TECHNISCHE UNIVERSITÄT MÜNCHEN

Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt

Professur für Biotechnologie der Naturstoffe

# Improving an *Escherichia coli* based biocatalyst for the production of small molecule glucosides

Julian Rüdiger

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Karl-Heinz Engel

Prüfer der Dissertation:

1. Prof. Dr. Wilfried Schwab
2. Prof. Dr.-Ing. Dirk Weuster-Botz

Die Dissertation wurde am 28.06.2019 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 11.11.2019 angenommen.

# I. Acknowledgement

I would like to extend my deepest gratitude to Prof. Dr. Wilfried Schwab for giving me the opportunity to work on this project and for his guidance during the project. I am also grateful to Dr. Thomas Hoffmann for being my mentor during my time at the Technical University Munich. Many thanks to all my colleagues who have been supporting me in the laboratory.

I wish to thank Prof. Dr.-Ing. Dirk Weuster-Botz for his constructive advice during the project. I would also like to extend my gratitude to Xenia Priebe for her contributions to this project and her insightful suggestions.

Special thanks to the TUM International Graduate School of Science and Engineering (IGSSE) for funding this project and thereby bringing the Institute of Biochemical Engineering (BVT) and the Associate Professorship of Biotechnology of Natural Products (BINA) together.

I am also grateful to Prof. Thorson for letting me work in his laboratory at the University of Kentucky, and for sharing his insights into directed evolution and high-throughput screening with me. Furthermore, I would like to thank his team for their practical suggestions and their helpful advice. In addition, I would like to extend my sincere thanks to Kelli and Tyler Huber for making my research stay abroad in Lexington, Kentucky more enjoyable.

I am extremely grateful to my friends and family for supporting me from far away and I would like to express my deepest appreciation to Annika Claus for supporting me every day.

# II. Abstract

Aroma molecules are used by different industries because of their pleasant features. These molecules are volatile compounds and usually insoluble in water. Aroma molecules containing an alcohol group can be coupled to a sugar moiety by biotransformation using a whole-cell biocatalysis. The resulting glycoside has a higher water solubility, is no longer volatile and is exported by the cells into the supernatant.

The aim of this study was to modify an *Escherichia coli* based production strain, expressing the plant glucosyltransferase VvGT14a, to increase the yield of small molecule glycosides such as terpenyl glucosides including geranyl glucoside. Therefore, the use of molecular chaperones was explored, different parts of the expression system of VvGT14a were varied, and the enzyme itself was subjected to a directed evolution approach. To test varying expression systems, different vectors were modified and linked to *VvGT14ao*, a codon-optimized version of *VvGT14a*. Directed evolution was done by screening a random mutant library, generated by an error-prone polymerase chain reaction, using two alternative methods, a coupled enzymatic reaction and a method based on liquid chromatography mass spectrometry. The biotransformation efficiencies of the production strains were evaluated with a novel plate-based biotransformation system.

The production strains *E. coli* BL21(DE3)pLysS carrying the plasmids pET29a_VvGT14ao and pGEX-4T-1_VvGT14ao, which produced 0.011 and 0.054 mM of glucoside, respectively, served as references for this work.

The use of molecular chaperones showed mixed results with only the production strain *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao in combination with the chaperone plasmid pGro7 leading to a $10^2$-fold increase in geranyl glucoside concentration.

The variation of the expression system resulted in strain-vector combinations, which produced significant higher glucoside titers. The strains *E. coli* BL21(DE3)pLysS /pET32a_VvGT14ao, *E. coli* BL21(DE3)pLysS /pET-SUMO_VvGT14ao, and *E. coli* BL21 /pGEX-K_VvGT14ao showed 0.046, 0.070, and 0.058 mM of geranyl glucoside, respectively in the supernatants of the biocatalysts. The highest titers were determined for *E. coli* BL21(DE3) /pET29a_VvGT14ao and the strain-vector combination of *E. coli* BL21(DE3) with pET-SUMO_VvGT14ao, which produced concentrations of 0.078 mM and 0.098 mM of the target molecule, respectively.

The mutant screening revealed three variant strains, K60, K43, and K163 producing 0.051, 0.203, and 0.848 mM geranyl glucoside, respectively. Sequencing results of K60 showed that this candidate contained at least three different vectors at different concentrations. Thus, this strain was not further analyzed. The mutant strain K43 expressed the wild type VvGT14a however exhibited several mutations in its genome resulting in increased product titers. The

mutant strain K163 expressed a variant gene coding for the double mutant VvGT14a-S168N-W353R. The positive effects of K43 and K163 were transferrable to other strain-vector combinations. While *E. coli* BL21(DE3) /pET29a_VvGT14ao produced only 0.064 mM of geranyl glucoside, the product yield in the background of K43 and with the double mutation of K163 increased to 0.220 and 0.234 mM, respectively. Both alterations together increased the product yield even further to 0.656 mM of geranyl glucoside. Therefore, the beneficial variations could be combined to further increase the geranyl glucoside yield of *Escherichia coli*.

The mutant library screening also revealed the N-truncated protein VvGT14a-M1_P241del-V431M. Even though the shortened enzyme lacked the acceptor binding domain and no longer produced glycosides, it still acted on UDP-glucose but catalyzed a hydrolase reaction.

The results show that the relationships and interactions of regulatory elements, gene and protein modifications, and supplemental enzymes are very complex and their impact on product yield cannot be reliably predicted. Therefore, it is of the utmost importance to test different expression systems at the beginning of a strain engineering project before considering other alterations.

# III. Zusammenfassung

Aromastoffe werden in der Industrie wegen ihrer angenehmen Eigenschaften eingesetzt. Die Moleküle sind oftmals leichtflüchtige Substanzen und meist schlecht wasserlöslich. Aromamoleküle welche über eine Alkoholfunktion verfügen können durch eine Ganzzell-Biokatalyse mittels Biotransformation mit einem Zucker verbunden werden. Die daraus entstehenden Glykoside weisen eine höhere Wasserlöslichkeit auf, sind nicht mehr leichtflüchtig und werden von den Zellen ins Medium abgegeben.

Das Ziel dieser Studie war es, einen auf *Escherichia coli* basierenden Produktionsstamm zu modifizieren, der die pflanzliche Glukosyltransferase VvGT14a exprimiert, um die Ausbeute an niedermolekularen Glykosiden wie Terpenylglukosiden einschließlich Geranylglukosid zu erhöhen. Dazu wurde der Einsatz von molekularen Chaperonen untersucht, verschiedene Teile des Expressionssystems von VvGT14a variiert und das Enzym selbst mittels gerichteter Evolution verändert. Um verschiedene Expressionssysteme zu testen, wurden verschiedene Vektoren modifiziert und mit dem Codon-optimierten Gen *VvGT14ao* ausgestattet. Zur gerichteten Evolution wurde eine zufällige Mutantenbibliothek, welche mit einer fehleranfälligen Polymerase-Kettenreaktion hergestellt wurde, durchmustert. Zwei alternative Methoden der Durchmusterung wurden verwendet, eine gekoppelte, enzymatische Reaktion und eine Methode basierend auf Flüssigchromatographie-Massenspektrometrie. Die Biotransformationseffizienz der Produktionsstämme wurde mit einem neuartigen plattenbasierten Biotransformationssystem evaluiert.

Die beiden Produktionsstämme *E. coli* BL21(DE3)pLysS /pET29a_VvGT14ao und *E. coli* BL21(DE3)pLysS /pGEX-4T-1_VvGT14ao dienten als Referenzen für diese Arbeit. Sie produzierten 0,011 beziehungsweise 0,054 mM Glukosid.

Die Verwendung von molekularen Chaperonen zeigte uneinheitliche Ergebnisse, nur der Produktionsstamm *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao in Kombination mit dem Chaperon-Plasmid pGro7 produzierte eine $10^2$-fach höhere Konzentration an Geranylglukosid.

Die Variation des Expressionssystems führte zu Stamm-Vektor-Kombinationen, mit denen deutlich höhere Glukosidtiter erzielt werden konnten. Die Stämme *E. coli* BL21(DE3)pLysS /pET32a_VvGT14ao, *E. coli* BL21(DE3)pLysS /pET-SUMO_VvGT14ao und *E. coli* BL21 /pGEX-K_VvGT14ao produzierten 0,046, 0,070 bzw. 0,058 mM Geranylglukosid im Überstand des Biokatalysators. Die höchsten Titer wurden bei *E. coli* BL21(DE3) /pET29a_VvGT14ao und der Stamm-Vektor-Kombination aus *E. coli* BL21(DE3) mit pET-SUMO_VvGT14ao nachgewiesen. Sie produzierten 0,078 bzw. 0,098 mM des Zielmoleküls.

Die Durchmusterung der Mutantenbibliothek brachte drei interessante Variantenstämme hervor, K60, K43 und K163, mit Produktausbeuten von 0,051, 0,203 bzw. 0,848 mM

Geranylglukosid. Die Sequenzierungsergebnisse von K60 zeigten, dass der Stamm mindestens drei verschiedene Plasmide in verschiedenen Konzentrationen enthält. Daher wurde dieser Stamm nicht weiter untersucht. Der Mutanten-Stamm K43 exprimierte die Wildtyp-VvGT14a, enthielt allerdings einige Mutationen im Genom, welche zu den gesteigerten Produkttitern führten. Der Stamm K163 exprimierte eine Genvariante, die für die Doppelmutante VvGT14a-S168N-W353R kodiert. Die positiven Effekte waren auf andere Stamm-Vektor-Kombinationen übertragbar. *E. coli* BL21(DE3) /pET29a_VvGT14ao produzierte 0,064 mM Geranylglukosid. Durch die Veränderungen von K43 und K163 wurde die Ausbeute auf 0,220 beziehungsweise 0,234 mM erhöht. Wurden beide Veränderungen eingebracht, erhöhte sich die Ausbeute auf 0,656 mM Geranylglukosid. Damit konnte der Glukosidtiter durch die Kombination der vorteilhaften Veränderungen weiter erhöht werden.

Die Durchmusterung der Mutantenbibliothek brachte auch das N-ständig verkürzte Protein VvGT14a-M1_P241del-V431M hervor. Obwohl dem verkürzten Enzym die Akzeptorbindedomäne fehlte und es keine Glykoside mehr produzierte, setzte es dennoch UDP-Glukose um, fungierte aber als Hydrolase.

Die Ergebnisse zeigten, dass die Beziehungen und das Zusammenspiel von regulatorischen Elementen, Gen- und Proteinmodifikationen, und Hilfsenzymen sehr komplex sind und ihr Einfluss auf die Produktausbeute nicht zuverlässig vorhergesagt werden kann. Deshalb ist es von größter Bedeutung verschiedene Expressionssysteme zu Beginn eines Stammentwicklungsprojekts zu testen, bevor weitere Änderungen vorgenommen werden.

# IV. Table of contents

# V. List of figures

# VI. List of tables

# VII. List of abbreviations

| Abbr. | Word / phrase |
|---|---|
| AAE | Amino acid exchange |
| ATP | Adenosine triphosphate |
| BB | Backbone |
| CAT | Chloramphenicol acetyltransferase |
| cfu | Colony forming unit |
| ClPNP | 2-Chloro-4-nitrophenol |
| cPCR | Colony PCR |
| CWW | Cell wet weight |
| DM | Double mutant |
| DNA | Deoxyribonucleic acid |
| epPCR | Error-prone PCR |
| EtOH | Ethanol |
| GH | Glycoside hydrolase |
| GP | Glycoside phosphorylase |
| GRAS | Generally recognized as safe |
| GST | Glutathione S-transferase |
| GT | Glycosyltransferase |
| HP | Hitplate 25 |
| HPLC | High-performance liquid chromatography |
| HSP | Heat shock protein |
| IBs | Inclusion bodies |
| LB | Lysogeny broth |
| LIC | Ligation independent cloning |
| LM | Length marker (DNA ladder) |
| MAI | Minimal auto induction |
| MBP | Maltose-binding protein |
| MIC | Minimum inhibitory concentration |
| mRNA | Messenger RNA |
| MS | Mass spectrometry |
| MTP | Microtiter plate |
| OD | Optical density |
| ori | Origin of replication |
| PCR | Polymerase chain reaction |

| | |
|---|---|
| PSPG | Plant secondary product glycosyltransferase |
| RBS | Ribosomal binding site |
| RNA | Ribonucleic acid |
| SD | Shine-Dalgarno |
| SDS-PAGE | Sodium dodecyl sulfate-polyacrylamide gel electrophoresis |
| SN | Serine-asparagine variant |
| SNP | Single-nucleotide polymorphism |
| SOC | Super optimal broth with catabolite repression |
| SUMO | Small ubiquitin-like modifier |
| SuSy | Sucrose synthase |
| TB-AIM | Terrific broth - auto induction medium |
| TG | Transglycosidase |
| tRNA | Transfer RNA |
| TRX | Thioredoxin |
| Ub | Ubiquitin |
| UDR | UDP detection reagent |
| UGT | UDP-glycosyltransferase |
| UV | Ultraviolet |
| WR | Tryptophan-arginine variant |
| WT | Wild type |

# 1 Introduction

The odor of food and beverages plays a huge role for people. Oftentimes, the consumer associates the scent of a food with its quality (Lawless 1991). Countless products are therefore refined with artificial or natural aromas. Many pleasant flavors and fragrances are natural substances and are produced in the secondary metabolism of plants (Hussain et al. 2012). Plant secondary metabolites are substances that are not needed for the growth or development of the plant. However, they are used by plants to interact with their environment. Plants produce compounds to attract insects for pollination, to repel predators, or to warn other plants of hazards (Stevenson et al. 2017, Ueda et al. 2012).

## 1.1 Aroma molecules

Many aroma molecules are terpenoids, but many other molecule classes such as phenylpropanoids, or furanones also contain aroma molecules. Terpenoids consist of isoprene units and therefore contain usually a multiple of five carbon atoms (Abdallah and Quax 2017). The phenylpropanoids derive from the shikimate pathway and have at least one aromatic phenyl group and a carbon-tail of three atoms (Herrmann and Weaver 1999). In strawberry fruit, furanones derive from D-fructose 6-phosphate and contain a furan-like backbone with a ketone group (Pérez et al. 1999). Geraniol and nerol are examples of aroma chemicals with a terpenoid structure. Geraniol and nerol can be found in flowers, such as roses and in some spices (Banthorpe et al. 1972, Francis and Allcock 1969). Nerol has a fresh, rose, and citrus odor, whereas geraniol has an intense floral odor (Chen and Viljoen 2010). Eugenol, a phenylpropanoid, is found in many spices and has a strong clove odor (Imark et al. 2000). Furaneol, a furanone, is found in strawberry fruit and other fruits. Furaneol at low concentrations has a sweet strawberry-like aroma (Zabetakis et al. 1999) while the aliphatic alcohol n-hexanol has a sweet, pleasant odor (Verschueren 2001) and a fruity taste (Furia 1972).

Flavors and fragrances are volatile substances and therefore disappear over time in open containers (Carvalho et al. 2016). Hand creams lose their fragrances because the odor molecules evaporate over a longer period of time. Another challenge is the low water-solubility of many aroma compounds as many fragrances and flavors are used in water-based consumer products. Citronellol, which has a floral, rose, and citrus odor, is only soluble in water up to a concentration of 200 mg/L (Samuel et al. 2003). One way to overcome these problems is the chemical modification of the volatile chemicals by glycosylation. Aroma molecules containing a hydroxyl group (Figure 1), such as terpenols or phenols, can be coupled with a sugar. The sugar moiety conveys solubility to the compounds and makes them less volatile and chemically more stable (Ohgami et al. 2015).

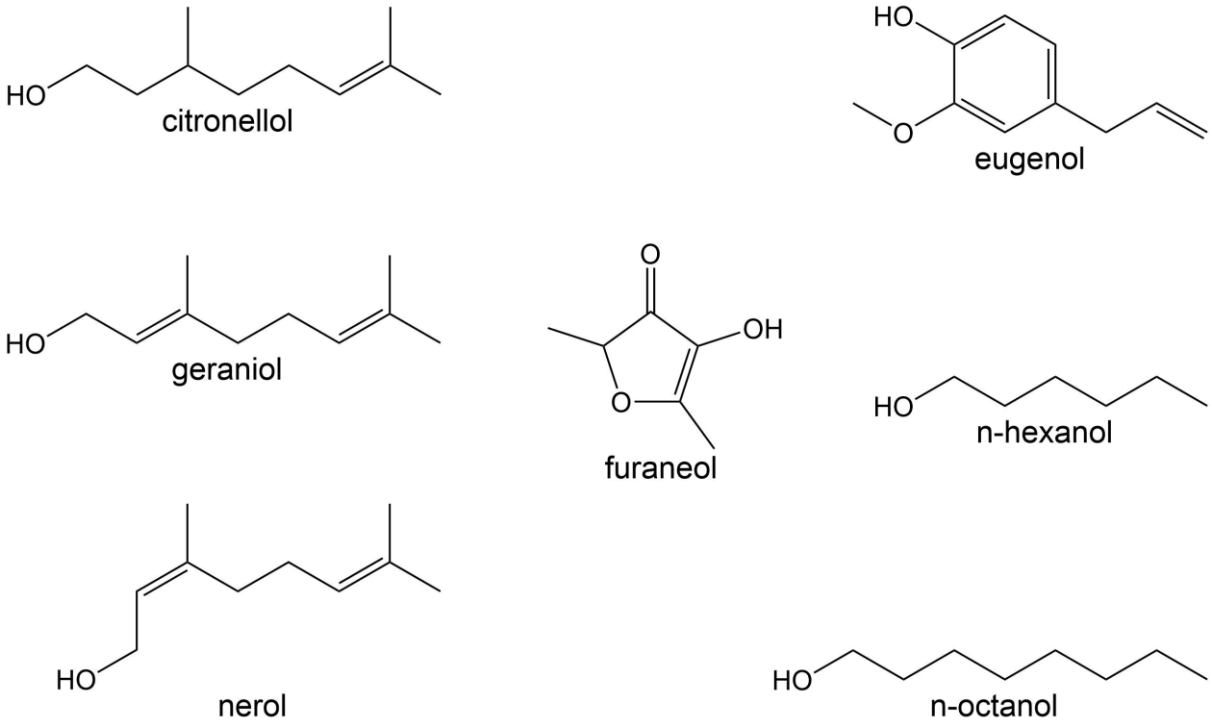


*Figure 1: Selected aroma molecules*

*Citronellol, eugenol, furaneol, geraniol, n-hexanol, nerol, and n-octanol are aroma molecules found in nature. The fragrances also occur as glycosides in nature.*

The quality of many consumer products is determined by their smell. For this reason, flavors and fragrances are used by numerous industries to create pleasant odors that are characteristic of each product group. Many foods and beverages are enhanced by the addition of flavors, while cosmetics, shampoo/body wash, deodorant, detergents, and washer/dryer additives are often perfumed with fragrances to improve consumer experience. Therefore, there are many possible industrial applications for aroma molecules or their glycosides (Goldmann 2018, Schwab et al. 2015).

## 1.2 Glycosides

Glycosides are produced by coupling a sugar (glycone) to an alcohol, amine, or thiol (aglycone) to form a glycosidic bond. The second molecule is also called an acceptor molecule because the sugar was transferred from a donor (usually a nucleotide-diphosphate sugar) to it (Figure 2). Glycosides are naturally occurring substances and are omnipresent in every living organism. The nucleobases in the DNA are linked to the phosphate-sugar backbone by an N-glycosidic bond (Watson and Crick 1953). Bacterial cell walls contain a grid of peptidoglycans called murein. Murein is a macromolecule composed of different amino acids and sugars (Vollmer et al. 2008). Eukaryotic cells use glycosylation as a post-translational modification of proteins. The glycosylation pattern highly influences the efficacy and lifetime of therapeutic proteins (Goettig 2016, Solá and Griebenow 2010). The human body can detoxify substances

by coupling them to glucuronic acid (Zheng et al. 2002), which makes the toxins tangible for urinary elimination (Perreault et al. 2013). Plants use glycosylation for many purposes. They can glycosylate toxins, making them accessible for storage in the vacuole. These toxins can come from the environment (Härtl et al. 2017), or can be produced by the plant itself. Many plants produce toxin precursors and activating enzymes that are able to form the final toxin as part of their defense mechanism against herbivores. Both components are stored in separate compartments. Upon damage by herbivores, the components are combined and the toxin is released. Many plants also store aroma and other plant secondary metabolites as glycosides in the vacuole (Liu et al. 2017, Marlatt et al. 1992, Sarry and Gunata 2004). Glycosides of citronellol, eugenol, furaneol, geraniol, and nerol have been identified, among others, in *Rosa damascena, Vitis vinifera, Fragaria x ananassa, Nepeta cadmea, Zingiber officinale,* and *Camellia sinensis* (Table 1).

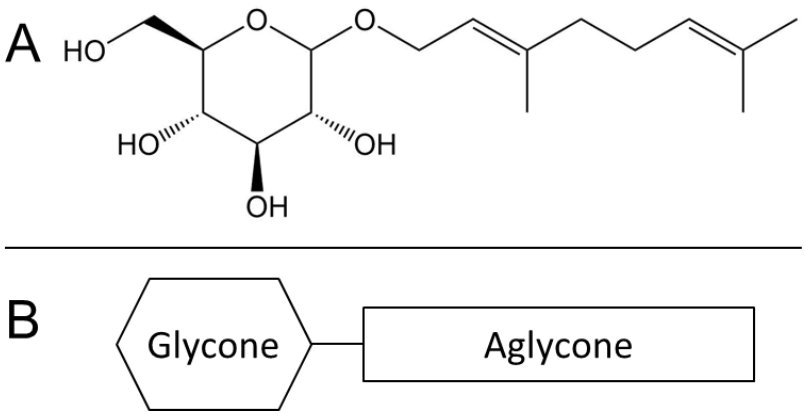


*Figure 2: Example of a small molecule glucoside*

*(A) Geranyl glucoside is composed of geraniol and glucose, which are connected by a glycosidic bond. (B) A schematic representation of a glycoside consisting of an aglycone and a glycone. Many sugars can be used as glycone, the range is not limited to hexose sugars.*

The properties of the aglycone are changed manifold by glycosylation. The aglycone can be made accessible to other enzymes or transporters by glycosylation, which allows further modification, transport, or storage (Shitan 2016). The aglycone gains polarity from the glycone, whereby the solubility is increased. The bioavailability of substances is altered by glycosylation (Kaşıkcı and Bağdatlıoğlu 2016, Moradi et al. 2016) and the effects of drug molecules can be modified (Jain et al. 2012). The addition of different sugar moieties results in a variety of new drug molecules differing in efficacy (Nandurkar et al. 2014, Zhang et al. 2015). Aglycones can be released from the glycoside by many factors such as heat, acid, or enzymes (Herrmann 2007). These effects open up countless opportunities for the use of glycosides in consumer products and therefore make glycosides very interesting for various industries such as food, cosmetics, and pharmaceutical industries.

*Table 1: Naturally occurring glycosides of citronellol, eugenol, furaneol, geraniol, and nerol*

| Glycoside | Occurrence | Literature |
|---|---|---|
| Citronellyl 2-O-β-D-glucopyranosyl-β-D-glucopyranoside | *Rosa damascena* bulgaria | (Oka et al. 1998) |
| Citronellyl β-D-glucoside | *Vitis vinifera* | (Hemingway et al. 1999) |
| Eugenyl 6-O-α-L-arabinofuranosyl-β-D-glucopyranoside | *Rosa damascena* Mill | (Straubinger et al. 1999) |
| Eugenyl 6-O-α-L-arabinopyranosyl-β-D-glucopyranoside | *Rosa damascena* Mill | (Straubinger et al. 1999) |
| Eugenyl 6-O-β-D-xylopyranosyl-β-D-glucopyranoside | *Rosa damascena* Mill | (Straubinger et al. 1999) |
| Eugenyl O-β-D-glucoside | *Nepeta cadmea* | (Takeda et al. 1998) |
| Furaneoyl O-β-D-glucoside | *Fragaria ananassa* | (Mayerl et al. 1989) |
| Geranyl 6-O-α-L-arabinofuranosyl-β-D-glucopyranoside | *Rosa danascena* Bulgaria | (Oka et al. 1997) |
| | *Vitis vinifera* (Muscat of Alexandria) | (Williams et al. 1982) |
| Geranyl 6-O-α-L-arabinopyranosyl-β-D-glucopyranoside | *Zingiber officinale* Roscoe | (Sekiwa et al. 2001a) |
| | *Camellia sinensis* var. sinensis cv. Yabukita | (Nishiktani et al. 1996) |
| Geranyl 6-O-α-L-rhamnopyranosyl-β-D-glucopyranoside | *Vitis vinifera* (Muscat of Alexandria) | (Williams et al. 1982) |
| Geranyl 6-O-β-D-apiofuranosyl-β-D-glucopyranoside | *Zingiber officinale* Roscoe | (Sekiwa et al. 2001a) |
| | *Rosa danascena* bulgaria | (Oka et al. 1997) |
| Geranyl 6-O-β-D-xylopyranosyl-β-D-glucopyranoside | *Zingiber officinale* Roscoe | (Sekiwa et al. 2001b) |
| | *Camellia sinensis* var. sinensis cv. Shuixian | (Guo et al. 1993) |
| Geranyl β-D-glucoside | *Vitis vinifera* | (Hemingway et al. 1999) |
| Neryl 6-O-α-L-arabinofuranosyl-β-D-glucopyranoside | *Vitis vinifera* (Muscat of Alexandria) | (Williams et al. 1982) |
| Neryl 6-O-α-L-rhamnopyranosyl-β-D-glucopyranoside | *Vitis vinifera* (Muscat of Alexandria) | (Williams et al. 1982) |
| Neryl β-D-glucoside | *Vitis vinifera* | (Hemingway et al. 1999) |

### 1.2.1 Chemical production

The chemical production of glycosides can be achieved by using different synthesis strategies (Demchenko 2008). These strategies can be classified according to the form of the glycosyl donor. The glycosyl donors can be (1) alkenyl glycosides, (2) alkynyl glycosides, (3) glycals, (4) glycosyl halides, (5) glycosyl imidates, (6) thioglycosides, or (7) unprotected sugars (Das and Mukhopadhyay 2016). Most donors need to be modified to prevent the formation of

byproducts because they usually display multiple functionalities. The last steps of many chemical syntheses of glycosides are the removal of protecting groups and subsequent product purification.

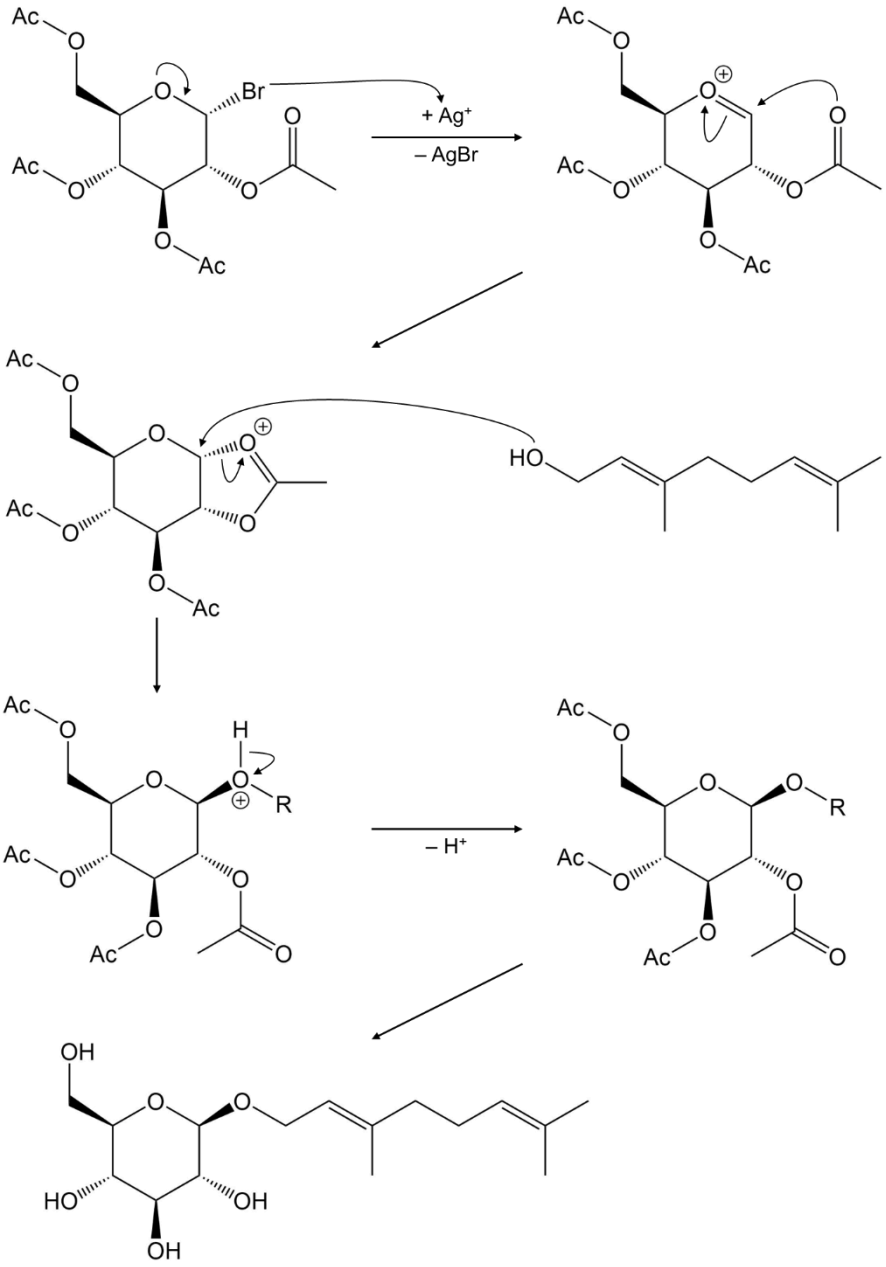


*Figure 3: Chemical synthesis of glycosides by the Koenigs-Knorr reaction*

*The chemists Wilhelm Koenigs and Eduard Knorr discovered the glycosylation reaction in 1901 when they treated alcohols with acetobromglucose and silver(I) carbonate (Königs and Knorr 1901).*

Terpenyl glucosides can be synthesized with the Koenigs-Knorr reaction (Figure 3) in moderate yields. Acetobromglucose (2,3,4,6-Tetra-O-acetyl-α-D-glucopyranosylbromid) reacts with silver(I) oxide, forming silver bromide and an intermediate susceptible for a nucleophilic attack. The oxygen of the hydroxyl group of the terpenol acts as a nucleophile on the C1-carbon of the intermediate forming an oxonium ion. The proton is released and the

terpenyl tetraacetyl glucoside is obtained. Finally, the protection groups are hydrolyzed by barium oxide and the terpenyl glucoside is produced (Ishag et al. 1985).

With the right protective group chemistry, almost all substrates can be glycosylated with a wide variety of sugars. However, low yields and the need to protect substrates with multiple functional groups, make the chemical synthesis labor-intensive and costly (De Bruyn et al. 2015b).

### 1.2.2 Carbohydrate-active enzymes

Glycosides are abundant in nature and there are also countless enzymes that produce said glycosides. Therefore, an alternative for the chemical synthesis of glycosides is to exploit the catalytic capabilities of theses enzymes (Figure 4). There are four major enzyme types, which can be used to produce glycosides: glycoside hydrolases (GHs) (Van Rantwijk et al. 1999), phosphorylases (GPs) (O'Neill and Field 2015), glycosyltransferases (GTs) (Vogt and Jones 2000), and transglycosidases (TGs) (Seibel et al. 2006). Hexyl glucoside was produced with a yield of 58 % (mol/mol) by a glycoside hydrolase (Svasti et al. 2003), hydroquinone-O-α-D-glucopyranoside (α-arbutin; 2.3 g) was synthesized with a glycoside phosphorylase (Kitao and Sekine 1994) and transglycosidases were used to produce resveratrol α-glucosides and isomaltulose with a yield of 50 % and 85 % (mol/mol), respectively (Rose and Kunz 2002, Torres et al. 2011).

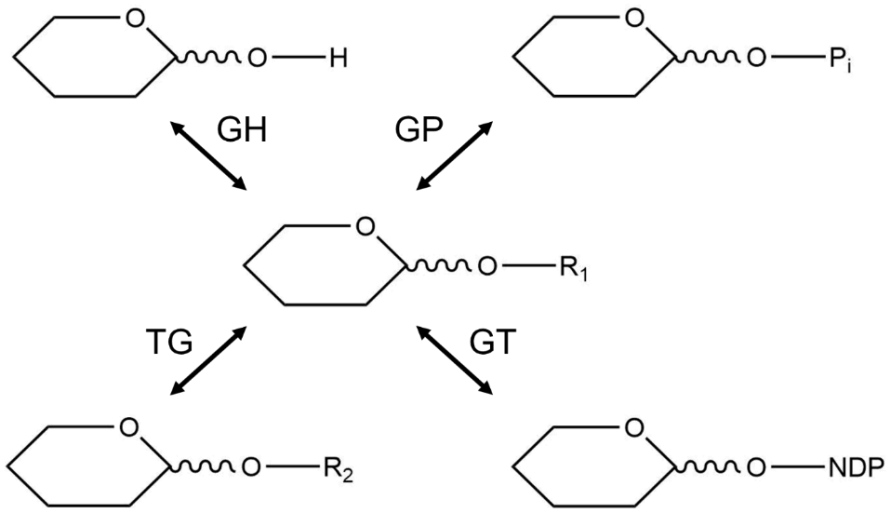


*Figure 4: Types of glycoside-forming enzymes*

*$R_1$ and $R_2$ are aglycones. Glycoside hydrolases (GH) or phosphorylases (GP) produce glycosides by their reverse reactions. Glycosyltransferases (GT) use activated sugars (nucleotide sugars; NDP-sugars) as sugar donors for the glycosylation of substrates. Transglycosidases (TG) transfer the sugar moieties from one glycoside to another acceptor molecule. Adapted from Desmet et al. (2012).*

Since GHs, GPs, and TGs require very high substrate concentrations and specific reaction conditions in order to obtain the desired product, they are mostly used for *in-vitro* glycoside production (Desmet et al. 2012). However, GTs can also produce a wide variety of products at

rather low substrate concentrations (De Bruyn et al. 2015a), therefore GTs are used for the biotechnological glycoside production both *in-vitro* and *in-vivo*.

Uridine-diphosphate (UDP) dependent glycosyltransferases (UGTs) make up the largest group of GTs and use UDP-sugars as glycosyl donors (Yonekura-Sakakibara and Hanada 2011). UGTs share structural similarities and a motive highly conserved in all plant UGTs (Figure 5), the plant secondary product glycosyltransferase (PSPG) box (Vogt and Jones 2000).

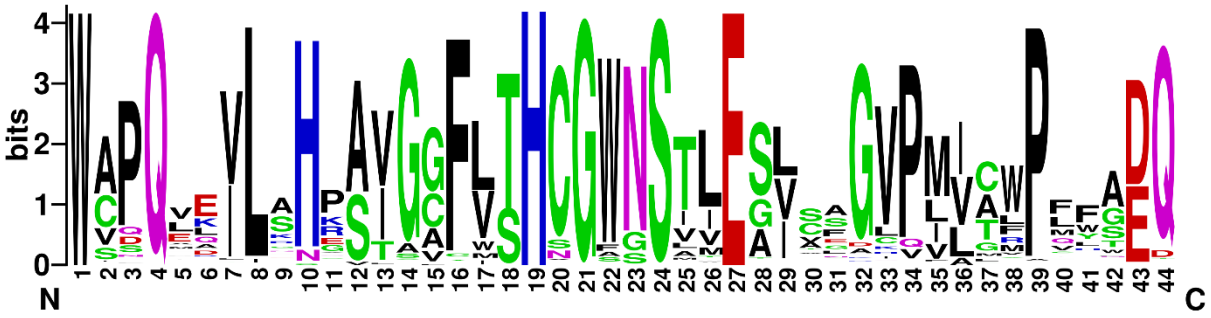


*Figure 5: The plant secondary product glycosyltransferase box*

*The plant secondary product glycosyltransferase (PSPG) box is a highly conserved region in UGTs* near the C-terminus *(Paquette et al. 2009).*

### 1.2.3 Glycoside production using glycosyltransferases

Biotechnological production strategies use the catalytic capabilities of GTs in order to obtain the glycoside product. Enzymes for conversion are produced by heterologous expression, typically in *E. coli* strains, and can either be purified for *in-vitro* transformation (Hyung Ko et al. 2006) or are used in a whole-cell biocatalysis approach (Pandey et al. 2013).

*In-vitro* biotransformations can be tailored to the needs of the enzyme in terms of buffer, pH, substrate concentration, and temperature. However, activated sugars like UDP-glucose, are expensive fine chemicals and therefore need to be regenerated (De Bruyn et al. 2015b). This regeneration can be performed with oligo-sugar synthases, which are driven to carry out the reverse reaction (Figure 6). In such a process, UDP-glucose can be produced from UDP and inexpensive sucrose by the sucrose synthase (Schmolzer et al. 2016). However, the *in-vitro* reaction requires an initial concentration of UDP or UDP-glucose and purified enzyme, which makes the *in-vitro* reactions costly. To reduce the costs, the GTs can be immobilized by different measures in order to reuse them in multiple reaction cycles. The immobilization can be done e.g. by fixation on a superabsorber (Heidlindemann et al. 2014) or by producing the enzymes with special protein tags, which induce the formation of catalytically-active inclusion bodies (Krauss et al. 2017).

