



Turun yliopisto  
University of Turku

# GLYCOSYLATION AND GLYCODIVERSIFICATION IN POLYKETIDE ANTIBIOTICS: UNRAVELING BIOSYNTHETIC STEPS IN NOGALAMYCIN FORMATION

---

Vilja Siitonen

## University of Turku

---

Faculty of Mathematics and Natural Sciences  
Department of Biochemistry  
Biochemistry  
Doctoral Programme in Molecular Life Sciences

## Supervised by

---

Docent Mikko Metsä-Ketelä, Ph.D.  
Department of Biochemistry  
University of Turku

## Reviewed by

---

Docent David Fewer, Ph.D.  
Department of Food and Environmental Sciences  
University of Helsinki  
Helsinki, Finland

Gerald Lackner, Ph.D.  
The Leibniz Institute for Natural Product Research  
and Infection Biology -  
Hans Knöll Institute  
Jena, Germany

## Opponent

---

Prof. Andriy Luzhetskyy, Ph.D.  
Helmholtz Institute for Pharmaceutical Research,  
Saarbrücken, Germany

Cover image by Vilja Siitonen

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-6543-4 (PRINT)

ISBN 978-951-29-6544-1 (PDF)

ISSN 0082-7002 (Print)

ISSN 2343-3175 (Online)

Painosalama Oy - Turku, Finland 2016

For the next generation

**TABLE OF CONTENTS**

<b>LIST OF ORIGINAL PUBLICATIONS.....</b>	<b>6</b>
<b>ABSTRACT .....</b>	<b>7</b>
<b>TIIVISTELMÄ .....</b>	<b>8</b>
<b>ABBREVIATIONS.....</b>	<b>9</b>
<b>AMINO ACID ABBREVIATIONS.....</b>	<b>11</b>
<b>1. INTRODUCTION .....</b>	<b>12</b>
<b>2. LITERATURE REVIEW .....</b>	<b>14</b>
2.1. GLYCOSYLTRANSFERASES .....	14
2.1.1. <i>Glycosyltransferase Folds</i> .....	15
2.1.2. <i>Sugar Linkage</i> .....	18
2.1.3. <i>Substrates for the Glycosyltransferase</i> .....	18
2.1.3.1. The Acceptor Substrate .....	18
2.1.3.2. The Donor Substrate: the Activation of the Sugar.....	19
2.1.3.3. The Donor Substrate: the Modification of the Sugar.....	21
2.1.4. <i>Mode of Action of the Glycosyltransferases</i> .....	24
2.1.5. <i>Auxiliary Proteins of Glycosyltransferases</i> .....	27
2.2. GLYCODIVERSIFICATION .....	29
2.2.1. <i>Engineering Enzymes Involved in the Sugar Biosynthesis</i> .....	30
2.2.1.1. Anomeric Position Phosphorylating Enzymes .....	30
2.2.1.2. Nucleotidyltransferases .....	31
2.2.1.3. Glycosyltransferases .....	31
2.2.2. <i>Approaches in Glycodiversification</i> .....	33
2.2.2.1. <i>In vitro</i> Approaches in Glycodiversification .....	33
2.2.2.2. <i>In vivo</i> Approaches in Glycodiversification .....	35
2.2.2.3. Comparison of the <i>in vitro</i> and <i>in vivo</i> Approaches in Glycodiversification.....	38
<b>3. AIMS OF STUDY.....</b>	<b>40</b>
<b>4. MATERIALS AND METHODS .....</b>	<b>41</b>
4.1. MOLECULAR BIOLOGY .....	41
4.2. HETEROLOGOUS EXPRESSION, PURIFICATION AND ANALYSIS OF PROTEINS .....	41
4.3. ENZYMATIC ASSAYS .....	42
4.4. HETEROLOGOUS EXPRESSION OF THE NOGALAMYCIN GENE CLUSTER AND GENE INACTIVATION MUTANTS.....	42
4.5. EXPRESSION, ISOLATION AND PURIFICATION OF COMPOUNDS .....	43
4.6. ANALYSIS OF THE COMPOUNDS.....	43
4.6.1. <i>Thin Layer Chromatography</i> .....	43

## Table of Contents

---

4.6.2. High Performance Liquid Chromatography.....	44
4.6.3. Mass Spectrometry.....	44
4.6.4. Nuclear Magnetic Resonance .....	44
<b>5. RESULTS AND DISCUSSION .....</b>	<b>45</b>
5.1. HETEROLOGOUS EXPRESSION OF THE NOGALAMYCIN GENE CLUSTER (I) .....	45
5.2. UNUSUAL MONOOXYGENASE SYSTEM: SNOAW AND SNOAL2 (II).....	49
5.3. STRUCTURE OF THE GLYCOSYLTRANSFERASE SNOGD (III) .....	50
5.4. A-KG-DEPENDENT DIOXYGENASES: SNOK AND SNOB (IV).....	52
5.4.1. Carbon–carbon Bond Formation.....	52
5.4.2. Stereochemistry of C4'' .....	53
5.4.3. Comparison of <i>SnoK</i> and <i>SnoB</i> .....	55
<b>6. CONCLUDING REMARKS AND FUTURE PERSPECTIVE.....</b>	<b>56</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>58</b>
<b>REFERENCES .....</b>	<b>60</b>
<b>REPRINTS OF THE ORIGINAL PUBLICATIONS .....</b>	<b>75</b>

## LIST OF ORIGINAL PUBLICATIONS

- I **Siitonen, V.**, Claesson, M., Patrikainen, P., Aromaa, M., Mäntsälä, P., Schneider, G., and Metsä-Ketelä, M. (2012) Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *ChemBioChem*, 13(1): 120–128.
- II **Siitonen, V.**, Blauenburg, B., Kallio, P., Mäntsälä, P., and Metsä-Ketelä, M. (2012) Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chemistry & Biology*, 19(5): 638–646.
- III Claesson, M., **Siitonen, V.**, Dobritzsch, D., Metsä-Ketelä, M., and Schneider, G. (2012) Crystal structure of the glycosyltransferase SnogD from the biosynthetic pathway of nogalamycin in *Streptomyces nogalater*. *FEBS journal*, 279(17): 3251–3263.
- IV **Siitonen V.**, Selvaraj, B., Niiranen, L., Lindqvist, Y., Schneider, G., and Metsä-Ketelä, M. (2016) Divergent non-heme iron enzymes in the nogalamycin biosynthetic pathway. *Proc. Natl. Acad. Sci. U. S. A.*, 113(19):5251–5256.

The original articles I and III were printed with the permission from © John Wiley and Sons and II from © Elsevier, IV with the permission from National Academy of Sciences, USA.

## ABSTRACT

Soil-dwelling *Streptomyces* bacteria are known for their ability to produce biologically active compounds such as antimicrobial, immunosuppressant, antifungal and anticancer drugs. *S. nogalater* is the producer of nogalamycin, a potential anticancer drug exhibiting high cytotoxicity and activity against human topoisomerases I and II. Nogalamycin is an anthracycline polyketide comprising a four-ring aromatic backbone, a neutral deoxy sugar at C7, and an amino sugar attached via an O–C bond at C1 and a C–C bond between C2 and C5'. This kind of attachment of the amino sugar is unusual thus making the structure of the compound highly interesting. The sugar is also associated with the biological activity of nogalamycin, as it facilitates binding to DNA. Furthermore, the sugar moieties of anthracyclines are often crucial for their biological activity. Together the interesting attachment of the amino sugar and the general reliance of polyketides on the sugar moieties for bioactivity have made the study of the biosynthesis of nogalamycin attractive.

The sugar moieties are typically attached by glycosyltransferases, which use two substrates: the donor and the acceptor. The literature review of the thesis is focused on the glycosylation of polyketides and the possibilities to alter their glycosylation patterns.

My own thesis work revolves around the biosynthesis of nogalamycin. We have elucidated the individual steps that lead to its rather unique structure. We reconstructed the whole biosynthetic pathway in the heterologous host *S. albus* using a cosmid and a plasmid. In the process, we were able to isolate new compounds when the cosmid, which contains the majority of the nogalamycin gene cluster, was expressed alone in the heterologous host. The new compounds included true intermediates of the pathway as well as metabolites, which were most likely altered by the endogenous enzymes of the host. The biological activity of the most interesting new products was tested against human topoisomerases I and II, and they were found to exhibit such activities.

The heterologous expression system facilitated the generation of mutants with inactivated biosynthetic genes. In that process, we were able to identify the functions of the glycosyltransferases SnogE and SnogD, solve the structure of SnogD, discover a novel C1-hydroxylase system comprising SnoaW and SnoaL2, and establish that the two homologous non-heme  $\alpha$ -ketoglutarate and  $\text{Fe}^{2+}$  dependent enzymes SnoK and SnoN catalyze atypical reactions on the pathway. We demonstrated that SnoK was responsible for the formation of the additional C–C bond, whereas SnoN is an epimerase. A combination of *in vivo* and *in vitro* techniques was utilized to unravel the details of these enzymes. Protein crystallography gave us an important means to understand the mechanisms. Furthermore, the solved structures serve as platforms for future rational design of the enzymes.

## TIIVISTELMÄ

Maaperässä esiintyvät *Streptomyces*-suvun bakteerit tunnetaan niiden kyvystä tuottaa biologisesti aktiivisia yhdisteitä. Nogalamysiini on potentiaalinen syöpälääke, joka on solumyrkyllinen ja aktiivinen ihmisen topoisomeraasi I:tä ja II:ta vastaan. *S. nogalater*-kanta tuottaa sitä. Nogalamysiini on antrasykliinipolyketidi, joka koostuu nelirenkaisesta aromaattisesta rungosta ja siihen C7-asemaan liittyneestä neutraalista sokerista sekä aminosokerista, joka on liittynyt sekä O–C-sidoksella asemaan C1 että C–C-sidoksella C2- ja C5''-aseman välillä. Yhdisteen rakenne on hyvin mielenkiintoinen, sillä aminosokeri on kiinnittynyt hyvin epätyypillisellä tavalla ja on omalta osaltaan vastuussa nogalamysiinin biologisesta aktiivisuudesta, sillä se auttaa molekyylin sitoutumisessa DNA:han. Myös yleisellä tasolla sokeriosat ovat usein vastuussa polyketidien biologisesta aktiivisuudesta. Yhdessä rakenteen epätavallisuus sekä sokeriosien tärkeys ovat tehneet nogalamysiinin tutkimisesta olennaista.

Sokeriosien kiinnittämisestä vastaa tyypillisesti sokerinsiirtäjäentsyymi glykosyyli-transferaasi. Tämä entsyymi käyttää substraatteinaan kahta molekyyliä, aktivoitua sokeria ja vastaanottajamolekyyliä, johon sokeri liitetään. Työni kirjallisuuskatsaus keskittyy sokerien liittämiseen polyketideihin sekä mahdollisuuksiin muuttaa näitä sokeroimitalleja.

Omat tutkimustulokseni liittyvät nogalamysiinin biosynteesiin, jonka yksittäisiä askeleita olemme tutkineet selvittääksemme, mistä yhdisteen lähes ainutlaatuinen rakenne johtuu. Siirsimme nogalamysiinin biosynteesistä vastaavan geeniklusterin heterologiseen isäntään, *S. albukseen*, kosmidin ja plasmidin avulla. Kun siirsimme isäntäkantaan ainoastaan kosmidin, joka sisälsi suurimman osan klusterista, mutta ei kaikkia geenejä, havaitsimme kannan tuottavan uusia yhdisteitä. Työssä mittasimme uusien yhdisteiden bioaktiivisuuden ihmisen topoisomeraaseja I:tä ja II:ta vastaan, ja niissä huomattiin olevan aktiivisuutta.

Biosynteesiin liittyvien geenien ilmentäminen heterologisessa isännässä mahdollisti inaktivointimutanttien teon. Työn aikana pystyimme selvittämään kahden sokerinsiirtäjäentsyymin, SnogE:n ja SnogD:n, toimintaa sekä SnogD:n rakenteen. Lisäksi löysimme uudenlaisen C1-aseman hydroksylaatiosysteemin, joka koostuu SnoaW:stä ja SnoaL2:sta, sekä kaksi homologista hemitöntä  $\alpha$ -ketoglutaraatti- ja rauta<sup>2+</sup>-riippuvaista entsyymiä: SnoK:n ja SnoN:n. Osoitimme, että SnoK on vastuussa epätyypillisen hiili–hiilidoksen muodostumisesta, kun taas SnoN on epimeraasi. Entsyymien toiminnan yksityiskohtien selvittämiseksi käytimme sekä *in vivo*- että *in vitro* -tekniikoita. Proteiinkristallografian avulla selvitimme kolmen entsyymin kolmiulotteiset rakenteet, joiden avulla pystyimme ymmärtämään entsyymien reaktiomekanismeja paremmin. Lisäksi rakenteet helpottavat proteiinien muokkausta.



## **ABBREVIATIONS**

1/2/3D	One/two/three-dimensional
AAT-fold	Aspartate aminotransferase fold
ABTS	2, 2'-azino-bis (3-ethylbenzthiazoline-6-sulfonic acid)
ATP	Adenosine triphosphate
CO <sub>2</sub>	Carbon dioxide
CoA	Coenzyme A
COSY	Correlation spectroscopy
dTDP	Thymidine diphosphate
dUDP	2'-deoxyuridine-5'-diphosphate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionization mass spectrometry
Fe <sup>2+</sup> / Fe(II)	Iron (2+)
GFOR	Glucose-fructose oxidoreductase fold
GT	Glycosyltransferase
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HRP	Horseradish peroxidase
HSQC	Heteronuclear single-quantum correlation spectroscopy
HSQCDE	Heteronuclear single-quantum and distortionless enhancement correlation
NAD <sup>+</sup>	Nicotinamide adenine dinucleotide, oxidized

## *Abbreviations*

---

NADH	Nicotinamide adenine dinucleotide, reduced
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced
NDP	Nucleoside diphosphate
NMP	Nucleoside monophosphate
NTP	Nucleoside triphosphate
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
OD <sub>600</sub>	Optical density at 600 nm
PBS	Phosphate-buffered saline
PKS	Polyketide synthase
PLP	Pyridoxal phosphate
SAM	<i>S</i> -adenosylmethionine
SDR	Short chain alcohol dehydrogenase
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
S <sub>N</sub> 1	Unimolecular nucleophilic substitution
S <sub>N</sub> 2	Bimolecular nucleophilic substitution
S <sub>N</sub> i	Nucleophilic internal substitution
TLC	Thin layer chromatography
TOCSY	Total correlation spectroscopy
UDP	Uridine diphosphate
UV/Vis	Ultraviolet/visible light
wt	Wild type
α-KG	Alpha-ketoglutarate/alpha-oxoglutarate

## **AMINO ACID ABBREVIATIONS**

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
X	Any amino acid	

## 1. INTRODUCTION

Natural products are an interesting research topic because of their broad diversity both in their chemical structures as well as in their biological functions. Even with the advances made in modern organic chemistry, natural products cannot be replaced. The complex structures of natural products continue to inspire synthetic and analytical chemists. Natural products are an important source for human medicine, they work as potential biological probes, and there are still many to be discovered (Walsh and Fischbach, 2010). Significant producers of natural products are the soil-dwelling Gram-positive *Streptomyces* bacteria (Kieser *et al.*, 2000). In addition to the compounds found in nature, substantial effort is being made to alter the chemical structures of natural products, so that they would work better as drugs.

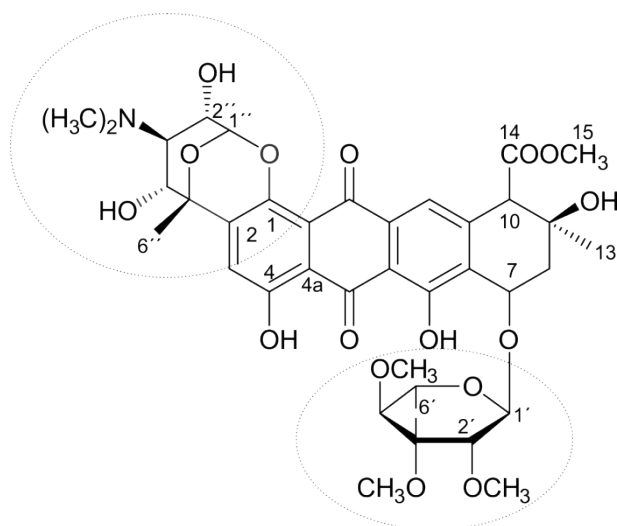
In the literature review, the focus is on polyketides, which are synthesized by polyketide synthases (PKSs). Three different types of PKSs are known, and they can be further divided into subtypes. They act either as modules or iteratively. A commonality with all the systems is their resemblance with fatty acid synthesis, where simple precursors are processed in successive decarboxylative condensation reactions (Khosla *et al.*, 1999).

Polyketides are often decorated with sugar moieties, which are attached by glycosyltransferases (GT) that use two substrates: the acceptor substrate, aglycone (derived from Greek: *a*: without and *glycone*: sweet), and the donor substrate, which donates the sugar moiety to be attached. There is a huge variety of aglyca and sugars in nature (Elshahawi *et al.*, 2015).

In many cases, carbohydrate moieties play a critical role in the biological activities of the compounds (Weymouth-Wilson, 1997). They facilitate the attachment of the compound to its target and may affect the stability as well as change the solubility properties of the compounds, which is an important factor in drug design (De Bryun *et al.*, 2015). Additional carbohydrates have even been used in synthetic drugs to increase solubility (Williams *et al.*, 2013). The carbohydrate moieties are typically attached via an O-, C-, N-, or S-glycosidic bond to the backbone, the most prevalent being the O-glycosylation. The sugar moieties go through various modifications, such as dehydration, isomerization, and group transfers. Most modifications happen prior to the attachment, but modifications may also occur after the glycosylation event (Thibodeaux *et al.*, 2008). One facile way to diversify natural products is to focus on altering the sugar moieties, their amount and/or their properties. This is because the biological activity of compounds often relies on them, and because the sugar-modifying enzymes from different pathways share structural and functional aspects, which may facilitate the combination of enzymes (Thibodeaux *et al.*, 2007).

The focus of my experimental work is on nogalamycin (Fig. 1), an aromatic polyketide belonging to the anthracyclines made by type II PKS. In these metabolites, the aglycone onto which the sugar moieties are attached is a tetracyclic polyphenolic anthraquinone (Metsä-Ketelä *et al.*, 2008). One of the sugar moieties of nogalamycin is attached in a very atypical manner, as the amino sugar is not only attached by the canonical O-glycosylation between the C1'' and C1 positions but also via a C–C bond between C5'' and C2. Although the substructure is not entirely unique in nature (Kawai *et al.*, 1987; Searle *et al.*, 1991; 2003; Ubukata *et al.*, 1991; 1993a–b), it is the only one studied in detail with regard to its biosynthesis (Beinker *et al.*, 2006; Grocholski *et al.*, 2010; Kantola *et al.*, 2000; Shao *et al.*, 2015; Siitonen *et al.*, 2012a; 2012b; 2016; Sultana *et al.*, 2004; Torkkell *et al.*, 1997; 2001; Ylihonko *et al.*, 1996)

In this work, the late stage tailoring steps in nogalamycin biosynthesis were elucidated in detail and now it is understood where the special features of this compound originate. Firstly, prior to the attachment of the amino sugar, the aglycone is hydroxylated at position C1 by a novel two-component cofactorless monooxygenase system *SnoaW/SnoaL2* (Siitonen *et al.*, 2012a). Secondly, we identified and solved the crystal structure of the *SnogD* GT responsible for the transfer of the amino sugar, whereas the *SnogE* GT was shown to catalyze the transfer of the neutral sugar (Claesson *et al.*, 2012; Siitonen *et al.*, 2012b). Thirdly, we demonstrated that the C5''–C2 bond is formed by *SnoK* and that an unexpected epimerization of the C4'' position of the amino sugar occurs through the action of *SnoN* (Siitonen *et al.*, 2016).



**Figure 1.** Nogalamycin with the studied positions highlighted in purple

## 2. LITERATURE REVIEW

Natural products containing carbohydrate moieties are chemically diverse, not only because there are hundreds of different sugar moieties, but also because the aglyca onto which the sugar moieties are attached vary significantly. Furthermore, sugar moieties are found in most of the various classes of natural products (Elshahawi *et al.*, 2015). The sugars are derived from the common carbohydrates of primary metabolism, but typically undergo a range of unusual modifications. Carbohydrate moieties attached to natural products are in most cases associated with the enhancement or alteration of biological potency. Furthermore, they act as solubility modulators and aid in penetration through membranes (Weymouth-Wilson, 1997). Sugar moieties can also be associated with self-resistance against the produced compounds (Cundliffe and Demain, 2010).

### 2.1. Glycosyltransferases

Glycosyltransferases (GT) are found in each domain of life. These diverse enzymes are important in attaching carbohydrates to biomolecules. GTs use two substrates: the donor substrate, which donates the sugar moiety to be attached and the acceptor substrate, which is decorated by the carbohydrate moiety attached in the process. The reaction happens in a regio- and stereoselective manner (Liang *et al.*, 2015). When the reaction is completed a glycosidic linkage is formed between the sugar and the backbone. The carbohydrate moiety is activated by a leaving group to facilitate the attachment. In the case of Leloir GTs – named after the Nobel laureate Luis Federico Leloir – the activation is achieved by a nucleoside diphosphate nucleotide or a nucleoside monophosphate, whereas with non-Leloir GTs, lipid phosphates or unsubstituted phosphates are utilized (Lairson *et al.*, 2008).

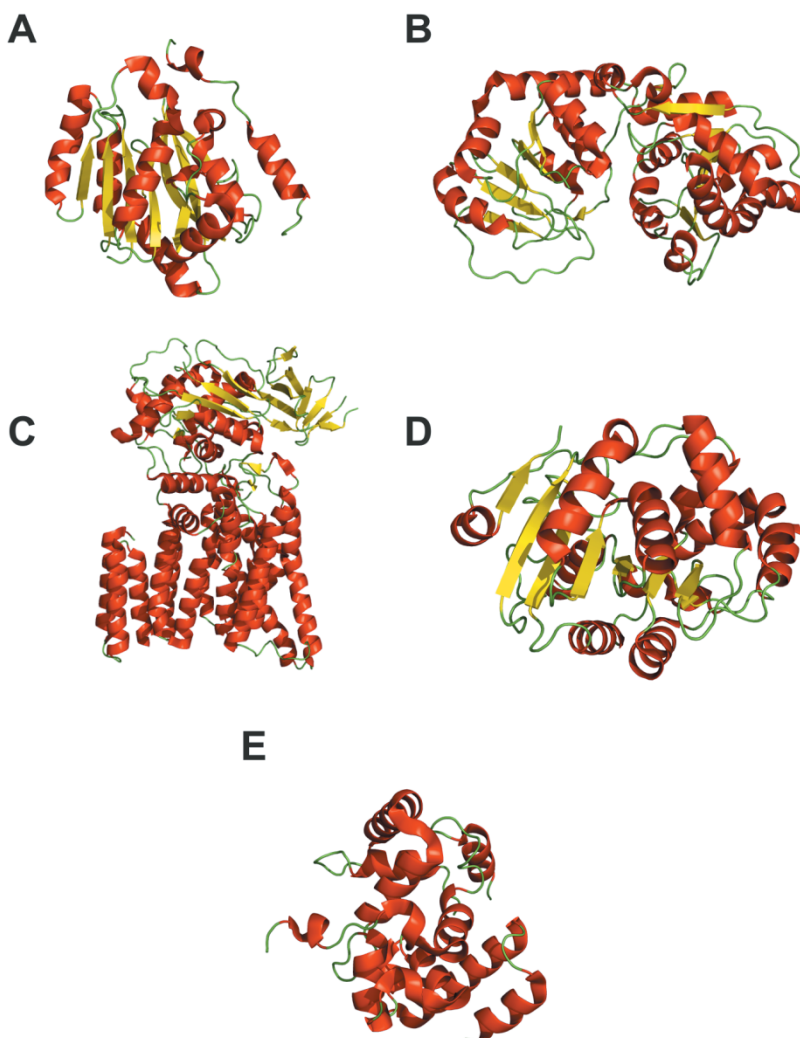
There are multiple classification schemes for dividing GTs into different groups. The traditional division is the Leloir and non-Leloir division as mentioned above. Division based on the fold is the most common, with further subdivision into clans based on the utilized mechanism (I and II: inverting, III and IV: retaining) and finally to families based on the sequences (Coutinho *et al.*, 2003; Liu and Mushgenian, 2003). GTs can also be divided into processive and non-processive GTs, which denotes whether or not the acceptor substrate is released between additional reaction cycles that attach subsequent sugars (Price *et al.*, 2002). In total, there are nearly 100 families of GTs. The sequence similarity between GTs is generally low, except within a specific GT family. The CAZy-database of GTs contains current information on GTs and other enzymes involved in carbohydrate metabolism, such as glycoside hydrolases, polysaccharide lyases, and carbohydrate esterases (Lombard *et al.*, 2014; [www.cazy.org](http://www.cazy.org)). Correspondingly, the 3D structures of GTs that have been solved are gathered into another database (Pérez *et al.*, 2015; [glyco3d.cermav.cnrs.fr](http://glyco3d.cermav.cnrs.fr)).

The accuracy and substrate specificity varies greatly. In many cases, GTs have been shown to exhibit flexible substrate specificity, especially regarding the donor sugar but also the acceptor aglycone (e.g. Borisova *et al.*, 2006; 2008; Minami and Eguchi, 2007; Minami *et al.*, 2005; Trefzer *et al.*, 2002; Yang, M. *et al.*, 2005). Nevertheless, there are GTs that are not shown to be promiscuous or that show very limited promiscuity under the tested conditions (e.g. Albermann *et al.*, 2003; Freel Meyers *et al.*, 2003; Oberthür *et al.*, 2005).

