starmerella bombicola* is used in the experiments and the medium used is Lang-medium. The cells are incubated in flasks of 200mL, at 30°C and 200rpm. The hydrophobic carbon source is added after 96h after the inoculation. Glucose is added once when its levels are close to zero. The incubation is stopped after 20 days.

2.2 Sampling

2.2.1 Biomass formation

The biomass formation is determined by counting of the colony forming units (CFU). In 3C-medium plates, dilutions till 10-7 of the culture broth were plated (10µL) in 3C medium. These dilutions were prepared in physiological water. The plates were incubated at 30°C for three days.

2.2.2 Glucose concentration

1 mL of culture broth is analyzed with the 2700 Select Biochemistry Analyzer (YSI Inc.). Samples are diluted, so that the concentration was within the equipment range detection (0.1-7.0g/L).

2.2.3 Sophorolipid formation

At 1mL of culture broth, 3mL of ethanol are added. The samples are left at room temperature for 30min, and then centrifuged at 45000rpm for 5min. 1.5mL of the supernatant is added to High Performance Liquid Chromatography (HPLC) vials. The

samples were analyzed by HPLC. If the solutions aren't clear, centrifugation and/or filtration is used.

At the end of the incubation period, a final extraction is performed. 3 volumes of ethanol are added to the residual fermentation medium and the cultures are put at room temperature for 30min. 7mL of each culture broth are collected and centrifuged at 45000rpm for 5min. Half of the supernatant is transferred to a falcon tube and the ethanol evaporated in a rotavapor. After 30min, the rest of the supernatant is added to the falcon tube in the rotavapor. The use of the rotavapor allows the concentration of the sample. In the case of this experiment, the evaporation is stopped when the volume of the samples reaches 1mL, volume which was transferred to HPLC vials and analyzed.

2.3 Extraction of sophorolipids from culture broth

2.3.1 Removal of yeast cells

3 volumes of technical ethanol are added to the fermentation medium. The sample is centrifuged for 10min at 4000rmp. The supernatant is put in a 2L boiling flask, which weight had already been determined. 1mL of sample is taken to analysis in HPLC.

2.3.2 Removal of water soluble components

The sample is evaporated for 1h30min to 2h in a rotavapor. The temperature starts at 30°C (azeotropic mixture) and slowly increases till 80°C. Evaporation is completed when all the water is gone and the residue more viscous. The final weight is determined.

400mL of ethanol (or 2 volumes) is added, which is then rotated in a rotavapor warm water bath without vacuum for 10 to 15min, to dissolve the sophorolipids. The sample is taken out of the rotavapor and let stand for 30min. The ethanol mixture is put over a Whatman filter with vacuum filtration, and the filtrate brought to a new 1L boiling flask. A sample is taken from the filtrate to be analyzed in HPLC. The filtrate is evaporated, process that takes 2h. At the end of the process, temperatures between 80 and 90°C are used. The weight of the sample is determined and, if it is bigger than the previous one, there is still water present.

2.3.3 Separation of sophorolipids and residual hydrophobic carbon source

100mL of water is added to the boiling flask and all dry matter dissolved by establishing the pH at 6.5. 100 mL of Hexane are added, mixed by shacking vigorously and brought in a separatory funnel. The sample is allowed to stabilize. The upper hexane layer will contain residual hydrophobic carbon source. The lower water layer will contain sophorolipids. The two layers are collected in separate 0.25L boiling flasks. A third sample is taken from the water phase and form the hexane phase. Both phases are evaporated and the weights of the residues determined. The residue is dissolved again in all limited amount of ethanol and put in a falcon tube. The samples are stored at 4°C.

The samples collected along the sophorolipid extraction are analyzed by HPLC.

2.4 HPLC analysis

The samples are analyzed by HPLC on a Varian Prostar HPLC system using a Chromolith® Performance RP-18e 100-4.6 mm column from Merck KGaA atc 30°C and Evaporative Light Scattering Detection (ELSD, Alltech). A gradient of two eluents is used: a 0.5% acetic acid aqueous solution and acetonitrile. The gradient started at 5% acetonitrile and linearly increased till 95 % in 40min. The mixture is kept this way for 10min and was then brought back to 5% acetonitrile in 5min. A flow rate of 1mL/min is used.

2.5 LC-MS analysis

Liquid chromatography mass spectroscopy (LC-MS) analysis was performed by a Micromass Quattro LC. The detection range was set at m/z 200 to 900 and the negative ion mode was applied. The same column and LC conditions as described for the HPLC analysis was used.

3. Enzyme test

A culture of *S. bombicola* cells is inoculated in Lang-medium for 72h, at 30°C and 200rpm. For the proteins extraction, the OD at 600nm (X) of the cell culture broth is measured. 50mL of the culture are put it in a 50mL falcon and centrifuged for 5min at 4500rpm. The pellet is washed with 50mL of distilled water. The sample is again centrifuged for 5min at 4500rpm. The pellet is resuspended in lyse buffer. The amount

of lyse buffer can be calculated from the formula: (OD X / OD100) * 50mL = Lyse buffer mL

The amount of glass beads added is equal to the amount of lyse buffer. The sample is vortexed for 15min. The sample was vortexed for 30sec, followed by 30sec on ice. The sample was centrifuged for 5min at 4500rpm and the supernatants used in the enzyme reaction.

For the lyse buffer preparation, KH2PO4 (6.805g/L) and K2HPO4 (8.705g/L) are used to prepare the phosphatebuffer (50mM, pH 8.5). The following additives are added: Glycerol (5% final concentration), DTT (1000x stock solution, 0.5M), PMSF (1000x stock solution, 1M) and MgCl2 (1000x stock solution, 0.5M). All the additives are solubilize in MQ water, with the exception of PMSF for which DMSO is used.
500µL of solution is used for the enzyme reaction (in duplicate for each substrate). The following reagents are used: 100µL of UDP-glucose (final 20mM), 25µL of substrate (2mM), 200µl of supernatants, 175µL of phosphatebuffer (50mM), and MQ water until 500µL.

The reactions are placed at 30°C for 4h and 24h. The samples are placed for 5min at 95°C, to stop the reaction. One volume of ethanol is added to each sample to dissolve the components. The samples are analysed by HPLC.

4. Polymerase chain reaction

4.1 High fidelity PCR

The PCR conditions and reagents are displayed in tables 1 and 2.

Table 1 High fidelity PCR reagents

PCR COMPONENTS	V (µL)
5X Q5 Reaction Buffer	10
dNTPs (2mM)	5
Forward Primer (10µM)	2.5
Reverse Primer (10µM)	2.5
DNA template (10ng)	1
Q5 High-Fidelity DNA Pol	0.5
MQ water	Adjust V until 25
	25

Table 2 High fidelity PCR conditions

PCR CODITIONS	T(°C)	Cycles
30 seconds	98	1
5 seconds	98	30
20 seconds	60[1]	
2 minutes 15 seconds	72[2]	
2 minutes	72	1
infinity	16	∞

[1]The annealing temperature depends on the primers used for the reaction
[2]Time and temperature of the elongation step depends on the length of the DNA template

4.2 Colony PCR

In this method a colonies grown in a solid medium with a selection marker are used. Alongside the colony PCR, the colonies used for the reaction were inoculated in a plate with LB medium and ampicillin, at 37 °C overnight. If colonies with the right DNA fragment are obtained, they are inoculated in liquid LB medium with ampicillin at 37°C overnight, for further plasmid extraction.

The PCR conditions and reagents are displayed in the tables 3 and 4.

Table 3 Colony PCR reagents

PCR COMPONENTS	V(µL)
10x *Taq* DNA polymerase Buffer	2.5
dNTPs-Mix (2mM)	2.5
Forward Primer (10µmol)	1.0
Reverse Primer (10µmol)	1.0
Taq DNA polymerase	0.1
MQ water	Adjust V until 25
	25

Table 4 Colony PCR conditions

PCR CONDITIONS	T (°C)	Cycles
5 minutes	94	1
30 seconds	94	30
30 seconds	50[1]	
1 minutes 50 seconds	70[2]	
10 minuts	70	1
infinity	16	∞

4.3 Degenerated PCR

The degenerated PCR conditions and reagents are displayed in tables 5 and 6.

Table 5 Degenerate PCR reagents

PCR COMPONENT	V (µL)
10x *Taq* DNA polymerase Buffer	5.0
dNTPs-Mix (2mM)	5.0
Forward Primer (10µM)	5.0
Reverse Primer (10µM)	5.0
DNA template	X(100ng)
Taq DNA polymerase	0.4
MQ water	Adjust V till 50
	50

Table 6 Degenerate PCR conditions

PCR CONDITIONS	T (°C)	Cycles
10 minutes	94	1
30 seconds	94	40
60 seconds	45[3]	
2 minutes 50 seconds [4]	68	
7 minuts	68	1
infinity	16	∞

[3]45°C to allow mismatches. If possible, use a gradient from 40 to 55°C
[4]5sec are added at each cycle

4.4 Gene walking

The reagents and conditions for the Primary PCR are given in tables 7 and 8.

Table 7 Primary PCR reagents – Gene walking

PCR COMPONENT	V(µL)
10x *Taq* polymerase Buffer	5.0
dNTPs (2mM)	5.0
Taq DNA polymerase	0.25
gDNA (50ng)	0.10
primary-PCR-single-tag (100ng)	1.86
Forward Primer (100ng)	1.47
Reverse Primer (100ng)	1.43
Primers Gene-walking A, T, C, G (100ng)	1.30
MQ water	Adjust V till 50
	50

Table 8 Primary PCR conditions – Gene walking

PCR CONDITIONS	T (°C)	Cycles
3 minutes	94	1
30 seconds	94	5
60 seconds	35	
3 minutes	72	
30 seconds	94	15
3 minutes	55	
3 minutes	72	
10 minutes	72	1
infinity	16	∞

The reagents and conditions for the SHP-PCR, first round, are presents in tables 9 and 10.

Table 9 SHP-PCR first round reagents – Gene walking

PCR COMPONENTS	V (µL)
10x *Taq* polymerase Buffer	5.0
dNTPs (2nM)	5.0
Taq DNA polymerase	0.25
gDNA (x100 diluted)	0.50
SHP PCR1 single-tag (100ng)	1.52
SHP PCR1 hybrid-tag (1ng)	0.01
Forward Primer 2(100ng)	1.50
Reverse Primer 2 (100ng)	1.40
MQ water	Adjust V till 50
	50

Table 10 SHP-PCR first round conditions – Gene walking

PCR CONDITIONS	T (°C)	Cycles
3 minutes	94	1
30 seconds	94	35
60 seconds	55	
3 minutes	72	
10 minuts	72	1
infinity	16	∞

The reagents and conditions for the SHP-PCR, second round, are presents in tables 11 and 12 (Martin-Harris *et al.*, 2010).

Table 11 SHP-PCR second round reagents – Gene walking

PCR COMPONENT	V(µL)
10x *Taq* polymerase Buffer	5.0
dNTPs (2mM)	5.0
Taq DNA polymerase	0.25
gDNA (x100 diluted)	0.50
SHP PCR2 single-tag (100ng)	1.47
SHP PCR2 hybrid-tag (1ng)	0.01
Foward Primer 3(100ng)	1.45
Reverse Primer 3 (100ng)	1.43
MQ water	Adjust V till 50
	50

Table 12 SHP-PCR second round conditions – Gene walking

PCR CONDITIONS	T (°C)	Cycles
3 minutes	94	1
30 seconds	94	35
60 seconds	54	
1 minute and 30 seconds	72	
10 minutes	72	1
infinity	16	∞

4.5 Sanchis method

Sanchis method is used to correct point mutations in DNA sequences. In the first PCR stage, two mutagenic primers or a mutagenic primer and an antiprimer are used to generate a megaprimer. The annealing temperature, in the second stage of the PCR is increased to eliminate priming by the oligonucleotide primers, and 20 cycles are carried out to amplify the mutated plasmid. Finally, the template plasmids are digested using *DpnI*: 2µL of buffer, 1µL of *Dpn*I enzyme and 3µL of MQ water; incubation for 1h30min on 37°C. The sample is put at 80°C for 20min for the denaturation of *Dpn*I. The resulting library is transformed in competent cells (Sanchis *et al.*, 2008).

Table 13 PCR components of Sanchis method

PCR COMPONENTS	V (µL)
Q5 High-fidelity DNA polymerase (2U/µL)	0.5
5x Q5 reaction buffer	10
dNTP mix (2mM)	5
Template (0.5ng/µl)	2
Foward Primer (10µM)	2.5
Reverse Primer (10µM)	2.5
MQ water	Adjust V till 50
	50

Table 14 PCR conditions of Sanchis method

PCR CONDITIONS	T (°C)	Cycles
30 sec	98	1x
10 sec	98	5x
20 sec	55[7]	
30 sec/kb[5]	72	
10 sec	98	25x
1 min/kb[6]	72	
2 min/kb[6]	72	1x
infinity	4-16	∞

[5]Size of the megaprimer
[6]Size of the whole plasmid
[7]Lowest melting temperature plus 3°C

5. Molecular tools

- GenElute™ Bacterial Genomic DNA Kit (Sigma) - Gram-positive bacterial preparation protocol
- Purification and concentration of PCR products from PCR reaction up to 50µL (innuPREP PCRpure Kit)
- Isolation of plasmid DNA from bacterial lysates (innuPREP Plasmid Mini Kit)
- Isolation of plasmid DNA (QIAprep Spin MiniprepKit)
- Gel Extraction Kit (QIAEX II)

6. Gel electrophoresis

For the gel electrophoresis a solution of 1% agarose is prepared and poured in an electrophoresis tray, at room temperature, until the gel is polymerized (approximately 30 minutes). The gel is transferred to an electrophoresis tray with TAE 1x buffer. After the samples are loaded, the gel was run at 120V, 400A and 25 minutes for most of the experiences, except for the gel extraction which was run at 90V, 400A and for 55 minutes. The gel is transferred to an ethidium bromite solution, where it stays at least for 20 minutes. Then the gel is reveled and analyzed using UV light. The marker used during the electrophoresis is the 2-log DNA ladder (0.1-10kb) (figure 11).

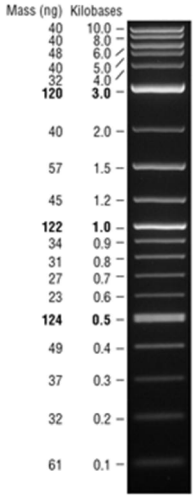


Figure 11 2-Log DNA Ladder visualized by ethidium bromide staining on a 1.0% TBE agarose gel. Mass values are for 1 µg/lane.

7. Transformation of *S. bombicola* cells

For the transformation of *S. bombicola*, a 5mL pre-culture in YPP must be made, and then one in 50mL.

The first step in turning the *S. bombicola* cells competent is to check if the OD600 of the culture is between 0.6 and 1. Then, 50mL of the culture is centrifuged at 4000rpm at 4°C for 5 minutes. The pellet is resuspended in 50mL of ice-cold sterile water. The sample is again centrifuged with the same conditions. The pellet is resuspended in 25mL of ice-cold sterile water. The sample is again centrifuged with the same conditions. The pellet is resuspended in 2mL of ice-cold sorbitol (1M). The sample is centrifuged with the same conditions. The pellet is resuspended in 2mL of lithium-acetate (0.1M) and ditioteritol (2.5mM). The sample is incubated at room temperature for 10 to 15 minutes. The sample is centrifuged with the same conditions. The pellet is resuspended in 2mL of icecold sorbitol (1M). The sample is centrifuged with the same conditions. The pellet is finally resuspended in 250mL of ice-cold sorbitol (1M).

The first step for the transformation of cells is by adding 5µL of linearized DNA (100-500 ng for plasmid DNA, 1 µg for linear DNA) to 50µL of electrocompetent cells. The sample is incubated on ice for 2 minutes. The sample is transferred to an ice-cold 0.2cm electroporation cuvette and incubated on ice for 5 minutes. The cells are transformed using 1.5kV, 25µF and 200Ω. 1mL of ice-cold YPD-media is added immediately to the cuvette and mixed by pipetting. The cells are incubated for 1 hour at 30°C. The sample is transferred to a sterile tube and it's centrifuged at 4000rpm for 2

minutes. The supernatant is discarded and the pellet resuspended in 200µL of sorbitol (1M). The sample is spread on SD/CSM-Ura-plates. The cells are incubated at 30°C for 3 to 10 days until colonies are formed.

8. Transformation of DH5α cells

2µL of the DNA is added to 50µL of competent cells and mixed gently. The tubes are incubated on ice for 30min, and then introduced in a 42°C water bath for 20sec. The tubes are placed on ice for 2min.

950µL of LB -medium is added to the tube and the samples incubated at 37°C for 1h at 200rpm. After the incubation, 50µL, 150µL and 200µL are incubated in three plates with LB medium and ampicillin (0.1mg/mL). The spear volume is centrifuged for 5 minutes at 6000rpm and, after removing most of supernatant and ressuspending the pellet, the cells are also plated in LB plates. The plates are incubated at 37°C overnight.

9. DNA insert ligation into pGEM-T

The pGEM-T vector tube is briefly centrifuged to collect contents at the bottom and the 2x Raid Ligation Buffer was mixed vigorously. The regents in table 15 are added in an Eppendorf tube mix.

Table 15 Reagents of DNA insert ligation into pGEM-T

COMPONENTS	V(µL)
2x Rapid Ligation Buffer	5.0
pGEm-T (25ng)	0.5
T4 DNA Ligase (3 units)	1.0
xµL PCR product	*
MQ water	Adjust V till 10
	10

*The volume required of the PCR product is calculated using the following equation, considering their DNA concentration:

$$\frac{ng\ of\ vector\ x\ kb\ size\ of\ insert}{kb\ size\ of\ the\ vector}\ x\ insert{:}\ vector\ molar\ ratio = ng\ of\ insert$$

The pGEM-T Vector System is optimized for a 1:1 molar ratio, but also ratios of 3:1, 1:3, 8:1 and 1:8 have been used successfully.

The reactions were incubated overnight at 4°C our 1h at 22°C. The resulting reaction mixture can be used directly for transformation.

10. Gibson assembly

The Gibson reagents are displayed in table 16.

Table 16 Gibson reagents

COMPONENTS	V (µL)
Fragment	*
Backbone	*
Gibson Assembly Master Mix	10
MQ water	Adjust V till 20
	20

*The amount of fragment and backbone is calculated using the following formula: pmols = (weight in ng) x 1,000 / (base pairs x 650 daltons)

The reaction is taking place at 50°C during 1h (Gibson *et al.*, 2009).

