Applications of Glycosyltransferases in Drug Discovery

by

Richard W. Gantt

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Pharmaceutical Sciences)

at the

UNIVERSITY OF WISCONSIN-MADISON

2012

Date of oral examination: Feb. 28th, 2012

The dissertation is approved by the following member of the Final Oral Committee:

Jon S. Thorson, Professor, Pharmaceutical Sciences
Timothy S. Bugni, Assistant Professor, Pharmaceutical Sciences
Charles T. Lauhon, Associate Professor, Pharmaceutical Sciences
Ronald T. Raines, Professor, Biochemistry
Michael G. Thomas, Professor, Bacteriology

Applications of Glycosyltransferases in Drug Discovery

Richard W. Gantt
Under the supervision of Prof. Jon S. Thorson
University of Wisconsin – Madison

Sugar nucleotides play a critical role in biology as substrates for glycosyltransferase enzymes, which are ubiquitous throughout biology and involved in a variety of essential processes which include energy storage, carbohydrate metabolism, intra- and extracellular trafficking and recognition, protein glycosylation, cell wall biosynthesis, and the synthesis of medicinally relevant natural products. The study and application of glycosyltransferases in the biosynthesis of natural products and drug discovery efforts and, indeed, the role of glycosylation in biology in general, continues to be limited by both the lack of glycosyltransferases promiscuous in terms of both nucleotide sugar and aglycon recognition and easily accessible routes to sugar nucleotides. The *Streptomyces antibioticus* glycosyltransferase OleD, however, provides an excellent model system for both understanding general promiscuity and catalyzing formation of nucleotide sugars through application of underutilized reverse reactions. Additionally, both the wild type enzyme and its engineered variants are capable of serving as general catalysts for drug discovery efforts by recognizing over 70 unique aglycon scaffolds, afford facile access to over 40 unique natural and 'unnatural' nucleotide sugars in a combinatorial manner, serve as a model with which to understand and exploit the typical thermodynamic landscape of these reactions, and provide high throughput colorimetric assays for general glycosyl transfer and sugar nucleotide utilization. These discoveries and technology developments are directly applicable to drug discovery, protein engineering, and other sugar nucleotide dependent investigations and promise to inspire future efforts into understanding and harnessing glycosylation's unique role in biology.

Acknowledgements

First, I would like to thank Prof. Jon Thorson for the opportunity to work in his lab on a project that continually challenged and expanded my skills as a scientist. My appreciation for his unwavering support will not be forgotten. My additional committee members Asst. Prof. Tim Bugni, Prof. Charles Lauhon, Prof. Ronald Raines, and Prof. Michael Thomas provided excellent guidance during both my research and preparation of this thesis. I would also like to acknowledge all of the members of the Thorson group over the years who have contributed to this research in whatever small manner. A special 'Thank You' is in order for Dr. Gavin Williams, Dr. Pauline Peltier-Pain, and Dr. Randy Goff. Without the mentorship, training, collaborations, encouragement, and long (and often tedious!) scientific discussions from these individuals, many of the results presented herein would have eluded me. Finally, I would like to thank my friends and family, in particular my parents, Alice and Larry Gantt, for their unmatched levels of encouragement, sacrifice, and support throughout my graduate studies. This thesis is dedicated to Martha Jo Lecroy Gantt (my grandmother) and Addison Ann Gantt (my first niece), whose death and birth, respectively, during the final days of writing served as a reminder of how fragile and fantastic life remains.

Table of Contents

List of Tables

List of Figures

Abbreviations

A.U. ..... absorbance units

GlcNAc ..... *N*-acetyl-D-glucosamine

ADP ..... adenosine 5'-diphosphate

ATP ..... adenosine 5'-triphosphate

A ..... alanine

R ..... arginine

N ..... asparagine

D ..... aspartic acid

C (temperature) ..... Celcius

C (amino acid) ..... cysteine

CDP ..... cytidine 5'-diphosphate

CTP ..... cytidine 5'-triphosphate

Da ..... Dalton

° ..... degree

DNA ..... deoxyribonucleic acid

dATP ..... 2'-deoxyadenosine 5'-triphosphate

dGTP ..... 2'-deoxyguanosine 5'-triphosphate

dCTP ..... 2'-deoxycytidine 5'-triphosphate

TDP ..... 2'-deoxythymidine 5'-diphosphate

TTP ..... 2'-deoxythymidine 5'-triphosphate

ESI-MS ..... electrospray ionization mass spectrometry

EtOAc .....ethyl acetate

g (force).....g-force (non-gravitational force)

g (mass).....gram

GDP.....guanosine 5'-diphosphate

GTP.....guanosine 5'-triphosphate

Glc.....glucose

E .....glutamic acid

Q.....glutamine

G.....glycine

GT .....glycosyltransferase

Hz.....hertz

HPLC .....high pressure liquid chromatography

HRMS .....high resolution mass spectrometry

H.....histidine

hr .....hour

I .....isoleucine

IPTG.....isopropyl β-D-1thiogalactopyranoside

kDa.....kilodalton

L (amino acid).....leucine

L (volume) .....liter

LB .....Luria-Bertani

LC .....liquid chromatography

K.....lysine

m/*z* .....mass / charge

MS.....mass spectrometry

MHz .....megahertz

T_m .....melting temperature

MeOH .....methanol

M.....methionine

µg .....microgram

min .....minute

mL.....milliliter

MES .....4-morpholineethanesulfonic acid

MOPS.....4-morpholinepropanesulfonic acid

nm .....nanometer

NADH.....nicotinamide adenine dinucleotide

NADPH.....nicotinamide adenine dinucleotide phosphate

NDP.....nucleotide 5'-diphosphate

NMP.....nucleotide 5'-monophosphate

NTP.....nucleotide 5'-triphosphate

NMR .....nuclear magnetic resonance

OD.....optical density

PCR.....polymerase chain reaction

F .....phenylalanine

P .....proline

R_f.....retention factor

RT .....room temperature

S .....serine

TLC .....thin layer chromatography

T .....threonine

Tris .....tris(hydroxymethyl)aminomethane

UDP.....uridine 5’-diphosphate

UTP .....uridine 5’-triphosphate

V.....valine

W.....tryptophan

Y.....tyrosine

Chapter 1:

Enzymatic methods for glyco(diversification/randomization) of drugs and small molecules

Portions of this chapter have been previously published as:

Gantt, R.W., Peltier-Pain, P.P. & Thorson, J.S. Enzymatic methods for glyco(diversification/randomization) of drugs and small molecules. *Nat. Prod. Rep*. **28**, 1811-1853 (2011).

1.1. Introduction

Simple and complex carbohydrates are ubiquitous in nature where they play a multitude of functions ranging from simple sources of energy to molecular recognition scaffolds critical to the interactions/communication among a wide array of biomolecules, cells, tissues and organisms[1-5]. While complex carbohydrates or glycosides clearly fall outside the accepted realm of Lipinski's 'rule of five'[6, 7], many carbohydrate-based therapeutics are both orally bioavailable and highly effective[8-13]. Among these, glycosylated natural product-based drugs remain a predominant influence, particularly among existing and emerging anticancer and anti-infective agents[14, 15]. Yet, despite the validated importance of complex carbohydrates and glycosides in drug discovery, studies to systematically understand and/or exploit this class of biomolecule remains limited in part due to a lack of practical synthetic tools. This review attempts to highlight emerging *in vitro* and *in vivo* enzyme-based platforms amenable to the 'glycorandomization' (or 'glycodiversification') of small molecule-based therapeutics. In the context of this review, it is important to note the distinction between 'glyco(randomization/diversification)' – a method explicitly focused upon the generation of sets of analogues of a core drug scaffold wherein the sole diversity element is a carbohydrate - and 'targeted' (bio)syntheses of single glycosylated analogues. The scope of this review is limited to

the former and specifically highlights emerging catalysts/platforms amenable to glycorandomization/diversification.

1.2. The influence of glycosylation upon drug properties

Within naturally-occurring glycoside-based therapeutics, the contribution of the appended sugars has, for many natural product classes, been studied in-depth and reviewed in a number of excellent compilations[10, 16-18]. This section briefly highlights a few select examples of the specific contribution by drug-appended sugars to important drug properties including pharmacokinetics, pharmacodynamics, solubility, mechanism, and potency. The examples presented herein, in conjunction with the expanding impact of complex carbohydrates and glycosides in drug discovery, supports the future development of practical glycorandomization/diversification platforms amenable to high throughput discovery-scale manipulations and/or large-scale production processes.

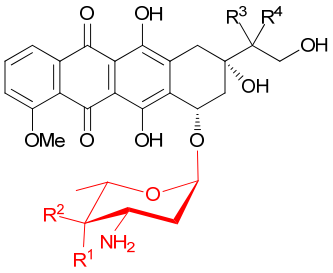
1.2.1. The influence of glycosylation upon drug pharmacokinetics

Anthracyclines are naturally-occurring glycosides which have been used to treat cancer in the clinic for decades. Yet, while >2000 anthracycline derivatives have been synthesized to date[19, 20], the precise mechanism(s) of anthracycline action remain under active debate. A predominant primary mechanism of anthracyclines is topoisomerase II inhibition[20, 21]. Within the context of this review, the sugars attached to anthracyclines are critical to both *in vitro* and *in vivo* activities where aglycons (lacking sugars) typically suffer a 2- to 3-order of magnitude decrease in activity. Importantly, subtle alterations in the attached sugar (*e.g.*, a simple axial to equatorial epimerization of the sugar C-4-OH as exemplified in 4'-epi-doxorubicin (**1**); **Figure**

1.1), renders minimal effect upon the mode of action or spectrum of activity but drastically impacts the drug's pharmacokinetics (PK). Specifically, compared to doxorubicin (**2**), 4'-epi-doxorubicin (**1**) displays an increased volume of distribution (V_d) as well as an enhanced rate of 4-*O*-glucuronidation for generation of **3** and **4** *in vivo*, the latter of which contributes to a marked enhancement of total body clearance (*i.e.*, a shorter terminal $T_{1/2}$)[20]. In addition, doxorubicin (**2**) is more susceptible than 4'-epi-doxorubicin (**1**) to NADPH-dependent cytoplasmic aldo/keto- or carbonyl-reductase-catalyzed reduction at C-13 and the corresponding secondary alcohol metabolites (DOXol (**5**) or EPIol (**6**)) have been proposed as contributors to cardiomyocyte apoptosis and chronic cardiomyopathy[20, 22]. While newer liposomal doxorubicin formulations have offered the ability to also increase the cumulative dose[23], a consequence of the notably distinct PK between doxorubicin and epirubicin is the ability to extend the use of epirubicin to cumulative doses nearly double that for doxorubicin[20, 22, 24]. This example suggests the potential for glycodiversification to improve a drug's pharmacokinetics and/or reduce unwanted dose-limiting toxicities.

1.2.2. The influence of glycosylation upon drug solubility and membrane transport

Tight junctions of the blood–brain barrier (BBB) provide an efficient block to solutes, including blood-borne glucose, from crossing between cells into brain extracellular space. To maintain the brain's high rate of aerobic metabolism, highly regulated transport of glucose, a polar hydrophilic molecule, is accomplished via regionally selective, facilitated transport. The predominant passage of BBB glucose, both luminally and abluminally, is facilitated by an energy-independent hexose transporter GLUT1[25, 26]. As exemplified by studies focused upon the use of glucose-conjugation as a means to facilitate the GLUT1-mediated transport of

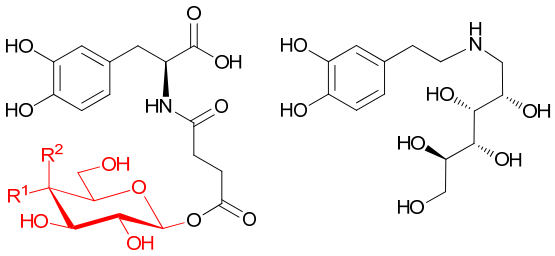


1: R^1 = H, R^2 = OH, R^3, R^4 = oxo
2: R^1 = OH, R^2 = H, R^3, R^4 = oxo
3: R^1 = H, R^2 = β-D-glucuronic acid, R^3, R^4 = oxo
4: R^1 = β-D-glucuronic acid, R^2 = H, R^3, R^4 = oxo
5: R^1 = OH, R^2 = H, R^3 = H, R^4 = OH
6: R^1 = H, R^2 = OH, R^3 = H, R^4 = OH

Figure 1.1. Anthracycline analogues and metabolites.

dopamine and L-DOPA prodrugs into the brain to treat Parkinson's disease[27, 28], sugar conjugation can greatly enhance drug solubility (up to >2 orders of magnitude), enhance uptake *in vitro* (albeit mostly via GLUT1-independent mechanisms) and present analogues (*e.g.*, L-DOPA glucose/galactose-based conjugates (**7a-b**); **Figure 1.2**) with improved efficacy in preclinical animal models[29, 30]. In addition, non-hydrolyzable linear carbohydrate conjugates (such as IPX-750 (**8**)) also present improved properties including solubility, oral bioavailability, efficacy in preclinical animal models and BBB penetration (albeit via GLUT1-independent processes)[31-33]. In a similar manner, sugar conjugation has been employed to improve BBB penetration of neuropeptides, the non-selective mechanism of which was proposed to derive through sugar-dependent alterations in amphipathcity of the corresponding conjugates (referred to as the 'biousian' hypothesis of membrane hopping)[34, 35].

The taxanes (e.g., Taxol® (**9**); **Figure 1.3**) are anticancer drugs that prevent disassembly of microtubules by binding to sites on polymeric β-tubulin[36-38]. The main obstacles in development of taxanes proved to be limited availability from natural sources, formulation problems predominantly resulting from their extremely low solubility and, to a lesser extent, cancer selectivity[39, 40]. Glycosylation has proven to be a successful strategy to improve taxane solubility, inspired in part by the discovery of naturally occurring glycosides such as 7β-xylosyl-10 deacetyltaxol (**10**)[41]. Among the many taxane glycoconjugate prodrugs synthesized to date, dramatic improvements in solubility have been achieved (*e.g.*, >2500-fold over the parent aglycon for the mannose-based conjugate **11**) without drastic reductions in potency[42]. Such sugar glycoconjugates are also advantageous in the context of the well-established facilitative sugar transporter (GLUT) overexpression observed in tumors[43-45]. As a representative example of GLUT-mediated tumor targeting of taxane glycoconjugates, Chen and collaborators revealed



7a: R^1 = OH, R^2 = H
7b: R^1 = H, R^2 = OH

8

Figure 1.2. Examples of compounds with enhanced solubility and membrane transport.

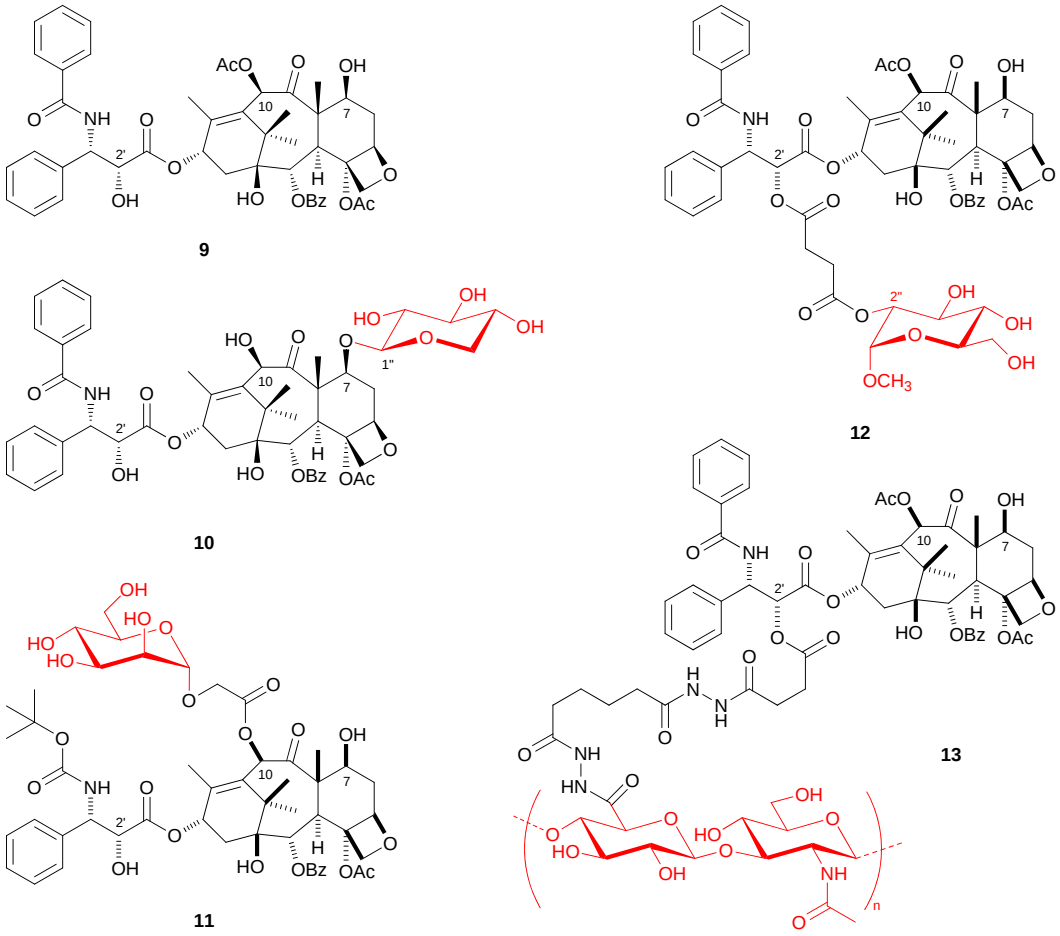


Figure 1.3. Taxol and various glycosylated analogues.

improvements in specificity of the methyl glucoside (**12**) for cancer cell lines versus normal epithelial cell lines that ranged from >19.5- to 1750-fold compared to the parent drug[46, 47]. This analogue was also greater than 2-orders of magnitude more soluble than paclitaxel. Prestwick and colleagues were the first to apply hyaluronan (HA) conjugation (to enable tumor-targeting via CD44-mediated endocytosis)[48, 49] to paclitaxel via the generation of amide-linked HA-conjugates (*e.g.*, **13**)[50, 51]. This pioneering study revealed paclitaxel solubility, potency and tumor cell line specificity could be enhanced via HA-conjugation and these factors were dependent upon the extent of paclitaxel loading. In summary, the examples highlighted within this section support the contention that glycodiversification can be used to improve a drug's solubility and selective (or non-selective) uptake into cells/organs of interest.

1.2.3. The influence of glycosylation upon drug pharmacodynamics and mechanism

While the sugars appended to naturally-occurring glycosylated natural products are commonly key to the biological activity of the parent structure, the impact of glycosylation upon the activity of non-glycosylated natural products and/or drugs is difficult to predict *a priori*. A recent study to examine the potential of differential glycosylation of a non-glycosylated natural product focused upon the plant alkaloid colchicine (**14**) produced by *Colchicum autumnale* (also known as 'meadow saffron'). Colchicine inhibits tubulin polymerization, causing metaphasic mitotic arrest, which leads to rapid cell death[52-55]. Toxicity limits its clinical use to the treatment of severe inflammatory episodes of gout (with Colcrys®), familial Mediterranean fever, and Behcet's disease[56, 57]. Only a few colchicine 2-demethyl-2- or 3-demethyl-3-*O*-glycosides have been reported[58-61], and thus, the effects of glycosylation upon this natural product remain largely unknown. Using a chemoselective neoglycosylation reaction between a free reducing

sugar and a methoxyamine-appended aglycon, Thorson and co-workers conducted the rapid synthesis of a 58-member differentially glycosylated colchicine library (see **Figure 1.4** for representative members)[62]. Cytotoxicity screens and subsequent microtubulin-based mechanistic assays revealed neoglycosylation to modulate the mechanism and potency of colchicine (**14**). Notably, the attachment of D-glucorono-3-6-lactonide (to provide **15a**) provided a less toxic microtubulin destablizer that dramatically synergized with taxol, the attachment of 2-deoxy-L-ribose or 2-deoxy-D-ribose (to provide **15b** or **15c**, respectively) gave analogues that were equipotent and functionally identical to taxol (a microtubulin stabilizer), while the attachment of 7-deoxy-D-galacto-heptopyranos-6-ulose or D-digitoxose (to provide **15d** or **15e**, respectively) led to analogues which, while relatively potent, did not influence tubulin polymerization. This study clearly highlights the potential for glycodiversification to alter the mechanism of a non-glycosylated parent drug and this contention is supported by more recent studies of glycodiversified libraries based upon betulinic acid[63], warfarin[64] and podophyllotoxin[65].

Naturally-occurring high affinity ligands for the Na^+,K^+-ATPase α subunit have been used for centuries to treat congestive heart failure where they function as ionotropic agents[66, 67]. In contrast, the binding of steroidal glycosides to the α subunit of the cancer Na^+,K^+-ATPase 'signalosome' releases Src and initiates kinase-dependent events involving EGFR, MAPK, PKC or TGF-beta signaling pathways that stimulate cell apopotosis[68-72]. The balance between ionotropic versus apoptotic effects of steroidal glycosides are influenced by key pharmacophores within steroidal glycoside structure notably including the appended carbohydrates. Based upon this precedent, Thorson and co-workers generated a glycorandomized library with 78 members based upon the cardenolide digitoxin (**16**) (from the foxglove plant, *Digitalis lanata*) from which

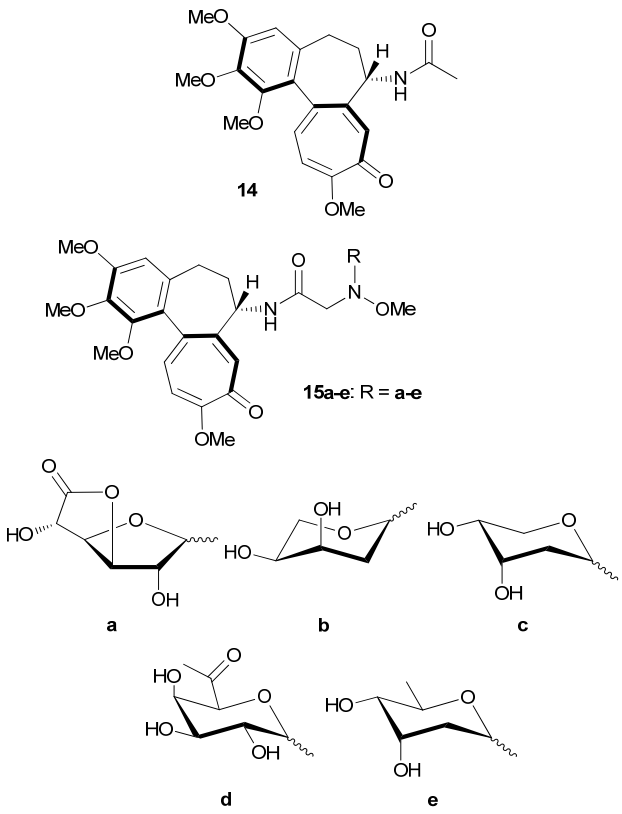


Figure 1.4. Colchicine and chemically glycosylated analogues.

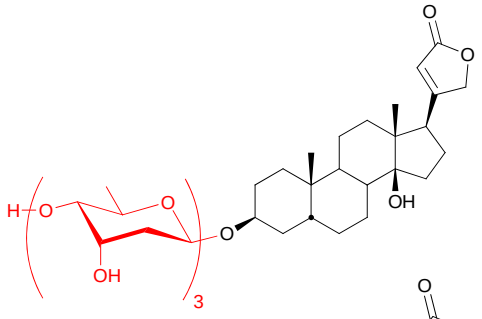
compounds were identified (*e.g.*, **17**) with enhanced anticancer efficacy (>50-fold) and reduced ionotropic effects (>4-fold reduction) (**Figure 1.5**)[73]. Inspired by this preliminary study, additional studies have focused upon extending modifications of the sugar, steroid and alkoxyamine[74-77]. Cumulatively, the examples presented in this section clearly highlight the potential of drug glycodiversification to expedite the discovery of novel leads with unique mechanisms of action and/or enhancements in desired potency.

1.3. A summary of glycosidic bond-forming enzymes

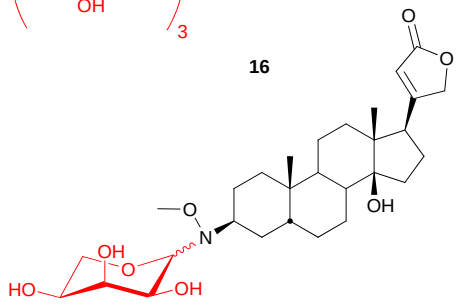
At the core of any enzymatic glycodiversification strategy is a glycosidic-bond forming enzyme (**Figure 1.6**). A number of factors must be considered in selecting a suitable catalyst, most notably: ***i)*** practical access to the corresponding substrates (and catalyst), ***ii)*** ability of the catalyst to act upon a broad range of substrates, and ***iii)*** efficiency of glycosidic bond formation. This section describes a variety of enzyme classes capable of glycosidic bond formation and, although not all enzyme classes discussed here have been specifically applied in the context of glycodiversification, we hope their inclusion herein may serve to inspire potential future applications and inquiry.

1.3.1. Glycosidases and transglycosylation

In contrast to the small molecule/drug glycosylation focus of this review, the bulk of the work with glycosidases to date has focused upon complex saccharides, much of which has been highlighted in a number of excellent recent reviews[78-81]. Glycosidase-catalyzed transglycosylation reactions mediate the transfer of sugars from unprotected di-, tri-, oligo-saccharide-based donors or glycosides bearing an appropriate leaving group at the anomeric



16



17

Figure 1.5. Glycosylated cardenolides.

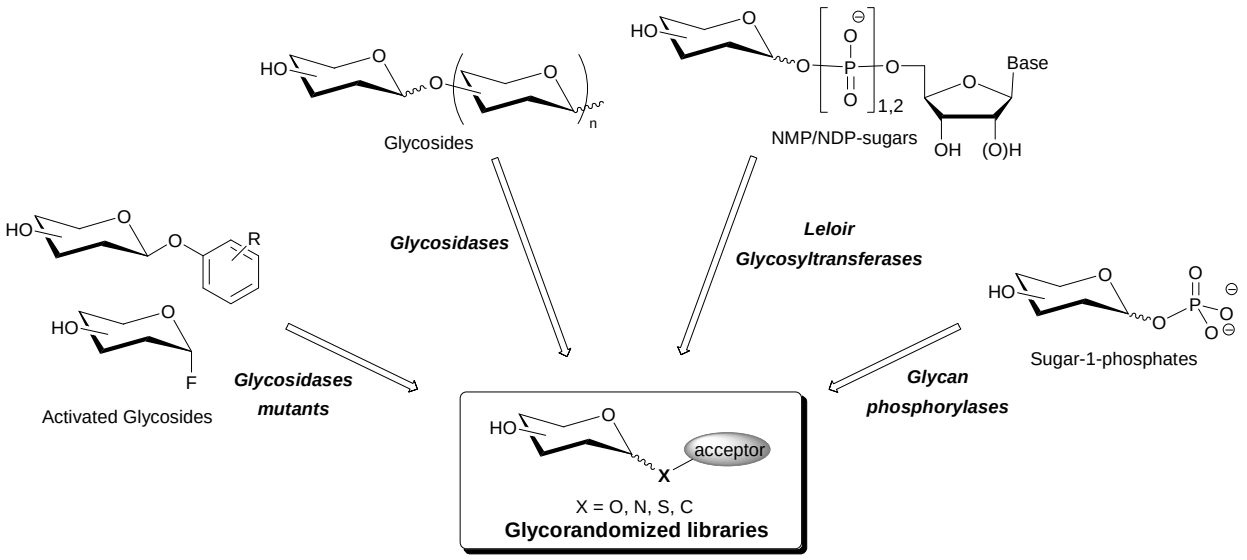


Figure 1.6. General routes to glycorandomized libraries.

position to targeted alycon acceptors. While glycosidases are able to generate very regio- and stereo-specific glycosidic bonds, desired product formation is constrained by both thermodynamic parameters and competition from corresponding hydrolytic reactions[82]. Therefore, synthetic transglycosylation reactions using wild-type glycosidases are generally conducted under tight kinetic control to maximize reaction yields. Although reports of transglycosylation of small molecules continue to be reported with wild-type glycosidases[83-86], such reactions remain under-utilized in glycorandomization efforts due to their overall limited recognition of drug-like or natural product-based small molecule acceptors[87]. To address the fundamental limitations of using glycosidases in the context of synthesis, several classes of modified glycosidases (glycosynthases, thioglycoligases, thioglycosynthases and *O*-glycoligases) have been developed. Common among these glycosidase variants is the elimination of the key active-site nucleophile (exclusively aspartic or glutamic acid) to prevent the parental hydrolytic activity. To further favor transglycosylation, these modified catalysts have been combined with the use of various 'activated' synthetic glycosyl donors (*e.g.*, glycosyl halides or aromatic glycosides) to enable efficient glycosidase-catalyzed *O*- and *S*-glycosidic bond formation. The next few paragraphs highlight key modified glycosidases and applications relevant to glycorandomization.

First introduced by the Withers group[88], glycosynthases, derived from both retaining and inverting glycosidases[80], have emerged as one of the most widely used of the modified glycosidases. They catalyze formation of glycosidic bonds primarily, but not exclusively, through the use of glycosyl fluoride donors[89]. Derived from glycosynthases, thioglycosynthases also function in a similar manner but an additional active mutation favors *S*-glycosidic bond formation[90]. While glycosynthase studies have largely focused upon glycosidic bond formation

between saccharides, glycosynthases capable of glycosylation of non-saccharides have also been reported. For example, an *Escherichia coli* glucuronylsynthase capable of transferring glucuronic acid to 15 small molecules containing primary, secondary and aryl alcohols, including the steroid hormone dehydroepiandrosterone[91], has been reported. The Cel7B glycosynthase mutant E197S from *Humicola insolens* also glycosylated several flavonoids[89]. Additionally, variant glycosynthases capable of generating highly complex glycosphingolipids on the milligram scale were identified through directed evolution[92] and glycosynthases with expanded glycosyl donor promiscuity are known[93]. These examples, coupled with the numerous screening methods available for glycosynthase activity[80, 92, 94-96], hold notable promise for the discovery of future glycosynthase variants amenable to glycorandomization.

Thioglycoligases are enzymes derived from retaining glycosidases and engineered for the synthesis of *S*-glycosidic bond formation. A distinguishing feature of thioglycoligases catalysis is the formation of a transient covalent glycosyl-enzyme intermediate in the absence of an acid/base catalyst (exclusively aspartic or glutamic acid mutated to an amino acid incapable of acid/base catalysis; *e.g.*, alanine). Acceptor moieties containing a thiosugar then perform a nucleophilic attack on the glycosyl-enzyme complex, catalyzing stereospecific formation of an *S*-glycosidic bond containing product[97]. The recently described *O*-glycoligases function in a similar manner, but do not require a thiol-bearing acceptor[98]. While both α-[99] and β-glycosidases[97, 100-104] have been successfully converted into thioglycoligases, all reported examples of thioglycoligase and *O*-glycoligase activity to date have focused upon saccharide synthesis. However, given the precedent for glycosidase-catalyzed modification of small molecules described in the previous paragraph, thioglycoligase and *O*-glycoligase enzymes are anticipated to be equally amenable for drug or small molecule modification.

1.3.2. Glycan phosphorylases

Glycan phosphorylases catalyze formation of sugar-1-phosphates via saccharide phosphorylysis. While glycan phosphorylases play a predominantly catabolic role *in vivo*, these catalysts can be exploited *in vitro* for regio- and stereoselective glycosidic bond formation. Compared to glycosidase-catalyzed transglycosylation, a notable advantage of glycan phosphorylase-based glycosidic-bond formations are the minimal competing side reactions, such as hydrolysis (in noted cases, reported to be nearly two orders of magnitude slower than the phosphorolytic reaction)[105, 106]. While recent reports of glycan phosphorylase-catalyzed glycosylation have largely been limited to the formation of specific saccharide products[107-111], transglycosylation to alternative acceptors[112], including aromatic alcohols[113-116], have been reported. Additionally, recent precedent for engineering glycan phosphorylase specificity includes the use of an engineered α-(1,4)-maltose phosphorylase from *Lactobacillus acidophilus* to produce both α-(1,2)-kojibiose and (α1,α1)-trehalose[117]. While future glycan phosphorylase engineering efforts are anticipated to continue the expansion of the substrate scope for this unique enzyme class, access to large numbers and quantities of the sugar-1-phosphate substrates necessary for glycorandomization efforts still remains challenging (see section **1.4.3.1**).

1.3.3. Glycosyltransferases

In comparison to the enzyme classes described above, glycosyltransferases (GTs; EC 2.4.x.y) accommodate a much larger array of substrate structural diversity and, as a result, remain the catalyst of choice for glycorandomization. GTs constitute one of the most diverse enzyme families with more than 65,000 confirmed and putative GT sequences classified in the

Carbohydrate Active enZyme database (CAZy, http://www.cazy.org/). GTs catalyze the formation of *O*-, *N*-, *S*- and *C*-glycosidic bonds by the attachment of a sugar moiety to a range of acceptors (including small molecules, peptides, proteins, glycons and lipids) with exquisite regio- and stereocontrol (*i.e.*, retention or inversion of anomeric stereochemistry). GTs are involved in many different cellular functions, including but not limited to energy storage, immune response, cellular structure, protein trafficking, and intra- and extra-cellular signaling[118-122]. GTs involved in secondary metabolism (see section **1.5**) often display notable promiscuity for structurally altered donors and/or acceptors. There are many excellent comprehensive reviews on various aspects of GTs (*e.g.*, mechanisms, function, structural biology and engineering)[123-131]. Notably, in the context of this review, the overall broad applicability and potential inherent flexibility of GTs is advantageous when choosing potential catalysts for glycorandomization. Additionally, the recent development of high throughput screens applicable to GT-catalyzed reactions has enabled the discovery of more permissive catalysts that may further strengthen GT-based glycorandomization[132-139].

The majority of characterized GTs are Leloir-type (sugar nucleotide-dependent) and utilize either nucleotide di- or monophosphate sugars as donors wherein the nucleotide phosphate component serves both as a recognition element for the GT and a leaving group for the glycotransfer reaction[123]. Compared to their sugar nucleotide-dependent counterparts, the less prevalent non-Leloir GTs use alternative sugar donors such as sucrose, glucose-1-phosphate or polyprenyl phosphate sugars and generally display a higher degree of substrate specificity[140]. The dependence upon sugar nucleotides, a substrate class with limited (and typically expensive) commercial availability which is also often difficult to access synthetically[141, 142], remains one of the main limitations of GT-catalyzed glycorandomization. Nevertheless, as described in the

next section this limitation is being addressed through a combination of engineering enzymes, pathways and substrates in an effort to provide a streamlined GT-based platform for the glycorandomization of drug-like and natural product-based small molecules.

1.4. Enzymatic strategies for GT donor (sugar nucleotide) production

1.4.1. Overview of NDP-sugars

While an enzymatic glycorandomization platform is advantageous for a variety of reasons (efficiency, regio- and stereoselectivity, no protection/deprotection chemistries required, amenable to *in vivo* strain development), NDP-sugar accessibility is critical to the success of any Leloir GT-catalyzed approach. Specifically, GT-catalyzed glycorandomization combines the inherent or engineered substrate flexibility of GTs with NDP-sugar donor libraries and target natural product/drug acceptors to enable the rapid generation of differentially glycosylated product arrays (**Figure 1.7**). While NDP-sugars can be chemically synthesized (see a recent review by Wagner and co-workers[142]), the many practical challenges imposed via the chemical syntheses of sugar nucleotide libraries compared to the numerous potential advantages of complementary general enzymatic strategies has promoted the development of the latter in the context of glycorandomization. The following section highlights the state of the art for the enzymatic generation of NDP-sugars as a core component of GT-catalyzed glycorandomization and focuses specifically upon three main strategies - ***i)*** native multi-enzyme pathways, ***ii)*** engineered permissive two enzyme pathways, and ***iii)*** altering the equilibrium of GT-catalyzed reactions – for the practical syntheses of novel sugar nucleotide libraries.

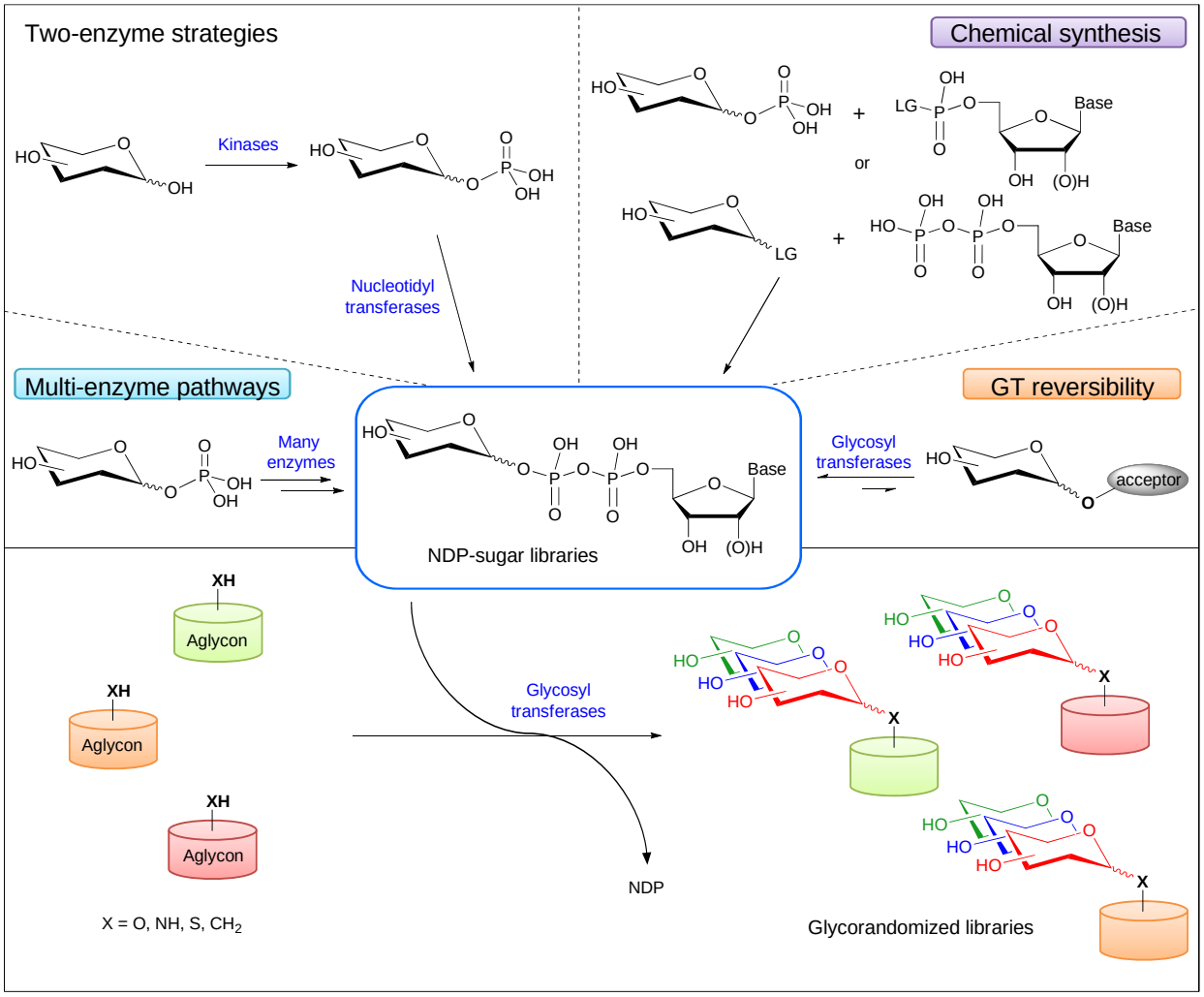


Figure 1.7. General approaches to NDP-sugar library synthesis for glycorandomization.

1.4.2. Multi-enzyme pathways

Inspired by the proof of concept of metabolic pathway engineering in the 1980's[143], the *in vivo* reconstruction of either native or designed multi-enzyme NDP-sugar pathways has successfully enabled the production of many differentially-glycosylated natural product analogues (**Figure 1.7**; see also section **1.4.5**). In work by Salas and co-workers, heterologous plasmids encoding for the biosynthesis of a variety of novel sugar nucleotides have been incorporated into a series of secondary metabolite-producing organisms to provide differentially-glycosylated derivatives of the corresponding target secondary metabolites[144-150]. Heterologous expression of sugar nucleotide pathways have also been applied in bioconversion as exemplified by the work of McDaniel and co-workers for desosaminylation of an array of modified macroloactones using an engineered *Streptomyces lividans* host which carried the genes encoding for the biosynthesis of TDP-desosamine and the corresponding macrolide desosaminyltransferase[151]. *In vitro* reconstruction of sugar nucleotide pathways has largely focused upon TDP-sugar analogues, most of which diverge from the common progenitor TDP-4-keto-6-deoxy-α-D-glucose[121, 141, 152]. For example, TDP-L-epivancosamine from the chloroeremomycin pathway was synthesized in a sequential series of *in vitro* reactions with 5 separate enzymes to produce the final product[153]. Takahashi *et al.* reported a two stage, one pot enzymatic synthesis of TDP-L-mycarose (from the tylosin pathway) utilizing 11 separate enzymes with a final yield of 16%[154]. Zhang *et al.* used purified enzymes from both the tylosin and kijanimicin biosynthetic clusters in a series of sequential reactions for synthesis of TDP-L-digitoxose[155]. Synthesis of TDP-L-forosamine was also demonstrated with a series of 5 purified enzymes from the spinosyn biosynthetic cluster[156]. Most recently, a one-pot enzymatic

synthesis of TDP-D-ravidosamine combining 7 enzymes from 3 organisms was reported[157]. Although these examples are limited when compared to the total number of NDP-sugars postulated to exist *in vivo*, they show a clear demonstration that enzymes can be readily combined both *in vivo* or *in vitro* to access highly novel NDP-sugars and highlight the importance of continued efforts to study the biosynthesis of novel sugar nucleotides[121].

1.4.3. Permissive two-enzyme strategies

Nature commonly employs more than two enzymes to generate sugar nucleotides, typically initiated via a two-step hexokinase (for sugar C-6 phosphorylation) - phosphoglucomutase (for interconversion of C-6/C-1 phosphoryl substitutions) series of reactions in primary metabolism[158]. However, the shortest feasible natural enzymatic route to a NDP-sugar can be reduced to two steps: kinase-catalyzed anomeric phosphorylation of a free unprotected reducing sugar to give the corresponding sugar-1-phosphate and subsequent coupling to a nucleotidyl monophosphate moiety by a nucleotidyltransferase (**Figure 1.7**). Highlighted herein are recent efforts to discover and engineer permissive sugar kinases and sugar-1-phosphate nucleotidyltransferases to enable the rapid production of NDP-sugar collections.

1.4.3.1. Enzyme 1 - anomeric kinases

Naturally-occurring anomeric kinases are surprisingly rare. For example, the reported hexose C-1 kinases are limited to phosphorylases[159], D-glucuronic acid kinase[160], D-galacturonic acid kinase[161], D-galactokinase[162, 163], *N*-acetyl-D-hexosamine kinase[164, 165], and L-fucokinases[166]. Distinct from phosphorylases which utilize disaccharides (*e.g.*, sucrose,

maltose, trehalose) or oligosaccharides (*e.g.*, glycogen, laminaritriose, β-1,3-glucan) for D-glucose-1-phosphate production, anomeric kinases catalyze the transfer of the γ-phosphate of a NTP (usually ATP) to the anomeric position of a sugar to provide sugar-1-phosphate and NDP as products. As discussed herein, a number of representative anomeric kinases have been studied in depth, demonstrating 59 distinct sugars (**18-76**) as substrates (**Figure 1.8** and **Table 1.1**), and it is also important to note that structures for phosphorylases (*e.g.*, glycogen, sucrose, maltose, cellobiose phosphorylases), galactokinases and *N*-acetyl-D-galactosamine kinases have been reported[162, 164, 174].

Among the anomeric kinases highlighted above, galactokinase is a catalyst commonly used for the primary metabolism of galactose and has been characterized from a variety of sources including bacteria[162, 168, 175], yeast[176-178], plants[179, 180], and mammals[181, 182]. Of these, the native *E. coli* enzyme was found to be among the most permissive with preliminary studies revealing tolerance to various substitutions at C-2, C-3 and C-6 of the sugar substrate (**Figure 1.8** and **Table 1.1**) [183]. This preliminary study also revealed the first high throughput anomeric kinase colorimetric assay which enabled a subsequent pioneering directed evolution study that led to the discovery of a more permissive GalK variant (Y371H) that included among its expanded substrate repertoire two L-sugars[183, 184]. While the early *E. coli* GalKs retained stringent D-galacto C-4 specificity, substrate specificity studies with *L. lactis* GalK surprisingly revealed tolerance toward D-gluco-configured substrates[170]. Engineering of the *E. coli* GalK based upon a *L. lactis* homology model led to the *E. coli* variant GalK M173L which displayed a particular promiscuity on C-4 and C-6[162, 167]. When combined with the original mutation discovered via directed evolution, the GalK M173L/Y371H double mutant presented even further enhancements in substrate flexibility over predecesors[167]. Specifically, the substrate set

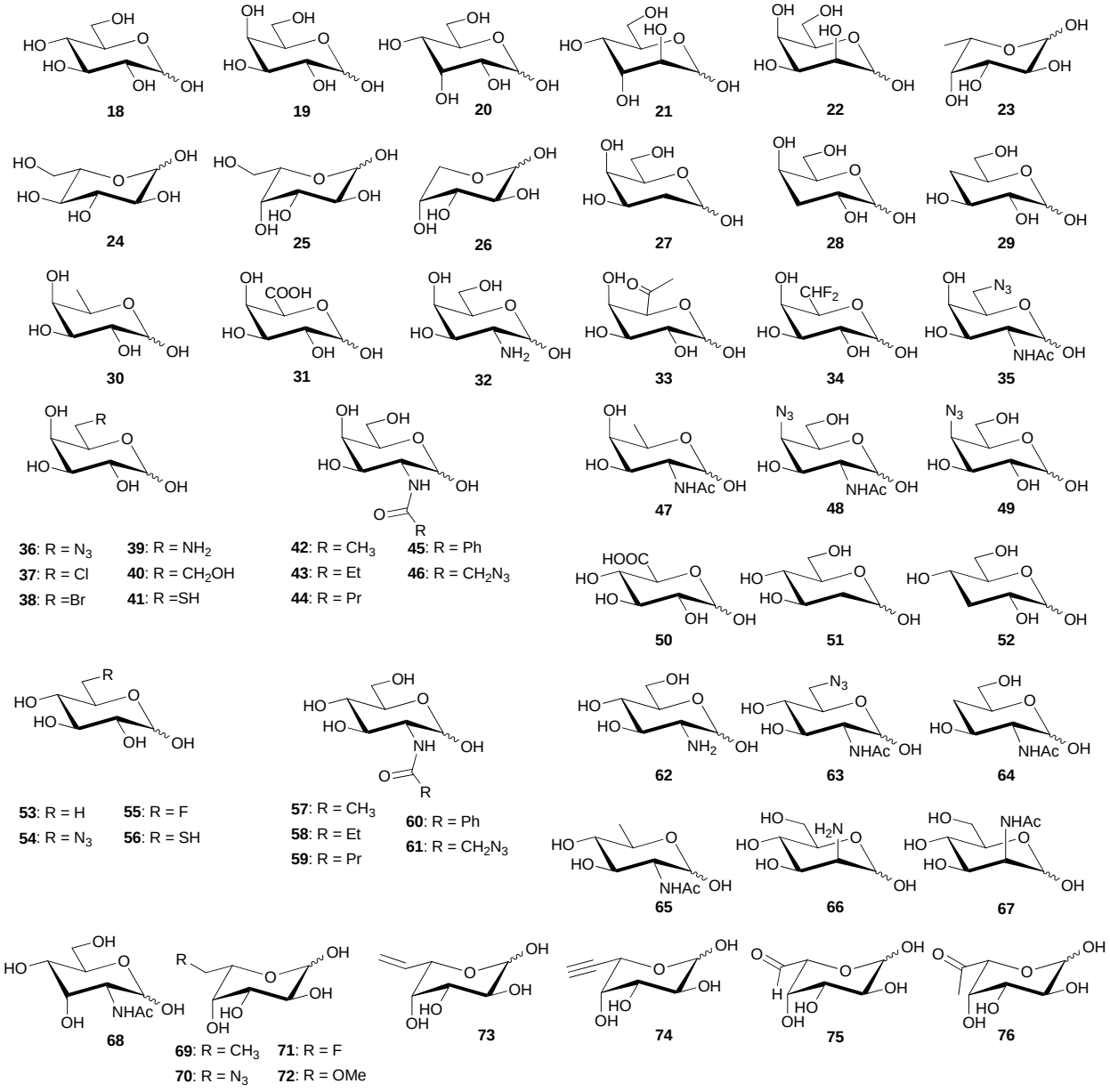


Figure 1.8. Sugars recognized by various sugar-1-kinases (see also **Table 1.1**).

Substrate	Enzyme	Conversion (%)[#]	Ref
D-glucose (**18**)	GalK (*E.coli*, M173L/Y371H)	42	167
D-galactose (**19**)	GalK (*E.coli*, Y223F)	89	168
D-allose (**20**)	GalK (*E.coli*, M173L/Y371H)	15	167
L-altrose (**21**)	GalK (*E.coli*, Y371H)	78	167
D-talose (**22**)	GalK (*E.coli*, Y223F)	59	168
L-fucose (**23**)	Fkp	92[a]	169
L-glucose (**24**)	GalK (*E.coli*, Y371H)	68	167
L-galactose (**25**)	Fkp	77[a]	169
D-arabinose (**26**)	Fkp	85[a]	169
2-deoxy-D-galactose (**27**)	GalK (*E.coli*, wt)	80	168
3-deoxy-D-galactose (**28**)	GalK (*L.lactis*, Y385H)	65	170
4-deoxy-D-galactose (**29**)	GalK (E.coli, Y223F)	55	168
6-deoxy-D-galactose (**30**)	GalK (*E.coli*, wt)	27	168
D-galacturonic acid (**31**)	GalK (*E.coli*, Y371H)	35	167
D-galactosamine (**32**)	GalK (*E.coli*, wt)	73	168
7-deoxy-D-galacto-heptopyranos-6-ulose (**33**)	GalK (*E.coli*, Y371H)	13	170
6-deoxy-6,6-difluoro-D-galactose (**34**)	GalK (*E.coli*, Y371H)	26	167
N-acetyl-6-azido-6-deoxy-D-galactosamine (**35**)	NahK	42[b]	171
6-azido-6-deoxy-D-galactose (**36**)	GalK (*E.coli*, Y371H)	44	167
6-chloro-6-deoxy-D-galactose (**37**)	GalK (*E.coli*, M173L/Y371H)	12	167
6-bromo-6-deoxy-D-galactose (**38**)	GalK (*E.coli*, M173L/Y371H)	14	167
6-amino-6-deoxy-D-galactose (**39**)	GalK (*E.coli*, Y371H)	15	170
6-hydroxymethylene-D-galactose (**40**)	GalK (*E.coli*, M173L)	15	167
6-deoxy-6-thio-D-galactose (**41**)	GalK (*E.coli*, M173L/Y371H)	25	167
N-acetyl-D-galactosamine (**42**)	NahK	78[b]	172
N-propyryl-D-galactosamine (**43**)	NahK	85[b]	171
N-butyryl-D-galactosamine (**44**)	NahK	86[b]	171
N-benzoyl-D-galactosamine (**45**)	NahK	77[b]	171
N-azidoacetyl-D-galactosamine (**46**)	NahK	65[b]	171
N-acetyl-6-deoxy-D-galactosamine (**47**)	NahK	37[b]	171
N-acetyl-4-azido-4-deoxy-D-galactosamine (**48**)	NahK	73[b]	171
4-azido-4-deoxy-D-galactose (**49**)	GalK (*E.coli*, M173L/Y371H)	10	167
D-glucuronic acid (**50**)	GalK (*E.coli*, M173L/Y371H)	9	167
2-deoxy-D-glucose (**51**)	GalK (*E.coli*, M173L/Y371H)	18	167
3-deoxy-D-glucose (**52**)	GalK (*E.coli*, M173L/Y371H)	15	167
6-deoxy-D-glucose (**53**)	GalK (*E.coli*, M173L/Y371H)	12	167
6-azido-6-deoxy-D-glucose (**54**)	GalK (*E.coli*, M173L/Y371H)	14	167
6-deoxy-6-fluoro-D-glucose (**55**)	GalK (*E.coli*, M173L/Y371H)	12	167
6-deoxy-6-thio-D-glucose (**56**)	GalK (*E.coli*, M173L/Y371H)	11	167
N-acetyl-D-glucosamine (**57**)	NahK	90[b]	172
N-propyryl-D-glucosamine (**58**)	NahK	86[b]	172
N-butyryl-D-glucosamine (**59**)	NahK	87[b]	172
N-benzoyl-D-glucosamine (**60**)	NahK	88[b]	172
N-azidoacetyl-D-glucosamine (**61**)	NahK	87[b]	172
D-glucosamine (**62**)	GalK (*E.coli*, M173L/Y371H)	10	167
N-acetyl-6-azido-6-deoxy-D-glucosamine (**63**)	NahK	34[b]	172
N-acetyl-4-deoxy-D-glucosamine (**64**)	NahK	70[b]	171
N-acetyl-6-deoxy-D-glucosamine (**65**)	NahK	75[b]	172
D-mannosamine (**66**)	NahK	(+)	165
N-acetyl-D-mannosamine (**67**)	NahK	(+)	165
N-acetyl-D-allosamine (**68**)	NahK	22[b]	172
6,7-dideoxy-L-galacto-heptopyranose (**69**)	Fkp	50[a]	173
6-azido-6-deoxy-L-galactose (**70**)	Fkp	94[a]	169
6-fluoro-6-deoxy-L-galactose (**71**)	Fkp	94[a]	169
6-*O*-methyl-L-galactose (**72**)	Fkp	77[a]	169
6,7-deoxy-L-galacto-hept-6-enopyranose (**73**)	Fkp	89[a]	169
6,7-deoxy-L-galacto-hept-6-ynopyranose (**74**)	Fkp	90[a]	169
L-galacto-hexodialdo-pyranose (**75**)	Fkp	9[a]	173
7-deoxy-L-galacto-heptopyranos-6-ulose (**76**)	Fkp	15[a]	173

Table 1.1. Sugar substrates converted into sugar-1-phosphates by C-1 kinases and mutants. [#]Best conversion reported to date, [a]Conversion to the corresponding GDP-sugar by the bifunctional enzyme, [b]Isolated yield, (+) product detected but conversion not reported.

for the new GalK double mutant encompassed 28 sugars, including three azidosugars and two thiosugars, which set the stage for downstream chemical modification of glycorandomized product via chemoselective strategies (see section **1.5.1.4**)[185-187].

The *N*-acetylhexosamine kinase NahK from *Bifidobacterium longum* was the first native D-gluco anomeric kinase reported[165]. This enzyme was initially found to accept both GlcNAc and GalNAc, as well as nine other monosaccharides with variable levels of efficiency (**Table 1.1** and **Figure 1.8**). Further studies revealed relaxed specificity toward *N*-acyl modifications, C-3 epimerization and C-4 modification (azido and deoxy) but a more limited tolerance toward C-6 modification was observed[171, 172]. Recently, Comstock and co-workers identified for the first time a bifunctional L-fucokinase and GDP-fucose pyrophosphorylase enzyme, designated Fkp, from the mammalian symbiont *Bacteroides fragilis* 9343 (**Table 1.1** and **Figure 1.8**)[188]. A related bifunctional enzyme from *Arabidopsis* possessing a similar activity than Fkp has also shown to efficiently form GDP-L-fucose[189]. Combining the inherent promiscuity of Fkp toward C-5 modified L-fucose analogues and a recombinant α-1,3-fucosyltransferase activity, Wu and co-workers generated a library of Le^x trisaccharides[169] and this catalyst has also been exploited for *in vivo* biotransformation (see section **1.4.5**)[173]. While systematic studies to compare (dis)advantages of such fused bi-functional catalysts to their stand alone counterparts have yet to be pursued, we expect continued catalyst discovery and optimization to fuel future glycorandomization efforts via increasing the availability of novel sugar-1-phosphates.

1.4.3.2. Enzyme 2 - nucleotidyltransferases

Nucleotidyltransferases, also refered to as pyrophosphorylases, are ubiquitous in nature with currently ~15,000 known and putative nucleotidyltransferase sequences in GenBank[190].

Physiologically, these enzymes are the key initiator of most sugar biosynthetic pathways and sugar attachment (*e.g.*, GT-catalyzed) processes where they catalyze the reversible interconversion of sugar-1-phosphates/NTPs and NDP-sugars (usually ADP-, CDP-, UDP-, GDP-, TDP-glucoses, as well as GDP-mannose, GDP-fucose, and UDP-*N*-acetylglucosamine)/pyrophosphate. Nucleotidyltransferases are often allosterically controlled and generally proceed via an ordered bi-bi mechanism (where the NTP binds first and the NDP-sugar is released last) which leads to the formation of a trigonal bipyrimidal phosphoryl ternary complex[191, 192]. While there is great diversity among nucleotidyltransferase members and the chemistries they enable, this section attempts to distill recent efforts specifically directed toward the use nucleotidyltransferases for the generation of sugar nucleotide arrays (as reagents for downstream GT-catalyzed glycorandomization reactions). The reader is referred to **Figure 1.9** and **Table 1.2** for an overview of substrates (**77**-**151**) recognized by nucleotidyltransferases as discussed below.

One of the most permissive and extensively studied nucleotidyltransferase is the α-D-glucose thymidyltransferase from *Salmonella enterica* LT2 (also known as RmlA or E_p), which displayed inherent flexibility toward a range natural and unnatural sugar-1-phosphates and NTPs[193, 194, 211, 213]. The inherent substrate promiscuity of RmlA was further expanded through both structure-based engineering[198, 199] and, aided by the development of the first high throughput screen for nucleotidyltransferase activity[220], directed evolution[201]. Key mutations identified to enhance the permissive nature of RmlA via these various studies include L89T, W224H and T201A (which expanded sugar variability)[198, 199] and Q83D/S (which altered the inherent NTP purine/pyrimidine bias)[194]. RmlA variants now available can utilize more than 40 sugar-1-phosphates ranging from all the epimers of D-glucose to substituted (amino, azido, thiol,

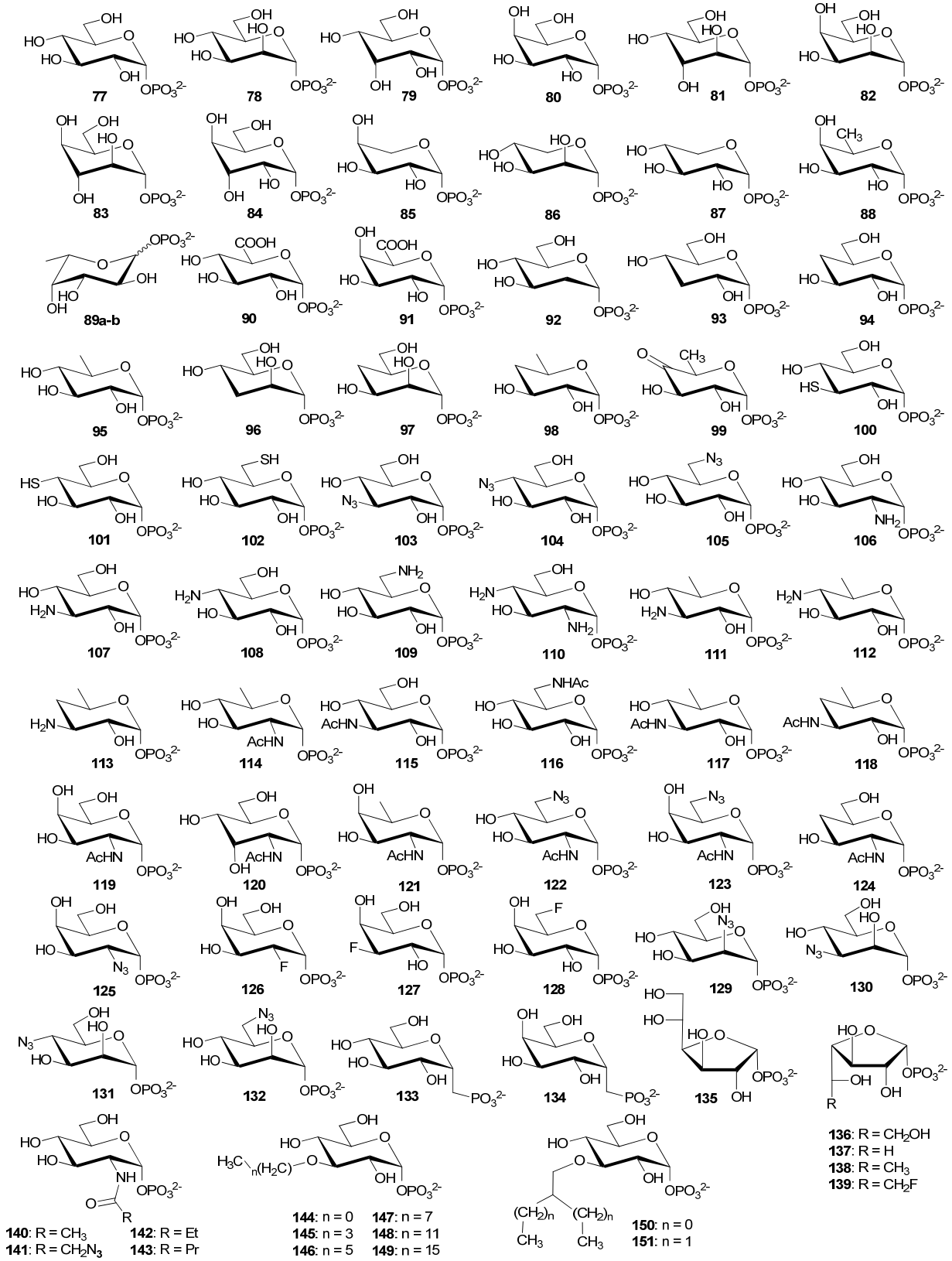


Figure 1.9. Sugar-1-phosphates recognized by various nucleotidyltransferases (see also **Table 1.2**).

Substrate	Conversion (%)[a]								Ref
	TTP	UTP	GTP	CTP	ATP	dGTP	dCTP	dATP	
α-D-glucose-1-phosphate (**77**)	99 [b]	99 [b]	>75 [m]	17 [g]	28 [g]	(+) [b]	(+) [b]	(+) [b]	193-197
α-D-mannose-1-phosphate (**78**)	99 [b]	>85 [m]	>85 [m]	>85 [m]	>85 [m]	(+) [b]	(+) [b]	(+) [b]	193, 194, 197
α-D-allose-1-phosphate (**79**)	53 [d]	37 [d]	n.d	n.d	n.d	n.d	n.d	n.d	198, 199
α-D-galactose-1-phosphate (**80**)	99 [l]	33 [c]	30 [m]	(-)	(+) [b]	(+) [b]	(+) [b]	(+) [b]	193, 194, 197, 200, 201
α-D-altrose-1-phosphate (**81**)	99 [d]	30 [d]	n.d	n.d	n.d	n.d	n.d	n.d	198
α-D-talose-1-phosphate (**82**)	99 [d]	78 [d]	n.d	n.d	n.d	n.d	n.d	n.d	198
α-D-idose-1-phosphate (**83**)	29 [d]	(-) [b]	n.d	n.d	n.d	n.d	n.d	n.d	198
α-D-gulose-1-phosphate (**84**)	(+) [b]	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	202, 203
β-L-arabinose-1-phosphate (**85**)	n.d	47 [k]	n.d	n.d	n.d	n.d	n.d	n.d	204
α-D-lyxose-1-phosphate (**86**)	n.d	n.d	(+) [n]	n.d	n.d	n.d	n.d	n.d	205
α-D-xylose-1-phosphate (**87**)	93 [c]	89 [k]	(+) [o]	n.d	n.d	n.d	n.d	n.d	18, 204, 206
α-D-fucose-1-phosphate (**88**)	n.d	57 [k]	n.d	n.d	n.d	n.d	n.d	n.d	204
β-L-fucose-1-phosphate (**89a**)	10 [l]	12 [k]	37 [m]	n.d	n.d	n.d	n.d	n.d	197, 204, 207
α-L-fucose-1-phosphate (**89b**)	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	201
α-D-glucuronic acid-1-phosphate (**90**)	88 [f]	(+) [h]	n.d	n.d	n.d	n.d	n.d	n.d	199, 208
α-D-galacturonic acid-1-phosphate (**91**)	n.d	(+) [i]	n.d	n.d	n.d	n.d	n.d	n.d	209
2-deoxy-α-D-glucose-1-phosphate (**92**)	26 [b]	22 [b]	29 [n]	n.d	n.d	n.d	n.d	n.d	193, 210
3-deoxy-α-D-glucose-1-phosphate (**93**)	96 [b]	7 [b]	(+) [o]	n.d	n.d	n.d	n.d	n.d	193, 206
4-deoxy-α-D-glucose-1-phosphate (**94**)	98 [b]	99 [b]	(+) [o]	n.d	n.d	n.d	n.d	n.d	193, 206
6-deoxy-α-D-glucose-1-phosphate (**95**)	98 [b]	99 [b]	(+) [o]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	193, 194, 206
3-deoxy-α-D-mannose-1-phosphate (**96**)	n.d	n.d	(+) [n]	n.d	n.d	n.d	n.d	n.d	205
4-deoxy-α-D-mannose-1-phosphate (**97**)	n.d	n.d	(+) [n]	n.d	n.d	n.d	n.d	n.d	205
4,6-dideoxy-α-D-glucose-1-phosphate (**98**)	99 [b]	(+) [b]	(+) [o]	n.d	n.d	n.d	n.d	n.d	206, 211, 212
6-deoxy-4-keto-α-D-glucose-1-phosphate (**99**)	(+) [b]	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	202, 203
3-deoxy-3-thio-α-D-glucose-1-phosphate (**100**)	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	202
4-deoxy-4-thio-α-D-glucose-1-phosphate (**101**)	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	202
6-deoxy-6-thio-α-D-glucose-1-phosphate (**102**)	(+) [b]	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	202, 212
3-azido-3-deoxy-α-D-glucose-1-phosphate (**103**)	(+) [b]	(+) [b]	(+) [b]	(-) [b]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	194, 212
4-azido-4-deoxy-α-D-glucose-1-phosphate (**104**)	(+) [b]	(+) [b]	(+) [o]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	194, 206, 212
6-azido-6-deoxy-α-D-glucose-1-phosphate (**105**)	(+) [b]	(+) [b]	(+) [o]	n.d	n.d	n.d	n.d	n.d	202, 206, 212
α-D-glucosamine-1-phosphate (**106**)	99 [b]	99 [b]	82 [m]	(+) [b]	5 [l]	(+) [b]	(+) [b]	(+) [b]	194, 197, 200, 213
3-amino-3-deoxy-α-D-glucose-1-phosphate (**107**)	99 [b]	78 [b]	(+) [o]	n.d	n.d	n.d	n.d	n.d	206, 213
4-amino-4-deoxy-α-D-glucose-1-phosphate (**108**)	99 [b]	90 [b]	n.d	n.d	n.d	n.d	n.d	n.d	213
6-amino-6-deoxy-α-D-glucose-1-phosphate (**109**)	99 [b]	15 [b]	(+) [o]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	194, 206, 213
4-amino-2,4-dideoxy-α-D-glucosamine-1-phosphate (**110**)	n.d	n.d	(+) [o]	n.d	n.d	n.d	n.d	n.d	206
3-amino-3,6-dideoxy-α-D-glucose-1-phosphate (**111**)	99 [b]	(+) [b]	n.d	n.d	n.d	n.d	n.d	n.d	211, 212
4-amino-4,6-dideoxy-α-D-glucose-1-phosphate (**112**)	99 [b]	15 [b]	(+) [o]	n.d	n.d	n.d	n.d	n.d	206, 213
3-amino-3,4,6-trideoxy-α-D-glucose-1-phosphate (**113**)	10 [b]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	211
N-acetyl-2,6-dideoxy-α-D-glucosamine-1-phosphate (**114**)	n.d	50 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
3-acetamido-3-deoxy-α-D-glucose-1-phosphate (**115**)	70 [b]	3 [b]	n.d	n.d	n.d	n.d	n.d	n.d	213
6-acetamido-6-deoxy-α-D-glucose-1-phosphate (**116**)	37 [f]	20 [b]	n.d	n.d	n.d	n.d	n.d	n.d	199, 213
3-acetamido-3,6-dideoxy-α-D-glucose-1-phosphate (**117**)	99 [e]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	211
3-acetamido-3,4,6-trideoxy-α-D-glucose-1-phosphate (**118**)	(+) [c]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	202
N-acetyl-α-D-galactosamine-1-phosphate (**119**)	(+) [b]	65 [j]	n.d	n.d	n.d	n.d	n.d	n.d	203, 214
N-acetyl-α-D-allosamine-1-phosphate (**120**)	n.d	20 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
N-acetyl-2,6-dideoxy-α-D-galactosamine-1-phosphate (**121**)	n.d	55 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
N-acetyl-6-azido,2,6-dideoxy-α-D-glucosamine-1-phosphate (**122**)	n.d	20 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214

Table 1.2. Sugar-1-phosphates substrates of nucleotidyltransferases and their conversions into corresponding NDP-sugars.

Substrate	Conversion (%)[a]								Ref
	TTP	UTP	GTP	CTP	ATP	dGTP	dCTP	dATP	
N-acetyl-2,4-dideoxy-α-D-glucosamine-1-phosphate (**124**)	n.d	59 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
2-azido-2-deoxy-α-D-galactose-1-phosphate (**125**)	n.d	89 [k]	n.d	n.d	n.d	n.d	n.d	n.d	204
2-deoxy-2-fluoro-α-D-galactose-1-phosphate (**126**)	n.d	97 [k]	n.d	n.d	n.d	n.d	n.d	n.d	204
3-deoxy-3-fluoro-α-D-galactose-1-phosphate (**127**)	n.d	64 [k]	n.d	n.d	n.d	n.d	n.d	n.d	204
6-deoxy-6-fluoro-α-D-galactose-1-phosphate (**128**)	n.d	37 [k]	n.d	n.d	n.d	n.d	n.d	n.d	204
2-azido-2-deoxy-α-D-mannose-1-phosphate (**129**)	n.d	n.d	41 [n]	n.d	n.d	n.d	n.d	n.d	215
3-azido-3-deoxy-α-D-mannose-1-phosphate (**130**)	n.d	n.d	52 [n]	n.d	n.d	n.d	n.d	n.d	215
4-azido-4-deoxy-α-D-mannose-1-phosphate (**131**)	n.d	n.d	55 [n]	n.d	n.d	n.d	n.d	n.d	215
6-azido-6-deoxy-α-D-mannose-1-phosphate (**132**)	n.d	n.d	63 [n]	n.d	n.d	n.d	n.d	n.d	215
C-(1-Deoxy-α-D-glucopyranosyl) methane phosphonate (**133**)	99 [l]	61 [l]	16 [l]	16 [l]	19 [l]	n.d	n.d	n.d	216
C-(1-Deoxy-α-D-galactopyranosyl) methane phosphonate (**134**)	42 [l]	1 [l]	n.d	n.d	n.d	n.d	n.d	n.d	216
α-D-glucofuranose-1-phosphate (**135**)	7 [l]	n.d	n.d	n.d	n.d	n.d	n.d	n.d	217
α-D-galactofuranose-1-phosphate (**136**)	98 [l]	79 [k]	n.d	n.d	n.d	n.d	n.d	n.d	200, 204
β-L-arabinofuranose-1-phosphate (**137**)	51 [l]	50 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200
α-D-fucofuranose-1-phosphate (**138**)	41 [l]	12 [k]	n.d	n.d	n.d	n.d	n.d	n.d	200, 218
6-deoxy-6-fluoro-α-D-galactofuranose-1-phosphate (**139**)	51 [l]	42 [k]	n.d	n.d	n.d	n.d	n.d	n.d	200, 218
N-acetyl-α-D-glucosamine-1-phosphate (**140**)	99 [e]	47 [b]	58 [m]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	194, 197, 199, 213
N-azidoacetyl-D-glucosamine (**141**)	n.d	44 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
N-propyryl-D-glucosamine (**142**)	n.d	57 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
N-butyryl-D-glucosamine (**143**)	n.d	27 [j]	n.d	n.d	n.d	n.d	n.d	n.d	214
3-*O*-methyl-D-glucose-1-phosphate (**144**)	98 [l]	97 [l]	(+) [o]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	(+) [b]	194, 200, 206, 211
3-*O*-butyl-α-D-glucose-1-phosphate (**145**)	99 [l]	50 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200
3-*O*-hexyl-α-D-glucose-1-phosphate (**146**)	(+) [l]	10 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200, 219
3-*O*-octyl-α-D-glucose-1-phosphate (**147**)	99 [l]	27 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200
3-*O*-dodecyl-α-D-glucose-1-phosphate (**148**)	97 [l]	73 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200
3-*O*-hexadecyl-α-D-glucose-1-phosphate (**149**)	32 [l]	7 [l]	n.d	n.d	n.d	n.d	n.d	n.d	219
3-*O*-(2-methylpropyl)-α-D-glucose-1-phosphate (**150**)	98 [l]	51 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200
3-*O*-(2-ethylbutyl)- α-D-glucose-1-phosphate (**151**)	96 [l]	39 [l]	n.d	n.d	n.d	n.d	n.d	n.d	200

Table 1.2 (continued). Sugar-1-phosphates substrates of nucleotidyltransferases and their conversions into corresponding NDP-sugars. [a] Best conversions with: [b] RmlA wild-type and/or mutants, [c] RmlA wild-type, [d] RmlA L89T, [e] RmlA T201A, [f] RmlA W224H, [g] Tca wild-type, [h] PsUSP wild-type, [i] *Lmj*USP wild-type , [j] GlmU wild-type, [k] Gal-1-PUT wild-type, [l] Cps2L wild-type or mutants, [m] manC wild-type, [n] GDP-ManPP wild-type, [o] RmlA Q83D (+): product detected but conversion not reported, (-): no reaction observed, n.d: not determined.

N-acetyl) and/or deoxy sugars and also the two anomers of L-fucose. Other glucose-1-phosphate thymidyltransferase homologs, like *Streptoccocus pneumonia* Cps2L, have also demonstrated the ability to accept a broad range of sugars, including an L-sugar (β-L-fucose-1-phosphate)[207], a series of alkyl sugar-1-phosphates[219], phosphonate analogues of both glucose and galactose-1-phosphates[216], and most notably, five furanosyl-1-phosphates[217]. Other homologs with base flexibility have also been observed and/or engineered. Within this context, the thermostable thymidyltransferase ST0452 from *Sulfolobus tokodaii* utilized all four deoxynucleotides and UTP[196]. Jakeman and co-workers reported similar activities with *Streptococcus pneumoniae* Cps2L *en route* to the preparation of UDP-, CDP-, ADP-, GDP- and TDP-sugars[207]. Further engineering of the active site of Cps2L augmented the inherent activity toward ribo-configured nucleoside triphosphates[200].

Inherent sugar-1-phosphate and/or NTP promiscuity has also been observed among uridyltransferases from a variety of sources including bacteria[195, 221, 222], plants[208, 223], and parasites[209, 224, 225]. Galactose-1-phosphate uridyltransferase (E.C. 2.7.7.12), responsible for the reversible transfer of the UMP moiety from UDP-Glc as part of cascade of events for UDP-Glc/UDP-Gal interconversion in primary metabolism, has also been demonstrated to accommodate a remarkably wide range of sugar-1-phosphates including the C-2 epimer (D-mannose), C-2, C-3 and C-5 substituted D-galactose, two L-sugars (L-fucose and L-arabinose) and even a set of four furanose 1-phosphates[204, 218]. Reminiscent of the bifunctional anomeric fucokinase/fucose-1-phosphate guanidyltransferase described in the section **1.4.3.1**, other nucleotidyltransferase-based bifunctional catalysts have been reported. For example, the archaeal hyperthermophile *Pyrococcus furiosus* glucose-1-phosphate uridyltransferase was demonstrated to catalyze both sugar *N*-acylation and subsequent nucleotidyltransfer to enable the production of

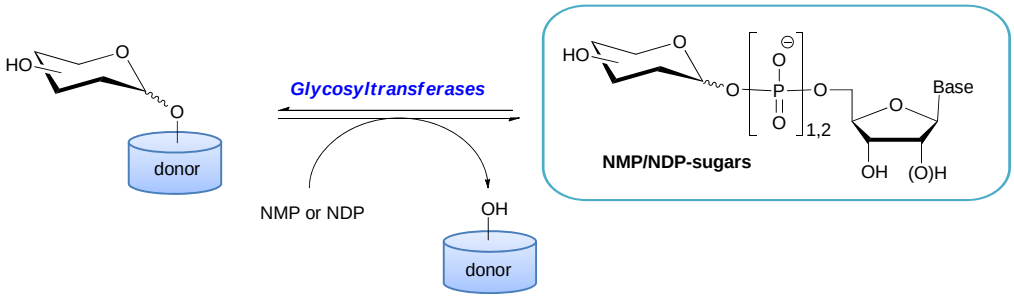
a small set of *N*-acyl glucosamine nucleotides[226]. GDP-sugars were similarly obtained with a bifunctional phosphomannose isomerase/mannose-1-phosphate guanidyltransferase from *P. furiosus* which accommodated a range of sugar-1-phosphates and nucleotide substrates[197].

Cumulatively, efforts such as those highlighted above have dramatically increased the number of NDP-sugars (> 175 NDP-sugars; **Table 1.2**) available for glycorandomization and notably include sugars which bear uniquely reactive functional groups for further downstream chemoselective diversification of library members (see also section **1.5.1.4**)[141, 202]. To further augment *in vitro* enzymatic production and utilization of NDP-sugars in coupled Leloir GT-based systems, platforms for *in situ* NDP-sugar regeneration have also been developed[227-230]. NDP-sugar regeneration is advantageous as it ***i)*** eliminates the potential for feedback inhibition (from NDP), ***ii)*** helps drive desired glycosylation reactions (via maintaining an excess of available NDP-sugar), and ***iii)*** helps reduce the cost associated with certain NTPs. In addition, enzyme immobilization has also been used in the context of nucleotidyltransferases to simplify sugar nucleotide purification and even facilitate high-throughput screening[194, 231, 232].

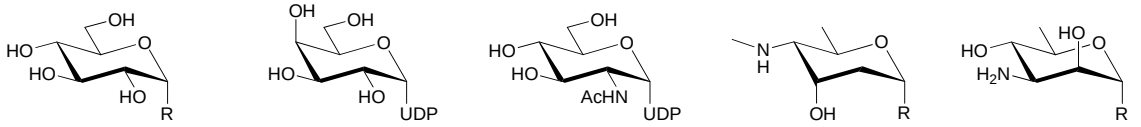
1.4.4. Permissive one-enzyme strategies

Another approach for generation of NDP-sugars is the exploitation of glycosyltransferase (GT) reversibility (**Figure 1.10a**). This approach attempts to co-opt a standard GT-catalyzed reaction and bias the inherent equilibrium of the overall reaction in favor of the reverse direction (*i.e.*, the conversion of a complex glycoside and an NDP into a corresponding aglycon and NDP-sugar). While there are many reports of the reversibility of GT-catalyzed reactions (yielding **152-162**, **171b**; **Figure 1.10b**)[141, 206, 233-246], the production of NDP-sugars has been almost exclusively restricted by the availability of complex natural product glycosides, the

a)



b)



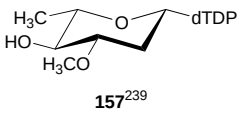
152a[139,141,233,236, 244,245]: R = UDP
152b[141]: R = dUDP
152c[139,141,236]: R = dTDP
152d[141]: R = CDP
152e[139,141]: R = ADP
152f[139]: GDP

153[240]

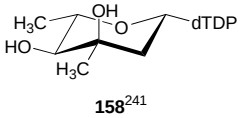
154[234,246]

155a[237]: R = UDP
155b[237]: R = dTDP

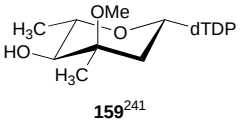
156a[206]: R = GDP
156b[206]: R = dGDP



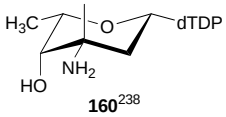
157[239]



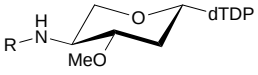
158[241]



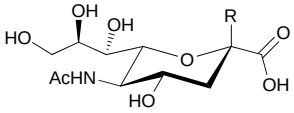
159[241]



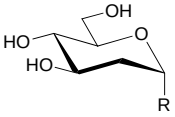
160[238]



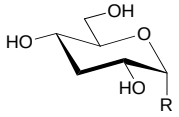
161a[238]: R = $CH_2(CH_3)_2$
161b[238]: R = CH_2CH_3
161c[238]: R = CH_3



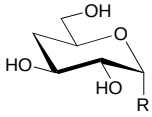
162a[235, 240, 243]: R = CMP
162b[235]: R = dCMP
162c[235]: R = UMP
162d[235]: R = CDP



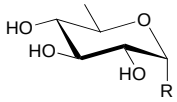
163a[139]: R = UDP
163b[139]: R = dTDP



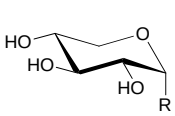
164a[139]: R = UDP
164b[139]: R = dTDP



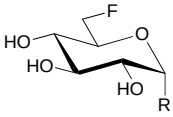
165a[139]: R = UDP
165b[139]: R = dTDP



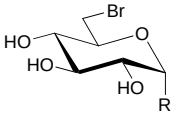
166a[139]: R = UDP
166b[139]: R = dTDP



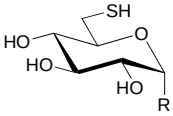
167a[139]: R = UDP
167b[139]: R = dTDP



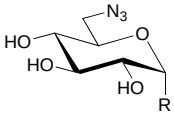
168a[139]: R = UDP
168b[139]: R = dTDP



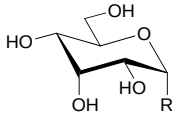
169a[139]: R = UDP
169b[139]: R = dTDP



170a[139]: R = UDP
170b[139]: R = dTDP



171a[139]: R = UDP
171b[139,238]: R = dTDP



172a[139]: R = UDP
172b[139]: R = dTDP

Figure 1.10. NMP- and NDP-sugars accessed through GT reversibility. a) GT reverse reaction scheme. **b)** NMP- and NDP-sugars catalyzed through GT reversibility. Compound numbers are followed by the corresponding literature reference(s).

corresponding GTs that catalyze their formation, and reaction thermodynamics which disfavor formation of the desired sugar nucleotides. However, recent work from the Thorson group promises to expanded this approach. Utilizing a small bank of engineered mutant variants of the GT OleD from *Streptomyces antibioticus*[133, 134, 212, 247, 248], aromatic β-D-glycosides were studied in reverse GT reactions to generate various NDP-sugars (**Figure 1.10b**)[139]. The results of this study form the basis of Chapter 3 and are discussed at length therein.

1.4.5. Format: *In vitro* versus *in vivo* 'synthetic biology'

While the bulk of 'classical' glycorandomization efforts have focused upon *in vitro* transformations (**Figure 1.11a**), it is important to note that any enzyme-based strategy also offers the potential for the development of *in vivo* platforms (**Figure 1.11b-g**). The initial *in vivo* strategies that successfully achieved alteration of natural product glycosylation utilized random mutagenesis, as exemplified by disruption of various tylosin (**173**) biosynthetic genes to yield strains defective in either the biosynthesis or attachment of all three native sugar moieties (**Figure 1.11c**)[249]. Rohr and co-workers were the first to expose inherent GT flexibility by isolating a glycosylated analogue of tetracenomycin from a variant of the urdamycin-producing strain *Streptomyces fradiae* Tü2353 engineered to also carry genes for production of tetracenomycin (**174**)or elloramycin (**175**), respectively. The glycosylated hybrid metabolite was proposed to derive from a combination of the tetracenomycin core, an urdamycin NDP-sugar donor and the elloramycin (**175**) GT (**Figure 1.11d**)[250]. Among other early examples of GT flexibility, Baltz and co-workers demonstrated *in vitro* and *in vivo* glycosylation of the heptapeptide aglycon A47934 (related to the glycopeptide teicoplanin (**176**, **Figure 1.26**)) used glycopeptide GTs (GtfE' and GtfB from the vancomycin (**177**) producer *Amycolatopsis*

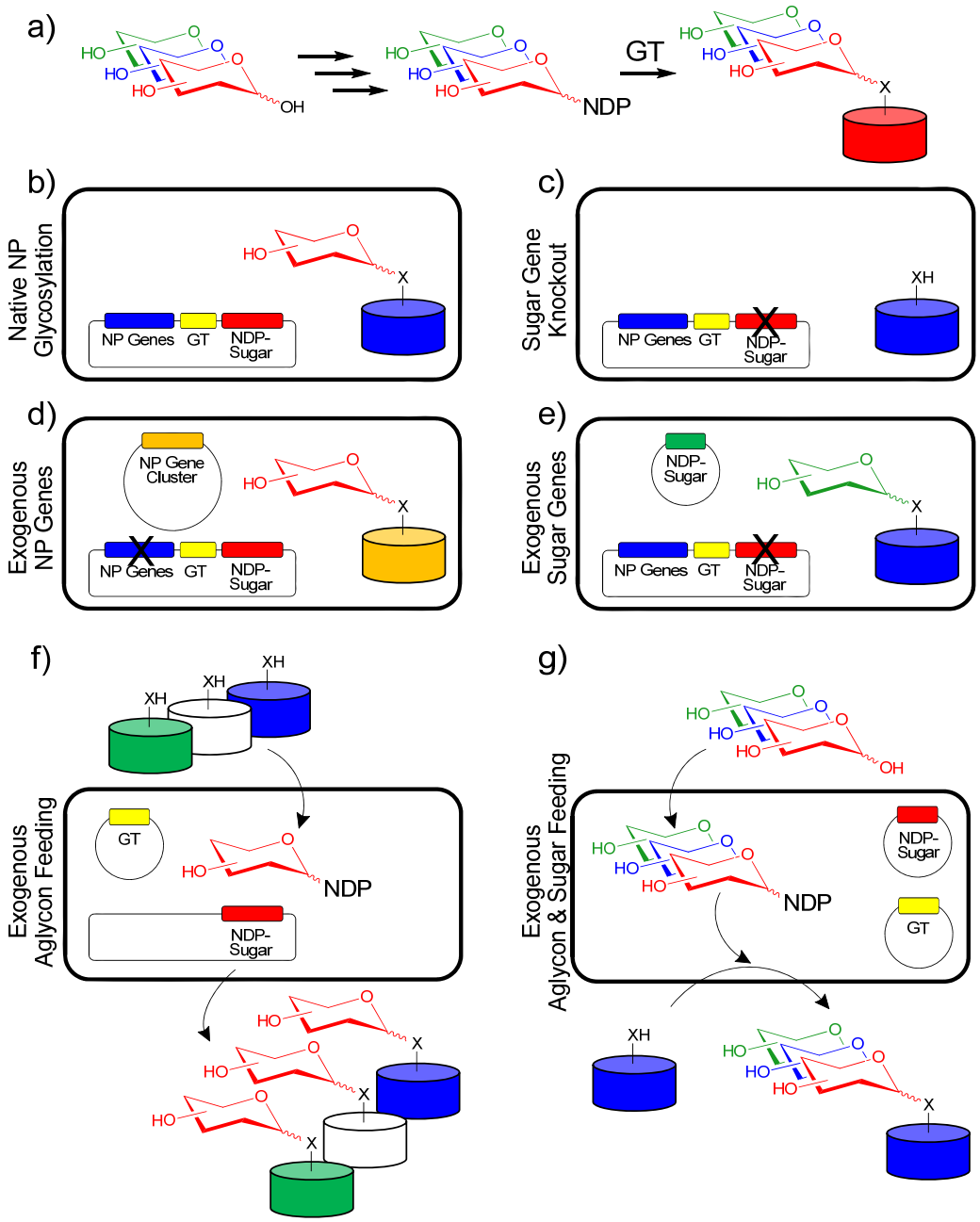
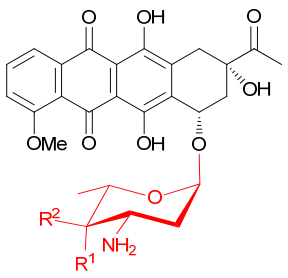


Figure 1.11. Various general approaches to glycorandomization: a) *In vitro* approach where glycosylation takes place completely outside the cell. **b)** Native natural product (NP) glycosylation where all the biosynthetic machinery is intact. **c)** Sugar gene knock-out wherein the biosynthetic machinery to generate the required NDP-sugar is disrupted, resulting in aglycon production. **d)** Exogenous NP pathways and/or genes are introduced to exploit native NDP-sugar biosynthesis and transfer. **e)** NDP-sugar biosynthesis pathways are introduced to complement knockouts and generate new glycosylated analogues. **f)** *In vivo* approach where aglycons are fed to a strain containing the biosynthetic machinery necessary to transfer a desired sugar. **g)** *In vivo* approach where both aglycons and non-natural sugars are fed to a strain containing the biosynthetic machinery necessary to both generate the non-native NDP-sugar and transfer it to a desired target.

orientalis C329.4) to enable among the first differentially glycosylated hybrid glycopeptide antibiotics[251]. Shortly thereafter, Hutchinson and co-workers reported the production of 4'-epi-doxorubicin (**1**) and 4'-epi-daunorubicin (**178**), analogues of doxorubicin (**2**) and daunorubicin (**179**), respectively, through a combination of gene inactivation and complementation with an inverting sugar 4'-ketoreductase gene (**Figure 1.1**, **Figure 1.11e** and **Figure 1.12**)[252]. These pioneering studies paved the way for a variety of subsequent *in vivo* strategies including complex metabolic pathway engineering for production of small molecules[253-255] and complex polysaccharides[173, 256-258], precursor-directed biosynthesis and mutasynthesis[151, 259], and the development of plasmids carrying a range of sugar nucleotide biosynthesis/attachment genes[144, 145] to ultimately modulate the sugar moieties on a wide range of targets[121, 260].

Inspired by these pioneering *in vivo* studies, investigators have also began to explore the potential of strains carrying more permissive sugar nucleotide machineries (*e.g.*, those described in section **1.4.3**) and GTs as general glycosylation hosts. The advantage of such *in vivo* platforms include translation of the 'cost' (for expensive reagents such as nucleotides) to the host organism, the elimination of a need for enzyme purification/storage and the potential for large scale production of desired products via standard fermentation. Such an approach has been sucessfully applied by a number of groups to generate relatively large quantities of complex saccharide targets using various recombiant organisms[256-258, 261, 262]. The *in vivo* glucosylation of various small molecules via bioconversion has also been accomplished using bacterial strains engineered to overexpress various plant GTs[263-265]. As an initial assessment of whether unnatural sugars could be utilized by such strains, Thorson and co-workers revealed the *in vivo* production of unnatural sugar-1-phosphates using an *E. coli* host engineered to over express M173L/Y371H GalK (see section **1.4.3.1**) in the presence of various modified novel free sugars[167, 266]. Two



178: R^1 = H, R^2 = OH
179: R^1 = OH, R^2 = H

Figure 1.12. 4'-epi-daunorubicin and daunorubicin.

prototype *E. coli* strains were subsequently developed for the first demonstration of *in vivo* glycorandomization[248]. In this pioneering study, 'glucoside'-strains (expressing only GT OleD variants)[133, 134, 212] led to the *in vivo* glucosylation of a series of acceptors[248] previously identified as substrates *in vitro*[212, 247] (**Figure 1.11f**). The 'non-natural sugar'-strain (expressing engineered anomeric kinase, sugar-1-phosphate nucleotidyltransferase, and GT OleD variant TDP-16), led to the glycodiversification of 4-methylumbelliferone (**180**, **Figure 1.38**) with all 9 sugars examined[248] (**Figure 1.11g**). This work enabled, for the first time, the ability to process both sugars and aglycons not accessible via known biosynthetic pathways. It is anticipated that future efforts directed toward strain improvement for sugar/aglycon uptake and/or reduction of competing *in vivo* side reactions will provide glycodiversification strains of significant utility.

1.5 Representative glycorandomized libraries

The bulk of glycorandomized small molecule libraries to date have derived from *in vitro* and *in vivo* GT-catalyzed strategies with a diverse range of novel NDP-sugars donors (see section **1.4**). Libraries highlighted in this section are limited to biologically active small molecules (bioactive secondary metabolites and drugs) and are classified based upon the GT employed to include ***i)*** libraries generated using a GT inherent to the biosynthesis of a targeted aglycon and ***ii)*** libraries constructed using uniquely promiscuous GTs.

1.5.1. Libraries based upon inherent biosynthetic glycosyltransferases

1.5.1.1. Aminocoumarins

The aminocoumarins consist of a common core structure of 3-amino-4-hydroxy-coumarin, a prenylated 4-hydroxybenzoic acid moiety, and a modified deoxy-sugar which is

essential for biological activity. These antibiotics are potent inhibitors of bacterial DNA gyrase[267] and members of this family, such as novobiocin (**181**), also inhibit Hsp90, a validated cancer target[268], by binding to the *C*-terminal ATP binding site[269]. Three glycosyltransferases, NovM, CloM, and CouM, have been characterized as the catalysts for glycosylation of novobiocin (**181**), clorobiocin (**182**), and coumermycin A_1 (**183**), respectively (**Figure 1.13**)[270]. All three GTs catalyze the attachment of 5-*O*-methyl-L-rhamnose (commonly referred to as L-noviose) prior to final sugar maturation via downstream tailoring enzymes. *In vivo* studies probing the bioconversion of synthetic substrate analogues revealed CloM as permissive toward 15 aminocoumarin aglycons (**184-198**) to afford a set of clorobiocin analogues bearing L-noviose (**Figure 1.14**)[270]. *In vitro* studies revealed a similar aglycon flexibility for NovM including the transfer L-noviose to novobiocic acid (**199**), cyclonovobiocic acid (**200**), a simocyclinone analogue (**201**) and a host of other related coumarin analogues and phenols (**202-205**) to produce **199a-205a**, respectively, with varying efficiencies (**Figure 1.15**)[271, 272]. Subsequent efforts to probe sugar promiscuity (using 34 UDP- and TDP-sugars) exposed a surprisingly narrow donor specificity for NovM, with only 3 sugars (6-deoxy-D-glucose, D-xylose and 4-keto-6-deoxy-D-glucose) transferred to novobiocic acid (**199**) to produce the aminocoumarin analogues **199a-199d**[203] (**Figure 1.15**). Whether related aminocoumarin GTs, such as CloM and CouM[273], display similar sugar donor stringency has yet to be investigated. To circumvent the limited NDP-sugar donor promiscuity of NovM, Thorson and co-workers developed an engineered variant of the oleandomycin GT OleD which enabled the production of 10 differentially glycosylated novobicin analogues (see section **1.5.2.1**)[212, 248].

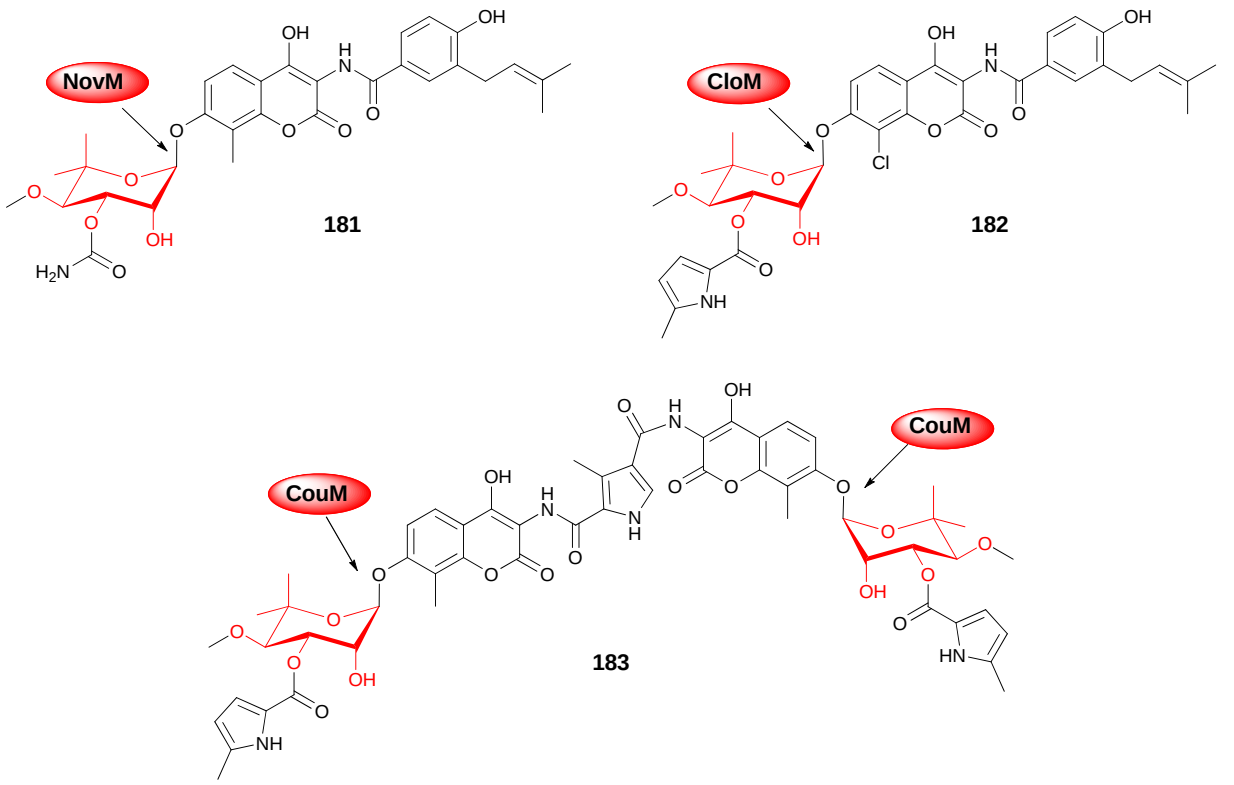


Figure 1.13. Aminocoumarins and the glycosyltransferases responsible for attachment of L-noviose (highlighted in red) **during their biosynthesis.**

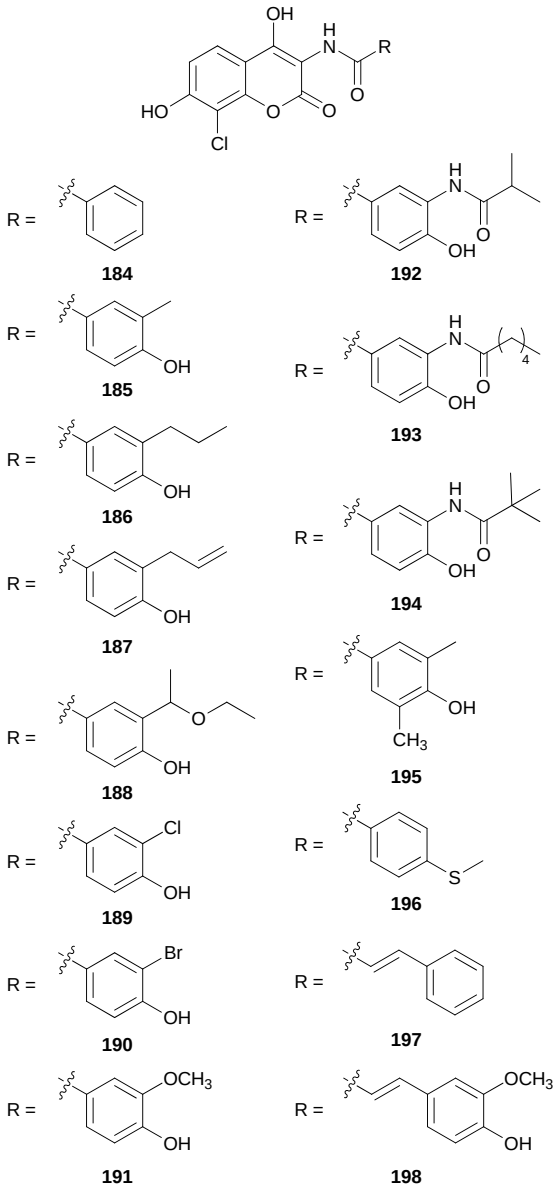


Figure 1.14. Clorobiocin analogues recognized by the GT CloM.

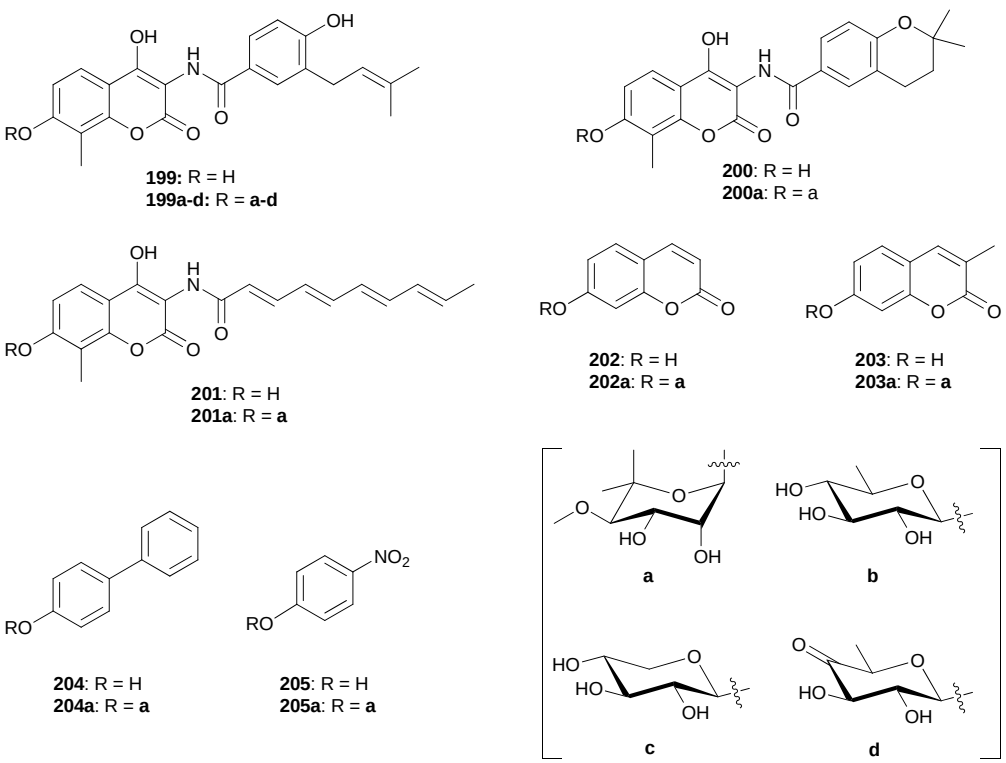


Figure 1.15. Aglycons and compounds recognized and catalyzed, respectively, by the GT NovM.

1.5.1.2. Aromatic polyketides

Aclarubicins are members of the anthracycline class of antitumor natural products ((*e.g.* doxorubicin (**2**), aclarubicin A (**206**), idarubicin (**207**)) structurally defined by a tetracyclic aromatic polyketide planar scaffold appended by one or more unique sugars critical for both drug-DNA interactions[274] and *in vivo* PK (**Figure 1.16**; see section **1.2.1**). The annotation of the aclarubicin A (**206**) biosynthetic gene cluster from *Streptomyces galilaeus* and corresponding pathway engineering revealed two GTs, AknS and AknK, as responsible for synthesizing the trisaccharide and also displayed flexibility toward alternative sugar donors[275, 276]. AknS was demonstrated *in vitro* to catalyze the transfer of the first sugar L-rhodosamine to the aglycon aklavinone (**208**) to yield **209**[277, 278], and to require AknT as an activating partner, as previously demonstrated for DesVII/DesVIII enzyme pair[279] (see section **1.5.1.7**). In addition to the native sugar L-rhodosamine[278], AknS/AknT also accommodated 2-deoxy-L-fucose and 4-amino-2-deoxy-L-rhamnose to afford **210** and **211**[277] (**Figure 1.16**), as well as a non-natural aglycon, ε-rhodomycinone (**212**), *en route* to novel glycosides **213** and **214**[277,278]. In a series of *in vitro* experiments, AknK was shown to catalyze sugar transfer to **209** (forming **215** or **216**), **210** (forming **217** or **218**), **214** (forming **219** or **220**), **215** (forming **221** or **222**) and **207** (forming **223**)[280] (**Figure 1.16**). Although the sugar donor substrate specificity of AknS/AknT- and AknK-catalyzed reactions requires more extensive investigation, these preliminary studies illustrate a promising role of AknS/AknT pair and AknK in future glycorandomization/diversification applications.

Aranciamycin (**224**; **Figure 1.17**) is an anthracycline produced by *Streptomyces echinatus*[281] which, in addition to the DNA-targeted activities of most anthracyclines, also uniquely inhibits collagenase[282]. *In vitro* characterization of AraGT, encoded by the only

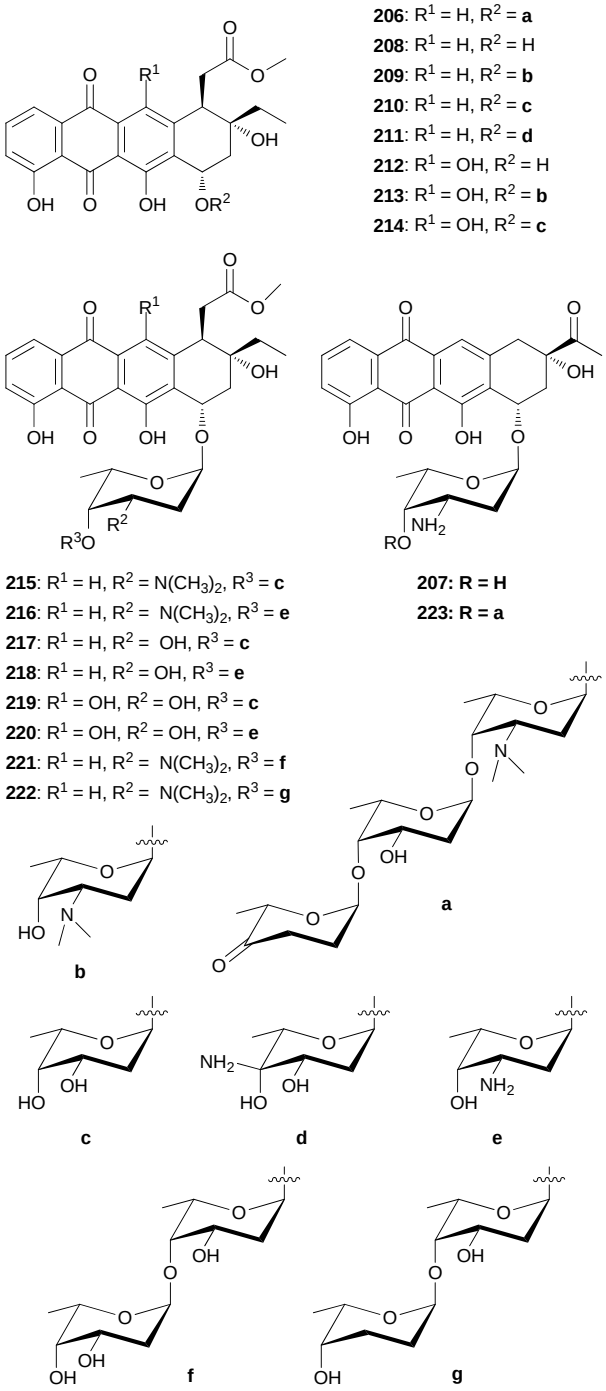


Figure 1.16. Substrates and products of AknS/AknT and AknK.

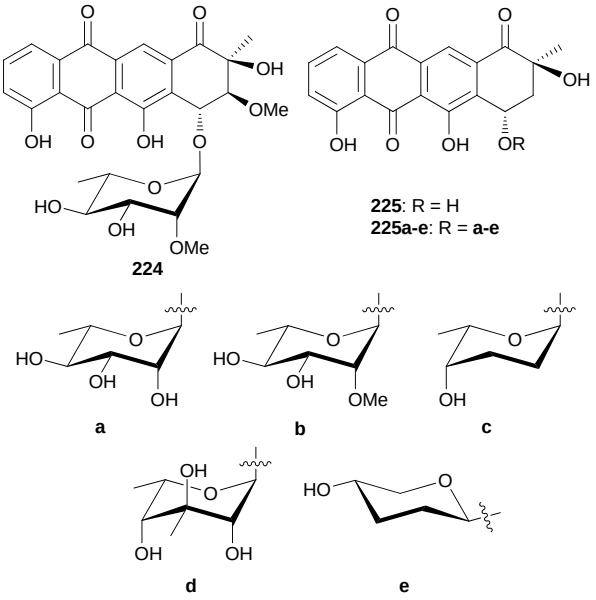


Figure 1.17. Substrates and products of the GT AraGT.

annotated GT gene within the aranciamycin (**224**) biosynthetic gene cluster, confirmed AraGT-catalyzed transfer of L-rhamnose to demethoxy-aranciamycinone (**225**) to yield **225a**[282] (**Figure 1.17**). The sugar donor flexibility of AraGT is supported by *in vivo* studies which led to 4 additional novel glycosylated aranciamycin derivatives (**225b-e**; **Figure 1.17**)[283].

A tetracycline-like compound possessing anti-tumor and anti-bacterial properties, elloramycin (**175**) is produced by *Streptomyces olivaceus* Tü2353 (**Figure 1.18**). Based upon a number of *in vivo* studies, the GT required for elloramycin (**225**) biosynthesis, ElmGT, displayed unique flexibility toward an array of NDP-sugar donors and aglycon acceptors[144-146, 284-289]. Specifically, ElmGT accomodated the transfer of 14 sugar moieties *in vivo* to either 8-demethyl-tetracenomycin C (**226**) or steffimycin analogues (**227** and **228**). This, in turn, generated **226a-n**, **227b-d**, **227h**, **227i**, and **227m** or **228c** and **228m** from the respective parent aglycons (**Figure 1.18**). Interestingly, the sugars transferred by ElmGT included both L- (7 total) and D-sugars (6 total) and formation of a disaccharide moiety. In a follow-up study, mutation of the postulated sugar donor site (L309 and N312) revealed the ability to further modulate the deoxy-sugar flexibility of ElmGT[149]. Collectively, these studies showcase the substrate promiscuity of ElmGT, demonstrate the augmentation of sugar donor promiscuity via targeted mutagenesis and also highlight an excellent example of introducing foreign sugar pathways into heterologous hosts for *in vivo* glycorandomization/diversification.

The gilvocarcin-type anticancer natural products share a polyketide-derived benzo[*d*]naphto[1,2-*b*]pyran-6-one moiety and often possess novel *C*-linked furanosides or pyranosides[150]. *In vivo* studies confirmed *gilGT* from the gilvocarcin V (**229**) gene cluster to encode a GT for *C*-glycosylation of C-4 with D-fucofuranose[290]. GilGT subsequently displayed NDP-sugar flexibility in the context of both native pathway engineering (to yield analogues **230b**

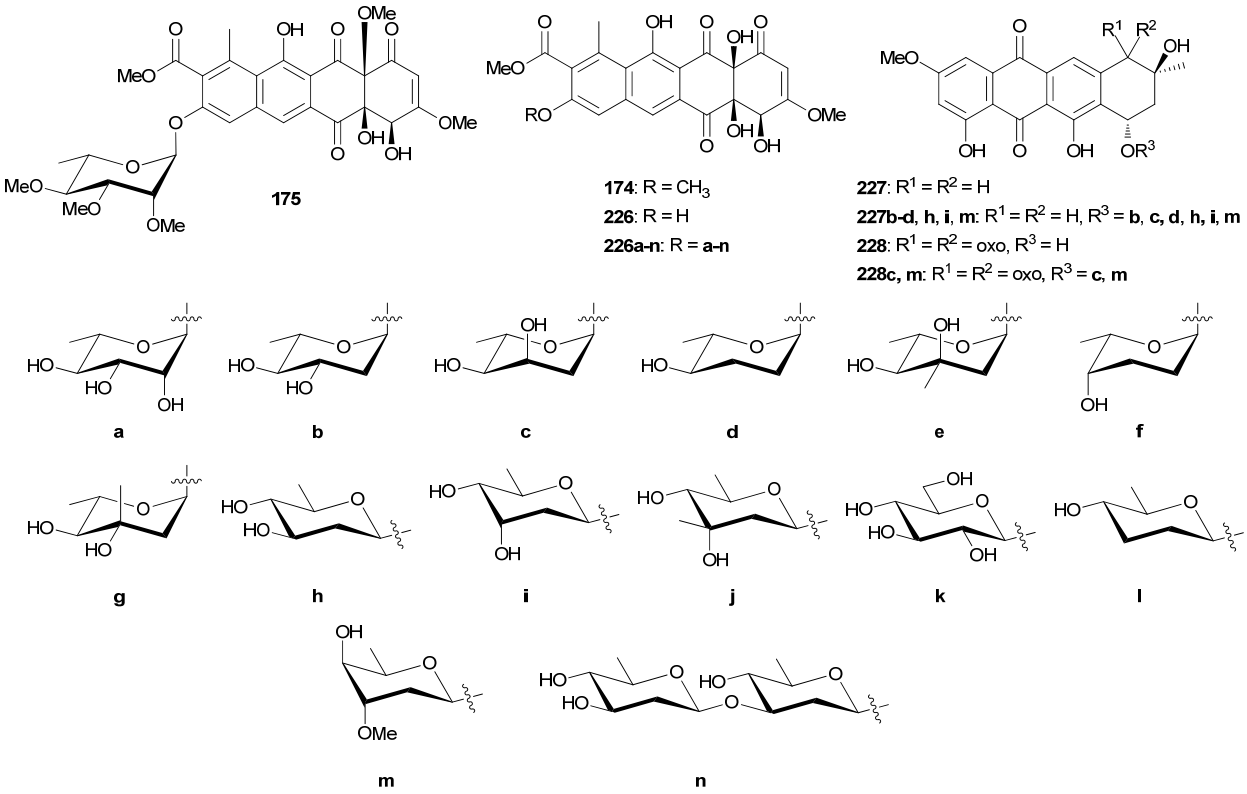


Figure 1.18. The secondary metabolite tetracenomycin along with the substrates and products of the GT ElmGT.

to **232b**)[291] or endogenous sugar nucleotide plasmids (to enable the transfer of two additional sugars to generate **230c-d** to **232c-d**; **Figure 1.19**)[150]. Two of the new analogues (**230c** and **232d**) demonstrated anticancer activity comparable to gilvocarcin V (**229**)[150]. While this stands as one of the first successful applications of a *C*-GT in glycorandomization/diversification, it is important to note that some GTs are innately capable or can be engineered for both *O*- and *C*-glycosidic bond formation[292, 293], suggesting the potential to redirect more permissive *O*-GTs toward *C*-glycosylation.

1.5.1.3. Enediynes

Enediynes are characterized by a signature nine- or ten-membered enediyne core decorated via an array of modifications including novel sugars[294, 295]. All family members are remarkably efficient DNA-damaging agents and the potent cytotoxicity of enediynes has served as the basis for the development of enediyne-based anticancer therapeutics. Calicheamicin (**233**), a flagship member of the 10-membered enediyne family, binds DNA via an aryltetrasaccharide moieity containing four unique deoxy-sugars (**Figure 1.20**)[294]. The identification of biosynthetic gene cluster of **233** revealed four putative GT-encoded genes[296], *calG1*, *calG2*, *calG3* and *calG4*, the products of which have all been both characterized[238, 242] and crystallized (http://www.pdb.org; accession numbers 3OTH, 3IAA, 3D0R, and 3IA7 for CalG1, 2, 3, and 4, respectively). Importantly, these studies illuminated the reversibility of GT-catalyzed reactions and highlighted the potential for exploiting this under-appreciated phenomenon in the context of complex natural product glycodiversification.

While CalG2 was found to accept only one (TDP-α-D-4,6-dideoxyglucose) of a set of 22 putative TDP-sugar donors in the presence of **234** as acceptor *in vitro* (to provide **235**; **Figure**

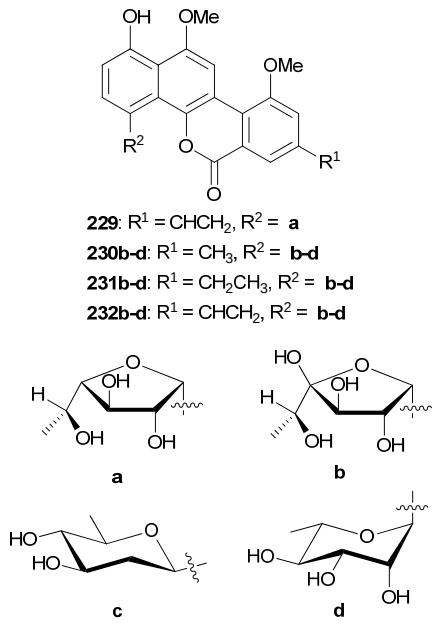
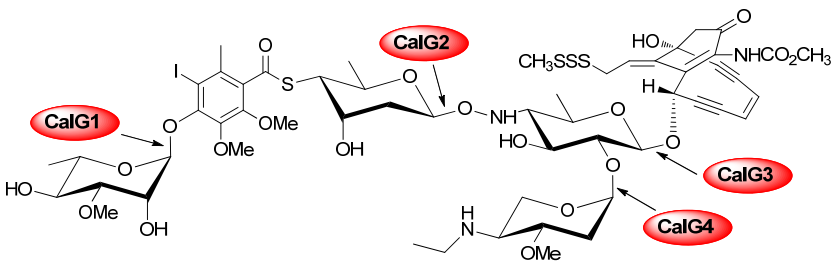


Figure 1.19. Substrates and products of GilGT.



233

Figure 1.20. Structure of the 10-membered enediyne calicheamicin.

1.21)[242], a similar study with CalG1 using **236** (**Figure 1.22**) as an acceptor revealed 10 non-natural sugar donors[238]. The studies also revealed the CalG1 catalyzed reaction to be readily reversible and to provide TDP-3-*O*-methyl-rhamnose in the presence of TDP and various calicheamicins (**236a-243a**) (**Figure 1.22**)[238]. Thorson and co-workers demonstrated the utility of GT-catalyzed *in situ* 'sugar exchange' via swapping the native 3-*O*-methyl-rhamnose moiety of **236a** to **243a** with 10 alternative sugars to provide >70 differentially glycosylated calicheamicins (**236b-k**, **237b-k**, **238b-k**, **239b-k**, **240b-k**, **241b-k**, **242b-k**, **243b-c**; **Figure 1.22**)[238]. Similar studies were conducted with both the requisite aminopentosyltransferase CalG4[238], and the novel hydroxylaminohexosyltransferase CalG3. In the latter study, CalG3 mediated the removal of the **234** hydroxylaminohexose to produce calicheamincinone **244** (**Figure 1.23**) in the presence of various NDPs and could also accommodate 10 alternative TDP-sugar donors via 'sugar exchange' to provide **245a-j** (**Figure 1.23**)[242].

CalG1 was also the centerpiece for demonstration of novel 'aglycon exchange' reactions, in which the 3-*O*-methylrhamnose was transferred among different enediyne aglycon acceptors *in situ*[238]. This demonstration was expanded to also include a CalG1-GtfE coupled reaction, in which the non-natural azido-sugar of **246** was excised by GtfE and then transferred *in situ* to the calicheamicin aglycon **236** by CalG1 to yield **247** and **236g** (**Figure 1.24**)[238]. Aglycon exchange reactions were also demonstrated using CalG4 where the aminopentose moiety present in **237a**, **238a**, **239a** and **240a** was excised (in turn, providing **236a**, **248**, **249** and **250** as by-products, respectively) and transferred to aglycon **236** to produce **251**, **252** and **253**, respectively (**Figure 1.25**)[238]. These examples notably demonstrated that the exotic sugars appended to complex natural products could be efficiently moved from one natural product to another in simple one pot GT-catalyzed reactions. This work notably inspired recent studies to use simple activated

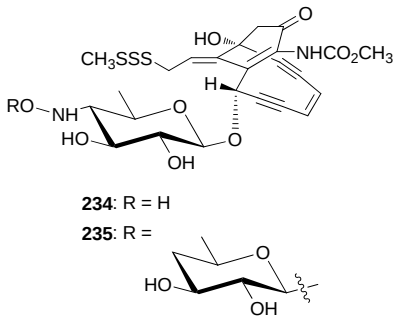


Figure 1.21. The calicheamicin analogue generated by CalG2.

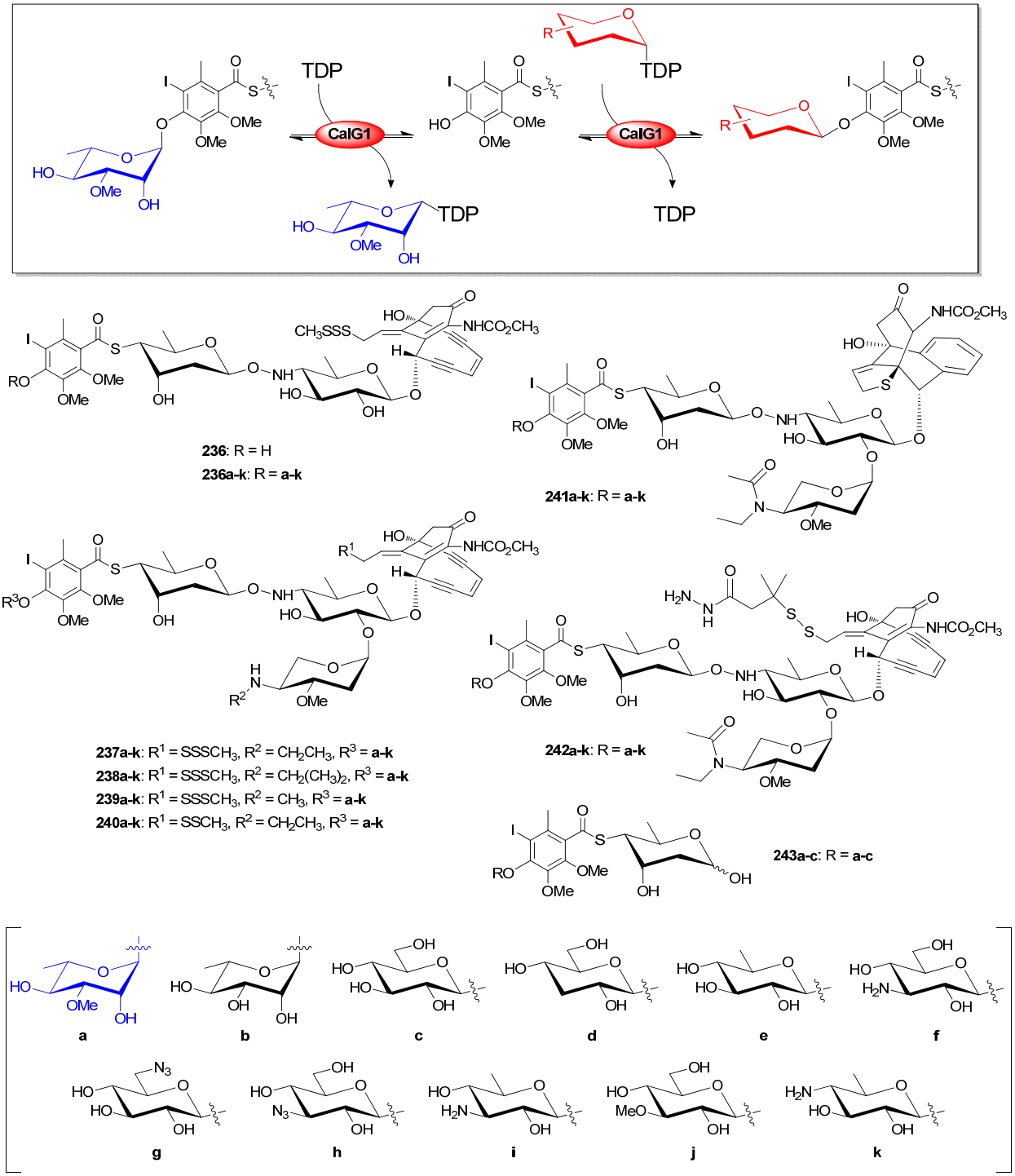


Figure 1.22. The calicheamicin analogues generated by CalG1.

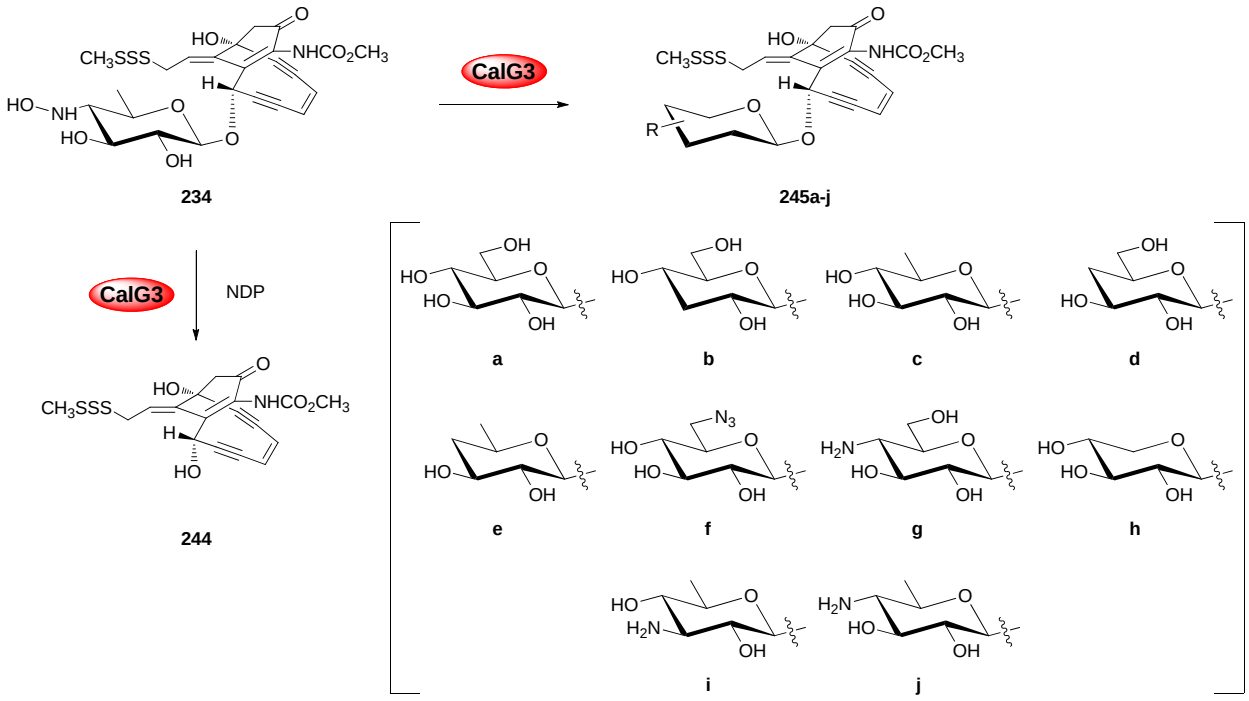


Figure 1.23. CalG3 reversibility and sugar exchange reactions.

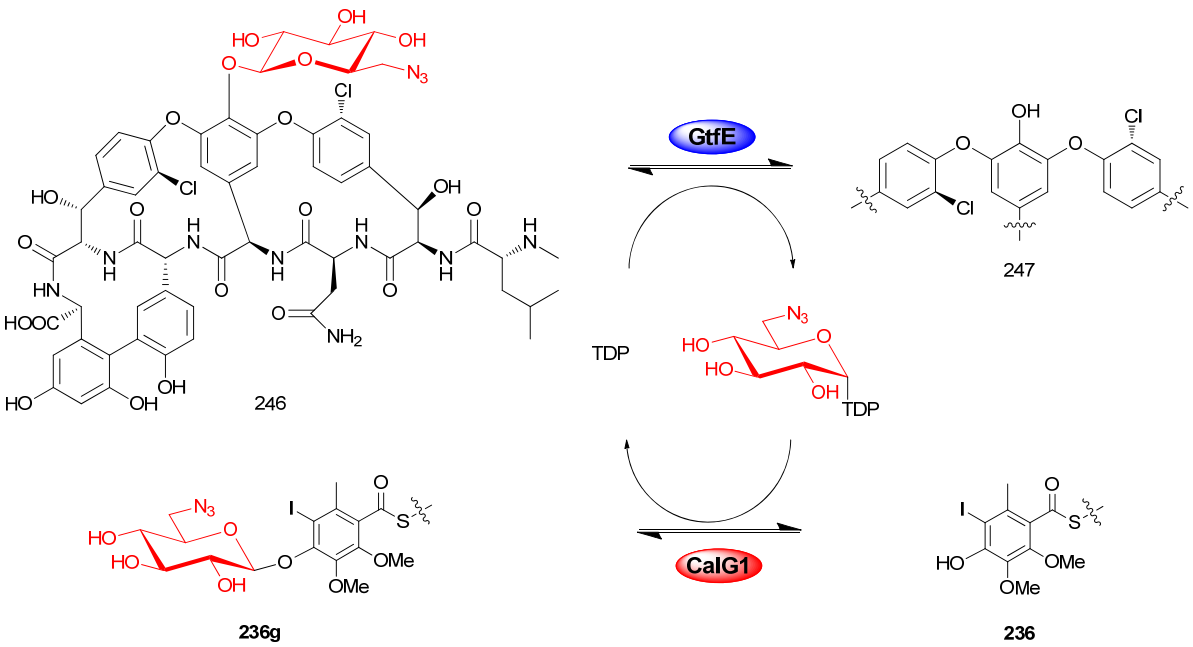


Figure 1.24. Aglycon exchange catalyzed by the dual action of GtfE (from vancomycin biosynthesis) **and CalG1** (from calicheamicin biosynthesis).

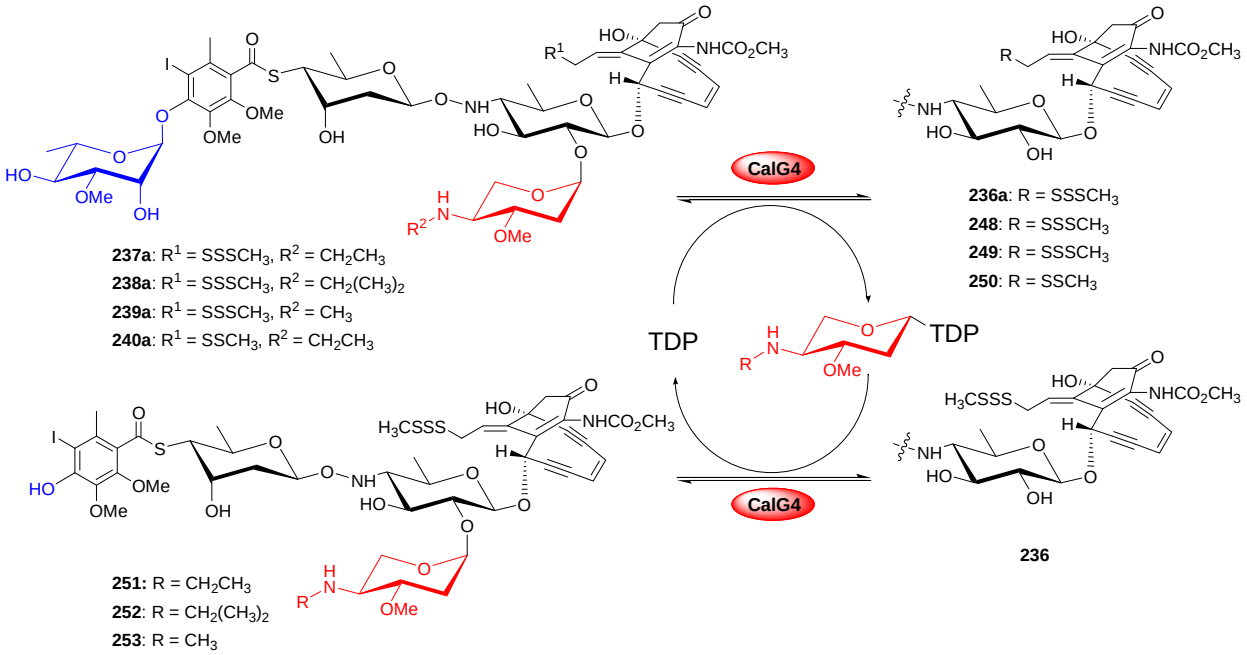


Figure 1.25. Aglycon exchange reactions catalyzed by CalG4.

glycoside donors to drive targeted transglycosylation reactions toward the synthesis of novel sugar nucleotides and corresponding glycosides (see section **1.4.4**).

1.5.1.4. Glycopeptides

Teicoplanin (**176**), vancomycin (**177**) and chloroeremomycin (**254**) are non-ribosomal glycopeptide antibiotics which serve as a basis for a dwindling set of antibiotics to treat life-threatening infections caused by multi-resistant Gram-positive bacteria (**Figure 1.26**)[297]. The rapid global emergence of vancomycin-resistant enterococci (VRE) and vancomycin-resistant *Staphylococcus aureus* (VRSA) has inspired the development of novel glycopeptides with improved activity against resistant strains[298-300], highlighting the influence of vancomycin's (**177**) α-L-vancosaminyl-(1,2)-β-D-glucosyl disaccharide modification in circumventing key resistance mechanisms[238, 251, 301-304]. Whether this established precedent for improving glycopeptide therapeutics via glycosylation will hold true for other non-ribosomal peptide antibiotics, such as the lipopeptide daptomycin[305], remains to be determined.

Both GTs required for vancomycin (**177**) biosynthesis, GtfE and GtfD, have been characterized *in vitro*[251, 306] and utilized in glycorandomization and 'aglycon exchange' schemes. A systematic evaluation of GtfE sugar nucleotide specificity demonstrated GtfE to accept thirty-one out of thirty-three NDP-sugars tested in the presence of vancomycin aglycon (**247**). Collectively, the novel sugars incorporated displayed changes at C-2, C-3, and/or C-4 of the sugar[302, 306, 307] to provide an array of differentially monoglycosylated vancomycin analogues (**246**, **255-284**) (**Figure 1.27**) some of which (e.g., azidosugar substituted **246**) were further diversified via chemoselective chemistries[307, 308]. In a similar manner, Walsh and co-workers evaluated the sugar donor substrate specificity of GtfD, GtfC and GtfA with 10 chemically

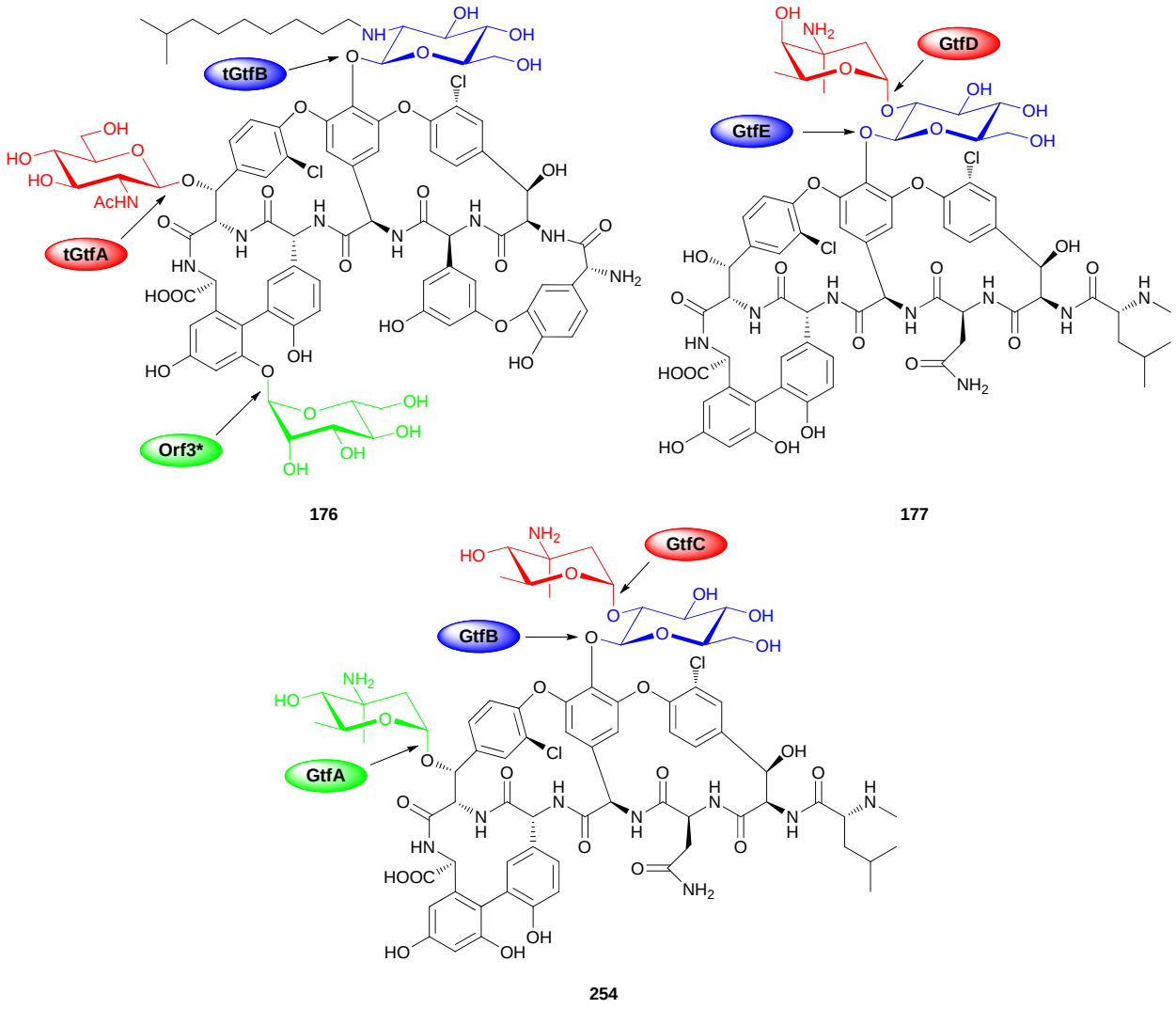


Figure 1.26. GTs involved in the biosynthesis of representative glycopeptides.

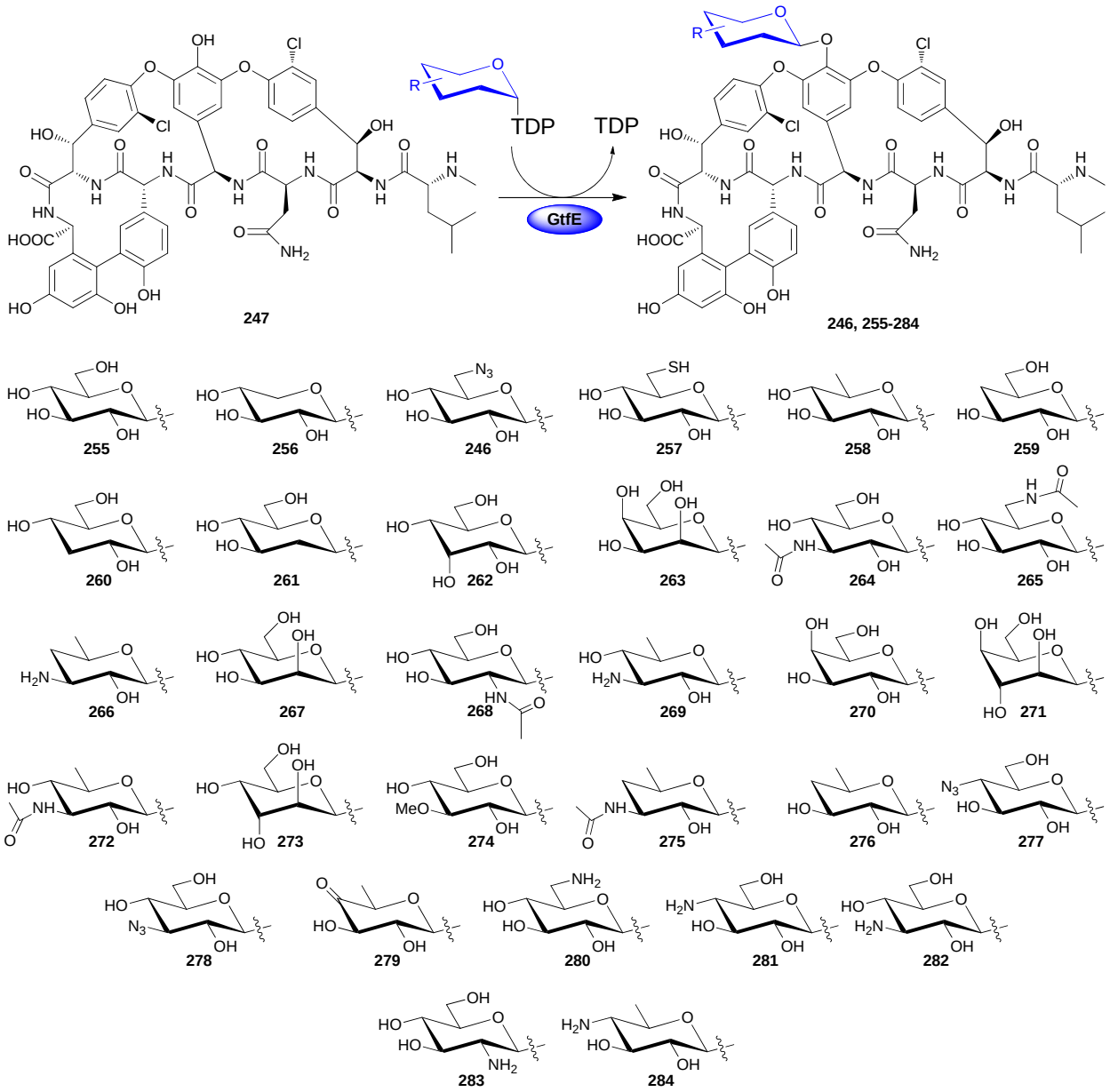


Figure 1.27. Vancomycin analogues generated by GtfE.

synthesized TDP-2-deoxy-β-L-sugar donors[309]. In this study, GtfD displayed the ability to transfer an additional 7 sugars to its natural substrate vancomycin pseudo-aglycon (**255**) to generate **177** and **285b-g** (**Figure 1.28**). GtfC efficiently catalyzed the transfer of 6 sugars to **256** for formation of trace amounts of **177** and **285b-f** and 7 sugars to **286** for formation of **254**, **287a** and **287c-g** (**Figure 1.28**). GtfA was capable of transferring 3 of the tested sugars onto **255** to yield **288b**, **288d** and trace amount of **288f** (**Figure 1.28**). Additionally, both GtfD and GtfE catalyzed reactions were revealed to be reversible and amenable to 'aglycon exchange' reactions (as described in section **1.5.1.3**)[238].

The biosynthesis of teicoplanin (**176**) requires three GTs, tGtfA, tGtfB and Orf3*, two of which, tGtfA and tGtfB, display significant homology to GtfA and GtfB, respectively[310, 311]. While their native activities have been confirmed[310, 312], the sugar flexibility of tGtfA and tGtfB has not been reported. However, both tGtfA and tGtfB were utilized (along with the vancomycin GTs GtfA and GtfB) for the creation of GT chimeras for the alteration of both teicoplanin and vancomycin glycosylation patterns[313]. When compared to their wild-type counterparts, the functional chimeras generated in this study demonstrated altered sugar specificity but maintained regioselectivity of the parental GT utilized for the *N*-terminus portion of the chimera (responsible for aglycon recognition)[313]. This study notably highlights the potential for GT domain swapping as another mechanism for extending the substrate scope of biosynthetic GTs. A recent study has also demonstrated glycodiversification of teicoplanin by modifying pre-existing sugars through a combination of sugar modification enzymes (deacetylase, acyltransferase and oxidase) and compatible alkylation chemistries[314].

Post-translationally modified peptides are a rapidly expanding class of natural products with a range of biological activities that include instances of analgesic, antitumor and

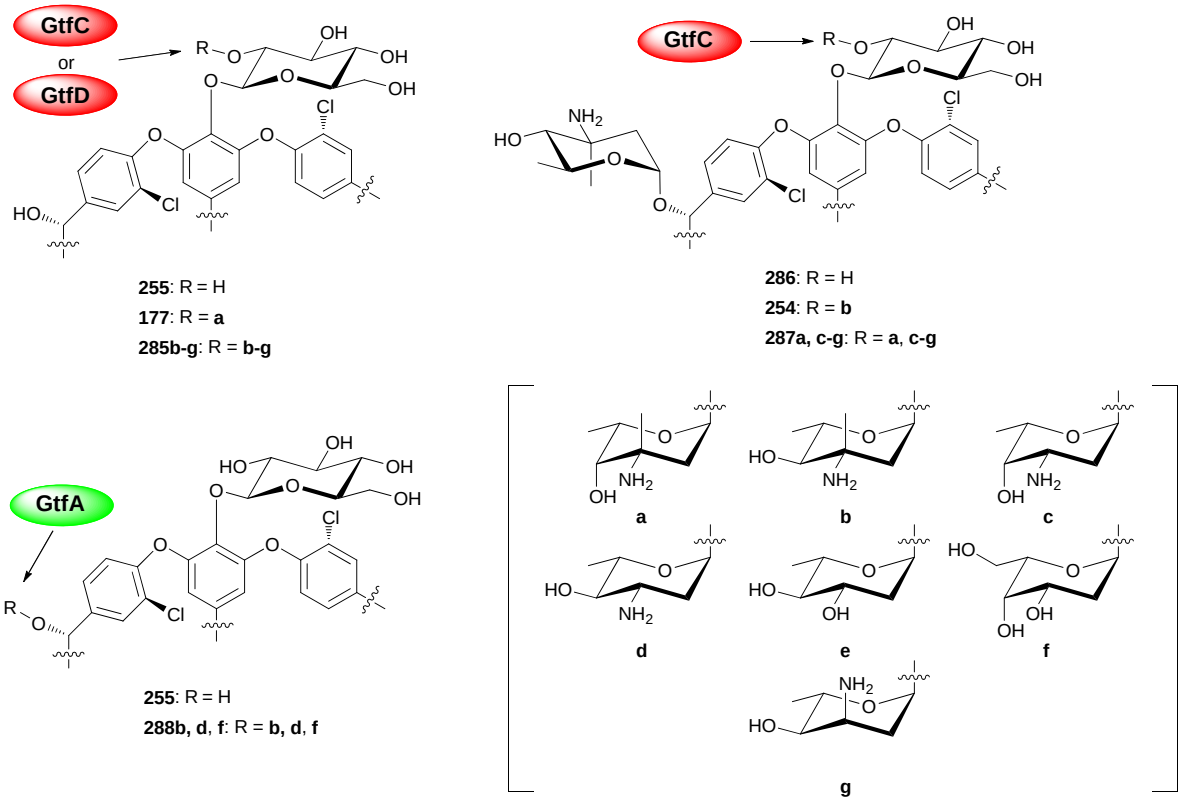


Figure 1.28. Various glycopeptide analogues generated by GtfC, GtfD and GtfA.

antibacterial properties[315, 316]. Synthesized by the ribosome, almost all members of this class are characterized by an *N*-terminal leader sequence and a *C*-terminal core peptide which is processed and modified into the mature compound[316]. During processing of the *C*-terminal core peptide, over two dozen different types of biosynthetic modifications can occur, including both *O*-[316, 317] and *S*-glycosylation[318, 319]. As an example of the latter, sublancin (**289**) contains a *S*-linked β-D-glucosyl cysteine residue. The GT responsible, SunS, was characterized *in vitro* by van der Donk and colleagues[318]. In addition to the native substrate UDP-α-D-glucose, SunS was demonstrated to accept UDP-α-D-*N*-acetylglucosamine, UDP-α-D-galactose, UDP-α-D-xylose and GDP-α-D-mannose in the presence of the sublancin aglycon (**290**) to generate **291b-e** (**Figure 1.29**). While antimicrobial assays of the resulting products highlighted glycosylation as critical for activity, the specific nature of the attached sugar was not important (perhaps reminiscent of some of the glycosylated peptides briefly mentioned in section **1.2.2**). The flexibility of SunS in terms of both base and sugar specificity, combined with the emerging rise in post-translationally modified peptide discovery[316], suggests a potential new opportunity for glycorandomization.

1.5.1.5. Indolocarbazoles

Isolated from actinomycetes, cyanobacteria, fungi, slime moulds and marine invertebrates, the indolocarbazoles are a subclass of alkaloids with a range of bioactivities[320]. Two key members, rebeccamycin (**292**) and staurosporine (**293**), are well characterized for their abilities to inhibit DNA topoisomerase I and protein kinase C, respectively[321]. GTs involved in the biosynthesis of rebeccamycin and staurosporine, RebG and StaG, respectively, have been utilized in attempts for glycorandomization/diversification of indolocarbazoles (**Figure 1.30**). The GT-encoding genes *atmG* and *atmG1* required for biosynthesis of AT2433 (**294**), the

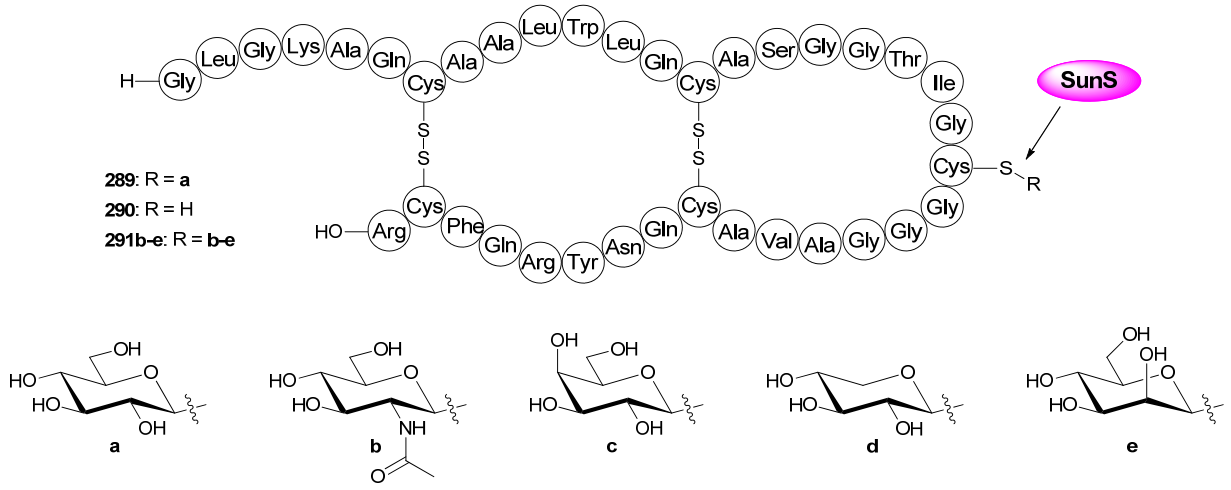


Figure 1.29. Analogues of sublancin generated by the GT SunS.

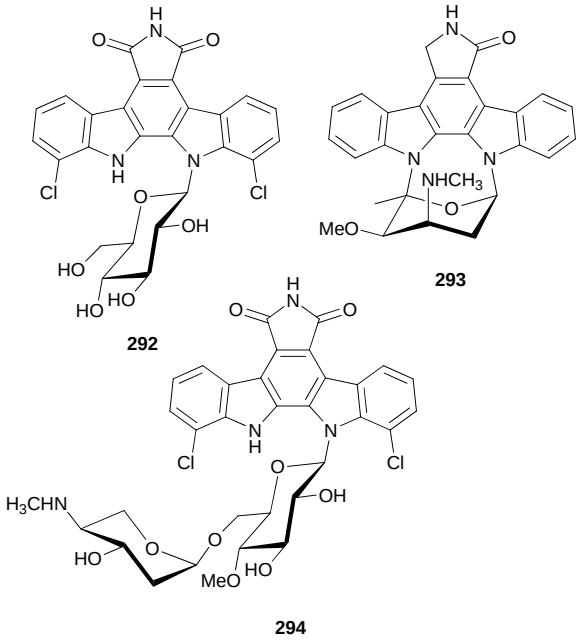
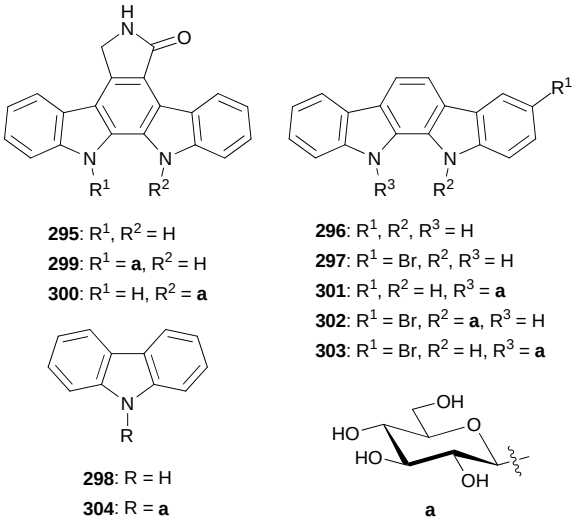


Figure 1.30. Representative indolocarbazoles.

disaccharide-substituted analogue of rebeccamycin, have also been reported[322]. Of the 19 putative aglycons screened for *in vivo* biotransformation assays by a RebG-expressing strain, only 4 aglycons (**295-298**) led to the respective glycosylated products **299-304** (**Figure 1.31**)[323]. Alternatively, through the use of a 'two-plasmid' system for heterologous co-expression of aglycon and sugar nucleotide production, StaG (with additional sugar functionalization by the P450 oxyenase StaN) was demonstrated to accommodate both D- and L-sugars *in vivo* to provide **305a-g** from **295** (**Figure 1.32**)[321]. Cumulatively, these studies suggest that certain indolocarbazole associated GTs may be beneficial to glycorandomization strategies targeting this class of natural product.

1.5.1.6. Macrolactams

VinC, the GT involved in the biosynthesis of the antitumor polyketide glycoside vicenistatin (**306**) from *Streptomyces halstedii* HC-34, is responsible for the transfer of the aminosugar vicenisamine to the 20-membered macro-lactam aglycone **307** (**Figure 1.33**)[324]. An additional panel of aglycons were identified as VinC acceptors, including neovicenilactam (**308**), which is structurally related to **307**, and brefeldin A (**309**), α- and β-zearalenol (**310** and **311**), a racemic substrate mimic (**312**), β-estradiol (**313**), pregnenolone (**314**), 3-*O*-acetyl-β-estradiol (**315**), and a simple alcohol substrate mimic (**316**) (**Figure 1.33**)[325]. In these studies, VinC efficiently catalyzed the transfer of vicenisamine to **308-316** to produce **308a-315a** and trace **316a** (**Figure 1.33**)[325]. The VinC reaction was also amenable to both reverse and aglycon exchange reactions[237]. In a study to probe NDP-sugar donor substrate specificity, VinC utilized both the α- and β-anomers of several chemically synthesized 'unnatural' TDP-2-deoxy-sugars and, in conjugation with different algycon acceptors, retained stereospecificity of transfer



295: R^1, R^2 = H
299: R^1 = **a**, R^2 = H
300: R^1 = H, R^2 = **a**

296: R^1, R^2, R^3 = H
297: R^1 = Br, R^2, R^3 = H
301: R^1, R^2 = H, R^3 = **a**
302: R^1 = Br, R^2 = **a**, R^3 = H
303: R^1 = Br, R^2 = H, R^3 = **a**

298: R = H
304: R = **a**

a

Figure 1.31. Products generated by RebG during biotransformations and their respective aglycons.

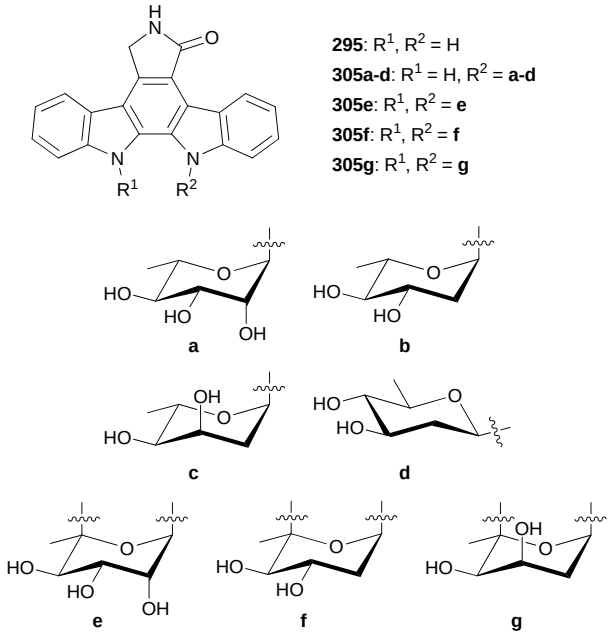


Figure 1.32. Products generated *in vivo* from their parent aglycon by StaG and/or StaN.

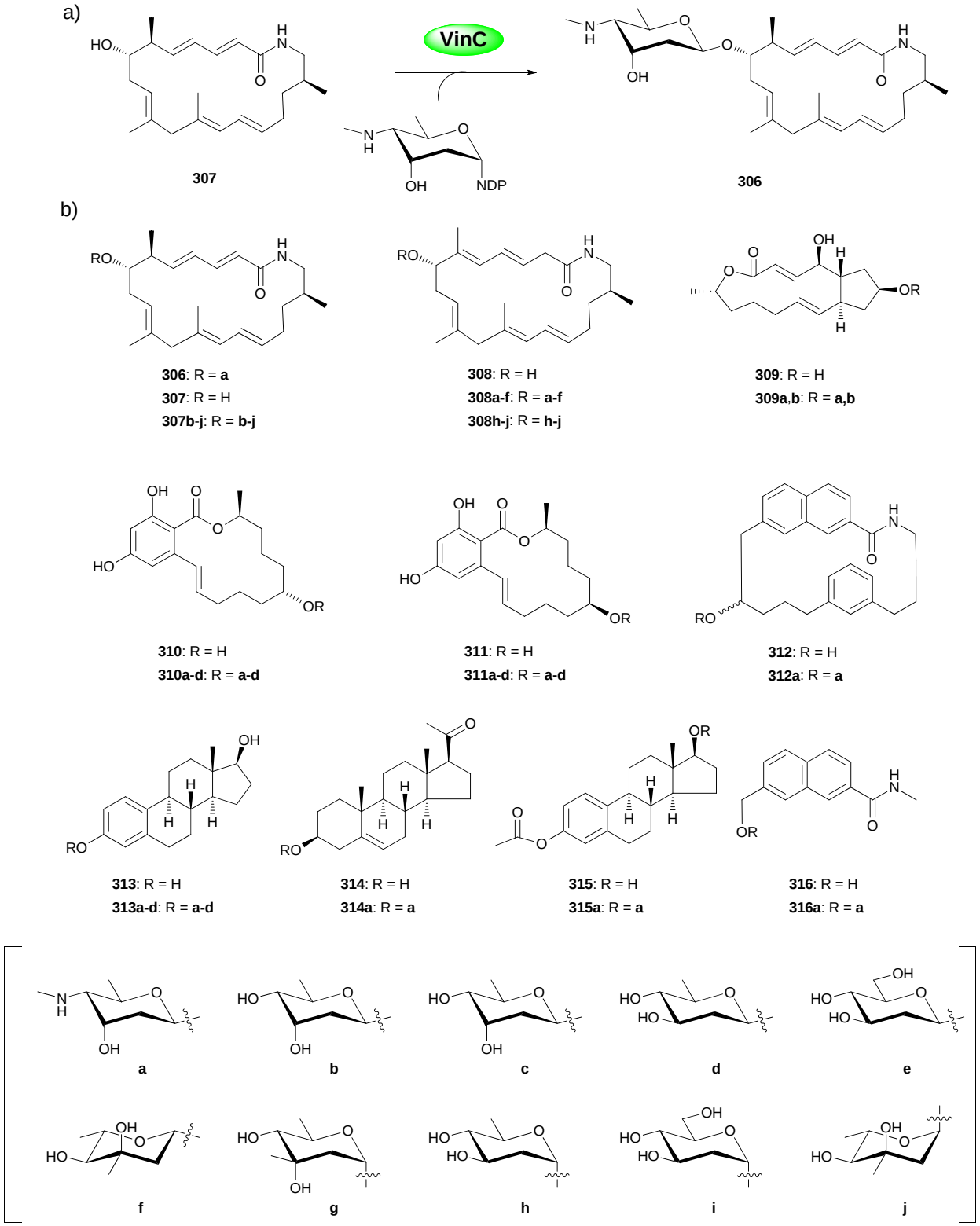


Figure 1.33. Aglycons recognized and products generated by the GT VinC. a) Wild-type VinC reaction. **b)** Alternative acceptors and their respective products generated by VinC.

(inversion) to yield the respective β- and α-glycosides **307b-j**, **308a-f**, **308h-j**, **309a**, **309b**, **310a-d**, **311a-d** and **313a-d** (**Figure 1.33**)[326]. Cumulatively, these studies reveal VinC to be promiscuous to both aglycon acceptors and NDP-sugar donors and thereby suggest VinC to have great potential in terms of glycorandomization/diversification.

1.5.1.7. Macrolactones

Macrolides are macrocyclic lactones that typically carry one or more attached deoxy sugars as represented by tylosin (**173**), methymycin (**317**), neomethymycin (**318**), erythromycin A (**319**), megalomicin A (**320**), oleandomycin (**321**), avermectin (**322**), glucosylated sorangicin A (**323**), and amphotericin B (**324**) (**Figure 1.34**). The classical macrolides (exemplified by erythromycins) are 12-, 14-, or 16-membered macrocyclic broad spectrum antibiotics which inhibit bacterial translation, a function heavily influenced, in most cases, by the macrolide's sugar moieties[327-331]. Certain macrolides have also more recently been found to show excellent immunomodulation for chronic pulmonary disorders (such as diffuse panbronchiolitis), implicating an added non-antibacterial anti-inflammatory effect[332]. Notably, glycosylation of classical macrolides can also either inactivate or alter the fundamental mechanism of these molecules. For example, macrolide glucosyltransferases (as exemplified by the GT OleD, discussed within section **1.5.1.7**) are a prototypical inactivation mechanism in macrolide-producing bacteria[333-335], while the addition of a third sugar (L-megosamine) to erythromycin A (**319**) yields megalomicin A (**320**), a macrolide which inhibits protein trafficking in the Golgi[336-338]. The following section describes the characterization of various macrolide GTs and applications for macrolide glycodiversification.

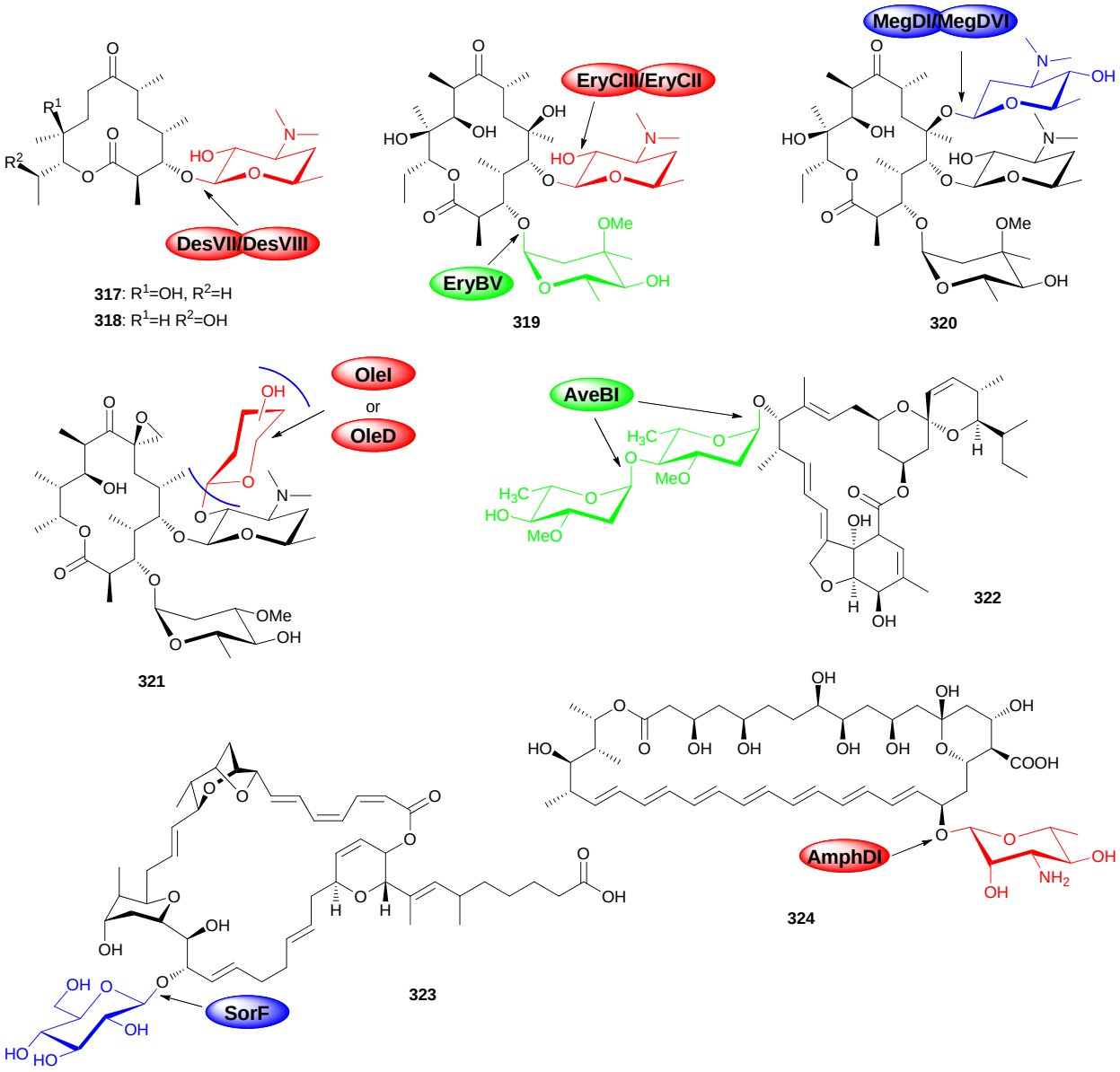


Figure 1.34. GTs involved in the biosynthesis of representative macrolides.

DesVII and EryCIII, macrolide GTs which require an additional protein partner for reconstitution of activity, have both been utilized in glycorandomization efforts. DesVII, in coordination with its partner protein DesVIII, catalyzes the transfer of desosamine to both 12- and 14-membered macrolactone acceptors during the biosynthesis of methymycin (**317**), neomethymycin (**318**) and pikromycin (**325**). Through both *in vivo* and *in vitro* experiments, DesVII/DesVIII was demonstrated to transfer D-desosamine to a wide range of aglycons including, but not limited to, macrolide scaffolds (*e.g.*, **326-332**), linear precursors (*e.g.*, **333-334**) and substrate mimics (*e.g.*, **335**), leading to the production of **326a**, **317**, **318**, **329a-332a**, **336-338**, **334a** and **335a** (**Figure 1.35**; see also section **1.4.2**)[151, 339-341]. Of the 24 sugar nucleotides investigated *in vitro* with DesVII/DesVIII, 7 were found to be substrates[339], highlighting an ability to process both D- and L- sugars and a requirement for 6-deoxysugars. In conjunction with various aglycons, this sugar specificity study ultimately led to the production of **317**, **318**, **326a-g**, **327b-g**, **328b-g**, **330a-g**, **331a-g** (**Figure 1.35**). Additionally, *in vivo* pathway engineering efforts further support the permissive nature of DesVII/DesVIII[151, 253, 255, 342-344]. In a similar fashion, the erythronolide GT EryCIII, along with its protein partner EryCII, was demonstrated to accommodate the native donor for D-desosamine (to produce **319**) as well as non-native sugar donors for D-mycaminose[345], D-angolosamine and L-mycarose[346], enabling the production of the new macrolides **339b-d** (**Figure 1.36**).

In contrast to the other macrolide GTs highlighted here, EryBV (from erythromycin biosynthesis) and OleI and OleD (of oleandomycin biosynthesis) do not require a partner protein for reconstitution of activity. Due to the lack of the native substrate, TDP-β-L-mycarose, *in vitro* EryBV activity was established via a 'sugar swapping' format wherein the incubation of

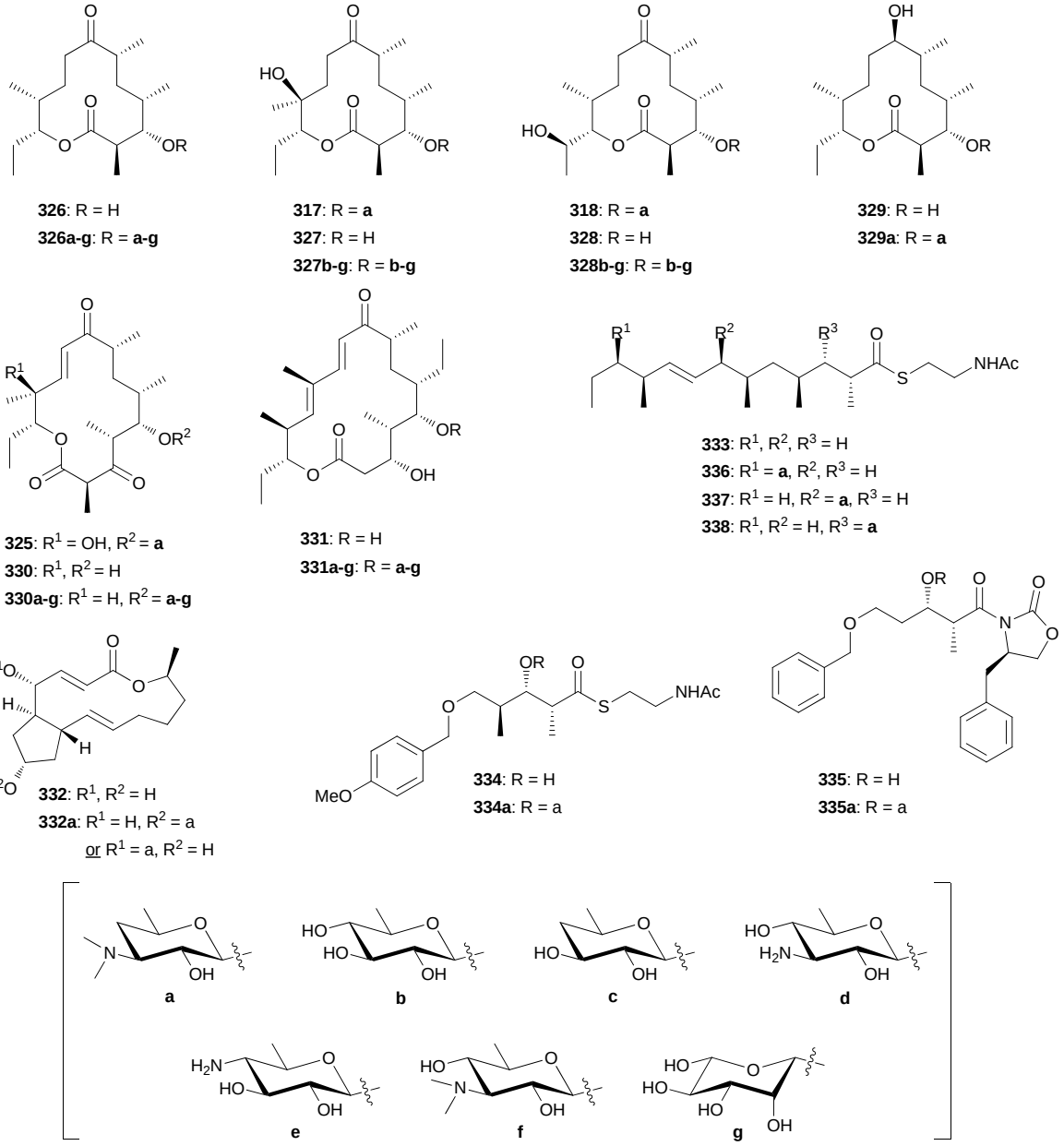


Figure 1.35. Products catalyzed by DesVII/DesVIII and their respective aglycons.

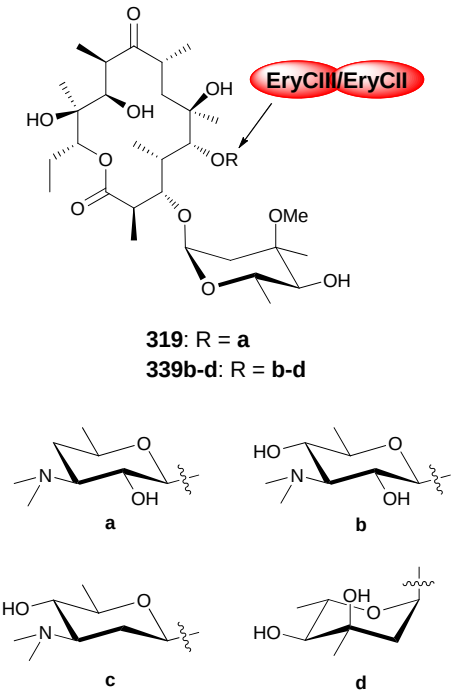


Figure 1.36. Analogues generated by EryCIII/EryCII.

EryBV with 22 putative NDP-sugars in the presence of aglycon **340** revealed the low level production of 11 glycosides (**340a-k**) (**Figure 1.37**)[241]. Application of EryBV in one- and two-GT 'aglycon exchange' reactions also led to corresponding erythromycin analogues[241]. Corresponding studies on the macrolide-inactivating GTs OleD and OleI, responsible for macrolide desosamine 2'-OH glucosylation[333, 334, 347], revealed subtle distinctions in substrate specificity between these GTs. Specifically, while OleD and OleI recognized **321**, OleD was found to also act on tylosin (**173**), erythromycin (**319**), lankamycin (**341**) and carbomycin (**342**) (**Figure 1.38**)[334, 347]. In a subsequent study, OleD and OleI were also demonstrated to glucosylate a number of small molecules (**180**, **202**, **343-349**) and accept sugars in addition to glucose for formation of **321a-e** or **321a-d** and **321f**, respectively (**Figure 1.38**)[348]. This notably permissive nature of OleD also has served as a basis for engineered variants that display remarkable flexibility toward a wide range of both donors and aglycons (see sections **1.5.2.1** and **1.5.2.2** and Chapter 2).

Avermectins (*e.g.*, **322**, **Figure 1.34**) are 16-membered macrocyclic lactones targeting the γ-aminobutyric acid related chloride ion channels unique to nematodes, insects, ticks and arachnids[349]. Thus, avermectins have been widely used as veterinary antiparasitic agents and expanded in clinical applications including treatment of onchocerciasis and lymphatic filariasis[349]. *In vivo* studies suggested that the two avermectin oleandrosyl residues in **322** were iteratively assembled in a step-wise, tandem manner by a single GT AveBI[284, 350, 351]. The AveBI-catalyzed iterative addition of L-oleandrose was unequivocally demonstrated *in vitro* via an aglycon exchange reaction, capitalizing upon the *in situ* generation of TDP-β-L-oleandrose by AveBI reverse catalysis[239]. In a designed two-GT aglycon exchange reaction, the L-mycarosyl residue was excised from an erythromycin analogue by the GT EryBV to afford

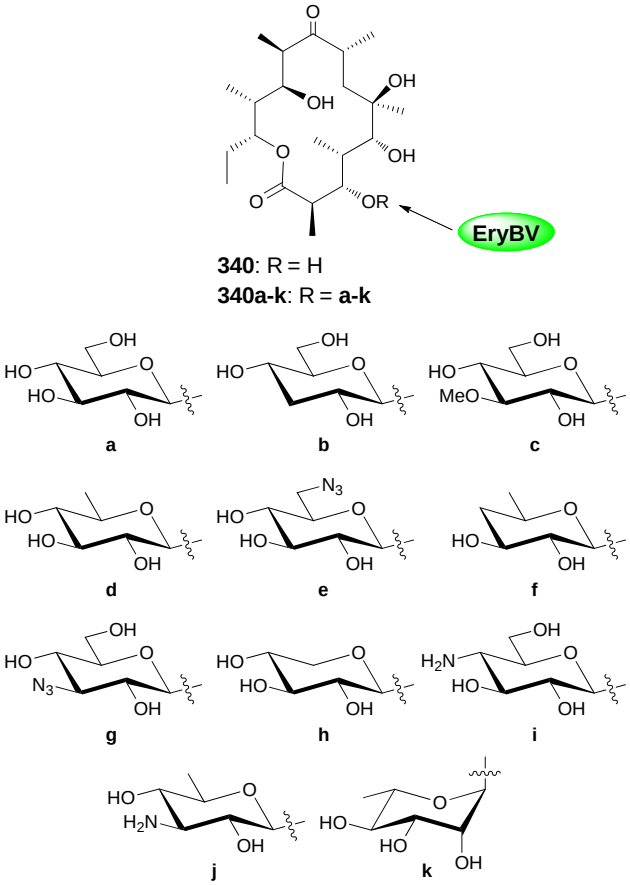


Figure 1.37. Analogues generated by EryBV.