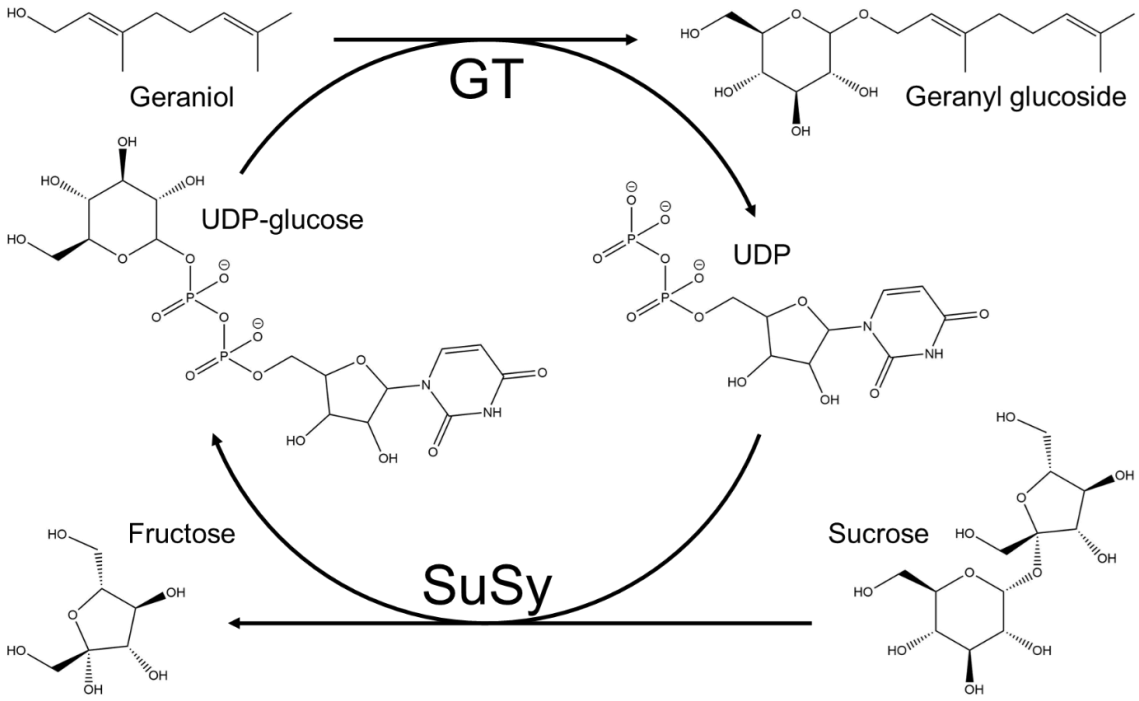


*Figure 6: UDP-glucose regeneration using sucrose synthase*

*Enzymatic glucosylation of geraniol with UDP-glucose regeneration by sucrose synthase (SuSy). Adapted from Schmolzer et al. (2016).*

*In-vivo* biotransformations use heterologously expressed GTs to produce glycosides. The main advantage of a whole-cell biotransformation system is that the UDP-sugar is produced and regenerated by the production host itself. Furthermore, it is advantageous that no enzyme has to be extracted and purified. As with *in-vitro* biotransformation, an enzyme is required that acts on the desired substrate and produces sufficient quantities of the glycoside. With the whole-cell system, however, further restrictions apply. The substrate used must be able to enter the cells and must not be used in toxic concentrations. Many interesting substrates are toxic towards *E. coli* at rather low concentrations. Citronellol has a minimum inhibitory concentration (MIC) of only 0.015 g/L (Lopez-Romero et al. 2015). The MIC of geraniol for *E. coli* BL21(DE3)pLysS is roughly 0.35 g/L (Alexandra Hermenau, personal communication). It is also important that the product is exported from the cells into the culture medium in order to keep the downstream processing simple.

## 1.3 The glycosyltransferase VvGT14a

The enzyme VvGT14a is a glycosyltransferase from *Vitis vinifera* (White Riesling, grape vine) and involved in the glucosylation of terpenols in grapes (Bönisch et al. 2014). The gene was cloned and studied because the expression of the gene *VvGT14a* correlated with the accumulation of terpenyl glycosides in grapes and leaves of different *V. vinifera* cultivars. VvGT14a was later renamed UGT85A69 according to the UGT nomenclature. VvGT15a and VvGT16 were also isolated from *V. vinifera* and associated with glucoside accumulation during

grape ripening. *In-vitro* experiments showed that VvGT14a, VvGT15a, and VvGT16 glucosylate citronellol, geraniol, 1-hexanol, and nerol. VvGT15a and VvGT16 showed catalytic activity towards eight and nine of 22 tested substrates, respectively. However, VvGT14a even glucosylated 18 of the 22 substrates tested, including terpenol, phenols and other small alcohols (Bönisch et al. 2014). Due to this broad substrate spectrum, VvGT14a was chosen as a model enzyme in this study.

## 1.4 Starting point for improvement of a whole-cell biocatalyst

The creation of a whole-cell biocatalyst involves many factors. Starting with the selection of the host organism, further features have to be considered such as the type of heterologous gene expression, modification of the enzyme e.g. with fusion tags, the choice of the enzyme itself, and the use of additional enzymes (Figure 7).

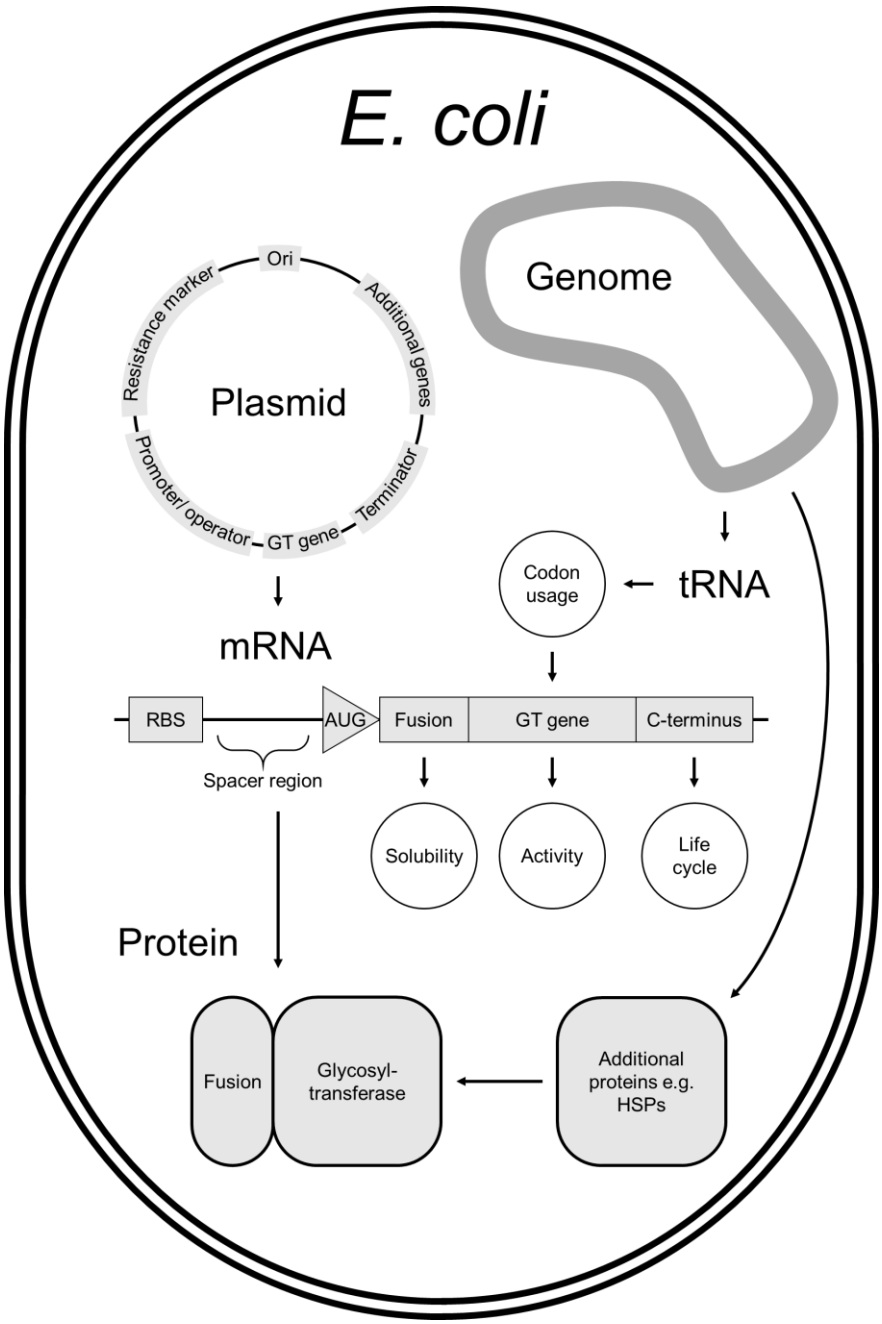


*Figure 7: Features for improving a whole-cell biocatalyst*

*Not only the choice of production host and production enzyme impact the final activity, the different kinds of regulation, fusion partners, codon usage, the life cycle of the GT and additional enzymes like heat shock proteins also strongly influence the activity of the strain.*

*Table 2: Alternative glycoside production hosts*

| Organism | Advantage / Disadvantage | Glycosyltransferase | Product | Reference |
|---|---|---|---|---|
| *Agrobacterium sp.* | + Highly efficient UDP-glucose regeneration system | N/A | N-Acetyllactosamine (7.5 g/L) | (Ruffing et al. 2006) |
| *Pichia pastoris* | + No endotoxin production<br>+ Usually correct folding of heterologous enzymes | bovine α-1,3-galactosyl transferase | α-Galactosyl (α-Gal) trisaccharide (Gal 1,3Gal1,4Glc) | (Shao et al. 2003) |
| *Saccharomyces cerevisiae* | + Co-expression cytochrome P450 reductase | OsCGT (Rice C-GT) | Flavone-C-glycoside (259.3 mg, 79.3 %) | (Brazier-Hicks and Edwards 2013) |
| *Saccharomyces cerevisiae* | + De novo synthesis of Vanillin<br>- Poor respiratory capacity<br>- High ethanol production | UGT72E2 (AtGT) | Vallinin β-D-glucoside (500 mg/L Vallinin glucoside) | (Brochado et al. 2010) |
| *Saccharomyces cerevisiae* | + GRAS status<br>+ Flavonoid glycosides are transported out of the yeast cells<br>- β-glucosidase activity | DicGT4 | Naringenin-7-O-glucoside (13.2 mg/L) | (Werner and Morgan 2009) |
| *Saccharomyces cerevisiae* | + GRAS status<br>+ Large UDP-glucose pool<br>- Poor permeability | UGT76G1 | Rebaudioside A (1.16 g/L) | (Li et al. 2016) |
| *Schizosaccharomyces pombe* | + De novo synthesis of Vanillin<br>+ ACAR activation (aromatic carboxylic acid reductase)<br>- Low yield | UGT72E2 (AtGT) | Vallinin β-D-glucoside (65 mg/L Vallinin)<br>(80 % (mol/mol) Vallinin glucoside) | (Hansen et al. 2009) |

*Escherichia coli* is frequently used as a host organism because it is genetically very well accessible (Rosano and Ceccarelli 2014). This species can be modified with ease, thanks to decades of molecular biological research (Sambrook et al. 1989). Other production hosts can

be utilized as well (Table 2), however byproduct formation, low yields, and product degradation are common problems (Brochado et al. 2010, Hansen et al. 2009, Werner and Morgan 2009). The choice of the specific strain should take into account the construction, regulation, and expression of the GT gene (Rosano and Ceccarelli 2014). *De novo* synthesis can rely on specific enzymes requiring eukaryotic cells, in those cases yeast strains are oftentimes used as production hosts (Brochado et al. 2010, Hansen et al. 2009).

### 1.4.1 Heterologous gene expression

Genes can be introduced into the host either by integration into the genome, or by an extrachromosomal expression plasmid (Martinez-Morales et al. 1999, Rosano and Ceccarelli 2014). The integration into the genome has the advantage that the final production process does not require antibiotics and the gene is relatively stable in the cells (Tyo et al. 2009). The omission of antibiotics reduces the costs of cultivation and causes less burden for the environment (Lenski 1998). However, microbial contaminations can occur more easily. The DNA integration into the genome can be done via attB-attP-site integration at very specific entry points (St-Pierre et al. 2013) or via recombination at virtually any site (Sabri et al. 2013). The gene of interest must be equipped with a suitable promoter, a ribosomal binding site (RBS), and a terminator (Iverson et al. 2016). These elements are already present on most expression vectors (Joseph et al. 2015). The introduction of the GT gene on an extrachromosomal expression vector enables the overexpression through strong promoters and a higher gene dosage. There are many expression vectors available differing in origin of replication (ori), selective marker, promoter, RBS, fusion tags, and terminator (Terpe 2006). In particular, the T7 promoter influences the host strain selection enormously (Joseph et al. 2015).

### 1.4.2 Regulation of expression

The regulation of the target gene strongly influences the final enzyme concentration inside the cell and the resulting activity of the production strain. The promoter can vary in two main attributes, the operator, and the strength of expression. Promoters can include an operator, which enables an inducible expression. Many promoters utilize the operator of the lac operon, which is repressed by the lac repressor encoded by *lacI*. The repressor disengages when lactose or isopropyl β-D-1-thiogalactopyranoside (IPTG) is introduced to the medium, allowing expression of the target gene (Ullmann 2009). The strength of the promoter is another factor that influences protein production through transcription. RNA polymerases have preferred DNA binding sequences and bind better to promoter sequences, which are similar to this sequence. A greater similarity leads to a stronger binding energy and thus a higher expression and more mRNA molecules (Brewster et al. 2012).

### 1.4.3 Regulation of translation

The efficiency of the translation heavily depends on two factors, the ribosome binding site (RBS), and the codon usage of the strain. The upstream region of the mRNA is responsible for ribosome recruitment, which is highly dependent on the adenine content (Laursen et al. 2005). The translation initiation is dependent on the similarity of the Shine-Dalgarno (SD) sequence to the anti-SD sequence of the ribosome. If the complementarity is too high, the binding is too strong, which hinders the translation start. However, too low complementarity results in little or no translation. Therefore, a decent sequence fit is required for an efficient initiation (De Boer and Hui 1990).

### 1.4.4 Codon Usage

The codon usage is the other important factor for efficient translation (Robinson et al. 1984). The genetic code offers multiple codons for all amino acids except for methionine and tryptophan. Each codon requires its own tRNA with the corresponding anti-codon. The tRNAs of highly used codons are more abundant in cells than tRNAs of rarely used codons. The size of the individual tRNA pools varies strongly between different domains (Tats et al. 2008). The pool sizes can even vary between closely related species. Therefore, the heterologous expression of genes may be impeded by a different codon bias. The excessive use of rare codons results in a slow translation or even an early termination of the translation. The translation can be improved by adjusting difficult codons or recreating the whole gene. The codon-optimized sequence should reflect the tRNA pool sizes of the host organism and its codon usage (Puigbo et al. 2007). A codon-optimized version of *VvGT14a* (*VvGT14ao*) was available at the beginning of this study.

### 1.4.5 Protein modification

Protein can be modified on its N- or C-terminus to increase the concentration of active enzyme inside the cell (Malhotra 2009). The protein can be combined with a fusion protein that conveys a higher water solubility. Commonly used protein tags include the glutathione S-transferase (GST), maltose-binding protein (MBP), NusA protein (NusA), small ubiquitin-related modifier (SUMO), thioredoxin (TRX), and ubiquitin (Ub) (Costa et al. 2014). Some fusion proteins are only effective if they are located at the N-terminus of the target protein e.g. NusA (Terpe 2003). The other end of the protein, the C-terminus, determines the turnover time of the protein inside the *E. coli* cells. Charged and polar amino acids within the last five residues of a protein significantly extend its half-life (Parsell et al. 1990). In this way, enzymes can be protected from proteolytic degradation.

### 1.4.6 Choice of enzyme

The choice of the enzyme used in the biocatalyst is also very important and depends on the substrates to be converted. *In-vitro* data of biochemically characterized enzymes can be used

to preselect enzymes. However, they still need to be tested *in-vivo* for their suitability to determine the best one, as the *in-vitro* parameters usually differ from the actual *in-vivo* parameters. Available GT libraries can also be screened with novel substrates to find enzymes with the desired activity (Brazier-Hicks et al. 2018).

### 1.4.7 Directed evolution

The application of enzymes is often limited by their lack of activity towards specific substrates or their promiscuity. These hurdles can be tackled by a technique which is gaining more and more importance, the so-called directed evolution (Arnold 2018). Directed evolution is composed of three stages: (1) generation of mutants, (2) screening of mutants, and (3) combination of mutations (Gantt et al. 2013). The mutant libraries are produced by error-prone PCR (epPCR) or other genetic techniques and should display a diverse mixture of mutants. The epPCR can span over a whole gene, or only over a pre-selected region. Hotspots can be identified through the comparison to homologous enzymes or through previous rounds of mutagenesis (Turner 2009). In order to obtain mutants displaying beneficial traits, a screening method is needed for the specific activity. It is favorably, when the product of the target reaction can be detected directly by photometry or fluorometry. In order to evaluate the mutants quickly, the target reaction is often coupled to a better detectable, second reaction, when neither the product nor the substrate are easily quantifiable (Gantt and Thorson 2012). This is necessary for high-throughput screening, which is needed in order to identify beneficial mutants in the vast library. Other methods of evaluation are also possible, however the screening process is the most time consuming step during directed evolution even with a high-throughput screening method. Finally, beneficial mutations can be analyzed and combined, in order to test their compatibility. The final gene can be used as template for further cycles of directed evolution (Arnold and Volkov 1999).

### 1.4.8 Utilization of heat shock proteins

The concentration of correctly folded, active enzyme can also be further improved by co-expression of heat shock proteins (HSPs) or chaperones. These proteins promote the production of active enzymes and even the degradation of misfolded ones (Meyer and Baker 2011). This helps to prevent the formation of inactive protein aggregates called inclusion bodies (IBs). Heat shock proteins are usually expressed when the organism experiences stresses like extreme temperatures, heavy metals, toxic substances, and UV radiation (Urban-Chmiel et al. 2013). The HSP production can be induced by adding ethanol or other irritating substances to the culture medium. The chaperones can also be heterologously expressed from additional plasmids without applying extra stresses (Nishihara et al. 1998, Thomas and Baneyx 1997).

The molecular chaperones and heat-shock-proteins GroES and GroEL are able to promote proper folding to misfolded proteins (Ashraf et al. 2017). Two GroEL heptamers form a barrel-like structure which displays hydrophobic residues near the openings. Misfolded proteins also display hydrophobic residues on their surface, which are hidden in correctly folded proteins, this allows them to bind to the GroEL complex. The additional binding of ATP to the GroEL complex induces a conformational change, which pulls the bound protein inside the chaperone barrel. Next a GroES heptamer binds to the barrel, closing the cavity with the misfolded protein. The inside of the chaperone barrel displays highly hydrophilic residues. This promotes the hiding of hydrophobic residues of the misfolded protein and encourages its re-folding. The hydrolysis of the bound ATP and the binding of another misfolded protein on the other side of the chaperone barrel induces another conformational change, which detaches the GroES heptamer and also releases the captured protein (Horwich et al. 2007). The co-expression of GroES/GroEL and human p50csk protein-tyrosine kinase in *E. coli* yielded 250 mg of pure, active enzyme from 100 g of cells (Amrein et al. 1995). Likewise, the utilization of the chaperones DnaJ, DnaK, and GrpE increased the expression of soluble anti-CD20 scFv by 50 % (Yousefi et al. 2018).

## 1.5 Objective

The objective of this work was to improve an *E. coli* based biocatalyst for the production of small molecule glycosides by varying different parts of the expression system. The glycosyltransferase gene *VvGT14a* from *Vitis vinifera* in its codon-optimized form (*VvGT14ao*) was used as catalytically active protein for the optimization of the biocatalyst and the glycosylation process (Schmideder et al. 2016). The production strains *E. coli* BL21(DE3)pLysS /pET29a-VvGT14ao and *E. coli* BL21(DE3)pLysS /pGEX-4T-1_VvGT14ao were already available (Bönisch et al. 2014, Schmideder et al. 2016) and were therefore used as starting point and as a reference.

In order to increase the glucoside yield of the whole-cell biocatalysis, several variants of the production strains were created and tested. Promising variations were combined and tested as well. In addition to the *E. coli* reference strains other *E. coli* strains were explored as alternatives. Since the host's ability to glycosylate the substrate depends on the presence of correctly formed active enzyme, the influence of several molecular chaperones was tested. The use of different fusion proteins and expression vectors also influences the amount of correctly formed protein, therefore, the expression system of *VvGT14ao* was varied by using different expression vectors in order to find the most suitable fusion protein or rather vector for this application. Directed evolution has proven to be a powerful tool for increasing the fitness of enzymes for various applications, and glycosyltransferases have also been subjected to directed evolution (Arnold 2018). Therefore, VvGT14a was subjected to a directed evolution

approach to increase the product yield of small molecule glucosides of the whole-cell biotransformation process.

# 2 Material and methods

## 2.1 Resources

### 2.1.1 Chemicals

*Table 3: Chemicals*

| Chemical name | Synonym | Supplier |
|---|---|---|
| **1-Hexanol** | n-Hexanol | Fluka |
| **1-Octanol** | n-Octanol | Carl Roth |
| **2-Chloro-4-nitrophenyl-beta-D-gluco-pyranoside** | ClPNP-glucose | University of Kentucky College of Pharmacy |
| **2-Mercaptoethanol** | 2-ME | Carl Roth |
| **2x Laemmli Sample Buffer** | | Bio-Rad |
| **30 % Acrylamide/bisacrylamide (37.5:1)** | Rotiphorese® Gel 30 | Carl Roth |
| **4-(2-Hydroxyethyl)-1-piperazine-ethanesulfonic acid** | HEPES | Carl Roth |
| **5,5'-Dithiobis(2-nitrobenzoic acid)** | DTNB | Sigma-Aldrich |
| **5-Bromo-4-chloro-3-indolyl phosphate** | BCIP | Carl Roth |
| **Acetic acid** | | Carl Roth |
| **Acetyl coenzyme A** | Acetyl-CoA | AppliChem |
| **Agar-agar** | Agar | Carl Roth |
| **Agarose** | | Sigma-Aldrich |
| **Ammonium chloride** | $NH_4Cl$ | Carl Roth |
| **Ammonium persulfate** | APS | Carl Roth |
| **Ammonium sulfate** | $(NH_4)_2SO_4$ | AppliChem |
| **Ampicillin** | Amp | Carl Roth |
| **Anti-Goat IgG – Alkaline Phosphatase antibody** | | Sigma-Aldrich |
| **Anti-GST Antibody** | | Amersham |
| **Benzyl alcohol** | BA | Riedel-de Haën |
| **Biotin** | Vitamin B7 | Calbiochem |
| **Boric acid** | $H_3BO_3$ | Merck |
| **Bovine serum albumin** | BSA | Sigma-Aldrich |
| **BugBuster®** | | Merck |
| **Calcium chloride** | $CaCl_2$ | Carl Roth |
| **Chloramphenicol** | Chl | Carl Roth |
| **Citronellol** | | Sigma-Aldrich |

| | | |
|---|---|---|
| **Citronellyl beta-D-glucopyranoside** | Citronellyl glucoside | BINA |
| **Cobalt(II) chloride hexahydrate** | $CoCl_2 * 6 H_2O$ | Carl Roth |
| **Coomassie Brilliant Blue G-250** | | Carl Roth |
| **Copper(II) chloride dihydrate** | $CuCl_2 * 2 H_2O$ | Carl Roth |
| **D-Glucose** | Glc | Carl Roth |
| **Dimethylformamide** | DMF | Carl Roth |
| **Dimethyl sulfoxide** | DMSO | Merck |
| **Disodium phosphate dihydrate** | $Na_2HPO_4 * 2 H_2O$ | Carl Roth |
| **Ethanol** | EtOH | VWR |
| **Ethylenediaminetetraacetic acid** | EDTA | Carl Roth |
| **Eugenol** | | Fluka |
| **Eugenyl beta-D-glucopyranoside** | Eugenyl glucoside | BINA |
| **Formic acid** | HCOOH | Carl Roth |
| **Furaneol** | | Fluka |
| **Geraniol** | | Carl Roth |
| **Geranyl beta-D-glucopyranoside** | Geranyl glucoside | BINA |
| **Glutathione** | GSH | Carl Roth |
| **Glycerol** | | Carl Roth |
| **Glycine** | Gly | Carl Roth |
| **GST•Bind™ Resin** | GST resin | Merck |
| **Hydrogen chloride** | HCl | Carl Roth |
| **Imidazole** | | Carl Roth |
| **Iron(III) chloride** | $FeCl_3$ | Merck |
| **Isopropyl β-D-1-thiogalactopyranoside** | IPTG | Carl Roth |
| **Kanamycin** | Kan | Carl Roth |
| **Lactose monohydrate** | | Carl Roth |
| **L-Arabinose** | Ara | Fluka |
| **Magnesium chloride hexahydrate** | $MgCl_2 * 6 H_2O$ | ACROS Organics |
| **Magnesium sulfate** | $MgSO_4$ | Carl Roth |
| **Manganese(II) chloride** | $MnCl_2$ | ACROS Organics |
| **Methanol** | MeOH | Riedel-de Haën |
| **Monopotassium phosphate** | $KH_2PO_4$ | Carl Roth |
| **Nerol** | | Sigma-Aldrich |
| **Neryl beta-D-glucopyranoside** | Neryl glucoside | BINA |
| **Nitro blue tetrazolium chloride** | NBT | Carl Roth |
| **Phenylmethylsulfonyl fluoride** | PMSF | Sigma-Aldrich |
| **Polysorbate 20** | Tween 20 | Carl Roth |

| | | |
|---|---|---|
| **Potassium chloride** | KCl | Carl Roth |
| **Profinity™ IMAC Resin, Ni-charged** | His resin | Bio-Rad |
| **Roti®-GelStain** | Roti-Safe | Carl Roth |
| **Roti®-Nanoquant** | | Carl Roth |
| **Sodium chloride** | NaCl | Carl Roth |
| **Sodium dodecyl sulfate** | SDS | Carl Roth |
| **Sodium hydroxide** | NaOH | Carl Roth |
| **Tetracycline** | Tet | Sigma-Aldrich |
| **Tetramethylethylenediamine** | TEMED | Carl Roth |
| **Thiamine** | Vitamin B1 | Sigma-Aldrich |
| **Tris(hydroxymethyl)aminomethane** | Tris | Carl Roth |
| **Tryptone** | | Carl Roth |
| **Uridine diphosphate** | UDP | Promega |
| **Uridine diphosphate alpha-D-glucose** | UDP-glucose | Sigma-Aldrich |
| **Yeast extract** | YE | Carl Roth |
| **Zinc chloride** | $ZnCl_2$ | Carl Roth |

### 2.1.2 Consumables

*Table 4: Consumables*

| Function | Name / model | Manufacturer |
|---|---|---|
| **Blotting membrane** | Roti®-PVDF, Pore size 0.45 µm | Carl Roth |
| **Cell spreader** | L-shaped bacterial cell spreaders | Carl Roth |
| **Centrifuge tubes** | 15 mL tubes | Sarstedt |
| **Centrifuge tubes** | 50 mL centrifuge tubes | Carl Roth |
| **Culture plates** | Disposable petri dishes | greiner bio-one |
| **Cuvettes** | Rotilabo®-single-use cells, PMMA | Carl Roth |
| **Enzyme spatulas** | Disposable plastic spatulas Micro | Carl Roth |
| **Filter paper** | Rotilabo®-Blotting Papers, Thickness 0.35 mm | Carl Roth |
| **Gloves** | Nitril® NextGen® | Meditrade® |
| **Grid sticker** | PetriSticker™ Square grid 100 | Diversify Biotech |
| **HPLC vial caps** | Screw Caps 9 mm with PTFE/Butylrubber Septa | WICOM |
| **HPLC vials** | 2 mL Screw Vials 12*32 | WICOM |
| **Insert** | 15 mm tip, 0.2 mL | Macherey-Nagel |
| **Micro reaction tubes** | Micro Tube 1.5 mL SafeSeal | Sarstedt |

| | | |
|---|---|---|
| **Micro reaction tubes** | SafeSeal micro tube 2 mL, PP | Sarstedt |
| **Microtiter plates** | 384-well Low Volume White Round Bottom PS NBS | Corning® |
| **Microtiter plates** | hitplate 25 | HJ-BIOANALYTIK |
| **Microtiter plates** | Nunc™ MicroWell™ F96 Plates | Thermo Scientific |
| **PCR tubes** | Multiply®-Pro 0.2 mL Biosphere® | Sarstedt |
| **Pipette tips** | 10 µL, clear | Omnitip™ |
| **Pipette tips** | 200 µL, Yellow | Omnitip™ |
| **Pipette tips** | Combitips advanced® | eppendorf |
| **Pipette tips** | Pipette Tips 50-1000 µL | Brand |
| **Plastic bags** | Disposable Bags | Carl Roth |
| **Sealing film** | PARAFILM® M | Pechiney Plastic Packaging |
| **Sealing foil** | Nonwoven foil | HJ-BIOANALYTIK |
| **Serological pipettes** | CELLSTAR® Serological Pipettes | greiner bio-one |
| **Syringe filters** | Rotilabo®-syringe filters, sterile | Carl Roth |
| **Syringes** | 2 mL luer | AMEFA |
| **Syringes** | Omnifix® | B.Braun |
| **Weighing paper** | Weighing paper | Macherey-Nagel |

### 2.1.3 Equipment

*Table 5: Devices*

| Function | Model | Manufacturer |
|---|---|---|
| **Autoclave** | FV13/11 | Fedegari Autoklaven |
| **Bunsen burner** | gasprofi 1 SCS | WLD-TEC |
| **Bunsen burner** | gasprofi 2 SCS accu | WLD-TEC |
| **Centrifuge** | 1-14 | Sigma |
| **Centrifuge** | 2-5 | Sigma |
| **Centrifuge** | 2K15 | Sigma |
| **Centrifuge** | 4K15 | Sigma |
| **Centrifuge** | 5145R | eppendorf |
| **Centrifuge** | miniSpin | eppendorf |
| **Centrifuge** | miniSpin plus | eppendorf |
| **Clean bench** | HERAsafe | Heraeus |
| **Dispenser** | seripettor® | BRAND |
| **Drying cabinet** | ULE 700 | Memmert |
| **Electrophoresis unit** | MIDI PLUS Electrophoresis Unit | Carl Roth |

| | | |
|---|---|---|
| **Electrophoresis unit** | Midi-l Bufferflow | neoLab |
| **Electrophoresis unit** | Rotiphorese® Unit PROfessional II | Carl Roth |
| **Electroporator** | MicroPulser™ | Bio-Rad |
| **Freezer** | Energiesparer | Privileg |
| **Freezer** | GTS 2112 | Liebherr |
| **Fridge** | Export 332 | Kirsch |
| **Fridge** | MediLine | Liebherr |
| **Gel documentation** | G:Box | Syngene |
| **Ice machine** | AF80 | Scotsman |
| **Incubator** | ThermoForma | Thermo Electron Corporation |
| **Laboratory dishwasher** | G 7783 CD MIELABOR | Miele |
| **Magnetic stirrer with heating plate** | MR 2002 | Heidolph |
| **Microwave** | R-4V10 | Sharp |
| **Mini centrifuge** | Sprout® | Heathrow Scientific |
| **PCR cycler** | Labcycler | SensoQuest |
| **pH meter** | pH 50+ DHS | Dostmann |
| **Photometer** | BioPhotometer | eppendorf |
| **Photometer** | DU®-62 Spectrophotometer | Beckman |
| **Photometer** | NICOLET evolution 100 | Thermo Electron Corporation |
| **Pipette** | DISCOVERY Comfort | HTL Lab Solutions |
| **Pipette** | Multipette® E3 | eppendorf |
| **Pipette** | Research | eppendorf |
| **Pipette** | Research plus | eppendorf |
| **Pipette** | Xplorer | eppendorf |
| **Plate reader** | CLARIOstar | BMG Labtech |
| **Power supply** | E132 | CONSORT |
| **Power supply** | E835 | CONSORT |
| **Power supply** | EV245 | CONSORT |
| **Power supply** | PowerPac 200 | Bio-Rad |
| **Rotary Mixer** | Rotary Mixer | LABINCO |
| **Rotor** | 11030 | Sigma |
| **Rotor** | 11150 | Sigma |
| **Rotor** | 12094 | Sigma |
| **Rotor** | 12169-H | Sigma |

| | | |
|---|---|---|
| **Rotor** | FA-45-12-11 | eppendorf |
| **Rotor** | FA-45-24-11 | eppendorf |
| **Scale** | 125A | Precisa |
| **Scale** | 1702004 | Sartorius |
| **Scale** | Scout Pro | Ohaus |
| **Scale** | SPB61 | SCALTEC |
| **SDS-PAGE unit** | MINI Vertical Dual Plate Electrophoresis Unit | Carl Roth |
| **Semi-Dry-Blotter** | V10-SDB | Scie Plas |
| **Shaker** | DOS-10L | neoLab |
| **Shaker** | Orbital-Shaker | neoLab |
| **Shaker** | Polymax 1040 | Heidolph |
| **Shaking Incubator** | 3033 | GFL |
| **Sonificator** | GM 2070 | BANDELIN |
| **Sonotrode** | MS 73 | BANDELIN |
| **Thermomixer** | Digital Heating Shaking Drybath | Thermo Scientific |
| **Thermomixer** | Thermomixer comfort | eppendorf |
| **Ultra low freezer** | MDF-794 | Sanyo |
| **Ultrasound transducer** | UW 2070 | BANDELIN |
| **Vacuum sealer** | 9937 | Finether |
| **Vortex mixer** | 2x³ | VELP Scientifica |
| **Vortex mixer** | Vortex Genie 2 | Scientific Industries |
| **Vortex mixer** | VV3 | VWR |
| **Water bath** | HC 5 | Julabo |
| **Water purification system** | Purelab classic | ELGA |

### 2.1.4 Software

*Table 6: Software*

| Name | Function | Company |
|---|---|---|
| **7-Zip** | File archiver | Igor Pavlov |
| **Adobe Acrobat Reader DC** | PDF reader | Adobe Systems, USA |
| **Adobe Illustrator** | Vector graphics software | Adobe Systems |
| **Agilent ChemStation** | HPLC – Control | Agilent Technologies |
| **ChemDraw Professional 16.0** | Chemical Drawing | PerkinElmer |
| **ChromPass** | HPLC – JASCO | Jasco |
| **CLARIOstar** | Plate reader control | BMG LABTECH |

| | | |
|---|---|---|
| **DataAnalysis for 6300 Series Ion Trap LC/MS Version 4.0 (Build 234)** | HPLC-MS – Analysis | Bruker Daltonik |
| **EndNote 7.8** | Reference management software | Thomson Reuters |
| **FileZilla** | FTP client | FileZilla |
| **Fusion 360** | 3D CAD/CAM design | Autodesk |
| **ImageJ** | Image Analysis | National Institutes of Health |
| **Inkscape** | Vector graphics software | Inkscape |
| **Microsoft Access** | Database management system | Microsoft Corporation |
| **Microsoft Excel** | Spreadsheet editor | |
| **Microsoft Outlook** | Personal information manager | |
| **Microsoft PowerPoint** | Presentation program | |
| **Microsoft Word** | Word processor | |
| **NanoDrop 1000** | NanoDrop software | Thermo Fisher Scientific |
| **Paint.NET** | Image processing software | dotPDN |
| **pDRAW** | Molecular Biology | AcaClone software |
| **Serial Cloner 2.6.1** | Molecular Biology | SerialBasics |
| **Snipping Tool** | Screenshot tool | Microsoft Corporation |
| **TeamViewer 14** | Remote control software | TeamViewer |

### 2.1.5 Enzymes

*Table 7: DNA-modifying enzymes*

| **Name** | **Supplier** |
|---|---|
| **FastAP Thermosensitive Alkaline Phosphatase** | Thermo Fisher Scientific |
| **FastDigest *Aat*II** | Thermo Fisher Scientific |
| **FastDigest *Bam*HI** | Thermo Fisher Scientific |
| **FastDigest *Bsa*I** | Thermo Fisher Scientific |
| **FastDigest *Eco*31I** | Thermo Fisher Scientific |
| **FastDigest *Eco*RI** | Thermo Fisher Scientific |
| **FastDigest *Kpn*I** | Thermo Fisher Scientific |
| **FastDigest *Ssp*I** | Thermo Fisher Scientific |
| **FastDigest *Xho*I** | Thermo Fisher Scientific |
| **Platinum II Taq Hot-Start DNA Polymerase** | Thermo Fisher Scientific |

| | |
|---|---|
| **Platinum SuperFi DNA Polymerase** | Invitrogen |
| **Q5 DNA Polymerase** | New England Biolabs |
| **T4 DNA Ligase** | Promega |
| **T4 DNA Polymerase** | Thermo Fisher Scientific |

### 2.1.6 Primers

*Table 8: Primers*

| Name | Sequence 5' → 3' | Function |
|---|---|---|
| **GT14ao-Bam_+1_fw** | ATAGGATCCGATGGGTAGTATGGAGAAGCC | *VvGT14ao* for pRSFDuet-1 |
| **GT14ao-BamHI_fwd** | ATAGGATCCATGGGTAGTATGGAGAAGCC | *VvGT14ao* |
| **GT14ao-NotI_rev** | TATGCGGCCGCTTACAGCAGTACCTGCTCGA | *VvGT14ao* |
| **p32aK_KanR2_Fw** | GGTCTCTTTATGCTGCATGCCTATTTGTTTAT | *KanR2* for pET32a + *Eco*31I |
| **p32aK_KanR2_Rv** | GGTCTCAAGTTGCAAGTGGCACTTTTCG | *KanR2* for pET32a + *Eco*31I |
| **pET_LIC_14ao_fw** | TACTTCCAATCCAATGCAATGGGTAGTATGGAGAAGCC | *VvGT14ao* LIC |
| **pET_LIC_14ao_rv** | TTATCCACTTCCAATGTTATTATTACAGCAGTACCTGCTCG | *VvGT14ao* LIC |
| **pET32a_BB_Fw** | GGTCTCAATAAACAAATAGGGGTTCCG | pET32a Backbone + *Eco*31I |
| **pET32a_BB_Rv** | GGTCTCTAACTGTCAGACCAAGTTTAC | pET32a Backbone + *Eco*31I |
| **PrKR2.1_AatII** | ATCGACGTCTTCAAATATGTATCCGCTCATGAGA | *KanR2* for pGEX + *Aat*II |
| **PrKR2.2_BsaI** | ATCACCGCGAGACCTTAGAAAAACTCATCGAGCATCAAA | *KanR2* for pGEX + *Bsa*I |
| **GM2_RSF-14ao_fwd** | CACAGCCAGGATCCGATG | Primer for epPCR |
| **GM2_RSF-14ao_rev** | AAGCATTATGCGGCCGCTTA | Primer for epPCR |
| **SDM-VvGT14ao-SN-fd** | CACCGCTGAAAGACGAAAACTATCTGACCAATGGC | Mutagenesis primer S168N |
| **SDM-VvGT14ao-SN-rv** | GCCATTGGTCAGATAGTTTTCGTCTTTCAGCGGTG | Mutagenesis primer S168N |
| **SDM-VvGT14ao-WR-fw** | GTGGCCTGTTAGCTGGTCGGTGCCCACAGGAACAG | Mutagenesis primer W353R |
| **SDM-VvGT14ao-WR-rv** | CTGTTCCTGTGGGCACCGACCAGCTAACAGGCCAC | Mutagenesis primer W353R |
| **VvGT15a-Fwd** | CGTGGATCCATGCCGAGGAGAGTGGTAC | *VvGT15a* |
| **VvGT15a-SN-fwd** | CACATTCCCCTGCAAGGCAACACCCTACATGATCCAGTG | Mutagenesis primer S152N |
| **VvGT15a-SN-rev** | CACTGGATCATGTAGGGTGTTGCCTTGCAGGGGAATGTG | Mutagenesis primer S152N |
| **VvGT15a-WR-fwd** | GAAAGATGCCACATTGTGAAACGGGCACCCCAGAAGGAAG | Mutagenesis primer W321R |

| | | |
|---|---|---|
| **VvGT15a-WR-rev** | CTTCCTTCTGGGGTGCCCGTTTCACAATGTGGCATCTTTC | Mutagenesis primer W321R |
| **VvGT15a-Rev** | GATGCGGCCGCCTAGAATGAAGAGATATACTC | *VvGT15a* |
| **VvGT16-Fwd** | CGTGGATCCATGGGTGATAAGCCTCATG | *VvGT16* |
| **VvGT16-SN-fwd** | GTCCCACTAAAAGATTTGAACTATCTCACAAATGGGTATTTG | Mutagenesis primer S170N |
| **VvGT16-SN-rev** | CAAATACCCATTTGTGAGATAGTTCAAATCTTTTAGTGGGAC | Mutagenesis primer S170N |
| **VvGT16-WR-fwd** | GAGGCCTACTGGCGGACCGGTGCCCGCAAGAGAAAG | Mutagenesis primer W354R |
| **VvGT16-WR-rev** | CTTTCTCTTGCGGGCACCGGTCCGCCAGTAGGCCTC | Mutagenesis primer W354R |
| **VvGT16-Rev** | GATGCGGCCGCTTATGGTTTTGTAAGCAGTATGTC | *VvGT16* |
| **Amp-fwd** | ATCGGTACCAACCCCTATTTGTTTATTTTTCTAAATACA | AmpR for pLysSA + *Kpn*I |
| **Amp-rev** | ATCCTCGAGATATGAGTAAACTTGGTCTGACAG | AmpR for pLysSA + *Xho*I |
| **pLysS-BB-fwd** | ATCGGTACCCAAAAAATACGCCCGGTAGTGATC | pLysS-Backbone w/o CAT |
| **pLysS-BB-rev** | ATCCTCGAGGGTGCTACGCCTGAATAAGTG | pLysS-Backbone w/o CAT |

### 2.1.7 Plasmids

*Table 9: Plasmids*

| Name | Function | Source |
|---|---|---|
| **pET29a** | Expression with S-tag | Merck Millipore |
| **pET32a** | Expression with TRX-tag | Merck Millipore |
| **pET32aK** | Expression with TRX-tag | This study |
| **pET-SUMO** | Expression with SUMO-tag | Addgene |
| **pGEX-4T-1** | Expression with GST-tag | GE Healthcare |
| **pGEX-K** | Expression with GST-tag | This study |
| **pRSFDuet-1** | Expression with His-tag | Merck Millipore |
| **pLysSA** | Suppression of basal expression | This study |
| **pKJE7** | Chaperone expression: dnaK-dnaJ-grpE | Takara Bio Inc. |
| **pG-KJE8** | Chaperone expression: dnaK-dnaJ-grpE-groES-groEL | Takara Bio Inc. |
| **pTf16** | Chaperone expression: tig | Takara Bio Inc. |
| **pG-Tf2** | Chaperone expression: groES-groEL-tig | Takara Bio Inc. |
| **pGro7** | Chaperone expression: groES-groEL | Takara Bio Inc. |

pET His6 Sumo TEV LIC cloning vector (1S) (pET-SUMO) was a gift from Scott Gradia (Addgene plasmid # 29659 ; http://n2t.net/addgene:29659 ; RRID:Addgene_29659).