Results

Synthesis of sophorolipids using unconventional substrates

For the production of sophorolipids, hydrophobic sources rich in fatty acids with 16 and 18 carbons are used. The reason is that the sophorolipids produced by *S. bombicola* usually contain 16 or 18 carbon atoms. Rapeseed oil, which is reach in oleic acid, is widely used as a hydrophobic source for the sophorolipid production by *S. bombicola* cultures.

To induce *S. bombicola* to produce sophorolipids with a branched hydrophobic tail, unconventional hydrophobic sources will be tested: Guerbet alcohols and diacyl glycerols. Theoretically, the sophorose molecule would be added at the position of the hydroxyl group, forming a sophorolipid with a branched hydrophobic tail.

The first strategy requires the use of three Guerbet alcohols: 2-hexyl-1-decanol (Euthanol G16), 2-octyl-1-dodecanol (Euthanol G20) and 2-decyl-1-tetradecnol (Euthanol G24) (figure 12). Guerbet alcohols are branched molecules, originating from the dimerization of two alcohols with the formation of water. They have an hydroxyl group in the middle of the chain (Sulzbacher, 2007). The second strategy relies on the use of diacyl glycerols (figure 13): 1,2-dioctanoyl-sn-glycerol (C8:0 glyceride dicaprylin), 1,2-didecanoyl-sn-glycerol (C10:0 glyceride dicaprin), 1,2-dimyristoyl-sn-glycerol (C14:0 glyceride dimyristin) and 1,2-dioleoyl-sn-glycerol (C18:1 glyceride diolein). Diacyl glycerols are derivates of glycerol, in which the hydroxyl groups were substituted with fatty acids (Goñi & Alonso, 1999). These are interesting molecules since they have a free hydroxyl group.

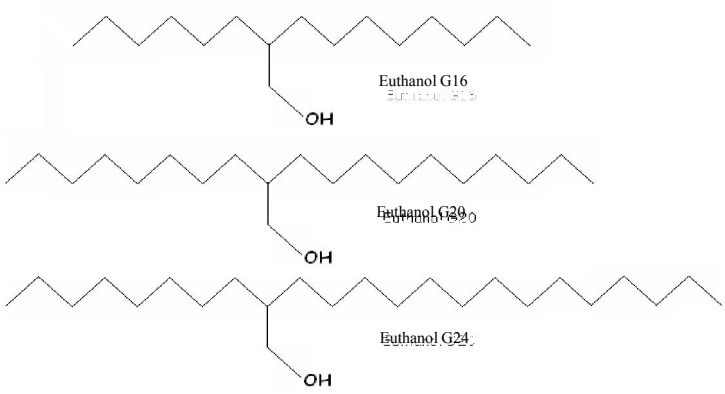


Figure 12 Guerbet alcohols

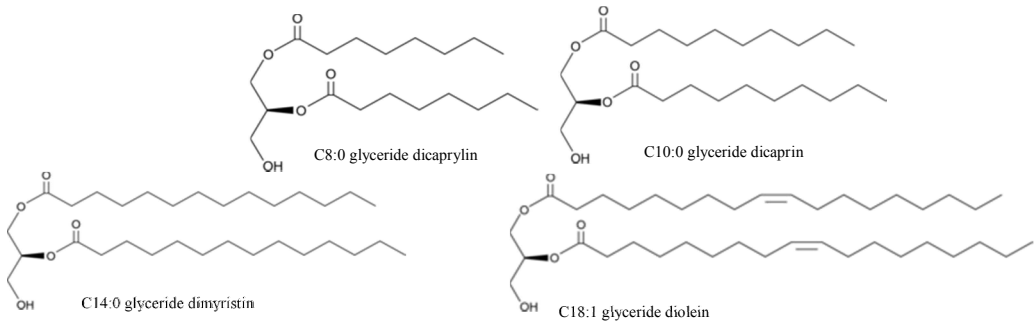


Figure 13 Diacyl glycerols

1. Production of sophorolipids on Guerbet alcohols

Two yeast strains were used: a wild type (WT) and a CYP52M1 knock-out (ΔCyp) strain. The wild-type culture will perform *de novo* synthesis of the fatty acids which will lead to the production of the classic sophorolipids produced by the yeast *S. bombicola*. By adding the referred substrates with a hydroxyl group in the middle of the chain, it's expected that the sophorose molecules will be glycosidically added at this position of the fatty acid. Therefore, wild-type cultures are also expected to produce sophorolipids with a branched hydrophobic tail.

The CYP52M1 knock-out culture will not produce the classic sophorolipids molecules, although *de novo* synthesis of the fatty acids still occurs. This is because the Cyp gene, that codes for the cytochrome P450 monooxygenase enzyme CYP52M1, is knocked-out. This enzyme is responsible for the hydroxylation of the terminal (ω) or the subterminal (ω-1) position of the fatty acid. Since the CYP52M1 enzyme is suppressed, the hydroxylation on the therminal or subterminal of the fatty acid produced by the cell will not occur and the culture will not produce the classic sophorolipid molecules. However, if the Guerbet alcohols are incorporated in the sophorolipid production, sophorolipids with a branched hydrophobic tail will be produced by the CYP52M1 knock-out culture.

1.1 Guerbet alcohol G16

1.1.1 Growth curve

Figure 14 depicts the growth curve of the yeast cultures wild-type and ΔCyp with G16 as substrate.

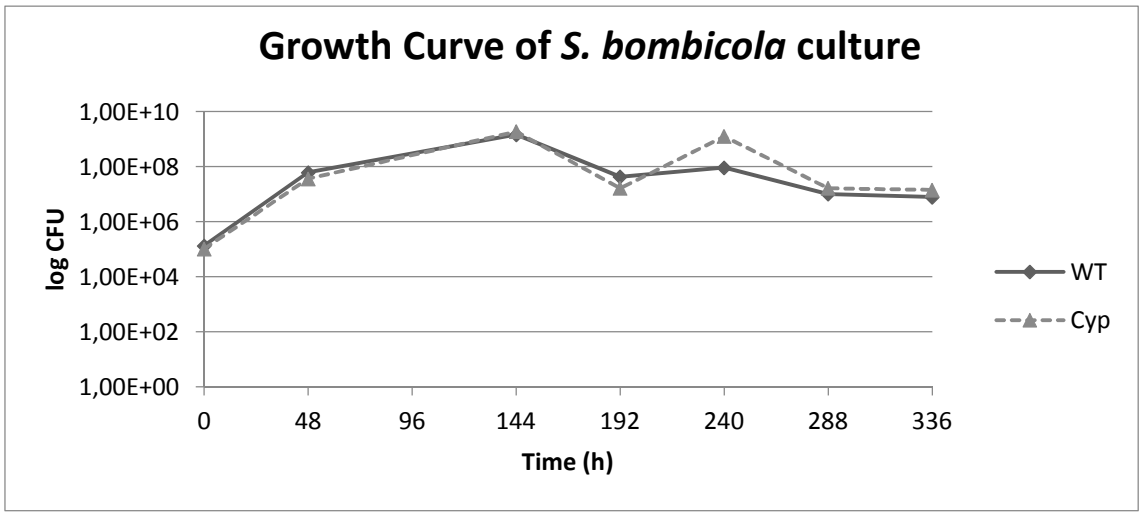


Figure 14 Growth curve of *Starmerella bombicola* cultures with G16 as substrate

The growth curve of both cultures shows a similar pattern, increasing the CFU until 150h, decreasing due to lack of nutrients, and increasing again due to the administration of glucose at 200h.

1.1.2 Glucose consumption

Figure 15 depicts the glucose consumption of the wild-type and ΔCyp with G16 as substrate.

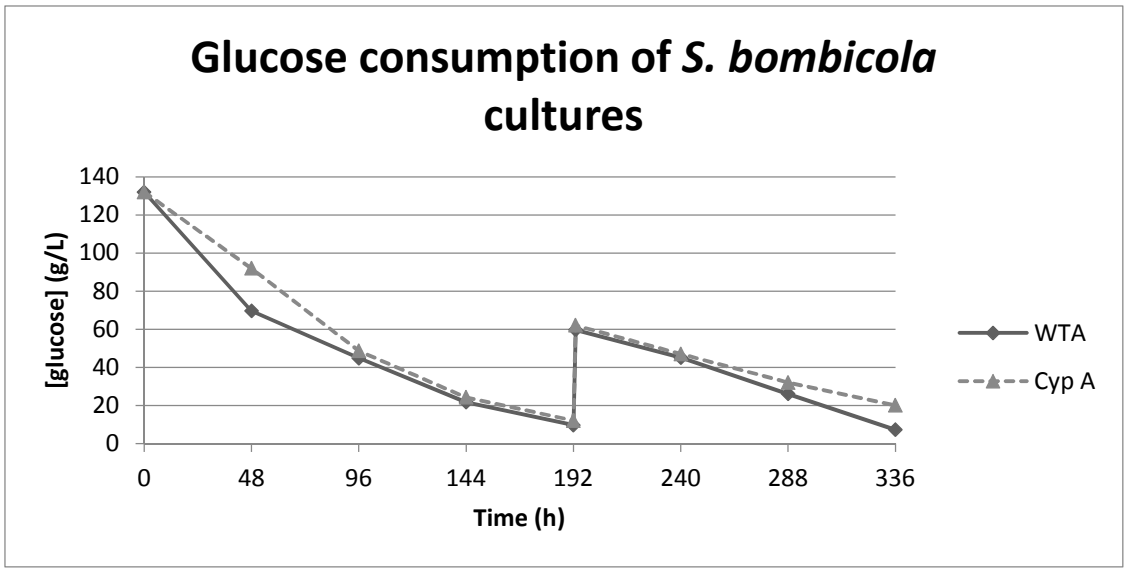


Figure 15 Glucose consumption of *Starmerella bombicola* cultures with G16 as substrate

The glucose consumption of both cultures shows a similar pattern, which is consistent with the growth curves in figure 14.

1.1.3 HPLC analysis

The following graphics, obtained through HPLC analysis, depict the production of sophorolipids during the WT culture growth.

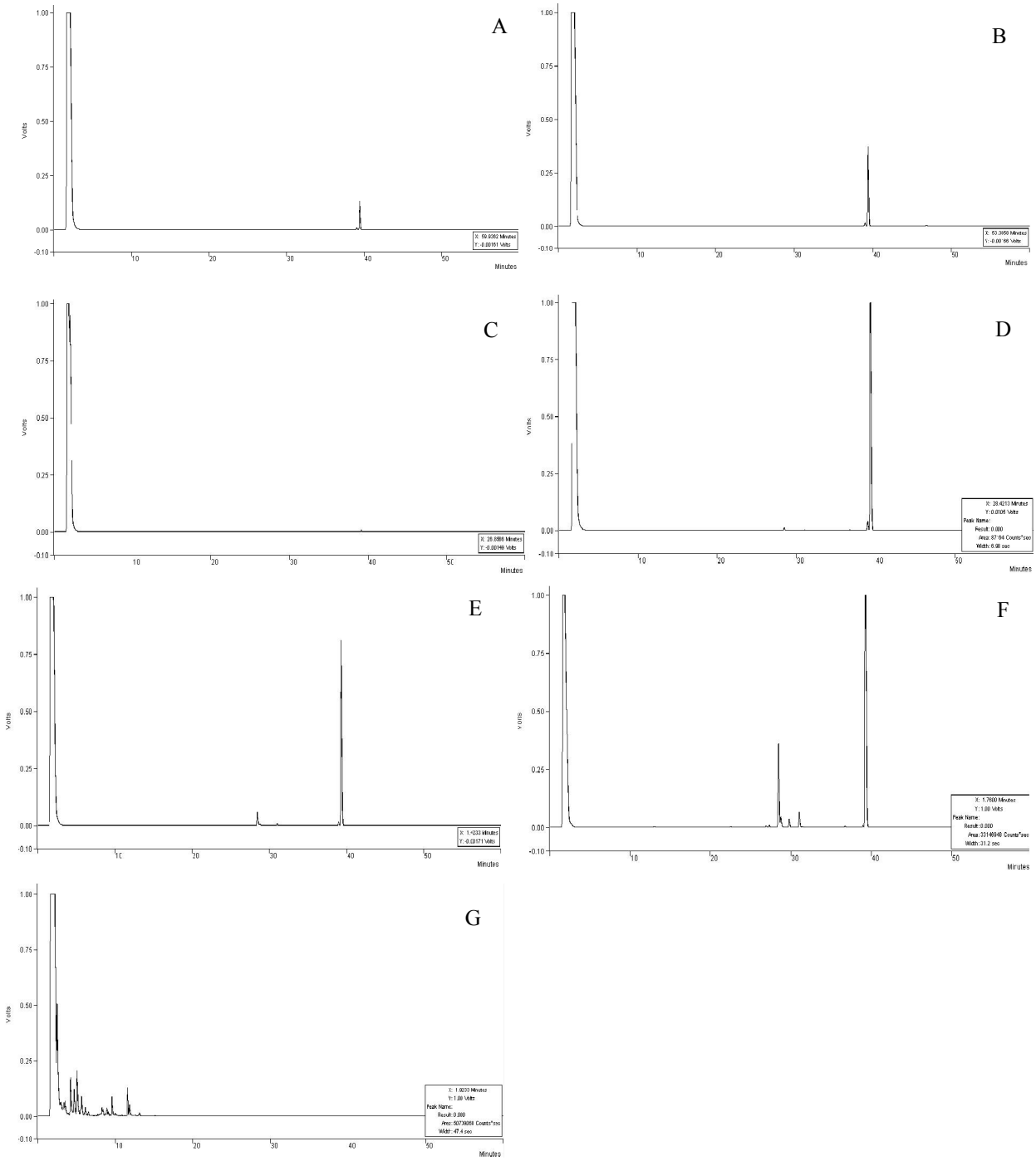


Figure 16 Analysis of the production of sophorolipids by WT at (A) 48h, (B) 96h, (C) 144h, (D) 192h, (E) 240h, (F) 288 and (G) 336h.

Through the analysis of the graphics, it can be deduced that the WT culture only produced the classic sophorolipid molecules. These sophorolipids corresponds to the peaks with an approximately elution time of 28 minutes in graphic F. The substrate G16

corresponds to the peak with 40 minutes elution time. These results indicate that the substrate was not incorporated by the yeast in the sophorolipids production. The peaks present in graph G may correspond to the degradation of the substrate after the final extraction.

There was no production of sophorolipids by the ΔCyp culture, since no peaks appeared on the graphs, besides the one corresponding to the substrate.

1.2 Guerbet alcohol G20

1.2.1 Growth curve

Figure 17 depicts the growth curve of the wild-type and ΔCyp with G20 as substrate.

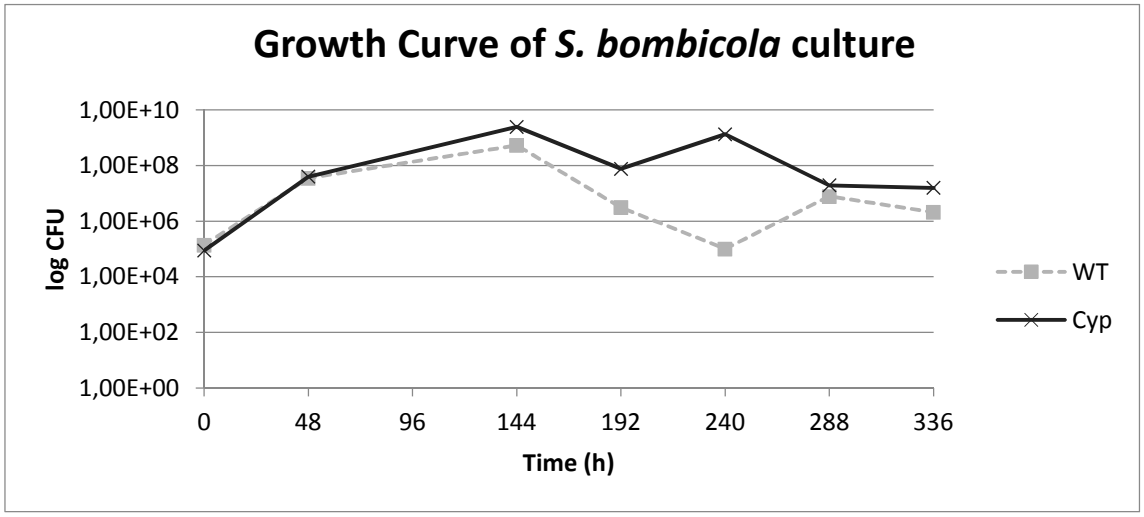


Figure 17 Growth curve of Starmerella bombicola cultures with G20 as substrate

The growth curve of both cultures shows a similar pattern, till 150h. After this, there is a decrease probably due to lack of one or more of the nutrients or a sampling error. There is an increase again due to the administration of glucose at 192h. However, the growth of culture ΔCyp is delayed, compared with the WT, after the addition of glucose.

1.2.2 Glucose consumption

Figure 18 depicts the glucose consumption of the yeast cultures wild-type and ΔCyp with G20 as substrate.