### 2.1.1. Glycosyltransferase Folds

The amino acid sequence similarity between GTs is generally low. However, protein crystallography has revealed that there is limited variation in 3D structures, and there are only two main folds, denoted as GT-A and GT-B (Bourne and Henrissat, 2001; Coutinho *et al.*, 2003). The GT-A fold (Fig. 2A) has been described as a single-domain enzyme (Chang *et al.*, 2011b; Thibodeaux *et al.*, 2008), although others have noted that it is formed by two closely adjoined domains that resemble  $\beta/\alpha/\beta$  Rossmann folds (Lairson *et al.*, 2008; Ünlügil and Rini, 2000). The catalytic site is situated between two distinct binding sites for the donor and acceptor molecules and there is a conformational change upon binding (Liang *et al.*, 2015). Enzymes containing the GT-A fold generally rely on a divalent metal ion, such as  $Mg^{2+}$  or  $Mn^{2+}$ , and consequently, a metal binding motive – DXD or EXD – is found in most sequences. The metal is proposed to stabilize the charged phosphate group of the donor substrate (Fig. 6–7; Charnock and Davies, 1999; Ünlügil and Rini, 2000). However, there are examples of GT-A proteins where the metal binding motive is not found (Pak *et al.*, 2011; Qasba *et al.*, 2005), and in these cases the stabilization of the transition state is most likely achieved in a manner similar to the metal-independent GT-Bs described below. Some GT-A enzymes have the globular catalytic domain linked to a transmembrane region and a small cytoplasmic domain in the N-terminus (Breton and Imberty, 1999; Morgan *et al.*, 2013).

Proteins of the GT-B family (Fig. 2B) are composed of two domains with Rossmann-type folds. This fold clearly consists of two distinct domains: the N-terminal acceptor (aglycone) binding and the C-terminal donor (NDP-sugar) binding domain. The domains are facing each other and are joined together by a flexible linker. The formation of the glycosidic bond takes place between these domains (Moncrieffe *et al.*, 2012; Mulichak *et al.*, 2003; Qasba *et al.*, 2005). Structural and biochemical analysis of GT-Bs has shown that the binding of the substrate induces the protein loops to change their conformation, shifting the relative locations of the Rossmann-type domains. This indicates a conformational selection or induced fit; thus the protein has an open and a closed conformation (Chang *et al.*, 2011a; Hu *et al.*, 2003; Mulichak *et al.*, 2003; Quirós *et al.*, 2000; Tam *et al.*, 2015). Typically, the enzymes of this fold



**Figure 2.** Examples of the different glycosyltransferase folds. A) GT-A fold, the example is SpsA, the GT involved in the formation of the spore coat of *Bacillus subtilis* (PDB: 1H7L, Tarbouriech *et al.*, 2001), B) GT-B fold, the example is SnogD, the GT involved in the nogalamycin biosynthesis (PDB: 4AMG, Claesson *et al.*, 2012), C) GT-C fold, the example is PglB, the GT that attaches oligosaccharyl moieties in *Campylobacter lari* (PDB: 3RCE, Lizak *et al.*, 2011), D) GT-D fold found in *Streptococci parasanguinis* in the domain that glycosylates *Streptococcal* adhesins (PDB: 4PFX, Zhang, H. *et al.*, 2014), E) lysozyme-like fold, the example is the peptidoglycan glycosyltransferase domain involved in peptidoglycan synthesis from *Aquifex aeolicus* PBP1A (PDB: 2OQO, Yuan *et al.*, 2007).



catalyze reactions in the absence of a metal-ion (Lairson *et al.*, 2008; Gloster, 2014), which is compensated by positively charged amino acids or a helix dipole (Fig. 6–7; Breton *et al.*, 2012). The fact that GT-A and GT-B folds both contain Rossmann-type domains indicates that they belong to the Leloir GTs.

In addition to these most prevalent folds, there are rarer folds emerging as more structural studies are progressing. The GT-C fold (Fig. 2C) is found in membrane-bound enzymes and thus contains transmembrane helices spanning the membrane (Igura *et al.*, 2008; Lizak *et al.*, 2011; Maita *et al.*, 2010). Many of the structures are limited to the truncated soluble C-terminal domains (Igura *et al.*, 2008; Maita *et al.*, 2010) due to difficulties in unraveling of 3D structures of membrane proteins, and only one structure of an entire protein is available to date (Lizak *et al.*, 2011). The lack of a Rossmann-type fold together with the fact that the protein resides on a membrane suggest that the sugars transferred by this type of GT are all activated by lipid phosphates (Igura *et al.*, 2008; Maita *et al.*, 2010). Similarly to the GT-A fold, the GT-C fold utilizes a divalent metal, but it appears to have a dual role, not only stabilizing the leaving group, but also orienting acidic residues to interact with the acceptor molecule (Lizak *et al.*, 2011).

The GT-D fold (Fig. 2D) is found in pathogenic *Streptococci* and fusobacteria (Zhang, H. *et al.*, 2014). It is related to the bacterial adhesin synthesis, and thus the GTs exhibiting this fold may play a role in the pathogenicity of these bacteria and hence be considered targets for drugs (Zhou and Wu, 2009). The fold contains three regions, a central Rossmann-type fold, flanking the C-terminal region and an N-terminal metal-binding motive, DXE, which is distinct from the GT-A metal-binding motive and cannot be replaced by another metal-binding motive (Zhang, H. *et al.*, 2014).

The lysozyme-like fold (Fig. 2E) has thus far been found in enzymes related to peptidoglycan synthesis (Lovering *et al.*, 2007; Yuan *et al.*, 2007). The fold resembles the fold found in  $\lambda$ -lysozyme, but instead of hydrolyzing the carbohydrates of the peptidoglycan surface, it attaches them and forms the carbohydrate chains on the bacterial cell wall to form peptidoglycan (Lovering *et al.*, 2007; Yuan *et al.*, 2007).

Based on sequence prediction, there are GTs that do not seem to belong to any of these folds, for example the *in silico* proposed GTS-D superfamily fold (Kikuchi *et al.*, 2003), but upon experimental data the conclusions may still change. All known GTs involved in polyketide biosynthesis - and most GTs in natural product biosynthesis in general - are Leloir GTs and possess the GT-B fold. A recently found exception is Ram29, which belongs to the GT-C fold. It is a polyprenyl phosphomannose dependent GT, which is involved in the biosynthesis of enduracidin, a polypeptide antibiotic, but the structure of the GT has not been elucidated (Wu, M.-C. *et al.*, 2015).

### 2.1.2. Sugar Linkage

The sugar moieties can be attached by different linkages to the aglycone moiety. The most common mode of attachment is the O-glycosidic bond in which the oxygen is derived from the sugar moiety. The rarer linkages are the C-glycosidic, N-glycosidic, S-glycosidic, hydroxylamine or orthoester bonds (Thibodeaux *et al.*, 2008). The type of bond is expected to affect stability; for example the O–C bond is the most prone to spontaneous hydrolysis, and the C–C bond is the most stable. For these reasons, the type of linkage is important for the stability of the final compound and also influences applicability in medicine (Bililign *et al.*, 2005). For instance, some O-glycosylated compounds are prone to deglycosylation in the acidic environment of the stomach, which prevents oral administration of the drugs (Hultin, 2005).

A specific GT typically makes just one kind of bond. Nevertheless, there are examples of natural or engineered GTs exhibiting a bifunctional nature in terms of the linkage. These examples are found in plants and yeast making both O- and N-linkages (Loutre *et al.*, 2003; Xu *et al.*, 2013), or O/C-glycosylating enzymes found in e.g. bacteria and plants (Dürr *et al.*, 2004; Gandia-Herrero *et al.*, 2008) and even trifunctional (O-/S-/N-) GT from bacteria (Gantt *et al.*, 2008). This feature makes it more difficult to understand what causes a specific linkage to be formed. Nevertheless, studies have given insight into the preference of linkage formation. The Bechthold group altered a strictly C–O bond forming GT to a strictly C–C bond forming GT (Härle *et al.*, 2011). The O–C forming GT, UrdGT2, from the urdamycin pathway was altered by rational design to mimic the C–C forming GT, LanGT2, from the landomycin pathway (Härle *et al.*, 2011). Further structural studies led to a model where the cause of the different linkage forming capacity lies in the conformation of the substrate in relation to the protein and, more specifically, to the nucleophile needed for catalysis (Tam *et al.*, 2015).

### 2.1.3. Substrates for the Glycosyltransferase

There are several different aglyca that act as acceptor substrates in GT reactions. A comprehensive presentation of aglyca including a wide range of different polyketides and other aglyca such as nucleosides and peptides as well as sugars is demonstrated in the detailed review by Elshahawi *et al.*, 2015.

#### 2.1.3.1. The Acceptor Substrate

Polyketide aglyca are built by PKS from simple compounds bound to coenzyme A, which are called the starter and the extender units (Fischbach and Walsh, 2006). In the case of type II PKS, the starter and extender units can be for example acetyl-CoA and malonyl-CoA, respectively, as in the example shown in Figure 3A. Through multiple

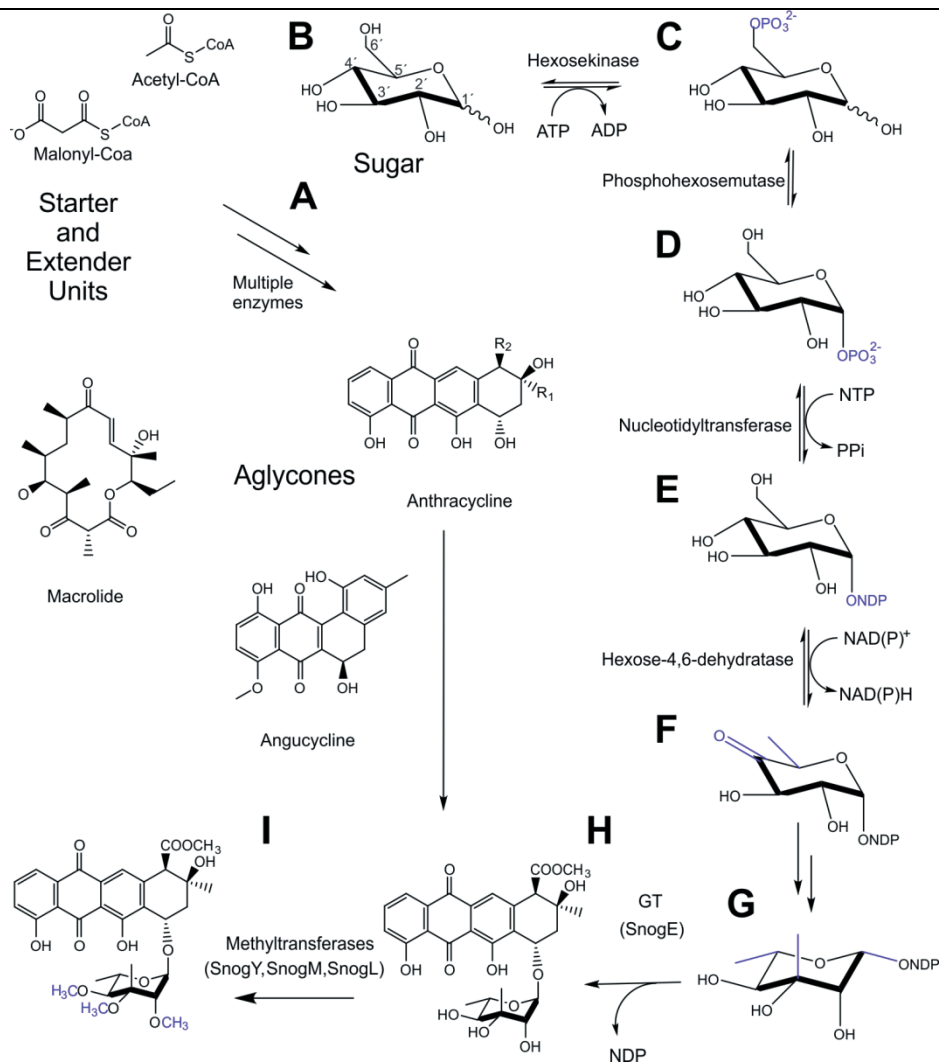
condensation reactions, a carbon chain is built by the minimal PKS, which consist of two ketosynthase units and an acyl carrier protein. Further modifications to the carbon chain, such as cyclizations, are made depending on the final structure (Hertweck *et al.*, 2007; Khosla *et al.*, 1999). As an example of possible aglycone structures, only a few possible aglyca are shown in Figure 3 (Elshahawi *et al.*, 2015).

### 2.1.3.2. *The Donor Substrate: the Activation of the Sugar*

The second substrate for the GT is the donor substrate, which is the attached sugar. The precursor for it is a monosaccharide-1-phosphate derived from primary metabolism. One well-studied route is the Leloir pathway, which converts  $\beta$ -D-galactose via multiple steps to D-glucose-1-phosphate (Frey, 1996; Holden *et al.*, 2003). Glycogen, fructose or glucose can also be utilized as the precursor for the biosynthesis of the attached sugar. In the example shown in Figure 3, the biosynthesis starts from glucose, which is phosphorylated by a hexose kinase to glucose-6-phosphate and converted to glucose-1-phosphate by a phosphohexomutase (Fig. 3B–D; Lu and Kleckner, 1994).

After the phosphate group has been transferred to the anomeric position, the sugar has to be activated with a leaving group in order for it to be a suitable substrate for a GT (Lairson *et al.*, 2008). In the case of Leloir GTs, the activation is achieved by a nucleoside diphosphate nucleotide or a nucleoside monophosphate, whereas with non-Leloir GTs, lipid phosphates or unsubstituted phosphates are utilized (Lairson *et al.*, 2008). Typically, GTs involved in the glycosylation of polyketides use activated diphosphate nucleotides as the leaving group of the donor substrate. Consequently, the nucleotide activated sugars are the focus of this thesis. The typical precursor, glucose-6-phosphate, is the sugar moiety for dTDP-, CDP- and UDP-activated sugars (Fig. 3E; Singh *et al.*, 2012). In natural product biosynthesis, the most prevalent nucleotide used by bacterial GTs is dTDP (Thibodeaux *et al.*, 2008). All bacterial GTs with available crystal structures that are involved in polyketide biosynthesis use either dTDP or UDP as the leaving group. Furthermore, the structural data on GTs suggest that, in the case of dTDP-utilizing enzymes, the preference of dTDP over UDP is rather stringent, because the nucleotide-binding pocket has a steric obstruction of the 2'-OH of the ribose, which hinders the binding of UDP (Isiorho *et al.*, 2014; Wang *et al.*, 2013).

The NDP-hexose nucleotidyltransferase, which can also be called NDP-hexose synthetase or NDP-hexose pyrophosphorylase, is responsible for the attachment of an NTP to the sugar moiety (Fig. 3E). The reaction is proposed to happen in a Bi-Bi ordered manner, where first the NTP molecule binds to the enzyme and subsequently a phosphate or a pyrophosphate is removed. Finally, in a typical scenario, an  $\alpha$ -sugar is linked to the corresponding NDP or NMP (Barton *et al.*, 2001; Blankenfeldt *et al.*, 2000; Zuccotti *et al.*, 2001). Similarly to the GTs, the nucleotidyltransferase fold contains distinct binding sites for the two substrates: the sugar and the nucleotide.



**Figure 3.** A schematic overview of the sugar modifications and attachment to the aglycone. A) The aglycone is synthesized. B) The sugar is derived from primary metabolism, in this example D-glucose, C) the sugar is phosphorylated by a hexose kinase, resulting in glucose-6-phosphate D) and the phosphate group position is changed by a phosphohexomutase, resulting in glucose-1-phosphate E) and activated by a nucleotidyltransferase, F) altered by a hexose-4,6-dehydratase G) and modified further. H) The activated sugar is attached to the aglycone I) and finally possible post-glycosylation events may take place. In the picture, as an example, is the deoxy sugar of nogalamycin; L-nogalose, which is attached to the aglycone by SnogE, and the methylation reactions are achieved by the actions of SnogY, SnogM and SnogL, which are proposed to happen after the attachment to the aglycone (Singh *et al.*, 2012, Thibodeaux *et al.*, 2008; Torkkell *et al.*, 2001; Siitonen *et al.*, 2012b).

As in the GT-A fold of GTs, a divalent metal stabilizes the leaving group. The activation of sugars by the attachment of a nucleotide is also part of primary metabolism, where it is utilized, for example, in cell wall and glycogen synthesis (Holden *et al.*, 2003; Zuccotti *et al.*, 2001). The congruence of the primary and secondary metabolism is highlighted by the fact that there are cases where the nucleotidyltransferases associated with secondary metabolism may not be included in the gene clusters of the corresponding polyketides but are in fact the same genes that are expressed for primary metabolism elsewhere in the genome (Thibodeaux *et al.*, 2008). However, this may also be pathway-dependent, since there is also evidence for a nucleotidyltransferase that is specific for secondary metabolism (Kudo *et al.*, 2005).

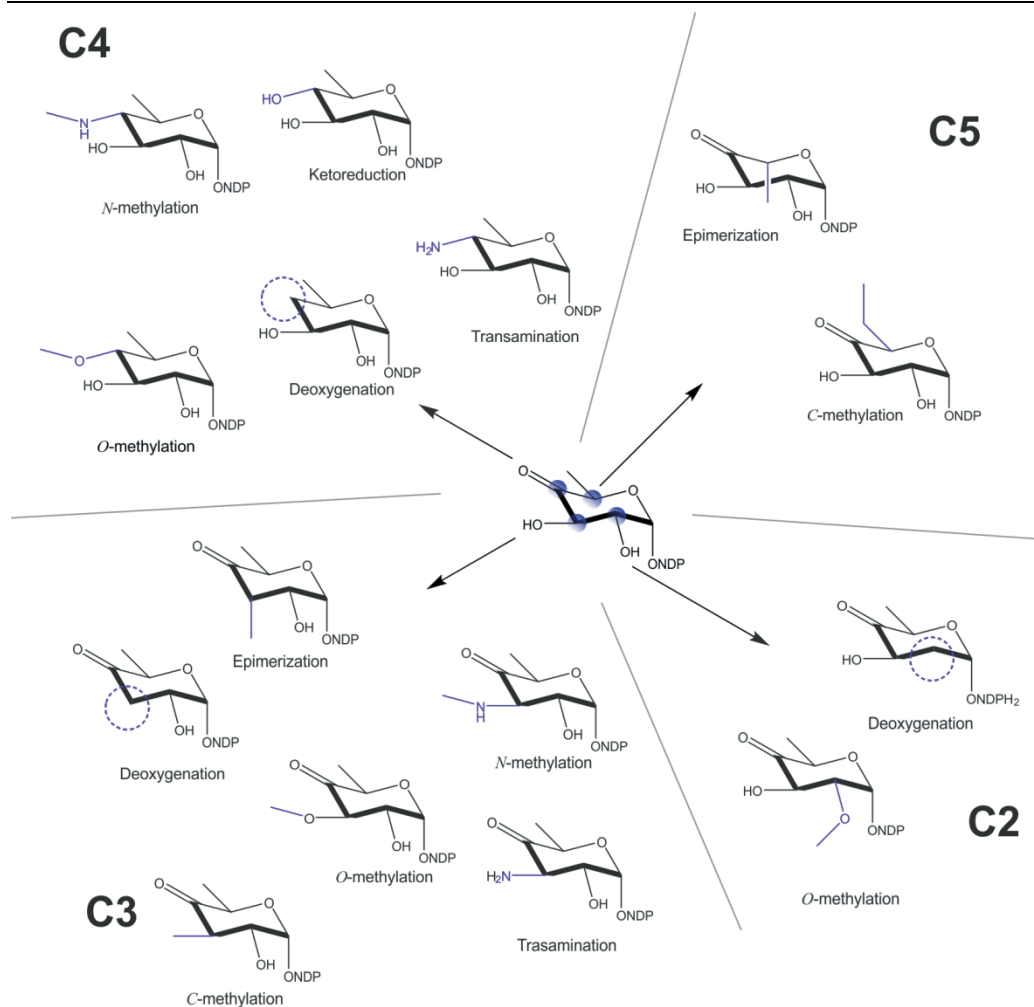
### **2.1.3.3. The Donor Substrate: the Modification of the Sugar**

After the activation of the sugar, the next shared step in the biosynthesis of unusual sugars is the dehydration of the sugar (Fig. 3F). The SDR fold containing NDP-hexose-4, 6-dehydratase creates an NDP-4-keto-6-deoxy sugar in a stepwise manner. The protein is NAD(P)<sup>+</sup>-dependent and catalyzes C4' oxidation, C5'/C6' dehydration and C5'/C6' ene reduction of the sugar (Allard *et al.*, 2002; Singh *et al.*, 2012). This reaction is considered irreversible and is the final common step for the sugar moieties involved in the biosynthesis of sugars attached to polyketides (He and Liu, 2002; Lombó *et al.*, 2009). As in the case of nucleotidyltransferases, the gene for the NDP-hexose-4, 6-dehydratase has also been speculated in some cases to reside outside the corresponding gene cluster, for example in the case of the sugar moiety L-daunosamine found in daunorubicin and doxorubin (Trefzer *et al.*, 1999).

Most of the modifications to carbohydrates happen to the nucleotide bound sugars before the attachment to the aglycone (Field and Naismith, 2003; Thibodeaux *et al.*, 2008). These include epimerizations, isomerizations, hydroxylations, transaminations, deoxygenations, oxidations, amino oxidations, pyranose/furanose interconversions, dehydrations, reductions, and group transfers such as methylations (Singh *et al.*, 2012).

The further modifications made to the sugars after the attachment of the NDP – some of which are briefly discussed in the following sections – are not universal to all unusual sugars (Fig. 4; Salas and Méndez, 2007). More information on the sugar-modifying enzymes can be found in the detailed reviews (e.g. Field and Naismith, 2003; He *et al.*, 2000; He and Liu, 2002; Singh *et al.*, 2012; Thibodeaux *et al.*, 2008).

One important and prevalent modification made to the sugars is dehydration at different positions of the sugar. Dehydrations are needed to achieve a reactive ketosugar, which is required for other modifications, such as the addition of an amino group. The dehydratases belong to two different classes: the SDR-fold dehydratases (Allard *et al.*, 2002) and the AAT-fold dehydratases, which resemble



**Figure 4.** Examples of possible modifications made to the common intermediate (from Fig. 3E) of the sugar moiety biosynthesis. Modifications are made to different positions (C2–C5) marked with purple spheres and divided into different corners of the figure. The figure was inspired by Salas and Méndez, 2007.

aminotransferases (Smith *et al.*, 2008). Amino group containing sugars are generated by aminotransferases, which attach an amino group from an amino donor (L-Glu, L-Gln or L-Asp) to an NDP-ketosugar by utilizing PLP as the cofactor. The stereochemistry of the attached amino group may differ depending on the pathway. These differences are suggested to result from different orientations of the sugar in the active site of the enzymes (Burgie and Holden, 2007).

The sugar epimerases or isomerases invert the stereochemistry of an asymmetric carbon of the sugar. There are several types of enzymes catalyzing epimerization reactions, and they use different strategies (Allard *et al.*, 2001). Cupin-type epimerases remove a proton from one side of the sugar, creating an unstable enolate, which is subsequently protonated from the other side of the sugar (Dong *et al.*, 2007; Dunwell *et al.*, 2001; Giraud *et al.*, 2000). The enzymes of the second epimerase family have an SDR-fold and require NAD(P)<sup>+</sup> as cofactors. In the proposed reaction mechanism, a hydride is removed from one side of the sugar and introduced to the other side, thus changing its stereochemistry (Frey and Hegeman, 2013; Lau and Tanner, 2008; Major *et al.*, 2005). A different approach is used by the GT-B type epimerase, which eliminates the attached nucleotide and reintroduces it to the molecule from the opposite side, resulting in an altered stereochemistry (Campbell *et al.*, 2000). The newest example of a sugar epimerase is the  $\alpha$ -KG-dependent epimerase, which we found to be involved in nogalamycin biosynthesis, where it creates the epimer of the amino sugar at position C4'' [Fig. 11: **10**, **12**, 16]. The mechanism differs from the other known sugar epimerases, as it is proposed to create a radical intermediate. Moreover, the reaction takes place only after the attachment of the carbohydrate (Siitonen *et al.*, 2016). A similar enzyme was first discovered in carbapenem biosynthesis, but there the epimerization reaction does not take place on a sugar molecule (Chang *et al.*, 2014).

Sugar ketoreductases reduce the reactive 3- or 4-ketosugar intermediates to the corresponding secondary alcohols with either stereochemistry (Thibodeaux *et al.*, 2008). The reduction results in deactivation of the carbohydrate and takes place after other chemical transformations, which have required unstable intermediates with carbonyl functional groups. Probably because of this, the ketoreductases are widespread among the biosynthetic pathways of sugars. The enzymes are found to have two folds: the SDR (Blankenfeldt *et al.*, 2002) and the glucose-fructose oxidoreductase fold (GFOR) (Kubiak and Holden, 2011). Both of these protein families require NAD(P)H for activity and, in addition, the SDR enzymes are Mg<sup>2+</sup>-dependent. The reduction is achieved by a hydride transfer from NAD(P)H to the carbonyl of the NDP-bound sugar (Blankenfeldt *et al.*, 2002; Kubiak and Holden, 2011).

Methylations are generally achieved through the addition of methyl groups to the sugars through O-, N- or C-bonds. Typically, O-methylations are thought to occur only after the attachment of the sugar moieties, whereas N- and C-methylations happen before the glycosylation event (Thibodeaux *et al.*, 2008). Methylases require SAM, which supplies the methyl group to be transferred. Thus these enzymes contain a binding site for SAM. Furthermore, these enzymes display low sequence similarity to each other but share a highly conserved fold. Most sugar methyltransferases require a divalent metal ion, such as Mg<sup>2+</sup> (Singh *et al.*, 2012). There is an example of a

bifunctional ketoreductase/methyltransferase, MtmC, from the mithramycin pathway, which catalyzes either reaction by using different cofactors (Chen *et al.*, 2015; Wang *et al.*, 2012).

In addition to modifications made to the NDP-bound sugar, examples of post-glycosylation events are known to exist, but they are much less diverse and mostly include group transfers, such as methylation e.g. in steffimycin (Olano *et al.*, 2008) and acetylation for example in chromomycin A<sub>3</sub> (Menéndez *et al.*, 2004) reactions. In Figure 3, L-nogalose from the pathway of nogalamycin serves as an example for modifications made to the sugar prior to and following the attachment. Furthermore, Figure 3 shows the attachment of L-nogalose to the aglycone, nogalamycinone (Fig. 3G–I; Siitonen *et al.*, 2012b; Torkkell *et al.*, 2001).

#### 2.1.4. Mode of Action of the Glycosyltransferases

As previously mentioned (2.1.), GTs use two substrates, the acceptor and the donor, which are attached to each other, while the leaving group, a nucleotide in the Leloir type GTs, detaches from the donor substrate. The reaction usually happens in an ordered Bi-Bi fashion (Liang *et al.*, 2015). There are examples of GTs where the donor substrate binds first (Qasba *et al.*, 2005) as well as of GTs where the acceptor substrate binds first (Quirós *et al.*, 2000). To make the matter more complicated, GtfA, the GT from the vancomycin biosynthesis, appears to enable the binding of both substrates independently without a strict Bi-Bi fashion (Mulichak *et al.*, 2003).