### 2.1.8 Microorganisms

*Table 10: Microorganisms*

| Strain | Plasmid | Source |
|---|---|---|
| ***E. coli* BL21** | | New England Biolabs |
| ***E. coli* BL21(DE3)** | | TUM |
| ***E. coli* BL21(DE3)K43** | | This study |
| ***E. coli* BL21(DE3)K60** | | This study |
| ***E. coli* BL21(DE3)pLysS** | | BINA collection |
| ***E. coli* C43(DE3)** | | TUM |
| ***E. coli* K12** | | BINA collection |
| ***E. coli* NEB 10β** | | BINA collection |
| ***E. coli* NEB 5α** | | BINA collection |
| ***E. coli* Waksman** | | BINA collection |
| ***E. coli* BL21(DE3)** | pET29a_VvGT14ao | This study |
| ***E. coli* BL21(DE3)** | pET29a_VvGT14ao + pLysSA | This study |
| ***E. coli* BL21(DE3)K43** | pET29a_VvGT14ao | This study |
| ***E. coli* BL21(DE3)pLysS** | pET29a_VvGT14ao | BINA collection |
| ***E. coli* BL21(DE3)** | pET29a_VvGT14ao-S168N | This study |
| ***E. coli* BL21(DE3)pLysS** | pET29a_VvGT14ao-S168N | This study |
| ***E. coli* NEB 10β** | pET29a_VvGT14ao-S168N | This study |
| ***E. coli* BL21(DE3)** | pET29a_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21(DE3)K43** | pET29a_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21(DE3)pLysS** | pET29a_VvGT14ao-S168N-W353R | This study |
| ***E. coli* NEB 10β** | pET29a_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21(DE3)** | pET29a_VvGT14ao-W353R | This study |
| ***E. coli* BL21(DE3)pLysS** | pET29a_VvGT14ao-W353R | This study |
| ***E. coli* NEB 10β** | pET29a_VvGT14ao-W353R | This study |
| ***E. coli* BL21(DE3)** | pET32a_VvGT14ao | This study |
| ***E. coli* BL21(DE3)pLysS** | pET32a_VvGT14ao | This study |
| ***E. coli* NEB 10β** | pET32a_VvGT14ao | This study |
| ***E. coli* NEB 10β** | pET32aK | This study |
| ***E. coli* BL21(DE3)** | pET32aK_VvGT14ao | This study |
| ***E. coli* BL21(DE3)K43** | pET32aK_VvGT14ao | This study |
| ***E. coli* BL21(DE3)pLysS** | pET32aK_VvGT14ao | This study |
| ***E. coli* BL21(DE3)** | pET-SUMO_VvGT14ao | This study |
| ***E. coli* BL21(DE3)K43** | pET-SUMO_VvGT14ao | This study |

| | | |
|---|---|---|
| ***E. coli* BL21(DE3)pLysS** | pET-SUMO_VvGT14ao | This study |
| ***E. coli* NEB 10β** | pET-SUMO_VvGT14ao | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT14ao | This study |
| ***E. coli* BL21(DE3)** | pGEX-4T-1_VvGT14ao | This study |
| ***E. coli* BL21(DE3)** | pGEX-4T-1_VvGT14ao + pGro7 | This study |
| ***E. coli* BL21(DE3)** | pGEX-4T-1_VvGT14ao + pKJE7 | This study |
| ***E. coli* BL21(DE3)** | pGEX-4T-1_VvGT14ao + pG-KJE8 | This study |
| ***E. coli* BL21(DE3)** | pGEX-4T-1_VvGT14ao + pTf16 | This study |
| ***E. coli* BL21(DE3)** | pGEX-4T-1_VvGT14ao + pG-Tf2 | This study |
| ***E. coli* BL21(DE3)pLysS** | pGEX-4T-1_VvGT14ao | This study |
| ***E. coli* Waksman** | pGEX-4T-1_VvGT14ao | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT15a-S152N | This study |
| ***E. coli* NEB 10β** | pGEX-4T-1_VvGT15a-S152N | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT15a-S152N-W321R | This study |
| ***E. coli* NEB 10β** | pGEX-4T-1_VvGT15a-S152N-W321R | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT15a-W321R | This study |
| ***E. coli* NEB 10β** | pGEX-4T-1_VvGT15a-W321R | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT16-S170N | This study |
| ***E. coli* NEB 10β** | pGEX-4T-1_VvGT16-S170N | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT16-S170N-W354R | This study |
| ***E. coli* NEB 10β** | pGEX-4T-1_VvGT16-S170N-W354R | This study |
| ***E. coli* BL21** | pGEX-4T-1_VvGT16-W354R | This study |
| ***E. coli* NEB 10β** | pGEX-4T-1_VvGT16-W354R | This study |
| ***E. coli* NEB 10β** | pGEX-K | This study |
| ***E. coli* BL21** | pGEX-K_VvGT14ao | This study |
| ***E. coli* BL21(DE3)** | pGEX-K_VvGT14ao | This study |
| ***E. coli* BL21(DE3)K43** | pGEX-K_VvGT14ao | This study |
| ***E. coli* BL21(DE3)pLysS** | pGEX-K_VvGT14ao | This study |
| ***E. coli* NEB 10β** | pGEX-K_VvGT14ao | This study |
| ***E. coli* Waksman** | pGEX-K_VvGT14ao | This study |
| ***E. coli* BL21** | pGEX-K_VvGT14ao-S168N | This study |
| ***E. coli* XL-1 Blue** | pGEX-K_VvGT14ao-S168N | This study |
| ***E. coli* BL21** | pGEX-K_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21(DE3)K43** | pGEX-K_VvGT14ao-S168N-W353R | This study |
| ***E. coli* NEB 10β** | pGEX-K_VvGT14ao-S168N-W353R | This study |
| ***E. coli* XL-1 Blue** | pGEX-K_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21** | pGEX-K_VvGT14ao-W353R | This study |

| | | |
|---|---|---|
| ***E. coli* XL-1 Blue** | pGEX-K_VvGT14ao-W353R | This study |
| ***E. coli* NEB 10β** | pLysSA | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1 | This study |
| ***E. coli* BL21(DE3)pLysS** | pRSFDuet-1 | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-(M1_P241)del-V431M | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao-(M1_P241)del-V431M | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-H371N | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao-H371N | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-M211K | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao-M211K | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-N45D-T442A | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao-N45D-T442A | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao | This study |
| ***E. coli* BL21(DE3)K43** | pRSFDuet-1_VvGT14ao | This study |
| ***E. coli* BL21(DE3)pLysS** | pRSFDuet-1_VvGT14ao | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-S168N | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao-S168N | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21(DE3)K43** | pRSFDuet-1_VvGT14ao-S168N-W353R | This study |
| ***E. coli* BL21(DE3)** | pRSFDuet-1_VvGT14ao-W353R | This study |
| ***E. coli* NEB 10β** | pRSFDuet-1_VvGT14ao-W353R | This study |

## 2.2 Buffers and solutions

All buffers and solutions have been prepared according to the following recipes unless otherwise stated.

### 2.2.1 Lysis buffer

*Table 11: Lysis buffer*

| amount/volume | | ingredient |
|---|---|---|
| 12 | mL | 1x Wash/bind buffer |
| 6 | mg | Lysozyme |
| 60 | µL | PMSF (100 mM) |
| 240 | µL | $MgCl_2$ (50 mM) |
| 1 | µL | DNAse I |

### 2.2.2 Reaction buffers

The buffer can be stored at room temperature. This buffer has been used for *in-vitro* glycosyltransferase reactions.

*Table 12: Reaction buffer (10x)*

| amount/volume | | ingredient |
|---|---|---|
| 30.3 | g | Tris |
| 4.8 | g | $MgCl_2$ |
| up to 1 | L | milliQ water adjusted to pH 8.0 with 1 N HCl |

The 2-ME buffer can be stored at room temperature. This buffer has been used for the coupled enzyme assays.

*Table 13: Reaction buffer with 2-ME (10x)*

| amount/volume | | ingredient |
|---|---|---|
| 30.3 | g | Tris |
| 4.8 | g | $MgCl_2$ |
| 7.8 | g | 2-Mercaptoethanol |
| up to 1 | L | milliQ water adjusted to pH 8.0 with 1 N HCl |

The 2x screening buffer cannot be stored and must be freshly prepared for each use.

*Table 14: 2x Screening Buffer*

| concentration | | ingredient |
|---|---|---|
| 100 | mM | Tris |
| 20 | mM | $MgCl_2$ |
| 40 | mM | 2-Mercaptoethanol |
| 0.2 | mM | UDP |
| 1 | mM | Geraniol |
| 5 | mM | ClPNP-glucose |
| 100 | nM | OleD |

### 2.2.3 Protein purification solutions

The 10x buffer can be stored at room temperature.

*Table 15: GST wash/bind buffer (10x)*

| amount/volume | | ingredient |
|---|---|---|
| 6.1 | g | $Na_2HPO_4$ |
| 2.0 | g | $KH_2PO_4$ |
| 80.1 | g | NaCl |
| 2.0 | g | KCl |
| up to 1 | L | milliQ water |

The 10x buffer can be stored at -20 °C.

*Table 16: GST elution buffer (10x)*

| amount/volume | | ingredient |
|---|---|---|
| 30.7 | g | Reduced glutathione |
| 60.6 | g | Tris |
| up to 1 | L | milliQ water |

The His wash/bind buffer can be stored at room temperature.

*Table 17: His wash/bind buffer (1x)*

| amount/volume | | ingredient |
|---|---|---|
| 6.0 | g | Sodium dihydrogen phosphate |
| 17.5 | g | NaCl |
| 2.0 | g | Imidazole |
| up to 1 | L | milliQ water |

The buffer can be stored at 4 °C.

*Table 18: His elution buffer (1x)*

| amount/volume | | ingredient |
|---|---|---|
| 6.0 | g | Sodium dihydrogen phosphate |
| 17.5 | g | NaCl |
| 34.0 | g | Imidazole |
| up to 1 | L | milliQ water adjusted to pH 8.0 |

### 2.2.4 Western blotting solutions

The semidry blotting buffer can be stored at room temperature.

*Table 19: Semidry blotting buffer*

| amount/volume | | ingredient |
|---|---|---|
| 3.0 | g | Tris |
| 15.0 | g | Glycine |
| 200 | mL | Methanol |
| up to 1 | L | milliQ water |

The western blot washing buffer can be stored at room temperature.

*Table 20: Western blot washing buffer*

| amount/volume | | ingredient |
|---|---|---|
| 2.4 | g | Tris |
| 8.2 | g | NaCl |
| 1 | mL | Tween 20 |
| up to 1 | L | milliQ water |

The western blot blocking buffer was set up fresh before use in WB washing buffer. This buffer cannot be stored.

*Table 21: Western blot blocking buffer*

| amount/volume | | ingredient |
|---|---|---|
| 5 | g | BSA |
| 100 | mL | WB washing buffer |

The Western blot detection buffer can be stored at room temperature.

*Table 22: Western blot detection buffer*

| amount/volume | | ingredient |
|---|---|---|
| 12.1 | g | Tris |
| 5.8 | g | NaCl |
| 0.48 | g | $MgCl_2$ |
| up to 1 | L | milliQ water adjusted to pH 9.5 with 1 N HCl |

The nitro blue tetrazolium chloride (NBT) solution can be stored at 4 °C.

*Table 23: NBT solution*

| amount/volume | | ingredient |
|---|---|---|
| 5 | g | NBT |
| 100 | mL | 70 % DMF |

The 5-bromo-4-chloro-3-indolyl phosphate (BCIP) solution can be stored at -20 °C.

*Table 24: BCIP solution*

| amount/volume | | ingredient |
|---|---|---|
| 5 | g | BCIP |
| 100 | mL | milliQ water |

### 2.2.5 Electrophoresis buffers and solutions

The 50x TAE buffer can be stored at room temperature.

*Table 25: 50x TAE buffer*

| amount/volume | | ingredient |
|---|---|---|
| 242 | g | Tris |
| 57 | mL | Acetic acid |
| 14.6 | g | EDTA |
| up to 1 | L | milliQ water adjusted to pH 8.0 with 1 N HCl |

The 10x SDS running buffer can be stored at room temperature.

*Table 26: 10x SDS running buffer*

| amount/volume | | ingredient |
|---|---|---|
| 30.3 | g | Tris |
| 144 | g | Glycine |
| 10 | g | SDS |
| up to 1 | L | milliQ water |

The Coomassie staining solution was always freshly prepared. Acetic acid and distilled water was added after the dissolution of the coomassie blue G250.

*Table 27: Coomassie staining solution*

| amount/volume | | ingredient |
|---|---|---|
| 0.33 | g | Coomassie blue G250 |
| 120 | mL | Ethanol |
| 24 | mL | Acetic acid |
| 120 | mL | Distilled water |

The destaining solution can be stored at room temperature and recycled several times by filtration through activated charcoal.

*Table 28: Destaining solution*

| amount/volume | | ingredient |
|---|---|---|
| 70 | mL | Acetic acid |
| 100 | mL | Ethanol |
| up to 1 | L | milliQ water |

## 2.3 Microbiological methods

### 2.3.1 Culture media

The LB medium can be stored at room temperature.

*Table 29: Lysogeny broth (LB medium)*

| amount/volume | | ingredient |
|---|---|---|
| 10 | g | Tryptone |
| 5 | g | Yeast extract |
| 10 | g | NaCl |
| up to 1 | L | milliQ water |

*Table 30: LB extended-log phase medium (Plank 2011)*

| amount/volume | | ingredient |
|---|---|---|
| 50 | mL | 1 M Potassium phosphate |
| 50 | mL | 10 % (v/v) Glycerol |
| up to 1 | L | LB medium |

The LB agar medium has been sterilized by autoclaving and can be stored at room temperature. After cooling the LB agar to approx. 60 °C, the antibiotics were added. Ampicillin, chloramphenicol, and kanamycin were used with final concentrations of 100, 34, and 50 µg/mL. A volume of 250 mL LB agar medium resulted in 13 thin LB agar plates.

*Table 31: LB agar*

| amount/volume | | ingredient |
|---|---|---|
| 15 | g | Agar-agar |
| up to 1 | L | LB medium |

Super optimal broth with catabolite repression (SOC medium) can be stored at room temperature for several weeks. The medium was sterilized by autoclaving before adding glucose. The glucose solution was sterilized by passing through a 0.22 µm filter and then added under sterile conditions.

*Table 32: SOC medium*

| amount/volume | | ingredient |
|---|---|---|
| 5 | g | Yeast extract |
| 20 | g | Tryptone |
| 0.58 | g | NaCl |
| 0.19 | g | KCl |
| 0.95 | g | $MgCl_2$ |
| 1.2 | g | $MgSO_4$ |
| 3.6 | g | Glucose |
| up to 1 | L | milliQ water |

2xYT medium was sterilized by autoclaving. The medium can be stored at room temperature.

*Table 33: 2xYT medium*

| amount/volume | | ingredient |
|---|---|---|
| 16 | g | Tryptone |
| 10 | g | Yeast extract |
| 5 | g | NaCl |
| up to 1 | L | milliQ water |

The 5x M9 salt solution was sterilized by autoclaving.

*Table 34: 5x M9 salt solution*

| amount/volume | | ingredient |
|---|---|---|
| 42.5 | g | $Na_2HPO_4$ x $H_2O$ |
| 15 | g | $KH_2PO_4$ |
| 2.5 | g | NaCl |
| 5 | g | $NH_4Cl$ |
| up to 1 | L | milliQ water |

The 5x M9 salt solution, the 20 % (w/v) glucose solution and water were sterilized by autoclaving. The 1 M $MgSO_4$ solution and the 1 M $CaCl_2$ solution were sterilized by passing through a 0.22 µm filter. All solutions can be stored separately at room temperature.

*Table 35: M9 minimal medium*

| amount/volume | | ingredient |
|---|---|---|
| 748 | mL | $H_2O$ |
| 200 | mL | 5x M9 salt solution |
| 50 | mL | 20 % (w/v) glucose |
| 2 | mL | 1 M $MgSO_4$ |
| 100 | µL | 1 M $CaCl_2$ |

The 10x $M9^+$ salt solution was sterilized by autoclaving.

*Table 36: 10x $M9^+$ salt solution*

| amount/volume | | ingredient |
|---|---|---|
| 75.2 | g | $Na_2HPO_4$ x 2*$H_2O$ |
| 30 | g | $KH_2PO_4$ |
| 5 | g | NaCl |
| 5 | g | $NH_4Cl$ |
| up to 1 | L | milliQ water adjusted to pH 7.2 with 1 N NaOH |

The trace elements $CuCl_2$, $CoCl_2$, $H_3BO_3$, and $MnCl_2$ were added to the 100x trace element solution from stock solutions with concentrations of 0.1, 0.2, 0.1, and 1 M, respectively.

*Table 37: 100x trace element solution*

| concentration | | ingredient |
|---|---|---|
| 13.4 | mM | EDTA |
| 3.1 | mM | $FeCl_3$ |
| 0.62 | mM | $ZnCl_2$ |
| 76 | µM | $CuCl_2$ |
| 42 | µM | $CoCl_2$ |
| 162 | µM | $H_3BO_3$ |
| 8.1 | µM | $MnCl_2$ |

The 10x $M9^+$ salt solution, the 20 % (w/v) glucose solution, the 1 M $MgSO_4$ solution, the 1 M $CaCl_2$ solution, and water were sterilized by autoclaving. The 0.1 % (w/v) biotin solution, the 0.1 % (w/v) thiamine, and the 100x trace elements solution were sterilized by passing through a 0.22 µm filter.

*Table 38: M9 minimal medium plus ($M9^+$) (Arie Geerlof)*

| volume | | ingredient |
|---|---|---|
| 867 | mL | $H_2O$ |
| 0.3 | mL | 1 M $CaCl_2$ |
| 100 | mL | $M9^+$ salt solution (10x) |
| 20 | mL | 20 % (w/v) Glucose |
| 1 | mL | 1 M $MgSO_4$ |
| 1 | mL | 0.1 % (w/v) Biotin |
| 1 | mL | 0.1 % (w/v) Thiamine |
| 10 | mL | Trace elements solution (100x) |

Minimal auto induction media (MAI and $MAI^+$) were prepared similarly to the M9 and $M9^+$ medium, respectively. The carbon source was modified, instead of 20 mL of a 20 % (w/v) glucose solution per liter, three carbon sources were used, glycerol, glucose, and lactose (Grabski et al. 2005, Studier 2005).

*Table 39: Carbon source for minimal auto induction medium (MAI)*

| volume | | ingredient |
|---|---|---|
| 10 | mL | 60 % (v/v) Glycerol |
| 5 | mL | 10 % (w/v) Glucose |
| 25 | mL | 8 % (w/v) Lactose |
| up to 1 | L | medium |

Terrific broth - auto induction medium (TB-AIM) was sterilized by autoclaving (Studier 2005).

*Table 40: Terrific broth - auto induction medium (TB-AIM)*

| amount/volume | | ingredient |
|---|---|---|
| 12 | g | Tryptone |
| 24 | g | Yeast extract |
| 0.15 | g | $MgSO_4$ |
| 3.3 | g | $(NH_4)_2SO_4$ |
| 6.5 | g | $KH_2PO_4$ |
| 7.1 | g | $Na_2HPO_4$ |
| 0.5 | g | Glucose |
| 2.0 | g | Lactose |
| up to 1 | L | milliQ water |

### 2.3.2 Culturing techniques

#### *2.3.2.1 Cultivation in liquid culture*

The use of a complex medium such as LB or 2xYT medium led to a rapidly growing culture. Minimal media such as M9 or $M9^+$ were preferred for biotransformation experiments, as the medium did not interfere with product analysis and purification. Depending on the strain, the media were supplemented with suitable antibiotics. Ampicillin, chloramphenicol, and kanamycin were used at final concentrations of 100, 35, and 50 µg/ml, respectively. Shaking flasks were used for protein, and cell production. The working volume was between 10 and 20 % of the total volume of the shaking flask in order to ensure adequate good aeration. A small volume (up to 5 mL) of medium was inoculated and incubated in test tubes. Shaking flasks and culture tubes were cultivated at 150 to 250 rpm, microtiter plates were cultivated at speeds of more than 250 rpm.

#### *2.3.2.2 Preparing cryo cultures*

Newly created strains were stored as cryo cultures at -80 °C. A cryo tube was labeled and filled with 500 µL of a 70 % (w/v) glycerol solution. Five hundred µL of fresh liquid culture was added and the glycerol cultures were immediately frozen after mixing.

#### *2.3.2.3 Minimum inhibitory concentration (MIC)*

The MIC represents the highest concentration of a toxin, which does not visibly inhibit the culture growth. To determine MIC of potential substrates, pre-cultures were cultivated and diluted to an OD of 0.02. Culture tubes were filled with 3 mL of this dilution and were supplemented with different amounts of the test substance. After overnight cultivation, turbidity of the cultures was determined. The highest concentration of the chemical at which a visible decrease in turbidity was observed was defined as the MIC.

### 2.3.3 Whole-cell biotransformation

The glycosylation performance of new strains had to be compared with previous strains to detect improvements. The glycosylation was done in shaking flasks and in 24-well plates (hitplate 25).

#### *2.3.3.1 Biotransformation in shaking flasks*

Shaking flasks with 50 mL of 2xYT medium were inoculated from cryo cultures, culture material from fresh agar plates, or from liquid pre-cultures. These cultures were incubated at 37 °C and 150 rpm overnight. The cultures were pelleted by centrifugation (5 min at 5,000 g and 4 °C) and were then suspended in 1 mL of M9 medium. The cultures were transferred to shaking flasks with 50 mL of M9 medium pre-chilled to a temperature of 18 °C. Ten µL of a 1 M IPTG solution was added to each shaking flask to induce GT expression. The cultures were incubated at 18 °C and 200 rpm for protein production. After 6 hrs, the substrate was added

and the cultivation was continued overnight. For product analysis, 500 µL of each culture were centrifuged twice (5 min at maximum speed and 4 °C) to obtain the cell-free supernatant. One hundred µL of supernatant were transferred to HPLC vials equipped with 200 µL inserts. These samples were used for the HPLC analysis.

#### *2.3.3.2 Hitplate 25*

Each deep well of a hitplate was filled with 2 mL of $M9^+$ medium as a pre-culture. The pre-cultures were inoculated with either fresh cell material (agar plate or liquid culture) or from cryo cultures. The hitplate was incubated at 37 °C and 250 rpm overnight. The main culture hitplate was prepared by filling two wells per test culture with $M9^+$ medium, each well with 2 mL. The cultures of the main plate were inoculated with 50 µL of the corresponding pre-culture and were incubated at 37 °C and 250 rpm for 4 hrs. The cultures were induced by adding 25 µL of 0.1 M IPTG solution to each well, resulting in a final IPTG concentration of 1.25 mM. The cultivation was continued at 18 °C and 500 rpm for 20 hrs. After GT expression, the biotransformation was started by adding 0.5 mL of the substrate solution (0.1 % (v/v) substrate in fresh $M9^+$ medium) to each well. The cultivation was continued for 24 hrs for product formation. The OD of each well was measured after biotransformation and samples of 1 mL were taken from each well. The samples were centrifuged for 10 min at maximum speed at 4 °C. The supernatant was transferred to an HPLC vial and was used for product quantification.

### 2.3.4 Cell disruption by sonication

Bacterial cell suspensions were subjected to sonication to disrupt the cell membrane. The suspension was kept in an ice bath to ensure proper cooling. The sonicator was set to a frequency of 50 % and a power of 50 %. Three cell disruption cycles were done with 60 seconds of sonication followed by 30 seconds of rest.

## 2.4 Molecular biological methods

### 2.4.1 Preparation of plasmid DNA

Plasmid extractions were done with the PureYield™ Plasmid Miniprep System from Promega (Mannheim, Germany) according to the manufacturer's manual.

### 2.4.2 Polymerase chain reaction

Polymerase chain reaction (PCR) was used (1) to amplify DNA for cloning, (2) for qualitative analysis, and (3) to introduce specific point mutations. Primers were modified and adapted to the different applications. The PCR conditions were adjusted for each reaction according to the manufacturer's recommendations in order to meet the requirements of each polymerase and primer pair used.

#### *2.4.2.1 Colony PCR*

Colony PCR (cPCR) was used to determine the presence of a gene or DNA construct in a bacterial colony. This PCR method was different from the standard protocol in terms of initial denaturation and the template type. A tiny speck of a bacterial colony from an agar plate or 1 µL of a liquid culture was dissolved in 50 µL of distilled water. Only 1 µL of the cell suspension was used as template for the PCR reaction. The initial denaturation step was prolonged up to 3 min to ensure proper cell lysis. The PCR product was detected by gel electrophoresis. When checking for an insert, it was preferable to use primers up- and downstream of the gene of interest. In this way, negative colonies are identified by detecting a much smaller PCR product.

#### *2.4.2.2 Error-prone PCR*

The epPCR was performed with the GeneMorph II Random Mutagenesis Kit from Agilent (Santa Clara, USA). The reactions were done according to the manufacturer's instructions.

#### *2.4.2.3 PCR product purification*

PCR products were separated from excess primers and other PCR components by applying the NucleoSpin® Gel and PCR Clean-up kit from Machery-Nagel (Düren, Germany). The purification was done according to the user manual.

### 2.4.3 Agarose gel electrophoresis

The operating parameters for electrophoresis using a mini gel were a voltage of 120 V and a separation time of 35 min. The separation was carried out on 1 % (w/v) agarose gel stained with Roti-Safe. A DNA ladder from GeneRuler DNA Ladder Mix, Thermo Fisher Scientific, Planegg, Germany was used. DNA bands were extracted and purified from the gel with the NucleoSpin® Gel and PCR Clean-up kit (Machery-Nagel, Düren, Germany).

*Table 41: Agarose gel composition*

| amount/volume | | ingredient |
|---|---|---|
| 100 | mL | 1x TAE buffer |
| 1 | g | Agarose |
| 5 | µL | Roti-Safe Gelstain |

### 2.4.4 Restriction enzyme digestion

The cutting of DNA was performed to analyze DNA samples, or clone new constructs for PCR or transformation. The reaction conditions recommended by the manufacturer have been met.

### 2.4.5 Dephosphorylation

FastAP Thermosensitive Alkaline Phosphatase was used to dephosphorylate DNA ends.

### 2.4.6 Ligation

DNA fragments were joined together using the T4 DNA ligase.

### 2.4.7 Transformation

*E. coli* cells were transformed with new DNA constructs by either electroporation or heat shock. Electroporation usually provides higher transformation rates and was therefore the method of choice when only low amounts of DNA were available or when a large number of transformants was required.

The heat shock transformation of *E. coli* BL21 was performed according to the transformation protocol C2530 from New England Biolabs GmbH (Frankfurt am Main, Germany). One tube of competent cells was thawed on ice for 10 min. One pg to 100 ng of DNA in 1 to 5 µL was added to a tube and mixed gently. The mixture was then incubated on ice for 30 min. The tube was dipped in water with a temperature of exactly 42 °C for precisely 10 seconds. The tube was incubated on ice for 5 min before adding 950 µL of SOC medium. The transformation culture was incubated at 37 °C and 200 rpm for 60 to 90 min to develop antibiotic resistance. Dilutions were plated on agar plates and were incubated to obtain transformants.

The electroporation was done with the MicroPulser™ using the pre-set program Ec3 with a 2 mm gap electroporation cuvette. The electro-competent cells were put on ice together with the DNA, a micro reaction tube, and the electroporation cuvette. After thawing, 40 µL of electro-competent cells were mixed with 1 ng DNA in the pre-chilled micro reaction tube. The DNA-cell-mixture was transferred to the bottom of the electroporation cuvette and was immediately subjected to one pulse of 3.0 V in the MicroPulser™. The cell suspension was immediately diluted with 1 mL of pre-warmed LB, or SOC medium and was transferred to a 2 mL micro reaction tube. The tube was incubated for 60 to 90 min at 37 °C and 200 rpm to allow the transformants to develop antibiotic resistance. Afterwards, 50 µL of the transformation culture was plated on an agar plate with the corresponding antibiotics and incubated overnight.

#### *2.4.7.1 Preparation of electro-competent cells*

Electro-competent cells were produced according to the manual of the MicroPulser™. The main culture (400 mL) was inoculated with 1 % (v/v) from a fresh pre-culture and was incubated under appropriate conditions until an OD of 0.6 was reached. The culture was then chilled on ice for at least 15 min and was kept cold throughout the rest of the procedure. The cells were harvested by centrifugation (15 min at 4,000 g and 4 °C) and were washed three times with decreasing volumes of a cold 10 % (w/v) glycerol solution. After washing with 400, 200, and 16 mL of cold glycerol solution, the cells were suspended in 2 mL of the cold 10 % (w/v) glycerol solution. This high-viscosity cell suspension was split into aliquots and stored at -80 °C.

#### *2.4.7.2 Preparation of electro-competent cells - quick protocol*

Electro-competent cells were also produced on a smaller scale for a few transformations on the same day. This method was adapted from a protocol of the Department of Biological Sciences, University of Texas Brownsville (Gonzales et al. 2013). The desired strain was streaked from a cryo culture onto an agar-plate. One hundred µL of LB medium was added to the plate and was distributed with an L-shaped cell spreader. After drying, the plate was incubated for 4 to 6 hrs at the appropriate temperature. The cells were then suspended in 1 mL of cold 10 % (w/v) glycerol solution by adding the glycerol solution onto the plate and scrubbing the agar surface carefully with an L-shaped cell spreader. The cell suspension was transferred to a micro reaction tube with a pipette and was centrifuged for 30 sec at maximum speed and 4 °C. The supernatant was discarded, the cells were suspended in 1 mL of new, cold glycerol solution and centrifuged again. This step was repeated twice with only 500 µL of cold glycerol solution. After the last centrifugation step, the supernatant was discarded again and the cell pellet was suspended in the required volume of 10 % (w/v) glycerol solution.

### 2.4.8 Plasmid curing

In order to remove plasmids from an *E. coli* strain, the strain was cultivated under stress conditions without the selective pressure required by the vector. *E. coli* cells were cultivated at 42 °C without antibiotics for several generations (up to 500 generations). Each day two cultures were started in shaking flasks with a starting concentration of $10^3$ cfu/mL or a theoretical optical density of $10^{-6}$ by inoculation with material of the previous culture. This low cell concentration resulted in more generations per flask.

An aliquot of each culture was plated onto an agar plate without antibiotics. Fifty µL of a $10^4$ and $10^5$ culture dilution were plated and incubated overnight at 37 °C. Single colonies from those plates were transferred to two collection plates equipped with PetriStickers displaying a 100 square grid. One pipet tip was used to transfer a single colony to the first collection plate and then to the second collection plate. The first collection plate did not contain antibiotics, the second collection plate contained the corresponding antibiotic of the expression vector to be removed. The collection plates were incubated overnight at 37 °C.

### 2.4.9 Ligation independent cloning

The cloning of *VvGT14ao* into the vector “pET His6 Sumo TEV LIC cloning vector (1S)” (pET-SUMO) was done with ligation independent cloning (LIC). The primers for *VvGT14ao* amplification were extended with the LIC tags. The forward primer was extended with TACTTCCAATCCAATGCA, and the reverse primer was extended with TTATCCACTTCCAATGTTATTA. The PCR product was treated with T4 DNA polymerase and dCTP. The vector was linearized with *Ssp*I, purified from an agarose gel, and subsequently,

treated with T4 DNA polymerase and dGTP. The vector and insert were annealed with a ratio of 1:3 regarding to the amount of substance and were directly used for transformation.

## 2.5 Biochemical methods

### 2.5.1 Protein quantification

Protein concentrations of samples were determined in microtiter plates with Roti®-Nanoquant according to the manufacturer's manual.