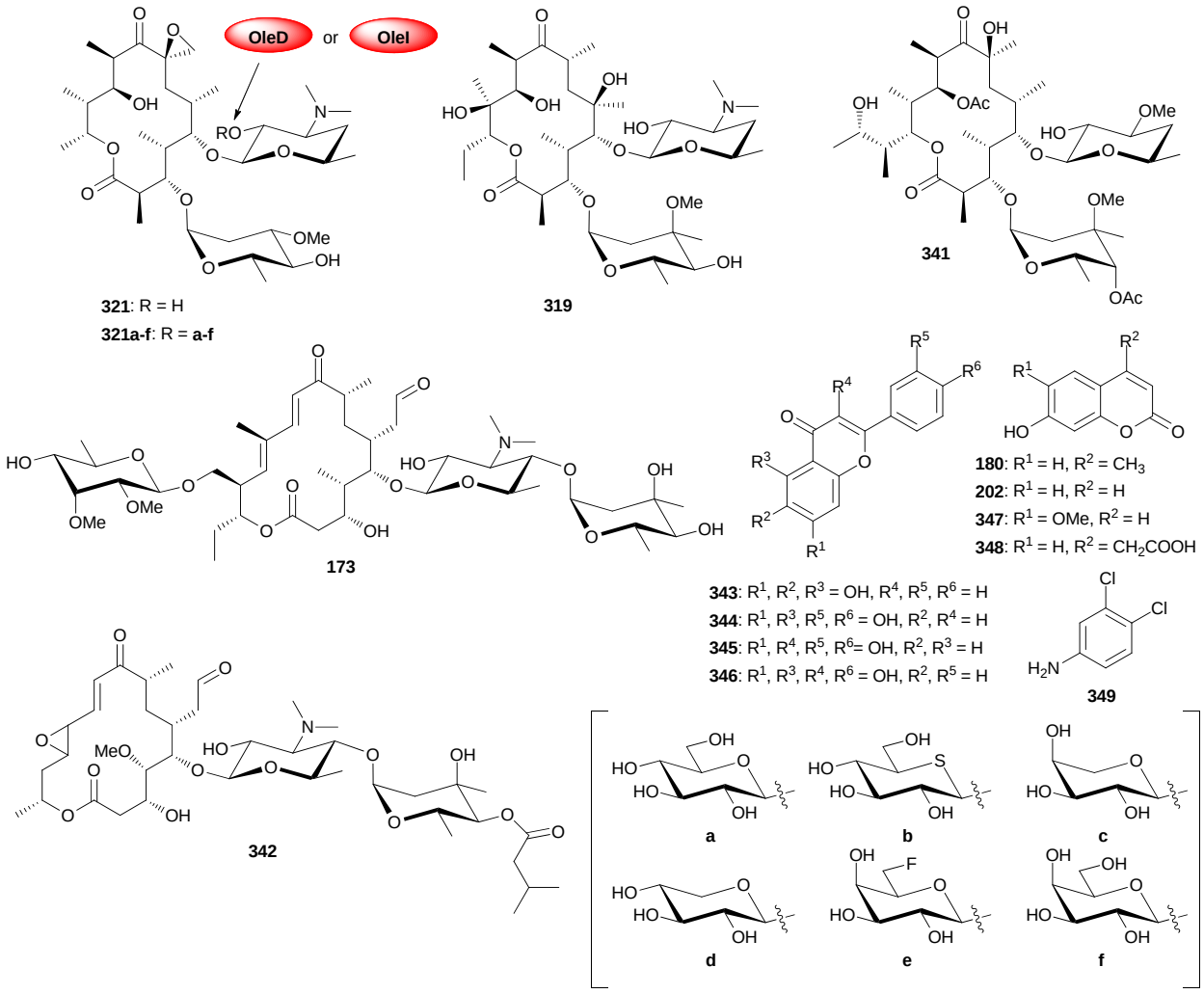


Figure 1.38. Aglycons recognized and products generated by OleD and/or OleI.

TDP-β-L-mycarose, which was then utilized *in situ* by AveBI to produce both mono-mycarosyl and di-mycarosyl avermectin analogues **350** and **351** from aglycon **352** (**Figure 1.39**)[241]. Subsequently, the donor substrate specificity of AveBI was probed with a small set of 22 TDP-sugars, revealing 10 non-native sugar substrates and suggesting a moderate promiscuity of AveBI toward sugar nucleotides[239]. Cumulatively, AveBI, in conjunction with 5 aglycons (**352-356**) and 10 sugar nucleotide donors, led to the production of 50 avermectin variants (**352a-j** to **356a-j**; **Figure 1.39**)[239], clearly illustrating the potential of this enzyme for glycorandomization/diversification.

Sorangicin A (**357**; **Figure 1.40**), a macrolide of polyketide origin, is produced by *Sorangium cellulosum* (strain So ce12) and represents one of the most potent myxobacterial antibiotics identified to date[352], the mechanism of action of which derives from a rifampicin-like inhibition of bacterial RNA-polymerase[353]. Biochemical characterization of the sorangicin GT SorF revealed an ability to glucosylate sorangicin A *in vitro* using both UDP- and TDP-α-D-glucose to afford a biologically inactive product (**323**) (**Figure 1.40**)[354], implicating SorF as a putative self-resistance GT reminiscent to the macrolide-inactivating GTs OleD and OleI[334]. SorF also accepted 4 additional sugar nucleotide donors, comprising of both D- and L-sugars, to produce a total of 5 sorangicin analogues (**323**, **358b-e**; **Figure 1.40**)[354]. Thus, the demonstrated permissive nature of SorF suggests potential in the context of future glycorandomization/diversification strategies.

Polyene antibiotics, predominantly produced by *Streptomyces*, comprise a vast family of natural products and are represented by amphotericin B (**324**), nystatin A1 (**359**), pimaricin (**360**) and candicidin/FR008 (**361**) (**Figure 1.41**)[355, 356]. Members of this class are clinically used for their broad spectral antifungal properties but some of them also exhibit antivirus, antiprotozoal,

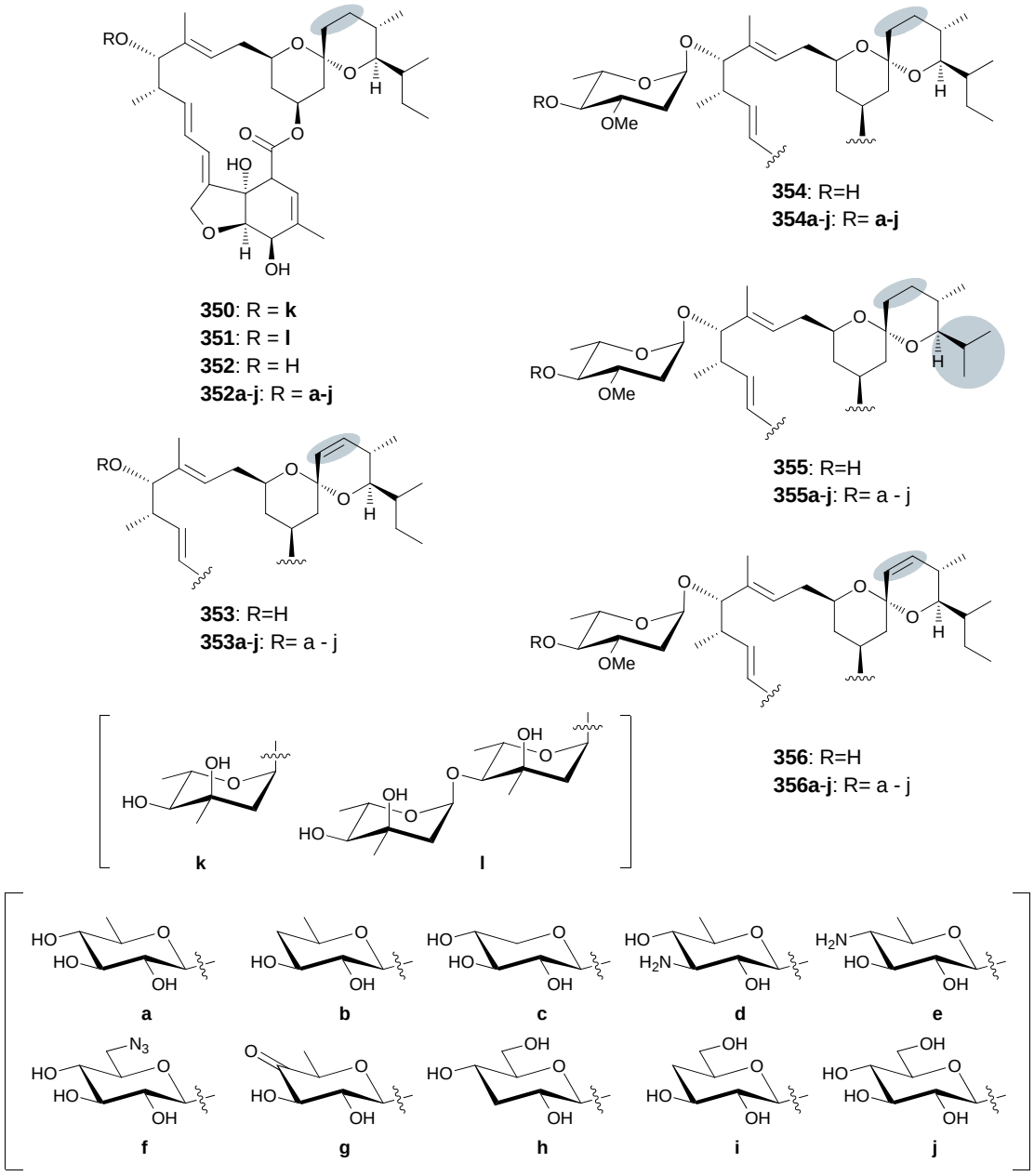


Figure 1.39. Aglycons recognized and products generated by AveBI. Subtle differences between substrates are shaded.

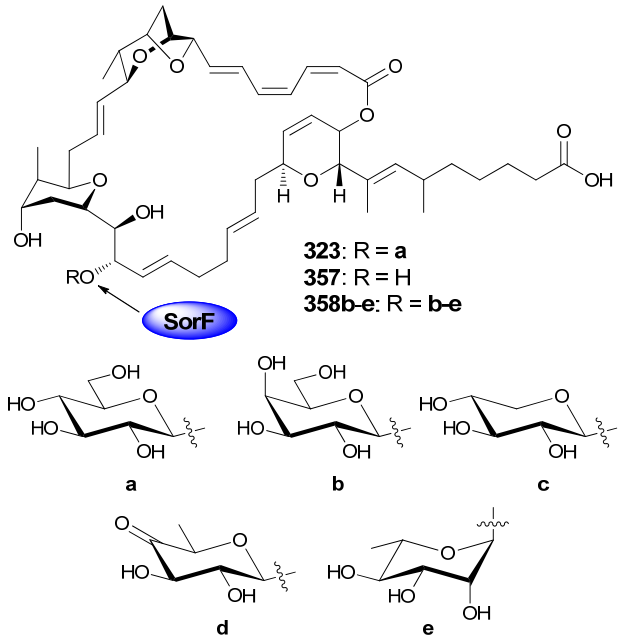


Figure 1.40. Products generated by SorF.

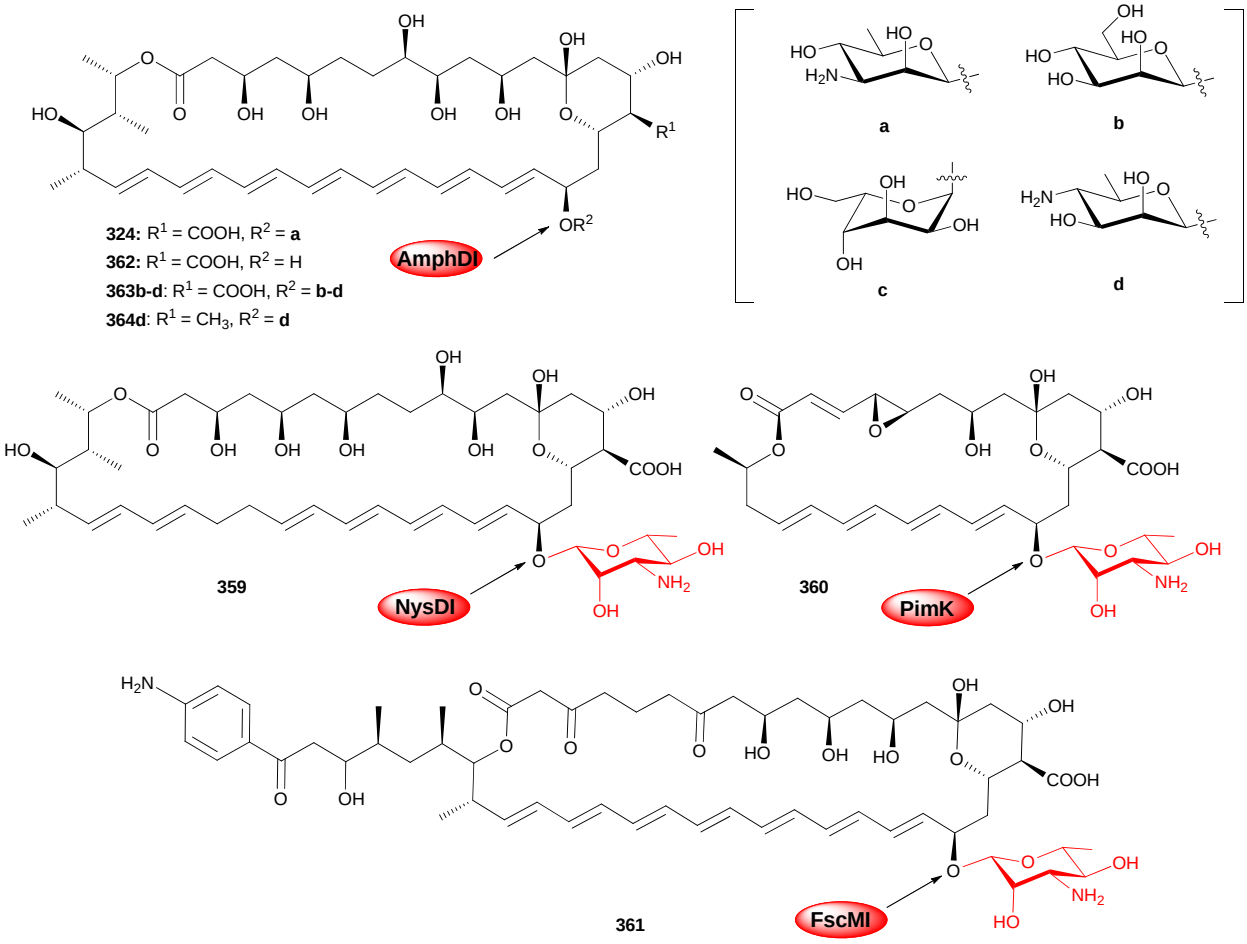


Figure 1.41. GTs involved in the biosynthesis of polyenes and generated analogues.

and antiprion activity[357, 358]. Due to their mechanism of action, even after a half century of clinical use, there exist few reported examples of clinical resistance to polyenes[359]. Similar to the importance of sugars appended to most glycosylated macrolides, the polyene aminosugar is critical for antifungal activity where it contributes key molecular interactions with the targeted fungal membrane sterols[360-363]. Biosynthetic gene loci encoding for **324** and **359-361** have been fully or partially characterized[364-368] and the corresponding role of encoded GTs (AmphDI, NysDI, FscMI and PimK) as the D-mycosaminyltransferase was first deduced via *in vivo* studies[368, 369]. Initial biochemical characterization of AmphDI and NysDI relied upon GT reverse catalysis wherein both GTs were found to remove the d-mycosamine of **324**, **359-361** in a GDP-dependent manner to provide the corresponding aglycons (*e.g.*, **362**, **Figure 1.41**)[206]. Of 21 corresponding GDP-sugar donors tested, AmphDI and NysDI accepted only GDP-D-mannose and GDP-L-gulose to afford novel polyene glycosides **363b-c** (**Figure 1.41**)[206]. In a subsequent *in vivo* study, a hybrid GT comprised of the aglycon recognition (*N*-terminal) domain of AmphDI and the sugar donor (*C*-terminal) recognition domain of the polyene perimycin GT PerDI enabled the transfer of perosamine to two amphotericin aglycons to generate **363d** and **364d**[370]. Collectively, these studies suggest the inherent stringency of polyene GTs may require additional engineering to provide catalysts capable of broader polyene glycodiversification.

1.5.1.8. Tetramic acids

The structural distinction of tetramic acids is based upon the presence of a pyrrolidine-2,4-dione heterocyclic core. Members of this class display a diverse variety of biological activities that includes antibiotic, antiviral, antifungal and antitumor activities[371]. Streptolydigin (**365a**), produced by *Streptomyces lydicus*[372], is synthesized through a hybrid type I polyketide

synthase non-ribosomal peptide synthetase system and displays potent activity against both bacterial RNA polymerase and eukaryotic terminal deoxynucleotidyl transferase[371]. Only a single GT gene, *slgG*, was identified within the streptolydigin gene cluster. Through an *in vivo* engineering studies, SlgG was demonstrated to catalyze *N*-glycosylation of the streptolydigin tetramic acid core with 3 distinct l- and 2-deoxy-D-sugars to yield **365a-e** (**Figure 1.42**), although the specific anomeric stereochemistry among all species was not determined[371]. This study is one of only a few examples of *N*-glycosyl-based glycorandomization and suggests SlgG to possess a permissive sugar donor profile beneficial to future glycorandomization efforts.

1.5.2. Libraries based upon uniquely permissive glycosyltransferases

While the examples of glycorandomization described above cover an impressive array of natural product classes and naturally occurring glycosidic bonds (e.g. *O*-, *S*-, *N*- and *C*-glycosides), GT specificity in nearly all cases to date remains closely tied to a catalyst's inherent biosynthetic function. To circumvent this restriction, recent efforts have focused on the generation of catalysts capable of glycosylating molecules for which native biosynthetic GTs do not exist. Herein we describe efforts toward both the generation and application of such broadly permissive GTs.

1.5.2.1. Generation of permissive GTs

While there are many screens available for GT activity[132, 135, 138, 373], relatively few have been specifically applied toward the generation of uniquely promiscuous GTs. High throughput screening for glycosyltransfer presents a challenge as most GT substrates/reactants lack

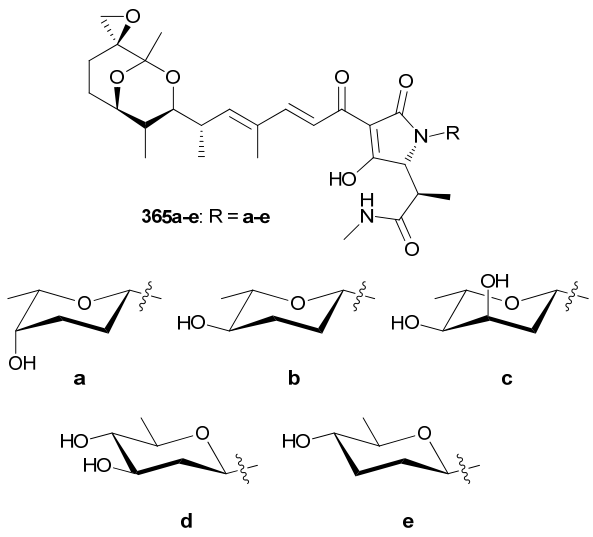


Figure 1.42. Streptolydigin and glycosylated analogues generated by SlgG.

measurable fluorescence or absorbance changes upon glycosyltransfer. Thus, the development of high throughput GT screens typically have relied upon modified substrates, which bias the outcome toward a narrowly defined GT reaction. Furthermore, most GT screens to date have been limited to screening against commercially available NDP-sugars or specially designed substrates in an initial high throughput screen wherein promiscuity against broader substrate panels must be subsequently assessed via more standard low throughput assays. Herein we summarize the assays and approaches utilized to identify or generate permissive GTs which act upon small molecules.

The Green-Amber-Red (GAR) mass spectrometry based assay developed by Davis and co-workers enabled the ability to rapidly delineate positive/negative glycosyltransfer in the presence of libraries of sugar donors or aglycon acceptors[348]. This method enabled the evaluation of the sugar donor and/or aglycon acceptor promiscuity of 6 individual wild-type GTs (UGT72B1, β-GalT from bovine, OleI, OleD, MGT and *Vv*GT1)[348, 374, 375]. While this approach was relatively costly and suffered from the standard limitations of mass spectrometry (*e.g.*, buffer suppression and the need for optimization for adequate detection of certain molecular species), the method presented the potential for a rapid end point assay amendable to a nearly limitless array of substrates.

The net release of protons during the glycosyltransferase reaction cycle provides the basis for pH indicator-based colorimetric assays amenable to high throughput screening. In an assay based upon phenol red, the substrate specificities of three separate GTs were investigated against 7 NDP-sugar donors and 2 acceptors for formation of disaccharides[376]. Palcic and Persson utilized a bromothymol blue-based colorimetric assay to identify variants from a single amino acid saturation library of the α-1,3-galactosyltransferase GTB which favored the non-native

substrate UDP-GlcNAc[136]. Utilizing a screen based upon cresol red, Kim and co-workers screened a library of incremental GT chimera, based upon the kanamycin GT KanF and vancomycin GT GtfE, for catalysts capable of glycosylating the 2-deoxystreptamine core of kanamycin[377]. This colorimetric screen identified HMT31, a GT chimera with expanded sugar donor promiscuity over the parental 2-deoxystreptamine GT KanF. While pH-based enzyme assays can by limited by sensitivity of the chosen indicator and/or a lack of sufficient signal-to-noise, these assays are notably based upon a general property of all GT reactions (net release of protons) and thus offer broad applicability.

A fluorescence-based assay developed by Thorson and co-workers was key to generating proficient (and permissive) variants of the GT OleD. This simple assay was based upon the masking the inherent fluorescence of 4-methylumbelliferone (**180**) upon C-7 *O*-glycosylation[133]. Using this assay, three improved OleD variants were identified from a library of ~1000 OleD random mutants, which ultimately led to a triple mutant (P67T/S132F/A242V; variant 'ASP') with 60-fold more efficiency toward the unnatural acceptor **180** and UDP-α-D-glucose through combination of functional mutations[133]. This increase in proficiency also led to a dramatic increase in promiscuity toward both donors and acceptors. Specifically, from 22 examined NDP-sugar donors, the 'ASP' variant accepted 15 NDP-sugars to produce **366-379**, while wild-type OleD only catalyzed detectable production of **366**, **369** and **370** (**Figure 1.43**)[133]. Moreover, the 'ASP' variant displayed improvements of glucosylation toward several acceptors including **199**, **346**, **349** and **380-382** (**Figure 1.43**)[133]. Saturation of position 242 (in context of OleD double mutant P67T/S132F) in a follow-up investigation identified OleD variant '3-1H12' (P67T/S132F/A242L) that was three-fold more active with UDP-α-D-glucose (in terms of

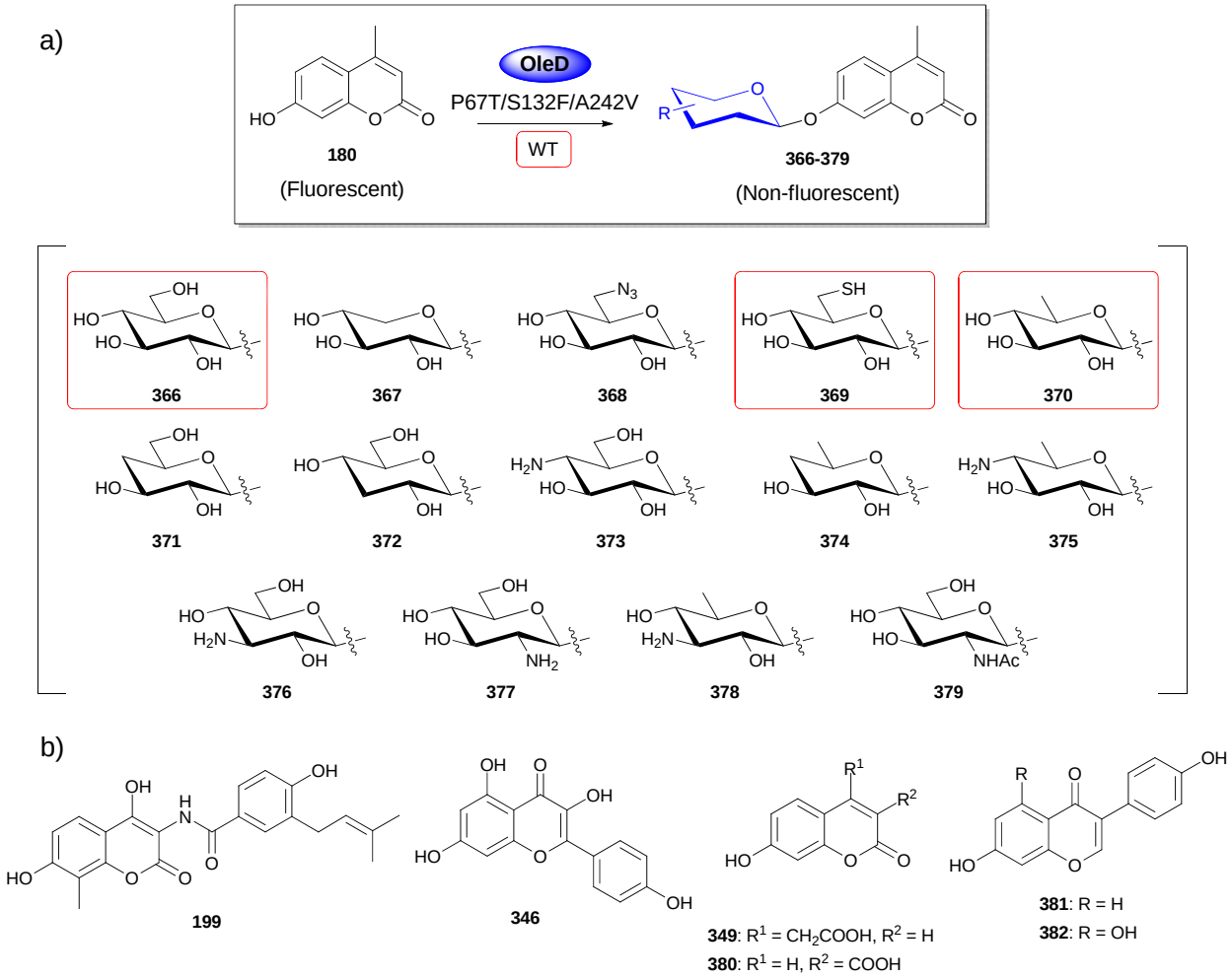


Figure 1.43. Screening and aglycon promiscuity of the OleD variant 'ASP'. a) The highthroughput fluorescent screen utilized to identify OleD variant 'ASP'. **b)** Aglycons recognized by the OleD variant 'ASP'.

specific activity) than the parent 'ASP' (P67T/S132F/A242V) and ten-fold more active than wild-type OleD[134]. Additional saturation mutagenesis and low throughput screening led to the identification of the OleD variant 'TDP-16' (P67T/132F/A242L/Q268V) that was 1.5-fold improved towards TDP-α-D-glucose[248].

Of particular interest, the observed improvement of OleD variant 'ASP' toward the novobiocin aglycon (**199**)[133] also presented a potential starting point for circumventing the stringency of the native noviosyltransferase NovM for novobiocin glycorandomization (see section **5.1.1.**)[203]. A comprehensive 'hot spot' saturation mutagenesis of the OleD residues identified via the umbelliferone-based study[133, 134], in conjunction with a low throughput HPLC-based screen for **199** glucosylation, led to identification of OleD variant '1C9' (P67T/I112T/A242V) with activity improvements of several hundred-fold toward production of **383a**[212]. Remarkably, the NDP-sugar donor promiscuity of 'ASP' was retained in the context of **199** to ultimately provide 10 differentially glycosylated coumarin analogues (**199b**, **199c**, **383a**, **383d-j**; **Figure 1.44**), only 4 (**199b**, **199c**, **383a**, **383d**; **Figure 1.44**) or 2 (**199b**, **199c**; **Figure 1.15** and **Figure 1.44**) of which were accessible by wild-type OleD or NovM-based catalysis, respectively[212].

1.6. Summation

Given the validated impact of complex carbohydrates and glycosides in drug discovery, practical platforms to expedite and simplify natural product and/or drug glycosylation are anticipated to continue to enhance drug discovery efforts. In addition, the tools, reagents and technologies developed for such purposes are expected to advance other areas of fundamental and applied glycobiology research. In addition to the continued incremental fundamental

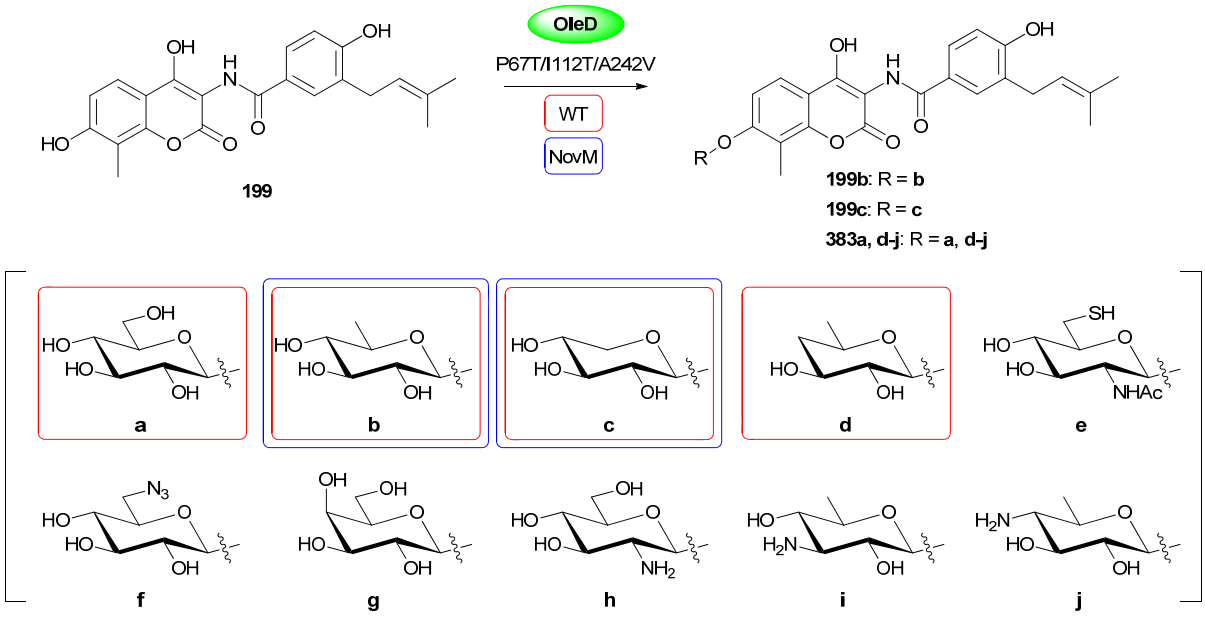


Figure 1.44. Novobioicn analogues generated by the OleD variant '1C9'. Compounds in red and/or blue boxes were capable of being generated by either wild-type OleD and/or NovM, respectively.

glycorandomization catalyst/platform advancements, studies to specifically decipher novel 'sugar codes' in the context of drug absorption, distrubution, metabolism, excretion and toxicity (*i.e.*, ADMET properties) toward predictive chemoinformatics as well as innovative new strategies to integrate the synthesis and evaluation process are expected to dramatically influence the field. Furthermore, the conceptual framework for glycorandomization should serve to inspire an expansion of natural product/drug diversification using many other potential tailoring enzyme-catalyzed processes including, but not limited to, acylation, alkylation, halogenation, oxidationn, phosphorylation and/or sulfation.

1.7. Thesis Overview

While numerous routes to NDP-sugars and efforts towards glyco(diversification/randomization) are discussed in this introduction, examining transfer with more than a dozen NDP-sugars within a single study remains the exception rather than the rule and there are only three GTs, ElmGT[144, 145, 146, 285], GtfE[238, 302, 307], and OleD[133, 375] that have been characterized as capable of transferring more than a dozen sugars to a single aglycon. These small numbers prevent efforts into investigating the true breadth of biological diversity found throughout sugars appended to natural products[121, 142]. The simple fact remains that the general synthesis of NDP-sugars through chemical, enzymatic, and chemoenzymatic routes remains non-trivial and requires significant effort to access even small numbers of these important reagents. Therefore, glyco(diversification/randomization) efforts, and indeed, investigations into glycosylation events in general, continue to remain largely limited by access to NDP-sugars.

Out of all of the approaches to the synthesis of NDP-sugars discussed in this introduction, perhaps the most underappreciated is the utilization of GT-catalyzed reverse reactions. Such

single enzyme systems would greatly expedite the identification of new catalysts towards generation of high value NDP-sugars by eliminating the need for whole pathway or multi-enzyme engineering. Although there are numerous examples of GT reversibility throughout the literature[141, 206, 233–246], a number of significant challenges restrict the general application of this strategy – most notably the typical thermodynamic nature of these reactions is such that large numbers of NDP equivalents are needed to drive fromation of the targeted NDP-sugar to acceptable levels. Consistent with this, the application of GT reverse reactions has mainly been limited production of single, unique NDP-sugars from complex natural products during biochemical characterization of GTs. However, an inverting GT identified from *Streptomyces antibioticus*, OleD, has the potential to address these issues and establish GT-catalyzed reverse reactions as a standard chemoenzymatic approach for NDP-sugar synthesis. A member of the GT-B superfamily and initially characterized as glucosylating the macrolide oleandomycin as a self-protection mechanism for the host organism[347, 378], further investigation suggests the ability of the wild type enzyme to recognize a range of aglycon substrates in forward GT reactions[375]. With the aid of an available crystal structure[378], several directed evolution and engineering studies paved the way to improved variants with expanded proficiency and promiscuity towards both NDP-sugars and aglycon acceptors[133, 134, 212, 248]. Taken together, these previous reports set the stage to address the current limitations of GT-catalyzed reverse reactions.

Towards this goal, the following chapters begin with a study of the potential of OleD to serve as a general catalyst for forward GT reacions in the context of drug discovery (Chapter 2). Subsequently, the specific ability of OleD to participate in GT-catalyzed reverse reactions as a general chemoenzymatic approach for NDP-sugar synthesis, and, as part of this study, the development and application of general glycosyl transfer screens is described (Chapter 3).

Finally, the development and application of a high throughput screen to identify OleD variants with improved proficiencies toward both specific glycoside donors and NDP acceptors is presented, culminating in a demonstration of the capacity for combinatorial synthesis of NDP-sugars in in GT-catalyzed reverse reactions (Chapter 4). Taken together, these investigations provide enormous inroads toward providing facile routes to NDP-sugar synthesis, general screens for glycosyl transfer, and enhanced catalysts relevant to both drug discovery and, more generally, the field of glycobiology.

1.8. References

1. Cipolla, L. *et al.* Discovery and design of carbohydrate-based therapeutics. *Exp. Opin. Drug Discov.* **5**, 721-737 (2010).

2. Kiessling, L.L. & Splain, R.A. Chemical approaches to glycobiology. *Annu. Rev. Biochem.* **79**, 619-653 (2010).

3. Lepenies, B., Yin, J. & Seeberger, P.H. Applications of synthetic carbohydrates to chemical biology. *Curr. Opin. Chem. Biol.* 14, 404-411 (2010).

4. Agard, N.J. & Bertozzi, C.R. Chemical approaches to perturb, profile, and perceive glycans. *Acc. Chem. Res.* **42**, 788-797 (2009).

5. Boltje, T.J., Buskas, T., & Boons, G.J. Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nat. Chem.* **1**, 611-622 (2009).

6. Zhang, M.Q. & Wilkinson, B. Drug discovery beyond the 'rule-of-five'. *Curr. Opin. Biotechnol.* **18**, 478-488 (2007).

7. Ganesan, A. The impact of natural products upon modern drug discovery. *Curr. Opin. Chem. Biol.* **12**, 306-317 (2008).

8. Ferla, B.L. *et al.* Natural glycoconjugates with antitumor activity. *Nat. Prod. Rep.* **28**, 630-648 (2011).

9. Schitter, G. & Wrodnigg, T.M. Update on carbohydrate-containing antibacterial agents. *Exp. Opin. Drug Discov.* **4**, 315-356 (2009).

10. Křen, V. & Řezanka, T. Sweet antibiotics - the role of glycosidic residues in antibiotic and antitumor activity and their randomization. *FEMS Microbiol. Rev.* **32**, 858-889 (2008).

11. Luzhetskyy, A., Méndez, C., Salas, J.A. & Bechthold, A. Deoxysugars in bioactive natural products: development of novel derivatives by altering the sugar pattern. *Curr. Top. Med. Chem.* **8**, 680-709 (2008).

12. Cipolla, L. & Peri, F. Carbohydrate-based bioactive compounds for medicinal chemistry applications. *Mini Rev. Med. Chem.* **11**, 39-54 (2011).

13. Ernst, B. & Magnani, J.L. From carbohydrate leads to glycomimetic drugs. *Nat. Rev. Drug Discov.* **8**, 661-677 (2009).

14. Newman, D.J. & Cragg, G.M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **70**, 461-477 (2007).

15. Butler, M.S. Natural products to drugs: natural product-derived compounds in clinical trials. *Nat. Prod. Rep.* **25**, 475-516 (2008).

16. Weymouth-Wilson, A.C. The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* **14**, 99-110 (1997).

17. Křen, V. & Martinkova, L. Glycosides in medicine: "The role of glycosidic residue in biological activity". *Curr. Med. Chem.* **8**, 1303-1328 (2001).

18. Thorson, J.S., Hosted Jr., T.J., Jiang, J., Higgins, J.B. & Ahlert, J. Nature's carbohydrate chemists: the enzymatic glycosylation of bioactive bacterial metabolites. *Curr. Org. Chem.* **5**, 139-167 (2001).

19. Giannini, G. Fluorinated anthracyclines: synthesis and biological activity. *Curr. Med. Chem.* **9**, 687-712 (2002).

20. Minotti, G., Menna, P., Salvatorelli, E., Cairo, G. & Gianni, L. Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacol. Rev.* **56**, 185-229 (2004).

21. Larsen, A.K., Escargueil, A.E. & Skladanowski, A. Catalytic topoisomerase II inhibitors in cancer therapy. *Pharmacol. Ther.* **99**, 167-181 (2003).

22. Carvalho, C. Doxorubicin: the good, the bad and the ugly effect. *Curr. Med. Chem.* **16**, 3267-3285 (2009).

23. Leonard, R.C., Williams, S., Tulpule, A., Levine, A.M. & Oliveros, S. Improving the therapeutic index of anthracycline chemotherapy: focus on liposomal doxorubicin (Myocet). *Breast* **18**, 218-224 (2009).

24. Smith, L.A. *et al.* Cardiotoxicity of anthracycline agents for the treatment of cancer: systematic review and meta-analysis of randomised controlled trials. *BMC Cancer* **10**, 337 (2010).

25. Thorens, B. & Mueckler, M. Glucose transporters in the 21st century. *Am. J. Physiol. Endocrinol. Metab.* **298**, E141-E145 (2010).

26. Qutub, A.A. & Hunt, C.A. Glucose transport to the brain: a systems model. *Brain Res. Brain Res. Rev.* **49**, 595-617 (2005).

27. Fernández, C. *et al.* Synthesis and biological studies of glycosyl dopamine derivatives as potential antiparkinsonian agents. *Carbohydr. Res.* **327**, 353-365 (2000).

28. Fernández, C. *et al.* Synthesis of glycosyl derivatives as dopamine prodrugs: interaction with glucose carrier GLUT-1. *Org. Biomol. Chem.* **1**, 767-771 (2003).

29. García-Alvarez, I., Garrido, L. & Fernández-Mayoralas, A. Studies on the uptake of glucose derivatives by red blood cells. *ChemMedChem* **2**, 496-504 (2007).

30. Bonina, F. *et al.* Glycosyl derivatives of dopamine and L-dopa as anti-Parkinson prodrugs: synthesis, pharmacological activity and *in vitro* stability studies. *J. Drug Target* **11**, 25-36 (2003).

31. Christian, S.T. US. Pat., 6,548,484, 2003.

32. Christian, S.T. US. Pat., 7,345,031, 2003.

33. Jiang, C. *et al.* Dopaminergic properties and experimental anti-parkinsonian effects of IPX750 in rodent models of Parkinson disease. *Clin. Neuropharmacol.* **27**, 63-73 (2004).

34. Polt, R., Dhanasekaran, M. & Keyari, C.M. Glycosylated neuropeptides: a new vista for neuropsychopharmacology? *Med. Res. Rev.* **25**, 557-585 (2005).

35. Lowery, J.J. *et al.* Glycosylation improves the central effects of DAMGO. *Chem. Biol. Drug Des.* **69**, 41-47 (2007).

36. Schiff, P.B., Fant, J. & Horwitz, S.B. Promotion of microtubule assembly *in vitro* by taxol. *Nature* **277**, 665-667 (1979).

37. Altmann, K.H. & Gertsch, J. Anticancer drugs from nature--natural products as a unique source of new microtubule-stabilizing agents. *Nat. Prod. Rep.* **24**, 327-357 (2007).

38. McGrogan, B.T., Gilmartin, B., Carney, D.N. & McCann, A. *Biochim. Biophys. Acta* **1785**, 96-132 (2008).

39. Cragg, G.M. & Newman, D.J. A tale of two tumor targets: topoisomerase I and tubulin. The Wall and Wani contribution to cancer chemotherapy. *J. Nat. Prod.* **67**, 232-244 (2004).

40. Kingston, D.G.I. in *Anticancer Agents from Natural Products*. Cragg, G.M., Kingston, D.G.I. & Newman, D.J. Eds., pp. 89-122 (CRC Press LLC, Boca Raton, Fl, 2005).

41. Sénilh, V. *et al.* Mise en évidence de nouveau analogues du taxol extraits de *Taxus baccata*. *J. Nat. Prod.* **47**, 131-137 (1984).

42. Mandai, T. *et al.* Synthesis and biological evaluation of water soluble taxoids bearing sugar moieties. *Heterocyles* **32**, 561-566 (2001).

43. Kim, J.W. & Dang, C.V. Cancer's molecular sweet tooth and the Warburg effect. *Cancer Res.* **66**, 8927-8930 (2006).

44. Airley, R.E. & Mobasheri, A. Hypoxic regulation of glucose transport, anaerobic metabolism and angiogenesis in cancer: novel pathways and targets for anticancer therapeutics. *Chemotherapy* **53**, 233-256 (2007).

45. Macheda, M.L., Rogers, S. & Best, J.D. Molecular and cellular regulation of glucose transporter (GLUT) proteins in cancer. *J. Cell. Physiol.* **202**, 654-662 (2005).

46. Liu, D.Z., Sinchaikul, S., Reddy, P.V., Chang, M.Y. & Chen, S. T. Synthesis of 2'-paclitaxel methyl 2-glucopyranosyl succinate for specific targeted delivery to cancer cells. *Bioorg. Med. Chem. Lett.* **17**, 617-620 (2007).

47. Lin, Y.S. *et al.* Targeting the delivery of glycan-based paclitaxel prodrugs to cancer cells via glucose transporters. *J. Med. Chem.* **51**, 7428-7441 (2008).

48. Toole, B.P. & Slomiany, M.G. Hyaluronan: a constitutive regulator of chemoresistance and malignancy in cancer cells. *Semin. Cancer Biol.* **18**, 244-250 (2008).

49. Naor, D., Wallach-Dayan, S.B., Zahalka, M.A. & Sionov, R.V. Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. *Semin. Cancer Biol.* **18**, 260-267 (2008).

50. Luo, Y. & Prestwich, G.D. Synthesis and selective cytotoxicity of a hyaluronic acid-antitumor bioconjugate. *Bioconjug. Chem.* **10**, 755-763 (1999).

51. Luo, Y., Ziebell, M.R. & Prestwich, G.D. A hyaluronic acid-taxol antitumor bioconjugate targeted to cancer cells. *Biomacromolecules* **1**, 208-218 (2000).

52. Dumontet, C. & Jordan, M.A. Microtubule-binding agents: a dynamic field of cancer therapeutics. *Nat. Rev. Drug Discov.* **9**, 790-803 (2010).

53. Zhou, J. & Giannakakou, P. Targeting microtubules for cancer chemotherapy. *Curr. Med. Chem. Anticancer Agents* **5**, 65-71 (2005).

54. Jordan, M.A. & Wilson, L. Microtubules as a target for anticancer drugs. *Nat. Rev. Cancer.* **4**, 253-265 (2004).

55. Graening, T. & Schmalz, H.G. Total syntheses of colchicine in comparison: a journey through 50 years of synthetic organic chemistry. *Angew. Chem. Int. Ed Engl.* **43**, 3230-3256 (2004).

56. Wertheimer, A.I., Davis, M.W. & Lauterio, T.J. A new perspective on the pharmacoeconomics of colchicine. *Curr. Med. Res. Opin.* **27**, 931-937 (2011).

57. Cocco, G., Chu, D.C.C. & Pandolfi, S. Colchicine in clinical medicine. A guide for internists. *Eur. J. Intern. Med.* **21**, 503-508 (2010).

58. Suri, O.P., Gupta, B.D., Suri, K.A., Sharma A.K. & Satti, N.K. A new glycoside, 3-*O*-demethylcolchicine-3-*O*-α-D-glucopyranoside, from *Gloriosa superba* seeds. *Nat. Prod. Lett.* **15**, 217-219 (2001).

59. Bombardelli, E. US. Pat., 5,777,136, 1998.

60. Bombardelli, E. & Fontana, G. PCT Int. Appl., PIXXD2 WO 2004111068 A1 20041223, 2004.

61. Bombardelli, E. Eur. Pat. Appl., EPXXDW EP 789028 A2 19970813, 1997.

62. Ahmed, A. *et al.* Colchicine glycorandomization influences cytotoxicity and mechanism of action. *J. Am. Chem. Soc.* **128**, 14224-14225 (2006).

63. Goff, R.D. & Thorson, J.S. Enhancing the divergent activities of betulinic acid via neoglycosylation. *Org. Lett.* **11**, 461-464 (2009).

64. Peltier-Pain, P., Timmons, S.C., Grandemange, A., Benoit E. & Thorson, J.S. Warfarin glycosylation invokes a switch from anticoagulant to anticancer activity. *ChemMedChem* **6**, 1347-1350 (2011).

65. Goff, R.D., Watson Jr., J.A. & Thorson, J.S. unpublished data.

66. Wade, O.L. Digoxin 1785-1985. I. Two hundred years of digitalis. *J. Clin. Hosp. Pharm.* **11**, 3-9 (1986).

67. Repke, K.R.H. Toward the discovery of digitalis derivatives with inotropic selectivity *Drug Discov. Today* **2**, 110-116 (1997).

68. Kometiani, P., Liu, L. & Askari, A. Digitalis-induced signaling by Na+/K+-ATPase in human breast cancer cells. *Mol. Pharmacol.* **67**, 929-936 (2005).

69. Xie, Z. & Xie, J. The Na/K-ATPase-mediated signal transduction as a target for new drug development. *Front. Biosci.* **10**, 3100-3109 (2005).

70. Tian, J. *et al.* Binding of Src to Na+/K+-ATPase forms a functional signaling complex. *Mol. Biol. Cell* **17**, 317-326 (2006).

71. Liang, M. *et al.* Identification of a pool of non-pumping Na/K-ATPase. *J. Biol. Chem.* **282**, 10585-10593 (2007).

72. Li, Z. & Xie, Z. The Na/K-ATPase/Src complex and cardiotonic steroid-activated protein kinase cascades. *Pflugers Arch.* **457**, 635-644 (2009).

73. Langenhan, J.M., Peters, N.R., Guzei, I.A., Hoffmann, F.M. & Thorson, J.S. Enhancing the anticancer properties of cardiac glycosides by neoglycorandomization. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12305-12310 (2005).

74. Langenhan, J.M. *et al.* Modifying the glycosidic linkage in digitoxin analogs provides selective cytotoxins. *Bioorg. Med. Chem. Lett.* **18**, 670-673 (2008).

75. Juncker, T., Schumacher, M., Dicato, M. & Diederich, M. UNBS1450 from Calotropis procera as a regulator of signaling pathways involved in proliferation and cell death. *Biochem. Pharmacol.* **78**, 1-10 (2009).

76. Iyer, A.K. *et al.* A direct comparison of the anticancer activities of digitoxin *MeON*-neoglycosides and *O*-glycosides: oligosaccharide chain length-dependent induction of caspase-9-mediated apoptosis. *ACS Med. Chem. Lett.* 1, 326-330 (2010).

77. Hutchinson, C.R., Shekhani, M.C. & Prudent, J.R. WIPO Pat. Appl., WO/2010/017480 A1, 2010.

78. Kadokawa, J.I. Precision polysaccharide synthesis catalyzed by enzymes. *Chem. Rev.* **111**, 4308-4345 (2011).

79. Rakić, B. & Withers, S.G. Recent developments in glycoside synthesis with glycosynthases and thioglycoligases. *Aust. J. Chem.* **62**, 510-520 (2009).

80. Hancock, S.M., Vaughan, M.D. & Withers, S.G. Engineering of glycosidases and glycosyltransferases. *Curr. Opin. Chem. Biol.* **10**, 509-519 (2006).

81. Jahn, M. & Withers, S.G. New Approaches to Enzymatic Oligosaccharide Synthesis: Glycosynthases and Thioglycoligases. *Biocatal. Biotransfor.* **21**, 159-166 (2003).

82. Wang, L.X. & Huang, W. Enzymatic transglycosylation for glycoconjugate synthesis. *Curr. Opin. Chem. Biol.* **13**, 592-600 (2009).

83. Bertrand, A. *et al. Leuconostoc mesenteroides* glucansucrase synthesis of flavonoid glucosides by acceptor reactions in aqueous-organic solvents. *Carbohydr. Res.* **341**, 855-863 (2006).

84. Y. H. Moon, *et al.* Enzymatic synthesis and characterization of arbutin glucosides using glucansucrase from *Leuconostoc mesenteroides* B-1299CB. *Appl. Microbiol. Biotechnol.* **77**, 559-567 (2007).

85. Kim, Y.M. *et al.* Enzymatic synthesis of alkyl glucosides using *Leuconostoc mesenteroides* dextransucrase. *Biotechnol. Lett.* **31**, 1433-1438 (2009).

86. Kang, J. *et al.* Synthesis and characterization of hydroquinone fructoside using *Leuconostoc mesenteroides* levansucrase. *Appl. Microbiol. Biotechnol.* **83**, 1009-1016 (2009).

87. André, I., Potocki-Véronèse, G., Morel, S., Monsan, P. & Remaud-Siméon, M. in *Topics in Current Chemistry*, Rauter, A.P., Vogel, P. & Queneau, Y. Eds., pp. 25-48 (Springer, New York, 2010).

88. Mackenzie, L.F., Wang, Q., Warren, R.A. & Withers, S.G. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.*, **120**, 5583-5584 (1998).

89. Yang, M., Davies, G.J. & Davis, B.G. A glycosynthase catalyst for the synthesis of flavonoid glycosides. *Angew. Chem. Int. Ed Engl.* **46**, 3885-3888 (2007).

90. Jahn, M., *et al.* Thioglycosynthases: double mutant glycosidases that serve as scaffolds for thioglycoside synthesis. *Chem. Commun. (Camb)* **3**, 274-275 (2004).

91. Wilkinson, S.M., *et al. Escherichia coli* glucuronylsynthase: an engineered enzyme for the synthesis of beta-glucuronides. *Org. Lett.* **10**, 1585-1588 (2008).

92. Hancock, S.M., Rich, J.R., Caines, M.E., Strynadka, N.C. & Withers, S.G. Designer enzymes for glycosphingolipid synthesis by directed evolution. *Nat. Chem. Biol.* **5**, 508-514 (2009).

93. Kim, Y.W., Lee, S.S., Warren, R.A. & Withers, S.G. Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire. *J. Biol. Chem.* **279**, 42787-42793 (2004).

94. Shaikh, F.A. & Withers, S.G. Teaching old enzymes new tricks: engineering and evolution of glycosidases and glycosyl transferases for improved glycoside synthesis. *Biochem. Cell Biol.* **86**, 169-177 (2008).

95. Ben-David, A., Shoham, G. & Shoham, Y. A universal screening assay for glycosynthases: directed evolution of glycosynthase XynB2(E335G) suggests a general path to enhance activity. *Chem. Biol.* **15**, 546-551 (2008).

96. Kittl, R. & Withers, S.G. New approaches to enzymatic glycoside synthesis through directed evolution. *Carbohydr. Res.* **345**, 1272-1279 (2010).

97. Armstrong, Z., Reitinger, S., Kantner, T. & Withers, S.G. Enzymatic thioxyloside synthesis: characterization of thioglycoligase variants identified from a site-saturation mutagenesis library of *Bacillus circulans* xylanase. *Chembiochem* **11**, 533-538 (2010).

98. Kim, Y.W., Zhang, R., Chen, H. & Withers, S.G. *O*-glycoligases, a new category of glycoside bond-forming mutant glycosidases, catalyse facile syntheses of isoprimeverosides. *Chem. Commun. (Camb)* **46**, 8725-8727 (2010).

99. Kim, Y., *et al.* Expanding the thioglycoligase strategy to the synthesis of alpha-linked thioglycosides allows structural investigation of the parent enzyme/substrate complex. *J. Am. Chem. Soc.* **128**, 2202-2203 (2006).

100. Jahn, M., Marles, J., Warren, R.A. & Withers, S.G. Thioglycoligases: mutant glycosidases for thioglycoside synthesis. *Angew. Chem. Int. Ed Engl.* **42**, 352-354 (2003).

101. Müllegger, J., Jahn, M., Chen, H.M., Warren, R.A. & Withers, S.G. Engineering of a thioglycoligase: randomized mutagenesis of the acid-base residue leads to the identification of improved catalysts. *Protein Eng. Des. Sel.* **18**, 33-40 (2005).

102. Müllegger, J. *et al.* Thermostable glycosynthases and thioglycoligases derived from *Thermotoga maritima* beta-glucuronidase. *Chembiochem* **7**, 1028-1030 (2006).

103. Kim, Y.W., Chen, H., Kim, J.H. & Withers, S.G. Catalytic properties of a mutant beta-galactosidase from Xanthomonas manihotis engineered to synthesize galactosyl-thio-beta-1,3 and -beta-1,4-glycosides. *FEBS Lett.* **580**, 4377-4381 (2006).

104. Kim, Y.W. *et al.* Thioglycoligase-based assembly of thiodisaccharides: screening as beta-galactosidase inhibitors. *Chembiochem* **8**, 1495-1499 (2007).

105. Mieyal, J.J. & Abeles, R.H. in *The Enzymes*, Boyer, P.D. Ed., pp. 515-532 (Academic Press, New York, 1972).

106. Silverstein, R., Voet, J., Reed, D. & Abeles, R.H. Purification and mechanism of action of sucrose phosphorylase. *J. Biol. Chem.* **242**, 1338-1346 (1967).

107. Suzuki, M., Kaneda, K., Nakai, Y., Kitaoka, M. & Taniguchi, H. Synthesis of cellobiose from starch by the successive actions of two phosphorylases. *N. Biotechnol.* **26**, 137-142 (2009).

108. Nakajima, M., Nishimoto, M. & Kitaoka, M. Practical preparation of D-galactosyl-beta1-->4-L-rhamnose employing the combined action of phosphorylases. *Biosci. Biotechnol. Biochem.* **74**, 1652-1655 (2010).

109. Hiraishi, M. *et al.* Synthesis of highly ordered cellulose II *in vitro* using cellodextrin phosphorylase. *Carbohydr. Res.* **344**, 2468-2473 (2009).

110. Kawazoe, S., Izawa, H., Nawaji, M., Kaneko, Y. & Kadokawa, J. Phosphorylase-catalyzed *N*-formyl-alpha-glucosaminylation of maltooligosaccharides. *Carbohydr. Res.* **345**, 631-636 (2010).

111. Nakai, H., Dilokpimol, A., Abou Hachem, M. & Svensson, B. Efficient one-pot enzymatic synthesis of alpha-(1-->4)-glucosidic disaccharides through a coupled reaction

catalysed by *Lactobacillus acidophilus* NCFM maltose phosphorylase. *Carbohydr. Res.* **345**, 1061-1064 (2010).

112. Goedl, C., Schwarz, A., Minani, A. & Nidetzky, B. Recombinant sucrose phosphorylase from *Leuconostoc mesenteroides*: characterization, kinetic studies of transglucosylation, and application of immobilised enzyme for production of alpha-D-glucose 1-phosphate. *J. Biotechnol.* **129**, 77-86 (2007).

113. Kitao, S., *et al.* Kinetic studies on p-nitrophenyl-cellobioside hydrolyzing xylanase from *Cellvibrio gilvus*. *Biosci. Biotechnol. Biochem.* **57**, 2010-2015 (1993).

114. Kitao, S. & Sekine, H. α-D-Glucosyl transfer to phenolic compounds by sucrose phosphorylase from *Leuconostoc mesenteroides* and production of α-arbutin. *Biosci. Biotechnol. Biochem.* **58**, 38-42 (1994).

115. Kitao, S. & Sekine, H. Syntheses of two kojic acid glucosides with sucrose phosphorylase from *Leuconostoc mesenteroides*. *Biosci. Biotechnol. Biochem.* **58**, 419-420 (1994).

116. Shin, M.H., Cheong, N., Lee, J. & Kim, K.H. Transglucosylation of caffeic acid by a recombinant sucrose phosphorylase in aqueous buffer and aqueous-supercritical CO2 media. *Food Chem.* **115**, 1028-1033 (2009).

117. Nakai, H. *et al.* Rational engineering of *Lactobacillus acidophilus* NCFM maltose phosphorylase into either trehalose or kojibiose dual specificity phosphorylase. *Protein Eng. Des. Sel.* **23**, 781-787 (2010).

118. Varki, A. & Chrispeels, M.J. *Essentials of glycobiology* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1999).

119. Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A. & Dwek, R.A. Glycosylation and the immune system. *Science* **291**, 2370-2376 (2001).

120. Zhang, X.L. Roles of glycans and glycopeptides in immune system and immune-related diseases. *Curr. Med. Chem.* **13**, 1141-1147 (2006).

121. Thibodeaux, C.J., Melançon III, C.E. & Liu, H.W. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed Engl.*, **47**, 9814-9859 (2008).

122. Hart, G.W., Slawson, C., Ramirez-Correa, G. & Lagerlof, O. Cross talk between *O*-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. *Annu. Rev. Biochem.* **80**, 825-858 (2011).

123. Lairson, L.L., Henrissat, B., Davies, G.J. & Withers, S.G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521-555 (2008).

124. Roychoudhury, R. & Pohl, N.L. New structures, chemical functions, and inhibitors for glycosyltransferases. *Curr. Opin. Chem. Biol.* **14**, 168-173 (2010).

125. Chang, A., Singh, S., Phillips, G.N. & Thorson, J.S. Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Curr. Opin. Biotechnol.* **22**, 800-808 (2011).

126. Breton, C., Snajdrova, L., Jeanneau, C., Koca, J. & Imberty, A. Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**, 29R-37R (2006).

127. Schuman, B., Alfaro, J.A. & Evans, S.V. Glycosyltransferase structure and function. *Top. Curr. Chem.* **272**, 217-257 (2007).

128. Soya, N., Fang, Y., Palcic, M.M. & Klassen, J.S. Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology* **21**, 547-552 (2010).

129. Errey, J.C. *et al*. Mechanistic insight into enzymatic glycosyl transfer with retention of configuration through analysis of glycomimetic inhibitors. *Angew. Chem. Int. Ed Engl.* **49**, 1234-1237 (2010).

130. Weijers, C.A., Franssen, M.C. & Visser, G.M. Glycosyltransferase-catalyzed synthesis of bioactive oligosaccharides. *Biotechnol. Adv.* **26**, 436-456 (2008).

131. Palcic, M.M. Glycosyltransferases as biocatalysts. *Curr. Opin. Chem. Biol.* **15**, 226-233 (2011).

132. Wagner, G.K. & Pesnot, T. Glycosyltransferases and their assays. *Chembiochem* **11**, 1939-1949 (2010).

133. Williams, G.J., Zhang, C. & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

134. Williams, G.J. & Thorson, J.S. A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Protoc.* **3**, 357-362 (2008).

135. Aharoni, A. *et al.* High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* **3**, 609-614 (2006).

136. Persson, M. & Palcic, M.M. A high-throughput pH indicator assay for screening glycosyltransferase saturation mutagenesis libraries. *Anal. Biochem.* **378**, 1-7 (2008).

137. Lee, H.S. & Thorson, J.S. Development of a universal glycosyltransferase assay amenable to high-throughput formats. *Anal. Biochem.* **418**, 85-88 (2011).

138. Placic, M.M. & Sujino, K. Assays for glycosyltransferases. *Trends Glycosci Glyc* **13**, 361-370 (2001).

139. Gantt, R.W., Peltier, P., Cournoyer, W.J. & Thorson, J.S. Using simple donors to drive the equilibria of glycosyltransferase-catalyzed reactions. *Nat. Chem. Biol.* **7**, 685-691 (2011).

140. Seibel, J., Jördening, H. & Buchholz, K. Glycosylation with activated sugars using glycosyltransferases and transglycosidases. *Biocatal. Biotransfor.* **24**, 311-342 (2006).

141. Rupprath, C., Schumacher, T. & Elling, L. Nucleotide deoxysugars: essential tools for the glycosylation engineering of novel bioactive compounds. *Curr. Med. Chem.* **12**, 1637-1675 (2005).

142. Wagner, G.K., Pesnot, T. & Field, R.A. A survey of chemical methods for sugar-nucleotide synthesis. *Nat. Prod. Rep.* **26**, 1172-1194 (2009).

143. Hopwood, D.A. *et al.* Production of 'hybrid' antibiotics by genetic engineering. *Nature* **314**, 642-644 (1985).

144. Rodriguez, L. *et al.* Engineering deoxysugar biosynthetic pathways from antibiotic-producing microorganisms. A tool to produce novel glycosylated bioactive compounds. *Chem. Biol.* 9, 721-729 (2002).

145. Lombó, F. *et al.* Engineering biosynthetic pathways for deoxysugars: branched-chain sugar pathways and derivatives from the antitumor tetracenomycin. *Chem. Biol.* **11**, 1709-1718 (2004).

146. Pérez, M. *et al.* Combinatorial biosynthesis of antitumor deoxysugar pathways in *Streptomyces griseus*: reconstitution of "unnatural natural gene clusters" for the biosynthesis of four 2,6-D-dideoxyhexoses. *Appl. Environ. Microbiol.* **72**, 6644-6652 (2006).

147. Baig, I. *et al.* Mithramycin analogues generated by combinatorial biosynthesis show improved bioactivity. *J. Nat. Prod.* **71**, 199-207 (2008).

148. Salas, J.A. & Méndez, C. Indolocarbazole antitumour compounds by combinatorial biosynthesis. *Curr. Opin. Chem. Biol.* **13**, 152-160 (2009).

149. Ramos, A., Olano, C., Braña, A.F., Méndez, C. & Salas, J.A. Modulation of deoxysugar transfer by the elloramycin glycosyltransferase ElmGT through site-directed mutagenesis. *J. Bacteriol.* **191**, 2871-2875 (2009).

150. Shepherd, M.D., Liu, T., Méndez, C., Salas, J.A. & Rohr, J. Engineered biosynthesis of gilvocarcin analogues with altered deoxyhexopyranose moieties. *Appl. Environ. Microbiol.* **77**, 435-441 (2011).

151. Tang, L. & McDaniel, R. Construction of desosamine containing polyketide libraries using a glycosyltransferase with broad substrate specificity. *Chem. Biol.* **8**, 547-555 (2001).

152. Elling, L. *et al.* An enzyme module system for the synthesis of dTDP-activated deoxysugars from dTMP and sucrose. *Chembiochem* **6**, 1423-1430 (2005).

153. Chen, H. *et al.* Deoxysugars in glycopeptide antibiotics: enzymatic synthesis of TDP-L-epivancosamine in chloroeremomycin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 11942-11947 (2000).

154. Takahashi, H., Liu, Y.N. & Liu, H.W. A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate. *J. Am. Chem. Soc.* **128**, 1432-1433 (2006).

155. Zhang, H. *et al.* Elucidation of the kijanimicin gene cluster: insights into the biosynthesis of spirotetronate antibiotics and nitrosugars. *J. Am. Chem. Soc.* **129**, 14670-14683 (2007).

156. Hong, L., Zhao, Z., Melançon III, C.E., Zhang, H. & Liu, H.W. In vitro characterization of the enzymes involved in TDP-D-forosamine biosynthesis in the spinosyn pathway of *Saccharopolyspora spinosa*. *J. Am. Chem. Soc.* **130**, 4954-4967 (2008).

157. Kharel, M.K., Lian, H. & Rohr, J. Characterization of the TDP-D-ravidosamine biosynthetic pathway: one-pot enzymatic synthesis of TDP-D-ravidosamine from thymidine-5-phosphate and glucose-1-phosphate. *Org. Biomol. Chem.* **9**, 1799-1808 (2011).

158. Shackelford, G.S., Regni, C.A. & Beamer L.J. Evolutionary trace analysis of the alpha-D-phosphohexomutase superfamily. *Protein Sci.* **13**, 2130-2138 (2004).

159. Kitaoka, M. & Hayashi, K. Carbohydrate-porcessing phosphorolytic enzymes. *Trends Glycosci. Glycotechnol.* **14**, 35-50 (2002).

160. Pieslinger, A.M., Hoepflinger, M.C. & Tenhaken, R. Cloning of glucuronokinase from *Arabidopsis thaliana*, the last missing enzyme of the myo-inositol oxygenase pathway to nucleotide sugars. *J. Biol. Chem.* **285**, 2902-2910 (2010).

161. Yang, T., Bar-Peled, L., Gebhart, L., Lee S.G. & Bar-Peled, M. Identification of galacturonic acid-1-phosphate kinase, a new member of the GHMP kinase superfamily in plants, and comparison with galactose-1-phosphate kinase. *J. Biol. Chem.* **284**, 21526-21535 (2009).

162. Thoden, J.B. & Holden, H.M. Molecular structure of galactokinase. *J. Biol. Chem.* **278**, 33305-33311 (2003).

163. Holden, H.M., Thoden, J. B., Timson, D.J. & Reece, R.J. Galactokinase: structure, function and role in type II galactosemia. *Cell Mol. Life Sci.* **61**, 2471-2484 (2004).

164. Thoden, J.B. & Holden, H.M. The molecular architecture of human *N*-acetylgalactosamine kinase. *J. Biol. Chem.* **280**, 32784-32791 (2005).

165. Nishimoto, M. & Kitaoka, M. Identification of *N*-acetylhexosamine 1-kinase in the complete lacto-*N*-biose I/galacto-*N*-biose metabolic pathway in *Bifidobacterium longum*. *Appl. Environ. Microbiol.* **73**, 6444-6449 (2007).

166. Park, S.H., Pastuszak, I., Drake, R. & Elbein, A.D. Purification to apparent homogeneity and properties of pig kidney L-fucose kinase. *J. Biol. Chem.* **273**, 5685-5691 (1998).

167. Yang, J., Fu, X., Liao, J., Liu, L. & Thorson, J.S. Structure-based engineering of *E. coli* galactokinase as a first step toward *in vivo* glycorandomization. *Chem. Biol.* **12**, 657-664 (2005).

168. Hoffmeister, D. & Thorson, J.S. Mechanistic implications of *Escherichia coli* galactokinase structure-based engineering. *Chembiochem.* **5**, 989-992 (2004).

169. W. Wang, *et al.* Chemoenzymatic synthesis of GDP-L-fucose and the Lewis X glycan derivatives. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16096-16101 (2009).

170. Yang, J., Liu, L. & Thorson, J.S. Structure-based enhancement of the first anomeric glucokinase. *Chembiochem.* **5**, 992-996 (2004).

171. Cai, L. *et al.* Substrate specificity of *N*-acetylhexosamine kinase towards *N*-acetylgalactosamine derivatives. *Bioorg. Med. Chem. Lett.* **19**, 5433-5435 (2009).

172. Cai, L. *et al.* A chemoenzymatic route to *N*-acetylglucosamine-1-phosphate analogues: substrate specificity investigations of *N*-acetylhexosamine 1-kinase. *Chem. Commun. (Camb)* (20), 2944-2946 (2009).

173. Yi, W. *et al.* Remodeling bacterial polysaccharides by metabolic pathway engineering. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 4207-4212 (2009).

174. Thoden, J.B., Timson, D.J., Reece, R.J. & Holden, H.M. Molecular structure of human galactokinase: implications for type II galactosemia. *J. Biol. Chem.* **280**, 9662-9670 (2005).

175. Hartley, A. *et al.* Substrate specificity and mechanism from the structure of *Pyrococcus furiosus* galactokinase. *J. Mol. Biol.* **337**, 387-398 (2004).

176. Schell, M.A. & Wilson, D.B. Purification and properties of galactokinase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **252**, 1162-1166 (1977).

177. Timson, D.J. & Reece, R.J. Kinetic analysis of yeast galactokinase: implications for transcriptional activation of the GAL genes. *Biochimie* **84**, 265-272 (2002).

178. Heinrich, M.R. The purification and properties of yeast galactokinase. *J. Biol. Chem.* **239**, 50-53 (1964).

179. Dey, P.M. Galactokinase of *Vicia faba* seeds. *Eur. J. Biochem.* **136**, 155-159 (1983).

180. Foglietti, M.J. & Percheron, F. Purification and mechanism of action of a plant galactokinase. *Biochimie* **58**, 499-504 (1976).

181. Ballard, F.J. Purification and properties of galactokinase from pig liver. *Biochem. J.* **98**, 347-352 (1966).

182. Timson, D.J. & Reece, R.J. Sugar recognition by human galactokinase. *BMC Biochem.* **4**, 16 (2003).

183. Yang, J. *et al.* Studies on the substrate specificity of *Escherichia coli* galactokinase. *Org. Lett.* **5**, 2223-2226 (2003).

184. Hoffmeister, D., Yang, J., Liu, L. & Thorson, J.S. Creation of the first anomeric D/L-sugar kinase by means of directed evolution. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13184-13189 (2003).

185. Langenhan, J.M. & Thorson, J.S. Recent carbohydrate-based chemoselective ligation applications. *Curr. Org. Synth.* **2**, 59-81 (2005).

186. Baskin, J.M. & Bertozzi, C.R. Copper-free click chemistry: bioorthogonal reagents for tagging azides. *Aldrichimica Acta* **43**, 15-23 (2010).

187. Sletten, E.M. & Bertozzi, C.R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed Engl.* **48**, 6974-6998 (2009).

188. Coyne, M.J., Reinap, B., Lee, M.M. & Comstock, L.E. Human symbionts use a host-like pathway for surface fucosylation. *Science* **307**, 1778-1781 (2005).

189. Kotake, T. *et al.* A bifunctional enzyme with L-fucokinase and GDP-L-fucose pyrophosphorylase activities salvages free L-fucose in *Arabidopsis*. *J. Biol. Chem.* **283**, 8125-8135 (2008).

190. Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Wheeler, D.L. GenBank. *Nucleic Acids Res.* **34**, D16-D20 (2006).

191. Blankenfeldt, W., Asuncion, M., Lam, J.S. & Naismith, J.H. The structural basis of the catalytic mechanism and regulation of glucose-1-phosphate thymidylyltransferase (RmlA). *EMBO J.* **19**, 6652-6663 (2000).

192. S. Zuccotti, *et al.* Kinetic and crystallographic analyses support a sequential-ordered bi bi catalytic mechanism for *Escherichia coli* glucose-1-phosphate thymidylyltransferase. *J. Mol. Biol.* **313**, 831-843 (2001).

193. Jiang, J., Biggins, J.B. & Thorson, J.S. A general enzymatic method for the synthesis of natural and "unnatural" UDP- and TDP-nucleotide sugars. *J. Am. Chem. Soc.* **122**, 6803-6804 (2000).

194. Moretti, R. & Thorson, J.S. Enhancing the latent nucleotide triphosphate flexibility of the glucose-1-phosphate thymidylyltransferase RmlA. *J. Biol. Chem.* **282**, 16942-16947 (2007).

195. Bae, J. *et al.* A practical enzymatic synthesis of UDP sugars and NDP glucoses. *Chembiochem* **6**, 1963-1966 (2005).

196. Zhang, Z. Identification of an extremely thermostable enzyme with dual sugar-1-phosphate nucleotidylyltransferase activities from an acidothermophilic archaeon, *Sulfolobus tokodaii* strain 7. *J. Biol. Chem.* **280**, 9698-9705 (2005).

197. Mizanur, R.M. & Pohl, N.L. Phosphomannose isomerase/GDP-mannose pyrophosphorylase from *Pyrococcus furiosus*: a thermostable biocatalyst for the synthesis of guanidinediphosphate-activated and mannose-containing sugar nucleotides. *Org. Biomol. Chem.* **7**, 2135-2139 (2009).

198. Barton, W.A., Biggins, J.B., Jiang, J., Thorson, J.S. & Nikolov, D.B. Expanding pyrimidine diphosphosugar libraries via structure-based nucleotidylyltransferase engineering. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 13397-13402 (2002).

199. Barton, W.A. *et al.* Structure, mechanism and engineering of a nucleotidylyltransferase as a first step toward glycorandomization. *Nat. Struct. Biol.* **8**, 545-551 (2001).

200. Jakeman, D.L. *et al.* Engineering ribonucleoside triphosphate specificity in a thymidylyltransferase. *Biochemistry* **47**, 8719-8725 (2008).

201. Moretti, R. *et al.* Expanding the nucleotide and sugar 1-phosphate promiscuity of nucleotidyltransferase RmlA via directed evolution. *Biol. Chem.* **286**, 13235-13243 (2011).

202. Thorson, J.S., Barton, W.A., Hoffmeister, D., Albermann, C. & Nikolov, D.B. Structure-based enzyme engineering and its impact on *in vitro* glycorandomization. *Chembiochem* **5**, 16-25 (2004).

203. Albermann, C. *et al.* Substrate specificity of NovM: implications for novobiocin biosynthesis and glycorandomization. *Org. Lett.* **5**, 933-936 (2003).

204. Errey, J.C., Mukhopadhyay, B., Kartha, K.P. & Field, R.A. Flexible enzymatic and chemo-enzymatic approaches to a broad range of uridine-diphospho-sugars. *Chem. Commun. (Camb)* (23), 2706-2707 (2004).

205. Watt, G.M., Flitsch, S.L., Fey, S., Elling, L. & Kragl, U. The preparation of deoxy derivatives of mannose-1-phosphate and their substrate specificity towards recombinant GDP-mannose pyrophosphorylase from *Salmonella enterica*, group B. *Tetrahedron: Asymmetry* **11**, 621-628 (2000).

206. Zhang, C., Moretti, R., Jiang, J. & Thorson, J.S. The *in vitro* characterization of polyene glycosyltransferases AmphDI and NysDI. *Chembiochem* **9**, 2506-2514 (2008).

207. Timmons, S.C., Mosher, R.H., Knowles, S.A. & Jakeman, D.L. Exploiting nucleotidylyltransferases to prepare sugar nucleotides. *Org. Lett.* **9**, 857-860 (2007).

208. Kotake, T. *et al.* UDP-sugar pyrophosphorylase with broad substrate specificity toward various monosaccharide 1-phosphates from pea sprouts. *J. Biol. Chem.* **279**, 45728-45736 (2004).

209. Damerow, S. *et al. Leishmania* UDP-sugar pyrophosphorylase: the missing link in galactose salvage? *J. Biol. Chem.* **285**, 878-887 (2010).

210. Y. Yang, *et al.* Characterization of GDP-mannose pyrophosphorylase from *Escherichia coli* O157:H7 EDL933 and its broad substrate specificity. *J. Molec. Catal. B* **37**, 1-8 (2005).

211. Jiang, J., Albermann, C. & Thorson, J.S. Application of the nucleotidylyltransferase Ep toward the chemoenzymatic synthesis of dTDP-desosamine analogues. *Chembiochem* **4**, 443-446 (2003).

212. Williams, G.J., Goff, R.D., Zhang, C. & Thorson, J.S. Optimizing glycosyltransferase specificity via "hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem. Biol.* **15**, 393-401 (2008).

213. Jiang, J., Biggins, J.B. & Thorson, J.S. Expanding the pyrimidine diphosphosugar Repertoirer: the chemoenzymatic synthesis of amino- and acetamidoglucopyranosyl derivatives. *Angew. Chem. Int. Ed Engl.* **40**, 1502-1505 (2001).

214. Guan, W., Cai, L., Fang, J., Wu, B. & George Wang, P. Enzymatic synthesis of UDP-GlcNAc/UDP-GalNAc analogs using *N*-acetylglucosamine 1-phosphate uridyltransferase (GlmU). *Chem. Commun. (Camb)* (45), 6976-6978 (2009).

215. Marchesan, S. & Macmillan, D. Chemoenzymatic synthesis of GDP-azidodeoxymannoses: non-radioactive probes for mannosyltransferase activity. *Chem. Commun. (Camb)* (36), 4321-4323 (2008).

216. Beaton, S.A., Huestis, M.P., Sadeghi-Khomami, A., Thomas, N.R. & Jakeman, D.L. Enzyme-catalyzed synthesis of isosteric phosphono-analogues of sugar nucleotides. *Chem. Commun. (Camb)* **2**, 238-240 (2009).

217. Timmons, S.C. *et al.* Enzyme-catalyzed synthesis of furanosyl nucleotides. *Org. Lett.* **10**, 161-163 (2008).

218. Peltier, P., Guégan, J., Daniellou, R., Nugier-Chauvin, C. & Ferrières, V. Stereoselective chemoenzymatic synthesis of UDP-1,2-cis-furanoses from α,β-furanosyl 1-phosphates. *Eur. J. Org. Chem.* **2008**, 5988-5994 (2008).

219. Huestis, M.P., Aish, G.A., Hui, J.P., Soo, E.C. & Jakeman, D.L. Lipophilic sugar nucleotide synthesis by structure-based design of nucleotidylyltransferase substrates. *Org. Biomol. Chem.* **6**, 477-484 (2008).

220. Moretti, R. & Thorson, J.S. A comparison of sugar indicators enables a universal high-throughput sugar-1-phosphate nucleotidyltransferase assay. *Anal. Biochem.* **377**, 251-258 (2008).

221. Ko, K.S., Zea, C.J. & Pohl, N.L. Surprising bacterial nucleotidyltransferase selectivity in the conversion of carbaglucose-1-phosphate. *J. Am. Chem. Soc.* **126**, 13188-13189 (2004).

222. Ko, K.S., Zea, C.J. & Pohl, N.L. Strategies for the chemoenzymatic synthesis of deoxysugar nucleotides: substrate binding versus catalysis. *J. Org. Chem.* **70**, 1919-1921 (2005).

223. Kotake, T. *et al.* Properties and physiological functions of UDP-sugar pyrophosphorylase in *Arabidopsis*. *Biosci. Biotechnol. Biochem.* **71**, 761-771 (2007).

224. Dickmanns, A. *et al.* Structural basis for the broad substrate range of the UDP-sugar pyrophosphorylase from *Leishmania major*. *J. Mol. Biol.*, 2011, 405, 461-478.

225. Yang, T. & Bar-Peled, M. Identification of a novel UDP-sugar pyrophosphorylase with a broad substrate specificity in *Trypanosoma cruzi*. *Biochem. J.* **429**, 533-543 (2010).

226. Mizanur, R.M., Jaipuri, F.A. & Pohl, N.L. One-step synthesis of labeled sugar nucleotides for protein *O*-GlcNAc modification studies by chemical function analysis of an archaeal protein. *J. Am. Chem. Soc.* **127**, 836-837 (2005).

227. Hanson, S., Best, M., Bryan, M.C. & Wong, C. Chemoenzymatic synthesis of oligosaccharides and glycoproteins. *Trends Biochem. Sci.* **29**, 656-663 (2004).

228. Koeller, K.M. & Wong, C. Synthesis of complex carbohydrates and glycoconjugates: enzyme-based and programmable one-pot strategies. *Chem. Rev.* **100**, 4465-4494 (2000).

229. Ichikawa, Y. *et al.* Chemical-enzymic synthesis and conformational analysis of sialyl Lewis X and derivatives. *J. Am. Chem. Soc.* **114**, 9283-9298 (1992).

230. De Luca, C. *et al.* Enzymic synthesis of hyaluronic acid with regeneration of sugar nucleotides *J. Am. Chem. Soc.* **117**, 5869-5870 (1995).

231. Liu, Z., Zhang, J., Chen, X. & Wang, P.G. Combined biosynthetic pathway for *de novo* production of UDP-galactose: catalysis with multiple enzymes immobilized on agarose beads. *Chembiochem* **3**, 348-355 (2002).

232. Chen, X. *et al.* Sugar nucleotide regeneration beads (superbeads): a versatile tool for the practical synthesis of oligosaccharides. *J. Am. Chem. Soc.* **123**, 2081-2082 (2001).

233. Cardini, C.E., Leloir, L.F. & Chiriboga, J. The biosynthesis of sucrose. *J. Biol. Chem.* **214**, 149-155 (1955).

234. Glaser, L. & Brown, D.H. The synthesis of chitin in cell-free extracts of *Neurospora crassa*. *J. Biol. Chem.* **228**, 729-742 (1957).

235. Sutter, A. & Grisebach, H. Free reversibility of the UDP-glucose: flavonol 3-*O*-glucosyltransferase reaction. *Arch. Biochem. Biophys.* **167**, 444-447 (1975).

236. Wang, S.X. & Ellis, B.E. Enzymology of UDP-glucose:sinapic acid glucosyltransferase from *Brassica napus*. *Phytochemistry* **49**, 307-318 (1998).

237. Minami, A., Kakinuma, K. & Eguchi, T. Aglycon switch approach toward unnatural glycosides from natural glycoside with glycosyltransferase VinC. *Tetrahedron Lett.* **46**, 6187-6190 (2005).

238. Zhang, C., *et al.* Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science* **313**, 1291-1294 (2006).

239. Zhang, C., Albermann, C., Fu, X. & Thorson, J.S. The *in vitro* characterization of the iterative avermectin glycosyltransferase AveBI reveals reaction reversibility and sugar nucleotide flexibility. *J. Am. Chem. Soc.* **128**, 16420-16421 (2006).

240. Lairson, L.L., Wakarchuk, W.W. & Withers, S.G. Alternative donor substrates for inverting and retaining glycosyltransferases. *Chem. Commun. (Camb)* (4), 365-367 (2007).

241. Zhang, C., Fu, Q., Albermann, C., Li, L. & Thorson, J. S. The in vitro characterization of the erythronolide mycarosyltransferase EryBV and its utility in macrolide diversification. *Chembiochem* **8**, 385-390 (2007).

242. Zhang, C. *et al.* Biochemical and structural insights of the early glycosylation steps in calicheamicin biosynthesis. *Chem. Biol.* **15**, 842-853 (2008).

243. Chandrasekaran, E.V., Xue, J., Xia, J., Locke, R.D., Matta, K.L. & Neelamegham, S. Reversible sialylation: synthesis of cytidine 5'-monophospho-*N*-acetylneuraminic acid from cytidine 5'-monophosphate with alpha2,3-sialyl *O*-glycan-, glycolipid-, and macromolecule-based donors yields diverse sialylated products. *Biochemistry* **47**, 320-330 (2008).

244. Modolo, L.V., Escamilla-Treviño, L.L., Dixon, R.A. & Wang, X. Single amino acid mutations of *Medicago* glycosyltransferase UGT85H2 enhance activity and impart reversibility. *FEBS Lett.* **583**, 2131-2135 (2009).

245. Modolo, L.V., *et al.* Crystal structures of glycosyltransferase UGT78G1 reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids. *J. Mol. Biol.* **392**, 1292-1302 (2009).

246. Okada, T., Ihara, H, Ito, R., Taniguchi, N. & Ikeda, Y. Bidirectional *N*-acetylglucosamine transfer mediated by beta-1,4-*N*-acetylglucosaminyltransferase III. *Glycobiology* **19**, 368-374 (2009).

247. Gantt, R.W., Goff, R.D., Williams, G.J. & Thorson, J.S. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem. Int. Ed Engl.* **47**, 8889-8892 (2008).

248. Williams, G.J., Yang, J., Zhang, C. & Thorson, J.S. Recombinant *E. coli* prototype strains for *in vivo* glycorandomization. *ACS Chem. Biol.* **6**, 95-100 (2011).

249. Baltz, R.H. & Seno, E.T. Genetics of *Streptomyces fradiae* and tylosin biosynthesis. *Annu. Rev. Microbiol.* **42**, 547-574 (1988).

250. Decker, H., Haag, S., Udvarnoki, G. & Rohr, J. Novel genetically engineered tetracenomycins. *Angew. Chem. Int. Ed Engl.* **34**, 1107-1110 (1995).

251. Solenberg, P.J. *et al.* Production of hybrid glycopeptide antibiotics *in vitro* and in *Streptomyces toyocaensis*. *Chem. Biol.* **4**, 195-202 (1997).

252. K. Madduri, *et al.* Production of the antitumor drug epirubicin (4'-epidoxorubicin) and its precursor by a genetically engineered strain of *Streptomyces peucetius*. *Nat. Biotechnol.* **16**, 69-74 (1998).

253. Borisova, S.A., Zhao, L., Sherman, D.H. & Liu, H.W. Biosynthesis of desosamine: construction of a new macrolide carrying a genetically designed sugar moiety. *Org. Lett.* **1**, 133-136 (1999).

254. Hong, J.S., Park, S.H., Choi, C.Y., Sohng, J.K. & Yoon, Y.J. New olivosyl derivatives of methymycin/pikromycin from an engineered strain of *Streptomyces venezuelae*. *FEMS Microbiol. Lett.* **238**, 391-399 (2004).

255. Pageni, B.B., Oh, T.J., Liou, K., Yoon, Y.J. & Sohng, J.K. Genetically engineered biosynthesis of macrolide derivatives including 4-amino-4,6-dideoxy-L-glucose from *Streptomyces venezuelae* YJ003-OTBP3. *J. Microbiol. Biotechnol.* **18**, 88-94 (2008).

256. Antoine, T. *et al.* Large-scale *in vivo* synthesis of the carbohydrate moieties of gangliosides GM1 and GM2 by metabolically engineered *Escherichia coli*. *Chembiochem* **4**, 406-412 (2003).

257. Drouillard, S., Driguez, H., & Samain, E. Large-scale synthesis of H-antigen oligosaccharides by expressing *Helicobacter pylori* α1,2-fucosyltransferase in metabolically engineered *Escherichia coli* cells. *Angew. Chem. Int. Ed. Engl.* **45**, 1778-1780 (2006).

258. Drouillard, S., Mine, T., Kajiwara, H., Yamamoto, T. & Samain, E. Efficient synthesis of 6'-sialyllactose, 6,6'-disialyllactose, and 6'-KDO-lactose by metabolically engineered *E. coli* expressing a multifunctional sialyltransferase from the *Photobacterium* sp. JT-ISH-224. *Carbohydr. Res.* **345**, 1394-1399 (2010).

259. Jung, W.S. *et al.* Bioconversion of 12-, 14-, and 16-membered ring aglycones to glycosylated macrolides in an engineered strain of *Streptomyces venezuelae*. *Appl. Microbiol. Biotechnol.* **76**, 1373-1381 (2007).

260. Salas, J.A. & Méndez, C. Engineering the glycosylation of natural products in actinomycetes. *Trends Microbiol.* **15**, 219-232 (2007).

261. Chen, X. *et al.* Reassembled biosynthetic pathway for large-scale carbohydrate synthesis: alpha-Gal epitope producing "superbug". *Chembiochem* **3**, 47-53 (2002).

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

263. Willits, M.G. *et al.* Bio-fermentation of modified flavonoids: an example of *in vivo* diversification of secondary metabolites. *Phytochemistry* **65**, 31-41 (2004).

264. Lim, E.K., Ashford, D.A., Hou, B., Jackson, R.G. & Bowles, D.J. *Arabidopsis* glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. *Biotechnol. Bioeng.* **87**, 623-631 (2004).

265. Kim, J.H., Shin, K.H., Ko, J.H. & Ahn, J.H. Glucosylation of flavonols by *Escherichia coli* expressing glucosyltransferase from rice (*Oryza sativa*). *J. Biosci. Bioeng.* **102**, 135-137 (2006).

266. Hui, J.P., Yang, J., Thorson, J.S. & Soo, E.C. Selective detection of sugar phosphates by capillary electrophoresis/mass spectrometry and its application to an engineered *E. coli* host. *Chembiochem* **8**, 1180-1188 (2007).

267. Gormley, N.A., Orphanides, G., Meyer, A., Cullis, P.M. & Maxwell, A. The interaction of coumarin antibiotics with fragments of DNA gyrase B protein. *Biochemistry* **35**, 5083-5092 (1996).

268. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).

269. Marcu, M.G., Chadli, A., Bouhouche, I., Catelli, M. & Neckers, L. M. The heat shock protein 90 antagonist novobiocin interacts with a previously unrecognized ATP-binding domain in the carboxyl terminus of the chaperone. *J. Biol. Chem.* **275**, 37181-37186 (2000).

270. Heide, L. Genetic engineering of antibiotic biosynthesis for the generation of new aminocoumarins. *Biotechnol. Adv.* **27**, 1006-1014 (2009).

271. Freel Meyers, C.L., Oberthur, M., Anderson, J.W., Kahne, D. & Walsh, C.T. Initial characterization of novobiocic acid noviosyl transferase activity of NovM in biosynthesis of the antibiotic novobiocin. *Biochemistry* **42**, 4179-4189 (2003).

272. Pacholec, M., Freel Meyers, C.L., Oberthur, M., Kahne, D. and Walsh, C.T. Characterization of the aminocoumarin ligase SimL from the simocyclinone pathway and tandem incubation with NovM,P,N from the novobiocin pathway. *Biochemistry* **44**, 4949-4956 (2005).

273. Li, S.M. & Heide, L. The biosynthetic gene clusters of aminocoumarin antibiotics. *Planta Med.* **72**, 1093-1099 (2006).

274. Temperini, C., Cirilli, M., Aschi, M. & Ughetto, G. Role of the amino sugar in the DNA binding of disaccharide anthracyclines: crystal structure of the complex MAR70/d(CGATCG). *Bioorg. Med. Chem.* **13**, 1673-1679 (2005).

275. K. Raty, *et al.* Cloning and characterization of *Streptomyces galilaeus* aclacinomycins polyketide synthase (PKS) cluster. *Gene* **293**, 115-122 (2002).

276. Raty, K., Kunnari, T., Hakala, J., Mantsala, P. & Ylihonko, K. A gene cluster from *Streptomyces galilaeus* involved in glycosylation of aclarubicin. *Mol. Gen. Genet.* **264**, 164-172 (2000).

277. Lu, W. *et al.* AknT is an activating protein for the glycosyltransferase AknS in L-aminodeoxysugar transfer to the aglycone of aclacinomycin A. *Chem. Biol.* **12**, 527-534 (2005).

278. Leimkuhler, C. *et al.* 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 (2007).

279. Borisova, S.A., Zhao, L., Melancon III, C.E., Kao, C.L. & Liu, H.W. Characterization of the glycosyltransferase activity of desVII: analysis of and implications for the biosynthesis of macrolide antibiotics. *J. Am. Chem. Soc.* **126**, 6534-6535 (2004).

280. Lu, W., Leimkuhler, C., Oberthur, M., Kahne, D. & Walsh, C.T. AknK is an L-2-deoxyfucosyltransferase in the biosynthesis of the anthracycline aclacinomycin A. *Biochemistry* **43**, 4548-4558 (2004).

281. Keller-Schierlein, W., Sauerbier, J., Vogler, U. & Zahner, H. Metabolites of microorganisms. Aranciamycin. *Helv. Chim. Acta* **53**, 779-789 (1970).

282. Bols, M., Binderup, L., Hansen J. & Rasmussen, P. Inhibition of collagenase by aranciamycin and aranciamycin derivatives. *J. Med. Chem.* **35**, 2768-2771 (1992).

283. Sianidis, G. *et al.* Cloning, purification and characterization of a functional anthracycline glycosyltransferase. *J. Biotechnol.* **125**, 425-433 (2006).

284. Wohlert, S. *et al.* Insights about the biosynthesis of the avermectin deoxysugar L-oleandrose through heterologous expression of *Streptomyces avermitilis* deoxysugar genes in *Streptomyces lividans*. *Chem. Biol.* **8**, 681-700 (2001).

285. Rodriguez, L. *et al.* Generation of hybrid elloramycin analogs by combinatorial biosynthesis using genes from anthracycline-type and macrolide biosynthetic pathways. *J. Mol. Microbiol. Biotechnol.* **2**, 271-276 (2000).

286. Blanco, G. *et al.* Identification of a sugar flexible glycosyltransferase from *Streptomyces olivaceus*, the producer of the antitumor polyketide elloramycin. *Chem. Biol.* **8**, 253-263 (2001).

287. Fischer, C. *et al.* Digitoxosyltetracenomycin C and glucosyltetracenomycin C, two novel elloramycin analogues obtained by exploring the sugar donor substrate specificity of glycosyltransferase ElmGT. *J. Nat. Prod.* **65**, 1685-1689 (2002).

288. Pérez, M. *et al.* Combining sugar biosynthesis genes for the generation of L- and D-amicetose and formation of two novel antitumor tetracenomycins. *Chem. Commun. (Camb)* (12), 1604-1606 (2005).

289. Olano, C. *et al.* Glycosylated derivatives of steffimycin: insights into the role of the sugar moieties for the biological activity. *Chembiochem* **9**, 624-633 (2008).

290. Liu, T., Kharel, M.K., Fischer, C., McCormick, A. & Rohr, J. Inactivation of gilGT, encoding a *C*-glycosyltransferase, and gilOIII, encoding a P450 enzyme, allows the details of the late biosynthetic pathway to gilvocarcin V to be delineated. *Chembiochem* **7**, 1070-1077 (2006).

291. Liu, T., *et al.* Inactivation of the ketoreductase gilU gene of the gilvocarcin biosynthetic gene cluster yields new analogues with partly improved biological activity. *Chembiochem* **10**, 278-286 (2009).

292. Durr, C. *et al.* The glycosyltransferase UrdGT2 catalyzes both *C*- and *O*-glycosidic sugar transfers. *Angew. Chem. Int. Ed Engl.* **43**, 2962-2965 (2004).

293. Härle, J. *et al.* Rational design of an aryl-*C*-glycoside catalyst from a natural product *O*-glycosyltransferase. *Chem. Biol.* **18**, 520-530 (2011).

294. Thorson, J.S. *et al.* Understanding and exploiting nature's chemical arsenal: the past, present and future of calicheamicin research. *Curr. Pharm. Des.* **6**, 1841-1879 (2000).

295. Galm, U. *et al.* Antitumor antibiotics: bleomycin, enediynes, and mitomycin. *Chem. Rev.* **105**, 739-758 (2005).

296. Ahlert, J. *et al.* The calicheamicin gene cluster and its iterative type I enediyne PKS. *Science* **297**, 1173-1176 (2002).

297. Kahne, D., Leimkuhler, C., Lu, W. & Walsh, C. Glycopeptide and lipoglycopeptide antibiotics. *Chem. Rev.* **105**, 425-448 (2005).

298. Weigel, L.M. *et al.* Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science* **302**, 1569-1571 (2003).

299. S. Chang, *et al.* Infection with vancomycin-resistant *Staphylococcus aureus* containing the vanA resistance gene. *N. Engl. J. Med.* **348**, 1342-1347 (2003).

300. Malabarba, A. & Ciabatti, R. Glycopeptide derivatives. *Curr. Med. Chem.* **8**, 1759-1773 (2001).

301. Ge, M., Thompson, C. & Kahne, D. Reconstruction of Vancomycin by Chemical Glycosylation of the Pseudoaglycon. *J. Am. Chem. Soc.* **120**, 11014-11015 (1998).

302. Losey, H.C. *et al.* Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides. *Chem. Biol.* **9**, 1305-1314 (2002).

303. Thayer, D.A. & Wong, C.H. Vancomycin analogues containing monosaccharides exhibit improved antibiotic activity: a combined one-pot enzymatic glycosylation and chemical diversification strategy. *Chem. Asian J.* **1**, 445-452 (2006).

304. Oh, T.J., Kim, D.H., Kang, S.Y., Yamaguchi, T. & Sohng, J.K. Enzymatic synthesis of vancomycin derivatives using galactosyltransferase and sialyltransferase. *J. Antibiot. (Tokyo)* **64**, 103-109 (2011).

305. Baltz, R.H. Biosynthesis and genetic engineering of lipopeptide antibiotics related to daptomycin. *Curr. Top. Med. Chem.* **8**, 618-638 (2008).

306. Losey, H.C. *et al.* Tandem action of glycosyltransferases in the maturation of vancomycin and teicoplanin aglycones: novel glycopeptides. *Biochemistry* **40**, 4745-4755 (2001).

307. Fu, X. *et al.* Antibiotic optimization via *in vitro* glycorandomization. *Nat. Biotechnol.* **21**, 1467-1469 (2003).

308. Fu, X., Albermann, C., Zhang, C. & Thorson, J.S. Diversifying vancomycin via chemoenzymatic strategies. *Org. Lett.* **7**, 1513-1515 (2005).

309. Oberthur, M. *et al.* A systematic investigation of the synthetic utility of glycopeptide glycosyltransferases. *J. Am. Chem. Soc.* **2005**, 10747-10752 (2005).

310. Li, T.L. *et al.* Biosynthetic gene cluster of the glycopeptide antibiotic teicoplanin: characterization of two glycosyltransferases and the key acyltransferase. *Chem. Biol.* **11**, 107-119 (2004).

311. Sosio, M. *et al.* Organization of the teicoplanin gene cluster in *Actinoplanes teichomyceticus*. *Microbiology* **150**, 95-102 (2004).

312. Howard-Jones, A.R. *et al.* Kinetic analysis of teicoplanin glycosyltransferases and acyltransferase reveal ordered tailoring of aglycone scaffold to reconstitute mature teicoplanin. *J. Am. Chem. Soc.* **129**, 10082-10083 (2007).

313. Truman, A.W. *et al.* Chimeric glycosyltransferases for the generation of hybrid glycopeptides. *Chem. Biol.* **16**, 676-685 (2009).

314. Liu, Y.C. *et al.* Interception of teicoplanin oxidation intermediates yields new antimicrobial scaffolds. *Nat. Chem. Biol.* **7**, 304-309 (2011).

315. Westerlind, U. & Norberg, T. Chemical synthesis of analogs of the glycopeptide contulakin-G, an analgetically active conopeptide from *Conus geographus*. *Carbohydr. Res.* **341**, 9-18 (2006).

316. Oman, T.J. & van der Donk, W.A. Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **6**, 9-18 (2010).

317. Craig, A.G. *et al.* Contulakin-G, an O-glycosylated invertebrate neurotensin. *J. Biol. Chem.* **274**, 13752-13759 (1999).

318. Oman, T.J., Boettcher, J.M., Wang, H., Okalibe, X.N. & van der Donk, W.A. Sublancin is not a lantibiotic but an *S*-linked glycopeptide. *Nat. Chem. Biol.* **7**, 78-80 (2011).

319. Stepper, J. *et al.* Cysteine *S*-glycosylation, a new post-translational modification found in glycopeptide bacteriocins. *FEBS Lett.* **585**, 645-650 (2011).

320. Sanchez, C., Méndez, C. & Salas, J.A. Indolocarbazole natural products: occurrence, biosynthesis, and biological activity. *Nat. Prod. Rep.* **23**, 1007-1045 (2006).

321. Salas, A.P. *et al.* Deciphering the late steps in the biosynthesis of the anti-tumour indolocarbazole staurosporine: sugar donor substrate flexibility of the StaG glycosyltransferase. *Mol. Microbiol.* **58**, 17-27 (2005).

322. Gao, Q., Zhang, C., Blanchard, S. & Thorson, J.S. Deciphering indolocarbazole and enediyne aminodideoxypentose biosynthesis through comparative genomics: insights from the AT2433 biosynthetic locus. *Chem. Biol.* **13**, 733-743 (2006).

323. Zhang, C. *et al.* RebG- and RebM-catalyzed indolocarbazole diversification. *Chembiochem* **7**, 795-804 (2006).

324. Ogasawara, Y. *et al.* Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in *Streptomyces halstedii*. *Chem. Biol.* **11**, 79-86 (2004).

325. Minami, A., Uchida, R., Eguchi, T. & Kakinuma, K. Enzymatic approach to unnatural glycosides with diverse aglycon scaffolds using glycosyltransferase VinC. *J. Am. Chem. Soc.* **127**, 6148-6149 (2005).

326. Minami, A. & Eguchi, T. Substrate flexibility of vicenisaminyltransferase VinC involved in the biosynthesis of vicenistatin. *J. Am. Chem. Soc.* **129**, 5102-5107 (2007).

327. Harms, J.M., Bartels, H., Schlunzen, F. & Yonath, A. Antibiotics acting on the translational machinery. *J. Cell. Sci.* **116**, 1391-1393 (2003).

328. McDaniel, R., Welch, M. & Hutchinson, C.R. Genetic approaches to polyketide antibiotics. 1. *Chem. Rev.* **105**, 543-558 (2005).

329. Katz, L. & Ashley, G.W. Translation and protein synthesis: macrolides. *Chem. Rev.* **105**, 499-528 (2005).

330. Wilson, D.N., Harms, J.M., Nierhaus, K.H., Schlunzen, F. & Fucini, P. Species-specific antibiotic-ribosome interactions: implications for drug development. *Biol. Chem.* **386**, 1239-1252 (2005).

331. Tenson, T. & Mankin, A. Antibiotics and the ribosome. *Mol. Microbiol.* **59**, 1664-1677 (2006).

332. Amsden, G.W. Anti-inflammatory effects of macrolides--an underappreciated benefit in the treatment of community-acquired respiratory tract infections and chronic inflammatory pulmonary conditions? *J. Antimicrob. Chemother.* **55**, 10-21 (2005).

333. Vilches, C., Hernandez, C., Méndez, C. & Salas, J.A. Role of glycosylation and deglycosylation in biosynthesis of and resistance to oleandomycin in the producer organism, *Streptomyces antibioticus*. *J. Bacteriol.* **174**, 161-165 (1992).

334. Quirós, L.M., Aguirrezabalaga, I., Olano, C., Méndez, C. & Salas, J.A. Two glycosyltransferases and a glycosidase are involved in oleandomycin modification during its biosynthesis by Streptomyces antibioticus. *Mol. Microbiol.* **28**, 1177-1185 (1998).

335. Zhao, L., Beyer, N.J., Borisova, S.A. & Liu, H.W. Beta-glucosylation as a part of self-resistance mechanism in methymycin/pikromycin producing strain Streptomyces venezuelae. *Biochemistry* **42**, 14794-14804 (2003)..

336. Bonay, P., Munro, S., Fresno, M. & Alarcón, B. Intra-Golgi transport inhibition by megalomicin. *J. Biol. Chem.* **271**, 3719-3726 (1996).

337. Bonay, P., Duran-Chica, I., Fresno, M., Alarcón, B. & Alcina, A. Antiparasitic effects of the intra-Golgi transport inhibitor megalomicin. *Antimicrob. Agents Chemother.* **42**, 2668-2673 (1998).

338. Salanueva, I.J. *et al.* Polymorphism and structural maturation of bunyamwera virus in Golgi and post-Golgi compartments. *J. Virol.* **77**, 1368-1381 (2003).

339. Borisova, S.A. *et al.* Substrate specificity of the macrolide-glycosylating enzyme pair DesVII/DesVIII: opportunities, limitations, and mechanistic hypotheses. *Angew. Chem. Int. Ed Engl.* **45**, 2748-2753 (2006).

340. Kao, C.L., Borisova, S.A., Kim, H.J. & Liu, H.W. Linear aglycones are the substrates for glycosyltransferase DesVII in methymycin biosynthesis: analysis and implications. *J. Am. Chem. Soc.* **128**, 5606-5607 (2006).

341. Borisova, S.A., Kim, H.J., Pu, X., & Liu, H.W. Glycosylation of acyclic and cyclic aglycone substrates by macrolide glycosyltransferase DesVII/DesVIII: analysis and implications. *Chembiochem* **9**, 1554-1558 (2008).

342. Zhao, L., Sherman, D.H. & Liu, H.W. Biosynthesis of desosamine: construction of a new methymycin/neomethymycin analogue by deletion of a desosamine biosynthetic gene. *J. Am. Chem. Soc.* **120**, 10256-10257 (1998).

343. Melançon III, C.E., Yu, W.L. & Liu, H.W. TDP-mycaminose biosynthetic pathway revised and conversion of desosamine pathway to mycaminose pathway with one gene. *J. Am. Chem. Soc.* **127**, 12240-12241 (2005).

344. Melançon III, C.E. & Liu, H.W. Engineered biosynthesis of macrolide derivatives bearing the non-natural deoxysugars 4-epi-D-mycaminose and 3-*N*-monomethylamino-3-deoxy-D-fucose. *J. Am. Chem. Soc.* **129**, 4896-4897 (2007).

345. Y. Yuan, *et al. In vitro* reconstitution of EryCIII activity for the preparation of unnatural macrolides. *J. Am. Chem. Soc.* **127**, 14128-14129 (2005).

346. U. Schell, *et al.* Engineered biosynthesis of hybrid macrolide polyketides containing D-angolosamine and D-mycaminose moieties. *Org. Biomol. Chem.* **6**, 3315-3327 (2008).

347. Quirós, L.M., Carbajo, R.J., Braña, A.F. & Salas, J.A. Glycosylation of macrolide antibiotics. Purification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*. *J. Biol. Chem.* **275**, 11713-11720 (2000).

348. Yang, M., Brazier, M., Edwards, R. & Davis, B.G. High-throughput mass-spectrometry monitoring for multisubstrate enzymes: determining the kinetic parameters and catalytic activities of glycosyltransferases. *Chembiochem* **6**, 346-357 (2005).

349. Omura, S. & Crump, A. The life and times of ivermectin - a success story. *Nat. Rev. Microbiol.* **2**, 984-989 (2004).

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

351. Schulman, M.D., Acton, S.L., Valentino, D.L. & Arison, B.H. Purification and identification of dTDP-oleandrose, the precursor of the oleandrose units of the avermectins. *J. Biol. Chem.* **265**, 16965-16970 (1990).

352. Irschik, H., Jansen, R., Gerth, K., Hofle, G. & Reichenbach, H. The sorangicins, novel and powerful inhibitors of eubacterial RNA polymerase isolated from myxobacteria. *J. Antibiot. (Tokyo)* **40**, 7-13 (1987).

353. Campbell, E.A. *et al.* Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J.* **24**, 674-682 (2005).

354. Kopp, M. *et al.* SorF: a glycosyltransferase with promiscuous donor substrate specificity *in vitro*. *Chembiochem* **8**, 813-819 (2007).

355. Gil, J.A. & Martin, J. F. in *Biotechnology of Antibiotics*, Strohl, W. Ed., pp. 551-576 (Marcel Dekker, New York, 1997).

356. Aparicio, J.F., Mendes, M.V., Antón, N., Recio, E. & Martin, J.F. Polyene macrolide antibiotic biosynthesis. *Curr. Med. Chem.* **11**, 1645-1656 (2004).

357. Zotchev, S.B. Polyene macrolide antibiotics and their applications in human therapy. *Curr. Med. Chem.* **10**, 211-223 (2003).

358. Baginski, M., Czub, J. & Sternal, K. Interaction of amphotericin B and its selected derivatives with membranes: molecular modeling studies. *Chem. Rec.* **6**, 320-332 (2006).

359. Cereghetti, D.M. & Carreira, E.M. Amphotericin B: 50 years of chemistry and biochemistry. *Synthesis*, 914-942 (2006).

360. Szlinder-Richert, J., Mazerski, J., Cybulska, B., Grzybowska, J. & Borowski, E. MFAME, *N*-methyl-*N*-D-fructosyl amphotericin B methyl ester, a new amphotericin B derivative of low toxicity: relationship between self-association and effects on red blood cells. *Biochim. Biophys. Acta* **1528**, 15-24 (2001).

361. Carmody, M. *et al.* Biosynthesis of amphotericin derivatives lacking exocyclic carboxyl groups. *J. Biol. Chem.* **280**, 34420-34426 (2005).

362. Paquet, V. & Carreira, E.M. Significant improvement of antifungal activity of polyene macrolides by bisalkylation of the mycosamine. *Org. Lett.* **8**, 1807-1809 (2006).

363. Palacios, D.S., Dailey, I., Siebert, D.M., Wilcock, B.C. & Burke, M.D. Synthesis-enabled functional group deletions reveal key underpinnings of amphotericin B ion channel and antifungal activities. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 6733-6738 (2011).

364. Caffrey, P., Lynch, S., Flood, E., Finnan S. & Oliynyk, M. Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem. Biol.* **8**, 713-723 (2001).

365. T. Brautaset, *et al.* Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. *Chem. Biol.* **7**, 395-403 (2000).

363. Aparicio, J.F., Colina, A.J., Ceballos, E. & Martin, J.F. The biosynthetic gene cluster for the 26-membered ring polyene macrolide pimaricin. A new polyketide synthase organization encoded by two subclusters separated by functionalization genes. *J. Biol. Chem.* **274**, 10133-10139 (1999).

367. Aparicio, J.F., Fouces, R., Mendes, M.V., Olivera, N. & Martín, J.F. A complex multienzyme system encoded by five polyketide synthase genes is involved in the biosynthesis of the 26-membered polyene macrolide pimaricin in *Streptomyces natalensis*. *Chem. Biol.* **7**, 895-905 (2000).

368. Chen, S. *et al.* Organizational and mutational analysis of a complete FR-008/candicidin gene cluster encoding a structurally related polyene complex. *Chem. Biol.* **10**, 1065-1076 (2003).

369. Nedal, A. *et al.* Analysis of the mycosamine biosynthesis and attachment genes in the nystatin biosynthetic gene cluster of *Streptomyces noursei* ATCC 11455. *Appl. Environ. Microbiol.* **73**, 7400-7407 (2007).

370. Hutchinson, E. *et al.* Redesign of polyene macrolide glycosylation: engineered biosynthesis of 19-(*O*)-perosaminyl-amphoteronolide B. *Chem. Biol.* **17**, 174-182 (2010).

371. Olano, C. *et al.* Deciphering biosynthesis of the RNA polymerase inhibitor streptolydigin and generation of glycosylated derivatives. *Chem. Biol.* **16**, 1031-1044 (2009).

372. Deboer, C., Dietz, A., Savage, G.M. & Silver, W.S. Streptolydigin, a new antimicrobial antibiotic. I. Biologic studies of streptolydigin. *Antibiot. Annu.*, **3**, 886-892 (1955).

373. Williams, G.J., Gantt, R.W. & Thorson, J.S. The impact of enzyme engineering upon natural product glycodiversification. *Curr. Opin. Chem. Biol.* **12**, 556-564 (2008).

374. Offen, W. *et al.* Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J.* **25**, 1396-1405 (2006).

375. Yang, M. *et al.* 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 (2005).

376. Deng, C. & Chen, R.R. A pH-sensitive assay for galactosyltransferase. *Anal. Biochem.* **330**, 219-226 (2004).

377. Park, S.H. *et al.* Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. *Biotechnol. Bioeng.* **102**, 988-994 (2009).

378. Bolam, D.N. *et al.* 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 (2007).

Chapter 2:

Probing the Aglycon Promiscuity of an Engineered Glycosyltransferase

Portions of this chapter have been previously published as:

Gantt, R.W., Goff, R.D., Williams, G.J., & Thorson, J.S. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem. Int. Ed.* **47**, 8889-8892 (2008).

2.1 Introduction

As discussed in Chapter 1, sugars appended to pharmaceutically important natural products influence key pharmacological properties and/or molecular mechanism of action[1-3]. However, studies designed to systematically understand and/or exploit the role of carbohydrates in drug discovery are often limited by the availability of practical synthetic and/or biosynthetic tools[4-6]. Among the contemporary options to address this limitation[7-10], chemoenzymatic glycorandomization utilizes a set of flexible enzymes consisting of an anomeric kinase, sugar-1-phosphate nucleotidylytransferase, and natural product glycosyltransferase (GT) [9-12]. While chemoenzymatic glycorandomization has been successfully applied to alter the natural sugar moieties of numerous natural products[9-14], the process remains primarily restricted by enzyme specificity and availability of suitable GTs for the target of interest. Thus, although there is precedent for improving non-glycosylated therapeutics via glycoconjugation, including colchicine[15], mitomycin[16], podophyllotoxin[17], rapamycin[18], isophosphoramide mustards[19], or taxol[20], such targets remain beyond chemoenzymatic strategies.

Studies on OleD, the oleandomycin (**1**) GT from *Streptomyces antibioticus* (**Figure 2.1a**), revealed an enhanced triple mutant (A242V/S132F/P67T, referred to herein as 'ASP') that displayed marked improvement in proficiency and substrate promiscuity[9]. Additionally, wild-type (WT) OleD had also been reported to glucosylate a small set of non-macrolide aglycon acceptors[21]. Thus, to further probe the fundamental differences and/or similarities

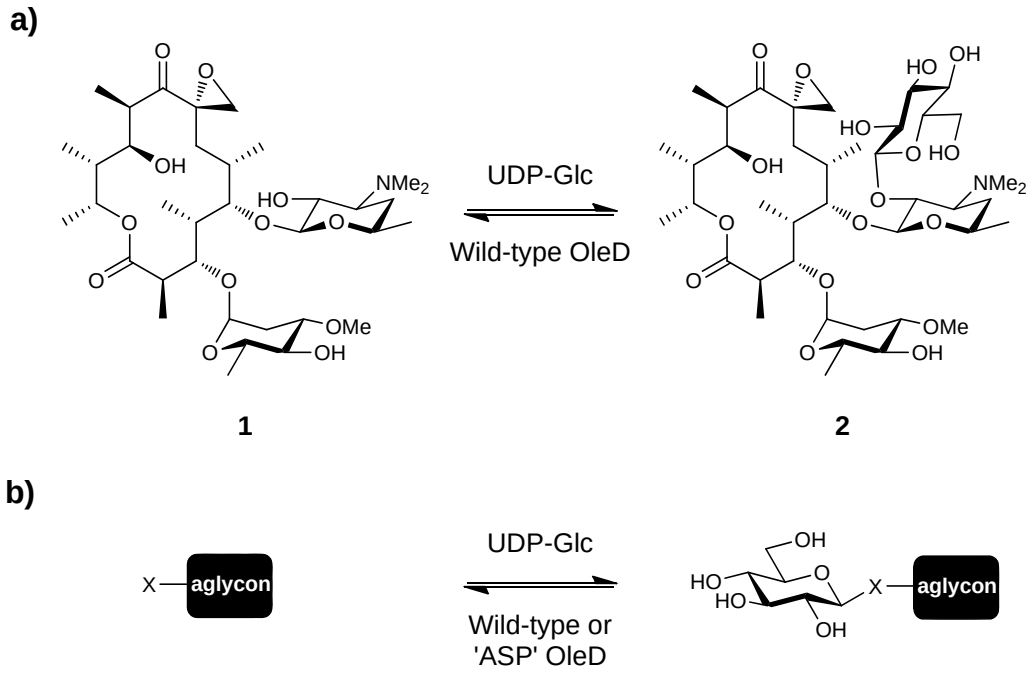


Figure 2.1. OleD reaction schemes. a) Natural reaction catalyzed by WT OleD **b)** General reaction for probing aglycon promiscuity *in vitro* against a panel of 137 library members, X = OH, SH, NH_2, or NHR.

of these two catalysts, a comparison of the aglycon specificities and proficiencies of the WT and 'ASP' OleD variants toward 137 drug-like acceptors (**3**-**139**) was investigated. As a result, the OleD variants investigated were found to glucosylate a total of 71 diverse acceptors, enable putative glycosidic bond formation with 12 distinct types of nucleophiles, and catalyze iterative glycosylation with numerous substrates – thereby establishing OleD as the first multifunctional GT capable of generating *O*-, *S*- and *N*-glycosides and perhaps the most permissive GT characterized to date.

2.2 Results and Discussion

2.2.1. Protein Expression and Purification. The OleD wild-type and 'ASP' variant were expressed at high levels (>20 mg L^{-1} of culture), and following purification, purity of each catalyst was confirmed by SDS-PAGE to be >95% (**Figure 2.2**).

2.2.2. Acceptor Library Screening. 137 drug-like acceptors (**3**-**139**) were screened with purified OleD WT or 'ASP' *in vitro* and evaluated by HPLC (**see Materials and Methods**). Reactions generating new products by HPLC were selected for product confirmation by LC-MS. From this analysis, enzyme-catalyzed glucosylation of 71 of the 137 library members (52%) was observed (**Figure 2.3**). 'ASP' provided higher conversion with 56 of the 71 substrates and, in 10 cases (**14, 18, 47, 53, 66, 67** and **70**-**73**), product was observed only with 'ASP'. In contrast, only polyene **57** was a unique substrate of WT OleD. Notably, of the 71 new substrates, 4 (**6**, **20**, **22**, **42**) and 20 (**13**, **19**, **23**, **26**, **29**, **33**, **34**, **37**, **43**-**45**, **47**, **48**, **50**, **52**, **53**, **55**, **56**, **61**, **63**, and **72**) library members exclusively contained either *S*- or *N*-based nucleophiles, respectively. While the

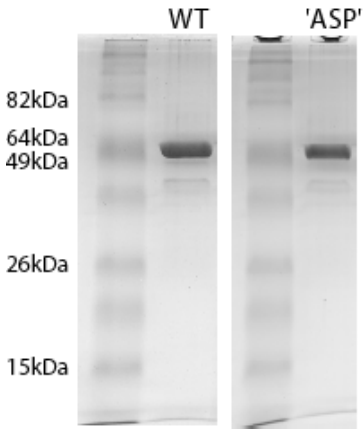


Figure 2.2. SDS-PAGE analysis of purified WT and 'ASP' OleD. Molecular weights for the standard ladder are labeled at left.

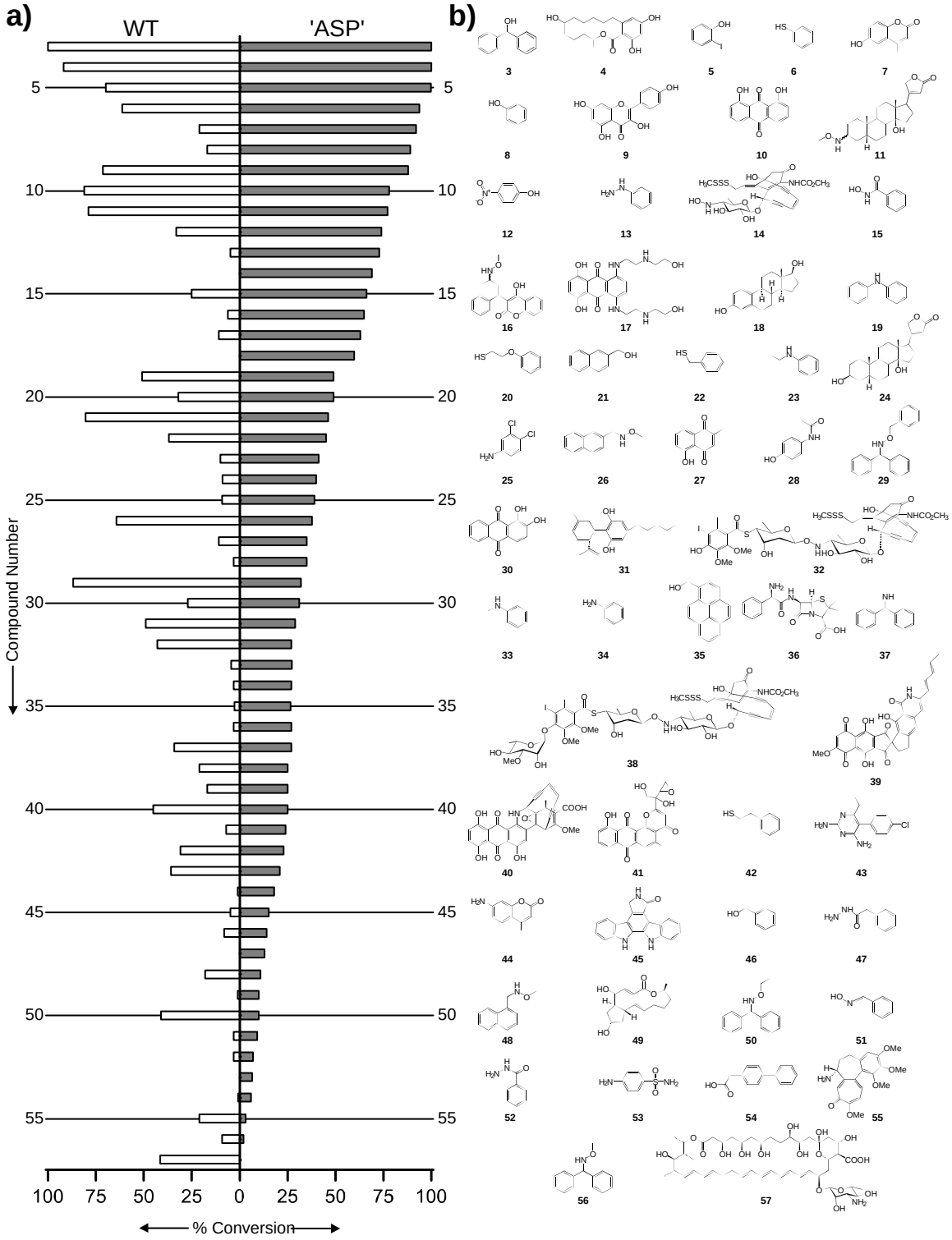


Figure 2.3. Aglycon library screening results. a) Percent conversion of each library member with both WT and 'ASP' OleD. Members are listed in descending order of 'ASP' conversion with numbering corresponding to the structures in (**b**). **b)** Structures of the corresponding library members. Compounds leading to trace products (**58-73**, >5% conversion) or no conversion (**74-139**) are listed in **Appendix 1**.

first-pass LC-MS analysis could not distinguish regio- or stereoselectivity, it is important to note that among the subgroup of library members containing multiple nucleophiles (34 members), 20 led to a single, chromatographically-distinct, monoglucosylated product. Perhaps most surprising was that 13 substrates (**3, 4, 6, 9, 18-21, 23, 26, 29, 33,** and **35**) led to products with masses corresponding to the addition of multiple glucose moieties and within this group, 10 (**3**, **6**, **19-21**, **23**, **26**, **29**, **33**, **35**) contained only a single heteroatom, implicating disaccharide formation via iterative glycosylation.

2.2.3. Confirmation and kinetic evaluation of *O*-, *S*- and *N*-glucoside formation. To confirm *O*-, *S*- and *N*-glucoside formation, iterative glycosylation, and determine anomeric stereoselectivity, a select set of OleD-catalyzed reactions with the simple aromatic model acceptors phenol (**8**), thiophenol (**6**), and aniline (**34**) were studied in depth. NMR characterization of products isolated from large scale reactions was consistent with the β-*O*-, *S*- and *N*-glucosides (**140-143**; **Figure 2.4**, **2.5**; J = 6.7, 7.8, and 9.6 Hz for H-1, respectively). Iterative glycosylation of model acceptor **6** was also determined to be both regio- and stereoselective to provide the 2'-O-(β-D-glucosyl)-glucoside **143** (J = 7.8 and 9.8 Hz for H-1 and H-1', respectively). Interestingly, while the kinetic parameters for all three model substrates fell within the range of previously reported values[9, 11], the ranked order of WT OleD catalytic efficiency (k_{cat}/K_m; thiophenol > phenol ≈ aniline) differed from that for 'ASP' (phenol > aniline > thiophenol) (**Figure 2.6**, **Table 2.1**). Consistent with the enhanced proficiency of 'ASP'[9, 10], the relative k_{cat}/K_m of this mutant was improved 25-, 5-, and 4- fold, respectively, toward phenol (**8**), aniline (**34**), and thiophenol (**6**) (**Table 2.1**).

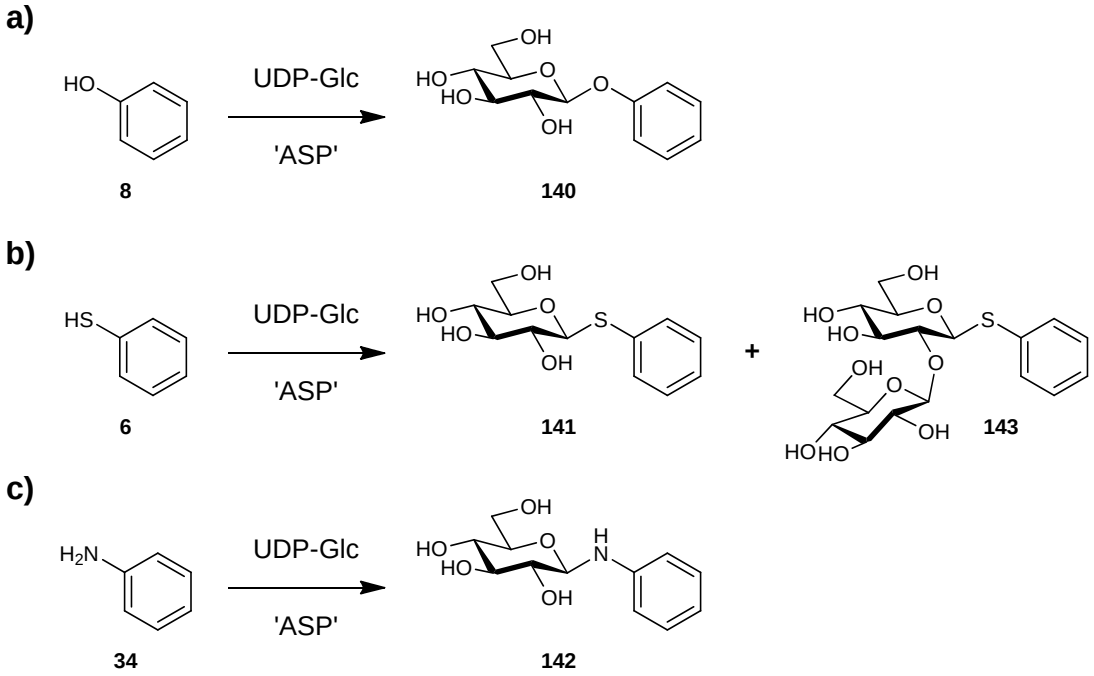


Figure 2.4. Products isolated from large scale enzymatic reactions of OleD variant 'ASP'. Schemes represent products generated from **a)** phenol (**8**), **b)** thiophenol (**6**), and **c)** aniline (**34**).

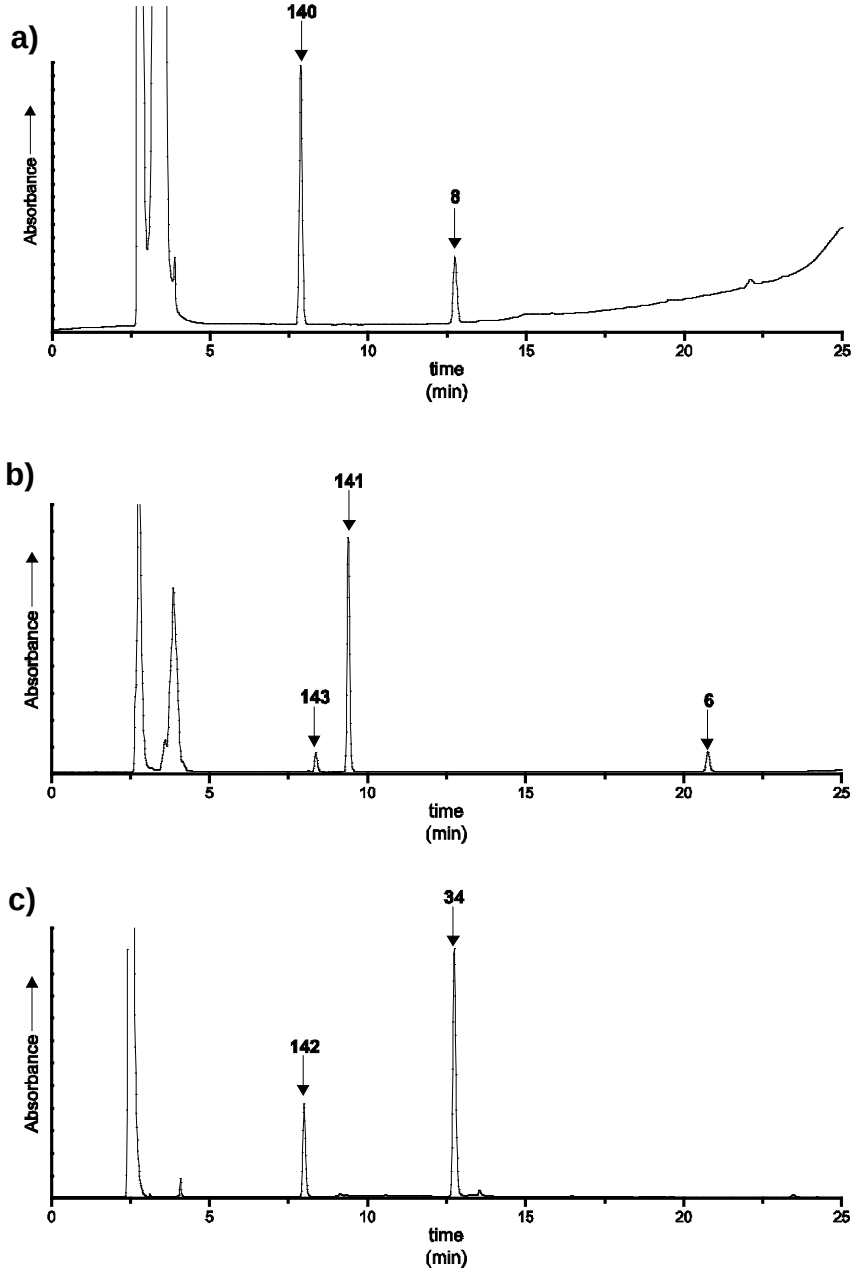


Figure 2.5. Representative HPLC traces for OleD reactions. Traces represent reactions with **a)** phenol (**8**), **b**) thiophenol (**6**), and **c**) aniline (**34**).

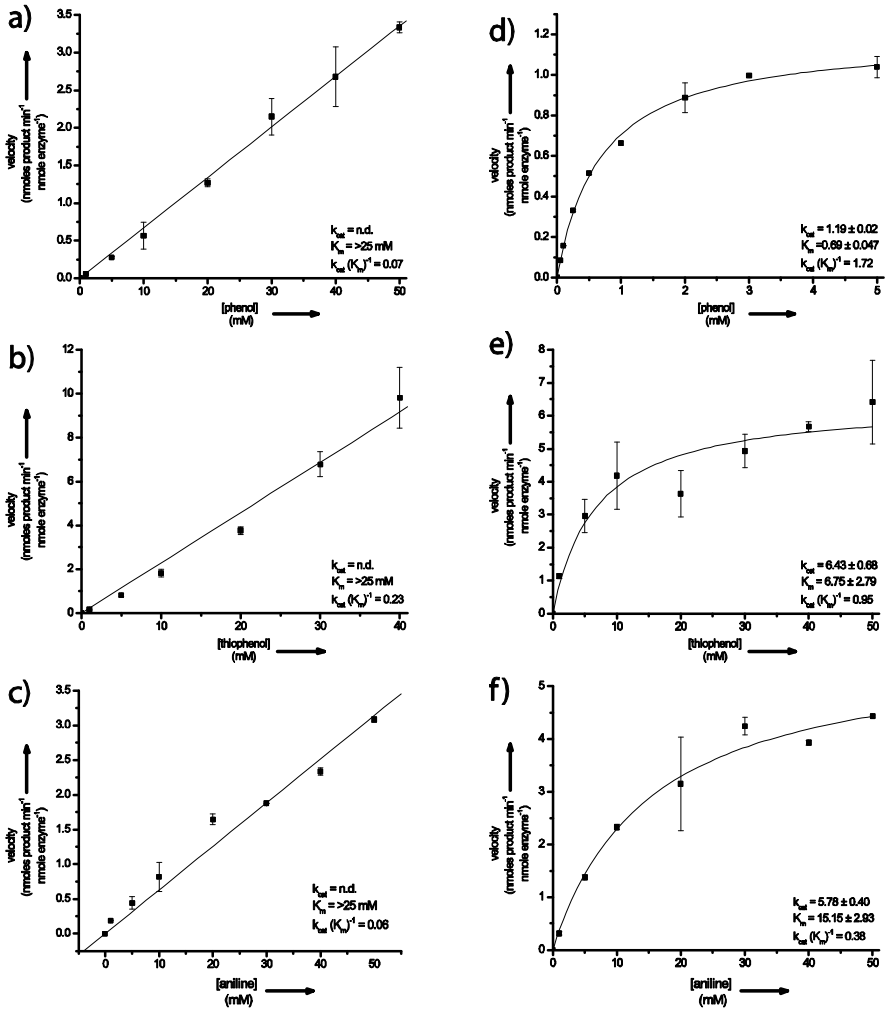


Figure 2.6. Determination of kinetic parameters for WT and 'ASP' OleD. Experiments were performed in triplicate with standard deviation noted. WT with varied acceptors **a)** phenol (**8**), **b)** thiophenol (**6**), and **c)** aniline (**34**). 'ASP' with varied acceptor **d)** phenol (**8**), **e)** thiophenol (**6**), and **f)** aniline (**34**).

Substrate	Enzyme	k_{cat} (min^{-1})	K_m (mM)	k_{cat} / K_m ($mM^{-1}min^{-1}$)	$(k_{cat} / K_m)_{ASP}/(k_{cat} / K_m)_{WT}$
phenol (**8**)	WT	n.a.[a]	n.a.	0.07	---
	'ASP'	1.19 ± 0.02	0.69 ± 0.05	1.72	24.6
thiophenol (**6**)	WT	n.a.	n.a.	0.23	---
	'ASP'	6.43 ± 0.68	6.75 ± 2.79	0.95	4.1
aniline (**34**)	WT	n.a.	n.a.	0.08	---
	'ASP'	5.78 ± 0.40	15.15 ± 2.93	0.38	4.8

[a] n.a., not available

Table 2.1. Kinetic parameters determined for WT and 'ASP' OleD. Parameters were determined with phenol (**8**), thiophenol (**6**), or aniline (**34**) and saturating UDP-glucose.

2.3 Conclusions

In terms of potential for combinatorial applications, this study revealed WT OleD and 'ASP' to glucosylate a diverse range of 'drug-like' scaffolds including anthraquinones, indolocarbozoles, polyenes, cardenolides, steroids, macrolides, beta-lactams, and enediynes (**Figure 2.1.b, Appendix 1**). Of particular note is that library members (or closely related compounds) are clinical analgesic (**28**), gout (**55**), congestive heart failure (**24**), hormone replacement (**18**), antifungal (**57**), antiparasitic (**43**), antibacterial (**36, 53, 63, 69**), and anticancer (**14, 17, 32, 38, 62**) agents[22]. The impressive range of scaffold structural diversity accessible by just a few GT variants suggests the potential to greatly expand the capabilities of chemoenzymatic glycorandomization from traditional natural product targets[9-14] to include fully synthetic scaffolds.

Although there exist a few reported natural or engineered 'bifunctional' GTs, capable of forming two types of glycosidic bonds (*O*/*N*-, *O*/*S*-, or *O*/*C*-)[23–26], this is the first example of a GT capable of catalyzing *O*-, *S*-, and *N*-glycosidic bond formation. Moreover, putative 'ASP'-catalyzed glycosidic bond formation was observed with aliphatic alcohols, aliphatic thiols, *N*-substituted anilines, oximes, hydrazines, hydrazides, *N*-hydroxyamides, *O*-substituted oxyamines, and carboxylic acids (**Figure 2.1.b**, **Appendix 1**), the heteroatom nucleophiles of which represent a pK_a range of ~ 4 – 14[27]. Although glycosides of oximes[28], hydrazines[29], hydrazides[30], *N*-hydroxyamides[31], and *O*-substituted oxyamines[15, 32, 33] have been chemically synthesized, to the best of this author's knowledge, this is the first enzyme-catalyzed route to these types of glycosidic linkages. While a few naturally-occurring iterative GTs also exist[14], this is the first report of OleD-catalyzed iterative glycosylation. This iterative activity is not entirely unexpected, as WT OleD typically catalyzes a (1→2)-β-D-glucosidic bond with the C-2

of deosamine for formation of a disaccharide moiety (**Figure 2.1a**)[34]. Cumulatively, the scaffold, nucleophile, and iterative adaptability of OleD clearly sets the stage for further engineering of custom GT catalysts for many applications which include, but are not limited to, *in vitro* drug discovery[9-14], *in vivo* fermentations or biotransformations to introduce or alter glycosylation patterns[35], and high throughput screening for catalytic enhancement[10, 36, 37].

In the context of GT promiscuity, the role of OleD in macrolide self-resistance (**1**-**2**, **Figure 2.1.a**)[34, 38] may require this enzyme to uniquely adapt to accommodate alternative acceptors. Yet, even with library members containing multiple nucleophiles, the majority of reactions studied were remarkably regio- and stereoselective. Although other reports have highlighted moderate promiscuity of natural product GTs[12, 16, 17, 39-41], the underlying structural basis for maintaining specificity with variant acceptors is not well understood. Studies designed to understand the promiscuity of other enzyme classes implicate dramatic enzyme conformational changes[42], voluminous active sites[43], or similar conformational binding of unrelated compounds within the active site[44] as potential contributors to promiscuity. Consistent with this, substrate modelling of the promiscuous GT VinC suggested that gross molecular size and hydrogen-bonding interactions play significant roles[41]. Future structural studies of enhanced OleD variants bound to diverse substrates are critical to extend our currently limited understanding of GT specificity and catalysis. Additionally, given the high conservation of the GT-B structural fold among natural product GTs[45], including OleD[46], other examples of dramatically promiscuous GTs are likely to emerge.

2.4 Materials and Methods

2.4.1 General materials. All chemicals and reagents were purchased from Sigma-Aldrich, Fluka, or New England Biolabs unless otherwise stated. The putative aglycons (**11**, **118-119**) [47] (**14**, **32**, and **38**)[48], (**16** and **136**)[49], (**45**, **70,** and **139**)[39], **55**[15], **73**[14], and **132-134** [50] have been previously reported. **29**, **50**, and **56** were a gift from Prof. Joe Langenhan (Seattle Univ., USA); **39** was a gift from Prof. Ben Shen (Scripps Research Institute, USA); **74** and **127** were a gift from Prof. Dr. Sarah O'Connor (Massachusetts Inst. of Tech., USA); **85** was a gift from Prof. Dr. Michael Thomas (Univ. of Wisconsin-Madison, USA); **105-112** were a gift from Prof. Dr. David L. Jakeman (Dalhousie University, Nova Scotia); **41** and **64** were isolated from fermentation. Plasmid pET28/OleD was obtained from Prof. Hung-Wen Liu (Univ. of Texas-Austin, USA).

2.4.1 General methods. LC/ESI-MS mass spectra were obtained using electrospray ionization on an Agilent 1100 HPLC-MSD SL quadrupole mass spectrometer connected to a UV/Vis diode array detector. High resolution mass spectra were determined utilizing electrospray ionization on a Waters (Micromass) LCT instrument (Beverly, MA, USA) with a time-of-flight analyzer and all HRMS samples contained an aliquot of a known compound (lock mass). Routine TLC analyses were accomplished on aluminum TLC plates coated with 0.2 mm silica gel from Sigma-Aldrich and monitoring at 254 nm. Flash column chromatography was achieved on 40 – 63 μm, 60 Å silica gel (Silicycle, Quebec, Canada). Unless otherwise noted, compounds were characterized by NMR with a ^{UNITY}INOVA 400 MHz instrument (Varian, Palo Alto, CA, USA) in conjunction with a QN Switchable BB probe (Varian, Palo Alto, CA, USA). ^{1}H and ^{13}C chemical shifts were referenced to internal solvent resonances and reported relative to TMS.

Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), qn (quintet), m (multiplet) and br (broad). Italicized elements or groups are those that are responsible for the shifts. Chemical shifts are reported in parts per million (ppm) and coupling constants *J* are given in Hz. NMR assignments were performed with the aid of COSY, TOCSY, HSQC, and HMBC experiments.

2.4.2. Protein expression and purification. Single colonies of *E. coli* BL21(DE3)pLysS (Stratagene, La Jolla, CA, USA) transformed with either pET28a/OleD or pET28a/OleD[P67T/ S132F/A242V] vector[9] were used to inoculate 3 mL Luria-Bertani (LB) medium supplemented with 50 μg mL^{-1} kanamycin and cultured overnight at 37 °C. The entire starter culture was then transferred to 1 L LB medium supplemented with 50 μg mL^{-1} kanamycin and grown at 37 °C until the OD_{600} reached ~0.7. Isopropyl β-D-thiogalactoside (IPTG) was subsequently added to a final concentration of 0.4 mM and the culture was incubated at 28ºC for approximately 18 hours. Cell pellets were collected by centrifugation at 10,000 *g* at 4 °C for 20 min. and the supernatant discarded. Pellets were resuspended in 10 mL lysis buffer (20 mM phosphate buffer, pH 7.4, 0.5 M NaCl, 10 mM imidazole) and were lysed by sonication. Cell debris was removed by centrifugation at 10,000 *g* at 4 °C for 20 min. and the cleared supernatant immediately applied to 2 mL of nickel nitrilotriacetic acid (Ni-NTA) resin (QIAgen Valencia, CA, USA) pre-equilibrated with lysis buffer. Protein was allowed to bind for 30 min. at 4 °C with gentle agitation, and the resin washed with 50 mL lysis buffer (x 4). Finally, the enzyme was eluted by incubation of the resin with 2 mL lysis buffer containing 250 mM imidazole for 10 min. at 4 °C with gentle agitation. The purified protein was applied to a PD-10 desalting column (Amersham Biosciences, Piscataway, NJ, USA), equilibrated with 50 mM Tris-HCl (pH 8.0), and eluted as

described by the manufacturer. Protein aliquots were immediately flash frozen in liquid nitrogen and stored at -80 °C. Protein purity was confirmed by SDS-PAGE to be >95% and protein concentration for all studies was determined using the Bradford Protein Assay Kit from Bio-Rad (Hercules, CA, USA).

2.4.3. Acceptor library screening. Reactions were conducted in a final volume of 100 µL and contained 50 µg of purified enzyme, 2.5 mM UDP-glucose, 50 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, and 1 mM of aglycon unless otherwise noted. Two separate control reactions for each aglycon that withheld either enzyme or UDP-glucose were performed in parallel. Reactions were allowed to proceed at 25^{o}C for ~16 hr, quenched with an equal volume of MeOH, centrifuged at 10,000 g for 10 min and the supernatant removed for analysis. The clarified reaction mixtures were analyzed by analytical reverse-phase HPLC with a 250 mm x 4.6 mm Gemini 5µ C18 column (Phenomenex, Torrance, CA, USA) using various methods (**Appendix 1**). HPLC peak areas were integrated with Star Chromatography Workstation Software (Varian, Palo Alto, CA, USA) and the total percent conversion calculated as a percent of the total peak area of substrate and product(s). Reactions which displayed potential new product(s) via HPLC were further analyzed by tandem LC/ESI-MS using a 250 mm x 4.6 mm Gemini 5µ C18 analytical column. If both WT and 'ASP' reactions resulted in chromatographically identical product(s), only the reaction with the highest percent conversion was analyzed via LC-MS. Chromatographic methods, reactant and product retention times, and MS determinations for all positive substrates are summarized in **Appendix 1**.

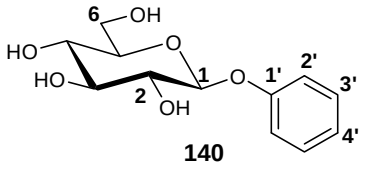
2.4.4. Determination of Kinetic Parameters. Assays were performed in a final volume of 200 µL 50 mM Tris-HCl (pH 8.0), and contained constant concentrations of enzyme (40 µg for 'ASP' reactions or 120 µg for WT reactions) and saturating UDP-glucose (2.5 mM) while varying the phenol (**8**), thiophenol (**6**), or aniline (**34**) concentration (0-50 mM). Aliquots (50 µL) were removed at 10 min (where the rate of product formation was determined to be linear), mixed with an equal volume of ice cold MeOH, and centrifuged at 10,000 *g* for 10 min. Supernatants were analyzed by analytical reverse-phase HPLC and percent conversion determined as described in the preceding section. All experiments were performed in triplicate. Initial velocities were fitted to the Michaelis-Menten equation using Origin Pro 7.0 software. OleD wild-type enzyme could not be saturated with acceptors **8**, **6**, and **34**. Consequentially, k_{cat}/K_m for wild-type was determined by linear regression.

2.4.5. Scale-Up and Characterization of Representative Glucosides.

2.4.5.1. General Reaction Procedure. (NMR characterization performed by Dr. R.D. Goff) Reactions were conducted in 10 mL Tris-HCl (50 mM, pH 8.0) containing substrate **6**, **8**, or **34** (20 mM), UDP-glucose (100 mM), and $MgCl_2$ (5 mM) at 25°C with agitation. Aliquots of 'ASP' enzyme (4.4 mg) were added to each reaction at 0, 7, and 19 hours. The reactions allowed to proceed for a total of 42 hours and were subsequently frozen (-80ºC), lyophilized, resuspended in 2 mL of ice cold MeOH, and filtered. Glucosides were isolated by collecting fractions from analytical reverse-phase HPLC with a 250 mm x 4.6 mm Gemini 5µ C18 column (Phenomenex, Torrance, CA, USA) using the appropriate method (**Appendix 1**). The product-containing fractions were subsequently collected and lyophilized. Products were confirmed by HRMS as

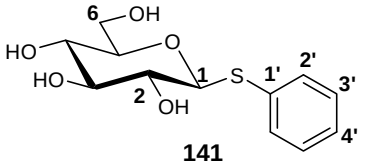
previously described and via ^{1}H and ^{13}C NMR using a Varian ^{UNITY}INOVA 500 MHz instrument (Palo Alto, CA, USA) with a Protasis/MRM CapNMR capillary probe (Savoy, IL, USA). ^{13}C assignments were attained through gHSQC and gHMBC methods.

2.4.5.2. Phenyl β-D-glucopyranoside (140). ^{1}H NMR (CD_3OD, 500 MHz) δ 7.32 (m, 2 H, H-3'), 7.14 (d, *J* = 8.2 Hz, 2 H, H-2'), 7.05 (m, 1 H, H-4'), 4.96 (d, *J* = 6.7 Hz, 1 H, H-1), 3.94 (d, *J* = 12.0 Hz, 1 H, H-6a), 3.75 (dd, *J* = 12.0, 5.0 Hz, 1 H, H-6b), 3.55-3.45 (m, 4 H, H-2, H-3, H-4, H-5); ^{13}C NMR (CD_3OD, 500 MHz) δ 158.0 (C-1'), 129.3, (C-3'), 122.3 (C-4'), 116.6 (C-2'), 101.2 (C-1), 76.9 (C-5), 73.8 (C2), 70.3 (C-4), 63.4 (C-3), 61.4 (C-6); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}NaO_6$, 279.0840; found 279.0833.



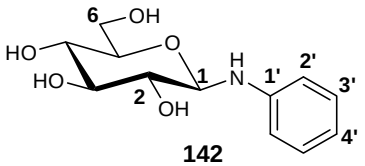
140

2.4.5.3. Phenyl 1-thio-β-D-glucopyranoside (141). ^{1}H NMR (CD_3OD, 500 MHz) δ 7.64 (d, *J* = 7.3 Hz, 2 H, H-2'), 7.38 (m, 2 H, H-3'), 7.35-7.30 (m, 1 H, H-4'), 4.68 (d, *J* = 9.6 Hz, 1 H, H-1), 3.95 (d, *J* = 12.0 Hz, 1 H, H-6a), 3.74 (dd, *J* = 12.0, 5.2 Hz, 1 H, H-6b), 3.46 (m, 1 H, H-3), 3.43-3.24 (m, 2 H, H-4, H-5), 3.29 (m, 1 H, H-2); ^{13}C NMR (CD_3OD, 500 MHz) δ 134.2 (C-1'), 131.5 (C-2'), 128.8 (C-3'), 127.2 (C-4'), 88.2 (C-1), 80.9 (C-4), 78.6 (C-3), 72.6 (C-2), 70.2 (C-5), 61.7 (C-6); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}NaO_5S$, 295.0611; found 295.0624.



141

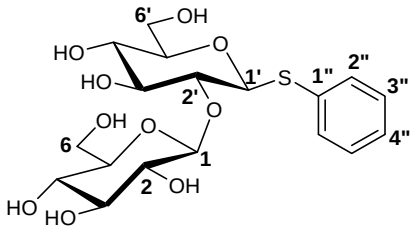
2.4.5.4. Phenyl 1-amino-β-D-glucopyranoside (142). ^{1}H NMR (CD_3OD, 500 MHz) δ 7.16 (m, 2 H, H-3'), 6.81 (d, *J* = 7.6 Hz, 2 H, H-2'), 6.73 (m, 1 H, H-4'), 4.57 (d, *J* = 7.8 Hz, 1 H, H-1), 3.87 (d, *J* = 11.7 Hz, 1 H, H-6a), 3.73-3.66 (m, 1 H, H-6b), 3.49 (t, *J* = 7.9 Hz, 1 H, H-3), 3.44-3.33 (m, 3 H, H-2, H-4, H-5); ^{13}C NMR (CD_3OD, 500 MHz) δ 147.0 (C-1'), 128.8, (C-3'), 118.4 (C-4'), 113.9 (C-2'), 85.8 (C-1), 77.9 (C-3), 77.2 (C-5), 73.5 (C-4), 70.7 (C-2), 61.6 (C-6); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{16}NNaO_5$, 278.0999; found 278.1006.



142

2.4.5.5. Phenyl β-D-glucopyranose-(1→2)-1-thio-β-D-glucopyranoside (143). ^{1}H NMR (acetone-d_6, 500 MHz) δ 7.61 (m, 2 H, H-2"), 7.33 (m, 2 H, H-3"), 7.28 (m, 1 H, H-4"), 4.77 (d, *J* = 9.8 Hz, 1 H, H-1'), 4.65 (d, *J* = 7.8 Hz, 1 H, H-1), 3.91-3.85 (m, 2 H, H-6A, H-6A'), 3.73-3.66 (m, 3 H, H-3', H-6B, H-6B'), 3.51-3.47 (m, 1 H, H-2'), 3.46-3.37 (m, 5 H, H-3, H-4, H-5, H-4', H-5'), 3.31-3.27 (m, 1 H, H-2); ^{13}C NMR (acetone-d_6, 500 MHz) δ 131.5 (C-2"), 128.9 (C-3"), 127.1 (C-4"), 104.6 (C-1), 85.7 (C-1'), 81.8 (C-2'), 80.7, 78.5 (C-3'), 77.3, 76.9, 75.3 (C-2),

71.0, 70.3, 62.4 (C-6, C-6'), 62.3 (C-6, C-6'); MS-ESI (*m/z*): $[M+Cl]^-$ calcd for $C_{18}H_{26}ClO_{10}S$, 469.1; 469.0 observed.[<1>]



143

2.4.6. Chemical syntheses of aglycons

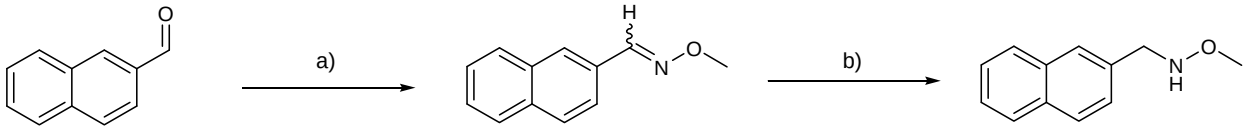
2.4.6.1. General reductive amination procedure. Aldehyde (12.5 mmol) was dissolved in CH_2Cl_2 to a final concentration of 0.45 M. To this was added 1.4 equivalents of $MeONH_2$•HCl and 2.2 equivalents of pyridine and the mixture was stirred for 2 hours at RT. TLC analysis at this stage revealed the substrate to be completely consumed with two products being formed. The reaction mixture was subsequently washed with 5% aqueous HCl (3 x 50 mL) and saturated NaCl (2 x 50 mL). The resulting organic layer was dried over Na_2SO_4 and concentrated under reduced pressure to provide the crude oxime (>80%) which was used in subsequent reactions without further purification.

Crude oxime was dissolved in EtOH to a final concentration of 1.5 M. The reaction mixture was cooled to 0 ºC, 3 equivalents of $NaBH_3CN$ were added, and the solution was stirred for 15 min. An equal volume of 20% HCl in EtOH chilled to 0 °C was subsequently added in a

[<1>] The disaccharide linkage of **143** was determined to be 1→2 through comparison of COSY and TOCSY NMR experiments of the peracetylated analogs of **143** and its monosaccharide, **141**. The H-2 proton of peracetylated **141** appears at 4.97 ppm ($CDCl_3$, 500 MHz), while the H-2 proton of the proximal sugar of peracetylated **143** is at 3.86 ppm, indicating the position of glycosylation.

drop-wise fashion over 10 min. The reaction was then allowed to warm to room temperature and stirred overnight. TLC analysis revealed complete consumption of substrate and a single new product. The reaction was neutralized with the addition of Na_2CO_3 until the evolution of gas halted, concentrated under reduced pressure, and CH_2Cl_2 (20 mL) was added. The resulting mixture was washed with saturated $NaHCO_3$ (2 x 50 mL), dried over Na_2SO_4, and the collected organics concentrated under reduced pressure. The concentrate was purified by flash chromatography (1:1 hexanes:CH_2Cl_3) to yield the desired methoxyamine product in > 50% yield.

2.4.6.2. Synthesis of *N*-methoxy-2-naphthalenemethanamine (26).



144 **145** **26**

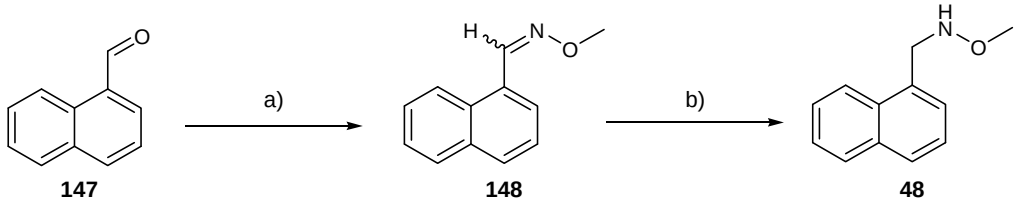
a) $MeONH_2 \cdot HCl$, CH_2Cl_2, pyridine; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C

2-Naphthaldehyde (**144**; 2.0 g, 12.5 mmol) provided the desired oxime **145** (2.3 g, 99% crude yield) as a white solid. TLC R_f = 0.48, 0.60 (1:8 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (s, 1 H, NCH), 7.92-7.80 (m, 5 H, Ph), 7.56-7.48 (m, 2 H, Ph), 4.058 (s, 3 H, OCH_3),[<2>] 4.055 (s, 3 H, OCH_3)2; ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.0, 134.4, 133.5, 130.2, 128.9, 128.6, 128.6, 128.2, 127.2, 126.9, 62.4; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{12}NO$, 186.0841; found 186.0920.

[<2>] Signals observed in ^{1}H spectrum at 4.058 and 4.055 for OCH_3 are from the presence of (*E*) and (*Z*)-isomers.

Oxime **145** (2.3 gm, 12.4 mmol) yielded **26** (1.5 g, 63% yield) as an orange oil. TLC R_f = 0.36 (1:4 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.84-7.74 (m, 4 H, Ph), 7.52-7.39 (m, 3 H, Ph), 5.80 (br s, 1 H, NH), 4.18 (s, 2 H, NHC*H*$_2$), 3.50 (t, *J* = 0.4 Hz, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 135.5, 133.7, 133.1, 128.4, 128.1, 128.0, 127.9, 127.2, 126.3, 126.1, 62.2, 56.6; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}NO$, 188.1070; found 188.1069.

2.4.6.3. Synthesis of *N*-methoxy-1-naphthalenemethanamine (48).



147 **148** **48**

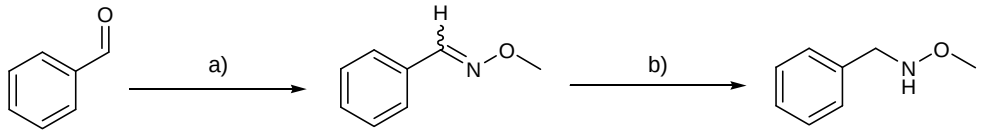
a) $MeONH_2$·HCl, CH_2Cl_2, pyridine; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C

1-Naphthaldehyde (**147**; 2.0 g, 12.7 mmol) provided the desired oxime **148** (2.0 g, 83% crude yield) as a yellow oil. TLC R_f = 0.56, 0.65 (1:4 EtOAc:hexanes,); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.71 (s, 1 H, NCH), 8.52 (m, 1 H, Ph), 7.84 (m, 2 H, Ph), 7.74 (m, 1 H, Ph), 7.58-7.42 (m, 3 H, Ph), 4.06 (s, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.5, 133.9, 130.8, 130.5, 128.8, 128.1, 127.4, 127.1, 126.2, 125.4, 124.6, 62.2; HRMS-ESI (*m/z*): $[M]^{+\bullet}$ calcd for $C_{12}H_{11}NO$, 185.0836; found 185.0842.

Oxime **148** (2.0 gm, 10.5 mmol) yielded **48** (1.2 g, 63% yield) as a yellow oil. TLC R_f = 0.41 (1:4::EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.15 (d, *J* = 8.8 Hz, 1 H, Ph), 7.84 (d, *J* = 8.4 Hz, 1 H, Ph), 7.77 (d, *J* = 8.4 Hz, 1 H, Ph), 7.56-7.36 (m, 4 H, Ph), 5.77 (br s, 1 H, NH), 4.51 (s, 2 H, NHC*H*$_2$), 3.53 (d, *J* = 0.4 Hz, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.1, 133.0,

132.3, 129.0, 128.7, 127.8, 126.5, 126.0, 125.7, 124.0, 62.1, 54.1; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}NO$, 188.1070; found 188.1070.

2.4.6.4. Synthesis of *N*-methoxybenzylamine (115).



149 **150** **115**

a) $MeONH_2 \cdot HCl$, CH_2Cl_2, pyridine; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C

Benzaldehyde (**149**; 4.8 g, 49.2 mmol) provided the desired oxime **150** (6.1 g, 91% crude yield) as a colorless oil. TLC R_f = 0.82 (1:8 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.03 (s, 1 H, NCH), 7.58-7.51 (m, 2 H, Ph), 7.37-7.32 (m, 3 H, Ph), 3.95 (s, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.8, 132.5, 130.0, 129.0, 127.3, 62.2[<3>]; HRMS-ESI (*m/z*): $[M]^{+\bullet}$ calcd for C_8H_9NO, 135.0679; found 135.0684.

Oxime **150** (6.1 g, 44.9 mmol) provided desired methoxyamine **115** (3.3 g, 53% yield) as a colorless oil. TLC R_f = 0.31 (1:8 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38-7.22 (m, 5 H, Ph), 5.71 (br s, 1 H, NH), 4.04 (s, 2 H, C*H*$_2$NH), 3.50 (d, *J* = 0.4 Hz, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 137.9, 129.1, 128.7, 127.7, 62.1, 56.5[<4>]; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_8H_{12}NO$, 138.0914; found 138.0916.

[<3>] Not all phenyl ^{13}C resonances were observed, presumably due to spectral overlap in the 132.5-127.3 ppm region.
[<4>] Not all phenyl ^{13}C resonances were observed, presumably due to spectral overlap in the 129.1-127.1 ppm region.

2.5 References

1. Křen, V. & Řezanka, T. Sweet antibiotics – the role of glycosidic residues in antibiotic and antitumor activity and their randomization. *FEMS Microbiol. Rev.* **32**, 858-889 (2008).

2. Thorson, J.S., Hosted Jr., T. J., Jiang, J., Biggins, J.B., & Ahlert, J. Nature's carbohydrate chemists: the enzymatic glycosylation of bioactive materials. *Curr. Org. Chem.* **5**, 139-167 (2001**)**.

3. Weymouth-Wilson, A.C. The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* **14**, 99-110 (1997).

4. Salas, J.A. & Méndez, C. Engineering the glycosylation of natural products in actinomycetes. *Trends Microbiol.* **15**, 119-232 (2007).

5. Thibodeaux, C.J., Melançon, C.E., & Liu, H. Unusual sugar biosynthesis & natural product glycodiversification. *Nature* **446**, 1008-1016 (2007).

6. Blanchard, S. & Thorson, J.S. Enzymatic tools for engineering natural product glycosylation. *Curr. Opin. Chem. Biol.* **10**, 263-271 (2006).

7. Griffith, B.R., Langenhan, J.M., & Thorson, J.S. 'Sweeting' natural products via glycorandomization. *Curr. Opin. Biotech.* **16**, 622-630 (2005).

8. Langenhan, J.M., Griffith, B.R., & Thorson, J.S. Neoglycorandomization and chemoenzymatic glycorandomization: two complementary tools for natural product diversification. *J. Nat. Prod.* **68**, 1696-1711 (2005).

9. Williams, G.J., Zhang, C., & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

10. Williams, G.J. & Thorson, J.S. A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Protoc.* **3**, 357-362 (2008).

11. Williams, G.J., Goff, R.D., Zhang, C., & Thorson, J.S. Optimizing glycosyltransferase specificity via ''hot spot'' saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem. Biol.* **15**, 393-401 (2008).

12. Fu, X., *et al.* Antibiotic optimization via *in vitro* glycorandomization. *Nat. Biotech.* **21**, 1467-1469 (2003).

13. Zhang, C., *et al.* Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science* **313**, 1291-1294 (2006).

14. Zhang, C., Albermann, C., Fu, X., & Thorson, J.S. The *in vitro* characterization of the iterative avermectin glycosyltransferase AveBI reveals reaction reversibility and sugar nucleotide flexibility. *J. Am. Chem. Soc.* **128**, 16420-16421 (2006).

15. Ahmed, A., *et al.* Colchicine glycorandomization influences cytotoxicity and mechanism of action. *J. Am. Chem. Soc.* **128**, 14224-14225 (2006).

16. Ghiorghis, A., Talebian, A., & Clarke, R. *In vitro* antineoplastic activity of C7-substituted mitomycin C analogues MC-77 and MC-62 against human breast-cancer cell lines. *Cancer Chemother. Pharmacol.* **29**, 290-296 (1992).

17. Imbert, T.F. Discovery of podophyllotoxins. *Biochimie* **80**, 207-222 (1998).

18. Abel, M., Szweda, R., Trepanier, D., Yatscoff, R.W., & Foster, R.T. Rapamycin carbohydrate derivatives. *U.S. Patent No. US 7,160,867* (2007**)**.

19. Veyhl, M., *et al.* Transport of the new chemotherapeutic agent beta-D-glucosylisophosphoramide mustard (D-19575) into tumor cells is mediated by the Na^+-D-glucose co-transporter SAAT1. *Proc. Natl. Acad. Sci. USA* **95**, 2914-2919 (1998).

20. Liu, D., Sinchaikeul, S., Reddy, P.V.G., Chang, M., & Chen, S. Synthesis of 2'-paclitaxel methyl 2-glucopyranosyl succinate for specific targeted delivery to cancer cells. *Bioorg. Med. Chem. Lett.* **17**, 617-620 (2007).

21. Yang, M., *et al.* 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 (2005).

22. Lacy, C.F., Armstrong, L.L., Goldman, M.P., & Lance, L.L. (eds.) *Drug Information Handbook*, 16th Ed. (Lexi-Comp, Hudson, Ohio, USA, 2008) pp. 22-30, 112-114, 314-315, 381-382, 427-430, 580-585, 723-725, 1151-1152, 1350-1351, 1609-1610, 1065-1067.

23. Loutre, C., *et al.* Isolation of a glucosyltransferase from Arabidopsis thaliana active in the metabolism of the persistent pollutant 3,4-dichloroaniline. *Plant J.* **34**, 485-493 (2003).

24. Brazier-Hicks, M., *et al.* Characterization and engineering of the bifunctional *N*- and *O*-glucosyltransferase involved in xenobiotic metabolism in plants. *Proc. Natl. Acad. Sci. USA* **104**, 20238-20243 (2007).

25. Patana, A., Kurkela, M., Finel, M., & Goldman, A. Mutation analysis in UGT1A9 suggests a relationship between substrate and catalytic residues in UDP-glucuronosyltransferases. *Protein Eng. Des. Sel.* **21**, 537-543 (2008).

26. Dürr, C., *et al.* The glycosyltransferase UrdGT2 catalyzes both *C*- and *O*-glycosidic sugar transfers. *Angew. Chem. Int. Ed. Engl.* **43**, 2962-2965 (2004).

27. Data available at http://www.chem.wisc.edu/areas/reich/pkatable/

28. Rodriguez, E.C., Marcaurelle, L.A., & Bertozzi, C.R. Aminooxy-, hydrazide-, and thiosemicarbazide-functionalized saccharides: versatile reagents for glycoconjugate synthesis. *J. Org. Chem.* **63**, 7134-7135 (1998).

29. J.P. Holland, *et al.* Functionalized bis(thiosemicarbazonato) complexes of zinc and copper: synthetic platforms toward site-specific radiopharmaceuticals. *Inorg. Chem.* **46**, 465-485 (2007).

30. Takeda, Y. Determination of sugar hydrazide and hydrazone structures in solution. *Carb. Res.* **77**, 9-23 (1979).

31. Thomas, M., Gesson, J., & Papot, S. First *O*-glycosylation of hydroxamic acids. *J. Org. Chem.* **72**, 4262-4264 (2007).

32. Peri, F., Deutman, A., La Ferla, B., & Nicotra, F. Solution and solid-phase chemoselective synthesis of (1-6)-amino(methoxy) di- and trisaccharide analogues. *ChemComm* **14**, 1504-1505 (2002).

33. Carrasco, M.R., Brown, R.T. A versatile set of aminooxy amino acids for the synthesis of neoglycopeptides. *J. Org. Chem.* **68**, 8853-8858 (2003).

34. Quirós, L.M., Carbajo, R.J., Braña, A.F., & Salas, J.A. Glycosylation of macrolide antibiotics: purification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*. *J. Biol. Chem.* **275**, 11717-11720 (2000).

35. Williams, G.J., Yang, J., Zhang, C., & Thorson, J.S. Recombinant *E. coli* prototype strains for *in vivo* glycorandomization. *ACS Chem. Biol.* **6**, 95-100 (2011).

36. Aharoni, A., *et al.* High-throughput screening methodology for the directed evolution of glycosyltransferases. *Nat. Methods* **3**, 609-614 (2006).

37. Persson, M., & Palcic, M.M. A high-throughput pH indicator assay for screening glycosyltransferase saturation mutagenesis libraries. *Anal. Biochem.* **378**, 1-7 (2008).

38. Salas, J.A., *et al.* Intracellular glycosylation and active efflux as mechasnisms for resistance to oleandomycin in *Streptomyces antibioticus*, the producer organism. *Microbiologia* **10**, 37-48 (1994).

39. Zhang, C., *et al.* RebG- and RebM-catalyzed indolocarbazole diversification. *ChemBioChem* **7**, 795-804 (2006).

40. Meyers, C.L.F., Oberthür, M., Anderson, J.W., Kahne, D., & Walsh, C.T. Initial characterization of novobiocic acid noviosyl transferase activity of NovM in biosynthesis of the antibiotic novobiocin. *Biochem.* **42**, 4179-4189 (2003).

41. Minami, A., Uchida, R., Eguchi, T., & Kakinuma, K. Enzymatic approach to unnatural glycosides with diverse aglycon scaffolds using glycosyltransferase VinC. *J. Am. Chem. Soc.* **127**, 6148-6149 (2005).

42. Ekroos, M. & Sjögren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA* **103**, 13682-13687 (2006).

43. Kuzuyama, T., Noel, J.P., & Richard, S.B. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* **435**, 983-987 (2005).

44. Liu, C., *et al.* Structural basis for dual functionality of isoflavonoid *O*-methyltransferases in the evolution of plant defense responses. *Plant Cell* **18**, 3656-3669 (2006).

45. Lairson, L.L., Henrissat, B., Davies, G.J., & Withers, S.G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521-555 (2008).

46. Bolam, D.N., *et al.* The crystal structure of two macrolide glycosyltransferases provides a blueprint for host cell antibiotic immunity. *Proc. Natl. Acad. Sci. USA* **104**, 5336-5341 (2007).

47. Langenhan, J.M., Peters, N.R., Guzei, I.A., Hoffmann, F.M., & Thorson, J.S. Enhancing the anticancer properties of cardiac glycosides by neoglycorandomization. *Proc. Natl. Acad. Sci. USA* **102**, 12305-12310 (2005).

48. Zhang, C., *et al.* Biochemical and structural insights of the early glycosylation steps in calicheamicin biosynthesis. *Chem. Biol.* **15**, 842-853 (2008).

49. Peltier-Pain, P., Timmons, S.C., Grandemange, A., Benoit, E., & Thorson, J.S. Warfarin glycosylation invokes a switch from anticoagulant to anticancer activity. *ChemMedChem* **6**, 1347-1350 (2011).

50. Griffith, B.R., *et al.* Model for antibiotic optimization via neoglycosylation: synthesis of liponeoglycopeptides active against VRE. *J. Am. Chem. Soc.* **129**, 8150-8155 (2007).

Chapter 3:

Using Simple Donors to Drive the Equilibria of Glycosyltransferase-Catalyzed Reactions

Portions of this chapter have been previously published as:

Gantt, R.W., Peltier-Pain, P.P., Cournoyer, W.J., and Thorson, J.S. Using simple donors to drive the equilibria of glycosyltransferase-catalyzed reactions. *Nat. Chem. Biol.* **7**, 685-691 (2011).

3.1. Introduction

Glycosyltransferases (GTs) constitute a predominant enzyme superfamily responsible for the attachment of carbohydrate moieties to a wide array of acceptors that include nucleic acids, polysaccharides, proteins, lipids, carbohydrates and medicinally relevant secondary metabolites [1, 2]. The majority of GTs are LeLoir (sugar nucleotide-dependent) enzymes and utilize nucleoside diphosphate sugars (NDP-sugars) as 'donors' for glycosidic bond formation (**Figure 3.1a**). Recent studies have revealed certain GT-catalyzed reactions from bacterial secondary metabolism to be reversible, presenting new GT-catalyzed methods for NDP-sugar synthesis as well as the GT-catalyzed exchange (**Figure 3.1b**) or transfer (**Figure 3.1c**) of sugars attached to both simple glycoside sugar donors[3, 4] and complex natural product glycosides including glycopeptides, enediynes[5], macrolides[6], macrolactams[7], (iso)flavonoids[8] and polyenes[9]. Of the few examples that employed 'activated' sugar donors (*e.g.*, glycosyl halides or aromatic glycosides) for GT-catalyzed transglycosylation[3, 4], only indirect evidence for the intermediacy of sugar nucleotides was provided. Among GT-catalyzed 'reverse' reactions where NDP-sugar formation has been confirmed, NDP-sugar formation was thermodynamically disfavored (*i.e.*, with NDP-sugar as product, $K_{eq} < 1$)[5, 10, 11] and typically required a 10- to 100-fold excess of NDP for sugar nucleotide production.

To address the severe thermodynamic limitations of reverse GT-catalyzed reactions, a series of small activated aromatic glucosides were investigated for their ability to participate

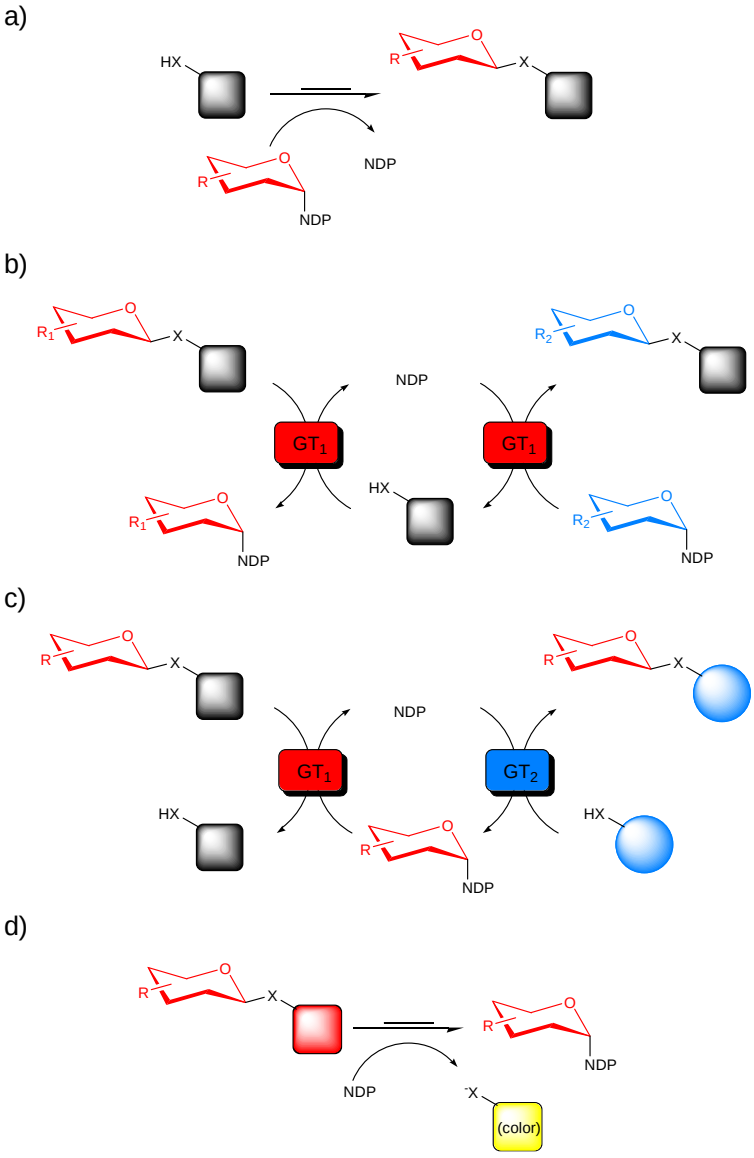


Figure 3.1. Representative GT-catalyzed reactions. a) Classical GT-catalyzed transformation wherein the sugar, presented in the form of a sugar nucleotide donor, is conjugated to an acceptor target of interest to provide a thermodynamically-favored glycoside product. **b)** A GT-catalyzed 'sugar exchange' reaction. In this reaction, a small amount of NDP is used to 'prime' the removal of the endogenous sugar appendage of a target complex natural product thus enabling the exchange of a native sugar for an endogenous sugar supplied in vast excess as a sugar nucleotide. **c)** A GT-catalyzed 'aglycon exchange' reaction where the sugar from one complex natural product is excised (using excess NDP) and subsequently attached to a structurally distinct target aglycon. **d)** The present study demonstrates the use of simple activated glycosides to dramatically shift the thermodynamics of GT-catalyzed reactions and thereby drive GT-catalyzed NDP-sugar synthesis, sugar exchange and/or aglycon exchange reactions while also offering a convenient colorimetric screen for glycosylation. X = O-, S-, or NH-.

as donors in the formation of NDP-sugars (**Figure 3.1d**). While a number of donors were identified, a particular subset were found to dramatically alter the equilibrium of GT-catalyzed reactions and thereby enabled a variety of novel transformations including: ***i)*** a unique platform for the efficient enzymatic syntheses of novel NDP-sugars, ***ii)*** a coupled GT-catalyzed platform for the differential glycosylation of small molecules (including natural products and synthetic drugs/targets), and ***iii)*** a colorimetric readout upon glycosyl-transfer amenable to high throughput formats for glycodiversification and glycoengineering that can be coupled to nearly any downstream sugar nucleotide-utilizing enzyme.

3.2. Results and Discussion

3.2.1. Protein Expression and Purification. All GT enzymes utilized in this study were expressed at high levels (>10 mg L^{-1} of culture) and, following purification, purity of each catalyst was confirmed by SDS-PAGE to be >95% (**Figure 3.2**, see section **3.4**).

3.2.2. Screening of β-D-glucoside donors. The availability of a GT capable of utilizing a wide array of both simple aromatic acceptors and sugar nucleotides set the stage for this systematic study. With the aid of a crystal structure[12], previous directed evolution and engineering of an inverting macrolide-inactivating GT (OleD) from *S. antibioticus* identified several highly permissive variants for both sugar nucleotides (14 known sugar substrates) and acceptors (>70 structurally diverse known substrates) in the context of the forward reaction[13-18]. Utilizing the aglycons recognized in forward reactions as a template, a set of 32 putative β-D-glucopyranoside donors (**1-32**, **Figure 3.3**) were synthesized (see section **3.4**) and tested against a

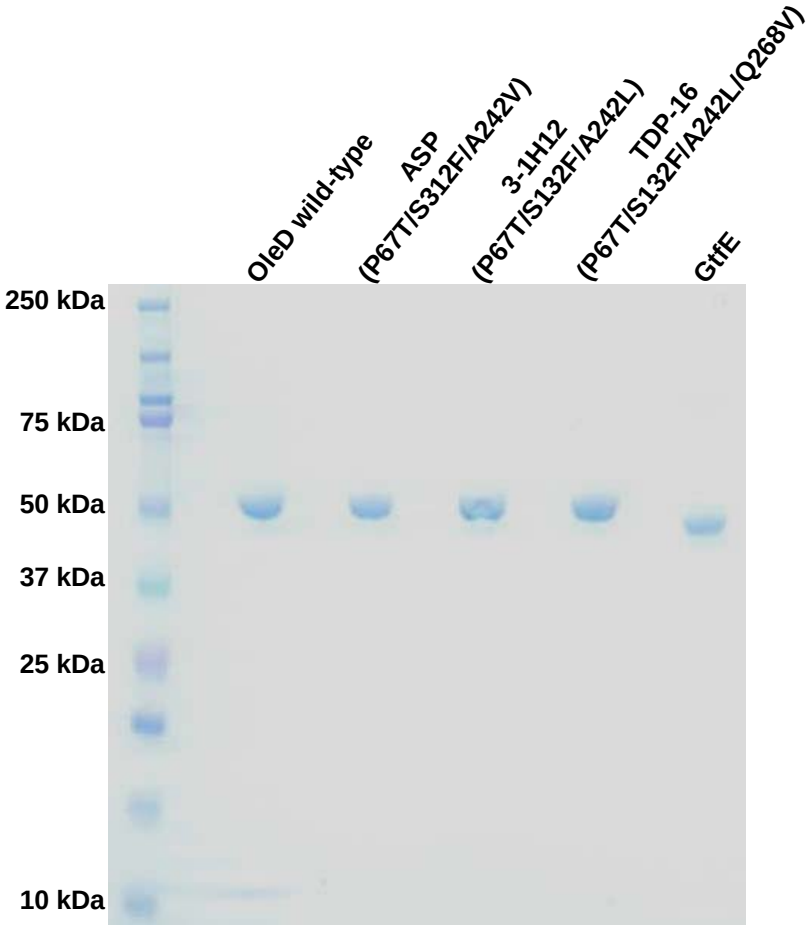


Figure 3.2. SDS-PAGE analysis of purified proteins. The expected molecular weights of OleD (and variants) and GtfE are 48 and 46 kDa, respectively, and correlate with the observed bands.

