Building a Glycosylation Platform in *E. coli* through Metabolic Engineering

Frederik De Bruyn

Examination committee

Prof. dr. ir. Paul Van Der Meeren (Ghent University, Chair)
Prof. dr. ir. Matthias D'hooghe (Ghent University)
Prof. dr. Magda Faijes (Universitat Ramon Llull, Spain)
Prof. dr. ir. Els Van Damme (Ghent University)
dr. ir. Manu De Groeve (Ablynx)
Prof. dr. ir. Wim Soetaert (Ghent University)
Prof. dr. ir. Marjan De Mey (Ghent University)

Supervisors

Prof. dr. ir. Wim Soetaert (Ghent University)
Prof. dr. ir. Marjan De Mey (Ghent University)
Department of Biochemical and Microbial Technology, Center of Expertise – Industrial Biotechnology and Biocatalysis, Ghent University

Dean

Prof. dr. ir. Guido Van Huylenbroeck

Rector

Prof. dr. Anne De Paepe

Ghent University
Faculty of Bioscience Engineering
Department of Biochemical and Microbial Technology

Building a Glycosylation Platform in *E. coli* through Metabolic Engineering

---

Frederik De Bruyn

Thesis submitted for the fulfillment of the requirements
For the degree of Doctor (PhD) in Applied Biological Sciences
Academic year: 2014-2015

Dutch translation of the title:

Ontwikkeling van een glycosyleringsplatform in *E. coli* door *Metabolic Engineering*

To refer to this thesis:

De Bruyn, F. (2014) Building a Glycosylation Platform in *E. coli* through Metabolic Engineering. PhD thesis, Faculty of Bioscience Engineering, Ghent University, Ghent.

Cover illustration:
©iStockphoto.com
Edited by Brecht De Paepe and Frederik De Bruyn

ISBN 978-90-5989-744-1

Frederik De Bruyn was supported by a fellowship of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen)

Woord vooraf

In de loop van de voorbije jaren heb ik vaak gemijmerd hoe een ideaal woord vooraf er zou kunnen uitzien. Het is meer dan zomaar een proloog of een dankwoord. Het zou eerder een metastabiele balans moeten zijn tussen het oproepen van de sfeer waarin ik vier jaar heb vertoefd, doorspekt met stilistische fijnzinnigheden en een portie karikaturale humor om de lezer niet te vervelen, en tussen een terechte dankzegging. Deze finale versie op papier krijgen bleek echter een werk van lange adem en is hetgeen u nu leest. Een kortverhaal waarin de beslommeringen en wederwaardigheden die mij te beurt zijn gevallen uit de doeken worden gedaan – vol herkenbare situaties, oneliners en een vleugje weemoedigheid – en waarin vele mensen een hoofd- of bijrol hebben gespeeld.

Alvorens alles teveel verbloemd wordt, dien ik er wel de nadruk op te leggen dat de voorbije vier jaar bij wijlen een aaneenschakeling waren van frustratiemomenten, gevloek en eindeloos geploeter omdat 'de beestjes' niet wilden groeien of besloten om pas tegen de avond in gang te schieten. Uiteraard zijn er nog andere topbelevenissen – *horresco referens* – zoals 's nachts met goede moed een staal nemen uit de reactor die blijkt overgeschuimd te zijn, zaterdagochtend katerig de warme kamer binnenwandelen en ontdekken dat je toch geen kolonies hebt of snel nog een groeiproef willen opzetten en vaststellen dat je geliefkoosde mediumfles weer gecontamineerd is met een schimmelbol… Misschien heb ik daarom wel vooraan dit proefschrift de woorden van Hannibal neergepend. *Ofwel vind ik een weg, ofwel maak ik er een*. Dat was zijn antwoord toen één van zijn generaals zei dat het onmogelijk was om de Alpen over te steken met olifanten. Welja, de vader der strategie achterna heb ik ook steeds een weg proberen vinden wanneer er zich verschillende obstakels op mijn pad bevonden.

In tegenstelling tot Hannibal had ik twee andere generaals aan mijn zijde die met veel graagte de moeilijkst begaanbare paden wilden opzoeken en wiens visie en overredingskracht mij aanspoorden om de rug te rechten en door te gaan tijdens de donkerste uren van dit proefschrift. Daarom ben ik ook veel dank verschuldigd aan Marjan De Mey en Wim Soetaert omdat ze mij de mogelijkheid hebben gegeven om mij te vervolmaken in dit boeiende onderzoeksdomein in een ongedwongen setting. Ze zorgden ervoor dat InBio (en zeker de MEMO-groep!) op de kaart werden gezet en het ultieme voorwerp van afgunst waren van menig vakgroep. Waar anders werd gezellig ontbeten, gebarbecued of gepintelierd? Waar anders werd lief en leed gedeeld of geschaterlacht? Waar anders stond men schouder aan schouder om ons territorium af te bakenen in de koepuur? Ik zal er niet ver naast zitten als ik stel dat *ons labo* een van de fijnste plekken is om onderzoek te verrichten. Het is een dorp waar iedereen elkaar (te

goed) kent en maandag staat te popelen om met een koffie in de hand elkaars geheugen op te frissen omtrent de doldwaze avonturen van afgelopen vrijdag…

Het echte orgelpunt van vier jaar doctoreren was misschien wel het "zorgvuldig" samenstellen van het beruchte UDP-suiker team. Deze *League of Extraordinary Gentlemen* werden als thesisstudenten in het heetste vuur gesmeed tot keurtroepen die mij hebben bijgestaan om de demonen van dit onderzoek te bedwingen. Zonder hun inzet en veerkracht was dit boekje niet geworden tot wat het nu is. Zonder morren werden groeiproeven herhaald, PCR's opnieuw ingezet of vermaledijde *vials* op de mysterieuze Shimadzu gestoken. Het moreel van het klein labo's *finest* werd steevast opgekrikt met een verraderlijk streekbier of een feestdis in de Pampa's. Hoewel hun namen later vast wel vereeuwigd en bezongen zullen worden, lijst ik ze hier toch nogmaals op. Brecht – tovenaar der TLC's en productiemedia – met jou had ik de eer en het genoegen om samen de eerste ongelooflijke glucosyleringsresultaten te behalen die tevens het begin van een nieuw doctoraatstijdperk inluidden. Maar al die wetenschap verbleekte bij de talloze *side-experiments* die we daarnaast hebben opgezet ("ik doe da gewoon he") of hoe we op een ongeëvenaarde manier legio wereldmysteries en samenzweringstheorieën oplosten, doorprikten of zelfs versterkten. Een pluim voor het *out of the box*-denken, het verdragen van mijn Michael Scott imitaties en je vriendschap. De volgende in rij was Jarno, die na ettelijke worstelpogingen met SuSy het enzym deels op de knieën kreeg. Zijn goedgemutstheid en kalmte waren legendarisch, alsook zijn alomgekende *jeeps* ochtend- en avondgroet. Aansluitend brak het jaar aan van de *dweepkes* Pieter (Cockeeer!) en Maarten (Van Breeempt!), die het begrip '*in vivo* glycosyleren' naar een ander niveau tilden. Maarten, ik sta nog steeds versteld hoeveel constructen iemand kan maken én volledig testen. Hoeveel keer namen we ons niet voor om een paar Ferrari's te kopen met onze bereide globotriose? En Pieter, hoeveel keer heb jij niet gevloekt op *Beverly* (nvdr: reactor B+1) of gebeld met de ontstellende woorden "ja lap, da medium is weer verkleurd"? Gespannen momenten werden dan ook geventileerd met spontane *dance show-offs* in het labo, de meeste bizarre imitaties ("*Don't you die on me Billy*") of een volmondig KHA-KHAAAA! Gasten, merci!

Veel dank ben ik eveneens verontschuldigd aan het gouden drietal dat de MEMO-groep jaren in goeie banen heeft geleid en mij steevast met raad en daad heeft bijgestaan. Vooreerst wil ik nogmaals Marjan in de bloemetjes zetten omwille van de uren en uren tijd die je in mij geïnvesteerd hebt tijdens het samen vorm geven, nalezen en zeer kritisch evalueren van dit werkstuk. Het kostte soms wat tijd, maar de violen werden steeds gelijk gestemd. Ook Jo, dag en nacht bereikbaar voor "Bruin", wil ik bedanken. Jouw *vage* schematische pentekeningen op maagdelijk wit papier bleken meer dan eens een waardevolle leidraad voor het structureren van papers en presentaties. Jouw onblusbare inzet alsook liefde voor lekker eten en cultuur lieten mij snel inzien dat zich achter de cynische humor een warme vriend schuilhield, wiens blikken en korte tussenwerpsels zoals *koepuur!*, *urrrgh* en *ow fuck* genoeg waren om te snappen wat je wilde zeggen. Tot slot mag ook de immer enthousiaste Joeri niet in de rij ontbreken. Hoewel je het veelal erg druk had, nam je de tijd om je licht te laten schijnen of wilde

verklaringen te zoeken voor de rare fenomenen die optraden tijdens mijn onderzoek. Tezamen hebben jullie een sterke basis gelegd en mij behoed voor de valkuilen van het doctoreren.

Niet in het minst wil ik ook alle andere InBio-kameraden in de schijnwerpers stellen, temeer omdat zij het labo écht maken tot wat het is. Mijn courante habitat (afgezien van de koffieruimte) was de bureau beneden waar een deel van de harde MEMO-kern gestationeerd was. Ondanks de grote *turnover* bleven enkelen de voorbije vier jaar tot het vaste meubilair behoren. Pieter en Gert – *de boyz* – waren altijd paraat voor overleg, nieuwe ideeën of een vroege aperitief. Met ons drie wilden we een jeugdig regime installeren, een activiteitenloos MEMO-weekend organiseren en goud behalen met het iGEM team… Achja, de lat mag soms hoog genoeg liggen. Gaspard, wiens practical jokes en sarcastisch geplaag vaak mythische proporties aannamen, bood graag hulp bij statistische problemen en bleef maar aandringen om op Linux over te stappen. Ook Sofie DM maakte graag tijd voor een goed gesprek en leverde een wijsheid aan die volledig in mijn kraam paste: “als het twee keer niet lukt, doe dan iets anders”. Ook de recentere aanwinsten Thomas, Bob, Maarten DM, Tom en David zullen ongetwijfeld bijdragen tot voortzetting van de topsfeer beneden en een nieuwe wind doen waaien. Ook ben ik zeer verheugd dat niemand minder dan Wouter nu plaats heeft genomen achter de bureau waar ik vier jaar gesleten heb. Jouw onvermoeibaar enthousiasme, werkethiek en grenzeloze sfeermakerij zijn niet te beschrijven, tenzij je het zelf hebt meegemaakt. Señor, hoed af! Samen met je kompaan Dries D hebben jullie cruciaal werk verricht in de lastigste en laatste uren van mijn doctoraat. Zonder jullie kon ik dit woord vooraf waarschijnlijk maar een maand later neerpennen.

Om nog maar te zwijgen over de *diehards* van het klein labo. Jullie bureau was een vaste halte tijdens de ochtendkoffie (sorry voor de achtergebleven tassen!), een ware oase die soms ontsierd werd door een welriekende autoclaaf en waar de nieuwste nieuwtjes werden uitgewisseld. Hoe dikwijls kwam ik niet bij Maarten, Dries, Eric, Gilles en Catherine binnen met een onoplosbaar PCR-mysterie, een spoedbestelling, een vastgelopen HPLC of een vraag tot medium-uitlening. En tegelijk moest ik jullie ontgoochelen met het fameuze *venting*-experiment of verzette ik de post al eens naar Topradio. Ik zal niet snel de blik in Eric zijn ogen vergeten als we weer *weekly supply sh*t manager* bleken te zijn, alsook de ongeëvenaarde tijd waarop we dit toch maar weer klaarspeelden à la monkeystyle. Of de talloze weddenschappen over geografische- en dierenweetjes waarbij ik jammerlijk de duimen moest leggen tegen de *disco lion*. Of hoe Vrijdag-Dries mij vier jaar geleden de twijfel ontnam om te doctoreren. Aansluitend wil ik de goedlachse MEMO-dames Lien en Isabelle bedanken, alsook deze van het coördinatieteam (Hilde, Dominique, Barbara, Anneleen en Liesbet) voor de administratieve rompslomp in goede banen te leiden. Ook de Glyco-collega’s en in het bijzonder Karel, Tom, Magali en Mareike stonden paraat als het over enzymen of een goed feestje ging! En dan heb ik het nog niet gehad over de kleurrijke figuren van Biosurf zoals daar zijn: Sophie – *eierkopke* – Roelants, HPLC/dj Stijn, de overenthousiaste Isabelle en Robin – de broer die ik nooit gehad heb. De soms iets te amicale avonturen

die Robin en ik beleefd hebben zijn meer dan memorabel, en ik zie er dan ook zeer naar uit om de *Blue Lagoon*, de kelder van een niet nader genoemd café en de Bourgoyen nog eens op stelten te zetten!

Overigens wens ik ook mijn vriendengroep buiten de labomuren te bedanken, die al 9 jaar het vaste decor vormen voor feestjes op ons *boerekot*, Brusseldrinks en heuse roadtrips. Stijn, Kèvn, Langen, Claire, Jef, Spekkie, Ruben en Ken: twas plezant! Ook Tom wil ik bedanken, een globetrotter van het eerste uur, met wie ik de eer heb gehad een groot deel van de Oude Wereld te verkennen, deels uit avonturenlust, deels uit weetgierigheid, zonder ons de allure van ontdekkingsreizigers aan te meten. Het maken van de meest absurde Slavische woordgroepen, de Kaukasische kookinitiaties en de talloze zoo-bezoeken op de Balkan staan allemaal in ons collectief geheugen geprent.

Het mooiste en dierbaarste dat deze doctoraatsjaren mij echter gebracht hebben, is de vrouw van mijn leven, Margo. Ik zal niet gauw vergeten hoe we naar elkaar zijn toegegroeid tijdens de *summerschool* die plaatsvond in de zwoele zomer van 2013. Ik denk dat onze genegenheid zelfs aanstekelijk heeft gewerkt op het labo… Dat we twee handen op een buik zijn, is meer dan eens gebleken tijdens de laatste loodjes van mijn doctoraat. Bedankt dat ik kon klagen, zagen en mijn hart uitstorten wanneer het allemaal onmogelijk leek, bedankt dat jij de stress voor mij ventileerde, bedankt voor je kritische kijk en SuSy tips, bedankt voor de momenten van puur geluk, bedankt om wie je bent. Er wachten ons nog vele avonturen, de wereld is aan ons!

Als laatste, en het hoogst van al in het vaandel gedragen, wil ik mijn ouders bedanken. Zij hebben ervoor gezorgd dat ik sta waar ik nu sta door hun onvoorwaardelijke steun en motivatie. Twee bijzondere mensen die meermaals (ja, zelfs wekelijks) op het labo geciteerd werden, waarbij ik meestal aanhief met de woorden “mijn moeder/vader zei altijd …”. Meestal volgde hierop een pakkende uitspraak die menig omstaander met verstomming sloeg en dikwijls maar het topje van de kennisijsberg waren die zij mij hebben meegegeven. Deze mensen hebben thuis een warm nest gecreëerd waar vele onderwerpen werden aangesneden (gaande van gouden reistips, ideeën voor *self-sufficiency* en het maken van wijn, tot uitdieping van kennis over de klassieke auteurs), en waar ik altijd met raad en daad werd bijgestaan. Daarnaast waren de *kleine* hulpstukjes zoals een meegekregen bakje ingevroren appelmoes, een proper gestreken hemd of een telefoontje om te vragen hoe het ging meer dan welkom en vormden zij tegelijk de steunpilaren voor dit *grote* bolwerk en de mens die ik geworden ben. Woorden kunnen niet genoeg beschrijven hoeveel ik jullie verschuldigd ben. Mama, Papa, bedankt! Ook andere familieleden speelden een sleutelrol, zoals mijn zus Caroline, met wie ik dikwijls samen de Pinuts onveilig maakte, of mijn grootouders, die met grote fierheid hun streekgebonden wijsheden doorgaven vergezeld van koffie en gebak.

Ziezo, dit zal de laatste maal zijn dat ik de hand leg aan *mijn boekje*. Geschreven aan een bureau die er al vier jaar hetzelfde uitziet en getooid is met een gelukSprentje gekregen van mijn tante bij aanvang van mijn studies, een Oekraïens vlagje van toen de Krim nog

niet Russisch was en een moleculaire structuur van UDP-glucose terwijl ik uitkijk op een beduimeld papiertje met aminozuren. De boeken kunnen toe, de blik kan op oneindig… Stiekem hopend dat dit werkstuk of één van mijn alternatieve levenswijsheden als leidraad kan dienen voor anderen. Want *sed omnia praeclara tam difficilia quam rara sunt*.[1]

Frederik

10 november 2014

[1] Baruch Spinoza (1632-1677)

Abbreviations

μ	specific growth rate
ADP	adenosine diphosphate
AMP	adenosine monophosphate
ATP	adenosine triphosphate
BGG	β-glucogallin
CDW	cell dry weight
F6P	fructose 6-phosphate
GA	gallic acid
gal1P	galactose 1-phosphate
GC	gas chromatography
glc1P	glucose 1-phosphate
glc6P	glucose 6-phosphate
GNB	galacto-*N*-biose
GT	glycosyltransferase
HBA	hydroxybenzoic acid
HCA	hydroxycinnamic acid
HMO	human milk oligosaccharide
HPLC	high performance liquid chromatography
KI	knockin
KO	knockout
LB	Luria Bertani
LC	liquid chromatography
LNB	lacto-*N*-biose
MES	2-(*N*-morpholino)ethanesulfonic acid
MOPS	3-(*N*-morpholino)propanesulfonic acid
MS	mass spectrometry
NAD	nicotinamide adenine dinucleotide
OD	optical density
PEP	phosphoenolpyruvate
PPi	pyrophosphate
q_p	specific production rate
q_S	specific substrate uptake rate
RC	regeneration cycles

r_p	volumetric production rate
TB	terrific broth
TLC	thin layer chromatography
UDP	uridine diphosphate
UDP-gal	UDP-galactose
UDP-glc	UDP-glucose
UDP-rha	UDP-rhamnose
UGT	uridine glycosyltransferase
UMP	uridine monophosphate
UTP	uridine triphosphate
WT	wild type
Y	yield

Contents

Abbreviations ..... **x**

Introduction **1**

literature overview **5**

Unravelling the Leloir Pathway of *Bifidobacterium bifidum* **28**

Development of an *in vivo* glucosylation platform by coupling production to growth **48**

Building a versatile glycosylation platform **73**

Unlocking the potential of Sucrose Synthase for *in vivo* glycosylation **97**

Conclusions and perspectives **115**

Appendices **124**

Bibliography **148**

Summary **174**

Samenvatting **177**

Curriculum vitae **181**

Chapter 1

Introduction

During the last couple of decades, industrial biotechnology has evolved from an infant technology into a cornerstone of today's global economy and was recently identified as a Key Enabling Technology[2] by the EU. This biotechnology branch – also called white biotechnology – uses micro-organisms or components thereof (such as enzymes) as industrial catalysts to produce valuable chemicals out of renewable carbon sources or waste streams. Prominent examples are the production of succinic acid[1], lactic acid[2,3], ethanol[4-6], glutamic acid[7], lysine[8-10] and citric acid[11] as bulk chemicals with various applications as solvents, fuel, food additives, precursors for synthesis or as building blocks for creating biodegradable polymers. These processes have replaced the traditional chemical manufacturing, hence supporting the economy to become more independent from fossil resources. On the other hand, biotechnology can aid with its complementary techniques in the existing chemical synthesis, thus shifting from stand-alone technologies towards the beneficial integration and expansion of these two fields.[12] Organic synthesis has the major advantages of broad reactivity and rapid optimization, while biological synthesis is characterized by high selectivity, mild conditions and sustainability.

Interesting compounds in this perspective are "small molecules", which often refers to specialized (secondary) metabolites such as alkaloids, flavonoids, phenolic acids, antibiotics or terpenoids. These molecules and analogues thereof are in high demand due to their interesting properties and applications in pharmaceutical, cosmetic or food industries, and are usually produced by physicochemical extraction or chemical synthesis. However, the majority of specialized metabolites in nature occurs in its glycosylated form, hereby greatly altering the solubility, stability or bioactivity of these molecules, which are desirable properties for their application in various fields. Since selective attachment of a sugar residue is chemically a daunting task, a myriad of biotechnological processes have been developed using enzymes (*in vitro*) or a whole-cell bioconversion (*in vivo*) approach. Key players in this context are glycosyltransferases (GTs), which are enzymes that regio- and stereoselectively transfer a sugar residue from

[2] http://ec.europa.eu/enterprise/sectors/ict/key_technologies/index_en.htm

an activated nucleotide sugar donor (often UDP-sugar) to various small molecules.[13] These *in vitro* methods are however often plagued by scalability issues, the need for purified proteins and the requirement for (equimolar) amounts of expensive donors (e.g. UDP-glucose: 150 €/g; UDP-galactose: 1500 €/g). Fueled by these challenges and a high demand, various *in vivo* processes have been developed whereby the UDP-sugars and enzymes are formed by an (engineered) micro-organism from cheap substrates. These recent biotechnological advances are critically reviewed in **Chapter 2** and the major hurdles for economic viability of the processes are identified. These include low conversion yields and titers, inefficient UDP-sugar formation, limited acceptor flexibility and problems regarding scale-up.

Hence, the objective of this doctoral research is to design a novel *in vivo* glycosylation platform which is generic, easily scalable and actively couples growth and production, thus overcoming the aforementioned hurdles. To this end, the model organism *Escherichia coli* W was metabolically engineered to create a completely novel glycosylation strain for various small molecules using GTs. This new approach should enable the bacterium to grow and produce glycosides simultaneously, hereby using only sucrose as an extremely cheap (250 €/ton) and sustainable carbon source. An overview of the envisaged approaches to obtain these objectives is presented as a chapter based flowchart (Figure 1.1).

To efficiently generate UDP-sugars in the cell as a donor for GTs, the expression of a promiscuous uridylyltransferase enzyme is necessary. In **Chapter 3**, the search for such an enzyme was started based on research from the late 70's using crude protein extract from the probiotic *Bifidobacterium bifidum*.[14] Since no sequence information was present on the enzyme, three uridylyltransferase candidates were cloned and investigated with a newly developed chemo-enzymatic assay. UgpA was identified as the enzyme sought after, exhibiting activity towards both glucose 1-phosphate and galactose 1-phosphate resulting in the formation of their respective UDP-sugars.

Chapter 4 describes the first engineering of *E. coli* W for the creation of a glucosylation platform. The strategy makes use of the introduction of an alternative sucrose metabolism in the form of a sucrose phosphorylase. This enzyme splits sucrose into fructose, which is used as a carbon source and glucose 1-phosphate, which is ideally converted into UDP-glucose by the UgpA enzyme selected in Chapter 3. A trimodular engineering approach is applied to obtain a split metabolism where both growth and production (glucosylation) can occur simultaneously. Expression of a glucosyltransferase from *Vitis vinifera* (VvGT2) enables the strain to efficiently produce

the promising compound β-glucogallin (1-*O*-galloyl-β-D-glucose) starting from only gallic acid and sucrose. Due to the broad activity of VvGT2, various other phenolic acids were tested to demonstrate the potential of this platform.

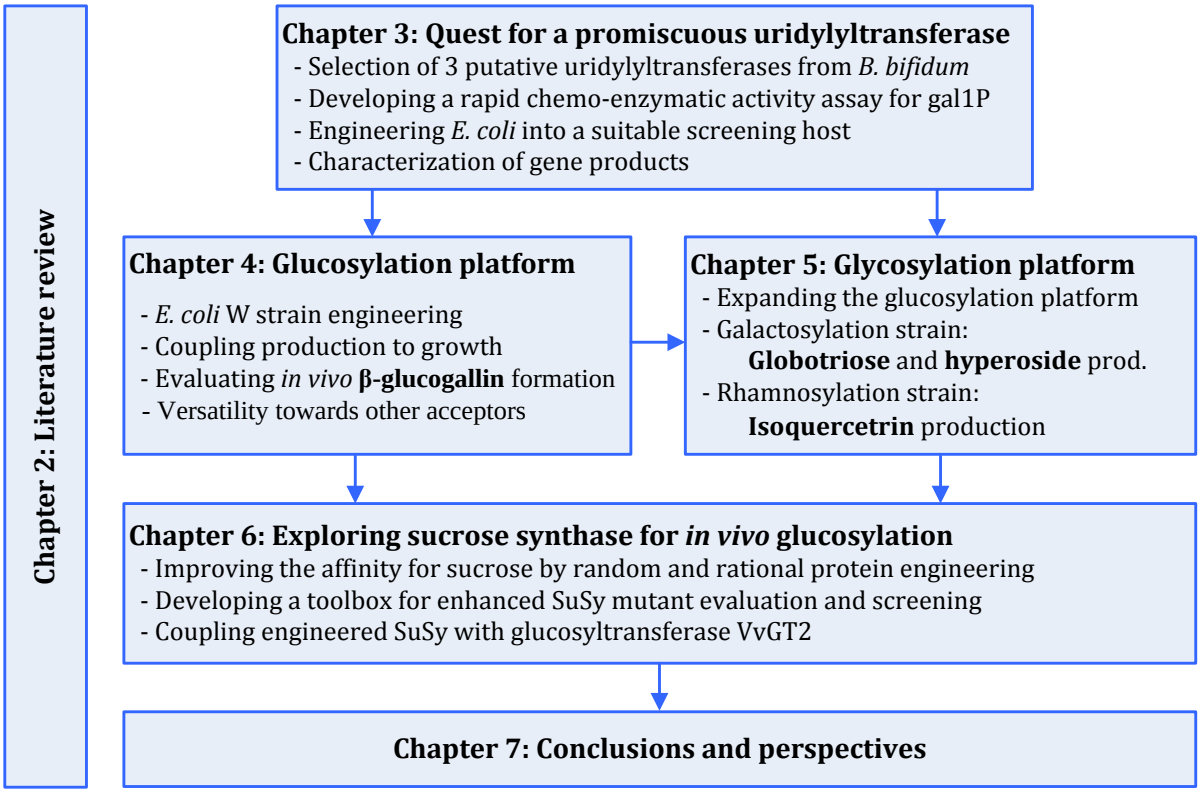


Figure 1.1: Flowchart of the different chapters covered in this thesis and the central objectives to obtain a generic and versatile *in vivo* glycosylation platform.

Because of its generic nature, the glucosylation platform can be easily transformed into a general glycosylation platform. In **Chapter 5** the possibilities are investigated to create a platform for the galactosylation and rhamnosylation of small molecules by expanding the production strain from Chapter 4 with an interconverting enzyme. Overexpression of either a UDP-glucose epimerase (*galE*) or a UDP-rhamnose synthase (*MUM4*) results in the formation of UDP-galactose or UDP-rhamnose respectively. As proof of concepts, the *in vivo* formation of the highly demanded galactosides hyperoside (quercetin 3-*O*-galactoside) and globotriose (α-galactosyl-1,4-lactose) and the rhamnoside quercitrin (quercetin 3-*O*-rhamnoside) will be evaluated and discussed.

Next, in **Chapter 6**, the potential of the promising enzyme sucrose synthase (SuSy) for the one-step generation of UDP-glucose from sucrose is explored. Coupling of SuSy with a glycosyltransferase results in an efficient UDP recycling system, which makes it attractive for industrial applications. However, due to its unfavorable kinetics (low

affinity for sucrose), it has been rarely used for *in vivo* generation of UDP-glucose. In this chapter, SuSy from *Solanum tuberosum* will be engineered to increase its affinity for sucrose and the obtained mutants will be evaluated using a developed screening system. The most active SuSy exhibiting the highest affinity for sucrose will be selected and coupled with the glucosyltransferase VvGT2 to assess the *in vivo* glucosylation potential of this regeneration system.

In a final chapter (**Chapter 7**), the approaches followed in this doctoral research are assessed and the versatile 'plug and play' glycosylation platforms are evaluated. Suggestions are given to further optimize the platforms and expand their industrial applicability.

Chapter 2

Biotechnological Advances in UDP-sugar based Glycosylation of Small Molecules

Literature overview

Contents

2.1 Abstract 6
2.2 Glycosylation: orchestrating life 7
2.3 State of the art glycosylation processes 9
2.4 Formation of common UDP-sugars: sweet routes to synthesis 16
2.5 Interconverting common UDP-sugars 20
2.6 Glycorandomization: beyond nature's glycosylation toolbox 23
2.7 Future outlook 26

This chapter has been submitted as:

De Bruyn, F., Maertens, J., Beauprez, J., Soetaert, W., and De Mey, M. (2014) **Biotechnological advances in UDP-sugar based glycosylation of small molecules.** *Biotechnol Rev, under review.*

2.1 Abstract

Glycosylation of small molecules like specialized (secondary) metabolites has a profound impact on their solubility, stability or bioactivity, making glycosides attractive compounds as food additives, therapeutics or nutraceuticals. The subsequently growing market demand has fuelled the development of various biotechnological processes, which can be divided in the *in vitro* (using enzymes) or *in vivo* (using whole cells) production of glycosides. In this context, uridine glycosyltransferases (UGTs) have emerged as promising catalysts for the regio- and stereoselective glycosylation of various small molecules, hereby using uridine diphosphate (UDP) sugars as activated glycosyldonors. This review gives an extensive overview of the recently developed *in vivo* production processes using UGTs and discusses the major routes towards UDP-sugar formation. Furthermore, the use of interconverting enzymes and glycorandomization is highlighted for the production of unusual or new-to-nature glycosides. Finally, the technological challenges and future trends in UDP-sugar based glycosylation are critically evaluated and summarized.

2.2 Glycosylation: orchestrating life

Glycosylation is one of the most prevalent and important modifications in nature as it plays a central role in cellular communication and the orchestration of life in general. Many diseases are associated with the unusual glycosylation of proteins[15-17] and the recognition of specific cell surface glycans often forms the basis of viral and bacterial infections[18] or cancer metastasis[19]. Sugar residues can be attached to a variety of molecules ranging from macromolecules like proteins, lipids or cell wall glycans to rather small molecules[13] such as specialized (secondary) metabolites and oligosaccharides. Glycosylation of these small molecules often has a profound impact since it can greatly alter their solubility, stability or bioactivity.[20] Remarkable examples are the glycosylation of vitamin C[21] or anthocyanidins[22] for increased stability, the detoxification of xenobiotics from the body by means of glucuronidation[23,24], the glycosylation of steviol resulting in sweet tasting steviosides[25], or the possibility to target galactosylated compounds towards the liver, making them interesting for site-specific drug delivery[26]. Furthermore, alteration of the sugar residue of the same compound often affects its pharmacological properties: quercetin 3-*O*-glucoside and quercetin 3-*O*-glucuronide have antioxidant properties[27,28], whereas their 3-*O*-galactoside, rhamnoside or xyloside exhibit anti-inflammatory[29], antiviral[30] or anticarcinogenic[31] activities, respectively. In addition, the type of sugar moiety of quercetin glycosides has been proven to be the major determinant in their small intestinal absorption.[32]

The majority of glycosylation reactions in nature are mediated by glycosyltransferases, which can transfer the sugar residue from an activated sugar donor to various acceptors.[33] These *gatekeepers of glycodiversity* use nucleotide sugars in 90 % of cases as activated donors and are often referred to as Leloir glycosyltransferases.[34] The uridine diphosphate (UDP) sugars form the largest group of nucleotide sugars[35] and consequently give rise to the large class of uridine glycosyltransferases (UGTs), which are characterized by a unique carboxy-terminal consensus sequence[36]. Predominant UDP-sugars in nature for the glycosylation of specialized metabolites are UDP-glucose, UDP-glucuronate, UDP-xylose, UDP-arabinose and UDP-rhamnose, while UDP-galactose and UDP-*N*-acetylglucosamine mostly occur in the formation of oligosaccharides. These UDP-sugars and corresponding UGTs are capable of efficiently glycosylating various compounds from diverse chemical classes in a regio- and stereoselective way.[13,37] Some examples of the glycosylation potential and versatility of UGTs towards small molecules are schematically shown in Figure 2.1, together with their applications.

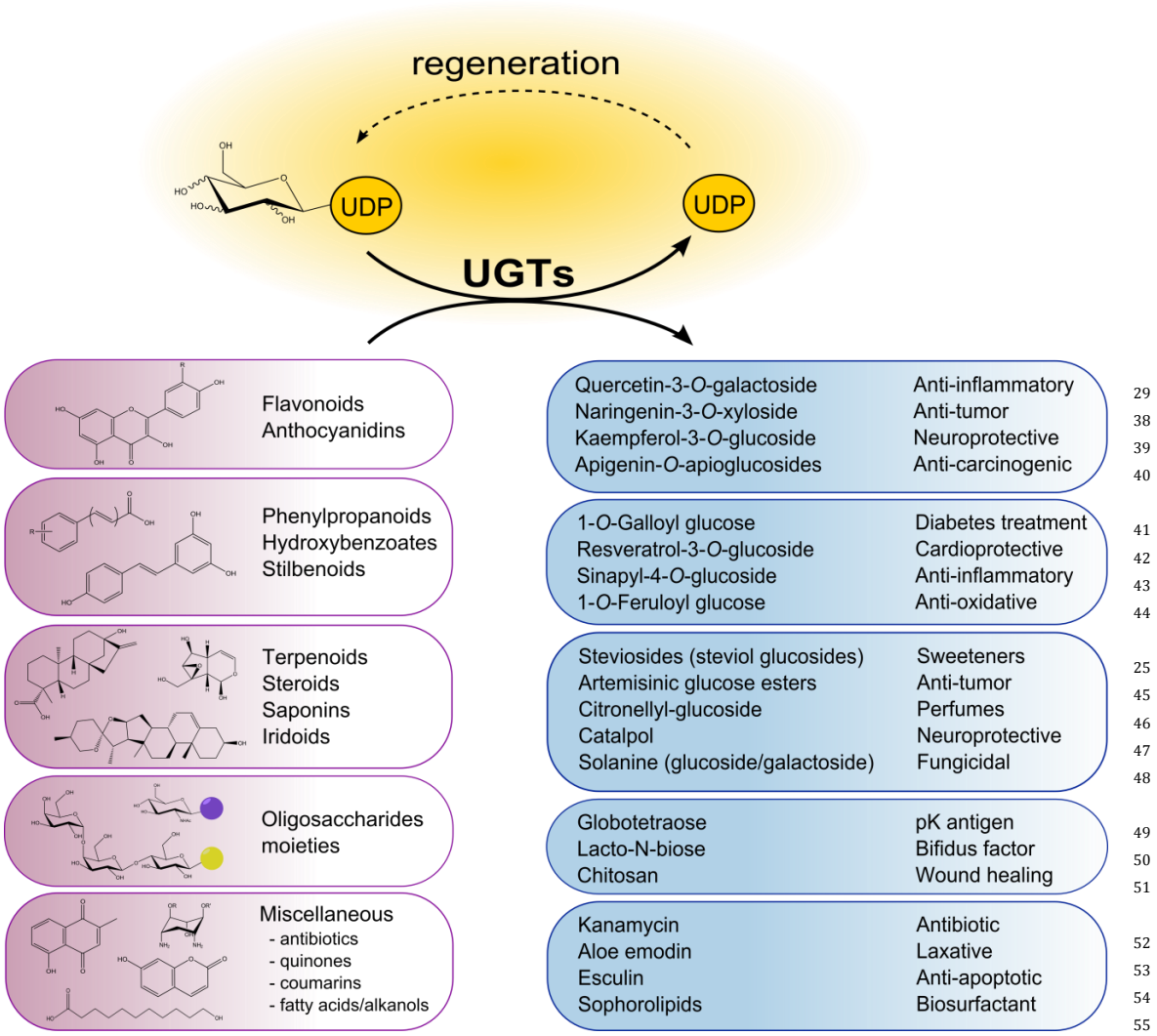


Figure 2.1: Schematic representation of the glycosylation of diverse classes of small molecules by uridine glycosyltransferases (UGTs) resulting in glycosides with various applications. The UDP-sugar is used as an activated donor for the regio- and stereoselective glycosylation, hereby releasing UDP which can be recycled via different routes further discussed in this chapter.

The microbial production of these small molecules has made an enormous progress in the last two decades.[56] This was largely stimulated by the increasing market demand and by a mentality shift in replacing chemically synthetized molecules by 'natural' ones. In addition, due to vast metabolic engineering[57] efforts in model organisms like *Escherichia coli*[58,59] and *Saccharomyces cerevisiae*[60], the sustainable production of flavonoids[61,62], anthocyanidins[63], terpenoids[61,62], phenolic acids[62], coumarins[64] and phenylpropanoids[65] has been made possible. Since their corresponding glycosides display attractive properties (Figure 2.1), many methodologies have been developed for the production thereof, which will be critically revised and compared in this review.

2.3 State of the art glycosylation processes

The seemingly easiest way to obtain naturally occurring glycosides is the extraction from native producing hosts (which are mostly plants). However, this process is often a low-yielding, laborious task and requires a specific method for each host. In addition, the yield can be highly dependent on geographical, seasonal or even political factors. Nevertheless, extraction remains to date a widely used technique, especially for obtaining anthocyanins[66], triterpenoid glycosides[67], polyphenol glycosides[68] and steviosides[69]. Seasonal factors and low yields can be tackled by using *in vitro* plant cell cultures and metabolic engineering thereof, thus guaranteeing or improving production.[70-72] However, cell cultures often suffer from economic issues and their use is currently restricted towards relatively highly priced metabolites.[73]

Alternatively, organic synthesis of natural glycosides remains a daunting task, even though it has become a sport for many chemists to synthesize every molecule imaginable. The stereospecific formation of glycosidic linkages is frequently compromised by the presence of many reactive groups[74], which results in the need for various protecting and deprotecting steps[75]. Furthermore, a low atom-efficiency, the generation of toxic waste and use of very expensive catalysts[76,77] make this method unfit for large-scale and economically viable production.

Hence, fuelled by a high demand and sustainability issues, numerous biotechnological solutions have been established. Most of them can be divided in two main categories: *in vitro* (enzymatic) and *in vivo* (bioconversion and fermentation) production. These aim at converting an acceptor (aglycon) into its corresponding glycoside (glycon), hereby only differing in the localization of the enzyme(s) involved and the metabolic state of the host. In general four types of carbohydrate-active enzymes are used for *in vitro* glycosylation: glycoside hydrolases (GHs), transglycosidases (TGs), glycoside phosphorylases (GPs) and (Leloir) glycosyltransferases (GTs).[78] These enzymes have

been classified in the Carbohydrate-Active Enzymes database (CAZy), which provides updated access to more accurate sequence-based family classification linking the sequence to the specificity and 3D structure of the enzymes.[79] Each of them have their own characteristics and drawbacks concerning substrate usage, yields and scale-up. GHs are naturally hydrolytic enzymes but can use various molecules as acceptor instead of water under controlled conditions. Their activity in the synthesis direction can be often boosted by using high substrate concentrations or adding an organic solvent to lower the water content, as for example in some industrially promising processes: production of rutin and various phenolic acid rhamnosides by an α-L-rhamnosidase from *Aspergillus terreus*[80] and large-scale production of prebiotic galactooligosaccharides (GOS) with different galactosidases[81,82]. TGs on the other hand are retaining GHs which do not use water as an acceptor. They are widely used since they can transfer the glucosyl moiety from cheap carbon sources such as sucrose towards other carbohydrates or small molecules. Interesting examples of multigram synthesis are the production of ampelopsin glucosides with a dextransucrase from *Leuconostoc mesenteroides*[83] or the biosynthesis of catechin glycosides with an amylosucrase from *Deinoccocus geothermalis*[84]. The less well-studied GPs require more expensive glycosyl phosphate donors and have been proven to be efficient catalysts for glycoside synthesis.[85,86] Well known examples are the glycosylation with sucrose phosphorylase of vitamin C[87], stilbenoids[88], hydroxybenzoates[89] and catechins[90]. In spite of their broad activity, the affinity of GHs, TGs and GPs towards new acceptors is generally very low and consequently require high concentrations of these (expensive) aglycons[91]. In addition, high substrate concentrations are evenly necessary to combat their reverse (and preferred) reaction, making the process overall often very costly. Due to their low conversion yield, restriction to mainly form *O*-glycosides, and hydrolytic/phosphorolytic nature, they are generally of no practical use when applied in *in vivo* systems.

The last type of carbohydrate-active enzymes are the GTs – with the aforementioned UGTs as largest group – and display superior conversion efficiencies (up to 100 %) towards an enormous variety of small molecules (Figure 2.1). They are able to stereo- and regiospecifically form *O*-, *N*-, *S*- and even *C*-glycosides[37] hereby requiring low acceptor concentrations. Furthermore, they can be readily applied in both *in vitro* and *in vivo* glycosylation systems. However, the main constraint to their use in an *in vitro* context is the need for UDP-sugars, which are extremely expensive and rarely available in large quantities.[92] One way to overcome this is to overexpress UGTs in a micro-organism, thus making use of their intracellular UDP-sugar pool. This methodology is the basis for *in vivo* UDP-sugar based glycosylation and eliminates the need for extensive enzyme purification and the addition of expensive cofactors. Figure 2.2 depicts a

schematic comparison of *in vitro* (A) and various *in vivo* (B to D) approaches. The latter can be divided in three major types depending on the addition of the aglycon and the metabolic state of the host(s). One of the first techniques developed was bacterial coupling (B), which basically consists of combining different hosts, each fulfilling a specific step in the formation of the glycosides. Typically, *Corynebacterium glutamicum* is used for the formation of UTP starting from the cheap precursor orotic acid (step 1). UTP is subsequently used by a recombinant strain (often *E. coli*) for the production of the UDP-sugar (step 2) and finally consumed by another *E. coli* host performing the glycosylation by expressing a UGT (step 3).[93] Such a system has been frequently used for large-scale production of oligosaccharides and is characterized by high product titers and rates.[93,94] However, the need for cell permeabilization and separate fermentations to obtain high cell densities of the various hosts involved, makes the overall process instable and less cost effective. The development of processes where these individual steps take place in one organism was, hence, a next logical step.

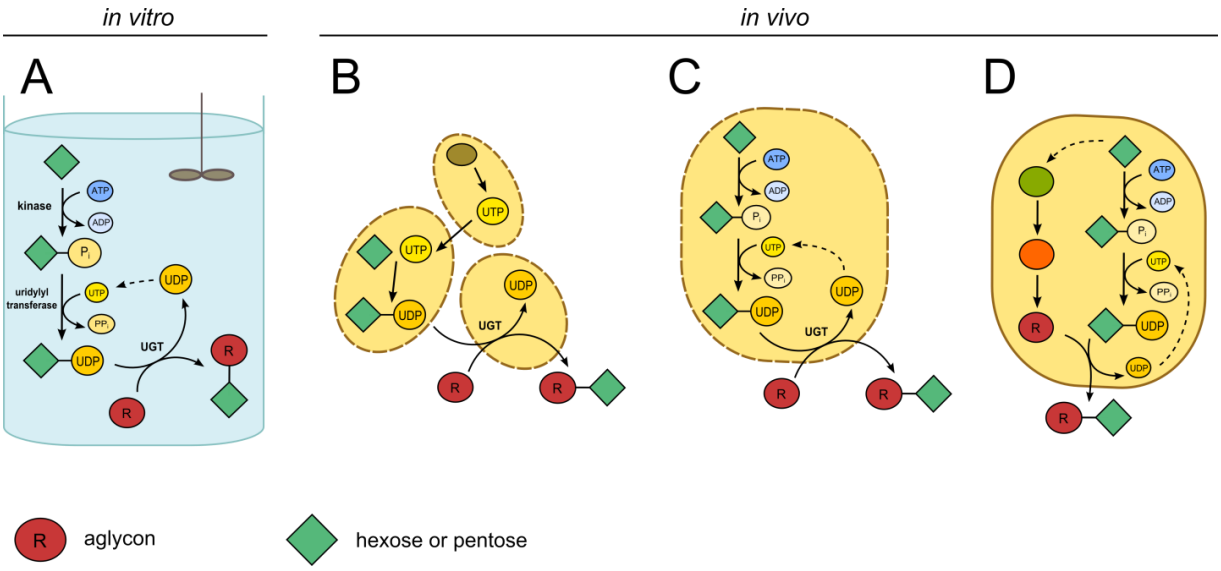


Figure 2.2: Schematic representation of different UGT based glycosylation mechanisms. Based on the location of the enzymes involved, *in vitro* (A) and *in vivo* (B,C,D) processes can be identified. While *in vitro* glycosylation requires purified enzymes and expensive cofactors, these are generated by the host itself from cheap substrates during *in vivo* production. The evolution of different types of *in vivo* glycosylation is also depicted. Bacterial coupling (B) combines different permeabilized (dashed line) hosts, each carrying out a step towards glycosylation of an added acceptor. A single cell glycosylation (C) process on the other hand combines all these steps in one host, which may be permeabilized to facilitate conversion of the acceptor. Cells that actively grow and glycosylate simultaneously are designated as a fermentation, resting cells as a bioconversion. Finally, the recently developed *de novo* fermentation (D) processes produces both the acceptor and subsequently its glycoconjugate from a cheap carbon source, thus eliminating the need for external addition of the acceptor.

The advances in the fields of metabolic and genetic engineering made it possible to merge various different pathway steps inside a single cell. This approach can be as effective as combining multiple hosts, provided that the released UDP can be efficiently regenerated. The major advantages of using a single cell are the need for only one fermentation and the use of a cheap carbon source like glucose or glycerol as only precursor for UDP-sugars. Depending on the metabolic state of the cell, a bioconversion (resting cells) or fermentation (actively growing cells) approach can be discerned. Bioconversion often requires permeabilization of the host and high initial cell densities to ensure proper conversion. In addition, the production rate frequently decreases in time due to cell decay, in contrast to actively growing cells where the reaction speeds up. The potential of producing glycosides with a fermentation based system and the recent expansion of metabolic engineering towards microbial production of many small molecules gave rise to a final and novel approach: *de novo* fermentation of glycosides (Figure 2.2 D). This system eliminates the need for addition of expensive acceptors such as specialized metabolites and effectively couples growth and production. Few examples exist to date with only vanillin glucoside[95,96], kaempferol 3-*O*-rhamnoside[97], resveratrol glucosides[98], anthocyanins[99] and steviol glycosides[100] being reported in the traditional hosts *E. coli* and *S. cerevisiae*.

A comprehensive overview of all *in vivo* systems for UDP-sugar based glycosylation focussing on small molecules like specialized metabolites and oligosaccharides is given in Table 2.1. To give a fair comparison of the processes and to make an estimation of their economic viability, the production rate, conversion efficiency, titer and medium composition is given for each process. Remarkably, UGTs have only been successfully applied for the *in vivo* production of oligosaccharides in terms of rates and titers (up to 188 g/L), while thus far glycosides derived from specialized metabolites are produced about three orders of magnitude lower. One explanation is the low solubility and toxicity of some of these metabolites, thus hampering efficient formation of the corresponding glycosides. On the other hand, it is striking how many *in vivo* approaches primarily concentrate on the expression of novel UGTs or medium optimization, while generally little effort has been put into the efficient generation and regeneration of UDP-sugars. In contrast, engineering strategies for oligosaccharide production have primarily focused on building an efficient UDP-sugar synthesizing module hereby exploiting the whole range of UDP-sugar formation techniques. The different methodologies and their applications in each system (*in vitro*/*in vitro*) are discussed in the next section.

Table 2.1: Overview of all existing *in vivo* processes for the glycosylation of natural small molecules using overexpressed UGTs.

Glycoconjugate	Process	Hosts	Culture conditions; Medium; Cofactors	Acceptor	Phases [a]	Conv [b]	r_{gc} (g/L/h)	Titer (g/L)	Ref
Flavonoids									
Quercetin 3-*O*-rhamnoside	bioconv	EC	High initial cell density; M9, Glc, def	quercetin	Gr → Pr	0.84	0.003	0.15	[101]
Quercetin 3-*O*-glucoside	bioconv	EC	LB-medium	quercetin	GrPr	0.03	0.005	0.10	[102]
Quercetin 3-*O*-*N*-acetylglucosamine	bioconv/ ferm	EC	High initial cell density; M9, Glc, def	quercetin	Gr → GrPr	0.84	0.015	0.38	[103]
Quercetin 3-*O*-xyloside	bioconv	EC	High initial cell density; TB-medium, Mann, Glc	quercetin	Gr → GrPr	0.97	0.001	0.04	[104]
Quercetin 3-*O*-xyloside	bioconv	EC	High initial cell density; M9, Glc, def	quercetin	Gr → Pr	0.86	0.005	0.15	[105]
Quercetin 3-*O*-arabinoside	bioconv	EC	High initial cell density; M9, Glc, def	quercetin	Gr → Pr	0.91	0.002	0.16	[105]
Quercetin 3,7-*O*-bisrhamnoside	bioconv	EC	High initial cell density; M9, Glc, def	quercetin	Gr → Pr	1.13 [d]	0.003	0.07	[106]
Apigenin 7-*O*-glucoside	ferm	EC	TB-medium, Glc	apigenin	GrPr	0.91	0.001	0.039	[107]
Baicalein 7-*O*-glucoside	ferm	EC	TB-medium, Glc	baicalein	GrPr	0.77	0.001	0.033	[107]
Biochanin A 7-*O*-glucoside [f]	bioconv	EC	Unknown cell density; TB-medium, Mann, Glc, Fru, Glyc	biochanin A	Gr → Pr	0.75	0.001	0.067	[108]
Naringenin 7-*O*-glucoside	ferm	SC	Modified SGI medium; 10 mM oro	naringenin	GrPr	0.71	0.002	0.155	[109]
Kaempferol 3-*O*-glucoside	bioconv	EC	Unknown cell density; TB-medium, Manno, Glc, Fru, Glyc	naringenin	Gr → Pr	0.49	0.002	0.109	[110]
Kaempferol 3-*O*-rhamnoside	ferm *de novo*	EC	M9, Glc, YE, def	-	GrPr	-	0.002	0.057	[97]
Stilbenoids									
Resveratrol 3-*O*-glucoside	ferm *de novo*	EC	M9, Glc, $CaCO_3$, def	-	GrPr	-	0.0001	0.003	[98]
Resveratrol 4'-*O*-glucoside	ferm *de novo*	EC	M9, Glc, $CaCO_3$, def	-	GrPr	-	0.0002	0.007	[98]
Chalcones									
Phloretin 2',4',4'-*O*-triglucoside [e]	bioconv	EC	High initial cell density; TB-medium	phloretin	Gr → Pr	0.11	0.003	0.163	[111]
Phloretin *C*-glucoside	ferm	SC	YPGE medium, Raff, def; 10 mM oro	phloretin	GrPr	0.79	0.002	0.052	[112]

Glycoconjugate	Process	Hosts	Culture conditions; Medium; Cofactors	Precursor	Phases [a]	Conv [b]	r_{gc} (g/L/h)	Titer (g/L)	Ref
Anthocyanidins									
Cyanidin 3-*O*-Glucoside	ferm *de novo*	EC	Modified M9, def; 0.1 mM oro	catechin	Gr → GrPr	-	ND	0.104	99
Pelargonidin 3-*O*-Glucoside	ferm *de novo*	EC	Modified M9, def; 0.1 mM oro	afzelechin	Gr → GrPr	-	ND	0.113	99
Phenolic compounds									
Vanillin 4-*O*-glucoside	ferm *de novo*	SC/SP	YES medium	-	GrPr	0.56 [g]	0.002	0.12	96
Vanillin 4-*O*-glucoside	ferm *de novo*	SC	Verduyn medium, Glc, def	-	GrPr	-	0.008	0.50	95
Vanillin 4-*O*-glucoside	ferm *de novo*	SC	SC medium, Glc, Scr, def	-	GrPr	-	0.006	0.45	113
Arbutin	ferm	EC	LB-medium	hydroquinone	GrPr	0.51	0.007	0.25	114
Terpenoids									
Rebaudioside A	ferm *de novo*	SC	SC medium w/o methionine, Glc	-	GrPr	-	0.0002	0.015	100
Plant hormones									
(+)-abscisic acid glucose ester	bioconv	EC	Sodium phosphate buffer, Glc, def	(±)-AA	Gr → Pr	0.20	0.004	0.085	115
Oligosaccharides									
Globotriose	BC - perm	CG+EC	250 g/L cells [c]; Glc, Fru, Gal, def; 50 mM oro	lactose	Gr → Pr	0.79	5.22	188	93
Globotriose	ferm	EC	High initial cell density; Glyc, def	lactose	Gr → GrPr	0.96	1.41	7.1	49
Globotriose	bioconv - perm	EC	200 g/L cells [c]; Scr, def; 2 mM UDP	lactose	Gr → Pr	0.22	0.61	22	116
Globotriose	BC	EC	Unknown cell amount; Tre; 7.4 mM UDP	lactose	Gr → Pr	0.05	0.16	2	117
Globotetraose	ferm	EC	High initial cell density; Glyc, def	lactose	Gr → GrPr	0.44	0.115	4.6	49
α-Gal epitope	bioconv - perm	EC	200 g/L cells [c]; Scr, def; 2 mM UDP	lactose	Gr → Pr	0.18	0.50	18	116
α-Gal epitope	bioconv - perm	PP	250 g/L cells [c]; Scr, def; 2 mM UDP	lactose	Gr → Pr	0.37	0.353	14.1	118
α-Gal epitope	bioconv	A	100 g/L cells [c]; Scr, def	lactose	Gr → Pr	0.03	0.006	1	119
α-Gal epitope	bioconv - perm	EC	130 g/L cells [c]; Gal, Glc, def; 2 mM Glc1P, 0.5 mM UDP-glc, 2 mM ATP	lactose	Gr → Pr	0.57	0.36	7.16	120
α-Gal epitope	BC	EC	Unknown cell amount; Tre; 7.4 mM UDP	lactose	Gr → Pr	0.04	0.15	1.76	117

Glycoconjugate	Process	Hosts	Culture conditions; Medium; Cofactors	Precursor	Phases [a]	Conv [b]	r_{gc} (g/L/h)	Titer (g/L)	Ref
N-acetyllactosamine	bioconv	A	100 g/L cells [c]; Glc, def	GlcNAc	Gr → Pr	0.19	0.15	7.3	[121]
N-acetyllactosamine	bioconv	EC	100 g/L cells [c]; Glc, def	GlcNAc	Gr → Pr	0.16	0.20	3.14	[122]
N-acetyllactosamine	BC - perm	CG+EC	250 g/L cells [c]; Glc, Fru, Gal, def; 50 mM oro	GlcNAc	Gr → Pr	0.96	2.81	107	[94]
P1 trisaccharide	bioconv - perm	EC	100 g/L cells [c]; Scr, def; 5 mM UDP-glc	GlcNAc	Gr → Pr	0.67	0.38	27	[123]

Process: bioconv: bioconversion, ferm: fermentation, *de novo*: acceptor was not added but also microbially produced, BC: bacterial coupling, perm: permeabilized cells; **Hosts:** EC: *Escherichia coli*, SC: *Saccharomyces cerevisiae*, SP: *Schizosaccharomyces pombe*, CG: *Corynebacterium glutamicum*, PP: *Pichia pastoris*, A: *Agrobacterium sp.*; **Culture conditions/Medium/Cofactors:** M9: M9 mineral medium, Glc: glucose, def: defined medium, TB: terrific broth, Mann: mannitol, Fru: fructose, Glyc: glycerol, oro: orotic acid, Manno: mannose, YE: yeast extract, Raff: raffinose, YES: yeast extract with supplements, SC: synthetic complete, Scr: sucrose, Gal: galactose, Tre: trehalose; **Precursor:** AA: abscisic acid, GlcNAc: *N*-acetylglucosamine; **Phases:** Gr: growth phase, Pr: production phase, GrPr: simultaneous growth and production; **r_{gc}:** volumetric production rate of glycoconjugates, ND: not determined.

[a] Bioconversion and bacterial coupling require a growth phase for the accumulation of enzymes in contrast to fermentative processes often exhibiting simultaneous growth and production (GrPr).

[b] Conversion (Conv) is expressed as mM glycoconjugate formed/mM acceptor added

[c] Cells expressed as wet weight

[d] Conversion exceeds 100 % without obvious explanation

[e] The triglucoside was the major glucoside formed next to 3 other mono- and diglucosides

[f] Other isoflavonoids (genistein, formononetin and daidzein) showed lower conversions and yielded glucoside mixtures

[g] Conversion was based on mM glucoside/mM vanillin and precursors thereof formed

2.4 Formation of common UDP-sugars: sweet routes to synthesis

In contrast to *in vitro* approaches, which make use of a vast arsenal of cofactor regeneration techniques or interesting methods for UDP-sugar generation from cheap substrates, most whole cell (*in vivo*) approaches apply the same methodology. Despite their reported process variety, only three major routes exist for the formation of all natural UDP-sugars: a synthase, a phosphorylase and a kinase route. A schematic representation of these is shown in Figure 2.3 together with a UGT coupling reaction to illustrate their different cofactor requirements (ATP, UTP, …) for sugar activation and UDP recycling.

Synthase route

This first route has gained a lot of attention since it directly forms a highly energetic UDP-sugar from a disaccharide. To date, only sucrose synthase (EC 2.4.1.13) and trehalose synthase (EC 2.4.1.245) have been extensively used for the generation of UDP-glucose from sucrose[124,125] and trehalose[126,127] respectively, due to their readily reversible character. Many other synthases such as cellulose synthase or lactose synthase have been described, but their application is restricted to the synthesis reaction because of their unfavorable equilibrium constants.[128] The reported trehalose synthases originate from bacteria[126,127], while the majority of sucrose synthases used typically originate from plants[125,129,130], with a cyanobacterial homolog as only exception to date[116]. Due to the availability of sucrose and trehalose at low cost and in bulk quantities, various processes using the corresponding synthases can be very cost effective when coupled with a UGT. Moreover, coexpressing a synthase prevents accumulation of UDP, which has been shown to cause the frequently encountered problem of UGT product inhibition.[124] This coupling principle has been applied both in *in vitro* and *in vivo* processes for the production of various oligosaccharides[116,123,131] and glycosylated specialized plant metabolites[130,132]. Moreover, a fusion protein of sucrose synthase and a UGT has been created and used for the production of quercetin 3-*O*-glucoside.[133] Despite the clear potential of synthases, most reported *in vitro* coupling systems still rather illustrate only a low-yielding proof of concept. Thus far only three efficient bioconversion processes have been developed, producing globotriose (22 g/L)[116], α-gal epitope (18 g/L)[116,118] and P1 trisaccharide (27 g/L)[123] as shown in Table 2.1. The key feature for an economically viable *in vitro* process, as recently pointed out by Nidetzky and coworkers[130], is a high number of regeneration cycles (RC) for the UDP-sugar. Through controlled feeding of the substrate phloretin and repeated co-addition of enzymes, they could obtain an impressive 20 g/L of the *C*-glucoside nothofagin with an RC_{max} of 50.

A major challenge when applying synthases as catalysts is improving their stability and kinetic parameters. For example, the affinity constant (K_m) of sucrose synthase for sucrose often dramatically increases when recombinantly overexpressed in a prokaryotic host[134], easily reaching up to 100 mM[125] or more[135]. Similarly, the K_m values for UDP and trehalose are 30 and 25 mM, respectively, for trehalose synthase from *Pyrococcus horikoshii*[126]. This feature makes synthases less suitable for fermentation systems since growing cells cannot accumulate the respective disaccharides intracellularly at the required concentrations, thus impeding UDP-sugar generation. The only example to date is the coexpression of sucrose synthase from *Coffea arabica* in yeast, which resulted in a 35 % increased *de novo* vanillin 4-*O*-glucoside production.[113] Extensive (structure based) protein engineering efforts could beneficially influence these kinetic parameters[118,120] and subsequent production.

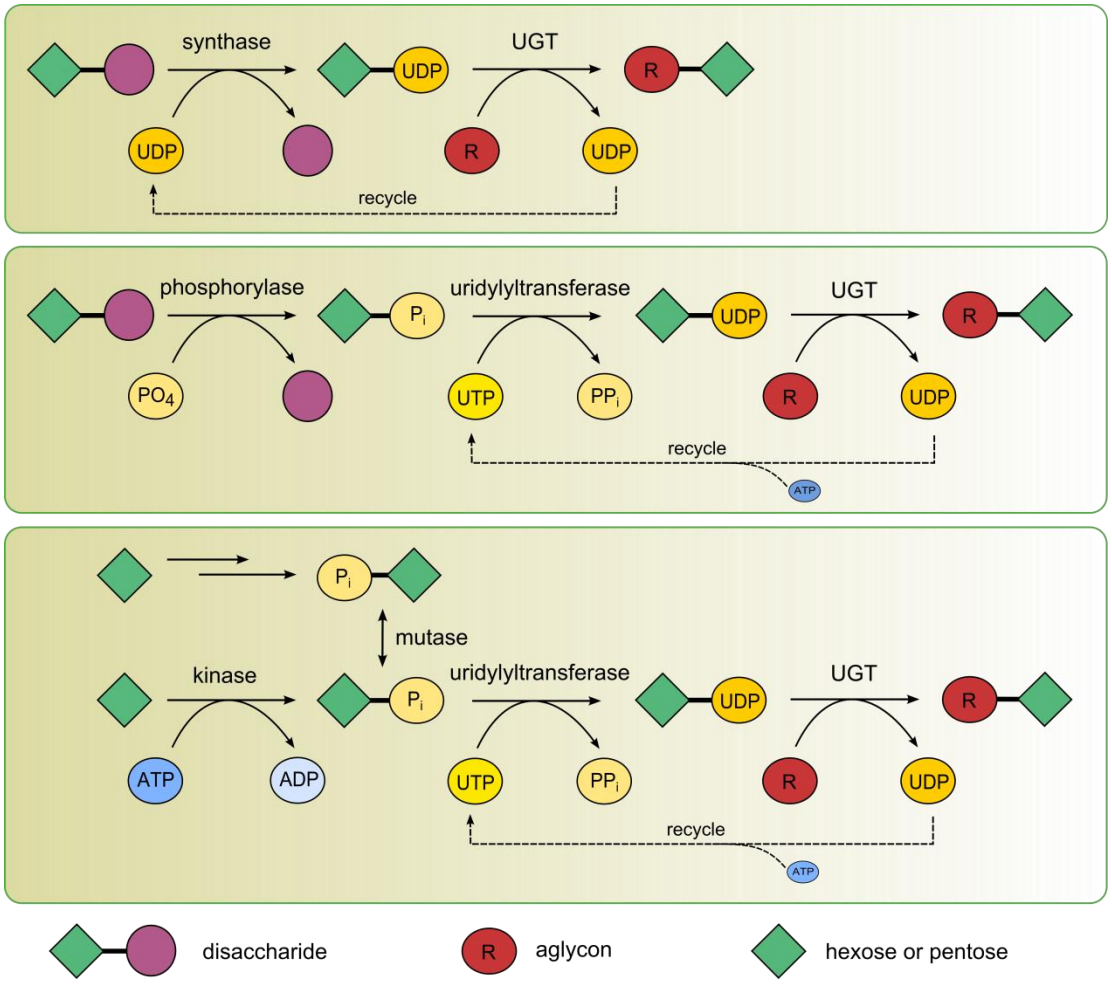


Figure 2.3: Schematic overview of the three major routes towards UDP-sugar formation, each coupled with a uridine glycosyltransferase (UGT). The synthase, phosphorylase and kinase route are shown and differ in their need for cofactors and the regeneration of UDP.

Phosphorylase route

Another less elaborate route comprises the use of glycoside phosphorylases, which have the ability to cleave disaccharides with inorganic phosphate (Pi) hereby producing an activated sugar 1-phosphate.[136] This reaction alleviates the need for ATP to generate a sugar 1-phosphate, which can be subsequently coupled with a uridylyltransferase to yield the corresponding UDP-sugar. Although many types of phosphorylases have been reported, α-glucose 1-phosphate is the most prevalent phosphate sugar[136] and can be formed through the action of a sucrose phosphorylase[137] (EC 2.4.1.7), a cellobiose phosphorylase[138] (EC 2.4.1.231) or a maltodextrin phosphorylase[139,140] (EC 2.4.1.1) from their respective substrates. Furthermore, the cellobiose phosphorylase from *Cellulomonas uda* was engineered by directed evolution to result in a novel lactose phosphorylase enzyme that can efficiently generate α-galactose 1-phosphate from the cheap disaccharide lactose.[141]

Although the only documented example of the phosphorylase route is the *in vitro* one pot synthesis of *N*-acetyllactosamine[142] using sucrose phosphorylase, this methodology holds great promise for *in vivo* glycosylation. In contrast to sucrose synthase, overexpression of sucrose phosphorylase can enable growth[143] on sucrose as only carbon source due to its high affinity[144] for this substrate (K_m ≈ 1 mM). Metabolic engineering of the host cell can direct the increased intracellular glucose 1-phosphate pool completely towards UDP-glucose, thus creating an efficient strain for *in vivo* glucosylation reactions. A first attempt based on this principle was described by Heinzle and coworkers[145], who created an engineered *E. coli* strain expressing a sucrose phosphorylase from *L. mesenteroides* which efficiently produced UDP-glucose after permeabilization. Interestingly, the use of disaccharides as starting substrate for UDP-sugars (phosphorylase and synthase route) generates a monosaccharide 'byproduct' (often fructose) which can be used as additional carbon source, thus unlocking the potential to simultaneously grow and produce glycosides using only a cheap substrate like sucrose.

Kinase route

The last route discussed is probably the most widely used both in *in vitro* and *in vivo* systems and is based on the combined action of a kinase and a uridylyltransferase to generate the required UDP-sugar (Figure 2.3). Many ATP dependent kinases have been discovered with specific or promiscuous activities towards a variety of monosaccharides for the generation of sugar 1-phosphates. The galactokinase (GalK, EC 2.7.1.6) from *E. coli* and *Streptococcus pneumoniae* naturally produces galactose 1-phosphate, but also

has activity towards other galactose derivatives[146,147] and glucose[147]. Furthermore, GalK has been the subject of site-directed and random mutagenesis to further expand the activity spectrum and has been extensively used for the phosphorylation of galacturonic acid, galactosamine, talose, and various L-sugars such as L-glucose.[148-150] Another kinase which gained a lot of attention is the *N*-acetylhexosamine kinase (NahK, EC 2.7.1.162) from *Bifidobacterium longum*, which predominantly acts on *N*-acetylglucosamine and *N*-acetylgalactosamine[151], but it also shows activity towards other hexosamine derivatives[152-154]. The latter kinases are mainly used for the *in vitro* synthesis of UDP-sugars, whereas most *in vivo* routes naturally generate a sugar 6-phosphate pool which can be easily interconverted through a phosphosugar mutase as depicted in Figure 2.3. The best documented example is phosphoglucomutase (Pgm) which forms glucose 1-phosphate from glucose 6-phosphate. The formed sugar 1-phosphates are the substrates for specific or promiscuous sugar 1-phosphate uridylyltransferases and this process functions as salvage pathway for UDP-sugar formation in many organisms. Regarding *in vivo* systems, GalU (EC 2.7.7.9) and GlmU (EC 2.7.7.23) have been frequently used for the formation of UDP-glucose and UDP-*N*-acetylglucosamine respectively[103], while *in vitro* synthesis more often relies on uridylyltransferases with broad activity. These promiscuous uridylyltransferases are often referred to as SLOPPY (EC 2.7.7.64)[155] and have been cloned from various organisms like *B. bifidum*[156], *B. longum*[146], *H. sapiens*[157], *P. sativum*[158] or the extremophile *P. furiosus*[159] for the formation of UDP-sugars derived from galactose, arabinose, xylose, glucuronic and galacturonic acid and *N*-acetylgalactosamine.

