Harnessing the Substrate Promiscuity of Enzymes Involved in Natural Product Glycosylation Towards Better Engineered Bioactive Small Molecules

Fathima Ifthiha Mohideen

A Thesis
In the Department of
Biology

Presented in Partial Fulfillment of the Requirements
For the Degree of
Doctor of Philosophy (Biology) at
Concordia University
Montreal, Quebec, Canada

July 2021

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Fathima Ifthiha Mohideen

Entitled: Harnessing the Substrate Promiscuity of Enzymes Involved in Natural Product Glycosylation Towards Better Engineered Bioactive Small Molecules

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Rachael Mansbach	Chair
Dr. Christopher Boddy	External Examiner
Dr. Peter Pawelek	External to Program
Dr. Malcolm Whiteway	Examiner
Dr. Brandon Findlay	Examiner
Dr. David H. Kwan	Thesis Supervisor

Approved by

Dr. Robert Weladji, Graduate Program Director

August 19, 2021

Dr. Pascale Sicotte, Dean of the Faculty of Arts and Science

Abstract

Harnessing the Substrate Promiscuity of Enzymes Involved in Natural Product Glycosylation Towards Better Engineered Bioactive Small Molecules

Fathima Ifthiha Mohideen, Ph.D.

Concordia University, 2021

Many small molecule natural products are decorated with sugar moieties that are essential for their biological activity. A considerable number of natural product glycosides and their derivatives are clinically important therapeutics. The biosynthesis of these natural product glycosides involves the action of sugar biosynthetic enzymes and glycosyltransferases (GTs), where many of these enzymes show relaxed substrate specificities. This can make them valuable biocatalytic tools for altering glycosylation as part of a strategy called "glycodiversification."

The success of glycodiversification greatly depends on the screening method, many of which have limitations with respect to their throughput and ease of use. To address this, we have developed a screening tool for assaying GTs in a high-throughput fashion enabled by rapid isolation and detection of chromophoric or fluorescent glycosylated natural products. Using our novel high-throughput assay, we screened a collection of natural product GTs against a panel of precursors to therapeutically important molecules. A number of these enzymes showed novel acceptor and donor specificities. Interestingly, three GTs catalyzed the synthesis of novel anthracycline glycosides. Our findings are particularly important towards the glycodiversification of therapeutics in this class, given the clinical value of anthracycline glycosides, and the importance of the sugar residue on the biological activity of these therapeutics.

Epirubicin is a high value anticancer anthracycline drug, as it is reported to have a lower cardiotoxicity than the parental compound, doxorubicin. The sugar moiety, L-acosamine and L-daunosamine is the only difference between epirubicin and doxorubicin. The corresponding nucleotide activated sugar donors are TDP-L-acosamine and TDP-L-daunosamine. Despite the importance of TDP-L-acosamine, a natural biosynthetic pathway for this precursor has not been described. Chemists have synthesized this donor via organic synthetic methods, however, it is a tedious process with multiple chemical conversions. Furthermore, the *in vivo* production of TDP-L-daunosamine is reported to be low. We have addressed this by engineering an *in vitro* enzymatic

pathway for the synthesis of TDP-L-acosamine and TDP-L-daunosamine based on the substrate flexibility of sugar biosynthetic enzymes. Starting from a common precursor, we have synthesized these nucleotide deoxysugars and a few intermediates of the enzymatic pathway and used one of the produced intermediates in enzymatic characterization. Moreover, these could be used in *in vitro* glycodiversification and as the starting material for the *in vitro* enzymatic synthesis of some other valuable deoxysugars.

Acknowledgments

I would like to express my sincere gratitude to my supervisor Dr. David H. Kwan for letting me be part of the Kwan lab, for his guidance, and continuous support. Thank you for being a great supervisor. I am thankful to my committee members Dr. Malcolm Whiteway and Dr. Brandon Findlay for their valuable inputs and support.

I am also grateful to Dr. Marcos Di Falco for helping me with performing multiple HPLC-MS runs, Dr. Pham Thi Thanh My for her valuable advices and analyzing NMR data, and Dr. Constantin Yannopoulos for performing NMR analysis.

I would like to thank all the current and past Kwan lab members for their friendship and help. I will always cherish those Kwan lab memories. A special thanks to Mohamed Nasr, Dr. Lan Huong Nguyen, Trisha Ghosh, and Jamaac Ahmed for always helping and encouraging me.

I am extremely thankful to my parents, who set me off on the road to this journey years ago. I am also thankful for my brother and in-laws for their continuous blessings. A big thanks to my loving husband Luke, without whom I would not have made it through my graduate life. Thank you, Luke, for your love and support. Finally, I thank God for his blessings and guidance.

Contribution of Authors

This study was designed by Dr. David H. Kwan.

Chapter 1:

Text from a chapter of Comprehensive Glycoscience (Second Edition) which I co-authored was adapted (Mohideen *et al*., 2021).

Chapter 2:

Text from a chapter of Comprehensive Glycoscience (Second Edition) which I co-authored was adapted (Mohideen *et al*., 2021). *VvGT1* gene had been previously cloned by Christopher Ford (Ford *et al*., 1998).

Chapter 3:

UGT72B1 gene had been previously cloned by Gideon Davies at University of York (Brazier-Hicks *et al*., 2007a). Dr. David H. Kwan contributed to cloning. Dr. David H. Kwan and Nathalia Kravchenko contributed to protein expression and purification. Nathalia Kravchenko performed a few enzyme screening assays.

Chapter 4:

Dr. David H. Kwan, Joel Richard, Dr. Lan Huong Nguyen contributed to cloning, and protein expression and purification. Joel Richard synthesized and purified an initial batch of TDP-4-keto-6-deoxy-D-glucose.

Table of Contents

List of Figures

List of Abbreviations

3,4-DCA	3,4-dichloroaniline
ADP	Adenosine diphosphate
aKG	α-ketoglutarate
Amp^R	Ampicillin resistance
ATP	Adenosine triphosphate
BiG HiT	Biphasic Glycosyltransferase High Throughput
CAZy	Carbohydrate-Active enZymes
CMP	Cytidine monophosphate
Cmp^R	Chloramphenicol resistance
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
FACS	Fluorescence-activated cell sorting
gDNA	Genomic DNA
GDP	Guanosine diphosphate
GSH	Glutathione
GSSG	Glutathione disulfide
GTs	Glycosyltransferases
His tagged	Histidine tagged
HPLC	High performance liquid chromatography
IPTG	Isopropyl β- d-1-thiogalactopyranoside
Kan^R	Kanamycin resistance
kDa	Kilodalton
L-GDH	L-glutamate dehydrogenase
L-Glu	L-glutamate
LB	Luria broth
LC-MS	Liquid chromatography – mass spectrometry
MS	Mass spectrometry
NAD^+	Nicotinamide adenine dinucleotide
NADH	Nicotinamide adenine dinucleotide (reduced)

$NADP^+$	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NDP	Nucleotide diphosphate
Ni-NTA	Nickel-nitrilotriacetic acid
NMP	Nucleotide monophosphate
NMR	Nuclear Magnetic Resonance
NTP	Nucleotide triphosphate
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PKS	Polyketide synthase
PLP	Pyridoxal 5'-phosphate
PPi	Pyrophosphate
PSPG	Putative Secondary Plant Glycosyltransferase
ROS	Reactive oxygen species
SDS	Sodium dodecyl-sulphate
S_N1	Substitution nucleophilic unimolecular
S_N2	Substitution nucleophilic bimolecular
S_Ni	Substitution nucleophilic internal
TCP	2,4,5-trichlorophenol
TDP	Thymidine diphosphate
TKDG	TDP-4-keto-6-deoxy-D-glucose
TLC	Thin layer chromatography
UDP	Uridine diphosphate
UV	Ultraviolet

CHAPTER ONE

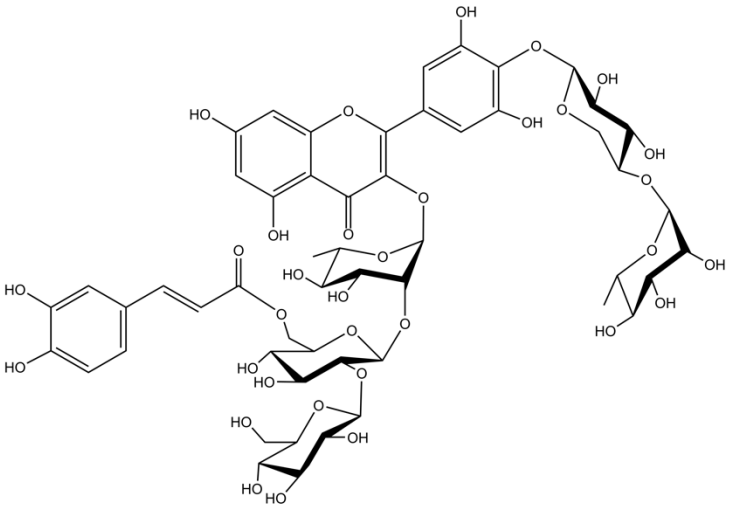
Introduction

1.1 Glycosylated natural products

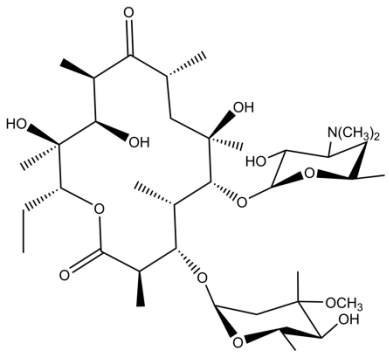
Small molecule natural products and their derivatives originate from a range of organisms, including bacteria, fungi, lichens, insects, marine invertebrates, and plants. These structurally diverse small organic molecules are synthesized through secondary metabolic pathways and production of specific secondary metabolites is limited to specific species categories (Davies and Ryan, 2012). Although natural products aren't vital for the growth and development of the producer organism, many of them support increasing the survival fitness of the producer by acting on certain molecular targets of competitors. For example, some microbes produce specific secondary metabolites that bind and cleave the DNA of the competing microbes, thus inhibiting the growth of rival organisms (Williams *et al*., 1989).

Since ancient times, humans have exploited these natural products for various purposes such as pigments, fragrances, poisons, and drugs (Davies and Ryan, 2012). In the 19th century, the synthesis of aspirin (acetylsalicylic acid) from salicylic acid obtained from the willow tree was a significant landmark for the use of natural products as therapeutics (Mahdi *et al*., 2006).

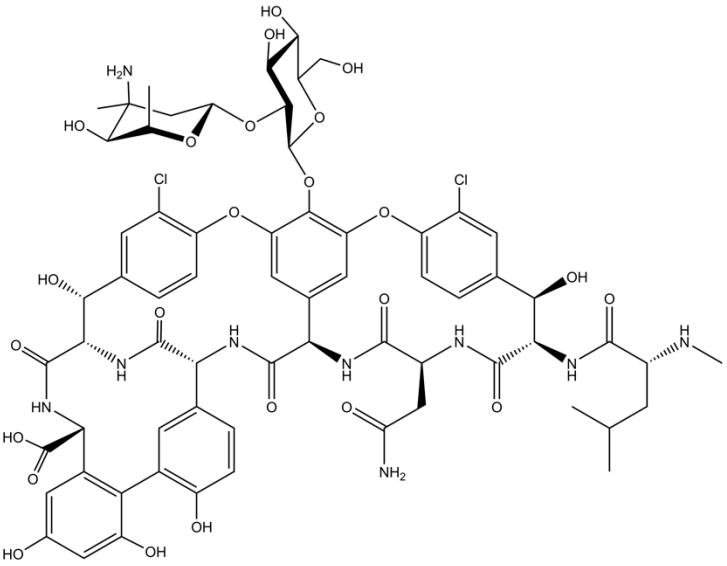
Many of these small molecule natural products are decorated with one or more sugar moieties that have a great effect on their biological activity (Salas and Méndez, 2007; Weymouth-Wilson, 1997). Examples include montbretin A, a potential anti-diabetic (Asada *et al*., 1988), antibiotics such as erythromycin (Staunton and Wilkinson, 1997) and vancomycin (Zmijewski and Briggs, 1989), and anticancer agents such as daunorubicin and its derivative doxorubicin (Di Marco *et al*., 1975) (**Figure 1**). These structurally diverse glycosylated molecules are biosynthesized through the orchestration of a cascade of enzymes that can be broadly categorized as enzymes responsible for the biosynthesis of sugar residues (sugar biosynthetic enzymes) and enzymes that transfer these sugar residues to accepting molecules (glycosyltransferases). Knowing the significance of sugar moieties on secondary metabolites and a comprehensive awareness of sugar biosynthetic machineries are necessary to harness them towards the production of better glycosides.



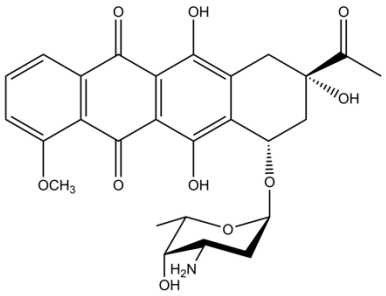
Montbretin A



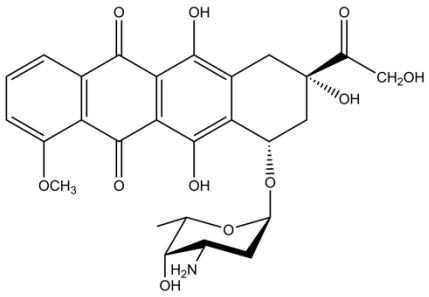
Erythromycin A



Vancomycin



Daunorubicin



Doxorubicin

Figure 1: Chemical structures of a few glycosylated natural products with therapeutic value.

1.2 Glycosyltransferases (GTs)

Glycosyltransferases (GTs) are enzymes that catalyze the assembly of glycans and glycoconjugates and play many diverse roles in biology. GTs catalyze the transfer of sugar residues from activated sugar donor substrates to acceptor molecules, forming glycosidic bonds (Campbell *et al*., 1997). The acceptor nucleophiles can be either other sugars or non-sugar aglycones such as proteins, lipids, nucleic acids, or small molecules (*e.g*., antibiotics and other secondary metabolites) (Coutinho *et al*., 2003; Lairson *et al*., 2008). Although the most common glycosylated products are *O*-linked, *N*-, *C*-, and *S*-linked are also produced (Lairson *et al*., 2008). Considering the biological significance of glycans, an understanding of the enzyme structure and mechanism allows researchers to use them towards desired applications.

1.2.1 Leloir versus non-Leloir GTs and their donor substrates

Activated sugar donors used by GTs can be in the form of nucleoside diphosphosugars (*e.g*., UDP-glucose, UDP-galactose, and GDP-mannose), nucleoside monophosphosugars (e.g., CMP-*N*-acetyl-neuraminic acid), lipid phosphosugars (*e.g*., dolichol phosphomannose, decaprenol phosphoarabinose), lipid diphosphosugars (*e.g*., dolichyl-diphosphooligosaccharide), sugar-1-phosphates (*e.g*., glucose-1-phosphate), or sugar-1-diphosphates (*e.g*., phosphoribosyl pyrophosphate) (**Figure 2**) (Albuquerque-Wendt *et al*., 2019; Berg *et al*., 2007; Lairson *et al*., 2008). Enzymes that use nucleotide-activated sugar donors are the most prevalent, consisting of approximately 65% of known GTs (Ardèvol and Rovira, 2015) and are termed Leloir GTs after Nobel laureate, Luis F. Leloir, who was the first to discover a nucleotide-sugar (Cardicini *et al*., 1950). Conversely GTs that use other sugar donors are defined as non-Leloir GTs, and those that use sugar-1-phosphates and sugar-1-diphosphates also categorized as phosphorylases and pyrophosphorylases, respectively (Kornberg *et al*., 1955; Lairson *et al*., 2008; Lieberman *et al*., 1955). An overview on sugar donors of natural product GTs is presented in **Section 1.3**.

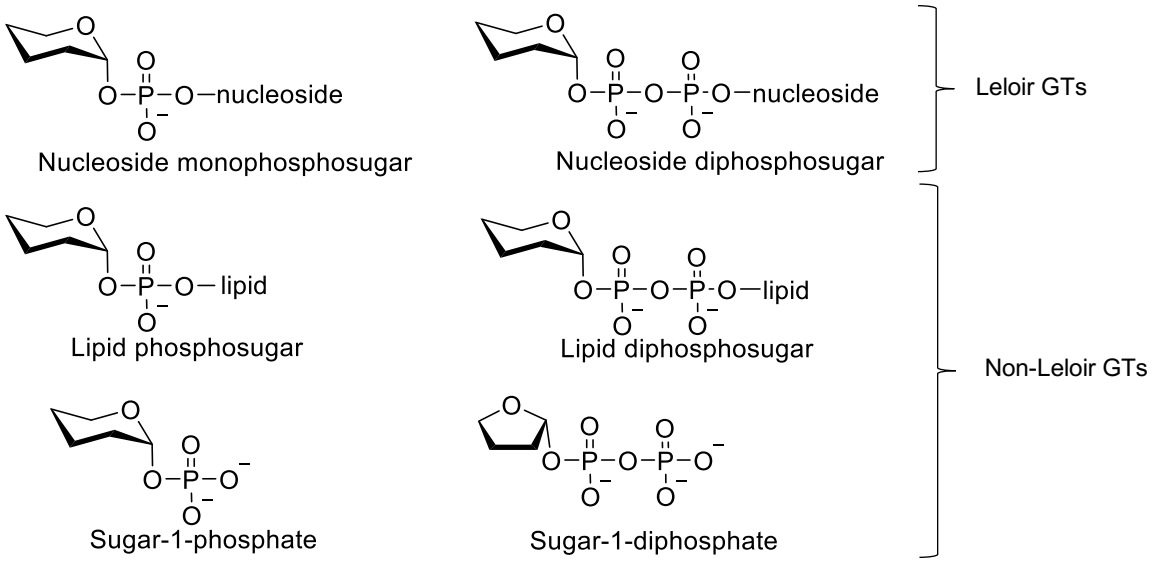


Figure 2: Sugar donors utilized by Leloir and non-Leloir GTs. Leloir GTs use nucleotide activated donors and non-Leloir GTs use other donors.

1.2.2 Sequence-based CAZy families and structural categorization of GTs

In 1997, Henrissat and colleagues classified GTs into families based on the similarities of their amino acid sequences (Campbell *et al*., 1997). As of 2021, > 800,000 GTs have been classified into > 110 GT families that are regularly updated in the Carbohydrate-Active enZymes (CAZy) database (www.cazy.org). Additionally, > 21,000 non-classified sequences are also reported. Of these families, only families GT2 and GT4 are represented in primitive archaea, suggesting that these might be the ancestral families from all other GT families evolved (Coutinho *et al*., 2003).

Despite the great number of GT families, the protein topology is limited to only a few structurally different folds, which in addition to the GT-A, GT-B, and GT-C folds includes lysozyme-type, and 5-bladed propeller folds (Breton *et al*., 2006; Lovering *et al*., 2007; Sernee *et al*., 2019). Structural analysis of Leloir GTs has revealed that GT-A and GT-B folds possess Rossmann-like folds. The two β/α/β Rossmann-like domains are compactly associated in GT-A fold enzymes, and thus often referred as a single Rossmann-like domain (Breton *et al*., 2006; Coutinho *et al*., 2003; Lairson *et al*., 2008). Generally, the catalytic activity of GT-A enzymes requires a divalent metal cation, such as Mg^{2+} or Mn^{2+}, which is coordinated along with the nucleotide-activated donor by a DXD (Asp-Xaa-Asp) amino acid sequence motif (Breton *et al*.,

2006; Ünligil and Rini, 2000). However, some GT-A enzymes do not rely on a metal ion for activity, and not all GT-A enzymes contain the DXD motif, nor is this motif diagnostic for GTs (Coutinho *et al*., 2003; Lairson *et al*., 2008).

The two Rossmann-like domains of GT-B fold enzymes face one another and are less compactly linked. The catalytic activity of these enzymes is metal independent, thus do not require a DXD motif (Coutinho *et al*., 2003; Lairson *et al*., 2008). Nevertheless, both GT-A and GT-B folds contain separate acceptor and nucleotide binding domains and these two folds have been observed to have either inverting or retaining mechanisms (**Section 1.2.3**) (Coutinho *et al*., 2003).

Based on sequence searches, Liu and Mushegian proposed a third GT-fold, called GT-C, that has multiple transmembrane α-helices (Liu and Mushegian, 2003). GTs that use lipid phosphate- and lipid diphosphate-activated sugars possess this fold and so far, all known GT-C enzymes operate with an inverting mechanism (**Section 1.2.3**) (Lairson *et al*., 2008; Moremen and Haltiwanger, 2019). Considering these structural folds and the stereochemistry of glycosylated products, GT families can be categorized into clans (**Figure 3**).

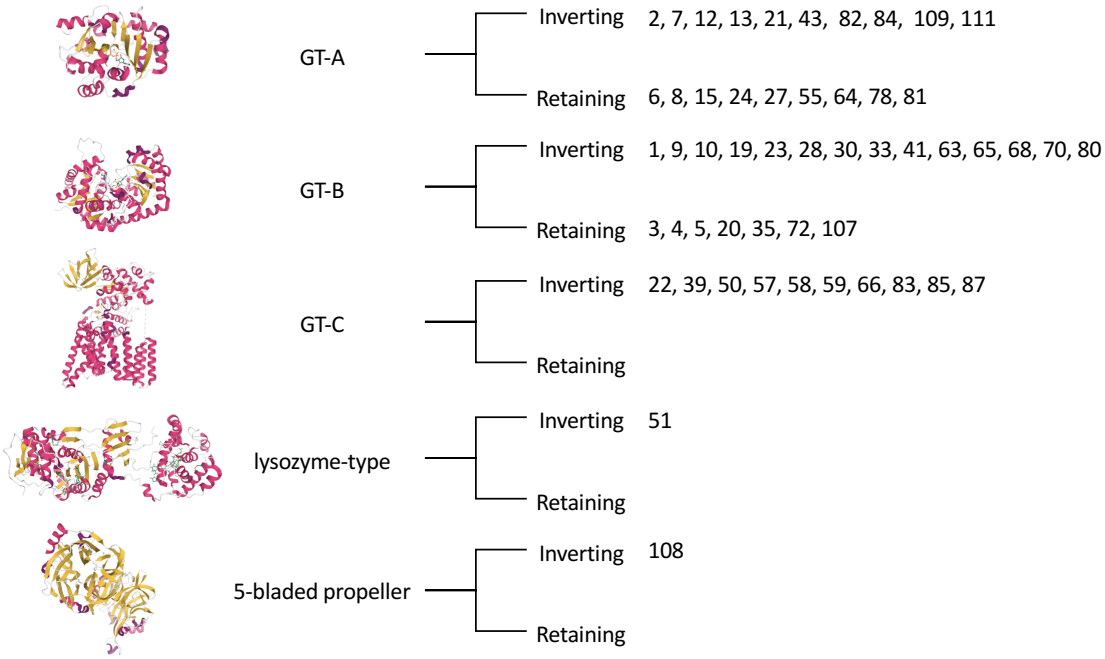


Figure 3: GT families are divided into clans based on fold and reaction mechanism (data from the CAZy database). GT-A, GT-B, GT-C, lysozyme-type, and 5-bladed propeller folds are represented by protein structures of LgtC from *Neisseria meningitidis* (PDB: 1G9R), VvGT1 from *Vitis vinifera* (PDB: 2C1X), and PglB from *Campylobacter lari* (PDB: 3RCE), Pbp2 from *Staphylococcus aureus* (PDB:2OLV), and MTP2 from *Leishmania mexicana* (PDB: 6Q4X), respectively.

1.2.3 Mechanism of GTs

The glycoside products of GT action can have the same (retained) or opposite (inverted) anomeric configuration, relative to the sugar donor. Based on the product stereochemistry, GTs are classified as inverting or retaining (**Figure 4**). The difference in the stereochemical outcome is related to the reaction mechanism (Lairson *et al*., 2008).

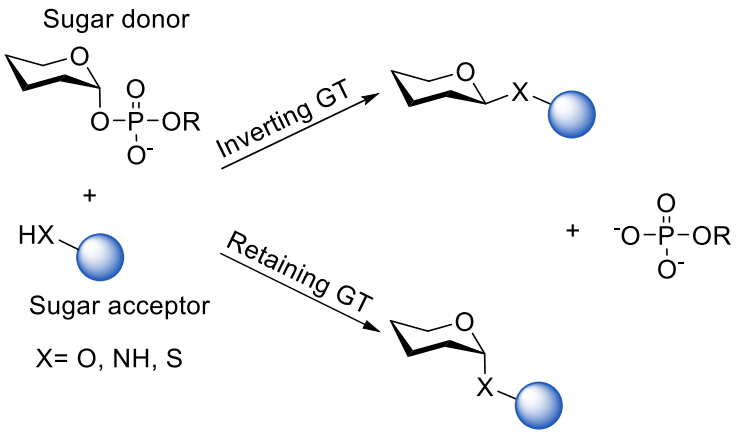


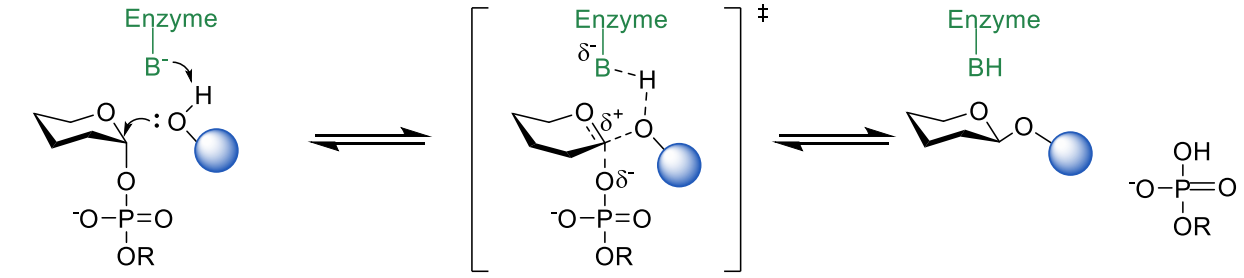
Figure 4: The two stereochemical products of GT catalyzed reactions. With respect to the sugar donor, an inverting GT gives a product with inversion of the anomeric configuration while the anomeric configuration is retained by a retaining GT.

1.2.3.1 Inverting GT mechanisms

Most inverting GTs use a single displacement S_N2 (substitution nucleophilic bimolecular) mechanism through an oxocarbenium ion-like transition state that results in inversion of configuration at the anomeric carbon of the sugar donor (**Figure 5A**). A side chain of an active site aspartic acid or glutamic acid assists as a general base catalyst, deprotonating the nucleophilic hydroxyl group of the acceptor molecule. At the same time, the nucleophile attacks the anomeric carbon of the donor and the nucleotide moiety leaves from the opposite side. Most inverting GT-A enzymes use a divalent metal ion to stabilize the negative charge of the phosphate leaving group, and in metal-independent inverting GT-B enzymes a positively-charged amino acid side-chain performs this function (Lairson *et al*., 2008; Rini and Esko, 2017).

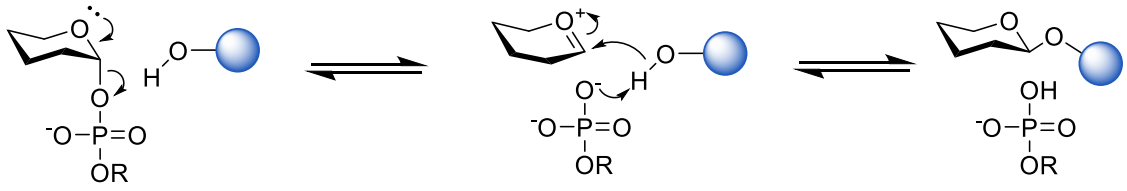
Not all inverting GTs contain an amino acid general base. In some cases, the β-phosphate oxygen atom of the donor molecule is proposed to fulfil this function. Examples include POFUT1 and FUT1, which are proposed to use an S_N1 (substitution nucleophilic unimolecular) reaction mechanism as illustrated by **Figure 5B**. Herein, glycosidic bond cleavage would result in a transient oxocarbenium ion. The close association of the β-phosphate group of the donor with the oxocarbenium ion leads to the attack of the incoming acceptor from the opposite direction (Li *et al*., 2017; Lira-Navarrete *et al*., 2011; Urbanowicz *et al*., 2017)

A



S_N2 (substitution nucleophilic bimolecular) mechanism

B



S_N1 (substitution nucleophilic unimolecular) mechanism

Figure 5: Inverting GT reaction mechanisms. (A) S_N2 reaction mechanism, (B) S_N1 mechanism.

1.2.3.2 Retaining GT mechanisms

Through detailed structural and kinetic studies, compelling evidence of retaining glycoside hydrolases supports a double-displacement substitution reaction involving a covalently linked glycosyl-enzyme intermediate (Zechel and Withers, 2000). The same mechanism has been suggested for retaining GTs (Lairson *et al*., 2008; Rini and Esko, 2017). As shown by **Figure 6A**, this proposed two-step double-displacement mechanism (going through two transition states with oxocarbenium character) requires a properly situated active site nucleophile to attack the anomeric carbon of the donor molecule (Lairson *et al*., 2008). Structural studies of several GTs have suggested that the leaving phosphate group may act as the base catalyst to activate the incoming acceptor nucleophile (Pedersen *et al*., 2003; Persson *et al*., 2001). Similar to inverting GTs, a divalent cation or a positively charged side chain assists the removal of the leaving group in many retaining GT-A or GT-B enzymes, respectively (Lairson *et al*., 2008).

In 2004, Withers and colleagues observed the first covalent intermediate of a mutant retaining GT, but were unable to demonstrate a covalent intermediate for the wildtype enzyme (Lairson *et al*., 2004). As of 2021, confirmation of a covalent glycosyl-enzyme intermediate has not been obtained for any wild-type retaining GT. Moreover, the relative location of a catalytic enzyme nucleophile seems poorly conserved among many retaining GTs suggesting an alternative mechanism for at least some of these enzymes (Lairson *et al*., 2008; Moremen and Haltiwanger, 2019).

An S_Ni (substitution nucleophilic internal)-like reaction with an oxocarbenium ion-like transition state has been proposed as for some retaining GTs. In this mechanism, the leaving group deprotonates the nucleophilic group of the acceptor and the retention of the stereochemistry is an outcome of the same side departure of the leaving group and attack by the nucleophile (**Figure 6B**). It is proposed that the front-face attack by the acceptor nucleophile is supported by hydrogen bonds between acceptor nucleophile and the leaving group, as first proposed for UDP-Glc dependent trehalose-6-phosphate synthase from *E. coli* (Lairson *et al*., 2008; Lee *et al*., 2011; Sinnott and Jencks, 1980). This mechanism has since been assigned to retaining GTs including human polypeptide N-acetylgalactosaminyltransferase 2 (GalNAc-T2) (Lira-Navarrete *et al*., 2014), mammalian xyloside xylosyltransferase 1 (XXYLT1) (Yu *et al*., 2015), and mycobacterial

glucosyl-3-phosphoglycerate synthase (GpgS) (Albesa-Jové *et al.*, 2015) by crystallographic, molecular dynamic, and mechanistic studies.

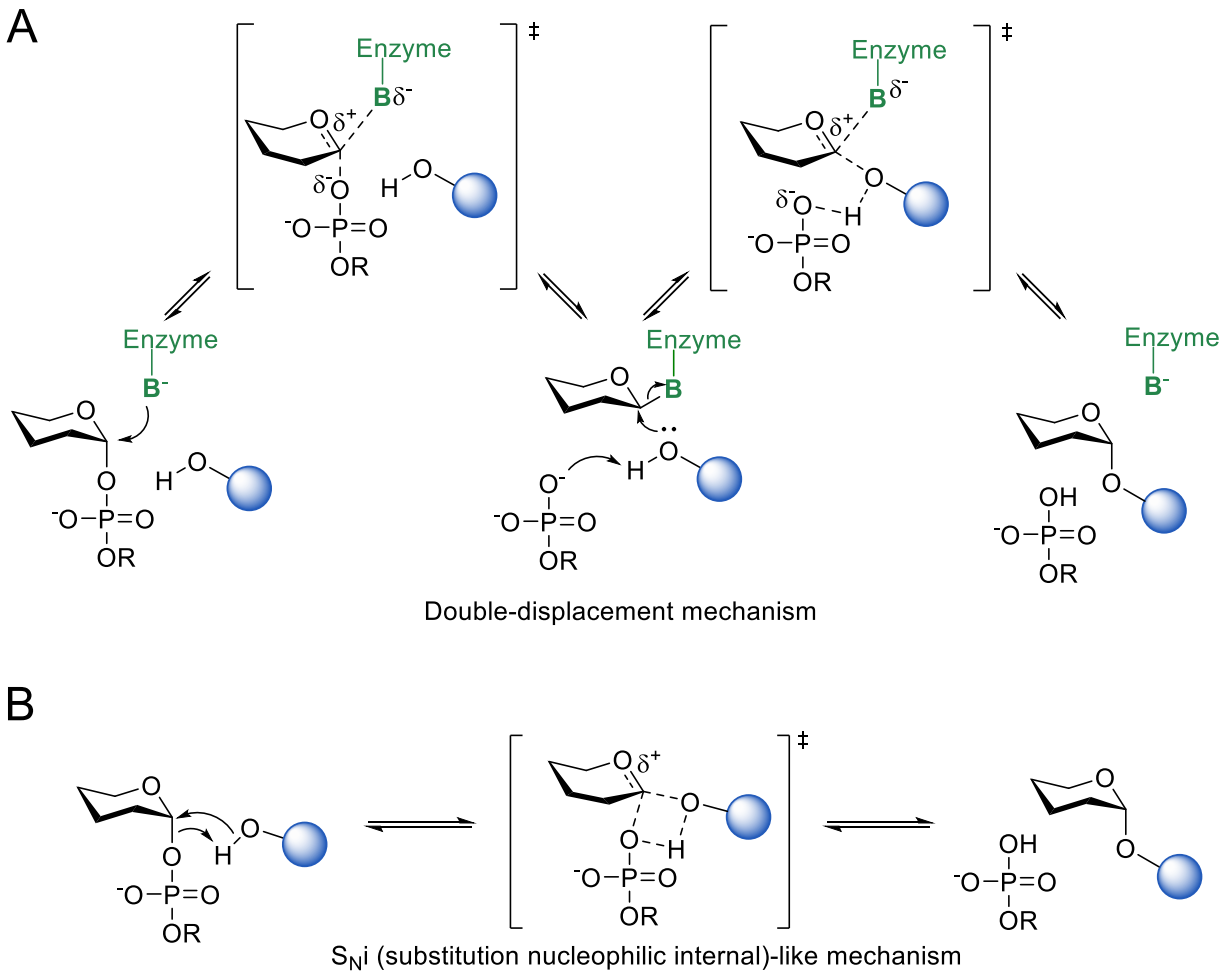


Figure 6: Retaining GT reaction mechanisms. The proposed (A) double-displacement reaction mechanism with a covalently linked enzyme-glycosyl intermediate (and two oxocarbenium ion-like transitions states) and (B) S_Ni-type mechanism with an oxocarbenium ion-like transition state; whether this reaction occurs in a concerted or non-concerted manner is an open question.

1.3 Sugar precursors of natural product glycosides

Natural product GTs use sugar donors that are nucleotide activated and many of them are derived from glucose-6-phosphate or fructose-6-phosphate through a series of enzymatic reactions catalyzed by sugar biosynthetic enzymes (Thibodeaux *et al*., 2007). A restricted number of nucleotide activated sugar donors are utilized by eukaryotic natural product GTs. For example, many plant natural product GTs typically use UDP-D-glucose, UDP-D-galactose, and UDP-D-xylose (**Figure 7A**) (Ross *et al*., 2001). In contrast, glycosylated bacterial natural products can include sugar units drawn from a large pool of highly diverse and unusual sugar moieties (Thibodeaux *et al*., 2007; Thibodeaux *et al*., 2008).

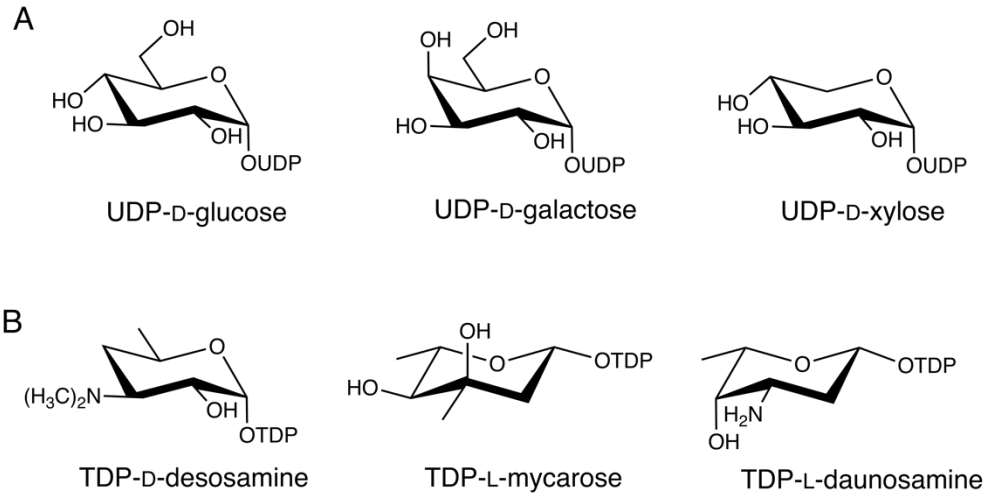


Figure 7: Examples of a few sugar precursors used by (A) plant natural product GTs and (B) bacterial natural product GTs

Most of the sugar precursors of prokaryotes are heavily modified and deoxygenated, thus would present a vast collection of glycosylated natural products. Many of them are thymidine diphosphate (TDP) activated and are deoxygenated at the C-6 position. Apart from being 6-deoxyhexoses, a considerable fraction of these are di or tri-deoxyhexoses, where deoxygenation at C-2, C-3, or C-4 is also observed (Thibodeaux *et al*., 2008). Some of the naturally occurring bacterial deoxysugar donors include TDP-D-desosamine and TDP-L-mycarose: the sugar precursors of several macrolide antibiotics (Summers *et al*., 1997), and TDP-L-daunosamine: the sugar donor of anticancer anthracycline daunorubicin and doxorubicin (Lomovskaya *et al*., 1999) (**Figure 7B**).

1.3.1 Unusual sugar donor biosynthesis

A significant characteristic of sugar precursors involved in natural product glycosylation is the activation of monosaccharides as nucleotide monophosphate (NMP) or diphosphate (NDP) derivatives. This is essential as the nucleotide appendage functions as an identification unit for sugar biosynthetic enzymes, and during the GT reaction it serves as a good leaving group (Thibodeaux *et al*., 2008). As illustrated in **Figure 8**, nucleotide activation is catalyzed through a nucleotidylyltransferase reaction using sugar-1-phosphate and a relevant nucleotide triphosphate (NTP) (Barton *et al*., 2001). The required sugar-1-phosphate is generally produced by a phosphohexose mutase from sugar-6-phosphate (e.g., glucose-6-phosphate or fructose-6-phosphate). However, the conversion of a hexose into its corresponding sugar-1-phosphate via anomeric hexose kinase is also observed (Thibodeaux *et al*., 2007; Thibodeaux *et al*., 2008).

Subsequently, the activated NDP-sugar is 4,6-dehydrogenated by NDP-hexose-4,6-dehydratase to produce NDP-4-keto-6-deoxy-D-hexose, the common precursor of most 6-deoxyhexoses. Starting from this immediate precursor, a multitude of different enzymatic reaction combinations including ketoreduction, deoxygenation, dehydration, transamination, methylation, and epimerization results in a diverse collection of unusual sugar donors. Remarkably, sugar biosynthetic enzymes that catalyze similar enzymatic reactions in different unusual sugar biosynthetic pathways are shown to have a high amino acid sequence similarity (Thibodeaux *et al*., 2007). A biosynthetic pathway of a deoxysugar involved in anticancer anthracycline glycosylation is discussed in **Section 1.4.2.3**.