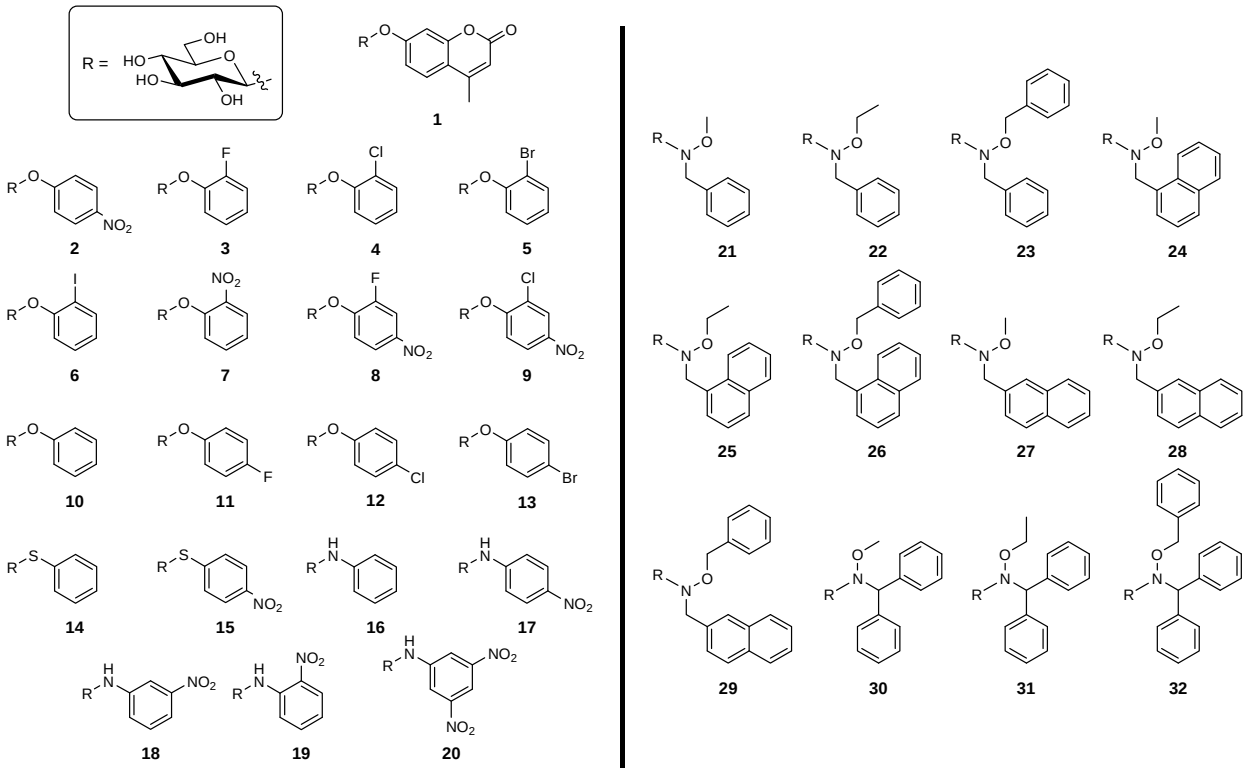


Figure 3.3. Structures of all evaluated β-D-glucopyranosides. Syntheses and commercial sources are noted in section **3.4**.

series of OleD variants in reverse reactions for production of UDP-α-D-glucose (UDP-Glc, **33a**) in the presence of UDP (**Figure 3.4a**). The syntheses of these putative donors required 1-3 steps (37% average overall yield) and, in all but one case (**19**), provided the desired β-anomer exclusively. Of the 32 putative donors evaluated, 9 (**1-9**) led to UDP-Glc (**33a**) formation with all OleD variants examined (**Figure 3.4b**). This systematic analysis revealed a clear correlation between the leaving group ability of the sugar donor and the production of desired sugar nucleotide wherein the combination of OleD variant TDP-16 (containing the mutations P67T/S132F/A242L/Q268V)[17] and 2-chloro-4-nitrophenyl β-D-glucopyranoside (**9**) provided the best overall yields of UDP-Glc (**33a**) or TDP-Glc (**33b**) (**Figure 3.4c**).

3.2.3. Optimized pH for the reverse reaction. In reverse reactions with 2-chloro-4-nitrophenyl β-D-glucopyranoside (**9**) as donor, UDP as acceptor, and TDP-16 as catalyst, maximal turnover was observed at pH 7.0-8.5 in Tris buffer (**Figure 3.5**, see section **3.4**). Significant drops in the rate of 2-chloro-4-nitrophenolate formation were observed outside of this pH range or when the buffer type was varied. This pH range is consistent with the previously reported pH-rate profile for the wild-type OleD in the forward direction[10].

3.2.4. Alternative NDP donors for the reverse reaction. In reverse reactions with 2-chloro-4-nitrophenyl β-D-glucopyranoside (**9**) as donor, various NDPs as acceptors, and TDP-16 as catalyst, NDP-sugar formation was also observed in the presence of ADP and GDP, albeit with much lower efficiency than with UDP or TDP (**Figure 3.6** and **3.7,** see section **3.4**). No turnover was observed in the presence of CDP. Thus, four of the five standard nucleotide moieties utilized by all LeLoir GTs (including not only natural product GTs but also those which

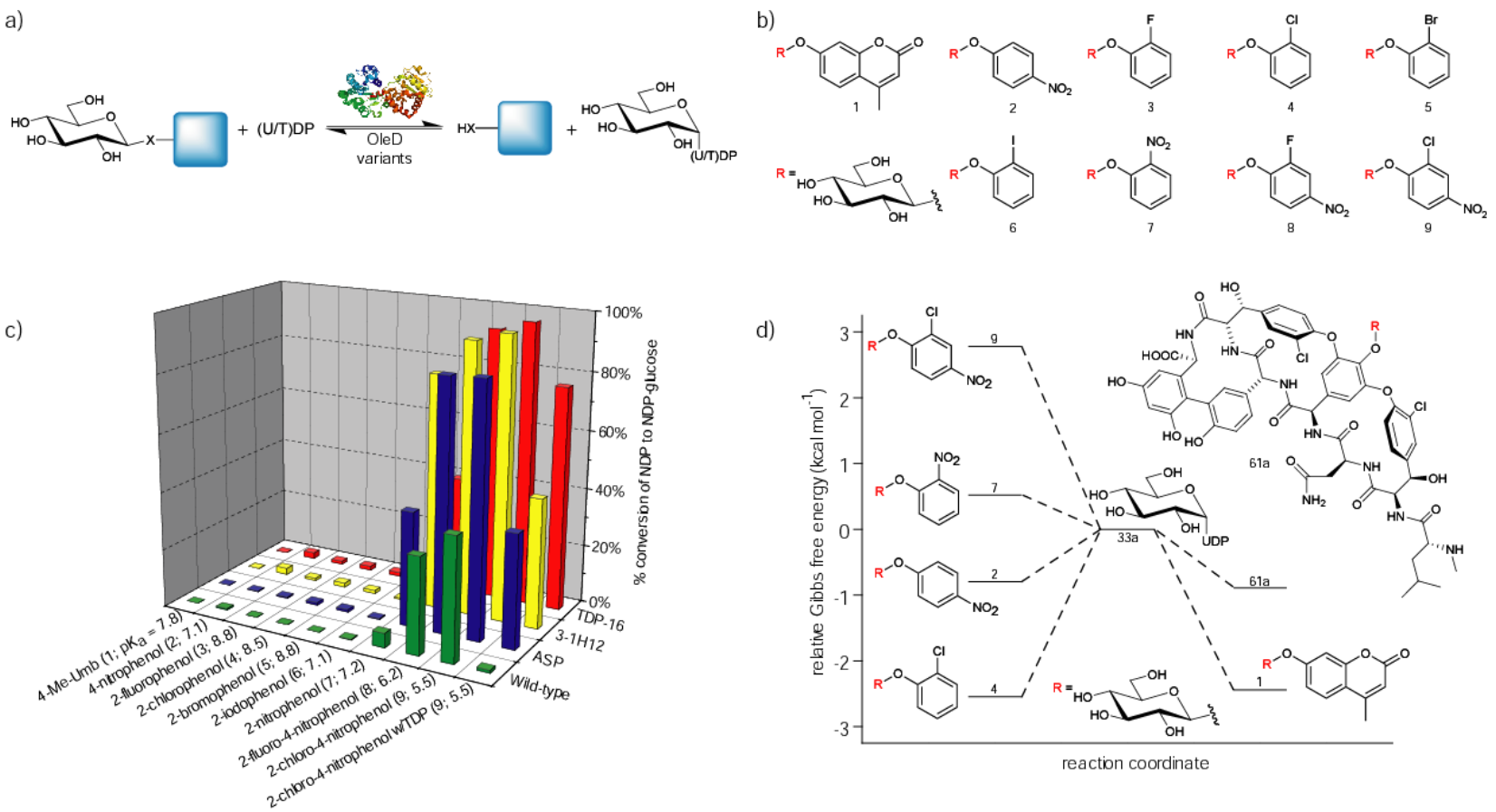


Figure 3.4. Evaluation of putative donors for sugar nucleotide synthesis. a) General reaction scheme. **b**) Structures of the β-D-glucopyranoside donors which led to (U/T)DP-glucose formation. **c**) Percent conversion of (U/T)DP to (U/T)DP-glucose with various donors. Reactions contained 2.1 μM (10 μg) OleD variant, 1 mM of (U/T)DP, and 1 mM of aromatic donor (**1**–**9**) in Tris-HCl buffer (50 mM, pH 8.5) with a final volume of 100 μl. After one hour at 25 °C, reactions were flash frozen and analyzed by HPLC (see section **3.4**). The pK_a for each corresponding donor aglycon is highlighted in parentheses. **d**) Plot depicting the relative Gibbs free energy of selected donors/acceptors in relation to **33a**. Small glycoside donors display large shifts in relative free energy, transforming formation of UDP-Glc (**33a**) from an endo- to an exothermic process. The $\Delta G°_{pH8.5}$ for **1**, **2, 4, 7,** and **9** with UDP in Tris-HCl buffer (50 mM, pH 8.5) at 298K relative to **33a** were determined in this study (see section **3.4**). The ΔG° for **61a** was previously determined (at pH 9.0 and 310K)[(5)].

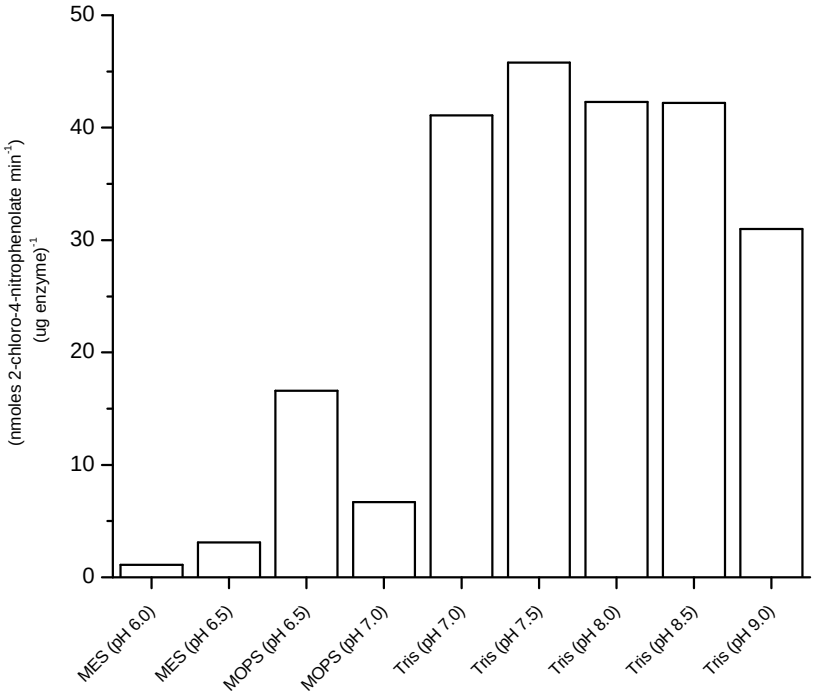


Figure 3.5. The rate of 2-chloro-4-nitrophenolate production as a function of buffer and pH.

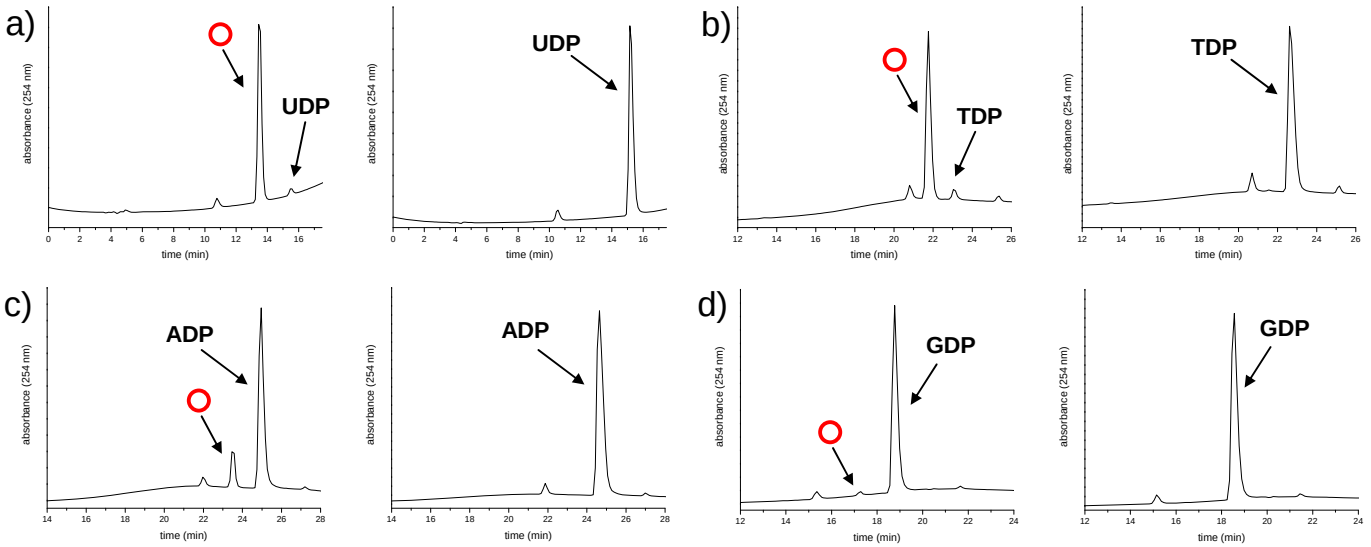


Figure 3.6. HPLC traces of reverse reactions with NDPs. Data for **a)** UDP, **b)** TDP, **c)** ADP, **d)** GDP containing reverse reactions. HPLC traces are displayed on the left with the putative NDP-α-D-glucose product labeled a red circle; NDP standards are displayed on the right. No new product peaks were observed in the presence of CDP.

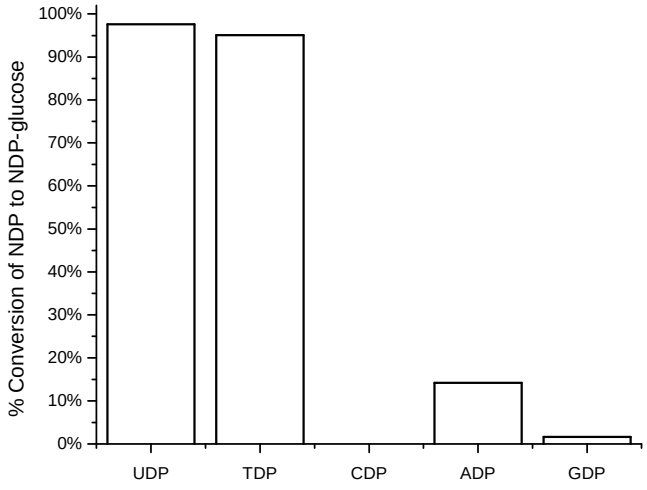


Figure 3.7. Percent conversions of NDP to NDP-glucose. Conversion as determined by HPLC (see **Figure 3.6**).

catalyze the formation of glycoproteins[19-21], oligosaccharides[21-25], glycolipids[26], glycoconjugates[1], etc.) are accessible via this method.

3.2.5. Scale-up and characterization of NDP-sugars To demonstrate preparative scale and provide material for full characterization, reverse GT reactions were conducted with a 1:1 molar ratio of glucoside donor to NDP using 9 mg of UDP or TDP to provide 6.9 (55% isolated yield) and 7.7 mg (61% isolated yield), of UDP-Glc (**33a**) and TDP-Glc (**33b**) (**Figure 3.8 ,** see section **3.4**).

3.2.6. Kinetic evaluations. Under saturating donor **9**, kinetic analysis (**Materials and Methods**) revealed the k_{cat}/K_m of TDP-16 to be improved by a factor of 25 or 315 (varied UDP or TDP, respectively) compared to wild-type OleD (**Figure 3.9 and 3.10, Table 3.1**). These results are consistent with this mutant's enhanced proficiency toward TDP-sugars[17].

3.2.7. Thermodynamic evaluations. Equilibrium constants ($K_{eq,pH8.5}$) were determined for **1**, **2**, **4**, **7**, and **9** and utilized to calculate the corresponding Gibbs free energy according to equation (*1*) (**Figure 3.11**, **Table 3.2**, see section **3.4**).

$$(1) \qquad \Delta G°_{pH8.5} = -RT\ln(K_{eq,pH8.5})$$

In agreement with the typically observed thermodynamics for these reactions[5, 10, 11], the **4-**, **1-**, or **2-** ($\Delta G°_{pH8.5}$ = +2.55, +2.44, and +0.92 kcal mol^{-1}, respectively) UDP-Glc transformations were endothermic. In stark contrast, **7-** or **9-** ($\Delta G°_{pH8.5}$ = -0.52 and -2.78 kcal mol^{-1}, respectively) UDP-Glc transformations were notably exothermic (**Figure 3.4d**) and thereby correspond to a dramatic shift of GT-catalyzed reaction K_{eq} which markedly favors NDP-sugar formation.

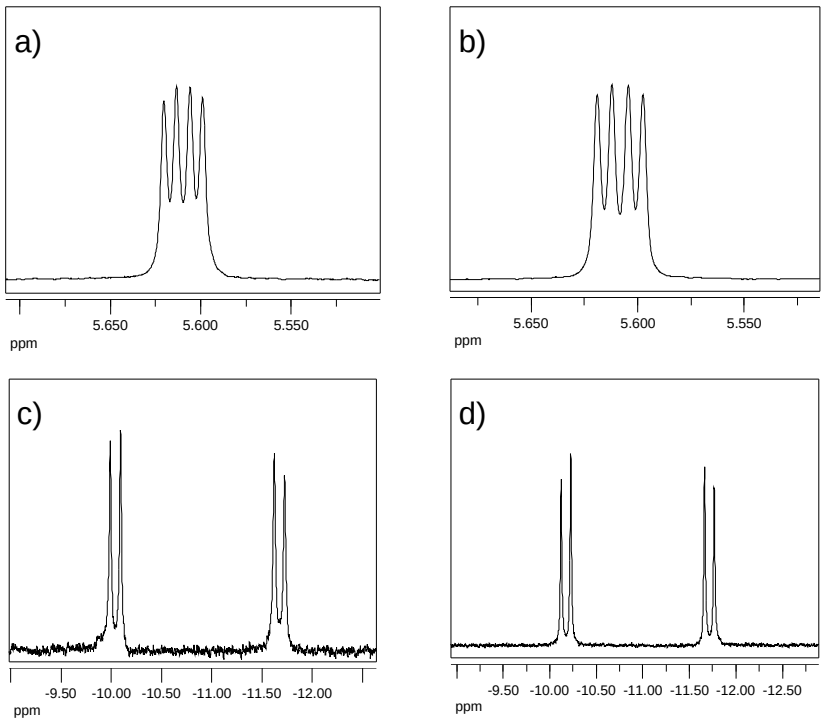


Figure 3.8. Characteristic NMR spectra of UDP- and TDP-Glc. a) ^{1}H NMR of the H-1 on UDP-Glc (**33a**) with a doublet of doublets signal and coupling constants ($J_{H\text{-}1,H\text{-}2}$ = 3.5 Hz, $J_{H\text{-}1,P}$ = 7.2 Hz) characteristic of NDP-sugars. **b)** ^{1}H NMR of the H-1 on TDP-Glc (**33b**) with a doublet of doublets signal and coupling constants ($J_{H\text{-}1,H\text{-}2}$ = 3.5 Hz, $J_{H\text{-}1,P}$ = 7.2 Hz) characteristic of NDP-sugars. **c)** ^{31}P NMR of UDP-Glc (**33a**) with doublets and identical coupling constants ($J_{P,P}$ = 20.7 Hz) characteristic of NDP-sugars. **d)** ^{31}P NMR of TDP-Glc (**33b**) with doublets and identical coupling constants ($J_{P,P}$ = 20.9 Hz) characteristic of NDP-sugars.

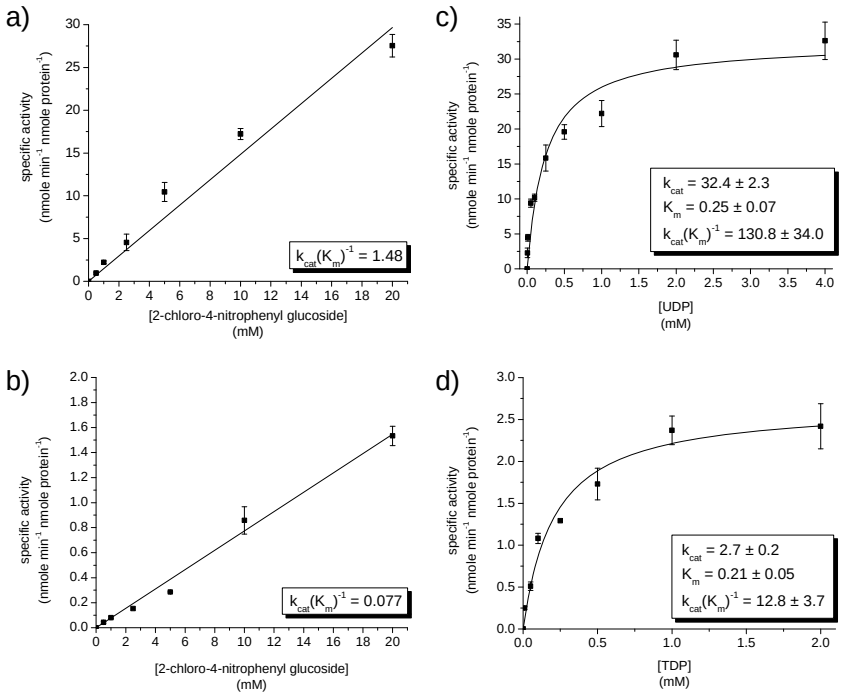


Figure 3.9. Wild-type OleD kinetic data. Graphs represent **a)** UDP saturating (2.5 mM) with **9** varied (0-20 mM), **b)** TDP saturating (2.5 mM) with **9** varied (0-20 mM), **c) 9** saturating (20 mM) with UDP varied (0-4 mM), and **d) 9** saturating (20 mM) with TDP varied (0-2 mM).

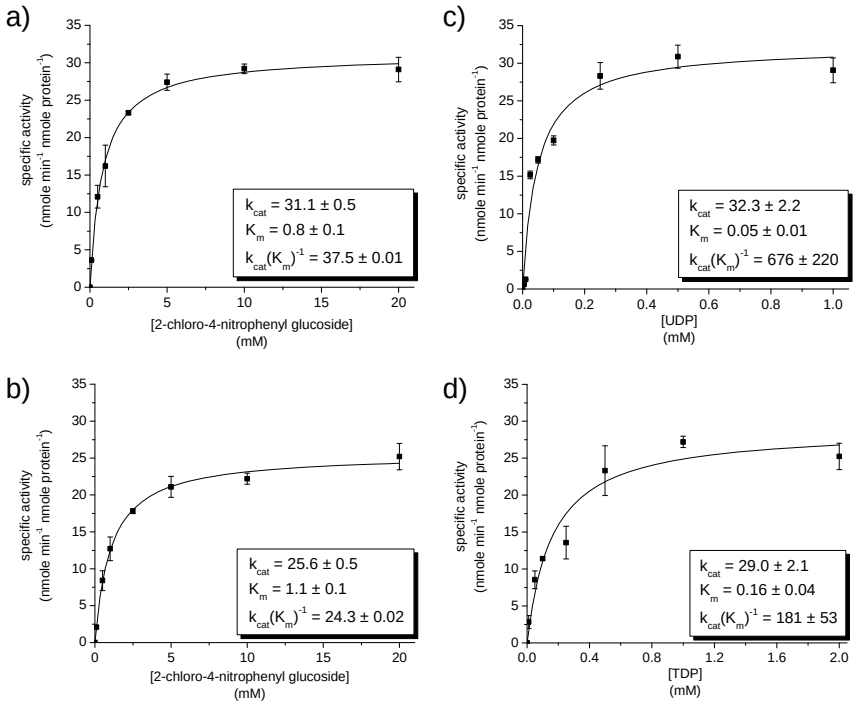


Figure 3.10. OleD variant TDP-16 kinetic data. Graphs represent **a)** UDP saturating (1.0 mM) with **9** varied (0-20 mM), **b)** TDP saturating (2.0 mM) with **9** varied (0-20 mM), **c) 9** saturating (20 mM) with UDP varied (0-1 mM), and **d) 9** saturating (20 mM) with TDP varied (0-2 mM).

Enzyme	Constant Substrate	Varied Substrate	k_{cat} (min^{-1})	K_m (mM)	k_{cat} / K_m ($mM^{-1}min^{-1}$)
WT	UDP	**9**	n.a.[1]	n.a.	1.48
	TDP	**9**	n.a.	n.a.	0.077
TDP-16	UDP	**9**	31.1 ± 0.5	0.8 ± 0.1	37.5
	TDP	**9**	25.6 ± 0.5	1.1 ± 0.1	24.3
WT	**9**	UDP	32.4 ± 2.3	0.25 ± 0.07	131
	9	TDP	2.7 ± 0.2	0.21 ± 0.05	12.8
TDP-16	**9**	UDP	32.3 ± 2.2	0.05 ± 0.01	676
	9	TDP	29.0 ± 2.1	0.16 ± 0.04	181

Table 3.1. Wild-type and TDP-16 OleD kinetic parameters. [1] n.a., not available.

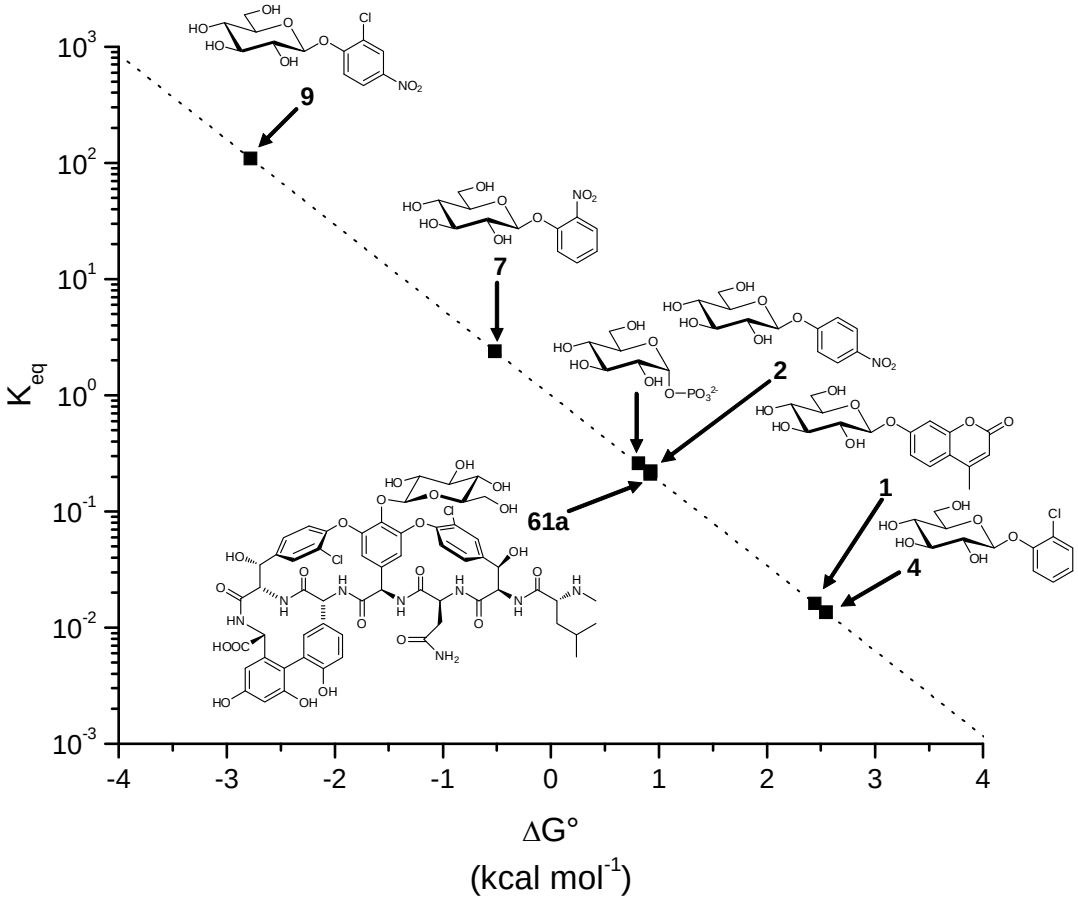


Figure 3.11. Plot of K_{eq} and their respective ΔG°. The dotted line represents the equation ΔG° = -RTln(K_{eq}). Data taken from **Table 3.2**.

Entry	Compound	$K_{eq,pH8.5}$	$\Delta G°_{pH8.5}$ (kcal mol^{-1})
4	2-chlorophenyl-β-D-glucopyranoside	1.36E-1 ± 8.50E-5	+2.55
1	4-methylumbelliferone-β-D-glucopyranoside	1.61E-2 ± 4.10E-4	+2.44
2	4-nitrophenyl-β-D-glucopyranoside	2.11E-1 ± 2.40E-4	+0.92
61a	vancomycin pseudoaglycon	2.22E-1[1]	+0.93
---	glucose-1-phosphate	2.60E-1[2]	+0.81
7	2-nitrophenyl-β-D-glucopyranoside	2.40 ± 7.40E-5	-0.52
9	2-chloro-4-nitrophenyl-β-D-glucopyranoside	1.09E+2 ± 6.00E-5	-2.78

Table 3.2. Determined K_{eq} and their respective ΔG°. K_{eq} is related to ΔG° through the equation ΔG° = -RTln(K_{eq}), where R = 1.987 cal/(mol K) and T = temperature in Kelvin (298K for this study). [1] Zhang, *et al.*(5), K_{eq} determined at pH 9.0 and 310K, uncorrected for pH. [2] Tsuboi, *et al.*(76), K_{eq} determined at pH 8.5 and 303K.

3.2.8. Screening of 2-chloro-4-nitrophenyl glycoside donors. To further assess the utility of this reaction toward novel sugar nucleotide synthesis, 15 additional 2-chloro-4-nitrophenyl glycosides (**34**-**47**) were synthesized (see section **3.4**) and evaluated for production of the corresponding sugar nucleotides in presence of UDP and TDP (**Figure 3.12a**-**b**). This set of putative donors represents a series of uniquely functionalized gluco-configured sugars as well as corresponding epimers (C-2, C-3, C-4), deoxy (C-2, C-3, C-4 and C-6) analogues and even L-sugars. The syntheses of these putative donors required 2-7 steps (35% average overall yield) and, in all cases, exclusively provided the desired anomeric stereochemistry (see section **3.4**).

Of the 15 glycoside donors evaluated (**9**, **34**-**47**) *in vitro* with TDP-16, 11 (**9**, **34**-**42**, **44**) resulted in the formation of the desired sugar nucleotide with both UDP and TDP (**Figure 3.12c, Table 3.3**, see section **3.4**). With a 1:1 molar ratio of (U/T)DP to glycoside donor in these 11 reactions, an average maximum conversion of 66% was observed (**Figure 3.13**-**3.15)**, once again highlighting the thermodynamic driving force provided by the aromatic sugar donor. For a small subset of donors (**40**, **41**, **44**), a shift of the ratio to 1:10 of UDP to glycoside donor drastically increased yields of UDP-sugars (**54a**, **55a**, **57a**) to >85%, while yields of TDP-sugars (**54b**, **55b**, **57b**) remained low (<25%) (**Figure 3.12c, Figure 3.16**). Distinct from the prior observation of extensive donor hydrolysis when attempting to use reactive *p*-nitrophenol glycoside donors for GT-catalyzed transglycosylation reactions[3], the current study revealed little or no background donor hydrolysis. While the few previously reported GT-catalyzed transglycosylation reactions using 'activated' sugar donors (*e.g.*, glycosyl halides or aromatic glycosides) required NDP or NMP for transglycosylation[3, 4], the present study is the first to establish a robust GT-based platform for the syntheses of NDP-sugars, including an array of

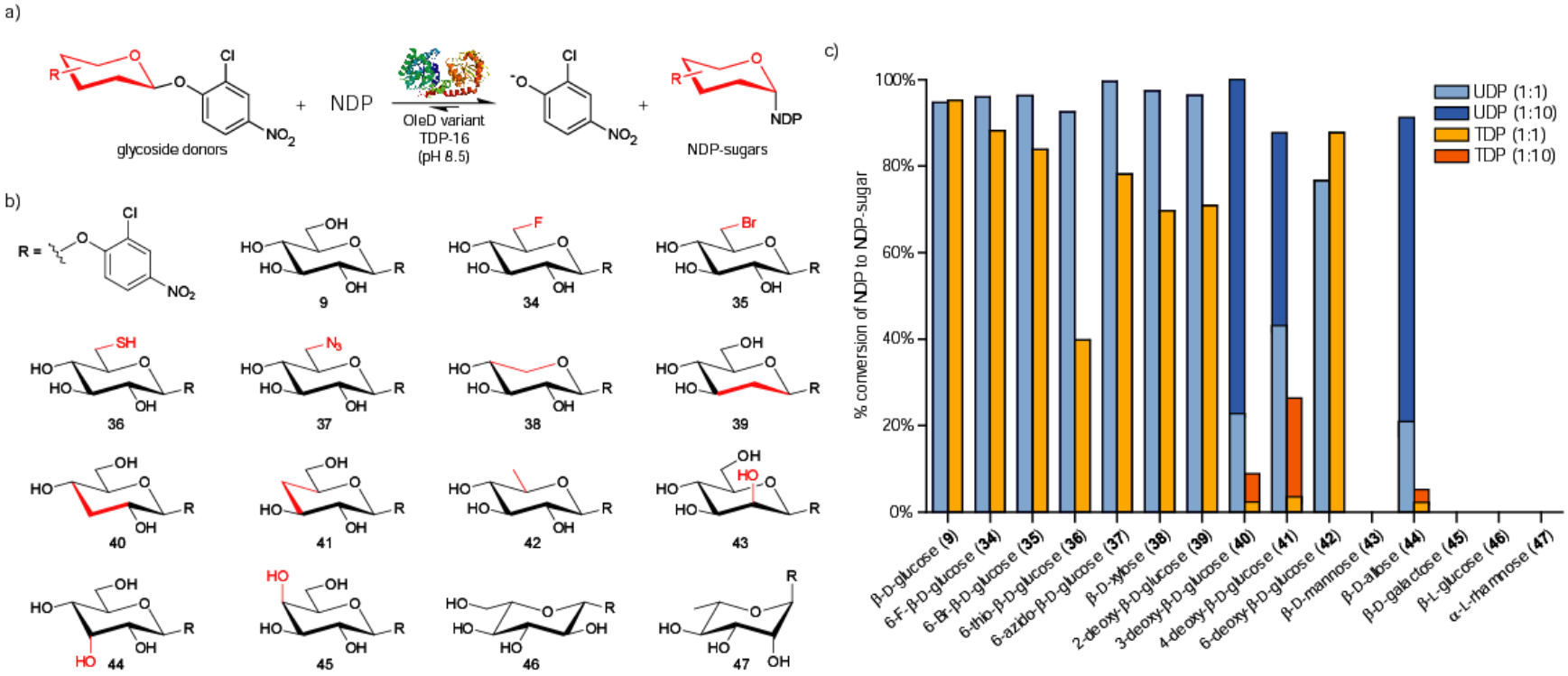


Figure 3.12. The synthesis of sugar nucleotides from 2-chloro-4-nitrophenyl glucosides. a) General reaction scheme. **b)** Structures of 2-chloro-4-nitrophenyl glycoside donors evaluated for D-sugars within this series, the differences between each member and the native OleD sugar substrate (β-D-glucose) are highlighted in red. **c)** Maximum observed percent conversion of (U/T)DP to (U/T)DP-glucose within a 21 hour time course assay for each donor. Standard reactions contained 7 μM TDP-16, 1 mM (U/T)DP, and 1 mM of 2-chloro-4-nitrophenyl glycoside donor (**9**, **34**-**47**) in Tris-HCl buffer (50 mM, pH 8.5) with a final volume of 300 μl. For reactions with UDP yielding <45% conversion under standard conditions (**40**, **41**, **43**-**47**), identical assays using 10-fold less (U/T)DP (0.1 mM) were also conducted and, where relevant, the percent conversions for the modified reactions are represented by the darker colors. HPLC chromatograms, full time course data, and product characterization are presented in **Figures 3.13-3.16** and **Table 3.3**.

Entry	Compound	Retention Time (min)	Elemental Composition $[M-2H]^{2-}$	HRMS-ESI (*m/z*) Calculated Mass $[M-2H]^{2-}$	HRMS-ESI (*m/z*) Observed Mass $[M-2H]^{2-}$
33a	UDP-α-D-glucose	8.9	$C_{15}H_{22}N_2O_{17}P_2^{2-}$	282.0202	282.0191
48a	UDP-6-deoxy-6-fluoro-α-D-glucose	9.5	$C_{15}H_{21}FN_2O_{16}P_2^{2-}$	283.0181	283.0184
49a	UDP-6-bromo-6-deoxy-α-D-glucose	11.2	$C_{15}H_{21}BrN_2O_{16}P_2^{2-}$	312.9780	312.9785
50a	UDP-6-deoxy-6-thio-α-D-glucose	10.7	$C_{15}H_{22}N_2O_{16}P_2^{2-}$	290.0081	290.0092
51a	UDP-6-azido-6-deoxy-α-D-glucose	9.4	$C_{15}H_{21}N_5O_{16}P_2^{2-}$	294.5235	294.5238
52a	UDP-α-D-xylose	9.1	$C_{14}H_{20}N_2O_{16}P_2^{2-}$	267.0150	267.0151
53a	UDP-2-deoxy-α-D-glucose	8.8	$C_{15}H_{22}N_2O_{16}P_2^{2-}$	274.0228	274.0229
54a	UDP-3-deoxy-α-D-glucose	11.0	$C_{15}H_{22}N_2O_{16}P_2^{2-}$	274.0228	274.0240
55a	UDP-4-deoxy-α-D-glucose	9.4	$C_{15}H_{22}N_2O_{16}P_2^{2-}$	274.0228	274.0234
56a	UDP-6-deoxy-α-D-glucose	9.5	$C_{15}H_{22}N_2O_{16}P_2^{2-}$	274.0228	274.0232
57a	UDP-α-D-allose	9.2	$C_{15}H_{22}N_2O_{17}P_2^{2-}$	282.0202	282.0205
33b	TDP-α-D-glucose	10.9	$C_{16}H_{24}N_2O_{16}P_2^{2-}$	281.0306	281.0302
48b	TDP-6-deoxy-6-fluoro-α-D-glucose	11.6	$C_{16}H_{23}FN_2O_{15}P_2^{2-}$	282.0284	282.0287
49b	TDP-6-bromo-6-deoxy-α-D-glucose	13.2	$C_{16}H_{23}BrN_2O_{15}P_2^{2-}$	311.9884	311.9887
50b	TDP-6-thio-6-deoxy-α-D-glucose	12.6	$C_{16}H_{24}N_2O_{15}P_2S^{2-}$	289.0192	289.0195
51b	TDP-6-azido-6-deoxy-α-D-glucose	13.0	$C_{16}H_{23}N_5O_{15}P_2^{2-}$	293.5338	293.5345
52b	TDP-α-D-xylose	11.0	$C_{15}H_{22}N_2O_{15}P_2^{2-}$	266.0253	266.0257
53b	TDP-2-deoxy-α-D-glucose	10.8	$C_{16}H_{24}N_2O_{15}P_2^{2-}$	273.0331	273.0340
54b	TDP-3-deoxy-α-D-glucose	11.5	$C_{16}H_{24}N_2O_{15}P_2^{2-}$	273.0331	273.0338
55b	TDP-4-deoxy-α-D-glucose	11.2	$C_{16}H_{24}N_2O_{15}P_2^{2-}$	273.0331	273.0336
56b	TDP-6-deoxy-α-D-glucose	11.5	$C_{16}H_{24}N_2O_{15}P_2^{2-}$	273.0331	273.0337
57b	TDP-α-D-allose	11.1	$C_{16}H_{24}N_2O_{16}P_2^{2-}$	281.0306	281.0305

Table 3.3. Characterization data for NDP-sugar products.

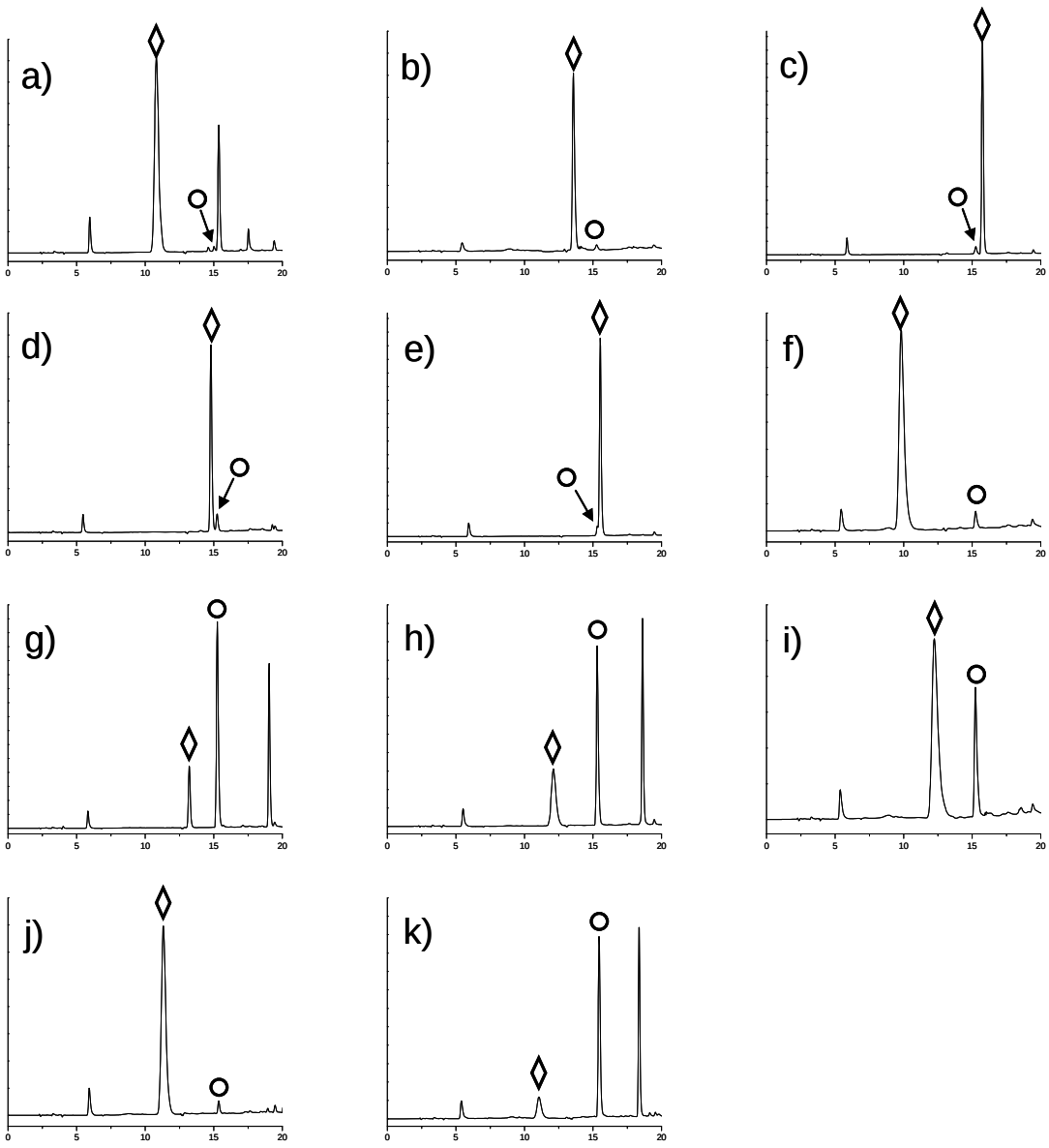


Figure 3.13. HPLC chromatograms for UDP-sugar formation. Selected chromatograms represent the maximum observed UDP-sugar formation for the corresponding sugar in a 21 hour time course experiment (see section **3.4**): **a) 33a** (UDP-glucose), **b) 48a** (UDP-6-fluoroglucose), **c) 49a** (UDP-6-bromoglucose), **d) 50a** (UDP-6-thioglucose), **e) 51a** (UDP-6-azidoglucose), **f) 52a** (UDP-xylose), **g) 53a** (UDP-2-deoxyglucose), **h) 54a** (UDP-3-deoxyglucose), **i) 55a** (UDP-4-deoxyglucose), **j) 56a** (UDP-6-deoxyglucose), **k) 57a** (UDP-allose). Open circles (○) denote UDP, open diamonds (◊) denote the respective UDP-sugar.

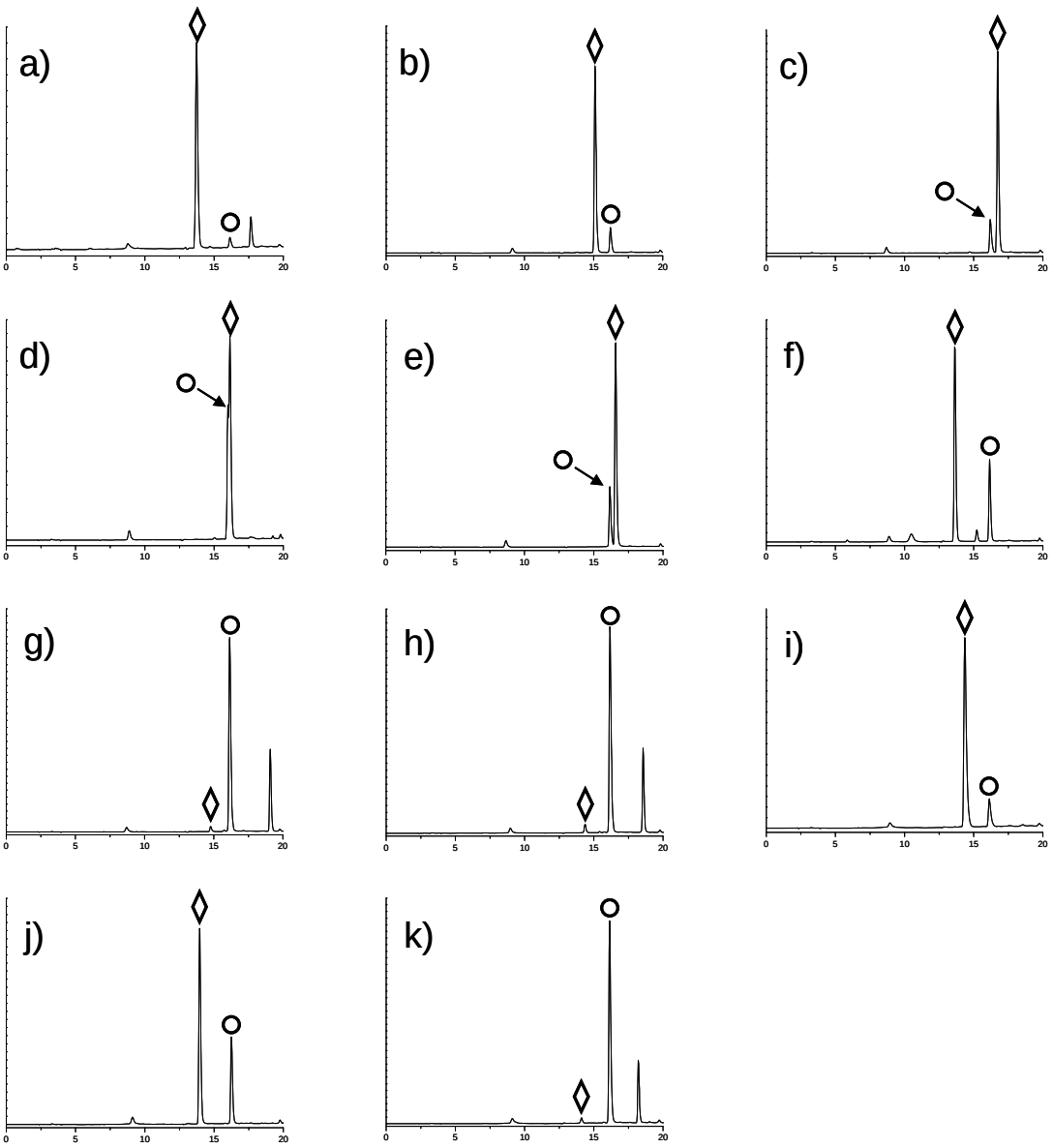


Figure 3.14. HPLC chromatograms for TDP-sugar formation. Selected chromatograms represent the maximum observed TDP-sugar formation for the corresponding sugar in a 21 hour time course experiment (see section **3.4**): **a) 33b** (TDP-glucose), **b) 48b** (TDP-6-fluoroglucose), **c) 49b** (TDP-6-bromoglucose), **d) 50b** (TDP-6-thioglucose), **e) 51b** (TDP-6-azidoglucose), **f) 52b** (TDP-xylose), **g) 53b** (TDP-2-deoxyglucose), **h) 54b** (TDP-3-deoxyglucose), **i) 55b** (TDP-4-deoxyglucose), **j) 56b** (TDP-6-deoxyglucose), **k) 57b** (TDP-allose). Open circles (○) denote TDP, open diamonds (◊) denote the respective TDP-sugar.

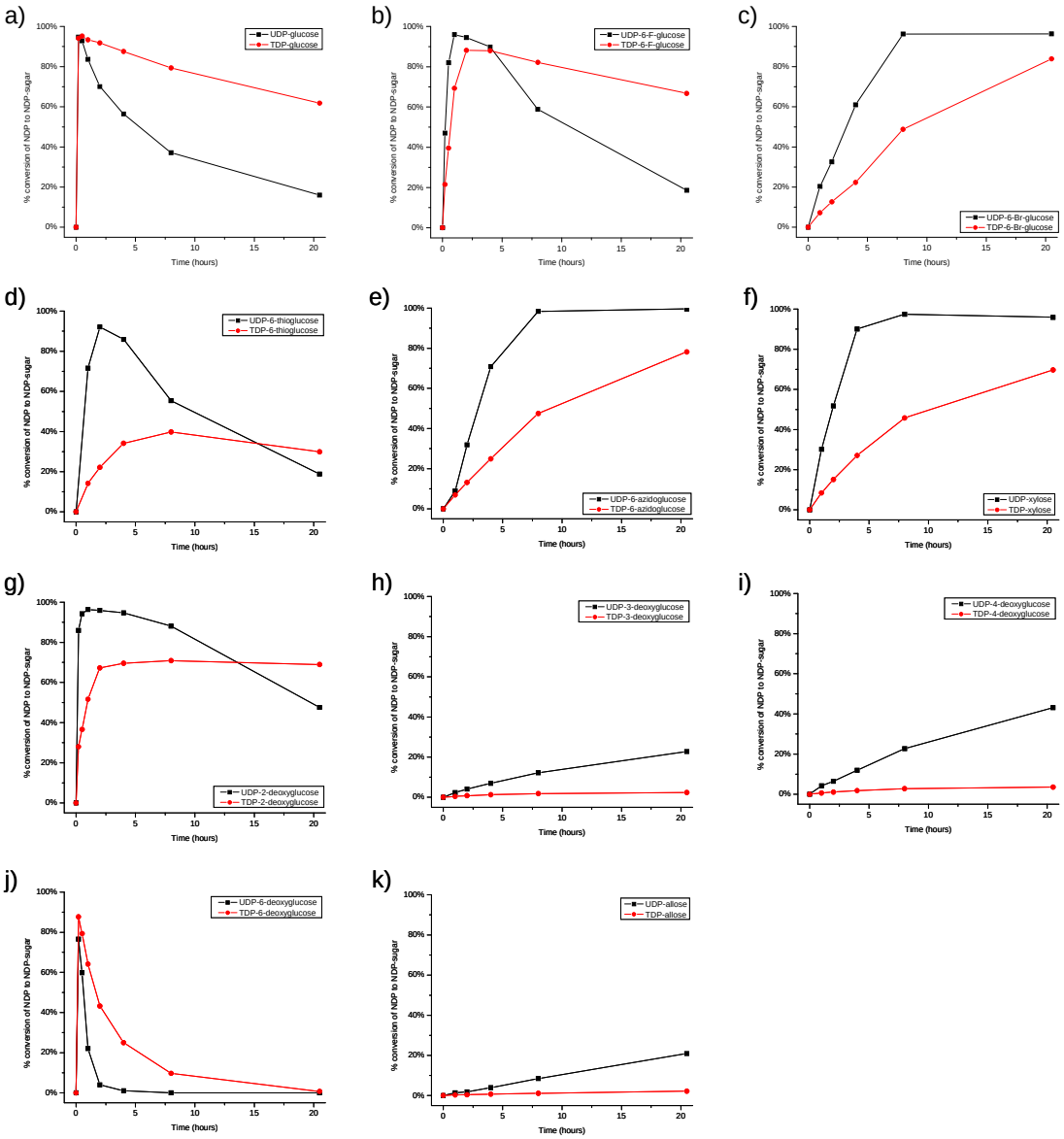


Figure 3.15. NDP-sugar formation time course (1:1). Graphs reflect a period of 21 hours for reactions containing a 1:1 ratio of NDP to 2-chloro-4-nitrophenyl glycoside (see section **3.4**) and highlight percent conversion of UDP and TDP into the respective NDP-sugars: **a) 33a-b** (NDP-glucose), **b) 48a-b** (NDP-6-fluoroglucose), **c) 49a-b** (NDP-6-bromoglucose), **d) 50a-b** (NDP-6-thioglucose), **e) 51a-b** (NDP-6-azidoglucose), **f) 52a-b** (NDP-xylose), **g) 53a-b** (NDP-2-deoxyglucose), **h) 54a-b** (NDP-3-deoxyglucose), **i) 55a-b** (NDP-4-deoxyglucose), **j) 56a-b** (NDP-6-deoxyglucose), **k) 57a-b** (NDP-allose). Multiple experiments with **9** as donor and UDP as acceptor for production of **33a** showed the error at each time point to be less than 10%. UDP-sugars are denoted by black squares (■) and TDP-sugars are denoted by red circles (●).

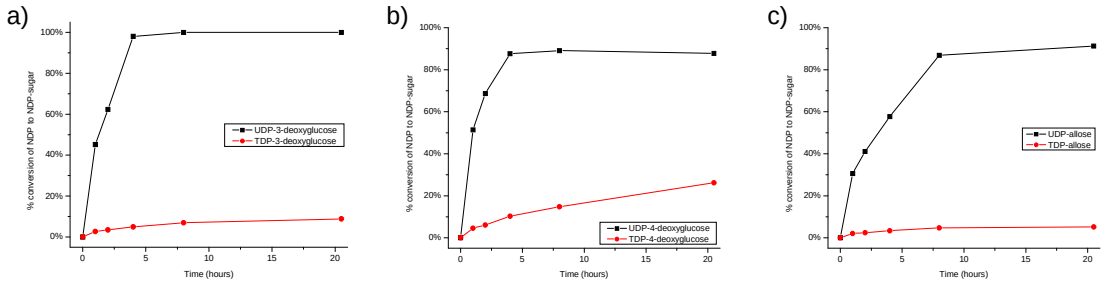


Figure 3.16. NDP-sugar formation time course (1:10). Graphs reflect a period of 21 hours for reactions containing a 1:10 ratio of NDP to 2-chloro-4-nitrophenyl glycoside (see section **3.4**) and highlight percent conversion of UDP and TDP into the respective NDP-sugars: **54a-b** (NDP-3-deoxyglucose), **b) 55a-b** (NDP-4-deoxyglucose), **c) 57a-b** (NDP-allose). Multiple experiments with **9** as donor and UDP as acceptor for production of **33a** showed the error at each time point to be less than 10%. UDP-sugars are denoted by black squares (■) and TDP-sugars are denoted by red circles (●).

uniquely-functionalized non-metabolic sugar nucleotides.

Also notable in contrast to the prior use of GT-catalyzed reactions for the synthesis of single sugar nucleotides (wherein a molar ratio of up to 1:100 sugar donor to NDP was required and, in all cases, <50% desired sugar nucleotide product was observed)[5-7, 9, 11], this study reveals a truly useful synthetic transformation (wherein a 1:1 molar ratio of sugar donor to UDP provides >70% yield desired sugar nucleotide product for 8 out of 11 examples examined). This study revealed TDP-16 to tolerate deoxygenation at C-2, C-3, C-4 or C-6; C-3 epimerization (as related to D-Glc); and an array of novel functionality at C-6. While the C-2 and C-4 epimers did not turnover in this pilot study, the production of 2-chloro-4-nitrophenolate in this reaction offers a convenient high throughput screen as a basis for rapidly expanding the sugar scope of this reaction (as subsequently discussed).

In the context of sugar nucleotide synthesis, this GT-catalyzed method offers a noteworthy alternative to both conventional chemoenzymatic approaches (requiring 1 to 11 enzymes with typical overall yields from unprotected sugars ranging from 10% to 35%)[27-32] and multi-step chemical syntheses (requiring 3-11 total steps for coupling sugar-1-phosphates and activated NMPs with typical overall yields from peracetylated sugars ranging from 9-66%)[33, 34]. As a specific comparison, the previously reported chemical synthesis of **48b** (TDP-6-deoxy-6-fluoro-α-D-glucose) from peracetylated 6-fluoro-D-glucose required 4 chemical steps including 9% yield for the final morpholidate coupling reaction,[33] while the present study reports 4 steps (3 chemical, 1 enzymatic) and 46% overall yield from the same starting material.

3.2.9. Single GT coupled reactions. The previously reported promiscuity of OleD variants in forward reactions[13-18], coupled with the newly demonstrated ability to synthesize large numbers

of NDP-sugars *in situ*, raised the question of whether TDP-16 could enable a one-pot transglycosylation wherein the sugar nucleotide formed from **9** (via the 'reverse reaction') could serve as a donor for a subsequent glycoside-forming reaction (via a simultaneous 'forward reaction') (**Figure 3.17a**). Such single and double enzyme 'aglycon exchange' reactions have been previously reported in the context of complex natural products[5-7, 9, 11], but again, the K_{eq} of such reactions has restricted their general utility.

To assess the potential of a single GT-catalyzed transglycosylation, a series of model reactions, each containing the aglycon acceptor 4-methylumbelliferone (**58**; 1 mM), one member of the 2-chloro-4-nitrophenyl glycoside donor series (**9**, **34-42**, **44**; 1mM), UDP (1 mM) and OleD variant TDP-16 (11 μM), revealed all 11 expected products (**1**, **59a-59j**) with an average yield of 45% (**Figure 3.17a, Figure 3.18 and 3.19** and **Table 3.4**). For comparison, the yield of **1** in the single GT coupled reaction was 62% (n=1), while the average yield of **1** via a standard OleD catalyzed forward reaction (using 1 equivalent of UDP-Glc donor) was 60% ± 3% (n=3). Given the established ability of OleD variants to glycosylate a wide array of structurally-diverse small molecules, drugs and natural products[13-18], the extension of this OleD-catalyzed single pot transglycosylation (or 'aglycon exchange') reaction is anticipated to offer a variety of opportunities for the glycodiversification of bioactive molecules including a number of clinically-approved drugs and complex natural products.

3.2.10. Dual GT coupled reactions. To further probe the potential of *in situ* NDP-sugars from synthetic donors to ultimately serve as donors for GTs other than OleD variants, a series of dual GT-catalyzed model reactions were performed. For this set of model reactions, GtfE was selected because of its known NDP-sugar promiscuity[35, 36] as well as the clinical potential of

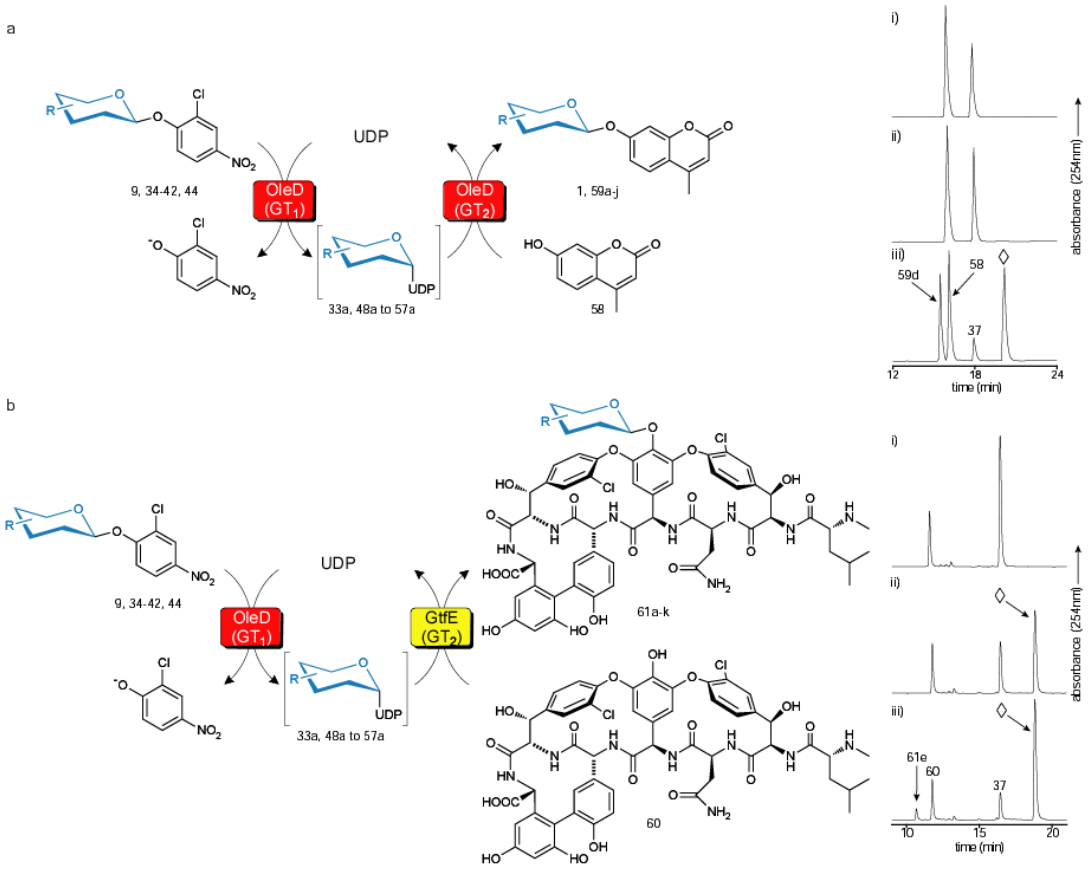


Figure 3.17. Evaluation of 2-chloro-4-nitrophenyl glycosides as sugar donors in coupled GT-catalyzed transglycosylation reactions. (**a**) The scheme for a single enzyme (TDP-16) coupled system with 4-methylumbelliferone (**58)** as the final acceptor (left) and a representative HPLC analysis (right) using the donor for 6-azido-6-deoxy-D-glucose (**37**). For the representative reaction: (i) control reaction lacking TDP-16; (ii) control reaction lacking UDP; (iii) full reaction where **37** is donor, **58** is acceptor, **59d** is desired product and ◊ represents 2-chloro-4-nitrophenolate. (**b**) The scheme for a double enzyme (TDP-16 and GtfE) coupled system with vancomycin aglycon (**60**) as the final acceptor (left) and a representative HPLC analysis (right) using the donor for 6-azido-6-deoxy-D-glucose (**37**). For the representative reaction: (i) control reaction lacking TDP-16; (ii) control reaction lacking GtfE; (iii) full reaction where **37** is donor, **60** is acceptor, **61e** is desired product and ◊ represents 2-chloro-4-nitrophenolate. See **Figures 3.18-3.21** and **Tables 3.4** and **3.5** for product structures, chromatograms, conversion rates, and mass characterization.

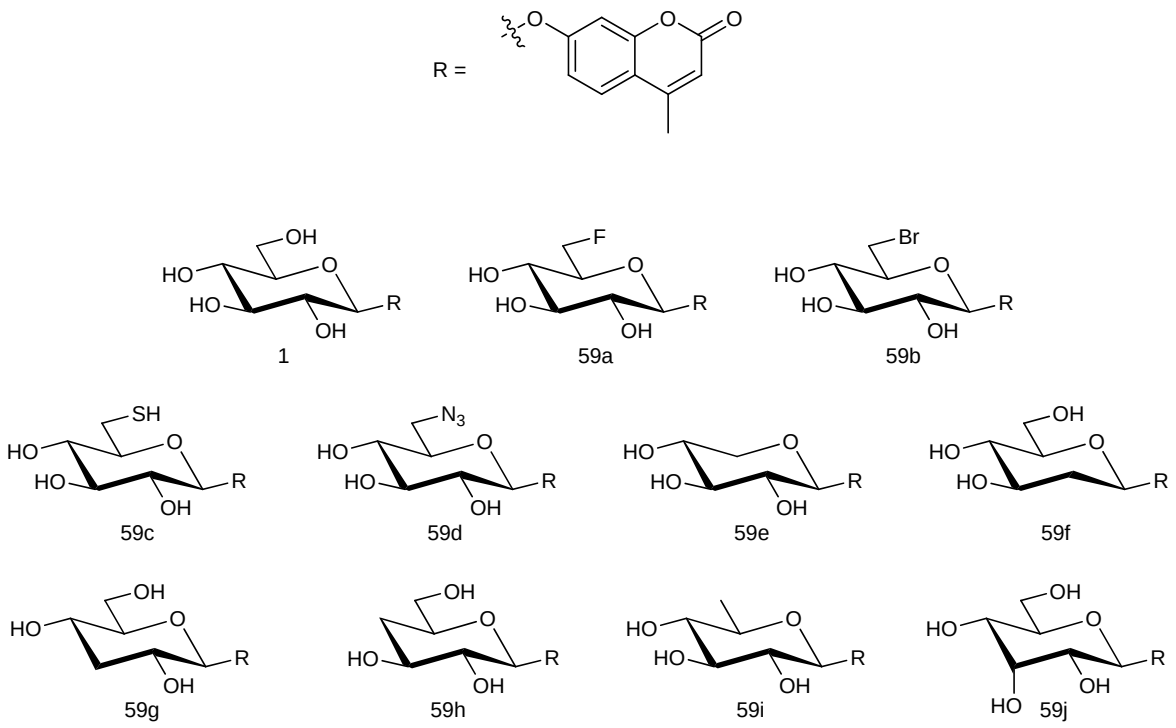


Figure 3.18. 4-methylumbelliferone glycoside products (1, 59a-j).

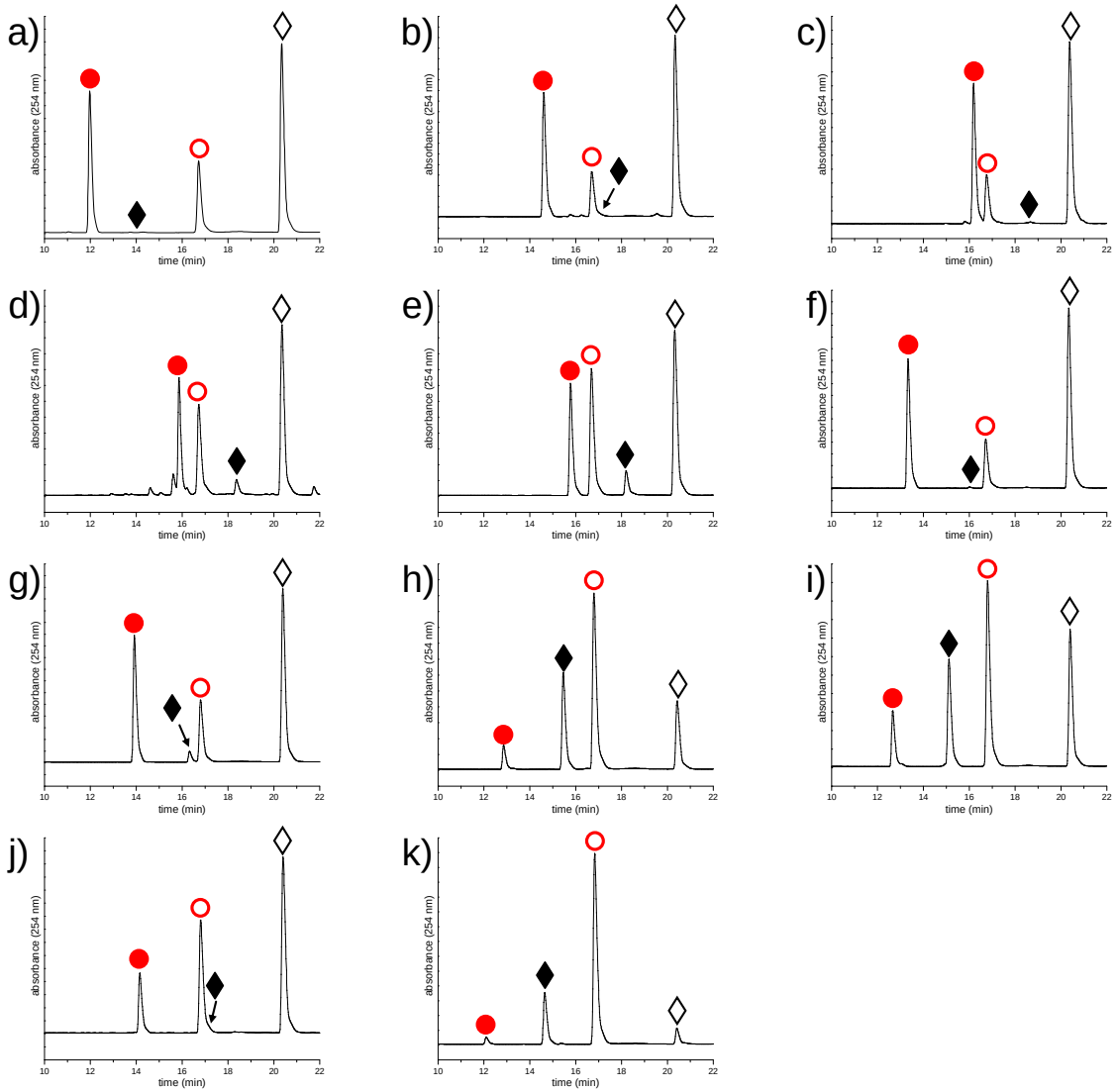


Figure 3.19. **HPLC of single GT reactions 4-Me-Umb** (**58**). Chromatograms represent the observed conversion in a 24 hour endpoint assay of **58** to the following glycosides: **a) 1** (glucoside), **b) 59a** (6-fluoroglucoside), **c) 59b** (6-bromoglucoside), **d) 59c** (6-thioglucoside), **e) 59d** (6-azidoglucoside), **f) 59e** (xyloside), **g) 59f** (2-deoxyglucoside), **h) 59g** (3-deoxyglucoside), **i) 59h** (4-deoxyglucoside), **j) 59i** (6-deoxyglucoside), **k) 59j** (alloside). Black closed diamonds (◆) denote 2-chloro-4-nitrophenyl glycosides, black open diamonds (◇) denote 2-chloro-4-nitrophenol, red open circles (○) denote **58**, and red closed circles (●) denote **1**, **59a-j** in their respective panels.

Entry	Conjugated Sugar	Percent Conversion	Retention Time (min)	Chemical Formula[1]	HRMS-ESI (*m* /z)	
					Calculated Theoretical Mass (*m* /z)	Observed Mass (*m* /z)
1	β-D-glucopyranose	62%	12.0	$C_{16}H_{18}O_8$	339.1074 $[M+H]^+$	339.1071 $[M+H]^+$
59a	6-deoxy-6-fluoro-β-D-glucopyranose	70%	14.6	$C_{16}H_{17}FO_7$	341.1031 $[M+H]^+$	341.1031 $[M+H]^+$
59b	6-bromo-6-deoxy-β-D-glucopyranose	73%	16.2	$C_{16}H_{17}BrO_7$	401.0230 $[M+H]^+$	401.0225 $[M+H]^+$
59c	6-deoxy-6-thio-β-D-glucopyranose	47%	15.9	$C_{16}H_{18}O_7S$	355.0846 $[M+H]^+$	355.0841 $[M+H]^+$
59d	6-azido-6-deoxy-β-D-glucopyranose	43%	15.8	$C_{16}H_{17}N_3O_7$	364.1139 $[M+H]^+$	364.1136 $[M+H]^+$
59e	β-D-xylopyranose	70%	13.3	$C_{15}H_{16}O_7$	309.0968 $[M+H]^+$	309.0968 $[M+H]^+$
59f	2-deoxy-β-D-glucopyranose	64%	13.9	$C_{16}H_{18}O_7$	323.1125 $[M+H]^+$	323.1125 $[M+H]^+$
59g	3-deoxy-β-D-glucopyranose	10%	12.9	$C_{16}H_{18}O_7$	323.1125 $[M+H]^+$	323.1121 $[M+H]^+$
59h	4-deoxy-β-D-glucopyranose	21%	12.7	$C_{16}H_{18}O_7$	323.1125 $[M+H]^+$	323.1121 $[M+H]^+$
59i	6-deoxy-β-D-glucopyranose	31%	14.2	$C_{16}H_{18}O_7$	323.1125 $[M+H]^+$	323.1124 $[M+H]^+$
59j	β-D-allopyranose	3%	12.0	$C_{16}H_{18}O_8$	339.1074 $[M+H]^+$	339.1085 $[M+H]^+$

Table 3.4. Data summaries for 4-Me-Umb glycosides (1, 59a-j). [1]Chemical formula before ionization.

glycodiversified vancomycin analogues[37, 38].

Typical reactions for this assessment contained a NDP-sugar generating component consisting of a 2-chloro-4-nitrophenyl glycoside donor (**9**, **34**-**42** or **44**; 1 mM), UDP (1 mM) and OleD variant TDP-16 (11 μM) coupled to a glycoside-forming component with vancomycin aglycon (**60**; 0.1 mM) and the vancomycin aglycon glucosyltransferase GtfE (11 μM). Remarkably, this series of dual GT-catalyzed transglycosylation (or 'aglycon exchange') reactions also led to the formation of all 11 expected products (**61a**-**61k**) with an average yield of 36% (**Figure 3.17b, Figure 3.20** and **3.21** and **Table 3.5**). As a comparison, the yield of **61a** in the dual-GT coupled reaction was 53% (n=1), while the average yield of **61a** via a standard GtfE catalyzed forward reaction (using 10 equivalents of native substrate UDP-Glc) was 53% ± 0.5% (n = 3). Thus, this convenient 2-chloro-4-nitrophenyl glycoside donor-driven coupled format presents synthetically useful novel sugar nucleotides *in situ* without the need for the tedious *a priori* sugar nucleotide synthesis and/or purification. Furthermore, the formation of the colorimetric product 2-chloro-4-nitrophenolate (λ_{max} = 398 nm; λ_{410} = 2.4 x 10^4 M^{-1} cm^{-1}; pH 8.5) upon GT-catalyzed glycoside formation in these single or dual GT-catalyzed coupled reactions also offers a unique opportunity for high throughput screening as described below.

3.2.11. High throughput assay for glycosyltransfer. The 2-chloro-4-nitrophenolate released during the course of GT-catalyzed NDP-sugar formation directly, or in the context of the coupled reactions formats presented above, can be followed spectrophotometrically at 410 nm in real-time (**Figure 3.22**). The ability to do so presents one of the first truly general continuous GT assays as the colorimetric read-out directly correlates to NDP-sugar usage in such reactions and thereby avoids the need for additional manipulations or specialized probes commonly associated

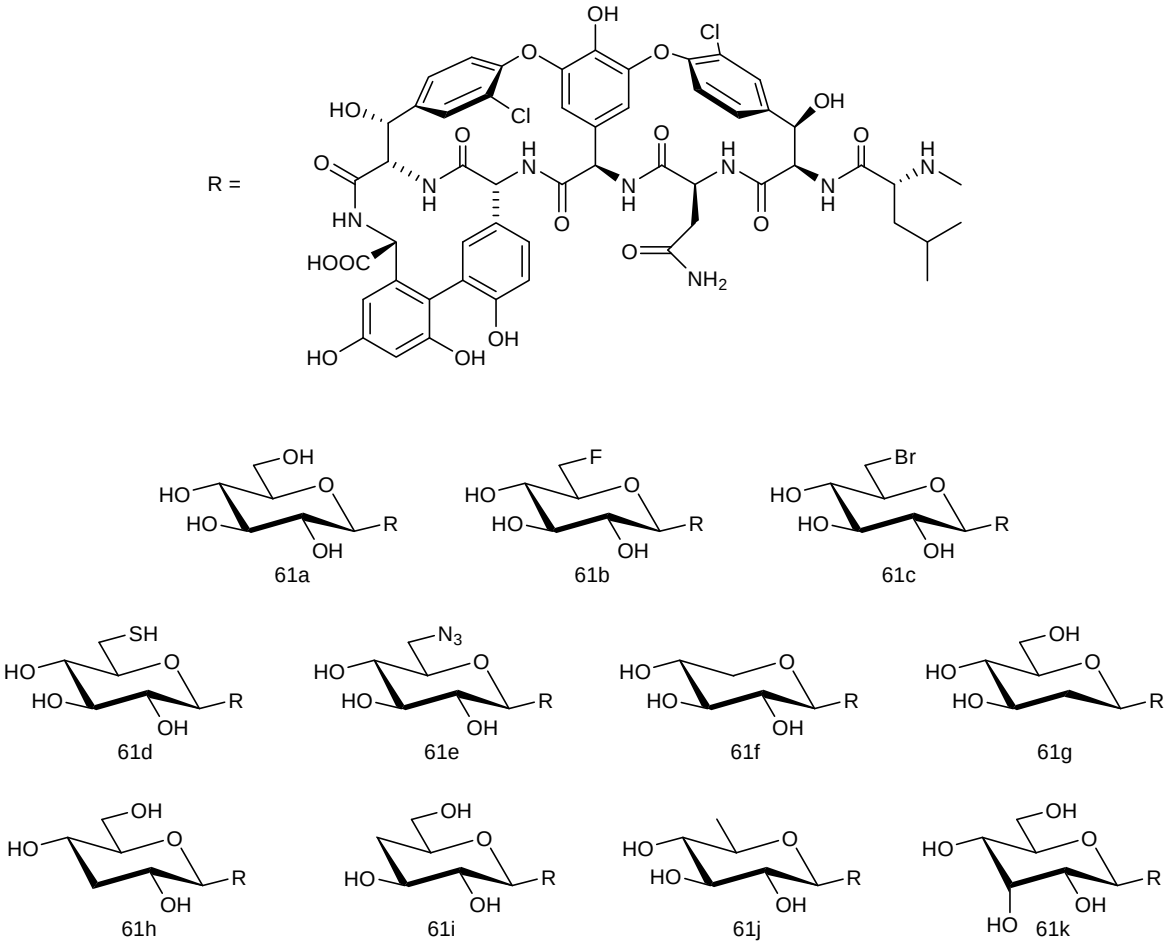


Figure 3.20. **Vancomycin glycoside products** (**61a-k**).

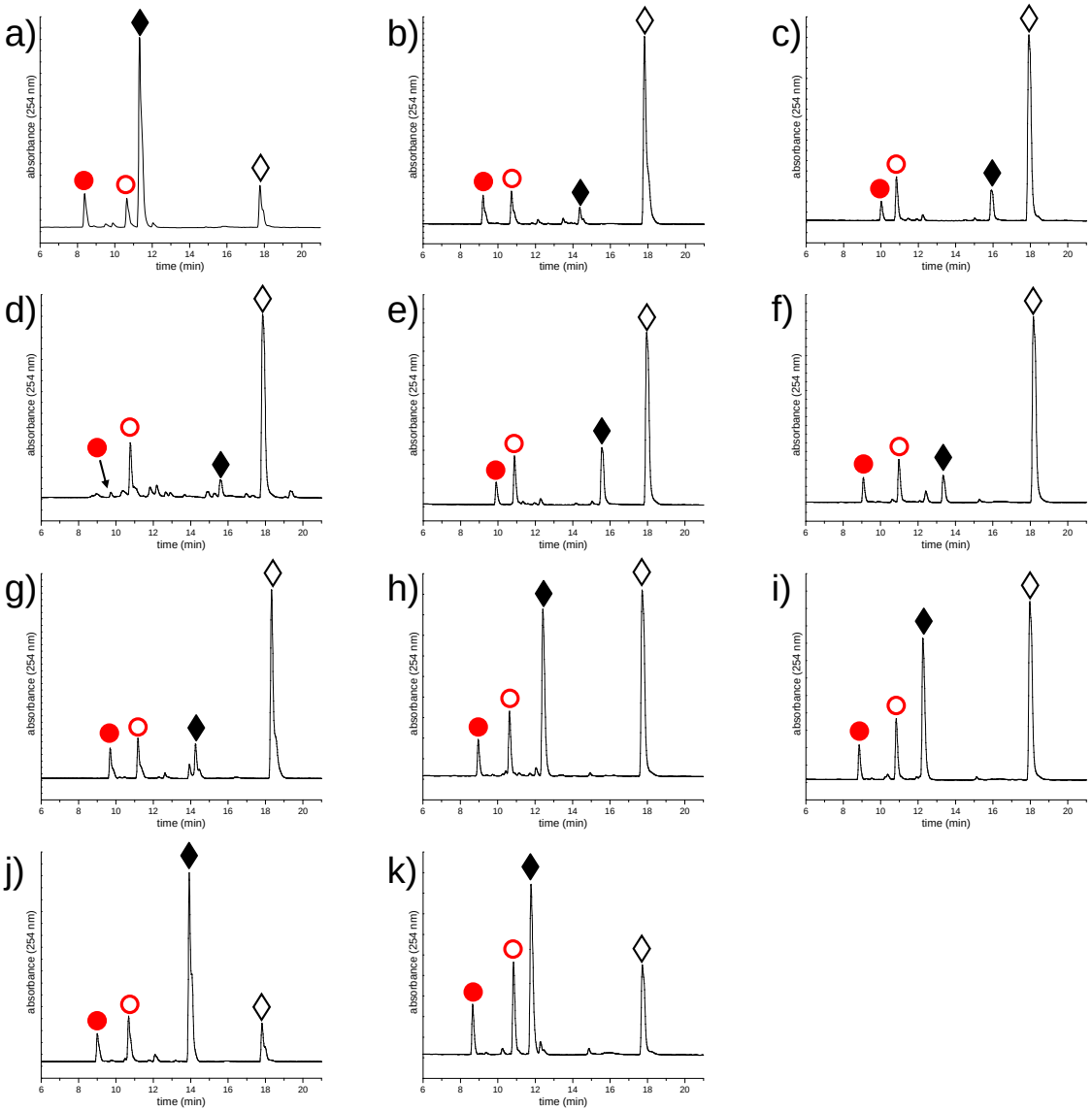


Figure 3.21. **HPLC of double GT reactions with vancomycin aglycon** (**60**). Chromatograms represent the observed conversion in a 24 hour endpoint assay of **60** to the following glycosides: **a) 61a** (glucoside), **b) 61b** (6-fluoroglucoside), **c) 61c** (6-bromoglucoside), **d) 61d** (6-thioglucoside), **e) 61e** (6-azidoglucoside), **f) 61f** (xyloside), **g) 61g** (2-deoxyglucoside), **h) 61h** (3-deoxyglucoside), **i) 61i** (4-deoxyglucoside), **j) 61j** (6-deoxyglucoside), **k) 61k** (alloside). Black closed diamonds (◆) denote 2-chloro-4-nitrophenyl glycosides, black open diamonds (◇) denote 2-chloro-4-nitrophenol, red open circles (○) denote **60**, and red closed circles (●) denote **61a-k** in their respective panels.

Entry	Conjugated Sugar	Percent Conversion	Retention Time (min)	Chemical Formula[1]	HRMS-ESI (*m*/z)	
					Calculated Theoretical Mass	Observed Mass
61a	β-D-glucopyranose	53.2%	8.4	$C_{59}H_{62}Cl_2N_8O_{22}$	1305.3428 $[M+H]^+$	1305.3408 $[M+H]^+$
61b	6-deoxy-6-fluoro-β-D-glucopyranose	45.1%	9.2	$C_{59}H_{61}BrCl_2N_8O_{21}$	1367.2584 $[M+H]^+$	1367.256 $[M+H]^+$
61c	6-bromo-6-deoxy-β-D-glucopyranose	30.0%	10.0	$C_{59}H_{61}Cl_2FN_8O_{21}$	654.1728 $[M+2H]^{2+}$	654.1710 $[M+2H]^{2+}$
61d	6-deoxy-6-thio-β-D-glucopyranose	8.6%	9.7	$C_{59}H_{62}Cl_2N_8O_{21}S$	1321.3200 $[M+H]^+$	1321.3157 $[M+H]^+$
61e	6-azido-6-deoxy-β-D-glucopyranose	32.2%	9.9	$C_{59}H_{61}Cl_2N_{11}O_{21}$	1330.3493 $[M+H]^+$	1330.3472 $[M+H]^+$
61f	β-D-xylopyranose	35.8%	9.1	$C_{58}H_{60}Cl_2N_8O_{21}$	1275.3322 $[M+H]^+$	1275.3307 $[M+H]^+$
61g	2-deoxy-β-D-glucopyranose	41.6%	9.7	$C_{59}H_{62}Cl_2N_8O_{21}$	1289.3479 $[M+H]^+$	1289.3472 $[M+H]^+$
61h	3-deoxy-β-D-glucopyranose	36.2%	9.0	$C_{59}H_{62}Cl_2N_8O_{21}$	1289.3479 $[M+H]^+$	1289.3451 $[M+H]^+$
61i	4-deoxy-β-D-glucopyranose	36.2%	8.8	$C_{59}H_{62}Cl_2N_8O_{21}$	1289.3479 $[M+H]^+$	1289.3458 $[M+H]^+$
61j	6-deoxy-β-D-glucopyranose	37.4%	9.0	$C_{59}H_{62}Cl_2N_8O_{21}$	1289.3479 $[M+H]^+$	1289.3462 $[M+H]^+$
61k	β-D-allopyranose	34.6%	8.7	$C_{59}H_{62}Cl_2N_8O_{22}$	1305.3428 $[M+H]^+$	1305.3422 $[M+H]^+$

Table 3.5. Data summaries for vancomycin glycosides (61a-k). [1]Chemical formula before ionization.

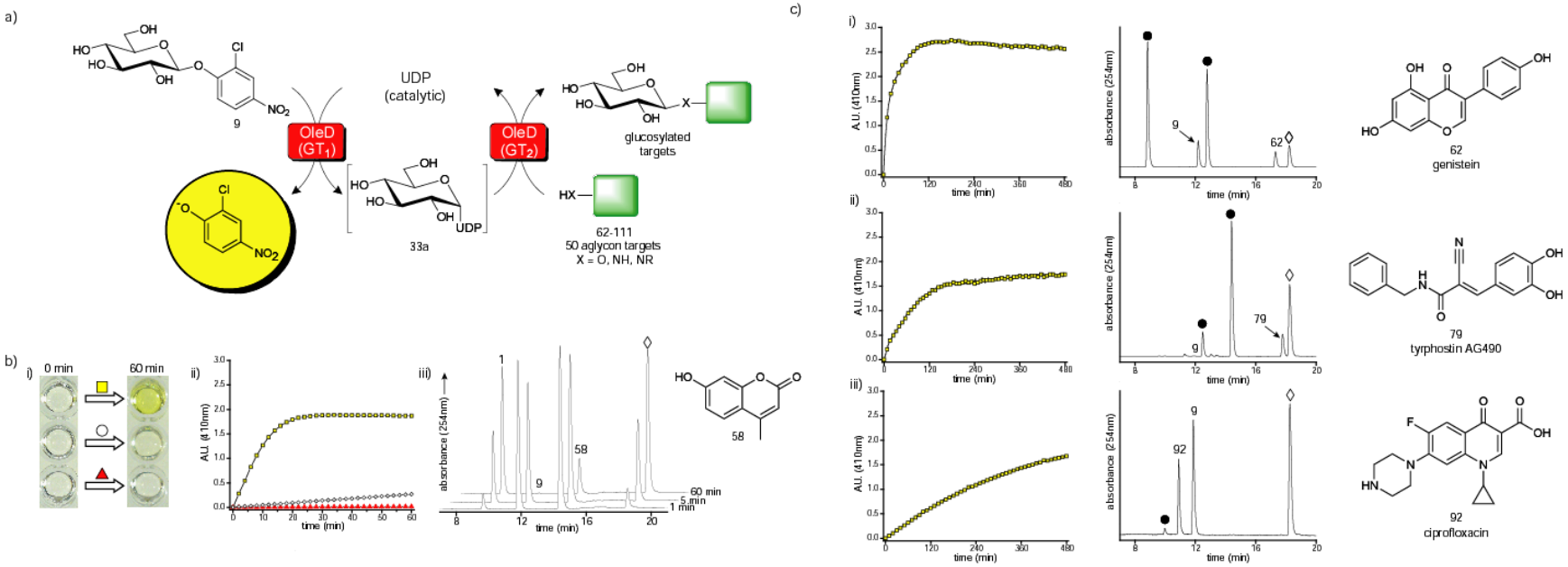


Figure 3.22. Representative data for 50 compound screen with single enzyme coupled reactions. (**a**) Scheme for colorimetric screen using the single enzyme (TDP-16) coupled format. (**b**) Evaluation of the colorimetric assay with **58** as the final acceptor. The reactions contained 0.5 mM **9** as donor, 0.5 mM **58** as acceptor, 5 µM UDP, and 11µM TDP-16 in a final total volume of 100 µl with Tris-HCl buffer (50 mM, pH 8.5) in a 96-well plate incubated at 25 °C for one hour. (**i**) Qualitative color change after one hour for the full reaction (yellow square), a control lacking the final acceptor **58** (white circle), and a control lacking UDP (red triangle). (**ii**) Δ410 nm over one hour for the full reaction (yellow squares), a control lacking the final acceptor **58** (white circles), and a control reaction lacking UDP (red triangles). (**iii**) HPLC chromatograms of full reaction at 1, 5, and 60 min where **1** is desired product, **9** is the donor, **58** is the target aglycon and ◊ represents 2-chloro-4-nitrophenolate. **(c)** The absorbance data and HPLC chromatograms of three representative hits [(**i**) **62** (genistein), (**ii**) **79** (tyrphostin), or (**iii**) **92** (ciprofloxacin)] from the broad 50 compound panel screen using the single enzyme (TDP-16) coupled format. In HPLC chromatograms **9** indicates donor; **62**, **79** or **92** represent target aglycon; ◊ indicates 2-chloro-4-nitrophenolate; and ● depicts glucosylated product(s). For overall results of the 50 compound screen see **Figure 3.23**. Additional representative absorbance plots and chromatograms, combined HPLC and LC/MS characterization, and structures of all library members are presented in **Appendix 2**.

with conventional assays for glycosidic bond formation[39-41]. To demonstrate this approach, a set of 50 (**62-111**) medicinally relevant compounds were screened with the single GT-catalyzed reaction in a high throughput format. Specifically, each 100 μl reaction in the 96-well plate contained the sugar donor **9** (0.5 mM), a putative aglycon acceptor (0.5 mM), a catalytic amount of UDP (5 μM) and TDP-16 (11 μM) and reaction progress was monitored at 410 nm over 480 min (**Figure 3.22a-b**, **Appendix 2**). Notably, the use of UDP as a limiting reagent within this coupled system reduces the potential for the various types of inhibition commonly observed in forward GT-catalyzed reactions with NDP and NDP analogues[1, 10].

Based upon this cumulative rapid analysis, 43 compounds (**62-103**) led to a positive response (designated as three standard deviations above the mean for control reactions; **Figure 3.23**), 37 of which (**62-93**, **96**, **97**, **99**, **101**, **103**) were subsequently confirmed by HPLC and/or LC/MS to lead to products consistent with glucoside formation (**Figure 3.22c, Appendix 2**). Within this group are compounds which are (or are closely related to) anti-cancer (**71**, **72**, **79**, **80**, **83**, **85**, **87**, **90**, **96**, **99**)[42-51], chemopreventive (**65-67**)[52, 53], antibacterial (**64**, **69**, **74**, **89**, **92**, **100**, **102**)[54-57], antiparasidic (**84**)[58], antiviral (**78**)[59], analgesic (**81**)[56], gout (**91**)[56], hormonal (**68**, **70**, **77**)[56, 60], immunosuppressive (**73**)[61], neurotransmitter (**86**)[62], antineurodegenerative (**75**)[63], cirrhosis (**76**)[64], conjestive heart failure (**88**)[56], and antiemetic (**97**)[56] agents. With the large variety of scaffolds identified and their cooresponding disease applications, these results clearly establish this colorimetric assay approach as an invaluable tool in the identification of new glycoconjugates for drug discovery.

This study highlights a high throughput assay to identify novel acceptors that can be glycosylated by a given GT and greatly expands upon the use of simple ‘activated’ glycosides as sugar donors in coupled reactions. Additionally, the demonstrated ability to couple this assay to

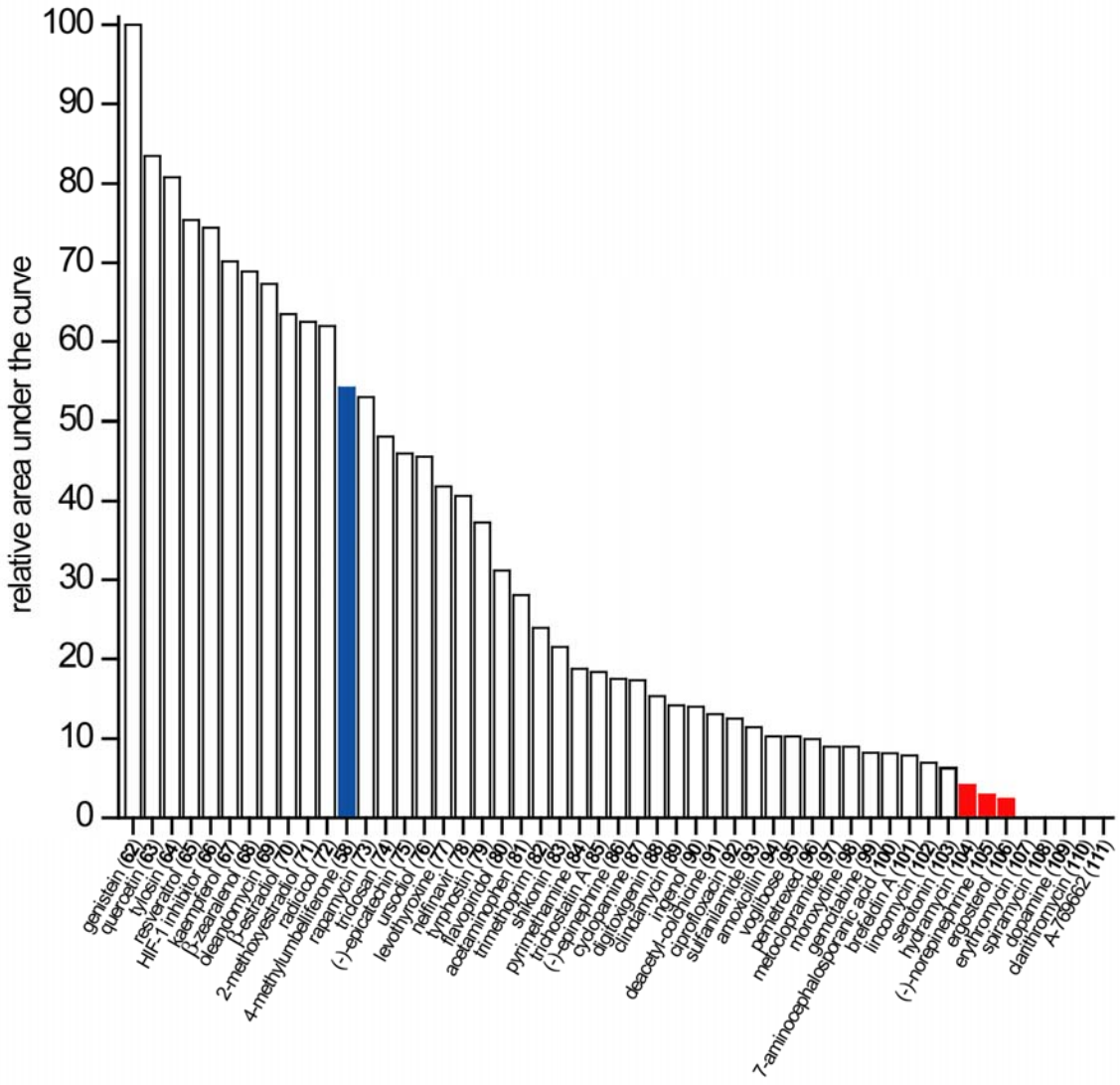


Figure 3.23. **Drug glycosylation screen ranking.** Compounds were ranked relative to the compound with the highest area under the curve (set at an arbitrary value of 100). The positive control, **58** (4-methylumbellilferone), is highlighted in blue, and compounds not identified as 'hits' are highlighted in red. Standard deviation of area under the curve for both the positive (with **58** as final acceptor) and negative controls (lacking final acceptor) was 3%.

essentially any downstream sugar-utilizing enzyme/process is also anticipated to have a broad range of fundamental applications including screens for GT inhibitors, GT engineering/evolution (toward utilization of novel NDPs, glycoside donors, and/or acceptors)[39, 40, 65-67] and/or engineering/evolution/investigations of additional NDP-sugar utilizing enzymes[19-26, 68, 69].

3.3. Conclusions

In summary, this study directly challenges the general notion that NDP-sugars are 'high-energy' sugar donors when taken out of their traditional biological context, revealing the equilibria of GT-catalyzed reactions to be highly substrate-dependent and adaptable. The flexibility of the GT thermodynamic landscape, in turn, enabled general NDP-sugar syntheses, *in situ* formation of NDP-sugars to drive coupled LeLoir GT-catalyzed reactions for glycoconjugate formation, and the first general high throughput colorimetric assay for glycosyltransfer. Given the power of the screen presented, these preliminary data suggest both the ability to enable the rapid optimization (via directed evolution) of new OleD prodigy for nearly any desired NDP/sugar pair as well as the ability to couple this screen to nearly any downstream sugar-utilizing for engineering/evolution or biochemical analysis. While substrates providing the greatest thermodynamic advantage may not always provide an equivalent kinetic advantage, this study highlights the merit of optimizing enzyme-catalyzed reactions based upon thermodynamic constraints. As such, future attempts to exploit and/or engineer other novel enzyme-catalyzed reactions may benefit from similar considerations.

3.4. Materials and Methods

3.4.1. General materials and methods.

3.4.1.1. General materials. Unless otherwise stated, all general chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) or New England Biolabs (Ipswich, MA, USA). Compounds **1**, **2**, **7**, **10**, **58**, **62**, **64**, **65**, **68-70**, **77**, **78**, **81-86**, **88**, **89**, **92-95**, **97**, **98**, **100-102** and **105-110** were purchased from Sigma-Aldrich (St. Louis, MO, USA). Compounds **63** and **67** were obtained from Indofine Chemicals (Hillsborough, NJ, USA). Compounds **65**, **66**, **71**, **72**, **74** and **83** were obtained from EMD Chemicals (Darmstadt Germany). Compounds **75** and **103** were obtained from Fisher Scientific (Pittsburgh, PA, USA). Compound **76** was obtained from MP Biochemicals (Solon, OH, USA). Compounds **73**, **79**, **90**, **96**, **99** and **111** were obtained from LC Laboratories (Woburn, MA, USA). Compound **80** was obtained from Selleck Chemicals (Houston, TX, USA). Compound **87** was obtained from Toronto Research Chemicals (Toronto, ON, Canada). Compound **91** was previously synthesized[70]. Compound **104** was isolated from fermentation. All commercially obtained compounds were utilized as received.

3.4.1.2 General methods. High resolution mass spectra were acquired on a Bruker MaXis ultra-high resolution quadrupole time of flight mass spectrometer by negative ionization electrospray with a source potential of 2800 V, drying gas at 200 °C flowing at 4 L/min and a nebulizing gas pressure of 0.4 bar. Samples were infused at 3 μL/min and spectra collected for 2 min. Routine TLC analyses were performed on aluminum TLC plates coated with 0.2 mm silica gel (from Sigma-Aldrich, St. Louis, MO, USA) and monitored at 254 nm. Flash column chromatography

was achieved on 40 – 63 μm, 60 Å silica gel (from Silicycle, Quebec, Canada).

Unless otherwise noted, analytical reverse-phase HPLC was conducted with a Gemini-NX C-18 (5 μm, 250 x 4.6 mm) column (from Phenomenex, Torrance, California, USA) with a gradient of 10% B to 75% B over 20 min, 75% B to 95% B over 1 min, 95% B for 5 min, 95% B to 10% B over 3 min, 10% B for 6 min (A = dH_2O with 0.1% TFA; B = acetonitrile; flow rate = 1 mL min^{-1}) and detection monitored at 254 nm. Regardless of method, HPLC peak areas were integrated with Star Chromatography Workstation Software (from Varian, Palo Alto, CA, USA) and the percent conversion calculated as a percent of the total peak area.

NMR spectra were obtained using a ^{UNITY}INOVA 400 MHz instrument (from Varian, Palo Alto, CA, USA) in conjunction with a QN Switchable BB probe (from Varian) or ^{UNITY}INOVA 500 MHz instrument in conjunction with a qn6121 probe (from Nalorac, Martinez, CA, USA). ^{1}H and ^{13}C chemical shifts were referenced to internal solvent resonances. ^{31}P chemical shifts were not referenced. Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), qn (quintet), m (multiplet) and br (broad). Italicized elements or groups are those that are responsible for the shifts. Chemical shifts are reported in parts per million (ppm) and coupling constants (*J*) are given in Hz. NMR assignments were performed with the aid of gCOSY and gHSQC experiments.

3.4.2. Protein production and purification. A single protocol based upon previously published methods for OleD[16] and GtfE[5] was utilized for all purifications. Specifically, single isolates of *E. coli* BL21(DE3)pLysS (Stratagene, La Jolla, CA, USA) transformed with pET28a/*oleD*[13], pET28a/*oleD*[P67T/S132F/A242V] (produces OleD variant ASP)[13], pET28a/*oleD*[P67T/S132F/A242L] (produces OleD variant 3-1H12)[15],

pET28a/*oleD*[P67T/S132F/A242L/Q268V] (produces OleD variant TDP-16)[17], or pET22b/*gtfE* vector[5] were utilized for protein production and purification. Briefly, single colonies were used to inoculate 5 mL starter cultures with 50 μg mL^{-1} kanamycin (for pET28a) or 100 μg mL^{-1} ampicillin (for pET22b) and incubated overnight at 37 °C and 250 rpm. 4 mL of saturated starter culture was transferred to 1 L cultures of Luria-Bertani medium supplemented with 50 μg mL^{-1} kanamycin (for pET28a) or 100 μg mL^{-1} ampicillin (for pET22b) and grown at 37 °C until the OD_{600} reached ~0.7. Isopropyl β-D-thiogalactoside (0.4 mM final concentration) was added and cultures were incubated at 28ºC for approximately 18 hours at 250 rpm. Cell pellets were collected by centrifugation (6,000 *g* at 4 °C for 20 min), resuspended in 10 mL of chilled lysis buffer (20 mM phosphate buffer, pH 7.4, 0.5 M NaCl, 10 mM imidazole), and lysed by sonication (5 pulses of 30 seconds each) in an ice bath. Cell debris was removed by centrifugation (10,000 *g* at 4 °C for 20 min) and the cleared supernatant was incubated with alkaline phosphatase (4 U ml^{-1}; from Roche, Basel, Switzerland) on ice for 2.5 hours with occasional agitation to degrade contaminating nucleotide diphosphates.

Following, the supernatant was applied to 2 mL of nickel nitrilotriacetic acid resin (from QIAgen Valencia, CA, USA) pre-equilibrated with wash buffer (20 mM phosphate buffer, pH 7.4, 0.3 M NaCl, 10 mM imidazole). Protein was allowed to bind for 30 min at 4 °C and the resin washed with 4 x 50 mL wash buffer. Finally, the enzyme was eluted with 2 mL of chilled wash buffer containing an additional 240 mM imidazole for 10 min at 4 °C. Purified protein was applied to a PD-10 desalting column (from Amersham Biosciences, Piscataway, NJ, USA), equilibrated with 50 mM Tris-HCl (pH 8.0), and eluted as described by the manufacturer to typically provide 2 mL of desired protein at typical concentrations ranging from 5-12 mg/mL. Final purified proteins were flash frozen drop-wise in liquid nitrogen and stored at -80 °C.

Protein purity was confirmed by SDS-PAGE to be >95% and protein concentration was determined using the Bradford Protein Assay Kit (from Bio-Rad, Hercules, CA, USA). Small aliquots of protein were thawed for experiments as required and did not undergo multiple freeze/thaw cycles.

3.4.3. Syntheses of β-D-glucosides (1-32).

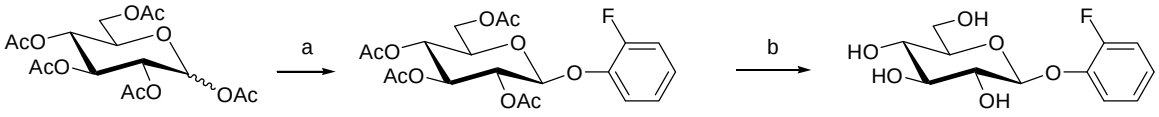
3.4.3.1. Substituted *O*-phenyl-β-D-glucosides (3-6, 8, 11-13).

3.4.3.1.1. General glycosylation procedure. According to a procedure from Lee, *et al.*[71], penta-*O*-acetyl-β-D-glucose (1 equiv.) and substituted phenol (2 equiv.) were added to a round bottom flask flushed with argon. Triethylamine (1 equiv.) in 9 mL of anhydrous CH_2Cl_2 was added. Boron trifluoride diethyl etherate (5 equiv.) in 1 mL of anhydrous CH_2Cl_2 was added dropwise to the reaction over 30 minutes. The mixture was kept under argon and allowed to proceed at room temperature. After the reaction was determined to be complete by TLC, an equal volume of saturated aqueous $NaHCO_3$ was added and the reaction was stirred until the evolution of gas halted. The organic layer was recovered and the aqueous layer was extracted 2X with an equal volume of CH_2Cl_2. The combined organic layers were dried over sodium sulfate and concentrated with reduced pressure. The peracetylated intermediate was purified by flash chromatography with EtOAc/Hexanes (1:2).

3.4.3.1.2. General deprotection procedure. The purified peracetylated glucoside intermediate was dissolved in MeOH (20 mL $mmol^{-1}$), a catalytic amount of sodium methoxide powder was

added, and the reaction was allowed to proceed overnight with stirring at room temperature. Neutralization was then performed by adding Amberlite IR-120 (H^+ form) ion-exchange resin (from Sigma Aldrich, St. Louis, Missouri, USA). The resin was filtered off using a small column of Celite 545 (from Fisher Scientific, Pittsburgh, Pennsylvania, USA), and then concentrated with reduced pressure to yield the final product without further purification.

3.4.3.1.3. Synthesis of 2-fluorophenyl-β-D-glucopyranoside (3).



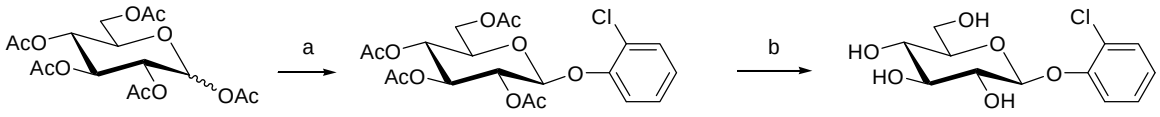
a) 2-fluorophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 36h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(2-fluorophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (112). General procedure **3.4.3.1.1** with 2-fluorophenol (0.22 g, 2.0 mmol) yielded **112** (0.37 g, 83% yield) as a white solid in a 36 hour reaction. TLC R_f = 0.15 (EtOAc/Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^-$ calcd for $C_{20}H_{22}FO_{10}$, 441.1; found 441.1.

2-fluorophenyl-β-D-glucopyranoside (3). General procedure **3.4.3.1.2** with **112** (0.37 g, 0.83 mmol) yielded **3** (0.21 g, 93% yield) as white crystals. TLC R_f = 0.22 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 7.27 (dt, $J_{Ph,Ph}$ = 1.6 Hz, $J_{Ph,Ph}$ = 8.3 Hz, 1 H, Ph), 7.14-7.05 (m, 2 H, Ph), 7.02-6.95 (m, 1 H, Ph), 4.96 (d, $J_{H1,H2}$ = 7.6 Hz, 1 H, H-1), 3.87 (dd, $J_{H5,H6a}$ = 1.8 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.2 Hz, 1 H, H-6b), 3.53-3.37 (m, 4 H, H-2, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 155.0 (d, $^1J_{C,F}$ = 244.0 Hz),146.6 (d, $^2J_{C,F}$ = 10.3

Hz), 125.6 (d, $^{3}J_{C,F}$ = 4.0 Hz), 123.9 (d, $^{3}J_{C,F}$ = 7.4 Hz), 119.3, 117.2 (d, $^{2}J_{C,F}$ = 19.2 Hz), 102.7, 78.3, 78.0, 74.9, 71.3, 62.5; HRMS (*m/z*): $[M+Na]^{+}$ calcd for $C_{12}H_{15}FNaO_6$ 297.0745, found 297.0751.

3.4.3.1.4. Synthesis of 2-chlorophenyl-β-D-glucopyranoside (4).

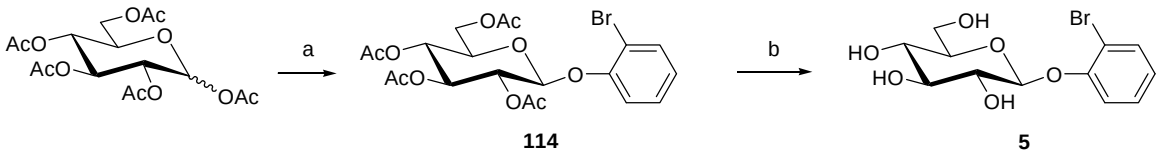


a) 2-chlorophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 36h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(2-chlorophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (113). General procedure **3.4.3.1.1** with 2-chlorophenol (0.26 g, 2.0 mmol) yielded **113** (0.29 g, 62% yield) as a white powder in a 36 hour reaction. TLC R_f = 0.13 (EtOAc/Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^{-}$ calcd for $C_{20}H_{22}ClO_{10}$, 457.1; found 457.2.

2-chlorophenyl-β-D-glucopyranoside (4). General procedure **3.4.3.1.2** with **113** (0.29 g, 0.62 mmol) yielded **4** (0.18 g, 99% yield) as white crystals. TLC R_f = 0.20 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 7.35 (dd, $J_{Ph,Ph}$ = 1.5 Hz, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.29-7.21 (m, 2 H, Ph), 7.01-6.95 (m, 1 H, Ph), 4.99 (d, $J_{H1,H2}$ = 7.5 Hz, 1 H, H-1), 3.88 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.2 Hz, 1 H, H-6b), 3.56-3.50 (m, 1 H, H-2), 3.50-3.40 (m, 3 H, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 153.2, 130.0, 127.8, 123.2, 122.8, 116.7, 101.0, 77.1, 76.9, 73.6, 70.0, 61.2; HRMS (*m/z*): $[M+Na]^{+}$ calcd for $C_{12}H_{15}ClNaO_6$ 313.0450, found 313.0456.

3.4.3.1.5. Synthesis of 2-bromophenyl-β-D-glucopyranoside (5).

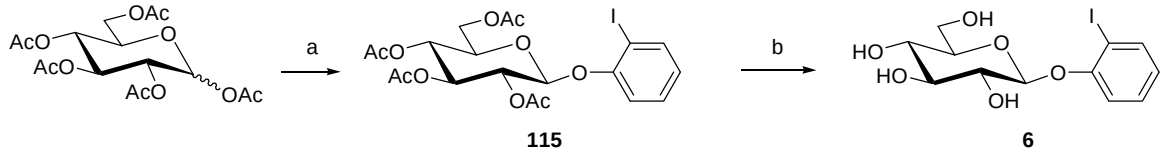


a) 2-bromophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 48h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(2-bromophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (114). General procedure **3.4.3.1.1** with 2-bromophenol (0.35 g, 2.0 mmol) yielded **114** (0.24 g, 47% yield) as a white powder in a 48 hour reaction. TLC R_f = 0.10 (EtOAc:Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^-$ calcd for $C_{20}H_{22}BrO_{10}$, 501.0; found 501.1.

2-bromophenyl-β-D-glucopyranoside (5). General procedure **3.4.3.1.2** with **114** (0.24 g, 0.47 mmol) yielded **5** (0.16 g, 99% yield) as white crystals. TLC R_f = 0.25 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 7.52 (dd, $J_{Ph,Ph}$ = 1.5 Hz, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.31-7.20 (m, 2 H, Ph), 6.94-6.87 (m, 1 H, Ph), 4.99 (d, $J_{H1,H2}$ = 7.6 Hz, 0.93 H, H-1), 3.87 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.69 (dd, $J_{H5,H6b}$ = 5.1 Hz, 1 H, H-6b), 3.53 (dd, $J_{H2,H3}$ = 8.8 Hz, 1 H, H-2), 3.49-3.35 (m, 3 H, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 154.2, 133.1, 128.5, 123.2, 116.5, 112.2, 101.0, 77.1, 76.9, 73.6, 70.0, 61.2; HRMS (*m/z*): $[M+NH_4]^+$ calcd for $C_{12}H_{19}BrNO_6$ 352.0391, found 352.0386.

3.4.3.1.6. Synthesis of 2-iodophenyl-β-D-glucopyranoside (6).

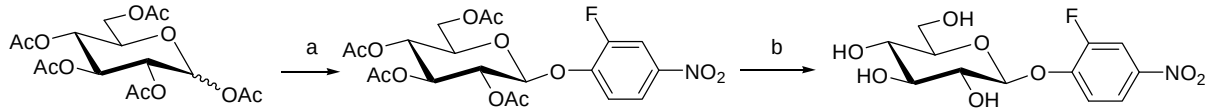


a) 2-iodophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 36h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(2-iodophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (115). General procedure **3.4.3.1.1** with 2-iodophenol (0.30 g, 2.0 mmol) yielded **115** (0.27 g, 53% yield) as a white powder in a 36 hour reaction. TLC R_f = 0.15 (EtOAc/Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^-$ calcd for $C_{20}H_{22}IO_{10}$, 549.0; found 549.0.

2-iodophenyl-β-D-glucopyranoside (6). General procedure **3.4.3.1.2** with **115** (0.27 g, 0.49 mmol) yielded **6** (0.15 g, 84% yield) as white crystals; less than 10% of α-anomer was observed during characterization. TLC R_f = 0.27 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.76 (dd, $J_{Ph,Ph}$ = 1.6 Hz, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.33-7.28 (m, 1 H, Ph), 7.16 (dd, $J_{Ph,Ph}$ = 1.6 Hz, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 6.78-6.74 (m, 1 H, Ph), 5.00 (d, $J_{H\text{-}1b,H\text{-}2}$ = 7.6 Hz, 0.9 H, H-1), 3.88 (dd, $J_{H\text{-}6a,H\text{-}5}$ = 2.0 Hz, $J_{H\text{-}6a,H\text{-}6b}$ = 12.1 Hz, 1 H, H-6a), 3.69 (dd, $J_{H\text{-}6b,H\text{-}5}$ = 5.4 Hz, $J_{H\text{-}6b,H\text{-}6a}$ = 12.1 Hz, 1 H, H-6b), 3.55 (m, 1 H, H-2), 3.49-3.37 (m, 3 H, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3O) δ 156.6, 139.4, 129.4, 123.7, 115.3, 101.1, 86.0, 77.1, 77.0, 73.7, 70.0, 61.3; HRMS (*m/z*): $[M+NH_4]^+$ calcd for $C_{12}H_{19}INO_6$ 400.0252, found 400.0246.

3.4.3.1.7. Synthesis of 2-fluoro-4-nitrophenyl-β-D-glucopyranoside (8).



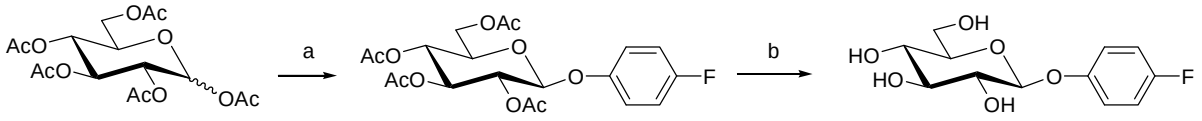
116 **8**

a) 2-fluoro-4-nitrophenol, $BF_3 \cdot OEt_2$, Et_3N in CH_2Cl_2, RT, 36h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(2-fluoro-4-nitrophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (116). General procedure **3.4.3.1.1** with 2-fluoro-4-nitrophenol (0.31 g, 2.0 mmol) yielded **116** (0.35 g, 72% yield) as a white solid in a 36 hour reaction. TLC R_f = 0.19 (EtOAc/Hexanes, 1:3); MS-ESI (*m/z*): $[M]^-$ calcd for $C_{20}H_{22}FNO_{12}$, 487.1; found 487.1.

2-fluoro-4-nitrophenyl-β-D-glucopyranoside (8). General procedure **3.4.3.1.2** with **116** (0.35 g, 0.73 mmol) yielded **8** (0.13 g, 40% yield) as yellow crystals. TLC R_f = 0.16 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.10-8.03 (m, 2 H, Ph), 7.49-7.43 (m, 1 H, Ph), 5.15 (d, $J_{H1,H2}$ = 7.4 Hz, 1 H, H-1), 3.90 (dd, $J_{H5,H6a}$ = 3.9 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.7 Hz, 1 H, H-6b), 3.57-3.37 (m, 4 H, H-2, H-3, H-4, H-5); ^{13}C NMR (100 MHz, $CDCl_3$) δ 153.0 (d, $^1J_{C,F}$ = 250.8Hz), 152.2 (d, $^2J_{C,F}$ = 10.5Hz), 143.4 (d, $^3J_{C,F}$ = 7.5 Hz), 121.8 (d, $^3J_{C,F}$ = 3.5 Hz), 117.9, 113.2 (d, $^2J_{C,F}$ = 23.4 Hz), 102.0, 78.6, 80.0, 74.6, 71.1, 62.4; HRMS (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}FNNaO_8$ 342.0596, found 342.0606.

3.4.3.1.8. Synthesis of 4-fluorophenyl-β-D-glucopyranoside (11).

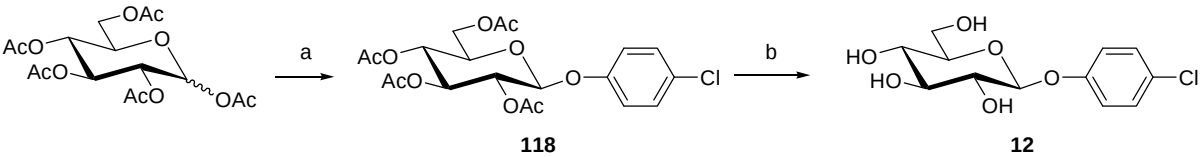


a) 4-fluorophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 30h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(4-fluorophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (117). General procedure **3.4.3.1.1** with 4-fluorophenol (0.22 g, 2.0 mmol) yielded **117** (0.38 g, 87% yield) as a white powder in a 30 hour reaction. TLC R_f = 0.17 (EtOAc:Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^-$ calcd for $C_{20}H_{22}FO_{10}$, 441.1; found 441.1.

4-fluorophenyl-β-D-glucopyranoside (11). General procedure **3.4.3.1.2** with **117** (0.38 g, 0.87 mmol) yielded **11** (0.23 g, 96% yield) as pale yellow crystals. TLC R_f = 0.23 ($MeOH/CH_2Cl_2$, 1:9); ^1H NMR (400 MHz, CD_3OD) δ 7.15-7.07 (m, 2 H, Ph), 7.05-6.97 (m, 2 H, Ph), 4.83 (d, $J_{H1,H2}$ = 7.6 Hz, 1 H, H-1), 3.90 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.71 (dd, $J_{H5,H6b}$ = 5.3 Hz, 1 H, H-6b), 3.51-3.36 (m, 4 H, H-2, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 159.5 (d, $^1J_{C,F}$ = 237.2 Hz), 155.4 (d, $^4J_{C,F}$ = 2.3 Hz), 119.4 (d, $^3J_{C,F}$ = 8.1 Hz), 116.7 (d, $^2J_{C,F}$ = 23.4 Hz), 103.1, 78.2, 78.0, 74.9, 71.4, 62.6; HRMS-ESI (*m/z*): $[M+NH_4]^+$ calcd for $C_{12}H_{19}FNO_6$ 292.1191, found 292.1193.

3.4.3.1.9. Synthesis of 4-chlorophenyl-β-D-glucopyranoside (12).

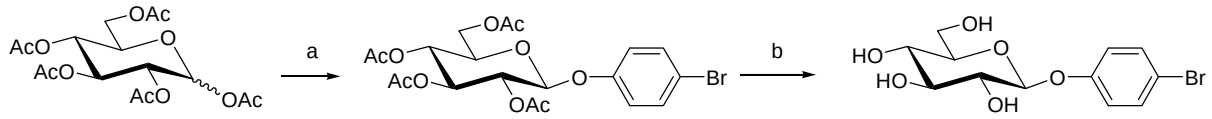


a) 4-chlorophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 30h; b) 0.1 M NaOMe in MeOH, RT, 18h.

(4-chlorophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (118). General procedure **3.4.3.1.1** with 4-chlorophenol (0.26 g, 2.0 mmol) yielded **118** (0.37 g, 82% yield) as white crystals in a 30 hour reaction. TLC R_f = 0.13 (EtOAc/Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^-$ calcd for $C_{20}H_{22}ClO_{10}$, 457.0; found 457.0.

4-chlorophenyl-β-D-glucopyranoside (12). General procedure **3.4.3.1.2** with **118** (0.37 g, 0.82 mmol) yielded **12** (0.23 g, 99% yield) as white crystals. TLC R_f = 0.24 (CH_2Cl_2/MeOH, 9:1); ^1H NMR (400 MHz, CD_3OD) δ 7.30-7.24 (m, 2 H, Ph), 7.11-7.05 (m, 2 H, Ph), 4.88 (d, $J_{H1,H2}$ = 7.2 Hz, 1 H, H-1), 3.90 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.4 Hz, 1 H, H-6b), 3.52-3.36 (m, 4 H, H-2, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 157.9, 130.3, 128.3, 119.4, 102.5, 78.2, 78.0, 74.9, 71.4, 62.5; HRMS-ESI (*m/z*): $[M+NH_4]^+$ calcd for $C_{12}H_{19}ClNO_6$ 308.0896, found 308.0902.

3.4.3.1.10. Synthesis of 4-bromophenyl-β-D-glucopyranoside (13).



a) 4-bromophenol, $BF_3{\cdot}OEt_2$, Et_3N in CH_2Cl_2, RT, 17h; b) 0.1 M NaOMe in MeOH, RT, 18h.

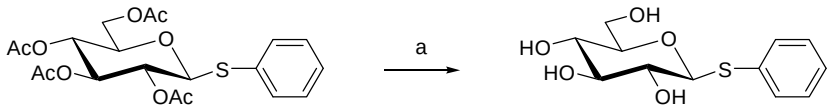
(4-bromophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (119). General procedure **3.4.3.1.1** with 4-bromophenol (0.35 g, 2.0 mmol) yielded **119** (0.09 g, 20.3% yield) as a white

powder in a 17 hour reaction. TLC R_f = 0.13 (EtOAc:Hexanes, 1:3); MS-ESI (*m/z*): $[M-H]^-$ calcd for $C_{20}H_{22}BrO_{10}$, 501.0; found 501.1.

4-bromophenyl-β-D-glucopyranoside (13). General procedure **3.4.3.1.2** with **119** (0.09 g, 0.2 mmol) yielded **13** (0.07 g, 99% yield) as pale yellow crystals. TLC R_f = 0.24 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 7.44-7.38 (m, 2 H, Ph), 7.06-7.00 (m, 2 H, Ph), 4.88 (d, $J_{H1,H2}$ = 7.6 Hz, 1 H, H-1), 3.90 (dd, $J_{H5,H6a}$ = 2.1 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.4 Hz, 1 H, H-6b), 3.53-3.36 (m, 4 H, H-2, H-3, H-4, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 158.37, 133.3, 119.8, 115.6, 102.4, 78.3, 78.0, 74.9, 71.4, 62.5; HRMS-ESI (*m/z*): $[M+NH_4]^+$ calcd for $C_{12}H_{19}BrNO_6$ 352.0391, found 352.0393.

3.4.3.2. Substituted *S*-phenyl-1-thio-β-D-glucosides (14, 15).

3.4.3.2.1. Synthesis of *S*-phenyl-1-thio-β-D-glucopyranoside (14).



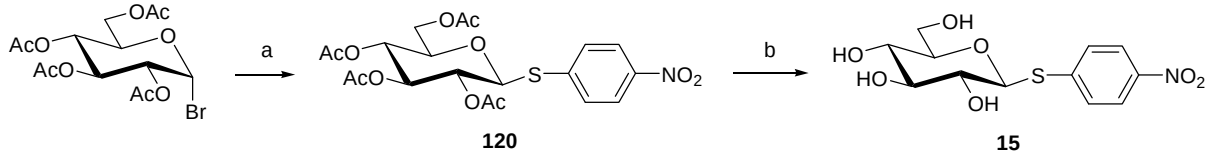
14

a) NaOMe 0.1M in MeOH, RT, 90 min.

S-phenyl-1-thio-β-D-glucopyranoside tetracetate (1.02 g, 2.31 mmol) was dissolved in MeOH (7.57 mL) to a final concentration of 300 mM. The solution was stirred on an ice bath and then sodium methoxide powder (1.25 g, 23.1 mmol) dissolved in 7.57 mL of MeOH was added to the reaction and allowed to proceed for 1.5 hours. Compound **14** (0.51 g, 81% yield) was purified by

flash chromatography with a 1% to 20% gradient of MeOH in CH_2Cl_2. TLC R_f = 0.12 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (500 MHz, CD_3OD) δ 7.60-7.57 (m, 2 H, Ph), 7.35-7.30 (m, 2 H, Ph), 7.29-7.26 (m, 1 H, Ph), 4.63 (d, $J_{H1,H2}$ = 9.8 Hz, 1 H, H-1), 3.89 (dd, $J_{H6a,H6b}$ = 12.1 Hz, $J_{H5,H6a}$ = 1.8 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.3 Hz, 1 H, H-6b), 3.42 (dd, $J_{H2,H3}$ = 8.6 Hz, $J_{H3,H4}$ = 8.6 Hz, 1 H, H-3), 3.38-3.30 (m, 2 H, H-4, H-5), 3.25 (dd, 1 H, H-2); ^{13}C NMR (125 MHz, CD_3OD) δ 135.21, 132.6, 129.8, 128.3, 89.3, 82.0, 79.6, 73.7, 71.3, 62.8; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}NaO_5S$, 295.06107, found 295.06113.

3.4.3.2.2. Synthesis of *S*-(4-nitrophenyl)-1-thio-β-D-glucopyranoside (15).



a) 4-nitrothiophenol, tetrabutylammonium bisulfate, Na_2CO_3, EtOAc, RT, 18h; b) NaOMe 0.1M in MeOH, RT, 8h.

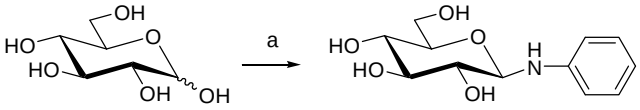
***S*-(4-nitrophenyl)-1-thio-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside** (**120**). Using a modified protocol described by D. Carrière *et al.*[72], 2,3,4,6-tetra-*O*-acetyl-α-D-bromoglucose (0.1 g, 0.24 mmol) was dissolved in 3 mL of EtOAc. Tetrabutylammonium bisulfate (83 mg, 0.24 mmol), 4-nitrothiophenol (0.19 g, 1.22 mmol) and 3 mL of a 1M solution of Na_2CO_3 were successively added, yielding **120** (0.03 g, 25% yield) as white crystals. TLC R_f = 0.11 (EtOAc/Hexanes,1:3); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.16 (d, $J_{Ph,Ph}$ = 9.0 Hz, 2 H, Ph), 7.60 (d, $J_{Ph,Ph}$ = 9.0 Hz, 2 H, Ph), 5.28 (t, J = 9.4 Hz, 1 H, H-3), 5.12-5.02 (m, 2 H, H-2, H-4), 4.88 (d, $J_{H1,H2}$ = 10.0 Hz, 1 H, H-1), 4.26 (dd, $J_{H5,H6a}$ = 5.6 Hz, $J_{H6a,H6b}$ = 12.3 Hz, 1 H, H-6a), 4.20 (dd, $J_{H5,H6b}$ = 2.4 Hz, 1 H, H-6b),

3.87-3.80 (m, 1 H, H-5), 2.11 (s, 3 H, OCH_3), 2.09 (s, 3 H, OCH_3), 2.05 (s, 3 H, OCH_3), 2.01 (s, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.5, 170.2, 169.5, 169.3, 147.2, 141.8, 131.2, 124.0, 84.5, 76.2, 73.3, 69.7, 68.1, 62.2, 20.9, 20.8, 20.7; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{20}H_{23}NNaO_{11}S$, 508.0885, found 508.0880.

***S*-(4-nitrophenyl)-1-thio-β-D-glucopyranoside (15).** Compound **120** (0.03 g, 0.06 mmol) was dissolved in MeOH (20 mL $mmol^{-1}$), sodium methoxide powder (4.9 mg, 0.09 mmol) was added, and the reaction allowed to proceed for 8 hours. The reaction was filtered over a small column of Celite 545 (Fisher Scientific, Pittsburgh, Pennsylvania, USA) to yield **15** (0.017 g, 90% yield) as a yellow solid. TLC R_f = 0.44 (15% MeOH in CH_2Cl_2); ^1H NMR (400 MHz, DMSO-d_6) δ 8.12 (d, $J_{Ph,Ph}$ = 9.0 Hz, 2 H, Ph), 7.62 (d, $J_{Ph,Ph}$ = 9.0 Hz, 2 H, Ph), 5.68 (br s, 1 H, OH), 5.50 (br s, 1 H, OH), 5.28 (br s, 1 H, OH), 4.92 (d, $J_{H1,H2}$ = 9.6 Hz, 1 H, H-1), 4.66 (br s, 1 H, OH), 3.71 (d, $J_{H6a,H6b}$ = 11.6 Hz, 1 H, H-6a), 3.47 (dd, $J_{H5,H6b}$ = 5.6 Hz, 1 H, H-6b), 3.42-3.24 (m, 2 H, H-3, H-5), 3.22-3.12 (m, 2 H, H-2, H-4); ^{13}C NMR (100 MHz, DMSO-d_6) δ 146.2, 144.9, 127.7, 123.8, 85.1, 81.1, 78.1, 72.5, 69.6, 60.8; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{15}NNaO_7S$, 340.0462, found 340.0467.

3.4.3.3. Substituted *N*-phenyl-β-D-glucosylamines (16-20).

3.4.3.3.1. Synthesis of *N*-phenyl-β-D-glucopyranosylamine (16).

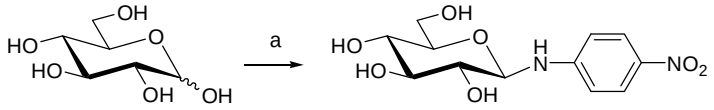


16

a) aniline, 100mM sodium phosphate buffer (pH 6.5), 40 °C, 5h.

This compound was synthesized as previously described[73], and yielded **16** (0.43 g, 58% yield) as a white solid. TLC R_f = 0.28 (CH_2Cl_2/MeOH, 8:2); ^{1}H NMR (500 MHz, CD_3OD) δ 7.16 (dd, $J_{Ph,Ph}$ = 7.4 Hz, $J_{Ph,Ph}$ = 8.4 Hz, 2 H, Ph), 6.82 (d, $J_{Ph,Ph}$ = 7.4 Hz, 2 H, Ph), 6.74 (t, $J_{Ph,Ph}$ = 7.4 Hz, 1 H, Ph), 4.58 (d, $J_{H1,H2}$ = 8.7 Hz, 1 H, H-1), 3.88 (dd, $J_{H6a,H6b}$ = 11.7, $J_{H5,H6a}$ = 1.1 Hz, 1 H, H-6a), 3.73-3.68 (m, 1 H, H-6b), 3.54-3.49 (m, 1 H, H-5), 3.42-3.32 (m, 3 H, H-2, H-3, H-4); ^{13}C NMR (125 MHz, CD_3OD) δ 148.0, 130.0, 119.5, 115.1, 86.8 (C-1), 79.0 (C-5), 78.3 (C-3), 74.6 (C-2), 71.7 (C-4), 62.6 (C-6). Spectral data were consistent with those previously reported[73]. HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{17}NNaO_5$, 278.09989, found 278.1001.

3.4.3.3.2. Synthesis of *N*-(4-nitrophenyl)-β-D-glucopyranosylamine (17).



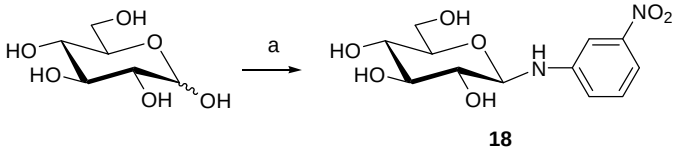
17

a) 4-nitroaniline, H_2SO_4 (1M) in MeOH, 50 °C, 2h.

A quantity of 4-nitroaniline (0.15 g, 1.1 mmol) was dissolved to a final concentration of 250 mM in MeOH and heated to 50 °C. D-glucose (0.13 g, 0.7 mmol) and then 3 μL of concentrated sulfuric acid were added and the reaction was allowed to proceed for 2 hours. A small scoop of sodium bicarbonate was added and the reaction was filtered. Following filtration, the desired product was recrystallized from the recovered filtrate. The resulting crystals were stored at -80 °C for 4 hours, washed with diethyl ether, dissolved with MeOH, and concentrated via

reduced pressure to yield **17** (0.05 g, 21% yield) as yellow crystals. TLC R_f = 0.30 (15% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, D_2O) δ 8.13 (d, $J_{Ph,Ph}$ = 9.2 Hz, 2 H, Ph), 6.89 (d, $J_{Ph,Ph}$ = 9.2 Hz, 2 H, Ph), 4.86 (d, $J_{H1,H2}$ = 8.7 Hz, 1 H, H-1), 3.91 (dd, $J_{H5,H6a}$ = 2.1 Hz, $J_{H6a,H6b}$ = 12.3 Hz, 1 H, H-6a), 3.74 (dd, $J_{H5,H6b}$ = 5.5 Hz, 1 H, H-6b), 3.66-3.58 (m, 2 H, H-3, H-5), 3.54-3.44 (m, 2 H, H-2, H-4); ^{13}C NMR (100 MHz, D_2O) δ 153.2, 139.4, 127.1, 113.6, 84.0, 77.3, 73.0, 70.1, 61.2; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}N_2NaO_7$, 323.0850, found 323.0844.

3.4.3.3.3. Synthesis of *N*-(3-nitrophenyl)-β-D-glucopyranosylamine (18).

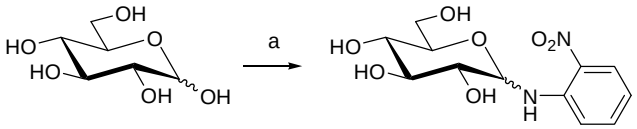


18

a) 3-nitroaniline, H_2SO_4 (1M) in MeOH, 70 °C, 3h.

Utilizing sulfuric acid instead of glacial acetic acid as previously described[73], yielded **18** (0.09 g, 17% yield) as bright yellow crystals. TLC R_f = 0.32 (15% MeOH in CH_2Cl_2); ^{1}H NMR (500 MHz, CD_3OD) δ 7.60 (t, $J_{Ph,Ph}$ = 2.2 Hz, 1 H, Ph), 7.53 (ddd, $J_{Ph,Ph}$ = 0.8 Hz, $J_{Ph,Ph}$ = 2.2 Hz, $J_{Ph,Ph}$ = 8.1 Hz, 1 H, Ph), 7.33 (t, $J_{Ph,Ph}$ = 8.1 Hz, 1 H, Ph), 7.13 (ddd, $J_{Ph,Ph}$ = 0.8 Hz, $J_{Ph,Ph}$ = 2.2 Hz, $J_{Ph,Ph}$ = 8.1 Hz, 1 H, Ph), 4.61 (d, $J_{H1,H2}$ = 8.7 Hz, 1 H, H-1), 3.86 (dd, $J_{H5,H6a}$ = 2.3 Hz, $J_{H6a,H6b}$ = 12.0 Hz, 1 H, H-6a), 3.69 (dd, $J_{H5,H6b}$ = 5.3 Hz, 1 H, H-6b), 3.51-3.47 (m, 1 H, H-3), 3.46-3.40 (m, 1 H, H-5), 3.39 (m, 2 H, H-4, H-2); ^{13}C NMR (125 MHz, CD_3OD) δ 150.5, 149.6, 130.8, 120.8, 113.5, 109.0, 86.2, 79.1, 78.6, 74.5, 71.6, 62.6; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}N_2NaO_7$, 323.0850, found 323.0844.

3.4.3.3.4. Synthesis of *N*-(2-nitrophenyl)-D-glucopyranosylamine (19).

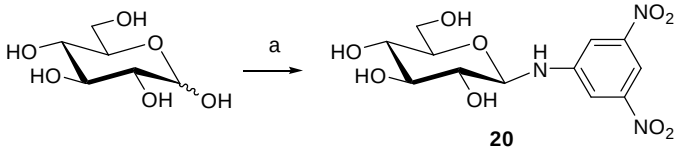


19

a) 2-nitroaniline, H_2SO_4 (1M) in MeOH, 40 °C, 1h.

A quantity of 2-nitroaniline (0.15 g, 1.1 mmol) was dissolved to a final concentration of 250 mM in MeOH and heated to 40 °C. D-glucose (0.130 g, 0.72 mmol) was added and then 36 μL of 1 M H_2SO_4 in MeOH (36 mmol) was added over 1 hour. The reaction was concentrated and then purified by flash chromatography with a 10% to 15% gradient of MeOH in CH_2Cl_2 to yield **19** (0.04 g, 17% yield) as yellow crystals with α- and β-anomers present in a 1:2 ratio. TLC R_f = 0. 48 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 8.16 (t, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.55 (t, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.23 (d, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 6.84 (t, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 5.30 (d, $J_{H1\alpha,H2}$ = 4.6 Hz, 1 H, H1α), 4.72 (d, $J_{H1\beta,H2}$ = 8.4 Hz, 1 H, H-1β), 3.89 (dd, $J_{H5,H6a}$ = 1.9 Hz, $J_{H6a,H6b}$ = 11.9 Hz, 1 H, H-6a), 3.82-3.66 (m, 2 H, H-6b, H-3), 3.54-3.35 (m, 3 H, H-2, H-4, H-5); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{16}N_2NaO_7$, 323.0850, found 323.0858.

3.4.3.3.5. Synthesis of *N*-(3,5-dinitrophenyl)-β-D-glucopyranosylamine (20).



20

a) 3,5-dinitroaniline, H_2SO_4 (1M) in MeOH, 50 °C, 2h.

A quantity of 3,5-dinitroanline (0.4 g, 2.2 mmol) was dissolved to a final concentration of 250 mM in MeOH and heated to 50 °C. D-glucose (0.39 g, 2.2 mmol) and then 3 μL of concentrated sulfuric acid were added and the reaction was allowed to proceed for 2 hours. The desired product recrystallized from solution as the reaction proceeded. The resulting crystals were stored at -80 °C for 4 hours, washed with diethyl ether, dissolved with MeOH, and concentrated with reduced pressure to yield **20** (0.21 g, 28% yield) as yellow crystals. TLC R_f = 0.36 (15% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 8.25 (t, $J_{Ph,Ph}$ = 2.0 Hz, 1 H, Ph), 7.92 (d, $J_{Ph,Ph}$ = 2.0 Hz), 2 H, Ph), 4.67 (d, $J_{H1,H2}$ = 8.6 Hz, 1 H, H-1), 3.88 (dd, $J_{H5,H6a}$ = 2.4 Hz, $J_{H6a,H6b}$ = 12.0 Hz, 1 H, H-6a), 3.68 (dd, $J_{H5,H6b}$ = 5.8 Hz, 1 H, H-6b), 3.53-3.45 (m, 2 H, H-3, H-5), 3.43-3.34 (m, 2 H, H-2, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 150.74, 150.65, 114.0, 107.6, 85.7, 79.1, 78.9, 74.5, 71.6; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{15}N_3NaO_9$, 368.07005, found 368.06973.

3.4.3.4. *O*-substituted oxyamines (**122**, **124**, **126**, **128**, **130**, **132**, **134**, **136**, **138**, **140**, **142**, **144**).

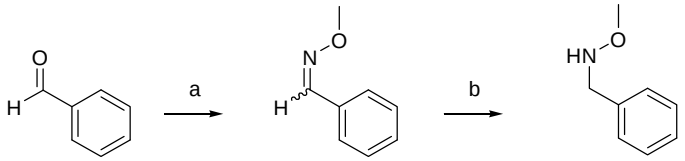
3.4.3.4.1. General reductive amination procedure.

a) Aldehyde was dissolved in CH_2Cl_2 to a final concentration of 0.45 M. Unless noted, to this was added 1.5 equivalents of the appropriate *O*-substituted oxyamine hydrochloride salt and 2.2 equivalents of pyridine. The mixture was stirred for 2 hours at room temperature. TLC analysis revealed the substrate to be completely consumed with two products (presumably *E*- and *Z*-oximes) being formed. The reaction mixture was subsequently washed with 5% aqueous HCl (3 x 50 mL) and saturated NaCl (2 x 50 mL). The resulting organic layer was dried over Na_2SO_4

and concentrated under reduced pressure to provide the crude oxime which was used in subsequent reactions without further purification.

b) Crude oxime was dissolved in EtOH to a final concentration of 1.5 M. The reaction mixture was cooled to 0 ºC, 3 equivalents of $NaBH_3CN$ were added, and the solution was stirred for 15 min. An equal volume of 20% HCl in EtOH chilled to 0 °C was subsequently added in a drop-wise fashion over 10 min. The reaction was then allowed to warm to RT and stirred overnight. TLC analysis revealed complete consumption of substrate in all cases. The reaction was neutralized with the addition of Na_2CO_3 until the evolution of gas halted, concentrated under reduced pressure, and CH_2Cl_2 (20 mL) was added. The resulting mixture was washed with saturated $NaHCO_3$ (2 x 50 mL), dried over Na_2SO_4, and the collected organics concentrated under reduced pressure. The concentrate was purified by flash chromatography to yield the desired *O*-substituted oxyamine product.

3.4.3.4.2. Synthesis of *N*-methoxybenzylamine (122).



121 **122**

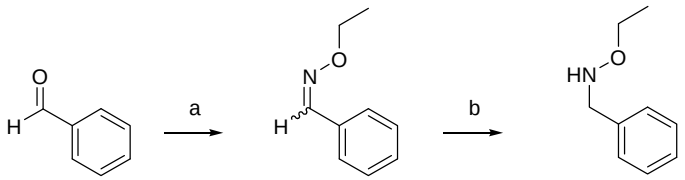
a) $MeONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

benzaldehyde-*O*-methyloxime (121). According to general procedure **3.4.3.4.1.a**, benzaldehyde (4.8 g, 49.2 mmol) afforded oxime **121** (6.1 g, 91% crude yield) as a colorless oil. TLC R_f = 0.82

(EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.03 (s, 1 H, NCH), 7.58-7.51 (m, 2 H, Ph), 7.37-7.32 (m, 3 H, Ph), 3.95 (s, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.8, 132.5, 130.0, 129.0, 127.3, 62.2; HRMS-ESI (*m/z*): $[M]^{+\bullet}$ calcd for C_8H_9NO, 135.0679, found 135.0684.

***N*-methoxybenzylamine (122).** According to general procedure **3.4.3.4.1.b**, oxime **121** (6.1 g, 44.9 mmol) provided desired product **122** (3.3 g, 53% yield) as a colorless oil. TLC R_f = 0.31 (EtOAc:hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.38-7.22 (m, 5 H, Ph), 5.71 (br s, 1 H, NH), 4.04 (s, 2 H, C*H*$_2$NH), 3.50 (d, J = 0.4 Hz, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 137.9, 129.1, 128.7, 127.7, 62.1, 56.5; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_8H_{12}NO$, 138.0914, found 138.0916.

3.4.3.4.3. Synthesis of *N*-ethoxybenzylamine (124).



123 **124**

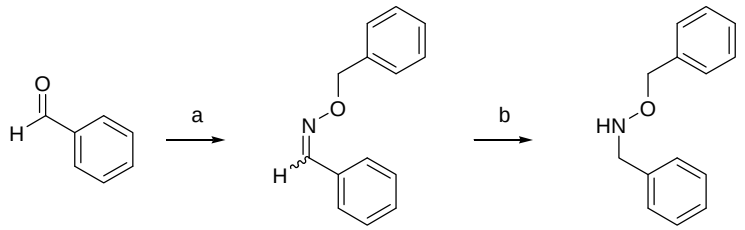
a) $EtONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

benzaldehyde-*O*-ethyloxime (123). According to general procedure **3.4.3.4.1.a**, benzaldehyde (1.0 g, 9.4 mmol) afforded oxime **123** (1.32 g, 94% crude yield) as a colorless oil. TLC R_f = 0.52, 0.62 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.07 (s, 1 H, NCH), 7.62-7.53 (m, 2 H, Ph), 7.42-7.23 (m, 3 H, Ph), 4.23 (q, ^{3}J = 7.2 Hz, 2 H, OC*H*$_2$), 1.32 (t, ^{3}J = 7.0 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 148.6, 132.8, 130.0, 129.0, 127.3, 70.1, 14.9;

HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_9H_{12}NO$, 150.0914, found 150.0919.

***N*-ethoxybenzylamine (124).** According to general procedure **3.4.3.4.1.b**, oxime **123** (1.32 g, 8.8 mmol) provided desired product **124** (0.886 g, 66% yield) as a colorless oil. TLC R_f = 0.37 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.37-7.22 (m, 5 H, Ph), 5.59 (br s, 1 H, NH), 4.03 (s, 2 H, NHC*H*$_2$), 3.69 (q, ^3J = 7.0 Hz, 2 H, OCH_2), 1.13 (t, ^3J = 7.0 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 137.9, 129.2, 128.6, 127.7, 69.5, 56.9, 14.5; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_9H_{14}NO$, 152.1070, found 152.1066.

3.4.3.4.4. Synthesis of *N*-benzoxybenzylamine (126).



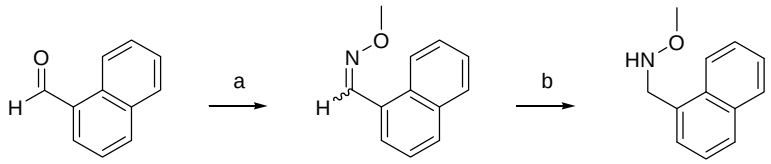
125 **126**

a) $BnONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

benzaldehyde-*O*-benzyloxime (125). According to general procedure **3.4.3.4.1.a**, benzaldehyde (1.5 g, 14.1 mmol) afforded oxime **125** (2.51 g, 84% crude yield) as a colorless oil. TLC R_f = 0.47, 0.56 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.13 (d, $J_{NCH,Ph}$ = 0.8 Hz, 1 H, NCH), 7.63-7.52 (m, 2 H, Ph), 7.46-7.25 (m, 8 H, Ph), 5.21 (s, 2 H, OCH_2); ^{13}C NMR (100 MHz, $CDCl_3$): δ 149.4, 137.8, 132.6, 130.2, 129.0, 128.8, 128.7, 128.3, 127.4, 76.7; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{14}H_{14}NO$, 212.1070, found 212.1073.

***N*-benzoxybenzylamine (126).** According to general procedure **3.4.3.4.1.b**, oxime **125** (2.45 g, 11.6 mmol) provided desired product **126** (0.60 g, 24% yield) as a colorless oil. TLC R_f = 0.25 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.37-7.22 (m, 10 H, Ph), 5.71 (br s, 1 H, NH), 4.65 (t, $J_{OCH_2,Ph}$ = 1.4 Hz, 2 H, OC*H*$_2$), 4.04 (d, $J_{NHCH_2,Ph}$ = 0.4 Hz, 2 H, NHC*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$): δ 138.2, 137.9, 129.3, 128.8, 128.7, 128.7, 128.1, 127.7, 76.6, 56.8; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{14}H_{16}NO$, 214.1227, found 214.1219.

3.4.3.4.5. Synthesis of *N*-methoxy-1-naphthalenemethanamine (128).



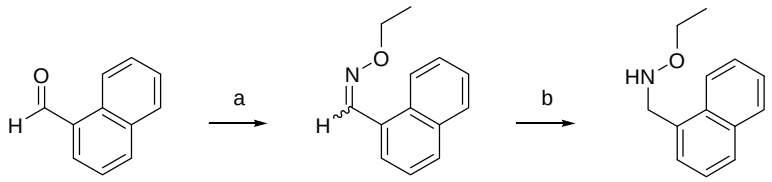
127 **128**

a) $MeONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

1-naphthaldehyde-*O*-methyloxime (127). According to general procedure **3.4.3.4.1.a**, 1-naphthaldehyde (2.0 g, 12.7 mmol) gave oxime **127** (2.0 g, 83% crude yield) as a yellow oil. TLC R_f = 0.56, 0.65 (EtOAc/hexanes, 1:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.71 (s, 1 H, NCH), 8.52 (m, 1 H, Ph), 7.84 (m, 2 H, Ph), 7.74 (m, 1 H, Ph), 7.58-7.42 (m, 3 H, Ph), 4.06 (s, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.5, 133.9, 130.8, 130.5, 128.8, 128.1, 127.4, 127.1, 126.2, 125.4, 124.6, 62.2; HRMS-ESI (*m/z*): $[M]^{+\bullet}$ calcd for $C_{12}H_{11}NO$, 185.0836, found 185.0842.

***N*-methoxy-1-naphthalenemethanamine (128).** According to general procedure **3.4.3.4.1.b**, oxime **127** (2.0 gm, 10.5 mmol) yielded **128** (1.2 g, 63% yield) as a yellow oil. TLC R_f = 0.41 (EtOAc/hexanes, 1:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.15 (d, $J_{Ph,Ph}$ = 8.8 Hz, 1 H, Ph), 7.84 (d, $J_{Ph,Ph}$ = 8.4 Hz, 1 H, Ph), 7.77 (d, $J_{Ph,Ph}$ = 8.4 Hz, 1 H, Ph), 7.56-7.36 (m, 4 H, Ph), 5.77 (br s, 1 H, NH), 4.51 (s, 2 H, NHC*H*$_2$), 3.53 (d, J = 0.4 Hz, 3 H, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.1, 133.0, 132.3, 129.0, 128.7, 127.8, 126.5, 126.0, 125.7, 124.0, 62.1, 54.1; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}NO$, 188.1070, found 188.1070.

3.4.3.4.6. Synthesis of *N*-ethoxy-1-naphthalenemethanamine (130).



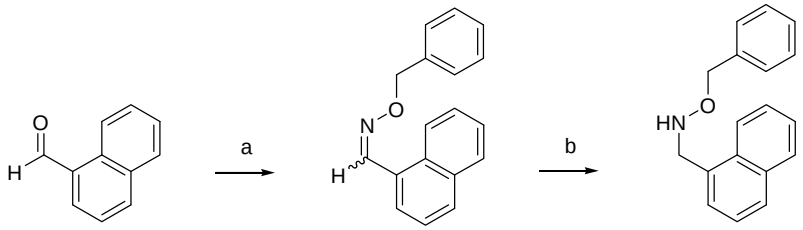
129 **130**

a) $EtONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

1-naphthaldehyde-*O*-ethyloxime (129). According to general procedure **3.4.3.4.1.a**, 1-naphthaldehyde (0.5 g, 3.2 mmol) gave oxime **129** (0.57 g, 90% crude yield) as a yellow oil. TLC R_f = 0.60, 0.65 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.77-8.74 (m, 1 H, NCH), 8.58 (d, $J_{Ph,Ph}$ = 8.3 Hz, 1 H, Ph), 7.90-7.86 (m, 2 H, Ph), 7.79 (d, $J_{Ph,Ph}$ = 7.2 Hz, 1 H, Ph), 7.62-7.46 (m, 3 H, Ph), 4.40-4.32 (m, 2 H, OC*H*$_2$), 1.45-1.39 (m, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 148.3, 134.0, 130.8, 130.4, 128.8, 128.4, 127.3, 127.1, 126.2, 125.4, 124.7, 70.0, 14.9; HRMS-ESI (*m/z*): $[M]^{+\bullet}$ calcd for $C_{13}H_{13}NO$, 199.0992, found 199.0993.

***N*-ethoxy-1-naphthalenemethanamine (130).** According to general procedure **3.4.3.4.1.b**, oxime **129** (0.45 gm, 2.2 mmol) yielded **130** (0.28 g, 62% yield) as a slightly yellow oil. TLC R_f = 0.38 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.17 (d, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.84 (dd, $J_{Ph,Ph}$ = 0.8, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.78 (d, $J_{Ph,Ph}$ = 8.0 Hz, 1 H, Ph), 7.56-7.36 (m, 4 H, Ph), 5.65 (br s, 1 H, NH), 4.51 (s, 2 H, NHCH_2) 3.73 (q, ^3J = 7.0 Hz, 2 H, OCH_2), 1.14 (t, ^3J = 7.0 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 134.1, 133.2, 132.4, 129.0, 128.7, 127.9, 126.5, 126.0, 125.7, 124.1, 69.6, 54.5, 14.6; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{13}H_{16}NO$, 202.1227, found 202.1217.

3.4.3.4.7. Synthesis of *N*-benzoxy-1-naphthalenemethanamine (132).



131 **132**

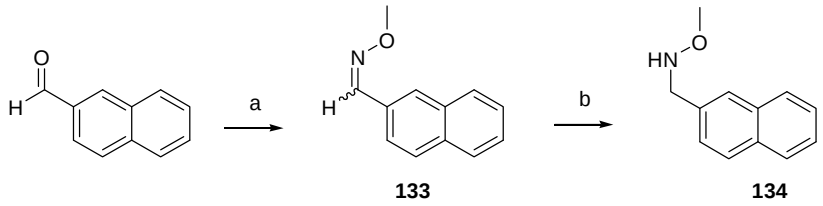
a) $BnONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

1-naphthaldehyde-*O*-benzyloxime (131). According to general procedure **3.4.3.4.1.a**, 1-naphthaldehyde (0.5 g, 3.2 mmol) gave oxime **131** (0.64 g, 77% crude yield) as a yellow oil. TLC R_f = 0.54, 0.61 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.77 (s, 1H, NCH), 8.52 (dd, $J_{Ph,Ph}$ = 0.7 Hz, $J_{Ph,Ph}$ = 8.3 Hz, 1 H, Ph), 7.84 (dd, $J_{Ph,Ph}$ = 0.6 Hz, $J_{Ph,Ph}$ = 8.2 Hz, 2 H, Ph), 7.73 (d, $J_{Ph,Ph}$ = 7.1 Hz, 1 H, Ph), 7.56-7.29 (m, 8 H, Ph), 5.30 (s, 2 H, OCH_2); ^{13}C NMR (100 MHz, $CDCl_3$): δ 149.2, 137.8, 134.1, 131.0, 130.7, 129.0, 128.8, 128.8, 128.3, 128.3, 127.8,

127.3, 126.4, 125.5, 124.9, 76.8; HRMS-ESI (*m/z*): [M+Na] calcd for $C_{18}H_{15}NaNO$, 284.1046, found 284.1053.

***N*-benzoxy-1-naphthalenemethanamine (132).** According to general procedure **3.4.3.4.1.b**, oxime **131** (0.52 gm, 2.0 mmol) yielded **132** (0.29 g, 55% yield) as a slightly yellow oil. TLC R_f = 0.40 (1:8::EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.04-7.98 (m, 1 H, Ph), 7.86-7.72 (m, 2 H, Ph), 7.56-7.14 (m, 9H, Ph), 5.75 (br s, 1 H, NH), 4.66 (s, 2 H, OCH_2), 4.48 (s, 2 H, NHC*H*$_2$); ^{13}C NMR (100 MHz, $CDCl_3$): δ 138.3, 134.1, 133.0, 132.4, 128.94, 128.89, 128.7, 128.6, 128.1, 128.0, 126.4, 126.0, 125.6, 124.3, 76.6, 54.6; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{18}H_{18}NO$, 264.1283, found 264.1389.

3.4.3.4.8. Synthesis of *N*-methoxy-2-naphthalenemethanamine (134).



133 **134**

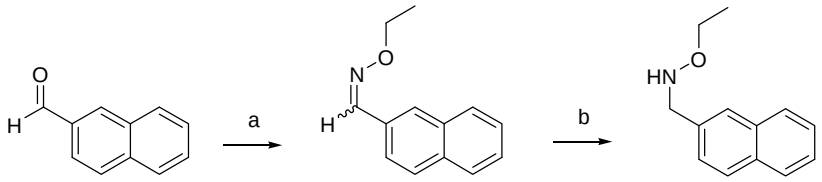
a) $MeONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

2-naphthaldehyde-*O*-methyloxime (133). According to general procedure **3.4.3.4.1.a**, 2-naphthaldehyde (2.0 g, 12.8 mmol) provided the desired oxime **133** (2.3 g, 99% crude yield) as a white solid. TLC R_f = 0.48, 0.60 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (s, 1 H, NC*H*), 7.92-7.80 (m, 5 H, Ph), 7.56-7.48 (m, 2 H, Ph), 4.058 (s, 3 H, OC*H*$_3$), 4.055 (s, 3 H, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ 148.0, 134.4, 133.5, 130.2, 128.9, 128.6, 128.6, 128.2,

127.2, 126.9, 123.3, 62.4; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{12}NO$, 186.0841, found 186.0920.

***N*-methoxy-2-naphthalenemethanamine (134).** According to general procedure **3.4.3.4.1.b**, oxime **133** (2.3 g, 12.4 mmol) yielded **134** (1.5 g, 63% yield) as an orange oil. TLC R_f = 0.36 (EtOAc/hexanes, 1:4); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.84-7.74 (m, 4 H, Ph), 7.52-7.39 (m, 3 H, Ph), 5.80 (br s, 1 H, NH), 4.18 (s, 2 H, NHC*H*$_2$), 3.50 (t, *J* = 0.4 Hz, 3 H, OC*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ 135.5, 133.7, 133.1, 128.4, 128.1, 128.0, 127.9, 127.2, 126.3, 126.1, 62.2, 56.6; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}NO$, 188.1070, found 188.1069.

3.4.3.4.9. Synthesis of *N*-ethoxy-2-naphthalenemethanamine (136).



135 **136**

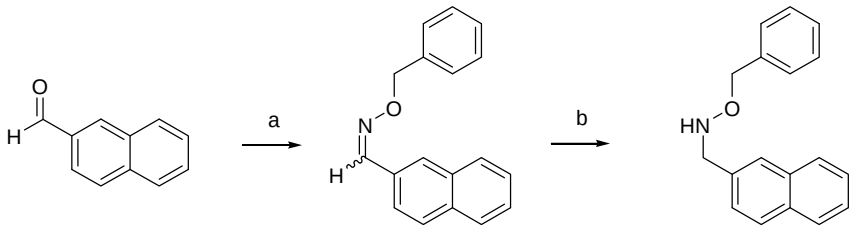
a) $EtONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

2-naphthaldehyde-*O*-ethyloxime (135). According to general procedure **3.4.3.4.1.a**, 2-naphthaldehyde (2.0 g, 12.8 mmol) provided the desired oxime **135** (2.46 g, 85% crude yield) as an off-white solid. TLC R_f = 0.50, 0.60 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.21 (s, 1 H, NCH), 7.89-7.75 (m, 5 H, Ph), 7.52-7.42 (m, 2 H, Ph), 4.272 (q, ^3J = 7.0 Hz, 2 H, OCH_2), 4.269 (q, ^3J = 7.0 Hz, 2 H, OCH_2), 1.355 (t, ^3J = 7.0 Hz, 3 H, OCH_2CH_3), 1.353 (t, ^3J = 7.0 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 148.8, 134.4, 135.5, 130.5, 128.8,

128.6, 128.5, 128.2, 127.1, 126.8, 123.3, 70.2, 15.0; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{13}H_{14}NO$, 200.1075, found 200.1080.

***N*-ethoxy-2-naphthalenemethanamine (136).** According to general procedure **3.4.3.4.1.b**, oxime **135** (2.46 g, 12.4 mmol) yielded **136** (1.52 g, 59% yield) as a yellow oil. TLC R_f = 0.43 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.84-7.75 (m, 4 H, Ph), 7.52-7.40 (m, 3 H, Ph), 5.66 (br s, 1 H, NH), 4.19 (s, 2H, NHC*H*$_2$), 3.70 (q, ^3J = 6.8 Hz, 2 H, OCH_2), 1.12 (t, ^3J = 6.8 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 135.5, 133.7, 133.1, 128.3, 128.1, 128.0, 127.4, 129.3, 126.1, 69.7, 57.1, 14.5; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{13}H_{16}NO$, 202.1227, found 202.1222.

3.4.3.4.10. Synthesis of *N*-benzoxy-2-naphthalenemethanamine (138).



137 **138**

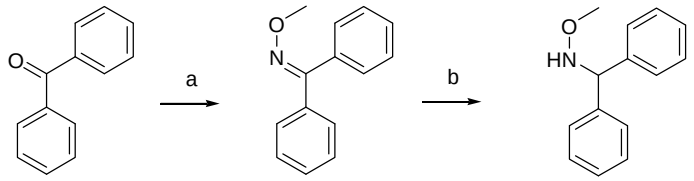
a) $BnONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

2-naphthaldehyde-*O*-benzyloxime (137). According to general procedure **3.4.3.4.1.a**, 2-naphthaldehyde (2.0 g, 12.8 mmol) provided of the desired oxime **137** (3.48 g, 87% crude yield) as an off-white solid. TLC R_f = 0.61, 0.74 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.29 (s, 1 H, NCH), 7.89-7.76 (m, 5 H, Ph), 7.54-7.43 (m, 4 H, Ph), 7.42-7.29 (m, 3 H, Ph), 5.26

(s, 2 H, OCH_2); ^{13}C NMR (100 MHz, $CDCl_3$): δ 149.5, 137.8, 134.5, 133.5, 130.3, 128.9, 128.8, 128.8, 128.8, 128.6, 128.3, 128.2, 127.2, 126.9, 123.4, 76.9; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{18}H_{16}NO$, 262.1227, found 262.1234.

***N*-benzoxy-2-naphthalenemethanamine (138).** According to general procedure **3.4.3.4.1.b**, oxime **137** (3.48 g, 13.3 mmol) yielded **138** (0.58 g, 17% yield) as a yellow solid. TLC R_f = 0.51 (EtOAc/hexanes, 1:8); ^1H NMR (400 MHz, $CDCl_3$): δ 7.87-7.72 (m, 4 H, Ph), 7.52-7.40 (m, 3 H, Ph), 7.35-7.21 (m, 5 H, Ph), 5.81 (br s, 1 H, NH), 4.65 (dd, $J_{OCH2,Ph}$ = 2.2 Hz, J = 3.0 Hz, 2 H, OCH_2), 4.18 (br s, 2 H, $NHCH_2$); ^{13}C NMR (100 MHz, $CDCl_3$): δ 138.2, 135.5, 133.7, 133.2, 128.8, 128.8, 128.6, 128.3, 128.1, 128.0, 128.0, 127.4, 126.3, 126.1, 76.7, 57.0; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{18}H_{18}NO$, 264.1383, found 264.1393.

3.4.3.4.11. Synthesis of *N*-methoxybenzhydrylamine (140).



139 **140**

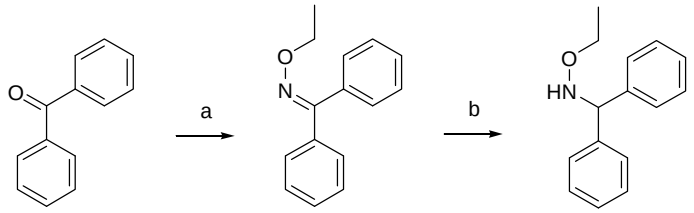
a) $MeONH_2.HCl$, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

benzophenone-*O*-methyloxime (139). According to general procedure **3.4.3.4.1.a**, benzophenone (0.36 g, 2.0 mmol) with 25 equiv. of pyridine provided the desired oxime **139** (0.33 g, 78% crude yield) as a white solid. TLC R_f = 0.71 (EtOAc/hexanes, 1:8); ^1H NMR (400 MHz, $CDCl_3$): δ 7.56-7.32 (m, 10 H, Ph), 4.02-4.01 (m, 3 H, OCH_3); ^{13}C NMR (100 MHz,

$CDCl_3$): δ 157.0, 136.7, 133.6, 129.6, 129.5, 129.1, 128.6, 128.4, 128.2, 62.7; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{14}H_{14}NO$, 212.1070, found 212.1079.

***N*-methoxybenzhydrylamine (140).** According to general procedure **3.4.3.4.1.b**, oxime **139** (0.32 g, 1.49 mmol) with 6 equivalents of $NaBH_3CN$ yielded **140** (0.17 g, 54% yield) as an oil. TLC R_f = 0.43 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.40-7.16 (m, 10 H, Ph), 5.84 (br s, 1 H, NH), 5.20 (s, 1 H, NHC*H*), 3.48 (s, 3 H, OCH_3); ^{13}C NMR (100 MHz, $CDCl_3$) δ 141.4, 128.7, 127.9, 127.7, 69.6, 62.6[1]; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{14}H_{15}NaNO$, 236.1046, found 236.1057.

3.4.3.4.12. Synthesis of *N*-ethoxybenzhydrylamine (142).



141 **142**

a) $EtONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

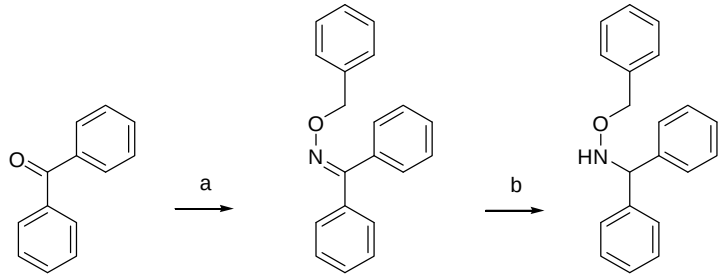
benzophenone-*O*-ethyloxime (141). According to general procedure **3.4.3.4.1.a**, benzophenone (0.36 g, 2.0 mmol) with 25 equiv. of pyridine provided the desired oxime **141** (0.41 g, 91% crude yield) as a colorless oil. TLC R_f = 0.68 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.52-7.26 (m, 10H, Ph), 4.24 (q, ^3J = 7.0 Hz, 2 H, OCH_2), 1.30 (t, ^3J = 7.0 Hz, 3 H, OCH_2CH_3);

[1] Not all phenyl ^{13}C resonances were observed, presumably due to spectral overlap in the 128.7-127.7 ppm region.

^{13}C NMR (100 MHz, $CDCl_3$): δ 156.6, 137.1, 133.8, 129.7, 129.4, 129.0, 128.5, 128.3, 128.2, 70.4, 15.1; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{15}H_{16}NO$, 226.1227, found 226.1235.

***N*-ethoxybenzhydrylamine (142).** According to general procedure **3.4.3.4.1.b**, oxime **141** (0.32 g, 1.49 mmol) with 6 equivalents of $NaBH_3CN$ yielded **142** (0.17 g, 54% yield) as an oil. TLC R_f = 0.50 (EtOAc/hexanes, 1:8); ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.44-7.17 (m, 10 H, Ph), 5.78 (br s, 1 H, NH), 5.21 (s, 1 H, NHC*H*), 3.70 (q, ^3J = 7.0 Hz, 2 H, O$CH_2$$CH_3$), 1.07 (t, ^3J = 7.0 Hz, OCH_2CH_3); ^{13}C NMR (100 MHz, $CDCl_3$): δ 141.6, 128.7, 128.1, 127.7, 70.0, 69.7, 14.4[2]; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{15}H_{17}NaNO$, 250.1203, found 250.1207.

3.4.3.4.13. Synthesis of *N*-benzoxybenzhydrylamine (144).



143 **144**

a) $BnONH_2$.HCl, CH_2Cl_2, pyridine, RT, 2h; b) $NaBH_3CN$, 20% HCl in EtOH, 0 °C, 18h.

benzophenone-*O*-benzyloxime (143). According to general procedure **3.4.3.4.1.a**, benzophenone (0.36 g, 2.0 mmol) with 20 equiv. of pyridine provided the desired oxime **143** (0.55 g, 96% crude yield) as a white solid. TLC R_f = 0.53 (EtOAc/hexanes, 1:8); ^{1}H NMR (400

[2] Not all phenyl ^{13}C resonances were observed, presumably due to spectral overlap in the 141.6 and 128.7-127.7 ppm regions.

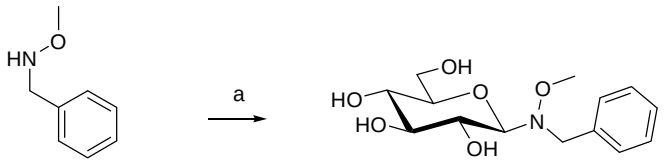
MHz, $CDCl_3$): δ 7.52-7.18 (m, 15 H, Ph), 5.23 (s, 2 H, OCH_2); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{20}H_{18}NO$, 288.1388, found 288.1398.

***N*-benzoxybenzhydrylamine (144).** According to general procedure **3.4.3.4.1.b**, oxime **143** (0.52 g, 1.8 mmol) with 6 equivalents of $NaBH_3CN$ yielded **144** (0.086 g, 17% yield) as a colorless oil. TLC R_f = 0.51 (1:8 EtOAc:hexanes); ^1H NMR (400 MHz, $CDCl_3$): δ 7.46-7.16 (m, 15 H, Ph), 5.87 (br s, 1 H, NH), 5.24 (br s, 1 H, NHC*H*), 4.65 (s, 2 H, OCH_2); ^{13}C NMR (100 MHz, $CDCl_3$): δ 141.4, 138.0, 128.9, 128.8, 128.6, 128.1, 128.1, 127.8, 77.0, 69.8; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{20}H_{20}NO$, 290.1540, found 290.1543.

3.4.3.5. *N,N*-disubstituted-β-D-glucopyranosylamines (21-32).

3.4.3.5.1. General neoglycosylation procedure. According to a modified procedure from Goff *et al.*[74], *O*-substituted oxyamine was dissolved in MeOH to a final concentration of 100 mM. 3 equivalents of D-glucose and 1.5 equivalents of acetic acid were added (unless otherwise noted). Reactions were allowed to proceed with stirring at 40 °C and monitored by TLC. Compounds were purified utilizing Extract-Clean SPE SI columns (from Alltech) pre-equilibrated with 1% MeOH in CH_2Cl_2. A gradient of 4 column volumes of 1% MeOH in CH_2Cl_2, 8 column volumes with 5% MeOH in CH_2Cl_2, and 12 column volumes of 10% MeOH in CH_2Cl_2 was sufficient for all purifications. Desired fractions were concentrated with reduced pressure to yield the final β-D-glucoside product.

3.4.3.5.2. Synthesis of *N*-(*N*-benzyl-*N*-methoxy)-β-D-glucopyranosylamine (21).

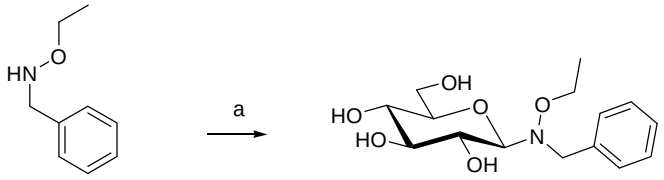


122 **21**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-benzyl-*N*-methoxy)-β-D-glucopyranosylamine (21).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **122** (76 mg, 0.55 mmol) yielded **21** (90 mg, 54% yield) as a colorless syrup. TLC R_f = 0.31 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.38-7.31 (m, 2 H, Ph), 7.25-7.13 (m, 3 H, Ph), 4.08 (d, ^2J = 12.8 Hz, 1 H, NCH_2), 3.95 (d, ^2J = 12.8 Hz, 1 H, NCH_2), 3.85 (d, $J_{H1,H2}$ = 9.2 Hz, 1 H, H-1), 3.79 (dd, $J_{H5,H6a}$ = 1.6 Hz, $J_{H6a,H6b}$ = 12.2 Hz, 1 H, H-6a), 3.62 (dd, $J_{H5,H6b}$ = 5.4 Hz, 1 H, H-6b), 3.48-3.40 (m, 1 H, H-2), 3.31 (s, 3 H, OCH_3), 3.28-3.18 (m, 2 H, H-3, H-4), 3.12-3.02 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 138.4, 131.0, 129.2, 128.4, 93.2, 79.6, 79.4, 71.5, 71.2, 62.8, 62.5, 57.6; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{14}H_{21}NaNO_6$, 322.1262, found 322.1268.

3.4.3.5.3. Synthesis of *N*-(*N*-benzyl-*N*-ethoxy)-β-D-glucopyranosylamine (22).

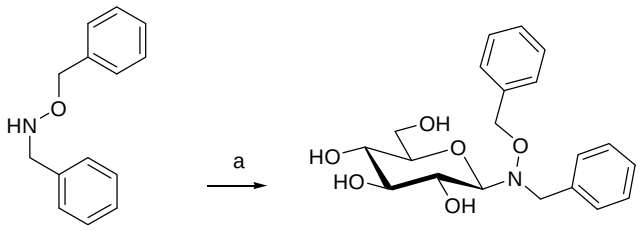


124 **22**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-benzyl-*N*-ethoxy)-β-D-glucopyranosylamine (22).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **124** (73 mg, 0.48 mmol) yielded **22** (114 mg, 75% yield) as a colorless syrup. TLC R_f = 0.36 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.46-7.38 (m, 2 H, Ph), 7.34-7.20 (m, 3 H, Ph), 4.15 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 4.04 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 3.94 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.88 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.0 Hz, 1 H, H-6a), 3.76-3.62 (m, 2 H, H-6b, OCH_2), 3.60-3.46 (m, 2 H, OCH_2, H-2), 3.36-3.28 (m, 2 H, H-3, H-4), 3.20-3.12 (m, 1 H, H-5), 0.96 (t, ^3J = 7.2 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, CD_3OD) δ 138.5, 131.1, 129.2, 128.4, 93.3, 79.5, 97.4, 71.5, 71.1, 70.8, 62.7, 58.0, 14.0; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{15}H_{23}NaNO_6$, 336.1418, found 336.1431.

3.4.3.5.4. Synthesis of *N*-(*N*-benzoxy-*N*-benzyl)-β-D-glucopyranosylamine (23).



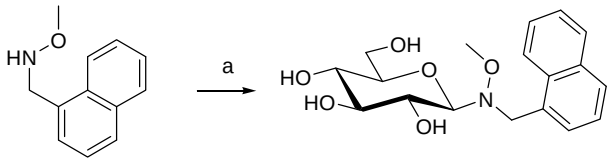
126 **23**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-benzoxy-*N*-benzyl)-β-D-glucopyranosylamine (23).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **126** (92 mg, 0.43 mmol) yielded **23** (123 mg, 76% yield) as a colorless syrup. TLC R_f = 0.50 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.48-7.40 (m, 2 H, Ph), 7.36-7.20 (m, 6 H, Ph), 7.16-7.08 (m, 2 H, Ph), 4.61 (d, ^2J = 9.8 Hz, 1

H, OCH_2), 4.40 (d, ^2J = 9.8 Hz, 1 H, OCH_2), 4.17 (d, ^2J = 12.8 Hz, 1 H, NCH_2), 4.05 (d, ^2J = 12.8 Hz, 1 H, NCH_2), 4.00 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.84 (dd, $J_{H5,H6a}$ = 2.2 Hz, $J_{H6a,H6b}$ = 12.2 Hz, 1 H, H-6a), 3.72-3.60 (m, 2 H, H-6b, H-2), 3.40-3.30 (m, 2 H, H-3, H-4), 3.20-3.10 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 138.4, 137.5, 131.3, 130.6, 129.3, 129.3, 129.3, 128.5, 93.4, 79.5, 79.5, 78.0, 71.6, 71.0, 62.6, 58.1; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{20}H_{25}NaNO_6$, 398.1575, found 398.1569.

3.4.3.5.5. Synthesis of *N*-(*N*-methoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (24).



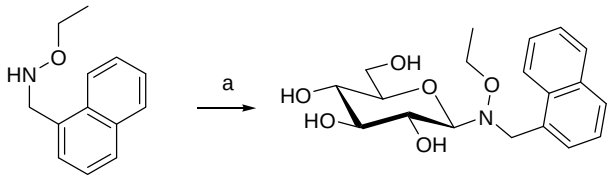
128 **24**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-methoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (24).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **128** (82 mg, 0.44 mmol) yielded **24** (113 mg, 74% yield) as a colorless syrup. TLC R_f = 0.28 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 8.39 (d, $J_{Ph,Ph}$ = 8.4 Hz, 1 H, Ph), 7.81 (dd, $J_{Ph,Ph}$ = 8.0 Hz, $J_{Ph,Ph}$ = 15.6 Hz, 2 H, Ph), 7.59-7.37 (m, 4 H, Ph), 4.64 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 4.53 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 3.97 (dd, $J_{H5,H6a}$ = 2.4 Hz, $J_{H6a,H6b}$ = 12.2 Hz, 1 H, H-6a), 3.96 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.79 (dd, $J_{H5,H6b}$ = 5.6 Hz, 1 H, H-6b), 3.60 (t, $J_{H2,H3}$ = 8.8 Hz, 1 Hz, H-2), 3.38 (s, 3 H, OCH_3), 3.42-3.27 (m, 2 H, H-3, H-4), 3.22-3.14 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ

135.2, 133.9, 133.7, 130.0, 129.5, 129.4, 127.0, 126.6, 126.3, 125.8, 93.0, 79.6, 79.4, 71.5, 71.2, 62.8, 62.5, 55.2; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{23}NaNO_6$, 372.1418, found 372.1420.