The kinase-uridylyltransferase route has been frequently applied both in *in vitro* and *in vivo* systems coupled with a UGT for the production of various glycosides. In spite of the need for expensive cofactors (ATP and UTP) during the *in vitro* formation of UDP-sugars, these (one pot) multi-enzymatic conversion processes have been repeatedly applied for multigram-scale production of oligosaccharides and derivatives such as globotriose[160], *N*-acetyllactosamine oligomers[161] and T-MUC1 glycopeptide[162]. To reduce the high costs, efficient systems for the regeneration of the UDP-sugar have been developed based on the coupling with pyruvate kinase.[163,164] This versatile enzyme catalyzes the transfer of a phosphate group from PEP to a nucleoside diphosphate (NDP), yielding pyruvate and the corresponding nucleoside triphosphate (NTP). This system is of particular interest since it eliminates product inhibition by accumulating UDP. In the case of these tailor made oligosaccharides, their market price justifies the costs of cofactors, regeneration systems and large-scale enzyme purification.

Taking everything into account, *in vitro* application of the kinase route is only advantageous for rapid, small-scale production of (novel) glycosides or UDP-sugars. In response to this fact, many efficient *in vivo* production processes have been developed, mostly relying on bacterial coupling or bioconversion while genuine (*de novo*) fermentation approaches are rather uncommon. In contrast to bacterial coupling and bioconversion where the kinase and uridylyltransferase are actively overexpressed, few engineering efforts have been made regarding UDP-sugar formation in the hosts for fermentation based glycosylation. As pointed out by Chen and coworkers[122], glucose 1-phosphate formation (mediated by Pgm) is a major bottleneck and overexpression of Pgm and GalU can result in a 10-fold higher glycoside production. Furthermore, fermentation systems could benefit from an improved recycling of the released UDP towards UTP, which is often built-in in a bacterial coupling[93] or bioconversion process[120]. A nucleotide diphosphate kinase (Ndk) naturally catalyzes this reaction and overexpression of this enzyme has been shown to enhance the formation of cyanidin 3-*O*-glucoside in growing *E. coli* cells[99].

2.5 Interconverting common UDP-sugars

In principle, production of a required UDP-sugar can be achieved by using one of the three routes discussed in Figure 2.3. The most common UDP-sugars are however formed through the kinase route since no synthase or phosphorylase alternatives exist for their generation. This in contrast to UDP-glucose, which can be rapidly generated via all three routes from various substrates, thus acting as a central pivot and starting point for the production of a large variety of UDP-sugars. These can be elegantly formed through the (combined) action of a large variety of epimerases, dehydrogenases or decarboxylases as shown in Figure 2.4. While epimerases can be used for both *in vivo* and *in vitro* systems, dehydrogenases and decarboxylases are (economically) restricted to *in vivo* processes due to the need for expensive cofactors like $NAD(P)^+$.

The first UDP-sugars that can be generated from UDP-glucose are UDP-galactose, UDP-glucuronate and UDP-rhamnose. The formation of UDP-galactose by a UDP-glucose 4-epimerase (GalE) is by far the best studied epimerization reaction which has been exploited in many *in vitro* or *in vivo* processes[116,119,123]. Despite its widespread use, it is often identified as the rate-limiting step in enzymatic systems so that the (galacto)kinase route is a preferred alternative.[164,165] In most cases, GalE from *E. coli* is used due to its broad pH range, general stability and high specific activity[166], with the highly thermostable GalE from *P. horikoshii* as a notable exception[167].

A less known multifunctional and interesting enzyme is rhamnose synthase (RHM), which catalyzes the conversion from UDP-glucose to UDP-rhamnose. RHM is only present in plants and the N-terminus of the enzyme carries out a dehydratase reaction while the C-terminus is responsible for a sequential epimerase and reductase reaction.[168] Recently, expression of RHM from *A. thaliana* in *E. coli* resulted in an efficient bioconversion process for the production of quercetin 3-*O*-rhamnoside and kaempferol 3-*O*-rhamnoside.[101] In bacteria, rhamnosylation occurs via dTDP-rhamnose, requiring three enzymes (RmlBCD) for its generation starting from dTDP-glucose[169] which often yields lower titers[170]. The UDP-rhamnose pathway thus proves to be a promising alternative for *in vivo* rhamnosylation and can be efficiently exploited for example when coupled with sucrose synthase.

Finally, UDP-glucuronate can be formed by a dehydrogenase enzyme Ugd, using two NAD^+ molecules which are activated by a tyrosine kinase.[171] Expression of UGD from *A. thaliana* in yeast has shown to elevate the UDP-glucuronate pool[172], whereas a fission yeast strain was metabolically engineered for the production of various (sterol) glucuronides.[173] Besides the metabolic importance of glucuronidation and the central role of UDP-glucuronate as a major constituent of glycosaminoglycans (GAG)[174], it is also a hub towards the formation of UDP-galacturonate, UDP-apiose and UDP-xylose, which itself acts as an intermediate for UDP-arabinose formation.

Decarboxylation of UDP-glucuronate results in the formation of UDP-apiose and UDP-xylose in a 2 to 1 ratio through the plant enzyme UDP-apiose/UDP-xylose synthase (AXS).[175] Depending on whether the decarboxylated intermediate undergoes a reduction or a ring rearrangement, UDP-xylose or UDP-apiose are formed, respectively, with the latter being highly unstable.[176] Recently, bacterial UDP-xylose synthases (Uxs) have also been characterized from *Sinorhizobium meliloti*[177], *Bacteroides fragilis*[178] and *Micromonospora echinospora*[104]. The latter Uxs and AXS from *A. thaliana* have both been successfully expressed in *E. coli* for the *in vivo* production of quercetin 3-*O*-xyloside.[104,105] Coupling UDP-xylose formation with a UDP-xylose 4-epimerase (UXE) efficiently generates UDP-arabinose and similarly produces quercetin 3-*O*-arabinoside when expressed in *E. coli*. This enzyme is present in green plants but has recently been cloned from *S. meliloti*[177]. Finally, UDP-glucuronate can be easily epimerized into UDP-galacturonate with a UDP-glucuronate 4-epimerase (Gla), an enzyme very similar to GalE, which is present in plants but has also been discovered in *Klebsiella pneumonia*[179] and *S. pneumoniae*[180].

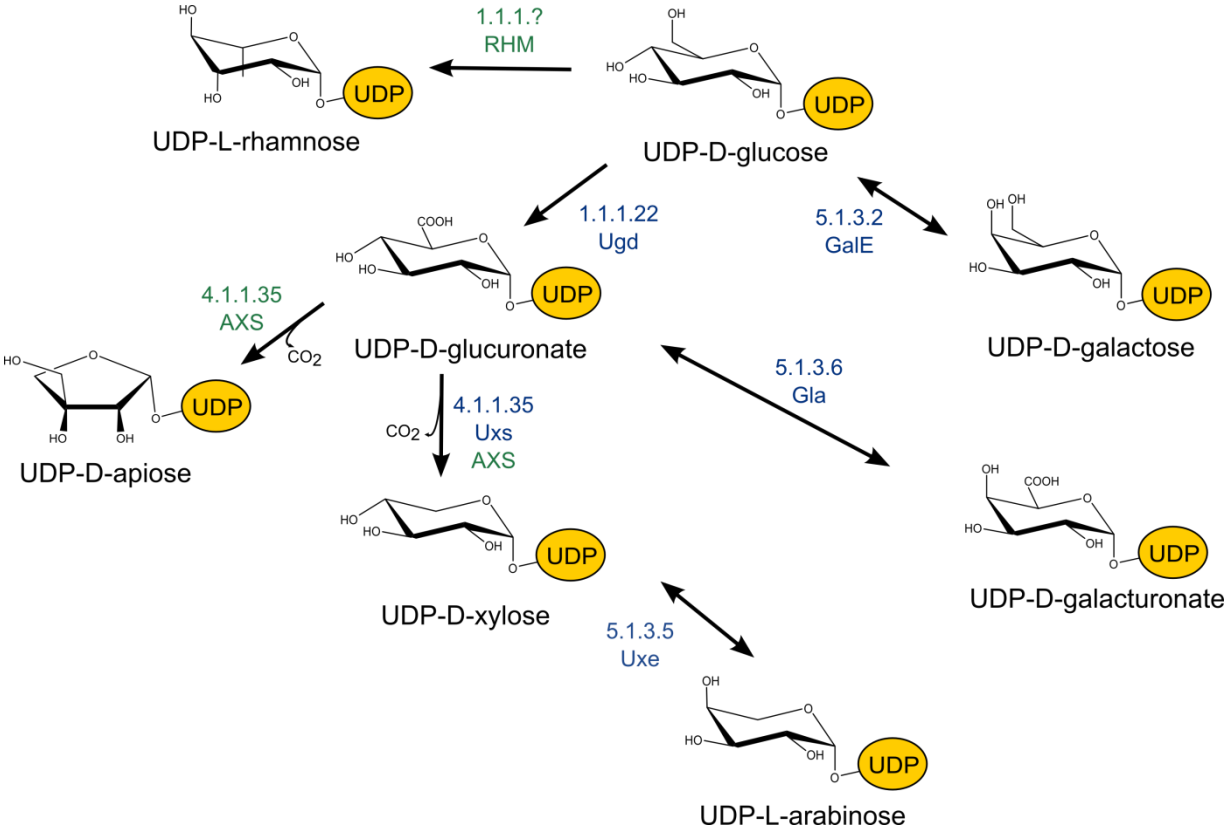


Figure 2.4: Overview of the interconverting enzymes for the generation of the most common UDP-sugars found in nature derived from the central molecule UDP-glucose. EC numbers and the abbreviated gene products are shown next to the reaction. Names colored in green are only present in plants, names in blue are (also) present in bacteria.

Next to UDP-glucose, UDP-*N*-acetylglucosamine is another core metabolite which can be easily formed *in vivo* and *in vitro*. It is the building block for chitin, various GAG[174] and is omnipresent in the sugar moiety of many glyconjugates such as 2-deoxystreptamine-containing aminoglycosides[52]. As shown in Figure 2.5, UDP-*N*-acetylglucosamine can be interconverted into UDP-*N*-acetylgalactosamine by a UDP-*N*-acetylglucosamine 4-epimerase (GalE2). This bacterial enzyme has been used for the enzymatic production of the oligosaccharides globotetraose and isoglobotetraose[181] and resulted in the *in vivo* production of globotetraose when overexpressed in an *E. coli* mutant[49]. Both UDP-*N*-acetylglycosamines can also be converted into their corresponding glycosaminuronate by the dehydrogenases WbpO[182] and WbpA[183]. Although these enzymes have thus far not been used for the production of rare UDP-sugars, they could become increasingly popular since glycoconjugates containing a carboxyl moiety are ligands to the activation receptors of human natural killer cells[184].

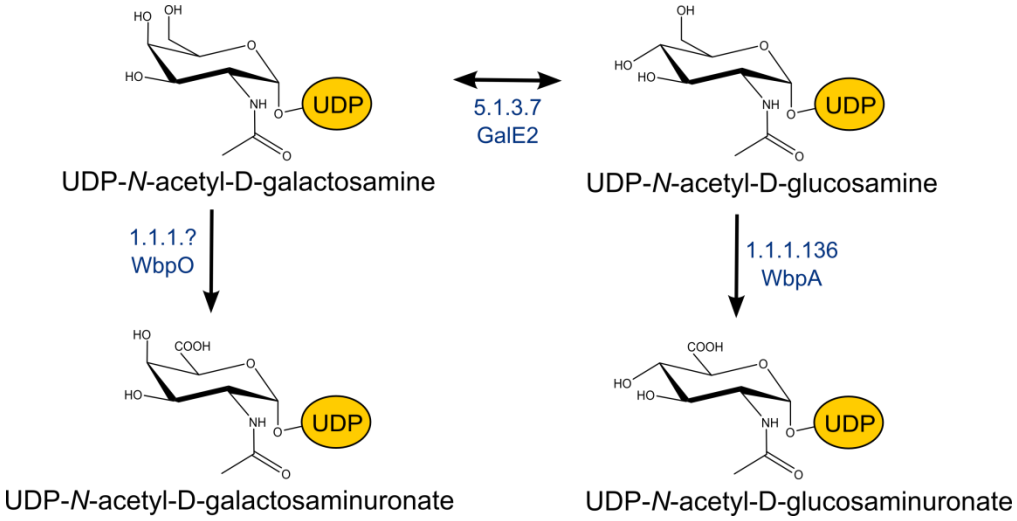


Figure 2.5: Overview of the interconverting enzymes for the generation of the most common UDP-sugars found in nature derived from the central molecule UDP-*N*-acetylglucosamine. EC numbers and the abbreviated gene products are shown next to the reaction. Names in blue are present in bacteria.

2.6 Glycorandomization: beyond nature's glycosylation toolbox

As illustrated in the previous sections, UGTs have the capability to glycosylate a myriad of acceptors, hereby using a large variety of UDP-sugars. Due to the importance of specialized metabolites as major drug leads and the profound effects of adding a sugar residue to these compounds, there is a growing interest in unusual or new-to-nature glycosides. The latter form the basis for the next generation glycosylated therapeutics[185] by combining new or unusual sugars with chemically synthesized aglycon libraries. Producing such glycosides with glycosyltransferases results in two key challenges: the formation of exotic UDP-sugars and the search for or engineering of promiscuous UGTs.

Unusual or new UDP-sugars can be formed by either exploiting the diversity provided by nature through the use of rare UDP-sugar modifying enzymes or by exploiting the chemical synthesis potential to produce new and diverse sugar libraries (Figure 2.6).[186] Such unusual NDP-sugar modifying enzymes are mainly found in (soil) bacteria, which mostly produce exotic glycosides for communication or for their never-ceasing chemical warfare. While the arsenal of unusual UDP-sugars may seem overwhelming, the most common starting point are UDP-4-keto-6-deoxy-sugars which are formed by a 4,6-dehydratase from UDP-glucose or UDP-*N*-acetylglucosamine.[187,188] These deoxysugars may subsequently undergo eight different types of modifications: epimerization, isomerization, transamination, group-transfer (formylation, acetylation, sulfation, methylation,…), reduction and oxidation, amino oxidation, pyranose/furanose

interconversion and dehydration reactions, whereby the underlying mechanisms and types of NDP-sugars formed have been excellently reviewed.[187,189,190]

These useful classes of modifying enzymes have only scarcely been used to create *in vitro* diversity due to the need for equimolar amounts of expensive cofactors like NAD(P)H or *S*-adenosyl methionine.[191,192] Notwithstanding the advances in enzymatic synthesis, these processes are often plagued by low yields, which has greatly stimulated an *in vivo* approach. Generally, the cluster coding for the unusual sugar biosynthesis is altered or expanded in the native host or introduced in a recombinant host by means of metabolic engineering.[193,194] Interesting examples thereof are the formation of a UDP–4-amino-4-deoxy-L-arabinose (UDP-Ara4FN)[195] and UDP–2-acetamido-2-deoxy-xylose (UDP–XylNAc)[196] pool in *E. coli* or the expansion of the natural deoxysugar metabolism of various *Streptomyces* species towards 4-epi-L-daunosamine[197] and 4-deacetyl-L-chromose B[198] decorated anthracyclines. It should however be noted that most of the unusual deoxysugars are dTDP-sugars, but due to the structural resemblance between (d)TDP and UDP and similarity between the mechanisms of their modifying enzymes, they can easily be interchanged.[199] Furthermore, the existing enzyme machinery will probably be greatly expanded or become more specific due to recent advances in the field of (functional) metagenomics which gives access to the enormous biodiversity.[200]

Another approach to produce new-to-nature sugar (libraries) is the exploitation of the vast array of chemical methodologies. These methods make use of selective aldol reactions[201,202] or ring-closing metathesis[203] for the formation of novel monosaccharides, but commonly available sugars may also be chemically modified[152,153,204]. Activation of these sugars is achieved by coupling with promiscuous (anomeric) kinases and uridylyltransferases to yield the corresponding UDP-sugars (Figure 2.6). As pointed out in section 2.4, both GalK and NahK display broad kinase activity[147,152-154], and have been the subject of protein engineering as well, resulting in an *E. coli* Y371H/M173L GalK which literally opened the gate for unnatural sugars[205]. Regarding promiscuous uridylyltransferases, the SLOPPY enzyme from *B. longum* has proven to be an efficient *in vitro* catalyst towards an enormous range of unusual or modified sugar 1-phosphates.[146] In addition, the uridylyl/thymidylyltransferase (RmlA; EC2.7.7.24) from *Salmonella enterica* has been frequently applied for both *in vivo* and *in vitro* generation of various UDP-sugars.[150,206] Various engineering efforts (rational design and directed evolution) have improved the activity and promiscuity of RmlA towards C2 to C4 modified sugars and nucleotides.[206-208] Another remarkable example is the mutagenesis of a Cps2L thymidylyltransferase from *S. pneumonia*, resulting in a

novel (uridylyl) catalyst for the production of a series of UDP-furanoses and UDP-3-*O*-alkylglucoses and UDP-mannose.[199]

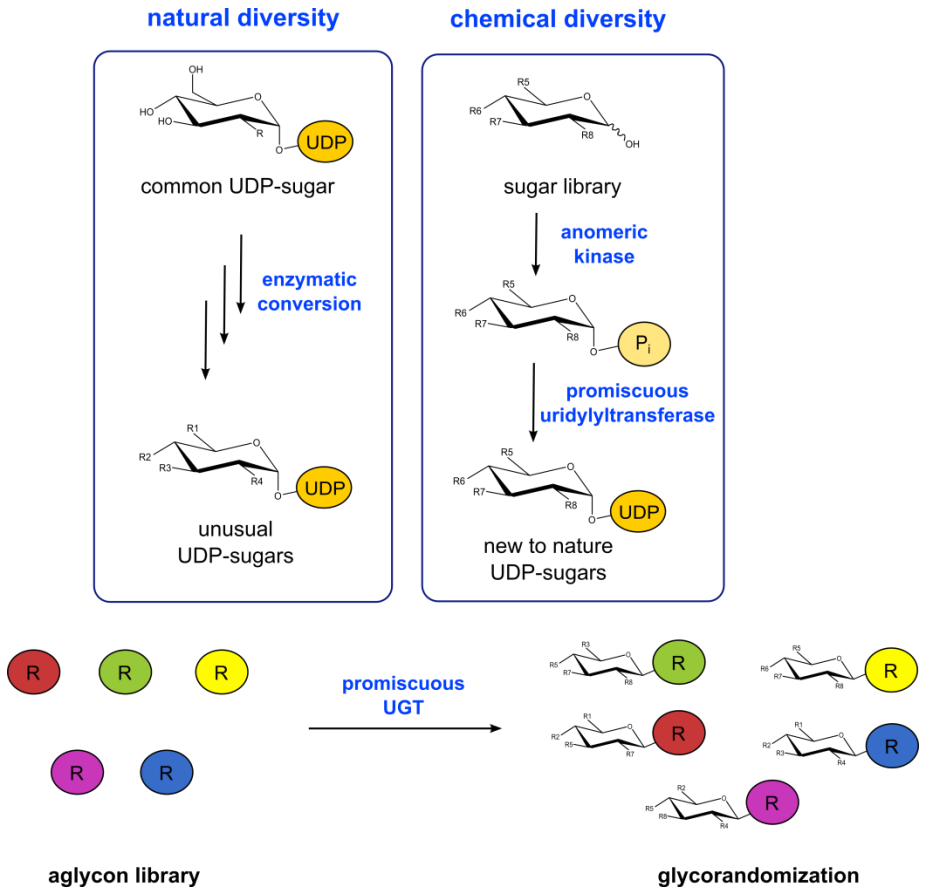


Figure 2.6: Chemo-enzymatic glycorandomization is based on the combination of an aglycon library with various (new or unusual) UDP-sugars. These can be formed by either exploiting the diversity of nature, using unusual UDP-sugar modifying enzymes or by exploiting the potential of chemistry to create new sugar libraries. The latter can be converted into the corresponding UDP-sugar via the combined action of a promiscuous kinase and uridylyltransferase. A chemically diverse aglycon library can be subsequently glycosylated with a promiscuous UGT using the new or unusual UDP-sugars, which results in a glycorandomized library.

The culminating step in glycorandomization is coupling the diversity of UDP-sugars with an aglycon (library). Despite the enormous wealth of UGTs, the majority of them is highly specific towards both the UDP-sugar and the aglycon, often rendering them unable to glycosylate even closely resembling analogues. An exception are some bacterial UGTs – characterized by a rather open and large aglycon binding pocket – which show promiscuous activity towards an enormous variety of chemical classes. To date, two bacterial UGTs have been extensively used: YjiC from *Bacillus licheniformis*[108,209] and OleD from *Streptomyces antibioticus*[150,210], the latter being extensively engineered into an enhanced multifunctional biocatalyst generating *O*-, *S*- and *N*-glycosides of well over 100 extremely diverse compounds.[210-214] Interestingly, the reversibility of some GTs (including OleD as best studied example) has been employed

to generate 22 nucleotide sugars from corresponding 2-chloro-4-nitrophenyl β-D-glycosides.[215] Since the released 2-chloro-4-nitrophenolate can be readily detected at 410 nm, this resulted in the first generally applicable high throughput screen for UGTs, which is an indispensable tool in screening their glycosylation potential in glycorandomization trials and eliminated the need for tedious UDP-sugar synthesis.[128,215]

2.7 Future outlook

The merging and recent advances in the fields of metabolic engineering, protein engineering and chemo-enzymatic glycorandomization have paved the way for UGT mediated production of virtually any glycoside. In this respect, *in vitro* glycosylation approaches are of great importance for discovery-scale synthesis of glycosides. These systems generally benefit from a facilitated downstream processing, and don't suffer from acceptor toxicity or solubility issues when compared to *in vivo* processes. However, scale-up towards economically viable processes is hampered by the need for purified proteins and expensive cofactors. Key requirements for successful *in vitro* application of UGTs in industry will consist of (thermo)stable enzymes, a high regeneration of UDP (RC_{max} of 50 or more), efficient conversion of the aglycon and cheap starting substrates. From this point of view, sucrose synthase is emerging as an efficient (re)generator of UDP-glucose from cheap sucrose as recently reported by Nidetzky and coworkers.[130] Also coupling two UGTs for the biocatalytic rearrangement of glycosides holds promise, as illustrated by the conversion of inexpensive phlorizin (*O*-glucoside) into nothofagin (*C*-glucoside).[216]

To overcome the issues regarding scalability and process costs, *in vivo* production of glycosides has become an interesting alternative since the microbial host produces its enzymes and endogenous UDP-sugar pool out of cheap substrates. Although some highly efficient producer strains have been developed for the multigram-scale production of oligosaccharides, the glycosylation of specialized metabolites lacks far behind. A major reason is the inadequate formation of UDP-sugars, which results in rapid depletion of these UDP-sugars when UGTs are overexpressed. To this end, the productivity of many strains could be greatly improved when most engineering strategies are revised as a function of the metabolic state of the host and the introduction of active UDP-sugar formation routes. In the latter case, sucrose synthase or sucrose phosphorylase could be of particular interest, since these directly form the activated precursors from the bio-economically attractive and more sustainable substrate sucrose. In this way, the classic engineering strategies relying on the bottleneck enzyme phosphoglucomutase (Pgm) for the generation of glucose 1-phosphate (as a precursor for UDP-glucose) are skipped.

Moreover, intracellularly released fructose (from sucrose) can function as a substrate for growth, thus opening the possibility for simultaneous growth and production, which results in a fermentative system. In addition, the formed UDP-glucose can subsequently act as a versatile pivot for many other natural and unusual UDP-sugars by overexpressing UDP-sugar interconverting or modifying genes. The wealth of these exotic enzymes and UGTs provided by nature may even be further expanded by using the powerful tools developed in metagenome research[217,218] and structure based protein engineering[149,206,219,220]. These methodologies will converge in the development of an easily scalable glycosylation host which is industrially applicable. However, some technological challenges remain regarding the evaluation of undesired events or side-reactions taking place in the cell: degradation of UDP-sugars, modification or metabolization of the aglycon, insolubility of overexpressed enzymes, aglycon or glycoside transport problems, or a redox imbalance by overconsumption of NAD(P)H. In addition, the toxicity of the aglycon can be a major hurdle for *in vivo* approaches, but can be partly resolved by controlled feeding or the *de novo* production of the aglycon, thus detoxifying the cell through glycosylation[96].

Although large-scale glycosylation is still in its infancy, many progress has been made in the microbial production of an overwhelming variety of specialized metabolites. Stimulated by a more systematic metabolic engineering approach, production pathways have been divided in modules which can be varied on various levels such as transcription, translation and even enzyme properties. This multivariate modular metabolic engineering (MMME) technique has proven to be very promising[221] as it alleviates the need for high throughput screens. To boost *in vivo* glycosylation and more specifically the *de novo* production of glycosides, UDP-sugar formation (and concomitant glycosylation) should be considered and implemented as a new module. In this way, aglycon and UDP-sugars modules can be varied to determine the optimal pathway conditions which can greatly enhance the product yields and titers. Eventually, the modular merging of novel and robust glycosylation strategies with microbial production of aglycons will unlock a new era for a myriad of glycosides, as extraction techniques will be increasinly replaced by *de novo* production.

Chapter 3

Unraveling the Leloir pathway of *Bifidobacterium bifidum*

Significance of the uridylyltransferases

Contents

3.1 Abstract **30**

3.2 Introduction **31**

3.3 Materials and methods **34**

3.3.1 Bacterial strains and plasmids 34

3.3.2 Reagents 34

3.3.3 Culture media 34

3.3.4 DNA isolation, manipulation and construction of uridylyltransferase vectors 34

3.3.5 Construction of deletion mutants and expression strains 35

3.3.6 Preparation of crude enzyme extracts 35

3.3.7 Purification of GalT enzymes 36

3.3.8 Assay for UTP-hexose-1-phosphate uridylyltransferase activity 36

3.3.9 Assay for UDP-glucose-hexose-1-phosphate uridylyltransferase 36

3.3.10 Kinetic analysis 37

3.3.11 Malachite green assay 37

3.3.12 LC-MS 37

3.3.13 Nucleotide sequences accession number 37

3.4 Results **38**

3.4.1 Cloning of the uridylyltransferase genes 38

3.4.2 Assay validation 38

3.4.3 Expression host validation 39

3.4.4 UTP-hexose-1-phosphate uridylyltransferase activity 40

3.4.5 Activity of GalT1 and GalT2 40

3.5 Discussion 42
3.5.1 Assay and expression host validation 42
3.5.2 Specificity of UgpA 42
3.5.3 Specificity of GalT enzymes 42
3.5.4 *galT* gene clusters 43
3.5.5 Metabolic implications 43
3.6 Conclusions 47

Part of this chapter has been published as:

De Bruyn, F., Beauprez, J., Maertens, J., Soetaert, W., and De Mey, M. (2013) **Unraveling the Leloir Pathway of *Bifidobacterium bifidum*: Significance of the Uridylyltransferases.** Appl. Environ. Microbiol. 79, 7028-7035.

3.1 Abstract

Sugars (cellulose, pectin, glycogen, starch) are the building blocks of life and play a central role at the same time in the organization thereof. Adding a sugar residue to a molecule can greatly alter its stability or bioactivity, and can even aid in organ specific (drug) delivery. However, sugars usually need to be 'activated' into a uridine-diphosphate (UDP) sugar before they can be linked by glycosyltransferases to a molecule of interest. This activation is usually mediated through uridylyltransferases, which reversibly catalyze the transfer of a UDP group to a sugar 1-phosphate. Many specific uridylyltransferases have been described in great detail, but those having a broad substrate range are absent in most bacteria. One exotic enzyme activity was reported in the late 70's from the crude extract of *Bifidobacterium bifidum* exhibiting activity towards both glucose 1-phosphate (glc1P) and galactose 1-phosphate (gal1P), but detailed information lacked about the sequence.

In this chapter, 3 uridylyltransferase candidates (*galT1*, *galT2* and *ugpA*) were cloned from *B. bifidum*, expressed in an engineered *E. coli* mutant and investigated with a newly developed chemo-enzymatic assay. GalT1 and GalT2 showed UDP-glucose-hexose-1-phosphate uridylyltransferase activity (EC 2.7.7.12), whereas UgpA showed the sought after promiscuous UTP-hexose-1-phosphate uridylyltransferase activity (EC 2.7.7.10). The activity of UgpA towards glc1P was about 33-fold higher than towards gal1P. Furthermore, the physiological roles of the other GalT's were also investigated and revealed that GalT1, as part of the bifidobacterial Leloir pathway, was about 357-fold more active than GalT2, the functional analog in the galacto-*N*-biose/lacto-*N*-biose (GNB/LNB) pathway. These results suggest that GalT1 plays a more significant role than previously thought and predominates when *B. bifidum* grows on lactose and human milk oligosaccharides. GalT2 activity is only required during growth on substrates with a GNB core such as mucin glycans.

3.2 Introduction

In 1899 a bacterium was isolated by Henri Tissier from the faeces of a breast-fed infant.[222] He introduced the name *Bacillus bifidus* (later reclassified as *Bifidobacterium bifidum*) after the Latin word *bifidus*, meaning forked, because of its Y-shaped morphology. Bifidobacteria are considered to benefit human health through inhibition of pathogens and regulation of intestinal microbial homeostasis.[223] They are a predominant part of the gut microbiota of breastfed infants and occur naturally in the lower part of the human gastrointestinal tract, where common mono- and disaccharides are scarce.[224] Therefore, bifidobacteria have developed alternative and exotic pathways that enable them to utilize various oligosaccharides such as mucin glycans and the two types of human milk oligosaccharides (HMO) based on their core sugars: lacto-*N*-biose I (LNB, type I) and *N*-acetyllactosamine (LacNAc, type II).[151,225,226] Growth of *B. bifidum* on these oligosaccharides by action of extracellular glycosidases has been investigated in great detail, demonstrating rapid release and subsequent uptake of LNB.[227] LNB is hypothesized to be the bifidus factor in HMO and thus a key factor in intestinal colonization.[50,228] Degradation of this disaccharide and the related galacto-*N*-biose (GNB) occurs through the GNB/LNB pathway.[151]

Table 3.1: Locus tags of the genes investigated in this study based on the *B. bifidum* PRL2010 genome[229] and their gene products.

Gene	Locus Tag	Gene product
ugpA	BBPR_0976	UTP-glucose-1-phosphate uridylyltransferase
galE1	BBPR_1456	UDP-glucose 4-epimerase
galT1	BBPR_0406	UDP-glucose-hexose-1-phosphate uridylyltransferase
galK	BBPR_0407	galactokinase
lnbp	BBPR_1055	galacto-*N*-biose/lacto-*N*-biose I phosphorylase
nahK	BBPR_1052	*N*-acetylhexosamine kinase
galT2	BBPR_1051	UDP-glucose-hexose-1-phosphate uridylyltransferase
galE2	BBPR_1050	UDP-glucose 4-epimerase

The GNB/LNB pathway was discovered in *Bifidobacterium longum* and is encoded by the *lnpABCD* operon. This operon codes for a galacto-*N*-biose/lacto-*N*-biose I phosphorylase (LNBP), a *N*-acetylhexosamine 1-kinase (NahK), a UDP-glucose-hexose-1-phosphate uridylyltransferase (GalT2), and a UDP-glucose 4-epimerase (GalE2).[151] Metabolic profiling and genome sequencing of various *B. bifidum* strains[229-232] reveal that a similar GNB/LNB gene cluster is present in this species and the corresponding locus tags are listed in Table 3.1. However, this cluster is organized differently as two sugar kinases lie

between the coding sequences for LNBP and NahK (Figure 3.1). LNBP catalyzes the first reaction of this pathway, the phosphorolytic cleavage of a galactosyl-beta-1,3-*N*-acetylhexosamine (GNB or LNB) into an *N*-acetylhexosamine (HexNAc) and galactose 1-phosphate (gal1P).[233] Due to its similarity to the Leloir pathway and the direct generation of gal1P without action of a galactokinase (GalK), the GNB/LNB pathway is considered an energy-saving variant of the Leloir pathway (Figure 3.2). Therefore, it has been suggested that this is the main pathway for galactose metabolism in bifidobacteria and that it prevails over the more common Leloir pathway.[151,234,235]

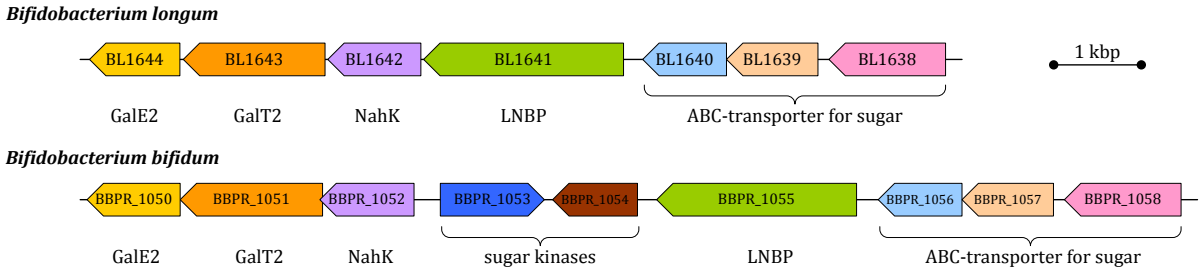


Figure 3.1: Schematic representation of the GNB/LNB gene cluster in *B. longum* NCC2705 (gene BL1638 to BL1644) and *B. bifidum* PRL2010 (gene BBPR_1050 to BBPR_1058). Two sugar kinases interrupt the GNB/LNB pathway of *B. bifidum* and it is not organized as in the *lnpABCD* operon (BL1641 to BL1644) of *B. longum*.

However, based on the presence of other uridylyltransferases and transcriptome profiling of *B. bifidum*, there are indications that the key intermediate gal1P may be metabolized via routes other than the GNB/LNB pathway. In this study we investigate the function of three annotated yet poorly characterized uridylyltransferases – GalT1, GalT2 and UgpA – of *B. bifidum* and their activity towards gal1P. GalT1 and GalT2 are both annotated as UDP-galactose-1-phosphate uridylyltransferases, yet share little (12.1 %) sequence identity, while UgpA is annotated as an UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9). In this paper, the enzymes were cloned and overexpressed in an engineered *E. coli* strain in order to eliminate interference with the enzymes investigated, and were screened for both activities, using a new chemo-enzymatic assay. More detailed characterization of these uridylyltransferases provides new insights in gal1P degradation and its metabolic implications in *B. bifidum*.

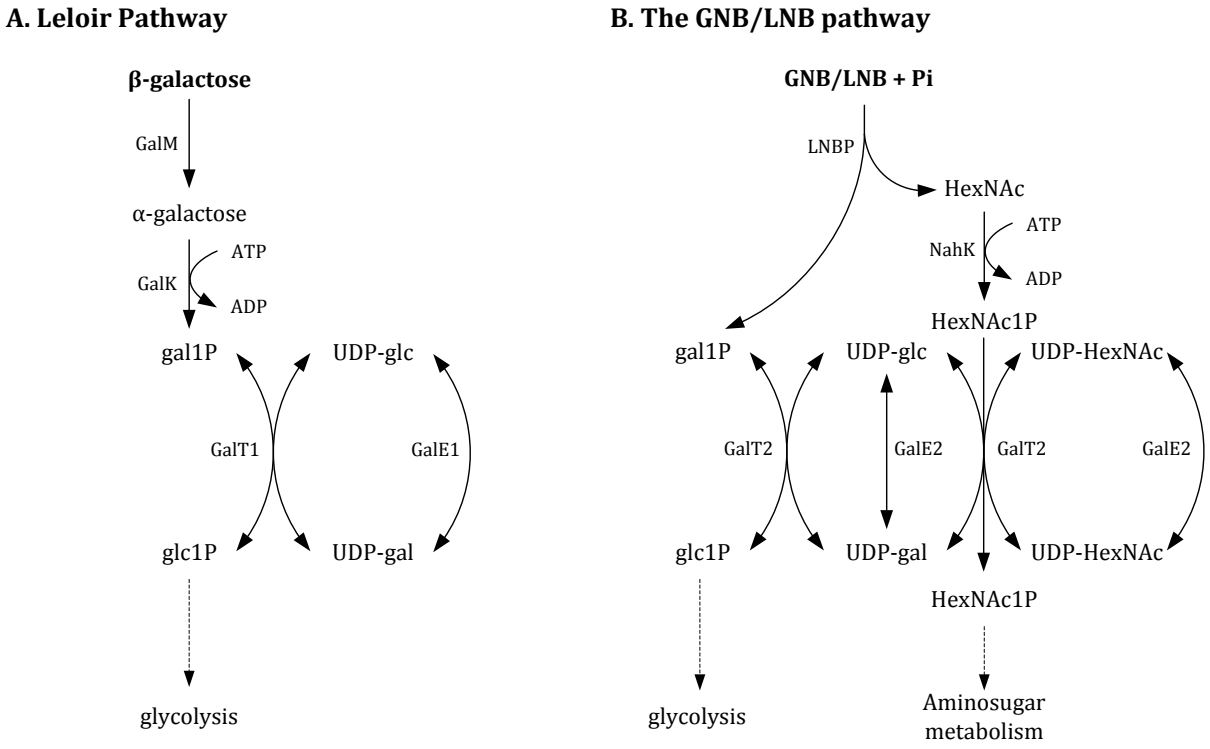


Figure 3.2: Overview of the Leloir pathway (A) and GNB/LNB pathway (B) in *Bifidobacterium bifidum*. LNBP, NahK, GalT2 and GalE2 are encoded in the same gene cluster.[151] In contrast, the classic Leloir pathway (A) is encoded by scattered genes and evidence is lacking for the existence of metabolic crossover between both pathways. Gene locus tags of the corresponding enzymes are listed in Table 3.1.

3.3 Materials and methods

3.3.1 Bacterial strains and plasmids

All plasmids and strains used in this study are listed in Appendix 1 and 2 respectively. *E. coli* DH5α (subcloning) was used for plasmid cloning and propagation, while *E. coli* MG1655 Δ*galETKM* Δ*galU* Δ*ushA* Δ*ugd* Δ*agp* (hereafter named sMEMO_WT) was used for the expression and crude extract preparation of the uridylyltransferases. *E. coli* BL21 (DE3) was used for the expression of the His_6 tagged uridylyltransferases. The λ-red pKD46 plasmid, plasmid pKD4 for amplification of the kanamycin resistant marker and plasmid pCP20 for removal of this marker were used for the one step inactivation in *E. coli* MG1655 as described by Datsenko and Wanner.[236]

3.3.2 Reagents

T4 DNA-ligase and all restriction enzymes were purchased from New England Biolabs (Ipswich, Massachusetts). PrimeSTAR polymerase was purchased from Takara Bio (Japan). All chemicals used in crude extract preparation and assays were purchased from Sigma Aldrich (Germany), except for gal1P and UTP purchased from Merck (Darmstadt, Germany) and UDP-glucose and *N*-acetylglucosamine 1-phosphate from Carbosynth (Berkshire, UK).

3.3.3 Culture media

E. coli cultures were grown on Luria-Bertani (LB) medium (Difco) with the necessary antibiotics (50 μg/mL kanamycin, 100 μg/mL ampicillin, 25 μg/mL chloramphenicol) for maintenance and selection of the plasmids. *B. bifidum* (ATCC 29521) was grown anaerobically at 37 °C on LMG Medium 144 (per liter of medium: 23 g special peptone (Difco), 1 g soluble starch, 5 g NaCl, 0.3 g cysteine hydrochloride and 5 g glucose). The anaerobic environment, consisting of 5 % H_2 and 95 % N_2, was created by an anaerobic chamber (Concept 1000, Ruskinn Technology Ltd, UK).

3.3.4 DNA isolation, manipulation and construction of uridylyltransferase vectors

Genomic DNA (gDNA) was obtained by harvesting the cells of the culture by centrifugation at 22000 × g (Heraeus Biofuge, Thermo). Cells were washed with saline and heated for 5 min at 100 °C. The cell debris was removed by centrifugation at 22000 × g for 10 min and the supernatant was used as gDNA. All primers used in this study are listed in Appendices Appendices A5 (deletion primers) and A6 (cloning primers) and the gDNA of *B. bifidum* was used for the amplification of *ugpA*, *galT1* and *galT2* genes. gDNA

of *E. coli* MG1655 was used for the amplification of *galU*. All the genes were amplified two times independently and sequenced to ensure no mutations occurred in the constructs. The amplified *galU* and *ugpA* fragments were cut with EcoRI-HF and SacI-HF and ligated into the pCX-Kan-P22 vector resulting in pCX-Kan-P22-galU and pCX-Kan-P22-ugpA respectively. The *galT1* and *galT2* fragments were cut with NdeI-HF and PmeI and ligated into the p10-Trc vector resulting in the p10-Trc-galT1 and p10-Trc-galT2 plasmids. In addition, both vectors were also redesigned so that an N-terminal His_6 tag was added to allow purification, resulting in the p10-Trc-His-galT1 and p10-Trc-His-galT2 plasmids. These plasmids were created by using a 2 pieces Gibson assembly method[237] using cloning primers for the amplification of the p10-Trc backbone and His-galT1 and His-galT2 inserts listed in Appendix A6. All plasmids were sequenced and showed no mutations.

3.3.5 Construction of deletion mutants and expression strains

To eliminate interference of the crude enzyme extract with the substrates used in the assay an *E. coli* MG1655 knockout mutant was created by application of the one step deletion system of Datsenko and Wanner[236]. The linear DNA for the deletion of each target gene was amplified using deletion primers (Fw/Rv_gene_del) as mentioned in Appendix A5. The *galETKM* operon and the genes *galU*, *ushA*, *ugd* and *agp* were deleted resulting in the sMEMO_WT mutant (Appendix A2) as an expression host for the recombinant uridylyltransferases. Transformation of the constructed pCX-Kan-P22-galU, pCX-Kan-P22-ugpA, p10-Trc-galT1 and p10-Trc-galT2 plasmids in this sMEMO_WT mutant resulted in the sMEMO_GalU, sMEMO_UgpA, sMEMO_GalT1 and sMEMO_GalT2 strains respectively. The p10-Trc-His-galT1 and p10-Trc-His-galT2 plasmids were transformed in BL21(DE3), resulting in sMEMO_His_GalT1 and sMEMO_His_GalT2 strains.

3.3.6 Preparation of crude enzyme extracts

The sMEMO_GalU, sMEMO_UgpA, sMEMO_GalT1 and sMEMO_GalT2 strains were grown in tubes containing 5 mL LB (plus kanamycin or chloramphenicol when required) for 8 h at 37 °C. The culture served as 2 % inoculum for 250 mL Erlenmeyer flasks containing 50 mL LB medium with 1 % glucose (and kanamycin or chloramphenicol when required). Strains sMEMO_GalT1 and sMEMO_GalT2 were induced with 0.2 mM IPTG after inoculation. Shake flasks were incubated at 37 °C and 200 rpm for 16 h. Cells were harvested by centrifugation (Heraeus Biofuge, Thermo) for 10 min at 5000 × g. The cell pellet was washed first with saline and then with 50 mM MOPS buffer (pH 6.5 or pH 7 depending on the assay). Finally, the pellet was dissolved in 5 mL of the above MOPS buffer and disrupted by sonication for 2 x 4 min (Branson sonifier, 50 % duty cycle,

output 3). The crude enzyme extract was obtained by centrifugation for 20 min at 22000 × g. Protein concentration was determined via Bradford assay.[238]

3.3.7 Purification of GalT enzymes

For the purification of both GalT1 and GalT2 enzymes, strains sMEMO_His_GalT1 and sMEMO_His_GalT2 were cultivated on 100 mL LB medium and chloramphenicol at 30 °C and 200 rpm. When absorbance at 600 nm reached 0.6, IPTG was added at a final concentration of 0.4 mM. The crude extract was obtained in the same way as described above and both enzymes were purified on Ni-nitrilotriacetic acid agarose gel (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. After purification a buffer exchange was performed to 50 mM MOPS buffer (pH 7) using an Amicon Ultra-4 centrifugal filter unit (30 kDa NMWL) from Merck (Darmstadt, Germany). Protein concentration was determined via Bradford assay [238].

3.3.8 Assay for UTP-hexose-1-phosphate uridylyltransferase activity

The assay was based on the article of Wu and coworkers[239], but redesigned for pyrophosphate quantification. UTP-hexose-1-phosphate uridylyltransferase activity was measured in a 20 µL reaction mixture containing 50 mM MOPS buffer (pH 6.5), 2 mM UTP, 2 mM hexose 1-phosphate (glucose 1-phosphate (glc1P) or gal1P), 2 mM $MgCl_2$, 1 U of inorganic pyrophosphatase from *E. coli*, and various concentrations of crude enzyme extract. The reaction mixture was incubated at 37 °C for 15 min and was stopped by adding 180 µL of a 10 mM EDTA solution. Final activity was determined by quantifying the released phosphate during the reaction using a malachite green assay as described below. One unit of UTP-hexose-1-phosphate uridylyltransferase activity was defined as the amount of enzyme that formed 1 µmol of UDP-hexose per minute under these conditions.

3.3.9 Assay for UDP-glucose-hexose-1-phosphate uridylyltransferase

UDP-glucose-hexose-1-phosphate uridylyltransferase activity was measured by using a continuous coupled assay for glc1P quantification. The assay mixture consisted of 100 µL assay solution (50 mM MOPS buffer (pH 7), 2 mM NAD^+, 10 mM $MgCl_2$, 30 µM glucose-1,6-diphosphate, 1.2 U/mL phosphoglucomutase from rabbit muscle and 1.2 U/mL glucose 6-phosphate dehydrogenase from *L. mesenteroides*), 50 µL of substrate solution (50 mM MOPS buffer (pH 7), 4 mM gal1P and 4 mM UDP-glucose) and a 50 µL dilution of crude enzyme extract or purified enzyme in 50 mM MOPS buffer (pH 7). Reaction was performed at 37 °C in a microtiter plate and NADH formation was monitored continuously by measuring the absorbance at 340 nm. One unit of UDP-

glucose-hexose-1-phosphate uridylyltransferase activity was defined as the amount of enzyme that formed 1 µmol of glc1P per minute under these conditions.

3.3.10 Kinetic analysis

The kinetic constants were derived from initial rate analysis by varying the concentration of individual substrate. For the UTP-hexose-1-phosphate uridylyltransferase assay, UTP was varied from 0 to 2 mM in the presence of 2 mM hexose 1-phosphate (glc1P or gal1P). Hexose 1-phosphate was subsequently varied from 0 to 2 mM at 2 mM saturation of UTP. For the UDP-glucose-hexose-1-phosphate uridylyltransferase assay, gal1P was varied from 1 to 0.02 mM in the presence of 1 mM UDP-glucose. Kinetic parameters were calculated from an S-V plot by non-linear regression analysis using the Michaelis-Menten kinetic equations in R ('nlstools' package).

3.3.11 Malachite green assay

Phosphate concentration of the samples in the micromolar range was determined using a malachite green assay: to 50 µL of sample, 30 µL reagent A, 100 µL milliQ water and 30 µL reagent B was added (in this order). The mixture was incubated for 20 min in a microtiter plate at room temperature and absorbance was measured at 630 nm. Reagent A consists of 50 mM of ammonium heptamolybdate in 3 M of sulphuric acid. Reagent B consists of 0.093 % (w/v) malachite green and 0.93 % (w/v) polyvinyl alcohol (M_w 14000). Concentrations were determined using a phosphate standard (serial dilutions ranging from 0 – 100 µM).

3.3.12 LC-MS

The products – phosphorylated and nucleotide sugars – of the enzymatic reactions were also verified by Liquid Chromatography coupled to a Mass Spectrometer (LC-MS) using a Cosmosil Hilic (Nacalai USA, San Diego) column (4.6 x 250 mm) with isocratic separation (0.1 M ammonium acetate (50 %) and acetonitrile (50%) at a flow rate of 1 mL/min at 35 °C for 30 min). The LC system was coupled to a Micromass Quattro LC (McKinley Scientific, USA). Detection was performed in negative mode ESI-MS with a capillary voltage of 2.53 kV, a cone voltage of 20 V, cone and desolvation gas flows of 93 and 420 L/h, and source and cone temperatures of 150 and 350°C, respectively.

3.3.13 Nucleotide sequences accession number

The accession numbers of the cloned genes are listed in Appendix A3.

3.4 Results

3.4.1 Cloning of the uridylyltransferase genes

The nucleotide sequences of *ugpA*, *galT1* and *galT2* from *B. bifidum* ATCC 29521 were 99.4 %, 99.6 % and 99.5 % identical to the BBPR_0976, BBPR_0406 and BBPR_1051 genes of *B. bifidum* PRL2010[229]. The predicted functions of these genes are UTP-glucose-1-phosphate uridylyltransferase (*ugpA*) and galactose-1-phosphate uridylyltransferase (*galT1* and *galT2*), respectively. These genes were expressed in the *E. coli* MG1655 mutant and an assay was performed to determine their function as described in the Materials and Methods section.

3.4.2 Assay validation

A new UTP-hexose-1-phosphate uridylyltransferase assay was developed based on the principle of phosphate detection as described by Wu and coworkers[239]. Crude extracts with this uridylyltransferase activity release pyrophosphate, which is converted to organic phosphate (Pi) when coupled with a pyrophosphatase. The subsequent reaction of phosphate with a malachite green reagent is spectrophotometrically detected in the micro molar range. A schematic overview of this assay is given in Figure 3.3. Since the detailed composition of the reagents for malachite green detection of phosphate was not available, a reverse engineering effort was done to determine this. The result is the 2-component mixture described in 3.3.11 and was able to quantify inorganic phosphate up to 100 µM.

Validation of the UTP-hexose-1-phosphate uridylyltransferase assay was performed using commercially available uridine-5′-diphosphoglucose pyrophosphorylase from baker's yeast (Sigma-Aldrich, 40 – 60 % protein concentration, ≥ 50 U/mg protein). One unit was defined as forming 1 µmol glucose 1-phosphate (glc1P) at 25 °C at pH 7.6 in 1 minute. Enzyme dilutions were incubated for 10 min with 2 mM of UTP and glc1P under the conditions described in section 3.3.8 and released phosphate was measured with the malachite green assay. A specific activity of 65.2 ± 1.7 U/mg protein was found. This coupled assay with subsequent malachite green phosphate detection proved to be a stable and precise method for determination of the specific activity. Validation of the UDP-glucose-hexose-1-phosphate uridylyltransferase assay was performed by adding 100 µL of 2 mM glc1P to the assay solution.

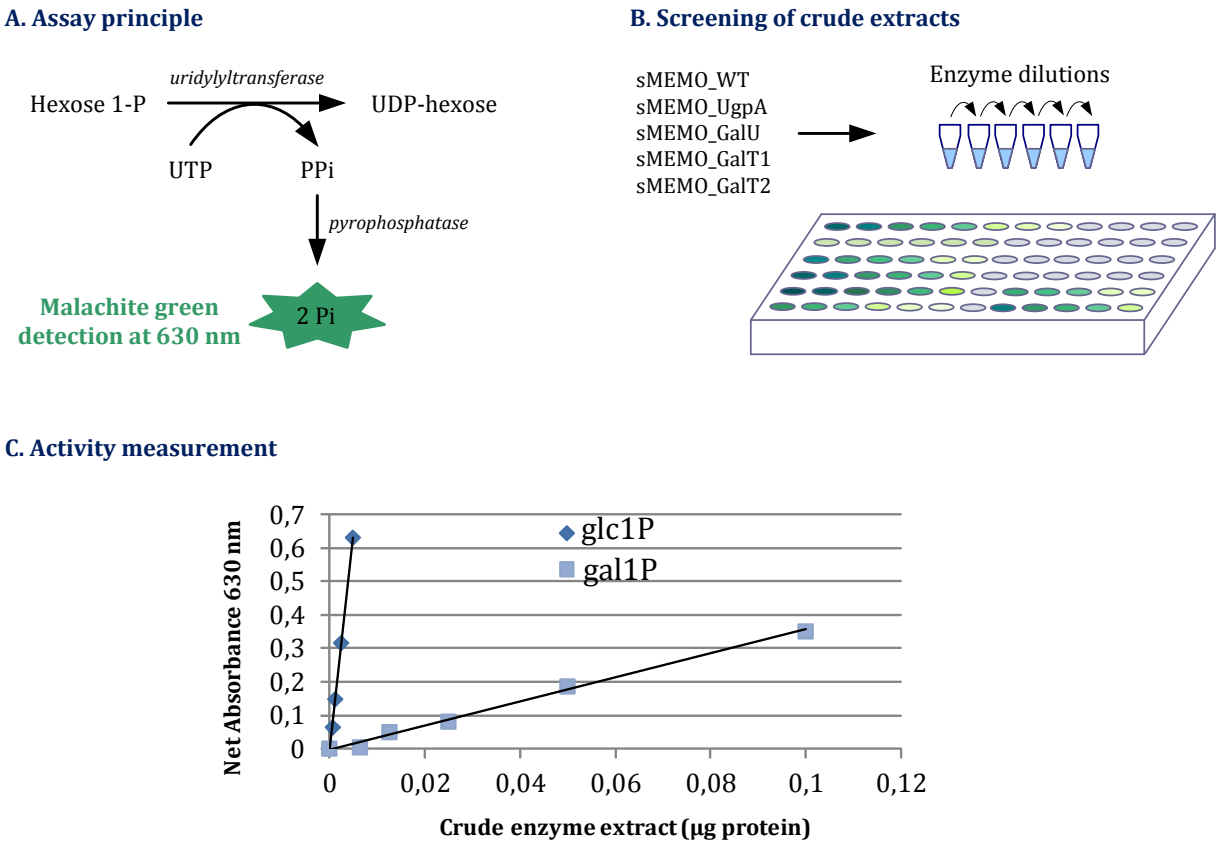


Figure 3.3: Schematic overview of the developed UTP-hexose-1-phosphate uridylyltransferase activity assay (A), screening of the crude enzyme extract dilutions for this activity (B) and the measurement of the specific activity towards glc1P and gal1P (C).

3.4.3 Expression host validation

An *E. coli* MG1655 mutant (sMEMO_WT) was created as an efficient screening host for the different uridylyltransferases from *B. bifidum*. The Leloir pathway (*galETKM*) and the gene coding for UTP-glucose-1-phosphate uridylyltransferase (*galU*) were deleted to prevent interference with the investigated enzymes. The degradation of hexose 1-phosphate or UDP-hexose substrates was also prevented by knocking out UDP-sugar hydrolase (*ushA*), UDP-glucose 6-dehydrogenase (*ugd*) and an acid glucose-1-phosphatase (*agp*). The crude extract of sMEMO_WT was tested against the substrates of both assays to detect possible interference. No activity was observed for the UDP-glucose-hexose-1-phosphate uridylyltransferase assay. Testing of the crude extract against a hexose 1-phosphate and UTP showed no activity for the UTP-hexose-1-phosphate uridylyltransferase assay. However minor activity was observed with UTP as sole substrate after extended incubation times (> 30 min) and using undiluted crude extracts. This is probably due to UTP hydrolysis by an unidentified hydrolase. Since

incubation times were always lower than 15 min and dilutions starting from 0.05 were used, sMEMO_WT was considered a suitable screening host.

3.4.4 UTP-hexose-1-phosphate uridylyltransferase activity

Crude enzyme extract dilutions of UgpA, GalT1 and GalT2 were screened for UTP-hexose-1-phosphate uridylyltransferase activity against 2 mM hexose 1-phosphate (glc1P or gal1P) and 2 mM UTP. Only UgpA showed activity against both substrates. Formation of the corresponding UDP-hexose (UDP-glucose or UDP-galactose respectively) was confirmed by LC-MS. The specific activity of UgpA crude extract towards glc1P was 13 U mg^{-1}, which was about 33-fold higher than towards gal1P. The apparent K_m values were also calculated and seemed to be similar for both substrates (Table 3.2). As a reference point for the substrate usage spectrum, the UgpA equivalent of *E. coli* (GalU) was also assayed against both substrates, showing only activity towards glc1P. The specific activity of the GalU crude extract was about 6.9-fold lower compared to UgpA.

Table 3.2: Kinetic parameters for UgpA and GalU crude protein extracts [a]

		H1P	UTP
	Sp act (U mg^{-1})	**K_m (mM)**	**K_m (mM)**
UgpA (G1P)	13.0 ± 0.08	0.098 ± 0.011	0.042 ± 0.006
UgpA (gal1P)	0.40 ± 0.01	0.148 ± 0.034	0.032 ± 0.010
GalU (G1P)	1.88 ± 0.03	0.010 [b]	0.070 [b]
GalU (gal1P)	ND	-	-

[a] Reactions carried out at pH 6.5 and 37 °C. ND: not detected. Experiments were carried out in triplicate and standard deviations are shown.

[b] Data from Weissborn *et al.* [240]. Reaction carried out at pH 7.8 and 37 °C.

3.4.5 Activity of GalT1 and GalT2

The 3 recombinant uridylyltransferase crude extracts were screened for UDP-glucose-hexose-1-phosphate uridylyltransferase activity against 2 mM UDP-glucose and 2 mM gal1P at pH 7 and 37 °C. GalT1 showed high activity; a 0.01 dilution of the crude extract converted all gal1P into glc1P within a 5 min incubation. The same dilution of GalT2, however, displayed a very weak activity which was 417-fold lower compared to GalT1. The activity did not improve after adding cysteine, Fe^{2+}, or Zn^{2+}, as proposed in literature.[241,242] Altering the expression conditions (0.1 mM IPTG induction), lowering of the assay temperature to 25 °C or raising the buffer pH to 8.5 had no effect either. To confirm these results, the His_6 tagged GalT1 and GalT2 were purified from the

sMEMO_His_GalT1 and sMEMO_His_GalT2 strains and different dilutions were used to measure activity. Other acceptors than gal1P were also tested, including *N*-acetylgalactosamine 1-phosphate (galNAc1P) and *N*-acetylglucosamine 1-phosphate (glcNAc1P) in the presence of 2 mM UDP-glucose and the specific activities are given in Table 3.3. GalT2 showed a 357-fold lower activity than GalT1, which is consistent with the results based on the crude extracts. The apparent K_m of GalT1 for gal1P in the presence of 2 mM UDP-glucose was 0.29 ± 0.08 mM, which is in the range of previous findings.[243] The apparent K_m of GalT2 for gal1P was 0.61 ± 0.1 mM, while the K_m of GalT2 towards galNAc1P was only 65.1 ± 15.9 µM.

Table 3.3: Kinetic parameters of the purified GalT1 and GalT2 enzymes as function of various hexose 1-phosphates [a]

Enzyme (H1P)	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
GalT1 (gal1P)	73 ± 1.8	0.29 ± 0.08	251240
GalT1 (galNAc1P)	ND	-	-
GalT1 (glcNAc1P)	ND	-	-
GalT2 (gal1P)	0.24 ± 0.01	0.61 ± 0.1	400
GalT2 (galNAc1P)	0.31 ± 0.03	0.07 ± 0.02	4802
GalT2 (glcNAc1P)	0.22 ± 0.01	-	-

[a] Reactions carried out at pH 7 and 37 °C for various acceptors. Experiments were carried out in triplicate and standard deviations are shown. ND: not detected.

3.5 Discussion

3.5.1 Assay and expression host validation

A new and reliable assay was developed that allows fast screening and characterization of enzymes with UTP-hexose-1-phosphate uridylyltransferase activity. The same assay can be used to investigate various other enzymes that liberate pyrophosphate, such as mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) or isoprene synthase (EC 4.2.3.27). The development of the described expression host, lacking the Leloir pathway and relevant degradation reactions, bypasses the need to purify the enzymes investigated. These assays proved to be indispensable tools to characterize UgpA, GalT1 and GalT2 and gain new insights in the LNB/GNB metabolism of bifidobacteria.

3.5.2 Specificity of UgpA

Our results indicate that UgpA is a promiscuous UTP-hexose-1-phosphate uridylyl-transferase that catalyzes the formation of UDP-hexose from hexose 1-phosphate. Based on this activity and the size of the monomer (55 kDa), this enzyme is probably homologous to the enzyme purified from *B. bifidum* by Lee *et al.* but was never properly annotated.[14,244] They described the reverse (pyrophosphorylase) action of UgpA, which showed the same reactivity towards UDP-galactose as towards UDP-glucose. However, our findings indicate a 33-fold lower activity in the synthesis direction of UDP-galactose. Because of this promiscuity, a new EC number (EC 2.7.7.10) was proposed to distinguish it from UTP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9; GalU) that predominantly acts on glc1P in prokaryotes.[240,245] However, due to the activity on different sugar 1-phosphates, the absence of similarity with GalU-type enzymes, and a high level of identity with promiscuous eukaryotic uridylyltransferases, we suggest that UgpA should be classified as an UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64, synonym of UDP-sugar pyrophosphorylase, USP)[158,246,247]. Moreover, a UDP-sugar pyrophosphorylase from *B. longum* (BLUSP) was recently cloned, showing 82 % amino acid identity with UgpA, and was used for the efficient synthesis of a variety of UDP-sugars[146].

3.5.3 Specificity of GalT enzymes

We observed that both GalT1 and GalT2 possessed UDP-glucose-hexose-1-phosphate uridylyltransferase activity (EC 2.7.7.12; GalT), yet the activity of GalT2 was 357-fold lower compared to GalT1. A similar result was reported in the large scale preparation of LNB, where GalT1 of *B. longum* was used instead of GalT2.[50,151] GalT2 shows a broad acceptor specificity, with a 9-fold higher affinity towards galNAc1P than gal1P. The

amino acid sequences of both GalT enzymes shared almost no identities (12.1 %), which could be explained by convergent evolution of similar enzymatic function. However, it was shown that the majority of catalytic residues are under different evolutionary constraints in both type of enzymes and they likely have different functions.[248] This is supported by the fact that purified GalT2 was 28 % more active towards galNAc1P than gal1P as substrate, providing a link between growth on GNB/LNB and aminosugar metabolism.

3.5.4 *galT* gene clusters

GalT enzyme activity, which is essential for the transfer of the uridine nucleotide moiety from UDP-glucose to gal1P, is widespread in many organisms. However, based on the enzyme dissimilarities between GalT1 and GalT2, two classes seem to exist. Class I enzymes (GalT1 type) are found in various eukaryotic organisms and bacteria, while Class II enzymes (GalT2 type) are almost exclusively present in Firmicutes and Actinobacteria.[249,250] Coding sequences for GalT1 or GalT2 do not co-occur in the same organism, with some Clostridiales and bifidobacteria as exceptions. Based on comparative bifidobacterial genome analysis, only *B. bifidum*, *B. longum* and *B. breve* strains possess both *galT1* and *galT2*, being part of a *galTK* gene cluster and a GNB/LNB degradation gene cluster, respectively.[229,231,251] A schematic phylogenetic overview of these gene clusters is given in Figure 3.4. Metabolic and genetic explanations for the unique coexistence of both GalT enzymes in *B. bifidum*, *B. longum* and *B. breve* strains remain unclear, but is coupled to the occurrence of the GNB/LNB pathway, which is apparently absent in other bifidobacteria[227,232].

3.5.5 Metabolic implications

The occurrence of three uridylyltransferases in *B. bifidum* that act on gal1P can have metabolic implications that have been overlooked. While the Leloir pathway is in most organisms the major, if not the only, pathway for the degradation of galactose, an alternative route was discovered in *B. longum* under the form of the GNB/LNB pathway. The same gene cluster is also present in *B. bifidum* and it is suggested that this energy-saving variant of the Leloir pathway is the main pathway for galactose metabolism.[151] However, glycoprofiling of *B. bifidum* during growth on HMO shows the release of lactose and LNB as prominent disaccharides. While LNB enters the GNB/LNB pathway, galactose is released by action of a β-galactosidase[227] from lactose. Our findings suggest that galactose is primarily metabolized via the Leloir pathway, together with gal1P originating from LNB or GNB.

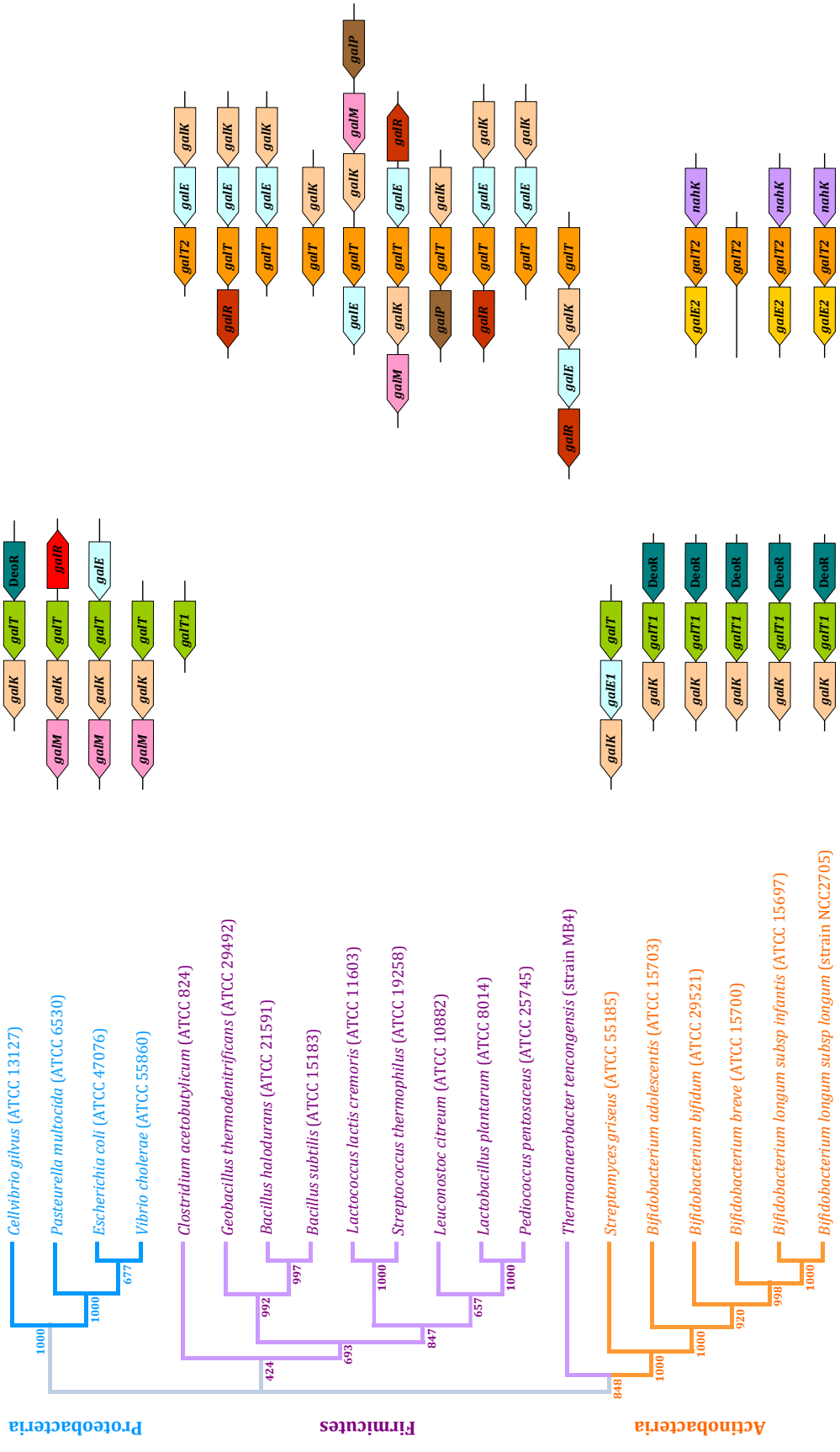


Figure 3.4: 16S rRNA-based phylogenetic tree with gene cluster organization and prevalence of Class I enzymes (GalT1 type) and Class II enzymes (GalT2 type) in three representative bacterial phyla. The bacterial strain codes and bootstrapping values are included. Orthologues that match the *galT1* or *galT2* genes of *B. bifidum* are aligned vertically. GalT2 type enzymes are present only in Firmicutes and Actinobacteria and show a high degree of rearrangement with the neighboring genes.