GTs are found to be either inverting or retaining, which corresponds to the stereochemistry of the anomeric carbon of the attached sugar (Fig. 5A–B). Thus if the sugar is attached to the nucleotide in an  $\alpha$ -configuration, after the attachment by a retaining GT the sugar is also an  $\alpha$ -anomer, whereas the action of inverting GT results in a  $\beta$ -anomer (Breton *et al.*, 2012; Lairson *et al.*, 2008; Sinnott, 1990). Both inverting and retaining GTs are found in the two main fold types, GT-A and GT-B, which are further divided into clans based on their mechanism (Coutinho *et al.*, 2003). Because the fold does not give insight into what kind of a mechanism the GT is using, it has proven to be difficult to predict the mechanism based on sequence information alone, but there have been attempts to do this (Rosén *et al.*, 2004).

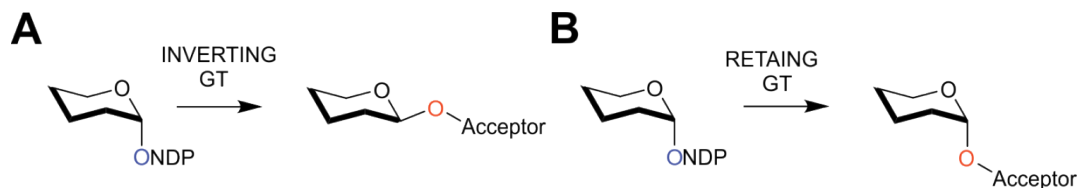
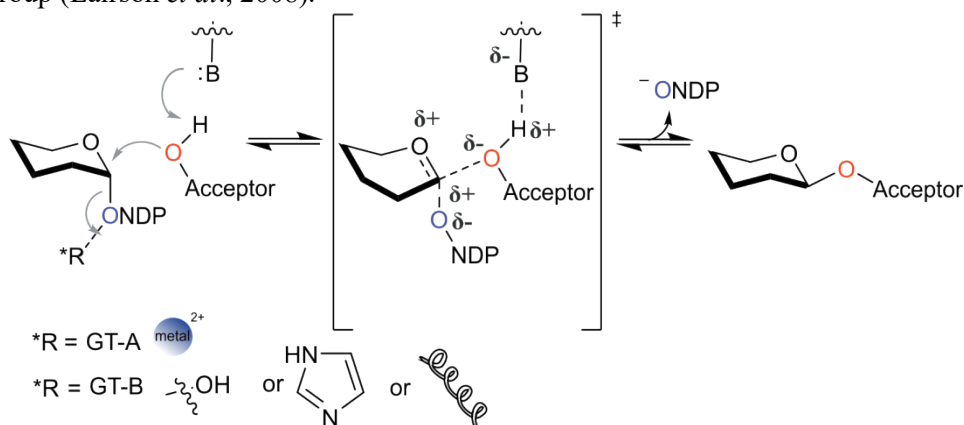


Figure 5. A) inverting GT, B) retaining GT



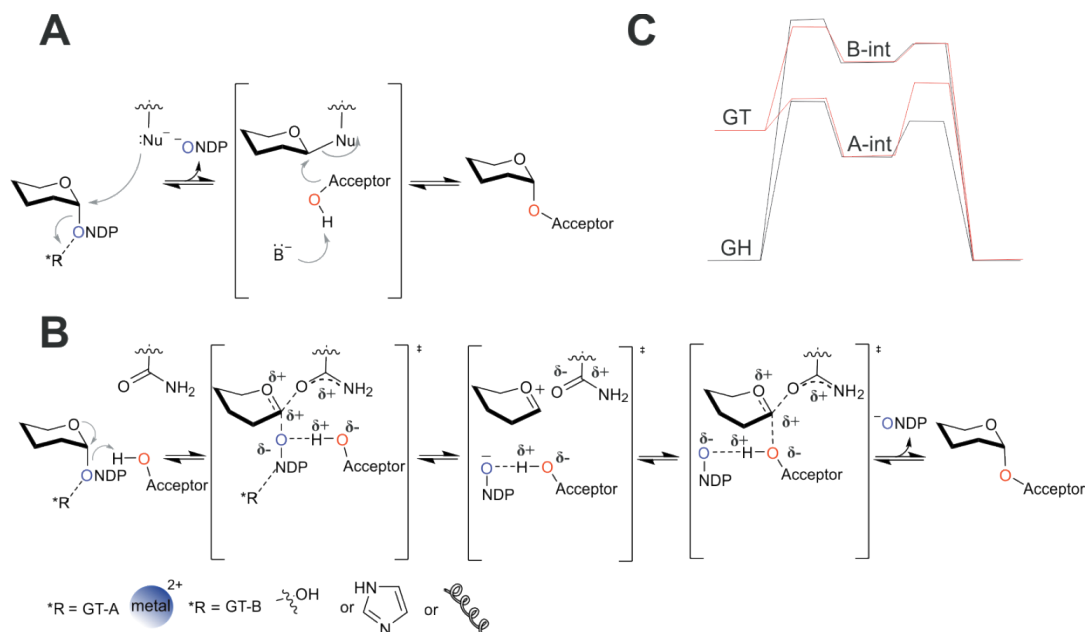
The mechanism of inverting GTs is proposed to follow a single displacement ( $S_N2$ ) mechanism with an oxocarbenium-ion transition state. The proposal is supported by structural data (Breton *et al.*, 2012) and quantum mechanics/molecular mechanics methods (Tvaroška, 2015). Overall, it is thought to follow a similar mechanism to inverting glycoside hydrolases (Lairson *et al.*, 2008). In this mechanism (Fig. 6), the acceptor substrate makes a nucleophilic attack on the anomeric carbon of the donor substrate, which is aided by the abstraction of a proton from the accepting hydroxyl group by a general base, which has been shown to be Asp (Mulichak *et al.*, 2004), Glu (Sun *et al.*, 2007), or His (Isiorha *et al.*, 2014). The departure of the leaving group is facilitated by different means dependent on the fold. The GT-A fold contains a metal binding site, thus a divalent metal assists the phosphate group departure, acting as a Lewis acid. In the GT-B fold, there is no metal binding site, thus the departure has to be facilitated by other means – by a helix dipole, a side chain hydroxyl, or an imidazole group (Lairson *et al.*, 2008).



**Figure 6. Mechanism of the inverting GT according to current knowledge (Lairson *et al.*, 2008).**

There is evidence for two mechanisms used by retaining GTs: the double-displacement mechanism ( $S_N2$ , Fig. 7A) or a special case of the  $S_N1$  mechanism, the internal return mechanism ( $S_{Ni}$ , Fig. 7B). The first proposed mechanism was the double-displacement mechanism, which is analogous to the mechanism utilized by retaining glycosidases (Lairson *et al.*, 2008). It consists of a covalently bound glycosyl-enzyme intermediate (Lairson *et al.*, 2004) and requires a catalytic nucleophile for the reaction. As in the inverting clan, the divalent metal is proposed to assist in the departure of the leaving group, acting as a Lewis acid. However, there is limited experimental evidence for this mechanism at present. Emerging structural data on these clans of GTs have revealed that there are cases in which the structures lack elements needed for the proposed double-displacement mechanism (Flint *et al.*, 2005). This has reinforced the proposal

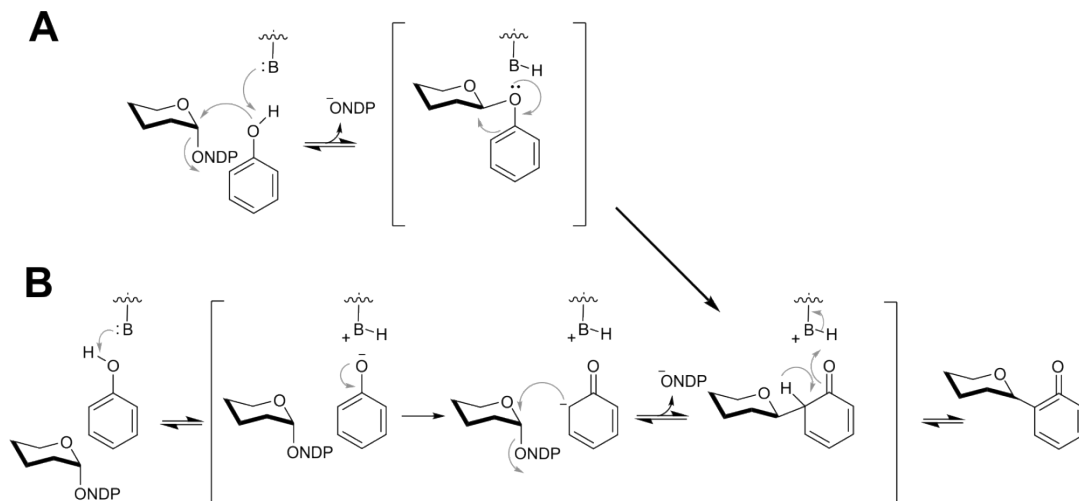
of the alternative  $S_Ni$  mechanism also. Instead of covalent intermediates, the  $S_Ni$  mechanism involves ion pair intermediates, which make it possible to return to the starting conditions without the loss of the compounds by possible hydrolysis. The nucleophilic hydroxyl group of the acceptor attacks the donor, which is followed by the departure of the leaving group to the same side of the attack (Breton *et al.*, 2012).



**Figure 7.** Two possible mechanisms for the retaining GTs according to current knowledge **A)** double-displacement mechanism, with the covalently bound intermediate (A-int in panel C) **B)** internal return mechanism with the ionic intermediates (B-int in panel C) (Breton *et al.*, 2012; Chang *et al.*, 2011b; Lairson *et al.*, 2008). **C)** Schematic presentation of the potential energy barriers for the different intermediates adapted from (Lairson *et al.*, 2008) GT: glycosyltransferase, GH: glycosyl hydrolase or glycosidase, A-int: intermediate from mechanism A and B-int: intermediate from mechanism B.

The intermediate of the alternative  $S_Ni$  mechanism has a higher energy level than the intermediate of the double displacement mechanism (Fig. 7C), but at the same time, the alternative  $S_Ni$  mechanism makes the hydrolysis of the intermediate less likely. Furthermore, in the case of GTs, the substrate carries a high energy leaving group, and thus the energy barrier in both cases ( $S_Ni$  and the double displacement mechanism) is lower than that of the corresponding retaining glycosidases, which do not have a high energy substituent. There is currently more supporting evidence for the  $S_Ni$  mechanism (Ardévol *et al.*, 2016; Gómez *et al.*, 2012; Lairson *et al.*, 2008). However, there may be a continuum between the two mechanisms (Lairson *et al.*, 2008).

The mechanism of the O–C bond formation has been well studied, as it is the most prevalent linkage. However, there have also been investigations into the mechanism regarding the formation of the C–C bond forming GTs (Bililign *et al.*, 2004; 2005; Dürr *et al.*, 2004; Gutmann and Nidetzky, 2012; Härle *et al.*, 2011; Tam *et al.*, 2015). Two mechanisms have been proposed for that the C–C bond formation: the *ortho/para* directed O- to C-glycosyl rearrangement (Fig. 8A) and a direct Friedel–Craft alkylation (Fig. 8B). In the first mechanism mentioned, the phenolic hydroxyl group is O-glycosylated, which is followed by O–C rearrangement. Usually the *ortho*-C-substituents are the major products, whereas *para*-C-substituents are found in minor



**Figure 8.** Possible mechanisms for the C–C bond forming GTs A) O- to C-glycosyl rearrangement B) a direct Friedel–Craft aromatic substitution (Bililign *et al.*, 2004; Thibodeaux *et al.*, 2007; 2008).

amounts. In the second mechanism mentioned, the phenolate anion of the acceptor substrate attacks the anomeric carbon of the NDP-sugar, thus directly forming a C-glycoside (Bililign *et al.*, 2004; Dürr *et al.*, 2004; Thibodeaux *et al.*, 2008). The studies on UrdGT2 and LanGT2 gave evidence for a direct Friedel–Craft alkylation, as mutants with an altered function (O–C bond forming to C–C bond forming enzyme) do not produce C–O linked compounds even in small amounts (Härle *et al.*, 2011).

### 2.1.5. Auxiliary Proteins of Glycosyltransferases

Some GTs have been shown to require an auxiliary protein for activity, e.g. from the pathway of tylosin, indolocarbazole and daunorubicin (Melançon III *et al.*, 2004; Otten *et al.*, 1995; Salas *et al.*, 2005). DesVIII, a protein expressed by *S. venezuelae*, which is the producer of for example pikromycin and methymycin, was the first discovered

auxiliary protein required for GT activity (Borisova *et al.*, 2004). Later, more putative auxiliary proteins were found based on sequence homology and, in most cases, the genes coding for the auxiliary proteins were found immediately upstream of the gene encoding the corresponding GT (Melançon III *et al.*, 2004; Otten *et al.*, 1995). Replacement of the auxiliary protein by a homologous enzyme from another species *in vitro* (Yuan *et al.*, 2005) and *in vivo* (Hong *et al.*, 2007; Wu, H. *et al.*, 2015) showed that many of the auxiliary proteins can work with an unnatural partner. Furthermore the study showed that the specificity of the partner is in the interaction between the proteins, the GT and the auxiliary protein, not between the auxiliary protein and the substrate or product (Borisova and Liu, 2010). On the other hand, the study on the spiramycin gene cluster leaves room for speculation, as there are two auxiliary proteins each associated with a different GT. While one of the auxiliary proteins can activate both GTs, the other cannot (Nguyen *et al.*, 2010). Although the auxiliary proteins are typically associated with the transfer of amino sugars, there is *in vitro* evidence that the transferred moiety may be a sugar without an amino group (Borisova *et al.*, 2006), but such examples from *in vivo* studies have not emerged.

Initial *in vitro* findings suggested that the auxiliary protein is needed for the activation of the GT rather than for the transfer reaction itself; this has been shown for EryCIII (Yuan *et al.*, 2005) and for DesVIII (Borisova *et al.*, 2006). Later, it was proposed that the auxiliary protein together with the GT form a tight complex, which is only formed by the co-expression of the two proteins in the same host. This complex was shown to significantly increase activity compared to the action of an “activated GT”, where a separately purified auxiliary protein and GT are mixed together (Borisova and Liu, 2010). Biochemical characterization of AknT and AknS, the auxiliary protein and GT from the aclacinomycin A pathway, respectively, demonstrated that the two proteins form a stable complex (Leimkuhler *et al.*, 2007; Lu *et al.*, 2005). The interaction between the partners was confirmed by crystallographic studies with the first 3D structure of an auxiliary protein (EryCII) together with the GT (EryCIII) from the biosynthetic pathway of erythromycin D (Moncrieffe *et al.*, 2012). The structural data showed a tight complex of a dimer of heterodimers and indicated a stabilizing effect of the auxiliary protein to the GT. Moreover, the authors proposed that it would specifically change the acceptor site of the GT, thus being an allosteric activator. The fold of EryCII resembles the fold of cytochrome P450 proteins. The auxiliary proteins lack the conserved Cys residue, which is required for coordinating to the heme, but contain an additional N-terminal helix, which provides an interface for the interaction between the GT and the auxiliary protein (Moncrieffe *et al.*, 2012). When the SpnG GT, which is involved in the biosynthesis of spinosyn A was investigated a motif for the interaction between the GT and an auxiliary protein was proposed (H-X-R-X5-D-X5-R-X12-20-D-P-X3-W-L-X12-18-E-X4-G). However, in the case of SpnG, the

auxiliary protein has not been identified, so it is possible that there is a distinct class of auxiliary proteins not yet recognized (Isiorho *et al.*, 2014).

## 2.2. Glycodiversification

Glycodiversification is *derivatization of natural product sugar moieties through exploitation of the sugar biosynthetic machinery* (Thibodeaux *et al.*, 2008). Sugar moieties play a significant role in the biological activity of polyketides, and thus there is a huge interest in altering the glycosylation patterns, as it might lead to more potent compounds. Nonglycosylated compounds can be glycoconjugated to improve them (Gantt *et al.*, 2008), or the intrinsically attached sugar moieties can be altered to change their pharmacological or pharmacokinetic properties (Ge *et al.*, 1999). The added or altered sugar moieties can change for instance the solubility, toxicity or  $K_D$  of binding to their target (Křen and Řezanka, 2008). In a study on warfarin, it was shown that even the target of the compound was altered because the compound was converted from an anticoagulant to a cytotoxic agent upon glycosylation (Peltier-Pain *et al.*, 2011). Altering the sugar moieties in polyketides may also help overcome bacterial resistance (Nguyen *et al.*, 2010). In many cases, drug resistance arises from mutations to the drug target. Drug resistance can be overcome by changing the drug so that it will recognize the mutated target (Liu and Douthwaite, 2002; Nguyen *et al.*, 2010). Another way for bacteria to become drug-resistant is by developing the ability to discharge the drugs out of the cells. By altering the drugs, for example by altering the glycosylation patterns, this efflux of the drug may be inhibited (Fang *et al.*, 2006). In addition to altering the individual sugars, there is also interest in altering the length of the sugar chain (Cipollone *et al.*, 2002; Pratesi *et al.*, 1998; Zhang, G. *et al.*, 2005) to enhance cytotoxicity and, on the other hand, to lessen side effects, as for example in the case of aclacinomycin A (Oki *et al.*, 1975). It contains three sugars and it has been noted to be less cardiotoxic than daunomycin, which contains only one sugar (Weymouth-Wilson, 1997).

Furthermore, as previously mentioned, there are cases where the carbohydrate units are involved in self-resistance (Bolam *et al.*, 2007; Cundliffe and Demain, 2010; Kwan *et al.*, 2010; Vilches *et al.*, 1992). This phenomenon can be utilized in the drug industry by creating prodrugs with more sugar moieties than the final drug to improve solubility and to lessen toxicity at the time of consumption (Bagshawe, 1987; Křen and Řezanka, 2008; Senter *et al.*, 1988). When the drug reaches its intended environment, the surplus sugars are cleaved off either by increased endogenous activity of  $\beta$ -glucuronidase in the inflamed tissue in prodrug monotherapy or by the use of an antibody-directed prodrug therapy, where  $\alpha/\beta$ -glycosidases are directed to the site of the tumor while the prodrug is consumed. The glycosidases then catalyze the cleavage of the additional

sugars on the tumor site, and the prodrug becomes an active drug (Bagshawe, 1987; Křen and Řezanka, 2008; Senter *et al.*, 1988).

In order to change the glycosylation patterns of a given molecule, one approach is to utilize the inherent flexibility of the enzymes involved in the biosynthesis. Nevertheless, in many cases there is only very limited or no inherent promiscuity, and thus one important approach is to alter the proteins related to the synthesis of glycosylated products. Examples of alterations to the enzymes involved in sugar biosynthesis will be discussed in 2.2.1. To create novel compounds with altered glycosylation patterns, one can approach the matter *in vitro* or *in vivo* or with a combination of both approaches. These will be discussed in 2.2.2. Different forms of glycodiversification have been reviewed in several articles (e.g. Griffith *et al.*, 2005; Luzhetskyy and Bechthold, 2008; Luzhetskyy *et al.*, 2008; Méndez *et al.*, 2008; Salas and Méndez, 2007; Thibodeaux *et al.*, 2008).

## 2.2.1. Engineering Enzymes Involved in the Sugar Biosynthesis

### 2.2.1.1. Anomeric Position Phosphorylating Enzymes

The phosphorylation of the sugar can be considered the first step in the sugar biosynthesis (Fig. 3C) for glycosylation of polyketides. As discussed in 2.1.3, the phosphorylating enzymes typically use naturally occurring sugars for example glycogen phosphorylases creating D-glucose-1-phosphate, fucokinase forming  $\beta$ -l-fucose-1-phosphate, and galactokinase generating  $\alpha$ -D-galactose-1-phosphate (Holden *et al.*, 2003; Johnson and Braford, 1990; Park *et al.*, 1998). The natural substrates for these enzymes are thus sugars derived from the primary metabolism. If the natural substrate has been altered prior to phosphorylation, for example by chemical synthesis, the original anomeric kinase may not recognize the sugar as a substrate, as the anomeric kinases usually act on the early stage of sugar biosynthesis. In approaches where the sugar has been altered prior to phosphorylation it would be beneficial to have an anomeric kinase, which can tolerate variation in mass or additional groups in the sugar moiety. In an example, the *E.coli* enzyme GalK, which naturally uses D-galactose and ATP as its substrates, was randomly mutated to be more permissive. Initially, the authors generated a variant that was able to accept differences in most positions of the sugar, but the C4 position of the sugar had to resemble the galactose structure (Hoffmeister *et al.*, 2003b; Yang *et al.*, 2003). However, a permissive variant was finally found in further rounds of mutagenesis (Hoffmeister and Thorson, 2004; Yang, J. *et al.*, 2005). In these studies mentioned, the authors used a combination of random mutagenesis and rational design.

### 2.2.1.2. Nucleotidyltransferases

The attachment of the nucleotide by the sugar-1-phosphate nucleotidyltransferase (Fig. 3E) follows phosphorylation of the sugar. The alteration of the nucleotidyltransferase substrate specificity can lead to an expanded pool of available NDP-sugars for attachment by the GT. Similarly to the GTs, the nucleotidyltransferases are known to be inherently promiscuous (Jiang *et al.*, 2000; 2001; Zhang, Z. *et al.*, 2005). One extensively studied target for alterations is RmlA (also known as E<sub>p</sub>), a thymidyltransferase from *Salmonella enterica* typhimurium LT2 (Lindquist *et al.*, 1993; Barton *et al.*, 2001). Another target has been the thymidyltransferase Cps2L from *Streptococcus pneumoniae* (Jakeman *et al.*, 2008). The wt RmlA is already capable of accepting all naturally occurring NTPs, but at varying level of catalytic efficiency. Mutant and wt RmlA have been used to make NDP-sugar libraries (Jiang *et al.*, 2000; 2001; 2003). There have been efforts in expanding the promiscuity of the enzyme by single point mutations targeting i) the sugar and ii) the nucleotide moiety acceptance (Barton *et al.*, 2002; Moretti and Thorson, 2007) and whole gene error prone PCR and site saturation mutagenesis to expand the activity towards non-native substrates (Moretti *et al.*, 2011). Finally, the altered activities can be assessed by an MS-based method (Zea and Pohl, 2004).

According to Moretti and Thorson (2007), the ultimate goal is to produce a single universal nucleotidyltransferase, which can accept all nucleotides as well as sugars. On the other hand, the nucleotide is important for binding to the GT, as mentioned in 2.1.3. Thus, if the nucleotidyltransferase attaches all nucleotides, the risk is that the downstream working GT only accepts one kind of nucleotide, so the entire pool of the newly created sugar nucleotides would not be accepted by the GT. Hence, one goal might also be to achieve a promiscuous nucleotidyltransferase with regard to the sugar moiety but stringent towards the attached nucleotide in order to assist the downstream working GT.

### 2.2.1.3. Glycosyltransferases

One major target for achieving altered glycosylation patterns is the GT. GTs exhibit inherent promiscuity in many cases. It is more typical that the GTs are flexible towards the sugar donor, but there are studies where the GT has displayed promiscuity towards both the donor sugar and the acceptor aglycone. These include the macrolide GTs DesVII, discussed earlier in 2.1. (Borisova *et al.*, 2006; 2008), OleD, a GT from the oleandomycin pathway (Yang, M. *et al.*, 2005), and VinC, a GT from the biosynthesis of vicenistatin (Minami and Eguchi, 2007; Minami *et al.*, 2005). There has been speculation that the promiscuity towards aglycone comes from the fact that the protein only recognizes certain factors, such as planarity, hydrophobicity and that the compound is cyclic, rather than distinguishes individual factors (Yang, M. *et al.*, 2005).

A more thorough study concerning important factors for the acceptance of aglycone was done by molecular mechanics calculations and a crystal structure (Minami *et al.*, 2005). The conclusions were similar; the size of the compound, the spatial arrangement of polar groups and the glycosyl accepting position, either a hydroxyl group or an amine, were important for substrate recognition and utilization (Minami *et al.*, 2005). The inherent promiscuity still seems to be limited to certain structural features (Park *et al.*, 2009).

As discussed in 2.1.2., GTs that attach the sugars by more than one kind of linkage can be called multifunctional. GTs that link together more than one kind of sugar but with the same kind of linkage are also called multifunctional. As an example, the MtmGIV from the mithramycin pathway is able to recognize two distinct donor and acceptor substrates (Wang *et al.*, 2012). GTs can also act in an iterative way, attaching identical sugars into a single compound. These enzymes may also be called multifunctional, as the acceptor varies (Lu *et al.*, 2004; Luzhetsky *et al.*, 2005; Zhang *et al.*, 2006a). However, given the known promiscuity of GTs in many cases, most GTs can act in a multifunctional way, even if the biological function is limited to the attachment of only one sugar to one aglycone.

In recent years, there has been an enormous increase in available genomic data, which has also affected the available sequences of GTs and enzymes involved in sugar biosynthesis, but structural investigations are lagging behind. The structural data are important for example for mutational studies, which have followed the identifications and characterizations of GTs. OleD, a GT from the oleandomycin pathway, has been found to be intrinsically promiscuous (Yang, M. *et al.*, 2005) and has been used as a template for improved GTs in many studies (Williams *et al.*, 2007; 2008; 2011). There are multiple examples of domain swapping to create chimeric GTs with altered specificities (e.g. Brazier-Hicks *et al.*, 2007; Cartwright *et al.*, 2008; Hoffmeister *et al.*, 2001; 2002; Krauth *et al.*, 2009; Park *et al.*, 2009; Truman *et al.*, 2009). Domain swapping can lead to different or lax specificity, and it can even lead to an enhanced catalytic performance (Truman *et al.*, 2009). One goal can also be to make the GT less promiscuous in order to achieve a more efficient GT towards a desired sugar (Ramos *et al.*, 2009). Other interesting targets for modifications are to alter the chain length of the attached sugars (Krauth *et al.*, 2009) or to change the bond forming capability of the GT (Gantt *et al.*, 2008; Härle *et al.*, 2011).