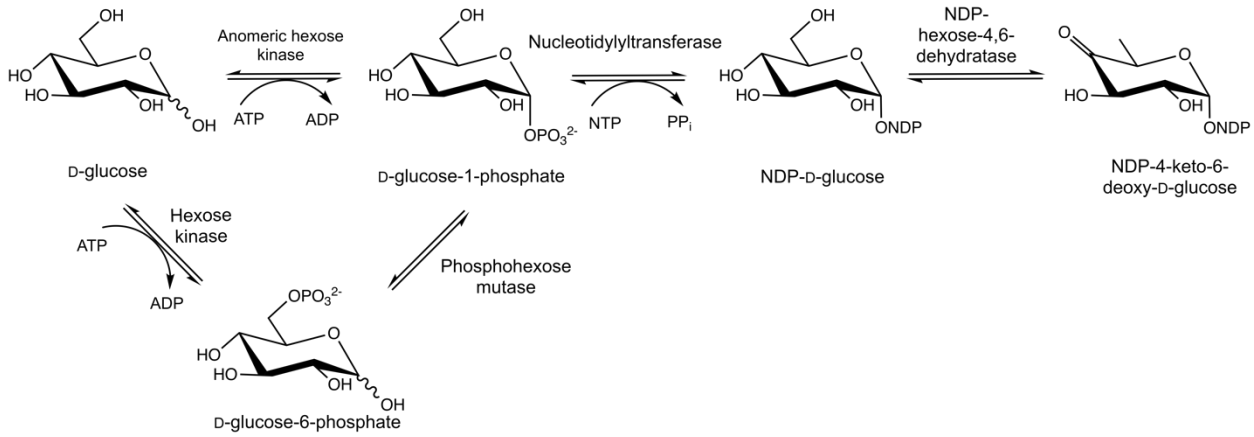


Figure 8: Biosynthesis of the common sugar intermediate NDP-4-keto-6-deoxy-D-glucose towards the production of diverse deoxysugars.

1.4 Anticancer anthracyclines

Anthracyclines are a class of glycosylated microbial natural products that possess anticancer properties. Two clinically important anthracyclines include daunorubicin and doxorubicin. These were the first discovered anthracyclines from the pigment producing soil-dwelling actinobacterium *Streptomyces peucetius* (Di Marco *et al*., 1975). The chemical structure of anthracyclines includes two components, a tetracyclic polyketide aglycone and a sugar residue. As shown in **Figure 9**, ring C and B of the aglycone contain neighboring quinone-hydroquinone functional groups and the sugar portion is attached to ring A through an *O*-glycosidic bond at C-7 (Minotti *et al*., 2004).

The sugar moiety of doxorubicin and daunorubicin is identical, whereas the structural difference is observed in the aglycone moiety. The extra hydroxyl group of doxorubicin at the C-14 of the side chain attached to ring A shows a significant consequence on its range of anticancer activity. Doxorubicin is used for treating a number of solid and blood cancers such as breast, lung, thyroid, gastric and ovarian cancer, soft tissue sarcomas, and non- Hodgkin's and Hodgkin's lymphoma. The use of daunorubicin is mainly limited to acute lymphoblastic leukemia and myeloblastic leukemia (Minotti *et al*., 2004; Weiss *et al*., 1986). The success of anthracyclines as anticancer agents depends on its ability to kill proliferating cancer cells.

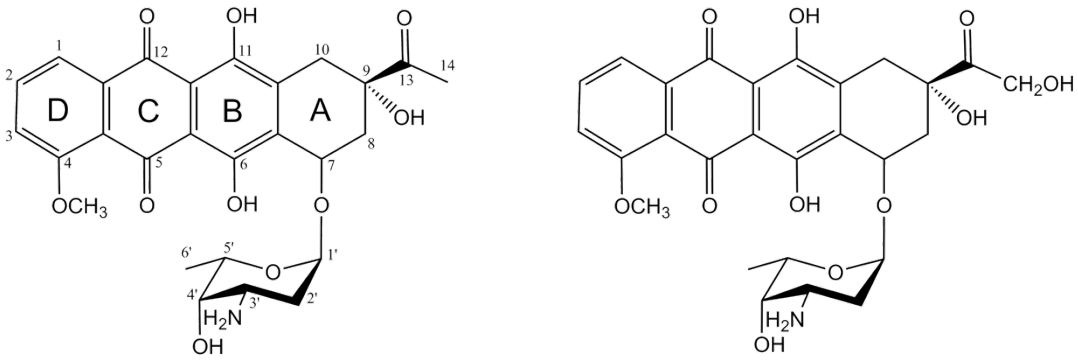


Figure 9: Chemical Structures of the anticancer anthracycline daunorubicin and doxorubicin. Doxorubicin contains an extra hydroxyl group at C-14 position.

1.4.1 Molecular mechanism of anticancer anthracyclines

A number of studies highlight the molecular basis of the anticancer activity of anthracyclines as topoisomerase II poisons.

Topoisomerases change the DNA topology by making temporary cleavages of DNA strands, alter the twisting of the DNA double helix, and re-ligate the cut sites. Depending on the structure and mechanism, there are two types of human topoisomerases, topoisomerase I and topoisomerase II, that cause single-strand and double-strand DNA breaks, respectively (Chen *et al*., 2013). Anthracyclines are reported as effective DNA topoisomerase II poisons. Upon interaction with DNA bases and amino acid residues of topoisomerase II active sites, anthracyclines stabilize a DNA-topoisomerase II complex, where the DNA strands are cleaved and covalently associated with the enzyme. This will inhibit the resealing activity of the enzyme and ultimately results in apoptosis (Capranico *et al*., 1997; Moro *et al*., 2004).

1.4.2 Daunorubicin and doxorubicin biosynthesis

Daunorubicin and its C-14 hydroxylated derivative doxorubicin are produced in nature by Gram-positive actinomycete *S. peucetius*. Including the wild-type *S. peucetius* ATCC 29050, the biosynthesis of daunorubicin is observed in several *Streptomyces* species (Grein, 1987). Initial studies stated that the production of doxorubicin was exclusive in the mutant *S. peucetius* subsp. *caesius* ATCC 27952. However, the wild-type *S. peucetius* was also reported to produce doxorubicin when an optimum medium was used (Lomovskaya *et al*., 1999). Regardless of the clinical importance of doxorubicin, in nature it is produced in minimal amounts (Wang *et al*., 2018b). Thus, unraveling the biosynthetic pathway of this anthracycline is extremely important to enhance the yield to reach the industrial necessity.

1.4.2.1 The biosynthetic gene cluster of daunorubicin/doxorubicin

Considering the therapeutic importance of daunorubicin and doxorubicin, thorough studies on the *S. peucetius* gene cluster were performed. In 1999, Hutchinson and colleagues published the *S. peucetius* ATCC 29050 daunorubicin/ doxorubicin gene cluster (Lomovskaya *et al*., 1999). Further genomic sequence analysis of the *S. peucetius* ATCC 27952 variant disclosed 37 open-reading frames in the biosynthetic gene cluster of daunorubicin/ doxorubicin that spans around 40 kb of the 8.1 Mb total genome (Niraula *et al*., 2010). As shown in **Figure 10**, the gene cluster contains structural genes encoding for the biosynthesis of the anthracycline precursors (polyketide ε-rhodomycinone and the deoxysugar TDP-L-daunosamine), genes encoding for glycosylation and post-modifications, regulatory genes, and resistance and efflux genes (Lomovskaya *et al*., 1999; Niraula *et al*., 2010). The biosynthesis of daunorubicin and doxorubicin follows three steps: production of ε-rhodomycinone aglycone, production of TDP-L-daunosamine sugar donor, and glycosylation of ε-rhodomycinone and post-glycosylation modifications (Hutchinson and Colombo, 1999).

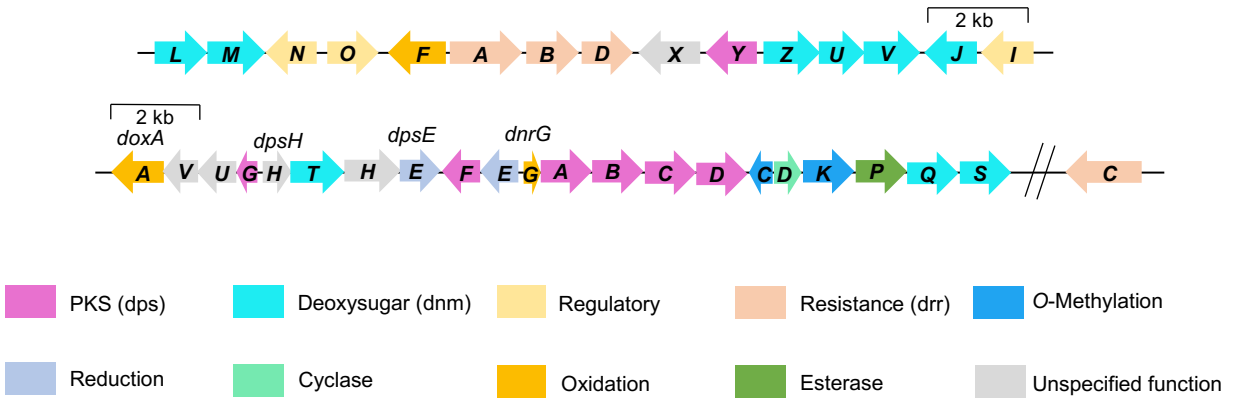


Figure 10: Daunorubicin/ doxorubicin biosynthetic gene cluster. The pointed boxes are a representation of the relative sizes of the open reading frames and gene transcription direction.

1.4.2.2 Biosynthesis of the aglycone precursor: ε-rhodomycinone

The tetracyclic ε-rhodomycinone is the aglycone precursor of daunorubicin and doxorubicin. The biosynthesis of this aglycone precursor begin from one propionyl-CoA starter, and nine malonyl-CoA extenders towards the production of the 12-deoxyaklanonic acid intermediate (**Figure 11A**). *dpsABCDGEFY* genes of the daunorubicin/ doxorubicin gene cluster (**Figure 10**) encodes a type II polyketide synthase (PKS) that is involved in synthesizing this intermediate (Grimm *et al*., 1994; Ye *et al*., 1994). Herein, the successive transfer of acetyl units from malonyl-CoA to propionyl-CoA towards the production of a 21-carbon decaketide is catalyzed by DpsA 3-oxoacyl ACP synthase, DpsB and DpsC ketosynthases (Bao *et al*., 1999; Grimm *et al*., 1994), DpsD acyltransferase, and DpsG acyl carrier protein (Lomovskaya *et al*., 1999). The resulting decaketide is further modified by DpsE ketoreductase, and DpsF and DpsY cyclases to yield 12-deoxyaklanonic acid (Grimm *et al*., 1994; Lomovskaya *et al*., 1998).

The synthesis of ε-rhodomycinone from a 12-deoxyaklanonic acid intermediate involves five sequential enzymatic reactions. As shown in **Figure 11A**, these include keto group addition, methylation, intramolecular aldol addition, reduction, and hydroxylation steps catalyzed by DnrG monooxygenase, DnrC methyltransferase, DnrD cyclase, DnrE aklaviketone reductase, and DnrF aklavinone-11-hydroxylase, respectively (Filippini *et al*., 1995; Grimm *et al*., 1994; Madduri and Hutchinson, 1995).

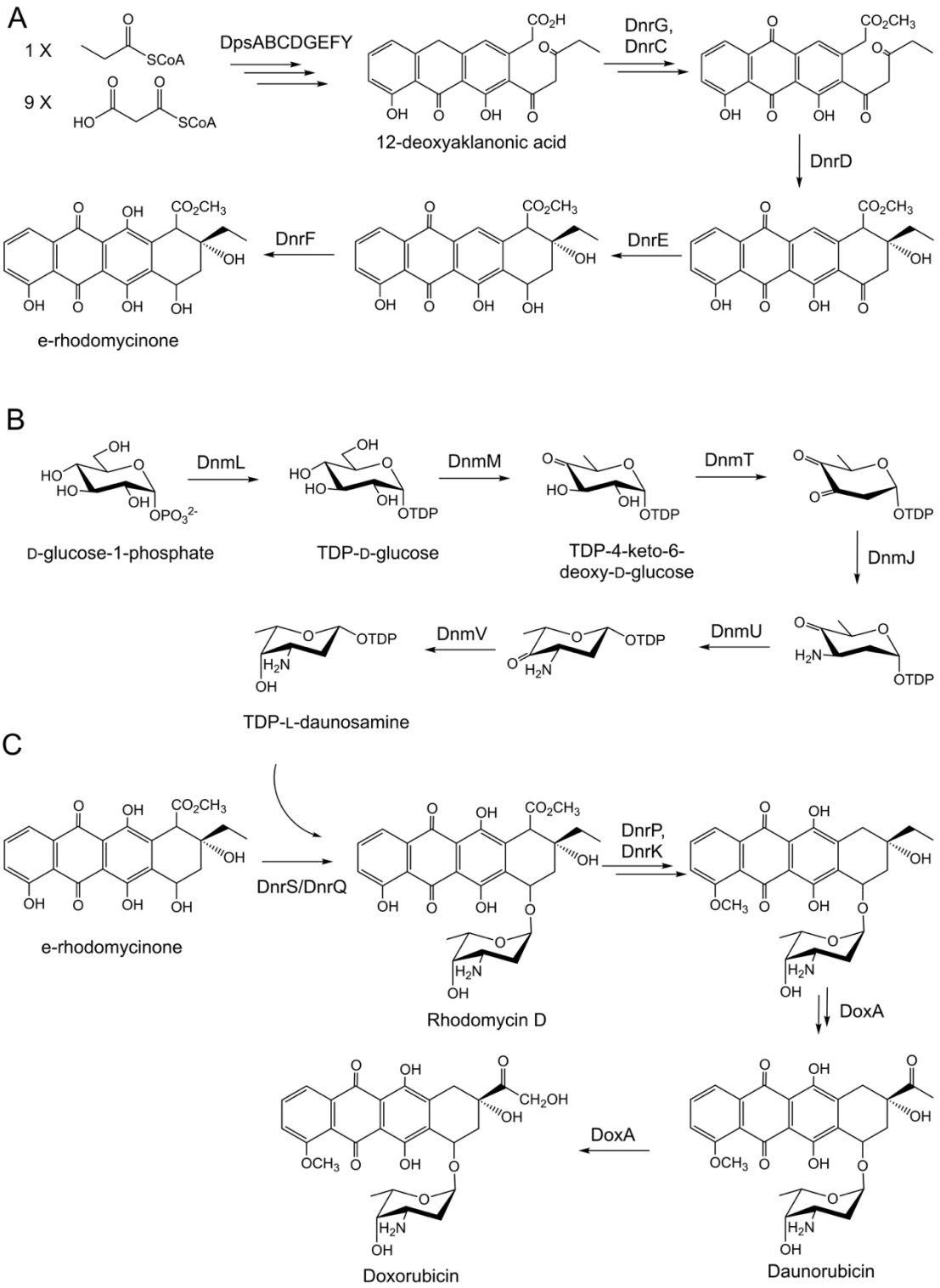


Figure 11: Biosynthetic pathway of daunorubicin and doxorubicin. (A) Synthesis of ε-rhodomycinone aglycone from 1 propionyl-CoA and 9 malonyl-CoA. (B) Synthesis of TDP-L-daunosamine and (C) glycosylation and post modification reactions towards producing daunorubicin and doxorubicin.

1.4.2.3 Biosynthesis of the sugar precursor: TDP-L-daunosamine

The sugar constituent of daunorubicin and doxorubicin is daunosamine. It is a heavily modified 2,3,6-trideoxy-3-aminohexose and the glycosyl precursor of daunosamine is TDP activated. Gene products of *dnmL*, *dnmM*, *dnmT*, *dnmJ*, *dnmU*, and *dnmV* of the *S. peucetius* daunorubicin/ doxorubicin gene cluster are involved in the production of TDP-L-daunosamine (**Figure 10**). The activation step initiates from the addition of thymidine monophosphate (TMP) moiety to D-glucose-1-phosphate catalyzed by DnmL glucose-1-phosphate thymidylyl transferase (**Figure 11B**). Then, by the action of DnmM TDP-D-glucose 4,6-dehydratase, TDP-D-glucose is converted to the key sugar donor intermediate, TDP-4-keto-6-deoxy-D-glucose (TKDG) (Gallo *et al*., 1996). Subsequently, through four enzymatic reactions this key intermediate is further modified to produce TDP-L-daunosamine. This conversion involves C-2,3 dehydration, C-3 amino group addition, C-3,5 epimerization, and C-4 reduction catalyzed by DnmT 2,3-dehydratase, DnmJ 3-aminotransferase, DnmU 3,5-epimerase, and DnmV 4-ketoreductase, respectively (Otten *et al*., 1997).

1.4.2.4 Glycosylation and post modifications to produce daunorubicin and doxorubicin

Following the synthesis of the ε-rhodomycinone aglycone precursor and TDP-L-daunosamine sugar donor, DnrS, which functions as a GT, transfers the daunosamine moiety to the C-7 position of ε-rhodomycinone with the assistance of DnrQ auxiliary protein. This results in the production of rhodomycin D, the first glycosylated intermediate of daunorubicin/doxorubicin biosynthetic pathway (Otten *et al*., 1995) (**Figure 11C**). This glycosylated intermediate undergoes two sequential modifications to produce 13-deoxydaunorubicin through the activity of DnrP esterase and DnrK methyltransferase, respectively (Dickens *et al*., 1997). Subsequently, DoxA monooxygenase converts 13-deoxydaunorubicin to daunorubicin through C-13 oxidation in two enzymatic steps. Hydroxylation of daunorubicin at C-14 by DoxA monooxygenase synthesizes the high value daunorubicin derivative, doxorubicin (Walczak *et al*., 1999).

1.4.3 Need for a better anthracycline analog

Prolonged administration of clinically important doxorubicin and daunorubicin anthracyclines results in dose-dependent cardiotoxicity observed as chronic cardiomyopathy and congestive heart failure (Lahtinen *et al*., 1991; Weiss *et al*., 1986). A few mechanisms have been proposed to explain the anthracycline cardiotoxicity, where oxidative stress caused by doxorubicin is clearly described.

1.4.3.1 Cardiomyocyte damage by oxidative stress

Doxorubicin is reported to build-up efficiently in the mitochondria of cardiomyocytes (Sarvazyan, 1996). Previous studies proposed that the oxidative stress caused by reactive oxygen species (ROS) as a result of redox cycling of the quinone moiety of anthracyclines leads to cardiotoxicity. As illustrated in **Figure 12**, in the presence of an NAD(P)H- dependent enzymes such as NADH dehydrogenase or NADPH-dependent glutathione reductase, the quinone moiety is reduced to semiquinone. Subsequent conversion of semiquinone to quinone produces superoxide anion ($\cdot O_2^-$), that is converted to H_2O_2 by the activity of superoxide dismutase 2. Then, H_2O_2 is broken into H_2O and O_2 by the activity of catalase and glutathione peroxidase (Jungsuwadee, 2016).

Generally, the level of detoxification of H_2O_2 caused in cardiomyocytes is considered low due to the reduced catalase concentration (Jungsuwadee, 2016), and inhibition of catalase and glutathione peroxidase by doxorubicin (Mukhopadhyay *et al*., 2009). H_2O_2 is also converted to hydroxyl radical ($\cdot OH$) by endogenous Fe^{2+} ions leading to membrane lipid oxidization producing lipid radicals. The interaction of $\cdot OH$ with membrane lipids and mitochondrial proteins, and lipid radical interaction with proteins damages the cardiac tissue (Jungsuwadee, 2016). Considering the toxicity of these anthracyclines, it is necessary to have better anticancer anthracycline analogs with reduced cardiotoxicity and a similar or improved antitumor activity to doxorubicin.

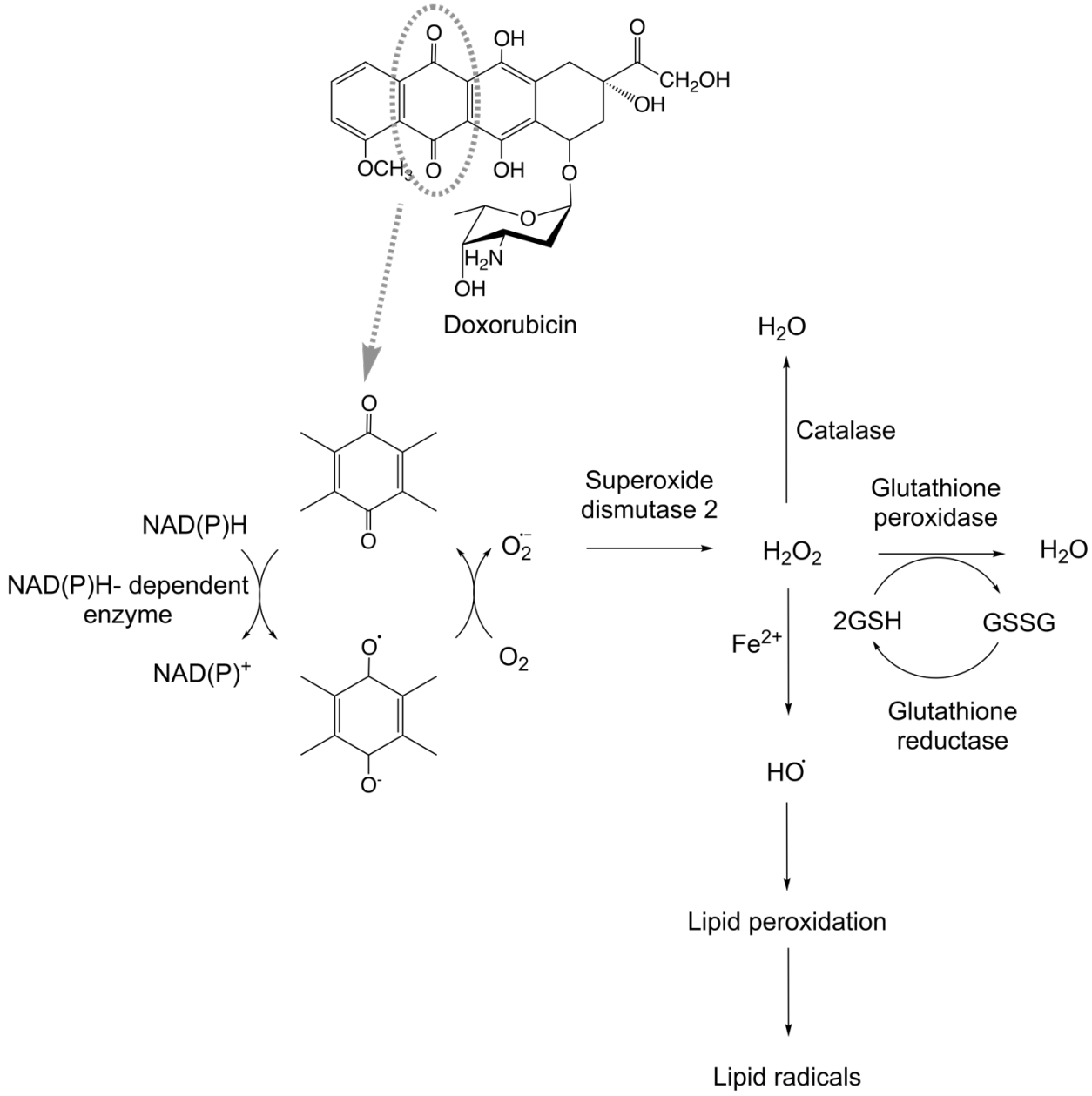


Figure 12: Redox-cycling of doxorubicin leading towards ROS. Herein, the quinone moiety is reduced to semiquinone by a NAD(P)H-dependent enzyme, and recycling of semiquinone produces $^{\cdot}O_2^-$ that is subsequently converted to H_2O_2 by superoxide dismutase 2. The resulting H_2O_2 is detoxified by catalase or glutathione peroxidase or is involved in generating lipid radicals through $^{\cdot}OH$. GSH: glutathione, GSSG: glutathione disulfide. The figure was adapted from Jungsuwadee, 2016.

1.4.3.2 The doxorubicin analog epirubicin

Although more than 1000 structurally different analogs were developed by altering the anthracycline chemical structure, only a few of them were clinically approved and this includes the doxorubicin analog, epirubicin (Weiss *et al*., 1986). Epirubicin structurally differs from doxorubicin by the equatorial C-4' hydroxyl group of the sugar moiety, where it is axial in doxorubicin (**Figure 13**). Currently, a number of cancers, including breast, lung, ovary, gastric, prostate, esophagus, and soft tissue sarcomas are treated using epirubicin (Cortés-Funes and Coronado, 2007). Interestingly, this analog showed similar antitumor activity as doxorubicin with a lower level of cardiotoxicity (Lahtinen *et al*., 1991; Weiss *et al*., 1986).

The difference in metabolism and pharmacokinetics leads to the reduced cardiotoxic profile of epirubicin. The epimerization of C4'-OH group on epirubicin favours the glucuronidation of the sugar moiety, where it increases the bile and urine excretion resulting in the increased total plasma clearance (Weiss *et al*., 1986). Thus, the total recommended cumulative dose of epirubicin is between 900-1000 mg/m^2 that is twice the dose of doxorubicin (Bonadonna *et al*., 1993).

L-acosamine is the sugar moiety of epirubicin and the corresponding sugar donor is TDP-L-acosamine. However, a natural biosynthetic pathway for epirubicin or its sugar precursor is not reported (Thibodeaux *et al*., 2007). Therefore, these two compounds are produced by chemical routes.

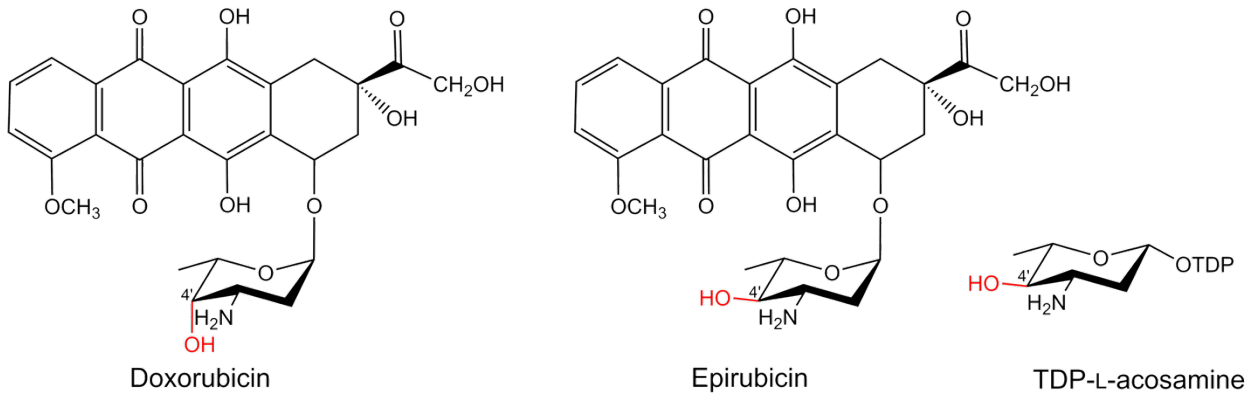


Figure 13: Structural difference between doxorubicin and epirubicin and the chemical structure of TDP-L-acosamine. The stereochemical difference between sugar residues are indicated in red.

1.4.3.3 Chemical synthesis of epirubicin and its sugar donor, TDP-L-acosamine

Epirubicin is semisynthetically produced using doxorubicin aglycone or daunorubicin through a number of chemical conversions.

In 1975, Soranzo and coworkers synthesized epirubicin starting from the doxorubicin aglycone. Herein, the precursor molecule was protected to prevent its interference with subsequent steps. It was then linked to a protected sugar donor derivative that was synthesized through a series of chemical reactions. Subsequently, the resulting anomeric mixture was chromatographically separated and the desired compound, epirubicin was obtained upon deprotection (Arcamone *et al*., 1975).

In another method, the chemical production of epirubicin is initiated from daunorubicin. Briefly, daunorubicin is broken into the aglycone moiety and the sugar residue. Then, the aglycone is converted into 14-acetoxy daunomycinone. A number of chemical reactions are used to obtain the protected aminosugar 4'-epimer from the sugar residue. Following chemical glycosylation of the 14-acetoxy daunomycinone with the protected sugar epimer, deprotection is performed to produce epirubicin (van der Rijst *et al*., 1999).

Since TDP-L-acosamine is not reported to be produced in a wild-type organism (Thibodeaux *et al*., 2007), Kahne and colleagues described a chemical synthetic process towards the synthesis of this sugar donor. Herein, the α-glycosyl chloride is converted to the protected sugar donor through numerous chemical steps. Subsequent chromatographic techniques and deprotection reactions lead to the TDP-L-acosamine sugar donor (Oberthür *et al*., 2004).

However, these studies demonstrated that the chemical synthesis of epirubicin and its nucleotide activated sugar donor as a tedious process as it involves many chemical conversions, blocking and deblocking the functional groups, and chromatographic purification of the desired anomeric mixture. Thus, use of sugar biosynthetic enzymes and GTs would be an attractive alternative for the production of glycosides and its respective sugar donors.

1.5 Glycodiversification of small molecule natural products

The biological value of sugar appendages on small molecule natural products emphasizes the importance in altering glycosylation or glycodiversification towards improving or modifying the activity of parental glycosides. Even though this can be done by enzymatic or chemical approaches, total chemical synthesis is not an attractive strategy as the structural complexity of several small molecule glycosides require multiple chemical reactions and protection and deprotection of functional groups (Thibodeaux *et al*., 2007; Thibodeaux *et al*., 2008).

An extensive knowledge on GTs, sugar biosynthetic enzymes and pathways, and the substrate flexibility of many of these enzymes have made enzymatic glycodiversification a feasible approach. The production of a desired stereo and/or regio product rather than a mixture, ease of scalability through fermentation, and the possibility in combinatorial glycodiversification are some of the key features that highlights the importance of enzyme-based glycodiversification as oppose to chemical derivatization (Thibodeaux *et al*., 2007; Thibodeaux *et al*., 2008). Several studies have highlighted the use of enzymes in altering glycosylation or in the synthesis of new glycoforms in host cells and *in vitro*.

1.5.1 *In vivo* natural product glycodiversification

In vivo glycodiversification is achieved by several glycodiversification strategies including combinatorial biosynthesis and metabolic pathway engineering. Herein, in a wild-type or mutant host, intermediates of a pathway are re-directed to produce novel glycoproducts by introducing and expressing genes of different origin (Thibodeaux *et al*., 2007; Thibodeaux *et al*., 2008). Previous studies have used this technique towards the production of the high value anthracycline epirubicin and its glycosyl donor TDP-L-acosamine (**Figure 14**). This is accomplished by the manipulation of genes involved in the doxorubicin biosynthetic pathway and heterologous expression of substrate promiscuous sugar biosynthetic enzymes and GTs (Han *et al*., 2011; Madduri *et al*., 1998; Wang *et al*., 2018a).

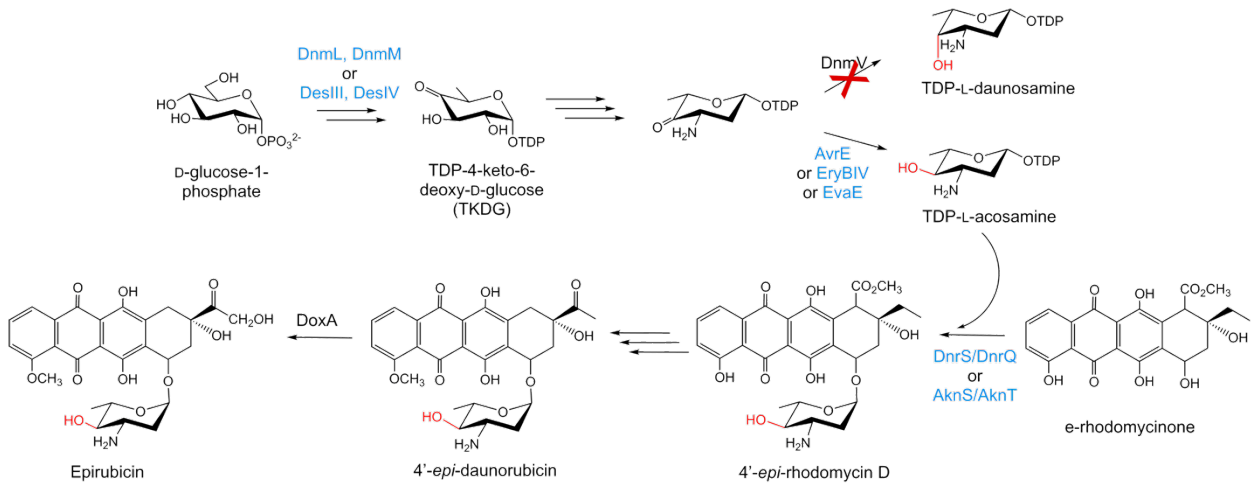


Figure 14: *In vivo* glycodiversification to produce TDP-L-acosamine and epirubicin. In one of the studies, *dnmV* gene was disrupted in *S. peucetius* strain and the sugar donor intermediate was converted to TDP-L-acosamine using AvrE and EryBIV separately. Herein, glycosylation was performed using DnrS/ DnrQ (Madduri *et al*., 1998). In another study, DesIII and DesIV was used to synthesize TKDG, AvrE was the 4-ketoreductase, and AknS/AknT was used to produce epirubicin in a *S. venezuelae* mutant (Han *et al*., 2011). In a third study, overexpression of DesIII, DesIV, EvaE, DnrS/DnrQ produced better titers of epirubicin in a doxorubicin overproducing *S. peucetius dnmV* mutant (Wang *et al*., 2018a).

In an early study, Hutchinson and colleagues reported the *in vivo* biosynthesis of TDP-L-acosamine and epirubicin in a *S. peucetius dnmV* mutant strain. Disruption of the *dnmV* gene terminated the production of TDP-L-daunosamine and its glycosylated anthracycline products, daunorubicin and doxorubicin. The accumulating sugar donor intermediate was re-directed to produce TDP-L-acosamine by introducing one of two substrate-flexible stereospecific ketoreductases (Madduri *et al*., 1998). These include AveBIV (AvrE) from *S. avermitilis* and EryBIV from *Saccharopolyspora erythraea* that are involved in the biosynthesis of avermectin and erythromycin, respectively (Ikeda *et al*., 1999; Summers *et al*., 1997). Fermentation of the engineered *S. peucetius* resulted in a considerable amount of 4'-*epi*-daunorubicin, one of the glycosylated intermediates of the proposed epirubicin biosynthetic pathway. However, the yield of epirubicin was low (Madduri *et al*., 1998).

In another study, Yoon and coworkers used a *S. venezuelae* mutant as the heterologous host for the production of TDP-L-acosamine and epirubicin. Herein, the complete antibiotic pikromycin gene cluster of *S. venezuelae* was deleted and doxorubicin resistant genes were introduced to the mutant strain. The mutant was transformed with a plasmid containing a gene set for the production

of TDP-L-acosamine, genes encoding a GT and its auxiliary protein, and postglycosylation enzymes. DesIII thymidylyl transferase and DesIV TDP-D-glucose 4,6-dehydratase from *S. venezuelae* were used to produce the key intermediate TKDG. AvrE was selected as the stereospecific 4-ketoreductase for the production of TDP-L-acosamine. The substrate flexible AknS from *S. galilaeus* was selected as the GT capable of transferring the L-acosamine residue to the exogenously fed ε-rhodomycinone with the aid of AknT auxiliary protein. Though they were able to produce 0.5 mg of epirubicin per liter of culture (Han *et al*., 2011), the yield was extremely low compared to the yield of doxorubicin by native and engineered producers (Wang *et al*., 2018b).

A recent work described by Chen and coworkers presented an engineered *S. peucetius* strain with improved epirubicin yields. Herein, the *dnmV* gene in a doxorubicin overproducer was inactivated and EvaE from *Amycolatopsis orientalis* was used as the 4-ketoreductase (Wang *et al*., 2018a). EvaE in its native organism is involved in the synthesis of TDP-L-epivancosamine sugar precursor towards the production of chloroeremomycin (Ikeda *et al*., 1999). Overexpression of EvaE, DesIII, DesIV, DnrS/DnrQ resulted in 270 mg/L of epirubicin (Wang *et al*., 2018a). These studies highlight the potential of *in vivo* combinatorial synthesis towards the production of sugar donor derivatives and the corresponding glycosides.

1.5.2 *In vitro* natural product glycodiversification

Small molecule natural product glycodiversification is not restricted to *in vivo* metabolic engineering and combinatorial synthetic approaches, but also performed *in vitro* using glycosylation enzymes. In comparison with *in vivo* routes, *in vitro* glycosylation has become an attractive alternative as the absence of host cells eliminates the concern on the toxicity of novel glycosides to the host. Furthermore, the diversity of the novel compounds can be expanded easily as the use of precursor molecules is not limited by the ability to produce or uptake by the host cells. This method also enables the optimization and control of enzymatic reactions towards maximizing the desired product yield while eliminating shunt pathways (Thibodeaux *et al*., 2008). This approach was used in a number of studies in diversifying the glycosylation of clinically important secondary metabolites.

Liu and coworkers reported the use of the substrate flexible DesVII GT along with DesVIII auxiliary protein from the macrolide antibiotic methymycin/pikromycin biosynthetic pathway of

S. venezuelae towards *in vitro* glycodiversification. The polyketide macrolactone 10-deoxymethynolide and narbonolide are the native acceptor substrates of DesVII/DesVIII, whereas its native donor is TDP-D-desosamine. Screening with a number of TDP-sugar donors and acceptors revealed the substrate tolerance of DesVII/DesVIII in synthesizing 19 novel macrolides (Borisova *et al*., 2006). This emphasizes the feasibility of the use of substrate flexible GTs in *in vitro* combinatorial biosynthesis leading to glycoform derivatives with therapeutic significance.

Not all GTs are inherently substrate flexible and the stringent substrate specificity of many GTs hampers their use in glycodiversification. To address this challenge, Spencer and colleagues used a structure-guided approach to engineer chimeras from GTs encoded in the vancomycin group of glycopeptide antibiotics chloroemomycin and teicoplanin biosynthetic gene clusters from *A. orientalis* and *Actinoplanes teichomyceticus*, respectively (Truman *et al*., 2009). For one of the chimeras, GtfA enzyme from the chloroemomycin pathway (Lu *et al*., 2004) and Orf1 from the teicoplanin pathway (Li *et al*., 2004) were used, and these two GTs share 66% sequence identity. These enzymes glycosylate their respective macrocyclic peptide acceptor substrates: desvancosaminyl vancomycin (DVV) for GtfA, and teicoplanin glucosaminyl-pseudoaglycone for Orf1; with GtfA transferring L-β-4-epi-vancosamine and Orf1 transferring *N*-acetylglucosamine from TDP- and UDP-donors, respectively. The GtfAH1 chimera, a hybrid GT consisting of the GtfA N-terminal domain and Orf1 C-terminal domain, can transfer *N*-acetylglucosamine from UDP-GlcNAc or glucose from UDP-Glc onto DVV as well as several related macrocyclic peptides (Truman *et al*., 2009). Interestingly, this hybrid GT possessed a substrate scope beyond that of its parents for the acceptors and donors that it can utilize and can synthesize a variety of hybrid vancomycin derivatives. Obtaining a substrate flexible GtfAH1 chimera by fusing domains of parent GTs that have strict substrate specificity is an interesting development with respect to the potential for glycodiversification.

1.6 Purpose and objectives

Sugar moieties attached to small molecule natural products show a significant effect on their biological activities. Therefore, changing the glycosylation of natural products through glycodiversification could affect or improve their biological properties. The overall purpose of this thesis project is to achieve *in vitro* glycodiversification of small molecules using substrate flexible GTs and sugar biosynthetic enzymes to produce novel high-value glycosylated products and heavily modified sugar precursors. The objectives towards this purpose are as follows.

In vitro glycodiversification necessitates a high-throughput assay towards identifying the activity of substrate flexible GT enzymes. However, the existing screening methods have some limitations with respect to the throughput and the ease of use. In **Chapter 2**, we have addressed this by developing a simple and easy to perform high-throughput screening tool for GTs based on the physiochemical changes of the aglycone moiety upon glycosylation.

In **Chapter 3**, we used this novel assay in *in vitro* glycodiversification, where a panel of substrate promiscuous plant GTs was combinatorially screened with a collection of acceptor and donor molecules. This enabled us to characterize the substrate flexibility of GTs and resulted in novel glycosides that might have therapeutic significance.