Although GalT2 (as part of the GNB/LNB pathway) is put forward to play the key role in gal1P degradation[151], we suggest that GalT1 (as part of the Leloir pathway) predominates when grown on lactose, galactooligosaccharides (GOS) and HMO. In addition to the 357-fold lower activity of GalT2, transcription data showed the upregulation of *galT1* during growth on these substrates relative to glucose[252,253]. It is often stated that the genes of the GNB/LNB pathway are upregulated during growth on oligosaccharides or mucin, but whether a part or the whole cluster is upregulated is highly dependent on the substrate. Growth on LNB-containing oligosaccharides (HMO) only requires LNBP and NahK activity in principle, yielding gal1P and glcNAc1P. The latter is metabolized into fructose 6-phosphate (F6P) by action of a phosphoglucosamine mutase, a *N*-acetylglucosamine 6-phosphate deacetylase, and a glucosamine 6-phosphate deaminase, which then enters the bifid shunt[254]. A significant upregulation of these genes was observed when grown on HMO.[229,255]

Substrates with a GNB core (such as mucin glycans) require LNBP and NahK activity, yielding gal1P and galNAc1P, yet also require GalT2 and GalE2 activity to be converted to glcNAc1P in order to enter the bifid shunt. This hypothesis is supported by the substrate preference of GalT2 towards galNAc1P and by transcriptional data of bifidobacteria grown on lactose, HMO, GOS or mucin based media.[229,252,253] A proposed metabolic route for lactose, LNB and GNB degradation is depicted in Figure 3.5. The metabolic implications of UgpA are probably limited to UDP-sugar generation, to provide the Leloir pathway with initial amounts of UDP-glucose.[235,245] This is supported by the fact that the K_m for hexose 1-phosphate is more than 2-fold lower compared to GalT1.

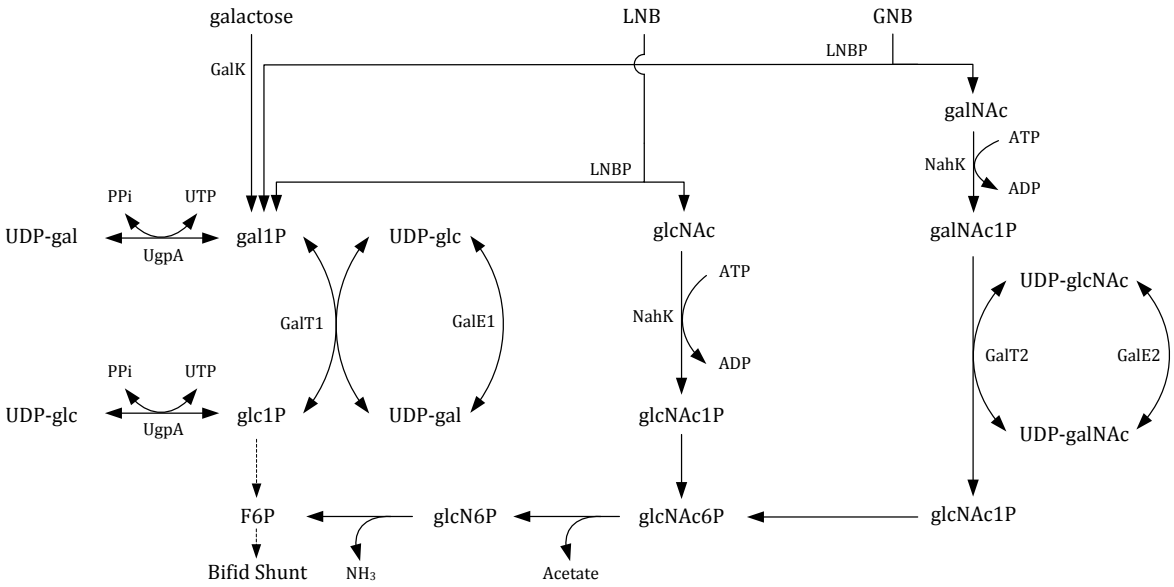


Figure 3.5: Proposed routes for the metabolization of galactose, lacto-*N*-biose I (LNB) and galacto-*N*-biose (GNB) in *B. bifidum*. All 3 substrates yield gal1P which is solely metabolized by the Leloir pathway (left). Depending on the presence of LNB or GNB, *N*-acetylglucosamine (glcNAc) or *N*-acetylgalactosamine (galNAc) are respectively formed through the action of LNBP. Utilization of glcNAc only requires *N*-acetylhexosamine kinase (NahK) activity to generate glcNAc1P. However, GNB rich substrates such as mucin need the full set of enzymes encoded by the GNB/LNB gene cluster (right). GalNAc is phosphorylated to GalNAc1P by NahK, which is subsequently converted by GalT2 and GalE2 via a Leloir like pathway to glcNAc1P which can then enter the Bifid Shunt via fructose 6-phosphate (F6P) by consequent action of a glucosamine mutase, a deacetylase and a deaminase. glcN6P: glucosamine 6-phosphate.

3.6 Conclusions

In this chapter 3 uridylyltransferases (UgpA, GalT1 and GalT2) from *B. bifidum* were characterized by expression in an engineered *E. coli* strain and using the crude extracts, thus avoiding the need to purify the enzymes. The extracts were easily screened for hexose-1-phosphate uridylyltransferase activity due to the development of a novel chemo-enzymatic assay based on phosphate detection in the micromolar range. UgpA was identified as the promiscuous uridylyltransferase with activity towards both glc1P and gal1P. This interesting feature together with its 7-fold higher activity compared to GalU has enabled the implementation in the next chapters. GalT1 and GalT2 exhibited UDP-glucose-hexose-1-phosphate uridylyltransferase activity and their function in respect to substrate usage in *B. bifidum* was discussed based on the data obtained *in vitro*.

Chapter 4

Development of an *in vivo* glucosylation platform by coupling production to growth

Production of phenolic glucosides by a GT of V. vinifera

Contents

4.1 Abstract **50**

4.2 Introduction **51**

4.3 Materials and methods **54**

4.3.1 Materials and Molecular Agents 54

4.3.2 Creation of the production plasmids 54

4.3.3 Creation of the *E. coli* W mutant strains 54

4.3.4 Media 55

4.3.5 Growth in shake flasks 55

4.3.6 Growth in bioreactors 56

4.3.7 Product analysis and quantification 56

4.3.8 Purification and structural elucidation of compounds 56

4.3.9 Mathematical methodologies 57

4.4 Results and discussion **58**

4.4.1 Introduction 58

4.4.2 Altering the sucrose metabolism of *E. coli* W to efficiently generate glc1P 58

4.4.3 Coupling growth and production 60

4.4.4 Optimization of production medium for glucosylation of gallic acid 63

4.4.5 Production of β-glucogallin 65

4.4.6 Screening for the glucosylation potential of different acceptors 68

4.4.7 Productivity comparison with the WT strain 70

4.5 Conclusions **72**

Part of this chapter was submitted as:

De Bruyn, F., De Paepe, B., Beauprez, J., Maertens, J., De Cocker, P., Stevens, C., Mincke, S., Soetaert, W., and De Mey, M. (2014) **Development of an *in vivo* glucosylation platform by coupling production to growth: Glucosylation of phenolic acids by GT of *V. vinifera*.**

Spectral analyses for the determination of the glucosides were carried out by C. Stevens and S. Mincke.

4.1 Abstract

Glycosylation of small molecules can significantly alter their properties such as solubility, stability and/or bioactivity, making glycosides attractive and highly demanded compounds. Consequently, many biotechnological glycosylation approaches have been developed, with enzymatic synthesis and whole-cell biocatalysis as the most prominent techniques. However, most processes still suffer from low yields, production rates and inefficient UDP-sugar formation. To this end, a novel metabolic engineering strategy is presented for the *in vivo* glucosylation of small molecules in *Escherichia coli* W. This strategy focuses on the introduction of an alternative sucrose metabolism using sucrose phosphorylase for the direct and efficient generation of glucose 1-phosphate as precursor for UDP-glucose, and fructose, which serves as a carbon source. By targeted gene deletions, a split metabolism is created whereby glucose 1-phosphate is rerouted from glycolysis to product formation (i.e. glucosylation). Further, the production pathway was enhanced by increasing and preserving the intracellular UDP-glucose pool. Expression of a versatile glucosyltransferase from *Vitis vinifera* (VvGT2) enabled the strain to efficiently produce 14 glucose esters of various hydroxycinnamic and hydroxybenzoic acids with conversion yields up to 100 %. To our knowledge, this fast growing (and simultaneously producing) *E. coli* mutant is the first versatile host described for the glucosylation of phenolic acids in a fermentative way using only sucrose as a cheap and sustainable carbon source.

4.2 Introduction

In nature, the majority of specialized (secondary) metabolites exist in its glycosylated form. Addition of a sugar residue to such molecules significantly alters the solubility, stability and/or bioactivity[20], which are desirable properties when glycoconjugates are applied as nutraceuticals, therapeutics or cosmetics. Small (lipophilic) molecules such as flavonoids, phenolic acids, alkaloids or terpenoids are an interesting class within plant specialized metabolites due to their myriad of applications ranging from cardioprotective[42], anticarcinogenic[40] and antiviral[30] to antioxidative[44] or sweetening[25] effects. Furthermore, glucosylation of these compounds, which is catalyzed by glucosyltransferases (GTs), proves to be the key factor in metabolic homeostasis[13] of plants. These enzymes transfer a glucose residue regio- and stereoselectively from the activated sugar donor UDP-glucose (UDP-glc) to the respective acceptors.[37]

Chemical synthesis of glucosides (and glycosides in general) has proven to be a very daunting task, requiring many protecting and deprotecting steps.[75] In addition, large-scale synthesis is hampered by a low yield and a low atom efficiency. To tackle these difficulties, many enzymatic processes have been developed using either glycoside hydrolases, or phosphorylases or the mentioned GTs for the generation of various glucosides.[78] However, whereas glycoside hydrolases often suffer from an unfavorable equilibrium, typically shifted towards hydrolysis rather than the synthesis reaction which results in low yields[86], glycoside phosphorylases require high concentrations of acceptor molecules for efficient catalysis[91]. The main constraint to the use of GTs in biocatalysis is the need for equimolar amounts of UDP-glc, which is very expensive and rarely available in large quantity.[92] To overcome this issue, a lot of effort has been invested both in the enzymatic regeneration of these nucleotide sugars[163,181] and in coupling GTs with sucrose synthase[130]. Despite these efforts, production rates remain low and the overall processes costly. An alternative approach is the use of microbial whole-cell systems for the *in vivo* glucosylation of various acceptors.[98,102,124,256] The used (engineered) cells generate the aforementioned expensive cofactors *in situ* and the expression of GT enzymes, thus bypassing the need for extensive protein purification.

Most of these microbial processes are bioconversions which use resting cells, hence resulting in inefficient UDP-sugar formation.[114,257] They typically require an initial growth phase and expensive growth media to obtain high cell densities that function as passive biocatalysts during the subsequent production phase. Furthermore, the majority of glycosylation systems to date rely on phosphoglucomutase (Pgm) for the generation of glucose 1-phosphate (glc1P) as precursor for UDP-glc, despite the fact that this

reaction has been frequently identified as a major bottleneck.[122] As a result, most *in vivo* glycosylation strains are plagued with low product titers and rates, rendering them unfit for large-scale applications. However, production rates and space-time yields could be greatly improved when formation of glycosides occurs simultaneously with biomass accumulation. Reassessing existing engineering strategies in such a way that the carbon source is equally invested in growth and production bypasses the need for a trade-off between both. Since growth is the main driving force of a cell, a steady supply of the UDP-sugars and its precursors is ensured.

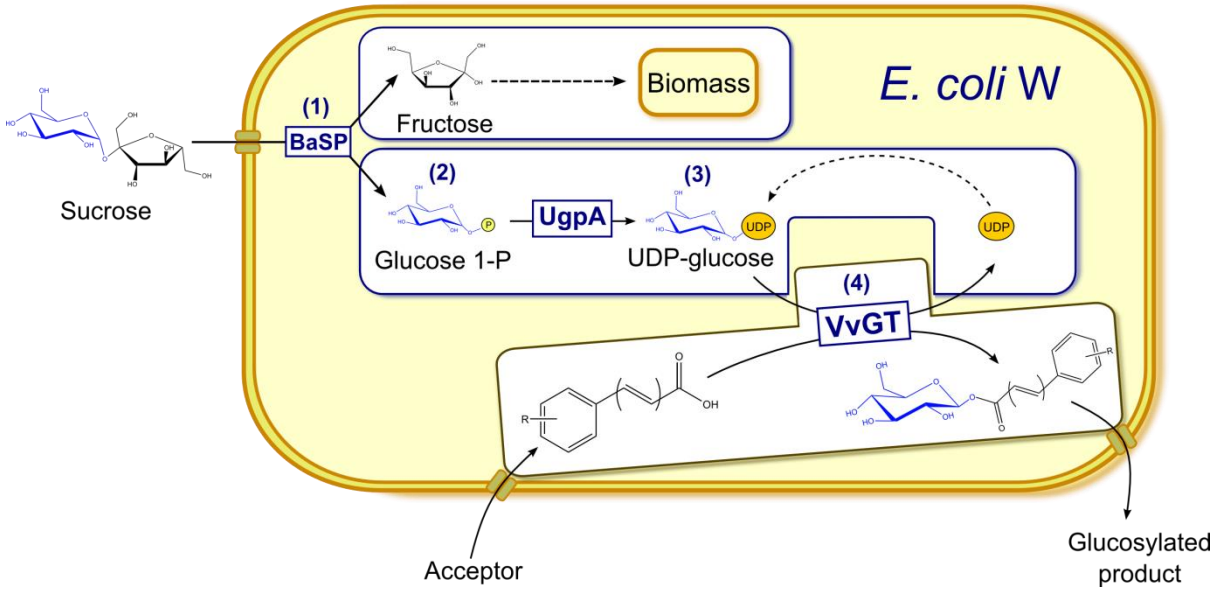


Figure 4.1: Overview of the novel metabolic engineering strategy which consists of (1) introducing sucrose phosphorylase (BaSP) in *E. coli* W for the direct generation of glucose 1-phosphate; (2) creation of a split metabolism by preventing glucose 1-phosphate degradation and rerouting it towards UDP-glucose; (3) formation and preservation of an enlarged UDP-glucose pool; and (4) the expression of a versatile glucosyltransferase VvGT2 for the glucosylation of various phenolic acids.

In this contribution, a novel metabolic engineering strategy that couples production to growth is presented for the *in vivo* glucosylation of small molecules in *Escherichia coli* W. This strategy, depicted in Figure 4.1, focuses on (1) the introduction of an alternative sucrose metabolism, using sucrose phosphorylase, for the efficient generation of glc1P, the precursor of UDP-glc; (2) the creation of a split *E. coli* W metabolism, resulting in the reduced or eliminated use of glc1P for growth, preserving it for UDP-glc formation; (3) enhancing and securing UDP-glucose formation; and (4) the expression of the versatile glucosyltransferase VvGT2 from *Vitis vinifera* for the glucosylation of various phenolic acids. To our knowledge, these compounds have not been the subject of *in vivo*

glucosylation in spite of their interesting properties as (dietary) antioxidants, which hold a large and steadily growing market.[62] Due to this engineering strategy, sucrose is split in fructose, which is used as a carbon source, and glc1P, which fuels the formation of glucosides, resulting in a strain that simultaneously grows and produces.

4.3 Materials and methods

4.3.1 Materials and Molecular Agents

All plasmids used were constructed using Gibson assembly[237] and all PCR fragments were amplified using Q5 polymerase from New England Biolabs (Ipswich, Massachusetts). Oligonucleotides were purchased from IDT (Leuven, Belgium). The plasmids and bacterial strains used in this study are listed in appendices A1 and A2. A list of primers for the creation of gene knockouts and a list of primers for the cloning of the expression plasmids are given in appendices A5 and A6, respectively. *E. coli* DH5α was used for plasmid cloning and propagation, while *E. coli* W was used for expression of the production plasmids and creation of gene knockouts. β-glucogallin was purchased from Carbosynth (Berkshire, UK). All other chemicals used were purchased from Sigma Aldrich (Germany) unless otherwise indicated.

4.3.2 Creation of the production plasmids

The plasmids for the production of glucosylated compounds were constructed by cloning the genes of interest (sucrose phosphorylase from *Bifidobacterium adolescentis* (*BaSP*), glucosyltransferase from *Vitis vinifera* (*vvGT2*), and glucose-1-phosphate uridylyltransferase from *E. coli* (*galU*) and *Bifidobacterium bifidum* (*ugpA*)) first in a pCX-Kan expression vector with medium-strength constitutive promoter P22 originating from a previously constructed promoter library[59,156,258]. This resulted in the corresponding pCX-Kan-P22-*gene* (pBaSP, pVvGT2, pUgpA and pGalU) plasmids. The gene accession numbers are listed in appendix A3 and an overview of the total cloning strategy is given in appendix A7. The genes with P22 promoter were then amplified and ligated following Gibson assembly[237] into an amplified pUC57 backbone containing 200 bp homologous sequential linkers (L_4 to L_7) flanking the insert, which permits rapid assembly and switching of genes. Assembly of the amplified L4-P22-BaSP-L5 insert and pCX-Kan backbone led to the construction of pCX-Kan-L4-P22-BaSP-L5. This plasmid was used for the consequent (homology based) stitching and creation of the final constitutive production plasmids pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6 (pBaSP/VvGT2) and pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-GalU/UgpA-L7 (pBaSP/VvGT2/GalU or pBaSP/VvGT2/UgpA).

4.3.3 Creation of the *E. coli* W mutant strains

All *E. coli* W knockout mutants were created by using the one step deletion method of Datsenko and Wanner.[236] The genes *cscAR*, *pgm*, *agp*, *ushA* and *glgC* were subsequently deleted. Each knockout mutant was transformed with the plasmid pCX-Kan-P22-BaSP

(pBaSP) for the reintroduction and evaluation of the sucrose metabolism and with the plasmids pBaSP/VvGT2, pBaSP/VvGT2/GalU or pBaSP/VvGT2/UgpA for assessing the production potential of glucosylated compounds. *E. coli* W was transformed with pCX-Kan-P22-VvGT2 (pVvGT2) to compare production between the engineered strains and the wild type (WT).

4.3.4 Media

Luria Broth (LB) medium consisted of 10 g L^{-1} tryptone peptone (Difco, Belgium), 5 g L^{-1} yeast extract (Difco) and 10 g L^{-1} NaCl and was autoclaved for 21 min at 121 °C.

Minimal medium contained 2 g L^{-1} NH_4Cl, 5 g L^{-1} $(NH4)_2SO_4$, 3 g L^{-1} KH_2PO_4, 7.3 g L^{-1} K_2HPO_4, 8.4 g L^{-1} MOPS, 0.5 g L^{-1} NaCl, 0.5 g L^{-1} $MgSO_4 \cdot 7H_2O$, 20 g L^{-1} sucrose, 1 mL L^{-1} trace element solution and 100 µL L^{-1} molybdate solution. The medium was set to a pH of 7 with 1 M KOH and filter sterilized.

Trace element solution consisted of 3.6 g L^{-1} $FeCl_2 \cdot 4H_2O$, 5 g L^{-1} $CaCl_2 \cdot 2H_2O$, 1.3 g L^{-1} $MnCl_2 \cdot 2H_2O$, 0.38 g L^{-1} $CuCl_2 \cdot 2H_2O$, 0.5 g L^{-1} $CoCl_2 \cdot 6H_2O$, 0.94 g L^{-1} $ZnCl_2$, 0.0311 g L^{-1} H_3BO_4, 0.4 g L^{-1} $Na_2EDTA \cdot 2H_2O$, 1.01 g L^{-1} thiamine · HCl. The molybdate solution contained 0.967 g L^{-1} $Na_2MoO_4 \cdot 2H_2O$.

Production medium has the same composition as the minimal medium but the pH is set to 6.5 and MES buffer is used instead of MOPS. Gallic acid (0.5 g L^{-1}) was added as acceptor and 1 g L^{-1} citric acid to prevent medium auto-oxidation.

4.3.5 Growth in shake flasks

E. coli W mutant precultures were grown in 50 mL glass tubes containing 5 mL LB medium with the necessary antibiotics (50 µg mL^{-1} kanamycin) for maintenance and selection of the various plasmids used. Cultures were grown for 16 h at 37 °C and 200 rpm and used for the 2 % inoculation of 100 mL minimal medium in 500 mL shake flasks. Experiments for glucogallin production were carried out in production medium. Glucosylation of various hydroxycinnamic acids (HCA) and hydroxybenzoic acids (HBA) was evaluated by adding 1 g L^{-1} of these compounds to the minimal medium. All shake flask cultures were grown at 37 °C and 200 rpm and samples were collected at regular intervals for the measurement of extracellular metabolites and optical density at 600 nm (OD_{600}) using a Jasco V-630Bio spectrophotometer (Easton, UK).

4.3.6 Growth in bioreactors

To evaluate the metabolic parameters, 100 mL cultures were grown overnight on minimal medium and used to inoculate (5 %) Biostat B+ reactors (Sartorius Stedim, Germany) with a 1 L working volume. The fermentor culture was maintained at 37 °C and 4 M of KOH and 0.5 M H_2SO_4 were used to keep the pH to 6.8. The gas flow was fixed at 1 vvm and the agitation ranged from 200 to 1200 rpm as function of the dissolved oxygen (kept above 40 %). Off-gas analysis (EGAS-1, Sartorius AG, Germany) provided on-line measurement of O_2 and CO_2. Growth experiments were performed in minimal medium without MOPS buffer and all parameters described above were monitored using MFCS/win software (Sartorius AG).

4.3.7 Product analysis and quantification

Culture samples containing HCA or HBA were primarily analyzed by TLC on Silica gel 60 F_{254} precoated plates (Merck, Germany). The plates were run in a solvent system of ethyl acetate/butanone/formic acid/water in a volume ratio of 5:3:1:1 in a closed TLC chamber and were developed using standard visualization techniques or agents: UV fluorescence (254 nm) or by staining with 10 % (v/v) H_2SO_4 and subsequent charring. If no glucoside standard was available, acceptor and product spot intensities were processed and quantified using ImageJ[259]. β-glucogallin and gallic acid were simultaneously detected with the method of Takemoto *et al.*[260] using a Prevail Organic Acids column (5 µm, Grace, USA) at 30 °C and a Varian HPLC system (Agilent technologies, California). Sucrose, fructose and glucose were detected using an X-bridge Amide column (35 µm, Waters, USA) at 35 °C. The mobile phase consisted of a mixture of 75 % acetonitrile, 24.8 % water and 0.2 % triethylamine at a flow rate of 1 mL min^{-1}. The accumulation of extracellular glucose 1-phosphate (glc1P) was analyzed by an enzymatic assay as previously described by De Bruyn *et al.*[156]

4.3.8 Purification and structural elucidation of compounds

Glucosylated compounds were purified from the culture broth using centrifugation of the culture broth (10 min at 5000 × g) and adding 1.5 g of activated carbon to 50 mL of the supernatant. After 3 min of adsorption, the mixture was poured onto a glass sintered filter and subsequently washed with 300 mL of water to remove salts and monosaccharides, 120 mL of 10 % EtOH to remove disaccharides, 90 mL of 20 % EtOH and 60 mL of 30 % EtOH. The HBA or HCA glucose esters were eluted with 3 times 30 mL of 50 % EtOH. Ethanol was removed from the purest fractions by evaporation at 40 °C and the samples were freeze dried (Christ Alpha 1-4 LSC Freeze Dryer, SciQuip, UK). The only exception to this procedure was the glucoside of syringic acid which

crystallized at 4 °C and was filtered off and washed with cold water. ^{1}H NMR (300 MHz) and ^{13}C NMR (75 MHz) spectra were run with a Bruker Avance III 400 MHz spectrometer (Bruker, Germany) at room temperature. The compounds were diluted in deuterated water, DMSO or acetone, quoted in parts per million (ppm) and referenced to tetramethylsilane (TMS, δ=0) or the appropriate residual solvent peak. The NMR spectra of novel glucosides are listed in appendix A12.

4.3.9 Mathematical methodologies

SigmaPlot (Systat Software, California) was used for fitting the data and calculation of standard errors. Maximal specific growth rate (μ_{max}) was calculated using a three parameter exponential growth model. Biomass yield (Y_{XS}) and product yield (Y_{PX}) were calculated as slopes of a linear regression model. Specific production rate (q_p) was obtained as the product of Y_{PX} and μ during the exponential growth phase.

4.4 Results and discussion

4.4.1 Introduction

In this work, we describe the transformation of *E. coli* W into an efficient glucosylation platform for phenolic acids. *E. coli* W was the bacterial host of choice since it is a safe, fast-growing strain[261] that has a good tolerance for acidic conditions and osmotic stress[262]. Furthermore, it can be easily grown to high cell density cultures[263] and uniquely possesses exotic gene clusters to cope with a wide variety of aromatic compounds[261], thus rendering it a very attractive production host for industrial purposes. A detailed overview of the engineering strategy, which couples production to growth, is shown in Figure 4.1 and will focus on the interesting class of phenolic acids as target for glucosylation. At first, production of β-glucogallin (BGG, 1-*O*-galloyl-β-D-glucose), a glucose ester derived from gallic acid, was chosen as a proof of concept since it has many beneficial properties besides a highly effective and stabile antioxidant. BGG is commonly found in oak (*Quercus sp.*)[264] as the first step towards gallotannin biosynthesis[265] and is proven to be a therapeutic lead to treat diabetic complications[41]. Furthermore it exhibits antibacterial activity[266] and is an efficient UV-photoprotectant[267]. In the following sections, an overview will be given of the various steps of the metabolic engineering strategy to transform *E. coli* W in a platform organism for *in vivo* glucosylation of gallic acid and other phenolic acids.

4.4.2 Altering the sucrose metabolism of *E. coli* W to efficiently generate glc1P

Sucrose is an industrially important and sustainable carbon source for microbial fermentation. In this context, *E. coli* W recently gained attention since it is a fast-growing strain that can utilize sucrose due to the presence of the *cscBKAR* regulon.[261] This regulon, shown in Figure 4.2, consists of a sucrose permease (*cscB*), a fructokinase (*cscK*) and a sucrose hydrolase (*cscA*). *cscR*, however, codes for a protein that represses expression of these sucrose utilization genes at low sucrose concentrations.[268] The *cscA* and *cscR* genes were simultaneously inactivated with the method of Datsenko and Wanner[236] resulting in an *E. coli* W Δ*cscAR* mutant which is unable to grow on sucrose as sole carbon source. Transforming this strain with plasmid pBaSP, which constitutively overexpresses a sucrose phosphorylase from *Bifidobacterium adolescentis*[137], restored growth on sucrose when grown on minimal medium.

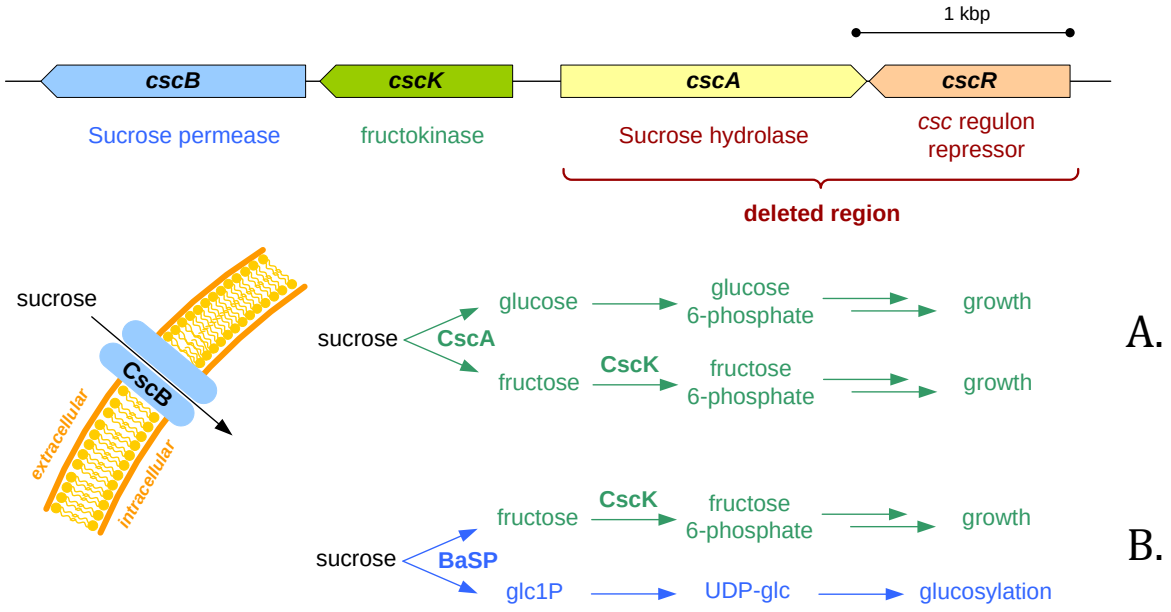


Figure 4.2: Metabolic representation of the *cscBKAR* regulon of *E. coli* W and the various reactions it codes for. A comparison is shown between the native sucrose metabolism (A) and the introduction of an alternative sucrose metabolism by deleting *cscAR* and overexpressing the sucrose phosphorylase BaSP (B).

Both *E. coli* W WT and Δ*cscAR* + pBaSP were cultured and characterized on 1-L bioreactors to ensure no limitation of oxygen. The obtained growth profiles and parameters (Figure 4.3 and Table 4.1) show that the WT strain reaches the upper limit of growth rates reported in *E. coli* W under controlled conditions[261,269]. The maximal specific growth rate μ_{max} of the Δ*cscAR* + pBaSP mutant on the other hand was half the one of the WT strain. However, when the sucrose concentration fell below 10 g/L, the WT slowed down and started to accumulate equimolar amounts of fructose and glucose (not shown), which were rapidly cometabolized when the sucrose concentration fell below 3 g/L. This is caused by the repression effect of CscR at low sucrose concentrations and has been described by Nielsen and coworkers[268]. Conversely, the Δ*cscAR* + pBaSP mutant displays a more constant growth.

Table 4.1: Kinetic parameters of WT and Δ*cscAR* + pBaSP grown on bioreactor scale in minimal medium.

	WT	**Δ*cscAR* + pBaSP**
μ_{max} (h^{-1})	0.92 ± 0.08	0.48 ± 0.04
CDW/OD	0.284 ± 0.001	0.249 ± 0.005
Y_{XS} (g/g)	0.33 ± 0.02	0.29 ± 0.05
q_S (g/g/h)	2.68 ± 0.03	1.75 ± 0.31

In addition, Table 4.1 shows a comparable biomass yield (Y_{XS}) for both strains. This proves the simultaneous metabolization of fructose/glucose for the WT, and fructose/glc1P for the Δ*cscAR* + pBaSP strain, respectively. The Δ*cscAR* + pBaSP mutant initially accumulated 1.8 g/L of fructose, which was subsequently cometabolized with sucrose as depicted in Figure 4.3. This fructose overflow is due to the constitutive expression of BaSP, resulting in the continuous splitting of sucrose into intracellular glc1P and fructose. Since extracellular glc1P levels are much lower, this indicates that glc1P is the preferred substrate and is rerouted towards glycolysis. The extracellular fructose is reconsumed by action of a phosphotransferase system (PTS) for fructose[270], thus bypassing the fructokinase CscK. After consumption of fructose, extracellular glc1P (absent in the WT) started to increase, reaching a maximal concentration of 0.58 g/L. Finally, all glc1P was metabolized and invested as biomass.

4.4.3 Coupling growth and production

Due to the consumption of glc1P, which is intended as a precursor for the formation of UDP-glc and not for growth, various genes need to be knocked out as depicted in Figure 4.1. The major biochemical routes directing glc1P towards the formation of biomass are encoded by the genes *pgm*, *agp* and *glgC*, and the enzymatic reactions they code for are shown in Appendix A4. Pgm is the key switch for converting glc1P into glucose 6-phosphate, rerouting it towards glycolysis, while GlgC forms the precursor ADP-glucose for the synthesis of the storage metabolite glycogen.[271] Agp is an acid glucose-1-phosphatase which is believed to be the main scavenger for glc1P in *E. coli*.[271] Deleting these genes preserves glc1P for UDP-glc formation, hence improving the glucosylation capacity. However, to prevent degradation of accumulated UDP-glc, also *ushA*, which encodes a UDP-sugar hydrolase, was knocked out. The resulting mutant (sGLC) strains are shown in Table 4.2 and were transformed with pBaSP, to gain more insight in the substrate flow and biomass yield.

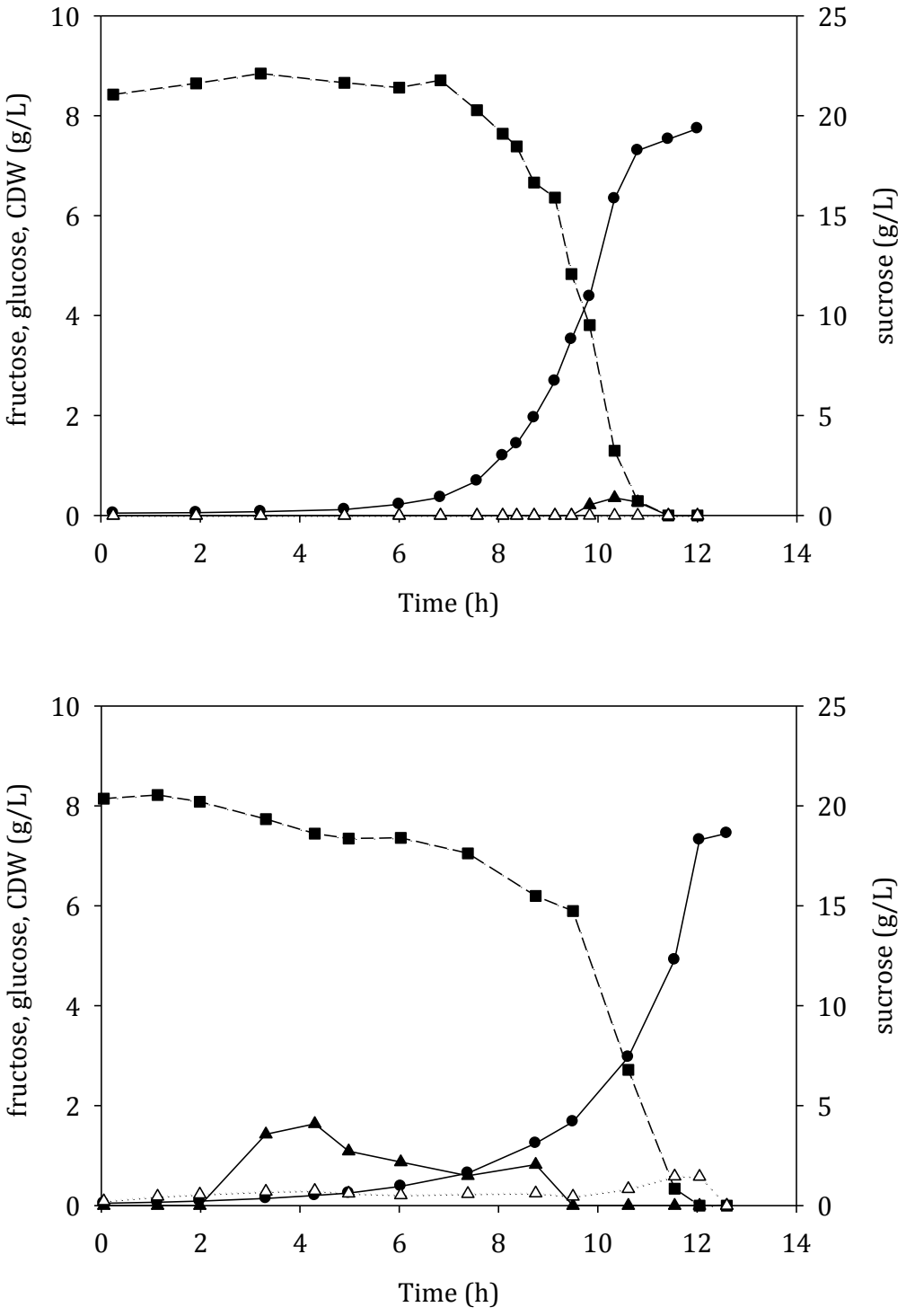


Figure 4.3: Culture growth and sugar utilization of *E. coli* W WT (top) and Δ*cscAR* + pBaSP (bottom) grown in minimal medium on bioreactor scale. Cell dry weight (●) was measured together with the extracellular sugars: fructose (▲), glucose 1-phosphate (△) and sucrose (■).

Table 4.2: Overview of the *E. coli* W mutants used in this chapter. Detailed information on the deleted genes is given in Appendix A2.

Strain	Code
E. coli W Δ*cscAR*	sGLC1
E. coli W Δ*cscAR* Δ*pgm*	sGLC2
E. coli W Δ*cscAR* Δ*pgm* Δ*agp*	sGLC3
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA*	sGLC4
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC*	sGLC5

The transformants were grown and characterized on shake flasks in minimal medium. All strains exhibited growth on sucrose with a μ_{max} comparable to sGLC1 + pBaSP as shown in Figure 4.4, except for sGLC5 + pBaSP which grew significantly faster (μ_{max} = 0.62 ± 0.02). Interestingly, mutants sGLC1 to 3 displayed a similar overall biomass yield (Y_{XS} ≈ 0.31 g/g), indicating that glc1P was still metabolized in spite of the *agp* gene deletion. Strains sGLC4 and sGLC5 displayed a significantly lower Y_{XS} of 0.26 ± 0.01 g/g, illustrating an altered metabolism. This could be explained by deletion of a futile cycle catalyzed by UshA, which hydrolyzes UDP-glc again into glc1P and by inhibiting (Δ*glgC*) glycogen formation, thus preventing glc1P to re-enter glycolysis during glycogen degradation. However, extracellularly accumulated glc1P during growth of sGLC5 on sucrose was again metabolized when the strain entered the stationary phase (data not shown).

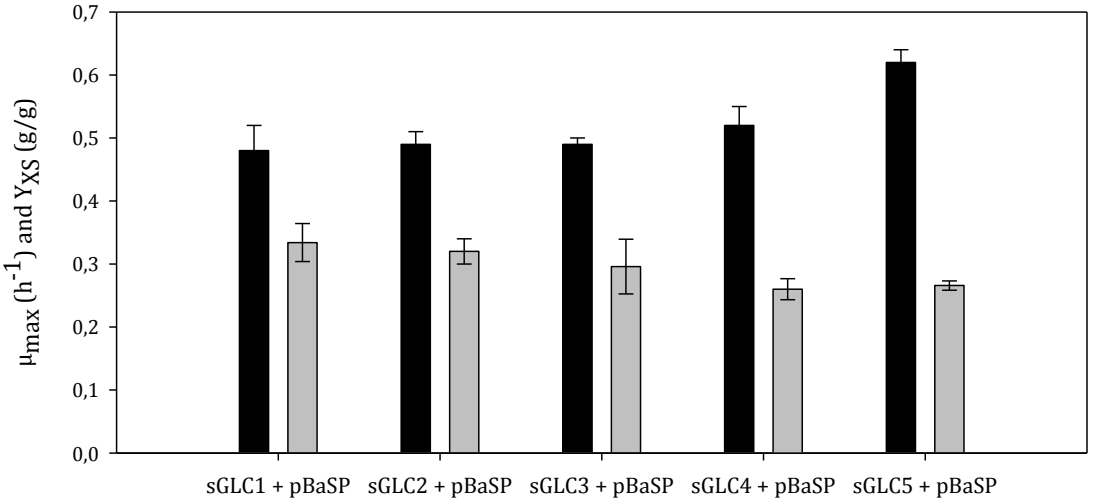


Figure 4.4: Comparison of the maximal specific growth rates μ_{max} (black) and calculated biomass yields Y_{XS} (gray) of the sGLC transformants expressing BaSP when grown on minimal medium. Error bars represent the standard deviation (n = 3).

This observation suggests that other sugar phosphatases are present besides Agp. Possible alternative candidates for glc1P degradation are the sugar phosphatases YihX and YidA, which were first discovered in a high-throughput screen of purified proteins of the haloacid dehalogenase (HAD)-like hydrolases.[272] To verify this, the putative genes *yihX* and *yidA* were additionally knocked out in sGLC5. These mutants were also transformed with pBaSP, resulting in the strains sGLC5 Δ*yihX* + pBaSP and sGLC5 Δ*yidA* + pBaSP. Both strains were cultured in minimal medium and exhibited a decreased μ_{max} of 0.42 ± 0.02 h^{-1} and 0.45 ± 0.01 h^{-1}, respectively. However, the observed biomass yields (0.25 ± 0.02 g/g and 0.26 ± 0.03 g/g) were not significantly different from sGLC5 + pBaSP and glc1P was equally degraded at the end of the exponential phase. Despite these parallel putative sugar phosphatase knockouts, the metabolism was not fully split. It is possible that both phosphatases complement each other so that a double knockout would truly prevent glc1P degradation. In addition, two other poorly studied enzymes may divert glc1P towards glycolysis. A first candidate is the phosphoglucosamine mutase (GlmM) enzyme, which has proven to effectively catalyze the formation of glucose 6-phosphate from glc1P *in vitro*, albeit at a rate 1400-fold lower than that observed with its natural substrate glucosamine 1-phosphate (glc1NP).[273] In addition, it was shown that expression of *glmM* could partially complement a Δ*pgm* mutant strain when grown on galactose, indicating that GlmM catalyzes at least to some extent the *in vivo* interconversion of glucose phosphate isomers.[274] However, since the K_m of GlmM towards glc1P is 10-fold higher than towards glc1NP[273], the role of GlmM as substitute for Pgm should not be overestimated. A second and more interesting candidate is the β-phosphoglucomutase YcjU, which presumably catalyzes the interconversion of β-glucose 1-phosphate (βglc1P) to β-glucose 6-phosphate. It was observed that overexpression of *ycjU* in *E. coli* led to a higher glucoside production than combined overexpression of *pgm*/*galU* when grown on glucose.[275] However, it is very unlikely that GalU (or other nucleotidyltransferases) would exhibit anomeric flexibility as this is highly conserved in nature.[207] A plausible explanation is either a spontaneous interconversion of αglc1P into βglc1P or the existence of a mutatrotase in *E. coli*, thus rerouting glc1P (mediated by YcjU) towards glycolysis.

4.4.4 Optimization of production medium for glucosylation of gallic acid

Prior to evaluating the production potential of the various developed mutants, the ability of the WT strain to produce BGG was investigated. To this end, the WT was transformed with the constructed plasmid pVvGT2, which constitutively expresses a glucosyltransferase from *Vitis vinifera*. VvGT2 catalyzes the formation of 1-*O*-acyl-glucose esters of phenolic acids *in vitro*, with gallic acid (GA) as the preferred substrate.[276] As a preliminary test, 1 g/L of GA was added to minimal medium and

inoculated with the transformant. After 16 h of growth, the supernatant was investigated using TLC and a spot was observed with the same retention as BGG (R_f = 0.5). The compound was purified by preparative TLC and its structure was confirmed by both NMR and MS (687 m/z as a [2M+Na]), thus proving the likelihood of *in vivo* BGG production.

On the other hand, the medium had acquired a dark green to off-brown color after 16h of incubation due to auto-oxidation of GA. This is likely due to Fe^{2+} present in the minimal medium which is complexated by GA, causing a significant drop in the reduction potential with rapid oxidation into Fe^{3+} (in the presence of O_2) as a result.[277] At pH 6, Fe^{2+} is typically complexated by 2 or 3 gallic acid ligands generating a dark blue complex (λ_{max} from 542 to 561 nm) and an oxygen radical species[277] initiating degradation of other GA molecules. Since the optical cell density is measured at 600 nm, this discoloration greatly interferes with a correct estimation of the growth characteristics. To minimize these undesirable effects, different GA concentrations were tested in minimal medium at different pH levels as well as the addition of citric acid as a chelating agent for Fe^{2+}. After 16 h, the absorbance at 400 nm was evaluated. Media containing 5 g/L GA showed significantly more discoloration compared to 0.5 g/L GA (p=0.008, data not shown) and lowering the pH, with addition of citric acid, proved to be the best composition as shown in Figure 4.5.

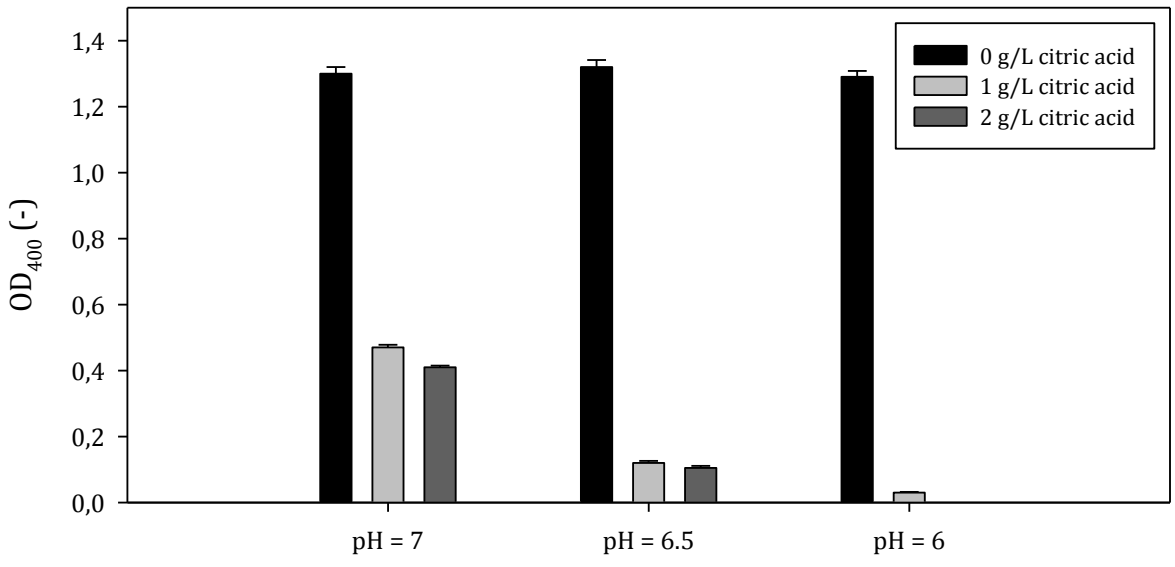


Figure 4.5: Optical density at 400 nm of different minimal media containing 0.5 g/L of GA at pH=7, 6.5 and 6 and addition of 1 or 2 g/L of citric acid (C) after 16 h.

Since GA is a phenolic acid with reported antimicrobial activity against *E. coli* (ATCC 25922)[278,279], a final experiment was conducted for the determination of the maximal specific growth rate at different gallic acid concentrations. The results are shown in Figure 4.6 and indicate no significant inhibition of growth up to 10 g/L. This in contrast to a minimal inhibitory concentration reported of 1 g/L in *E. coli* UB1005[279], which illustrates the adaptation of *E. coli* W to phenolic compounds. Therefore, the composition of the production medium was defined as minimal medium put on pH 6.5 with 1 g/L of citric acid and 0.5 g/L of GA. MES was used instead of MOPS buffer, since its buffer ranges from 5,5 to 6,7. Although pH 6 gave less discoloration, it is less close to the physiological optimum (pH 7). *E. coli* W + pVvGT2 was able to produce BGG at a specific rate of 14.7 ± 4.5 mg/g CDW/h when grown in this production medium and was used as a reference for the engineered strains described hereafter.

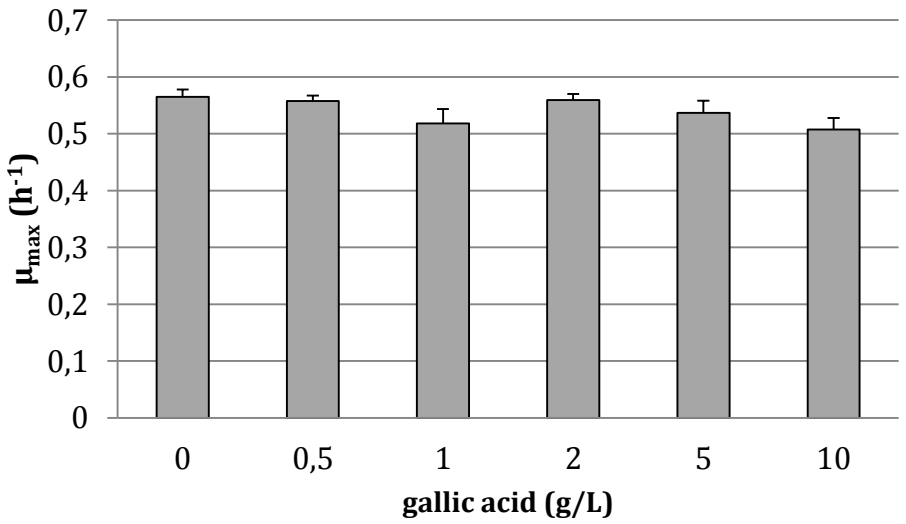
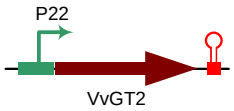
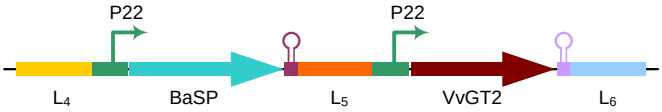
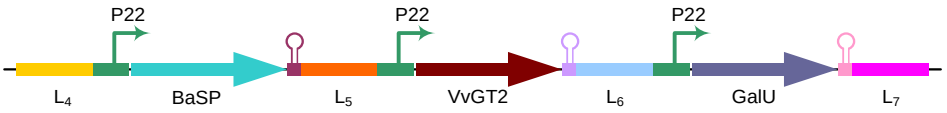
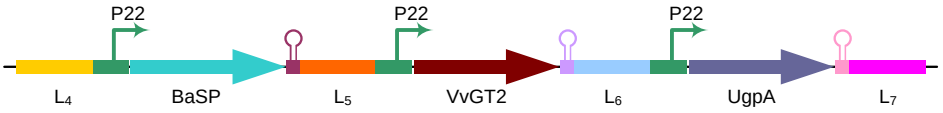


Figure 4.6: Maximal specific growth rates of *E. coli* W at different concentrations of gallic acid in minimal medium pH 6.5 and citric acid.

4.4.5 Production of β-glucogallin

In order to merge the effects of a split metabolism and BGG production, various production plasmids were constructed and are schematically shown in Table 4.3. The use of a constitutive P22 promoter sequence[59] results in uniform expression of the genes and does not require expensive inducers like IPTG. Furthermore, it is very difficult to fine-tune multiple genes carrying inducible promoters, since most expression systems exhibit an all-or-nothing effect.[59]

Table 4.3: Schematic representations of the gene organization in each production plasmid used in this chapter. P22: constitutive promoter, L_n: homologous sequential linker, hairpin: terminator sequence. Gene function is described in Appendix A4.

Plasmid name	Genetic diagram
pVvGT2	
pBaSP/VvGT2	
pBaSP/VvGT2/GalU	
pBaSP/VvGT2/UgpA	

To present a first and fair comparison, all sGLC strains were transformed with the pBaSP/VvGT2 plasmids and grown on production medium. When the cultures reached an OD of 1, 0.5 g/L of GA was added and production of BGG was measured. The calculated specific production rates are shown in Figure 4.7 and unexpectedly depict a higher productivity for the WT compared to sGLC1 to sGLC3. Deletion of *ushA* in sGLC4 however, coding for a promiscuous UDP-sugar hydrolase, resulted in a 10-fold productivity increase. This can be explained by prevention of UDP-glc hydrolysis, which gives rises to an enlarged UDP-glucose pool. Although the precise regulation mechanisms of UshA are not known, its deletion has a profound effect regarding *in vivo* glucosylation reactions. Indeed, a Δ*ushA* mutant is commonly used for the *in vivo* glycosylation of compounds[104,108] though a comparison between WT and this mutant lacks in these studies. Additionally, deletion of *glgC* did not significantly improve the productivity compared to sGLC4, but showed less biological variation.

Up to this point, the UDP-glucose pool in these strains is only formed by action of the native GalU enzyme. Therefore, all sGLC strains were transformed with either the constructed pBaSP/VvGT2/GalU or pBaSP/VvGT2/UgpA plasmid. These plasmids constitutively co-express a UTP-glucose-1-phosphate uridylyltransferase (GalU or UgpA), thus ensuring the formation of UDP-glucose from the accumulated glc1P.[240] Both GalU, native in *E. coli*, and UgpA from *B. bifidum* were studied, since the latter has been proven to exhibit a 7-fold higher activity in crude extracts (Chapter 3) and since the regulation of GalU is not fully understood. The transformants were again grown on

production medium in shake flasks and GA was added when the OD reached 1. Remarkably, sGLC2, sGLC4 and sGLC5 mutants containing pBaSP/VvGT2/GalU didn't reach OD 1 within 24 h. To this end, only the growing strains were evaluated and are

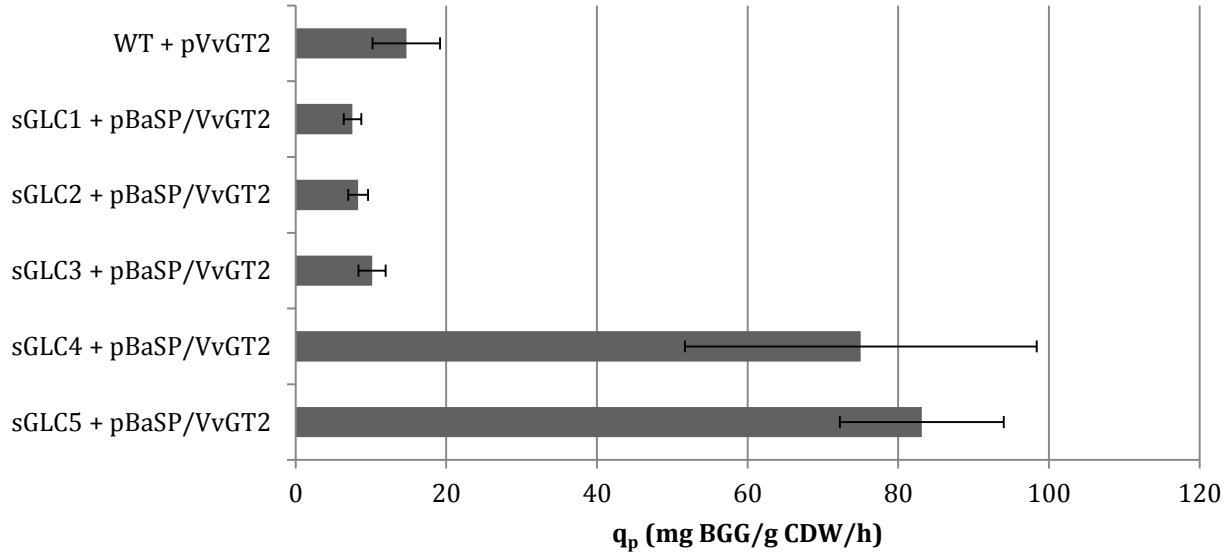


Figure 4.7: Comparison of WT + VvGT2 and sGLC strains expressing both BaSP and VvGT2, grown in production medium on shake flasks. Errors are standard errors (n=2).

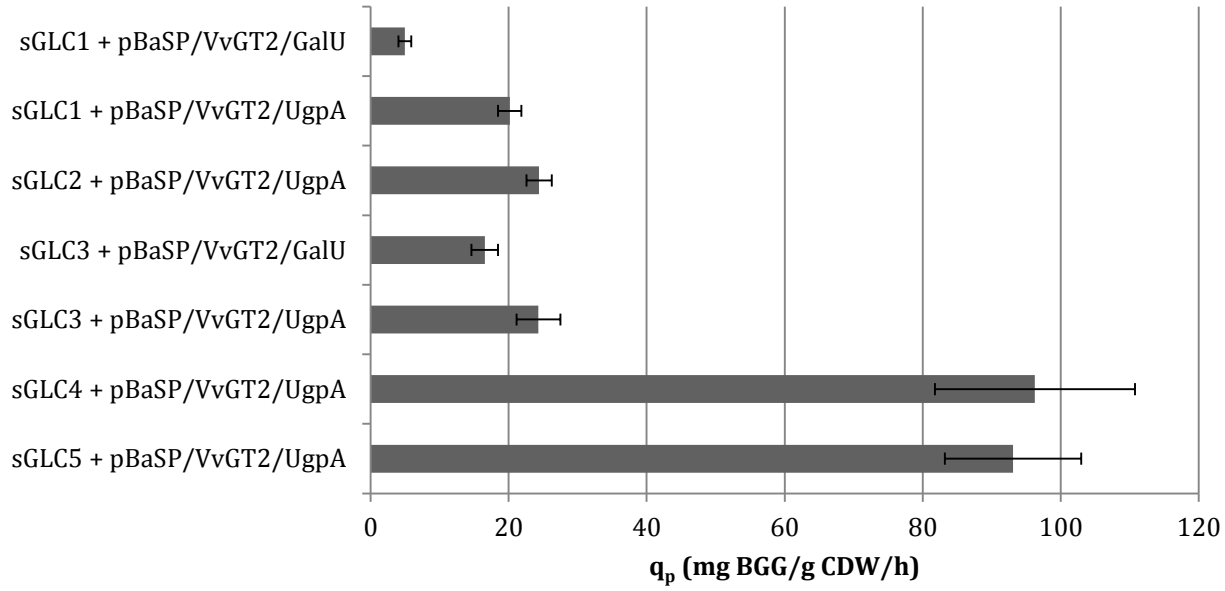


Figure 4.8: Evaluation of the specific productivity of the sGLC strains expressing BaSP,VvGT2 and a uridylyltransferase GalU or UgpA. Strains were grown on production medium on shake flasks. Errors are standard errors (n=2).

are depicted in Figure 4.8. When comparison was available, it was clear that UgpA led to a significantly higher production than GalU. Overexpression of UgpA in mutants sGLC1 to sGLC3 significantly increased the BGG production on average by 2.7-fold, probably resupplying the UDP-glucose pool depleted by VvGT as hypothesized above. However, this was not visible in sGLC4 and sGLC5. Since sGLC5 + pBaSP/VvGT/UgpA showed in general a shorter lag phase and less biological variation when compared to sGLC4, this strain was used for further experiments.

4.4.6 Screening for the glucosylation potential of different acceptors

The used VvGT2 glucosyltransferase was first characterized by Terrier and coworkers[276] and showed broad *in vitro* activity towards many different hydroxycinnamic and hydroxybenzoic acids. The derived glucose esters of these hydroxycinnamic and hydroxy-benzoic acids are often applied as efficient acyl donors for the biosynthesis and thus decoration of numerous plant specialized metabolites.[280,281] Attachment of phenolic residues to plant polymers such as arabinoxylan can induce crosslinking in cell walls via dehydrodimerization and is said to be the initiator for lignification in grasses.[282] Furthermore, these phenolic aglycons have proven to be powerful antioxidants. The trend in replacing synthetic antioxidant food additives with natural ones has boosted the search for biotechnological production alternatives with many success stories.[62] Addition of a sugar moiety to these phenolates can confer extra stability, improved solubility[20] or even organ specific drug delivery[26]. Despite these attractive properties, efficient and economically viable production of these glucosides does not exist and relies on expensive chemical or enzymatic synthesis.[265,283]

Due to the ability of the final production mutant (sGLC5 + pBaSP/VvGT/UgpA) to produce BGG in an efficient way with full conversion, various phenolic acceptors were screened for their glucosylation potential *in vivo*. The glucosylation mutant was grown in shake flasks containing 100 mL minimal medium with 1 g/L of acceptor mentioned in Table 4.4. The acceptor range tested by Terrier and coworkers[276] was extended with vanillic, nicotinic, *p*-aminobenzoic, ferulic, cinnamic, isophthalic, terephthalic, cyclohexanecarboxylic and salicylic acid. The production and growth were monitored during 48 h and 13 out of 15 acceptors showed a spot on TLC with a retention comparable to BGG, except for salicylic and *p*-hydroxybenzoic acid. A newly developed and quick method based on activated carbon adsorption was developed as described in 4.3.8 for the efficient separation of the putative glucose esters. Structures were elucidated and confirmed by NMR sprectoscopy as β-D-glucose esters of the added acceptors. The references to the spectra of previously reported glucosides are shown in Table 4.4 next to the acceptor name. To undoubtedly prove that the acceptors were

glucosylated solely through the action of VvGT2 and not by cryptic genes in *E. coli* W, the WT strain was grown in the same media. The supernatant was visualized on TLC after 48 h and no spot with a retention comparable to the glucose esters was observed.

Table 4.4: Glucosylation potential of the final production mutant towards various phenolic acids.

Acceptor	μ_{max} (h^{-1})	$r_{p,max}$ (mg/L/h)	q_p (mM/g CDW/h)	Conversion [a] (%)
caffeic acid [284]	0.24 ± 0.02	171 ± 9	0.55 ± 0.02	100
sinapic acid [283]	0.22 ± 0.03	111 ± 6	0.53 ± 0.02	82
protocatechuic acid [285]	0.23 ± 0.05	183 ± 12	0.51 ± 0.04	100
p-coumaric acid [284]	0.23 ± 0.01	128 ± 5	0.55 ± 0.01	100
vanillic acid [285]	0.28 ± 0.03	126 ± 6	0.62 ± 0.03	100
nicotinic acid [286]	0.25 ± 0.04	34.6 ± 1.6	0.28 ± 0.02	45
p-aminobenzoic acid [287]	0.11 ± 0.02	137 ± 11	0.63 ± 0.03	100
ferulic acid [283]	0.25 ± 0.03	166 ± 8	0.54 ± 0.03	100
cinnamic acid [288]	0.13 ± 0.02	75.7 ± 4.5	0.45 ± 0.02	47
isophthalic acid	0.19 ± 0.03	84.1 ± 6.3	0.62 ± 0.05	100
terephthalic acid	0.23 ± 0.02	89.7 ± 4.9	0.69 ± 0.02	100
cyclohexanecarboxylic acid	0.11 ± 0.02	144 ± 13	0.58 ± 0.04	100
syringic acid [285]	0.23 ± 0.02	160 ± 8	0.55 ± 0.02	100
salicylic acid	0.03 ± 0.02	-	-	-
p-hydroxybenzoic acid	0.21 ± 0.02	-	-	-

[a] Maximal conversion is expressed as mM glucoside/mM acceptor added after 48 h hours

Apparently the potential of VvGT2 is as broad *in vivo* as it is *in vitro*, and may be explained by the similarity in the conjugated system of hydroxycinnamic and hydrobenzoic acids. The newly tested nicotinic acid showed 45 % conversion after 48 h, indicating the promiscuity towards a heterocyclic aromatic ring. Remarkably, VvGT2 also showed a high activity towards cyclohexanecarboxylic acid, which was completely converted, proving that even aromaticity is no necessity for VvGT2. To our knowledge, the β-D-glucose esters of terephthalic, isophthalic and cyclohexanecarboxylic acid have not been described before and details about the ^{1}H and ^{13}C NMR spectra are shown in Appendix A12.

To make a fair comparison of the glucosylation potential of the production strain towards various acceptors, the specific production rate q_p was calculated and expressed in mM of glucoside/g CDW/h. Although there is a big difference in growth rate, probably

caused by toxicity of some compounds[289], the calculated q_p values are quite comparable with nicotinic acid as sole exception. The obtained profiles (Appendix A13) showed the coupling of growth to production, which results in a non-linear production rate. The $r_{p,max}$ shown is the maximal production slope during the growth phase. Interestingly, all acceptors could enter the *E. coli* cell and the glucose esters were subsequently secreted back into the medium. An explanation lies in the fact that *E. coli* W is the most versatile host in the degradation of aromatic compounds in comparison to *E. coli* B, C or K12.[261] Degradation routes are not necessarily a drawback since the associated gene clusters often express a promiscuous transporter. Table 4.5 shows the gene clusters responsible for the metabolization or transportation of compounds comparable to the acceptors investigated. The mechanisms of secretion of the glucosides are not known in *E. coli* W, but uptake of β-D-glucosides is often linked with cellobiose uptake (and degradation)[290], which is encoded by the *bgl* regulon[261]. Interestingly, the *bglX* gene codes for a β-glucosidase and has been shown to hydrolyze o-nitrophenyl-β-D-glucopyranoside.[291] Possible hydrolysis (decreasing spot intensity) due to this undesirable glucosidase was observed with production of protocatechuyl-glucose (data not shown).