### 2.5.2 SDS-PAGE

The resolving gel was poured directly after the addition of polymerization catalyst, TEMED. The resolving gel was overlaid with 1 mL of isopropyl alcohol in order to avoid air bubbles between the resolving and stacking gel. After polymerization of the resolving gel, isopropyl alcohol was poured off and the stacking gel was prepared and poured on top of the resolving gel. A sample comb was inserted into the not yet polymerized stacking gel to obtain wells for the sample application. After polymerization, the gel was transferred to a SDS-PAGE chamber and the sample comb was removed carefully. The samples were boiled beforehand with an equal volume of 2x Laemmli Sample Buffer. The SDS gel was loaded with a protein ladder (PageRuler™ Prestained Protein Ladder, 10 to 180 kDa, Thermo Fisher Scientific, Planegg, Germany) and the samples. The running conditions were a constant voltage of 100 V and a running time of 3 hrs. The SDS gel was washed afterwards with water in a small plastic box to remove excess SDS. The SDS gel was then overlaid with 25 mL of the coomassie staining solution. The staining was done on a rocker incubator for 1 to 2 hrs. The gel was destained with 50 mL destaining solution. A folded paper towel was laid on top to absorb excess coomassie. The gel was documented when the background staining was sufficiently reduced.

*Table 42: SDS gel composition*

| | **Resolving gel 12 % SDS** | | **Stacking gel 5 % SDS** | |
|---|---|---|---|---|
| **ingredient** | volume | | volume | |
| **Distilled water** | 2.45 | mL | 2.03 | mL |
| **30 % Acrylamide/ bisacrylamide (37.5:1)** | 3.0 | mL | 0.5 | mL |
| **1 M Tris pH 6.8** | - | - | 375 | µL |
| **1.5 M Tris pH 8.8** | 1.9 | mL | - | - |
| **10 % (w/v) SDS** | 75 | µL | 30 | µL |
| **10 % (w/v) APS** | 75 | µL | 30 | µL |
| **TEMED** | 3 | µL | 3 | µL |

### 2.5.3 Western blot

Two identical SDS gels were prepared, one was stained and the other one was used for the western blot. The gel was incubated in semidry blotting buffer for 5 to 10 min and then transferred to the blotting sandwich. The blotting sandwich was assembled from bottom to top with two pieces of filter papers drenched in semidry blotting buffer, a PVDF membrane, the unstained SDS gel, and again two pieces of soaked filter papers. The PVDF membrane was activated beforehand by incubating it with ethanol for 1 min and then equilibrating it with semidry blotting buffer for 5 min. The blotting took place in a semidry blotter for 60 min with a voltage of 10 V and 0.8 mA/cm². After blotting, the membrane was transferred to a plastic box and was incubated with 20 mL of blocking buffer overnight. The membrane was then washed three times with 20 mL of washing buffer, 5 min each time. Subsequently, the membrane was incubated for 60 min at 4 °C with the first antibody in 20 mL of blocking buffer (1:5,000 Anti-GST). The membrane was washed three more times with washing buffer, 10 min each time. Then, the membrane was incubated for 60 min at room temperature with the second antibody in 20 mL of blocking buffer (1:30,000 Anti-Goat+AP). The membrane was washed again three times with washing buffer, for 10 min each time. The membrane was then equilibrated with 20 mL of detection buffer twice, each time for 5 min. Finally, the color reaction was performed for 5 to 60 min with 50 µL of NBT and 25 µL of BCIP in 20 mL of detection buffer. The reaction took place on a rocker incubator in the dark and was stopped by washing away the substrate with distilled water when the bands were clearly visible.

*Table 43: Incubation steps during western blotting*

| Time | Solution | Repetitions |
|---|---|---|
| Overnight | Blocking buffer | 1x |
| 5 min | Washing buffer | 3x |
| 60 min | Blocking buffer + 1$^{st}$ antibody | 1x |
| 10 min | Washing buffer | 3x |
| 60 min | Blocking buffer + 2$^{nd}$ antibody | 1x |
| 10 min | Washing buffer | 3x |
| 5 min | Detection buffer | 2x |
| 5 to 60 min | Detection buffer + NBT & BCIP | 1x |

### 2.5.4 Glycosyltransferase assay

The reaction was performed in micro reaction tubes and was started by adding UDP-glucose. The reaction was either stopped by adding an equal volume of methanol or UDP Detection Reagent (UDR). The samples inactivated with methanol were centrifuged and then quantified by HPLC. The samples stopped with UDR were measured with the UDP-Glo™ Assay.

*Table 44: Glycosyltransferase reaction*

| volume | | ingredient |
|---|---|---|
| 40 – X – Y | µL | Water |
| 5 | µL | 10x Reaction buffer |
| X | µL | Sample solution |
| Y | µL | Substrate solution |
| 5 | µL | 1 mM UDP-glucose |

### 2.5.5 UDP-Glo™ Assay

The UDP-Glo™ Assay determines the UDP concentration in a sample. Since the glucosylation reaction releases equimolar amounts of glucoside and UDP, the glucoside formation can be quantified by measuring the UDP release. Five µL of the GT reaction was added to a well of a 384 well round-bottom microtiter plate containing 5 µL of UDR. The UDR was prepared according to the manufacturer's information. The plate's luminescence was measured after 60 min of incubation at room temperature in the dark. A calibration curve was created with varying concentrations of UDP and was used to estimate the UDP and the glycoside concentration of the sample.

### 2.5.6 Coupled enzyme assay

VvGT14a uses UDP-glucose and geraniol to produce geranyl glucoside and UDP. The second enzyme, OleD, uses UDP and ClPNP-glucose to form UDP-glucose and ClPNP. ClPNP-glucose is present during the reaction in excess and is a colorless substance. The free ClPNP however displays a yellow color, which can be quantified by a photometer. The geranyl glucoside formation was measured indirectly by the increase in absorbance due to ClPNP accumulation.

#### *2.5.6.1 Cultures grown and induced in 96 deep well plates*

Pre-cultures were grown in LB medium overnight at 37 °C and 280 rpm. The main culture was grown in 1 mL 2xYT medium with 2 % (w/v) of glucose and inoculated with 100 µL of the pre-culture. The induction was done after 4 hrs of cultivation with 1 mM of IPTG. The GT production was carried out overnight at 18 °C and 280 rpm. The cells were harvested by centrifugation for 20 min at 5,000 rpm and 4 °C, and were suspended in 150 µL of reaction buffer with 2-mercaptoethanol and 10 mg/mL lysozyme. After 20 min of incubation at 37 °C, the plates were frozen at -80 °C overnight. The plates were thawed at 37 °C for 15 min and the cell debris was pelleted by centrifugation for 45 min at 5,000 rpm and 4 °C. One hundred µL of the supernatant was transferred to the wells of a 96 well plate containing 100 µL of 2x screening buffer. The plate was immediately placed into the plate reader, where the absorbance was monitored at a wavelength of 410 nm and logged for 150 min. The plate reader was pre-heated

to 30 °C for the reaction. The activity of the samples was represented by the slope (m) of the increase in absorbance at 410 nm. The relative wild type (WT) activity (Q) was estimated through the slope of the positive and negative control carrying the wild type vector and the empty vector, respectively. The relative wild type activity was obtained from the sample-wild type quotient of the negative control differentials (equation below).

$$Relative\ wild\ type\ activity\ =\ Q_{\Delta_{sample}/\Delta_{wild\ type}} = \frac{m_{sample} - m_{empty\ vector}}{m_{wild\ type} - m_{empty\ vector}}$$

### 2.5.7 Protein purification

In order to purify His- or GST-tagged proteins a culture was grown, induced, and incubated overnight. The cells were pelleted by centrifugation for 10 min at 5,000 g and 4 °C, and frozen without the supernatant at -80 °C for at least 15 min. The pellet was suspended in 10 mL of lysis buffer (basis: GST wash/bind buffer 1x for GST-tagged proteins, and His wash/bind buffer 1x for His-tagged proteins) and lysed by sonication. The total protein was separated from the cell debris by centrifugation for 20 min at 10,000 g and 4 °C. The supernatant was transferred to a 15 mL falcon tube with 500 µL resin (GST•Bind™ Resin for GST-tagged proteins, and Profinity™ IMAC Resin, Ni-charged for His-tagged proteins) and incubated for 60 min at 4 °C on a vertical rotator. The resin was equilibrated beforehand by washing it twice with the corresponding 1x wash/bind buffer. The resin was pelleted after each step by centrifugation for 5 min at 800 g and 4 °C. The rest of the cell lysate, containing unbound protein, was poured off and the resin was washed twice with 5 mL of wash/bind buffer. The bound proteins were eluted by adding 300 µL of elution buffer to the resin and incubating for 5 min at room temperature. GST elution buffer 1x was used for GST-tagged proteins, and His elution buffer 1x for His-tagged proteins. The elution step was repeated four more times resulting in five elution fractions.

### 2.5.8 Chloramphenicol acetyltransferase assay

The chloramphenicol acetyltransferase (CAT) assay from Sigma-Aldrich (retrieved from https://www.sigmaaldrich.com/technical-documents/protocols/biology/enzymatic-assay-of-chloramphenicol-acetyltransferase.html) was used in a modified form. In order to detect side-activities of CAT, the reaction time was increased up to 24 hrs. 5,5'-Dithiobis-(2-nitrobenzoic acid) (DTNB or Ellman's reagent) was not stable for that long period in the reaction mixture and was therefore added afterwards. In this way, the color formation was performed after the acetylation reaction. The reaction was performed at 30 °C with soft shaking. The absorbance was measured for a few min after the addition of DTNB at a wavelength of 412 nm. The

measurement was complete when the color formation reaction was finished and the absorbance had stabilized.

*Table 45: Chloramphenicol acetyltransferase reaction*

| volume | | ingredient |
|---|---|---|
| 0.1 | mL | CAT (500 U/mL) |
| 0.2 | mL | 5 mM Acetyl-CoA |
| 0.7 | mL | 100 mM Tris pH 7.8 |
| 1 | µL | Substrate |
| 0.2 | mL | 2.5 mM DTNB |

## 2.6 Analytical methods

The following analytical methods were used to determine cell, DNA, and product concentrations.

### 2.6.1 DNA quantification

The DNA concentration of a sample was determined with a NanoDrop 1000 according to the manufacturer's manual. Alternatively, the DNA concentration of a sample was estimated from an agarose gel by comparing the band intensity to the DNA ladder.

### 2.6.2 HPLC

The glycoside concentrations in culture supernatants were determined by HPLC-UV and HPLC-UV-MS analyses. One HPLC-UV system by Jasco (Pfungstadt, Germany) with an isocratic method, and two gradient HPLC-UV-MS systems from Agilent Technologies (Santa Clara, USA) were used. All three systems were equipped with a Luna® 3 µm C18 LC Column 150 x 2 mm (part number: 00F-4251-B0) purchased from phenomenex (Aschaffenburg, Germany).

#### *2.6.2.1 Isocratic elution*

The Jasco HPLC-UV system was composed of an autosampler: AS-1555; a degasser: DG-1580-54; an interface: LC-Net II/ADC; a mixing chamber: LG-1580-04; a pump: PU-1580; and an UV detector: UV-1575. A volume of 10 µL cell free culture supernatant was used for the injection. Isocratic elution was performed with 45 % (v/v) aqueous methanol with 0.1 % (v/v) formic acid and the substances citronellyl glucoside, eugenyl glucoside, geranyl glucoside, and neryl glucoside were eluted at approximately 8.3, 3.2, 5.4 and 5.5 min, respectively. Substances were quantified by UV detection at a wavelength of 210 nm. Authentical reference material was used for calibration.

*2.6.2.2 Gradient elution*

The first HPLC-UV-MS system was composed of an Agilent autosampler G1313A, a column compartment G1316A, a degasser: G1379A, a diode Array Detector (DAD) G1315B, a quaternary pump G1311A and a mass spectrometer (Esquire 3000 Plus Ion Trap) from Bruker Daltonics (Bremen, Germany). The MS parameters are shown in Table 46.

The second HPLC-UV-MS system was composed of an Agilent autosampler G1313A, a capillary pump G1376A, a column compartment G1316A, a degasser G1379A, a variable Wavelength Detector (VWD), G1314A and a mass spectrometer 6340 Ion Trap LC/MS: G2499A. The MS parameters are shown in Table 46.

Solvent A was distilled water and solvent B was methanol, both with 0.1 % (v/v) formic acid. The gradient started with 10 % (first HPLC-UV-MS) and 0 % (second HPLC-UV-MS) solvent B, respectively. The proportion of solvent B was increased to 50 % over the first 3 min then to 100% within 3 min. Pure solvent B was kept up for 8 min and then reduced within 0.1 min to 10 % and 0 % solvent B for the first and the second system, respectively. This concentration was kept for the remaining 10.9 min of the methods.

Citronellol, geraniol, nerol and the corresponding glucosides were quantified at 210 nm. Furaneol and furaneoyl glucoside was quantified at 254 nm. In the MS analysis the formic acid adduct ions $[M+HCOOH-H]^-$ of the products were used for the quantification, except for the furaneoyl glucoside where the $[M-H]^-$ ion was used.

*Table 46: Mass spectrometry parameters*

| **Parameter** | **value** | | **unit** |
|---|---|---|---|
| Machine | 6340 Ion Trap | Esquire 3000 Plus Ion Trap | |
| Manufacturer | Agilent Technologies | Bruker Daltonics | |
| Spray gas | nitrogen | nitrogen | |
| Spray gas pressure | 30.0 | 30.0 | psi |
| Dry gas | nitrogen | nitrogen | |
| Dry gas temperature | 330 | 330 | °C |
| Dry gas flow | 9 | 9 | L/min |
| Scan range | 50 to 800 | 50 to 800 | m/z |
| Polarity | positive/negative alternating | positive/negative alternating | |
| ICC target | 100,000 | 30,000 | |
| ICC target (max.) | 200 | 200 | ms |
| Ionization voltage | ± 3,500 | ± 4,000 | V |
| end plate voltage | ± 500 | ± 500 | V |
| MS/MS | auto | auto | |
| Collision gas | helium | helium | |
| Collision voltage | 1.0 | 1.0 | V |
| **Settings of the ion trap** | | | |
| Skimmer | ± 40 | ± 40 | V |
| Cap exit | ± 121 | ± 121 | V |
| Oct 1 DC | ± 8 | ± 12 | V |
| Oct 2 DC | ± 1.7 | ± 1.7 | V |
| Trap drive | 50.3 | 45.6 | |
| Ocr RF | 171 | 150 | Vpp |
| Lens 1 | ± 5 | ± 5 | V |
| Lens 2 | ± 60 | ± 60 | V |

### 2.6.3 Cell density quantification

The optical density (OD) of cultures was measured at a wavelength of 600 nm with a photometer in semi-micro cuvettes using different dilutions. The corresponding dilution of fresh culture medium was used as an appropriate blank reference. The cell dry weight was calculated by multiplying the OD with 0.39 g/L (Glazyrina et al. 2010).

# 3 Results

The two geranyl glucoside production strains that were available at the beginning of this work were *E. coli* BL21(DE3)pLysS /pET29a_VvGT14ao and /pGEX-4T-1_VvGT14ao. They produced 0.010 and 0.037 mM of geranyl glucoside from geraniol by biotransformation. The aim of this study was to provide an improved production system. A production organism contains many parts, which can be varied in order to increase the product yield of the biotransformation. However, before any improvements can be made on these parts, the choice of the production host should be reconsidered. The choice of the host strain can limit the available production plasmids or substrates.

## 3.1 Testing of the geraniol tolerance of *E. coli* strain C43(DE3)

Many terpenols have toxic effects on microorganisms and geraniol also shows antimicrobial properties. The minimal inhibitory concentration (MIC) of geraniol acting on the common expression host, *E. coli* BL21(DE3)pLysS, was determined to be 0.35 g/L (Alexandra Hermenau, personal communication). To find a more resistant strain, *E. coli* C43(DE3), a mutant derived from *E. coli* BL21(DE3), was tested for geraniol tolerance. *E. coli* C43(DE3) can produce and tolerate toxic proteins and membrane proteins (Miroux and Walker 1996).

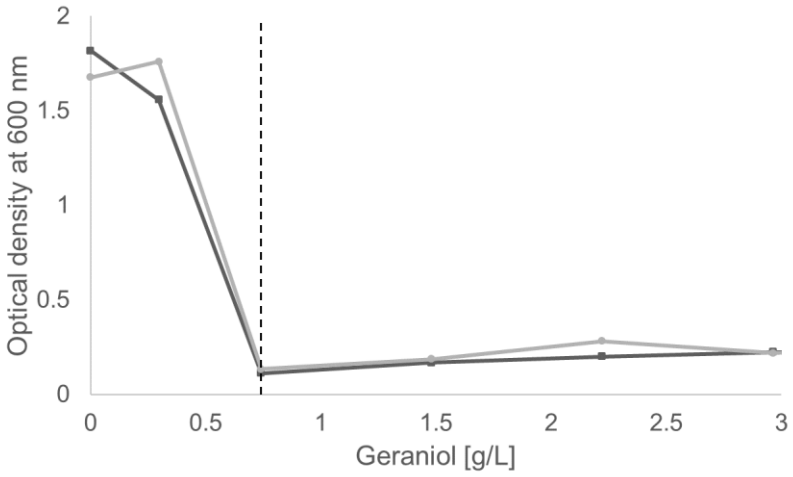


*Figure 8: Growth of E. coli BL21(DE3) and C43(DE3) in the presence of geraniol*

*E. coli strains BL21(DE3) (dark grey) and C43(DE3) (light grey) were cultivated overnight in LB medium at 37 °C and 180 rpm with different geraniol concentrations. Optical density of the cultures is shown. The dashed line indicates the maximum water solubility of geraniol.*

Geraniol concentrations of 0.3 to 3 g/L were chosen (Figure 8). An oily layer was visible at higher concentrations indicating an excess of geraniol. Both strains showed a very similar growth profile (Figure 8). Therefore, *E. coli* C43(DE3) offers no advantage since geraniol is equally toxic for both tested *E. coli* strains.

## 3.2 Testing of molecular chaperones

The overexpression of foreign proteins in host organisms often results in the formation of inactive protein aggregates called inclusion bodies. A common way to improve the concentration of correctly formed proteins inside *E. coli* cells is to employ molecular chaperones promoting proper folding of active proteins or degradation of inactive protein aggregates. These chaperones are introduced via expression plasmids. The chaperone plasmid set from Takara Bio contains five plasmids expressing different chaperones and chaperone combinations. Plasmid pG-KJE8 codes for dnaK-dnaJ-grpE and groES-groEL, pG-Tf2 encodes groES-groEL and tig, and pGro7, pKJE7, and pTf16 code for groES-groEL, dnaK-dnaJ-grpE, and for tig, respectively (Nishihara et al. 1998, Nishihara et al. 2000). Another way to employ chaperones is to elicit their natural production by introducing stress to the culture, for example by adding ethanol to a final concentration of 3 % (v/v) to the culture (Thomas and Baneyx 1996).

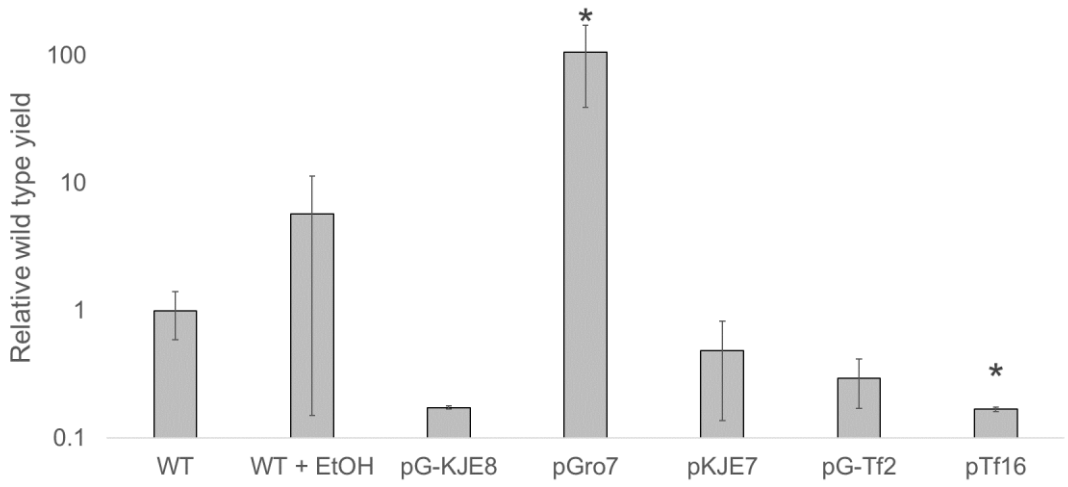


*Figure 9: Relative geranyl glucoside titers produced by E. coli BL21(DE3) /pGEX-4T-1_VvGT14ao strains equipped with different chaperone plasmids in comparison with the concentration produced by the strain lacking the chaperones (logarithmic scale).*

*The production strain E. coli BL21(DE3) /pGEX-4T-1_VvGT14ao (WT) was equipped with plasmids coding for different chaperones (pGro7, pKJE7, and pTf16) and chaperone combinations (pG-KJE8 and pG-Tf2). The production strain was also supplemented with 3 % (v/v) ethanol at protein induction in order to trigger the natural chaperone production in E. coli cells as a second reference (WT + EtOH). The induction of the chaperone plasmids was done according to the manufacturer's manual. Samples with a significant deviation from the wild type reference are marked with an asterisk. Experiments were performed in shaking flasks by Marion Koop.*

The introduction of chaperone expression plasmids into the production strain *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao showed mixed results (Figure 9). Most of the used plasmids did not increase the geranyl glucoside titer. However, the expression of pGro7, significantly raised the product yield. The induction of the natural stress response by addition of ethanol showed a positive effect, although the difference was statistically not significant.

Geranyl glucoside production of *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao with and without pGro7 was also studied at different temperatures (Figure 10). The strain *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao transformed with the pGro7 plasmid, produced significantly higher amounts of the terpenyl glucoside than the original strain without the chaperone, regardless of the temperature.

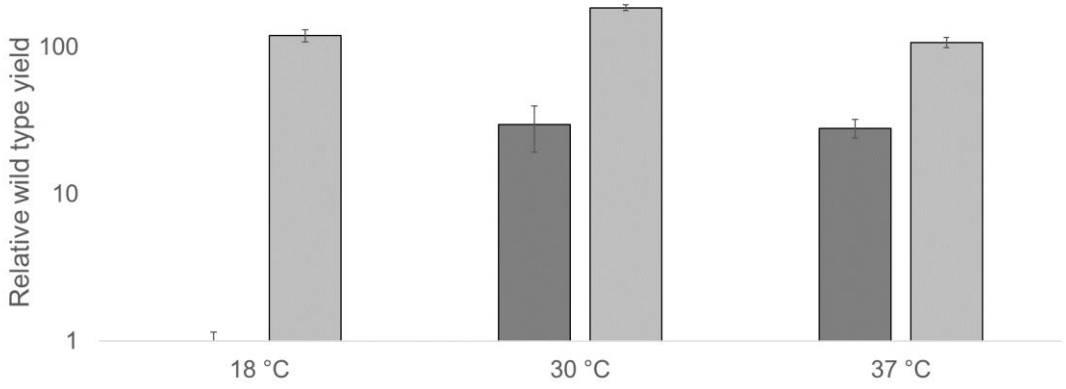


*Figure 10: Relative geranyl glucoside concentrations produced by E. coli BL21(DE3) /pGEX-4T-1_VvGT14ao with and without pGro7 at different temperatures (logarithmic scale)*

*The production strain E. coli BL21(DE3) /pGEX-4T-1_VvGT14ao was used with (light bars) and without (dark bars) the second plasmid pGro7 at different temperatures to produce geranyl glucoside. Protein production and biotransformation was performed at 18 °C, 30 °C, and 37 °C. The experiments were carried out in shaking flasks for 16 hrs by Marion Koop.*

## 3.3 Cloning of new expression vectors

The production of active enzyme can be influenced by its mechanism of regulation, the gene dosage, and the addition of a fusion tag (Sørensen and Mortensen 2005). Many expression vectors provide fusion tags for different purposes, mainly for protein purification, but also for higher protein solubility, or both. The vectors pET32a, pET-SUMO, and pGEX-4T-1 were chosen because of their fusion proteins TRX, SUMO, and GST, respectively. These fusion proteins are supposed to convey higher solubility to a heterologous expressed protein. The expression vector pET29a was chosen for its rather small fusion tag, the S-tag, as a reference and pRSFDuet-1 was chosen for its high-copy promoter.

Most expression vectors contain a kanamycin or ampicillin resistance cassette. However, since ampicillin is less stable in high cell-density cultures, it is not suitable for large scale industrial applications (Sivashanmugam et al. 2009). Therefore, the vectors pET32a and pGEX-4T-1 were modified to carry a kanamycin resistance cassette instead of an ampicillin resistance cassette (Figure 11). Hence, the modified versions of pGEX-4T-1 and pET32a only differed in their antibiotic resistance cassette and were termed pGEX-K and pET32aK, respectively.

The choice of possible production hosts was limited by the T7 promoter used in pET29a, pET32a, pET32aK, pET-SUMO, and pRSFDuet-1. These expression vector can only be utilized by a host carrying the λ DE3 insertion and were therefore combined with *E. coli* BL21(DE3) and *E. coli* BL21(DE3)pLysS. The vectors pGEX-4T-1 and pGEX-K use the tac promoter instead of the T7 promoter. Therefore, they can be read by virtually any *E. coli* strain and were combined with *E. coli* BL21, *E. coli* BL21(DE3), *E. coli* BL21(DE3)pLysS, and *E. coli* Waksman. In total 18 strain-vector combinations could be created out of the four hosts and seven vectors (Table 47).

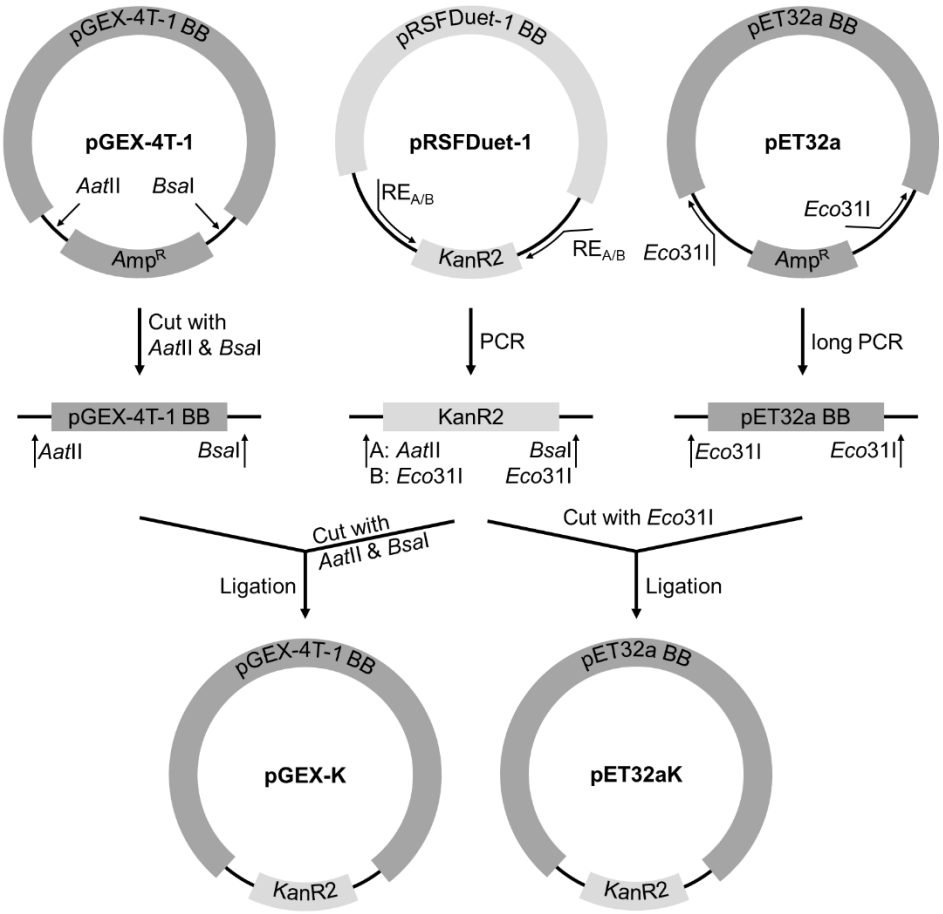


*Figure 11: Cloning of pGEX-K and pET32aK*

*Expression vectors pGEX-K and pET32aK are kanamycin resistant versions of pGEX-4T-1 and pET32a, respectively. The previous ampicillin resistance cassette was exchanged with the kanamycin resistance cassette of pRSFDuet-1. The pGEX-4T-1 vector contains suitable restriction enzyme recognition sites for cloning. The kanamycin resistance cassette was amplified using PCR and the corresponding restriction sites were added via primer extensions. The entry vector and the PCR product were digested with AatII and BsaI, gel-purified and ligated to obtain pGEX-K. The pET32a vector does not contain suitable restriction sites. Therefore, the vector backbone was amplified using long PCR and flanked by suitable restriction sites. The kanamycin cassette was also flanked by the corresponding restriction sites. Both PCR products were digested with Eco31I, gel-purified and ligated in order to obtain pET32aK.*

*Table 47: Strain-vector combinations with VvGT14ao*

| Species | Strain | Vector | Antibiotic resistance | Fusion tag | Origin of replication | Promoter |
|---|---|---|---|---|---|---|
| ***E. coli*** | BL21<br>BL21(DE3)<br>BL21(DE3)pLysS<br>Waksman | pGEX-4T-1 | Amp | GST | pBR322 | tac |
| | | pGEX-K | Kan | | | |
| | BL21(DE3)<br>BL21(DE3)pLysS | pET29a | Kan | S | pBR322 | T7 |
| | | pET32a | Amp | TRX | | |
| | | pET32aK | Kan | TRX | | |
| | | pET-SUMO | Kan | SUMO | | |
| | BL21(DE3)<br>BL21(DE3)pLysS | pRSFDuet-1 | Kan | His | RSF | T7 |

The origin of replication (ori) determines the plasmid copy number (PCN; Table 47). Plasmids with the pBR322 ori resulted in a PCN of approximately 15–20 (Boros et al. 1984), whereas the high-copy ori RSF yielded PCN of over 100 copies (Chakravartty and Cronan 2015).

A codon-optimized version of *VvGT14a*, *VvGT14ao*, was cloned into all seven expression vectors. The cloning of *VvGT14ao* into the expression vectors was done with standard cloning techniques (Figure 12), except for pET-SUMO where *VvGT14ao* was cloned via ligation independent cloning (LIC).

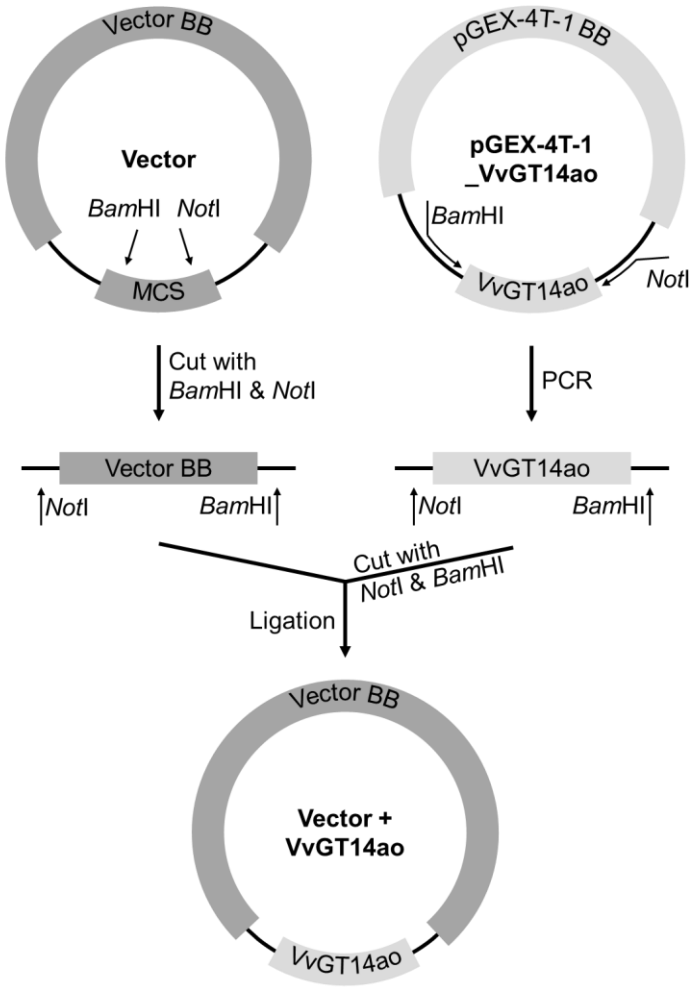


*Figure 12: Cloning of VvGT14ao into expression vectors using BamHI and NotI*

*The gene VvGT14ao was amplified from pGEX-4T-1_VvGT14ao using PCR with primers conveying the restriction sites BamHI and NotI. The entry vector and the PCR product were digested with BamHI and NotI, gel-purified, and ligated to obtain the final expression vector. All expression vectors containing VvGT14ao were obtained this way, except for pET-SUMO, which was obtained by ligation independent cloning (LIC). The final vectors were verified via sequencing.*

The *E. coli* hosts were transformed with suitable expression vectors (Table 47) and tested for their glycosylation ability *in-vivo*.

## 3.4 Developing a novel small scale biotransformation system

The initial cultivation system available to test the glycosylation ability of *E. coli* strains used shaking flasks and employed a medium exchange. The cells were harvested from a complex medium prior to induction and were suspended in a simple minimal medium. This ensured optimal cell growth and an adequate cell density at induction. However, experiments with replicates required a lot of space. Therefore, a plate-based system was developed using a 24 well microtiter plate called hitplate 25 (HP).

The new system should allow the testing of multiple strains and substrates as well as replicates on the same plate. For the first experiments the cells were propagated in shaking flasks and

the cultures were distributed onto the plates after the medium exchange and the addition of the inducer, isopropyl β-D-1-thiogalactopyranoside (IPTG). The first assays showed that the production of glucosides inside hitplates was possible (Figure 13). In addition, the standard deviations of the quadruplicates measured on one plate were low (less than 26 %). Subsequently, multiple determinations were performed on one plate using twelve wells per substrate. The wells A1 through C4 were used to produce hexyl glucoside, and D1 through F4 were used to produce furaneoyl glucoside. Product formation was quantified by liquid chromatography-ultraviolet detection (LC-UV) and liquid chromatography-mass spectrometry (LC-MS). The deviations of the MS peak areas were more pronounced than the deviations of the UV peak areas (Figure 14). Therefore, the UV data were better suited for glucoside quantification.

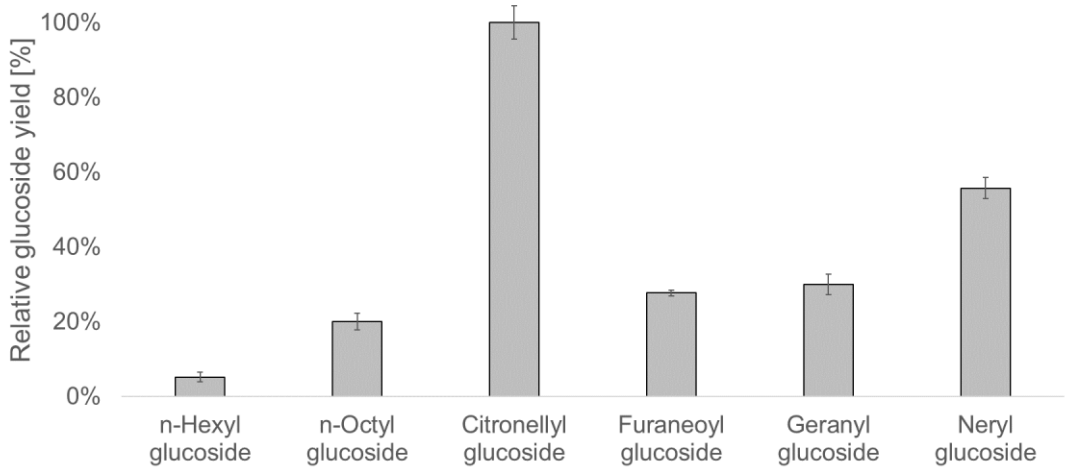


*Figure 13: Glucoside production using HP cultivation*

*E. coli BL21(DE3)pLysS /pGEX-K_VvGT14ao was used to produce the glucosides of 1-hexanol, 1-octanol, citronellol, furaneol, geraniol, and nerol in a hitplate 25. The pre-cultivation was done in a shaking flask in 2xYT medium at 37 °C and 180 rpm until an optical density of 0.6 was reached. The culture was harvested, suspended in M9 medium, induced with 1 mM of IPTG, and was distributed into the wells of the hitplate. Each well was filled with 3.75 mL of induced culture. After an overnight incubation at 18 °C and 500 rpm, the substrates were added with 1.25 mL of a 0.35 g/L substrate solution and the cultivation was continued for 24 hrs. Analysis was performed by LC-MS, signal intensity of citronellyl glucoside was set 100 %.*

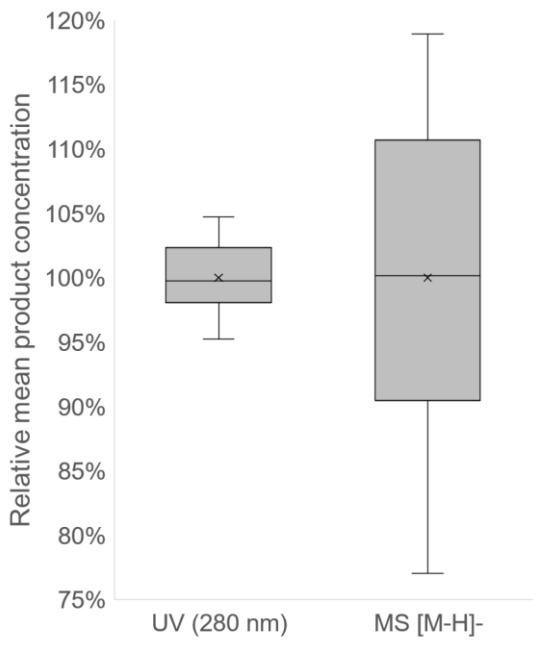


*Figure 14: Analytical variations of the furaneoyl glucoside production on a hitplate*

*E. coli BL21(DE3)pLysS /pGEX-K_VvGT14ao was used to produce furaneoyl glucoside by HP cultivation. The pre-cultivation was done in a shaking flask in 2xYT medium at 37 °C and 180 rpm until an optical density of 0.6 was reached. The culture was harvested, suspended in M9 medium, induced with 1 mM of IPTG, and was distributed into the wells of the hitplate. Each well was filled with 3.75 mL of induced culture. After an overnight incubation at 18 °C and 500 rpm, 0.08 g/L of substrate was added and the cultivation was continued for 24 hrs. Boxplot of the measured product concentrations relative to the plates mean product concentration.*

The HP cultivation system can be used to not only observe differences between production strains but also to quantify the difference using UV data. To facilitate the handling of the system a new medium ($M9^+$) was tested for its growing and producing capabilities. $M9^+$ medium is a minimal medium like M9 medium, however it contains trace elements and vitamins. It was tested whether the $M9^+$ medium could replace the 2xYT and M9 medium in the biotransformation experiments using hitplates. This would eliminate the need to change the medium and simplify the testing method.