Considering the clinical importance of TDP-L-daunosamine and its epimer TDP-L-acosamine on small molecule glycosides, in **Chapter 4** we have reconstituted an *in vitro* enzymatic pathway towards producing these sugar donors and some of the intermediates of the pathway. Towards this we used heterologously expressed and purified substrate flexible sugar biosynthetic enzymes. In the reconstituted pathway, we have closely analyzed one of the enzymatic steps, which constitutes a major bottleneck in the biosynthesis of the desired sugar donors. Using one of the purified sugar donor intermediates, we have determined the kinetic parameters of the sugar biosynthetic enzyme that catalyzes this key reaction. This highlights the importance of having purified sugar donors not only as precursors towards producing novel bioactive molecules through *in vitro* glycodiversification, but also as tools to study glycosylated natural product biosynthesis at the level of enzyme catalysis.

CHAPTER TWO

Development of a Biphasic Glycosyltransferase High-Throughput (BiG HiT) assay

2.1 Background

The sugar moieties of many small molecule natural products are essential for their biological activity (Weymouth-Wilson, 1997). Several natural product GTs are inherently promiscuous to substrates, and thus can be challenged by non-native donor and acceptor molecules. This makes them valuable tools towards enhancing or altering the bioactivity of natural products through changing their glycosylation, otherwise known as glycodiversification (Williams *et al*., 2007). The success of glycodiversification greatly depends on effective screening methods to identify promiscuous enzymes. However, screening for GT activity is difficult, since the formation of glycosidic bonds does not produce a conveniently measurable signal like a change in absorbance or fluorescence (Aharoni *et al*., 2006).

Therefore, the catalytic activity of GTs is traditionally assayed through high performance liquid chromatography (HPLC) and/or mass spectrometry (MS) analysis. Dixon and coworkers used a high-throughput HPLC method to screen the enzyme activity of a panel of plant natural product GTs with a precursor library (Modolo *et al*., 2007). Despite their characterization of this as a high-throughput tool, the throughput of this screening approach is limited by the long operation time of HPLC instrumentation.

Consequently, researchers have come up with different high-throughput assays that use modified substrates or fluorescent sensors that facilitate the screening of GT activity. Withers and colleagues used fluorescently tagged acceptors to screen the activity of a bacterial sialyltransferase mutant library *in vivo* using fluorescence-activated cell sorting (FACS). In this strategy, with the aid of membrane transporters the exogenously provided fluorescently tagged acceptors are readily transported across the cell. Functional GT mutant enzymes catalyze the synthesis of fluorescent product that is not a substrate for membrane transport proteins, and which therefore gets entrapped and links the genotype (the gene encoding the mutant GT) and the resulting fluorescent phenotype within the cell (Aharoni *et al*., 2006).

Lee and Thorson designed an *in vitro* high-throughput assay employing a xanthene-based NDP sensor. This complex fluorescent sensor showed selective binding to the releasing NDP byproduct of Leloir-GT catalyzed reactions over NDP-sugar donor substrates. Even though crude cell lysates and purified GTs are compatible with this assay, synthesis of the complex NDP sensor is laborious as it involves multiple chemical reactions (Lee and Thorson, 2011). Instead of the complex xanthene-based sensor, Lee and colleagues improved the assay with the use of an anthracene-based fluorescent sensor where the chemical preparation is not difficult. The fluorescent sensor is bound to a quencher, and the presence of NDP replaces the quencher and emits fluorescence (Ryu *et al*., 2014). Although these GT screening tools are attractive in the context of its throughput and sensitivity, each assay necessitates chemical synthesis of fluorescently labeled acceptors, or fluorescence sensors.

We have addressed this by developing a high-throughput screening tool based on the changes in physical properties of the aglycone acceptors upon glycosylation. Various lipophilic small molecules with chromogenic or fluorescent spectral properties are poorly soluble in polar solvents and the polarity significantly increases upon glycosylation (Modolo *et al*., 2007). The high-throughput assay we developed utilizes a simple liquid-liquid extraction to separate the polar glycoside product of the GT-catalyzed reaction which can be detected by the inherent absorbance or fluorescence of the aglycone residue (**Figure 15**). The assay we have devised is relatively simple and can be easily used in a high-throughput fashion on microtiter plates in combinatorial screening or directed evolution.

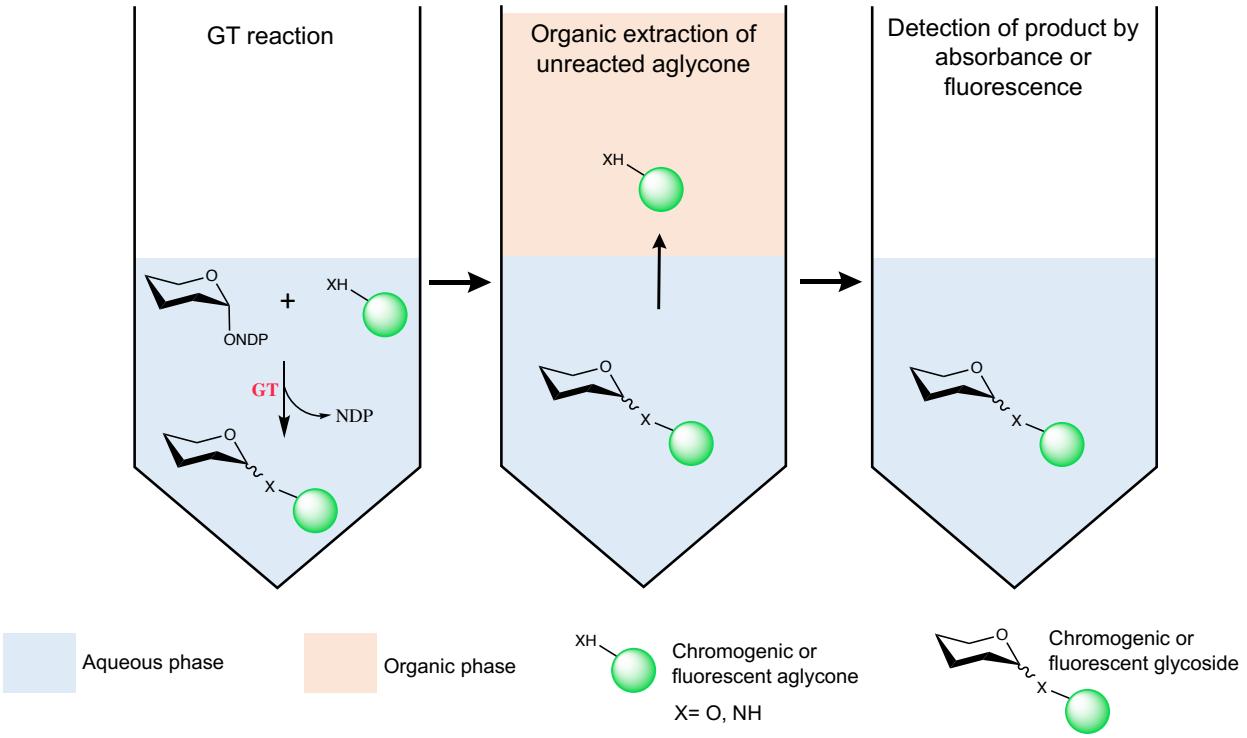


Figure 15: Schematic representation of the developed Biphasic Glycosyltransferase High-Throughput (BiG HiT) assay. A catalytically active GT transfers a sugar moiety from a nucleotide activated donor to a non-polar naturally chromogenic or fluorescent acceptor and results in a chromogenic or fluorescent polar product (*O*-, *C*- or *N*-linked). The unreacted aglycone is separated from the product by performing an organic solvent extraction. Subsequently, the product in the aqueous phase is detected by measuring the absorbance or fluorescence.

2.2 Materials and Methods

2.2.1 Chemical reagents and enzymes

Chemicals used in this study were of analytical grade and were commercially purchased from Sigma-Aldrich, Carbosynth, Bio Basic Canada, and Bio-Rad. DNase I, RNase A, and lysozyme were purchased from Bio Basic Canada.

2.2.2 Plasmids, bacterial strains, and culture conditions

Plasmids and bacterial strains of this chapter are summarized in **Table A2** and **A3**, respectively. Growth and maintenance of *E. coli* strains were done using LB broth or agar media (**Table A4**) with 100 µg/mL of ampicillin. The cultures were incubated at 37 ºC and liquid cultures were shaken at 220 rpm, unless otherwise noted.

The chemically competent *E. coli* cells used in this study were prepared following a previous procedure described by Hanahan (Hanahan, 1983) with modifications. A 1 % of starter culture that had grown overnight at 37 ºC was inoculated into 500 mL LB broth and was incubated at 37 ºC until the OD_{600} is between 0.4- 0.6. Following centrifugation at 4500 x g for 5 min at 4 ºC, the harvested cells were resuspended in 200 mL ice-cold TBF1 buffer (**Table A5**). Next, a 5 min incubation on ice was done and cells were harvested by another centrifugation step (4500 x g for 5 min at 4 ºC). Subsequently, the cells were gently resuspended in 20 mL ice-cold TBF2 buffer (**Table A5**), and 100 µL aliquots were stored at -80 ºC.

2.2.3 Expression and purification of recombinant VvGT1 glycosyltransferase

The *VvGT1* gene had been previously cloned into the pET14b T7 expression vector by Christopher Ford (Ford *et al*., 1998). This construct was transformed into chemically competent *E. coli* BL21 (DE3) cells and VvGT1 protein was heterologously expressed and purified following a previously established protocol (Ford *et al*., 1998) with slight modifications. The starter culture was grown overnight at 37 ºC and 1% of it was inoculated into 800 mL of LB broth supplemented with 100 µg/mL of ampicillin. Cells were grown at 16 ºC until OD_{600} is between 0.4- 0.6 and protein expression was induced with 0.4 mM IPTG. The cultures were further incubated for an additional 24 hours at 16 ºC and the cells were harvested by centrifugation (10,000 x g, 30 min, 4 ºC). The

cell pellet was resuspended in 40 mL of VvGT1 lysis buffer. DNase I, RNase A, and lysozyme (5 μg/mL of each), and one mini tablet of protease inhibitor cocktail (EDTA-free) were added. Following cell disruption by sonication with a 5 s pulse on, 15 s pulse off for a total pulse time of 3 min at 25% amplitude (Fisherbrand™ Model 505 Sonic Dismembrator), the insoluble components were pelleted by centrifugation at 15,000 x g for 30 min at 4 ºC.

A 1 mL Ni-NTA resin column (MCLAB) was pre-equilibrated with 10 mL of VvGT1 wash buffer and the filtered supernatant (0.22 μm filter, UltiDent) was loaded to the column. Subsequently, the column was washed with 10 column volumes of VvGT1 wash buffer and the expressed VvGT1 protein was isocratic eluted by gravity through 8 column volumes of VvGT1 elution buffer. The fractions containing VvGT1 protein were pooled and buffer exchanged using 10DG desalting column (Bio-Rad) with 4 mL of VvGT1 storage buffer, and aliquots of protein were stored at -80 ºC.

Note: The buffers used in this section were listed in **Table A5**.

2.2.4 Biphasic glycosyltransferase assay validation using purified VvGT1 enzyme

The biphasic glycosyltransferase high throughput assay (BiG HiT) validation was carried out in 96-well PCR plate (VWR) with 8 replicates. 10 μL of purified VvGT1 protein (0.25 mg/mL in total reaction volume) was added to a 20 μL substrate mixture containing 10 mM Tris pH 7.8, 1.5 mM quercetin (3 % DMSO), 3 mM UDP-glucose, and the total reaction volume was 30 μL. As the no enzyme control, 10 μL of assay control buffer (**Table A5**) was added to the substrate mixture. The reactions were incubated at 30 ºC for 2 hours and terminated by incubating at 90 ºC for 3 min. The insoluble components were pelleted out by centrifugation at 3,000 x g for 5 min. 25 μL of the supernatant was transferred to a new well and one volume of ethyl acetate: chloroform (3:1 v/v) was added as the organic phase. The content was mixed using an 8-channel pipettor (Corning Lambda™ Plus), and liquid-liquid extraction was performed on a Whatman 1PS phase separator filter paper (GE Healthcare). The solvent extraction step was repeated and 15 μL of aqueous phase was transferred to wells of 96-well half area plate (Corning). The sample was diluted with 1 volume of distilled water and 2 volumes of 0.2 % 2-aminoethyl diphenylborinate fluorescent probe in 20 % ethanol. The absorbance at 394 nm was measured using a CLARIOstar

monochromator microplate reader (BMG Labtech) and the data were received from MARS data analysis software 3.20 R2.

2.2.5 Biphasic glycosyltransferase assay validation using crude VvGT1 enzyme

To assay the compatibility of crude lysate with the developed high throughput screen, VvGT1 protein lysate was prepared as follows: pET14b plasmid encoding VvGT1 was transferred into chemically competent XJb (DE3) autolysis cells. The cells were grown in 1 mL of LB broth containing 100 µg/mL of ampicillin at 37 ºC for 24 hours and as the control XJb (DE3) cells harbouring an empty pET14b vector was used. 7.5 µL of inoculum was added into V-bottom 96-well microplates (NUNC) having 150 µL LBE-5052 autoinduction medium (**Table A4**) supplemented with 100 µg/mL of ampicillin and 3 mM arabinose (8 replicates). The cells were grown at 30 ºC for an additional 24 h. The cells were pelleted by centrifuging at 3,000 x g for 10 min and the cell pellets were resuspended in 30 µL of 10 mM Tris pH 7.8. The cells were frozen at -80 ºC for 1 hour, then thawed at 30 ºC. Following the addition of 10 µL of a mixture containing 20 µg/mL DNase I, 20 µg/mL RNase A, and 20 µg/mL lysozyme in 10 mM Tris buffer, pH 7.8, the cell suspension was incubated at room temperature for 15 min. Cell debris were removed by centrifugation (3,000 x g for 10 min at 4 ºC) and 10 µL of crude lysate was used for the above explained high throughput assay validation procedure (**Section 2.2.4**).

2.3 Results and discussion

2.3.1 Development and validation of a Biphasic Glycosyltransferase High-Throughput (BiG HiT) assay

The high-throughput assay that we devised for glycosyltransferase screening is based on the change in physical characteristics of the aglycone acceptor upon glycosylation. In this strategy, the GT activity is detected by the shift in the polarity of the inherently chromogenic or fluorescent acceptor.

As a proof of concept, a promiscuous plant GT enzyme, UDP-glucose: flavonoid 3-*O*-glycosyltransferase (VvGT1) from *Vitis vinifera* (red grape) (Ford *et al*., 1998; Offen *et al*., 2006) was used in the assay we developed. Herein, we used the crude lysate from *E. coli* clones

expressing VvGT1 and His-tagged purified protein towards assay validation. Despite the fact that natural glycosylation reaction that it catalyzes involves the conversion of the unstable cyanidin acceptor to the stable cyanidin 3-*O*-glucoside, *in vitro* studies have demonstrated the potential of VvGT1 to use other flavonoid acceptors such as quercetin (Ford *et al*., 1998). VvGT1 transfers the glucose moiety from the nucleotide activated sugar donor, UDP-glucose to the non-polar aglycone acceptor, quercetin. This enzymatic reaction yields polar coloured compound quercetin 3-*O*-β-glucoside (**Figure 16A**).

To validate our proposed strategy for high-throughput screening, we performed the assay for VvGT1 purified enzyme (0.25 mg/mL) with 8 replicates. The reactions were carried out in Tris buffer, pH 7.8 (6.7 mM) using quercetin (1 mM) as the acceptor and UDP-glucose (2 mM) as the sugar donor. Herein, to determine any non-enzymatic background reactions, an assay control buffer (**Table A5**) without VvGT1 enzyme was used. Glycosylation reactions were performed for 2 h at 30 ºC and stopped by heating at 90 ºC for 3 min to denature the glycosyltransferase enzyme. Subsequently, unreacted nonpolar quercetin acceptor was removed from the polar quercetin 3-*O*-β-glucoside by an organic solvent extraction using ethyl acetate: chloroform (3:1 v/v). Herein, the solvent extraction was facilitated by the use of silicone impregnated filter paper that is impermeable to the aqueous phase but permeable to the organic phase. The quercetin aglycone is partitioned in the organic phase which is absorbed into the silicone impregnated filter paper, while the aqueous phase containing the glycosylated product remains as a droplet. The separated polar quercetin 3-*O*-β-glucoside product was detected by measuring the absorbance of the aqueous phase at 394 nm. Herein, 2-aminoethyl diphenylborinate was used to enhance the absorbance signal of flavonoid glycoside (Saslowsky *et al*., 2005).

The Z-factor is a statistical parameter that is used to measure the quality of high-throughput assays. This is calculated based on the variations between groups of positive and negative controls, and the value ranges from <0 – 1 (Zhang *et al*., 1999). We evaluated the robustness of our developed high-throughput assay by calculating the Z-factor. As illustrated by **Figure 16B**, the calculated Z-factor value for the purified VvGT1 enzyme was 0.83.

The compatibility with crude enzyme extracts is necessary towards utilizing our proposed high-throughput assay in enzyme engineering strategies such as directed evolution. Thus, the above-described assay was repeated with VvGT1 crude extract and crude extract without the

enzyme of interest was used as the no GT control. This resulted in a Z factor value of 0.63 (**Figure 16C**). Since the obtained Z-factor values for the purified VvGT1 and the crude cell extracts were greater than 0.5, the developed assay is considered to have an excellent assay quality demonstrating its utility for high-throughput screening. Therefore, we are confident that our proposed screen is well-suited in assaying crude enzyme extracts and purified GTs.

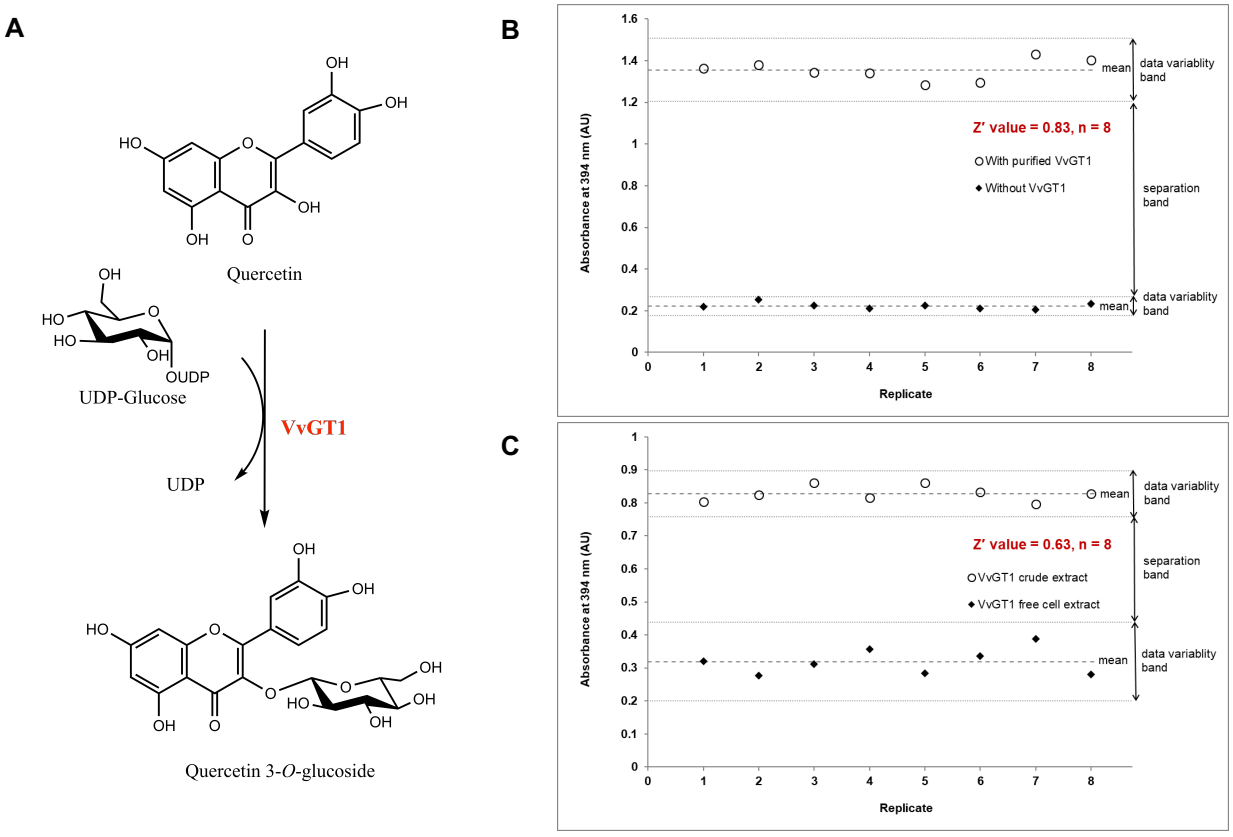


Figure 16: (A) Glycosylation reaction of VvGT1 in the presence of UDP-glucose and quercetin, and statistical validation of the proposed BiG HiT assay using **(B)** VvGT1 purified enzyme (0.25 mg/mL) and **(C)** crude extract. Following solvent extraction, the glycosylated quercetin 3-*O*-β-glucoside in aqueous phase was detected at 394 nm. Z' was calculated considering the variability between positive and negative controls. n= 8.

2.4 Conclusions

The diversity and the abundance of glycans in nature emphasize their importance in biological applications. Specifically, many therapeutically valuable compounds are decorated with sugars highlighting the pivotal function of GTs that catalyze their assembly. High-throughput enzyme screening platforms are valuable tools in identifying useful activities from naturally diverse pools of enzymes, exploring substrate scope and combinatorial synthetic activity in promiscuous enzymes, and in evolving existing enzymes by directed evolution (Kwan and Withers, 2011). Although several high-throughput GT assays have been described, many of them require chemically manipulated substrates or fluorescent sensors, and some of them are limited in their throughput. This brought our attention towards the development of a high-throughput screening tool for GTs without the need of chemically modified precursors or sensors.

We were inspired by the natural chromogenic and fluorescent properties of many polyphenolic aglycones of glycosylated natural products. Glycosylation of these hydrophobic aglycone molecules shifts its polarity (Modolo *et al*., 2007) and this is the basis of our novel GT assay. In this study, we have developed a screening method for assaying natural product GTs in a high-throughput fashion enabled by rapid isolation and detection of chromophoric or fluorescent glycosylated natural products. Previously, Thorson and colleagues developed a fluorescence-quenching assay utilizing the changes associated with the fluorescent properties of a coumarin acceptor upon glycosylation (Williams *et al*., 2007). To the best of our knowledge, the GT assay we designed is the first screening tool developed that links the spectral and polarity characteristics of the aglycone moiety. Herein, we have demonstrated the compatibility of our devised plate-based high-throughput screen with both crude lysate and purified GTs, using VvGT1 as a representative enzyme.

Although, our assay is attractive in the context of its simplicity and ease of use, this tool can only be used with GTs that use chromogenic or fluorescent lipophilic acceptors. The use of an appropriate organic phase that removes the unreacted aglycone but not the glycoside is extremely important as improper separation would result in false positive or false negative results. However, a complete separation of the aglycone from its glycoside without any cross-contamination in organic and aqueous phases may not be possible. Therefore, subtle glycosylation activities of GTs might not be detected by our BiG HiT assay. Despite these limitations, our plate-based method

could allow for the screening of 1000s of combinations/ conditions/ mutants per day. Therefore, our novel assay will be an invaluable tool for discovering, analyzing, and engineering GTs through directed evolution.

In the following chapter, we have used our novel high-throughput assay in screening a collection of natural product GTs against a panel of precursors to therapeutically important molecules.

CHAPTER THREE

Exploring the catalytic scope of a plant UGT collection towards the production of novel glycosides

3.1 Background

UDP-dependent GTs (UGTs) are a ubiquitous class of enzymes that use UDP activated sugar donors, including UDP-glucose, UDP-galactose, UDP-glucuronic acid, and UDP-xylose towards glycosylation (Mackenzie *et al*., 1997). According to the carbohydrate-active enzyme (CAZy) database classification, GTs and UGTs that usually use small molecule acceptors belong to GT1 family (http://www.cazy.org/GT1.html), and these are categorized as inverting GTs since their glycoside products result in an inverted anomeric configuration with respect to the sugar donors. Structurally, these UGTs of the GT1 family fit to the GT-B fold enzymes that possess two β/α/β Rossmann-like domains and have definite acceptor and donor binding sites (Coutinho *et al*., 2003; Lairson *et al*., 2008).

An enormous collection of UGTs was found in plants, including in *Arabidopsis thaliana* and *Medicago truncatula* (Modolo *et al*., 2007; Ross *et al*., 2001). The well conserved Putative Secondary Plant Glycosyltransferase (PSPG) motif near the C-terminal region is an exclusive characteristic of plant UGTs. This motif is reported to represent the UDP-sugar binding site (Hughes and Hughes, 1994; Paquette *et al*., 2003). Plant UGTs are involved in glycosylating small molecules such as plant hormones (Jackson *et al*., 2002), flavonoids (Ford *et al*., 1998), terpenoids (Caputi *et al*., 2008), saponins (Shibuya *et al*., 2010), and xenobiotics (Loutre *et al*., 2003). Glycosylation of these small molecules affects their solubility, stability, bioactivity, compartmentalization, and storage (Coleman *et al*., 1997; Hollman *et al*., 1999; Smith *et al*., 2000). In view of the importance of sugar residues on small molecules, glycodiversification—that is changing the sugar groups attached to these metabolites through glycosylation—would result in novel compounds with modified chemical properties and bioactivities. The feasibility of this strategy mainly relies upon the substrate flexibility of GTs (Thibodeaux *et al*., 2008; Williams *et al*., 2007).

Several *in vitro* studies revealed the substrate promiscuity of many plant UGTs participating in secondary metabolism (Caputi *et al*., 2012). Using our novel high-throughput assay

(**Chapter 2**), we screened a collection of plant UGTs against a panel of precursors to therapeutically important molecules. Six substrate-flexible plant UGTs: VvGT1 from *V. vinifera*, UGT71G1 and UGT78G1 from *M. truncatula*, UGT78K6 from *Clitoria ternatea*, UGT708A6 from *Zea mays*, and UGT72B1 from *A. thaliana* were selected as our enzyme collection. In nature, these UGTs exhibit different roles ranging from pigment production, plant defense mechanism, and detoxification of environmental pollutants through glycosylation (Achnine *et al*., 2005; Ford *et al*., 1998; Loutre *et al*., 2003).

Glycosylation is an important reaction in the synthesis of water-soluble anthocyanin pigments in plants that defines the colours of flowers and fruits. In red grape, VvGT1, a flavonoid 3-*O*-glycosyltransferase is involved in the production of red wine pigment through glycosylating a cyanidin acceptor towards the formation of cyanidin 3-*O*-glucoside (Ford *et al*., 1998; Offen *et al*., 2006). UGT78G1 and UGT78K6 are two other flavonoid *O*-glycosyltransferases that are involved in anthocyanin pigment biosynthesis (Hiromoto *et al*., 2013; Peel *et al*., 2009). In butterfly pea, UGT78K6 participates in glycosylating a delphinidin acceptor using UDP-glucose sugar donor (Hiromoto *et al*., 2013). This is an early biosynthetic step in producing the anthocyanin pigment ternatin that makes the petal colour blue (Kazuma *et al*., 2004).

Triterpene saponins are plant secondary metabolites that are essential players in plant defense and many of these small molecules are glycosylated. UGT71G1 is an *O*-UGT that is reported in glycosylating triterpene aglycones such as medicagenic acid and hederagenin towards the biosynthesis of saponins in *M. truncatula* (Achnine *et al*., 2005).

Apart from molecules with *O*-linked glycosylation (the most common type of glycosylated natural product), plant UGTs are capable of synthesizing *C*- and *N*-linked glycosides. UGT708A6 encompass a dual-function by forming both *O*- and *C*-linked glycosylated flavonoids. This enzyme is demonstrated to use UDP-glucose as the sugar donor towards glycosylating flavones and flavanones (Ferreyra *et al*., 2013). UGT72B1 is another dual-functional enzyme involved in detoxifying xenobiotics such as 2,4,5-trichlorophenol (TCP) and 3,4-dichloroaniline (3,4-DCA), by *O*- and *N*- glycosylation, respectively (Loutre *et al*., 2003).

Our novel high-throughput assay was used in *in vitro* combinatorial screening of these UGTs with a selected substrate collection. Through this screen, certain enzymes exhibited novel acceptor and donor substrate specificities. Importantly, a few of the enzymes produced novel

glycosides demonstrating its ability in glycodiversification towards therapeutically important compounds.

3.2 Materials and Methods

3.2.1 Chemical reagents and enzymes

Chemicals used in this study were of analytical grade and were commercially purchased from Sigma-Aldrich, Carbosynth, Bio Basic Canada, and Bio-Rad. ε-rhodomycinone was provided by Dr. Kristiina Ylihonko. Restriction endonucleases, T4 DNA ligase, Antarctic phosphatase, and Phusion DNA polymerase were from New England Biolabs. DNase I, RNase A, and lysozyme were purchased from Bio Basic Canada.

3.2.2 Plasmids, bacterial strains, and culture conditions

Plasmids and bacterial strains of this chapter are summarized in **Table A2** and **A3**, respectively. Growth and maintenance of *E. coli* strains with appropriate antibiotics and preparation of chemically competent cells were done as in **Section 2.2.2**.

3.2.3 Construction of plasmids encoding UGTs and enzymes involved in UDP-xylose *in vitro* pathway

Separate constructs of *UGT71G1*, *UGT78G1*, *UGT78K6*, *UGT708A6, hUGDH*, *CtXR*, and *hUXS* with *Nde*I and *Xho*I restriction sites were designed in PUC57-Kan vector and the synthesized constructs were purchased from Bio Basic. Following standard procedures, each gene was excised at *Nde*I and *Xho*I restriction sites, subcloned into pET28a vector, and individual constructs were sequence verified.

The truncated *hUXS* (Δ*84hUXS*) gene in pET28a was built as follows. Δ*84hUXS* gene was PCR amplified in a 25 µL buffered reaction containing <250 ng of pET28a_ *hUXS* (full length) template, 0.5 µM of hUXS_F and 0.5 µM T7 terminator primers (**Table A1**), 200 µM dNTPs, 0.02 to 0.2 U/ µL of Phusion polymerase, and 3% DMSO. Herein, the PCR buffer was diluted according to the manufacturer's instructions and the temperature profile was: 98 °C for 30 s, followed by 30 cycles of 98 °C for 10 s, 55–71 °C for 30 s, and 72 °C for 1 min, followed by 72 °C for 10 min.

The PCR product was restriction digested at *Nde*I and *Xho*I, cloned into pET28a vector, and sequence verified.

The *UGT72B1* gene had been previously cloned into the pET30a vector by Gideon Davies at University of York (Brazier-Hicks *et al*., 2007a).

3.2.4 Expression and purification of recombinant proteins

3.2.4.1 Recombinant plant UGTs

Each of the UGT constructs with N terminal His tag in pET vectors were individually transformed into chemically competent *E. coli* BL21 (DE3) cells and UGT proteins were expressed as described below. 1 % of starter cultures that had grown overnight at 37 ºC were inoculated into separate 800 mL volumes of LB broths containing 50 µg/mL of kanamycin. The cultures were incubated at 37 ºC until OD_{600} of 0.4 - 0.8 was reached and induced with 0.5 mM IPTG. All the cultures were overnight grown at 16 ºC and bacterial cells were harvested by centrifugation at 10,000 x g for 30 min at 4 ºC.

Cell lysis, protein purification, buffer exchange, and storage were performed as explained in **Section 2.2.3** using UGT buffers (**Table A5**).

3.2.4.2 Recombinant proteins involved in UDP-xylose *in vitro* pathway

Separate 800 mL cultures of *E. coli* BL21 (DE3) cells containing *hUGDH*, *CtXR*, and Δ*84hUXS* pET28a expression constructs were grown as explained in Section **3.2.4.1** until OD_{600} was between 0.6 - 0.8. Protein expression was initiated by adding 0.5 mM IPTG for hUGDH and *Ct*XR, and 1 mM IPTG for Δ84hUXS, and the cells were induced overnight at 16 ºC. Following centrifugation (10,000 x g for 30 min at 4 ºC) the cell pellets were resuspended in lysis/ wash buffer I (**Table A5**), and the cells were disrupted as described in **Section 2.2.3**.

The isocratic purification of Δ84hUXS protein was performed by gravity as explained in **Section 2.2.3** using elution buffer I (**Table A5**). hUGDH and *Ct*XR protein purification were performed using a 1 mL Ni-NTA resin column (MCLAB) equipped on an AKTA purifier 10 system (GE Healthcare Life Sciences). Herein, the proteins were gradient eluted from 10 mM (wash buffer I: **Table A5**) to 500 mM imidazole (elution buffer I: **Table A5**) at a flow rate of 1

mLmin^{-1} over 25 column volumes. The fractions containing the recombinant hUGDH, *Ct*XR, and Δ84hUXS proteins were separately concentrated to 3 mL using centrifugal protein concentrators (6 mL, 30 kDa MWCO, GE Healthcare), were dialyzed using 10DG desalting columns (Bio-Rad) against storage buffer I (**Table A5**), and protein aliquots were stored at -80 ºC.

3.2.5 *In vitro* enzymatic synthesis and purification of UDP-xylose

UDP-xylose was enzymatically prepared from UDP-glucose in a one-pot two-step reaction following a previously described procedure (Eixelsberger and Nidetzky, 2014) with necessary modifications. Briefly, to a 4 mL reaction containing 50 mM potassium phosphate buffer pH 7.5, 10 mM UDP-glucose, 6 mM NAD^+, 0.025 mM 9,10-phenanthrenequinone (PQ), and 0.10 w/v % BSA, the following two were added in order: 100 U/mL bovine liver catalase and 0.072 mg/mL *Ct*XR. The reaction was initiated by adding 0.825 mg/mL hUGDH, and 100 mM H_2O_2, mixed and incubated at 37 ºC for 24 h. Subsequently, 4 mL of 1.43 mg/mL of hUXS was added and incubated for an additional 26 h. The reaction was terminated at 99 ºC for 5 min and denatured proteins were removed by centrifugation at 3,000 x g for 5 min.

Synthesized UDP-xylose was purified in two batches (4 mL each) by anion exchange chromatography on an AKTA purifier 10 system (GE Healthcare Life Sciences) with an 8 mL Macro-prep high Q column (Bio-Rad). Preceding sample loading, the column was equilibrated with 5 column volumes of buffer A (buffer II: **Table A5**). The compounds were eluted at 2 mLmin^{-1} flow rate with a gradient of 0 % to 11.5 % B over 4 column volumes, 11.5 % to 21.2 % B over 5 column volumes, and 21.2 % to 100 % B over 4 column volumes where buffer B is buffer III (**Table A5)**. The fractions containing UDP-xylose were monitored at UV 254 nm and verified by mass spectrometry (MS) (Thermo-Finnigan 7 Tesla LTQ-FT) in negative mode with a source voltage of 2800 V and a scan range from 100 to 1000 *m/z*. The acquired data from MS were analyzed using the Xcalibur software package.

The pooled fractions were concentrated to 2 mL using a rotary evaporator at 40 ºC (BUCHI rotavapor R-114) and applied to a pre-equilibrated (with 2 column volumes of deionized water) 75 mL Sephadex G-10 (Sigma Aldrich) column connected to an AKTA purifier 10 system (GE Healthcare Life Sciences) to remove ammonium formate from UDP-xylose. Herein, deionized water was used to elute the compounds at 0.2 mLmin^{-1} flow rate over 3 column volumes and UDP-

xylose was monitored by UV absorption at 254 nm. Fractions were verified by MS as explained above and lyophilized using a ModulyoD-115 freeze dryer at -50 ºC and a pressure less than 0.51 mbar.

3.2.6 High-throughput combinatorial screening of the promiscuous plant UGT collection

The purified UGT enzyme collection (along with VvGT1) and a panel of substrates were combinatorially screened using the high throughput assay we developed. Herein, quercetin, kaempferol, myricetin, emodin, and ε-rhodomycinone were used as acceptors and UDP-glucose, UDP-galactose, and UDP-xylose were used as sugar donors. For each acceptor substrate, a 96-well PCR plate (VWR) was used with different combinations of UGTs and sugar donors, and the screening was performed as explained in **Section 2.2.4** with a minimum of 2 replicates. Two negative controls, either without sugar donors or without UGTs were performed. The organic solvent mix for flavonoids and anthraquinones were ethyl acetate: chloroform (3:1) and ethyl acetate: chloroform (1:3), respectively. For flavonol acceptors, 0.2% of 2-aminoethyl diphenylborinate was used.

3.2.7 Validation of positive hits by HPLC−MS/MS Analysis

The positive reactions of the combinatorial high-throughput screen were analyzed by performing HPLC-MS/MS analysis using an Acquity BEH C18 column (2.1 x 50 mm, 1.7 µm) (Waters, CA) connected to an Agilent 1260 Infinity II UHPLC system (Agilent technologies, CA) equipped with a Thermo-Finnigan 7 Tesla LTQ-FT mass spectrometer. 5 µL of the diluted sample (1:1 using distilled water) was injected and chromatographic separation was performed at a flow rate of 0.3 mL/min with a gradient of 5% to 80 % buffer B over 5 min, followed by two isocratic steps with 90% buffer B over 0.5 min, and 5% buffer B over 2.5 min where buffer A was buffer IV (**Table A5**) and buffer B was buffer V (**Table A5**). The analytes were detected using mass spectrometry in negative mode and MS/MS fragmentation was done for selected *m/z* ratios at an isolation window of 2.5 Da, activation time of 30 s, and using 15 and 20 normalized collision energies (cid).

3.3 Results and discussion

3.3.1 Expression of the plant UGT collection

The broad substrate specificity of many natural product GTs that act on secondary metabolites allows them to be screened against a variety of donor and acceptor molecules. This paves the path for glycodiversification that would allow for the synthesis of novel glycosylated small molecules with modified or enhanced biological activities. Towards *in vitro* glycodiversification, six well-characterized and substrate promiscuous plant UGTs were selected. These include five UGTs that act on flavonoids (VvGT1, UGT71G1, UGT78G1, UGT78K6, and UGT708A6) and one UGT that detoxifies xenobiotics (UGT72B1). These UGTs were cloned into pET expression vectors by traditional cloning methods, soluble and active proteins were heterologously expressed in *E. coli* BL21 (DE3) cells, and hexa-His N-terminal tagged proteins were Ni affinity purified (**Figure A1**). Approximately 6 mg of purified UGTs were obtained from 800 mL of culture.

3.3.2 *In vitro* enzymatic synthesis of UDP-xylose

Certain UGTs involved in natural product glycoside biosynthesis are capable of utilizing UDP-xylose as the sugar donor and transferring the xylose moiety to a range of acceptors. Therefore, along with UDP-glucose and UDP-galactose, UDP-xylose was also used as a sugar donor in our high-throughput combinatorial screen towards *in vitro* glycodiversification.

A previously established procedure was used with necessary modifications to produce UDP-xylose in a one-pot two-step enzymatic reaction (Eixelsberger and Nidetzky, 2014). The pathway involves three enzymes, human UDP-glucose 6-dehydrogenase (hUGDH), *Candida tenuis* xylose reductase (*Ct*XR), and human UDP-xylose synthase 1 (hUXS) (**Figure 17**). Individual gene constructs encoding these proteins were built in pET28a expression vector, soluble and active proteins were overexpressed in *E. coli* BL21 (DE3) cells, and His tagged proteins were purified by Ni affinity chromatography (**Figure A2**). For 800 mL of culture, around 8 mg, 28 mg, and 4 mg of purified hUGDH, *Ct*XR, and hUXS proteins were obtained, respectively.

As shown in **Figure 17**, the proposed pathway was similar to that of the natural biosynthetic pathway of UDP-xylose, where UDP-xylose was enzymatically synthesized from UDP-glucose through an UDP-glucuronic acid intermediate. The sequential reactions of the suggested enzymatic

scheme were catalyzed by hUGDH and hUXS, respectively (Eixelsberger and Nidetzky, 2014). The coupling of a coenzyme nicotinamide adenine dinucleotide (NAD^+) recycling cascade was also used in the first step to elevate the efficiency of UDP-glucuronic acid production (Pival *et al*., 2008). According to **Figure 17**, the reduction of 9,10-phenanthrenequinone (PQ) to the respective hydroquinone (PQH_2) by *Ct*XR results in the regeneration of NAD^+. Subsequently, molecular oxygen converts PQH_2 to PQ while the resulting H_2O_2 was decomposed by bovine liver catalase. Following the complete conversion of UDP-glucose to UDP-glucuronic acid intermediate, the successive step of the reaction was performed by the introduction of hUXS. As reported by Eixelsberger and Nidetzky, it was extremely important to synthesize UDP-xylose in two steps to avoid the inhibition of hUGDH by UDP-xylose (Eixelsberger and Nidetzky, 2014).