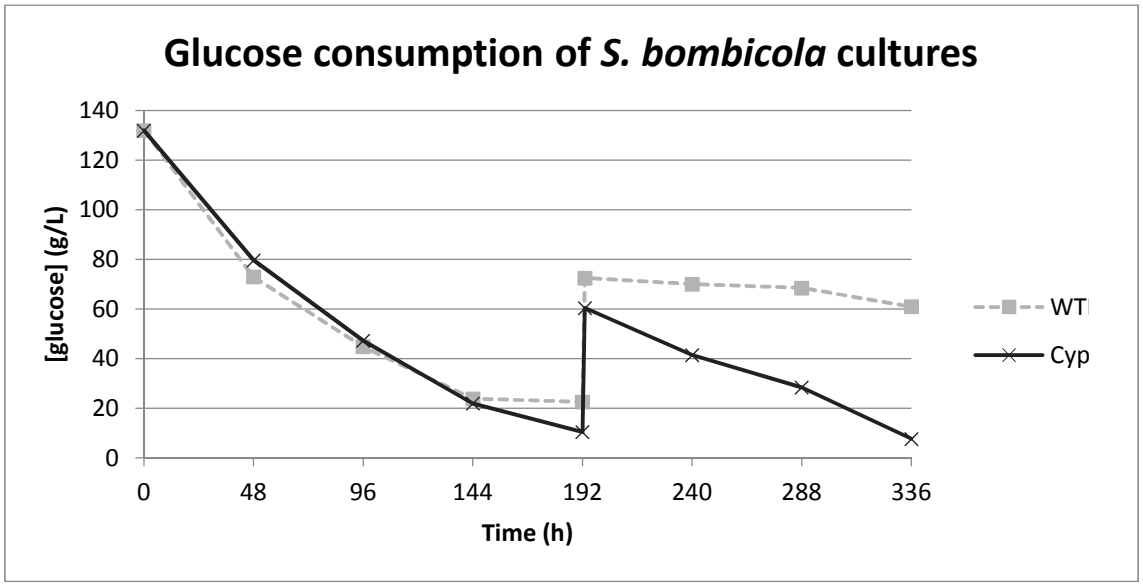
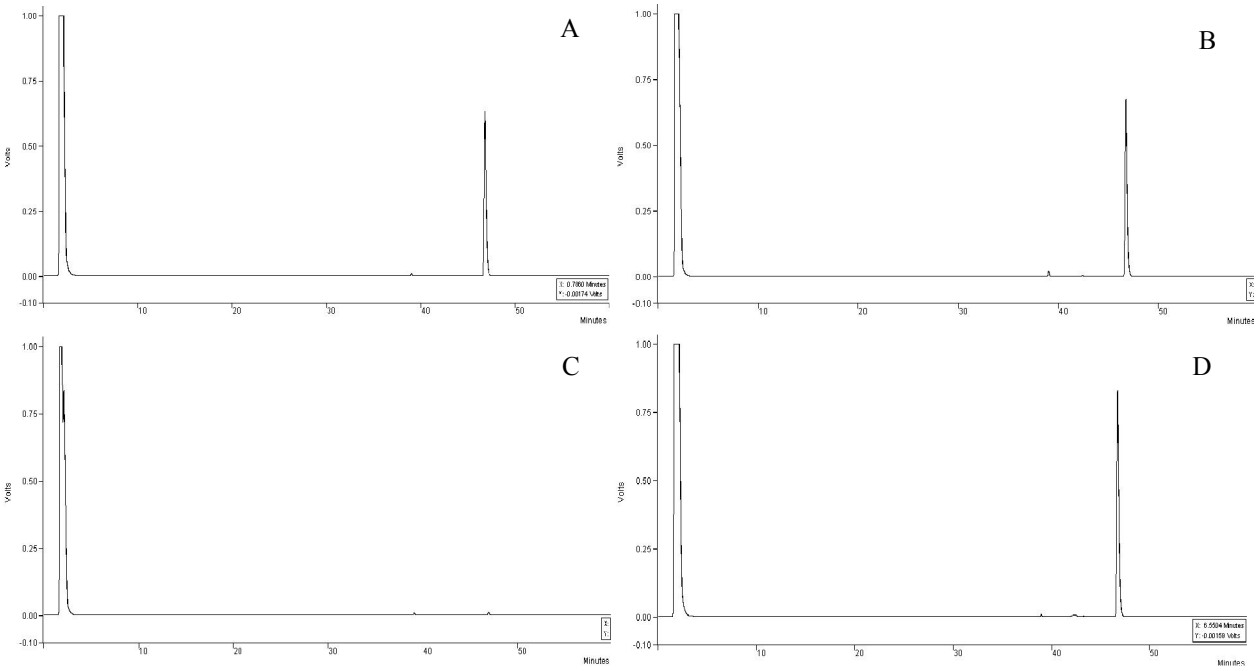


Figure 18 Glucose consumption of *Starmerella bombicola* cultures with G20 as substrate

The glucose consumption of both cultures shows a similar pattern, which is consistent with the growth curves in figure 17.

1.2.3 HPLC analysis

The following graphics, obtained through HPLC analysis, depict the production of sophorolipids during the WT culture growth.



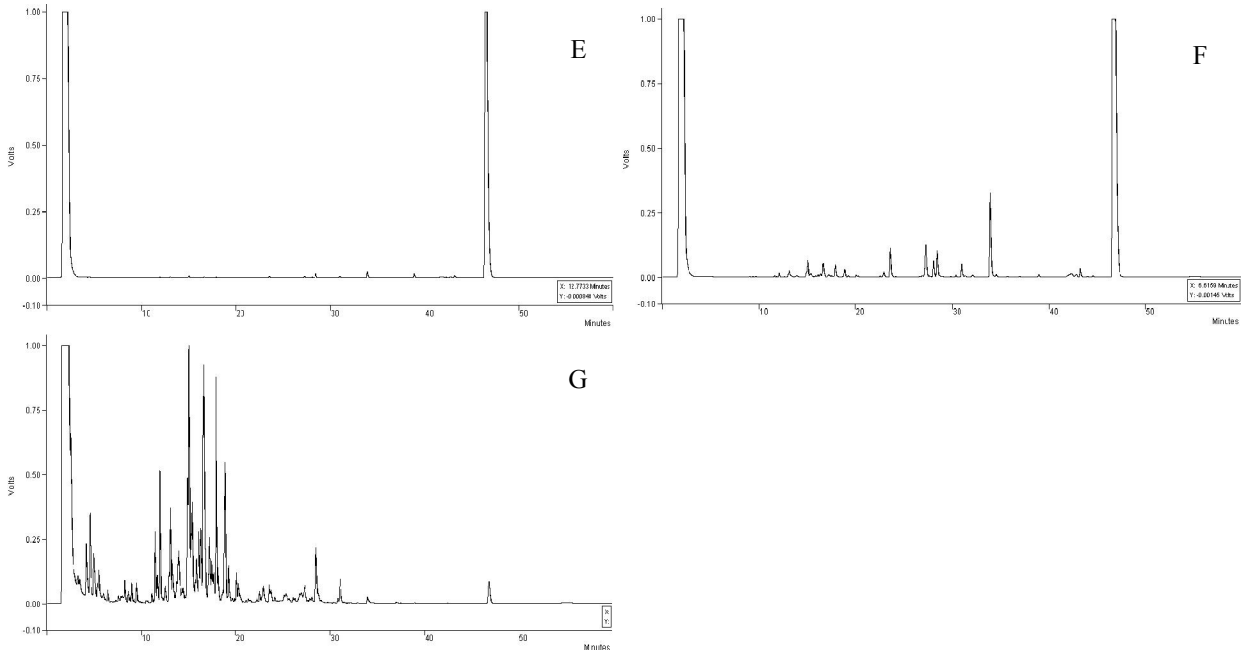


Figure 19 Analysis of the production of sophorolipids by WT at (A) 48h, (B) 96h, (C) 144h, (D) 192h, (E) 240h, (F) 288 and (G) 336h.

The HPLC analysis of the WT showed that the *de novo* synthesis of sophorolipids occurred (elution time near 27 minutes in graphic E), and also the synthesis of possible new molecules with an elution time near 32 minutes in graphic E and between 10 and 15 minutes in graphic. The peak with 45 minutes of elution correspond to the substrate. The peaks present in graph G may correspond to the degradation of the substrate after the final extraction. Graph C shows no peaks, result which could have be due to sampling errors. There was no production of sophorolipids by the ΔCyp culture.

1.3 Guerbet alcohol G24

1.3.1 Growth curve

Figure 20 depicts the growth curve of the yeast cultures wild-type and ΔCyp with G24 as substrate.

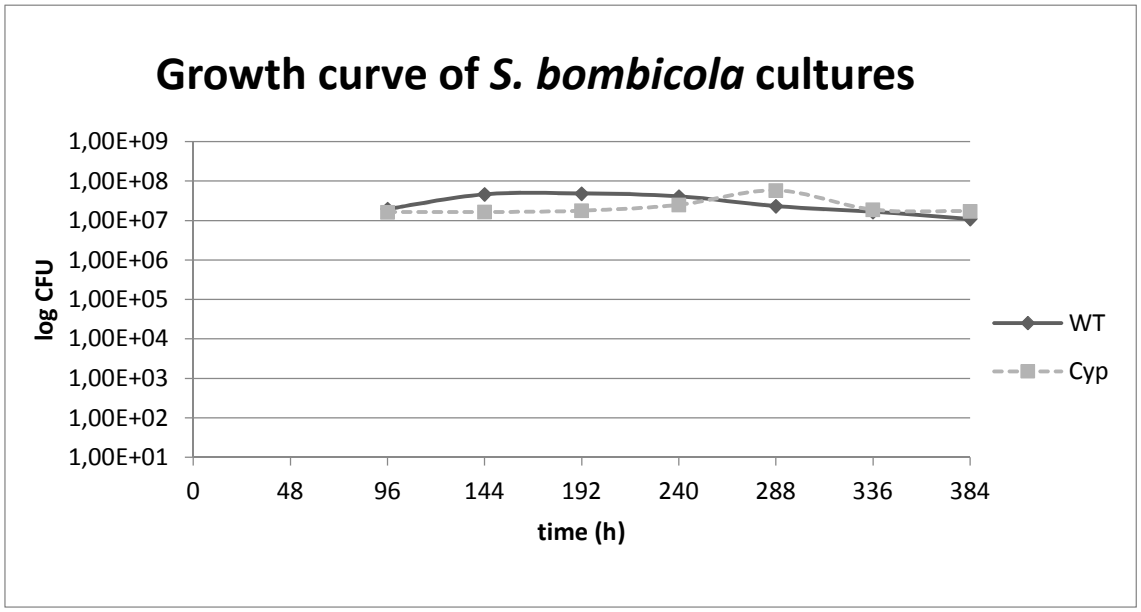


Figure 20 Growth curve of *Starmerella bombicola* cultures with G24 as substrate

The growth curve of both cultures show a slightly different pattern, which can be due to the production of new sophorolipid molecules by the WT culture. There is no information of the CFU counting at 48h due to errors committed during the dilution of the samples.

1.3.2 Glucose consumption

Figure 21 depicts the glucose consumption of the yeast cultures wild-type and ΔCyp with G24 as substrate.

The glucose consumption of both cultures show similar patterns, however the WT culture consumes more glucose after 144h.

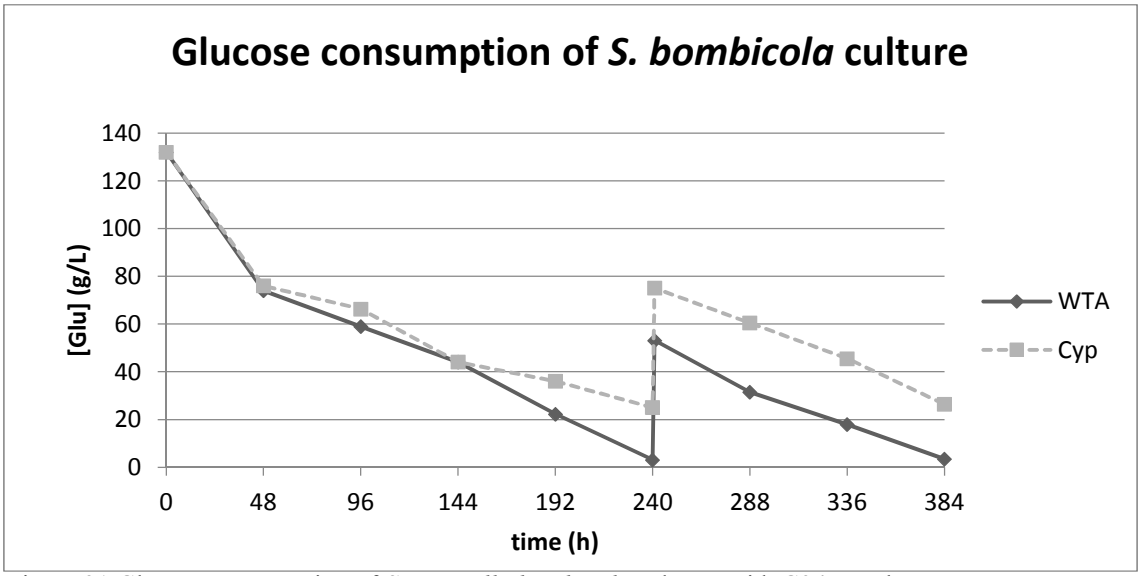


Figure 21 Glucose consumption of *Starmerella bombicola* cultures with G24 as substrate

1.3.3 HPLC analysis

The following graphics, obtained through HPLC analysis, depict the production of sophorolipids during the WT culture growth.

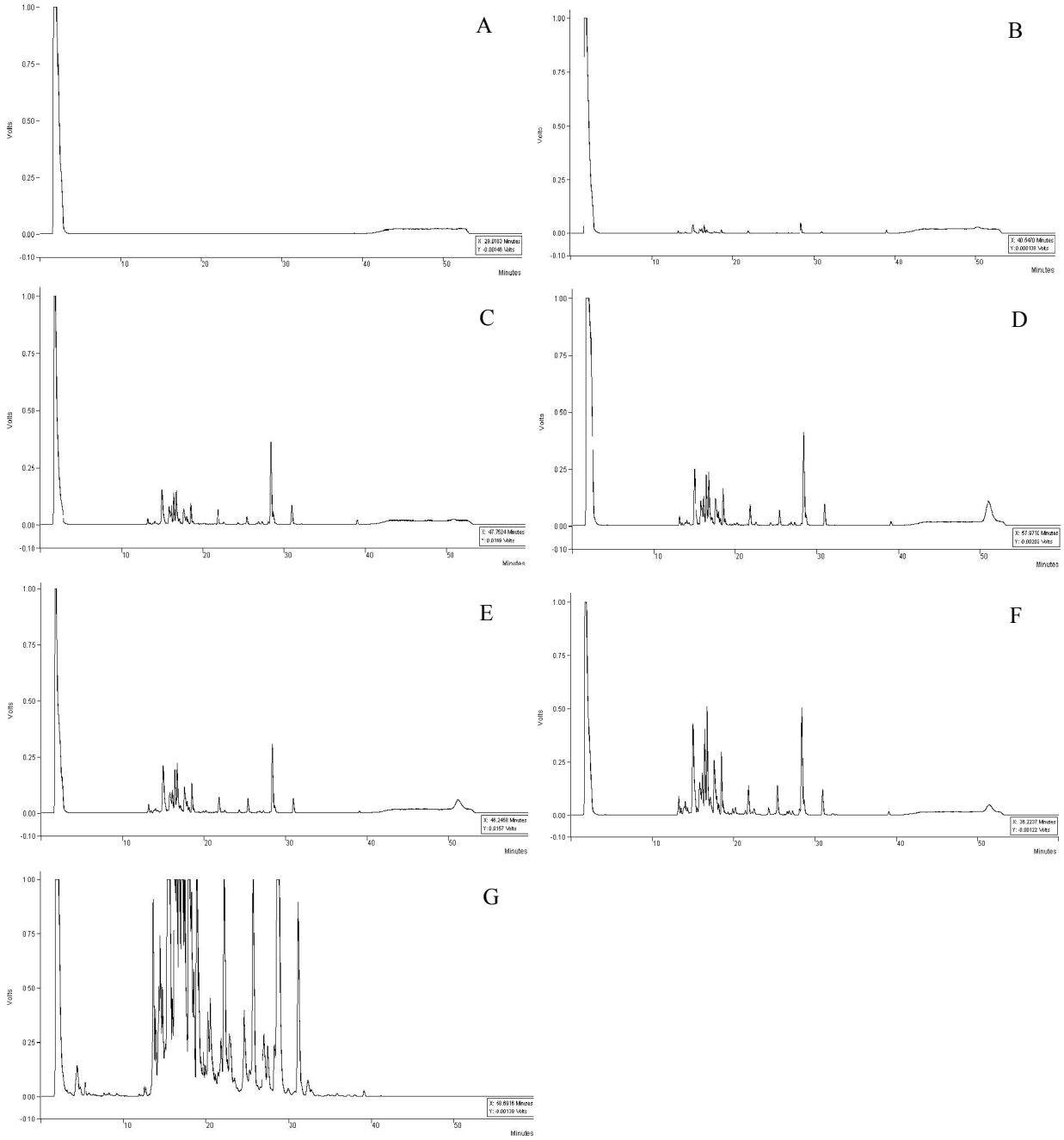


Figure 22 Analysis of the production of sophorolipids by WT at (A) 144h, (B) 192h, (C) 240h, (D) 288h, (H) 336h, (I) 384h and final extraction at 384h (G).

The HPLC analysis of the WT shows that the *de novo* synthesis of sophorolipids occurred (elution time of 28.3, 29.0 and 30.9 minutes from graphic B to F), and also the synthesis of possible new sophorolipid molecules. These sophorolipids have an elution between 13.1 and 27.2 minutes form graphic B to G. These results indicate that the G24

substrate (elution time 51.0 minutes) may have been incorporated in the sophorolipids, allowing the production of new molecules, potentially sophorolipids with branched hydrophobic tail. During the development of the ΔCyp cultures, there was also no production of sophorolipids.

1.3.4 LC-MS analysis

Since the WT culture showed promising production of new sophorolipids molecules, the sample from 384h final extraction was analyzed by LC-MS. The results are shown in figure 23.

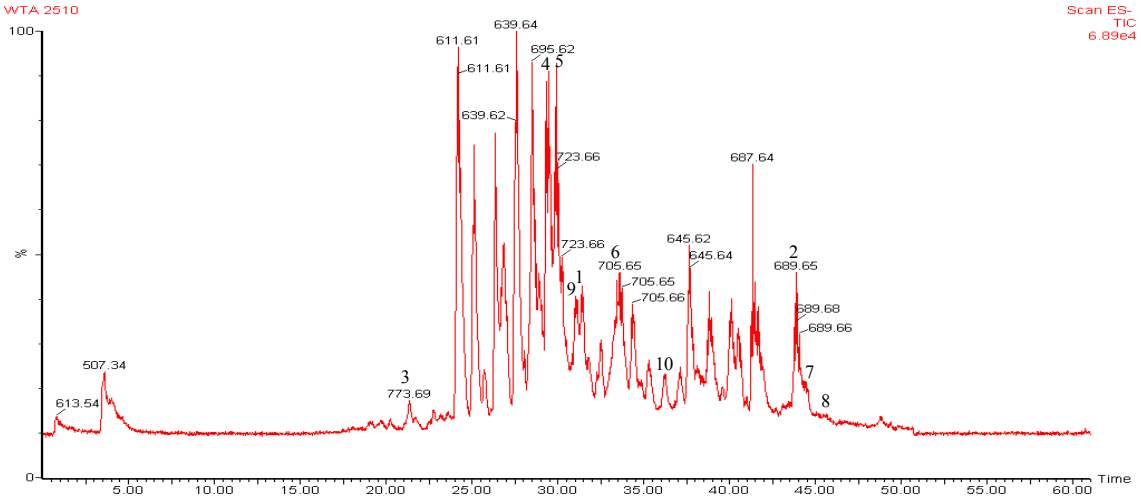


Figure 23 Mass spectrometry chromatogram of the WT sample colected at 384h. The numbers 1 to 10 represent the 10 peaks that correspond to new-to-nature sophorolipids produced.