Table 4.5: Gene clusters and number of genes responsible for the metabolization or active transport of aromatic compounds in *E. coli* W.[292]

Gene cluster	Aromatische componenten	# genes
hpa	3- or 4- hydroxyphenylacetic acid (HPA) and 3,4-diHPA	11
paa	Phenylacetic acid	14
mhp	3-hydroxyphenylpropionic and 3-hydroxycinnamic acid	8
hca	Phenylpropionic and cinnamic acid	5

4.4.7 Productivity comparison with the WT strain

To validate the engineering strategy which envisioned coupling growth and production for the glucosylation of various phenolic acids, the performance of the WT + pVvGT2 was compared to sGLC5 + pBaSP/VvGT2/UgpA. To this end, three representative compounds (caffeic, terephthalic and *p*-aminobenzoic acid) from different phenolic subclasses were added to minimal medium, whereafter the growth and production of the WT + pVvGT2 were followed up over time. The growth and conversion profiles are shown in Appendix A13, clearly illustrating the improved glucosylation potential of the engineered strain sGLC5 + pBaSP/VvGT2/UgpA. As depicted in Figure 4.9, the specific productivity (q_p) increased 3.5 to 35-fold compared to the WT, and all 3 acceptors were fully converted within 48 h. Expression of VvGT2 in the WT strain probably rapidly consumes the intracellular UDP-glc pool, which is not as efficiently regenerated as in the engineered

strain. Furthermore, production is linear over time and continues in the stationary phase, thus resembling a bioconversion process.

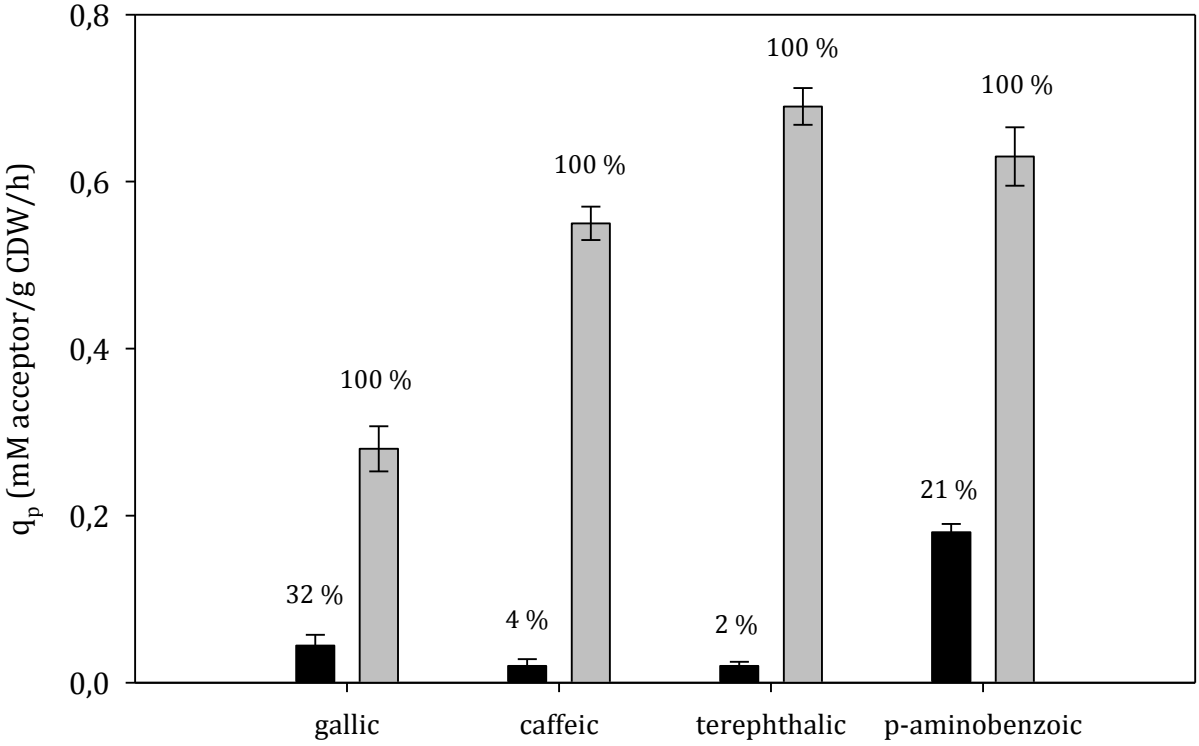


Figure 4.9: Comparison of the specific productivities q_p between the WT + pVvGT2 (black) and the final production strain sGLC5 + pBaSP/VvGT/UgpA (gray). The conversion of the added acceptor after 48 h of incubation is shown above the bars.

When the created platform is compared to other reported *in vivo* processes for the glucosylation of specialized metabolites, the importance of engineering UDP-glc formation becomes clear. All processes display glucoside production rates between 0.001 and 0.01 g/L/h, which is at least a factor 10 to 100 lower than our developed process. As a result, most reported glucoside titers are around 100 mg/L (Table 2.1), with 500 mg/L of vanillin 4-*O*-glucoside[95] and 250 mg/L of arbutin[114] as notable exceptions. Moreover, these systems mostly use expensive growth media (TB, LB) and enzyme inducers to obtain proper glucosylation. Strains with an *ushA* deletion[107,108,110] generally show a 2-fold increase in glucoside production, yet this effect is less significant than the data we obtained, due to indirect glc1P generation from glucose via the bottleneck enzyme Pgm. Despite the possibilities of this novel in vivo platform, economic feasibility and large-scale production should be evaluated. Both on strain level

(complete split of the metabolism) and process level (fed-batch high cell density fermentations and downstream processing), more research will be needed to fully optimize this process. Furthermore, glucosylation of (other) acceptors or increasing their concentration could give rise to issues regarding toxicity, low solubility, transport or by degradation or metabolization of the acceptors by the host. This would require tailor-made solutions on a strain level, by deleting unwanted side reactions or enhancing transport. Toxicity on the other hand can be countered by controlled feeding, continuous extraction of the glucoside from the broth or adaptive evolution of the strain. Moreover, the rationale of the split metabolism is not restricted to E. coli W and can be effectively applied to other industrial hosts.

4.5 Conclusions

In this contribution the systematic creation and evaluation of a glucosylation platform in the form of an engineered *E. coli* strain is described. To this end, the existing sucrose metabolism of *E. coli* W was deleted (Δ*cscAR*) and successfully replaced by overexpressing a sucrose phosphorylase (BaSP). In this way, the released fructose is used as a carbon source for growth and glucose 1-phosphate as a precursor for UDP-glucose. Through specific gene deletions (*pgm*, *agp*, *ushA* and *glgC*), the organism could effectively channel glucose 1-phosphate to UDP-glucose, which was consumed by the overexpressed glucosyltransferase (VvGT2) from *Vitis vinifera*. This enabled the final production strain to efficiently glucosylate gallic acid *in vivo*, yielding the interesting glucose ester β-glucogallin at a specific production rate (q_p) of 93.1 ± 9.9 mg/g CDW/h. Due to the promiscuity of VvGT2, the created strain could easily convert 14 other phenolic acids into their corresponding glucose esters (up to 100 % conversion), resulting in three new-to-nature glucosides. The engineering strategy which coupled production to growth proved to be very effective as the q_p's were up to 35-fold higher when compared to the wild type strain. This novel glucosylation approach could overcome existing economic hurdles and pave the way for large-scale production of glycosides.

Chapter 5

Building a versatile glycosylation platform

Production of bio-active flavonoid glycosides

Contents

5.1 Abstract **75**

5.2 Introduction **76**

5.3 Materials and methods **79**

5.3.1 Materials and Molecular Agents 79

5.3.2 Creation of the production plasmids 79

5.3.3 Creation of the *E. coli* W production mutants 80

5.3.4 Media 81

5.3.5 Growth in shake flasks and sampling 81

5.3.6 Growth in bioreactors 82

5.3.7 Product analysis and quantification 82

5.3.8 Purification and structural elucidation of compounds 82

5.4 Results and discussion **83**

5.4.1 Using *E. coli* W as a host for *in vivo* glycosylation 83

5.4.2 Construction of an advanced sucrose-based glycosylation strain 85

5.4.3 Enhanced production of bio-active quercetin glycoforms 86

5.4.4 Beyond quercetin: *in vivo* glycosylation of other flavonols 89

5.4.5 Production of the oligosaccharide globotriose 91

5.5 Conclusions **96**

Part of this chapter was submitted as:

De Bruyn, F., Van Brempt, M., Maertens, J., Beauprez, J., Duchi, D., Van Bellegem, W., Soetaert, W., and De Mey, M. (2014) **Building a versatile glycosylation platform: production of bio-active quercetin glycosides.**

Spectral analyses were carried out by C. Stevens and S. Mincke.

5.1 Abstract

Flavonoids are bio-active specialized plant metabolites which mainly occur as different glycosides. Due to the increasing market demand, various biotechnological approaches have been developed which use *Escherichia coli* as a microbial catalyst for the stereoselective glycosylation of flavonoids. Despite many metabolic engineering efforts, most processes still display low production rates and titers, thus rendering them unsuitable for large-scale applications. In this contribution, we expanded a previously developed *in vivo* glucosylation platform in *E. coli* W, into a strain for galactosylation and rhamnosylation of the flavonol quercetin. The main rationale of the novel metabolic engineering strategy relied on the introduction of an alternative sucrose metabolism in the form of a sucrose phosphorylase, which divided sucrose into fructose for growth and glc1P as precursor for UDP-glucose. Due to the pivotal role of UDP-glucose, overexpression of interconverting enzymes *galE* and *MUM4* ensured the formation of both UDP-galactose and UDP-rhamnose, respectively. By coupling this *in vivo* UDP-sugar formation with a flavonol galactosyltransferase (F3GT) and rhamnosyltransferase (RhaGT), 0.94 g/L hyperoside (quercetin 3-*O*-galactoside) and 1.12 g/L quercitrin (quercetin 3-*O*-rhamnoside) could be obtained, respectively. These strains, effectively coupling growth and production, also showed activity towards other promising dietary flavonols like kaempferol, fisetin, morin and myricetin. Finally, the flexibility of the galactosylation platform to produce specialty oligosaccharides from lactose was also assessed. Replacing F3GT with a lipo-oligosaccharide galactosyltransferase (LgtC), enabled the strain to efficiently produce 2.8 g/L of globotriose (α-galactosyl-1,4-lactose).

5.2 Introduction

Flavonoids are a large and diverse class of specialized plant metabolites which have attracted much attention in recent years because they have various beneficial effects on human health. Among other, biological activities ranging from anticancer[293] and antioxidant[294] to anti-inflammatory[295], antimicrobial[296] and antiviral[296,297] effects have been attributed to them. As a final step in their biosynthesis, flavonoids are often glycosylated which has a profound effect on their solubility, stability or bio-activity.[13,20] For example, the best studied flavonol quercetin predominantly occurs as different glycosides, which make up to 75 % of our daily flavonoid intake. Over 350 different quercetin glycoforms have been reported to date with varying pharmacological properties.[298,299] In this context, hyperoside (quercetin 3-*O*-galactoside) and quercitrin (quercetin 3-*O*-rhamnoside) have gained a lot of attention as valuable products for the pharmaceutical industry. They are the major glycosides found in apple peels[299,300] and are powerful antioxidants resulting in cytoprotective effects[301-304]. Both have shown to be promising antiviral agents by blocking replication of the influenza virus[305] or inhibiting the viruses hepatitis B[306] and SARS[307]. Furthermore, they have been attributed with anti-inflammatory[29,308], antidepressant[309,310], apoptotic[311] and antifungal[312] activities, rendering them interesting therapeutics resulting in a steadily increasing market demand.

To date, the majority of quercetin and its glycosides are extracted from plant material, which is generally a laborious and low-yielding process requiring many purification steps.[313] *In vitro* plant cell cultures or engineered plants can be used to overcome the low yields and improve production[70-72], however since metabolic engineering of plants is controversial and still in its infancy[73], this approach is often restricted to small-scale production. Although chemical synthesis of quercetin (glycosides) has proven to be feasible[314-316], stereo- or regiospecific formation of glycosidic linkages is often hampered by the presence of various reactive groups[74]. The need for many protecting and deprotecting steps[75], the generation of toxic waste and a low atom-efficiency[77] make these production processes not sustainable nor economically viable.

As a result, the microbial production of specialized (secondary) plant metabolites has made an enormous progress in the last two decades.[56] Advances in the fields of metabolic engineering and synthetic biology have accelerated the sustainable production of flavonoids[61,62,317] in model organisms like *Escherichia coli* and *Saccharomyces cerevisiae*, rendering them in real microbial cell factories. Subsequently, strategies for the *in vivo* glycosylation of flavonoids have also been developed. These are

based on both the overexpression of specific glycosyltransferases, which transfer a sugar residue from an activated nucleotide sugar to a molecule in a stereo- and regioselective way, and the engineering or introduction of the targeted nucleotide sugar pathway. In this way, various quercetin glycosides have already been produced in *E. coli* such as the naturally occurring 3-*O*-glucoside[102], 3-*O*-xyloside[104] and 3,7-*O*-bisrhamnoside[106], or the new-to-nature quercetin 3-*O*-(6-deoxytalose)[318]. Despite these engineering efforts, the reported product rates and titers are still in the milligram range, rendering these microbial hosts unsuitable as industrial catalysts. In addition, since the developed strategies typically rely on bioconversion processes using resting cells, which require a growth phase to gain biomass with a subsequent production phase, it is difficult to improve the production rates or the active generation of the UDP-sugar involved. Furthermore, such systems often entail expensive growth media or the addition of enzyme inducers, making the overall process very costly.

To tackle these problems, we previously developed an efficient platform for the glucosylation of small molecules in *E. coli* W.[319] Through metabolic engineering, a mutant was created which couples production of glucosides to growth, using only sucrose as a cheap and sustainable carbon source. The rationale of the engineering strategy is depicted in Figure 5.1. Here, sucrose is split by a sucrose phosphorylase from *Bifidobacterium adolescentis* (BaSP) into fructose to be used for growth purposes and a glucose 1-phosphate (glc1P) as precursor for UDP-glucose (UDP-glc) formation. To impede the conversion of glc1P into biomass, the endogenous genes for phosphoglucomutase (*pgm*) and an acid glucose-1-phosphatase (*agp*) were knocked out. Glc1P is subsequently rerouted by overexpression of a uridylyltransferase from *Bifidobacterium bifidum* (UgpA) towards UDP-glc, whose metabolization is prevented by knocking out the UDP-sugar hydrolase (*ushA*) and the galactose operon (*galETKM*). This system can however easily be expanded towards other UDP-sugars such as UDP-galactose (UDP-gal), UDP-rhamnose (UDP-rha) or UDP-glucuronate in view of the pivotal role of UDP-glc, which is an ideal starting point for the production of a large variety of UDP-sugars. In this contribution, the system's flexibility is demonstrated by the possibility to galactosylate and rhamnosylate exogenously fed quercetin in *E. coli* W, yielding hyperoside and quercitrin, respectively.

Unlike other systems, our engineered hosts display growth-coupled *in vivo* glycosylation, resulting in gram-scale production of hyperoside and quercitrin. In addition, this novel approach can be used for the production of various other galactosides or rhamnosides. This is of particular interest since rhamnosides play an important role in

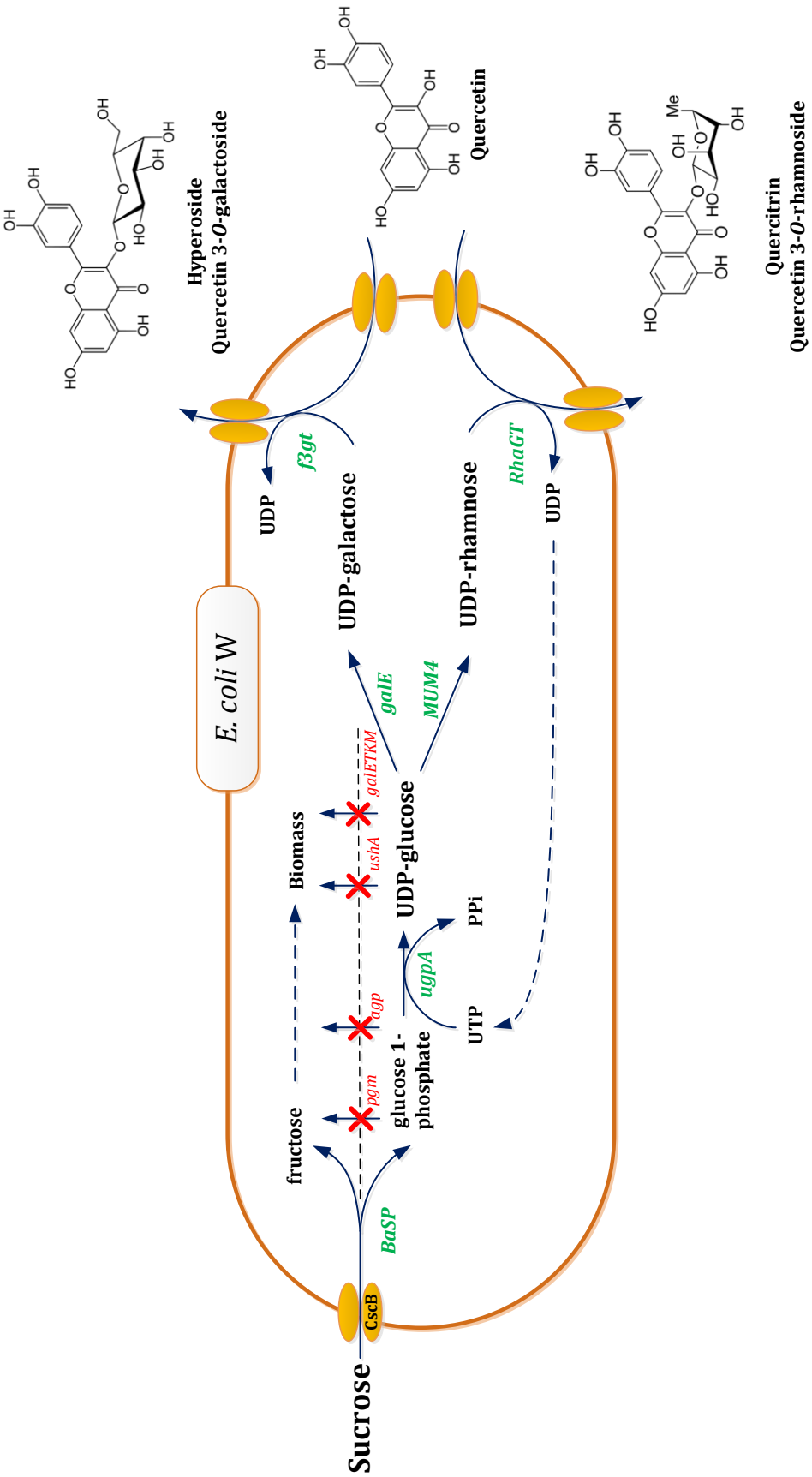


Figure 5.1: Metabolic engineering strategy for the sucrose based glycosylation platform in E. coli W. The production of the bioactive quercetin glycosides hyperoside and quercitrin is chosen to evaluate the versatility of the engineered strains. The same rationale of a split metabolism is applied, whereby sucrose is divided by sucrose phosphorylase (BaSP) in a fructose part for growth and a glucose 1-phosphate part as activated precursor for UDP-glucose. The latter is a universal pivot molecule towards the formation of UDP-galactose and UDP-rhamnose, catalyzed by the interconverting enzymes GalE and MUM4, respectively. To ensure a coupling of production to growth, various genes were knocked out (shown in red), thus channeling the UDP-sugars towards the glycosyltransferases F3GT and RhaGT.

cell and tissue aging[320] and galactosides are targeted towards the liver, making them interesting for site-specific drug delivery[26]. Harnessing the power of this glycosylation strategy into production strains of specialized metabolites will unlock new possibilities for the *de novo* production of glycosides from renewable resources.

5.3 Materials and methods

5.3.1 Materials and Molecular Agents

All plasmids used were constructed using Gibson assembly method[237] or CLIVA[321]. All PCR fragments were amplified using Q5 polymerase from New England Biolabs (Ipswich, Massachusetts). Oligonucleotides were purchased from IDT (Leuven, Belgium). The plasmids and bacterial strains used in this study are listed in appendices A1 and A2. A list of primers for the creation of gene knockouts/knockins and a list for the cloning of the expression plasmids are given in appendices A5 and A6, respectively. *E. coli* DH5α was used for plasmid cloning and propagation, while *E. coli* W was used for expression of the production plasmids and creation of gene knockouts and knockins. Globotriose, hyperoside, quercitrin, isoquercitrin, kaempferol and myricetin were purchased from Carbosynth (Berkshire, UK). All other chemicals used were purchased from Sigma Aldrich (Germany) unless otherwise indicated.

5.3.2 Creation of the production plasmids

First, the expression plasmid for the production of the galactosides hyperoside and globotriose were constructed as depicted in Appendix A9. The constitutive expression plasmid pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7 (pBaSP/VvGT2/UgpA) constructed in Chapter 4 was used as a PCR template for the generation of the backbones. Plasmids pKan-L5-P22-LgtC-L6 (pLgtC) and pKan-L5-P22-F3GT-L6 (pF3GT) were synthesized by GeneArt® (Life Technologies) and used as PCR template for the inserts. LgtC was a C-terminal 25 amino acid truncation of the *lgtC* gene from *N. meningitidis* and the *f3gt* sequence from *Petunia hybrida* was codon optimized. The accession numbers and functions of all the genes used are listed in Appendices A3 and A4, respectively. Gibson or CLIVA assembly of backbones and inserts resulted in the construction of the production plasmids pCX-Kan-L4-P22-BaSP-L5-P22-LgtC-L6-P22-UgpA-L7 (pBaSP/LgtC/UgpA) and pCX-Kan-L4-P22-BaSP-L5-P22-F3GT-L6-P22-UgpA-L7 (pBaSP/F3GT/UgpA). These were subsequently used for the amplification of a pBaSP/LgtC or pBaSP/F3GT backbone. The *galE* and *galE2* sequences were amplified from the genomic DNA of *E. coli* and *Bifidobacterium bifidum*, respectively, and inserted in a pUC-P22 backbone with linkers L4 and L5 flanking the gene as shown in Appendix A9. These intermediate pUC plasmids served as template for the amplification of the L4-

P22-gene-L5 fragment and were inserted using Gibson assembly into the backbones at the position of BaSP. This resulted in the final galactosylation plasmids pGalE/GT/UgpA and pGalE2/GT/UgpA, with GT either LgtC or F3GT.

Plasmids for rhamnosylation were similarly constructed and are shown in Appendix A11. The plasmids pKan-L5-P22-RhaGT-L6 (pRhaGT) and pUC-L4-P22-MUM4-L5 (pMUM4) were synthesized by GeneArt® (Life Technologies) and used as PCR template for the inserts. Sequences *MUM4* and *RhaGT* (*AtUGT78D1*) from *A. thaliana* were codon optimized. pBaSP/VvGT2/UgpA was used for backbone amplification, which was subsequently used for a 3-pieces Gibson assembly with MUM4 and RhaGT, resulting in the rhamnosylation plasmid pCX-Kan-L4-P22-MUM4-L5-P22-RhaGT-L6-P22-UgpA-L7 (pMUM4/RhaGT/UgpA).

5.3.3 Creation of the *E. coli* W production mutants

The overall *E. coli* W knockout mutants were created by using the one step deletion system of Datsenko and Wanner.[236] The sGLC4 strain created in Chapter 4 was used as a starting point to delete the *lacZYA* operon. Subsequently, *lacY* was constitutively overexpressed by knocking in a truncated P22 promoter (Appendix A8). Subsequently, *BaSP,* with constitutive promoter P22, was knocked in flanked by L4 and L5 with constitutive promoter P22 at two different sites resulting in Δ*melA*::L4-P22-*BaSP*-L5 and Δ*glgC*::L4-P22-*BaSP*-L5. The strategy is schematically shown in Appendix A10. L4-P22-*BaSP*-L5 was amplified from its corresponding pUC vector and the homologous regions adjacent to *melA* and *glgC* were amplified from the gDNA of *E. coli*. Using built-in homologous overlaps, a 3-pieces Gibson assembly was carried out and the resulting mixture was used as a template for amplification. The obtained knockout/knockin linear DNA was used to transform sGLC1 and sGLYC1 (*E. coli* W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY*, Appendix A2) according to the Datsenko and Wanner protocol. The transformants were plated on minimal sucrose medium agar plates and grown overnight. Finally, the *galETKM* operon was deleted, yielding the final base strain *E. coli* W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*melA*::L4-P22-*BaSP*-L5 Δ*galETKM*.

These strains and the *E. coli* W wild type were transformed with the production plasmids described above, resulting in the galactosylation (sGAL) and rhamnosylation (sRHA) strains given in Table 5.1. Additionally, strains for the production of globotriose as a proof of concept were also created and denoted as sGLOBO.

Table 5.1: Overview of the quercetin galactosylation (sGAL) and rhamnosylation (sRHA) mutants used in this chapter. Strains for globotriose production are denoted as sGLOBO.

Strain	Code
E. coli W + pF3GT	sGAL1
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE/F3GT/UgpA	sGAL2
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE2/F3GT/UgpA	sGAL3
E. coli W + pRhaGT	sRHA1
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pMUM4 + pRhaGT	sRHA2
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pMUM4/RhaGT/UgpA	sRHA3
E. coli W + pLgtC	sGLOBO1
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* + pBaSP/LgtC/UgpA	sGLOBO2
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galTKM* + pBaSP/LgtC/UgpA	sGLOBO3
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE/LgtC/UgpA	sGLOBO4
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE2/LgtC/UgpA	sGLOBO5

5.3.4 Media

LB medium and minimal (sucrose) medium are described in Chapter 4. Minimal medium agar plates with sucrose (50 g/L) had the same composition as minimal medium, but contained 15 g/L of agarose. The agarose and salts were autoclaved separately at 121 °C for 21 min. Sucrose was filter sterilized through a 0.22 µm corning filter (Fisher, Belgium) and heated for 1 min in a microwave oven at 800 W prior to mixing it with the warm agarose and salt solutions. 1 mL/L of mineral solution (described in Chapter 4) was sterilely added prior to pouring the plates.

5.3.5 Growth in shake flasks and sampling

E. coli W mutant precultures were grown in 5 mL LB medium with the necessary antibiotics (50 µg/mL kanamycin or carbenicillin) for maintenance and selection of the plasmids. The cultures were grown for 16 h at 37 °C and 200 rpm and used for the 2 % inoculation of 100 mL minimal sucrose or lactose medium in 500 mL shake flasks. For the production of hyperoside and quercitrin, quercetin was added to the minimal medium at a concentration of 0.15 or 1.5 g/L. Production of globotriose was achieved by adding lactose to the minimal medium. Growth conditions were the same as described in Chapter 4. Samples were taken at regular intervals from the broth and the supernatant was used for the analysis and quantification of sugars. For the simultaneous analysis of quercetin and its glycosides, 200 µL of the culture was collected and extracted with 800 µL ethyl acetate. The organic layer was collected, evaporated in a SpeedVac™ vacuum concentrator (Thermo Fisher, USA) and dissolved in 200 µL of DMSO for HPLC quantification.

5.3.6 Growth in bioreactors

The description of the bioreactor set-up and conditions is given in Chapter 4. Production experiments were performed on minimal sucrose medium without MOPS buffer and with the addition of quercetin or lactose as acceptors.

5.3.7 Product analysis and quantification

Culture samples were primarily analyzed by TLC on Silica gel 60 F_{254} precoated plates (Merck, Germany). All plates were run in a closed TLC chamber and developed using standard visualization techniques or agents: UV fluorescence (254 nm) or by staining with 10 % (v/v) H_2SO_4 and subsequent charring. The mobile phase for simultaneous detection of the various flavonols and corresponding glycosides consisted of an ethyl acetate:acetic acid:formic acid:water (100:11:11:27 v/v) mixture.[322] If no flavonol glycoside standard was available, product spot intensities were processed and quantified using ImageJ[259]. For the detection of globotriose, lactose and sucrose, an isopropanol:water:NH_4OH mixture was used in a 7:3:2 volume ratio.[93] HPLC quantification of globotriose, lactose, sucrose, fructose and glucose was performed using an X-bridge Amide column (35 μm, Waters, USA) as described in Chapter 4. Quercetin, hyperoside, quercitrin and isoquercitrin were detected with the method described by Pandey *et al*.[104] using a Varian HPLC system (Agilent technologies, California). Mass spectrometry for determination of the various flavonol glycosides was performed with a Micromass Quattro LC (McKinley Scientific, USA). Detection was performed in negative mode ESI-224 MS with a capillary voltage of 2.53 kV, a cone voltage of 20 V, cone and desolvation gas flows of 93 and 420 L/h, and source and cone temperatures of 150 and 350 °C, respectively.

5.3.8 Purification and structural elucidation of compounds

Supernatant from the bioreactors was used for the purification of globotriose and quercetin glycosides. Purification of globotriose consisted of heating the supernatant first to denature the remaining proteins. Activated carbon was added (2 g to 50 mL) to the cleared supernatant and globotriose was separated and purified according to the method of Koizumi *et al*.[93] Quercetin glycosides were extracted from the broth with an equal volume of ethyl acetate after which the organic layer was evaporated to dryness. The remaining products were dissolved in the solvent system described in 5.3.7 and run on a preparative TLC plate. The band containing hyperoside (R_f 0.53) or quercitrin (R_f 0.75) was scraped off, extracted with ethyl acetate and evaporated to yield a bright yellow powder. Products were confirmed by NMR. Spectra were reported elsewhere.[101,323,324]

5.4 Results and discussion

5.4.1 Using *E. coli* W as a host for *in vivo* glycosylation

E. coli W is a fast-growing non-pathogenic strain which tolerates osmotic stress, acidic conditions, and can be cultured to high cell densities, making it an attractive host for industrial fermentations.[261] Moreover, *E. coli* W is able to grow on sucrose as sole carbon source[261], which is an emerging feedstock for the production of bio-products, due to its widespread availability and exceptional performance in life-cycle assessments.[325] In the recent light of developing sustainable processes, *E. coli* W was selected as host for sucrose-based *in vivo* glycosylation reactions. Prior to the production of the glycosides hyperoside and quercitrin in *E. coli* W, the toxicity of the acceptor quercetin was evaluated. To this end, the wild type (WT) strain was grown on minimal sucrose medium containing different concentrations of quercetin ranging from 0.05 to 5 mM. No significant ($p < 0.05$) growth inhibition was observed and all cultures reached a biomass concentration of 1.28 ± 0.09 g CDW/L at the beginning of the stationary phase. This is in contrast to the gradual addition of quercetin to avoid (supposed) lysis of *E. coli* cells as reported by Kim *et al*.[103]

To evaluate the potential of *in vivo* glycosylation, the WT strain was transformed with the plasmids pF3GT or pRhaGT, which constitutively expressed the flavonol 3-*O*-galactosyltransferase from *Petunia hybrid*a or the flavonol 3-*O*-rhamnosyltransferase from *A. thaliana*, respectively. In a first test, 5 mM of quercetin was added to minimal medium and inoculated with the new transformants WT + pF3GT (sGAL1) and WT + pRhaGT (sRHA1). After 16 h of growth, the supernatants of both cultures showed a new yellow product spot when analyzed on TLC. The spot of sGAL1 had the same retention time as a hyperoside standard (R_f = 0.5), which was subsequently purified and confirmed by both NMR and MS to be quercetin 3-*O*-galactoside. However, the product spot obtained after incubation of sRHA1 with quercetin had a different retention factor (R_f = 0.55) than the quercitrin standard (R_f = 0.74). Purification and analysis showed that the compound was isoquercitrin (quercetin 3-*O*-glucoside). As opposed to other reports on wild type *E. coli* strains expressing RhaGT, which simultaneously produced quercitrin (quercetin 3-*O*-rhamnoside) and iso-quercitrin[101,170], no rhamnoside could be detected. Examination of the *E. coli* W genome revealed that the gene cluster responsible for the endogenous production of dTDP-rhamnose is not present[261,271], which functions as an alternative rhamnosyldonor for RhaGT in *E. coli* B and K12 derivatives.[101]

In a more detailed experiment, sGAL1 and sRHA1 were grown on minimal medium with two different concentrations (0.5 and 5 mM) of quercetin. Growth and glycoside

formation were monitored during 30 h. The final titers and specific productivities (q_p) are shown in Figure 5.2. Remarkably, an increase in quercetin concentration resulted in a 2 to 3-fold increase in productivity and titer, indicating that quercetin transport is rate-limiting and crucial for efficient *in vivo* glycosylation. sGAL1 continuously produced hyperoside during the exponential phase, which is reflected in the relatively high specific productivity. On the other hand, sRHA1 only started to accumulate significant amounts of isoquercitrin at the end of the exponential phase. At this stage, the specific growth rate dropped from 0.35 ± 0.04 h^{-1} to 0.06 ± 0.01 h^{-1}, which subsequently affected the corresponding productivities for isoquercitrin.

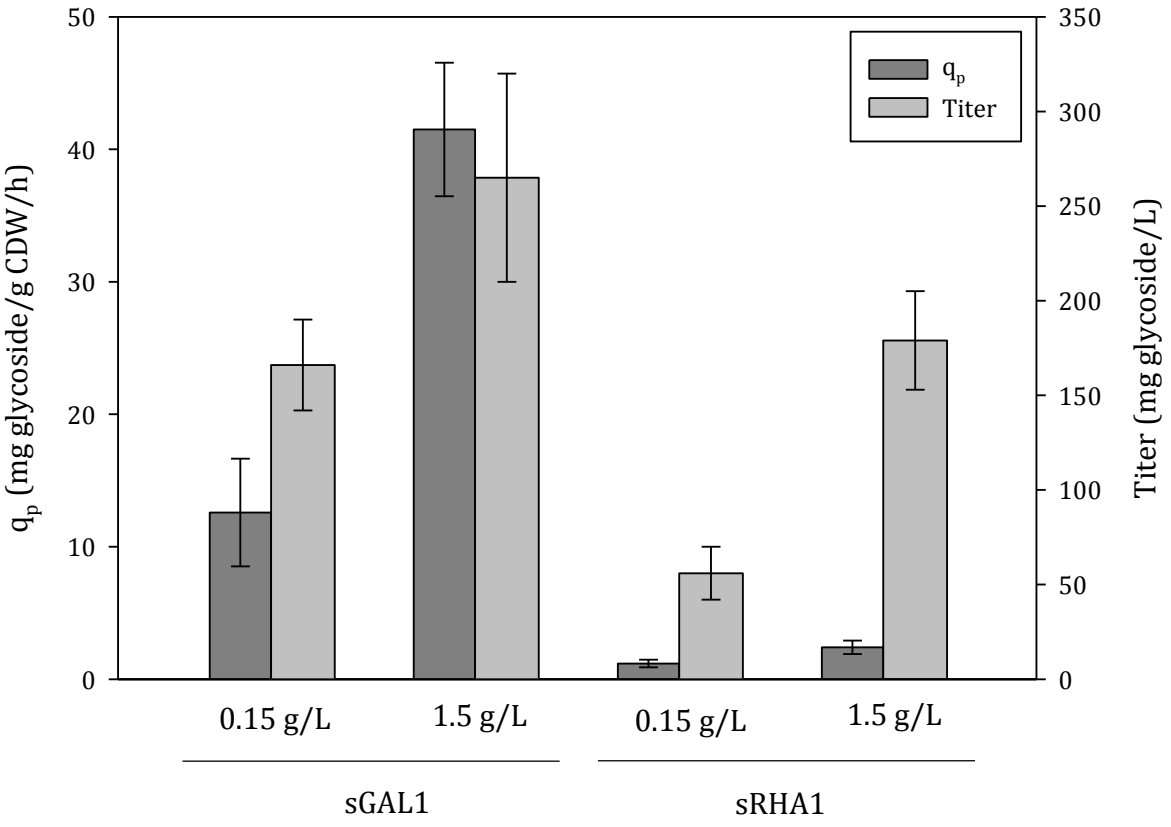


Figure 5.2: Comparison of the specific productivities (q_p) and glycoside titers reached after 30 h of incubation of strains sGAL1 and sRHA1. Strains were grown on minimal medium containing 0.15 or 1.5 g/L (0.5 or 5 mM) of quercetin. The 3-*O*-galactoside and 3-*O*-glucoside of quercetin were produced by sGAL1 and sRHA1, respectively. Errors represent standard deviations (n=2).

5.4.2 Construction of an advanced sucrose-based glycosylation strain

We previously created an *in vivo* glucosylation platform by metabolic engineering of *E. coli* W which was able to simultaneously grow and produce various phenolic glucosides. This growth coupled production is a highly desirable feature for industrial fermentation processes since the need for a carbon source trade-off between growth and production is bypassed. The engineering strategy relies on the introduction of an alternative sucrose metabolism in the form of a sucrose phosphorylase from *Bifidobacterium adolescentis*, BaSP, which splits sucrose into fructose, to be used for growth, and glc1P as precursor for UDP-glc formation. The genes *pgm*, *agp* and *ushA* were deleted to prevent metabolization of glc1P and UDP-glc, thus dividing the routes for growth and production.

In the basic glucosylation strain, BaSP is expressed on a medium-copy plasmid under control of a medium-strong constitutive promoter (P22). To allow a more modular assembly and variation of the genes encoding the glycosyltransferase and the UDP-sugar pathway, these were combined on a single plasmid, while BaSP was integrated in the genome of *E. coli* W. Chromosomal integration may be advantageous because of a significant increase in gene stability, no need for antibiotics to retain the plasmid and a reduced metabolic load. However, expression levels can differ between different integration sites due to structural differences such as supercoiling DNA regions[326,327] which can impair growth. Subsequently, two different DNA sites were assessed for BaSP integration, being *melA* and *glgC*, which encode an α-galactosidase and a glucose-1-phosphate adenylyltransferase, respectively. To this end, an adapted knockin-knockout procedure for large DNA fragments was developed, which is schematically shown in Appendix A10. The BaSP sequence, under control of promoter P22 and flanked by homologous linkers L_4 and L_5, was first knocked in in *E. coli* W Δ*cscAR* to evaluate the effect of the integration site. This resulted in the respective *E. coli* W strains Δ*cscAR* Δ*melA*::L4-P22-BaSP-L5 and Δ*cscAR* Δ*glgC*::L4-P22-BaSP-L5, which were grown on minimal sucrose medium and compared to strain Δ*cscAR* + pBaSP. The resulting maximal specific growth rates (μ_{max}) are shown in Figure 5.3. The influence of the knockin locus on the expression of sucrose phosphorylase and on the maximal specific growth rate is clear. Interestingly, integration at the *melA* position resulted in a strain with a μ_{max} which was not significantly different from its plasmid bearing equivalent. Given these results, the same knockin-knockout method was performed at the *melA* site to yield the final production base strain *E. coli* W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA::*L4-P22-*BaSP*-L5 (sGLYC2, Appendix A2).

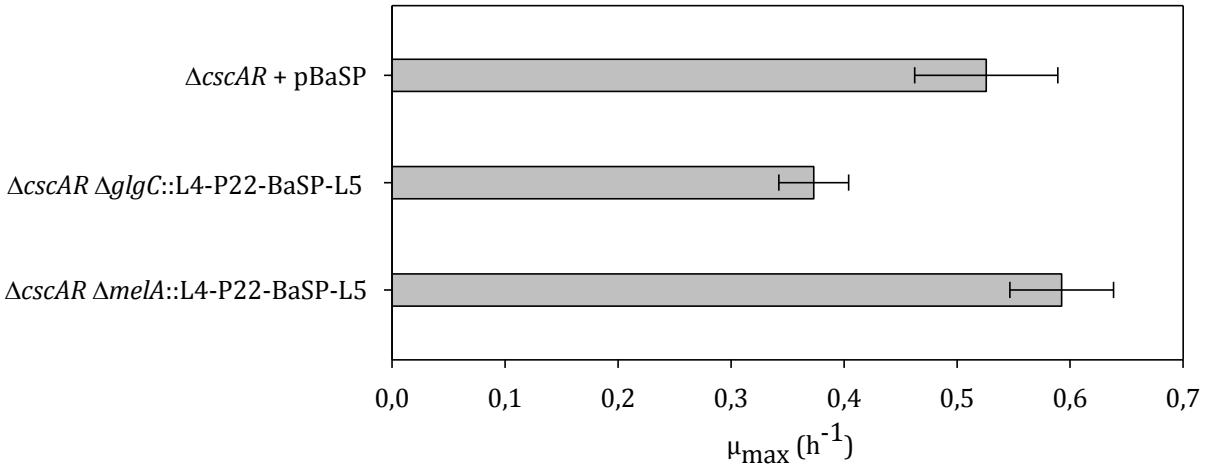


Figure 5.3: Effect of the chromosomal integration site on the expression of the knocked in *BaSP* gene, under control of promoter P22 and flanked by homologous linkers L_4 and L_5. Strains were grown in shake flasks (n=2) and the resulting μ_{max} was compared to *E. coli* W Δ*cscAR* + pBaSP.

5.4.3 Enhanced production of bio-active quercetin glycoforms

In nature, UDP-glc serves a pivotal role in the formation of other UDP-sugars by using various interconverting enzymes.[155] UDP-gal is generated by a UDP-glucose 4-epimerase (GalE), while UDP-rha synthesis is catalyzed by a UDP-rhamnose synthase (MUM4). GalE naturally occurs in *E. coli* W, but due to its tight and complex regulation, an alternative homologous epimerase (GalE2) from *B. bifidum* was also selected and cloned. On the other hand, UDP-rhamnose synthesis is restricted to plants, whereas *E. coli* W is unable to form endogenous dTDP-rhamnose as alternative rhamnosyl donor due to lack of the *rfb* cluster[261,271]. To this end, the *MUM4* sequence from *A. thaliana* was codon optimized and ordered as plasmid pMUM4 to achieve UDP-rhamnose formation in *E. coli* W (Figure 5.1).

Two production plasmid variants pGalE/F3GT/UgpA and pGalE2/F3GT/UgpA were created, as described in materials and methods, to ensure the formation of hyperoside. Both plasmids were transformed in sGLYC2 to yield the galactosylation strains sGAL2 and sGAL3. In a similar way, the rhamnosylation plasmid pMUM4/RhaGT/UgpA was constructed and transformed in sGLYC2, resulting in strain sRHA3. Alternatively, sRHA2 was also created, carrying plasmids pRhaGT and pMUM4. The obtained galactosylation and rhamnosylation strains were grown on minimal medium with two levels (0.5 and 5

mM) of quercetin. Growth and production were monitored to calculate the specific productivities, as shown in Figure 5.4. Again, higher extracellular quercetin concentrations resulted in a 5-fold increase in q_p. However, no significant difference in productivity was observed between sGAL2 and sGAL3 at 5 mM quercetin, indicating that UDP-galactose formation is as efficient with both GalE homologs and not the rate limiting step. The highest hyperoside titer obtained was with sGAL3 (0.94 g/L), which was 3.5-fold higher compared to sGAL1.

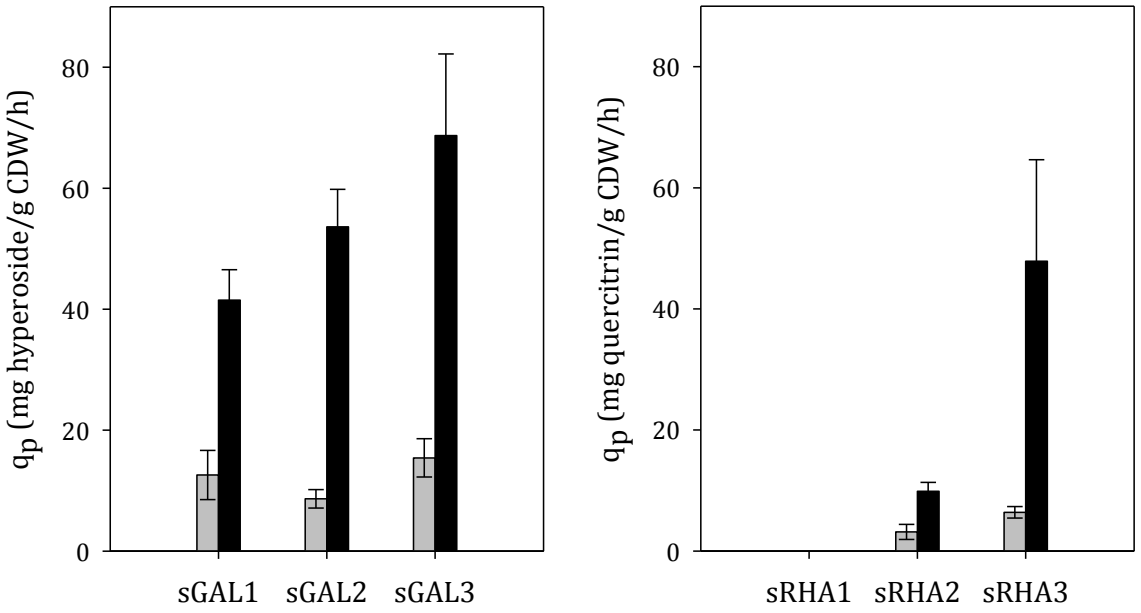


Figure 5.4: Comparison of the specific productivities q_p of the sGAL (hyperoside formation) and sRHA (quercitrin formation) strains. Strains were grown on minimal medium containing 0.5 mM (gray bars) and 5 mM (black bars) quercetin.

In contrast to sRHA1, incubation of sRHA2 and sRHA3 showed a product spot on TLC with an altered retention factor corresponding to quercitrin. Analysis of the compound using MS proved the formation of the 3-*O*-rhamnoside and consequently the *in vivo* activity of MUM4. A quercitrin titer of 1.18 g/L was obtained after 30 h incubation of sRHA3 when 5 mM quercetin was added to the medium, which corresponded to a 53 % conversion. Detailed analysis showed that also 51 mg/L of isoquercitrin was produced extracellularly, with a specific productivity of 3.17 ± 1.01 mg/g CDW/h, showing the preference of RhaGT for UDP-rhamnose when different UDP-sugar donors are present. Remarkably, the specific productivity of sRHA2 was significantly lower (5-fold decrease) compared to sRHA3. Possible explanations are either a higher metabolic burden caused by a two plasmid system or a too low activity of the native GalU, which is not sufficient for adequate formation of UDP-glc. Furthermore, the potential of the novel metabolic

engineering strategy, which couples growth and production, becomes apparent in the form of a 20-fold increase in specific productivity of sRHA3 compared to sRHA1.

To our knowledge, this is the first reported *in vivo* production of hyperoside. To demonstrate the scalability of the engineered strain, strain sGAL3 was cultured in a 1-L bioreactor to ensure a constant pH set at 6.80 and avoid oxygen limitation. A detailed overview of growth, production and consumption of sucrose is given in Figure 5.5. After an observed lag-phase of 5 h (possibly due to shear stress from the impellers), the strain displayed a growth rate of 0.32 ± 0.02 h^{-1} while simultaneously producing hyperoside. A specific productivity was observed of 65.9 ± 2.6 mg/g CDW/h which is comparable to the one on shake flask. When nearly all quercetin was converted, hyperoside formation slowed down. This can be explained by the observed correlation between quercetin concentration and q_p, but also by the reported reversibility of F3GT[328]. A continuous supply of quercetin using a fed-batch mode will likely result in a constant q_p and improve the existing titer.

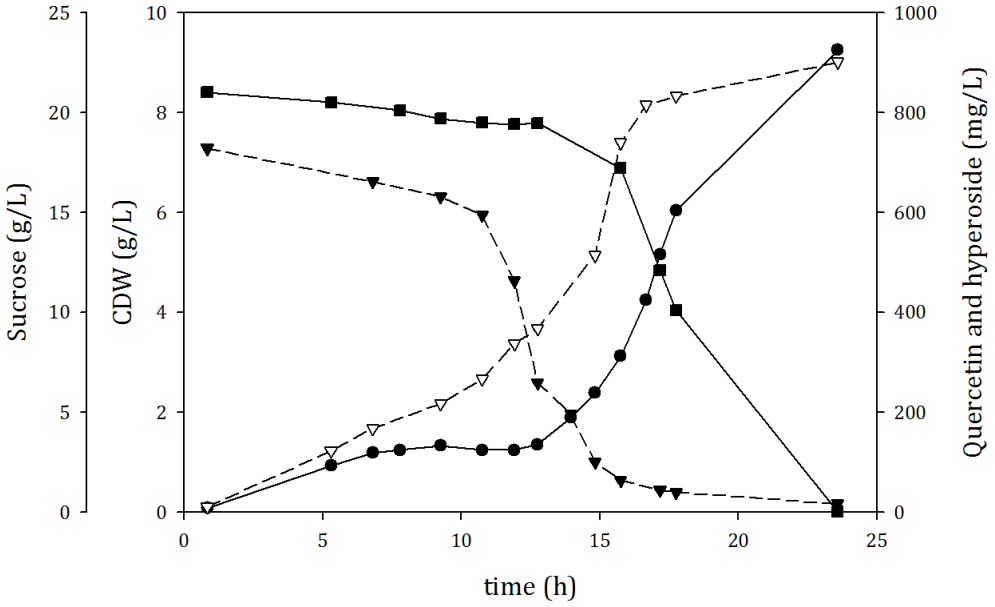


Figure 5.5: Production of hyperoside (▽) in a 1-L bioreactor on minimal medium containing 2.5 mM (755 mg/L) quercetin (▼) using strain sGAL3. Cell dry weight (●) was measured together with extracellular sucrose (■).

To the best of our knowledge, the obtained hyperoside and quercitrin titers are the highest reported to date, which proves the potential of this novel metabolic engineering strategy. A more fair comparison with other *in vivo* processes can be made based on production rates. To this end, the maximal production rates ($r_{p,max}$) of reported processes were compared to those obtained in this contribution, the here obtained

maximal production rate was 6 to 50-fold higher as illustrated in Figure 5.6. However, in view of the growth coupled production obtained with the strains developed, even higher production rates are likely achievable.

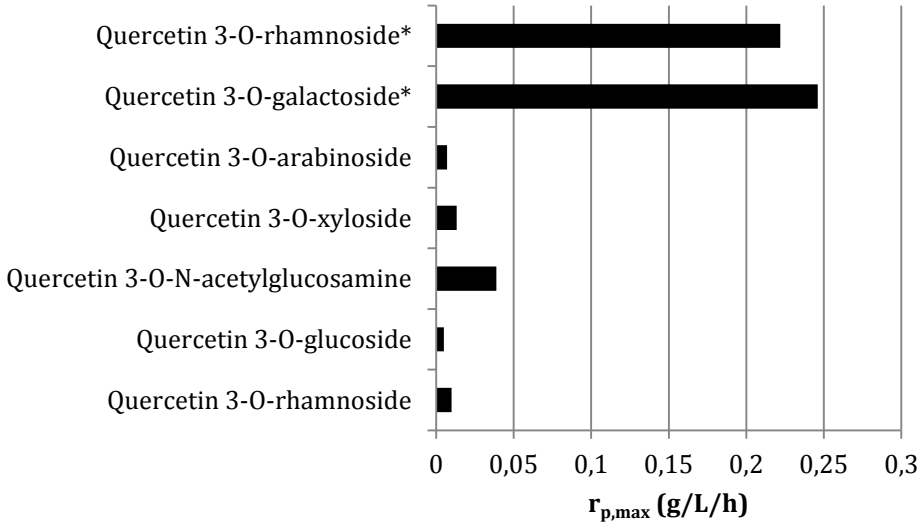


Figure 5.6: Comparison of the maximal production rates obtained using the developed galactosylation and rhamnosylation platform (*) with reported maximal rates found in literature[101-103,105].

5.4.4 Beyond quercetin: *in vivo* glycosylation of other flavonols

Quercetin is the most abundant and best studied dietary flavonol. However, other flavonols such as kaempferol, fisetin, morin and myricetin significantly contribute to our daily flavonoid intake and their beneficial effects are extremely diverse.[329,330] Kaempferol 3-*O*-rhamnoside for example inhibits breast cancer cell proliferation through activation of the caspase cascade pathway[331] while its galactoside has a remarkable anti-inflammatory effect[332]. Myricetin glycosides on the other hand, which naturally occur in the mediterranean shrub *Myrtus communis*, display antioxidant and antimutagenic activity[333], but also exert potent antipsychotic and anxiolytic-like effects[334]. In addition, fisetin can act as a novel inhibitor in prostate cancer management[335] and has proven to be an interesting antiseptic[336]. Finally, morin glycosides have shown to be responsible for the antimicrobial effect in the leaves of *Psidium guajava*[337], whereas its aglycon has been recently identified as a promising insulin-mimetic for diabetes treatment.[338] As the sugar moiety is a major determinant of the intestinal absorption of dietary flavonoids and their subsequent bioactivity[32,339], we investigated the potential of the created *E. coli* W mutants towards both galactosylation and rhamnosylation of various flavonols.

To this end, strains sGAL3 and sRHA3 were grown in tubes with 5 mL minimal medium, each containing 5 mM of kaempferol, myricetin, morin or fisetin. Growth and production were monitored over 48 h and various spots were observed on TLC with similar retention factors as hyperoside and quercitrin. Mass spectrometry was used to identify the compounds, which confirmed the *in vivo* galactosylation of myricetin, kaempferol and fisetin (Table 5.2). All compounds occurred mainly with an m/z of [M+114], illustrating the complexation with trifluoroacetic acid from the mobile phase. The galactoside of morin was produced at a slow rate, which is in accordance to the very low *in vitro* activity of F3GT towards this flavonol.[328] A possible explanation lies in the presence of an unusual hydroxyl group at the 2' position, which may sterically hinder the deprotonation and consequent galactosylation of morin at hydroxyl group 3.[340]

Table 5.2: Galactosylation and rhamnosylation potential of strains sGAL3 and sRHA3 respectively towards other flavonols. Specific productivity q_p and titer reached after 48 h of incubation are shown. ND = not detected.

	3-*O*-galactoside		3-*O*-rhamnoside	
Flavonol	**titer (mg/L)**	**q_p (mg/gCDW/h)**	**titer (mg/L)**	**q_p (mg/gCDW/h)**
Kaempferol	84 ± 14	3.46 ± 0.86	416 ± 37 [a]	12.1 ± 1.4
Myricetin	52 ± 7.1	2.88 ± 0.22	72.3 ± 9.1	2.8 ± 0.5
Morin	34 ± 5.8	1.65 ± 0.15	116 ± 21 [b]	2.5 ± 1.5
Fisetin	134 ± 22	9.32 ± 0.55	403 ± 31 [c]	11.3 ± 0.9

[a] 3-*O*-glucoside was also detected at 52 ± 17 mg/L

[b] 3-*O*-glucoside was also detected at 21.7 ± 6.2 mg/L

[c] 3-*O*-glucoside concentration was lower than 5 mg/L

Incubation of sRHA3 with the different flavonols investigated showed two distinct glycoside spots on TLC, which corresponded to the 3-*O*-rhamnoside and 3-*O*-glucoside. Kaempferol proved to be the best substrate for RhaGT and was predominantly rhamnosylated (8:1 ratio), with a titer exceeding 400 mg/L, which is 2-fold higher than previously reported[101]. Fisetin on the other hand was efficiently glucosylated, yet the formation of its rhamnoside was not as efficient, with a titer below 5 mg/L. A similar preference towards glucoside formation was also observed with myricetin and morin, which indicates that the positioning of the hydroxyl groups is the determining factor for glycosylation with RhaGT. The production of preferred rhamnosides, galactosides or glucosides may be largely improved by using UGTs that are more specific towards certain flavonols and UDP-sugars. Subsequent implementation of these UGTs in the developed *in vivo* glycosylation strains presents an efficient alternative for the large-

scale production of various flavonol glycoforms, which are to date mainly extracted from plant material.

5.4.5 Production of the oligosaccharide globotriose

Besides glycosylated flavonols, the developed glycosylation platform is a flexible tool for the production of other molecules. For example, galactose residues widely occur in diverse oligosaccharide structures such as blood group antigens[341], globosides[342], Lewis antigens[343], galactomannan[344], Galili epitopes[345] and human milk oligosaccharides[346]. Hence, the galactosylation potential of the metabolically engineered strains was investigated for the production of the interesting oligosaccharide globotriose (galactosylated lactose; α-D-Gal-(1→4)-β-D-Gal-(1→4)-D-Glc). Globotriose has promising pharmaceutical properties due to its involvement in numerous pathogenic recognition processes. In addition, it naturally occurs as globotriaosyl-ceramide, is a part of the P1PK Blood Group System, which consists out of three membrane associated antigens and has a function in the immune system related to B cell differentiation.[347] Furthermore, the globotriose moiety of the antigen binds specifically to the Shiga toxin from *Shigella dysenteriae* or Shiga like toxins from pathogenic *E. coli* strains.[348] This process forms the base for diseases such as hemorrhagic colitis and hemolytic uremic syndrome. Hence, the clinical applications of freely soluble (or conjugated) globotriose are as pathogen-capturing drugs to combat for example urinary tract infections.[349,350]

Globotriose production can be achieved by replacing the flavonol galactosyltransferase (F3GT) on the production plasmids by LgtC, a lipo-oligosaccharide α-1,4-galactosyl-transferase from *Neisseria meningitidis*, which is able to galactosylate lactose moieties[351]. When overexpressed however, LgtC is prone to proteolysis[323], so by deleting the C-terminal 25 amino acids of the *lgtC* gene, the stability and expression of this enzyme can be increased. This truncated sequence was flanked by the homologous linkers L_5 and L_6 to allow rapid assembly of the globotriose production plasmids, which were subsequently transformed in the mutant strains as described in Materials and Methods, yielding the sGLOBO strains. A metabolic overview of the globotriose production strategy is shown in Figure 5.7. To avoid hydrolysis of the acceptor lactose, *lacZ* was additionally knocked out. On the other hand, since lactose permease (LacY) activity is necessary for adequate influx of the acceptor lactose[271], the *lacZYA* operon was replaced with a constitutively expressed *lacY* gene (Δ*lacZYA*::P22-*lacY*) as shown in Appendix A8.

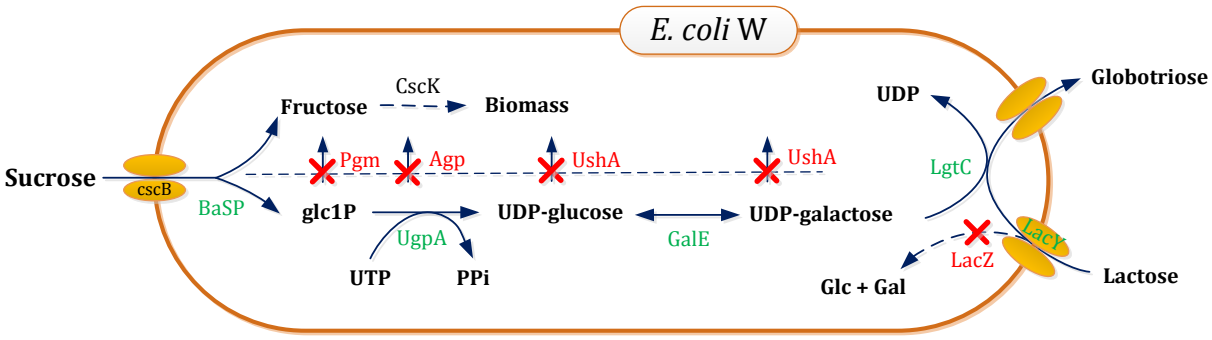


Figure 5.7: Schematic representation of the globotriose production strategy in *E. coli* W. Exogenously fed lactose is transported into the cell by constitutive expression of *lacY* and galactosylated by the galactosyltransferase LgtC. The sucrose phosphorylase BaSP splits sucrose in a fructose part for growth and a glucose 1-phosphate part for the formation of UDP-galactose through UgpA and GalE. Targeted gene deletions (red) prevent the consumption of these intermediates for biomass formation while the *lacZ* knockout prevents degradation of the acceptor lactose.

In a first experiment, sGLOBO1 (*E. coli* W + pLgtC) and sGLOBO2 (*E. coli* W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* + pBaSP/LgtC/UgpA) were grown overnight on minimal medium with 1 g/L lactose. Only visualization of the supernatant of sGLOBO2 on TLC showed a production spot with a similar retention factor as the globotriose standard. This is probably caused by hydrolytic activity of LacZ in sGLOBO1, which prevents the intracellular accumulation of lactose and concomitant globotriose formation. The product was purified from the supernatant with activated carbon and confirmed to be globotriose by MS.

To determine the optimal lactose concentration, sGLOBO2 was grown on minimal medium with lactose ranging from 1 to 10 g/L. The average maximal specific growth rate, which was independent of the different lactose concentrations used, was 0.32 ± 0.02 h^{-1} as shown in Figure 5.8. Conversely, the specific productivity q_p showed a 49 % increase from 1 to 10 g/L lactose, which may indicate that higher extracellular lactose concentrations result in an increased inflow of lactose in the cell. Furthermore, the product to biomass yield (Y_{PX}) was linear during exponential growth, proving growth-coupled production. However, increasing the lactose concentration even further to 50 g/L did not show a significant productivity increase when compared to 10 g/L. Therefore, the lactose concentration in the following experiments was set to 10 g/L.

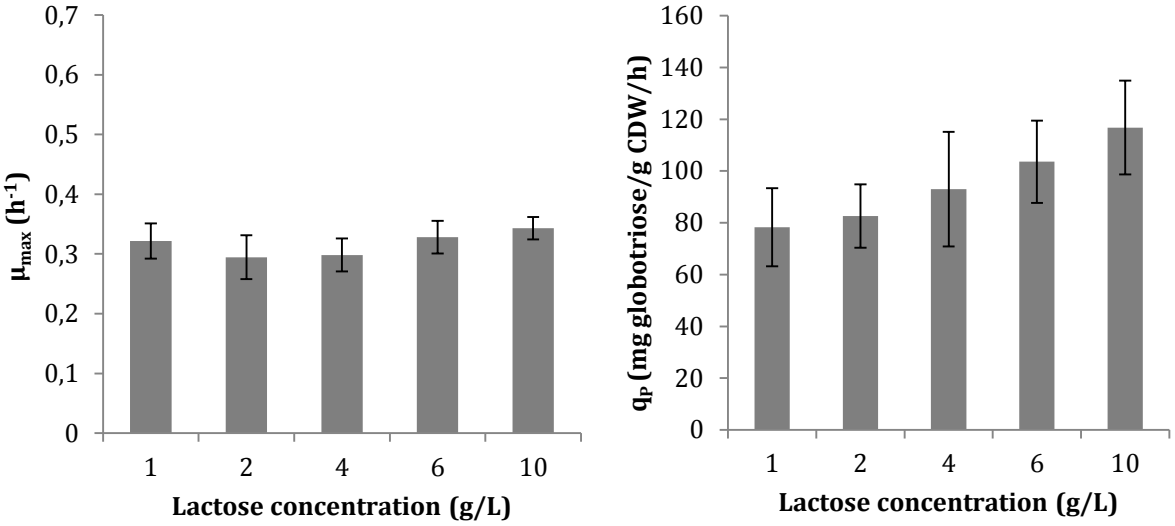


Figure 5.8: Maximal specific growth rates (left) and specific production rates of globotriose (right) of sGLOBO2 in minimal sucrose medium containing different lactose concentrations.

In sGLOBO2, the native GalE activity is solely responsible for the formation of UDP-galactose. To investigate the effects of this operon, three other sGLOBO strains were created whereby either the *galTKM* part (sGLOBO3) or the whole operon *galETKM* (sGLOBO4) was deleted. In the latter strain, *galE* was additionally overexpressed yielding sGLOBO5. The latter two have a constitutive knockin of *BaSP* at the *melA* position, which is believed to code for an α-galactosidase that can hydrolyze globotriose.[49] The strains were grown on minimal medium containing 10 g/L of lactose and the obtained specific productivities are shown in Figure 5.9. Comparison of sGLOBO4 and 5 with sGLOBO3 clearly shows that overexpression of *galE* from *E. coli* or *galE2* from *B. bifidum* does not result in a significantly increased globotriose production.

A plausible explanation could be that constitutive overexpression of *galE(2)* does not result in a higher UDP-galactose pool and only complements for the chromosomal *galE*. It has indeed been observed that expression of *galE* is higher than that of the promoter-distal genes *galTKM*.[352] Another possibility may be that LgtC is the rate limiting step in globotriose formation. On the other hand, it is remarkable that a Δ*galTKM* deletion has a profound impact on globotriose production, which was at least 15-fold lower. This may indicate that ribosomal translation or the stability and/or synthesis of the corresponding mRNA is hampered in this strain, which is supported by the fact that the mRNA termination is disrupted. Furthermore, GalE is an endogenously SsrA-tagged protein, resulting in its degradation to rescue stalled ribosomes.[353]

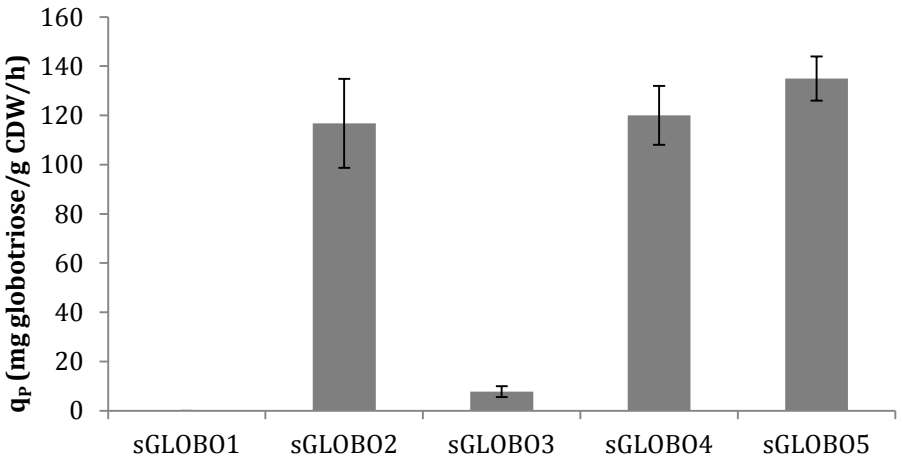


Figure 5.9: Comparison of the specific globotriose productivity of the created sGLOBO strains. Strains were grown on minimal medium containing 10 g/L of lactose.

Since sGLOBO5 showed the least biological variation, it was chosen to investigate the potential for large-scale globotriose production. Subsequently, the strain was grown in a 1-L bioreactor on minimal medium with 10 g/L lactose. The obtained fermentation profiles are shown in Figure 5.10 and clearly illustrate growth coupled globotriose production. At the end of the exponential phase, 1.9 g/L globotriose had accumulated extracellularly which increased further to 2.8 g/L during the stationary phase. The product to biomass yield Y_{PX} was 360 ± 14 mg/g CDW, which is comparable to the one observed on shake flask. However, the maximal specific growth rate of 0.22 ± 0.01 h^{-1} was significantly lower (possibly due to shear stress), resulting in a lower productivity and accumulation of fructose. Fine-tuning of this fermentation set-up will likely improve the production rates (r_p) which are, in contrast to existing processes (r_p = *constant*)[49,323], exponentially increasing ($r_p = Y_{PX} . \mu . X$) due to growth coupled production.