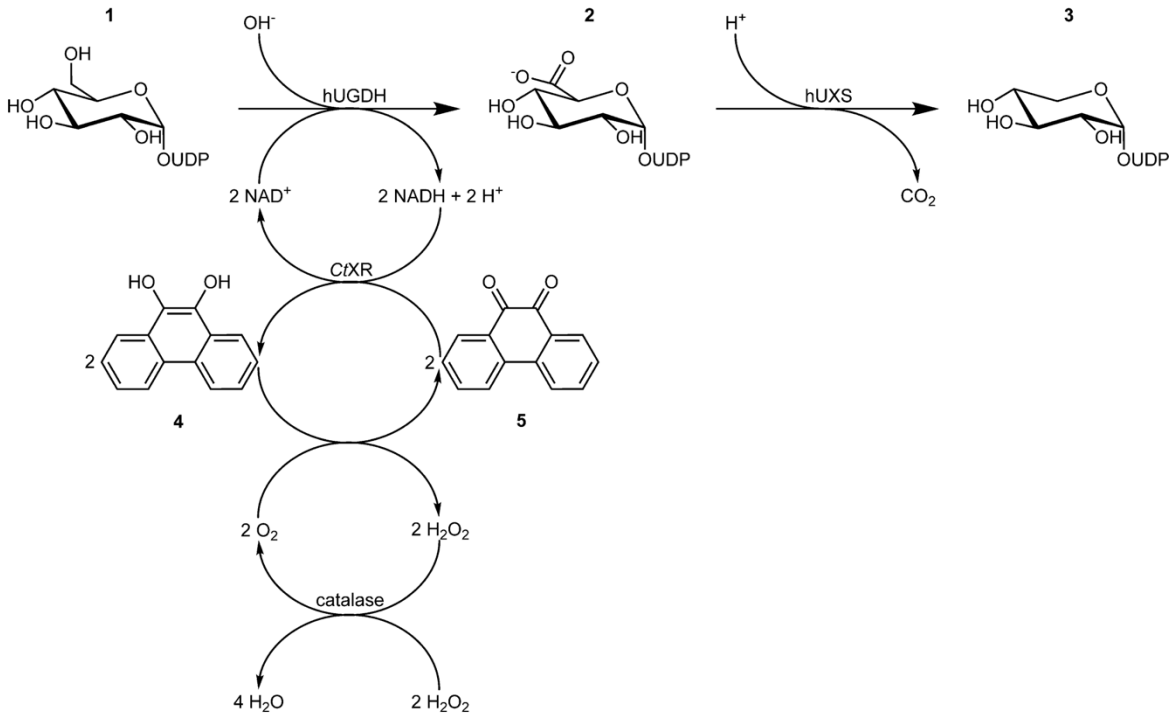


Figure 17: Previously published *in vitro* enzymatic route for UDP-Xylose synthesis. UDP-Glucose (1) was converted to UDP-Glucuronic acid (2) by hUGDH enzyme. The efficiency of the reaction was enhanced by combining the NAD^+ regeneration cascade that involves *Ct*XR and bovine liver catalase. Next, UDP-xylose was synthesized employing hUXS enzyme. Compound 4: PQH_2, compound 5: PQ. (Eixelsberger and Nidetzky, 2014).

Next, the produced UDP-xylose was purified using a strong anion exchange resin (Macro-prep high Q) by anion exchange chromatography with a gradient elution profile of NH_4HCO_2 pH 4.2 (20 mM – 500 mM) (**Figure A3A**). Subsequently, a second purification step using Sephadex G-10 gel filtration medium was performed to remove NH_4HCO_2 and other salts (**Figure A3B**). Finally, UDP-xylose was lyophilized, verified by MS analysis (**Figure 18**), and a total yield of 54% (13.2 mg) was obtained (starting with 24.4 mg of UDP-glucose). The purified UDP-xylose was added to our substrate panel towards screening our UGT collection.

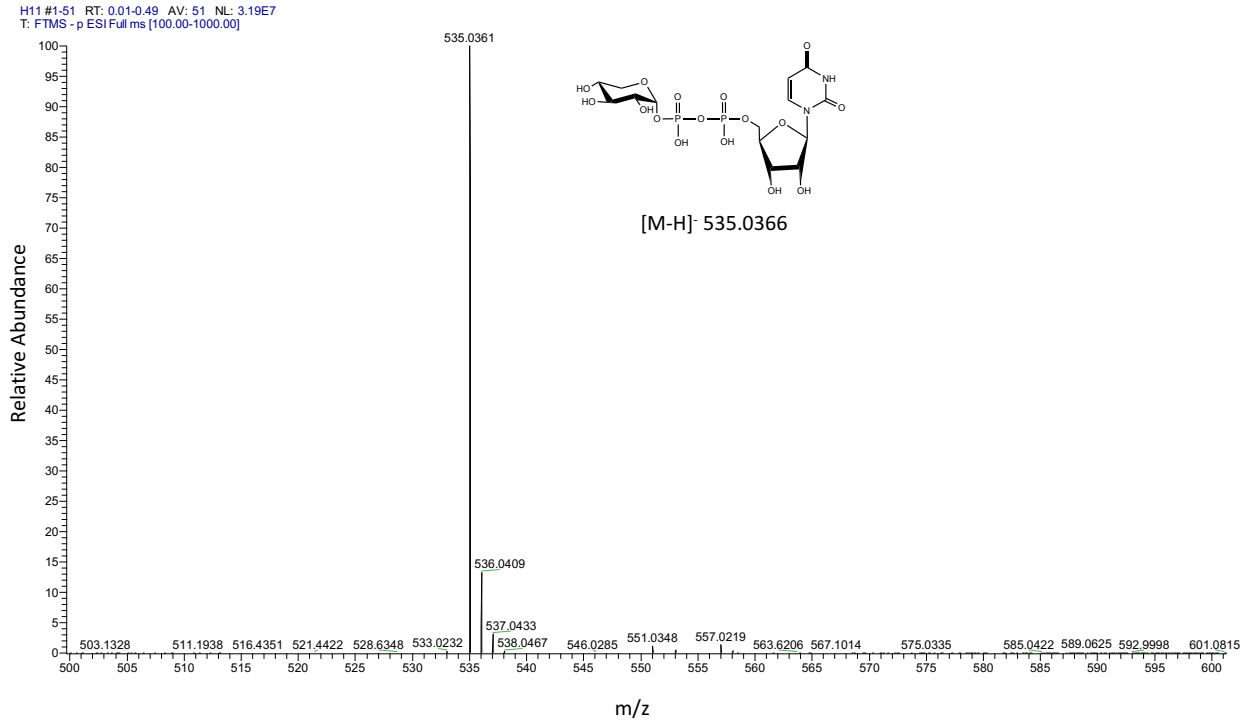


Figure 18: Mass spectrum of the *in vitro* enzymatically synthesized and purified UDP-xylose sugar donor. The sample was analyzed in negative mode and the calculated *m/z* for UDP-xylose was 535.0366.

3.3.3 High-throughput combinatorial screening of the UGT collection with flavonol/ anthraquinone acceptors and UDP-sugar donors

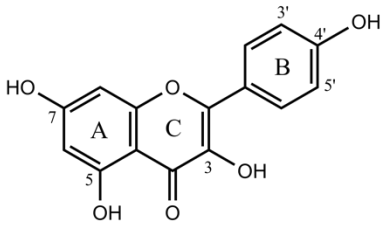
The novel high-throughput assay that was developed for GTs acting on lipophilic small molecule acceptors with natural fluorescent or chromogenic properties (**Chapter 2**) was used to combinatorially screen the purified recombinant UGT enzyme collection against a group of selected substrates. Herein, three flavonoids that belong to the flavonol subcategory (kaempferol, quercetin, and myricetin) and two anthraquinones (emodin and ε-rhodomycinone) were used as acceptors and UDP-glucose, UDP-galactose, and UDP-xylose were the sugar donors (**Figure 19**).

In the high-throughput combinatorial screen, for each UGT (0.25 mg/mL) reactions were carried out in a buffered mix (6.7 mM Tris buffer, pH 7.8) containing different acceptor (1 mM) and donor (2 mM) combinations with a minimum of two replicates. Two negative controls, either without UGTs or sugar donors were carried out. The reactions were performed for 2 h at 30 ºC and terminated by denaturing the enzymes at 90 ºC for 3 min. If the tested UGTs can glycosylate the given flavonol or anthraquinone acceptors using UDP sugar donors, this would result in polar glycosylated products. Next, the unreacted nonpolar acceptor was removed from the polar glycoside by using ethyl acetate: chloroform (3:1 v/v) or (1:3 v/v) for flavonol or anthraquinone, respectively. Herein, the organic solvent extraction was facilitated by the use of a silicone impregnated filter paper that is impermeable to the aqueous phase but permeable to the organic phase. Subsequently, absorbance measurements of the aqueous phases were performed at selected wavelength ranges to detect any glycosylated product formation (**Figure A4**). When flavonol acceptors were used in the high-throughput screen, the absorbance signal of the glycosylated flavonols were enhanced by using 2-aminoethyl diphenylborinate (Saslowsky *et al*., 2005).

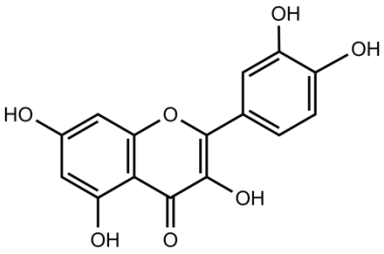
Figure 20 summarizes the results of the high-throughput combinatorial screen. According to the screen, if UGT enzymes were active in glycosylating the given acceptors with donors, an increase in absorbance signal was observed compared to the no enzyme and no sugar donor controls. The reactions with positive signals were then analyzed by HPLC-MS/MS for further confirmation and false positives were excluded. Using our high-throughput assay, we were able to identify previously unreported donor and acceptor specificities for certain UGTs of our UGT collection. Interestingly, our screening tool enabled us to detect three UGTs that were flexible in

synthesizing novel anthraquinone glycosides emphasizing the significance of our developed assay in *in vitro* glycodiversification, given the therapeutic value of anthraquinone drugs.

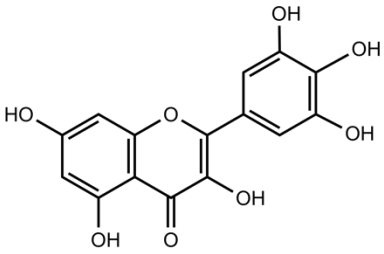
Flavonol acceptors



Kaempferol

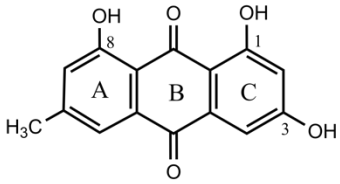


Quercetin

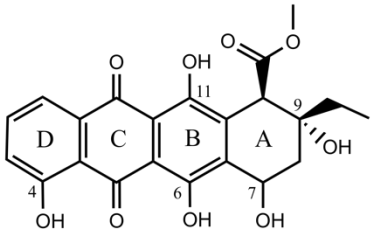


Myricetin

Anthraquinone acceptors

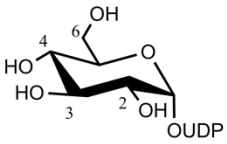


Emodin

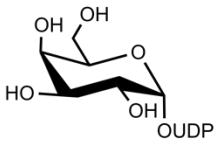


epsilon-Rhodomycinone

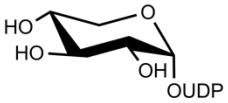
UDP-Sugar donors



UDP-Glucose



UDP-Galactose



UDP-Xylose

Figure 19: Chemical structures of the aglycone acceptors and sugar donors used in combinatorial high-throughput screen. The three flavonols differ from one another by the number of hydroxyl groups on B-ring, and the numbering of quercetin and myricetin is similar to that of kaempferol. Similarly, the numbering of UDP-galactose and UDP-xylose is similar to that of UDP-glucose.

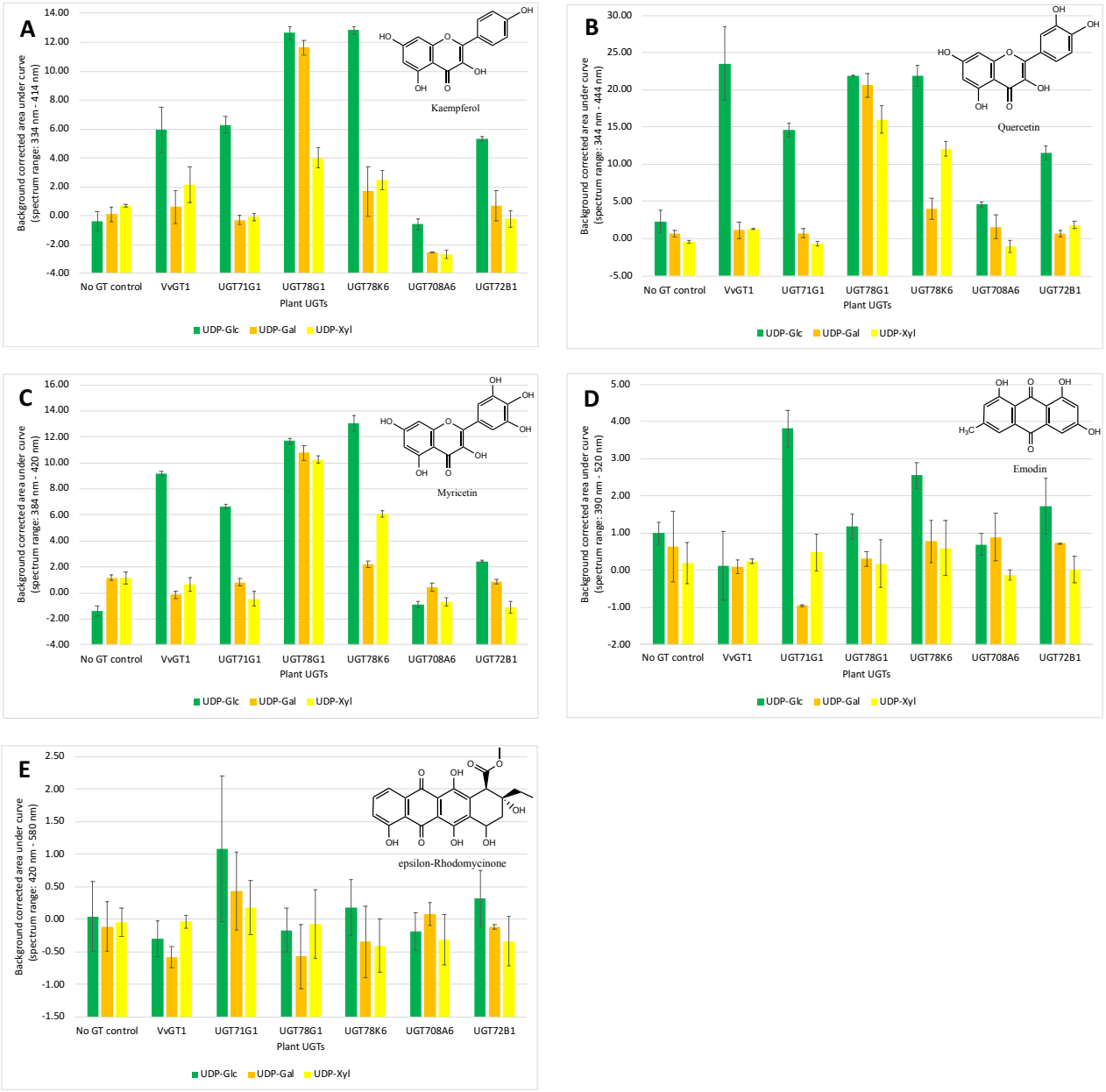


Figure 20: *In vitro* glycosylation activity of the recombinant UGT collection towards different donor and acceptor **(A)** kaempferol, **(B)** quercetin, **(C)** myricetin, **(D)** emodin, and **(E)** ε-rhodomycinone combinations. Screening was performed using the novel BiG HiT assay and 50 mM Tris pH 7.5, 150 mM NaCl was used as the no GT control. Each reaction was performed with a minimum of two replicates and error bars represent standard deviation.

3.3.3.1 Novel sugar donor specificities of the UGT collection

Sugar donor specificities of the UGT collection were analyzed using UDP-glucose, UDP-galactose, and UDP-xylose against different flavonol and anthraquinone acceptors.

The preferred sugar donor substrate of many plant UGTs is UDP-glucose and this is in line with our observations. When the glycosylation activity of our UGT collection was screened using the combinatorial high-throughput assay, as illustrated by **Figure 20**, all six UGTs used UDP-glucose in glycosylating a few or many tested acceptor molecules. The *in vitro* glycosylation activities of VvGT1, UGT71G1, UGT78G1, and UGT78K6 in the presence of UDP-glucose as the sugar donor and a number of flavonol acceptors were previously reported (Achnine *et al*., 2005; Ford *et al*., 1998; Hiromoto *et al*., 2015; Modolo *et al*., 2007). *In vitro* assays of UGT708A6 revealed its potential in glycosylating flavanones using UDP-glucose (Ferreyra *et al*., 2013). Similarly, Edwards and coworkers described the use of UDP-glucose by UGT72B1 in conjugating aniline and phenolic acceptors (Loutre *et al*., 2003).

Our screen suggested that the substrate specificity of certain UGTs in the collection was not limited to the favoured sugar donor, UDP-glucose, as we observed that UDP-galactose was also used by three out of six enzymes at least in glycosylating two tested flavonol acceptors (**Figure 20A-C**). Earlier studies reported that the substitution of UDP-glucose with UDP-galactose significantly reduced the *in vitro* glycosylation activity of VvGT1 and UGT71G1 in the presence of quercetin (He *et al*., 2006; Offen *et al*., 2006) and we were able to detect a similar observation for VvGT1. Nevertheless, our screen did not reveal the reported subtle glycosylation activity of UGT71G1 that was previously observed using UDP-galactose and quercetin as substrates by a relatively long run of reverse phase HPLC analysis (He *et al*., 2006). The extremely low level of glycosylation activity of UGT71G1 and the higher sensitivity of HPLC detection could be a possible explanation towards this discrepancy. Conversely, for UGT78G1, substantial glycosylation activity was detected with UDP-glucose and UDP-galactose against flavonols as described in prior work (Modolo *et al*., 2007). Interestingly, UGT78K6 showed novel substrate specificity towards UDP-galactose with all three flavonol acceptors (**Figure 20A-C and Figure A5A**). However, compared to its favoured sugar donor UDP-glucose, the glycosylation activity was lower in the presence of UDP-galactose. The orientation of the 4-OH group of the galactose moiety is different from that in UDP-glucose and Wang and colleagues suggested that this opposite

positioning could interfere with the enzyme-donor interaction (Shao *et al*., 2005). While this was proposed for UGT71G1, we can speculate a similar rationale for UGT78K6.

In the tested UGT collection, only two enzymes were previously reported to use UDP-xylose *in vitro*. This includes VvGT1 that was demonstrated to have a reduced catalytic efficiency in the formation of quercetin-xyloside (Offen *et al*., 2006) and UGT72B1 that produced DCA-xyloside and TCP-xyloside (Brazier-Hicks and Edwards, 2005). However, with the exception of UGT708A6, we were able to detect the potency of the remaining three UGTs: UGT71G1, UGT78G1, and UGT78K6 towards UDP-xylose *in vitro* (**Figure 20 and Figure A5B-C**). The latter two enzymes exhibited a considerable activity with this novel donor substrate at least in the presence of some flavonol acceptors. The 6-OH of UDP-glucose may be needed for the overall glycosylation activity of UGT78G1 and UGT78K6, yet it may not be vital as these UGTs tolerate UDP-xylose.

3.3.3.2 Novel acceptor specificities of the UGT collection

The high-throughput combinatorial screening of the recombinant UGT collection against three flavonol and two anthraquinone acceptors enabled characterizing the acceptor substrate flexibilities.

Many UGTs of our enzyme collection were characterized as flavonoid-glycosyltransferases since these enzymes showed activity with several flavonoids *in vitro* (Ford *et al*., 1998; Hiromoto *et al*., 2015; Modolo *et al*., 2007). As summarized in **Figure 20A-C**, all six UGTs were shown to accept a minimum of two of the flavonols tested at least when UDP-glucose was the sugar donor. These observations were consistent with previous studies on VvGT1, UGT71G1, UGT78G1, and UGT78K6 (Ford *et al*., 1998; Hiromoto *et al*., 2015; Modolo *et al*., 2007). In spite of having multiple hydroxyl groups on flavonols, VvGT1, UGT78G1, and UGT78K6 were reported to synthesize flavonol-3-*O*-glycosides which was reinforced by crystallographic studies (Hiromoto *et al*., 2015; Modolo *et al*., 2009; Offen *et al*., 2006). However, in terms of regioselectivity, UGT71G1 was identified to be versatile where *in vitro* glycosylation with UDP-glucose and quercetin synthesized all 5 mono *O*-glucosylated products with quercetin-3'-*O*-glucoside being the preferred glycoside. Molecular docking studies presented that all 5

possible *O*-glycosylation sites of quercetin could be docked in the vicinity of the catalytic base His-22 due to the spacious acceptor binding pocket of the enzyme (Shao *et al*., 2005).

The tested flavonols in our combinatorial screen were different from one another by the number of hydroxyl groups on the ring B (**Figure 19**). Davies and colleagues provided a detailed account of VvGT1 crystal structures with quercetin and kaempferol separately. According to their observations, the accommodation of structurally similar flavonols is possible as quercetin and kaempferol were reported to have similar interactions with the enzyme (Offen *et al*., 2006). As described by Wang and colleagues, the acceptor binding pattern of UGT78G1 was similar to that of VvGT1 underlining its ability to have a similar acceptor profile (Modolo *et al*., 2009). The observations for UGT78K6 were supported by the previous structural analysis with kaempferol acceptor that revealed the enzyme's possibility to occupy flavonols with altered substitutions to ring B (Hiromoto *et al*., 2015).

Interestingly, UGT708A6 and UGT72B1 used in our screen showed novel *in vitro* substrate tolerance towards at least two tested flavonols, yet the activity was subtle compared to the above explained UGTs (**Figure 20A-C and Figure A5D-E**). However, Casati and colleagues were not able to observe quercetin or kaempferol glycoconjugates when these acceptors were fed exogenously to UGT708A6-expressing *E.coli* cells (Ferreyra *et al*., 2013). In our own analysis we were also unable to detect glycosyltransferase activity with kaempferol, yet we observed that the enzyme showed low levels of *in vitro* flexibility towards quercetin and myricetin. UGT708A6 was reported to utilize flavanones such as naringenin and flavones such as apigenin towards producing *O*- and *C*-linked glycosides, respectively (Ferreyra *et al*., 2013). Thus, it would be interesting to know the glycosylation pattern of the resulted flavonol glycosides. In a previous *in vitro* study on UGT72B1, Edwards and coworkers did not observe glycosylated flavonol products using quercetin, kaempferol, and myricetin as acceptor substrates (Loutre *et al*., 2003). Yet, our screen showed some level of tolerance of this non-flavonoid GT towards the same three flavonols. This discrepancy could be due to several factors such as the differences in used enzyme and substrate concentrations, and the incubation time.

To our surprise, five out of six UGTs of our UGT collection showed novel *in vitro* acceptor specificities at least with one of the anthraquinones (**Figure 20D-E and Figure A5F-G**). VvGT1, UGT71G1, UGT78G1, UGT78K6, and UGT72B1 were capable in glycosylating the plant derived

anthraquinone emodin that has some structural similarities to flavonols where both comprise of a tricyclic skeleton with multiple phenolic hydroxyl groups. These structural resemblances may enable the accommodation of emodin in the acceptor pocket of the enzyme. Furthermore, our screening results confidently identified UGT71G1, UGT78K6, and UGT72B1 as capable of catalyzing the *in vitro* synthesis of glycosides of the anthraquinone ε-rhodomycinone. Earlier structural analysis of UGT71G1 demonstrated its spacious acceptor binding cavity that can occupy structurally large native acceptors such as medicagenic acid and hederagenin triterpenes (Shao *et al*., 2005). Therefore, the tetracyclic ε-rhodomycinone may fit into the acceptor binding site leading towards the production of its corresponding glycoconjugates.

3.3.3.3 *In vitro* glycodiversification of anthracyclines

The significance of sugar moieties on glycosylated natural products highlights the importance of glycodiversification towards the production of enhanced or altered chemical characteristics and biological activities. The high-throughput screening of our substrate flexible UGT collection against the ε-rhodomycinone anthraquinone acceptor with the UDP-sugar donor panel resulted in the production of novel glycosides. As depicted in **Figure 20E**, ε-rhodomycinone-glucoside was produced by UGT71G1, UGT78K6, and UGT72B1 and ε-rhodomycinone-xyloside was produced by UGT71G1. When these glucosides and xyloside were analyzed by LC-MS/MS, distinctive fragmentation profiles were observed by the loss of glucose (162 *m/z*) (**Figure 21A-C**) and xylose (132 *m/z*) (**Figure 21D**) moieties. The ε-rhodomycinone has received clinical significance as it is the aglycone precursor of the anticancer anthracyclines including doxorubicin and epirubicin (Minotti *et al*., 2004). The only structural difference between these two chemotherapeutic is on the sugar residue that accounts for the reduced cardiotoxicity of one over the other. Hence, the resultant novel ε-rhodomycinone glycosides by our high-throughput combinatorial screen may possess therapeutic importance.

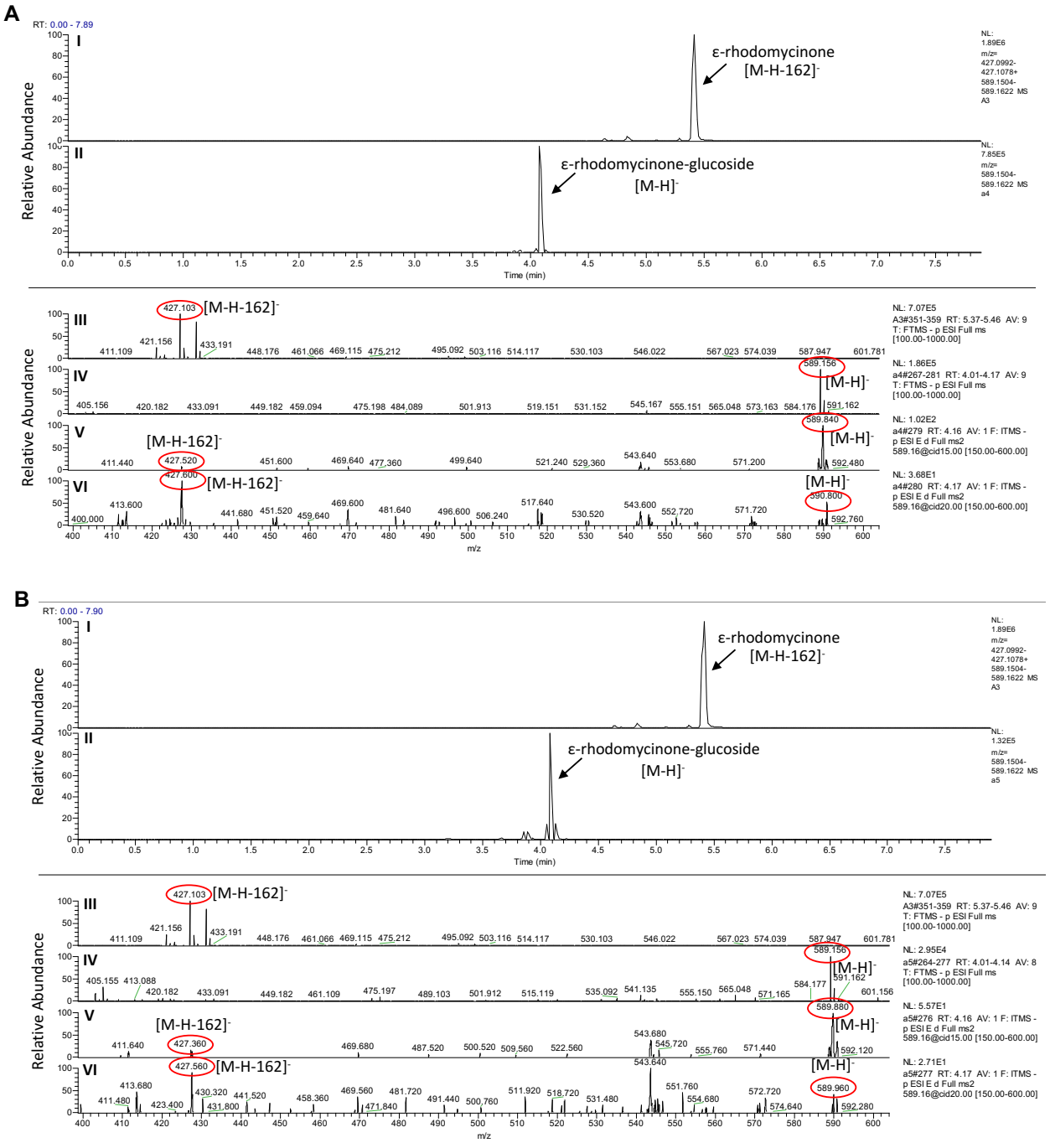


Figure 21: LC-MS/MS analysis of **(A)** UGT71G1 and **(B)** UGT78K6 activity with ε-rhodomycinone and UDP-glucose. Panel I and II: LC chromatograms of no UGT control and in the presence of UGT, respectively. Panel III-VI: mass spectra of ε-rhodomycinone (of no GT control), ε-rhodomycinone-glucoside, and MS/MS fragmentation profiles of ε-rhodomycinone-glucoside at 15 and 20 normalized collision energies (cid), respectively. Sample were analyzed in negative mode and the calculated *m/z* for ε-rhodomycinone and ε-rhodomycinone-glucoside were 427.103 and 589.156, respectively.

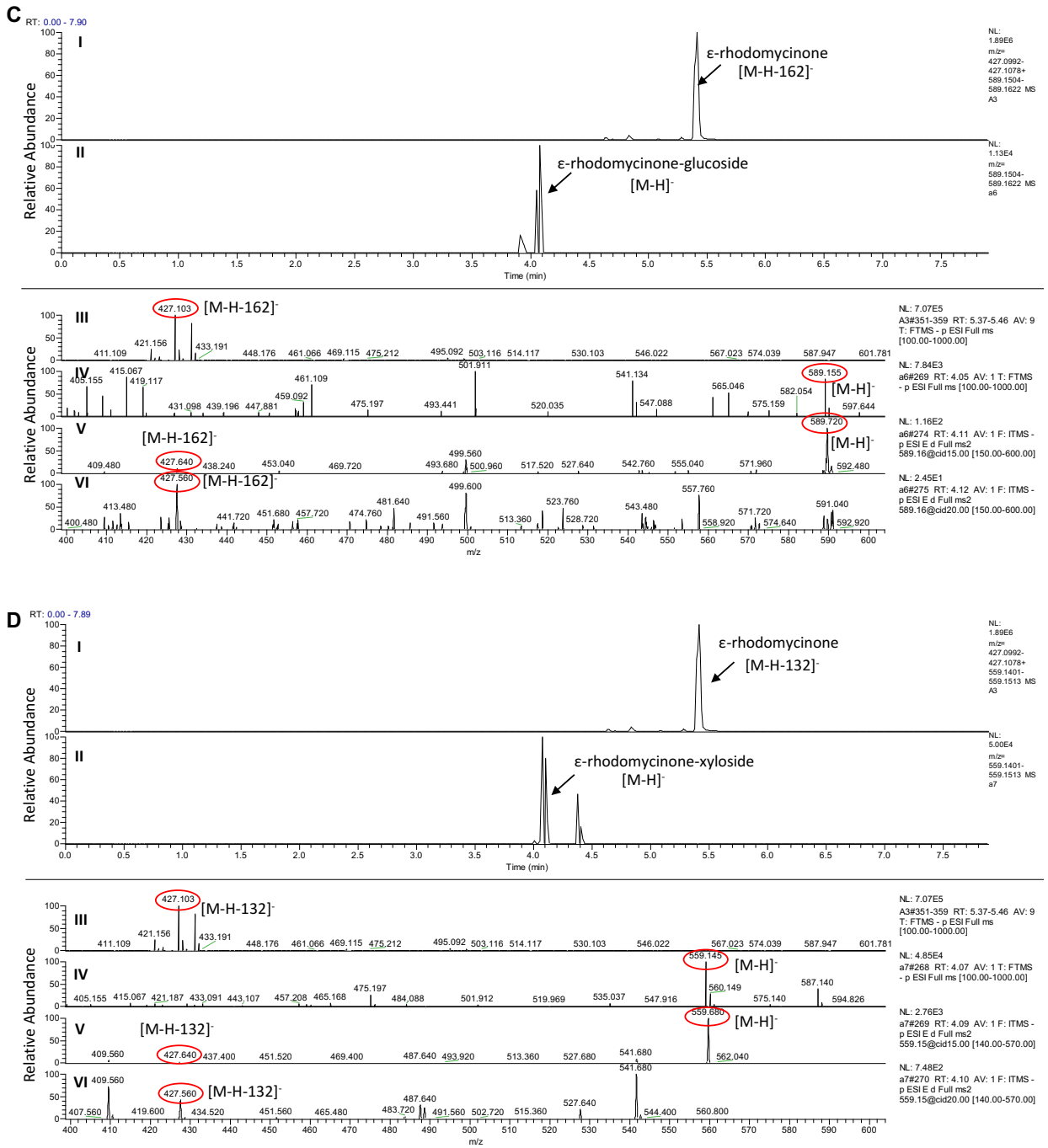


Figure 21: LC-MS/MS analysis of **(C)** UGT72B1 activity with ε-rhodomycinone and UDP-glucose and **(D)** UGT71G1 activity with ε-rhodomycinone and UDP-xylose. Panel I and II: LC chromatograms of no UGT control and in the presence of UGT, respectively. Panel III-VI: mass spectra of ε-rhodomycinone (of no GT control), ε-rhodomycinone-glycoside, and MS/MS fragmentation profiles of ε-rhodomycinone-glycoside at 15 and 20 cid, respectively. Sample were analyzed in negative mode and the calculated *m/z* for ε-rhodomycinone, ε-rhodomycinone-glucoside, ε-rhodomycinone-xyloside were 427.103, 589.156, and 559.145, respectively.

3.4 Conclusions

The inherent substrate flexibility of many natural product GTs could be harnessed towards glycodiversification that aids in the development and improvement of small molecule natural product therapeutics. Herein, using our high-throughput assay, screening a promiscuous plant UGT collection against a panel of flavonol and anthraquinone acceptors and UDP-sugars enabled us to characterize novel substrate specificities of certain UGTs. These include novel sugar donor tolerance towards UDP-galactose, UDP-xylose and novel acceptor tolerance towards flavonols and anthraquinones. Though we have made a considerable improvement in characterizing the substrate flexibility of our UGT collection, further studies are needed to identify the regioselectivity of the produced glycosides.

One of the key findings of our study is the identification of three plant UGTs: UGT71G1, UGT78K6, and UGT72B1 as being useful for the production of novel ε-rhodomycinone glycosides. Considering the pharmaceutical value of the clinically used anthraquinone glycosides that are biosynthesized from ε-rhodomycinone precursor, and the significance that the sugar moiety has on the biological activity of these drugs, our results are of particular importance towards the glycodiversification of therapeutics in this class. The GTs identified through our strategy and the novel compounds that they produce show promise as biocatalytic tools and new therapeutics.

The substrate flexibility is not limited to certain GTs but also found in a number of sugar biosynthetic enzymes involved in the production of nucleotide activated sugar donors. The subsequent chapter highlights the importance of these promiscuous sugar biosynthetic enzymes towards the *in vitro* engineering of an enzymatic pathway leading to the production of sugar precursors of therapeutic value.

CHAPTER FOUR

An engineered *in vitro* enzymatic pathway for the anticancer anthracycline deoxysugars: TDP-L-daunosamine and TDP-L-acosamine

4.1 Background

A considerable number of small molecule natural products of clinical therapeutic importance are glycosylated (Weymouth-Wilson, 1997). Interestingly, the sugar moieties that decorate many of the secondary metabolites derived from bacteria are heavily modified and deoxygenated (Thibodeaux *et al*., 2008). For example, anthracyclines such as daunorubicin and doxorubicin are decorated with a 2,3,6-trideoxy-3-aminohexose, L-daunosamine (**Figure 22**) (Otten *et al*., 1997). Doxorubicin is used for treating a number of solid and blood cancers whereas the use of daunorubicin is mainly limited to acute lymphoblastic leukemia and myeloblastic leukemia (Minotti *et al*., 2004; Weiss *et al*., 1986). The unusual glycosyl residues of small molecule natural products including in these anthracyclines have an important effect on their biological properties.

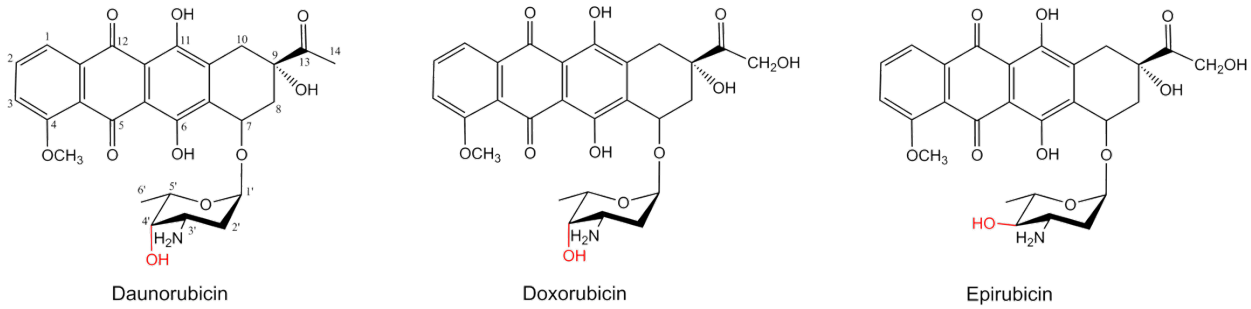


Figure 22: Chemical Structures of daunorubicin, doxorubicin, and epirubicin anticancer anthracyclines. The tetracyclic aglycone residue of doxorubicin and epirubicin are similar whereas daunorubicin lacks the C-14 hydroxyl group. The sugar moiety of doxorubicin and daunorubicin is the deoxyhexose daunosamine whereas epirubicin has an epimer of daunosamine. The stereochemical difference between sugar residues are indicated in red.

The high-value anticancer drug, epirubicin shows a broad-spectrum anticancer activity against a number of different cancers (Cortés-Funes and Coronado, 2007). It has become an attractive chemotherapeutic compared to its structural counterpart doxorubicin due to the reported lower cardiotoxicity (Lahtinen *et al*., 1991; Weiss *et al*., 1986). As shown in **Figure 22**, the sole difference between epirubicin and doxorubicin is in the sugar moiety, L-acosamine and L-daunosamine, respectively. The corresponding nucleotide-activated sugar precursors of these anthracyclines are TDP-L-acosamine and TDP-L-daunosamine. The vital functions played by heavily modified deoxysugars on secondary metabolites have spurred interest in unraveling their underlying biosynthetic pathways.

The daunorubicin/doxorubicin biosynthetic pathway of *S. peucetius* was extensively studied and structural gene products for the synthesis of donor TDP-L-daunosamine and aglycone acceptor ε-rhodomycinone were identified (Lomovskaya *et al*., 1999). Although fermentation of the *S. peucetius* produced a considerable yield of the aglycone precursor, the *in vivo* production of TDP-L-daunosamine is reported to be low (Malla *et al*., 2010).

The therapeutic value of epirubicin has motivated efforts in medicinal chemistry towards developing synthetic schemes for its production. Conventionally, it is produced semisynthetically through several organic synthetic steps (Arcamone *et al*., 1975; van der Rijst *et al*., 1999), however, an economical biological route would be of considerable value. Enzymatic synthesis of epirubicin, either in engineered bacterial strains or *in vitro*, could be achieved by using ε-rhodomycinone aglycone, TDP-L-acosamine sugar donor, and a potent glycosyltransferase that can act on these substrates. Nevertheless, a natural biosynthetic pathway for TDP-L-acosamine has not been described (Thibodeaux *et al*., 2007). Chemists have synthesized this precursor via organic synthetic routes (Oberthür *et al*., 2004). However, it is a tedious process with multiple chemical conversions. Even though *in vivo* enzymatic pathways for TDP-L-acosamine has been engineered in previous works by altering the *in vivo* natural biosynthetic pathway of TDP-L-daunosamine, the authors of these studies did not attempt to isolate the sugar donor (Han *et al*., 2011; Madduri *et al*., 1998; Wang *et al*., 2018a).