A specific molecular mass scan was performed on the obtained chromatogram. The following new sophorolipid molecules were identified.

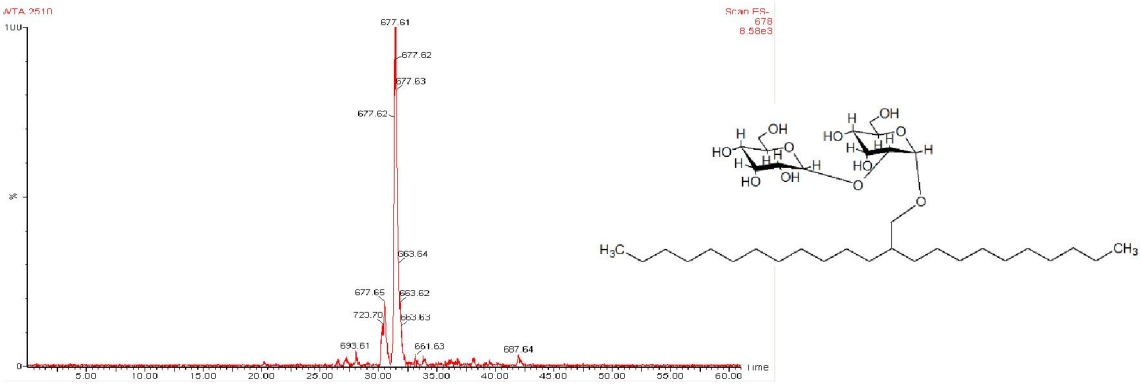


Figure 24 LC-MS peak and illustration of a sophorolipid molecule with a branched hydrophobic tail with a molecular mass of 678m/z. This peak is the number 1 in figure 23.

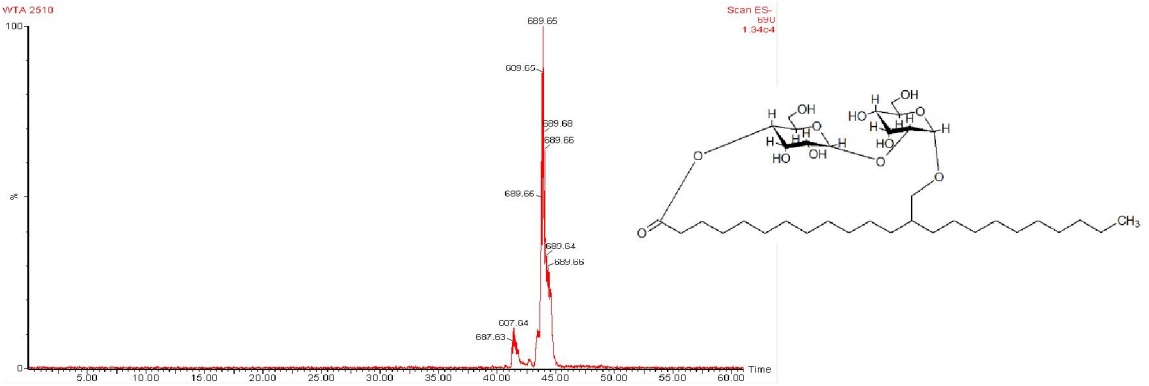


Figure 25 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 690m/z. This peak is the number 2 in figure 23.

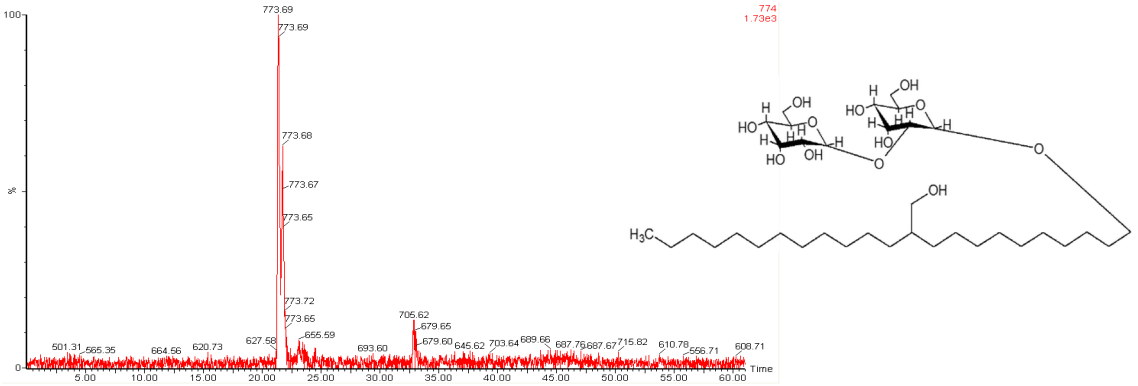


Figure 26 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 774m/z. This peak is the number 3 in figure 23.

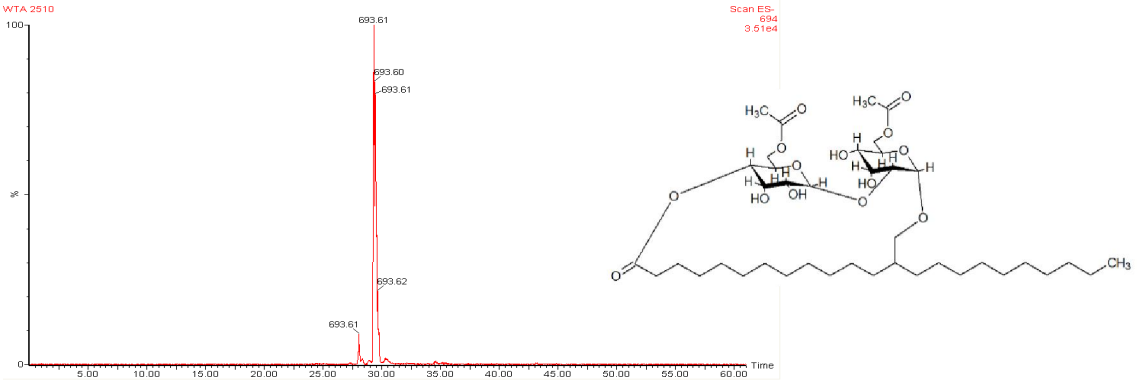


Figure 27 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 694m/z. This peak is the number 4 in figure 23.

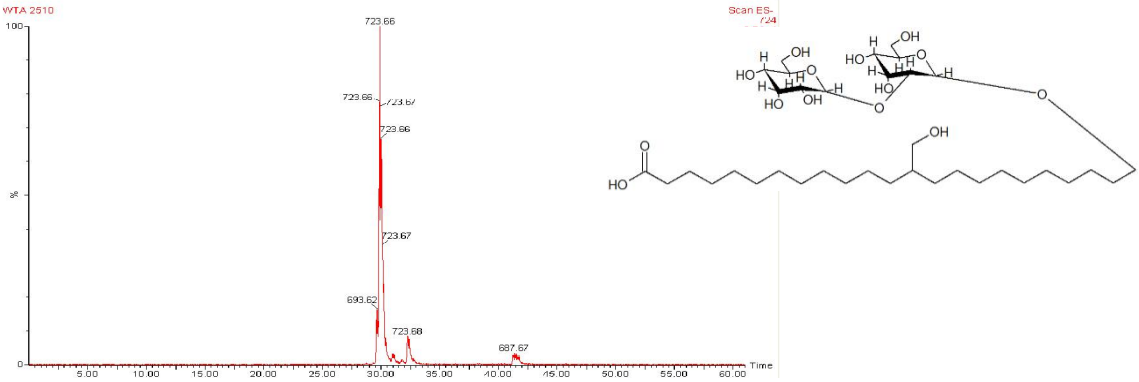


Figure 28 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 724m/z. This peak is the number 5 in figure 23.

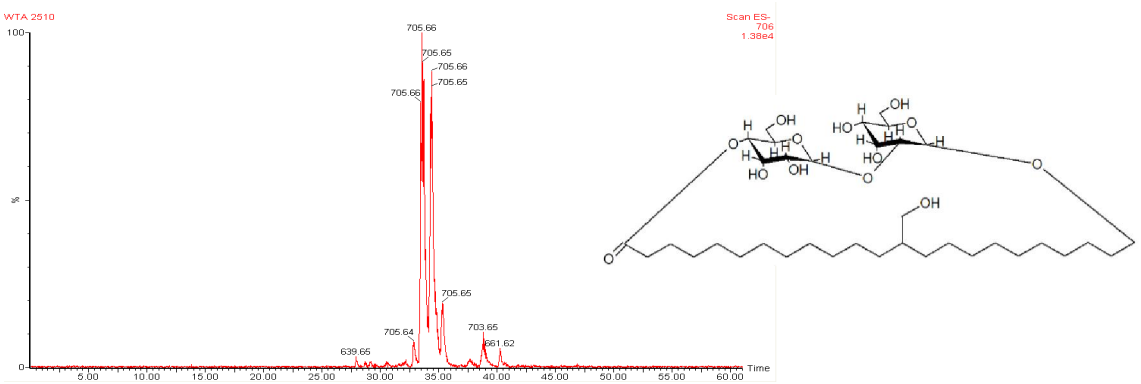


Figure 29 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 706m/z. This peak is the number 6 in figure 23.

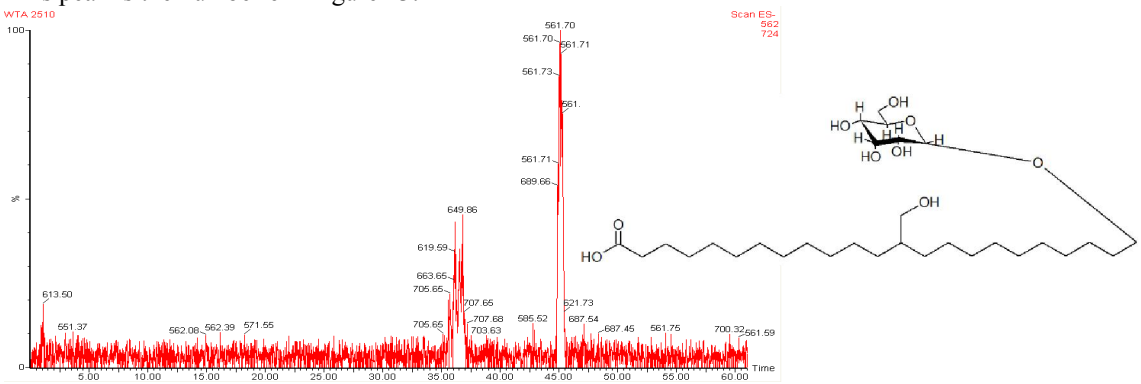


Figure 30 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 562m/z. This peak is the number 7 in figure 23.

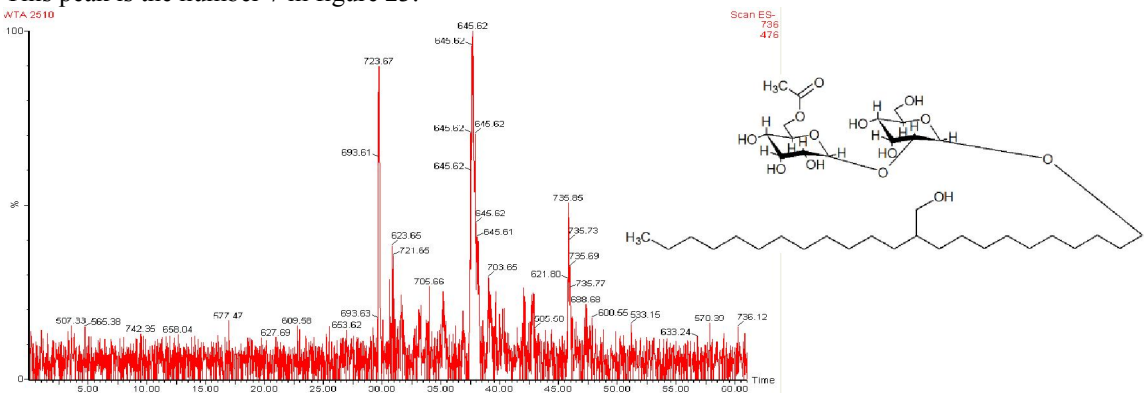


Figure 31 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 736m/z. This peak is the number 8 in figure 23.

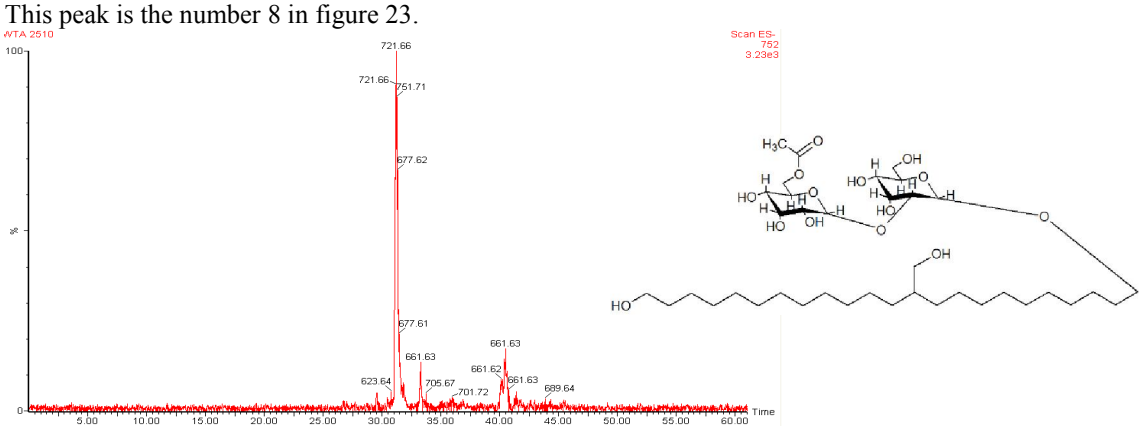


Figure 32 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 752m/z. This peak is the number 9 in figure 23.

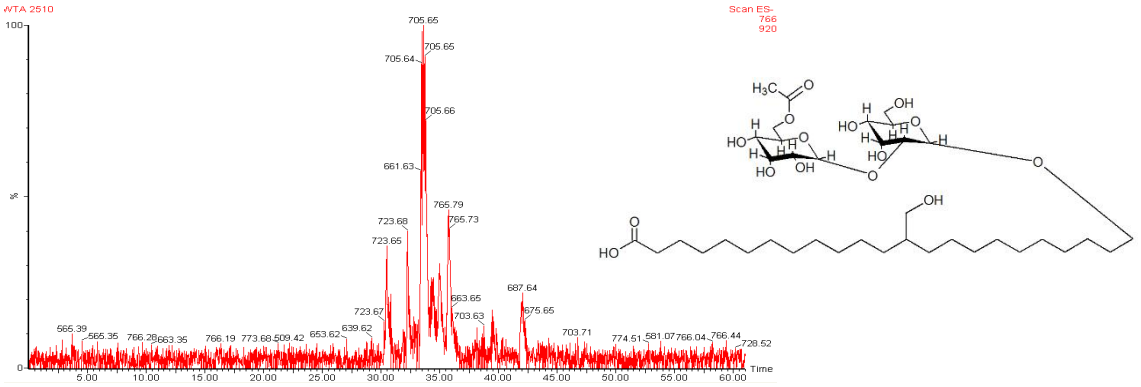


Figure 33 LC-MS peak and illustration of a new-to-nature sophorolipid with a molecular mass of 766m/z. This peak is the number 10 in figure 23.

There is a strong suggestion that ten new-to-nature sophorolipid molecules were produced, being only one molecule a sophorolipid with direct incorporation of the substrate. This molecule (figure 24) has an in-chain hydroxylation and no acetylation or lactonization. Two similar molecules were identified, one with a lactonization (figure 25) and a lactonization and two acetylation (figure 26). Lactonic molecules were not expected, since the substrate has no carboxyl group at the end of the chain. However enzymes present in the cell must have introduced the carboxyl group, allowing the formation of lactonic forms.

As expected from the results obtained from the ΔCyp cultures, the glucosyltransferases (Ugt) appear to be more specific for the hydroxylation at the end of the hydrophobic chain. These explains the molecules represented in figures 27 to 33. As referred previously, the UgtA1 and UgtB1 enzymes are responsible for the glycosylation of the first glucose unit to the hydroxyl group of the fatty acid, and the glycosylation of the second glucose unit to the first one, respectively (Saerens *et al.*, 2011a,b).

Although the suggested molecules have the correct molecular mass, some masses may correspond to more than one molecule. The molecular masses in figures 27 to 33 may also be explained by molecules with in-chain hydroxylation and a glycosylation at the terminal or sub-terminal of the chain. This indicates that other sophorolipid with a branched hydrophobic tail may have been produced by the culture. Molecule 3 (figure 26) may not correspond to the attributed peak, since it has a low retention time for a lactonic molecule. Molecules 7 (figure 30), 8 (figure 31), and 10 (figure 33) have a low abundance, indicating that the molecules may not correspond to the assessed molecular masses. The identity of some of the molecules from the non-identified peaks is unknown. Some peaks correspond to the sophorolipids produced by

de novo synthesis, and others can be explained by molecular adduct ions for sophorolipid, as Na+. Some molecular masses may also correspond to molecular structures which are not expected, as degraded sophorolipid molecules.

1.3.5 Final sophorolipid analysis

A final extraction of the sophorolipids was performed on the culture broth. The yield of the sophorolipid production was 16.5g/L, which is an acceptable value for flasks cultures.

This work leaded to the production of a branched sophorolipid molecule with two glucoses with a fatty acid of 24 carbons, being the first work to do so, using *S. bombicola* and unconventional substrates. The production yield obtained for this molecules was 0.21g/L (1.285% of the total production). This calculation was made by taking into account the peaks areas from the LC-MS graph.