3.4.3.5.6. Synthesis of *N*-(*N*-ethoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (25).

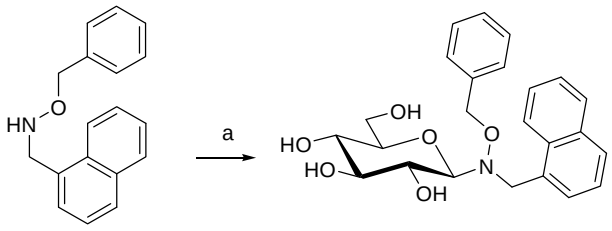


130 **25**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-ethoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (25).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **130** (77 mg, 0.38 mmol) yielded **25** (71 mg, 52% yield) as a colorless syrup. TLC R_f = 0.33 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 8.38 (d, $J_{Ph,Ph}$ = 8.4 Hz, 1 H, Ph), 7.81 (dd, $J_{Ph,Ph}$ = 8.2 Hz, $J_{Ph,Ph}$ = 14.6 Hz, 2H, Ph), 7.59-7.38 (m, 4 H, Ph), 4.62 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 4.53 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 3.97 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.96 (dd, $J_{H5,H6a}$ = 2.4 Hz, $J_{H6a,H6b}$ = 12.2 Hz, 1 H, H-6a), 3.79 (dd, $J_{H5,H6b}$ = 5.2 Hz, 1 H, H-6b), 3.83-3.75 (m, 1 H, OCH_2), 3.58 (t, J = 9.0 Hz, 1 H, H-2), 3.54-3.44 (m, 1 H, OCH_2), 3.42-3.27 (m, 2 H, H-3, H-4), 3.20-3.13 (m, 1 H, H-5), 0.92 (t, ^3J = 7.0 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, CD_3OD) δ 135.2, 134.0, 134.0, 130.0, 129.4, 129.4, 127.0, 126.6, 126.2, 125.8, 93.2, 79.6, 79.5, 71.6, 71.1, 70.7, 62.8, 55.6, 14.1; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{19}H_{25}NaNO_6$, 386.1575, found 386.1585.

3.4.3.5.7. Synthesis of *N*-(*N*-benzoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (26).

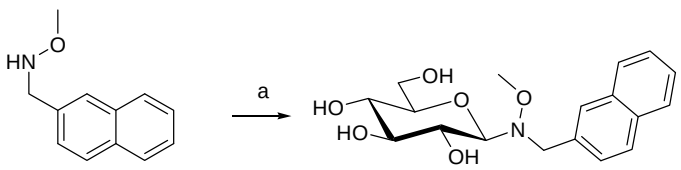


132 **26**

a) D-glucose, AcOH, MeOH, 40 °C, 96h.

***N*-(*N*-benzoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (26).** According to general procedure **3.4.3.5.1**, a reaction time of 96 hours with **132** (88 mg, 0.33 mmol) yielded **26** (69 mg, 49% yield) as a colorless syrup. TLC R_f = 0.40 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 8.36-8.24 (m, 1 H, Ph), 7.90-7.77 (m, 2 H, Ph), 7.60-7.38 (m, 4 H, Ph), 7.30-7.18 (m, 3 H, Ph), 7.14-7.04 (m, 2 H, Ph), 4.67 (d, ^2J = 12.4 Hz, 1 H, NCH_2), 4.62-4.50 (m, 2 H, NCH_2, OCH_2), 4.33 (d, ^2J = 9.6 Hz, 1 H, OCH_2), 4.03 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.98-3.88 (m, 1 H, H-6a), 3.80-3.64 (m, 2 H, H-2, H-6), 3.42-3.28 (m, 2 H, H-3, H-4), 3.23-3.12 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 137.7, 135.3, 134.1, 134.0, 130.7, 130.3, 129.6, 129.4, 129.3, 129.3, 127.1, 126.7, 126.3, 126.1, 93.4, 79.65, 79.57, 77.9, 71.7, 71.1, 62.7, 55.9; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{24}H_{27}NaNO_6$, 448.1731, found 448.1746.

3.4.3.5.8. Synthesis of *N*-(*N*-methoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (27).

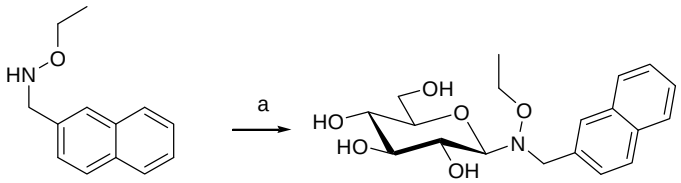


134 **27**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-methoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (27).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **134** (78 mg, 0.42 mmol) yielded **27** (49 mg, 33% yield) as a white solid. TLC R_f = 0.23 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.78-7.68 (m, 4 H, Ph), 7.55-7.49 (m, 1 H, Ph), 7.38-7.30 (m, 2 H, Ph), 4.23 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 4.11 (d, ^2J = 12.6 Hz, 1 H, NCH_2), 3.88 (d, $J_{H1,H2}$ = 9.2 Hz, 1 H, H-1), 3.83 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.0 Hz, 1 H, H-6a), 3.64 (dd, $J_{H5,H6b}$ = 5.6 Hz, 1 H, H-6b), 3.50-3.42 (m, 1 H, H-2), 3.26-3.18 (m, 2 H, H-3, H-4), 3.12-3.04 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 136.0, 134.8, 134.4, 134.0, 129.8, 129.0, 128.8, 128.6, 127.0, 126.8, 93.4, 79.8, 79.5, 71.6, 71.3, 62.9, 62.6, 57.8; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{23}NaNO_6$, 372.1418, found 372.1408.

3.4.3.5.9. Synthesis of *N*-(*N*-ethoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (28).

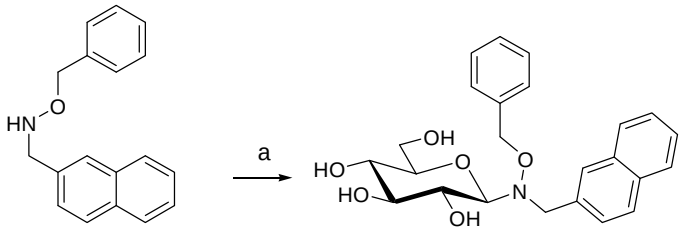


136 **28**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-ethoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (28).** According to general procedure **3.4.3.5.1**, a reaction time of 36 hours with **136** (90 mg, 0.45 mmol) yielded **28** (99 mg, 61% yield) as a white solid. TLC R_f = 0.30 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.84-7.74 (m, 4 H, Ph), 7.55-7.60 (m, 1 H, Ph), 7.44-7.36 (m, 2 H, Ph), 4.29 (d, ^2J = 12.8 Hz, 1 H, NCH_2), 4.18 (d, ^2J = 12.8 Hz, 1 H, NCH_2), 3.96 (d, $J_{H1,H2}$ = 9.2 Hz, 1 H, H-1), 3.90 (dd, $J_{H5,H6a}$ = 2.0 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.73 (dd, $J_{H5,H6b}$ = 5.4 Hz, 1 H, H-6b), 3.70-3.62 (m, 1 H, OCH_2), 3.58-3.46 (m, 2 H, OCH_2, H-2), 3.36-3.25 (m, 2 H, H-3, H-4), 3.20-3.12 (m, 1 H, H-5), 0.96-0.88 (m, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, CD_3OD) δ 136.1, 134.8, 134.3, 129.8, 129.1, 128.76, 128.75, 128.6, 127.0, 126.8, 93.5, 79.6, 79.5, 71.6, 71.2, 70.9, 62.8, 58.1, 14.0; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{19}H_{25}NaNO_6$, 386.1575, found 386.1579.

3.4.3.5.10. Synthesis of *N*-(*N*-benzoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosyl-amine (29).



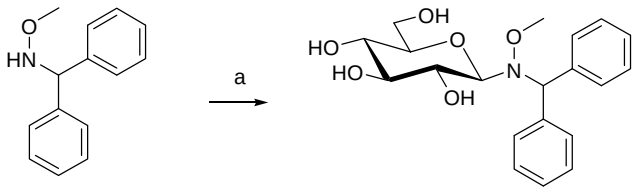
138 **29**

a) D-glucose, AcOH, MeOH, 40 °C, 36h.

***N*-(*N*-benzoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (29).** According to

general procedure **3.4.3.5.1**, a reaction time of 36 hours with **138** (68 mg, 0.26 mmol) yielded **29** (88 mg, 80% yield) as a white solid. TLC R_f = 0.40 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, DMSO-d_6) δ 7.98-7.88 (m, 4 H, Ph), 7.66-7.62 (m, 1 H, Ph), 7.57-7.49 (m, 2 H, Ph), 7.32-7.24 (m, 3 H, Ph), 7.18-7.12 (m, 2 H, Ph), 4.63 (d, ^2J = 10.0 Hz, 1 H, OCH_2), 4.43 (d, ^2J = 10.0 Hz, 1 H, OCH_2), 4.26 (s, 2 H, NCH_2), 3.89 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.82-3.74 (m, 1 H, H-6a), 3.56-3.46 (m, 2 H, H-2, H-6b), 3.22-3.14 (m, 1 H, H-3), 3.10-3.02 (m, 2 H, H-4, H-5); ^{13}C NMR (100 MHz, DMSO-d_6) δ 146.7, 144.9, 142.8, 142.2, 138.8, 138.4, 138.2, 138.1, 137.8, 137.5, 137.4, 136.0, 135.8, 101.8, 88.5, 87.5, 87.9, 85.6, 80.0, 79.9, 71.2, 66.2, HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{24}H_{27}NaNO_6$, 448.1731, found 448.1726.

3.4.3.5.11. Synthesis of *N*-(*N*-benzhydryl-*N*-methoxy)-β-D-glucopyranosylamine (30).



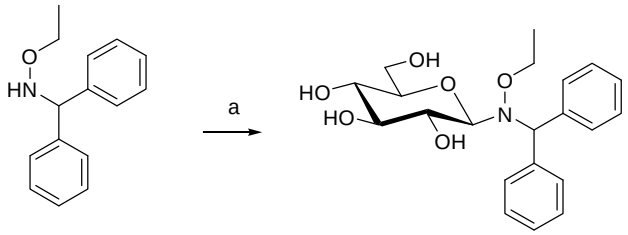
140 **30**

a) D-glucose, AcOH, MeOH, 40 °C, 10 days.

***N*-(*N*-benzhydryl-*N*-methoxy)-β-D-glucopyranosylamine (30).** According to general procedure **3.4.3.5.1**, a reaction time of 10 days with **140** (113 mg, 0.53 mmol) yielded **30** (92 mg, 46% yield) as a yellow oil. TLC R_f = 0.32 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.65 (d, $J_{Ph,Ph}$ =7.6 Hz, 2 H, Ph), 7.55 (d, $J_{Ph,Ph}$ =7.6 Hz, 2 H, Ph), 7.44-7.16 (m, 6 H, Ph), 5.44 (s, 1 H, NCH), 3.88-3.83 (m, 1 H, H-6a), 3.81 (d, $J_{H1,H2}$ = 9.6 Hz, 1 H, H-1), 3.73-3.64 (m, 1 H, H-

6b), 3.63-3.55 (m, 1 H, H-2), 3.39 (s, 3 H, OCH_3), 3.32-3.25 (m, 1 H, H-4), 3.15 (t, J = 9.0 Hz, 1 H, H-3), 2.76-2.70 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 145.5, 142.1, 129.9, 129.7, 129.6, 129.1, 128.6, 128.2, 91.8, 79.74, 79.71, 73.3, 71.7, 71.1, 64.2, 62.8; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{20}H_{25}NaNO_6$, 398.1575, found 398.1584.

3.4.3.5.12. Synthesis of *N*-(*N*-benzhydryl-*N*-ethoxy)-β-D-glucopyranosylamine (31).

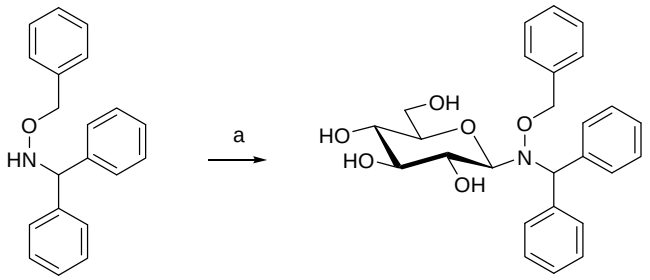


142 **31**

a) D-glucose, AcOH, MeOH, 40 °C, 7 days.

***N*-(*N*-benzhydryl-*N*-ethoxy)-β-D-glucopyranosylamine (31).** According to general procedure **3.4.3.5.1**, a reaction time of 7 days with **142** (29 mg, 0.13 mmol) yielded **31** (15 mg, 30% yield) as a yellow oil. TLC R_f = 0.37 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.64 (d, $J_{Ph,Ph}$ = 7.6 Hz, 2 H, Ph), 7.56 (d, $J_{Ph,Ph}$ = 7.6 Hz, 2 H, Ph), 7.32-7.25 (m, 4 H, Ph), 7.24-7.16 (m, 2 H, Ph), 5.44 (s, 1 H, NCH), 3.92-3.81 (m, 2 H, OCH_2, H-1), 3.70 (dd, $J_{H5,H6a}$ = 4.6 Hz, $J_{H6a,H6b}$ = 12.2 Hz, 1 H, H-6a), 3.56 (t, J = 8.8 Hz, 1 H, H-2), 3.52-3.42 (m, 1 H, OCH_2), 3.36-3.27 (m, 1 H, H-4, H-6b), 3.16 (t, J = 8.8 Hz, 1 H, H-3), 2.75-2.69 (m, 1 H, H-5), 0.70 (t, ^3J = 7.2 Hz, 3 H, OCH_2CH_3); ^{13}C NMR (100 MHz, CD_3OD) δ 143.9, 142.1, 130.0, 129.6, 129.0, 128.6, 128.5, 128.2, 91.7, 79.7, 79.6, 73.3, 72.4, 71.7, 70.9, 62.6, 13.7; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{21}H_{27}NaNO_6$, 412.1731, found 412.1731.

3.4.3.5.13. Synthesis of *N*-(*N*-benzhydryl-*N*-benzoxy)-β-D-glucopyranosylamine (32).



144 **32**

a) D-glucose, AcOH, MeOH, 40 °C, 7 days.

***N*-(*N*-benzhydryl-*N*-benzoxy)-β-D-glucopyranosylamine (32).** According to general procedure **3.4.3.5.1**, a reaction time of 7 days with **144** (115 mg, 0.40 mmol) yielded **32** (57 mg, 32% yield) as a yellow solid. TLC R_f = 0.49 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, CD_3OD) δ 7.78-7.73 (m, 2 H, Ph), 7.65-7.60 (m, 2 H, Ph), 7.46-7.10 (m, 9 H, Ph), 7.76-7.70 (m, 2 H, Ph), 5.49 (s, 1 H, NCH), 4.78 (d, ^2J = 9.0 Hz, 1 H, OCH_2), 4.43 (d, ^2J = 9.0 Hz, 1 H, OCH_2), 3.94 (d, $J_{H1,H2}$ = 8.8 Hz, 1 H, H-1), 3.81 (dd, $J_{H5,H6a}$ = 2.2 Hz, $J_{H6a,H6b}$ = 12.2 Hz, 1 H, H-6a), 3.77 (t, J = 8.8 Hz, 1 H, H-2), 3.67 (dd, $J_{H5,H6b}$ = 2.2 Hz, 1 H, H-6b), 3.38-3.29 (m, 1 H, H-4), 3.21 (t, J = 8.8 Hz, 1 H, H-3), 2.79-2.70 (m, 1 H, H-5); ^{13}C NMR (100 MHz, CD_3OD) δ 143.4, 142.1, 137.1, 130.6, 130.3, 129,7, 129.6, 129.5, 129.3, 129.2, 129.1, 128.6, 128.5, 91.7, 79.7, 79.5, 79.2, 73.3, 71.7, 70.7, 62.3; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{26}H_{29}NaNO_6$, 474.1888, found 474.1877.

3.4.4. β-D-glucoside screening.

3.4.4.1 *In vitro* β-D-glucoside screening with OleD variants. Reactions containing 2.1 μM (10 μg) of purified OleD variant, 1 mM of UDP or TDP, and 1 mM of β-D-glucopyranoside (**1**-**32**) in Tris-HCl (50 mM, pH 8.5) with a final volume of 100 μl were incubated at room temperature for 1 hour. Samples were frozen in a bath of dry ice and acetone and stored at -20 °C. Following, samples were thawed at 4 °C and filtered through a MultiScreen Filter Plate (from Millipore, Billerica, MA, USA) according to manufacturer's instructions and evaluated for formation of UDP- (**33a**) or TDP-α-D-glucose (**33b**) by analytical reverse-phase HPLC with a 250 mm x 4.6 mm Gemini-NX 5μ C18 column (Phenomenex, Torrance, CA, USA) using a linear gradient of 0% to 15% CH_3CN (solvent B) over 15 minutes (solvent A = aqueous 50 mM triethylammonium acetate buffer [from Sigma-Aldrich, St. Louis, MO, USA], flow rate = 1 ml min^{-1}, with detection monitored at 254 nm).

3.4.4.2 pH optimization of reverse reaction All reactions were performed in a final volume of 500 μl dH_2O buffered with 50 mM MES (pH 6.0 or 6.5) 50 mM MOPS (pH 6.5 or 7.0) or 50 mM Tris-HCl (pH 7.0, 7.5, 8.0, 8.5 or 9.0) with 0.21 μM (5 μg) OleD variant TDP-16, 0.05 mM of 2-chloro-4-nitrophenyl-β-D-glucoside (**9**) as donor and 0.05 mM of UDP as acceptor. Absorbance measurements were taken at t = 0 and 60 minutes and Δ410 nm was calculated. Rates of 2-chloro-4-nitrophenolate production were then calculated by comparing them against standard curves at the corresponding pH and buffer. Collectively, the standard deviation in rates from pH 7.0 to pH 8.5 with Tris-HCl as buffer was <5%. Rates dropped sharply outside of this pH range or with change in buffer. These observations are consistent with those previously reported for forward reactions with wild-type OleD[10].

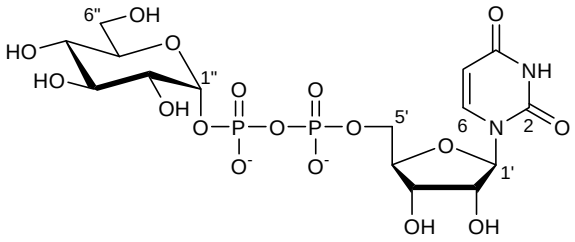
3.4.4.3. NDP screening with 2-chloro-4-nitrophenyl glucoside (**9**) All reactions were performed in a final volume of 200 µl Tris-HCl buffer (50 mM, pH 8.5) with 2.1 µM (20 µg) OleD variant TDP-16, 1 mM of 2-chloro-4-nitrophenyl-β-D-glucoside (**9**) as donor and 1 mM of either UDP, TDP, CDP, ADP or GDP as acceptor. Reactions were allowed to proceed for 5 hours, quenched with an equal volume of 40 mM phosphoric acid (adjusted to pH 6.5 with triethylamine) and heated for 30 seconds on a heat block. Following, samples were centrifuged at 10,000 g for 30 min at 0 °C and the supernatant removed for analysis. The clarified reaction mixtures were analyzed by analytical reverse-phase HPLC. HPLC was conducted with a Supelcosil LC18-T (3 µm, 150 x 4.6 mm) column (Sigma-Aldrich, St. Louis, MO, USA) with a gradient of 0% B to 100% B over 20 min (A = 40 mM phosphoric acid [adjusted to pH 6.5 with triethylamine]; B = 10% MeOH in 40 mM phosphoric acid [adjusted to pH 6.5 with triethylamine]; flow rate = 0.5 mL min^{-1}) and detection monitored at 254 nm.

3.4.5. (U/T)DP-α-D-glucose (33a-b) scale-up/characterization.

3.4.5.1. General Reaction Procedure. Reactions were conducted at 25 °C in 2 mL Tris-HCl (50 mM, pH 8.5) with 2-chloro-4-nitrophenyl β-D-glucopyranoside (**9**), either UDP or TDP, and 4.2 µM (~400 µg) of OleD variant P67T/S312F/A242L/Q268V. At 6 hours, 4 mL of 50 mM Tris-HCl (50 mM, pH 7.0) and 5 µL of alkaline phosphatase (100 U, Roche) were added. At 7.5 hours, the reaction was passed through a 10K MWCO filter, frozen at -80 °C, and lyophilized. The dried reaction was dissolved in 2 mL of ddH_2O and the desired product purified by semi-preparative HPLC with a Supelcosil LC18, 5 µm, 25 cm x 10 mm column (Supelco) using a gradient of 0% to 12.5% CH_3CN (solvent B) over 12.5 min, 12.5% to 90% B over 1 min, 90% B

for 5 min (A = 50 mM PO_4^{-2}, 5 mM tetrabutylammonium bisulfate, 2% acetonitrile, pH adjusted to 6.0 with KOH; flow rate = 5 mL min^{-1}; A_{254} nm). Desired fractions were concentrated under reduced pressure, frozen at -80 °C, and lyophilized. The resulting products were dissolved in 2 mL of ddH_2O and purified by semi-preparative HPLC with the column mentioned above using linear gradient of 0% B to 10% B over 10 min (A = 50 mM triethylammonium acetate buffer; B = acetonitrile; flow rate = 5 mL min^{-1}; A_{254} nm). The desired fractions were concentrated via reduced pressure, frozen at -80 °C, and lyophilized multiple times. Products were confirmed by mass spectrometry and via ^{1}H, ^{13}C, and ^{31}P NMR using a Varian ^{UNITY}INOVA 500 MHz instrument (Palo Alto, CA, USA) with a Nalorac qn6121 probe (Martinez, CA, USA). Assignments were aided with gCOSY and gHSQC methods. See **Appendix 2** for ^{1}H, ^{13}C, and ^{31}P spectra.

3.4.5.2. Synthesis of uridine 5'-diphosphate α-D-glucopyranoside (33a).

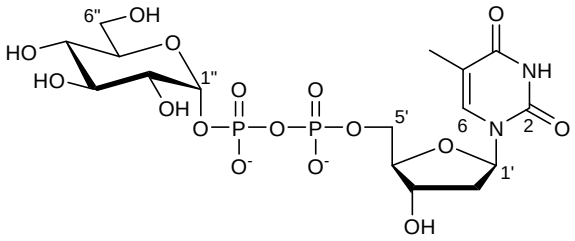


33a

UDP (9.0 mg, 0.022 mmol) and 8.9 mg of **9** (0.025 mmol) in the above method yielded 6.9 mg of **33a** (0.012 mmol, 55% isolated yield). ^{1}H NMR (500 MHz, D_2O) δ 7.97 (d, $J_{H\text{-}6,H\text{-}5}$ = 8.1 Hz, 1 H, H-6), 6.05-5.90 (m, 2 H, H-1', H-5), 5.61 (dd, $J_{H\text{-}1",H\text{-}2"}$ = 3.5 Hz, $J_{H\text{-}1",P}$ = 7.2 Hz, 1 H, H-1"), 4.38 (m, 2 H, H-2', H-3'), 4.31-4.28 (m, 1 H, H-4'), 4.27-4.20 (m, 2 H, H-5a', H-5b'), 3.91 (ddd, $J_{H\text{-}5",H\text{-}6a"}$ = 2.2 Hz, $J_{H\text{-}5",H\text{-}6b"}$ = 4.3 Hz, $J_{H\text{-}5",H\text{-}4"}$ = 9.9 Hz, 1 H, H-5"), 3.87 (dd, $J_{H\text{-}6a",H\text{-}5"}$ = 2.2

Hz, $J_{H\text{-}6a'',H\text{-}6b''}$ = 12.5 Hz, 1 H, H-6a''), 3.81-3.75 (m, 2 H, H-3'', H-6b''), 3.56-3.51 (m, 1 H, H-2''), 3.47 (dd, $J_{H\text{-}4'',H\text{-}3''}$ = 9.9 Hz, $J_{H\text{-}4'',H\text{-}5''}$ = 9.9 Hz, 1 H, H-4''); ^{13}C NMR (126 MHz, D_2O) δ 167.2 (C-4), 152.7 (C-2), 142.5 (C-6), 103.5 (C-5), 96.4 ($J_{C\text{-}1'',P}$ = 6.7 Hz, C-1''), 89.2 (C-1'), 84.1 ($J_{C\text{-}4',P}$ = 9.2 Hz, C-4'), 74.6 (C-2'), 73.7 (C-5''), 73.6 (C-3''), 72.5 ($J_{C\text{-}2'',P}$ = 8.5 Hz, C-2''), 70.5 (C-3'), 70.0 (C-4'), 65.8 ($J_{C\text{-}5',P}$ = 5.5 Hz, C-5'), 61.2 (C-6''); ^{31}P NMR (202 MHz, D_2O) δ -10.0 (d, $J_{P,P}$ = 20.7 Hz), -11.7 (d, $J_{P,P}$ = 20.7 Hz); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{15}H_{22}N_2NaO_{17}P_2$ 587.02969, found 587.02954; spectral data are consistent with those reported by Bae *et al*.[75].

3.4.5.3. Synthesis of thymidine 5'-diphosphate α-D-glucopyranoside (33b).



33b

TDP (9.0 mg, 0.022 mmol) and 9.0 mg of **9** (0.025 mmol) in the above method yielded 7.7 mg of **33b** (0.013 mmol, 61% isolated yield). ^{1}H NMR (500 MHz, D_2O) δ 7.75 (d, $J_{H\text{-}6,5\text{-}CH3}$ = 1.0 Hz, 1 H, H-6), 6.36 (t, $J_{H\text{-}1',H\text{-}2's}$ = 7.0 Hz, 1 H, H-1'), 5.61 (dd, $J_{H\text{-}1'',H\text{-}2''}$ = 3.5 Hz, $J_{H\text{-}1'',P}$ = 7.2 Hz, 1 H, H-1''), 4.65-4.62 (m, 1 H, H-3'), 4.20-4.17 (m, 3 H, H-4', H-5a', H-5b'), 3.91 (ddd, $J_{H\text{-}5'',H\text{-}6a''}$ = 2.3 Hz, $J_{H\text{-}5'',H\text{-}6b''}$ = 4.4 Hz, $J_{H\text{-}5'',H\text{-}4''}$ = 9.8 Hz, 1 H, H-5''), 3.87 (dd, $J_{H\text{-}6a'',H\text{-}5''}$ = 2.3 Hz, $J_{H\text{-}6a'',H\text{-}6b''}$ = 12.4 Hz, 1 H, H-6a''), 3.81-3.75 (m, 2 H, H-3'', H-6b''), 3.56-3.50 (m, 1 H, H-2''), 3.47 (dd, $J_{H\text{-}4'',H\text{-}3''}$ = 9.8 Hz, $J_{H\text{-}4'',H\text{-}5''}$ = 9.8 Hz, 1 H, H-4''), 2.41-2.35 (m, 2 H, H-2a', H-2b'), 1.94 (d, $J_{5\text{-}CH_3,H\text{-}6}$ = 1.0 Hz, 1 H, 5-CH_3); ^{13}C NMR (126 MHz, D_2O) δ 167.2 (C-4), 152.4 (C-2), 138.0 (C-6),

112.4 (C-5), 96.3 ($J_{C\text{-}1'',P}$ = 6.7 Hz, C-1''), 86.1 ($J_{C\text{-}4',P}$ = 9.1 Hz, C-4'), 85.7 (C-1'), 73.6 (C-3''), 73.5 (C-5''), 72.4 ($J_{C\text{-}2'',P}$ = 8.6 Hz, C-2''), 71.8 (C-3'), 69.9 (C-4''), 66.2 ($J_{C\text{-}5',P}$ = 5.7 Hz, C-5'), 61.1 (C-6''), 39.4 (C-2'), 12.5 (5-CH_3); ^{31}P NMR (202 MHz, D_2O) δ -10.18 (d, $J_{P,P}$ = 20.9 Hz), -11.72 (d, $J_{P,P}$ = 20.9 Hz); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{16}H_{24}N_2NaO_{16}P_2$ 585.05042, found 585.05105; spectral data are consistent with those reported by Bae *et al*.[75].

3.4.6. Determination of kinetic parameters. Assays were performed in a final volume of 500 µL of 50 mM Tris-HCl (pH 8.5) using 0.42 µM (10 µg) of enzyme (either wild-type or TDP-16). Reactions were prepared with either UDP (2.5 mM for wild-type, 1.0 mM for variant OleD), TDP (2.5 mM for wild-type, 2.0 mM for variant OleD), or **9** (20 mM) saturating and the corresponding reactant varied (from 0 mM until saturation conditions or solubility limits were met). Reactions were followed at 410 nm on a DU800 spectrophotometer (Beckman Coulter, Brea, CA, USA) where the rate of 2-chloro-4-nitrophenolate formation was determined to be linear (<10 min). Initial rates were converted to product formation per unit time by comparing values to a standard curve. All experiments were performed in triplicate. Initial velocities were fit to the Michaelis-Menten equation using Origin Pro 7.0 software. OleD wild-type enzyme could not be saturated with donor **9** due to limited solubility (~25 mM under the stated conditions). Consequentially, k_{cat}/K_m for wild-type was determined by linear regression. Values obtained are in agreement with kinetic parameters from previous studies with OleD wildtype and numerous variants[13, 16, 17].

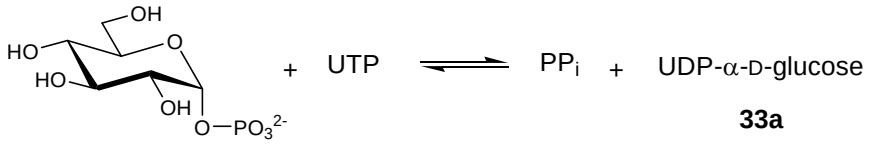
3.4.7.1. Determination of Equilibrium Constants. Reactions contained 21 µM (200 µg) OleD variant P67T/S312F/A242L/Q268V, 1 mM UDP, and 1 mM β-D-glucopyranoside donor (**1**, **2**, **4**,

7, or **9**) in Tris-HCl buffer (50 mM, pH 8.5) in a final total volume of 200 µl. Multiple time course evaluations were conducted to determine the time at which each reaction reached equilibrium (<2 min for **7**, **9**; <90 min for **1**, **4**; 200 min for **2**). Following, each analysis was conducted in triplicate to the minimum observed time point for equilibrium and samples were processed and evaluated as described below to determine concentrations of UDP to **33a** (UDP-α-D-glucose). HPLC conditions consisted of a Gemini-NX C-18 (5 µm, 250 x 4.6 mm) column (from Phenomenex, Torrance, California, USA) with a gradient of 0% B to 20% B over 20 min, 20% B to 80% B over 1 min, 80% B for 6 min (A = 50 mM triethylammonium acetate buffer; B = acetonitrile; flow rate = 1 mL min^{-1}), and detection monitored at 254 nm. Glucoside and aglycon concentrations were inferred from the determined concentrations of UDP and **33a**. $K_{eq,pH8.5}$ was determined according to the following equation:

$$K_{eq,pH8.5} = ([\text{UDP-glucose}][\text{aglycon}]) / ([\text{UDP}][\text{glucoside}])$$

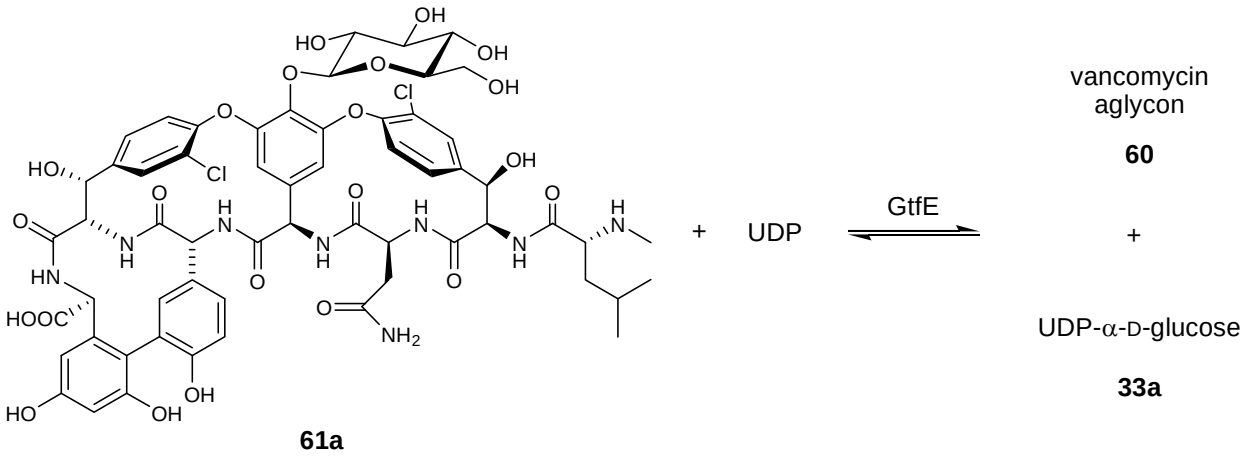
Related examples for equilibrium determinations via identical methodology include:

i) UDP-glucose pyrophosphorylase (which catalyzes the following reaction)[76]:



33a

ii) Glucosyltransferase GtfE, which catalyzes the following reaction[5]:



3.4.8. Syntheses of *2*-chloro-4-nitrophenyl glycosides. (Syntheses and characterization by R.W. Gantt and Dr. P. Peltier-Pain)

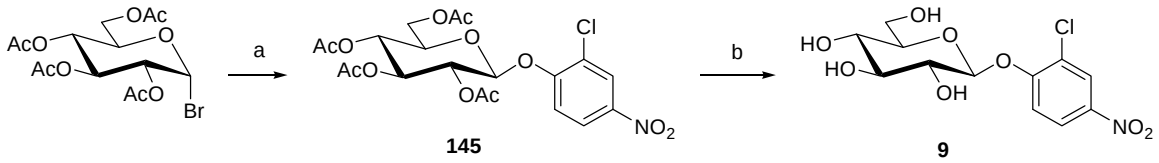
3.4.8.1. General procedure for bromination. Per *O*-acetylated glycopyranose (0.50 mmol) was dissolved in CH_2Cl_2 (1mL) and treated with a 33% solution of HBr in glacial acetic acid (1 mL) at 0 °C for 30 min. The reaction was subsequently allowed to warm to room temperature and the stirring was continued until no starting compound was detected by TLC. Following, the mixture was diluted with CH_2Cl_2 (50 mL), washed with $NaHCO_3$ sat solution (3 x 25 mL) and with brine (25mL). The organic phase was dried over $MgSO_4$, and the solvent removed under reduced pressure. The residue obtained was used directly without purification.

3.4.8.2. General procedure for phase transfer catalyzed glycosylation of *2*-chloro-4-nitrophenol. Per *O*-acetylated glycopyranosyl bromide was dissolved in CH_2Cl_2 to a final concentration of 125 mM. Were successively added 1.5 equiv. of tetrabutylammonium bromide and 3 equiv. of 2-chloro-4-nitrophenol. An equal volume of 1M NaOH solution was added at

0 °C and the reaction mixture was stirred vigorously at room temperature overnight. After dilution with 2.5 volumes of EtOAc, the organic phase was washed three times with 0.2 volumes of a 1M NaOH solution and finally with 0.2 volumes of brine. The organic phase was dried over $MgSO_4$, and the solvent removed under reduced pressure. Purification was carried out by chromatography on silica gel.

3.4.8.3. General procedure for deacetylation. The acetylated glycoside (0.10 mmol) was dissolved in dry MeOH (2 mL) and treated at room temperature with a 0.1 M solution of sodium methoxide (150 µL). The mixture was stirred until no starting compound was detected by silica gel TLC with 9:1 (v/v) CH_2Cl_2/MeOH. Neutralization was then performed by adding Amberlite IR-120 (H^+ form). The resin was removed via filtration and the solvent removed under reduced pressure. Purification was carried out by chromatography on silica gel.

3.4.8.4. Synthesis of 2-chloro-4-nitrophenyl-β-D-glucopyranoside (9).



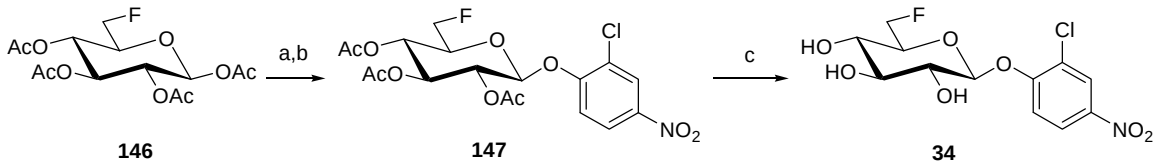
a) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; b) NaOMe 0.1M in MeOH, RT, 18h.

(2-chloro-4-nitrophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranoside (145). This compound was prepared from 2,3,4,6 tetra-*O*-acetyl-α-D-glucopyranosyl bromide (1g, 2.4 mmol) according

the general procedure **3.4.8.2**. The purification on silica gel (hexanes/EtOAc, 7:3) afforded **145** (1.08 mg, 89%) as a white powder. TLC R_f = 0.45 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.13 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.25 (d, 1 H, H-6'), 5.38 (dd, $J_{H1, H2}$ = 7.5 Hz, $J_{H2, H3}$ = 9.4 Hz, 1 H, H-2), 5.32 (dd, $J_{H3, H4}$ = 9.1 Hz, 1 H, H-3), 5.19 (dd, $J_{H4, H5}$ = 9.9 Hz, 1 H, H-4), 5.14 (d, 1 H, H-1), 4.27 (dd, $J_{H5, H6a}$ = 5.3 Hz, $J_{H6a, H6b}$ = 12.4 Hz, 1 H, H-6a), 4.20 (dd, $J_{H5, H6b}$ = 2.6 Hz, 1 H, H-6b), 3.92 (ddd, 1 H, H-5), 2.10, 2.09, 2.07, 2.06 (4s, 12 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.5, 170.3, 169.4, 169.2, (C=O), 157.4 (C-1'), 143.4 (C-4'), 126.4 (C-3'), 125.2(C-2'), 123.7 (C-5'), 116.6 (C-6'), 99.4 (C-1), 72.7 (C-5), 72.2 (C-3), 70.6 (C-2), 68.1(C-4), 61.9 (C-6), 20.8, 20.7, 20.7, 20.7 (*C*H_3CO).; HRMS-ESI (*m/z*): $[M+NH_4]^+$ calcd for $C_{20}H_{22}ClNO_{12}$, 521.1169, found 521.1158.

2-chloro-4-nitrophenyl-β-D-glucopyranoside (9). A solution of **145** (750 mg, 1.49 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **9** (448 mg, 90%) as a white powder. TLC R_f = 0.26 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.30 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.41 (d, 1 H, H-6'), 5.17 (d, $J_{H1, H2}$ = 7.6 Hz, 1 H, H-1), 3.89 (dd, $J_{H5, H6a}$ = 2.2 Hz, $J_{H6a, H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5, H6b}$ = 5.7 Hz, 1 H, H-6b), 3.58 (dd, $J_{H2, H3}$ = 8.8 Hz, 1 H, H-2), 3.55-3.50 (m, 1 H, H-5), 3.50 (dd, $J_{H3, H4}$ = 8.8 Hz, 1 H, H-3), 3.42 (dd, $J_{H4, H5}$ = 10.1 Hz, 1 H, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 159.4 (C-1'), 143.6 (C-4'), 126.7 (C-3'), 124.9(C-2'), 124.8 (C-5'), 116.9 (C-6'), 101.8 (C-1), 78.6 (C-5), 78.1 (C-3), 74.6 (C-2), 71.0 (C-4), 62.4 (C-6). HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNO_8$, 358.0301, found 358.0313.

3.4.8.4. Synthesis of (2-chloro-4-nitrophenyl)-6-deoxy-6-fluoro-β-D-glucopyranoside (34).



a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

1,2,3,4,-tetra-*O*-acetyl-6-deoxy-6-fluoro-β-D-glucopyranose (146). This compound was prepared as previously described[77] from 1,2,3,4-tetra-*O*-acetyl-β-D-glucopyranose (300 mg, 0.86mmol) in 83% yield (250 mg, 83%). TLC R_f = 0.45 (EtOAc:hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.73 (d, $J_{H1, H2}$ = 8.2 Hz, 1 H, H-1), 5.28 (dd, $J_{H2, H3}$ = 9.4 Hz, $J_{H3, H4}$ = 9.3 Hz, 1 H, H-3), 5.18-5.09 (m, 2 H, H-2,H-4), 4.48 (dddd, $J_{H5, H6a}$ = 2.4 Hz, $J_{H5, H6b}$ = 4.1 Hz, $J_{H6a, H6b}$ = 10.6 Hz, $J_{H6, F}$ = 46.9 Hz, 2 H, H-6a, H-6b), 3.83 (dddd, $J_{H4, H5}$ = 6.5 Hz, $J_{H5, F}$ = 22.4 Hz, 1 H, H-5), 2.11, 2.05, 2.03, 2.02 (4s, 12 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.2, 169.4, 169.3, 169.1 (C=O), 91.7 (C-1), 80.2 (d, $J_{C6, F}$ = 176.6 Hz, C-6), 73.5 (d, $J_{C5, F}$ = 19.5 Hz C-5), 72.9 (C-3), 70.3 (C-2), 67.7 (d, $J_{C4, F}$ = 6.5 Hz, C-4), 20.9, 20.7, 20.7 (*C*H_3CO).

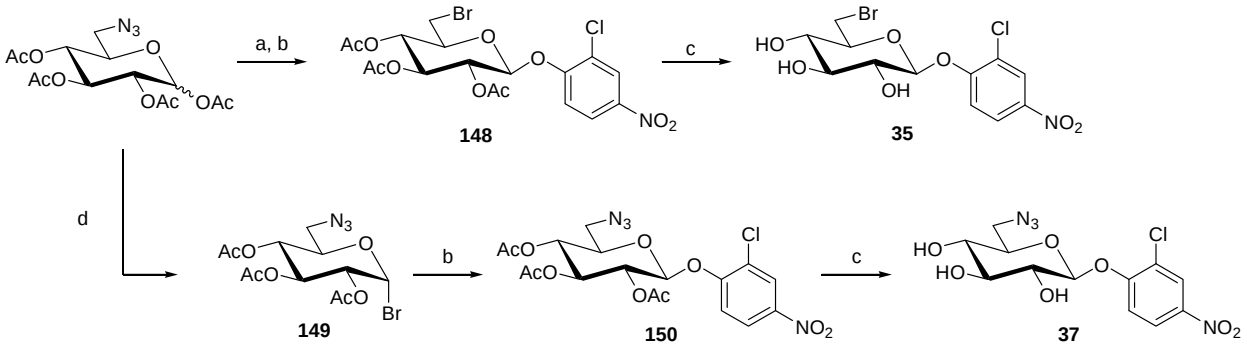
(2-chloro-4-nitrophenyl)-2,3,4-tri-*O*-acetyl-6-deoxy-6-fluoro-β-D-glucopyranoside (147). A solution of **146** (60 mg, 0.17 mmol) was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following the general procedure **3.4.8.2**. The purification on silica gel (hexanes/EtOAc 7:3) afforded **147** (41mg, 52%) as a white powder. TLC R_f = 0.43 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.29 (d,

$J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.13 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.28 (d, 1 H, H-6'), 5.40-5.32 (m, 2 H, H-2, H-3), 5.21 (d, $J_{H1, H2}$ = 7.2 Hz, 1 H, H-1), 5.12 (dd, $J_{H3, H4}$ = 9.8 Hz, $J_{H4, H5}$ = 9.8 Hz, 1 H, H-4), 4.58-4.45 (m, 2 H, H-6a, H-6b), 3.99 (dddd, $J_{H5, H6a}$ = 1.8 Hz, $J_{H5, H6b}$ = 4.1 Hz, $J_{H5, F}$ = 23.2 Hz, 1 H, H-5), 2.08, 2.07, 2.05 (3s, 9 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.4, 169.3, 169.0 (C=O), 157.2 (C-1'), 143.2 (C-4'), 126.1 (C-3'), 124.9 (C-2'), 123.7 (C-5'), 116.6 (C-6'), 99.2 (C-1), 80.5 (d, $J_{C6, F}$ = 176.1 Hz, C-6), 73.3 (d, $J_{C5, F}$ = 19.9 Hz C-5), 71.9 (C-3), 70.4 (C-2), 67.5 (d, $J_{C4, F}$ = 7.0 Hz, C-4), 20.5, 20.5, 20.5 (CH_3CO).; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{19}ClFNO_{10}$, 486.05737, found 486.05697.

(2-chloro-4-nitrophenyl)-6-deoxy-6-fluoro-β-D-glucopyranoside (34). A solution of **147** (38 mg, 0.10 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **147** (26 mg, 93%) as a white powder. TLC R_f = 0.40 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.31 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.17 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.39 (d, 1 H, H-6'), 5.20 (d, $J_{H1,H2}$ = 7.4 Hz, 1 H, H-1), 4.64 (dddd, $J_{H5,H6a}$ = 1.8 Hz, $J_{H5,H6b}$ = 4.7 Hz, $J_{H6a,H6b}$ = 10.4 Hz, $J_{H6,F}$ = 47.7 Hz, 2 H, H-6a, H-6b), 3.73 (dddd, $J_{H4,H5}$ = 9.7 Hz, $J_{H5,F}$ = 24.2 Hz, 1 H, H-5), 3.57 (dd, $J_{H2,H3}$ = 8.9 Hz, 1 H, H-2), 3.51 (dd, $J_{H3,H4}$ = 9.4 Hz, 1 H, H-3), 3.45 (dd, 1 H, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 159.2 (C-1'), 143.7 (C-4'), 126.8 (C-3'), 124.9 (C-2'), 124.9 (C-5'), 116.7 (C-6'), 101.64 (C-1), 83.1 (d, $J_{C6, F}$ = 172.0 Hz, C-6), 77.9 (C-3), 76.8 (d, $J_{C5, F}$ = 18.1 Hz C-5), 74.5 (C-2), 69.9 (d, $J_{C4, F}$ = 7.0 Hz, C-4); ^{19}F NMR (376 MHz, CD_3OD) δ -236 (dt, $J_{H5,F}$ = 24.2 Hz, $J_{H6,F}$ = 47.7 Hz). HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{13}ClFNO_7$, 360.02568, found 360.02556.

3.4.8.5. Synthesis of (2-chloro-4-nitrophenyl)-6-bromo-6-deoxy-β-D-glucopyranoside (35)

and (2-chloro-4-nitrophenyl)-6-azido-6-deoxy-β-D-glucopyranoside (37).



a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h; d) $TiBr_4$ in AcOEt/CH_2Cl_2, RT, 48h;

(2-chloro-4-nitrophenyl)-2,3,4-tri-*O*-acetyl-6-bromo-6-deoxy-β-D-glucopyranose (148). A solution of the previously described 1,2,3,4-tetra-*O*-acetyl-6-azido-6-deoxy-D-glucopyranose[78] (200 mg, 0.54 mmol) in CH_2Cl_2 (500 µL) was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following the general procedure **3.4.8.2**. The purification on silica gel (hexanes/EtOAc, 7:3) afforded **148** (135 mg, 48%) as a white powder. TLC R_f = 0.60 (EtOAc:hexanes, 5:5). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.28 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.14 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.44 (d, 1 H, H-6'), 5.38 (dd, $J_{H1, H2}$ = 7.6 Hz, $J_{H2, H3}$ = 9.5 Hz, 1 H, H-2), 5.31 (dd, $J_{H3, H4}$ = 9.0 Hz, 1 H, H-3), 5.13 (d, 1 H, H-1), 5.04 (dd, $J_{H4, H5}$ = 9.8 Hz, 1 H, H-4), 3.95-3.88 (m, 1 H, H-5), 3.50 (dd, $J_{H5, H6a}$ = 2.5 Hz, $J_{H6a, H6b}$ = 11.4 Hz, 1 H, H-6a), 3.41 (dd, $J_{H5, H6b}$ = 8.6 Hz, 1 H, H-6b), 2.08, 2.07, 2.04 (3s, 9 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.0, 169.4, 169.0 (C=O), 157.2 (C-1'), 143.3 (C-4'), 126.0 (C-3'), 124.8 (C-2'), 123.8 (C-5'), 116.9 (C-6'), 99.3 (C-1), 74.7 (C-5), 71.9

(C-3), 70.6 (C-2), 70.5 (C-4), 30.2 (C-6), 20.6, 20.6, 20.5 (CH_3CO).; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{19}BrClNO_{10}$, 545.9779, found 545.9759.

(2-chloro-4-nitrophenyl)-6-bromo-6-deoxy-β-D-glucopyranose (35). A solution of **148** (130 mg, 0.27 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **35** (90 mg, 85%) as a white powder. TLC R_f = 0.44 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.30 (d, $J_{H3',H5'}$ = 2.8 Hz, 1 H, H-3'), 8.17 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.47 (d, 1 H, H-6'), 5.18 (d, $J_{H1,H2}$ = 7.6 Hz, 1 H, H-1), 3.82 (dd, $J_{H5,H6a}$ = 2.1 Hz, $J_{H6a,H6b}$ = 11.0 Hz, 1 H, H-6a), 3.71 (ddd, $J_{H4,H5}$ = 9.5 Hz, $J_{H5,H6b}$ = 7.4 Hz, 1 H, H-5), 3.59 (dd, $J_{H2,H3}$ = 9.2 Hz, 1 H, H-2), 3.54 (dd, 1 H, H-6b), 3.50 (dd, $J_{H3,H4}$ = 9.1 Hz 1 H, H-3), 3.37 (dd, 1 H, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 159.2 (C-1'), 143.7 (C-4'), 126.7 (C-3'), 124.8 (C-2'), 124.7 (C-5'), 117.2 (C-6'), 101.7 (C-1), 77.7 (C-3), 77.4 (C-5), 74.6 (C-2), 73.4 (C-4), 33.3 (C-6); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{13}BrClNO_7$, 419.94561, found 419.94527.

1-bromo-2,3,4,-tri-*O*-acetyl-6-azido-6-deoxy-α-D-glucopyranose (149). This compound was prepared from the previously described 1,2,3,4-tetra-*O*-acetyl-6-azido-6-deoxy-D-glucopyranose[78] (200 mg, 0.54 mmol) which was dissolved in anhydrous CH_2Cl_2 (3.5 mL) and EtOAc (300 μL). Titanium tetrabromide (492 mg, 1.34 mmol) was added and the reaction was stirred at room temperature under nitrogen for 48 hours. The reaction was stopped by addition of NaOAc (200 mg) and stirred for 15 min, then filtered through Celite, and the Celite pad was washed with CH_2Cl_2 (20 mL). The organic filtrate was washed with cold water (15 mL) and the organic phase was then dried over $MgSO_4$, and the solvent removed under reduced pressure.

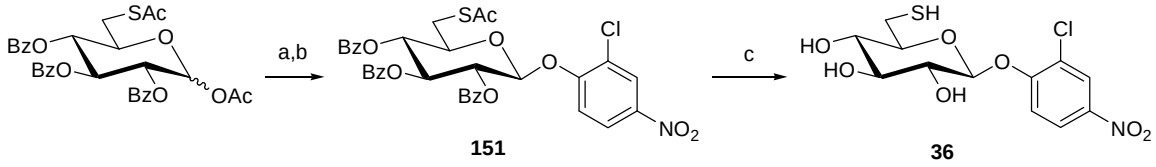
Purification was carried out by chromatography on silica gel with hexanes/EtOAc 9:1 to 5:5 and afforded **149** (109 mg, 52%) as an oil. TLC R_f = 0.68 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.63 (d, $J_{H1, H2}$ = 4.0 Hz, 1 H, H-1), 5.55 (dd, $J_{H2, H3}$ = 9.8 Hz, $J_{H3, H4}$ = 9.6 Hz, 1 H, H-3), 5.15 (dd, $J_{H4, H5}$ = 10.0 Hz, 1 H, H-4), 4.84 (dd, 1 H, H-2), 4.29-4.24 (m, 1 H, H-5), 3.49 (dd, $J_{H5, H6a}$ = 2.7 Hz, $J_{H6a, H6b}$ = 13.7 Hz, 1 H, H-6a),3.36 (dd, $J_{H5, H6b}$ = 5.1 Hz, 1 H, H-6b), 2.08, 2.06, 2.04 (3s, 9 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 169.7, 169.6, 169.3 (C=O), 86.0 (C-1), 72.9 (C-5), 70.4 (C-3), 69.9 (C-2), 68.1 (C-4), 50.1 (C-6), 20.5, 20.5, 20.5 (CH$_3$CO).; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{16}BrN_3O_7$, 394.02444, found 394.02453.

(2-chloro-4-nitrophenyl)-2,3,4-tri-*O*-acetyl-6-azido-6-deoxy-β-D-glucopyranoside (150). This compound was prepared following the general procedure **3.4.8.2** from **149** (100 mg, 0.25 mmol). The purification on silica gel (hexanes/EtOAc, 7:3) afforded **150** (76 mg, 61%) as a white powder. TLC R_f = 0.50 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.29 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.16 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.29 (d, 1 H, H-6'), 5.40 (dd, $J_{H1, H2}$ = 7.6 Hz, $J_{H2, H3}$ = 9.5 Hz, 1 H, H-2), 5.33 (dd, $J_{H3, H4}$ = 9.1 Hz, 1 H, H-3), 5.19 (d, 1 H, H-1), 5.10 (dd, $J_{H4, H5}$ = 9.9 Hz, 1 H, H-4), 3.87 (ddd, $J_{H5, H6a}$ = 2.6 Hz, $J_{H5, H6b}$ = 7.5 Hz, 1 H, H-5), 3.48 (dd, $J_{H6a, H6b}$ = 13.4 Hz, 1 H, H-6a), 3.68 (dd, 1 H, H-6b), 2.09, 2.08, 2.06 (3s, 9 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.0, 169.4, 169.0 (C=O), 157.0 (C-1'), 143.3 (C-4'), 126.1 (C-3'), 124.9 (C-2'), 123.7 (C-5'), 116.7 (C-6'), 99.2 (C-1), 73.9 (C-5), 71.9 (C-3), 70.4 (C-2), 69.1 (C-4), 51.2 (C-6), 20.5, 20.5, 20.5 (CH$_3$CO).; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{19}ClN_4O_{10}$, 509.06819, found 509.06898.

(2-chloro-4-nitrophenyl)-6-azido-6-deoxy-β-D-glucopyranoside (37). A solution of **150** (40

mg, 0.08 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **37** (25 mg, 85%) as a white powder. TLC R_f = 0.40 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.32 (d, $J_{H3',H5'}$ = 2.8 Hz, 1 H, H-3'), 8.20 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.45 (d, 1 H, H-6'), 5.24 (d, $J_{H1,H2}$ = 7.7 Hz, 1 H, H-1), 3.71 (ddd, $J_{H4,H5}$ = 9.5 Hz, $J_{H5,H6a}$ = 2.4 Hz, $J_{H5,H6b}$ = 6.9 Hz, 1 H, H-5), 3.60 (dd, $J_{H2,H3}$ = 9.1 Hz, 1 H, H-2), 3.57 (dd, $J_{H6a,H6b}$ = 13.3 Hz, 1 H, H-6a), 3.50 (dd, $J_{H3,H4}$ = 9.0 Hz, 1 H, H-3), 3.45 (dd, 1 H, H-6b), 3.38 (dd, 1 H, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 159.1 (C-1'), 143.8 (C-4'), 126.8 (C-3'), 125.0 (C-2'), 124.9 (C-5'), 116.9 (C-6'), 101.6 (C-1), 77.7 (C-3), 77.4 (C-5), 74.6 (C-2), 72.0 (C-4), 52.7 (C-6); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{13}ClN_4O_7$, 383.03650, found 383.03555.

3.4.8.6. Synthesis of (2-chloro-4-nitrophenyl)-6-deoxy-6-thio-β-D-glucopyranoside (36).



a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

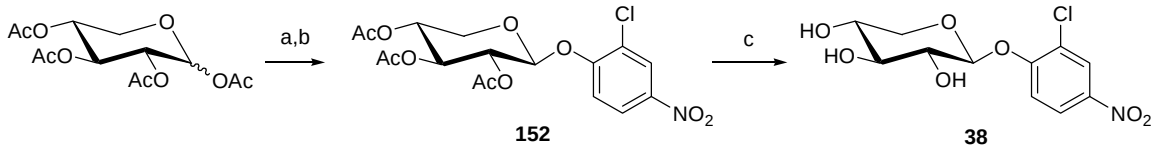
(2-chloro-4-nitrophenyl)-2,3,4-tri-*O*-benzoyl-6-*S*-acetyl-6-deoxy-β-D-glucopyranoside (151). This compound was prepared following the general procedure **3.4.8.1** and **3.4.8.2** from the previously described acetyl-2,3,4-tri-*O*-benzoyl-6-*S*-acetyl-6-deoxy-β-D-glucopyranoside[79] (200 mg, 0.34 mmol). The purification on silica gel (hexanes/EtOAc, 7:3) afforded **151** (191 mg, 80%)

as a white powder. TLC R_f = 0.53 (EtOAc:hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.20 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.01-7.85 (m, 6 H, Ph), 7.55-7.26 (m, 10 H, Ph), 5.95 (dd, $J_{H2, H3}$ = $J_{H3, H4}$ = 9.5 Hz, 1 H, H-3), 5.87 (dd, $J_{H1, H2}$ = 7.5 Hz, 1 H, H-2), 5.60 (dd, $J_{H4, H5}$ = 9.5 Hz, 1 H, H-4), 5.40 (d, 1 H, H-1), 4.13-4.07 (m, 1 H, H-5), 3.50 (dd, $J_{H5, H6a}$ = 2.9 Hz, $J_{H6a, H6b}$ = 14.4 Hz, 1 H, H-6a), 3.09 (dd, $J_{H5, H6b}$ = 8.2 Hz, 1 H, H-6b), 2.37 (CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 194.1, 165.6, 165.4, 164.8 (C=O), 157.2 (C-1'), 143.2 (C-4'), 133.6, 133.4, 129.8, 129.8, 129.7, 129.7, 128.9, 128.6, 128.5, 128.5, 128.3, 128.3 (C-aro)126.0 (C-3'), 125.1 (C-2'), 123.5 (C-5'), 116.9 (C-6'), 99.6 (C-1), 74.4 (C-5), 72.1 (C-3), 71.2 (C-2), 71.1 (C-4), 30.4 (C-6), 30.4 (*C*H_3CO).; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{35}H_{28}ClNO_{11}$, 728.0963, found 728.0944.

(2-chloro-4-nitrophenyl)-6-deoxy-6-thio-β-D-glucopyranoside (36). A solution of **151** (100 mg, 0.14 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **36** (38 mg, 78%) as a white powder. TLC R_f = 0.48 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.31 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.46 (d, 1 H, H-6'), 5.18 (d, $J_{H1,H2}$ = 7.7 Hz, 1 H, H-1), 3.58 (dd, $J_{H2,H3}$ = 9.2 Hz, 1 H, H-2), 3.57-3.52 (m, 1 H, H-5), 3.49 (dd, $J_{H3,H4}$ = 9.1 Hz, 1 H, H-3), 3.38 (dd, $J_{H4,H5}$ = 9.3 Hz, 1 H, H-4), 2.99 (dd, 1 H, $J_{H5,H6a}$ = 2.3 Hz, $J_{H6a,H6b}$ = 14.2 Hz, H-6a), 2.66 (dd, $J_{H5,H6b}$ = 7.9 Hz, 1 H, H-6b); ^{13}C NMR (100 MHz, CD_3OD) δ 159.2 (C-1'), 143.7 (C-4'), 126.8 (C-3'), 124.9 (C-2'), 124.9 (C-5'), 117.0 (C-6'), 101.8 (C-1), 78.8 (C-3), 77.8 (C-5), 74.7 (C-2), 73.5 (C-4), 26.8 (C-6) [3]; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_7S$, 374.00717, found 374.00681.

[3] Less than 10% of dimer was observed by ^{1}H NMR.

3.4.8.7. Synthesis of 2-chloro-4-nitrophenyl-β-D-xylopyranoside (38).



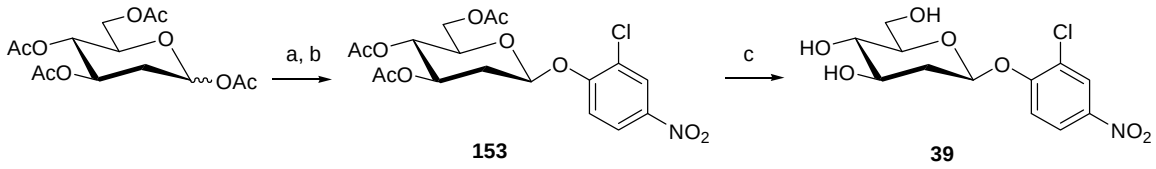
a) HBr in AcOH, CH_2Cl_2, 0 °C, 2h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

(2-chloro-4-nitrophenyl)-2,3,4-tri-*O*-acetyl-β-D-xylopyranoside (152). A solution of per-*O*-acetylated D-xylopyranose (355 mg, 1.1 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 7:3) afforded **152** (180 mg, 38% yield) as a white powder. TLC R_f = 0.56 (EtOAc:hexanes, 5:5). The product was carried forward without additional characterization.

2-chloro-4-nitrophenyl-β-D-xylopyranoside (38). A solution of **152** (180 mg, 0.42 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **38** (96 mg, 75% yield, β-anomer) as white crystals. TLC R_f = 0.53 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.31 (d, $J_{H3',H5'}$ = 2.8 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.38 (d, 1 H, H-6'), 5.15 (d, $J_{H-1,H-2}$ = 7.3 Hz, 1 H, H-1), 3.95 (dd, $J_{H-5a,H-4}$ = 5.1 Hz, $J_{H-5a,H-5b}$ = 11.4 Hz, 1 H, H-5a), 3.64-3.54 (m, 2 H, H-2, H-4), 3.49-3.41 (m, 2 H, H-3, H-5b); ^{13}C NMR (100 MHz, CD_3OD) δ 159.3 (C-1'), 143.7 (C-4'), 126.8 (C-3'), 125.0 (C-2'), 124.9 (C-5'), 116.8 (C-6'), 102.5 (C-1), 77.6 (C-3), 74.4 (C-2), 70.8 (C-4), 67.2 (C-5). HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNO_8$, 358.0301, found 358.0313.

An alternative reaction following sequential reaction procedures outlined in sections **3.4.8.1** and **3.4.8.2** yielded **38** as a mixture of anomers. Subsequent HPLC purification of a portion of the crude product yielded 19 mg of α- and 68 mg of β-anomer, suggesting they were present in a α:β ratio of 1:3 in the crude reaction. Characterization of the α-anomer of **38** was as follows: TLC R_f = 0.54 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.30 (d, $J_{H3',H5'}$ = 2.8 Hz, 1 H, H-3'), 8.19 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.45 (d, 1 H, H-6'), 5.80 (d, $J_{H-1,H-2}$ = 3.5 Hz, 1 H, H-1), 3.88 (dd, $J_{H-5a,H-4}$ = 8.5 Hz, $J_{H-5a,H-5b}$ = 9.8 Hz, 1 H, H-5a), 3.66-3.56 (m, 3 H, H-2, H-3, H-4), 3.48 (t, $J_{H-5b,H-a}$ = 9.8 Hz, 1 H, H-5b).

3.4.8.8. Synthesis of (2-chloro-4-nitrophenyl)-2-deoxy-β-D-glucopyranoside (39).



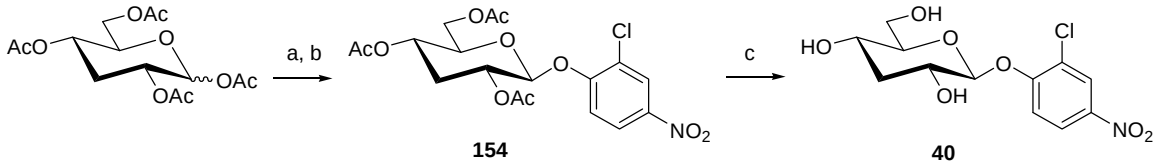
a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

(2-chloro-4-nitrophenyl)-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranoside (153). A solution of per-*O*-acetylated 2-deoxy-D-glucopyranose (150 mg, 0.45 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was then directly used in the glycosylation reaction following the general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 8:2) afforded **153** (45 mg, 22%) as a white powder. TLC R_f = 0.39

(EtOAc:hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.29 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.12 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.25 (d, 1 H, H-6'), 5.37 (dd, $J_{H1, H2a}$ = 2.5 Hz, $J_{H1, H2b}$ = 8.8 Hz, 1 H, H-1), 5.17-5.11 (m, 1 H, H-3), 5.09 (dd, $J_{H3, H4}$ = $J_{H4, H5}$ = 8.8 Hz, 1 H, H-4), 4.31 (dd, $J_{H5, H6a}$ = 5.8 Hz, $J_{H6a, H6b}$ = 12.2 Hz, 1 H, H-6a), 4.19 (dd, $J_{H5, H6b}$ = 2.9 Hz, 1 H, H-6b), 3.87 (ddd, 1 H, H-5), 2.65 (ddd, $J_{H1, H2a}$ = 2.5 Hz, $J_{H2a, H3}$ = 4.7 Hz, $J_{H2a, H2b}$ = 12.9 Hz, 1 H, H-2a), 2.26-2.18 (m, 1 H, H-2b), 2.09, 2.09, 2.06 (3s, 9 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.4, 170.1, 169.6, (C=O), 157.2 (C-1'), 142.6 (C-4'), 126.1 (C-3'), 124.4 (C-2'), 123.4 (C-5'), 115.6 (C-6'), 97.0 (C-1), 72.6 (C-5), 69.2 (C-3), 68.2 (C-4), 62.3 (C-6), 34.8 (C-2), 20.8, 20.7, 20.6, (CH_3CO); HRMS (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{20}ClNO_{10}$, 468.0668, found 468.0689.

(2-chloro-4-nitrophenyl)-2-deoxy-β-D-glucopyranoside (39). A solution of **153** (43 mg, 0.10 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 95:5) to give **39** (19 mg, 61%) as a white powder. TLC R_f = 0.45 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.28 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.16 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.42 (d, 1 H, H-6'), 5.46 (dd, $J_{H1, H2a}$ = 2.2 Hz, $J_{H1, H2b}$ = 9.7 Hz, 1 H, H-1), 3.91 (dd, $J_{H5, H6a}$ = 2.3 Hz, $J_{H6a, H6b}$ = 12.0 Hz, 1 H, H-6a), 3.71 (dd, $J_{H5, H6b}$ = 5.8 Hz, 1 H, H-6b), 3.69 (ddd, $J_{H2a, H3}$ = 5.1 Hz, $J_{H2b, H3}$ = $J_{H3, H4}$ = 12.0 Hz, 1 H, H-3), 3.45 (ddd, $J_{H4, H5}$ = 12.0 Hz, 1 H, H-5), 3.31 (dd, 1 H, H-4), 2.40 (ddd, $J_{H1, H2a}$ = 2.1 Hz, $J_{H2a, H3}$ = 5.1 Hz, $J_{H2a, H2b}$ = 12.3 Hz, 1 H, H-2a), 1.86 (ddd, $J_{H1, H2b}$ = 9.7 Hz, $J_{H2b, H3}$ = 12.0 Hz, 1 H, H-2b); ^{13}C NMR (100 MHz, CD_3OD) δ 159.1 (C-1'), 143.5 (C-4'), 126.6 (C-3'), 124.9 (C-2'), 124.5 (C-5'), 117.1 (C-6'), 98.8 (C-1), 78.7 (C-5), 72.6 (C-4), 72.1 (C-3), 62.6 (C-6), 39.7 (C-2); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_7$, 342.03510, found 342.03495.

3.4.8.9. Synthesis of (2-chloro-4-nitrophenyl)-3-deoxy-β-D-glucopyranoside (40).

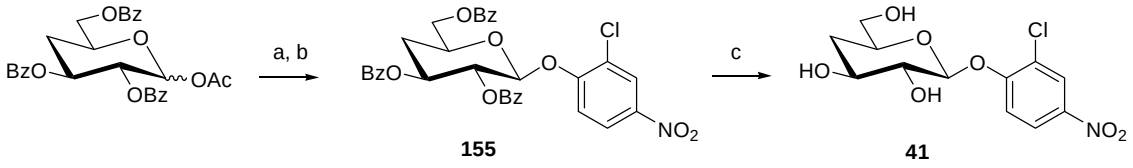


a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

(2-chloro-4-nitrophenyl)-2,4,6-tri-*O*-acetyl-3-deoxy-β-D-glucopyranoside (154). A solution of 1,2,4,6-tetra-*O*-acetyl-3-deoxy-β-D-glucopyranoside synthesized as previously described[80] (60 mg, 0.18 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was used without further purification following the general procedure **3.4.8.2**. The purification on silica gel (hexanes/EtOAc, 8:2) afforded **154** (41 mg, 50%) as a white powder. TLC R_f = 0.49 (EtOAc:hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.28 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.11 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.26 (d, 1 H, H-6'), 5.21-5.17 (m, 2 H, H-1, H-2), 4.94 (ddd, $J_{H3a, H4}$ = 4.9 Hz, $J_{H3b, H4}$ = 9.9 Hz $J_{H4, H5}$ = 8.9 Hz, 1 H, H-4), 4.25 (dd, $J_{H5, H6a}$ = 3.5 Hz, $J_{H6a, H6b}$ = 12.2 Hz, 1 H, H-6a), 4.19 (dd, $J_{H5, H6b}$ = 5.8 Hz, 1 H, H-6b), 3.94 (ddd, 1 H, H-5), 2.67 (ddd, $J_{H2, H3a}$ = 12.7 Hz, $J_{H3a, H3b}$ = 9.5 Hz, $J_{H3a, H3b}$ = 9.5 Hz, 1 H, H-3a), 2.08, 2.08, 2.05 (3s, 9 H, CH_3CO), 1.80 (m, 1 H, H-3b); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.4, 169.4, 169.3, (C=O), 157.3 (C-1'), 142.8 (C-4'), 126.1 (C-3'), 124.7 (C-2'), 123.5 (C-5'), 115.9 (C-6'), 99.9 (C-1), 75.4 (C-5), 67.3 (C-2), 65.2 (C-4), 62.4 (C-6), 31.7 (C-3), 20.8, 20.8, 20.6, (*C*H_3CO); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{18}H_{20}ClNO_{10}$, 468.06680, found 468.06699.

(2-chloro-4-nitrophenyl)-3-deoxy-β-D-glucopyranoside (40). A solution of **154** (40 mg, 0.09 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 95:5) to give **40** (25 mg, 86%) as a white powder. TLC R_f = 0.50 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.29 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.43 (d, 1 H, H-6'), 5.12 (d, $J_{H1, H2}$ = 7.5 Hz, 1 H, H-1), 3.87 (dd, $J_{H5, H6a}$ = 2.5 Hz, $J_{H6a, H6b}$ = 12.1 Hz, 1 H, H-6a), 3.80 (ddd, $J_{H2, H3a}$ = 12.4 Hz, $J_{H2, H3b}$ = 11.8 Hz, 1 H, H-2), 3.67 (dd, $J_{H5, H6b}$ = 5.9 Hz, 1 H, H-6b), 3.68-3.61 (m, 1 H, H-4), 3.50 (ddd, $J_{H4, H5}$ = 8.9, H, H-5), 2.42 (ddd, $J_{H3a, H4}$ = 4.9 Hz, $J_{H3a, H3b}$ = 9.8 Hz, 1 H, H-3a), 1.64 (ddd, $J_{H3b, H4}$ = 11.4 Hz, 1 H, H-3b); ^{13}C NMR (100 MHz, CD_3OD) δ 159.4 (C-1'), 143.5 (C-4'), 126.7 (C-3'), 124.9 (C-2'), 124.8 (C-5'), 116.9 (C-6'), 103.7 (C-1), 82.3 (C-5), 68.7 (C-2), 65.6 (C-4), 62.4 (C-6), 40.5 (C-3); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_7$, 342.03510, found 342.03511.

3.4.8.10. Synthesis of (2-chloro-4-nitrophenyl)-4-deoxy-β-D-glucopyranoside (41).



a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 16h.

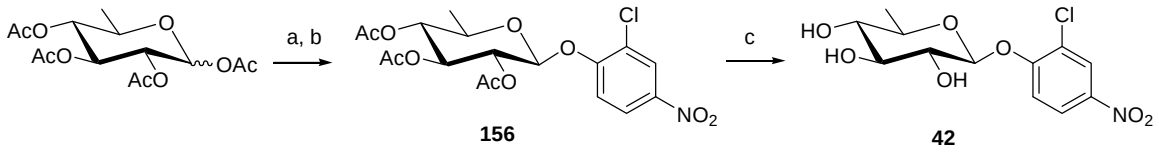
(2-chloro-4-nitrophenyl)-2,3,6-tri-*O*-benzoyl-4-deoxy-β-D-glucopyranoside (155). A solution of acetyl 2,3,6-tri-*O*-benzoyl-4-deoxy-β-D-glucopyranoside (130 mg, 0.25 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was used without

further purification following the general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 9:1) afforded **155** (156 mg, 98%) as a white powder. TLC R_f = 0.46 (EtOAc:hexanes, 1:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.12 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.05-7.97 (m, 6 H, Ph), 7.80 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.64-7.34 (m, 9 H, Ph), 7.26 (d, 1 H, H-6'), 5.82 (dd, $J_{H1, H2}$ = 7.5 Hz, $J_{H2, H3}$ = 9.3 Hz, 1 H, H-2), 5.58 (ddd, $J_{H3, H4a}$ = 5.4 Hz, $J_{H3, H4b}$ = 11.1 Hz, 1 H, H-3), 5.38 (d, 1 H, H-1), 4.60 (dd, $J_{H5, H6a}$ = 3.7 Hz, $J_{H6a, H6b}$ = 11.8 Hz, 1 H, H-6a), 4.53 (dd, $J_{H5, H6b}$ = 6.9 Hz, 1 H, H-6b), 4.36-4.31 (m, 1 H, H-5), 2.60 (ddd, $J_{H4a, H4b}$ = 11.5 Hz, $J_{H4a, H5}$ = 7.3 Hz, 1 H, H-4a), 2.04 (m, 1 H, H-4b); ^{13}C NMR (100 MHz, $CDCl_3$) δ 165.8, 165.7, 165.1, (C=O), 157.3 (C-1'), 142.7 (C-4'), 133.5, 133.4, 133.3, 129.7, 129.6, 129.5, 129.3, 129.1, 129.0, 128.5, 128.4, 128.3 (C-aro), 125.9 (C-3'), 124.7 (C-2'), 123.2 (C-5'), 116.4 (C-6'), 99.6 (C-1), 71.6 (C-5), 70.7 (C-2), 70.6 (C-3), 65.2 (C-6), 32.0 (C-4). HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{33}H_{26}ClNO_{10}$, 654.11374, found 654.11225.

(2-chloro-4-nitrophenyl)-4-deoxy-β-D-glucopyranoside (41). A solution of **155** (150 mg, 0.09 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **41** (62 mg, 82%) as a white powder. TLC R_f = 0.42 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.30 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.17 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.42 (d, 1 H, H-6'), 5.11 (d, $J_{H1, H2}$ = 7.6 Hz, 1 H, H-1), 3.84-3.77 (m, 1 H, H-5), 3.74 (ddd, $J_{H2, H3}$ = 9.0 Hz, $J_{H3, H4a}$ = 5.2 Hz, $J_{H3, H4b}$ = 11.5 Hz, 1 H, H-3), 3.62 (dd, $J_{H5, H6a}$ = 4.0 Hz, $J_{H6a, H6b}$ = 11.8 Hz, 1 H, H-6a), 3.58 (dd, $J_{H5, H6b}$ = 5.9 Hz, 1 H, H-6b), 3.47 (dd, 1 H, H-2), 2.00 (ddd, $J_{H4a, H5}$ = 1.9 Hz, $J_{H4a, H4b}$ = 12.9 Hz, 1 H, H-4a), 1.52 (dd, $J_{H4b, H5}$ = 11.7 Hz, 1 H, H-4b); ^{13}C NMR (100 MHz, CD_3OD) δ 159.5 (C-1'), 143.5 (C-4'), 126.7 (C-3'), 124.9 (C-2'), 124.8 (C-5'), 116.9 (C-6'), 102.3 (C-1), 76.3 (C-2), 74.8 (C-5), 72.1 (C-3), 65.2 (C-

6), 35.9 (C-4); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_7$, 342.03510, found 342.03525.

3.4.8.11. Synthesis of (2-chloro-4-nitrophenyl)-6-deoxy-β-D-glucopyranoside (42).

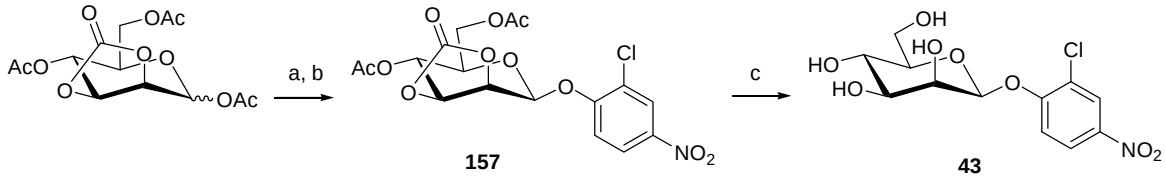


a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

(2-chloro-4-nitrophenyl)-2,3,4-tri-*O*-acetyl-6-deoxy-β-D-glucopyranoside (156). A solution of per-*O*-acetylated 6-deoxy-glucopyranose (120 mg, 0.37 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following the general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 8:2) afforded **156** (132 mg, 81%) as a white powder. TLC R_f = 0.55 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.25 (d, $J_{H3',H5'}$ = 2.7 Hz, 1 H, H-3'), 8.11 (dd, $J_{H5',H6'}$ = 9.1 Hz, 1 H, H-5'), 7.21 (d, 1 H, H-6'), 5.34 (dd, $J_{H1,H2}$ = 7.7 Hz, $J_{H2,H3}$ = 9.7 Hz, 1 H, H-2), 5.25 (dd, $J_{H3,H4}$ = 9.3 Hz, 1 H, H-3), 5.12 (d, 1 H, H-1), 4.93 (dd, $J_{H4,H5}$ = 9.5 Hz, 1 H, H-4), 3.82-3.76 (m, 1 H, H-5), 2.05, 2.05, 2.02 (3s, 9 H, CH_3CO); 1.31 (d, $J_{H5,H6}$ = 6.2 Hz, 3 H, H-6) ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.2, 169.5, 169.1, (C=O), 157.4 (C-1'), 143.0 (C-4'), 126.1 (C-3'), 124.8 (C-2'), 123.6 (C-5'), 116.2 (C-6'), 99.2 (C-1), 72.7 (C-3), 72.1 (C-4), 70.8 (C-2), 70.8 (C-5), 20.6, 20.6, 20.5, (*C*H_3CO), 17.4 (C-6).; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{20}ClNO_{10}$, 468.0673, found 468.0677.

(2-chloro-4-nitrophenyl)-6-deoxy-β-D-glucopyranoside (42). A solution of **156** (130 mg, 0.29 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **42** (81 mg, 87%) as a white powder. TLC R_f = 0.43 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.27 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.16 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.35 (d, 1 H, H-6'), 5.15 (dd, $J_{H1, H2}$ = 7.7 Hz, 1 H, H-1), 3.62-3.55 (m, 2 H, H-2, H-5), 3.46 (dd, $J_{H2, H3} = J_{H3, H4}$ = 9.1 Hz, 1 H, H-3), 3.13 (dd, $J_{H4, H5}$ = 9.3 Hz, 1 H, H-4), 1.31 (d, $J_{H5, H6}$ = 6.2 Hz, 3 H, H-6);^{13}C NMR (100 MHz, CD_3OD) δ 159.3 (C-1'), 143.5 (C-4'), 126.8 (C-3'), 124.9 (C-2'), 124.8 (C-5'), 116.6 (C-6'), 101.5 (C-1), 77.7 (C-3), 76.4 (C-4), 74.8 (C-2), 73.8 (C-5), 18.0 (C-6); HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNO_7$, 342.0356, found 342.0356.

3.4.8.12. Synthesis of 2-chloro-4-nitrophenyl-β-D-mannopyranoside (43).



a) HBr in AcOH, CH_2Cl_2, 0 °C, 90 min; b) potassium 2-chloro-4-nitrophenol, dicyclohexyl-18-crown-6, ACN, RT, 12h; c) 4 Å molecular sieves, MeOH, RT, 12h.

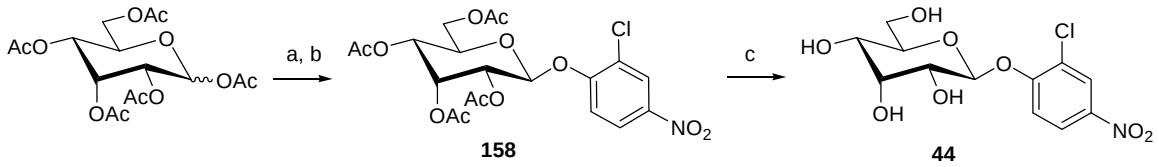
(2-chloro-4-nitrophenyl)-4,6-di-*O*-acetyl-2,3-carbonyl-β-D-mannopyranoside (157). A solution of the previously described 4,6-di-*O*-acetyl-2,3-carbonyl-β-D-mannopyranoside[81] (100 mg, 0.30 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1** and the crude glycosyl bromide obtained used without further purification. For glycosylation, 2-chloro-4-

nitrophenol (100 mg, 0.58 mmol) was converted to the potassium salt by dissolving it in an equimolar solution of potassium hydroxide, followed by evaporation under reduced pressure and finally freeze-drying. In a dry flask at room temperature were combined the above prepared crude glycosyl bromide, 2-chloro-4-nitrophenol (potassium salt, 94 mg, 0.45 mmol), dicyclohexyl-18-crown-6 (8 mg, 0.30 mmol) in anhydrous acetonitrile (3 mL). The reaction was stirred at room temperature for 12 hours. The mixture was filtered and the filtrate evaporated. The residue was diluted with CH_2Cl_2 (40 mL), washed with $NaHCO_3$ sat solution (2 x 15 mL) and with brine (15 mL). The organic phase was dried over $MgSO_4$, and the solvent removed under reduced pressure. Purification on silica gel (hexanes/EtOAc, 2:8) afforded **157** (117 mg, 87%) as a white powder. TLC R_f = 0.38 (EtOAc:hexanes, 8:2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.33 (d, $J_{H3',H5'}$ = 2.7 Hz, 1 H, H-3'), 8.17 (dd, $J_{H5',H6'}$ = 9.1 Hz, 1 H, H-5'), 7.28 (d, 1 H, H-6'), 5.96 (dd, $J_{H3,H4}$ = 6.4 Hz, $J_{H4,H5}$ = 10.4 Hz, 1 H, H-4), 5.85 (d, $J_{H1,H2}$ = 3.2 Hz, 1 H, H-1), 5.08 (dd, $J_{H2,H3}$ = 9.2 Hz, 1 H, H-2), 5.03 (dd, 1 H, H-3), 4.17-4.03 (m, 3 H, H-5, H-6a, H-6b), 2.13, 1.63 (2s, 6 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.1, 168.5, (C=O), 156.1 (C-1'), 142.8 (C-4'), 126.0 (C-3'), 124.6 (C-2'), 123.6 (C-5'), 114.7 (C-6'), 93.0 (C-1), 76.7 (C-3), 71.6 (C-5), 70.9 (C-2), 66.0 (C-4), 61.4 (C-6), 20.6, 19.9 (*C*H_3CO).; HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{17}H_{16}ClNO_{11}$, 468.0304, found 468.0294.

2-chloro-4-nitrophenyl-β-D-mannopyranoside (43). A mixture of **157** (10 mg, 0.02 mmol) and powdered activated 4 Å molecular sieves (10 mg) in methanol (500μL) was stirred at room temperature for 12 hours. The reaction was then filtered, concentrated and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **43** (5.5 mg, 73%) as a white powder. TLC R_f = 0.39 (CH_2Cl_2/MeOH, 85:15); ^{1}H NMR (500 MHz, CD_3OD) δ 8.30 (d, $J_{H3',H5'}$ = 2.8 Hz,

1 H, H-3'), 8.18 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.42 (d, 1 H, H-6'), 5.40 (s, 1 H, H-1), 4.15 (d, $J_{H2, H3}$ = 3.0 Hz, 1 H, H-2), 3.92 (dd, $J_{H5, H6a}$ = 2.2 Hz, $J_{H6a, H6b}$ = 12.0 Hz, 1 H, H-6a), 3.74 (dd, $J_{H5, H6b}$ = 6.1 Hz, 1 H, H-6b), 3.67 (dd, $J_{H3, H4} = J_{H4, H5}$ = 9.5 Hz, 1 H, H-4), 3.59 (dd, 1 H, H-3), 3.47 (ddd, 1 H, H-5); ^{13}C NMR (125 MHz, CD_3OD) δ 159.2 (C-1'), 143.7 (C-4'), 126.7 (C-3'), 124.9(C-2'), 124.8 (C-5'), 117.3 (C-6'), 99.7 (C-1), 79.0 (C-5), 75.0 (C-3), 72.1 (C-2), 68.2 (C-4), 62.7 (C-6). HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_8$, 358.0300, found 358.0293.

3.4.8.13. Synthesis of 2-chloro-4-nitrophenyl-β-D-allopyranoside (44).



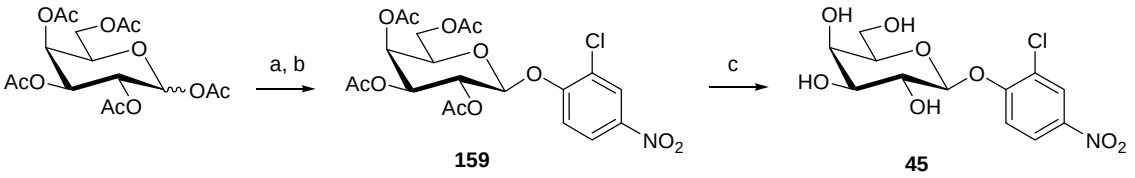
a) HBr in AcOH, CH_2Cl_2, 0 °C, 90 min; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 12h.

(**2-chloro-4-nitrophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-allopyranoside (158).** A solution of per-*O*-acetylated allopyranose (172 mg, 0.44 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following the general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 7:3) afforded **158** (63mg, 29%) as a white powder. TLC R_f = 0.37 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (d, $J_{H3', H5'}$ = 2.7 Hz, 1 H, H-3'), 8.14 (dd, $J_{H5', H6'}$ = 9.1 Hz, 1 H, H-5'), 7.31 (d, 1 H, H-6'), 5.75 (dd, $J_{H2, H3}$ = 3.0 Hz, $J_{H3, H4}$ = 2.9 Hz, 1 H, H-3), 5.41 (d, $J_{H1, H2}$ = 8.1 Hz, 1 H, H-1), 5.31 (dd, 1 H, H-2), 5.06 (dd, $J_{H4, H5}$ = 9.6 Hz 1 H, H-4), 4.34-4.20 (m, 3 H, H-5, H-6a, H-6b),

2.19, 2.11, 2.08, 2.04 (4s, 12 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.5, 169.5, 168.9, 168.8, (C=O), 157.5 (C-1'), 142.9 (C-4'), 126.1 (C-3'), 124.8 (C-2'), 123.5 (C-5'), 116.2 (C-6'), 97.8 (C-1), 71.0 (C-5), 68.2 (C-3), 68.1 (C-2), 65.9 (C-4), 62.0 (C-6), 20.7, 20.6, 20.5, 20.4 (CH_3CO).; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{20}H_{22}ClNO_{12}$, 526.07227, found 526.07234.

2-chloro-4-nitrophenyl-β-D-allopyranoside (44). A solution of **158** (50 mg, 0.10 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **44** (27 mg, 83%) as a white powder. TLC R_f = 0.28 (CH_2Cl_2/MeOH, 9:1); ^1H NMR (400 MHz, CD_3OD) δ8.29 (d, $J_{H3',H5'}$ = 2.7 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.42 (d, 1 H, H-6'), 5.47 (dd, $J_{H1, H2}$ = 7.8 Hz, 1 H, H-1), 4.16 (dd, $J_{H2, H3}$ = 3.0 Hz, $J_{H3, H4}$ = 2.9 Hz, 1 H, H-3), 3.94 (ddd, $J_{H4, H5}$ = 9.7 Hz, $J_{H5, H6a}$ = 2.4 Hz, $J_{H5, H6b}$ = 5.7 Hz, 1 H, H-5), 3.88 (dd, $J_{H6a, H6b}$ = 12.1 Hz, 1 H, H-6a), 3.72 (dd, 1 H, H-2), 3.70 (dd, 1 H, H-6b), 3.63 (dd, 1 H, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 159.7 (C-1'), 143.5 (C-4'), 126.7 (C-3'), 124.9(C-2'), 124.7 (C-5'), 116.8 (C-6'), 100.2 (C-1), 76.2 (C-5), 73.1 (C-3), 71.8 (C-2), 68.4 (C-4), 62.7 (C-6). HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNO_8$, 358.03002, found 358.03008.

3.4.8.14. Synthesis of 2-chloro-4-nitrophenyl-β-D-galactopyranoside (45).



a) HBr in AcOH, CH_2Cl_2, 0 °C, 1h; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide,

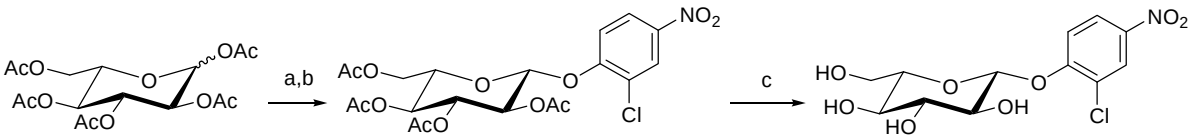
CH_2Cl_2/NaOH (1:1), RT, 18h; c) Et_3N/MeOH/H_2O (4:10:10), RT, 18h.

(2-chloro-4-nitrophenyl)-2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranoside (159). A solution of per-*O*-acetylated galactopyranose (300 mg, 0.77 mmol) in was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following the general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 7:3) afforded **159** (349 mg, 90%) as a white powder. TLC R_f = 0.45 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.31 (d, $J_{H3', H5'}$ = 2.5 Hz, 1 H, H-3'), 8.13 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.29 (d, 1 H, H-6'), 5.63 (dd, $J_{H1, H2}$ = 7.9 Hz, $J_{H2, H3}$ = 10.5 Hz, 1 H, H-2), 5.50 (dd, $J_{H3, H4}$ = 3.4 Hz, $J_{4,5}$ = 0.9 Hz, 1 H, H-4), 5.15 (dd, 1 H, H-3), 5.12 (d, 1 H, H-1), 4.23 (dd, $J_{H5, H6a}$ = 7.0 Hz, $J_{H6a, H6b}$ = 11.0 Hz, 1 H, H-6a), 4.17-4.13 (m, 1 H, H-5), 4.14 (dd, $J_{H5, H6b}$ = 6.1 Hz, 1 H, H-6b), 2.21, 2.10, 2.09, 2.03 (4s, 12 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.2, 170.0, 169.9, 169.1, (C=O), 157.3 (C-1'), 143.1 (C-4'), 126.1 (C-3'), 124.8 (C-2'), 123.5 (C-5'), 116.3 (C-6'), 99.9 (C-1), 71.6 (C-5), 70.3 (C-3), 67.8 (C-2), 66.6 (C-4), 61.3 (C-6), 20.6, 20.6, 20.5, 20.5 (*C*H_3CO). HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{20}H_{22}ClNO_{12}$, 526.07227, found 526.07208. Spectral data are consistent with those previously reported[82].

2-chloro-4-nitrophenyl-β-D-galactopyranoside (45). A solution of **159** (50 mg, 0.10 mmol) in 2.5 mL of a mixture Et_3N/MeOH/H_2O (4:10:10) was stirred at room temperature for two hours. The reaction was concentrated under reduced pressure and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **45** (29 mg, 88%) as a white powder. TLC R_f = 0.28 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.30 (d, $J_{H3', H5'}$ = 2.8 Hz, 1 H, H-3'), 8.17 (dd, $J_{H5', H6'}$ = 9.2 Hz, 1 H, H-5'), 7.45 (d, 1 H, H-6'), 5.13 (dd, $J_{H1, H2}$ = 7.7 Hz, 1 H, H-1), 3.95-

3.87 (m, 2 H, H-2, H-4), 3.80-3.73 (m, 3 H, H-4, H-6a, H-6b), 3.62 (dd, $J_{H2,\ H3}$ = 9.7 Hz, $J_{H3,\ H4}$ = 3.4 Hz, 1 H, H-3), ^{13}C NMR (100 MHz, CD_3OD) δ 159.5 (C-1'), 143.6 (C-4'), 126.7 (C-3'), 124.9 (C-2'), 124.8 (C-5'), 116.9 (C-6'), 102.5 (C-1), 77.5 (C-5), 74.9 (C-3), 71.9 (C-2), 70.2 (C-4), 62.4 (C-6). HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNO_8$, 358.03002, found 358.02994. Spectral data are consistent with those previously reported[82].

3.4.8.15. Synthesis of 2-chloro-4-nitrophenyl-β-L-glucopyranoside (46).



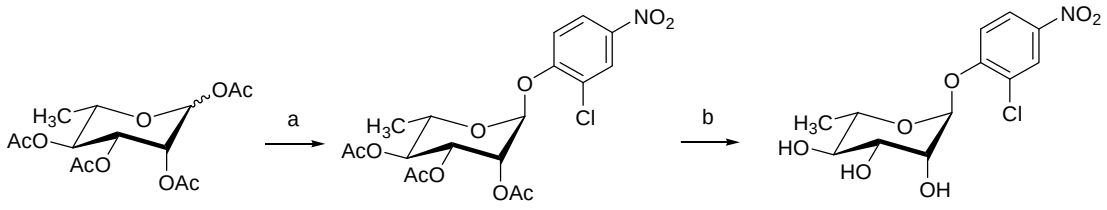
a) HBr in AcOH, CH_2Cl_2, 0 °C, 90 min; b) 2-chloro-4-nitrophenol, tetrabutylammonium bromide, CH_2Cl_2/NaOH (1:1), RT, 18h; c) NaOMe 0.1M in MeOH, RT, 18h.

(2-chloro-4-nitrophenyl)-2,3,4,6-tetra-*O*-acetyl-β-L-glucopyranoside (160). A solution of per-*O*-acetylated L-glucopyranose (150 mg, 0.39 mmol) in CH_2Cl_2 was treated as described in the general procedure **3.4.8.1**. The residue obtained was directly used in the glycosylation reaction following the general procedure **3.4.8.2**. Purification on silica gel (hexanes/EtOAc, 7:3) afforded **160** (95 mg, 49%) as a white powder.TLC R_f = 0.46 (EtOAc/hexanes, 5:5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.30 (d, $J_{H3',\ H5'}$ = 2.7 Hz, 1 H, H-3'), 8.13 (dd, $J_{H5',\ H6'}$ = 9.1 Hz, 1 H, H-5'), 7.27 (d, 1 H, H-6'), 5.40 (dd, $J_{H1,\ H2}$ = 7.5 Hz, $J_{H2,\ H3}$ = 9.3 Hz, 1 H, H-2), 5.33 (dd, $J_{H3,\ H4}$ = 9.1 Hz, 1 H, H-3), 5.19 (dd, $J_{H4,\ H5}$ = 9.8 Hz, 1 H, H-4), 5.17 (d, 1 H, H-1), 4.29 (dd, $J_{H5,\ H6a}$ = 5.2 Hz, $J_{H6a,\ H6b}$ = 12.4 Hz, 1 H, H-6a), 4.22 (dd, $J_{H5,\ H6b}$ = 2.6 Hz, 1 H, H-6b), 3.95 (ddd, 1 H, H-5), 2.10, 2.09,

2.07, 2.06 (4s, 12 H, CH_3CO); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.3, 170.1, 169.2, 168.9, (C=O), 157.2 (C-1'), 143.2 (C-4'), 126.1 (C-3'), 124.9 C-2'), 123.5 (C-5'), 99.2 (C-1), 72.5 (C-5), 72;1 (C-3), 70.4 (C-2), 67.9 (C-4), 61.7 (C-6), 20.6, 20.5, 20.5, 20.5 (*C*H_3CO). HRMS-ESI (*m/z*): $C_{20}H_{22}ClNO_{12}$, 526.07227, found 526.07263.

2-chloro-4-nitrophenyl-β-L-glucopyranoside (46). A solution of **160** (75 mg, 0.15 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **46** (35 mg, 70%) as a white powder. TLC R_f = 0.26 (CH_2Cl_2/MeOH, 9:1); ^1H NMR (400 MHz, CD_3OD) δ 8.30 (d, $J_{H3',H5'}$ = 2.8 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.42 (d, 1 H, H-6'), 5.17 (d, $J_{H1,H2}$ = 7.6 Hz, 1 H, H-1), 3.89 (dd, $J_{H5,H6a}$ = 2.2 Hz, $J_{H6a,H6b}$ = 12.1 Hz, 1 H, H-6a), 3.70 (dd, $J_{H5,H6b}$ = 5.7 Hz, 1 H, H-6b), 3.58 (dd, $J_{H2,H3}$ = 9.0 Hz, 1 H, H-2), 3.55-3.50 (m, 1 H, H-5), 3.50 (dd, $J_{H3,H4}$ = 8.7 Hz, 1 H, H-3), 3.42 (dd, $J_{H4,H5}$ = 9.4 Hz, 1 H, H-4); ^{13}C NMR (100 MHz, CD_3OD) δ 159.4 (C-1'), 143.6 (C-4'), 126.7 (C-3'), 124.9(C-2'), 124.8 (C-5'), 116.9 (C-6'), 101.8 (C-1), 78.6 (C-5), 78.1 (C-3), 74.6 (C-2), 71.1 (C-4), 62.4 (C-6). HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_8$, 358.03002, found 358.02997.

3.4.8.16. Synthesis of 2-chloro-4-nitrophenyl-α-L-rhamnopyranoside (47).



a) 2-chloro-4-nitrophenol, BF_3•OEt_2, Et_3N, CH_2Cl_2, RT, 5 days; b) NaOMe 0.1M in MeOH, RT, 18h.

2-chloro-4-nitrophenyl-2,3,4-tri-*O*-acetyl-α-L-rhamnopyranose (161). To a solution of 1,2,3,4-tetra-*O*-acetyl-L-rhamnopyranose (530 mg, 1.50 mmol) in anhydrous CH_2Cl_2 (5 mL), were added 2 equiv. of 2-chloro-4-nitrophenol (520 mg), 1 equiv. of triethylamine, and 10 equiv. of boron trifluoride etherate in 4 mL of anhydrous CH_2Cl_2. The reaction was allowed to proceed for 5 days. Purification on silica gel (hexanes/EtOAc, 7:3) afforded **161** (557 mg, 83%) as a yellow powder. TLC R_f = 0.68 (EtOAc:hexanes, 5:5). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.33 (d, $J_{H3',H5'}$ = 2.7 Hz, 1 H, H-3'), 8.15 (dd, $J_{H5',H6'}$ = 9.1 Hz, 1 H, H-5'), 7.44 (d, 1 H, H-6'), 5.65 (d, $J_{H1,H2}$ = 1.6 Hz, 1 H, H-1) 5.57-5.51 (m, 2 H, H-2, H-3), 5.20 (dd, J = 9.7 Hz, 1 H, H-4), 3.95-3.90 (m, 1 H, H-5), 2.22, 2.08, 2.05 (3s, 9 H, CH_3CO), 1.22 (d, $J_{H5,H6}$ = 6.2 Hz, 1 H, H-6).; ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.0, 169.9, 169.8 (C=O), 156.1 (C-1'), 142.7 (C-4'), 126.3 (C-3'), 124.6 (C-2'), 123.7 (C-5'), 115.0 (C-6'), 96.2 (C-1), 70.3 (C-4), 69.2 (C-2), 68.5 (C-3), 68.3 (C-5), 20.8, 20.8, 20.7 (CH_3CO), 17.4 (C-6) ; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{18}H_{19}ClNO_{10}$, 468.06795, found 468.06597.

2-chloro-4-nitrophenyl-α-L-rhamnopyranose (47). A solution of **161** (525 mg, 1.18 mmol) was treated as described in the general procedure **3.4.8.3** and purified by chromatography on silica gel (CH_2Cl_2/MeOH, 9:1) to give **47** (327 mg, 87%) as a white powder. TLC R_f = 0.39 (CH_2Cl_2/MeOH, 9:1); ^{1}H NMR (400 MHz, CD_3OD) δ 8.28 (d, $J_{H3',H5'}$ = 2.8 Hz, 1 H, H-3'), 8.18 (dd, $J_{H5',H6'}$ = 9.2 Hz, 1 H, H-5'), 7.47 (d, 1 H, H-6'), 5.71 (d, $J_{H-1,H-2}$ = 1.8, 1 H, H-1), 4.13 (dd, $J_{H-2,H-3}$ = 3.4 Hz, 1 H, H-2), 5.44 (dd, $J_{H-3,H-4}$ = 9.1 Hz, 1 H, H-3), 3.58-3.49 (m, 2 H, H-4, H-5),

1.24 (d, $J_{H\text{-}6,H\text{-}5}$ = 5.8 Hz, 3 H, H-6); ^{13}C NMR (100 MHz, CD_3OD) δ 156.9 (C-1'), 142.2 (C-4'), 125.6 (C-3'), 123.7 (C-2'), 123.7 (C-5'), 115.4 (C-6'), 99.3 (C-1), 72.2 (C-4), 70.9 (C-3), 70.5 (C-5), 70.4 (C-2), 16.8 (C-6); HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{12}H_{14}ClNO_7$, 342.03510, found 342.03391.

3.4.9. 2-chloro-4-nitrophenyl glycoside screening.

3.4.9.1. *In vitro* 2-chloro-4-nitrophenyl glycoside screening. Reactions containing 7.0 μM (100 μg) of OleD variant TDP-16, 1 mM or 0.1 mM of (U/T)DP, and 1 mM of glycoside member (**9**, **34**-**47**) in 50 mM Tris (pH 8.5) with a final volume of 300 μl were incubated at room temperature. Aliquots were removed at various time points, mixed with an equal volume of ddH_2O, frozen in a bath of dry ice and acetone, and stored at -20 °C. Following, samples were thawed at 4 °C and filtered through a MultiScreen Filter Plate (Millipore, Billerica, MA, USA) according to manufacturer's instructions. Samples were evaluated for formation of NDP-sugar by analytical reverse-phase HPLC with a 250 mm x 4.6 mm Gemini-NX 5μ C18 column (Phenomenex, Torrance, CA, USA) using a linear gradient of 0% to 50% CH_3CN (solvent B) over 25 minutes (solvent A = 50 mM PO_4^{-2}, 5 mM tetrabutylammonium bisulfate, 2% acetonitrile [pH adjusted to 6.0 with KOH]; flow rate = 1 ml min^{-1}; A_{254} nm). In the case of **51a** and **50b** formation, percent conversion was calculated from peak height due to co-elution of NDP and product. Screening of the α-anomer of **38** yielded no turnover in any reactions (see **3.4.8.7** for synthesis and characterization), demonstrating that TDP-16 is only capable of recognizing the β-anomers of D-sugars.

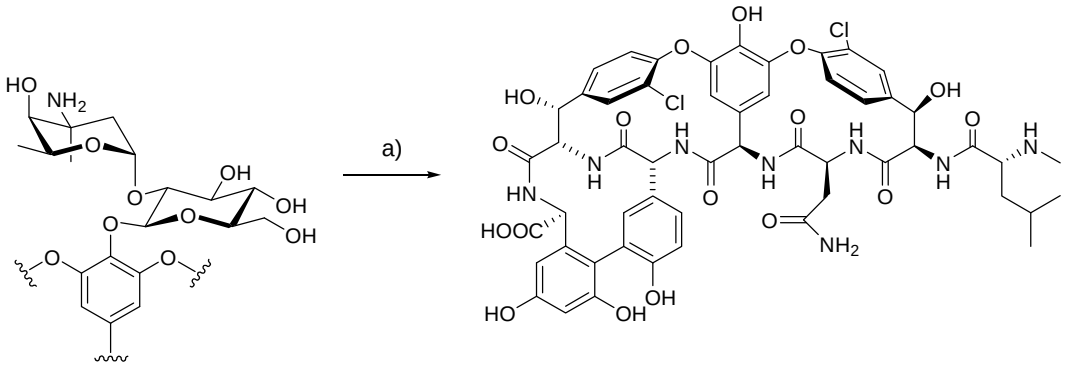
3.4.9.2. Purification and characterization of NDP-sugars. Reactions containing 2.5 mM UDP or TDP, 1 mM 2-chloro-4-nitrophenyl glycoside (**9, 34-42,** or **44**), 4.2 μM (20 μg) of OleD variant TDP-16, and 50 mM Tris (pH 8.5, total volume of 100 μL) were prepared and allowed to proceed at room temperature for 12 hours. A volume of 100 μL of ddH_2O was added to each reaction and samples were filtered through a MultiScreen Filter Plate (from Millipore, Billerica, Maryland, USA) for 2 hours at 2000 *g*. The recovered filtrate for each sample was injected onto a Gemini-NX C-18 (5 μm, 250 x 4.6 mm) column (from Phenomenex, Torrance, California, USA) with a gradient of 0% to 20% CH_3CN (solvent B) over 20 min (A = 50 mM triethylammonium acetate buffer; flow rate = 1 mL min^{-1}; A_{254} nm). Fractions corresponding to the desired products were collected, frozen, and lyophilized. Following, samples were dissolved in 1 mL of ddH_2O, frozen, and lyophilized (x3). Final products were dissolved in 1:1 ddH_2O/acetonitrile to a final concentration of 1 μg mL^{-1} and submitted for mass analysis.

3.4.10. Single and double coupled system protocols.

3.4.10.1. Evaluation of single enzyme coupled system. Reactions containing 10.5 μM (50 μg) of purified OleD variant TDP-16, 1 mM of UDP, 1 mM 4-methylumbelliferone (**58**) and 1 mM of 2-chloro-4-nitrophenyl glycoside (**9, 34-42,** or **44**) in Tris-HCl buffer (50 mM, pH 8.5) at a final volume of 100 μl were incubated in a 30 °C water bath for 24 hours. Samples were subsequently mixed with an equal volume of MeOH, centrifuged at 10,000 *g* for 30 min at 0 °C, and the supernatant removed for analysis. The clarified reaction mixtures were analyzed by analytical reverse-phase HPLC with a Gemini-NX C-18 (5 μm, 250 x 4.6 mm) column (Phenomenex, Torrance, California, USA) with a gradient of 10% B to 75% B over 20 min, 75%

B to 95% B over 1 min, 95% B for 5 min, 95% B to 10% B over 3 min, 10% B for 6 min (A = dH_2O with 0.1% TFA; B = acetonitrile; flow rate = 1 mL min^{-1}) and detection monitored at 254 nm. Fractions corresponding to the desired products were collected, frozen, lyophilized, dissolved in 1:1 acetonitrile/water to a final concentration of 1 μg mL^{-1}, and submitted for mass analysis.

3.4.10.2. Synthesis of vancomycin aglycon (60).



vancomycin

60
vancomycin aglycon

a) HCl in H_2O, 10 min

From vancomycin, **60** was prepared as described by Thompson, *et al.*[83]. Purification performed by analytical reverse-phase HPLC (see section **1.2**) yielded **60** (>98% pure by peak area) (**Figure 3.24**). HRMS-ESI (*m/z*): $[M+H]^+$ calcd for $C_{53}H_{52}Cl_2N_8O_{17}$ 1143.2900, found 1143.2889.

3.4.10.3. Evaluation of dual enzyme coupled system. All reactions were performed in a final volume of 100 μl Tris-HCl buffer (50 mM, pH 8.5) with 10.8 μM (50 μg) purified GtfE, 1 mM vancomycin aglycon (**60**) and 1 mM of 2-chloro-4-nitrophenyl glycoside (**9**, **34**-**42**, or **44**). Reactions with **35**-**38**, **40**-**41**, or **44** as donor contained 10.5 μM (50 μg) OleD variant TDP-16

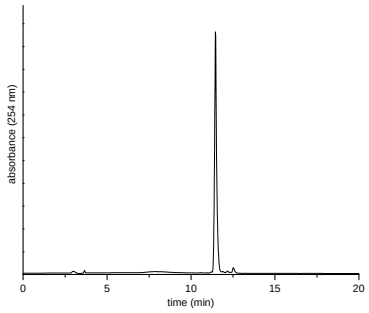


Figure 3.24. **HPLC chromatogram of vancomycin aglycon** (**60**).

and 1 mM UDP. Reactions with **9**, **34**, or **39** as donor contained 1.1 μM (5 μg) OleD variant TDP-16 and 1 μM UDP. A reaction with **42** as donor contained 0.1 μM (0.5 μg) OleD variant TDP-16 and 0.001 mM UDP. All components of the reaction(s) were added at time equals zero hours. Reactions were then incubated in a 30 °C water bath for 24 hours. Samples were then prepared and analyzed as described for the single enzyme coupled reactions (see above).

3.4.10.4 Coupled system for drug screening

3.4.10.4.1 Colorimetric evaluation of single enzyme coupled system for drug screening. Reactions containing 10.5 μM (50 μg) OleD variant TDP-16, 5 μM of UDP, 0.5 mM final acceptor (**58** [as a positive control] and **62-111**; see **Appendix 2** for structures), and 0.5 mM **9** in Tris-HCl buffer (50 mM, pH 8.5) with a final volume of 100 μl were prepared in a 96 well flat bottom Bacti plate (0.4 mL well^{-1}; Nagle Nunc International, Rochester, NY, USA). Absorbance measurements were recorded every 2 min at 410 nm for 8 hours on a FLUOstar Optima plate reader (BMG, Durham, NC, USA) with the plate shaken within the reader for 5 seconds before collection of each time point. Reactions containing final acceptor were run at n = 1, control reactions lacking final acceptor were run at n = 6 and control reactions lacking both final acceptor and UDP were run at n = 3. At 8 hours, reactions were filtered through a MultiScreen Filter Plate with a 10 kDa molecular weight cut-off (Millipore, Billerica, MA, USA) according to manufacturer's instructions at 4 °C, frozen at -20 °C, and thawed for additional analysis.

3.4.10.4.2 Data processing of colorimetric data. All absorbance data was normalized to an absorbance value of zero ($y_{k=1} = 0$) at time zero (t= 0). Area under the curve for each reaction

was determined with the following equation (*2*):

$$A = y_{k=1} + 2\left(\sum_{k=2}^{n-1} y_k\right) + y_{k=n} \qquad (2)$$

where A equals the total area of the curve and y_k equals the absorbance value at the k^{th} timepoint. The mean and standard deviation of the area under the curve for the control reactions lacking final acceptor was then determined. Reactions demonstrating positive area 3 standard deviations from the mean of the control reactions lacking final acceptor were identified as 'hits'.

3.4.10.4.3. Evaluation of single enzyme coupled system for drug screening. Following the initial screen, 'hits' (**62-103**) based upon area under the curve (see **3.4.10.4.2**) were advanced for further confirmation via HPLC and/or LC-MS. Representative time course plots of absorbance, endpoint HPLC chromatograms, and full analysis of results are presented in **Appendix 2**. LC/ESI-MS mass spectra of 'hits' were obtained using electrospray ionization in both (+) and (-) mode on an Agilent 1100 HPLC-MSD SL quadrupole mass spectrometer connected to a UV/Vis diode array detector. A 4.6 mm x 2.0 mm C18 column (Phenomenex, Torrance, CA, USA) for separation with a gradient of 3% CH_3CN (solvent B) for 1 min, 3% to 75% B over 8 min, 75% to 3% B over 1 min, 3% B for 1 min (A = ddH_2O; flow rate = 1 mL min^{-1}; A_{254} nm) were utilized for all analyses. LC/ESI-MS results are presented in **Appendix 2**.

3.5 References

1. Varki, A. *et al.* Eds., *Essentials of Glycobiology* (Cold Spring Harbor, New York, ed. 2, 2009).

2. Thibodeaux, C.J., Melançon III, C.E., & Liu, H. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**, 1008-1016 (2007).

3. Lairson, L.L., Wakarchuk, W.W., & Withers, S.G. Alternative donor substrates for inverting and retaining glycosyltransferases. *ChemComm*, 365-367 (2007).

4. Lougheed, B., Ly, H.D., Wakarchuk, W.W., & Withers, S.G. Glycosyl fluorides as substrates for nucleotide phosphosugar-dependent glycosyltransferases. *J. Biol. Chem.* **274**, 37717-37722 (1999).

5. Zhang, C. *et al.* Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science* **313**, 1291-1294 (2006).

6. Zhang, C., Albermann, C., Fu, X., & Thorson, J.S. The *in vitro* characterization of the iterative avermectin glycosyltransferase AveBI reveals reaction reversibility and sugar nucleotide flexibility. *J. Am. Chem. Soc.* **128**, 16420-16421 (2006).

7. Minami, A., Kakinuma, K., & Eguchi, T. Algycon switch approach toward unnatural glycosides from natural glycoside with glycosyltransferase VinC. *Tet. Lett.* **46**, 6187-6190 (2005).

8. Modolo, L.V., Escamilla-Treviño, L.L., Dixon, R.A, & Wang, X. Single amino acid mutations of *Medicago* glycosyltransferase UGT85H2 enhance activity and impart reversibility. *FEBS Lett.* **583**, 2131-2135 (2009).

9. Zhang, C., Moretti, R., Jiang, J., & Thorson, J.S. The *in vitro* characterization of polyene glycosyltransferases AmphDI and NysDI. *ChemBioChem* **9**, 2506-2514 (2008)**.**

10. Quirós, L.M., Carbajo, R.J., Braña, A.F., & Salas, J.A. Glycosylation of macrolide antibiotics: purification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*. *J. Biol. Chem.* **275**, 11713-11720 (2000).

11. Okada, T. *et al.* Bidirectional *N*-acetylglucosamine transfer mediated by β-1,4-*N*-acetylglucosaminyltransferase III. *Glycobiol.* **19**, 368-374 (2009).

12. Bolam, D.N. *et al.* 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 (2007).

13. Williams, G.J., Zhang, C., & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

14. Williams, G.J., Goff, R.D., Zhang, C., & Thorson, J.S. Optimizing glycosyltransferase specificity via "hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem. Biol.* **15**, 393-401 (2008).

15. Williams, G.J. & Thorson, J.S. A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Protocols* **3**, 357-362 (2008).

16. Gantt, R.W., Goff, R.D., Williams, G.J., & Thorson, J.S. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem. Int. Ed.* **47**, 8889-8892 (2008).

17. Williams, G.J., Yang, J., Zhang, C., & Thorson, J.S. Recombinant *E. coli* prototype strains for *in vivo* glycorandomization. *ACS Chem. Biol.* **6**, 95-100 (2011).

18. Yang, M. *et al.* 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 (2005)

19. Rexach, J.E., Clark, P.M., & Hsieh-Wilson, L.C. Chemical approaches to understanding *O*-GlcNAc glycosylation in the brain. *Nat. Chem. Biol.* **4**, 97-106 (2008).

20. Sakabe, K. & Hart, G.W. O-GlcNAc Transferase regulates mitotic chromatin dynamics. *J. Biol. Chem.* **285**, 34460-34468 (2010).

21. Hanson, S., Best, M., Bryan, M.C., & Wong, C. Chemoenzymatic synthesis of oligosaccharides and glycoproteins. *Trends Biochem. Sci.* **29**, 656-663 (2007).

22. Weigel, P.H. & DeAngelis, P.L. Hyaluronan synthases: a decade-plus of novel glycosyltransferases. *J. Biol. Chem.* **282**, 36777-36781 (2007).

23. Bojarová, P. *et al.* Synthesis of LacdiNAc-terminated glyconjugates by mutant galactosyltransferase – a way to new glycodrugs and materials. *Glycobiol.* **19**, 509-517 (2009).

24. Liu, R. *et al.* Chemoenzymatic design of heparin sulfate oligosaccharides. *J. Biol. Chem.* **285**, 34240-34249 (2010).

25. Lau, K. *et al.* Highly efficient chemoenzymatic synthesis of β1-4-linked galasctosides with promiscuous bacterial β1-4-galactosyltransferases. *Chem. Commun.* **46**, 6066-6068 (2010).

26. Maccioni, H.J.F., Quiroga, R., & Ferrari, M.L. Cellular and molecular biology of glycosphingolipid glycosylation. *J. Neurochem.* In press (2011) doi: 10.1111/j.1471-4159.2011.07232.x

27. Rupprath, C., Schumacher, T., & Elling, L. Nucleotide deoxysugars: essential tools for the glycosylation engineering of novel bioactive compounds. *Curr. Med. Chem.* **12**, 1637-1675 (2005).

28. Thibodeaux, C.J., Melançon III, C.E., & Liu, H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed.* **47**, 9814-9859 (2008).

29. Yang, J., Hoffmeister, D., Liu, L., Fu, X., & Thorson, J.S. Natural product glycorandomization. *Bioorg. Med. Chem.* **12**, 1577-1584 (2004).

30. Takahashi, H., Liu, Y., & Liu, H. A two-stage one-pot enzymatic synthesis of TDP-L-mycarose from thymidine and glucose-1-phosphate. *J. Am. Chem. Soc.* **128**, 1432-1433 (2006).

31. Rupprath, C., Kopp, M., Hirtz, D., Muller, R., & Elling, L. An enzyme module system for *in situ* regeneration of deoxythymidine 5'-diphosphate (dTDP)-activated deoxy sugars. *Adv. Synth. Catal.* **349**, 1489-1496 (2007).

32. Timmons, S.C., Mosher, R.H., Knowles, S.A., & Jakeman, D.L. *Org. Lett.* **9**, 857-860 (2007).

33. Wagner, G.K., Pesnot, T., & Field, R.A. A survey of chemical methods for sugar-nucleotide synthesis. *Nat. Prod. Rep.* **26**, 1172-1174 (2009).

34. Ko, H. *et al.* Molecular recognition in the P2Y14 receptor: probing the structurally permissive terminal sugar moiety of uridine-5'-diphosphoglucose. *Bioorg. Med. Chem.* **17**, 5298-5311 (2009).

35. Losey, H.C. *et al.* Incorporation of glucose analogs by GtfE and GtfD from the vancomycin biosynthetic pathway to generate variant glycopeptides. *Chem. Biol.* **9**, 1305-1314 (2002).

36. Fu, X. *et al.* Antibiotic optimization via *in vitro* glycorandomization. *Nat. Biotech.* **21**, 1467-1469 (2003).

37. Kahne, D., Leimkuhler, C., Lu, W., & Walsh, C. Glycopeptide and lipoglycopeptide antibiotics. *Chem. Rev.* **105**, 425-448 (2005).

38. Schitter, G. & Wrodnigg, T.M. Update on carbohydrate-containing antibacterial agents. *Expert Opin. Drug Discov.* **4**, 315-356 (2009).

39. Wagner, G.K. & Pesnot, T. Glycosyltransferases and their assays. *ChemBioChem* **11**, 1939-1949 (2010).

40. Kittl, R., & Withers, S.G. New approaches to enzymatic glycoside synthesis through directed evolution. *Carbohydr. Res.* **345**, 1272-1279 (2010).

41. Pesnot, T., Palcic, M.M., & Wagner, G.K. A novel fluorescent probe for retaining galactosyltransferases. *ChemBioChem* **11**, 1392-1398 (2010).

42. Schumacher, G. & Neuhaus, P. The physiological estrogen metabolite 2-methoxyestradiol reduces tumor growth and induces apoptosis in human solid tumors. *J. Cancer Res. Clen. Oncol.* **127**, 405-410 (2001).

43. Taldone, T., Sun, W., & Chiosis, G. Discovery and development of heat shock protein 90 inhibitors. *Bioorg. Med. Chem.* **17**, 2225-2235 (2009).

44. Levitzki, A. Protein tyrosine kinase inhibitors as novel therapeutic agents. *Pharmacol. Ther.* **82**, 231-239 (1999).

45. Krystof, V. & Uldrijan, S. Cyclin-dependent kinase inhibitors as anticancer drugs. *Curr. Drug Targets* **11**, 291-302 (2010).

46. Chen, X., *et al.* Shikonin, a component of chinese herbal medicine, inhibits chemokine receptor function and suppresses human immunodeficiency virus type 1. *Antimicrob. Agents Chemother.* **47**, 2810-2816 (2003).

47. Codd, R., Braich, N., Liu, J., Soe, C.Z., & Pakchung, A.A. Zn(II)-dependent histone deacetylase inhibitors: suberoylanilide hydroxamic acid and trichostatin A. *Int. J. Biochem. Cell Biol.* **41**, 736-739 (2009).

48. Bisht, S., Brossart, P., Maitra, A., & Feldmann, G. Agents targeting the Hedgehog pathway for pancreatic cancer treatment. *Curr. Opin. Investig. Drugs* **11**, 1378-1398 (2010).

49. Li, L., *et al.* The skin cancer chemotherapeutic agent ingenol-3-angelate (PEP005) is a substrate for the epidermal multidrug transporter (ABCB1) and targets tumor vasculature. *Cancer Res.* **70**, 4509-4519 (2010).

50. Joerger, M., Omlin, A., Cerny, T., & Früh, M. The role of pemetrexed in advanced non small-cell lung cancer: special focus on pharmacology and mechanism of action. *Curr. Drug Targets* **11**, 37-47 (2010).

51. Shelley, M.D., Cleves, A., Wilt, T.J., & Mason, M.D. Gemcitabine chemotherapy for the treatment of metastatic bladder carcinoma. *BJU Int.* **108**, 168-179 (2011).

52. Baur, J.A., & Sinclair, D.A. Therapeutic potential of resveratrol: the *in vivo* evidence. *Nature Rev. Drug Discov.* **5**, 493–506 (2006).

53. Calderón-Montaño, J.M., Burgos-Morón, E., Pérez-Guerrero, C., & López-Lázaro, M. A review on the dietary flavonoid kaempferol. *Mini Rev. Med. Chem.* **11**, 298-344 (2011).

54. Anadón, A., & Reeve-Johnson, L. Macrolide antibiotics, drug interactions and microsomal enzymes: implications for veterinary medicine. *Res. Vet. Sci.* **66**, 197-203 (1999).

55. McMurry, L.M., Oethinger, M., & Levy, S.B. Triclosan targets lipid synthesis. *Nature* **394**, 531-532 (1998).

56. Lacy, C.F., Armstrong, L.L., Goldman, M.P., & Lance, L.L. (eds.) *Drug Information Handbook*, 16th Ed. (Lexi-Comp, Hudson, Ohio, USA, 2008).

57. Bartlett, J.G., Sutter, V.L., & Finegold, S.M. Treatment of Anaerobic Infections with Lincomycin and Clindamycin. *N. Engl. J. Med.* **287**, 1006-1010 (1972).

58. Eastman, R.T., & Fidock, D.A. Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. *Nat. Rev. Microbiol.* **7**, 864-874 (2009).

59. De Celercg, E. Antiviral drugs in current clinical use. *J. Clin. Virol.* **30**, 115-133.

60. Fitzpatrick, D.W., Picken, C.A., Murphy, L.C., & Buhr, M.M. Measurement of the relative binding affinity of zearalenone, alpha-zearalenol and beta-zearalenol for uterine and oviduct estrogen receptors in swine, rats and chickens: an indicator of estrogenic potencies. *Comp. Biochem. Physiol. C* **94**, 691-694 (1989).

61. McKenna, G.J., *et al.* Limiting hepatitis C virus progression in liver transplant recipients using sirolimus-based immunosuppression. *Am. J. Transplant* **11**, 2379-2387 (2011).

62. Berecek, K.H. & Brody, M.J. Evidence for a neurotransmitter role for epinephrine derived from the adrenal medulla. *Am. J. Physiol.* **242**, H593-H601 (1982).

63. Hou, W.C., Lin, R.D., Chen, C.T., & Lee, M.H. Monoamine oxidase B (MAO-B) inhibition by active principles from Uncaria rhynchophylla. *J. Ethnopharmacol.* **100**, 216-220 (2005).

64. Jackson, H., *et al.* Influence of ursodeoxycholic acid on the mortality and malignancy associated with primary biliary cirrhosis: a population-based cohort study. *Hepatology* **46**, 1131-1137 (2007).

65. Romero, P.A. & Arnold, F.H. Exploring protein fitness landscapes by directed evolution. *Nature Rev. Mol. Cell Biol.* **10**, 866-876 (2009).

66. Williams, G.J., Gantt, R.W., & Thorson, J.S. The impact of enzyme engineering upon natural product glycodiversfication. *Curr. Opin. Chem. Biol.* **12**, 556-564 (2008).

67. Hancock, S.M., Vaughan, M.D., & Withers, S.G. Engineering of glycosidases and glycosyltransferases. *Curr. Opin. Chem. Biol.* **10**, 509-519 (2006).

68. Zhang, B. *et al.* Golgi nucleotide sugar transporter modulates cell wall biosynthesis and plant growth in rice. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5110-5115 (2011).

69. Dickmanns, A. *et al.* Structural basis for the broad substrate range of the UDP-sugar pyrophosphorylase from *Leishmania major*. *J. Mol. Biol.* **405**, 461-478 (2011).

70. Ahmed, A. *et al.*, Colchicine glycorandomization influences cytotoxicity and mechanism of action. *J. Am. Chem. Soc.* **128**, 14224-14225 (2006).

71. Lee, Y.S. *et al*., Pratical β-stereoselective *O*-glycosylation of phenols with penta-*O*-acetyl-β-D-glucopyranose. *J. Carbohydr. Chem*. **20**, 503-506 (2001).

72. Carrière, D. *et al*., Phase transfer catalysis toward the synthesis of *O*-, *S*-, *Se*- and *C*-glycosides. *J. Mol. Cat. A. Chem*. **154**, 9-22 (2000).

73. Bridiau, N., Benmansour, M., Legoy, M.D., & Maugard, T. One-pot stereoselective synthesis of β-*N*-aryl-glycosides by *N*-glycosylation of aromatic amines: application to the synthesis of tumor-associated carbohydrate antigen building blocks. *Tetrahedron*. **63**, 4178-4183 (2007).

74. Goff, R.D. & Thorson, J.S. Assessment of chemoslective neoglycosylation methods using chlorambucil as a model. *J. Med. Chem.* **53**, 8129-8139 (2010).

75. Bae, J. *et al*. A pratical enzymatic synthesis of UDP sugars and NDP glucoses. *ChemBioChem* **6**, 1963-1966 (2005).

76. Tsuboi, K.K., Fukunaga, K., & Petricciani, J.C. Purification and specific kinetic properties of erythrocyte uridine diphosphate glucose pyrophosphorylase. *J. Biol. Chem.* **244**, 1008-1015 (1969).

77. Whiters, S.J., MacLennan, D.J., & Street, I.P. The synthesis and hydrolysis of a series of deoxyfluoro-D-glucopyranosyl phosphates. *Carbohydr. Res*. **154**, 127-144 (1986).

78. Maunier, V., Boullanger, P., Lafont, D., & Chevalier, Y. Synthesis and surface-active properties of amphiphilic 6-aminocarbonyl derivatives of D-glucose. *Carbohydr. Res*. **299**, 49-57 (1997).

79. Goddard-Borger, E.D., & Stick, R.V. The synthesis of various 1,6-disulfide-bridged D-hexopyranoses. *Aust. J. Chem.* **58**, 188-198 (2005).

80. Hirschmann, R. *et al*., De novo design and synthesis of somatostatin non-peptide peptidomimetics utilizing β-D-glucose as a novel scaffolding. *J. Am. Chem. Soc.* **115**, 12550-12568 (1993).

81. Kleine, H. P., & Sidhu, R.S. Synthesis of aryl β-D-mannopyranosides. *Carbohydr. Res*. **182**, 307-312 (1988).

82. Kröger, L., & Thiem, J. Synthesis and evaluation of glycosyl donors with novel leaving groups for transglycosylations employing β-galactosidase from bovine teste. *Carbohydr. Res*. **342**, 467-481 (2007).

83. Thompson, C., Ge, M., & Kahne, D. Synthesis of vancomycin from the aglycon. *J. Am. Chem. Soc.* **121**, 1237-1244 (1999).

Chapter 4:

Engineering of a Glycosyltransferase for Combinatorial Catalysis of NDP-sugars

4.1. Introduction

Nucleotide diphosphate sugars (NDP-sugars) are critically important to living organisms because of their role as substrates for Leloir glycosyltransferases (GTs), which are involved in a variety of essential processes including energy storage, carbohydrate metabolism, protein glycosylation, cell wall biosynthesis, and the synthesis of bioactive natural products[1, 2]. Although there are many reported chemical, enzymatic, and chemoenzymatic routes to this class of molecules (see Chapter 1 and references therein)[3, 4], their synthesis and isolation remains a non-trivial matter due to the low solubility of sugar nucleotides in traditional organic solvents, the presence of multiple polar and/or charged functional groups, and the susceptibility of both glycosidic and pyrophosphate bonds to hydrolytic cleavage[3]. As a result, the poor availability of sugar nucleotides continues to limit investigations into the innate role of glycosylation as a whole in biology[3], the elucidation of glycosylated natural product biosynthetic pathways[2, 3, 5], and high impact applications such as drug discovery[4, 6, 7].

Recently, great advances have been made in accessing NDP-sugars through glycosyltransferase-catalyzed reverse reactions (see Chapter 3)[8]. These breakthroughs have demonstrated milligram quantity synthesis of NDP-sugars, *in situ* formation and utilization of NDP-sugars without the usual tedious *a priori* isolation, and simple, high throughput colorimetric screens for NDP-sugar formation and general glycosyl transfer. These advancements are largely dependent upon two factors: ***i)*** utilizing thermodynamically favorable

2-chloro-4-nitrophenyl glycosides as donors and ***ii)*** identification of GT enzymes capable of recognizing these donors in reverse reactions for catalysis of NDP-sugar formation.

The current 'bottle-neck' for this technology platform is the identification of catalysts for generation of unique NDP-sugars. While formation of 22 natural and non-natural NDP-sugars with 11 distinct 2-chloro-4-nitrophenyl glycoside donors was recently reported with a single GT enzyme[8], this number is dwarfed by the number and variety of NDP-sugars utilized throughout biology as a whole[2-5]. Certain sugars, specifically 3- or 4-deoxy-D-glucose and D-allose (the C-3 epimer of D-glucose) were reported as relatively poor substrates (when compared to the native sugar substrate, D-glucose) and led to low yields unless numerous equivalents of NDP were utilized to drive the GT-catalyzed reverse reaction. Others, such as D-mannose (the C-2 epimer of D-glucose) or L-sugars, were not recognized as substrates. In terms of the NDPs recognized, while UDP and TDP were recognized as excellent substrates, CDP, ADP, and GDP were reported as either very poor or non-substrates[8]. Additionally, sugars with deoxygenation at multiple positions or those modified by amino-, *N*-methyl, *O*-methyl, or *C*-methyl groups were not initially investigated. As these chemical moieties and motifs, especially dideoxygenation, are found throughout sugars appended to medicinally relevant natural products[2, 4-6], it is imperative to demonstrate their utilization in GT-catalyzed reverse reactions.

With the limitations described above in mind, a study was undertaken to develop a general catalyst for NDP-sugar production in GT-catalyzed reverse reactions. The identification of such a catalyst would provide the proof-of-concept that the simple colorimetric screen provided by 2-chloro-4-nitrophenyl glycosides could be utilized for rapid optimization of GT variants for the generation of virtually any desired NDP-sugar pair and also expand the number and variety of NDP-sugars available for biochemical investigations and applications. The

selected parental catalyst, the GT OleD variant TDP-16, has been previously engineered in the context of forward GT-catalyzed reactions for increases in both NDP-sugar and aglycon recognition, promiscuity, and proficiency[9] and characterized as the most advantageous variant identified in GT-catalyzed reverse reactions[8]. The engineering effort described herein should be considered a first step towards providing facile access to a wide range of NDP-sugars with unique motifs and chemical modifications.

4.2. Results and Discussion

4.2.1. TDP-16 Saturation Libraries.

4.2.1.1 Library Design. Positions targeted for saturation mutagenesis were selected by comparing positions identified through previous engineering efforts[9-12], OleD variant characterization reports[8, 13], and a study of the wild type OleD crystal structure[14]. While other substrates have been co-crystallized with wild type OleD[14], binding position(s) of the putative phenyl glycoside donors is/are unknown and difficult to model due to OleD's relatively large donor binding site. Therefore, positions for saturation within the donor site were selected based upon a number of criteria which included: ***i)*** positions within characterized OleD variants with enhanced activity toward small, phenolic substrates[8-13] ***ii)*** positions capable of interacting with those from the first group ***iii)*** positions closest to the desosamine moiety of the co-crystallized substrate erythromycin which participates in forward OleD catalyzed reactions[14, 15] and ***iv)*** positions capable of interacting with residue H20, a known catalytic participant[14]. The UDP acceptor has been co-crystallized with OleD and its contacts are well defined. To account for the

presence of a full NDP-sugar into account, the ligand UDP-2-deoxy-2-fluoroglucose from UGT72B1 (a GT possessing a similar GT-B fold to that of OleD; PDB ID 2VCE)[16] was modeled into the OleD NDP binding site. Positions that contacted the nucleotide base, diphosphate bridge, and sugar moiety (which would be present following a GT-catalyzed reverse reaction) on this modeled substrate were selected for saturation. Special care was taken to ensure that the substrate binding pockets for both the putative 2-chloro-4-nitrophenyl glycoside donors (**1**-**9**) and NDP acceptors were equally represented (**Figure 4.1**, **4.2**).

4.2.1.2. Creation of Saturation Libraries. Using standard molecular biology techniques, a total of 15 saturation libraries were generated based upon the parental sequence of the gene *oleD*[P67T/S132F/A242L/Q268V] (generates the OleD variant TDP-16)[9] were created. A total of 269 out of a possible 285 variants (94% coverage) and confirmed by DNA sequencing (see **Appendix 3**). Single isolates of *E. coli* BL21(DE3)pLysS transformed with pET28a vectors containing these individual sequence variants and a number of previously reported OleD variants[9-12] were utilized for the creation of 3 separate 96-well master plates (see **Appendix 3**) for high throughput screening.

4.2.1.3. Saturation Library Expression Test. To assess the overall expression and solubility of the generated TDP-16 saturation libraries, strains containing empty pET28a vector, pET28a/*oleD*[P67T/S132F/A242L/Q268V], or plasmid variants possessing mutations (either alanine or glycine) at each of the targeted saturation positions within the parental gene sequence were selected for expression testing. An analysis of the purified variants by SDS-PAGE demonstrated that mutations can be introduced into each of the 15 selected saturation positions

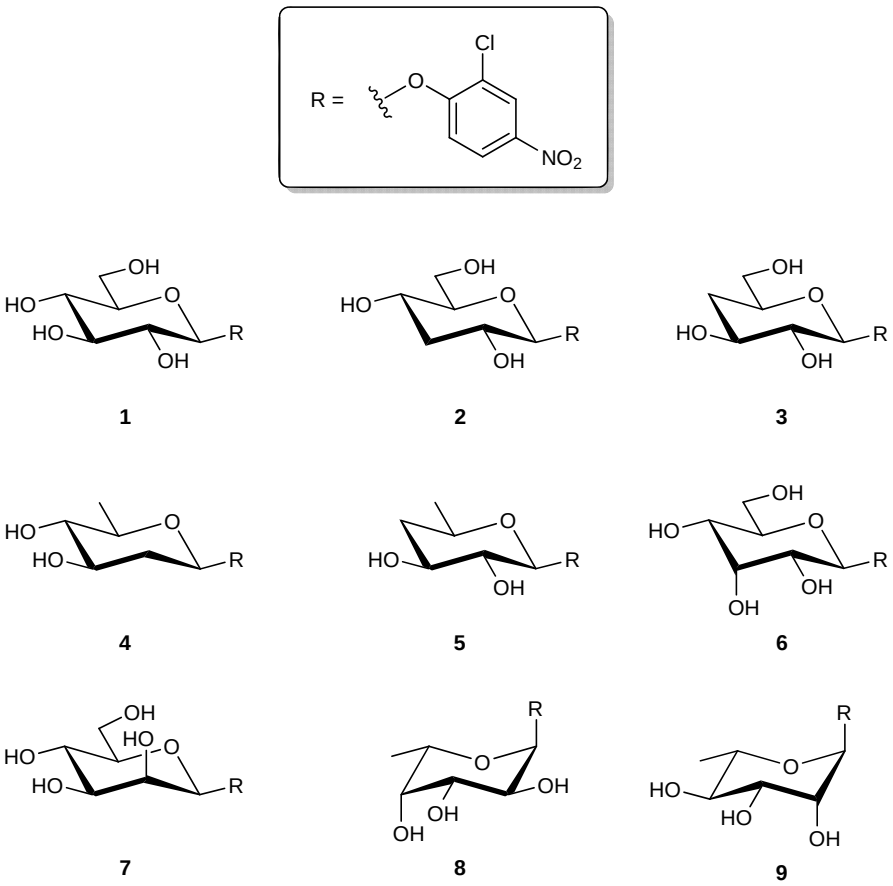


Figure 4.1. Structures of 2-chloro-4-nitrophenyl glycoside donors used in this study.

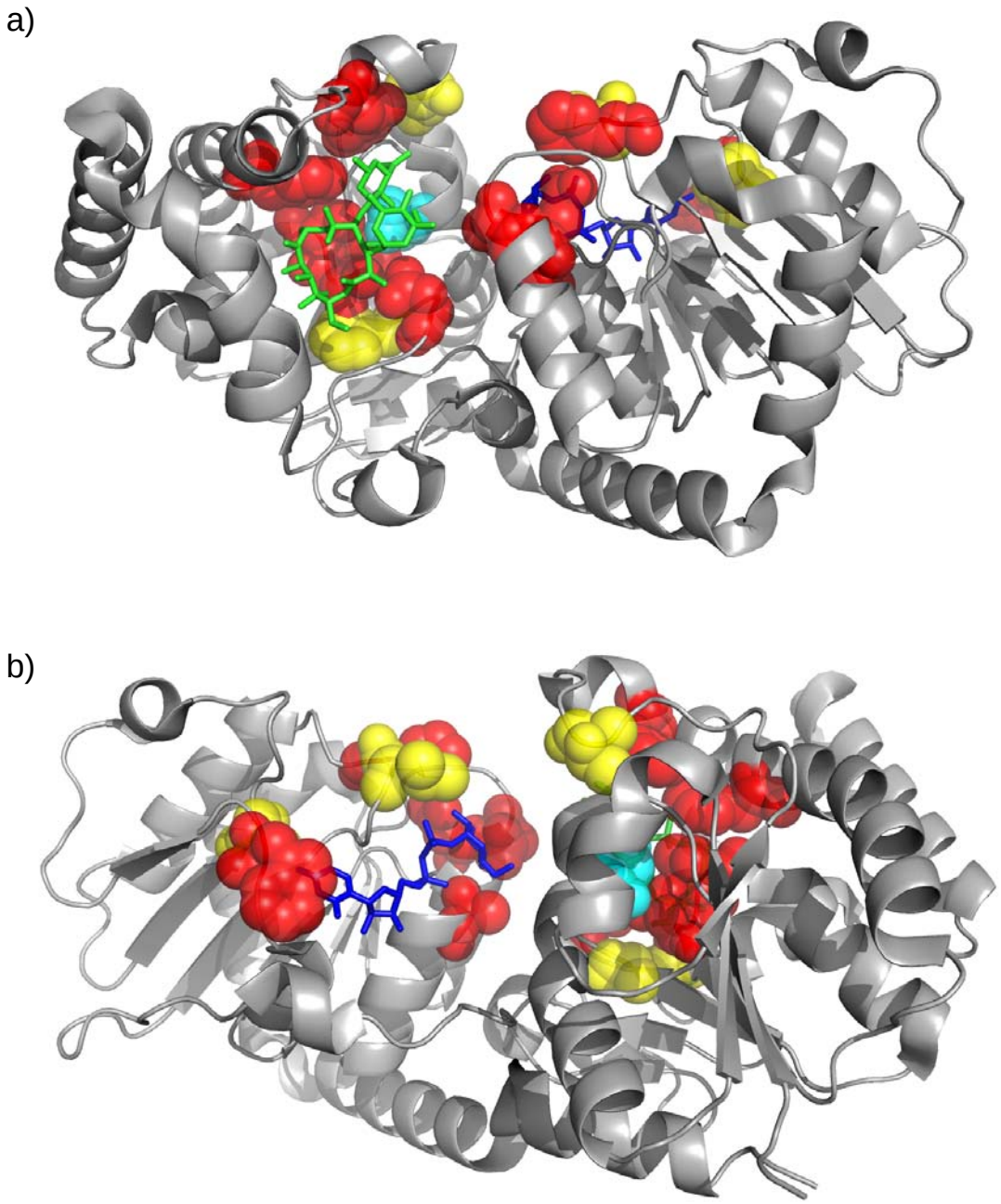


Figure 4.2. Crystal **structure representation of positions selected for saturation mutagenesis.** In the context of a GT-catalyzed reverse reaction, the different views highlight the **a)** 'donor' pocket or **b)** 'acceptor' pocket of the OleD variant TDP-16. All of the existing mutations within the TDP-16 variant (P67T/S132F/A242L/Q268V) were selected for saturation mutagenesis and modeled into the wild type crystal structure (highlighted in yellow). Additional mutations selected for mutagenesis are highlight in red. Erythromycin (co-crystallized) and UDP-2-fluoroglucose (modeled) are highlighted in green and blue, respectively. The catalytic participant residue H20 is highlighted in cyan.

without detriment to the overall expression or solubility of the enzyme (**Figure 4.3**).

4.2.2. High Throughput Screening. Using the previously reported colorimetric assay for NDP-sugar formation (see Chapeter 3)[8], master plates containing the saturation libraries were grown, induced for protein expression, processed, and the resulting cell lysate screened against a total of 13 distinct substrate combinations in reverse GT reactions for generation of NDP-sugars. There were two categories of screens: ***i)*** 2-chloro-4-nitrophenyl β-D-glucopyranoside (**1**) as donor and various NDPs (UDP, TDP, CDP, ADP, or GDP) as acceptors and ***ii)*** various 2-chloro-4-nitrophenyl glycosides (**2**-**9**) as donors and UDP as acceptor (**Figure 4.1**). Following the addition of cell lysate to the prepared reaction mixtures, absorbance at 410 nm was monitored over a period of several hours. Representative data from a single assay of an individual master plate were provided in **Figure 4.4** and **4.5**. No activity was observed in screens with **7**-**9** as donor and UDP as acceptor. Following statistical analysis of the primary data, a subset of hits identified as top performers across multiple assays or substrates types was selected for additional screening. from the substrate combinations for which positive response was observed in the high throughput assays.

4.2.3. NDP-sugar Product Confirmation. Before moving forward with the hits identified from the primary screen, the parental catalyst TDP-16 was investigated for its ability to generate previously unconfirmed NDP-sugar combinations. Confirmation of these previously uncharacterized NDP-sugar products ensured that observed increases in absorbance from the high throughput screen were attributable to desired NDP-sugar formation and not from undesired hydrolytic events. Formation of all 10 expected NDP-sugar products (**10**-**19**) by the parental catalyst TDP-16 was confirmed with a combination of HPLC

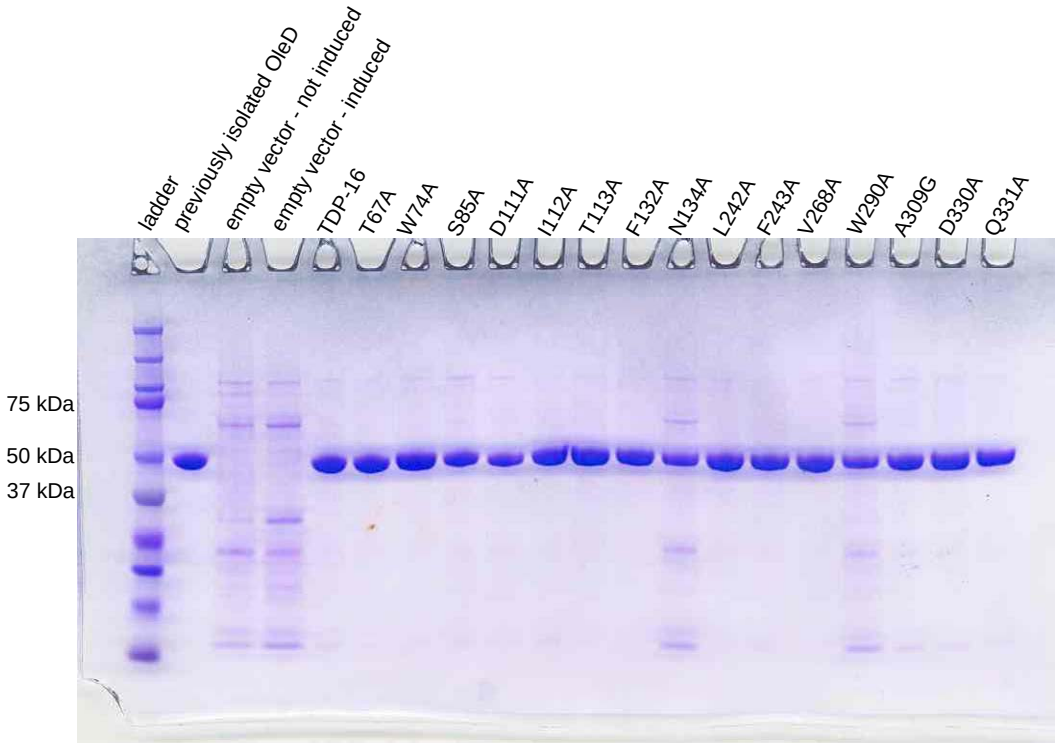


Figure 4.3. SDS-PAGE analysis of selected library variants. The expected molecular weight of TDP-16 (and variants) is 48 kDa and correlated with the observed bands. The label above each well is the amino acid mutation introduced into the TDP-16 parental gene sequence of *oleD*[P67T/S132F/A242L/Q268V].

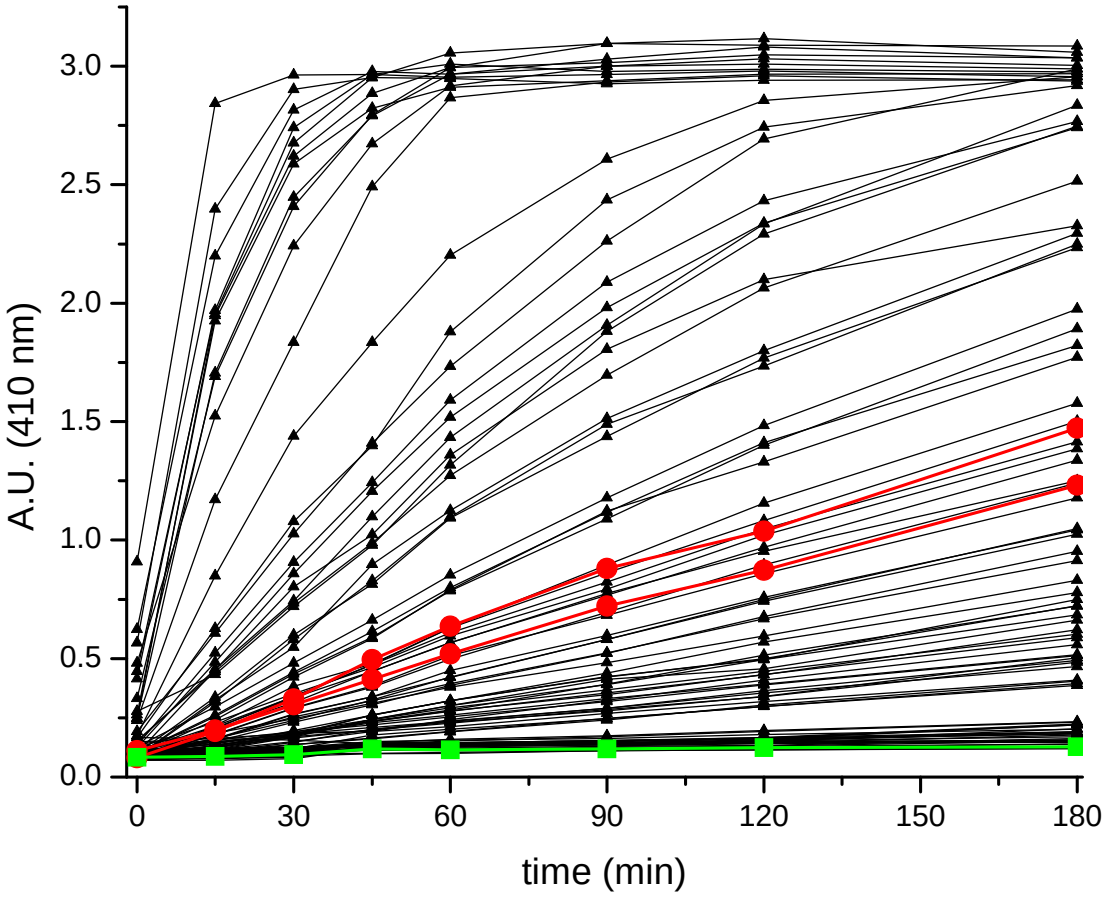


Figure 4.4. Representative data of variants screened in real time. Variants present on master plate X (see **Appendix 3**) were screened with 2-chloro-4-nitrophenyl-β-D-glucoside (**1**; 0.5 mM) and UDP (0.5 mM) as final acceptor for formation of UDP-glucose (**10**). Absorbance measurements (410 nm) were taken at various time points for 3 hours and were uncorrected. Empty vector control (n = 1), TDP-16 parental catalyst (n = 2), and individual catalyst variants (n = 1) correspond to green squares, red circles, and black triangles, respectively. A.U., absorbance units.

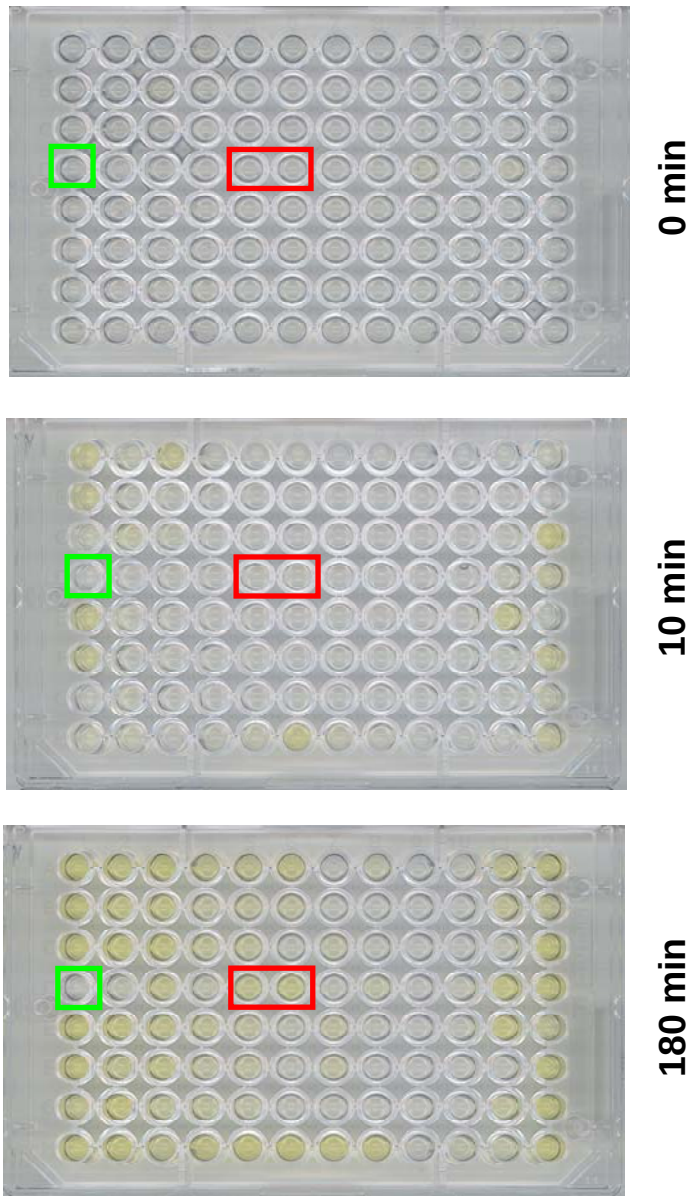


Figure 4.5. Photographic representation of variants screened in real time. Variants present on master plate X were screened with 2-chloro-4-nitrophenyl-β-D-glucoside (**1**; 0.5 mM) and UDP (0.5 mM) as final acceptor for formation of UDP-glucose (**10**). Photographs were taken at 0 (before catalyst addition), 10, and 180 minutes. Wells corresponding to the empty vector control and the TDP-16 parental catalyst are highlighted in green and red squares, respectively. Photographs were taken on an HP Scanjet 7400c (Hewlett-Packard Company, Palo Alto, CA, USA) and utilized without color correction.

and HRMS-ESI techniques (**Table 4.1**).

4.2.4. Secondary Screening. A total of 19 'hits' identified from the primary screen (7% of variants screened) were expressed and purified for secondary screening on the 0.5 L scale (see section **4.4**). Analysis by SDS-PAGE demonstrated the isolated enzymes to exhibit purity of >95% homogeneity (**Figure 4.6**). All purified catalysts were screened against a total of 10 assays with either 2-chloro-4-nitrophenyl-β-D-glucopyranoside (**1**) as donor and various NDPs as acceptors or various 2-chloro-4-nitrophenyl glycosides (**2-6**) as donors and UDP as acceptor. Following the addition of equal amounts of catalyst, the absorbance at 410 nm for each reaction was followed for a minimum of 3 hours.

Statistical analysis of the initial slopes for each assay set revealed four specific mutations (I112P, T113M, L242I, and V268Q) in the context of the TDP-16 parental sequence to be most advantageous overall towards the formation of the targeted NDP-glucose and/or UDP-sugar combinations.

4.2.5. Recombination and Tertiary Screening. The four mutations identified through secondary screening were recombined in the context of the TDP-16 parental sequence to yield 8 additional variants (see section **4.4**). The recombinants were expressed and purified on the 0.5 L scale (see section **4.4**) and analyzed by SDS-PAGE for purity (**Figure 4.7**). Following, the hits from the primary screen (*i.e.*, I112P, T113M, L242I, and V268Q) and all of the obtained recombinant catalysts were screened against a total of 10 assays with either 2-chloro-4-nitrophenyl-β-D-glucopyranoside (**1**) as donor and various NDPs as acceptors or various 2-

Entry	Compound	Retention Time (min)	Elemental Composition $[M-2H]^{2-}$	HRMS-ESI (*m/z*)	
				Calculated Mass $[M-2H]^{2-}$	Observed Mass $[M-2H]^{2-}$
10	UDP-α-D-glucose†	8.9	$C_{15}H_{22}N_2O_{17}P_2^{2-}$	282.02020	282.0191
11	TDP-α-D-glucose†	10.9	$C_{16}H_{24}N_2O_{16}P_2^{2-}$	281.03060	281.0302
12	CDP-α-D-glucose	17.3	$C_{15}H_{23}N_3O_{16}P_2^{2-}$	281.52822	281.5285
13	ADP-α-D-glucose	16.1	$C_{16}H_{23}N_5O_{15}P_2^{2-}$	293.53384	293.5341
14	GDP-α-D-glucose	13.2	$C_{16}H_{23}N_5O_{16}P_2^{2-}$	301.53130	301.5314
15	UDP-3-deoxy-α-Dglucose†	11.0	$C_{15}H_{22}N_5O_{16}P_2^{2-}$	274.02280	274.0228
16	UDP-4-deoxy-α-D-glucose†	9.4	$C_{15}H_{22}N_5O_{16}P_2^{2-}$	274.02280	274.0234
17	UDP-2,6-dideoxy-α-D-glucose	12.9	$C_{16}H_{22}N_2O_{15}P_2^{2-}$	266.02532	266.0255
18	UDP-4,6-dideoxy-α-D-glucose	14.1	$C_{15}H_{22}N_2O_{15}P_2^{2-}$	266.02532	266.0254
19	UDP-α-D-allose†	9.2	$C_{15}H_{22}N_2O_{17}P_2^{2-}$	282.02020	282.0205

Table 4.1. Characterization data for NDP-sugar products generated by TDP-16. HPLC retention times reported are those observed with phosphate buffer and triethylammonium bisulfate and the respective method utilized for their isolation. Retention times for CDP, ADP, GDP, and UDP were 31.2, 18.7, 17.4, and 15.8 minutes, respectively. †Previously characterized from TDP-16 enzymatic reactions by both HPLC and HRMS-ESI (retention times of UDP and TDP from this previous study were 15.2 and 16.0, respectively)[8].

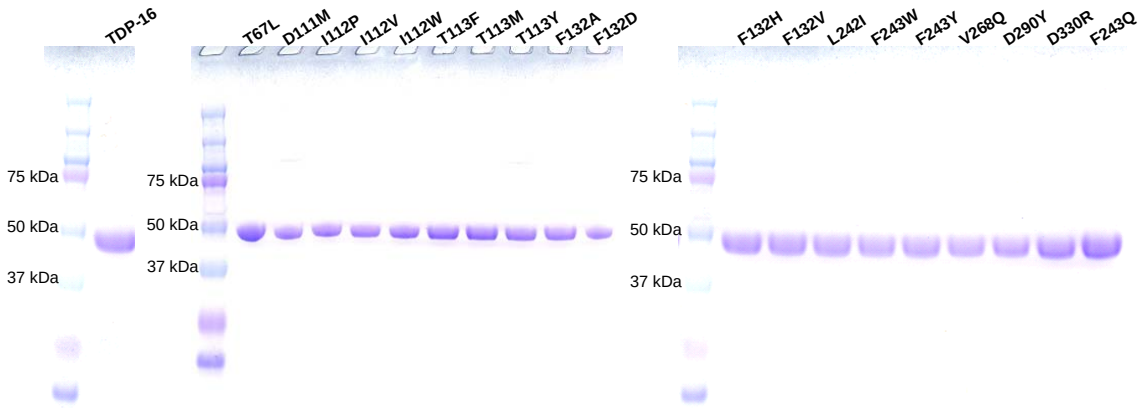


Figure 4.6. SDS-PAGE analysis of 'hits' purified for secondary screening. The mutation listed above each well is the amino acid mutation introduced into the parental gene sequence of *oleD*[P67T/S132F/A242L/Q268V] (generates OleD variant TDP-16). The expected molecular weight for TDP-16 (and variants) is 48 kDa and is consistent with the observed bands. The molecular weight marker for each gel is labeled at left.

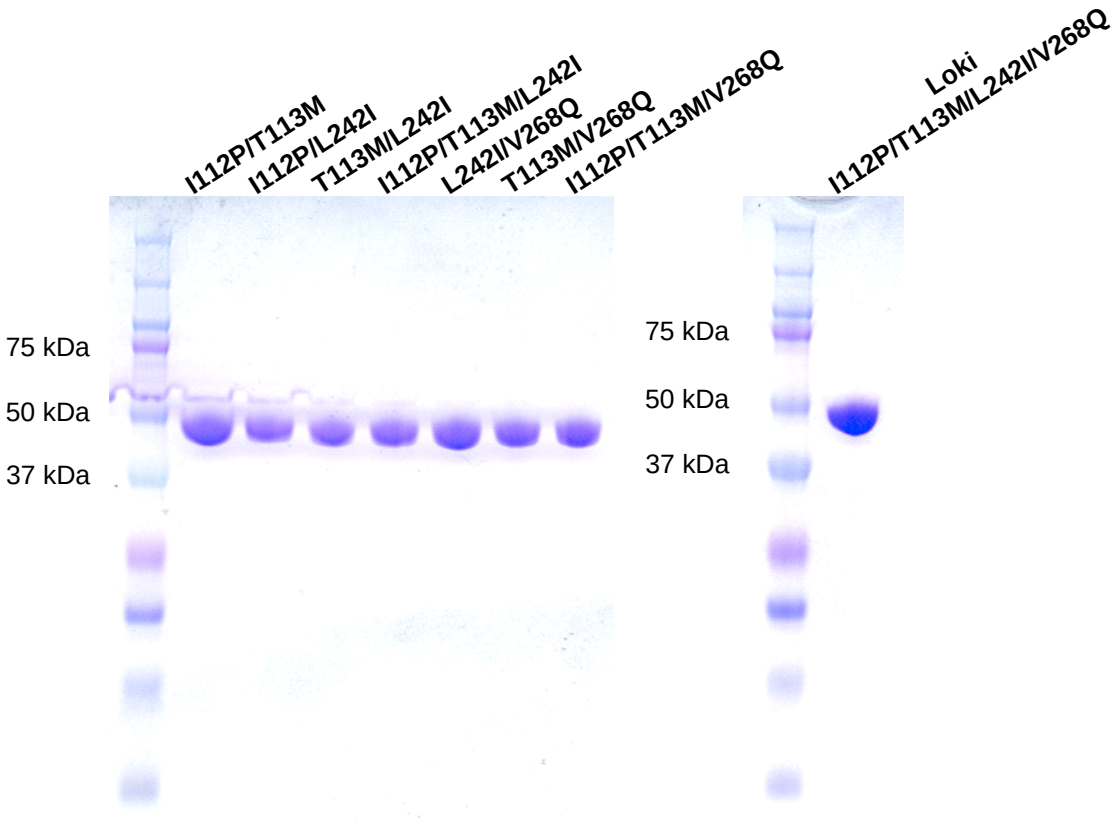


Figure 4.7. SDS-PAGE analysis of purified recombinants for tertiary screening. The mutations listed above each well are the amino acid mutations introduced into the parental gene sequence of *oleD*[P67T/S132F/A242L/Q268V] (generates OleD variant TDP-16). The expected molecular weight for all variants is 48 kDa and is consistent with the observed bands. The molecular weight marker for each gel is labeled at left.

chloro-4-nitrophenyl glycosides (**2-6**) as donors and UDP as acceptor. Following the addition of equal amounts of catalyst, the absorbance at 410 nm for each reaction was monitored for a minimum of 3 hours.

Comparison of the initial slopes for each reaction revealed that a few combinations of mutations in the context of the TDP-16 variant generated the best catalyst for particular combinations of NDP-glucose (*e.g.*, I112P/T113M for formation of UDP-glucose (**10**)) or UDP-sugar (*e.g.*, T113M/V268Q for the formation of UDP-4-deoxyglucose (**16**)). However, the recombined variant introducing the mutations I112P, T113M, L242I, and V268Q into the TDP-16 parental backbone (hereafter referred to as the Loki variant; named after the Norse god of mischief and trickery due to this variant's promiscuities toward both glycoside donors and NDP acceptors) displayed the best overall catalytic advantage for the formation of the targeted NDP-glucose and/or UDP-sugar combinations (**Figure 4.8**, **4.9**). This was especially clear in the recognition of purine-based NDPs (**Figure 4.8**). While the goal of this engineering effort was to identify the best overall general catalyst for NDP-sugar formation, these results demonstrated that the best catalyst in terms of promiscuity may not provide the optimum proficiency towards a specific target. Overall, when compared to the OleD wild type sequence, the Loki variant contains a total of 5 mutations (P67T/I112P/T113M/S132F/A242I; **Figure 4.10**). The possible contributions of each mutation are discussed below.

4.2.6 Discussion of Identified Mutations. The P67T mutation was initially identified from a random mutagenesis library screened in the context of a GT-catalyzed forward reaction[10] and has been analyzed in additional reports and characterizations of OleD variants[8-13]. Within the reported OleD crystal structure[14], residue 67 is situated within a loop region (amino acids 60-76,

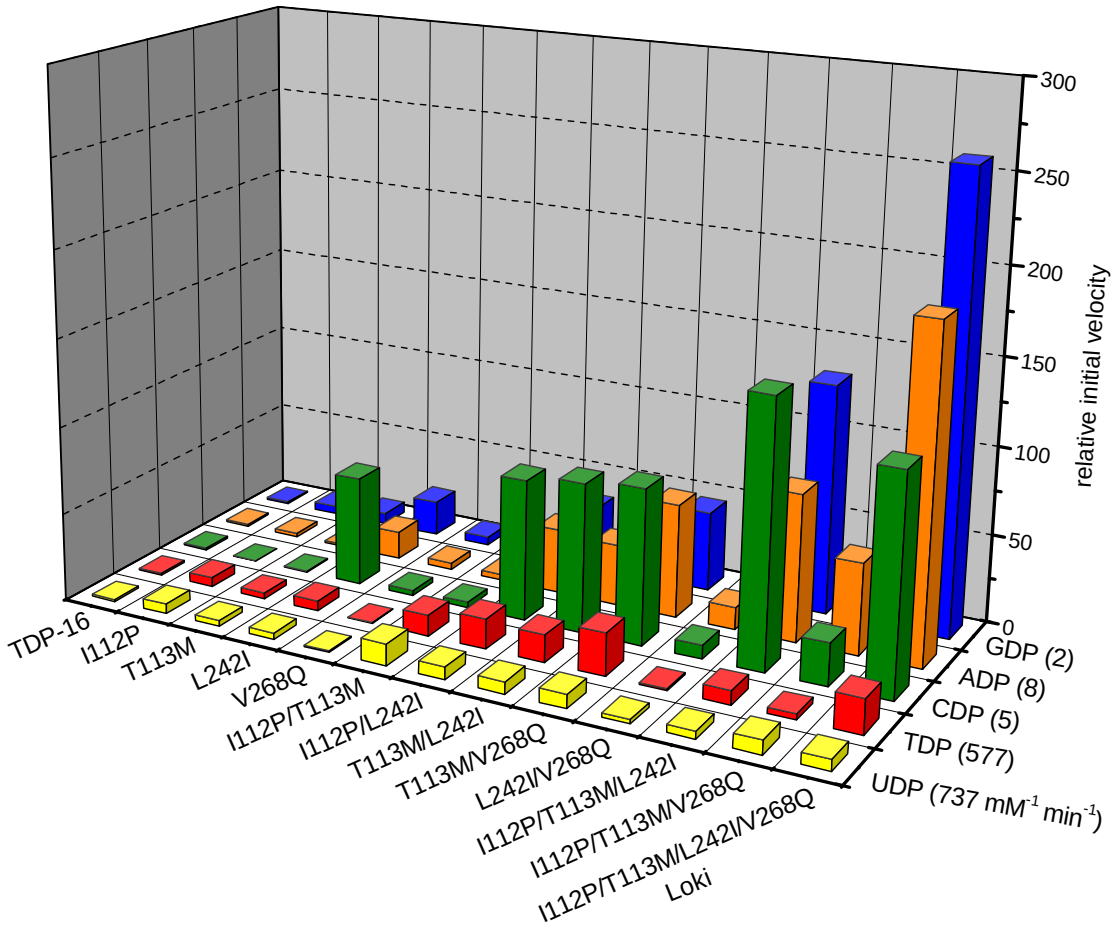


Figure 4.8. Tertiary screening results with NDPs. Relative initial velocities of OleD variants with 2-chloro-4-nitrophenyl glucoside (**1**) as donor and NDPs as acceptors. The relative initial velocity for TDP-16 was set equal to 1 for all NDPs. The label for each variant are the amino acid mutation(s) introduced into the parental gene sequence of *oleD*[P67T/S132F/A242L/Q268V] (the template for OleD variant TDP-16). The determined $k_{cat}(K_m)^{-1}$ for TDP-16 with **1** saturating and the respective NDP varied are listed after each NDP label. Standard deviations of absorbance readings for TDP-16 (n = 2) typically varied by < 5% for each data point over the course of each assay.

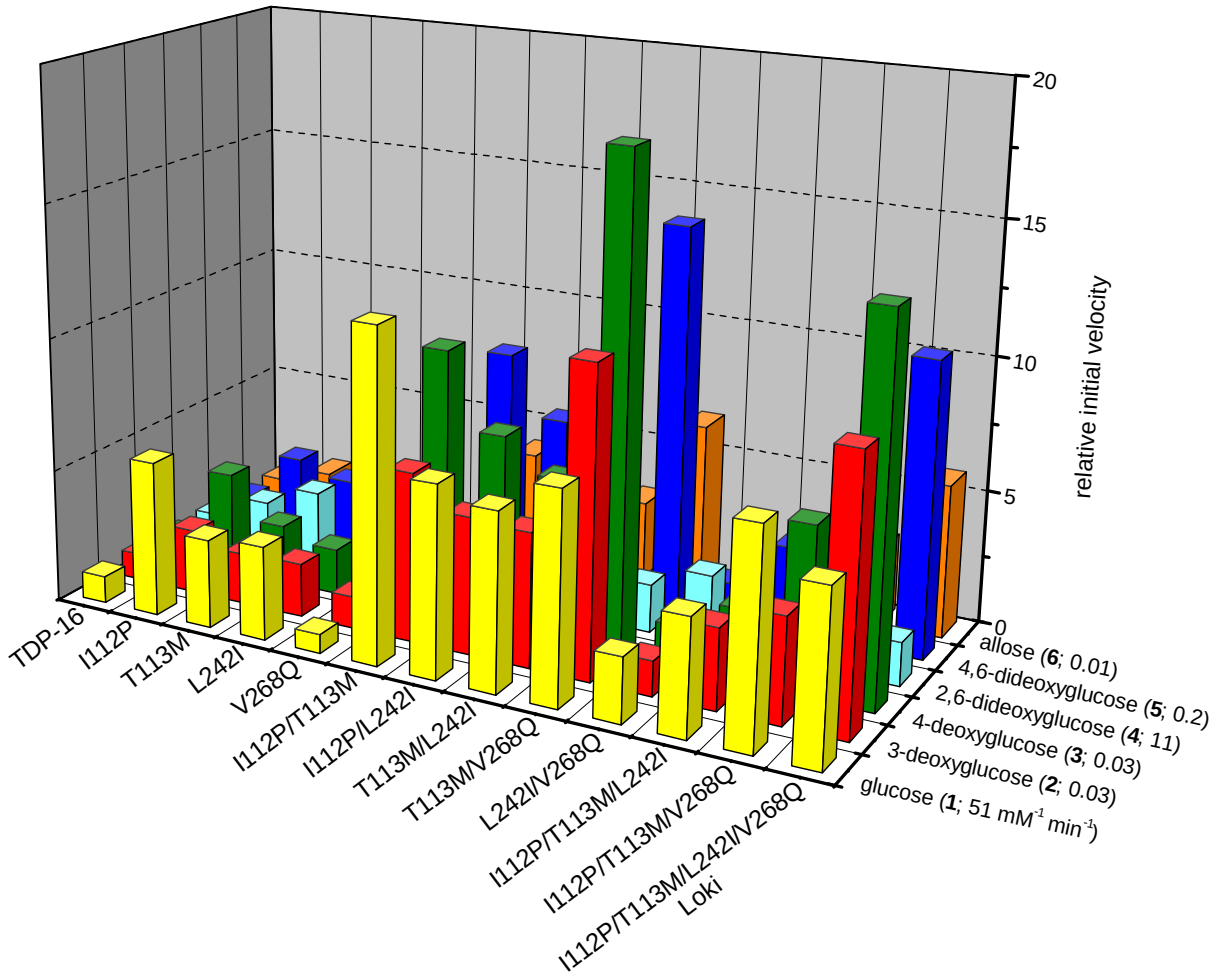


Figure 4.9. Tertiary screening results with varied sugar donors. Relative initial velocities of OleD variants with various 2-chloro-4-nitrophenyl glycoside donors (**1**-**6**) and UDP as acceptor. The relative initial velocity for TDP-16 was set equal to 1 for all glycoside donors. The mutations listed are the amino acid mutations introduced into the parental gene sequence of *oleD*[P67T/S132F/A242L/Q268V] (generates OleD variant TDP-16). The determined $k_{cat}(K_m)^{-1}$ for TDP-16 with UDP saturating and the respective 2-chloro-4-nitrophenyl glycoside (**1**-**6**) varied are listed after the label of each donor sugar. Standard deviations of absorbance readings for TDP-16 (n = 2) typically varied by < 5% for each data point over the course of each individual assay.

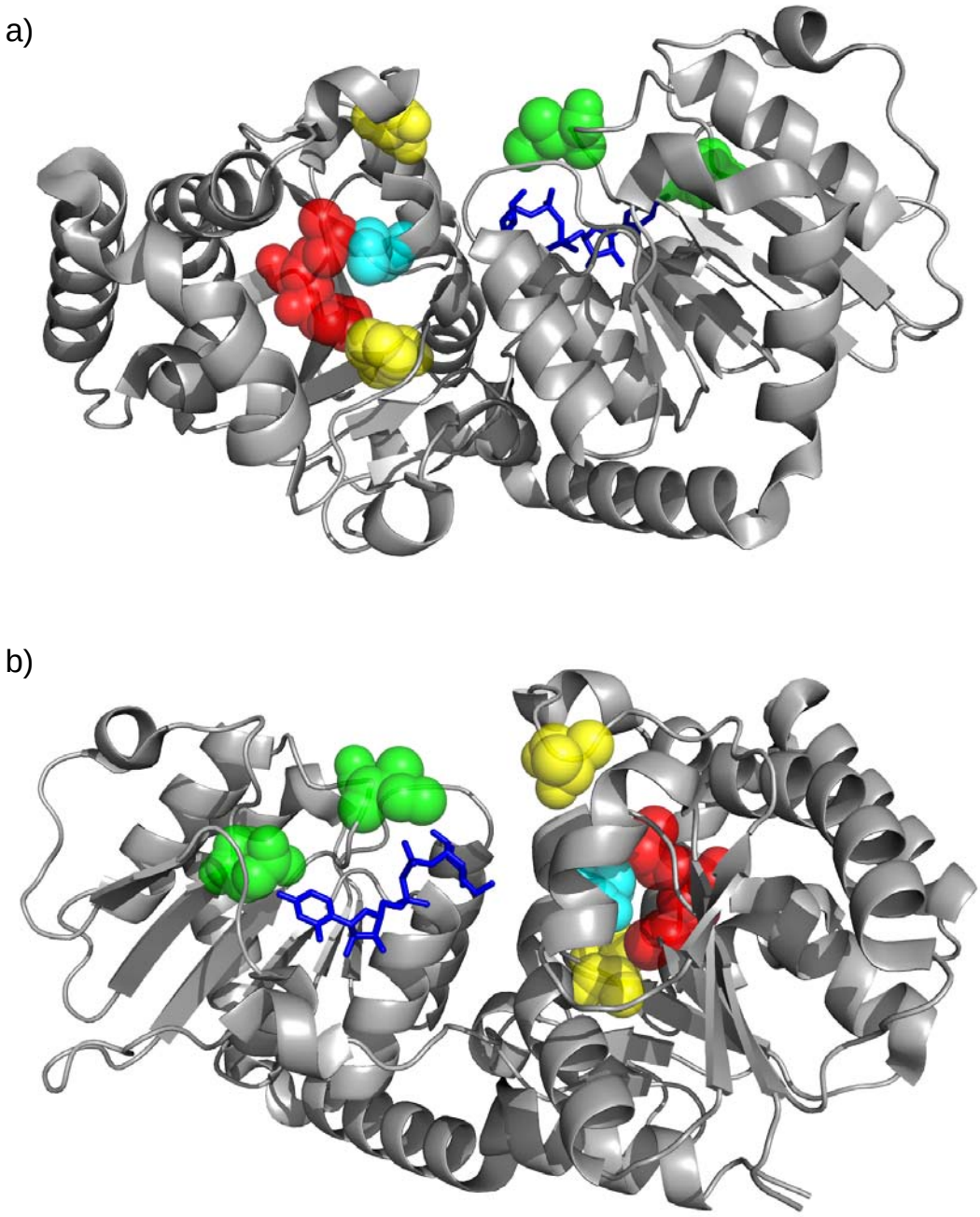


Figure 4.10. Crystal structure representation of mutations within the OleD variant Loki. In the context of a GT-catalyzed reverse reaction, the different views highlight the **a)** 'donor' pocket or **b)** 'acceptor' pocket of the OleD variant Loki. All of the differences between wild type OleD and the Loki variant (*i.e.*, P67T/I112P/T113M/S132F/A242I) were modeled into the wild type crystal structure. Residues 67 and 132 possess identical mutations in both the TDP-16 and Loki variants (highlighted in yellow). Residues 242 and 268 are mutated to different amino acids in the TDP-16 and Loki variants (highlighted in green). Residues 112 and 113 are mutated only in the Loki variant (highlighted in red). UDP-2-fluoroglucose (modeled) and the catalytic residue H20 are highlighted in blue and cyan, respectively.

loop N3), which is hypervariable in other GTs possessing the GT-B fold and contributes to forming the 'donor' site in the context of GT-catalyzed reverse reactions[10, 17]. Investigations of other GTs have highlighted additional positions within the N3 loop which are responsible for defining enzyme specificity[18].

Mutations at position 112 have been characterized in several previous studies[9-11]. While the single mutation I112T was initially reported as non-beneficial in the context of a GT-catalyzed forward reaction for the glycosylation of 4-methylumbelliferone[10], additional investigations and saturation mutagenesis reported substantial activity gains (i.e., k_{cat}/K_m) for the glycosylation of novobiocic acid with both I112T and I112K[11]. In the context of a GT-catalyzed reverse reaction, residue 112 is located deep within the 'donor' pocket, and the isoleucine within the crystallized wild type enzyme provides direct interaction with the co-crystallized erythromycin and may indirectly influence the positioning of the catalytic residue H20 by interacting with position 113. The mutation of position 112 from isoleucine to proline in this study was interesting in several regards. First, the distinctive cyclic structure of proline may impart exceptional conformational rigidity, limiting the conformational changes available to the peptide backbone in neighboring residues. As a consequence, there may be substantial positioning changes of surrounding residues which would include residues 113 and, indirectly, residue H20, a catalytic participant. Secondly, the planar nature of proline at residue 112 may be providing additional hydrophobic surface for the binding of the investigated phenolic glycoside donors within the active site.

Advantageous mutations at position 113 within OleD have not been previously reported. Mutation of this residue from threonine to methionine could substantially increase the hydrophobic nature of the binding pocket. This mutation, with its longer side chain, is likely

directly affecting the positions of residue 112 and the catalytic residue H20. Additionally, the sulfur within methionine, while not traditionally viewed as a participant in hydrogen bond formation, can potentially act as a weak hydrogen bond acceptor to generate a range of interactions with surrounding residues or substrates[19, 20]. These potential interactions are further complicated by the synergy observed for GT-catalyzed reverse reactions when both I112P and T113M were introduced into the parental sequence of TDP-16 (**Figure 4.8**, **4.9**).