The significance of TDP-L-daunosamine and TDP-L-acosamine highlights their potential to serve as precursors towards *in vitro* glycodiversification of small molecules employing GTs via either *in vitro* combinatorial screening or enzymatic engineering. Towards this, we have

reconstituted an *in vitro* enzymatic pathway for the synthesis of these two sugar donors, taking advantage of the substrate flexibility of sugar biosynthetic enzymes. Our *in vitro* pathway was built on the natural biosynthetic pathway of TDP-L-daunosamine.

As illustrated in **Figure 23**, the *in vitro* pathway was divided into two main parts, where the part shown in Scheme **A** comprises of four enzymatic reactions leading to the synthesis of the key intermediate TDP-4-keto-6-deoxy-D-glucose (TKDG or compound **1**). Herein, glucose-1-phosphate and TMP were used as precursors. Subsequently, starting from compound **1**, the pathway was extended to produce TDP-L-daunosamine and TDP-L-acosamine (Scheme **B)**. Except for the difference in the orientation of the C-4′ hydroxyl group, the amino hexose in TDP-L-acosamine is structurally similar to that in TDP-L-daunosamine. Therefore, excluding the last ketoreductase reaction, the remaining three enzymatic steps towards the production of these two sugar donors were identical. In total, nine overexpressed and purified sugar biosynthetic enzymes were used for the engineered *in vitro* pathway. First, compound **1** was enzymatically synthesized in a one pot reaction and two purification steps were performed. Next, starting from this purified key intermediate, TDP-L-daunosamine and TDP-L-acosamine were synthesized separately in a one-pot two-step reaction. Along with TDP-L-daunosamine, we were able to synthesize and purify few milligrams of the first amino hexose intermediate (compound **3**) of the pathway and it was used in enzymatic characterization.

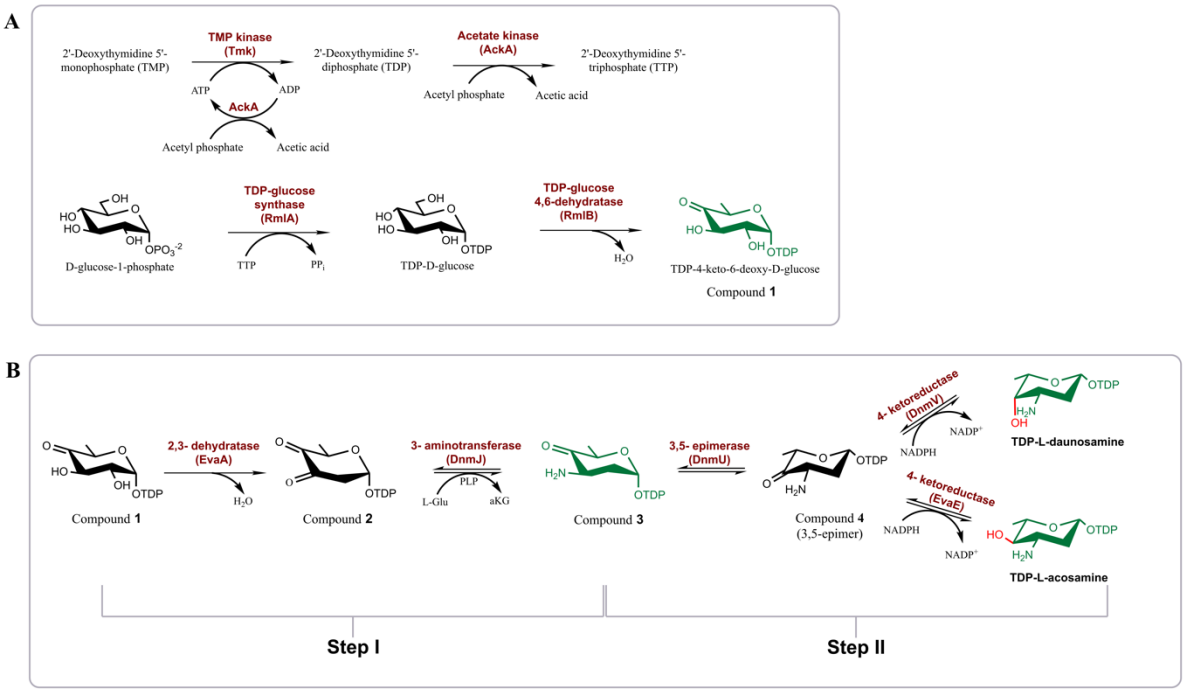


Figure 23: The reconstituted enzymatic pathway for the *in vitro* synthesis of TDP-L-daunosamine and TDP-L-acosamine. Scheme **A:** The key intermediate TKDG (compound **1**) was synthesized based on a previously published protocol with necessary modifications (Oh *et al*., 2003). Scheme **B** illustrates the synthesis of TDP-L-daunosamine and TDP-L-acosamine separately in a one-pot two-step reaction. The sugar donors indicated in green were synthesized in this study. PPi: pyrophosphate, L-Glu: L-glutamate, aKG: α-ketoglutarate, PLP: pyridoxal 5'-phosphate

4.2 Materials and Methods

4.2.1 Chemical reagents and enzymes

Chemicals used in this study were of analytical grade and were commercially purchased from Sigma-Aldrich, Carbosynth, Bio Basic Canada, and Bio-Rad. Restriction endonucleases, T4 DNA ligase, Antarctic phosphatase, and Phusion DNA polymerase were from New England Biolabs. DNase I, RNase A, and lysozyme were purchased from Bio Basic Canada.

4.2.2 Plasmids, bacterial strains, and culture conditions

Plasmids and bacterial strains of this chapter are summarized in **Table A2** and **A3**, respectively. Growth and maintenance of *E. coli* strains with appropriate antibiotics and preparation of chemically competent cells were done as **Section 2.2.2**.

4.2.3 Construction of plasmids encoding TDP-L-daunosamine and TDP-L-acosamine biosynthetic enzymes

Individual plasmid constructs encoding TDP-L-daunosamine and TDP-L-acosamine biosynthetic enzymes in pET28a expression vector were prepared as follows. Genomic DNA (gDNA) from *E. coli* K12, *A. orientalis*, and *S. peucetius* were isolated using the standard phenol/ chloroform extraction protocol. Desired genes were PCR amplified (**Section 3.2.3**) using gDNA as the template, and the used forward and reverse primers were listed in **Table A1**. Herein, the templates for *tmk*, *ackA*, and *rmlA* genes were *E. coli* K12 gDNA, for *rmlB* gene was a synthesized *rmlB* gene, for *evaA* and *evaE* genes were *A. orientalis* gDNA, and for *dnmJ*, *dnmU*, and *dnmV* genes were *S. peucetius* gDNA. The PCR products were restriction digested at *Nde*I and *Xho*I sites, cloned into individual pET28a, and sequence verified.

4.2.4 Expression and purification of recombinant proteins of TDP-L-daunosamine and TDP-L-acosamine enzymatic pathway

Except for *evaA* and *dnmJ* constructs, the rest of genes coding N-terminally His_6-tagged sugar biosynthetic enzymes in pET28a vector were individually used to transform chemically competent *E. coli* BL21(DE3) cells, whereas chemically competent *E. coli* BL21(DE3) codon plus

L1SL2/pETcoco-2 cells (Moncrieffe *et al*., 2012) were transformed with plasmids for *evaA* and *dnmJ* expression. For protein expression, 1 % of starter cultures that had grown overnight at 37 °C were inoculated into individual 800 mL volumes of LB broths supplemented with appropriate antibiotics (50 µg/mL of kanamycin was used for BL21(DE3) cells harbouring pET28a constructs, and 100 µg/mL ampicillin, 25 µg/mL chloramphenicol, and 50 µg/mL of kanamycin were used for the codon plus cells with pET28a constructs). The culture conditions for each protein were as follows. For Tmk, AckA, RmlA, RmlB, and DnmU cultures were incubated at 37 ºC until OD_{600} of 0.4 - 0.8 was reached and induced with 0.5 mM IPTG. When EvaA culture reached an OD_{600} of 0.2 and 0.6 upon incubation at 37 ºC, 0.01% L-arabinose and 1 mM IPTG were added, respectively. The culture conditions of DnmJ were same as EvaA except 0.02 mM IPTG was used. For DnmV, cultures were incubated at 37 ºC until OD_{600} of 0.6 was reached and induced with 0.05 mM IPTG. Upon IPTG addition, all the cultures were grown overnight at 16 ºC. EvaE culture was incubated at 15 ºC for 72 h without IPTG induction.

Bacterial cell harvest and lysis, protein purification, protein concentration (using appropriate MWCO), buffer exchange, and storage were performed as explained in **Section 2.2.3** and **3.2.4.2**. Tmk, AckA, RmlA, RmlB, and EvaE proteins were isocratic purified by gravity and EvaA, DnmJ, DnmU, and DnmV were gradient purified using AKTA purifier 10 system (GE Healthcare Life Sciences). Buffer I was used for Tmk, AckA, RmlA, RmlB, DnmU, DnmV, and EvaE, and buffer VI was used for EvaA and DnmJ (**Table A5**).

4.2.5 *In vitro* enzymatic synthesis and purification of TDP-L-daunosamine and two intermediates of the engineered pathway

4.2.5.1 TDP-4-keto-6-deoxy-D-glucose (TKDG)

Enzymatic synthesis of TKDG (compound **1**) was done in a one-pot reaction according to the procedure described by Lee and colleagues (Oh *et al*., 2003), with minor modifications. A 30 mL buffered reaction (50 mM Tris-HCl pH 7.5), containing 10 mM TMP, 15 mM glucose-1-phosphate, 20 mM $MgCl_2$, 1 mM ATP, 40 mM lithium acetyl phosphate, and 0.1 mg/mL of each four enzymes (TmK, AckA, RmlA, and RmlB) were incubated at 37 ºC for 1.5 h. Subsequently, the reaction was terminated by denaturing the proteins at 90 ºC for 3 min, and insoluble contents were pelleted out at 3,000 x g for 10 min.

The synthesized TKDG was purified by anion exchange chromatography using an 8 mL Macro-prep high Q column (Bio-Rad) connected to an AKTA purifier 10 system (GE Healthcare Life Sciences). The column was equilibrated with 5 column volumes of deionized water (buffer A) and the sample was loaded (10 mL loading volume) at a flow rate of 1 $mLmin^{-1}$. Next, the compounds were eluted at the same flow rate with a gradient of 0 % to 100 % B over 5 column volumes, and 100 % B over 1 column volume where buffer B was buffer VII (**Table A5**). The eluting compounds were monitored by UV absorption at 254 nm and 280 nm. Next, the desired compound was verified by thin layer chromatography (TLC) using TLC Silica gel 60 plates (EMD millipore) with ethyl acetate: methanol: dH_2O: acetic acid (3:2:1:0.5 v/v) mobile phase and stained by *p*-anisaldehyde stain (**Table A5**). Subsequently, TKDG containing pure fractions were pooled, concentrated to 1 mL using a rotary evaporator at 40 ºC (BUCHI rotavapor R-114) and further verified by MS as described in **Section 3.2.5**.

Then, size exclusion chromatography was performed for desalting. Herein, the concentrated TKDG sample was loaded to a pre-equilibrated (with 2 column volumes of deionized water) 100 mL Bio-Gel P-2 (Bio-Rad) column attached to an AKTA purifier 10 system (GE Healthcare Life Sciences). Using deionized water as the eluent, elution was done at a flow rate of 0.2 $mLmin^{-1}$ over 2.2 column volumes and fractions were monitored at 254 nm and 280 nm absorbance. Following TLC verification and MS analysis as mentioned above, TKDG containing fractions were pooled and lyophilized (**Section 3.2.5**).

4.2.5.2 Compound 3

A 20 mL reaction containing 50 mM Bis-Tris Propane pH 7.5, 1.5 mM compound **1**, 75 mM glutamate, 0.15 mM pyridoxal-5-phosphate (PLP), 0.038 mg/mL of EvaA, and 0.75 mg/mL of DnmJ was incubated for 3.5 h at 24 ºC. Next, termination of the reaction and removal of the insoluble contents were done as **Section 4.2.5.1**.

Subsequently, an AKTA purifier 10 system (GE Healthcare Life Sciences) equipped with a 22 mL Sephadex Q (Pharmacia) column was used for anion exchange purification using buffer VIII and IX (**Table A5**) as buffer A and B, respectively. Using 2 column volumes of buffer A, the column was pre-equilibrated, and 1 mL of sample was loaded. Fractions were eluted at 1 $mLmin^{-1}$ flow rate with the following elution profile: 0% B, 0% to 20% B, 20% to 100 % B, one column

volume each. The eluting compounds were monitored by UV absorption at 254 nm and 280 nm and the fractions containing compound **3** was neutralized with 500 mM NH_4OH. Following compound verification by MS (**Section 3.2.5**), pooled fractions were concentrated to 1-2 mL using a rotary evaporator (**Section 4.2.5.1**).

Desalting of the concentrated compound **3** was done using a 75 mL Sephadex G-10 (Sigma Aldrich) column as described in **Section 3.2.5**. Herein, elution profile was monitored by UV absorption at 254 nm and 280 nm. Subsequent verification of Compound **3** by MS and lyophilization of the pooled fractions were done as **Section 3.2.5**.

4.2.5.3 TDP-L-daunosamine

As explained in **Section 4.2.5.2**, a 20 mL reaction was setup for the production of compound **3** and incubated at 24 ºC for 3.5 h. Next, a 20 mL buffered reaction (50 mM Bis-Tris Propane pH 7.5) containing 30 mM NADPH, DnmU and DnmV (0.6 mg/mL each) was added to the same reaction and incubated for an additional 16 h at 24 ºC. Subsequently, reaction termination, insoluble content removal, anion exchange chromatography, compound verification, and lyophilization were done as mentioned in **Section 4.2.5.2**.

4.2.6 TDP-L-acosamine synthesis and MS verification

A 10 µL reaction was setup for the production of compound **3** (components and concentrations were same as **Section 4.2.5.2**) and incubated at 24 ºC for 3.5 h. Next, a 10 µL buffered reaction (50 mM Bis-Tris Propane pH 7.5) containing 30 mM NADPH, DnmU and EvaE (0.6 mg/mL each) was added to the same reaction and incubated at 24 ºC for an additional 1.5 h. Upon reaction termination and insoluble component removal (**Section 4.2.5.1**), MS analysis was performed for product verification (**Section 3.2.5**).

4.2.7 DnmJ aminotransferase kinetic assay

The kinetic parameters of DnmJ were determined by coupling the aminotransferase reaction to an NAD^+/NADH and resazurin/resorufin reactions in a 96-well half area black plate (Corning) with

three replicates. In a total of 60 μL buffered reaction (50 mM Bis-Tris Propane pH 7.5) containing 1 mM α-ketoglutarate, 0.1 mM PLP, 0.5 mM NAD^+, 12.5 μM resazurin, 5 U/mL L-GDH, 0.02 mg/mL diaphorase with purified compound **3** ranging from 0.025 -0.50 mM were incubated at 37 ºC for 10 min. The reaction was initiated upon addition of 1 mg/mL DnmJ and fluorescence readings (525 nm excitation and 598 nm emission) were obtained at 37 ºC over a period of 140 min using a CLARIOstar monochromator microplate reader (BMG Labtech). Two controls, one without compound **3** and one without DnmJ (for each compound **3** concentration) were performed. The data were received from MARS data analysis software 3.20 R2 and data were analyzed using GraFit version 7 software.

4.3 Results and discussion

4.3.1 Cloning and recombinant expression of TDP-L-daunosamine and TDP-L-acosamine biosynthetic enzymes

An important characteristic of many sugar biosynthetic enzymes is the capacity to accept non-native substrates. Our proposed pathway towards the production of TDP-L-daunosamine and TDP-L-acosamine was driven by nine substrate flexible sugar biosynthetic enzymes. Three enzymes were from daunorubicin/doxorubicin biosynthetic pathway of *S. peucetius* whereas the rest were from different strains including *E. coli* K12, *S. enterica*, and *A. orientalis*. We cloned the corresponding sugar biosynthetic genes into pET expression vectors using traditional cloning methods. The soluble and active proteins were overexpressed in *E. coli* strains, and His_6 tagged proteins were nickel affinity purified (**Figure 24**). For each protein, an 800 mL of culture resulted in a protein yield of 18 mg of Tmk, 35 mg of AckA, 24 mg of RmlA, 31 mg of RmlB, 3 mg of EvaA, 12 mg of DnmJ, 3 mg of DnmU, 1.6 mg of DnmV, and 1.2 mg of EvaE. These purified enzymes were used in subsequent deoxysugar biosynthesis.

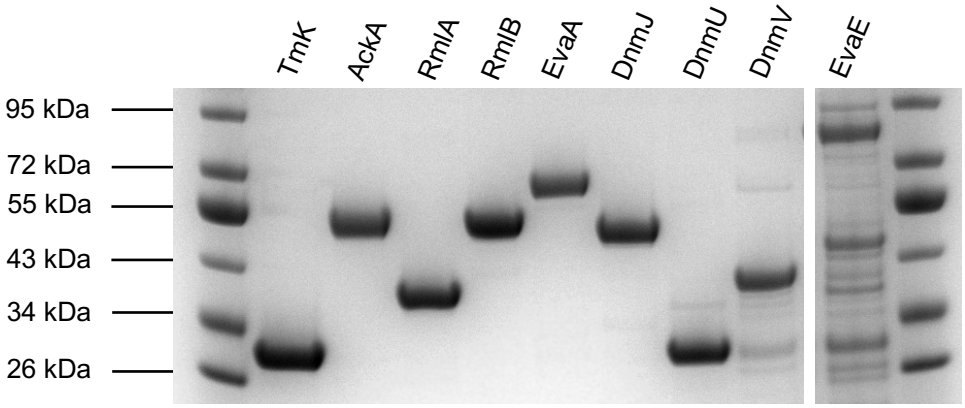


Figure 24: SDS-PAGE (4-12% Bis-Tris gel) of purified sugar biosynthetic proteins. The proteins were overexpressed in *E. coli* strains and Ni-NTA purified. Following SDS-PAGE (0.25 – 0.50 mg/mL of final protein concentrations were loaded), the proteins were visualized by staining with Coomassie blue (**Table A5**). The molecular weights of the proteins were as follows. TmK: 24 kDa, AckA: 43 kDa, RmlA: 32 kDa, RmlB: 40 kDa, EvaA: 53 kDa, DnmJ: 41 kDa, DnmU: 23 kDa, DnmV: 33 kDa, and EvaE: 34 kDa.

4.3.2 *In vitro* enzymatic synthesis and purification of TDP-4-keto-6-deoxy-D-glucose (TKDG or compound 1)

TDP-4-keto-6-deoxy-D-glucose (compound **1**) functions as an important common intermediate of many TDP-deoxysugar biosynthetic pathways (Thibodeaux *et al*., 2007). Lee and co-workers proposed a cost-effective *in vitro* enzymatic pathway for the production of compound **1** using crude enzyme lysates (Oh *et al*., 2003). As illustrated in **Figure 23**, the reaction begins from TMP, where it is converted to thymidine triphosphate (TTP) through TDP by two sequential reactions catalyzed by two enzymes that originate from *E. coli*: TMP kinase (Tmk) and acetate kinase (AckA), respectively. In addition to catalyzing the conversion of TDP to TTP, AckA is also used to re-generate adenosine triphosphate (ATP). Next, the TDP-glucose synthase enzyme from *E. coli* (RmlA) transfers a TMP moiety from TTP onto the glucose-1-phosphate substrate, producing TDP-glucose. Then, dehydration of TDP-glucose by TDP-glucose 4,6-dehydratase (RmlB from *S. enterica*) results in the desired compound **1**. In our study, crude enzyme lysates were replaced with

overexpressed and purified His-fusion proteins and compound **1** was produced in a 30 mL reaction with minor alterations to the proposed protocol of Lee and co-workers (Oh *et al*., 2003).

Following the synthesis of compound **1**, it was purified by two chromatographic techniques. First, the remaining reactants and intermediates were removed from compound **1** by applying the material to a Macro-prep high Q strong anion exchange column using a gradient elution profile of 0 mM – 1 M NaCl (**Figure A6A**). Subsequently, desalting was done using Bio-Gel P-2 (polyacrylamide) resin (**Figure A7A**). At each purification step the product was monitored via thin layer chromatography (TLC) using the mobile phase of ethyl acetate: methanol: dH_2O: acetic acid (3:2:1:0.5 v/v) (**Figure A6B** and **A7B**, respectively) and MS analysis. The lyophilized compound **1** resulted in 130.6 mg equivalent to a total yield of 80% (starting with 110 mg of TMP). Two peaks each corresponding to compound **1** (*m/z* 545.058), and the hydrated form of compound **1** at the C-4 position (*m/z* 563.068) was observed in the MS spectrum (**Figure 25A**). This purified intermediate was used as the starting precursor in the following deoxysugar enzymatic synthesis.

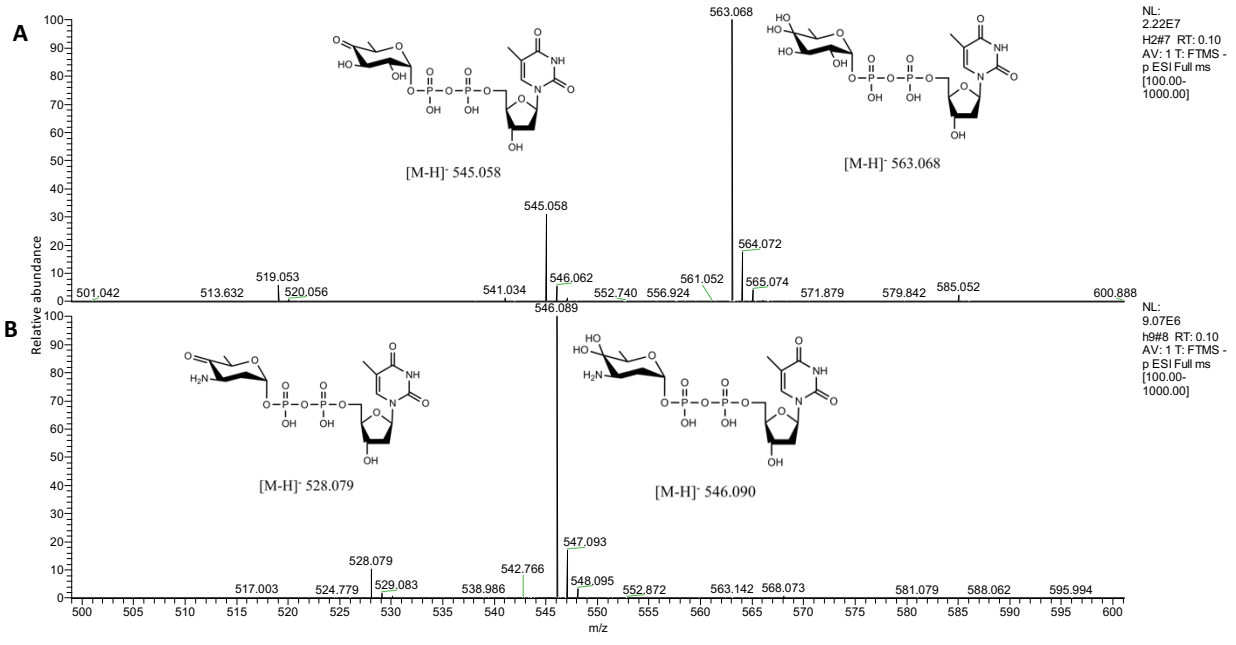


Figure 25: MS spectra of **(A)** compound **1** and **(B)** compound **3** synthesized using the reconstituted *in vitro* pathway. The samples were analyzed in negative mode and as detected by *m/z*, both compounds exist in its non-hydrated and hydrated form at C4 position of the sugar moiety.

4.3.3 The DnmJ catalyzed transamination reaction constitutes a bottleneck in the TDP-L-daunosamine/ TDP-L-acosamine pathway *in vitro*, which can lead to the shunt product formation

Starting from compound **1**, we have proposed an *in vitro* enzymatic pathway to synthesize TDP-L-daunosamine and TDP-L-acosamine based on the natural biosynthetic pathway of TDP-L-daunosamine in *S. peucetius* (**Figure 23B)**. Herein, the initial three enzymatic reactions towards the production of both sugar donors are identical. First, compound **1** is further dehydrated by a 2,3-dehydratase (EvaA from *A. orientalis*) towards the production of 2,6-dideoxy-3,4-diketo-TDP-sugar (compound **2**). EvaA is a homologue of DnmT, from the *S. peucetius* daunorubicin/doxorubicin pathway, which was used since the latter could not be expressed recombinantly from *E. coli* in soluble form. Next, the first TDP-aminosugar of the pathway (compound **3**) is synthesized from compound **2** by the action of a 3-aminotransferase from *S. peucetius* (DnmJ) in the presence of the coenzyme PLP and cosubstrate L-glutamate. Subsequently, compound **3** is epimerized by a 3,5-epimerase (DnmU from *S. peucetius*). The final step in our proposed pathway plays the key role in differentiating the stereochemistry of the C-4 hydroxyl group of the sugar donors. TDP-L-daunosamine has an axial C4-OH group whereas it is equatorial in TDP-L-acosamine. Herein, in a NADPH dependent reaction, DnmV 4-ketoreductase from *S. peucetius* or EvaE 4-ketoreductase from *A. orientalis* results in TDP-L-daunosamine or TDP-L-acosamine, respectively.

Our initial attempt in synthesizing TDP-L-daunosamine was done in a one-pot one-step reaction starting from compound **1** and adding all four enzymes (EvaA, DnmJ, DnmU, and DnmV). However, along with TDP-L-daunosamine (*m/z* of 530.095), an undesired product was observed by MS analysis with a *m/z* of 547.074 (**Figure A8A**). Analysis of different combinations of these four enzymes revealed that the production of this undesired sugar donor was driven by the action of three enzymes: EvaA, DnmU, and DnmV (**Figure A8B and C**). This suggests that the substrate flexibility of DnmU allows it to accept compound **2** in addition to compound **3** and the flexibility of DnmV in acting on the 5-epimer or 3,5-epimer of either compound **2** or compound **3** (which result from conversion of those intermediates by DnmU), respectively.

The shunt product formation indicates that the aminotransferase reaction is a critical bottleneck in driving the reaction towards TDP-L-daunosamine synthesis. Moreover, the product

of EvaA dehydratase, compound **2**, was reported to be unstable and prone to spontaneous decomposition into maltol and TDP (**Figure A9**) (Chen *et al*., 1999). Therefore, to favour the equilibrium of the reversible aminotransferase reaction towards compound **3**, the concentration of DnmJ enzyme was increased 20-fold compared to EvaA enzyme. Furthermore, with the decomposition of compound **2,** the equilibrium of the DnmJ-catalyzed transamination will favour reversal of the reaction, resulting in consumption of compound **3.** Hence, it is important to limit the incubation time of the transaminase reaction (Chen *et al*., 2000).

4.3.4 A one-pot two-step *in vitro* enzymatic pathway for the synthesis of TDP-L-daunosamine

Considering the shunt product formation, scheme **B** of the *in vitro* pathway (**Figure 23**) was separated into two sequential steps (**step I** and **II**) where **step I** involves the addition of EvaA and DnmJ to ensure the conversion of compound **1** to compound **3**. Subsequent addition of DnmU and DnmV to the same reaction mixture results in the production of the desired sugar donor, TDP-L-daunosamine.

4.3.4.1 *In vitro* enzymatic synthesis and purification of compound 3

Compound **3** is the first TDP-aminosugar intermediate of the pathway that could lead to the production of a range of important TDP-aminosugar donors. Apart from TDP-L-daunosamine and TDP-L-acosamine, this could serve as a precursor for TDP-4-*epi*-L-vancosamine, TDP-L-ristosamine, TDP-3-*N*-methyl-L-ristosamine, and TDP-L-megosamine (Han *et al*., 2011).

Compound **3** was synthesized in a 20 mL reaction by incubating compound **1**, PLP, and glutamate with EvaA and DnmJ in a buffered reaction. The reaction time was limited to 3.5 hours in order to minimize the loss of compound **3** due to back-conversion to compound **2**, which is unstable and can spontaneously decompose.

Similar to compound **1** purification, the synthesized compound **3** was purified by two chromatographic methods. First, anion exchange chromatography was performed using Sephadex Q resin with 20 mM NH_4HCO_2 (pH 3.5) – 1 M NH_4HCO_2 (pH 6.7) gradient elution profile (**Figure A10A**). Subsequently, desalting was performed using Sephadex G-10 column (**Figure A10B**). The

product was verified by MS at each purification step and following lyophilization. Same as compound **1**, compound **3** also exists in its non-hydrated and hydrated form at 4-keto position of the sugar moiety as detected by the *m/z* of 528.079 and 546.090, respectively (**Figure 25B**). Using our *in vitro* pathway, we were able to synthesize 1.5 mg of compound **3** with a yield of 10% (starting with 16.4 mg of compound **1**).

Aside from being precursors, purified sugar donors can be used in enzyme characterization. We used the purified compound **3** for kinetic characterization of the DnmJ aminotransferase (**Section 4.3.6**) and co-crystallization with DnmU as a collaboration with Dr. Rong Shi at University of Laval.

4.3.4.2 *In vitro* enzymatic synthesis and purification of TDP-L-daunosamine

We were able to synthesize TDP-L-daunosamine in a 40 mL reaction using our one-pot, two-step enzymatic pathway (**Figure 23:** Scheme **B**). Herein, the first step was performed to synthesize compound **3** and following the 3.5 h incubation, the second step was initiated by the addition of DnmU and DnmV along with NADPH co-substrate and was incubated for an additional 16 h at 24 ºC. Subsequently, the synthesized TDP-L-daunosamine was purified using the same anion exchange chromatographic technique that was used to purify compound **3** (**Figure A11**). Lyophilization of the purified TDP-L-daunosamine resulted in 2.2 mg with an overall yield of 14 % (starting with 16.4 mg of compound **1**) and analysis of TDP-L-daunosamine by MS resulted in a *m/z* of 530.094 in negative mode (**Figure 26**). Though we were able to synthesize and purify TDP-L-daunosamine, structural analysis through Nuclear magnetic resonance (NMR) studies is needed for TDP-L-daunosamine characterization.

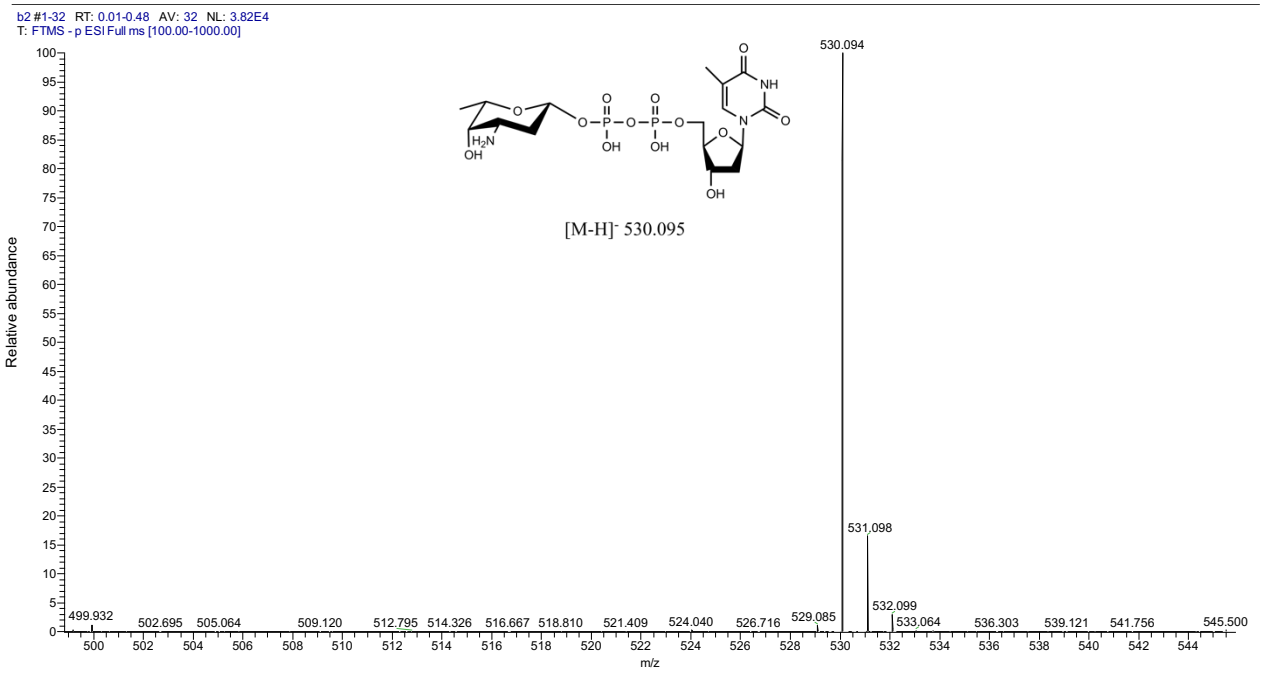


Figure 26: MS spectrum of TDP-L-daunosamine synthesized using the reconstituted *in vitro* pathway. The sample was analyzed in negative mode and the calculated *m/z* for TDP-L-daunosamine was 530.095.

4.3.5 Varying the downstream ketoreductase leads to the production of stereoisomers of TDP-L-daunosamine

The orientation of the hydroxyl group at C4 of TDP-L-acosamine is equatorial as opposed to the axial hydroxyl group of TDP-L-daunosamine. This is defined by the stereospecific nature of the ketoreductases used in our engineered pathway. Towards producing TDP-L-acosamine, three stereospecific ketoreductase candidates were selected: AveBIV (AvrE) from *S. avermitilis*, RmlD from *S. enterica*, and EvaE from *A. orientalis.* These three enzymes are involved in the biosynthetic pathways of TDP-L-oleandrose, TDP-L-rhamnose, and TDP-L-epivancosamine, respectively (Giraud *et al*., 1999; Ikeda *et al*., 1999; Williams *et al*., 1998). Previous studies reports the potential of AveBIV in the *in vivo* production of TDP-L-acosamine towards the synthesis of epirubicin (Han *et al*., 2011; Madduri *et al*., 1998). However, out of the three ketoreductases, soluble His_6-tagged AveBIV protein could not obtained by recombinant expression in *E. coli* for *in vitro* experiments. Even though soluble *N*-His_6-RmlD was obtained, the enzyme showed stringent specificity and did not accept compound **4** as a substrate. Therefore, AveBIV and RmlD were not used in further studies.

The potential of EvaE in producing TDP-L-acosamine in engineered *Streptomyces* strains was previously reported (Han *et al*., 2011; Wang *et al*., 2018a). Interestingly, we were able to reproduce this result *in vitro* by replacing DnmV with EvaE in our one-pot, two-step enzymatic pathway and MS spectrum of the product resulted in a *m/z* of 530.094 corresponding to TDP-L-acosamine (**Figure 27A**).

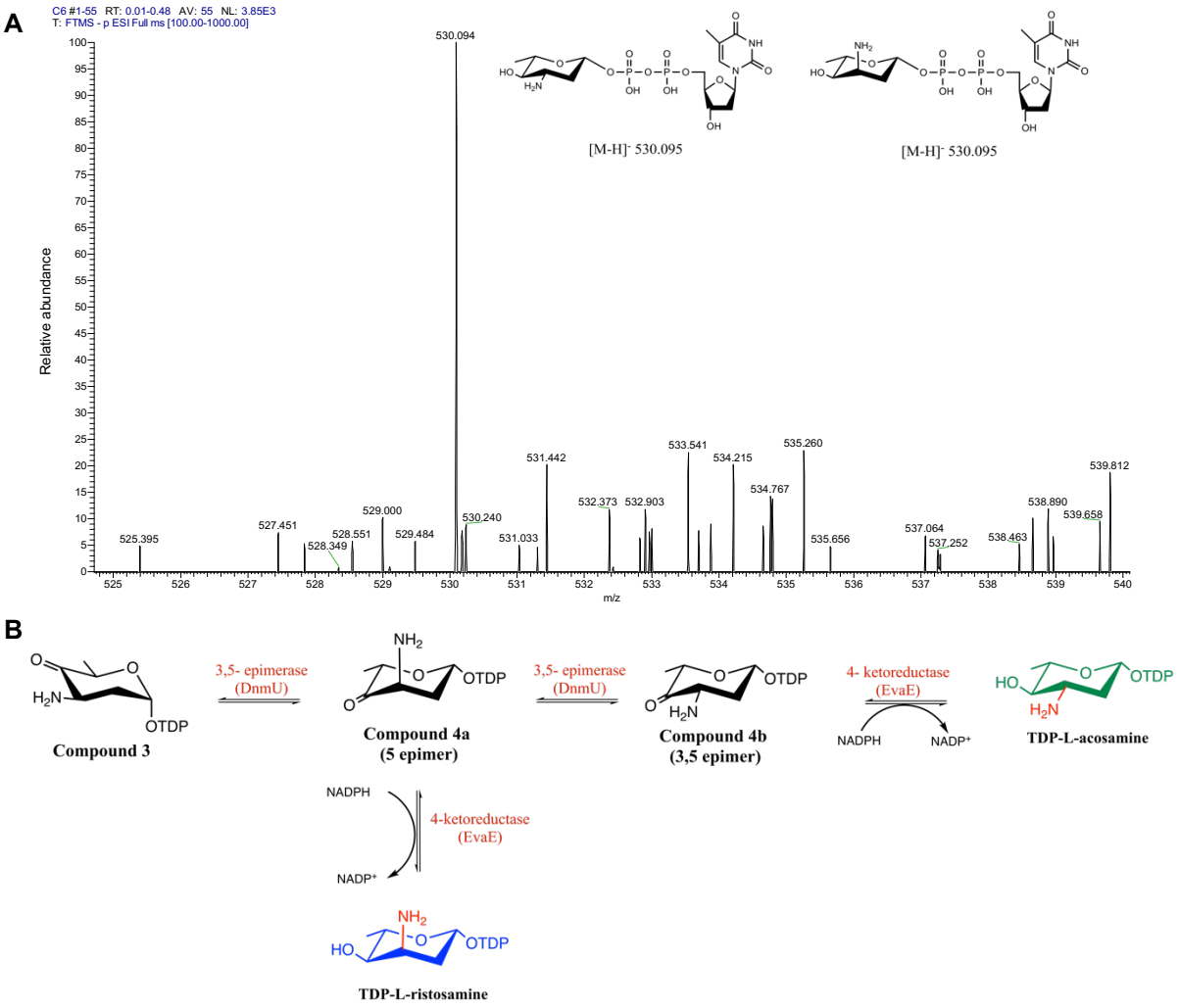


Figure 27: (A) MS spectrum of EvaE product(s) synthesized using the reconstituted *in vitro* pathway. The samples were analyzed in negative mode and the calculated *m/z* for TDP-L-acosamine and ristosamine was 530.095. **(B)** EvaE ketoreductase is capable of using compound **4a** epimer to produce TDP-L-ristosamine apart from converting compound **4b** to our desired product, TDP-L-acosamine.

However, as illustrated in **Figure 27B**, Yoon and coworkers reported the potential of EvaE to be substrate flexible with both 5- and 3,5-epimer of compound **3** (compound **4a** and **4b**, respectively) producing TDP-L-ristosamine and TDP-L-acosamine, respectively, as inferred by the incorporation of L-ristosamine and L-acosamine into a modified anthracycline produced by engineered streptomyces strains (Han *et al*., 2011). Therefore, in addition from producing the desired sugar donor, TDP-L-acosamine, we anticipate that our *in vitro* enzymatic reaction may have resulted in the production of TDP-L-ristosamine contaminant as well. Since these two stereoisomers cannot be differentiated by MS analysis, further analyses would be necessary to determine product stereochemistry (*e.g.* by NMR).

4.3.6 Determination of kinetic parameters of DnmJ aminotransferase

The aminotransferase reaction in our reconstituted pathway is a vital step as DnmJ seems to be a bottleneck towards the production of our desired donor sugars. This motivated us to carry out kinetic analysis of the DnmJ-catalyzed transamination. However, due to the low stability of compound **2**, it cannot be used as a substrate to assay the forward aminotransferase reaction. As reported by Krebs, the equilibrium constant (K_{eq}) for most aminotransferases is close to unity (Krebs, 1953). Assuming this to be true of the reaction catalyzed by DnmJ, the k_{cat}/K_M values for the forward and reverse reactions are likely to be very similar. Therefore, kinetic studies were performed for the reverse aminotransferase reaction using the purified compound **3**.