2. Production of sophorolipids using diacyl glycerols as substrates

The following figures depict the HPLC results from the enzyme test performed using diacyl glycerols. As referred previously, diacyl glycerols have a free hydroxyl group, so a position where the sophorose molecule can be added, possibly leading to the production of sophorolipids with a branched hydrophobic tail.

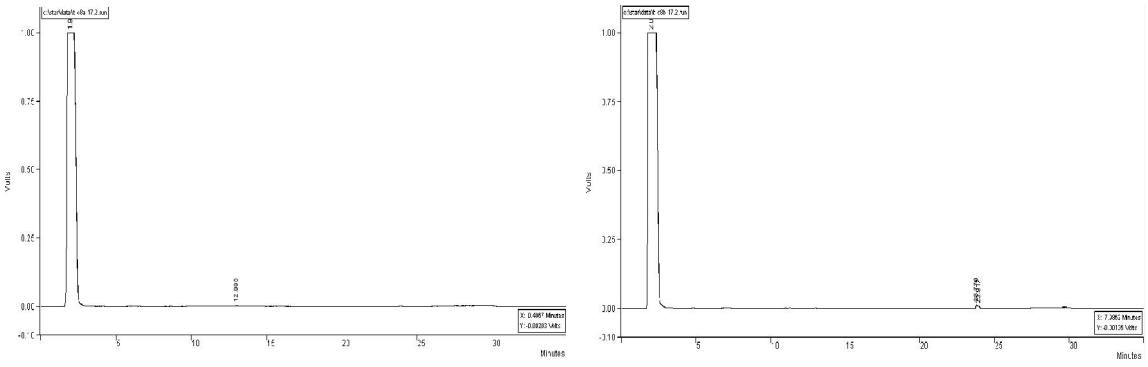


Figure 34 Analysis of the production of sophorolipids using C8:0 glyceride dicaprylin as substrate. The samples analyzed are replicas.

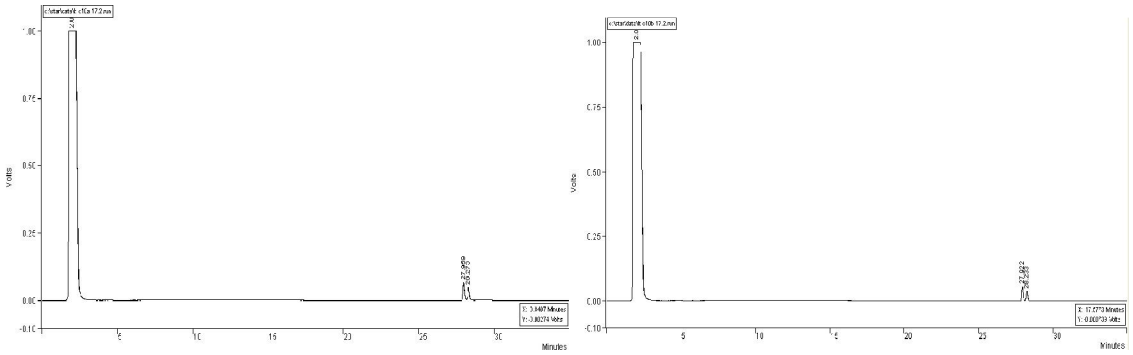


Figure 35 Analysis of the production of sophorolipids using C10:0 glyceride dicaprin as substrate. The samples analyzed are replicas.

By using C8:0 glyceride dicaprylin and C10:0 glyceride dicaprin as substrates, there is a minor production of molecules with the elution time of 13, 23.8 and 23.9 (figure 34) and 27.9 and 28.2 (figure 35). For the other two substrates no molecules were detected. This may be because there is no production of sophorolipids by using the referred substrates.

In conclusion, diacyl glycerols are not the ideal substrates for the yeast *S. bombicola* to produce sophorolipids with a branched hydrophobic tail.

Heterologous expression of genes performing in-chain hydroxylation of a fatty acid

The first step of sophorolipid biosynthesis is the hydroxylation on the terminal or the subterminal position of a fatty acid. This reaction is mediated by the cytochrome P450 monooxygenase enzyme CYP52M1.

In order to produce sophorolipids with a branched hydrophobic tail, the CYP52M1 gene of *S. bombicola* is substituted with a gene from another microorganism, *Elizabethkingia meningoseptica* or *Rhodococcus rhodochrous*. The referred genes encode for an oleate hydratase which takes care of in-chain hydroxylation of a fatty acid, preferentially oleic acid.

Elizabethkingia meningoseptica is a gram negative bacteria ubiquitous in many environments. It is associated with meningitis in infants (Kim *et al.*, 2005). The enzyme of *Elizabethkingia meningoseptica* is already described and the sequence identified. However, since its codon usage is not compatible with the *S. bombicola*, the oleate hydratase gene was ordered. This bacteria produces R-configuration products, which can constitute a problem because in the native sophorolipid production only the S-enantiomers get incorporated. The R-enantiomers of the subterminal hydroxylated fatty acids don't get incorporated. Since it's not known if the same happens for the in-chain hydroxylation, the same gene of *Rhodococcus rhodochrous* was isolated. *Rhodococcus rhodochrous* is a gram positive bacteria, which is usually present in contaminated soils and it's known to produce a mix of both enantiomers. Some strains of this bacteria are pathogenic to plants and mammals. The genome of these bacteria is not fully known (Larkin *et al.*, 2010), including the oleate hydratase gene.

The goal of this strategy is to construct a DNA cassette, containing the oleate hydratase gene, which is going to be used to transform *S. bombicola* cells. The vector used (CYP52M1trunc) is a derivate of the cloning vector pGEM-T (figure 29). Besides the lactose operon and the gene for resistance to ampicillin (selection in *E. coli* cells), it contains the URA3 gene for selection in *S. bombicola* cells. It contains the promoter and terminator for the gene CYP52M1, the Ura3 gene and respective promoter and terminator. This vector also contains homologous regions, with the gene that encodes for the enzyme CYP52M1, which flank the region where the oleate hydratase gene is introduced. By homology, the oleate hydratase gene is introduced in the middle of the gene that encodes for the enzyme CYP52M1 in *S. bombicola*, knocking-out its

expression. This way, the oleate hydratase performs the in-chain hydroxylation of a fatty acid, leading to the production of sophorolipids with a branched hydrophobic tail.

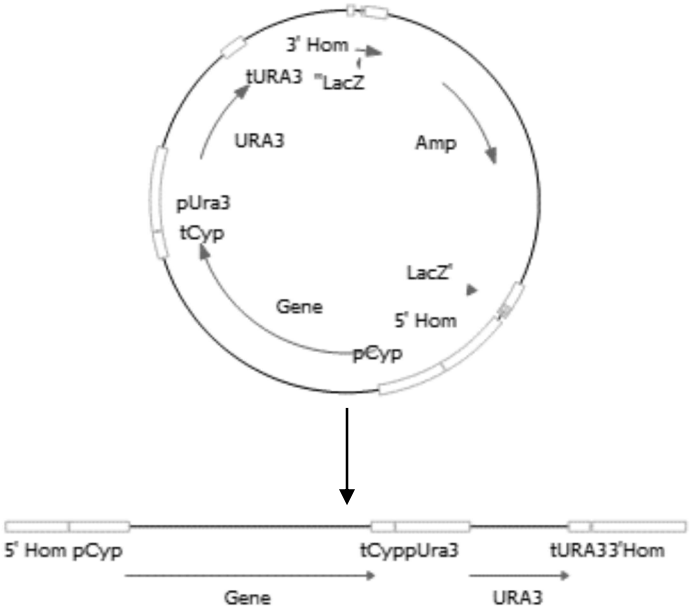


Figure 36 Transformation vector and transformation cassete for S. bombicolla. The vector is constituted by the Cyp promoter and terminator (pCyp and tCyp, respectevely), the Ura3 gene and respective promoter and terminator (pUra3 and tUra3), 5' and 3'

1. *Elizabethkingia meningoseptica* gene

The *Elizabethkingia meningoseptica* gene that encodes for the oleate hydratase protein is known. However, the codon usage of this bacteria is not compatible with *S. bombicola*. For this reason, the referred gene was ordered and ligated into CYP52M1trunc, using Gibson assembly. The DNA cassette, containing the homology regions to the Cyp gene, and the Ura3 gene, is used to transform *S. bombicola* cells.

The first step was performing a high fidelity PCR using the ordered *Elizabethkingia meningoseptica* gene. The primers used have the following sequence:5'-CTCCGTTTATCATATTGAAAAGTACATATATGAATCCTATCACC TCCAA-3' as forward primer and 5'-TCCATGGGAGGCTAAGAAACTTA ACCTCTGATACCT TTGAC-3' as reverse primer. Figure 37 depicts the result.

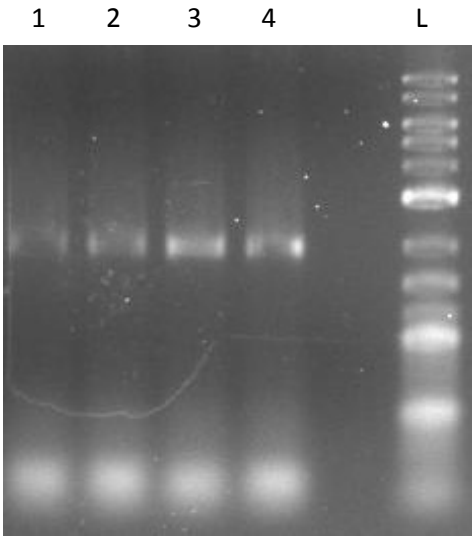


Figure 37 Electrophoresis of high fidelity PCR of *Elizabethkingia menigoseptica* gene. The first four lines correspond to the four PCR replicas of the gene of interest (1, 2, 3, 4), and the last line corresponds to the mass ladder (L).

The aim of the high fidelity PCR was to guarantee overlaping regions between the *Elizabethkingia meningoseptica* gene and the vector CYP52M1trunc for a Gibson reaction. For the same reason, a high fidelity PCR was also performed for the referred vector using the following primers: 5'-GTTTCTTAGCCTCCCATGGAAG-3' as forward primer and 5'-ATATGTACTTTTCAATATGATAAACGGAGAAATAACG-3' as reverse primer.

The Gibson reaction was used to transform DH5α cells and a colony PCR was performed using the following primers: 5'-GAGCTGCAAATTTTTAAAA CTGCAGTTAAGAAGCTAATTCACTAATTG-3' as forward and 5'-AATC CCATAAACGACTACTC-'3 as reverse. After extensive work, one positive colony was identified, but the sequence had four point mutation. A Sanchis reaction was performed in order to correct these mutations, but no successful results were obtained. For this reason, the original *Elizabethkingia meningoseptica* gene was introduced in a pJET vector and the construct used to transform DH5α cells. A colony PCR was performed with the following primers: 5'-CGACTCACTATAGGGAGAGCGGC-3' as forward primer and 5'-AAGAACATCGATTTTCCATGGCAG-3' as reverse. Colonies which showed the correct size (2060bp) were send to sequencing.

Following the procedure mentioned previously, colonies that showed the presence of a 1332bp fragment were send to sequencing. The result of the colony PCR is depicted in figure 38.

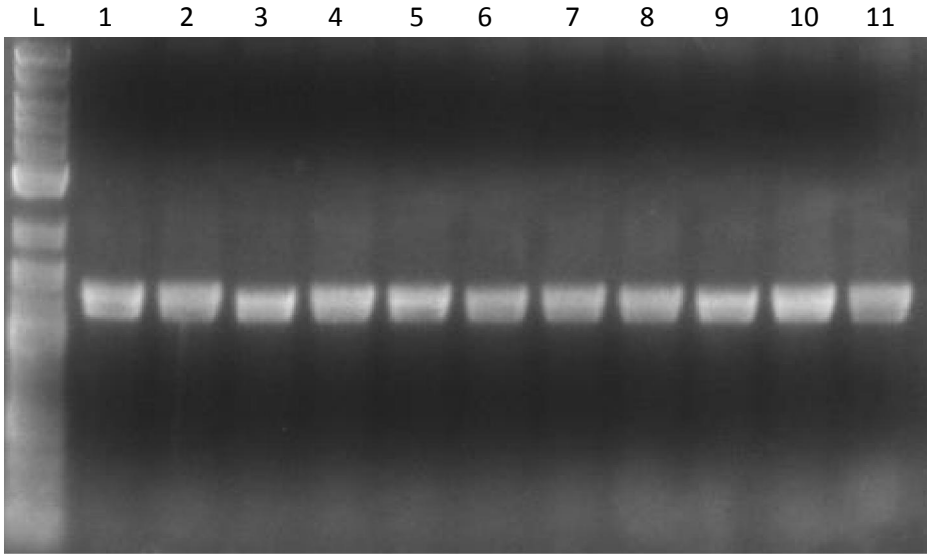


Figure 38 Electrophoresis of the colony PCR of DH5α cells containing the *Elizabethkingia menigoseptica* gene and the vector cyp52M1trunc. The numbers correspond to the 11 colonies analyzed and the L corresponds to the ladder.

A plasmid extraction was performed for 9 and 10 and the samples were sent for sequencing. A high-fidelity PCR was performed using the referred samples, with the intent of isolating cassette for transformation in *S. bombicola* cells. This PCR was performed with the following primers: 5'-CGTTGTCAAG TCCTAAGGTAT-3' as forward and 5'-AAGCGTGAAGCTCCTCTGACAATC-3'.The results obtained are depicted in figure 39.

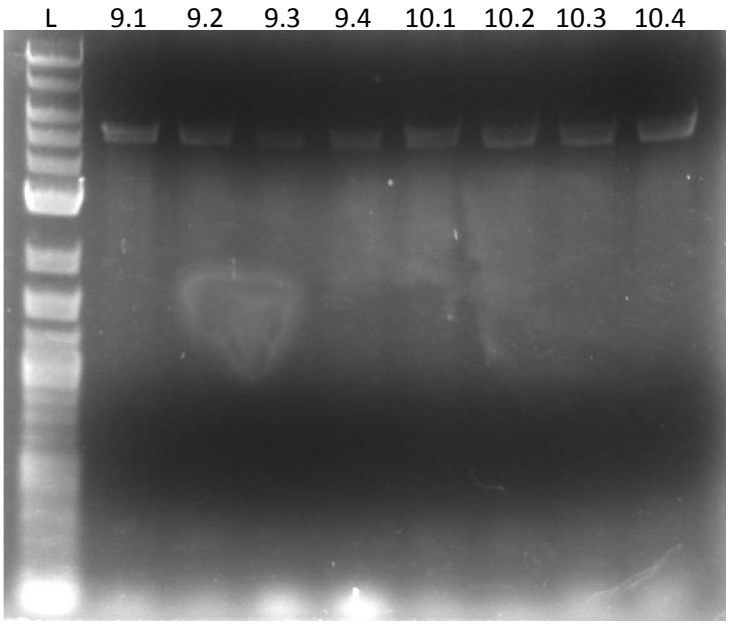


Figure 39 Electrophoresis of the transformation cassette containing the Elizabethkingia menigoseptica gene. The numbers correspond to the four replicas of sample 9 and 10, and the L corresponds to the ladder.

All samples presented the correct length of 5453bp.The samples were used to transform *S. bombicola* PT36 Ura3 negative competent cells, which were plated in SD Ura negative medium.

2. *Rhodococcus rhodochrous* gene

The genome of *Rhodococcus rhodochrous* is not fully known, including the gene that encodes for the oleate hydratase enzyme. For this reason it was necessary to isolate the genome of the bacterium and then perform a degenerated PCR. The primers used were designed based on conserved regions screened between different bacteria, including *Elizabethkingia meningoseptica*. The next step is gene walking, using primers which were designed based on the results obtained though the degenerated PCR. After the gene that encodes for the oleate hydratase is fully sequenced it is ligated (using Gibson assembly) in CYP52M1trunc for selection in *E. coli*. Following the same procedure as for the *Elizabethkingia meningoseptica* gene, the DNA cassette is used to transform *S. bombicola* cells.

The first step for obtaining the gene that encodes for the oleate hydratase protein is genomic DNA isolation of the referred bacteria. The electrophoresis result shows that the extraction was successfully performed.

The sample containing the genomic DNA of *Rhodococcus rhodochrous* was used to perform a degenerated PCR. The results are shown in figure 40. The annealing temperatures used were 40°C, 44°C, 51.2°C and 55.9°C (1st, 2nd, 3rd and 4th samples, respectively). The primers used had the following sequence: 5'-ATGATHMGNGAYGGNCARATGGC-3' for the forward primer and 5'-TGNCCDATYTCRTTYTCRTC-3' for the reverse primer.

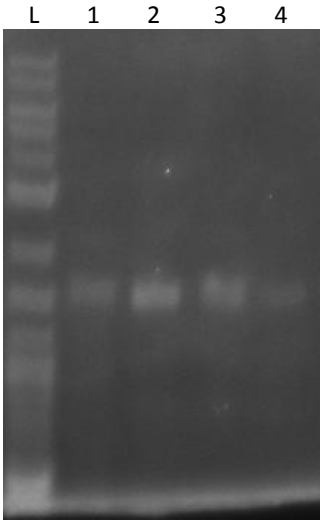


Figure 40 Electrophoresis of degenerate PCR of the Rhodococcus rhodochrous genome. The first line corresponds to the ladder. The annealing temperature of the first sample was performed at 40.0°C, the second at 44.0°C, the third at 51.2°C and the fourt

The DNA present in the above samples was ligated into a pGEM-T vector, which was used to transform DH5α cells. A plasmid isolation was performed on colonies which shown the presence of the *Rhodococcus rhodochrous* DNA fragment, and the samples sent to sequencing. The BLAST alignment indicates that the sequenced gene has an identity of 99% with a hypothetical protein of *Rhodococcus*, and 82% of identity with the myosin-cross-reactive antigen of *Mycobacterium farcinogenes*. This results suggest that part of the *Rhodococcus rhodochrous* gene that encodes for the oleate hydratase was sequenced.

Since only a fraction of the gene was sequenced, the next step was gene walking. The sample containing the genomic DNA of *Rhodococcus rhodochrous* was used. The result from the primary PCR is depicted in figure 42.