The alteration of position 132 from serine to phenylalanine was initially reported by Williams, *et al.*[10] and the same alteration was present in a number of other characterized OleD variants[8, 9, 11-13]. Comparison to the sequence and structure of the plant flavonoid GT VvGT1 reveals a likely role for Ser132 in binding the sugar transferred during the reaction[21, 22]. The VvGT1 equivalent to the OleD Ser132 (Thr141) forms a hydrogen bond with the C-6 hydroxyl group of UDP-2-deoxy-2-fluoroglucose. Lack of a hydrogen bond donor/acceptor at this position may explain the large increases in activity observed in previous studies[13] and this investigation with sugars either lacking a hydroxyl group or modified at C-6.

Mutation of residue Ala242 has been reported in several cases to be advantageous in the context of both forward and reverse GT-catalyzed reactions[8-13]. All of the reported advantageous mutations at this position (valine, leucine, and isoleucine) have been hydrophobic in nature. In the OleD crystal structure, Ala242 follows Ser241, which forms a hydrogen bond to the α-phosphate of the UDP acceptor in the context of a GT-catalyzed reverse reaction (this interaction was also observed in the VvGT1 structure)[10, 21]. Hydrophobic mutations at this position most likely affect binding of the NDP acceptor and/or sugar moiety after transfer from a 2-chloro-4-nitrophenyl glycoside donor.

The residue Q268 interacts directly with the nitrogenous base of UDP co-crystallized in the wild type OleD crystal structure. The Q268V mutation within the TDP-16 variant was originally identified in a screen for utilization of TDP-glucose (**11**) in forward GT-catalyzed reactions[9]. However, reversion of the valine at 268 in the TDP-16 sequence back to the wild type residue glutamine is beneficial to GT-catalyzed reverse reactions, especially with purine NDPs and the various glycoside donors screened in this study (**Figure 4.8, 4.9**). Until more information is obtained regarding the binding of small phenolic glycosides within GTs, providing in depth information on the specific interactions generated by these advantageous mutations within both substrate binding pockets remains beyond the scope of these investigations.

4.2.7. Kinetic Evaluation. Kinetic evaluations comparing the parental catalyst TDP-16 and the final Loki variant were performed for all of the screened substrates (**Figure 4.11-4.14**, **Table 4.2**, **4.3**). Overall the greatest increases in efficiency were observed towards the purine-based NDPs. In cases where 2-chloro-4-nitrophenyl glucoside (**1**) was saturating and ADP or GDP was varied (for formation of **13** and **14**, respectively), both substantial increases and decreases in k_{cat} and K_m, respectively, contributed to improving the catalytic efficiency >400-fold when directly comparing TDP-16 and Loki (**Table 4.2**). Interestingly, there was a substantial change to the rank order of the substrates when comparing k_{cat}/K_m across these variants (UDP>TDP>ADP>CDP>GDP for TDP-16; ADP>TDP>CDP>UDP>GDP for Loki). The largest contributing factors to this change in rank order include: ***i)*** only a doubling of k_{cat} was observed under variable UDP while all of the other determined k_{cat}s with variable NDP increased by at least an order of magnitude and ***ii)*** the minimum 5-fold decrease in K_m observed toward CDP, ADP, and GDP while the K_m for both UDP and TDP slightly increased (**Table 4.2**).

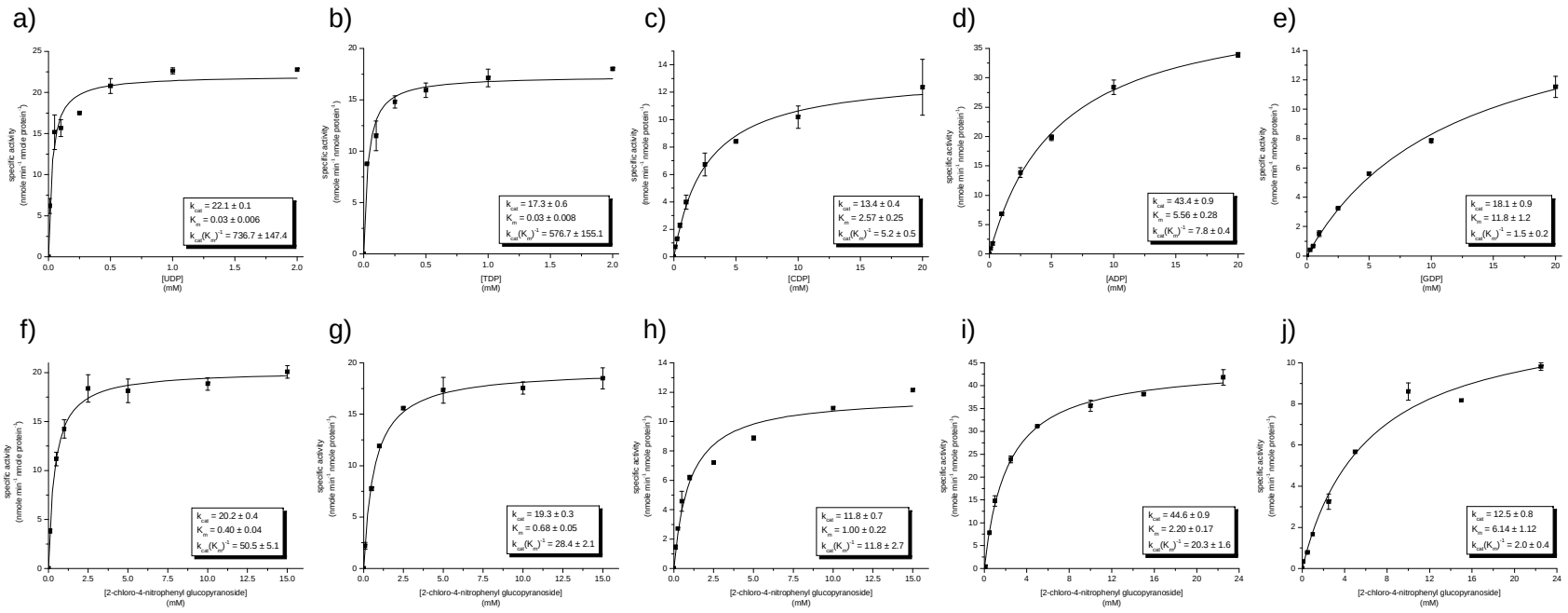


Figure 4.11. Determination of kinetic parameters for OleD variant TDP-16. Graphs represent **a)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with UDP varied (0 to 2 mM), **b)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with TDP varied (0 to 2 mM), **c)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with CDP varied (0 to 20 mM), **d)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with ADP varied (0 to 20 mM), **e)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with GDP varied (0 to 20 mM), **f)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-15 mM), **g)** TDP saturating (2 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-15 mM), **h)** CDP saturating (20 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-15 mM), **i)** ADP saturating (20 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-22.5 mM), **j)** GDP saturating (20 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-22.5 mM). Error bars denote standard deviation.

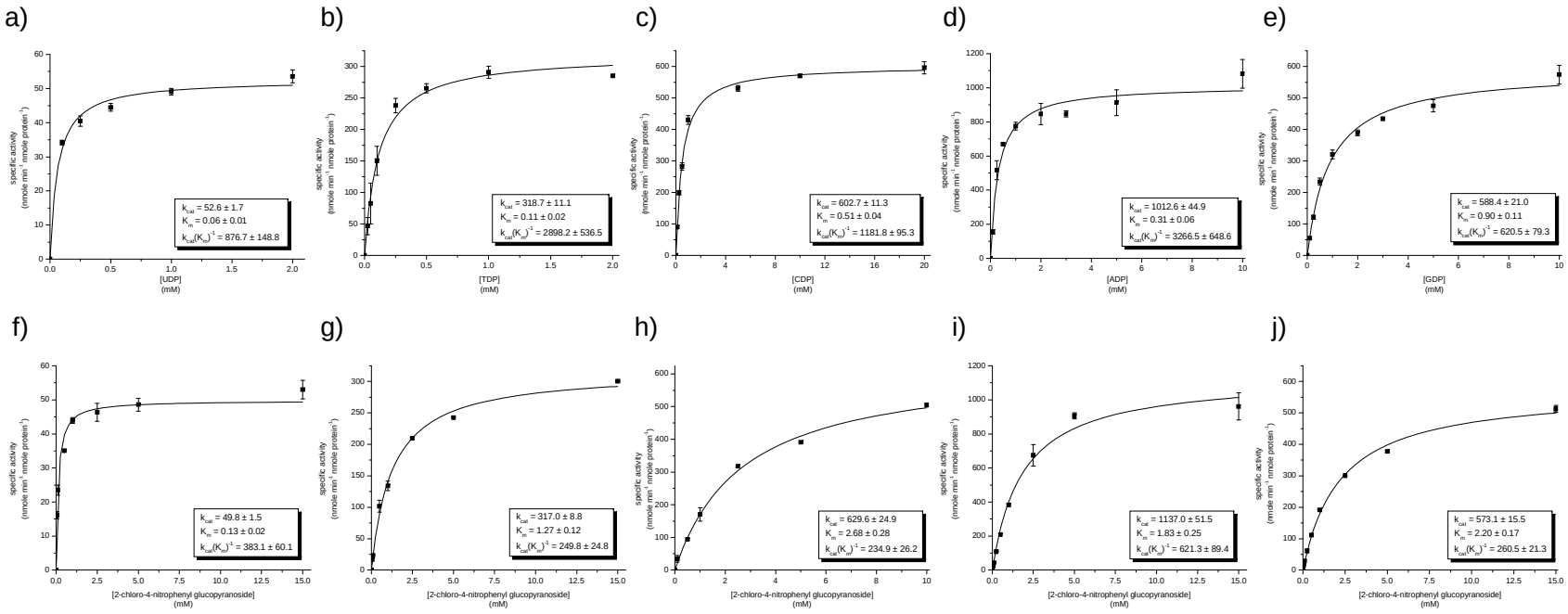


Figure 4.12. Determination of kinetic parameters for OleD variant Loki. Graphs represent **a)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with UDP varied (0 to 2 mM), **b)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with TDP varied (0 to 2 mM), **c)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with CDP varied (0 to 20 mM), **d)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with ADP varied (0 to 10 mM), **e)** 2-chloro-4-nitrophenyl glucopyranoside (**1**) saturating (15 mM) with GDP varied (0 to 10 mM), **f)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-15 mM), **g)** TDP saturating (2 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-15 mM), **h)** CDP saturating (20 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-15 mM), **i)** ADP saturating (10 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-22.5 mM), **j)** GDP saturating (10 mM) with 2-chloro-4-nitrophenyl glucopyranoside (**1**) varied (0-22.5 mM). Error bars denote standard deviation.

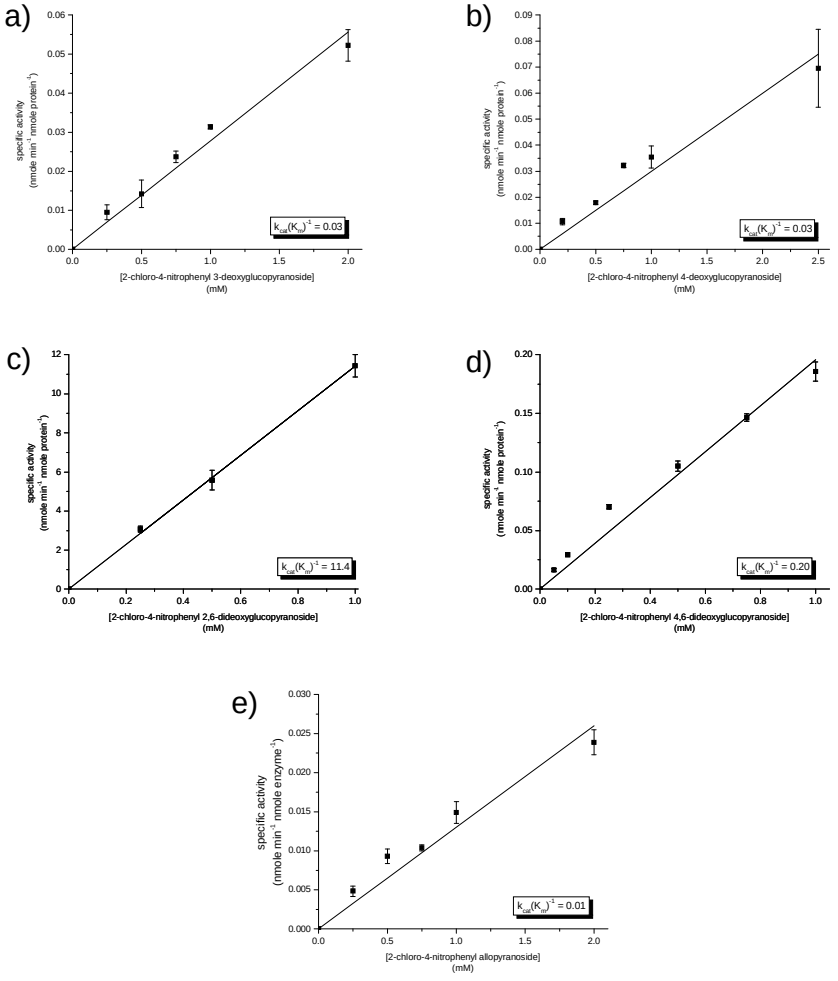


Figure 4.13. Determination of kinetic parameters for OleD variant TDP-16 with various sugar donors. Graphs represent **a)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 3-deoxyglucopyranoside (**2**) varied (0-2 mM), **b)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 4-deoxyglucopyranoside (**3**) varied (0-2.5 mM), **c)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 2,6-dideoxyglucopyranoside (**4**) varied (0-1 mM), **d)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 4,6-dideoxyglucopyranoside (**5**) varied (0-1 mM), **e)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl allopyranoside (**6**) varied (0-2 mM). Error bars denote standard deviation.

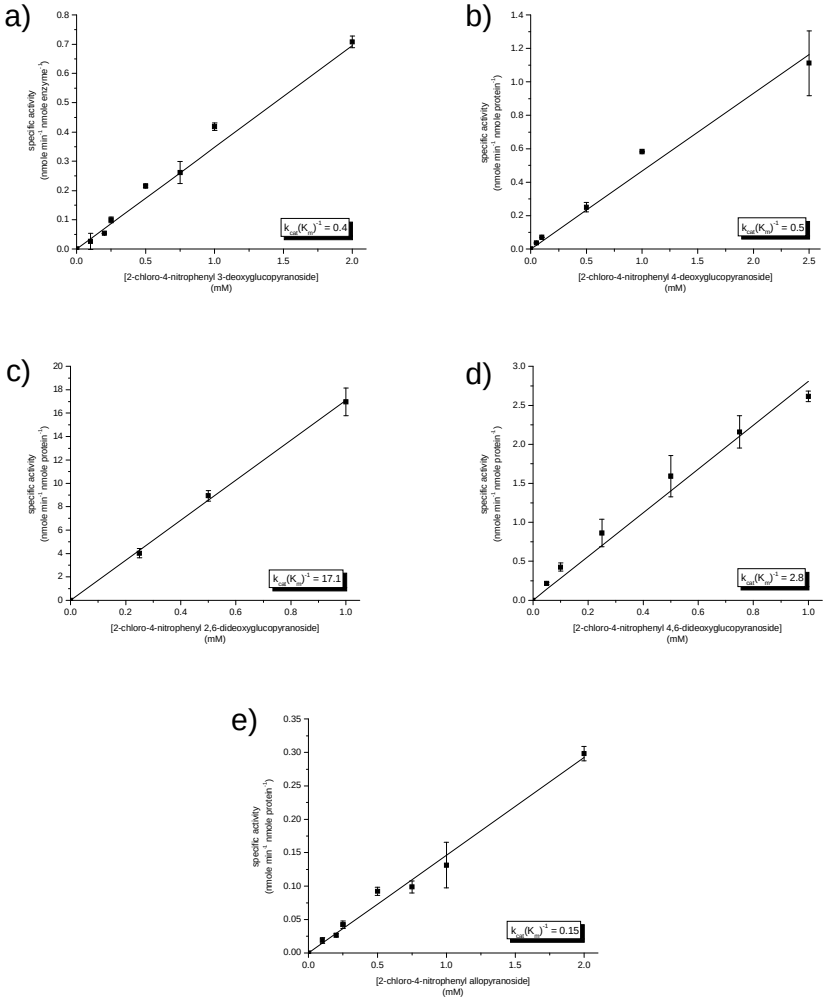


Figure 4.14. Determination of kinetic parameters for OleD variant Loki with various sugar donors. Graphs represent **a)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 3-deoxyglucopyranoside (**2**) varied (0-2 mM), **b)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 4-deoxyglucopyranoside (**3**) varied (0-2.5 mM), **c)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 2,6-dideoxyglucopyranoside (**4**) varied (0-1 mM), **d)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl 4,6-dideoxyglucopyranoside (**5**) varied (0-1 mM), **e)** UDP saturating (2 mM) with 2-chloro-4-nitrophenyl allopyranoside (**6**) varied (0-2 mM). Error bars denote standard deviation.

Enzyme	Saturating Substrate	Varied Substrate	k_{cat} $(min)^{-1}$	K_m (mM)	k_{cat} / K_m $(mM^{-1} min^{-1})$	$(k_{cat} / K_m)_{Loki} / (k_{cat} / K_m)_{TDP-16}$
TDP-16	UDP	**1**	20.2 ± 0.4	0.40 ± 0.04	50.5 ± 5.1	---
	TDP	**1**	19.3 ± 0.3	0.68 ± 0.05	28.4 ± 2.1	---
	CDP	**1**	11.8 ± 0.7	1.00 ± 0.22	11.8 ± 2.7	---
	ADP	**1**	44.6 ± 0.9	2.20 ± 0.17	20.3 ± 1.6	---
	GDP	**1**	12.5 ± 0.8	6.14 ± 1.12	2.0 ± 0.4	---
	1	UDP	22.1 ± 0.1	0.03 ± 0.006	736.7 ± 147.4	---
	1	TDP	17.3 ± 0.6	0.03 ± 0.008	576.7 ± 155.1	---
	1	CDP	13.4 ± 0.4	2.57 ± 0.25	5.2 ± 0.5	---
	1	ADP	43.4 ± 0.9	5.56 ± 0.28	7.8 ± 0.4	---
	1	GDP	18.1 ± 0.9	11.80 ± 1.18	1.5 ± 0.2	---
Loki	UDP	**1**	49.8 ± 1.5	0.13 ± 0.02	383.1 ± 60.1	**7.6**
	TDP	**1**	317.0 ± 8.8	1.27 ± 0.12	249.6 ± 24.6	**8.8**
	CDP	**1**	629.6 ± 24.9	2.68 ± 0.28	234.9 ± 26.2	**19.9**
	ADP	**1**	1137.0 ± 51.5	1.83 ± 0.25	621.3 ± 89.4	**30.6**
	GDP	**1**	573.1 ± 15.5	2.20 ± 0.17	260.5 ± 21.3	**128.0**
	1	UDP	52.6 ± 1.7	0.06 ± 0.01	876.7 ± 148.8	**1.2**
	1	TDP	318.8 ± 11.1	0.11 ± 0.02	2898.2 ± 536.5	**5.0**
	1	CDP	602.7 ± 11.3	0.51 ± 0.04	1181.8 ± 95.3	**226.7**
	1	ADP	1012.6 ± 44.9	0.31 ± 0.06	3266.5 ± 648.6	**418.5**
	1	GDP	588.41 ± 21.0	0.90 ± 0.11	620.5 ± 79.3	**404.5**

Table 4.2. Determined kinetic parameters for the OleD variants TDP-16 and Loki. Parameters were determined with 2-chloro-4-nitrophenyl glucopyranoside (1) as donor and various NDPs as acceptors.

Modest increases in catalytic efficiency were observed for all sugar substrates. The largest improvements were observed in substrates containing chemical motifs (i.e., **2, 3, 6**) that are characterized with TDP-16 as relatively poor substrates when compared to the parental sugar D-glucose (**Table 4.3**)[8]. There was no substantial change to the rank order of the substrates screened when comparing k_{cat}/K_m for TDP-16 and Loki as the parental sugar D-glucose remains the preferred substrate for the Loki variant by at least an order of magnitude in these investigations.

4.2.8. Combinatorial NDP-sugar Production. To demonstrate the true combinatorial potential of the Loki variant as a general catalyst for NDP-sugar formation, 6 glycoside donors and 5 NDPs were mixed in all available combinations, and the formation of the expected sugar was evaluated by HPLC and HRMS techniques. Amazingly, formation of all 30 expected NDP-sugar products was confirmed (**Figure 4.15**, **4.16**, **Table 4.4**). The observed product peaks appeared in a time dependent manner by HPLC and did not possess retention times similar to 2-chloro-4-nitrophenolate, NDP, or the respective 2-chloro-4-nitrophenyl glycoside donor. Additionally, all NDP-sugar products display similar absorbance profiles over the 200-800 nm wavelength range when compared to their respective NDP (data not shown).

As a comparison, a number of routes for the synthesis of TDP-2,6-dideoxy-α-D-glucose (**22**) have been previously reported. Beginning from 2-deoxy-D-glucose, Minami *et al.* reported a 9 step chemical synthesis of **22** with a 2.5% overall isolated yield[23]. From D-glucose-6-phosphate, Amann *et al.* reported a 3 step synthesis (1 chemical, 2 enzymatic) requiring 5 enzymes with a 15% overall yield and a putative mixture of C-4 epimers[24]. Most recently, Wang *et al.* reported a 3 step synthesis from α-D-glucose-1-phosphate and TMP requiring 8 enzymes from 4 separate organisms, 2 cofactors (NAD^+, and NADPH), and 2 separate substrate

Enzyme	Saturating Substrate	Varied Substrate	k_{cat} $(min)^{-1}$	K_m (mM)	k_{cat} / K_m $(mM^{-1}\ min^{-1})$	$(k_{cat} / K_m)_{Loki}$ / $(k_{cat} / K_m)_{TDP-16}$
TDP-16	UDP	2-chloro-4-nitrophenyl-β-D-glucopyranoside (**1**)	20.2 ± 0.4	0.40 ± 0.04	50.5 ± 5.1	---
		2-chloro-4-nitrophenyl-3-deoxy-β-D-glucopyranoside (**2**)	---	---	0.03	---
		2-chloro-4-nitrophenyl-4-deoxy-β-D-glucopyranoside (**3**)	---	---	0.03	---
		2-chloro-4-nitrophenyl-2,6-dideoxy-β-D-glucopyranoside (**4**)	---	---	11.4	---
		2-chloro-4-nitrophenyl-4,6-dideoxy-β-D-glucopyranoside (**5**)	---	---	0.20	---
		2-chloro-4-nitrophenyl-β-D-allopyranoside (**6**)	---	---	0.01	---
Loki	UDP	2-chloro-4-nitrophenyl-β-D-glucopyranoside (**1**)	49.8 ± 1.5	0.13 ± 0.02	383.1 ± 60.1	**7.6**
		2-chloro-4-nitrophenyl-3-deoxy-β-D-glucopyranoside (**2**)	---	---	0.37	**12.3**
		2-chloro-4-nitrophenyl-4-deoxy-β-D-glucopyranoside (**3**)	---	---	0.47	**15.7**
		2-chloro-4-nitrophenyl-2,6-dideoxy-β-D-glucopyranoside (**4**)	---	---	18.10	**1.6**
		2-chloro-4-nitrophenyl-4,6-dideoxy-β-D-glucopyranoside (**5**)	---	---	2.81	**14.1**
		2-chloro-4-nitrophenyl-β-D-allopyranoside (**6**)	---	---	0.15	**15.0**

Table 4.3. Determined kinetic parameters for the OleD variants TDP-16 and Loki. Parameters were determined with various 2-chloro-4-nitrophenyl glycosides (**1-6**) as donors and UDP as acceptor.

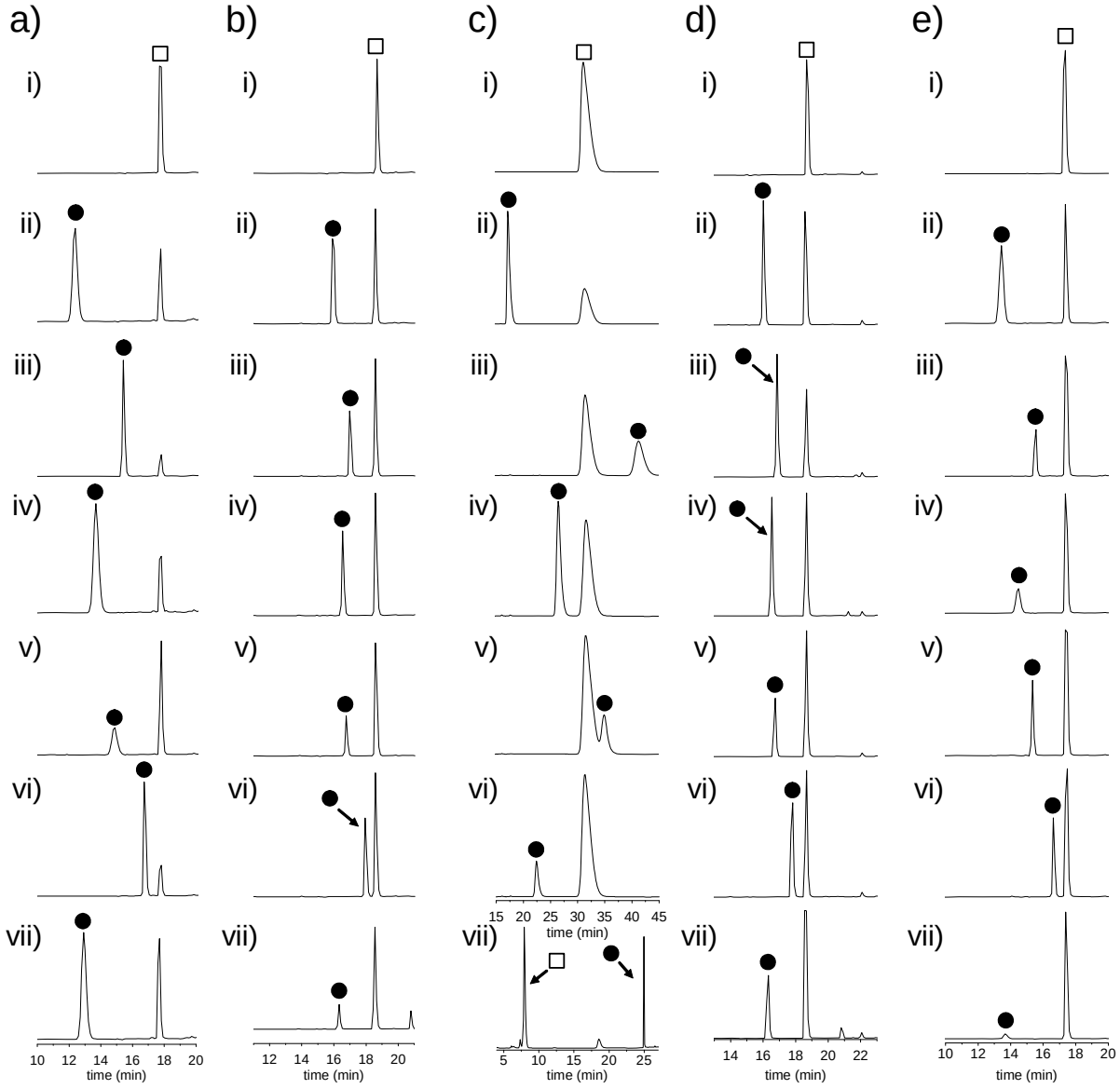


Figure 4.15. Representative HPLC chromatograms of combinatorial NDP-sugar synthesis with the OleD variant Loki as catalyst. Graphs represent **a) i)** UDP control and formation of **ii)** UDP-glucose (**10**), **iii)** UDP-3-deoxyglucose (**15**), **iv)** UDP-4-deoxyglucose (**16**), **v)** UDP-2,6-dideoxyglucose (**17**), **vi)** UDP-4,6-dideoxyglucose (**18**), **vii)** UDP-allose (**19**); **b) i)** TDP control and formation of **ii)** TDP-glucose (**11**), **iii)** TDP-3-deoxyglucose (**20**), **iv)** TDP-4-deoxyglucose (**21**), **v)** TDP-2,6-dideoxyglucose (**22**), **vi)** TDP-4,6-dideoxyglucose (**23**), **vii)** TDP-allose (**24**); **c) i)** CDP control and formation of **ii)** CDP-glucose (**12**), **iii)** CDP-3-deoxyglucose (**25**), **iv)** CDP-4-deoxyglucose (**26**), **v)** CDP-2,6-dideoxyglucose (**27**), **vi)** CDP-allose (**29**) **vii)** CDP-4,6-dideoxyglucose (**28**); **d) i)** ADP control and formation of **ii)** ADP-glucose (**13**), **iii)** ADP-3-deoxyglucose (**30**), **iv)** ADP-4-deoxyglucose (**31**), **v)** ADP-2,6-dideoxyglucose (**32**), **vi)** ADP-4,6-dideoxyglucose (**33**), **vii)** ADP-allose (**34**); **e) i)** GDP control and formation of **ii)** GDP-glucose (**14**), **iii)** GDP-3-deoxyglucose (**35**), **iv)** GDP-4-deoxyglucose (**36**), **v)** GDP-2,6-dideoxyglucose (**37**), **vi)** GDP-4,6-dideoxyglucose (**38**), **vii)** GDP-allose (**39**). NDPs and NDP-sugars are represented by □ and ●, respectively. Vertical axis represents absorbance at 254 nm.

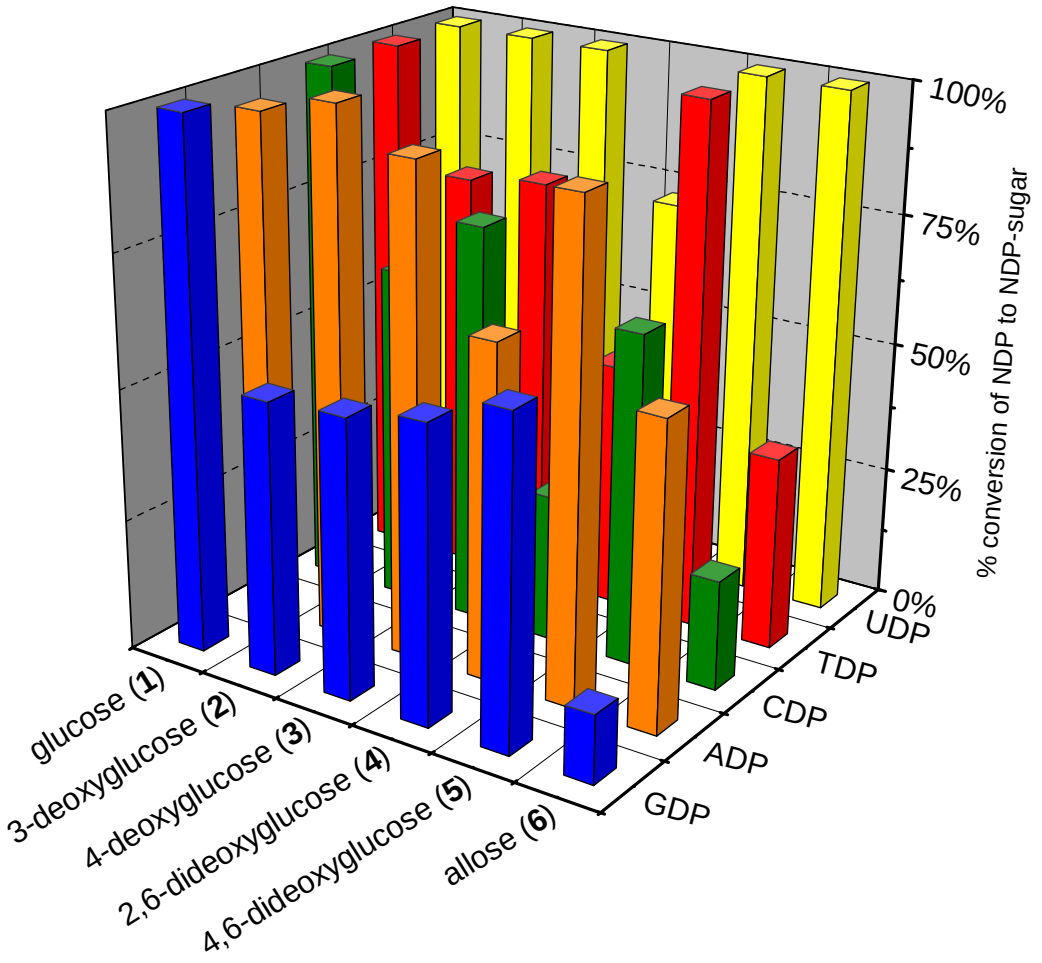


Figure 4.16. Percent conversions of NDP to NDP-sugar with the OleD variant Loki. Reactions contained 0.5 mM of 2-chloro-4-nitrophenyl glycoside donor (**1**-**6**), 1.0 mM NDP, and 10.8 μM (50 μg) of purified Loki variant in Tris-HCl buffer (50 mM, pH 8.0) and a final volume of 100 μL. Absorbance at 410 nm was followed and reactions were halted when they reached completion, increase in absorbance halted, or total reaction time reached 24 hours (see section **4.4**). Percent conversions were determined by HPLC. The standard deviation for production of UDP-α-D-glucose (**10**) across multiple runs (n=3) was <10%.

Entry	Compound	Reaction Length (hours)	Retention Time (min)	Observed Conversion by HPLC	Elemental Composition $[M-2H]^{2-}$	HRMS-ESI (*m/z*)	
						Calculated Mass $[M-2H]^{2-}$	Observed Mass $[M-2H]^{2-}$
10	UDP-α-D-glucose†	0.25	12.4	100%	$C_{15}H_{22}N_2O_{17}P_2^{2-}$	282.02023	282.0200
11	TDP-α-D-glucose†	0.25	16.0	100%	$C_{16}H_{24}N_2O_{16}P_2^{2-}$	281.03060	281.0303
12	CDP-α-D-glucose	1.75	17.3	100%	$C_{15}H_{23}N_3O_{16}P_2^{2-}$	281.52823	281.5285
13	ADP-α-D-glucose	0.25	16.0	96%	$C_{16}H_{23}N_5O_{15}P_2^{2-}$	293.53385	293.5343
14	GDP-α-D-glucose	0.25	13.5	100%	$C_{16}H_{23}N_5O_{16}P_2^{2-}$	301.53130	301.5317
15	UDP-3-deoxy-α-D-glucose†	10	15.3	100%	$C_{15}H_{22}N_5O_{16}P_2^{2-}$	274.02278	274.0236
16	UDP-4-deoxy-α-D-glucose†	8	13.7	100%	$C_{15}H_{22}N_5O_{16}P_2^{2-}$	274.02278	274.0229
17	UDP-2,6-dideoxy-α-D-glucose	1.25	14.8	74%	$C_{16}H_{22}N_2O_{15}P_2^{2-}$	266.02532	266.0255
18	UDP-4,6-dideoxy-α-D-glucose	2.25	16.7	100%	$C_{15}H_{22}N_2O_{15}P_2^{2-}$	266.02532	266.0252
19	UDP-α-D-allose†	24	12.9	100%	$C_{15}H_{22}N_2O_{17}P_2^{2-}$	282.02023	282.0208
20	TDP-3-deoxy-α-D-glucose	24	17.0	77%	$C_{16}H_{24}N_2O_{15}P_2^{2-}$	273.03314	273.0326
21	TDP-4-deoxy-α-D-glucose	24	16.5	79%	$C_{16}H_{24}N_2O_{15}P_2^{2-}$	273.03314	273.0337
22	TDP-2,6-dideoxy-α-D-glucose	1.25	16.8	47%	$C_{16}H_{24}N_2O_{14}P_2^{2-}$	265.03569	265.0362
23	TDP-4,6-dideoxy-α-D-glucose	6	18.0	100%	$C_{16}H_{24}N_2O_{14}P_2^{2-}$	265.03569	265.0362
24	TDP-α-D-allose	24	16.3	36%	$C_{16}H_{24}N_2O_{16}P_2^{2-}$	281.03060	281.0313
25	CDP-3-deoxy-α-D-glucose	24	41.3	64%	$C_{15}H_{23}N_3O_{15}P_2^{2-}$	273.53077	273.5316
26	CDP-4-deoxy-α-D-glucose	24	26.6	76%	$C_{15}H_{23}N_3O_{15}P_2^{2-}$	273.53077	273.5312
27	CDP-2,6-dideoxy-α-D-glucose	4	35.0	28%	$C_{15}H_{23}N_3O_{14}P_2^{2-}$	265.53331	265.5326
28	CDP-4,6-dideoxy-α-D-glucose	24	25.0	63%	$C_{15}H_{23}N_3O_{14}P_2^{2-}$	265.53331	265.5341
29	CDP-α-D-allose	24	22.4	21%	$C_{15}H_{23}N_3O_{16}P_2^{2-}$	281.52823	281.5282
30	ADP-3-deoxy-α-D-glucose	21	16.9	100%	$C_{16}H_{23}N_5O_{14}P_2^{2-}$	285.53639	285.5360
31	ADP-4-deoxy-α-D-glucose	10	16.5	93%	$C_{16}H_{23}N_5O_{14}P_2^{2-}$	285.53639	285.5370
32	ADP-2,6-dideoxy-α-D-glucose	1.25	16.7	64%	$C_{16}H_{23}N_5O_{13}P_2^{2-}$	277.53893	277.5393
33	ADP-4,6-dideoxy-α-D-glucose	2.25	17.8	93%	$C_{16}H_{23}N_5O_{13}P_2^{2-}$	277.53893	277.5395
34	ADP-α-D-allose	24	16.3	58%	$C_{16}H_{23}N_5O_{15}P_2^{2-}$	293.53385	293.5345
35	GDP-3-deoxy-α-D-glucose	24	15.5	53%	$C_{16}H_{23}N_5O_{15}P_2^{2-}$	293.53384	293.5343
36	GDP-4-deoxy-α-D-glucose	24	14.5	53%	$C_{16}H_{23}N_5O_{15}P_2^{2-}$	293.53384	293.5347
37	GDP-2,6-dideoxy-α-D-glucose	1.25	15.4	56%	$C_{16}H_{23}N_5O_{14}P_2^{2-}$	285.53639	285.5369
38	GDP-4,6-dideoxy-α-D-glucose	24	16.7	62%	$C_{16}H_{23}N_5O_{14}P_2^{2-}$	285.53639	285.5370
39	GDP-α-D-allose	24	13.7	13%	$C_{16}H_{23}N_5O_{16}P_2^{2-}$	301.53130	301.5319

Table 4.4. Characterication of NDP-sugars generated by the Loki variant. Retention times for chemical entries refer to those observed in **Figure 4.15**. Retention times for UDP, TDP, ADP, GDP, and CDP were 17.7, 18.7, 18.7, 17.3, and 31.2 (or 7.9 min in the case of entry **28**), for their respective methods. † Product formation previously confirmed[(C8)].

or cofactor regeneration systems with a 70% isolated yield for the final enzymatic step[25]. In comparison, the current study reports the synthesis of **22** from peracetylated 2-deoxy-D-glucose in 6 steps (5 chemical, 1 enzymatic) with a 47% observed yield from **4** and TDP for the final GT-catalyzed reverse reaction. Importantly, this method can be easily scaled given the reliance upon only a single enzymatic transformation and the final sugar nucleotide formation reaction can be accomplished in a coupled *in situ* format (with a colorimetric readout) for added convenience (see Chapter 3). Furthermore, this chemoenzymatic approach can be easily modified to enable strategically-placed isotopic labels for biosynthetic mechanistic studies and is clearly advantageous over existing enzymatic strategies which are limited to commercially available sugar phosphates.

Taken together with previous results[8], NDP-sugars containing 2-, 3-, 4-, or 6-deoxy, and 2,6- or 4,6-dideoxy motifs are now readily accessible. The impact of being able to access such a wide variety of NDP-deoxy- and, especially, NDP-dideoxysugars with GT-catalyzed reverse reactions should not be underestimated, as these motifs appear consistently in an enormous variety of sugars throughout natural products[2, 7] and are considered essential for drug discovery efforts[6]. Indeed, TDP-2,6-dideoxysugars and their derivatives account for the majority of TDP-sugars used in natural-product biosynthetic pathways[5]. Specific examples include 2,6-dideoxy-D-glucose (*i.e.*, D-olivose), which is utilized in the biosynthesis of natural products that include landomycin A[26], various urdamycin analogs[27], mithramycin[28], various chromomycin analogs[29], chlorothricin[30], avilamycin[31], and concanamycin[32] (**Figure 4.17**). In terms of the 4,6-dideoxy motif, perhaps the most widely recognized sugars are D-chalcose and D-desosamine, which are appended to a wide variety of macrolides. D-chalcose is present in lankamycin[33], chalcomycin[34], and dihydrochalcomycin[35], while D-desosamine can be found appended to

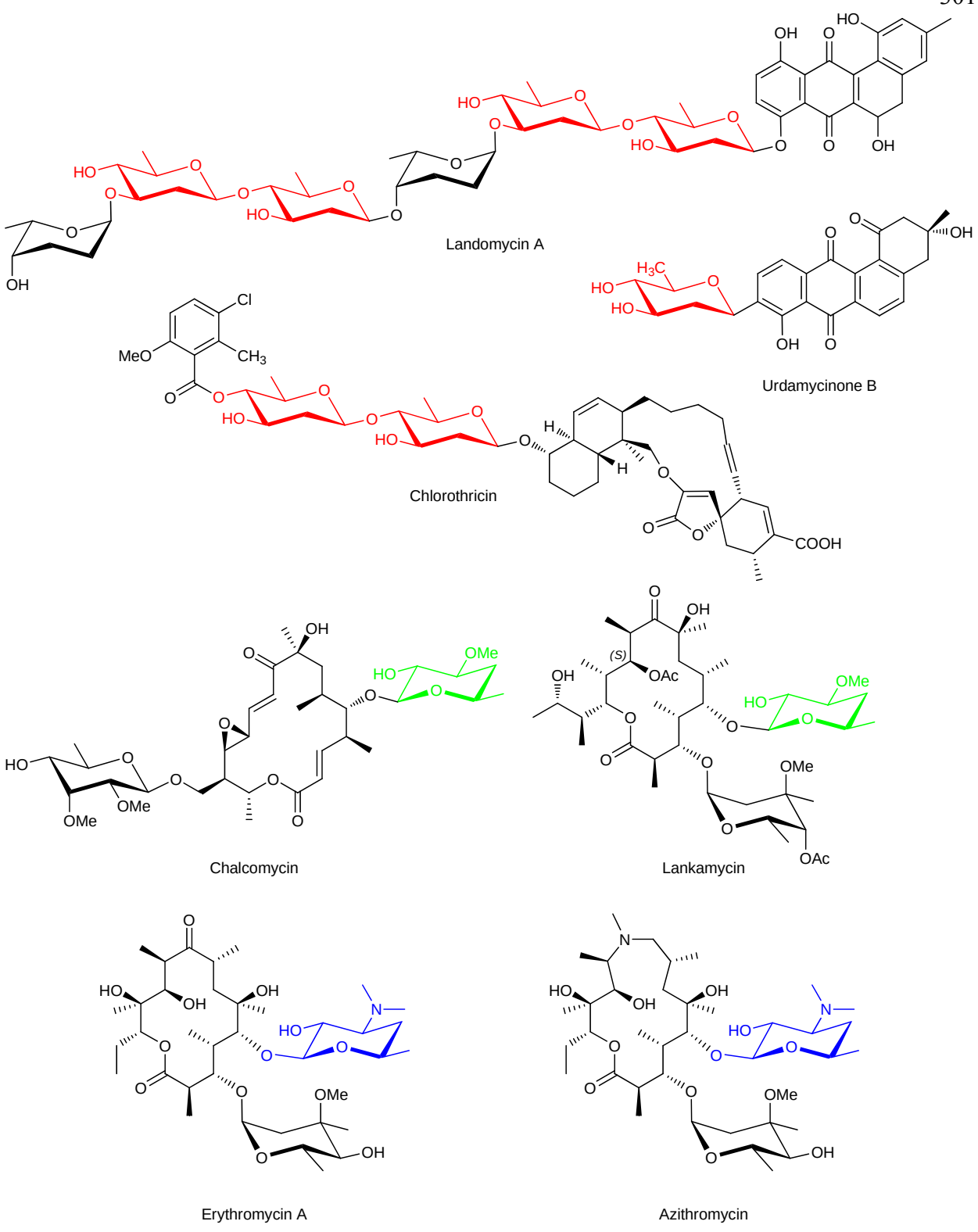


Figure 4.17. Natural products and semi-synthetic derivatives containing dideoxysugars. 2,6-dideoxy-D-glucose (*i.e.*, D-olivose), D-chalcose, and D-desosamine moieties are highlight in red, green, and blue, respectively.

clinically relevant natural product macrolides and semi-synthetic derivatives that include erythromycin[36], megalomicin[37], azithromycin, clarithromycin, and telithromycin[38] (**Figure 4.17**). Collectively, these examples clearly highlight the importance of access to NDP-deoxy-and NDP-dideoxysugars in the context of drug discovery and the elucidation of natural product biosynthetic pathways.

4.3. Conclusions

This study provides the first example of protein engineering in the context of a GT-catalyzed reverse reaction and clearly demonstrates the power of the colorimetric high throughput screen provided by 2-chloro-4-nitrophenyl glycosides to rapidly identify improved catalysts for any desired NDP-sugar pair. The identified OleD Loki variant is capable of utilizing 5 NDPs and 6 sugars for combinatorial catalysis of 30 distinct NDP-sugars (**10**-**39**). Given the even broader characterized sugar promiscuity of similar OleD variants in both forward and reverse GT reactions[8-10], the results here likely offer only a glimpse of the true sugar promiscuity landscape for the Loki variant. Additionally, given the ability of the Loki variant to recognize both pyridine- and purine-based acceptors, generation of NDP-sugars with modified nucleotide moieties possessing unique chemical and/or analytical properties should also not be overlooked[3]. Combined with previous studies[8], >40 distinct NDP-sugars have now been generated utilizing 2-chloro-4-nitrophenyl glycosides as donors in GT-catalyzed reverse reactions. This expansion of substrates also greatly increases the applicability of single and dual GT coupled reactions for biological investigations and drug discovery efforts[4, 8]. Taken together, these factors establish GT-catalyzed reverse reactions, in concert with 2-chloro-4-nitrophenyl

glycosides, as a highly advantageous chemoenzymatic approach for the generation of NDP-sugars.

While many of the unique core sugar modifications associated with natural products were addressed in this study (specifically, deoxy- and dideoxy- motifs), others, such as amino-, *N*-methyl, *O*-methyl, and *C*-methyl groups, as well as chemoselective substitutions (*e.g.*, azides, thiols, etc.), remain to be investigated. Additionally, catalysts for D-mannose (the C-2 epimer of D-glucose) or the L-sugars screened in this study remain elusive. However, with the characterization of GTs capable of recognizing both D- and L-sugars[4], increasing numbers of GTs characterized with small, phenolic substrates[4, 8, 39-41], generation of GT chimeras with modified NDP-sugar substrate recognition[42, 43], and continuing reports of crystal structures[14, 16, 17, 21, 44-47], it is likely only a matter of time until GT variants are identified to address these current deficiencies. Nevertheless, identification of the Loki variant provides an important first step towards expanding applications of GT-catalyzed reverse reactions for facile access to both biologically significant NDP-sugars and those with a wide variety of chemical modifications.

4.4. Materials and Methods

4.4.1. General materials and methods.

4.4.1.1 General materials. Unless otherwise stated, all general chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) or New England Biolabs (Ipswich, MA, USA). The compounds 2-chloro-4-nitrophenyl-β-D-glucopyranoside (**1**), 2-chloro-4-nitrophenyl-β-D-3-deoxy-glucopyranoside (**2**), 2-chloro-4-nitrophenyl-β-D-4-deoxy-glucopyranoside (**3**), 2-chloro-4-nitrophenyl-β-D-allopyranoside (**6**), 2-chloro-4-nitrophenyl-β-D-mannopyranoside (**7**), and 2-chloro-4-nitrophenyl-α-L-rhamnopyranoside (**9**) were previously synthesized[8]. 2-chloro-4-nitrophenyl-α-L-fucopyranoside (**8**) was purchased from Carbosynth (United Kingdom). All commercially obtained compounds were utilized as received.

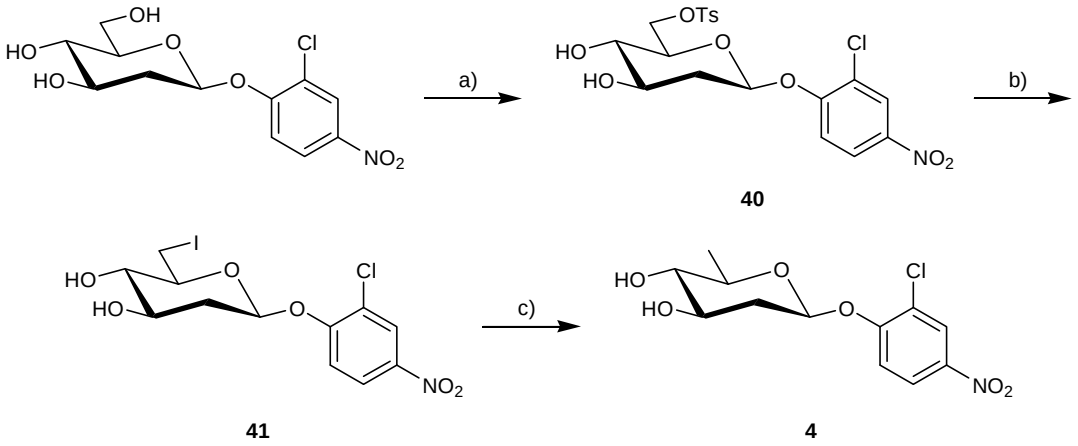
4.4.1.2 General methods. The sequence of wild-type OleD can be found within the Entrez Protein database under accession code ABA42119. The examined OleD crystal structure can be found under the Protein Data Bank ID 2IYF[14]. The modeled ligand UDP-2-deoxy-2-fluoroglucose was taken from UGT72B1 (PDB ID: 2VCE)[16]. All crystal structure figures were generated with Pymol Molecular Graphics System software. Primers were obtained from Integrated DNA Technologies, Inc. (Coralville, IA, USA). All DNA sequencing was conducted by the Univ. of Wisconsin Biotechnology Center with the primers T7 Promoter (5'-TAA TAC GAC TCA CTA TAG GG-3') or T7 Terminator (5'-GCT AGT TAT TGC TCA GCG G-3'). All Luria-Bertani (LB) media and agar were supplemented with 50 μg mL^{-1} kanamycin. Liquid manipulations for high throughput assays were accomplished with a Biomek FX Liquid

Handling Workstation (Beckman Coulter, Brea, CA, USA). Absorbance readings were taken on a FLUOstar Optima plate reader (BMG Labtechnologies, Cary, NC, USA). High resolution mass spectra (HRMS) were acquired on a Bruker MaXis ultra-high resolution quadrupole time of flight mass spectrometer by negative ionization electrospray with a source potential of 2800 V, drying gas at 200 °C flowing at 4 L/min and a nebulizing gas pressure of 0.4 bar. Samples were infused at 3 μL/min and spectra collected for 2 min. Routine TLC analyses were performed on aluminum TLC plates coated with 0.2 mm silica gel (from Sigma-Aldrich, St. Louis, MO, USA) and monitored at 254 nm. Flash column chromatography was achieved on 40 – 63 μm, 60 Å silica gel (Silicycle, Quebec, Canada). Regardless of method, HPLC peak areas were integrated with Star Chromatography Workstation Software (Varian, Palo Alto, CA, USA) and the percent conversion calculated as a percent of the total peak area.

NMR spectra were obtained using a ^{UNITY}INOVA 400 MHz instrument (from Varian, Palo Alto, CA, USA) in conjunction with a QN Switchable BB probe (from Varian) or ^{UNITY}INOVA 500 MHz instrument in conjunction with a qn6121 probe (from Nalorac, Martinez, CA, USA). ^{1}H and ^{13}C chemical shifts were referenced to internal solvent resonances. ^{31}P chemical shifts were not referenced. Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), qn (quintet), m (multiplet) and br (broad). Italicized elements or groups are those that are responsible for the shifts. Chemical shifts are reported in parts per million (ppm) and coupling constants (*J*) are given in Hz. NMR assignments were performed with the aid of gCOSY and gHSQC experiments.

4.4.2. Chemical Syntheses

4.4.2.1. 2-chloro-4-nitrophenyl-β-D-2,6-dideoxy-glucopyranoside (4). This compound was synthesized and characterized by Dr. Maoquan Zhou.



a) *p*-TsCl, anhydrous pyridine, CH_2Cl_2; **b)** NaI, 2-butanone; **c)** Bu_3SnH, AIBN, 1,2-dimethoxyethane, 80 °C

2-chloro-4-nitrophenyl-2-deoxy-6-*O*-tosyl-β-D-glucopyranoside (40). The previously described 2-chloro-4-nitrophenyl-2-deoxy-β-D-glucopyranoside[8] (735 mg, 2.3 mmol) was dissolved in anhydrous pyridine (10 mL) and anhydrous CH_2Cl_2 (6 mL). The mixture was cooled to 0 °C and *p*-toluenesulfonyl chloride (573mg, 3 mmol) was added dropwise over period of 10 minutes with vigorous stirring. The reaction was allowed to warm up to room temperature and stirred overnight. The solvent was evaporated under reduced pressure and remaining residue dissolved in CH_2Cl_2 (70 mL) and washed with saturated $NaHCO_3$ (2 x 20 mL) and dH_2O (20 mL). The recovered organics were dried over $NaSO_4$ and concentrated with reduced pressure. Purification by flash chromatography on silica gel (6% MeOH in CH_2Cl_2) afforded **40** (675 mg, 1.43 mmol, 62%) as a white powder. TLC R_f = 0.43 (10% MeOH in CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (d, J = 2.8 Hz, 1 H), 7.97 (dd, = 9.2, 2.4 Hz, 1 H), 7.74 (m, 1 H), 7.72 (m, 1

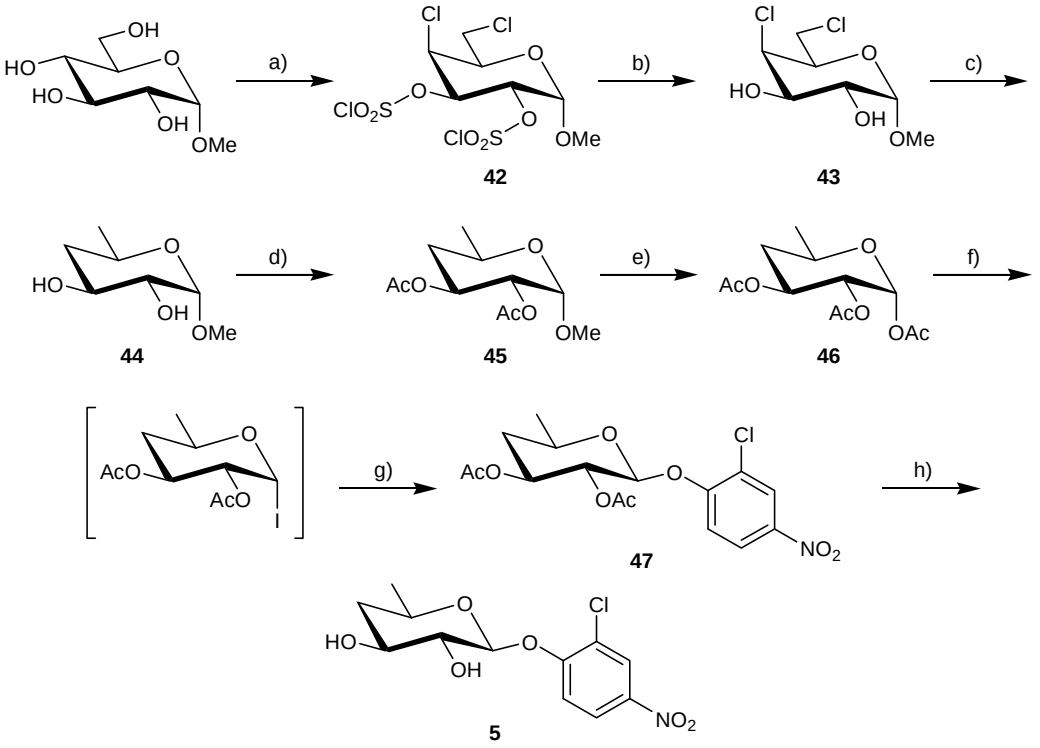
H), 7.25 (m, 1 H), 7.23 (m, 1 H), 7.10 (d, *J* = 9.2 Hz, 1 H), 5.29 (d, *J* = 3.2 Hz, 1 H), 4.43 (dd, *J* = 10.8, 5.2 Hz, 1 H), 4.37 (dd, *J* = 11.2, 2.4 Hz, 1 H), 3.86 (ddd, *J* = 11.6, 8.8, 5.2 Hz, 1 H), 3.69 (ddd, *J* = 9.6, 4.8, 2.0 Hz, 1 H), 3.57 (dd, *J* = 9.2, 9.2 Hz, 1 H), 3.30 (s, 2 H), 2.53 (ddd, *J* = 12.8, 5.2, 2.4 Hz, 1 H), 2.38 (s, 3H), 2.03 (ddd, *J* = 11.6, 11.2, 9.6 Hz, 1 H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 157.5, 145.7, 142.6, 132.5, 130.1 (2), 128.1 (2), 126.2, 124.3, 123.8, 115.9, 97.8, 74.3, 71.2, 71.0, 68.8, 38.0, 21.8. HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{19}H_{20}ClNNaO_6S$, 496.0440, found 496.0433.

2-chloro-4-nitrophenyl-2-deoxy-6-iodo-β-D-glucopyranoside (41). A solution of **40** (474.5 mg, 1.0 mmol) in 2-butanone (10 mL) was refluxed with sodium iodide (450 mg, 3.0 mmol) for 4 hours. The reaction mixture was concentrated with reduced pressure and the residue purified by flash chromatography (7% MeOH in CH_2Cl_2) to afford **41** (250 mg, 0.58 mmol, 58%) as a white powder. TLC R_f = 0.40 (10% MeOH in CH_2Cl_2); ^{1}H NMR (500 MHz, 1:4 $CD_3OD/CDCl_3$) δ 8.71 (d, *J* = 2.5 Hz, 1 H), 8.59 (dd, = 9.0, 2.5 Hz, 1 H), 7.96 (d, *J* = 9.0 Hz, 1 H), 5.76 (dd, *J* = 9.5, 2.0 Hz, 1 H), 4.17 (dd, *J* = 11.0, 2.0 Hz, 1 H), 4.13 (ddd, *J* = 8.5, 7.0, 4.0 Hz, 1 H), 3.88 (ddd, *J* = 9.0, 9.0, 2.0 Hz, 1 H), 3.81 (s, 1H), 3.77 (m, 1H), 3.67 (dd, *J* = 10.5, 9.0 Hz, 1 H), 3.63 (dd, *J* = 9.0, 9.0 Hz, 1 H), 2.91 (ddd, *J* = 12.5, 5.0, 2.0 Hz, 1 H), 2.42 (ddd, *J* = 12.5, 10.0, 10.0 Hz, 1 H); ^{13}C NMR (125 MHz, 1:4 $CD_3OD/CDCl_3$) δ 157.3, 142.1, 125.5, 123.6, 123.5, 116.3, 97.5, 76.9, 75.1, 70.2, 38.0, 4.9. HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{13}ClINO_6Na$, 451.9368, found 451.9359.

2-chloro-4-nitrophenyl-2,6-dideoxy-β-D-glucopyranoside (4). A solution of **41** (51 mg, 0.12 mmol) in 1,2-dimethoxyethane (1 mL) was refluxed with tri-*n*-butyltin hydride (55 mg, 0.19

mmol) and catalytic 2,2'-azobis(2-methylpropinonitrile) (2 mg, 0.012 mmol) at 80 °C for 5 hours. The reaction mixture was concentrated with reduced pressure and the residue was purified by flash chromatography on silica gel (10% MeOH in CH_2Cl_2) to afford **4** (11 mg, 0.036 mmol, 30%) as a white powder. TLC R_f = 0.35 (10% MeOH in CH_2Cl_2); ^{1}H NMR (500 MHz, CD_3OD) δ 8.31 (d, $J_{H\text{-}3',H\text{-}5'}$ = 3.0 Hz, 1 H, H-3'), 8.20 (dd, $J_{H\text{-}5',H\text{-}6'}$ = 9.0 Hz, $J_{H\text{-}5',H\text{-}3'}$ = 2.5 Hz, 1 H, H-5'), 7.38 (d, $J_{H\text{-}6',H\text{-}5'}$ = 8.5 Hz, 1 H, H-6'), 5.46 (dd, $J_{H\text{-}1,\,H\text{-}2_{ax}}$ = 8.5 Hz, $J_{H\text{-}1,\,H\text{-}2_{eq}}$ = 2 Hz, 1 H, H-1), 3.68 (ddd, $J_{H\text{-}3,H\text{-}2_{ax}}$ = 12.0 Hz, $J_{H\text{-}3,H\text{-}4}$ = 9.0 Hz, $J_{H\text{-}1,H\text{-}2_{eq}}$ = 5.5 Hz, 1 H, H-3), 3.51 (dd, $J_{H\text{-}5,H\text{-}4}$ = 9.0 Hz, $J_{H\text{-}5,H\text{-}6}$ = 6.5 Hz, 1 H, H-5), 3.46 (dd, $J_{H\text{-}4,H\text{-}3}$ = 9.0 Hz, $J_{H\text{-}4,H\text{-}5}$ = 9.0 Hz, 1 H, H-4), 2.41 (ddd, $J_{H\text{-}2_{eq},H\text{-}2_{ax}}$ = 12.5 Hz, $J_{H\text{-}2_{eq},H\text{-}1}$ = 5.0 Hz, $J_{H\text{-}2_{eq},H\text{-}3}$ = 2.0 Hz, 1 H, H-2_{eq}), 1.87 (ddd, $J_{H\text{-}2_{ax},H\text{-}2_{eq}}$ = 12.0 Hz, $J_{H\text{-}2_{ax},H\text{-}1}$ = 10.0 Hz, $J_{H\text{-}2_{ax},H\text{-}3}$ = 10.0 Hz, 1 H, H-2_{ax}), 1.34 (d, $J_{H\text{-}6,H\text{-}5}$ = 6.5 Hz, 3 H, H-6); ^{13}C NMR (125 MHz, CD_3OD) δ 159159.0, 143.5, 126.7, 124.9, 124.6, 116.8, 98.4, 78.1, 74.0, 71.8, 39.9, 18.2. HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNO_6Na$, 326.0402, found 326.0398.

4.4.2.2. 2-chloro-4-nitrophenyl-β-D-4,6-dideoxy-glucopyranoside (5). This compound was synthesized and characterized in a collaboration with Dr. Maoquan Zhou.



a) excess SO_2Cl_2, anhydrous pyridine, $CHCl_3$; **b)** NaI, MeOH; **c)** Bu_3SnH, AIBN, reflux at 105 °C in toluene; **d)** excess Ac_2O, pyridine; **e)** concentrated H_2SO_4, Ac_2O, glacial acetic acid; **f)** TMSI, DCM, 0 °C; **g)** potassium 2-chloro-4-nitrophenol salt, 18-crown-6, THF, 0 °C; **h)** 4 Å molecular sieves, MeOH

1-methyl-2,3-disulfochloro-4,6-dideoxy-4,6-dichloro-α-D-galactopyranoside (42). From previously reported protocols[48, 49], 1-methyl-α-D-glucopyranoside (21.0 g, 108.1 mmol) was dissolved in 90 mL of anhydrous pyridine and 600 mL of anhydrous $CHCl_3$. The mixture was cooled on a bath of dry ice and acetone (-70 °C) and an excess of sulfuryl chloride (60 mL) was

added dropwise over period of 30 minutes with vigorous stirring. The reaction was allowed to proceed for 2 hours at -70 °C and then warmed to room temperature over several hours. TLC analysis (25% MeOH in DCM) revealed complete consumption of starting material (R_f = 0.45). The reaction was filtered to remove precipitants formed during the course of the reaction and then washed with 10% H_2SO_4 in dH_2O (600 mL) and saturated sodium bicarbonate (2 x 500 mL). The organic solution was dried over $NaSO_4$ and concentrated with reduced pressure. The resulting crude orange syrup of **42** (46.0 g, 107 mmol, >99% yield) was carried forward without additional characterization.

1-methyl-4,6-dideoxy-4,6-dichloro-α-D-galactopyranoside (43). Crude **42** (46.0 g, 107 mmol)[48, 49] was dissolved in MeOH (400 mL). To this reaction mixture, a solution of sodium iodide (24 g dissolved in 60 mL of 50% MeOH in H_2O) was added dropwise over the course of 30 minutes with stirring at RT. The reaction was neutralized at 4 hours with the addition of $NaHCO_3$ until the evolution of gas halted. Following filtration, the resulting precipitant was thoroughly washed with MeOH. The collected filtered organics were then concentrated under reduced pressure. The concentrate was dissolved in 250 mL of dH_2O and then a few crystals of thiosulphate were added to remove trace iodine. The solution was then extracted with DCM (3 x 250 mL). The resulting organics were dried over $MgSO_4$ and concentrated with reduced pressure. Purification with flash chromatography on silica gel (5% MeOH in DCM) afforded **43** (10.2 g, 44.1 mmol, 39% yield) as white crystals. TLC R_f = 0.68 (10% MeOH in DCM); ^{1}H NMR (400 MHz, CD_3OD) δ 4.73 (d, $J_{H\text{-}1,H\text{-}2}$ = 3.9 Hz, 1 H, H-1), 4.45 (dd, $J_{H\text{-}4,H\text{-}5}$ = 1.1, $J_{H\text{-}4,H\text{-}3}$ = 3.6, 1 H, H-4), 4.17-4.13 (m, 1 H, H-5), 3.99 (dd, $J_{H\text{-}3,H\text{-}4}$ = 3.6, $J_{H\text{-}3,H\text{-}2}$ = 9.9, 1 H, H-3), 3.79 (dd, $J_{H\text{-}2,H\text{-}1}$ = 3.9, $J_{H\text{-}2,H\text{-}3}$ = 9.9, 1 H, H-2), 3.72-3.61 (m, 2H, H-6a, H-6b), 3.43 (s, 3 H, OCH_3); ^{13}C NMR (100

MHz, CD_3OD) δ 101.4, 70.9, 69.8, 69.5, 64.8, 55.7, 44.3; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_7H_{12}Cl_2NaO_4$, 253.000485, found 252.99969. Observed spectra are consistent with those reported by van Summeren, *et al.*[50]

1-methyl-4,6-dideoxy-α-D-glucopyranoside (44). According to a previously described protocol[50], **43 (**5.0 g, 21.7 mmol) in the presence of tributyltin hydride (32.94 g, 103 mmol) and 2,2'-azobis(2-methylpropinonitrile) (1.0 g, 6.0 mmol; dissolved in 5 mL of toluene) in toluene (35 mL) was refluxed at 105 °C overnight. The reaction mixture was cooled to RT and concentrated with reduced pressure. The concentrate (yellowish oil) was dissolved with acetonitrile (10 mL) and washed with MeOH (2 x 4 mL). Following, the solution was concentrated with reduced pressure and the resulting white crystals were washed with diethyl ether to afford **44** (2.42 g, 14.9 mmol, 68% yield). TLC R_f = 0.58 (10% MeOH in DCM); ^{1}H NMR (400 MHz, $CDCl_3$) δ 4.7 (d, $J_{H\text{-}1,H\text{-}2}$ = 3.7 Hz, 1 H, H-1), 3.95-3.79 (m, 2 H, H-3, H-5), 3.42-3.37 (m, 1 H, H-2), 1.97 (ddd, J = 2.0, 4.9, 12.9 Hz, 1 H, H-4), 1.35 (q, J = 11.6 Hz, 1 H, 4-H'), 1.21 (d, $J_{H\text{-}6,H\text{-}5}$ = 6.3 Hz, 3 H, H-6); ^{13}C NMR (100 MHz, $CDCl_3$) δ 100.0, 74.3, 68.4, 64.0, 55.1, 40.0, 20.7; HRMS-ESI (*m/z*): calcd for $C_7H_{14}NaO_4$ 185.07843, found 185.07786.

1-methyl-2,3-di-*O*-acetyl-4,6-dideoxy-α-D-glucopyranoside (45). In anhydrous pyridine (30 mL), **44** (2.42 g, 14.9 mmol) was dissolved and cooled to 0 °C on an ice bath. Acetic anhydride (14.1 mL, 15.2 g, 149 mmol) was added to the reaction mixture dropwise at 0 °C over 15 minutes with stirring. After the final addition of acetic anhydride, the reaction was stirred for 5 minutes at 0 °C, warmed to room temperature, and then sonicated in a water bath for 2 hours. Following, reaction mixture was concentrated with reduced pressure and the concentrate purified

with flash chromatography on silica gel (1:2 EtOAc:hexanes) to afford **45** (3.52 g, 14.2 mmol, 96% yield) as a colorless oil. TLC R_f = 0.55 (1:1 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.31-5.22 (m, 1 H, H-3), 4.89-4.82 (m, 2 H, H-1, H-2), 4.02-3.95 (m, 1 H, H-5), 3.38 (s, 3 H, OCH_3), 2.20-2.13 (m, 1 H, H-4), 2.09 (s, 3 H, $COCH_3$), 2.02 (s, 3 H, $COCH_3$), 1.45 (q, J = 11.7 Hz, 1 H, H-4'), 1.21 (d, $J_{H\text{-}6,H\text{-}5}$ = 6.3 Hz, 3 H, 6-H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.5, 170.2, 97.4, 72.1, 68.0, 63.0, 54.9, 38.1, 20.9, 20.8, 20.5; HRMS (*m/z*): $[M+Na]^+$ calcd for $C_{11}H_{18}NaO_6$ 269.099559, found 269.09987.

1,2,3-tri-*O*-acetyl-4,6-dideoxy-α-D-glucopyranoside (46). With conditions adapted from a previously reported protocol[51], the methyl glycoside **45** (3.95 g, 16.0 mmol) was dissolved in glacial acetic acid (45 mL) and acetic anhydride (45 mL). The reaction was cooled to 0 °C on an ice bath with stirring and concentrated H_2SO_4 (5 mL) was added dropwise over 10 minutes. The reaction mixture was then allowed to warm to RT and proceed for 18 hours. Following, the reaction was cooled to 0 °C on an ice bath and triethylamine (20 mL) was added to the reaction in small aliquots over 30 minutes. The neutralized reaction was concentrated with reduced pressure and the resulting residue dissolved in DCM (200 mL). The resulting solution was washed with saturated $NaHCO_3$ (3 x 200 mL) and the recovered organics concentrated with reduced pressure. Purification by flash chromatography on silica gel (gradient of 1:3 to 1:2 EtOAc:hexanes) afforded **46** (3.57 g, 81% yield) as a colorless oil. TLC R_f = 0.42 (1:2 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.29 (d, $J_{H\text{-}1,H\text{-}2}$ = 3.6 Hz, 1 H, H-1), 5.30-5.23 (m, 1 H, 3-H), 5.01 (dd, $J_{H\text{-}2,H\text{-}1}$ = 3.6 Hz, $J_{H\text{-}2,H\text{-}3}$ = 10.3 Hz, 1 H, H-2), 4.17-4.09 (m, 1 H, H-5), 2.24-2.19 (m, 1 H, H-4), 2.14 (s, 3 H, $COCH_3$), 2.05 (s, 3 H, $COCH_3$), 2.02 (s, 3 H, $COCH_3$), 1.53 (q, J = 12.0 Hz, 1 H, H-4'), 1.22 (d, $J_{H\text{-}6,H\text{-}5}$ = 6.2 Hz, 3 H, 6-H); ^{13}C NMR (100 MHz,

$CDCl_3$) δ 170.6, 170.3, 169.6, 90.6, 70.3, 67.9, 66.4, 37.9, 21.23, 21.17, 20.87, 20.85; HRMS (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{18}NaO_7$ 297.09447, found 297.09301.

2-chloro-4-nitrophenyl-2,3-di-*O*-acetyl-4,6-dideoxy-β-D-glucopyranoside (47). A solution of **46** (738 mg, 2.7 mmol) in CH_2Cl_2 (8 mL) was purged with argon and cooled to 0 °C for 15 min. Trimethylsilyl iodide (440 µl) was added by syringe to the cooled solution. The mixture was stirred for 30 minutes, during which the starting material was consumed based upon TLC analysis. Anhydrous toluene was then added and the reaction mixture was concentrated under reduced pressure (3 x 50 mL). The crude product was dissolved in anhydrous THF (8 mL) and cooled to 0 °C for 15 minutes To a cooled solution of potassium 2-chloro-4-nitrophenol (348 mg, 1.65 mmol), 18-crown-6 (436 mg, 1.65 mmol), and 4 Å Molecule sieves (300 mg) in THF (8 mL) was added the iodo-sugar solution. The mixture was stirred for 1 hour, diluted with EtOAc (100 mL) and filtered over a pad of Celite. The filtrate was washed with saturated $Na_2S_2O_3$ (2 x 100 mL) and brine (2 x 100 mL). The organic phase was dried over $MgSO_4$ and concentrated with reduced pressure. Purification by flash chromatography on silica gel (3:7 EtOAc:hexanes) afforded **47** (378 mg, 1.62 mmol, 60%) as a white powder. TLC R_f = 0.5 (1:1 EtOAc:hexanes); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.27 (d, $J_{H\text{-}3',H\text{-}5'}$ = 2.8 Hz, 1 H, H-3'), 8.12 (dd, $J_{H\text{-}5',H\text{-}6'}$ = 9.2 Hz, $J_{H\text{-}5',H\text{-}3'}$ = 2.8 Hz, 1 H, H-5'), 7.23 (d, $J_{H\text{-}5',H\text{-}6'}$ = 9.2 Hz, 1 H, H-6'), 5.26 (dd, $J_{H\text{-}2,H\text{-}3}$ = 9.6 Hz, $J_{H\text{-}2,H\text{-}1}$ = 7.6 Hz, 1 H, H-2), 5.08 (ddd, $J_{H\text{-}3,H\text{-}4_{ax}}$ = 11.2 Hz, $J_{H\text{-}3,H\text{-}2}$ = 9.6 Hz, $J_{H\text{-}3,H\text{-}4_{eq}}$ = 5.2 Hz, 1 H, H-3), 5.04 (d, $J_{H\text{-}1,H\text{-}2}$ = 7.6 Hz, 1 H, H-1), 3.88 (ddd, $J_{H\text{-}5,H\text{-}4_{ax}}$ = 11.2 Hz, $J_{H\text{-}5,H\text{-}6}$ = 6.4 Hz, $J_{H\text{-}5,H\text{-}4_{eq}}$ = 2.0 Hz, 1 H, H-5), 2.23 (ddd, $J_{H\text{-}4_{eq},H\text{-}4_{ax}}$ = 13.2 Hz, $J_{H\text{-}4_{eq},H\text{-}3}$ = 5.2 Hz, $J_{H\text{-}4_{eq},H\text{-}5}$ = 2.0 Hz, 1 H, H-4_{eq}), 2.07 (s, 3 H, H-CH_3CO), 2.06 (s, 3 H, H-CH_3CO), 1.64 (ddd, $J_{H\text{-}4_{ax},H\text{-}4_{eq}}$ = 13.2 Hz, $J_{H\text{-}4_{ax},H\text{-}3}$ = 11.6 Hz, $J_{H\text{-}4_{ax},H\text{-}5}$ = 11.6 Hz, 1 H, H-4_{ax}), 1.36 (d, $J_{H\text{-}6,H\text{-}5}$ = 6.4 Hz, 3 H, H-6); ^{13}C

NMR (100 MHz, $CDCl_3$) δ 170.5, 169.6, 157.9, 142.9, 126.2, 124.8, 123.8, 116.1, 100.0, 71.6, 70.5, 69.1, 37.5. 21.1, 20.9, 20.8; HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{16}H_{18}ClNNaO_8$, 410.0613, found 410.0608.

2-chloro-4-nitrophenyl-4,6-dideoxy-β-D-glucopyranoside (5). The acetylated glycoside **47** (342 mg, 0.88 mmol) was dissolved in dry MeOH (2 mL) and treated at room temperature with a 0.1 M solution of sodium methoxide (1 mL). The mixture was stirred until no starting compound was detected by TLC (10% MeOH in CH_2Cl_2). Neutralization was performed by adding Amberlite IR-120 (H+ form). The resin was removed via filtration and the filtrate concentrated with reduced pressure. Purification by flash chromatography on silica gel (10% MeOH in CH_2Cl_2) afforded **5** (76 mg, 0.25 mmol, 28%) as a white powder. TLC R_f = 0.43 (10% MeOH in CH_2Cl_2); ^{1}H NMR (500 MHz, 1:3 CD_3OD:$CDCl_3$) δ 8.17 (d, $J_{H\text{-}3',H\text{-}5'}$ = 2.5 Hz, 1 H, H-3'), 8.03 (dd, $J_{H\text{-}5',H\text{-}6'}$ = 9.0Hz, $J_{H\text{-}5',H\text{-}3'}$ = 2.5 Hz, 1 H, H-5'), 7.11 (d, $J_{H\text{-}6',H\text{-}5'}$ = 9.0 Hz, 1 H, H-6'), 4.87 (d, $J_{H\text{-}1,H\text{-}2}$ = 8.5 Hz, 1 H, H-1), 3.72 (m, 3 H, C-2-O*H*, C-3-O*H*, H-5), 3.62 (ddd, $J_{H\text{-}3,H\text{-}2}$ = 9.0 Hz, $J_{H\text{-}3,H\text{-}4_{eq}}$ = 5.0 Hz, $J_{H\text{-}3,H\text{-}4_{ax}}$ = 11.5 Hz, 1 H, H-3), 3.46 (dd, $J_{H\text{-}2,H\text{-}3}$ = 8.0 Hz, $J_{H\text{-}2,H\text{-}1}$ = 8.5 Hz, 1 H, H-2), 1.95 (ddd, $J_{H\text{-}4_{eq},H\text{-}4_{ax}}$ = 13.0 Hz, $J_{H\text{-}4_{eq},H\text{-}3}$ = 5.0 Hz, $J_{H\text{-}4_{eq},H\text{-}5}$ = 2.0 Hz, 1 H, H-4_{eq}), 1.41 (ddd, $J_{H\text{-}4_{ax},H\text{-}4_{eq}}$ = 13.0 Hz, $J_{H\text{-}4_{ax},H\text{-}3}$ = 11.5 Hz, $J_{H\text{-}4_{ax},H\text{-}5}$ = 11.5 Hz, 1 H, H-4_{ax}), 1.20 (d, $J_{H\text{-}6,H\text{-}5}$ = 6.5 Hz, 3 H, H-6); ^{13}C NMR (125 MHz, 1:3 CD_3OD:$CDCl_3$) δ 158.0, 142.1, 125.9, 123.6, 123.6, 115.2, 101.0, 74.8, 70.6, 69.1, 39.6, 20.6. HRMS-ESI (*m/z*): $[M+Na]^+$ calcd for $C_{12}H_{14}ClNNaO_6$, 326.0402, found 326.0397.

4.4.3. Creation of Saturation Libraries and Recombinants. Primers were designed such that the annealing portions of the primer had a T_m of 63-65 °C. Primer phosphorylation reactions

containing 2.5 μL of 100 μM 'forward' primer, 2.5 μL of 100 μM 'reverse' primer, 1 μL of 10X T4 DNA Ligase buffer (New England Biolabs Ipswich, MA, USA), 0.5 μL of T4 polynucleotide kinase (New England Biolabs), and 4 μL of sterile ddH_2O were allowed to proceed for 2 hours at 37 °C and then heat inactivated at 90 °C for 10 minutes.

PCR reactions containing 5 μL of phosphorylated primer mixture, 1 μL (17 ng total) methylated template plasmid DNA (pET28a/*oleD*[P67T/S132F/A242L/Q268V] – the template for OleD variant TDP-16)[9], 20 μL Phusion Flash PCR Master Mix (New England Biolabs), and 14 μL of ddH_2O were prepared. Reactions for recombination contained the appropriate combination of methylated template plasmid DNA and primers to generate the desired combination of mutations. See **Appendix 3** for combinations of template plasmid DNA and primers utilized to generate all variants described herein. PCR reactions were subjected to 1 cycle of 98 °C for 10 seconds, 30 cycles of 98 °C for 1 second, 63 °C for 5 seconds, 72 °C for 1 minute 45 seconds, and 1 cycle of 72 °C for 1 minute. Following confirmation of product formation by separation on agarose gel, the methylated template plasmid DNA was degraded with 20 U of DpnI (New England Biolabs). The digestion was allowed to proceed at 37 °C for 3 hours and then separated on a 0.8% low melting point agarose gel. The desired linear DNA product was excised and purified with the PureLink gel extraction kit according to the manufacturer's protocol (Invitrogen, Grand Island, NY, USA).

Following, a ligation reaction containing 8 μL of gel extraction product, 1 μL of T4 DNA ligase (New England Biolabs), and 1 μL of the supplied 10X ligase buffer was incubated overnight at 25 °C. 4 μL of ligation reaction was transformed into *E. coli* NovaBlue chemically competent cells (EMD Biosciences, Darmstadt, Germany) according to the manufacturer's protocol, plated on LB agar, and incubated at 37 °C. The resulting colonies were pooled and

utilized to inoculate a single 5 mL culture of LB medium and incubated overnight at 37 °C and 250 RPM. Plasmid DNA was isolated from each culture with the QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions.

The isolated plasmid DNA was transformed into chemically competent *E. coli* BL21(DE3)pLysS (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions, plated on LB agar, and incubated at 37 °C. Isolated single colonies were used to inoculate 5 mL cultures of LB medium and incubated overnight at 37 °C and 250 RPM. A glycerol stock (35% glycerol (v/v)) of each culture was prepared and frozen at -80 °C and plasmid DNA was isolated from the remainder of the culture as described above. Isolated plasmid DNA was submitted for sequencing to confirm the presence of the desired mutations. See **Appendix 3** for sequencing results of individual variants.

Glycerol stocks of the *E. coli* BL21(DE3)pLysS strains containing the desired mutated plasmids (prepared above) were then utilized to inoculate 1 mL cultures of LB medium in 96 deep-well plates (VWR International, Radnor, PA, USA) and incubated overnight at 37 °C and 300 RPM. Following, library master glycerol stocks (35% glycerol (v/v)) were prepared in 96-well microtiter plates and frozen at -80 °C. See **Appendix 3** for layouts of the master plates.

4.4.4. Library Expression Test. Strains containing empty pET28a vector, pET28a/*oleD*[P67T/S132F/A242L/Q268V][9], or plasmid variants possessing mutations which result in an amino acid alteration to either alanine or glycine at each of the targeted saturation positions within the parental gene sequence were selected for expression. Briefly, single colonies were used to inoculate 2 mL LB starter cultures and incubated overnight at 37 °C and 250 RPM. 0.5 mL of saturated starter culture was transferred to 50 mL cultures of LB medium grown at

37 °C until the OD_{600} reached ~0.7. IPTG (0.4 mM final concentration) was added and cultures were incubated at 28ºC for approximately 18 hours at 250 RPM.

Cell pellets were collected by centrifugation (6,000 *g* at 4 °C for 20 min), resuspended in 2.5 mL of chilled lysis buffer (20 mM phosphate buffer, pH 7.4, 0.5 M NaCl, 10 mM imidazole), and lysed by sonication (5 pulses of 1 minute each) in an ice bath. Cell debris was removed by centrifugation (10,000 *g* at 4 °C for 20 min). Following, the supernatant was applied to 150 µL of nickel nitrilotriacetic acid resin (QIAgen Valencia, CA, USA) pre-equilibrated with wash buffer (20 mM phosphate buffer, pH 7.4, 0.3 M NaCl, 10 mM imidazole). Protein was allowed to bind for 30 min at 4 °C and the resin washed with 3.5 mL wash buffer. Finally, the enzyme was eluted with 250 µL of chilled wash buffer containing an additional 240 mM imidazole for 10 min at 4 °C. Protein eluted from the nickel nitrilotriacetic acid resin was not subjected to desalting or concentration. Following quantification of each sample by Bradford assay, equal amounts of protein (~3 µg) were evaluated by SDS-PAGE to assess the expression and solubility of the selected variants.

4.4.5. Primary Screening and Analysis

4.4.5.1. Primary High Throughput Screen. Glycerol master plates (see section **4.4.3**) were thawed at 4 °C and then 10 µL was transferred to 96-deep-well microtiter plates containing 1 mL of LB in each well. The glycerol master plates were refrozen at -80 °C while the culture plates were sealed with Breathe-Easy Sealing Membrane (Research Products International Corp., Mount Prospect, IL, USA) and incubated at 37 °C and 350 RPM overnight. Following, 50 µL of each culture was transferred to a fresh deep-well plate containing 950 µL of LB medium.

The freshly inoculated plates were incubated for 3 hours at 37 °C and 350 RPM. Expression of the *N*-terminal His_6-tagged OleD variants was induced via the addition of IPTG to a final concentration of 0.4 mM and the plate was incubated for 12-15 hours at 28 °C and 350 RPM. Cells were harvested by centrifugation at 3,000 *g* for 20 min at 4 °C, the cell pellets were thoroughly resuspended in 250 µL of 50 mM Tris-HCl (pH 8.0) containing 5 mg mL^{-1} lysozyme (Sigma Aldrich, St. Louis, MO) at 4 °C, and the plates were subjected to a single freeze-thaw cycle. Cell debris was then collected by centrifugation at 3,000g for 20 min at 4 °C, and either 25 or 100 µL of the cleared supernatant was utilized for the various enzyme assays (**Table 4.5**).

All assays were conducted in 50 mM Tris-HCl buffer (pH 8.0) and a final volume of 200 µL. For the assays, cleared lysate was added to various concentrations and combinations of 2-chloro-4-nitrophenyl glycoside donor and nucleotide diphosphate acceptor (**Table 4.5**). Upon mixing, the production of 2-chloro-4-nitrophenolate was followed by absorbance at 410 nm over a period of up to 48 hours at 25 °C.

4.4.5.2. Statistical Analysis. The initial slope and area under the curve for the absorbance data were calculated individually for each OleD variant. Area under the curve for each reaction was determined with the following equation (*1*):

$$(1) \qquad \mathrm{A} = y_{k=1} + 2\left(\sum_{k=2}^{n-1} y_k\right) + y_{k=n}$$

where A equals the total area of the curve and y_k equals the absorbance value at the k^{th} timepoint.

2-chloro-4-nitrophenyl glycoside donor (final conc.)	NDP acceptor (final conc.)	Volume of cleared lysate added to assay	Assay length
1 (0.5 mM)	UDP (0.5 mM)	25 µL	3 hours
1 (0.5 mM)	TDP (0.5 mM)	25 µL	3 hours
1 (0.5 mM)	CDP (0.25 mM)	25 µL	3 hours
1 (0.5 mM)	ADP (0.5 mM)	25 µL	3 hours
1 (0.5 mM)	GDP (0.5 mM)	25 µL	3 hours
2 (0.25 mM)	UDP (0.125 mM)	25 µL	48 hours
3 (0.5 mM)	UDP (0.75 mM)	100 µL	33 hours
4 (0.5 mM)	UDP (0.75 mM)	100 µL	0.75 hours
5 (0.5 mM)	UDP (0.75 mM)	100 µL	3 hours
6 (0.25 mM)	UDP (0.125 mM)	25 µL	48 hours
7 (0.25 mM)	UDP (0.125 mM)	25 µL	48 hours
8 (0.5 mM)	UDP (0.125 mM)	25 µL	48 hours
9 (0.5 mM)	UDP (1.5 mM)	100 µL	48 hours
10 (0.5 mM)	UDP (1.5 mM)	100 µL	48 hours

Table 4.5. Components and concentrations of primary assay mixtures. All reactions were run in Tris-HCl buffer (50 mM, pH 8.0) and a final total volume of 100 µL.

'Hits' which performed above the TDP-16 parental sequence and ranked high under both initial rate and area under the curve for individual substrate sets were selected for secondary screening and validation. Intra-plate standard deviations for the TDP-16 parental sequence absorbance data across all substrate combinations were typically < 15% (n=2) over the course of the respective assay.

4.4.6. Protein Expression and Purification. A single protocol based upon previously published methods for OleD was utilized for all purifications[8, 10, 11, 13]. Specifically, single isolates of *E. coli* BL21(DE3)pLysS (Stratagene, La Jolla, CA, USA) transformed with pET28a/*oleD*[P67T/S132F/A242L/Q268V] (produces OleD variant TDP-16) or any additional *oleD* variant investigated in this study, were utilized for protein expression and purification. Briefly, single colonies were used to inoculate 5 mL LB starter cultures and incubated overnight at 37 °C and 250 RPM. A 4 mL aliquot of the saturated starter culture was transferred to 0.5 L cultures of Luria-Bertani medium grown at 37 °C until the OD_{600} reached ~0.7. IPTG (0.4 mM final concentration) was added and cultures were incubated at 28 °C for approximately 18 hours at 250 RPM.

Cell pellets were collected by centrifugation (6,000 *g* at 4 °C for 20 min), resuspended in 25 mL of chilled lysis buffer (20 mM phosphate buffer, pH 7.4, 0.5 M NaCl, 10 mM imidazole), and lysed by sonication (5 pulses of 1 minute each) in an ice bath. Cell debris was removed by centrifugation (10,000 *g* at 4 °C for 20 min) and the cleared supernatant was incubated with alkaline phosphatase (0.4 U ml^{-1} of supernatant; from Roche, Basel, Switzerland) on ice for 1 hour with agitation to degrade contaminating nucleotide diphosphates.

Following, the supernatant was applied to 1.5 mL of nickel nitrilotriacetic acid resin (from QIAgen Valencia, CA, USA) pre-equilibrated with wash buffer (20 mM phosphate buffer, pH 7.4, 0.3 M NaCl, 10 mM imidazole). Protein was allowed to bind for 30 min at 4 °C and the resin washed with 35 mL wash buffer. Finally, the enzyme was eluted with 2.5 mL of chilled wash buffer containing an additional 240 mM imidazole for 10 min at 4 °C. Purified protein was applied to a PD-10 desalting column (Amersham Biosciences, Piscataway, NJ, USA), equilibrated with Tris-HCl buffer (10 mM, pH 8.0), and eluted as described by the manufacturer to typically provide 3.5 mL of desired protein. The purified protein was then concentrated with an Amicon Ultracel 30 kDa molecular weight cut-off centrifugal filter (Millipore, Billerica, MA, USA) to a final volume ranging from 750 to 1000 μL. Final purified proteins were flash frozen dropwise in liquid nitrogen and stored at -80 °C. Purity was confirmed by SDS-PAGE to be >95% homogeneous and concentration was determined using the Bradford Protein Assay Kit (Bio-Rad, Hercules, CA, USA). Typical concentrations of final proteins ranged from 7 to 25 mg mL^{-1}. Small aliquots of protein were thawed for experiments as required and did not undergo multiple freeze/thaw cycles.

4.4.7. Purification and characterization of NDP-sugars. Reactions containing ***a)*** 1 mM UDP and 1 mM of 1 mM 2-chloro-4-nitrophenyl glycoside (**4** or **5**) or ***b)*** 1 mM NDP (CDP, ADP or GDP) and 1 mM 2-chloro-4-nitrophenyl-β-D-glucopyranoside (**1**) with 10.5 μM (50 μg) of purified OleD variant TDP-16 in Tris buffer (50 mM, pH 8.0) and a final volume of 100 μL were prepared. Reactions were allowed to proceed at room temperature for 1.5 hours (with **4** as donor) or 12 hours (remaining reactions). Following, a volume of 100 μL of ddH_2O was added to each

reaction and samples were filtered through a MultiScreen Filter Plate (Millipore, Billerica, Maryland, USA) for 2 hours at 2000 *g*.

NDP-sugar products (**13-14**, **17**, **18**) from the cleared supernatant were isolated by analytical reverse-phase HPLC with a 250 mm x 4.6 mm Gemini-NX 5μ C18 column (Phenomenex, Torrance, CA, USA) using a linear gradient of 0% to 50% CH_3CN (solvent B) over 25 minutes (solvent A = 50 mM PO_4^{-2}, 5 mM tetrabutylammonium bisulfate, 2% acetonitrile [pH adjusted to 6.0 with KOH]; flow rate = 1 mL min^{-1}; A_{254} nm). CDP-glucose (**12**) was isolated a 250 mm x 4.6 mm Inertsil ODS-4 3μ C18 column (GL Sciences, Inc., Torrance, CA, USA) using a gradient of 0% B for 30 min, 0% to 100% B over 3 min, and 100% B for 22 min (solvent A = 100 mM PO_4^{-2}, 8 mM tetrabutylammonium bisulfate [pH adjusted to 6.4 with KOH]; solvent B = 70% buffer A, 30% CH_3CN [after mixing components, pH was adjusted to 6.4]; flow rate = 0.5 mL min^{-1}; A_{254} nm). In each case, fractions corresponding to the desired product were collected, frozen, and lyophilized.

Following, all samples were dissolved in 200 μL of ddH_2O and then re-injected onto a Gemini-NX C-18 (5 μm, 250 x 4.6 mm) column (Phenomenex, Torrance, California, USA) with a gradient of 0% to 12.5% CH_3CN (solvent B) over 25 min (A = 50 mM triethylammonium acetate buffer; flow rate = 1 mL min^{-1}; A_{254} nm). Again, fractions corresponding to the desired products were collected, frozen, and lyophilized. Samples were then dissolved in 1 mL of ddH_2O, frozen, and lyophilized (x2). Final products were dissolved in 1:1 ddH_2O/acetonitrile to a final concentration of 1 μg mL^{-1} and submitted for mass analysis. NDP-sugar products **10**, **11**, **15**, **16**, and **19** were previously characterized with TDP-16 as catalyst[8].

4.4.8. Secondary and Tertiary Screening. In a final volume of 200 μL, reactions containing 2-chloro-4-nitrophenyl glycoside donor, NDP acceptor, and purified catalyst in Tris-HCl buffer (50 mM, pH 8.0) were prepared in 96-well microtiter plates (**Table 4.6**, **4.7**). Following the addition of equivalent amounts of each catalyst, absorbance readings at 410 nm were recorded for a minimum of 3 hours. Assay conditions were optimized for each substrate pair to maximize the difference in absorbance signal between the parental catalyst TDP-16 (n = 2) and the best performing variants (n = 1). Statistical analysis (see section **4.4.5.2**) was utilized to identify and quantify the top performers. Standard deviations of absorbance readings for TDP-16 (n = 2) typically varied by < 5% for each data point over the course of the assay.

4.4.9. Determination of kinetic parameters. Assays were performed in a final volume of 100 μL in Tris-HCl (50 mM, pH 8.0) with either 2-chloro-4-nitrophenyl glycoside donor or NDP acceptor saturating and the corresponding reactant varied (from 0 mM until saturation conditions or solubility limits were met) and a defined amount of enzyme (either TDP-16 or Loki variant) for each substrate pair. A summary of the assay conditions utilized for determination of kinetic parameters is provided in **Table 4.8**, **4.9**. Reactions were followed at 410 nm where the rate of 2-chloro-4-nitrophenolate formation was determined to be linear (<20 min for all reactions). Initial rates were converted to product formation per unit of enzyme per unit time by comparing values to a standard curve. At a minimum, all experiments were performed in duplicate. Initial velocities were fit to the Michaelis-Menten equation using Origin Pro 7.0 software. Neither the TDP-16 or Loki variants could be saturated with glycoside donors **2**-**6** due to limited solubility (under the stated conditions: 5 mM for **2** and **3**, 2 mM for **4** and **5**, and 10 mM for **6**). Consequentially, k_{cat}/K_m for compounds **2**-**6** was determined by linear regression. Reactions

2-chloro-4-nitrophenyl glycoside donor (final conc.)	NDP acceptor (final conc.)	Catalyst concentration	Assay length
1 (0.5 mM)	UDP (0.5 mM)	0.2 μM (2 μg)	3 hours
1 (0.5 mM)	TDP (0.5 mM)	0.2 μM (2 μg)	3 hours
1 (0.5 mM)	CDP (0.25 mM)	2.2 μM (20 μg)	3 hours
1 (0.5 mM)	ADP (0.5 mM)	2.2 μM (20 μg)	3 hours
1 (0.5 mM)	GDP (0.5 mM)	2.2 μM (20 μg)	3 hours
2 (0.25 mM)	UDP (0.375 mM)	2.2 μM (20 μg)	6 hours
3 (0.5 mM)	UDP (0.75 mM)	2.2 μM (20 μg)	3 hours
4 (0.5 mM)	UDP (0.75 mM)	0.2 μM (2 μg)	3 hours
5 (0.5 mM)	UDP (0.75 mM)	2.2 μM (20 μg)	3 hours
6 (0.25 mM)	UDP (0.375 mM)	5.4 μM (50 μg)	6 hours

Table 4.6. Assay conditions utilized for secondary screening. All reactions were run in Tris-HCl buffer (50 mM, pH 8.0) and a final total volume of 200 μL.

2-chloro-4-nitrophenyl glycoside donor (final conc.)	NDP acceptor (final conc.)	Catalyst concentration	Assay length
1 (0.5 mM)	UDP (0.5 mM)	0.1 μM (1 μg)	3 hours
1 (0.5 mM)	TDP (0.5 mM)	0.1 μM (1 μg)	3 hours
1 (0.5 mM)	CDP (0.5 mM)	0.1 μM (1 μg)	3 hours
1 (0.5 mM)	ADP (0.5 mM)	0.1 μM (1 μg)	3 hours
1 (0.5 mM)	GDP (0.5 mM)	0.1 μM (1 μg)	3 hours
2 (0.25 mM)	UDP (1 mM)	2.2 μM (20 μg)	6 hours
3 (0.5 mM)	UDP (1 mM)	2.2 μM (20 μg)	3 hours
4 (0.5 mM)	UDP (1 mM)	0.1 μM (1 μg)	3 hours
5 (0.5 mM)	UDP (1 mM)	1.1 μM (10 μg)	3 hours
6 (0.25 mM)	UDP (1 mM)	5.4 μM (50 μg)	6 hours

Table 4.7. Assay conditions utilized for tertiary screening. All reactions were run in Tris-HCl buffer (50 mM, pH 8.0) and a final total volume of 200 μL.

Enzyme	Enzyme Concentration	Saturating Substrate	Saturating Substrate Concentration	Varied Substrate	Varied Substrate Concentration Range
TDP-16	0.4 μM (2 μg)	UDP	2 mM	**1**	0 to 15 mM
	0.4 μM (2 μg)	dTDP	2 mM	**1**	0 to 15 mM
	5.4 μM (25 μg)	CDP	20 mM	**1**	0 to 15 mM
	1.1 μM (5 μg)	ADP	20 mM	**1**	0 to 22.5 mM
	1.1 μM (5 μg)	GDP	20 mM	**1**	0 to 22.5 mM
	0.4 μM (2 μg)	**1**	15 mM	UDP	0 to 2 mM
	0.4 μM (2 μg)	**1**	15 mM	dTDP	0 to 2 mM
	5.4 μM (25 μg)	**1**	15 mM	CDP	0 to 20 mM
	1.1 μM (5 μg)	**1**	15 mM	ADP	0 to 20 mM
	1.1 μM (5 μg)	**1**	15 mM	GDP	0 to 20 mM
Loki	0.2 μM (1 μg)	UDP	2 mM	**1**	0 to 15 mM
	0.2 μM (1 μg)	dTDP	2 mM	**1**	0 to 15 mM
	0.2 μM (1 μg)	CDP	20 mM	**1**	0 to 15 mM
	0.2 μM (1 μg)	ADP	10 mM	**1**	0 to 15 mM
	0.2 μM (1 μg)	GDP	10 mM	**1**	0 to 15 mM
	0.2 μM (1 μg)	**1**	15 mM	UDP	0 to 2 mM
	0.2 μM (1 μg)	**1**	15 mM	dTDP	0 to 2 mM
	0.2 μM (1 μg)	**1**	15 mM	CDP	0 to 20 mM
	0.2 μM (1 μg)	**1**	15 mM	ADP	0 to 10 mM
	0.2 μM (1 μg)	**1**	15 mM	GDP	0 to 10 mM

Table 4.8. Assay conditions for determining kinetic parameters with various NDPs as acceptors. All reactions were run with saturating 2-chloro-4-nitrophenyl glucopyranoside (**1**) as donor and in Tris-HCl buffer (50 mM, pH 8.0) and a final total volume of 100 μL.

Enzyme	Enzyme Concentration	Saturating Substrate	Saturating Substrate Concentration	Varied Substrate	Varied Substrate Concentration Range
TDP-16	0.4 μM (2 μg)	UDP	2 mM	**1**	0 to 15 mM
	10.8 μM (50 μg)	UDP	2 mM	**2**	0 to 2 mM
	10.8 μM (50 μg)	UDP	2 mM	**3**	0 to 2.5 mM
	0.2 μM (1 μg)	UDP	2 mM	**4**	0 to 1 mM
	10.8 μM (50 μg)	UDP	2 mM	**5**	0 to 1 mM
	10.8 μM (50 μg)	UDP	2 mM	**6**	0 to 2 mM
Loki	0.2 μM (1 μg)	UDP	2 mM	**1**	0 to 15 mM
	2.2 μM (10 μg)	UDP	2 mM	**2**	0 to 2 mM
	2.2 μM (10 μg)	UDP	2 mM	**3**	0 to 2.5 mM
	0.2 μM (1 μg)	UDP	2 mM	**4**	0 to 1 mM
	2.2 μM (10 μg)	UDP	2 mM	**5**	0 to 1 mM
	5.4 μM (25 μg)	UDP	2 mM	**6**	0 to 2 mM

Table 4.9. Assay conditions for determining kinetic parameters with various sugar donors. 2-chloro-4-nitrophenyl glycosides (**1**-**6**) were utilized as donors with saturating UDP as acceptor. All reactions were run in Tris-HCl buffer (50 mM, pH 8.0) and a final total volume of 100 μL.

containing **4** were corrected for a competing hydrolytic reaction by comparing them to identical reactions lacking catalyst. Errors associated with k_{cat}/K_m were calculated according to standard methods of error propagation; the fractional error of k_{cat}/K_m,

$$\mu\frac{k_{cat}}{K_m},$$

was determined with the equation (*2*):

$$\mu\frac{k_{cat}}{K_m} = \sqrt{\left(\frac{\mu k_{cat}}{k_{cat}}\right)^2 + \left(\frac{\mu K_m}{K_m}\right)^2} \quad (2)$$

where μk_{cat} is the standard deviation of k_{cat} and μK_m is the standard deviation of K_m. Values obtained are in agreement with kinetic parameters from previous studies with numerous variants of OleD[8, 10, 11, 13].

4.4.10. Combinatorial NDP-sugar Synthesis. Reactions containing 5.25 μM (50 μg) of Loki variant and each individual combination of 0.5 mM 2-chloro-4-nitrophenyl donor (**1**-**6**) and 1 mM NDP acceptor (UDP, TDP, CDP, ADP, or GDP) were prepared in Tris-HCl buffer (50 mM, pH 8.0) with a final volume of 200 μL in a flat bottom 96-well microtiter plate. Following the addition of catalyst, the change in absorbance was followed at 410 nm. Identical reactions containing triethylammonium acetate buffer (50 mM, pH 7.0) in place of Tris-HCl buffer were run simultaneously.

When reactions with Tris-HCl buffer reached completion (determined by comparing the absorbance reading to a standard curve), absorbance readings remained constant over 2 or more readings, or the total reaction time reached 24 hours, reactions with triethylammonium acetate

buffer were treated with 20 U of alkaline phosphatase (Roche, Basel, Switzerland) for 30 minutes at RT. Following, reactions from both buffer systems were frozen at -80 °C. After thawing at 4 °C, samples were filtered through a MultiScreen Filter plate with Ultracel-10 Membrane (Millipore, Billerica, MA, USA) according to manufacturer's instructions. The cleared supernatant from each sample with Tris-HCl was evaluated for formation of the expected NDP-sugars with analytical reverse phase HPLC. Formation of UDP- (**10**, **15**-**19**), TDP- (**11**, **20**-**24**), ADP- (**13**, **30**-**34**), and GDP-sugars (**14**, **35**-**39**) were analyzed with a 250 mm x 4.6 mm Gemini-NX 5µ C18 column while CDP-sugars (**12**, **25**-**27**, **29**) were analyzed with a 250 mm x 4.6 mm Inertsil ODS-4 3µ C18 column (see section **4.4.7** for HPLC methods). Formation of **28** was confirmed on a 250 mm x 4.6 mm Gemini-NX 5µ C18 column with a gradient of 0% CH_3CN (solvent B) for 15 min, 0% to 25% B over 10 min, and 25% to 90% B over 5 min (A = 50 mM triethylammonium acetate buffer; flow rate = 1 mL min^{-1}; A_{254} nm). The standard deviation for production of UDP-α-D-glucose (**10**) across multiple runs (n=3) was <10%.

The cleared supernatant from each sample with triethylammonium acetate buffer was lyophilized and resuspended in water (3 x 500 µL). Following the final lyophilization, the crude reactions were resuspended in 30 µL of dH_2O, diluted 1:10 with 1 mM ammonium acetate in 85% CH_3CN / 15% dH_2O (pH 5.5), and then 5 µL was injected for LC/MS analysis. The LC/MS method consisted of an Acquity BEH 1.7 µm Amide column (2.1 x 100 mm; Waters Corp., Milford, MA, USA) with a gradient of 75% to 25% B over 5 min, 25% B for 3 min, and 25% to 75% B over 0.2 min, 75% B for 10 min (solvent A = 1.0 mM ammonium acetate in 65% CH_3CN / 35% dH_2O [pH 5.5]; solvent B = 1.0 mM ammonium acetate in 85% CH_3CN / 15% dH_2O [pH 5.5]; flow rate = 1.0 mL min^{-1}; A_{254} nm) while HRMS spectra were collected over the course of the separation (see section **4.4.1.2**). Retention times for NDP-sugars were typically 4-7 min.

4.5. References

1. Varki, A. *et al.* Eds., *Essentials of Glycobiology* (Cold Spring Harbor, New York, ed. 2, 2009).

2. Thibodeaux, C.J., Melançon III, C.E., & Liu, H. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**, 1008-1016 (2007).

3. Wagner, G.K., Pesnot, T., & Field, R.A. A survey of chemical methods for sugar-nucleotide synthesis. *Nat. Prod. Rep.* **26**, 1172-1174 (2009).

4. Gantt, R.W., Peltier-Pain, P., & Thorson, J.S. Enzymatic methods for glyco(diversification/randomization) of drugs and small molecules. *Nat. Prod. Rep.* **28**, 1811-1853 (2011).

5. Thibodeaux, C.J., Melançon III, C.E., & Liu, H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed.* **47**, 9814-9859 (2008).

6. Rupprath, C., Schumacher, T., & Elling, L. Nucleotide deoxysugars: essential tools for the glycosylation engineering of novel bioactive compounds. *Curr. Med. Chem.* **12**, 1637-1675 (2005).

7. Pérez, M. *et al.* Combinatorial biosynthesis of antitumor deoxysugar pathways in *Streptomyces griseus*: reconstitution of ”unnatural natural gene clusters” for the biosynthesis of four 2,6-dideoxyhexoses. *Appl. Envir. Microbiol.* **72**, 6644-6652 (2006).

8. Gantt, R.W., Peltier-Pain, P.P., Cournoyer, W.J., and Thorson, J.S. Using simple donors to drive the equilibria of glycosyltransferase-catalyzed reactions. *Nat. Chem. Biol.* **7**, 685-691 (2011).

9. Williams, G.J., Yang, J., Zhang, C., & Thorson, J.S. Recombinant *E. coli* prototype strains for *in vivo* glycorandomization. *ACS Chem. Biol.* **6**, 95-100 (2011).

10. Williams, G.J., Zhang, C., & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

11. Williams, G.J., Goff, R.D., Zhang, C., & Thorson, J.S. Optimizing glycosyltransferase specificity via “hot spot” saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem. Biol.* **15**, 393-401 (2008).

12. Williams, G.J. & Thorson, J.S. A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Protocols* **3**, 357-362 (2008).

13. Gantt, R.W., Goff, R.D., Williams, G.J., & Thorson, J.S. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem. Int. Ed.* **47**, 8889-8892 (2008).

14. Bolam, D.N. *et al.* 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 (2007).

15. Quirós, L.M., Carbajo, R.J., Braña, A.F., & Salas, J.A. Glycosylation of macrolide antibiotics: purification and kinetic studies of a macrolide glycosyltransferase from *Streptomyces antibioticus*. *J. Biol. Chem.* **275**, 11713-11720 (2000).

16. Brazier-Hicks, M., *et al.* 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 (2007).

17. Mulichak, A.M., Lu, W., Losey, H.C., Walsh, C.T., & Garavito, R.M. Crystal structure of vancosaminyltransferase GtfD from the vancomycin biosynthetic pathway: interactions with acceptor and nucleotide ligands. Biochemistry 43, 5170–5180 (2004).

18. Hoffmeister, D. *et al.* Engineered urdamycin glycosyltransferases are broadened and altered in substrate specificity. Chem. Biol. **9**, 287–295 (2002).

19. Gregoret, L.M., Rader, S.D., Fletterick, R.J., & Cohen, F.E. Hydrogen bonds involving sulfur atoms in proteins. *Proteins* **9**, 99-107 (1991).

20. Zhou, P., Tian, F., Lv, F., & Shang, Z. Hydrogen bonds involving sulfur atoms in proteins. *Proteins* **76**, 151-163 (2009).

21. Shao, H., *et al.* Crystal structures of a multifunctional triterpene/flavanoid glycosyltransferase from *Medicago truncatula*. *Plant Cell* 17, 3141–3154 (2005).

22. Offen, W. *et al.* Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. *EMBO J.* **25**, 1396–1405 (2006).

23. Minami, A. & Eguchi, T. Substrate flexibility of vicenisaminyltransferase VinC involved in the biosynthesis of vicenistatin. *J. Am. Chem. Soc.* **129**, 5102-5107 (2007).

24. Amann, S., Dräger, G., Rupprath, C., Kirschning, A., & Elling, L. (Chemo)enzymatic synthesis of dTDP-activated 2,6-dideoxysugars as building blocks of polyketide antibiotics. *Carb. Res.* **335**, 23-32 (2001).

25. Wang, G., Kharel, M.K., Pahari, P., & Rohr, J. Investigating mithramycin deoxysugar biosynthesis: enzymatic tool synthesis of TDP-D-olivose. *ChemBioChem* **12**, 2568-2571 (2011).

26. Westrich, L., *et al.* Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis. *FEMS Microbiol. Lett.* **170**, 381-387 (1999).

27. Hoffmeister, D., *et al.* 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 (2000).

28. Lombó, F., Menéndez, N., Salas, J.A., & Méndez, C. The aureolic acid family of antitumor compounds: structure, mode of action, biosynthesis, and novel derivatives. *Appl. Microbiol. Biotechnol.* **73**, 1-14 (2006).

29. Menendez, N., *et al.* Biosynthesis of the antitumor chromomycin A3 in *Streptomyces griseus*: analysis of the gene cluster and rational design of novel chromomycin analogs. *Chem. Biol.* **11**, 21-32 (2004).

30. Jia, X.Y., *et al.* Genetic characterization of the chlorothricin gene cluster as a model for spirotetronate antibiotic biosynthesis. *Chem. Biol.* **13**, 575-585 (2006).

31. Weitnauer, G., *et al.* Biosynthesis of the orthosomycin antibiotic avilamycin A: deductions from the molecular analysis of the avi biosynthetic gene cluster of *Streptomyces viridochromogenes* Tü57 and production of new antibiotics. *Chem. Biol.* **8**, 569-581 (2001).

32. Haydock, S.F., *et al.* Organization of the biosynthetic gene cluster for the macrolide concanamycin A in Streptomyces neyagawaensis ATCC 27449. *Microbiology* **151**, 3161-3169 (2005).

33. Mochizuki, S., *et al.* The large linear plasmid pSLA2-L of *Streptomyces rochei* has an unusually condensed gene organization for secondary metabolism. *Mol. Microbiol.* 48, 1501-1510 (2003).

34. Ward, S.L., *et al.* Chalcomycin biosynthesis gene cluster from *Streptomyces bikiniensis*: novel features of an unusual ketolide produced through expression of the chm polyketide synthase in *Streptomyces fradiae*. *Antimicrob. Agents Chemother.* **48**, 4703-4712 (2004).

35. Jaishy, B.P., *et al.* Cloning and characterization of a gene cluster for the production of polyketide macrolide dihydrochalcomycin in *Streptomyces* sp. KCTC 0041BP. *J. Microbiol. Biotechnol.* **16**, 764-770 (2006).

36. Summers, R.G., *et al.* Sequencing and mutagenesis of genes from the erythromycin biosynthetic gene cluster of *Saccharopolyspora erythraea* that are involved in L-mycarose and D-desosamine production. *Microbiology* **143**, 3251-3262 (1997).

37. Volchegursky, Y., Hu, Z., Katz, L., & McDaniel, R. Biosynthesis of the anti-parasitic agent megalomicin: transformation of erythromycin to megalomicin in *Saccharopolyspora erythraea*. *Mol. Microbiol.* **37**, 752-762 (2000).

38. Asaka, T., Manaka, A., & Sugiyama, H. Recent developments in macrolide antimicrobial research. *Curr. Top. Med. Chem.* **3**, 961-989 (2003)

39. Brazier-Hicks, M., Edwards, L.A., & Edwards, R. Selection of plants for roles in phytoremediation: the importance of glucosylation. *Plant Biotech. J.* **5**, 627-635 (2007).

40. Gandia-Herrero, F., *et al.* Detoxification of the explosive 2,4,6-trinitrotoluene in *Arabidopsis*: discovery of bifunctional O- and C-glucosyltransferases. *Plant J.* **56**, 963-974 (2008).

41. Yang, M., *et al.* Probing the breadth of macrolide glycosyltransferases: *in vitro* remodeling of a polyketide antibiotic creates active bacterial uptake and enhances potency. *J. Amer. Chem. Soc.* **127**, 9336-9337 (2005).

42. Truman, A.W., *et al.* Chimeric glycosyltransferases for the generation of hybrid glycopeptides. *Chem. Biol.* **26**, 676-685 (2009).

43. Park, S.-H., *et al.* Reconstitution of antibiotics glycosylation by domain exchanged chimeric glycosyltransferase. *J. Mol. Catal. B: Enzym.* **60**, 29-35 (2009).

44. Mulichak, A.M., Losey, H.C., Walsh, C.T., & Garavito, R.M. Structure of the UDP-glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure* **9**, 547–557 (2001).

45. Mulichak, A.M., *et al.* Structure of the TDP-epi-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9238–9243 (2003).

46. Wang, X. Structure, mechanism and engineering of plant natural product glycosyltransferases. *FEBS Lett.* **583**, 3303-3309 (2009).

47. Chang, A., *et al.* 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 (2011).

48. Hanessian, S., Ed., *Preperative carbohydrate chemistry* (Marcel Dekker, Inc., New York, 1996).

49. Lawton, B.T., Szarek, W.A., & Jones, J.K.N., A facile synthesis of 4,6-dideoxy-D-xylo-hexose. *Carb. Res.* **14**, 255-258 (1970).

50. van Summeren, R.P., Feringa, B.L., & Minnaard, A.J. New approaches towards the synthesis of the side-chain of mycolactones A and B. *Org. Biomol. Chem.* **3**, 2524-2533 (2005).

51. Lowary, T.L., Eichler, E., & Bundle, D.R. Oligosaccharide recognition by antibodies: synthesis and evaluation of talose oligosaccharide analogues. *Can. J. Chem.* **80**, 1112-1130 (2002).

Chapter 5:

Concluding Remarks and Future Directions

5.1. Concluding Remarks

Glycosyltransferases (GTs) play a critical role throughout biology as a whole[1, 2] and in drug discovery efforts. Investigations and applications of GTs, however, are largely limited by a lack of availability to their substrates, NDP-sugars. While many chemical, enzymatic, and chemoenzymatic routes are described in the literature[2-4], limited access to a range of structurally diverse members, particularly in meaningful quantities, remains a major restriction. Furthermore, GT characterizations which move beyond screening putative wild type aglycons and NDP-sugars to provide glimpses of substrate landscapes remain the exception and not the rule (Chapter 1 and references therein). Reported within this dissertation are studies aimed at identifying GTs capable of serving in forward reactions as general catalysts for drug discovery efforts, providing facile access to NDP-sugars through chemoenzymatic routes, and development of general glycosyl transfer screens.

Over 70 structurally diverse aglycons, which span a range of natural products and small molecule drug classes and include unique nucleophiles not previously known to participate in GT reactions, were confirmed as substrates for GT-catalyzed forward reactions with either OleD wild type or the variant ASP as catalyst and UDP-glucose as donor (Chapter 2). This not only established OleD as among the most promiscuous GTs studied to date, but also provided a brand new ‘toolbox’ of nucleophiles to consider for the inclusion of unique glycosidic bonds into future drug design and discovery efforts. Additionally, OleD was characterized as possessing iterative glycosylation activity for formation of disaccharide moieties (previously unreported) and was

confirmed to catalyze *O*-, *S*-, and *N*-glycosidic bond formation. Collectively, these studies of GT-catalyzed forward reactions with OleD set the stage for a thorough investigation of GT-catalyzed reverse reactions.

GT-catalyzed reverse reactions, one of newer routes for NDP-sugar synthesis (Chapter 1), have been largely underutilized due to unfavorable thermodynamics and are typically limited to generation of single NDP-sugar from a complex glycosylated natural product and its respective GT. Using knowledge gained through investigations of GT-catalyzed forward reactions (Chapter 2), studies were designed to expand and generalize GT-catalyzed reverse reactions as an approach for accessing NDP-sugars in a facile manner by addressing both the thermodynamic and substrate availability issues. By screening a library of potential glycoside donors, a range of suitable leaving groups were identified which enabled the production of desired NDP-sugars through a systematic modulation of the GT-catalyzed reaction thermodynamics. As a direct result of these findings, milligram scale syntheses of final products was demonstrated (Chapter 3) and over 40 distinct natural and 'unnatural' NDP-sugars were accessed through GT-catalyzed reverse reactions (Chapter 3 and 4). Additionally, single and double enzyme coupled systems were utilized for *in situ* NDP-sugar formation for glycorandomization of simple phenolic and complex natural product targets, respectively. Most importantly, thermodynamically-favorable 2-chloro-4-nitrophenyl glycoside donors enabled the development of high throughput colorimetric assays for both NDP-sugar formation and general glycosyl transfer (Chapter 3). This screen, along with the known aglycon recognition plasticity of OleD variants (Chapter 2), allowed for demonstration of a high throughput colorimetric single enzyme coupled assay for drug discovery with a library consisting of 50 medicinal compounds (Chapter 3).

During the optimization of GT-catalyzed reverse reactions in these studies, a decisive

factor in demonstrating NDP-sugar formation was identification of substrates which provided the greatest thermodynamic advantage. While this thermodynamic advantage will not always translate to an equivalent kinetic advantage, thermodynamic parameters are not as plastic as those of a kinetic nature in biological systems. Thermodynamic parameters can only be addressed through alteration of the starting substrates or physical environment of reactions, the latter of which, in biological systems, is largely limited by the relatively small ranges of physical parameters (*e.g.*, temperature, pH, ionic strength, etc.) where these systems typically function. Kinetic parameters (*i.e.*, k_{cat}, K_m), however, can be easily modified by identification of improved enzyme variants towards a selected substrate with directed evolution or engineering techniques. Future studies which attempt to exploit and/or engineer other novel enzyme-catalyzed reactions should benefit from these considerations.

With the initial success of GT-catalyzed reverse reactions, additional investigations were designed to identify improved catalysts against NDP acceptors and poor performing or previously uninvestigated sugar donors. Through applying the aforementioned colorimetric assay, three rounds of high throughput screening and recombination identified the OleD variant Loki, which was characterized with kinetic parameters markedly improved toward a range of glycoside donors and NDP acceptors (Chapter 4). Within these glycoside donors were deoxy- and dideoxy-motifs which are representative of sugars appended to medicinally relevant natural products. The Loki variant was also found to be capable of combinatorial synthesis with 6 glycoside donors and 5 NDP acceptors resulting in catalysis of 30 NDP-sugar products. While the sugar recognition of closely related OleD variants is known to be larger than the 6 glycoside donors investigated, these initial efforts reveal a glimpse of an anticipated broad sugar recognition landscape for the Loki variant.

5.2. Future Directions

More important than the direct knowledge gained during these studies are the new lines of inquiry and potential applications which have emerged. The most straightforward applications include screening against glycoside donors and NDP acceptors for formation of any desired NDP-sugar with any given GT. While over 40 distinct natural and 'unnatural' NDP-sugars are described in this work, motifs that are present in sugars appended to natural products at almost a ubiquitous level, such as amino-, *N*-methyl, *O*-methyl, or *C*-methyl groups were not investigated. Examples of galacto- and manno- configured D-sugars (C-2 and C-4 epimers of D-glucose, respectively) and L-sugars were also elusive. As addressed in Chapter 4, screening for formation of NDP-sugars with modified nucleotide moieties possessing unique chemical and/or analytical properties should also not be overlooked[1]. Given the validation of the colorimetric high throughput screen for identifying improved variants for NDP-sugar formation in this dissertation, the current limitations of the OleD sugar promiscuity can be addressed in a straightforward manner with additional engineering programs (**Figure 5.1a**).

The observed plasticity and known limitations of OleD and its variants in the context of forward and reverse GT-catalyzed reactions also raises a number of questions. First, which other naturally occurring GTs are capable of recognizing small phenolic glycosides as donors in reverse reactions for generation of NDP-sugars? If other GTs are identified, will they exclusively possess the inverting reaction mechanism of OleD or will they possess retaining or other proposed mechanisms characterized within the GT superfamily[2, 5]? Secondly, what are the effects on substrate recognition in the context of both forward and reverse reactions when the numerous advantageous mutations identified from OleD engineering efforts are introduced into other GTs? Thirdly, can comparisons of OleD to other GT sequences and crystal structures help

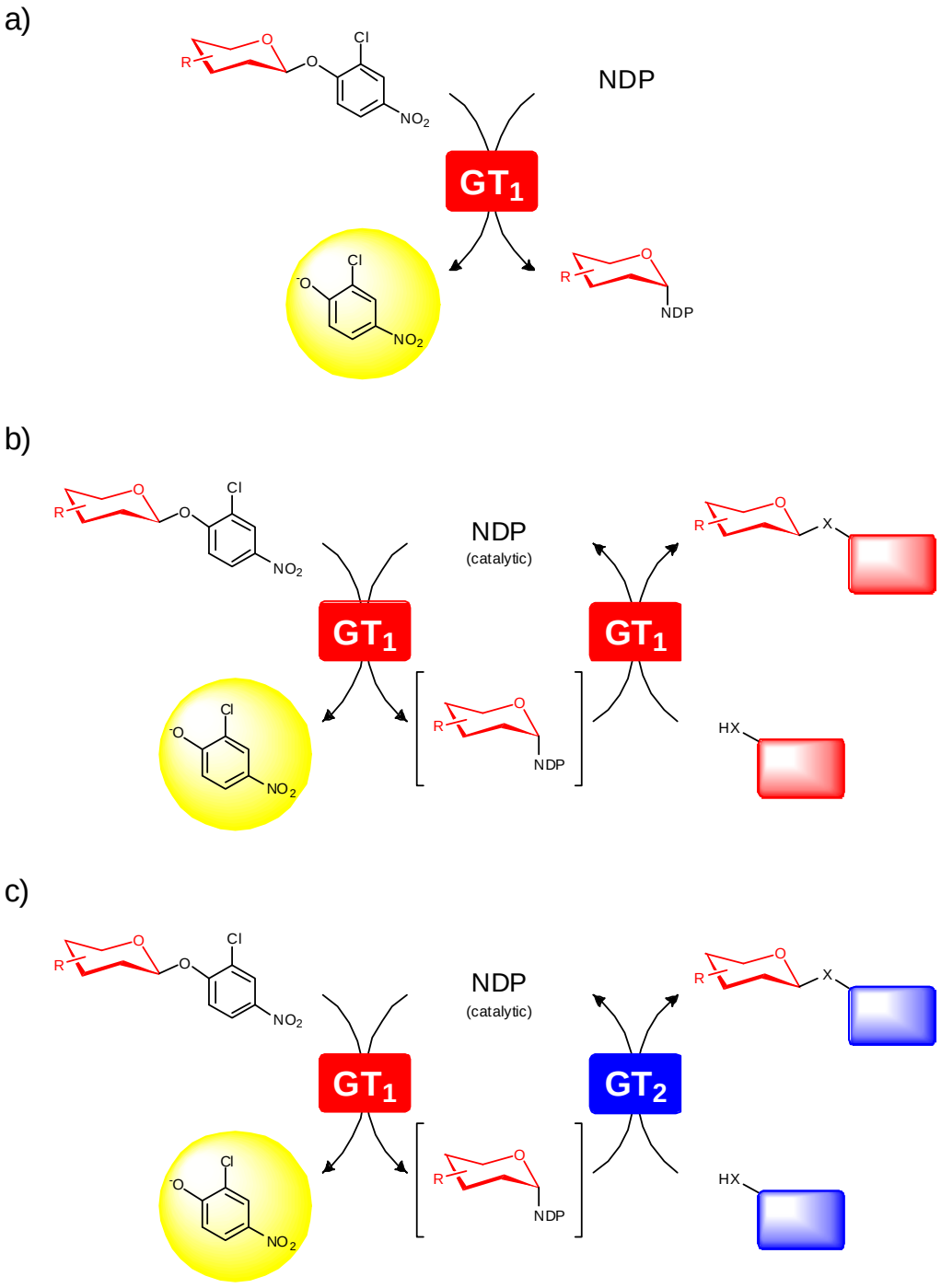


Figure 5.1. GT-catalyzed reverse reactions and single and dual coupled enzyme systems. **a)** Single enzyme system for detection of NDP-sugar formation. **b)** Single enzyme coupled reaction for detection of glycosyl transfer to any potential final aglycon acceptor. **c)** Dual enzyme coupled reaction for detection of glycosyl transfer to any potential final aglycon acceptor. X = O-, S-, NH-, or NHR-.

identify the current limitations of sugar recognition or are other approaches, such as generation of GT-chimeras or identification of other GT wild type scaffolds, necessary? With the colorimetric assay systems described herein, these questions can be probed to help explain not only why OleD and its variants are so promiscuous, but also how to impart this promiscuity into other GTs.

The potential applications available for the coupled reaction systems described in this work are enormous. The approach for both the single and double enzyme coupled reactions can be easily generalized for glycorandomization efforts of any targeted aglycon (**Figure 5.1b-c**). Additionally, with arising methods for identification of compounds in complex mixtures, screening against multiple sugars and/or multiple aglycons in a single reaction is now a feasible proposition (**Figure 5.2a-b**). Such an approach would greatly expedite drug discovery screening efforts, especially with the large sizes of contemporary drug libraries (>1 x 10^5 compounds).

The double enzyme coupled reaction system has perhaps the most powerful range of potential applications. In addition to screening for any potential downstream glycosyl transfer event, this approach should also be applicable to NDP-sugar modifying enzymes (**Figure 5.2c**), greatly expediting the elucidation of natural product biosynthetic pathways. Such approaches might include *in situ* formation and utilization of innately unstable NDP-sugar pathway intermediates for real time mass spectrometry or NMR studies. Directed evolution or engineering of any downstream enzyme with this approach is also now possible, potentially allowing facile identification of improved variants for entire classes of enzymes which may have relatively few or no high throughput screens (**Figure 5.2c**). The insights generated from investigations presented in this dissertation should serve to inspire future efforts into understanding and harassing glycosylation's unique role in biology and drug discovery.

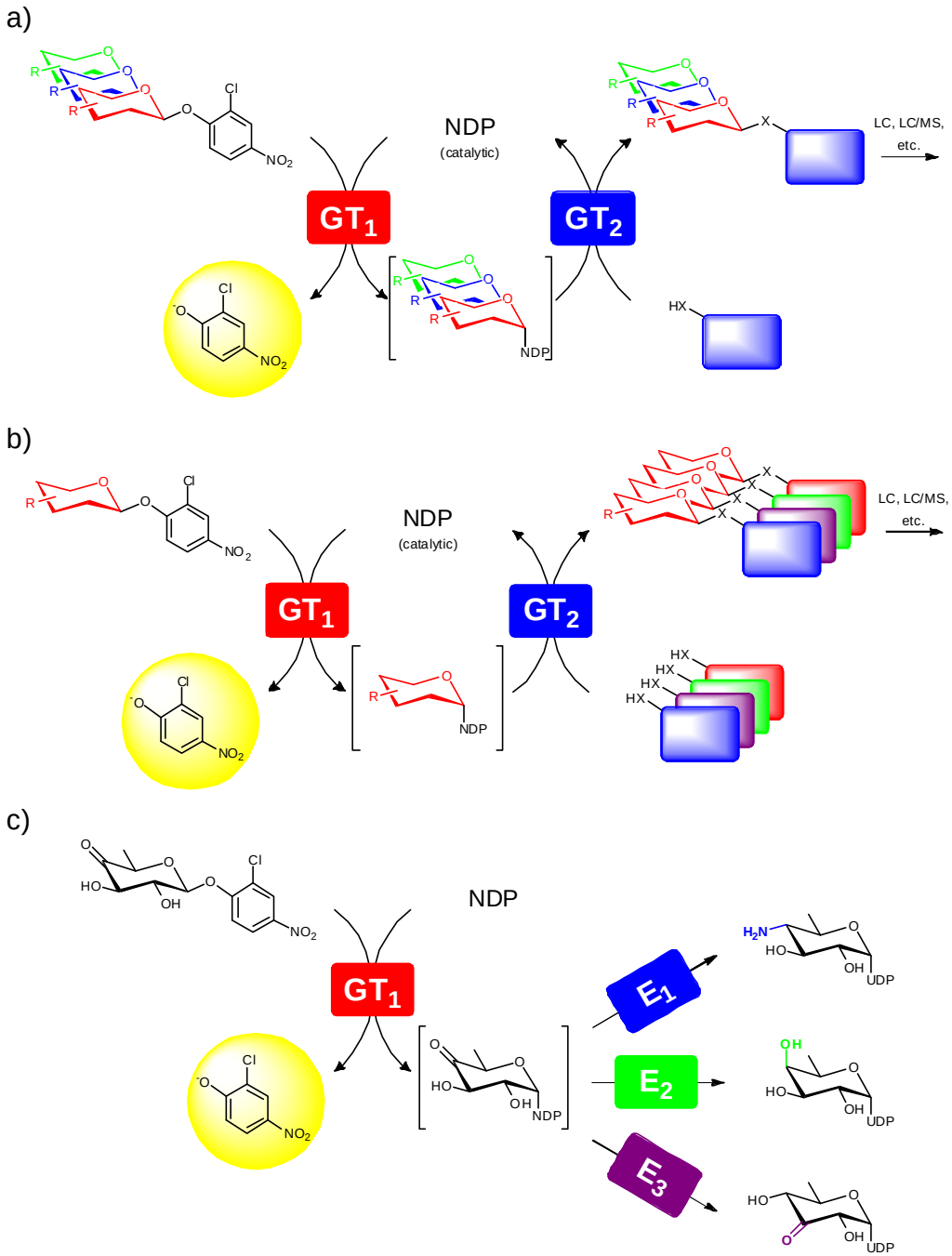


Figure 5.2. Potential applications of coupled enzyme systems. **a)** Screening for glycosyl transfer of numerous sugars to a single aglycon target in a single reaction. Confirmation of which sugars are transferred is carried out with various analytical techniques (*e.g.*, MS, LC/MS). **b)** Screening for glycosyl transfer of a single sugar to a numerous aglycon targets in a single reaction. Confirmation of which aglycons are glycosylated is carried out with various analytical techniques (*e.g.*, MS, LC/MS). **c)** Screening for activity of NDP-sugar utilizing enzymes. A theoretical scheme of using a glycoside donor to generate NDP-4-keto-6-deoxy-α-D-glucose for biosynthetic investigations or engineering of an aminotransferase (E_1), ketoreductase (E_2), or ketoisomerase (E_3) enzymes. X = O-, S-, NH-, or NHR-.

5.3. References

1. Varki, A. *et al.* Eds., *Essentials of Glycobiology* (Cold Spring Harbor, New York, ed. 2, 2009).

2. Thibodeaux, C.J., Melançon III, C.E., & Liu, H. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**, 1008-1016 (2007).

3. Wagner, G.K., Pesnot, T., & Field, R.A. A survey of chemical methods for sugar-nucleotide synthesis. *Nat. Prod. Rep.* **26**, 1172-1174 (2009).

4. Thibodeaux, C.J., Melançon III, C.E., & Liu, H. Natural-product sugar biosynthesis and enzymatic glycodiversification. *Angew. Chem. Int. Ed.* **47**, 9814-9859 (2008).

5. Lairson, L.L., Henrissat, B., Davies, G.J. & Withers, S.G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521-555 (2008).

Appendix 1:

Supplementary Data for Chapter 2

A1.1. Supporting Data for 137 Member Acceptor Library Screen

Compound Number	Compound Name	WT % conversion[a]	'ASP' % conversion[a]	substrate retention time	calculated substrate mass (*m/z*)	observed substrate mass (*m/z*)		product retention time	expected product mass (*m/z*)	observed product mass (*m/z*)	product designation[b]	HPLC method[c]	comments
3	benzhydrol	>99.9	>99.9	19.7	184.2	n.f.		12.9	508.2	531.0 $[M+Na]^+$	DI	A	
								13.3	508.2	531.0 $[M+Na]^+$	DI		
								14.0	346.4	381.2 $[M+Cl]^-$	MONO		
4	β-zearalenol	91.8	>99.9	18.1	320.2	321.2 $[M+H]^+$		8.1	482.2	483.2 $[M+H]^+$	MONO	A	
								9.6	482.2	483.2 $[M+H]^+$	MONO		
								9.7	644.3	679.2 $[M+Cl]^-$	DI		
								10.5	644.3	667.7 $[M+Na]^+$	DI		
								11.7	482.2	483.2 $[M+H]^+$	MONO		
								12.7	644.3	643.2 $[M-H]^-$	DI		
								12.8	644.3	643.2 $[M-H]^-$	DI		
								13.9	482.2	483.2 $[M+H]^+$	MONO		
5	2-iodophenol	70	99.7	18.6	219.9	243.8 $[M+H]^+$		11.7	382.0	404.4 $[M+Na]$	MONO	A	
6	thiophenol	61.3	93.7	20.8	110.2	n.f.		8.4	434.1	469.0 $[M+Cl]^-$	DI	A	
								9.4	272.1	307.2 $[M+Cl]^-$	MONO		
7	6-hydroxy-4-methylcoumarin	21	92	13.0	176.2	175.2 $[M-H]^-$		8.7	338.3	373.2 $[M+Cl]^-$; 339.0 $[M+H]^+$	MONO	A	0.2 mM final concentration
8	phenol	17	89	12.8	94.0	129.5 $[M+Cl]^-$		7.9	256.3	291.2 $[M+Cl]^-$	MONO	A	
9	kaempferol	71.3	88.0	17.4	286.0	285.0 $[M-H]^-$		7.1	448.1	483.0 $[M+Cl]^-$	MONO	A	
								8.2	610.2	609.0 $[M-H]^-$	DI		
								8.6	610.2	609.0 $[M-H]^-$	DI		
								8.8	610.2	609.0 $[M-H]^-$	DI		
								9.5	610.2	609.0 $[M-H]^-$; 645.0 $[M+Cl]^-$	DI		
								10.0	610.2	609.0 $[M-H]^-$; 645.0 $[M+Cl]^-$	DI		
								10.9	610.2	609.0 $[M-H]^-$	DI		
								11.2	610.2	645.0 $[M+Cl]^-$	DI		
								11.8	610.2	645.0 $[M+Cl]^-$	DI		
								12.3	448.1	447.0 $[M-H]^-$	MONO		
								12.4	448.1	447.0 $[M-H]^-$	MONO		
								12.8	448.1	447.0 $[M-H]^-$	MONO		
10	1,8-dihydroxyanthraquinone	81	78	24.2	240.0	240.8 $[M+H]^+$		13.3	402.1	403.0 $[M+H]^+$	MONO	A	
11	digitoxin neoaglycon	79	77	14.8	403.3	404.0 $[M+H]^+$		12.8	565.3	566.0 $[M+H]^+$	MONO	A	0.25 mM final concentration
12	4-nitrophenol	33	74	15.4	139.0	138.4 $[M-H]^-$		9.6	301.3	336.4 $[M+Cl]^-$	MONO	A	
13	phenylhydrazine	5	73	13.3	108.1	109.2 $[M+H]^+$		11.3	270.1	270.8 $[M+H]^+$; 292.8 $[M+Na]^+$; 268.8 $[M-H]^-$; 304.8 $[M+Cl]^-$	MONO	C	
14	T_0 - calicheamicin fragment	0	69	12.3	584.1	582.8 [M-H]-		11.89	746.1	781.0 $[M+Cl]^-$	MONO	B	0.1 final concentration
15	*N*-hydroxybenzamide	25	66	11.8	137.1	138.2 $[M+H]^+$		13.8	299.1	300.2 $[M+H]^+$	MONO	J	
16	warfarin neoaglycon	6.2	64.9	21.4	339.2	340.6 $[M+H]^+$		13.7	501.2	524.2 $[M+Na]^+$	MONO	D	
								13.9	501.2	524.0 $[M+Na]^+$	MONO		
								14.2	501.2	524.0 $[M+Na]+$	MONO		
17	mitoxantrone	11	63	16.1	444.2	445.2 $[M+H]^+$		15.6	606.3	607.0 $[M+H]^+$	MONO	E	0.25 mM final concentration
18	β-estradiol	0	59.7	20.2	272.2	273.2 $[M+H]^+$		11.8	596.3	619.2 $[M+Na]^+$	DI	A	
								13.1	596.3	619.4 $[M+Na]^+$	DI		
								14.7	434.2	435.4 $[M+H]^+$	MONO		
19	diphenylamine	51.0	48.9	24.1	169.1	170.0 $[M+H]^+$		13.3	493.2	494.2 $[M+H]^+$; 516.0 $[M+Na]^+$; 491.8 $[M-H]^-$; 528.2 $[M+Cl]^-$	DI	G	
								14.0	331.1	331.6 $[M+H]^+$; 353.6 $[M+Na]^+$;	MONO		
20	phenoxyethanethiol	32.0	48.9	21.9	154.0	n.f.		8.2	478.2	513.0 $[M+Cl]^-$	DI	A	
								11.7	316.1	351.0 $[M+Cl]^-$	MONO		
21	2-naphthalenemethanol	80.3	46.1	21.5	158.1	n.f.		16.2	482.2	505.0 $[M+Na]^+$	DI	E	
								17.1	482.2	504.8 $[M+Na]^+$	DI		
								17.5	320.1	342.6 $[M+Na]$	MONO		
22	benzylmercaptan	37	45	21.4	286.1	n.f.		10.3	286.3	321.2 $[M+Cl]^-$	MONO	A	
23	*N*-ethylaniline	10.2	41.2	20.1	121.1	122.2 $[M-H]^-$		9.4	445.2	446.2 $[M+H]^+$; 468.2 $[M+Na]^+$	DI	G	
								10.7	283.1	283.8 $[M+H]^+$; 305.8 $[M+Na]^+$	MONO		
24	digitoxigenin	8.9	39.9	17.9	374.2	375.0 $[M+H]^+$		12.2	536.3	537.0 $[M+H]^+$	MONO	A	
								13.5	536.3	537.0 $[M+H]^+$	MONO		
25	3,4-dichloroaniline	9	39	15.3	161.0	162.2 $[M+H]^+$		13.0	323.0	328.8 $[M+H]^+$	MONO	D	0.2 mM final concentration
26	*N*-methoxy-2-naphthalenemethanamine	64.2	37.9	19.5	187.1	210.2 $[M+Na]^+$		11.6	511.2	510.2 $[M+H]^+$	DI	G	
								12.1	511.2	534.2 $[M+Na]^+$	DI		
								13.6	349.1	372.2 $[M+Na]$	MONO		
27	plumbagin	11	35	20.9	188.1	189.0 $[M+H]^+$		11.1	350.1	348.6 $[M-H]^-$; 384.8 $[M+Cl]^-$	MONO	A	
28	acetaminophen	3	35	13.4	151.1	152.1 $[M+H]^+$		11.5	313.2	314.1 $[M+H]^+$	MONO	E	
29	*N*-benzoxybenzhydrylamine	87	32	26.2	289.2	290.0 $[M+H]+$		15.9	613.3	613.8 $[M+H]^+$;635.6 $[M+Na]^+$	DI	D	
								17.2	451.1	452.2 $[M+H]^+$; 474.0 $[M+Na]^+$	MONO		

Table A1.1. Summary of HPLC and LC/MS data for the compound library.

Compound Number	Compound Name	WT % Conversion	'ASP' % Conversion	substrate retention time	calculated substrate mass (*m/z*)	observed substrate mass (*m/z*)		product retention time	expected product mass (*m/z*)	observed product mass (*m/z*)	product designation	HPLC method	comments
30	alazarin	27	31	19.6	240.0	238.6 $[M-H]^-$		13.4	402.1	425.0 $[M+Na]^+$; 400.8 $[M-H]^-$	MONO	A	
31	cannabidiol	49	29	24.8	314.2	415.2 $[M+H]^+$		24.5	476.3	477.3 $[M+H]^+$	MONO	E	0.1 mM final concentration
32	pseudo-aglycon - calicheamicin fragment	43	27	19.4	1050.1	1049.1 $[M-H]^-$		9.9	1212.1	1247.1 $[M+Cl]^-$	MONO	B	0.1 mM final concentration
33	*N*-methylaniline	4.4	27.2	18.1	107.1	n.f.		8.1	431.2	432.2 $[M+H]^+$; 454.0 $[M+Na]^+$	DI	G	0.2 mM final concentration
								9.1	269.1	269.8 $[M+H]^+$; 291.8 $[M+Na]^+$	MONO		
34	aniline	3	27	12.8	93.1	94.2 $[M+H]^+$		8.0	255.1	256.0 $[M+H]^+$	MONO	G	
35	1-pyrenemethanol	2.6	26.7	22.5	232.1	255.0 $[M+Na]^+$		14.0	556.2	591.2 $[M+Cl]^-$	DI	A	
								15.2	556.2	591.2 $[M+Cl]^-$	DI		
								16.0	394.1	429.0 $[M+Cl]^-$	MONO		
36	ampicillin	3	27	9.8	349.1	350.2 $[M+H]^+$		15.4	511.2	511.8 $[M+H]^+$; 533.8 $[M+Na]^+$	MONO	A	
								16.1	511.2	511.8 $[M+H]^+$; 533.8 $[M+Na]^+$	MONO		
37	benzhydrylamine	34	27	19.1	183.1	206.0 $[M+Na]^-$		18.1	345.2	345.8 $[M+H]^-$; 367.6 $[M+Na]^-$	MONO	E	
38	α3I - calicheamicin fragment	21	25	17.4	1210.2	1245.1 $[M+Cl]^-$		15.2	1372.2	1407.1 $[M+Cl]^-$	MONO	B	0.1 mM final concentration
39	fredericamycin A	17	25	23.6		540.0 $[M+H]^+$		10.0	701.2	700.8 $[M-H]^-$	MONO	A	0.25 mM final concentration
40	dynemicin A	45	25	24.9	537.1	537.8 $[M+H]^+$		14.8	699.6	722.2 $[M+Na]^+$	MONO	M	0.1 mM final concentration
41	hydramycin	7	24	19.1	408.1	407.0 [M-H]-		15.7	570.1	569.0 [M-H]	MONO	L	
42	2-phenylethanethiol	31	23	22.8	138.2	172.2 $[M+Cl]^-$		11.8	300.1	335.0 $[M+Cl]^-$	MONO	A	
43	pyrimethamine	36	21	20.2	248.1	249.2 $[M+H]^+$		17.8	410.1	411.2 $[M+H]^+$	MONO	J	
44	7-amino-4-methylcoumarin	1	18	12.3	175.1	176.2 $[M+H]^+$		9.2	337.3	338.0 $[M+H]^+$	MONO	A	0.2 mM final concentration
45	staurosporine aglycon	5	15	19.5	311.1	312.2 $[M+H]^+$		14.6	473.2	474.8 $[M+H]^+$	MONO	A	0.05 mM final concentration
46	benzylalcohol	8	14	11.4	108.1	107.0 $[M-H]^-$		8.75	270.1	305.0 $[M+Cl]^-$	MONO	A	
47	phenylacetic hydrazide	0	13	12.8	150.1	151.2 $[M+H]^+$		15.2	312.1	313.2 $[M+H]^+$	MONO	J	
48	*N*-methoxy-1-naphthalenemethanamine	18	11	19.8	187.1	188.2 $[M+H]^+$		13.5	349.1	350.2 $[M+H]^+$	MONO	H	
49	brefeldin A	1	10	16.0	280.2	281.2 $[M+H]^+$			442.2	477.2 $[M+Cl]^-$	MONO	A	
50	*N*-ethoxybenzhydrylamine	41	10	24.8	227.1	250.0 $[M+Na]^+$		14.2	389.2	412.2 $[M+Na]^+$	MONO	D	
51	benzaldehyde oxime	3	9	14.8	121.1	122.2 $[M+H]^+$		9.9	283.1	284.2 $[M+H]^+$	MONO	A	
52	benzoic hydrazide	3	7	5.7	136.1	137.0 $[M+H]^+$		7.2	298.1	299.2 $[M+H]^+$	MONO	A	
53	sulfanilamide	0	6.7	10.6	172.0	173.1 $[M+H]^+$		3.7	334.1	335.1 $[M+H]^+$	MONO	E	
								5.9	334.1	335.1 $[M+H]^+$; 369.0 $[M+Cl]^-$	MONO		
54	4-biphenylacetic acid	0.8	5.9	20.3	212.1	213.2 $[M+H]^+$		15.9	374.1	397.0 $[M+Na]^+$	MONO	A	
								16.0	374.1	397.0 $[M+Na]^+$	MONO		
								16.2	374.1	397.0 $[M+Na]^+$	MONO		
55	deacetyl-colchicine	21	3	11.0	357.2	358.0 $[M+H]^+$		11.4	519.0	520.0 $[M+H]^+$	MONO		0.5 mM final concentration
56	*N*-methoxybenzhydrylamine	9	2	23.6	213.2	248.6 $[M+Cl]$		13.4	375.2	398.0 $[M+Na]^+$	MONO	D	
57	nystatin A1	41.6	0.0	15.2	925.5	926.0 $[M+H]^+$		12.7	1087.6	1088.4 $[M+H]^+$	MONO	F	0.25 mM final concentration
								13.9	1087.6	1087.8 $[M+H]^+$	MONO		
								14.1	1087.6	1088.2 $[M+H]^+$	MONO		
58	2-naphthoic acid	0.2	4.6	17.4	172.1	173.2 $[M+H]^+$		12.0	334.1	357.2 $[M+Na]^+$	MONO	A	
								12.3	334.1	357.2 $[M+Na]^+$	MONO		
								12.6	334.1	357.2 $[M+Na]^+$	MONO		
59	1-hydroxyanthraquinone	3	4	24.1	224.1	224.8 $[M+H]^+$		13.1	386.1	386.8 $[M+H]^+$	MONO	A	
60	3-iodobenzoic acid	0.6	3.2	18.7	248.0	249.0 $[M+H]^+$		13.4	410.0	445.0 $[M+Cl]^-$	MONO	A	
								13.6	410.0	445.0 $[M+Cl]^-$	MONO		
								13.9	410.0	445.0 $[M+Cl]^-$	MONO		
								15.0	410.0	445.0 $[M+Cl]^-$	MONO		
61	sulfapyridine	1	3	15.4	249.1	250.1 $[M+H]^+$		12.6	411.1	412.2 $[M+H]^+$	MONO	E	
62	daunorubicin	0.4	2.8	14.5	527.2	528.2 [M+H]+		11.1	689.2	690.2 $[M+H]^+$	MONO	A	
								13.3	689.2	690.2 $[M+H]^+$	MONO		
63	trimethoprim	2	2	15.4	290.1	291.3 $[M+H]^+$		12.0	452.2	453.2 $[M+H]^+$	MONO	E	
64	calicheamicin γ	3	1	15.6	1367.3	1402.2 $[M+Cl]^-$		9.9	1529.3	1564.3 $[M+Cl]^-$	MONO	B	0.1 mM final concentration
65	3-chlorobenzoic acid	0.1	1.8	17.2	156.0	157.0 $[M+H]^+$		11.9	318.1	317.0 [M-H]	MONO	A	
								12.4	318.1	317.0 [M-H]	MONO		
								12.8	318.1	317.0 [M-H]	MONO		

Table A1.1 (continued). Summary of HPLC and LC/MS data for the compound library.

Compound Number	Compound Name	WT % Conversion	'ASP' % Conversion	substrate retention time	calculated substrate mass (*m/z*)	observed substrate mass (*m/z*)		product retention time	expected product mass (*m/z*)	observed product mass (*m/z*)	product designation	HPLC method	comments
66	ivermectin	0	2	33.2	874.5	897.2 $[M+Na]^+$		23.4	1036.6	1059.2 $[M+Na]^+$	MONO	K	0.1 mM final concentration
67	phenylacetic acid	0	2	14.1	136.1	137.2 $[M+H]^+$		9.9	298.1	333.2 $[M+Cl]^-$	MONO	A	
68	2-benzoylbenzoic acid	1	1	18.2	226.1	227.2 $[M+H]^+$		13.2	388.1	411.2 $[M+Na]^+$	MONO	A	
								13.4	388.1	411.0 $[M+Na]^+$	MONO		
								13.6	388.1	411.0 $[M+Na]^+$	MONO		
								13.7	388.1	411.2 $[M+Na]^+$	MONO		
								13.9	388.1	411.2 $[M+Na]^+$	MONO		
69	chloramphenicol	1	1	14.4	322.0	321.0 [M-H]-		12.4	484.1	483.0 $[M-H]^-$	MONO	A	
70	staurosporine glucoside	0	1	14.6	473.2	472.2 $[M-H]^-$		7.6	635.2	670.0 $[M+Cl]^-$	MONO	A	0.05 mM final concentration
71	avermectin	0	1	26.3	872.5	895.2 $[M+Na]^+$		18.2	1034.5	1058.0 $[M+Na]^+$	MONO	K	0.1 mM final concentration
72	indole	0	0.4	18.4	117.1	118.2 $[M+H]^+$		10.6	279.1	280.0 $[M+H]^+$	MONO	A	
73	ivermectin psuedo-aglycon	0	0.2	27.8	730.4	753.0 $[M+Na]^+$		18.65	892.5	915.6 $[M+Na]^+$	MONO	K	0.01 mM final concentration

Table A1.1 (continued). Summary of HPLC and LC/MS data for the compound library. [a] % conversion is calculated as the sum of all observed products; **[b]** MONO – mass corresponds to monosaccharide product; DI – mass corresponds to disaccharide product; **[c]** refer to **Table A1.2**; **[d]** n.f. – not found.

Method	Solvent A[a]	Solvent B	Gradient	Detection (nm)
A	0.01% TFA	CH_3CN	10-75% B, 20 min; 75-100% B, 1min; 100% B, 5 min	254
B	0.01% TFA	CH_3CN	10-100% B, 20 min	280
C	50 mM Tris, pH 8.0	CH_3CN	5% B, 2.5 min; 5-60% B, 18.5 min; 60-100% B, 1 min; 100% B, 6 min	254
D	ddH_2O	CH_3CN	10-75% B, 20 min; 75-100% B, 1min; 100% B, 14 min	254
E	ddH_2O	CH_3CN	2% B, 2.5 min; 2-60% B, 17.5 min; 60-100% B, 1 min; 100% B, 4 min	254
F	0.01% TFA	CH_3CN	10-75% B, 20 min; 75-100% B, 1min; 100% B, 5 min	305
G	10 mM NH_4CO_3, pH 10.0	CH_3CN	10-75% B, 20 min; 75-100% B, 1min; 100% B, 5 min	254
H	ddH_2O	CH_3CN	10-75% B, 20 min; 75-100% B, 1min; 100% B, 5 min	254
J	ddH_2O	CH_3CN	5% B, 5 min; 5-75% B, 15 min; 75-100% B, 1 min; 100% B, 5 min	254
K	ddH_2O	CH_3CN	30-70% B, 5 min; 70-100% B, 30 min; 100% B, 3 min	243
L	0.01% TFA	CH_3CN	10-100% B, 15 min; 100% B, 20 min	254
M	0.01% TFA	CH_3CN	10-75% B, 20 min; 75-100% B, 1min; 100% B, 5 min	320

Table A1.2. HPLC conditions for the compound library. [a] 0.1% formic acid in ddH_2O was utilized for LC/ESI-MS methods when solvent A was ddH_2O or 0.1% trifluoroacetic acid in ddH_2O. 5 mM ammonium bicarbonate (pH 8.0) was utilized for LC/ESI-MS methods when solvent A was 50 mM Tris-HCl (pH 8.0) in ddH_2O or 10 mM ammonium bicarbonate (pH 10.0) in ddH_2O.

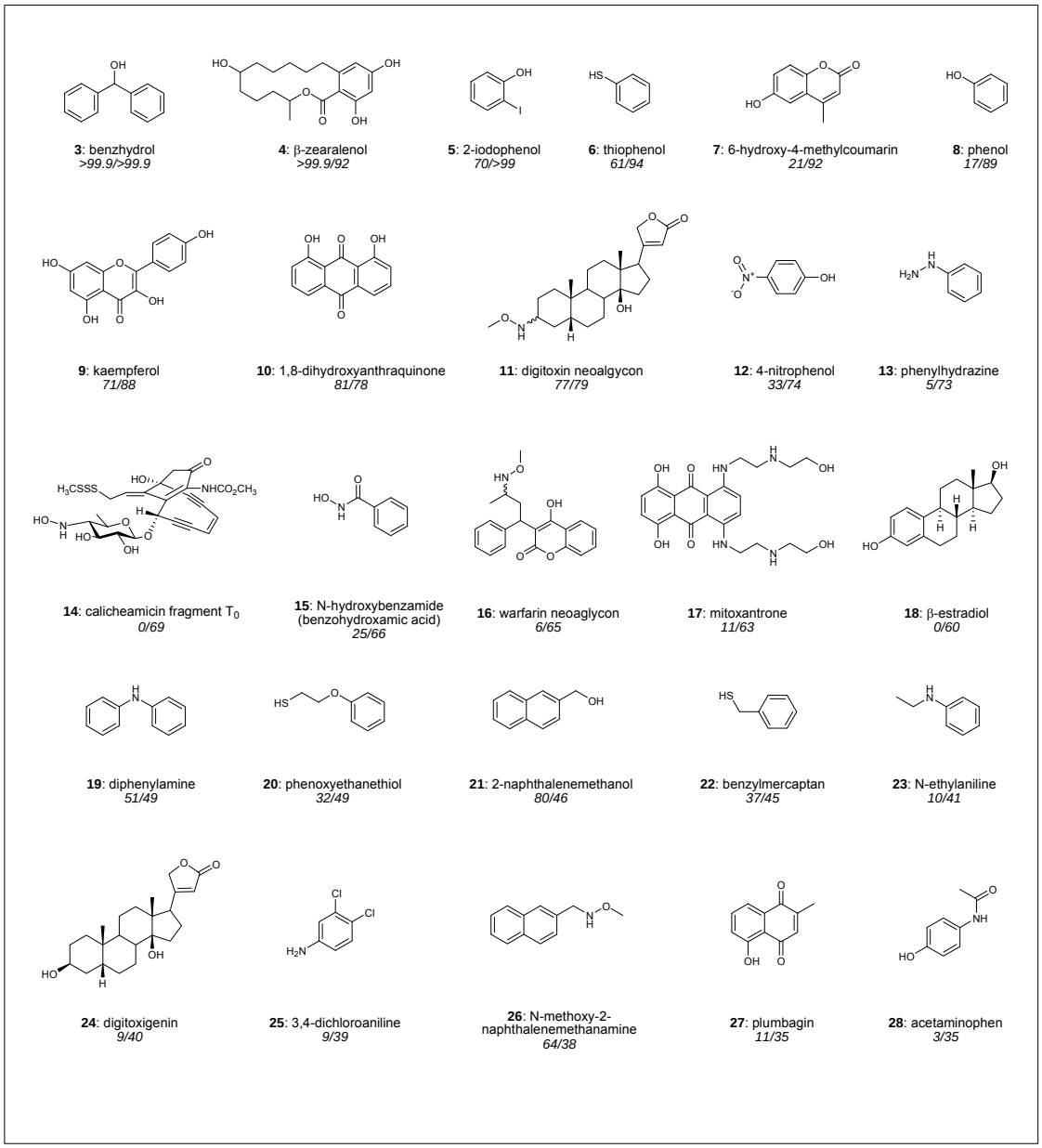
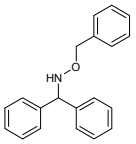
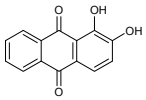


Figure A1.1. Structures of aglycon library members. Compounds that resulted in observable glucosylation (**3-73**) are listed in descending order of observed 'ASP' conversion. Compounds with no observable glucosylation (**74-139**) are listed in alphabetical order. Italic numbers noted under compound number and name denote observed percent conversion for *WT/'ASP'* (continued on subsequent pages).