As shown in **Figure 28A**, kinetic parameter determination was done using a coupling reaction. The conversion of compound **3** to compound **2** transfers an amine group to α-ketoglutarate and produces glutamate. Since this reaction does not release a measurable signal, it was coupled to an NAD^+/ NADH redox reaction catalyzed by L-glutamate dehydrogenase (L-GDH). A subsequent diaphorase catalyzed coupling reaction with resazurin was performed to favour NAD^+ recycling and provide a strong fluorescence signal. Herein, the DnmJ enzyme activity was measured at varying concentrations of compound **3** by monitoring the fluorescence intensity of the producing resorufin (525 nm excitation and 598 nm emission) (Davis *et al*., 2016).

Briefly, 60 µL of buffered reactions (50 mM bis-tris propane pH 7.5) contained α-ketoglutarate (1 mM), PLP (0.1 mM), NAD^+ (0.5 mM), resazurin (12.5 µM), L-GDH (5 U/mL), diaphorase (0.02 mg/mL), with varying compound **3** concentrations (0.025 -0.50 mM). The

reactions were initiated upon addition of DnmJ (1 mg/mL) and were performed at 37 ºC. For each compound **3,** control reactions without DnmJ were performed to account for the background caused by traces of glutamate present in the purified compound **3**. Since the background was too high for compound **3** concentrations above 0.125 mM, these were eliminated from kinetic parameter determination.

Using the Michaelis-Menten curve, for DnmJ, we determine K_M value of (120 ± 72) µM with compound **3** and a k_{cat} of (2.6 ± 0.9) x 10^{-4} s^{-1} (**Figure 28B**). However, the level of confidence on the determined kinetic parameters was limited since the obtained Michaelis-Menten curve do not plateau and the highest compound **3** concentration was similar to that of the determined K_M value. Therefore, for DnmJ we determined the k_{cat}/ K_M of 1.64 x 10^{-6} s^{-1} μM^{-1} using a linear fit of the points, assuming the substrate concentration is negligible in comparison to the K_M value (**Figure A12**). In comparison with the catalytic efficiencies determined by the Michaelis-Menten curve (k_{cat}/ K_M of 2.17 x 10^{-6} s^{-1} μM^{-1}), the linear fit resulted in similar catalytic efficiency. Thus, this provided us with reasonable confidence about our kinetic measurements. Furthermore, we observed a similar K_M value to that reported for a homolog aminotransferase WecE from *E. coli* MG1655 (Cui *et al*., 2020).

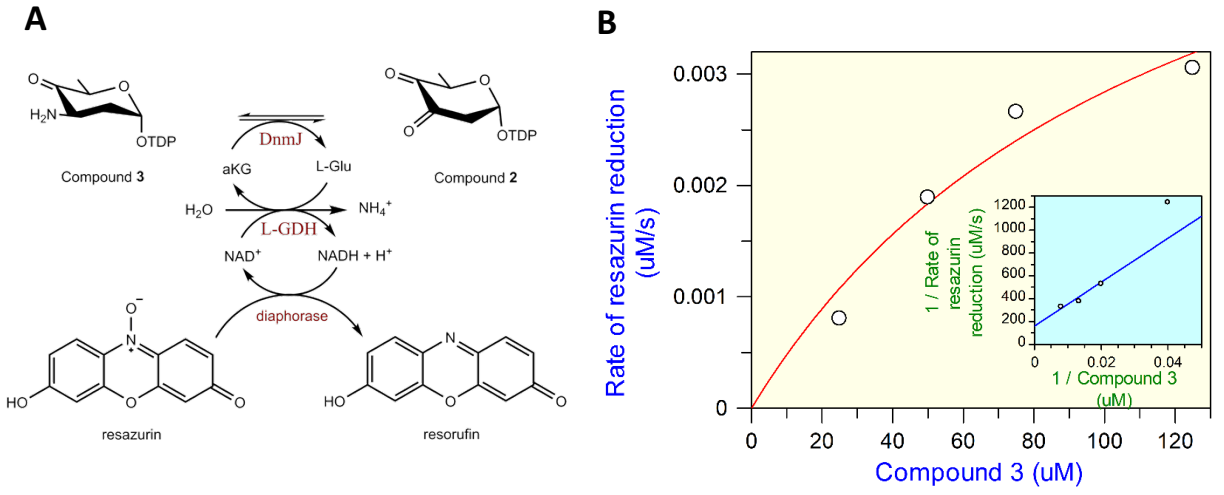


Figure 28: **(A)** The coupling reaction used to measure DnmJ kinetic parameters. aKG: α-ketoglutarate, L-Glu: L-glutamate, L-GDH: L-glutamate dehydrogenase. The reduction of resazurin to resorufin was observed at 525 nm excitation and 598 nm emission. **(B)** The Michaelis-Menten curve for DnmJ enzyme obtained from the resazurin coupled reaction.

4.4 Conclusions

The clinical relevance of L-daunosamine and its stereoisomers of anticancer anthracyclines inspired us in reconstituting an *in vitro* enzymatic pathway for the synthesis of the corresponding nucleotide activated sugar donors. An important characteristic of many sugar biosynthetic enzymes is their capacity to accept non-native substrates, where it assists the enzymatic production of novel donor molecules.

In this study using nine sugar biosynthetic enzymes overexpressed and purified, we have reconstituted an *in vitro* pathway for the production of TDP-L-daunosamine and its stereoisomer TDP-L-acosamine. Following the synthesis and purification of compound **1**, a one-pot two-step pathway was proposed towards the desired deoxyhexose donor production. Herein, we were able to synthesize and purify a few milligrams of TDP-L-daunosamine and the aminosugar intermediate, compound **3**. Furthermore, using our proposed pathway, we were able to synthesize TDP-L-acosamine that may have contaminated with TDP-L-ristosamine. Therefore, production of pure TDP-L-acosamine needs to be addressed in further studies.

The use of compound **3** towards determining the kinetic parameters of DnmJ highlights the importance of having purified sugar donors and intermediates in enzymatic characterization. Moreover, these could also be used in *in vitro* glycodiversification and as precursors for the *in vitro* enzymatic synthesis of some other valuable deoxysugars. The work presented here shed light towards the synthesis of many other unnatural deoxyhexoses that might have therapeutic importance.

CHAPTER 5

General conclusions and future directions

The biological significance of sugar moieties on small molecule natural products emphasizes the importance of glycodiversification or changing glycosylation towards enhancing or altering their function. The inherent or engineered substrate promiscuity of sugar biosynthetic enzymes and GTs are the driving tools of glycodiversification and this strategy can be performed *in vivo* or *in vitro*.

In vitro glycodiversification necessitates high-throughput screening platforms in order to screen a large number of GT-substrate libraries. As an alternative to the limitations of the existing high-throughput GT screening tools, we have developed a novel assay based on the polarity shift caused by glycosylation of small molecule aglycone acceptors that are inherently chromogenic or fluorescent. A key feature of this assay is its simplicity and ease of use as it doesn't require chemically altered substrates or sensors. Our plate-based high-throughput assay was statistically validated using a plant UGT and the assay showed compatibility with the purified enzyme and crude lysate. This enables it to be used in high-throughput GT characterization, *in vitro* glycodiversification, as well as in enzyme engineering approaches.

Subsequently, employing our assay, six substrate flexible plant UGTs were screened with a panel of flavonol and anthraquinone aglycones and three UDP-sugar donors. Through this screen, we identified previously unreported acceptor and donor substrate flexibilities of many of these UGTs. This demonstrates the potential of our screening tool in enzyme characterization. Furthermore, validation of these findings through LC-MS/MS analysis confirmed the reliability of our assay.

An interesting finding of our combinatorial screen was the identification of three substrate flexible plant UGTs (UGT71G1, UGT78K6, and UGT72B1) that synthesized novel ε-rhodomycinone-glycosides including glucosides and a xyloside. To the best of our knowledge, these UGTs were not reported to act on anthraquinone acceptors. While we have confirmed these glycosides through MS/MS fragmentation profiles, NMR studies would provide further details to characterize the structures of these novel glycosides. Towards NMR analysis, we may produce and purify (through reverse-phase prep-HPLC using a C-18 column) a few milligrams of these novel glycosides. The ε-rhodomycinone serves as the aglycone precursor towards the production of

anticancer drugs including daunorubicin, doxorubicin, and epirubicin. Considering the biological significance of ε-rhodomycinone on these anticancer anthracyclines and the presence of altered sugar moieties, we speculate that novel ε-rhodomycinone-glycosides may have therapeutic significance. Therefore, we may analyze these compounds for preliminary anticancer activity using an *in vitro* cytotoxic assay such as 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay.

So far, we performed the combinatorial screen using a limited number of acceptors and common sugar donors. Being inspired by our findings, it would be interesting to expand the combinatorial screen with a panel of other aglycone acceptors such as flavones, flavanones, anthocyanidins, and other anthraquinones using heavily modified deoxysugar donors knowing their biological impact on natural product glycosides. Towards this, the purified sugar donors and intermediates of Chapter **4** could be used.

TDP-L-daunosamine and TDP-L-acosamine are sugar precursors of the clinically important doxorubicin and its analog epirubicin, respectively. Considering the value of these sugar donors, we have reconstituted an *in vitro* enzymatic pathway towards the production of these precursors employing substrate flexible sugar biosynthetic enzymes. Using our *in vitro* pathway, we synthesized and purified TDP-L-daunosamine, and two other sugar intermediates, compound **1** and **3**. Despite our success in producing and purifying compound **3** and TDP-L-daunosamine using the reconstituted *in vitro* pathway, the obtained yield was between 10- 15 %. This can be addressed by coupling the proposed pathway to two co-substrate recycling cascades as illustrated in Figure **29**. The coupling of NADH/ L-glutamate dehydrogenase with formate/ formate dehydrogenase and NADH/SthA with formate/ formate dehydrogenase would regenerate L-glutamate and NADPH co-substrates, respectively. This would not only drive the reaction towards the desired sugar donor products but also allows the limited use of co-substrates that would otherwise interfere with the chromatographic purification of sugar donors and intermediates by co-eluting.

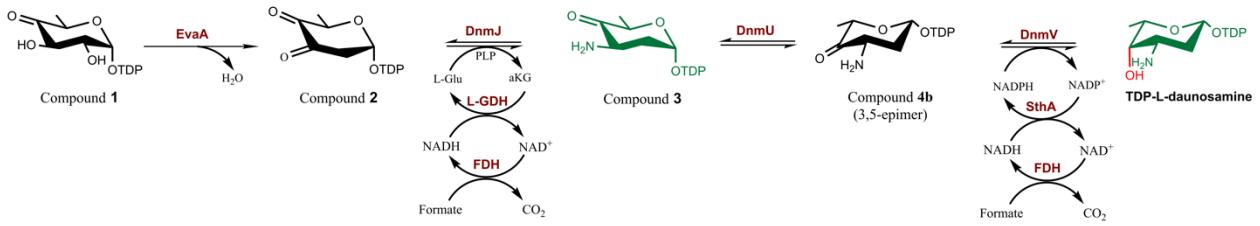


Figure 29: The proposed L-glutamate (L-Glu) and NADPH co-substrate recycling cascades of TDP-L-daunosamine pathway. aKG: α-ketoglutarate, PLP: pyridoxal 5'-phosphate

Using our proposed pathway, we were able to synthesize TDP-L-acosamine employing the substrate flexible EvaE 4-ketoreductase. However, considering its ability to act on compound **4a** and compound **4b** as reported by Yoon and colleagues (Han *et al*., 2011), we speculate that the synthesized product may be a mixture of TDP-L-acosamine and TDP-L-ristosamine (**Figure 27**). As shown in **Figure 30**, a possible alternative to have pure TDP-L-acosamine would be to perform the reverse reaction of DnmV enzyme starting from purified TDP-L-daunosamine to produce compound **4b**. Following isolation, compound **4b** could be converted to TDP-L-acosamine. Another approach would be to use a different 4-ketoreductase that shows stringent substrate specificity to compound **4b** but not compound **4a**. A potential candidate towards this purpose would be EryBIV from *S. erythraea* that was previously used in *in vivo* production of TDP-L-acosamine towards epirubicin synthesis (Madduri *et al*., 1998). The native substrate of EvaE is a TDP-deoxysugar intermediate that has two substituents at the C-3 position, and this may afford it some flexibility when challenged with compound **4a** or **4b**, which are epimers each with opposite configuration at the C-3 position bearing only a single amino group as a substituent. Similarly, the native substrate of EryBIV has two substituents at the C-3 position. Therefore, following that rational EryBIV may also have the same lack of specificity.

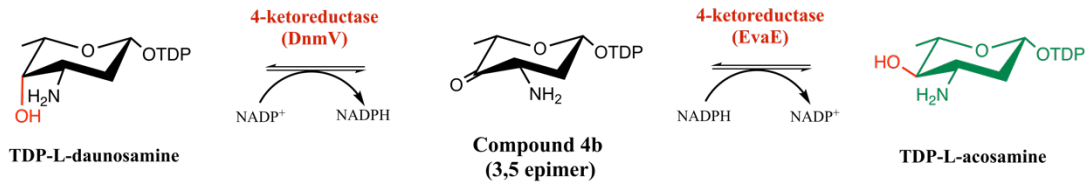


Figure 30: The proposed pathway for the production of pure TDP-L-acosamine sugar donor starting from purified TDP-L-daunosamine.

The production and purification of these heavily modified sugar donors and its precursors can be used in enzyme characterization. Compound **3** was used to determine the kinetic parameters of DnmJ enzyme that catalyze the aminotransferase reaction of the re-constituted pathway. Furthermore, compound **3** serves as a precursor for the DnmU 3,5-epimerase. Therefore, a few milligrams of purified compound **3** was sent to Dr. Rong Shih at University of Laval towards co-crystalizing with DnmU enzyme. Having purified TDP-L-daunosamine would enable us to obtain insights about the kinetic parameters of DnmV 4-ketoreductase. This may be achieved by performing the reverse reaction of the 4-ketoreductase where the reaction can be monitored by measuring the production of NADPH at 340 nm.

Apart from their use in enzyme characterization, these sugar donors and their intermediates could serve as precursors towards the *in vitro* enzymatic production of a number of diverse sugar donors of therapeutic value. For example, compound **3** acts as a precursor for the production of TDP-4-*epi*-L-vancosamine: the sugar precursor of antibiotic chloroeremomycin (Ikeda *et al*., 1999), whereas TDP-L-daunosamine can be converted to TDP-L-rhodosamine: one of the sugar precursors of anticancer aclarubicin (Räty *et al*., 2000).

Despite the clinical importance of epirubicin, it is conventionally produced semisynthetically through several organic synthetic steps. To circumvent the laborious chemical synthesis, efforts have been made for the heterologous production of epirubicin. Nevertheless, known GT enzymes cannot catalyze this enzymatic reaction efficiently and only poor glycosylation of ε-rhodomycinone acceptor with a TDP-L-acosamine donor has been observed in engineered biosynthetic pathways (Han *et al*., 2011; Madduri *et al*., 1998). Wang and colleagues reported epirubicin production in a *S. peucetius* mutant that was constructed by UV irradiation and metabolic pathway engineering. Although the obtained epirubicin yield through this study was significant, details on mutations caused by UV irradiation were not reported. Furthermore, the reported epirubicin yield could be a mixture of epirubicin and its epimer with ristosamine sugar moiety as they used EvaE towards TDP-L-acosamine production (Wang *et al*., 2018a).

An attractive solution to increase epirubicin production would be to engineer a GT enzyme through directed evolution that could efficiently act on the precursors of epirubicin. Two potential candidates to be used as the target GT genes would be *dnrS* from *S. peucetius* and *aknS* from *S. galilaeus* that were reported to be substrate flexible towards producing epirubicin in engineered

Streptomyces mutants (Han *et al*., 2011; Madduri *et al*., 1998; Wang *et al*., 2018a). In its native organisms, DnrS is involved in glycosylating ε-rhodomycinone using TDP-L-daunosamine towards daunorubicin/ doxorubicin production (Otten *et al*., 1995) and AknS glycosylates aklavinone using TDP-L-rhodosamine towards the production of the anticancer anthracycline aclarubicin (Räty *et al*., 2000). Aklavinone is the immediate intermediate of ε-rhodomycinone and it lacks a hydroxyl group at C-11 position (Filippini *et al*., 1995) and TDP-L-rhodosamine is the 3-*N*,*N*-methyl form of TDP-L-daunosamine (Räty *et al*., 2000).

Since crystal structures of these GTs are not available, mutant library construction could be performed by applying a random mutagenesis strategy such as error-prone PCR. However, this would result in a larger library size which would be tedious in subsequent screening. Therefore, it would be of interest in future work to build a homology model for the target GT that would give an understanding about the specific sites involved in substrate binding and catalytic activity. Once the target residues for mutation are determined, iterative saturation mutagenesis (ISM) may be performed until a desired phenotype is obtained (Reetz and Carballeira, 2007). Towards screening the mutant library, our novel high-throughput assay could serve as the screening tool using ε-rhodomycinone aglycone and TDP-L-acosamine that we would synthesize and purify. The best mutants will be subjected to further rounds of directed evolution until an engineered GT that could efficiently glycosylate epirubicin precursors is attained.

References

Achnine L, Huhman D V., Farag MA, Sumner LW, Blount JW, Dixon RA. 2005. Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume Medicago truncatula. *Plant J.* **41**:875–887.

Aharoni A, Thieme K, Chiu CPC, Buchini S, Lairson LL, Chen H, Strynadka NCJ, Wakarchuk WW, Withers SG. 2006. High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* **3**:609–614.

Albesa-Jové D, Mendoza F, Rodrigo-Unzueta A, Gomollõn-Bel F, Cifuente JO, Urresti S, Comino N, Gõmez H, Romero-García J, Lluch JM, Sancho-Vaello E, Biarnés X, Planas A, Merino P, Masgrau L, Guerin ME. 2015. A Native Ternary Complex Trapped in a Crystal Reveals the Catalytic Mechanism of a Retaining Glycosyltransferase. *Angew. Chemie - Int. Ed.* **54**:9898–9902.

Albuquerque-Wendt A, Hütte HJ, Buettner FFR, Routier FH, Bakker H. 2019. Membrane topological model of glycosyltransferases of the GT-C superfamily. *Int. J. Mol. Sci.* **20**:4842.

Arcamone F, Penco S, Vigevani A, Redaelli S, Franchi G, Di Marco A, Casazza AM, Dasdia T, Formelli F, Neceo A, Soranzo C. 1975. Synthesis and Antitumor Properties of New Glycosides of Daunomycinone and Adriamycinone. *J. Med. Chem.* **18**:703–707.

Ardèvol A, Rovira C. 2015. Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J. Am. Chem. Soc.* **137**:7528–7547.

Asada Y, Hirayama Y, Furuya T. 1988. Acylated flavonols from Crocosmia crocosmiiflora. *Phytochemistry* **27**:1497–1501.

Bao W, Sheldon PJ, Wendt-Pienkowski E, Richard Hutchinson C. 1999. The Streptomyces peucetius dpsC gene determines the choice of starter unit in biosynthesis of the daunorubicin polyketide. *J. Bacteriol.* **181**:4690–4695.

Barton WA, Lesniak J, Biggins JB, Jeffrey PD, Jiang J, Rajashankar KR, Thorson JS, Nikolov DB. 2001. Structure, mechanism and engineering of a nucleotidylyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* **8**:545–551.

Berg S, Kaur D, Jackson M, Brennan PJ. 2007. The glycosyltransferases of Mycobacterium tuberculosis - Roles in the synthesis of arabinogalactan, lipoarabinomannan, and other

glycoconjugates. *Glycobiology* **17**:35R–56R.

Bonadonna G, Gianni L, Santoro A, Bonfante V, Bidoli P, Casali P, Demicheli R, Valagussa P. 1993. Drugs ten years later: Epirubicin. *Ann. Oncol.* **4**:359–369. https://doi.org/10.1093/oxfordjournals.annonc.a058514.

Borisova SA, Zhang C, Takahashi H, Zhang H, Wong AW, Thorson JS, Liu HW. 2006. Substrate specificity of the macrolide-glycosylating enzyme pair DesVII/DesVIII: Opportunities, limitations, and mechanistic hypotheses. *Angew. Chemie - Int. Ed.* **45**:2748–2753.

Brazier-Hicks M, Edwards R. 2005. Functional importance of the family 1 glucosyltransferase UGT72B1 in the metabolism of xenobiotics in Arabidopsis thaliana. *Plant J.* **42**:556–566.

Brazier-Hicks M, Offen WA, Gershater MC, Revett TJ, Lim EK, Bowles DJ, Davies GJ, Edwards R. 2007a. Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc. Natl. Acad. Sci. U. S. A.* **104**:20238–20243.

Brazier-Hicks M, Offen WA, Gershater MC, Revett TJ, Lim EK, Bowles DJ, Davies GJ, Edwards R. 2007b. Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. *PNAS* **104**:20238–20243.

Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A. 2006. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**:29R-37R.

Campbell JA, Davies GJ, Bulone V, Henrissat B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem. J.* **326**:929–939.

Capranico G, Binaschi M, Borgnetto ME, Zunino F, Palumbo M. 1997. A protein-mediated mechanism for the DNA sequence-specific action of topoisomerase II poisons. *Trends Pharmacol. Sci.* **18**:323–329.

Caputi L, Lim EK, Bowles DJ. 2008. Discovery of new biocatalysts for the glycosylation of terpenoid scaffolds. *Chem. - A Eur. J.* **14**:6656–6662.

Caputi L, Malnoy M, Goremykin V, Nikiforova S, Martens S. 2012. A genome-wide phylogenetic reconstruction of family 1 UDP- glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. *Plant J.* **69**:1030–1042.

Cardicini CE, Paladini AC, Caputto R, Leloir LF. 1950. Uridine Diphosphate Glucose: the

Coenzyme of the Galactose-Glucose Phosphate lsomerization. *Nature* **165**:191–192.

Chen H, Agnihotri G, Guo Z, Que NLS, Chen XH, Liu HW. 1999. Biosynthesis of Mycarose: Isolation and Characterization of Enzymes Involved in the C-2 Deoxygenation. *J. Am. Chem. Soc.* **121**:8124–8125.

Chen H, Thomas MG, Hubbard BK, Losey HC, Walsh CT, Burkart MD. 2000. Deoxysugars in glycopeptide antibiotics: Enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **97**:11942–11947.

Chen SH, Chan N-L, Hsieh T. 2013. New Mechanistic and Functional Insights into DNA Topoisomerases. *Annu. Rev. Biochem.* **82**:139–170.

Coleman JOD, Blake-Kalff MMA, Davies TGE. 1997. Detoxification of xenobiotics by plants: Chemical modification and vacuolar compartmentation. *Trends Plant Sci.* **2**:144–151.

Cortés-Funes H, Coronado C. 2007. Role of anthracyclines in the era of targeted therapy. *Cardiovasc. Toxicol.* **7**:56–60.

Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328**:307–317.

Cui L, Wei X, Wang X, Bai L, Lin S, Feng Y. 2020. A validamycin shunt pathway for valienamine synthesis in engineered streptomyces hygroscopicus 5008. *ACS Synth. Biol.* **9**:294–303.

Davies J, Ryan KS. 2012. Introducing the parvome: Bioactive compounds in the microbial world. *ACS Chem. Biol.* **7**:252–259.

Davis MI, Shen M, Simeonov A, Hall MD. 2016. Diaphorase Coupling Protocols for Red-Shifting Dehydrogenase Assays. *Assay Drug Dev. Technol.* **14**:207–212.

Dickens ML, Priestley ND, Strohl WR. 1997. In vivo and in vitro bioconversion of ε-rhodomycinone glycoside to doxorubicin: Functions of DauP, DauK, and DoxA. *J. Bacteriol.* **179**:2641–2650.

Eixelsberger T, Nidetzky B. 2014. Enzymatic redox cascade for one-pot synthesis of Uridine 5'-Diphosphate xylose from Uridine 5'-Diphosphate glucose. *Adv. Synth. Catal.* **356**:3575–3584.

Ferreyra MLF, Rodriguez E, Casas MI, Labadie G, Grotewold E, Casati P. 2013. Identification of a bifunctional Maize C- and O-glucosyltransferase. *J. Biol. Chem.* **288**:31678–31688.

Filippini S, Solinas MM, Breme U, Schluter MB, Gabellini D, Biamonti G, Colombo AL,

Garofano L. 1995. Streptomyces peucetius daunorubicin biosynthesis gene, dnrF: Sequence and heterologous expression. *Microbiology* **141**:1007–1016.

Ford CM, Boss PK, Bordier Høj P. 1998. Cloning and Characterization of Vitis vinifera UDP-Glucose:Flavonoid 3-O-Glucosyltransferase, a Homologue of the Enzyme Encoded by the Maize Bronze-1 Locus That May Primarily Serve to Glucosylate Anthocyanidins in Vivo*. *J. Biol. Chem.* **273**:9224–9233. http://www.jbc.org/.

Gallo MA, Wardlt J, Hutchinson CR. 1996. The dnrM gene in Streptomyces peucetius contains a naturally occurring frameshift mutation that is suppressed by another locus outside of the daunorubicin-production gene cluster. *Microbiology* **142**:269–275.

Giraud MF, McMiken HJ, Leonard GA, Messner P, Whitfield C, Naismith JH. 1999. Overexpression, purification, crystallization and preliminary structural study of dTDP-6-deoxy-L-lyxo-4-hexulose reductase (RmlD), the fourth enzyme of the dTDP-L-rhamnose synthesis pathway, from Salmonella enterica serovar typhimurium. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **D55**:2043–2046.

Grein A. 1987. Antitumor Anthracyclines Produced by Streptomyces peucetius. *Adv. Appl. Microbiol.* **32**:203–214.

Grimm A, Madduri K, Ali A, Hutchinson CR. 1994. Characterization of the Streptomyces peucetius ATCC 29050 genes encoding doxorubicin polyketide synthase. *Gene* **151**:1–10.

Han AR, Park JW, Lee MK, Ban YH, Yoo YJ, Kim EJ, Kim E, Kim BG, Sohng JK, Yoon YJ. 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.

Hanahan D. 1983. Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* **166**:557–580.

He XZ, Wang X, Dixon RA. 2006. Mutational analysis of the Medicago glycosyltransferase UGT71G1 reveals residues that control regioselectivity for (Iso)flavonoid glycosylation. *J. Biol. Chem.* **281**:34441–34447.

Hiromoto T, Honjo E, Noda N, Tamada T, Kazuma K, Suzuki M, Blaber M, Kuroki R. 2015. Structural basis for acceptor-substrate recognition of UDP-glucose: anthocyanidin 3-O-glucosyltransferase from Clitoria ternatea. *Protein Sci.* **24**:395–407.

Hiromoto T, Honjo E, Tamada T, Noda N, Kazuma K, Suzuki M, Kuroki R. 2013. Crystal

structure of UDP-glucose:anthocyanidin 3-O-glucosyltransferase from Clitoria ternatea. *J. Synchrotron Radiat.* **20**:894–898.

Hollman PCH, Bijsman MNCP, Van Gameren Y, Cnossen EPJ, De Vries JHM, Katan MB. 1999. The sugar moiety is a major determinant of the absorption of dietary flavonoid glycosides in man. *Free Radic. Res.* **31**:569–573.

Hughes J, Hughes MA. 1994. Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (Manihot esculenta Crantz) cotyledons. *DNA Seq.* **5**:41–49.

Hutchinson CR, Colombo AL. 1999. Genetic engineering of doxorubicin production in Streptomyces peucetius: A review. *J. Ind. Microbiol. Biotechnol.* **23**:647–652.

Ikeda H, Nonomiya T, Usami M, Ohta T, Omura S. 1999. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in Streptomyces avermitilis. *Proc. Natl. Acad. Sci. U. S. A.* **96**:9509–9514.

Jackson RG, Kowalczyk M, Li Y, Higgins G, Ross J, Sandberg G, Bowles DJ. 2002. Over-expression of an Arabidopsis gene encoding a glucosyltransferase of indole-3-acetic acid: Phenotypic characterisation of transgenic lines. *Plant J.* **32**:573–583.

Jungsuwadee P. 2016. Doxorubicin-induced cardiomyopathy: An update beyond oxidative stress and myocardial cell death. *Cardiovasc. Regen. Med.* **3**:1–8.

Kazuma K, Kogawa K, Noda N, Kato N, Suzuki M. 2004. Identification of delphinidin 3-O-(6′-O-Malonyl)-β-glucoside- 3′-O-β-glucoside, a postulated intermediate in the biosynthesis of ternatin C5in the blue petals of clitoria ternatea (butterfly pea). *Chem. Biodivers.* **1**:1762–1770.

Kornberg A, Lieberman I, Simms ES. 1955. Enzymatic synthesis of purine nucleotides. *J. Biol. Chem.* **215**:417–427.

Krebs HA. 1953. Equilibria in transamination systems. *Biochem. J.* **54**:82–86.

Kwan DH, Withers SG. 2011. Toward efficient enzymatic glycan synthesis: Directed evolution and enzyme engineering. *J. Carbohydr. Chem.* **30**:181–205.

Lahtinen R, Kuikka J, Nousiainen T, Uusitupa M, Lansimies E. 1991. Cardiotoxicity of epirubicin and doxorubicin: A double-blind randomized study. *Eur. J. Haematol.* **46**:301–305.

Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyltransferases: Structures,

Functions, and Mechanisms. *Annu. Rev. Biochem.* **77**:521–555.

Lairson LL, Chiu CPC, Ly HD, He S, Wakarchuk WW, Strynadka NCJ, Withers SG. 2004. Intermediate trapping on a mutant retaining α-galactosyltransferase identifies an unexpected aspartate residue. *J. Biol. Chem.* **279**:28339–28344.

Lee HS, Thorson JS. 2011. Development of a universal glycosyltransferase assay amenable to high-throughput formats. *Anal. Biochem.* **418**:85–88. http://dx.doi.org/10.1016/j.ab.2011.06.016.

Lee SS, Hong SY, Errey JC, Izumi A, Davies GJ, Davis BG. 2011. Mechanistic evidence for a front-side, SNi-type reaction in a retaining glycosyltransferase. *Nat. Chem. Biol.* **7**:631–638. http://dx.doi.org/10.1038/nchembio.628.

Li T-L, Huang F, Haydock SF, Mironenko T, Leadlay PF, Spencer JB. 2004. Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase. *Chem. Biol.* **11**:107–119.

Li Z, Han K, Pak JE, Satkunarajah M, Zhou D, Rini JM. 2017. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat. Chem. Biol.* **13**:757–763. http://dx.doi.org/10.1038/nchembio.2381.

Lieberman I, Kornberg A, Simms ES. 1955. Enzymatic synthesis of pyrimidine nucleotides. Ortidine-5'-phosphate and uridine-5'-phosphate. *J. Biol. Chem.* **215**:403–415.

Lira-Navarrete E, Iglesias-Fernández J, Zandberg WF, Compañón I, Kong Y, Corzana F, Pinto BM, Clausen H, Peregrina JM, Vocadlo DJ, Rovira C, Hurtado-Guerrero R. 2014. Substrate-guided front-face reaction revealed by combined structural snapshots and metadynamics for the polypeptide N-acetylgalactosaminyltransferase 2. *Angew. Chemie - Int. Ed.* **53**:8206–8210.

Lira-Navarrete E, Valero-González J, Villanueva R, Martínez-Júlvez M, Tejero T, Merino P, Panjikar S, Hurtado-Guerrero R. 2011. Structural insights into the mechanism of protein O-fucosylation. *PLoS One* **6**:e25365.

Liu J, Mushegian A. 2003. Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.* **12**:1418–1431.

Lomovskaya N, Doi-Katayama Y, Filippini S, Nastro C, Fonstein L, Gallo M, Colombo AL, Hutchinson CR. 1998. The Streptomyces peucetius dpsY and dnrX genes govern early and late steps of daunorubicin and doxorubicin biosynthesis. *J. Bacteriol.* **180**:2379–2386.

Lomovskaya N, Otten SL, Doi-Katayama Y, Fonstein L, Liu XC, Takatsu T, Inventi-Solari A, Filippini S, Torti F, Colombo AL, Hutchinson CR. 1999. Doxorubicin overproduction in Streptomyces peucetius: Cloning and characterization of the dnrU ketoreductase and dnrV genes and the doxA cytochrome P-450 hydroxylase gene. *J. Bacteriol.* **181**:305–318.

Loutre C, Dixon DP, Brazier M, Slater M, Cole DJ, Edwards R. 2003. Isolation of a glucosyltransferase from Arabidopsis thaliana active in the metabolism of the persistent pollutant 3,4-dichloroaniline. *Plant J.* **34**:485–493.

Lovering AL, De Castro LH, Lim D, Strynadka NCJ. 2007. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science (80-.).* **315**:1402–1405.

Lu W, Oberthür M, Leimkuhler C, Tao J, Kahne D, Walsh CT. 2004. Characterization of a regiospecific epivancosaminyl transferase GtfA and enzymatic reconstitution of the antibiotic chloroeremomycin. *Proc. Natl. Acad. Sci. U. S. A.* **101**:4390–4395.

Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF, Nebert DW. 1997. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* **7**:255–269.

Madduri K, Hutchinson CR. 1995. Functional characterization and transcriptional analysis of a gene cluster governing early and late steps in daunorubicin biosynthesis in Streptomyces peucetius. *J. Bacteriol.* **177**:3879–3884.

Madduri K, Kennedy J, Rivola G, Inventi-Solari A, Filippini S, Zanuso G, Colombo AL, Gewain KM, Occi JL, MacNeil DJ, Hutchinson CR. 1998. Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of Streptomyces peucetius. *Nat. Biotechnol.* **16**:69–74.

Mahdi JG, Mahdi AJ, Mahdi AJ, Bowen ID. 2006. The historical analysis of aspirin discovery, its relation to the willow tree and antiproliferative and anticancer potential. *Cell Prolif.* **39**:147–155.

Malla S, Prasad Niraula N, Singh B, Liou K, Kyung Sohng J. 2010. Limitations in doxorubicin production from Streptomyces peucetius. *Microbiol. Res.* **165**:427–435. http://dx.doi.org/10.1016/j.micres.2009.11.006.

Di Marco A, Arcamone F, Zunino F. 1975. Daunomycin (Daunorubicin) and Adriamycin and

Structural Analogues: Biological Activity and Mechanism of Action. In: Corcoran, JW, Hahn, FE, Snell, JF, Arora, KL, editors. *Mech. Action Antimicrob. Antitumor Agents*. Springer, Berlin, Heidelberg, pp. 101–128.

Minotti G, Menna P, Salvatorelli E, Cairo G, Gianni L. 2004. Anthracyclines: Molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacol. Rev.* **56**:185–229.

Modolo L V., Blount JW, Achnine L, Naoumkina MA, Wang X, Dixon RA. 2007. A functional genomics approach to (iso)flavonoid glycosylation in the model legume Medicago truncatula. *Plant Mol. Biol.* **64**:499–518.

Modolo L V., Li L, Pan H, Blount JW, Dixon RA, Wang X. 2009. Crystal Structures of Glycosyltransferase UGT78G1 Reveal the Molecular Basis for Glycosylation and Deglycosylation of (Iso)flavonoids. *J. Mol. Biol.* **392**:1292–1302. http://dx.doi.org/10.1016/j.jmb.2009.08.017.

Mohideen FI, Ghosh T, Nguyen LH, Kwan DH. 2021. Controlling Glycosyltransferase Activity: Inhibition and Enzyme Engineering. In: Barchi, J, editor. *Compr. Glycosci. (Second Ed.*, pp. 204–232.

Moncrieffe MC, Fernandez MJ, Spiteller D, Matsumura H, Gay NJ, Luisi BF, Leadlay PF. 2012. Structure of the glycosyltransferase EryCIII in complex with its activating P450 homologue EryCII. *J. Mol. Biol.* **415**:92–101. http://dx.doi.org/10.1016/j.jmb.2011.10.036.

Moremen KW, Haltiwanger RS. 2019. Emerging structural insights into glycosyltransferase-mediated synthesis of glycans. *Nat. Chem. Biol.* **15**:853–864. http://dx.doi.org/10.1038/s41589-019-0350-2.

Moro S, Beretta GL, Dal Ben D, Nitiss J, Palumbo M, Capranico G. 2004. Interaction model for anthracycline activity against DNA topoisomerase II. *Biochemistry* **43**:7503–7513.

Mukhopadhyay P, Rajesh M, Bátkai S, Kashiwaya Y, Haskó G, Liaudet L, Szabó C, Pacher P. 2009. Role of superoxide, nitric oxide, and peroxynitrite in doxorubicin-induced cell death in vivo and in vitro. *Am. J. Physiol. - Hear. Circ. Physiol.* **296**:1466–1483.

Niraula NP, Kim SH, Sohng JK, Kim ES. 2010. Biotechnological doxorubicin production: Pathway and regulation engineering of strains for enhanced production. *Appl. Microbiol. Biotechnol.* **87**:1187–1194.

Oberthür M, Leimkuhler C, Kahne D. 2004. A practical method for the stereoselective

generation of β-2-deoxy glycosyl phosphates. *Org. Lett.* **6**:2873–2876.

Offen W, Martinez-Fleites C, Yang M, Kiat-Lim E, Davis BG, Tarling CA, Ford CM, Bowles DJ, Davies GJ. 2006. Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J.* **25**:1396–1405.

Oh J, Lee SG, Kim BG, Sohng JK, Liou K, Lee HC. 2003. One-pot Enzymatic Production of dTDP-4-keto-6-deoxy-D-glucose from dTMP and Glucose-1-phosphate. *Biotechnol. Bioeng.* **84**:452–458.

Otten SL, Liu X, Ferguson J, Hutchinson CR. 1995. Cloning and characterization of the Streptomyces peucetius dnrQS genes encoding a daunosamine biosynthesis enzyme and a glycosyl transferase involved in daunorubicin biosynthesis. *J. Bacteriol.* **177**:6688–6692.

Otten SL, Gallo MA, Madduri K, Liu X, Hutchinson CR. 1997. Cloning and characterization of the Streptomyces peucetius dnmZUV genes encoding three enzymes required for biosynthesis of the daunorubicin precursor thymidine diphospho-L-daunosamine. *J. Bacteriol.* **179**:4446–4450.

Paquette S, Møller BL, Bak S. 2003. On the origin of family 1 plant glycosyltransferases. *Phytochemistry* **62**:399–413.

Pedersen LC, Dong J, Taniguchi F, Kitagawa H, Krahn JM, Pedersen LG, Sugahara K, Negishi M. 2003. Crystal structure of an α1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. *J. Biol. Chem.* **278**:14420–14428.

Peel GJ, Pang Y, Modolo L V., Dixon RA. 2009. The LAP1 MYB transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in Medicago. *Plant J.* **59**:136–149.

Persson K, Ly HD, Dieckelmann M, Wakarchuk WW, Withers SG, Strynadka NCJ. 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.

Pival SL, Klimacek M, Nidetzky B. 2008. Novel chemo-enzymatic mimic of hydrogen peroxide-forming NAD(P)H oxidase for efficient regeneration of NAD+ and NADP+. *Adv. Synth. Catal.* **350**:2305–2312.

Räty K, Kunnari T, Hakala J, Mäntsälä P, Ylihonko K. 2000. A gene cluster from Streptomyces galilaeus involved in glycosylation of aclarubicin. *Mol. Gen. Genet.* **264**:164–172.

Reetz MT, Carballeira JD. 2007. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2**:891–903.

van der Rijst M, Scheeren JW, de Vos D. 1999. Process for preparing epirubicin or acid addition salts thereof from daunorubicin.

Rini JM, Esko JD. 2017. Glycosyltransferases and Glycan--Processing Enzymes. In: Varki, A, Cummings, RD, Esko, JD, Stanley, P, Hart, GW, Aebi, M, Darvill, AG, Kinoshita, T, Packer, NH, Prestegard, JH, Schnaar, RL, Seeberger, PH, editors. *Essentials Glycobiol.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), pp. 65–75.

Ross J, Li Y, Lim E, Bowles DJ. 2001. Higher plant glycosyltransferases. *genomebiology* **2**:3004.1–3004.6.

Ryu J, Eom MS, Ko W, Han MS, Lee HS. 2014. A fluorescence-based glycosyltransferase assay for high-throughput screening. *Bioorganic Med. Chem.* **22**:2571–2575. http://dx.doi.org/10.1016/j.bmc.2014.02.027.