The test with the $M9^+$ medium showed that a medium exchange was not necessary for the assessment of a production strain since using the $M9^+$ medium achieved titer similar to a change of medium from 2xYT to M9 (data not shown). An attempt was also made to increase the substrate concentration from 0.02 % (v/v) to 0.04 % (v/v). However, the increased geraniol concentration did not lead to a higher glucoside concentration, probably due to the poor water solubility of geraniol. The titers could be quantified after determination of calibration curves by LC-UV analysis. Both test series yielded approximately 58 mM geranyl glucoside after 24 hrs and 0.109 mM after 96 hrs. The following experiments were therefore done in $M9^+$ medium, with 0.02 % (v/v) of substrate, and with a biotransformation time of 24 hrs.

The glucosylation activity of the culture decreased continuously when the biotransformation was performed over a longer period of time. The calculated activity after 24 hrs of biotransformation was 99 µU or 32.0 µU/mg considering the approximately 3.1 mg calculated dry cell mass. The decrease in activity appeared to be linear. A linear regression allowed the calculation of the theoretical initial activity of 107 µU or 34.5 µU/mg for *E. coli* BL21(DE3) /pET29a_VvGT14ao (Figure 15).

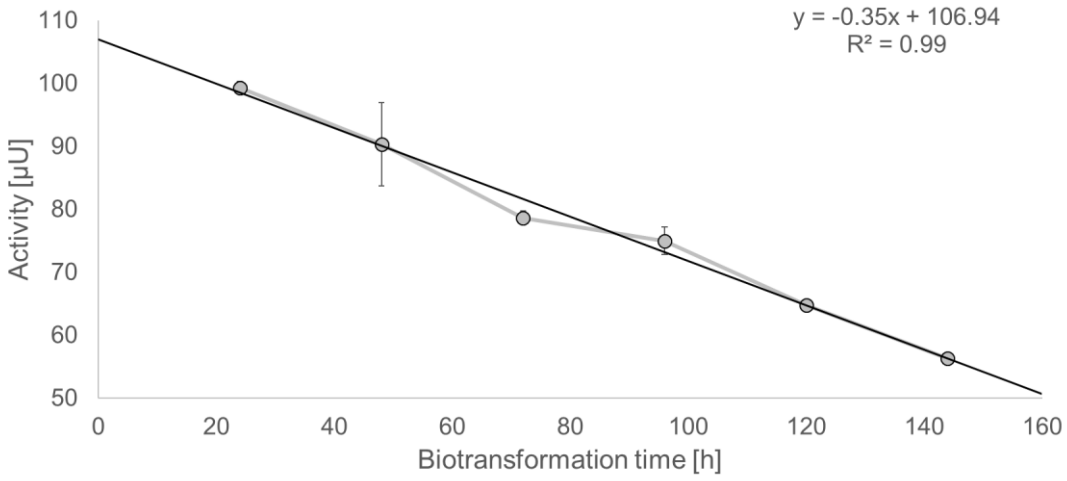


*Figure 15: Activity development during biotransformation*

*E. coli BL21(DE3) /pET29a_VvGT14ao was used to produce geranyl glucoside in an HP cultivation system. The pre-cultivation was done in a shaking flask in M9⁺ medium at 37 °C and 180 rpm until an optical density of 0.6 was reached. The culture was induced with 1.25 mM of IPTG and dispensed into hitplates. After an overnight incubation at 18 °C and 500 rpm, 0.18 g/L of substrate was added and the cultivation was continued for 144 hrs. Samples were taken every 24 hrs.*

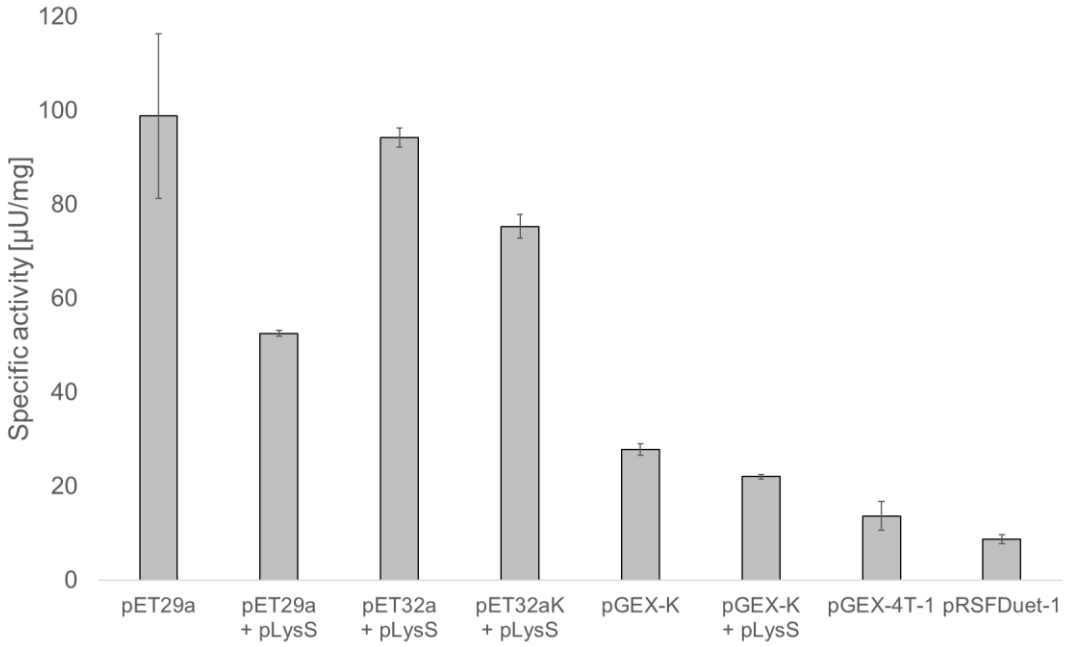


*Figure 16: Specific activity of E. coli BL21(DE3) strains*

*The specific activity for geranyl glucoside formation of E. coli BL21(DE3) strains containing different expression vectors was tested in an HP culture system The pre-cultivations were done in hitplates as well. The cultures were grown in M9$^+$ medium at 37 °C and 250 rpm until the last culture reached an optical density of 0.6. The remaining cultures were diluted with fresh medium to a similar optical density. The cultures were induced with 1.25 mM of IPTG and then incubated overnight at 18 °C and 500 rpm. 0.18 g/L of substrate was added and the cultivation was continued for 24 hrs.*

The specific glucosylation activity of different *E. coli* BL21(DE3) strains towards geraniol was determined by HP cultivation with adjusted optical density prior to induction (Figure 16). The strain with pET29a_VvGT14ao showed the highest specific activity with 98.8 µU/mg, followed by pET32a_VvGT14ao + pLysS with 94.2 µU/mg. The lowest specific activity was observed for pRSFDuet-1_VvGT14ao with a value of 8.7 µU/mg. In particular, strains without pLysS achieved higher specific activities than their pLysS counterparts. For example, the strain with pGEX-K_VvGT14ao showed a specific activity of 27.8 µU/mg without pLysS and only 22.0 µU/mg with pLysS.

In order to investigate other methods of induction for the HP cultivation, auto-inducing culture media were used. An auto-induction medium (AIM) based on a minimal medium, and TB (terrific broth)-AIM, a complex medium, were tested for their usability in the biotransformation.

*E. coli* BL21(DE3) /pET29a_VvGT14ao showed only little activity in the used auto-induction media. The highest activity was 2.1 µU/mg in AIM with 3 % (v/v) of ethanol, while in TB-AIM, no product could be detected. The same strain showed a specific activity of 32.0 µU/mg when used in M9$^+$ medium with a separate growth phase at 37 °C and an induction after 4 hrs with 1.25 mM IPTG.

The following experiments were performed with the improved cultivation parameters. The main cultures were inoculated with 2.5 % (v/v) of fresh pre-culture. The cultures were incubated for 4 hrs at 37 °C and 250 rpm before induction. Most strains reached an optical density of 0.4 to 0.6 during these 4 hrs. The protein production was induced with 1.25 mM IPTG and took place overnight at 18 °C and 500 rpm.

## 3.5 Testing of novel strain-vector combinations using the HP cultivation system

The generated strain-vector combinations (Table 47) were tested with the novel HP cultivation system for their glucosylation abilities towards geraniol, nerol, citronellol, and eugenol (Figure 17). In this experiment, the molar concentrations of the products were determined for direct comparison of glucosylation activities towards the different substrates.

The different host-vector combinations revealed minor preferences of host organisms towards specific substrates. The host organism *E. coli* BL21(DE3) produced the highest eugenyl glucoside titer (Figure 17, strains 1, 3, 5, 7, and 13), except when combined with a pGEX vector (Figure 17, strains 11, and 17). *E. coli* BL21(DE3)pLysS formed the most citronellyl glucoside (strains 4, 6, 8, 12, and 18), except when used with pRSFDuet-1_VvGT14ao (strain 2) or pET32a_VvGT14ao (strain 14). *E. coli* Waksman (strains 9, and 15) and *E. coli* BL21 (strains 10, and 16) seemed to prefer nerol and citronellol as a substrate, respectively. However, the sample size was small as both strains functioned only with the two pGEX vectors.

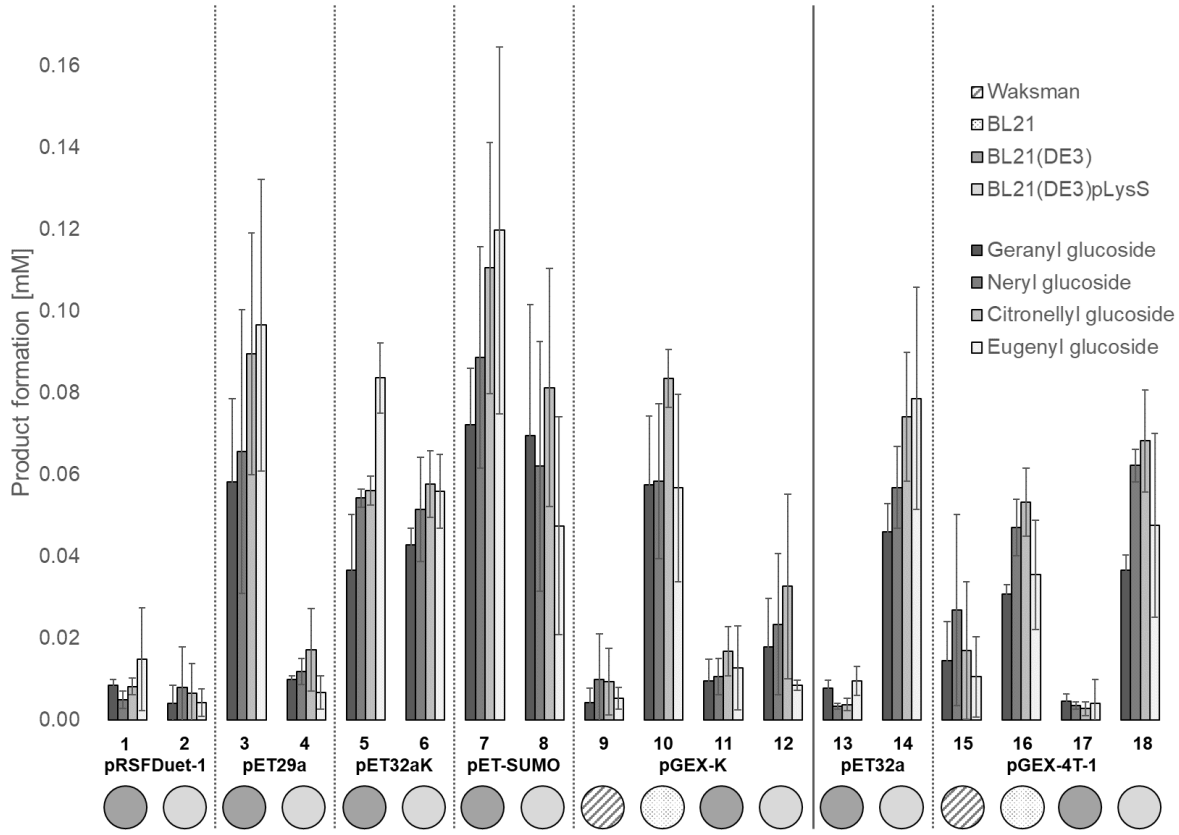


*Figure 17: Geranyl, neryl, citronellyl, and eugenyl glucoside formation using 18 different E. coli strains*

*E. coli BL21(DE3) and E. coli BL21(DE3)pLysS were combined with pRSFDuet-1, pET29a, pET32aK, pET-SUMO, pGEX-K, pET32a, and pGEX-4T-1. E. coli Waksman and E. coli BL21 were combined with pGEX-K, and pGEX-4T-1. All expression vectors contained the codon-optimized gene coding for VvGT14a. The 18 strains were used to produce geranyl, neryl, citronellyl, and eugenyl glucoside in the HP cultivation system.*

The expression vector did not seem to influence the substrate preference. Only two vectors showed a minor preference towards a substrate. Most hosts combined with a pGEX vector preferred citronellol as substrate. However, *E. coli* Waksman (strains 9 and 15) preferred nerol and *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao (strain 17) geraniol. Production strains containing a pET32 vector (strains 5, 13, and 14) seemed to prefer eugenol as substrate, except for *E. coli* BL21(DE3)pLysS /pET32aK_VvGT14ao (strain 6) which preferentially glucosylated citronellol.

*E. coli* BL21(DE3) /pET-SUMO_VvGT14ao (strain 7) preferred eugenol as a substrate like most *E. coli* BL21(DE3) production strains (strains 1, 3, 5, and 13) and produced the highest glucoside titers from all substrates tested. The preferences are not statistically significant, as the error bars were quite large, and overlapped to a large extent.

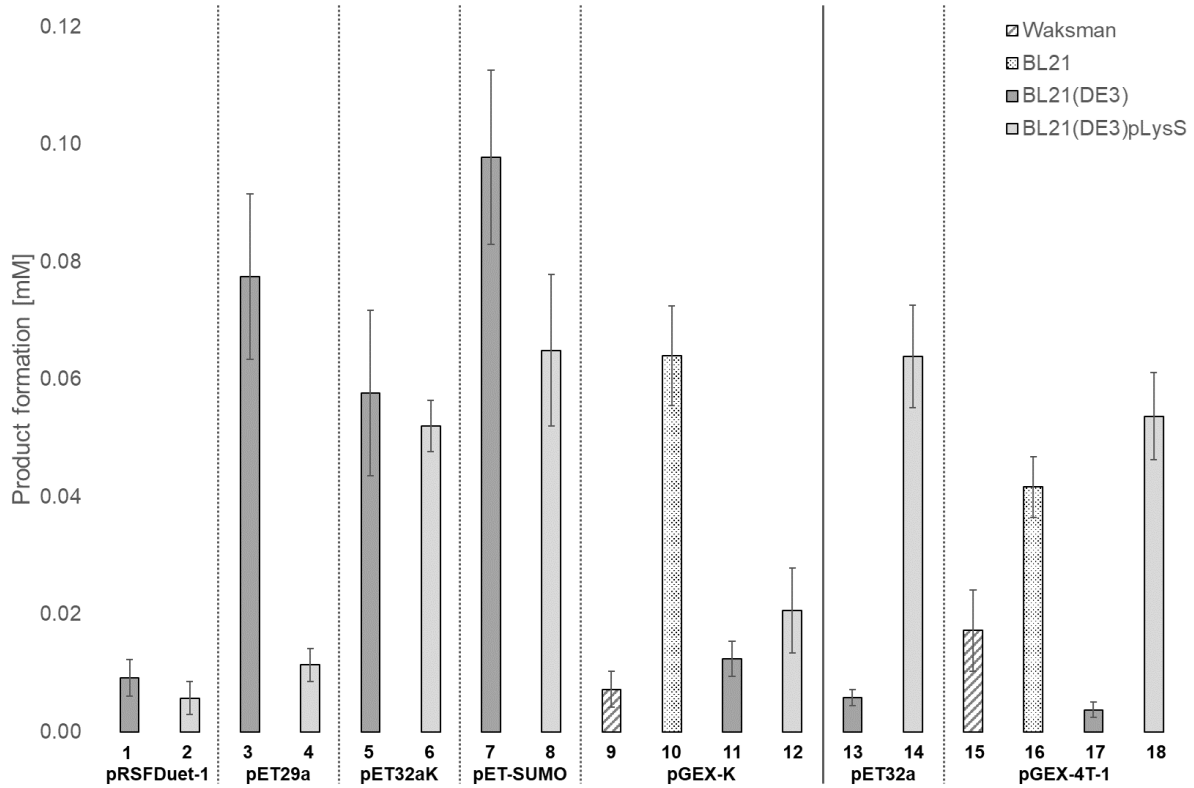


*Figure 18: Comparative total glucoside formation properties of 18 different E. coli strains*

*E. coli BL21(DE3) and E. coli BL21(DE3)pLysS were combined with pRSFDuet-1, pET29a, pET32aK, pET-SUMO, pGEX-K, pET32a, and pGEX-4T-1. E. coli Waksman and E. coli BL21 were combined with pGEX-K, and pGEX-4T-1. All expression vectors contained the codon-optimized gene for VvGT14a. The 18 strains were used to produce geranyl, neryl, citronellyl, and eugenyl glucoside in the HP cultivation system. The product formation from all four substrates was combined into one value per strain.*

Since VvGT14a did not seem to prefer any particular substrate (Figure 17) the molar glucoside amounts formed from the different substrates were added and converted to general product values (Figure 18). This allowed the categorization of the production strains into five performance groups according to their product yield: 1.) Low, 0 to 0.02 mM, 2.) Low-medium, 0.02 to 0.04 mM, 3.) Medium, 0.04 to 0.06 mM, 4.) Medium-high, 0.06 – 0.08 mM, and 5.) High, 0.08 mM glucoside yield and above.

Both production strains combined with pRSFDuet-1_VvGT14ao (strains 1 and 2) produced only low glucoside concentrations with 0.009 and 0.006 mM, respectively. In most cases, the glucoside titers produced with the various vectors were strongly dependent on the production host. *E. coli* BL21(DE3) (strain 7) produced the highest concentrations with 0.098 mM, whereas *E. coli* BL21(DE3)pLysS (strain 8) produced medium-high titers with 0.065 mM. The vector pGEX-K_VvGT14ao in combination with *E. coli* BL21 (strain 10) achieved a glucoside concentration of 0.064 mM, whereas the other strains only formed low to low-medium glucosides titers (strains 9, 11, and 12) with 0.007 to 0.021 mM. The combinations with *E. coli*

Waksman (strain 15) and *E. coli* BL21(DE3) (strain 17) produced only 0.017 and 0.004 mM of glucosides, respectively.

## 3.6 Testing of different inducer concentrations

The induction of gene expression during the HP cultivation was done by adding 25 µL of a 100 mM IPTG solution to 2 mL of culture, resulting in an inducer concentration of 1.25 mM before and 1.00 mM after substrate addition. All tested strains (Figure 17) were induced in this way. In order to test whether these concentrations are within the optimal range, concentrations between 0 and 1.5 mM IPTG were tested with two vectors in *E. coli* BL21(DE3) (Figure 19). The vector pRSFDuet-1 was chosen because of its overall poor performance in the strain-vector comparison (Figure 18).

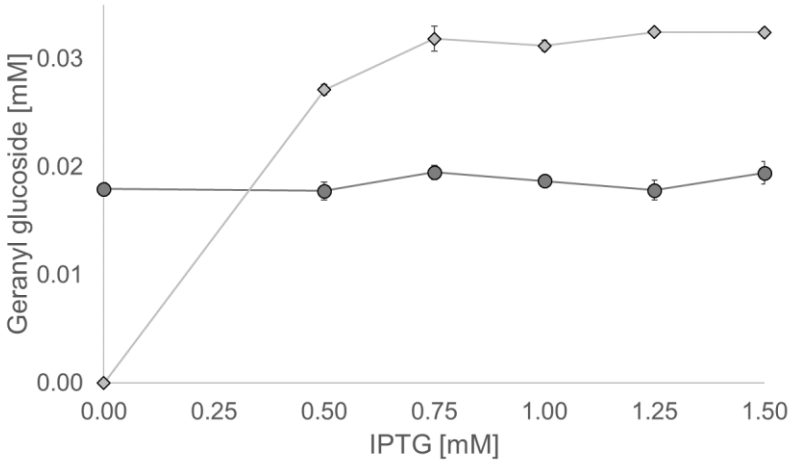


*Figure 19: Impact of the IPTG concentration on the geranyl glucoside formation*

*E. coli BL21(DE3) /pRSFDuet-1_VvGT14ao (dark grey, circles) and E. coli BL21(DE3) /pET29a_VvGT14ao (light grey, diamonds) were used to produce geranyl glucoside in the HP cultivation system. The protein production was induced with 0, 0.50, 0.75, 1.00, 1.25, and 1.50 mM of IPTG.*

The product formation of *E. coli* BL21(DE3) /pRSFDuet-1_VvGT14ao appeared to be independent of the inducer concentration. The strain displayed glycosylation activity even without protein induction by IPTG. The product concentration varied in the narrow interval between 0.018 and 0.020 mM of geranyl glucoside. *E. coli* BL21(DE3) /pET29a_VvGT14ao showed no glucosylation activity without induction. An inducer concentration of 0.75 mM IPTG seemed to be sufficient for maximal activity and resulted in 0.032 mM geranyl glucoside. When using 0.5 mM of IPTG only 0.027 mM of geranyl glucoside was obtained. Both strains did not show negative effects at higher IPTG concentrations.

## 3.7 Side activity of chloramphenicol acetyltransferase towards glycosylation substrates

The chloramphenicol acetyltransferase (CAT) is encoded on the plasmid pLysS and conveys the antibiotic resistance against chloramphenicol. CAT reportedly acts also on low molecular weight alcohols (White et al. 2017) and could thereby interfere during the whole-cell biocatalysis. To determine whether *cat* poses a problem, CAT isolated from *E. coli* was purchased from Sigma-Aldrich and its acetylation activity was experimentally tested with chloramphenicol, 1-octanol and geraniol. The enzyme displayed a specific activity of 33, 3, and 1.5 mU/mg enzyme for the substrates chloramphenicol, 1-octanol, and geraniol, respectively. In order to prevent the formation of geranyl acetate as a by-product and at the same time to ensure strict regulation by pLysS, the chloramphenicol resistance cassette was replaced with the ampicillin resistance cassette of pGEX-4T-1 (Figure 20). The resulting plasmid was termed pLysSA.

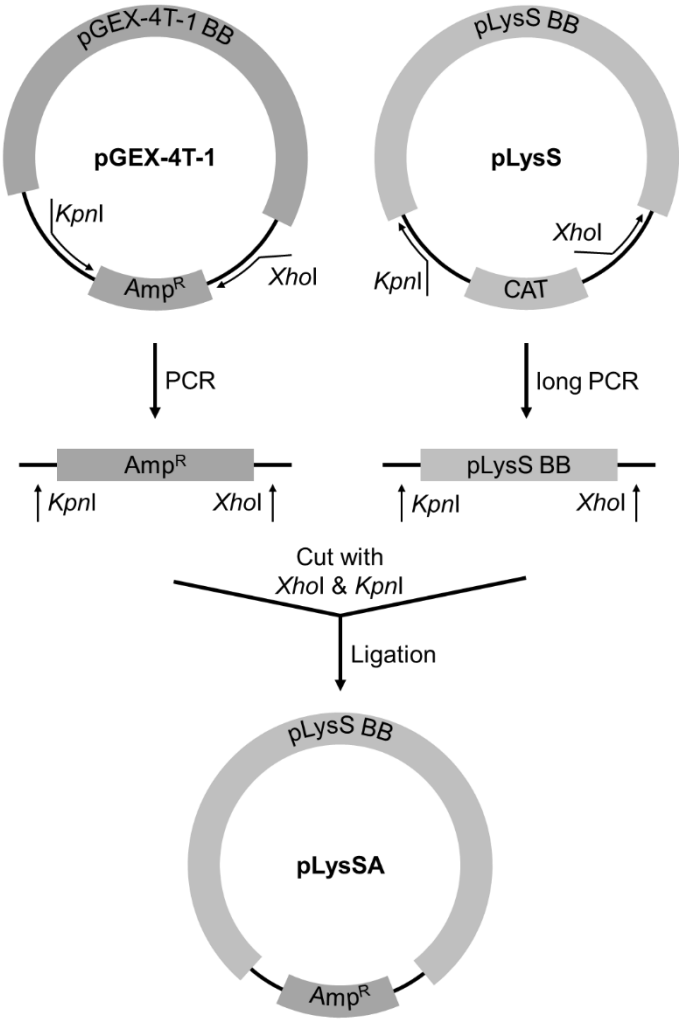


*Figure 20: Cloning of pLysSA from pGEX-4T-1 and pLysS*

*The resistance cassette from pGEX-4T-1 and the vector backbone from pLysS were amplified via PCR, digested with KpnI and XhoI, gel-purified and ligated. The final vector was termed pLysSA and confirmed by cPCR tests for the ampicillin resistance cassette and the pLysS vector backbone.*

The impact of pLysSA on acetylation side reaction and glycosylation performance was tested and compared to the corresponding *E. coli* BL21(DE3)pLysS strain (Figure 21). The elimination of the *cat* gene significantly reduced the formation of geranyl acetate and no byproduct could be detected anymore.

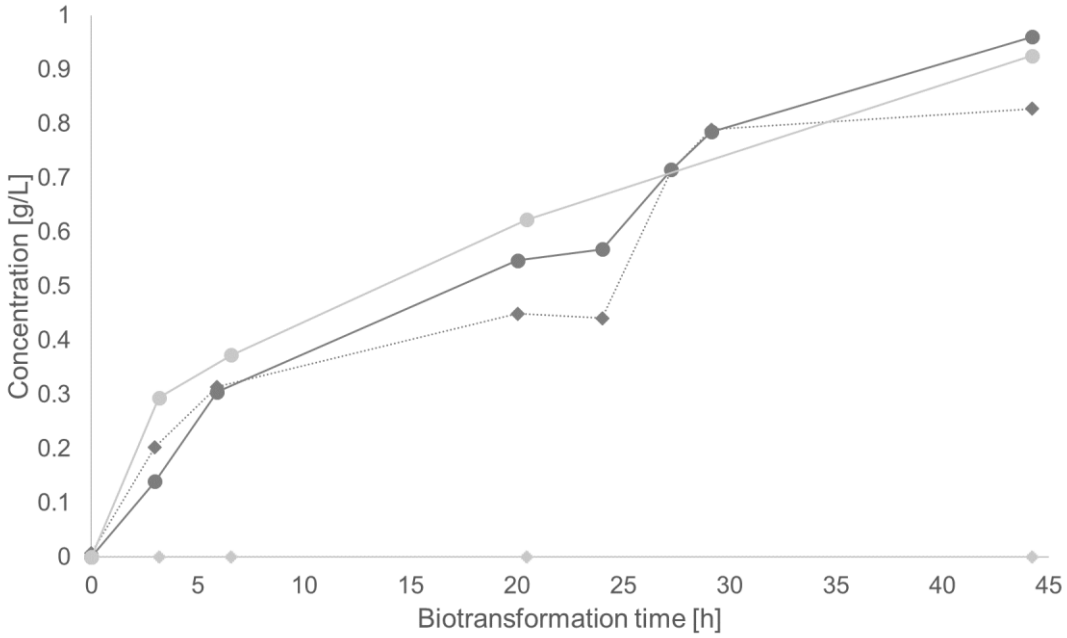


*Figure 21: Geranyl acetate and glucoside concentrations during biotransformations with E. coli BL21(DE3) /pET29a_VvGT14ao using pLysS and pLysSA*

*Geranyl acetate (diamonds) and glucoside (circles) concentrations during biotransformations with E. coli BL21(DE3) /pET29a_VvGT14ao using pLysS (dark grey) and pLysSA (light grey) at L-scale in a stirred-tank reactor. Reaction parameters pLysS strain: V = 1 L, t = 72 h, T = 30 °C, n = 1,000 rpm, pH 7, 6 g/L biocatalyst, 20 g/L batch glucose, glucose pulses (24 h: 30 g/L, 48 h: 40 g/L), 0.8 g/L batch geraniol, geraniol pulses (24/28 h: 0.8 g/L), 20 % (v/v) isopropyl myristate. Reaction parameters pLysSA strain: V = 0.4 L, t = 47 h, T = 37 °C, n = 1,400 rpm, pH 7, 16 g/L biocatalyst, 0.6 g/L batch glucose, glucose feed (3.5-7 h: 1.7 g/L*h, 7-26 h: 2 g/L*h, 26-47 h: 1.7 g/L*h), 1.6 g/L geraniol, 5 % (v/v) isopropyl myristate. This experiment was designed and performed by Xenia Priebe (BVT project partner).*

## 3.8 Directed evolution

In order to improve the biocatalyst itself, a directed evolution approach was employed. The approach consisted of the generation of a mutant library and the screening for improved mutants.

### 3.8.1 Random mutagenesis

The mutant libraries of *VvGT14ao* were created using the GeneMorph II Random Mutagenesis Kit (Agilent Technologies) according to the manufacturer's manual. Libraries 1 and 2 were created by using 100 ng of target DNA, for libraries 3, 4, 5, and 6 target DNA amounts of 500, 1000, 600, and 800 ng were used, respectively. Libraries 1, and 2 were produced with 20, and

25 PCR cycles, respectively. Libraries 3 through 6 were produced with 30 PCR cycles. The mutant genes were cloned into pRSFDuet-1 using the restriction enzymes *Bam*HI and *Not*I, resulting in an N-terminally His-tagged enzyme. *E. coli* NEB10β was transformed with the libraries in order to analyze the mutation rates (Figure 22) and to select the most suitable library for the screening.

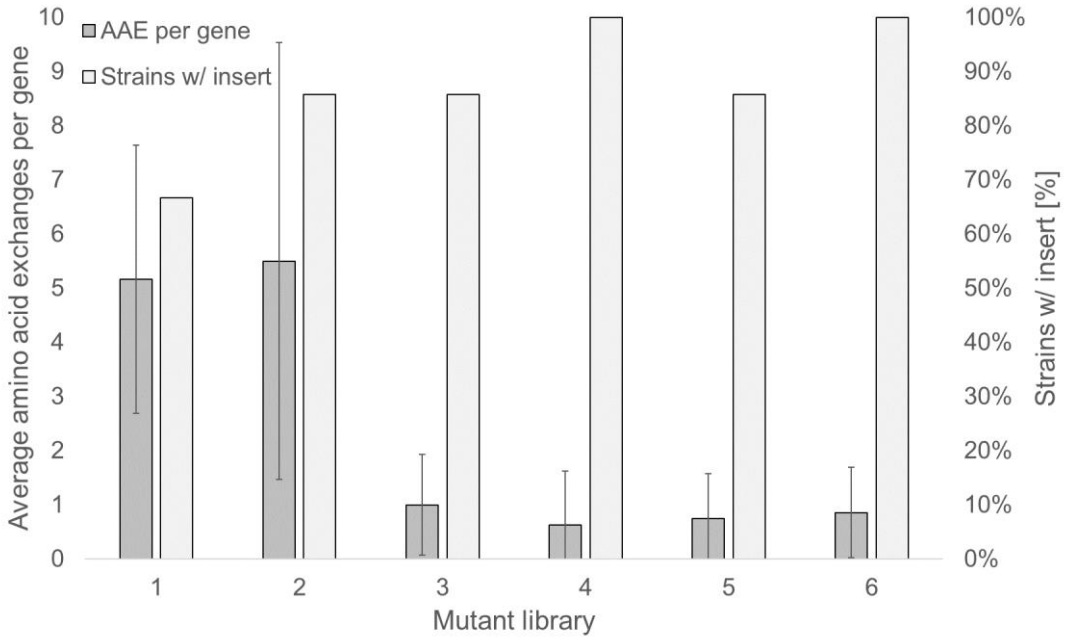


*Figure 22: Analysis of mutant libraries*

*Single colonies of the mutant libraries (E. coli NEB10β/pRSFDuet-1_VvGT14ao-variants) were picked and sequenced. The sequencing results were aligned with the VvGT14ao sequence and analyzed in terms of insert presence and exchanges in the amino acid sequence. Displayed are the average amino acid exchanges (AAE) per gene and the percentage of colonies analyzed with an insert.*

Library 4 was chosen for the screening, since the mutation rate was close to 1 AAE per gene. Furthermore, the library showed a high proportion of strains with an insert. In order to screen for better producing variants, the production host *E. coli* BL21(DE3) was transformed with library 4. Two screening approaches were employed to find better producing variants: i) a high-throughput screening using a coupled enzyme assay (*in-vitro*), and ii) a medium-throughput screening using whole-cell biocatalysis (*in-vivo*).

### 3.8.2 Screening using a coupled enzyme assay

The glycosylation of geraniol by VvGT14a cannot be measured in real-time since neither the substrate nor the product are photometrically quantifiable. The idea behind the coupled enzyme assay was to combine the glycosylation reaction tightly with a color-forming reaction and thereby enabling the photometric, real-time analysis of the GT activity. Furthermore, the UDP-glucose had to be recycled during the reaction in order to create a feasible *in-vitro* assay. Both features, color-formation and UDP-glucose recycling, were achieved in one reaction by

one additional enzyme. OleD, a glycosyltransferase, can perform the reverse glycosylation reaction under the applied reaction conditions. OleD uses UDP and 2-chloro-4-nitrophenyl-β-D-glucopyranoside (ClPNP-glucose) to form UDP-glucose and 2-chloro-4-nitrophenol (ClPNP), a yellow compound which can be quantified via photometer. The yellow color intensifies depending on the activity of the VvGT14ao variant as long as enough substrates are available for both enzymes. A weakness of the coupled enzyme assay was that any side-activity which releases UDP mimics GT activity. These side-activities can stem from the crude extract used as sample material or from OleD which is added prior to the assay. The coupled enzyme assay (Figure 23) was developed and performed at the University of Kentucky – College of Pharmacy in Lexington, KY, USA.