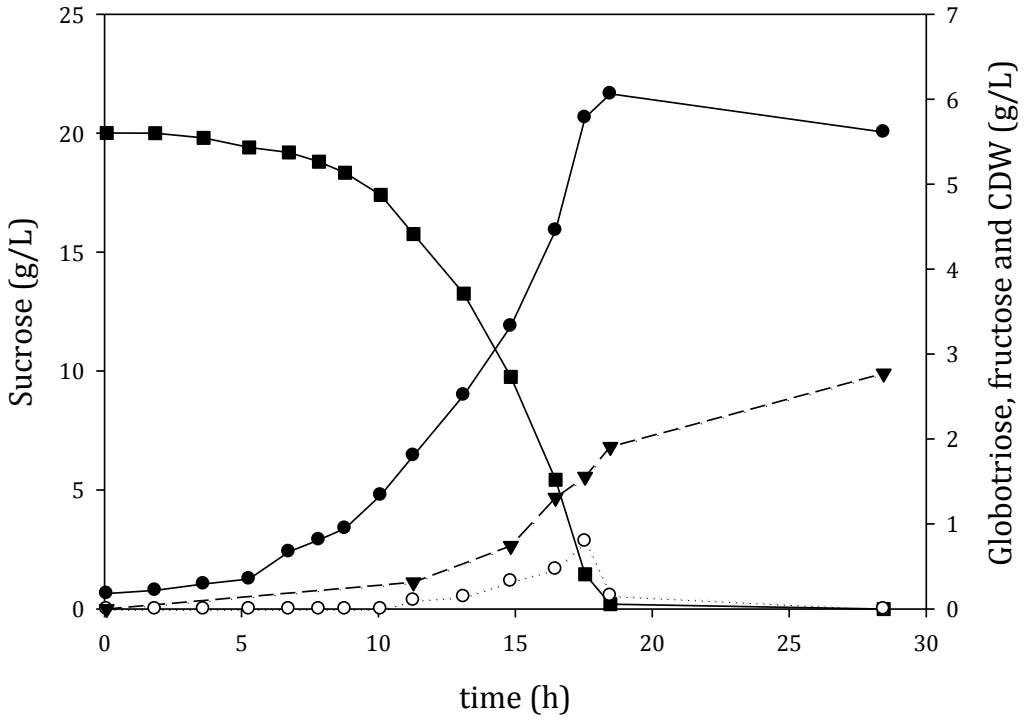


Figure 5.10: Production of globotriose in a 1-L bioreactor on minimal medium containing 10 g/L of lactose using strain sGLOBO5. Cell dry weight (●) was measured together with the extracellular sugars globotriose (▾), sucrose (■) and fructose (○).

5.5 Conclusions

In this contribution, the generic nature of the previously created glucosylation host was demonstrated by expanding it towards a galactosylation and rhamnosylation platform. To this end, UDP-glucose was converted into UDP-galactose or UDP-rhamnose by expressing a UDP-glucose epimerase (*galE*) or a UDP-rhamnose synthase (*MUM4*), respectively. As a proof of concept, the bio-active flavonol quercetin was selected for galactosylation and rhamnosylation, yielding hyperoside (quercetin 3-*O*-galactoside) and quercitrin (quercetin 3-*O*-rhamnoside), respectively. To achieve this, the flavonol 3-*O*-galactosyltransferase (F3GT) from *Petunia hybrida* and the flavonol 3-*O*-rhamnosyltransferase from *Arabidopsis thaliana* (RhaGT) were overexpressed in the engineered *E. coli* W mutants, respectively. The created strains were able to produce 940 mg/L of hyperoside and 1176 mg/L of quercitrin at specific production rates of 68.7 mg/g CDW/h and 47.8 mg/g CDW/h, respectively, which are the highest reported to date. Interestingly, both GTs showed *in vivo* activity towards other dietary flavonols, whereby over 400 mg/L of kaempferol 3-*O*-rhamnoside could be formed extracellularly. Finally, the flexibility of the engineered strain towards other acceptors than flavonols was illustrated by overexpressing the UGT LgtC, which galactosylates lactose. In this way, 2.8 g/L of the therapeutic oligosaccharide globotriose was produced in a 1-L bioreactor with a specific productivity of 117 mg globotriose/g CDW/h.

Chapter 6

Unlocking the potential of Sucrose Synthase for *in vivo* glycosylation

Contents

6.1 Abstract 98
6.2 Introduction 99
6.3 Materials and methods 101
6.3.1 Bacterial strains and plasmids 101
6.3.2 Materials and Molecular Agents 101
6.3.3 Structural modelling 101
6.3.4 Culture media and culture conditions 102
6.3.5 Preparation of crude extracts and enzyme assays 102
6.3.6 Bioconversion experiments 102
6.3.7 Product analysis and quantification 102
6.3.8 Kinetic analysis 103
6.4 Results and discussion 104
6.4.1 Selection of a suitable sucrose synthase 104
6.4.2 Developing a SuSy screening system 105
6.4.3 Engineering the affinity for sucrose 107
6.4.4 Coupling SuSy to UGT 112
6.5 Conclusions 114

6.1 Abstract

Sucrose Synthase (SuSy) is a promising enzyme which catalyzes the reversible conversion of sucrose and UDP into fructose and UDP-glucose. Since SuSy efficiently generates UDP-glucose from a cheap carbon source, it has been frequently used in bioconversion experiments or enzymatic synthesis of various glycoconjugates. When SuSy is coupled with a (uridine) glycosyltransferase (UGT), UDP is recycled which makes the overall process economically attractive. However, the use of SuSy in (*in vivo*) glycosylation systems has been hampered by its low affinity for sucrose, low stability or difficult expression.

To this end, a SuSy from *Solanum tuberosum* was selected due to its excellent overall stability and was engineered to obtain a robust enzyme with increased affinity for sucrose. To rapidly screen and characterize various SuSy mutants, a screening system was developed, which consisted of developing a growth-based high-throughput screen and creating a continuous assay to quickly identify altered kinetic parameters. Introduction of an S11E mutation resulted in a 3-fold higher affinity for sucrose due to mimicking phosphorylation of the serine residue. This improved enzyme was coexpressed with the glucosyltransferase VvGT2 in an engineered *E. coli* W host to produce the glucose ester of vanillic acid. The permeabilized cells efficiently produced vanilloyl-β-D-glucose at a rate of 0.41 ± 0.11 g/L/h with 68 % conversion. This newly developed system for the glucosylation of specialized metabolites alleviates the need to purify the proteins involved and is easily scalable.

6.2 Introduction

Sucrose synthase (SuSy – EC 2.4.1.13) is a key enzyme in the sucrose catabolism of green plants.[354] It catalyzes the reversible conversion of sucrose and a nucleoside diphosphate (preferably uridine diphosphate, UDP) into fructose and a nucleoside diphosphate glucose (often UDP-glucose, UDP-glc).[355] The function of SuSy shows a strong correlation with cellulose synthesis[356], sugar import[357], environmental stress response[358], nitrogen fixation and phloem unloading[359], yet still remains the subject of much debate in the field of plant physiology. Recently SuSy has been cloned from cyanobacteria such as *Anabaena* and *Synechocystis* species, being the first prokaryotic SuSy's elucidated. These SuSy's exhibit a preference for adenosine diphosphate (ADP) and have a non-homologous N-terminal region of the amino acid sequence.[360]

Since SuSy has the possibility to directly generate the expensive nucleotide sugar UDP-glc from the cheap and sustainable carbon source sucrose, it has been used in bioconversion experiments or enzymatic synthesis for the production of nucleotide sugars[361] and glycoconjugates[123,130,362]. When coupling SuSy with a (uridine) glycosyltransferase (UGT), UDP is recycled which makes the overall process economically attractive. However, the use of SuSy in (*in vivo*) glycosylation systems has been hampered by its low affinity for sucrose, low stability or difficult expression of this large (eukaryotic) protein (805 residues). Moreover, expression of several SuSy's in *E. coli* showed a 10-fold increase of the K_m in comparison to the native enzyme purified from the plant.[125,134] Native SuSy is phosphorylated on a conserved serine (Ser^{11}) residue near the N-terminus which has a significant impact on the kinetic parameters.[134,363] This feature is disadvantageous since *E. coli* is unable of phosphorylating SuSy. Cyanobacterial SuSy has no phosphorylation site and a high affinity constant for sucrose of 303 mM.[360]

Despite the recent structure elucidation of SuSy from *A. thaliana* by Garavito and coworkers[364], SuSy has never been the subject of extensive protein engineering to improve its kinetic parameters, nor the target for introduction of an alternative sucrose metabolism in micro-organisms. To this end, the main goal of this chapter is selecting a SuSy for subsequent engineering to obtain a stable, active enzyme with an improved affinity for sucrose, which can be used as an efficient *in vivo* catalyst. The pinnacle of this engineering strategy would be a SuSy that enables growth on sucrose when expressed in *E. coli* and is coupled with a UGT, thus giving rise to an effective *in vivo* glycosylation host.

To achieve this challenging goal, a novel SuSy enzyme will be selected as a starting point for both rational and random mutagenesis. The obtained mutants or mutant library will be screened for altered kinetic parameters using a developed high-throughput *in vitro* and *in vivo* screening system. The overall strategy is depicted in Figure 6.1, whereby the most active SuSy (exhibiting the lowest K_m) will be coexpressed with a UGT in an engineered *E. coli* W production host. In this perspective, VvGT2 (from Chapter 4) will be used as UGT of choice due to its *in vivo* versatility towards various phenolic acids. The production of vanilloyl-β-D-glucose was chosen as a proof of concept, since this compound has remarkable antioxidant properties[365] and a promising therapeutic potential[366].

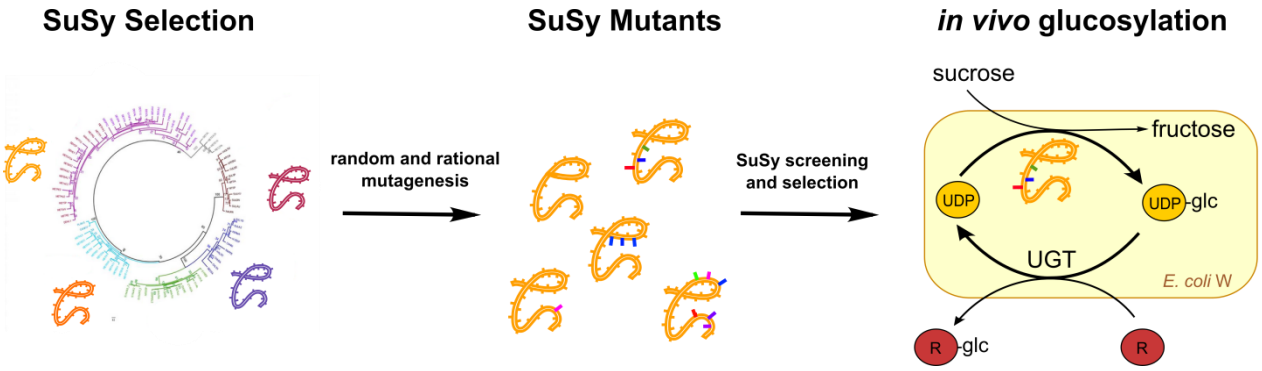


Figure 6.1: Overview of the strategy for the development of an *in vivo* glycosylation platform based on an engineered sucrose synthase and coexpression of an uridine glycosyltransferase (UGT) of choice. First, a novel SuSy will be selected based on reported activities, which will serve as a template for both random and rational mutagenesis. The obtained mutants are subsequently screened for altered kinetic parameters using a developed in vitro and in vivo screening system. Finally, the SuSy exhibiting the highest affinity for sucrose is coexpressed with a UGT in an engineered host to produce a glucosylated acceptor molecule.

6.3 Materials and methods

6.3.1 Bacterial strains and plasmids

All plasmids and strains used in this study are listed in Appendices A1 and A2, respectively. A codon optimized *susy* gene (accession number P49039) from *Solanum tuberosum* was synthesized by GeneArt® (Life Technologies, Massachusetts) and was used as a template for the amplification of *susy* using primers Fw_SuSy_EcoRI and Rv_SuSy_SacI, listed in Appendix A6. The amplified fragment was cut with EcoRI-HF and SacI-HF and ligated into the pCX-Kan-P22[258] vector resulting in the plasmid pCX-Kan-P22-SuSy (pSuSy). This plasmid was used as a template for both random and rational mutagenesis using primers from Appendix A6. All resulting fragments (truncated genes, S11E and S11D mutants, and randomly mutated genes) were subsequently religated in the pCX-Kan-P22 backbone following the Gibson assembly method[237]. Error prone PCR (ePCR) was performed according to the protocol of Savilahti and coworkers.[367] The obtained pCX-Kan-P22-SuSy-S11E was used as template for the amplification of *susy*-S11E sequence, whereas pBaSP/VvGT2 (Chapter 4) was used for amplification of the backbone. Gibson assembly of the SuSy-S11E insert at the BaSP position resulted in the production plasmid pCX-Kan-L5-P22-VvGT2-L6-P22-SuSy-S11E-L7 (pSuSy-S11E/VvGT2). *E. coli* W Δ*cscAR* was used as general expression host for pSuSy mutants, while *E. coli* W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* was used as production host for pSuSy-S11E/VvGT2.

6.3.2 Materials and Molecular Agents

T4-ligase, all restriction enzymes and Q5 polymerase were purchased from New England Biolabs (Ipswich, Massachusetts). All chemicals used were purchased from Sigma Aldrich (Germany) unless otherwise stated. Oligonucleotides were purchased from IDT (Leuven, Belgium).

6.3.3 Structural modelling

SuSy homology models were created on the I-TASSER server.[368] All figures were created with PyMOL.[369] Protein and ligands (sucrose, UDP, fructose) were converted when needed to pdb files and were further prepared using AutoDock Tools by converting them to pdbqt files.[370] Docking was performed using Autodock Vina[371] and affinities (kcal/mol) of the ligands to the receptor were calculated automatically. An estimation of the route(s) to the active site was calculated by using CAVER.[372]

6.3.4 Culture media and culture conditions

General growth media (LB and minimal medium) are described in Chapter 4. MacConkey agar base was purchased from Difco and was used to create sucrose based MacConkey agar according to the manufacturer's instructions. Minimal medium agar plates with sucrose (50 g/L) where created according to Aerts (2012)[373], but with addition of 1 mL L^{-1} of mineral solution as described in Chapter 4. Cells for permeabilization experiments were grown in tubes containing 5 mL LB with kanamycin (50 µg/mL) and served as a 2 % inoculum for 2 L shake flasks containing 500 mL LB medium and kanamycin . The cultures were grown at 37 °C at 200 rpm until the end of the exponential phase was reached. Cells were sterilely harvested by centrifugation for 20 min at 5000 × g and washed with 50 mM MOPS buffer (pH 7). Cell pellet was stored at -80 °C for permeabilization experiments. Conversion medium consisted of 50 mM MOPS buffer (pH 7) containing 200 mM sucrose, 0.5 mM UDP-glc, 1 mM $MgCl_2$ and 1 g/L vanillic acid.

6.3.5 Preparation of crude extracts and enzyme assays

Preparation of crude enzyme extracts has been described in Chapter 3. Determination of the enzyme activity was based on the protocol described by Salerno and coworkers[374], but fructose accumulation was continuously measured by adding hexokinase, phosphoglucose isomerase, glucose 6-phosphate dehydrogenase, ATP and NAD^+ directly to the reaction mixture. UDP-glucose was continuously measured (340 nm) using a newly developed assay by adding 1.2 U UDP-glucose pyrophosphorylase, 1.2 U phosphoglucomutase, 1.2 U glucose 6-phosphate dehydrogenase, 1 mM pyrophosphate (PPi), 2 mM NAD^+ and 0.2 mM glucose 1,6-bisphosphate.

6.3.6 Bioconversion experiments

Cells stored at -80 °C were thawed and dissolved in cold 50 mM MOPS buffer (pH 7) to a concentration of 500 g/L. Aliquots of 1 mL cell paste were used for cell permeabilization by sonication for 4 min at suboptimal conditions (Branson sonifier, 50 % duty cycle, output 3). Conversion experiments were initiated by adding 10 mL of conversion medium to the permeabilized cells.

6.3.7 Product analysis and quantification

Vanilloyl-β-D-glucose was primarily analyzed on TLC and afterwards quantified in detail using the same method described in Chapter 4 for glucogallin measurements. Purified vanilloyl-β-D-glucose from Chapter 4 was used as a standard.

6.3.8 Kinetic analysis

The kinetic constants were derived from initial rate analysis by varying the concentration of individual substrate. Using the continuous fructose assay, sucrose was varied from 0 to 1 M in the presence of 2 mM UDP. Kinetic parameters were calculated from an S-V plot by non-linear regression analysis using the Michaelis-Menten kinetic equations in R ('nlstools' package).

6.4 Results and discussion

6.4.1 Selection of a suitable sucrose synthase

Since a large number of sucrose synthases have been described to date and a lot of isoforms exist, choosing a suitable SuSy as a starting point for engineering is challenging. The goal of this chapter is to identify or design a SuSy with favorable kinetic parameters (high affinity for sucrose) when expressed in *E. coli*, which can coupled to a UGT, thus resulting in an efficient *in vivo* glucosylation platform. However, heterologous expression of SuSy in *E. coli* often results in an increased K_m when compared to the native enzyme due to the impossibility of phosphorylation.[134] An exception is the cyanobacterial SuSy of *Anabaena* which does not possess this conserved serine residue.[135] Although many SuSy's have been kinetically characterized, the obtained data are often inconsistent due to the fact that K_m estimation methods, protein expression and purification, and assay conditions significantly vary. Table 6.1 attempts to give an overview of frequently reported SuSy's and compares these affinity constants for different isoforms and expression methods.

Table 6.1: Overview of the affinity constants for sucrose of native SuSy's ($K_{m,scr}$) and recombinant SuSy's ($K_{m,scr,rec}$) expressed in *E. coli*. All K_m values were determined at saturation with UDP. A distinction is made between different isoforms when applicable. ND: not determined.

origin	isoform	$K_{m,scr}$ (mM)	$K_{m,Scr,rec}$ (mM)	reference
Solanum tuberosum	1	35	80.3	375,376
Solanum tuberosum	2	ND	105 ± 10	This study
Vigna radiata	2	17	161	134
Arabidopsis thaliana	1	ND	53 ± 6	355
Arabidopsis thaliana	3	ND	48 ± 3	355
Glycine max	1	31.3 ± 7.1	55.4	377,378
Pisum sativum	1	33 ± 5	32.1 ± 2.1	379,380
Pisum sativum	2	ND	42.7 ± 2.2	379
Pisum sativum	3	ND	67.5 ± 1.6	379
Zea mays	1	52 to 40	ND	381-383
Zea mays	2	63	ND	381
Anabaena sp.	2	303	270	135,360

Although their affinity constants are rather high, the SuSy's from *S. tuberosum* were selected since they have been proven to be very stable at 37 °C and show no activity loss during repeated freeze-thawing[125], making them suitable for industrial applications and

the goals of this chapter. Furthermore, they are well documented[125,375,376], exhibit a high affinity[375] for the cosubstrate UDP (K_m = 0.2 mM) when expressed in *E. coli* and they already have been deployed as efficient catalysts for the large scale production of ADP-glucose[384] and various oligosaccharides[131,385]. Isoform 1 (SuSy1) has been comprehensively studied, but isoform 2 (SuSy2, hereafter named SuSy) from *S. tuberosum* was not yet characterized in detail. The native sequence of *susy* was screened for rare codon clusters, which can drastically impede ribosomal translation and thus have a negative impact on protein expression.[386] As shown in Appendix A14, the number of rare codon clusters increased from 16 to 58 % when the expression host switched from *S. tuberosum* to *E. coli*. Based on this information the *susy* sequence was codon optimized and cloned in the constitutive expression vector pCX-Kan-P22[59], resulting in the plasmid pCX-Kan-P22-SuSy that was used throughout this chapter as a template for mutagenesis. Expression of SuSy in *E. coli* BL21(DE3) showed activity and had a K_m of 105 ± 10 mM based on the crude enzyme extract (Table 6.1). As a preliminary test, this strain was inoculated on minimal medium with 20, 200 and 500 g/L sucrose as only carbon source, but none of the cultures displayed growth after 72 h.

6.4.2 Developing a SuSy screening system

Obtaining an enzyme with desired properties solely by rational design is a very difficult challenge. Hence, a lot of enzyme engineering strategies rely on evolutionary processes in which a selective pressure is applied to a population of mutants, and those expressing a particular or desired function are preferentially propagated.[387] Due to the relative ease with which a mutant library can be generated nowadays, high-throughput screening efforts are needed. On the other hand, the mutagenesis process is typically applied in an iterative fashion until a satisfactory level of functional performance has been attained.[387] The probability that a first round of mutagenesis yields a SuSy with a 10-fold higher sucrose affinity is very low, which implies the need for developing an *in vitro* screen to evaluate the kinetic parameters of the obtained mutants. Therefore, a typical sceening system consists of two parts: an *in vivo* high-throughput screen to assess the possibility to grow on sucrose and a continuous assay for rapid determination of the kinetic parameters.

Developing a high-throughput screen

To develop a growth-based high-throughput screen, it is important to select a suitable screening host for SuSy mutants. A good host is sucrose negative, has a high transformation efficiency, expresses an active SuSy and is able to import sucrose. Three commonly used *E. coli* strains BL21(DE3), DH5α and MG1655 were screened together with the created *E. coli* W mutant Δ*cscAR* from Chapter 4 for these desired properties

(Table 6.2). All strains showed no growth on minimal medium with sucrose as only carbon source and had a high transformation efficiency, which is important to maintain the diversity generated by random mutagenesis. The occurrence of a sucrose transporter in many *E. coli* strains is still not fully elucidated, but it is suggested that *lacY* and *lamB* respectively code for a lactose permease and a maltose porin, which have a co-affinity for sucrose.[388] *E. coli* W, which grows particularly fast on sucrose, expresses a sucrose permease (CscB) which belongs like LacY to the oligosaccharide/H^+ symporter subfamily of the major facilitator superfamily.[389] It was shown that a *cscB* deletion mutant was unable to grow on sucrose, proving that no other transporter in *E. coli W* is capable of transporting sucrose.[269] On the other hand, MG1655 that constitutively overexpresses a sucrose phosphorylase (BaSP) is able to grow on sucrose while a DH5α transformant is not.[269,390] This suggests that LacY probably plays a significant role in sucrose transport, as the *lacZYA* operon is absent in *E. coli* DH5α.

Table 6.2: Overview of desired properties in some frequently used *E. coli* genotypes concerning the expression of (mutant) SuSy's. Transf. eff.: transformation efficiency.

***E. coli* strain**	**Growth on sucrose**	**Transf. eff. [a] (CFU/µg DNA)**	**Sucrose transporter**	**Reference**
BL21 (DE3)	No	1×10^9	LacY?	[271]
DH5α	No	2.54×10^9	-	[269]
K12 MG1655	No	1×10^9	LacY?	[390]
W Δ*cscAR*	No	1.56×10^9	CscB	This study

[a] transformation by electroporation

Since *E. coli* W Δ*cscAR* displays all the necessary characteristics, it was chosen as the expression host for SuSy screening. Furthermore, *E. coli* W Δ*cscAR* is deprived of its sucrose regulator which encodes the CscR protein that represses expression of the sucrose utilization genes at low sucrose concentrations.[268] In addition, *E. coli* W expresses a unique fructokinase CscK, absent in the other investigated *E. coli* strains, that effectively channels intracellular fructose towards the glycolysis.[269] This feature is of particular interest since cleavage of sucrose by SuSy or a sucrose phosphorylase releases fructose in the cell, which is normally converted by the manno(fructo)kinase Mak. This enzyme is shown to be promiscuous and exhibits a rather low activity[270,391], thus hampering the screen for an active SuSy in the other hosts.

Assay development

Various enzymatic assays have been developed for SuSy in both the synthesis and cleavage reaction. Since a quick screen for improved kinetics of SuSy is required, there will be only focused on continuously coupled assays which greatly reduce the assay time and cost. Figure 6.2 gives an overview of three possible continuous assays. An adaptation of the discontinuous fructose assay described by Salerno and coworkers[374] was identified as the best of these three, since the coupled enzymes are cheap, widely available and form a stabile mixture. Therefore, this assay was used for the determination of the kinetic constants of SuSy.

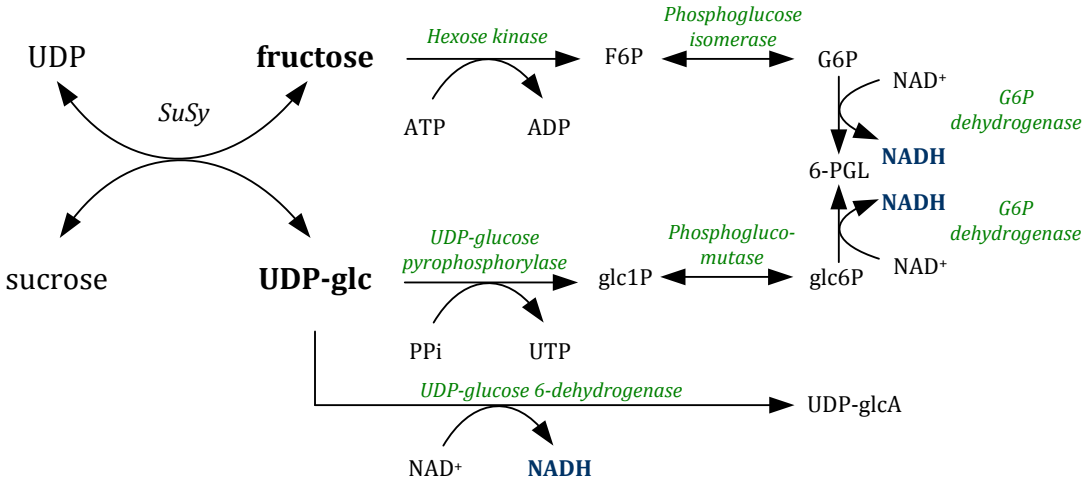


Figure 6.2. Possible continuous assays for the cleavage reaction of SuSy. Coupling enzymes are shown in green and NADH (absorbance at 340 nm) is depicted in blue. PPi: pyrophosphate, glc1P: glucose 1-phosphate, glc6P: glucose 6-phosphate, 6-PGL: 6-phosphogluconolactone.

Conversely, the UDP-glucose 6-dehydrogenase assay has been used by some research groups[392,393], but the current availability, price and overall stability of the enzyme make it however very expensive. Also, the possibility of coupling UDP-glucose pyrophosphorylase to the SuSy reaction was explored.[156] Though the new assay performed well in measuring SuSy activity, combination of Mg and pyrophosphate (PPi) resulted in the precipitation of an insoluble MgPPi complex (K_d = 55.3 µM)[394] which interfered with the absorbance measurements.

6.4.3 Engineering the affinity for sucrose

Finalization of the screening system permits the start of engineering a SuSy with an increased affinity for sucrose. First, the general structure of SuSy will be discussed whereafter some structure based engineering strategies will be evaluated.

General structure considerations

Although SuSy has been the subject of extensive (physiological) characterization for the past 50 years, any attempts towards protein engineering are scarce.[134,378] Only in 2011, the structure of SuSy1 from *A. thaliana* was elucidated by Garavito and coworkers[364] and showed that the structure was a flat, donut-like tetramer with two types of subunit interfaces. Each monomer has a tri-lobed structure with 4 distinct domains shown in Figure 6.3: a cellular targeting domain (CTD, residues 11-127), an ENOD10 peptide binding domain (EPBD, residues 157-276), and 2 glycosyltransferase B domains (GT-B_N, residues 277-526; GT-B_C, residues 527-754).[364] An interesting characteristic is the presence of a unique Nα1 helix which is 10 Å longer than any other homologous helix in GT-B structures reported.

The SuSy structure of *Arabidopsis* (AtSuSy) is 78 % identical to isoform 2 of *Solanum* and a monomeric homology model was created using I-TASSER[368], which was used to identify interesting regions for mutagenesis. Plotting of the overall electrostatic potential φ(r)[395,396] by solving the Poisson-Boltzmann equation revealed that SuSy has a negatively charged surface (Figure 6.4). This is in accordance to the theoretical isoelectric point of 5.93 and the slightly more acidic p*I* of 5.5 as determined experimentally[125]. Figure 6.4 also displays the presence of a positively charged (blue) zone which probably acts as an entrance for the negatively charged UDP or UDP-glucose. The route towards the active site of SuSy has never been investigated and the 3D structure of AtSuSy is presumably in the "closed" and thus active conformation[364], a typical feature of glycosyltransferases[34,397]. The substrates are captured in between the cleft of GT-B_N (light green) and GT-B_C (light yellow) as shown in Figure 6.4. The CAVER tool[372] was used to search for possible routes towards the active center and are depicted in white, which correspond to the positively charged cleft.

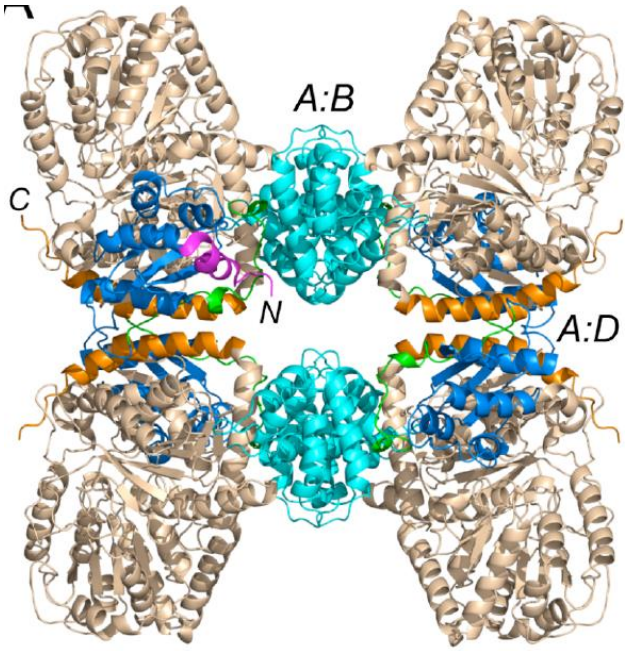


Figure 6.3: Ribbon representation of the tetrameric SuSy from *A. thaliana* (PDB 3S27) displaying the subunit interface interactions (A:B and A:D). The cellular targeting domain (CTD) is colored in marine and the ENOD40 peptide binding domain (EPBD) is cyan. The GTB domains are depicted in wheat together with the C-terminus in orange. Figure adapted from Zheng *et al.* (2011).[364]

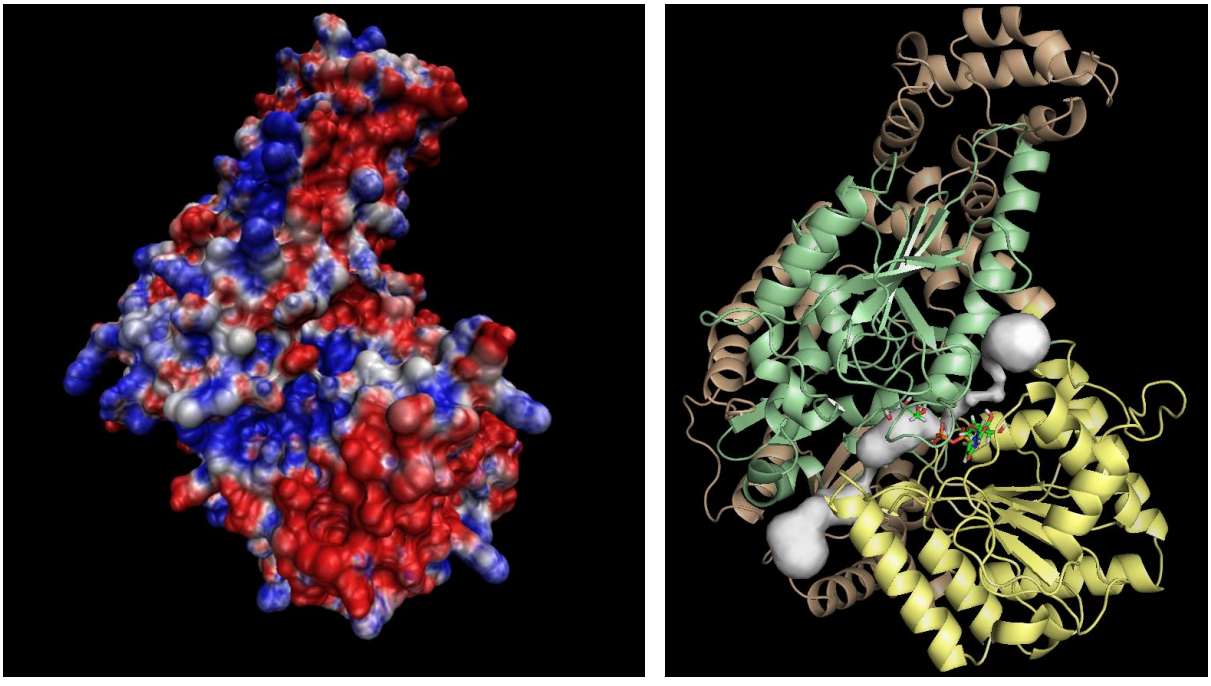


Figure 6.4: Structural overview of the electrostatic potential (left) and closed conformation with routes to the active site (right) of SuSy from *S. tuberosum* (PDB 3S29). The electrostatic surface potential shows a positively charged zone (blue) which corresponds to the cleft between the GTB_N (green) and GTB_C (yellow) domains. Two possible routes near the active site were visualized using CAVER[372] (white).

Truncation mutants of SuSy

A first attempt to gain insight in the function and necessity of the N-terminal domains (CTD and EPBD) of SuSy, is to make truncations. These domains cover 34 % of SuSy without any contribution to the active site of the GTB. A previous study created an N-terminal truncation (A2 to R13) of soybean SuSy and concluded that the activity and K_m were not significantly different compared to the wild type.[378] In addition, precise truncation of proteins can result in a higher specific activity[398] or the identification of the minimal active protein[399]. Figure 6.5 shows the location of various selected truncations whereby SuSy-T111 codes for SuSy with amino acids 2 to 110 deleted. The selection of deleted residues was based on the starting position of a new helix or limitation of the hydrophobic effect.

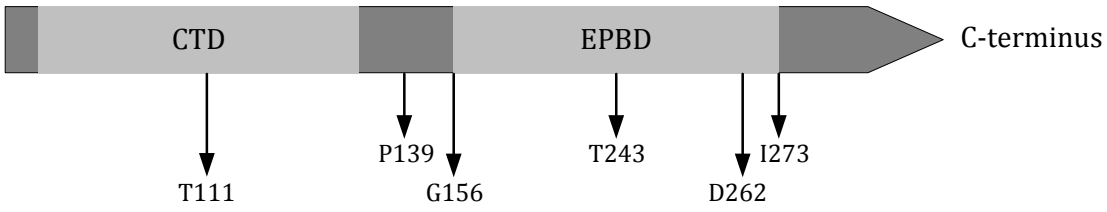


Figure 6.5: Location of the selected truncations in SuSy of *S. tuberosum* in the N-terminal region, consisting of a cellular targeting domain (CTD) and an ENOD40 peptide binding domain (EPBD). Arrows indicate the position where the truncation was made, whereby T111 is SuSy with amino acids 2 to 110 deleted.

Analysis of the crude protein extracts revealed that all activity in both cleavage and synthesis direction was lost. These findings illustrate the importance of the cellular targeting domain (CTD) and ENOD40 peptide binding domain (EPBD). Possible explanations for the inactivity of these mutants are the loss of its tetrameric conformation, deletion of highly conserved residues and destabilization of the structure. As shown in Figure 6.3, the A:B interface originates almost entirely from interactions between adjacent EPBDs, while the A:D interface is formed between EPBD and the C-terminus. It has been shown that a C-terminal truncation results in the direct loss of activity[378] and that a dimeric SuSy is not active[400]. Furthermore, the highly conserved residue Q77 makes a polar contact with E491 from the GTB domain thus ensuring proper folding and holding both domains together. Deletion of this residue probably results in a destabilized structure.

Mimicking phosphorylation

The S11 residue is highly conserved in plants and is phosphorylated, which results in a lowering of the K_m.[125,134] The N-terminus of cyanobacterial SuSy's is completely different as depicted in Figure 6.6 and has no comparable residue. Two mutations (S11D and S11E) were introduced to mimic a constitutive phosphorylation, resulting in the pCX-Kan-P22-SuSy-S11D/E plasmids. The plasmids were transformed in the developed *E. coli* W Δ*cscAR* host and the crude extracts showed an apparent increase in sucrose affinity from 105 mM to 54.8 ± 1.2 mM for S11D and 35.7 ± 1.1 mM for the S11E mutant, respectively. The S11E mutation is more beneficial since its *phosphomimetic* potential (charge and residue length) is higher. These findings are inconsistent with the data obtained by Elling and coworkers[375], which showed no significant increase in affinity between isoform 1 of SuSy from *S. tuberosum* and the corresponding S11D mutant.

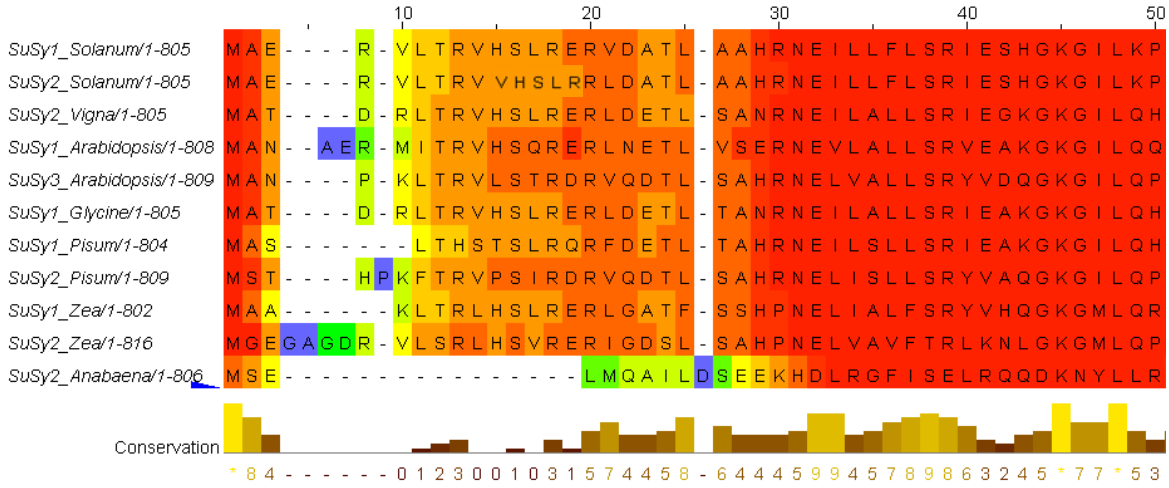


Figure 6.6: Alignment of the SuSy amino acid sequences from Table 6.1 at the N-terminus. The serine (S11) residue is conserved in plants and is not present in *Anabaena*, while K39 and L42 are conserved within all described SuSy's.

Random mutagenesis

Rational design of SuSy is very challenging and resulted till now in an S11E mutant with a 3-fold higher affinity for sucrose. Since other educated guesses to improve the kinetic constants of this enzyme are difficult to make, some effort will be put in (semi) random mutagenesis. The activity of SuSy S11E was not significantly different from the native SuSy, so that the corresponding sequence was used as a template for mutagenesis. Error prone PCR (ePCR) using a mutagenic buffer based on the method of Rasila *et al.*[367] was used to generate mutant libraries of the complete *SuSy*-S11E gene or large subdomains thereof. The latter regions of interest were the Nα1 helix (stretching from G301 to Q324), an outer loop (from L338 to Q349) and a combination thereof (from G301 to

Q349). The complementary backbones were amplified using a high fidelity polymerase and the degenerated inserts were ligated using the Gibson assembly method[237]. The created mutant libraries were transformed in *E. coli* W Δ*cscAR* and 80 % of the transformation mixture was plated on sucrose minimal medium agar, whereas 10 % was plated out on sucrose based MacConkey agar and the remaining 10 % on LB agar with kanamycin. No colonies were detected on minimal medium after extended incubation. In addition, no mutants were found on MacConkey agar that could coferment sucrose. The colonies on LB+Kan plates were grown on liquid medium and the crude extracts were tested for activity. The positives were rapidly screened for an improved affinity for sucrose and are listed in Table 6.3. A second round of ePCR on the Nα1 helix resulted in loss of activity or activity starting from 100 mM sucrose, thus illustrating the challenge of lowering the K_m value by both rational and random mutagenesis.

Table 6.3. Results of the random mutagenesis.

Mutation	Result
Complete gene	All active. K_m higher than WT.
Nα1 helix G301 to Q324	Active mutants with K_m as WT.
Outer loop L338 to Q349	Mutants only active from 100 mM sucrose
Combined V303 to Q349	Mutants only active from 100 mM sucrose

6.4.4 Coupling SuSy to UGT

Although this affinity constant was greatly reduced with the S11E mutation, it proved to be too high to enable growth on sucrose when expressed in *E. coli*. As the intracellular concentration of UDP is almost 10-fold higher than the K_m for UDP of this enzyme[125,401], sucrose concentration in the cell is likely to be the limiting factor. When taken into account the K_m of other sucrose degrading enzymes like invertases and phosphorylases (SP)[373] which do permit growth when expressed, an estimation of the affinity range needed is between 0.5 and 5 mM. This is supported by the fact that SP enzymes with a K_m above 5 mM were unable to grow on sucrose when expressed in *E. coli*.[390] Since the affinity constant of SuSy-S11E is over 7-fold higher than this threshold, further engineering is required when a strain needs to be obtained with production coupled to growth as described in Chapters 4 and 5.

Nevertheless, this engineered SuSy has become more suitable and efficient for biocatalysis, especially when coupled with a UGT. To this end, SuSy-S11E was coexpressed with the glucosyltransferase VvGT2, resulting in the constitutive expression plasmid pSuSy/VvGT2, which was transformed in the sGLC4 host (*E. coli* W Δ*cscAR*

Δ*pgm* Δ*agp* Δ*ushA*) from Chapter 4. The strain was grown on LB after which the cells were harvested, washed and stored at -80 °C. The cell pellet was resolved in 50 mM MOPS (pH 7), permeabilized and used at a concentration of 50 g/L for conversion experiments. Vanillic acid was used as acceptor, since the resulting glucose ester vanilloyl-β-D-glucose has promising therapeutic potential[366]. A first bioconversion was carried out in 50 mM MOPS buffer (pH 7) containing 200 mM sucrose, 2 mM UDP, 1 mM $MgCl_2$ and 1 g/L vanillic acid, but after 8 h of incubation at 37 °C, no glucoside could be detected. Analysis of SuSy in the permeabilized extracts showed an activity of 61 ± 12 U/L, thus proving that SuSy was not the bottleneck. Repeating the experiment in the presence of 0.5 mM of UDP-glc instead of 2 mM UDP showed a rapid formation of glucoside as shown in Figure 6.7.

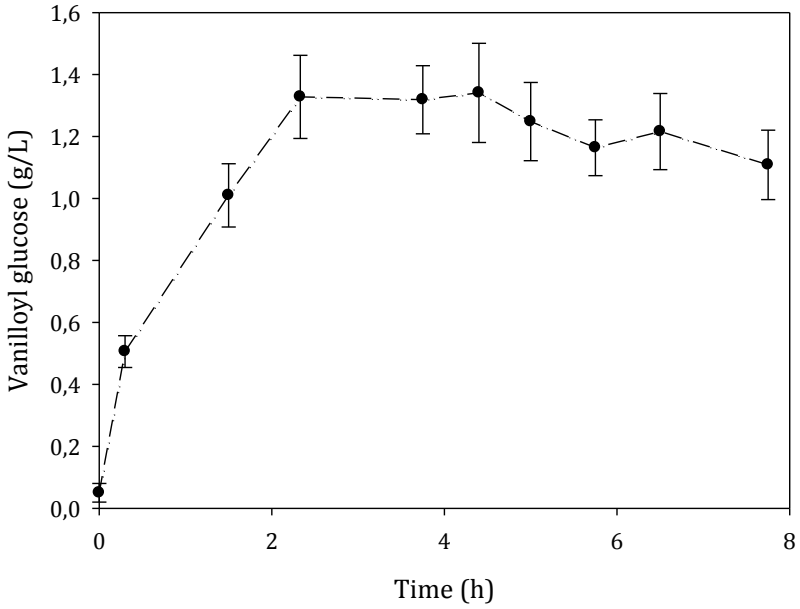


Figure 6.7: Bioconversion of 1 g/L of vanillic acid into vanilloyl-β-D-glucose by coupling SuSy-S11E and VvGT2 in presence of 200 mM sucrose and 0.5 mM UDP-glucose at pH 7 and 37 °C.

This indicates that UDP inhibits VvGT2 as an end product, a phenomenon frequently encountered with *O*-glucosyltransferases[124,130]. The vanillic acid was almost completely converted in 2 h, resulting in 8.13 regeneration cycles of UDP-glc. Remarkably, the product degraded over time and resulted in a maximal conversion of 69.1 ± 9.7 %. An average production rate of 0.41 ± 0.11 g/L/h of glucoside could be reached, which is to our knowledge in the same line as the most efficient *in vitro* process reported regarding the glucosylation of specialized metabolites[130]. This system could efficiently convert the dihydrochalcon phloretin in its corresponding C-glucoside nothofagin at an estimated rate of 0.34 g/L/h. Despite the potential of this *in vitro* process, it should be mentioned

that this was carried out on a 1.5 mL scale and that scale-up will be far from cost-effective due to the need for extensive enzyme purification. Moreover, the process required repeated enzyme feed to avoid slowdown of the production rate, thus making this system even more expensive. In contrast, our process alleviates the need for extensive enzyme purification and is easily scalable, thus illustrating an improved *in vivo* method for the glucosylation of small molecules.

6.5 Conclusions

In this chapter, a structural analysis is made of the Sucrose Synthase (SuSy) from *Solanum tuberosum* to generate both rational and random mutants with an increased affinity towards sucrose. To efficiently screen and characterize these mutants, a screening system was developed which consisted of a growth-based high-throughput screen and an adapted continuous assay. The newly designed *E. coli* W Δ*cscAR* proved to be the most suitable screening host for SuSy due to its high transformation efficiency and presence of the sucrose transporter CscB. N-terminal truncation mutants (up to I237) showed no activity, thus illustrating the importance of these domains. The native affinity constant (K_m) for sucrose could be reduced 3-fold to 35.7 ± 1.1 mM by introducing an S11E mutation, which mimics phosphorylation of the serine residue. Random mutagenesis efforts starting from this SuSy-S11E mutant yielded only mutants with the same K_m or enlarged values. Coexpression of the improved SuSy-S11E with the glucosyltransferase VvGT2 in an engineered *E. coli* W host (sGLC4) resulted in a strain that could efficiently glucosylate vanillic acid after permeabilization. The glucoside vanilloyl-β-D-glucose was produced at a rate of 0.41 ± 0.11 g/L/h with 8.13 regeneration cycles of UDP-glucose, which illustrates the potential of this new system for the glucosylation of specialized metabolites.

Chapter 7

Conclusions and perspectives

Glycosylation is an important modification in nature which can significantly improve the physicochemical and biological properties of small molecules like specialized (secondary) plant metabolites or antibiotics. The majority of glycosylation reactions are mediated by glycosyltransferases (GTs), which transfer the sugar residue from an activated sugar donor (like UDP-sugars) to various acceptors such as specialized (secondary) metabolites or oligosaccharides. The addition of a sugar residue to these molecules can greatly alter their solubility, stability or bioactivity, which are desirable properties when they are applied as nutraceuticals, therapeutics or cosmetics. Since the majority of specialized metabolites occurs as glycosides, extraction from natural resources remains the most widely used method to deal with the ever increasing demand for glycosylated compounds. However, this is often a low-yielding and laborious task using toxic solvents and generating much waste.

Fuelled by these challenges, many 'green' alternatives have been developed for the production of glycosides as illustrated in **Chapter 2**. These can be divided in *in vitro* (using enzymes) and *in vivo* (using whole cells) systems, each having their specific advantages and challenges. Since *in vitro* processes are difficult to scale up and need purified enzymes and expensive cofactors, the key factors towards their economic viability are engineering the enzymatic (thermo)stability and increasing the number of regeneration cycles (RC) of nucleotide sugars. Furthermore, the search for tailor-made enzymes with novel or improved functions is greatly stimulated by the expanding fields of enzyme engineering[402] and metagenomics[218]. Sequence based metagenomics can aid in the discovery of more stabile or active alternatives for existing enzymes, while function based screens can unravel completely novel enzyme activities.

On the other hand, *in vivo* processes have been proven to be very useful for the cheap and large scale production of various oligosaccharides and glycosylated specialized metabolites. Notwithstanding the promising conversion yields obtained in hosts like *E. coli* and *S. cerevisiae*, the corresponding production rates and titers of many glycosides are often very low. This is primarily caused by metabolic engineering strategies which mostly achieve indirect formation of UDP-sugars via the bottleneck enzyme Pgm. To this

end, alternative routes could be introduced in the form of sucrose synthase or sucrose phosphorylase, thus resulting in the direct formation of activated precursors from the bio-economically attractive and more sustainable substrate sucrose. In addition, the released fructose can function as a substrate for growth, which opens the possibility to simultaneously produce and grow. A final and recently emerging perspective for *in vivo* glycosylation is the complete *de novo* synthesis of glycosides, whereby the aglycon is equally formed by the host. In this way, glycosylation becomes a 'decoration' module in the total biotechnological synthesis of natural or new to nature glycosides.

In the light of efficient UDP-sugar formation and the search for more flexible enzymes, a promiscuous uridylyltransferase from *B. bifidum* was identified in **Chapter 3**. Three putative gene candidates (*galT1*, *galT2* and *ugpA*) were cloned in a newly engineered *E. coli* strain. Expression in this host enabled to screen and characterize the enzymes only by their crude extracts since major interference reactions were deleted, thus alleviating the need for enzyme purification. A newly developed chemo-enzymatic assay based on phosphate detection in the micromolar range enabled to identify UgpA as the promiscuous uridylyltransferase with UTP-monosaccharide-1-phosphate uridylyltransferase activity (EC 2.7.7.64). This enzyme catalyzed both the formation of UDP-glucose and UDP-galactose from glucose 1-phosphate (glc1P) and galactose 1-phosphate (gal1P), respectively. Remarkably, UgpA showed a 7-fold higher activity compared to the homologous *E. coli* uridylyltransferase GalU. Hence, UgpA proved to be a versatile catalyst for *in vitro* UDP-sugar generation. To what extent other sugar 1-phosphates may be converted by UgpA could be the subject of future research, hereby optionally using the developed screening host.

To address the aforementioned inadequate UDP-sugar formation in most *in vivo* platforms and to harness the power of the newly discovered UgpA, a novel glucosylation platform was designed in **Chapter 4**. Through metabolic engineering of *Escherichia coli* W, a robust plug and play strain was developed that couples production of glucosides directly to growth, hereby only using sucrose as a cheap carbon source. To this end, the existing sucrose metabolism (encoded by *cscARKB*) was modified by knocking out *cscAR* combined with introduction of the sucrose phosphorylase from *Bifidobacterium adolescentis* (BaSP), which directly generates glc1P as an activated precursor for UDP-glucose. To ensure proper glucoside production, three consequent modules were elaborated: (1) the creation of a split *E. coli* W metabolism, resulting in the reduced or eliminated use of glc1P for growth, preserving it for UDP-glc formation; (2) enhancing and securing UDP-glucose formation; and (3) the expression of the versatile glucosyltransferase VvGT2 from *Vitis vinifera* for the glucosylation of various phenolic

acids. This VvGT2 specifically enabled the strain to produce the therapeutic compound β-glucogallin (1-*O*-galloyl-β-D-glucose) starting from gallic acid and sucrose while actively growing. Specific productivities were obtained up to 100 mg β-glucogallin/g CDW/h, which was 20-fold higher compared to the wild type strain harboring VvGT2. This effect was largely caused by the deletion of *ushA*, which prevented degradation of the UDP-sugar pool. Due to the promiscuous character of VvGT2, 14 other glucose esters of various hydroxycinnamic and hydroxybenzoic acids were produced in a stereospecific way with high conversion yields up to 100 %, including three newly reported glucose esters. To our knowledge, this fast growing (and simultaneously producing) *E. coli* mutant is the first versatile host described for the glucosylation of phenolic acids in a fermentative way solely relying on sucrose.

However, as discussed in Chapter 4, the envisioned split metabolism (characterized by a 50 % decrease in biomass yield) was not fully achieved. In this context, investigating the effect of a double phosphatase (encoded by *yihX* and *yidA*) knockout or an *ycjU* knockout in sGLC5 could be very useful. Alternatively, ^{13}C metabolic flux analysis can provide fundamental information[403] on the intracellular fluxes in the central metabolism of the *E. coli* W mutants, which can result in an educated guess for engineering a truly split metabolism. In this way, product yields and concomitantly specific productivities will be even further increased, which gives rises to a more robust strain.

Expression of other GTs in this platform would result in a true plug and play strain that glucosylates any desired small molecule, thus demonstrating the versatility towards different acceptors. However, this could require additional engineering efforts regarding transport, degradation, scalability and toxicity of each tested acceptor and corresponding glucoside. This was clearly demonstrated when more than 1 g/L of vanillic acid was added, which proved toxic and caused inhibition of growth, whereas the glucoside could be accumulated over 2 g/L without any effects (data not shown). Some ways to overcome this are the controlled (exponential) feed of the acceptor or the use of a continuous fermentation. Besides alteration of the process conditions, strain improvements can be realized on molecular level as well. At this moment, the genes encoding BaSP, VvGT2 and UgpA are constitutively expressed under the same promoter P22, possibly resulting in an imbalanced pathway with accumulation of key intermediates. In order to minimize the metabolic burden and to maximize the flux through the pathway, resulting in a fine-tuned and increased production, a combination of different promoters with different strengths and ribosome binding sites preceding the constitutively overexpressed genes can be tested. A schematic representation of the combinatorial pathway optimization for selecting improved β-glucogallin producer

strains is shown in Figure 7.1. The resulting mutant library can be rapidly screened by a newly developed assay whereby rhodanine reacts with residual gallic acid, resulting in a bright red complex (540 nm)[404].

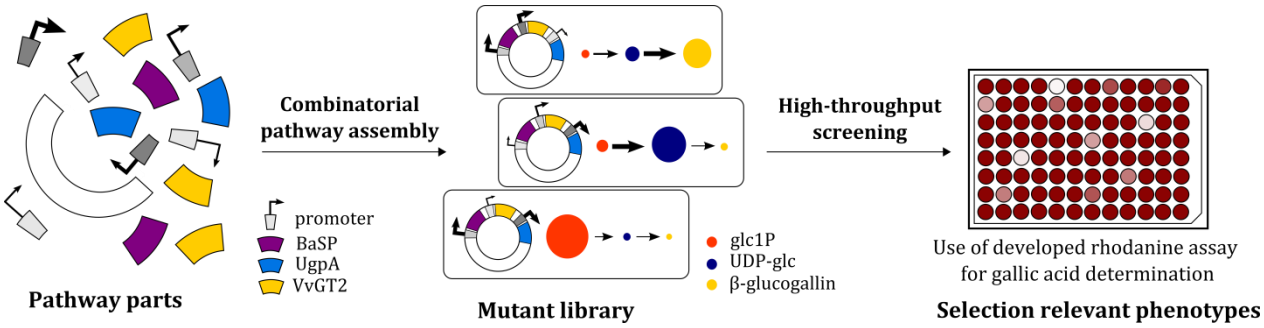


Figure 7.1: Schematic representation of a combinatorial pathway optimization. A mutant library can be easily obtained by combining the fixed pathway parts (BaSP, UgpA, VvGT2) with a promoter or RBS library. The selection of the best producer strains for β-glucogallin is done using a high-throughput screen based on the coloration of residual gallic acid (rhodanine assay).

In **Chapter 5** the generic nature of the created glucosylation host was demonstrated by expanding it towards a galactosylation and rhamnosylation platform. To this end, UDP-glucose was converted into UDP-galactose or UDP-rhamnose by simply expressing a UDP-glucose epimerase (*galE*) or a UDP-rhamnose synthase (*MUM4*), respectively. As a proof of concept, the bio-active flavonol quercetin was selected for galactosylation and rhamnosylation, yielding hyperoside (quercetin 3-*O*-galactoside) and quercitrin (quercetin 3-*O*-rhamnoside), respectively. To achieve this, the flavonol 3-*O*-galactosyltransferase (F3GT) from *Petunia hybrida* and the flavonol 3-*O*-rhamnosyltransferase from *Arabidopsis thaliana* (RhaGT) were overexpressed in the engineered *E. coli* W mutants. The created strains were able to produce 940 mg/L of hyperoside and 1176 mg/L of quercitrin at specific production rates of 68.7 mg/g CDW/h and 47.8 mg/g CDW/h, respectively, which are the highest reported to date. Interestingly, both GTs showed *in vivo* activity towards other dietary flavonols, whereby over 400 mg/L of kaempferol 3-*O*-rhamnoside could be formed extracellularly. Finally, the flexibility of the galactosylation platform to produce specialty oligosaccharides from lactose was also assessed. However, to prevent metabolization of this acceptor, the *lac* operon, containing an β-galactosidase (*lacZ*) and lactose permase (*lacY*) was knocked out and lactose transport was restored by knocking a constitutively expressed *lacY* back in at the same locus. Replacing F3GT with the lipo-oligosaccharide galactosyltransferase (LgtC) from *Neisseria meningitidis*, enabled our mutant strain to efficiently produce 2.8 g/L of globotriose (α-galactosyl-1,4-lactose) at a rate of 117 mg/g CDW/h.

Due to the pivotal role of UDP-glc, various other UDP-sugars can be formed *in vivo* such as UDP-glucuronate, UDP-xylose, UDP-arabinose, UDP-apiose and UDP-galacturonate. Figure 7.2 gives an overview of the UDP-sugars produced in this dissertation (coloured in green) and the most common derivatives thereof. When implementing a dehydrogenase like Ugd, effects of NAD^+ consumption should be taken into account to co-engineer an optimal redox balance.

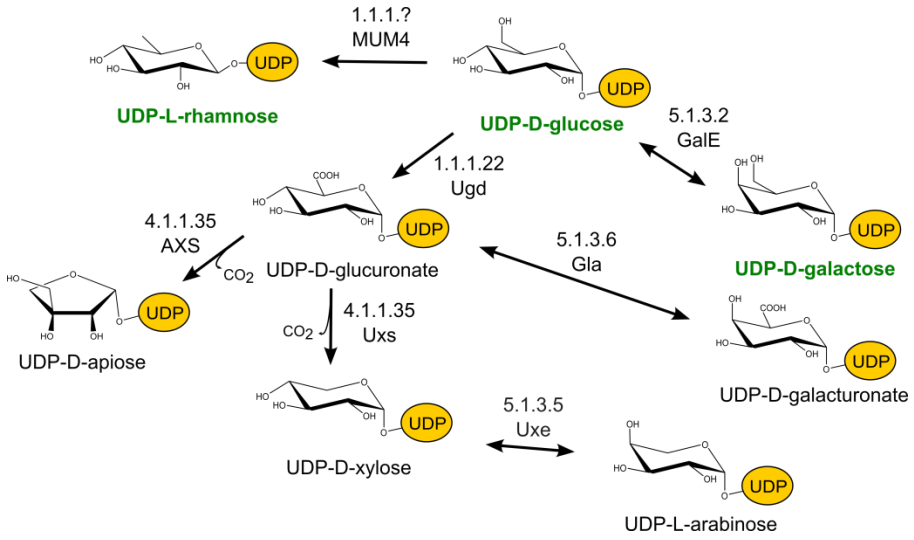


Figure 7.2: Overview of the interconverting reactions starting from the central molecule UDP-glucose for the generation of the most common UDP-sugars found in nature. UDP-sugars depicted in green were successfully produced in this dissertation. EC numbers and the abbreviated gene products are shown next to each reaction.

In combination with the modularity of the developed glycosylation strains, which permits rapid introduction of any GT or UDP-sugar pathway, virtually any glycoside can be produced. Together with the novel metabolic engineering strategy which established coupling of growth and production, over 50-fold increase in productivity can be achieved compared to existing processes. To fully exploit this strategy, the fermentations carried out in this dissertation should be expanded into a fed-batch set-up. In this way, high cell densities can be obtained which equally result in exponentially increasing product titers. Interesting target molecules for future research with industrial applications are the antioxidative bisglycoside rutin (quercetin 3-*O*-rutinoside), the anticarcinogenic baicalin (baicalein 7-*O*-glucuronide) or the cardioprotective piceid (resveratrol 3-*O*-glucoside).

Finally in **Chapter 6**, we explored the potential of the promising enzyme sucrose synthase (SuSy), which is able to directly generate UDP-glucose from sucrose. Since combination of SuSy with a GT gives rise to a rapid and energy efficient recycling of UDP,

exploiting this principle in actively growing cells is industrially very attractive. However, due to its unfavorable kinetics (low affinity for sucrose), SuSy has been rarely used for the *in vivo* formation of UDP-glucose. To create a more suitable SuSy, a new screening system was developed to rapidly evaluate and characterize SuSy mutants. This system consists of a growth-based high-throughput screen and a continuous assay to quickly identify altered kinetic parameters. The SuSy of *Solanum tuberosum* was used as a starting point and displayed an affinity constant (K_m) of 105 ± 10 mM. This value could be significantly reduced to 35.7 ± 1.1 mM by introducing an S11E mutation, which mimics phosphorylation of the serine residue. Other attempts (random mutagenesis and N-terminal truncations) to lower this constant remained unfruitful and expression of SuSy-S11E in *E. coli* W Δ*cscAR* did not confer growth on sucrose to this strain. However, coexpression of this SuSy mutant with the glucosyltransferase VvGT2 in the engineered host created in chapter 4 resulted in an *in vivo* (bioconversion) host. The permeabilized cells were able to efficiently glucosylate vanillic acid at a rate of 0.41 ± 0.11 g/L/h with 68 % conversion and an RC_{max} of 8. This system alleviates the need to purify the proteins involved and is easily scalable. Production rates, RC_{max} and titers can be further increased by (stepwise) addition of more acceptor, which will result in a new economically viable and versatile glycosylation host.

This methodology could be further improved by engineering a SuSy that permits growth on sucrose, thus effectively creating a split metabolism with production (glycosylation) coupled to growth. In this way, the cells do not need to be cultured before, since growth and production will be coupled. An estimation of the K_m to be achieved for growth will be in the range of 1 mM, which is 20 to 300-fold higher than the naturally occurring SuSy's. However, engineering the affinity constant of SuSy will be very challenging, especially since no attempts have been made to date to approach this enzyme in a systematic and structural way. Indeed, the only available 3D structure from *A. thaliana* gives no information about the combined position of UDP and sucrose ligands, but only shows UDP combined with a possible glucosyl intermediate.[364] To this end, a homology model was created of Susy from *S. tuberosum*, which was subsequently used to dock the ligands sucrose and UDP. Figure 7.3 shows the most representative docking with a binding affinity of 7 kcal/mol whereby the residues located in the active site are colored in orange. Other possible conformations were useful in the identification of other catalytic residues or residues that influence the affinity for sucrose or UDP. However, these results are estimations since the structure is only available in its closed conformation, yet give an interesting insight in the molecular organization of SuSy. Sucrose is stabilized by the pyrophosphate moiety of UDP and the conserved residues H436 and E673. The glycines (G300 and G301) are conserved in all GTB type

glycosyltransferases and possibly have the same 'hinge' function during closing of the domains in their active forms as described in sucrose 6-phosphate synthase[405]. Furthermore, G301 forms the beginning of the interesting Nα1 helix, which pushes into the active centre and probably has a major influence on the affinity constant for sucrose. Alanine scanning or site saturation of the residues surrounding the docked sucrose ligand could provide valuable information on their contribution to the activity, specificity and affinity of SuSy.

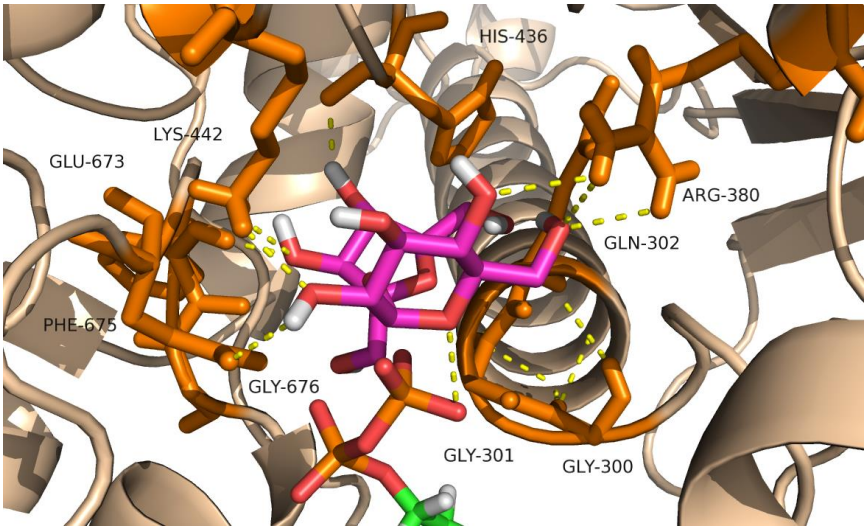


Figure 7.3: Docking of sucrose in the modeled SuSy of *S. tuberosum* with UDP (green) ligand. Residues that participate in bonding interactions are shown in orange. Polar interactions are displayed as yellow dashes.

Remarkably, all research concerning SuSy has been focused on eukaryotic enzymes, with two cyanobacterial homologs (*Anabaena sp.* and *Thermosynechococcus elongatus*) as only exceptions. However, since many proteobacterial genomes have been sequenced in the past years, certain gene clusters exhibiting strong homology with SuSy have been identified, but not yet characterized.[406] To this end, we cloned a putative SuSy gene originating from *Nitrosomonas europaea* and confirmed its function as genuine sucrose synthase for the first time. This SuSy preferred ADP over UDP and displayed a K_m value towards sucrose of only 25 ± 5 mM, which is even lower as the engineered SuSy of *S. tuberosum*. These proteobacterial SuSy's could unlock new possibilities when used as industrial catalysts.