Salas JA, Méndez C. 2007. Engineering the glycosylation of natural products in actinomycetes. *Trends Microbiol.* **15**:219–232.

Sarvazyan N. 1996. Visualization of doxorubicin-induced oxidative stress in isolated cardiac myocytes. *Am. J. Physiol.* **271**:H2079-2085.

Saslowsky DE, Warek U, Winkel BSJ. 2005. Nuclear localization of flavonoid enzymes in Arabidopsis. *J. Biol. Chem.* **280**:23735–23740. http://dx.doi.org/10.1074/jbc.M413506200.

Sernee MF, Ralton JE, Nero TL, Sobala LF, Kloehn J, Vieira-Lara MA, Cobbold SA, Stanton L, Pires DEV, Hanssen E, Males A, Ward T, Bastidas LM, van der Peet PL, Parker MW, Ascher DB, Williams SJ, Davies GJ, McConville MJ. 2019. A Family of Dual-Activity Glycosyltransferase-Phosphorylases Mediates Mannogen Turnover and Virulence in Leishmania Parasites. *Cell Host Microbe* **26**:385–399. https://doi.org/10.1016/j.chom.2019.08.009.

Shao H, He X, Achnine L, Blount JW, Dixon RA, Wang X. 2005. Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from Medicago truncatula. *Plant Cell* **17**:3141–3154.

Shibuya M, Nishimura K, Yasuyama N, Ebizuka Y. 2010. Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in Glycine max. *FEBS Lett.* **584**:2258–2264. http://dx.doi.org/10.1016/j.febslet.2010.03.037.

Sinnott ML, Jencks WP. 1980. Solvolysis of D-Glucopyranosyl Derivatives in Mixtures of Ethanol and 2,2,2-Trifluoroethanol. *J. Am. Chem. Soc.* **102**:2026–2032.

Smith GJ, Thomsen SJ, Markham KR, Andary C, Cardon D. 2000. The photostabilities of naturally occurring 5-hydroxyflavones, flavonols, their glycosides and their aluminium complexes. *J. Photochem. Photobiol. A Chem.* **136**:87–91.

Staunton J, Wilkinson B. 1997. Biosynthesis of erythromycin and rapamycin. *Chem. Rev.* **97**:2611–2629.

Summers RG, Donadio S, Staver MJ, Wendt-Pienkowski E, Hutchinson CR, Katz L. 1997. Sequencing and mutagenesis of genes from the erythromycin biosynthetic gene cluster of Saccharapolyspora erythraea that are involved in L-mycarose and D-desosamine production. *Microbiology* **143**:3251–3262.

Thibodeaux CJ, Melançon CE, Liu H. 2008. Natural Product Sugar Biosynthesis and Enzymatic Glycodiversification. *Angew. Chemie Int. Ed.* **47**:9814–9859.

Thibodeaux CJ, Melançon CE, Liu HW. 2007. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**:1008–1016.

Truman AW, Dias MVB, Wu S, Blundell TL, Huang F, Spencer JB. 2009. Chimeric Glycosyltransferases for the Generation of Hybrid Glycopeptides. *Chem. Biol.* **16**:676–685. http://dx.doi.org/10.1016/j.chembiol.2009.04.013.

Ünligil UM, Rini JM. 2000. Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* **10**:510–517.

Urbanowicz BR, Bharadwaj VS, Alahuhta M, Peña MJ, Lunin V V., Bomble YJ, Wang S, Yang J-Y, Tuomivaara ST, Himmel ME, Moremen KW, York WS, Crowley MF. 2017. Structural, Mutagenic and In Silico studies of Xyloglucan Fucosylation in Arabidopsis thaliana Suggest a Water-Mediated Mechanism. *Plant J.* **91**:931–949.

Walczak RJ, Dickens ML, Priestley ND, Strohl WR. 1999. Purification, properties, and characterization of recombinant Streptomyces sp. strain C5 DoxA, a cytochrome P-450 catalyzing multiple steps in doxorubicin biosynthesis. *J. Bacteriol.* **181**:298–304.

Wang X, Tian X, Shen X, Wu Y, Yang S, Chen S, Road G. 2018a. Enhanced One-Step Fermentative Production of Epirubicin by Combination of Mutagenesis and Genetic Engineering in Doxorubicin-Producing Streptomyces peucetius. *J. Microbiol. Biotechnol.* **7**:22–30.

Wang X, Tian X, Wu Y, Shen X, Yang S, Chen S. 2018b. Enhanced doxorubicin production by Streptomyces peucetius using a combination of classical strain mutation and medium optimization. *Prep. Biochem. Biotechnol.* **48**:514–521. https://doi.org/10.1080/10826068.2018.1466156.

Weiss RB, Sarosy G, Clagett-Carr K, Russo M, Leyland-Jones B. 1986. Anthracycline analogs: The past, present, and future. *Cancer Chemother. Pharmacol.* **18**:185–197.

Weymouth-Wilson AC. 1997. The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* **14**:99–110.

Williams DH, Stone MJ, Hauck PR, Rahman SK. 1989. Why are secondary metabolites (Natural Products) biosynthesized. *J. Nat. Prod.* **52**:1189–1208.

Williams DH, Ma A, Wageningen V, Kirkpatrick PN, Harris BR, Kershaw JK, Lennard NJ, Jones M, Jm S, Solenberg PJ. 1998. Sequencing and anlysis of genes involved in the biosynthesis a vancomycin group antibiotic. *Chem. Biol.* **5**:155–162.

Williams GJ, Zhang C, Thorson JS. 2007. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**:657–662.

Ye J, Dickens ML, Plater R, Li Y, Lawrence J, Strohl WR. 1994. Isolation and sequence analysis of polyketide synthase genes from the daunomycin-producing Streptomyces sp. strain C5. *J. Bacteriol.* **176**:6270–6280.

Yu H, Takeuchi M, LeBarron J, Kantharia J, London E, Bakker H, Haltiwanger RS, Li H, Takeuchi H. 2015. Notch-modifying xylosyltransferase-substrate complexes support an SNi-like retaining mechanism. *Nat. Chem. Biol.* **11**:847–854.

Zechel DL, Withers SG. 2000. Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **33**:11–18.

Zhang J-H, Chung TDY, Oldenburg KR. 1999. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4**:67–73.

Zmijewski MJ, Briggs B. 1989. Biosynthesis of vancomycin: identification of TDP-glucose: aglycosyl-vancomycin glucosyltransferase from Amycolatopsis orientalis. *FEMS Microbiol. Lett.* **59**:129–134.

Appendix

List of Figures

List of Tables

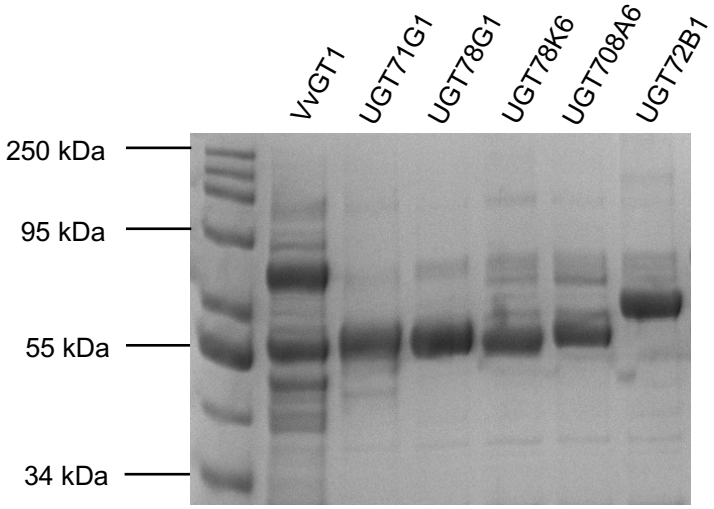


Figure A1: SDS-PAGE (4-12% Bis-Tris gel) of purified UGT collection. The proteins were overexpressed in *E. coli* strains and Ni-NTA purified. Following SDS-PAGE (0.25 – 0.50 mg/mL of final protein concentrations were loaded), the proteins were visualized by staining with Coomassie blue (**Table A5**). The molecular weights of the proteins were as follows: VvGT1: 50 kDa, UGT71G1: 52 kDa, UGT78G1: 50 kDa, UGT78K6: 49 kDa, UGT708A6: 51 kDa, and UGT72B1: 53 kDa.

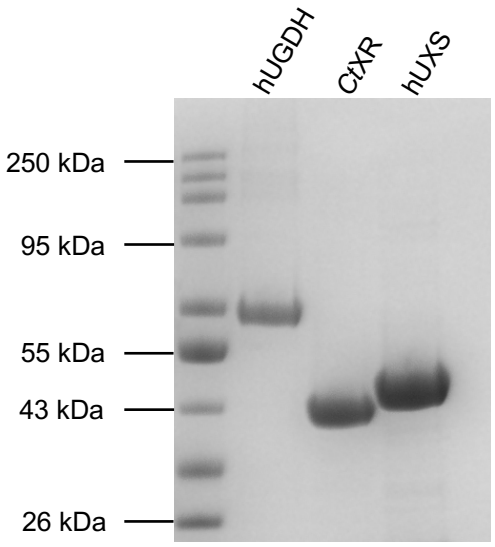


Figure A2: SDS-PAGE (4-12% Bis-Tris gel) of purified enzymes involved in UDP-xylose *in vitro* pathway. The proteins were overexpressed in *E. coli* strains and Ni-NTA purified. Following SDS-PAGE (0.25 – 0.75 mg/mL of final protein concentrations were loaded), the proteins were visualized by staining with Coomassie blue (**Table A5**). The molecular weights of the proteins were as follows: hUGDH: 55 kDa, *Ct*XR: 36 kDa, and hUXS: 48 kDa.

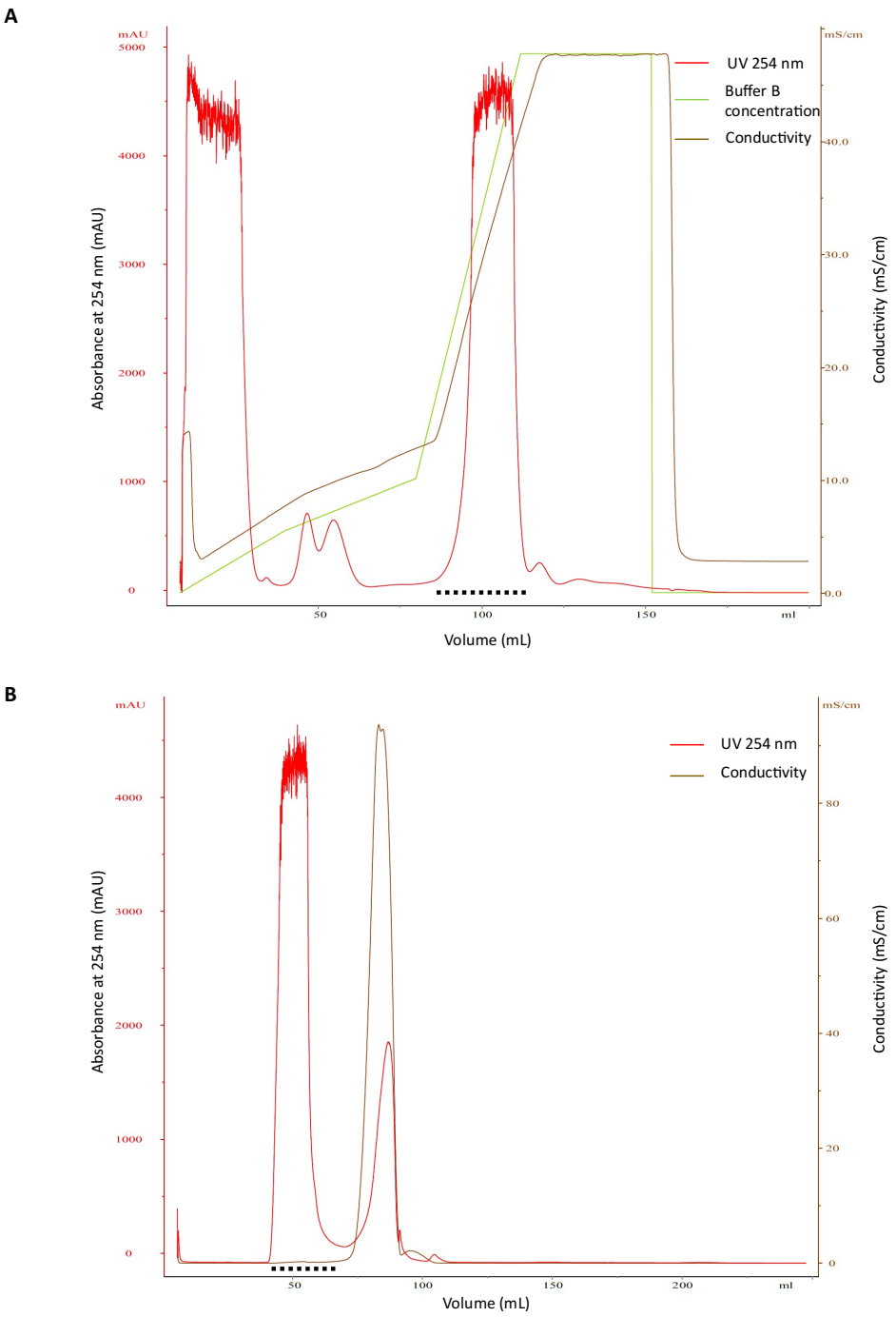


Figure A3: (A) Anion exchange and **(B)** size exclusion chromatograms of UDP-xylose purification. **(A)** UDP-xylose was purified using a Macro-prep high Q column with a 20-500 mM NH_4HCO_2 elution gradient. Subsequently, **(B)** salt was removed from UDP-xylose using a Sephadex G-10 column and eluted with dH_2O. The dotted lines denote fractions containing UDP-xylose.

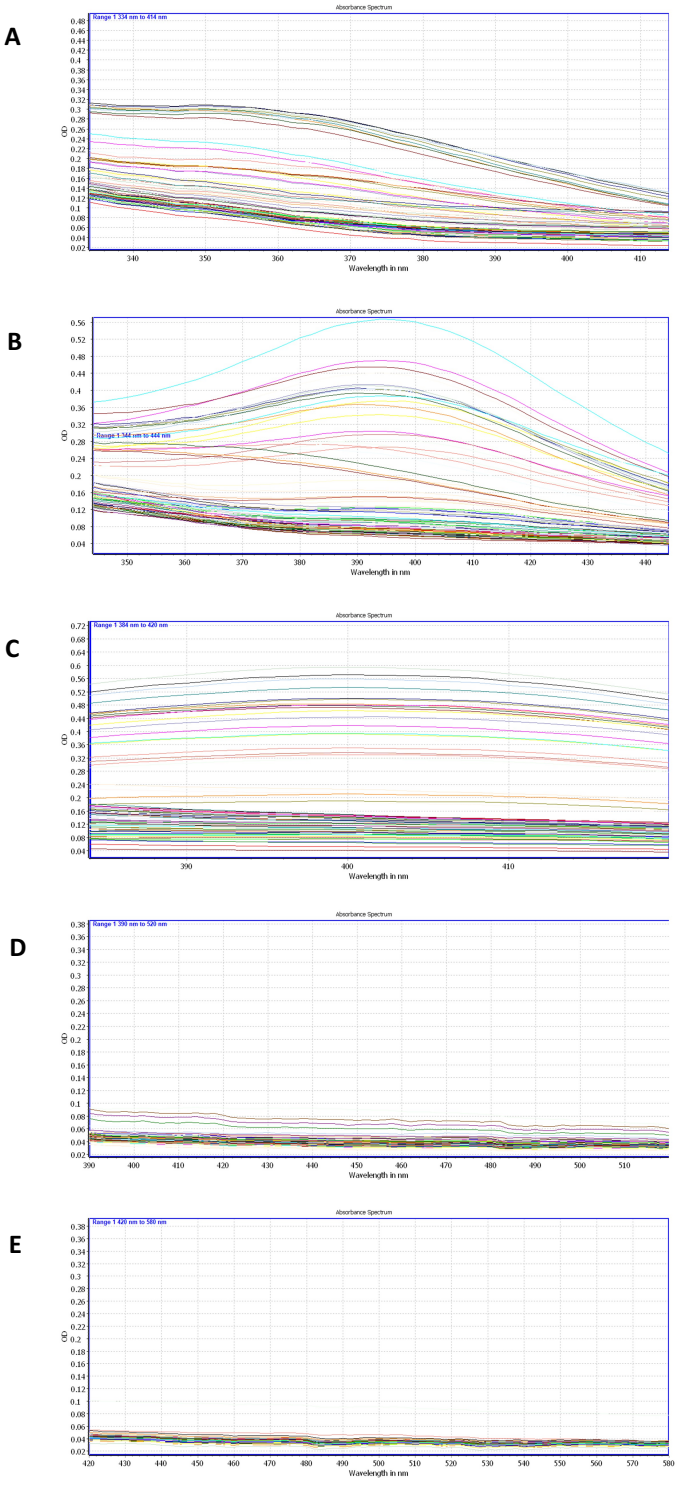


Figure A4: Raw absorbance spectrum data for combinatorially screened UGT enzyme collection with UDP-sugar donors and **(A)** kaempferol, **(B)** quercetin, **(C)** myricetin, **(D)** emodin, and **(E)** ε-rhodomycinone.

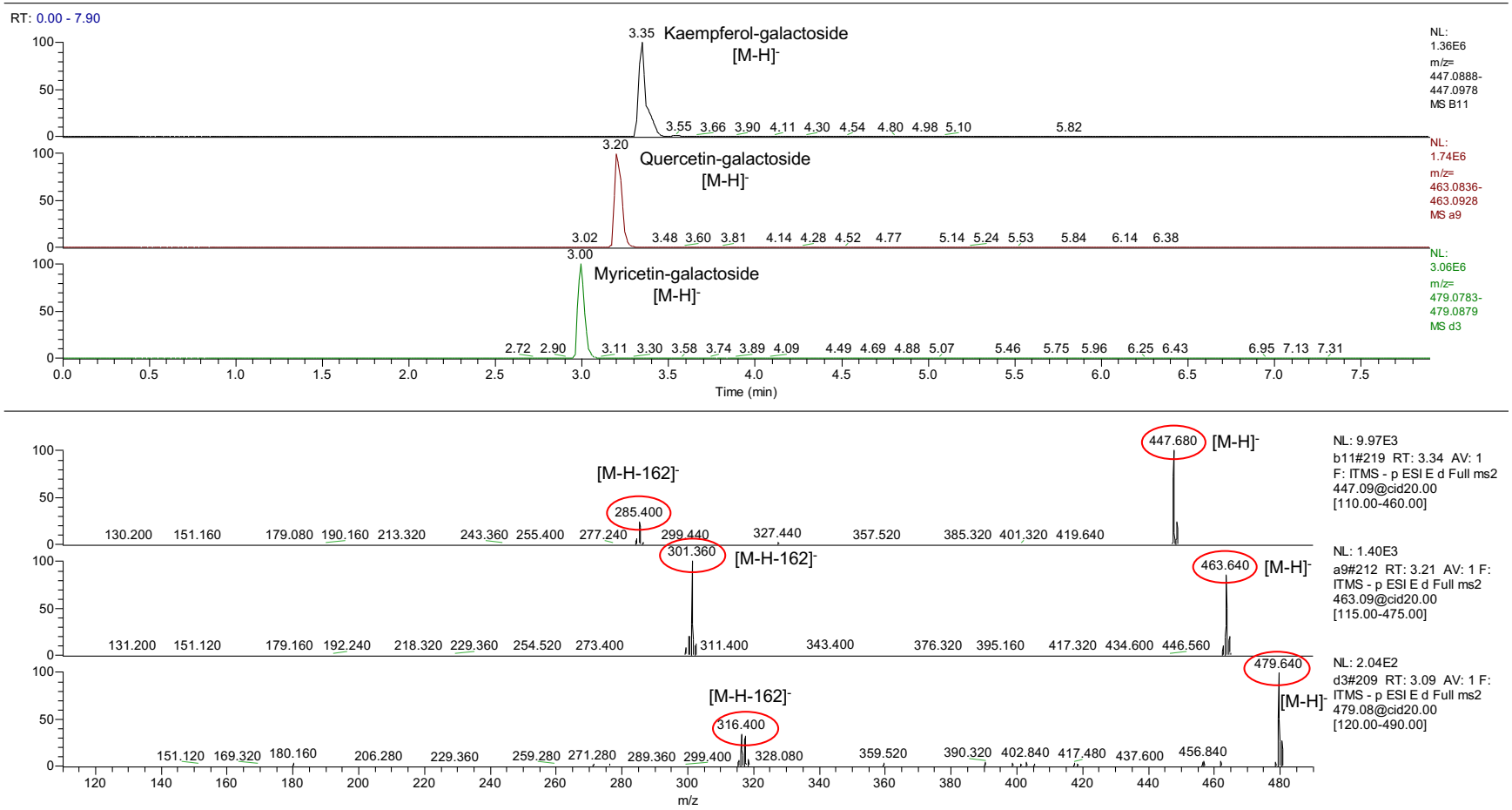


Figure A5: (A) LC-MS/MS analysis of UGT78K6 activity with kaempferol, quercetin, and myricetin acceptors in the presence of UDP-galactose. Top 3 panels: LC chromatograms. Bottom 3 panels: respective MS/MS fragmentation profiles at 20 cid. Sample were analyzed in negative mode and loss of galactose is indicated by -162.

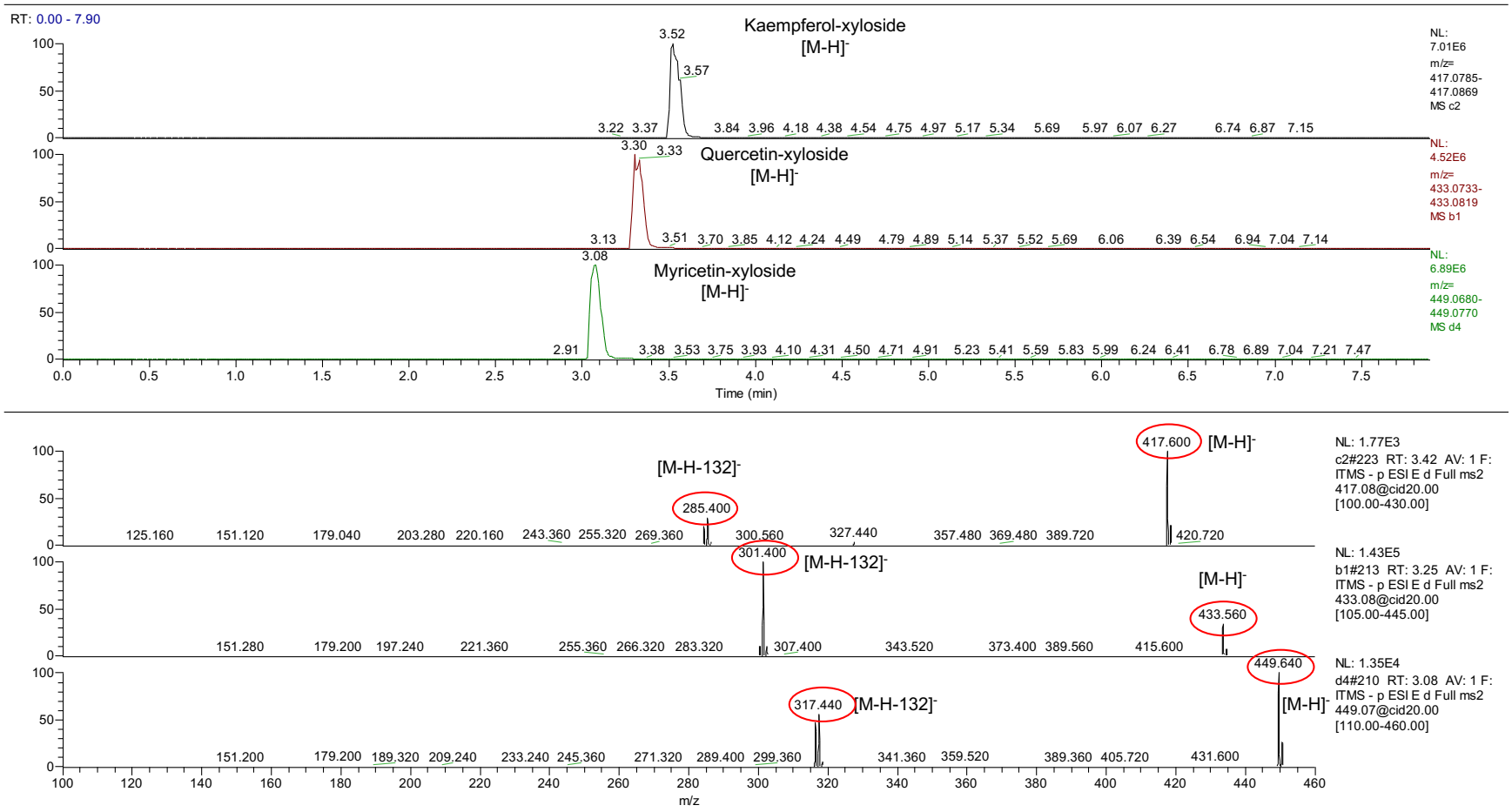


Figure A5: (B) LC-MS/MS analysis of UGT78G1 activity with kaempferol, quercetin, and myricetin acceptors in the presence of UDP-xylose. Top 3 panels: LC chromatograms. Bottom 3 panels: respective MS/MS fragmentation profiles at 20 cid. Sample were analyzed in negative mode and loss of xylose is indicated by -132.

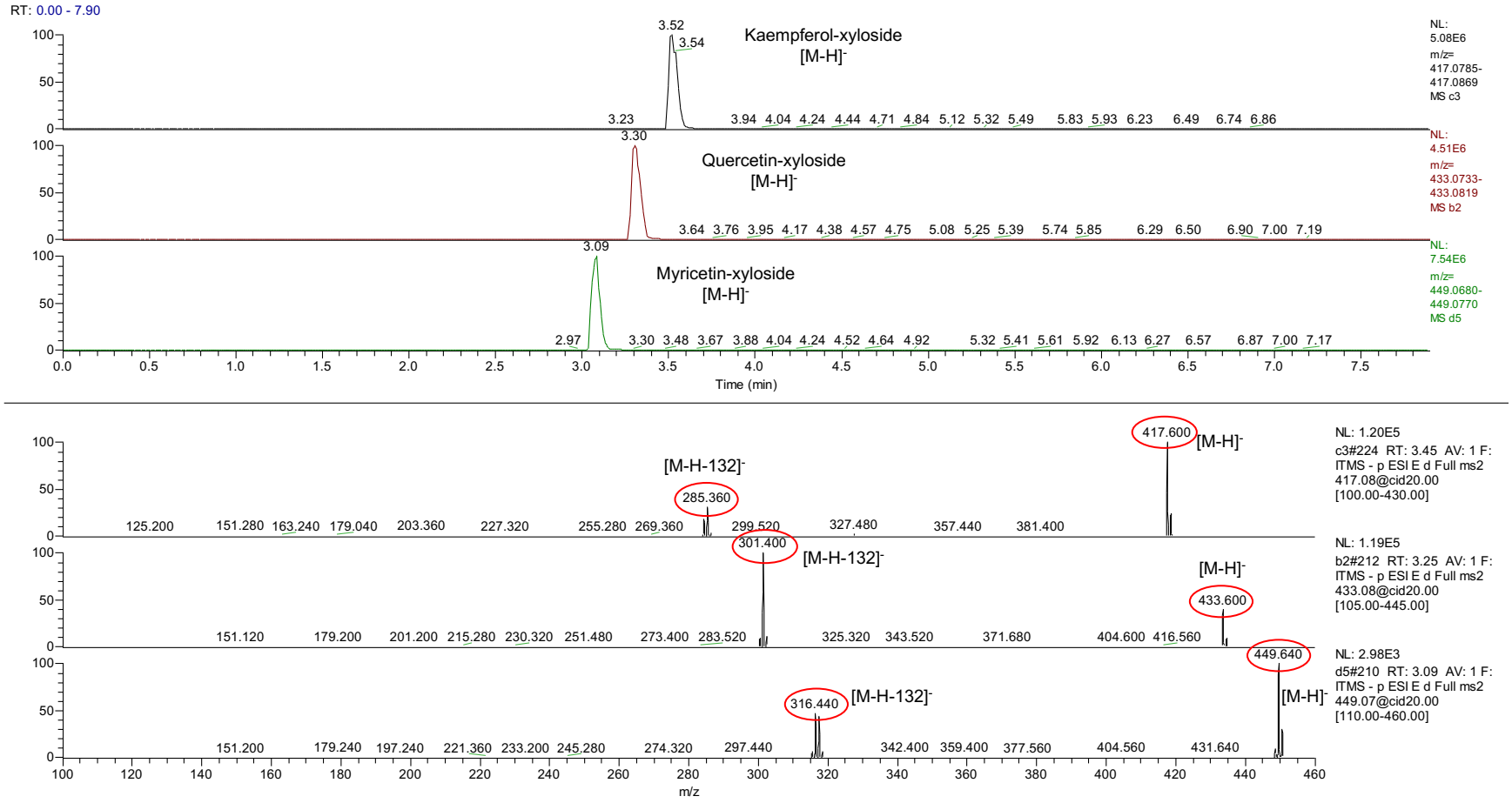


Figure A5: (C) LC-MS/MS analysis of UGT78K6 activity with kaempferol, quercetin, and myricetin acceptors in the presence of UDP-xylose. Top 3 panels: LC chromatograms. Bottom 3 panels: respective MS/MS fragmentation profiles at 20 cid. Sample were analyzed in negative mode and loss of xylose is indicated by -132.

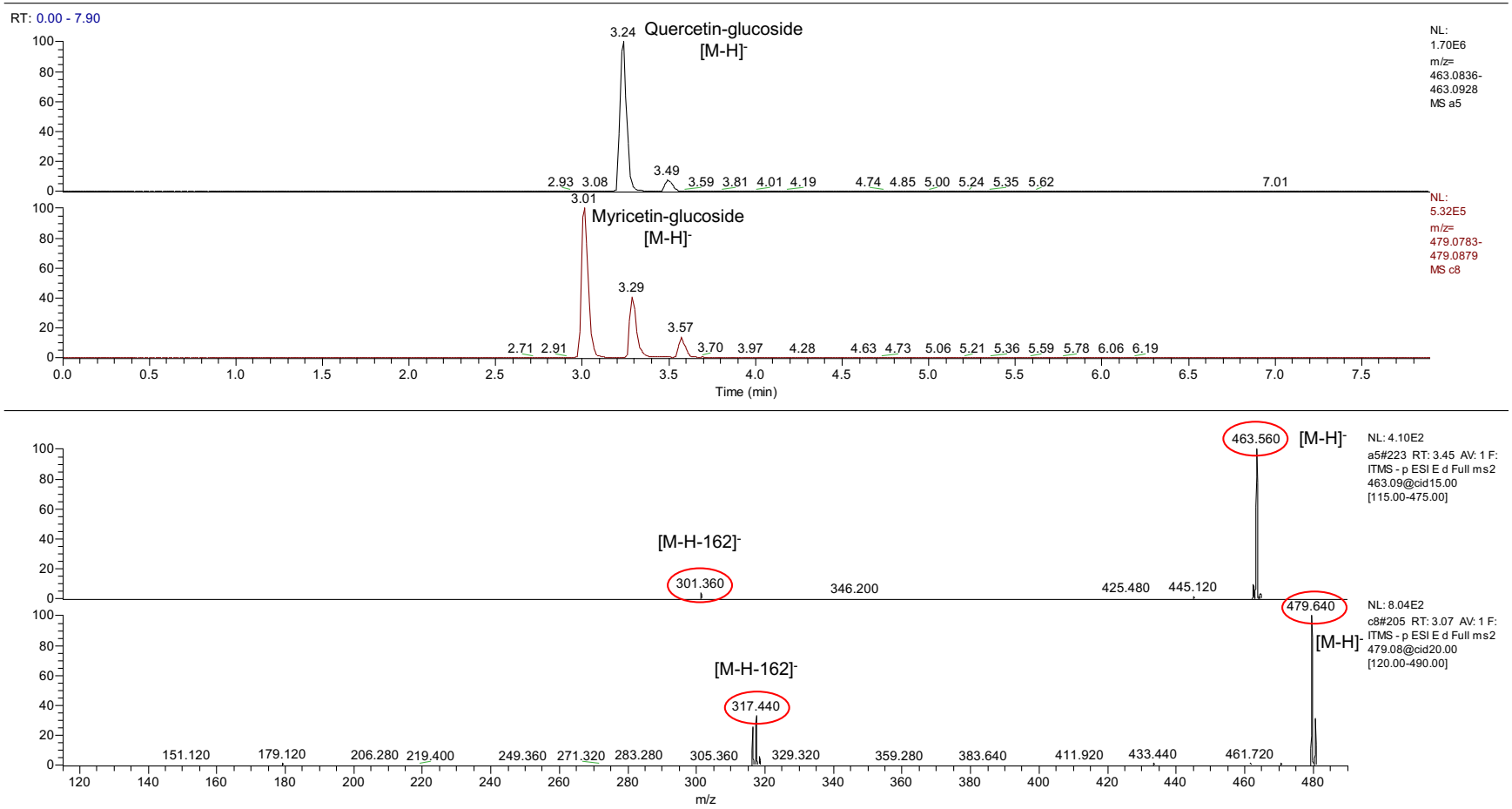


Figure A5: (D) LC-MS/MS analysis of UGT708A6 activity with quercetin and myricetin acceptors in the presence of UDP-glucose. Top 2 panels: LC chromatograms. Bottom 2 panels: respective MS/MS fragmentation profiles at 15 and 20 cid. Sample were analyzed in negative mode and loss of glucose is indicated by -162.

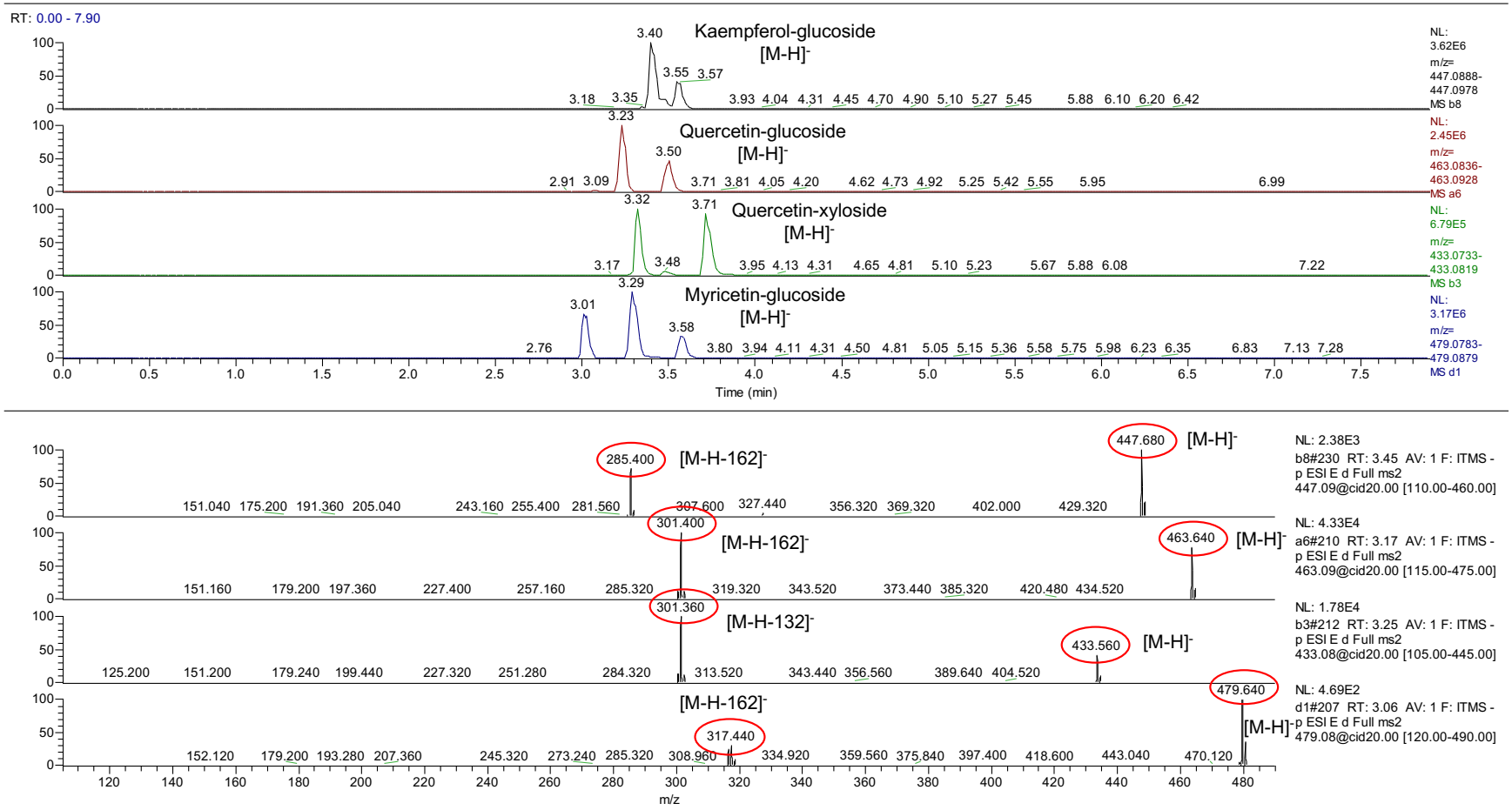


Figure A5: (E) LC-MS/MS analysis of UGT72B1 activity with kaempferol, quercetin, and myricetin acceptors in the presence of UDP-glucose/ UDP-xylose. Top 4 panels: LC chromatograms. Bottom 4 panels: respective MS/MS fragmentation profiles at 20 cid. Sample were analyzed in negative mode and loss of glucose and xylose are indicated by -162 and -132, respectively.

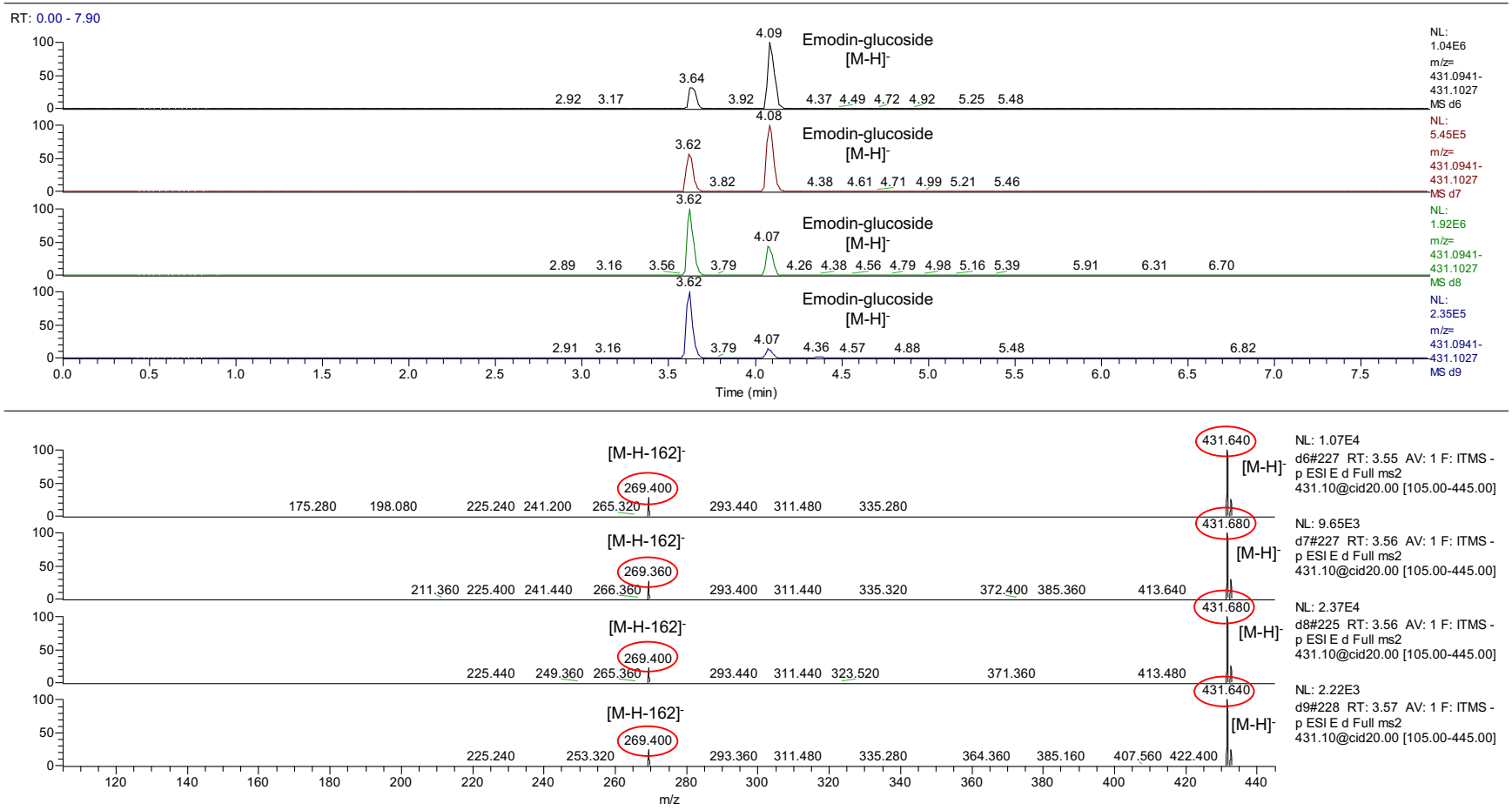


Figure A5: (F) LC-MS/MS analysis of UGT activity with emodin acceptor in the presence of UDP-glucose. Top 4 panels: LC chromatograms of UGT71G1, UGT78G1, UGT78K6, and UGT72B1 activities (from top to bottom). Bottom 4 panels: respective MS/MS fragmentation profiles at 20 cid. Sample were analyzed in negative mode and loss of glucose is indicated by -162.