In order to recognize altered GT activities, robust GT assays are required. One challenge is often the lack of an authentic donor. One way to overcome this limitation is to investigate the GT activity in a reversed reaction and by monitoring the accumulation of the aglycone; instead of attaching the sugar, it is detached from the aglycone (Zhang *et al.*, 2006a–b). As libraries of GTs expand, there is a growing need



for high throughput screening methods (Wagner and Pesnot, 2010). Examples of different assays include an assay relying on pH change, which utilizes the fact that a proton is released during the formation of the glycosidic bond (Deng and Chen, 2004; Park *et al.*, 2009), and another assay relies on the detection of the released inorganic phosphate (Wu *et al.*, 2011). There are also fluorescent-based methods, e.g. a fluorescent coupled acceptor is used in a FACS-based method (Aharoni *et al.*, 2006; Yang *et al.*, 2010), or assays to detect free nucleotides in fluorescent coupled reactions (Kumagai *et al.*, 2014; Lee and Thorson, 2011).

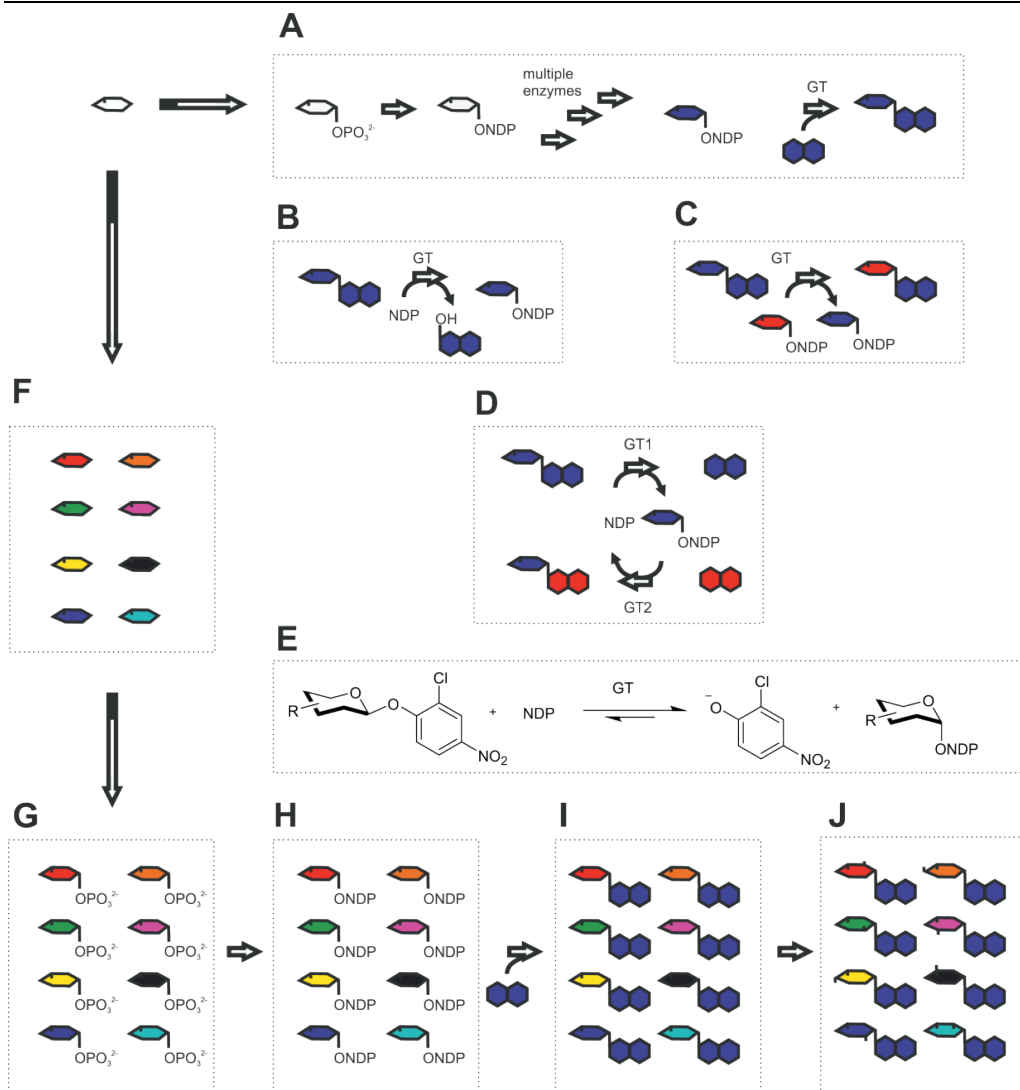
## 2.2.2. Approaches in Glycodiversification

### 2.2.2.1. *In vitro* Approaches in Glycodiversification

There are few examples of *in vitro* reconstruction of glycosylation or sugar pathways (Borisova *et al.*, 2006; 2008; Chen *et al.*, 2000; Takahashi *et al.*, 2006) (simplified example Fig. 9A). There are multiple challenges in the approach, such as the amount of different cosubstrates needed for the sugar modification steps, the need for unknown auxiliary proteins (Kim *et al.*, 2014) as well as the production and purification of the individual proteins. Also problematic are the scalability of the reactions, generalization of the process, and low yields (Yang *et al.*, 2004).

One way to generate diverse sugar moieties is by utilizing the reversibility of GT reactions (Fig. 9B–E), so that the GT detaches the desired sugar from a compound and attaches it to a nucleotide (Fig. 9B). This novel sugar nucleotide can then be used as a substrate to glycosylate another aglycone (Zhang *et al.*, 2006a–b). Transglycosylation can also be used to move a desired sugar from one aglycone to another (Fig. 9C; Zhang *et al.*, 2007), or the aglycone can be changed. This is typically achieved by utilizing two GTs, the first detaching the sugar from the original aglycone and the second attaching the sugar to the novel aglycone (Fig. 9D; Zhang *et al.*, 2006b). The reversibility of the GT reaction can also be utilized with unnatural substrates, as in the example where the authors synthesized 2-chloro-4-nitrophenyl azido attached to amino sugars and a mutated GT (Fig. 9E; Gantt *et al.*, 2013; Williams *et al.*, 2011). Together with the unnatural substrates, a permissive GT and free nucleotides, it was possible to create new sugar nucleotides, which in turn could be used to glycosylate the intended substrate (Zhang, J. *et al.*, 2014).

Another way to achieve structural diversity of sugars is to use chemical synthesis (Andreana *et al.*, 2002; Northrup *et al.*, 2004a–b). In this approach, the sugars are often in a free form without nucleotide activation (Fig. 9F), because this has been proven to be laborious to achieve by organic synthesis (Hoffmeister *et al.*, 2003b). There are different methods to overcome the problem. One is the approach where the synthesis is made in aqueous solution, which is easier to perform than conventional organic



**Figure 9.** Examples of *in vitro* glycosylations. **A)** Typical reaction cascade: The sugar is first phosphorylated and activated by a nucleotide, followed by enzymatic alterations and attachment by a GT to an aglycone. **B)** Utilizing the reversibility of the GT by detaching the sugar from the aglycone and attaching it to a nucleotide, thus creating a novel NDP-sugar. **C)** Exchange of the attached sugar by a supplied NDP-sugar. **D)** Exchange of the aglycone usually by two GTs. **E)** Example of a chemically synthesized donor 2-chloro-4-nitrophenyl azido amino sugars in a reverse reaction with a permissive GT to achieve NDP-sugars. **F)** Using for example chemical synthesis to generate altered sugars **G)** the sugars are phosphorylated and **H)** activated and attached to an aglycone. **J)** Possible additional modifications after the attachment. Figure inspired by Griffith *et al.*, 2005; Zhang *et al.*, 2006b; Zhang, J. *et al.*, 2014.

synthesis and is easier to combine with enzymatic synthesis (Tanaka *et al.*, 2012). A usable substrate for attachment could also be achieved by using for example chemically synthesized sugars, a promiscuous anomeric kinase (Fig. 9G) and a nucleotidyltransferase *in vitro* (Fig. 9H). This approach was used to create vancomycin derivatives (Fu *et al.*, 2003). These NDP-sugars can then be attached to an aglycone by a GT (Fig. 9I). One problem in using anomeric kinases *in vitro* has been the low yields due to substrate (ATP) inhibition. The authors overcame the problem by combining an ATP regeneration system to the synthesis (Liu *et al.*, 2015). Other cosubstrate regeneration systems have been used in *in vitro* enzymatic synthesis, as well (Koeller and Wong, 2000; 2001). Another problem is the need for large amounts of pure enzymes. In one example, this was overcome by recycling the GTs by immobilizing OleD GT on hybrid magnetic nanoparticles (Choi *et al.*, 2012). If the sugar is generated enzymatically, a multienzyme cascade can be difficult to track and measure. For that, an NMR-based method was created (Singh *et al.*, 2014).

Finally, novel sugars can be achieved by altering the sugar moieties only after their attachment to the aglycone. One way of doing this is through the addition of functional groups to the sugar moieties of the glycosylated products by chemical synthesis (Fig. 9J; Nicolau *et al.*, 2001). The other alternative is to use sugar modifying enzymes, which naturally work after the glycosylation, but in that case the substrate specificity of the enzymes may prove problematic, if the novel substrate is considerably different from the natural substrate.

#### **2.2.2.2. *In vivo* Approaches in Glycodiversification**

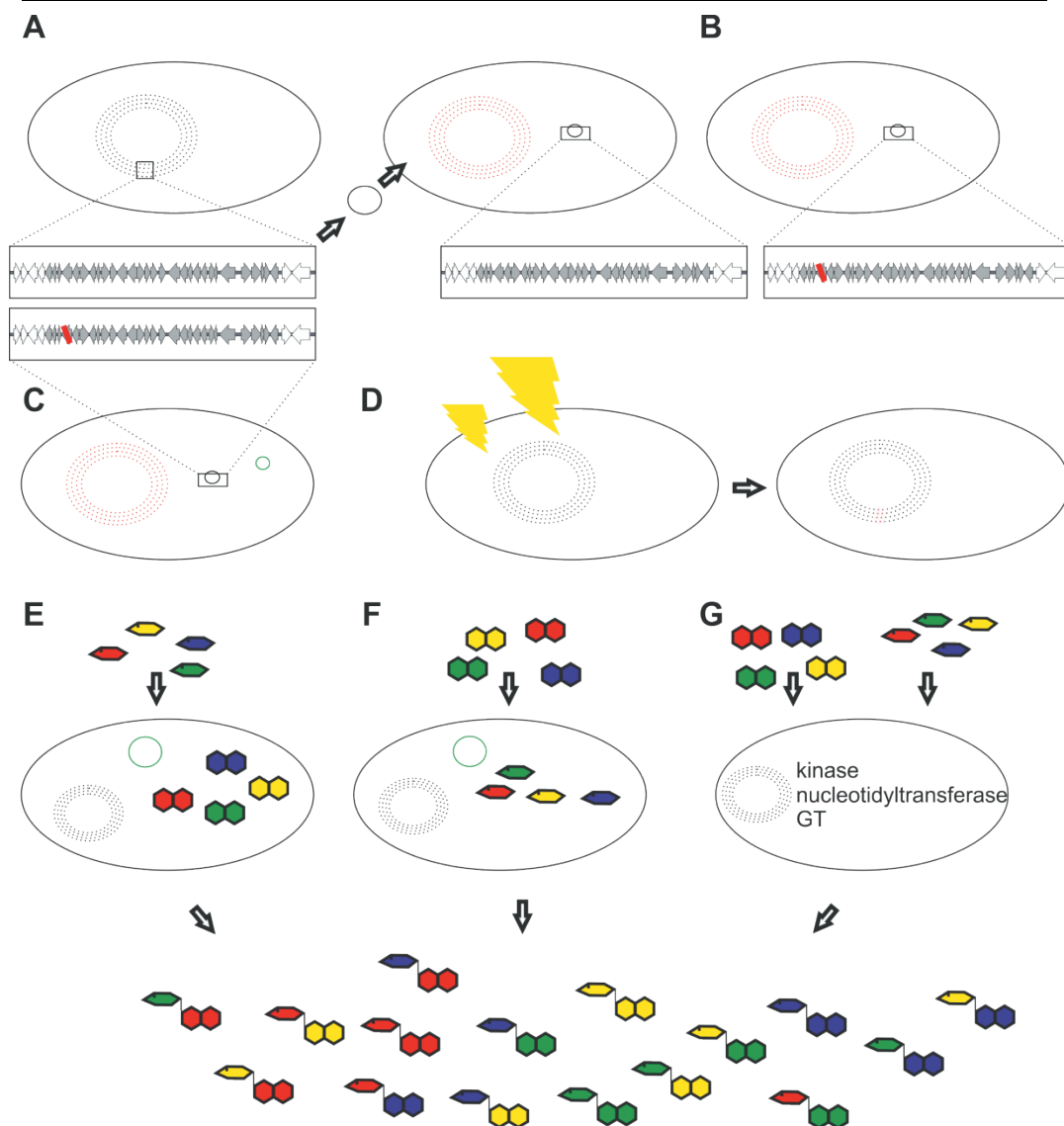
One approach to achieve altered glycosylation patterns *in vivo* is to utilize a heterologous host expression system (Fig. 10A). The heterologous host facilitates gene inactivation experiments (Fig. 10B) and is in many cases needed for efficient genetic modifications. The heterologous host expression system can also help to verify the boundaries of the gene cluster and to determine the functions of genes and the enzymes. The first examples of cloning and expressing the whole cluster for polyketide production in a heterologous host are from the 1980s (Malpartida and Hopwood, 1984; Motamedi and Hutchinson, 1987).

The strains which are lacking *in vivo* gene inactivation strains may produce compounds with altered glycosylation patterns; the biosynthesis may stop at different points, leading to compounds with less sugar moieties than in the final product. The sugar moieties may be altered by inactivating genes encoding sugar-modifying enzymes, as in the example of the pathways of urdamycin A and B. In these studies, multiple genes responsible for the alteration of the sugar units were inactivated, and novel glycosylated products were obtained (Hoffmeister *et al.*, 2000; 2003a). When generating mutants with inactivated genes, possible polar effects should be taken into

account to make sure that changes seen in the phenotype of the mutant are indeed due to the missing product of the inactivated gene and not for example caused by possible problems in expressing the neighboring genes (Thibodeaux *et al.*, 2008).

In addition to inactivating genes, introducing additional genes to the strains with inactivated genes can be used as an *in vivo* glycodiversification tool (Fig. 10C). The first example of an altered sugar moiety in such a way was the alteration of epirubicin; it was achieved by replacing the C4-reductase *dnrV* from the pathway of daunorubicin with the corresponding genes *avrB* or *eryBIV* from the avermectin or erythromycin pathways. The strains produced an epimer of the original compound, a compound not found in nature (Madduri *et al.*, 1998). In many cases when altering the compounds by *in vivo* approaches, the yields of the novel compounds can be modest compared to the original producers (Borisova *et al.*, 2004). As an example, the inactivation of the C5-methyltransferase (CloU) from the biosynthesis of clorobiocin led to very low yields of a novel compound without the methyl group, which was due to the limited substrate specificity of the C4-ketoreductase (CloS), which catalyzes the subsequent reaction on the pathway. The problem was solved by introducing another C4-ketoreductase from another pathway (OleU from the oleandomycin pathway) to the strain with the inactivated gene, thus leading to satisfactory yields of the novel compound (Freitag *et al.*, 2006). In addition, the host can alter the glycosylation profile through the action of endogenous genes that may interfere with the biosynthesis and lead to further diversification of the metabolites produced (Borisova *et al.*, 1999; Hoffmeister *et al.*, 2000; 2003a; Siitonen *et al.*, 2012b). The problem of this kind of glycodiversification is that the pattern may differ between cultivations, as the expression of individual genes may differ, and thus reproducibility may be compromised. Furthermore, if the GT shows insufficient or no promiscuity, it might lead to a situation where novel sugars are created in the cell, but they are not attached to the aglyca. It is also possible to replace the original GT by another to create more diversity. In an example, mutant strains of *Saccharopolyspora erythraea* lacking a GT were complemented by GTs from another strain, leading to novel compounds in addition to restoring the biosynthesis of the original compounds (Doumith *et al.*, 1999).

Randomly mutating the strain that produces glycosylated products can also lead to altered glycosylation patterns (Fig. 10D), as in the case of aclacinomycins produced by *S. galilaeus* (ATCC 31615) (Räty *et al.*, 2000; Ylihonko *et al.*, 1994). The mutations were further investigated by complementation with gene products from the original wt strain but also from another strain, *S. nogalater* (Räty *et al.*, 2002). There have been similar approaches where sugar biosynthetic genes have been combined in gene cassettes to achieve varied combinations of sugars and aglycones. As an example, the aglycone of elloramycin was produced from a cosmid in the heterologous host *S. lividans* in which different combinations of plasmids expressing sugar-altering gene



**Figure 10.** Schematic representation of examples of *in vivo* approaches. A) Heterologous host expression of an antibiotic biosynthetic gene cluster. B) A gene inactivated mutant. C) Introducing additional genes to a gene inactivated strain. D) Random mutagenesis. E) Feeding a strain capable of constructing the aglycone with sugars. F) Feeding a strain capable of making diverse sugars with aglyca or intermediates. G) Feeding a strain expressing a promiscuous anomeric kinase, a nucleotidyltransferase and a GT with sugars and aglyca derived by diverse means.

products were combined, resulting in novel compounds (Fischer *et al.*, 2002; Lombó *et al.*, 2004; Pérez *et al.*, 2005).

All of the examples mentioned above describe strains that contain all of the components required for the production of novel compounds. Another well-established approach for the generation of novel compounds is to conduct biotransformation experiments, where the strains are fed with either sugars, aglyca (or intermediates) or both. It is possible to use a strain that is capable of producing an aglycone and only sugars are fed to the strain (Fig. 10E; Yang *et al.*, 2004). Alternatively, only aglyca or intermediates are fed to a strain that is producing and making modifications to the sugars (Fig. 10F). This approach has been utilized for example with *Streptomyces* (Han *et al.*, 2011) and *E.coli* as the hosts (Peirú *et al.*, 2008). Doxorubicin analogs were generated in the former example and erythromycin D analogs in the latter (Han *et al.*, 2011; Peirú *et al.*, 2008). The limitation of these kinds of approaches is that only the aglycone or the sugar that the strains are capable of producing can be used as the scaffolds for the final products. This is overcome by the approach where an *E. coli* strain contains a modified anomeric kinase (Hoffmeister *et al.*, 2003b; Yang, J. *et al.*, 2005) and a modified nucleotidyltransferase (Barton *et al.*, 2002) alongside a modified GT (Fig. 10G; Williams *et al.*, 2011). The sugars and aglyca are both fed to the strain and, as a result, the sugars are attached to the aglyca *in vivo*. The source of the sugars and aglyca can vary, for example they can be generated by fermenting *in vivo* or *in vitro* enzymatically or chemically (Williams *et al.*, 2011). The advantages of these methods are that it is possible to use nonphosphorylated sugars without a nucleotide as the starting material, and the attached sugar is not determined by the endogenous genes of the host. Disadvantages may be the difficulty of getting the aglyca into the cells (Williams *et al.*, 2011), but this challenge may be overcome with the use of transporters. A universal transporter, which is able to move a broad range of compounds out of the cell, could help solve this particular issue (Fernández-Moreno *et al.*, 1998; Salas and Méndez, 1998). Another problem may be the toxicity of the proteins and the compounds for the heterologous host, which may also cause the strains to favor the production of uninteresting, biologically inactive, derivatives (Thorson *et al.*, 2004).

### **2.2.2.3. Comparison of the *in vitro* and *in vivo* Approaches in Glycodiversification**

The benefit of the *in vitro* approach is that it makes it possible to understand the individual steps and the sequence of the steps in more detail compared to studies *in vivo*. Furthermore, by combining enzymes from different pathways, *in vitro* reconstruction allows engineering the reaction cascades to produce new natural products.

The *in vivo* approaches typically give better yields and are cheaper but are less controllable than the *in vitro* approaches. Furthermore, the utilized combinations may work *in vitro*, but *in vivo* the enzymes, substrates, cosubstrates and cofactors may be in different locations and the concentrations may be insufficient for the reactions to take place (Offen *et al.*, 2006).

In conclusion, the *in vitro* approach typically works better as an initial step for better understanding of the system, but when one wants to produce these new compounds on a larger scale, the *in vivo* approach or a combination of *in vitro* and *in vivo* tackles some of the aforementioned problems. Reasons in favor of the *in vitro* system are control over the components and their ratios and easier combination with chemical synthesis leading to chemoenzymatic approaches.

### **3. AIMS OF STUDY**

The aim of the study was to gain an understanding of the late stage tailoring steps in the biosynthesis of the polyketide antibiotic nogalamycin. The thesis focuses especially on the amino sugar moiety of the compound, which is attached in a highly unusual fashion by an oxygen–carbon and a carbon–carbon bond.

More specifically the aims were:

- I to express the nogalamycin cluster in a heterologous host, elucidate the structures of new compounds and identify the genes responsible for the glycosyl transfer reactions;
- II to understand how position C-1 of nogalamycin is hydroxylated prior to glycosylation;
- III to solve the structure of the SnogD GT and conduct structural and functional investigations;
- IV to identify and investigate the enzyme responsible for the C–C bond formation and to elucidate how the nogalamine sugar of nogalamycin is formed.



## 4. MATERIALS AND METHODS

The experimental procedures are described in more depth in the original publications (I–IV). A brief overview is presented here.

### 4.1. Molecular Biology

Standard molecular biology methods were used; PCR, molecular cloning using digestion of DNA with restriction enzymes and ligation with T4 ligase (Thermo Scientific). The *E.coli* TOP10 cloning strain and the pBAD vector system were used for cloning and protein production. E.Z.N.A plasmid mini kit and E.Z.N.A gel extraction kit (Omega Bio-Tek) were used for DNA isolation and purification. Phusion and DyNAzyme II (New England Biolabs) were used as polymerases in the PCR reactions, and the oligonucleotide primers utilized are presented in the corresponding articles.

### 4.2. Heterologous Expression, Purification and Analysis of Proteins

*E.coli* TOP10 cells were grown either in Erlenmeyer flasks or a Fermenter (Bioengineering). After the cells reached the density of OD<sub>600</sub> 0.5–0.8, protein production was induced with 0.02 % (w/v) L-arabinose, and the cells were grown for 16–19 h, typically at room temperature. The cells were collected by centrifugation and, depending on the case, were either frozen or used straight away for purification. In some applications, the cells were washed prior to purification with PBS.

The purification was typically performed with TALON resin (Clontech laboratories Inc.) and gravity flow with a low imidazole buffer during the beginning and wash steps and with a high imidazole concentration during the elution step. Alternatively, a 5 ml HisTrap column (GE Healthcare) attached to an ÄKTA system was used. In most cases, the buffer was exchanged to an imidazole-free buffer by using a PD10 column (GE Healthcare). The purity and size of the proteins were estimated by SDS-PAGE, and the concentration was measured by absorbance (NanoDrop2000, Thermo Scientific).

The structures of the proteins were solved by protein crystallography at the laboratory of our collaborators at Karolinska Institut. Diffraction data were collected at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Details regarding structure refinement are given in the articles. Protein figures were prepared by PyMOL (DeLano Scientific LLC).

### 4.3. Enzymatic Assays

Typical end point measurements were performed in an Eppendorf tube in an end volume of 200–1000  $\mu$ l. The conditions, i.e. the enzyme, substrate and cosubstrate concentrations and the buffer conditions, were varied; the details are given in the original publications. The reaction mixtures were extracted after the completion of the reactions by using chloroform. The chloroform phase was dried under pressurized air or *in vacuo* and dissolved in methanol prior to HPLC analysis. Alternatively, the compounds were extracted by solid phase extraction by Discovery DSC-18 columns (Supelco, Sigma-Aldrich).

Reactions demonstrating oxygen dependence were performed in a Thunberg cuvette. The enzyme and substrate were kept in separate compartments until oxygen was removed by flushing the cuvette multiple times with nitrogen. After the reaction mixture was depleted of air, the enzyme was mixed with the other reaction components. The reaction was recorded using a spectrophotometer [Multiskan GO spectrophotometer (Thermo Scientific)] after which oxygen was reintroduced and measured again.

The consumption of oxygen was measured with the MI-730 Oxygen electrode (Microelectrodes). The formation of H<sub>2</sub>O<sub>2</sub> was determined using the HRP/ABTS system (Szutowicz *et al.*, 1984).

The kinetic assays were performed by changing the amount of one component (substrate or cosubstrate) and keeping the other components constant. The initial velocities were recorded and plotted against the concentration of the varied component. Origin was used for fitting the curves (OriginLab Corporation).

### 4.4. Heterologous Expression of the Nogalamycin Gene Cluster and Gene Inactivation Mutants

The nogalamycin gene cluster was expressed in a heterologous host *S. albus* using a cosmid pCK505 with an added oriT to allow conjugation sequence as the vector, and genes from the nogalamycin gene cluster. The genes from the cluster that did not reside in the cosmid were added in an additional vector pIJT486 (Siitonen *et al.*, 2012b). The gene inactivated constructs were made on pSnogaori by using the  $\lambda$  Red recombinase (Datsenko and Wanner, 2000) and the pFLP2 system (Hoang *et al.*, 1998) using *E. coli* K12 for gene inactivation steps and ET12567/pUZ8002 (Kieser *et al.*, 2000) for intergeneric conjugation.

## 4.5. Expression, Isolation and Purification of Compounds

For analysis or purification of the compounds, the strains were grown in E1soy medium (Siitonen *et al.*, 2012b). For small scale analyses, the strains were grown in 30 ml medium in a 250 ml Erlenmeyer and for large scale in 50 ml (250 ml Erlenmeyer) and 100 ml (500 ml Erlenmeyer).

The compounds were purified in multiple steps. The first step was usually the addition of XAD7-resin (Rohm and Haas, 20 gL<sup>-1</sup>) to the supernatant. The compounds produced by the bacteria were absorbed by the hydrophobic resin and then extracted by methanol after washing the resin with water. The next step was either open column chromatography using normal phase purification and a methanol:chloroform gradient or a liquid-liquid extraction either with chloroform and water or with ethyl acetate supplemented with 1 % acetic acid and water. The final step before HPLC purification was either another liquid-liquid extraction with chloroform or concentration of the aqueous phase. The solvents were evaporated using a Rotavapor RII (Büchi). In some cases, size-exclusion chromatography in methanol was included [Sephadex LH-20 (GE Healthcare)] in the purification scheme.