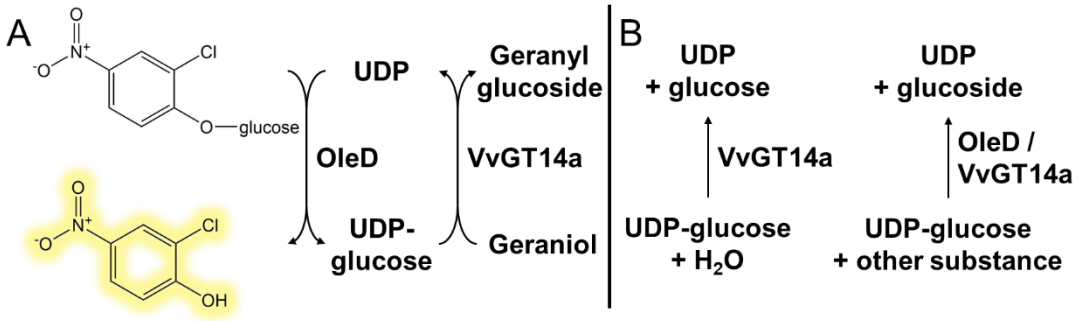


*Figure 23: Principle of the coupled enzyme assay*

*A.) The coupled enzyme assay employed two enzymes, for two separate purposes. The first glycosyltransferase, OleD, produced UDP-glucose and 2-chloro-4-nitrophenol (ClPNP) from UDP and 2-chloro-4-nitrophenyl glucoside (ClPNP-glucose). The second glycosyltransferase, VvGT14a, utilized the UDP-glucose for the glycosylation of geraniol. ClPNP-glucose is a colorless compound, the produced ClPNP however can be measured at 410 nm. The UDP pool was limited, so the color formation could only continue when the UDP-glucose was consumed. B.) Side reactions can falsely contribute to the color formation by releasing UDP without the production of geranyl glucoside.*

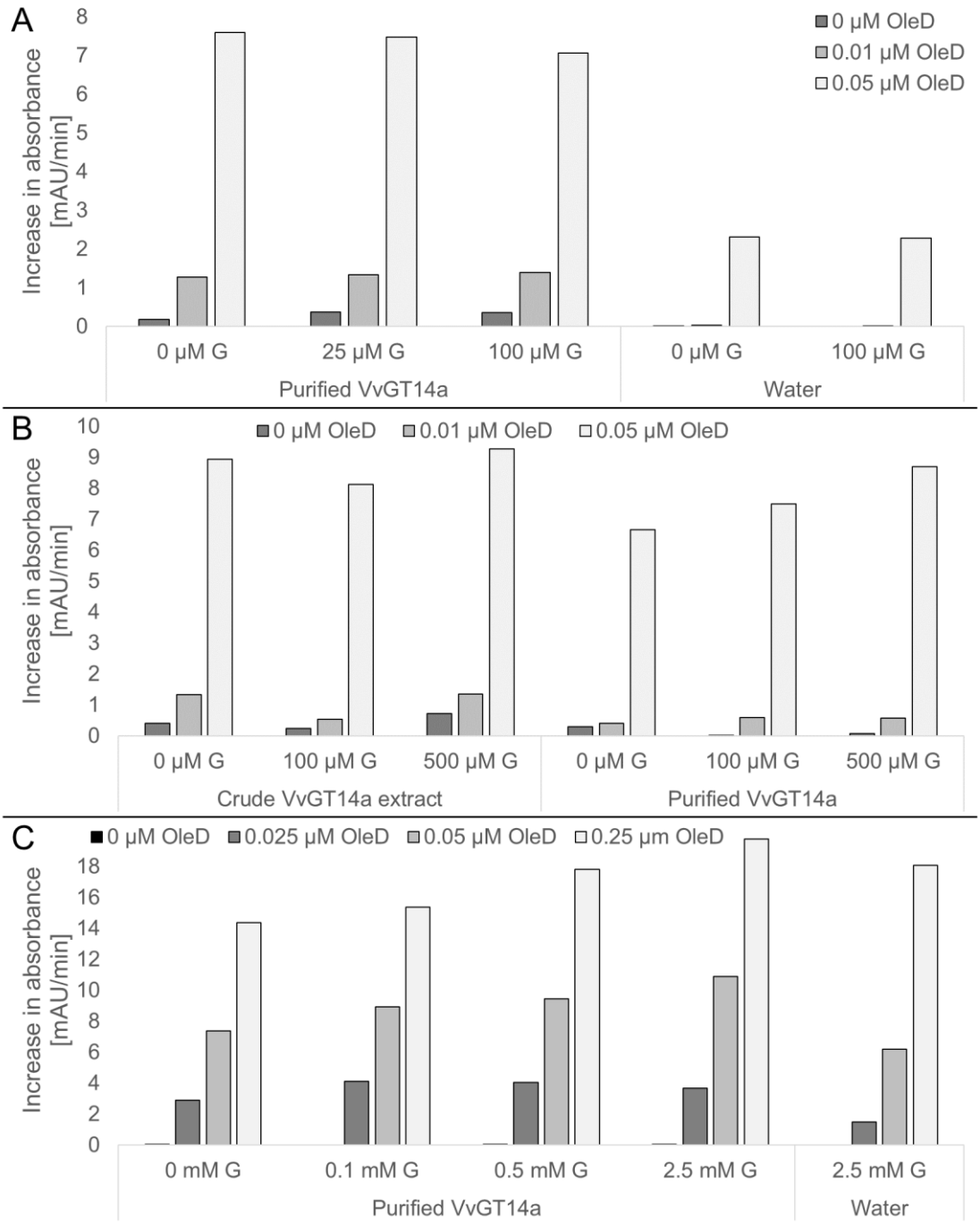


*Figure 24: Successive testing of assay conditions - OleD and geraniol concentrations*

*The concentrations of OleD and geraniol were varied in order to improve upon the assay. The concentrations were successively increased and tested without replicates. All reactions were performed with 500 µM ClPNP-glucose. A: The coupled enzyme assay was performed with varying concentrations of OleD, and geraniol (G), and 0.75 µM VvGT14a. The assay was performed with purified VvGT14a. Negative controls with water instead of VvGT14a were tested as well. The slope of the increase in absorbance was determined between 50 and 150 min reaction time. B: The coupled enzyme assay was performed with varying concentrations of OleD and geraniol (G). The assay was performed with 10 % (v/v) crude extract and purified enzyme 0.75 µM. The slope of the increase in absorbance was determined between 24 and 90 min reaction time. C: The coupled enzyme assay was performed with varying concentrations of OleD and geraniol (G). The assay was performed with 0.75 µM purified VvGT14a. A negative control with water instead of VvGT14a was tested as well. The slope of the increase in absorbance was determined over the first 90 min reaction time.*

VvGT14a was tested with ClPNP to make sure that there were no cross-activities disturbing the assay. Preliminary tests showed that VvGT14a could not glucosylate ClPNP and could therefore be used in the coupled enzyme assay. Since, the coupled enzyme assay had never been used before with VvGT14a as second enzyme and geraniol as substrate, the reaction conditions for this application had to be adjusted. VvGT14a was purified in order to optimize the assay under ideal conditions and test various variations.

The first test with geraniol concentrations of 0, 25, and 100 µM showed no increase in absorption due to the formation of geranyl glucoside. However, the background signal increased depending on the concentration of OleD and VvGT14a (Figure 24 A). The increase in absorbance was not reliable during the first few min of the reaction, since the slope was not constant. The reaction also slowed down continuously. Therefore, the linear section in the middle of the absorbance curve was used for the analysis of the increase in absorbance. When using higher concentrations of geraniol together with 0.05 µM OleD, a geraniol concentration dependent increase in absorbance could be observed (Figure 24 B). Therefore, the geraniol concentration did not yet saturate the enzyme. The crude VvGT14a samples however did not show this effect.

Even higher geraniol and OleD concentrations seemed to further intensify the increase in absorbance dependent on the geraniol concentration. The reactions without geraniol, and without VvGT14a showed, that VvGT14a contributed roughly 2 to 4 mAU/min of background gain (Figure 24 C). However, the OleD concentration had a stronger impact on the background reaction.

Crude VvGT14a extract and purified VvGT14a were used in two concentrations to test the system. The slope of the absorbance increase did not raise linearly when increasing the enzyme content from 0 to 0.38 and 0.75 µM (Figure 25 B). In order to further improve the assay, even higher geraniol concentrations were tested (Figure 25 A).

Geraniol at a certain concentration also increased the background signal without VvGT14a being present (Figure 25 A). The sensitivity of the assay was therefore not really further increased. Since samples with VvGT14a could produce UDP from UDP-glucose and increase absorbance even in the absence of geraniol (Figure 25 A), it was assumed that VvGT14a had UDP-glucose hydrolase activity. Coupled enzyme analyses were performed to test for UDP-glucose hydrolase activity and the crude VvGT14a extract showed the strongest color-formation with 5.0 mAU/min. The purified enzyme still displayed a high background with 3.6 mAU/min. The crude extract from the empty vector strain also showed an increase in absorbance, however the increase was not as pronounced with only 1.6 mAU/min. The water control displayed a slight increase in absorbance with 0.3 mAU/min.

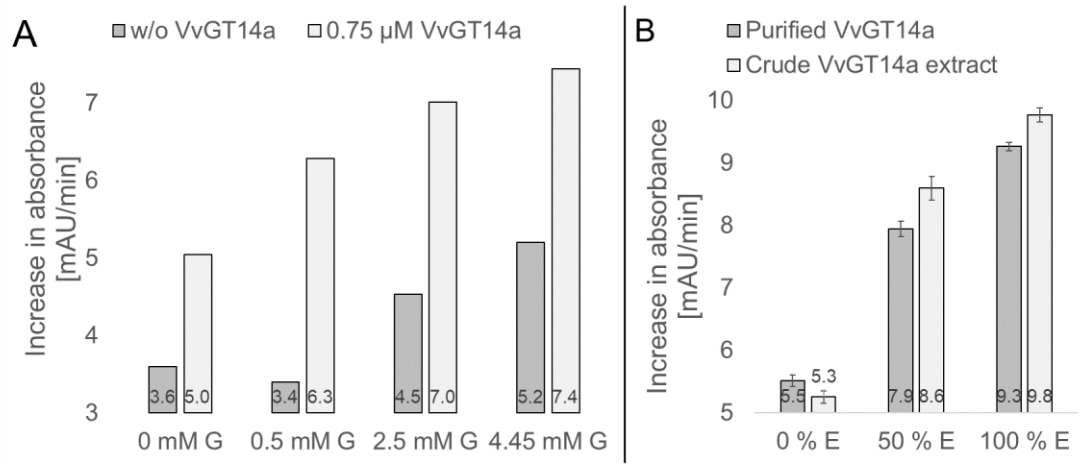


*Figure 25: Testing of assay conditions - geraniol and enzyme concentrations*

*The slope of absorbance increase was determined within the first 90 min. A: The coupled enzyme assay was performed with varying concentrations of geraniol (G) and 0.75 µM purified VvGT14a. Negative controls with water instead of VvGT14a were tested as well. The reaction was carried out with 500 µM ClPNP-glucose. B: The coupled enzyme assay was performed with varying concentrations of VvGT14a (E), crude extract (10 % (v/v) crude extract solution) and purified enzyme (0.75 µM). The enzyme solutions were used undiluted (100 %) and 1:1 diluted (50 %). The reaction was carried out with 500 µM ClPNP-glucose, and 2.5 mM geraniol.*

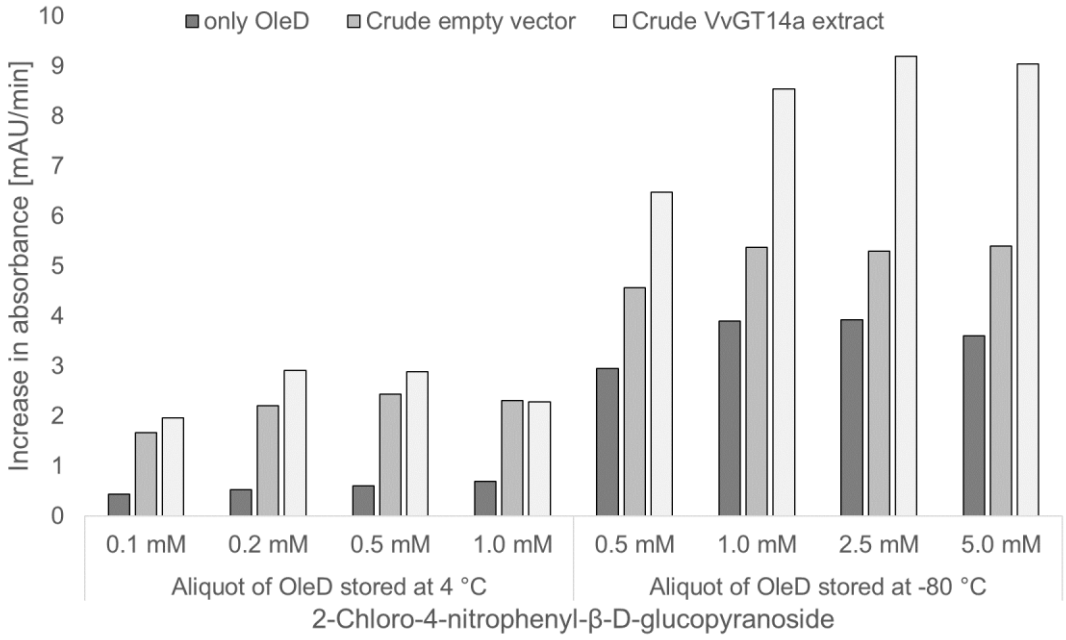


*Figure 26: Testing of assay conditions - ClPNP-glucose concentration*

*The coupled enzyme assay was performed with crude extracts of an empty vector strain and VvGT14a using varying concentrations of ClPNP-glucose. A negative control with water was tested as well (only OleD). One series was performed with OleD stored at 4 °C and one with OleD stored at -80 °C. The slope of the increase in absorbance was determined within 30 and 90 min.*

Experiments with differently stored OleD showed that the enzyme lost a significant part of its activity over four weeks of storage at 4 °C (Figure 26). Higher ClPNP-glucose concentrations showed only a small impact on the reaction speed. The highest difference in absorbance between empty vector and wild type enzyme could be observed at 2.5 mM ClPNP-glucose

using freshly thawed OleD (Figure 26). The coupled enzyme assay with optimized parameters displayed a clear difference between crude extract samples from empty vector and VvGT14a wild type strains grown and induced in a deep well plate with a slope of 5.3 and 10.4 mAU/min, respectively (Figure 26). Thus, the slope of the VvGT14a wild type was approximately 5.1 mAU/min higher than the slope of the empty vector. The absorbance increase of the water control was 3.1 mAU/min. The parameters tested within this experiment were used for the screening of VvGT14ao variants.

Several cultivation media were tested to determine their suitability for direct cultivation from a cryogenic stock. This was attempted to avoid the need for a freshly prepared pre-culture for the inoculation. Experiments in lysogeny broth (LB) medium with IPTG induction using 1 mM of inducer did not show sufficient separation between the positive control (wild type strain) and the negative control (empty vector strain). The 2xYT medium induced with 1 mM of IPTG and LB auto-induction medium showed no better separation. A pre-culture seemed to be necessary. The use of an overnight pre-culture grown in LB medium resulted in a visible difference. The main culture was grown in 1 mL of 2xYT medium supplemented with 2 % (w/v) of glucose and was inoculated with 100 µL of the pre-culture grown in LB medium. The induction was done after 4 hrs of cultivation with 1 mM IPTG. The distinction between the positive (A1) and negative (B1) control was now visible at about 2.4 mAU/min (Figure 27 B).

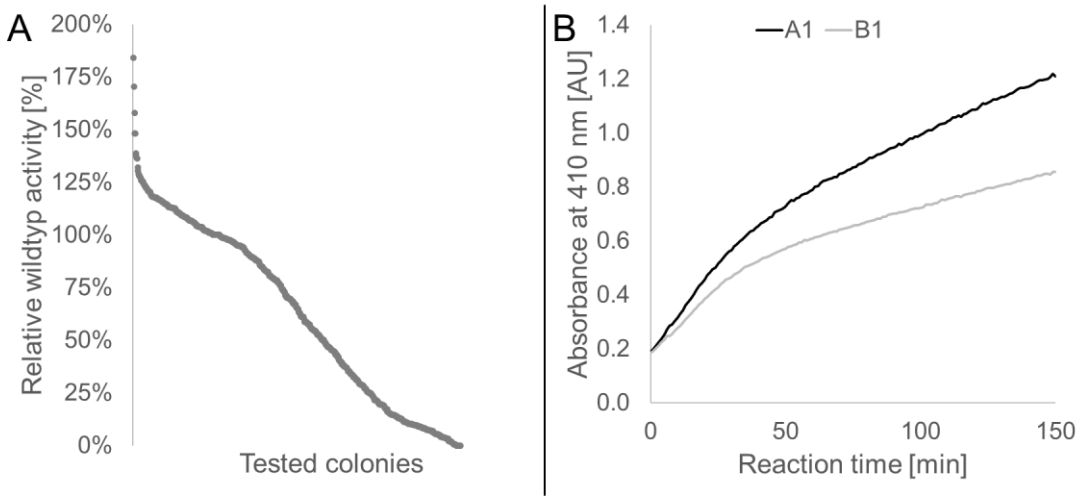


*Figure 27: Screening result of the coupled enzyme assay*

*A: Accumulated result. The VvGT14a wild type activity was set 100%. The samples were ordered by their value from highest to lowest. B: Positive (A1) and negative (B1) control from the first screening plate. The optimized coupled enzyme assay was used for the screening of VvGT14ao variants. Shown are the positive control (black; E. coli BL21(DE3) /pRSFDuet-1_VvGT14ao) and the negative control (grey; E. coli BL21(DE3) /pRSFDuet-1) of the first screening plate. The slope of the increase in absorbance was determined over the reaction time between 30 and 90 min.*

*E. coli* BL21(DE3) colonies from the screening library agar plates were picked and collected in 96 deep well screening plates. Each screening plate was equipped with a positive control in

well A1 (*E. coli* BL21(DE3) /pRSFDuet-1_VvGT14ao) and a negative control in well B1. (*E. coli* BL21(DE3) /pRSFDuet-1). The VvGT14ao variants were grown in the screening plates, lysed, and the centrifuged. The cleared supernatants were then used as sample material in the coupled enzyme assay.

In total, 846 putative mutants were screened with the coupled enzyme assay and sorted by their calculated, relative activity (Figure 27 A). The screening results showed that none of the variants showed an activity twice (200 %) as high as the VvGT14a wild type (Figure 27 A). The four cultures with the highest activities in the screening were picked for further analysis. Gene sequencing showed that the screening hits were VvGT14ao variants: VvGT14ao-N45D-T442A, VvGT14ao-H371N, VvGT14ao-M211K, and VvGT14ao-M1_P241del-V431M. The variants were further investigated by whole-cell biotransformation, however the geranyl glucoside yield of all variants was lower or in case of the truncated enzyme VvGT14ao-M1_P241del-V431M (deletion mutant) not detectable.

The deletion mutant was investigated further since it showed elevated (UDP-forming) activity in the screening but absolutely no product formation in whole-cell biocatalysis. Therefore, the shortened enzyme was purified and fractions tested *in-vitro* for UDP-glucose hydrolase activity (Figure 28).

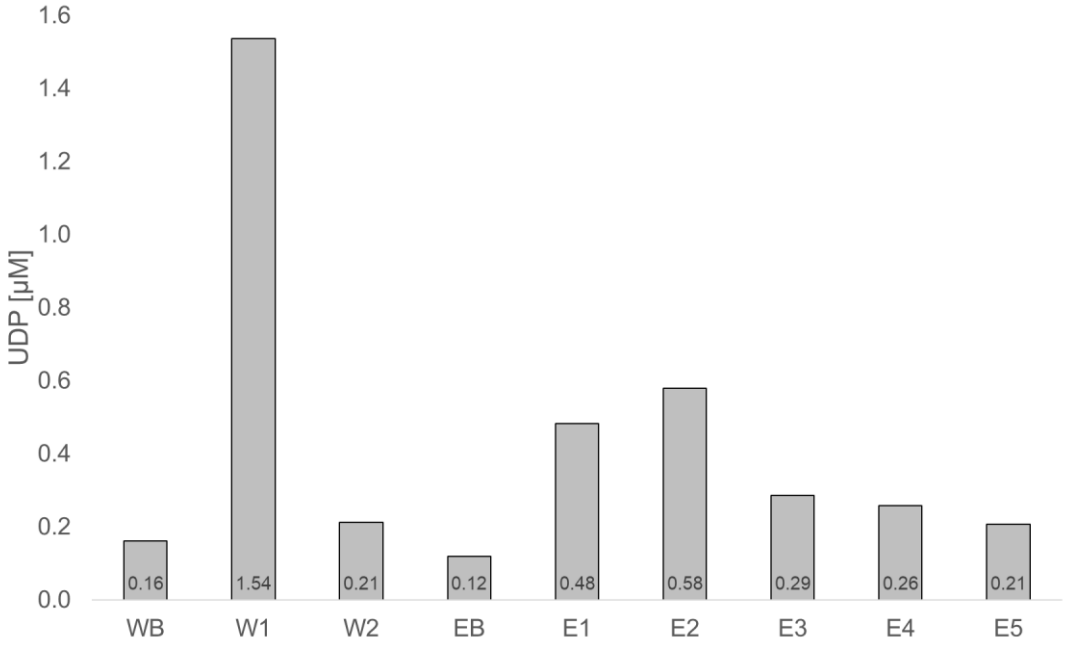


*Figure 28: Hydrolase activity of protein purification fractions from VvGT14ao-M1_P241del-V431M*

*The UDP-Glo assay was used for the determination of UDP-glucose hydrolase activity of protein purification fractions of VvGT14ao-M1_P241del-V431M (deletion mutant). The analyzed samples were wash buffer (WB), wash fraction 1 and 2 (W1 and W2), elution buffer (EB), and elution fractions 1 through 5 (E1 – E5). The reaction was performed at 30 °C for 60 min with 100 µM of UDP-glucose.*

The analysis of the protein purification fractions showed that the His-tagged protein, VvGT14ao-M1_P241del-V431M, does indeed have UDP-glucose hydrolase activity. The elution fraction E2 produced 0.6 µM of UDP. The highest total activity was found in the sample of wash fraction 1, which hydrolyzed 1.5 µM of the substrate (Figure 28).

### 3.8.3 Screening using the HP cultivation system and HPLC analysis

Since the first screening of the mutant library revealed a number of false positive variant strains, a different screening approach was performed. The second screening of the mutant library was done with the HP cultivation system and product analysis via HPLC. The advantage of the second screening was the detection method, since only strains with a proven increase of product yield were identified as hits. The same mutant library DNA as in the first screening was used for the transformation of *E. coli* BL21(DE3). Transformants were collected on agar plates for short-term storage and the inoculation of pre-cultures. Leftover pre-culture material was used to prepare cryo cultures of the tested strains, for long-term storage and potential further analysis. In total, 170 putative variants were tested with this method (Figure 29).

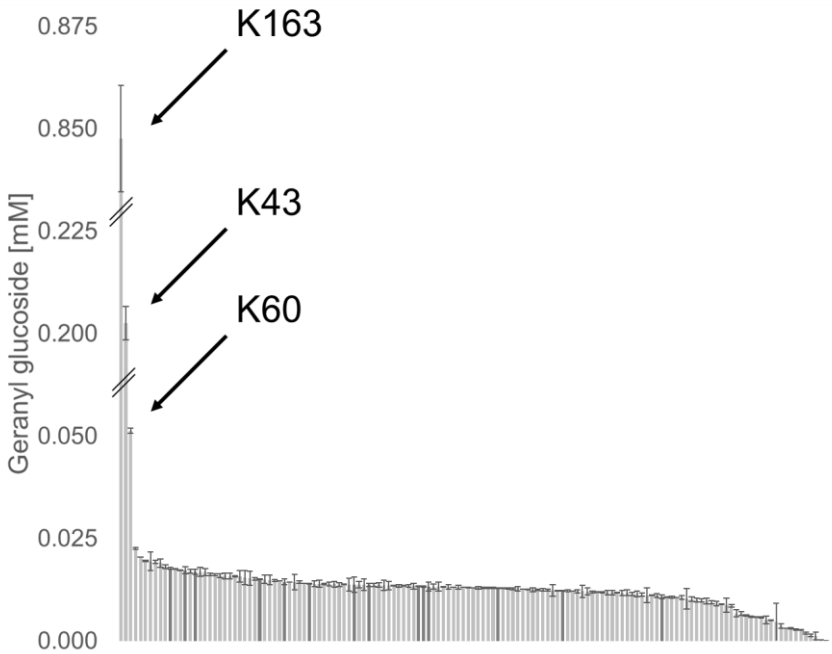


*Figure 29: Summarized result of the screening using the HP cultivation system*

*The product formation of putative mutants was tested with the HP cultivation system. The tested strains were sorted by their product yield from highest to lowest. Standard deviations are shown. The dark grey bars indicate the product formation by the wild type strain E. coli BL21(DE3) /pRSFDuet-1_VvGT14ao.*

Screening with HP cultivation and subsequent HPLC analysis resulted in three putative screening hits with an increased geranyl glucoside formation (Figure 29). Similar to the first screening, a number of empty vector, wild type and non-productive mutant strains were found in the mutant library. The strains K43, K60, and K163 were analyzed further since they produced concentrations of 0.203, 0.051, and 0.848 mM geranyl glucoside, respectively

compared to less than 0.025 mM for the VvGT14a wild type strain. The expression plasmids from the putative mutants were extracted and sequenced. The plasmid sequencing of K43 revealed the wild type sequence. The sequencing result of K60 revealed the empty vector sequence, and the result of K163 revealed a double mutant, VvGT14ao-S168N-W353R. The increased activities of K43, K60, and K163 were confirmed by additional biotransformations and the sequencing reactions of K43, and K60 were repeated. However, the second sequencing result yielded the same outcome.

### 3.8.4 Candidates K43, K60, and K163

In order to rule out plasmid related effects, the strains K43, and K60 were cured of their plasmids and tested for their glucoside production capabilities. The cured strains *E. coli* BL21(DE3)K43, and *E. coli* BL21(DE3)K60 were unable to produce geranyl glucoside. The cured strains were transformed with the wild type vector, pRSFDuet-1_VvGT14ao, and were tested again for glucoside production. The newly transformed strains *E. coli* BL21(DE3)K43 /pRSFDuet-1_VvGT14ao, and *E. coli* BL21(DE3)K60 /pRSFDuet-1_VvGT14ao produced geranyl glucoside. The new K43 strain showed increased geranyl glucoside concentrations like the original K43 screening hit but the new K60 strain produced only geranyl glucoside titers equal to the wild type VvGT14ao strain.

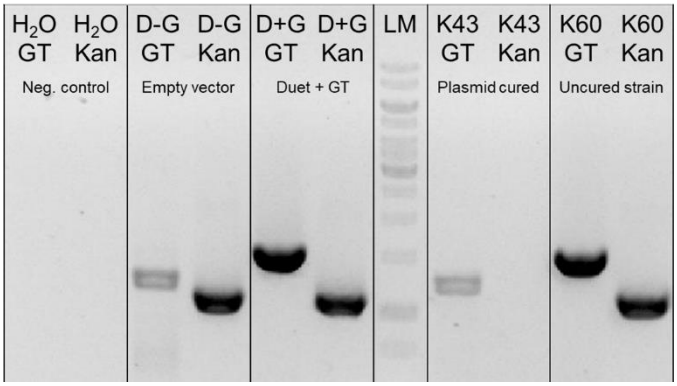


*Figure 30: Testing for VvGT14ao and KanR2 in K43 and K60 using cPCR*

*A colony PCR was performed with the screening hit K60 and the cured K43 strain to test for the glycosyltransferase gene VvGT14ao (GT) and the kanamycin resistance cassette of pRSFDuet-1 (Kan) with specific primers. Controls were run with water ($H_2O$), empty vector strain (D-G), and wild type strain (D+G) as sample. GeneRuler DNA Ladder Mix was used as Ladder (LM) in order to appraise DNA bands.*

The original K60 strain from the screening and the cured K43 strain were tested for the presence of *KanR2* (the resistance cassette from pRSFDuet-1) and *VvGT14ao* with cPCR using specific primers (Figure 30). The results revealed that the cured K43 strain contained neither the gene coding for VvGT14a nor the pRSFDuet-1 vector. The result of K60 not only indicated that the pRSFDuet-1 vector was present in the strain, as the sequencing reaction had already shown, but also contained a glycosyltransferase gene coding for a supposed

VvGT14ao variant. Therefore, the K60 strain must contain both empty vector and pRSFDuet-1 with a GT gene. The cPCR product of K60 obtained with specific *VvGT14ao* primers was sequenced and revealed the DNA sequence coding for a VvGT14ao variant, VvGT14ao-T398S. In addition, a closer look at the sequencing results also showed that the wild type sequence was also present in the sample, i.e. K60 contains at least three different vectors at different concentrations. Therefore, this mutant was not further investigated.

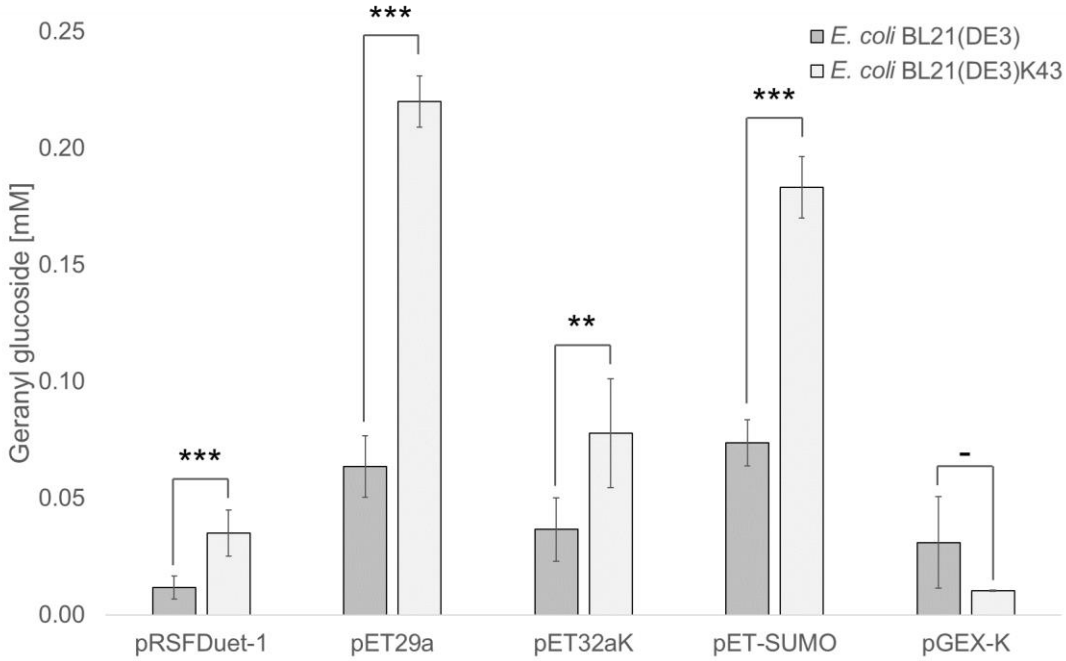


*Figure 31: Geranyl glucoside titers produced by different E. coli BL21(DE3)K43 strains*

*E. coli BL21(DE3)K43 was transformed with pRSFDuet-1_VvGT14ao, pET29a_VvGT14ao, pET32aK_VvGT14ao, pET-SUMO_VvGT14ao, and pGEX-K_VvGT14ao and tested for its geranyl glucoside production capability against the original E. coli BL21(DE3) strains. The results were obtained with the HP cultivation system. The statistical analysis was done with an unpaired t test. - not statistically significant, * p<0.05, ** p<0.01, and *** p<0.001.*

From the results so far, we concluded that the strain *E. coli* BL21(DE3)K43 had a beneficial mutation in its genomic background. To test this hypothesis, K43 was equipped with different vectors. Strain-vector combinations with expression vector other than pRSFDuet-1_VvGT14ao also showed increased geranyl glucoside production in comparison with the original strain (Figure 31). The only exception was the combination with the pGEX-K vector, where the original strain *E. coli* BL21(DE3) /pGEX-K_VvGT14ao produced more product than the corresponding K43-strain. However, the difference was statistical insignificant (Figure 31).

The genomes of K43 and the original laboratory strain were sequenced and compared to the published *E. coli* BL21(DE3) genome AM946981 (Krempl et al. 2008). The genome of K43 contained five regions with one or more mutations that were not present in the laboratory strain (Figure 32). The first region contained one single nucleotide polymorphism (SNP) at position

110,370, which is far outside of any coding sequence. The second region was located between positions 2,929,276 and 2,929,309, and contained five SNPs and two indels. This region was downstream of two flanking coding sequences. The third region was located at position 2,964,343, 82 base pairs upstream of *yghE* (coding for a putative type II secretion system L-type protein) and contained only one SNP. The fourth and biggest region was located in the coding sequence for *yghJ* (coding for a predicted inner membrane lipoprotein), between positions 2,975,823 and 2,977,059, and showed 41 SNPs and six indels. The last region contained only one indel, downstream of the coding sequence of rho (transcription termination factor) at the position 3,857,938.

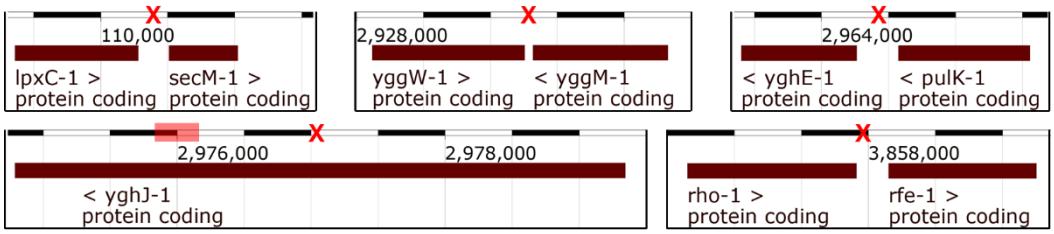


*Figure 32: The five regions of the K43 genome containing mutations*

*The loci containing mutations (red) were elucidated with the E. coli BL21(DE3) genome AM946981 used as a reference. One black and white bar together represent 1,000 base pairs of the genome. All mutation loci are in between coding sequences, except for the mutation region four within yghJ.*

The coding sequence of *yghJ* from the original laboratory strain differed strongly from the literature sequence of *E. coli* BL21(DE3) AM946981 (Krempl et al. 2008) and showed 186 SNPs and indels. The *yghJ* gene from K43 showed 233 SNPs and indels in comparison to the literature genome data. The *yghJ* genes from the laboratory strain and K43 both code for a full-length protein, containing multiple variations in the amino acid sequence. Most mutations in *yghJ* from K43 were located between 2,975,874 and 2,976,172, where a peptidase M60 domain is located.

The third screening hit K163, the double mutant VvGT14ao-S168N-W353R, was further investigated and the corresponding single mutations S168N and W353R were generated by recombination to analyze the source of the increased product yield.

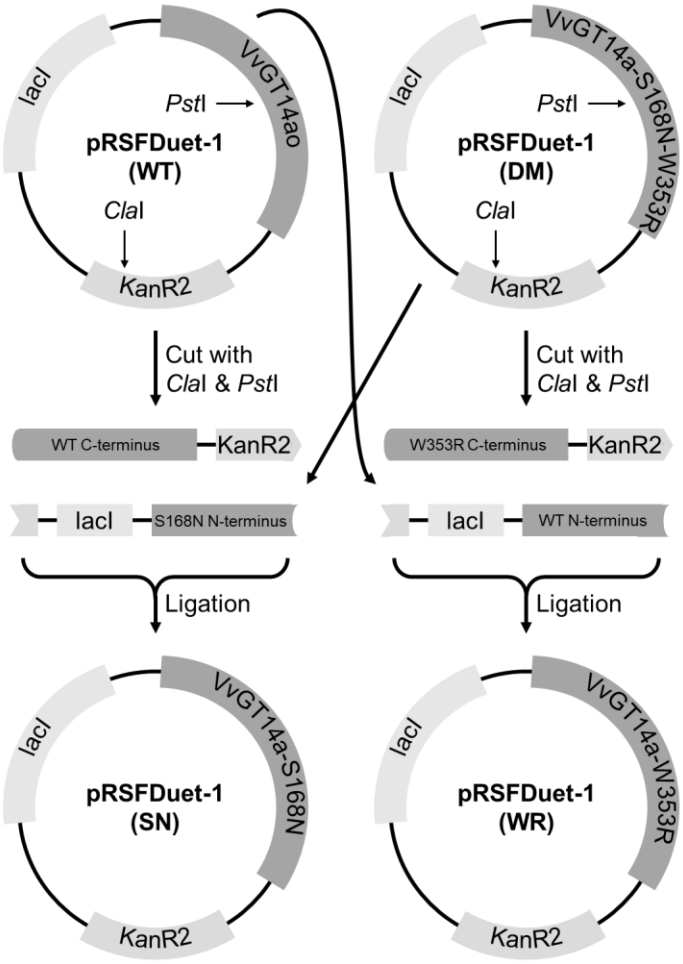


*Figure 33: Cloning of VvGT14ao variants S168N and W353R by combining VvGT14ao-S168N-W353R with VvGT14ao*

*The single mutants from VvGT14ao-S168N-W353R were created by combining the different vector parts of the double mutant (DM; pRSFDuet-1_VvGT14ao-S168N-W353R) with the complementary parts of the original vector (WT; pRSFDuet-1_VvGT14ao). The vectors were cleaved twice, once between the two mutation positions in the gene with PstI and once more in the vector's backbone with ClaI. The vector parts were separated and combined with the corresponding counterpart. The final vector products, pRSFDuet-1_VvGT14ao-S168N (SN) and pRSFDuet-1_VvGT14ao-W353R (WR) were obtained after ligation.*

The expression vectors of the single mutants from K163 were obtained by combining different parts of the mutant vector with the complementary parts of the wild type vector. The vector parts were obtained by cutting with *Pst*I and *Cla*I, followed by gel-purification. The DNA fragment coding for the C-terminus of the double mutant was ligated to the DNA fragment coding for the N-terminus of the original vector to obtain the S168N mutant. The remaining two DNA fragments were ligated to generate the W353R mutant (Figure 33). The produced vectors pRSFDuet-1_VvGT14ao-S168N and pRSFDuet-1_VvGT14ao-W353R were verified by sequencing. *E. coli* BL21(DE3) was transformed with the single mutant vectors and their biotransformation capabilities compared with those of the double mutant and the original strain. The *VvGT14ao* variants of K163 were also cloned into pET29a to test the transferability of

these mutations. *E. coli* BL21(DE3) was transformed with the new vectors and the glycosylation activities were compared between the variants and the wild type vectors (Figure 34).

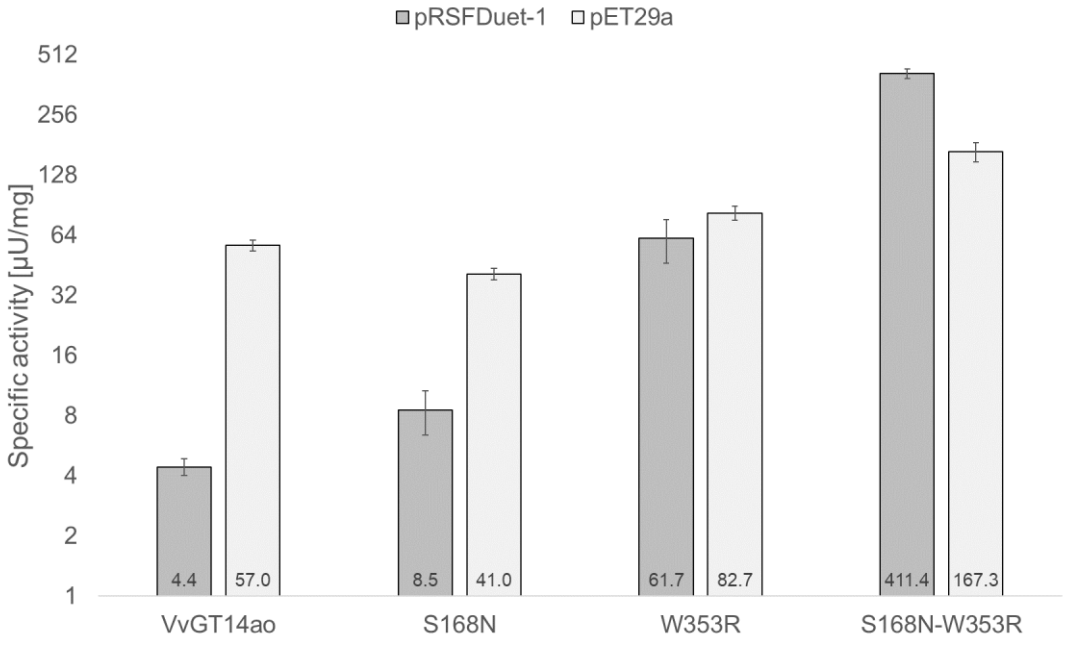


*Figure 34: Specific strain activity of the original VvGT14ao, single mutants and double mutant in pRSFDuet-1 and pET29a (logarithmic scale)*

*E. coli BL21(DE3) was transformed with pRSFDuet-1, and pET29a containing VvGT14ao, VvGT14ao-S168N, VvGT14ao-W353R, and VvGT14ao-S168N-W353R. The combinations were tested for their geranyl glucoside production capabilities. The results were obtained with a HP cultivation.*

The VvGT14ao mutants W353R and S168N-W353R in the expression vectors pRSFDuet-1, and pET29a combined with *E. coli* BL21(DE3) showed higher glucosylation activity than the original VvGT14ao strains. The VvGT14ao-S168N variant showed only increased activity in the pRSFDuet-1 vector. In pET29a the activity seemed to be decreased (Figure 34).