Up to now, the created glycosylation strains were sucrose based, hereby exploiting sucrose phosphorylase to generate glc1P. Theoretically, various other disaccharides may be 'split up' by various carbohydrate active enzymes to yield a part for growth and a part for the formation of the required UDP-sugar. In this perspective, Table 7.1 gives an overview of disaccharides commonly used in growth media and which activated intermediates they form when split by a synthase, phosphorylase or hydrolase. For

example, UDP-galactose can be generated in different ways, other than by overexpressing the interconverting enzyme GalE. However, its direct formation by lactose synthase (EC 2.4.1.22) is not feasible since the equilibrium is completely shifted towards synthesis of lactose. Alternatively, lactose phosphorylase may be used to generate gal1P, but due to the high K_m value (228 mM) and low activity[141], strains expressing this enzyme show slow growth on lactose. Finally, lactose may be hydrolyzed by the naturally occurring β-galactosidase LacZ in *E. coli*, which yields intracellular galactose and glucose. The latter is used as a substrate for growth, while galactose is converted to UDP-galactose by combined action of GalKM (galactose kinase/mutarotase) and the promiscuous UgpA.

Table 7.1: Overview of other metabolic strategies for coupling production to growth, represented as a function of the disaccharide substrates and the carbohydrate active enzyme class. EC numbers of the enzymes involved are also shown. N.D.= not described.

	Synthase	Phosphorylase	Hydrolase
Sucrose	UDP-glucose *EC 2.4.1.13* fructose	glucose 1-phosphate *EC 2.4.1.7* fructose	glucose *EC 3.2.1.26* fructose
Lactose	UDP-galactose *EC 2.4.1.22* glucose	galactose 1-phosphate *EC 2.4.1.20* glucose	galactose *EC 3.2.1.23* glucose
Trehalose	UDP-glucose *EC 2.4.1.245* glucose	glucose 1-phosphate *EC 2.4.1.231* glucose	glucose *EC 3.2.1.28* glucose
Maltose	N.D.	β-glucose 1-phosphate *EC 2.4.1.8* glucose	glucose *EC 3.2.1.20* glucose

To evaluate the potential of this newly designed galactosylation pathway starting from lactose (depicted in Figure 7.4), two *E. coli* mutants were created. *E. coli* W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*galETKM* was used as base strain for the expression plasmids pGalKM/LgtC/UgpA and pGalKM/F3gt/UgpA, resulting in a globotriose and hyperoside producer respectively. Both strains grew on lactose as only carbon source, albeit relatively slow, but showed no production of the respective galactosides. In the case of globotriose, this could be attributed due to competition between the substrate and acceptor, which are both lactose. In the case of hyperoside, a glucose/galactose and glycerol/galactose medium were tested, but without success. A probable explanation lies in the fact that the intermediate gal1P accumulates in the cell, which is toxic at very low concentrations (galactosemia). A possible solution for the reduction of intracellular

gal1P is fine-tuning of the production pathway, by altering the promoter strengths of *galK* and *ugpA*, for example. An unconfirmed hypothesis about the toxicity of gal1P is that it competitively inhibits the essential enzyme phosphoglucosamine mutase (GlmM), thus preventing peptidoglycan synthesis and growth. Another solution in this context could be the search for an alternative GlmM that is not inhibited[407].

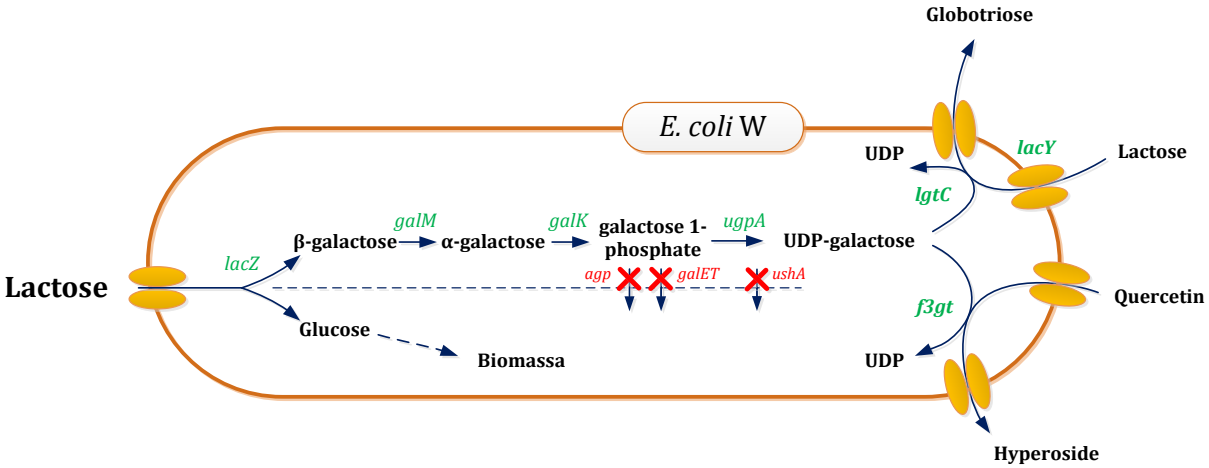


Figure 7.4: Alternative strategy for the galactosylation of small molecules starting from lactose. The endogenous β-galactosidase (LacZ) is used to generate intracellular galactose, which is converted to UDP-galactose by the combined action of GalKM and the promiscuous UgpA. Deletion of *galET*, *agp* and *ushA* again give rise to a split metabolism whereby glucose is used as a carbon source.

In conclusion, all routes towards a novel *in vivo* glycosylation platform were explored whereby a sucrose based approach was identified as the most promising. Metabolic engineering of *E. coli* W resulted in a versatile plug and play strain for the efficient glucosylation, galactosylation and rhamnosylation of small molecules. This new approach enabled the engineered host to grow and produce glycosides simultaneously, hereby using only sucrose as a cheap and sustainable carbon source. As a result, we obtained the highest product rates and titers reported to date, which outcompetes the majority of existing glycosylation processes. The ease of scaling up the created glycosylation strains truly paves the way for an economically viable production of various glycosides, thus meeting the envisioned objectives of this PhD thesis. Merging these novel insights and engineering approaches with existing microbial cell factories for production of aglycons will unlock a new era for sustainable *de novo* glycoside formation.

Appendices

A1. Plasmids created and used

Plasmid vectors	Description
pCX-Kan-P22 [a]	Constitutive expression vector with P22 promoter, Amp^r, Kan^r
pCX-Kan-P22-BaSP [a]	pCX-Kan-P22 vector carrying *BaSP* (sucrose phosphorylase) from *B. adolescentis*
pCX-Kan-P22-GalU	pCX-Kan-P22 vector carrying *galU* from *E. coli*
pCX-Kan-P22-UgpA	pCX-Kan-P22 vector carrying *ugpA* from *B. bifidum*
pCX-Kan-P22-GalK	pCX-Kan-P22 vector carrying *galK* from *E. coli*
pCX-Kan-P22-VvGT2	pCX-Kan-P22 vector carrying *vvGT2* from *Vitis vinifera*
pKan-L5-P22-LgtC-L6	pUC57-Kan vector carrying C-25 truncated *lgtC* from *N. meningitidis* flanked by homologous linkers L5 and L6 and P22
pKan-L5-P22-F3GT-L6	pUC57-Kan vector carrying codon optimized *f3gt* from *Petunia hybrida* flanked by homologous linkers L5 and L6 and P22
pKan-L5-P22-RhaGT-L6	pUC57-Kan vector carrying codon optimized *AtUGT78D1* from *A. thaliana* flanked by homologous linkers L5 and L6 and P22
pCX-Kan-P22-SuSy	pCX-Kan-P22 vector carrying codon optimized *susy* isoform 2 from S*olanum tuberosum*
pCX-Kan-P22-SuSy-S11D	pCX-Kan-P22-SuSy vector with Ser^{11} changed to aspartic acid
pCX-Kan-P22-SuSy-S11E	pCX-Kan-P22-SuSy vector with Ser^{11} changed to glutamic acid
p10-Trc [a]	Inducible expression vector, Cm^r, low copy number (± 10)
p10-Trc-galT1	p10-Trc carrying *galT1* from *B. bifidum*
p10-Trc-galT2	p10-Trc carrying *galT2* from *B. bifidum*
p10-Trc-His-galT1	p10-Trc carrying *galT1* from *B. bifidum* with His_6 tag
p10-Trc-His-galT2	p10-Trc carrying *galT2* from *B. bifidum* with His_6 tag
pUC-L_n-RFP-L_{n+1} [a]	pUC57 vector carrying RFP sequence flanked by 200 bp homologous linkers, Amp^r
pUC-L4-P22-BaSP-L5	pUC57 vector constitutively expressing *BaSP* flanked by homologous linkers L4 and L5
pUC-L4-P22-GalE-L5	pUC57 vector constitutively expressing *galE* from *E. coli* flanked by homologous linkers L4 and L5
pUC-L4-P22-GalE2-L5	pUC57 vector constitutively expressing *galE2* from *B. bifidum* flanked by homologous linkers L4 and L5
pUC-L4-P22-GalK-L5	pUC57 vector constitutively expressing *galK* flanked by homologous linkers L4 and L5
pUC-L4-P22-GalKM-L5	pUC57 vector constitutively expressing *galKM* from *E. coli* flanked by homologous linkers L4 and L5
pUC-L5-P22-VvGT2-L6	pUC57 vector constitutively expressing *vvGT2* flanked by homologous linkers L5 and L6
pUC-L6-P22-GalU-L7	pUC57 vector constitutively expressing *galU* flanked by homologous linkers L6 and L7
pUC-L6-P22-UgpA-L7	pUC57 vector constitutively expressing *ugpA* flanked by homologous linkers L6 and L7
pUC-L4-P22-MUM4-L5	pUC57 vector constitutively expressing codon optimized *MUM4* from *A. thaliana* flanked by homologous linkers L4 and L5

pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6	pCX-Kan vector constitutively expressing *BaSP* and *vvGT2* flanked by homologous linkers
pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-GalU-L7	pCX-Kan vector constitutively expressing *BaSP*, *vvGT2* and *galU* flanked by homologous linkers
pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *BaSP, vvGT2* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-BaSP-L5-P22-LgtC-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *BaSP, lgtC* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalE-L5-P22-LgtC-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galE, lgtC* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalE2-L5-P22-LgtC-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galE2*, *lgtC* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-BaSP-L5-P22-F3GT-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *BaSP*, *f3gt* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalE-L5-P22-F3GT-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galE, f3gt* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalE2-L5-P22-F3GT-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galE2*, *f3gt* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-MUM4-L5-P22-RhaGT-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *MUM4, AtUGT78D1* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-SuSy-S11E-L5-P22-VvGT2-L6	pCX-Kan vector constitutively expressing *susy* S11E and *vvGT2* flanked by homologous linkers
pCX-Kan-L4-P22-GalK-L5-P22-LgtC-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galK*, *lgtC* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalK-L5-P22-F3GT-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galK*, *f3gt* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalKM-L5-P22-LgtC-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galKM*, *lgtC* and *ugpA* flanked by homologous linkers
pCX-Kan-L4-P22-GalKM-L5-P22-F3GT-L6-P22-UgpA-L7	pCX-Kan vector constitutively expressing *galKM*, *f3gt* and *ugpA* flanked by homologous linkers
pKD46 [b]	λ Red recombinase expression, Amp^r
pCP20 [b]	FLP recombinase expression, Amp^r, Cm^r
pKD3 [b]	Cm cassette template, Cm^r, Amp^r
pKD4 [b]	Kan cassette template, Kan^r, Amp^r

[a] plasmids were in house

[b] plasmids obtained from Datsenko and Wanner (2000)

A2. Strains

Strain (CODE)	Description
E. coli DH5α [a]	General cloning host
E. coli DH5α Competent Cells [a]	Subcloning Efficiency Competent Cells
E. coli MG1655 [a]	*Escherichia coli* λ^-, F^-, rph-1
E. coli BL21 (DE3) [a]	*Escherichia coli* F^- λ(DE3 [lacI lacUV5-T7])
E. coli W	*Escherichia coli* W LMG 11080
B. bifidum	*Bifidobacterium bifidum* (Tissier 1900) LMG 11041
E. coli MG1655 Δ*galETKM* Δ*galU* Δ*ushA* Δ*ugd* Δ*agp*(**sMEMO_WT**)	*E. coli* MG1655 with *galETKM, galU, ushA, ugd,* and *agp* -deleted
sMEMO_UgpA	sMEMO_WT carrying pCX-Kan-P22-UgpA
sMEMO_GalU	sMEMO_WT carrying pCX-Kan-P22-GalU
sMEMO_GalT1	sMEMO_WT carrying p10-Trc-GalT1
sMEMO_GalT2	sMEMO_WT carrying p10-Trc-GalT2
sMEMO_His_GalT1	*E. coli* BL21 (DE3) carrying p10-Trc-His-GalT1
sMEMO_His_GalT2	*E. coli* BL21 (DE3) carrying p10-Trc-His-GalT2
E. coli W Δ*cscAR* (**sGLC1**)	*E. coli* W with *cscAR*-deleted
E. coli W Δ*cscAR* Δ*melA*::L4-P22-*BaSP*-L5	*E. coli* W with *cscAR* and *melA* deleted and L4-P22-BaSP-L5 integrated in genome
E. coli W Δ*cscAR* Δ*glgC*::L4-P22-*BaSP*-L5	*E. coli* W with *cscAR* and *glgC* deleted and L4-P22-BaSP-L5 integrated in genome
E. coli W Δ*cscAR* Δ*pgm* (**sGLC2**)	*E. coli* W with *cscAR*, and *pgm*-deleted
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* (**sGLC3**)	*E. coli* W with *cscAR*, *pgm*, and *agp*-deleted
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* (**sGLC4**)	*E. coli* W with *cscAR*, *pgm*, *agp*, and *ushA*-deleted
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* (**sGLC5**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA* and *glgC*-deleted
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* Δ*yihX::cat* (**sGLC6**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA* , *glgC and yihX* –deleted with Cm cassette at *yihX* position
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* Δ*yidA::cat* (**sGLC7**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA* , *glgC and yidA* –deleted with Cm cassette at *yidA* position
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*galETKM* (**sGLYC4**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA* and *galETKM*-deleted
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* (**sGLYC1**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA* and *lacZYA* deleted and P22-*lacY* integrated in genome
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galTKM* (**sGLYC3**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA, lacZYA* and *galTKM* deleted and P22-*lacY* integrated in genome
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*melA*::L4-P22-*BaSP*-L5	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA, lacZYA* and *melA* deleted and P22-*lacY* and L4-P22-BaSP-L5 integrated in genome

E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*melA*::L4-P22-*BaSP*-L5 Δ*galETKM* (**sGLYC2**)	*E. coli* W with *cscAR*, *pgm*, *agp*, *ushA, lacZYA, melA* and *galETKM* deleted and P22-*lacY* and L4-P22-BaSP-L5 integrated in genome
E. coli W + pVvGT2	*E. coli* W carrying pCX-Kan-P22-VvGT2
E. coli W + pLgtC	*E. coli* W carrying pKan-L5-P22-LgtC-L6
E. coli W + pF3GT	*E. coli* W carrying pKan-L5-P22-F3GT-L6
E. coli W + pRhaGT	*E. coli* W carrying pKan-L5-P22-RhaGT-L6
E. coli W Δ*cscAR* + pBaSP	sGLC1 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* + pSuSy	sGLC1 carrying pCX-Kan-P22-SuSy
E. coli W Δ*cscAR* + pSuSy-S11D	sGLC1 carrying pCX-Kan-P22-SuSy-S11D
E. coli W Δ*cscAR* + pSuSy-S11E	sGLC1 carrying pCX-Kan-P22-SuSy-S11E
E. coli W Δ*cscAR* Δ*pgm* + pBaSP	sGLC2 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* + pBaSP	sGLC3 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* + pBaSP	sGLC4 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* + pBaSP	sGLC5 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* Δ*yihX::cat* + pBaSP	sGLC6 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* Δ*yidA::cat* + pBaSP	sGLC7 carrying pCX-Kan-P22-BaSP
E. coli W Δ*cscAR* + pBaSP/VvGT2	sGLC1 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* + pBaSP/VvGT2	sGLC3 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* + pBaSP/VvGT2	sGLC4 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6
E. coli W Δ*cscAR* + pBaSP/VvGT2/GalU	sGLC1 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-GalU-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* + pBaSP/VvGT2/GalU	sGLC3 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-GalU-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* + pBaSP/VvGT2/GalU	sGLC4 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-GalU-L7
E. coli W Δ*cscAR* + pBaSP/VvGT2/UgpA	sGLC1 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* + pBaSP/VvGT2/UgpA	sGLC3 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* + pBaSP/VvGT2/UgpA	sGLC4 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* + pBaSP/VvGT2	sGLC5 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* + pBaSP/VvGT2/GalU	sGLC5 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-GalU-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* + pBaSP/VvGT2/UgpA	sGLC5 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7

E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* Δ*yihX::cat* + pBaSP/VvGT2/UgpA	sGLC6 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*glgC* Δ*yidA::cat* + pBaSP/VvGT2/UgpA	sGLC7 carrying pCX-Kan-L4-P22-BaSP-L5-P22-VvGT2-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* + pBaSP/LgtC/UgpA	sGLYC1 carrying pCX-Kan-L4-P22-BaSP-L5-P22-LgtC-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galTKM* + pBaSP/LgtC/UgpA	sGLYC3 carrying pCX-Kan-L4-P22-BaSP-L5-P22-LgtC-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE/LgtC/UgpA	sGLYC2 carrying pCX-Kan-L4-P22-GalE-L5-P22-LgtC-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE2/LgtC/UgpA	sGLYC2 carrying pCX-Kan-L4-P22-GalE2-L5-P22-LgtC-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE/F3GT/UgpA	sGLYC2 carrying pCX-Kan-L4-P22-GalE-L5-P22-F3GT-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pGalE2/F3GT/UgpA	sGLYC2 carrying pCX-Kan-L4-P22-GalE2-L5-P22-F3GT-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pRhaGT + pMUM4	sGLYC2 carrying pCX-Kan-L5-P22-RhaGT-L6 and pUC-L4-P22-MUM4-L5
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*lacZYA*::P22-*lacY* Δ*galETKM* Δ*melA*::P22-BaSP + pMUM4/RhaGT/UgpA	sGLYC2 carrying pCX-Kan-L4-P22-MUM4-L5-P22-RhaGT-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*galETKM* + pGalK/LgtC/UgpA	sGLYC4 carrying pCX-Kan-L4-P22-GalK-L5-P22-LgtC-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*galETKM* + pGalK/F3GT/UgpA	sGLYC4 carrying pCX-Kan-L4-P22-GalK-L5-P22-F3GT-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*galETKM* + pGalKM/LgtC/UgpA	sGLYC4 carrying pCX-Kan-L4-P22-GalKM-L5-P22-LgtC-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* Δ*galETKM* + pGalKM/F3GT/UgpA	sGLYC4 carrying pCX-Kan-L4-P22-GalKM-L5-P22-F3GT-L6-P22-UgpA-L7
E. coli W Δ*cscAR* Δ*pgm* Δ*agp* Δ*ushA* + pSuSy/VvGT2	sGLC4 carrying pCX-Kan-L4-P22-SuSy-S11E-L5-P22-VvGT2-L6

[a] strains were in house

A3. Accession numbers of cloned genes

Gene	Acc. number	Remarks
galU	NC_000913	Amplified from gDNA
ugpA	KC261357	Amplified from gDNA. Deposited at GenBank.
galT1	KC261358	Amplified from gDNA. Deposited at GenBank.
galT2	KC261359	Amplified from gDNA. Deposited at GenBank.
BaSP	NC_008618	Sequence has a N-terminal hexahistidine tag
vvGT2	JN164680	Sequence was provided by N. Terrier (INRA)
lgtC	U65788	Sequence was C-terminal truncated (25 AA) and ordered
f3gt	AF165148	Sequence was codon optimized and ordered
MUM4	NC_003070	Sequence was codon optimized and ordered
AtUGT78D1	NC_003070	Sequence was codon optimized and ordered
galK	NC_000913	Amplified from gDNA.
galM	NC_000913	Amplified from gDNA.
galE	NC_000913	Amplified from gDNA.
galE2	KJ543703	Amplified from gDNA. Deposited at GenBank.
susy	P49039	Sequence was codon optimized and ordered

A4. Reaction overview

Enzyme	Reaction			EC code
Agp	Glucose 1-phosphate + H_2O	⟶	Glucose + Phosphate	3.1.3.10
BaSP	Sucrose + phosphate	⟷	Fructose + Glucose 1-phosphate	2.4.1.7
CscA	Sucrose + H_2O	⟶	Glucose + Fructose	3.2.1.26
CscB	$Sucrose_{[extracellular]}$ + H^+	⟶	$Sucrose_{[intracellular]}$ + H^+	-
CscK	Fructose + ATP	⟶	Fructose 6-phosphate + H_2O + ADP	2.7.1.4
F3GT	Flavonol + UDP-galactose	⟷	Flavonol-3-*O*-galactoside + UDP	2.4.1.234
GalE	UDP-glucose	⟷	UDP-galactose	5.1.3.2
GalE2	UDP-glucose	⟷	UDP-galactose	5.1.3.2
GalK	Galactose + ATP	⟶	Galactose 1-phosphate + ADP	2.7.1.6
GalM	α-D-galactose	⟷	β-D-galactose	5.1.3.B1
GalT(1)	Galactose 1-phosphate + UDP-glc	⟷	UDP-gal + glucose 1-phosphate	2.7.7.12
GalT2	galNAc1P + UDP-glcNAc	⟷	glcNAc1P + UDP-galNAc	2.7.7.12
GalU	UTP + Glucose 1-phosphate	⟷	UDP-glucose + pyrophosphate	2.7.7.9
GlgC	ATP + Glucose 1-phosphate	⟷	ADP-glucose + pyrophosphate	2.7.7.27
Glk	$Glucose_{[intracellular]}$ + ATP	⟶	Glucose 6-phosphate + ADP	2.7.1.2
LgtC	$Lactose_{[moiety]}$ + UDP-galactose	⟶	$Globotriose_{[moiety]}$ + UDP	2.4.1.44
Mak	$Fructose_{[intracellular]}$ + ATP	⟶	Fructose 6-phosphate + ADP	2.7.1.4
MelA	Globotriose + H_2O	⟶	Lactose + Galactose	3.2.1.22
MUM4	UDP-glucose + NADPH + H^+	⟶	UDP-rhamnose + $NADP^+$ + H_2O	-
Ndk	Nucleoside diphosphate + ATP	⟷	Nucleoside triphosphate + ADP	2.7.4.6
Pgi	Glucose 6-phosphate	⟷	Fructose 6-phosphate	5.3.1.9
Pgm	Glucose 1-phosphate	⟷	Glucose 6-phosphate	5.4.2.2
RhaGT	Flavonol + UDP-rhamnose	⟶	Flavonol-3-*O*-rhamnoside + UDP	2.4.1.-
SuSy	Sucrose + UDP	⟷	UDP-glucose + Fructose	2.4.1.13
Ugd	UDP-glucose + 2 NAD^+ + H_2O	⟷	UDP-glucuronate + 2 NADH + 3 H^+	1.1.1.22
UgpA	UTP + sugar 1-phosphate	⟷	pyrophosphate + UDP-sugar	2.7.7.64
UshA	UDP-glucose + H_2O	⟷	UMP + glucose 1-phosphate + 2 H^+	3.6.1.45
	UMP + H_2O	⟷	uridine + phosphate	3.1.3.5
VvGT2	UDP-glucose + gallic acid	⟶	UDP + 1-galloyl-β-D-glucose	2.4.1.136

A5. Primers for knockouts and knockins

Primers [a]	Oligonucleotide sequences (5'-3') [b]
Fw_galETKM_KO_MG1655	CTGGTGATTTGAACAATATGAGATAAAGCCCTCATGACGAGGGCGTAACAgtgtaggctggagctgcttc
Fw_galETKM_KO_W	GTGATTTGAGCAATATAAGGATAAAGCCCTCATTACGAGGGCTTCAGGTGgtgtaggctggagctgcttc
Rv_galETKM_KO	CTTTGTTATGCTATGGTTATTTCATACCATAAGCCTAATGGAGCGAATTATGcatatgaatatcctccttag
Fw_galU_KO	GCGATACAGAAATATGAACACGTTCAAAACACGAACAGTCCAGGAGAATTTAAgtgtaggctggagctgcttc
Rv_galU_KO	ACGGCGTCGATTGCTCAACGCCGTTTCGTGGATAACACCGATACGGATGcatatgaatatcctccttag
Fw_ushA_KO	TCGCGTCATACTATTTTTCAACACGTTGAAATCAGGTCAGGGAGAGAAGTgtgtaggctggagctgcttc
Rv_ushA_KO	CCCGCCGCGATTAAGCATTGTGCCGGATGCAAACATCCGGCACTTTCGGAcatatgaatatcctccttag
Fw_ugd_KO	TGTAAGTAACAAAAGACAATCAGGGCGTAAATAGCCCTGATAACAAGATGgtgtaggctggagctgcttc
Rv_ugd_KO	GATGCTAAAAACATCATGATTCACAGTTAAGTTAATTCTGAGAGCATGAAcatatgaatatcctccttag
Fw_agp_KO	CATATTTCTGTCACACTCTTTAGTGATTGATAACAAAAGAGGTGCCAGGAgtgtaggctgagctgcttc
Rv_agp_KO_MG1655	TAAAAACGTTTAACCAGCGACTCCCCCGCTTCTCGCGGGGGAGTTTTCTGcatatgaatatcctccttag
Rv_agp_KO_W	TAAAAACGTTTAACCAGCGACTCCCCCACTTCTCGCGGGGGAGTTTTCTGcatatgaatatcctccttag
Fw_cscAR_KO	AAAAGTTAACGTTAACAATTCACCAAATTTGCTTAACCAGGATGATTAAAgtgtaggctggagctgcttc
Rv_cscAR_KO	GCTTGGGTATGGCTTCATTAAAGGATGTCGCACGCCTGGCGGGAGTGTCGcatatgaatatcctccttag
Fw_pgm_KO	TGAGAAGGTTCGCGGAACTATCTAAAACGTTGCAGACAAAGGACAAAGCAgtgtaggctggagctgcttc
Rv_pgm_KO	CAAAGTAAAAAAGGGCGATCTTGCGACCGCCCTTTTTTTATTAAATGTGTcatatgaatatcctccttag
Fw_glgC_KO	AGACCGCCGGTTTTAAGCAGCGGGAACATCTCTGAACATACATGTAAAACCTGCAgtgtaggctggagctgcttc
Rv_glgC_KO	GTCTGGCAGGGACCTGCACACGGATTGTGTGTGTTCCAGAGATGATAAAAAAGGAGTTAGTCcatatgaatatcctccttag
Fw_lacZYA_KO	GCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTCGCCTACCTGTGACGGAAG
Rv_lacZYA_KO	GCTGAACTTGTAGGCCTGATAAGCGCAGCGTATCAGGCAATTTTTATAATCTTCATTTAAATGGCGCGC
Fw_galTKM_KO_W	GGCACTGGCAGTCACGCCATCCACAGGGATATCCCGATTAAGGAACGACCgtgtaggctgagctgcttc
Rv_galTKM_KO_W	GTGATTTGAGCAATATAAGGATAAAGCCCTCATTACGAGGGCTTCAGGTGcatatgaatatcctccttag
Fw_yihX_KO	CCATCCTTTCCCGCTACAGTTAATTTCTTGTGGCGCGAAAGGAGGCAAAAgtgtaggctgagctgcttc

Rv_yihX_KO	CCAGGCCCATAGTGGACGGGTACGGTGCCTGGCTTTGTCCTGAATGGTTTcatatgaatatcctccttag
Fw_yidA_KO	AATAATCAGTAAGCGGGCAAACGCGTTTATGCTGTTTGCCCGTCCACTGAgtgtaggctggagctgcttc
Rv_yidA_KO	GAGCGGAATCGCGTTAGCATGGGTCAGGAACTAATCTACCTGGGGAACTCcatatgaatatcctccttag
Fw_lacY_KI	CAGGTTTCCCGACTGGAAAG
Rv_lacY_KI	TGTGCGTCGTTGGGCTGATG
Fw_L4_start_BaSP_KI	GGCAGGAGTATCGTCCGTAG
Rv_L5_tail_BaSP_KI	CTGAAGGCGCTCCTCATAC
Fw_glgC::BaSP_L	GACCATGTTCGTCACCGGCCAGTAACATCGCCTTGATAGGCCAGGTTGTG
Rv_glgC::BaSP_L	ACAGTTGAAACTACGGACGATACTCCTGCCGAAATGCTACGGAAGTTAGG
Fw_glgC::BaSP_R	GTATTCAGCCAGTATGAGGAGCGCCTTCAGGACTAACTCCTTTTTTATCATCTCTGGAAC
Rv_glgC::BaSP_R	CGATGTTACTGGCCGGTGAC
Fw_melA::BaSP_L	GGATATTCCCTTCTGGTCGCTGGTTCCAAGCGTGATATGACCCTGATTG
Rv_melA::BaSP_L	ACAGTTGAAACTACGGACGATACTCCTGCCCTCCTGGCTTGCTTGAATAAC
Fw_melA::BaSP_R	GTATTCAGCCAGTATGAGGAGCGCCTTCAGAACGCGACTAAACGCTACTG
Rv_melA::BaSP_R	TTGGAACCAGCGACCAGAAG
Fw_glgC_KO_BaSP_KI	AAGCCCGCTGGCCATTTCTG
Rv_glgC_KO_BaSP_KI	CGCCGCGTTAATCCATCTGC
Fw_melA_KO_BaSP_KI	TGCTTCACGCAGGATCTGAG
Rv_melA_KO_BaSP_KI	ACAGACAGCCCAACGACATC

[a] Notations KO: knockout, KI: knockin, MG1655: primer specific for *E. coli* MG1655 genome, W: primer specific for *E. coli* W genome.

[b] Priming sites for the kanamycin resistance marker are denoted in lower case, flanking regions in upper case.

A6. Primers for plasmid constructions and SuSy mutagenesis

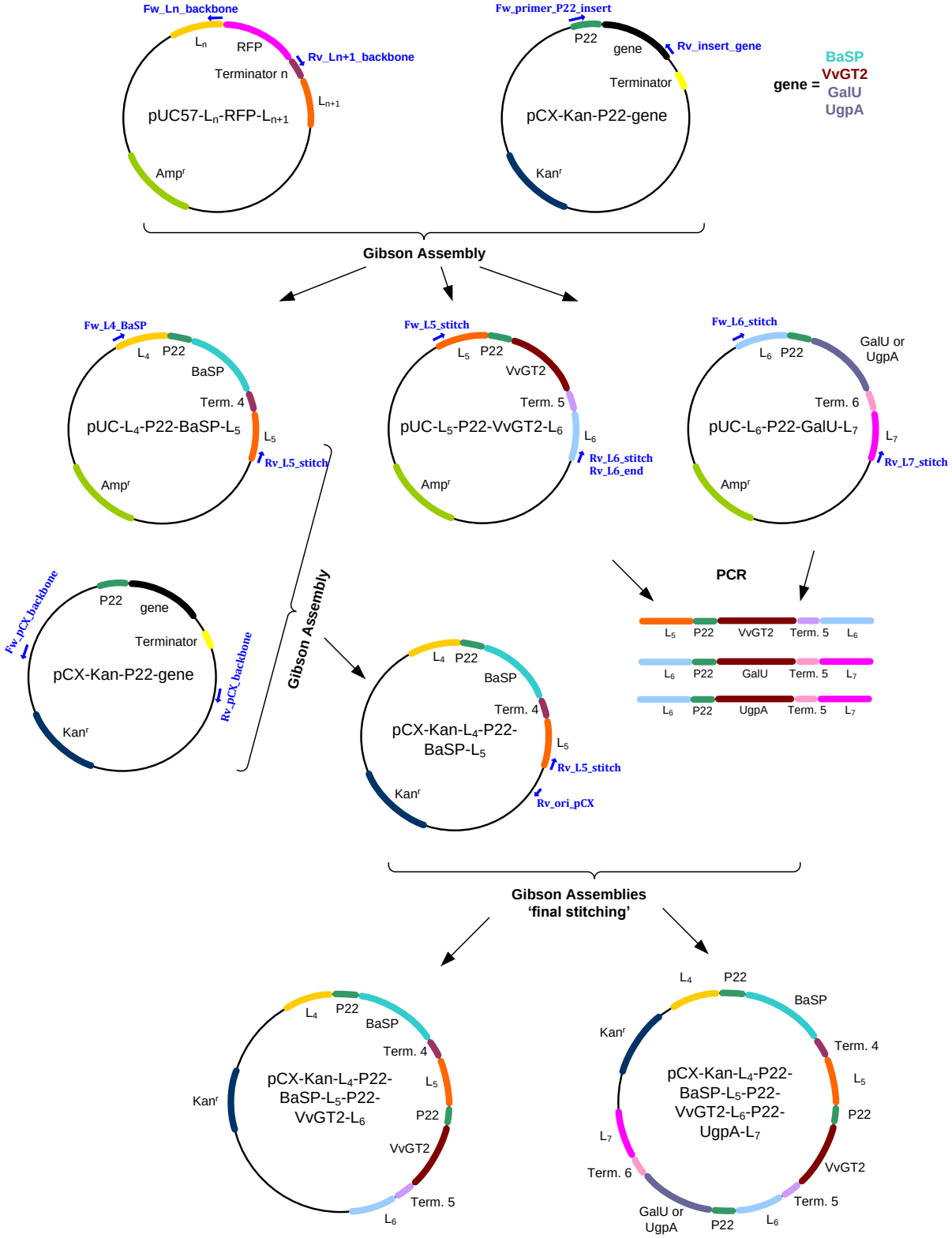
Primers	Oligonucleotide sequences (5'-3') [a]
Fw_galU_EcoRI	CCGGCGAATTCGGAGGAAACAAAGATGGCTGCCATTAATACGAAAG
Rv_galU_SacI	CGCCGAGCTCTTACTTCTTAATGCCCATCTC
Fw_ugpA_EcoRI	CCGGCGAATTCGGAGGAAACAAAGATGTTTGCCGAAGATCTGAAACG
Rv_ugpA_SacI	CGCCGAGCTCTCACACCCAATCACCGGGCTCGATG
Fw_galT1_NdeI	AGGTCGCGCATATGGCAGAAATCACCAACTACAC
Rv_galT1_PmeI	TATGTTTAAACGCTAGCCTCGAGGCGGCCGCGAGCTCTAAAGGGCCGGCCTAAACG AATTCTCAGTCGGAGATGTCGATCTG
Fw_galT2_NdeI	AGGTCGCGCATATGACCACGGAAGAGAAGAAGG
Rv_galT2_PmeI	TATGTTTAAACGCTAGCCTCGAGGCGGCCGCGAGCTCTAAAGGGCCGGCCTAAACG AATTCCTAATGCTGAGTATGGAATCCGAGGCCTTCC
Fw_galT1_his	ATGGGCGGCTCA**CACCACCACCACCACCAC**GGTATGGCGTCTATGGCAGAAATCAC CAACTAC
Fw_galT2_his	ATGGGCGGCTCA**CACCACCACCACCACCAC**GGTATGGCGTCTATGACCACGGAAGA GAAGAAG
Rv_galT1_his	CAATGATGATGATGATGATGGTCGACTCAGTCGGAGATGTCGATCTGTC
Rv_galT2_his	CAATGATGATGATGATGATGGTCGACCTAATGCTGAGTATGGAATC
Fw_backbone_p10	AGACGCCATACCGTGGTGGTGGTGGTGGTGTGAGCCGCCCATATGTTATTCCTCCT TATTTAATC
Rv_backbone_p10	GTCGACCATCATCATCATCATCATTG
Fw_L4_backbone	ATTTATAAATGAAGCGGCCGCCCGGGATAGACTTCAGGCAGACCACGCTTGAC
Rv_L5_Backbone	TAACCATGGGCTAGCATTGC
Fw_L5_backbone	ATTTATAAATGAAGCGGCCGCCCGGGATCTGAAGGCGCTCCTCATACTG
Rv_L6_Backbone	TAACCATGGGCTAGCGGCCAGCAAAG
Fw_L6_backbone	ATTTATAAATGAAGCGGCCGCCCGGGATGTTCCGATGGCGTGCATCAG
Rv_L7_Backbone	TAACCATGGGCTAGCAATGC
Univ_Fw_primer_P22 _insert	ATCCCGGGCGGCCGCTTCATTTATAAATTTC
Rv_insert_BaSP	AGTCGGAACGGCAATGCTAGCCCATGGTTATCAGGCGACGACAGGCGGATTG
Rv_insert_VvGT2	GTCCCTTTGCTGGCCGCTAGCCCATGGTTATTAAATTTTCTTTGACTTGCAAACCA GCTCCATAC
Rv_insert_galU	GGTTAAGGCCGTTTGGCATTGCTAGCCCATGGTTATTACTTCTTAATGCCCATCTC
Rv_insert_UgpA	GGTTAAGGCCGTTTGGCATTGCTAGCCCATGGTTATCACACCCAATCACCGGGCTC GATG
Fw_pCX_backbone	GAAGGCGGCGGTGGAATCGAAATC
Rv_pCX_backbone	CGTGAGTTTTCGTTCCACTGAGCGTCAGAC
Fw_L4_BaSP	TCACGAGATTTCGATTCCACCGCCGCCTTCAGACGAATTACTTATCTGGCAGGAGT ATC
Rv_BasP_L5	GTCTGACGCTCAGTGGAACGAAAACTCACGCTGAAGGCGCTCCTCATACTG
Fw_ori_pCX	CGTTCCACTGAGCGTCAGAC
Rv_L5_stitch	GCCGTTGCACTAATACAGGGTAATTC
Fw_L5_stitch	GAGGGAATTACCCTGTATTAG
Rv_L6_end	TTTCTACGGGGTCTGACGCTCAGTGGAACGGTTCCGATGGCGTGCATCAG
Rv_L6_stitch	GTAAAGCGGGCGCACCCTCTGAGAATTAAC
Fw_L6_stitch	GTTAATTCTCAGAGGGTGCGCCCGCTTTAC

Rv_L7_end	TTTCTACGGGGTCTGACGCTCAGTGGAACGCATCTTCTTCGAGTGGTCCCAGAAC
Fw_P22_galE	CCGTCGACCTCGAATTCGGAGGAAACAAAGATGAGAGTTCTGGTTACCGGTGGTAG
Rv_galE_T4	GAACGGCAATGCTAGCCCATGGTTATTAATCGGGATATCCCTGTGGATGG
Fw_P22_galE2	CCGTCGACCTCGAATTCGGAGGAAACAAAGATGACAGTTCTCGTTACCGGTGGGTG
Rv_galE2_T4	GAACGGCAATGCTAGCCCATGGTTATCAGCCGACCTGCTTCCAGGTG
Rv_bb_general	CTTTGTTTCCTCCGAATTCGAGGTC
Fw_L4_stitch	CGATGGTCAAGCGTGGTCTG
Rv_L4_stitch	AGACTTCAGGCAGACCACGC
Fw_L5-stitch_CLIVA	CGCGGA*CATGAT*TTTGATTG
Rv_L6-stitch_CLIVA	CGCTGC*GGCAAT*CTTATGG
Fw_L6-stitch_CLIVA	ATTGCC*GCAGCG*CTTTATTG
Rv_L5-stitch_CLIVA	ATCATG*TCCGCG*CCAGTTTC
Rv_SuSy_T6	CCGTTTGGCATTGCTAGCCCATGGTTATTATTCAACTGCCAGCGGAACC
Rv_L6_P22	ATTTATAAATGAAGCGGCCGCCCGGGATGTTCCGATGGCGTGCATCAG
Rv_L7_ori	TTTCTACGGGGTCTGACGCTCAGTGGAACGCATCTTCTTCGAGTGGTCCCAGAAC
Rv_L4_P22	ATTTATAAATGAAGCGGCCGCCCGGGATAGACTTCAGGCAGACCACGCTTGAC
Rv_galK_T4	GGAACGGCAATGCTAGCCCATGGTTATCAGCACTGTCCTGCTCCTTG
Fw_P22_galKM	CCGTCGACCTCGAATTCGGAGGAAACAAAGATGAGTCTGAAAGAAAAAACACAATCTC
Rv_galKM_T4	GGAACGGCAATGCTAGCCCATGGTTACTACTCAGCAATAAACTGATATTCC
Fw_pCXkanBackbone	CATCTTTGTTTCCTCCGAATTCG
Rv_pCXKanBackbone	TAAGAGCTCCCAACGCGTTGGATG
Fw_SuSy_EcoRI	GGTACCGAATTCGGAGGAAACAAAGATGGCAGAACGTGTTCTGAC
Rv_SuSy_SacI	GCCCAGGAGCTCTTATTATTCAACTGCC
Rv_backbone_S11E	ATGAACACGGGTCAGAACACG
Fw_backbone_S11E	GTAAACTGGCACAGCTGGTTCC
Fw_insert_S11E	GGAAACAAAGATGGCAGAACGTGTTCTGACCCGTGTTCATGAACTGCGTGAACG
Rv_insert_S11E	CTTATTATTCAACTGCCAGCGGAACCAGCTGTGCCAGTTTACG
Rv_Susy_end	TTGCATGCCTGCATCCAACGCGTTGGGAGCTCTTATTATTCAACTGCCAGCGGAACC
Fw_SuSyThr111	GATCCGTCGACCTCGAATTCGGAGGAAACAAAGATGACCGTTCCGGAATTTCTGCAGTTTAAAG
Fw_Susy_Pr139	GATCCGTCGACCTCGAATTCGGAGGAAACAAAGATGCCGTTTACCGCAAGCTTTCC
Fw_Susy_Gly157	GATCCGTCGACCTCGAATTCGGAGGAAACAAAGATGGGTGTGGAATTTCTGAATCGTCATC
Fw_Susy_Thr243	GATCCGTCGACCTCGAATTCGGAGGAAACAAAGATGACCGCAGAACGTGTGCTGGAAATG
Fw_Susy_Asp262	GATCCGTCGACCTCGAATTCGGAGGAAACAAAGATGGATAGCTGTACCCTGGAAAAATTTCTG
Fw_Susy_Ile273	GATCCGTCGACCTCGAATTCGGAGGAAACAAAGATGATTCCGATGGTGTTCAATGTTG
Fw_SuSy_begin	CTTGGGATCCGTCGACCTCGAATTC
Fw_dgBackboneA	GGCACTGGAACGTGAAATGC
Rv_dgBackboneA	CATTTTCTTGGGCAAAATAACCATGCGGACTC
Fw_dgloopA	CCGCATGGTTATTTTGCCCAAGAAAATGNNNNNNNNNNNNNNNNNNNNNNNNNNNGGTTGTTTATATTCTGGATCAGGTTCCG
Rv_dgloopA	GTTTCAGCATTTCACGTTCCAGTGCCGGAACCTGATCCAGAATATAAACAACC
Fw_ins_helixNa	TTCTGAGTCCGCATGGTTATTTTGCCCAAGAAAATGTTCTGGGTTATCCGGATACCGGTGGTC

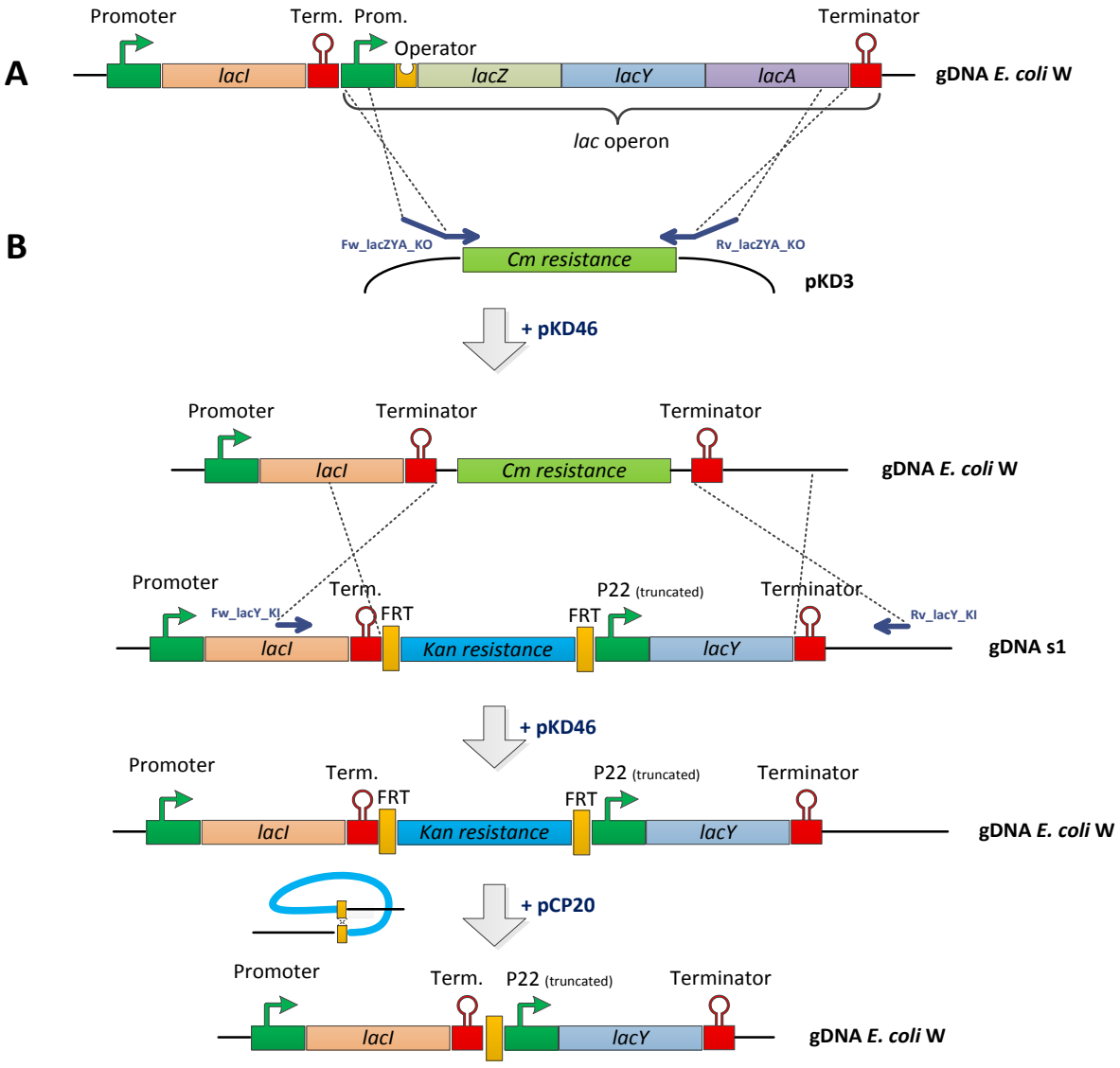
Rv_ins_helixNa	TGGTGCCAACTGCATCCGGCAGCAGACGGGTAACAATCAGAATACGCGGTTTAATG TCCAG
Rv_bb_helixNa	CCAGAACATTTTCTTGGGC
Fw_bb_helixNa	TGATTGTTACCCGTCTGCTG
Fw_ins_loopA	GCACCGGATAGCTGTACCCTGGAAAAATTTCTGGGTCGTATTCCGATGGTGTTCAA TGTTGTTATTCTG
Rv_ins_loopA	CGGTACGAAACGGCACACGCAGAATATGGCTATGTTCGGTGCCAAAAACTTTTTCC AGACGC
Fw_bb_loopA	CCGAACATAGCCATATTCTGCGTGTG
Rv_bb_loopA	AATACGACCCAGAAATTTTTCCAGGG

[a] **Restriction sites are underlined. His6 tags are underlined and in bold.**

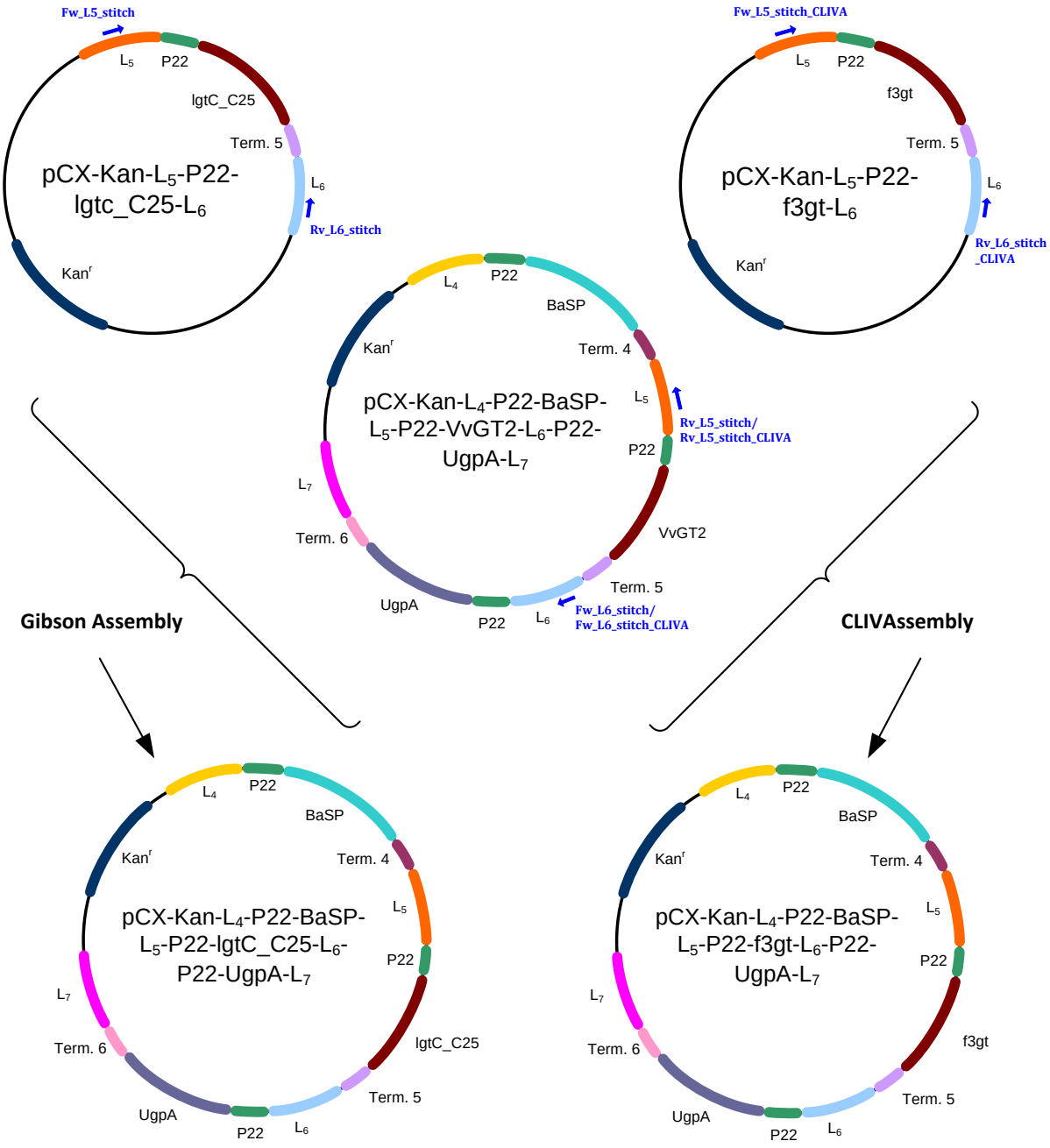
A7. Cloning strategy for plasmids of the glucosylation platform

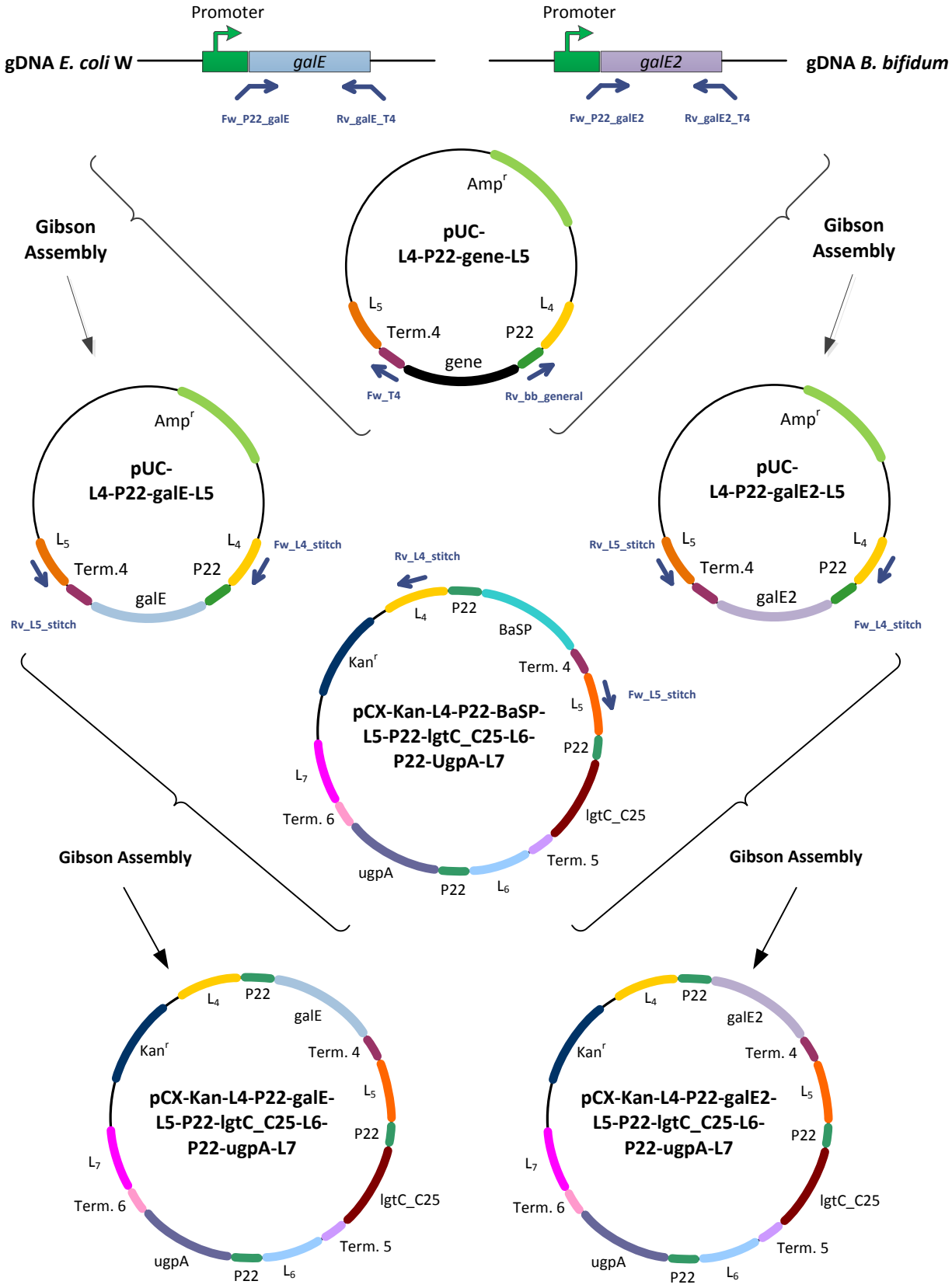


A8. Cloning strategy for constitutive *lacY* knockin



A9. Cloning strategy for plasmids of the galactosylation platform




Promoter
gDNA E. coli W
galE
Fw_P22_galE
Rv_galE_T4
Promoter
galE2
gDNA B. bifidum
Fw_P22_galE2
Rv_galE2_T4
Amp^r
pUC-
L4-P22-gene-L5
L5
L4
Term.4
P22
gene
Fw_T4
Rv_bb_general
Gibson
Assembly
Gibson
Assembly
Amp^r
pUC-
L4-P22-galE-L5
L5
L4
Term.4
P22
galE
Fw_L4_stitch
Rv_L5_stitch
Amp^r
pUC-
L4-P22-galE2-L5
Rv_L5_stitch
L5
L4
Term.4
P22
galE2
Fw_L4_stitch
Rv_L4_stitch
L4
P22
BaSP
Kan^r
Term. 4
L5
Fw_L5_stitch
pCX-Kan-L4-P22-BaSP-
L5-P22-lgtC_C25-L6-
P22-UgpA-L7
P22
L7
Term. 6
lgtC_C25
Term. 5
ugpA
P22
L6
Gibson Assembly
Gibson Assembly
L4
P22
galE
Kan^r
Term. 4
L5
pCX-Kan-L4-P22-galE-
L5-P22-lgtC_C25-L6-
P22-ugpA-L7
P22
L7
Term. 6
lgtC_C25
Term. 5
ugpA
P22
L6
L4
P22
galE2
Kan^r
Term. 4
L5
pCX-Kan-L4-P22-galE2-
L5-P22-lgtC_C25-L6-
P22-ugpA-L7
P22
L7
Term. 6
lgtC_C25
Term. 5
ugpA
P22
L6

A10. Cloning strategy for KO/KI of Δ*melA*::*BaSP* and Δ*glgC*::*BaSP*

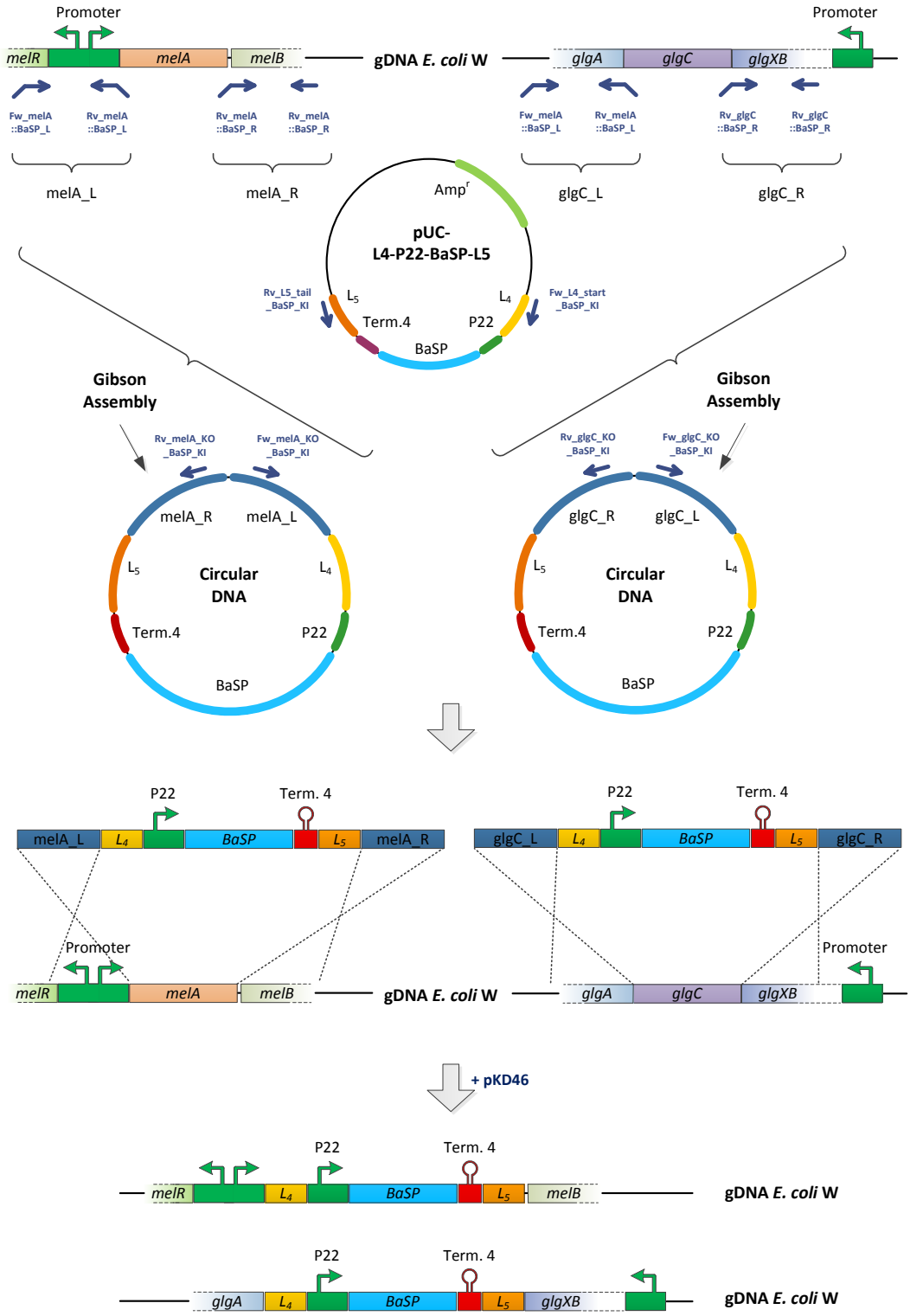
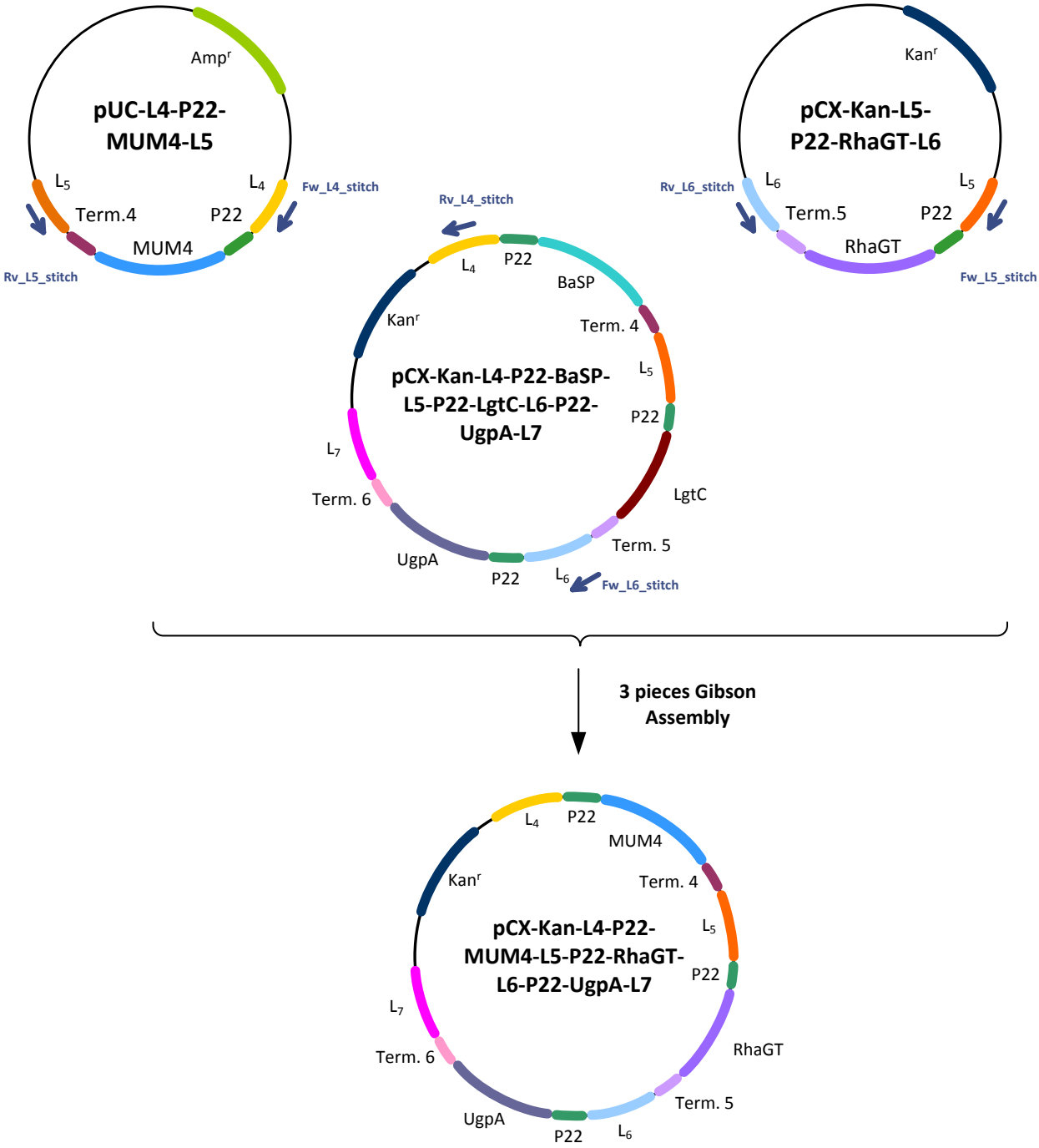


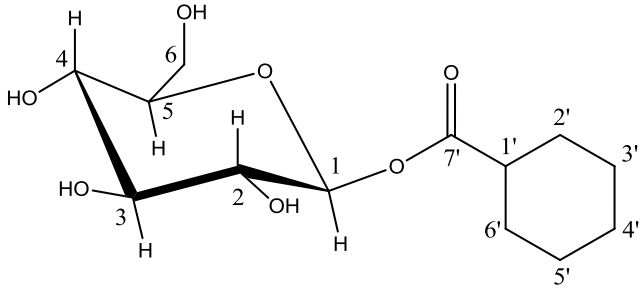
Figure A10. ***BaSP*** **knockin strategy at two different loci, resulting in two new strains. First,** ***BaSP*** **and its surrounding linker sequences, were amplified from the plasmid pUC-L4-P22-BaSP-L5 using PCR. Adjacent DNA regions of the integration spots, genes** ***melA*** **and** ***glgC*****, were also amplified from the genome. The primers were designed with the reverse primer of the left border and the forward primer of the right border having an overhanging sequence homologous to respectively linker L4 and L5. Also the forward primer of the left border has an overhang sequence homologous to the right end of the right border. For both the** ***melA*** **and** ***glgC*** **gene replacement, these 3 pieces were put together using Gibson Assembly. Subsequently, these circular DNA molecules were used as a template to pick up the knockin-knockout constructs containing prolonged homologous regions for recombination.**

A11. Cloning strategy for plasmid of the rhamnosylation platform



A12. NMR spectra of novel glucose esters

1-*O*-cyclohexanoyl-β-D-glucose



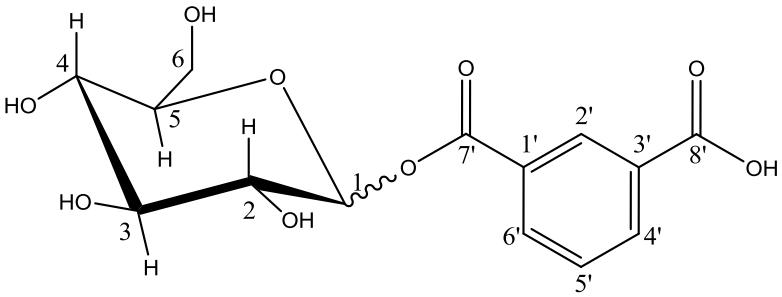
^{1}H-NMR (DMSO-d6, 400 MHz)

δ 1.13 – 1.41 (5H, m, H2'a, H3'a, H4'a, H5'a, H6'a),
1.56 – 1.60 (1H, m, H4'b)
1.66 – 1.69 (2H, m, H3'b, H5'b)
1.79 – 1.91 (2H, m, H2'b, H6'b)
2.24 – 2.36 (1H, m, H1')
3.06 – 3.13 (2H, m, H2, H4)
3.16 – 3.24 (2H, m, H3, H5)
3.40 – 3.46 (1H, m, H6a)
3.61 – 3.66 (1H, m, H6b)
4.53 (1H, dxd, J_1 = 5.9 Hz, J_2 = 5.9 Hz, C6-OH)
4.98 (1H, d, J = 5.4 Hz, OH)
5.07 (1H, d, J = 5.0 Hz, OH)
5.21 (1H, d, J = 5.5 Hz, OH)
5.31 (1H, d, J = 8.1 Hz, H1) ppm

^{13}C-NMR (DMSO-d6, 100.6 MHz)

δ 24.6 – 24.7 (C3', C5')
25.3 (C4')
28.3 – 28.5 (C2', C6')
42.1 (C1')
60.5 (C6)
69.5 (C4)
72.5 (C2)
76.4 (C5)
77.9 (C3)
94.1 (C1)
173.9 (C7') ppm

1-*O*-isophthaloyl-β-D-glucose



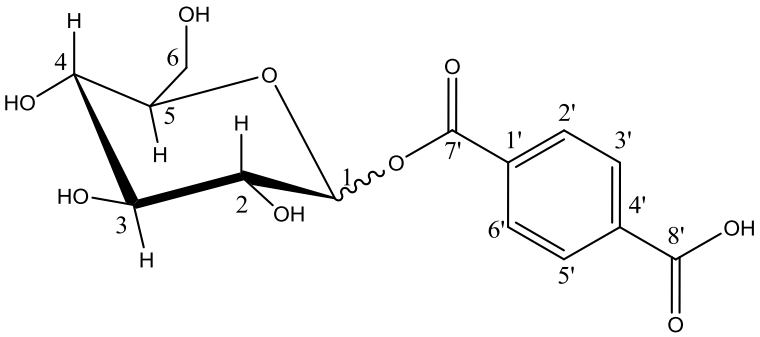
^{1}H-NMR (Acetone-d6, 400 MHz)

δ 2.82 (OH)
3.51 – 3.58 (4H, m, H2, H3, H4, H5)
3.69 – 3.73 (1H, m, H6a)
3.83 – 3.86 (1H, m, H6b)
5.80 (1H, d, $J_{equatoriaal}$ = 2.9 Hz, J_{axiaal} = 8.0 Hz, H1)*
7.72 (1H, dxd, J_1 = 7.8 Hz, J_2 = 7.8 Hz, H5')
8.30 – 8.33 (2H, m, H4', H6')
8.70 (1H, s, H2') ppm
* over 80 % of the compound had glucose in an equatorial position.