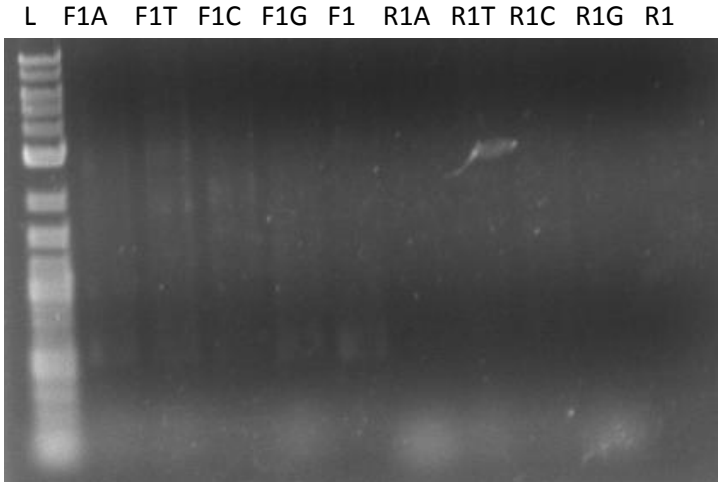


Figure 41 Electrophoresis of the primary PCR of *Rhodococcus rhodochrous* gene – gene walking. The first line corresponds to the mass ladder (L). The first five samples contain the first forward primer and a gene-walking A, T, C or G primer, being the fifth the blank. The last five samples contain the first reverse primer and a gene-walking-primer, being the last one also the blank.

The primary PCR was followed by the first sequential hybrid primer PCR (SHP-PCR). The results are shown in figure 43.

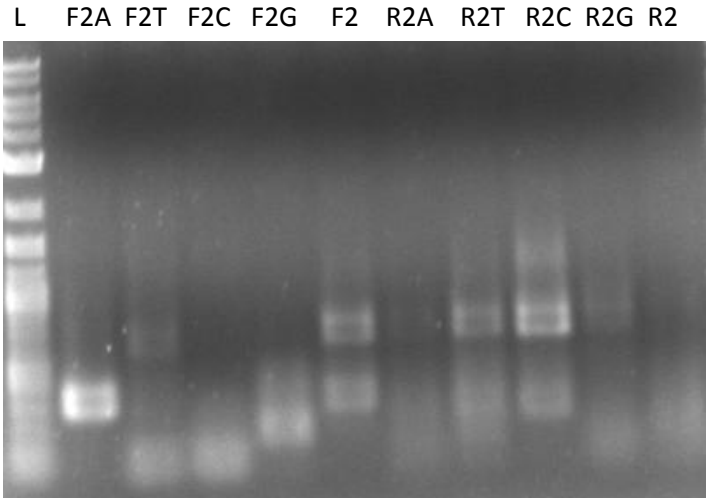


Figure 42 Electrophoresis of the first SHP-PCR of *Rhodococcus rhodochrous* gene – gene walking. The first line corresponds to the mass ladder (L). The first five samples contain the second forward primer and a gene-walking A, T, C or G primer, being the fifth the blank. The last five samples contain the second reverse primer and a gene-walking-primer, being the last one also the blank.

The results from the second SHP-PCR are shown in figure 44.

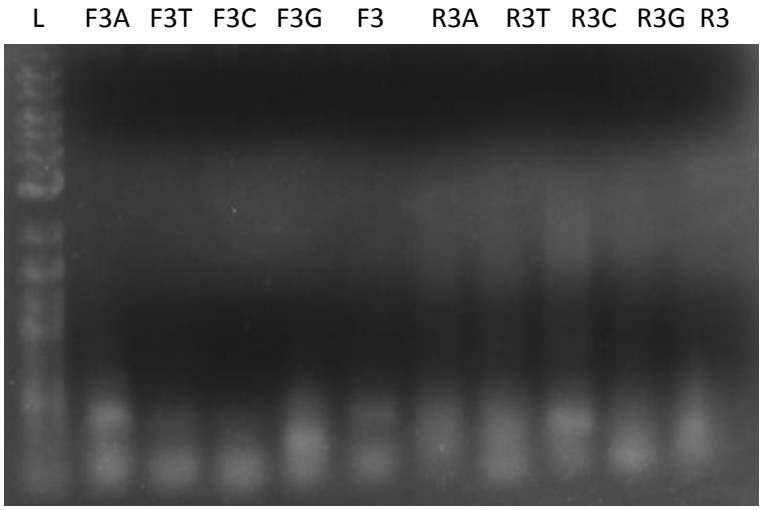


Figure 43 Electrophoresis of the second SHP-PCR of *Rhodococcus rhodochrous* gene – gene walking. The first line corresponds to the mass ladder (L). The first five samples contain the third forward primer and a gene-walking A, T, C or G primer, being the fifth the blank. The last five samples contain the third reverse primer and a gene-walking-primer, being the last one also the blank.

The primers used in gene walking are present in table 17.

Table 17 Primers, and respective sequence, used in Gene Walking

PCR	Name	Sequence
Primary PCR	Primary PCR single-tag	5'-TCATGCTAGGATCCAAGG-3'
	Gene walking A	5'-TGCTAGGATCCAAGGNNNNNATTCG-3'
	Gene walking T	5'-TGCTAGGATCCAAGGNNNNNTTTCG-3'
	Gene walking C	5'-TGCTAGGATCCAAGGNNNNNCTTCG-3'
	Gene walking G	5'-TGCTAGGATCCAAGGNNNNNGTTCG-3'
	Primer Forward 1	5'-GTCGTCTTTGTTGAGCCAGTAG-3'
	Primer Reverse 1	5'-CAACACCCGACTGCATCTTCAC-3'
SHP-PCR first round	SHP PCR1 single-tag	5'-GCAGTTCATGGTCCCGTCATC-3'
	SHP PCR1 hybrid-tag	5'-GCAGTTCATGGTCCCGTCATCCTCATGCTAGGATC CAAGG-3'
	Primer Forward 2	5'-GAACAGGTCCCAGAGACATTCG-3'
	Primer Reverse 2	5'-TACAGCCTGCTCGGGATCGAACG-3'
SHP-PCR second round	SHP PCR2 single-tag	5'-AAAGAAGAGTCGCAGAGGCTCG-3'
	SHP PCR2 hybrid-tag	5'-AAAGAAGAGTCGCAGAGGCTCGGCAGTTCATGGT CCCGTCATCC-3'
	Primer Forward 3	5'-TGATCTTCCATCTCGCGGCCACC-3'
	Primer Reverse 3	5'-CCACCTACGACATCCGCAAACTTC-3'

Gel extraction of samples F2A and R2C was performed. The DNA was inserted into a pGEM-T vector which was used to transform DH5α cells. The next step was performing a colony PCR with primers SP6 (5'-CGCCAAGCTATTTAGGTGAC-3') as forward and T7 (5'-GGCGATTAAGTTGGGTAACG-3') as reverse. Samples which had bands of the correct size were sent for sequencing with the previously mentioned primers. Complementary regions were found between the already sequenced portion of the gene with the sequencing results from sample F2A and R2C. The BLAST alignment results indicate that the protein encoded by the DNA sequence has a 99% of identity with a hypothetical protein of *Rhodococcus*, with a 77% of query cover and the frame -2. With this information, the DNA sequence was constructed with a 99% of identity with the hypothetical protein of *Rhodococcus*, 100% of query cover and the frame +1. These results indicate that the *Rhodococcus rhodochrous* gene that encodes for the oleate hydratase protein was successfully sequenced and identified. As referred previously, this gene was never sequenced and little is known about the *Rhodococcus rhodochrous* genome.

The next step was insertion of the sequenced gene in the vector CYP52M1trunc, using Gibson assembly. To guarantee the overlapping regions necessary for the assembly, a high fidelity PCR was performed, using the extracted genome of *Rhodococcus rhodochrous* and the following primers: 5'-CTCCGTTTATCATATTGAAAAGTACAT-3' as forward, and 5'-TCCATGGGAGGCTAAGAAACTCATTCT-3' as reverse. For the same reason, a high fidelity PCR was also performed for the referred vector using the following primers: 5'-GTTTCTTAGCCTCCCATGGAAG-3' as forward primer and 5'-

ATATGTACTTTTCAATATGATAAACGGAGAAATAACG-3' as reverse primer. Transformation of DH5α cells was performed using the Gibson reaction and a colony PCR with the following primers: 5'-CGTTGTCAAGTCCTAAGGTAT-3' as forward and 5'-GAACAGGTCCCAGAGACATTCG-3' as reverse. The results are shown in figure 45.

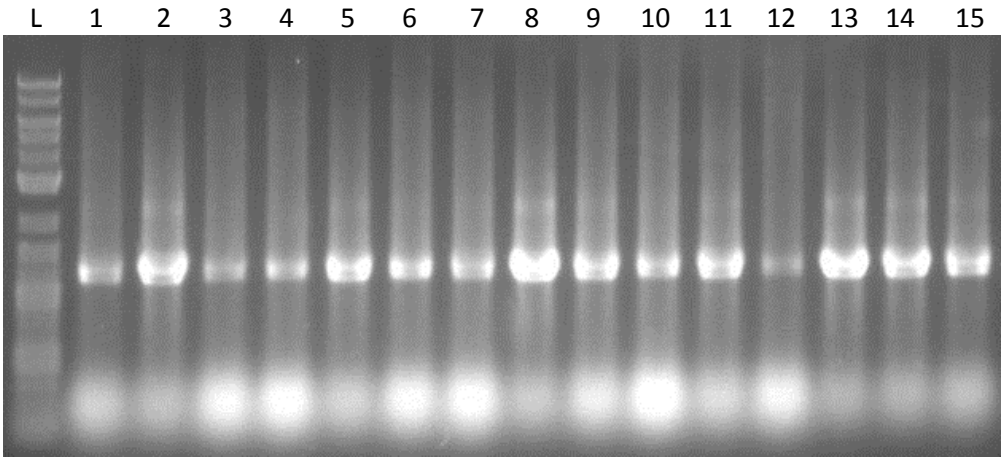


Figure 44 Electrophoresis of the colony PCR of DH5α cells containing the *Rhodococcus rhodochrous* gene and the vector CYP52M1trunc. The numbers correspond to the 15 colonies analyzed and the L corresponds to the ladder.

All colonies in figure 43 show the correct DNA length (1270bp). A plasmid extraction of sample 8 was performed and the sample sent to sequence. Good plasmids were used as template to amplify the integration cassette in a similar way as for the *Elizabethkingia meningoseptica* gene. The results are shown in figure 46.

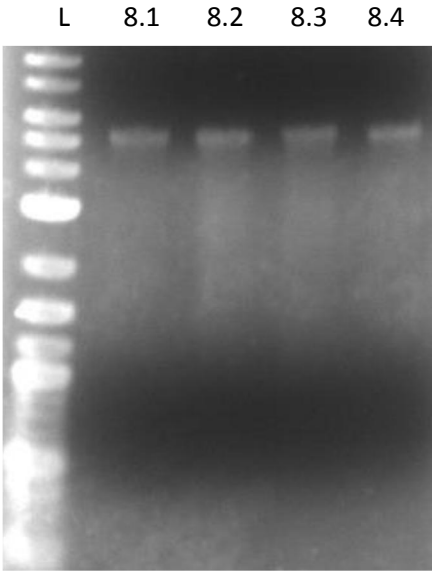


Figure 45 Electrophoresis of the transformation cassette containing the *Elizabethkingia menigoseptica* gene. The numbers correspond to the four replicas of sample 8, and the L corresponds to the ladder.

All samples presented the correct length of 5279bp.The samples were used to transform *S. bombicola* PT36 Ura3 negative competent cells, which were plated in SD Ura negative medium.

2.2 Analysis of the oleate hydratase gene of *Rhodococcus rhodochrous*

The sequence of the *Rhodococcus rhodochrous* gene that encodes for the oleate hydratase protein is the following:

ATGTATTACAGCAGTGGAAACTACGAAGCGTTCGCGCGGCCACGCAAGCCTGACGGGGTGG
ACGGCAAGACTGCATGGTTCGTCGGTTCGGGGCTTGCATCGCTTGCCGGTGCAGCATTCATG
ATCCGCGACGGCCGGATGGCAGGAAACAACATCACCGTCCTGGAGCGGTTGAAATTGCCCG
GCGGTGCACTCGACGGAATCAAGGAACCGGAGAAGGGCTTCGTCATTCGCGGTGGCCGCGA
GATGGAAGATCACTTCGAATGTCTCTGGGACCTGTTCCGCTCTGTCCCCTCGATCGAGGTCG
AAGACGCCAGCGTGCTCGACGAATTCTACTGGCTCAACAAAGACGACCCGAACTATTCGCTG
CAGCGCGTCACCGAACGTCAGGGTGAGGACGCGCACACAGACTTCAAGTTCAATCTGAACT
CGAAGGCGCAGAAGGACATCATGAAGGTCTTTCTCACTCCGCGTAGCGAACTCGAGAACAA
GCGCATCAACGAGGTGTTCGGCAAGGAATTCCTCGCAAGCAATTTCTGGCTGTACTGGCGCA
CAATGTTCGCGTTCGAGGAATGGCACAGTGCTCTCGAAATCAAGCTGTACCTGCATAGATTC
ATCCATCACATCAAAGGCCTGCCGGACCTGAGTACCCTCAAGTTCACCAAATACAACCAGTA
CGAGTCTCTCGTGCTGCCGCTGTACACGTGGCTGCTCGATCAGGGAGTGAATTTTCGGTTCG
ATACCGAAGTCACCGACATTGACTTCGATATCACGGGTGAAGTCAAGCGGGCCAAACGCAT
TCACTGGAAATCCGAGGGTGTTTTGGGTGGGGTCGACCTCGATGACGGCGACCTGGTGTTGA
CTACCATCGGCTCGCTCACCGAGAACTCGAACAACGGAGACCATCACACGCCGGCGAAACT
CGACGAGGGCCCCGCTCCAGCGTGGGATTTGTGGCGTCGTATAGCAACAAAAGACCCCGCA
TTCGGACGACCAGACGTCTTCGGTGGGCACGTCGCAGAAACCAAGTGGGAATCAGCGACAG
TGACCACGTTGGATGCCCGAATCCCAGAATACATCGAAAAGATATGCAAGCGTGACCCATTC
AGTGGACGCGTAGTCACCGGCGGCATCGTGACAGCACGAGATTCGAAGTGGTTGATGAGCT
GGACCGTAAACCGGCAACCGCACTTCAAACAACAACCCAAGGATCAGATCGTCGTCTGGGT
CTACTCACTGTTCGTCGACGTCCCCGGTGACTACGTCGACAAGCCGATGCAGGAATGCACCG
GCGAGGAAATCACCCAAGAGTGGCTCTATCACATGGGCGTTCCAGTCGAGGACATCCCGGA
ACTCGCGGCCAATGCGGCGAAGACCGTTCCAGTGATGATGCCCTATGTCACTTCCTTTTTCAT
GCCCCGTCAGGCCGGTGACCGGCCTGCGGTTGTACCGGATGGTGCCGTCAACTTCGCCTTCC
TCGGCCAGTTCGCGGAAACAACACCCGACTGCATCTTCACGACGGAGTACTCGGTCCGCACC
GGGATGGAGGCCGCGTACAGCCTGCTCGGGATCGAACGCGGAGTCCCGGAAGTTTTCGGCT
CCACCTACGACATCCGCAAACTTCTCCAAGCAACCTACTTCCTCAATGACAAGAAGGAGGAG
AGCGTGCCGTTGCCGAAGCTACTTCGCCGCCGACTCGGCAAGAAGCTCGACGAGAACGAGA
TCGGACAGCTGCTGCACGAATACCACATTCTCCCCCGAGAATGA

This gene is flanked downstream by a DNA sequence which encodes for a choloylglycine hydrolase. Both sequences are separated by 11bp. The BLAST alignment results indicate a 99% identity with the referred protein of *Rhodococcus erythropolis* and 79% with the one from *Rhodococcus rhodochrous*. The DNA sequence of the referred gene was only partially sequenced (529bp).The upstream sequence of the gene (164bp) couldn't be identified.

Using the genome walking technique, partial sequences of other genes of *Rhodococcus rhodochrous* were also identified. One of these genes (1091bp) corresponds to a beta-lactase, identified in BLAST as from an uncultured soil bacteria with an identity of 63%. The gene of an alkanesulfonate monooxygenase (1033bp) was identified as from *Rhodococcus* with an identity of 98%. The gene of beta-lactase is located downstream, from the one that encodes for the oleate hydratase, and the gene of alkanesulfonate monooxygenase is located upstream.

Using the ExPASy Translate tool, and assuming a standard codon usage, the DNA sequence of oleate hydratase encodes for the following 588 aminoacid sequence. The oleate hydratase protein has an estimated weight of 67097.1Da and a pI of 5.32 (data provided by ExPASy tool ProtParam):

MYYSSGNYEAFARPRKPDGVDGKTAWFVGSGLASLAGAAFMIRDGRMAGNNITVLERLKLPGG
ALDGIKEPEKGFVIRGGREMEDHFECLWDLFRSVPSIEVEDASVLDEFYWLNKDDPNYSLQRVTE
RQGEDAHTDFKFNLNSKAQKDIMKVFLTPRSELENKRINEVFGKEFLASNFWLYWRTMFAFEE
WHSALEIKLYLHRFIHHIKGLPDLSTLKFTKYNQYESLVLPLYTWLLDQGVNFRFDTEVTDIDFDI
TGEVKRAKRIHWKSEGVLGGVDLDDGDLVLTTIGSLTENSNNGDHHTPAKLDEGPAPAWDLWR
RIATKDPAFGRPDVFGGHVAETKWESATVTTLDARIPEYIEKICKRDPFSGRVVTGGIVTARDSK
WLMSWTVNRQPHFKQQPKDQIVVWVYSLFVDVPGDYVDKPMQECTGEEITQEWLYHMGVPV
EDIPELAANAAKTVPVMMPYVTSFFMPRQAGDRPAVVPDGAVNFAFLGQFAETTPDCIFTTEYS
VRTGMEAAYSLLGIERGVPEVFGSTYDIRKLLQATYFLNDKKEESVPLPKLLRRRLGKKLDENEI
GQLLHEYHILPRE

An alignment between the oleate hydratase aminoacid sequence of *Rhodococcus rhodochrous* and other 7 sequences from different microorganism found in GenBank was made. The Clustal tool was used and a phylogentic tree (figure 47) constructed based on the protein maximum likelihood principle.