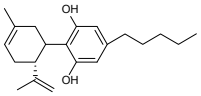
Structures of All Library members (con't)



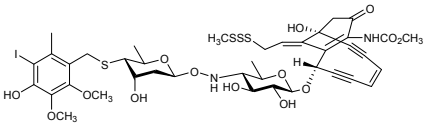
29: N-benzoxybenzhydrylamine
87/32



30: alizarin
27/31



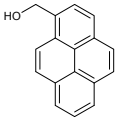
31: cannabidiol
49/29



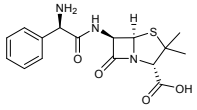
32: calicheamicin fragment psuedo-aglycon
43/27

33: N-methylaniline
4/27

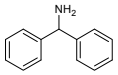
34:aniline
3/27



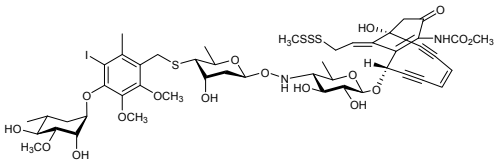
35: 1-pyrenemethanol
3/27



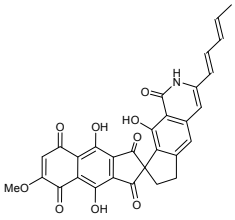
36: ampicillin
3/27



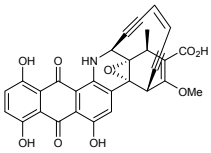
37: benzhydrylamine
(aminodiphenylmethane)
34/27



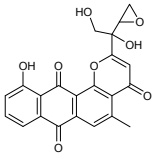
38: calicheamicin fragment α3I
21/25



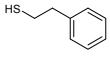
39: fredericamycin A
17/25



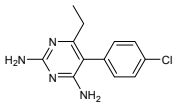
40: dynemicin A
24/25



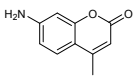
41: hydramycin
7/24



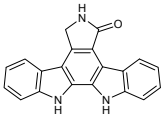
42: 2-phenylethanethiol
31/23



43: pyrimethamine
36/21

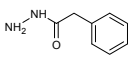


44: 7-amino-4-methylcoumarin
1/18

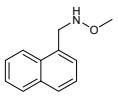


45: staurosporine aglycon
5/15

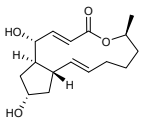
46: benzylalcohol
8/14



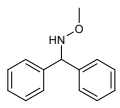
47: phenylacetic hydrazide
0/13



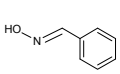
48: N-methoxy-1-naphthalenemethanamine
18/11



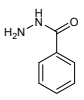
49: brefeldin A
1/10



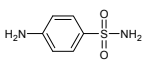
50: N-methoxybenzhydrylamine
9/2



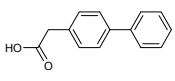
51: benzaldehyde oxime
3/9



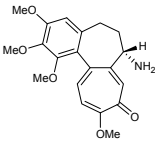
52: benzoic hydrazide
3/7



53: sulfanilamide
0/7

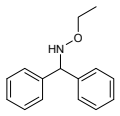


54: 4-biphenylacetic acid
1/6

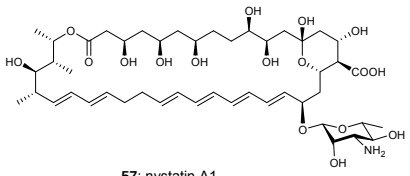


55: deacetyl-colchicine
21/3

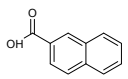
Structures of All Library members (con't)



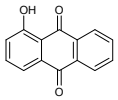
56: N-ethoxydiphenylmethylamine
41/10



57: nystatin A1
42/0

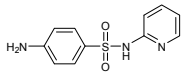


58: 2-naphthoic acid
0.2/5

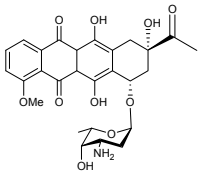


59: 1-hydroxyanthraquinone
3/4

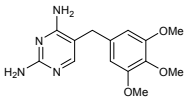
60: 3-iodobenzoic acid
1/3



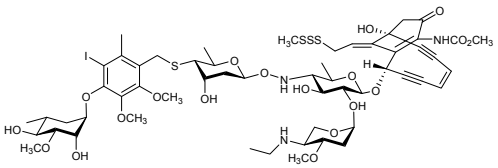
61: sulfapyridine
1/3



62: daunorubicin
0/3

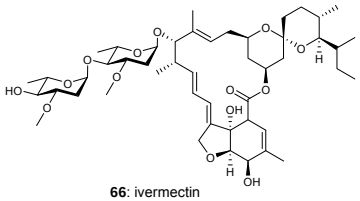


63: trimethoprim
2/2

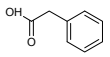


64: calicheamicin γ
3/1

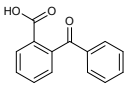
65: 3-chlorobenzoic acid
0/2



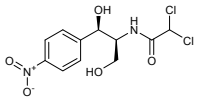
66: ivermectin
0/2



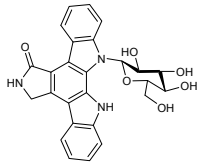
67: phenylacetic acid
0/2



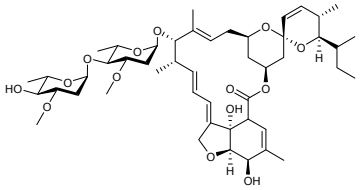
68: 2-benzoylbenzoic acid
1/1



69: chloramphenicol
1/1

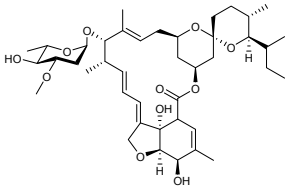


70: staurosporine glucoside
0.1/1

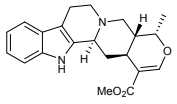


71: avermectin
0/1

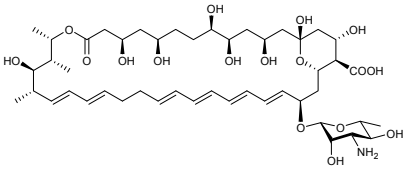
72: indole
0/0.4



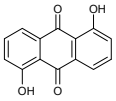
73: ivermectin psuedo-aglycon
0/0.2



74: ajmalicine
0/0



75: amphotericin B
0/0

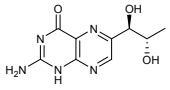


76: anthrarufin
0/0

77: benzaldehyde
0/0

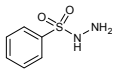
78: benzoic acid
0/0

Structures of All Library members (con't)

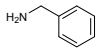


79: 6-biopterin
0/0

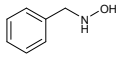
80: benzenesulfonamide
0/0



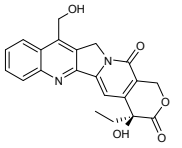
81: benzenesulfonyl hydrazide
0/0



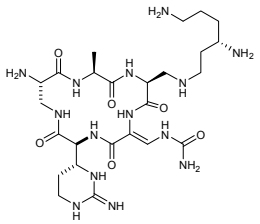
82: benzylamine
0/0



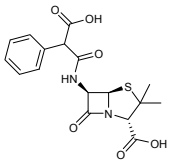
83: O-benzylhydroxylamine
0/0



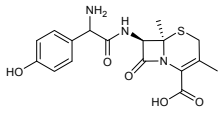
84: camptothecin
0/0



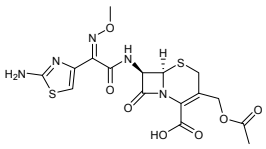
85: capreomycin 1B
0/0



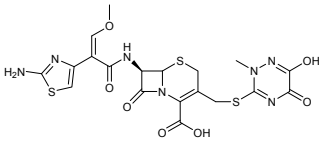
86: carbenicillin
0/0



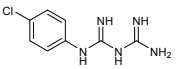
87: cefadroxil
0/0



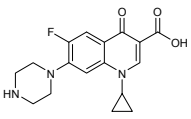
88: cefotaxime
0/0



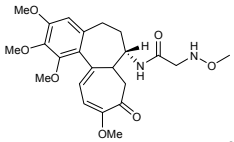
89: ceftriaxone
0/0



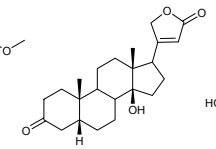
90: 1-(4-chlorophenyl)biguanide
0/0



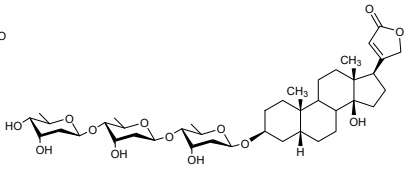
91: ciprofloxacin
0/0



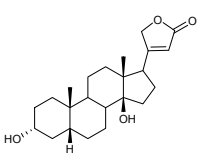
92: colchicine neoaglycon
0/0



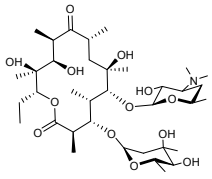
93: digitoxone
0/0



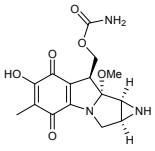
94: digitoxin (digoxin)
0/0



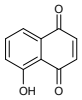
95: epi-digitoxigenin
0/0



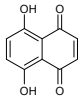
96: erythromycin
0/0



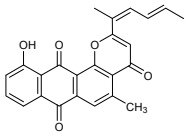
97: hydroxymitomycin C
0/0



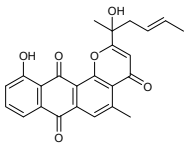
98: 5-hydroxy-1,4-naphthaquinone
0/0



99: 5,8-hydroxy-1,4-naphthaquinone
0/0



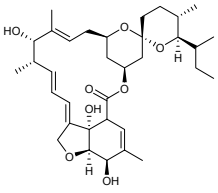
100: α-indomycinone
0/0



101: β-indomycinone
0/0

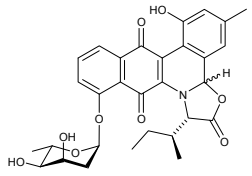
102: 2-iodobenzoic acid
0/0

103: isoniazid
0/0

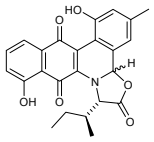


104: ivermectin aglycon
0/0

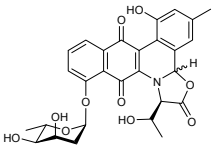
Structures of All Library members (con't)



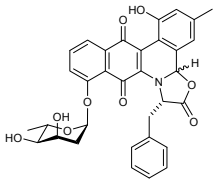
105: jadomycin B
0/0



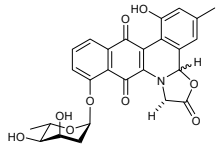
106: jadomycin B aglycone
0/0



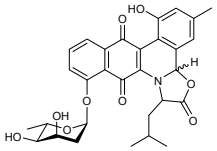
107: jadomycin DT
0/0



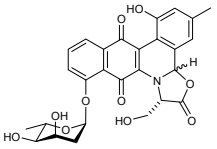
108: jadomycin F
0/0



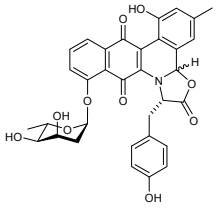
109: jadomycin G
0/0



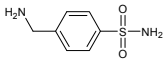
110: jadomycin L
0/0



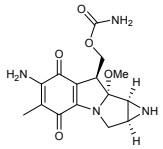
111: jadomycin S
0/0



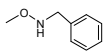
112: jadomycin Y
0/0



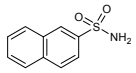
113: mafenide
0/0



114: mitomycin C
0/0

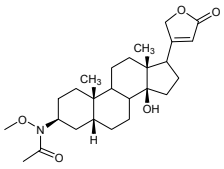


115: *N*-methoxybenzylamine
0/0

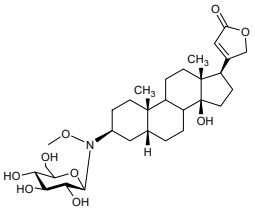


116: naphthalene-2-sulfonamide
0/0

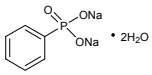
117: 1,8-naphthalimide
0/0



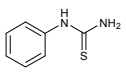
118: neodigitoxin 'E3'
0/0



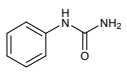
119: neodigitoxin glucoside
0/0



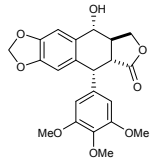
120: sodium phenyl phosphate dibasic dihydrate
0/0



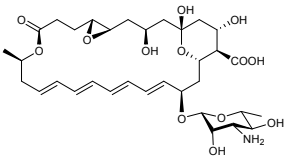
121: *N*-phenylthiourea
0/0



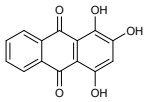
122: *N*-phenylurea
0/0



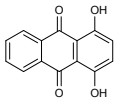
123: podophyllotoxin
0/0



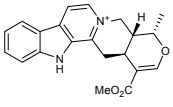
124: pimaricin
0/0



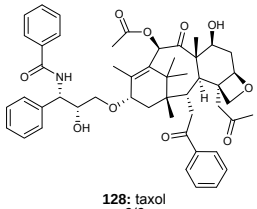
125: purpurin
0/0



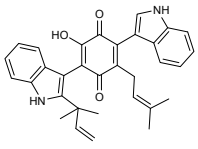
126: quinizarin
0/0



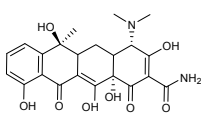
127: serpentine
0/0



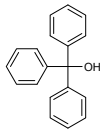
128: taxol
0/0



129: terrequinone A
0/0

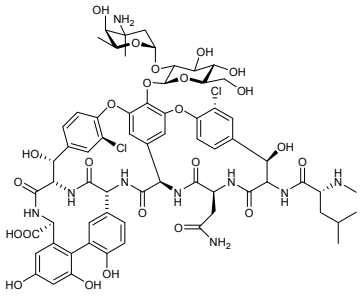


130: tetracycline
0/0

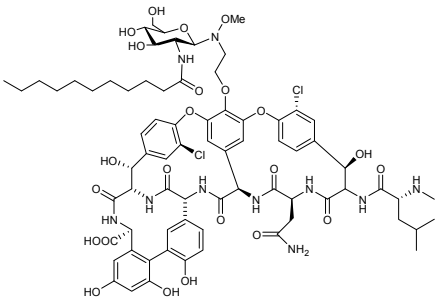


131: triphenolmethanol
0/0

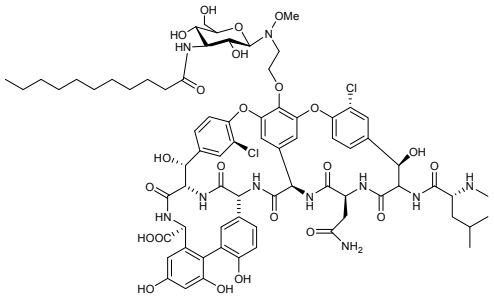
Structures of All Library members (con't)



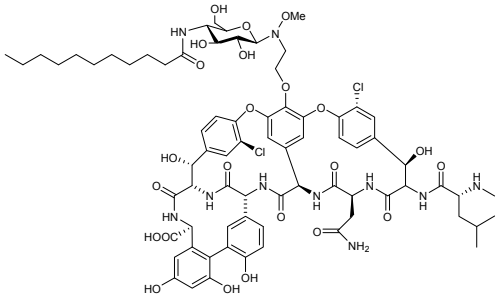
132: vancomycin
0/0



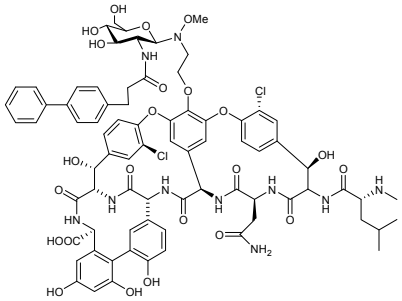
133: vancomycin 1-1
0/0



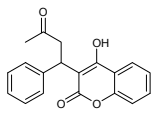
134: vancomycin 1-2
0/0



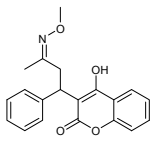
135: vancomycin 1-4
0/0



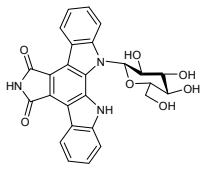
134: vancomycin 1-5
0/0



135: warfarin
0/0



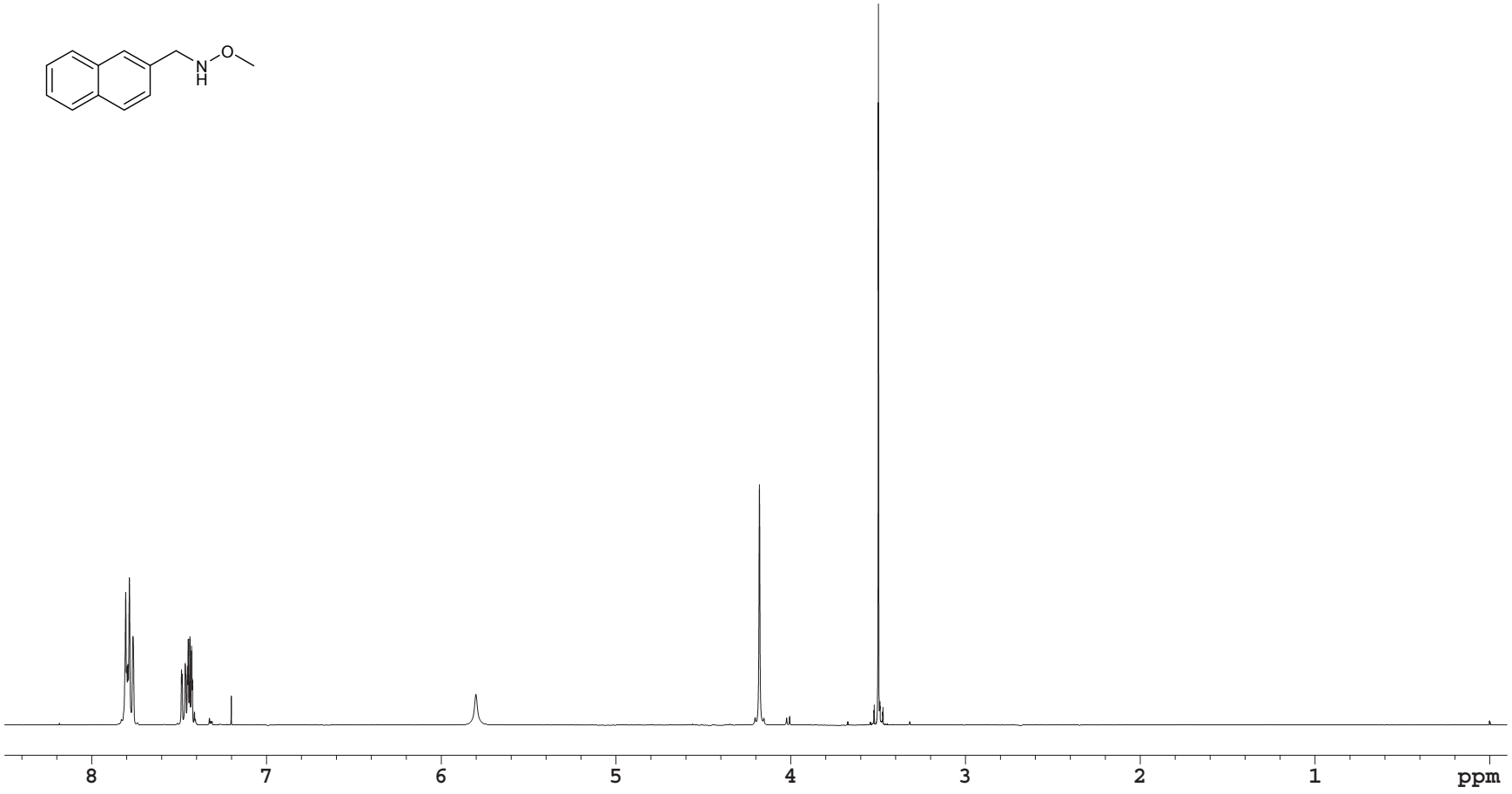
136: warfarin oxime
0/0



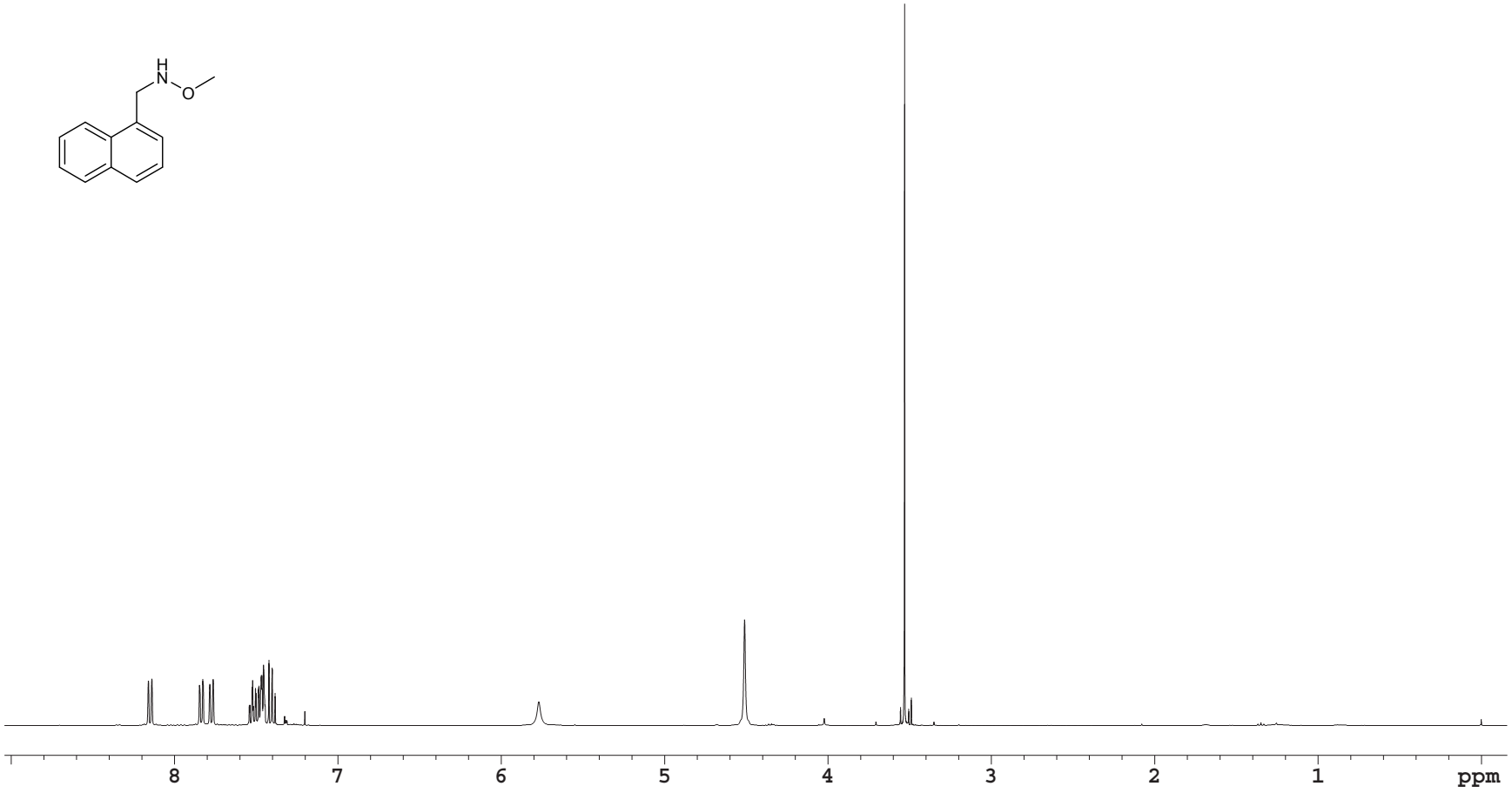
137: '131'
0/0

A1.2. ^{1}H NMR Spectra

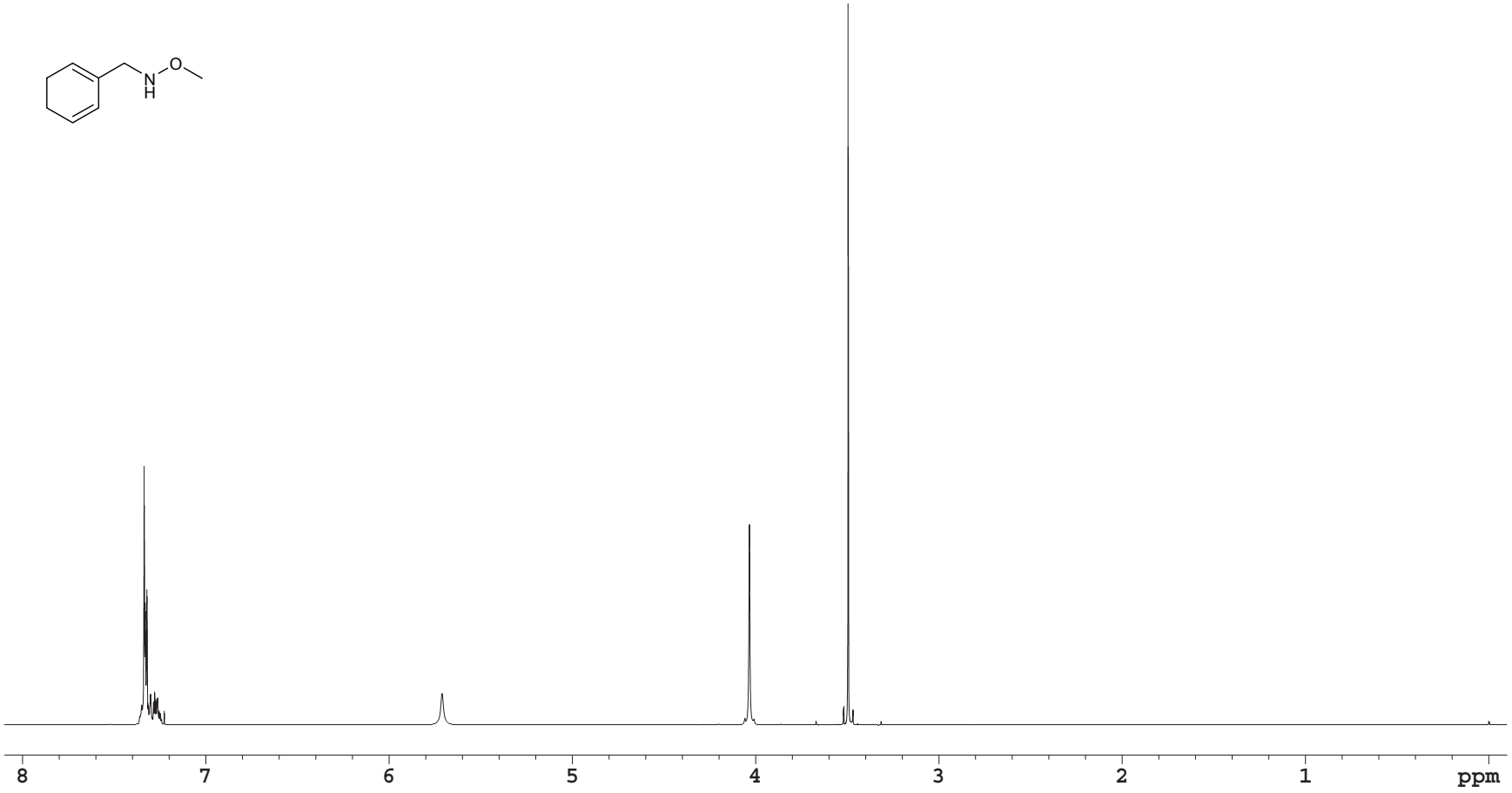
N-methoxy-2-naphthalenemethanamine (**26**): ^{1}H NMR (400 MHz, $CDCl_3$)



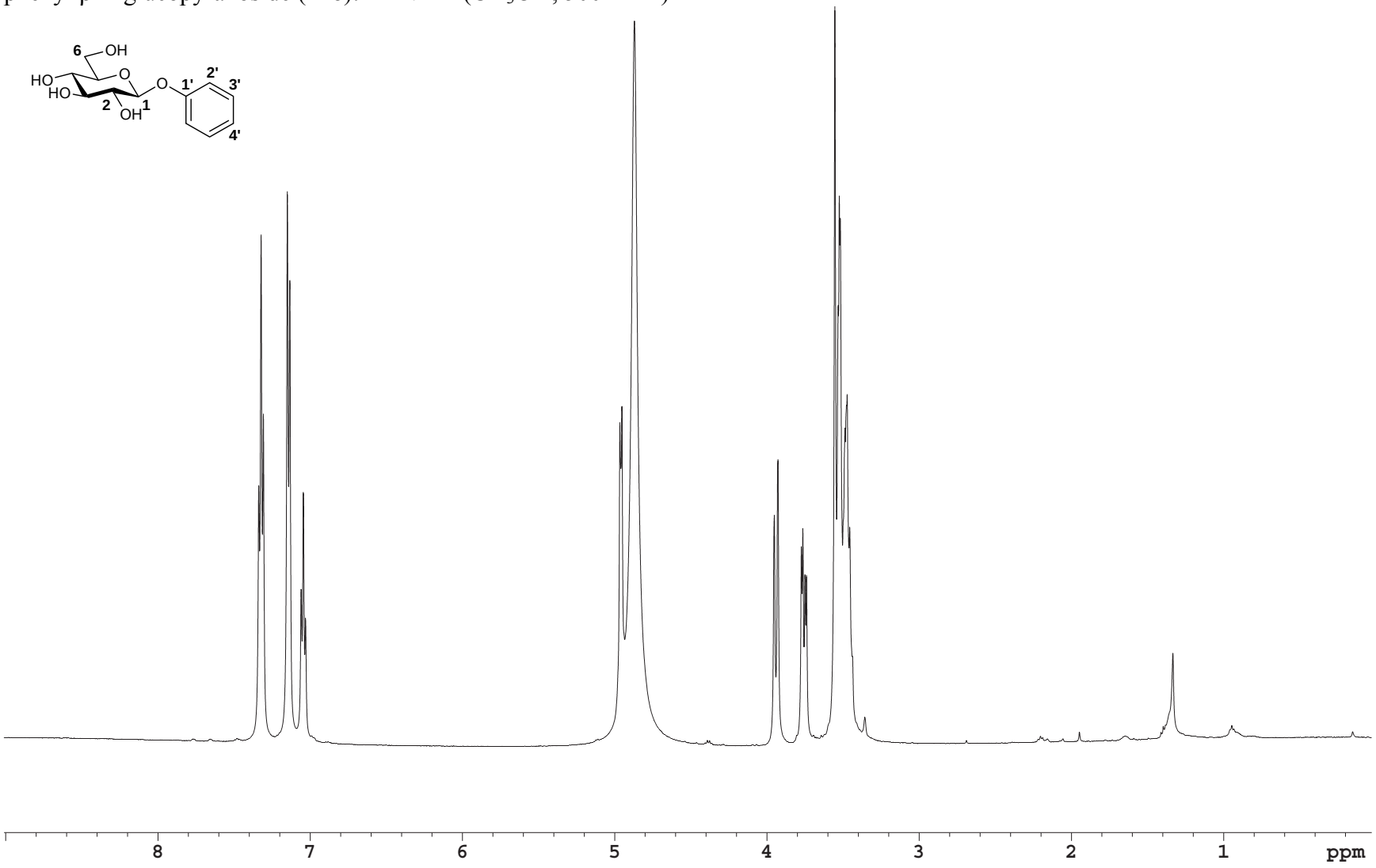
N-methoxy-1-naphthalenemethanamine (**48**): ^{1}H NMR (400 MHz, $CDCl_3$)



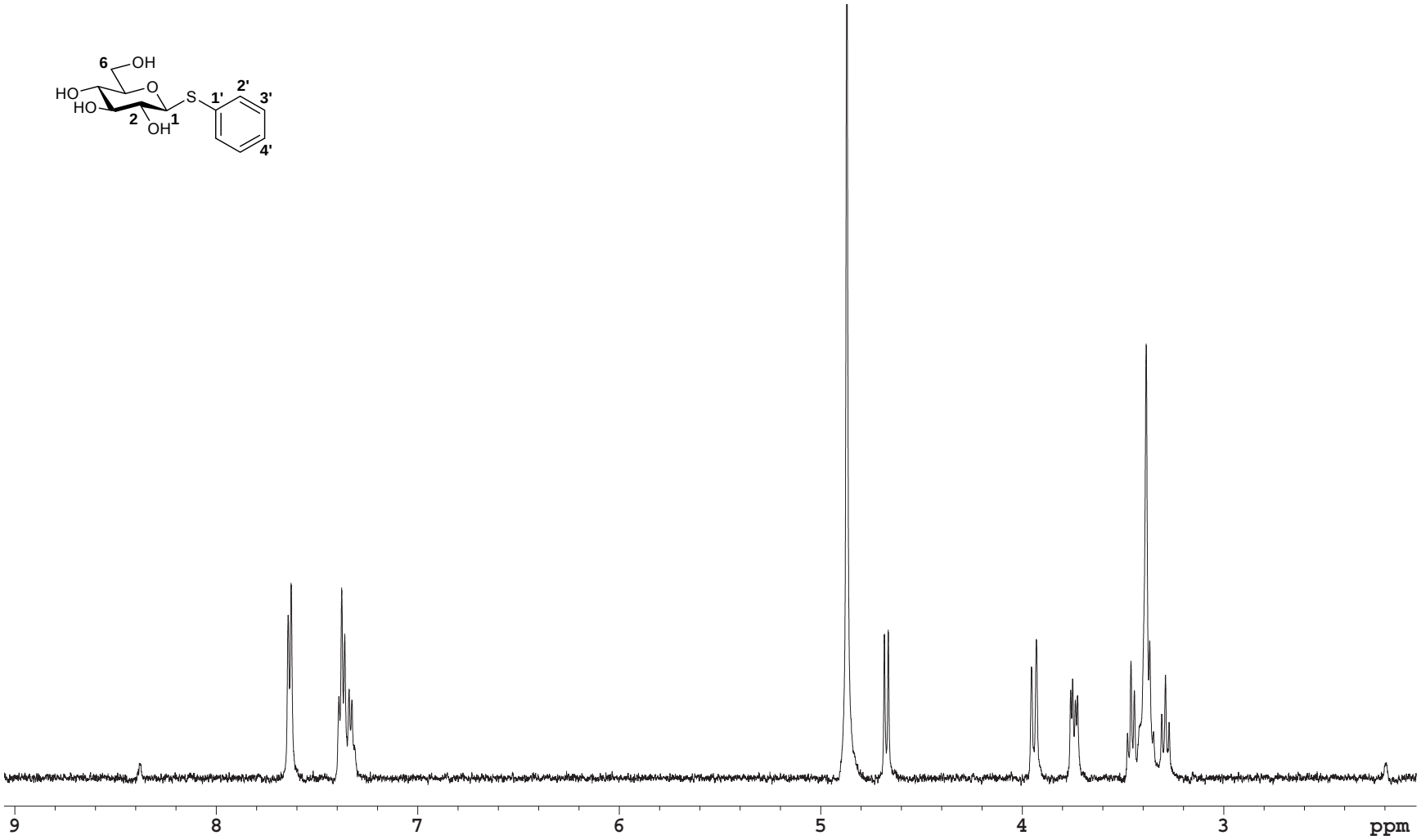
N-methoxybenzylamine (**115**): ^{1}H NMR (400 MHz, $CDCl_3$)



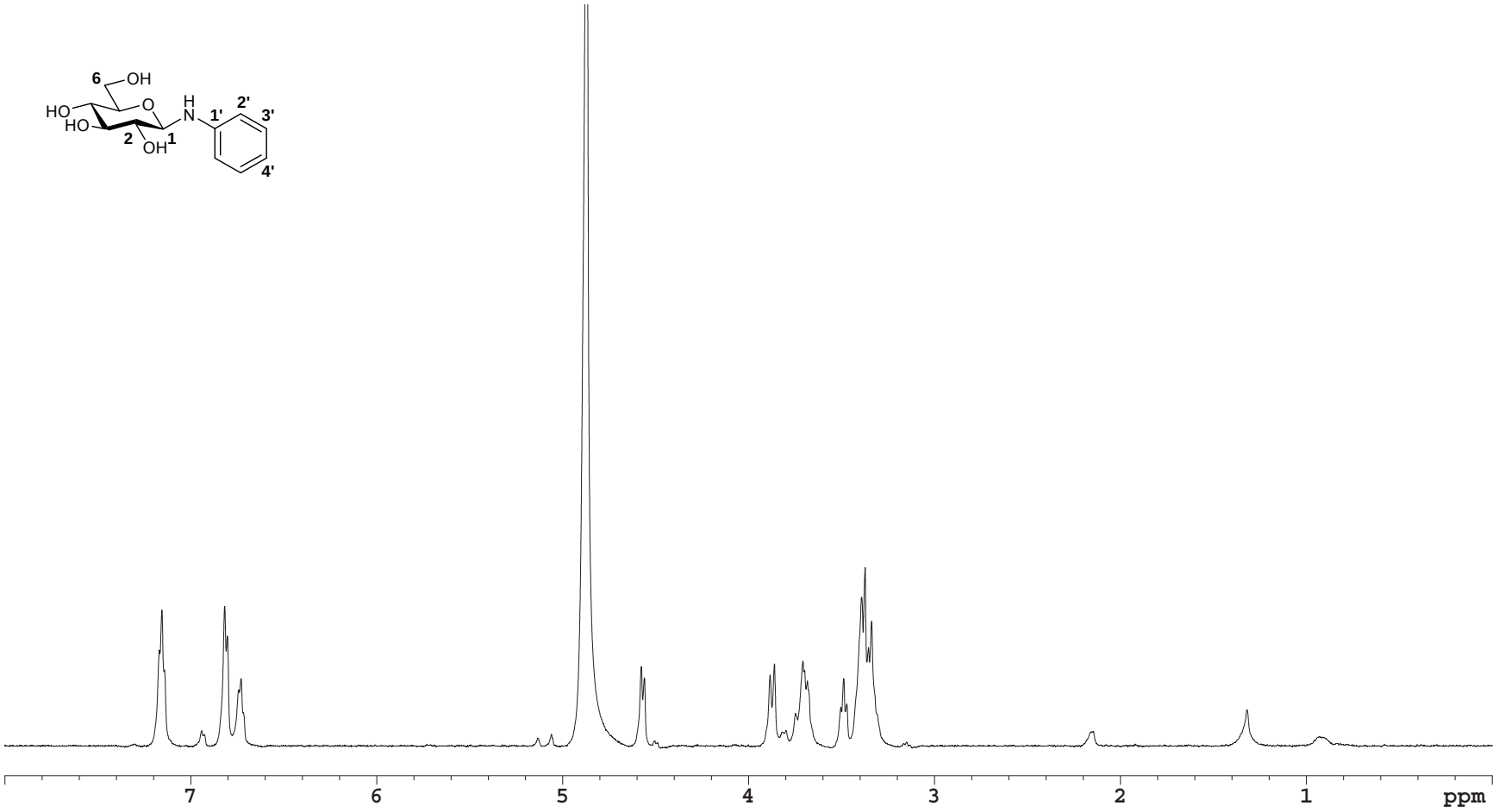
phenyl β-D-glucopyranoside (**140**): ^{1}H NMR (CD_3OD, 500 MHz)



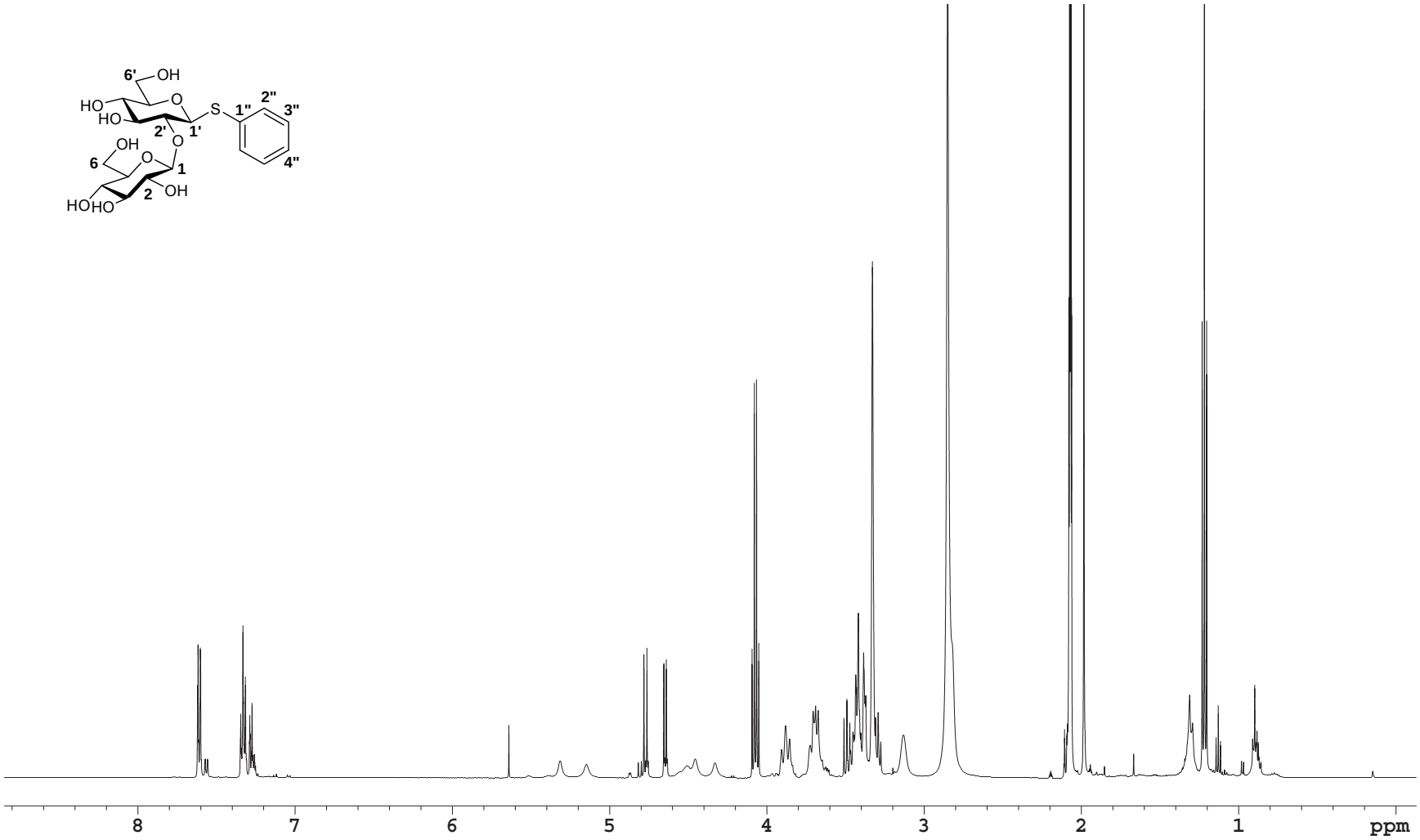
phenyl 1-thio-β-D-glucopyranoside (**141**): ^{1}H NMR (CD_3OD, 500 MHz)



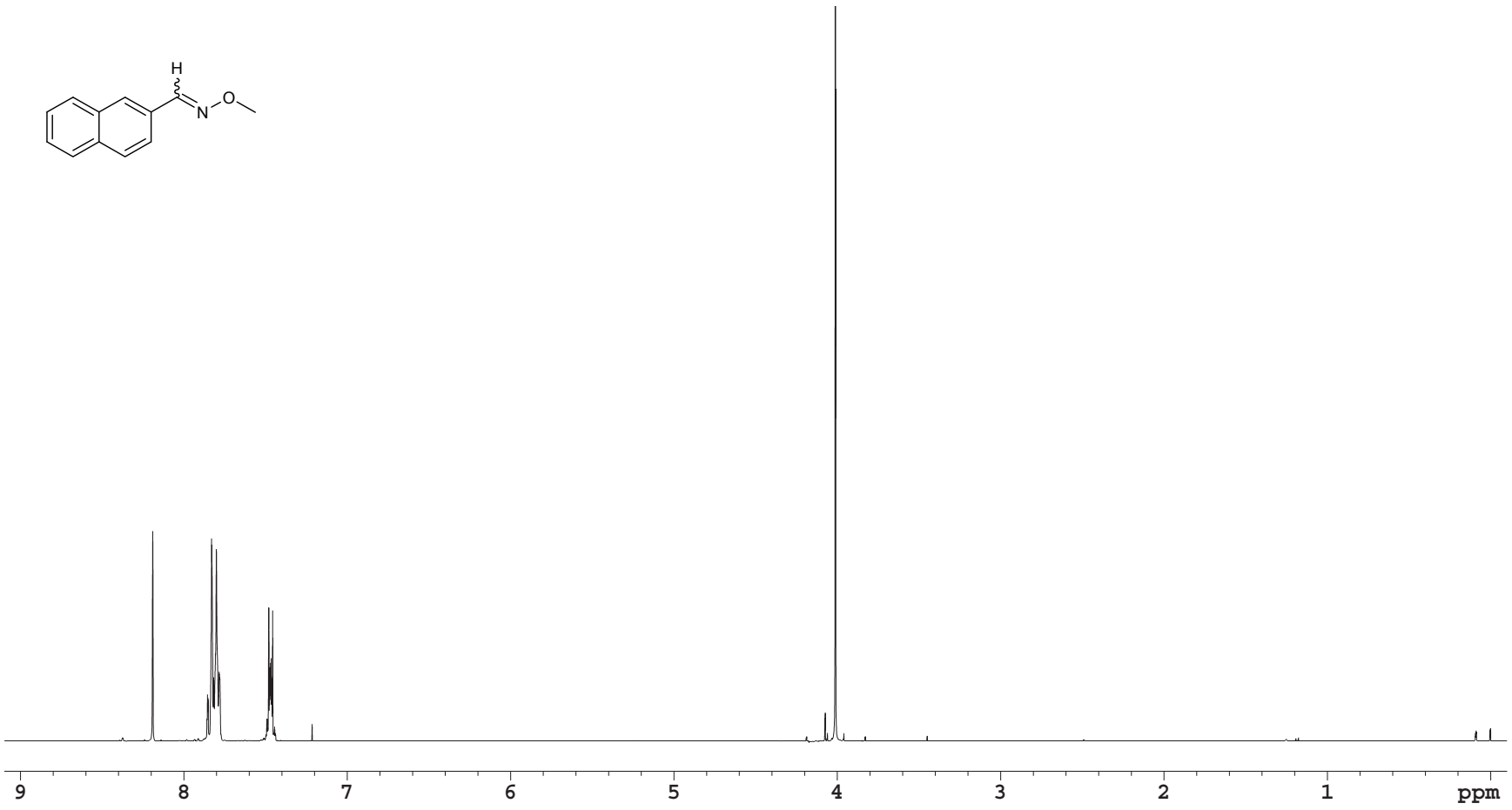
phenyl 1-amino-β-D-glucopyranoside (**142**): ^{1}H NMR (CD_3OD, 500 MHz)



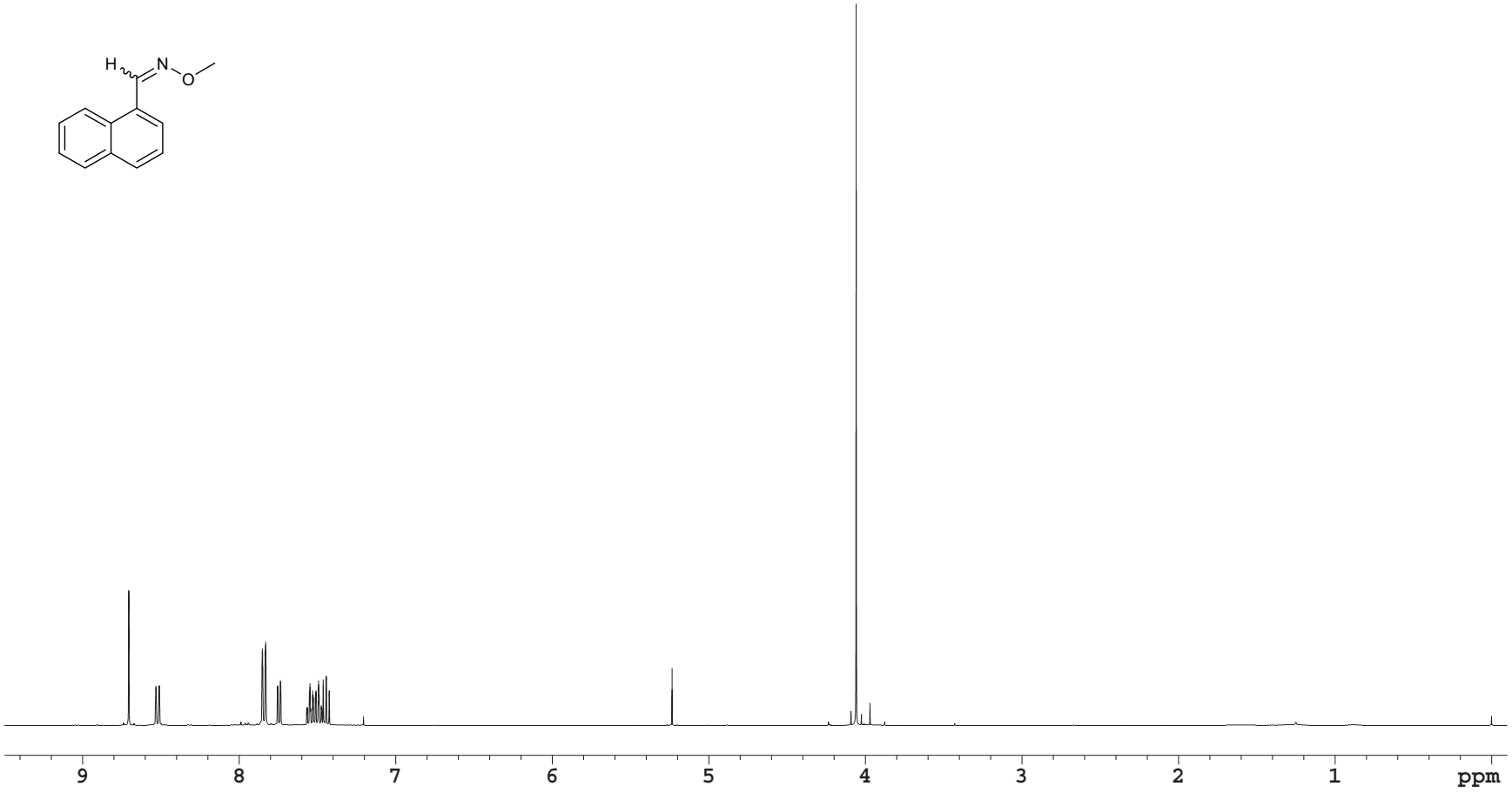
phenyl β-D-glucopyranose-(1→2)-1-thio-β-d-glucopyranoside (**143**): ^{1}H NMR (acetone-d_6, 500 MHz)



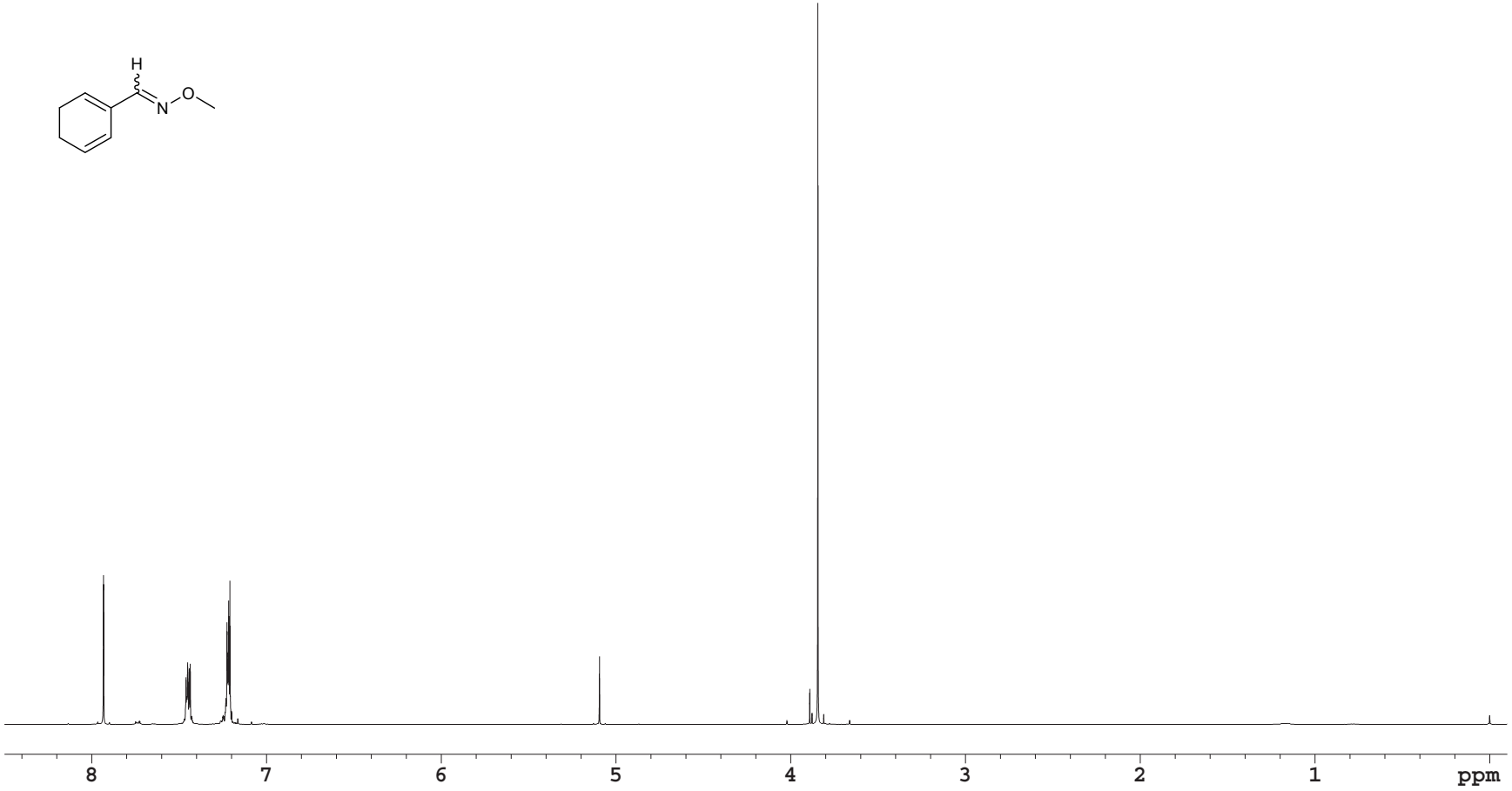
2-naphthaldehyde-*O*-methyloxime (**145**): ^{1}H NMR (400 MHz, $CDCl_3$)



1-naphthaldehyde-*O*-methyloxime (**148**): ^{1}H NMR (400 MHz, $CDCl_3$)

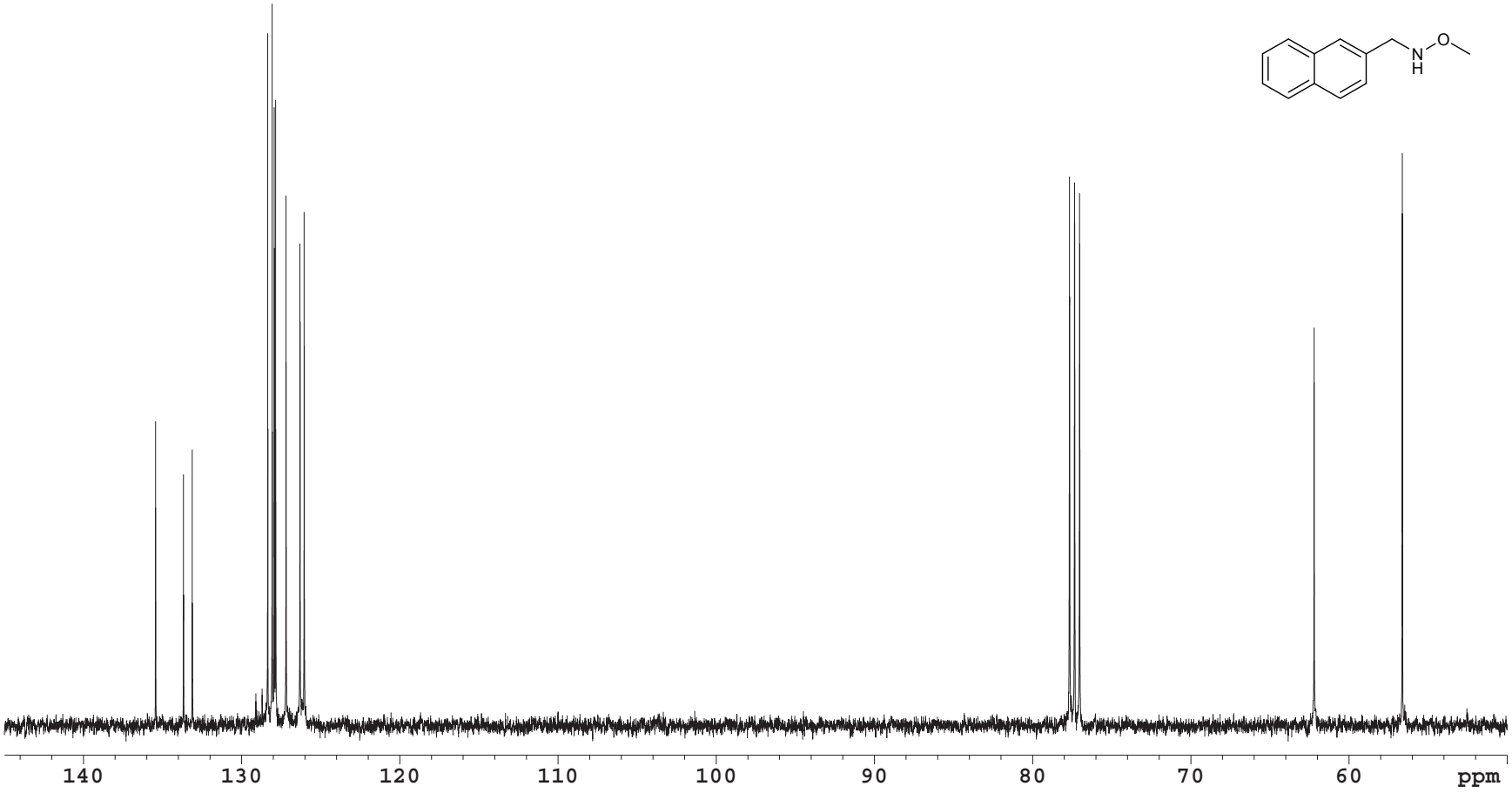


benzaldehyde-*O*-ethyloxime (**150**): ^{1}H NMR (400 MHz, $CDCl_3$)

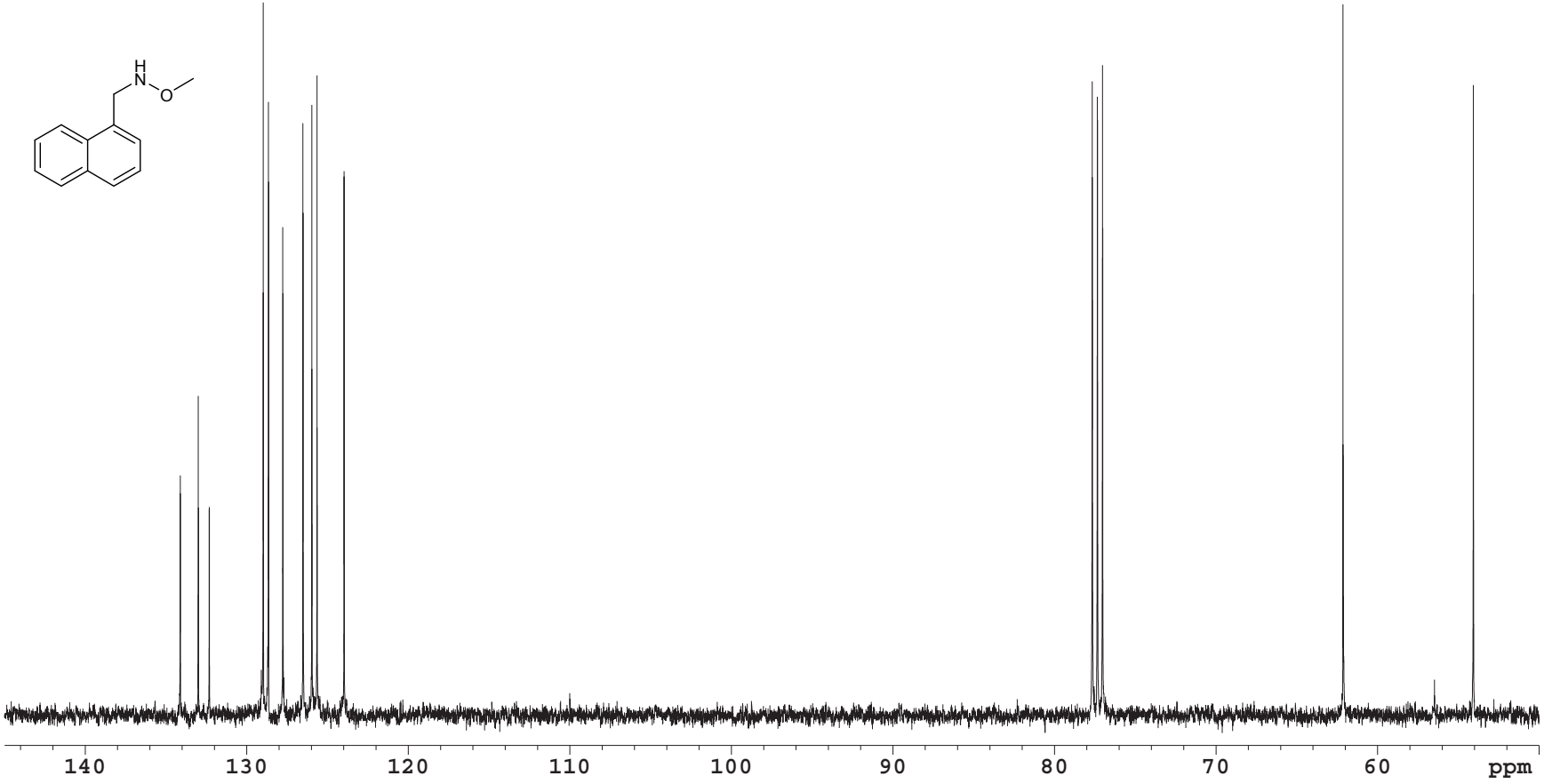


A1.2 ^{13}C NMR Spectra

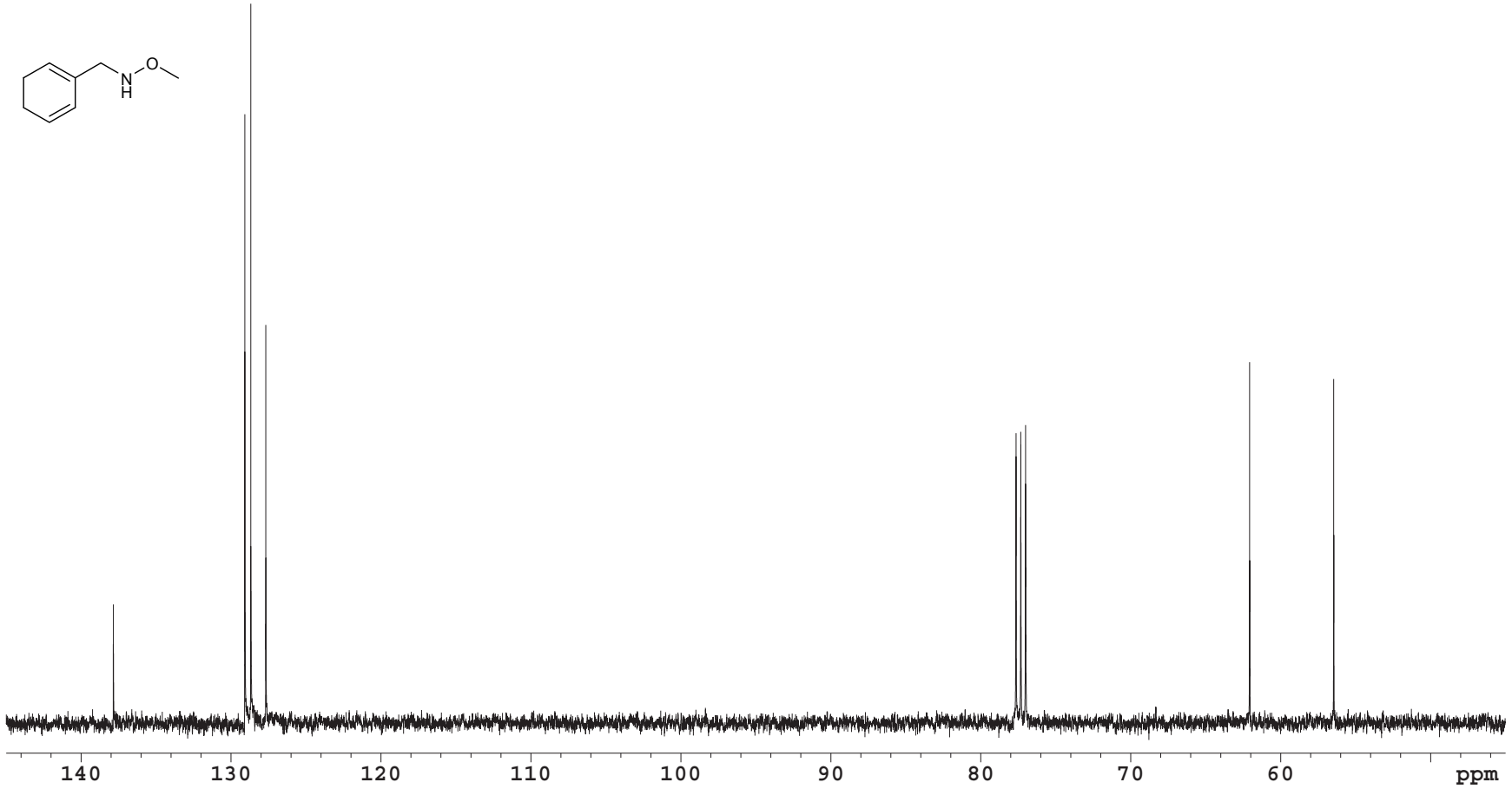
N-methoxy-2-naphthalenemethanamine (**26**): ^{13}C NMR (100 MHz, $CDCl_3$)



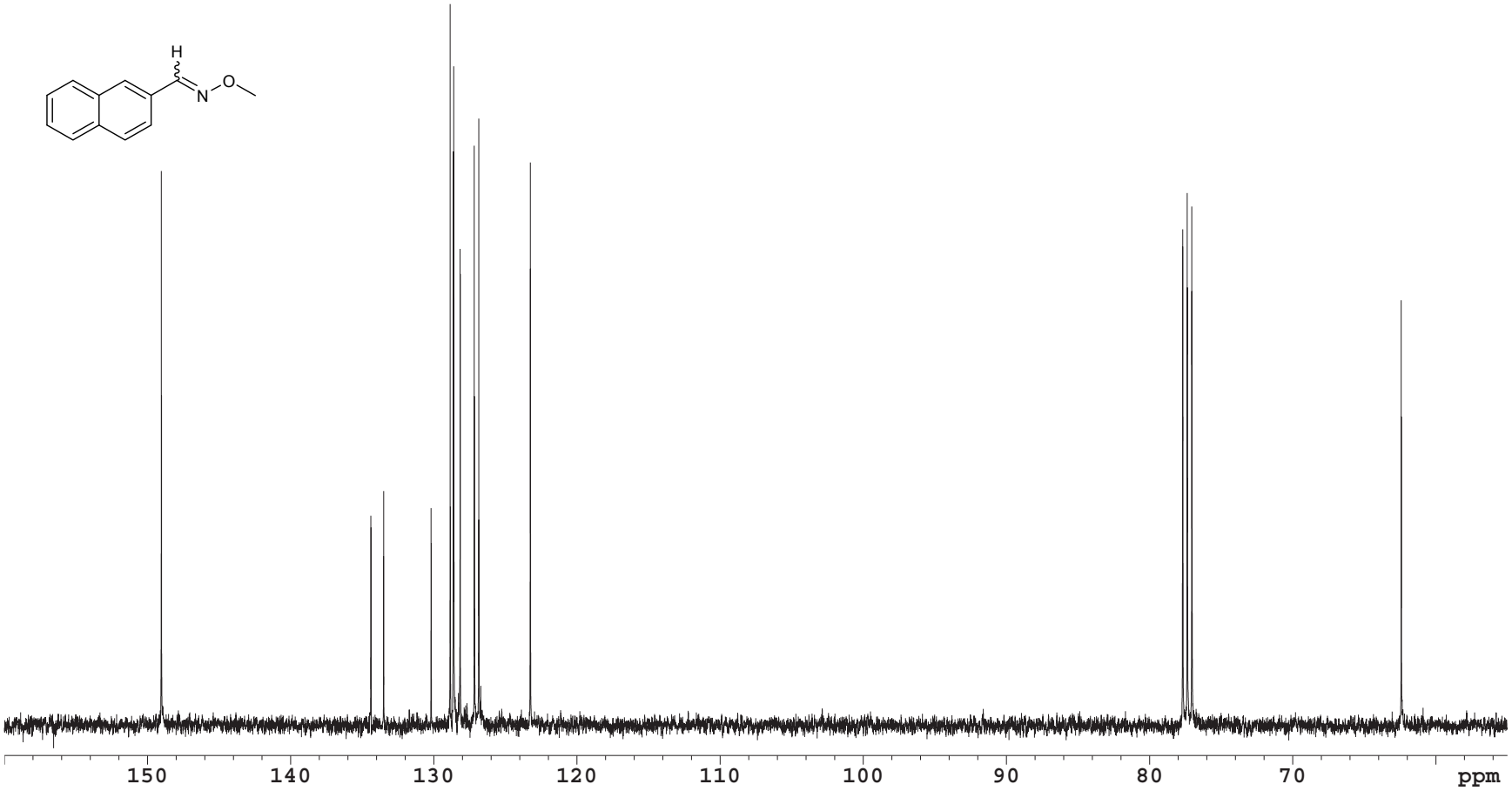
N-methoxy-1-naphthalenemethanamine (**48**): ^{13}C NMR (100 MHz, $CDCl_3$)



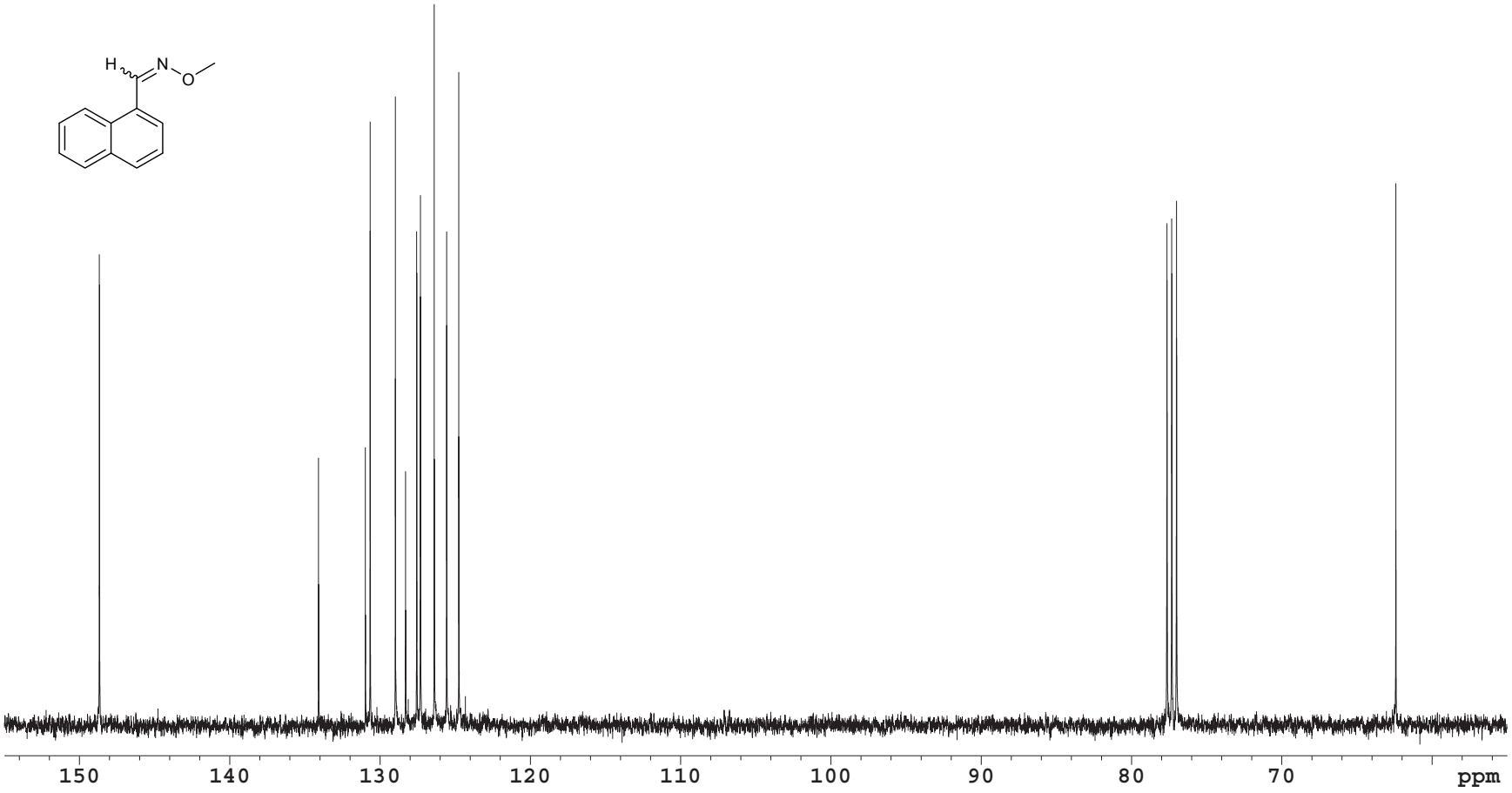
N-methoxybenzylamine (**115**): ^{13}C NMR (100 MHz, $CDCl_3$)



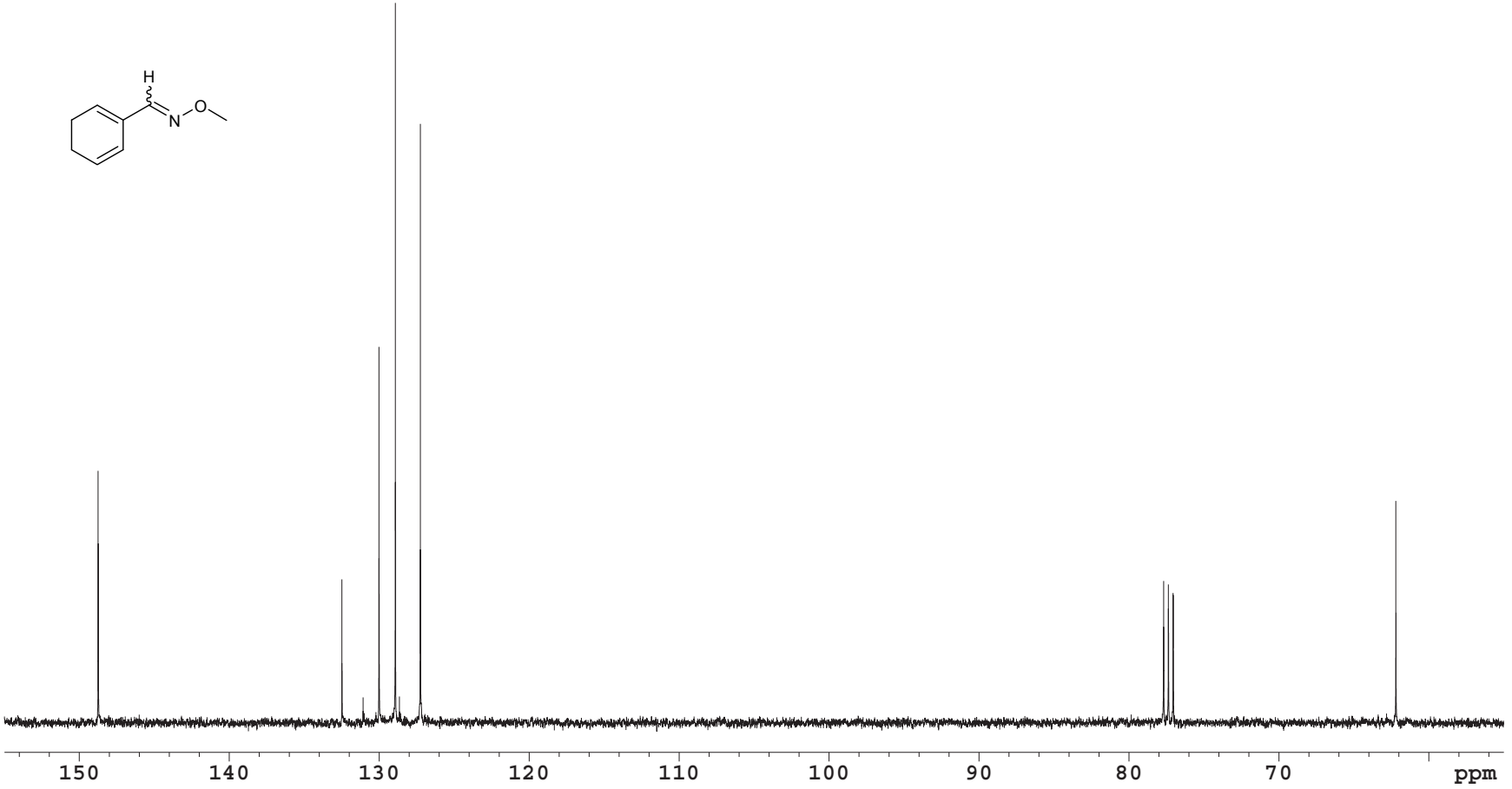
2-naphthaldehyde-*O*-methyloxime (**145**): ^{13}C NMR (100 MHz, $CDCl_3$)



1-naphthaldehyde-*O*-methyloxime (**148**): ^{13}C NMR (100 MHz, $CDCl_3$)



benzaldehyde-*O*-ethyloxime (**150**): ^{13}C NMR (100 MHz, $CDCl_3$)



Appendix 2:

Supplementary Data for Chapter 3

A2.1. Supporting Data for 50 Member Acceptor Library Screen

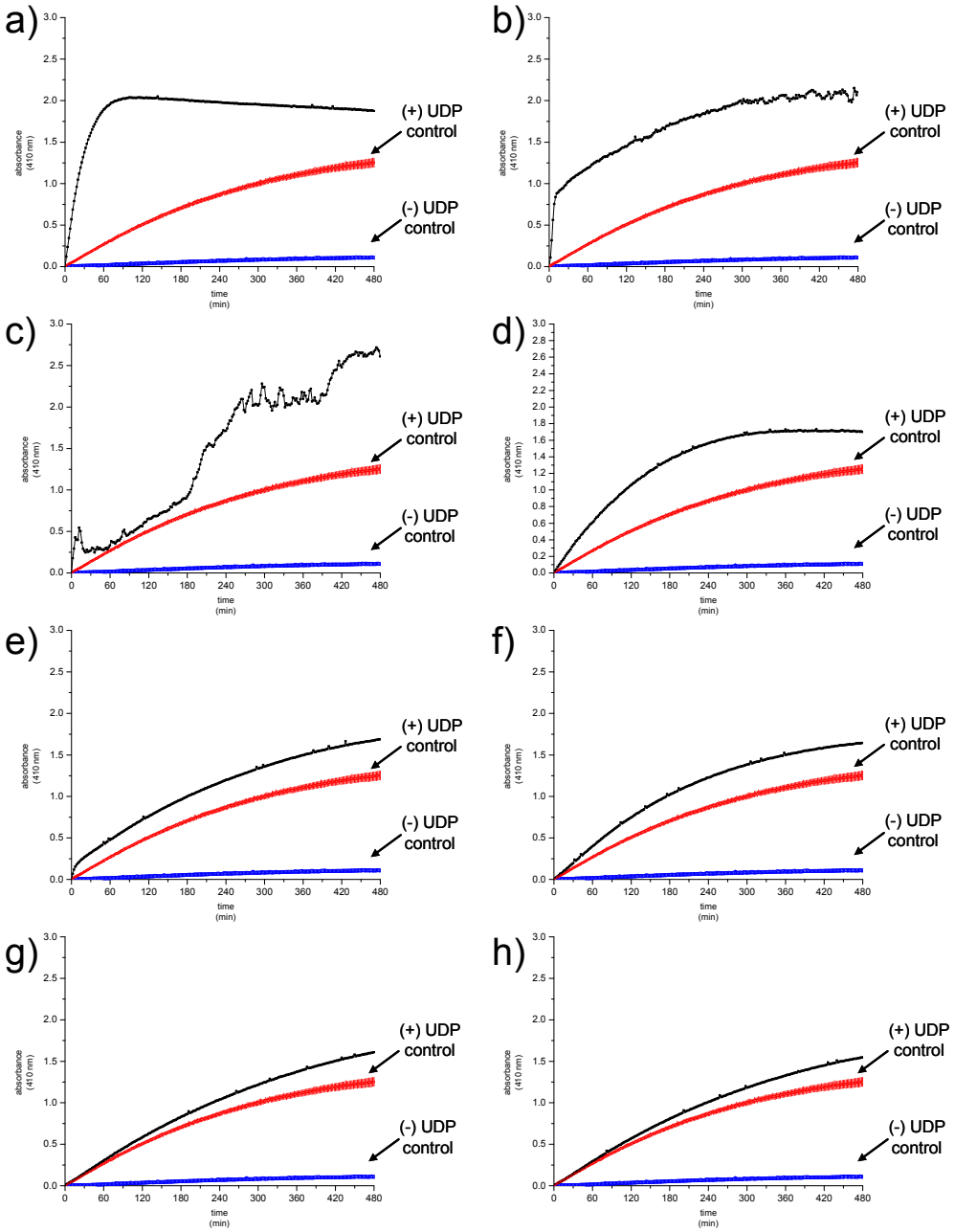


Figure A2.1. Representative drug screening data at 410 nm. a) 71 (2-methoxyestradiol), **b) 73** (rapamycin), **c) 78** (nelfinavir), **d) 80** (flavopiridol), **e) 85** (triclostatin A), **f) 86** ((-)-epinephrine), **g) 96** (pemetrexed), and **h) 99** (gemcitabine). Experimental data are colored black. Mean data of control reactions lacking final acceptor ((+) UDP control) or final acceptor and UDP ((-) UDP control) are labeled and colored red and blue, respectively, with standard deviation noted. **78** (nelfinavir) was insoluble at the concentration evaluated (0.5 mM), causing inconsistency with the absorbance reading. See **Figure A2.2** for product confirmation by HPLC.

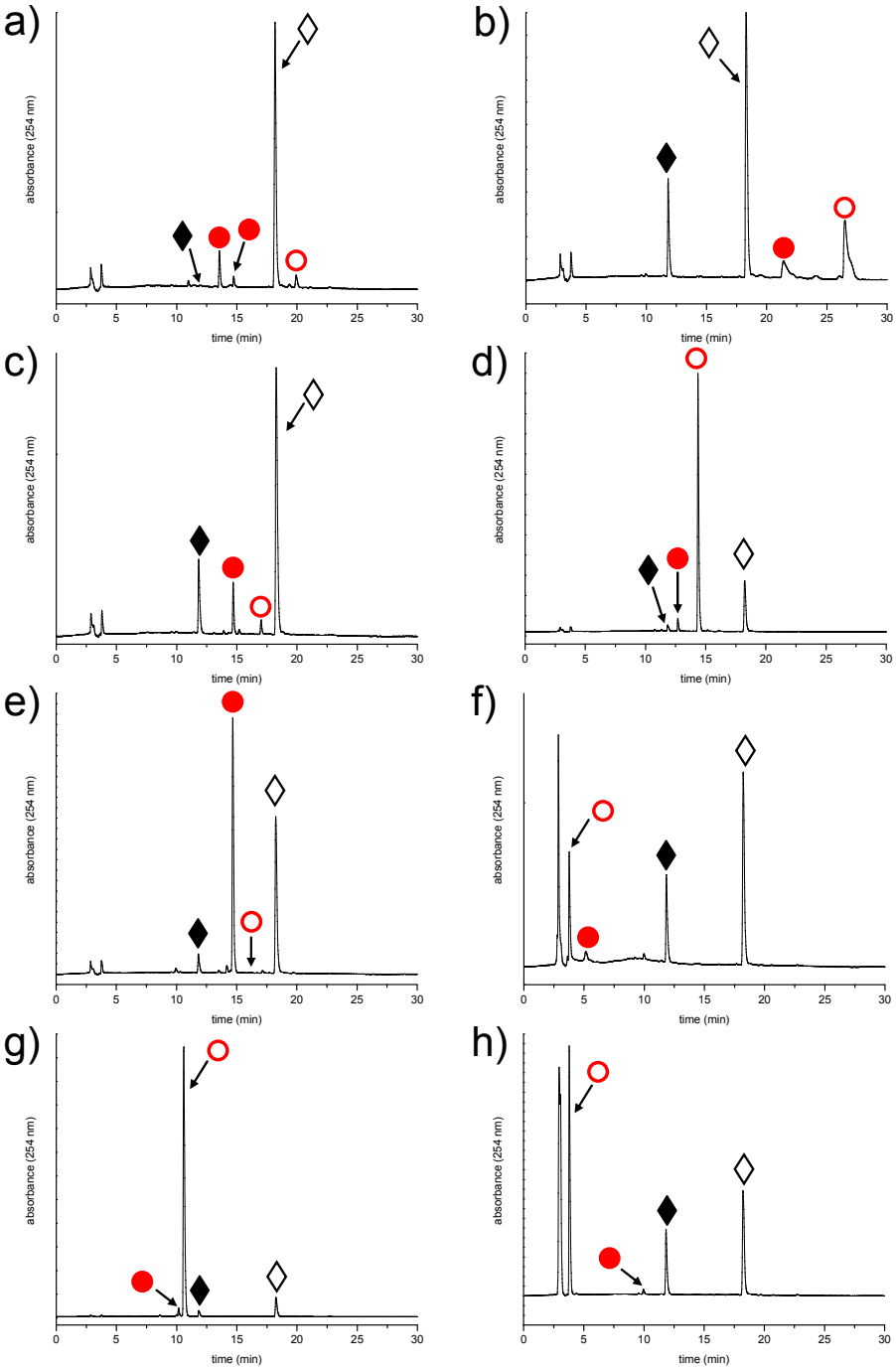


Figure A2.2. Representative drug screening HPLC data. a) 71 (2-methoxyestradiol), **b) 73** (rapamycin), **c) 78** (nelfinavir), **d) 80** (flavopiridol), **e) 85** (triclostatin A), **f) 86** ((-)-epinephrine), **g) 96** (pemetrexed), and **h) 99** (gemcitabine). Black closed diamonds (◆) denote **9**, black open diamonds (◇) denote 2-chloro-4-nitrophenol, red open circles (○) denote final acceptor aglycon, and red closed circles (●) denote glucosylated products.

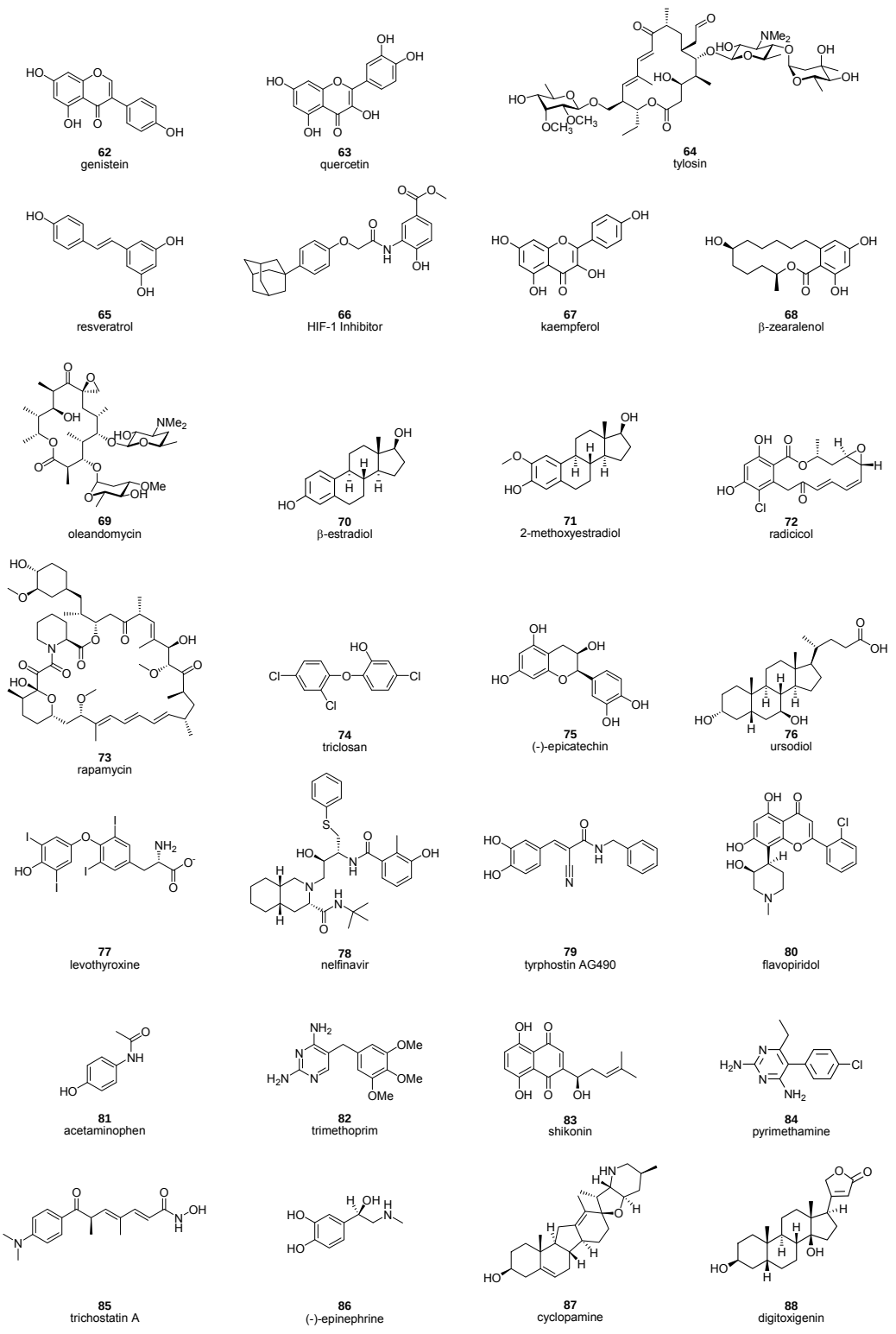


Figure A2.3. Compounds tested (62-111) in 50 member acceptor library screen. The structures of compounds not identified as 'hits' (**104-111**) are colored grey.

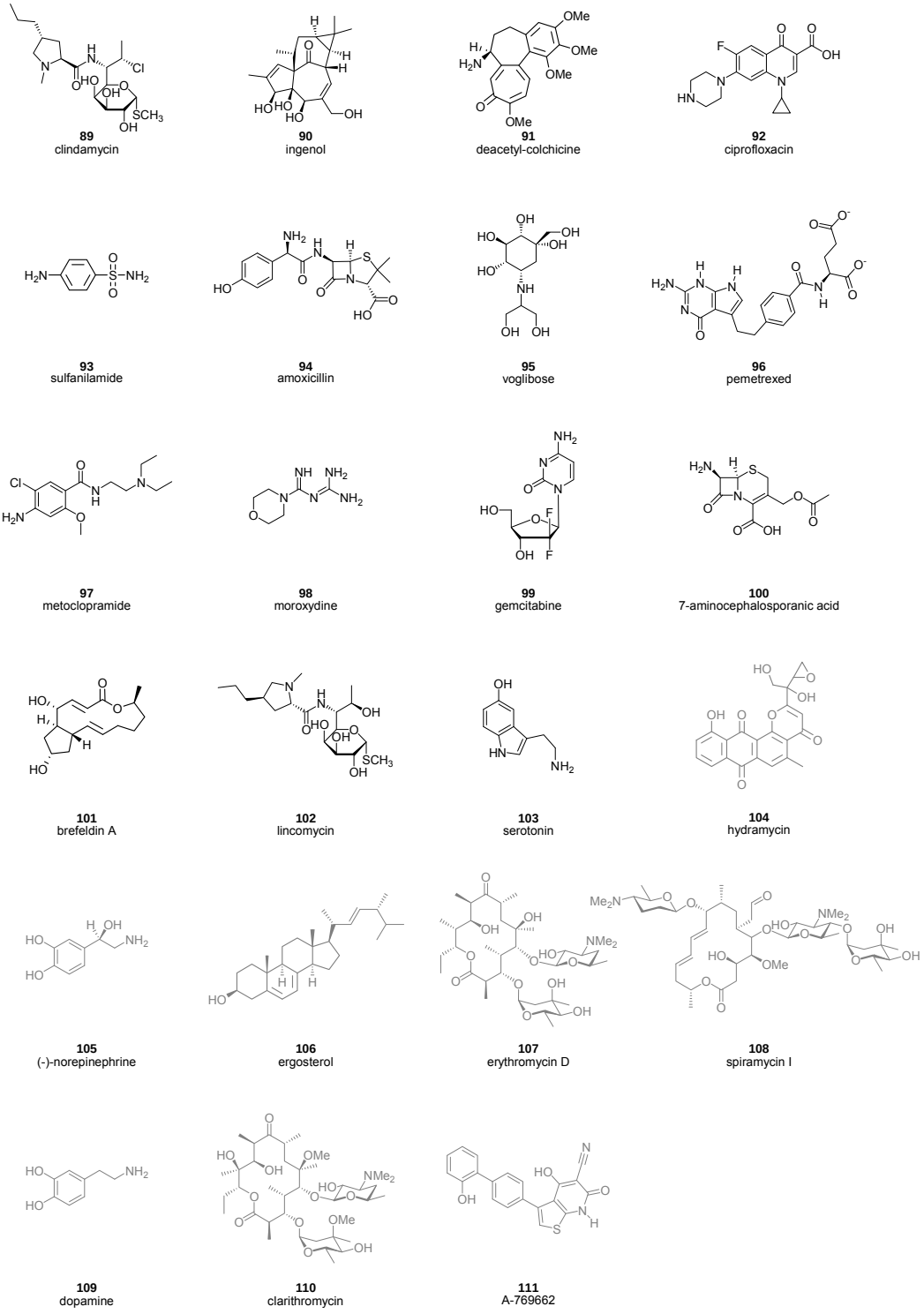


Figure A2.3 (continued). Compounds tested (62-111) in 50 member acceptor library screen. The structures of compounds not identified as 'hits' (**104-111**) are colored grey.

Entry	Compound	Rank by Area Under the Curve	Calculated Substrate Mass (*m/z*)[1]	Observed Substrate Mass (*m/z*)	Substrate Retention Time (minutes)	Calculated Product Mass (m/z)[1]	Observed Product Mass (*m/z*)	Product Retention Time	HPLC Evaluation	Substrate Retention Time (minutes)	Product Retention Time(s) (minutes)	Mass Previously Confirmed[2]
62	genistein	1	270.1	269.1 $[M-H]^-$	5.2	432.1	433.0 $[M+H]^+$; 467.0 $[M+Cl]^-$	3.9	Y	17.3	8.8; 12.2; 12.8	Y
							433.0 $[M+H]^+$; 431.0 $[M-H]^-$	4.2				
63	quercetin	2	302.0	n.f.[3]	---	464.1	465.0 $[M+H]^+$	3.5	Y	15.7	4.5; 6.1; 7.1; 8.1; 8.4; 9.7; 10.3; 10.6; 11.2; 11.6; 12.9	Y
							463.0 $[M-H]^-$	3.9				
							463.0 $[M-H]^-$	4.3				
64	tylosin	3	916.1	917.2 $[M-H]^-$	4.8	1078.0	1079.3 $[M+H]^+$	4.5	Y[4]	14.1	12.2; 10.5	Y
65	resveratrol	4	228.0	227.1 $[M-H]^-$	4.6	390.1	425.0 $[M+Cl]^-$	3.4	Y	15.1	8.2; 9.2; 10.8; 12.3; 12.7; 15.9	n.a.[5]
							425.0 $[M+Cl]^-$	3.6				
							425.0 $[M+Cl]^-$	3.9				
							425.0 $[M+Cl]^-$	4.2				
							425.0 $[M+Cl]^-$	4.4				
66	HIF-1 Inhibitor	5	435.2	434.0 $[M-H]^-$	3.5	597.3	595.8 $[M-H]^-$	3.4	Y	27.9	n.f.	n.a.
67	kaempferol	6	286.0	n.f.	---	448.1	449.2 $[M+H]^+$	3.2	Y	17.5	8.3; 8.8; 9.0; 9.9; 10.4; 12.5; 12.7; 13.0	Y
							449.2 $[M+H]^+$	3.6				
							449.2 $[M+H]^+$; 447.0 $[M-H]^-$	4.2				
68	β-zearalenol	7	322.2	n.f.	---	484.2	n.f.	---	Y	18.1	9.9; 10.8; 12.9; 14.0	Y
69	oleandomycin	8	687.4	686.3 $[M-H]^-$	4.1	849.5	850.2 $[M+H]^+$	4.0	Y[6]	14.3	13.3	Y
70	beta-estradiol	9	272.2	n.f.	---	434.2	433.2 $[M-H]^-$	4.5	Y	19.3	13.6	Y
71	2-methoxyestradiol	10	302.2	303.1 $[M+H]^+$	4.5	464.2	487.2 $[M+Na]^+$; 499.0 $[M+Cl]^-$	4.5	Y	19.9	13.6; 14.7	n.a.
							487.2 $[M+Na]^+$; 463.2 $[M-H]^-$	5.0				
72	radicicol	11	364.1	363.0 $[M-H]^-$	5.5	526.1	525.0 $[M-H]^-$	4.1	Y	17.6	10.6; 10.7; 11.1; 12.2; 12.5; 13.1; 13.4; 13.7; 14.4; 15.1; 16.3; 16.9; 18.9; 19.4; 20.1; 20.5; 20.8	n.a.
							525.0 $[M-H]^-$	4.5				
58	4-methylumbelliferone[7]	12	176.0	177 $[M+H]^+$	4.2	338.1	339.1 $[M+H]^+$	3.2	Y	14.2	9.6	Y
73	rapamycin	13	913.6	912.3 $[M+H]^+$	8.5	1075.6	1098.0 $[M+Na]^+$	7.1	Y	26.5	21.4	n.a.
							1098.0 $[M+Na]^+$	7.3				
74	triclosan	14	288.0	288.9 $[M+H]^+$	8.2	450.0	473.0 $[M+Na]^+$	3.2	Y	25.9	15.3	n.a.
75	(−)-epicatechin	15	290.1	289.0 $[M-H]^-$	3.4	452.1	451.0 $[M-H]^-$	3.0	Y	10.1	8.0; 9.1	n.a.
							451.0 $[M-H]^-$	3.3				
76	ursodiol	16	392.3	391.2 $[M-H]^-$	6.3	554.3	553.2 $[M-H]^-$	5.0	N	---	---	n.a.
							553.2 $[M-H]^-$	5.4				
							553.2 $[M-H]^-$	5.6				
77	levothyroxine	17	775.7	n.f.	---	937.7	n.f.	---	Y	16.3	12.3; 13.1	n.a.
78	nelfinavir	18	567.3	568.2 $[M+H]^+$	7.6	729.4	730.2 $[M+H]^+$	6.4	Y	17.0	14.7	n.a.
79	tyrphostin	19	294.1	293.1 $[M-H]^-$	5.6	456.2	457.0 $[M+H]^+$	4.8	Y	17.8	12.5; 14.4	n.a.
80	flavopiridol	20	401.1	402.1 $[M+H]^+$	4.7	563.2	564.3 $[M+H]^+$	3.7	Y	14.4	12.7	Y
							564.3 $[M+H]^+$	3.9				

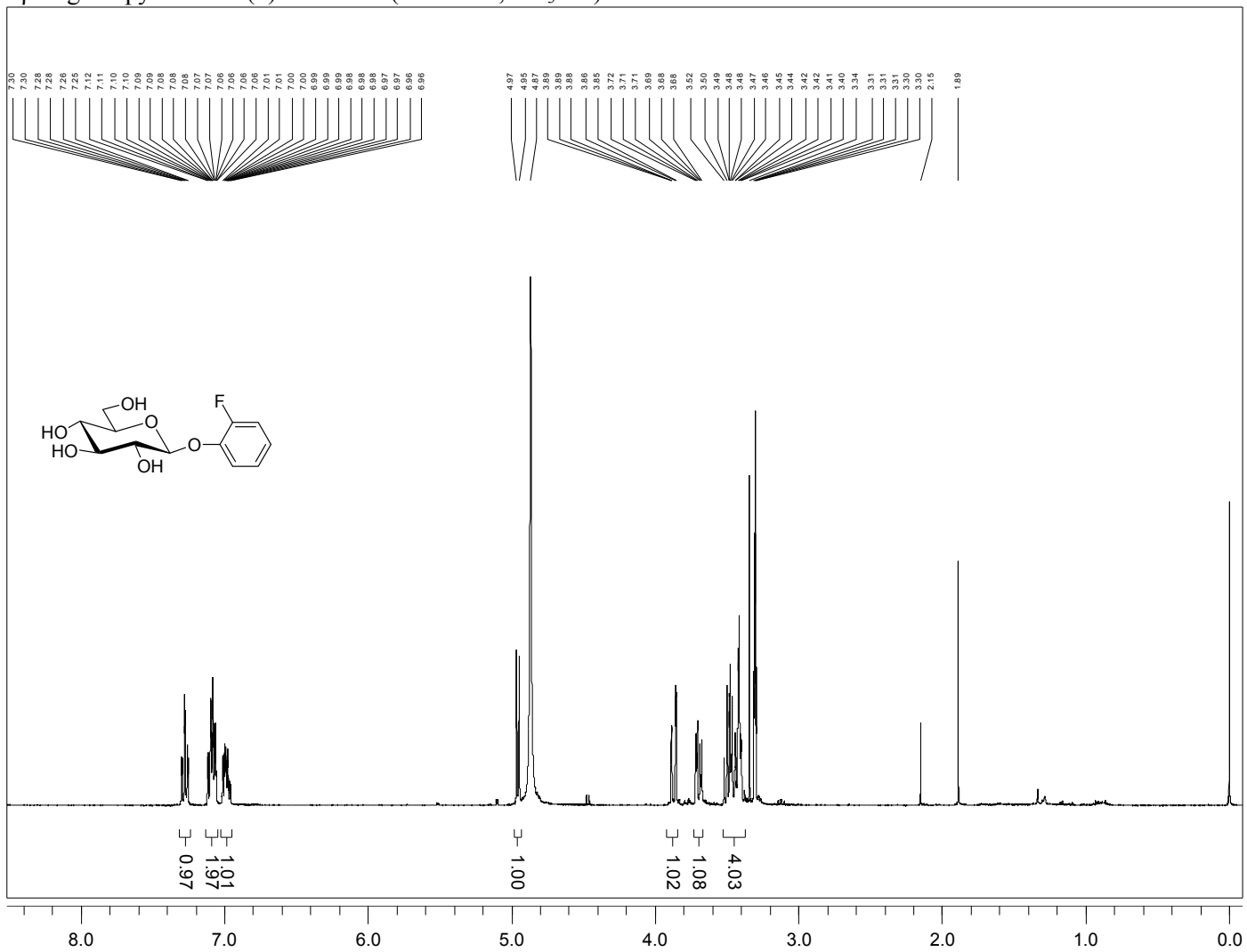
Table A2.1. Data summaries for drug screen hits. [1] Mass calculated before ionization. [2] Glucosylated products of substrate have been confirmed previously across various analytical techniques with other OleD variants(1-3). [3] n.f., not found. [4]Followed at 287nm. [5] n.a., not available. [6] Followed at 220nm and replaced solvent A with 50mM phosphate buffer (pH 6.3). [7]Utilized for positive control.

Entry	Compound	Rank by Area Under the Curve	Calculated Substrate Mass (*m/z*)[1]	Observed Substrate Mass (*m/z*)	Substrate Retention Time (minutes)		Calculated Product Mass (m/z)[1]	Observed Product Mass (*m/z*)	Product Retention Time		HPLC Evaluation	Substrate Retention Time (minutes)	Product Retention Time(s) (minutes)		Mass Previously Confirmed[2]
81	acetaminophen	21	151.1	150.6 $[M-H]^-$	0.9		313.1	348.0 $[M+Cl]^-$	0.6		Y	6.3	3.8		Y
82	trimethoprim	22	290.1	291.1 $[M+H]^+$	3.9		452.2	453.1 $[M+H]^+$	3.2		Y	10.3	10.0		Y
83	shikonin	23	290.1	291.0 $[M+H]^+$	2.6		452.2	453.1 $[M+H]^+$	3.2		Y	23.7; 26.1	12.3; 13.9; 14.1; 14.7; 15.0; 15.3; 16.5		n.a.[4]
84	pyrimethamine	24	248.1	249.0 $[M+H]^+$	4.2		410.1	411.0 $[M+H]^+$	3.5		Y	13.4	10.0; 11.1		Y
85	trichostatin A	25	302.2	337.2 $[M+Cl]^-$	4.9		464.2	465.1 $[M+H]^+$	5.1		Y	16.2	14.9		n.a.
86	(−)-epinephrine	26	183.1	n.f.[3]	---		345.1	n.f.	---		Y	3.8	5.2		n.a.
87	cyclopamine	27	411.3	413.3 $[M+H]^+$	4.7		573.4	574.3 $[M+H]^+$	3.8		N	---	---		n.a.
88	digitoxigenin	28	374.2	375.1 $[M+H]^+$	5.4		536.3	535.2 $[M-H]^-$; 571.2 $[M+Cl]^-$	4.5		Y	17.9	13.5		Y
89	clindamycin	29	424.2	425.1 $[M+H]^+$	5.8		586.2	609.3 $[M+Na]^+$	3.8		N	---	---		n.a.
90	lngenol	30	348.2	347.1 $[M-H]^-$	5.1		510.2	509.1 $[M-H]^-$	4.5		N	---	---		n.a.
91	deacetyl-colchicine	31	357.2	358.2 $[M+H]^+$	4.0		519.2	520.2 $[M+H]^+$	4.2		Y	11.5	11.7		Y
92	ciprofloxacin	32	331.1	332.3 $[M+H]^+$	6.3		493.2	n.f.	---		Y	10.9	10.0		Y
93	sulfanilamide	33	172.0	172.1 $[M-H]^-$	0.6		334.1	369.9 $[M+Cl]^-$	3.5		Y	4.1	10.0		Y
94	amoxicillin	34	365.1	n.f.	---		527.2	n.f.	---		Y	n.f.	n.f.		n.a.
95	voglibose	35	267.1	268.2 $[M+H]^+$	3.0		429.2	n.f.	---		N	---	---		n.a.
96	pemetrexed	36	425.1	448.2 $[M+Na]^+$	3.2		587.2	n.f.	---		Y	10.6	10.2		n.a.
97	metoclopramide	37	299.1	300.1 $[M+H]^+$	3.0		461.2	462.1 $[M+H]^+$	2.4		Y	11.3	10.1		n.a.
98	moroxydine	38	171.1	n.f.	---		333.2	n.f.	---		N	---	---		n.a.
99	gemcitabine	39	263.1	298.0 $[M+Cl]^-$	0.6		425.1	426.3 $[M+H]^+$	3.1		Y	3.8	10.0		n.a.
100	7-aminocephalosporanic acid	40	272.0	n.f.	---		434.1	n.f.	---		Y	n.f.	n.f.		n.a.
101	brefeldin A	41	280.2	303.2 $[M+Na]^+$	5.2		442.2	441.2 $[M-H]^-$	4.7		Y	16.1	13.6		Y
102	lincomycin	42	406.2	407.3 $[M+H]^+$	4.6		568.3	n.f.	---		N	---	---		n.a.
103	serotonin	43	176.1	177.1 $[M+H]^+$	3.8		338.1	373.4 $[M+Cl]^-$	3.5		Y	4.7	3.1		n.a.

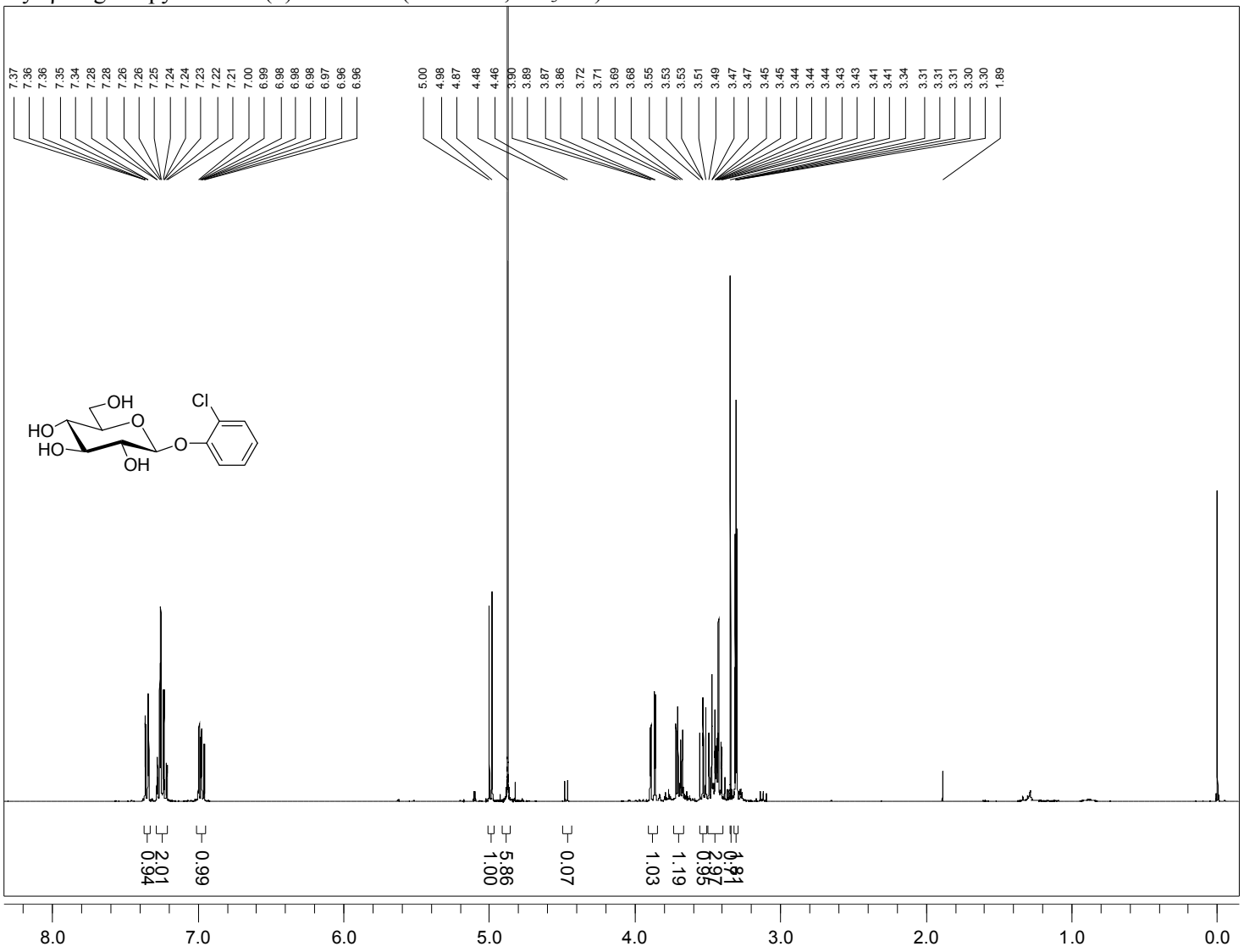
Table A2.1 (continued). Data summaries for drug screen hits. [1] Mass calculated before ionization. [2] Glucosylated product(s) of substrate have been confirmed previously across various analytical techniques with other OleD variants(1-3). [3] n.f., not found. [4] n.a., not available.

A2.2. ^{1}H NMR Spectra

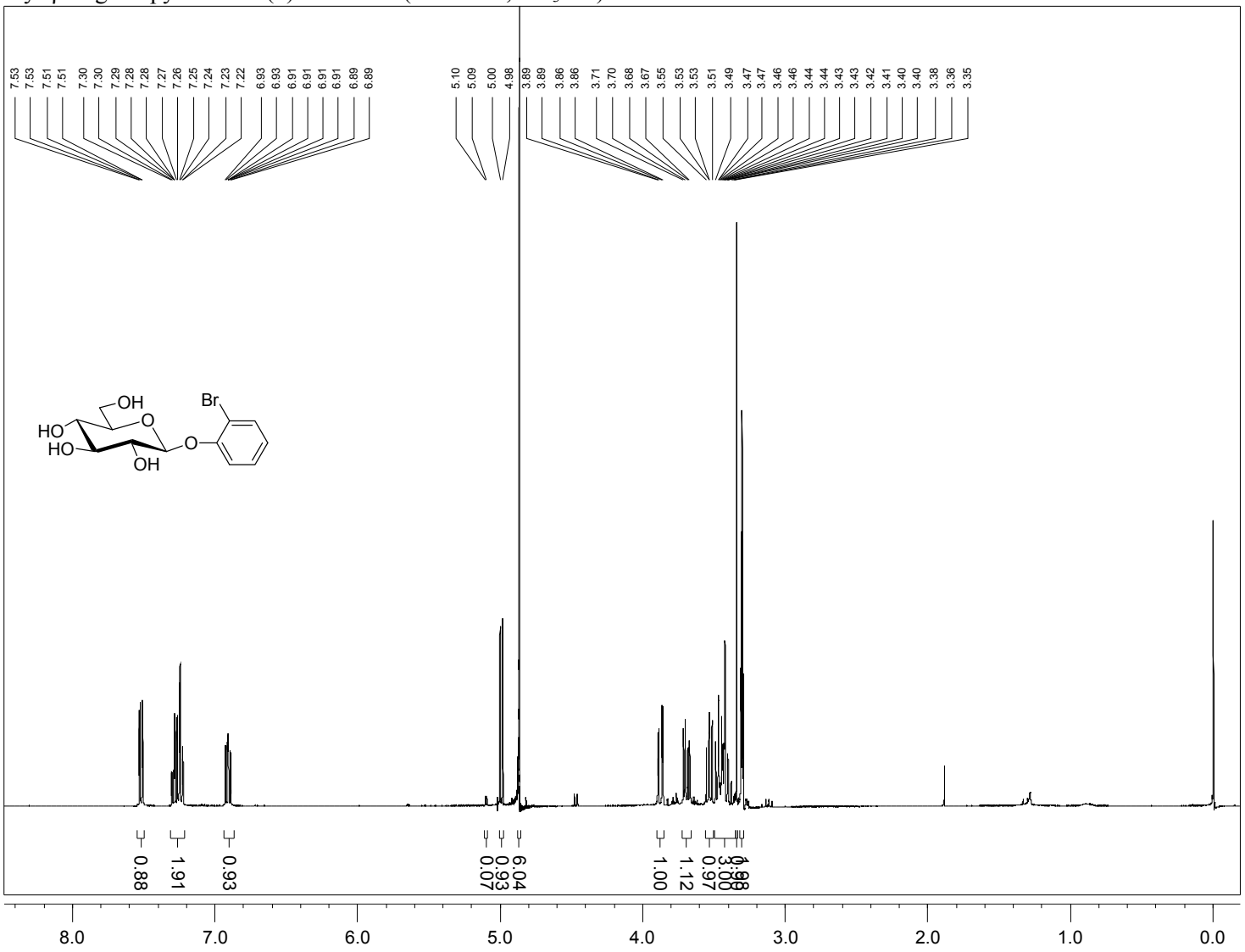
2-fluorophenyl-β-D-glucopyranoside (**3**): ^{1}H NMR (400 MHz, CD_3OD)



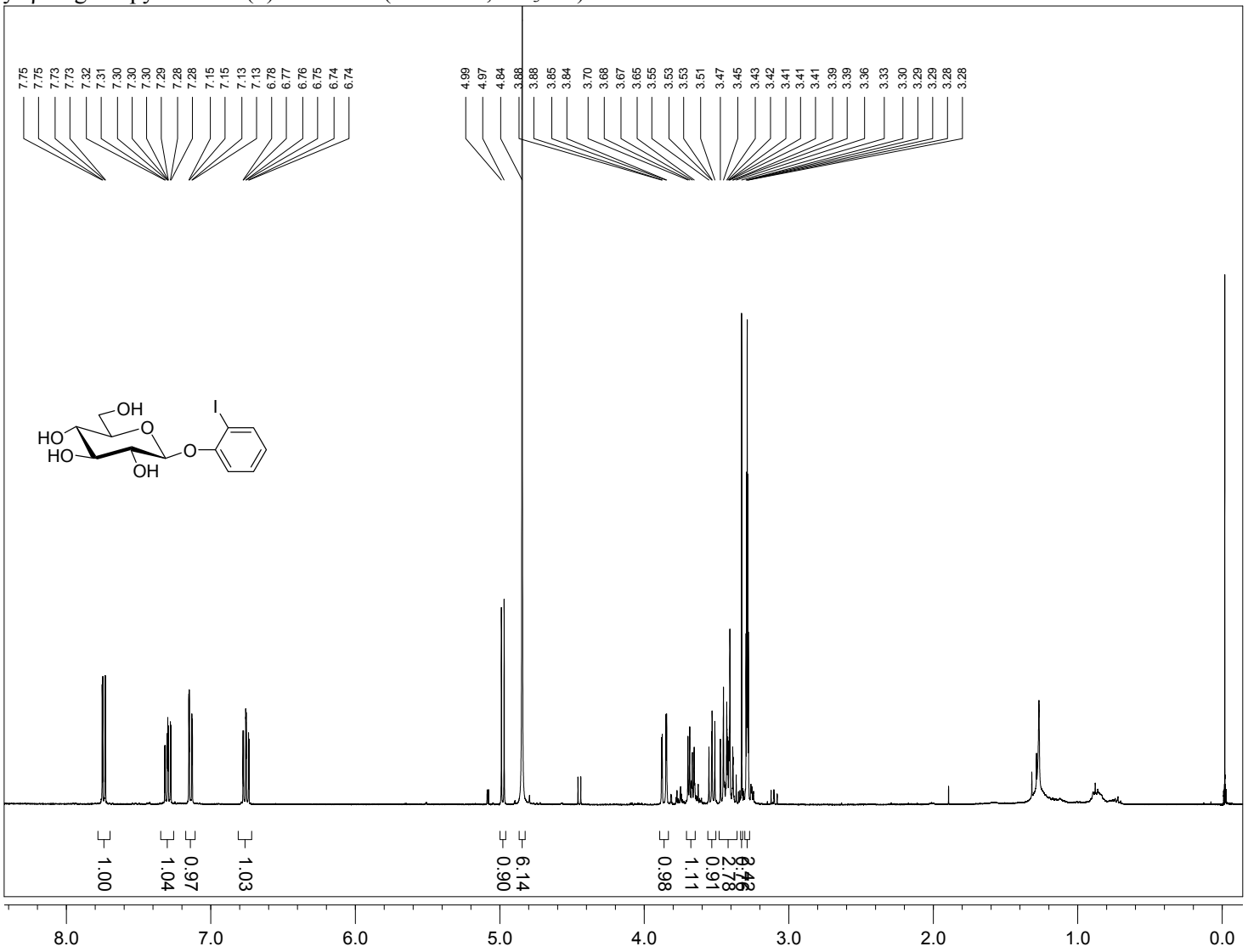
2-chlorophenyl-β-D-glucopyranoside (**4**): ^{1}H NMR (400 MHz, CD_3OD)



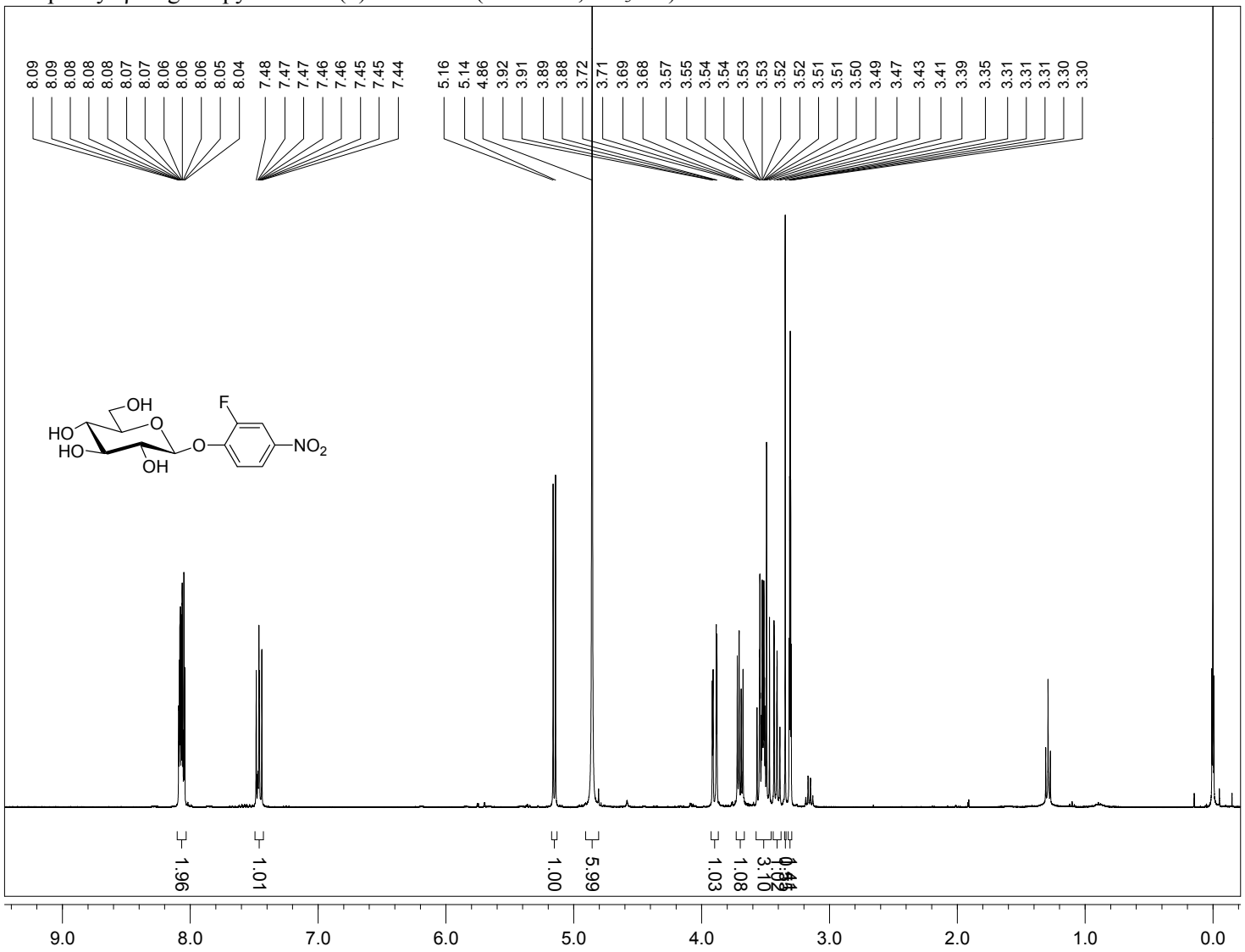
2-bromophenyl-β-D-glucopyranoside (**5**): ^{1}H NMR (400 MHz, CD_3OD)



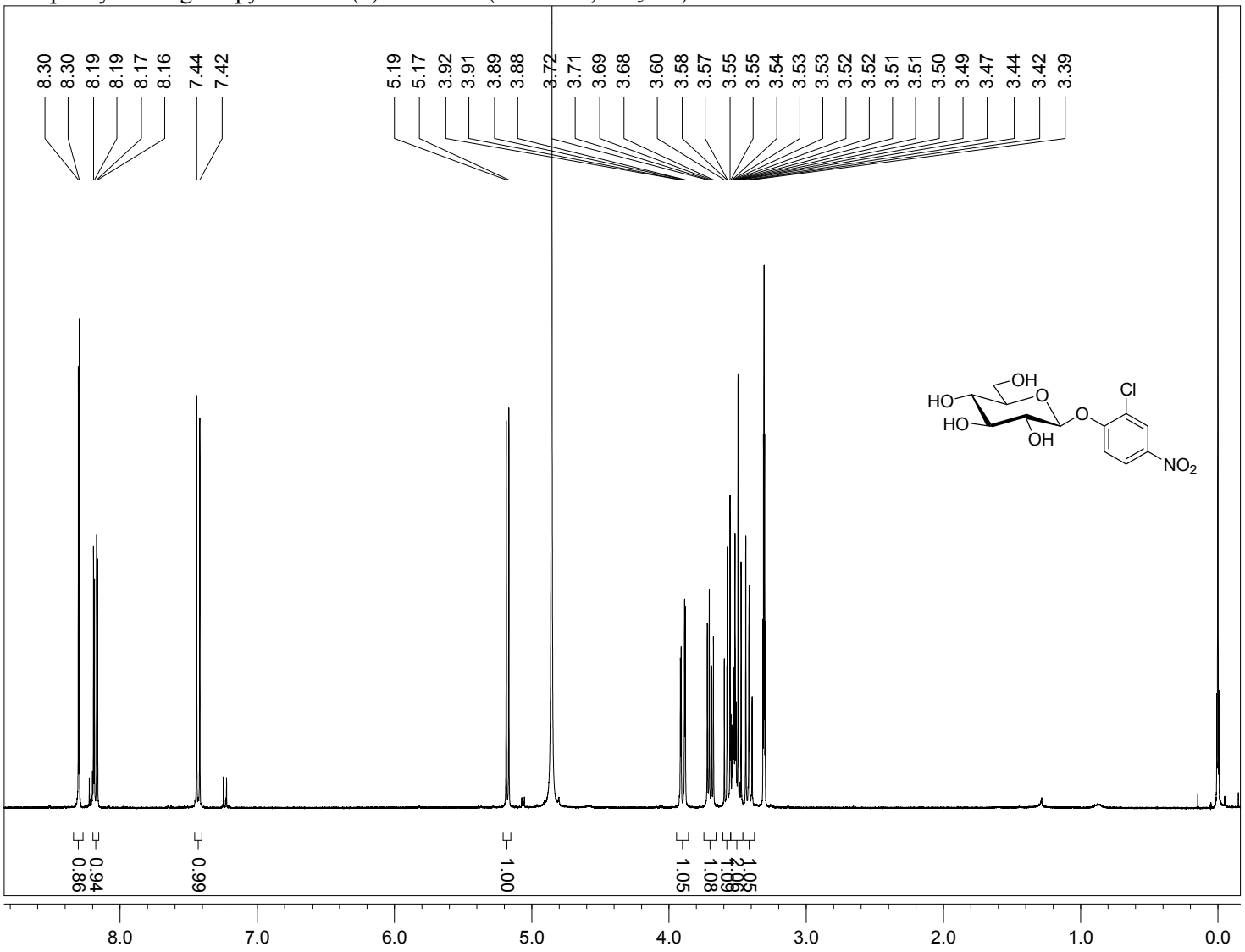
2-iodophenyl-β-D-glucopyranoside (**6**): ^{1}H NMR (400 MHz, CD_3OD)



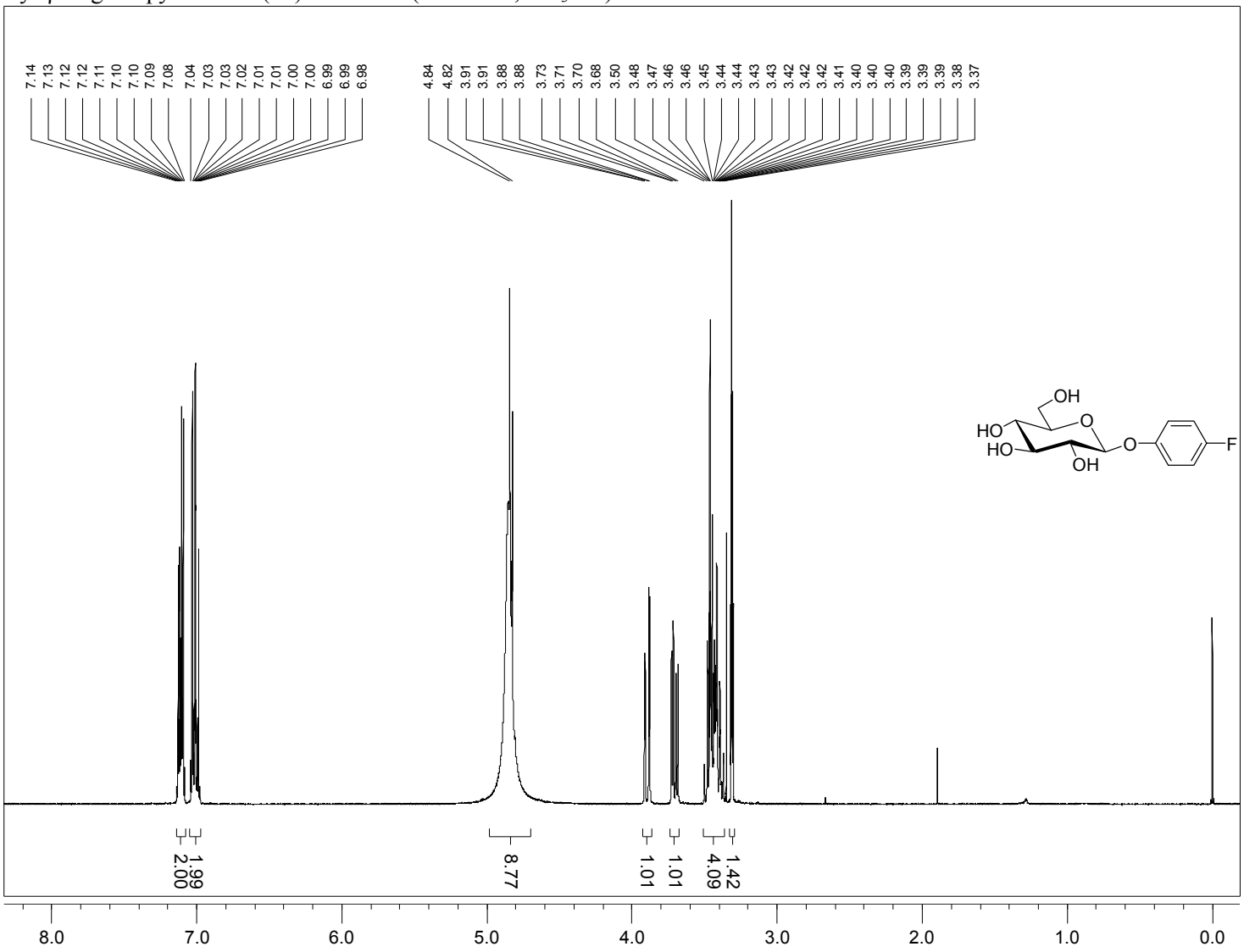
2-fluoro-4-nitrophenyl-β-D-glucopyranoside (**8**): ^{1}H NMR (400 MHz, CD_3OD)



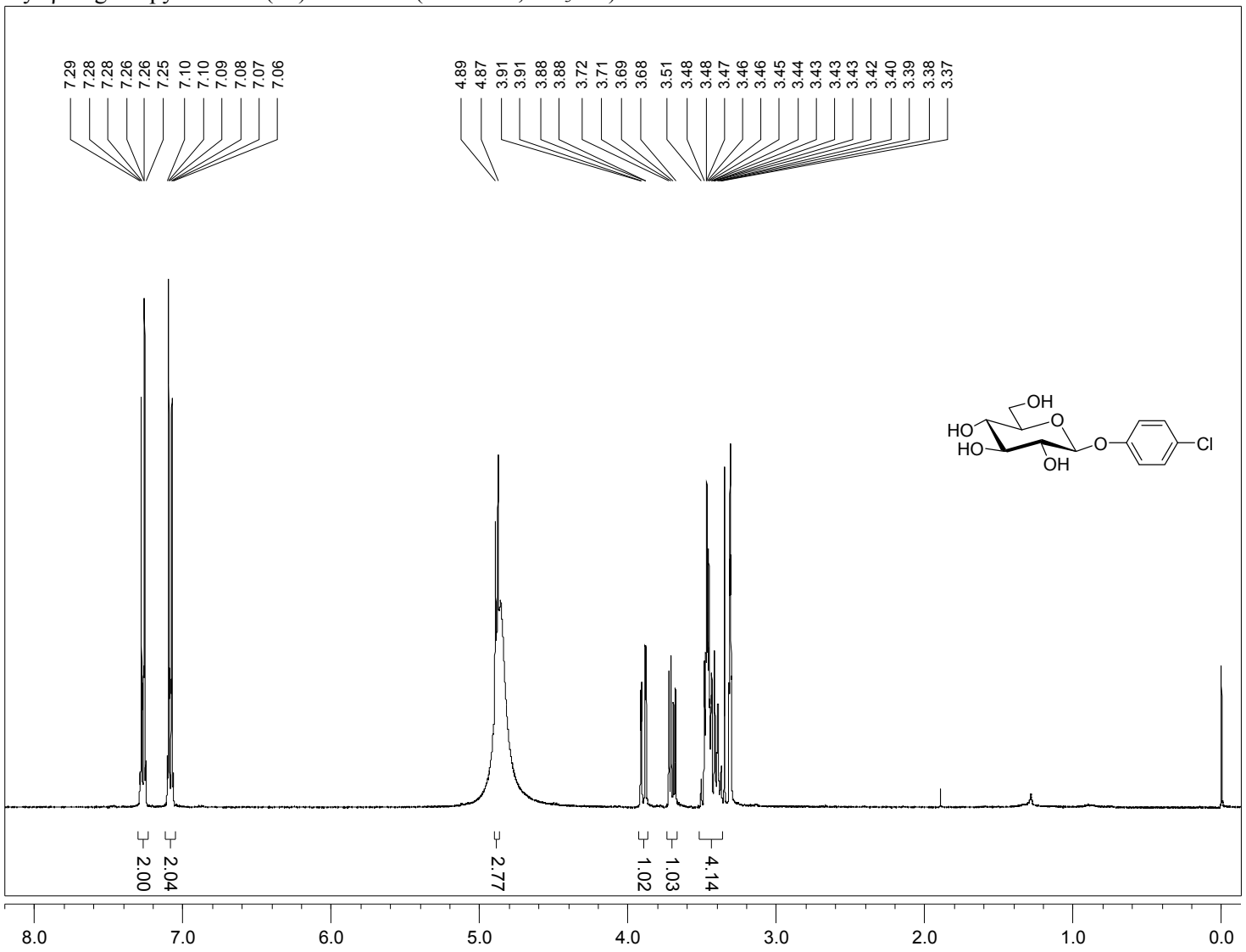
2-chloro-4-nitrophenyl-□-D-glucopyranoside (**9**): ^{1}H NMR (400 MHz, CD_3OD)



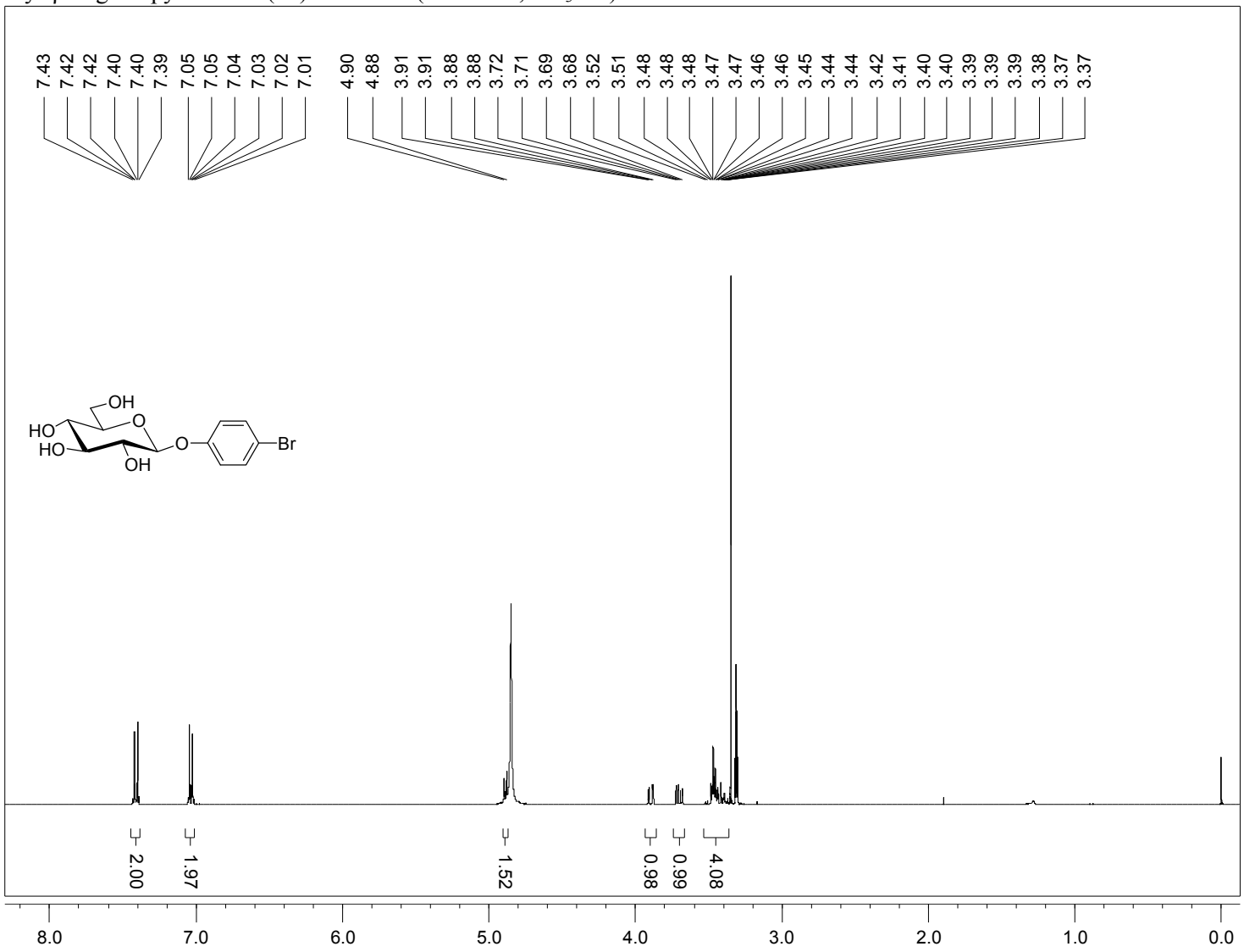
4-fluorophenyl-β-D-glucopyranoside (**11**): ^{1}H NMR (400 MHz, CD_3OD)



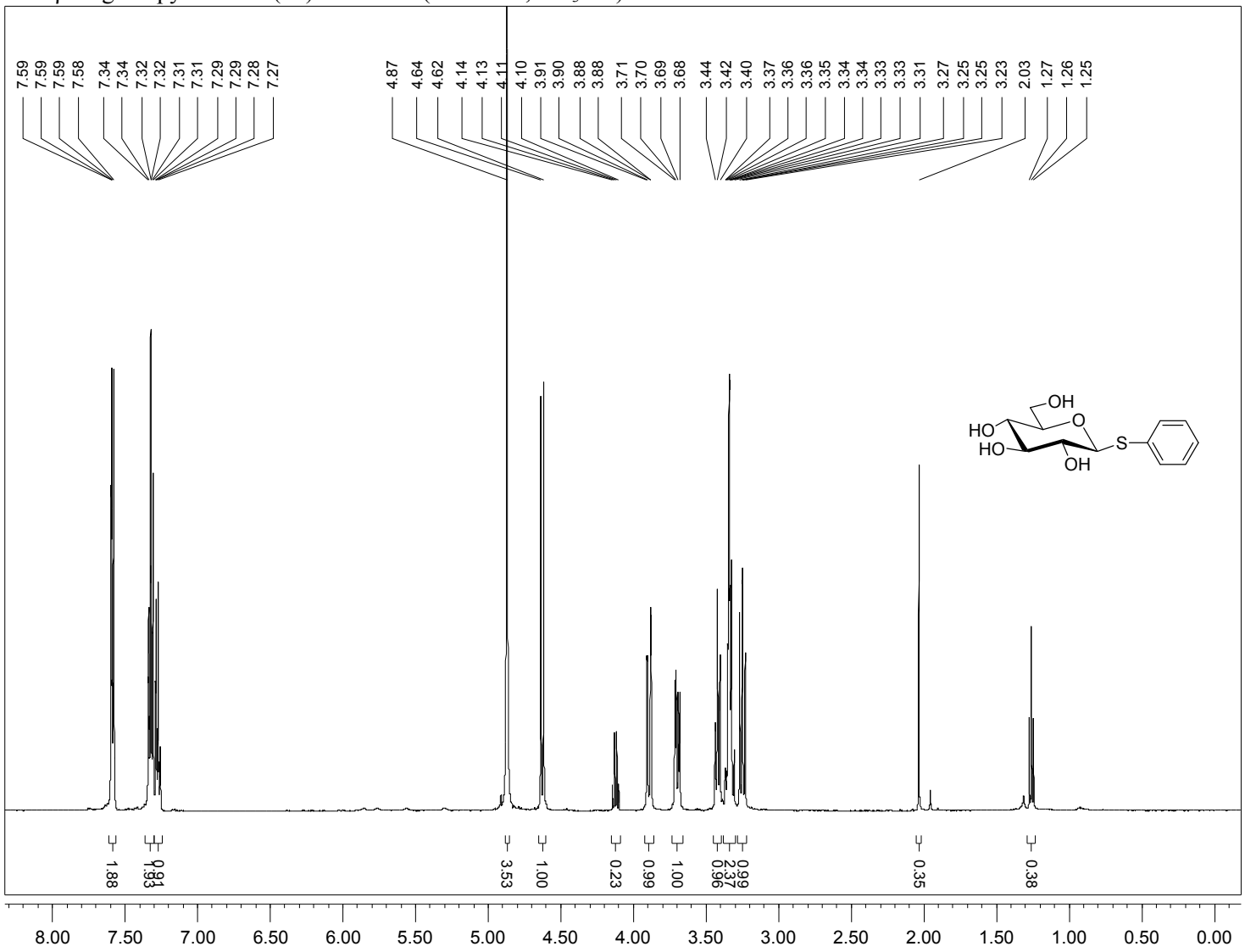
4-chlorophenyl-β-D-glucopyranoside (**12**): ^{1}H NMR (400 MHz, CD_3OD)



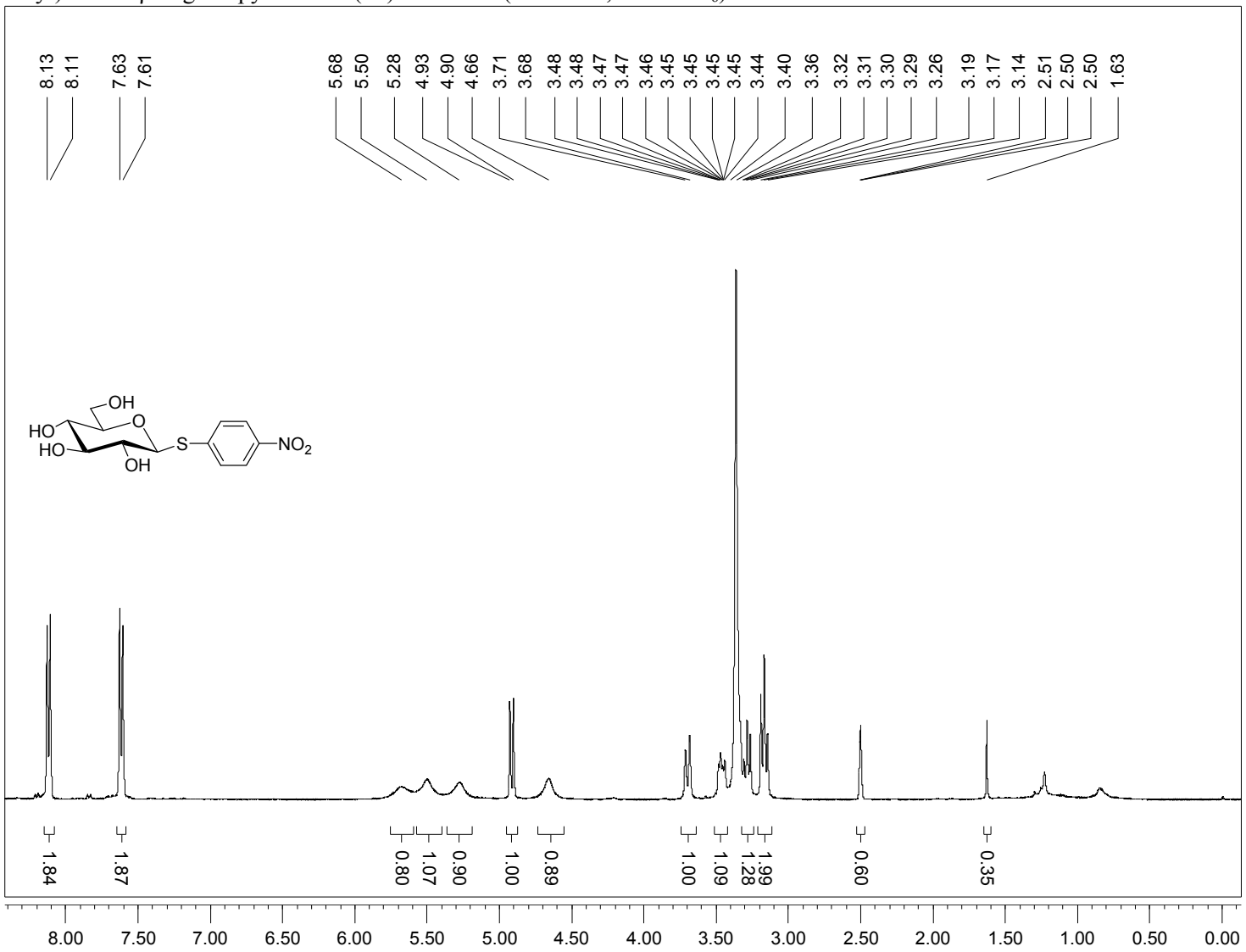
4-bromophenyl-β-D-glucopyranoside (**13**): ^{1}H NMR (400 MHz, CD_3OD)



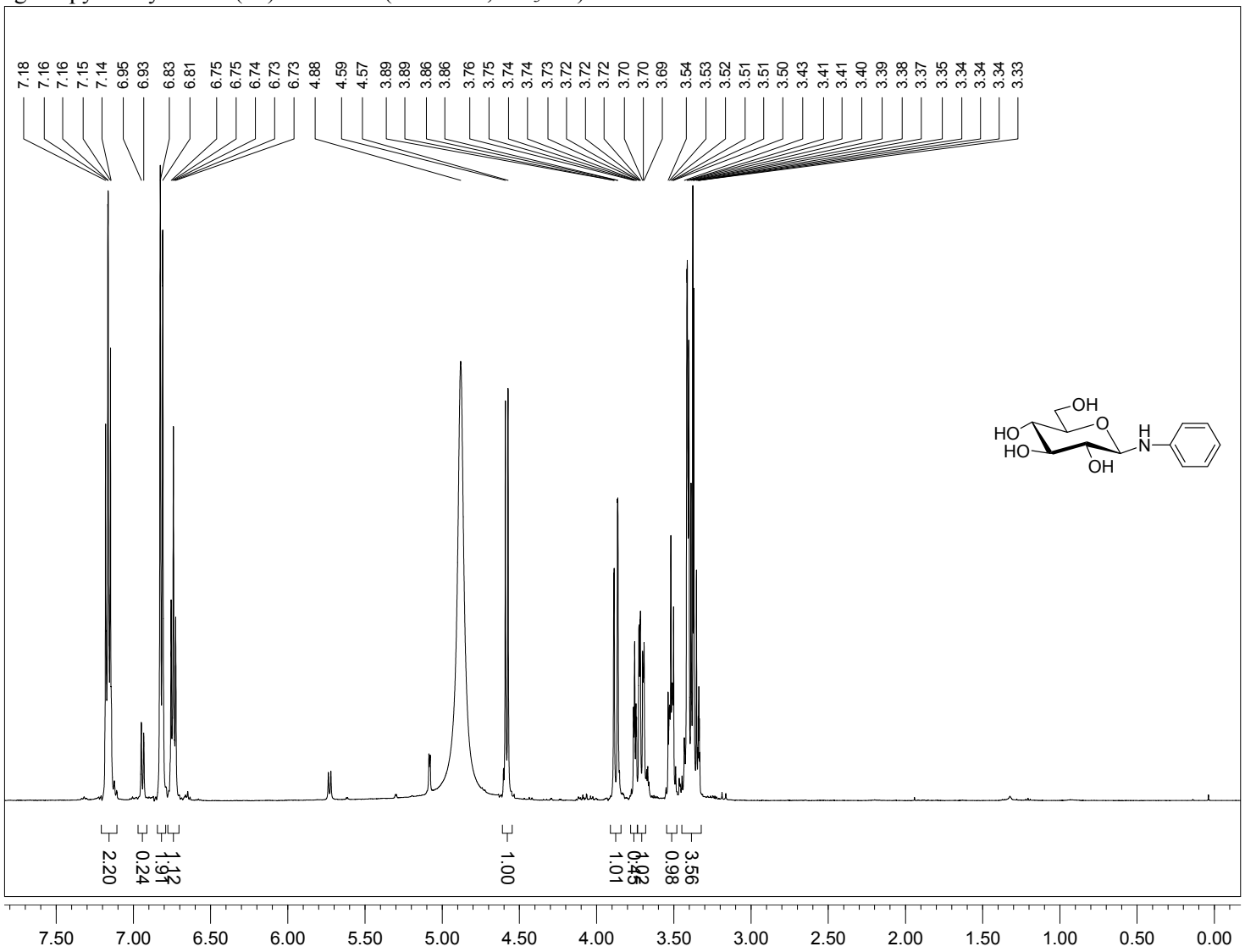
S-phenyl-1-thio-β-D-glucopyranoside (**14**): ^{1}H NMR (500 MHz, CD_3OD)



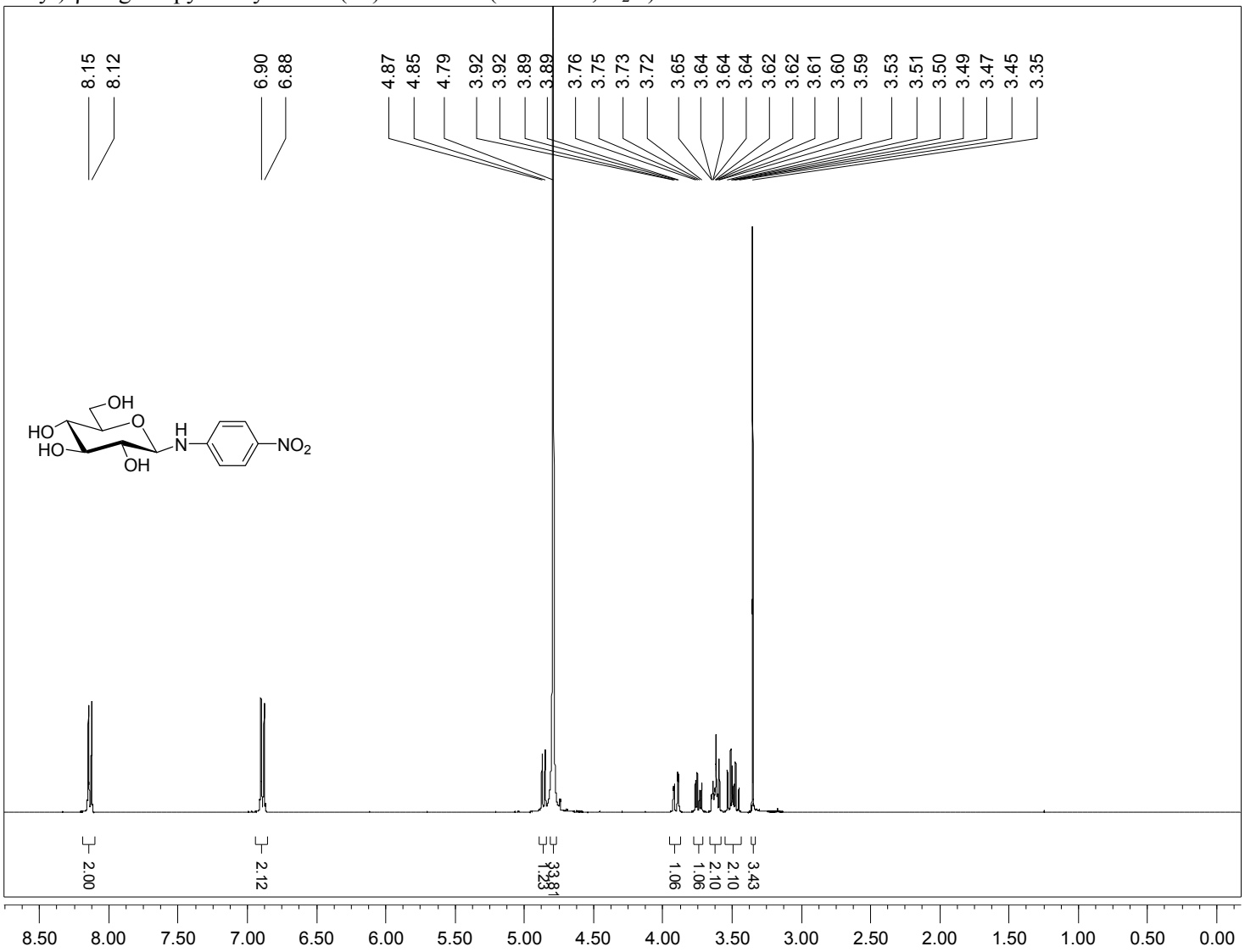
S-(4-nitrophenyl)-1-thio-β-D-glucopyranoside (**15**): ^{1}H NMR (400 MHz, DMSO-d_6)



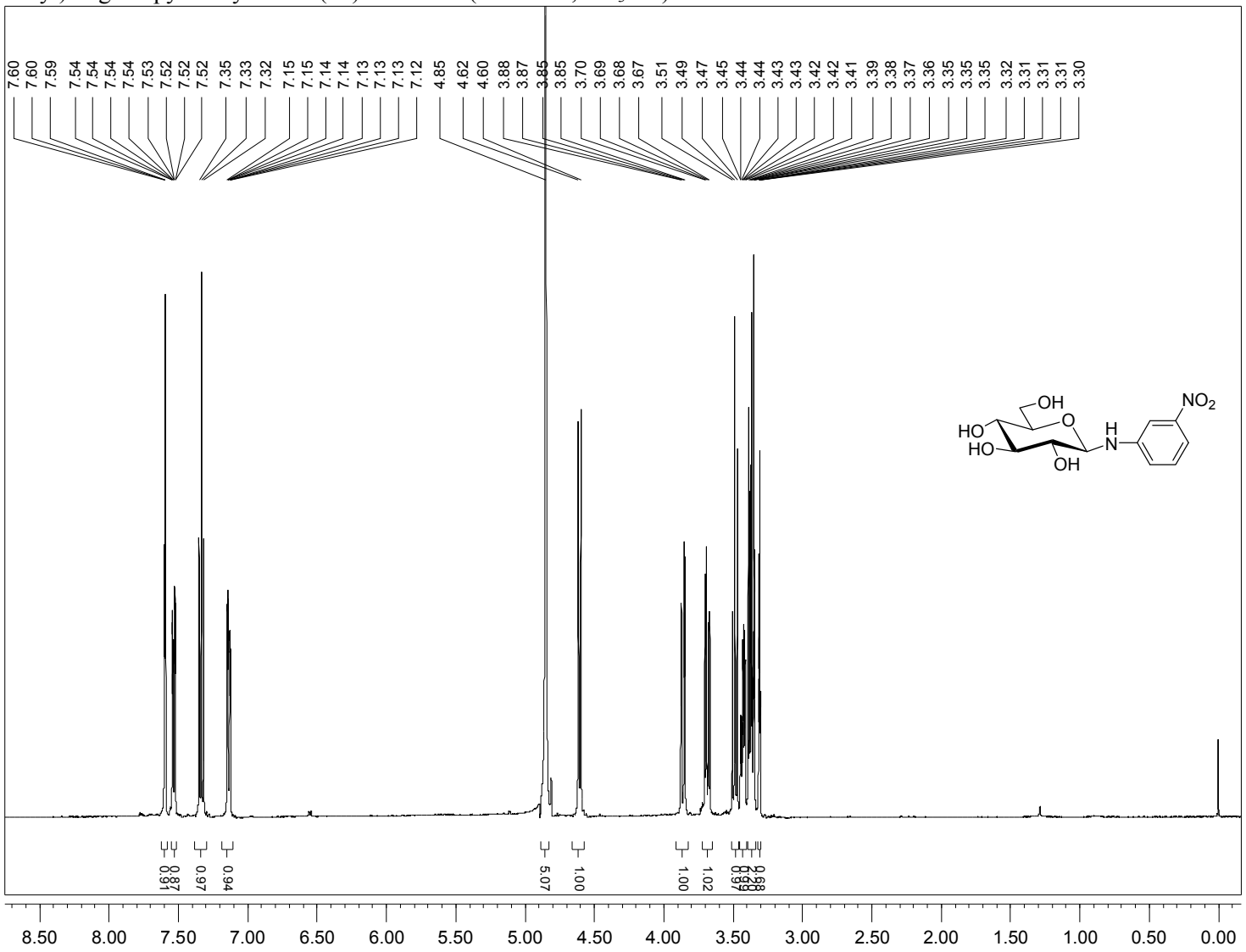
N-phenyl-D-glucopyranosylamine (**16**): ^{1}H NMR (500 MHz, CD_3OD)



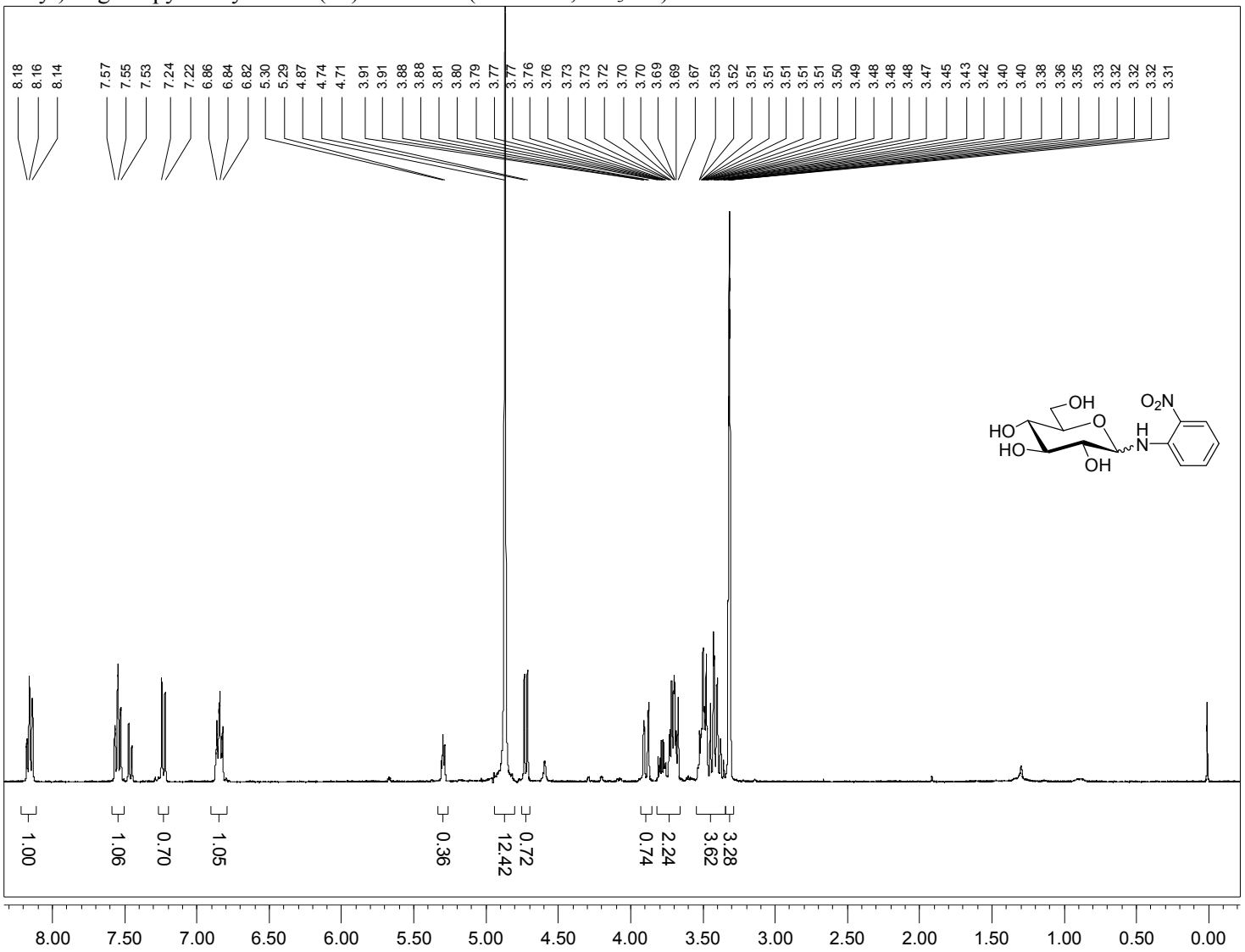
N-(4-nitrophenyl)-β-D-glucopyranosylamine (**17**): ^{1}H NMR (400 MHz, D_2O)



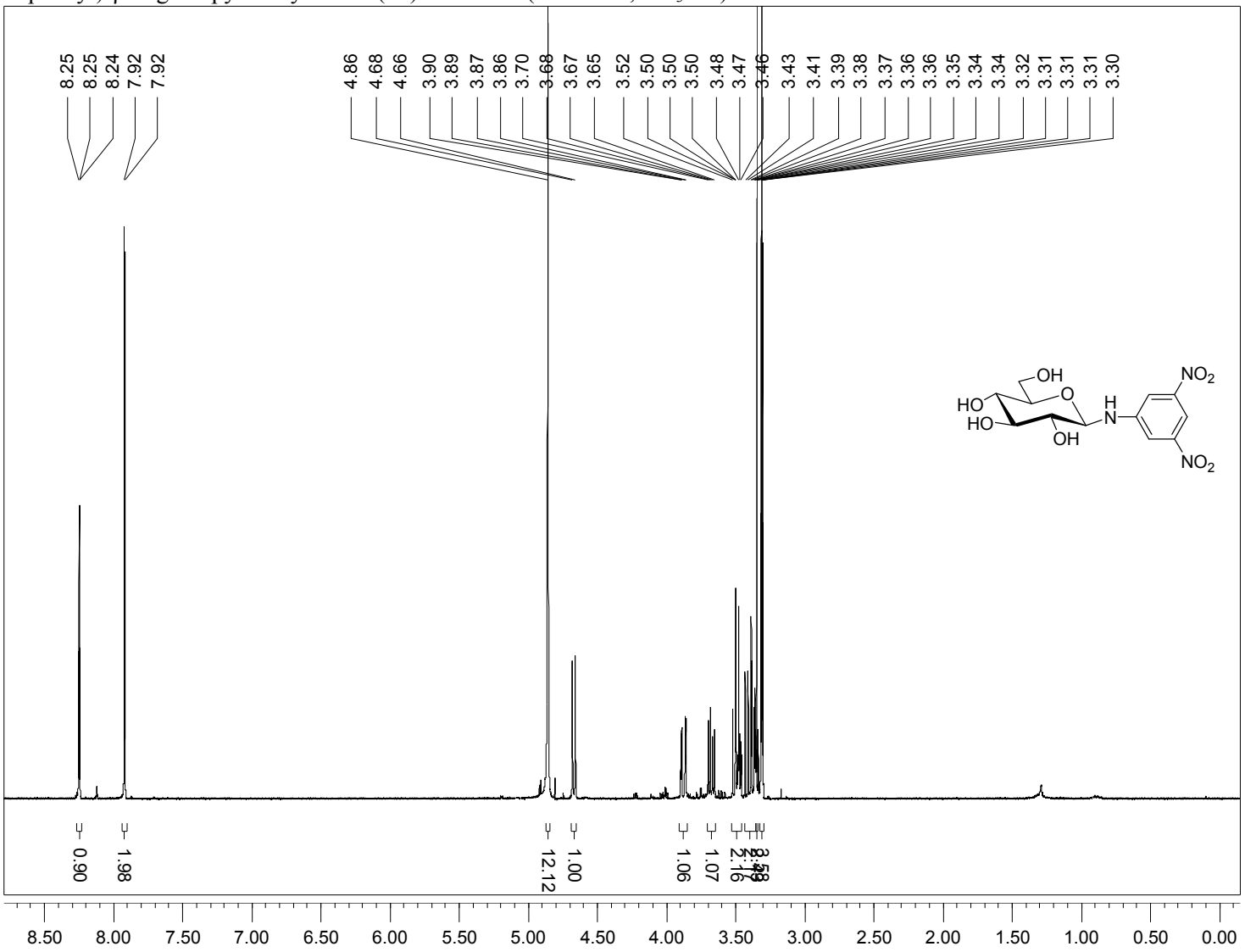
N-(3-nitrophenyl)-D-glucopyranosylamine (**18**): ^{1}H NMR (500 MHz, CD_3OD)



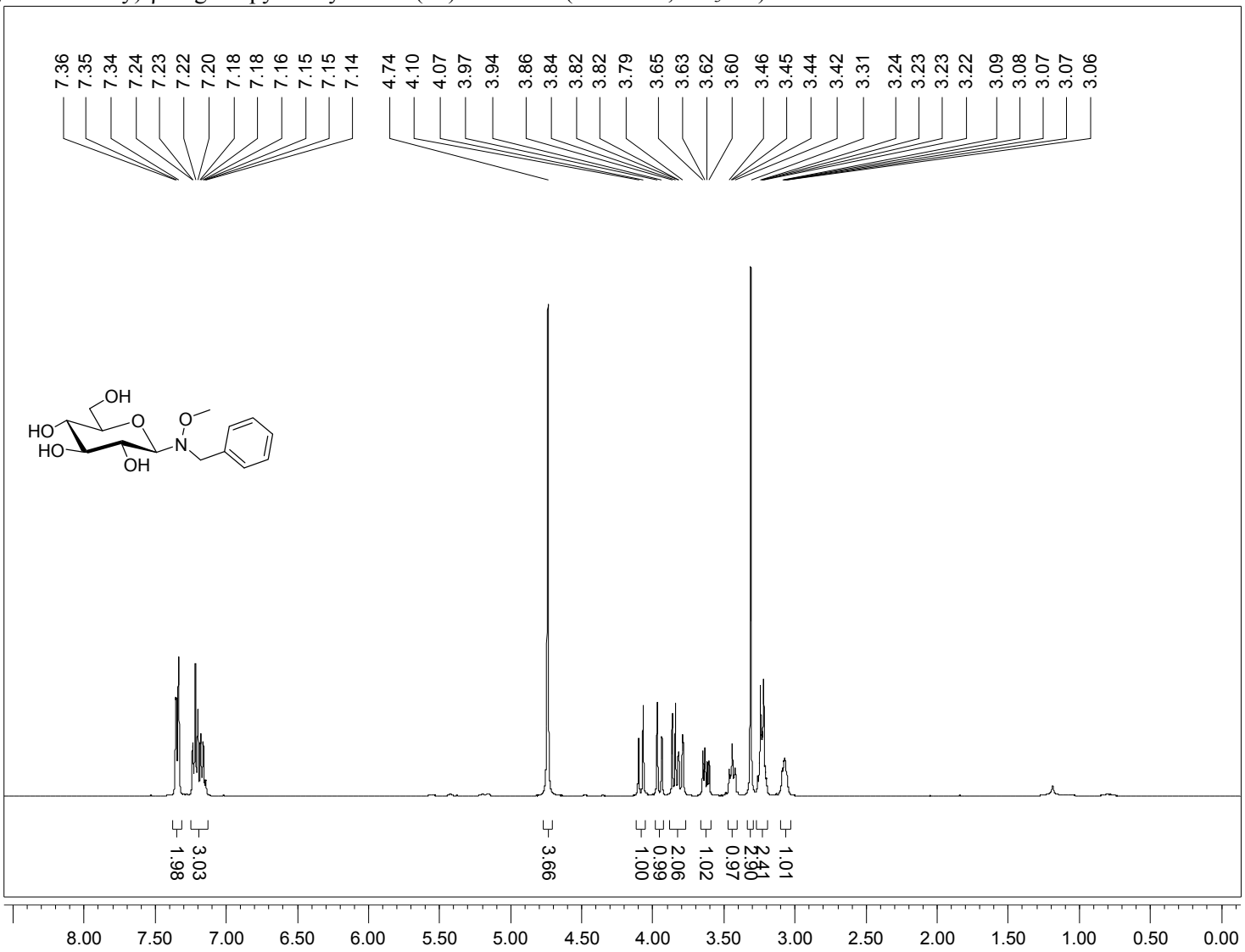
N-(2-nitrophenyl)-D-glucopyranosylamine (**19**): ^{1}H NMR (400 MHz, CD_3OD)



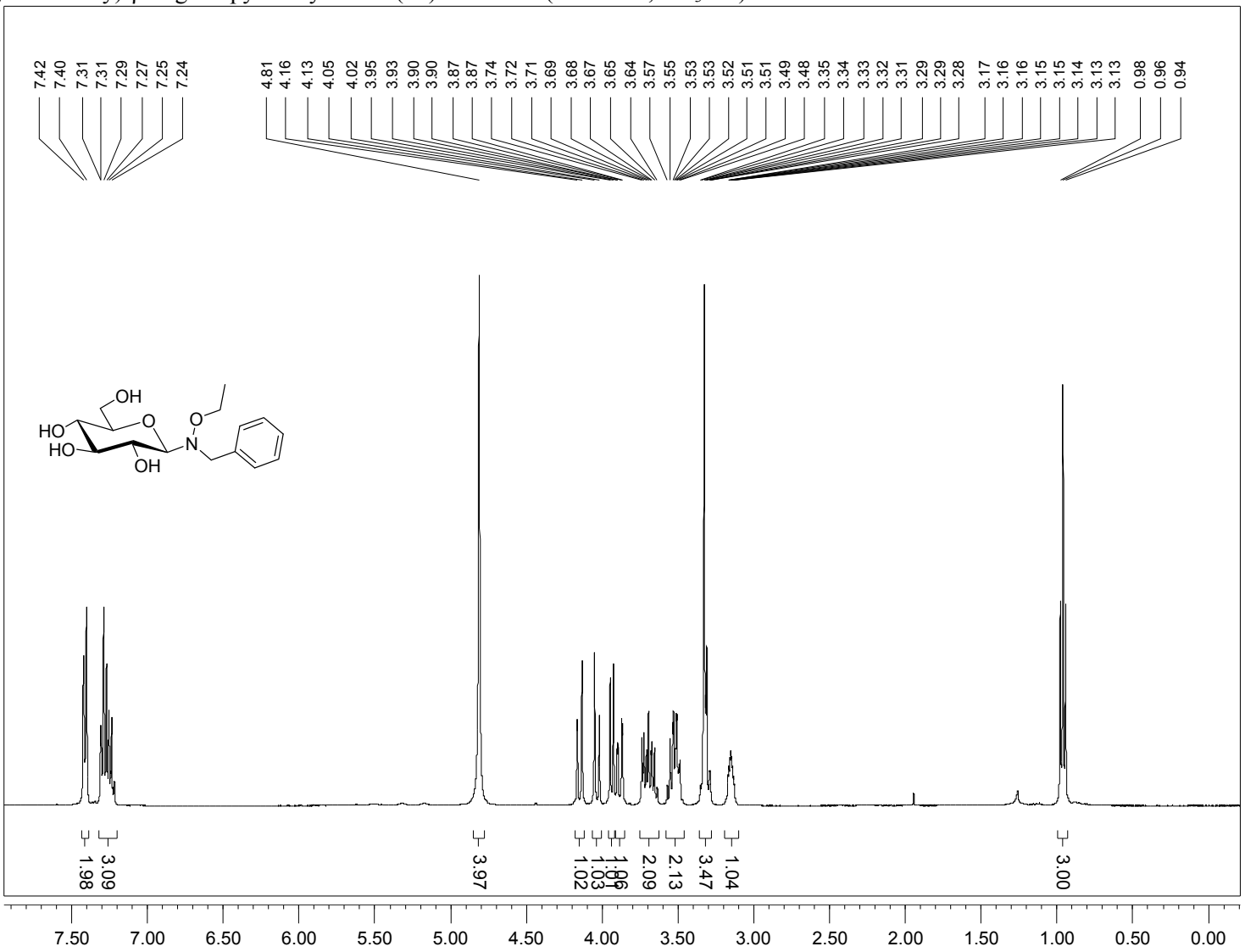
N-(3,5-dinitrophenyl)-β-D-glucopyranosylamine (**20**): ^{1}H NMR (400 MHz, CD_3OD)



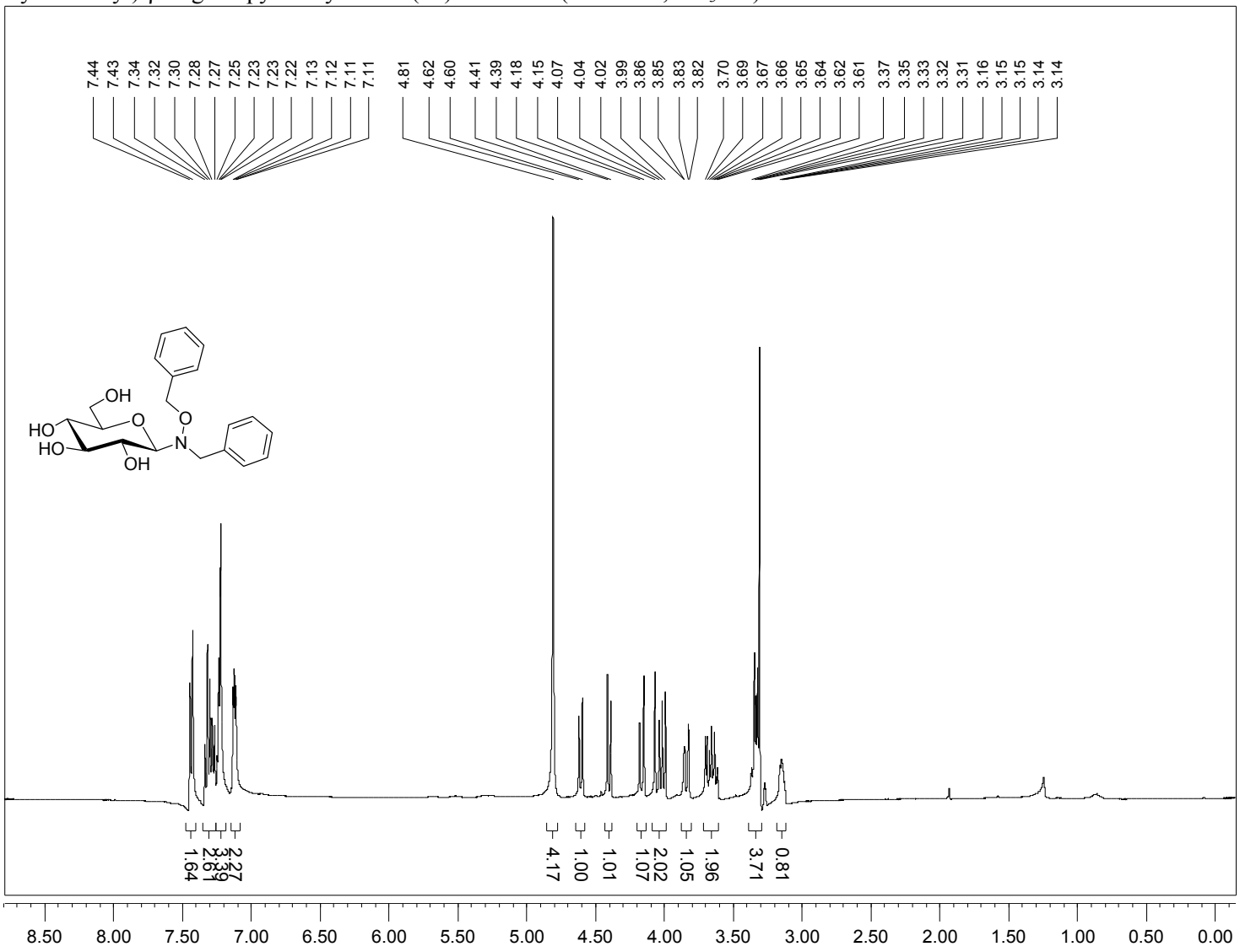
N-(*N*-benzyl-*N*-methoxy)-β-D-glucopyranosylamine (**21**): ^{1}H NMR (400 MHz, CD_3OD)



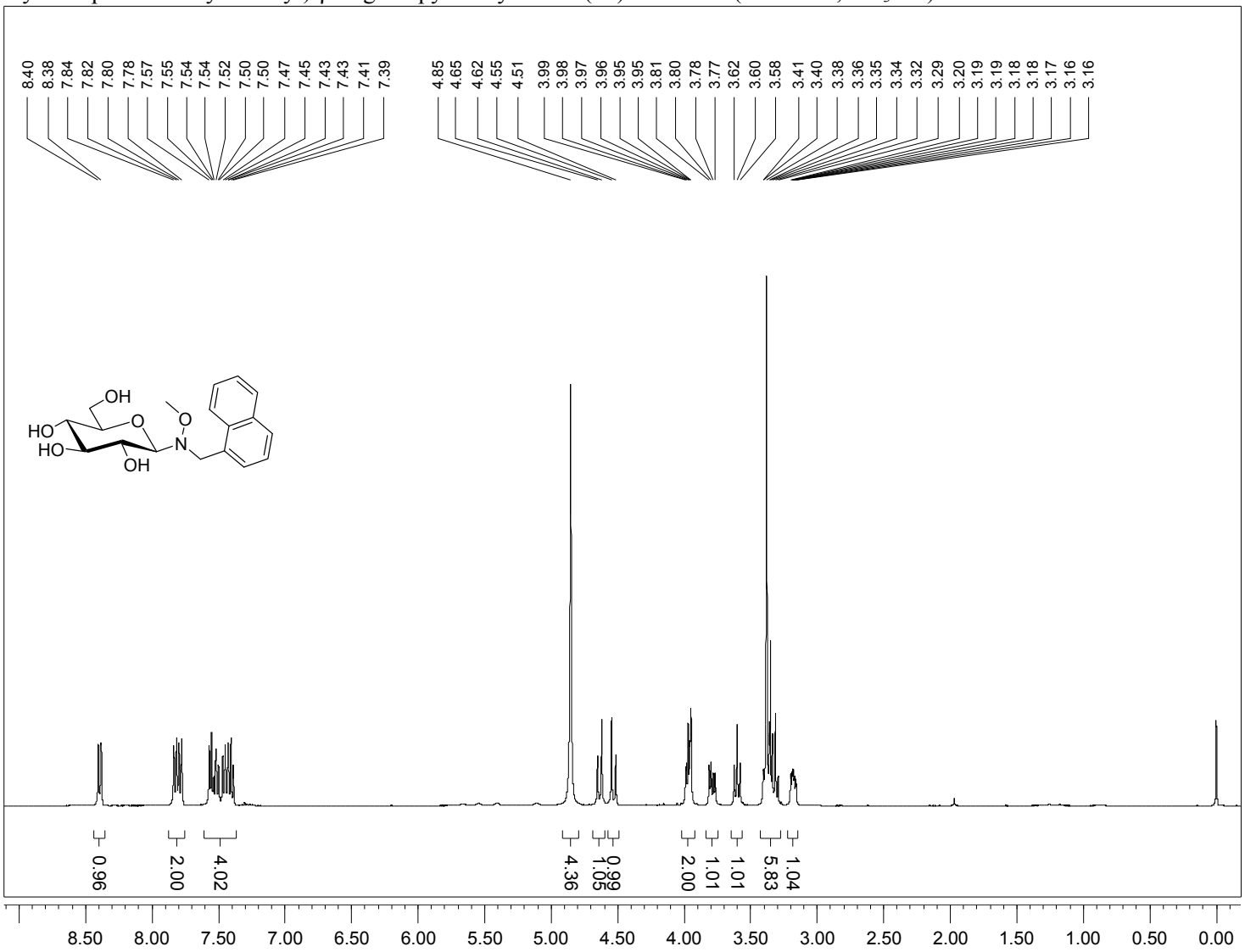
N-(*N*-benzyl-*N*-ethoxy)-β-D-glucopyranosylamine (**22**): ^{1}H NMR (400 MHz, CD_3OD)



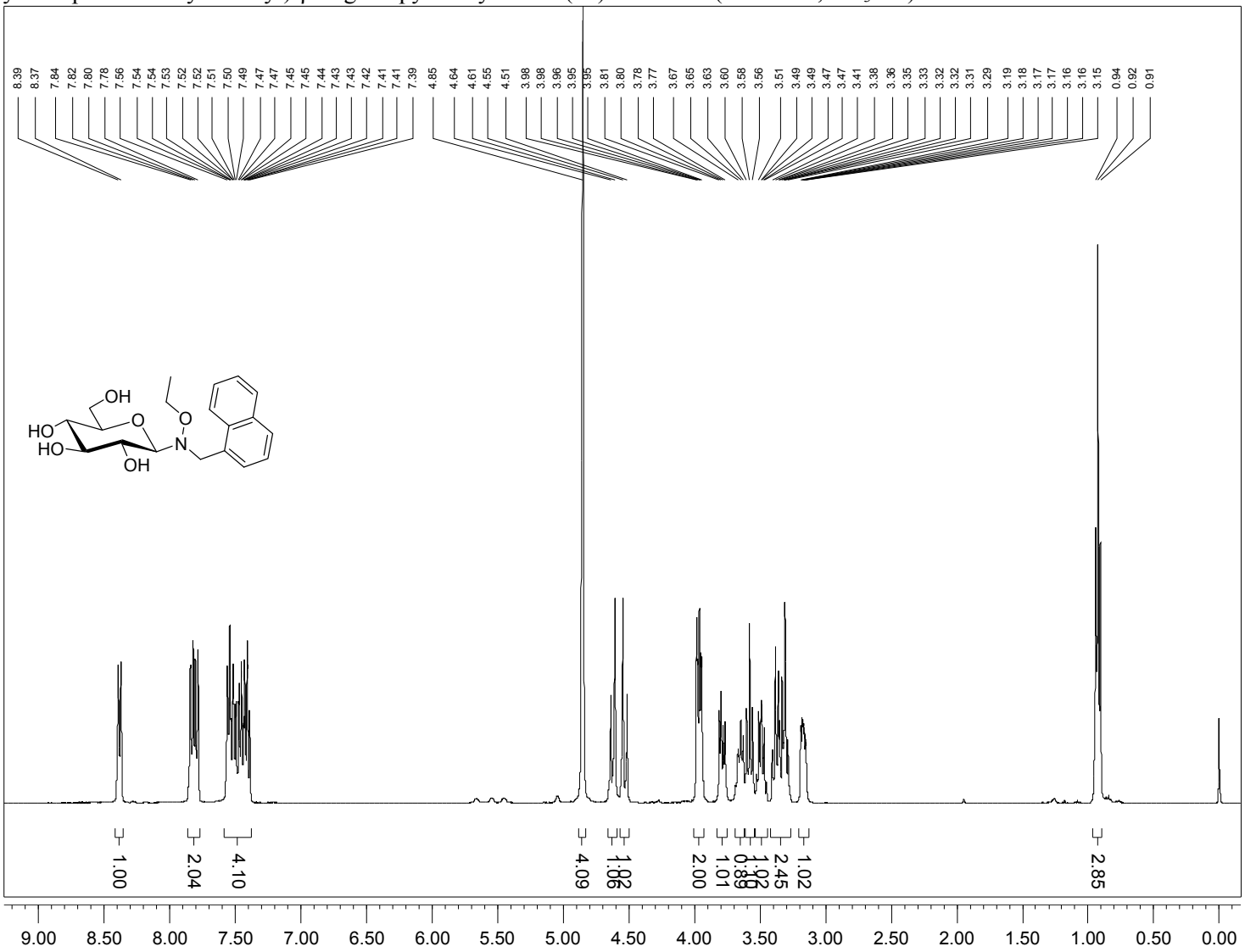
N-(*N*-benzoxy-*N*-benzyl)-β-D-glucopyranosylamine (**23**): ^{1}H NMR (400 MHz, CD_3OD)



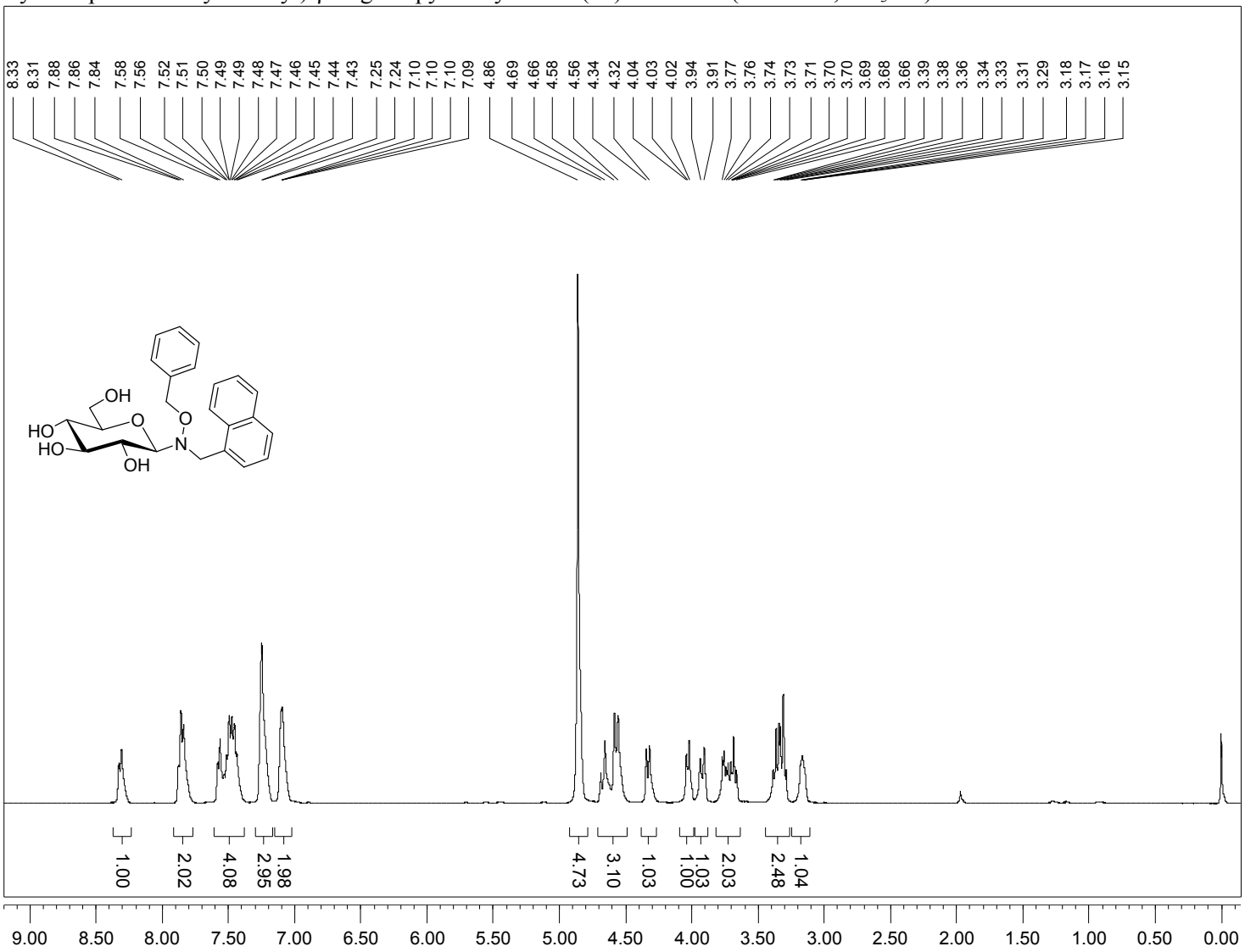
N-(*N*-methoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (**24**): ^{1}H NMR (400 MHz, CD_3OD)



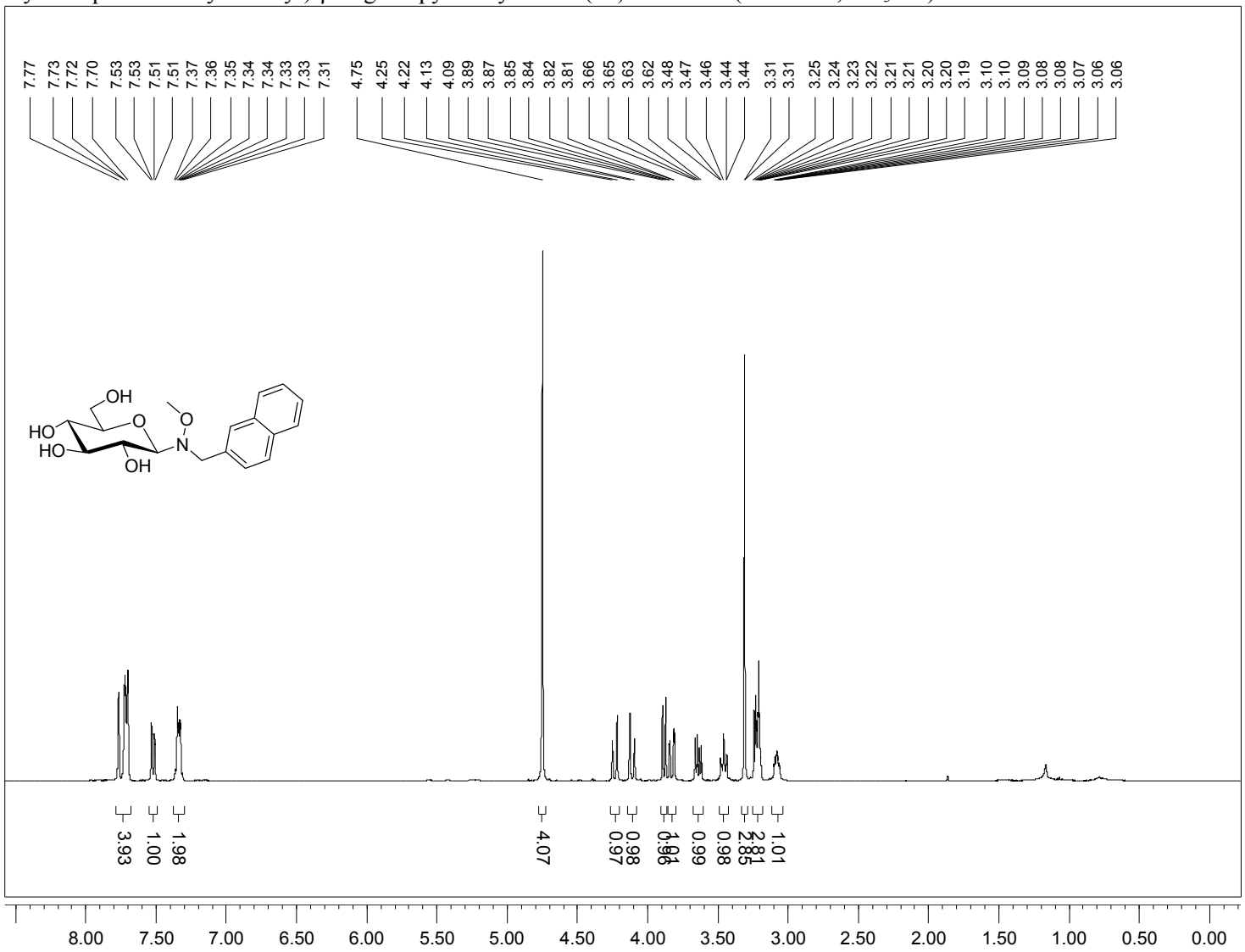
N-(*N*-ethoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (**25**): ^{1}H NMR (400 MHz, CD_3OD)



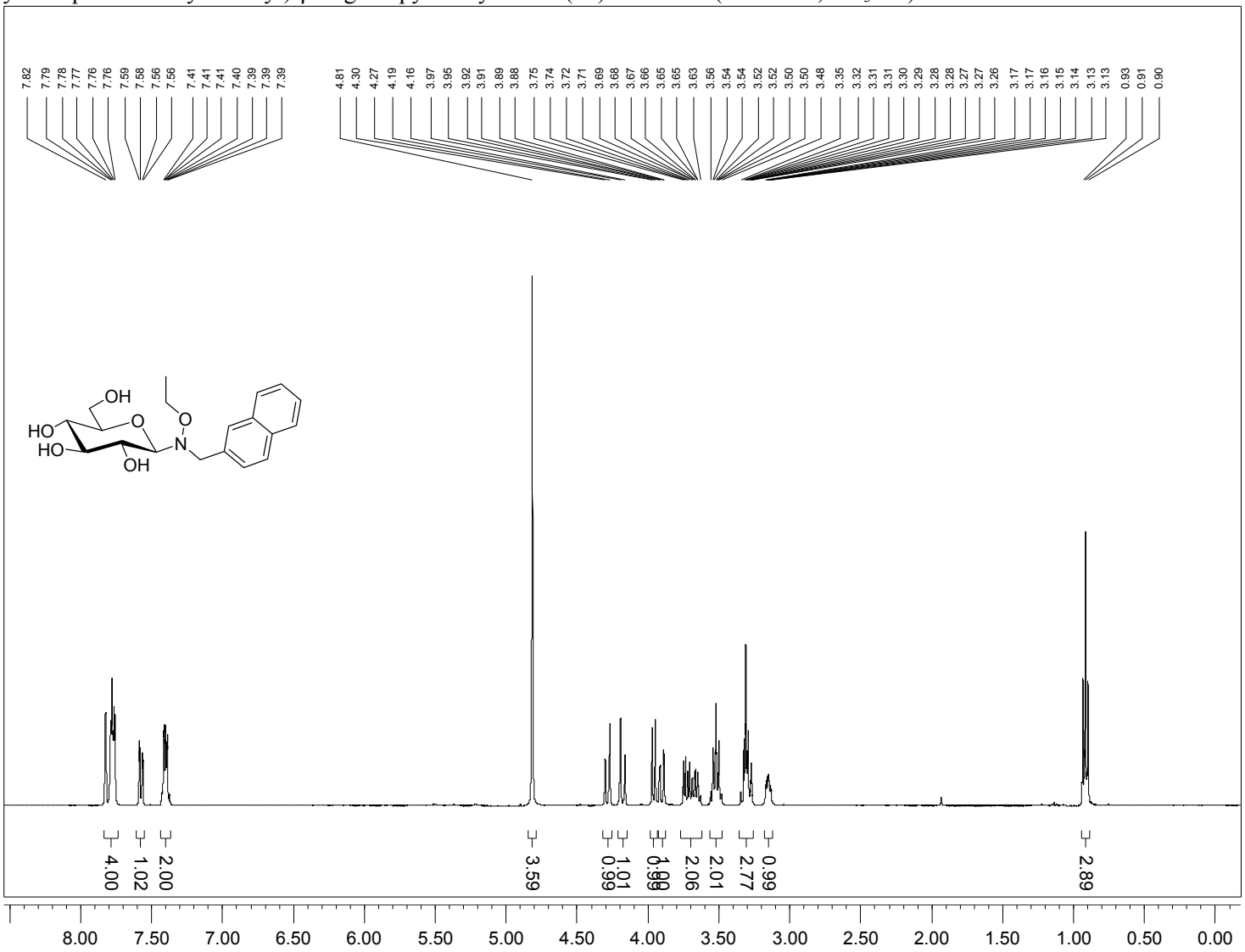
N-(*N*-benzoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (**26**): ^{1}H NMR (400 MHz, CD_3OD)



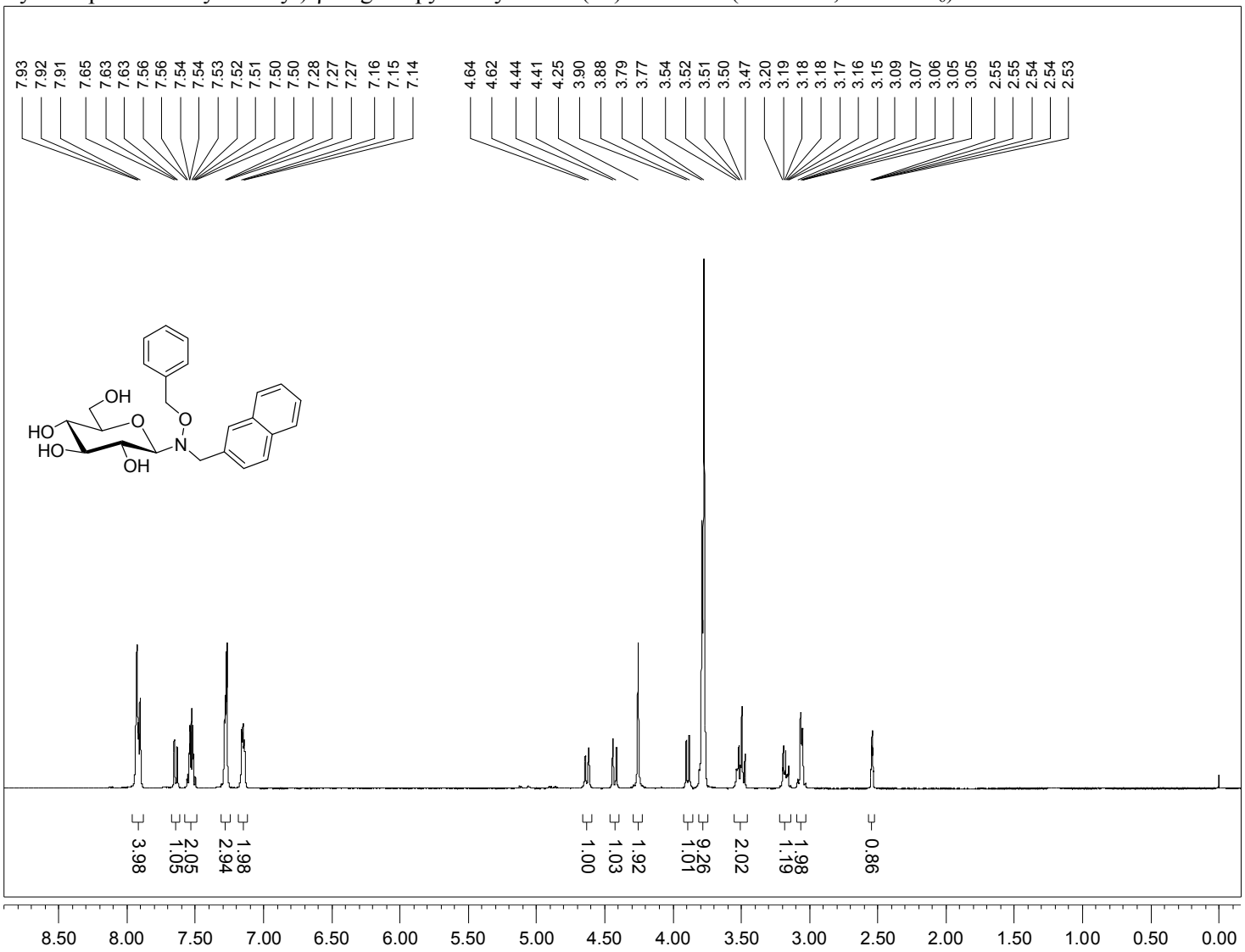
N-(*N*-methoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (**27**): ^{1}H NMR (400 MHz, CD_3OD)