To further investigate the K163-derived VvGT14ao variants, their *in-vitro* enzymatic activity was quantified with the UDP-Glo assay, which determines the release of UDP during the glucosylation reaction. Since the assay required purified enzyme as sample material, the variants were therefore cloned in pGEX-K to provide a GST-tag for easy purification. *E. coli* NEB10β was transformed with the new constructs. The new plasmids were verified by sequencing. *E. coli* BL21(DE3) was transformed with the new expression vectors and used for protein purification.

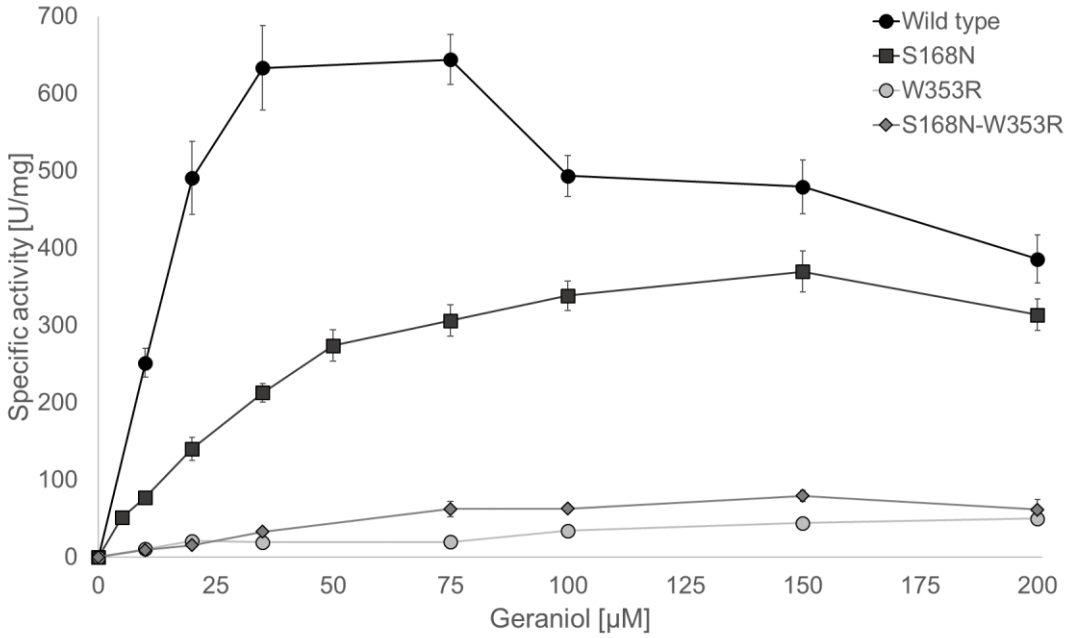


*Figure 35: Kinetic data of the K163-derived VvGT14ao variants*

*The UDP-Glo assay was used to determine the glucosylation activity of the GST-tagged VvGT14ao variants towards geraniol. The protein concentration of the purified enzymes was used to calculate the specific activities of the enzyme variants. The reaction was performed at 30 °C for 15 min with 100 µM of UDP-glucose and varying concentrations of geraniol.*

The *in-vitro* data obtained by the UDP-Glo assay showed reduced enzymatic activities for all the VvGT14ao variants (Figure 35). The calculated kinetic parameters of the variants also show reduced maximum specific activities compared to the wild type. The $K_M$ values of the single mutants were similar to the wild type, whereas the $K_M$ value of the double mutant was greatly impaired (Table 48).

*Table 48: Kinetic data of VvGT14ao variants*

| **VvGT14ao variant** | **$v_{max}$ [U/mg]** | **$K_M$ [µM]** | **$k_{cat}$ [1/s]** |
|---|---|---|---|
| WT | 2100 ± 1180 | 72 ± 49 | 1850 ± 1040 |
| S168N | 581 ± 39.3 | 64 ± 6 | 512 ± 35 |
| W353R | 58.2 ± 10.7 | 44 ± 12 | 51.2 ± 9.5 |
| DM | 219 ± 129 | 230 ± 148 | 193 ± 113 |

Even though the VvGT14ao variant K163 and its derivatives showed reduced enzymatic activities *in-vitro*, the whole-cell biocatalysts always yielded increased product concentrations. In order to further investigate the double mutation, the corresponding amino acids exchanges were performed in other glycosyltransferases. To transfer the K163-derived mutations of VvGT14a in VvGT15a and VvGT16, the loci for the mutations were determined by structure based sequence alignment (T-Coffee) by PROMALS3D (Pei et al. 2008). The alignment

showed highly similar sequences in the vicinity of the serine residue (S168) for VvGT14a and VvGT16 (Figure 36). The serine residues at position 152 and 170 in VvGT15a and VvGT16, respectively were exchanged with asparagine. The second amino acid of the double mutant, W353 is the tryptophan residue which determines the beginning of the highly conserved PSPG box of UGTs. The corresponding tryptophan residues of VvGT15a and VvGT16 at positions 321 and 354, respectively were exchanged with arginine.

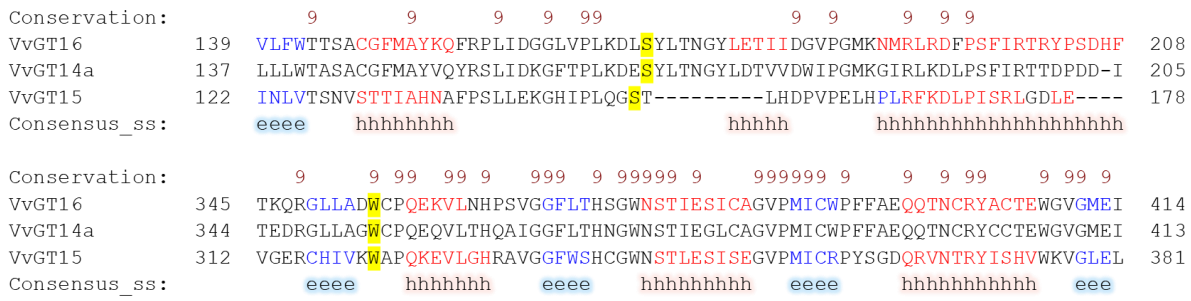


*Figure 36: Structure-based sequence alignment of VvGT14a, VvGT15a, and VvGT16*

*The amino acid sequences of VvGT14a, VvGT15a, and VvGT16 were aligned according to their predicted structure using PROMALS3D (Pei et al. 2008). The alignment was used to locate the corresponding amino acid sites of the VvGT14ao-S168N-W353R variant in VvGT15a and VvGT16. The chosen sites are marked in yellow. Red color indicate alpha-helices (h), blue color beta-strands (e). Complete conservation is indicated by a 9 above the corresponding amino acids.*

The variants of VvGT15a and VvGT16 were cloned by Jiajia Zheng during a research internship. The variants were inserted in the vector pGEX-4T-1, resulting in the following vectors: pGEX-4T-1_VvGT15a-S152N, pGEX-4T-1_VvGT15a-W321R, pGEX-4T-1_VvGT15a-S152N-W321R, pGEX-4T-1_VvGT16-S170N, pGEX-4T-1_VvGT16-W354R, and pGEX-4T-1_VvGT16-S170N-W354R. *E. coli* BL21(DE3) was transformed with the variant vectors and the resulting strains were tested for their ability to produce geranyl glucoside (Figure 37). The beneficial effect of the K163-derived variants is not transferable to VvGT15a and VvGT16. The product concentrations of the variants VvGT15a-W321R, VvGT15a-S152N-W321R, VvGT16-W354R, and VvGT16-S170N-W354R were extremely low compared to the corresponding wild type enzyme. The titers of variants VvGT15a-S152N, and VvGT16-S170N were not different from those of the original enzymes, however no positive effect could be observed either (Figure 37). The positive effects of VvGT14ao-W353R, and VvGT14ao-S168N-W353R seen in pRSFDuet-1 and pET29a, were also observed in *E. coli* BL21(DE3) using the pGEX-K vector (Figure 34 and 37). The K163-derived variants of VvGT15a and VvGT16 were also tested *in-vitro* for their glucosylation ability towards geraniol and benzyl alcohol, respectively. The enzymes were purified via their GST-tag from *E. coli* BL21(DE3) and tested with the UDP-Glo assay.

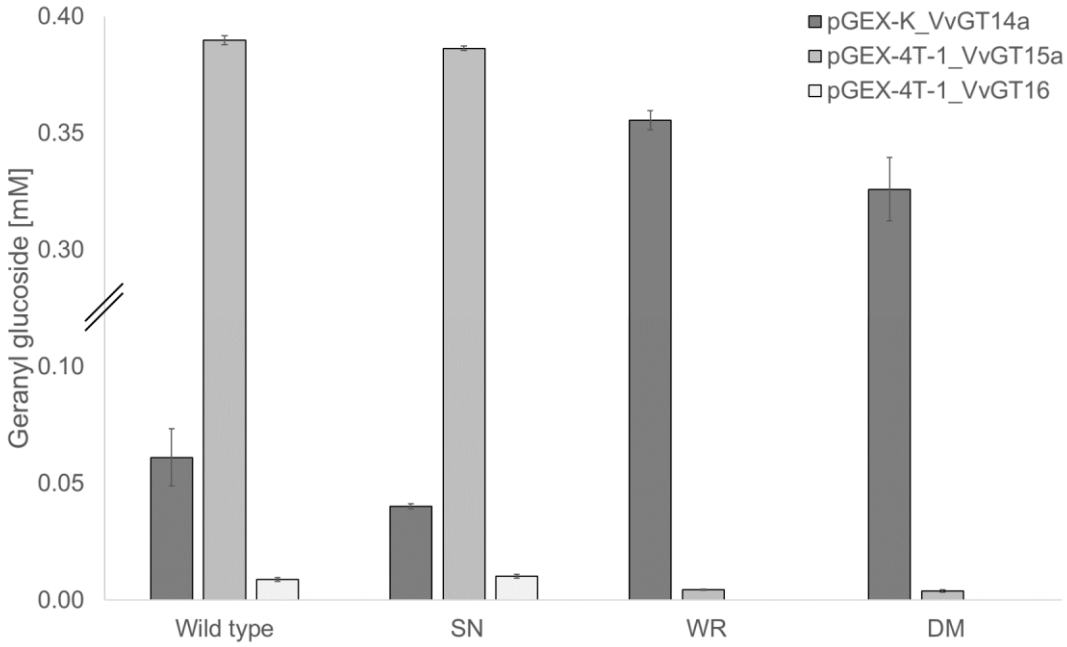


*Figure 37: Geranyl glucoside titers produced by E. coli BL21(DE3) transformed with K163-derived mutants of VvGT14ao, VvGT15a, and VvGT16 in-vivo*

*E. coli BL21(DE3) was transformed with pGEX-4T-1 vectors containing K163-derived mutants of VvGT15a, and VvGT16. The corresponding VvGT14ao variants in pGEX-K in E. coli BL21(DE3) were used as reference. The results were obtained with the HP cultivation system.*

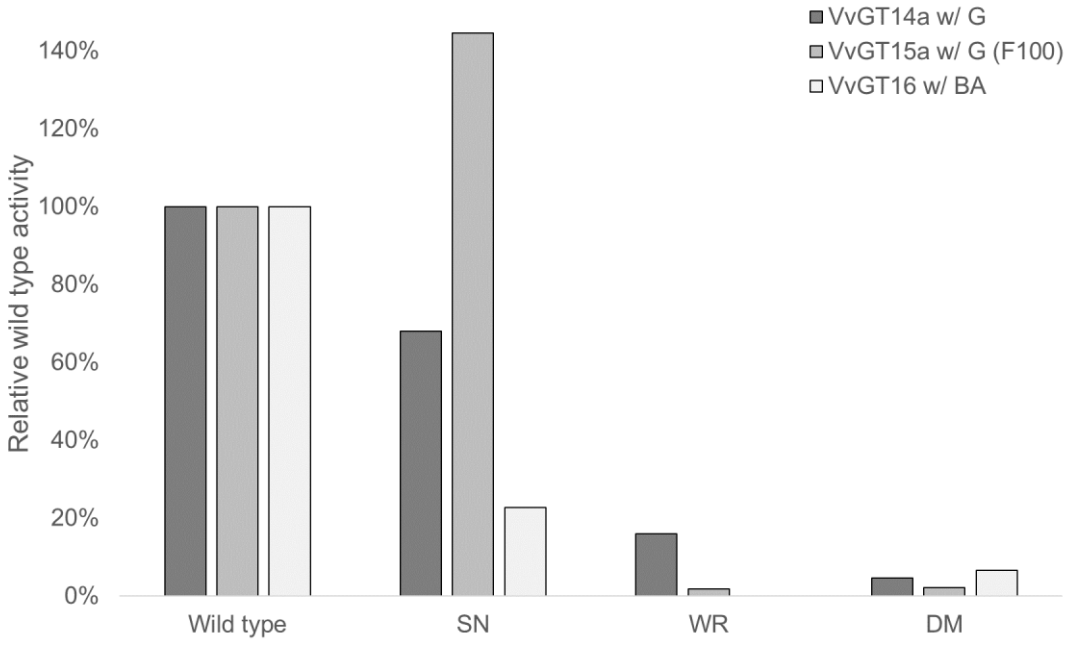


*Figure 38: Activity of K163-derived mutants of VvGT14ao, VvGT15a, and VvGT16 in-vitro*

*The UDP-Glo assay was used to determine the glucosylation activity of K163-derived variants from VvGT14ao, VvGT15a, and VvGT16. VvGT14ao, and VvGT15a variants were tested with geraniol (G), VvGT16 variants were tested with benzyl alcohol (BA) as substrate. The reactions were performed at 30 °C for 90 min with 100 µM of UDP-glucose and 100 µM of substrate.*

Similar to the *in-vivo* data (Figure 37), the activity of the tryptophan-arginine (WR) variants and double mutants was lower than that of the wild type enzymes *in-vitro* (Figure 38). The activity

of the serine-asparagine (SN) variants did not seem to vary significantly from the activity of the wild type enzymes. The enzyme variants of VvGT15a and VvGT16 were not further investigated. Since the K163-derived VvGT14ao variants showed elevated product concentrations *in-vivo* but not *in-vitro*, the protein levels of the glucosyltransferase variants were investigated.

The VvGT14ao variants were analyzed by SDS gel and western blot (Figure 39). The western blot clearly showed intensive bands for the GT variants VvGT14ao-W353R (WR) and VvGT14ao-S168N-W353R (DM) compared to the wild type enzyme (WT; Figure 39), indicating a higher abundance of these enzyme variants in the production host. Thus, the mutations promote the production of active enzymes in the cells.

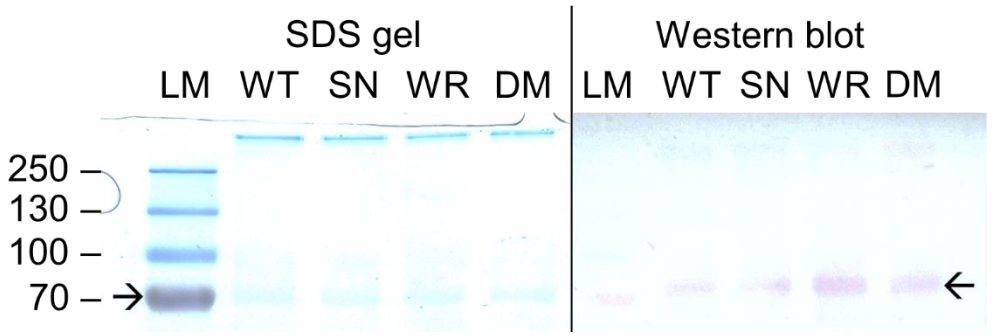


*Figure 39: SDS gel and western blot analysis of the K163-derived VvGT14ao variants*

*The K163-derived variants were expressed in E. coli BL21(DE3) using the pGEX-K vector. Same cell concentrations were used for the preparation of the SDS samples. The arrows indicate the height of the GST-tagged VvGT14ao variants inside the SDS gel (left)/ on the western blot (right).*

## 3.9 Combination of positive effects of K43 and K163

The screening hits, K43 and K163, showed reproducible, elevated glucoside concentrations. Their positive effects were caused by different modes of action and were transferable to other production strains (Figure 31 and 34). Therefore, the cured K43 strain was combined with expression plasmids coding for the double mutant in order to further increase the product yield. The new strains were tested against the corresponding original strains.

*E. coli* BL21(DE3) and *E. coli* BL21(DE3)K43 were transformed with three different expression vectors, pRSFDuet-1, pET29a, and pGEX-K, coding either for VvGT14ao or for VvGT14ao-S168N-W353R (Figure 40). The *E. coli* BL21(DE3)K43 strains carrying pRSFDuet-1 and pET29a vectors performed better than the corresponding *E. coli* BL21(DE3) strains expressing the original VvGT14ao (Figure 40). The positive effect of the K43 strain could not be seen with the pGEX-K vectors. The pGEX-K vectors showed reduced product concentrations in the K43 strain when compared to the corresponding *E. coli* BL21(DE3) strain. Each strain expressing the double mutant *VvGT14ao-S168N-W353R* performed better than the corresponding strains expressing *VvGT14ao*. The only strain that benefited from both variations, the genomic

background of the K43 strain and the double mutant from K163, was *E. coli* BL21(DE3)K43 /pET29a_VvGT14ao-S168N-W353R. This strain produced 0.656 mM geranyl glucoside. However, *E. coli* BL21(DE3) /pRSFDuet-1_VvGT14ao-S168N-W353R showed the highest concentration of 0.666 mM (210.5 mg/l) geranyl glucoside. Compared to the concentrations of 0.010 and 0.037 mM of geranyl glucoside obtained with the initial strains, a 67- and 18-fold increase, respectively in product titer was achieved.

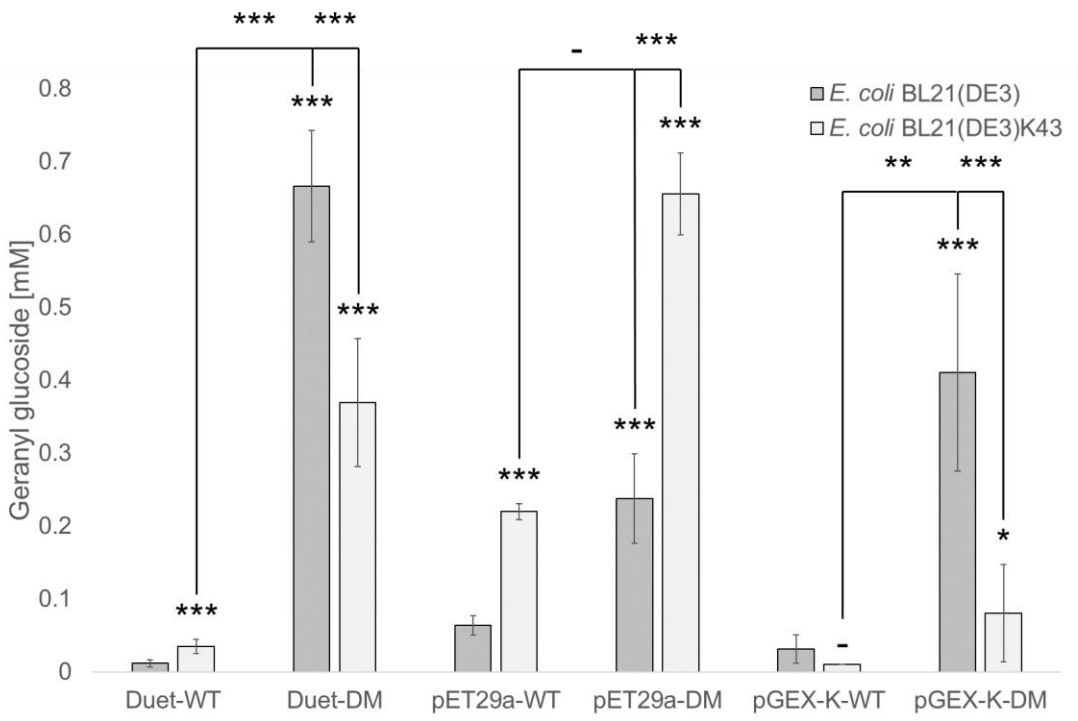


*Figure 40: Geranyl glucoside formation of production strains derived from K43 and K163*

*E. coli BL21(DE3)K43 was transformed with the double mutant, VvGT14ao-S168N-W353R, contained in different expression vectors. The vectors pRSFDuet-1 (Duet), pET29a, and pGEX-K were used. The K43 strains were compared to the corresponding E. coli BL21(DE3) strains, containing also the double mutant (DM) or the original (WT) enzyme in the appropriate vector. The results were obtained with the HP cultivation system. The statistical analysis was done with an unpaired t test. - not statistically significant, * p<0.05, ** p<0.01, and *** p<0.001. The indication directly above the bars depicts the level of significance between the respective sample and the E. coli BL21(DE3) strain expressing VvGT14ao with the same expression vector backbone.*

# 4 Discussion

The aim of the study was to improve a whole-cell biocatalyst for the production of geranyl glucoside. The strains *E. coli* BL21(DE3)pLysS /pET29a_VvGT14ao and *E. coli* BL21(DE3)pLysS /pGEX-4T-1_VvGT14ao were available at the project start and were used as references. The strains produced 0.010 mM and 0.037 mM of geranyl glucoside (Figure 18) and displayed a specific strain activity of 8.6 µU/mg and 31.6 mU/mg, respectively (data not shown).

## 4.1 Alternative production host: *E. coli* C43(DE3)

The selection of alternative expression hosts was limited because many common expression vectors use the T7 expression system. The initial production host *E. coli* BL21(DE3)pLysS contained an additional plasmid pLysS encoding T7 lysozyme. The T7 lysozyme inhibits basal expression and is required for the expression of toxic proteins. Strictly regulated expression is not required for this application as the target enzyme VvGT14a does not pose a risk to *E. coli*. Therefore, the production host *E. coli* BL21(DE3) without pLysS was also tested in this study.

The *E. coli* strain C43(DE3) can overexpress difficult and toxic proteins and membrane proteins like the subunit b of the bacterial F-ATPase (Miroux and Walker 1996). The strain was derived from *E. coli* BL21(DE3) and can therefore utilize the T7 expression system. The geraniol tolerance of *E. coli* C43(DE3) was tested in a concentration range around and above the reported tolerance level of *E. coli* BL21(DE3). However, *E. coli* C43(DE3) did not tolerate higher geraniol concentrations than *E. coli* BL21(DE3) or even saturated aqueous conditions (approximately 4.45 mM of geraniol; Figure 8). Therefore, *E. coli* strain C43(DE3) was no longer considered as a production strain.

The pGEX vectors used in this study use the tac promoter instead of the T7 promoter and can therefore be used by virtually any *E. coli* strain. Therefore, the strains *E. coli* BL21 and *E. coli* Waksman were used in this study together with pGEX-4T-1 and pGEX-K vectors. *E. coli* Waksman did not perform well under the conditions applied in this study and showed a maximum of 0.017 mM glucoside (Figure 18). In contrast, *E. coli* BL21 produced 0.042 mM with pGEX-4T-1 and 0.064 mM with pGEX-K as expression vector (Figure 18).

## 4.2 Molecular chaperones

The overexpression of proteins in *E. coli* oftentimes leads to the formation of misfolded proteins, and inactive protein aggregates called inclusion bodies. Molecular chaperones are supposed to promote proper folding, and prevent the formation of those protein aggregates. In order to improve protein production in the whole-cell biocatalyst, the production strain *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao was combined with a set of plasmids expressing

chaperones separately. The comparison with the original strain showed that only pGro7, allowed the strain to produce higher geranyl glucoside concentrations. pGro7 codes for GroEL and GroES. GroEL builds up a tetradecamer consisting of two rings. This barrel-shaped protein complex allows the binding of misfolded proteins at the rim. The binding of ATP and a GroES heptamer induces a conformational change that transfers the bound protein to the inner chamber and promotes the proper folding. The hydrolysis of ATP causes the protein to be released into the cytosol again. Misfolded VvGT14a could be transformed into an active form by this chaperone complex. pGro7 was also used to increase ribonuclease inhibitor activity from roughly 100 to 700 kU/$g_{CWW}$ (Siurkus and Neubauer 2011).

The positive effect from pGro7 was investigated further at different temperatures. The use of pGro7 allowed *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao to produce similar amounts of geranyl glucoside independent of the biotransformation temperature (Figure 10). Low temperatures such as 18 °C were used for glucoside production in batch cultivation since the production strains grow faster at higher temperatures e.g. 37 °C (Noor et al. 2013) and thereby deplete the available resources faster. This makes it necessary to supplement the culture with nutrients and to adjust the pH value of the culture medium more frequently. The stable product yield at lower temperatures could also be transferred to other biotransformation systems.

The inherent chaperone production in *E. coli* can also be elicited through chemical stresses e.g. by adding 3 % (v/v) of ethanol to the culture medium (Thomas and Baneyx 1996). The addition of ethanol to production strains showed promising results but the results of the biological replicates displayed too strong variations (Figure 9). Therefore, no further experiments with the chemical elicitation were performed.

## 4.3 Engineering of the HP cultivation system

In order to test several variables in parallel during the biotransformation experiments, a 24 well plate (hitplate 25) system was employed. Biotransformation tests in hitplates with *E. coli* BL21(DE3)pLysS /pGEX-K_VvGT14ao confirmed the production of glucosides (Figure 13). In order to further improve this system, the cell growth and protein production step was also transferred to the hitplate cultivation system. This allowed the testing of different production strains, culture media, and substrates or substrate concentrations in parallel. The HP cultivation system can also be transferred to other biotransformation applications in the future.

The use of auto-induction media was not successful since no activity with TB-AIM and only minor activities with MAI medium could be detected. The sequencing of the laboratory strain genome showed premature stop codon within *lacY*, the gene coding for β-galactoside permease. Therefore, lactose may not be able to enter the cells and induce protein expression

when using auto-induction media. Hence, the HP cultivations were performed in $M9^+$ medium with IPTG induction. The inoculation of the main cultures was done with 2.5 % (v/v) of pre-culture, which is a relatively high concentration. This high concentration was chosen so that all cultures achieved an OD of 0.4 to 0.6 in a short time. Most strains reached a high optical density of 2.5 to 3.0 after the first overnight cultivation. All strains were supplemented with new nutrients at the start of the biotransformation. Therefore, all biotransformations were performed with similar amounts of induced cells and identical amounts of nutrients.

### 4.3.1 Reproducibility of the HP cultivation system

The data obtained from three separate experiments proved the reproducibility of the HP cultivation system (Figure 17). The combination of more data sets further increased the method's precision (Figure 18) and thereby enabled the detection of smaller differences.

In addition to reproducibility, the culture system was also tested for its analytical variance by simultaneously analyzing the furaneoyl glucoside production capabilities of *E. coli* BL21(DE3)pLysS /pGEX-K_VvGT14ao in twelve wells. The reproducibility of the twelve biological replicates was acceptable and the variation of the UV data was very low (Figure 14). Therefore, UV detection was used throughout the study for the analysis of glucoside products.

### 4.3.2 Practicality of the specific activity

Calibration curves performed with authentic material were used to calculate molar concentrations of the products and the optical densities of the biotransformation culture were determined to calculate the specific strain activity. Since the testing system used overnight cultivations during the protein production and the biotransformation, almost all cultures reached the plateau phase and displayed similar optical densities at the sampling time. It was observed that the production cultures continuously lost their glucosylation activities over time, most likely due to depleting glucose and nutrient concentrations (Figure 15).

It is currently not possible to track the glycosylation of terpenols during the whole-cell biotransformation online. Therefore, it is only possible to calculate the initial activity by extrapolation of the course of activity over time. The decrease in activity of *E. coli* BL21(DE3) /pET29a_VvGT14ao was linear with a coefficient of determination of 0.99 (Figure 15). The starting activity could be taken directly from the linear regression model and was 107 µM at the beginning of the biotransformation. The calculated specific activity was 34.5 µU/mg.

In an effort to further improve the strain testing, the different strains were adjusted to an optical density of 0.6 prior to induction. This measure improved catalytic activity as *E. coli* BL21(DE3) /pET29a_VvGT14ao showed a specific activity of 98.8 µU/mg (Figure 15) instead of 32.0 µU/mg in a previous experiment. However, the effort required to adjust all test strains in parallel at an OD of 0.6 does not justify the benefits. The main purpose of the HP cultivation

system was to identify differences between production strains in terms of product formation. The system without adjustment of the cell density could also determine these differences.

## 4.4 Testing of novel strain-vector combinations

The 18 created strain-vector combinations were tested with the novel HP cultivation system. Even though, there was no dominant substrate preference throughout the samples (Figure 17), there were some minor, observable trends. When looking at the different hosts, some preferences for substrate use emerged. Most *E. coli* BL21(DE3) vector combinations produced higher quantities of eugenyl glucoside, except for strains with a pGEX vector. Whole-cell biocatalysts based on *E. coli* BL21(DE3)pLysS produced higher quantities of citronellyl glucoside, except for strains with pRSFDuet-1_VvGT14ao and pET32a_VvGT14ao. The preference of *E. coli* BL21(DE3)pLysS might be due to the chloramphenicol acetyltransferase (CAT) encoded on pLysS. CAT showed acetylation activity towards geraniol (Figure 21) and it is possible that CAT also acetylates other substrates used in this study. The activity of CAT towards citronellol may be lower than towards the other substrates. This would affect the substrate concentration in favor of citronellol during the biotransformation. *E. coli* Waksman and *E. coli* BL21 seemed to prefer nerol and citronellol as a substrate, respectively, however only two expression vectors were tested with these two host organisms.

When looking at different expression vectors, production strains with pGEX vector tended to produce slightly higher amounts of citronellyl glucoside (Figure 17). However both *E. coli* Waksman strains, and *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao did not follow this trend. On the other hand, strains containing the pET32 vector appeared to produce higher quantities of eugenyl glucoside, except for *E. coli* BL21(DE3)pLysS /pET32aK_VvGT14ao which preferred citronellol as a substrate like most of the other *E. coli* BL21(DE3)pLysS-based strains. *E. coli* BL21(DE3) /pGEX-4T-1_VvGT14ao was the only strain which produced more geranyl glucoside than any other glucoside. However, this strain showed the lowest product concentrations. Most of the observed preferences are not statistically significant due to high and overlapping variations in product concentrations and should not be overestimated.

The combined data (Figure 18) reveals two main results. The *E. coli* BL21 host performed better with the pGEX vectors than *E. coli* BL21(DE3), and *E. coli* BL21(DE3) produced higher glucoside concentrations with expression vectors containing a kanamycin resistance instead of an ampicillin resistance cassette. It is understandable that the pGEX vectors in *E. coli* BL21 worked better than in *E. coli* BL21(DE3) because the pGEX vectors do not require the T7 RNA polymerase encoded by the (DE3) insertion. However, the pGEX vectors in *E. coli* BL21(DE3)pLysS also performed better than in *E. coli* BL21(DE3). This was surprising, because *E. coli* BL21(DE3)pLysS carries an additional plasmid, which expresses T7 lysozyme. Since the T7 RNA polymerase is not used by the pGEX vectors, neither the (DE3) insertion

nor pLysS should directly affect the expression of the glucosyltransferase. Furthermore, it would be reasonable to assume, that these extra functionalities would divert production means needed for the overexpression of *VvGT14ao*. Surprisingly, *E. coli* BL21(DE3)pLysS /pGEX-4T-1_VvGT14ao produced higher glucoside titers than *E. coli* BL21 /pGEX-4T-1_VvGT14ao. Even though, *E. coli* BL21(DE3) worked better with plasmids using the kanamycin resistance cassette, *E. coli* BL21(DE3) with pLysS performed better with expression vectors using an ampicillin resistance cassette.

## 4.5 Impact of the inducer concentration

The induction test with pET29a_VvGT14ao, and pRSFDuet-1_VvGT14ao (Figure 19) showed that pRSFDuet-1 is very leaky. The inducer concentration did not have any influence on the final product concentration. There are many factors that might explain the leakiness of pRSFDuet-1. The pET29a and pRSFDuet-1 vectors do not differ in the sequence of the T7 promoter and the lac operator, so their transcription should be induced in the same way (Dubendorf and Studier 1991). Our sequencing results identified two SNPs upstream of the ribosomal binding site (RBS; Shine-Dalgarno sequence). In addition, another SNP was found downstream of the RBS in the spacer region and the spacer region of pRSFDuet-1 is shortened by one nucleotide. The closer SNP located upstream of the RBS core sequence is 14 and 15 nucleotides from the start codon in pRSFDuet-1 and pET29a, respectively. This SNP might influence the binding affinity of the ribosome, which in turn might affect translation. The two variations in the spacer region before the start codon could influence the translation initiation and thus also the translation. A major difference between pET29a and pRSFDuet-1 is their origin of replication and as a result their plasmid copy number per cell. The pET29a plasmid uses the pBR322 ori, which results in approximately 20 plasmid copies per cell while pRSFDuet-1 uses an RSF1030-derived ori instead, which results in over 100 plasmid copies per cell (Bagdasarian et al. 1981, Sørensen and Mortensen 2005). Both vectors express the *lacI* gene coding for the lac repressor. The lac repressor should be present in proportional amounts to the plasmid copy number (Dubendorf and Studier 1991). Nevertheless, *E. coli* BL21(DE3) /pRSFDuet-1_VvGT14ao produced the target enzyme prior to the induction. Consequently, some or all of these factors led to the extremely leaky and non-inducible expression of *VvGT14ao* from pRSFDuet-1.

High-copy vectors are reportedly leaky in their expression (Gruber et al. 2008). However, it could be shown that the protein production with pRSFDuet-1 as an expression vector and *E. coli* BL21(DE3) as an expression host does not produce the target protein without induction (Zou et al. 2014). On the other hand, some researchers also showed little effect of the inducer concentration on gene expression from pRSFDuet-1, and comparable activities even without induction (Chen et al. 2019). Neither pRSFDuet-1_VvGT14ao nor pET29a_VvGT14ao in

*E. coli* BL21(DE3) showed negative effects at higher IPTG concentrations. Therefore, a concentration of 1.25 mM IPTG was used for all experiments without individual investigation of test strains.

## 4.6 Directed evolution

A directed evolution approach was employed to improve the catalytic activity of the target enzyme, VvGT14a. Six mutant libraries of *VvGT14ao* were created for this purpose. The optimal screening library should contain only mutants with one amino acid exchange per enzyme. Therefore, library number 4 was chosen for the screening because of the narrow mutation range and the high abundance of insert in the tested transformants (Figure 22). Two screening methods were used to find improved mutants within the library, a coupled enzyme assay and the HP cultivation screening.

The two methods used for the screening have their advantages but also shortcomings. The coupled enzyme assay can be set up to run in a high-throughput format. The measuring of 94 putative mutants took 150 min, but the measuring time could be further reduced to allow more measurements per day. The assay could also be used with an end point determination, since a steeper slope resulted in a higher end value. However, the coupled enzyme assay displayed very high variations due to numerous working steps. In comparison, the HP cultivation screening produced highly reproducible results with rather low variations. However, the HP cultivation screening was hardly high-throughput since only eleven strains could be tested on one plate and the HPLC measurement took up to 8 hrs.

Both screenings revealed that only about 75 % of the tested strains seemed to carry an insert, even though the sequence analysis of the mutant library showed up to 100 % of strains carrying an insert (Figure 22). This discrepancy can be explained by the different strains used for the sequencing analysis and the screening itself. The sequencing analysis of the library was done with *E. coli* NEB10β. This strain does not carry the gene for the T7 RNA polymerase and cannot express the glucosyltransferase gene from pRSFDuet-1, which uses the T7 promoter. Therefore, the difference in colony size of empty vector strains and strains carrying a GT gene is not really noticeable. However, *E. coli* BL21(DE3) was used for the screening, which is able to express the target gene. Since pRSFDuet-1 is a high-copy vector and its expression is leaky (Figure 19), the growth of *E. coli* BL21(DE3) strains carrying a pRSFDuet-1_VvGT14ao variant was impaired. Since the colonies for the screening were picked by hand, it is possible, that disproportionally more empty vector strains were selected for the screening through colony size discrimination.