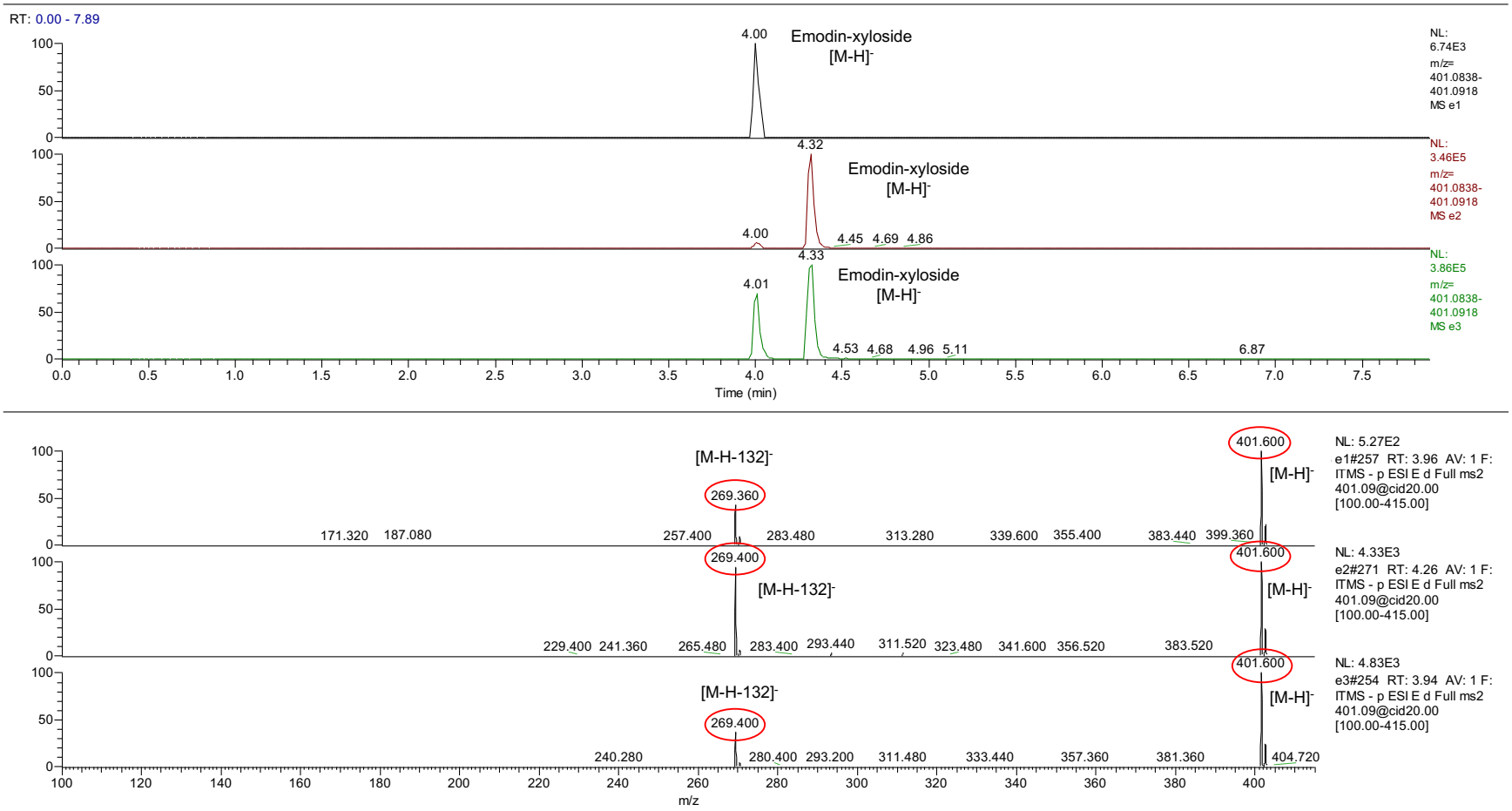


Figure A5: (G) LC-MS/MS analysis of UGT activity with emodin acceptor in the presence of UDP-xylose. Top 3 panels: LC chromatograms of VvGT1, UGT71G1, and UGT78K6 activities (from top to bottom). Bottom 3 panels: respective MS/MS fragmentation profiles at 20 cid. Sample were analyzed in negative mode and loss of xylose is indicated by -132.

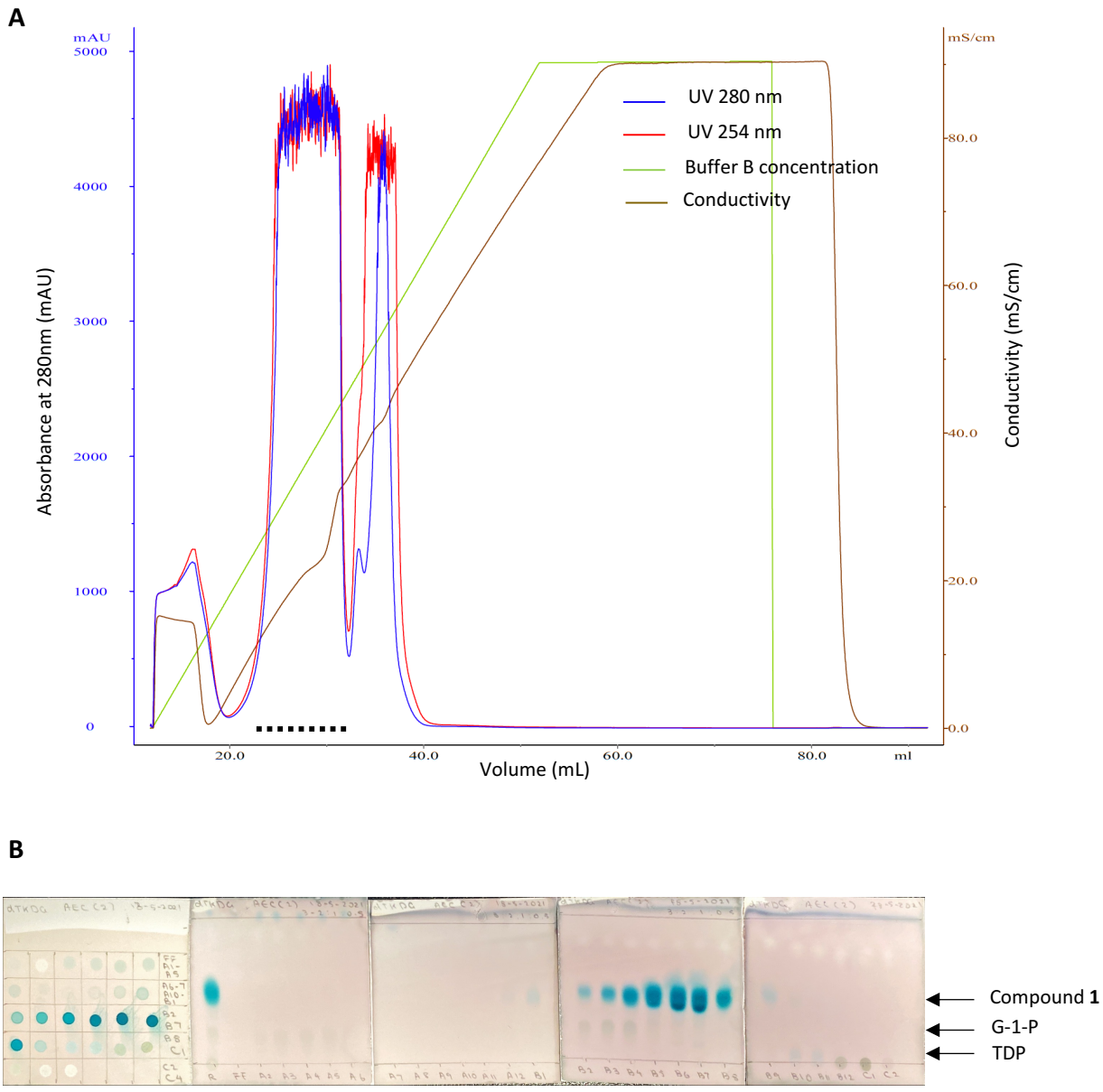


Figure A6: (A) Anion exchange chromatogram of compound **1** purified using a Macro-prep high Q column with a 0 - 1 M NaCl elution gradient. The dotted line represents fractions containing compound **1**. **(B)** Fractions of anion exchange chromatography were verified by TLC using ethyl acetate: methanol: dH_2O: acetic acid (3:2:1:0.5 v/v) mobile phase, which was stained using *p*-anisaldehyde. G-1-P: glucose-1-phosphate (a precursor), TDP: deoxythymidine diphosphate (an intermediate).

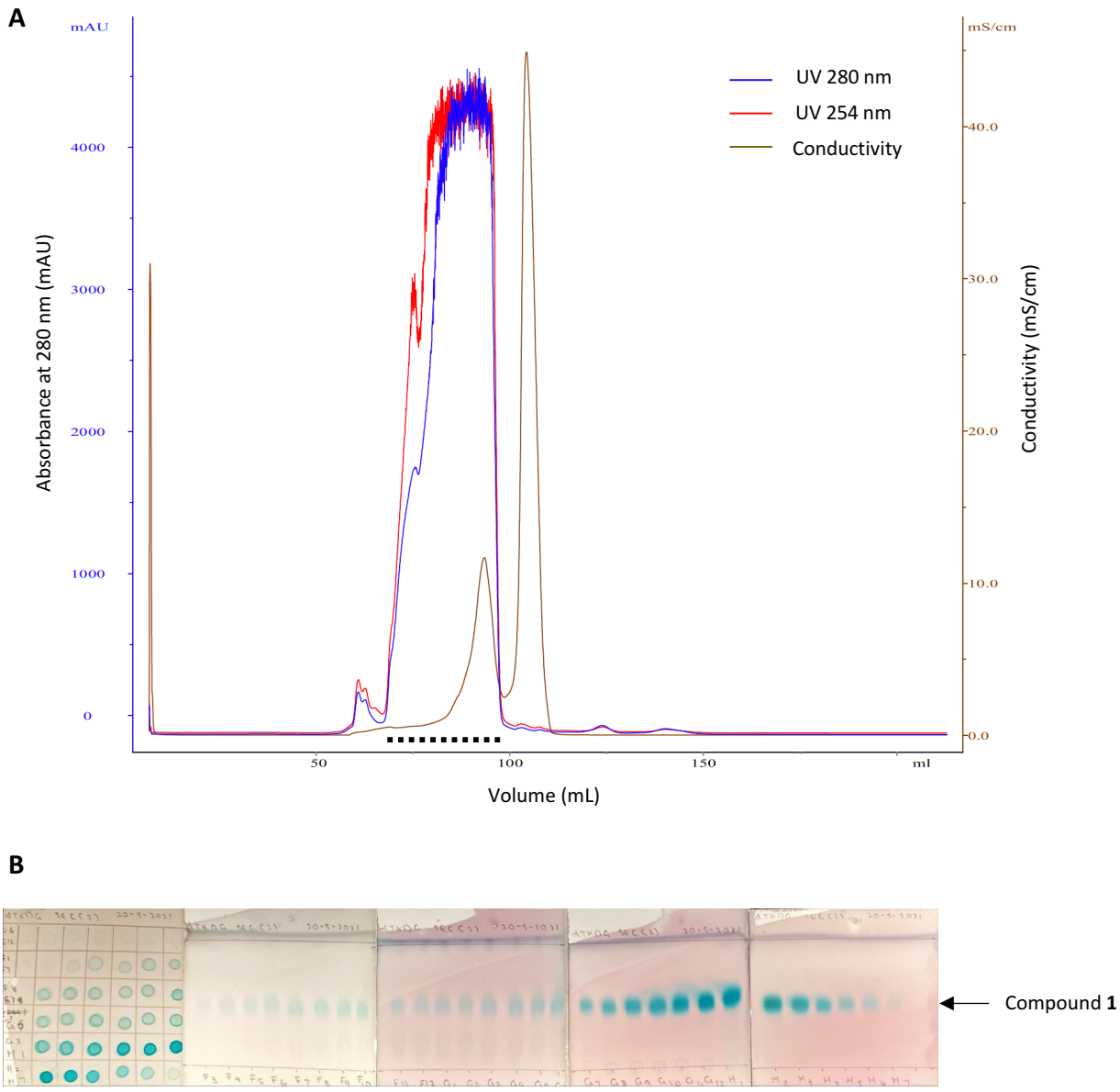


Figure A7: (A) Size exclusion chromatogram of compound **1** desalted using a Bio-Gel P-2 (polyacrylamide) resin and deionized water. The dotted line represents fractions containing compound **1**. **(B)** Fractions of size exclusion chromatography were verified by TLC using ethyl acetate: methanol: dH_2O: acetic acid (3:2:1:0.5 v/v) mobile phase, which was stained using *p*-anisaldehyde.

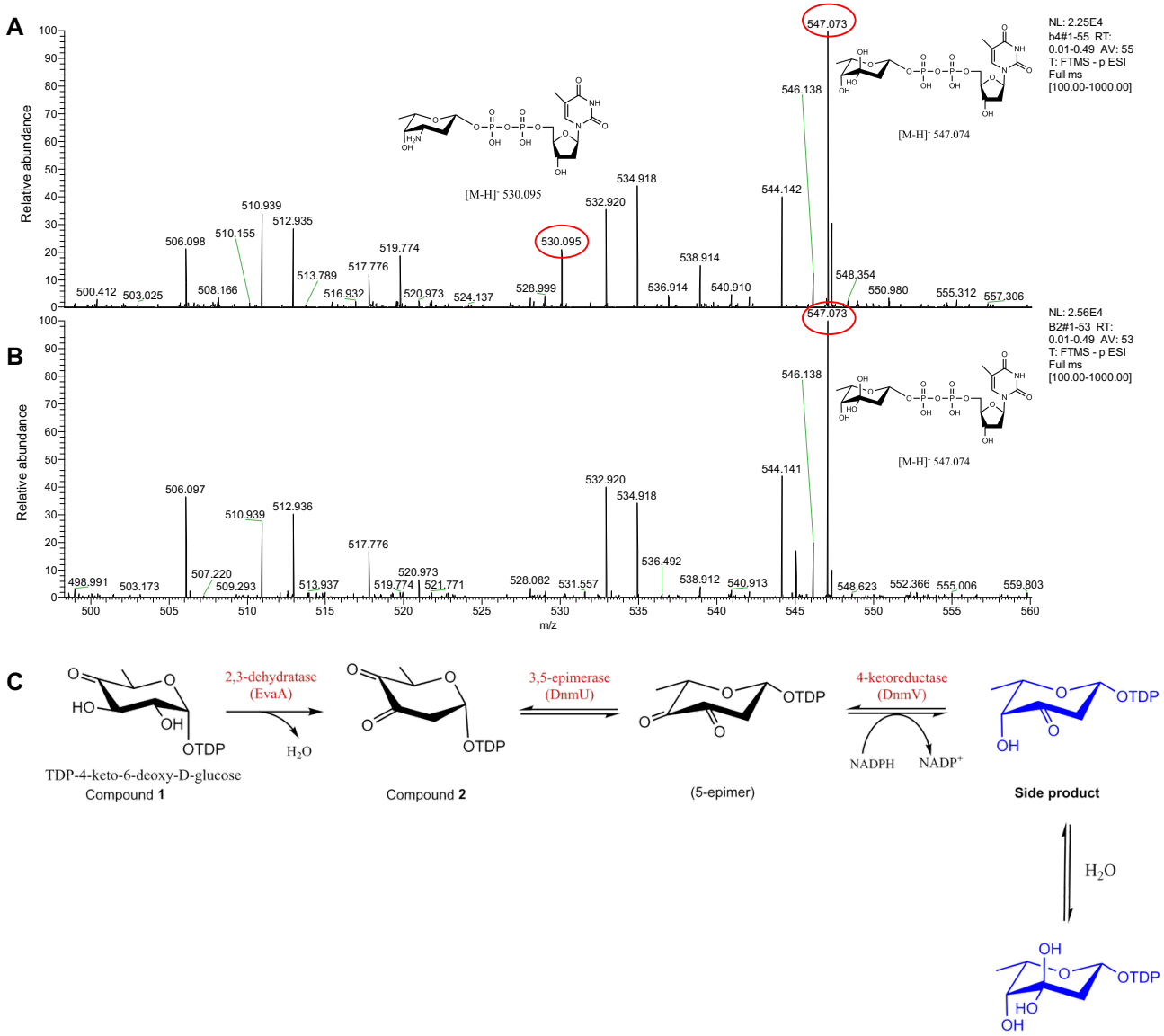


Figure A8: MS spectra of **(A)** TDP-L-daunosamine synthesis reaction (in a one-pot one-step) starting from compound **1** and adding EvaA, DnmJ, DnmU, and DnmV together, and **(B)** reaction with compound **1** and EvaA, DnmU, and DnmV enzymes. Samples were analyzed in negative mode and the calculated *m/z* ratios of TDP-L-daunosamine and the hydrated form of the side product were 530.095 and 547.074, respectively. **(C)** The sequential enzymatic activity of EvaA, DnmU, and DnmV results in the shunt product and its hydrated form that are shown in blue.

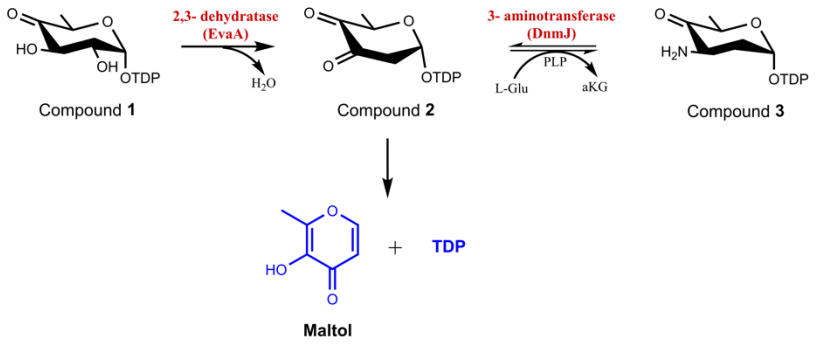


Figure A9: Schematic representation of the decomposition of compound **2** into maltol and TDP.

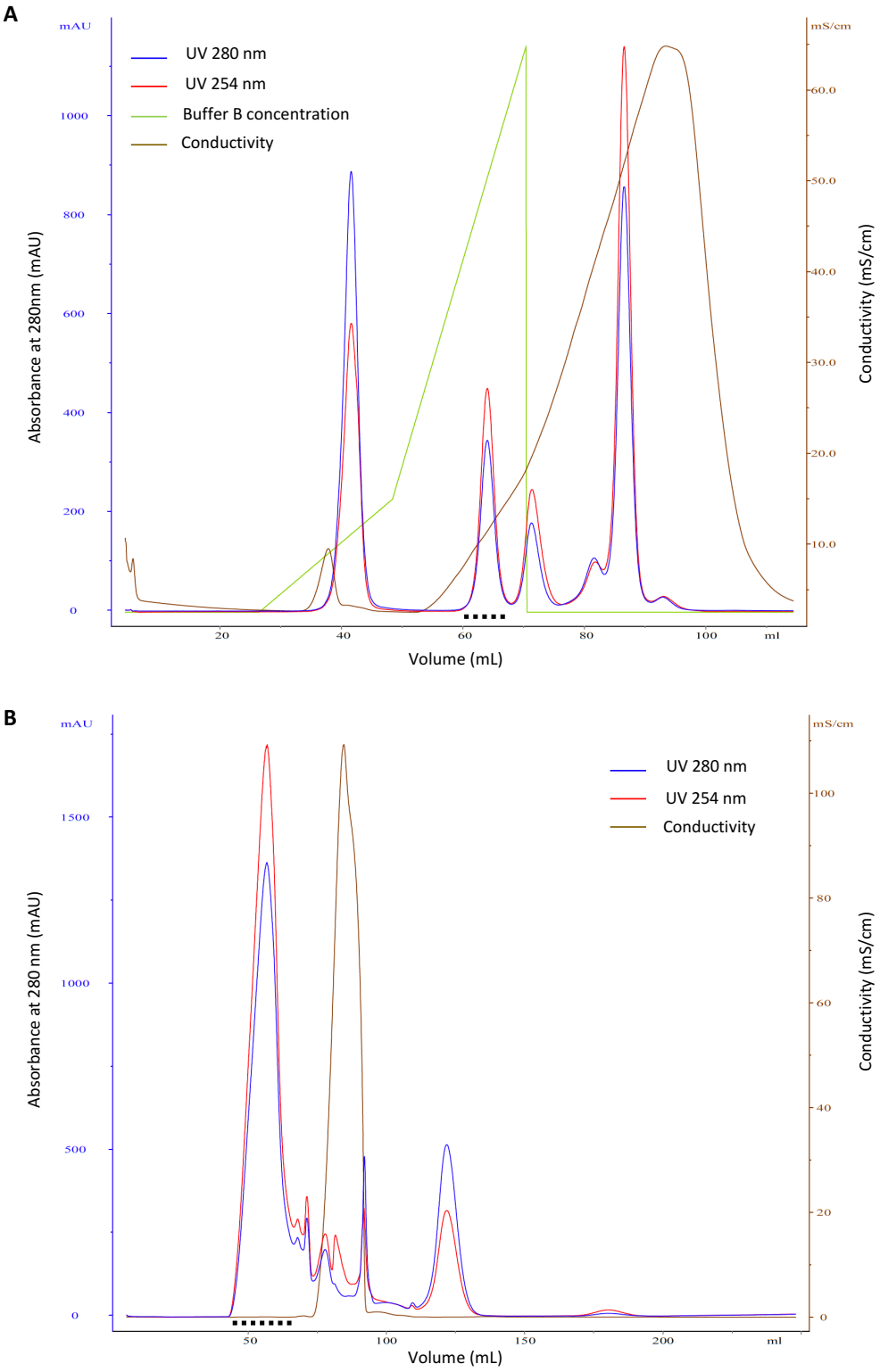


Figure A10: (A) Anion exchange chromatogram of compound **3** purified using Sephadex Q column with a NH_4HCO_2 gradient elution. **(B)** Size exclusion chromatogram of compound **3** desalted using Sephadex G-10 column and deionized water. The dotted lines represent fractions containing compound **3**.

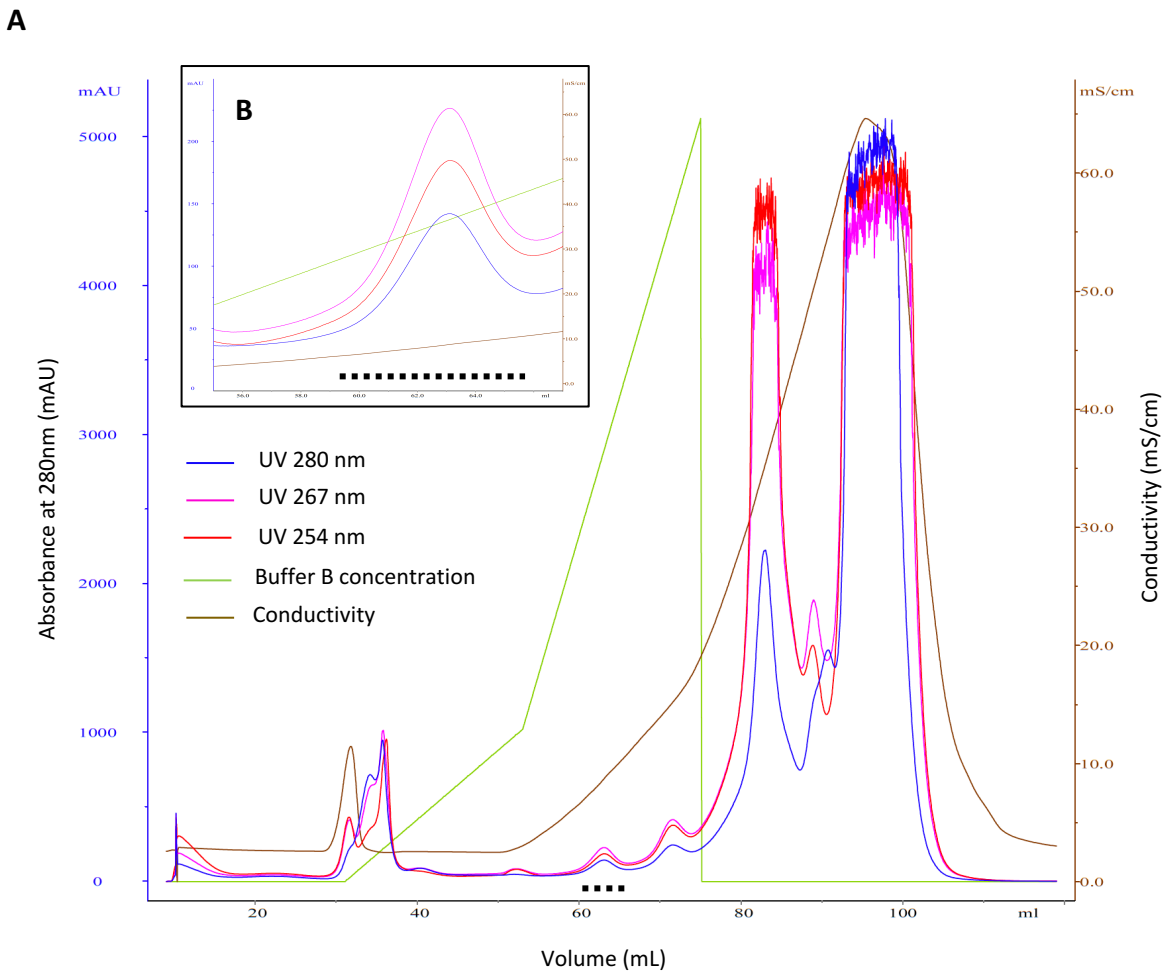


Figure A11: (A) Anion exchange chromatogram of TDP-L- daunosamine purified using Sephadex Q column with a NH_4HCO_2 gradient elution and **(B)** TDP-L-daunosamine elution region of the chromatogram. The dotted lines represent fractions containing TDP-L-daunosamine.

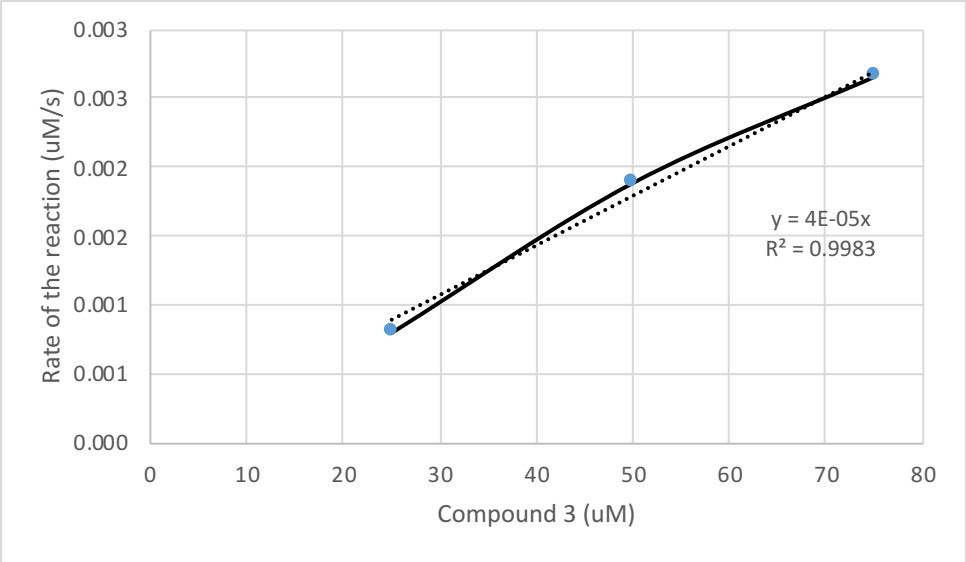


Figure A12: The linear fit for DnmJ enzyme assuming the substrate concentration is negligible in comparison to the K_M value.

Table A1 Oligonucleotides used in Chapter 3 and 4

Name	Sequence (5'-3')
hUXS_F	TATACATATGGAGAAAGACCGTAAACGCATCCTG
ACK_ex_for	TCGCATATGTCGAGTAAGTTAGTACTGGTT
ACK_ex_rev	TCGGAATTCTCAGGCAGTCAGGCGGCTCGCGTC
TMK_ex_for	GCTGAATTCATGCGCAGTAAGTATATCGTCATTGAG
TMK_ex_rev	CTGGAAGCTTCATGCGTCCAACTCCTTCACCCA
RmlB_for	TTGGATCCATGAAAATACTTGTTACTGGTGGCGCAGGA
RmlB_rev	ATTCTCGAGTTACTGGCGGCCCTCATAGTTCTGTTC
Sen_RmlA_NdeI.fw	GCCATATGAAAACGCGTAAGG
Sen_RmlA_EcoRI.rv	CTCGAATTCTTATAAACCTTTCACCATC
EvaA_fw	GGGAATTTCCATATGTCGTCCTTCGTCGTTCCA
EvaA-rv	AAACCGCTCGAGTCATGCACCTCCCCGAGGCTG
dnmJ-NdeI.fw	ATACATATGTCCACCTACGTCTGGCAATAC
dnmJ-XhoI.rv	ATACTCGAGTCACAGGCTTCCGACGACCTC
dnmU-NdeI.fw	ATACATATGAAGGCGCGGGAACTGG
dnmU-XhoI.rv	ATACTCGAGTCACCCGTCTCCGCGTGACC
dnmV-NdeI.fw	ATACATATGCGGGTCGTGGTTCTGG
dnmV-XhoI.rv	ATACTCGAGCTAGGCCGGGGCGCCGTG
EvaE_fw	GGAGCATATGATGAAGCTGATCACCGTGCTCGG
EvaE_rv	ACACGCGCTCGAGTCATGCGCGAGCCTTTCCATGAAC
T7 promoter primer	TAATACGACTCACTATAGGG
T7 terminator primer	GCTAGTTATTGCTCAGCGG

Table A2 Vectors and plasmids used in chapter 2 to 4

Plasmid	Marker	Description	Source
pET14b	Amp^R	Bacterial expression plasmid with T7 promoter and N terminal His tag	Novagen
pET28a	Kan^R	Bacterial expression plasmid with T7 promoter and N and C terminal His tag	Novagen
pET14b_ VvGT1	Amp^R	pET14b based expression plasmid encoding VvGT1	Christopher Ford (Ford *et al*., 1998)
pUC57-Kan_ UGT71G1	Kan^R	pUC57-Kan based cloning plasmid encoding UGT71G1	Bio Basic
pUC57-Kan_ UGT78G1	Kan^R	pUC57-Kan based cloning plasmid encoding UGT78G1	Bio Basic
pUC57-Kan_ UGT78K6	Kan^R	pUC57-Kan based cloning plasmid encoding UGT778K6	Bio Basic
pUC57-Kan_ UGT708A6	Kan^R	pUC57-Kan based cloning plasmid encoding UGT708A6	Bio Basic
pET28a_ UGT71G1	Kan^R	pET28a based expression plasmid encoding UGT71G1	This study
pET28a_ UGT78G1	Kan^R	pET28a based expression plasmid encoding UGT78G1	This study
pET28a_ UGT778K6	Kan^R	pET28a based expression plasmid encoding UGT778K6	This study
pET28a_ UGT708A6	Kan^R	pET28a based expression plasmid encoding UGT708A6	This study
pET30a_ UGT72B1	Kan^R	pET30a based expression plasmid encoding UGT72B1	Gideon Davies (Brazier-Hicks *et al*., 2007b)
pUC57-Kan_ hUGDH	Kan^R	pUC57-Kan based cloning plasmid encoding hUGDH	Bio Basic
pUC57-Kan_ hUXS	Kan^R	pUC57-Kan based cloning plasmid encoding hUXS	Bio Basic
pUC57-Kan_ CtXR	Kan^R	pUC57-Kan based cloning plasmid encoding CtXR	Bio Basic

pET28a_ hUGDH	Kan^R	pET28a based expression plasmid encoding hUGDH	This study
pET28a_ hUXS	Kan^R	pET28a based expression plasmid encoding hUXS	This study
pET28a_∆84hUXS	Kan^R	pET28a based expression plasmid encoding ∆84hUXS	This study
pET28a_ *Ct*XR	Kan^R	pET28a based expression plasmid encoding *Ct*XR	This study
pET28a_ TmK	Kan^R	pET28a based expression plasmid encoding TmK	This study
pET28a_ AckA	Kan^R	pET28a based expression plasmid encoding AckA	This study
pET28a_ RmlA	Kan^R	pET28a based expression plasmid encoding RmlA	This study
pET28a_ RmlB	Kan^R	pET28a based expression plasmid encoding RmlB	This study
pET28a_ EvaA	Kan^R	pET28a based expression plasmid encoding EvaA	This study
pET28a_ DnmJ	Kan^R	pET28a based expression plasmid encoding DnmJ	This study
pET28a_ DnmU	Kan^R	pET28a based expression plasmid encoding DnmU	This study
pET28a_ DnmV	Kan^R	pET28a based expression plasmid encoding DnmV	This study
pET28a_ EvaE	Kan^R	pET28a based expression plasmid encoding EvaE	This study

Table A3 Strains used in Chapter 2 to 4

Strain	Relevant characteristic
E. coli DH10B	F^-, *mcrA*, Δ(*mrr-hsd*RMS*mcr*BC), φ80d*lacZ*ΔM15, Δ*lac*X74, *deo*R, *rec*A1, *end*A1, *ara*D139, Δ (*ara-leu*)7697, *gal*U, *gal*K, λ^-, *rps*L, *nup*G Routine cloning
E. coli BL21(DE3)	F^-, *omp*T, $hsdS_B(r_B^- m_B^-)$ *gal*, *dcm*, λ(DE3) Protein expression host. Encodes T7 RNA polymerase (under control of lacUV5 promoter).
E. coli XJb(DE3)	F^-, *omp*T, $hsdS_B(r_B^- m_B^-)$ *gal*, dcm^+ *araB*::*R*,cat λDE3 Autolysis strain. Encodes for the bacteriophage λ endolysin protein, and T7 polymerase.
E. coli BL21-CodonPlus(DE3)-RP containing plasmid L1SL2/pETcoco-2 (Moncrieffe *et al*., 2012)	F– *ompT hsdS*($r_B^- m_B^-$) dcm^+ Tet^r *gal* λ(DE3) *end*A Hte [*argU proL* Cam^r] Protein expression host. Contains extra gene copies of rare tRNAs. Encodes T7 RNA polymerase (under control of lacUV5 promoter). L1SL2/pETcoco-2: pETcoco-2 expression vector encoding GroEL1, GroEL2 and GroES chaperones

Table A4 Culture media used in Chapter 2 to 4

LB	
Tryptone	10 g
Yeast extract	5 g
NaCl	10 g
(Agar)	(15 g)
Distilled water	To final volume of 1L
LBE-5052 autoinduction media	
Tryptone	10 g
Yeast extract	5 g
NH_4Cl	2.68 g
Na_2SO_4	0.71 g
Glycerol	5 g
Glucose	0.5 g
Lactose	2 g
Distilled water	898 mL
2 M $MgSO_4$ (*filter sterilized*)	1 mL
1000X Trace metals mix (*filter sterilized*)	1 mL
1 M K_2HPO_4 (*filter sterilized*)	40 mL
1 M KH_2PO_4 (*filter sterilized*)	10 mL
1000X Trace metals mix	50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2$, 10 mM $ZnSO_4$, 2 mM $CoCl_2$, 2 mM $CuCl_2$, 2 mM $NiCl_2$

Added after autoclaving

Table A5 Solutions, buffers, and stains used in chapter 2 to 4

Solution	Components
TFB1 (*filter sterilized*)	30 mM KOAc, 10 mM $CaCl_2$, 50 mM $MnCl_2$, 100 mM RbCl, 15% Glycerol (pH 5.8: adjust with AcOH)
TBF2 (*filter sterilized*)	100 mM MOPS or PIPES, 75 mM $CaCl_2$, 10 mM RbCl, 15% Glycerol (pH 6.5: adjust with KOH)
VvGT1 lysis buffer	20 mM Tris pH 8, 100 mM NaCl, 10 mM 2-mercaptoethanol, 10 % v/v glycerol
VvGT1 wash buffer	20 mM Tris pH 8, 100 mM NaCl, 10 mM 2-mercaptoethanol, 10 % v/v glycerol, 10 mM imidazole
VvGT1 elution buffer	20 mM Tris pH 8, 100 mM NaCl, 10 mM 2-mercaptoethanol, 10 % v/v glycerol, and 300 mM imidazole
VvGT1 storage buffer	50 mM Tris pH 7.5, 150 mM NaCl, 10 mM 2-mercaptoethanol, and 10 % v/v glycerol
UGT lysis/ wash buffer	50 mM Tris pH 8, 500 mM NaCl, 10 mM 2-mercaptoethanol, and 10 mM imidazole
UGT elution buffer	50 mM Tris pH 8, 500 mM NaCl, 10 mM 2-mercaptoethanol, and 250 mM imidazole
UGT storage buffer	50 mM Tris pH 7.5, 150 mM NaCl, and 5 % v/v glycerol
Lysis/ wash buffer I	50 mM Tris pH 8, 500 mM NaCl, and 10 mM imidazole
Elution buffer I	50 mM Tris pH 8, 500 mM NaCl, and 500 mM imidazole
Assay control buffer/ Storage buffer I	50 mM Tris pH 7.5, 150 mM NaCl
Buffer II	20 mM ammonium formate pH 4.2
Buffer III	500 mM ammonium formate pH 4.2
Buffer IV	Water with 0.1 % formic acid
Buffer V	Acetonitrile with 0.1% formic acid
Lysis buffer VI	50 mM BTP pH 7.5, 5 mM imidazole
Wash buffer VI* (*Pre-equilibration and wash step following* DnmJ *sample loading*)	50 mM BTP pH 7.5, 500 mM NaCl, 20 mM imidazole

*Excluded for EvaA

Wash buffer VI (*AKTA purifier 10 system*)	50 mM BTP pH 7.5, 150 mM NaCl, 20 mM imidazole
Elution buffer VI	50 mM BTP pH 7.5, 150 mM NaCl, 300 mM imidazole
Storage buffer VI	50 mM BTP pH 7.5, 150 mM NaCl
Buffer VII	1 M NaCl
Buffer VIII	20 mM NH_4HCO_2 pH 3.5
Buffer IX	1 M NH_4HCO_2 pH 6.7
p-anisaldehyde stain	135 mL of absolute ethanol, 5 mL of conc. sulfuric acid, 1.5 mL of glacial acetic acid, 3.7 mL of *p*-anisaldehyde
Coomassie blue stain	1 g of Coomassie brilliant blue in 1 L of 50% dH2O, 40 % methanol, and 10% acetic acid
Destaining solution	50% dH2O, 40 % methanol, and 10% acetic acid