The final purification step was done by a preparative HPLC with an UV-detector (Merck Hitachi L-6200A), a reverse-phase column (SunFire Prep C18, 5 mm, 10 3 250 mm, Waters), and a gradient from 10 % acetonitrile to 70 % acetonitrile, both containing 18 mM ammonium acetate pH 3.6, or alternatively 15 % acetonitrile containing 0.1 % formic acid to pure acetonitrile.

For the analysis of the *in vitro* product, multiple large scale *in vitro* reactions with a purified enzyme were performed (e.g. in 40 ml) and subsequently extracted with chloroform prior to purification with a combination of size-exclusion [Sephadex LH-20 (GE Healthcare)] and preparative HPLC.

## 4.6. Analysis of the Compounds

### 4.6.1. Thin Layer Chromatography

TLC was used for initial screening or for the fast evaluation of compounds. The mobile phase used was typically a mixture of methanol and chloroform (2:8). The TLC plates were silica-covered glass or aluminum (Merck). The results were evaluated under visible or UV light conditions.

#### 4.6.2. High Performance Liquid Chromatography

Analytical HPLC analyses were performed with a SCL-10Avp HPLC system equipped with an SPD-M10Avp diode array detector (Shimadzu) using a SunFire (C18, 3.5 mm, 2.1 x 150 mm, Waters) or a Kinetex (2.6  $\mu\text{m}$  C18 100  $\text{\AA}$ , 100 x 4.6 mm, Ea, Phenomenex) column. The HPLC analyses were performed as gradient runs. The solvents used were either 0.1 % formic acid in a water acetonitrile mixture to pure acetonitrile or a mixture containing 18 mM ammonium acetate (pH 3.6) in both eluents.

#### 4.6.3. Mass Spectrometry

Protein MS was performed at the Centre for Biotechnology (Turku). The samples were prepared by running an SDS-PAGE gel upon which the bands were visualized by Coomassie staining. The band of interest was isolated and sent for analysis.

The HPLC electrospray ionization (ESI)-MS analyses of the compounds were performed either with a MicrOTOF-Q high resolution MS (Bruker Daltonics) linked to an Agilent 1200 Series HPLC system or with a low resolution MS with a HPLC system (Agilent 1260 Infinity 6120 Quadrupole LC/MS), in most cases with both negative and positive modes. For most of the analyzed compounds, the collection of the data was more successful in the negative mode, so the values obtained in that mode were used in the publications.

#### 4.6.4. Nuclear Magnetic Resonance

NMR was performed with either a 400 MHz or 500 MHz instrument (Bruker Daltonics). The samples were prepared by drying and dissolving the compounds in deuterated solvents, such as methanol-*d*<sub>4</sub>, acetonitrile-*d*<sub>3</sub>, chloroform-*d*, acetone-*d*<sub>6</sub>, DMSO-*d*<sub>6</sub>, or a mixture of two solvents. The conducted experiments included 1D spectral analyses (<sup>13</sup>C, <sup>1</sup>H, 1D-TOCSY) and 2D measurements (COSY, HMBC, HSQC, HSQCDE, NOESY). TopSpin (Bruker Daltonics) was used for spectral analysis.

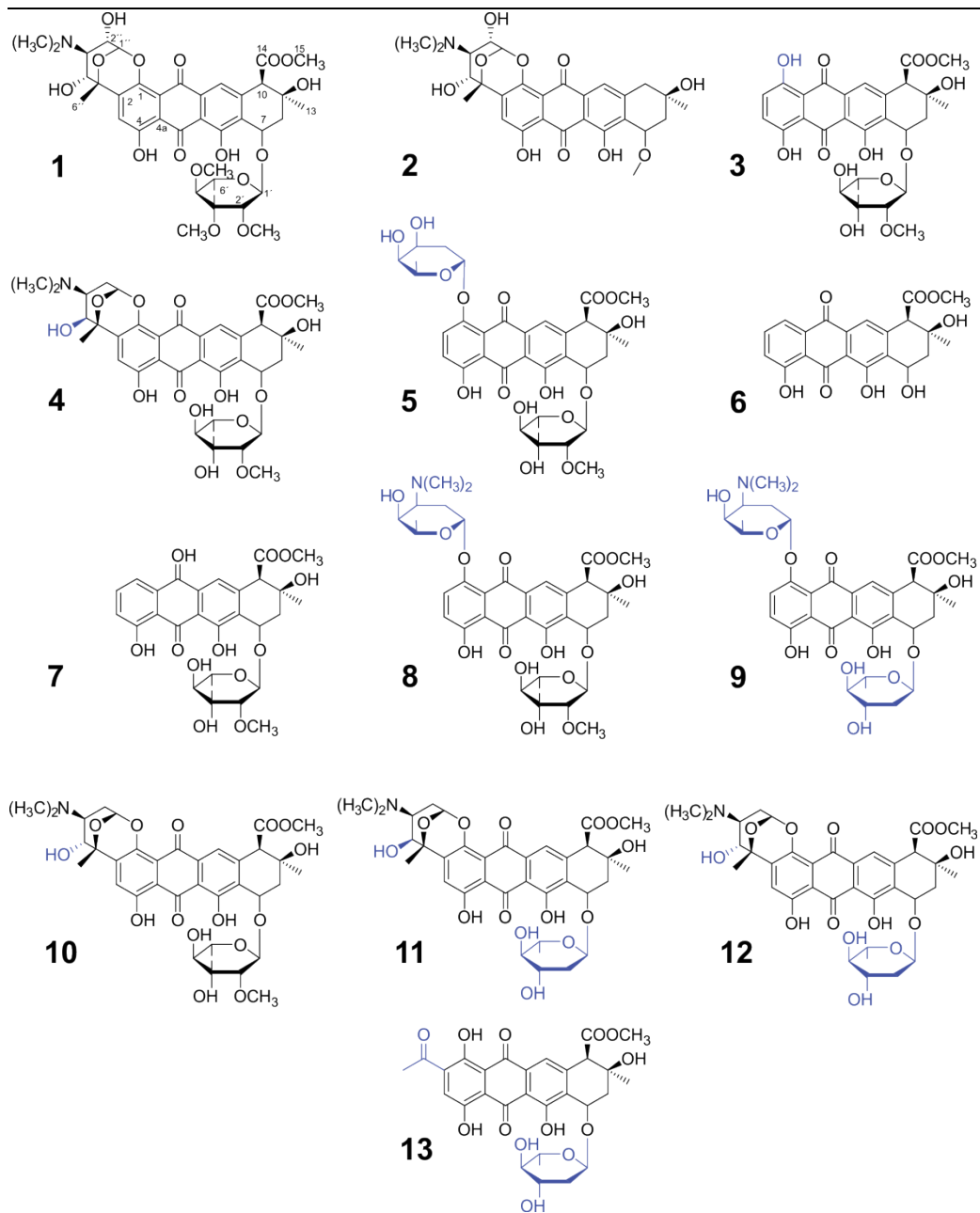
## 5. RESULTS AND DISCUSSION

### 5.1. Heterologous Expression of the Nogalamycin Gene Cluster (I)

The late stages of nogalamycin biosynthesis were the focus of the experimental studies of this thesis. Nogalamycin was first discovered in 1965 (Bhuyan and Dietz) and found to exhibit activity against Gram-positive bacteria and cancer cell lines, but high toxicity prevented its use as medication. Nevertheless, the interesting structural features prompted researchers to create a semisynthetic derivative, menogaril, with less toxicity (Fig. 11: **2**; Wiley, 1979). Furthermore, it inspired us to study the details of the biosynthesis of nogalamycin. *S. nogalater* is the wt producer of nogalamycin, but its restriction modification system makes the genetic engineering of the strain difficult (Ylihonko, personal communication).

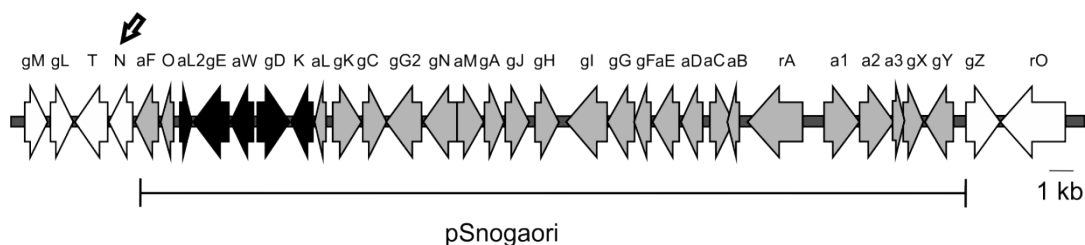
We utilized the heterologous host expression system to overcome the problems related to modifying *S. nogalater*; the heterologous host made it possible to perform genetic modifications on the nogalamycin gene cluster. *S. albus* (Chater and Wilde, 1980) was selected as the heterologous host, as it does not produce any anthracyclines, which might interfere with the analyses. We adopted a two-plasmid approach for the expression of the whole gene cluster due to the size constraints of the cosmid pKC505 denoted pSnogaori (Fig. 12). The end product of the strain that harbors the cosmid pSnogaori and the plasmid pIJTZOMLTN is nogalamycin (Fig. 11: **1**, 13A–C). The strain containing only the cosmid pSnogaori, harboring most of the genes of the nogalamycin gene cluster, denoted *S. albus*/pSnogaori is also capable of producing anthracyclines. All the examined compounds produced by the strain were found to be similar in structure to nogalamycin (Fig. 11: **3–5**, 13) with distinctions. Furthermore, all the compounds produced by this strain are lacking two methyl groups in the neutral deoxy sugar moiety, nogalose. This is logical, since the genes *snogM* and *snogL*, which are most probably responsible for these steps, are not included in the cosmid pSnogaori but reside in the auxiliary plasmid (Fig 12).

The compound 3', 4'-demethoxynogalose-1-hydroxynogalamycinone (Fig. 11: **3**, 13D) is probably an intermediate of the biosynthesis, as it was later also shown to be capable of being used as a substrate for sugar attachment (Siitonen unpublished). Nogalamycin R (Fig. 11, 13D, **5**) closely resembles nogalamycin but lacks an OH-group at the C2'' position and exhibits a different stereochemistry at the C4'' position of the amino sugar, which was determined from the coupling constants obtained from NMR analyses. The data revealed that the amino sugar is rhodosamine ( $J_{3'', 4''}$ : 3.4 Hz), an epimer of nogalamine ( $J_{3'', 4''}$ : 10.5 Hz; Wiley *et al.*, 1977). Based on the structure, we hypothesized that the change in stereochemistry was due to the activity of endogenous enzymes produced by the heterologous host, *S. albus*. Nevertheless, later studies

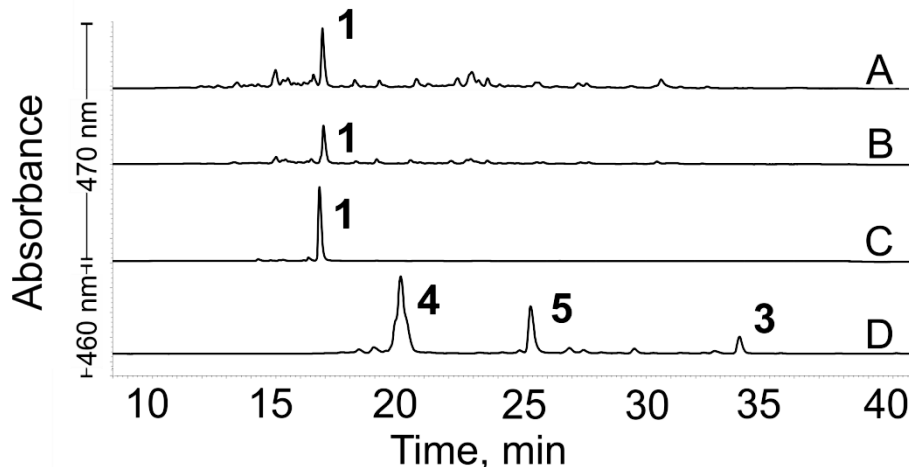


**Figure 11.** The compounds related to the study. 1) Nogalamycin, 2) Menogaril, 3) 3', 4'-demethoxynogalose-1-hydroxynogalamycinone, 4) Nogalamycin R, 5) Nogalamycin F, 6) Nogalamycinone, 7) 3', 4'-demethoxynogalose-nogalamycinone, 8) Nogalamycin K, 9) Nogalamycin KO, 10) Nogalamycin RE, 11) Nogalamycin RO, 12) Nogalamycin ROE and 13) Nogalamycin RON. 3-13 were analyzed by NMR during the course of this study.

confirmed that nogalamycin R may in fact be a true intermediate of the pathway, and the stereochemistry is inverted by a tailoring enzyme SnoN. Furthermore, the corresponding gene, *snoN*, resides outside the cosmid pSnogaori but is present in the nogalamycin gene cluster (Siitonen *et al.*, 2016). The structure of nogalamycin F (Fig. 11: 5, 13), on the other hand, underlies the promiscuity of the GT SnogD, as the sugar moiety attached to C1 is deoxy fucose, a neutral sugar. Moreover, the sugar moiety is



**Figure 12.** The nogalamycin gene cluster studied. The major part of the cluster is in the cosmid pSnogaori. The rest of the genes are present in the plasmid pIJTZOMLTN, marked in white. The genes, which were inactivated from the cosmid pSnogaori in different studies are colored black, and the gene *snoN*, which was studied in IV, is marked with an arrow. The figure is modified from Siitonen *et al.*, 2016.



**Figure 13.** The expression of the biosynthetic gene cluster of nogalamycin in the heterologous host A) The *S. albus*/pSnogaori + pIJTZOMLTN extract B) The co-elution of nogalamycin standard and the product of the strain *S. albus*/pSnogaori + pIJTZOMLTN C) *S. nogalater* crude extract D) *S. albus*/pSnogaori products. The numbers refer to the compounds (Fig. 11). The figure is modified from Siitonen *et al.*, 2012b.

only attached via an O-glycosidic bond, which gives indications of the C–C bond formation. The fact that SnogD attaches a neutral sugar only via an O-glycosidic bond indicates that the carbon–carbon bond formation might require a charged group in the attached sugar. We later showed that SnogE too exhibits promiscuity, as it can, in addition to nogalose, at least add olivose to the C7 position (Fig. 11: **9**, **11–12**) (Siitonen *et al.*, 2016).

While the new compounds produced by the strain *S. albus*/pSnogaori are as such interesting, the strain also made it possible to create gene inactivation mutants. In the first part of the study, we wanted to find out the genes responsible for the sugar transfers. The gene inactivation mutants were designed based on a combination of *in silico* studies and what was previously known about the gene cluster (Torkkell *et al.*, 1997; 2001). Because the structure of nogalamycin is especially interesting due to the mode of attachment of the amino sugar moiety, the first genes of interest were the GTs. Because one of the three annotated GTs (*snogZ*) resides in the plasmid and is missing from the cosmid, and the strain is still capable of attaching both sugars and creating the C–C bond, that gene seems to be redundant. One may speculate that the putative GT coded by the gene *snogZ* could be related to self-resistance, as is shown for example for oleandomycin (Cundliffe and Demain, 2010), but there is no experimental proof for this hypothesis.

The gene inactivation mutant missing the gene *snogE* produces nogalamycinone (Fig 11: **6**) as the main product, implying that SnogE is responsible for the attachment of the deoxy sugar at the C7 position. The gene inactivated mutant lacking the gene for SnogD produces 3', 4'-demethoxynogalose-1-hydroxynogalamycinone (Fig. 11: **3**), the same metabolite that was observed as a minor product of *S. albus*/pSnogaori, indicating that the enzyme is responsible for the attachment of the amino sugar. These gene inactivated mutants as well as the strain containing the cosmid pSnogaori also gave an indication of the sequence in which the sugar moieties are attached to the backbone. First the deoxy sugar is attached to the C7 position and only after that the amino sugar at position C1, as none of the compounds had a sugar moiety only at C1 position.

The new compounds were tested against human topoisomerase I and II, as nogalamycin is a known topoisomerase poison (Sim *et al.*, 1997). As nogalamycin F (Fig. 11: **5**) is structurally more closely related to nogalamycin than 3', 4'-demethoxynogalose-1-hydroxynogalamycinone (Fig. 11: **3**) it is surprising that the former showed less activity against topoisomerase I than the latter. In conclusion, all of the novel compounds showed less activity against both topoisomerases than nogalamycin.



## 5.2. Unusual Monooxygenase System: *SnoaW* and *SnoaL2* (II)

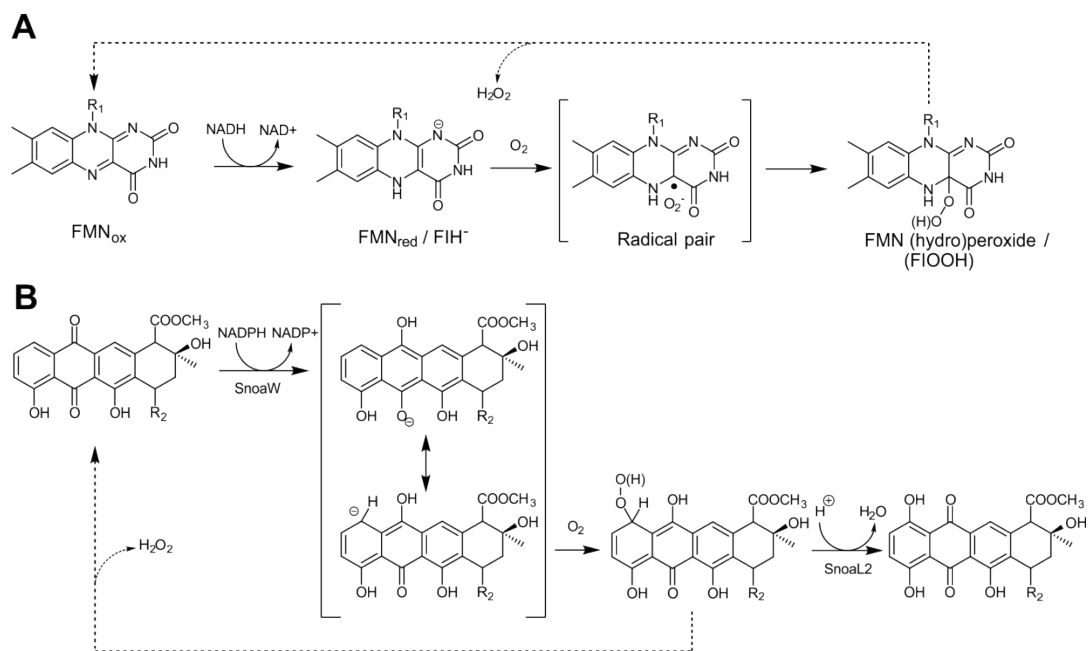
Before the amino sugar can be attached to the polyketide at position C1, the carbon has to be hydroxylated. Previously, it had been shown that *SnoaL2*, a putative monooxygenase, is involved in the C1-hydroxylation reaction in nogalamycin biosynthesis, but no *in vitro* activity could be observed at the time (Beinker *et al.*, 2006; Torkkell *et al.*, 2001). Now the heterologous host expression system made it possible to inactivate genes in the cluster to understand the roles of remaining unknown gene products, which could clarify the hydroxylation reaction. One of the unknown genes was *snoaW*, which encodes a putative SDR enzyme.

The two mutant strains with inactivated genes *snoaL2* and *snoaW* gave similar chromatogram profiles when extracts of the bacterial cultures were analyzed by HPLC. The products contained absorbance maxima at 430 nm, which implied that no hydroxyl groups were present at the C1 position. The additional hydroxyl group shifts the absorbance maximum to 490 nm, as in the case of 3', 4'-demethoxynogalose-nogalamycinone (Fig. 11: 3) due to the change in the conjugated system. The structures of the two main products were investigated and verified to be aglycone, nogalamycinone (Fig. 11: 6) and a previously unknown product: nogalamycinone with a deoxy sugar attached at the C7 position (Fig. 11: 7).

The finding implied that both of the enzymes are needed for the hydroxylation reaction, as the hydroxylation could not proceed *in vivo* if one of the components was missing. The finding was verified by *in vitro* reactions using heterologously expressed His-tagged proteins (*E. coli*, TOP10). We demonstrated that the C1-hydroxylation reaction proceeds in the presence of NADPH or NADH, the two enzymes, and molecular oxygen, but no other components are needed for the reaction. The main products of the gene inactivated strains were tested as substrates in the reaction, and both could be utilized as substrates. It was nevertheless clear that 3', 4'-demethoxynogalose-nogalamycinone (Fig 11:7) is the preferred substrate, because the turnover is nearly 100 % within minutes, whereas nogalamycinone (Fig. 11: 6) has only a modest turnover (20 %) in an overnight reaction.

This finding further gave confirmation for the order of the attachment of the sugar moieties *in vivo*, as the sugar is needed at the C7 position for effective turnover. It implies that the amino sugar is attached only after the deoxy sugar, which was already indicated by the fact that the gene inactivated strain of the deoxy sugar GT, *SnogE*, produces nogalamycinone (Fig. 11: 6) and not menogaril derivatives (Fig. 11: 2). The finding is a setback for the idea of synthesizing menogaril derivatives, lacking the deoxy sugar but containing the amino sugar moiety, in a heterologous host. Menogaril derivatives would be appealing due to the reduced toxicity of menogaril compared to nogalamycin (Wiley, 1979).

Both NADPH and NADH are possible cosubstrates with the preferred primary substrate, 3', 4'-demethoxynogalose-nogalamycinone (Fig 11:7). However, if the substrate nogalamycinone (Fig.11: 6) was utilized, only NADPH led to a reaction. For the completion of the reaction, oxygen is needed, but no additional cofactors or metals are required. If only SnoaW, without the addition of SnoaL2, is used, it results in a non-productive cycle where NAD(P)H is consumed and hydrogen peroxide is formed, but no new end product is obtained. Furthermore, it was possible to capture a short-lived intermediate from a reaction without SnoaL2 and detect it by an HPLC analysis. These data led to a model where the primary substrate itself is used as a cosubstrate in a similar fashion to flavin being used by flavin-dependent monooxygenases (Fig. 14). A similar mechanism was later supported in a study of 1*H*-3-hydroxy-4-oxoquinoline 2, 4-dioxygenase (Hod) from *Arthrobacter nitroguajacolicus* R 61a by electron paramagnetic studies (Thierbach *et al.*, 2014).



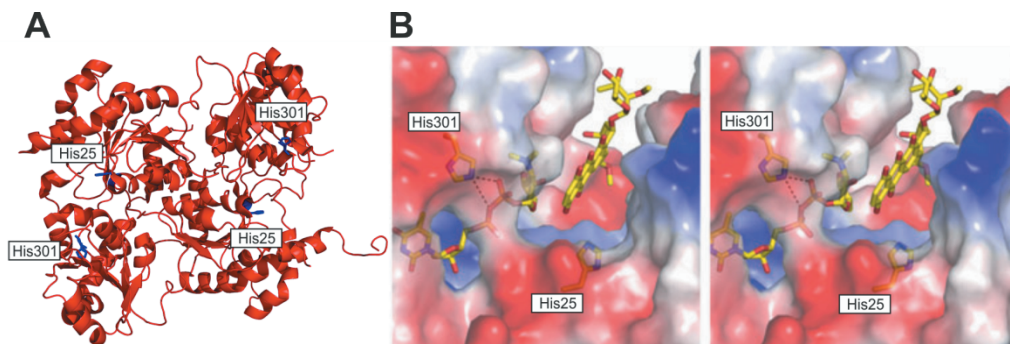
**Figure 14.** The proposed reaction mechanism for SnoaW/SnoaL2 compared to the flavin-dependent monooxygenases. A) Flavin-dependent monooxygenase adapted from (Massey, 1994; Mattevi, 2006; Valton *et al.*, 2008) compared to B) the cosubstrate-independent monooxygenase system SnoaW/SnoaL2 (Figure adapted from Siitonen *et al.*, 2012a).

### 5.3. Structure of the Glycosyltransferase SnogD (III)

In the first part of the studies, the roles of GTs responsible for the transfer of both sugar moieties of nogalamycin were established (Siitonen *et al.*, 2012b). Because the amino

sugar contains the intriguing C–C bond, the GT responsible for the attachment of the amino sugar, SnogD, was examined more closely. The 3D structure was solved by protein crystallography. The enzyme was found to be a dimer in solution as well as in the crystal (Fig. 15A). The crystal structures were obtained both in apo-form and with a nucleotide, dUDP, bound. Based on the crystal structures, a model was made with both ligands: the donor, TDP-nogalamine, and the acceptor 3', 4'-demethoxynogalose-1-hydroxynogalamycinone (Fig. 11: 3, 15B).

SnogD belongs to the GT1 family and thus has the GT-B fold and utilizes the inverting mechanism. Based on the modeled ternary complex, mutants were designed (H301A, H25A, H25N). The mutants were examined *in vitro* in the reverse reaction (Zhang *et al.*, 2006a–b), in which the sugar moiety is detached from the backbone and attached to NTP. The *in vitro* reactions were performed using nogalamycin F (Fig. 11: 5) as the substrate, which led to the accumulation of 3', 4'-demethoxynogalose-1-hydroxynogalamycinone (Fig. 11: 3) as the product. The mutants were also examined *in vivo* by using the gene inactivated strain *S. albus*/pSnoΔgD supplemented with the mutated or wt *snogD* genes in an additional plasmid. The extracts of these strains were examined by HPLC, and the accumulation of amino sugar containing polyketides was investigated. These studies confirmed the importance of His25 and His301 for catalysis, as the activities of the mutants were reduced significantly. His25 acts as a catalytic base in the transfer reaction, whereas His301 binds the diphosphate group of the donor sugar.



**Figure 15.** A) SnogD crystal structure as a dimer. The important histidines are labeled and marked in blue. B) The active site of SnogD in the model of SnogD Michaelis complex with acceptor substrate 3 and donor substrate TDP-nogalamine in the active site. The calculated electrostatic potential is colored from blue to red (increasing negative charge). The figure B) is modified from Claesson *et al.*, 2012.

The study on SnogD showed that it is responsible for the O–C bond formation of the amino sugar but that the C–C bond formation was not likely to be related to the sugar attachment by the GT SnogD. The backward reaction did not proceed with the

substrate with the intact C–C bond, but was only successful with nogalamycin F lacking the C–C bond but with the intact C–O bond. Furthermore, the structure of SnogD shows that the enzyme lacks the catalytic groups or cofactors that would facilitate the abstraction of the hydrogen, which is likely needed for the C–C bond formation.