^{13}C-NMR (Acetone-d6, 100.6 MHz)

δ 62.3 (C6)
71.0 (C4)
73.8 (C2)
77.7 (C5)
78.6 (C3)
96.3 (C1)
130.0 (C5')
131.2 (C1' or C3')
131.5 (C2')
132.0 (C1' or C3')
134.8 (C4')
135.2 (C6')
164.9 (C7')
166.7 (C8') ppm

1-*O*-terephthaloyl-β-D-glucose



^{1}H-NMR (DMSO-d6, 400 MHz)

δ 3.16 – 3.21 (1H, m, H4)

3.30 – 3.32 (3H, m, H2, H3, H5)

3.47 (1H, m, H6a)

3.64 – 3.69 (1H, m, H6b)

5.61 (1H, d, $J_{equatoriaal}$ = 3.5 Hz, J_{axiaal} = 7.8 Hz, H1)*

8.07 – 8.13 (4H, m, H2', H3', H5', H6') ppm

* over 90 % of the compound had glucose in an equatorial position.

^{13}C-NMR (DMSO-d6, 100.6 MHz)

δ 60.7 (C6)

69.6 (C4)

72.6 (C2)

76.4 (C5)

78.1 (C3)

95.4 (C1)

129.7 (C2', C6' or C3', C5')

129.9 (C2', C6' or C3', C5')

132.6 (C4')

136.0 (C1')

164.3 (C7')

167.0 (C8') ppm

A13. Growth and glucose ester formation

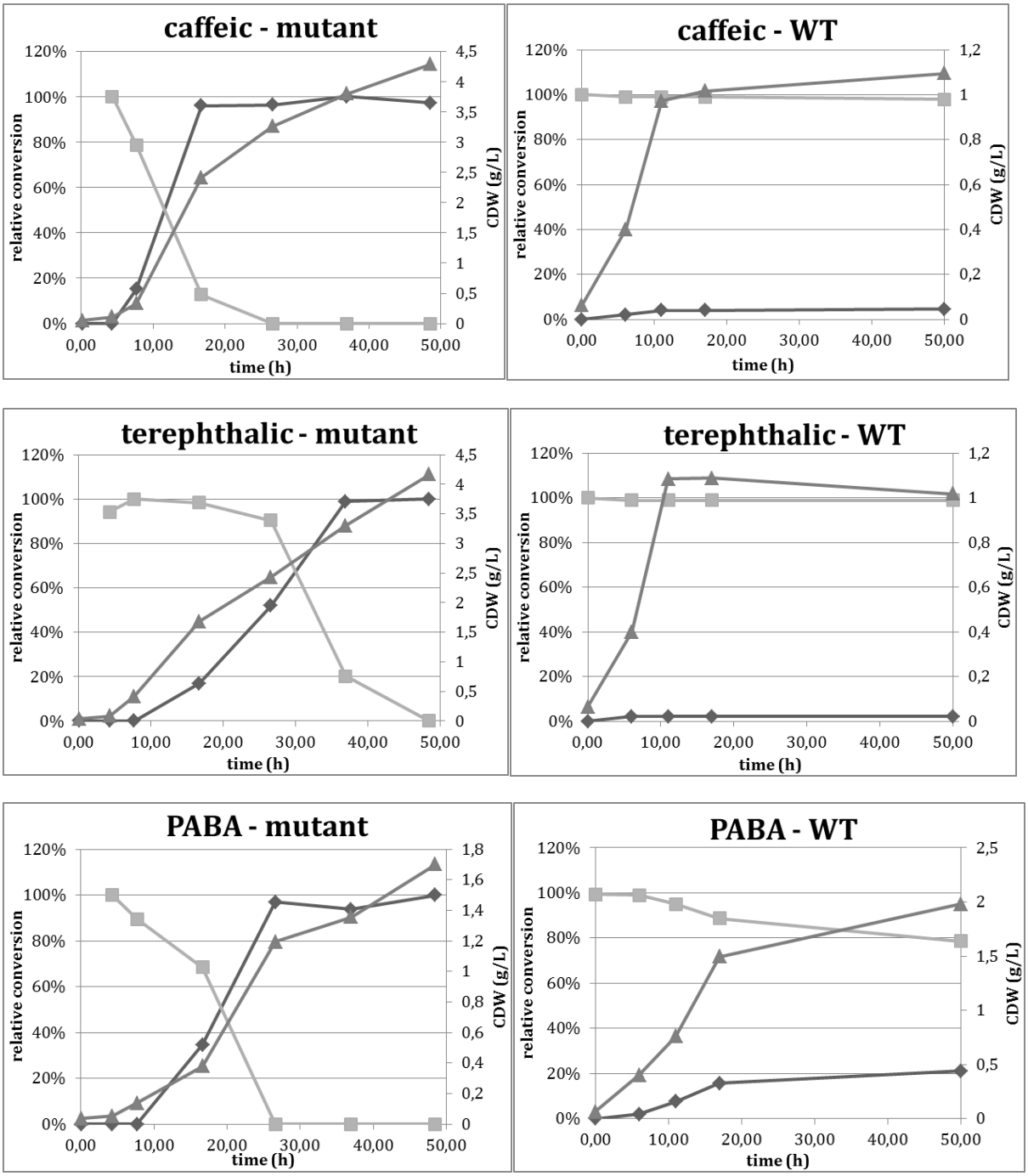
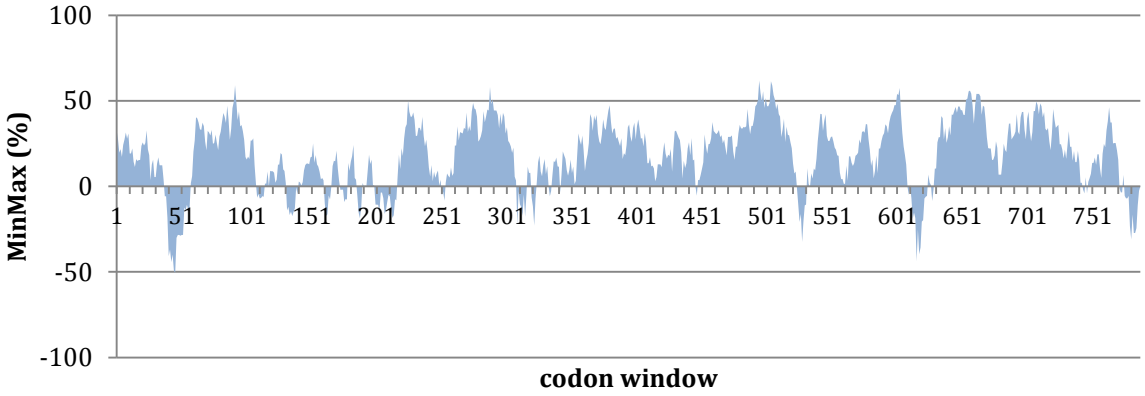


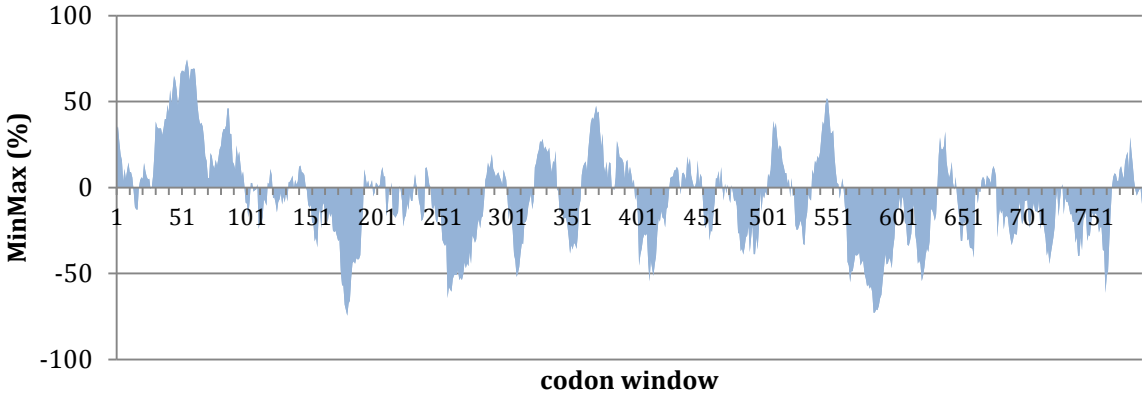
Figure A13. Overview of the growth (triangles) and conversion of a phenolic acid (squares) into the corresponding glucose ester (diamonds) over time. The engineered mutant sGLC5 + pBaSP/VvGT2/UgpA is compared with the WT + pVvGT2 for the formation of caffeic, terephthalic and *p*-aminobenzoic acid (PABA).

A14. Rare codon clusters of SuSy isoform 2 in different hosts

SuSy* native sequence – codon usage of *S. tuberosum



SuSy* native sequence – codon usage of *E. coli



SuSy* codon optimized – codon usage of *E. coli

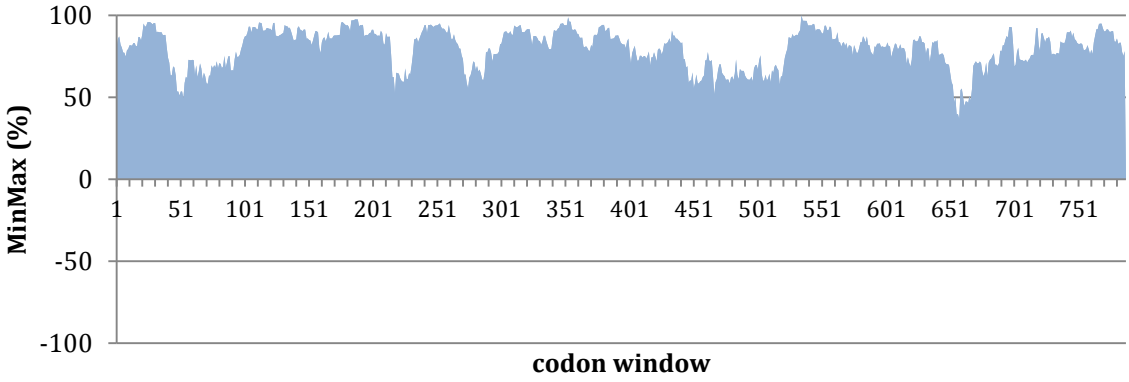


Figure A14. Frequency plots on the distribution of rare and common codons in the sequence of SuSy from *Solanum tuberosum*. A comparison is shown between the native sequence translated in *S. tuberosum* (top) and *E. coli* (middle) versus a codon optimized sequence for *E. coli* translated in *E. coli* (bottom).

Bibliography

(1) Zeikus, J. G., Jain, M. K., and Elankovan, P. (1999) Biotechnology of succinic acid production and markets for derived industrial products. *Appl Microbiol Biot 51*, 545-552.
(2) Datta, R., and Henry, M. (2006) Lactic acid: recent advances in products, processes and technologies - a review. *J Chem Technol Biot 81*, 1119-1129.
(3) Garlotta, D. (2001) A literature review of poly(lactic acid). *J Polym Environ 9*, 63-84.
(4) Hahn-Hagerdal, B., Galbe, M., Gorwa-Grauslund, M. F., Liden, G., and Zacchi, G. (2006) Bio-ethanol - the fuel of tomorrow from the residues of today. *Trends Biotechnol 24*, 549-556.
(5) Balat, M., and Balat, H. (2009) Recent trends in global production and utilization of bio-ethanol fuel. *Appl Energ 86*, 2273-2282.
(6) Mussatto, S. I., Dragone, G., Guimaraes, P. M. R., Silva, J. P. A., Carneiro, L. M., Roberto, I. C., Vicente, A., Domingues, L., and Teixeira, J. A. (2010) Technological trends, global market, and challenges of bio-ethanol production. *Biotechnol Adv 28*, 817-830.
(7) Kimura, E. (2003) Metabolic engineering of glutamate production. *Adv Biochem Eng Biotechnol*. *79*, 37-57.
(8) Leuchtenberger, W., Huthmacher, K., and Drauz, K. (2005) Biotechnological production of amino acids and derivatives: current status and prospects. *Appl Microbiol Biotechnol 69*, 1-8.
(9) Hermann, T. (2003) Industrial production of amino acids by coryneform bacteria. *J Biotechnol 104*, 155-172.
(10) Becker, J., and Wittmann, C. (2012) Systems and synthetic metabolic engineering for amino acid production - the heartbeat of industrial strain development. *Curr Opin Biotechnol 23*, 718-726.
(11) Dhillon, G. S., Brar, S. K., Verma, M., and Tyagi, R. D. (2011) Recent Advances in Citric Acid Bio-production and Recovery. *Food Bioprocess Tech 4*, 505-529.
(12) Wallace, S., and Balskus, E. P. (2014) Opportunities for merging chemical and biological synthesis. *Curr Opin Biotechnol 30C*, 1-8.
(13) Bowles, D., Isayenkova, J., Lim, E. K., and Poppenberger, B. (2005) Glycosyltransferases: managers of small molecules. *Curr Opin Plant Biol 8*, 254-263.
(14) Lee, L., Kimura, A., and Tochikura, T. (1979) Purification and properties of UDP-glucose (UDP-galactose) pyrophosphorylase from *Bifidobacterium bifidum. J Biochem*. *86*, 923-928.
(15) Durand, G., and Seta, N. (2000) Protein glycosylation and diseases: blood and urinary oligosaccharides as markers for diagnosis and therapeutic monitoring. *Clin Chem 46*, 795-805.
(16) Dennis, J. W., Granovsky, M., and Warren, C. E. (1999) Protein glycosylation in development and disease. *Bioessays 21*, 412-421.

(17) Varki, A., and Freeze, H. H. (2009) Glycans in Acquired Human Diseases, In *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., Eds.) 2nd ed., Cold Spring Harbor (NY).
(18) Schnaitman, C. A., and Klena, J. D. (1993) Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol Rev*. *57*, 655-682.
(19) Varki, A., Kannagi, R., and Toole, B. P. (2009) Glycosylation Changes in Cancer, In *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., Eds.) 2nd ed., Cold Spring Harbor (NY).
(20) Kren, V., and Martinkova, L. (2001) Glycosides in medicine: "The role of glycosidic residue in biological activity". *Curr Med Chem 8*, 1303-1328.
(21) Yamamoto, I., Muto, N., Nagata, E., Nakamura, T., and Suzuki, Y. (1990) Formation of a stable L-ascorbic acid alpha-glucoside by mammalian alpha-glucosidase-catalyzed transglucosylation. *Biochim Biophys Acta*. *1035*, 44-50.
(22) Zhao, C. L., Chen, Z. J., Bai, X. S., Ding, C., Long, T. J., Wei, F. G., and Miao, K. R. (2014) Structure-activity relationships of anthocyanidin glycosylation. *Mol Divers*.
(23) Ritter, J. K. (2000) Roles of glucuronidation and UDP-glucuronosyltransferases in xenobiotic bioactivation reactions. *Chem Biol Interact 129*, 171-193.
(24) King, C. D., Rios, G. R., Green, M. D., and Tephly, T. R. (2000) UDP-glucuronosyltransferases. *Curr Drug Metab 1*, 143-161.
(25) Geuns, J. M. (2003) Stevioside. *Phytochemistry 64*, 913-921.
(26) Wong, A., and Toth, I. (2001) Lipid, sugar and liposaccharide based delivery systems. *Curr Med Chem*. *8*, 1123-1136.
(27) Olthof, M. R., Hollman, P. C. H., Vree, T. B., and Katan, M. B. (2000) Bioavailabilities of quercetin-3-glucoside and quercetin-4 '-glucoside do not differ in humans. *J Nutr*. *130*, 1200-1203.
(28) Terao, J., Yamaguchi, S., Shirai, M., Miyoshi, M., Moon, J. H., Oshima, S., Inakuma, T., Tsushida, T., and Kato, Y. (2001) Protection by quercetin and quercetin 3-O-beta-glueuronide of peroxynitrite-induced antioxidant consumption in human plasma low-density lipoprotein. *Free Radical Res 35*, 925-931.
(29) Kim, S. J., Um, J. Y., Hong, S. H., and Lee, J. Y. (2011) Anti-Inflammatory Activity of Hyperoside Through the Suppression of Nuclear Factor-kappa B Activation in Mouse Peritoneal Macrophages. *Am J Chinese Med 39*, 171-181.
(30) Choi, H. J., Song, J. H., Park, K. S., and Kwon, D. H. (2009) Inhibitory effects of quercetin 3-rhamnoside on influenza A virus replication. *Eur J Pharm Sci 37*, 329-333.
(31) Sanchez-del-Campo, L., Saez-Ayala, M., Chazarra, S., Cabezas-Herrera, J., and Rodriguez-Lopez, J. N. (2009) Binding of natural and synthetic polyphenols to human dihydrofolate reductase. *Int J Mol Sci*. *10*, 5398-5410.
(32) Arts, I. C. W., Sesink, A. L. A., Faassen-Peters, M., and Hollman, P. C. H. (2004) The type of sugar moiety is a major determinant of the small intestinal uptake and subsequent biliary excretion of dietary quercetin glycosides. *Brit J Nutr 91*, 841-847.
(33) Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem 77*, 521-555.
(34) Breton, C., Fournel-Gigleux, S., and Palcic, M. M. (2012) Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struct Biol 22*, 540-549.

(35) Yonekura-Sakakibara, K., and Hanada, K. (2011) An evolutionary view of functional diversity in family 1 glycosyltransferases. *Plant J*. *66*, 182-193.
(36) Ross, J., Li, Y., Lim, E., and Bowles, D. J. (2001) Higher plant glycosyltransferases. Genome Biol. *2*.
(37) Bowles, D., Lim, E. K., Poppenberger, B., and Vaistij, F. E. (2006) Glycosyltransferases of lipophilic small molecules. *Annu Rev Plant Biol 57*, 567-597.
(38) Simkhada, D., Kim, E., Lee, H. C., and Sohng, J. K. (2009) Metabolic engineering of *Escherichia coli* for the biological synthesis of 7-O-xylosyl naringenin. *Mol Cells*. *28*, 397-401.
(39) Yu, L., Chen, C., Wang, L. F., Kuang, X., Liu, K., Zhang, H., and Du, J. R. (2013) Neuroprotective effect of kaempferol glycosides against brain injury and neuroinflammation by inhibiting the activation of NF-kappaB and STAT3 in transient focal stroke. *PloS one 8*, e55839.
(40) Shukla, S., Bhaskaran, N., Babcook, M. A., Fu, P., Maclennan, G. T., and Gupta, S. (2014) Apigenin inhibits prostate cancer progression in TRAMP mice via targeting PI3K/Akt/FoxO pathway. *Carcinogenesis 35*, 452-460.
(41) Puppala, M., Ponder, J., Suryanarayana, P., Reddy, G. B., Petrash, J. M., and Labarbera, D. V. (2012) The Isolation and Characterization of beta-Glucogallin as a Novel Aldose Reductase Inhibitor from *Emblica officinalis*. *PloS one 7*, e31399.
(42) Orallo, F., Alvarez, E., Camina, M., Leiro, J. M., Gomez, E., and Fernandez, P. (2002) The possible implication of trans-Resveratrol in the cardioprotective effects of long-term moderate wine consumption. *Mol Pharmacol 61*, 294-302.
(43) Choi, J., Shin, K. M., Park, H. J., Jung, H. J., Kim, H. J., Lee, Y. S., Rew, J. H., and Lee, K. T. (2004) Anti-inflammatory and antinociceptive effects of sinapyl alcohol and its glucoside syringin. *Planta Med*. *70*, 1027-1032.
(44) Miyake, Y., Mochizuki, M., Okada, M., Hiramitsu, M., Morimitsu, Y., and Osawa, T. (2007) Isolation of antioxidative phenolic glucosides from lemon juice and their suppressive effect on the expression of blood adhesion molecules. *Biosci Biotechnol Biochem 71*, 1911-1919.
(45) Zhu, J., Zeng, Z., Song, L., Hu, Y., Wen, W., and Yu, R. (2014) Stereo and regio-selective biosynthesis of two new dihydroartemisinic acid glycosides by suspension-cultured cells of *Artemisia annua*. *Pharmacogn Mag 10*, 110-114.
(46) Caputi, L., Lim, E. K., and Bowles, D. J. (2008) Discovery of new biocatalysts for the glycosylation of terpenoid scaffolds. *Chem-Eur J 14*, 6656-6662.
(47) Chen, W. N., Li, X. M., Jia, L. Q., Wang, J., Zhang, L., Hou, D. D., Wang, J. Y., and Ren, L. (2013) Neuroprotective activities of catalpol against CaMKII-dependent apoptosis induced by LPS in PC12 cells. *Brit J Pharmacol 169*, 1140-1152.
(48) Butnaru, C., Vlase, L., Lazar, D., Agoroaei, L., and Lazar, M. I. (2011) Hplc/Ms Analysis of Solanine in *Physalis Alkekengi* and *Solanum Dulcamara*. *Farmacia 59*, 172-178.
(49) Antoine, T., Bosso, C., Heyraud, A., and Samain, E. (2005) Large scale in vivo synthesis of globotriose and globotetraose by high cell density culture of metabolically engineered *Escherichia coli*. *Biochimie 87*, 197-203.
(50) Nishimoto, M., and Kitaoka, M. (2007) Practical preparation of lacto-N-biose I, a candidate for the bifidus factor in human milk. *Biosci Biotechnol Biochem 71*, 2101-2104.

(51) Dai, T. H., Tanaka, M., Huang, Y. Y., and Hamblin, M. R. (2011) Chitosan preparations for wounds and burns: antimicrobial and wound-healing effects. *Expert Rev Anti-Infe 9*, 857-879.
(52) Park, S. R., Park, J. W., Ban, Y. H., Sohng, J. K., and Yoon, Y. J. (2013) 2-Deoxystreptamine-containing aminoglycoside antibiotics: Recent advances in the characterization and manipulation of their biosynthetic pathways. *Nat Prod Rep 30*, 11-20.
(53) Muller, S. O., Eckert, I., Lutz, W. K., and Stopper, H. (1996) Genotoxicity of the laxative drug components emodin, aloe-emodin and danthron in mammalian cells: Topoisomerase II mediated? *Mutat Res-Genet Tox 371*, 165-173.
(54) Zhao, D. L., Zou, L. B., Lin, S., Shi, H. G., and Zhu, H. B. (2007) Anti-apoptotic effect of esculin on dopamine-induced cytotoxicity in the human neuroblastoma SH-SY5Y cell line. *Neuropharmacology 53*, 724-732.
(55) Van Bogaert, I. N. A., Saerens, K., De Muynck, C., Develter, D., Soetaert, W., and Vandamme, E. J. (2007) Microbial production and application of sophorolipids. *Appl Microbiol Biot*. *76*, 23-34.
(56) Barras, C. (2014) The synbio revolution: Made in microbes. *New Sci. 222*, 34-37.
(57) Curran, K. A., and Alper, H. S. (2012) Expanding the chemical palate of cells by combining systems biology and metabolic engineering. *Metab Eng 14*, 289-297.
(58) Coussement, P., Maertens, J., Beauprez, J., Van Bellegem, W., and De Mey, M. (2014) One step DNA assembly for combinatorial metabolic engineering. *Metab Eng 23*, 70-77.
(59) De Mey, M., Maertens, J., Lequeux, G. J., Soetaert, W. K., and Vandamme, E. J. (2007) Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. *BMC Biotechnol 7*.
(60) Liu, L., Redden, H., and Alper, H. S. (2013) Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Curr Opin Biotechnol 24*, 1023-1030.
(61) Putignani, L., Massa, O., and Alisi, A. (2013) Engineered *Escherichia coli* as new source of flavonoids and terpenoids. *Food Res Int 54*, 1084-1095.
(62) Lin, Y., Jain, R., and Yan, Y. (2014) Microbial production of antioxidant food ingredients via metabolic engineering. *Curr Opin Biotechnol 26*, 71-78.
(63) Renault, H., Bassard, J. E., Hamberger, B., and Werck-Reichhart, D. (2014) Cytochrome P450-mediated metabolic engineering: current progress and future challenges. *Curr Opin Plant Biol 19C*, 27-34.
(64) Lin, Y., Sun, X., Yuan, Q., and Yan, Y. (2013) Combinatorial biosynthesis of plant-specific coumarins in bacteria. *Metab Eng 18*, 69-77.
(65) Jansen, F., Gillessen, B., Mueller, F., Commandeur, U., Fischer, R., and Kreuzaler, F. (2014) Metabolic engineering for p-coumaryl alcohol production in *Escherichia coli* by introducing an artificial phenylpropanoid pathway. *Biotechnol Appl Biochem*.
(66) Schwarz, M., Hillebrand, S., Habben, S., Degenhardt, A., and Winterhalter, P. (2003) Application of high-speed countercurrent chromatography to the large-scale isolation of anthocyanins. *Biochem Eng J 14*, 179-189.
(67) Shen, R., Fuller, N., Osswald, G., Austin, W. F., Hubbs, J. L., Creaser, S. P., Findeis, M. A., Ives, J. L., and Bronk, A. (2014) Multikilogram-Scale Production of Cycloartenol Triterpenoid Glycosides as Synthetic Intermediates for a γ-Secretase Modulator. *Org Process Res Dev*.

(68) Bai, X. L., Yue, T. L., Yuan, Y. H., and Zhang, H. W. (2010) Optimization of microwave-assisted extraction of polyphenols from apple pomace using response surface methodology and HPLC analysis. *J Sep Sci 33*, 3751-3758.
(69) Philippe, R. N., De Mey, M., Anderson, J., and Ajikumar, P. K. (2014) Biotechnological production of natural zero-calorie sweeteners. *Curr Opin Biotechnol 26*, 155-161.
(70) Hussain, M. S., Fareed, S., Ansari, S., Rahman, M. A., Ahmad, I. Z., and Saeed, M. (2012) Current approaches toward production of secondary plant metabolites. J Pharm Bioallied Sci. *4*, 10-20.
(71) Rao, S. R., and Ravishankar, G. A. (2002) Plant cell cultures: Chemical factories of secondary metabolites. *Biotechnol Adv 20*, 101-153.
(72) Bertoli, A., Ruffoni, B., and Pistelli, L. (2010) Analytical methods for the extraction and identification of secondary metabolite production in 'in vitro' plant cell cultures. *Adv Exp Med Biol 698*, 250-266.
(73) Verpoorte, R., and Memelink, J. (2002) Engineering secondary metabolite production in plants. *Curr Opin Biotechnol 13*, 181-187.
(74) Demchenko, A. V. (2008) *Handbook of chemical glycosylation: advances in stereoselectivity and therapeutic relevance*, John Wiley & Sons.
(75) Seeberger, P. H., Finney, N., Rabuka, D., and Bertozzi, C. R. (2009) Chemical and Enzymatic Synthesis of Glycans and Glycoconjugates, In *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., Eds.) 2nd ed., Cold Spring Harbor (NY).
(76) Zhu, X., and Schmidt, R. R. (2009) New principles for glycoside-bond formation. *Angew Chem Int Ed Engl 48*, 1900-1934.
(77) de Roode, B. M., Franssen, M. C., van der Padt, A., and Boom, R. M. (2003) Perspectives for the industrial enzymatic production of glycosides. *Biotechnol Prog 19*, 1391-1402.
(78) Desmet, T., Soetaert, W., Bojarova, P., Kren, V., Dijkhuizen, L., Eastwick-Field, V., and Schiller, A. (2012) Enzymatic glycosylation of small molecules: challenging substrates require tailored catalysts. *Chemistry 18*, 10786-10801.
(79) Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res 42*, D490-495.
(80) De Winter, K., Simcikova, D., Schalck, B., Weignerova, L., Pelantova, H., Soetaert, W., Desmet, T., and Kren, V. (2013) Chemoenzymatic synthesis of alpha-L-rhamnosides using recombinant alpha-L-rhamnosidase from *Aspergillus terreus*. *Bioresource Technol 147*, 640-644.
(81) Ganzle, M. G. (2012) Enzymatic synthesis of galacto-oligosaccharides and other lactose derivatives (hetero-oligosaccharides) from lactose. *Int Dairy J 22*, 116-122.
(82) Maugard, T., Gaunt, D., Legoy, M. D., and Besson, T. (2003) Microwave-assisted synthesis of galacto-oligosaccharides from lactose with immobilized beta-galactosidase from *Kluyveromyces lactis*. *Biotechnol Lett 25*, 623-629.
(83) Woo, H. J., Kang, H. K., Thi, T. H. N., Kim, G. E., Kim, Y. M., Park, J. S., Kim, D., Cha, J., Moon, Y. H., Nam, S. H., Xia, Y. M., Kimura, A., and Kim, D. (2012) Synthesis and characterization of ampelopsin glucosides using dextransucrase from *Leuconostoc mesenteroides* B-1299CB4: Glucosylation enhancing physicochemical properties. *Enzyme Microb Tech*. *51*, 311-318.
(84) Cho, H. K., Kim, H. H., Seo, D. H., Jung, J. H., Park, J. H., Baek, N. I., Kim, M. J., Yoo, S. H., Cha, J., Kim, Y. R., and Park, C. S. (2011) Biosynthesis of (+)-catechin glycosides

using recombinant amylosucrase from *Deinococcus geothermalis* DSM 11300. *Enzyme Microb Tech*. *49*, 246-253.

(85) Luley-Goedl, C., and Nidetzky, B. (2010) Small-molecule glucosylation by sucrose phosphorylase: structure-activity relationships for acceptor substrates revisited. *Carbohydr Res*. *345*, 1492-1496.

(86) Desmet, T., and Soetaert, W. (2011) Enzymatic glycosyl transfer: mechanisms and applications. *Biocatal Biotransformation 29*, 1-18.

(87) Kwon, T., Kim, C. T., and Lee, J. H. (2007) Transglucosylation of ascorbic acid to ascorbic acid 2-glucoside by a recombinant sucrose phosphorylase from *Bifidobacterium longum*. *Biotechnol Lett 29*, 611-615.

(88) De Winter, K., Verlinden, K., Kren, V., Weignerova, L., Soetaert, W., and Desmet, T. (2013) Ionic liquids as cosolvents for glycosylation by sucrose phosphorylase: balancing acceptor solubility and enzyme stability. *Green Chem 15*, 1949-1955.

(89) Sugimoto, K., Nomura, K., Nishiura, H., Ohdan, K., Nishimura, T., Hayashi, H., and Kuriki, T. (2008) Sucrose phosphorylases catalyze transglycosylation reactions on carboxylic acid compounds. *Biologia 63*, 1015-1019.

(90) Kitao, S., Ariga, T., Matsudo, T., and Sekine, H. (1993) The Syntheses of Catechin-Glucosides by Transglycosylation with *Leuconostoc mesenteroides* Sucrose Phosphorylase. *Biosci Biotech Bioch*. *57*, 2010-2015.

(91) De Winter, K., Desmet, T., Devlamynck, T., Van Renterghem, L., Verhaeghe, T., Pelantova, H., Kren, V., and Soetaert, W. (2014) Biphasic Catalysis with Disaccharide Phosphorylases: Chemoenzymatic Synthesis of α-D-Glucosides Using Sucrose Phosphorylase. Org Process Res Dev.

(92) Ruffing, A., and Chen, R. R. (2006) Metabolic engineering of microbes for oligosaccharide and polysaccharide synthesis. *Microb Cell Fact 5*, 25.

(93) Koizumi, S., Endo, T., Tabata, K., and Ozaki, A. (1998) Large-scale production of UDP-galactose and globotriose by coupling metabolically engineered bacteria. *Nat Biotechnol 16*, 847-850.

(94) Endo, T., Koizumi, S., Tabata, K., Kakita, S., and Ozaki, A. (1999) Large-scale production of N-acetyllactosamine through bacterial coupling. Carbohydr Res. *316*, 179-183.

(95) Brochado, A. R., Matos, C., Moller, B. L., Hansen, J., Mortensen, U. H., and Patil, K. R. (2010) Improved vanillin production in baker's yeast through in silico design. *Microb Cell Fact 9*, 84.

(96) Hansen, E. H., Moller, B. L., Kock, G. R., Bunner, C. M., Kristensen, C., Jensen, O. R., Okkels, F. T., Olsen, C. E., Motawia, M. S., and Hansen, J. (2009) De Novo Biosynthesis of Vanillin in Fission Yeast (*Schizosaccharomyces pombe*) and Baker's Yeast (*Saccharomyces cerevisiae*). *Appl Environ Microb 75*, 2765-2774.

(97) Yang, S. M., Han, S. H., Kim, B. G., and Ahn, J. H. (2014) Production of kaempferol 3-O-rhamnoside from glucose using engineered *Escherichia coli*. *J Ind Microbiol Biotechnol*.

(98) Choi, O., Lee, J. K., Kang, S. Y., Pandey, R. P., Sohng, J. K., Ahn, J. S., and Hong, Y. S. (2014) Construction of artificial biosynthetic pathways for resveratrol glucoside derivatives. *J Microbiol Biotechnol*.

(99) Leonard, E., Yan, Y., Fowler, Z. L., Li, Z., Lim, C. G., Lim, K. H., and Koffas, M. A. (2008) Strain improvement of recombinant Escherichia coli for efficient production of plant flavonoids. *Mol Pharm 5*, 257-265.

(100) Kishore, G. M., Motion, M., Hicks, P. M., Hansen, J., Houghton-Larsen, J., Hansen, E. H., Mikkelsen, M. D., Tavares, S., and Blom, C. (2013) Production of steviol glycosides in microorganisms. Google Patents.
(101) Kim, B. G., Kim, H. J., and Ahn, J. H. (2012) Production of Bioactive Flavonol Rhamnosides by Expression of Plant Genes in *Escherichia coli*. *J Agr Food Chem*. *60*, 11143-11148.
(102) Lim, E. K., Ashford, D. A., Hou, B. K., Jackson, R. G., and Bowles, D. J. (2004) *Arabidopsis* glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. *Biotechnol Bioeng 87*, 623-631.
(103) Kim, B. G., Sung, S. H., and Ahn, J. H. (2012) Biological synthesis of quercetin 3-O-N-acetylglucosamine conjugate using engineered *Escherichia coli* expressing UGT78D2. *Appl Microbiol Biotechnol 93*, 2447-2453.
(104) Pandey, R. P., Malla, S., Simkhada, D., Kim, B. G., and Sohng, J. K. (2013) Production of 3-O-xylosyl quercetin in *Escherichia coli*. *Appl Microbiol Biotechnol 97*, 1889-1901.
(105) Han, S. H., Kim, B. G., Yoon, J. A., Chong, Y., and Ahn, J. H. (2014) Synthesis of Flavonoid O-Pentosides by *Escherichia coli* through Engineering of Nucleotide Sugar Pathways and Glycosyltransferase. *Appl Environ Microbiol 80*, 2754-2762.
(106) Kim, H. J., Kim, B. G., and Ahn, J. H. (2013) Regioselective synthesis of flavonoid bisglycosides using *Escherichia coli* harboring two glycosyltransferases. *Appl Microbiol Biotechnol 97*, 5275-5282.
(107) Thuan, N. H., Park, J. W., and Sohng, J. K. (2013) Toward the production of flavone-7-O-beta-D-glucopyranosides using *Arabidopsis* glycosyltransferase in *Escherichia*. *Process Biochem 48*, 1744-1748.
(108) Pandey, R. P., Parajuli, P., Koirala, N., Lee, J. H., Park, Y. I., and Sohng, J. K. (2014) Glucosylation of Isoflavonoids in Engineered *Escherichia coli.* Mol Cells. *37*, 172-177.
(109) Werner, S. R., and Morgan, J. A. (2010) Controlling selectivity and enhancing yield of flavonoid glycosides in recombinant yeast. *Bioprocess Biosystems Eng. 33*, 863-871.
(110) Malla, S., Pandey, R. P., Kim, B. G., and Sohng, J. K. (2013) Regiospecific modifications of naringenin for astragalin production in *Escherichia coli*. *Biotechnol Bioeng 110*, 2525-2535.
(111) Pandey, R. P., Li, T. F., Kim, E. H., Yamaguchi, T., Park, Y. I., Kim, J. S., and Sohng, J. K. (2013) Enzymatic Synthesis of Novel Phloretin Glucosides. *Appl Environ Microb*. *79*, 3516-3521.
(112) Brazier-Hicks, M., and Edwards, R. (2013) Metabolic engineering of the flavone-C-glycoside pathway using polyprotein technology. *Metab Eng 16*, 11-20.
(113) Houghton-Larsen, J., Hicks, P. M., Naesby, M., Ostergaard, T. T., Hansen, J., Dalgaard, M. M., Halkjaer, H. E., and Simon, E. (2013) Recombinant production of steviol glycosides. Google Patents.
(114) Arend, J., Warzecha, H., Hefner, T., and Stockigt, J. (2001) Utilizing genetically engineered bacteria to produce plant-specific glucosides. Biotechnol Bioeng. *76*, 126-131.
(115) Lim, E. K., Doucet, C. J., Hou, B., Jackson, R. G., Abrams, S. R., and Bowles, D. J. (2005) Resolution of (+)-abscisic acid using an *Arabidopsis* glycosyltransferase. . *Tetrahedron Asymmetry 16*, 143-147.

(116) Chen, X., Zhang, J., Kowal, P., Liu, Z., Andreana, P. R., Lu, Y., and Wang, P. G. (2001) Transferring a biosynthetic cycle into a productive *Escherichia coli* strain: large-scale synthesis of galactosides. *J Am Chem Soc 123*, 8866-8867.
(117) Ryu, S. I., and Lee, S. B. (2013) Synthesis of nucleotide sugars and alpha-galacto-oligosaccharides by recombinant *Escherichia coli* cells with trehalose substrate. *Enzyme Microb Tech*. *53*, 359-363.
(118) Shao, J., Hayashi, T., and Wang, P. G. (2003) Enhanced production of alpha-galactosyl epitopes by metabolically engineered *Pichia pastoris. Appl Environ Microb*. *69*, 5238-5242.
(119) Ruffing, A. M., and Chen, R. R. (2010) Metabolic engineering of *Agrobacterium* sp. strain ATCC 31749 for production of an alpha-Gal epitope. *Microb Cell Fact 9*, 1.
(120) Chen, X., Liu, Z., Zhang, J., Zhang, W., Kowal, P., and Wang, P. G. (2002) Reassembled biosynthetic pathway for large-scale carbohydrate synthesis: alpha-Gal epitope producing "superbug". *ChemBioChem 3*, 47-53.
(121) Ruffing, A., Mao, Z., and Chen, R. R. Z. (2006) Metabolic engineering of *Agrobacterium* sp for UDP-galactose regeneration and oligosaccharide synthesis. *Metab Eng*. *8*, 465-473.
(122) Mao, Z., Shin, H. D., and Chen, R. R. (2006) Engineering the E. coli UDP-glucose synthesis pathway for oligosaccharide synthesis. *Biotechnol Prog 22*, 369-374.
(123) Liu, Z., Lu, Y., Zhang, J., Pardee, K., and Wang, P. G. (2003) P1 Trisaccharide (Galalpha1,4Galbeta1,4GlcNAc) synthesis by enzyme glycosylation reactions using recombinant *Escherichia coli*. *Appl Environ Microbiol 69*, 2110-2115.
(124) Masada, S., Kawase, Y., Nagatoshi, M., Oguchi, Y., Terasaka, K., and Mizukami, H. (2007) An efficient chemoenzymatic production of small molecule glucosides with in situ UDP-glucose recycling. *FEBS Lett 581*, 2562-2566.
(125) Romer, U., Schrader, H., Gunther, N., Nettelstroth, N., Frommer, W. B., and Elling, L. (2004) Expression, purification and characterization of recombinant sucrose synthase 1 from *Solanum tuberosum* L. for carbohydrate engineering. *J Biotechnol 107*, 135-149.
(126) Ryu, S. I., Kim, J. E., Kim, E. J., Chung, S. K., and Lee, S. B. (2011) Catalytic reversibility of Pyrococcus horikoshii trehalose synthase: Efficient synthesis of several nucleoside diphosphate glucoses with enzyme recycling. *Process Biochem. 46*, 128-134.
(127) Qu, Q. H., Lee, S. J., and Boos, W. (2004) TreT, a novel trehalose glycosyltransferring synthase of the hyperthermophilic Archaeon *Thermococcus litoralis. J Biol Chem*. *279*, 47890-47897.
(128) Field, R. A. (2011) Glycobiology: Challenging reaction equilibria. *Nat Chem Biol 7*, 658-659.
(129) Elling, L., Grothus, M., Zervosen, A., and Kula, M. R. (1995) Application of sucrose synthase from rice grains for the synthesis of carbohydrates. *Ann N Y Acad Sci 750*, 329-331.
(130) Bungaruang, L., Gutmann, A., and Nidetzky, B. (2013) Leloir Glycosyltransferases and Natural Product Glycosylation: Biocatalytic Synthesis of the -Glucoside Nothofagin, a Major Antioxidant of Redbush Herbal Tea. *Adv Synth Catal 355*, 2757-2763.
(131) Brinkmann, N., Malissard, M., Ramuz, M., Romer, U., Schumacher, T., Berger, E. G., Elling, L., Wandrey, C., and Liese, A. (2001) Chemo-enzymatic synthesis of the Galili epitope Gal alpha(1 -> 3)Gal beta(1 -> 4)GlcNAc on a homogeneously

soluble PEG polymer by a multi-enzyme system. *Bioorg Med Chem Lett 11*, 2503-2506.
(132) Terasaka, K., Mizutani, Y., Nagatsu, A., and Mizukami, H. (2012) In situ UDP-glucose regeneration unravels diverse functions of plant secondary product glycosyltransferases. *FEBS Lett 586*, 4344-4350.
(133) Son, M. H., Kim, B. G., Kim, D. H., Jin, M., Kim, K., and Ahn, J. H. (2009) Production of Flavonoid O-Glucoside Using Sucrose Synthase and Flavonoid O-Glucosyltransferase Fusion Protein. *J Microbiol Biotechn 19*, 709-712.
(134) Nakai, T., Konishi, T., Zhang, X. Q., Chollet, R., Tonouchi, N., Tsuchida, T., Yoshinaga, F., Mori, H., Sakai, F., and Hayashi, T. (1998) An increase in apparent affinity for sucrose of mung bean sucrose synthase is caused by in vitro phosphorylation or directed mutagenesis of Ser11. *Plant Cell Physiol 39*, 1337-1341.
(135) Curatti, L., Porchia, A. C., Herrera-Estrella, L., and Salerno, G. L. (2000) A prokaryotic sucrose synthase gene (susA) isolated from a filamentous nitrogen-fixing cyanobacterium encodes a protein similar to those of plants. *Planta 211*, 729-735.
(136) Desmet, T., and Soetaert, W. (2012) Broadening the synthetic potential of disaccharide phosphorylases through enzyme engineering. *Process Biochem. 47*, 11-17.
(137) Sprogoe, D., van den Broek, L. A., Mirza, O., Kastrup, J. S., Voragen, A. G., Gajhede, M., and Skov, L. K. (2004) Crystal structure of sucrose phosphorylase from *Bifidobacterium adolescentis*. *Biochemistry 43*, 1156-1162.
(138) de Groeve, M. R. M., Desmet, T., and Soetaert, W. (2011) Engineering of cellobiose phosphorylase for glycoside synthesis. *J Biotechnol*. *156*, 253-260.
(139) Nahalka, J. (2008) Physiological aggregation of maltodextrin phosphorylase from *Pyrococcus furiosus* and its application in a process of batch starch degradation to alpha-D-glucose-1-phosphate. *J Ind Microbiol Biot 35*, 219-223.
(140) Mizanur, R. M., Griffin, A. K., and Pohl, N. L. (2008) Recombinant production and biochemical characterization of a hyperthermostable alpha-glucan/maltodextrin phosphorylase from *Pyrococcus furiosus*. *Archaea 2*, 169-176.
(141) De Groeve, M. R., De Baere, M., Hoflack, L., Desmet, T., Vandamme, E. J., and Soetaert, W. (2009) Creating lactose phosphorylase enzymes by directed evolution of cellobiose phosphorylase. *Protein Eng Des Sel*. *22*, 393-399.
(142) Ichikawa, M., Schnaar, R. L., and Ichikawa, Y. (1995) Application of Sucrose Phosphorylase Reaction in One-Pot Enzymatic Galactosylation - Scavenger of Phosphate and Generation of Glucose 1-Phosphate in-Situ. *Tetrahedron Lett. 36*, 8731-8732.
(143) Kim, M., Kwon, T., Lee, H. J., Kim, K. H., Chung, D. K., Ji, G. E., Byeon, E. S., and Lee, J. H. (2003) Cloning and expression of sucrose phosphorylase gene from *Bifidobacterium longum* in *E. coli* and characterization of the recombinant enzyme. *Biotechnol Lett 25*, 1211-1217.
(144) Aerts, D., Verhaeghe, T. F., Roman, B. I., Stevens, C. V., Desmet, T., and Soetaert, W. (2011) Transglucosylation potential of six sucrose phosphorylases toward different classes of acceptors. Carbohydr Res. *346*, 1860-1867.
(145) Krauser, S., Weyler, C., Blass, L. K., and Heinzle, E. (2013) Directed multistep biocatalysis using tailored permeabilized cells. Adv Biochem Eng Biotechnol. *137*, 185-234.
(146) Muthana, M. M., Qu, J., Li, Y., Zhang, L., Yu, H., Ding, L., Malekan, H., and Chen, X. (2012) Efficient one-pot multienzyme synthesis of UDP-sugars using a

promiscuous UDP-sugar pyrophosphorylase from *Bifidobacterium longum* (BLUSP). *Chem Commun 48*, 2728-2730.
(147) Zou, Y., Wang, W., Cai, L., Chen, L., Xue, M., Zhang, X., Shen, J., and Chen, M. (2012) Substrate specificity of galactokinase from *Streptococcus pneumoniae* TIGR4 towards galactose, glucose, and their derivatives. *Bioorg Med Chem Lett 22*, 3540-3543.
(148) Hoffmeister, D., Yang, J., Liu, L., and Thorson, J. S. (2003) Creation of the first anomeric D/L-sugar kinase by means of directed evolution. *Proc Natl Acad Sci U S A 100*, 13184-13189.
(149) Hoffmeister, D., and Thorson, J. S. (2004) Mechanistic implications of *Escherichia coli* galactokinase structure-based engineering. *ChemBioChem 5*, 989-992.
(150) Williams, G. J., Yang, J., Zhang, C., and Thorson, J. S. (2011) Recombinant *E. coli* prototype strains for in vivo glycorandomization. *Acs Chem Biol 6*, 95-100.
(151) Nishimoto, M., and Kitaoka, M. (2007) Identification of N-acetylhexosamine 1-kinase in the complete lacto-N-biose I/galacto-N-biose metabolic pathway in *Bifidobacterium longum. Appl Environ Microbiol*. *73*, 6444-6449.
(152) Cai, L., Guan, W., Kitaoka, M., Shen, J., Xia, C., Chen, W., and Wang, P. G. (2009) A chemoenzymatic route to N-acetylglucosamine-1-phosphate analogues: substrate specificity investigations of N-acetylhexosamine 1-kinase. *Chem Commun (Camb)*, 2944-2946.
(153) Cai, L., Guan, W., Wang, W., Zhao, W., Kitaoka, M., Shen, J., O'Neil, C., and Wang, P. G. (2009) Substrate specificity of N-acetylhexosamine kinase towards N-acetylgalactosamine derivatives. *Bioorg Med Chem Lett 19*, 5433-5435.
(154) Li, Y., Yu, H., Chen, Y., Lau, K., Cai, L., Cao, H., Tiwari, V. K., Qu, J., Thon, V., Wang, P. G., and Chen, X. (2011) Substrate promiscuity of N-acetylhexosamine 1-kinases. *Molecules 16*, 6396-6407.
(155) Bar-Peled, M., and O'Neill, M. A. (2011) Plant nucleotide sugar formation, interconversion, and salvage by sugar recycling. *Annu Rev Plant Biol 62*, 127-155.
(156) De Bruyn, F., Beauprez, J., Maertens, J., Soetaert, W., and De Mey, M. (2013) Unraveling the Leloir Pathway of *Bifidobacterium bifidum*: Significance of the Uridylyltransferases. Appl Environ Microb. *79*, 7028-7035.
(157) Guan, W., Cai, L., and Wang, P. G. (2010) Highly efficient synthesis of UDP-GalNAc/GlcNAc analogues with promiscuous recombinant human UDP-GalNAc pyrophosphorylase AGX1. *Chemistry 16*, 13343-13345.
(158) Kotake, T., Yamaguchi, D., Ohzono, H., Hojo, S., Kaneko, S., Ishida, H. K., and Tsumuraya, Y. (2004) UDP-sugar pyrophosphorylase with broad substrate specificity toward various monosaccharide 1-phosphates from pea sprouts. *J Biol Chem*. *279*, 45728-45736.
(159) Mizanur, R. M., Zea, C. J., and Pohl, N. L. (2004) Unusually broad substrate tolerance of a heat-stable archaeal sugar nucleotidyltransferase for the synthesis of sugar nucleotides. *J Am Chem Soc 126*, 15993-15998.
(160) Zhao, X., Zou, M., Ma, Z., Wang, S., Wang, P. G., and Chen, M. (2014) A one-pot approach to bio-synthesize globotriose and its derivatives from simpler substrates. *Eur J Med Chem*.
(161) Chien, W. T., Liang, C. F., Yu, C. C., Lin, C. H., Li, S. P., Primadona, I., Chen, Y. J., Mong, K. K., and Lin, C. C. (2014) Sequential one-pot enzymatic synthesis of oligo-N-acetyllactosamine and its multi-sialylated extensions. *Chem Commun (Camb)*.
(162) Malekan, H., Fung, G., Thon, V., Khedri, Z., Yu, H., Qu, J., Li, Y., Ding, L., Lam, K. S., and Chen, X. (2013) One-pot multi-enzyme (OPME) chemoenzymatic synthesis of

sialyl-Tn-MUC1 and sialyl-T-MUC1 glycopeptides containing natural or non-natural sialic acid. *Bioorg Med Chem*. *21*, 4778-4785.

(163) Chen, X., Fang, J. W., Zhang, J. B., Liu, Z. Y., Shao, J., Kowal, P., Andreana, P., and Wang, P. G. (2001) Sugar nucleotide regeneration beads (superbeads): A versatile tool for the practical synthesis of oligosaccharides. *J Am Chem Soc 123*, 2081-2082.

(164) Tsai, T. I., Lee, H. Y., Chang, S. H., Wang, C. H., Tu, Y. C., Lin, Y. C., Hwang, D. R., Wu, C. Y., and Wong, C. H. (2013) Effective sugar nucleotide regeneration for the large-scale enzymatic synthesis of Globo H and SSEA4. *J Am Chem Soc 135*, 14831-14839.

(165) Su, D. M., Eguchi, H., Yi, W., Li, L., Wang, P. G., and Xia, C. (2008) Enzymatic synthesis of tumor-associated carbohydrate antigen Globo-H hexasaccharide. *Org Lett 10*, 1009-1012.

(166) Wilson, D. B., and Hogness, D. S. (1964) Enzymes of Galactose Operon in *Escherichia coli* .I. Purification + Characterization of Uridine Diphosphogalactose 4-Epimerase. *J Biol Chem*. *239*, 2469-&.

(167) Chung, S. K., Ryu, S. I., and Lee, S. B. (2012) Characterization of UDP-glucose 4-epimerase from *Pyrococcus horikoshii*: Regeneration of UDP to produce UDP-galactose using two-enzyme system with trehalose. *Bioresource Technol*. *110*, 423-429.

(168) Kim, B. G., Jung, W. D., and Ahn, J. H. (2013) Cloning and characterization of a putative UDP-rhamnose synthase 1 from *Populus euramericana* Guinier. *J Plant Biol 56*, 7-12.

(169) Oka, T., Nemoto, T., and Jigami, Y. (2007) Functional analysis of *Arabidopsis thaliana* RHM2/MUM4, a multidomain protein involved in UDP-D-glucose to UDP-L-rhamnose conversion. *J Biol Chem*. *282*, 5389-5403.

(170) Simkhada, D., Lee, H. C., and Sohng, J. K. (2010) Genetic Engineering Approach for the Production of Rhamnosyl and Allosyl Flavonoids From *Escherichia coli. Biotechnol Bioeng*. *107*, 154-162.

(171) Grangeasse, C., Obadia, B., Mijakovic, I., Deutscher, J., Cozzone, A. J., and Doublet, P. (2003) Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J Biol Chem*. *278*, 39323-39329.

(172) Oka, T., and Jigami, Y. (2006) Reconstruction of de novo pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in *Saccharomyces cerevisiae*. *FEBS J*. *273*, 2645-2657.

(173) Dragan, C. A., Buchheit, D., Bischoff, D., Ebner, T., and Bureik, M. (2010) Glucuronide production by whole-cell biotransformation using genetically engineered fission yeast *Schizosaccharomyces pombe*. Drug Metab Dispos. *38*, 509-515.

(174) Esko, J. D., Kimata, K., and Lindahl, U. (2009) Proteoglycans and Sulfated Glycosaminoglycans, In *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., Eds.) 2nd ed., Cold Spring Harbor (NY).

(175) Choi, S. H., Mansoorabadi, S. O., Liu, Y. N., Chien, T. C., and Liu, H. W. (2012) Analysis of UDP-D-apiose/UDP-D-xylose synthase-catalyzed conversion of UDP-D-apiose phosphonate to UDP-D-xylose phosphonate: implications for a retroaldol-aldol mechanism. *J Am Chem Soc 134*, 13946-13949.

(176) Guyett, P., Glushka, J., Gu, X., and Bar-Peled, M. (2009) Real-time NMR monitoring of intermediates and labile products of the bifunctional enzyme UDP-apiose/UDP-xylose synthase. Carbohydr Res. *344*, 1072-1078.
(177) Gu, X., Lee, S. G., and Bar-Peled, M. (2011) Biosynthesis of UDP-xylose and UDP-arabinose in *Sinorhizobium meliloti* 1021: first characterization of a bacterial UDP-xylose synthase, and UDP-xylose 4-epimerase. *Microbiology 157*, 260-269.
(178) Coyne, M. J., Fletcher, C. M., Reinap, B., and Comstock, L. E. (2011) UDP-glucuronic acid decarboxylases of Bacteroides fragilis and their prevalence in bacteria. *J Bacteriol 193*, 5252-5259.
(179) Frirdich, E., and Whitfield, C. (2005) Characterization of Gla(KP), a UDP-galacturonic acid C4-epimerase from *Klebsiella pneumoniae* with extended substrate specificity. *J Bacteriol 187*, 4104-4115.
(180) Munoz, R., Lopez, R., de Frutos, M., and Garcia, E. (1999) First molecular characterization of a uridine diphosphate galacturonate 4-epimerase: an enzyme required for capsular biosynthesis in Streptococcus pneumoniae type 1. *Mol Microbiol 31*, 703-713.
(181) Shao, J., Zhang, J., Kowal, P., and Wang, P. G. (2002) Donor substrate regeneration for efficient synthesis of globotetraose and isoglobotetraose. *Appl Environ Microbiol 68*, 5634-5640.
(182) Zhao, X., Creuzenet, C., Belanger, M., Egbosimba, E., Li, J., and Lam, J. S. (2000) WbpO, a UDP-N-acetyl-D-galactosamine dehydrogenase from *pseudomonas aeruginosa* serotype O6. *J Biol Chem*. *275*, 39802.
(183) Miller, W. L., Wenzel, C. Q., Daniels, C., Larocque, S., Brisson, J. R., and Lam, J. S. (2004) Biochemical characterization of WbpA, a UDP-N-acetyl-D-glucosamine 6-dehydrogenase involved in O-antigen biosynthesis in *Pseudomonas aeruginosa* PAO1. *J Biol Chem*. *279*, 37551-37558.
(184) Fialova, P., Namdjou, D. J., Ettrich, R., Prikrylova, W., Rauvolfova, J., Krenek, K., Kuzma, M., Elling, L., Bezouska, K., and Kren, V. (2005) Combined application of galactose oxidase and beta-N-acetylhexosaminidase in the synthesis of complex immunoactive N-acetyl-D-galactosaminides. *Adv Synth Catal 347*, 997-1006.
(185) Chen, X. (2011) Fermenting Next Generation Glycosylated Therapeutics. *Acs Chem Biol 6*, 14-17.
(186) Griffith, B. R., Langenhan, J. M., and Thorson, J. S. (2005) 'Sweetening' natural products via glycorandomization. *Curr Opin Biotechnol 16*, 622-630.
(187) Thibodeaux, C. J., Melancon, C. E., and Liu, H. W. (2007) Unusual sugar biosynthesis and natural product glycodiversification. *Nature 446*, 1008-1016.
(188) Morrison, J. P., Schoenhofen, I. C., and Tanner, M. E. (2008) Mechanistic studies on PseB of pseudaminic acid biosynthesis: a UDP-N-acetylglucosamine 5-inverting 4,6-dehydratase. *Bioorg Chem 36*, 312-320.
(189) Holden, H. M., Cook, P. D., and Thoden, J. B. (2010) Biosynthetic enzymes of unusual microbial sugars. *Curr Opin Struct Biol 20*, 543-550.
(190) Singh, S., Phillips, G. N., Jr., and Thorson, J. S. (2012) The structural biology of enzymes involved in natural product glycosylation. *Nat Prod Rep 29*, 1201-1237.
(191) Takahashi, H., Liu, Y. N., and Liu, H. W. (2006) A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate. *J Am Chem Soc 128*, 1432-1433.
(192) Zhang, H., White-Phillip, J. A., Melancon, C. E., 3rd, Kwon, H. J., Yu, W. L., and Liu, H. W. (2007) Elucidation of the kijanimicin gene cluster: insights into the

biosynthesis of spirotetronate antibiotics and nitrosugars. *J Am Chem Soc 129*, 14670-14683.
(193) White-Phillip, J., Thibodeaux, C. J., and Liu, H. W. (2009) Enzymatic synthesis of TDP-deoxysugars. *Methods Enzymol 459*, 521-544.
(194) Salas, J. A., and Mendez, C. (2007) Engineering the glycosylation of natural products in actinomycetes. *Trends Microbiol 15*, 219-232.
(195) Kim, B. G., Jung, N. R., Joe, E. J., Hur, H. G., Lim, Y., Chong, Y., and Ahn, J. H. (2010) Bacterial synthesis of a flavonoid deoxyaminosugar conjugate in *Escherichia coli* expressing a glycosyltransferase of *Arabidopsis thaliana*. *ChemBioChem 11*, 2389-2392.
(196) Yang, T., Bar-Peled, Y., Smith, J. A., Glushka, J., and Bar-Peled, M. (2012) In-microbe formation of nucleotide sugars in engineered *Escherichia coli*. *Anal Biochem 421*, 691-698.
(197) Han, A. R., Park, J. W., Lee, M. K., Ban, Y. H., Yoo, Y. J., Kim, E. J., Kim, E., Kim, B. G., Sohng, J. K., and Yoon, Y. J. (2011) Development of a *Streptomyces venezuelae*-based combinatorial biosynthetic system for the production of glycosylated derivatives of doxorubicin and its biosynthetic intermediates. *Appl Environ Microbiol 77*, 4912-4923.
(198) Yuan, T., Xie, L., Zhu, B., and Hu, Y. (2014) Bioconversion of deoxysugar moieties to the biosynthetic intermediates of daunorubicin in an engineered strain of *Streptomyces coeruleobidus*. *Biotechnol Lett*.
(199) Jakeman, D. L., Young, J. L., Huestis, M. P., Peltier, P., Daniellou, R., Nugier-Chauvin, C., and Ferrieres, V. (2008) Engineering ribonucleoside triphosphate specificity in a thymidylyltransferase. *Biochemistry 47*, 8719-8725.
(200) Simon, C., and Daniel, R. (2009) Achievements and new knowledge unraveled by metagenomic approaches. *Appl Microbiol Biotechnol 85*, 265-276.
(201) Northrup, A. B., and MacMillan, D. W. (2004) Two-step synthesis of carbohydrates by selective aldol reactions. *Science 305*, 1752-1755.
(202) Mlynarski, J., and Gut, B. (2012) Organocatalytic synthesis of carbohydrates. Chem Soc Rev. *41*, 587-596.
(203) Andreana, P. R., McLellan, J. S., Chen, Y., and Wang, P. G. (2002) Synthesis of 2,6-dideoxysugars via ring-closing olefinic metathesis. *Org Lett 4*, 3875-3878.
(204) Lau, K., Thon, V., Yu, H., Ding, L., Chen, Y., Muthana, M. M., Wong, D., Huang, R., and Chen, X. (2010) Highly efficient chemoenzymatic synthesis of beta1-4-linked galactosides with promiscuous bacterial beta1-4-galactosyltransferases. *Chem Commun (Camb) 46*, 6066-6068.
(205) Yang, J., Fu, X., Liao, J. C., Liu, L., and Thorson, J. S. (2005) Structure-based engineering of *E. coli* galactokinase as a first step toward in vivo glycorandomization. *Chem Biol 12*, 657-664.
(206) Thorson, J. S., Barton, W. A., Hoffmeister, D., Albermann, C., and Nikolov, D. B. (2004) Structure-based enzyme engineering and its impact on in vitro glycorandomization. *ChemBioChem 5*, 16-25.
(207) Moretti, R., Chang, A., Peltier-Pain, P., Bingman, C. A., Phillips, G. N., Jr., and Thorson, J. S. (2011) Expanding the nucleotide and sugar 1-phosphate promiscuity of nucleotidyltransferase RmlA via directed evolution. *J Biol Chem*. *286*, 13235-13243.
(208) Barton, W. A., Biggins, J. B., Jiang, J. Q., Thorson, J. S., and Nikolov, D. B. (2002) Expanding pyrimidine diphosphosugar libraries via structure-based nucleotidylyltransferase engineering. *P Natl Acad Sci USA 99*, 13397-13402.

(209) Pandey, R. P., Gurung, R. B., Parajuli, P., Koirala, N., Tuoi le, T., and Sohng, J. K. (2014) Assessing acceptor substrate promiscuity of YjiC-mediated glycosylation toward flavonoids. *Carbohydr Res*. *393*, 26-31.
(210) Gantt, R. W., Goff, R. D., Williams, G. J., and Thorson, J. S. (2008) Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew Chem Int Ed Engl 47*, 8889-8892.
(211) Williams, G. J., Goff, R. D., Zhang, C., and Thorson, J. S. (2008) Optimizing glycosyltransferase specificity via "hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem Biol 15*, 393-401.
(212) Wang, L., Han, W., Xie, C., Hou, J., Fang, Q., Gu, J., Wang, P. G., and Cheng, J. (2013) Comparing the acceptor promiscuity of a *Rosa hybrida* glucosyltransferase RhGT1 and an engineered microbial glucosyltransferase OleD(PSA) toward a small flavonoid library. *Carbohydr Res*. *368*, 73-77.
(213) Zhou, M., Hamza, A., Zhan, C. G., and Thorson, J. S. (2013) Assessing the regioselectivity of OleD-catalyzed glycosylation with a diverse set of acceptors. *J Nat Prod 76*, 279-286.
(214) Zhou, M., Hou, Y., Hamza, A., Pain, C., Zhan, C. G., Bugni, T. S., and Thorson, J. S. (2012) Probing the regiospecificity of enzyme-catalyzed steroid glycosylation. *Org Lett 14*, 5424-5427.
(215) Gantt, R. W., Peltier-Pain, P., Cournoyer, W. J., and Thorson, J. S. (2011) Using simple donors to drive the equilibria of glycosyltransferase-catalyzed reactions. *Nat Chem Biol 7*, 685-691.
(216) Gutmann, A., Krump, C., Bungaruang, L., and Nidetzky, B. (2014) A two-step O- to C-glycosidic bond rearrangement using complementary glycosyltransferase activities. *Chem Commun (Camb) 50*, 5465-5468.
(217) Andre, I., Potocki-Veronese, G., Barbe, S., Moulis, C., and Remaud-Simeon, M. (2014) CAZyme discovery and design for sweet dreams. *Curr Opin Chem Biol 19*, 17-24.
(218) Rabausch, U., Juergensen, J., Ilmberger, N., Bohnke, S., Fischer, S., Schubach, B., Schulte, M., and Streit, W. R. (2013) Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. *Appl Environ Microbiol 79*, 4551-4563.
(219) Williams, G. J., Gantt, R. W., and Thorson, J. S. (2008) The impact of enzyme engineering upon natural product glycodiversification. *Curr Opin Chem Biol 12*, 556-564.
(220) Truman, A. W., Dias, M. V. B., Wu, S., Blundell, T. L., Huang, F. L., and Spencer, J. B. (2009) Chimeric Glycosyltransferases for the Generation of Hybrid Glycopeptides. *Chem Biol*. *16*, 676-685.
(221) Biggs, B. W., De Paepe, B., Santos, C. N., De Mey, M., and Kumaran Ajikumar, P. (2014) Multivariate modular metabolic engineering for pathway and strain optimization. *Curr Opin Biotechnol 29C*, 156-162.
(222) Tissier, H. (1900) Ph.D. thesis. University of Paris, Paris, France. Recherches sur la flore intestinale des nourrissons (état normal et pathologique).
(223) Guarner, F., and Malagelada, J. R. (2003) Gut flora in health and disease. *Lancet 361*, 512-519.
(224) Turroni, F., Foroni, E., Pizzetti, P., Giubellini, V., Ribbera, A., Merusi, P., Cagnasso, P., Bizzarri, B., de'Angelis, G. L., Shanahan, F., van Sinderen, D., and Ventura, M. (2009) Exploring the diversity of the bifidobacterial population in the human intestinal tract. *Appl Environ Microbiol*. *75*, 1534-1545.