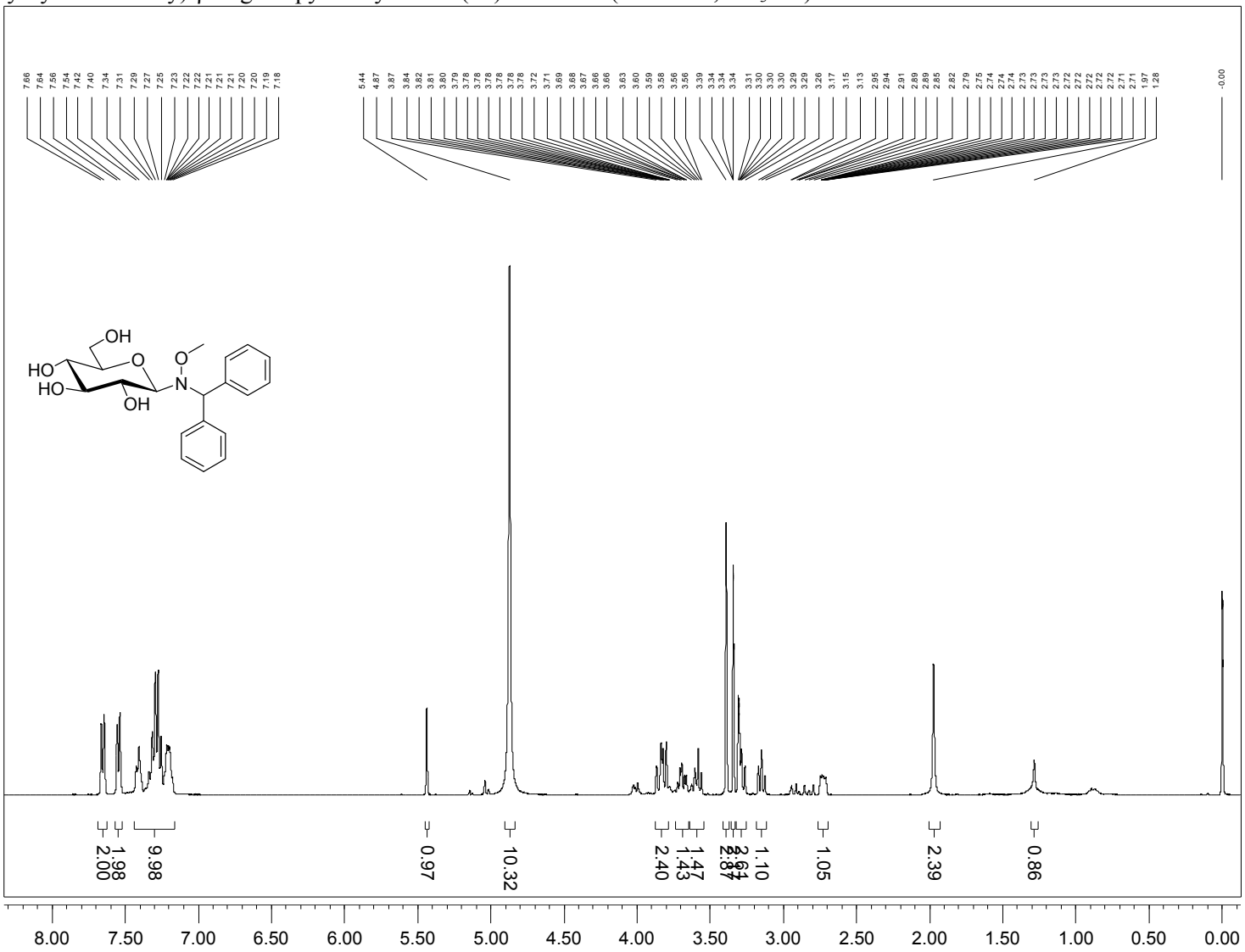
N-(*N*-ethoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (**28**): ^{1}H NMR (400 MHz, CD_3OD)



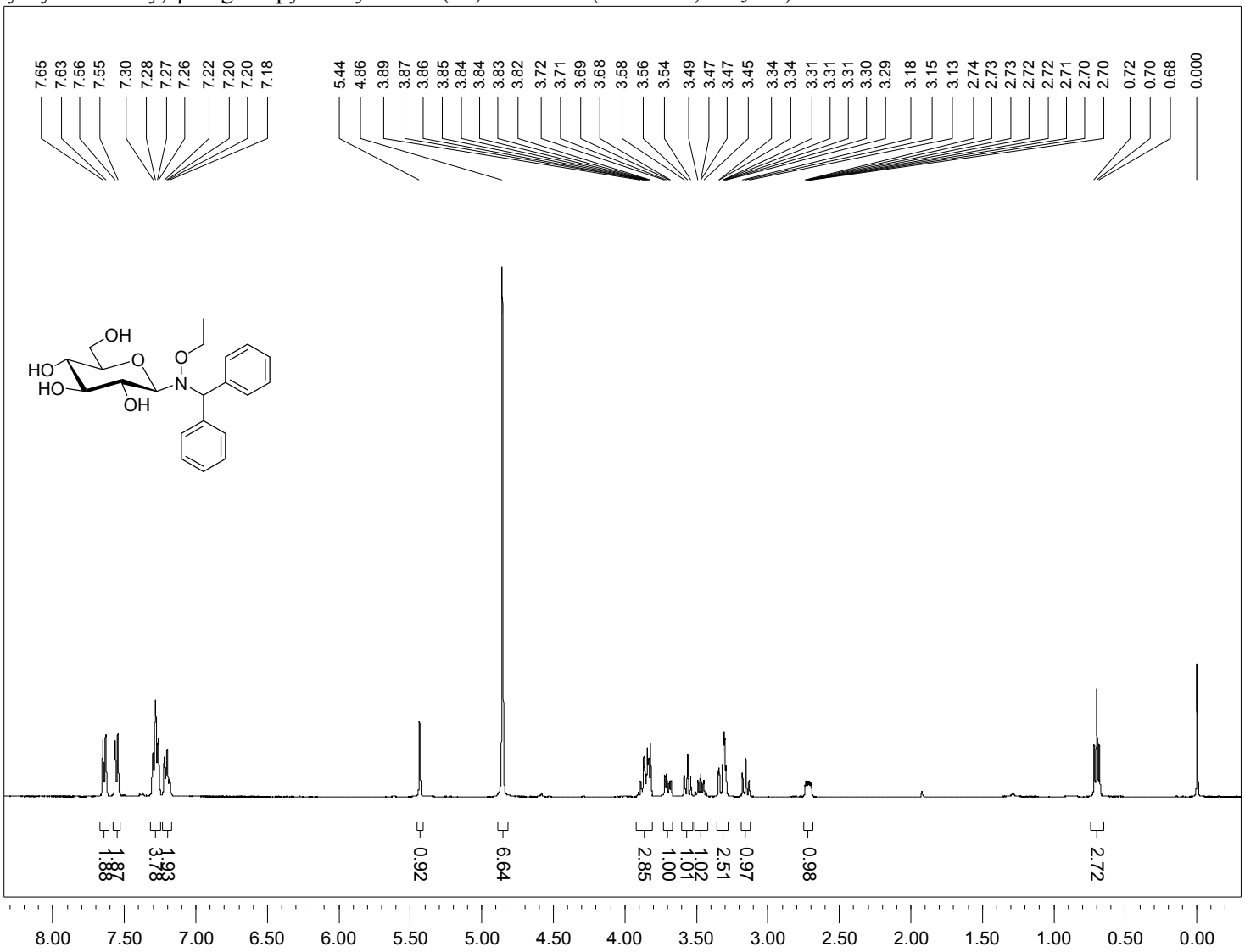
N-(*N*-benzoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (**29**): ^{1}H NMR (400 MHz, DMSO-d_6)



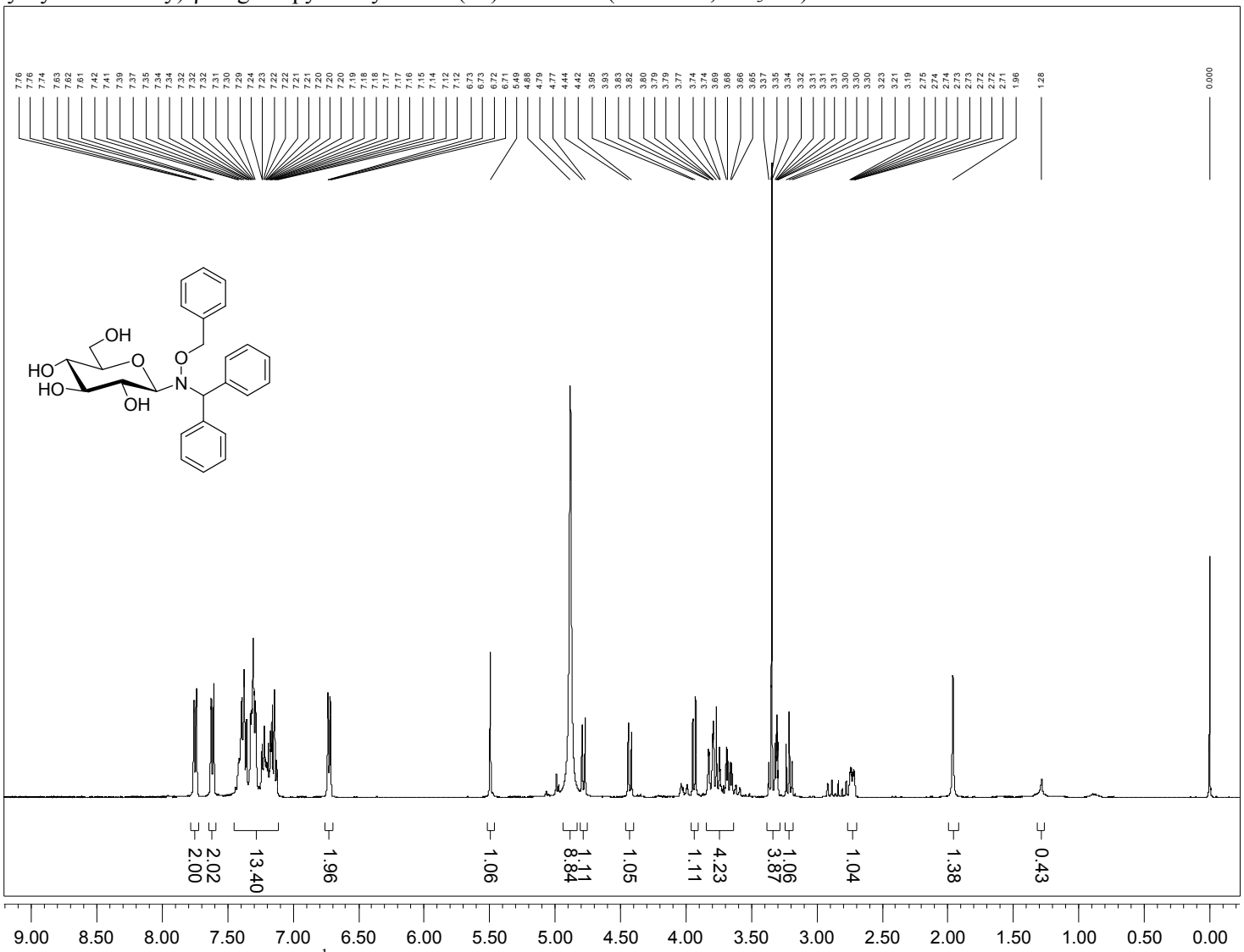
N-(*N*-benzhydryl-*N*-methoxy)-β-D-glucopyranosylamine (**30**): ^{1}H NMR (400 MHz, CD_3OD)



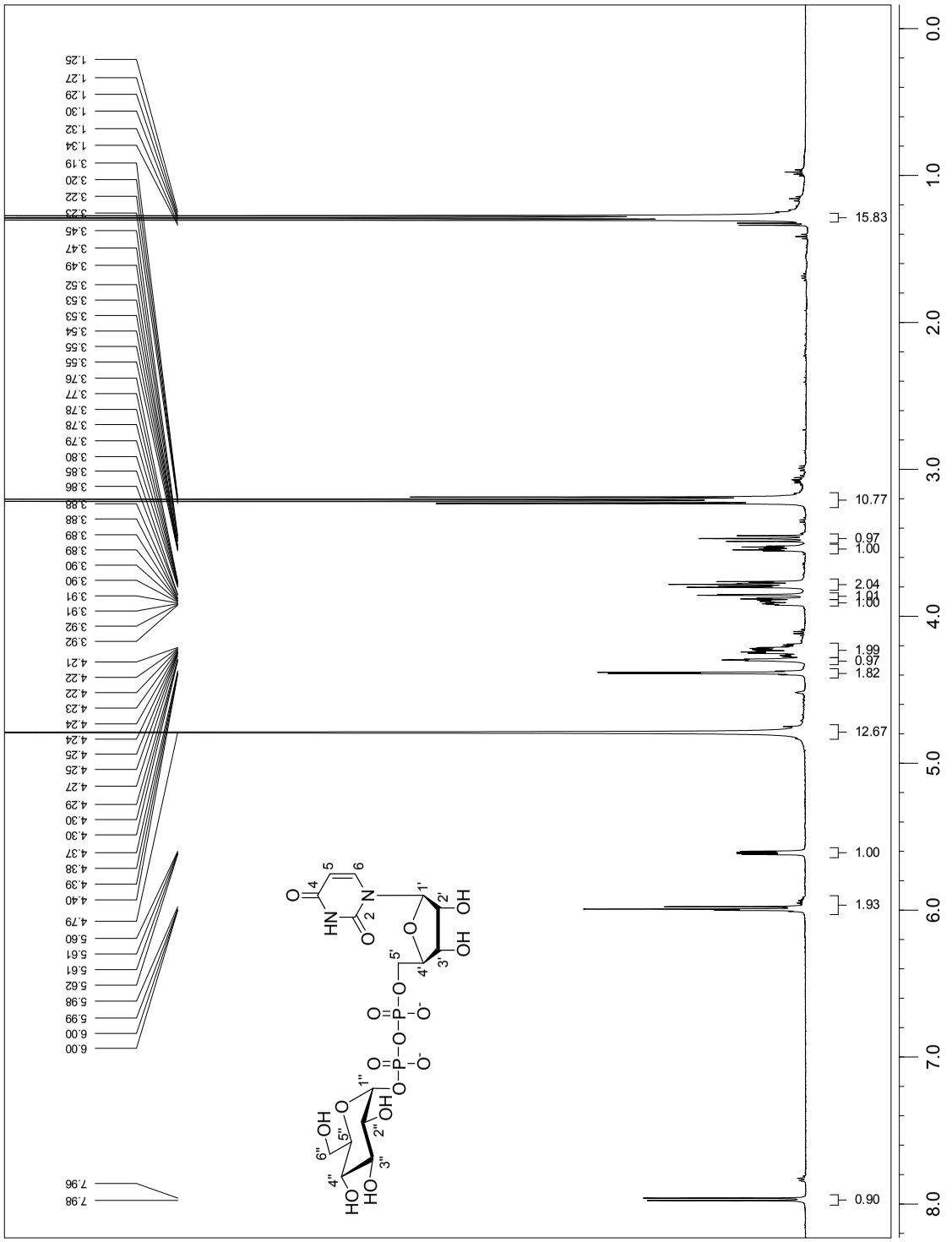
N-(*N*-benzhydryl-*N*-ethoxy)-β-D-glucopyranosylamine (**31**): ^{1}H NMR (400 MHz, CD_3OD)



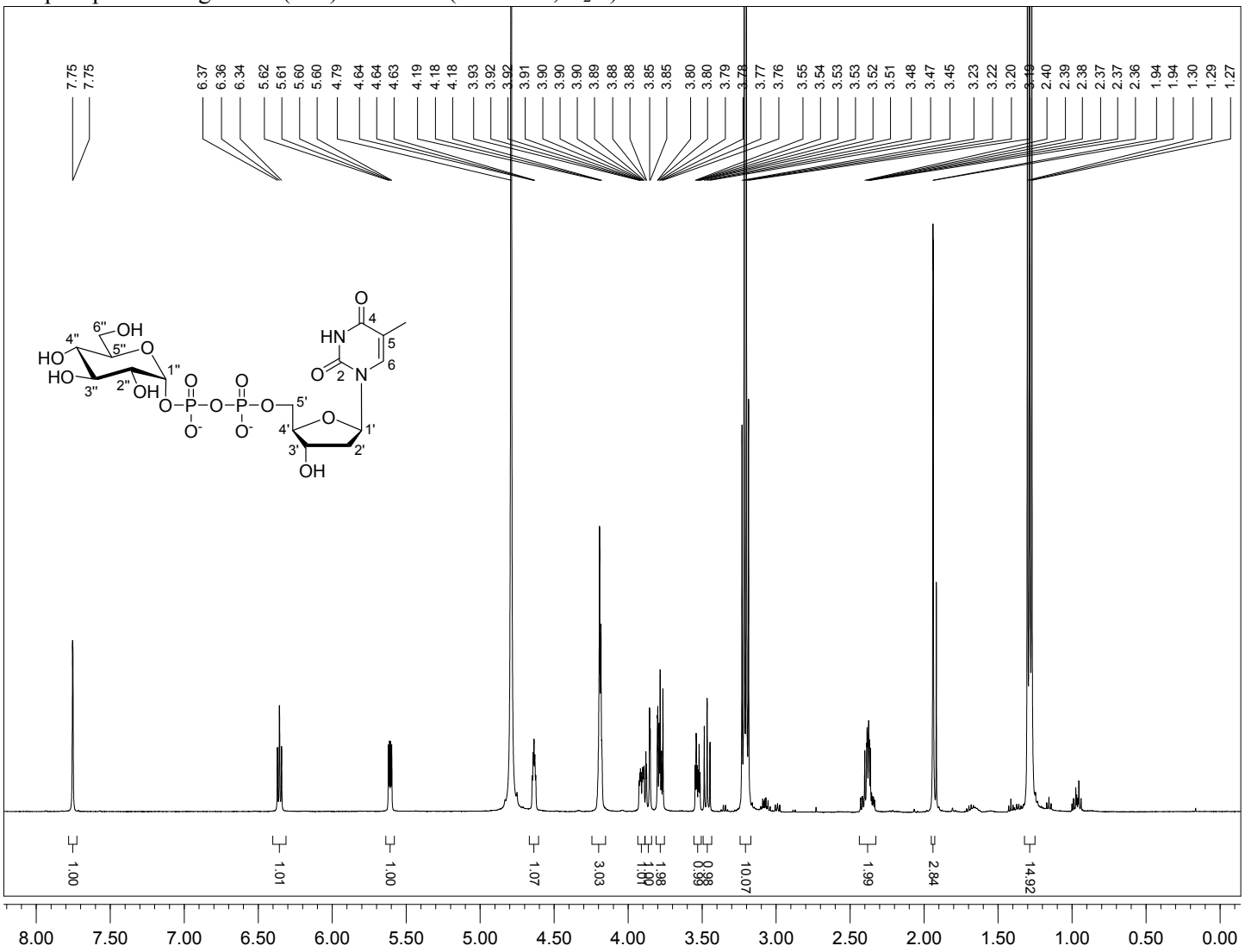
N-(*N*-benzhydryl-*N*-benzoxy)-β-D-glucopyranosylamine (**32**): ^{1}H NMR (400 MHz, CD_3OD)



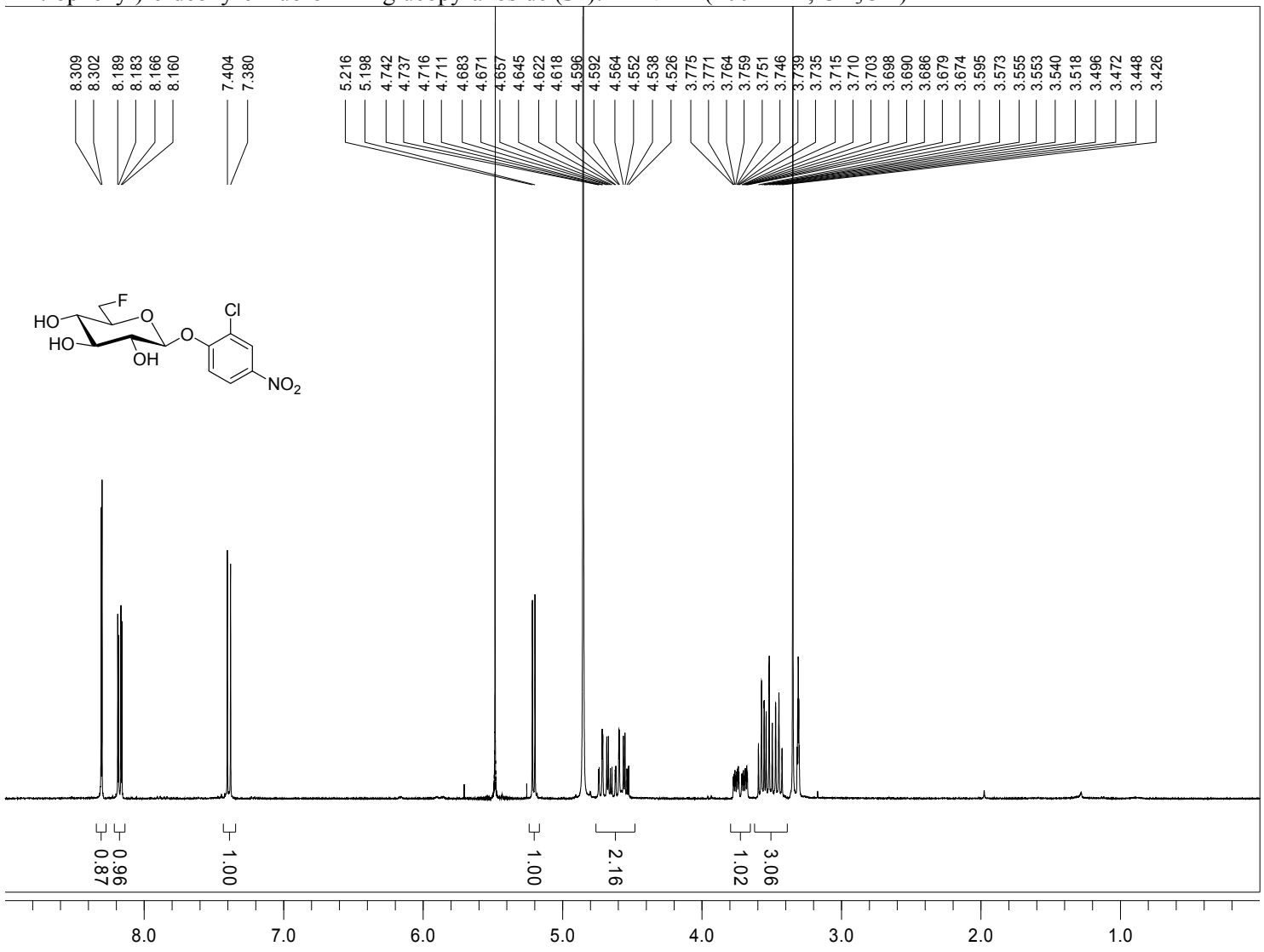
Uridine 5'-diphosphate α-D-glucose (**33a**): ^{1}H NMR (500 MHz, D_2O)


1.25
1.27
1.29
1.30
1.32
1.34
3.19
3.20
3.22
3.23
3.45
3.47
3.49
3.52
3.53
3.53
3.54
3.55
3.55
3.76
3.77
3.78
3.78
3.79
3.80
3.85
3.86
3.88
3.88
3.89
3.89
3.90
3.90
3.91
3.91
3.92
3.92
4.21
4.22
4.22
4.23
4.24
4.24
4.25
4.25
4.27
4.29
4.30
4.30
4.37
4.38
4.39
4.40
4.79
5.60
5.61
5.61
5.62
5.98
5.99
6.00
6.00
7.96
7.98
15.83
10.77
0.97
1.00
2.04
1.01
1.00
1.99
0.97
1.82
12.67
1.00
1.93
0.90
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0

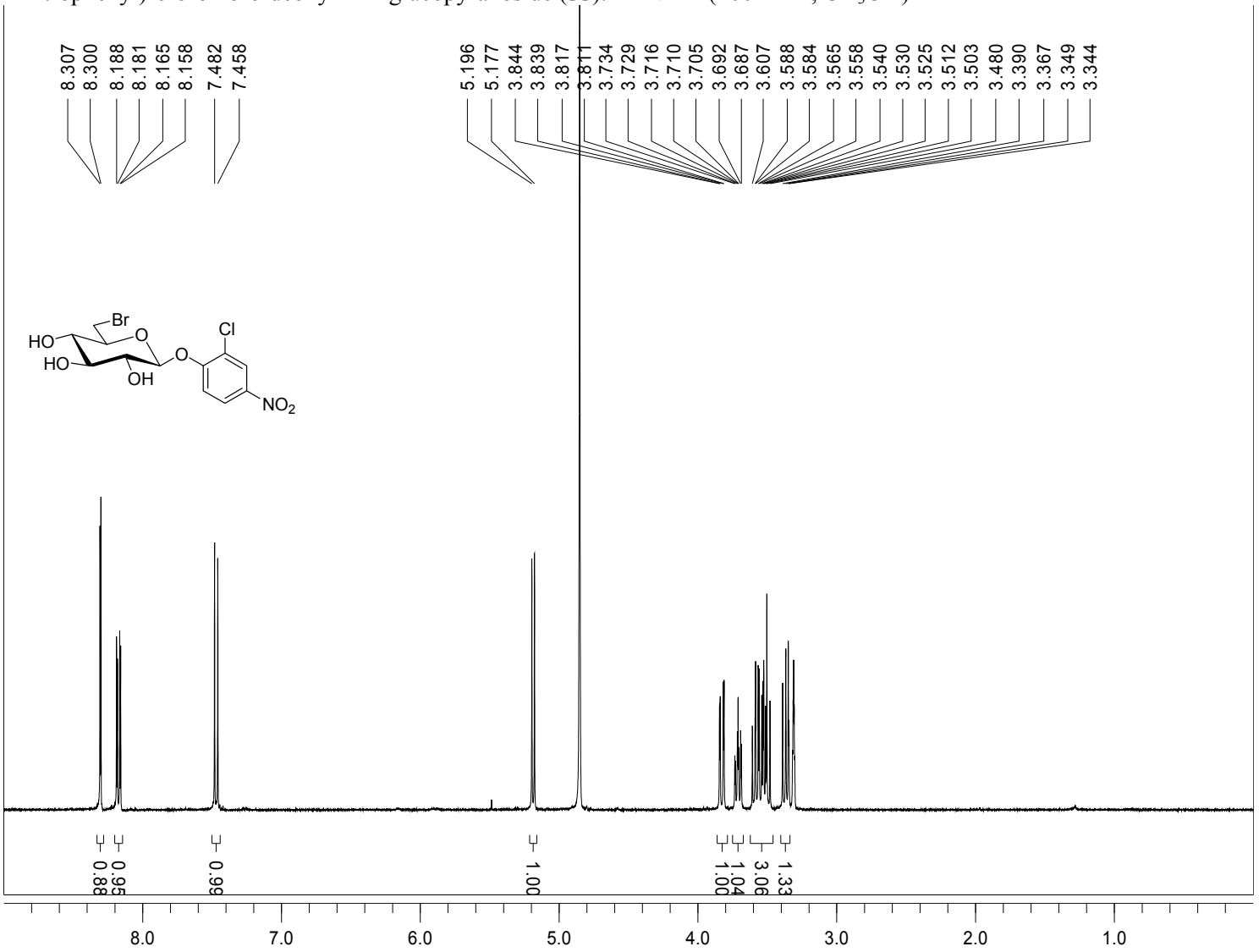
Thymidine 5’-diphosphate α-D-glucose (**33b**): ^{1}H NMR (500 MHz, D_2O)



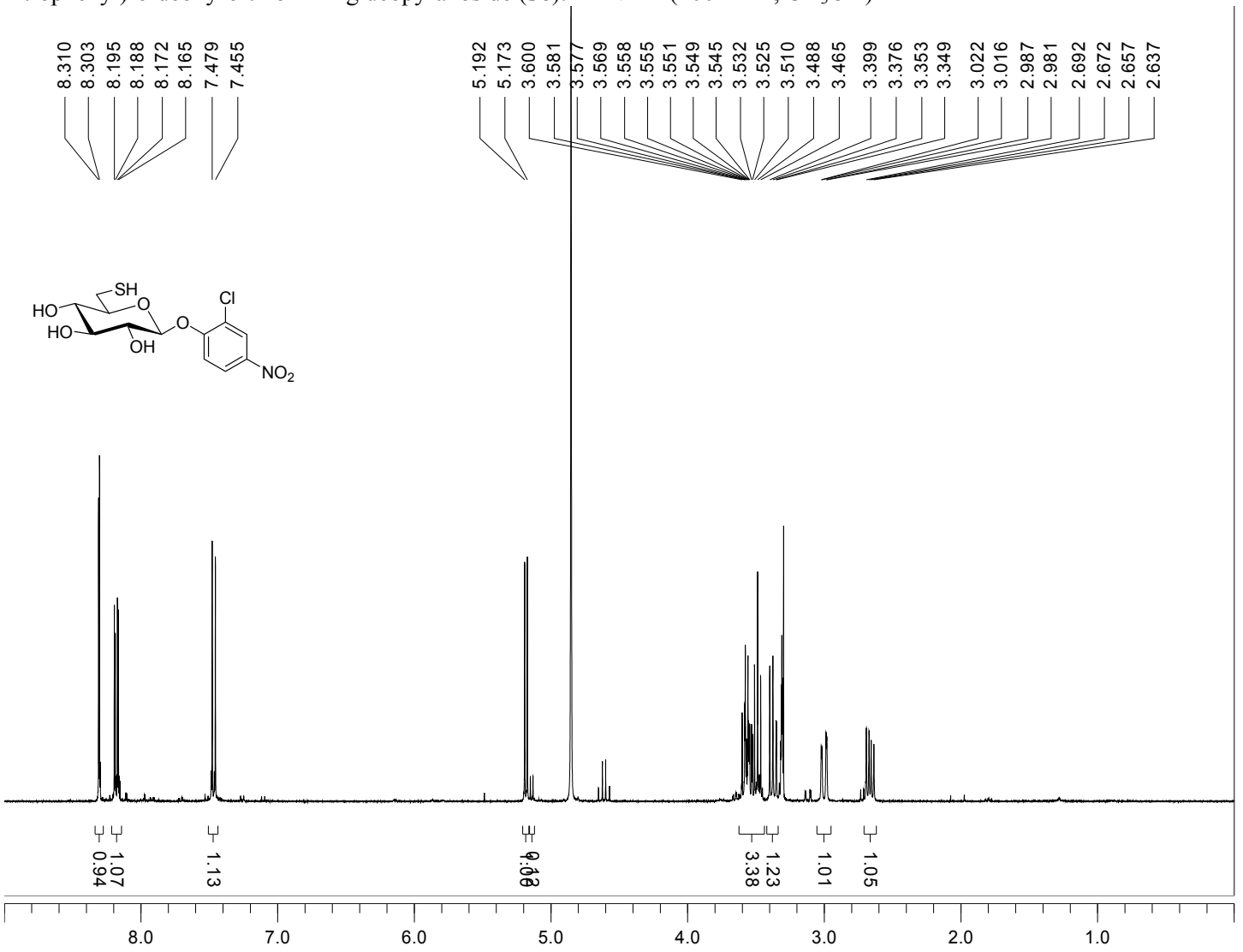
(2-chloro-4-nitrophenyl)-6-deoxy-6-fluoro-□-D-glucopyranoside (**34**): ^{1}H NMR (400 MHz, CD_3OD)



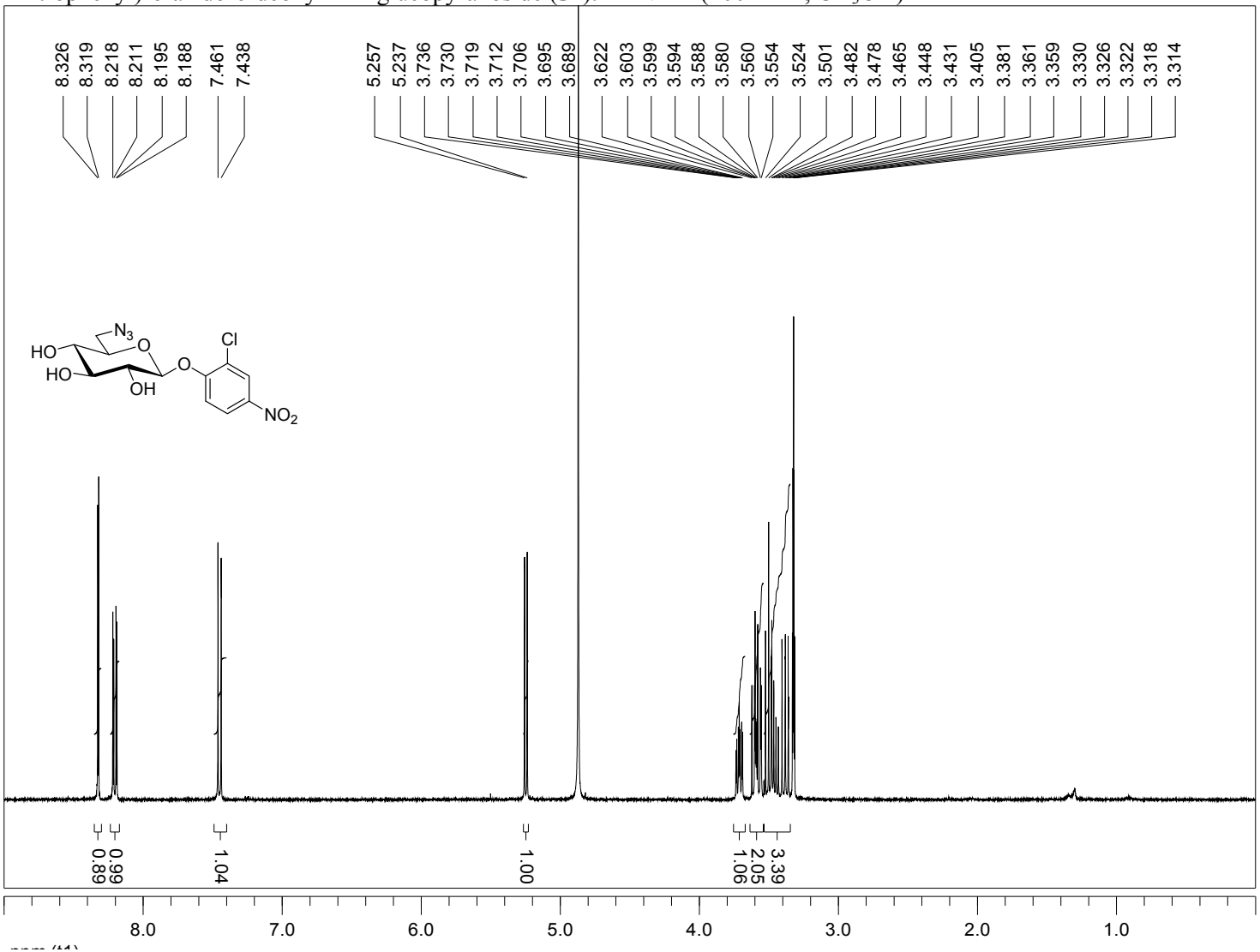
(2-chloro-4-nitrophenyl)-6-bromo-6-deoxy-□-D-glucopyranoside (**35**): ^{1}H NMR (400 MHz, CD_3OD)



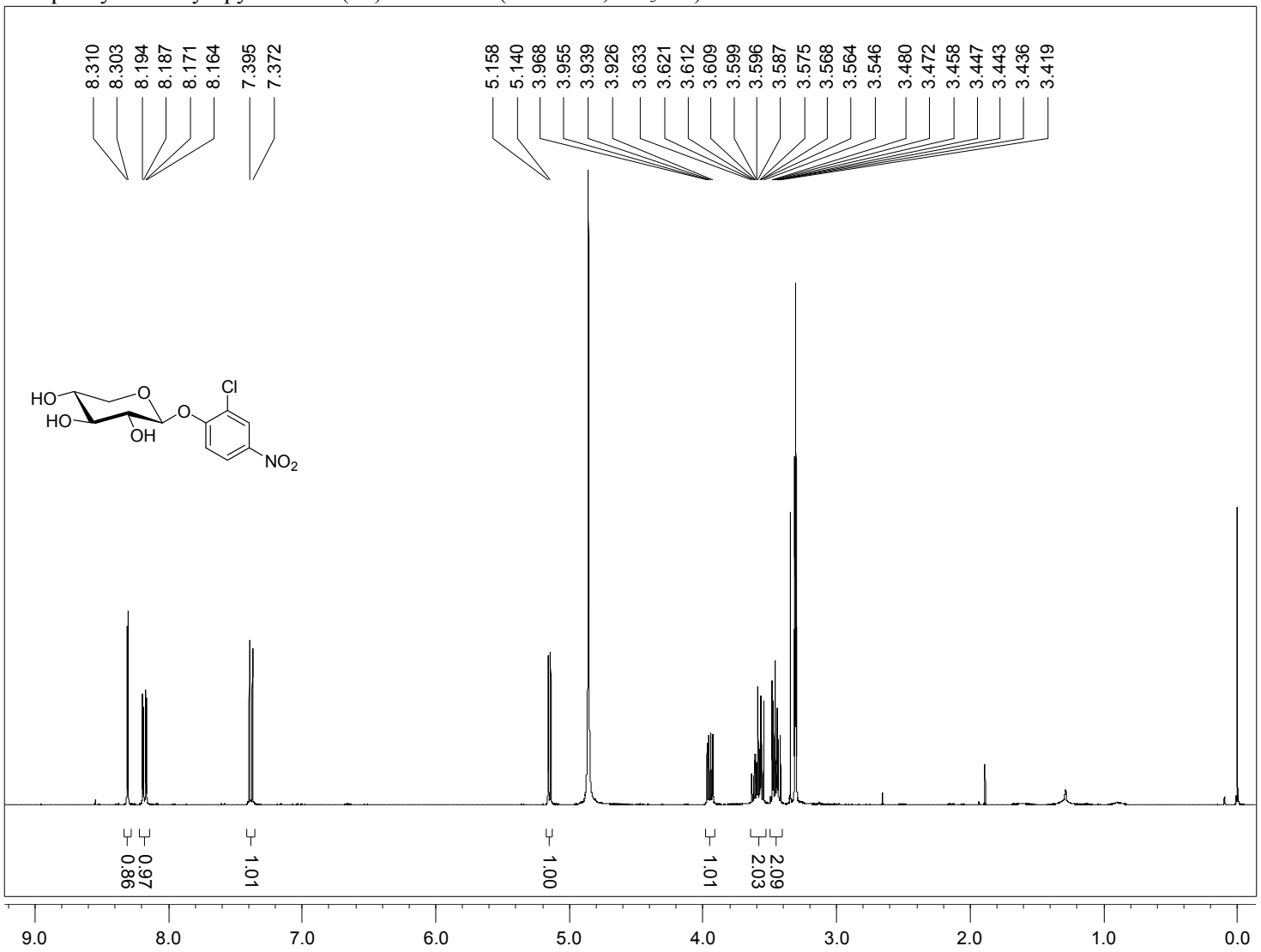
(2-chloro-4-nitrophenyl)-6-deoxy-6-thio-□-D-glucopyranoside (**36**): ^{1}H NMR (400 MHz, CD_3OD)



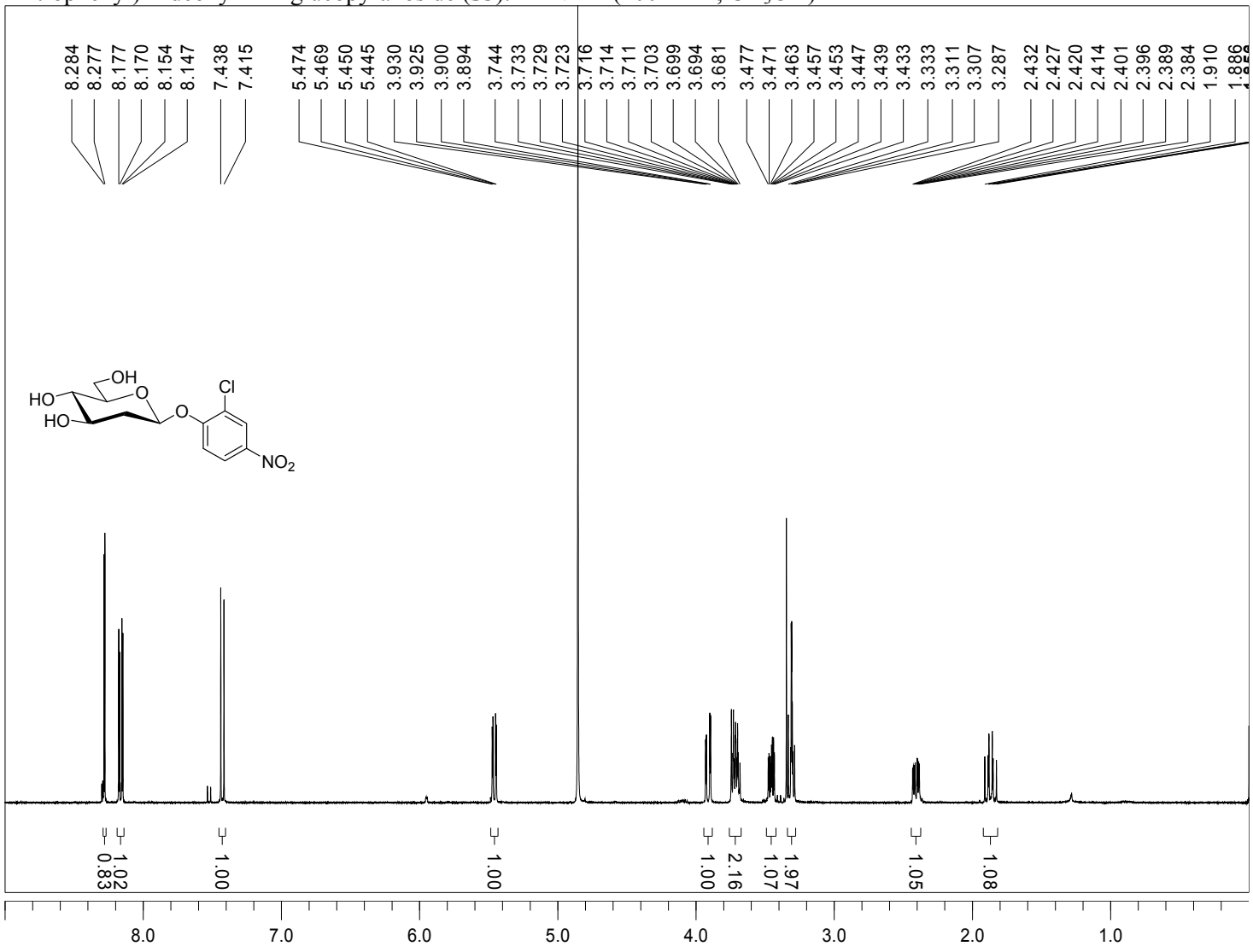
(2-chloro-4-nitrophenyl)-6-azido-6-deoxy-□-D-glucopyranoside (**37**): ^{1}H NMR (400 MHz, CD_3OD)



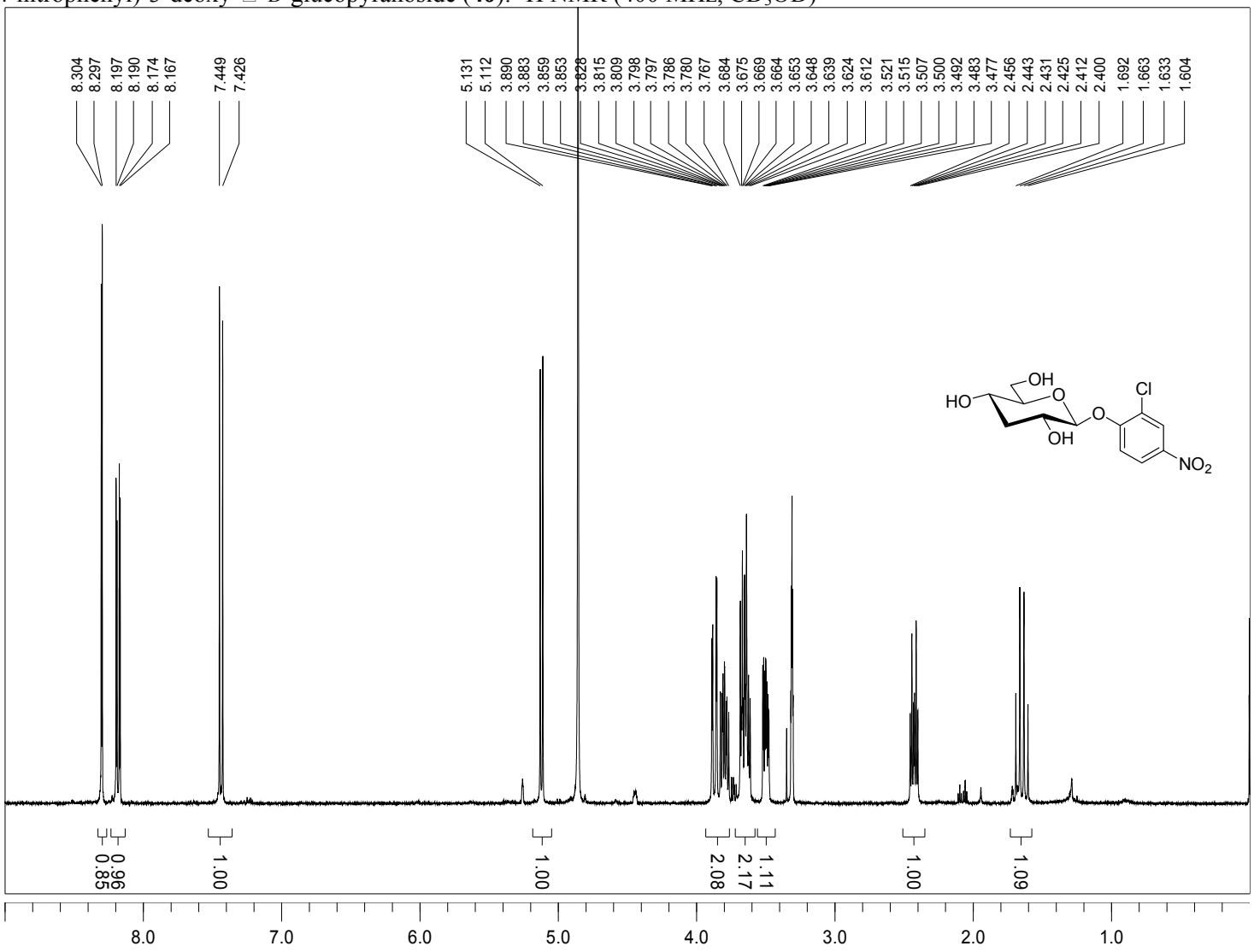
2-chloro-4-nitrophenyl-□-D-xylopyranoside (**38**): ^{1}H NMR (400 MHz, CD_3OD)



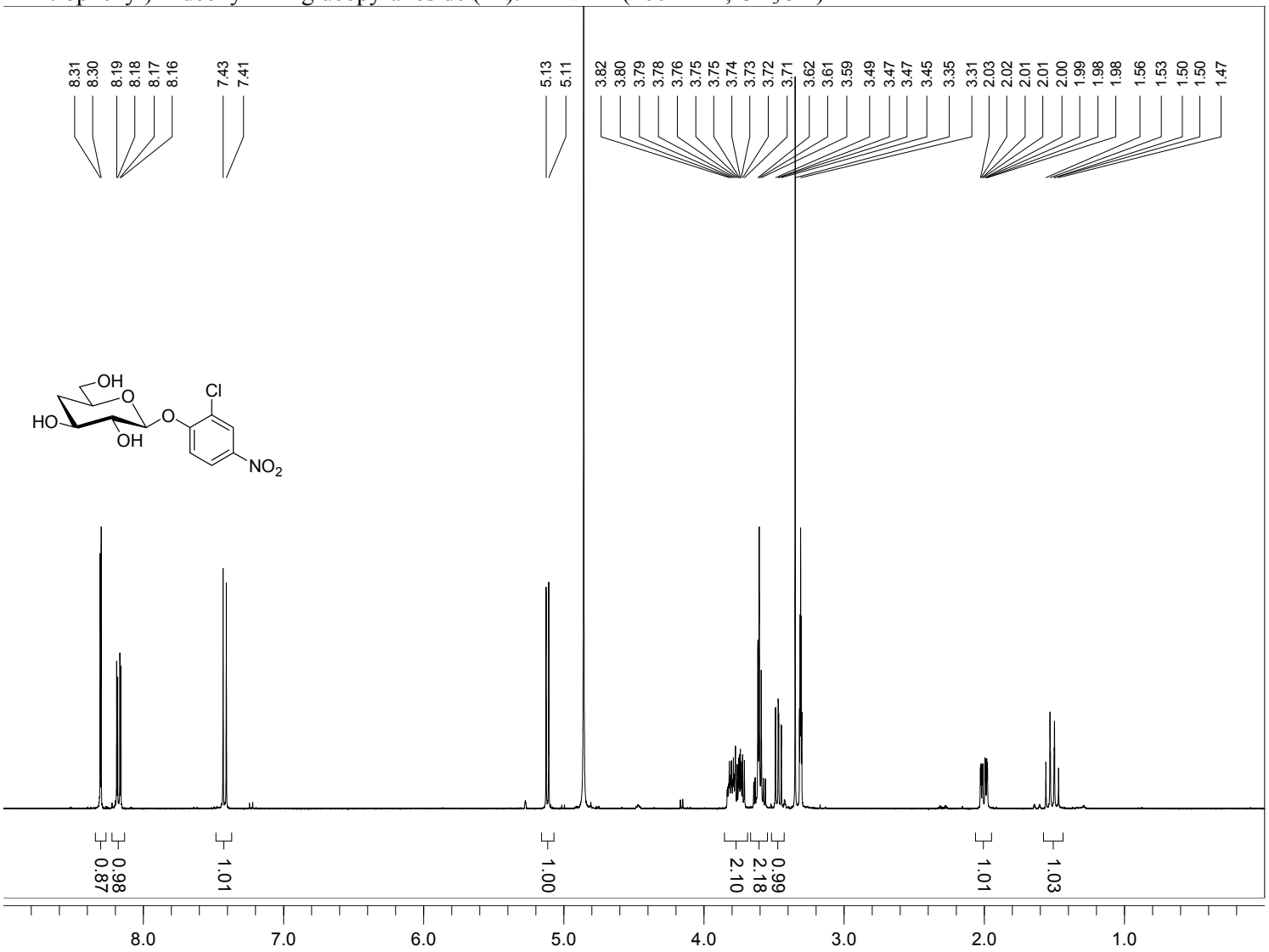
(2-chloro-4-nitrophenyl)-2-deoxy-□-D-glucopyranoside (**39**): ^{1}H NMR (400 MHz, CD_3OD)



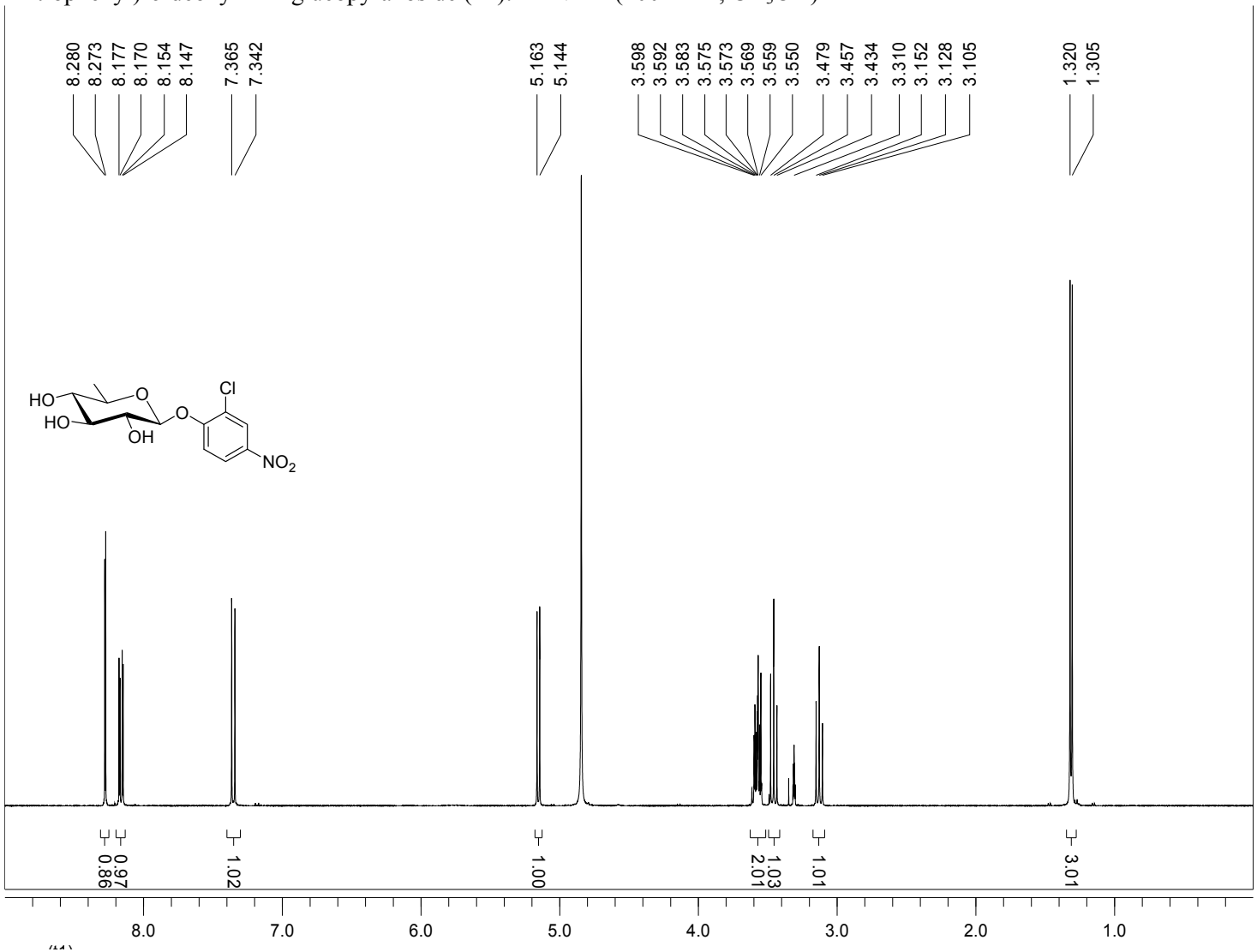
(2-chloro-4-nitrophenyl)-3-deoxy-□-D-glucopyranoside (**40**): ^{1}H NMR (400 MHz, CD_3OD)



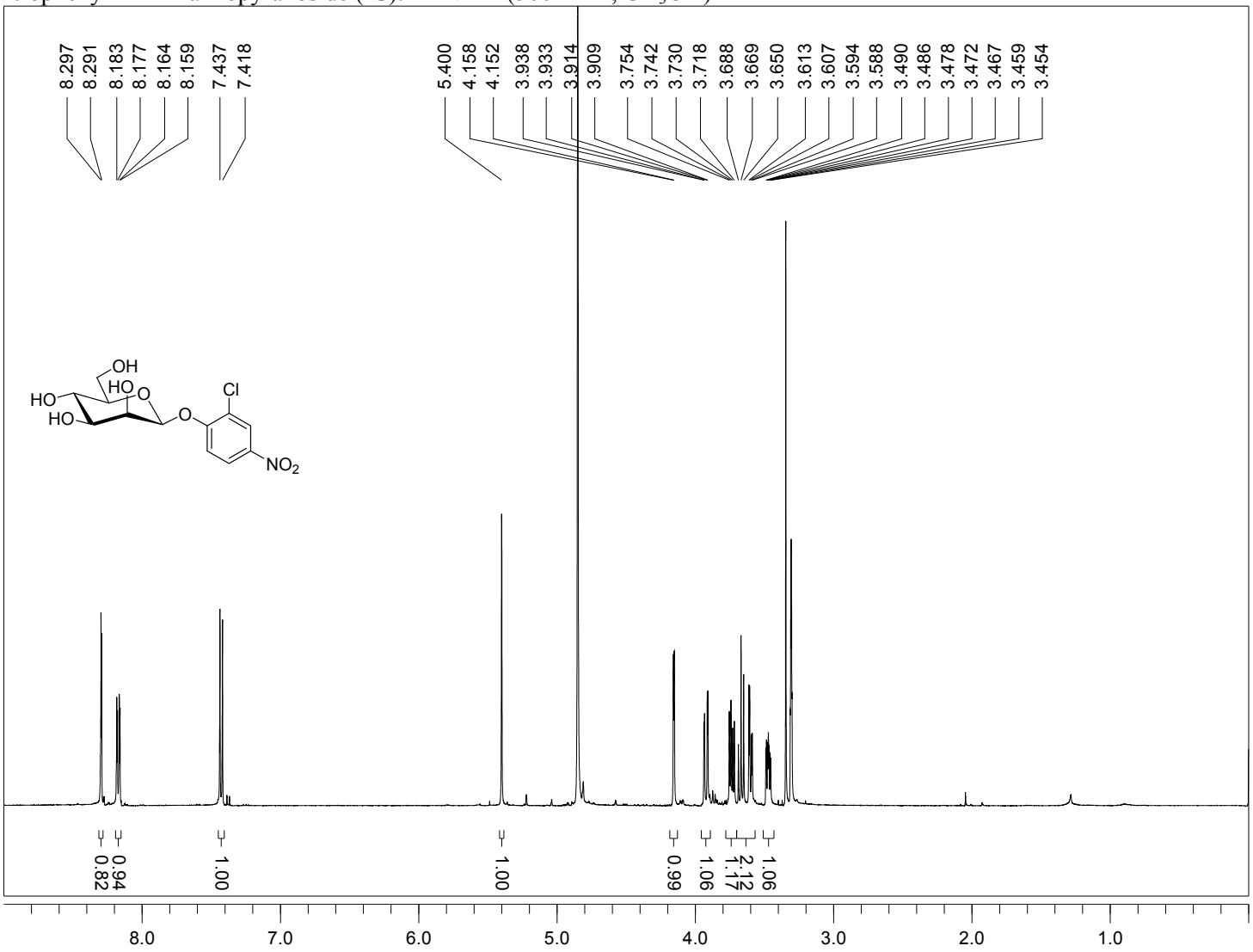
(2-chloro-4-nitrophenyl)-4-deoxy-□-D-glucopyranoside (**41**): ^{1}H NMR (400 MHz, CD_3OD)



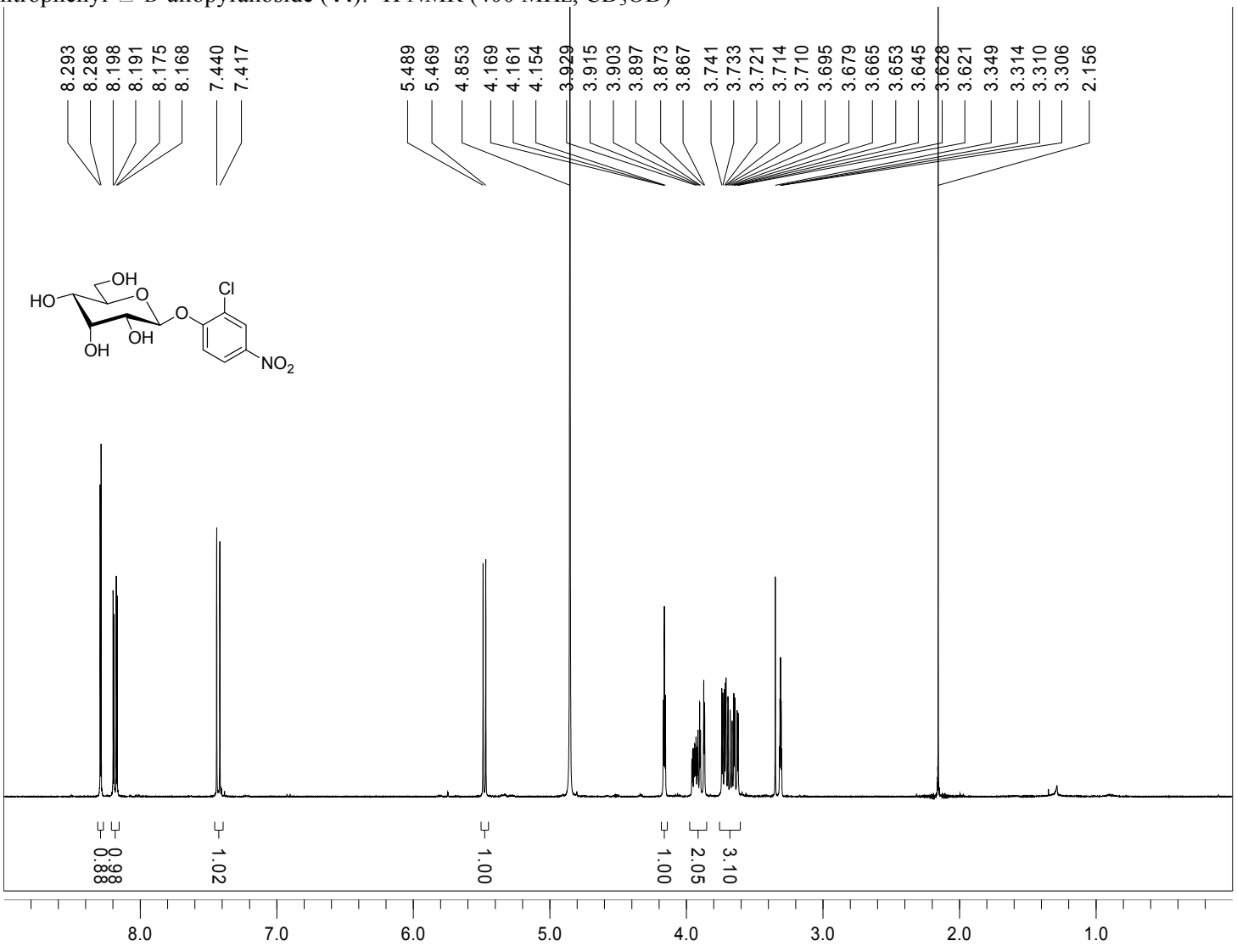
(2-chloro-4-nitrophenyl)-6-deoxy-□-D-glucopyranoside (**42**): ^{1}H NMR (400 MHz, CD_3OD)



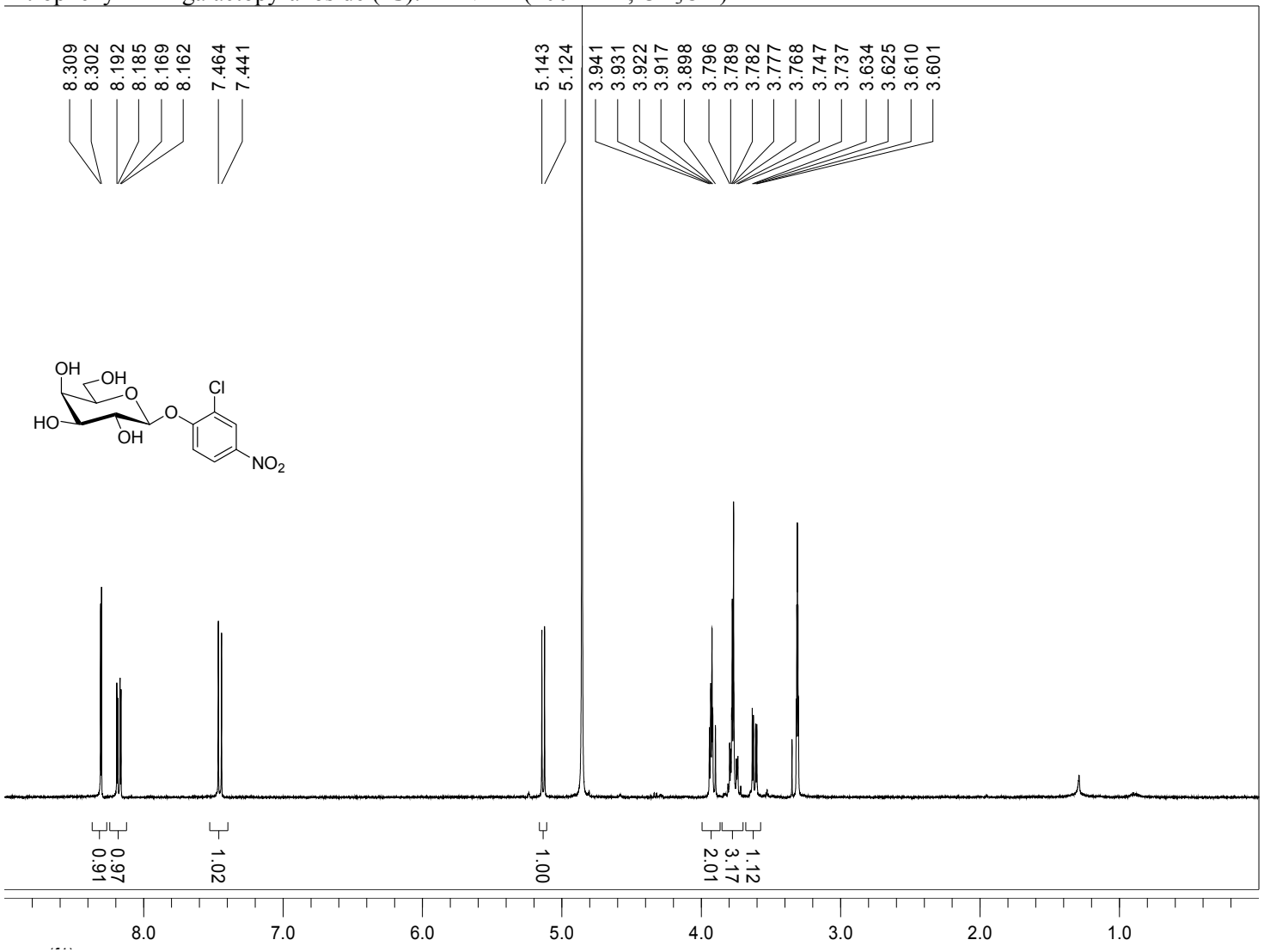
2-chloro-4-nitrophenyl-□-D-mannopyranoside (**43**): ^{1}H NMR (500 MHz, CD_3OD)



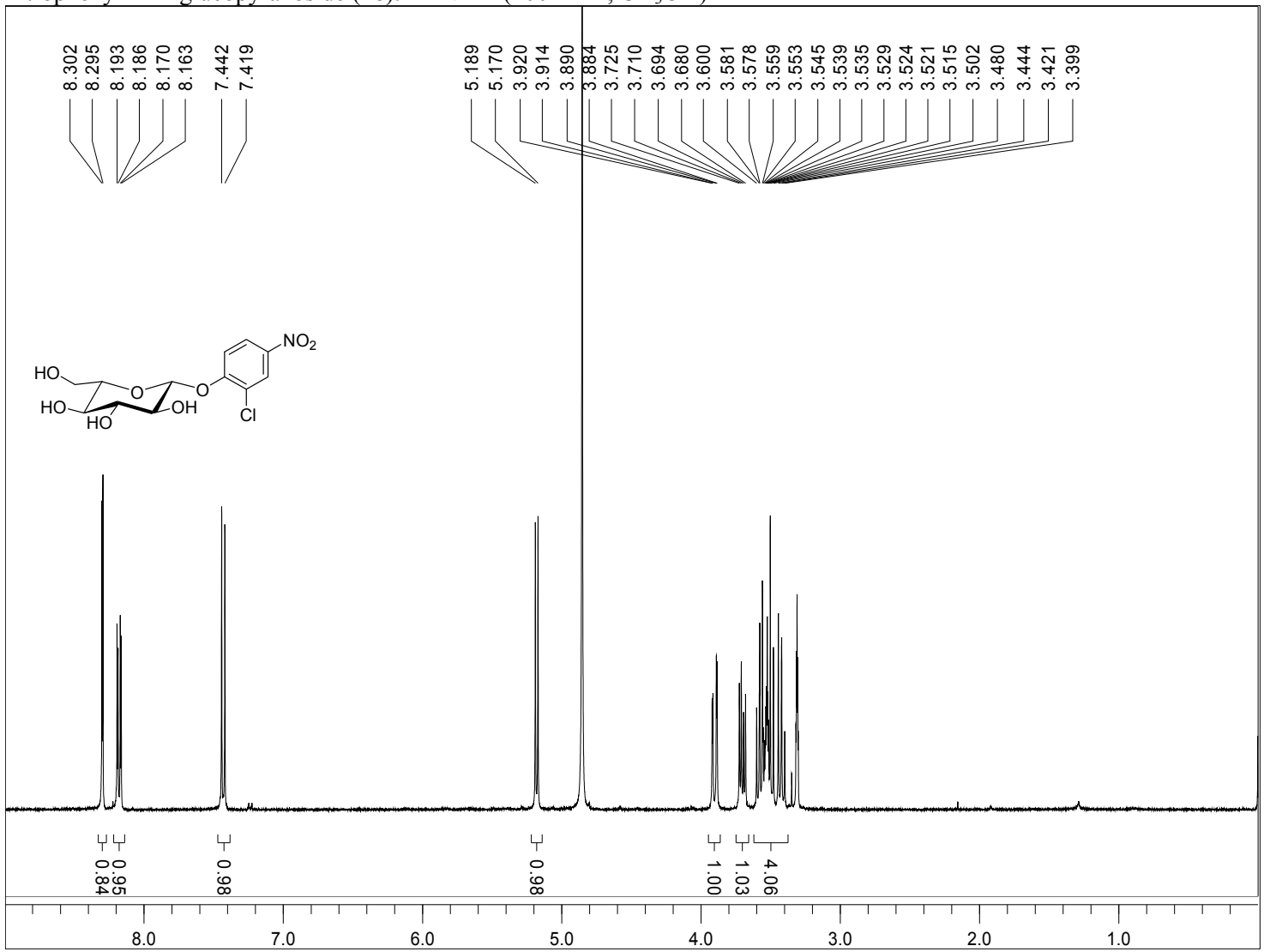
2-chloro-4-nitrophenyl-□-D-allopyranoside (**44**): ^{1}H NMR (400 MHz, CD_3OD)



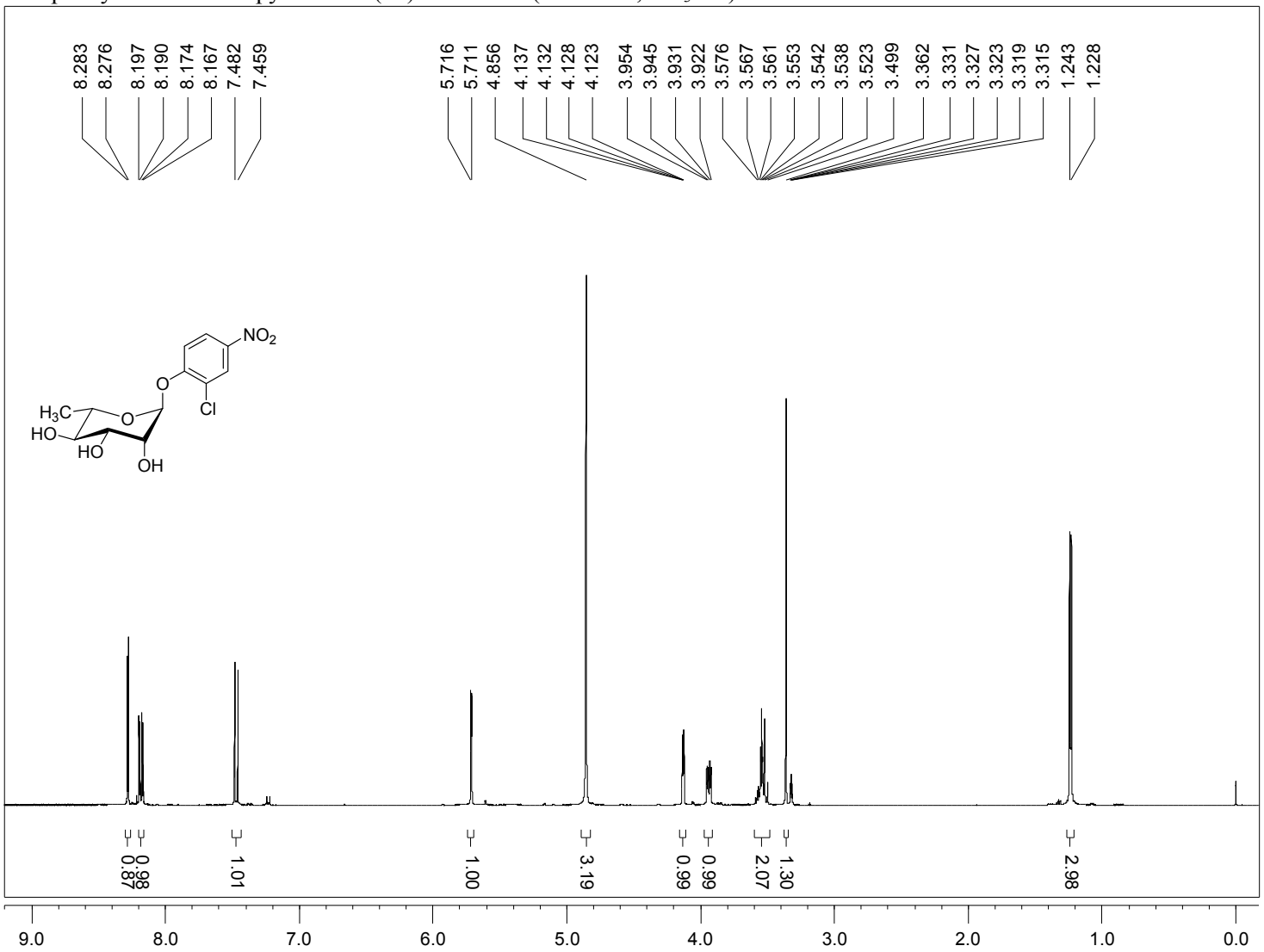
2-chloro-4-nitrophenyl-□-D-galactopyranoside (**45**): ^{1}H NMR (400 MHz, CD_3OD)



2-chloro-4-nitrophenyl-□-L-glucopyranoside (**46**): ^{1}H NMR (400 MHz, CD_3OD)

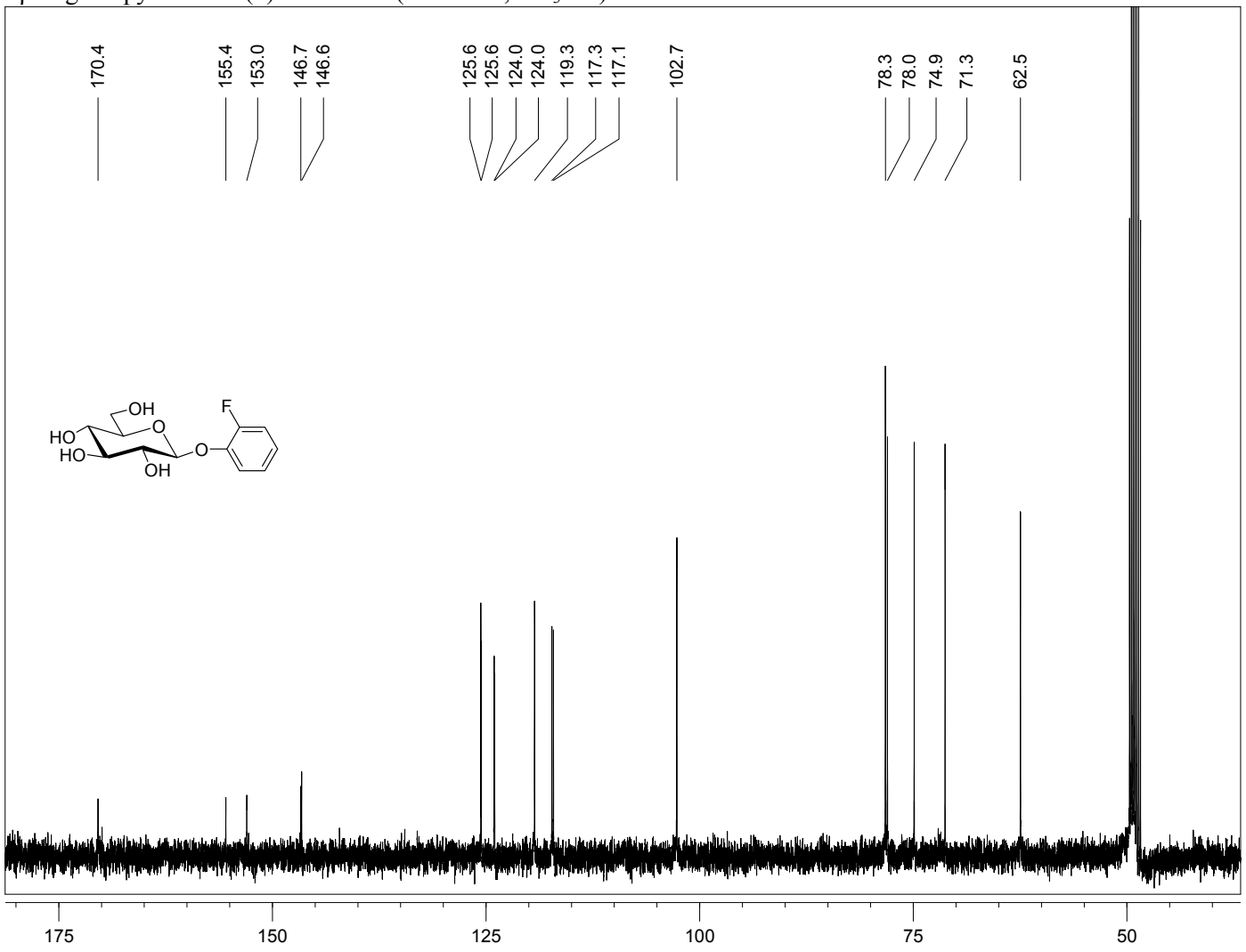


2-chloro-4-nitrophenyl-□-L-rhamnopyranoside (**47**): ^{1}H NMR (400 MHz, CD_3OD)

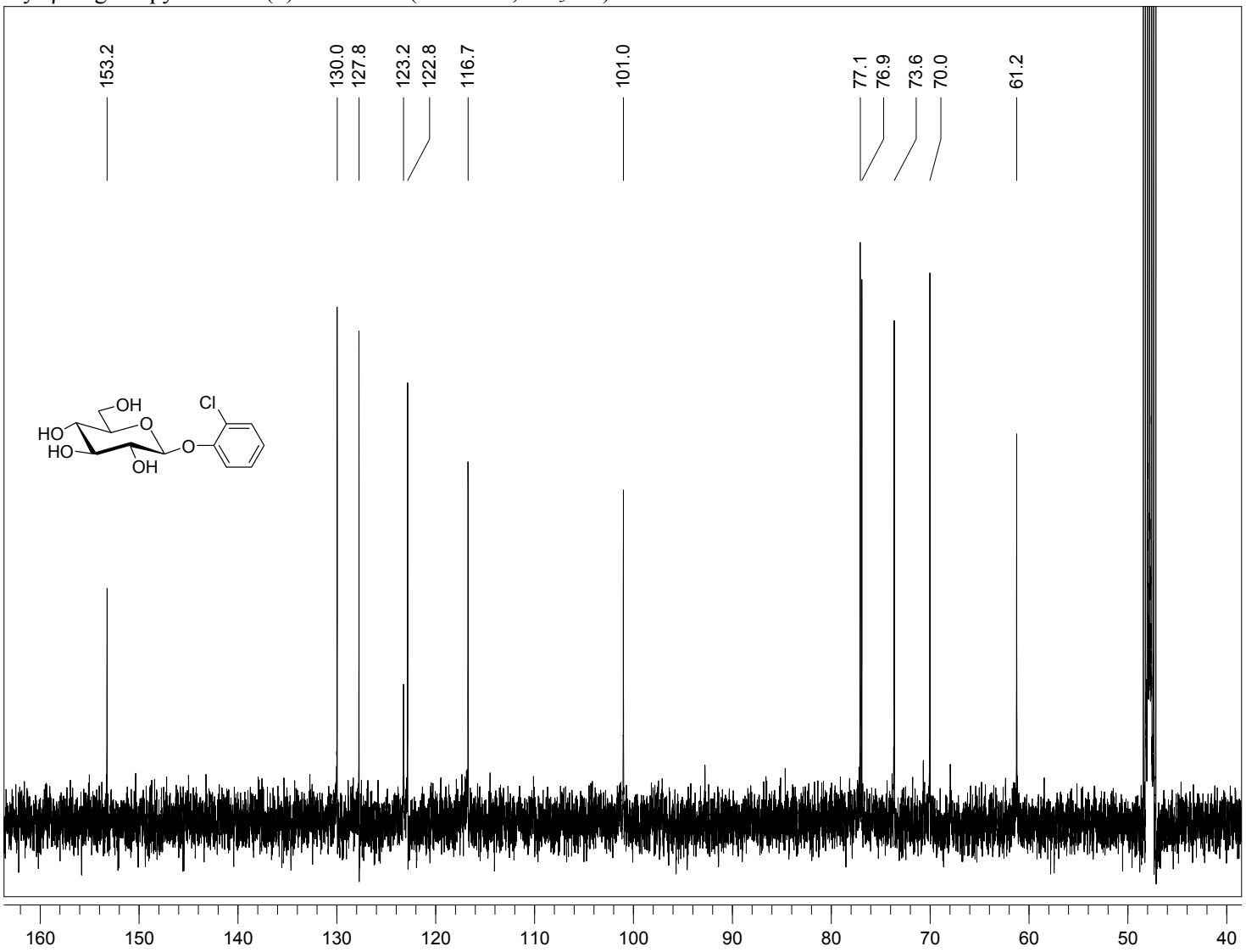


A2.3. ^{13}C NMR Spectra

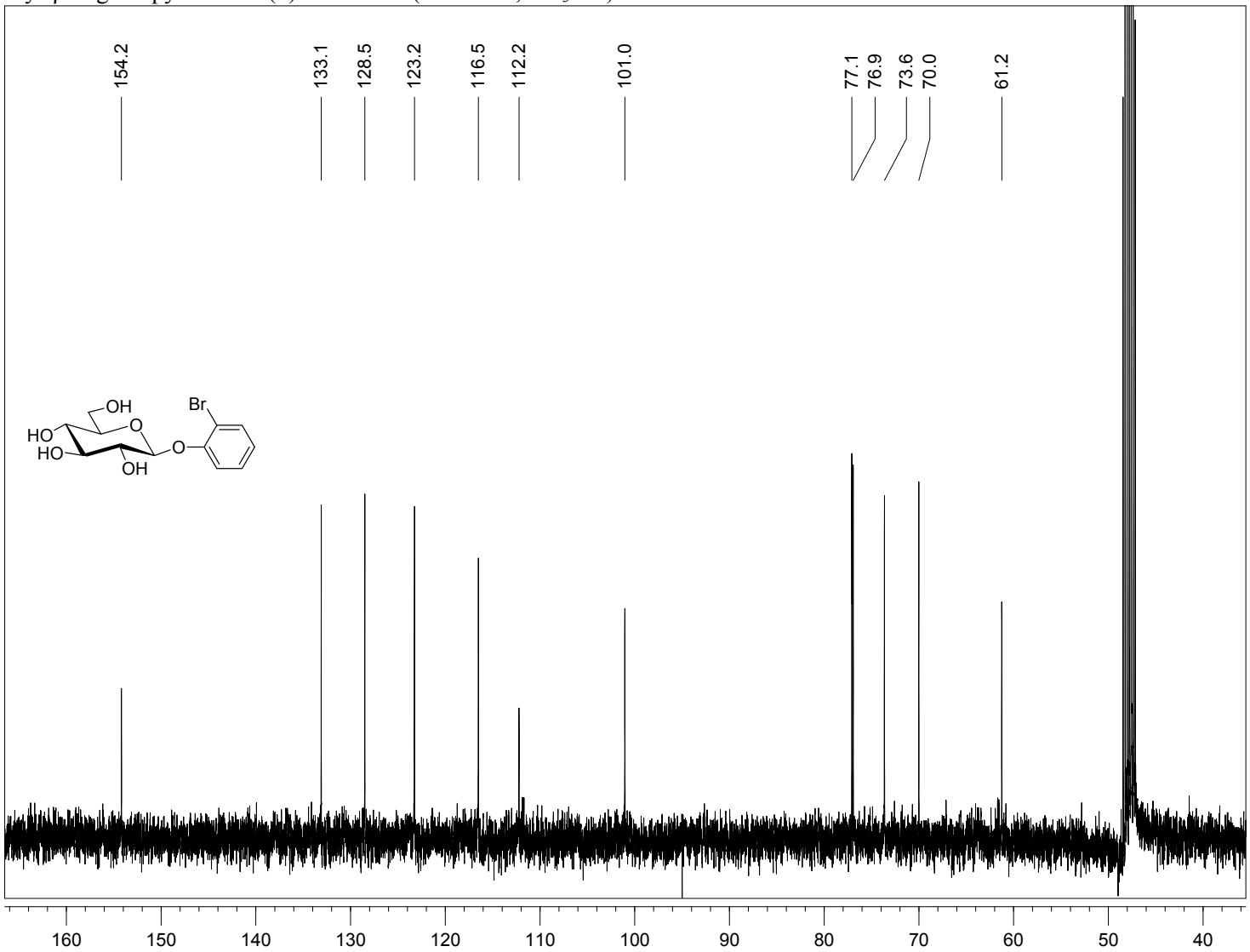
2-fluorophenyl-β-D-glucopyranoside (**3**): ^{13}C NMR (100 MHz, CD_3OD)



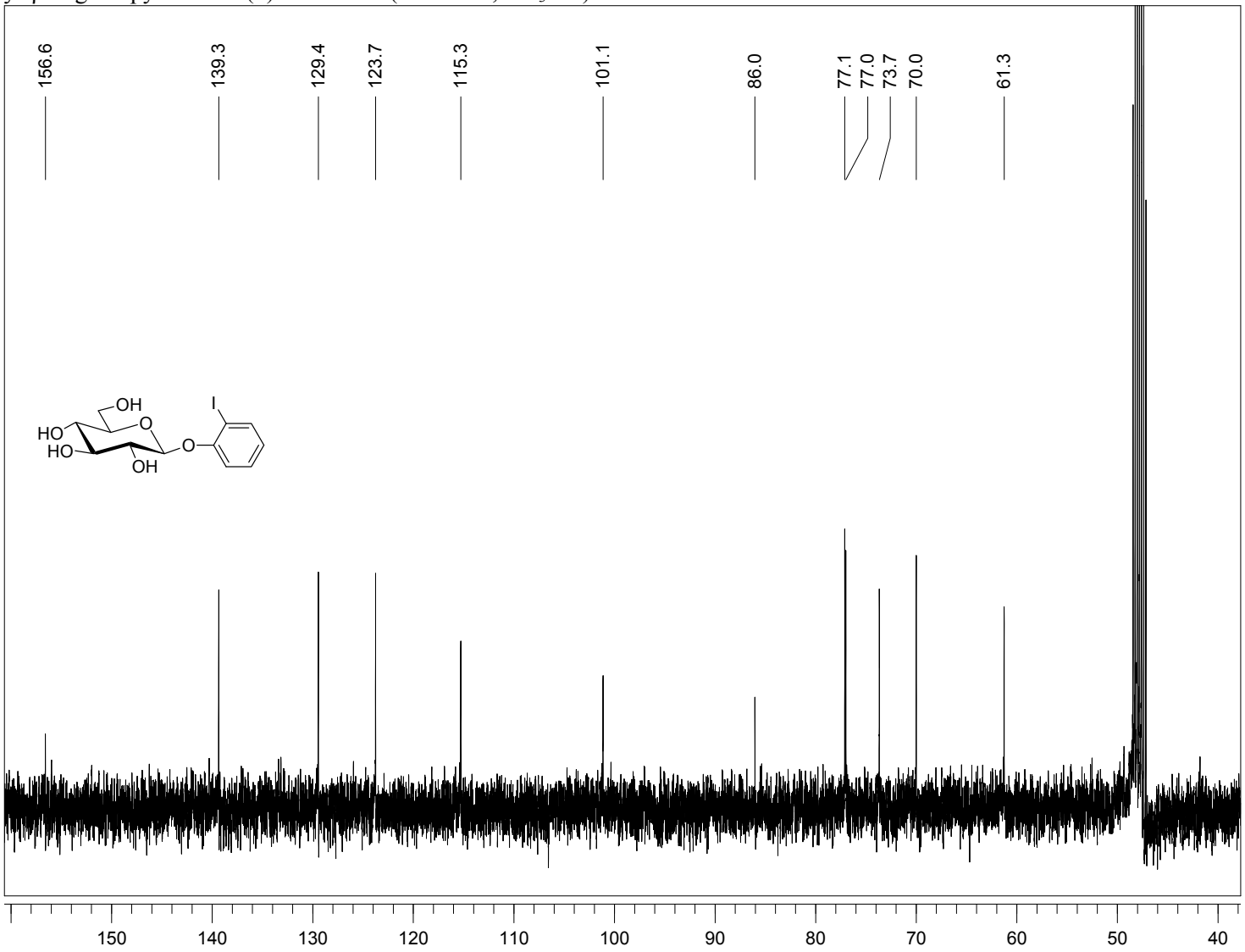
2-chlorophenyl-β-D-glucopyranoside (**4**): ^{13}C NMR (100 MHz, CD_3OD)



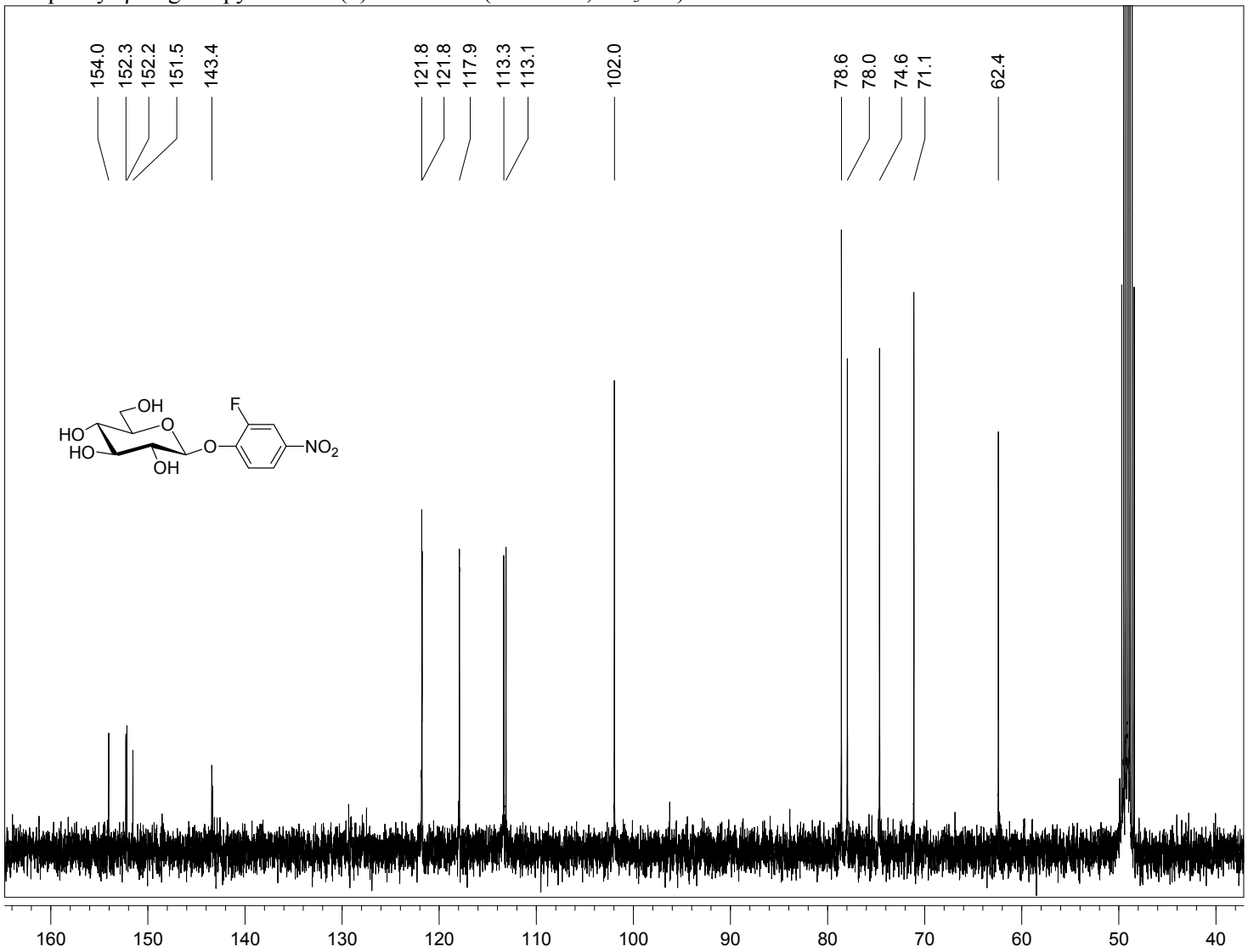
2-bromophenyl-β-D-glucopyranoside (**5**): ^{13}C NMR (100 MHz, CD_3OD)



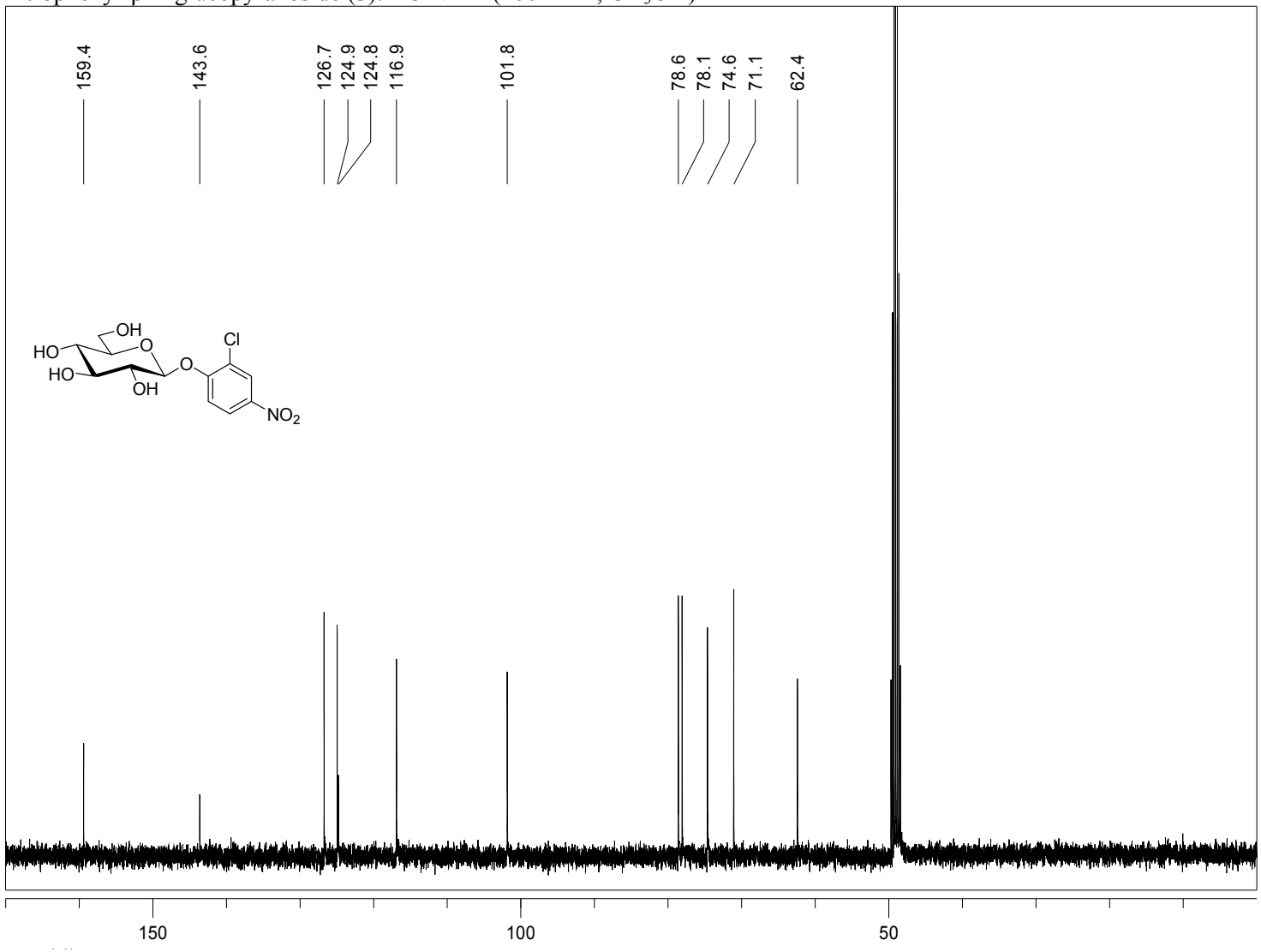
2-iodophenyl-β-D-glucopyranoside (**6**): ^{13}C NMR (100 MHz, CD_3OD)



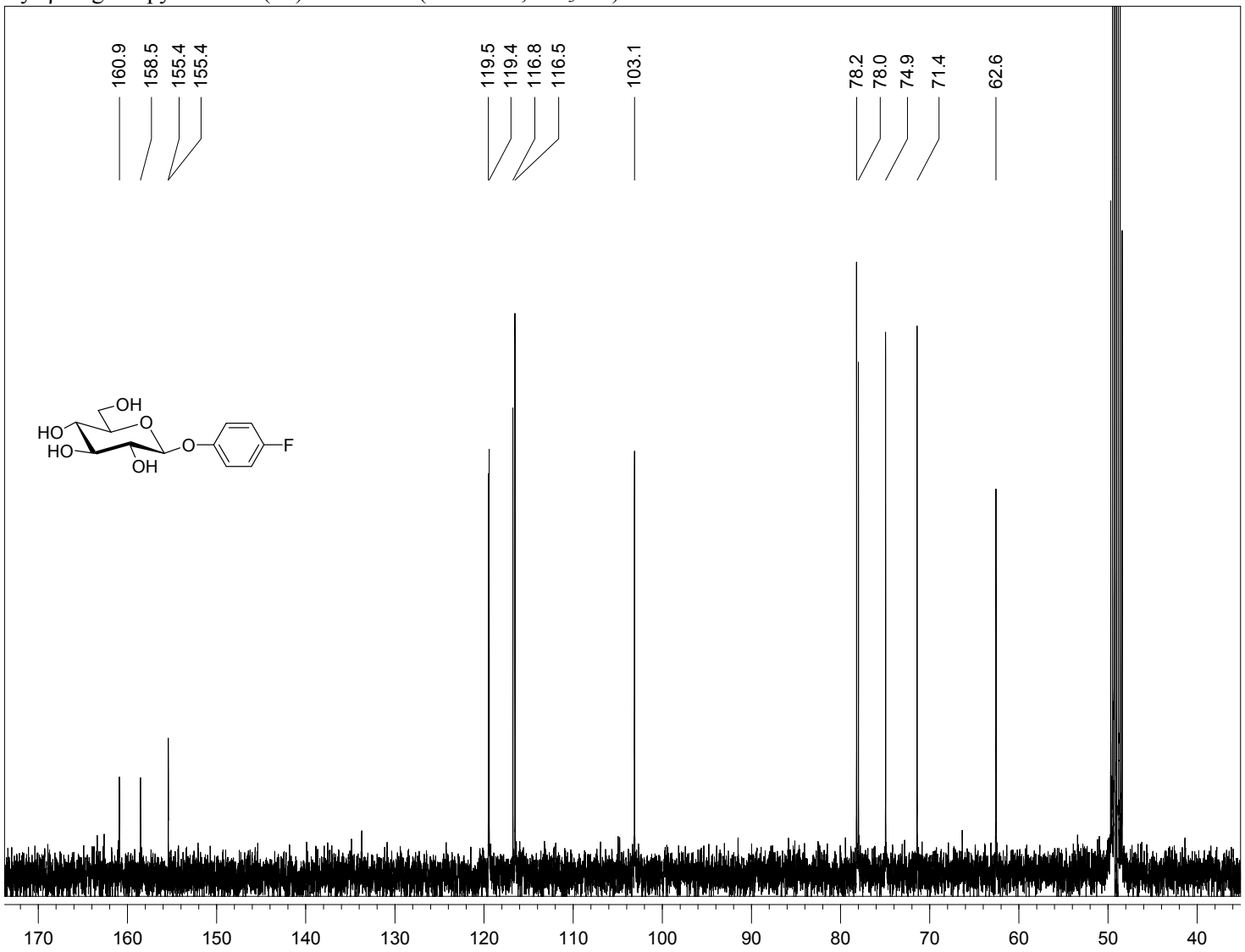
2-fluoro-4-nitrophenyl-β-D-glucopyranoside (**8**): ^{13}C NMR (100 MHz, CD_3OD)



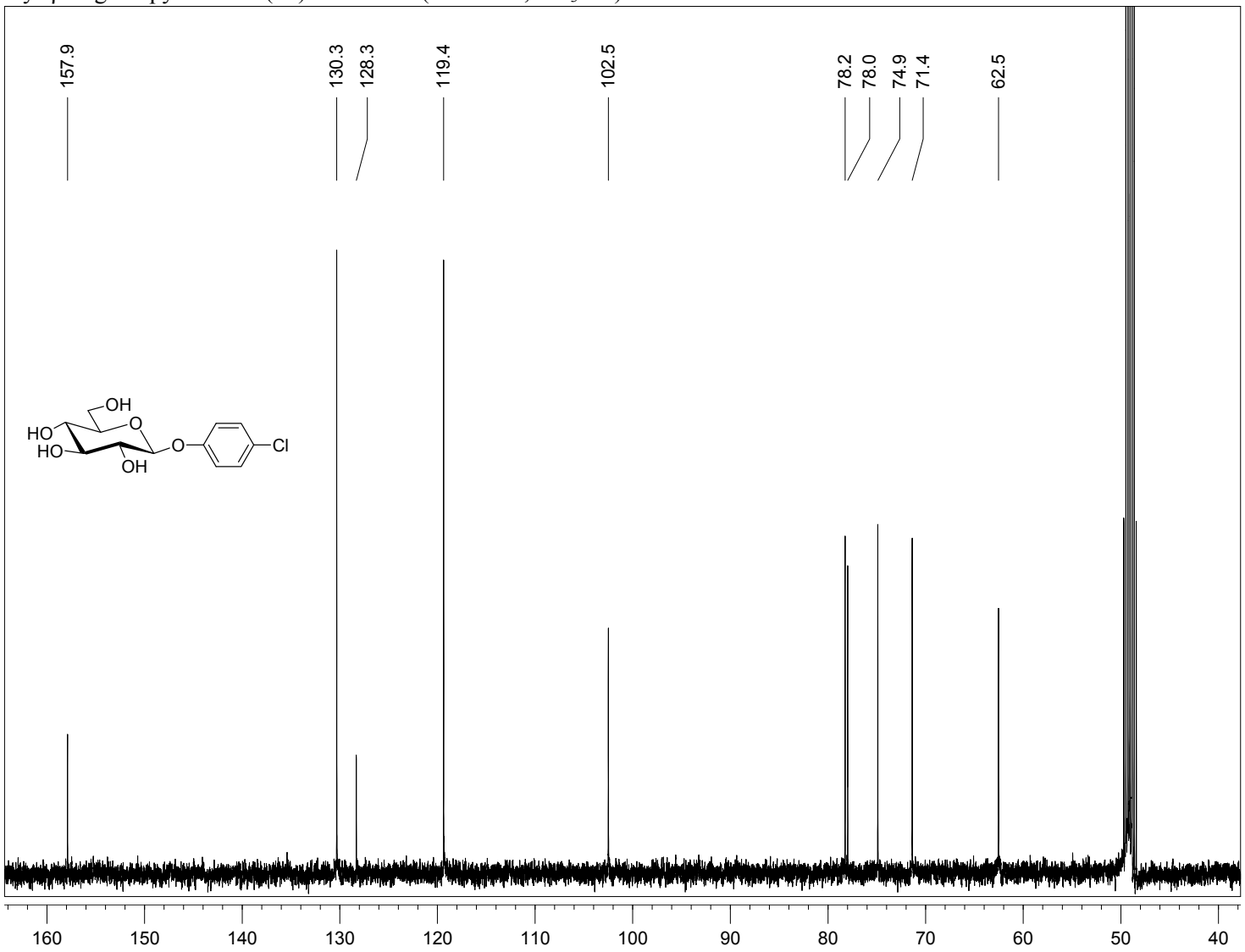
2-chloro-4-nitrophenyl-β-D-glucopyranoside (**9**): ^{13}C NMR (100 MHz, CD_3OD)



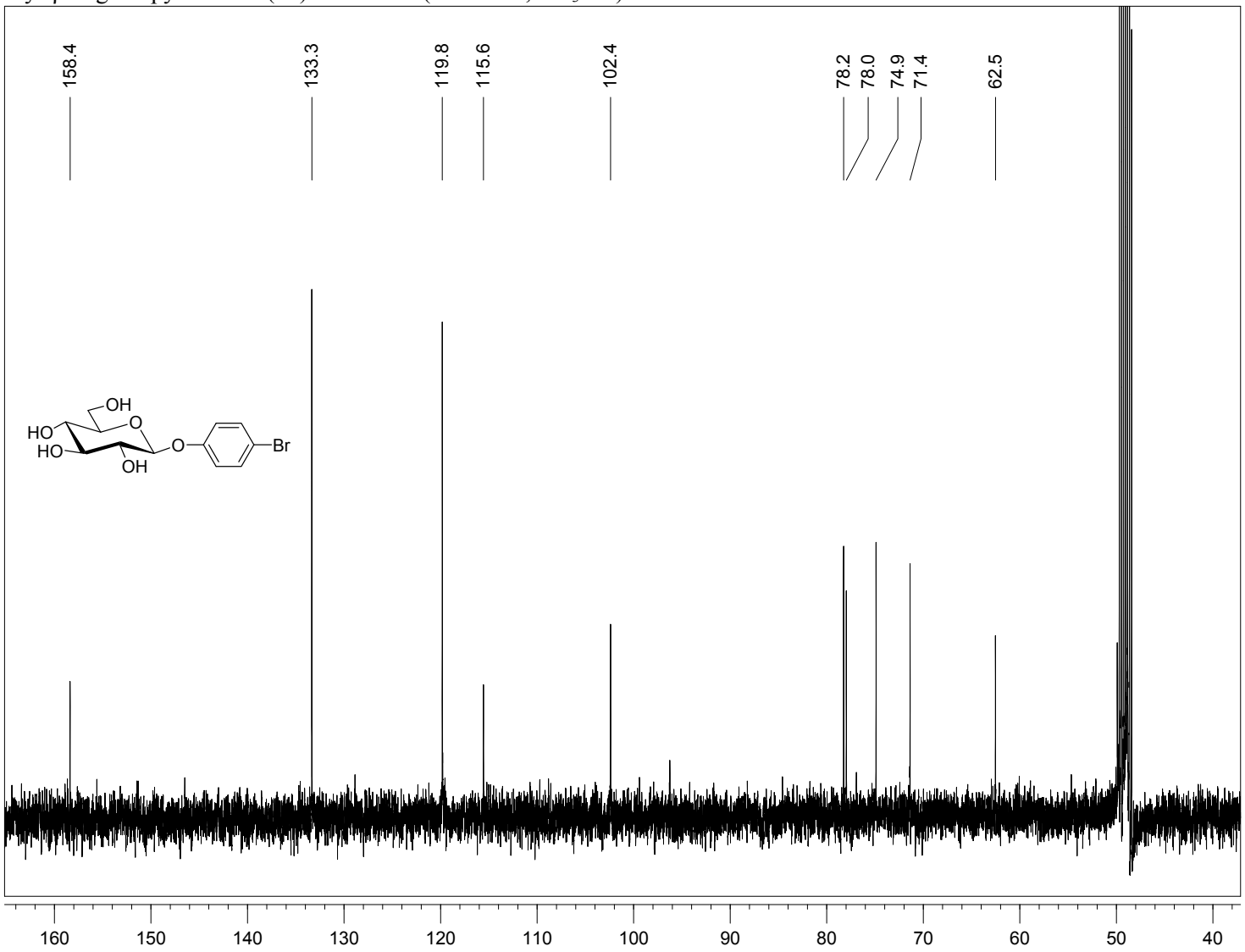
4-fluorophenyl-β-D-glucopyranoside (**11**): ^{13}C NMR (100 MHz, CD_3OD)



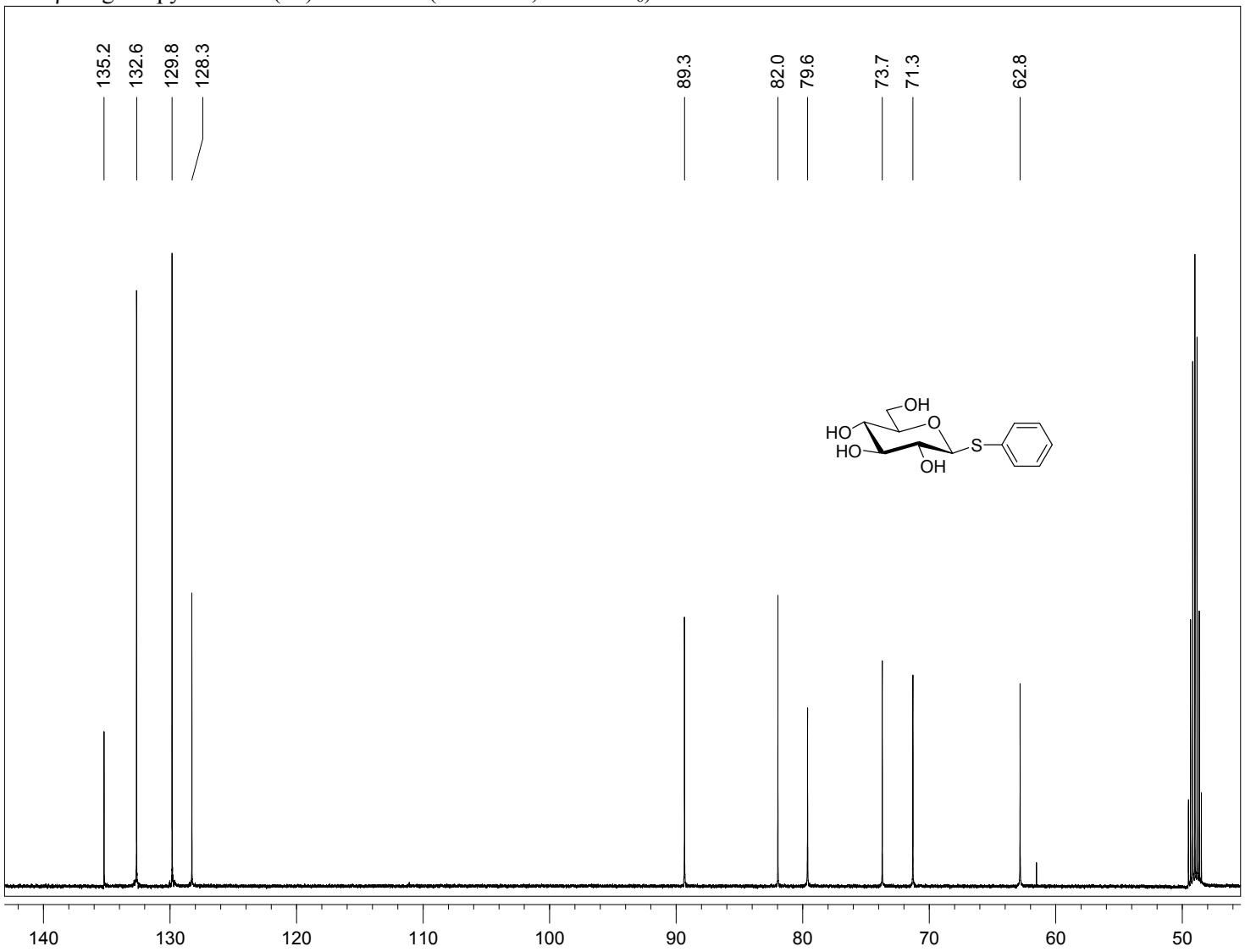
4-chlorophenyl-β-D-glucopyranoside (**12**): ^{13}C NMR (100 MHz, CD_3OD)



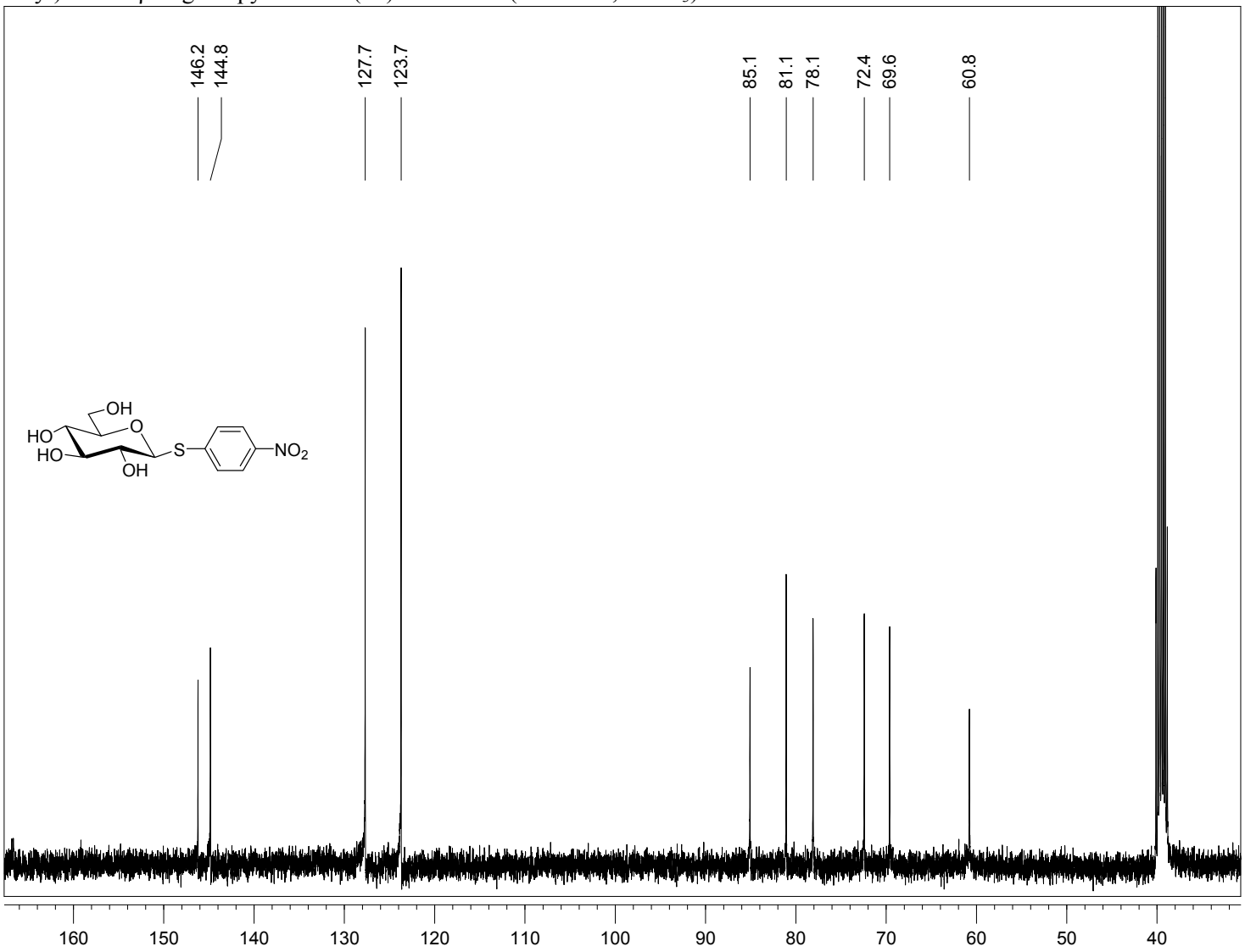
4-bromophenyl-β-D-glucopyranoside (**13**): ^{13}C NMR (100 MHz, CD_3OD)



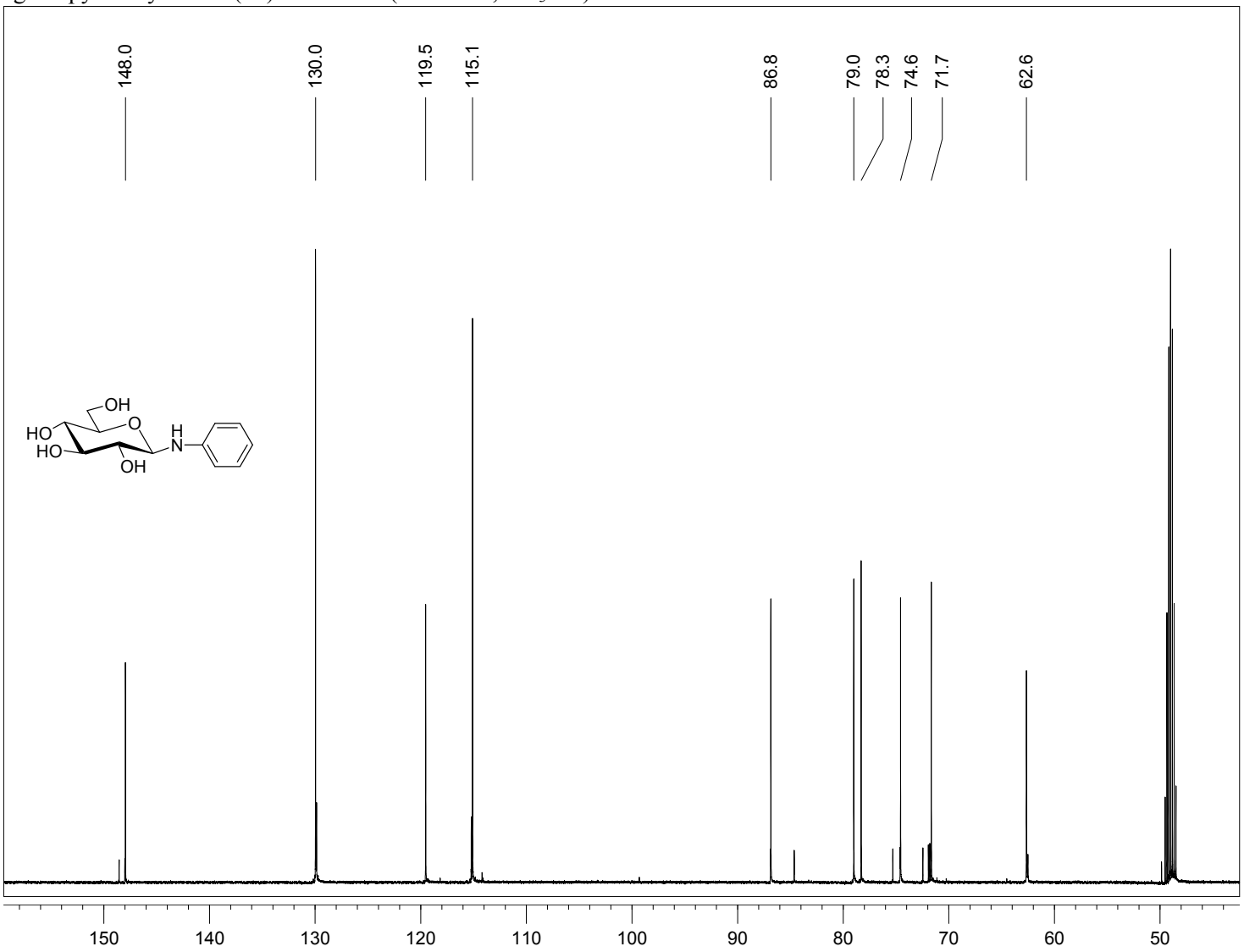
S-phenyl-1-thio-β-D-glucopyranoside (**14**): ^{13}C NMR (125 MHz, DMSO-d_6)



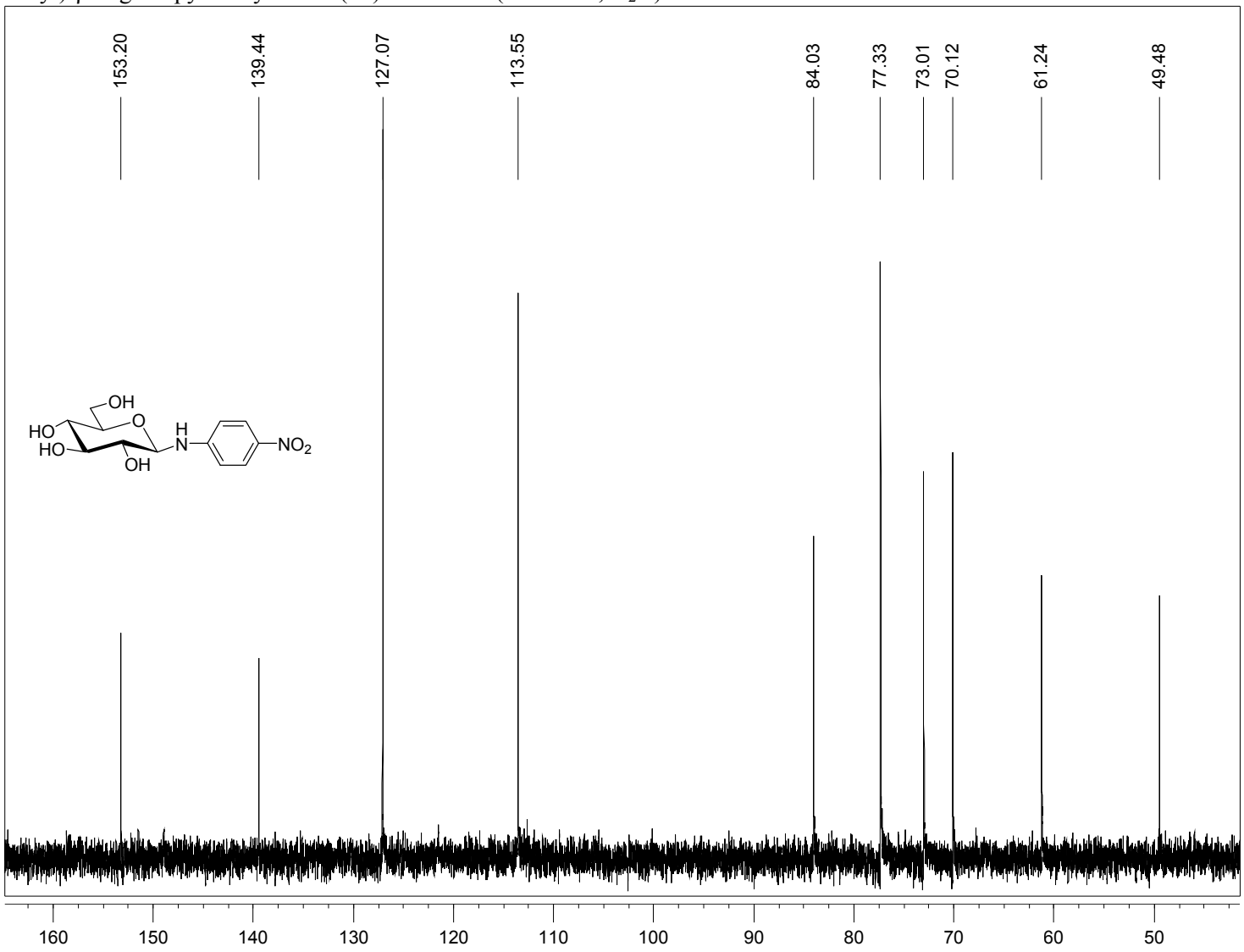
S-(4-nitrophenyl)-1-thio-β-D-glucopyranoside (**15**): ^{13}C NMR (100 MHz, $CDCl_3$)



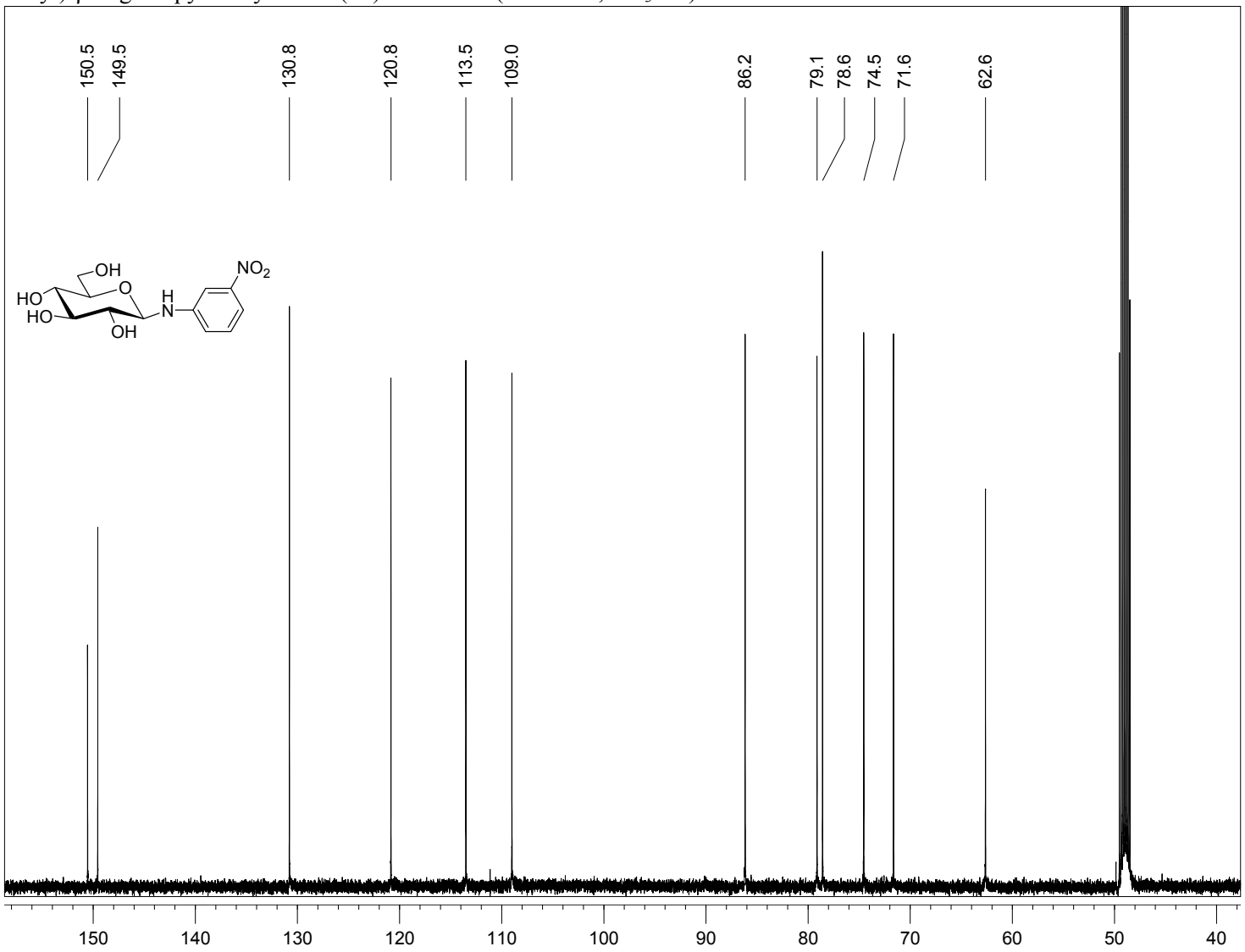
N-phenyl-D-glucopyranosylamine (**16**): ^{13}C NMR (125 MHz, CD_3OD)



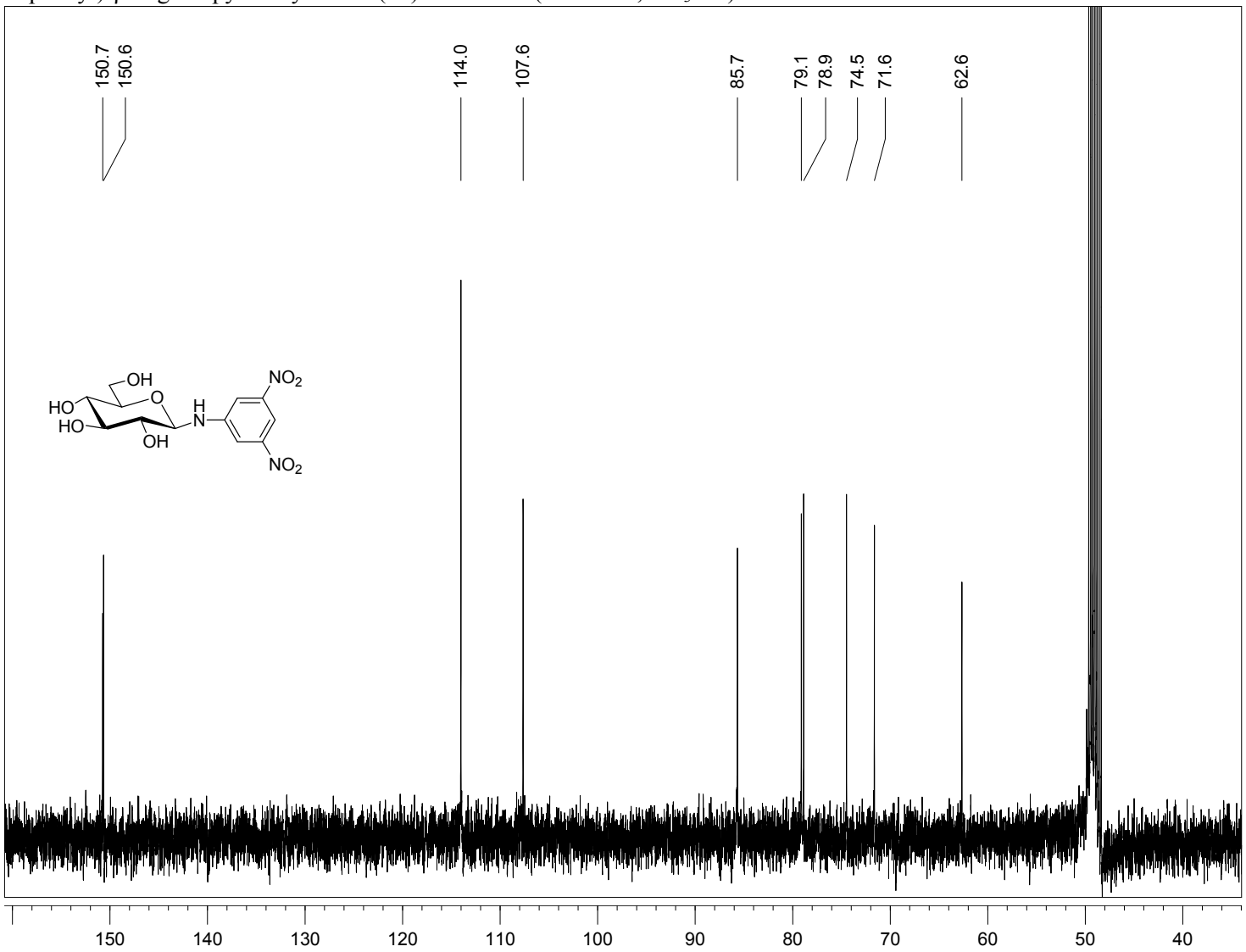
N-(4-nitrophenyl)-β-D-glucopyranosylamine (**17**): ^{13}C NMR (100 MHz, D_2O)



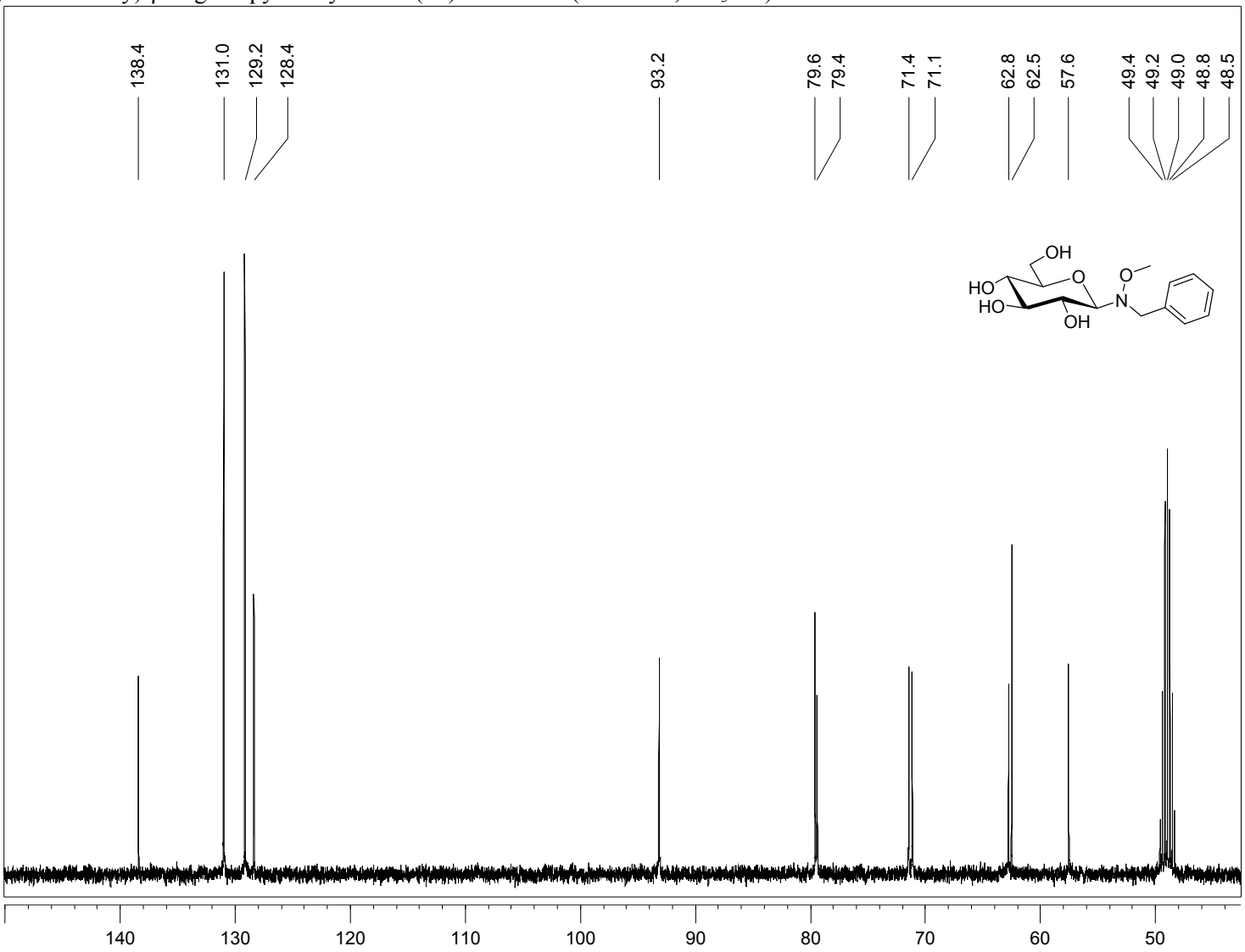
N-(3-nitrophenyl)-β-D-glucopyranosylamine (**18**): ^{13}C NMR (125 MHz, CD_3OD)



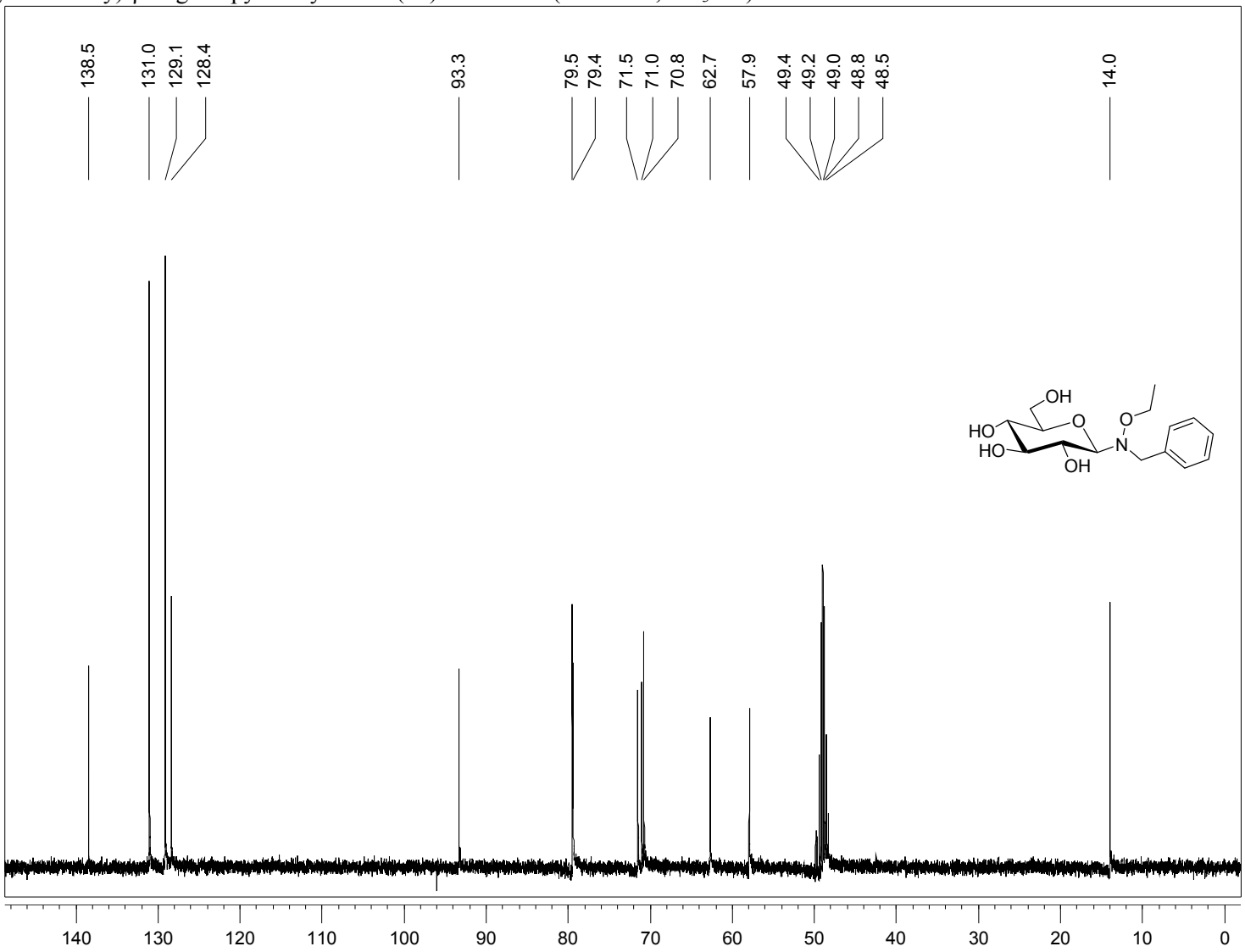
N-(3,5-dinitrophenyl)-β-D-glucopyranosylamine (**20**): ^{13}C NMR (100 MHz, CD_3OD)



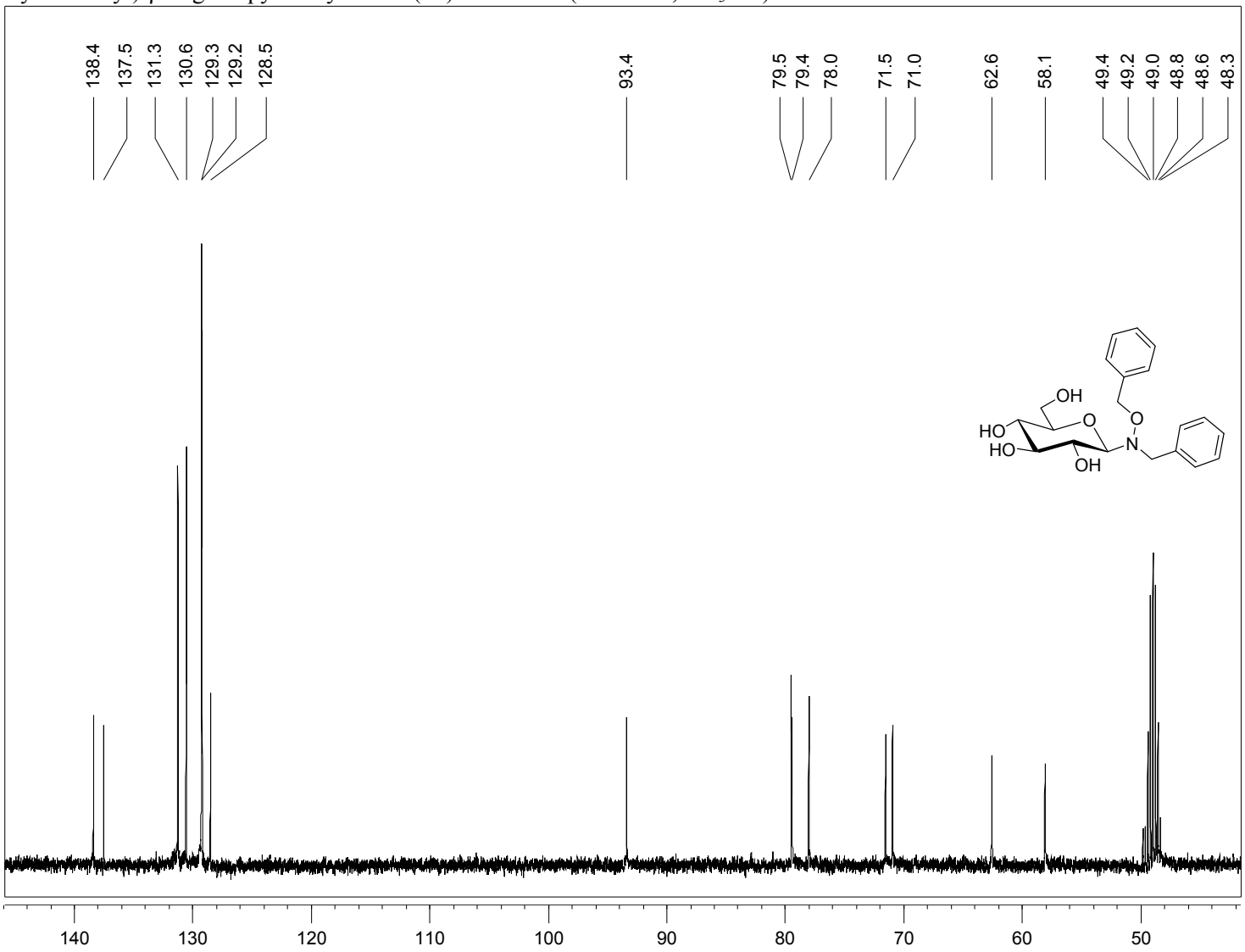
N-(*N*-benzyl-*N*-methoxy)-β-D-glucopyranosylamine (**21**): ^{13}C NMR (100 MHz, CD_3OD)



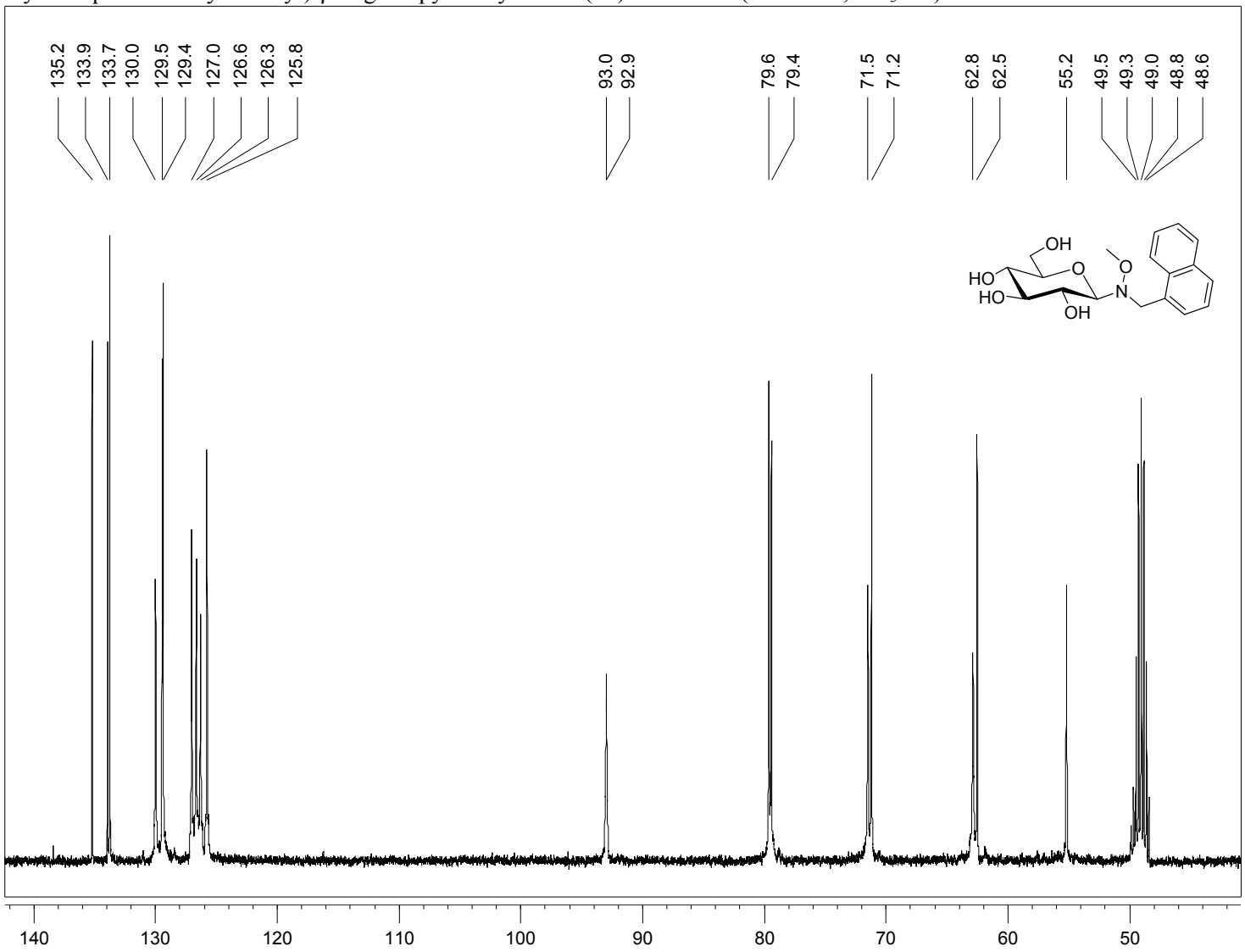
N-(*N*-benzyl-*N*-ethoxy)-β-D-glucopyranosylamine (**22**): ^{13}C NMR (100 MHz, CD_3OD)



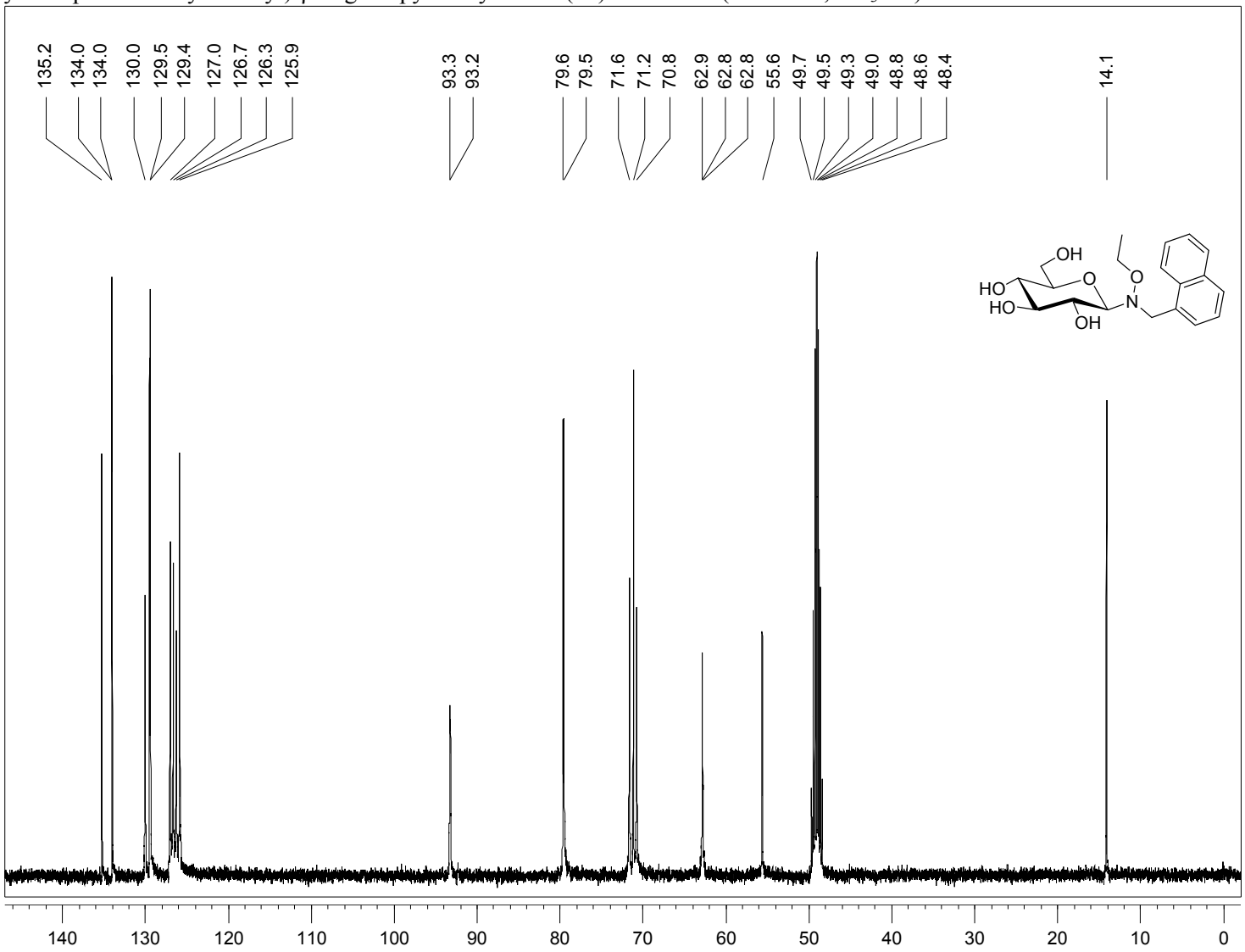
N-(*N*-benzoxy-*N*-benzyl)-β-D-glucopyranosylamine (**23**): ^{13}C NMR (100 MHz, CD_3OD)



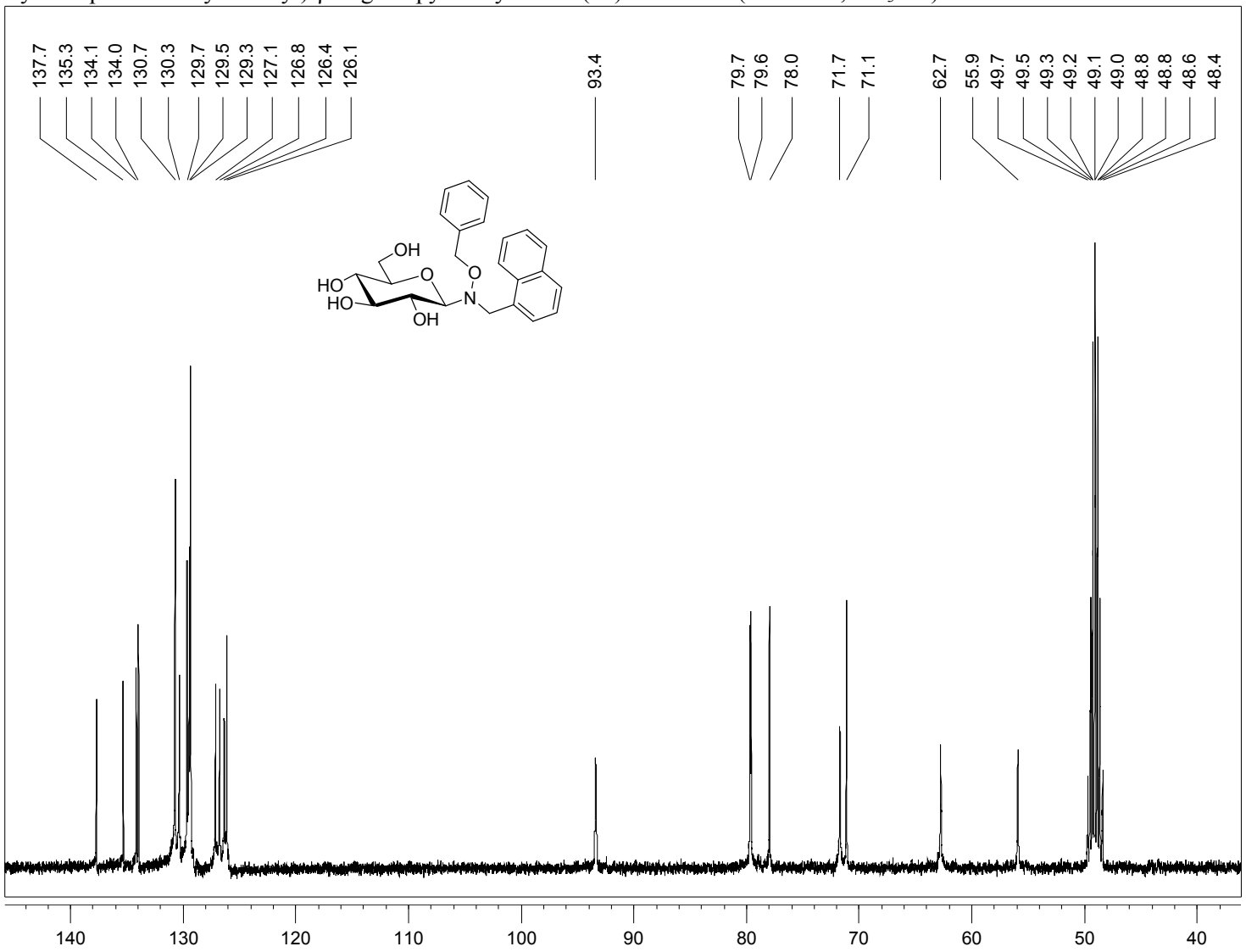
N-(*N*-methoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (**24**): ^{13}C NMR (100 MHz, CD_3OD)



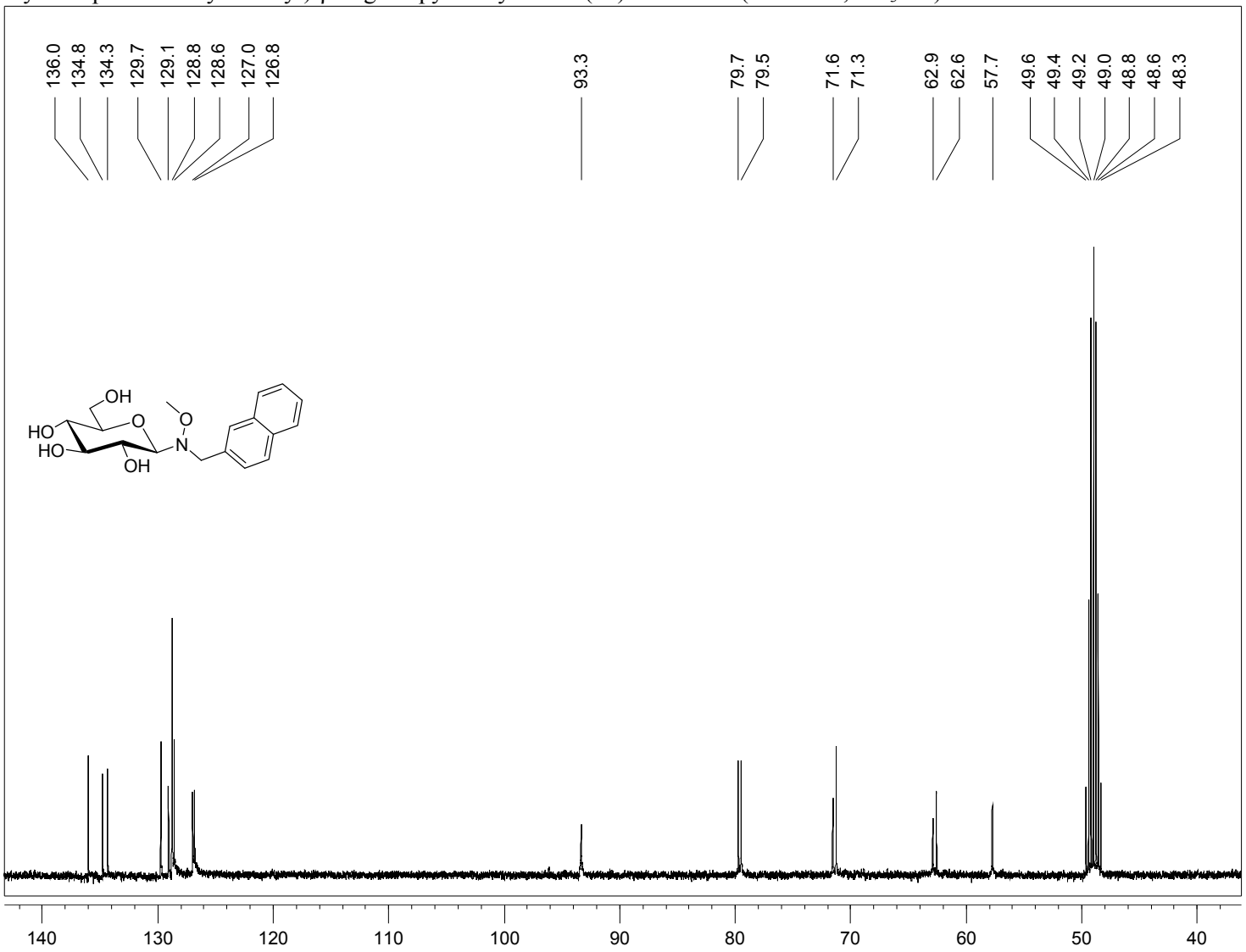
N-(*N*-ethoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (**25**): ^{13}C NMR (100 MHz, CD_3OD)



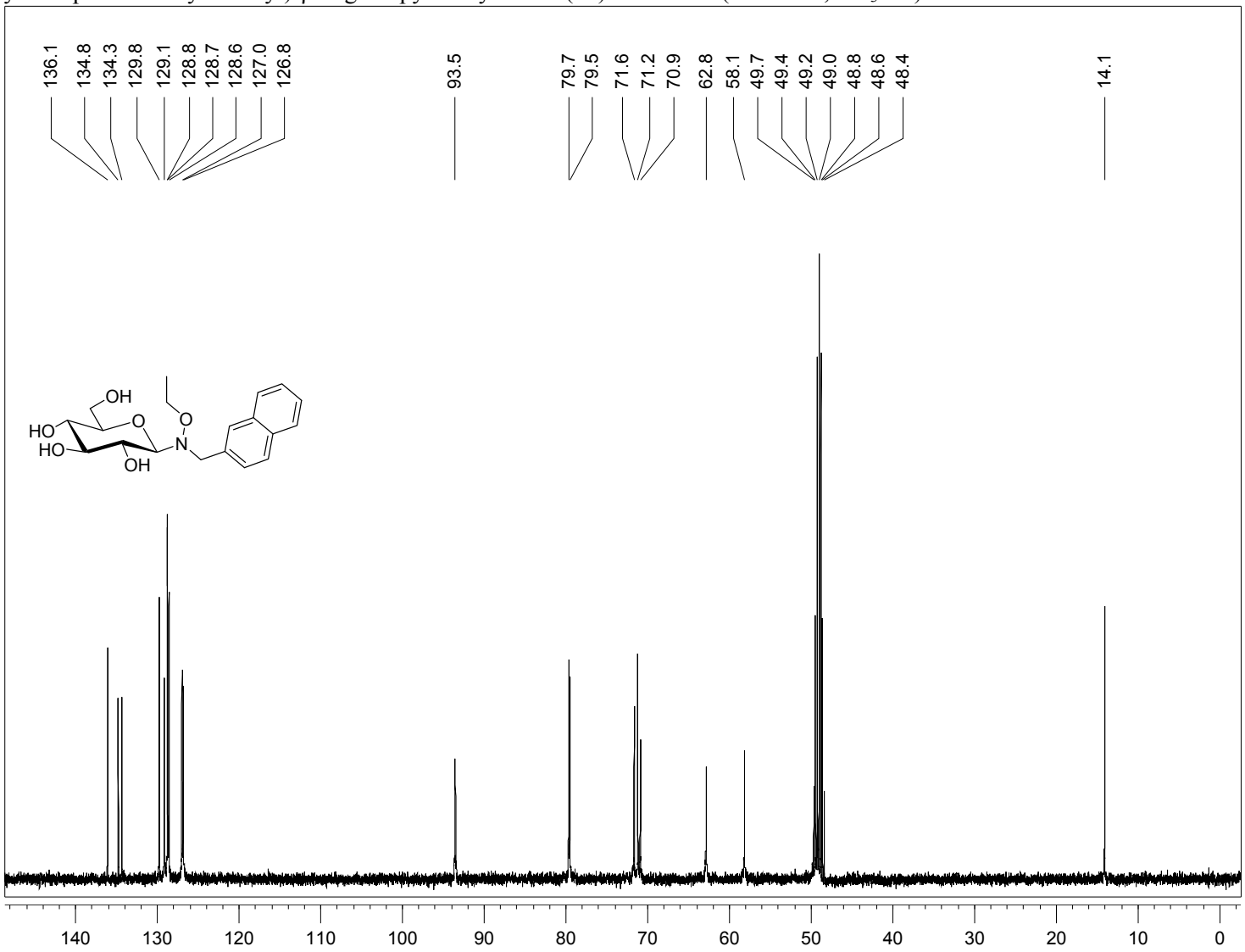
N-(*N*-benzoxy-*N*-naphthalen-1-yl-methyl)-β-D-glucopyranosylamine (**26**): ^{13}C NMR (100 MHz, CD_3OD)



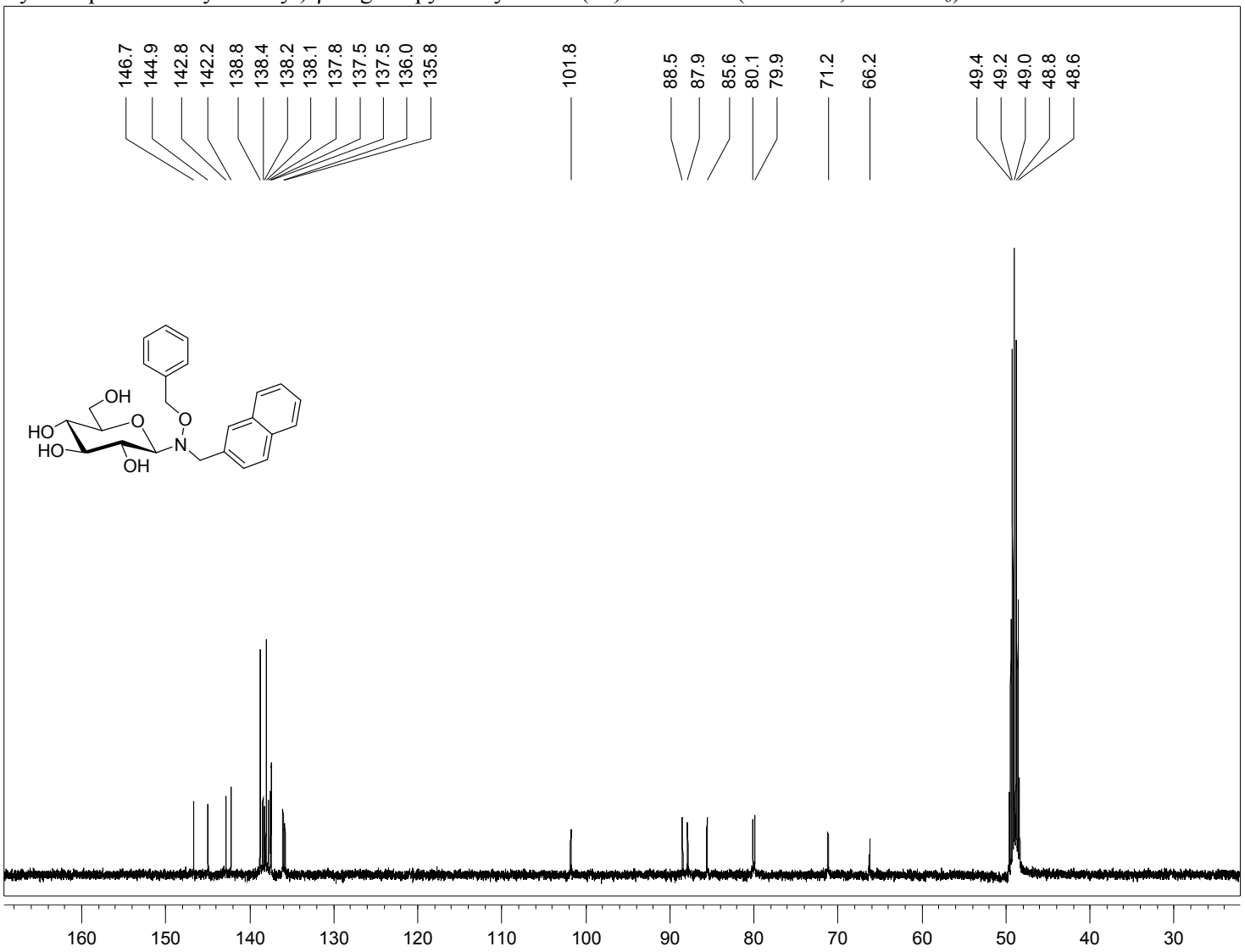
N-(*N*-methoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (**27**): ^{13}C NMR (100 MHz, CD_3OD)



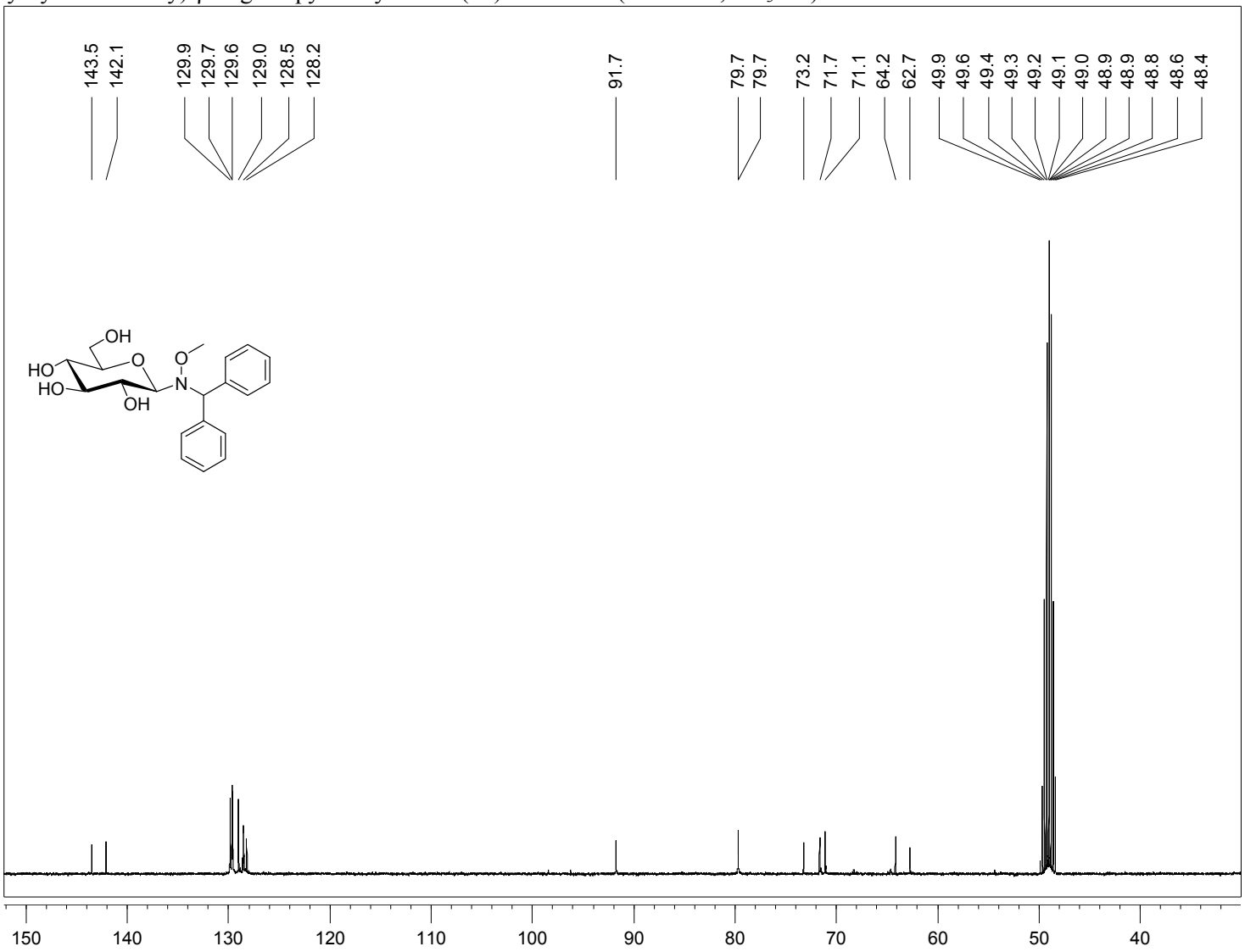
N-(*N*-ethoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (**28**): ^{13}C NMR (100 MHz, CD_3OD)



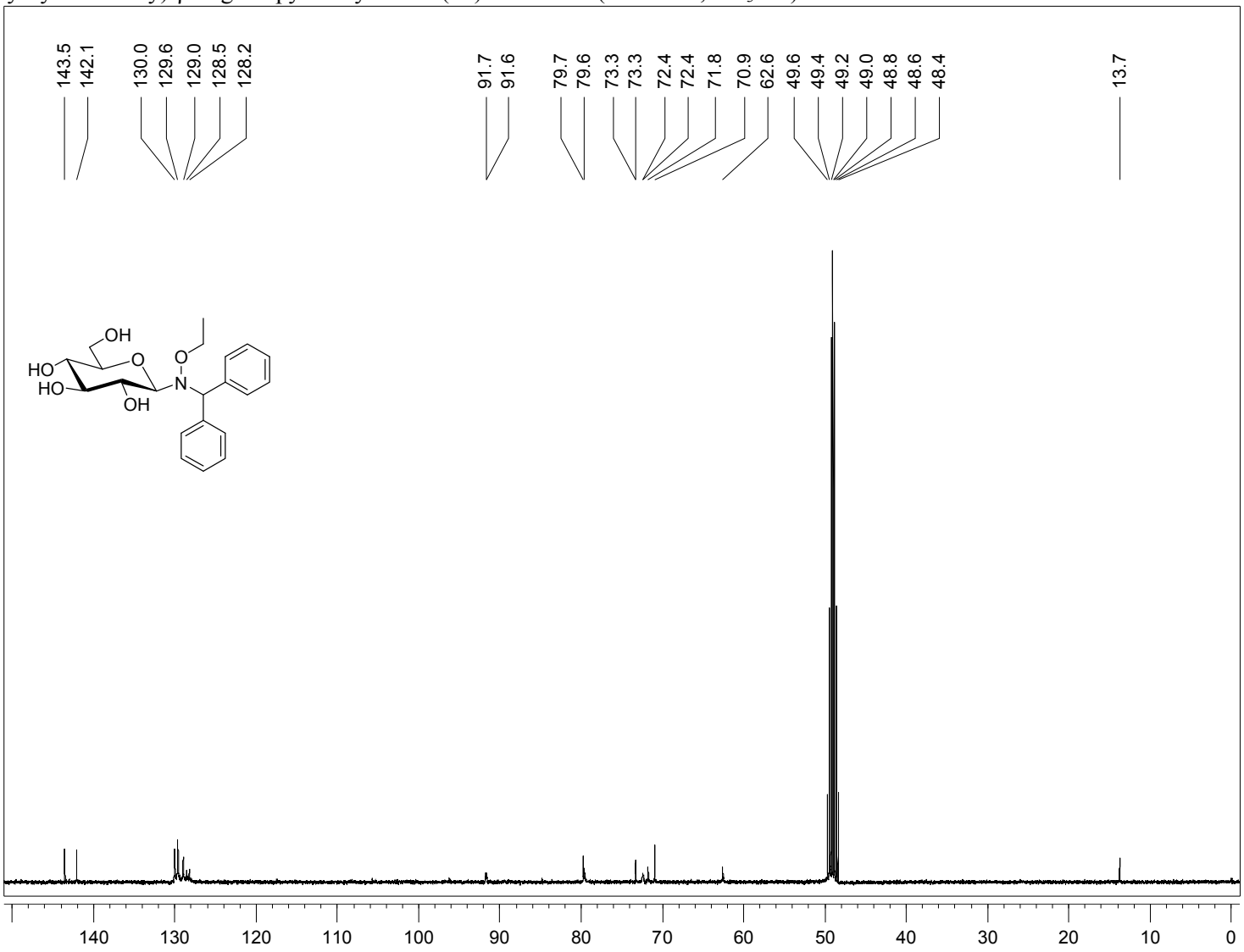
N-(*N*-benzoxy-*N*-naphthalen-2-yl-methyl)-β-D-glucopyranosylamine (**29**): ^{13}C NMR (100 MHz, DMSO-d_6)



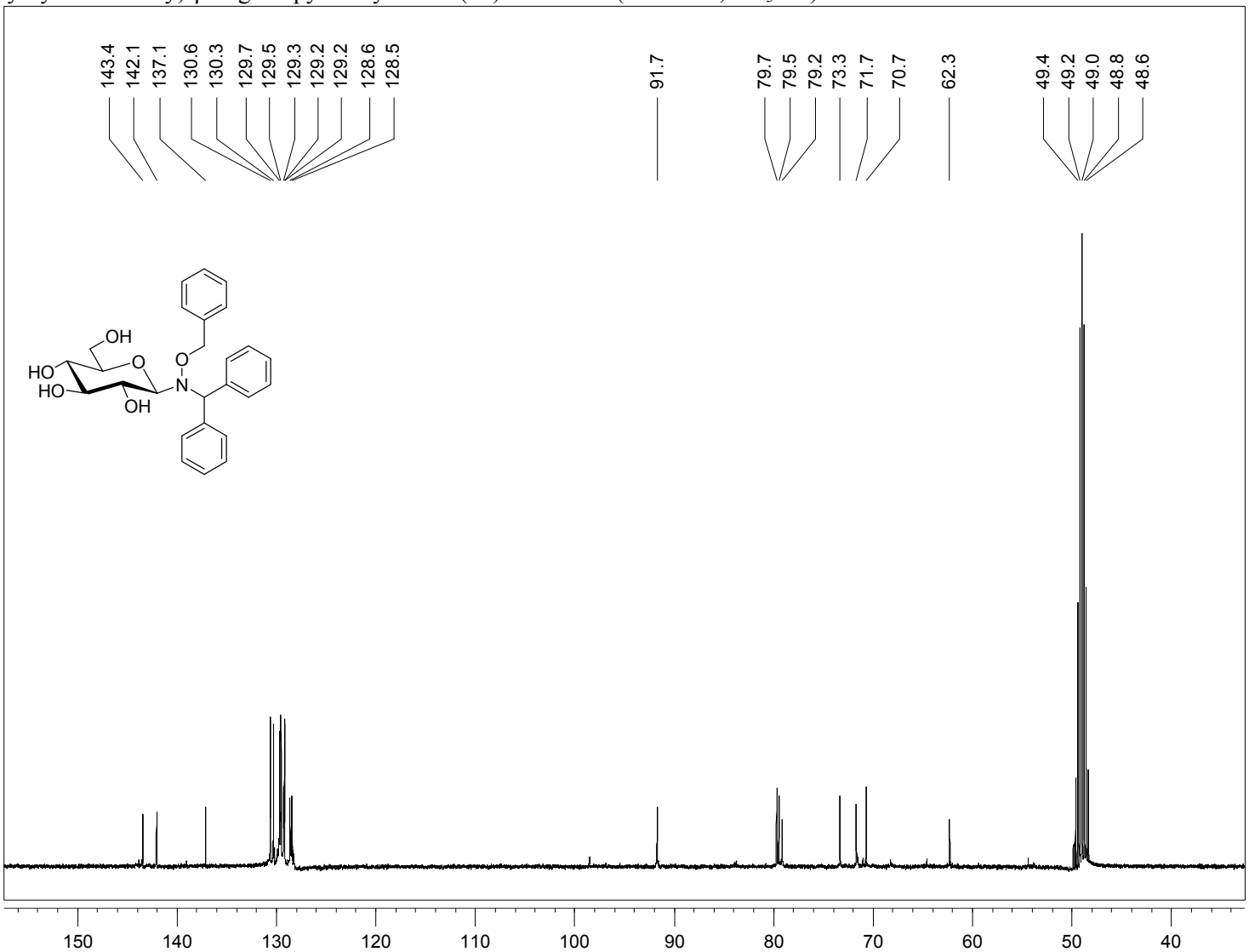
N-(*N*-benzhydryl-*N*-methoxy)-β-D-glucopyranosylamine (**30**): ^{13}C NMR (100 MHz, CD_3OD)