## 5.4. $\alpha$ -KG-Dependent Dioxygenases: SnoK and SnoN (IV)

### 5.4.1. Carbon–carbon Bond Formation

From the first two studies concerning the sugar attachment (Claesson *et al.*, 2012; Siitonen *et al.*, 2012b), there was evidence that the GT SnogD is not responsible for the C–C bond formation; consequently, there should be another enzyme accountable for it. Furthermore, the gene should reside in the cosmid pSnogaori, since compounds with an intact C–C bond could be detected in expression studies of *S. albus*/pSnogaori (Fig. 11: 4, 12D, 13).

One of the remaining genes of unknown function was *snoK* – a gene coding for an  $\alpha$ -KG and  $\text{Fe}^{2+}$  dependent dioxygenase. Enzymes belonging to this family perform a wide range of reactions (Hausinger, 2004), but this kind of carbocyclization is the first for a member of this family. From the unrelated Rieske family an enzyme responsible for a similar reaction from the streptorubicin B pathway has been found (Sydor *et al.*, 2011). The gene inactivated mutant with missing *snoK*, *S. albus*/pSno $\Delta$ K, produces two main products: nogalamycin K and KO (Fig. 11: 8 and 9), with hypsochromic shifts compared to nogalamycin R (Fig. 11: 4) (from 470 nm to 458 nm), slightly changed hydrophobic properties based on the retention times by HPLC analyses, and notable difference in stability in chloroform. Furthermore, structural elucidation by MS and NMR revealed that both compounds lacked the C–C bond. The formation of the C–C bond can be observed *in vitro* when nogalamycin K or KO (Fig. 11: 8 and 9) are used as substrates for the heterologously expressed His-tagged SnoK. The reaction can be monitored spectrophotometrically as well as with HPLC and MS. If nogalamycin K (Fig. 11: 8) is used as a substrate in the *in vitro* reaction containing  $\alpha$ -KG and  $\text{Fe}^{2+}$ , nogalamycin R (Fig. 11: 4) is formed. The reaction requires oxygen,  $\text{Fe}^{2+}$  and  $\alpha$ -KG. The reactions were typically conducted with the addition of L-ascorbate to prevent the oxidation state of iron from changing from two to three, but the reaction can also proceed in the absence of L-ascorbate. We propose a reaction mechanism where the binding of oxygen to  $\text{Fe}^{2+}$  leads to a  $\text{Fe}^{3+}$ -superoxide, which is converted to a cyclic ferryl-bridged peroxo species after a nucleophilic attack by the carbonyl carbon of  $\alpha$ -KG. The cleavage of the oxygen–oxygen bond and decarboxylation of  $\alpha$ -KG leads to the production of succinate, carbon dioxide and the reactive  $\text{Fe}^{4+}=\text{O}$  species. The steps until this point are typical for  $\alpha$ -KG and  $\text{Fe}^{2+}$  dependent oxygenases. We propose that

the reaction continues by abstraction of the H5''-atom by the Fe<sup>4+</sup>=O species. This creates a substrate radical. Finally, the C–C bond is formed by the attack of the C2 to the radical center (Fig. 16).

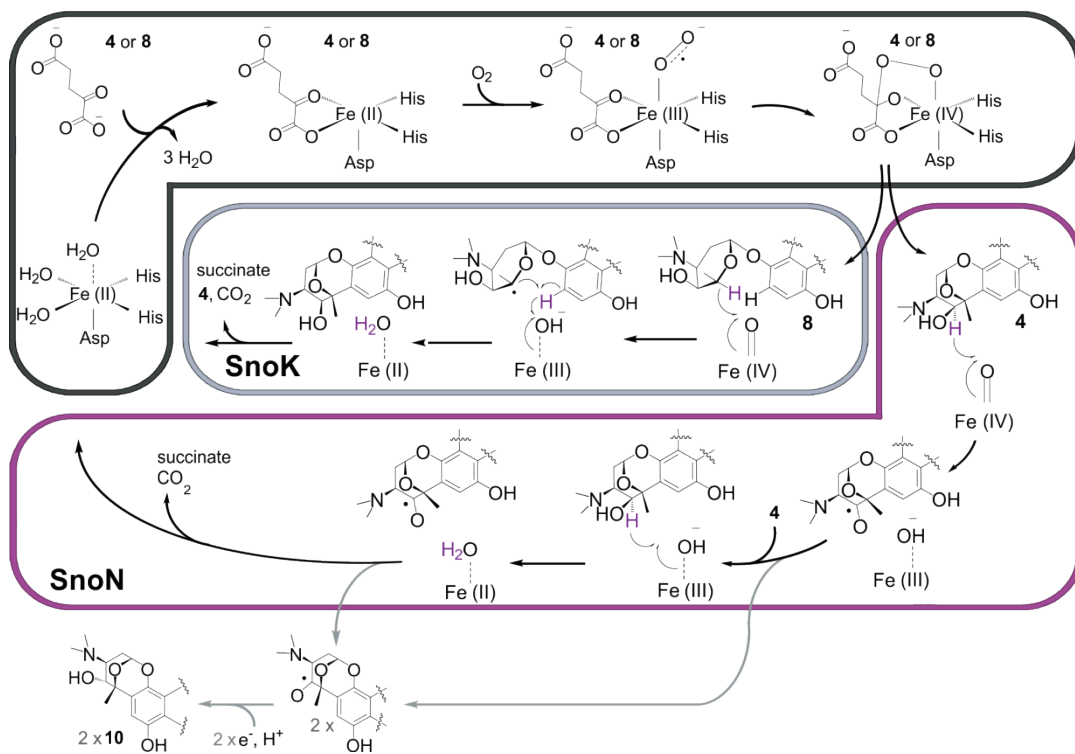
#### 5.4.2. Stereochemistry of C4''

The stereochemical difference between nogalamycin R (Fig. 11: **4**) and nogalamycin (Fig. 11: **1**) resides at the C4'' position, which was previously speculated to arise from gene products expressed by the heterologous host (Siitonen *et al.*, 2012b). The stereochemical change was proven to be the action of another  $\alpha$ -KG and Fe<sup>2+</sup> dependent enzyme, SnoN (Siitonen *et al.*, 2016). The remaining genes outside the cosmid pSnogaori included genes coding for methyltransferases (*snogY*, *snogM*), for resistance (*snorO*), and two genes of unknown function (*snoT*, *snoN*), which encode a Rieske-type dioxygenase SnoT and an  $\alpha$ -KG and Fe<sup>2+</sup> dependent dioxygenase SnoN, respectively.

The unknown gene *snoN*, which based on the sequence analysis produces an  $\alpha$ -KG and Fe<sup>2+</sup> dependent dioxygenase, was introduced to the heterologous host containing the cosmid pSnogaori. In this strain, new products arose with the identical stereochemistry of the C4'' position to that of nogalamycin based on the coupling constant of the hydrogens attached to the C4'' ( $J_{3'',4''}$ : 10.5 **1**,  $J_{3'',4''}$ : 10.6 **10**,  $J_{3'',4''}$ : 10.5 **12**). The strains without the addition of *snoN* produce compounds with significantly smaller coupling constants. This demonstrated *in vivo* that the epimerase changing the stereochemistry of the C4'' is SnoN and that the product nogalamycin R (Fig. 11: **4**) produced by the strain *S. albus*/pSnogaori may indeed be a true intermediate of the pathway, and it is not likely to originate from the interference of genes from the heterologous host *S. albus*, as speculated earlier (Siitonen *et al.*, 2012b).

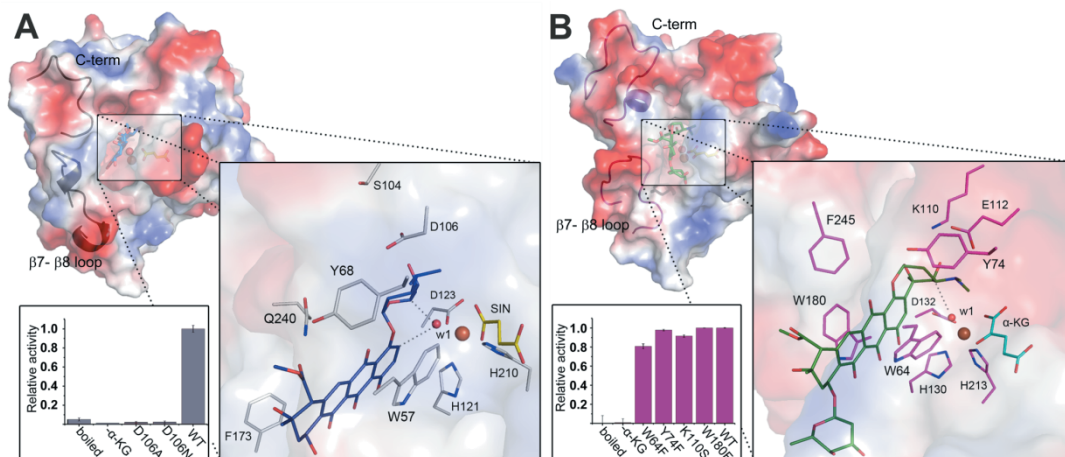
The gene was also expressed in a heterologous host (*E. coli*, TOP10) as a His-tagged protein for *in vitro* characterization. It appears that, instead of generating the biological reaction, the purified SnoN generates compounds with smaller masses when nogalamycin R (Fig. 11: **4**), nogalamycin RO (Fig. 11: **11**), nogalamycin K (Fig. 11: **8**) or nogalamycin KO (Fig. 11: **9**) are used as substrates. When nogalamycin RO (Fig. 11: **11**) is used as a substrate, instable intermediates are formed, which are readily converted to a stable end product. The structure of this product was solved by MS and NMR, and it shows that the amino sugar has been broken into a ketone group attached to the backbone by a C–C bond (Fig. 11: **13**). The extensive degradation of the amino sugar portion of the substrate is proposed to arise from the appearance of radicals in the reaction. Without the physiological context, the radicals are scavenged by the substrate, thus breaking the compound. A similar enzyme CarC from the pathway of carbapenem has been proposed to act in a similar manner (Chang *et al.*, 2014), but there are

differences in the proposed reaction mechanisms. In the case of CarC, the enzyme is inactivated by the reaction instead of the substrate.



**Figure 16.** The proposed reaction mechanisms of SnoK and SnoN. Figure modified from Siitonen *et al.*, 2016.

CarC achieves stereoinversion by creating a  $Fe^{4+}$ -oxo intermediate, which abstracts hydrogens from C5. Subsequently, the tyrosine Y165 donates the hydrogen to the other side of the substrate, creating an epimer of the original compound (Chang *et al.*, 2014). Based on the CarC structure and mechanism, SnoN mutants were designed (W64F, Y74F and W180F). None of the mutants diminished the consumption of the used substrate, nogalamycin K (Fig. 11: **8**, 17). Furthermore, these amino acids are not as well positioned in relation to the substrate as in CarC. We therefore suggest that amino acid radicals are not necessarily involved in the reaction, as in the case of CarC (Chang *et al.*, 2014). Further evidence for the difference between the CarC and SnoN reactions is given by the usage of  $\alpha$ -KG. We show that SnoN utilizes one  $\alpha$ -KG per two substrate molecules, suggesting that SnoN needs an additional component to donate a hydrogen atom back to the substrate, while SnoN is only responsible for the abstraction of the hydrogen (Fig. 16).



**Figure 17.** The binding of the primary substrates to the enzymes SnoK and SnoN. A) SnoK with nogalamycin K modelled into the active site with important amino acids shown and the activities of the mutants studied visualized. B) SnoN in complex with nogalamycin RO with important amino acids shown and the activities of the mutants studied visualized. Legend: boiled, heat inactivated enzyme; - $\alpha$ -KG, control reaction without  $\alpha$ -KG; WT, wild type enzyme. Figure adapted from Siitonen *et al.*, 2016.

### 5.4.3. Comparison of SnoK and SnoN

The crystal structures of both, SnoK and SnoN were solved. Both are in complex with  $\alpha$ -KG and mononuclear iron. Furthermore, a complex of SnoN with the substrate nogalamycin RO (Fig11: 11) was obtained. The main differences between SnoN and SnoK are highlighted in figure 17; overall, the structures are very similar. The figure shows SnoK with modeled nogalamycin K (Fig. 11: 8) in the active site (Fig. 17A) and SnoN with bound nogalamycin RO (Fig. 11: 11, 17B). The most important difference is in the alignment of the substrate in the active site. The positioning of the amino sugar moiety in front of the iron is different. In SnoK, the residues Ser104 and Asp106 bring the substrate deep into the cleft of the enzyme. The carbons C5'' of the sugar and C2 of the aglycone are positioned next to the iron, facilitating carbocyclization. The importance of the residue in substrate alignment was demonstrated by the mutants D106A/N, which both abolished the activity of SnoK (Fig. 17A). In SnoN, the corresponding residues are the longer residues Lys110 and Glu112, which prevent the substrate from going as deep in the cleft as in the case of SnoK. One hypothesis was that, if Lys is replaced by a shorter residue, it might compromise the catalytic activity of SnoN, but the mutant K110S did not considerably lower the activity (Fig. 17). Finally, two more notable differences affect the positioning of the primary substrate in the active site; the loop between  $\beta$ 7 and  $\beta$ 8 is six amino acids longer in SnoK than in SnoN, whereas the C-terminal section is seven amino acids longer in SnoN than SnoK (Fig 17).

## **6. CONCLUDING REMARKS AND FUTURE PERSPECTIVE**

Within the scope of this thesis, the functions of six previously unknown or unverified gene products from the nogalamycin biosynthesis were established. We now understand how C1-hydroxylation of the aglycone is achieved, what proteins are responsible for the attachment of both sugar moieties, and what proteins are responsible for the C–C bond formation as well as the C4'' epimerization.

The important prerequisite for this advancement was that we were able to produce nogalamycin in the heterologous host *S. albus*. The study provided novel enzyme mechanisms, which included the discovery of an unusual two-component monooxygenase system, and two  $\alpha$ -KG and iron dependent non-heme oxygenases, which catalyze unconventional reactions. During the course of the study, we solved the structures of three enzymes. The SnogD GT provides an important progression in structural biology in the field of unusual sugar biosynthesis. The other two enzymes were two homologous  $\alpha$ -KG and iron dependent non-heme oxygenases. Despite the diverse chemistries that these enzymes catalyze in comparison to their canonical family members, the structures proved to be very similar to those more conventional members of the family.

One fruitful and feasible future prospect is the possibility to make mutagenesis to SnoK and SnoN mutants, as they have very similar structures and somewhat similar sequences (38 % identity). The goal could be to alter the activity of both towards each other – or to change the site of the reaction, epimerization, or C–C bond formation. The studies now show that one important factor in the different chemistries they utilize is the orientation of the substrate in the active site; changing the cavity and the orientation could change the chemistries or the site of the reaction. Additionally, further work is required to elucidate the details of the SnoN reaction *in vitro* and the causes that lead to the degradation of the substrate. Even though we have extensively studied different possibilities regarding the SnoN *in vitro* reaction, it is still possible that future results will shed light on the unknown missing component of the SnoN reaction and enable *in vitro* epimerization of the substrate.

The advances in understanding the biosynthesis of unusual sugars in a broader sense has taken important steps forward in the past years, as was elaborated in the *Literature Review* part of this thesis. Not only knowledge on how the sugars are generated in nature but also knowledge and new techniques on how to create even more diverse sugars as well as how to utilize them as parts of novel compounds have been discussed, and the progression is clear. Nevertheless, the long-awaited breakthrough in achieving novel drugs produced in the ways presented in the thesis has not yet been made. However, with the growing need for finding novel antibiotics and other useful drugs,



### *Concluding Remarks and Future Perspective*

---

the possibility that more effort is made in this branch of biochemistry may one day convert basic science into useful applications. The findings of this thesis could also be used in the long-term in creating novel drugs. I would like to see enzymes and genes that we have studied used as components in elegant machineries when novel compounds are created through synthetic biology.

## **ACKNOWLEDGEMENTS**

The work was carried out to most parts in the University of Turku at the Department of Biochemistry (formerly the Department of Biochemistry and Food Chemistry). It was supported by the Academy of Finland, the Varsinais-Suomi Regional Fund of the Finnish Cultural Foundation and the Finnish Cultural Foundation, the MLS, UTUGS, Alfred Kordelin Foundation as well as the Vilho, Yrjö ja Kalle Väisälä Foundation. I want to thank Turku Centre for Systems Biology, Turku University Foundation and MLS, UTUGS for travel grants. I am deeply grateful for the financial support that I have gotten throughout the years.

Great thanks go to Professor Jyrki Heino as the head of biochemistry and the Dean Professor Reijo Lahti for providing the work environment.

I want to thank my opponent Prof. Dr. Andriy Luzhetskyy and reviewers Dr. Gerald Lackner and Docent David Fewer for investing time and efforts in my thesis.

I want to thank my supervisor Docent Mikko Metsä-Ketelä. You impressed me with your vast knowledge and creative thinking.

My co-authors from the University of Turku: Bastian Blauenburg, Pekka Mäntsälä, Laila Niiranen, Pauli Kallio, Pekka Patrikainen, and Maria Aromaa as well as from Karolinska Institutet: Magnus Claesson, Dobreen Dobritzsch, Brinda Selvaraj, Ylva Lindqvist, and Gunter Schneider are thanked for playing an important role in our publications.

Satu Jasu, Anu Hirvensalo, Teija Luotohaara, Heli Kalevo and Jani Sointusalo – Without you nothing would work. Thank you for all the help! Jari Sinkkonen, Maarit Karonen and Petri Tähtinen are thanked for helping with the equipment at the Instrument Centre and Jukka-Pekka Suomela is thanked for helping with the MS at Food chemistry. I also gratefully acknowledge access to the synchrotron radiation at the European Synchrotron Radiation Facility. Jarmo Käpylä, Pekka Rappu and Heidi Tuominen are thanked for helping out so many times on various issues.

I want to thank my current and former lab mates for making everyday life at work so much nicer, Pekka, Bastian, Pauli, Laila, Terhi, Thadée, Pekka, Jarmo, Kaisa, Keshav, Benjamin, Ämir, Maria, Laura, Laurie, Nadine, Aroa and all the students who have been with us for shorter or longer times especially those who I have supervised: Bastian, Linnea, Aiste, Enkhee, Henri, Erika, Janne, Olli-Pekka and Akke-Pekka. I also want to thank the former and current doctoral candidates at the department for peer support and other colleagues for making the work environment as it is.

## *Acknowledgements*

---

I want to thank the whole scientific community for making science where it is today. I express gratitude to all my friends and relatives for making my life cheerful and full. Thank you Maria, Sanna-Kaisa, Natalia, LeLe and all my other friends and relatives!

I want to thank my parents Arto and Kirsti for leading the way, and for loving and supporting me throughout my life and Aaretti for inspiring me and being my sparring partner my whole life.

Finally, thank you my love, Gosha, for compensating my weaknesses and for amplifying my strengths.

## REFERENCES

- Aharoni, A., Thieme, K., Chiu, C. P. C., Buchini, S., Lairson, L. L., Chen, H., Strynadka, N. C. J., Wakarchuk, W. W., and Withers, S. G. (2006) High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* 3, 609–614.
- Albermann, C., Soriano, A., Jiang, J., Vollmer, H., Biggins, J. B., Barton, W. A., Lesniak, J., Nikolov, D. B., and Thorson, J. S. (2003) Substrate specificity of NovM: implications for novobiocin biosynthesis and glycorandomization. *Org. Lett.* 5, 933–936.
- Allard, S. T. M., Beis, K., Giraud, M. F., Hegeman, A. D., Gross, J. W., Wilmouth, R. C., Whitfield, C., Graninger, M., Messner, P., Allen, A. G., Maskell, D. J., and Naismith, J. H. (2002) Toward a structural understanding of the dehydratase mechanism. *Structure* 10, 81–92.
- Allard, S. T. M., Giraud, M., and Naismith, J. H. (2001) Cellular and Molecular Life Sciences Epimerases: structure, function and mechanism. *C. Cell. Mol. Life Sci.* 58, 1650–1665.
- Andreana, P. R., McLellan, J. S., Chen, Y., and Wang, P. G. (2002) Synthesis of 2,6-Dideoxysugars via Ring-Closing Olefinic Metathesis. *Org. Lett.* 4, 3875–3878.
- Ardevol, A., Iglesias-Fernandez, J., Rojas-Cervellera, V., and Rovira, C. (2016) The reaction mechanism of retaining glycosyltransferases. *Biochem. Soc. Trans.* 44, 51–60.
- Bagshawe, K. D. (1987) Antibody directed enzymes revive anti-cancer prodrugs concept. *Br. J. Cancer* 56, 531–532.
- Barton, W. A., Lesniak, J., Biggins, J. B., Jeffrey, P. D., Jiang, J., Rajashankar, K. R., Thorson, J. S., and Nikolov, D. B. (2001) Structure, mechanism and engineering of a nucleotidyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* 8, 545–551.
- Barton, W. A., Biggins, J. B., Jiang, J., Thorson, J. S., and Nikolov, D. B. (2002) Expanding pyrimidine diphosphosugar libraries via structure-based nucleotidyltransferase engineering. *Proc. Natl. Acad. Sci. U. S. A.* 99, 13397–13402.
- Beinker, P., Lohkamp, B., Peltonen, T., Niemi, J., Mäntsälä, P., and Schneider, G. (2006) Crystal structures of SnoaL2 and AclR: two putative hydroxylases in the biosynthesis of aromatic polyketide antibiotics. *J. Mol. Biol.* 359, 728–740.
- Bhuyan, B. K., and Dietz, A. (1965) Fermentation, taxonomic, and biological studies on nogalamycin. *Antimicrob. Agents Chemother.* 5, 836–844.
- Bililign, T., Griffith, B. R., and Thorson, J. S. (2005) Structure, activity, synthesis and biosynthesis of aryl-C-glycosides. *Nat. Prod. Rep.* 22, 742–760.
- Bililign, T., Hyun, C.-G., Williams, J. S., Czisny, A. M., and Thorson, J. S. (2004) The hedamycin locus implicates a novel aromatic PKS priming mechanism. *Chem. Biol.* 11, 959–969.
- Blankenfeldt, W., Asuncion, M., Lam, J. S., and Naismith, J. H. (2000) The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidyltransferase (RmlA). *EMBO J.* 19, 6652–6663.
- Blankenfeldt, W., Kerr, I. D., Giraud, M. F., McMiken, H. J., Leonard, G., Whitfield, C., Messner, P., Graninger, M., and Naismith, J. H. (2002) Variation on a theme of SDR: dTDP-6-deoxy-L-lyxo-4-hexulose reductase (RmlD) shows a new Mg<sup>2+</sup>-dependent dimerization mode. *Structure* 10, 773–786.
- Bolam, D. N., Roberts, S., Proctor, M. R., Turkenburg, J. P., Dodson, E. J., Martinez-Fleites, C., Yang, M., Davis, B. G., Davies, G. J., and Gilbert, H. J. (2007) The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proc. Natl. Acad. Sci. U. S. A.* 104, 5336–5341.

## References

- Borisova, S. A., Zhao, L., Sherman, D. H., and Liu, H. (1999) Biosynthesis of Desosamine: Construction of a New Macrolide Carrying a Genetically Designed Sugar Moiety. *Org. Lett.* 1, 133–136.
- Borisova, S. A., Kim, H. J., Pu, X., and Liu, H.-W. (2008) Glycosylation of acyclic and cyclic aglycone substrates by macrolide glycosyltransferase DesVII/DesVIII: analysis and implications. *Chembiochem* 9, 1554–1558.
- Borisova, S. A., and Liu, H.-W. (2010) Characterization of glycosyltransferase DesVII and its auxiliary partner protein DesVIII in the methymycin/picromycin biosynthetic pathway. *Biochemistry* 49, 8071–8084.
- Borisova, S. A., Zhang, C., Takahashi, H., Zhang, H., Wong, A. W., Thorson, J. S., and Liu, H. (2006) Substrate specificity of the macrolide-glycosylating enzyme pair DesVII/DesVIII: opportunities, limitations, and mechanistic hypotheses. *Angew. Chem. Int. Ed. Engl.* 45, 2748–2753.
- Borisova, S. A., Zhao, L., Melançon III, C. E., Kao, C.-L., and Liu, H.-W. (2004) Characterization of the glycosyltransferase activity of *desVII*: analysis of and implications for the biosynthesis of macrolide antibiotics. *J. Am. Chem. Soc.* 126, 6534–6535.
- Bourne, Y., and Henrissat, B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr. Opin. Struct. Biol.* 11, 593–600.
- Brazier-Hicks, M., Offen, W. A., Gershtater, M. C., Revett, T. J., Lim, E.-K., Bowles, D. J., Davies, G. J., and Edwards, R. (2007) 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.
- Breton, C., Fournel-Gigleux, S., and Palcic, M. M. (2012) Recent structures, evolution and mechanisms of glycosyltransferases. *Curr. Opin. Struct. Biol.* 22, 540–549.
- Breton, C., and Imberty, A. (1999) Structure/function studies of glycosyltransferases. *Curr. Opin. Struct. Biol.* 9, 563–571.
- Burgie, E. S., and Holden, H. M. (2007) Molecular Architecture of DesI: A Key Enzyme in the Biosynthesis of Desosamine. *Biochemistry* 46, 8999–9006.
- Campbell, R. E., Mosimann, S. C., Tanner, M. E., and Strynadka, N. C. J. (2000) The Structure of UDP-*N*-Acetylglucosamine 2-Epimerase Reveals Homology to Phosphoglycosyl Transferases. *Biochemistry* 39, 14993–15001.
- Cartwright, A. M., Lim, E. K., Kleanthous, C., and Bowles, D. J. (2008) A kinetic analysis of regiospecific glycosylation by two glycosyltransferases of *Arabidopsis thaliana*: Domain swapping to introduce new activities. *J. Biol. Chem.* 283, 15724–15731.
- Chang, A., Singh, S., Helmich, K. E., Goff, R. D., Bingman, C. A., Thorson, J. S., and Phillips, G. N. (2011a) Complete set of glycosyltransferase structures in the calicheamicin biosynthetic pathway reveals the origin of regiospecificity. *Proc. Natl. Acad. Sci. U. S. A.* 108, 17649–17654.
- Chang, A., Singh, S., Phillips, G. N., and Thorson, J. S. (2011b) Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Curr. Opin. Biotechnol.* 22, 800–808.
- Chang, W., Guo, Y., Wang, C., Butch, S. E., Rosenzweig, A. C., Boal, A. K., Krebs, C., and Bollinger, J. M. (2014) Mechanism of the C5 stereoinversion reaction in the biosynthesis of carbapenem antibiotics. *Science* 343, 1140–1144.
- Charnock, S. J., and Davies, G. J. (1999) Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* 38, 6380–6385.
- Chater, K., and Wilde, L. C. (1980) *Streptomyces albus* G Mutants Defective in the *SalGI* Restriction-Modification System. *J. Gen. Microbiol.* 116, 323–334.