(225) Fushinobu, S. (2010) Unique sugar metabolic pathways of bifidobacteria. *Biosci Biotechnol Biochem 74*, 2374-2384.
(226) Ruas-Madiedo, P., Gueimonde, M., Fernandez-Garcia, M., de los Reyes-Gavilan, C. G., and Margolles, A. (2008) Mucin degradation by *Bifidobacterium* strains isolated from the human intestinal microbiota. *Appl Env Microbiol*. *74*, 1936-1940.
(227) Asakuma, S., Hatakeyama, E., Urashima, T., Yoshida, E., Katayama, T., Yamamoto, K., Kumagai, H., Ashida, H., Hirose, J., and Kitaoka, M. (2011) Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. *J Biol Chem*. *286*, 34583-34592.
(228) Kitaoka, M., Tian, J., and Nishimoto, M. (2005) Novel putative galactose operon involving lacto-N-biose phosphorylase in *Bifidobacterium longum. Appl Env Microbiol*. *71*, 3158-3162.
(229) Turroni, F., Bottacini, F., Foroni, E., Mulder, I., Kim, J. H., Zomer, A., Sanchez, B., Bidossi, A., Ferrarini, A., Giubellini, V., Delledonne, M., Henrissat, B., Coutinho, P., Oggioni, M., Fitzgerald, G. F., Mills, D., Margolles, A., Kelly, D., van Sinderen, D., and Ventura, M. (2010) Genome analysis of *Bifidobacterium bifidum* PRL2010 reveals metabolic pathways for host-derived glycan foraging. *Proc Natl Acad Sci U.S.A. 107*, 19514-19519.
(230) Zhurina, D., Zomer, A., Gleinser, M., Brancaccio, V. F., Auchter, M., Waidmann, M. S., Westermann, C., van Sinderen, D., and Riedel, C. U. (2011) Complete genome sequence of *Bifidobacterium bifidum* S17. *J bacteriol 193*, 301-302.
(231) Yu, D. S., Jeong, H., Lee, D. H., Kwon, S. K., Song, J. Y., Kim, B. K., Park, M. S., Ji, G. E., Oh, T. K., and Kim, J. F. (2012) Complete genome sequence of the probiotic bacterium *Bifidobacterium bifidum* strain BGN4. *J bacteriol 194*, 4757-4758.
(232) Xiao, J. Z., Takahashi, S., Nishimoto, M., Odamaki, T., Yaeshima, T., Iwatsuki, K., and Kitaoka, M. (2010) Distribution of in vitro fermentation ability of lacto-N-biose I, a major building block of human milk oligosaccharides, in bifidobacterial strains. *Appl Environ Microbiol*. *76*, 54-59.
(233) Hidaka, M., Nishimoto, M., Kitaoka, M., Wakagi, T., Shoun, H., and Fushinobu, S. (2009) The crystal structure of galacto-N-biose/lacto-N-biose I phosphorylase: a large deformation of a TIM barrel scaffold. *J Biol Chem*. *284*, 7273-7283.
(234) Kitaoka, M. (2012) Bifidobacterial enzymes involved in the metabolism of human milk oligosaccharides. *Adv. Nutr. 3*, 422S-429S.
(235) Frey, P. A. (1996) The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J. 10*, 461-470.
(236) Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U.S.A. 97*, 6640-6645.
(237) Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, and Smith, H. O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods 6*, 343-345.
(238) Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. *72*, 248-254.
(239) Wu, Z. L., Ethen, C. M., Prather, B., Machacek, M., and Jiang, W. (2011) Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology 21*, 727-733.

(240) Weissborn, A. C., Liu, Q., Rumley, M. K., and Kennedy, E. P. (1994) UTP: alpha-D-glucose-1-phosphate uridylyltransferase of *Escherichia coli*: isolation and DNA sequence of the *galU* gene and purification of the enzyme. *J bacteriol 176*, 2611-2618.

(241) Geeganage, S., and Frey, P. A. (1999) Significance of metal ions in galactose-1-phosphate uridylyltransferase: an essential structural zinc and a nonessential structural iron. *Biochemistry 38*, 13398-13406.

(242) Arabshahi, A., Ruzicka, F. J., Geeganage, S., and Frey, P. A. (1996) Standard free energies for uridylyl group transfer by hexose-1-P uridylyltransferase and UDP-hexose synthase and for the hydrolysis of uridine 5'-phosphoimidazolate. *Biochemistry 35*, 3426-3428.

(243) Lee, L., Kimura, A., and Tochikura, T. (1978) Purification and Properties of Hexose 1-Phosphate Uridylyltransferase from *Bifidobacterium bifidum*. *Agric Biol Chem 42, 723-730*.

(244) Lee, L. J., Kimura, A., and Tochikura, T. (1978) Presence of a single enzyme catalyzing the pyrophosphorolysis of UDP-glucose and UDP-galactose in *Bifidobacterium bifidum. Biochim Biophys Acta*. *527*, 301-304.

(245) Thoden, J. B., and Holden, H. M. (2007) The molecular architecture of glucose-1-phosphate uridylyltransferase. *Protein Sci. 16*, 432-440.

(246) Kleczkowski, L. A., Decker, D., and Wilczynska, M. (2011) UDP-sugar pyrophosphorylase: a new old mechanism for sugar activation. *Plant Physiol. 156*, 3-10.

(247) Dai, N., Petreikov, M., Portnoy, V., Katzir, N., Pharr, D. M., and Schaffer, A. A. (2006) Cloning and expression analysis of a UDP-galactose/glucose pyrophosphorylase from melon fruit provides evidence for the major metabolic pathway of galactose metabolism in raffinose oligosaccharide metabolizing plants. *Plant Physiol. 142*, 294-304.

(248) Abhiman, S., and Sonnhammer, E. L. (2005) Large-scale prediction of function shift in protein families with a focus on enzymatic function. *Proteins 60*, 758-768.

(249) Mollet, B., and Pilloud, N. (1991) Galactose utilization in *Lactobacillus helveticus*: isolation and characterization of the galactokinase (*galK*) and galactose-1-phosphate uridyl transferase (*galT*) genes. *J bacteriol 173*, 4464-4473.

(250) Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., and von Mering, C. (2009) STRING 8-a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res. 37*, D412-D416.

(251) Schell, M. A., Karmirantzou, M., Snel, B., Vilanova, D., Berger, B., Pessi, G., Zwahlen, M. C., Desiere, F., Bork, P., Delley, M., Pridmore, R. D., and Arigoni, F. (2002) The genome sequence of *Bifidobacterium longum* reflects its adaptation to the human gastrointestinal tract. *Proc Natl Acad Sci U.S.A. 99*, 14422-14427.

(252) Klaassens, E. S., Ben-Amor, K., Vriesema, A., Vaughan, E. E., and de Vos, W. (2011) The fecal bifidobacterial transcriptome of adults: a microarray approach. *Gut microbes 2*, 217-226.

(253) Gonzalez, R., Klaassens, E. S., Malinen, E., de Vos, W. M., and Vaughan, E. E. (2008) Differential transcriptional response of *Bifidobacterium longum* to human milk, formula milk, and galactooligosaccharide. *Appl Environ Microbiol*. *74*, 4686-4694.

(254) Sela, D. A., Chapman, J., Adeuya, A., Kim, J. H., Chen, F., Whitehead, T. R., Lapidus, A., Rokhsar, D. S., Lebrilla, C. B., German, J. B., Price, N. P., Richardson, P. M., and Mills, D. A. (2008) The genome sequence of *Bifidobacterium longum* subsp. *infantis*

reveals adaptations for milk utilization within the infant microbiome. *Proc Natl Acad Sci U.S.A. 105*, 18964-18969.
(255) Garrido, D., Ruiz-Moyano, S., and Mills, D. A. (2012) Release and utilization of N-acetyl-D-glucosamine from human milk oligosaccharides by *Bifidobacterium longum* subsp. *infantis*. *Anaerobe 18*, 430-435.
(256) Ozaki, S., Imai, H., Iwakiri, T., Sato, T., Shimoda, K., Nakayama, T., and Hamada, H. (2012) Regioselective glucosidation of trans-resveratrol in *Escherichia coli* expressing glucosyltransferase from *Phytolacca americana.* Biotechnol Lett. *34*, 475-481.
(257) Yan, Y., Chemler, J., Huang, L., Martens, S., and Koffas, M. A. (2005) Metabolic engineering of anthocyanin biosynthesis in *Escherichia coli*. *Appl Environ Microbiol 71*, 3617-3623.
(258) Aerts, D., Verhaeghe, T., De Mey, M., Desmet, T., and Soetaert, W. (2011) A constitutive expression system for high-throughput screening. *Eng Life Sci 11*, 10-19.
(259) Alam, J., Shaheen, A., and Anwar, M. S. (2014) Accessing select properties of the electron with ImageJ: an open-source image-processing paradigm. Eur J Phys. *35*.
(260) Takemoto, J. K., and Davies, N. M. (2011) Method development for beta-glucogallin and gallic acid analysis: application to urinary pharmacokinetic studies. *J Pharm Biomed Anal 54*, 812-816.
(261) Archer, C. T., Kim, J. F., Jeong, H., Park, J. H., Vickers, C. E., Lee, S. Y., and Nielsen, L. K. (2011) The genome sequence of *E. coli* W (ATCC 9637): comparative genome analysis and an improved genome-scale reconstruction of *E. coli*. *BMC Genomics 12*, 9.
(262) Nagata, S. (2001) Growth of *Escherichia coli* ATCC 9637 through the uptake of compatible solutes at high osmolarity. *J Biosci Bioeng 92*, 324-329.
(263) Lee, S. Y., and Chang, H. N. (1993) High Cell-Density Cultivation of *Escherichia coli* W Using Sucrose as a Carbon Source. *Biotechnol Lett*. *15*, 971-974.
(264) Mammela, P., Savolainen, H., Lindroos, L., Kangas, J., and Vartiainen, T. (2000) Analysis of oak tannins by liquid chromatography-electrospray ionisation mass spectrometry. *J Chromatogr A*. *891*, 75-83.
(265) Gross, G. G. (1982) Synthesis of Beta-Glucogallin from Udp-Glucose and Gallic Acid by an Enzyme Preparation from Oak Leaves. *FEBS Lett. 148*, 67-70.
(266) Zhang, J., Li, L., Kim, S. H., Hagerman, A. E., and Lu, J. (2009) Anti-cancer, anti-diabetic and other pharmacologic and biological activities of penta-galloyl-glucose. Pharm Res. *26*, 2066-2080.
(267) Majeed, M., Bhat, B., and Susmitha Anand, T. S. (2010) Inhibition of UV induced adversaries by β-glucogallin from Amla (*Emblica officinalis* Gaertn.) fruits. *Indian J Nat Prod Resour 1(4)*, 462-465.
(268) Arifin, Y., Sabri, S., Sugiarto, H., Kromer, J. O., Vickers, C. E., and Nielsen, L. K. (2010) Deletion of *cscR* in *Escherichia coli* W improves growth and poly-3-hydroxybutyrate (PHB) production from sucrose in fed batch culture. *J Biotechnol 156*, 275-278.
(269) Sabri, S., Nielsen, L. K., and Vickers, C. E. (2013) Molecular control of sucrose utilization in *Escherichia coli* W, an efficient sucrose-utilizing strain. *Appl Environ Microbiol 79*, 478-487.
(270) Kornberg, H. L. (2001) Routes for fructose utilization by *Escherichia coli*. *J Mol Microb 3*, 355-359.

(271) Keseler, I. M., Mackie, A., Peralta-Gil, M., Santos-Zavaleta, A., Gama-Castro, S., Bonavides-Martinez, C., Fulcher, C., Huerta, A. M., Kothari, A., Krummenacker, M., Latendresse, M., Muniz-Rascado, L., Ong, Q., Paley, S., Schroder, I., Shearer, A. G., Subhraveti, P., Travers, M., Weerasinghe, D., Weiss, V., Collado-Vides, J., Gunsalus, R. P., Paulsen, I., and Karp, P. D. (2013) EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res 41*, D605-612.

(272) Kuznetsova, E., Proudfoot, M., Gonzales, C. F., Brown, G., Omelchenko, M. V., Borozan, I., Carmel, L., Wolf, Y. I., Mori, H., Savchenko, A. V., Arrowsmith, C. H., Koonin, E. V., Edwards, A. M., and Yakunin, A. F. (2006) Genome-wide analysis of substrate specificities of the *Escherichia coli* haloacid dehalogenase-like phosphatase family. *J Biol Chem*. *281*, 36149-36161.

(273) Jolly, L., Ferrari, P., Blanot, D., Van Heijenoort, J., Fassy, F., and Mengin-Lecreulx, D. (1999) Reaction mechanism of phosphoglucosamine mutase from *Escherichia coli. Eur J Biochem*. *262*, 202-210.

(274) Mengin-Lecreulx, D., and van Heijenoort, J. (1996) Characterization of the essential gene *glmM* encoding phosphoglucosamine mutase in *Escherichia coli. J Biol Chem*. *271*, 32-39.

(275) Lim, C. G. (2011) Metabolic Engineering for high-level flavonoids and stilbenes production in microorganisms. University at Buffalo, State University of New York.

(276) Khater, F., Fournand, D., Vialet, S., Meudec, E., Cheynier, V., and Terrier, N. (2012) Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. *J Exp Bot 63*, 1201-1214.

(277) Perron, N. R., Wang, H. C., Deguire, S. N., Jenkins, M., Lawson, M., and Brumaghim, J. L. (2010) Kinetics of iron oxidation upon polyphenol binding. *Dalton Trans 39*, 9982-9987.

(278) Chung, K. T., Lu, Z., and Chou, M. W. (1998) Mechanism of inhibition of tannic acid and related compounds on the growth of intestinal bacteria. *Food Chem Toxicol*. *36*, 1053-1060.

(279) Saxena, G., McCutcheon, A. R., Farmer, S., Towers, G. H., and Hancock, R. E. (1994) Antimicrobial constituents of *Rhus glabra*. *J Ethnopharmacol 42*, 95-99.

(280) Milkowski, C., and Strack, D. (2010) Sinapate esters in brassicaceous plants: biochemistry, molecular biology, evolution and metabolic engineering. *Planta 232*, 19-35.

(281) Li, A. X., and Steffens, J. C. (2000) An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc Natl Acad Sci U S A 97*, 6902-6907.

(282) Ralph, J. (2010) Hydroxycinnamates in lignification. *Phytochem Rev 9*, 65-83.

(283) Zhu, Y. M., and Ralph, J. (2011) Stereoselective synthesis of 1-O-beta-feruloyl and 1-O-beta-sinapoyl glucopyranoses. *Tetrahedron Lett 52*, 3729-3731.

(284) Jiang, Z. H., Hirose, Y., Iwata, H., Sakamoto, S., Tanaka, T., and Kouno, I. (2001) Caffeoyl, coumaroyl, galloyl, and hexahydroxydiphenoyl glucoses from *Balanophora japonica*. *Chem Pharm Bull (Tokyo) 49*, 887-892.

(285) Klick, S., and Herrmann, K. (1988) Glucosides and Glucose Esters of Hydroxybenzoic Acids in Plants. *Phytochemistry 27*, 2177-2180.

(286) Ohshima, N., Nishitani, H., Nishikawa, S., Okumura, K., and Taguchi, H. (1997) 1-O-nicotinoyl-beta-D-glucopyranose in cultured tobacco cells. *Phytochemistry 44*, 1483-1484.

(287) Furuya, T., Asada, Y., Mizobata, S., Matsuura, Y., and Hamada, H. (1998) Biotransformation of p-aminobenzoic acid by cultured cells of *Eucalyptus perriniana*. *Phytochemistry 49*, 109-111.
(288) Michodjehoun-Mestres, L., Amraoui, W., and Brillouet, J. M. (2009) Isolation, characterization, and determination of 1-O-trans-cinnamoyl-beta-D-glucopyranose in the epidermis and flesh of developing cashew apple (*Anacardium occidentale* L.) and four of its genotypes. *J Agric Food Chem 57*, 1377-1382.
(289) Davis, B. D. (1951) Inhibition of *Escherichia coli* by p-aminobenzoic acid and its reversal by p-hydroxybenzoic acid. *J Exp Med*. *94*, 243-254.
(290) Hall, B. G., and Betts, P. W. (1987) Cryptic genes for cellobiose utilization in natural isolates of *Escherichia coli*. *Genetics 115*, 431-439.
(291) Yang, M., Luoh, S. M., Goddard, A., Reilly, D., Henzel, W., and Bass, S. (1996) The *bglX* gene located at 47.8 min on the *Escherichia coli* chromosome encodes a periplasmic beta-glucosidase. *Microbiology 142 (Pt 7)*, 1659-1665.
(292) Diaz, E., Ferrandez, A., Prieto, M. A., and Garcia, J. L. (2001) Biodegradation of aromatic compounds by *Escherichia coli. Microbiol Mol Biol Rev*. *65*, 523-569, table of contents.
(293) Ren, W., Qiao, Z., Wang, H., Zhu, L., and Zhang, L. (2003) Flavonoids: promising anticancer agents. *Med Res Rev*. *23*, 519-534.
(294) Prochazkova, D., Bousova, I., and Wilhelmova, N. (2011) Antioxidant and prooxidant properties of flavonoids. *Fitoterapia 82*, 513-523.
(295) Pan, M. H., Lai, C. S., and Ho, C. T. (2010) Anti-inflammatory activity of natural dietary flavonoids. *Food & function 1*, 15-31.
(296) Orhan, D. D., Ozcelik, B., Ozgen, S., and Ergun, F. (2010) Antibacterial, antifungal, and antiviral activities of some flavonoids. *Microbiol Res 165*, 496-504.
(297) Kang, S. Y., Kang, J. Y., and Oh, M. J. (2012) Antiviral activities of flavonoids isolated from the bark of *Rhus verniciflua* stokes against fish pathogenic viruses In Vitro. *J Microbiol 50*, 293-300.
(298) Jones, P., Messner, B., Nakajima, J., Schaffner, A. R., and Saito, K. (2003) UGT73C6 and UGT78D1, glycosyltransferases involved in flavonol glycoside biosynthesis in *Arabidopsis thaliana. J Biol Chem*. *278*, 43910-43918.
(299) Hollman, P. C. H., and Arts, I. C. W. (2000) Flavonols, flavones and flavanols - nature, occurrence and dietary burden. *J Sci Food Agr 80*, 1081-1093.
(300) Lister, C. E., Lancaster, J. E., Sutton, K. H., and Walker, J. R. L. (1994) Developmental-Changes in the Concentration and Composition of Flavonoids in Skin of a Red and a Green Apple Cultivar. *J Sci Food Agr*. *64*, 155-161.
(301) Piao, M. J., Kang, K. A., Zhang, R., Ko, D. O., Wang, Z. H., You, H. J., Kim, H. S., Kim, J. S., Kang, S. S., and Hyun, J. W. (2008) Hyperoside prevents oxidative damage induced by hydrogen peroxide in lung fibroblast cells via an antioxidant effect. *Bba-Gen Subjects 1780*, 1448-1457.
(302) Babujanarthanam, R., Kavitha, P., Rao, U. S. M., and Pandian, M. R. (2011) Quercitrin a bioflavonoid improves the antioxidant status in streptozotocin: induced diabetic rat tissues. *Mol Cell Biochem 358*, 121-129.
(303) Yin, Y. Q., Li, W. Q., Son, Y. O., Sun, L. J., Lu, J., Kim, D., Wang, X., Yao, H., Wang, L., Pratheeshkumar, P., Hitron, A. J., Luo, J., Gao, N., Shi, X. L., and Zhang, Z. (2013) Quercitrin protects skin from UVB-induced oxidative damage. *Toxicol Appl Pharm 269*, 89-99.

(304) Choi, J. H., Kim, D. W., Yun, N., Choi, J. S., Islam, M. N., Kim, Y. S., and Lee, S. M. (2011) Protective effects of hyperoside against carbon tetrachloride-induced liver damage in mice. *J Nat Prod 74*, 1055-1060.
(305) Choi, H. J., Song, J. H., and Kwon, D. H. (2012) Quercetin 3-rhamnoside Exerts Antiinfluenza A Virus Activity in Mice. *Phytother Res*. *26*, 462-464.
(306) Wu, L. L., Yang, X. B., Huang, Z. M., Liu, H. Z., and Wu, G. X. (2007) In vivo and in vitro antiviral activity of hyperoside extracted from *Abelmoschus manihot* (L) medik. *Acta Pharmacol Sin*. *28*, 404-409.
(307) Chen, L. L., Li, J., Luo, C., Liu, H., Xu, W. J., Chen, G., Liew, O. W., Zhu, W. L., Puah, C. M., Shen, X., and Jiang, H. L. (2006) Binding interaction of quereetin-3-beta-galactoside and its synthetic derivatives with SARS-CoV 3CL(pro): Structure-activity relationship studies reveal salient pharmacophore features. *Bioorgan Med Chem 14*, 8295-8306.
(308) Comalada, M., Camuesco, D., Sierra, S., Ballester, I., Xaus, J., Galvez, J., and Zarzuelo, A. (2005) In vivo quercitrin anti-inflammatory effect involves release of quercetin, which inhibits inflammation through down-regulation of the NF-kappa B pathway. *Eur J Immunol 35*, 584-592.
(309) Zheng, M. Z., Liu, C. M., Pan, F. G., Shi, D. F., and Zhang, Y. C. (2012) Antidepressant-like effect of hyperoside isolated from *Apocynum venetum* leaves: Possible cellular mechanisms. *Phytomedicine 19*, 145-149.
(310) Butterweck, V., Jurgenliemk, G., Nahrstedt, A., and Winterhoff, H. (2000) Flavonoids from *Hypericum perforatum* show antidepressant activity in the forced swimming test. *Planta Med*. *66*, 3-6.
(311) Cincin, Z. B., Unlu, M., Kiran, B., Bireller, E. S., Baran, Y., and Cakmakoglu, B. (2014) Apoptotic Effects of Quercitrin on DLD-1 Colon Cancer Cell Line. *Pathol Oncol Res*.
(312) Li, S. Y., Zhang, Z. Z., Cain, A., Wang, B., Long, M., and Taylor, J. (2005) Antifungal activity of camptothecin, trifolin, and hyperoside isolated from *Camptotheca acuminata. J Agr Food Chem*. *53*, 32-37.
(313) Dai, J., and Mumper, R. J. (2010) Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. *Molecules 15*, 7313-7352.
(314) Bouktaib, M., Atmani, A., and Rolando, C. (2002) Regio- and stereoselective synthesis of the major metabolite of quercetin, quercetin-3-O-beta-D-glucuronide. *Tetrahedron Lett 43*, 6263-6266.
(315) Kajjout, M., and Rolando, C. (2011) Regiospecific synthesis of quercetin O-beta-D-glucosylated and O-beta-D-glucuronidated isomers. *Tetrahedron 67*, 4731-4741.
(316) Yamauchi, K., Mitsunaga, T., and Batubara, I. (2014) Synthesis of quercetin glycosides and their melanogenesis stimulatory activity in B16 melanoma cells. *Bioorg Med Chem*. *22*, 937-944.
(317) Wang, Y. C., Chen, S., and Yu, O. (2011) Metabolic engineering of flavonoids in plants and microorganisms. *Appl Microbiol Biot*. *91*, 949-956.
(318) Yoon, J. A., Kim, B. G., Lee, W. J., Lim, Y., Chong, Y., and Ahn, J. H. (2012) Production of a novel quercetin glycoside through metabolic engineering of *Escherichia coli*. *Appl Environ Microbiol 78*, 4256-4262.
(319) De Bruyn, F., De Paepe, B., Beauprez, J., Maertens, J., De Cocker, P., Stevens, C., Soetaert, W., and De Mey, M. (2014) Building a Glucosylation Platform in *Escherichia coli* trough Metabolic Engineering: Coupling Production to Growth. *Biores Tech.*
(320) Faury, G., Molinari, J., Rusova, E., Mariko, B., Raveaud, S., Huber, P., Velebny, V., Robert, A. M., and Robert, L. (2011) Receptors and aging: structural selectivity of

the rhamnose-receptor on fibroblasts as shown by Ca(2+)-mobilization and gene-expression profiles. *Arch Gerontol Geriatr*. *53*, 106-112.

(321) Zou, R., Zhou, K., Stephanopoulos, G., and Too, H. P. (2013) Combinatorial engineering of 1-deoxy-D-xylulose 5-phosphate pathway using cross-lapping in vitro assembly (CLIVA) method. *PloS one 8*, e79557.

(322) Kaya, B., Menemen, Y., and Saltan, F. Z. (2012) Flavonoid compounds identified in *Alchemilla* L. species collected in the north-eastern Black Sea region of Turkey. *Afr J Tradit Complement Altern Med*. *9*, 418-425.

(323) Zhang, J., Kowal, P., Chen, X., and Wang, P. G. (2003) Large-scale synthesis of globotriose derivatives through recombinant *E. coli*. *Org Biomol Chem 1*, 3048-3053.

(324) Zhou, T., Chen, B., Fan, G., Chai, Y., and Wu, Y. (2006) Application of high-speed counter-current chromatography coupled with high-performance liquid chromatography-diode array detection for the preparative isolation and purification of hyperoside from *Hypericum perforatum* with online purity monitoring. *J Chromatogr A*. *1116*, 97-101.

(325) Renouf, M. A., Wegener, M. K., and Nielsen, L. K. (2008) An environmental life cycle assessment comparing Australian sugarcane with US corn and UK sugar beet as producers of sugars for fermentation. *Biomass Bioenerg 32*, 1144-1155.

(326) Mrazek, J. (2010) Comparative Analysis of Sequence Periodicity among Prokaryotic Genomes Points to Differences in Nucleoid Structure and a Relationship to Gene Expression. *J Bacteriol*. *192*, 3763-3772.

(327) Kravatskaya, G. I., Chechetkin, V. R., Kravatsky, Y. V., and Tumanyan, V. G. (2013) Structural attributes of nucleotide sequences in promoter regions of supercoiling-sensitive genes: How to relate microarray expression data with genomic sequences. *Genomics 101*, 1-11.

(328) Miller, K. D., Guyon, V., Evans, J. N., Shuttleworth, W. A., and Taylor, L. P. (1999) Purification, cloning, and heterologous expression of a catalytically efficient flavonol 3-O-galactosyltransferase expressed in the male gametophyte of *Petunia hybrida*. *J Biol Chem*. *274*, 34011-34019.

(329) Aherne, S. A., and O'Brien, N. M. (2002) Dietary flavonols: chemistry, food content, and metabolism. *Nutrition 18*, 75-81.

(330) Baranowska, I., and Magiera, S. (2011) Development and validation of a UHPLC method for the determination of flavonoids in red wine. *J AOAC Int 94*, 786-794.

(331) Diantini, A., Subarnas, A., Lestari, K., Halimah, E., Susilawati, Y., Supriyatna, Julaeha, E., Achmad, T. H., Suradji, E. W., Yamazaki, C., Kobayashi, K., Koyama, H., and Abdulah, R. (2012) Kaempferol-3-O-rhamnoside isolated from the leaves of *Schima wallichii* Korth. inhibits MCF-7 breast cancer cell proliferation through activation of the caspase cascade pathway. *Oncol Lett 3*, 1069-1072.

(332) Orhan, I., Kiipeli, E., Terzioglu, S., and Yesilada, E. (2007) Bioassay-guided isolation of kaempferol-3-O-beta-D-galactoside with anti-inflammatory and antinociceptive activity from the aerial part of *Calluna vulgaris* L. *J Ethnopharmacol 114*, 32-37.

(333) Hayder, N., Bouhlel, I., Skandrani, I., Kadri, M., Steiman, R., Guiraud, P., Mariotte, A. M., Ghedira, K., Dijoux-Franca, M. G., and Chekir-Ghedira, L. (2008) In vitro antioxidant and antigenotoxic potentials of myricetin-3-o-galactoside and myricetin-3-o-rhamnoside from *Myrtus communis*: Modulation of expression of genes involved in cell defence system using cDNA microarray. *Toxicol in Vitro 22*, 567-581.

(334) Pereira, M., Siba, I. P., Chioca, L. R., Correia, D., Vital, M. A. B. F., Pizzolatti, M. G., Santos, A. R. S., and Andreatini, R. (2011) Myricitrin, a nitric oxide and protein kinase C inhibitor, exerts antipsychotic-like effects in animal models. *Prog Neuro-Psychoph 35*, 1636-1644.

(335) Adhami, V. M., Syed, D. N., Khan, N., and Mukhtar, H. (2012) Dietary flavonoid fisetin: a novel dual inhibitor of PI3K/Akt and mTOR for prostate cancer management. *Biochem Pharmacol 84*, 1277-1281.

(336) Yoo, H., Ku, S. K., Han, M. S., Kim, K. M., and Bae, J. S. (2014) Anti-septic Effects of Fisetin In Vitro and In Vivo. *Inflammation*.

(337) Rattanachaikunsopon, P., and Phumkhachorm, P. (2007) Bacteriostatic effect of flavonoids isolated from leaves of Psidium guajava on fish pathogens. *Fitoterapia 78*, 434-436.

(338) Paoli, P., Cirri, P., Caselli, A., Ranaldi, F., Bruschi, G., Santi, A., and Camici, G. (2013) The insulin-mimetic effect of Morin: A promising molecule in diabetes treatment. *Bba-Gen Subjects 1830*, 3102-3111.

(339) Hollman, P. C. H., Bijsman, M. N. C. P., van Gameren, Y., Cnossen, E. P. J., de Vries, J. H. M., and Katan, M. B. (1999) The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. *Free Radical Res 31*, 569-573.

(340) Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., and Davies, G. J. (2006) Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *The EMBO journal 25*, 1396-1405.

(341) Heissigerova, H., Breton, C., Moravcova, J., and Imberty, A. (2003) Molecular modeling of glycosyltransferases involved in the biosynthesis of blood group A, blood group B, Forssman, and iGb3 antigens and their interaction with substrates. *Glycobiology 13*, 377-386.

(342) Bekri, S., Lidove, O., Jaussaud, R., Knebelmann, B., and Barbey, F. (2006) The role of ceramide trihexoside (globotriaosylceramide) in the diagnosis and follow-up of the efficacy of treatment of Fabry disease: a review of the literature. *Cardiovasc Hematol Agents Med Chem*. *4*, 289-297.

(343) Stanley, P., and Cummings, R. D. (2009) Structures Common to Different Glycans, In *Essentials of Glycobiology* (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., Eds.) 2nd ed., Cold Spring Harbor (NY).

(344) Reid, J. S., Edwards, M. E., Gidley, M. J., and Clark, A. H. (1992) Mechanism and regulation of galactomannan biosynthesis in developing leguminous seeds. *Biochem Soc Trans 20*, 23-26.

(345) Galili, U. (2005) The alpha-gal epitope and the anti-Gal antibody in xenotransplantation and in cancer immunotherapy. *Immunol Cell Biol 83*, 674-686.

(346) Bode, L. (2012) Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology 22*, 1147-1162.

(347) Hellberg, A., Westman, J. S., Thuresson, B., and Olsson, M. L. (2013) P1PK: the blood group system that changed its name and expanded. *Immunohematology / American Red Cross 29*, 25-33.

(348) Liu, M. Z., Fan, H. N., and Lee, Y. C. (2001) Chemical synthesis of a P(k)-antigenic globotriose analogue with a functionalized aglycon. *Biochimie 83*, 693-698.

(349) Leach, J. L., Garber, S. A., Marcon, A. A., and Prieto, P. A. (2005) In vitro and in vivo effects of soluble, monovalent globotriose on bacterial attachment and colonization. *Antimicrobial agents and chemotherapy 49*, 3842-3846.
(350) Li, X., Wu, P., Cheng, S., and Lv, X. (2012) Synthesis and assessment of globotriose-chitosan conjugate, a novel inhibitor of shiga toxins produced by *Escherichia coli*. *J Med Chem 55*, 2702-2710.
(351) Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2001) Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat Struct Biol 8*, 166-175.
(352) Bingham, A. H., and Busby, S. J. (1987) Translation of galE and coordination of galactose operon expression in Escherichia coli: effects of insertions and deletions in the non-translated leader sequence. *Mol Microbiol 1*, 117-124.
(353) Roche, E. D., and Sauer, R. T. (2001) Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons. *J Biol Chem*. *276*, 28509-28515.
(354) Subbaiah, C. C., Huber, S. C., Sachs, M. M., and Rhoads, D. (2007) Sucrose synthase: expanding protein function. *Plant Signal Behav*. *2*, 28-29.
(355) Baroja-Fernandez, E., Munoz, F. J., Li, J., Bahaji, A., Almagro, G., Montero, M., Etxeberria, E., Hidalgo, M., Sesma, M. T., and Pozueta-Romero, J. (2012) Sucrose synthase activity in the sus1/sus2/sus3/sus4 *Arabidopsis* mutant is sufficient to support normal cellulose and starch production. *Proc Natl Acad Sci U S A 109*, 321-326.
(356) Fujii, S., Hayashi, T., and Mizuno, K. (2010) Sucrose synthase is an integral component of the cellulose synthesis machinery. *Plant Cell Physiol 51*, 294-301.
(357) Klotz, K. L., Fingerb, F. L., and Shelverc, W. L. (2003) Characterization of two sucrose synthase isoforms in sugarbeet root. *Plant Physiol Biochem*. *41*, 107-115.
(358) Bieniawska, Z., Paul Barratt, D. H., Garlick, A. P., Thole, V., Kruger, N. J., Martin, C., Zrenner, R., and Smith, A. M. (2007) Analysis of the sucrose synthase gene family in *Arabidopsis. Plant J*. *49*, 810-828.
(359) Martin, T., Frommer, W. B., Salanoubat, M., and Willmitzer, L. (1993) Expression of an Arabidopsis sucrose synthase gene indicates a role in metabolization of sucrose both during phloem loading and in sink organs. *The Plant journal : for cell and molecular biology 4*, 367-377.
(360) Porchia, A. C., Curatti, L., and Salerno, G. L. (1999) Sucrose metabolism in cyanobacteria: sucrose synthase from *Anabaena* sp. strain PCC 7119 is remarkably different from the plant enzymes with respect to substrate affinity and amino-terminal sequence. *Planta 210*, 34-40.
(361) Bulter, T., and Elling, L. (1999) Enzymatic synthesis of nucleotide sugars. *Glycoconjugate journal 16*, 147-159.
(362) Zervosen, A., and Elling, L. (1996) A novel three-enzyme reaction cycle for the synthesis of N-acetyllactosamine with in situ regeneration of uridine 5'-diphosphate glucose and uridine 5'-diphosphate galactose. *J Am Chem Soc*. *118*, 1836-1840.
(363) Huber, S. C., Huber, J. L., Liao, P. C., Gage, D. A., McMichael, R. W., Jr., Chourey, P. S., Hannah, L. C., and Koch, K. (1996) Phosphorylation of serine-15 of maize leaf sucrose synthase. Occurrence in vivo and possible regulatory significance. *Plant Physiol 112*, 793-802.

(364) Zheng, Y., Anderson, S., Zhang, Y., and Garavito, R. M. (2011) The structure of sucrose synthase-1 from *Arabidopsis thaliana* and its functional implications. *J Biol Chem*. *286*, 36108-36118.
(365) Baderschneider, B., and Winterhalter, P. (2001) Isolation and characterization of novel benzoates, cinnamates, flavonoids, and lignans from Riesling wine and screening for antioxidant activity. *J Agric Food Chem 49*, 2788-2798.
(366) Kiss, R., Dubois, J., Neve, J., Lamoral-Theys, D., Dufrasne, F., and Pottier, L. (2011) Di-Vanilloyl And Tri-Vanilloyl Derivatives For Use In Anti-Cancer Therapy. Google Patents.
(367) Rasila, T. S., Pajunen, M. I., and Savilahti, H. (2009) Critical evaluation of random mutagenesis by error-prone polymerase chain reaction protocols, Escherichia coli mutator strain, and hydroxylamine treatment. *Anal Biochem 388*, 71-80.
(368) Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics 9*, 40.
(369) The PyMOL Molecular Graphics System, V. r. p., Schrödinger, LLC.
(370) Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem 30*, 2785-2791.
(371) Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem 31*, 455-461.
(372) Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., and Damborsky, J. (2006) CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics 7*, 316.
(373) Aerts, D. (2012) Screening and engineering of sucrose phosphorylase for the glycosylation of small molecules. PhD thesis, Faculty of Bioscience Engineering, Ghent University, Ghent.
(374) Figueroa, C. M., Diez, M. D. A., Kuhn, M. L., McEwen, S., Salerno, G. L., Iglesias, A. A., and Ballicora, M. A. (2013) The unique nucleotide specificity of the sucrose synthase from *Thermosynechococcus elongatus*. *FEBS Lett. 587*, 165-169.
(375) Sauerzapfe, B., Engels, L., and Elling, L. (2008) Broadening the biocatalytic properties of recombinant sucrose synthase 1 from potato (*Solanum tuberosum* L.) by expression in *Escherichia coli* and *Saccharomyces cerevisiae*. *Enzyme Microb Tech 43*, 289-296.
(376) Baroja-Fernandez, E., Munoz, F. J., Saikusa, T., Rodriguez-Lopez, M., Akazawa, T., and Pozueta-Romero, J. (2003) Sucrose synthase catalyzes the de novo production of ADPglucose linked to starch biosynthesis in heterotrophic tissues of plants. *Plant Cell Physiol 44*, 500-509.
(377) Morell, M., and Copeland, L. (1985) Sucrose synthase of soybean nodules. *Plant Physiol 78*, 149-154.
(378) Zhang, X. Q., Lund, A. A., Sarath, G., Cerny, R. L., Roberts, D. M., and Chollet, R. (1999) Soybean nodule sucrose synthase (nodulin-100): further analysis of its phosphorylation using recombinant and authentic root-nodule enzymes. *Arch Biochem Biophys*. *371*, 70-82.
(379) Barratt, D. H., Barber, L., Kruger, N. J., Smith, A. M., Wang, T. L., and Martin, C. (2001) Multiple, distinct isoforms of sucrose synthase in pea. *Plant Physiol 127*, 655-664.

(380) Dejardin, A., Rochat, C., Maugenest, S., and Boutin, J. P. (1997) Purification, characterization and physiological role of sucrose synthase in the pea seed coat (*Pisum sativum* L.). *Planta 201*, 128-137.
(381) Echt, C. S., and Chourey, P. S. (1985) A Comparison of Two Sucrose Synthetase Isozymes from Normal and shrunken-1 Maize. *Plant Physiol 79*, 530-536.
(382) Tsai, C. Y. (1974) Sucrose-Udp Glucosyltransferase of Zea-Mays Endosperm. *Phytochemistry 13*, 885-891.
(383) Su, J. C., and Preiss, J. (1978) Purification and properties of sucrose synthase from maize kernels. *Plant Physiol 61*, 389-393.
(384) Zervosen, A., Romer, U., and Elling, L. (1998) Application of recombinant sucrose synthase large scale synthesis of ADP-glucose. *J Mol Catal B-Enzym 5*, 25-28.
(385) Romer, U., Nettelstroth, N., Kockenberger, W., and Elling, L. (2001) Characterization of recombinant sucrose synthase 1 from potato for the synthesis of sucrose analogues. *Adv Synth Catal 343*, 655-661.
(386) Clarke, T. F. t., and Clark, P. L. (2008) Rare codons cluster. *PloS one 3*, e3412.
(387) Olsen, M., Iverson, B., and Georgiou, G. (2000) High-throughput screening of enzyme libraries. *Curr Opin Biotech 11*, 331-337.
(388) Van Gelder, P., Dutzler, R., Dumas, F., Koebnik, R., and Schirmer, T. (2001) Sucrose transport through maltoporin mutants of *Escherichia coli. Protein Eng*. *14*, 943-948.
(389) Sugihara, J., Smirnova, I., Kasho, V., and Kaback, H. R. (2011) Sugar recognition by CscB and LacY. *Biochemistry 50*, 11009-11014.
(390) De Bruyn, F. (2010) Ontwikkeling van een *E. coli* mutant voor de productie van zeldzame suikers. Ghent University. Master thesis.
(391) Sproul, A. A., Lambourne, L. T., Jean-Jacques, D. J., and Kornberg, H. L. (2001) Genetic control of manno(fructo)kinase activity in *Escherichia coli*. *P Natl Acad Sci USA 98*, 15257-15259.
(392) Zrenner, R., Salanoubat, M., Willmitzer, L., and Sonnewald, U. (1995) Evidence of the crucial role of sucrose synthase for sink strength using transgenic potato plants (*Solanum tuberosum* L.). *Plant J*. *7*, 97-107.
(393) Tang, G. Q., and Sturm, A. (1999) Antisense repression of sucrose synthase in carrot (*Daucus carota* L.) affects growth rather than sucrose partitioning. *Plant Mol Biol 41*, 465-479.
(394) Fraichard, A., Trossat, C., Perotti, E., and Pugin, A. (1996) Allosteric regulation by Mg2+ of the vacuolar H(+)-PPase from Acer pseudoplatanus cells. Ca2+/Mg2+ interactions. *Biochimie 78*, 259-266.
(395) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A 98*, 10037-10041.
(396) Unni, S., Huang, Y., Hanson, R. M., Tobias, M., Krishnan, S., Li, W. W., Nielsen, J. E., and Baker, N. A. (2011) Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J Comput Chem 32*, 1488-1491.
(397) Breton, C., Snajdrova, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology 16*, 29R-37R.
(398) Van Dijck, P., Mascorro-Gallardo, J. O., De Bus, M., Royackers, K., Iturriaga, G., and Thevelein, J. M. (2002) Truncation of *Arabidopsis thaliana* and *Selaginella lepidophylla* trehalose-6-phosphate synthase unlocks high catalytic activity and supports high trehalose levels on expression in yeast. *Biochem J*. *366*, 63-71.

(399) Henion, T. R., Macher, B. A., Anaraki, F., and Galili, U. (1994) Defining the minimal size of catalytically active primate alpha 1,3 galactosyltransferase: structure-function studies on the recombinant truncated enzyme. *Glycobiology 4*, 193-201.
(400) Hardin, S. C., Duncan, K. A., and Huber, S. C. (2006) Determination of structural requirements and probable regulatory effectors for membrane association of maize sucrose synthase. *Plant Physiol. 141*, 1106-1119.
(401) Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol 5*, 593-599.
(402) Bornscheuer, U. T., Huisman, G. W., Kazlauskas, R. J., Lutz, S., Moore, J. C., and Robins, K. (2012) Engineering the third wave of biocatalysis. *Nature 485*, 185-194.
(403) Wiechert, W. (2001) 13C metabolic flux analysis. *Metab Eng 3*, 195-206.
(404) Inoue, K. H., and Hagerman, A. E. (1988) Determination of gallotannin with rhodanine. *Anal Biochem 169*, 363-369.
(405) Chua, T. K., Bujnicki, J. M., Tan, T. C., Huynh, F., Patel, B. K., and Sivaraman, J. (2008) The structure of sucrose phosphate synthase from *Halothermothrix orenii* reveals its mechanism of action and binding mode. *Plant Cell 20*, 1059-1072.
(406) Arp, D. J., Chain, P. S., and Klotz, M. G. (2007) The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. *Annu Rev Microbiol 61*, 503-528.
(407) Neves, A. R., Pool, W. A., Solopova, A., Kok, J., Santos, H., and Kuipers, O. P. (2010) Towards enhanced galactose utilization by *Lactococcus lactis*. *Appl. Environ. Microbiol. 76*, 7048-7060.

Summary

Glycosylation is one of the most important and widespread modifications in nature. Adding a sugar residue to (lipophilic) small molecules greatly alters or enhances their solubility, stability or bioactivity, which are desirable properties when they are applied as nutraceuticals, therapeutics or cosmetics. The majority of specialized (secondary) metabolites exists in its glycosylated form, with well-known examples like steviosides (used as sweeteners), the skin lightener arbutine, the chemotherapeutic daunorubicin or the antiviral compound quercitrin. Glycosides can be extracted from natural sources such as plants, or can be chemically synthesized. However, this is often a low-yielding and laborious task using toxic solvents. In nature, glycosylation is mainly catalyzed by glycosyltransferases (GTs), which efficiently and selectively transfer a sugar residue from an activated sugar donor (such as UDP-sugars) to a myriad of acceptor molecules. Despite the potential of GTs, their use in *in vitro* (enzymatic) synthesis is often hampered by the exuberant price of nucleotide sugars (UDP-glucose: 150 €/g; UDP-galactose: 1500 €/g) and difficulties regarding scale-up. Fuelled by these challenges and a high demand for glycosides, various *in vivo* (whole cell) systems have been developed whereby the host cell produces its own UDP-sugars and enzymes *in situ* starting from cheap substrates. Although very promising, economically viable processes remain scarce due to low product yields, titers and rates.

Therefore, the main objective of this PhD research is to create a novel *in vivo* glycosylation platform which is easily scalable and couples production to growth. To this end, the model organism *Escherichia coli* W was metabolically engineered to obtain a generic yet versatile host that effectively glycosylates various small molecules, hereby using only sucrose (250 €/ton) as cheap carbon source.

To ensure adequate UDP-sugar formation in the *E. coli* host, a robust and flexible uridylyltransferase needs to be expressed. The native enzyme (GalU) 'activates' glucose 1-phosphate (glc1P) into the corresponding UDP-glucose, but is unable to catalyze the formation of other UDP-sugars. To this end, a promiscuous uridylyltransferase was searched for in *Bifidobacterium bifidum* and three gene candidates (*ugpA*, *galT1* and *galT2*) were selected and expressed in a newly engineered *E. coli* strain. Expression in this host enabled to screen and characterize the enzymes only by their crude extracts since major interference reactions were deleted, thus alleviating the need for enzyme

purification. The extracts were firstly screened for hexose-1-phosphate uridylyltransferase activity due to the development of a novel chemo-enzymatic assay based on phosphate detection in the micromolar range. UgpA was identified as the promiscuous uridylyltransferase with activity towards both glc1P and galactose 1-phosphate (gal1P). Interestingly, UgpA showed a 7-fold higher activity compared to the native GalU. On the other hand, GalT1 and GalT2 exhibited UDP-glucose-hexose-1-phosphate uridylyltransferase activity, with the latter being 375-fold less active, thus revealing that GalT1 plays a more significant role than previously thought regarding galactose metabolism in *B. bifidum*.

To harness the power of the newly discovered UgpA, it was used to build a glucosylation platform in the form of a metabolically engineered *E. coli* W strain. Although this strain is sucrose positive by the presence of a sucrose hydrolase (CscA), its metabolism was altered by introducing a sucrose phosphorylase from *Bifidobacterium adolescentis* (BaSP). This novel strategy is advantageous since the precursor glc1P is directly formed together with fructose, which serves as a carbon source. By introducing specific gene (*pgm*, *agp*, *ushA* and *glgC*) deletions, the formed glc1P is exclusively channeled towards UDP-glucose. Expression of a glucosyltransferase from *Vitis vinifera* (VvGT2) enabled the strain to efficiently produce β-glucogallin (1-*O*-galloyl-β-D-glucose) starting from gallic acid and sucrose. By implementing this novel metabolic engineering approach, growth and production occurred simultaneously. Indeed, specific productivities could be reached up to 100 mg β-glucogallin/g CDW/h, which was 20-fold higher than compared to the wild type strain. This effect was largely caused by the deletion of *ushA*, which prevents the degradation of the UDP-sugar pool. Due to the promiscuous character of VvGT2, 14 other glucose esters of various hydroxycinnamic and hydroxybenzoic acids were produced in a stereoselective way with high conversion yields up to 100 %, including 3 newly reported glucose esters. To our knowledge, this fast growing (and coupled production) *E. coli* mutant is the first versatile host described for the glucosylation of phenolic acids in a fermentative way using only sucrose as a cheap and sustainable carbon source.

In view of the pivotal role of UDP-glc as ideal starting point for the production of a large variety of UDP-sugars, the glucosylation strain was expanded towards formation of UDP-galactose (UDP-gal) and UDP-rhamnose (UDP-rha) by overexpressing a UDP-glucose epimerase (GalE) and UDP-rhamnose synthase (MUM4), respectively. Due to its beneficial effects on human health, the flavonol quercetin was chosen as target for the production of hyperoside (quercetin 3-*O*-galactoside) and quercitrin (quercetin 3-*O*-rhamnoside). The production of these bio-active flavonol glycosides was achieved by co-

expression with a galactosyltransferase from *Petunia hybrida* (F3GT) or a rhamnosyltransferase from *A. thaliana* (RhaGT). The created strains were able to produce 900 mg/L of hyperoside and 1176 mg/L of quercitrin at specific production rates of 68.7 mg/g CDW/h and 47.8 mg/g CDW/h, respectively, which are the highest reported to date. Furthermore, the glycosylation potential towards other flavonols was also evaluated, with the formation of over 400 mg/L of kaempferol 3-*O*-rhamnoside as most successful example. In addition, the galactosylation strain proved equally flexible for the production of therapeutic and extremely expensive oligosaccharide globotriose (α-galactosyl-1,4-lactose). To this end, the lipo-oligosaccharide galactosyltransferase from *Neisseria meningitidis* (LgtC) was expressed. However, to prevent metabolization of the acceptor lactose, the *lac* operon, containing an β-galactosidase (*lacZ*) and lactose permase (*lacY*) was knocked out and lactose transport was restored by knocking a constitutively expressed *lacY* back in at the same locus. By optimization of the added lactose concentrations, a specific production rate of 117 mg globotriose/g CDW/h was obtained, yielding a final titer of 2.8 g/L.

Finally, the potential of sucrose synthase (SuSy) for the one-step generation of UDP-glucose from sucrose was explored. Since combination of SuSy with a GT gives rise to a rapid and energy efficient recycling of UDP, exploiting this principle in actively growing cells is industrially very attractive. However, due to its unfavorable kinetics (low affinity for sucrose), SuSy has been rarely used for the *in vivo* formation of UDP-glucose. To create a more suitable SuSy, a new screening system was developed to rapidly evaluate and characterize SuSy mutants. SuSy from *Solanum tuberosum* was used as a starting point and displayed an affinity constant (K_m) of 105 mM. This value could be significantly reduced to 35.7 mM by introducing an S11E mutation, which mimics phosphorylation of the serine residue. Coexpression of this SuSy mutant with the glucosyltransferase VvGT2 in the engineered strain created in chapter 4 resulted in an *in vivo* (bioconversion) host. The permeabilized cells were able to efficiently glucosylate vanillic acid at a rate of 0.41 g/L/h with 68 % conversion and an RC_{max} of 8. This system alleviates the need to purify the proteins involved and is easily scalable.

To conclude, the most important outcome of this PhD thesis is the creation of versatile glycosylation platform using a completely novel metabolic strategy which couples growth and production. In this way, *in vivo* glucosylation, galactosylation and rhamnosylation of various molecules could be achieved, hereby outcompeting the majority of glycosylation processes described to date. Therefore, this unique fermentative approach using only sucrose as sustainable carbon source paved the way for economically viable production of glycosides.

Samenvatting

Glycosylatie, de additie van een suikergroep aan een andere molecule, is één van de belangrijkste en meest voorkomende modificaties in de natuur. Dit proces beïnvloedt of verandert de oplosbaarheid, stabiliteit of bio-activiteit van vele verbindingen, wat erg aantrekkelijke eigenschappen zijn wanneer deze moleculen worden gebruikt in de farmaceutische of voedingsindustrie. Bekende voorbeelden zijn steviosiden, die gebruikt worden als laag-calorische zoetstof, daunorubicine, een gevestigd chemotherapeuticum, of quercitrine, een veelbelovend antiviraal middel. De meerderheid van secundaire metabolieten komt voor in zijn geglycosyleerde vorm, waardoor tot op heden de meeste glycosiden geëxtraheerd worden uit natuurlijke bronnen zoals plantenmateriaal. Deze extractieprocessen zijn echter vaak zeer arbeidsintensief en gaan gepaard met lage opbrengsten. Anderzijds wordt chemische synthese van geglycosyleerde verbindingen sterk bemoeilijkt door de vele reactieve groepen, waardoor vele beschermings- en ontschermingsstappen nodig zijn die veel afval genereren. Als duurzaam alternatief wordt daarom tegenwoordig gebruikt gemaakt van glycosyltransferasen (GT's), die het merendeel van glycosyleringsreacties in de natuur katalyseren. Deze enzymen brengen zeer efficiënt en selectief een suikergroep over van een geactiveerde suikerdonor (bijvoorbeeld UDP-suikers) naar een groot aantal acceptormoleculen. Ondanks het potentieel van deze GT's wordt hun gebruik in enzymatische synthese (*in vitro*) vaak belemmerd door de exuberante prijs van de benodigde UDP-suikers (kostprijs UDP-glucose = 150 €/g; UDP-galactose = 1500 €/g). Door de steeds groter wordende vraag naar glycosiden zijn recent diverse cellulaire (*in vivo*) systemen ontwikkeld, waarbij een microbiële cel zelf zijn UDP-suikers en enzymen produceert vertrekkende van goedkope substraten. Hoewel deze nieuwe processen veelbelovend zijn, blijft hun uitvoering op grote schaal verre van economisch rendabel ten gevolge van lage productiviteiten, opbrengsten en titers.

Daarom is de belangrijkste doelstelling van dit onderzoek om een geheel nieuw *in vivo* glycosyleringsplatform te creëren dat eenvoudig opschaalbaar is en eveneens productie koppelt met groei. Hiertoe werd het modelorganisme *Escherichia coli* W metabolisch gemodificeerd zodat een generieke en veelzijdige gastheer werd bekomen die verschillende kleine moleculen kan glycosyleren, en hierbij bovendien enkel gebruik maakt van sucrose (250 € / ton) als goedkope koolstofbron.

Om adequate vorming van UDP-suikers in de *E. coli* gastheer te verzekeren, dient een flexibel uridylyltransferase tot expressie te worden gebracht. Het natieve *E. coli*-enzym GalU "activeert" glucose 1-fosfaat (glc1P) tot het overeenkomstig UDP-glucose, maar kan de vorming van andere UDP-suikers niet katalyseren. Daarom werd een promiscue uridylyltransferase gezocht in *Bifidobacterium bifidum* waarna drie kandidaat genen (*ugpA*, *galT1* en *galT2*) werden geselecteerd en tot expressie werden gebracht in een gemodificeerde *E. coli* mutant. Expressie in deze mutant maakte het mogelijk om de onderzochte enzymen enkel op basis van hun ruwe celextracten te screenen en karakteriseren doordat de belangrijkste afbraakreacties waren verwijderd en het bijgevolg niet nodig was om de enzymen op te zuiveren. De extracten werden eerst gescreend op hexose-1-fosfaat uridylyltransferase activiteit door middel van een nieuwe chemo-enzymatische assay, dewelke gebaseerd was op fosfaatdetectie op micromolair niveau. UgpA werd geïdentificeerd als promiscue uridylyltransferase dat zowel glc1P en galactose 1-fosfaat (gal1P) als substraat kon gebruiken. Bovendien vertoonde UgpA een 7-maal hogere activiteit vergeleken met het natieve GalU. GalT1 en GalT2 daarentegen vertoonden een UDP-glucose-hexose-1-fosfaat uridylyltransferase activiteit. GalT2 bleek 375 keer minder actief dan GalT1, waardoor dit enzym een veel belangrijker rol speelde in het complexe galactosemetabolisme van *B. bifidum* dan tot nu toe gedacht.

Het geïdentificeerde UgpA werd gebruikt om eerst een glucosyleringsplatform te bouwen in een metabolisch gemanipuleerde *E. coli* W stam. Hoewel dit organisme van nature in staat is om sucrose te gebruiken door de aanwezigheid van een sucrose hydrolase (CscA), werd het metabolisme veranderd door introductie van het sucrose fosforylase uit *Bifidobacterium adolescentis* (BaSP). Deze nieuwe strategie zorgt ervoor dat de precursor glc1P direct gevormd wordt samen met fructose, die als koolstofbron fungeert. Door gericht specifieke genen (*pgm*, *agp*, *ushA* en *glgC*) uit te schakelen, werd het gevormde glc1P voorbehouden voor de vorming van UDP-glucose. Expressie van een glucosyltransferase uit *Vitis vinifera* (VvGT2) stelde de stam in staat om efficiënt β-glucogalline (1-*O*-galloyl-β-D-glucose) te produceren uitgaande van galluszuur en sucrose. Implementatie van deze nieuwe *metabolic engineering* strategie zorgde ervoor dat groei en productie gelijktijdig plaatsvonden en aan elkaar gekoppeld waren. Op deze manier kon een specifieke productiviteit worden bereikt boven de 100 mg β-glucogalline/g CDW/h, wat 20 maal hoger was dan deze in de wild type stam. Dit effect werd grotendeels veroorzaakt door het uitschakelen van *ushA*, wat afbraak van de intracellulaire UDP-suikers verhindert. Door het promiscue karakter van VvGT2 konden 14 andere glucose esters van verschillende hydroxykaneelzuren en hydroxybenzoëzuren worden gevormd op een stereoselectieve wijze met hoge

conversies tot 100 %. Deze snel groeiende (en simultaan producerende) *E. coli* stam is de eerste veelzijdige mutant voor de glucosylering van fenolische zuren op een fermentatieve manier, waarbij enkel sucrose wordt verbruikt als een goedkope en duurzame koolstofbron.

In de natuur speelt UDP-glucose een centrale rol als startpunt voor de vorming van vele andere UDP-suikers. Hierdoor kon het ontworpen glucosyleringsplatform makkelijk worden uitgebreid tot het vormen van UDP-galactose en UDP-rhamnose door overexpressie van respectievelijk een UDP-glucose epimerase (GalE) en een UDP-rhamnose synthase (MUM4). Het flavonol quercetine werd omwille van zijn interessante gezondheidseffecten gekozen als startverbinding voor de productie van hyperoside (quercetine 3-*O*-galactoside) en quercitrine (quercetine 3-*O*-rhamnoside). De productie van deze bio-actieve flavonolglycosiden gebeurde door co-expressie met een galactosyltransferase van *Petunia hybrida* (F3GT) of een rhamnosyltransferase van *Arabidopsis thaliana* (RhaGT). De gecreëerde mutanten konden 900 mg/L hyperoside en 1176 mg/L quercitrine produceren met een specifieke productiviteit van 68,7 mg/g CDW/h en 47,8 mg/g CDW/h, respectievelijk, dewelke de hoogste zijn tot nu toe beschreven. Bovendien werd het glycosyleringspotentieel tegenover andere flavonolen eveneens geëvalueerd, waarbij de vorming van meer dan 400 mg/L kaempferol 3-*O*-rhamnoside het meest succesvol bleek. Voorts kon het ontworpen galactosyleringsplatform worden aangepast voor de productie van het therapeutische oligosacharide globotriose (α-galactosyl-1,4-lactose). Daartoe werd het lipo-oligosacharide galactosyltransferase van *Neisseria meningitidis* (LgtC) tot overexpressie gebracht en werd lactose toegevoegd als acceptormolecule. Hiertoe werd het *lac* operon uitgeschakeld (om degradatie van lactose te vermijden) en vervangen door een constitutieve *knock-in* van het lactose permease *lacY*. Door optimalisatie van de toegevoegde lactose concentratie kon een specifieke productiesnelheid van 117 mg globotriose/g CDW/h worden bekomen met een finale titer van 2,8 g/L.

Tenslotte werd ook de combinatie van een GT met sucrose synthase (SuSy) geëvalueerd. SuSy is immers in staat om in één stap UDP-glucose te vormen uit sucrose en vormt zo een aantrekkelijk alternatief voor het hierboven beschreven sucrose fosforylase-gebaseerde glycosylatie platform. Wegens de ongunstige kinetische eigenschappen van SuSy (lage affiniteit voor sucrose) is dit enzym echter zelden gebruikt voor *in vivo* vorming van UDP-glucose. Hiertoe werden verschillende SuSy mutanten (random en rationeel) gecreëerd die met behulp van een nieuw screeningssysteem snel konden worden gekarakteriseerd. Het SuSy van *Solanum tuberosum* werd gebruikt als startpunt voor mutagenese en had een initiële affiniteitsconstante van 105 mM. Deze waarde kon

significant verlaagd worden tot 36 mM door introductie van een S11E mutatie, dewelke fosforylatie van het serine residu nabootst. Co-expressie van deze verbeterde SuSy mutant met het glucosyltransferase VvGT2 in de gemodificeerde *E. coli* W stam, resulteerde in een nieuw *in vivo* (bioconversie) glucosylatieplatform. De gepermeabiliseerde cellen konden efficiënt vanillezuur glucosyleren met een productiesnelheid van 0,41 g/L/h en een RC_{max} van 8. Dit systeem omzeilt de noodzaak om de betrokken enzymen op te zuiveren en is gemakkelijk opschaalbaar.

De belangrijkste uitkomst van dit proefschrift is de creatie van een veelzijdig glycosyleringsplatform met behulp van een geheel nieuwe *metabolic engineering* strategie die groei en productie aan elkaar koppelt. Zo werd de *in vivo* glucosylering, galactosylering en rhamnosylering van verschillende moleculen bekomen, die efficiënter en economisch rendabeler is dan alle bestaande processen tot nu beschreven. Deze unieke fermentatieve benadering, die enkel sucrose als duurzame koolstofbron gebruikt, heeft de weg vrijgemaakt voor een grootschalige en duurzame productie van glycosiden.

Curriculum vitae

Frederik De Bruyn (° April 28th 1987)

Education

2010 – 2014	**PhD student**, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium 'Building a Glycosylation Platform in *E. coli* through Metabolic Engineering' Promotors: Wim Soetaert and Marjan De Mey
2005 – 2010	**Master in Bioscience Engineering**, option Chemistry and Bioprocess Technology, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium Master thesis: 'Ontwikkeling van een *E. coli* mutant voor de productie van zeldzame suikers'
1999 – 2005	**Latin-Mathematics**, St. Jozefscollege Aalst, Aalst, Belgium

Specialized courses

Protein Engineering, 31 October – 4 November 2011, Radboud University Nijmegen, the Netherlands

Biobased Economy Summer School, 19-22 August 2013, Ghent University, Belgium

Technology Transfer Skills, Spring course 2014, Ghent University, Belgium

Bioinformatics for Bio-process Optimization, 29-30 September 2014, Ghent University, Belgium

Transferable skills courses

Communication Skills, 15-16 May 2012, Ghent University, Belgium

Leadership Foundation, 12-19-26 June 2012, Ghent University, Belgium

Impact and Research Communication, 27 September 2012, Ghent University, Belgium

Publications

De Bruyn, F., Beauprez, J., Maertens, J., Soetaert, W., and De Mey, M. (2013) **Unraveling the Leloir Pathway of *Bifidobacterium bifidum*: Significance of the Uridylyltransferases**. Appl. Environ. Microbiol. 79, 7028-7035.

De Bruyn, F., Soetaert, W., and De Mey, M. (2014) **Biotechnological Advances in UDP-sugar based Glycosylation of Small Molecules**. Biotechnology Advances. *Submitted*

De Bruyn, F., De Paepe, B., Beauprez, J., Maertens, J., De Cocker, P., Mincke, S., Stevens, C., Soetaert, W., and De Mey, M. (2014) **Development of an *in vivo* glucosylation platform by coupling production to growth: Production of phenolic glucosides by a GT of *V. vinifera***. Bioresource technology. *Submitted*

De Bruyn, F., Van Brempt, M., Beauprez, J., Maertens, J., Duchi, D., Van Bellegem, W., Soetaert, W., and De Mey, M. (2014) **Building a versatile glycosylation platform: Production of bio-active quercetin glycosides.** *Submitted*

Attended conferences

Bioinformatics: Tools in Research, 28 September 2011, Ghent, Belgium

Metabolic Engineering IX, 3-7 June 2012, Biarritz, France
Poster presentation: **Unravelling the Leloir pathway in *Bifidobacterium bifidum***

CINBIOS Forum for Industrial Biotechnology, 23 October 2012, Mechelen, Belgium

Yeasterday, 24 May 2013, Ghent, Belgium

Knowledge for Growth, 30 May 2013, Ghent, Belgium

Eurocarb17, 7-11 July 2013, Tel-Aviv, Israel

Poster presentation: **Building a glucosylation platform in *E. coli* through Metabolic Engineering**

25th Joint Glycobiology Meeting, 14-16 September 2014, Ghent, Belgium
Poster presentation: **Creating a novel and versatile glycosylation platform in *E. coli***

Student guidance

Practical exercises General Microbiology (2011-2013)

Instructor iGEM Team UGent (2013).
Project: **A new model for chromosomal evolution: eliminating antibiotic resistance**

Tutor in master theses

Maarten Van Hyfte, B.Sc. thesis (2011-2012). **Screening en detectie van sachariden in biologische stalen**

Brecht De Paepe, M.Sc. thesis (2012-2013). **Ontwikkeling en evaluatie van een glucosyleringsplatform in *E. coli***

Jarno Van de Voorde, M.Sc. thesis (2012-2013). **Sucrose Synthase als interessante tool voor Metabolic Engineering**

Pieter De Cocker, M.Sc. thesis (2013-2014). **Optimization and scale up of a glucosylation platform in *E. coli*: coupling growth to production**

Maarten Van Brempt, M.Sc. thesis (2013-2014). **Evaluation and optimization of a galactosylation platform in *E. coli***

Other activities

International Judge at the iGEM Competition for Synthetic Biology, 11-13 October 2013, Lyon, France

Keynote speaker at the Synenergene Workshop on Synthetic Biology, 8-9 April 2014, Brussels, Belgium