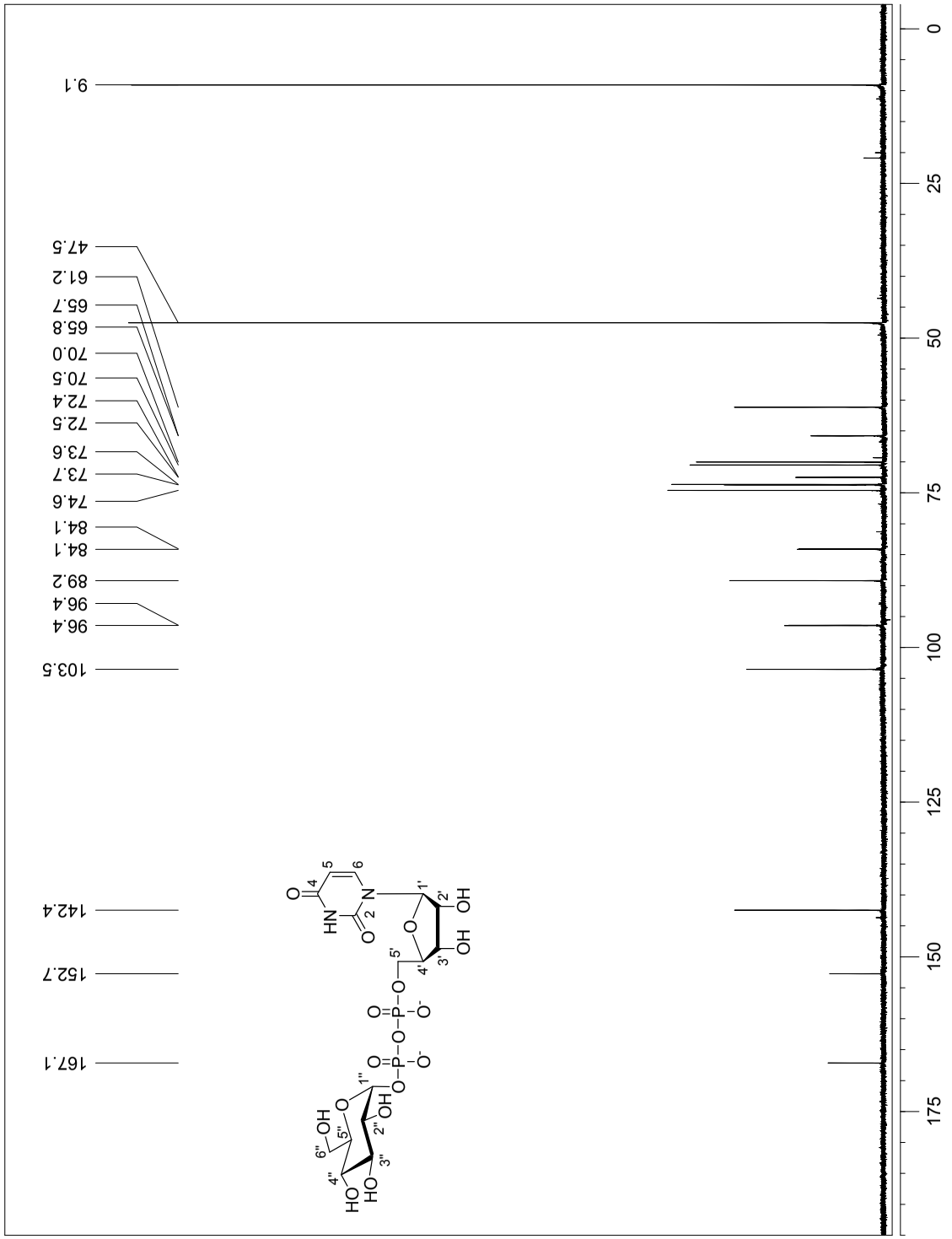
N-(*N*-benzhydryl-*N*-ethoxy)-β-D-glucopyranosylamine (**31**): ^{13}C NMR (100 MHz, CD_3OD)



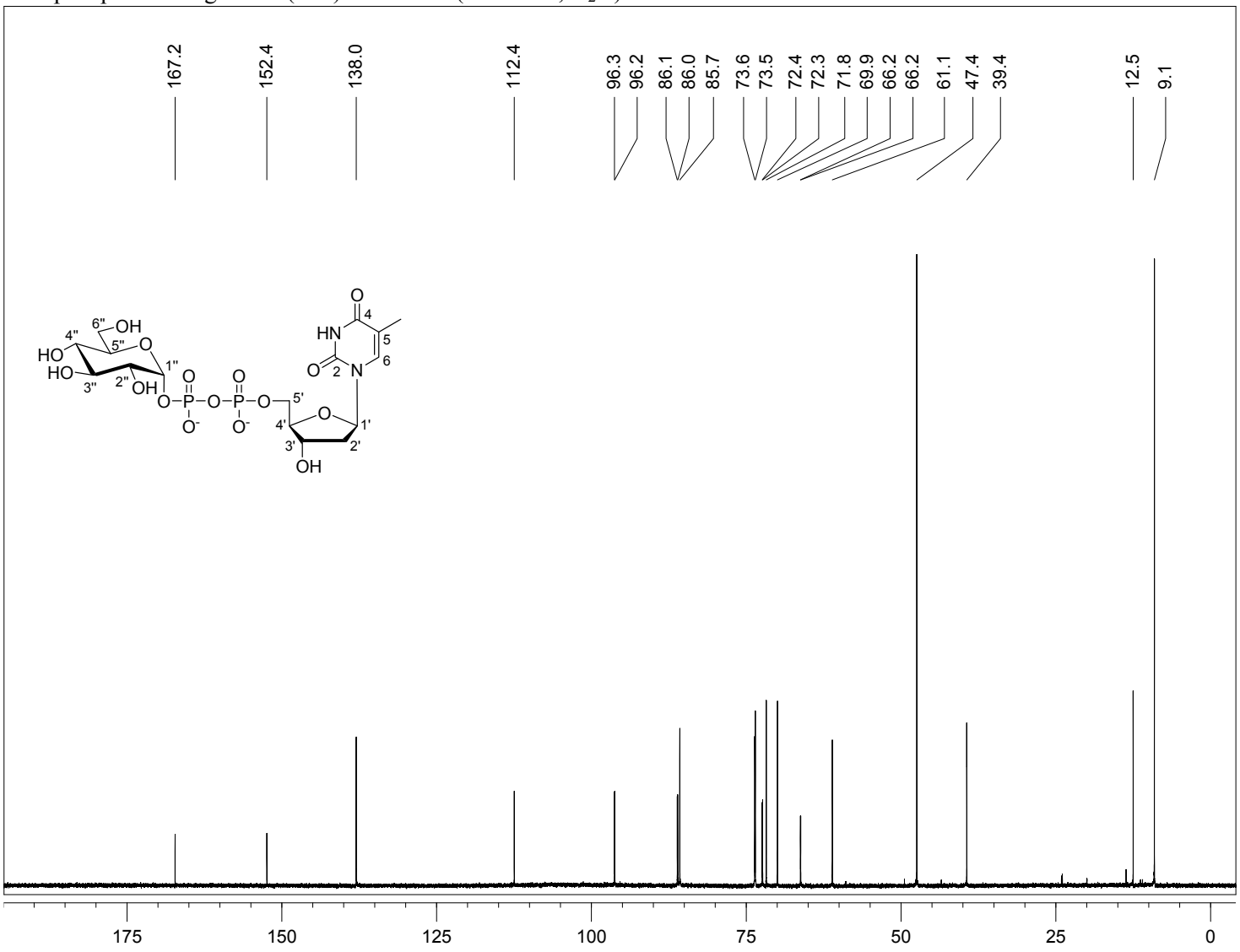
N-(*N*-benzhydryl-*N*-benzoxy)-β-D-glucopyranosylamine (**32**): ^{13}C NMR (100 MHz, CD_3OD)



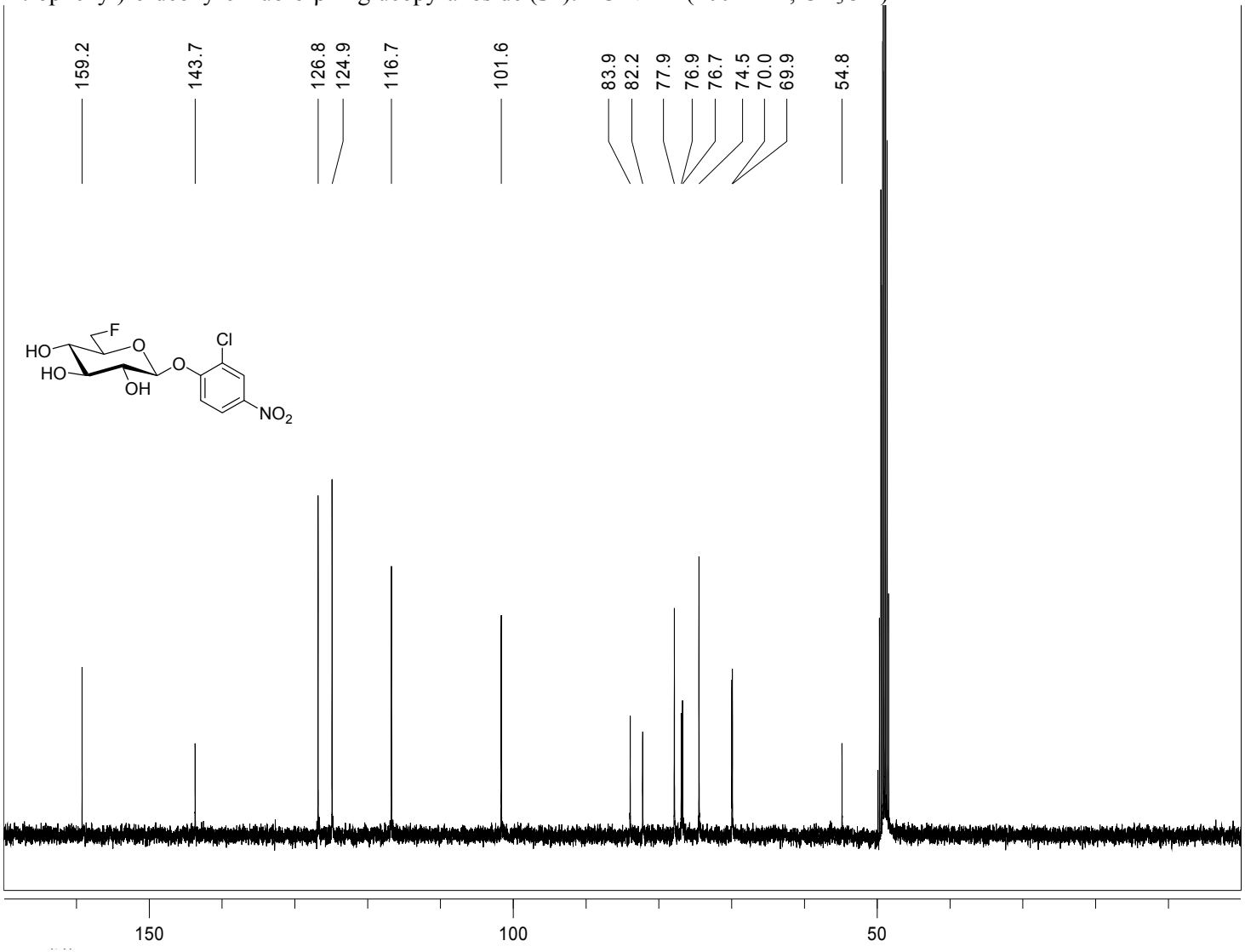
Uridine 5'-diphosphate α-D-glucose (**33a**): ^{13}C NMR (125 MHz, D_2O)


9.1
47.5
61.2
65.7
65.8
70.0
70.5
72.4
72.5
73.6
73.7
74.6
84.1
84.1
89.2
96.4
96.4
103.5
142.4
152.7
167.1
0
25
50
75
100
125
150
175

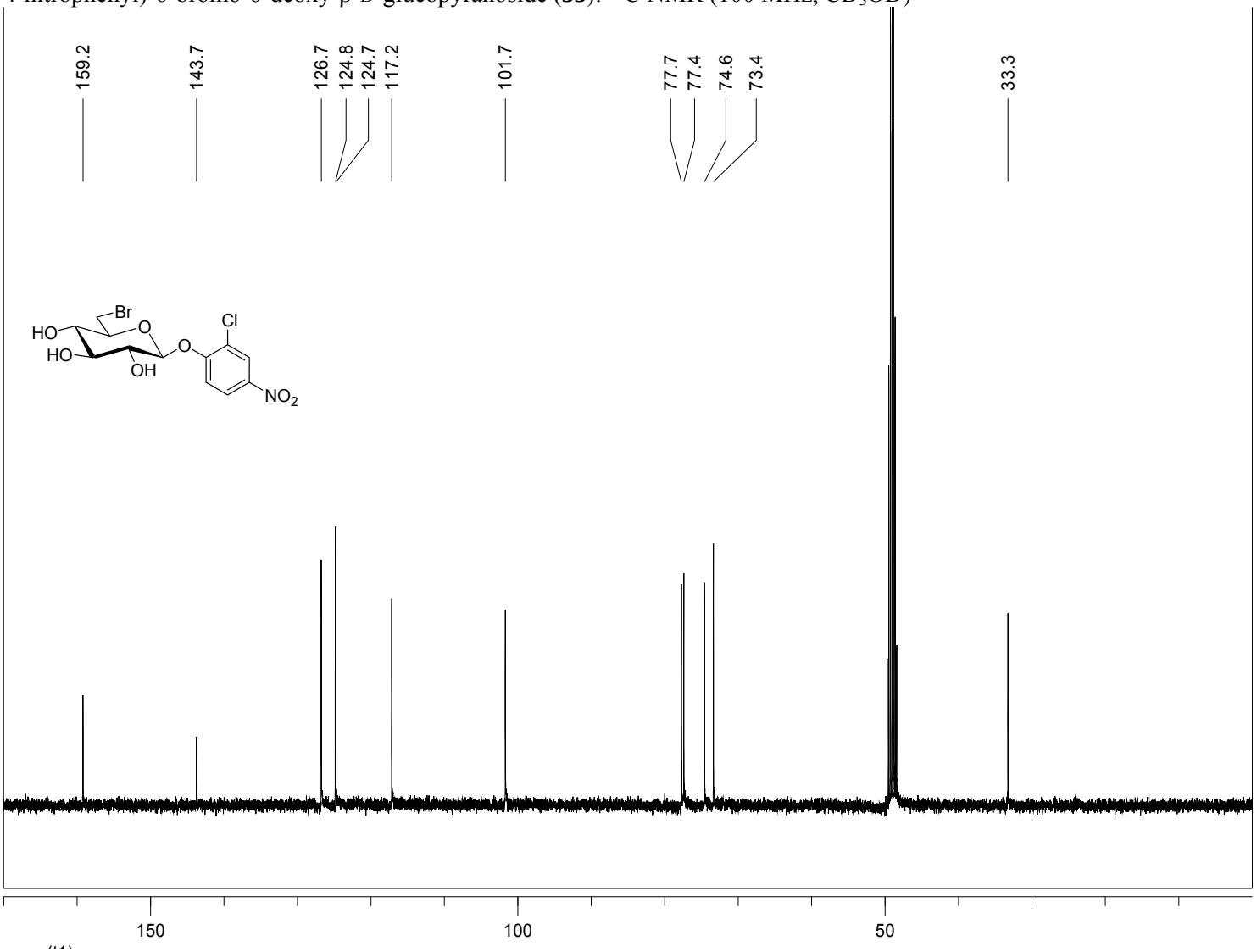
Thymidine 5’-diphosphate α-D-glucose (**33b**): ^{13}C NMR (125 MHz, D_2O)



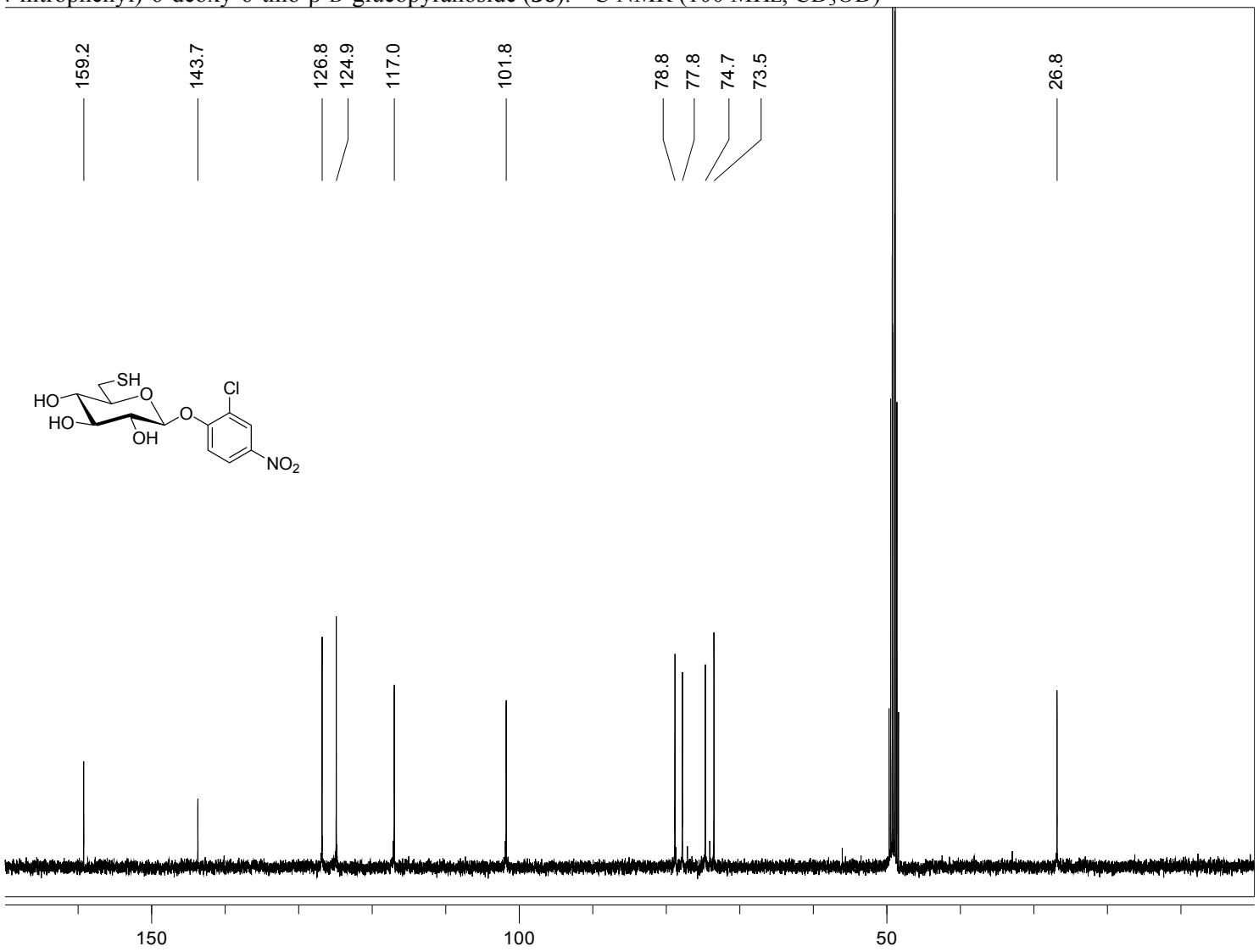
(2-chloro-4-nitrophenyl)-6-deoxy-6-fluoro-β-D-glucopyranoside (**34**): ^{13}C NMR (100 MHz, CD_3OD)



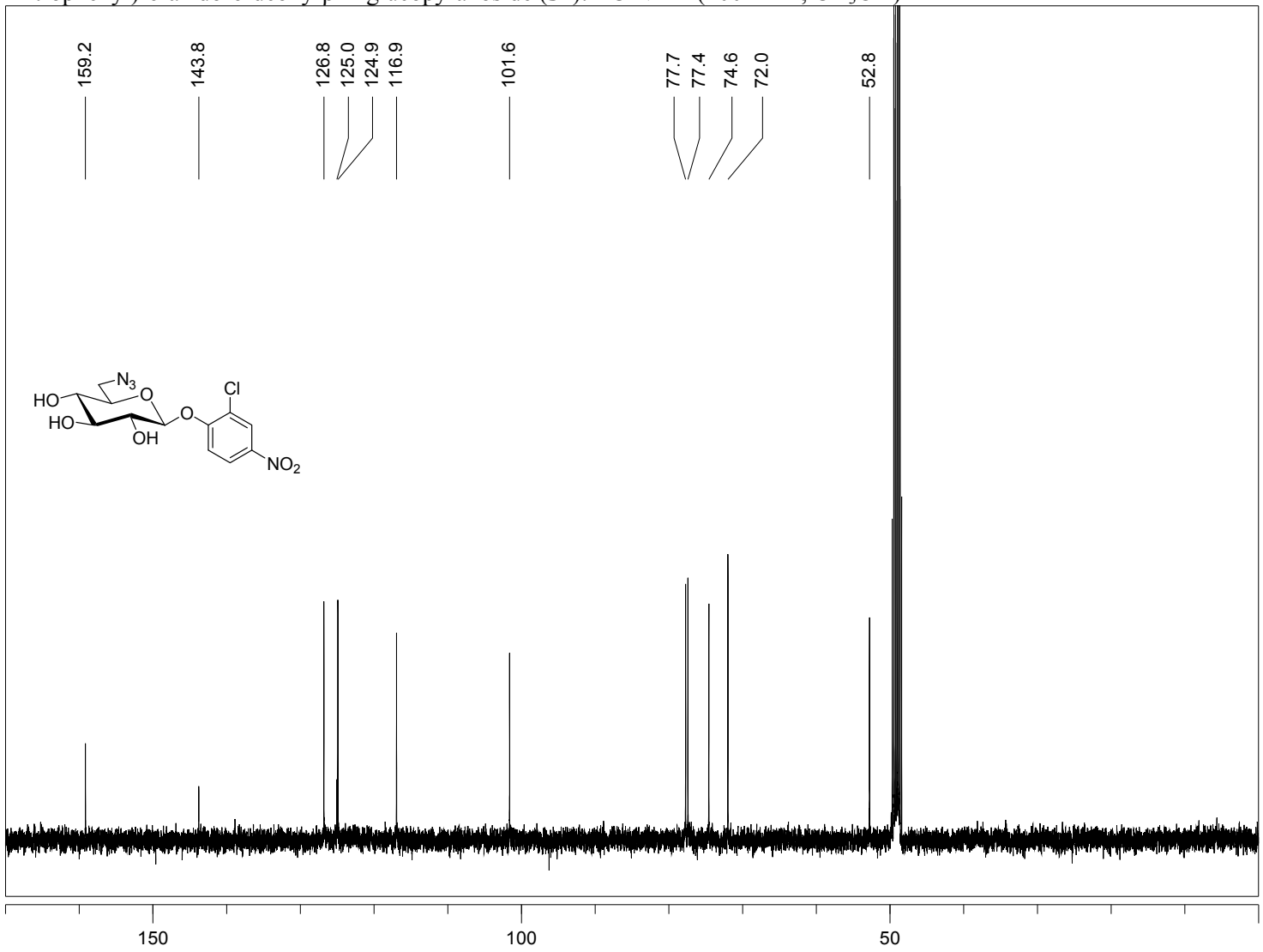
(2-chloro-4-nitrophenyl)-6-bromo-6-deoxy-β-D-glucopyranoside (**35**): ^{13}C NMR (100 MHz, CD_3OD)



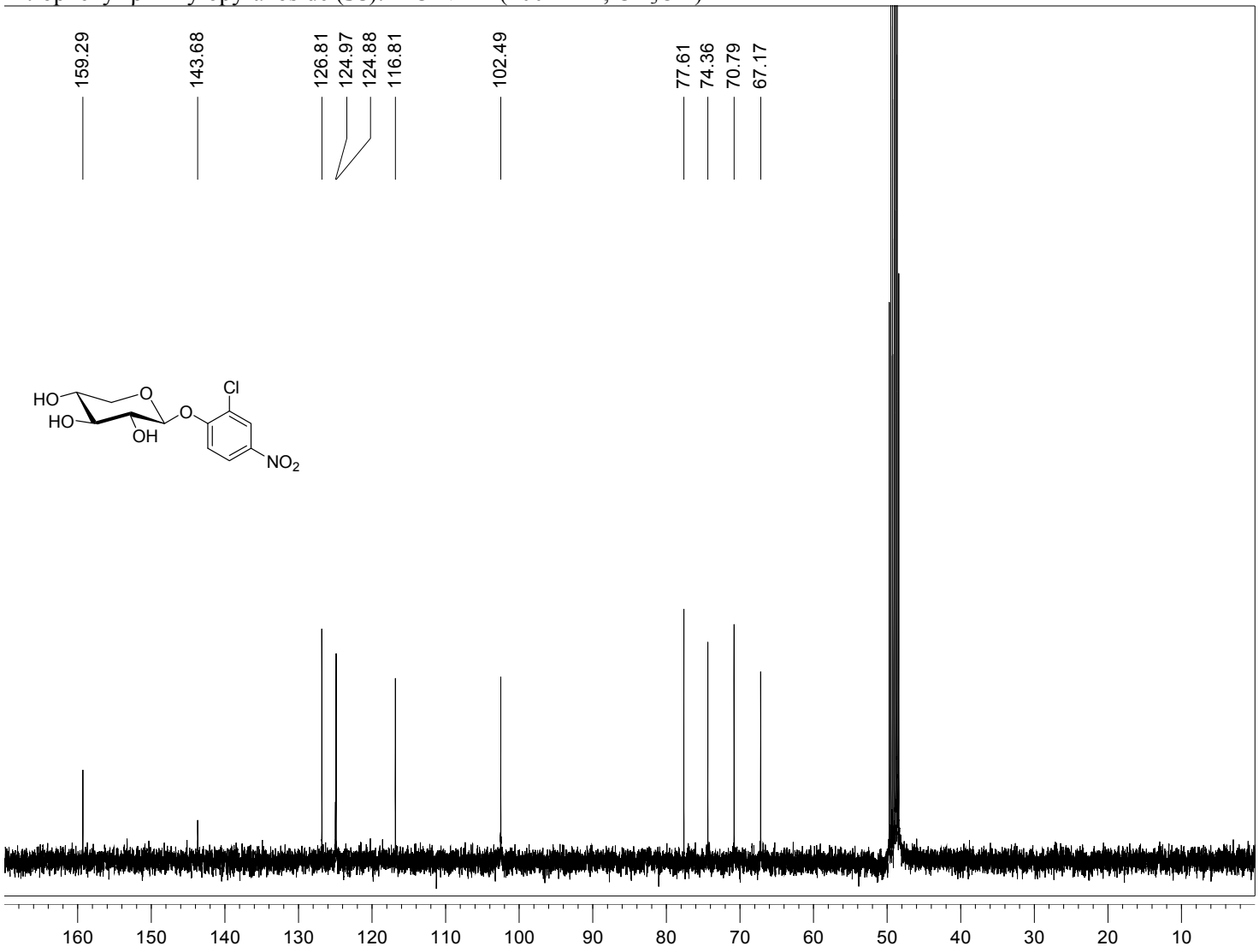
(2-chloro-4-nitrophenyl)-6-deoxy-6-thio-β-D-glucopyranoside (**36**): ^{13}C NMR (100 MHz, CD_3OD)



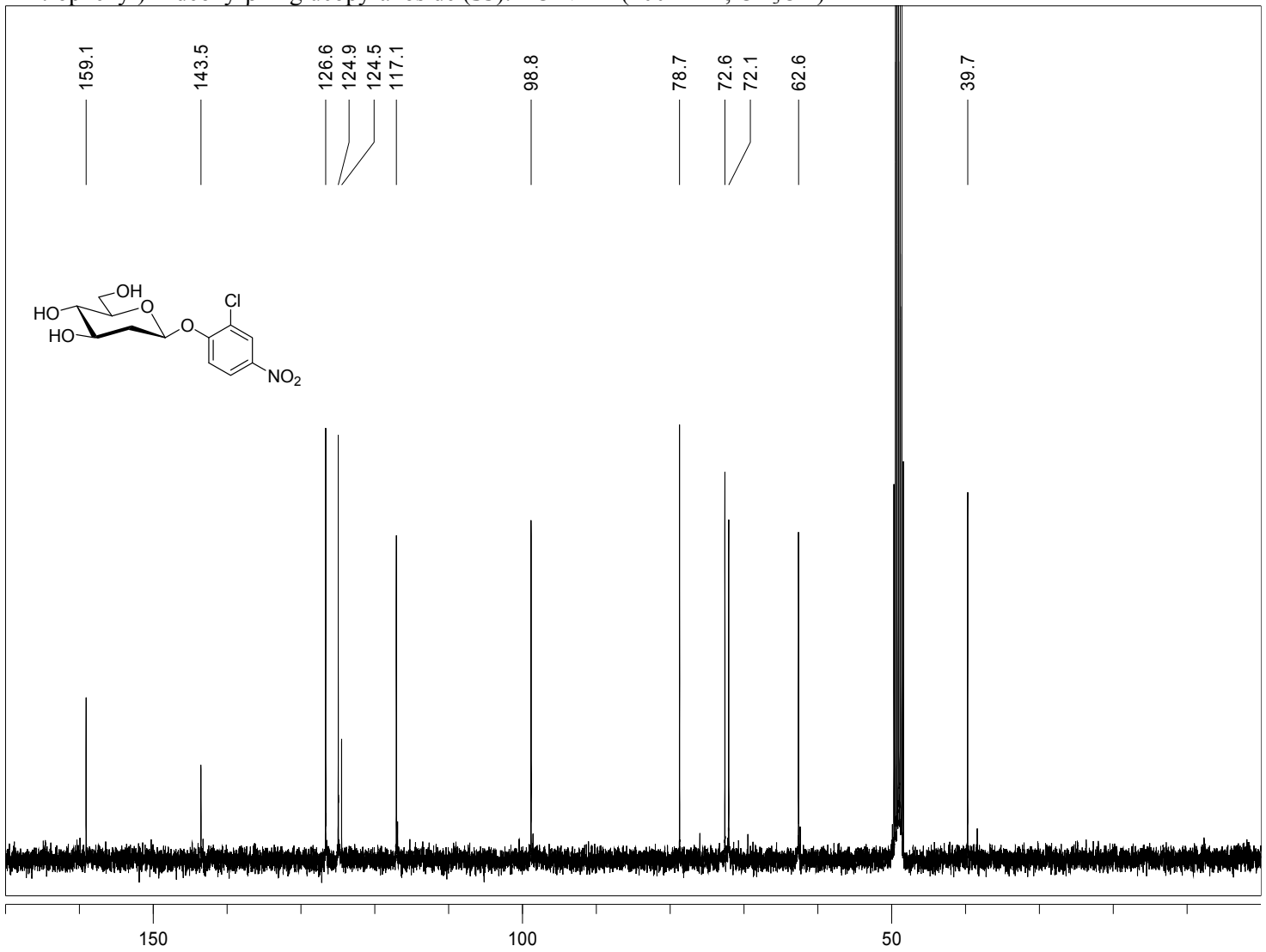
(2-chloro-4-nitrophenyl)-6-azido-6-deoxy-β-D-glucopyranoside (**37**): ^{13}C NMR (100 MHz, CD_3OD)



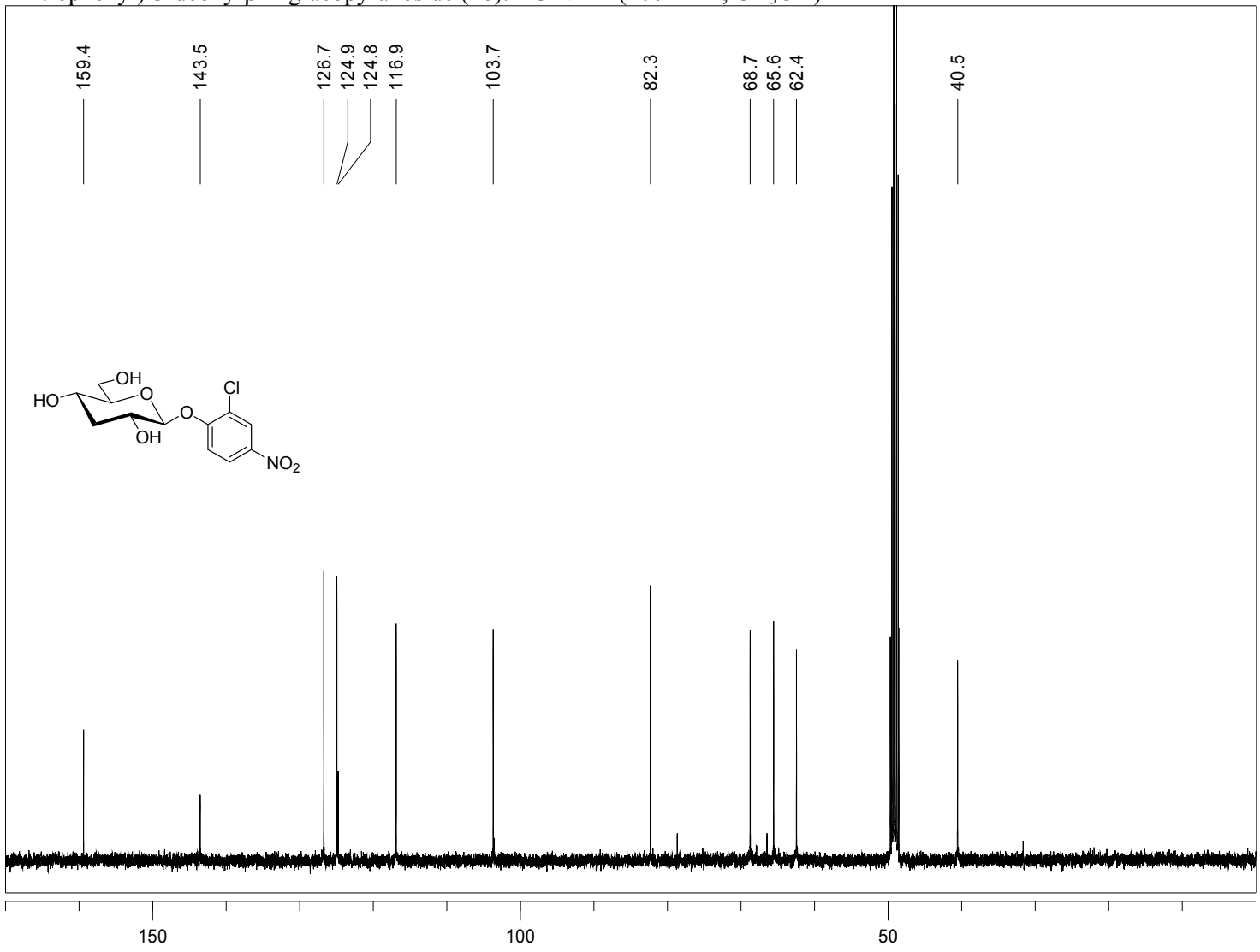
2-chloro-4-nitrophenyl-β-D-xylopyranoside (**38**): ^{13}C NMR (100 MHz, CD_3OD)



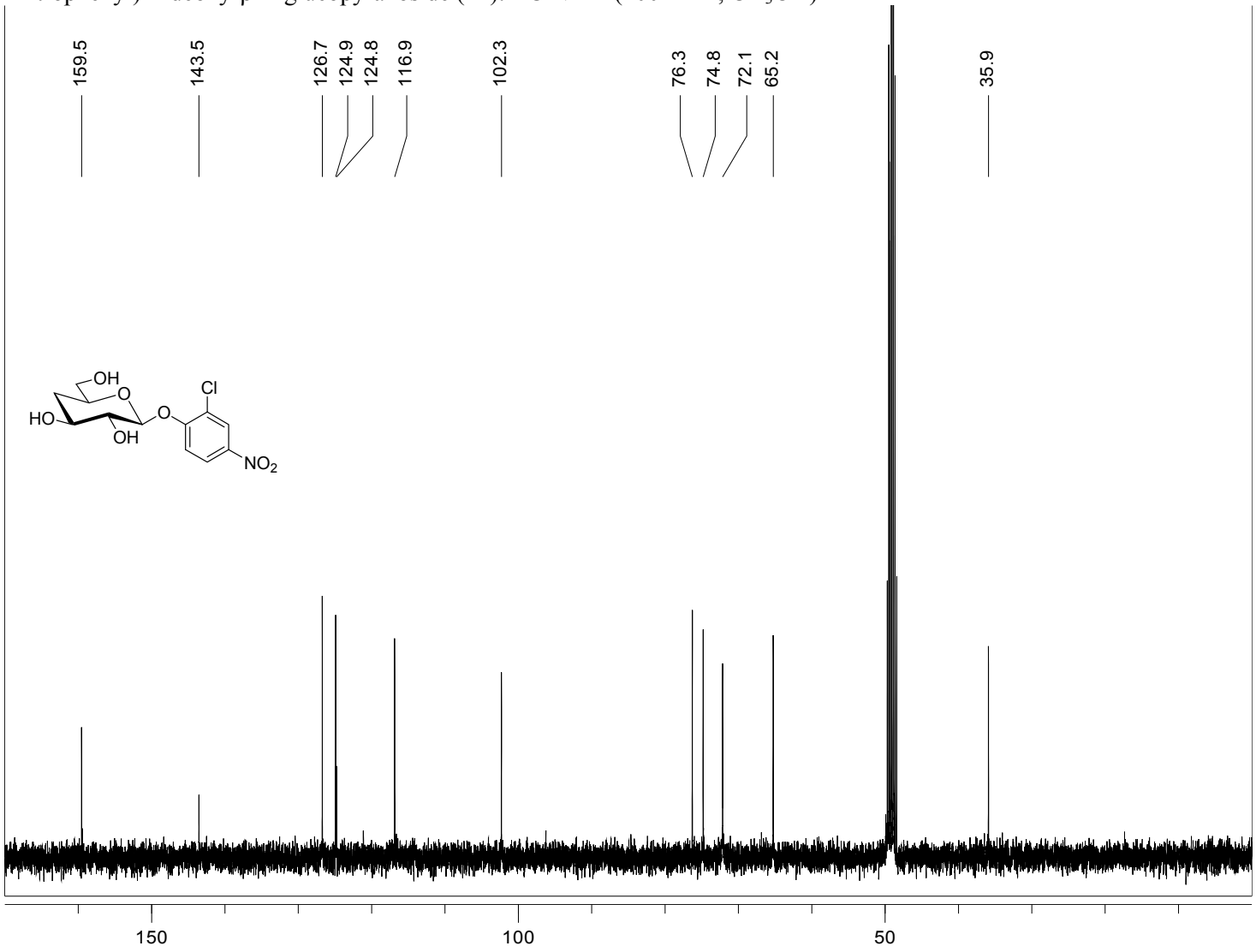
(2-chloro-4-nitrophenyl)-2-deoxy-β-D-glucopyranoside (**39**): ^{13}C NMR (100 MHz, CD_3OD)



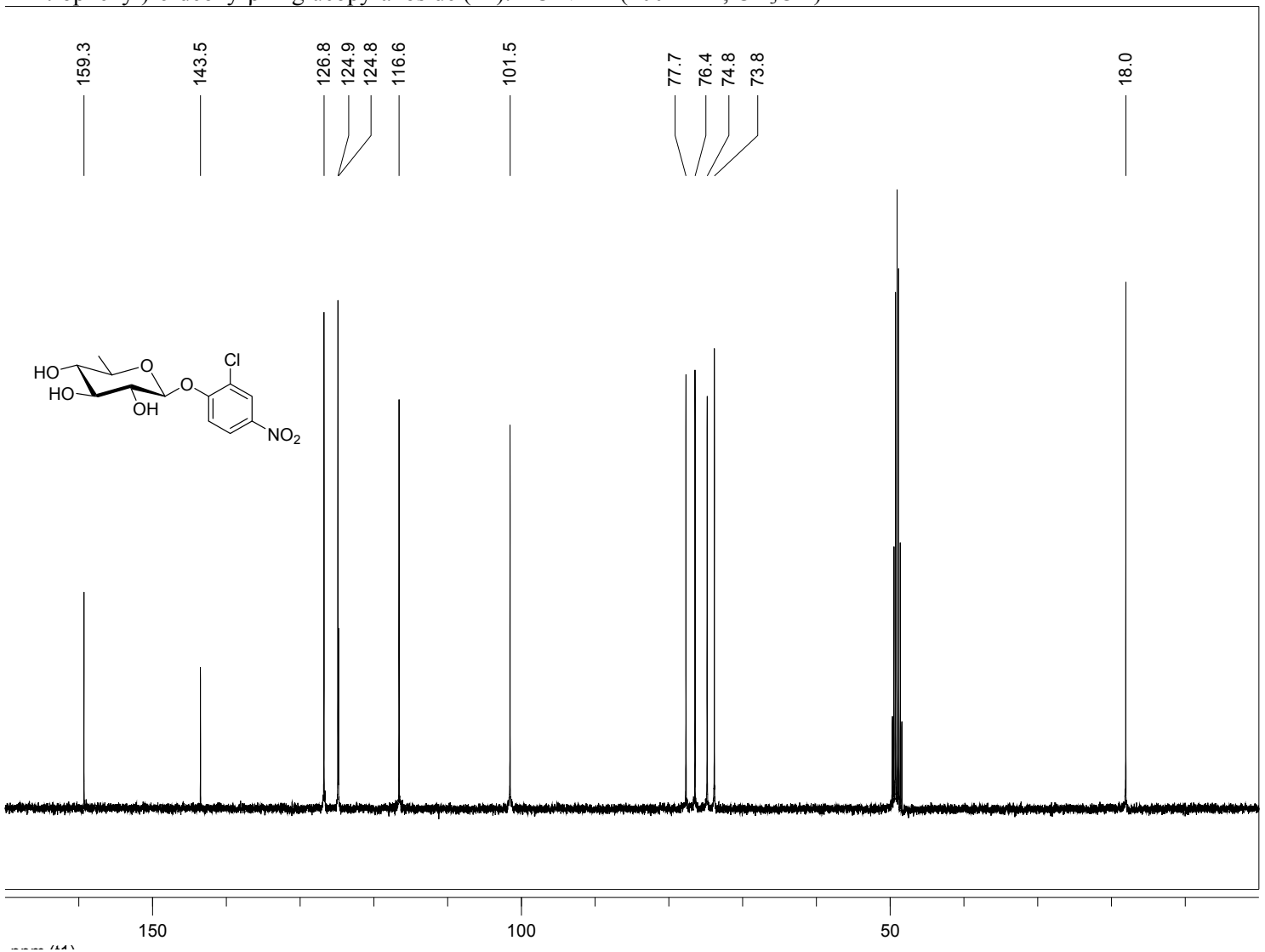
(2-chloro-4-nitrophenyl)-3-deoxy-β-D-glucopyranoside (**40**): ^{13}C NMR (100 MHz, CD_3OD)



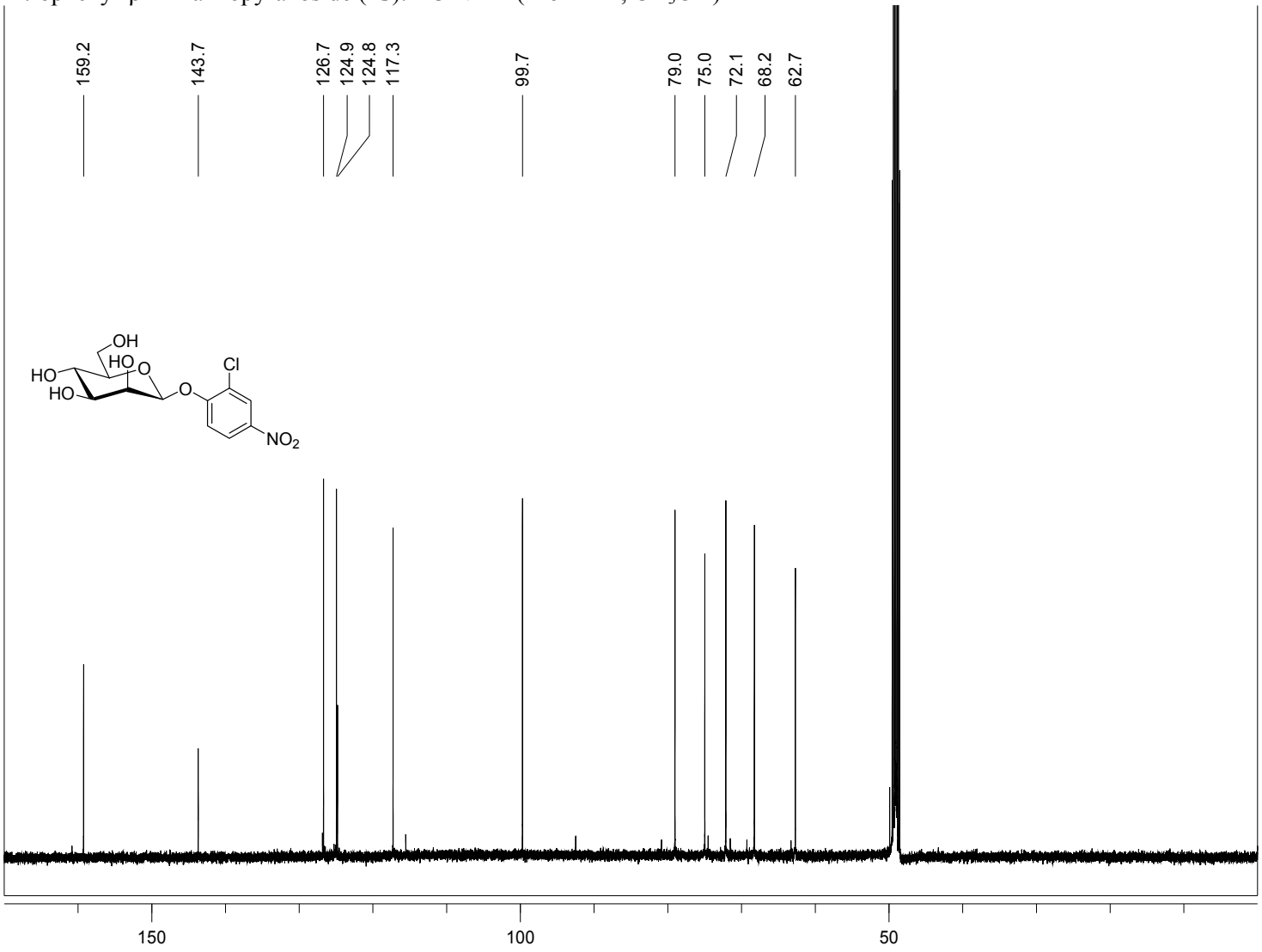
(2-chloro-4-nitrophenyl)-4-deoxy-β-D-glucopyranoside (**41**): ^{13}C NMR (100 MHz, CD_3OD)



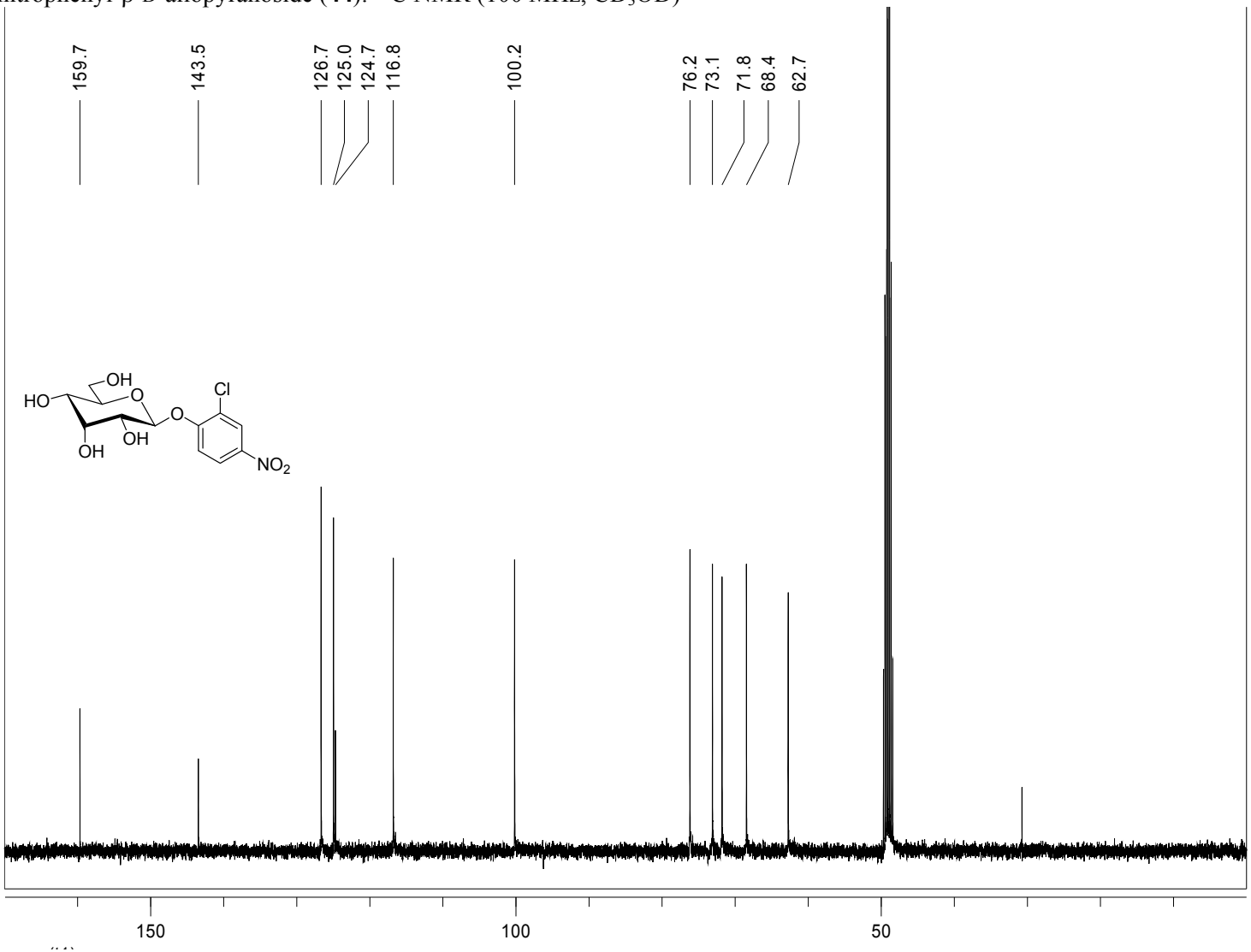
(2-chloro-4-nitrophenyl)-6-deoxy-β-D-glucopyranoside (**42**): ^{13}C NMR (100 MHz, CD_3OD)



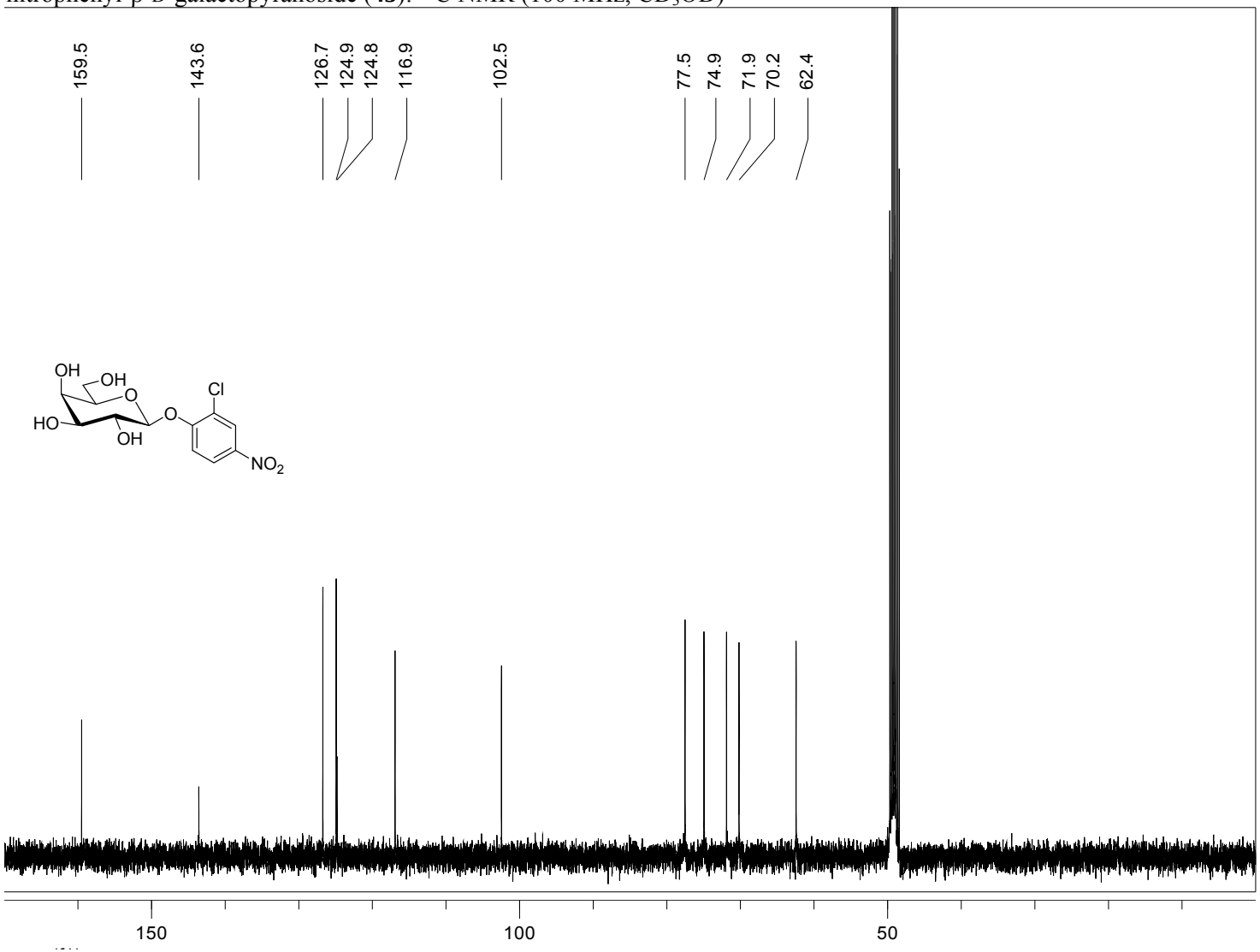
2-chloro-4-nitrophenyl-β-D-mannopyranoside (**43**): ^{13}C NMR (125 MHz, CD_3OD)



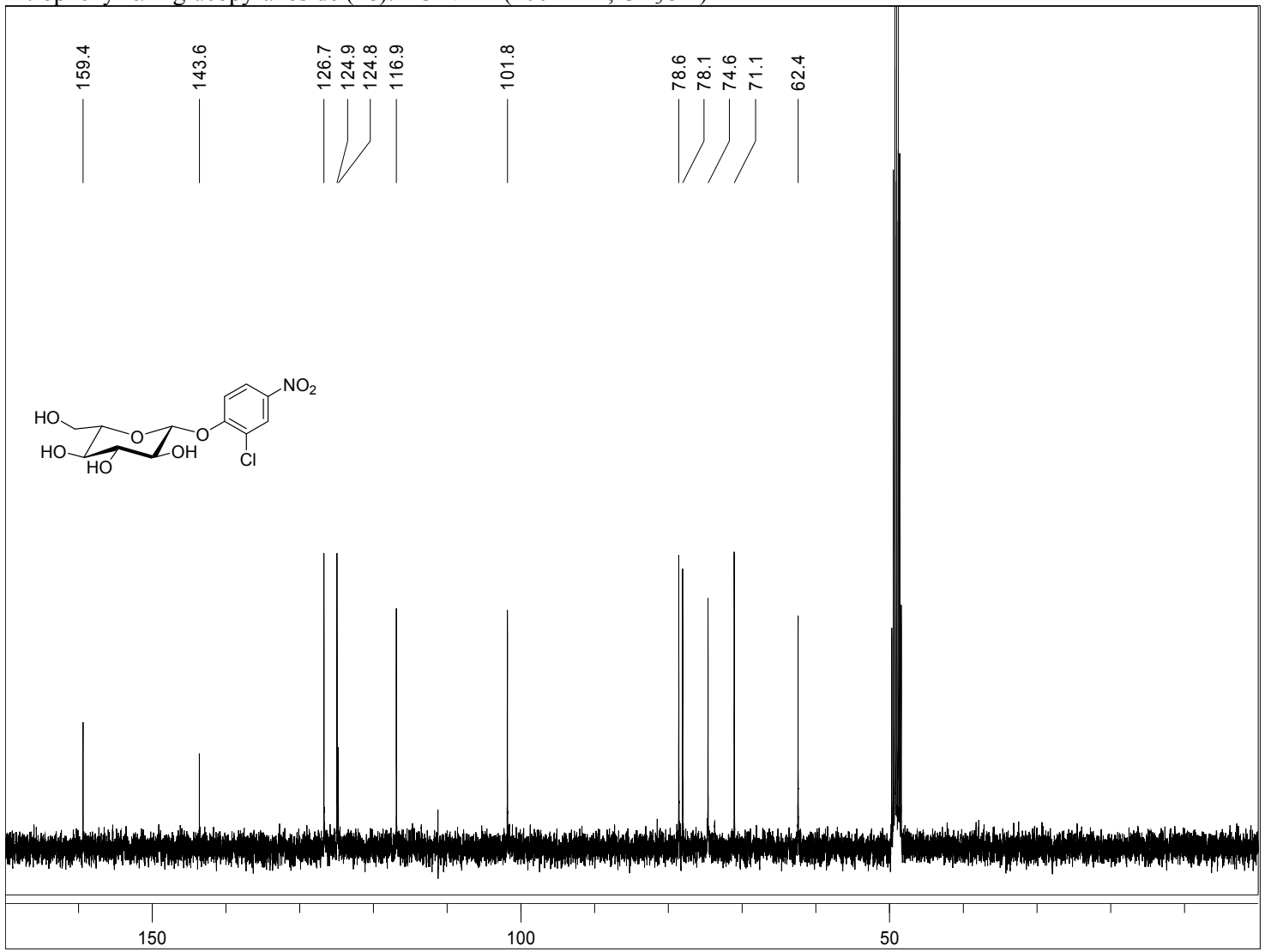
2-chloro-4-nitrophenyl-β-D-allopyranoside (**44**): ^{13}C NMR (100 MHz, CD_3OD)



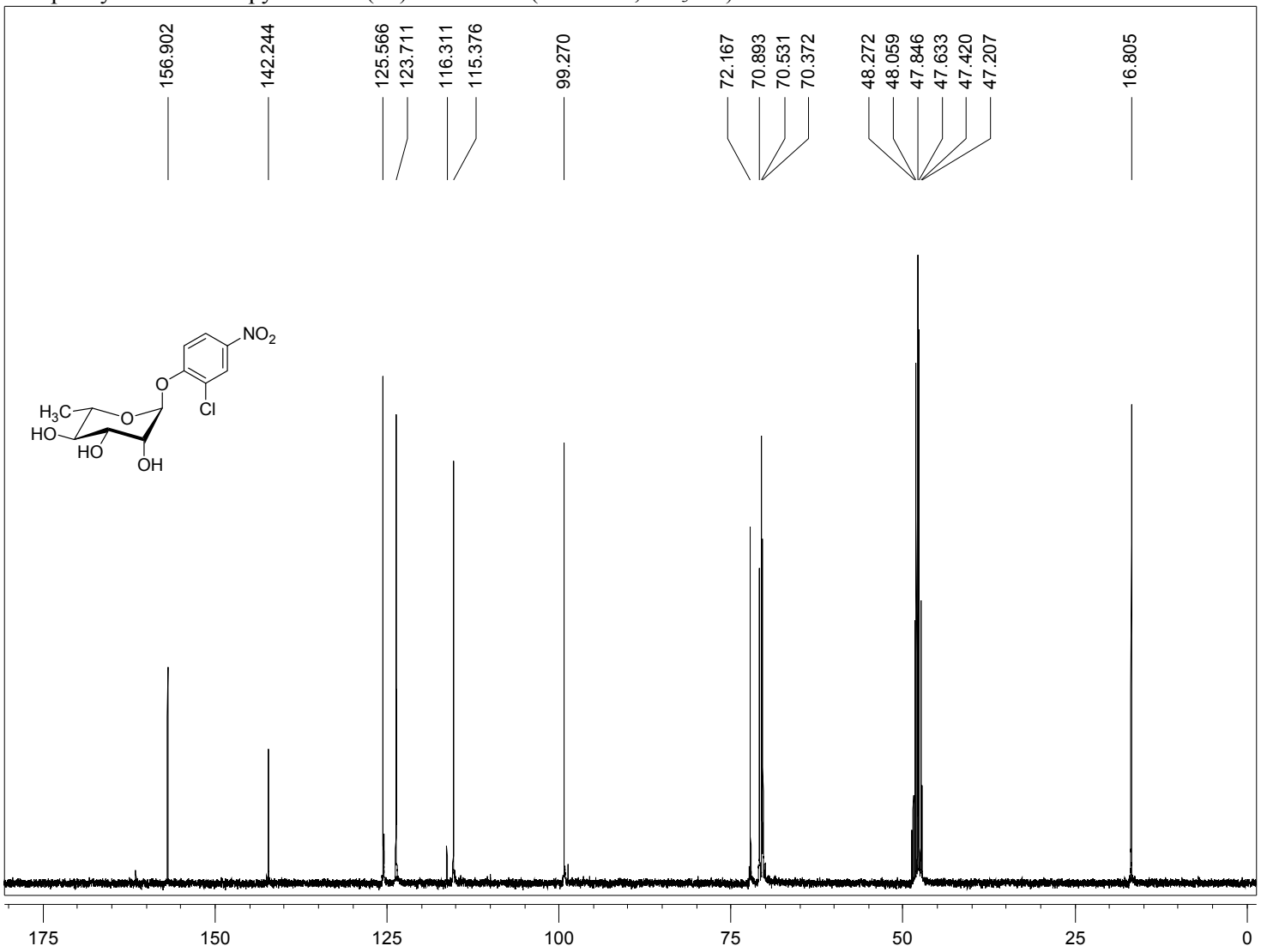
2-chloro-4-nitrophenyl-β-D-galactopyranoside (**45**): ^{13}C NMR (100 MHz, CD_3OD)



2-chloro-4-nitrophenyl-α-L-glucopyranoside (**46**): ^{13}C NMR (100 MHz, CD_3OD)

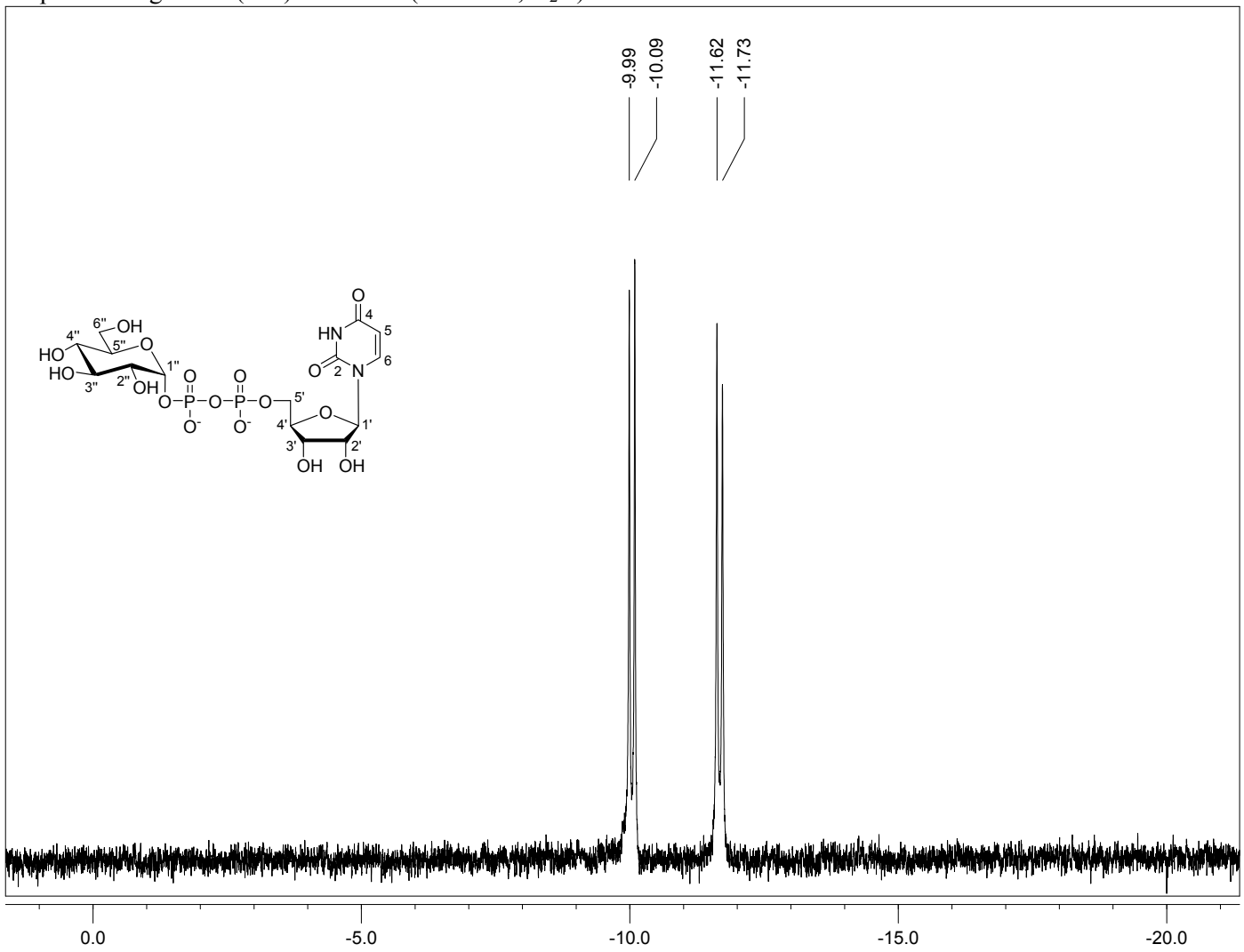


2-chloro-4-nitrophenyl-α-L-rhamnopyranoside (**47**): ^{13}C NMR (100 MHz, CD_3OD)

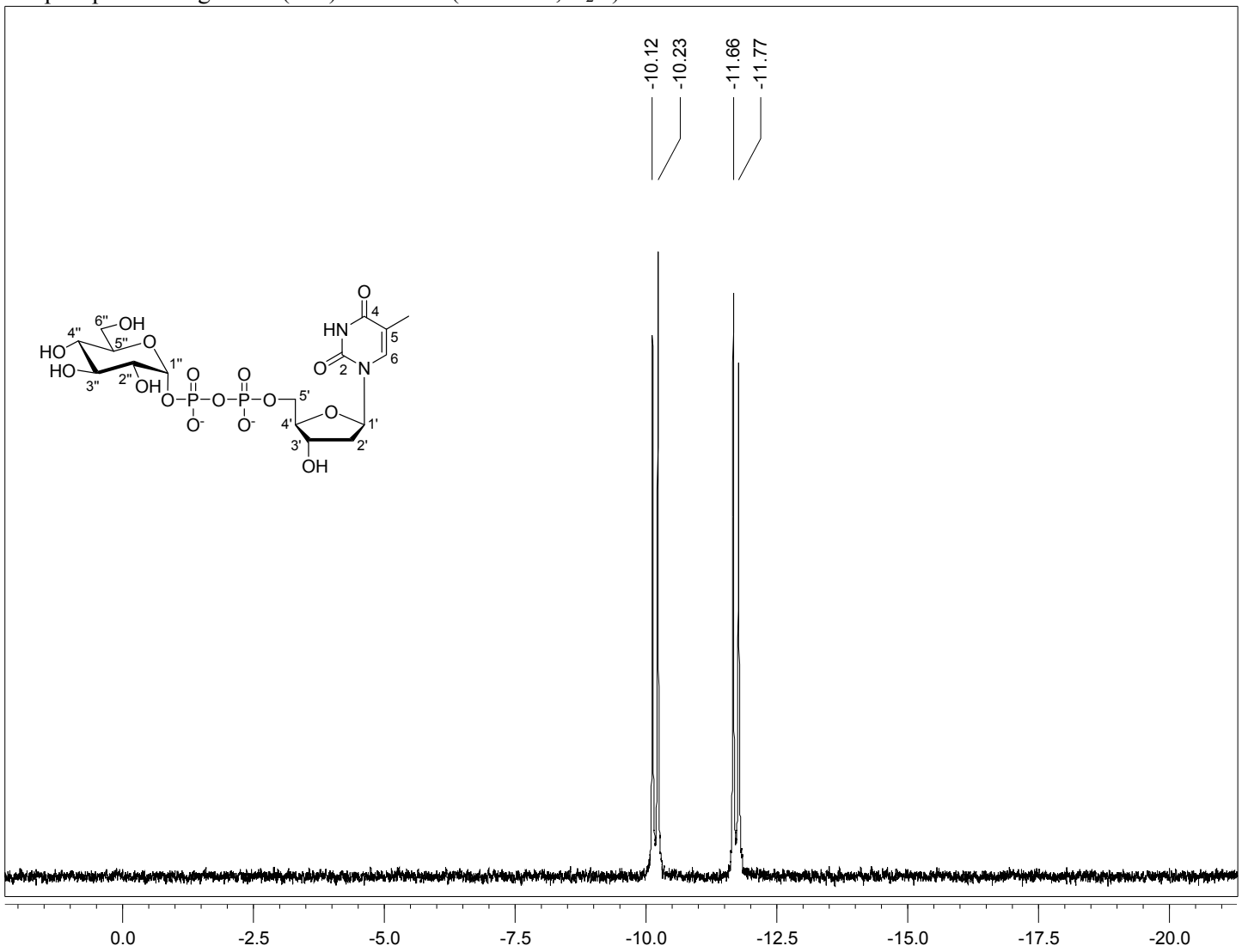


A2.4. ^{31}P NMR Spectra

Uridine 5'-diphosphate α-D-glucose (**33a**): ^{31}P NMR (500 MHz, D_2O)



Thymidine 5’-diphosphate α-D-glucose (**33b**): ^{31}P NMR (500 MHz, D_2O)



A2.5. Appendix 2 References

1. Yang, M. *et al.*, 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**, 9336-9937 (2005).

2. Williams, G.J., Zhang, C., & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

3. Gantt, R.W., Goff, R.D., Williams, G.J., & Thorson, J.S. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew. Chem. Int. Ed.* **47**, 8889-8892 (2008).

Appendix 3:

Supplementary Data for Chapter 4

A3.1. Master Plate Layouts

X	1	2	3	4	5	6	7	8	9	10	11
A	067A	067I	067R	074D	074V	074M	085G	085Q	111F	111Q	112A
B	067C	067K	067S	074E	074N	074Y	085H	085R	111G	111R	112C
C	067D	067L	067V	074G	074P	085A	085I	085T	111I	111S	112D
D	E.V.	WT	ASP	1C9	TDP-16	TDP-16	085K	085V	111K	111T	112E
E	067E	067M	067W	074H	074Q	085C	085L	085Y	111L	111V	112F
F	067F	067N	067Y	074I	074R	085D	085M	111A	111M	111W	112G
G	067G	067P	074A	074K	074S	085E	085N	111C	111N	111Y	112H
H	067H	067Q	074C	074L	074T	085F	085P	111E	111P	111M	112K

Y	1	2	3	4	5	6	7	8	9	10	11
A	112W	113G	113P	132A	132K	132S	134L	134H	134T	242I	242T
B	112Y	113H	113Q	132C	132L	132T	134E	134I	242A	242K	242V
C	113A	113I	113R	132D	132M	132V	134S	134M	242C	242M	242W
D	E.V.	WT	ASP	1C9	TDP-16	TDP-16	134R	134K	242D	242N	242Y
E	113C	113K	113S	132E	132N	132W	134C	134D	242E	242P	243A
F	113D	113L	113V	132F (silent mutation)	132P	132Y	134Y	134V	242F	242Q	243C
G	113E	113M	113W	132G	132Q	134G	134Q	134W	242G	242R	243D
H	113F	113N	113Y	132H	132R	134A	134P	134F	242H	242S	243E

Z	1	2	3	4	5	6	7	8	9	10	11
A	243Q	268A	268I	268R	309D	309M	309W	330H	330R	331D	331M
B	243R	268C	268K	268S	309E	309N	309Y	330I	330S	331E	331N
C	243S	268D	268L	268T	309F	309P	330A	330K	330T	331F	331P
D	E.V.	WT	ASP	1C9	TDP-16	TDP-16	3-1H12	330L	330V	331G	331R
E	243T	268E	268M	268W	309H	309Q	330C	330M	330W	331H	331S
F	243V	268F	268N	268Y	309I	309R	330E	330N	330Y	331I	331T
G	243W	268G	268P	309G	309K	309T	330F	330P	331A	331K	331V
H	243Y	268H	268Q	309C	309L	309V	330G	330Q	331C	331L	331W

Figure A3.1. Master plate designations for all strains carrying OleD variants. The amino acid position and mutation introduced into the parental gene sequence of *oleD*[P67T/S132F/A242L/Q268V] (generates OleD variant TDP-16)[1] is listed within each well. E.V., empty vector; WT, wild type OleD; ASP, P67T/S132F/A242V[2]; 1C9, P67T/I112K/A242V[3]; 3-1H12, P67T/S132F/A242L[4].

A3.2. Summary of OleD Variants

Note: Mixed bases denoted within primers in the following tables (**A3.1** to **A3.16**) conform to the nomenclature recommendations from the International Union of Biochemistry[5].

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
67	A	X	A01	ACC	GCG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	C	X	B01		TGT	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	D	X	C01		GAT	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	E	X	E01		GAG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	F	X	F01		TTT	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	G	X	G01		GGG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	H	X	H01		AAG	*YAY* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	I	X	A02		ATT	*AWY* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	K	X	B02		AAG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	L	X	C02		TTG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	M	X	E02		ATG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	N	X	F02		AAC	*AWY* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	P	X	G02		CCG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	Q	X	H02		CAG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	R	X	A03		CGG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	S	X	B03		TCT	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	T (parental)	---	---		---	---	---
67	V	X	C03		GTG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	W	X	E03		TGG	*NNK* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG
67	Y	X	F03		TAC	*YAY* GACGCCGACCCGGAGG	GCCGGGCAGGGTGGAGTG

Table A3.1. Summary of TDP-16 variants obtained with mutations at amino acid position 67. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
74	A	X	G03	TGG	GCG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	C	X	H03		TGT	*TDY* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	D	X	A04		GAT	*GAY* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	E	X	B04		GAG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	F	n.a.	n.a.		---	---	---
74	G	X	C04		GGG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	H	X	E04		CAC	*CMN* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	I	X	F04		ATC	*AYY* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	K	X	G04		AAA	*AAM* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	L	X	H04		TTG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	M	X	A06		ATG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	N	X	B05		AAC	*AAM* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	P	X	C05		CCG	*CMN* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	Q	X	E05		CAG	*CMN* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	R	X	F05		AGG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	S	X	G05		TCG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	T	X	H05		ACT	*AYY* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	V	X	A05		GTG	*NNK* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC
74	W (parental)	---	---		---	---	---
74	Y	X	B06		TAC	*TDY* GGAAGCACCCTGCTGGAC	TGCCTCCGGGTCGGC

Table A3.2. Summary of TDP-16 variants obtained with mutations at amino acid position 74. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
85	A	X	C06	TTC	GCG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	C	X	E06		TGC	*TRY* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	D	X	F06		GAT	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	E	X	G06		GAG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	F	X	H06		TTT	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	G	X	A07		GGG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	H	X	B07		CAT	*CAY* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	I	X	C07		ATT	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	K	X	D07		AAG	*AAN* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	L	X	E07		CTT	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	M	X	F07		ATG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	N	X	G07		AAC	*AAN* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	P	X	H07		CCG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	Q	X	A08		CAG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	R	X	B08		AGG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	S (parental)	---	---		---	---	---
85	T	X	C08		ACG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	V	X	D08		GTG	*NNK* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG
85	W	n.a.	n.a.		---	---	---
85	Y	X	E08		TAT	*TRY* CTGAACGACGCGATCCAGG	CGGTTCGACGTTGTCCAGCAGG

Table A3.3. Summary of TDP-16 variants obtained with mutations at amino acid position 85. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
111	A	X	F08	GAC	GCG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	C	X	G08		TGC	*TKY* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	D (parental)	---	---		---	---	---
111	E	X	H08		GAC	*GAR* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	F	X	A09		TTC	*TKY* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	G	X	B09		GGT	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	H	n.a.	n.a.		---	---	---
111	I	X	C09		ATT	*AWB* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	K	X	D09		AAG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	L	X	E09		CTG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	M	X	F09, H10		ATG	*AWB* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	N	X	G09		AAC	*AWB* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	P	X	H09		CCG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	Q	X	A10		CAG	*CAN* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	R	X	B10		AGG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	S	X	C10		AGT	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	T	X	D10		ACG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	V	X	E10		GTG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	W	X	F10		TGG	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC
111	Y	X	G10		TAT	*NNK* ATCACCTCCTACCCGGCCG	GTGCAGGACGAGATCGGGGATGTC

Table A3.4. Summary of TDP-16 variants obtained with mutations at amino acid position 111. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
112	A	X	A11	ATC	GCG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	C	X	B11		TGT	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	D	X	C11		GAT	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	E	X	D11		GAG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	F	X	E11		TTT	*TWY* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	G	X	F11		GGG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	H	X	G11		CAC	*CMY* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	I (parental)	---	---		---	---	---
112	K	X	H11		AAG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	L	X	A12		CTG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	M	n.a.	n.a.		---	---	---
112	N	X	B12		AAT	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	P	X	C12		CCT	*CMY* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	Q	X	D12		CAG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	R	X	E12		CGG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	S	X	F12		TCG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	T	X	G12		ACG	*AYG* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	V	X	H12		GTG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	W	Y	A01		TGG	*NNK* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG
112	Y	Y	A02		TAC	*TWY* ACCTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG

Table A3.5. Summary of TDP-16 variants obtained with mutations at amino acid position 112. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
113	A	Y	C01	ACC	GCG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	C	Y	E01		TGT	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	D	Y	F01		GAT	*GAY* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	E	Y	G01		GAG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	F	Y	H01		TTT	*TWY* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	G	Y	A02		GGG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	H	Y	B02		CAT	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	I	Y	C02		ATT	*AWB* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	K	Y	E02		AAG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	L	Y	F02		TTG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	M	Y	G02		ATG	*AWB* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	N	Y	H02		AAT	*AWB* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	P	Y	A03		CCG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	Q	Y	B03		CAG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	R	Y	C03		CGT	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	S	Y	E03		TCT	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	T (parental)	---	---		---	---	---
113	V	Y	F03		GTG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	W	Y	G03		TGG	*NNK* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG
113	Y	Y	H03		TAC	*TWY* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG

Table A3.6. Summary of TDP-16 variants obtained with mutations at amino acid position 113. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
132	A	Y	A04	TTC	GCG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	C	Y	B04		TCG	*TDY* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	D	Y	C04		GAC	*GAN* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	E	Y	E04		GAG	*GAN* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	F (parental)	Y	F04		TTT	*TDY* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	G	Y	G04		GGT	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	H	Y	H04		CAT	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	I	n.a.	n.a.		---	---	---
132	K	Y	A05		AAG	*AAN* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	L	Y	B05		CTG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	M	Y	C05		ATG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	N	Y	E05		AAC	*AAN* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	P	Y	F05		CCG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	Q	Y	G05		CAG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	R	Y	H05		CGT	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	S	Y	A06		AGT	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	T	Y	B06		ACG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	V	Y	C06		GTG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	W	Y	E06		TGG	*NNK* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA
132	Y	Y	F06		TAT	*TDY* CCGAACCTCGTCGCCTGGAAG	GAGGGAGACCGCCGGGA

Table A3.7. Summary of TDP-16 variants obtained with mutations at amino acid position 132. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
134	A	Y	H06	AAC	GCG	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	C	Y	E07		TGT	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	D	Y	E08		GAC	*GWY* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	E	Y	B07		GAG	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	F	Y	H08		TTT	*TWY* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	G	Y	G06		GGG	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	H	Y	A08		CAT	*CMN* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	I	Y	B08		ATC	*ATB* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	K	Y	D08		AAA	*AMR* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	L	Y	A07		CTG	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	M	Y	C08		ATG	*ATB* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	N (parental)	---	---		---	---	---
134	P	Y	H07		CCT	*CMN* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	Q	Y	G07		CAA	*CMN* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	R	Y	D07		AGG	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	S	Y	C07		AGT	*NNK* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	T	Y	A09		ACG	*AMR* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	V	Y	F08		GTT	*GWY* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	W	Y	G08		TGG	*TGG* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC
134	Y	Y	F07		TAC	*TWY* CTCGTCGCCTGGAAGGGTTACG	CGGGAAGAGGGAGACCGCC

Table A3.8. Summary of TDP-16 variants obtained with mutations at amino acid position 134. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
242	A	Y	B09	CTT	GCG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	C	Y	C09		TGT	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	D	Y	D09		GAT	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	E	Y	E09		GAG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	F	Y	F09		TTT	*THY* TTCACCAAGCAGCCCGCCTTCTAC	CGAGCCGAGCGACACCAGG
242	G	Y	G09		GGG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	H	Y	H09		CAT	*CAY* TTCACCAAGCAGCCCGCCTTCTAC	CGAGCCGAGCGACACCAGG
242	I	Y	A10		ATT	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	K	Y	B10		AAG	*AAN* TTCACCAAGCAGCCCGCCTTCTAC	CGAGCCGAGCGACACCAGG
242	L (parental)	---	---		---	---	---
242	M	Y	C10		ATG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	N	Y	D10		AAT	*AAN* TTCACCAAGCAGCCCGCCTTCTAC	CGAGCCGAGCGACACCAGG
242	P	Y	E10		CCG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	Q	Y	F10		CAG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	R	Y	G10		AGG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	S	Y	H10		TCT	*THY* TTCACCAAGCAGCCCGCCTTCTAC	CGAGCCGAGCGACACCAGG
242	T	Y	A11		ACG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	V	Y	B11		GTT	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	W	Y	C11		TGG	*NNK* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG
242	Y	Y	D11		TAC	*THY* TTCACCAAGCAGCCCGCCTTCTAC	CGAGCCGAGCGACACCAGG

Table A3.9. Summary of TDP-16 variants obtained with mutations at amino acid position 242. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
243	A	Y	E11	TTC	GCG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	C	Y	F11		TGT	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	D	Y	G11		GAC	*GWY* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	E	Y	H11		GAG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	F (parental)	---	---		---	---	---
243	G	Y	A12		GGG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	H	Y	B12		CAC	*CAN* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	I	Y	C12		ATC	*AWY* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	K	Y	D12		AAG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	L	Y	E12		CTG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	M	Y	F12		ATG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	N	Y	G12		AAT	*AWY* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	P	Y	H12		CCT	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	Q	Z	A01		CAG	*CAN* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	R	Z	B01		CGT	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	S	Z	C01		TCC	*TMY* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	T	Z	E01		ACG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	V	Z	F01		GTG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	W	Z	G01		TGG	*NNK* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG
243	Y	Z	H01		TAT	*TMY* ACCAAGCAGCCCGCCTTCTAC	AAGCGAGCCGAGCGACACCAG

Table A3.10. Summary of TDP-16 variants obtained with mutations at amino acid position 243. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
268	A	Z	A02	GTT	GCG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	C	Z	B02		TGC	*TDY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	D	Z	C02		GAC	*GAY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	E	Z	E02		GAG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	F	Z	F02		TTC	*TDY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	G	Z	G02		GGG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	H	Z	H02		CAC	*CAY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	I	Z	A03		ATT	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	K	Z	B03		AAG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	L	Z	C03		TTG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	M	Z	E03		ATG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	N	Z	F03		AAT	*AMY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	P	Z	G03		CCG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	Q	Z	H03		CAG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	R	Z	A04		CGG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	S	Z	B04		TCG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	T	Z	C04		ACC	*AMY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	V (parental)	---	---		---	---	---
268	W	Z	E04		TGG	*NNK* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG
268	Y	Z	F04		TAT	*TDY* ATCGGCCGGAAGGTGACCCC	GAGGACGAGGTGCCAGCCG

Table A3.11. Summary of TDP-16 variants obtained with mutations at amino acid position 268. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
290	A	Z	B12	TGG	GCG	*NNK* GTGCCGCAGCTCGCGATCCTG	GTCGTGTACCTCCACGTTGTCCGGC
290	C	Z	C12		TGT	*NNK* GTGCCGCAGCTCGCGATCCTG	GTCGTGTACCTCCACGTTGTCCGGC
290	D	n.a.	---		---	---	---
290	E	n.a.	---		---	---	---
290	F	n.a.	---		---	---	---
290	G	Z	D12		GGT	*NNK* GTGCCGCAGCTCGCGATCCTG	GTCGTGTACCTCCACGTTGTCCGGC
290	H	n.a.	---		---	---	---
290	I	Z	E12		ATT	*NNK* GTGCCGCAGCTCGCGATCCTG	GTCGTGTACCTCCACGTTGTCCGGC
290	K	n.a.	---		---	---	---
290	L	n.a.	---		---	---	---
290	M	n.a.	---		---	---	---
290	N	n.a.	---		---	---	---
290	P	n.a.	---		---	---	---
290	Q	n.a.	---		---	---	---
290	R	Z	F12		CGG	*NNK* GTGCCGCAGCTCGCGATCCTG	GTCGTGTACCTCCACGTTGTCCGGC
290	S	n.a.	---		---	---	---
290	T	Z	G12		ACT	*NNK* GTGCCGCAGCTCGCGAT	GTCGTGTACCTCCACGTTGTCCGGC
290	V	n.a.	---		---	---	---
290	W (parental)	---	---		---	---	---
290	Y	Z	H12		TAT	*NNK* GTGCCGCAGCTCGCGATCCTG	GTCGTGTACCTCCACGTTGTCCGGC

Table A3.12. Summary of TDP-16 variants obtained with mutations at amino acid position 290. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
309	A (parental)	---	---		---	---	---
309	C	Z	H04		TGT	*TNY* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	D	Z	A05		GAC	*GAY* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	E	Z	B05		GAG	*NNK* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	F	Z	C05		TTC	*TNY* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	G	Z	G04		GGG	*NNK* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	H	Z	E05		CAC	*CMN* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	I	Z	F05		ATC	*ATB* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	K	Z	G05		AAA	*AMM* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	L	Z	H05	GCC	CTG	*NNK* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	M	Z	A06		ATG	*ATB* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	N	Z	B06		AAC	*AMM* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	P	Z	C06		CCG	*CMN* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	Q	Z	E06		CAG	*CMN* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	R	Z	F06		AGG	*NNK* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	S	n.a.	---		---	---	---
309	T	Z	G06		ACC	*AMM* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	V	Z	H06		GTG	*NNK* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	W	Z	A07		TGG	*NNK* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG
309	Y	Z	B07		TAC	*TNY* GGCGGCAGCCAGGAG	GCCCGCGTGGGTGACG

Table A3.13. Summary of TDP-16 variants obtained with mutations at amino acid position 309. Within primers, nucleotide bases in *italics* denote targeted positions for mutation. n.a.; not available

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
330	A	Z	C07	GAC	GCG	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	C	Z	E07		TGC	*TDY* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	D (parental)	---	---		---	---	---
330	E	Z	F07		GAG	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	F	Z	G07		TTT	*TDY* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	G	Z	H07		GGT	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	H	Z	A08		CAC	*CAY* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	I	Z	B08		ATT	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	K	Z	C08		AAG	*AMN* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	L	Z	D08		CTG	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	M	Z	E08		ATG	*ATG* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	N	Z	F08		AAT	*AMN* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	P	Z	G08		CCG	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	Q	Z	H08		CAG	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	R	Z	A09		CGG	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	S	Z	B09		TCT	*NNK* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	T	Z	C09		ACC	*AMN* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	V	Z	D09		GTG	*GTN* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	W	Z	E09		TGG	*TGG* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG
330	Y	Z	F09		TAT	*TDY* CAGTTCGGCAACGCCGAC	GACGGCCTGCGGTACGG

Table A3.14. Summary of TDP-16 variants obtained with mutations at amino acid position 330. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Amino Acid Position	Amino Acid Change	Master Plate Designation	Master Plate Well	Parental Codon	Mutant Codon	Primer 1 (5' --> 3')	Primer 2 (5' --> 3')
331	A	Z	G09	CAG	GCA	*GMR* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	C	Z	H09		TGT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	D	Z	A10		GAT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	E	Z	B10		GAG	*GMR* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	F	Z	C10		TTT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	G	Z	D10		GGT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	H	Z	E10		CAT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	I	Z	F10		ATT	*AWY* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	K	Z	G10		AAG	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	L	Z	H10		CTG	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	M	Z	A11		AAT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	N	Z	B11		ATG	*AWY* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	P	Z	C11		CCT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	Q (parental)	---	---		---	---	---
331	R	Z	D11		CGG	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	S	Z	E11		TCT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	T	Z	F11		ACG	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	V	Z	G11		GTT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	W	Z	H11		TGG	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG
331	Y	Z	A12		TAT	*NNK* TTCGGCAACGCCGACATG	GTCGACGGCCTGCGGTACG

Table A3.15. Summary of TDP-16 variants obtained with mutations at amino acid position 331. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

Master Plate Designation (strain utilized for plasmid DNA isolation)								
Plate	**Well**	**Existing Mutant Amino Acid Positions in Template DNA**	**Desired Mutation(s)**	**Targeted Codon(s) in Template DNA**	**Primer 1 (5' --> 3')**	**Primer 2 (5' --> 3')**	**Confirmed Targeted Mutant Codon(s)**	**Final Mutated Amino Acid Positions in Desired Recombinant**
Y	G02	T113M	I112P	ATC	*CCN* ATGTCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG	CCG	I112P/T113M
X	C12	I112P	L242I	CTT	*ATH* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG	ATT	I112P/L242I
Y	A10	L242I	T113M	ACC	*ATG* TCCTACCCGGCCCGC	GATGTCGTGCAGGACGAGATCGGG	ATG	T113M/L242I
Z	H03	V268Q	T113M	ACC	*ATG* TCCTACCCGGCCCGCG	GATGTCGTGCAGGACGAGATCGGG	ATG	T113M/V268Q
Y	A10	L242I	V268Q	GTT	*ATH* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG	CAG	L242I/V268Q
Y	A10	L242I	I112P & T113M	ATCACC	*CCNATG* TCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG	CCGATG	I112P/T113M/L242I
Z	H03	V268Q	I112P & T113M	ATCACC	*CCNATG* TCCTACCCGGCCCG	GTCGTGCAGGACGAGATCGGGG	CCGATG	I112P/T113M/V268Q
n.a.	n.a.	I112P/T113M/V268Q	L242I	CTT	*ATH* TTCACCAAGCAGCCCGCCTTC	CGAGCCGAGCGACACCAGG	ATC	I112P/T113M/L242I/V268Q (*i.e.*, Loki variant)

Table A3.16. Summary of recombinants generated from secondary 'hit' mutations. Within primers, nucleotide bases in *italics* denote targeted positions for mutation.

A3.3. DNA Sequences

The sequence alignment of wild-type OleD can be found within the Entrez Protein database under accession code ABA42119. In the following sequences, differences from the OleD wild-type are highlighted in **bold** font and underlined.

A3.3.1. OleD variant TDP-16

```
1     ATG ACC ACC CAG ACC ACT CCC GCC CAC ATC GCC ATG TTC TCC ATC   45
1     Met Thr Thr Gln Thr Thr Pro Ala His Ile Ala Met Phe Ser Ile   15

46    GCC GCC CAC GGC CAT GTG AAC CCC AGC CTG GAG GTG ATC CGT GAA   90
16    Ala Ala His Gly His Val Asn Pro Ser Leu Glu Val Ile Arg Glu   30

91    CTC GTC GCC CGC GGC CAC CGG GTC ACG TAC GCC ATT CCG CCC GTC   135
31    Leu Val Ala Arg Gly His Arg Val Thr Tyr Ala Ile Pro Pro Val   45

136   TTC GCC GAC AAG GTG GCC GCC ACC GGC GCC CGG CCC GTC CTC TAC   180
46    Phe Ala Asp Lys Val Ala Ala Thr Gly Ala Arg Pro Val Leu Tyr   60

181   CAC TCC ACC CTG CCC GGC ACC GAC GCC GAC CCG GAG GCA TGG GGA   225
61    His Ser Thr Leu Pro Gly Thr Asp Ala Asp Pro Glu Ala Trp Gly   75

226   AGC ACC CTG CTG GAC AAC GTC GAA CCG TTC CTG AAC GAC GCG ATC   270
76    Ser Thr Leu Leu Asp Asn Val Glu Pro Phe Leu Asn Asp Ala Ile   90

271   CAG GCG CTC CCG CAG CTC GCC GAT GCC TAC GCC GAC GAC ATC CCC   315
91    Gln Ala Leu Pro Gln Leu Ala Asp Ala Tyr Ala Asp Asp Ile Pro   105

316   GAT CTC GTC CTG CAC GAC ATC ACC TCC TAC CCG GCC CGC GTC CTG   360
106   Asp Leu Val Leu His Asp Ile Thr Ser Tyr Pro Ala Arg Val Leu   120

361   GCC CGC CGC TGG GGC GTC CCG GCG GTC TCC CTC TTC CCG AAC CTC   405
121   Ala Arg Arg Trp Gly Val Pro Ala Val Ser Leu Phe Pro Asn Leu   135

406   GTC GCC TGG AAG GGT TAC GAG GAG GAG GTC GCC GAG CCG ATG TGG   450
136   Val Ala Trp Lys Gly Tyr Glu Glu Glu Val Ala Glu Pro Met Trp   150

451   CGC GAA CCC CGG CAG ACC GAG CGC GGA CGG GCC TAC TAC GCC CGG   495
151   Arg Glu Pro Arg Gln Thr Glu Arg Gly Arg Ala Tyr Tyr Ala Arg   165

496   TTC GAG GCA TGG CTG AAG GAG AAC GGG ATC ACC GAG CAC CCG GAC   540
166   Phe Glu Ala Trp Leu Lys Glu Asn Gly Ile Thr Glu His Pro Asp   180

541   ACG TTC GCC AGT CAT CCG CCG CGC TCC CTG GTG CTC ATC CCG AAG   585
181   Thr Phe Ala Ser His Pro Pro Arg Ser Leu Val Leu Ile Pro Lys   195

586   GCG CTC CAG CCG CAC GCC GAC CGG GTG GAC GAA GAC GTG TAC ACC   630
196   Ala Leu Gln Pro His Ala Asp Arg Val Asp Glu Asp Val Tyr Thr   210

631   TTC GTC GGC GCC TGC CAG GGA GAC CGC GCC GAG GAA GGC GGC TGG   675
211   Phe Val Gly Ala Cys Gln Gly Asp Arg Ala Glu Glu Gly Gly Trp   225

676   CAG CGG CCC GCC GGC GCG GAG AAG GTC GTC CTG GTG TCG CTC GGC   720
226   Gln Arg Pro Ala Gly Ala Glu Lys Val Val Leu Val Ser Leu Gly   240

721   TCG CTT TTC ACC AAG CAG CCC GCC TTC TAC CGG GAG TGC GTG CGC   765
241   Ser Leu Phe Thr Lys Gln Pro Ala Phe Tyr Arg Glu Cys Val Arg   255
```

```
766   GCC TTC GGG AAC CTG CCC GGC TGG CAC CTC GTC CTC GTT ATC GGC   810
256   Ala Phe Gly Asn Leu Pro Gly Trp His Leu Val Leu Val Ile Gly   270

811   CGG AAG GTG ACC CCC GCC GAA CTG GGG GAG CTG CCG GAC AAC GTG   855
271   Arg Lys Val Thr Pro Ala Glu Leu Gly Glu Leu Pro Asp Asn Val   285

856   GAG GTA CAC GAC TGG GTG CCG CAG CTC GCG ATC CTG CGC CAG GCC   900
286   Glu Val His Asp Trp Val Pro Gln Leu Ala Ile Leu Arg Gln Ala   300

901   GAT CTG TTC GTC ACC CAC GCG GGC GCC GGC GGC AGC CAG GAG GGG   945
301   Asp Leu Phe Val Thr His Ala Gly Ala Gly Gly Ser Gln Glu Gly   315

946   CTG GCC ACC GCG ACG CCC ATG ATC GCC GTA CCG CAG GCC GTC GAC   990
316   Leu Ala Thr Ala Thr Pro Met Ile Ala Val Pro Gln Ala Val Asp   330

991   CAG TTC GGC AAC GCC GAC ATG CTC CAA GGG CTC GGC GTC GCC CGG   1035
331   Gln Phe Gly Asn Ala Asp Met Leu Gln Gly Leu Gly Val Ala Arg   345

1036  AAG CTG GCG ACC GAG GAG GCC ACC GCC GAC CTG CTC CGC GAG ACC   1080
346   Lys Leu Ala Thr Glu Glu Ala Thr Ala Asp Leu Leu Arg Glu Thr   360

1081  GCC CTC GCT CTG GTG GAC GAC CCG GAG GTC GCG CGC CGG CTC CGG   1125
361   Ala Leu Ala Leu Val Asp Asp Pro Glu Val Ala Arg Arg Leu Arg   375

1126  CGG ATC CAG GCG GAG ATG GCC CAG GAG GGC GGC ACC CGG CGG GCG   1170
376   Arg Ile Gln Ala Glu Met Ala Gln Glu Gly Gly Thr Arg Arg Ala   390

1171  GCC GAC CTC ATC GAG GCC GAA CTG CCC GCG CGC CAC GAG CGG CAG   1215
391   Ala Asp Leu Ile Glu Ala Glu Leu Pro Ala Arg His Glu Arg Gln   405

1216  GAG CCG GTG GGC GAC CGA CCC AAC GGT GGG TGA                   1248
406   Glu Pro Val Gly Asp Arg Pro Asn Gly Gly End
```

A3.3.2. OleD variant Loki

```
1     ATG ACC ACC CAG ACC ACT CCC GCC CAC ATC GCC ATG TTC TCC ATC   45
1     Met Thr Thr Gln Thr Thr Pro Ala His Ile Ala Met Phe Ser Ile   15

46    GCC GCC CAC GGC CAT GTG AAC CCC AGC CTG GAG GTG ATC CGT GAA   90
16    Ala Ala His Gly His Val Asn Pro Ser Leu Glu Val Ile Arg Glu   30

91    CTC GTC GCC CGC GGC CAC CGG GTC ACG TAC GCC ATT CCG CCC GTC   135
31    Leu Val Ala Arg Gly His Arg Val Thr Tyr Ala Ile Pro Pro Val   45

136   TTC GCC GAC AAG GTG GCC GCC ACC GGC GCC CGG CCC GTC CTC TAC   180
46    Phe Ala Asp Lys Val Ala Ala Thr Gly Ala Arg Pro Val Leu Tyr   60

181   CAC TCC ACC CTG CCC GGC ACC GAC GCC GAC CCG GAG GCA TGG GGA   225
61    His Ser Thr Leu Pro Gly Thr Asp Ala Asp Pro Glu Ala Trp Gly   75

226   AGC ACC CTG CTG GAC AAC GTC GAA CCG TTC CTG AAC GAC GCG ATC   270
76    Ser Thr Leu Leu Asp Asn Val Glu Pro Phe Leu Asn Asp Ala Ile   90

271   CAG GCG CTC CCG CAG CTC GCC GAT GCC TAC GCC GAC GAC ATC CCC   315
91    Gln Ala Leu Pro Gln Leu Ala Asp Ala Tyr Ala Asp Asp Ile Pro   105

316   GAT CTC GTC CTG CAC GAC CCG ATG TCC TAC CCG GCC CGC GTC CTG   360
106   Asp Leu Val Leu His Asp Pro Met Ser Tyr Pro Ala Arg Val Leu   120

361   GCC CGC CGC TGG GGC GTC CCG GCG GTC TCC CTC TTC CCG AAC CTC   405
121   Ala Arg Arg Trp Gly Val Pro Ala Val Ser Leu Phe Pro Asn Leu   135

406   GTC GCC TGG AAG GGT TAC GAG GAG GAG GTC GCC GAG CCG ATG TGG   450
136   Val Ala Trp Lys Gly Tyr Glu Glu Glu Val Ala Glu Pro Met Trp   150

451   CGC GAA CCC CGG CAG ACC GAG CGC GGA CGG GCC TAC TAC GCC CGG   495
151   Arg Glu Pro Arg Gln Thr Glu Arg Gly Arg Ala Tyr Tyr Ala Arg   165

496   TTC GAG GCA TGG CTG AAG GAG AAC GGG ATC ACC GAG CAC CCG GAC   540
166   Phe Glu Ala Trp Leu Lys Glu Asn Gly Ile Thr Glu His Pro Asp   180

541   ACG TTC GCC AGT CAT CCG CCG CGC TCC CTG GTG CTC ATC CCG AAG   585
181   Thr Phe Ala Ser His Pro Pro Arg Ser Leu Val Leu Ile Pro Lys   195

586   GCG CTC CAG CCG CAC GCC GAC CGG GTG GAC GAA GAC GTG TAC ACC   630
196   Ala Leu Gln Pro His Ala Asp Arg Val Asp Glu Asp Val Tyr Thr   210

631   TTC GTC GGC GCC TGC CAG GGA GAC CGC GCC GAG GAA GGC GGC TGG   675
211   Phe Val Gly Ala Cys Gln Gly Asp Arg Ala Glu Glu Gly Gly Trp   225

676   CAG CGG CCC GCC GGC GCG GAG AAG GTC GTC CTG GTG TCG CTC GGC   720
226   Gln Arg Pro Ala Gly Ala Glu Lys Val Val Leu Val Ser Leu Gly   240

721   TCG ATC TTC ACC AAG CAG CCC GCC TTC TAC CGG GAG TGC GTG CGC   765
241   Ser Ile Phe Thr Lys Gln Pro Ala Phe Tyr Arg Glu Cys Val Arg   255

766   GCC TTC GGG AAC CTG CCC GGC TGG CAC CTC GTC CTC CAG ATC GGC   810
256   Ala Phe Gly Asn Leu Pro Gly Trp His Leu Val Leu Gln Ile Gly   270

811   CGG AAG GTG ACC CCC GCC GAA CTG GGG GAG CTG CCG GAC AAC GTG   855
271   Arg Lys Val Thr Pro Ala Glu Leu Gly Glu Leu Pro Asp Asn Val   285

856   GAG GTA CAC GAC TGG GTG CCG CAG CTC GCG ATC CTG CGC CAG GCC   900
286   Glu Val His Asp Trp Val Pro Gln Leu Ala Ile Leu Arg Gln Ala   300
```

```
901   GAT CTG TTC GTC ACC CAC GCG GGC GCC GGC GGC AGC CAG GAG GGG   945
301   Asp Leu Phe Val Thr His Ala Gly Ala Gly Gly Ser Gln Glu Gly   315

946   CTG GCC ACC GCG ACG CCC ATG ATC GCC GTA CCG CAG GCC GTC GAC   990
316   Leu Ala Thr Ala Thr Pro Met Ile Ala Val Pro Gln Ala Val Asp   330

991   CAG TTC GGC AAC GCC GAC ATG CTC CAA GGG CTC GGC GTC GCC CGG   1035
331   Gln Phe Gly Asn Ala Asp Met Leu Gln Gly Leu Gly Val Ala Arg   345

1036  AAG CTG GCG ACC GAG GAG GCC ACC GCC GAC CTG CTC CGC GAG ACC   1080
346   Lys Leu Ala Thr Glu Glu Ala Thr Ala Asp Leu Leu Arg Glu Thr   360

1081  GCC CTC GCT CTG GTG GAC GAC CCG GAG GTC GCG CGC CGG CTC CGG   1125
361   Ala Leu Ala Leu Val Asp Asp Pro Glu Val Ala Arg Arg Leu Arg   375

1126  CGG ATC CAG GCG GAG ATG GCC CAG GAG GGC GGC ACC CGG CGG GCG   1170
376   Arg Ile Gln Ala Glu Met Ala Gln Glu Gly Gly Thr Arg Arg Ala   390

1171  GCC GAC CTC ATC GAG GCC GAA CTG CCC GCG CGC CAC GAG CGG CAG   1215
391   Ala Asp Leu Ile Glu Ala Glu Leu Pro Ala Arg His Glu Arg Gln   405

1216  GAG CCG GTG GGC GAC CGA CCC AAC GGT GGG TGA                   1248
406   Glu Pro Val Gly Asp Arg Pro Asn Gly Gly End
```

A3.4. OleD 'Family Tree'

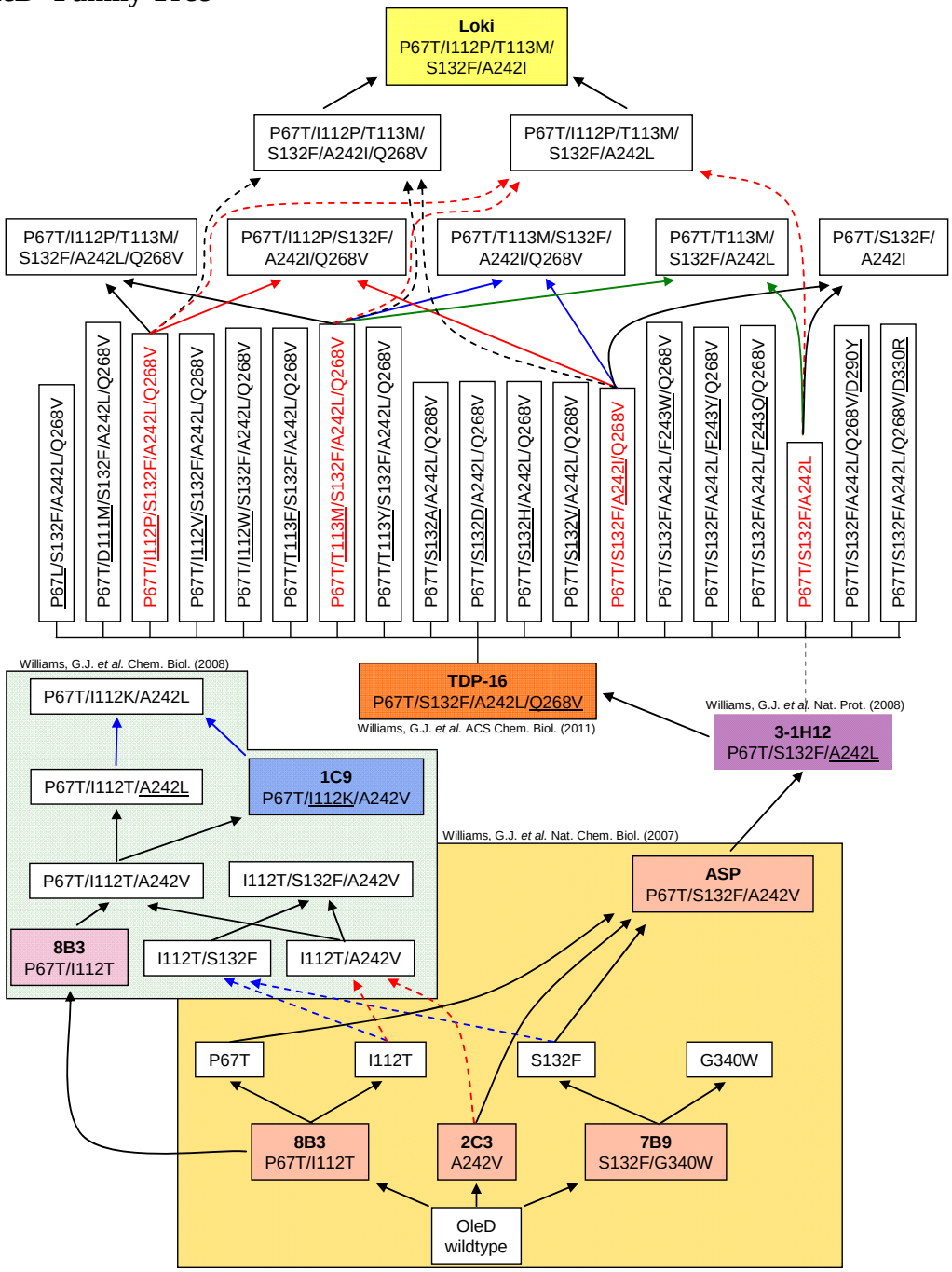
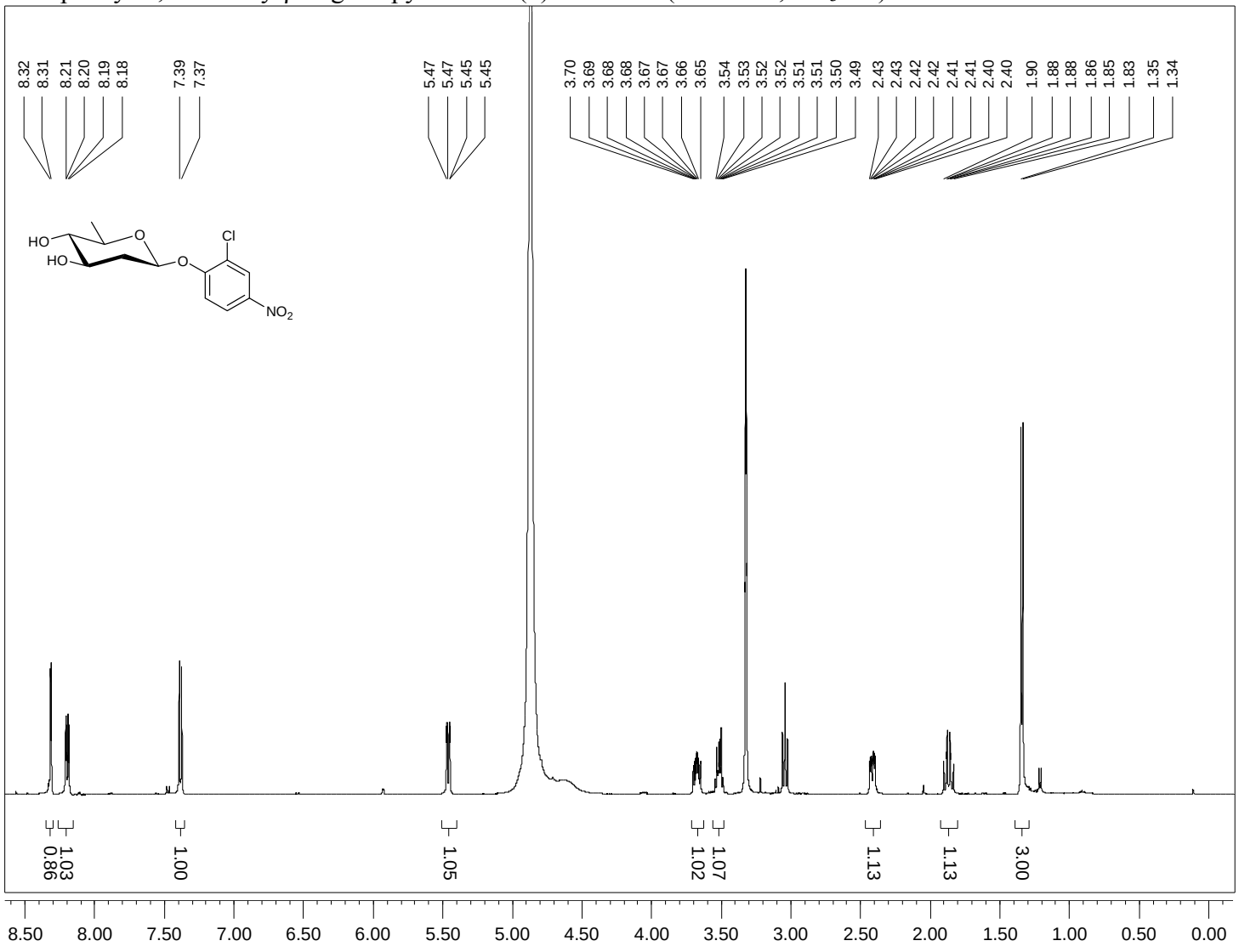


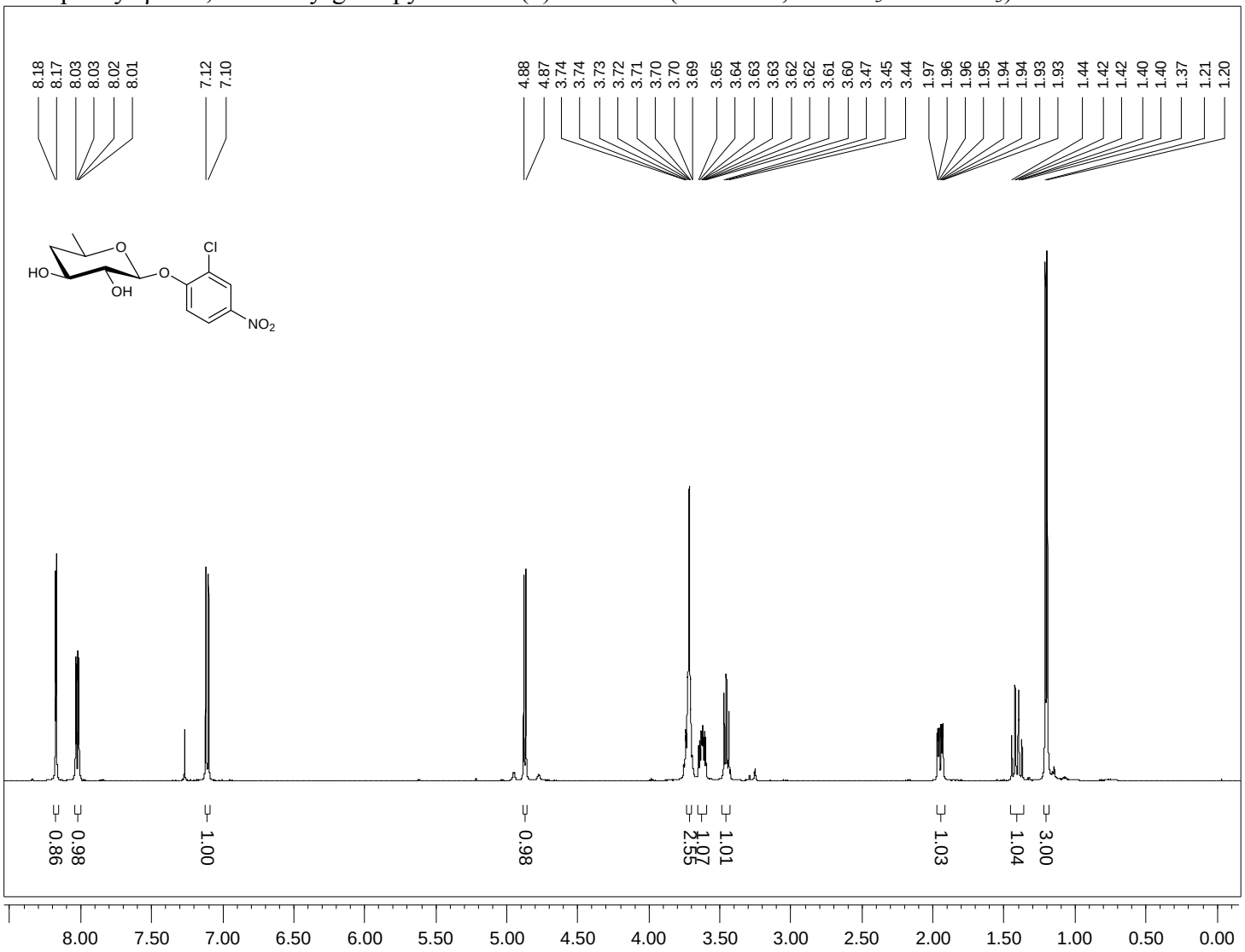
Figure A3.2. OleD 'Family Tree'. A scheme depicting a history of all of the major OleD variants which have been previously reported or generated in this study. All mutations listed are in the context of the OleD wild type sequence. Labels above mutations were the given 'nickname' for the respective sequence in the literature. Variants contained within larger colored boxes were generated and/or characterized within the same study (the literature reference is labeled on the outside of the respective boxes). Underlined mutations refer to an amino acid position which was identified or optimized through saturation mutagenesis.

A3.5. ^{1}H NMR Spectra

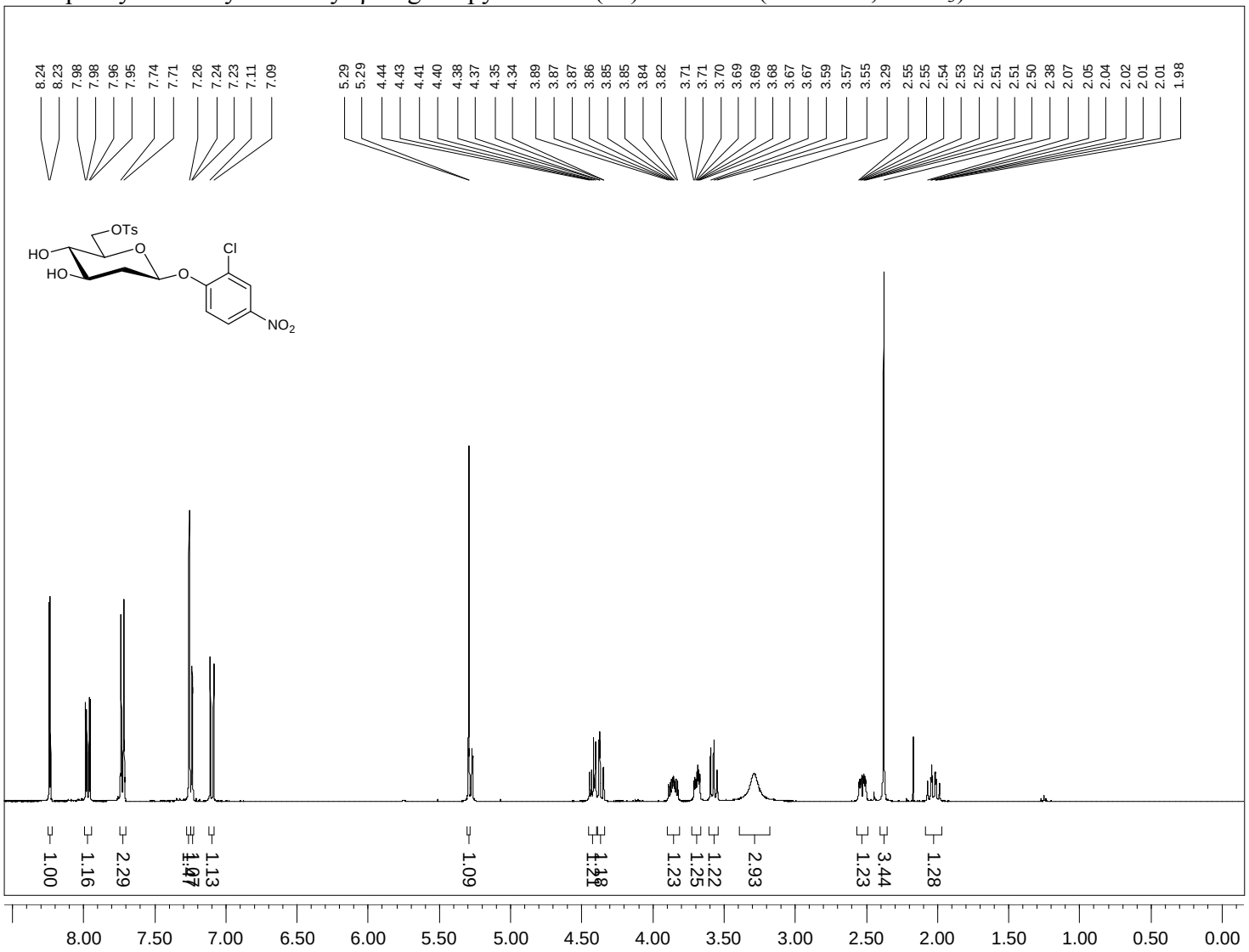
2-chloro-4-nitrophenyl-2,6-dideoxy-β-D-glucopyranoside (**4**): ^{1}H NMR (500 MHz, CD_3OD)



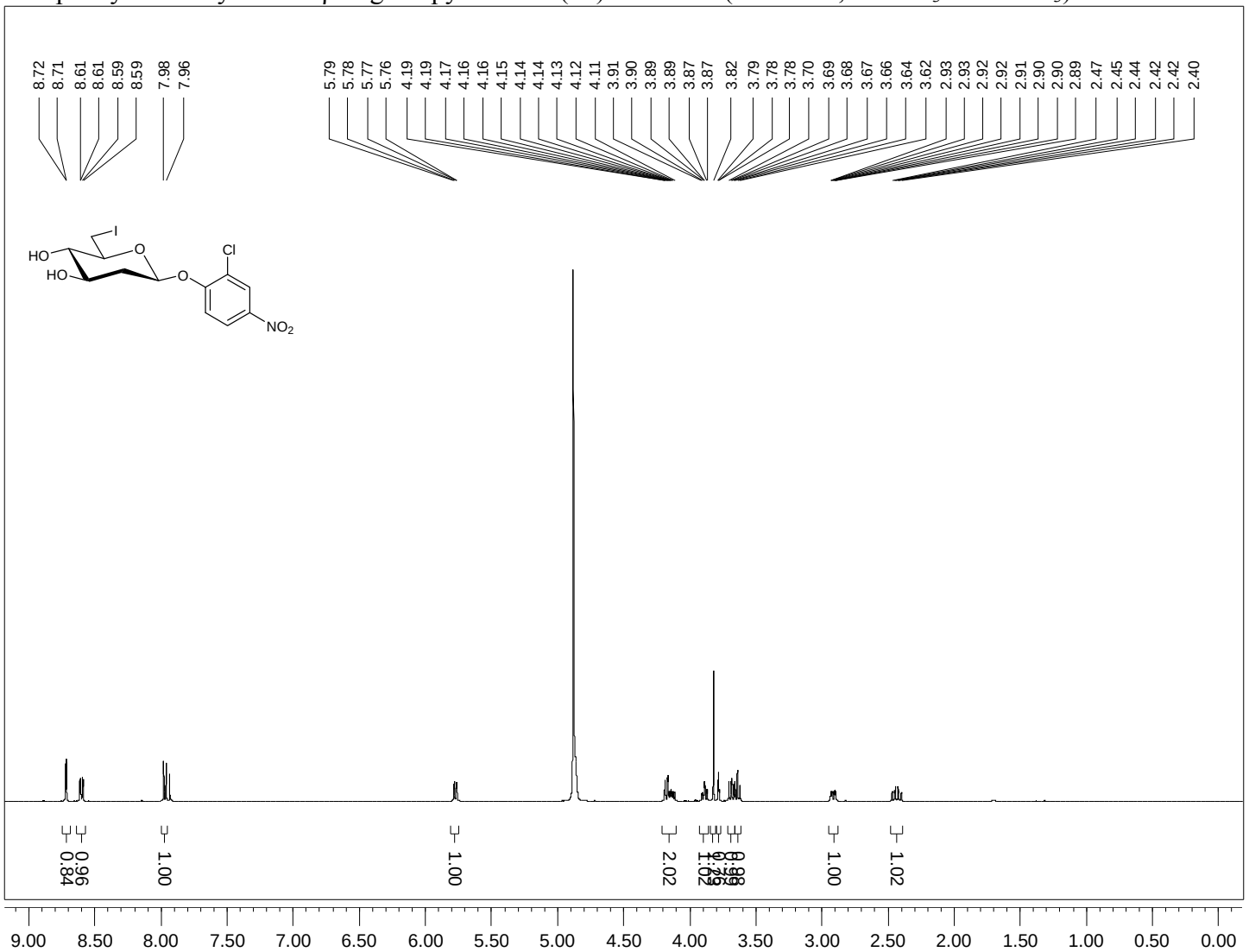
2-chloro-4-nitrophenyl-β-D-4,6-dideoxy-glucopyranoside (**5**): ^{1}H NMR (500 MHz, 1:3 CD_3OD:$CDCl_3$)



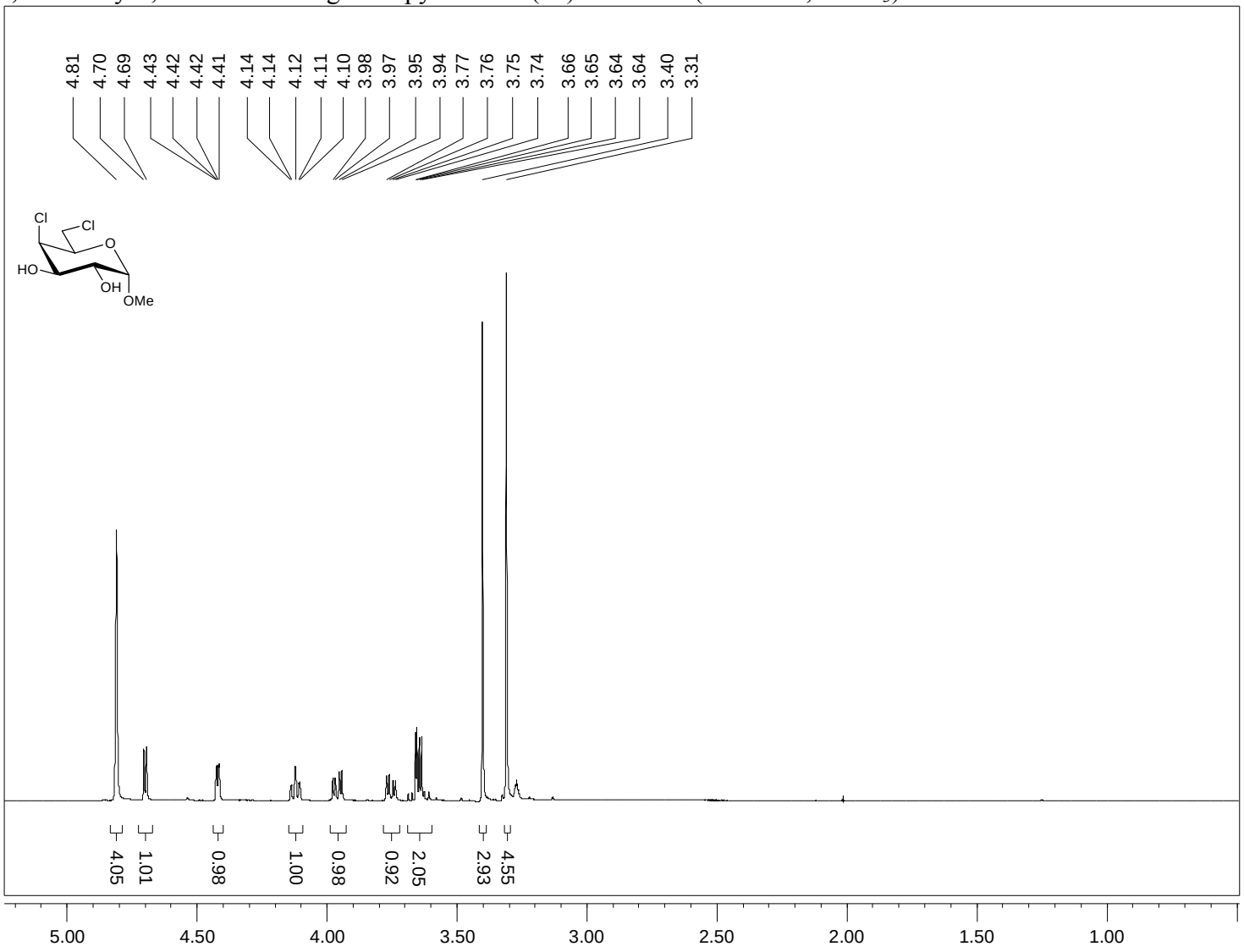
2-chloro-4-nitrophenyl-2-deoxy-6-*O*-tosyl-β-D-glucopyranoside (**40**): ^{1}H NMR (400 MHz, $CDCl_3$)



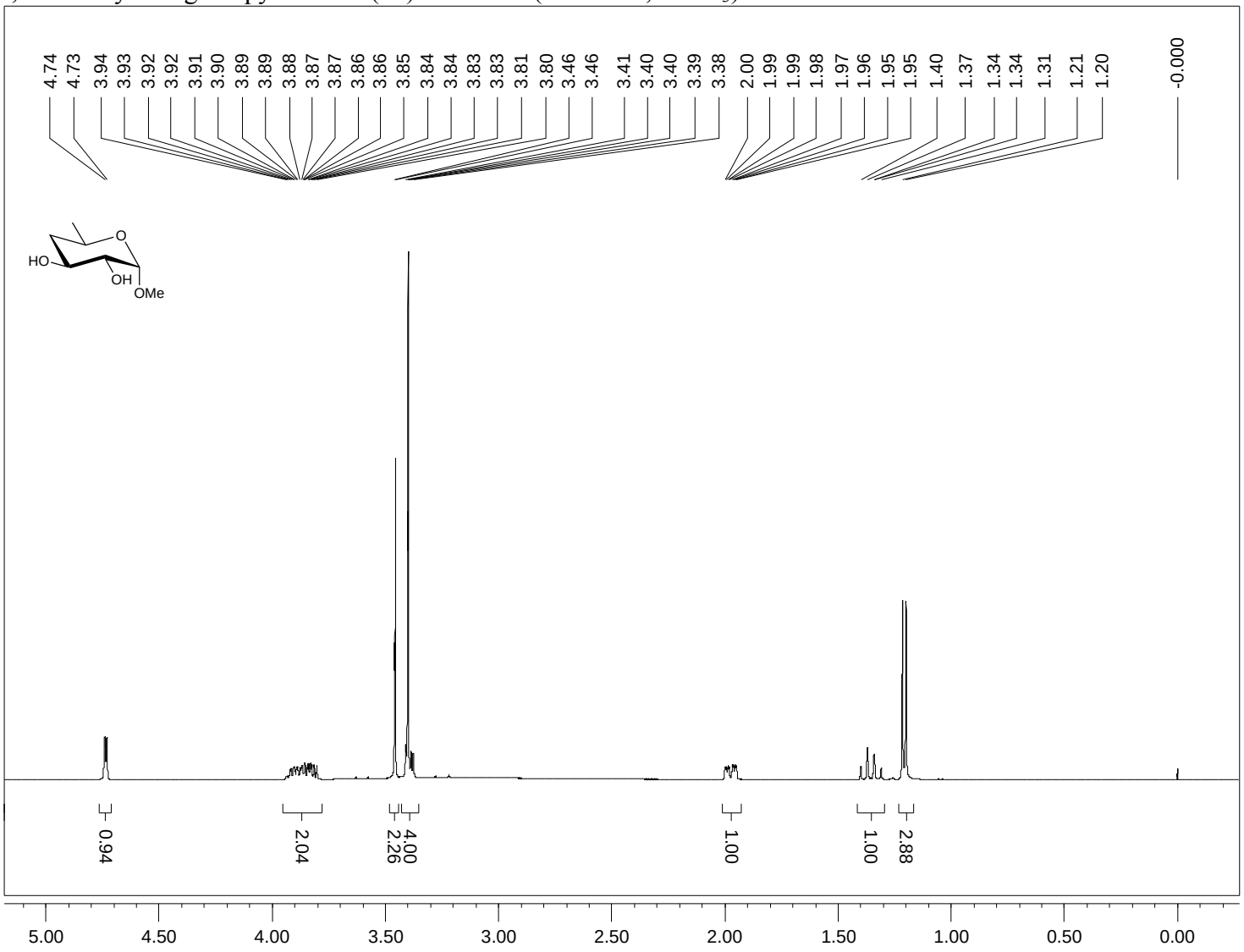
2-chloro-4-nitrophenyl-2-deoxy-6-iodo-β-D-glucopyranoside (**41**): ^{1}H NMR (500 MHz, 1:4 $CD_3OD/CDCl_3$)



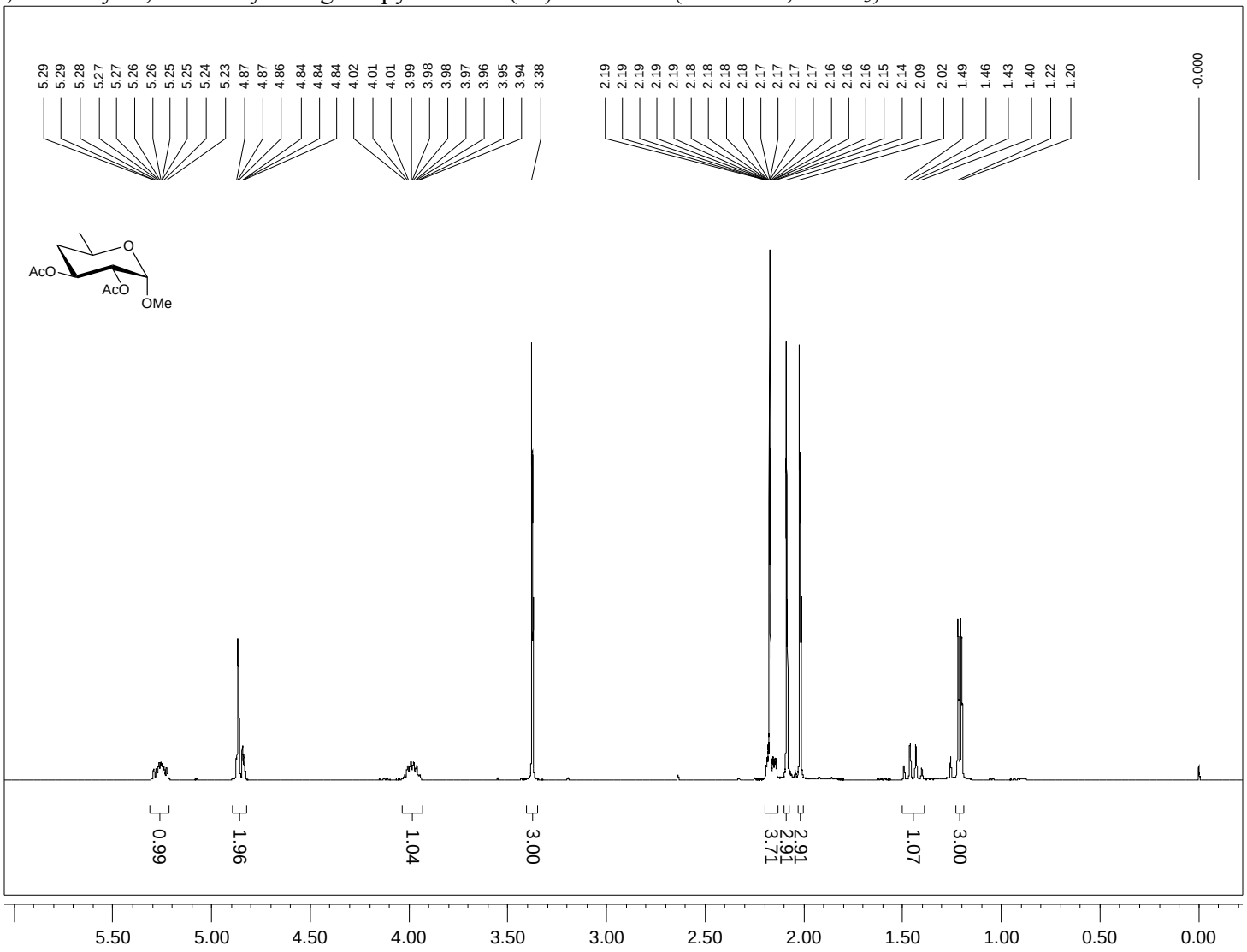
1-methyl-4,6-dideoxy-4,6-dichloro-α-D-galactopyranoside (**43**): ^{1}H NMR (400 MHz, $CDCl_3$)



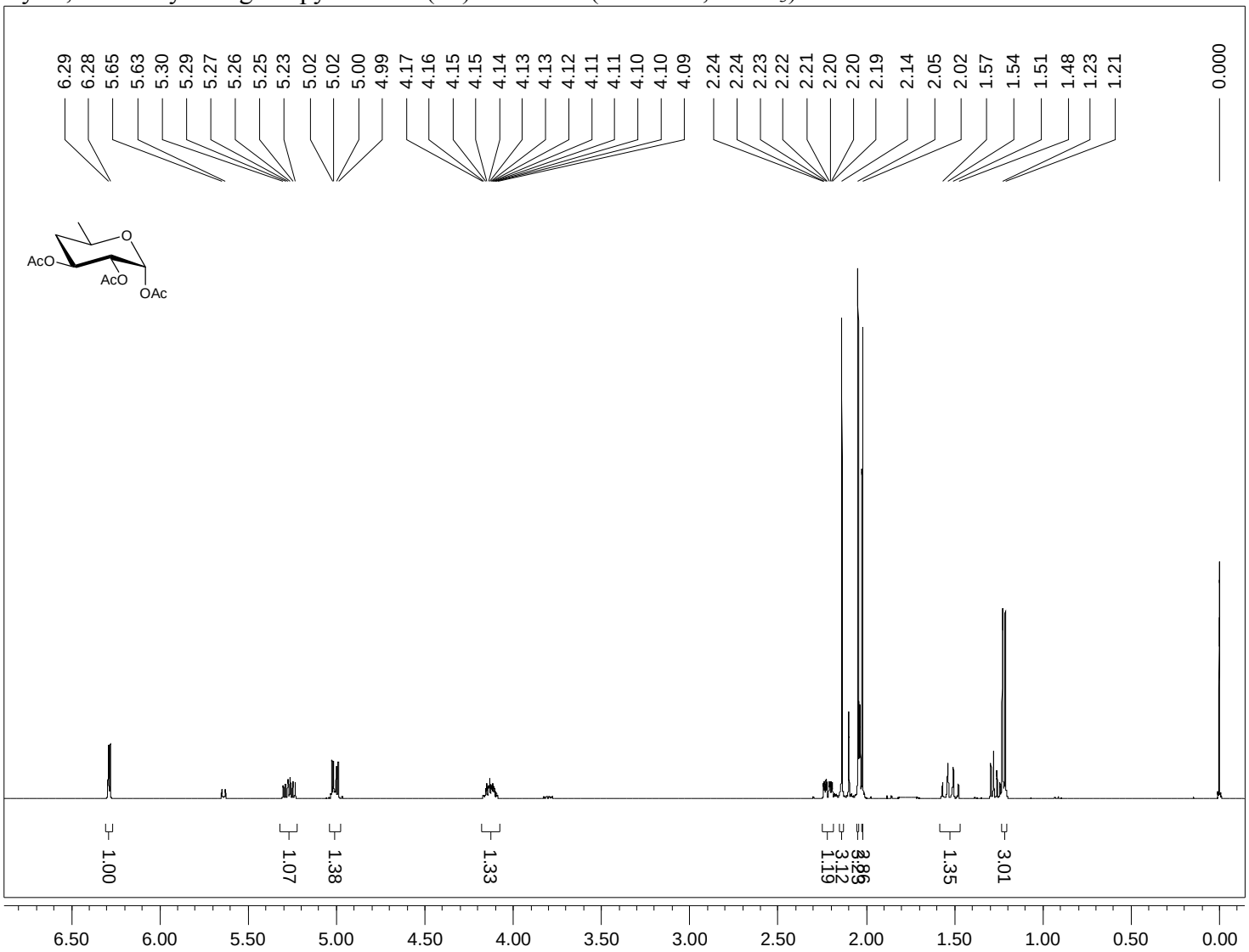
1-methyl-4,6-dideoxy-α-D-glucopyranoside (**44**): ^{1}H NMR (400 MHz, $CDCl_3$)



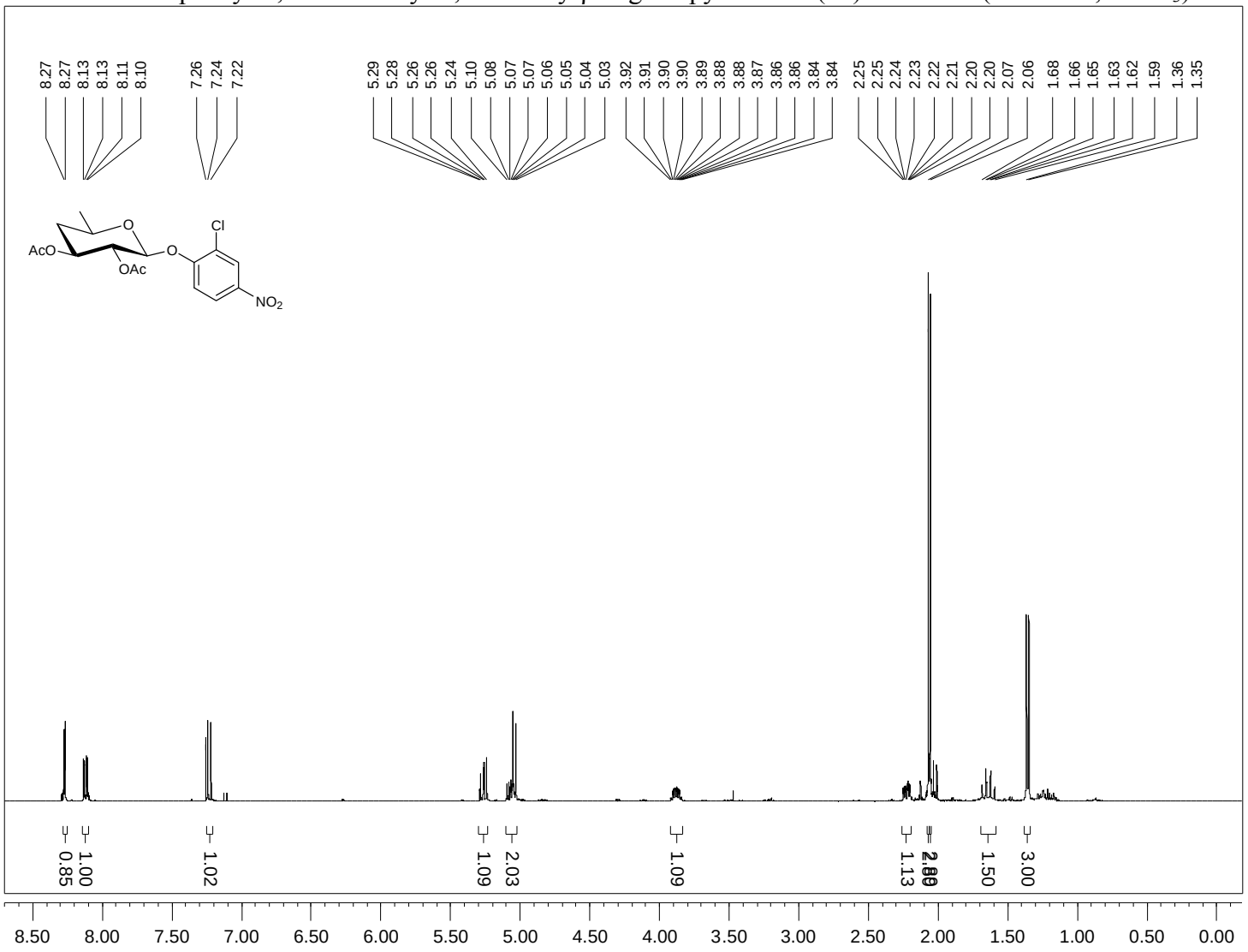
1-methyl-2,3-diacetyl-4,6-dideoxy-α-D-glucopyranoside (**45**): ^{1}H NMR (400 MHz, $CDCl_3$)



1,2,3-triacetyl-4,6-dideoxy-α-D-glucopyranoside (**46**): ^{1}H NMR (400 MHz, $CDCl_3$)

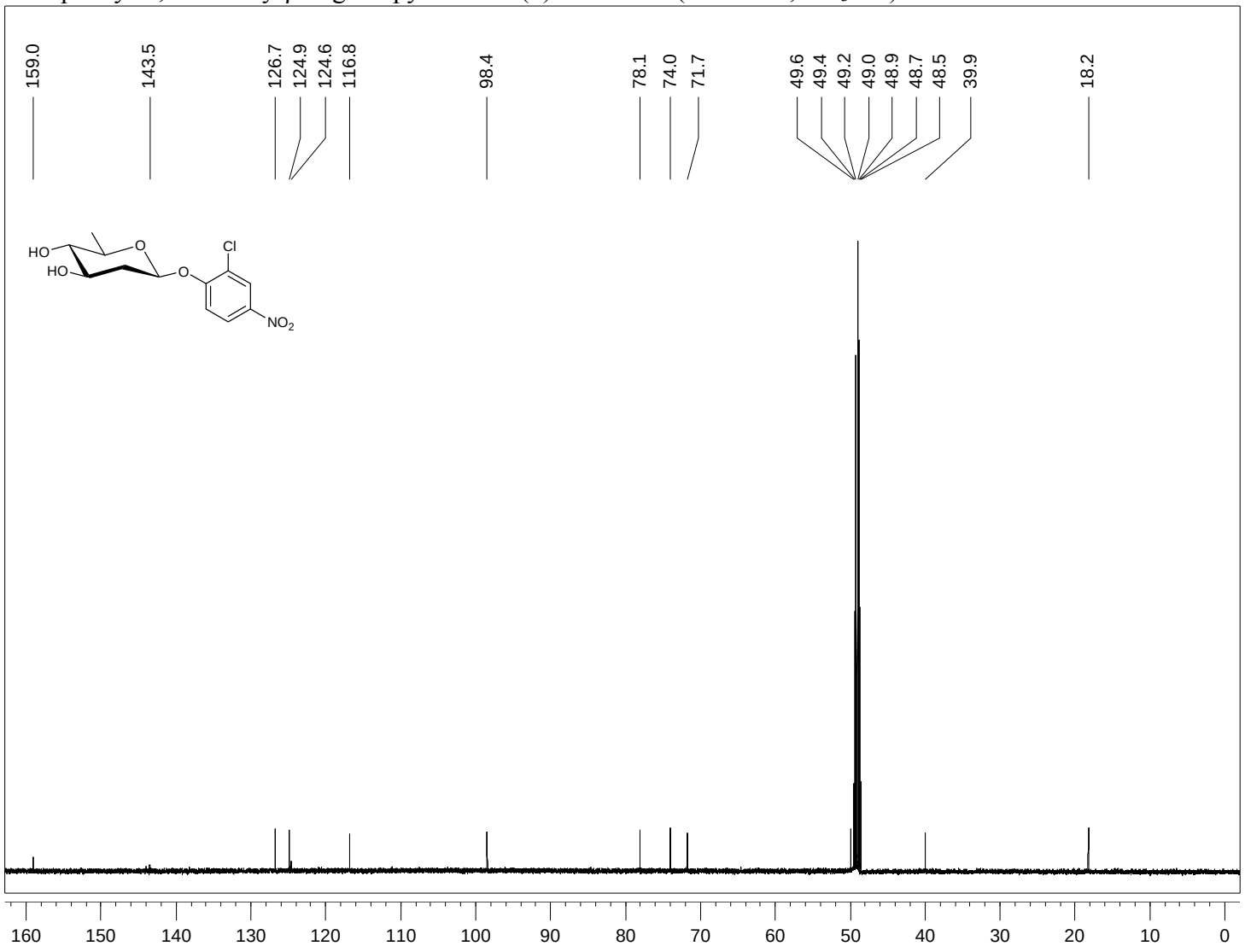


2-chloro-4-nitrophenyl-2,3-di-*O*-acetyl-4,6-dideoxy-β-D-glucopyranoside (**47**): ^{1}H NMR (400 MHz, $CDCl_3$)

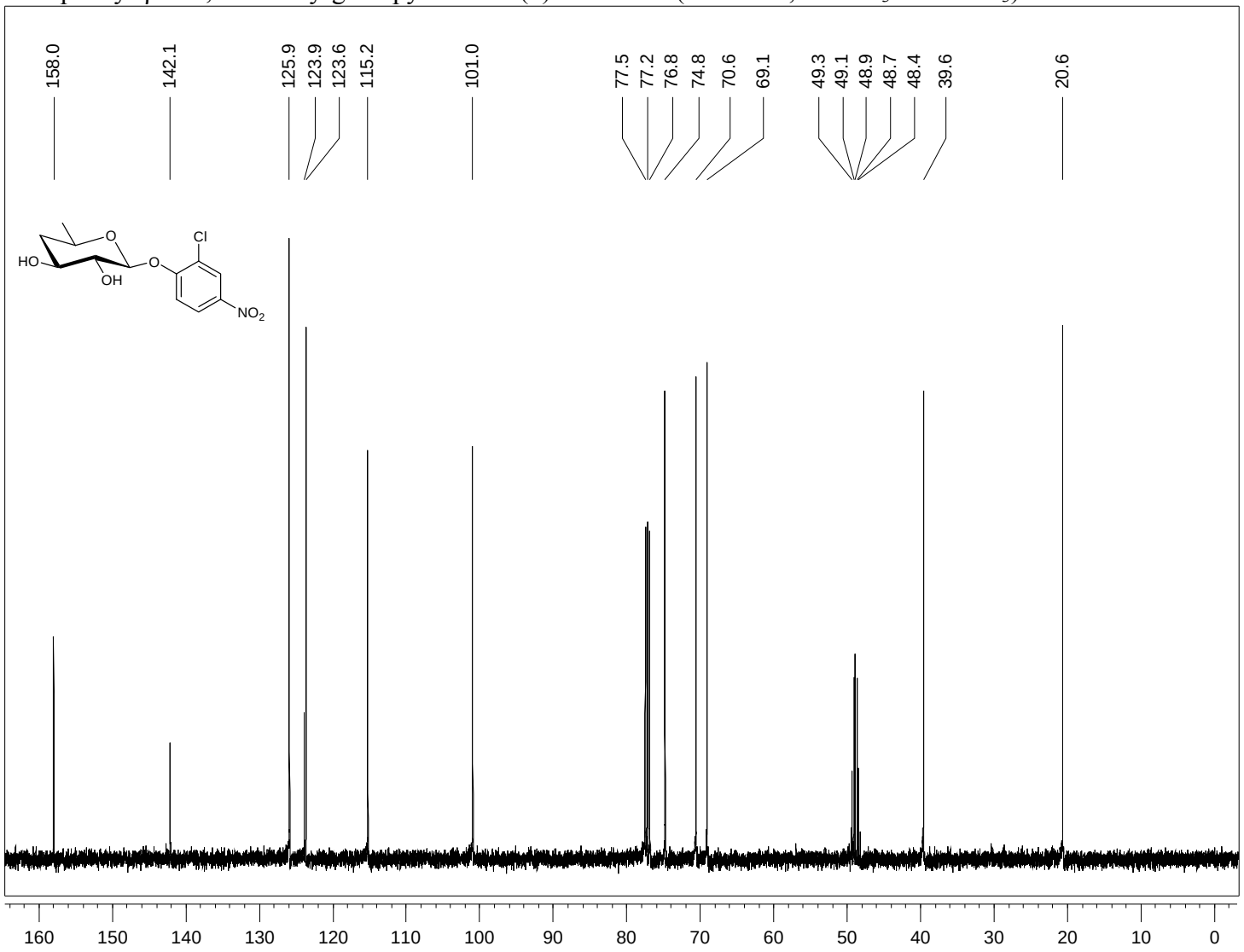


A3.6. ^{13}C NMR Spectra

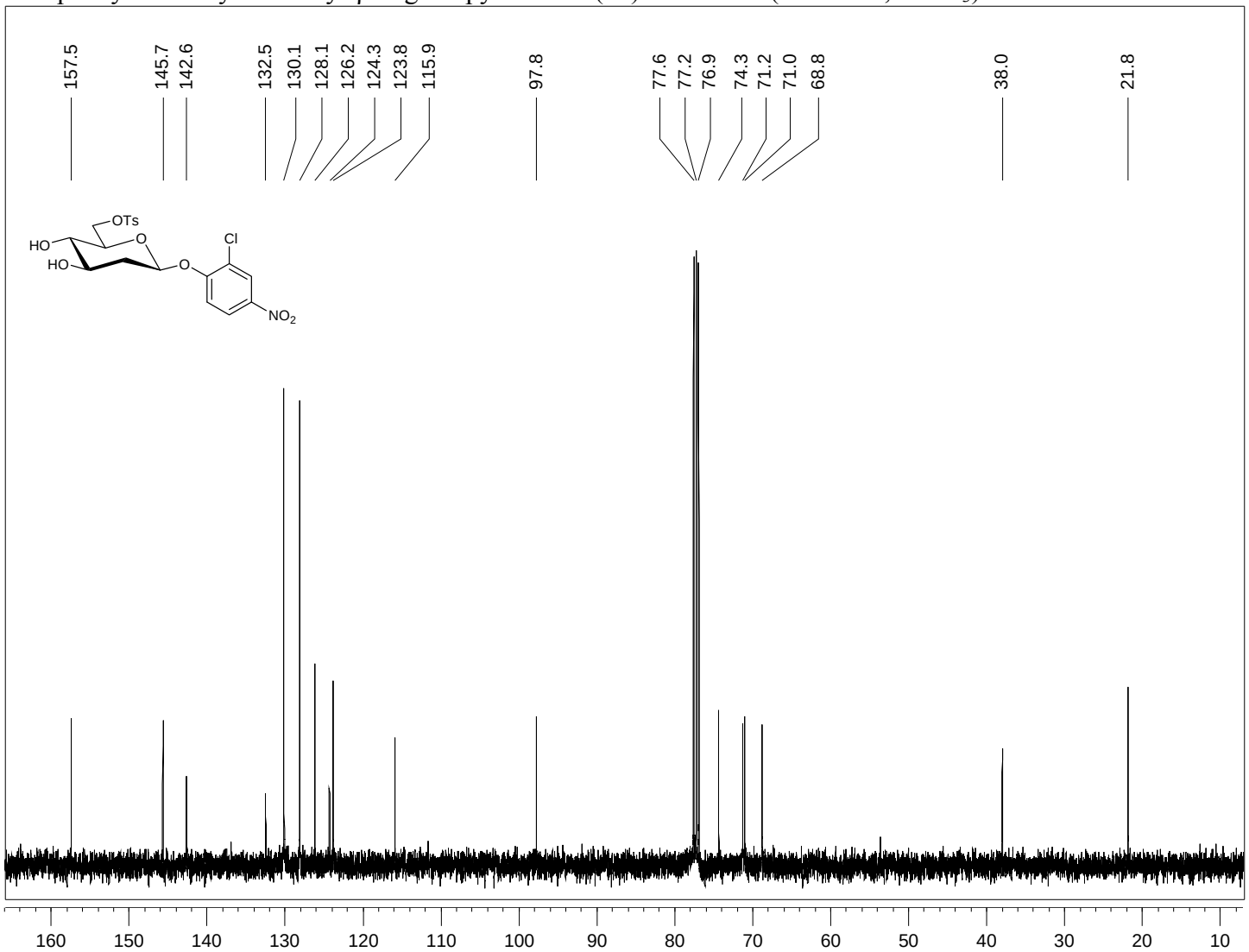
2-chloro-4-nitrophenyl-2,6-dideoxy-β-D-glucopyranoside (**4**): ^{13}C NMR (125 MHz, CD_3OD)



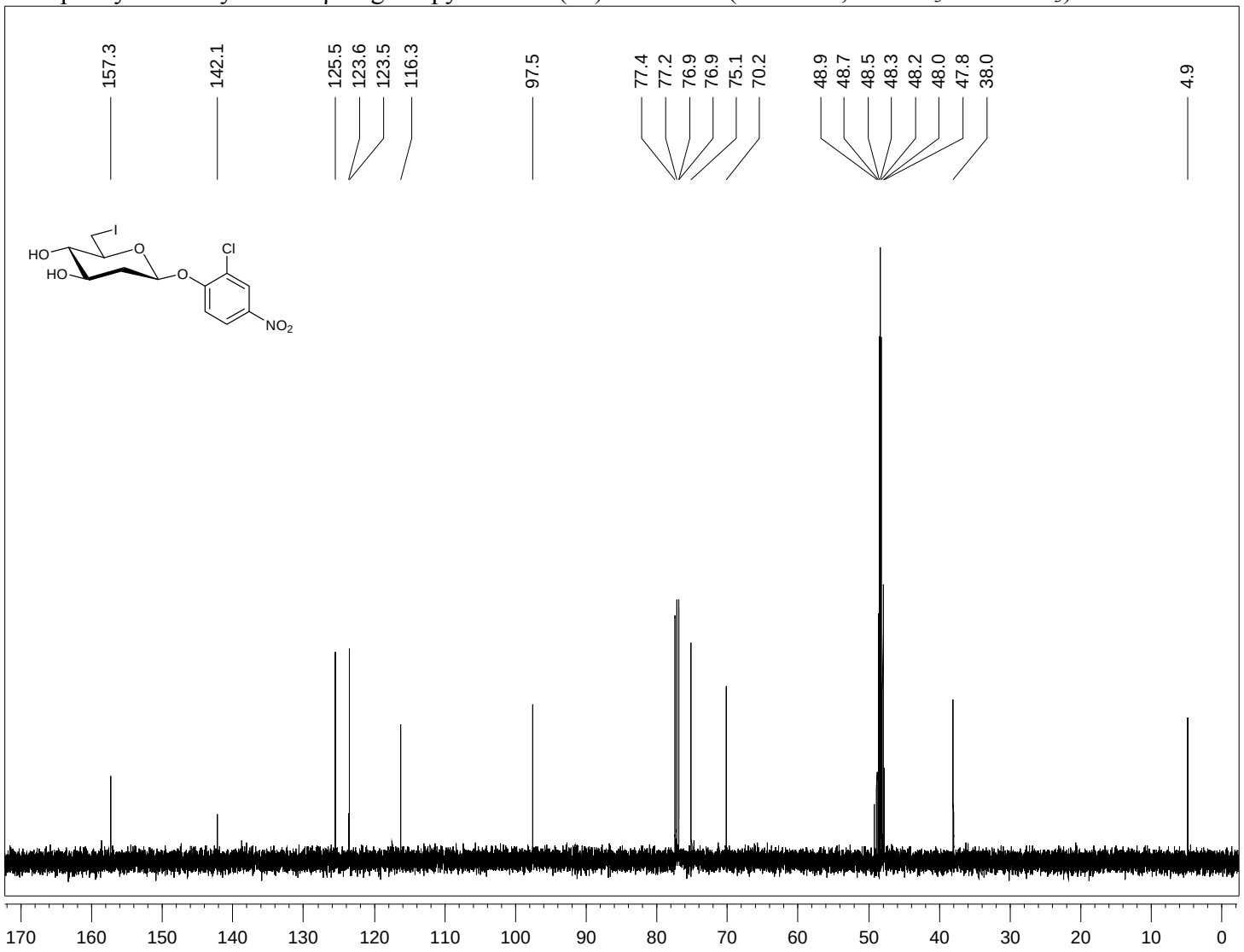
2-chloro-4-nitrophenyl-β-D-4,6-dideoxy-glucopyranoside (**5**): ^{13}C NMR (125 MHz, 1:3 CD_3OD:$CDCl_3$)



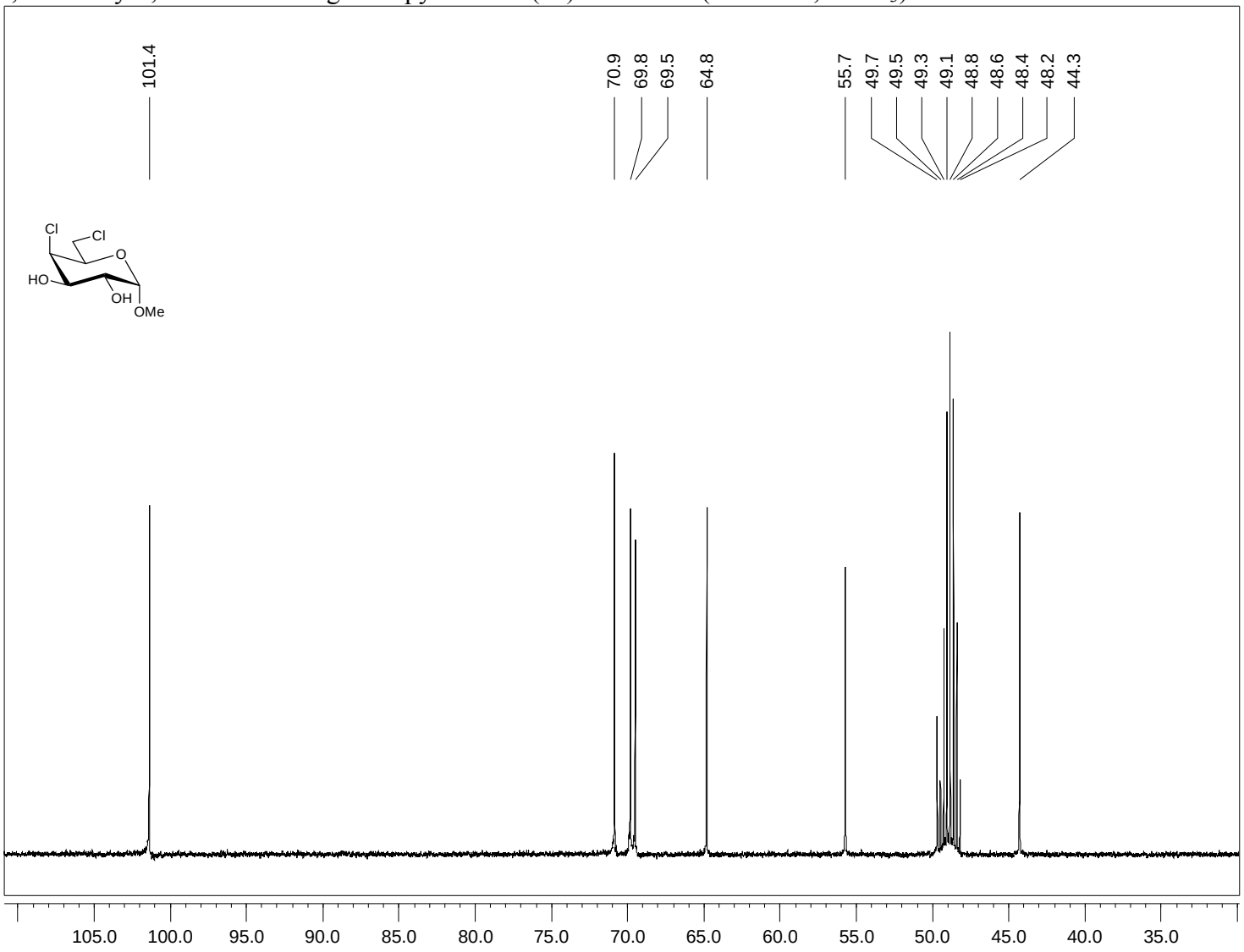
2-chloro-4-nitrophenyl-2-deoxy-6-*O*-tosyl-β-D-glucopyranoside (**40**): ^{13}H NMR (100 MHz, $CDCl_3$)



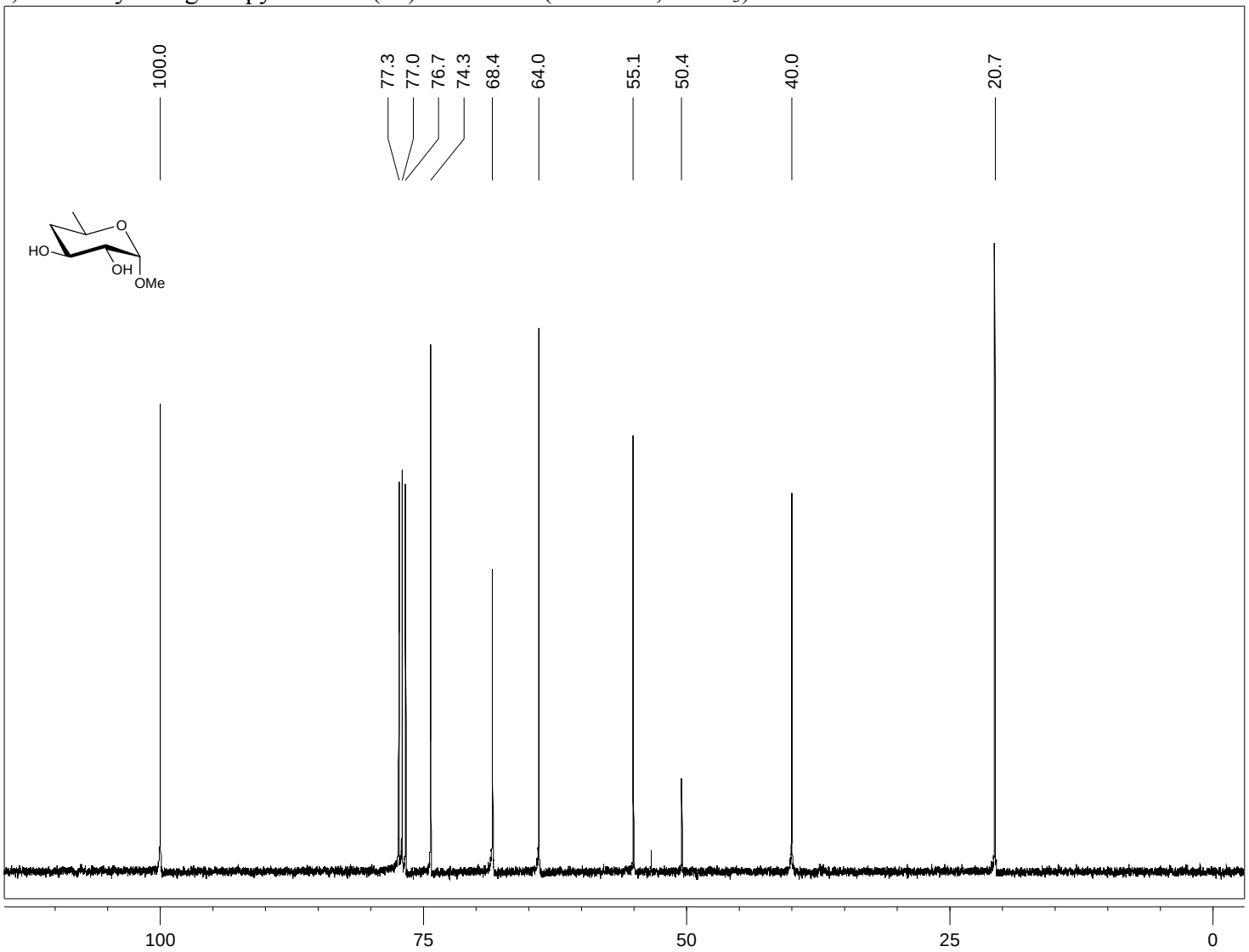
2-chloro-4-nitrophenyl-2-deoxy-6-iodo-β-D-glucopyranoside (**41**): ^{13}C NMR (125 MHz, 1:4 $CD_3OD/CDCl_3$)



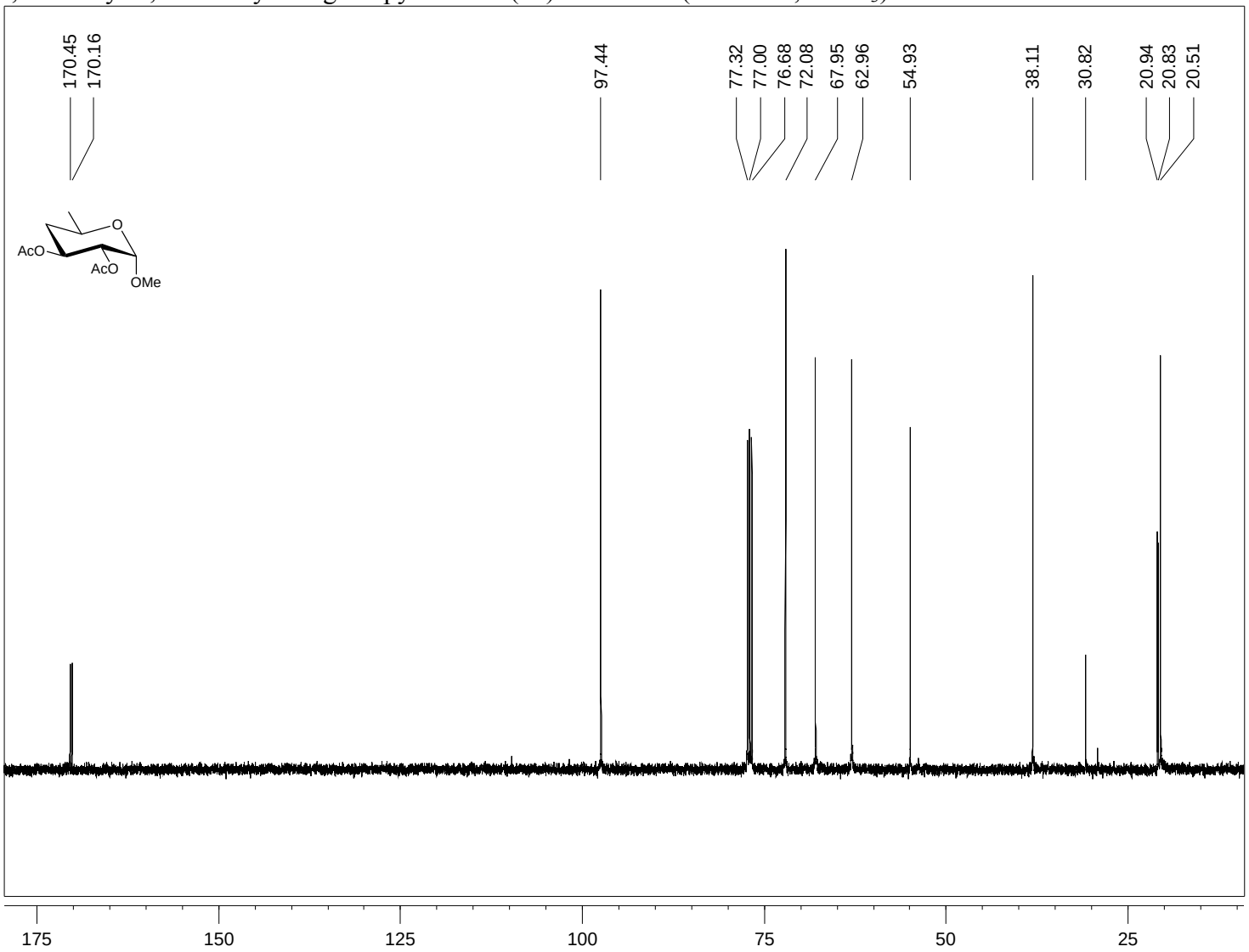
1-methyl-4,6-dideoxy-4,6-dichloro-α-D-galactopyranoside (**43**): ^{13}C NMR (100 MHz, $CDCl_3$)



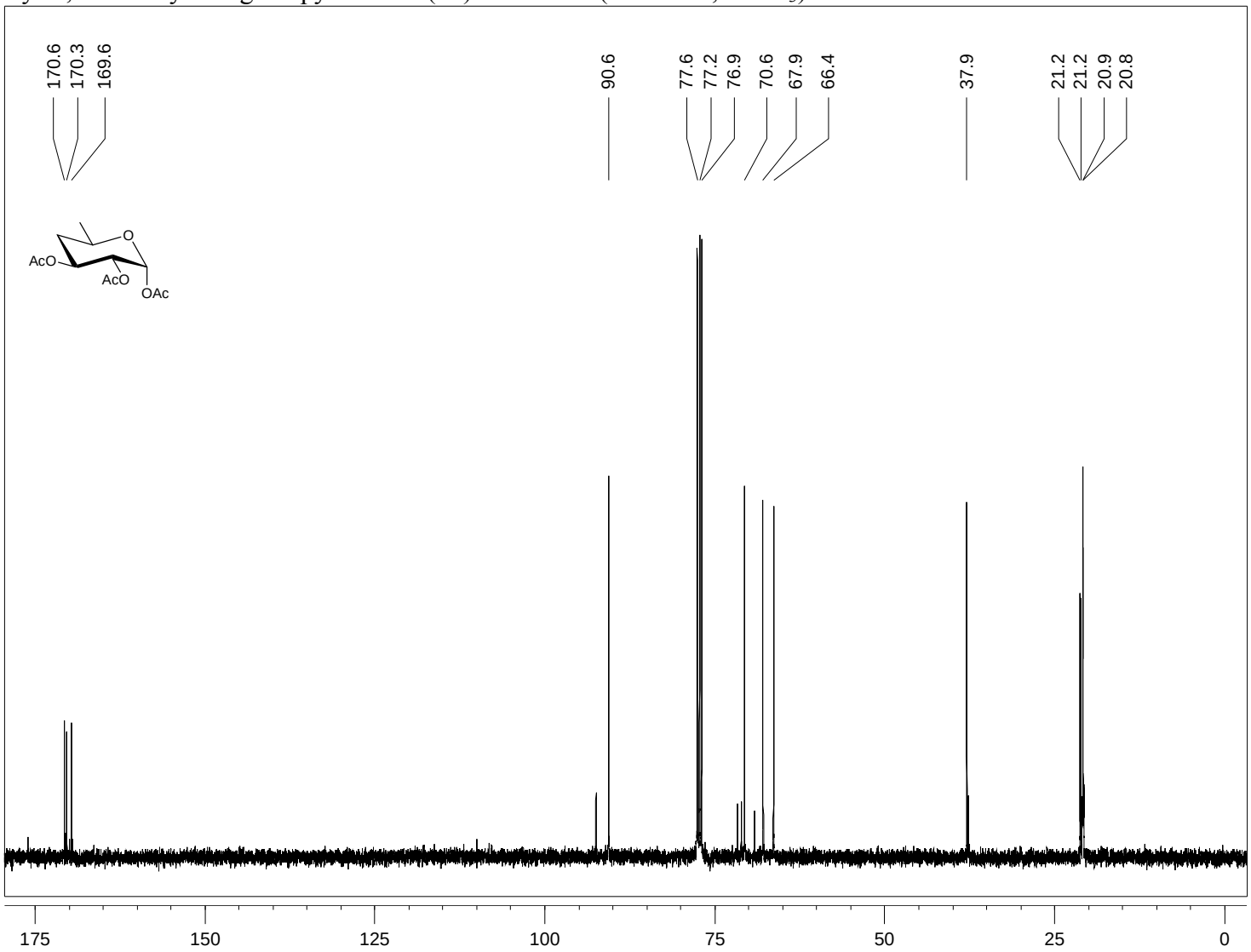
1-methyl-4,6-dideoxy-α-D-glucopyranoside (**44**): ^{13}C NMR (100 MHz, $CDCl_3$)



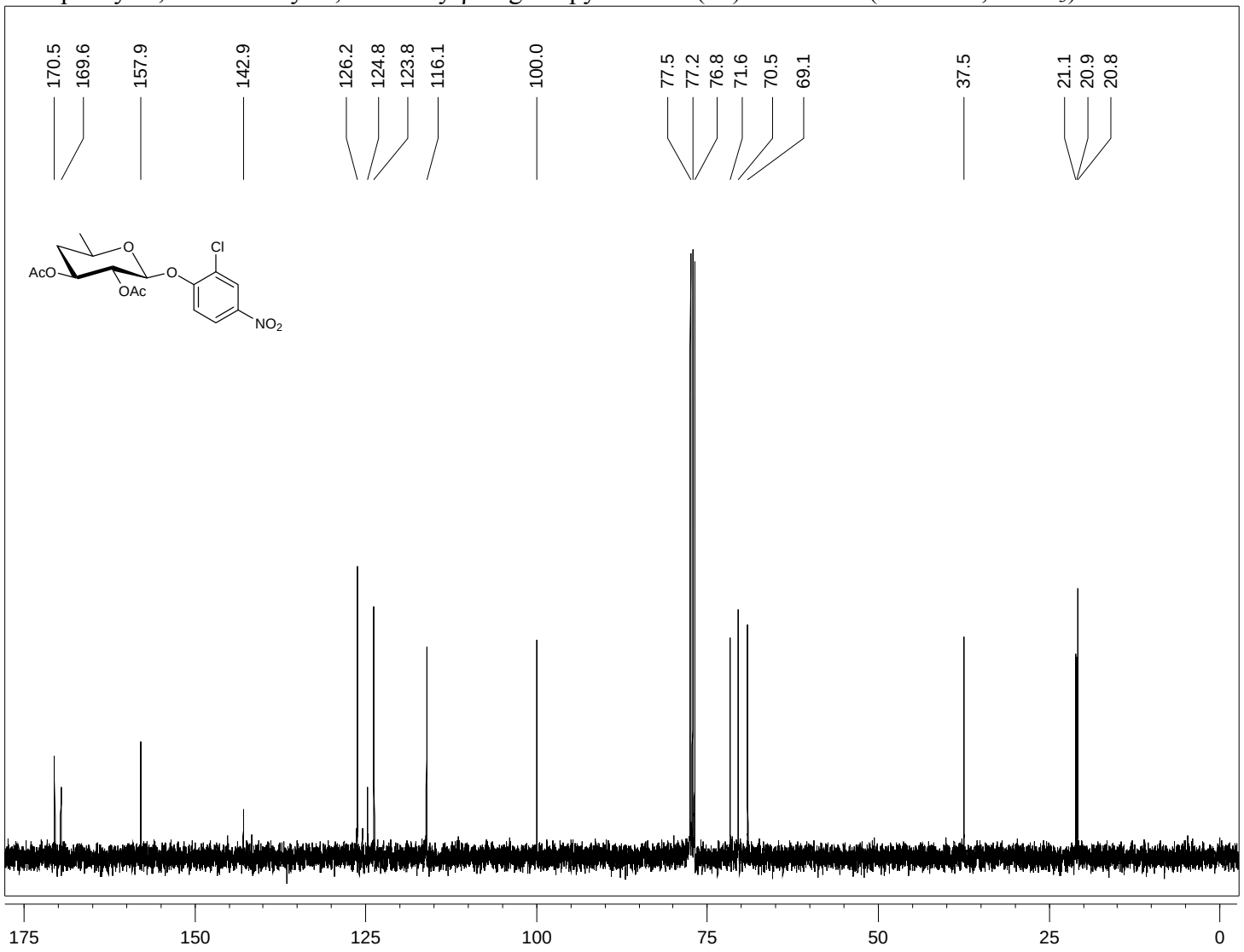
1-methyl-2,3-diacetyl-4,6-dideoxy-α-D-glucopyranoside (**45**): ^{13}C NMR (100 MHz, $CDCl_3$)



1,2,3-triacetyl-4,6-dideoxy-α-D-glucopyranoside (**46**): ^{13}C NMR (100 MHz, $CDCl_3$)



2-chloro-4-nitrophenyl-2,3-di-*O*-acetyl-4,6-dideoxy-β-D-glucopyranoside (**47**): ^{13}C NMR (100 MHz, $CDCl_3$)



A3.7. Appendix 3 References

1. Williams, G.J., Yang, J., Zhang, C., & Thorson, J.S. Recombinant *E. coli* prototype strains for *in vivo* glycorandomization. *ACS Chem. Biol.* **6**, 95-100 (2011).

2. Williams, G.J., Zhang, C., & Thorson, J.S. Expanding the promiscuity of a natural-product glycosyltransferase by directed evolution. *Nat. Chem. Biol.* **3**, 657-662 (2007).

3. Williams, G.J., Goff, R.D., Zhang, C., & Thorson, J.S. Optimizing glycosyltransferase specificity via "hot spot" saturation mutagenesis presents a catalyst for novobiocin glycorandomization. *Chem. Biol.* **15**, 393-401 (2008).

4. Williams, G.J. & Thorson, J.S. A high-throughput fluorescence-based glycosyltransferase screen and its application in directed evolution. *Nat. Prot.* **3**, 357-362 (2008).

5. Liébecq, C. Ed., *Biochemical Nomenclature and Related Documents* (Portland Press, Portland, OR, ed. 2, 1992).