## References

- Chen, H., Thomas, M. G., Hubbard, B. K., Losey, H. C., Walsh, C. T., and Burkart, M. D. (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, J.-M., Hou, C., Wang, G., Tsodikov, O. V., and Rohr, J. (2015) Structural Insight into MtmC, a Bifunctional Ketoreductase-Methyltransferase Involved in the Assembly of the Mithramycin Trisaccharide Chain. *Biochemistry* 54, 2481–2489.
- Choi, S. H., Ryu, M., Yoon, Y. J., Kim, D.-M., and Lee, E. Y. (2012) Glycosylation of various flavonoids by recombinant oleandomycin glycosyltransferase from *Streptomyces antibioticus* in batch and repeated batch modes. *Biotechnol. Lett.* 34, 499–505.
- Cipollone, A., Berettoni, M., Bigioni, M., Binaschi, M., Cermele, C., Monteagudo, E., Olivieri, L., Palomba, D., Animati, F., Goso, C., and Maggi, C. A. (2002) Novel anthracycline oligosaccharides: influence of chemical modifications of the carbohydrate moiety on biological activity. *Bioorg. Med. Chem.* 10, 1459–1470.
- Claesson, M., Siitonen, V., Dobritzsch, D., Metsä-Ketelä, M., and Schneider, G. (2012) Crystal structure of the glycosyltransferase SnogD from the biosynthetic pathway of nogalamycin in *Streptomyces nogalater*. *FEBS J.* 279, 3251–3263.
- Coutinho, P. M., Deleury, E., Davies, G. J., and Henrissat, B. (2003) An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* 328, 307–317.
- Cundliffe, E., and Demain, A. L. (2010) Avoidance of suicide in antibiotic-producing microbes. *J. Ind. Microbiol. Biotechnol.* 37, 643–672.
- Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* 97, 6640–6645.
- De Bruyn, F., Maertens, J., Beauprez, J., Soetaert, W., and De Mey, M. (2015) Biotechnological advances in UDP-sugar based glycosylation of small molecules. *Biotechnol. Adv.* 33, 288–302.
- Deng, C., and Chen, R. R. (2004) A pH-sensitive assay for galactosyltransferase. *Anal. Biochem.* 330, 219–226.
- Dong, C., Major, L. L., Srikannathasan, V., Errey, J. C., Giraud, M.-F., Lam, J. S., Graninger, M., Messner, P., McNeil, M. R., Field, R. A., Whitfield, C., and Naismith, J. H. (2007) RmlC, a C3' and C5' carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J. Mol. Biol.* 365, 146–159.
- Doumith, M., Legrand, R., Lang, C., Salas, J. A., and Raynal, M. C. (1999) Interspecies complementation in *Saccharopolyspora erythraea*: elucidation of the function of *oleP1*, *oleG1* and *oleG2* from the oleandomycin biosynthetic gene cluster of *Streptomyces antibioticus* and generation of new erythromycin derivatives. *Mol. Microbiol.* 34, 1039–1048.
- Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R., and Goodenough, P. W. (2001) Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* 26, 740–746.
- Dürr, C., Hoffmeister, D., Wohler, S.-E., Ichinose, K., Weber, M., von Mulert, U., Thorson, J. S., and Bechthold, A. (2004) The glycosyltransferase UrdGT2 catalyzes both C- and O-glycosidic sugar transfers. *Angew. Chem. Int. Ed. Engl.* 43, 2962–2965.
- Elshahawi, S. I., Shaaban, K. A., Kharel, M. K., and Thorson, J. S. (2015) A comprehensive review of glycosylated bacterial natural products. *Chem. Soc. Rev.* 44, 7591–7697.
- Fang, L., Zhang, G., Li, C., Zheng, X., Zhu, L., Xiao, J. J., Szakacs, G., Nadas, J., Chan, K. K., Wang, P. G., and Sun, D. (2006) Discovery of a daunorubicin analogue that exhibits potent antitumor activity and overcomes P-gp-mediated drug resistance. *J. Med. Chem.* 49, 932–941.
- Fernández-Moreno, M. A., Carbó, L., Cuesta, T., Vallín, C., and Malpartida, F. (1998) A silent ABC transporter isolated from *Streptomyces rochei* F20

## References

- induces multidrug resistance. *J. Bacteriol.* **180**, 4017–4023.
- Field, R. A., and Naismith, J. H. (2003) Structural and mechanistic basis of bacterial sugar nucleotide-modifying enzymes. *Biochemistry* **42**, 7637–7647.
- Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496.
- Fischer, C., Rodríguez, L., Patallo, E. P., Lipata, F., Braña, A. F., Méndez, C., Salas, J. A., and Rohr, J. (2002) Digitoxosyltetracenomyacin C and glucosyltetracenomyacin C, two novel elloramycin analogues obtained by exploring the sugar donor substrate specificity of glycosyltransferase ElmGT. *J. Nat. Prod.* **65**, 1685–1689.
- Flint, J., Taylor, E., Yang, M., Bolam, D. N., Tailford, L. E., Martinez-Fleites, C., Dodson, E. J., Davis, B. G., Gilbert, H. J., and Davies, G. J. (2005) Structural dissection and high-throughput screening of mannosylglycerate synthase. *Nat. Struct. Mol. Biol.* **12**, 608–614.
- Freel Meyers, C. L., Oberthu, M., Anderson, J. W., Kahne, D., and Walsh, C. T. (2003) Initial Characterization of Novobiocin Acid Noviosyl Transferase Activity of NovM in Biosynthesis of the Antibiotic Novobiocin. *Biochemistry* **42**, 4179–4189.
- Freitag, A., Méndez, C., Salas, J. A., Kammerer, B., Li, S. M., and Heide, L. (2006) Metabolic engineering of the heterologous production of clorobiocin derivatives and elloramycin in *Streptomyces coelicolor* M512. *Metab. Eng.* **8**, 653–661.
- Frey, P. A. (1996) The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB J.* **10**, 461–470.
- Frey, P. A., and Hegeman, A. D. (2013) Chemical and stereochemical actions of UDP-galactose 4-epimerase. *Acc. Chem. Res.* **46**, 1417–1426.
- Fu, X., Albermann, C., Jiang, J., Liao, J., Zhang, C., and Thorson, J. S. (2003) Antibiotic optimization via *in vitro* glycorandomization. *Nat. Biotechnol.* **21**, 1467–1469.
- Gandia-Herrero, F., Lorenz, A., Larson, T., Graham, I. A., Bowles, D. J., Rylott, E. L., and Bruce, N. C. (2008) Detoxification of the explosive 2,4,6-trinitrotoluene in Arabidopsis: discovery of bifunctional *O*- and *C*-glucosyltransferases. *Plant J.* **56**, 963–974.
- Gantt, R. W., Goff, R. D., Williams, G. J., and Thorson, J. S. (2008) Probing the Aglycon Promiscuity of an Engineered Glycosyltransferase. *Angew. Chem. Int. Ed. Engl.* **47**, 8889–8892.
- Gantt, R. W., Peltier-Pain, P., Singh, S., Zhou, M., and Thorson, J. S. (2013) Broadening the scope of glycosyltransferase-catalyzed sugar nucleotide synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 74648–7653.
- Ge, M., Chen, Z., Onishi, H. R., Kohler, J., Silver, L. L., Kerns, R., Fukuzawa, S., Thompson, C., and Kahne, D. (1999) Vancomycin Derivatives That Inhibit Peptidoglycan Biosynthesis Without Binding D-Ala-D-Ala. *Science* **284**, 507–511.
- Giraud, M. F., Leonard, G. A., Field, R. A., Berling, C., and Naismith, J. H. (2000) RmlC, the third enzyme of dTDP-L-rhamnose pathway, is a new class of epimerase. *Nat. Struct. Biol.* **7**, 398–402.
- Gloster, T. M. (2014) Advances in understanding glycosyltransferases from a structural perspective. *Curr. Opin. Struct. Biol.* **28C**, 131–141.
- Gómez, H., Polyak, I., Thiel, W., Lluch, J. M., and Masgrau, L. (2012) Retaining glycosyltransferase mechanism studied by QM/MM methods: Lipopolysaccharyl- $\alpha$ -1,4-galactosyltransferase C transfers  $\alpha$ -galactose via an oxocarbenium ion-like transition state. *J. Am. Chem. Soc.* **134**, 4743–4752.
- Griffith, B. R., Langenhan, J. M., and Thorson, J. S. (2005) “Sweetening” natural products via glycorandomization. *Curr. Opin. Biotechnol.* **16**, 622–630.

## References

- Grocholski, T., Koskiniemi, H., Lindqvist, Y., Mäntsälä, P., Niemi, J., and Schneider, G. (2010) Crystal structure of the cofactor-independent monooxygenase SnoaB from *Streptomyces nogalater*: implications for the reaction mechanism. *Biochemistry* 49, 934–944.
- Gutmann, A., and Nidetzky, B. (2012) Switching between *O*- and *C*-glycosyltransferase through exchange of active-site motifs. *Angew. Chem. Int. Ed. Engl.* 51, 12879–12883.
- Han, A. R., Park, J. W., Lee, M. K., Ban, Y. H., Yoo, Y. J., Kim, E. J., Kim, E., Kim, B.-G., Sohng, J. K., and Yoon, Y. J. (2011) Development of a *Streptomyces venezuelae*-Based Combinatorial Biosynthetic System for the Production of Glycosylated Derivatives of Doxorubicin and Its Biosynthetic Intermediates. *Appl. Environ. Microbiol.* 77, 4912–4923.
- Hausinger, R. P. (2004) FliI/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* 39 21–68
- He, X. M., and Liu, H.-W. (2002) FORMATION OF UNUSUAL SUGARS: Mechanistic Studies and Biosynthetic Applications. *Annu. Rev. Biochem.* 71, 701–754.
- He, X., Agnihotri, G., and Liu, H.-W. (2000) Novel enzymatic mechanisms in carbohydrate metabolism. *Chem. Rev.* 100, 4615–4661.
- Hertweck, C., Luzhetskyy, A., Rebets, Y., and Bechthold, A. (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* 24, 162–190.
- Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchna, A. J., and Schweizer, H. P. (1998) A broad-host-range FliP-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* 212, 77–86.
- Hoffmeister, D., Dräger, G., Ichinose, K., Rohr, J., and Bechthold, A. (2003a) The C-Glycosyltransferase UrdGT2 Is Unselective toward D- and L-Configured Nucleotide-Bound Rhodinoses. *J. Am. Chem. Soc.* 125, 4678–4679.
- Hoffmeister, D., Ichinose, K., Domann, S., Faust, B., Trefzer, A., Dräger, G., Kirschning, A., Fischer, C., Künzel, E., Bearden, D., Rohr, J., and Bechthold, A. (2000) The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster. *Chem. Biol.* 7, 821–831.
- Hoffmeister, D., Ichinose, K., and Bechthold, A. (2001) Two sequence elements of glycosyltransferases involved in urdamycin biosynthesis are responsible for substrate specificity and enzymatic activity. *Chem. Biol.* 8, 557–567.
- Hoffmeister, D., and Thorson, J. S. (2004) Mechanistic implications of *Escherichia coli* galactokinase structure-based engineering. *ChemBiochem* 5, 989–992.
- Hoffmeister, D., Wilkinson, B., Foster, G., Sidebottom, P. J., Ichinose, K., and Bechthold, A. (2002) Engineered Urdamycin Glycosyltransferases Are Broadened and Altered in Substrate Specificity. *Chem. Biol.* 9, 287–295.
- Hoffmeister, D., Yang, J., Liu, L., and Thorson, J. S. (2003b) Creation of the first anomeric D/L-sugar kinase by means of directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* 100, 13184–13189
- Holden, H. M., Rayment, I., and Thoden, J. B. (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J. Biol. Chem.* 278, 43885–43888.
- Hong, J. S. J., Park, S. J., Parajuli, N., Park, S. R., Koh, H. S., Jung, W. S., Choi, C. Y., and Yoon, Y. J. (2007) Functional analysis of DesVIII homologues involved in glycosylation of macrolide antibiotics by interspecies complementation. *Gene* 386, 123–130.
- Hu, Y., Chen, L., Ha, S., Gross, B., Falcone, B., Walker, D., Mokhtarzadeh, M., and Walker, S. (2003) Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural



## References

- principles of a superfamily of glycosyltransferases. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 845–849.
- Hultin, P. G. (2005) Bioactive C-glycosides from bacterial secondary metabolism. *Curr. Top. Med. Chem.* **5**, 1299–1331.
- Härle, J., Günther, S., Lauinger, B., Weber, M., Kammerer, B., Zechel, D. L., Luzhetskyy, A., and Bechthold, A. (2011) Rational Design of an Aryl-C-glycoside Catalyst from a Natural Product O-glycosyltransferase. *Chem. Biol.* **18**, 520–530.
- Igura, M., Maita, N., Kamishikiryo, J., Yamada, M., Obita, T., Maenaka, K., and Kohda, D. (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* **27**, 234–243.
- Isiorho, E. A., Jeon, B., Kim, N. H., Liu, H., and Keatinge-Clay, A. T. (2014) Structural Studies of the Spinosyn Forosaminyltransferase, SpnP. *Biochemistry* **53**, 4292–4301.
- Jakeman, D. L., Young, J. L., Huestis, M. P., Peltier, P., Daniellou, R., Nugier-Chauvin, C., and Ferrières, V. (2008) Engineering Ribonucleoside Triphosphate Specificity in a Thymidyltransferase. *Biochemistry* **47**, 8719–8725.
- Jiang, J., Albermann, C., and Thorson, J. S. (2003) Application of the nucleotidyltransferase Ep toward the chemoenzymatic synthesis of dTDP-desosamine analogues. *ChemBiochem* **4**, 443–446.
- Jiang, J., Biggins, J. B., and Thorson, J. S. (2001) Expanding the Pyrimidine Diphosphosugar Repertoire: The Chemoenzymatic Synthesis of Amino- and Acetamidoglucofuranosyl Derivatives. *Angew. Chemie* **113**, 1550–1553.
- Jiang, J., Biggins, J. B., and Thorson, J. S. (2000) A General Enzymatic Method for the Synthesis of Natural and “Unnatural” UDP- and TDP-Nucleotide Sugars. *J. Am. Chem. Soc.* **122**, 6803–6804.
- Johnson, L., and Bradford, D. (1990) Glycogen Phosphorylase. *J. Biol. Chem.* **265**, 2409–2412.
- Kantola, J., Kunnari, T., Hautala, A., Hakala, J., Ylihonko, K., and Mäntsälä, P. (2000) Elucidation of anthracyclonone biosynthesis by stepwise cloning of genes for anthracyclines from three different *Streptomyces* spp. *Microbiology* **146**, 155–163.
- Kawai, H., Hayakawa, Y., Nakagawa, M., Furihata, K., Furihata, K., Shimazu, A., Seto, H., and Otake, N. (1987) Arugomycin, a new anthracycline antibiotic I. Taxonomy, fermentation, isolation and physico-chemical properties. *J. Antibiot. (Tokyo)* **40**, 1266–1272.
- Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) Tolerance and specificity of polyketide synthases. *Annu. Rev. Biochem.* **68**, 219–253.
- Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K.F., Hopwood, D. A. (2000) *Practical Streptomyces Genetics* (The John Innes Foundation, Norwich, UK).
- Kikuchi, N., Kwon, Y.-D., Gotoh, M., and Narimatsu, H. (2003) Comparison of glycosyltransferase families using the profile hidden Markov model. *Biochem. Biophys. Res. Commun.* **310**, 574–579.
- Kim, H. J., Choi, S., Jeon, B., Kim, N., Pongdee, R., Wu, Q., and Liu, H. (2014) Chemoenzymatic synthesis of spinosyn A. *Angew. Chem. Int. Ed. Engl.* **53**, 13553–13557.
- Koeller, K. M., and Wong, C. (2001) Enzymes for chemical synthesis. *Nature* **409**, 232–240.
- Koeller, K. M., and Wong, C. (2000) Complex carbohydrate synthesis tools for glycobiologists: enzyme-based approach and programmable one-pot strategies. *Glycobiology* **10**, 1157–1169.
- Krauth, C., Fedoryshyn, M., Schleberger, C., Luzhetskyy, A., and Bechthold, A. (2009) Engineering a function into a glycosyltransferase. *Chem. Biol.* **16**, 28–35.
- Křen, V., and Řezanka, T. (2008) Sweet antibiotics - the role of glycosidic residues in antibiotic and antitumor activity and their randomization. *FEMS Microbiol. Rev.* **32**, 858–889.

## References

- Kubiak, R. L., and Holden, H. M. (2011) Combined structural and functional investigation of a C-3''-ketoreductase involved in the biosynthesis of dTDP-L-digitoxose. *Biochemistry* 50, 5905–5917.
- Kudo, F., Kawabe, K., Kuriki, H., Eguchi, T., and Kakinuma, K. (2005) A new family of glucose-1-phosphate/glucosamine-1-phosphate nucleotidyltransferase in the biosynthetic pathways for antibiotics. *J. Am. Chem. Soc.* 127, 1711–1718.
- Kumagai, K., Kojima, H., Okabe, T., and Nagano, T. (2014) Development of a highly sensitive, high-throughput assay for glycosyltransferases using enzyme-coupled fluorescence detection. *Anal. Biochem.* 447, 146–155.
- Kwan, J. C., and Luesch, H. (2010) Weapons in Disguise—Activating Mechanisms and Protecting Group Chemistry in Nature. *Chemistry* 16, 13020–13029.
- Lairson, L. L., Henrissat, B., Davies, G. J., and Withers, S. G. (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* 77, 521–555.
- Lairson, L. L., Chiu, C. P. C., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C. J., and Withers, S. G. (2004) Intermediate Trapping on a Mutant Retaining  $\alpha$ -Galactosyltransferase Identifies an Unexpected Aspartate Residue. *J. Biol. Chem.* 279, 28339–28344.
- Lau, S. T. B., and Tanner, M. E. (2008) Mechanism and active site residues of GDP-fucose synthase. *J. Am. Chem. Soc.* 130, 17593–17602.
- Lee, H. S., and Thorson, J. S. (2011) Development of a universal glycosyltransferase assay amenable to high-throughput formats. *Anal. Biochem.* 418, 85–88.
- Leimkuhler, C., Fridman, M., Lupoli, T., Walker, S., Christopher, T., and Kahne, D. (2007) Characterization of rhodosaminyl-transfer by the AknS/AknT glycosylation complex and its use in reconstituting the biosynthetic pathway of Aclacinomycin A. *J. Am. Chem. Soc.* 129, 10546–10550.
- Liang, D.-M., Liu, J.-H., Wu, H., Wang, B.-B., Zhu, H.-J., and Qiao, J.-J. (2015) Glycosyltransferases: mechanisms and applications in natural product development. *Chem. Soc. Rev.* 44, 8350–8374.
- Lindquist, L., Kaiser, R., Reeves, P. R., and Lindberg, A. A. (1993) Purification, characterization and HPLC assay of *Salmonella* glucose-1-phosphate thymidyltransferase from the cloned *rfa* gene. *Eur. J. Biochem.* 211, 763–770.
- Liu, J., and Mushegian, A. (2003) Three monophyletic superfamilies account for the majority of the known glycosyltransferases. *Protein Sci.* 12, 1418–1431.
- Liu, M., and Douthwaite, S. (2002) Resistance to the macrolide antibiotic tylosin is conferred by single methylations at 23S rRNA nucleotides G748 and A2058 acting in synergy. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14658–14663.
- Liu, Y., Nishimoto, M., and Kitaoka, M. (2015) Facile enzymatic synthesis of sugar 1-phosphates as substrates for phosphorylases using anomeric kinases. *Carbohydr. Res.* 401, 1–4.
- Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K. P. (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* 474, 350–355.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, 490–495.
- Lombó, F., Gibson, M., Greenwell, L., Braña, A. F., Rohr, J., Salas, J. A., and Méndez, C. (2004) Engineering biosynthetic pathways for deoxysugars: branched-chain sugar pathways and derivatives from the antitumor tetracenomycin. *Chem. Biol.* 11, 1709–1718.
- Lombó, F., Olano, C., Salas, J. A., and Méndez, C. (2009) Chapter 11 Sugar Biosynthesis and Modification. *Methods Enzymol.* 458, 277–307.
- Loutre, C., Dixon, D. P., Brazier, M., Slater, M., Cole, D. J., and Edwards, R., (2003) Isolation of a glucosyltransferase from *Arabidopsis thaliana*

## References

- active in the metabolism of the persistent pollutant 3,4-dichloroaniline. *Plant J.* 34, 485–493.
- Lovering, A. L., de Castro, L. H., Lim, D., and Strynadka, N. C. J. (2007) Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. *Science* 315, 1402–1405.
- Lu, M., and Kleckner, N. (1994) Molecular Cloning and Characterization of the *pgm* Gene Encoding Phosphoglucomutase of *Escherichia coli*. *J. Bacteriol.* 176, 5847–5851.
- Lu, W., Leimkuhler, C., Gatto, G. J., Kruger, R. G., Oberthür, M., Kahne, D., and Walsh, C. T. (2005) AknT is an activating protein for the glycosyltransferase AknS in L-aminodeoxysugar transfer to the aglycone of aclacinomycin A. *Chem. Biol.* 12, 527–534.
- Lu, W., Leimkuhler, C., Oberthür, M., Kahne, D., and Walsh, C. T. (2004) AknK is an L-2-deoxyfucosyltransferase in the biosynthesis of the anthracycline aclacinomycin A. *Biochemistry* 43, 4548–4558.
- Luzhetskyy A., and Bechthold, A. (2008) Features and applications of bacterial glycosyltransferases: current state and prospects. *Appl. Microbiol. Biotechnol.* 80, 945–952.
- Luzhetskyy, A., Méndez, C., Salas, J. A., and Bechthold, A. (2008) Glycosyltransferases, important tools for drug design. *Curr. Top. Med. Chem.* 8, 680–709.
- Luzhetskyy, A., Fedoryshyn, M., Dürr, C., Taguchi, T., Novikov, V., and Bechthold, A. (2005) Iteratively acting glycosyltransferases involved in the hexasaccharide biosynthesis of landomycin A. *Chem. Biol.* 12, 725–729.
- Madduri, K., Kennedy, J., Rivola, G., Inveni-Solari, A., Filippini, S., Zanuso, G., Colombo, A., Gewain, K., Occi, J., MacNeil, D., and Hutchinson, C. (1998) Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of *Streptomyces peucetius*. *Nat. Biotechnol.* 16, 291–294.
- Maita, N., Nyirenda, J., Igura, M., Kamishikiryo, J., and Kohda, D. (2010) Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. *J. Biol. Chem.* 285, 4941–4950.
- Major, L. L., Wolucka, B. A., and Naismith, J. H. (2005) Structure and function of GDP-mannose-3',5'-epimerase; an enzyme which performs three chemical reactions at the same active site. *J. Am. Chem. Soc.* 127, 18309–18320.
- Malpartida, F., and Hopwood D. A. (1984) Molecular cloning of the whole biosynthetic pathway of a *Streptomyces antibiotic* and its expression in a heterologous host. *Nature* 309, 462–464.
- Massey, V. (1994) Activation of Molecular Oxygen by Flavins and Flavoprotein. *J. Biol. Chem.* 269, 22459–22462.
- Mattevi, A. (2006) To be or not to be an oxidase: challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* 31, 276–283.
- Melançon III, C. E., Takahashi, H., and Liu, H. (2004) Characterization of *tylM3/tylM2* and *mydC/mycB* pairs required for efficient glycosyltransfer in macrolide antibiotic biosynthesis. *J. Am. Chem. Soc.* 126, 16726–16727.
- Méndez, C., Luzhetskyy, A., Bechthold, A., and Salas, J. A. (2008) Deoxysugars in bioactive natural products: development of novel derivatives by altering the sugar pattern. *Curr. Top. Med. Chem.* 8, 710–724.
- Menéndez, N., Nur-e-Alam, M., Braña, A. F., Rohr, J., Salas, J. A., and Méndez, C. (2004) Tailoring modification of deoxysugars during biosynthesis of the antitumor drug chromomycin A<sub>3</sub> by *Streptomyces griseus* ssp. *griseus*. *Mol. Microbiol.* 53, 903–915.
- Metsä-Ketelä, M., Niemi, J., Mäntsälä, P., Schneider, G. (2008) *Anthracycline Chemistry and Biology I: Biological Occurrence and Biosynthesis, Synthesis and Chemistry*, ed. Krohn K (Springer-Verlag, Berlin/Heidelberg), pp. 101–140.