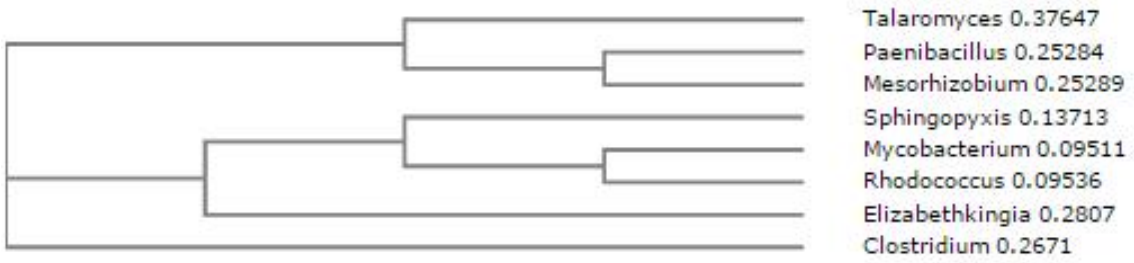


Figure 46 Phylogenic tree for oleate hydratase aminoacid sequence of 8 microorganisms, including Rhodococcus rhodochrous, Elizabethkingia meningoseptica, Mycobacterium farcinogenes, Paenibacillus riograndensis SBR5, Clostridium bornimense, Sphingopyxis sp. Kp5.2

The results obtained indicate that the oleate hydratase protein of *Rhodococcus rhodochrous* is closer in phylogeny with the one from *Mycobacterium farcinogenes*, secondly with *Sphingopyxis* sp. Kp5.2, and then with the protein of *Elizabethkingia*

meningoseptica. Figure 48 depicts the alignment of the aminoacid sequence of the four mentioned microorganisms.

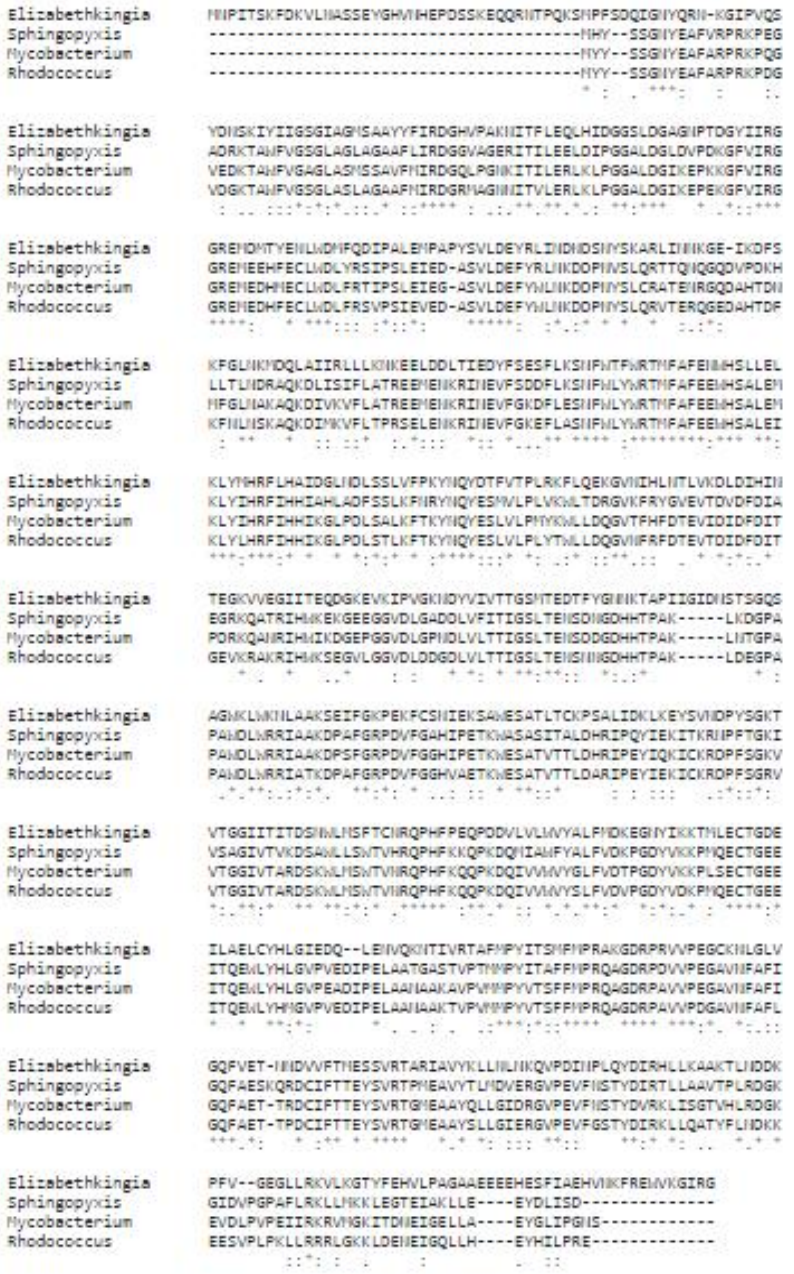


Figure 47 Alignment of the oleate hydratase aminoacid sequence of *Rhodococcus rhodochrous*, *Elizabethkingia meningoseptica, Mycobacterium farcinogenes* and *Sphingopyxis sp.* The symbol * indicates positions which have a single and fully conserved residue, : indicates conservation between groups of strongly similar properties, and . indicates conservation between groups of weakly similar

With the CLUSTAL alignment several conserved regions can be identified between the aminoacid sequences. The BLAST alignment shows that the aminoacid sequence of the oleate hydratase of *Rhodococcus rhodochrous* has 81% identity with the one from *Mycobacterium farcinogenes*, 72% of identity with the one from *Sphingopyxis* sp. Kp5.2 and 45% with the one from *Elizabethkingia meningoseptica*.

Conclusions and perspectives

Sophorolipids molecules are a class of biosurfactants, which are recently given great interest by industry. They have various applications, mainly in cleaning and cosmetic industries, and they are environmentally friendly.

The most studied and best producer of sophorolipids is *Starmerella bombicola*. The molecules produced by this yeast are constituted by a sophorose molecule that binds to the terminal or sub-terminal of a fatty acid, usually with sixteen or eighteen carbons. The sophorose molecule can have one or two acetylations, and lactonic forms are possible. *Rhodotorula bogoriensis* also produces sophorolipids, but the sophorose attaches to the middle of the fatty acid chain. Despite their interesting properties, this molecules are produced in low yields by the referred microorganism, making it difficult to assess their properties properly. For this reason, the aim of this thesis was to induce *S. bombicola* to produce sophorolipids with a branched hydrophobic tail.

Unconventional hydrophobic substrates were used for the cultures of *S. bombicola* cells: three Guerbet alcohols and four diacyl glycerols. All this molecules have a free hydroxyl group in common, prone for the binding of the sophorose molecule and for the production of sophorolipids with a branched hydrophobic tail.

Growth tests were performed using the three Guerbet alcohols and two stains of the yeast, a wild-type and a Cyp knock-out. The results of the wild-type culture indicate that it is possible to produce the desired sophorolipid molecules with G24 as substrate. LC-MS results suggest that the *S. bombicola* culture produced a branched sophorolipid molecule with no acetylations or lactonization. This thesis represent the first time such a molecule was produced using *S. bombicola* culture. This culture may also have produced other branched molecules, however is not possible to determine with certainty, since other molecules are possible for the same molecular mass. Other new-to-nature sophorolipid were produced by the culture, however they are not branched molecules. Lactonic forms were not expected, since G24 has no carboxyl group at the end of its chain, but they were detected by LC-MS. The production of this molecules can be explained by the action of enzymes present is the *S. bombicola* cells that must have introduced the carboxyl group. Most of the molecular masses identified can be explained by molecules with an additional hydroxyl function, and in some cases further presence of a carboxyl function. Probably, the sophorose unit bind to this additional

terminal or sub-terminal hydroxyl group. This may indicate that the Ugt enzymes are more specific for the glycosylation of the sophorose molecule to the referred position on the fatty acid chain. This supposition may also be the reason why the Cyp knock-out culture didn't produce sophorolipid molecules. Without the Cyp gene, the hydroxylation of the fatty acid at the terminal or sub-terminal is not possible. Another reason for the lack of production by this culture may be that the genes of the sophorolipid cluster are not expressed due to the suppression of the Cyp gene.

There was no production of new-to-nature sophorolipid molecules by the wild-type cultures with the substrates G16 and G20. This results indicate that this substrates were not incorporated in the sophorolipid production pathway.

In the case of the diacyl glycerols tests, no growth experiments were made, only an enzyme test. There was only a minor production of molecules using C8:0 glyceride dicaprylin and C10:0 glyceride dicaprin as substrates. This results indicate that diacyl glycerols are not the ideal substrates for the production of sophorolipids with a branched hydrophobic tail by *S. bombicola*.

As future perspectives, it would be interesting to study the properties of the new-to-nature sophorolipid molecules produced by the wild-type culture with G24 as substrate. Analysis with nuclear magnetic resonance (NMR) after purification of specific compounds could confirm the molecular structure. Increasing the fermentation scale would also be interesting to obtain higher production levels of these molecules, as only 16.5g/L were obtained using flasks experiments. A fed-batch experiment in a bio-reactor would be advisable.

The genes of two bacteria that encode for the oleate hydratase enzyme were intended to be used for the production of sophorolipids with a branched hydrophobic tail. As referred previously, the two bacteria from which this gene was isolated were: *Rhodococcus rhodochrous*, for which the gene sequence is not known, but the bacterium produces R and S-enantiomers; and *Elizabethkingia meningoseptica* for which the sequence is known but the codon is not the same as *S. bombicola* and the bacterium only produces R-enantiomers.

In the case of *Rhodococcus rhodochrous* gene it was necessary to isolate its genomic DNA and select the gene that encodes for the oleate hydratase through various PCR techniques. It was possible to successfully isolate the gene of interest and determine its DNA sequence. This represents the first time this gene and aminoacid

sequence were identified, which constitutes an important mark in better knowing the *Rhodococcus rhodochrous* genome.

In the case of the *Elizabethkingia meningoseptica*, the gene was ordered with codon optimization for *S. bombicola*. For both genes a transformation cassette for homologous recombination was created and transformed in *S. bombicola* cells.
In the timeframe of this thesis, no further test could be conducted, but next steps are the evaluation of the transformants in shake flask cultures for sophorolipid synthesis. HPLC and LC-MS analysis should be done to assess if the cultures produce sophorolipid molecules with a branched hydrophobic tail. One expects a higher probability of the production of this molecules by the cells transformed with the gene of *Rhodococcus rhodochrous*, since the products would be a mix of R and S-enantiomer. The reason is that only the S-enantiomers of the subterminal hydroxylated fatty acids get incorporated in the classic sophorolipid production pathway, but it's not known if the same happens for the in-chain hydroxylated fatty acids.

This master thesis reports the first production of sophorolipid with branched hydrophobic tail by *S. bombicola*, with a sophorose molecule and a twenty-four carbon fatty acid. This discovery represents an important mark for this yeast as a producing organism of tailor-made sophorolipids and other biomolecules. This thesis also reports the first time the DNA sequence of the oleate hydratase gene of *Rhodococcus rhodochrous* to be identified. This discovery contributes to the better knowledge of the genome of *Rhodococcus rhodochrous*, which is limitedly known.

References

Abdel-Mawgou A, Lépine F, Déziel E. (2010) Rhamnolipids: diversity of structures, microbial origins and roles. *Appl Microbiol Biotechnol*. doi: 10.1007/s00253-010-2498-2

Akari S, Akari Y (1987) Method of modifying quality of wheat flour product. Japanese patent 61205449

Allingham RP (1971) Sophoroside esters in prepared food products. US patent 3622344

Asmer HJ, Lang S, Wagner F, Wray V. (1988) Microbial-production. Structure elucidation and bioconversion of sophorose lipids. *J Am Oil Chem Soc., 65(9)*, 1460-1466. doi: 10.1007/BF02898308.

Bajaj VK & Annapure. (2015) Castor Oil as Secondary Carbon Source for Production of Sophorolipids Using *Starmerella bombicola* NRRL Y-17069, *323*(3), 315–323.

Banat IM, Franzetti A, Gandolfi I. Bestetti G, Martinotti MG, Fracchia L, Marchant R (2010) Microbial biosurfactants production, applications and future potential. *Applied Microbiology and Biotechnology*, *87*(2), 427–44. doi:10.1007/s00253-010-2589-0

Baviere M, Degouy D, Lecourtier J (1994) Process for washing solid particles comprising a sophoroside solution. US patent 5326407

Bertani G (1951) Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *Journal of Bacteriology*, *62*(3), 293–300.

Bevers LE, Pinkse MWH, Verhaert PDEM, Hagen WR. (2009. Oleate hydratase catalyzes the hydration of a nonactivated carbon-carbon bond. J*ournal of Bacteriology*, *191*(15), 5010–2. doi:10.1128/JB.00306-09

Bhangale A, Wadekar S, Kale S, Bhowmick D, Pratap, A. (2014) Production of sophorolipids synthesized on castor oil with glucose and glycerol by using *Starmerella bombicola* (ATCC 22214). *European Journal of Lipid Science and Technology*, *116*, 336–343. doi:10.1002/ejlt.201300236

Bisht KS, Gross RA, Kaplan DL (1999) Enzyme-mediated regioselective acylations of sophorolipids. *J Org Chem, 64*, 780–789. doi: 10.1021/jo981497m

Chen J, Song X, Zhang H, Qu YB, Miao JY (2006) Sophorolipid produced from the new yeast strain *Wickerhamiella domercqiae* induces apoptosis in H7402 human liver cancer cells. *Appl Microbiol Biotechnol*, *72 (1)*, 52–9. doi: 10.1007/s00253-005-0243-z

Chen J, Song X, Zhang H, Qu YB, Miao JY (2010) Production, structure elucidation and anticancer properties of sophorolipid from *Wickerhamiella domercqiae. FEMS Microbiol Lett, 311(3)*, 140–6. doi:10.1016/j.enzmictec.2005.12.022

Daniel H-J, Reuss M, Syldatk C (1998) Production of sophorolipids in high concentration from deproteinized whey and rapeseed oil in a two stage fed batch process using *Candida bombicola* ATCC 22214 and *Cryptococcus curvatus* ATCC 20509. *Biotechnol Lett*, *20 (12)*, 1153-1156. doi:10.1023/A:1005332605003

Davila AM, Marchal R, Vandecasteele JP (1992) Kinetics and balance of a fermentation free from product inhibition-sophorose lipid production by *Candida bombicola*. *Appl Microbiol Biotechnol, 38(1)*, 6–11. doi:10.1007/BF00169410

Davila AM, Marchal R, Monin N, Vandecasteele JP (1993) Identification and determination of individual sophorolipids in fermentation products by gradient elution high-performance liquid-chromatography with evaporative lightscattering detection. *J Chromatogr*, *648(1)*, 139–49. doi:10.1016/0021-9673(93)83295-4

Davila AM, Marchal R, Vandecasteele JP (1994) Sophorose lipid production from lipidic precursors-Predictive evaluation of industrial substrates. *J Ind Microbiol,* 13(4), 249–257. doi:10.1007/BF01569757

Davila AM, Marchal R, Vandecasteele JP (1997) Sophorose lipid fermentation with differentiated substrate supply for growth and production phases. *Appl Microbiol Biotechnol*, *47(5)*, 496–501. doi: 10.1007/s002530050962

De Koster C, Heerma W, Pepermans HAM, Groenewegen A, Peters H, Haverkamp J (1995) Tandem mass spectrometry and nuclear magnetic resonance spectroscopy studies of *Candida bombicola* sophorolipids and product formed on hydrolysis by cutinase. *Anal Biochem 230(1)*, 135–148. doi: 10.1007/s002530050962

Develter D, Renkin M, Jacobs I (2007) Detergent compositions. EP1445302

Develter D, Lauryssen LML (2010) Properties and industrial applications of sophorolipids. *Eur J Lipid Sci Technol, 112(6)*, 628–38. doi: 10.1002/ejlt.200900153

Ducreux J, Ballerini D, Baviere M, Bocard C, Monin N (1997) Composition containing a surface active compound and glycolipids and decontamination process for a porous medium polluted by hydrocarbons. US patent 5654192

Esders TW, Light RJ (1972) Glucosyl- and acetyltransferases involved in the biosynthesis of glycolipids from *Candida bogoriensis*. *J Biol Chem 247(5)*, 1375–86. doi: 10.1007/s002530050962

Felse PA, Shah V, Chan J, Rao KJ, Gross RA (2007) Sophorolipid biosynthesis by *Candida bombicola* from industrial fatty acid residues. *Enzyme Microb Technol*, 40(2), 316–23. doi: 10.1016/j.enzmictec.2006.04.013

Giani C, Meiwes J, Rothert R, Wullbrandt D. (1997). Pseudomonas aeruginosa and its use in a Process for the Biotechnological Preparation of L-rhamnose. Patent US5501966

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–5. doi:10.1038/nmeth.1318

Gobbert U, Lang S, Wagner F (1984) Sophorose lipid formation by resting cells of *Torulopsis bombicola. Biotechnol Lett, 6(4)*, 225–230. doi:10.1007/s002530050962

Goñi FM & Alonso A. (1999) Structure and functional properties of diacylglycerols in membranes1This work is dedicated to Professor Vittorio Luzzati on occasion of his 75th birthday.1. *Progress in Lipid Research*, *38*(1), 1–48. doi:10.1016/S0163-7827(98)00021-6

Gorin PAJ, Spencer JFT, Tulloch AP (1961) Hydroxy fatty acid glycosides of sophorose from *Torulopsis magnoliae*. *Can J Chem, 39(4)*, 846–855 doi: 10.1139/v61-104

Greenway DLA & Dyke KGH. (1979) Mechanism of the Inhibitory Action of Linoleic Acid on the Growth of *Staphylococcus aureus*. *Journal of General Microbiology*, *115*(1), 233–245. doi:10.1099/00221287-115-1-233

Gross RA, Shah V (2007) Anti-herpes virus properties of various forms of sophorolipids. World patent WO2007US63701.

Guilmanov V, Ballistreri A, Impallomeni G, Gross RA (2002) Oxygen transfer rate and sophorose lipid production by *Candida bombicola*. *Biotechnol Bioeng, 77(5)*, 489–494 doi: 10.1007/s002530050962

Hall PJ, Haverkamp J, Van Kralingen CG, Schmidt M (1996) Laundry detergent composition containing synergistic combination of sophorose lipid and nonionic surfactant. US patent 5520839.