### 4.6.1 Engineering of the coupled enzyme assay

The coupled enzyme assay (Figure 23 A) indirectly determined the activity of VvGT14a and its variants using a photometer. The reaction conditions had to be optimized and many reaction parameters e.g. buffer composition and reaction temperature were adopted from Bönisch et al. (2014). The concentrations of OleD, geraniol, and ClPNP-glucose were determined experimentally (Figure 24).

The geraniol concentration for the assay was varied between 25 µM and 4.45 mM to find the optimal substrate concentration. VvGT14ao should catalyze the reaction under substrate saturation condition to find mutants with higher maximum reaction rates and not with reduced $K_M$ values. The $K_M$ value of VvGT14a for geraniol is relatively low with only 9 µM (Bönisch et al. 2014). When testing high geraniol concentrations with purified enzyme and water control, it became clear that high geraniol concentrations increase the absorption signal even without VvGT14a (Figure 25 A). This might be due to a side activity of OleD towards geraniol. Therefore, 0.5 mM geraniol was used for the final control.

The coupled enzyme assay was used to produce vancomycin pseudoaglycone glycosides with an average yield of 36 %. The concentrations and conditions used in that study vary in almost every aspect. The VvGT14a assay required OleD with a concentration of 0.05 µM, whereas the study needed 10.5 µM of each enzyme (Gantt and Thorson 2012). VvGT14a has a rather low $K_M$ value for UDP-glucose with only 16 µM (Bönisch et al. 2014). Therefore, little as 100 µM UDP were used in the assay, while 1 mM of UDP was used for the vancomycin glycoside production. However, the ClPNP-glucose concentration had to be increased to 2.5 mM in order to improve upon the resolution of the VvGT14a assay, the other assay only used 1 mM ClPNP-glucose (Gantt and Thorson 2012). Despite the low $K_M$ value of 9 µM of VvGT14a towards geraniol (Bönisch et al. 2014), a concentration of 0.5 mM was determined to be optimal. The glycosylation of vancomycin pseudoaglycones only required 0.1 mM of the aglycones. Both assays were performed at 30 °C, however the VvGT14a assay was reduced to a reaction time of 2.5 h instead of 24 h (Gantt and Thorson 2012). Prolonged reaction times might improve upon the resolution of the assay, however short reaction times are needed for a high-throughput assay.

The developed screening method showed an absorption difference of approx. 5 mAU/min between the proteins, isolated from *E. coli* expressing the empty vector and those expressing VvGT14a. Even though the system does not seem to run linearly in terms of enzyme activity, the difference between crude extracts of empty vector strain and VvGT14ao strain was sufficient to identify putative positive mutants in the mutant screening.

### 4.6.2 Screening using the coupled enzyme assay

The screening with the coupled enzyme assay yielded no mutants exceeding the 2-times wild type activity (Figure 27 A). The expression vectors of putative positive mutants which exceeded 130 % wild type activity were sequenced. Some strains however revealed the wild type *VvGT14ao* gene and were not further investigated. The identified mutants were tested against the original VvGT14ao strain in HP cultivation. However no mutant showed an increased geranyl glucoside concentration and one mutant did not show any product formation. This mutant was the VvGT14ao-M1_P241del-V431M, which is lacking the first 241 amino acids.

The cloning of the mutant library was done with *Bam*HI and *Not*I. The epPCR completed an internal *Bam*HI recognition site consisting of 5 out of 6 base pairs that were already necessary for the *Bam*HI restriction. The shortened gene was cloned in-frame, allowing translation of the remaining 233 amino acid long C-terminus. Without the N-terminus, the shortened enzyme should not be able to bind an aglycone, the C-terminus however could still bind UDP-glucose.

The coupled enzyme assay showed an increase in absorbance, when UDP is released, not necessarily when geranyl glucoside was formed. Therefore, mutants with increased UDP-glucose hydrolase activity also show a steeper slope and cannot be distinguished by this assay from real screening hits that actually produce higher yields of geranyl glucoside.

The shortened enzyme, VvGT14ao-M1_P241del-V431M, was purified by immobilized metal affinity chromatography and the UDP-glucose hydrolase activity was confirmed by the UDP-Glo assay performed without a sugar acceptor present (Figure 28). The release of UDP from UDP-glucose indicates that the truncated enzyme indeed can bind UDP-glucose and furthermore can catalyze a reaction. Although, the reaction with UDP-glucose is now done with a water molecule instead of a sugar acceptor. However, there must be a basic amino acid nearby which can take up a proton from the water molecule. This is necessary to obtain a nucleophile, which is able to perform the nucleophilic attack on the C1 carbon of glucose and thereby split the glucose from UDP. The hydrolase activity towards NDP-sugars was also reported for other plant glycosyltransferases (Shao et al. 2005).

In the end, no real screening hits could be determined with the coupled enzyme assay. The obtained putative screening hits turned out to produce similar or less geranyl glucoside than the wild type strain. Normally $10^3$ to $10^6$ colonies of a high quality mutation library have to be tested in order to obtain a satisfying mutant (Turner 2009). Considering the rather small number of screened colonies, with only 846 putative mutations tested, it is not surprising that no superior mutants were obtained with this screening.

### 4.6.3 Screening using the HP cultivation system

The second screening of the mutant library used HPLC to directly detect product concentrations. This screening yielded three strains, which showed increased geranyl glucoside formation. A number of three screening hits from only 170 screened putative mutants appears to be unusually high. However, these screening hits do not code for more active enzymes, the increased product yields are due to other effects. The strains, K60, K43, and K163, were investigated in order to find the source of the improved product formation.

#### *4.6.3.1 The K60 strain*

The strain K60 produced approximately 3-times more glucoside than the original VvGT14ao strain. The strain was analyzed by cPCR (Figure 30) and multiple sequencing reactions. K60 contained at least the empty vector, a variant vector coding for VvGT14ao-T398S, and the wild type vector. It is highly likely that the boost in productivity is due to an expression effect caused by the lower gene dose or slower translation (Rosano and Ceccarelli 2009). The codon in the DNA sequence was changed from ACG to TCG. The tRNA responding to UCG in the mRNA is slightly less abundant than the tRNA responding to ACG (Kane 1995, Wada et al. 1992). In addition, it would also be possible, although unlikely, that the GT variant (VvGT14ao-T398S) is more active than the wild type. The mutation is located only two amino acids downstream of the PSPG box and would have to be about 9 times as active as the wild type enzyme to account for the lower proportion of the expression vector, and the higher product yield.

#### *4.6.3.2 The K43 strain*

The strain K43 was determined to carry the wild type expression vector. The curing and combination with different expression vectors showed that a mutation in the genomic background must be responsible for the increased product yield. The genomes of K43 and the wild type laboratory strain were sequenced. The genome of K43 showed a number of mutations in comparison to the laboratory strain (Figure 32). However, only one gene coding for a protein was changed, *yghJ*. This gene was already heavily mutated in the laboratory strain, nevertheless both genes still code for a full-length protein. YghJ is a secreted, cell surface associated lipoprotein and contains a M60 metalloprotease domain (Tapader et al. 2017). This zinc metalloprotease domain forms amyloid-like fibrils and is associated with intestinal mucin degradation and proinflammatory responses (Belousov et al. 2018). Therefore, yghJ plays an important role in the virulence of intestinal pathogenic *E. coli* strains (Tapader et al. 2017). However, no causal connection to the geranyl glucoside production could be made.

The other mutation loci are 82 or more base pairs upstream of the next genes coding for proteins. Although, this is far from the translation start, these mutations still could have an influence on the transcription (Mendoza-Vargas et al. 2009).

#### *4.6.3.3 The K163 strain*

The strain K163 carried a variant coding for a double mutant, VvGT14ao-S168N-W353R. Single mutants were generated from the double mutant to analyze which mutation is responsible for the increase in product yield. The K163-derived mutants were tested in pET29a, pGEX-K, and pRSFDuet-1. The double mutant and the W353R variant consistently produced more glucoside *in-vivo* than the wild type strain. The S168N variant only showed an advantage in pRSFDuet-1 (Figure 34). The double mutant was the most productive variant in the expression vectors pET29a, and pRSFDuet-1. The W353R variant was the most active one in the pGEX-K vector. However, the mutations were not transferable to other glycosyltransferases (Figure 37).

The *in-vitro* analysis of the VvGT14ao variants showed that the reaction speed at low substrate concentrations was drastically reduced, especially for the W353R variant, and the double variant. The wild type VvGT14ao showed a maximum specific activity of 2100 U/mg and a $K_M$ value of 72 µM geraniol. The $K_M$ values of the single mutants, S168N and W353R, were not strongly affected, their maximum specific activity however was drastically reduced to 581 and 58.2 U/mg, respectively. The double mutant was strongly affected, the $K_M$ value (230 µM) and the maximum specific activity (219 U/mg) were both impaired compared to the wild type enzyme (Table 48). This was surprising, since the *in-vivo* data clearly showed that the mutants produced higher product titers. However, there is a decrease in reaction speed visible at 75 µM geraniol for the wild type enzyme and at 150 µM geraniol for the S168N variant and the S168N-W353R variant (Figure 35). The reduced activity at higher substrate concentrations showed that the GTs were inhibited by the substrate. The *in-vitro* data displayed the activity of the variants up to 200 µM geraniol. However, a concentration of 1.155 mM geraniol was used in the biotransformation experiments. Since $K_M$ and $v_{max}$ were influenced by the mutations, it is also possible that the $K_i$ value of the substrate inhibition was altered as well, possibly in favor of the double mutation. So, the activities of the variants might approach the activity of the wild type enzyme at higher geraniol concentrations. This possibility together with the increased enzyme concentration could account for the higher geranyl glucoside yield of the K163-derived variants *in-vivo*.

The higher product levels were assumed to be due to increased amounts of active enzyme. Therefore, protein expression of the variants was studied in *E. coli* BL21(DE3) using the pGEX-K vector. The enzyme variants W353R, and S168N-W353R seemed to be produced in higher amounts (Figure 39). The DNA sequence of *VvGT14ao* was changed from AGC to AAC for the S168N mutation and from TGG to CGG for the W353R mutation. The tRNA responding to AAC in the mRNA is slightly more abundant than the tRNA required for the wild type. However, the tRNA responding to CGG is a lot less abundant than the tRNA responding to UGG in the

wild type (Wada et al. 1992). The translation might be slowed down by the W353R mutation and might result in a higher proportion of correctly formed enzyme (Kane 1995).

#### *4.6.3.4 Combining K43 and K163*

The attempt to combine the positive effects from K43 and K163 to further increase the geranyl glucoside yield resulted in a production strain producing 60 and 12 times as much product than the initial reference strains. Overall, the product concentrations and the specific strain activities were boosted from 0.011 mM and 8.6 µU/mg of *E. coli* BL21(DE3)pLysS /pET29a_VvGT14ao and from 0.054 mM and 31.6 µU/mg of *E. coli* BL21(DE3)pLysS /pGEX-4T-1_VvGT14ao to 0.656 mM and approximately 380 µU/mg of *E. coli* BL21(DE3)K43 /pET29a_VvGT14ao-S168N-W353R.

# 5 Bibliography


Abdallah II, Quax WJ. 2017. A Glimpse into the Biosynthesis of Terpenoids. KnE Life Sciences 3:81-98.

Amrein KE, Takacs B, Stieger M, Molnos J, Flint NA, Burn P. 1995. Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. Proceedings of the National Academy of Sciences 92:1048-1052.

Arnold FH. 2018. Directed evolution: bringing new chemistry to life. Angewandte Chemie International Edition 57:4143-4148.

Arnold FH, Volkov AA. 1999. Directed evolution of biocatalysts. Current opinion in chemical biology 3:54-59.

Ashraf R, Muhammad MA, Rashid N, Akhtar M. 2017. Cloning and characterization of thermostable GroEL/GroES homologues from *Geobacillus thermopakistaniensis* and their applications in protein folding. Journal of biotechnology 254:9-16.

Bagdasarian M, Lurz R, Rückert B, Franklin F, Bagdasarian M, Frey J, Timmis K. 1981. Specific-purpose plasmid cloning vectors II. Broad host range, high copy number, RSF 1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. Gene 16:237-247.

Banthorpe DV, Le Patourel GN, Francis MJ. 1972. Biosynthesis of geraniol and nerol and their β-d-glucosides in *Perlargonium graveolens* and *Rosa dilecta*. Biochemical Journal 130:1045-1054.

Belousov MV, Bondarev SA, Kosolapova AO, Antonets KS, Sulatskaya AI, Sulatsky MI, Zhouravleva GA, Kuznetsova IM, Turoverov KK, Nizhnikov AA. 2018. M60-like metalloprotease domain of the *Escherichia coli* YghJ protein forms amyloid fibrils. PLoS One 13:e0191317.

Bönisch F, Frotscher J, Stanitzek S, Ruhl E, Wust M, Bitz O, Schwab W. 2014. Activity-based profiling of a physiologic aglycone library reveals sugar acceptor promiscuity of family 1 UDP-glucosyltransferases from grape. Journal of Plant Physiology 166:23-39.

Boros I, Pósfai G, Venetianer P. 1984. High-copy-number derivatives of the plasmid cloning vector pBR322. Gene 30:257-260.

Brazier-Hicks M, Edwards R. 2013. Metabolic engineering of the flavone-C-glycoside pathway using polyprotein technology. Metabolic engineering 16:11-20.

Brazier-Hicks M, Gershater M, Dixon D, Edwards R. 2018. Substrate specificity and safener inducibility of the plant UDP-glucose-dependent family 1 glycosyltransferase super-family. Plant Biotechnology Journal 16:337-348.

Brewster RC, Jones DL, Phillips R. 2012. Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli*. PLOS Computational Biology 8:e1002811.

Brochado AR, Matos C, Møller BL, Hansen J, Mortensen UH, Patil KR. 2010. Improved vanillin production in baker's yeast through in silico design. Microbial cell factories 9:84.

Carvalho IT, Estevinho BN, Santos L. 2016. Application of microencapsulated essential oils in cosmetic and personal healthcare products–a review. International journal of cosmetic science 38:109-119.

Chakravartty V, Cronan JE. 2015. A series of medium and high copy number arabinose-inducible *Escherichia coli* expression vectors compatible with pBR322 and pACYC184. Plasmid 81:21-26.

Chen W, Viljoen A. 2010. Geraniol—a review of a commercially important fragrance material. South African Journal of Botany 76:643-651.

Chen Y, Ma B, Cao S, Wu X, Xu Y. 2019. Efficient synthesis of Ibrutinib chiral intermediate in high space-time yield by recombinant *E. coli* co-expressing alcohol dehydrogenase and glucose dehydrogenase. RSC Advances 9:2325-2331.

Costa S, Almeida A, Castro A, Domingues L. 2014. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. Frontiers in microbiology 5:63.

Das R, Mukhopadhyay B. 2016. Chemical O-Glycosylations: An Overview. ChemistryOpen 5:401-433.
De Boer HA, Hui AS. 1990. [9] Sequences within ribosome binding site affecting messenger RNA translatability and method to direct ribosomes to single messenger RNA species. Pages 103-114. Methods in Enzymology, vol. 185 Elsevier.
De Bruyn F, De Paepe B, Maertens J, Beauprez J, De Cocker P, Mincke S, Stevens C, De Mey M. 2015a. Development of an *in vivo* glucosylation platform by coupling production to growth: Production of phenolic glucosides by a glycosyltransferase of *Vitis vinifera*. Biotechnology and bioengineering 112:1594-1603.
De Bruyn F, Maertens J, Beauprez J, Soetaert W, De Mey M. 2015b. Biotechnological advances in UDP-sugar based glycosylation of small molecules. Biotechnology Advances 33:288-302.
Demchenko AV. 2008. Handbook of chemical glycosylation: advances in stereoselectivity and therapeutic relevance. John Wiley & Sons.
Desmet T, Soetaert W, Bojarova P, Kren V, Dijkhuizen L, Eastwick-Field V, Schiller A. 2012. Enzymatic glycosylation of small molecules: challenging substrates require tailored catalysts. Chemistry 18:10786-10801.
Dubendorf JW, Studier FW. 1991. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. Journal of Molecular Biology 219:45-59.
Francis M, Allcock C. 1969. Geraniol β-D-glucoside; occurrence and synthesis in rose flowers. Phytochemistry 8:1339-1347.
Furia TE. 1972. CRC handbook of food additives. CRC Press.
Gantt RW, Peltier-Pain P, Singh S, Zhou M, Thorson JS. 2013. Broadening the scope of glycosyltransferase-catalyzed sugar nucleotide synthesis. Proceedings of the National Academy of Sciences 110:7648-7653.
Gantt RW, Thorson JS. 2012. High-throughput colorimetric assays for nucleotide sugar formation and glycosyl transfer. Methods in Enzymology 516:345-360.
Glazyrina J, Materne E-M, Dreher T, Storm D, Junne S, Adams T, Greller G, Neubauer P. 2010. High cell density cultivation and recombinant protein production with *Escherichia coli* in a rocking-motion-type bioreactor. Microbial cell factories 9:42.
Goettig P. 2016. Effects of Glycosylation on the Enzymatic Activity and Mechanisms of Proteases. International Journal of Molecular Sciences 17.
Goldmann. 2018. Flavor on Demand - Adding superior aroma effects to your products.: S. Goldmann GmbH & Co. KG.
Gonzales MF, Brooks T, Pukatzki SU, Provenzano D. 2013. Rapid protocol for preparation of electrocompetent *Escherichia coli* and *Vibrio cholerae*. Journal of visualized experiments: JoVE.
Grabski A, Mehler M, Drott D. 2005. The overnight express autoinduction system: high-density cell growth and protein expression while you sleep. Nature methods 2:233.
Gruber DF, Pieribone VA, Porton B, Kao H-T. 2008. Strict regulation of gene expression from a high-copy plasmid utilizing a dual vector system. Protein expression and purification 60:53-57.
Guo W, Sakata K, Watanabe N, Nakajima R, Yagi A, Ina K, Luo S. 1993. Geranyl 6-O-β-D-xylopyranosyl-β-D-glucopyranoside isolated as an aroma precursor from tea leaves for oolong tea. Phytochemistry 33:1373-1375.
Hansen EH, Møller BL, Kock GR, Bünner CM, Kristensen C, Jensen OR, Okkels FT, Olsen CE, Motawia MS, Hansen J. 2009. De novo biosynthesis of vanillin in fission yeast (*Schizosaccharomyces pombe*) and baker's yeast (*Saccharomyces cerevisiae*). Applied and Environmental Microbiology 75:2765-2774.
Härtl K, Huang FC, Giri AP, Franz-Oberdorf K, Frotscher J, Shao Y, Hoffmann T, Schwab W. 2017. Glucosylation of Smoke-Derived Volatiles in Grapevine (*Vitis vinifera*) is Catalyzed by a Promiscuous Resveratrol/Guaiacol Glucosyltransferase. Journal of Agricultural and Food Chemistry 65:5681-5689.

Heidlindemann M, Rulli G, Berkessel A, Hummel W, Gröger H. 2014. Combination of Asymmetric Organo- and Biocatalytic Reactions in Organic Media Using Immobilized Catalysts in Different Compartments. ACS Catalysis 4:1099-1103.
Hemingway KM, Alston MJ, Chappell CG, Taylor AJ. 1999. Carbohydrate-flavour conjugates in wine. Carbohydrate polymers 38:283-286.
Herrmann A. 2007. Controlled release of volatiles under mild reaction conditions: from nature to everyday products. Angewandte Chemie International Edition in English 46:5836-5863.
Herrmann KM, Weaver LM. 1999. The shikimate pathway. Annual review of plant biology 50:473-503.
Horwich AL, Fenton WA, Chapman E, Farr GW. 2007. Two families of chaperonin: physiology and mechanism. Annual Review of Cell and Developmental Biology 23:115-145.
Hussain MS, Fareed S, Saba Ansari M, Rahman A, Ahmad IZ, Saeed M. 2012. Current approaches toward production of secondary plant metabolites. Journal of pharmacy & bioallied sciences 4:10.
Hyung Ko J, Gyu Kim B, Joong-Hoon A. 2006. Glycosylation of flavonoids with a glycosyltransferase from *Bacillus cereus*. FEMS microbiology letters 258:263-268.
Imark C, Kneubühl M, Bodmer S. 2000. Occurrence and activity of natural antioxidants in herbal spirits. Innovative Food Science & Emerging Technologies 1:239-243.
Ishag KEA, Jork H, Zeppezauer M. 1985. Mono-und Sesquiterpenalkoholglykoside Mono-and sesquiterpene alcohol glycosides. Fresenius' Zeitschrift für analytische Chemie 321:331-336.
Iverson SV, Haddock TL, Beal J, Densmore DM. 2016. CIDAR MoClo: Improved MoClo Assembly Standard and New *E. coli* Part Library Enable Rapid Combinatorial Design for Synthetic and Traditional Biology. ACS Synthetic Biology 5:99-103.
Jain K, Kesharwani P, Gupta U, Jain NK. 2012. A review of glycosylated carriers for drug delivery. Biomaterials 33:4166-4186.
Joseph BC, Pichaimuthu S, Srimeenakshi S, Murthy M, Selvakumar K, Ganesan M, Manjunath SR. 2015. An overview of the parameters for recombinant protein expression in *Escherichia coli*. Journal of Cell Science and Therapy 6:1.
Kane JF. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. Current opinion in biotechnology 6:494-500.
Kaşıkcı MB, Bağdatlıoğlu N. 2016. Bioavailability of quercetin. Current Research in Nutrition and Food Science Journal 4:146-151.
Kitao S, Sekine H. 1994. α-D-Glucosyl transfer to phenolic compounds by sucrose phosphorylase from *Leuconostoc mesenteroides* and production of α-arbutin. Bioscience, Biotechnology, and Biochemistry 58:38-42.
Königs W, Knorr E. 1901. Über einige Derivate des Traubenzuckers und der Galactose. Berichte der deutschen chemischen Gesellschaft 34:957-981.
Krauss U, Jäger VD, Diener M, Pohl M, Jaeger K-E. 2017. Catalytically-active inclusion bodies—Carrier-free protein immobilizates for application in biotechnology and biomedicine. Journal of biotechnology 258:136-147.
Krempl PM, Mairhofer J, Eisenkolb M, Specht T, Leparc GG, Kreil DP, Bayer K, Striedner G. 2008. *Escherichia coli* BL21(DE3), complete genome. (06-18-2019 2019; https://www.ncbi.nlm.nih.gov/nuccore/AM946981.2)
Laursen BS, Sorensen HP, Mortensen KK, Sperling-Petersen HU. 2005. Initiation of protein synthesis in bacteria. Microbiology and Molecular Biology Reviews 69:101-123.
Lawless H. 1991. The sense of smell in food quality and sensory evaluation. Journal of Food Quality 14:33-60.
Lenski RE. 1998. Bacterial evolution and the cost of antibiotic resistance. International Microbiology 1:265-270.
Li Y, Li Y, Wang Y, Chen L, Yan M, Chen K, Xu L, Ouyang P. 2016. Production of rebaudioside a from stevioside catalyzed by the engineered *Saccharomyces cerevisiae*. Applied biochemistry and biotechnology 178:1586-1598.
Liu J, Zhu XL, Ullah N, Tao YS. 2017. Aroma Glycosides in Grapes and Wine. Journal of Food Science 82:248-259.
Lopez-Romero JC, Gonzalez-Rios H, Borges A, Simoes M. 2015. Antibacterial Effects and Mode of Action of Selected Essential Oils Components against *Escherichia coli* and

*Staphylococcus aureus*. Evidence-based Complementary and Alternative Medicine 2015:795435.
Malhotra A. 2009. Tagging for protein expression. Pages 239-258. Methods in Enzymology, vol. 463 Elsevier.
Marlatt C, Ho CT, Chien M. 1992. Studies of aroma constituents bound as glycosides in tomato. Journal of Agricultural and Food Chemistry 40:249-252.
Martinez-Morales F, Borges A, Martinez A, Shanmugam K, Ingram L. 1999. Chromosomal integration of heterologous DNA in *Escherichia coli* with precise removal of markers and replicons used during construction. Journal of bacteriology 181:7143-7148.
Mayerl F, Näf R, Thomas AF. 1989. 2, 5-Dimethyl-4-hydroxy-3 (2H)-furanone glucoside: Isolation from strawberries and synthesis. Phytochemistry 28:631-633.
Mendoza-Vargas A, Olvera L, Olvera M, Grande R, Vega-Alvarado L, Taboada B, Jimenez-Jacinto V, Salgado H, Juárez K, Contreras-Moreira B. 2009. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. PLoS One 4:e7526.
Meyer AS, Baker TA. 2011. Proteolysis in the *Escherichia coli* heat shock response: a player at many levels. Current opinion in microbiology 14:194-199.
Miroux B, Walker JE. 1996. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. Journal of Molecular Biology 260:289-298.
Moradi SV, Hussein WM, Varamini P, Simerska P, Toth I. 2016. Glycosylation, an effective synthetic strategy to improve the bioavailability of therapeutic peptides. Chemical Science 7:2492-2500.
Nandurkar NS, Zhang J, Ye Q, Ponomareva LV, She QB, Thorson JS. 2014. The identification of perillyl alcohol glycosides with improved antiproliferative activity. Journal of Medicinal Chemistry 57:7478-7484.
Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T. 1998. Chaperone coexpression plasmids: Differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. Applied and Environmental Microbiology 64:1694-1699.
Nishihara K, Kanemori M, Yanagi H, Yura T. 2000. Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. Applied and Environmental Microbiology 66:884-889.
Nishikitani M, Kubota K, Kobayashi A, Sugawara F. 1996. Geranyl 6-O-α-L-arabinopyranosyl-β-D-glucopyranoside isolated as an aroma precursor from leaves of a green tea cultivar. Bioscience, Biotechnology, and Biochemistry 60:929-931.
Noor R, Islam Z, Munshi SK, Rahman F. 2013. Influence of temperature on *Escherichia coli* growth in different culture media. Journal of Pure and Applied Microbiology 7:899-904.
O'Neill EC, Field RA. 2015. Enzymatic synthesis using glycoside phosphorylases. Carbohydrate research 403:23-37.
Ohgami S, Ono E, Horikawa M, Murata J, Totsuka K, Toyonaga H, Ohba Y, Dohra H, Asai T, Matsui K. 2015. Volatile glycosylation in tea plants: sequential glycosylations for the biosynthesis of aroma β-primeverosides are catalyzed by two *Camellia sinensis* glycosyltransferases. Plant Physiology 168:464-477.
Oka N, Ikegami A, Ohki M, Sakata K, Yagi A, Watanabe N. 1998. Citronellyl disaccharide glycoside as an aroma precursor from rose flowers. Phytochemistry 47:1527-1529.
Oka N, Ohki M, Ikegami A, Sakata K, Watanabe N. 1997. First isolation of geranyl disaccharide glycosides as aroma precursors from rose flowers. Natural Product Letters 10:187-192.
Pandey RP, Malla S, Simkhada D, Kim B-G, Sohng JK. 2013. Production of 3-O-xylosyl quercetin in *Escherichia coli*. Applied microbiology and biotechnology 97:1889-1901.
Paquette SM, Jensen K, Bak S. 2009. A web-based resource for the Arabidopsis P450, cytochromes b5, NADPH-cytochrome P450 reductases, and family 1 glycosyltransferases (http://www.P450.kvl.dk). Phytochemistry 70:1940-1947.
Parsell DA, Silber KR, Sauer RT. 1990. Carboxy-terminal determinants of intracellular protein degradation. Genes & development 4:277-286.

Pei J, Kim BH, Grishin NV. 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic acids research 36:2295-2300.
Pérez AG, Olías R, Olías JM, Sanz C. 1999. Biosynthesis of 4-hydroxy-2, 5-dimethyl-3 (2 H)-furanone and derivatives in *in vitro* grown strawberries. Journal of Agricultural and Food Chemistry 47:655-658.
Perreault M, Bialek A, Trottier J, Verreault M, Caron P, Milkiewicz P, Barbier O. 2013. Role of glucuronidation for hepatic detoxification and urinary elimination of toxic bile acids during biliary obstruction. PLoS One 8:e80994.
Plank J. 2011. Eliminate the Growth Lag with Large *E. coli* Cultures. Bitesize Bio: Science Squared.
Puigbo P, Guzman E, Romeu A, Garcia-Vallve S. 2007. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic acids research 35:W126-W131.
Robinson M, Lilley R, Little S, Emtage J, Yarranton G, Stephens P, Millican A, Eaton M, Humphreys G. 1984. Codon usage can affect efficiency of translation of genes in *Escherichia coli*. Nucleic acids research 12:6663-6671.
Rosano GL, Ceccarelli EA. 2009. Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted *Escherichia coli* strain. Microbial cell factories 8:41.
---. 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. Frontiers in microbiology 5:172.
Rose T, Kunz M. 2002. Production of isomalt. Landbauforsch Völkenrode (2002) SH 241:75-80.
Ruffing A, Mao Z, Chen RR. 2006. Metabolic engineering of *Agrobacterium sp.* for UDP-galactose regeneration and oligosaccharide synthesis. Metabolic engineering 8:465-473.
Sabri S, Steen JA, Bongers M, Nielsen LK, Vickers CE. 2013. Knock-in/Knock-out (KIKO) vectors for rapid integration of large DNA sequences, including whole metabolic pathways, onto the *Escherichia coli* chromosome at well-characterised loci. Microbial cell factories 12:60.
Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold spring harbor laboratory press.
Samuel H, Yalkowsky H, Jain P. 2003. Handbook of aqueous solubility data: CRC Press: Boca Raton, FL.
Sarry J, Gunata Z. 2004. Plant and microbial glycoside hydrolases: Volatile release from glycosidic aroma precursors. Food Chemistry 87:509-521.
Schmideder A, Priebe X, Rubenbauer M, Hoffmann T, Huang F-C, Schwab W, Weuster-Botz D. 2016. Non-water miscible ionic liquid improves biocatalytic production of geranyl glucoside with *Escherichia coli* overexpressing a glucosyltransferase. Bioprocess and biosystems engineering 39:1409-1414.
Schmolzer K, Gutmann A, Diricks M, Desmet T, Nidetzky B. 2016. Sucrose synthase: A unique glycosyltransferase for biocatalytic glycosylation process development. Biotechnology Advances 34:88-111.
Schwab W, Fischer TC, Giri A, Wüst M. 2015. Potential applications of glucosyltransferases in terpene glucoside production: impacts on the use of aroma and fragrance. Applied microbiology and biotechnology 99:165-174.
Seibel J, Jördening H-J, Buchholz K. 2006. Glycosylation with activated sugars using glycosyltransferases and transglycosidases. Biocatalysis and biotransformation 24:311-342.
Sekiwa Y, Kobayashi A, Kubota K, Takenaka M. 2001a. First isolation of geranyl disaccharides from ginger and their relations to aroma formation. Natural Product Letters 15:267-274.
Sekiwa Y, Mikami N, Kubota K, Kobayashi A. 2001b. Formation of geraniol-related compounds in ginger. Pages 305-312 in Spanier AM, Shahidi F, Parliment TH, Mussinan C, Ho CT, Contis ET, eds. Food Flavors and Chemistry: Advances of the New Millennium, The Royal Society of Chemistry.
Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. The Plant Cell 17:3141-3154.
Shao J, Hayashi T, Wang PG. 2003. Enhanced Production of α-Galactosyl Epitopes by Metabolically Engineered *Pichia pastoris*. Applied and Environmental Microbiology 69:5238-5242.

Shitan N. 2016. Secondary metabolites in plants: transport and self-tolerance mechanisms. Bioscience, Biotechnology, and Biochemistry 80:1283-1293.
Siurkus J, Neubauer P. 2011. Heterologous production of active ribonuclease inhibitor in *Escherichia coli* by redox state control and chaperonin coexpression. Microbial cell factories 10:65.
Sivashanmugam A, Murray V, Cui C, Zhang Y, Wang J, Li Q. 2009. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. Protein Science 18:936-948.
Solá RJ, Griebenow K. 2010. Glycosylation of therapeutic proteins. BioDrugs 24:9-21.
Sørensen HP, Mortensen KK. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. Journal of biotechnology 115:113-128.
St-Pierre F, Cui L, Priest DG, Endy D, Dodd IB, Shearwin KE. 2013. One-step cloning and chromosomal integration of DNA. ACS Synthetic Biology 2:537-541.
Stevenson PC, Nicolson SW, Wright GA, Manson J. 2017. Plant secondary metabolites in nectar: impacts on pollinators and ecological functions. Functional Ecology 31:65-75.
Straubinger M, Knapp H, Watanabe N, Oka N, Washio H, Winterhalter P. 1999. Three novel eugenol glycosides from rose flowers, *Rosa damascena* Mill. Natural Product Letters 13:5-10.
Studier FW. 2005. Protein production by auto-induction in high-density shaking cultures. Protein expression and purification 41:207-234.
Svasti J, Phongsak T, Sarnthima R. 2003. Transglucosylation of tertiary alcohols using cassava β-glucosidase. Biochemical and biophysical research communications 305:470-475.
Takeda Y, Ooiso Y, Masuda T, Honda G, Otsuka H, Sezik E, Yesilada E. 1998. Iridoid and eugenol glycosides from *Nepeta cadmea*. Phytochemistry 49:787-791.
Tapader R, Bose D, Pal A. 2017. YghJ, the secreted metalloprotease of pathogenic *E. coli* induces hemorrhagic fluid accumulation in mouse ileal loop. Microbial pathogenesis 105:96-99.
Tats A, Tenson T, Remm M. 2008. Preferred and avoided codon pairs in three domains of life. BMC Genomics 9:463.
Terpe K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Applied microbiology and biotechnology 60:523-533.
Terpe K. 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Applied microbiology and biotechnology 72:211.
Thomas JG, Baneyx F. 1996. Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins. Journal of Biological Chemistry 271:11141-11147.
---. 1997. Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*. Protein expression and purification 11:289-296.
Torres P, Poveda A, Jimenez-Barbero J, Parra JL, Comelles F, Ballesteros AO, Plou FJ. 2011. Enzymatic Synthesis of α-Glucosides of Resveratrol with Surfactant Activity. Advanced Synthesis & Catalysis 353:1077-1086.
Turner NJ. 2009. Directed evolution drives the next generation of biocatalysts. Nature Chemical Biology 5:567.
Tyo KE, Ajikumar PK, Stephanopoulos G. 2009. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. Nature biotechnology 27:760.
Ueda H, Kikuta Y, Matsuda K. 2012. Plant communication: mediated by individual or blended VOCs? Plant Signaling & Behavior 7:222-226.
Ullmann A. 2009. *Escherichia coli* Lactose Operon. Encyclopedia of Life Sciences, John Wiley & Sons, Inc.
Urban-Chmiel R, Dec M, Puchalski A, Wernicki A. 2013. Characterization of heat-shock proteins in *Escherichia coli* strains under thermal stress *in vitro*. Journal of Medical Microbiology 62:1897-1901.
Van Rantwijk F, Woudenberg-van Oosterom M, Sheldon R. 1999. Glycosidase-catalysed synthesis of alkyl glycosides. Journal of Molecular Catalysis B: Enzymatic 6:511-532.

Verschueren K. 2001. Handbook of environmental data on organic chemicals: Vol. 1. John Wiley and Sons, Inc.
Vogt T, Jones P. 2000. Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends in plant science 5:380-386.
Vollmer W, Blanot D, de Pedro MA. 2008. Peptidoglycan structure and architecture. FEMS Microbiology Reviews 32:149-167.
Wada K-n, Wada Y, Ishibashi F, Gojobori T, Ikemura T. 1992. Codon usage tabulated from the GenBank genetic sequence data. Nucleic acids research 20:2111.
Watson JD, Crick FH. 1953. The structure of DNA. Pages 123-131. Cold Spring Harbor symposia on quantitative biology: Cold Spring Harbor Laboratory Press.
Werner SR, Morgan JA. 2009. Expression of a *Dianthus* flavonoid glucosyltransferase in *Saccharomyces cerevisiae* for whole-cell biocatalysis. Journal of biotechnology 142:233-241.
White BE, Fenner CJ, Smit MS, Harrison STL. 2017. Effect of cell permeability and dehydrogenase expression on octane activation by CYP153A6-based whole cell *Escherichia coli* catalysts. Microbial cell factories 16:156.
Williams PJ, Strauss CR, Wilson B, Massy-Westropp RA. 1982. Novel monoterpene disaccharide glycosides of *Vitis vinifera* grapes and wines. Phytochemistry 21:2013-2020.
Yonekura-Sakakibara K, Hanada K. 2011. An evolutionary view of functional diversity in family 1 glycosyltransferases. The Plant Journal 66:182-193.
Yousefi M, Farajnia S, Mokhtarzadeh A, Akbari B, Khosroshahi SA, Mamipour M, Dariushnejad H, Ahmadzadeh V. 2018. Soluble expression of humanized anti-CD20 single chain antibody in *Escherichia coli* by cytoplasmic chaperones co-expression. Avicenna journal of medical biotechnology 10:141.
Zabetakis I, Gramshaw J, Robinson D. 1999. 2, 5-Dimethyl-4-hydroxy-2H-furan-3-one and its derivatives: analysis, synthesis and biosynthesis—a review. Food Chemistry 65:139-151.
Zhang J, Ponomareva LV, Nandurkar NS, Yuan Y, Fang L, Zhan CG, Thorson JS. 2015. Influence of Sugar Amine Regiochemistry on Digitoxigenin Neoglycoside Anticancer Activity. ACS Medicinal Chemistry Letters 6:1053-1058.
Zheng Z, Fang JL, Lazarus P. 2002. Glucuronidation: an important mechanism for detoxification of benzo[a]pyrene metabolites in aerodigestive tract tissues. Drug Metabolism and Disposition 30:397-403.
Zou L, Zhao H, Wang D, Wang M, Zhang C, Xiao F. 2014. Expression and purification of a functional recombinant aspartate aminotransferase (AST) from *Escherichia coli*. Journal of Microbiology and Biotechnology 24:998-1003.