## References

- Minami, A., and Eguchi, T. (2007) Substrate Flexibility of Viceniaminyltransferase VinC Involved in the Biosynthesis of Vicenistatin. *J. Am. Chem. Soc.* 129, 5102–5107.
- Minami, A., Uchida, R., Eguchi, T., and Kakinuma, K. (2005) Enzymatic Approach to Unnatural Glycosides with Diverse Aglycon Scaffolds Using Glycosyltransferase VinC. *J. Am. Chem. Soc.* 127, 6148–6149.
- Moncrieffe, M. C., Fernandez, M.-J., Spittler, D., Matsumura, H., Gay, N. J., Luisi, B. F., and Leadlay, P. F. (2012) Structure of the glycosyltransferase EryCIII in complex with its activating P450 homologue EryCII. *J. Mol. Biol.* 415, 92–101.
- Moretti, R., Chang, A., Peltier-Pain, P., Bingman, C. A., Phillips, G. N., and Thorson, J. S. (2011) Expanding the nucleotide and sugar 1-phosphate promiscuity of nucleotidyltransferase RmlA via directed evolution. *J. Biol. Chem.* 286, 13235–13243.
- Moretti, R., and Thorson, J. S. (2007) Enhancing the latent nucleotide triphosphate flexibility of the glucose-1-phosphate thymidyltransferase RmlA. *J. Biol. Chem.* 282, 16942–16947.
- Morgan, J. L. W., Strumillo, J., and Zimmer, J. (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* 493, 181–186.
- Motamedi, H., and Hutchinson, C. R. (1987) Cloning and heterologous expression of a gene cluster for the biosynthesis of tetracenomycin C, the anthracycline antitumor antibiotic of *Streptomyces glaucescens*. *Proc. Natl. Acad. Sci. U. S. A.* 84, 4445–4449.
- Mulichak, A. M., Losey, H. C., Lu, W., Wawrzak, Z., Walsh, C. T., and Garavito, R. M. (2003) Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9238–9243.
- Mulichak, A. M., Lu, W., Losey, H. C., Walsh, C. T., and Garavito, R. M. (2004) Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. *Biochemistry* 43, 5170–5180.
- Nguyen, H. C., Karray, F., Lautru, S., Gagnat, J., Lebrihi, A., Huynh, T. D. H., and Pernodet, J.-L. (2010) Glycosylation steps during spiramycin biosynthesis in *Streptomyces ambofaciens*: involvement of three glycosyltransferases and their interplay with two auxiliary proteins. *Antimicrob. Agents Chemother.* 54, 2830–2839.
- Nicolaou, K. C., Cho, S. Y., Hughes, R., Winssinger, N., Smethurst, C., Labischinski, H., and Endermann, R. (2001) Solid- and solution-phase synthesis of vancomycin and vancomycin analogues with activity against vancomycin-resistant bacteria. *Chemistry* 7, 3798–3823.
- Northrup, A. B., Mangion, I. K., Hettche, F., and MacMillan, D. W. C. (2004a) Enantioselective Organocatalytic Direct Aldol Reactions of  $\alpha$ -Oxyaldehydes: Step One in a Two-Step Synthesis of Carbohydrates. *Angew. Chem. Int. Ed. Engl.* 43, 2152–2154.
- Northrup A. B., MacMillan, D. W. C. (2004b) Two-Step Synthesis of Carbohydrates by Selective Aldol Reactions. *Science* 305, 1752–1755.
- Oberthür, M., Leimkuhler, C., Krüger, R. G., Lu, W., Walsh, C. T., and Kahne, D. (2005) A Systematic Investigation of the Synthetic Utility of Glycopeptide Glycosyltransferases. *J. Am. Chem. Soc.* 127, 10747–10752.
- Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., and Davies, G. J. (2006) Structure of a flavonoid glycosyltransferase reveals the basis for plant natural product modification. *EMBO J.* 25, 1396–1405.
- Olano, C., Abdelfattah, M. S., Gullón, S., Braña, A. F., Rohr, J., Méndez, C., and Salas, J. A. (2008) Glycosylated derivatives of steffimycin: insights into the role of the sugar moieties for the biological activity. *Chembiochem* 9, 624–633.
- Oki, T., Matsuzawa, Y., Yoshimoto, A., Numata, K., Kitamura, I., Hori, S., Takamatsu, A., Umezawa,

## References

- H., Ishizuka, M., Naganawa, H., Suda, H., Hamada, M., and Takeuchi, T. (1975) New antitumor antibiotics aclacinomycins A and B. *J. Antibiot. (Tokyo)* 28, 830–834.
- Otten, S. L., Liu, X., Ferguson, J., and Hutchinson, C. R. (1995) Cloning and Characterization of the *Streptomyces peuceetius* *dnrQS* Genes Encoding a Daunomamine Biosynthesis Enzyme and a Glycosyl Transferase Involved in Daunorubicin Biosynthesis. *J. Bacteriol.* 177, 6688–6692.
- Pak, J. E., Satkunarajah, M., Seetharaman, J., and Rini, J. M. (2011) Structural and Mechanistic Characterization of Leukocyte-Type Core 2  $\beta$ 1,6-*N*-Acetylglucosaminyltransferase: A Metal-Ion-Independent GT-A Glycosyltransferase. *J. Mol. Biol.* 414, 798–811.
- Park, S. H., Pastuszak, I., Drake, R., and Elbein, A. D. (1998) Purification to Apparent Homogeneity and Properties of Pig Kidney L-Fucose Kinase. *J. Biol. Chem.* 273, 5685–5691.
- Park, S.-H., Park, H.-Y., Sohng, J. K., Lee, H. C., Liou, K., Yoon, Y. J., and Kim, B.-G. (2009) Expanding Substrate Specificity of GT-B Fold Glycosyltransferase Via Domain Swapping and High-Throughput Screening. *Biotechnol. Bioeng.* 102, 988–994.
- Peirú, S., Rodríguez, E., Menzella, H. G., Carney, J. R., and Gramajo, H. (2008) Metabolically engineered *Escherichia coli* for efficient production of glycosylated natural products. *Microb. Biotechnol.* 1, 476–486.
- Peltier-Pain, P., Timmons, S. C., Grandemange, A., Benoit, E., and Thorson, J. S. (2011) Warfarin glycosylation invokes a switch from anticoagulant to anticancer activity. *ChemMedChem* 6, 1347–1350.
- Pérez, M., Lombó, F., Zhu, L., Gibson, M., Braña, A. F., Rohr, J., Salas, J. A., and Méndez, C. (2005) Combining sugar biosynthesis genes for the generation of L- and D-amictose and formation of two novel antitumor tetracenomycins. *Chem. Commun. (Camb.)*, 1604–1606.
- Pérez, S., Sarkar, A., Rivet, A., Breton, C., and Imbert, A. (2015) Glyco3D: A portal for structural glycosciences, in *Glycoinformatics SE-18* (Lütteke, T., and Frank, M., Eds.), pp 241–258. Springer New York.
- Pratesi, G., De Cesare, M., Caserini, C., Perego, P., Dal Bo, L., Polizzi, D., Supino, R., Bigioni, M., Manzini, S., Iafrate, E., Salvatore, C., Casazza, A., Arcamone, F., and Zunino, F. (1998) Improved efficacy and enlarged spectrum of activity of a novel anthracycline disaccharide analogue of doxorubicin against human tumor xenografts. *Clin. Cancer Res.* 4, 2833–2839.
- Price, N. J., Reiter, W.-D., and Raikhel, N. V. (2002) Molecular genetics of non-processive glycosyltransferases. *The Arabidopsis Book / American Society of Plant Biologists*, 1, e0025.
- Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem. Sci.* 30, 53–62.
- Quirós, L. M., Carbajo, R. J., Braña, A. F., and Salas, J. A. (2000) Glycosylation of macrolide antibiotics. Purification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*. *J. Biol. Chem.* 275, 11713–11720.
- Ramos, A., Olano, C., Braña, A. F., Méndez, C., and Salas, J. A. (2009) Modulation of deoxysugar transfer by the elloramycin glycosyltransferase ElmGT through site-directed mutagenesis. *J. Bacteriol.* 191, 2871–2875.
- Räty, K., Kunnari, T., Hakala, J., Mäntsälä, P., and Ylihanko, K. (2000) A gene cluster from *Streptomyces galilaeus* involved in glycosylation of aclarubicin. *Mol. Gen. Genet.* 264, 164–172.
- Räty, K., Hautala, A., Torkkell, S., Kantola, J., Mäntsälä, P., Hakala, J., and Ylihanko, K. (2002) Characterization of mutations in aclacinomycin A-non-producing *Streptomyces galilaeus* strains with altered glycosylation patterns. *Microbiology* 148, 3375–3384.
- Rosén, M. L., Edman, M., Sjöström, M., and Wieslander, A. (2004) Recognition of fold and

## References

- sugar linkage for glycosyltransferases by multivariate sequence analysis. *J. Biol. Chem.* **279**, 38683–38692.
- Salas, J. A., and Méndez, C. (1998) Genetic manipulation of antitumor-agent biosynthesis to produce novel drugs. *Trends Biotechnol.* **16**, 475–482.
- Salas, J. A., and Méndez, C. (2007) Engineering the glycosylation of natural products in actinomycetes. *Trends Microbiol.* **15**, 219–232.
- Salas, A. P., Zhu, L., Sánchez, C., Braña, A. F., Rohr, J., Méndez, C., and Salas, J. A. (2005) Deciphering the late steps in the biosynthesis of the anti-tumour indolocarbazole staurosporine: sugar donor substrate flexibility of the StaG glycosyltransferase. *Mol. Microb.* **58**, 17–27.
- Searle, M. S., Bicknell, W., Wakelin, L. P. G., and Denny, W. A. (1991) Anthracycline antibiotic arugomycin binds in both grooves of the DNA helix simultaneously: an NMR and molecular modelling study. *Nucleic Acids Res.* **19**, 2897–2906.
- Searle, M. S., Maynard, A. J., and Williams, H. E. (2003) DNA recognition by the anthracycline antibiotic respinomycin D: NMR structure of the intercalation complex with d(AGACGTCT)<sub>2</sub>. *Org. Biomol. Chem.* **1**, 60–66.
- Senter, P. D., Saulnier, M. G., Schreiber, G. J., Hirschberg, D. L., Brown, J. P., Hellström, I., and Hellström, K. E. (1988) Anti-tumor effects of antibody-alkaline phosphatase conjugates in combination with etoposide phosphate. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 4842–4846.
- Shao, L., Shi, X., Liu, W., Gao, X., Pu, T., Ma, B., and Wang, S. (2015) Inactivation and identification of three genes encoding glycosyltransferase required for biosynthesis of nogalamycin. *Biotechnol. Appl. Biochem.* **62**, 765–771.
- Siitonen, V., Blauenburg, B., Kallio, P., Mäntsälä, P., and Metsä-Ketelä, M. (2012a) Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chem. Biol.* **19**, 638–646.
- Siitonen, V., Claesson, M., Patrikainen, P., Aromaa, M., Mäntsälä, P., Schneider, G., and Metsä-Ketelä, M. (2012b) Identification of late-stage glycosylation steps in the biosynthetic pathway of the anthracycline nogalamycin. *ChemBiochem* **13**, 120–128.
- Siitonen V., Selvaraj, B., Niiranen, L., Lindqvist, Y., Schneider, G., and Metsä-Ketelä, M. (2016) Divergent non-heme iron enzymes in the nogalamycin biosynthetic pathway. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5251–5256.
- Sim, S., Gatto, B., Yu, C., Liu, A. A., Li, T., Pilch, D. S., Lavoie, E. J., and Liu, L. F. (1997) Differential Poisoning of Topoisomerases by Menogaril and Nogalamycin Dictated by the Minor Groove-Binding Nugalose Sugar. *Biochemistry* **36**, 13285–13291.
- Singh, S., Phillips, G. N. Jr., and Thorson, J. S. (2012) The Structural Biology of Enzymes Involved in Natural Product Glycosylation. *Nat. Prod. Rep.* **29**, 1201–1237.
- Singh, S., Peltier-Pain, P., Tonelli, M., and Thorson, J. S. (2014) A General NMR-Based Strategy for the in Situ Characterization of Sugar-Nucleotide-Dependent Biosynthetic Pathways. *Org. Lett.* **16**, 3220–3223.
- Sinnott, M. L. (1990) Catalytic Mechanisms of Enzymic Glycosyl Transfer. *Chem Rev* **90**, 1171–1202.
- Smith, P., Szu, P., Bui, C., Liu, H., and Tsai, S. (2008) Structure and Mutagenic Conversion of E1 Dehydrase: At the Crossroads of Dehydration, Amino Transfer, and Epimerization. *Biochemistry* **47**, 6329–6341.
- Sultana, A., Kallio, P., Jansson, A., Wang, J.-S., Niemi, J., Mäntsälä, P., and Schneider, G. (2004) Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *EMBO J.* **23**, 1911–1921.
- Sun, H.-Y., Lin, S.-W., Ko, T.-P., Pan, J.-F., Liu, C.-L., Lin, C.-N., Wang, A. H.-J., and Lin, C.-H. (2007) Structure and Mechanism of *Helicobacter*

## References

- pylori* Fucosyltransferase. *J. Biol. Chem.* **282**, 9973–9982.
- Sydor, P. K., Barry, S. M., Odulate, O. M., Barona-gomez, F., Haynes, S. W., Corre, C., Song, L., and Challis, G. L. (2011) Regio and Stereodivergent Antibiotic Oxidative Carbocyclizations Catalyzed by Rieske Oxygenase-Like Enzymes. *Nat. Chem.* **3**, 388–392.
- Szutowicz, A., Kobes, R. D., and Orsulak, P. J. (1984) Colorimetric assay for monoamine oxidase in tissues using peroxidase and 2,2'-azinodi(3-ethylbenzthiazoline-6-sulfonic acid) as chromogen. *Anal. Biochem.* **94**, 86–94.
- Takahashi, H., Liu, Y.-N., and Liu, H.-W. (2006) A two-stage one-pot enzymatic synthesis of TDP- L-mycarose from thymidine and glucose-1-phosphate. *J. Am. Chem. Soc.* **128**, 1432–1433.
- Tam, H. K., Härle, J., Gerhardt, S., Rohr, J., Wang, G., Thorson, J. S., Bigot, A., Lutterbeck, M., Seiche, W., Breit, B., Bechthold, A., and Einsle, O. (2015) Structural Characterization of O- and C-Glycosylating Variants of the Landomycin Glycosyltransferase LanGT2. *Angew. Chem. Int. Ed. Engl.* **54**, 2811–2815.
- Tanaka, H., Yoshimura, Y., Jørgensen, M. R., Cuesta-Seijo, J. A., and Hindsgaul, O. (2012) A simple synthesis of sugar nucleoside diphosphates by chemical coupling in water. *Angew. Chem. Int. Ed. Engl.* **51**, 11531–11534.
- Tarbouriech, N., Charnock, S. J., and Davies, G. J. (2001) Three-dimensional structures of the Mn and Mg dTDP complexes of the family GT-2 glycosyltransferase SpsA: a comparison with related NDP-sugar glycosyltransferases. *J. Mol. Biol.* **314**, 655–661.
- Thibodeaux, C. J., Melançon, C. E., and Liu, H. (2008) Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed. Engl.* **47**, 9814–9859.
- Thibodeaux, C. J., Melançon, C. E., and Liu, H. (2007) Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**, 1008–1016.
- Thierbach, S., Bui, N., Zapp, J., Chhabra, S. R., Kappl, R., and Fetzner, S. (2014) Substrate-assisted O<sub>2</sub> activation in a cofactor-independent dioxygenase. *Chem. Biol.* **21**, 217–225.
- Thorson, J. S., Barton, W. A., Hoffmeister, D., Albermann, C., and Nikolov, D. B. (2004) Structure-Based Enzyme Engineering and Its Impact on *In Vitro* Glycorandomization. *Chembiochem* **5**, 16–25.
- Torkkell, S., Kunnari, T., Palmu, K., Mäntsälä, P., Hakala, J., and Ylihonko, K. (2001) The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures. *Mol. Genet. Genomics* **266**, 276–288.
- Torkkell, S., Kunnari, T., Palmu, K., Hakala, J., Mäntsälä, P., and Ylihonko, K. (2000) Identification of a cyclase gene dictating the C-9 stereochemistry of anthracyclines from *Streptomyces nogalater*. *Antimicrob. Agents Chemother.* **44**, 396–399.
- Torkkell, S., Ylihonko, K., Hakala, J., Skurnik, M., and Mäntsälä, P. (1997) Characterization of *Streptomyces nogalater* genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. *Mol. Gen. Genet.* **256**, 203–209.
- Trefzer, A., Salas, J. A., and Bechthold, A. (1999) Genes and enzymes involved in deoxysugar biosynthesis in bacteria. *Nat. Prod. Rep.* **16**, 283–299.
- Trefzer, A., Blanco, G., Remsing, L., Künzel, E., Rix, U., Lipata, F., Braña, A. F., Méndez, C., Rohr, J., Bechthold, A., and Salas, J. A. (2002) Rationally designed glycosylated premithramycins: hybrid aromatic polyketides using genes from three different biosynthetic pathways. *J. Am. Chem. Soc.* **124**, 6056–6062.
- Truman, A. W., Dias, M. V. B., Wu, S., Blundell, T. L., Huang, F., and Spencer, J. B. (2009) Chimeric Glycosyltransferases for the Generation of Hybrid Glycopeptides. *Chem. Biol.* **16**, 676–685.
- Tvaroška, I. (2015) Atomistic insight into the catalytic mechanism of glycosyltransferases by

## References

- combined quantum mechanics/molecular mechanics (QM/MM) methods. *Carbohydr. Res.* 403, 38–47.
- Ubukata, M., Uzawa, J., Osada, H., and Isono, K. (1993a) Respinomycins A1, A2, B, C and D, a novel group of anthracycline antibiotics. II. Physico-chemical properties and structure elucidation. *J. Antibiot. (Tokyo)* 46, 942–951.
- Ubukata, M., Uzawa, J., Osada, H., and Isono, K. (1991) Respinomycin A1, a new anthracycline antibiotic. *J. Antibiot. (Tokyo)* 44, 1274–1276.
- Ubukata, M., Uzawa, J., Osada, H., and Isono, K. (1993b) Respinomycins A1, A2, B, C and D, a novel group of anthracycline antibiotics. II. Taxonomy, fermentation, isolation and biological activities. *J. Antibiot. (Tokyo)* 46, 936–941.
- Ünlügil, U. M., and Rini, J. M. (2000) Glycosyltransferase structure and mechanism. *Curr. Opin. Struct. Biol.* 10, 510–517.
- Valton, J., Mathevon, C., Fontecave, M., Nivière, V., and Ballou, D. P. (2008) Mechanism and regulation of the Two-component FMN-dependent monooxygenase ActVA-ActVB from *Streptomyces coelicolor*. *J. Biol. Chem.* 283, 10287–10296.
- Vilches, C., Hernandez, C., Mendez, C., and Salas, J. A. (1992) Role of Glycosylation and Deglycosylation in Biosynthesis of and Resistance to Oleandomycin in the Producer Organism, *Streptomyces antibioticus*. *J. Bacteriol.* 174, 161–165.
- Wagner, G. K., and Pesnot, T. (2010) Glycosyltransferases and their assays. *ChemBiochem* 11, 1939–1949.
- Walsh, C. T., and Fischbach, M. A. (2010) Natural products version 2.0: connecting genes to molecules. *J. Am. Chem. Soc.* 132, 2469–2493.
- Wang, F., Zhou, M., Singh, S., Yennamalli, R. M., Bingman, C. A., Thorson, J. S., and Phillips, G. J. (2013) Crystal structure of SsfS6, the putative C-glycosyltransferase involved in SF2575 biosynthesis. *Proteins* 81, 1277–1282.
- Wang, G., Pahari, P., Kharel, M. K., Chen, J., Zhu, H., Van Lanen, S. G., and Rohr, J. (2012) Cooperation of Two Bifunctional Enzymes in the Biosynthesis and Attachment of Deoxysugars of the Antitumor Antibiotic Mithramycin. *Angew. Chem. Int. Ed. Engl.* 51, 10638–10642.
- Weymouth-Wilson, A. C. (1997) The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* 14, 99–110.
- Wiley, P. F., Kelly, R. B., Caron, E. L., Wiley, V. H., Johnson, J. H., Mackellar, F. A., and Mizsak, S. A. (1977) Structure of Nogalamycin. *J. Am. Chem. Soc.* 99, 542–549.
- Wiley, P. F. (1979) Improved antitumor activity by modification of nogalamycin. *J. Nat. Prod.* 42, 569–582.
- Williams, G. J., Goff, R. D., Zhang, C., and Thorson, J. S. (2008) Optimizing Glycosyltransferase Specificity via “Hot Spot” Saturation Mutagenesis Presents a Catalyst for Novobiocin Glycorandomization. *Chem. Biol.* 15, 393–401.
- Williams, G. J., Yang, J., Zhang, C., and Thorson, J. S. (2011) Recombinant *E. coli* Prototype Strains for *in Vivo* Glycorandomization. *ACS Chem. Biol.* 6, 95–100.
- Williams, G. J., Zhang, C., and Thorson, J. S. (2007) Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat Chem Biol* 3, 657–662.
- Williams, H. D., Trevaskis, N. L., Charman, S. A., Shanker, R. M., Charman, W. N., Pouton, C. W., and Porter, C. J. H. (2013) Strategies to address low drug solubility in discovery and development. *Pharmacol. Rev.* 65, 315–499.
- Wu, H., Li, W., Xin, C., Zhang, C., Wang, Y., Ren, S., Ren, M., Zhao, W., Yuan, L., Xu, Z., Yuan, H., Geng, M., Zhang, L., Weaver, D. T., and Zhang, B. (2015) *In vivo* investigation to the macrolide-glycosylating enzyme pair DesVII/DesVIII in *Saccharopolyspora erythraea*. *Appl. Microbiol. Biotechnol.* 100, 2257–2266.



## References

- Wu, M.-C., Styles, M. Q., Law, B. J. C., Struck, A.-W., Nunns, L., and Micklefield, J. (2015) Engineered biosynthesis of enduracidin lipoglycopeptide antibiotics using the ramoplanin mannosyltransferase Ram29. *Microbiology* 161, 1338–1347.
- Wu, Z. L., Ethen, C. M., Prather, B., Machacek, M., and Jiang, W. (2011) Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology* 21, 727–733.
- Xu, Z.-S., Xue, W., Xiong, A.-S., Lin, Y.-Q., Xu, J., Zhu, B., Zhao, W., Peng, R.-H., and Yao, Q.-H. (2013) Characterization of a bifunctional O- and N-glycosyltransferase from *Vitis vinifera* in glucosylating phenolic compounds and 3,4-dichloroaniline in *Pichia pastoris* and *Arabidopsis thaliana*. *PLoS One* 8, e80449.
- Yang, J., Fu, X., Jia, Q., Shen, J., Biggins, J. B., Jiang, J., Zhao, J., Schmidt, J. J., Wang, P. G., and Thorson, J. S. (2003) Studies on the Substrate Specificity of *Escherichia coli* Galactokinase. *Org. Lett.* 5, 2223–2226.
- Yang, G., Rich, J. R., Gilbert, M., Wakarchuk, W. W., Feng, Y., and Withers, S. G. (2010) Fluorescence activated cell sorting as a general ultra-high-throughput screening method for directed evolution of glycosyltransferases. *J. Am. Chem. Soc.* 132, 10570–10577.
- Yang, J., Fu, X., Liao, J., Liu, L., and Thorson, J. S. (2005) Structure-based engineering of *E. coli* galactokinase as a first step toward *in vivo* glycorandomization. *Chem. Biol.* 12, 657–664.
- Yang, J., Hoffmeister, D., Liu, L., Fu, X., and Thorson, J. S. (2004) Natural product glycorandomization. *Bioorg. Med. Chem.* 12, 1577–1584.
- Yang, M., Proctor, M. R., Bolam, D. N., Errey, J. C., Field, R. A., Gilbert, H. J., and Davis, B. G. (2005) Probing the breadth of macrolide glycosyltransferases: *in vitro* remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. *J. Am. Chem. Soc.* 127, 9336–9337.
- Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L., and Mäntsälä, P. (1996) A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater*: sequence analysis and complementation of early-block mutations in the anthracycline pathway. *Mol. Gen. Genet.* 251, 113–120.
- Ylihonko, K., Hakala, J., Niemi, J., and Lundel, J. (1994) Isolation and characterization of aclacinomycin A-non-producing *Streptomyces galilaeus* (ATCC 31615) mutants. *Microbiology* 140, 1359–1365.
- Yuan, Y., Barrett, D., Zhang, Y., Kahne, D., Sliz, P., and Walker, S. (2007) Crystal structure of a peptidoglycan glycosyltransferase suggests a model for processive glycan chain synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 5348–5353.
- Yuan, Y., Chung, H. S., Leimkuhler, C., Walsh, C. T., and Walker, S. (2005) *In Vitro* Reconstitution of EryCIII Activity for the Preparation of Unnatural Macrolides. *J. Am. Chem. Soc.* 127, 14128–14129.
- Zea, C. J., and Pohl, N. L. (2004) General assay for sugar nucleotidyltransferases using electrospray ionization mass spectrometry. *Anal. Biochem.* 328, 196–202.
- Zhang, C., Albermann, C., Fu, X., Thorson, J. S., and Highland, A. V. (2006a) The *in Vitro* Characterization of the Iterative Avermectin Glycosyltransferase AveBI Reveals Reaction Reversibility and Sugar Nucleotide Flexibility. *J. Am. Chem. Soc.* 128, 16420–16421.
- Zhang, C., Fu, Q., Albermann, C., Li, L., and Thorson, J. S. (2007) The *in vitro* characterization of the erythronolide mycarosyltransferase EryBV and its utility in macrolide diversification. *ChemBiochem* 8, 385–390.
- Zhang, C., Lee, I., Li, L., and Thorson, J. S. (2006b) Exploiting the Reversibility Glycosyltransferase-Catalyzed Reactions. *Science* 313, 1291–1294.
- Zhang, G., Fang, L., Zhu, L., Aimiwu, J. E., Shen, J., Cheng, H., Muller, M. T., Lee, G. E., Sun, D., and Wang, P. G. (2005) Syntheses and Biological

## References

---

- Activities of Disaccharide Daunorubicins. *J. Med. Chem.* 48, 5269–5278.
- Zhang, H., Zhu, F., Yang, T., Ding, L., Zhou, M., Li, J., Haslam, S. M., Dell, A., Erlandsen, H., and Wu, H. (2014) The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat. Commun.* 5, 4339.
- Zhang, J., Singh, S., Hughes, R. R., Zhou, M., Sunkara, M., Morris, A. J., and Thorson, J. S. (2014) A simple strategy for glycosyltransferase-catalyzed aminosugar nucleotide synthesis. *ChemBiochem* 15, 647–652.
- Zhang, Z., Tsujimura, M., Akutsu, J., Sasaki, M., Tajima, H., and Kawarabayasi, Y. (2005) Identification of an extremely thermostable enzyme with dual sugar-1-phosphate nucleotidyltransferase activities from an acidothermophilic archaeon, *Sulfolobus tokodaii* strain 7. *J. Biol. Chem.* 280, 9698–9705.
- Zhou, M., and Wu, H. (2009) Glycosylation and biogenesis of a family of serine-rich bacterial adhesins. *Microbiology* 155, 317–327.
- Zuccotti, S., Zanardi, D., Rosano, C., Sturla, L., Tonetti, M., and Bolognesi, M. (2001) Kinetic and Crystallographic Analyses Support a Sequential-ordered Bi Bi Catalytic Mechanism for *Escherichia coli* Glucose-1-phosphate Thymidyltransferase. *J. Mol. Biol.* 313, 831–843.