Hirata Y, Ryu M, Oda Y, Igarashi K, Nagatsuka A, Furuta T, et al. (2009) Novel characteristics of sophorolipids, yeast glycolipid biosurfactants, as biodegradable low-foaming surfactants. *J Biosci Bioeng, 108(2)*, 142–6. doi: 10.1016/j.jbiosc.2009.03.012.

Hommel R & Ratledge C. (1993) *Biosurfactants: Production: Properties: Applications*. (N. Kosaric, Ed.) (p. 4). Marcel Dekker, INC. Retrieved from http://www.google.pt/books?hl=pt-PT&lr=&id=Vt9fD5D4bNAC&pgis=1

Hu Y & Ju LK (2001). Sophorolipid production from different lipid precursors observed with LC-MS. *Enzyme and Microbial Technology*, *29*(10), 593–601. doi:10.1016/S0141-0229(01)00439-2

Huang N, Agrawal V, Giacomini KM, Miller WL. (2008) Genetics of P450 oxidoreductase: sequence variation in 842 individuals of four ethnicities and activities of 15 missense mutations. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(5), 1733–8. doi:10.1073/pnas.0711621105

Inoh Y, Furuno T, Hirashima N, Kitamoto D, Nakanishi M (2011) Rapid delivery of small interfering RNA by biosurfactant MEL-A-containing liposomes. *Biochem Biophys Res Commun, 414(3)*, 635-640 doi: 10.1016/j.bbrc.2011.09.147

Inoue S, Miyamoto N (1980) Process for producing a hydroxyfatty acid ester. US patent 4201844

Isoda H, Kitamoto D, Shinmoto H, Matsumura M, Nakahara T (1997) Microbial extracellular glycolipid induction of differentiation and inhibition of the protein kinase C activity of human promyelocytic leukemia cell line HL60. *Biosci Biotechnol Biochem, 61(4)*, 609–14. doi: 10.1271/bbb.61.609

Ivanković T & Hrenović J. (2010) Surfactants in the environment. *Arhiv Za Higijenu Rada I Toksikologiju*, *61*(1), 95–110. doi:10.2478/10004-1254-61-2010-1943

Joo YC, Jeong KW, Yeom SJ, Kim YS, Kim Y, Oh DK. (2012). Biochemical characterization and FAD-binding analysis of oleate hydratase from *Macrococcus caseolyticus*. *Biochimie*, *94*(3), 907–15. doi:10.1016/j.biochi.2011.12.011

Kasture M, Singh S, Patel P, Joy PA, Prabhune AA, Ramana CV, et al. (2007) Multiutility sophorolipids as nanoparticle capping agents: synthesis of stable and water dispersible Co nanoparticles. *Langmuir 23(23)*, 11409–12. doi: 10.1021/la702931j

Kawano J, Suzuki T, Inoue S, Hayashi S (1981a) Powdered compressed cosmetic material. US patent 4305931

Kawano J, Suzuki T, Inoue S, Hayashi S (1981b) Stick-shaped cosmetic material. US patent 4305929

Kim K, Kim M, Lim J, Park H, Lee S (2005) Transfer of *Chryseobacterium meningosepticum* and *Chryseobacterium miricola* to *Elizabethkingia gen*. nov. as *Elizabethkingia meningoseptica* comb. nov. and *Elizabethkingiamiricola* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, 55(3), 1287–1293. 10.1099/ijs.0.63541-0

Klekner V, Kosaric N, Zhou QH (1991) Sophorose lipids produced from sucrose. *Biotechnol Lett, 13(5)*, 345–348. doi: 10.1271/bbb.61.609

Konishi M, Fukuoka T, Morita T, Imura T, Kitamoto D (2008) Production of new types of sophorolipids by *Candida batistae*. *J Oleo Sci, 57(6)*, 359–69. doi: 10.1271/bbb.61.609

Kurtzman CP, Price NPJ, Ray KJ, Kuo TM. Production of sophorolipid Biosurfactants by multiple species of the *Starmerella (Candida) bombicola* yeast clade. *FEMS Microbiology Letters*, *311*(2), 140–6. doi:10.1111/j.1574-6968.2010.02082.x

Lang S, Brakemeier A, Heckmann R, Spockner S, Rau U (2000) Production of native and modified sophorose lipids. *Chim Oggi-Chem Today* 18, 76-79

Lang S, Katsiwela E, Wagner F (1989) Antimicrobial effects of biosurfactants. *Fett Wiss Technol-Fat Sci Technol, 91(9)*, 363–6. doi: 10.1002/lipi.19890910908

Larkin, M., Kulakov, L., & Allen, C. (2010) Genomes and plasmids in *Rhodococcus*. *Biology of Rhodococcus*, 16, 73-90. doi: 10.1007/978-3-642-12937-7-3.

Levinson MI (2009) In U. Zoller & P. Sosis (Eds.), *Handbook of Detergents, Part F: Production* (p. 1). CRC Press Taylor & Francis Group.

Lin SC (1996) Biosurfactants: recent advances. J Chem Technol Biotechnol 66, 109–120 Lottermoser K, Schunck WH, Asperger O (1996) Cytochromes P450 of the sophorose lipid-producing yeast *Candida apicola*: Heterogeneity and polymerase chain reaction-mediated cloning of two genes. Yeast 12, 565–575

Marchal R, Lemal J, Sulzer C, Davila AM (1999) Production of sophorolipid acetate acids from oils or esters. US patent 5900366

Martin-Harris MH, Bartley PA, Morley AA (2010) Gene walking using sequential hybrid primer polymerase chain reaction. *Analytical Biochemistry*, *399*(2), 308–10. doi:10.1016/j.ab.2010.01.005

Marliere P. (2011). Method for producing an alkene comprising the step of converting an alcohol by an enzymatic dehydration step. Patent: EP2336341 A1

Morita T, Fukuoka T, Imura T, Kitamoto D (2013) Production of mannosylerythritol lipids and their application in cosmetics. *Appl Microbiol Biotechnol,* 97(11), 4691-4700. doi: 10.1007/s00253-013-4858-1

MoritaT, Ishibashi Y, Fukuoka T, Imura T, Sakai H, Abe M, Kitamoto D. (2011) Production of Glycolipid Biosurfactants, Cellobiose Lipids, by *Cryptococcus humicola* JCM 1461 and Their Interfacial Properties. *Bioscience, Biotechnology, and Biochemistry*, *75*(8), 1597–1599. doi:10.1271/bbb.110036

Morita T, Konishi M, Fukuoka T, Imura T, Sakai H, Kitamoto D (2008). Efficient production of di-acylated and tri-acylated mannosylerythritol lipids as glycolipid biosurfactants by *Pseudozyma parantarctica* JCM 11752(T). *J Oleo Sci, 57(10)*, 557-565. doi: 10.1271/bbb.61.609

Mulligan C, Wang S. (2006). Remediation of a heavy metal contaminated soil by a rhamnolipid foam. *Eng Geol*, *85(1-2)*, 75-81 doi: 10.1016/j.enggeo.2005.09.029

Nebert DW, Gonzalez FJ (1987) P450 genes: structure, evolution, and regulation. *Annu Rev Biochem, 56*, 945–93. doi:10.1146/annurev.bi.56.070187.004501

Nguyen TTL, Edelen A, Neighbors B, Sabatini DA (2010) Biocompatible lecithinbased microemulsions with rhamnolipid and sophorolipid biosurfactants: formulation and potential applications. *J Colloid Interface Sci, 348(4)*, 498–504. doi: 10.1016/j.jcis.2010.04.053

Palme O, Comanescu G, Stoineva I, Radel S, Benes E, Develter D, et al. (2010) Sophorolipids from *Candida bombicola*: cell separation by ultrasonic particle manipulation. *Eur J Lipid Sci Technol, 112(6)*, 663–73. doi: 10.1002/ejlt.200900163

Palmer JM & Keller NP. (2010) Secondary metabolism in fungi: does chromosomal location matter? *Curr Opin Microbiol, 13(4)*, 431–436. doi:10.1016/j.mib.2010.04.008

Pesce L. (2002) A biotechnological method for the regeneration of hydrocarbons from dregs and muds, on the base of biosurfactants. World patent 02/062495

Piatesi, A., Baldenius, K.U., Ditrich, K., Kindler, A., Zajaczkowski-Fischer, M., Böhling, R. (2014). Process for preparing sebacic acid. World patent: WO2014037328 A1

Raychowdhury, M. K., Goswami, R., & Chakrabarti, P. (1985). Effect of unsaturated fatty acids in growth inhibition of some penicillin-resistant and sensitive bacteria. *Journal of Applied Bacteriology*, *59*(2), 183–188. doi:10.1111/j.1365-2672.1985.tb03319.x

Raza ZA, Khalid ZM, Banat IM (2009) Characterization of rhamnolipids produced by a *Pseudomonas aeruginosa* mutant strain grown on waste oils. *J Environ Sci Health Part A—Toxic/Hazard Substances Environ Eng; 44(13)*, 1367–1373. doi: 10.1080/10934520903217138.

Reis RS, Pacheco GJ, Pereira AG, Freire DMG. (2013). Biosurfactants : Production and Applications. doi: 10.57727 /56144

Ribeiro I, Bronze MR, Castro MF, Ribeiro MHL. (2012). Design of selective production of sophorolipids by *Rhodotorula bogoriensis* through nutritional requirements. *Journal of Molecular Recognition : JMR*, *25*(11), 630–40. doi:10.1002/jmr.2188

Roelants S, Saerens K, Derycke T, Lin YC, Van de Peer Y, Soetaert W. (2013). *Candida bombicola* as a platform organism for the production of tailor-made biomolecules. *Biotechnology and Bioengineering*, *110*(9), 2494–503. doi:10.1002/bit.24895

Rosa CA, Lachance MA, Silva JOC, Teixeira ACP, Marini MM, Antonini Y, Martins RP (2003) Yeast communities associated with stingless bees. Fems Yeast Res 4, 271-275. doi: 10.1016/S1567-1356(03)00173-9

Saerens KMJ, Roelants SLKW, Van Bogaert INA, Soetaert W (2011b) Identification of the UDP-glucosyltransferase gene UGTA1, responsible for the first glucosylation step in the sophorolipid biosynthetic pathway of *Candida bombicola* ATCC 22214. *FEMS Yeast Res* ,11(1), 123–32. doi:10.1111/j.1567-1364.2010.00695.x

Saerens KMJ, Zhang J, Saey L, Van Bogaert INA, Soetaert W (2011c) Cloning and functional characterisation of the UDP-glucosyltransferase UgtB1 involved in sophorolipid production by *Candida bombicola* and creation of a glucolipid producing yeast strain. *Yeast, 28(4),* 279-92. doi:10.1002/yea.1838.

Saleh H, El-Sharkawy, Yang W, Dostal L, Rosazza J (1992) Microbial oxidation of oleic acid. *Applied and environmental microbiology, 58(7), 2116-2122.*

Sanchis J, Fernández L, Carballeira JD, Drone J, Gumulya Y, Höbenreich H, Reetz MT (2008). Improved PCR method for the creation of saturation mutagenesis libraries in directed evolution: application to difficult-to-amplify templates. *Applied Microbiology and Biotechnology*, *81*(2), 387–97. doi:10.1007/s00253-008-1678-9

Scamehorn J, Sabati D, Harwell J. (2004) *Encyclopedia of Supramolecular Chemistry, Volume 2*. (J. Atwood & J. Steed, Eds.) (p. 1463). CRC Press. Retrieved from https://books.google.com/books?id=pcQSir2QXbMC&pgis=1

Schofield M, Renga T, Gross R. (2013, April 11). Modified sophorolipids for the inhibition of plant pathogens. Retrieved from http://www.google.com/patents/WO2013052615A1?cl=en

Shah V, Doncel GF, Seyoum T, Eaton KM, Zalenskaya I, Hagver R, et al. (2005) Sophorolipids, microbial glycolipids with anti-human immunodeficiency virus and sperm-immobilizing activities. *Antimicrob Agents And Chemother, 49(10)*, 4093–100. doi: 10.1128/AAC.49.10.4093-4100.2005

Sha R, Jiang L, Meng Q, Zhang G, Song Z. (2012). Producing cell-free culture broth of rhamnolipids as a cost-effective fungicide against plant pathogens. *J Basic Microbiol*, *52(4)*, 458-466. doi: 10.1002/jobm.201100295

Sleiman JN, Kohlhoff SA, Roblin PM, Wallner S, Gross R, Hammerschlag MR, et al. (2009) Sophorolipids as antibacterial agents. *Ann Clin Lab Sci,* 39(1), 60–3.

Soliance. Product guide – sopholiance: sophoro-lipids specific antibacterial biological agent; 2004.

Spoeckner S, Wray V, Nimtz M, Lang S (1999) Glycolipids of the smut fungus *Ustilago maydis* from cultivation on renewable resources. *Appl Microbiol Biotechnol* 51(1), 33-39. doi: 10.1007/s002530051359

Spencer JFT, Gorin PAJ, Tulloch AP (1970) *Torulopsis bombicola sp. n. Antonie VanLeeuwenhoek, 36(1)*, 129–33. doi:10.1007/BF02069014

Strieker M, Marahiel MA 2009) The structural diversity of acidic lipopeptide antibiotics. *ChemBioChem,* 10(4), 607-616. doi:10.1002/cbic.200800546.

Sulzbacher M. (2007). The Guerbet reaction of cetyl alcohol. *Journal of Applied Chemistry*, *5*(12), 637–641. doi:10.1002/jctb.5010051201

Teichmann B, Labbé C, Lefebvre F, Bolker M, Linne U, Bélanger RR (2011) Identification of a biosynthesis gene cluster for flocculosin a cellobiose lipid produced by the biocontrol agent *Pseudozyma flocculosa*. *Mol Microbiol, 79(6)*, 1483-1495. doi: 10.1111/j.1365-2958.2010.07533.x

Thavasi R, Jayalakshmi S, Balasubramanian T, Banat IM (2007) Biosurfactant production by *Corynebacterium kutscheri* from waste motor lubricant oil and peanut oil cake. *Lett Appl Microbiol, 45(6)*, 686–691 doi: 10.1111/j.1472-765X.2007.02256.x

Thavasi R, Jayalakshmi S, Balasubramanian T, Banat IM (2008) Production and characterization of a glycolipid biosurfactant from *Bacillus megaterium* using economically cheaper sources. World J Microbiol Biotechnol, *24(7)*, 917–925. doi:10.1111/j.1472-765X.2007.02256.x

Tulloch AP, Spencer JFT (1968) Fermentation of long-chain compounds by *Torulopsis apicola*. IV. Products from esters and hydrocarbons with 14 and 15 carbon atoms and from methyl palmitoleate. *Can J Chem, 46(9),* 1523–8. doi: 10.1139/v68-249

Tulloch AP, Spencer JFT, Deinema MH (1968) A new hydroxy fatty acid sophoroside from *Candida bogoriensis*. *Can J Chem, 46(3)*, 345–8. doi: 10.1139/v68-057

Van Bogaert I, Develter D, Soetaert W, Vandamme E. (2008) Unbound MEDLINE : Cerulenin inhibits de novo sophorolipid synthesis of *Candida bombicola*. *30*. doi:10.1007/s10529-008-9764-8

Van Bogaert I, Fleurackers S, Van Kerrebroeck S, Develter D, Soetaert, W. (2011) Production of new-to-nature sophorolipids by cultivating the yeast *Candida bombicola* on unconventional hydrophobic substrates. *Biotechnology and Bioengineering*, *108*(4), 734–41. doi:10.1002/bit.23004

Van Bogaert I, Holvoet K, Roelants SLKW, Li B, Lin YC, Van de Peer Y, Soetaert W. (2013) The biosynthetic gene cluster for sophorolipids: a biotechnological interesting biosurfactant produced by *Starmerella bombicola*. *Molecular Microbiology*, *88*(3), 501–9. doi:10.1111/mmi.12200

Van Bogaert I, Zhang J, Soetaert W. (2011) Microbial synthesis of sophorolipids. *Process Biochemistry*, *46*(4), 821–833. doi:10.1016/j.procbio.2011.01.010

Volkov A, Liavonchanka A, Kamneva O, Fiedler T, Goebel C, Kreikemeyer B, Feussner I. (2010). Myosin cross-reactive antigen of *Streptococcus pyogenes* M49 encodes a fatty acid double bond hydratase that plays a role in oleic acid detoxification and bacterial virulence. *The Journal of Biological Chemistry*, *285*(14), 10353–61. doi:10.1074/jbc.M109.081851

Wallen LL, Benedict RG, Jackson RW (1962) The microbiological production of 10-hydroxystearic acid from oleic acid. *Arch. Biochem. Biophys*. 99(2), 249–253. doi:10.1016/0003-9861(62)90006-1

Worakitkanchanakul W, Imura T, Fukuoka T, Morita T, Sakai H, Abe M, Rujiravanit R, Chavadej S, Minamikawa H, Kitamoto D (2008) Aqueous-phase behavior and vesicle formation of natural glycolipid biosurfactant, mannosylerythritol lipid-B. *Colloids Surf B Biointerfaces,* 65(1), 106-112. doi: 10.1016/j.colsurfb.2008.03.009

Yeh MS, Wei YH, Chang JS (2005) Enhanced production of surfactin from *Bacillus subtilis* by addition of solid carriers. *Biotechnology Progress*, 21(4), 1329-1334. doi: 10.1021/bp050040c

Yoo DS, Lee BS, Kim EK (2005) Characteristics of microbial biosurfactant as an antifungal agent against plant pathogenic fungus. *J Microbiol Biotechnol, 15(6),* 1164–1169.

Zhang, J., Saerens, K. M. J., Van Bogaert, I. N. a, & Soetaert, W. (2011). Vegetable oil enhances sophorolipid production by Rhodotorula bogoriensis. *Biotechnology Letters*, *33*, 2417–2423. doi:10.1007/s10529-011-0703-8

Zhang L, Somasundaran P, Singh SK, Felse AP, Gross R (2004) Synthesis and interfacial properties of sophorolipid derivatives. *Colloid Surf A Physicochem Eng Asp, 240(1-3)*, 75–82. doi: 10.1016/j.colsurfa.2004.02.016

Zheng, C. J., Yoo, J.-S., Lee, T.-G., Cho, H.-Y., Kim, Y.-H., & Kim, W.-G. (2005). Fatty acid synthesis is a target for antibacterial activity of unsaturated fatty acids. *FEBS Letters*, *579*(23), 5157–62. doi:10.1016/j.febslet.2005.08.028